

Hai Jin (Ed.)

Communications in Computer and Information Science

2451

Artificial Intelligence and Machine Learning

Second International Artificial Intelligence Conference, IAIC 2024
Jinyun, China, November 8–10, 2024
Revised Selected Papers, Part I

Part 1

 Springer



Communications in Computer and Information Science

2451

Series Editors

Gang Li , *School of Information Technology, Deakin University, Burwood, VIC, Australia*

Joaquim Filipe , *Polytechnic Institute of Setúbal, Setúbal, Portugal*

Zhiwei Xu, *Chinese Academy of Sciences, Beijing, China*

Rationale

The CCIS series is devoted to the publication of proceedings of computer science conferences. Its aim is to efficiently disseminate original research results in informatics in printed and electronic form. While the focus is on publication of peer-reviewed full papers presenting mature work, inclusion of reviewed short papers reporting on work in progress is welcome, too. Besides globally relevant meetings with internationally representative program committees guaranteeing a strict peer-reviewing and paper selection process, conferences run by societies or of high regional or national relevance are also considered for publication.

Topics

The topical scope of CCIS spans the entire spectrum of informatics ranging from foundational topics in the theory of computing to information and communications science and technology and a broad variety of interdisciplinary application fields.

Information for Volume Editors and Authors

Publication in CCIS is free of charge. No royalties are paid, however, we offer registered conference participants temporary free access to the online version of the conference proceedings on SpringerLink (<http://link.springer.com>) by means of an http referrer from the conference website and/or a number of complimentary printed copies, as specified in the official acceptance email of the event.

CCIS proceedings can be published in time for distribution at conferences or as post-proceedings, and delivered in the form of printed books and/or electronically as USBs and/or e-content licenses for accessing proceedings at SpringerLink. Furthermore, CCIS proceedings are included in the CCIS electronic book series hosted in the SpringerLink digital library at <http://link.springer.com/bookseries/7899>. Conferences publishing in CCIS are allowed to use Online Conference Service (OCS) for managing the whole proceedings lifecycle (from submission and reviewing to preparing for publication) free of charge.

Publication process

The language of publication is exclusively English. Authors publishing in CCIS have to sign the Springer CCIS copyright transfer form, however, they are free to use their material published in CCIS for substantially changed, more elaborate subsequent publications elsewhere. For the preparation of the camera-ready papers/files, authors have to strictly adhere to the Springer CCIS Authors' Instructions and are strongly encouraged to use the CCIS LaTeX style files or templates.

Abstracting/Indexing

CCIS is abstracted/indexed in DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago, Scopus. CCIS volumes are also submitted for the inclusion in ISI Proceedings.

How to start

To start the evaluation of your proposal for inclusion in the CCIS series, please send an e-mail to ccis@springer.com.

Hai Jin
Editor

Artificial Intelligence and Machine Learning

Second International Artificial Intelligence Conference, IAIC 2024
Jinyun, China, November 8–10, 2024
Revised Selected Papers, Part I

Editor

Hai Jin

Huazhong University of Science and Technology
Wuhan, China

ISSN 1865-0929

ISSN 1865-0937 (electronic)

Communications in Computer and Information Science

ISBN 978-981-96-7187-8

ISBN 978-981-96-7188-5 (eBook)

<https://doi.org/10.1007/978-981-96-7188-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Preface

These conference proceedings are a collection of the papers accepted by IAIC 2024 — the 2024 International Artificial Intelligence Conference, held on November 8–10, 2024, in Jinyun, China.

The conference was organized by Zhejiang Sci-Tech University and Tech Science Press. IAIC 2024 provided a vibrant platform for the exchange of ideas and discussions on the latest advancements in artificial intelligence. The conference encompassed a wide range of topics, including machine learning, natural language processing, computer vision, robotics, AI ethics, and more.

The reviewing process for **IAIC 2024** was rigorous and meticulously organized. A total of **108 submissions** were received, and each manuscript underwent a **single-blind peer-review process**, with **two independent review reports** provided for every paper. Despite the high overall quality of the submissions, only **39 papers** were accepted for publication. This selective acceptance rate reflects the conference's commitment to academic excellence and the outstanding caliber of research showcased at IAIC 2024.

We extend our heartfelt gratitude to the authors for their excellent contributions, as well as to the reviewers for their dedication to ensuring a thorough evaluation process, which has been instrumental in maintaining the conference's high quality.

We would also like to express our appreciation to the organizers and sponsors, whose support and commitment made IAIC 2024 a success. Their dedication to advancing the field of artificial intelligence is truly commendable. The logos of our esteemed sponsors can be found on the following pages.

We hope this volume serves as a valuable resource for researchers, academics, and practitioners, fostering continued innovation and dialogue in the dynamic field of artificial intelligence.

IAIC 2024 Organizing Committee

December 2024

Organization

General Chair

Hai Jin	Huazhong University of Science and Technology, China
---------	---

Technical Program Committee Chair

Zhihui Zhan	Nankai University, China
-------------	--------------------------

Technical Program Committee Members

Sohaib Latif	University of Chenab, Pakistan
Wei Liang	Hunan University of Science and Technology, China
Muhammad Aamir	Huanggang Normal University, China
Ziaur Rahman	Huanggang Normal University, China
Zhonglin He	Huanggang Normal University, China
Jameel Ahmed Bhutto	Huanggang Normal University, China
Tariq Hussain	Zhejiang Gongshang University, China
Haseeb Ur Rahaman	University of Malakand, Pakistan
Muhammad Shahid Anwar	Gachon University, South Korea
Khurshed Aurangzeb	King Saud University, Saudi Arabia
Javed Ali Khan	University of Hertfordshire, UK
Hameedur Rahman	Air University, Pakistan
Yutao Tang	Guilin University of Electronic Technology, China
Qiong Wu	Jiangnan University, China
Zhiguo Qu	Nanjing University of Information Science and Technology, China
Le Sun	Nanjing University of Information Science and Technology, China
Chenxi Huang	Xiamen University, China
Yudong Zhang	University of Leicester, UK
Nitin Goyal	Central University of Haryana, India
Nonita Sharma	Indira Gandhi Delhi Technical University for Women, India

Xingsi Xue	Fujian University of Technology, China
Haoxi Zhang	Chengdu University of Information Technology, China
Xia Du	Xiamen University of Technology, China
Qi Zhang	City University of Macau, China
Jizhe Zhou	Sichuan University, China
Xin Liu	Lanzhou University, China
Saadaldeen Rashid Ahmed	Bayan University, Iraq
Gang Xu	North China University of Technology, China
Mohammed Amin Almaiah	University of Jordan, Jordan
Said A. Salloum	University of Salford, UK
Fang Liu	Hunan University, China
Zhiping Cai	National University of Defense Technology, China
Yizhang Wang	Yangzhou University, China
Le Zou	Hefei University, China
Weiwei Jiang	Beijing University of Posts and Telecommunications, China
Weixi Gu	China Academy of Industrial Internet, China
Miao He	Beijing Institute of Mathematical Sciences and Applications, China
Jian Su	Nanjing University of Information Science and Technology, China
Zhengguo Sheng	University of Sussex, UK
Siguang Chen	Nanjing University of Posts and Telecommunications, China
Wei Fang	Nanjing University of Information Science and Technology, China
Victor S. Sheng	Texas Tech University, USA

Organization Committee Members

Heming Jia	Sanming University, China
Yi Cai	South China University of Technology, China
Huiling Chen	Wenzhou University, China
Wei Fang	Jiangnan University, China
Jingrui Zhang	Xiamen University, China
Wentao Li	Southwest University, China
Deming Lei	Wuhan University of Technology, China
Keping Yu	Hosei University, Japan
Hongfeng Wang	Northeastern University, China

Shijun Liu	Shandong University, China
Xingsi Xue	Fujian University of Technology, China
Taiyong Li	Southwestern University of Finance and Economics, China
Zhi-Hui Zhan	Nankai University, China
Bin Ma	Shandong University of Technology, China
Rencan Nie	Yunnan University, China
Shiwei Cheng	Shanghai Jiao Tong University, China
Haidong Shao	Hunan University, China
Yin Zhang	University of Electronic Science and Technology of China, China
Yan Pei	University of Aizu, Japan
Changyan Yi	Nanjing University of Aeronautics and Astronautics, China
Chunwei Tian	Northwestern Polytechnical University, China
Zhiping Cai	National University of Defense Technology, China
Pei-Wei Tsai	Swinburne University of Technology, Australia
Farman Ali	Sungkyunkwan University, South Korea
Adam Slowik	Koszalin University of Technology, Poland
Nitin Goyal	Central University of Haryana, India
Cesar Maldonado Sanin	University of Newcastle, Australia
Zhenhua Yu	Xi'an University of Science and Technology, China
Xiaozhi Gao	University of Eastern Finland, Finland
Yue Zhang	Drexel University, USA
Nadeem Sarwar	Bahria University Lahore Campus, Pakistan
Saeed Hamood Alsamhi	University of Galway, Ireland
Alexander Mikroyannidis	Open University, UK
Ashutosh Kumar Dubey	Chitkara University, India
Eugenio Vocaturro	University of Calabria, Italy
Fabrizio Marozzo	University of Calabria, Italy
Jawad Ahmad	Edinburgh Napier University, UK
Marios Avgeris	Carleton University, Canada
Noor Zaman Jhanjhi	Taylor's University, Malaysia
Muhammad Tariq Sadiq	University of Essex, UK
Batyrkhan Omarov	Al-Farabi Kazakh National University, Kazakhstan
Anand Nayyar	Duy Tan University, Vietnam
Seifedine Kadry	Noroff University College, Norway
Essam H. Houssein	Minia University, Egypt
Ali Safaa Sadiq	Nottingham Trent University, UK

Sultan Daud Khan	National University of Technology, Pakistan
Jeng-Shyang Pan	Shandong University of Science and Technology, China
Roseline Oluwaseun Ogundokun	Landmark University, Nigeria
Mukesh Prasad	University of Technology Sydney, Australia
Qiong Wu	Jiangnan University, China
Xin Tan	East China Normal University, China
Songjie Wei	Nanjing University of Science and Technology, China
Gokhan Altan	Iskenderun Technical University, Turkey
Yu-Cheng Hu	Tunghai University, Taiwan
Zhi-Hui Zhan	Nankai University, China
Shahid Anwar	Gachon University, South Korea
Tianyuan Liu	Donghua University, China

Organizers

Co-organized by



Zhejiang Sci-tech University



Tech Science Press

Supporting Institutes



Contents – Part I

Artificial Intelligence in Real-World Applications

An Adaptive Power Control Method for Power Communication Networks Based on Quantum Security	3
<i>Xingnan Li, Donghai Huang, Jiajia Fu, Zhongmiao Kang, and Jiangang Lu</i>	
From Machine Intelligence to Machine Ethics: An Analysis of the Practical Path of Artificial Morality	12
<i>Jingjing Zhao and Xingtong Liu</i>	
The Frequency Bounds for Travelling Salesman Problem Based on Frequency Quadrilaterals	23
<i>Yong Wang and WenFa Zheng</i>	
Speech Recognition Method Based on Improved Transformer	39
<i>Liqiang Wang, Dong Lv, Wenchen He, Haiyang Wang, Yifan Dong, and Mingshuang Sui</i>	
A Two-Stage Approach to Multi-Feature Acoustic Scene Classification	49
<i>Liqiang Wang, Yitao Li, Shan Ning, Haiyang Wang, Jie Chu, and Qirong Zhou</i>	
Star Extraction Based on Multi-scale Grey Morphology in Stray Light Backgrounds	61
<i>Shuhui Zhou, Yang Zheng, Chenxin Lu, Jinwei Cheng, and Wanyu Liu</i>	
Lightweight One-Stage Ship Detection Based on SAR Images for Embedded Platform	76
<i>Chenxin Lu, Ji Shen, Shuhui Zhou, Yang Zheng, and Jie Yuan</i>	
Multi-Modal False Information Identification Using BERT and ResNet	89
<i>Dong Lv, Yitao Li, Wenchen He, Zihan Wang, Jiawei Qin, and Liqiang Wang</i>	
Arithmetic Awareness Based Distributed Routing Mechanism for Power Dispatch Networks	99
<i>Zhifei Hao, Xiao Feng, Hang Zhang, Hanming Yan, and Guangwei Liu</i>	

A Dual-Mode Integrated Communication Method for Power Line Carrier and Wireless in Distribution Systems 109
Xiao Feng, Peiyao Zhang, Mingtao Cui, Yuhao Chang, Yong Sun, and Wenbin Liu

High-Performance Database Storage Technology for Power Smart Terminals 119
Lycho Huang, Xiao Feng, Jinguo Fang, Yuhao Chang, and Guangwei Liu

Quantum Key Distribution Strategy for Power Quantum Communication Networks Based on Trusted Relay Nodes 132
Donghai Huang, Xingnan Li, Jiajia Fu, Wenjuan Liang, and Zhongmiao Kang

Research on Network Traffic Anomaly Detection Method Integrating Graph Topology Optimization and Random Restart Walk Algorithm 142
Wen Luo, Haotian Li, Kunxiang Niu, and Longfei Yao

An Evaluation Method for Cutted Generated Video Based on Eigenvectors 158
Ji Yu, Haotian Gao, and Jundi Ding

Joint Optimization of Task Scheduling and Resource Allocation for Hierarchical Federated Learning 169
Nan Ding, Tianshu Huang, and Liutao Zhao

A Smart Grid Planning Data Decision Scoring Model Based on Random Forest 179
Lijun Xiao, Xiangwei Meng, Dezhi Han, and Dacheng He

Feature Importance on Axial Compressive Strength Prediction of Concrete-Filled Steel Tubular Columns Using Random Forest Algorithm 189
Bo Su, Zhenhong Wu, Dongqi Jiang, Gang Bi, Jiaxin Fan, and Kai Yan

AEDs Treatment and Cognition Analysis of BECTS Through EEG Features ... 201
Zhezha Zhao, Jiagui Chen, Dinghan Hu, Jiajia Fang, and Jiuwen Cao

Light Source Optimization Design of Intelligent Grading Equipment for Flue-Cured Tobacco 212
Chen Pengbiao, Hou Kaihu, Gong Peiyao, and Bai Fan

Multidimensional Attention Convolution Plugin for Ship Radiated Noise Classification 228
Zhongxiang Zheng and Peng Liu

Underwater Acoustic Target Detection Network Based on Comprehensive
Feature Extraction 239
 Denghui Sun, Kaiqiao Wang, Peng Liu, and Chun Zhang

Research on a Temperature Prediction Model for Meteorological Data
Integrating XGBoost and MLP 252
 Wenhua Xue, Yuxing Gao, Long Chen, and Enqi Lv

A Review of Research Advances in Image Segmentation of Skin Lesions 265
 Mengmeng He and Xiaoyan Zhang

Author Index 281

Contents – Part II

Artificial Intelligence in Network and Security Systems

Noise-Resilient Compressive Sensing via Orthogonality Constraints	3
<i>Yang Zheng, Zhen Zhang, Jie Yuan, Shuyi Feng, and Chenxin Lu</i>	
Joint Sensing of Computing Power and Network for Smart Grid: A Deep Matrix Completion Approach	18
<i>Yuxiang Qiu, Lei Zhang, Yuan Zhu, and Lan Gan</i>	
Rethinking RSSI-Based Key Extraction for UAVs and Ground Stations	31
<i>Yanbo Li and Jain Huang</i>	
Campus Card Security Depository System Based on State Secret Algorithm ...	43
<i>Yingzi Huo, Kefei Li, Meiguo Ke, Yanbing Wu, Yuxin Liu, and Ronglin Zhang</i>	
Network Traffic Classification via Deep Hashing Network	54
<i>Kunxiang Niu and Wen Luo</i>	
A Trust Management Integrated Bayesian Network Prediction in Vehicular Networks	66
<i>Hongbin Zhang, Zi Li, Dongmei Zhao, and Yuelin Liu</i>	
Research and Development of Distribution Line Defects and Violations Detection Method and Edge Computing System Based on Improved PPYOLOe	77
<i>Zaixin Chen, Peike Du, Tihua Yan, and Kaijun Wu</i>	
A Joint Optimization Algorithm for Computing Resource Allocation, UAV Trajectory, and Task Offloading in Remote Regional VANETs	93
<i>Geng Chen, Yiming Zhang, and Qingtian Zeng</i>	
LSTMR: Spatio-Temporal 3D Multiscale ResNet Model for Cellular Network Traffic Prediction	106
<i>Geng Chen, Weiming Wang, and Qingtian Zeng</i>	
An Efficient Name Lookup Approach Based on Character Indexing	121
<i>Dacheng He, Man Jiang, and Ke Xu</i>	

A Multi-channel Target Direction Speech Separation Model Based
on Efficient Convolutional Networks 132
Zaixin Chen, Yang Mao, Tihua Yan, and Biqing Li

Secure Cloud-Edge Collaborative Task Scheduling Framework Across
Data Centers 147
Ming Wan, Yuxiang Qiu, Yuan Zhu, and Lan Gan

Research on Security Sharing Mechanism of Cyber Threat Intelligence
Based on Consortium Blockchain 161
Fei Wang

Towards Automated Decoding of Vehicle CAN Data Using Deep Learning 181
Ankang Jiao, Wendi Li, Honglong Zhu, and Hao Han

Transfer Learning-Driven Surrogate-Assisted Differential Evolution
Algorithm with User Generated Content 201
Yumeng Peng, Lin Bao, Xiaoyan Sun, Jun Hao, Yu Wang, and Hang Yang

Author Index 217

Artificial Intelligence in Real-World Applications



An Adaptive Power Control Method for Power Communication Networks Based on Quantum Security

Xingnan Li^(✉), Donghai Huang, Jiajia Fu, Zhongmiao Kang, and Jiangang Lu

Guangdong Power Dispatch and Control Center, Guangdong 510000, China

2244697578@qq.com

Abstract. As an important component of modern industrial system, power system is re-lated to the lifeline of national economy and national energy security. With the increase of the number of the user input, its communication security has attracted more and more attention. Quantum key distribution (QKD) enables secure remote information transmission between two users theoretically. Therefore, the combination of QKD and the existing power communication network can better ensure the security of the existing power service transmiss-ion. However, since optical signals and quantum signals need to share a set of optical fiber transmission equipment, there will be resource competition between the two parts, and it is necessary to strike a balance between power service reliability and secure key rate (SKR). Therefore, a dynamic power regulation (DPR) scheme is designed in this paper. Based on the ε -greedy algorithm, the transmitted power can be dynamically adjusted according to the condition of quantum key pool (QKP), so as to jointly optimize the relia-bility of power service and SKR. The results show that DPR scheme can ef-fectively reduce the probability of key resource depletion or overflow in QKP while ensuring the reliability of power service.

Keywords: Quantum Communication · Quantum Key Distribution · Distribution Network · Noise Suppression

1 Introduction

With the advent of the Internet of Things era, the scale of the network is increasing, the boundaries are becoming increasingly blurred, and the threats are increasing. At the same time, with the emergence of new technologies such as supercomputers and quantum computers, the third-party eavesdroppers have greatly improved their computing power and password cracking ability.

The existing asymmetric encryption algorithms based on computational complexity will be unreliable in the future [1]. Therefore, there is an urgent need to research new encryption methods to resist the security threats and better ensure the communication security. The security of quantum communication technology is guaranteed by the three physical laws of quantum mechanics which are Heisenberg uncertainty principle, measurement collapse principle, and quantum state non cloning principle [2]. Among them,

QKD has made significant breakthroughs and progress in theoretical research, laboratory measurements, field deployment and networking, and many other aspects as one of the important communication methods.

According to the three laws of quantum mechanics, in the process of quantum key distribution, it is impossible for eavesdroppers to eavesdrop on the information carried in the transmitted signal without being discovered by the quantum communication system. In addition, the combination of quantum key distribution technology and “one secret at a time” technology can theoretically realize remote secure communication of information between two users [3]. In terms of the protocol of QKD, in 1984, Bennett and Brassard proposed the BB84 protocol [4], which was named after his two names and the year in which the protocol was proposed. In the following decades, scholars around the world have improved and proposed various QKD protocols, laying a solid theoretical foundation for the development and promotion of QKD technology. In 1991, Artur Eckert proposed E91 protocol based on quantum entangled states, the security of which was guaranteed by bell theorem [5]. In 1992, Bennett proposed B92 protocol, only two quantum states are used to achieve key distribution, which is two fewer than the number of quantum states in the BB84 protocol, and the process of quantum key distribution is simplified at the cost of 50% transmission efficiency [6]. In 2012, in response to the vulnerability of third-party eavesdroppers using the imperfect measuring devices in the OKD system to attack, Lo proposed the Measurement Device-Independent Quantum Key Distribution protocol (MD-QKD), In 2012, in response to the vulnerability of third-party eavesdroppers using the imperfect measuring devices in the OKD system to attack, Lo proposed the Measurement Device-Independent Quantum Key Distribution protocol (measurement-Device-independent Quantum Key Distribution). MD-QKD completely solves the vulnerability problem caused by the imperfect measurement end of the QKD system [7]. It breaks through the limit of key generation rate-distance of the original QKD protocol [8]. Xiongfeng Ma, Xiangbin Wang and Guangcan Guo respectively proposed PM-QKD, SNS-TF-QKD and phase-free TF-QKD protocol [9].

The power communication network is a specialized communication network within the power industry, originally built to meet the needs of power system operation and management, safety production control, dispatch and command automation, etc. It is an important infrastructure of the power system and plays an important role in the safe and reliable operation of the power system, information communication management, automation management, intelligent monitoring, [10]. There are many types of services in the power communication network, and the requirements for security and reliability are high. QKD technology has excellent security encryption performance. Integrating QKD technology with the power communication network will be an important direction for future power communication network technology.

In 2013, Ruina Zhang proposed a typical application example of fiber QKD in smart grid, and prospected the research direction and application scenarios of QKD in power system [11]. In 2017, Georgi Bebro proposed a solution to the security problem of the communication network between the user and the control center in the smart grid. The security encryption of smart grid is discussed from two aspects of security and resource economy by selecting a suitable QKD protocol [12]. In 2018, Maninder Kaur proposed a mutual authentication scheme in smart grid between users and servers based on QKD to

realize mutual authentication [13]. In 2019, William Lardier proposed an open-source co-simulation platform that researchers can use to conduct system simulations of different power network topologies, QKD protocols, and attack threats [14].

In the quantum secure distribution communication network, this paper proposes a dynamic power regulation (DPR) scheme. Based on the ε -greedy algorithm, the transmitted power can be dynamically adjusted according to the condition of quantum key pool (QKP). This reduces noise interference on the quantum signal, improves the security key generation rate, and achieves the goal of jointly optimizing the reliability and security key generation rate of power services.

2 System Model

The power communication access network provides a way for distribution terminal users and electricity business terminals to access the power system. The power communication access network mainly includes 10 kV power communication access network and 0.4 kV power communication access network. The existing fiber optic communication access network is mainly based on two technologies which are Ethernet Passive Optical Network (EPON) and Industrial Ethernet, Fig. 1 shows the EPON network structure. The Optical Line Terminal (OLT) is typically situated on the substation side, while the Optical Network Unit (ONU) is linked to the distribution terminal, which encompasses the feeder terminal unit (FTU), remote terminal unit (RTU), and other pertinent components.

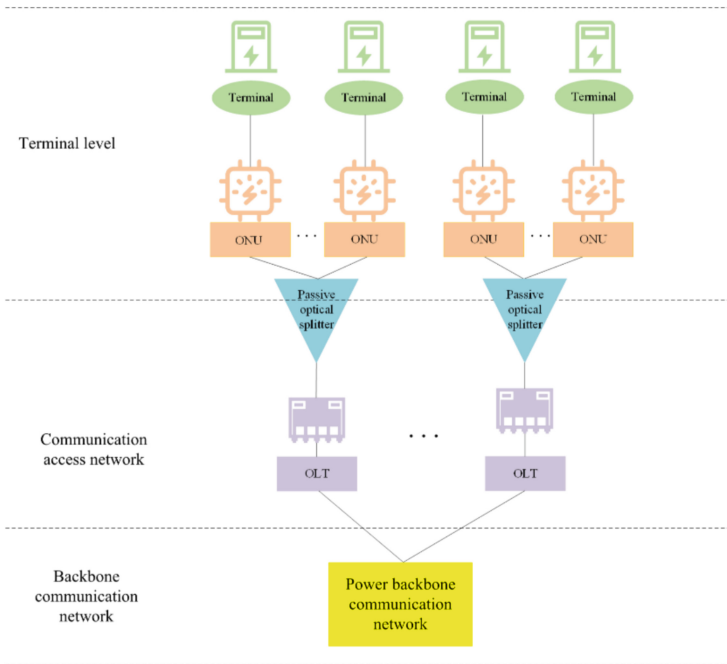


Fig. 1. System Model

3 The Proposed Scheme

In the process of integrating QKD with the distribution communication network, simply reducing the transmission power of classical signals to the minimum value will reduce the power budget reserved by the system for additional fiber attenuation, thereby affecting the reliability of power services. Meanwhile, due to the lack of consideration for the remaining key amount in the QKP, this can easily lead to overflow or depletion of key resources. To address these issues, assuming that the transmission power of classical signals can be reconfigured at different times, the goal of the DPR scheme is to adjust the transmission power of uplink classical signals based on the current key resource situation to optimize the reliability and security key generation rate of power services jointly.

In the distribution communication network, there are usually two working modes for power information collection, namely centralized collection and random collection. The size of the data signal has the characteristic of periodic variation. Therefore, this paper considers the characteristic of flow data and adjusts the signal transmission power corresponding to flow data according to the optimization objective during the period of collecting distribution terminal data in the substation. The optimization objectives of the proposed DPR scheme are as follows:

$$\text{Maximize : } \sum_{t=0}^T (1 - w) \cdot PI(P) + w \cdot PI(SKR) \quad (1)$$

where $PI(P) = \frac{P^T(BR) - P_{\min}^T(BR)}{P_{\min}^T(BR)}$, calculates the percentage increase in actual transmission power compared to the minimum error free transmission requirement, which can measure the additional power budget. $PI(SKR) = \frac{SKR(P^T) - SKR_{\min}}{SKR_{\min}}$ represents the percentage improvement of SKR compared to the system when the transmission power of the classical signal is adjusted to P^R . $P_{\min}^T(BR)$ is the minimum transmission power calculated when the corresponding data rate is BR and $Q=6$, where Q corresponds to the reliability of power services, and reflects the transmission quality of services. T is the period for data collection. $P^T(BR)$ is the transmission power of a classical signal with a data rate of BR , which will be adjusted according to the optimization objective of the scheme. The value of SKR depends on the channel allocation of the QKD system and the P^T under the premise of performance parameter determination. $w = \frac{S_{\max} - S}{S_{\max}}$ is a parameter set based on the current key resource which is represented as S .

The specific steps of the DPR algorithm are described as follows:

1. Firstly, determine the upper limit of classical signal transmission power P^T by referring to the performance parameters of the transmitter and the classical value of the transmission signal power, and calculate the minimum transmission power $P_{\min}^T$ corresponding to the current flow data size based on the theoretical formula. Each ONU corresponds to a transmission power set P , and the lower limit $P_{\min}^T$ in the power set is calculated based on the current flow data size of the ONU.
2. When the value of S is small, it means that the key resources are lack. In this case, priority should be given to optimizing the security key generation rate. Therefore, the value of w is relatively large to reduce the power of the classical signal and improve the key generation rate.

3. When the value of S is large, the current key resources are sufficient, and the probability of key resource over-flow in the QKP increases. At this time, the reliability of the classical signal will be prioritized, and the value of w is relatively small, in order to improve the power of the classical signal.
4. The improved ε -greedy algorithm is used to find the transmitting power that can maximize the optimization target. The purpose of the ε -greedy algorithm is to achieve a basic balance between exploration and exploitation, and in order to achieve this goal, the algorithm introduces the key parameter ε . In each experimental time slot, ε -greedy algorithm performs the exploration operation with the probability of ε , that is, it is possible for the algorithm to select an option whose average reward is not previously specified. The exploitation operation is performed with a probability of $1-\varepsilon$, that is, the algorithm has a probability of $1-\varepsilon$ to select the option with the currently defined maximum average reward.

The steps are as follows:

4.1. Divide the adjustable power range $[P_{\min}^T, P^T]$ of each ONU into d equal parts, and calculate the power adjustment gradient value $\Delta = \frac{(P_{TX} - P_{TX}^{\min})}{d}$. Therefore, the set of signal transmission powers that can be selected is $P = \{P_{TX}^{\min}, P_{TX}^{\min} + \Delta, P_{TX}^{\min} + 2\Delta, \dots, P_{TX}\}$, each item in the set is an action.

4.2 Let $B_t(a)$ be the average reward for choosing a for the t time, if the average reward after $t - 1$ attempts is $B_{t-1}(a)$, and the reward after t attempts is v_t , then the average reward is updated to

$$B_t(a) = \frac{1}{t}((t-1)B_{t-1}(a) + v_t) = B_{t-1}(a) + \frac{1}{t}(v_t - B_{t-1}(a)) \quad (2)$$

where $v_t = (1-w) \cdot PI(P) + w \cdot PI(SKR)$.

4.3 In the traditional ε -greedy algorithm, it is a fixed value. However, if the number of attempts is very large, the reward of each rocker arm can be approximated well after a period of time, and exploration is no longer needed, thus reducing the efficiency of the system. Based on this, we assign ε to $\frac{1}{t}$, so that as the number of attempts increases, it gradually decreases, reducing the system overhead while ensuring accuracy.

4.4 Select the Signal Transmitting Power that Can Make the Average Reward Maximum for Signal Transmission.

4 Performance Analysis

This paper compares the performance of DPR scheme, Minimum Power Control (MPC) scheme, and Without Power Control (w/o PC) scheme. The MPC ensures that the power of the power communication optical signal always meets the minimum theoretical requirement for error free transmission, while keeping $BER = 10^{-9}$, the transmission power of the power communication optical signal varies with the size of the collected traffic data. The w/oPC scheme ensures that the transmission power of a classical signal is always fixed at the typical power value such as 0dBm, and the transmission power does not change with the size of the collected traffic data from the end user.

This article simulates the network situation of 8 sets of ONU devices and 2 ONU groups, the QKD protocol used is the BB84 protocol, the QKD system parameter facilities are shown in Table 1. Assuming that the key consumption rate of each QKP in the network architecture is 1kbps.

Table 1. QKD system parameter facilities.

Parameter	Value
Single photon detector repetition rate	10 MHz
Single photon detector efficiency	0.1
Single photon detector dark counting	10–6
Single photon detector door width	1 ns
Insertion loss of AWG	5 dB
QKD receiver narrowband filter bandwidth	15 GHz
QKD protocol	BB84

Key resources: As shown in Fig. 2, the w/o PC scheme suffers from significant noise interference and the key generation rate cannot always meet the key consumption needs of the network. Within one collection cycle, the key capacity of the QKP is rapidly depleted. In the MPC scheme, the key generation rate is much higher than the key consumption rate, and the key resources overflow after one collection cycle. In the DPR scheme, the key resources are always kept at a relatively secure level, ensuring that the system can provide an appropriate amount of key resources at any time without causing waste of secure keys. Therefore, the DPR scheme has more advantages in the

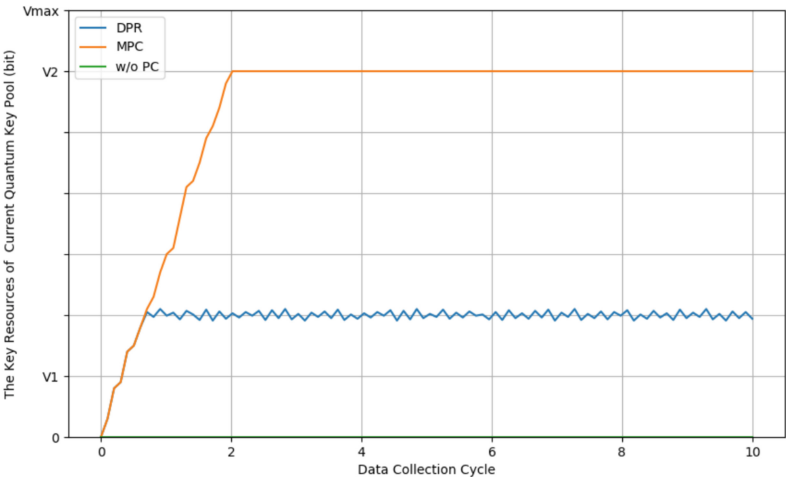


Fig. 2. Changes of key resources in different collection periods

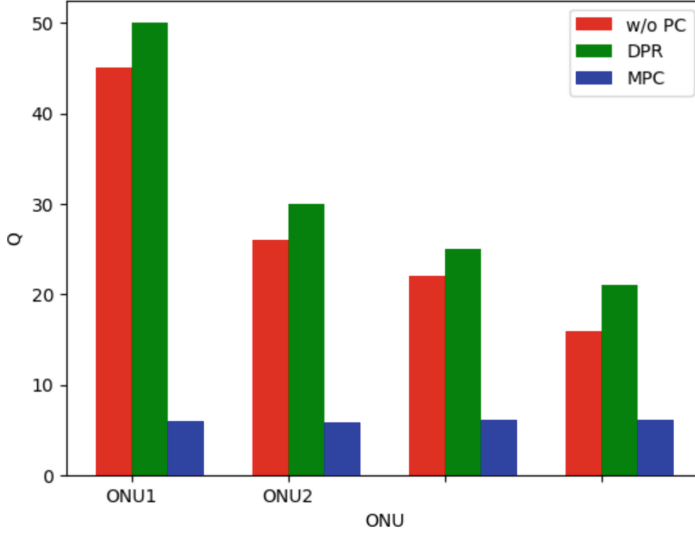


Fig. 3. Q factor of uplink signal in random collection mode under different schemes

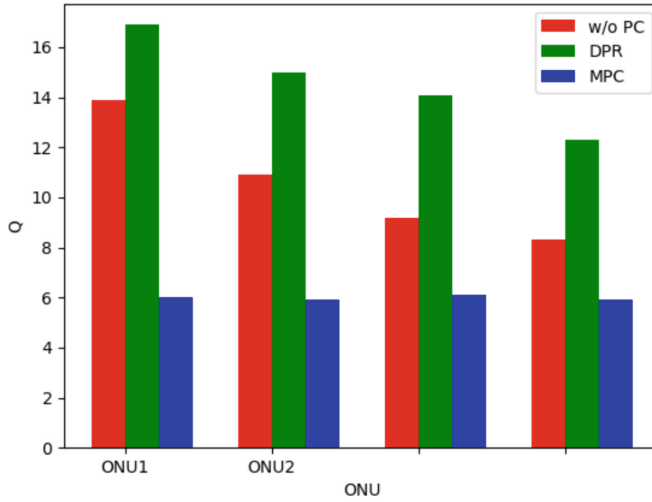


Fig. 4. Q factor of uplink signal in centralized collection mode under different schemes

management of QKP, always ensuring that the system can provide an appropriate amount of key resources without causing waste of key resources.

Reliability of Power Business: This paper evaluates the Q of uplink signals under the coexistence of quantum signals and power communication optical signals. We calculate the average Q for different acquisition modes. As shown in Fig. 3 and Fig. 4, the average zero factor of different ONU groups in the DPR scheme was improved by at least 81.24% and 327.23% in the centralized collection mode and random supplementary collection

mode, respectively, compared to the MPC scheme, with the highest being 129.29% and 649.56%.

5 Conclusion and Future Work

This paper proposes a dynamic power regulation (DPR) scheme. Based on the -greedy algorithm, the transmitted power can be dynamically adjusted according to the condition of quantum key pool (QKP). This reduces noise interference on the quantum signal, improves the security key generation rate, and achieves the goal of jointly optimizing the reliability and security key generation rate of power services. The simulation results show that compared with MPC and w/o PC schemes, the DPR scheme can effectively reduce the probability of key resource depletion or overflow in QKP, maintain key resources at a safe and available level, and ensure the reliability of power business.

Acknowledgement. This Paper is Supported by Research on Key Technologies of High Security Protection of Basic Information Facilities for New Power System (GDKJXM20230420).

References

1. Shor, P.W.: Algorithms for quantum computation: discrete logarithms and factoring. In: Proceedings 35th Annual Symposium on Foundations of Computer Science. IEEE, pp. 124–134 (1994)
2. Wei, C.: Experimental Study of Quantum Key Allocation in Optical Fiber. University of Science and Technology of China, Anhui (2008)
3. Wootters, W.K., Zurek, W.H.: A single quantum cannot be cloned. *Nature*, 1982299(5886), pp. 802–803
4. Bennett, C.H.: Quantum cryptography: public key distribution and coin tossing. In: Proceedings of IEEE International Conference on Computers, Institute of Electrical and Electronics Engineers, pp. 7–11 (1984)
5. Ekert, A.K.: Quantum cryptography and Bell's theorem. *Quantum Measurements in Optics*. Springer, Boston, MA, pp. 413–418 (1992)
6. Yang, Y.G., Sun, S.J., Xu, P., et al.: Flexible protocol for quantum private query based on B92 protocol. *Quantum Inf. Process.* **13**(3), 805–813 (2014)
7. Lo, H.K., Curty, M., Qi, B.: Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **108**(13), 130503 (2012)
8. Lucamarini, M., Yuan, Z.L., Dynes, J.F., et al.: Overcoming the rate-distance limit of quantum key distribution without quantum repeaters. *Nature* 2018557(7705), 400–403
9. Ma, X.F., Zeng, P., Zhou, H.: Phase-matching quantum key distribution. *Phys. Rev. X* **8**(3), 031043 (2018)
10. Gao, H.: Reliability Study of Electric Power Communication Network. North China Electric Power University, Hebei (2009)
11. Zhang, R.R., Chen, X.: Prospects of fiber quantum key distribution technology for power systems. In: International Conference & Exhibition on Electricity Distribution. IET, pp. 0154–0154 (2013)
12. Bebrov, G., Dimova, R., Pencheva, E.: Quantum approach to the information privacy in smart grid. In: 2017 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM) & 2017 Intl Aegean Conference on Electrical Machines and Power Electronics (ACEMP). IEEE, pp. 971–976 (2017)

13. Kaur, M., Kalra, S.: Security in IoT-based smart grid through quantum key distribution. *Advances in Computer and Computational Sciences*, Springer, Singapore, pp. 523–530 (2018)
14. Lardier, W., Varo, Q., Yan, J.: Quantum-sim: an open-source co-simulation platform for quantum key distribution-based smart grid communications. In: 2019 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm). IEEE, pp. 1–6 (2019)



From Machine Intelligence to Machine Ethics: An Analysis of the Practical Path of Artificial Morality

Jingjing Zhao¹(✉) and Xingtong Liu²

¹ Undergraduate School, National University of Defense Technology,
Changsha 410073, China
jjzhao1983@126.com

² College of Electronic Science, National University of Defense Technology,
Changsha 410073, China

Abstract. Intelligent machines are increasingly appearing in daily ethical scenarios, and there is an urgent need to consider the ethical module design of intelligent machines and give them moral considerations. In current research on machine ethics, the most discussed approaches are top-down and bottom-up approaches, both of which have their own rationality but have limitations in computability. Therefore, scholars have begun to study the design of hybrid paths, which involves how to better integrate the two approaches. This paper analyzes the problems encountered in the specific implementation of machine morality and proposes that both design and integration issues can adopt the method of transfer learning. It suggests improving the good ideas and suitable model frameworks applied in machine intelligence and then introducing them into the design and implementation of machine morality. Specifically, on one hand, the machine intelligence implementation method of “autonomous mental development” can be transferred to the framework design of “autonomous moral development”. On the other hand, considering the characteristics of moral issues themselves, factors such as emotion, sociality, semantic understanding, and the impact of consciousness on human moral decision-making should be taken into account and integrated into the machine morality module. Based on the above two considerations, this paper proposes a feasible modeling framework for machine ethics, providing a new design idea and practical reference for the realization of machine ethics.

Keywords: Machine Ethics · AI Ethics · Artificial Intelligence

1 Introduction

Machine intelligence and machine ethics are both important branches of artificial intelligence research. In today’s literature and papers discussing artificial intelligence, ethical issues related to intelligent machines are generally addressed. However, there is still controversy among scholars regarding the specific definition and delineation. Yanyong Du [1] believes that although Artificial Intelligence and Robotics are two different concepts and have many differences. However, from the perspective of technological ethics

research, the basic viewpoints, theories, and methods of the two are almost universal. Weiwen Duan [2] proposed the need to distinguish between machine ethics and robot ethics. He believes that robot ethics emphasizes the responsibility and ethical principles that humans should follow when interacting with robots without autonomy, while machine ethics attempts to endow machines with moral behavior and consider their potential to develop into moral agents. Xue Yu and Qian Wang [3] have systematically sorted out the division of ethical subjects related to machine ethics among scholars. They believe that emphasis should be placed on explicit moral embedding rather than implicit moral implications. They emphasize embedding ethical programs or principles in machines to achieve explicit moral effects of intelligent machines. Although scholars have different views on the specific definition and delineation, the analysis of ethical issues related to intelligent machines is generally based on the work of James Moor. James Moor [4] defined five types of ethical agents based on the degree of involvement of ethical factors. The first type of subject is normative agents, which can be any technical subject capable of performing tasks and completing predetermined work, and the user determines its moral attributes. The second type of subject is ethical-impact agents, who not only perform predetermined tasks but also generate moral influence. The third type is implicit ethical agents, whose behavior implies ethical and moral considerations. Its designer will design based on ethical and moral principles, and the moral subject itself can potentially exhibit moral behavior. Its design actually implies the designer's materialized morality, thus exhibiting moral intentions in its behavior. The fourth type is explicit ethical agents, which are able to identify information related to ethics and morality, and evaluate various possible behaviors based on the ethical and moral mechanisms built into them, calculating and selecting the optimal behavior. The fifth type is full ethical agents, which can highly simulate human thinking and make ethical judgments and moral choices in specific contexts. In theory, the operation of such agents can be completely independent of humans, and currently only exists in science fiction works. Wendell Wallach and Colin Allen further focus on the deep integration of "intelligent machines" and "morality" in reality [5]. They proposed the concept of ethical machines, believing that artificial intelligence expands the scope of autonomous agents, so when designing these agents, it is necessary to make them respect the values and laws that a wider range of human moral subjects need to respect. They further narrowed down the scope of their research and emphasized the need to increase the ethical and moral attributes of machines. The moral subject they focused on was actually what James Moor referred to as the explicit moral subject, which is, giving machines moral principles or moral procedures so that they can find solutions when faced with moral dilemmas, and making them operate in a morally responsible way, and making their own moral decisions. They believe that with the enhancement of Machine Intelligence and automation, humans can better utilize the intelligent functions of machines. In extremely dangerous environments, the advantages of machines will be further highlighted. Therefore, they emphasize the need to embed ethical programs in intelligent machines to achieve their ethical decision-making functions.

In summary, compared with general applied ethics, the research on ethical issues of intelligent machines shows stronger interdisciplinary characteristics. With the continuous improvement of the decision-making ability of intelligent machines and robots,

they are gradually forming new challenges to traditional moral philosophy. On the one hand, traditional ethical principles or moral norms have shown imperfections in artificial intelligence, prompting ethicists to think deeply and revise existing moral concepts and theories. Therefore, the theoretical research work of artificial intelligence moral philosophy needs to be strengthened. However, scholars [6] generally believe that the moral philosophy of artificial intelligence is not a negation or replacement of moral philosophy, but an extension and development of traditional moral philosophy. On the other hand, intelligent autonomous systems have increasingly appeared in daily ethical scenarios, and there is an urgent need to consider the design and implementation of machine ethics. When making decisions, intelligent machines must take into account both external and internal values in their choices and actions. From an execution perspective, all attempts at machine ethics need to evaluate the feasibility of transforming theory into computer programs, truly combining theoretical research and practice of machine ethics to make responsible artificial intelligence. This article mainly focuses on the specific implementation of ethical decision-making in intelligent machines, aiming to study the advantages and disadvantages of different approaches to machine ethics implementation from a feasibility perspective. It discusses the implementation of machine ethics from a comprehensive perspective, combining top-down and bottom-up approaches to construct a working system, in order to promote the practical implementation of “machine ethics”.

2 Practical Exploration of Machine Intelligence and Machine Ethics.

2.1 Three Major Paths of Artificial Intelligence

Human attention to intelligence has a long history. In ancient times, intelligence was usually considered as a general term for wisdom and ability. It is said that “wisdom” and “ability” are two relatively independent concepts, for example, in “Xunzi”, “What is known by people is called ‘knowledge’, and when knowledge combines with things, it is called ‘wisdom’. Therefore, what can be done by a person is called ‘instinct’, and what can be done in conjunction with others is also called ‘ability’”. Here, “wisdom” emphasizes cognitive activities more, while “ability” leans towards practical activities; Secondly, intelligence itself is a whole; The combination of intelligence and ability is regarded as a symbol for evaluating people. In recent times, “intelligence” emphasizes more on “brainpower, comprehension, and wisdom”, explaining intelligence from the perspective of power, which precisely demonstrates the dynamic mechanism by which “intelligence” can be transformed into “power”. This concept has also influenced the development of artificial intelligence. Currently, although there is no consensus in the academic community on the definition of artificial intelligence, according to Lucci and Koppc’s viewpoint in the classic textbook “Artificial Intelligence” [7], artificial intelligence is composed of people, ideas, methods, machines, and outcomes; Firstly, humans make up artificial intelligence. Humans have ideas and turn them into methods; These methods can be represented by algorithms, heuristics, programs, or systems that serve as the backbone of computation; Finally, the products obtained from these machines

programs are called “results”, and each result can be measured based on its value, effectiveness, efficiency, and other aspects. So, the artificial intelligence we are discussing today is mostly measured from the perspective of results in terms of its “intelligence” level, which is often the case in multiple disciplines involving artificial intelligence.

But in the early stages of artificial intelligence research, people’s ideas were more focused on the essence and simulation of intelligence. At first, there was a constant debate in the scientific and philosophical communities about whether machines could achieve intelligence. Amidst these debates and doubts, artificial intelligence is constantly maturing and developing. Today, looking back, we realize that research in artificial intelligence not only allows us to better understand ourselves, but also enables us to empower machines to solve problems that have relied solely on human intelligence for thousands of years. There have been some disagreements in academia since the birth of artificial intelligence on how to make machines intelligent, leading to a debate among the three major schools of thought. Connectionism believes that the foundation of intelligence is built upon the true nervous system. They hope to clarify the elusive characteristics of human intelligence through the study of neuronal interactions and aim to solve intelligence problems by replicating the connections between numerous neurons. They do not emphasize centralized storage of information, but rather hope that intelligent machines can distribute knowledge and memory across all connections like the human brain. Functionalists advocate for “only seeking results, not means”. They believe that intelligence is a characteristic of living organisms and has no essential connection with the organization of the body. Intelligence can exist in any system as long as there is a correct causal relationship between the components of the system, and these components can be neurons, silicon chips, or any other substitute. Behaviorists believe that intelligence does not require knowledge, representation, reasoning, and that artificial intelligence can evolve like human intelligence by interacting with its surrounding environment. They conducted research on adaptive, self-organizing, and self-learning issues by simulating the intelligent behavior and role of humans in the control process. Three different approaches have shown their respective potentials in solving different problems, while also exposing corresponding questions. On the development path of artificial intelligence, they complement each other’s strengths and weaknesses, gradually moving from separation to integration, thus giving birth to Mechanism. Mechanism [9] emphasizes exploring the essence of intelligence from the common mechanisms of intelligent generation, unifying the three in a common framework, and complementing each other to achieve artificial intelligence.

In many specific studies on artificial intelligence, we have found that the initial research on many fundamental issues usually manifests as general questions about intelligence, with a focus on understanding ourselves as humans. However, as research deepens, especially with the continuous improvement of Machine Intelligence performance, people begin to realize the importance of this issue in practical applications and shift their focus from the process to the results. That is to say, people pursue practical effects more than just understanding our own dimensions. The focus of understanding the issue of “intelligence” in artificial intelligence research has gradually shifted from leaning towards “wisdom” to pursuing “ability”, and the pursuit of “intelligence” that can be “implemented” has become an important dimension in related research. It is not only

necessary to explain intelligence from the perspective of “power”, but also to derive “power” from intelligence, which has gradually become a characteristic in many studies of artificial intelligence. This understanding has also played a guiding role in the development of artificial intelligence in the future. Especially in recent years, with the development of deep learning, treating intelligent problems as “black box” has become a widely adopted research model. Behind the performance improvement, it has also attracted people’s attention to the issue of “Knowing how something happens is more important than knowing what happens”. The interpretability of artificial intelligence has also attracted widespread attention, which is a key issue that must be faced in developing responsible artificial intelligence.

2.2 Three Approaches to Realizing Machine Ethics

The current consensus in academia regarding the development of intelligent machines is that a good intelligent machine should not only be able to detect potential harm or negligence of responsibility, but also take measures to avoid or reduce negative outcomes, comprehensively considering both external and internal values in its choices and actions. There are currently three main approaches to the feasibility of machine ethics: the “top-down” approach, the “bottom-up” approach, and the “hybrid” approach.

Different disciplines use the concepts of “top-down” and “bottom-up”, but there are differences in their specific research applications. The “top-down” approach in ethical research typically focuses on top-down norms, standards, and theoretical approaches to moral judgment. They generally focus on the application of broad moral principles to specific examples. Therefore, theoretical standards, norms, and principles are the basis for evaluating the morality of actions. In engineering, the concept of “top-down” is used in a different sense, usually referring to an analytical method for completing tasks, which decomposes complex tasks into simple subtasks in a top-down manner to achieve task objectives. In the discussion of machine ethics, two different “top-down” concepts are combined. From a fusion perspective, designing a machine ethics module means first determining a specific ethical theory, then analyzing the information and program needs to implement this theory in computer programs, and then applying the analysis to the design subsystems and their interconnections. This approach emphasizes that designers should anticipate various situations as much as possible and provide expected results guided by rules within the scope of an intelligent machine usage environment. From the existing research results, the discussions involved in this approach mainly include the computability and practicability of rule-based and responsibility based ethical concepts, as well as the possibility of calculating the net effect of an action required by the effects theory of ethics. As Wendell Wallach and Colin Allen proposed [5], the theory of ethics provides a solution. If moral principles or norms can be clearly stated, then the requirement for intelligent machines to take moral actions can be transformed into specific questions about how intelligent machines should comply with norms. At this point, the intelligent machine needs to calculate whether its behavior is allowed by the rules and convert moral rules into formal algorithms. Because intelligent machines need to rely on top-down rules to provide information and reasons for their behavior, designers need to capture the rules of human morality. However, from the current research and implementation situation, the prominent problem encountered by the “top-down” approach

is: how to implement the specific calculations required for theory. Is there a computer that can collect and compare all the necessary information for the full application of theory in real-time? This question is particularly sharp for the approach of Utilitarianism, as the outcome of any action is bound to be infinite in time or space. In the design based on Deontological ethics, conflicts of rules arise, necessitating the development of a framework to resolve these conflicts. Practically, it's challenging to envision that moral rules can provide a thorough and detailed decision-making process. Currently, top-down designs rely on heuristics to inform decision-making and provide evaluative experts with the information for critical analysis.

The bottom-up approach typically applies methods of learning, development, or evolution, with the goal of deriving moral competence from general intelligence characteristics. In the sense of ethics, the bottom-up approach to obtaining morality is to treat standardized values as implicit content in the activities of the subject, rather than being clearly expressed by abstract theories. In the bottom-up approach to obtaining machine ethics, the focus is on creating an environment where intelligent machines can explore their own behavior patterns. The ethical principles in this approach are discovered or constructed, rather than pre-designed. The underlying theory used in the bottom-up approach can be understood as a means to clarify tasks for the system, rather than finding execution methods or control structures. The current consensus in academia is that both genetic evolution and postnatal learning theories contribute to the development of machine moral modules, as the process of moral production itself is the result of the combined effects of innate nature and acquired learning. Bottom-up strategies are expected to generate indispensable technologies and standards for the overall design of intelligent machines, but these strategies still face difficulties in reality: firstly, they are full of trial and error and are a slow process; Secondly, it is not yet clear what the appropriate evolutionary goals are, what appropriate criteria can determine which intelligent machines can be complex and mutated, and how these goals effectively define a system; In addition, the initial learning ability of existing artificial intelligence systems is not yet as strong as the rich adaptive learning skills exhibited by young children, and there is still a long way to go in guiding intelligent machines to engage in moral learning. So, from a practical perspective, the learning approach of artificial intelligence has not yet been used to develop ethics at this stage. However, in terms of research trends, ethicists and engineering design researchers have begun to propose that the evolution of learning or simulation should be combined with more bottom-up approaches in design systems, adopting a hybrid development model.

It should be noted that the initial distinction between top-down and bottom-up approaches was too simple to cover all the challenges that machine ethics face in design. This is a fact, both in terms of engineering design and ethical theory. In practical applications, we need to combine top-down and bottom-up methods to construct a working system. However, scholars have not reached a consensus on how to integrate the two specifically. For example, from the perspective of human morality, morality is actually a mixture of top-down and bottom-up approaches. The acquisition of quality characteristics is a bottom-up process, while the emergence of regulatory awareness is mostly a top-down process. In recent years, many empirical literature on moral psychology have

also emphasized a person's innate physical or psychological abilities [5]. So, rationality, emotion, sociality, semantic understanding, and consciousness are all important for human moral decision-making. However, whether these factors are essential for intelligent machines, if so, whether they can be implemented in computers, and how they can be integrated are also questions that must be considered.

3 From Machine Intelligence to Machine Ethics

Alan Turing once had this thought: instead of trying to write a program to imitate the thinking of adults, why not directly try to write a program to imitate the thinking of children? If we subsequently undergo a suitable educational process, we will acquire the adult brain as a result. At that time, Turing did not specifically consider the issue of machine ethics. He was thinking about whether computers could produce original behavior. But today we will see that this approach is not only suitable for the study of machine intelligence problems but also for thinking about the feasibility of machine ethics. Because, whether it is natural evolution or artificial construction, intelligent machines with moral considerations require certain abilities to learn and discover the norms of the place they are placed in. Throughout their entire lifecycle, intelligent machines actually need to absorb feedback on the moral acceptability or unacceptability of their behavior. Currently, despite the rapid development of artificial intelligence, we must face the fact that there is no effective learning machine model that can approach the rich diversity of actual biological learning mechanisms in any situation. Therefore, the Turing ideal model faces great difficulties in practice, but we have found through research that existing attempts and achievements in artificial intelligence can be introduced into the implementation of machine ethics through transfer learning, as there are many similarities between the two in a specific implementation.

3.1 The Similarities Between Machine Intelligence and Machine Ethics at the Practical Level

In the design of artificial intelligence, on the one hand, we need to fully consider and improve its specific performance, that is, the problem of intelligence; On the other hand, it is also necessary for them to abide by human moral norms, that is, moral problems. An excellent artificial intelligence design requires a comprehensive consideration of both intelligence and ethics. At present, artificial intelligence systems are not yet able to evolve and develop a sense of morality on their own, and their instrumental nature may lead to deviations in their utilization. Therefore, researchers and designers of artificial intelligence should strive to reduce the possibility of their abuse and misuse before the product is officially launched. As proposed by Lun Li and Baoxue Sun [8], the design of artificial intelligence ethics needs to be carried out from two aspects: first, before designing and developing a certain artificial intelligence product or service, designers and manufacturers should have a clear value positioning and be able to have an overall prediction of its ethical and social impact; Secondly, if there is a value deviation in the process of providing services to consumers through artificial intelligence, the built-in correction mechanism of the system can effectively prevent the harm from continuing to

occur and prevent the occurrence of danger. If highly autonomous intelligent machines lack ethical constraints and value consistency, the consequences are even more worrying. The overall value positioning of ethical design for artificial intelligence is to design ethical guidelines for AI systems and embed good value standards into the system in order to ensure that AI better serves humanity, rather than making humans subject to AI. Therefore, the overall goal of intelligence and moral problems is consistent, both striving towards the ultimate goal of goodness, beauty, and excellence.

There is also a debate over whether morality can be calculated in the implementation of machine ethics. At present, the implementation of machine ethics is to transform human moral behavior into a computable form for representation, which is actually a recognition of ethical computability. On the issue of human intelligence, the debate over whether it is computable has a long history. Nowadays, this debate is widely present in fields such as intelligence, cognition, and ethics. Computationalism believes that the human mind can be understood as a form of computation. Human thinking and mind are both controlled by rules and can be explained using computational terms. According to computationalism, ethical reflection, as an important part of human mental activity, can also be demonstrated through computational methods. In the field of artificial intelligence, many algorithms are currently implemented through computation. When it comes to the question of whether intelligence can compute, in terms of the computational capacity of intelligence, John Searle had already used the Chinese room argument to raise doubts: while computers can perform symbolic processing of numerical, textual, and other data, they do not truly comprehend such information. Similarly, when dealing with ethical and moral problems, intelligent machines can use symbols and numbers to represent ethical and moral behavior, but do they really understand moral behavior and moral intentions? Alternatively, in the issue of machine ethics, humans do not necessarily require machines to understand ethical behavior, but rather hope to transform ethical behavior into a way that artificial intelligence systems can handle, using cognitive logic and moral logic calculation methods to enable machines to make ethical decisions, thereby further increasing public safety, health, and well-being. Although there is controversy among scholars regarding the computability of morality, there is a general consensus on designing artificial intelligence that conforms to human ethical standards. Therefore the current focus of research is still on how to design a more reasonable plan.

3.2 Drawing Inspiration from the Practice of Machine Intelligence to Improve Machine Ethics

As early as 2001, Weng and some other scientists jointly published paper in Science, proposing a new design concept for artificial intelligence called “autonomous mental development” [10]. Under the control of its internal developmental program, the machine system can autonomously interact with the external environment in real-time through the use of sensors and drivers, continuously accumulating knowledge in a learning manner, and growing into a mature intelligent machine capable of handling multiple different tasks to achieve the development of the machine’s mind, and accumulate memory and knowledge during the process of mental development. The concept of autonomous mental development is by no means groundless, and subsequently they developed a series

of robots such as SAIL and David, which have the ability of mental development. After more than 20 years of development, the design concept of “autonomous mental development” has not only been widely recognized in the industry, but its performance has also steadily improved, allowing us to see the possibility of directly trying to write a program to imitate children’s thinking and then receiving a suitable educational process to make machines intelligent.

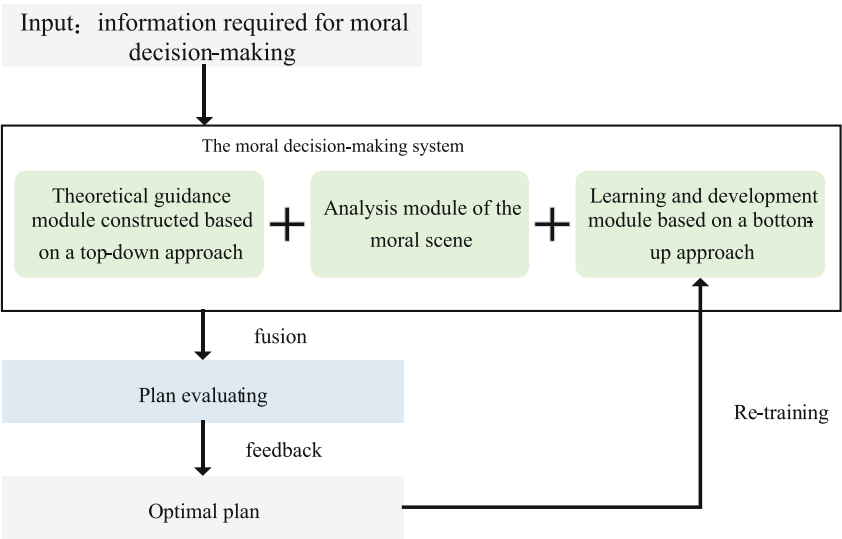


Fig. 1. The Framework of Machine Ethics Model based on Autonomous Mental Development.

In the development of the theory of autonomous mental development, a series of core algorithms have emerged, from Incremental hierarchical discriminant regression [11] (IHDR) to Developmental Network-2(DN-2) [12]. The core goal of this type of algorithm is to build decision systems in high-dimensional space incrementally through online real-time learning. Autonomous mental development robots are designed and developed according to the principles of learning and evolution and have achieved significant breakthroughs. It should be emphasized that due to its simulation of human growth process, it has unique advantages in the interpretability of artificial intelligence. This also provides a reference and inspiration for the implementation of machine ethics. Compared with pre-collected ‘big data’, autonomous imitation greatly reduces the complexity of teaching, especially without the need for annotation of training data. Therefore, it is very suitable for building machine ethics modules. Based on this, we have designed a training design framework for machine ethics, as shown in Fig. 1. Firstly, collect the relevant information required for moral decision-making and input it into the moral decision-making system. The moral decision-making system mainly consists of three modules: one is the theoretical guidance module constructed based on a top-down approach; The second is a learning and development module based on a bottom-up approach; The third is the analysis module of the moral scene in which it operates; The fusion of the three produces a plan to be evaluated. The evaluation plan usually consists of multiple

plans, and the trainer scores and provides feedback on the series of plans to determine the optimal plan. The optimal plan is then fed back to a learning and development module based on a bottom-up approach, which is dynamically updated through DN-2 learning. As the algorithm foundation, DN-2 has the characteristics of incremental, non-iterative, real-time learning mode, and optimality (in the sense of maximum likelihood) of super Turing machines emerging from open real-world learning. In the construction of moral decision-making frameworks in specific scenarios, we can use small samples for training and dynamically update criteria in real-time through learning, which can better solve the problem of different design requirements for machine moral modules with different functions in different scenarios.

4 Conclusion

How can machine ethics better align with human ethics? Scholars generally believe that there will be two forms of machine ethics before the development of widely trusted moral actors by humans: operational ethics and functional ethics. Operational ethics is completely controlled by tool designers and users, and its design process is carried out with full consideration of ethical values. The forms of functional morality are more diverse, ranging from systems that can only act within acceptable behavioral standards, to intelligent types that can evaluate the important moral significance of their own behavior, all of which belong to functional morality. They have the ability to make moral decisions and respond to moral challenges on their own. However, due to the limitations of current technology, the functional ethics of machines still face many challenges.

The existence of challenges has not reduced people's determination to develop machines with moral attributes, and extensive research has been conducted on this issue, resulting in different practical approaches and ideas. The main issues involved in the top-down approach include the computability and practicability of rule-based and responsibility-based ethical concepts, as well as the possibility of calculating behavioral effects in the effects theory of ethics; The bottom-up approach applies methods of learning, development, or evolution, with the goal of deriving moral competence from general intelligence characteristics, with a focus on the self-evolution and learning issues of moral competence. Both designs have their own rationality, but also face many practical challenges, giving rise to the idea of a hybrid approach. Our research is based on a hybrid approach, drawing inspiration from the study of autonomous mental development in machine intelligence, and designing a hybrid framework with the hope of providing a new reference for the implementation of machine ethics.

Acknowledgement. This work was supported by the National Social Science Fund of China (The algorithmic risks and ethical review of artificial intelligence).

References

1. Du, Y.: Artificial Intelligence Ethics. Shanghai Jiaotong University Press, Shanghai (2020)
2. Duan, W.: The basic approaches and meaning in the ethical construction of robot. *Science and Society* 5(2), 35–46 (2015)

3. Yu, X., Wang, Q.: The value and limitations of the concept of “machine ethics”. *Studies in Ethics* 2016(4), 109–114 (2016)
4. Moor, J.: The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems* **21**(4), 18–21 (2006)
5. Wallach, W., Allen, C.: *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press (2009)
6. Li, L.: *Ethics of Artificial Intelligence and Big Data*. Science Press, Beijing (2018)
7. Lucci, S., Musa, S.M., Kopec, D.: *Artificial Intelligence in the 21st Century*. Mercury Learning and Information, Berlin, Boston (2022)
8. Li, L., Sun, B.: Enduing artificial intelligence “moral chips (conscience)”——four dimensional of ethical research of artificial intelligence. *Teaching and Research* 2018(8), 72–79 (2018)
9. Zhong, Y., et al.: *Introduction to Intelligent Science and Technology*. Beijing University of Posts and Telecommunications Press, Beijing (2006)
10. Weng, J., et al.: Autonomous mental development by robots and animals. *Science* **291**(5504), 599–600 (2001)
11. Weng, J., Hwang, W.-S.: Incremental hierarchical discriminant regression. *IEEE Transaction on Neural Networks* **18**(2), 397–415 (2007)
12. Weng, J.: An algorithmic theory for conscious learning. In: *Proceedings of the 2022 3rd International Conference on Artificial Intelligence in Electronics Engineering*, pp. 1–8. ACM, New York, NY, USA (2022)



The Frequency Bounds for Travelling Salesman Problem Based on Frequency Quadrilaterals

Yong Wang^(✉) and WenFa Zheng

School of New Energy, North China Electric Power University, Beijing 102206, China
wangyong100@163.com

Abstract. Given a weighted graph for TSP, a corresponding frequency graph is computed using the optimal paths with given endpoints (optimal paths for short). The edges in optimal Hamiltonian cycle (OHC) show special frequencies much bigger than those of most other edges. The average frequencies for edges and paths in OHC are studied as a frequency graph is computed based upon the frequency quadrilaterals or optimal 4-vertex paths. The lower frequency bounds for an OHC edge, two and three immediate OHC edges are proven to be $35/9$, $22/3$ and $35/3$, respectively in the worst average case. Moreover, the average frequency for all OHC edges is bigger than 4 and it is bigger than $13/3$ for big and large TSP. These findings are verified with experimental results.

Keywords: Travelling salesman problem · frequency graph · optimal path with given endpoints · frequency quadrilateral

1 Introduction

Given n cities $\{1, 2, \dots, n\}$ and a distance $d(i, j) > 0$ is given for each pair of cities i, j ($1 \leq i \neq j \leq n$), a salesman starts from one initial city and visits each of the other cities exactly once, and then returns to the starting city. Then, a Hamiltonian cycle (HC) is found. To save travel time or cost, he must find the shortest HC called the optimal Hamiltonian cycle (OHC). This is the traveling salesman problem (TSP) which has been proven to be NP-complete [1]. Until now, no exactly polynomial-time algorithms appeared for TSP unless $P = NP$. TSP was regarded in computer science for searching the optimal solution as soon as possible [2].

A complete weighted graph (WG) K_n is usually representing TSP. A vertex v_i ($1 \leq i \leq n$) in K_n notes the i^{th} city and an edge (v_i, v_j) notes the route between v_i and v_j . Each edge (v_i, v_j) is assigned a value $d(v_i, v_j) > 0$ which is taken as the distance between v_i and v_j . The algorithms for TSP are explored and analyzed based on graph theory. However, special properties on OHC are hard to find according to the distances on edges. Various algorithms including exact, approximate and heuristic algorithms are designed for searching OHC. The exact algorithms usually consume $O(a^n)$ [3] time, where $a > 1$. In theory, the TSP instance of big size is difficult to resolve. Fortunately, the scholars have found some useful methods, such as branch-and-bounds [4] and cutting planes [5], for neglecting many ordinary edges and paths out of OHC so that the search space

and computation time are reduced significantly. Thus, the TSP containing thousands of nodes were resolved. For example, the VLSI application containing 85,900 nodes was resolved [6] on one 128 networked computer system. The experiments showed that the time is difficult to decrease for tackling large TSP [7]. To resolve TSP as quick as possible, the heuristics and metaheuristic algorithms are widely applied [8] whereas the solutions must be evaluated through experiments. In current state, the algorithms designed according to *WG* do not meet the demands for tackling real-life TSP.

Besides the *WG* for TSP, the frequency graph (*FG*) [9] is another representation for TSP. Different from *WG*, the distance on an edge is changed into a corresponding frequency in *FG*. The frequency for an edge is computed based upon the optimal paths with given endpoints (optimal path) in K_n . An optimal path containing i vertices is computed in this way. Given i vertices $\{1, 2, \dots, i\}$ in K_n for $i \geq 4$, appoint two vertices as the endpoints, such as 1 and 2, and the other $i-2$ vertices are taken as the intermediate vertices. Since the intermediate vertices can be exchanged with each other, there are total $(i-2)!$ paths for endpoints 1 and 2. Among the group of paths, there is one path with the minimal length. This path is the optimal i -vertex path for endpoints 1 and 2. $\binom{i}{2}$ pairs of endpoints exist according to the vertex set $\{1, 2, \dots, i\}$ and $\binom{i}{2}$ optimal i -vertex paths arise. K_n contains $\binom{n}{i}$ sets of i vertices and there are total $\binom{i}{2} \binom{n}{i}$ optimal i -vertex paths. After these paths are computed, one can enumerate the number of paths including an edge (v_i, v_j) . This number is the frequency for (v_i, v_j) . While the frequency for each edge is obtained, a *WG* is converted into one corresponding *FG*. As *FG* is computed based upon optimal 4-vertex paths and frequency quadrilaterals, it found the frequency for an edge in OHC is much higher than that of most ordinary edges and [10].

In 1920s, Menger pointed out the OHC edges and the ordinary edges are difficult to separate with respect to their distances. Differently, an edge in OHC shows a frequency much bigger than that of most other edges in an *FG*. Wang [11] proved that the edges in OHC have one lower frequency bound as it is computed based on the optimal 4-vertex paths or frequency quadrilaterals. Moreover, the frequency or average frequency for an edge in OHC increases according to n [12]. These results demonstrate that the edges in OHC are obviously different from most ordinary edges according to their frequencies. Furthermore, we guess a path in OHC has one special frequency different from that for an ordinary path containing the same set vertices. Thus, the frequency bounds for the OHC paths will be explored to reduce the search space and save time as TSP is resolved. In this paper, we prove the frequency bounds for an OHC edge, two and three immediate edges in OHC and OHC, respectively in case that *FG* is computed using the frequency quadrilaterals (or optimal 4-vertex paths). Firstly, the previous frequency bound 3 for the edges in OHC is derived based on a new method. The change of the frequency for an edge in OHC is analyzed with respect to rising n . Moreover, the lower frequency bound is improved from 3 to $35/9$ for an edge in OHC if the restrictions of optimal 4-vertex paths are taken into consideration. Secondly, the lower frequency bounds for two and three immediate OHC edges are proven to be $22/3$ and $35/3$, respectively in the worst average case. Moreover, we proved that the average frequency for all edges in OHC is greater than 4 and it is bigger than $13/3$ for big and large TSP. These findings illustrate

the edges and paths in OHC are better cultivated in FG . In the end, the experiments are executed and the results are illustrated.

The following contents are presented in the following. The optimal 4-vertex paths and FG are introduced briefly in Sect. 2. The previous lower frequency bound for an edge in OHC is rederived with a new method and the frequency bound is improved as well in Sect. 3. In the 4th section, the lower frequency bounds for two and three immediate edges and all edges in OHC are proven, respectively, based on the frequencies on edges constrained by optimal 4-vertex paths. In the fifth section, the experimental results are illustrated. In the end, the paper is summarized.

2 The FG Based on Optimal 4-vertex Paths

An i -vertex path is noted as a sequence of i vertices $P = (v_1, v_2, \dots, v_i)$ ($2 \leq i \leq n$). v_1, v_i are the endpoints and the other vertices are the intermediate vertices. As $i = 2$, $P = (v_1, v_2)$ represents an edge. For the symmetric TSP, P is the same as the reversed path. *Given an OHC path containing i vertices, it never becomes shorter whenever the intermediate vertices exchange their positions. This kind of paths are optimal i -vertex paths* (optimal path). As there is no confusion, an optimal path is also noted as OP . There are total $n(n-1)2^{n-3}$ optimal i -vertex paths in K_n [9]. OHC in K_n contains n OP s for a given number i . The OP s not in OHC are ordinary OP s including at least one ordinary edge not in OHC. Different from the other paths in K_n , the OP s containing bigger number of vertices are composed of the OP s containing smaller number of vertices rather than the other paths or edges. Moreover, only a part of the OP s containing less vertices are included in the OP s containing more vertices.

Given a number i , K_n contains $\binom{n}{i}$ K_i s and there are $\binom{i}{2}$ optimal i -vertex paths in each K_i . Thus, total $\binom{i}{2}\binom{n}{i}$ optimal i -vertex paths exist in K_n . As i is big, it is time-consuming to compute an optimal i -vertex path. If i is small, all optimal i -vertex paths can be computed in reasonable time. For example, if $i = 4$, one can compute all the optimal 4-vertex paths. In this case, each edge is contained in $\binom{n-2}{2}$ K_4 s each of which contains six optimal 4-vertex paths. For an edge (v_i, v_j) contained in $\binom{n-2}{2}$ K_4 s, the number of optimal 4-vertex paths including (v_i, v_j) can be figured out according to the $6\binom{n-2}{2}$ optimal 4-vertex paths. This number is the frequency for (v_i, v_j) . Once the frequency for every edge is computed, an FG is obtained. In such an FG , the experiments illustrated that the frequencies for edges in OHC are much bigger than those for most ordinary edges [9, 10]. Moreover, the lowest frequency for edges in OHC is bigger than $3\binom{n-2}{2}$ and it increases according to n [12]. Given a WG , the OHC edges generally do not have special distances different from those of the ordinary edges.

3 The Lower Frequency Bound for Edges in OHC

The lower frequency bound has been discussed in paper [11]. Here, a different method is used to prove the previous lower frequency bound 3. One can predict the change of the average frequency for an edge in OHC with respect to rising n . Moreover, the lower bound is improved to 35/9 under the restrictions of the optimal 4-vertex paths.

3.1 The Previous Lower Frequency Bound

Given a K_4 , it contains six edges and optimal 4-vertex paths as well. As the frequency for each edge is computed based upon the six OP s, the frequency f is 1, 3 or 5 for an edge in the corresponding frequency quadrilateral [10]. Given a frequency quadrilateral, each of the edges in OHC in the corresponding K_4 have the frequency 3 and 5, respectively. An edge (v_i, v_j) in K_n is included in $\binom{n-2}{2}$ K_4 s and $\binom{n-2}{2}$ frequency quadrilaterals. In the k^{th} frequency quadrilateral, (v_i, v_j) has a specific frequency $f_k = 1, 3$ or 5 . As the frequencies 1, 3 and 5 are uniformly distributed for (v_i, v_j) with respect to all the frequency quadrilaterals, the probabilities $p_1(f = 1) = \frac{1}{3} - \frac{2}{3(n-2)}$, $p_3(f = 3) = \frac{1}{3}$ and $p_5(f = 5) = \frac{1}{3} + \frac{2}{3(n-2)}$ that an edge in OHC has 1, 3 and 5 are derived in the worst average case [11]. As the frequency F_{ij} for edge (v_i, v_j) is computed using all frequency quadrilaterals, i.e., $F_{ij} = \sum_{k=1}^{\binom{n-2}{2}} f_k \geq 3\binom{n-2}{2}$ holds for an edge in OHC. Meanwhile, the average frequency $f_{ij} = \frac{F_{ij}}{\binom{n-2}{2}} \geq 3$ also holds for an OHC edge. In the next, another method is used to prove the frequency bound 3 for edges in OHC.

Theorem 1. As the average frequency f_{ij} for an edge (v_i, v_j) in OHC is computed using the frequency quadrilaterals containing (v_i, v_j) , f_{ij} is bigger than 3 in the worst average case and it increases according to n .

Proof: Given an OHC edge (v_i, v_j) in K_n ($n \geq 4$), it is included in $\binom{n-2}{2}$ frequency quadrilaterals. (v_i, v_j) has the frequency 5 with a probability p_5^n according to the $\binom{n-2}{2}$ frequency quadrilaterals. The number of such frequency quadrilaterals is computed as $M_n = \binom{n-2}{2} p_5^n$. As a certain vertex is added to K_n , K_{n+1} is generated and (v_i, v_j) is contained in $\binom{n-1}{2}$ frequency quadrilaterals. (v_i, v_j) has the frequency 5 with a probability p_5^{n+1} with respect to the $\binom{n-1}{2}$ frequency quadrilaterals. There are $M_{n+1} = \binom{n-1}{2} p_5^{n+1}$ such frequency quadrilaterals for (v_i, v_j) in K_{n+1} .

As K_n turns into K_{n+1} , $n-2$ more frequency quadrilaterals are added for (v_i, v_j) . The $n-2$ frequency quadrilaterals are one sample drawn from the $\binom{n-1}{2}$ frequency quadrilaterals. On average, there will be $M = (n-2)p_5^{n+1}$ more frequency quadrilaterals in which (v_i, v_j) has the frequency 5 according to the added frequency quadrilaterals. Because $M_n + M \leq M_{n+1}$ exists, $p_5^n \leq p_5^{n+1}$ is derived. It means (v_i, v_j) will have more percentage of 5s according to n . It also says (v_i, v_j) will be contained in more percentage of the OP s with respect to rising n . As the probability equality $p_{3,5}^n = p_3^n + p_5^n$

exists, $p_{3,5}^n \leq p_{3,5}^{n+1}$ also holds and the same conclusion is drawn. Since the equalities $F_{ij} = (5p_5^n + 3p_3^n + p_1^n) \binom{n-2}{2} = f_{ij} \binom{n-2}{2}$ and $p_5^n + p_3^n + p_1^n = 1$ hold, f_{ij} increases according to n and it will tend to 5 as n is sufficiently big.

According to OHC, (v_i, v_j) has $n-3$ vertex-disjoint OHC edges in K_n . (v_i, v_j) and one of these edges are contained in one K_4 . In each of these corresponding frequency quadrilaterals, (v_i, v_j) has the frequency 3 or 5 [10]. Given such one frequency quadrilateral, (v_i, v_j) has 3 in the worst case and the probabilities $p_1^4 = p_5^4 = 0$ and $p_3^4 = 1$ hold. As the other vertices are added to the K_4 one by one, (v_i, v_j) will be contained in a sequence of graphs K_k s ($4 < k \leq n$) containing more and more vertices. Because $p_{3,5}^k$ and p_5^k increase according to k and f_{ij} as well, $f_{ij} \geq 3$ holds as $k \geq 4$. $\square$

Since $F_{ij} \geq 3 \binom{n-2}{2}$ for an edge in OHC, $p_5(f = 5)$ must be bigger than $p_1(f = 1)$ for each OHC edge, i.e. $p_5(f = 5) \geq p_1(f = 1)$. In the average case, an OHC edge will have a 3 or 5 in a frequency quadrilateral including it, i.e. $p_{3,5}(f = 3 \text{ or } 5) = 1$. Moreover, the frequency for an edge in OHC increases according to n . Given an ordinary edge, it is generally contained in many frequency quadrilaterals where the frequency is 1. Based on Theorem 1, the probability $p_1(f = 1)$ for an ordinary edge will rise according to n . There are $\binom{n}{4}$ frequency quadrilaterals in K_n and each frequency quadrilateral includes six edges having 1, 1, 3, 3, 5 and 5. The total 1s, 3s and 5s reaches the number $6\binom{n}{4}$ and the average frequency for all edges is 3. Based on Theorem 1, an OHC edge will have more 3s and 5s than an edge not in OHC. It says the average frequency for an ordinary edge is smaller than 3. The distances related to the OHC edges generally do not have upper and lower bounds.

3.2 The Improved Lower Frequency Bound

As the probability Eq. (1) is derived [10] for an OHC edge, the authors assumed the frequencies 1, 3 and 5 in a frequency quadrilateral are independent to each other and uniformly distributed for all edges in K_n . The experiments illustrated that this lower frequency bound is too small for most TSP instances [10, 12]. In fact, the frequencies 1, 3 and 5 are not uniformly distributed for each edge due to the optimal 4-vertex paths. Based on the OP s in a K_4 , the lower frequency bound for edges in OHC can be improved to 35/9.

Theorem 2. As the frequency for each edge is computed using the frequency quadrilaterals in K_n , the average frequency for an edge in OHC is greater than 35/9 in the worst average case.

Proof: A K_4 contains six edges and six optimal 4-vertex paths. For example, a K_4 on A, B, C and D is shown in the left upper side in Fig. 1. $d(A,B)$, $d(B,C)$, $d(C,D)$, $d(A,D)$, $d(A,C)$ and $d(B,D)$ are the distances on the corresponding edges, respectively. The six optimal 4-vertex paths depend on the three summed distances $d(C,D) + d(A,B)$, $d(B,C) + d(A,D)$ and $d(B,D) + d(A,C)$. For example, according to the inequality $d(B,D) + d(A,C) > d(B,C) + d(A,D) > d(C,D) + d(A,B)$, the six optimal 4-vertex paths are (A,D,C,B) , (A,B,D,C) , (A,B,C,D) , (B,A,D,C) , (B,A,C,D) , and (C,B,A,D) which are shown under $ABCD$ in Fig. 1. As the frequency for every edge is enumerated from the six optimal 4-vertex paths, the frequency quadrilateral $ABCD$ is obtained and shown in the upper

right side in Fig. 1. The frequency beside each edge is the number of the *OPs* including the edge.

In the frequency quadrilateral, the vertex-disjoint edges have the same frequency, such as the frequency 5 for (A,B) and (C,D) , frequency 1 for (A,C) and (B,D) , and frequency 3 for (A,D) and (B,C) . Moreover, the three edges gathering at a vertex have the different frequencies. For example, (A,B) , (A,C) and (A,D) containing A have the frequencies 5, 1 and 3, respectively. Under the frequency $ABCD$, the six *OPs* are illustrated as well as the corresponding frequency is attached to each edge. There are three edges in each optimal 4-vertex path including one middle edge and two edges on both ends. Different from the diverse distances on edges, the frequencies on edges in each optimal 4-vertex path are restricted to 1, 3 and 5, and they conform to the following rules. In each optimal 4-vertex path, the two edges on both ends have one identical frequency. As the frequency of the middle edge is 5, the frequency of the edges on both ends must be 3. Else if the frequency of the middle edge is 1 or 3, the frequency of the edges on both ends must be 5. The frequency 1 never appears on the end edges in any one optimal 4-vertex path. Given two adjacent edges in an optimal 4-vertex path, the two adjacent edges may have two frequencies in $\{3,5\}$ or $\{1,5\}$ rather than $\{1,3\}$. It says two adjacent edges having the frequencies in $\{1,3\}$ in any one frequency quadrilateral are never contained in an optimal 4-vertex path. Thus, two edges having the frequencies in $\{1,3\}$ in all frequency quadrilaterals are never adjacent in OHC. $d(C,D) + d(A,B)$, $d(B,C) + d(A,D)$ and $d(B,D) + d(A,C)$ have six orderings. Based on each of the other five orderings, one can derive the corresponding six *OPs* as well as frequency quadrilateral, respectively. In total, six frequency quadrilaterals exist for $ABCD$ [10]. The frequencies for the three edges in every optimal 4-vertex path conform to the same rules.

There are $\binom{n-2}{2} K_4$ s for each edge in K_n and each K_4 includes twelve 4-vertex paths. Since OHC only contains *OPs*, the six *OPs* in each K_4 are concerned and the other paths are neglected. An OHC edge in a K_4 must be contained in one of the six optimal 4-vertex paths in Fig. 1. Moreover, the frequency for each OHC edge in an optimal 4-vertex path must conform to the frequencies on edges in the six *OPs*. An *OP* may contain the OHC edge or other ordinary edges. Based on Theorem 1, an OHC edge will have a bigger probability to own the 3 and 5 than an ordinary edge. In the worst case, the frequencies on edges according to the six *OPs* in Fig. 1 are uniformly distributed for the OHC edge as an ordinary edge. The probabilities that an OHC edge has 1, 3 and 5 will be $1/9$, $1/3$ and $5/9$, respectively. In this case, the average frequency for an edge in OHC is bigger than $35/9$. $\square$

Based on OHC, there are $n-3$ frequency quadrilaterals in each of which an OHC edge has a 3 or 5. Moreover, there are $\frac{2}{3}(n-3)$ 5s and $\frac{1}{3}(n-3)$ 3s in the average case [11]. In the remainder frequency quadrilaterals, the OHC edge will have 1, 3 and 5 with the probabilities $1/9$, $1/3$ and $5/9$, respectively based on the six *OPs* in Fig. 1. In this case, $p_1(f=1)$, $p_3(f=3)$ and $p_5(f=5)$ for an OHC edge are improved to those in the Eq. (1), respectively. As the frequency for an edge is computed using optimal 4-vertex

paths, the average frequency for an edge in OHC can be estimated according to Eq. (1).

$$\begin{cases} p_1(f=1) = \frac{1}{9} - \frac{2}{9(n-2)} \\ p_3(f=3) = \frac{1}{3} \\ p_5(f=5) = \frac{5}{9} + \frac{2}{9(n-2)} \end{cases} \quad (1)$$

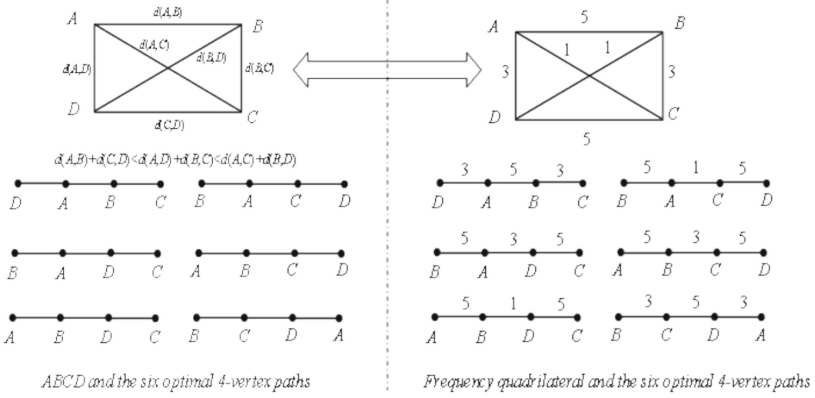


Fig. 1. The frequencies on edges in the six optimal 4-vertex paths

4 The Lower Frequency Bounds for Two, Three Immediate OHC Edges and OHC

For each TSP instance, every edge has a fixed frequency in FG computed with the frequency quadrilaterals. Given a P , OP or HC containing k ($2 \leq k \leq n$) edges, the average frequency for P , OP or HC is the sum of the average frequencies for the edges contained in the P , OP or HC , respectively. In similarity, each P or OP also has one specific frequency as well as an HC . An OHC in K_n contains n optimal i -vertex paths for given i . The frequencies for the OHC paths and OHC are concerned here. According to the distributions of the frequencies on edges with respect to the six OP s in Fig. 1, the lower frequency bounds for two, three immediate OHC edges and OHC will be derived, respectively.

Corollary 1. As the frequency for each edge is computed with the frequency quadrilaterals in K_n , the average frequency for two immediate edges in OHC is greater than $22/3$ in the worst average case.

Proof: OHC in K_n only contains OP s. Two immediate edges are contained in $n-3$ K_4 s and a K_4 includes six optimal 4-vertex paths. Moreover, one of the optimal 4-vertex paths in some K_4 is contained in OHC. In each corresponding frequency quadrilateral, the frequencies for the two edges must have the frequencies on two adjacent edges in the six OP s in Fig. 1. Otherwise, the two edges will not be contained in an optimal 4-vertex path and OHC as well. An optimal 4-vertex path may contain two immediate

OHC edges or not. In the worst case, the frequencies for two adjacent edges in the six OP s in Fig. 1 are uniformly distributed for two immediate OHC edges in K_n . In this case, the probabilities that two immediate edges have the frequencies 6 and 8 are $1/3$ and $2/3$, respectively. The average frequency for two immediate OHC edges is equal to $22/3$. As n rises, the probabilities $p_{3,5}^n$ and p_5^n increase for an OHC edge. Therefore, the average frequency for two immediate OHC edges will be bigger than $22/3$ if $n > 4$.

In addition, two adjacent edges in an optimal 4-vertex path have the minimum frequency 6. In the worst case, two immediate OHC edges own the lowest frequency 6 in each of the frequency quadrilaterals. In this case, the two immediate edges in OHC have the lowest average frequency 6. $\square$

Corollary 2. As the frequency for an edge is computed based upon the frequency quadrilaterals in K_n , three immediate edges in OHC have an average frequency bigger than $35/3$ in the worst average case.

The proof is the same as that for proving Corollary 1. It mentions that the minimum average frequency for three immediate OHC edges is 11 in the worst case. For big and large TSP, an edge in OHC will have a frequency bigger than 3 in a frequency quadrilateral. In this case, the average frequency for two and three immediate edges in OHC shall surpass 8 or 13 (see Fig. 1), respectively.

The average frequencies for two and three immediate edges in OHC in the worst case are too small in applications. Once TSP size becomes big, they are not helpful to filter out the ordinary OP s containing two and three edges, respectively. The bigger frequency bounds for them will be derived through frequency consistency [11]. In K_n , two adjacent edges are contained in $n-3$ K_4 s. It means that two adjacent edges are at most contained in $2(n-3)$ optimal 4-vertex paths. If two adjacent edges are contained in some optimal 4-vertex paths, the frequencies must be in $\{5,3\}$ or $\{5,1\}$. Besides the $n-3$ K_4 s, each of them is contained in the other $\binom{n-3}{2}$ K_4 s, respectively. If the frequency distributions for two adjacent edges are derived [11] according to the $n-3$ frequency quadrilaterals containing them, the frequency for the two edges can be estimated. As an edge e is noted as (v_i, v_j) , the frequency $f(e) = f(v_i, v_j)$ exists.

Theorem 3. Given a frequency quadrilateral $ABCD$ in K_n on set $\{1, 2, \dots, n\}$. If $f(A, C) > f(A, B)$ (or $f(A, B) > f(A, C)$) holds, $f(A, C) > f(A, B)$ (or $f(A, B) > f(A, C)$) exists with the probability $> 2/3$ in another frequency $ABCE$ and $E \in K_n \setminus \{A, B, C, D\}$.

Theorem 3 indicates an edge will keep the same frequency in two related frequency quadrilaterals where this edge and another identical vertex are contained, see the proof in paper [11]. Based on Theorem 3, the lower frequency bound will be derived for an OP containing several immediate edges in OHC.

Theorem 4. As the frequency for an edge is computed based upon the frequency quadrilaterals in K_n , the average frequencies for two and three immediate OHC edges are bigger than $20/3$ and $31/3$, respectively. For k immediate edges in OHC, the average frequency will be greater than $10k/3$ for even number k .

Proof: Given two immediate edges (A, B) and (A, C) in OHC in K_n , they are contained in $n-3$ frequency quadrilaterals $ABCD$. In every $ABCD$, there are two optimal 4-vertex path where (A, B) and (A, C) are in the middle, respectively. Since (A, B) and (A, C) are the unique pair of OHC edges containing A , the optimal 4-vertex paths in which (A, B) or (A, C) is in the middle must include the other edge, respectively. Otherwise, (A, B) or

(A,C) cannot be approved to be contained in OHC simultaneously if one of them is the OHC edge. It says the $n-3$ optimal 4-vertex paths in which (A,B) is in the middle must include (A,C) and vice versa.

The average frequency for (A,B) and (A,C) is 3 in the worst case. On average, (A,B) and (A,C) will have the frequency 3 or 5 in a frequency quadrilateral including them, respectively. In the worst case, one of the edges, such as (A,C) , has 3 with respect to all frequency quadrilaterals including it. Since (A,B) and (A,C) are contained in $n-3$ optimal 4-vertex paths, (A,B) must have 5 in the $n-3$ frequency quadrilaterals with respect to the optimal 4-vertex paths in Fig. 1.

As vertex C or D is substituted with a certain vertex E in K_n , another frequency quadrilateral $ABCE$ or $ABDE$ is generated. In every frequency quadrilateral $ABCE$ or $ABDE$, 1, 3 and 5 for (A,B) occurs with the probability $1/9$, $4/9$ and $4/9$, respectively based on Theorem 3. Total $\binom{n-3}{2}$ such frequency quadrilaterals $ABCE$ and $ABDE$ are generated for (A,B) . Plus the initial $n-3$ frequency quadrilaterals where (A,B) has 5, the total frequency of (A,B) is $\frac{11}{3}\binom{n-2}{2} + \frac{4}{3}(n-3)$. Thus, the average frequency for (A,B) and (A,C) is bigger than $20/3$. It mentions that there are the other $n-3$ frequency quadrilaterals including (A,B) and the other immediate edge in OHC. In each of them, (A,B) has 3 and 5 rather than 1 [10]. Moreover, (A,B) and (A,C) will be contained in more percentage of the OPs according to $n \geq 4$ as well as the path (B,A,C) based on Theorem 1. Therefore, the average frequency for two immediate edges in OHC will be greater than $20/3$.

In the same way, the other OHC edge adjacent to (A,C) must have an average frequency bigger than $11/3$. Thus, the average frequency for three immediate OHC edges is bigger than $31/3$. For an OHC path containing k immediate edges, the average frequency will be bigger than $10k/3$ for even number k .

On average, (A,B) has 3 or 5 in each of the frequency quadrilaterals $ABCE$ and $ABDE$. Thus, (A,B) will have 3 and 5 with the probabilities $1/3$ and $2/3$, respectively according to the $\binom{n-3}{2}$ $ABCEs$ and $ABDEs$ based on Theorem 3. In this manner, the average frequency for (A,B) and (A,C) becomes the same as that derived in Corollary 1. Moreover, the frequency computed for three immediate edges in OHC is the same as that derived in Corollary 2 and that for an OHC path containing k edges will be bigger than $11k/3$ for even number k . $\square$

As the frequencies for (A,B) and (A,C) in OHC are not restricted by the optimal 4-vertex paths, the lowest frequency derived for (A,B) is smaller than that derived in Theorem 2. $11/3$ can be viewed as the lower frequency bound in the worst case for medium and big TSP. The lowest frequency for three immediate edges in OHC is also smaller than that derived in Corollary 2. This lower frequency bound can be used for small TSP. In the next, the lower frequency bound for OHC will be proven.

Theorem 5. As the frequency of each edge is computed based on frequency quadrilaterals in K_n , OHC has an average frequency bigger than $4n$ in the worst average case.

Proof: Based on Theorem 1, the average frequency for an edge in OHC is greater than 3 in the worst case. Thus, an OHC edge will have 3 or 5 in every frequency quadrilateral including it. Given a frequency quadrilateral, it contains six optimal 4-vertex paths.

Based upon the six OPs in Fig. 1, four of them meet the demand of OHC edges (see the four optimal paths (D,A,B,C) , (B,A,D,C) , (A,B,C,D) and (B,C,D,A) in Fig. 1). It is known that an OHC edge will be contained in more percentage of the OPs according to n . In the worst case, the frequencies in the four OPs are uniformly distributed for an OHC edge. In this case, an edge in OHC has 3 and 5 with the probability $1/2$, respectively. An edge in OHC has the average frequency 4. Meanwhile, OHC has an average frequency bigger than $4n$.

For large and big TSP, $p_5(f = 5)$ will tend to 1 and an OHC edge will have an average frequency approaching 5 based on Theorem 1. In this case, the optimal 4-vertex paths with the maximum frequency are used to compute the average frequency for an OHC edge. An OHC edge will have the improved lower frequency bound $13/3$. Meanwhile, OHC will have an average frequency greater than $13n/3$. $\square$

Based on Theorem 5, two immediate edges in OHC have an average frequency bigger than 8 in most cases. This result can also be derived according to Theorem 1. As $n = 4$, two immediate edges in OHC have 3 and 5 in a frequency quadrilateral, respectively. The total frequency for them is 8. As n increases, the probabilities p_5^n and $p_{3,5}^n$ will rise for each of the two immediate edges in OHC. The average frequency $5p_5^n + 3p_3^n + p_1^n$ also become bigger according to n . Thus, two immediate edges in OHC will have an average frequency bigger than 8 as $n > 4$. In similarity, three immediate OHC edges have an average frequency bigger than 12 in most cases.

After WG is converted into FG based on frequency quadrilaterals, the lower frequency bounds for an OHC edge, two and three immediate OHC edges can be used to trim many ordinary edges and OPs for searching OHC. The lower frequency bounds for the above cases are presented in Table 1.

Table 1. The lower frequency bounds for OHC edges and paths

Edges and paths in OHC	Worst average case		Worst case	
	Small TSP	Big TSP	Small TSP	Big TSP
An OHC edge	35/9	4	3	11/3
Two adjacent OHC edges	22/3	8	6	20/3
Three immediate OHC edges	35/3	12	31/3 or 11	11
OHC	$4n$	$13n/3$	$4n$	$4n$

In applications, these frequency bounds can be selectively used for TSP of different sizes. For example, the smaller bounds are applied to small TSP and big bounds are used for big and large TSP. Moreover, as a path contains enough number of edges, the edges will have an average frequency greater than 4 or $13/3$. Since an edge in OHC has an increasing frequency according to n , the average frequency for a number of immediate edges in OHC will be bigger than 4 for big and large TSP. It mentions that all the detected paths should be OPs meeting the frequency bounds. If an HC is found based on these frequency bounds, its frequency can be taken as a new bound for finding the next better HC.

5 Experimental Results and Analysis

The real-life TSP instances in TSPLIB [14] are used to verify the above findings. TSPLIB includes TSP instances of Euclidean, GEO, ATT, CEIL and Matrix. One OHC for each of these instances has been computed with Concorde package [15]. Given an edge in a TSP instance, 1000 K_4 s containing this edge are randomly drawn for computing its average frequency. The average frequency for an edge does not change much in two or more experiments owing to the binomial distribution. The lowest frequencies are computed and recorded for an OHC edge, two, three and more number of immediate OHC edges, respectively. One can compare them with the derived lower frequency bounds for an edge and several immediate edges in OHC.

The results are illustrated in Table 4. The first column is TSP name. Two digits after integer are maintained for each average frequency. The experimental results are divided into two groups. In the first group, the sufficient condition [13] for an edge in OHC is not considered. For each instance, the lowest frequencies are illustrated for an OHC edge, two and more number of immediate OHC edges, respectively. In the second group, the sufficient condition [13] is used to find the OHC edges with an average frequency bigger than $13/3$. Then, the lowest frequencies are computed for two and three immediate OHC edges containing one confirmed OHC edge, respectively. In the first group, the minimum frequency is provided for an OHC path containing k ($k = 1, 2, 3, 4, 5$ and 6) edges as well as that for OHC. In the 9th column in Table 4, $f_{>4k/k}$ presents the minimum frequency bigger than $4k$ for k immediate OHC edges in case that k reaches the minimum number for each of the TSP instances. In the last column, $m_{f<35/9}$ and $m_{f<4}$ represent the numbers of OHC edges with an average frequency smaller than $35/9$ and 4 , respectively, and n notes the TSP size.

Firstly, the minimum frequency for an OHC edge is above 3 except brg180 and si1032. It says the bound 3 works well for nearly all these TSP instances. In addition, the OHC edges having an average frequency close to 3 merely happens to some small TSP instances. As TSP sizes is above 400, nearly all of them become bigger than $11/3$ except dsj1000, fl1400, fl3795, si535 and si1032. Moreover, the lowest frequency is greater than $35/9$ for most medium and big TSP instances. In the last column in Table 4, the numbers of the OHC edges having an average frequency smaller than $35/9$ and 4 are provided for each TSP instance, respectively. As n is small, some TSP instances contain several OHC edges having an average frequency below $35/9$ or 4 whereas these frequencies are much bigger than 3. However, the number of such edges is limited and much smaller than n . As n is bigger than a number, such as $n > 1000$, $m_{f<4}$ becomes very small and it is equal to 0, 1, 2 and 3 for most of the TSP instances except fl1400. Moreover, $m_{f<4}$ equals zero for nearly all the TSP instances as $n > 2000$. It means the OHC for most TSP instances only contains a small number of edges with an average frequency below 4. That's why the average frequency for the edges in most OHC paths is bigger than 4. Theorem 2 and Theorem 4 are totally verified by these experimental results. It mentions that there are some TSP instances having a lot of edges with equal distances, such brg180, si535, dsj1000, fl1400, fl3795 and si1032, etc. For these TSP instances, it is difficult to select the right OP s containing some known OHC edges. In our algorithm, the OP s are chosen according to the lexicographic order of vertices. For example, if the two paths (A,C,B,D) and (A,B,C,D) have the equal length in $ABCD$, the

second path is chosen and used whereas the first one is discarded. This manner will lower the frequency of some OHC edges for some instances, such as brg180 and si1032, etc. For the other TSP instances without many equal-distance edges, Theorems 1, 2 and 4 work well for each of them.

In similarity, the lowest frequency for two immediate edges in OHC is bigger than $22/3$ and that for three immediate OHC edges is bigger than $35/3$ for most of these TSP instances. For some small TSP instances, the minimum frequency for two and three immediate OHC edges are smaller than but close to $22/3$ and $35/3$, respectively. As TSP size is big enough, such as $n > 1000$, the minimum frequency for two immediate edges in OHC is bigger than 8 except fl1400, si1032 and u1060. Moreover, the minimum frequency for three immediate edges in OHC is greater than 12 for all the instances of big size. The frequency bounds work well for these real-life TSP instances. In addition, the corresponding minimum frequency are bigger than the bounds 6 and $31/3$ or 11 in the worst case, respectively for all these instances. For the medium or big TSP instances, the lowest frequency for two immediate edges in OHC is bigger than $20/3$, respectively. Some TSP instances, such as brg180, have a lot of edges of equal distance and they may contain many OHCs. The optimal 4-vertex paths that we use may be not right to compute a big frequency for some known OHC edges. For some small TSP instances, such as bays29 and gr24, the probabilities $p_5(f = 5)$ and $p_3(f = 3)$ do not increase much according to the small size. For all medium or big TSP, the results totally conform to the lower frequency bounds in Table 1.

Moreover, the difference between two immediate frequencies from k to $k + 1$ is bigger than 4 for each of these instances. Thus, the frequency 4 can be taken as a common difference for searching the candidate OHC paths containing quite a few edges. As $k = n$, the average frequency for all edges in OHC is greater than 4 for either small or big TSP. It means all the TSP instances comply with Theorem 5. In addition, there are only five small TSP instances having the average frequency smaller than $13/3$ for all OHC edges. For the other instances, the OHC edges have an average frequency bigger than $13/3$. As n rises, one also sees the average frequency for all OHC edges increases according to n , see the results for the same type of TSP instances, such as d198 ~ d2103, fl417 ~ fl3795, gr21 ~ gr48/gr120, gr96/gr137 ~ gr666, kroA100 ~ kroA200, pcb442 ~ pcb3038, pr76 ~ pr2392, u159 ~ u2319, etc. For the big TSP, the average frequency for all OHC edges approaches 5. It says Theorems 1 and 5 works well for all these instances.

The lowest frequency bigger than $4k$ for the minimum number k illustrates that k is smaller than 20 for all these instances except brg180. For the small TSP instances as $n < 200$, k is relatively big for some of them but smaller than 20. As n is above 300, k becomes smaller than 10 for nearly all of them. Moreover, k is equal to 1, 2 or 3 for most instances as $n > 400$. It means there are a small number of OHC edges having a small average frequency for medium or big instances. Thus, the average frequency for a few edges contained in an OHC path begins surpass 4. It mentions that the results for large instances are not provided since they totally conform to these frequency bounds, see the experiments [12, 13].

As one OHC edge is confirmed [13], the two immediate edges in OHC have the lowest frequency greater than $22/3$ except brg180 having the frequency bigger than 7. Two immediate edges in OHC have the lowest frequency greater than 8 for most

instances. As $k = 3$, the three immediate OHC edges have the lowest frequency bigger than $35/3$ except bays29, brg180, kroA100 and kroC100. For kroA100 and kroC100, the lowest frequency for the three immediate edges in OHC is bigger than $31/3$ in the worst case. For the other instances, the lowest frequency for three immediate edges in OHC is bigger than 12 in most cases. Moreover, it is bigger than 13 for TSP instances of scale $n > 2000$. If some OHC edges with an average frequency bigger than $13/3$ are found, the two and three immediate OHC edges containing one of them usually have the frequency bigger than 8 and 12, respectively (Table 2).

Table 2. The minimum average frequency for k ($=1 \sim 6, n$) immediate OHC edges in OHC

TSP	Without sufficient condition								With sufficient condition		$m_{f < 35/9}/m_{f < 4/n}$
	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = n$	$f_{>4k}/k$	$k = 2$	$k = 3$	
a280	3.65	7.59	11.78	15.87	20.02	24.19	4.64	20.02/5	8.50	12.65	1/2/280
ail535	3.98	8.20	12.66	16.92	21.09	25.39	4.71	8.20/2	8.43	12.66	0/1/535
att48	3.49	7.28	10.98	14.89	18.69	22.50	4.31	52.06/13	8.39	12.09	11/15/48
att532	3.89	8.07	12.31	16.68	20.93	25.19	4.65	8.07/2	8.36	12.41	0/2/532
bayg29	3.47	7.01	10.86	14.92	19.24	23.64	4.32	36.36/9	8.43	12.09	5/6/29
bays29	3.20	6.60	10.08	13.89	18.00	22.40	4.34	36.00/9	7.80	10.99	5/8/29
bier127	3.88	7.97	12.12	16.31	20.46	24.55	4.54	12.12/3	8.41	12.53	1/6/127
brazil58	3.65	7.72	11.66	15.74	19.92	24.33	4.40	24.33/6	8.25	12.15	5/10/58
brg180	2.95	6.61	10.31	14.02	17.69	21.43	4.30	104.31/26	7.32	10.98	12/14/180
ch130	3.67	7.71	11.61	15.44	19.54	23.81	4.59	32.12/8	8.18	12.12	3/5/130
ch150	3.58	7.67	11.67	15.98	20.15	24.28	4.60	20.15/5	8.24	12.30	2/5/150
d198	3.60	7.68	11.57	15.81	19.57	23.70	4.51	36.12/9	8.03	12.16	14/198
d493	3.93	7.96	12.03	16.07	20.36	24.48	4.65	12.03/3	8.35	12.35	0/3/493
d657	3.84	7.97	12.12	16.63	20.76	24.91	4.73	12.12/3	8.38	12.48	1/1/657
d1291	3.91	8.52	13.12	17.68	22.28	26.64	4.80	8.52/2	8.52	13.12	0/1/1291
d1655	3.70	8.27	12.23	16.80	21.32	25.90	4.80	8.27/2	8.27	12.79	2/3/1655
d2103	4.38	8.87	13.38	17.90	22.44	26.94	4.82	4.38/1	8.87	13.38	0/2103
dsj1000	3.49	7.74	12.10	16.49	18.61	20.96	4.76	12.10/3	7.87	12.10	1/1/1000
fl417	3.53	7.49	11.32	15.66	19.33	23.23	4.71	44.12/11	8.01	11.90	7/11/417
fl1400	3.50	7.70	12.03	16.38	20.89	24.89	4.77	12.03/3	8.23	12.45	6/12/1400
fl1577	3.77	8.37	12.93	17.56	22.12	26.76	4.83	8.37/3	8.37	12.93	2/3/1577
fl3795	3.61	8.43	13.13	17.82	22.50	27.24	4.88	8.43/2	8.43	13.13	2/3/3795
fnl4461	4.33	8.80	13.24	17.75	22.25	26.76	4.81	4.33/1	8.80	13.24	0/0/4461
fri26	3.03	6.81	10.95	14.64	18.73	22.22	4.27	60.22/15	8.15	12.00	6/6/26

(continued)

Table 2. (continued)

TSP	Without sufficient condition								With sufficient condition		$m_{f<35/9}/m_{f<4}/n$
	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = n$	$f_{>4k}/k$	$k = 2$	$k = 3$	
gil262	3.83	7.90	11.98	16.03	20.26	24.31	4.66	16.03/4	8.26	12.47	2/4/262
gr21	3.50	7.07	11.13	15.21	19.43	23.44	4.33	28.02/7	8.10	11.67	2/3/21
gr24	3.22	6.49	10.05	13.35	17.02	20.87	4.23	60.22/15	8.45	12.68	6/8/24
gr48	3.57	7.44	11.38	15.56	19.67	23.92	4.43	28.21/7	8.00	11.82	5/7/48
gr96	3.60	7.54	11.26	15.69	19.74	23.53	4.52	52.29/13	8.05	11.99	8/14/96
gr120	3.42	7.34	11.01	14.90	18.79	23.00	4.50	64.17/16	8.32	12.44	8/15/120
gr137	3.52	7.49	11.48	15.44	19.43	23.46	4.46	64.16/16	8.30	12.01	9/19/137
gr202	3.98	8.13	12.16	16.31	20.34	24.58	4.59	8.13/2	8.33	12.51	0/1/202
gr229	3.39	7.47	11.44	15.61	19.72	23.92	4.56	28.15/7	8.35	12.52	1/5/229
gr431	3.70	7.63	11.62	15.35	19.36	23.33	4.58	68.13/17	8.19	12.36	8/15/431
gr666	4.12	8.44	12.85	17.22	21.49	26.00	4.70	4.12/1	8.48	12.85	0/0/666
hk48	3.19	6.77	10.56	14.66	18.54	22.70	4.36	80.03/20	8.11	12.11	9/11/48
kroA100	3.34	6.84	10.80	14.84	18.54	22.39	4.51	76.05/19	7.69	11.18	9/10/100
kroA150	3.64	7.39	11.34	15.32	19.52	23.44	4.56	44.06/11	8.19	12.23	7/13/150
kroA200	3.56	7.34	11.24	15.48	19.26	23.52	4.60	44.04/11	8.09	11.92	7/10/200
p654	3.48	7.33	11.41	15.79	19.64	23.28	4.73	32.30/8	8.11	11.87	8/12/654
pa561	3.81	7.92	11.83	16.13	20.36	24.27	4.74	16.13/4	8.36	12.64	1/5/561
pcb442	3.87	8.19	12.52	16.82	21.02	25.31	4.70	8.19/2	8.26	12.61	1/1/442
pcb1173	4.04	8.41	12.82	17.25	21.48	25.89	4.76	4.04/1	8.41	12.82	0/0/1173
pcb3038	4.21	8.58	13.03	17.42	21.83	26.26	4.81	4.21/1	8.71	13.03	0/0/3038
pla7397	4.12	8.37	12.71	17.31	21.96	26.39	4.86	4.12/1	8.468	12.71	0/0/7397
pr76	3.42	7.59	11.23	15.33	19.48	23.66	4.52	60.23/15	8.51	12.74	4/5/76
pr107	3.50	8.33	12.91	17.42	21.92	26.42	4.62	8.33/2	8.33	12.91	2/2/107
pr124	3.29	7.60	11.88	15.98	19.89	23.91	4.63	28.37/7	8.09	12.02	5/7/124
pr136	3.89	7.89	12.09	16.22	20.28	24.49	4.60	12.09/3	8.51	12.61	0/4/136
pr144	3.15	7.63	11.58	16.13	20.77	25.36	4.70	16.13/4	7.63	12.19	4/5/144
pr152	3.52	8.00	12.47	16.89	21.20	25.25	4.69	8.00/2	8.00	12.47	2/4/152
pr226	3.45	7.92	12.47	17.12	21.12	25.59	4.75	12.47/3	8.17	12.47	3/3/226
pr264	3.49	7.92	12.44	16.69	21.55	26.38	4.70	12.44/3	7.92	12.44	2/3/264
pr299	3.70	7.82	11.79	15.76	19.73	23.98	4.60	28.18/7	8.39	12.27	4/10/299
pr439	3.93	7.92	12.06	16.21	20.34	24.84	4.70	12.06/3	8.46	12.51	0/4/439

(continued)

Table 2. (continued)

TSP	Without sufficient condition								With sufficient condition		$m_{f < 35/9}/m_{f < 4}/n$
	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = n$	$f_{>4k}/k$	$k = 2$	$k = 3$	
pr1002	3.84	8.28	12.73	17.07	21.41	25.66	4.76	8.28/2	8.31	12.73	1/1/1002
pr2392	4.09	8.52	12.88	17.34	21.70	26.23	4.81	4.09/1	8.52	12.88	0/0/2392
rat575	3.86	7.88	12.13	16.16	20.27	24.34	4.64	12.13/3	8.42	12.58	1/3/575
rat783	3.95	8.05	12.26	16.48	20.73	24.99	4.68	8.05/2	8.33	12.46	0/4/783
rd400	3.84	7.98	12.20	16.39	20.60	24.84	4.69	12.20/3	8.41	12.55	1/2/400
rl1304	3.95	8.24	12.74	17.36	21.90	26.41	4.81	8.24/2	8.36	12.74	0/1/1304
rl1323	4.06	8.33	12.58	16.74	21.07	25.43	4.80	4.06/1	8.54	12.85	0/0/1323
rl1889	3.92	8.38	12.65	17.23	21.63	25.89	4.82	8.38/2	8.38	12.92	0/2/1889
rl5915	4.33	8.86	13.37	17.95	22.51	27.08	4.86	4.33/1	8.86	13.37	0/0/5915
rl5934	4.24	8.86	13.31	17.77	22.29	26.90	4.86	4.24/1	8.86	13.31	0/0/5934
si175	3.97	8.61	13.19	17.95	22.95	27.62	4.85	8.61/2	8.61	13.41	0/2/175
si535	3.58	8.51	13.17	17.37	22.30	27.23	4.94	8.51/2	8.51	13.17	3/3/535
si1032	2.68	7.68	12.67	17.67	22.66	27.65	4.99	12.67/3	7.68	12.67	3/3/1032
st70	3.32	7.18	11.18	15.06	19.18	23.37	4.52	40.03/10	8.23	12.24	3/3/70
ts225	4.05	8.33	12.53	16.75	21.03	25.30	4.68	4.05/1	8.57	12.70	0/0/225
tsp225	3.89	8.03	12.23	16.26	20.55	24.58	4.60	8.03/2	8.31	12.41	0/3/225
u159	3.70	7.68	11.57	15.60	19.48	23.75	4.60	32.08/8	8.09	12.06	6/10/159
u574	3.90	8.02	12.26	16.41	20.53	24.76	4.69	8.02/2	8.34	12.60	0/2/574
u724	3.90	8.10	12.20	16.48	20.84	24.95	4.70	8.10/2	8.42	12.46	0/1/724
u1060	3.83	7.94	12.16	16.44	20.79	25.03	4.70	12.16/3	8.31	12.37	1/2/1060
u1432	4.06	8.40	12.80	17.19	21.56	25.96	4.76	4.06/1	8.40	12.80	0/0/1432
u1817	3.74	8.27	12.80	17.36	21.91	26.42	4.78	8.27/2	8.27	12.80	1/2/1817
u2152	4.10	8.53	13.05	17.37	21.89	26.40	4.79	4.10/1	8.58	13.05	0/0/2152
u2319	4.27	8.65	13.01	17.38	21.77	26.20	4.76	4.27/1	8.67	13.01	0/0/2319
vm1084	3.88	8.05	12.35	16.60	20.88	25.29	4.76	8.05/2	8.29	12.54	1/3/1084
vm1748	4.11	8.33	12.54	16.87	21.09	25.55	4.80	4.11/1	8.45	12.70	0/0/1748

6 Conclusions

Several frequency bounds for OHC edges and paths are derived for TSP based on frequency quadrilaterals. As the average frequency for an edge is computed based upon the frequency quadrilaterals, the lower frequency bounds for an edge, two and three immediate edges in OHC are $35/9$, $22/3$ and $35/3$, respectively in the worst average case. These frequency bounds are lowered to 3, 6 and 11, respectively in the worst case. Moreover, all OHC edges have an average frequency greater than 4 and even bigger than $13/3$ for

big and large TSP. As these frequency bounds are used to trim the *OPs* having a smaller frequency, a lot of the ordinary edges and *OPs* are trimmed. The experimental results illustrate the frequency bounds work well for the real-life instances.

Acknowledgments. The authors acknowledge financial support from the National Key Research and Development Program of China (No. 2022YFE0207000), the Fundamental Research Funds for the Central Universities (No. 2024JC006) and the support by the State Key Lab of Alternate Electrical Power System With Renewable Energy Sources.

Data Availability Statements. The datum originates from TSPLIB and the hyperlinks is TSPLIB (uni-heidelberg.de).

Conflict of Interest Statement. The authors declared that there are no competing interests with the other persons, institutions and companies.

References

1. Karp, R.: On the computational complexity of combinatorial problems. *Networks (USA)* **5**(1), 45–68 (1975)
2. Yang, L.R., Wang, X.Y., He, Z.Q., Wang, S.C., Lin, J.: Review of Traveling Salesman Problem Solution Methods. *CCIS* 2062, 3–16 (2024)
3. Held, M., Karp, R.: A dynamic programming approach to sequencing problems. *J. Sco. Indust. Appl. Math.* **10**(1), 196–210 (1962)
4. Zhang, W., Sauppe, J.J., Jacobson, S.H.: Results for the close-enough traveling salesman problem with a branch-and-bound algorithm. *Comput. Optim. Appl.* **85**, 369–407 (2023)
5. Vo, T.Q.T., Baiou, M., Nguyen, V.H.: A branch-and-cut algorithm for the balanced traveling salesman problem. *J Comb Optim* **47**(4) (2024). <https://doi.org/10.1007/s10878-023-01097-4>
6. Applegate, D., et al.: Certification of an optimal TSP tour through 85900 cities. *Oper. Res. Lett.* **37**(1), 11–15 (2009)
7. M.S. Finding the right cutting planes for the TSP. *Journal of Experimental Algorithmics (JEA)* **5**(6), 1–16 (2000)
8. Pop, P.C., Cosma, O., Sabo, C., Sitar, C.P.: A comprehensive survey on the generalized traveling salesman problem. *Eur. J. Oper. Res.* **314**(3), 819–835 (2024)
9. Wang, Y.: An approximate method to compute a sparse graph for traveling salesman problem. *Expert Syst. Appl.* **42**(12), 5150–5162 (2015)
10. Wang, Y., Rummel, J.B.: A binomial distribution model for the traveling salesman problem based on frequency quadrilaterals. *J. Graph Algorithms and Applications* **20**(2), 411–434 (2016)
11. Wang, Y.: The polynomial randomized algorithm to compute bounded degree graph for TSP based on frequency quadrilaterals. *CCIS* **1693**, 77–95 (2022)
12. Wang, Y.: Sufficient and necessary conditions for an edge in the optimal Hamiltonian cycle based on frequency quadrilaterals. *J. Optim. Theory Appl.* **181**(2), 671–683 (2019)
13. Wang, Y.: Finding the edges in optimal Hamiltonian cycles based on frequency quadrilaterals. *Theoretical Computer Science* **986**, article No.114323 (2024)
14. Mittelman, H.: NEOS Server for Concorde (2023). <http://neosserver.org/neos/solvers/co:concorde/TSP.html>
15. Reinelt, G.: <http://comopt.ifl.uni-heidelberg.de/software/TSPLIB95/> (2023)



Speech Recognition Method Based on Improved Transformer

Liqiang Wang¹, Dong Lv¹, Wencheng He¹(✉), Haiyang Wang², Yifan Dong²,
and Mingshuang Sui²

¹ National Computer Network Emergency Response Technical Team/Coordination Center of
China, Beijing 100029, China
1009156962@qq.com

² Institute of Network Technology, Yantai 264003, China

Abstract. As speech recognition technology rapidly advances and becomes widely applied, it has become a key technology in many fields. However, the problems such as information forgetting, context dependence and inaccurate semantic understanding under long speech fragments seriously affect the accuracy and efficiency of speech recognition, and restrict its promotion and application in a wider range of scenarios. In this paper, a speech recognition model based on improved Transformer is proposed, which introduces LSTM network, and pays more attention to the context characteristics of speech data to obtain good recognition results. Finally, the experiment is carried out using AISHELL-3 speech data set, and the experimental results at the end of this paper indicate that the method proposed has low Character Error Rate and Sentence Error Rate. Compared with the other two recognition models, the proposed method reduces the CER and SER by 5.73% and 13.37% respectively compared with Transformer, and reduces the CER and SER by 4.42% and 9.68% respectively compared with DeepSpeech2, verifying the effectiveness of the method.

Keywords: Speech Recognition · Deep Learning · Transformer · LSTM

1 Introduction

Speech Recognition, also referred to as Automatic Speech Recognition (ASR), focuses on converting human speech content into text through computers, so that computers can “dictate” continuous speech spoken by humans, and realize the conversion of sound to text. At present, ASR technology has been deeply involved in many fields, such as sentiment analysis, educational assistance, medical health, and cross-cultural communication, etc. Its wide application has been elaborated in [1–3].

As pointed out [4], traditional speech recognition models focused on extracting acoustic features from audio signals [5]. These acoustic features such as LPC, LPCC and MFCC play an important role in ASR system because they are able to capture key information related to human speech pronunciation in audio signals. At the same time, in view of the common noise problem in speech signals, the research [6] focused on

the single-channel time-domain denoising method, which further improved the performance of speech recognition system by optimizing the denoising strategy. Dua et al. [7] conducted a systematic study on noise challenges in ASR systems and sorted out related technologies. A lightweight speech enhancement algorithm [8] based on Time Frequency Convolutional Networks (TFCN) was proposed to effectively suppress background noise and improve speech quality. With the emergence and rise of deep learning technology, various techniques for combining it with acoustic features have emerged. For example, Kumar et al. [9] cleverly used recurrent neural networks (RNN) to mine high-level features in speech signals to improve the accuracy of recognition. For dialect speech recognition, an attention mechanism-based end-to-end encoder-decoder system [10] was constructed. An automatic hate speech recognition system based on Support Vector Machine was constructed [11]. Almadhor et al. [12] used SCNN to extract spatial features of speech signals, while MHAT was used to process and analyze the extracted features, and combined the information with visual features, bringing a new perspective to the field of ASR.

Since Transformer model [13] was proposed, breakthroughs have been made in natural language processing, image processing, multimodal and other fields such as text and video, and speech recognition is no exception. Significant progress was made in the application of Transformer to the recognition of Dravidian languages [14]. Shi et al. [15] applied Transformer to speech recognition and proposed a caching mechanism to save the calculation results of keys and values in attention calculation and reused them in subsequent processing, thus reducing the amount of calculation in Transformer. In addition, a series of improved models based on Transformer have also been heavily used in the speech recognition field [16][17], achieving outstanding recognition performance. An efficient Transformer: Squeezeformer [16] was proposed to simplify the recognition model architecture. A Dysarthric Speech Transformer [17] was developed, which used the improved Deep Transformer Architecture to improve the recognition ability of audio data with Dysarthric characteristics. However, in practice, the context information of speech signals often spans long time intervals, and the ability of the recognition model to handle long-term dependencies is not considered in the above work. This can reduce the continuity of speech recognition tasks and thus affect their accuracy, especially when dealing with longer-term audio data. To address the above issues, this paper introduces LSTM into the original Transformer model. The gate control mechanism (forgetting gate, input gate and output gate) designed by LSTM can capture the medium and long term dependency relationship of voice processing tasks and use the characteristics of its memory unit to allow the model to effectively learn and store long sequences of information. This improves continuity and the accuracy of ASR system. The primary contributions include the following three points in this paper:

1. Aiming at the common noise problem of voice data, the preprocessing methods such as filtering and echo suppression are used to reduce the noise of the original voice signal, so as to reduce the interference of background noise for the subsequent speech recognition to a certain extent.
2. This paper designs an optimized speech recognition model based on Transformer. In view of the limitations of the current speech recognition model in long-term dependency processing, LSTM network is creatively introduced into the original speech

recognition model. By considering the timing relationship in the speech recognition process, the recognition effect of the model is optimized to some extent in terms of long speech fragments.

3. This paper sets the comparative experiments to evaluate the feasibility and effectiveness of the method proposed. Based on AISHELL-3 data set, design comparative experiments and verify that the model of this paper constructed has 5.73% and 13.37% lower Character and Sentence Error Rate than Transformer speech recognition model, and 4.42% and 9.68% lower Character and Sentence Error Rate than DeepSpeech2 speech recognition model [18].

2 Speech Recognition System

This chapter mainly introduces the speech recognition technology scheme of this paper, which mainly includes the following parts: speech data preprocessing, feature extraction, and Transformer-LSTM speech recognition model. The architecture diagram of the whole scheme is illustrated in Fig. 1.

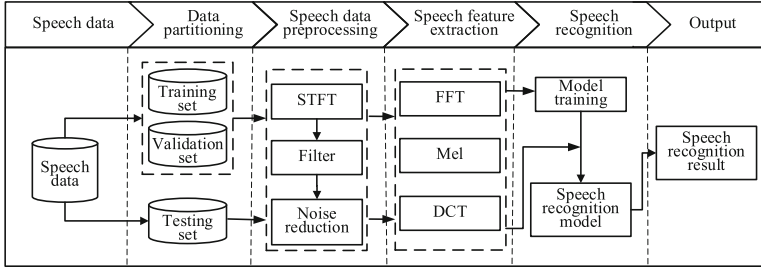


Fig. 1. Flow chart of speech recognition technology

2.1 Data Preprocessing

The purpose of the data preprocessing session of this chapter is to reduce the interference of background noise of the original input speech signal, so as to enhance recognition accuracy of the subsequent speech recognition process and achieve better recognition effects. It is an important part in ASR task. The preprocessing of speech data mainly includes the following four steps.

Step 1: Perform short-time spectrum conversion on the voice data to convert the time-domain signals into the frequency-domain signals for subsequent processing.

$$Y(f, t) = \sum_{n=0}^{N-1} y(n) \cdot h(n-t) \cdot e^{-\frac{j2\pi fn}{N}} \quad (1)$$

In Eq. (1), $y(n)$ is the time domain signal, $h(n-t)$ is the window function, N is the sampling number of short-time spectrum conversion, j is the imaginary number unit, $Y(f, t)$ is the frequency domain signal.

Step 2: Perform adaptive weight adjustment on the frequency domain signal to eliminate echoes and residual noise. The signal quality is optimized by dynamically adjusting the filter weight.

$$W_{k+1} = W_k + \eta \frac{Y^*(f, t)\varepsilon(t)}{\|Y(f, t)\|^2 + \alpha} \quad (2)$$

In Eq. (2), W_k is the current filter weight matrix, η is the step factor, $\varepsilon(t)$ is the error signal, α is the stability factor, and $Y^*(f, t)$ represents the conjugate complex number.

The signal $W_k \cdot Y(f, t)$ after adaptive weight adjustment is used to suppress the echo, the echo component in the signal is removed, and the signal $y_{res}(t)$ after echo suppression is obtained by subtracting it from the original signal.

$$y_{res}(t) = y(t) - W_k \cdot Y(f, t) + \lambda \sqrt{\sum_{i=0}^{I-1} |W_i \cdot Y_i(f, t)|^2} \quad (3)$$

In Eq. (3), $y_{res}(t)$ is the signal after echo suppression, λ is the regulatory factor, I is the cumulative weight and the number of input signals.

Step 3: The signal after echo suppression is enhanced to further enhance the clarity of the voice signal by reducing the background noise and enhancing the signal details.

$$Y_{aug}(f, t) = \begin{cases} \sqrt{\frac{|Y(f, t)|^2 - N(f, t)}{|Y(f, t)|^2}} |Y(f, t)| & \text{if } |Y(f, t)|^2 > \gamma |N(f, t)| \\ (1 - \lambda) |Y(f, t)| & \text{otherwise} \end{cases} \quad (4)$$

In Eq. (4), $Y_{aug}(f, t)$ is the upgraded frequency domain signal, $N(f, t)$ is the noise estimation, and λ, γ is the control parameter.

Step 4: Convert the enhanced frequency domain signal back to the time domain to obtain the enhanced speech signal.

$$y_{aug}(t) = \sum_{n=0}^{N-1} Y_{aug}(f, t) \cdot h(n - t) \cdot e^{\frac{j2\pi fn}{N}} \quad (5)$$

In Eq. (5), $y_{aug}(t)$ is the time domain signal after enhanced processing, representing the clear speech signal obtained after the final processing.

2.2 Feature Extraction

Speech feature extraction is an important step in speech recognition. Common ones include Perceptual Linear Predictive (PLP) and Mel-Frequency Cepstrum Coefficients (MFCC).

The pre-emphasis, framing, window addition, FFT, Mel filtering, DCT transformation are performed on the above preprocessed speech signal to finally obtain the MFCC acoustic characteristics, and the subsequent speech recognition tasks are carried out. Its flow chart can be seen in Fig. 2.

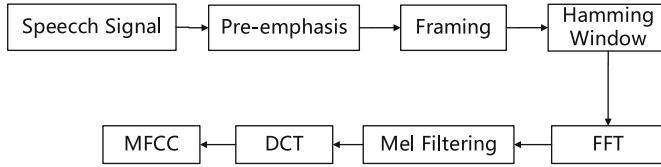


Fig. 2. Flow chart of speech feature extraction

2.3 Transformer-LSTM Speech Recognition Model

This paper aims to further optimize the timing features of speech signals, a multi-layer structure is constructed, which includes feedforward neural network, LSTM module and multi-head self-attention mechanism (see Fig. 3). Extracted speech features are input into the hybrid model based on Transformer and LSTM network and through the process of feature vector mapping, LSTM layer time series processing, feature transformation and layer normalization processing, etc. Finally, the model output is converted into character probability distribution through linear transformation and Softmax function, and the text sequence is gradually generated. Realize speech recognition.

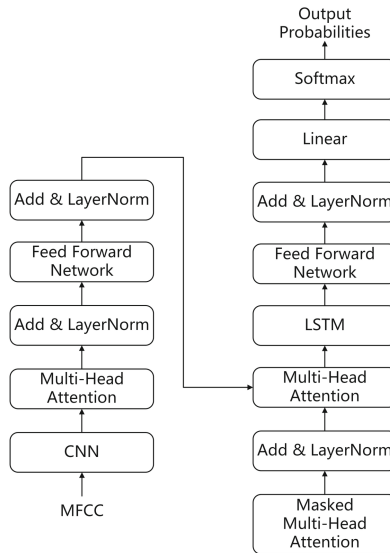


Fig. 3. Speech recognition model based on Transformer-LSTM

Multi-Head Attention Mechanism. When processing features in speech recognition, the multi-head attention mechanism can improve the model's ability to process speech features and extract richer audio features. It is expressed as multiple independent attention heads calculation, and then the results are concatenated and processed by linear transformations:

$$A(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

$$\text{head}_i = A\left(QW_i^Q, KW_i^K, VW_i^V\right) \quad (7)$$

$$MH(Q, K, V) = \text{Con}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) \cdot W_0 \quad (8)$$

where, Q refers the query matrix, which is used to determine which features to focus on for the current input, K refers the key matrix, which is used to help the attention mechanism determine which part of the input needs attention, and V refers the value matrix, which is weighted by attention weights to form the output of the attention mechanism. $\sqrt{d_k}$ refers a scaling factor and d_k represents the dimensional size of the bond matrix. Con refers to the concatenation operation, head_i is the output of the i attention head, W represents weight matrix, and h is the total number of attention heads.

By computing multiple attention heads in parallel, the Multi-Head Attention Mechanism is able to capture key features in different contexts, resulting in richer attention features:

$$A = \text{MultiHead}(Q, K, V) \quad (9)$$

LSTM Module. LSTM precisely controls the process of reading, writing and forgetting cell state through the gating mechanism (see Fig. 4), the process is as follows:

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_i) \quad (10)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (11)$$

$$h_t = o_t \odot \tanh(C_t) \quad (12)$$

where, W_C refers the weight matrix input and hidden to the candidate cell state, and b_i refers the bias. C_t refers the cell state at the current moment, f_t refers the output of the forgetting gate, C_{t-1} refers the cell state at the previous moment, i_t refers the output of the input gate, $\tilde{C}_t$ refers the candidate cell state, h_t refers the hidden state at the current moment and o_t refers the output of the output gate.

Feedforward Neural Network. The FNN consists of two linear layers that use the ReLU activation function for nonlinear transformation and the Dropout layer to prevent overfitting from occurring.

$$FFN(x) = \max(0, xW_1 + b_1)W_2 + b_2 \quad (13)$$

In Eq. (13), W_1 represents the first layer weight matrix of the FNN, b_1 represent the bias term of the first layer in the FNN, and W_2 represents the second layer weight matrix of the FNN, b_2 represents the bias term of the second layer of the FNN.

Finally, the model output is mapped to the space of output characters through the full connection layer, and the full connection layer output is converted into the probability distribution through the Softmax function to generate the final text sequence.

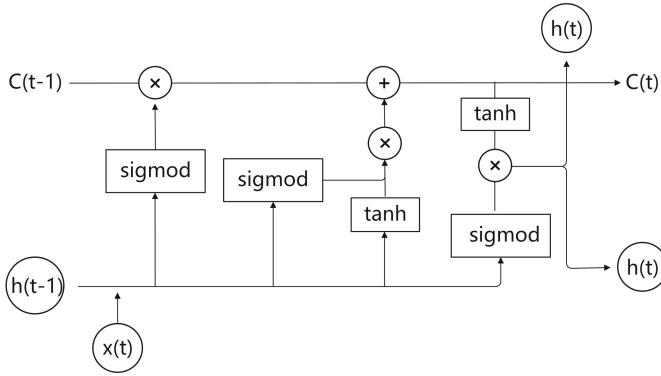


Fig. 4. Neural network model for long and short memory

3 Experimental Results and Analysis

3.1 Experimental Environment

Chinese Mandarin speech database AISHELL-3 data set was adopted in this paper. The speech duration of this data set was 85 h, with a total of 88,035 sentences, covering a variety of accents and different speech environments. After professional speech proof-reading, phonetic and prosodic labeling was carried out, and strict quality inspection was passed.

In this paper, the various experiments are conducted based on Pytorch deep learning framework, and other configurations are shown in Table 1. The dimension of the input module of the speech recognition model is set to 256, the dimension of the feedforward neural network layer is 1024, the attention heads are set to 4, the LSTM layer is 3, and the dimension of the network recursive layer is 1024, and each module is connected by residual error. Using the learning rate adaptive change strategy, the starting learning rate was established at 0.001, and the training process was optimized by Adam optimizer. A total of 100 epochs were trained.

Table 1. Hardware configuration and software environment.

Hardware configuration	CPU	Intel Core i9-12900H
	GPU	NVIDIA RTX 3060
	Internal memory	16 GB
Software environment	Operation system	Windows10
	Programming language	Python3.8
	Deep learning framework	Pytorch2.0.1

3.2 Evaluation Index

In the speech recognition system, evaluation indicators usually include Sentence Error Rate (SER), Word Error Rate (WER) and Character Error Rate (CER). As the evaluation is Chinese data set, this paper selects the Character Error Rate and Sentence Error Rate (SER) to evaluate the recognition effect. CER represents the proportion of error characters to the total number of characters in the speech recognition result, which reflects the recognition accuracy of speech recognition system at character level. SER represents the proportion of the number of sentences that are identified incorrectly to the total number of sentences, and the character recognition errors in a sentence is identified as the sentence recognition errors. Specific definitions are as follows:

$$CER = \frac{S_char + D_char + I_char}{N_char} \times 100\% \quad (14)$$

$$SER = \frac{S_Error}{S_Total} \times 100\% \quad (15)$$

where S_char , D_char , I_char are the total number of characters that have been substituted, deleted and inserted, respectively. N_char refers the total number of characters in the original sentence. Because the smallest unit in Chinese is the character, the statistical recognition result is also carried out by the character. S_Error refers the number of sentences that are incorrectly identified, and S_Total refers the total number of sentences in the evaluation process. It is important to note that the Sentence Error Rate treats each sentence as a single sample of whether it is correct or incorrect.

3.3 Results and Analysis

In this paper, to evaluate the recognition performance of the speech recognition model, current mainstream Transformer model [10] and DeepSpeech2 [18] speech recognition model are selected for comparison. Table 2 shows the respective calculation results of CER and SER of different speech recognition models based on the AISHELL-3 data set.

Table 2. Evaluation index results of different models.

ASR Model	CER	SER
Transformer [10]	13.41%	51.02%
DeepSpeech2 [18]	12.10%	47.33%
Transformer-LSTM	7.68%	37.65%

Experimental results indicate that the Character Error Rate of recognized characters of the model of this paper constructed on AISHELL-3 data set can reach 92.32%, which reflects the effectiveness of the model in speech recognition tasks. It can also be seen from the comparison of value views in Table 2 that the model of this paper constructed has 5.73% and 13.37% lower Character and Sentence Error Rate than Transformer speech

recognition model, and 4.42% and 9.68% lower Character and Sentence Error Rate than DeepSpeech2 speech recognition model, indicating that the proposed model has certain advantages in performance.

Due to the timing of speech signals, the Transformer-LSTM speech recognition model can effectively capture long-term dependency relationships in speech sequences by introducing LSTM network structure, further strengthen memory of information of a long time in the past, process and understand contextual semantics, and learn more accurate speech features. Thus improve the accuracy of speech to text conversion. During the statistical process, it was found that the Transformer would incorrectly recognize words in some speech fragments as words that sound similar to them, while the Transformer-LSTM could more accurately understand the semantics of the whole sentence and convert the speech into correct text expression.

4 Conclusion

In this paper, we have designed an optimized speech recognition model based on Transformer. LSTM module was introduced to focus on the context features of voice data. Based on the speech recognition model, the timing relationship of voice signals is strengthened and memory is strengthened in combination with voice data pre-processing and feature extraction. The recognition accuracy of long speech fragments is improved effectively. The proposed method was compared with Transformer and PaddleSpeech2 models on AISHELL-3 data set. Through the calculation and statistics of the experimental results, it is found that the model of this paper constructed has 5.73% and 13.37% lower Character and Sentence Error Rate than Transformer speech recognition model, and 4.42% and 9.68% lower Character and Sentence Error Rate than DeepSpeech2 speech recognition model. It indicates that the model of this paper constructed has certain application value, but it still needs to be perfected. In the future, the network structure will be further optimized to reduce the speech recognition model complexity, and the data set size will be expanded or data enhancement will be adopted to improve the normalization ability of speech recognition model. In addition, some post-recognition translation optimization technologies will be considered to realize the correction processing of similar pronunciation, rare characters or words, and further improve the recognition accuracy.

References

1. Shadiev, R., Liu, J.: Review of research on applications of speech recognition technology to assist language learning. *ReCALL* **35**(1), 74–88 (2023)
2. Kheddar, H., Hemis, M., Himeur, Y.: Automatic speech recognition using advanced deep learning approaches: a survey. *Information Fusion* **109**, 102422 (2024)
3. Al-Fraihat, D., Sharrah, Y., Alzyoud, F., et al.: Speech recognition utilizing deep learning: a systematic review of the latest developments. *HCIS* **14**, 15 (2024)
4. Basak, S., Agrawal, H., Jena, S., et al.: Challenges and limitations in speech recognition technology: a critical review of speech signal processing algorithms, tools and systems. *CMES-Computer Modeling in Engineering and Sciences* **135**(2), 1053–1089 (2023)
5. Alim, S.A., Rashid, N.: *Some Commonly Used Speech Feature Extraction Algorithms*. IntechOpen, London, UK (2018)

6. Kinoshita, K., Ochiai, T., Delcroix, M., et al.: Improving noise robust automatic speech recognition with single-channel time-domain enhancement network. In: ICASSP 2020–2020 IEEE International Conference on Acoustics. Speech and Signal Processing (ICASSP), pp. 7009–7013. IEEE, Barcelona (2020)
7. Dua, M., Akanksha, Dua, S.: Noise robust automatic speech recognition: review and analysis. *International Journal of Speech Technology* **26**(2), 475–519 (2023)
8. Li, L.: Design and implementation of intelligent interaction system: based on face recognition and speech recognition technology. *Computer Science and Application* **14**(6), 81–93 (2024)
9. Kumar, T., Rajest, S, Villalba-Condori, K., et al.: An evaluation on speech recognition technology based on machine learning. *Webology* **19**(1), 646–663 (2022)
10. Mao, Y.: Research on key technology of dialect speech recognition and its application in air conditioning. *J. Appliance Science & Technol.* (S1), 167–171 (2022)
11. William, P., Gade, R., Chaudhari, R., et al.: Machine learning based automatic hate speech recognition system. In: 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), pp. 315–318. IEEE, Erode (2022)
12. Almadhor, A., Irfan, R., Gao, J., et al.: E2E-DASR: end-to-end deep learning-based dysarthric automatic speech recognition. *Expert Syst. Appl.* **222**, 119797 (2023)
13. Vaswani, A., Shazeer, N., Parmar, N., et al.: Attention is all you need. in: advances in neural information processing systems. Neural Information Processing Systems Foundation, Long Beach, pp. 5999–6009 (2017)
14. Chowdary, D, Ganesan, R., Dabbara, H., et al.: Transformer-based multilingual automatic speech recognition (ASR) model for dravidian languages. *Automatic Speech Recognition and Translation for Low Resource Languages*, pp. 259–273 (2024)
15. Shi, Y., Wang, Y., Wu, C., et al.: Emformer: efficient memory transformer based acoustic model for low latency streaming speech recognition. In: ICASSP 2021–2021 IEEE International Conference on Acoustics. Speech and Signal Processing (ICASSP), pp. 6783–6787. IEEE, Toronto (2021)
16. Kim, S., Gholami, A., Shaw, A., et al.: Squeezeformer: an efficient transformer for automatic speech recognition. *Adv. Neural. Inf. Process. Syst.* **35**, 9361–9373 (2022)
17. Shahamiri, S, Lal, V., Shah, D.: Dysarthric speech transformer: a sequence-to-sequence dysarthric speech recognition system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **31**, 3407–3416 (2023)
18. Amodei, D., Ananthanarayanan, S., Anubhai, R., et al.: Deep Speech 2: end-to-end speech recognition in english and Mandarin. In: 33rd International Conference on Machine Learning, pp. 312–321. International Machine Learning Society, New York (2016)



A Two-Stage Approach to Multi-Feature Acoustic Scene Classification

Liqiang Wang¹, Yitao Li¹, Shan Ning¹(✉), Haiyang Wang², Jie Chu²,
and Qirong Zhou²

¹ National Computer Network Emergency Response Technical Team/Coordination Center of
China, Beijing 100029, China
406710907@qq.com

² Institute of Network Technology, Yantai 264003, China

Abstract. The analysis and classification of diverse noises in the urban environment have emerged as a research focus. Aiming at the limitations of existing acoustic feature classification algorithms in terms of accuracy and background noise processing, this paper proposes a two-stage multi-feature acoustic scene classification network. Firstly, a secondary feature extraction module is proposed to combine the Mel-Frequency Cepstral Coefficients (MFCC) and the first-order differential MFCC for filtering the noisy signals and preliminary feature refinement. Immediately after that, in order to further improve the classification performance, the double pooling residual module is proposed, which can deeply capture the contextual feature information in the acoustic scene and enhance the network's ability to understand the complex acoustic environment. Judging from the experimental outcomes, in contrast to the baseline network, this algorithm has boosted the classification accuracy of urban environmental sounds by 1.6%.

Keywords: Acoustic feature classification · Two-stage algorithms · Deep learning · Urban environmental sound classification

1 Introduction

Acoustic scene classification (ASC) constitutes a task where the features of audio signals are analyzed to attain the acoustic semantic categorizations of diverse scenes. For example, audio signals can be classified into labels from different life scenes, such as bus, office, or home. ASC is currently widely used and can be applied to many fields such as audio retrieval [1], smart home [2], city planning [3], and so on. For example, when the ASC system detects excessive traffic noise in the city, more green belts and soundproof walls can be set up in the city.

Traditional acoustic feature classification techniques [1, 4, 5] mainly rely on feature of sound signals to realize the classification of sound. These characteristics generally encompass features in the time domain, frequency domain, and time-frequency domain. However, as the problems pointed out in [6] and [7], traditional acoustic feature classification techniques rely on manual design, and their time-varying as well as nonlinear

feature processing capabilities are insufficient. Due to these limitations, traditional techniques often struggle to achieve the desired classification results when dealing with complex and variable acoustic environments.

In general, the acoustic scene contains numerous sound sources, which each have obvious unstructured characteristics in the audio segment of the environment, and the cross-class and intra-class variations of the environmental sounds are also complex and variable, which are difficult to be represented by a single feature [8]. During recent years, acoustic features that have gained considerable popularity comprise Constant-Q Transform (CQT) characteristics, Log-Mel Spectrogram, Gammatone Spectrogram, and MFCC. To address the above challenges, many studies have combined these several features. Wang et al. [9] fused four spectrogram representations such as log-Mel Spectrograms, CQTs, Gammatone Spectrograms, and MFCC features into a feature-level fusion to accommodate multi-category acoustic scene classification. With the rise of deep learning, there are also many studies [10–12] that combine traditional feature extraction techniques with deep learning techniques. For example, [3] classifies sound in two ways, one is to process the sound to obtain hand-crafted features by signal processing, and the other way is to represent the sound visually and classify the hybrid features of the two classes using Convolutional Neural Network (CNN).

However, there are few studies for background noise processing in audio. LOSTANLEN [13] employed single-channel energy normalization to reduce static noise and enhance transient sound events in environmental recordings. Wu [14] utilized two image processing edge detection methods based on differential Gaussian filtering and Sobel filtering to enhance the edge structure in the Mel spectrogram. This approach makes use of the structural characteristics of audio signals in the spectral domain to a certain extent. However, it also bears the risk of information loss and may weaken those frequency components that are crucial but not sufficiently important for classification.

LOSTANLEN [13] and Wu [14] directly noise-reducing the raw sound data mitigates to some extent the problem of classification errors caused by background noise. However, noise reduction on raw data may lose some useful information, especially when the noise is strong or the data volume is large. Based on the above considerations, this paper proposes a new acoustic feature processing scheme by designing a two-stage feature extraction method, where two acoustic features are acquired in the feature extraction stage and fused with secondary extraction, thus filtering the features that may contain noise or redundant information. Meanwhile, a double pooling residual block is designed after the secondary feature extraction for further feature learning. This approach can enable the model to focus on the most pertinent features and boost the efficacy of classification tasks. In comparison to approaches that filter the original noise, the feature extraction method involving secondary extraction of acoustic features can remove noise in the features more effectively without unduly influencing the information of the original data.

This paper proposes a two-stage acoustic scene classification strategy with the following main contributions:

1. a complete acoustic scene classification model is constructed, and a two-stage multi-feature acoustic scene classification network is designed for classification training and testing against input features.

2. Based on the work of primary feature extraction, a secondary feature extraction module is creatively added. The principal component analysis method is employed for dimensionality reduction to achieve the effect of noise reduction, thereby enhancing the accuracy of the extracted features. The accuracy of the downstream classification task is further improved by improving the accuracy of the upstream data.
3. Innovatively proposed dual-pooling residual blocks, aiming to further strengthen the feature learning capability by combining the dual-pooling mechanism of feature and residual learning.
4. The comparative experiment is intended to validate the credibility and efficacy of the suggested solution. For experimental verification, the UrbanSound8K dataset is employed. The verification outcomes indicate that, in comparison to the baseline network, the classification accuracy of the overall model is enhanced by 1.6%.

2 Method

This chapter focuses on the specific design of the two-stage acoustic feature classification model. The study can be divided into two core parts: first, the basic framework and core concepts of the proposed two-stage acoustic scene classification network are comprehensively introduced. Immediately after that, the specific design details of each part are elaborated, especially the secondary feature extraction part and the double pooling residual block.

2.1 Introduction to the Main Algorithm

In general, since speech signals are nonlinear, preprocessing operations must be performed to extract their spectral features. In this paper, based on the classification model of traditional acoustic features, a secondary feature extraction part is added for noise filtering. Next, deep learning is performed on the refined features by constructing a feature learning network. Finally, the classifier is used to judge and output the results. The specific flow chart is shown in Fig. 1.

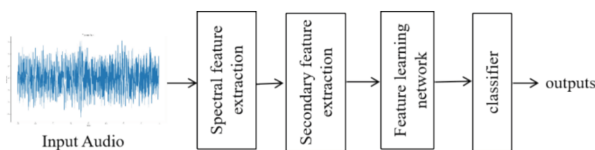


Fig. 1. Algorithm flowchart

2.2 Feature Introduction

The MFCC feature is a commonly used acoustic characteristic. In Fig. 2, the MFCC feature is obtained from the original sound through pre-emphasis, frame segmentation, Short-Time Fourier Transform (STFT), Mel filter bank, logarithmic operation, and Discrete Cosine Transform (DCT).

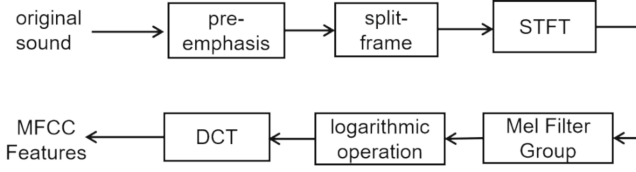


Fig. 2. Flowchart of MFCC feature generation

The detailed extraction procedure of MFCC characteristics is presented as follows.

Initially, a pre-emphasis operation is carried out. The aim is to boost the high-frequency portion of the speech signal, and this is primarily accomplished by means of a first-order high-pass filter.

$$H(z) = 1 - az^{-1} \quad (1)$$

where a is pre-emphasis factor, which assumes values ranging from 0.9 to 1.0.

After pre-emphasis, the frame-splitting and windowing operations are performed, by segmenting the speech signal and then processing each segment with a Hamming window. The primary objective of applying a window function is to avert spectral leakage and the picket fence phenomenon. The Hamming window can be expressed by the following equation:

$$w(n) = a_0 - (1 - a_0) \cos\left(\frac{2\pi n}{N - 1}\right) \quad (2)$$

where, $n \in \{0, 1, 2, \dots, N - 1\}$ is the index sample, N is the length of the frame and a_0 is taken as 0.54.

The STFT is then performed to convert the audio map to a spectrogram.

$$STFT(t, f) = \int_{-\infty}^{\infty} x(\tau) w(t - \tau) e^{-j2\pi f \tau} d\tau \quad (3)$$

where $x(\tau)$ is the input signal, $w(t - \tau)$ is the window function, t is the time displacement, f is the frequency, j is the imaginary unit, and $STFT(t, f)$ the output is f_s .

In general, the frequencies in the spectrum after the STFT are nonlinear, so a Mel filter bank is needed to convert them to nonlinear in order to simulate the perceptual properties of the human ear for sound frequencies. The conversion formula is:

$$m = 2595 * \log_{10}\left(1 + \frac{f_s}{700}\right) \quad (4)$$

Finally, the logarithmic output of the Meier filter bank is converted to cepstrum coefficients using DCT. The formula for DCT is shown below.

$$X[k] = \sum_{n=0}^{N-1} x[n] \cos\left(\frac{\pi k(2n + 1)}{2N}\right) \quad (5)$$

where $x[n]$ is the logarithmic energy spectrum, N is filter count, and is k the spectral cue lead of the DCT.

In addition, since the MFCC coefficients contain only spectral features, their first-order difference features are further extracted in order to obtain dynamic time-varying information, which $c_t[i]$ is realized as follows. Assuming that the i order MFCC coefficient at frame t is represented, the first-order difference feature of this coefficient is $\delta c_t[i]$.

$$\delta c_t[i] = \sum_{k=-2}^2 l c_{t+k}[i] \quad (6)$$

where, l stands for the dimension of the differential window., here taken as 2.

2.3 Two-Stage Multi-Feature Acoustic Scene Classification Network

This section introduces the overall framework of the two-stage multi-feature Acoustic Scene Classification Network (TM-ASCNet). Unlike traditional acoustic feature extraction, TM-ASCNet innovatively introduces secondary feature extraction, which reprocesses, fuses, and secondarily extracts the original acoustic features.

First, the MFCC features and first-order difference MFCC features are fed into TM-ASCNet for processing. Figure 3 shows the overall architecture of TM-ASCNet. In TM-ASCNet, the MFCC features and first-order difference MFCC features are deeply fused and processed by the secondary feature fusion module. Subsequently, the Principal Component Analysis (PCA) technology is utilized for additional feature extraction. This procedure can efficiently eliminate a portion of noise and enhance the accuracy of the feature signal. After the feature extraction operation, the features are fed into the N-layer double pooled residual block for feature learning. Finally, a classifier consisting of global average pooling and fully connected layers is employed for categorization.

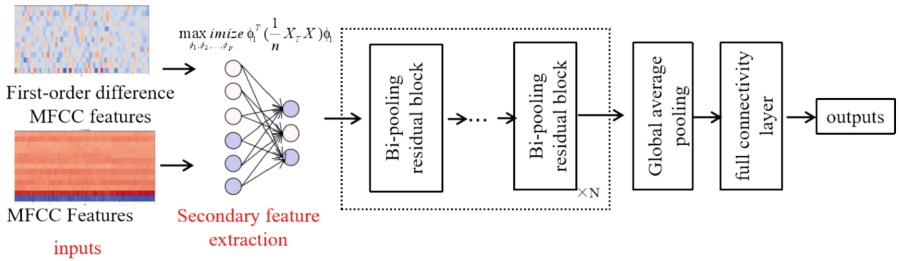


Fig. 3. Overall architecture of TM-ASCNet

2.3.1 Secondary Feature Extraction

Although MFCC exhibit robust capabilities in speech recognition, music information retrieval, and environmental sound classification, it unavoidably carries noise components to a certain extent. First of all, the original design of the Mayer filter is to simulate the nonlinear characteristics of the human ear in the perception of different frequencies,

and its smoothing process enhances the perceptual correlation of the signal, but also accidentally “amplifies” the effect of noise on the eigenvectors. This is because the spectral characteristics of the noise signal, especially the non-smooth noise, may be modulated by the shape of the filter as it passes through the filter bank, thus leaving more significant traces in the MFCC features. Second, the logarithmic operation and the DCT step in the MFCC computation process, although effective in compressing the feature dimensions and extracting key acoustic features, have the limitation of not being able to completely eliminate the noise interference. Although logarithmic operations enhance the dynamic range differences between signals, they may also amplify the tiny fluctuations in noise. And although the DCT performs excellently in energy concentration and decorrelation, in some cases, especially when the noise energy is relatively dispersed, it is very difficult to completely separate the signal from the noise.

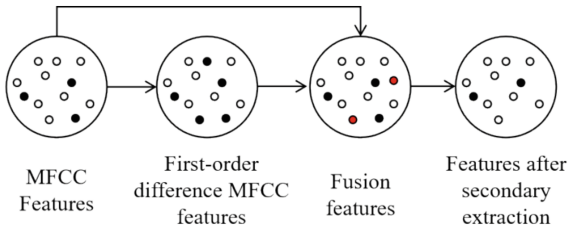


Fig. 4. Effect of secondary feature extraction

The white circles in Fig. 4 represent, the black circles represent the noise, and the red circles represent the noise that may increase after fusion. Generally, the initially extracted MFCC features contain some noise, which may increase after first-order differencing. Based on the above considerations, this paper proposes an innovative solution, i.e., a secondary feature extraction module is introduced after the traditional MFCC feature extraction process. The core idea of this module is to achieve deep optimization and enhancement at the feature level by combining spectral features and dynamic time features, so as to improve the robustness and differentiation ability of features.

Specifically, the module first performs the cascade fusion of MFCC features with their first-order difference features, which not only retains the static spectral information, but also incorporates the important temporal dynamic information, which helps to capture the subtle changes in the audio signal. Then, feature extraction and denoising are performed using PCA, which is a powerful data dimensionality reduction technique that can effectively remove redundant information caused by noise while preserving the main direction of variation in the data. In the secondary feature extraction module, PCA is skillfully applied to the fused feature set by identifying and retaining the most important feature dimensions while removing the noise-dominated low contribution dimensions.

2.3.2 Bi-pooling Residual Block

The intermediate feature learning module is the core part of the model structure. It employs N elaborately designed double-pooling residual blocks with the same structure

(the actual value of N in this paper is 4). The design of the double-pooling residual block is shown in Fig. 5. This design aims to extract and fuse features through multiple iterations and gradually enhance the model's ability to understand and describe the complexity of acoustic scenes. The double-pooling residual block realizes effective feature extraction and dimensionality reduction through the unique "pooling + convolution + pooling" structure while maintaining the transmission of key information. The design of the double-pooling residual block is as follows:

First, the first pooling layer uses average pooling with a pooling kernel of 3×3 and a stride of 1. This means that this layer will not change the spatial dimension of the feature map during the pooling process but can effectively decrease the redundant information of the feature map and enhance the local invariance of the features. This design helps to retain more detailed information and provides richer input for the subsequent convolutional layer.

Subsequently, in the middle part are two consecutive 3×3 convolutional layers. The main function of the first convolutional layer is to augment channel dimension (i.e., the number of features) of the feature map. This step aims to enhance the expressiveness of the model by increasing the diversity of features. Then, the second convolutional layer keeps the channel dimension unchanged and further refines and abstracts the features. At the same time, batch normalization (BN) and the nonlinear activation function ReLU are inserted between the two convolutional layers to accelerate the training process and introduce nonlinearity to boost the generalization capability of the model.

Finally, the second pooling layer is maximal pooling, which performs a downsampling operation to halve the spatial size of the feature maps (i.e., downsampling multiplicity of 2). This step not only reduces the computational effort, but also increases the abstraction level of the features, which helps the model to capture a larger range of contextual information.

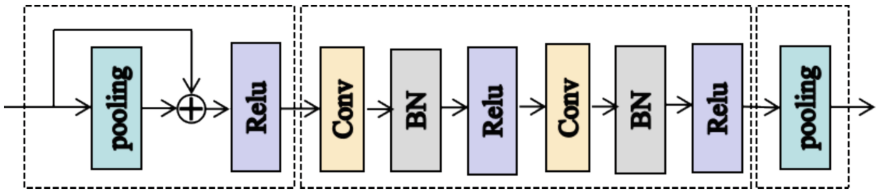


Fig. 5. Bi-pooling residual block design

2.3.3 Classifier

The classifier is responsible for mapping the learned features to specific category labels. In this paper, the classifier is composed of a Global Average Pooling Layer (GAPL) and three fully connected layers. The role of the GAPL is to perform effective scaling and feature aggregation on the feature maps. Specifically, it regards each channel of each feature map as an independent feature vector and calculates the average value of all elements within this vector, thereby obtaining a single value that only contains the global information of that channel. In addition, to prevent overfitting of the model, Dropout is

introduced after the GAPL to enhance the generalization ability of the model. Finally, three fully connected layers are used to gradually reduce the number of channels. The last fully connected layer adjusts the number of channels according to the classification categories to achieve the final classification.

3 Experimental

The experimental validation session is subdivided into two key components: experimental setup and experimental design. The experimental setup section describes the dataset on which the experiment is based and the parameter settings adopted for the experiment. The focal point of the experimental design segment lies in conducting a series of comparative experiments to validate the reasonableness of theoretical assumptions, assess the model's performance, and inspect the viability of technical solutions.

3.1 Experimental Setup

UrbanSound8K is a frequently utilized publicly accessible dataset for the classification of urban environmental sounds. Comprising 8,732 sound files in total, it encompasses ten classifications: the sound of air conditioners, car horns, children at play, dogs barking, drilling sounds, engine idling noises, gunshots, hand-held electric drills, police sirens, and street music.

The experiments were conducted on Ubuntu 20.04 operating system running on pytorch 2.1.0 based on python 3.10. The specific parameter settings were: optimizer using Adam, learning rate was 0.001, learning rate tuning strategy using WarmupCosineSchedulerLR. The batch size is 16 and epoch is set to 60. There are 7762 training sets and 970 test sets. The evaluation metrics used are accuracy rate, and the formula is as follows:

$$Accuracy = \frac{TP + TN}{P + N} \quad (7)$$

where TP stands for true positives, namely the count of positive samples that are predicted as positive by the model. TN represents true negatives, which are the number of negative samples predicted as negative by the model. P signifies the number of actual positive samples. N denotes the quantity of actual negative samples.

3.2 Experimental Design

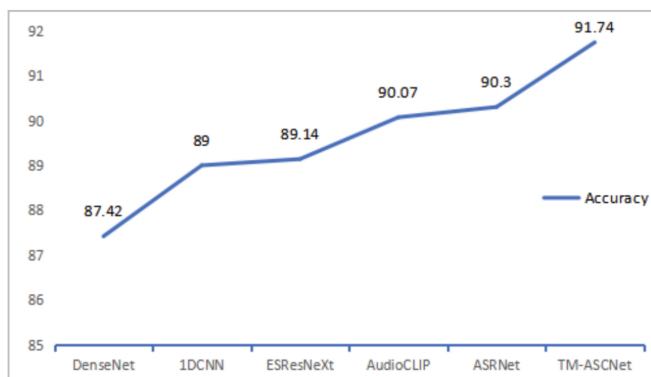
Table 1 provides an exhaustive list of the specific layer designs of the TM-ASCNet model, the data output dimensions of each layer, and the number of training parameters.

Furthermore, in order to comprehensively evaluate the performance of the model proposed in this paper on acoustic classification tasks and conduct a comparative analysis with the cutting-edge technologies in the current field of deep learning, the comparison results are plotted as a line graph (as shown in Fig. 6). In the comparison in Fig. 6, DenseNet [15], 1DCNN [16], AudioCLIP [17], and ASRNet [18] are selected for comparison with the model in this paper. The DenseNet model takes the lead in exploring

Table 1. TM-ASCNet model parameters

layer	Output data dimensions	Number of training parameters
convolution block	(1, 64, 20, 49)	1664
Bi-pooling residual block1	(1, 64, 10, 24)	73984
Bi-pooling residual block2	(1, 128, 5, 12)	221184
Bi-pooling residual block3	(1, 256, 2, 6)	884736
Bi-pooling residual block4	(1, 512, 1, 3)	3538944
Global average pooling layer	(1, 512, 1, 1)	0
full connectivity layer1	(1, 256)	131328
full connectivity layer2	(1, 256)	64792
full connectivity layer3	(1, 10)	2570

the effectiveness of CNN networks in the field of audio classification and by introducing three different types of classical CNN architectures for comparative experiments, it provides a new approach for audio feature deep mining provides new ideas. While 1DCNN applies the Gammatone filter bank to the initial extraction of audio features, followed by deep learning and classification of the extracted features through multiple well-designed convolutional layers. AudioCLIP seamlessly integrates audio signals into CLIP [19] (Contrastive Language-Image Pre-Training) This powerful cross-modal pre-training model not only broadens the boundaries of CLIP applications, but also opens up new avenues for correlation learning between audio and text. The ASRNet model achieves efficient comprehension and classification of speech signals by extracting the MFCC features and combining them with a convolutional plus pooled CNN network.

**Fig. 6.** Comparison of results with current methods

ASRNet was chosen for the baseline model. In terms of experimental design, a series of comparative experiments were designed with the aim of comprehensively evaluating the contribution of the different components of the model to the acoustic feature classification task. First, to ensure the fairness and reproducibility of the experiments, the baseline model was meticulously reproduced in strict accordance with the description in the original paper to ensure the consistency of the experimental conditions. Experiment 1 shows the baseline model reproduction results.

Then, in order to explore the possibility of feature deep mining, Experiment 2 further introduces a secondary feature extraction validation session. This session adds the secondary feature extraction module on the basis of the baseline model.

Finally, in order to verify the advantage of double-pooled residual blocks, Experiment 3 adds a classification network consisting of four consecutive double-pooled residual blocks on the basis of Experiment 2. By combining the advantages of global average pooling and maximum pooling, the dual-pooling residual block retains the important global information of the audio signal and highlights the importance of local features.

Table 2 lists the results of the ablation experiment. After adding secondary feature extraction to the baseline model, the accuracy is improved by 0.55%. After adding the double pooling residual block, the accuracy is improved by 1.6% with respect to the baseline.

Table 2. Module Design Validation Table

Number	secondary extraction	Bi-pooling residual block	accuracy
1	×	×	90.3
2	✓	×	90.8
3	✓	✓	91.74

To intuitively display the classification effect of each category, the confusion matrix of the prediction results is visualized as shown in Fig. 7. The horizontal axis stands for the prediction results, while the vertical axis denotes the true results. The values in the Fig. are normalized according to the true label values of each category. The 0–9 in the Figure represent the 9 sound categories. However, it is worth noting that the model still has some limitations when confronted with sounds that have similar acoustic characteristics and subtle differences between classes. For example, the low humming sound emitted by air conditioners, whose spectral characteristics partially overlap with those of drilling or engine idling, leads to occasional confusion in the model’s recognition process, misclassifying air conditioner sounds as drilling or engine idling sounds.

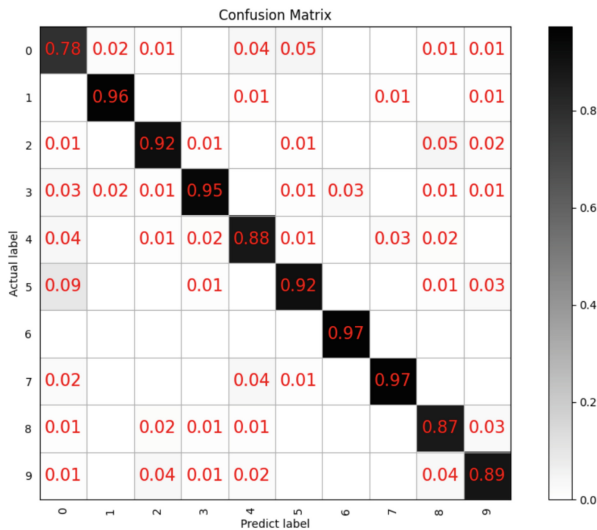


Fig. 7. Confusion matrix

4 Conclusion

In this paper, by designing a two-stage multi-feature acoustic scene classification model, the shortcomings of traditional acoustic feature classification algorithms in terms of accuracy and background noise processing are effectively overcome. Combined with the Mel frequency cepstrum coefficient and its secondary feature extraction of first-order difference features, as well as the design of double-pooling residual module, the accuracy of sound recognition is significantly improved. Experimental validation shows that the method shows significant improvement in classification accuracy over the baseline network, reaching an increase of 1.6%. The research in this paper provides certain technical support for urban environmental noise management, intelligent monitoring and other fields, and has certain application prospects and practical value, but there are still some limitations to be further extended. In future studies, one might contemplate delving deeper into ways to enhance the model’s generalization capability and mitigate the overfitting of the model to specific datasets by incorporating more diverse training samples and employing more advanced regularization technologies.

References

- 1. Cai, X., Zhang, H.: Music genre classification based on auditory image, spectral and acoustic features. *Multimedia Syst.* **28**(3), 779–791 (2022)
- 2. Papel, J.F., Munaka, T.: Home activity recognition by sounds of daily life using improved feature extraction method. *IEICE Trans. Inf. Syst.* **106**(4), 450–458 (2023)
- 3. Özseven, T.: Investigation of the effectiveness of time-frequency domain images and acoustic features in urban sound classification. *Appl. Acoust.* **211**, 109564 (2023)

4. Ye, J., Kobayashi, T., Murakawa, M., et al.: Robust acoustic feature extraction for sound classification based on noise reduction. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5944–5948 (2014)
5. Herrera, P., Yeterian, A., Gouyon, F.: Automatic classification of drum sounds: a comparison of feature selection methods and classification techniques. *International Conference on Music and Artificial Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 69–80 (2002)
6. Singh, V.K., Sharma, K., Sur, S.N.: A survey on preprocessing and classification techniques for acoustic scene. *Expert Syst. Appl.* **229**, 120520 (2023)
7. Ding, B., Zhang, T., Wang, C., et al.: Acoustic scene classification: a comprehensive survey. *Expert Systems with Applications*, 121902 (2023)
8. Kang, L., Ma, J., Zhang, Z.: Acoustic scene classification based on multi-feature late fusion. *Computer Engineering and Design* **44**(1), 141–147 (2023)
9. Wang, H., Zou, Y., Chong, D.: Acoustic Scene Classification with Spectrogram Processing Strategies. *arXiv preprint arXiv*, 2007.03781 (2020)
10. Mei, K., Liu, Z., Du, H., et al.: Considering temporal connection between turns for conversational speech synthesis. *ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 11426–11430 (2024)
11. Chen, L., Chen, J.: Deep neural network for automatic classification of pathological voice signals. *Journal of Voice* **36**(2), 288. e15–288. e24 (2022)
12. Sharma, A.K., Aggarwal, G., Bhardwaj, S., et al.: Classification of Indian Classical Music with Time-Series Matching Deep Learning Approach **9**, 102041–102052 (2021)
13. Lostanlen, V., Salamon, J., Cartwright, M., et al.: Per-channel energy normalization: why and how. *IEEE Signal Process. Lett.* **26**(1), 39–43 (2018)
14. Wu, Y., Lee, T.: Enhancing sound texture in CNN-based acoustic scene classification. *ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 815–819 (2019)
15. Palanisamy, K., Singhanian, D., Yao, A.: Rethinking CNN Models for Audio Classification. *arXiv preprint arXiv*. 2007.11154 (2020)
16. Abdoli, S., Cardinal, P., Koerich, A.L.: End-to-end environmental sound classification using a 1D convolutional neural network. *Expert Syst. Appl.* **136**, 252–263 (2019)
17. Guzhov, A., Raue, F., Hees, J., et al.: Audioclip: Extending clip to image, text and audio. *ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 976–980 (2022)
18. Zheng, W., He, W.: Multi-scene oriented sound recognition for environmental anomalies. *Science, Technology and Engineering* **23**(17), 7444–7449 (2023)
19. Radford, A., Kim, J.W., Hallacy, C., et al.: Learning transferable visual models from natural language supervision. *International conference on machine learning*. PMLR, pp. 8748–8763 (2021)



Star Extraction Based on Multi-scale Grey Morphology in Stray Light Backgrounds

Shuhui Zhou¹, Yang Zheng², Chenxin Lu², Jinwei Cheng², and Wanyu Liu¹(✉)

¹ Sino-European School of Technology of Shanghai University, Shanghai University,
Shanghai 200444, China
liuwanyu@shu.edu.cn

² Shanghai Aerospace Electronic Technology Institute, Shanghai 201109, China

Abstract. In the space-based space target monitoring system, stray light in the space environment causes very serious non-uniform distribution characteristics in the background of the star map, which greatly affects the accuracy in extracting star. Therefore, it is necessary to preprocess star images under stray light background. Traditional preprocessing methods have difficulty in dealing with complex stray light backgrounds, leading to poor star extraction performance. To address the issue, this paper introduces a multi-scale grayscale morphology method for star extraction under stray light background. First, median filtering is performed on the image to effectively suppress high-frequency noise and impulse noise and smooth the non-uniform background. Then, the image is further processed using dual-structure multi-scale grayscale morphological transformation to effectively suppress background noise and enhance the contrast of the target area. Then, adaptive threshold processing is performed to accurately separate star and background areas under different brightness conditions. Finally, the stars are extracted through connected domain analysis. The results of the semi-physical simulation experiment prove that the peak signal-to-noise ratio PSNR is 39 dB, the background suppression factor BSF is 5.729×10^{-4} , and the average running time of the method is 0.126 s. This method can greatly suppress background noise and greatly improve the accuracy of star extraction, enabling a higher processing speed.

Keywords: Space-based Star Maps · Non-uniform Background · Multi- scale Grey Scale Morphology · Star Extraction

1 Introduction

With the advancement of space exploration technology, the utilization and monitoring of outer space have become increasingly important. The space-based space target surveillance system has a large monitoring field of view, good on-board real-time monitoring effect, and good wartime survivability. Space-based space target surveillance will become an important development direction of future space target surveillance [1]. In astronomical image processing, star extraction is an important part of the space-based space target monitoring system. By extracting star, targets such as satellites and space debris can be identified and tracked, and help classify targets and calculate orbits. Due to

the influence of components, dark current, reflected light from the external environment, etc., the star map acquired by the camera contains a large amount of non-uniform system noise. The stray light and non-uniform background in the star map not only reduce the image quality, but also increase the difficulty of star extraction. The presence of complex noises such as detector noise and thermal noise further reduces the saliency of weak and small targets, resulting in the difficulty of separating weak and small targets. The traditional background and noise removal algorithms are not very targeted and are not suitable for star map background and noise removal. Therefore, how to reduce or eliminate the interference of background and noise in the preprocessing process is the primary problem that needs to be solved in space object detection.

In order to suppress non-uniform background and noise in the image and improve target saliency, preprocessing is required before target detection. Common early methods suppress image noise by designing spatial or frequency domain filters, such as median filtering, mean filtering and low-pass filters. These methods have good smoothing effects and anti-noise performance, but their computational complexity is high, the processing time is long when processing large-size images, and edge effects may be introduced. The grayscale morphology method is a nonlinear filtering method that uses the window to traverse the image by designing a structural window that is suitable for the target, and performs operations such as erosion and dilation in the window to accomplish background suppression. Bai [2] et al. improved the structure window of traditional morphological Top-hat, so that the method can better suppress the background and facilitate target detection. Ding Peng [3] et al. combined the Top-hat method to remove background interference and the Canny method to detect edges, and achieved good target detection results. Li Jie [4] et al. combined shear wave transform and top hat transform to achieve double suppression of the background, thereby completing the rough extraction of star. The above-mentioned method based on grayscale morphology has the advantages of simplicity, efficiency, and removal of background brightness changes. It can highlight the target and enhance the contrast between the target and the background. However, this category of method relies heavily on the shape and size of structural elements and is not suitable for processing complex backgrounds, which may lead to false detections or missed detections. In addition, it does not perform well when processing images with high noise levels and needs to be combined with other denoising methods. In recent years, methods based on local contrast measurement (Local Contrast Measure, LCM) have received widespread attention in the fields of image processing and target detection. The local contrast method (LCM) proposed by Chen [5] et al. can greatly improve the performance of target detection by designing a contrast calculation window to highlight the target area and suppress background interference. Han [6] et al. proposed a multi-scale relative local contrast measure (RLCM) to introduce relative contrast, so that the method can simultaneously enhance the target and suppress the background. Niu [7] et al. used median filtering to remove part of the noise, fast maximum filtering to screen target candidate areas, and then used a three-layer contrast window to suppress complex backgrounds and highlight target imaging features. The local contrast method improves the efficiency of the method by designing a complex contrast calculation window, but the complex contrast traversal calculation also greatly increases the method running time.

To address the above shortcomings, this paper introduces a method based on multi-scale grayscale morphology. In conclusion, our research presents the following key contributions:

- 1) The median filtering method was used to remove some noise and smooth the image, avoiding the effects of pretzel noise and highlighting noise.
- 2) The structural elements of the dilation and erosion operations are designed respectively to make full use of the grey scale difference between the target and the background, which greatly suppresses the stray light.
- 3) Adaptive thresholding method combined with connectivity domain analysis is used to complete star extraction.

2 Related Theories

2.1 Space-Based Star Image Imaging Characteristics

During the acquisition process, space-based star image data is affected by noise such as stray light, which will produce serious degradation effects, affecting the accuracy and performance of the entire detection system such as image segmentation, feature extraction, and target recognition [8]. When detecting space targets, the acquired space-based star image data contains a variety of information. According to its information source, the imaging characteristics of space-based star images can be described through the subsequent equation [9].

$$F(x, y) = T(x, y) + S(x, y) + B(x, y) + N(x, y) \quad (1)$$

where $F(x, y)$ represents the grayscale value of the pixel at the position (x, y) in the space-based star image, $T(x, y)$ is the space target, $S(x, y)$ is the star, $B(x, y)$ is the background in the space-based star image, and $N(x, y)$ is the noise in the space-based star image.

Ideally, a space-based star image should be composed of bright space targets and stars against a dark background. Under the influence of the point spread function in the optical system, the space targets and stars are approximately diffusely distributed in the form of a two-dimensional Gaussian function within a few or dozens of pixels [10], and the formula is as follows.

$$f(x, y) = \frac{S}{2\pi\sigma_{psf}^2} \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma_{psf}^2}\right) \quad (2)$$

Among them, (x_0, y_0) is the coordinate of the target center position, S is the total intensity of the target, and σ_{psf} is the diffusion standard deviation of the point spread function in the optical system. Ideally, deep sky backgrounds have almost zero brightness because they are in space far from light sources and there are no significant light sources illuminating these areas, which makes most pixels have near-zero values and the image appears black. Stars, although very bright, occupy very few pixels in the star map, often only a few pixels in area. The ideal star map distribution is shown in Fig. 1.

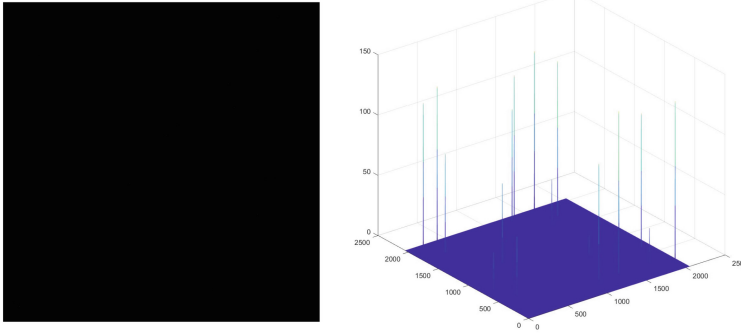


Fig. 1. Ideal star map distribution

2.2 Grayscale Morphological Filtering Method Theory

Matheron and Serra first proposed mathematical morphology and extended it to the field of image analysis [11]. Mathematical morphology provides a shape-based approach to image processing that operates in favour of simplifying image data, preserving basic shape features and eliminating distractions. Unlike filtering methods based on spatial or frequency domains, since morphological operators are directly related to shapes, they can intuitively obtain more essential characteristics such as the topological and structural information of objects. Therefore, compared with other forms of filtering methods, it has obvious advantages in the field of image processing [12]. Grayscale morphology is an extension of binary morphology in grayscale space. It mainly includes four basic operations: dilation, erosion, and opening and closing operations. The principle is to use structural elements for local max/min operations to preserve grey scale, suppress noise, remove background, and extract target shapes.

(1) Grayscale dilation

Let $B(m, n)$ be the structural element, $A(x, y)$ be the input grayscale image, the domain of $B(m, n)$ is D_B , the domain of $A(x, y)$ is D_A , then the grayscale image $A(x, y)$ is expanded by the structural element $B(m, n)$, denoted as: $(A \oplus B)$. It is defined as:

$$(A \oplus B)(x, y) = \max\{A(x - m, y - n) + B(m, n) | (m, n) \in D_B\} \quad (3)$$

(2) Grayscale erosion

Let $B(m, n)$ be the structural element, $A(x, y)$ be the input grayscale image, the domain of $B(m, n)$ is D_B , the domain of $A(x, y)$ is D_A , then the grayscale image $A(x, y)$ is eroded by the structural element $B(m, n)$, denoted as: $(A \ominus B)$. It is defined as:

$$(A \ominus B)(x, y) = \min\{A(x + m, y + n) - B(m, n) | (m, n) \in D_B\} \quad (4)$$

Although grayscale morphology-based methods have been widely used to correct non-uniform backgrounds in recent years [13], their processing capabilities are limited when the image background is complex and changeable and the target signal-to-noise

ratio is low. Additionally, such methods are also very sensitive to the size of the filter template. Fixed filter templates cannot fit all cases, and poor selection impacts performance. Fast and accurate correction of inhomogeneous backgrounds from stray light, while preserving star data, remains a key challenge.

3 Proposed Method

Due to sensor defects, data transmission errors, interference in the photoelectric conversion process, harsh environmental conditions, and damage during image compression and transmission, there will be some impulse noise and salt and pepper noise in the space-based star map. When performing grayscale morphological operations, impulse noise and high-frequency noise will introduce a large number of artifacts and false detections, reducing the accuracy of star detection. Through median filtering, high-frequency noise and impulse noise can be effectively suppressed, and the non-uniform background in the image can be smoothed, making the grayscale value difference between the star area and the background area more obvious, thereby improving the accuracy and robustness of subsequent morphological operations. This processing method can not only retain the edge information of the star, but also enhance the contrast between the star and the background, making the star detection more accurate and reliable. In order to reduce the influence of such noise on the subsequent star extraction, the median filtering method is used in this paper to smooth the image. The median filtering formula is shown below:

$$g(x_i, y_j) = \text{med} \{ f(x_{i+\zeta}, y_{j+\eta}) | (\zeta, \eta) \in W \} \quad (5)$$

Among them, *med* represents the median operation, *W* represents the two-dimensional filter template, and the filter window size set in this paper is 3×3 .

Due to the difference in grayscale values between the star area and the background area, it is possible for different methods to correct the non-uniform background and retain the star. However, although the previous methods took this difference information into account, due to the irrationality of the filter template, they could not effectively distinguish between the star and the background, and introduced the pixel values of the star area and the surrounding background area into the calculation. This greatly reduces the utilization of the difference information, resulting in poor accuracy of non-uniform background correction, and may cause the loss of low signal-to-noise ratio star.

In order to solve these problems, this paper combines the diffusion model of star and the distribution characteristics on the imaging surface to design two different but related structural elements for dilation and erosion, as shown in Fig. 2. Based on the experimental results, the structural element of the dilation operation consists of a circular structural element B_i nested in a square structural element B_o , where the B_i radius is 11 and the B_o side length is 29, as shown in Fig. 2 (a). The structural element B_m of the erosion operation consists of a circular structural element inscribed in a square structural element, where the circular structural radius is 20, as shown in Fig. 2 (b). In addition, the value of the circular structural element of the dilation structural element is 0, and the value of the remaining part is 1. The value of the circular structural element of the erosion structural element is 1, and the value of the remaining part is 0. The structural

element is essentially a small matrix used to determine the shape and range of the dilation and erosion operations, and the value of the structural element is a Boolean value. In the structural element, the pixel with a value of 1 represents the operation area, and the pixel with a value of 0 represents the non-operation area. Dilation is to expand the “1” in the structural element into the image, making the target object larger and filling holes or small breaks. Erosion is to shrink the target area in the image, remove small noise or the edge of the target object, and make the object smaller.

This structure makes full use of the smoothness of the circular structural element and the coverage of the square structural element to improve the accuracy and robustness of the morphological dilation operation.

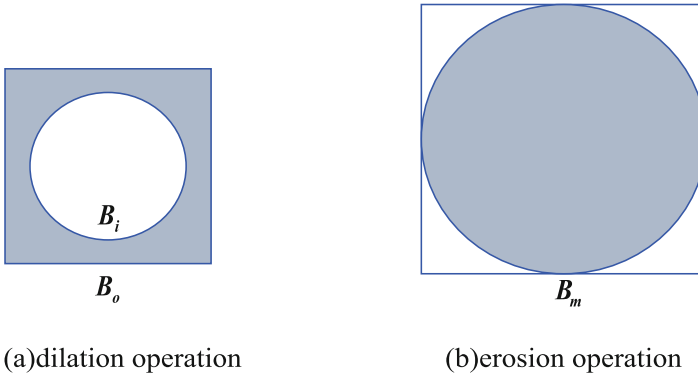


Fig. 2. Dilation and erosion operation

In order to protect the star area, the structural element B_{io} is first used to dilate the star map $f(x, y)$ containing the non-uniform background to obtain the processed image $F_1(x, y)$, which is expressed as follows:

$$F_1(x, y) = f(x, y) \oplus B_{io} \quad (6)$$

The star area is protected by the dilation operation of the structural element B_{io} , and the pixels in the star area are replaced by the pixels in the surrounding background area.

Then, the image $F_1(x, y)$ is eroded using a structure element B_m that is slightly larger than B_{io} to obtain the processed image $F_2(x, y)$, which is expressed as follows:

$$F_2(x, y) = F_1(x, y) \ominus B_m \quad (7)$$

After the above morphological operations, the star with a size smaller than B_i are retained, and the remaining star with a size larger than B_i still exist in the non-uniform background. After the erosion operation using the large-sized structural element, the lost large-sized stars are retrieved. Due to the size difference between B_{io} and B_m , the non-uniform background area becomes smaller after the two operations. Therefore, continue to use B_m to perform the dilation operation on $F_2(x, y)$ to restore the size of the non-uniform background, and obtain the processed image $F_3(x, y)$, which is expressed as

follows:

$$F_3(x, y) = F_2(x, y) \oplus B_m \quad (8)$$

Finally, due to the dilation operation using the structural element B_{io} , the pixels in the non-star area have uncertainty, which may lead to negative values when removing the non-uniform background. Therefore, the expression for correcting the stray light non-uniform background star map $F_4(x, y)$ is as follows:

$$F_4(x, y) = f(x, y) - \min[F_3(x, y), f(x, y)] \quad (9)$$

The dilation operation uses structural elements that can effectively suppress background noise, while the erosion operation uses structural elements that can retain the details of the star. This combination can more accurately separate the star and the background, and improve the robustness and accuracy of the method.

After obtaining the corrected star map, in order to verify the denoising effect and the retention effect of the star, the image segmentation method is used to segment and extract the star. The purpose of image segmentation: firstly, to separate the star points from the background; secondly, to separate the star points from each other. Compared with the fixed threshold, the adaptive threshold will be flexibly selected to adapt to the image according to the different image pixels. Finally, the segmentation effect is good and the speed is fast. Threshold T is as follows:

$$T = \mu + k\sigma \quad (10)$$

Among them, μ is the average value of the input image, σ is the standard deviation of the input image, and k is the average grayscale of the separated non-uniform background image.

To separate the star points, the star point boundaries need to be determined to complete the clustering division of the star points. The quadruple concatenation criterion [14] is used for the segmentation between the star points. According to the above segmentation results, only one scan through the image gray matrix is needed to complete the two major segmentation tasks of the entire star map, which can evaluate the denoising effect and fulfill the real-time demands of image processing in real application scenarios.

4 Experimental Results and Analysis

4.1 Experimental Data

To verify the proposed method's effectiveness in this paper, a semi-physical simulation star map experiment was carried out. The star map was generated by a star map simulator device, and the camera directly captured and imaged the star map for subsequent data processing. The experimental platform and star map are shown in Fig. 3.

All experiments were run on a personal computer with a 2.60 GHz Intel Core i5-13500H processor and 16 GB of RAM, and the test software used was Matlab2018.

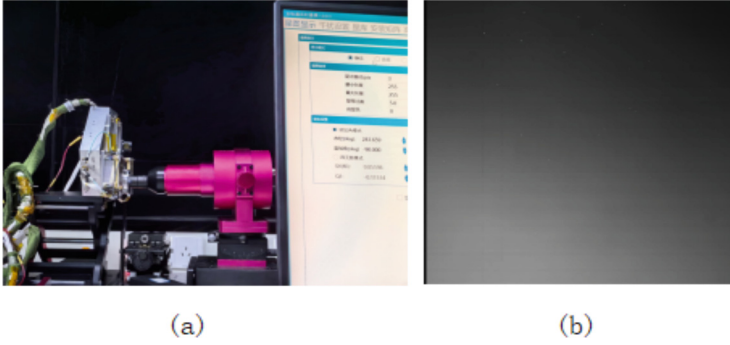


Fig. 3. Experimental platform and imaging star map

4.2 Evaluation Indicators

The evaluation indicators used in the experiment are peak signal-to-noise ratio (PSNR) and background suppression factor (BSF), which are specifically defined as.

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (11)$$

$$R_{BSF} = \frac{\sigma_{out}}{\sigma_{in}} \quad (12)$$

Among them, MSE stands for mean square error, and σ_{in} , σ_{out} are the standard deviations of the background area before and after processing. A higher PSNR means that the difference between the processed image and the original image is smaller and the image quality is higher. BSF is used to evaluate the effect of the method in suppressing background noise. A lower BSF means that the background noise of the processed image is better suppressed.

In addition, Precision, Recall and F1-Score are used to evaluate the quality of target detection. Their definitions are as follows:

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (15)$$

TP (True Positive): The actual number of stars that are correctly detected as star.

FP (False Positive): The number of non-stars that are incorrectly detected as star.

FN (False Negative): The actual number of stars that are falsely not detected as a star.

where *Precision* is used to measure the accuracy of the detection method. A higher precision means a lower false positive rate, that is, most of the points detected as star are

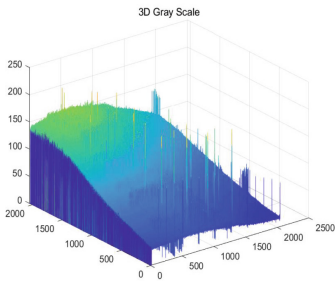
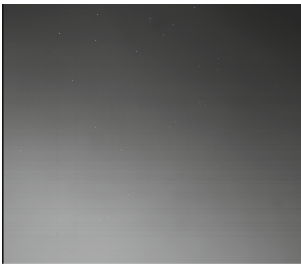
indeed star. *Recall* is used to measure the comprehensiveness of the detection method. A higher recall rate means a lower missed detection rate, that is, most of the actual star are detected. The *F1 – Score* is used to comprehensively measure the accuracy and comprehensiveness of the method, especially when there is a trade-off between precision and recall. A higher *F1 – Score* means that the method performs well in both accuracy and comprehensiveness. Time is used to measure the running time of various methods.

4.3 Experimental Results

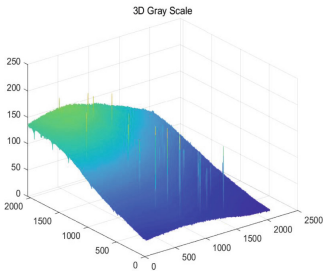
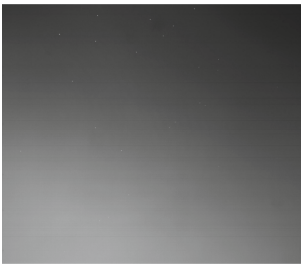
The proposed method is used for star map experimental verification, and Fig. 4 shows the detection results of the method. Figure 4 (a) shows the original image and its three-dimensional grayscale distribution, Fig. 4 (b) shows the smoothed image and its three-dimensional grayscale distribution, Fig. 4 (c) shows the non-uniform background image and its three-dimensional grayscale distribution, and Fig. 4 (d) shows the star image and its three-dimensional grayscale distribution after the non-uniform background is removed.

After obtaining the corrected star map, an adaptive threshold is applied to separate the star from the background, confirming the method's effectiveness in removing the non-uniform background. The four-connectivity criterion is used to determine the star boundary, complete the clustering of the star, and thus complete the segmentation between the stars. The following Fig. 5 shows the star extraction result, where 5 (a) is the original image and 5 (b) is the star extraction result after correcting the non-uniform background.

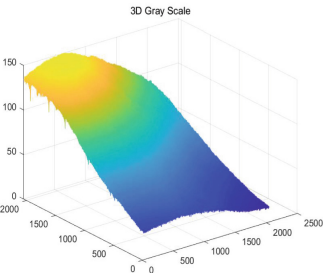
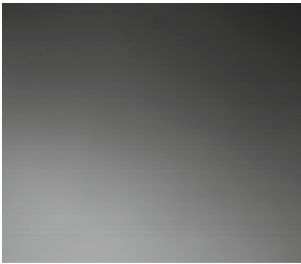
To further confirm the effectiveness of the proposed method, it is compared with the maximum mean filter based on spatial domain, the local contrast method (LCM) based on human vision, and the Top-hat method based on morphology. In order to avoid the influence of the size difference of the filter template on the non-uniform background correction performance of different methods, the window size of the maximum mean filter is 15, the window size of the LCM is 15, and the disk size of the Top-hat morphological structure element is 15. In order to more intuitively compare the performance of different methods, three groups of relatively typical simulation images are selected to demonstrate the non-uniform background correction effect of different methods. As shown in Fig. 6, the three-dimensional grayscale distribution of the uniform star map after correction by different methods, where (a) is the simulated original star map, (b) is the three-dimensional grayscale distribution of the original star map, (c) is the maximum mean filter method, (d) is the local contrast method, (e) is the Top-hat transformation method, and (f) is the non-uniform background correction method based on multi-scale grayscale morphology proposed in this paper. As can be seen from Fig. 7, different methods have a certain suppression effect on non-uniform background. The star structure of the maximum mean filtering method has changed, and the grayscale distribution of the star tends to be unified, which is not conducive to the subsequent centroid positioning of the star. Although the local contrast method and the Top-hat method have a certain ability to suppress the background, there are still some background residues, which is not conducive to the threshold selection in the subsequent star segmentation. The proposed method has a stronger ability to suppress the background. The background grayscale is basically smooth except for the star position, which minimizes the



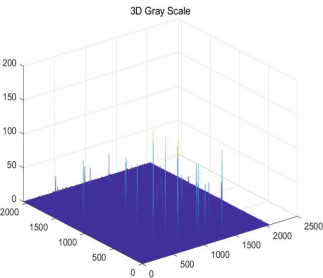
(a)



(b)



(c)



(d)

Fig. 4. Test results

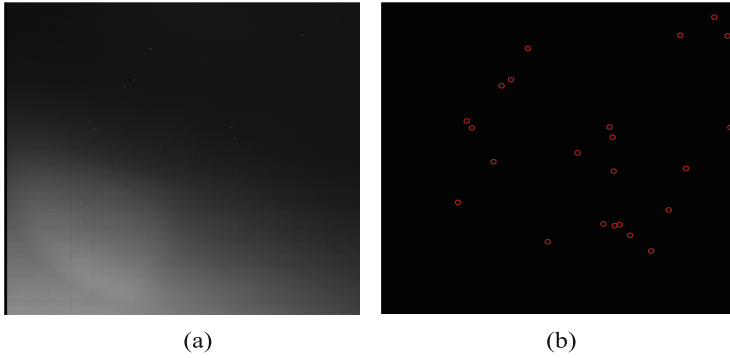


Fig. 5. Star extraction results

interference to the subsequent processing and has a larger threshold selection space. Figure 7 shows the three-dimensional grayscale distribution of the non-uniform background image eliminated after correction by different methods. The figure shows that the background removed by the maximum mean filtering method contains some star, and the removal effect is not good. The local contrast method has a better removal effect. The Top-hat method has a good removal effect on some star maps, and there are some stars in the background removed from some star maps. The proposed method can effectively correct the stray light and non-uniform background in star images with different signal-to-noise ratios, and has extremely high correction accuracy.

Furthermore, with the aim of analyzing the background suppression ability, star retention rate and running time of different methods, the data of P, R, F1, PSNR, BSF and Time of the four methods were statistically analyzed. The results are shown in Table 1, and the best result is indicated in bold.

Table 1 indicates that although the maximum mean filter method has slightly higher star extraction accuracy, the star recall rate and F1-Score are significantly lower than the proposed method. The star extraction accuracy, star recall rate and F1-Score of LCM are all lower than the proposed method. Although the recall rate of the Top-hat method is similar to the proposed method, the precision and F1-Score are lower than the proposed method. The F1-Score of the proposed method is optimal compared to other methods, indicating that the proposed method balances accuracy and comprehensiveness. In terms of method speed, the computational complexity of maximum mean filtering is high because the maximum and mean values of the local area need to be calculated at each pixel position. For large-size images and larger filter windows, the calculation time will increase significantly. The local contrast method improves the performance of the method by designing a complex contrast calculation window, but the complex contrast traversal calculation also greatly increases the method running time. The Top-hat method greatly saves method time due to its simple method structure, linear time complexity, efficient implementation, and does not involve global optimization. However, it uses the same structural elements for dilation and corrosion operations, ignoring the contrasting details between the target area and the surrounding background area, and the Top-hat method is sensitive to the shape of the target. If the shape of the target does not match the shape of

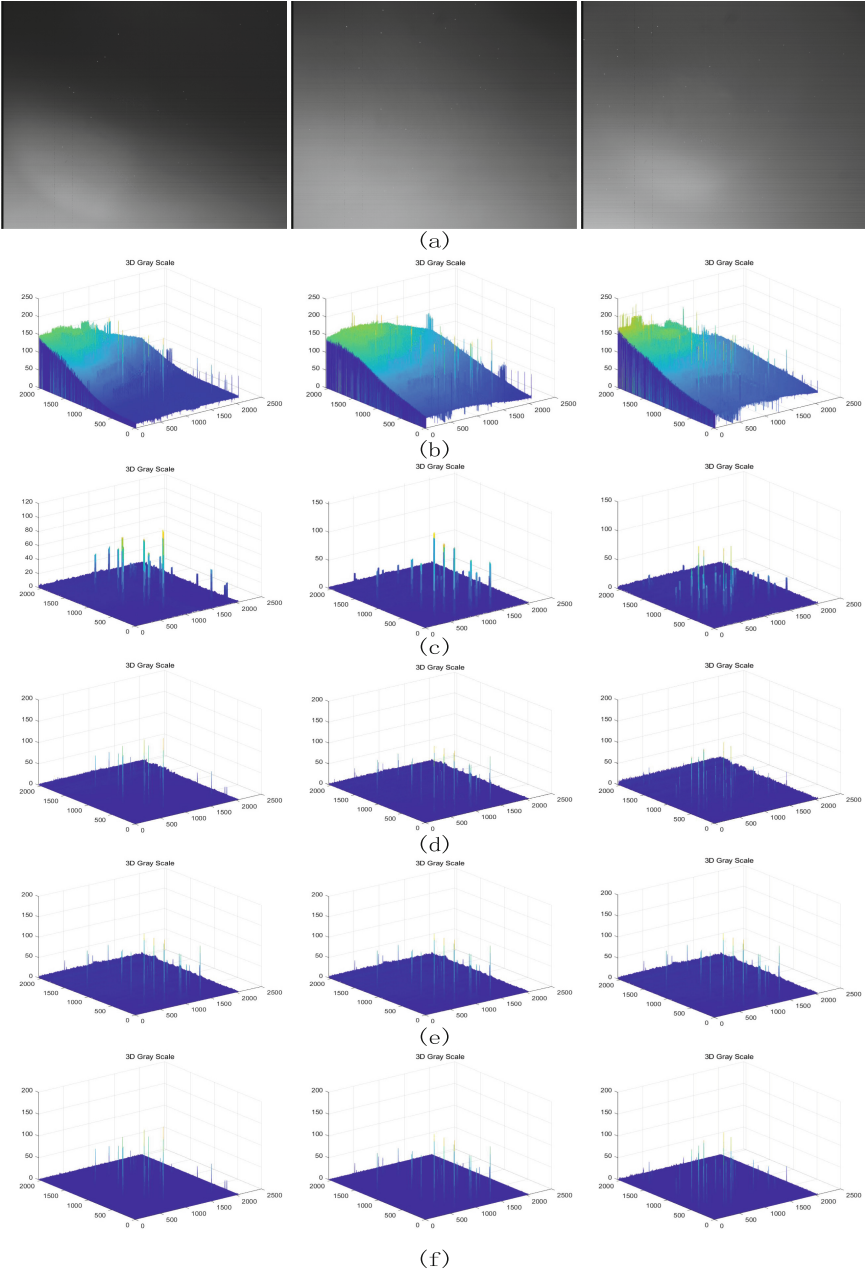


Fig. 6. Grayscale diagram of non-uniform background correction results of different methods

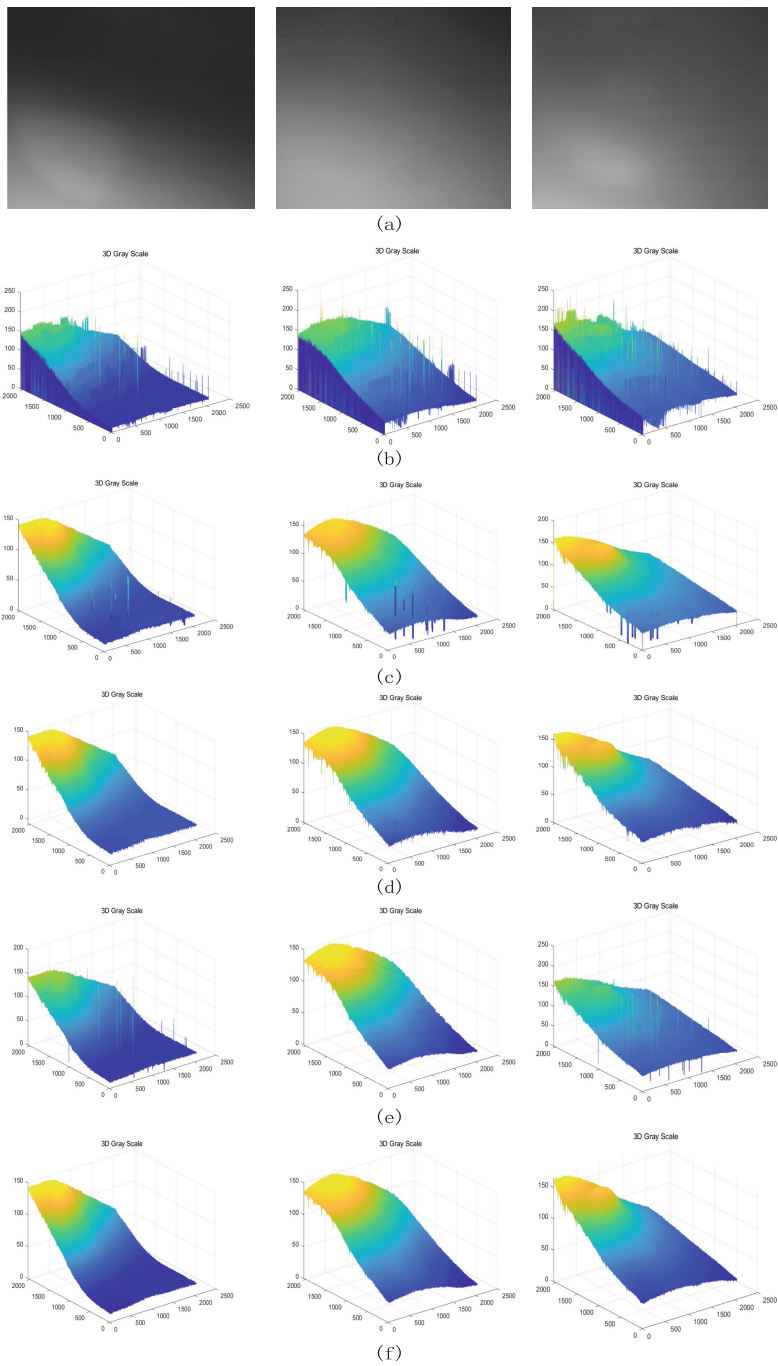


Fig. 7. Schematic diagram of non-uniform background grayscale eliminated by different methods

Table1. Comparison of detection results of different methods.

	Max-mean filter	LCM	Top-hat	Proposed
P	0.992	0.959	0.959	0.990
R	0.943	0.973	0.974	0.973
F1	0.966	0.965	0.966	0.981
PSNR	26.671	39.542	38.118	39.036
BSF	0.028	0.011	0.024	5.729e-4
Time	20.796	30.986	0.095	0.126

the structural element, the target detection effect may be unsatisfactory. The proposed method designs separate structural units for the morphological operations of dilation and corrosion, makes full use of the contrast between the target area and the surrounding background area, enhances the background suppression capability without increasing excessive time consumption, and can be used in actual non-uniform background star map prediction. The processing can effectively suppress the background while retaining the star, which has high practicality.

5 Conclusion

Star map preprocessing can effectively suppress stray light and noise interference, enhance target characteristics, and improve star center of mass positioning accuracy and target detection accuracy. Traditional methods are difficult to effectively distinguish the target from the background when faced with stray light background, resulting in insufficient accuracy of star extraction and reduced reliability of target recognition. This paper proposes a multi-scale grayscale morphology method, which effectively improves the accuracy and robustness of star extraction by introducing dual structural elements and adaptive threshold technology. The experimental outcomes demonstrate that this technique substantially improve the accuracy of star extraction under stray light background, and has important strategic significance for future space resource competition and space target monitoring technology.

References

1. Tang, Z., Huang, X., Cai, Z.: Development status and enlightenment of foreign space-based space surveillance systems. *Aerosp. Electron. Warf* **31**, 24–26 (2015)

2. Bai, X., Zhou, F.: Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recogn.* **43**(6), 2145–2156 (2010)

3. Din, P., Zhang, Y., Liu, R.: Infrared small target detection based on adaptive Canny algorithm and morphology. *Chinese J. Liquid Crystals and Displays* **31**(8), 793–800 (2016)

4. Li, J., Guo, P., Wang, C.: Application of shearlet transform in the star extraction. *Chinese Optics* **3**(3), 386–393 (2015)

5. Chen, C.L.P., Li, H., Wei, Y., Xia, T., Tang, Y.: A local contrast method for small infrared target detection. *IEEE Transactions on Geoscience and Remote Sensing* **52**(1), 574–581 (2013)
6. Han, J., Liang, K., Zhou, B., Zhu, X., Zhao, J., Zhao, L.: Infrared small target detection utilizing the multiscale relative local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **15**(4), 612–616 (2018)
7. Niu, H., Yan, C., Wang, Y., Guan, H., Wang, C., Shao, J.: Star map spatial target detection method based on fast local contrast and target features. *Chinese Journal of Liquid Crystals and Displays* (2024)
8. He, Y., Wang, H., Feng, L., You, S.: A novel method of eliminating stray light interference for star sensor. *IEEE Sens. J.* **20**(15), 8586–8595 (2020)
9. Liu, D., Wang, X., Li, Y., Xu, Z., Wang, J., Mao, Z.: Space target detection in optical image sequences for wide-field surveillance. *Int. J. Remote Sens.* **41**, 1–12 (2020)
10. Sun, T., Xing, F., Bao, J., Ji, S., Li, J.: Suppression of stray light based on energy information mining. *Appl. Opt.* **57**, 9239–9245 (2018)
11. Serra, J.: *Image Analysis and Mathematical Morphology*. Academic Press, New York (1982)
12. Haralick, R.M., Sternberg, S.R., Zhuang, X.: Image analysis using mathematical morphology. *IEEE Trans. Pattern Anal. Mach. Intell.* **9**(4), 532–550 (1987)
13. Wei, M.-S., Xing, F., You, Z.: A real-time detection and positioning method for small and weak targets using a 1D morphology-based approach in 2D images. *Light, Sci. Appl.* **7**(5), 18006 (2018)
14. Liu, X., Li, B., Su, Q.: New exact labeling algorithm of connected regions in binary images. *Comput. Eng. Appl.. Eng. Appl.* **42**(22), 76–78 (2007)



Lightweight One-Stage Ship Detection Based on SAR Images for Embedded Platform

Chenxin Lu¹, Ji Shen¹, Shuhui Zhou², Yang Zheng^{1(✉)}, and Jie Yuan¹

¹ Shanghai Aerospace Electronic Technology Institute, Shanghai 201109, China
zhengyang_058@126.com

² Sino-European School of Technology of Shanghai University (UTSEUS),
Pudong 200444, China

Abstract. Ship detection by satellite-carried synthetic aperture radar (SAR) imagery is currently a hot topic in the development of space remote sensing technology. Due to the limitation of hardware resources and the requirement of high-speed processing, the current deep learning model has the problems of low detection rate and high false alarm density in ship detection in complex scenes. To address the issues, we propose a lightweight method to detect ships in SAR images for embedded platforms based on the upgraded YOLOv7-tiny, called Lightweight Ship Detection-YOLO (LSD-YOLO). Firstly, the MobileNetV3Block module is introduced into the backbone network, and incorporating a ULSAM attention mechanism to enhance the capacity for the method to capture important features of ships in complex scenarios. Secondly, we utilize SPPFCSPC module and replace certain standard convolutions with GsConv. The design not just reduces computational complexity but also significantly enhances the method's detection performance. Notably, the average detection accuracy on the SSDD dataset is 97.7%, which is a 1.9% improvement over the yolov7-tiny model, with the number of parameters and computation decreasing by 21.7% and 24.4%, respectively, and significantly outperforming the other compared models. Finally, on the Cambri-con Siyuan220, the proposed method has a detection speed of 153 frame/s, and achieves the accurate and real-time detection in SAR image ship detection task.

Keywords: SAR image · Ship detection · Embedded platform · Lightweight module · High-accuracy

1 Introduction

With the in-depth promotion of China's task of building a strong oceanic country and the spaceflight remote, the quick growth of sensing technology puts forward higher requirements for ocean monitoring capability, among which ship detection is an important part of ocean monitoring with certain military value. Synthetic Aperture Radar (SAR) has the working characteristics of all-day and all-weather earth observation [1]. The identification of vessels using SAR imagery is currently a trending area of study within the realm of remote sensing.

The SAR image detection is a very crucial part of the SAR image detection process. However, SAR image detection also has many problems [2] (1) False Alarm due to complex background on land (2) Missed detection due to side-by-side docking (3) False Alarm due to sea clutter (4) Missed detection of small targets. Conventional approaches to detecting targets in SAR imagery, including the common Constant False Alarm Rate (CFAR) target detection method, are susceptible to noise in the complex background [3] are susceptible to noise interference in complex backgrounds, and tend to have long processing time, poor generalization ability and low detection accuracy, making it difficult to satisfy the demands for both immediacy and precision.

Recently, deep learning methods have been widely used within the domain of SAR image processing. Deep learning-based target detection techniques can be primarily categorized into two types: one is a two-stage detection method based on candidate regions, such as Fast RCNN [4], Faster RCNN [5] Mask RCNN [6] etc., the whole target detection work needs to be completed in two stages, first locating the candidate region and then sending it to the classifier for classification. The other class is single-stage detection methods, such as RetinaNet [7], CenterNet [8] FCOS [9] Nanodet [10] YOLO series, etc. The two-stage methods are too slow to meet the on-board processing requirements, and most of the lightweight target detection models are single-stage detection methods. Nanodet is a mobile target detection model, which is a FCOS single-stage anchor-free target detection model with obvious advantages in detection rate [11]. Liu et al. optimized the design of Nanodet by combining the cavity convolution, and at the same time, removed the central branch at the output end to reduce the computation amount of the model. The YOLO series of detection models include YOLOv5 [12], YOLOv6 [13] YOLOv7 [14] and YOLOv8. YOLOv7 uses an efficient aggregated feature extraction network, the RepBlock module to reparametrize the model, and coarse-to-fine to guide the label assignment strategy. YOLOv7-Tiny is a lightweight version of the YOLOv7 model, with higher detection accuracy and rate, and is easier to deploy due to its similarity to the overall YOLOv5 framework. YOLOv7-Tiny is a lightweight YOLOv7 model with higher detection accuracy and rate. On the COCO dataset, the detection performance of YOLOv7-Tiny is better than that of YOLOv5n, YOLOv6n and YOLOv8n. To ensure effective ship target detection in SAR images on a satellite-based system, this paper chooses the YOLOv7-Tiny model, which combines accuracy and speed, as the basis for the study.

Many scholars have written on YOLOv5, YOLOv7, and YOLO7-Tiny corresponding improvements are proposed on the basis of the model. Lu Peng [15] et al. proposed to use ShuffleNetv2 model is used as the backbone part, while fusing Transformer structure and dual attention mechanism to address the challenging issue of detecting dense distribution of multi-scale ships, the model is more complex and not easy to deploy on hardware. Li Yonggang [16] et al. proposed to introduce deformable convolutional DCNv2 in the backbone network to enhance the target feature extraction ability and enhance the detection rate of ship targets in SAR images under complex scenes, but from the experimental results, the mAP indicator is less enhanced, but the detection rate is reduced more. Chen [17] et al. in YOLOv7 method based on the combination of hollow space pyramid pooling and attention mechanism, to reduce the loss of ship features, hollow convolution in the premise of not increasing the amount of computation

to increase the sensory field, to enhance the different scales of image ship target feature extraction ability, but hollow convolution computation there is a cache miss problem, you need to visit the memory many times to get the data, and the actual detection rate of hardware will drop a lot. Will decrease a lot. Guo Wei [18] et al. introduced HOI-ELAN, a high-efficiency layer aggregation network module with high-order interaction function, based on YOLOv7-Tiny method into the backbone network, which enhances the feature extraction capability of small targets in SAR images.

However, the recursive gated convolution module RGC is time-consuming [19] et al. use Re-I-block module to build the backbone network part, which is composed of ELAN module and ResNeXt module to enhance the feature lifting ability of the backbone network and enhance the accuracy of target detection in scenarios with complex backgrounds, but the multi-branching increases the amount of computation, and the efficiency of the hardware execution will be reduced. Zeng [20] et al. use DsSPPF to replace the original SPPCSPC module, which adopts smaller scale max-pooling and depth-separable convolution to replace the original large-scale max-pooling operation, which is used to aggregate scale-varied feature details, and the computational volume of the model increases greatly, but the improvement of detection performance is limited. Although the above improvements improve the detection performance to a certain extent, the modified modules are complex and not suitable for deployment on hardware. Therefore, it is necessary to combine the hardware characteristics to design the relevant modules in an efficient and lightweight way.

In summary, this paper designs a lightweight method, LSD-YOLO, on the basis of YOLOv7-Tiny. In the backbone network part, a depth-separable convolutional network MobileNetV3Block is used to construct the lightweight model, and only part of the feature extraction layer is replaced with the MobileNetV3Block module to improve the board-end detection speed. In the feature enhancement network part, the SPPFCSPC module is used for feature compression and parallel computation, which significantly decreases computational requirements and enhances processing velocity, making the ship detection algorithm better adapted to the actual application cases. At the same time, all the convolutions in the non-ELAN module are replaced with GSConv, and the experiment proves that this operation can lead to better efficiency in ship detection under complex working conditions. In the application of the attention mechanism, the existing attention mechanism uses 1×1 convolution with higher cost and MLP with more parameters, thus generating higher computational overhead and parameter overhead. This model replaces the SE attention module in MobileNetV3Block with the ULSAM module, and this operation further enhances the model's feature extraction capability without increasing the computational overhead. Experiments show that the algorithm improves some modules in the backbone network and feature enhancement network, and the algorithm can be successfully deployed to Cambrian Siyuan 220 with better actual processing performance, which meets the high-performance processing requirements on the star. The main contribution of this paper is that we propose a lightweight method to detect ships in SAR images for embedded platforms.

2 LSD-YOLO Target Detection Method

2.1 Overall Structure of the Model

This method is based on the YOLOv7-Tiny model for efficient lightweighting. The model is divided into three parts: the backbone network, the feature enhancement network and the output. The backbone network is optimized by using the lightweight module and combining the operator characteristics; the feature enhancement network is optimized by using the efficient feature fusion module and the efficient feature enhancement module, and the output is optimized by using the attention mechanism and efficient activation function are used to significantly improve feature acquisition capability of the network, and the overall structure of the model is shown in Fig. 1.

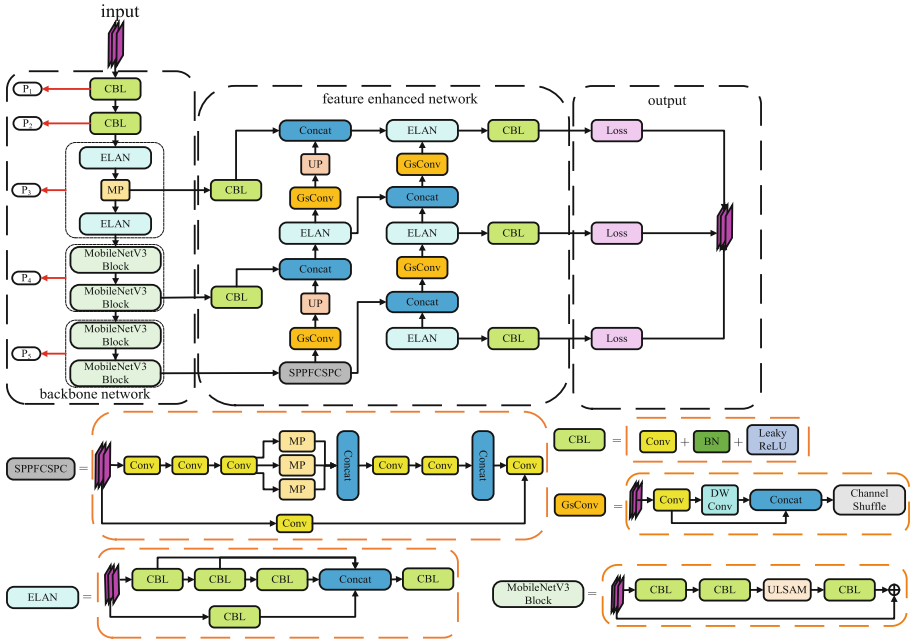


Fig. 1. Overall structure of the model.

2.2 Lightweight Backbone Network Design

Depthwise Separable Convolutional Convolution (DSC) can greatly reduce the amount of network computation and parameters, and is widely used in lightweight network design. Typical DSC networks are the MobileNet series and ShuffleNet series. The ShuffleNet series uses the Channel Split operation, which achieves higher detection accuracy and faster detection speed, but this operation is time-consuming and not friendly to most hardware.

MobileNetV3 is the latest version of this series, which combines the deep separable convolution of MobileNetV1 and the inverted residual structure of MobileNetV2, uses the network search method, and introduces the SE module and H-Swish SE module, H-Swish activation function, etc. MobileNetV3_small is a lightweight model suitable for embedded devices, and MobileNetV3_small outperforms ShuffleNetV2 in terms of model Top1 accuracy and inference time [21]. MobileNetV3_small is a lightweight model suitable for embedded devices.

Another type of lightweight module is the GhostNet series [22]. The GhostNet series is based on a set of intrinsic feature maps that are transformed by a series of linear transformations in order to achieve a low cost of ownership. GhostNet generate multiple feature maps that fully reveal intrinsic feature information.

The GhostNet family of models consists of multiple GhostBlockNeck modules, which significantly decreases the parameter count and computational effort of the models. However, from the test results of several SAR datasets, to attain the detection capabilities of the backbone network composed of MobileNetV3Block, it is necessary to stack more GhostBlockNeck modules, and the actual computation will increase a lot, and the real processing performance is not as good as that of MobileNetV3Block, so this paper adopts MobileNetV3Block to build the GhostNet series model. MobileNetV3Block to construct the backbone network of the lightweight model.

The MobileNetV3Block module uses 1×1 convolutional fusion channel information, which occupies a large amount of memory and requires high computing power for embedded devices. Although the embedded devices used in this paper can meet this requirement, the 1×1 convolutional fusion channel information will occupy a large amount of memory, which requires high computing power of the embedded devices.

The extensive use of convolution also reduces the processing speed of the board to some extent. In the experiment, we replaced all the feature extraction layers with MobileNetV3Block.modules, it is found that the board-end inference time is not improved, so the model in this paper only replaces p4, p5 with two MobileNetV3Block modules. As shown in Fig. 2, this design can enhance the board-end detection speed on the one hand, and on the other hand, due to the use of standard convolution in the early stage of the feature extraction phase, the feature extraction capability is improved, which to some degree enhances the model's detection precision.

2.3 Feature-Enhanced Network Segment Optimization

This model uses the SPPFCSPC module to replace the SPPCSPC module block, the SPPFCSPC module is an optimization of the SPPCSPC module with reference to the SPPF, which enhances the overall feature extraction capability of the model, enabling the model to have higher detection accuracy in complex scenes, and reduces the computational load and accelerates the speed of model inference while keeping the sense field unchanged. Compared with SPPCSPC module and SPPF module, SPPFCSPC module can significantly improve the model detection performance. In practical applications, ship detection methods need to process large-scale SAR image data sets, so efficient processing speed is very important. SPPFCSPC module greatly reduces computational load and accelerates the processing speed by means of effective feature compression and

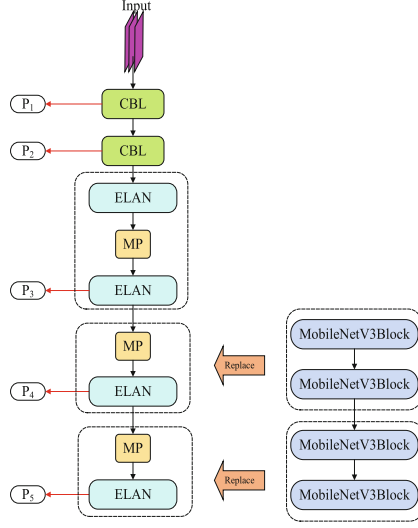


Fig. 2. Replacement of backbone.

parallel computation, so that the ship detection method can better adapt to the practical application scenarios. Figure 3 illustrates the architecture of the module.

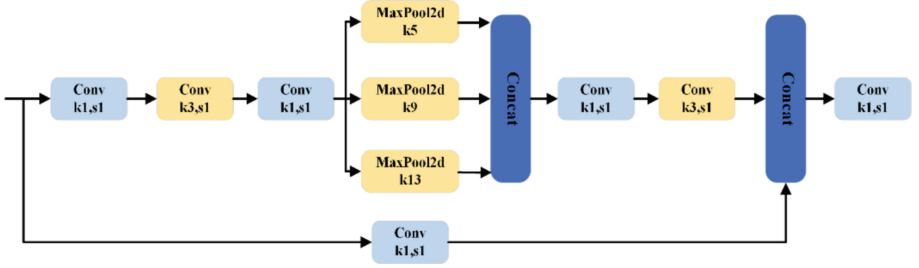


Fig. 3. SPPFCSPC module.

To address the problems of dense distribution accurately and the land scene that is too large and causes a false scene, both detection accuracy and detection speed need to be considered in the design of feature enhancement networks. GsConv is a lightweight convolutional module, which makes the result of the convolution kernel as close as possible to the standard convolutional product while reducing the amount of computation. The article [23] points out that if GsConv is used throughout the model, it will greatly deepen the model and reduce the inference speed of the model, while using GsConv only at Neck, where the dimension of the channel information is the largest and the dimension of the spatial information is the smallest, is a more preferable choice. To further enhance the functionality of the feature enhancement network part, the convolution in the non-ELAN module is replaced by GsConv, and the experiment proves that this

operation can enhance the overall detection performance to some degree, especially the performance of the detection of ships in complex working conditions.

2.4 Use of Attention Mechanisms

To enhance the feature extraction performance of the backbone network, a lightweight backbone add the attention module in the network part, the general attention module does not improve the model detection performance, this paper model selection Ultra-Lightweight Subspace Attention Module ULSAM [24] (Ultra- Lightweight Subspace Attention Module), ULSAM can be Inferring a separate attention map for each feature subspace and reducing spatial and channel redundancy in compact CNN feature maps enables efficient learning of cross-channel information as well as multi-scale and multi-frequency feature learning. ULSAM is end-to-end trainable, and experiments validate the effectiveness of ULSAM on the MobileNet family. The ULSAM module uses a subspace attention mechanism to improve the efficiency of utilizing the standard The ULSAM module uses the subspace attention mechanism to improve the efficiency of the standard convolution, which divides the extracted features into g groups, and spatially re-calibrates the sub-features of each group to connect the g groups together. Figure 4 depicts the configuration of ULSAM.

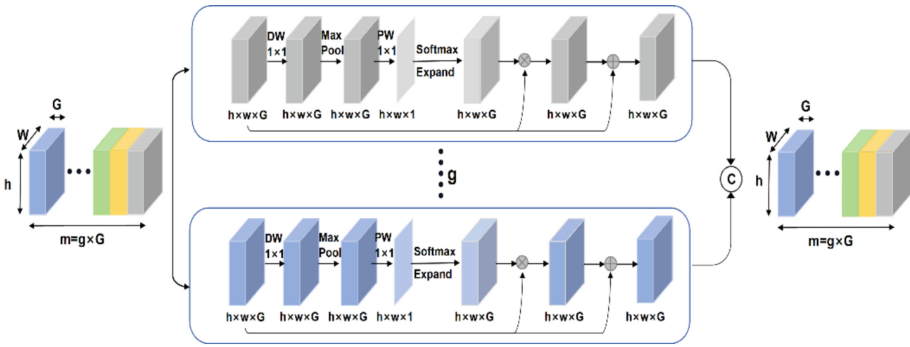


Fig. 4. ULSAM module.

Existing attention mechanisms such as non-local, A2Net, SENet, BAM, CBAM, incur high computational and parameter overheads due to the use of costly 1×1 convolutions and parameter-heavy MLPs. On the contrary, ULSAM uses only one 1×1 filter and exploits the linear relationship between the feature maps and avoids the use of MLP, with a computational overhead of the number of parameters is smaller, which is superior in lightweight network deployment and hardware implementation. This model replaces the SE attention module in MobileNetV3Block with ULSAM module, and this operation further enhances the feature extraction capability of the model without increasing the computation.

3 Implementation and Evaluation

Introduction to Data Sets and Evaluation Indicators

The experimental dataset is the SSDD public dataset [25]. The data set used for the experiments is the SSDD public dataset. 1160 images containing 2456 ships, with the maximum size of the ships being 211×298 pixels with a minimum size of 7×7 pixels. At the same time, the dataset contains different imaging modes and polarization modes, different sizes and shapes of ship targets, and covers multiple types of scenes. During the experiment, the training dataset and validation dataset are divided according to 7:3 ratio. The model evaluation metrics are detection precision (P), recall (R), and average mean average precision (mAP), the number of model parameters (M) and the amount of computation (GFLOPs).

3.1 Experimental Platform and Model Training Setup

The hardware platform of this experiment PC is intel®Core™ i7-97. CPU@3.60 GHz $\times$ 8, 128G RAM computer, the graphics card is RTX 4090Ti. The software environment is the operating system Ubuntu 18.04, cuda version 12.0, pytorch version 1.7. The input image size is 640×640 , epoch is 300, learning rate is 0.01, and optimizer is SGD. The data enhancement part of the training phase includes operations such as random pan, rotate, flip, crop, random HSV enhancement, Mosaic, and Mixup.

3.2 Results

Comparison of Ablation Experiments.

So as to verify the validity of the model improvement points in this paper, YOLOv7-Tiny was utilized as the baseline model, and P, R, mAP, number of parameters, and computation amount were used as the evaluation indexes, and the ablation experiments were conducted by combining several modules. The experimental results are shown in Table 1.

Table 1. Improved point ablation experiments.

Mobv3	SPPFCSPC	GsConv	ULSAM	P	R	mAP	M	GFLOPs
$\times$	$\times$	$\times$	$\times$	0.941	0.912	0.958	5.728	13.0
$\checkmark$	$\times$	$\times$	$\times$	0.928	0.914	0.966	3.490	9.1
$\checkmark$	$\checkmark$	$\times$	$\times$	0.959	0.912	0.971	4.678	10.1
$\checkmark$	$\checkmark$	$\checkmark$	$\times$	0.948	0.927	0.973	4.482	9.9
$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	0.959	0.922	0.977	4.485	9.9

Through Table 1, replacing the ELAN module p4,p5 with MobileNetv3 module improves the mAP by 0.8%; replacing SPPF with SPPFCSPC improves the mAP by 0.5%; replacing convolution in non-ELAN in the feature-enhanced network part

with GsConv improves the mAP by 0.5%; replacing the SE Attention module in MobileNetV3Block is replaced by ULSAM module, which improves the mAP by 0.4%; after applying all the improvements, the mAP of this model is advanced by 1.9% compared with the original YOLOv7-Tiny, and the parameter and computation decreased by 21.7% and 24.4% respectively, proving the effectiveness of each improvement point of the model in this paper, and the detection performance was improved substantially.

Comparative Experiments of Different Lightweighting Models

In order to validate LSD-YOLO in complex scenarios, SAR image ship, in this paper, typical images of various complex scenes are selected for detection, and the mainstream lightweight models Nanodet, YOLOv5n, YOLOv5s, YOLOv7- Tiny, YOLOv8n, and LSD-YOLO are compared in this paper, and the results are shown in Table 2.

Table 2. Comparison experiment of each detection method.

Model	M	GFLOPs	P	R	mAP	Layers
Nanodet	0.75	2.88	0.901	0.842	0.909	266
YOLOv5n	1.50	4.0	0.903	0.831	0.909	220
YOLOv5s	5.99	14.9	0.950	0.878	0.944	220
YOLOv7-Tiny	5.72	13.1	0.941	0.912	0.958	200
YOLOv8n	3.01	8.9	0.944	0.919	0.960	225
LSD-YOLO	4.48	9.9	0.967	0.939	0.977	228

In summary, it is apparent that the model LSD-YOLO in this paper has a significant effect on the mAP indicator. It is significantly higher than Nanodet, YOLOv5n, YOLOv5s, YOLOv7-Tiny and YOLOv8n. Although the number of parameters and computation amount increase a little bit compared with Nanodet and YOLOv5n, the detection performance is improved. Compared with the latest YOLOv8n, the model in this paper also has some advantages, although the number of parameters and computation amount indicate a tiny increment, but the detection performance is improved, which verifies the high performance of target detection of the proposed lightweight model in this paper.

Verification of Test Results

The actual detection results of the comparison model in Sect. 3.3.2 will be compared to the working conditions such as missed detections by docking and side-by-side arrangement, misdetections by complex land background and sea clutter, and missed detections by small ships. The display is shown in Fig. 5.

Analyzing the detection outcomes depicted in Fig. 5, it can be concluded that compared with other lightweight models, this model can better solve the detection problems of docking misdetection, complex background misdetection, scale difference leakage detection, side-by-side leakage detection, sea clutter misdetection, and large scene misdetection leakage detection and other complex working conditions, but there is still a certain degree of misdetection of highly similar non-ship objects, which needs to be researched in the future.

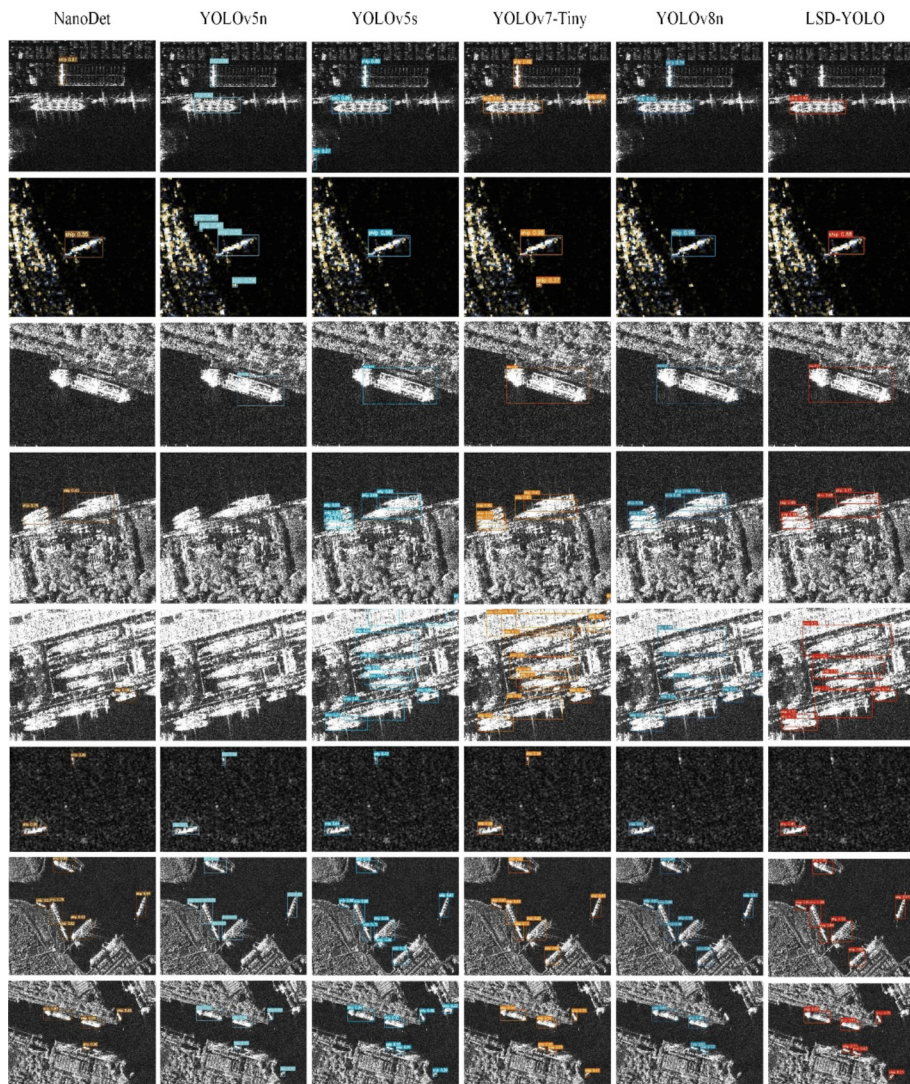


Fig. 5. Comparison of detection results of each method.

3.3 Hardware Implementation

The Cambrian Siyuan 220 is an intelligent SOC module that supports edge module adopts the new MLUv02 architecture. The module adopts the new MLUv02 architecture, the CPU is ARM A55 $\times$ 4, the main frequency is 1.5 GHz, the AI arithmetic power is 16TOPS, and the power consumption is less than 15W, which can satisfy the needs of the lightweight model for the detection of ships.

Offline Model Generation

The model LSD- needs to be generated in the format of cambrian before deploying model to Cambrian Siyuan 220. The practiced model weights are converted into offline models that can be run in Siyuan 220 through the steps of uncompressed saving, quantization, fusion inference and cross-compilation, which are carried out on the official development container. For the offline model with quantization mode INT8, the threshold error of the output frame is large, so we need to set the quantization mode to INT16.Considering the detection performance after migration, the model input needs to be set to half16, the batch is set to 4, and the core is set to 4.

Application Programming

Model calculations based on the CNRT interface, written in Bang c language Runner application. Based on the Pipeline model, the three parts of data pre-processor, model calculation Runner and output result post-processor, are multi-threaded, and the input nodes are set according to the data characteristics and the output node vectors are parsed.

Model Performance Testing

The offline models are deployed on the Cambrian Siyuan 220 and the performance tests are completed, and it is proved that the computational resource utilization is highest when the model batch size is 4, and the number of Preprocessor, Runner and Postprocessor threads are 2,4, and 1. Similarly, the performance of each model is also tested on the SSDD validation set, the input image size is 640×640 , and the confidence threshold of offline model inference is set to 0.20, and the performance of each model on the Cambrian Siyuan 220 is shown in Table 3.

Table 3. Comparison experiment of each detection method on Cambrian Siyuan 220.

model	P	R	mAP	FPS
Yolov7-Tiny	0.940	0.915	0.956	133
LSD-YOLO	0.971	0.938	0.980	153

As can be seen in Table 3, in terms of detection rate The model detection time for images of 640×640 is 6.55 ms, while the original YOLOv7-Tiny is 7.53ms. The new model reduces the inference time by 1ms per image, and the FPS improves by 20, which is an obvious improvement in the detection rate. In terms of detection accuracy, LSD-YOLO continues to maintain high detection accuracy on Cambrian Siyuan 220, while the original YOLOv7-Tiny has a slight decrease in detection accuracy, which verifies the high performance of the model in this paper for ship detection on embedded platforms. The frame rate of LSD-YOLO reaches 153 frames/s, which can satisfy the needs of SAR image real-time processing on-rail 1:1.

4 Conclusion

In order to address the limited computing power and memory resources of the hardware platforms on the star. To address the problem of low ship detection performance in complex SAR image scenes, a lightweight embedded-oriented SAR image ship detection method, LSD-YOLO, is proposed. The method is designed on the basis of YOLOv7-Tiny with a lightweight design, and uses ELAN module and MobileNetv3 module to construct a lightweight backbone network, and introduces the ULSAM attention module. to build a lightweight backbone network, while introducing the ULSAM attention module. Considering the efficient lightweight feature of GsConv, the standard convolution in the non-ELAN module is replaced by GsConv in the feature enhancement network, and the SPPFCSPC module is used to enhance the performance of multi-scale feature fusion. On the SSDD dataset, it has been demonstrated that the detection performance of the LSD-YOLO method is significantly reduced in computational load and the number of parameters, and the detection performance is improved a notable degree, which is better than other detection methods of the same type. Then, the method is deployed to Cambrian Siyuan 220, and based on the Pipeline, the three parts of data preprocessing, model calculation and output result post-processing are operated in multiple threads to further improve the performance of the model on the hardware side, and the frame rate of the 640×640 image processing is 153 frames per second, which meets the real-time processing rate requirement on the star. Experiments show that this model can better solve most of the complex working conditions of the ship detection problems, but there is still a certain amount of misdetection in the highly similar non-ship objects, which needs to followed up further research on this issue.

References

1. Moreira, A., Prats-Iraola, P., Younis, M., et al.: A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine* **1**(1), 6–43 (2013)
2. Zhang, D.D., Wang, C.P., Fu, Q.: Ship target detection in SAR image based on feature-enhanced network. *Systems Engineering and Electronics* **45**(4), 1032–1039 (2023)
3. Wackerman, C.C., Friedman, K.S., Pichel, W.G., et al.: Automatic detection of ships in RADARSAT-1 SAR imagery. *Can. J. Remote. Sens.* **27**(5), 568–577 (2001)
4. Girshick, R., Donahue, J., Darrell, T., et al.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, USA, June, pp. 580–587 (2014)
5. Ren, S.Q., He, K.M., Girshick, R., et al.: Faster RCNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2017)
6. Nie, X., Duan, M.Y., Ding, H.X., et al.: Attention mask R-CNN for ship detection and segmentation from remote sensing images. *IEEE Access* **8**, 9325–9334 (2020)
7. Lin, T.Y., Goyal, P., Girshick, R., et al.: Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, pp. 2999–3007 (2017)
8. Duan, K., Bai, S., Xie, L., et al.: CenterNet: Keypoint triplets for object detection. In: *IEEE International Conference on Computer Vision*, Seoul, Korea (South), pp. 6569–6578 (2019)
9. Tian, Z., Shen, C., Chen, H., et al.: FCOS: fully convolutional one-stage object detection. In: *IEEE International Conference on Computer Vision*, Seoul, Korea (South), pp. 9627–9636 (2019)

10. Rangi, L.: Super Fast and Lightweight Anchor-Free Object Detection Model Real-Time on Mobile Devices. [2020–12–01] <https://github.com/RangiLyu/nanodet> (2020)
11. Liu, F.J., Li, Y.: SAR remote sensing image ship detection Method Nanodet based on visual saliency. *Journal of Radars* **10**(6), 885–894 (2021)
12. Zhou, F., Zhao, H., Nie, Z.: Safety helmet detection based on YOLOv5. In: *IEEE International Conference on Power, Electronics and Computer Applications*, Shenyang, pp. 6–11 (2021)
13. Li, C.Y., Li, L.L., Jiang, H.L., et al.: YOLOv6: a Single-Stage Object Detection Framework for Industrial Applications. [2022–09–07] <https://arxiv.org/abs/2209.02976.2022>
14. Wang, C.Y., Bochkovskiy, A., Liao, H.Y. M.: YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In: *IEEE Conference on Computer Vision and Pattern Recognition*, Vancouver, BC, Canada, pp. 7464–7475 (2023)
15. Lu, P., Cao, Y., Zou, G.L., et al.: Improved shufflenetv2_YOLOv5 lightweight SAR image ship object real-time detection. *Hydrographic Surveying and Charting* **43**(01), 58–62, 82 (2023)
16. Li, Y.G., Zhu, W.G., Huang, Q.N., et al.: Near-shore ship target detection with SAR images in complex background. *Systems Engineering and Electronics* **44**(10), 3096–3103 (2022)
17. Chen, Z., Liu, C., Filaretov, V., et al.: Multi-scale ship detection algorithm based on YOLOv7 for complex scene SAR images. *Remote Sensing* **15**(8), 2071 (2023)
18. Guo, W., Wang, J.D., Wang, X.Z., et al.: Ship detection in SAR images based on high-order feature aggregation and convolutional modulation. *Science of Surveying and Mapping* **48**(8), 81–93 (2023)
19. Jia, K., Niu, Q., Wang, L., et al.: A new efficient multi-object detection and size calculation for blended tobacco shreds using an improved YOLOv7 network and LWC algorithm. *Sensors* **23**(20), 8380 (2023)
20. Zeng, Y., Zhang, T., He, W., et al.: YOLOv7-UAV: an unmanned aerial vehicle image object detection algorithm based on improved YOLOv7. *Electronics* **12**(14), 3141 (2023)
21. Howard, A., Sandler, M., Chu, G., et al.: Searching for MobileNetV3. In: *IEEE International Conference on Computer Vision*, Seoul, Korea (South), pp. 1314–1324 (2019)
22. Han, K., Wang, Y., Tian, Q., et al.: GhostNet: More features from cheap operations. In: *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, pp. 1580–1589 (2020)
23. Li, H., Li, J., Wei, H., et al.: Slim-neck by GsConv: A better design paradigm of detector architectures for autonomous vehicles. In: *IEEE Conference on Computer Vision and Pattern Recognition* New Orleans, LA, USA, pp. 2206–2424 (2022)
24. Saini, R., Jha, N.K., Das, B., et al.: ULSAM: ultra-lightweight subspace attention module for compact convolutional neural networks. In: *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Snowmass Village, CO, USA, pp. 1616–1625 (2020)
25. Zhang, T.W., Zhang, X.L., Li, J.W., et al.: SAR ship detection dataset (SSDD): official release and comprehensive data analysis. *Remote Sensing* **13**(18), 3690 (2021)



Multi-Modal False Information Identification Using BERT and ResNet

Dong Lv, Yitao Li, Wenchen He^(✉), Zihan Wang, Jiawei Qin, and Liqiang Wang

National Computer Network Emergency Response Technical Team/Coordination Center of
China, Beijing 100029, China
1009156962@qq.com

Abstract. In the context of rapid advancements in information technology, social media platforms such as Weibo, WeChat, and Douyin have become crucial channels for accessing information and understanding current affairs. Despite the convenience these platforms offer, the issue of false information has become increasingly severe due to the lack of effective regulatory mechanisms and rapid response measures. Traditional manual methods for detecting false information are inefficient and resource-intensive, necessitating accelerated digital development to enhance digital societal infrastructure and improve the level of digital and intelligent public services and social governance. To address this, this paper proposes a false information identification method that combines BERT and ResNet models for efficient and automatic detection of false information on social media. This approach aims to enhance the efficiency of false information detection, curb its spread, and promote the construction of a digital society.

Keywords: False Information · Identification · BERT · ResNet

1 Introduction

With the advancement of Internet information technology and the growing demand for diverse methods of acquiring information, social media platforms such as WeChat, Weibo, and Douyin have gradually become crucial channels for people to obtain external information and stay informed about current affairs. In recent years, the number of social media users in China has continued to rise. According to the latest data released by the China Internet Network Information Center, by the end of 2021, the scale of social media users in China had reached 1.007 billion, marking a 2.6% increase from the previous period. About 70% of the country's total population regularly uses social media [1]. On these platforms, users act as both creators and disseminators of information, leading to the rapid and widespread dissemination of content. Especially in social emergencies, the volume of information spread can even reach millions within a short period.

The rapid development of social media has provided a highly interconnected communication platform that transcends spatial limitations, greatly facilitating the publishing, acquiring, and dissemination of information. However, this also means that users face

minimal constraints on their information behaviors on social media platforms. Simultaneously, the lack of effective regulatory mechanisms and rapid response measures on these platforms makes it difficult to ensure the quality of information, fostering the proliferation and development of false information. False information on social media manifests in various forms, such as “clickbait” headlines that attract attention and provoke user emotions while the article content is misleading, “trumping” that manipulates public opinion and interferes with the presentation of real information, and “fake news” that exaggerates or distorts facts related to current events to incite negative emotions among the public [2]. Once false information spreads widely, it can have extremely serious negative effects on individuals and society as a whole.

Currently, the proliferation of false information on social media is considered a significant global risk. It not only disrupts normal information exchange among users on social platforms but also poses severe threats to social and economic development and even national political security [3]. In today’s social media environment, where there is an overwhelming amount of data, manually detecting false information is resource-intensive and inefficient. Leveraging relevant technical means to automatically identify false information on social media can enhance the efficiency of detection, quickly curb the spread of false information, and improve the digital and intelligent governance of social media platforms, thus strengthening the construction of a digital society. Presently, there are still certain limitations in methods for identifying false information on social media, primarily manifesting in three aspects: (1) Multi-source content: There are numerous types of current network platforms, and false information on the same topic spreads across different platforms, displaying multi-source content. (2) Multi-modal content: A significant amount of false information generated by intelligent algorithms has evolved from pure text to multimedia forms, including both text and images, showing multi-modal content. (3) High availability of content: The Internet features cases where images synthesized through deep synthesis technology are spread as false information or alongside false information text, increasing the perceived authenticity of the content and making it more accessible. In response, this study integrates false information identification methods using BERT and ResNet models, building on existing domestic and international research results, to achieve efficient and automatic detection of false information on social media.

2 Related Works

Since there is no universally accepted definition of the term “rumor”, the problem of online rumor detection can be categorized into two main targets: one is to detect information whose authenticity has not yet been determined, meaning the key feature of a “rumor” is that it is “unverified”. The other is to identify rumors that are ultimately verified to be false. Many experts and scholars argue that rumors which are widely spread and proven false are more harmful. Therefore, “rumors” can be equated with incorrect and false information. At the same time, it is believed that information from official channels is accurate, whereas information differing from official sources can be considered a rumor, or false information. This paper falls into the second category, equating “rumors” with incorrect or false information. Therefore, the research goal of this paper is to identify and detect false information in social media content.

In recent years, researchers have increasingly focused on the problem of multi-modal false information detection. To learn multi-modal representations of posts, Singhal et al. used BERT to extract text vector representations and VGG19 to extract image vector representations, which were then concatenated to form a joint representation. Yang et al. not only extracted explicit features from text and images but also used convolutional neural networks to learn their latent features. Although the concatenation-based approach is straightforward, not all features are equally important for false information detection. This means that when determining whether a post is false information or true information, the weight of visual features can vary compared to text features. Jin et al. proposed a recurrent neural network with an attention mechanism to fuse the image and text features of posts for false information detection. To further learn the shared representation between text and images, Wang et al. introduced an event adversarial neural network. This network includes a multi-modal feature extractor to extract text and visual features from posts, which works with a false information detector to learn discriminative representations. The event discriminator removes event-specific features and retains shared features across events, thus learning event-invariant features that represent various topics and fields, which aids in false information detection for new events. Khattar et al. proposed a multi-modal variational autoencoder that learns shared text and image representations. After training, this model can discover correlations between different modalities in tweets, and it pairs the variational autoencoder with a classifier to detect false information. Zhou et al. proposed a similarity-aware model to examine the role of similarity between text and image information in detecting false information. Meel et al. adopted an integrated learning model to capture the relevance of multi-modal information, using a multi-level attention network to learn the writing style of false information and employing noise variance inconsistency and error level analysis techniques to assess image authenticity. Wei et al. proposed a two-stage false information detection model, introducing a knowledge distillation method to capture cross-modal feature correlation. In the first stage, the cross-modal knowledge distillation function served as a soft target to guide the training of the single-modal network. In the second stage, the parameters of the two trained networks were fixed, and the extracted features were input into the fusion model to further enhance the performance of multi-modal fake news detection.

3 Theoretical Methods

BERT is a pre-trained language model based on multi-layer bidirectional Transformers, proposed by Devlin et al. in October 2018. BERT leverages large-scale unlabeled data to extract semantic information from text and further enhances the depth of natural language processing models. Due to its outstanding performance, it has achieved state-of-the-art results in various natural language processing (NLP) tasks, such as reading comprehension, machine translation, and text classification, establishing its significant position in NLP research.

From an overall structural perspective, the BERT model is essentially composed of multiple layers of stacked Transformer encoders. Each Transformer encoder consists of a multi-head self-attention mechanism and two layers of feed-forward neural networks. The self-attention mechanism is the core component of the entire Transformer model

framework. It calculates the correlations (attentions) between the state at a given moment in the current text sequence and the states at other moments, thereby representing the state at the current moment. Its calculation formula is shown in the accompanying formula.

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

In the formula, Q represents query; K represents key; V represents value; d_k represents the dimension of word vector; softmax represents softmax function.

Essentially, Q, K, and V are parameter vectors obtained by multiplying the input word vector with three different weight matrices. When calculating the attention weight of a word, the Q vector of the current word is first dot-multiplied with the K vectors of other words in the text sequence, and the result is divided by $\sqrt{d_k}$ for scaling.

The softmax function is then applied for calculation. Scaling is used here because the softmax function is more sensitive to larger input values. This scaling helps prevent the problem of gradient vanishing due to excessively large dot product results and ensures the stability of model training. After obtaining the corresponding weights from the softmax output, these weights are applied to the V vector to derive the final self-attention result, which is then fed into the feedforward neural network for further training. The application of self-attention allows the current word to encode the entire text sequence. Additionally, since multiple self-attention mechanisms do not interfere with each other during the calculation and can be computed in parallel, this approach addresses the long-term dependency problem present in traditional models such as RNNs, significantly enhancing the model's computational speed (Fig. 1).

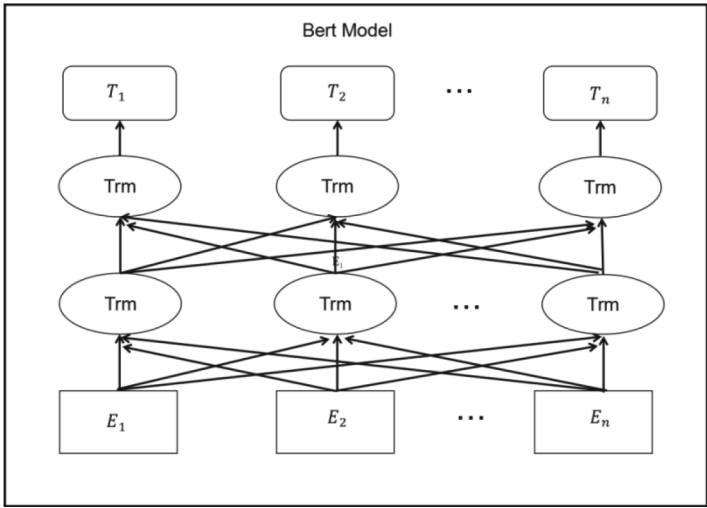


Fig. 1. BERT model

CNN (Convolutional Neural Network) is one of the most common and widely used network structures in deep learning, achieving impressive results in the field of image

recognition. CNN combines the feature extraction module with the classification module to enable automatic classification of target images. As illustrated in the figure, a standard CNN comprises multiple layers (Fig. 2).

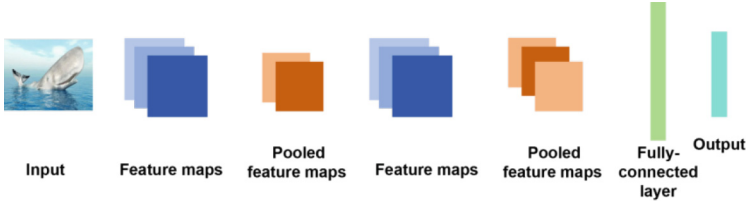


Fig. 2. CNN Model

In the early days, it was widely believed that the depth of a network was crucial for effective classification and recognition. To improve the model's performance on image classification tasks, researchers extended the depth of the network. As the network deepens, the features extracted become more complex and advanced, which can enhance model performance. However, increasing the number of layers can also lead to a decrease in classification effectiveness. This is due to issues such as gradient explosion, which make the model difficult to train once the network reaches a certain depth. To address the problem of deteriorating model performance with excessive network depth, Kaiming He and his team introduced the ResNet architecture in 2015. ResNet achieved remarkable results in various visual competitions that year, drawing considerable attention. Its innovative residual block structure significantly improved performance by allowing larger gradient values to be propagated back to the earlier layers of the network during the backpropagation process. Additionally, ResNet's flexible network design enables easy adjustments by modifying the number of stacked blocks and the number of channels within each block. This design mitigates concerns about gradient problems associated with deeper networks. With sufficient training data and gradual increases in network depth, better learning performance can be achieved. The basic structure of the residual network in ResNet is illustrated in the figure (Fig. 3).

The curve shortcut channel in the figure, the input x directly jumps two layers of network and completes the linear connection with $F(x)$. As can be seen from the figure, the working steps of the residual block can be clearly seen. x represents the input of the first layer, the network learning feature is recorded as $H(x)$, and the residual mapping is marked as $F(x)$. First, the input x is weighted for the first time, and then the activation operation is performed. In order not to compress the amplitude of the data, the activation function here uses relu , and finally the output is recorded as $F(x)$. Then, after the output after the first activation is weighted for the second time, before the second activation operation, the $H(x) = F(x) + x$ operation must be performed to obtain the final output of the module. It should be noted that after the second weighted processing operation and before the activation operation, a linear transformation operation is required, and this path becomes a shortcut connection. It is easier for the network to learn residual mapping than to learn direct mapping. Generally, in model training, the number of residual blocks reaches two or more to have a significant effect. Because the Resnet network does have

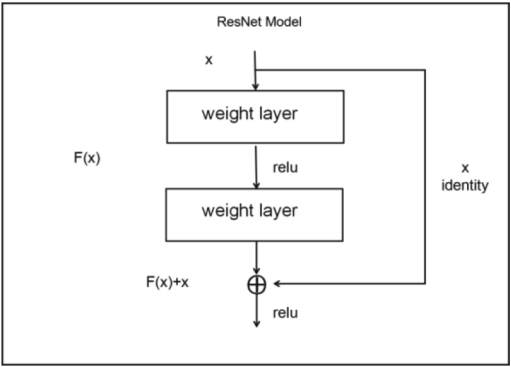


Fig. 3. ResNet Model

a good performance in dealing with the problem of poor network performance caused by deepening the number of layers. Therefore, Resnet is widely used in many fields[11].

4 Algorithm Design

The dataset was primarily sourced from public internet platforms. Specifically, for the Russian-Ukrainian conflict scenario, data was collected from nine platforms, namely Twitter, Weibo, Baidu, Wikipedia, WeChat, NetEase News, People’s Daily Online, CCTV.com, and Xinhuanet. A total of 47,982 text and image samples were gathered, consisting of 32,939 text samples and 15,043 image samples (Table 1).

Table 1. Data Introduction

Training set	27437, 0:19758,1:7679
Test Set	4420, 0:2980,1:1440
Validation set	1082, 0:374,1:708

The false information model is a binary classification model, and the data processing involves several key steps. Based on an official media list, a labeling function is constructed programmatically, and the collected data is labeled using a combination of manual and automated methods to create a sample set. Content published by the information publishing entities is labeled accordingly, with content from People’s Daily Online, CCTV.com, and Xinhuanet classified as positive samples. For the text data, we use the nltk library in Python to segment the text, remove stop words, and extract stems before applying the processed text to the model. For the image data, missing images in the sample dataset are uniformly filled with white images to ensure the model’s accuracy.

Bayesian classification is a statistical method that predicts the labels of test set data based on the statistical distribution of the training set samples. In simpler terms, a

Bayesian model classifies a sample according to the type that most closely resembles the known data. The term “Bayes” refers to Bayes’ theorem.

$$P(c|x) = \frac{p(x|c)p(c)}{p(x)}$$

where c is the category and x is the sample. Since $p(c)$, i.e. the proportion of each category in the data (prior distribution), is known; $p(x)$ is known (generally $1/N$); the solution of $p(x|c)$ is different in different Bayesian models, but all can be solved. In this way, the probability $p(c|x)$ (posterior distribution) that x belongs to each category can be obtained, and the argmax of the category is taken to obtain the classification prediction result. Through the Bayesian classification model, false information is classified.

The multi-modal classification model was constructed, and BERT was used for text feature extraction and Resnet was used for image extraction.

Under the trend of rich media in social media, false information has changed from pure text to a combination of text and pictures. Multi-modal content can verify the authenticity of information from different angles, and related research on false information identification based on multi-modal content mining has emerged. Deep fusion of content in different modalities is the key to using multi-modal content for false information detection. However, although related research has jointly represented vision and text, it has not fully mined semantic information when modeling features for visual and text information. The semantic features of vision mostly stay at the perceptual level, and the semantic features of text only stay at the entity level and contextual information. It is difficult to solve the semantic gap between the representations of different modal information, resulting in fusion only staying at a shallow level, and it is difficult to comprehensively and comprehensively judge the authenticity of information. In addition, existing research lacks the use of common sense for fact checking, making it difficult to identify counter-common sense false information. To address the above issues, the BERT and Resnet models are integrated to implement a multi-modal false information detection method that includes three modules: feature extraction, feature fusion, and false information discrimination, as shown in the figure (Fig. 4):

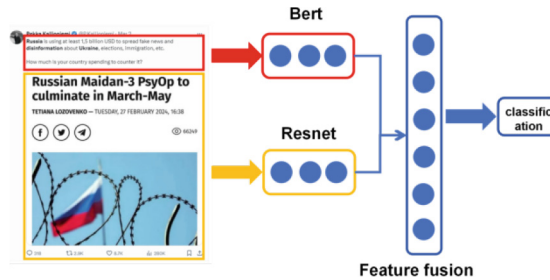


Fig. 4. Multi-modal false information detection method

Feature extraction is divided into feature extraction of text and feature extraction of images. For text feature extraction, firstly, entity information existing in the text is

mined based on named entity recognition technology, and entities are matched with concepts in the knowledge graph, and then the initial representation of text and concepts is obtained by introducing a pre-trained language model; for image feature extraction, a pre-trained object detection model is introduced to extract entities appearing in the image and match them with concepts in the knowledge graph, fully understand the potential semantic information behind the image, and obtain the initial representation of images and concepts through the self-supervised visual model MAE (Masked Auto-Encoders) and pre-trained language model respectively. Feature fusion aims to deeply fuse images and texts at the semantic level through entities and concepts in the knowledge graph. Firstly, a heterogeneous graph is constructed based on entities and concepts in text and images, and representation learning is performed through a heterogeneous graph neural network to obtain the final embedded representation of text and image.

False information discrimination is based on the image and text representations obtained in the first two steps. The image and text representations are spliced through the attention mechanism to obtain a unified expression of multi-modal features, and the distribution of classification labels is obtained based on this, that is, whether the output target text is false information and its confidence level. A total of three BERT models are used, namely BERT-base-chinese, chinese-BERT-wwm-ext and inirbt-h256 for model pre-training. In summary, for the classification of false information, logistic regression, Bayesian classification and multi-modal models are constructed respectively. The specific results are as follows (Table 2):

Table 2. Experimental results of machine learning models and multi-modal models

Classification Model	Accuracy
Logistic Regression	84.2%
Bayesian classification	81.8%
multi-modal Classification	82.3%

The logistic regression model demonstrates the highest accuracy, which may primarily be due to the sample data skew and the large amount of false sample data. The Bayesian classification model shows lower accuracy compared to logistic regression, likely due to the uneven distribution of the sample space. The multi-modal classification model exhibits the lowest accuracy among the three, which may be attributed to insufficient sample data. For text feature analysis, the model employs the deep learning model BERT. Specifically, it uses BERT-base-chinese, chinese-BERT-wwm-ext, and minirbt-h256t within the BERT framework. For image analysis, the model utilizes ResNet architectures, including ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152 (Fig. 5, 6 and 7).

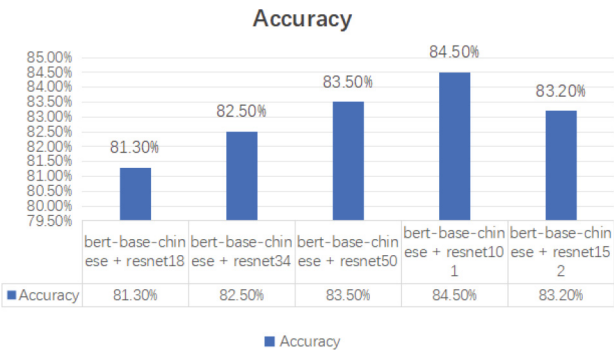


Fig. 5. BERT-base-chinese model and Resnet model experiments

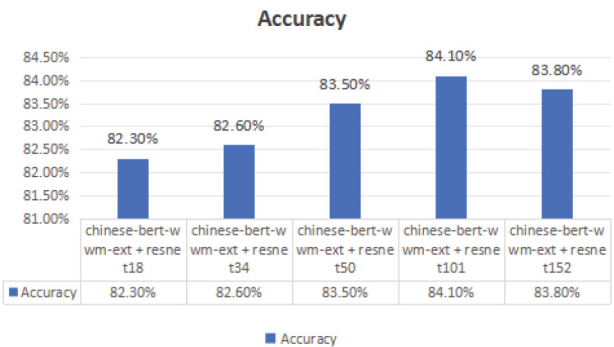


Fig. 6. Chinese-BERT-wwm-ext model and Resnet model experiments

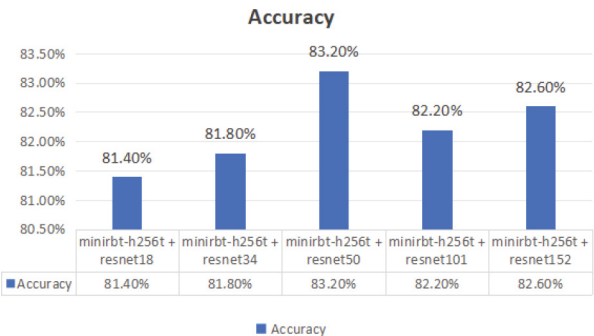


Fig. 7. Minirbt-h256t model and Resnet model experiments

5 Conclusion

In terms of data collection, gathering data from multiple platforms presents a significant challenge that greatly affects the construction of the dataset. During the model development phase, the strategy for text and image feature fusion required numerous attempts

and adjustments. The experimental process revealed that combining different BERT models, such as BERT-base-chinese, chinese-BERT-wwm-ext, and minirbt-h256t, with various ResNet versions including ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152, yielded varying results. In future work, the dataset will be rebuilt to further explore the model's capabilities and enhance its ability to detect and identify false information.

References

1. Nathan, F., Brent, C., Mark, L., et al.: Understanding landscape aesthetics using a novel viewshed assessment of social media locations within the troodos unesco global geopark, cyprus. *Frontiers in Environmental Science* **10** (2022)
2. 2022 Will See a Boom in Artificial Intelligence and Machine Learning in Media Entertainment; ABI Research's 2022 Trend Report identifies the one key metaverse market and technologies trend that will deliver in 2022—and the one that won't. M2 Presswire (2021)
3. Arega, L.K., Alasadi, K.M., Yaseen, J.A, et al.: Machine learning based detection of fake facebook profiles in afan oromo language. *Mathematical Modelling of Engineering Problems* **10**(6) (2023)
4. Bondielli, A., Marcelloni, F.: A survey on fake news and rumour detection techniques. *Inf. Sci.* **497**, 38–55 (2019)
5. Chen, W.L., Yeo, C.K., Lau, C.T., et al.: Behavior deviation: an anomaly detection view of rumor preemption. In: *Proceedings of the 7th Annual Information Technology, Electronics and Mobile Communication Conference*. IEEE, pp. 1–7 (2016)
6. Yang, F., Liu, Y., Yu, X.H., et al.: Automatic detection of rumor on sina weibo. In: *Proceedings of the 2012 ACM SIGKDD Workshop on Mining Data Semantics*, p. 13 (2012)
7. Wang, S.H., Terano, T.: Detecting rumor patterns in streaming social media. In: *Proceedings of the 2015 IEEE International Conference on Big Data*. IEEE, pp. 2709–2715 (2015)
8. Jain, S., Sharma, V., Kaushal, R.: Towards automated real-time detection of misinformation on twitter. In: *Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics*, pp. 2015–2020 (2016)
9. Singhal, S., Shah, R.R., Chakraborty, T., et al.: Spotfake: a multi-modal framework for fake news detection. In: *2019 IEEE fifth International Conference on Multimedia Big Data (Big MM)*. IEEE, pp. 39–47 (2019)



Arithmetic Awareness Based Distributed Routing Mechanism for Power Dispatch Networks

Zhifei Hao¹, Xiao Feng^{2(✉)}, Hang Zhang¹, Hanming Yan², and Guangwei Liu²

¹ State Grid Information & Telecommunication Co., Ltd., Beijing 102211, China

² State Grid Info-Telecom Great Power Science and Technology Co., Ltd., Xiamen 361008, China

wche@bupt.cn

Abstract. With the continuous development of power Internet of Things, the network has gradually become a bridge between intelligent power terminal equipment and intelligent computing communication. The data generated by power equipment will be processed through the network for intelligent computing. How to better balance the scheduling nodes and achieve the optimal scheduling routing mechanism under the limited arithmetic resources has become an urgent problem to be solved. To this end, this paper first builds an arithmetic-aware power dispatch network (AAPDN) architecture, which adopts a distributed routing control scheme to avoid the situation where large-scale power data are processed, parsed, stored and transmitted in a single centralised point, which leads to a single point of failure caused by an overburdened computational routing. On this basis, this paper further introduces an arithmetic routing notification and decision-making mechanism, which is used to share the arithmetic state information of the scheduling service nodes and formulate the optimal routing and scheduling decisions. Simulation experiments show that the mechanism proposed in this paper can maintain a high level of power scheduling service satisfaction while ensuring the arithmetic usage rate.

Keywords: Power Dispatch Network · Arithmetic Awareness · Distributed Routing

1 Introduction

In today's rapidly developing technological era, power IoT, as an important part of the smart grid, is gradually becoming a key bridge connecting intelligent power terminal equipment and realising intelligent computing [1, 2]. As the level of intelligence of power equipment continues to improve, the amount of data they generate is also increasing dramatically, which puts forward higher requirements for the data processing capability of the power system. How to achieve efficient data management and optimised scheduling strategies with limited arithmetic resources has become an important problem to be solved in the field of power IoT [3].

This paper aims to explore and solve this problem, and first proposes an arithmetic-aware power dispatch network (AAPDN) architecture. The AAPDN architecture adopts a distributed routing control scheme, which effectively avoids the computational overburden and the risk of a single point of failure caused by the centralised processing of large-scale power data. This distributed control scheme not only improves the reliability and stability of the system but also provides strong support for the efficient processing of power data.

Further, this paper introduces the arithmetic routing notification and decision-making mechanism, the core of this mechanism is to achieve the sharing of arithmetic state information of scheduling service nodes and to make the optimal routing and scheduling decisions based on this information. Through this mechanism, the arithmetic resources can be allocated more reasonably and the power scheduling process can be optimised, so that the satisfaction of the power scheduling service can be improved while ensuring the arithmetic usage rate.

In order to verify the effectiveness of the proposed mechanism, this paper tests the constructed model through simulation experiments. The experimental results show that the arithmetic-aware power scheduling network architecture and arithmetic routing decision-making mechanism proposed in this paper can effectively improve the scheduling efficiency and service quality of the power system in practical applications, and provide a strong technical support for the further development of the Internet of Things in power.

In summary, the research in this paper not only has important theoretical significance for the optimal scheduling of electric power Internet of Things, but also provides feasible solutions for the intelligent management and operation of the actual power system. With the continuous progress of technology and the depth of application, the research in this paper will contribute an important force to the construction and development of smart grid.

2 Related Works

2.1 Power Dispatch Network

As a key component of the power system, the technical development and innovation of the power dispatch network has been a hot research topic. The application of artificial intelligence techniques such as Deep Reinforcement Learning (DRL) in power dispatch optimisation provides new ideas to address the uncertainty and complexity of the power system [4]. The adoption of EPON-based data communication solutions for distribution network automation can improve network reliability and communication efficiency while reducing maintenance costs [5].

2.2 Arithmetic Awareness

Arithmetic awareness is a key technology in the development of smart grid, which involves real-time sensing, management and scheduling of computing resources in the grid to improve the operational efficiency and responsiveness of the grid. The arithmetic-aware network improves the utilisation of computing resources and network service

experience through the deep fusion of computing and networking as well as collaborative distributed computing resources [6]. The research focuses on arithmetic resource awareness, arithmetic task scheduling, and how to efficiently utilise the massive distributed computing resources at the edge of the network, which helps to realise the synergistic linkage between edge computing and cloud computing, and to provide the optimal user experience as well as the utilisation of computing and network resources [7].

2.3 Distributed Routing

Distributed energy access in smart grids poses technical challenges, such as grid stability issues and increased scheduling complexity. To cope with these challenges, researchers have proposed a series of solutions, including the development of intelligent scheduling strategies, the application of power quality control technologies, and the enhancement of communication and control infrastructure [8]. Multi-energy coordinated and optimal operation is an important direction in the study of energy internet. Distributed routing is generally achieved through distributed sending signalling, where multiple routing nodes have inter-arrival next-hop lists with each other, and nodes receive packets and forward them according to the next-hop list for guidance [9].

3 Arithmetic-Aware Power Dispatch Network Architecture

Aiming at the situation of soaring data scale and arithmetic demand of power dispatch network nowadays, the distributed route control scheme shifts the burden of core network computational routing to the edge, integrates with edge computing, sinks the arithmetic resources to the near-intelligent power end, and at the same time, retains the characteristics of stability and scalability, so this paper adopts the distributed route control scheme.

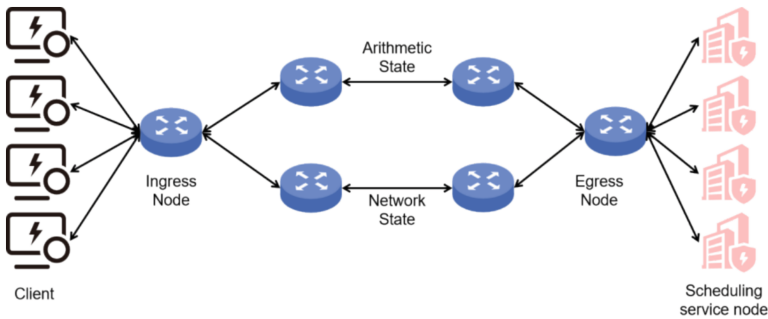


Fig. 1. Arithmetic-aware power dispatch network (AAPDN) architecture

As shown in Fig. 1, to describe the workflow of the arithmetic-aware power scheduling network more clearly, the power scheduling network nodes are classified into the following five types based on their different functions:

Distributed Service Request Node (Client): it is the initiator of the scheduling service request, and can use Anycast address to identify a specific service provided by multiple scheduling service nodes on the edge of the power scheduling network, i.e., Service ID (SID), which is used by the scheduling service request end to initiate the scheduling service request. It should be noted that the scheduling service will be assigned a unique SID based on the service type when it is registered, and in traditional power scheduling network scenarios, the Unified Platform Manager is generally responsible for distributing and updating the scheduling service registration information.

Ingress Node: it is an internal network element of the arithmetic-aware power scheduling network. The nearby ingress node faces the power scheduling terminal and guides the confirmation of the optimal scheduling service node based on the arithmetic routing table it generates.

Intermediate Node: It is only responsible for receiving and forwarding the arithmetic state information table of each scheduling service node in the arithmetic-aware power scheduling network as well as the detected network state information.

Egress Node: the export node faces the scheduling service nodes connected to it, and is responsible for obtaining and updating the real-time arithmetic state information of the scheduling service end connected to it, and notifying the whole network of this arithmetic state information, so as to provide support for the generation of the arithmetic routing table at the subsequent entrance node.

Scheduling Service Node: on the edge side of the power scheduling network, it is responsible for providing specific scheduling services for the scheduling service request end.

It is worth mentioning that the arithmetic routing entry node periodically generates the arithmetic routing table based on the arithmetic notification information and the network state information, and when the first packet of the scheduling request is received from the power scheduling end, the entry node will determine the target egress routing node based on the arithmetic routing table and the optimal service will be finally provided by the computing service node connected to it.

Arithmetic registration update means that the scheduling service node reports the quantitative arithmetic resources it has to its nearest router and periodically sends a POST request to update the arithmetic load information through RESTAPI, and thus the router notifies the arithmetic status information, and each router establishes and maintains the arithmetic status table after receiving the arithmetic information, and the router node serves as the originating point for the arithmetic notification. The routing decision process means that when a client node has a service scheduling request, the edge router measures the network state and fetches the arithmetic state table, takes into account the service attributes such as service priority, service type, etc., and based on the arithmetic state table and the network state table, obtains the routing information items and generates the forwarding table to guide the resource scheduling of the service.

The control plane of the arithmetic network mainly notifies the service state information of nodes, and the arithmetic-aware network needs the service state information table of the control plane to contain both computational performance and network performance data.

4 Arithmetic Routing Notification Mechanism

The key process of routing based on a distributed control scheme mainly includes the control plane and data plane process, in which the implementation of routing decision in the control plane requires that the arithmetic routing nodes notify each other about the available computational load corresponding to the support service information of relevant scheduling nodes, and each arithmetic routing node generates local routing table based on the arithmetic state information of scheduling service nodes on the edge side of the arithmetic-aware network and combines with the real-time network state. The core problem is how to achieve the notification of arithmetic routes under the distributed control scheme, and the control and selection of routes based on the distributed control scheme will generally be achieved by using appropriate routing protocols, so the problem lies in what kind of routing protocols should be selected to achieve the notification of arithmetic routes. Protocol to achieve the announcement of arithmetic routes.

4.1 Routing Protocol

Considering the introduction of arithmetic state information under the arithmetic-aware power scheduling network, a scalable routing protocol should be selected to carry the corresponding arithmetic state information to realise the arithmetic route announcement.

BGP is an autonomous system routing protocol running on TCP, with high scalability and better route announcement performance. BGP runs on routers divided into two types, IBGP and EBGP, which is different from IGP in that its purpose is not to discover and compute routes, but to achieve route filtering and obtain the best routes by setting up different route attributes. BGP refers to a variety of route policies to achieve the purpose of flexible route control through routing policies, while BGP only publishes updated routing information in real time to announce dynamic route changes in the network to reduce the bandwidth used. BGP references a variety of routing policies, which can be used to achieve the purpose of flexible routing control, and at the same time, BGP only publishes updated routing information, releases dynamic routing changes in the network in real time, and reduces the bandwidth used. The exchange of BGP routing information is accomplished through the interaction of BGP speakers, and autonomous domains (ASes) all designate a BGP spokesperson, who interacts with BGP speakers of other ASes to exchange routing information. The responsibility of the BGP spokesperson is to send BGP messages, and the BGP spokesperson of an AS can also be called a border gateway or border router. It receives routing information on behalf of the AS in the domain of other ASes or advertises the routing information updated by its own AS.

Border Gateway Protocol Link-State (BGP-LS) is a network topology information collection technique that allows the BGP protocol to aggregate topology information collected by IGP (e.g., OSPF or IS-IS) protocols and upload it to the upper layer controller. This technique emerged primarily to address the topology information collection problems of traditional IGP protocols in multi-domain or multi-AS (autonomous system) environments, which include high computational power requirements for upper layer controllers, incomplete cross-domain topology information, and complexity due to multiple IGP protocols.

The BGP-LS message format consists of two main parts: the BGP header and the BGP data. The BGP header consists of the Marker, Length, and Type fields. The Marker field takes up 16 bytes and is set to 1. The Length field takes up 2 bytes and indicates the length of the entire BGP message. The Type field takes up 1 byte and identifies the type of BGP message, such as Open, Keepalive, Update, and Notification. Each message has a fixed-size header to determine whether the data portion is included after the message header, depending on the type of message. The Type field takes up 1 byte and identifies the type of BGP message, such as Open, Keepalive, Update, and Notification. Each message has a fixed-size header, and it is based on the message type that you can determine whether or not the data portion is included after the message header. The message header fields are laid out as follows (Fig. 2):

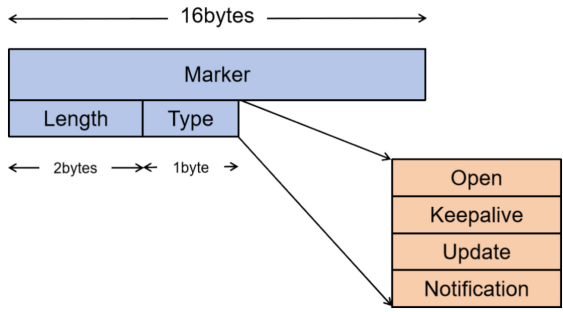


Fig. 2. BGP-LS Message Format

BGP-LS uses the new BGP Network Layer Reachability Information (NLRI) encoding format to describe node, link, and prefix information for IGP. This information is propagated through BGP Update messages. The BGP-LS NLRI includes the Node NLRI, Link NLRI, and IPv4/IPv6 Topology Prefix NLRI types. Each NLRI type contains a series of Type/Length/Value (TLV) triples to carry Node, Link, Route Prefix information, IPv6 Route Prefix information, SRv6 SID Routing information and TE Policy Routing information.

The Attribute of BGP-LS is an optional, non-passed BGP attribute used to carry parameters and attributes for links, nodes and prefixes. These attributes are also encoded in TLV format and should only be included with Link-State NLRIs.

In a real-world deployment, BGP-LS can pass topology information from IGP to BGP by configuring the IGP protocol to integrate with BGP-LS. This way, external components can connect to a BGP-LS router and obtain topology information about the entire network for path computation and traffic engineering.

The BGP-LS message format and TLV encoding method provide a flexible and scalable way to collect and distribute network topology information to support more complex network management and optimisation strategies.

4.2 BGP-LS Routing Rule

BGP-LS integrates route learning and routing policy into one, in which group attributes are extended to ensure the reasonable use of the routing policy function. BGP-LS itself does not do route calculation, but achieves route selection through rich routing rules. Therefore, it is necessary to design suitable routing rules according to the business requirements, i.e., to achieve routing control by setting reasonable routing policies. In this paper, we mention that routing policies are more widely used, which can complete the control of route announcement, reception, introduction of routes that satisfy the conditions, etc.

5 Arithmetic Routing Mechanism

In the arithmetic-aware distributed routing control scheme for power scheduling network mentioned in this paper, the arithmetic routing entry node periodically calculates the optimal arithmetic routing exit nodes corresponding to different SIDs (connected to the optimal computation service nodes) based on the state table of the power scheduling network and the arithmetic state table by the arithmetic routing algorithm, and updates the local arithmetic routing table. When the service request arrives, it is then forwarded and scheduled according to the guidance of the local arithmetic routing table of the nearest arithmetic routing entry node (Fig. 3).

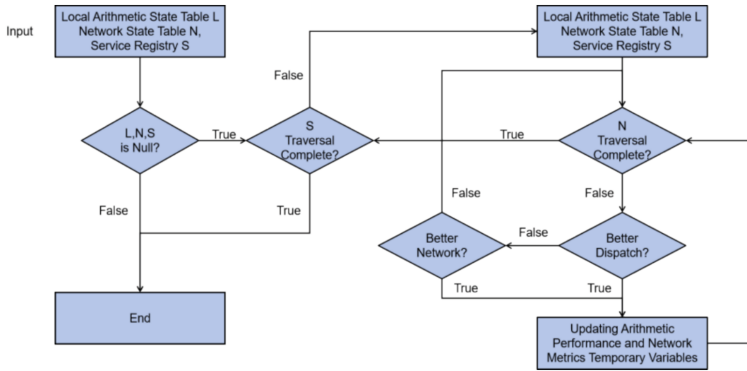


Fig. 3. AADRM-PDN flow chart

The arithmetic-aware power scheduling network aims to ensure the efficiency of arithmetic utilisation by optimising the arithmetic resource allocation scheduling, and at the same time guaranteeing the network performance metrics, which in turn improves the service satisfaction of power users. The arithmetic awareness-based distributed routing mechanism for power dispatch networks (AADRM-PDN) proposed in this paper is shown in the following figure, which aims to obtain the optimal scheduling node corresponding to each scheduling service through the input arithmetic state table, network state table and service registry.

Firstly, the initialisation of computing performance parameters and network performance parameters is carried out, and then the service registry is traversed to ensure that the optimal service nodes are calculated for different types of computing services, and for the nodes that can provide the corresponding computing services, the surplus of the computing performance of the service nodes is given priority and the service nodes with better network performance are selected under the approximation of arithmetic state, and the acquisition of the optimal nodes for computing services is completed by selecting the generation with the help of the temporary variables in the service registry. The optimal nodes for computing services are obtained by means of temporary variables, which ensures that each computing service node in the arithmetic network has load-balancing characteristics.

6 Experimental Validation

6.1 Simulation Environment

Arithmetic-aware network is oriented to various kinds of popular power services today, and it must ensure load balancing while satisfying services with high arithmetic demand. Avoid the situation that a single scheduling service node is overloaded while another similar scheduling service node is idle, and maximise the balance of arithmetic power utilisation. In this paper, in the arithmetic routing simulation, three types of service logos are registered, in which Server1 and Serve3 both support 10.10.10.1 logo service, while Server1 also supports 10.10.10.2 logo service. By deploying the worker helmet detection application, the computational load profile of Server1 and Server3 is observed by increasing the number of worker helmet detection request streams, where the computational load profile only considers the CPU load.

6.2 Results of the Experiment

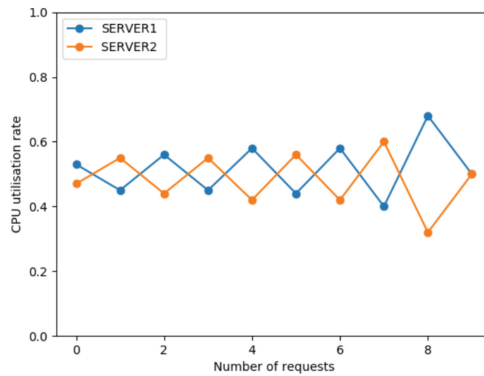


Fig. 4. AADRM Load Balancing Simulation Results

Figure 4 shows the CPU load changes of the scheduling service nodes in AADRM that support the same kind of power scheduling service demand for the surge of service request streams. It can be clearly seen that the loads of two scheduling service nodes, Server1 and Server2, are staggered with the increase of the number of service request streams, that is, the service order of the service request streams is alternating, and the two Servers maintain a good computational resource surplus to ensure that the service request streams respond to the service request streams at a very high speed and provide the corresponding services, which also verifies that the AADRM under study has a good load balancing capability. AADRM has good load-balancing capability.

In order to verify the impact of the AADRM proposed in this paper compared with the shortest path scheduling on the satisfaction of service requests, based on the simulation topology scenario of the power scheduling network constructed in this paper, the number of service request flows is detected by setting up four different groups and the number of satisfied users is counted. For a user, the user is defined as satisfied when and only when it can correctly provide the required power service resource demand and meet the maximum time delay of task processing, and the maximum time delay of task processing for this experiment is set to 1 s. It is worth mentioning that the shortest path scheduling algorithm in this experiment is to select the scheduling service node that can provide the corresponding service with the 'least number of hops' to provide service for the user, which can also be understood as the proximity scheduling.

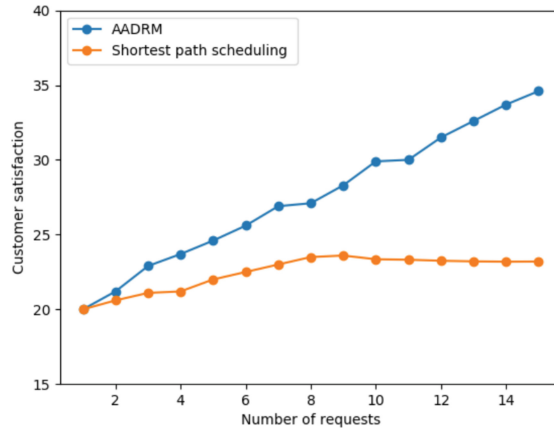


Fig. 5. Impact of different routing algorithms on user service satisfaction

As shown in Fig. 5, the AADRM studied in this paper improves user satisfaction compared to shortest path scheduling, plans optimal scheduling resources for service requests based on real-time arithmetic state, ensures arithmetic utilisation and avoids waste of resources caused by excessive pressure on individual service nodes in case of a large number of power scheduling service requests.

7 Conclusion

In this paper, we propose an arithmetic-aware power scheduling network architecture based on arithmetic awareness to address the limited arithmetic resources and scheduling optimisation needs encountered in the rapid development of power IoT. The architecture effectively solves the problems of single point of failure and excessive computational burden that may be caused by traditional centralised routing through a distributed routing control scheme. Further, the article introduces an arithmetic routing notification and decision-making mechanism, which achieves more accurate and optimised routing and scheduling decisions by sharing the arithmetic state information of each scheduling service node. Verified by simulation experiments, the proposed mechanism not only ensures the efficient use of arithmetic resources, but also meets the high satisfaction requirements of power scheduling services. Overall, this research provides an innovative solution for optimal allocation of arithmetic resources and intelligent scheduling in power IoT, which is expected to promote the further development and application of smart grid.

Acknowledgement. This Paper is Supported by Research and Development of Common Platform and Typical Application Products for Power Intelligent Terminal.

References

1. Moro, V.C., Trindade, F.C.L., Costa, F.B.B.B.D.: Distributed generation hosting capacity analysis: an approach using interval-affine arithmetic and power flow sensitivities. *Electric Power Systems Research* **226**(Jan.):1.1–1.15 (2024)
2. Raja, A.A., Mansoor, M., Zahid, F.: Optimal siting of distributed generation based on hosting capacity approach. In: 2018 International Conference on Engineering and Emerging Technologies (2018)
3. Thakur, G., Sohal, H., Jain, S.: Low-power Approximate Arithmetic Circuits for IoT devices. Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering) (2022). <https://doi.org/10.2174/2352096515666220627124337>
4. Shi, Z.: Service Flow Differentiated Priority-driven Power Line Carrier Routing Algorithm. In: 2023 3rd International Conference on New Energy and Power Engineering (ICNEPE), Huzhou, China, pp. 782–786 (2023). <https://doi.org/10.1109/ICNEPE60694.2023.10429098>
5. Reza, C.M.F.S., Lu, D.: Over-current protection for power packet dispatching system. *IEEE Trans. Circuits Syst. II Express Briefs* **67**(11), 2617–2621 (2020). <https://doi.org/10.1109/TCSII.2019.2957501>
6. Gao, X., Peng, M., Tse, C.K.: Cascading failure analysis of cyber-physical power systems considering routing strategy. *IEEE Trans. Circuits Syst. II Express Briefs* **70**(1), 136–140 (2023). <https://doi.org/10.1109/TCSII.2021.3071920>
7. Boroojeni, K., Amini, M.H., Nejadpak, A., Dragičević, T., Iyengar, S.S., Blaabjerg, F.: A novel cloud-based platform for implementation of oblivious power routing for clusters of microgrids. *IEEE Access* **5**, 607–619 (2017). <https://doi.org/10.1109/ACCESS.2016.2646418>
8. Zhang, P., Li, X., Liu, Y.: Routing optimization mechanism for srv6 based power data network. In: 2023 6th World Symposium on Communication Engineering (WSCE), Thessaloniki, Greece, pp. 61–67 (2023). <https://doi.org/10.1109/WSCE59557.2023.10365774>
9. Yoshida, N., Takahashi, R., Hikihara, T.: Power regeneration from dc motor with bidirectional router in power packet dispatching system. *IEEE Trans. Circuits Syst. II Express Briefs* **67**(12), 3088–3092 (2020). <https://doi.org/10.1109/TCSII.2020.2968384>



A Dual-Mode Integrated Communication Method for Power Line Carrier and Wireless in Distribution Systems

Xiao Feng¹(✉), Peiyao Zhang¹, Mingtao Cui¹, Yuhao Chang¹, Yong Sun²,
and Wenbin Liu²

¹ State Grid Information and Telecommunication Co., Ltd., Beijing 102211, China
wche@bupt.cn

² State Grid Shandong Electric Power Company, Electric Power Scientific Research Institute,
Jinan 250002, China

Abstract. The advent of the Internet of Things (IoT) and smart grids has placed significant strain on communication networks based on single-channel access methods, such as power line communication (PLC) and wireless communication. These networks are unable to meet the demands of distribution systems. The integration of communication between power lines and wireless channels can yield complementary advantages and enhance the performance of communication networks. In light of the discrepancies between PLC and wireless communication with regard to transmission medium, data rate, and interference resistance, this paper puts forth a dual-mode integrated communication method that fuses power line carrier and wireless communication. At the physical layer, the employment of orthogonal frequency-division multiplexing (OFDM) modulation technology facilitates the achievement of signal diversity processing in both the power line carrier communication channel and the wireless communication channel, thereby ensuring the transmission of signals over both channels. At the media access control (MAC) layer, a p-persistent CSMA algorithm is proposed as a means of selecting the channel for signal transmission. This is intended to reduce the probability of collisions through the use of backoff mechanisms, thereby improving channel utilisation. The experimental results demonstrate that the proposed method exhibits competitive performance in terms of throughput, delay, and other metrics.

Keywords: Collaborative Communication · Orthogonal Frequency-Division Multiplexing · Media Access Control

1 Introduction

The distribution system represents a critical component of the power network, responsible for the delivery of electricity from substations to end-users. In this process, the collection of information plays a pivotal role, involving the gathering of a range of data, including voltage, current, power, and energy consumption. This data is of paramount

importance for the monitoring and maintenance of the grid, ensuring the reliability of the power supply. By utilizing smart meters, sensors, and other monitoring devices, the distribution system can obtain real-time critical information and use data analysis to enhance efficiency, prevent failures, and support decision-making.

Bidirectional, real-time, and reliable communication networks are essential for distribution systems. Power line communication (PLC) are typically embedded within walls, making them relatively secure and less prone to damage. However, the quality of PLC communication can be variable due to factors such as line noise interference, time-varying input impedance, and signal attenuation. The utilisation of wireless communication, encompassing low-power wireless, Wi-Fi, wireless broadband and mobile communications, offers a convenient solution that obviates the necessity for wiring. However, the quality of wireless communication can also be affected by a number of other factors, including the frequency bands used, the geographic location of the transmission, and the environmental conditions prevailing at the time. Furthermore, the attenuation of wireless signals when encountering obstacles represents another significant challenge. In the context of smart grid applications, the environment is complex, the requirements for business support are diverse, and high reliability in transmission is demanded. No single communication method can fully satisfy the requirements for the construction of distribution system communication networks.

To ensure the reliability of distribution system communication networks, both academia and industry consider the combination of power line and wireless technologies to be an effective means of overcoming the limitations of individual communication modes. Integrated communication networks can leverage the strengths of each communication method to achieve cooperative gains and improve the performance of business transmission in distribution networks. This paper proposes a dual-mode integrated communication method with multiplexed access. At the physical layer, the method implements dual-mode integration of the power line channel and the wireless channel using OFDM. At the Media Access Control (MAC) layer, a p-persistent CSMA algorithm is designed. Under this algorithm, the device monitors the PLC channel and the wireless communication channel at the same time to seek opportunities to send data packets. Data can be transmitted on both channels and the probability of conflicts can be reduced through the p-persistence mechanism.

2 Related Works

PLC technology can be categorised into narrowband, midband and broadband technologies according to frequency and rate. Narrowband PLCs are mainly used for low-rate applications such as remote meter reading, while medium-band PLC technology is suitable for scenarios such as smart grids that require higher reliability and real-time performance. Broadband PLC, on the other hand, offers higher data transmission rates and is suitable for scenarios such as home broadband access.

In practice, PLC technology faces challenges such as complex grid environments, severe noise interference, and high time variability. To improve the communication quality, a variety of technical means are used, such as OFDM (Orthogonal Frequency Division Multiplexing) technology, anti-noise technology, anti-channel fading and load

variation technology. The application of these techniques makes PLC communication more stable and efficient.

In order to improve the stability and efficiency of PLC communication, the key lies in choosing a suitable communication protocol, such as Modbus, Profibus or Profinet, to ensure high efficiency and reliability of data transmission. At the same time, designing optimised communication interfaces and adopting efficient modulation techniques, such as OFDM, can enhance interference immunity and data transmission rates. Ensuring electromagnetic compatibility (EMC) is also an important measure to improve the stability of the PLC system, which involves reducing electromagnetic interference and power supply interference, as well as ensuring that the temperature and humidity of the working environment are appropriate and avoiding the influence of pollutants and corrosive gases. In addition, modular design of hardware and software helps to adapt to changing communication standards and market opportunities. The use of high-quality communication cables and equipment, as well as regular system maintenance and inspection, are also effective ways to ensure stable operation of PLC communications. Through these combined measures, the overall performance and reliability of a PLC system can be significantly improved.

PLC and wireless communication have been widely applied in distribution systems. In recent years, considerable research has focused on the collaborative operation of PLC and wireless communication. Gonzalez et al. [1] proposes a combined communication approach that integrates PLC and wireless communication, demonstrating its feasibility. Sayed et al. [2] investigates the interference resistance of hybrid PLC and wireless communication systems from the perspective of signal interference, suggesting the utilization of signal diversity between the power line and wireless links to enhance system robustness. Due to the different transmission media and significant differences in signal frequencies [3], it is challenging to achieve fusion at the physical layer between PLC and wireless communication. WANG et al. [4] presents a physical-layer fusion method based on signal-to-noise ratio (SNR) preference for PLC and wireless communication, where the receiver selects the higher SNR signal among the two paths, which results in underutilization of resources. The IEEE 802.11 standard [5] provides detailed specifications for the MAC layer and physical layer for local area networks. The IEEE 1901 standard [6] for PLC uses OFDM technology at specifies CSMA/CA and TDMA protocols at the MAC layer, enabling transmission rates of up to 200 Mbps and higher, thus supporting high-speed transmission in various scenarios. The IEEE 1905 standard [7] introduces a virtual MAC layer architecture that establishes an abstraction layer between the logical link layer and the media access layer to integrate Ethernet, wireless LAN, power line carrier, and coaxial cable communication technologies. Chen et al. [8] proposes a selective algorithm based on the CSMA/CA protocol and utilizes state decomposition and Markov models to analyze the system performance of dual-medium cooperative communication at the MAC layer. However, this algorithm is designed for a first idle-first contention scheme, where a station can attempt to transmit a packet as soon as only one of the backoff counters reaches zero. Furthermore, the algorithm is computationally complex, requiring iterative solutions to equation sets.

3 Power Line Carrier and Wireless Dual-Mode Integrated Communication Method

This paper considers a model of a collaborative communication system based on PLC and wireless communication, as shown in Fig. 1. The model includes N devices, each equipped with both PLC and wireless interfaces and using the same access algorithm.

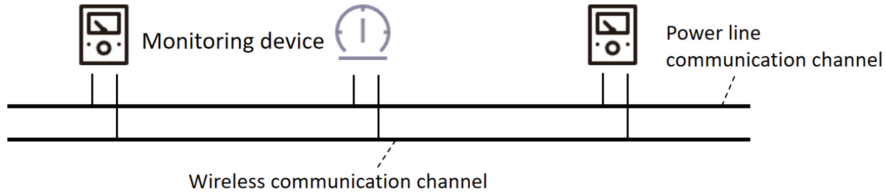


Fig. 1. Model of a Collaborative Communication System Based on Power Line and Wireless Communication

The comprehensive communication process proposed in this paper for the integration of power line and wireless communication in distribution systems starts at the sender and traverses the communication protocols at each layer to convey information to the application layer of the receiver. The sequence of data flow at the transmitter end is as follows: the application layer of the integrated dual-mode system, the network layer, the medium access control (MAC) layer and the physical layer. Conversely, the sequence of data flow at the receiving end is: the physical layer of the dual-mode integrated system, the MAC layer, the network layer and the application layer. The physical layer of the dual-mode integrated system provides the physical medium connections and pathways for data transmission in the distribution system communication. These include both wired and wireless media, and the layer continuously converts the transmitted ‘0’ and ‘1’ bit streams into physical signals and vice versa. The MAC layer of the dual-mode integrated system bears the primary responsibility for framing and deframing the ‘0’ and ‘1’ bit streams from the physical layer, while also providing access methods for the shared dual-mode medium. The objective of the network layer of the dual-mode integrated system is to facilitate transparent end-to-end data transmission within the communication system. This is typically achieved through the implementation of specific functions, including addressing, routing selection, and path maintenance, which collectively enable the delivery of the services associated with the dual-mode integrated network layer. The application layer of the dual-mode integrated system provides an interface between the application service programs and the underlying network layer for message transmission, functioning as both the initial point of contact and the final destination within a comprehensive dual-mode integrated communication process.

The dual-mode integrated communication method for distribution systems, as proposed in this paper, entails the processing of input signals by the physical layer for both the power line carrier communication channel and the wireless communication channel. This is achieved through the utilisation of Orthogonal Frequency Division Multiplexing (OFDM) modulation technology, with the resulting signals then being output. At present,

power line carrier communication in distribution systems utilises OFDM-based physical layer modulation signal technology. Additionally, there are numerous OFDM-based wireless communication technologies available, including Wi-Fi and 4G. In the context of advanced distribution system communication networks, the adoption of Orthogonal Frequency Division Multiplexing (OFDM) as a modulation technique presents a unique opportunity for the seamless integration of power line carrier communication (PLCC) with wireless communication channels. By leveraging the same OFDM technology for both channels, a novel physical layer integration approach can be conceptualized, as depicted in Fig. 2. This approach harnesses the concept of spatial diversity for OFDM signals, enabling a dual-mode communication paradigm that synergizes PLCC with wireless channels.

The integration process commences with the OFDM modulation of the input signal, which is then simultaneously transmitted across both the wireless and PLCC channels. This dual-channel transmission ensures that the signal is not only propagated through the conventional wireless medium but also through the power line infrastructure, thereby creating a redundant and robust communication pathway. The essence of this strategy lies in the replication of the input signal across two distinct spatial domains, enhancing the signal's resilience against channel-specific impairments and interference.

Upon reaching the receiver, the signals intercepted from both the wireless and PLCC channels are subjected to a sophisticated diversity combining algorithm. This process intelligently amalgamates the received signals, capitalizing on the spatial diversity to mitigate the effects of fading, noise, and other channel-induced distortions. The combined signal, which is a result of this synergistic integration, is then demodulated using OFDM techniques, allowing for the accurate reconstruction of the original information.

This dual-mode integration strategy not only enhances the reliability and robustness of the communication system but also optimizes the spectral efficiency by leveraging the inherent advantages of OFDM modulation. The method's innovative application of spatial diversity in the context of OFDM signals presents a significant advancement in the field of distribution system communication networks, offering a promising solution for future smart grid and IoT applications where reliability and efficiency are paramount. In summary, the proposed integration method represents a forward-thinking approach to communication system design, where the convergence of PLCC and wireless communication channels is achieved through the strategic application of OFDM modulation and spatial diversity techniques. This research contributes to the body of knowledge in the field, providing a foundation for further exploration and development of advanced communication strategies that can meet the demands of evolving power distribution networks.

In the realm of dual-medium communication systems, devices are endowed with the capability to operate over both power line and wireless interfaces, thus enabling concurrent access contention on these distinct channels. At the MAC layer, these devices adhere to the protocols specified by IEEE 1901 for power line communication and IEEE 802.11 for wireless communication, respectively. To augment channel efficiency, a p-persistent Carrier Sense Multiple Access (CSMA) algorithm has been ingeniously crafted, embracing the principles of spatial diversity for OFDM signals.

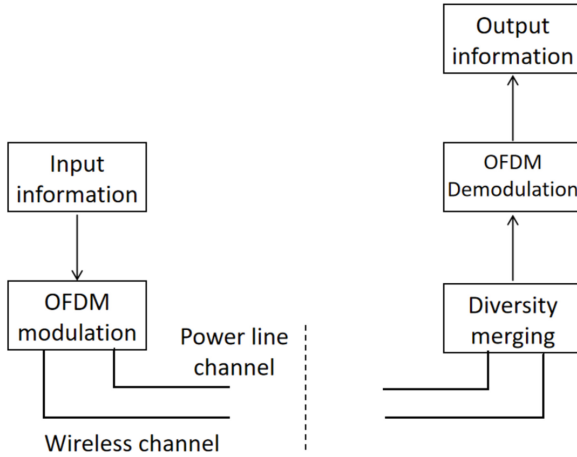


Fig. 2. OFDM-based physical layer signal modulation

This algorithm is orchestrated by a quintet of counters: BPC1, BPC2, BC1, BC2, and DC, which meticulously manage the backoff process. Specifically, BPC1, BC1, and DC are designated for the power line channel, serving as the backoff stage counter, backoff counter, and delay counter, respectively. In parallel, BPC2 and BC2 fulfill analogous roles for the wireless channel. Upon successfully securing access to a channel, such as the power line channel, within an allocated time slot, the device proceeds to transmit data over that medium, while the wireless channel remains poised for access by other devices. This strategic approach ensures the comprehensive utilization of both channels, thereby amplifying the system’s performance.

In accordance with the algorithm, when a device is poised to transmit packets, it initiates the counters and embarks on the backoff process. It is postulated that the backoff processes on both channels operate independently, without cross-interference. Upon achieving a successful transmission, all counters are reset, thus resetting the device’s state for subsequent communication cycles. The intricacies of the dual-channel p-persistent CSMA algorithm are delineated in Algorithm 1, providing a comprehensive framework for efficient channel access in dual-medium communication systems.

This approach not only enhances the reliability and robustness of the communication system but also optimizes spectral efficiency by leveraging the inherent advantages of OFDM modulation and the strategic employment of spatial diversity techniques. The integration of these advanced communication strategies paves the way for future-proof smart grid and IoT applications, where the demand for high reliability and efficiency is paramount.

Algorithm 1: Dual-Channel p-Persistent CSMA Algorithm

1. Step 1: Devices set BPC1 and BPC2 to 0, indicating the current backoff stage is 0. BC1 and BC2 are initialized to random values within their respective backoff windows. DC is assigned an empirical value.
 2. Step 2: Devices simultaneously monitor both the power line and wireless communication channels. If a channel is detected as busy, the corresponding BC counter is frozen; if the channel is sensed as idle in the current time slot, the BC value for that channel is decremented by one.
 3. Step 3: When BC1 or BC2 reaches zero, the device attempts to send a packet over the corresponding channel with probability p , and with probability $1 - p$ delays the transmission for one time slot before attempting again with probability pp . Other devices continue to compete for the other channel. In the event of a collision, the backoff stage counter BPC for the collided channel is incremented, moving the device to the next backoff stage, and the BC counter is reinitialized to a random value within the new backoff window. The process then returns to Step 2
 4. Step 4: If both BC1 and BC2 reach zero at the same time, the device randomly selects one channel and attempts to send a packet with probability p , and with probability $1 - p$ delays the transmission for one time slot. The other channel continues to be used by other devices. In the event of a collision on the selected channel, the BPC, BC, and DC counters for that channel are updated, and the process returns to Step 2.
 5. Step 5: Upon receiving an ACK signal, the device stops contending for the channel, indicating a successful packet transmission.
-

When a device successfully transmits a data packet, it is crucial to immediately initialize all five counters to ensure the system's readiness for the next transmission. This step is essential for maintaining overall fairness in the network and minimizing the coupling effect between different devices. If only the counters associated with the successfully transmitted packet are reset, the remaining counters could disrupt the system's fairness, leading to imbalanced resource allocation and potential communication delays.

The p-persistent Carrier Sense Multiple Access (CSMA) mechanism plays a vital role in reducing data conflicts compared to the traditional CSMA algorithm. This improvement is particularly significant for enhancing system capacity and reliability. By controlling when a device is allowed to transmit data based on a probability p , the p-persistent CSMA reduces the likelihood of collisions and ensures more efficient use of the communication channel.

The performance of the p-persistent CSMA algorithm is heavily influenced by the setting of the probability p . If p is set too high, multiple devices may attempt to transmit simultaneously, leading to increased collisions and reduced system performance. Conversely, if p is set too low, devices may wait unnecessarily long before transmitting, which can also degrade performance. Therefore, the value of p is typically adjusted dynamically according to the state of the communication network. Factors such as network load, traffic patterns, and the number of active devices are considered when setting p to optimize the balance between minimizing collisions and maximizing throughput.

This adaptive adjustment helps the system to operate efficiently under varying network conditions, thereby improving overall communication performance.

4 Experiments

In this chapter, MATLAB is employed for the analysis of the power line carrier and wireless dual-mode integrated communication method, with a view to verifying the performance of the aforementioned method. The simulation scenario represents a fusion of power line carrier communication and micro-power wireless communication. It is assumed that the device will not lose packets at the physical layer, and that collisions are caused by multiple devices transmitting simultaneously. In order to facilitate comparison, throughput and latency have been selected as the key indicators. The PLC method and wireless communication method have been identified as the comparison methods. The maximum backoff stage of both the wireless communication channel and the PLC channel is set to 3. In order to control the performance of the two channels to be as close as possible, the backoff window value of the PLC channel is set to $CW1 = [8, 16, 32, 64]$, the backoff window value of the PLC channel is set to $CW2 = [32, 64, 128, 256]$, and the PLC channel delay counter $DC = [0, 1, 3, 15]$. The average length of each data packet of the station is set to 8184 bits, the channel rate is set to 1 Mbit/s, the time slot is set to $50\mu s$, and other parameter settings are consistent with the IEEE 1901 and IEEE 802.11 standards. The p value of the p-persistence mechanism is set to 80%.

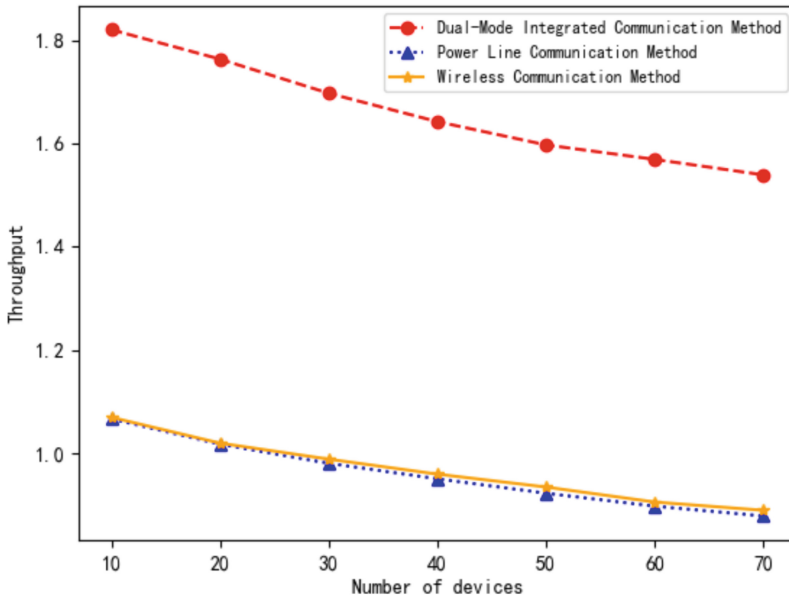


Fig. 3. Comparison of throughput under various methods

Figure 3 analyzes the trend of throughput changes with the number of accessing devices in the dual-medium communication system model, using the single wireless communication method and the single PLC method as comparative methods. It is observed that the throughput of the dual-mode integrated communication method is significantly higher than that of the single wireless communication method and the single PLC method. This is because in the multiplexed access algorithm, devices do not need to wait for both backoff counters to reach zero, making it easier for the device's backoff counters to reach zero and increasing the probability of the device winning the channel. Moreover, during the transmission process, a device occupies only one channel, while the other channel remains available for use by other devices. Therefore, the resources of both channels can be fully utilized, improving the throughput performance.

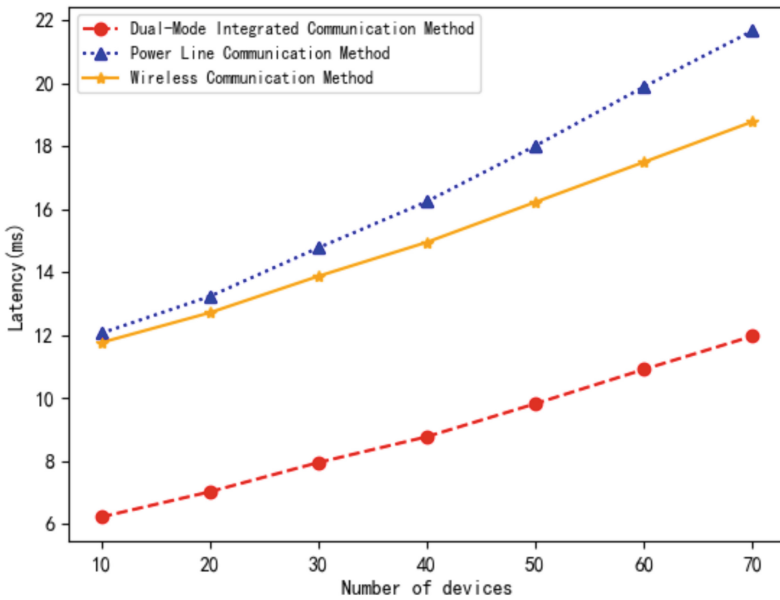


Fig. 4. Comparison of latency under various methods

Figure 4 presents an analysis of the trend of delay changes in the context of the dual-medium communication system model, with the single wireless communication method and the single PLC method serving as comparative benchmarks. It can be observed that as the number of devices increases, the delay for all methods demonstrates an upward trend. This is due to the fact that an increase in the number of devices increases the probability of conflicts. Nevertheless, the delay of the dual-mode integrated communication method is considerably less than that of the single wireless communication method and the single PLC method. The utilisation of parallel communication over dual channels has been demonstrated to markedly enhance communication efficiency. Furthermore, the configuration that enables devices to probabilistically transmit data upon reaching zero backoff counters serves to further reduce the probability of conflicts.

5 Conclusion

This paper puts forth a dual-mode integrated communication method that fuses power line carrier and wireless communication. At the physical layer, the method employs space diversity processing of input signals for both the power line carrier communication channel and the wireless communication channel, utilising OFDM, and subsequently outputs the processed signals. At the Media Access Control (MAC) layer, a multiplexed access algorithm based on both power line and wireless media is proposed. This allows devices to randomly transmit data packets across both channels, thereby improving system performance and channel utilisation. The method presented in this paper achieves heterogeneous integration of power line and wireless networks, optimises network performance, and enhances the success rate of data collection.

Acknowledgement. This Paper is Supported by Application and Demonstration of Minimised and Accurate Collection of Low Voltage Distribution System Based on Cloud-Edge-End Collaboration.

References

1. Gonzalez, O.A., Urminsky, J., Calvo, M., et al.: Performance analysis of hybrid broadband access technologies using PLC and Wi-Fi. In: 2005 International Conference on Wireless Networks, Communications and Mobile Computing. IEEE, 1, 564–569 (2005)
2. Sayed, M., Al-Dhahir, N.: Narrowband-PLC/wireless diversity for smart grid communications. In: 2014 IEEE Global Communications Conference. IEEE, pp. 2966–2971 (2014)
3. Sinha, T., Bhaumik, J.: Design of computationally efficient sharp FIR filter utilizing modified multistage FRM technique for wireless communications systems. *J. Electronic Science and Technol.* **17**(2), 185–192 (2019)
4. Wang, X., Wang, L., Song, Y., et al.: Power line carrier and wireless communication physical layer fusion method based on signal to noise ratio optimization. *Electric Power Information and Communication Technology* **16**(9), 14–20 (2018)
5. Feng, W., Nallanathan, A., Krishna, G.H.: Performance of physical (PHY) and medium access control (MAC) layers of IEEE 802.11 b in the presence of Bluetooth piconets. In: The 57th IEEE Semiannual Vehicular Technology Conference, 2003. VTC 2003-Spring. IEEE **2**, 1489–1492 (2003)
6. IEEE Standards Association. IEEE standard for broadband over power line networks: Medium access control and physical layer specifications. IEEE std, 1901, 2010, 1–1586 (2010)
7. IEEE 802 Working Group. IEEE Standard for a Convergent Digital Home Network for Heterogeneous Technologies; IEEE Std 1905.1–2013. IEEE: Manhattan, NY, USA, pp. 1–93 (2013)
8. Chen, Z., Liu, Y., Liu, R., et al.: Improved CSMA/CA algorithm based on alternative channel of power line and wireless and first-time idle first acquisition. *IEEE Access*, **7**, 41380–41394 (2019)



High-Performance Database Storage Technology for Power Smart Terminals

Lvchao Huang¹, Xiao Feng²(✉), Jinguo Fang¹, Yuhao Chang², and Guangwei Liu²

¹ State Grid Information & Telecommunication Co., Ltd., Beijing 102211, China

² State Grid Info-Telecom Great Power Science and Technology Co., Ltd.,
Xiamen 361008, China
wche@bupt.cn

Abstract. Smart terminals, with their portability and powerful computing capabilities, have been widely applied across various fields. They are particularly prevalent in new power systems, where they are extensively used for tasks such as equipment monitoring, data collection, and remote control. These application scenarios demand enhanced storage and data processing capabilities from databases. Given the increasingly complex deployment scenarios, smart terminals face challenges such as large data volumes, diverse data types, limited storage space, and constrained processing performance. Efficiently storing and processing large amounts of data under resource constraints has become a significant challenge. This paper investigates high-performance database storage technologies tailored for smart terminals and proposes an integrated database storage architecture and optimisation scheme that adapts to the characteristics of the smart terminal environment. We have developed a database framework specifically designed for the data processing characteristics of smart terminals, employing lightweight rule engine-based data cleaning techniques and the FVD-SDT algorithm-based data compression technology. Experimental results demonstrate that the proposed methods significantly improve storage space utilization and data processing efficiency, providing technical support for the application of smart terminals in new power systems, industrial IoT, and similar scenarios.

Keywords: Smart Terminal · Data Cleaning · Data Compression

1 Introduction

Smart terminals are highly integrated devices equipped with various functions, including data processing, communication, and control, making them a critical component of new power systems. They are responsible for generating and collecting most of the data within these systems. Typically, smart terminals utilize embedded system technologies, combining software and hardware to enable various power applications [1]. The widespread use of smart terminals has significantly enhanced the real-time processing and reliability of data in new power systems and other IoT environments. However, this has also raised the performance requirements for database systems deployed on smart terminals.

As database technology has evolved, particularly with the widespread adoption of embedded databases, databases have become an indispensable part of smart terminals. These databases must possess efficient data storage and processing capabilities to address the challenges posed by resource-constrained environments [2]. In particular, databases deployed on smart terminals must support high concurrency processing, low latency response, multi-language support, robust data compression and storage optimisation to meet the complex requirements of power smart terminals. However, power smart terminals often encounter a series of challenges in practical applications, such as large data volumes, diverse data types, limited storage space, and constrained processing performance. These challenges make it difficult for traditional database technologies to operate efficiently in smart terminal environments. Consequently, the requirements for databases in power smart terminals have become more stringent. Not only must these databases support efficient storage and processing of massive amounts of data, but they must also provide fast data access and robust data compression capabilities to maximize the utilization of limited system resources.

To address these issues, researchers have begun to explore database optimisation technologies tailored to the power smart terminal environment. Some researchers have opted to trim the functional modules of traditional databases to meet the data processing and storage needs of smart terminals in specific environments with lower resource consumption. To address the limited storage space in smart terminals, other researchers have applied data compression techniques to large files or bulk data storage. The aim of these studies is to improve the performance of smart terminal databases to meet the requirements of smart terminals in power systems. Based on the aforementioned background and challenges, this paper investigates high-performance database storage technologies for power smart terminals. Our focus is on enhancing data storage and access performance within the resource-constrained environment of smart terminals, providing robust support for their application in new power systems. Our specific contributions are as follows:

- (1) We designed a database framework tailored to the smart terminal environment that effectively addresses the need for efficient data processing in resource-constrained settings.
- (2) We proposed a lightweight rule engine-based data cleaning technique, which, through modular design and customizable rule configuration, significantly improves the efficiency and flexibility of data cleaning.
- (3) We developed an FVD-SDT-based data compression technology that incorporates a dynamic adjustment mechanism to optimize the data compression ratio and efficiency, thereby enhancing storage space utilization and data integrity.

2 Related Works

2.1 Embedded Databases

Due to their compact design and low resource consumption, embedded databases are widely used in smart terminals, IoT devices, and mobile devices, with the goal of operating efficiently in resource-constrained environments. In recent years, researchers have focused on enhancing the performance of embedded databases, particularly in terms of

high-concurrency processing and cross-language support in smart terminal applications. Some research has concentrated on trimming and optimizing database system functions to address the storage and computational resource limitations of smart terminals. For instance, certain studies have explored designing specialized databases for specific data scenarios. For example, paper [3] utilizes the graph database Neo4j to store a knowledge graph of the power grid model, leveraging the advantages of graph databases in handling complex network relationships. Other studies have designed data conversion tools to enhance database flexibility and accommodate data access needs in various application scenarios. Paper [4] introduces a tool called FlexiDC, a computerized tool that handles the conversion between different types and formats of databases, reducing the costs and time required for database enhancement, migration, integration, conversion, and new development, thereby helping users better manage the evolving demands of existing databases. These research outcomes provide more lightweight and efficient database solutions for smart terminals.

2.2 Data Cleaning Technology

Data cleaning technology plays a crucial role in improving data quality and reducing anomalies and noise. In recent years, there has been an increasing focus on research related to the complex data environments of power smart terminals. Some researchers have proposed rule-based or state-based cleaning algorithms to improve data consistency and accuracy. For example, paper [5] proposes a spatiotemporal correlation constraint-based method for cleaning power grid operational data to address the issues of poor convergence, high computational complexity, and difficulty in effectively handling significant error data found in traditional state estimation-based data cleaning methods when dealing with large-scale data. Paper [6] suggests performing targeted data cleaning steps before conducting association rule analysis by setting reasonable thresholds for support, confidence, and filtering to identify and remove anomalous data. Additionally, other studies have explored how to dynamically optimize the data cleaning process by incorporating machine learning techniques. Paper [7] applies a Deep Q-Network (DQN) in the feature selection process for multi-source data, using the DQN approach to automatically learn and optimize the feature selection strategy, thereby improving the efficiency and effectiveness of data cleaning.

2.3 Data Compression Technology

Data compression technology is critically important in various scenarios, especially in environments with limited storage and transmission resources, such as power smart terminals, IoT devices, and embedded systems. Through compression techniques, data storage space can be effectively reduced, transmission bandwidth requirements lowered, and overall system performance enhanced. Data compression can be categorized into two main types: lossless compression and lossy compression. Lossless compression retains all the original data during compression and decompression, making it highly suitable for applications where data accuracy is paramount, such as financial data storage and medical image storage. Common lossless compression algorithms include Huffman coding, LZ77, and SNAPPY compression [8]. Lossy compression, on the other hand,

achieves a higher compression ratio by discarding some data, which is suitable for scenarios where data accuracy is less critical, such as audio and video compression or environmental monitoring data compression [9]. In these scenarios, slight data loss does not significantly affect usability. Typical lossy compression algorithms include JPEG, MP3, and the SDT compression algorithm [10], where the SDT algorithm constructs parallelograms to select and retain key data points, achieving a high compression ratio when data changes are minimal.

Although the above works provide a detailed discussion of embedded databases, data cleansing techniques and data compression techniques, it lacks an in-depth analysis of the practical effects and applicable scenarios of data cleansing techniques; while the data compression techniques section distinguishes between lossless and lossy compression, it lacks an adequate description of the specific implementation and optimisation details of SDT compression algorithms.

3 Database High-Performance Storage Solution for Smart Terminals

3.1 Database Functional Framework

To meet the data storage requirements in smart terminal environments, this paper proposes a database design solution tailored for smart terminals, which aims to improve the storage performance of databases within the smart terminal operating environment. The framework is designed with a strong emphasis on optimising data processing capabilities and storage performance to ensure efficient operation in complex application scenarios, while taking into account the resource constraints typical of smart terminals.

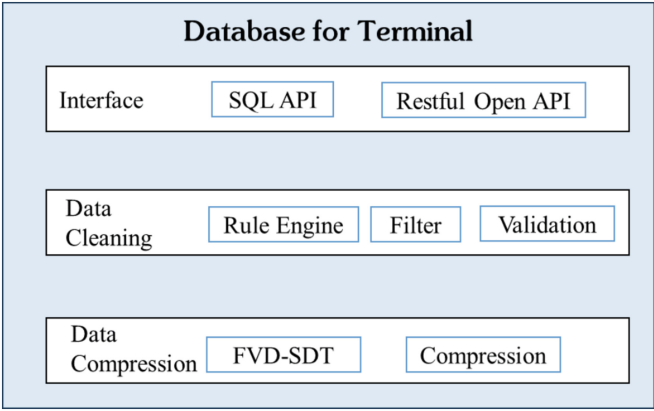


Fig. 1. Database Functional Framework

As shown in Fig. 1. The database functional framework is composed of the following core modules:

Data Access Engine: The data access engine supports multiple programming languages, including C, C++, Go, and Java, by extending the SQL API, providing a unified SQL data operation service. This multi-language access engine ensures seamless data interaction across different applications within the smart terminal. Additionally, a Restful-style OpenAPI interface is designed, allowing users to customize the data transmission protocol type (e.g., text, binary, JSON) between the client and server, according to the requirements of different scenarios.

Data Cleaning Module: The data cleaning module is a critical component of the smart terminal database functional framework. It is responsible for ensuring data quality throughout the entire process, from data collection to storage. To address common data quality issues in smart terminal environments—such as data noise, missing values, and duplicate data—this module integrates various cleaning algorithms, including duplicate value processing and outlier detection. The data cleaning module ensures the accuracy and consistency of the data stored in the database, thereby enhancing the reliability of data processing.

Data Compression Module: The data compression module is designed to optimize the limited storage resources and data transmission efficiency of smart terminals. Embedded databases often need to handle large volumes of sensor data, log data, and more. The use of efficient data compression algorithms can significantly reduce storage requirements and communication bandwidth. The data compression module supports an innovative data compression scheme based on the FVD-SDT (Flow Variability Degree-Swing Door Trending) algorithm, tailored to meet the varying precision needs of data compression in different scenarios.

Through its modular design, the database functional framework flexibly addresses the diverse needs of smart terminal environments. The data access engine ensures multi-language support and efficient data interaction, while the data cleaning and data compression modules provide essential data preprocessing and storage optimization techniques.

3.2 Data Cleaning Technology for Smart Terminal Databases Based on Lightweight Rule Engine

In the context of database storage for power smart terminals, data cleaning is a critical technology to ensure data quality and system performance. As smart terminals become more widely used, data collection processes are prone to various issues, such as sensor failures, network delays, formatting errors, and duplicate data, leading to degraded data quality. Uncleaned raw data can result in wasted storage space, reduced system performance, and even decision-making errors. Therefore, designing an efficient database data cleaning solution is of paramount importance. This paper introduces a data cleaning technology based on a lightweight rule engine, which ensures data accuracy and reliability in environments with limited computing resources. As shown in Fig. 2. The proposed rule engine for smart terminal databases comprises the following functional modules:

Data Deduplication Module: In the mass data collection process of smart terminals, duplicate data issues can lead to wasted storage space and reduced data quality. To address this, the rule engine is equipped with a data deduplication module, which by default uses

a Bloom Filter as the deduplication rule. The advantage of the Bloom Filter lies in its ability to perform efficient deduplication operations with minimal memory usage. Although it may have a certain false positive rate, its extremely low memory consumption and fast query performance make it highly suitable for resource-constrained smart terminal environments. Additionally, the data deduplication module can be replaced with other deduplication technologies, such as Roaring Bitmap, hash functions, fingerprint recognition, or similarity metrics, depending on the application scenario of the smart terminal.

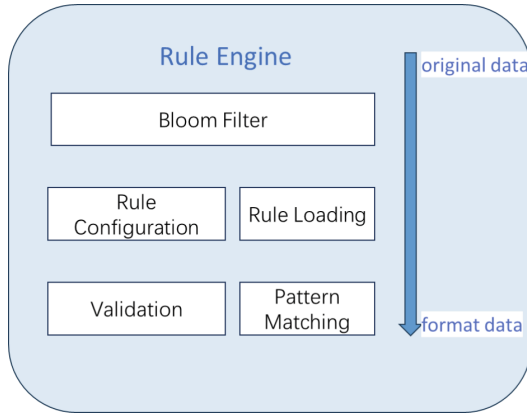


Fig. 2. The rule engine function module

Rule Configuration Module: In the lightweight rule engine, users can configure different rules according to the data scenarios faced by smart terminals to achieve personalized data cleaning. The rule configuration module allows operators to customize the complexity of the rules and the application scenarios based on actual needs, ensuring that the system can operate under different resource conditions.

Rule Loading Module: To further improve system efficiency, the lightweight rule engine includes a rule loading module. For more complex pattern matching rules, the system can preload or precompile them to avoid frequent pattern parsing during runtime, thus reducing unnecessary computational overhead. The rule loading module ensures the stability and efficiency of the system during operation, particularly in resource-limited smart terminal environments.

Rule-Based Validation Module: The rule-based validation module is one of the core components of the lightweight rule engine, primarily used to verify the format of collected data to ensure its integrity, consistency, and parsability. This module checks each data entry against a pre-defined set of rules to ensure it meets the required format standards. Validation typically includes data type checks, data range validation, and data dependency validation. Additionally, to adapt to the diverse needs of smart terminals, the rule-based validation module can be flexibly configured based on different deployment scenarios. For example, in a high-frequency data collection environment, validation rules can be optimized to reduce system resource consumption.

Pattern Matching Module: The pattern matching module is another critical function of the lightweight rule engine, used to identify and process complex data patterns. In the data collection process of power smart terminals, data often exhibits periodicity, repetition, and regularity. Therefore, the pattern matching module can help identify anomalies and predefined patterns in the data, such as sudden fluctuations in power load or daily cyclical changes. Traditional pattern matching methods often rely on complex algorithms and significant computational resources, which are limited in smart terminals. To address this, the proposed pattern matching module uses regular expression-based pattern matching algorithms to reduce computational overhead.

Traditional rule engines are feature-rich but come with high computational and resource costs. To accommodate the resource constraints of smart terminals, the proposed lightweight rule engine trims down traditional rule engine functionalities, focusing on essential data cleaning functions while eliminating complex inference and multi-layer rule processing. The lightweight rule engine adopts a modular design, allowing each functional module to be flexibly adjusted according to different deployment environments and data characteristics, ensuring efficient and stable operation in resource-constrained environments.

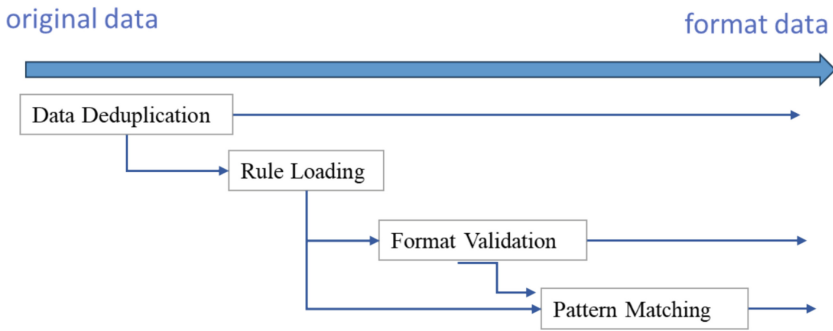


Fig. 3. Rules engine workflows

In Fig. 3, the operation mode of the lightweight rule engine follows these steps:

Data Deduplication: First, the system uses the data deduplication module (e.g., Bloom Filter) to remove duplicate data from the collected data, reducing storage space usage and improving data quality.

Reading Rule Configuration Information: Next, the rule engine reads and loads the predefined rule configuration information. This process includes obtaining the user-defined rule set from the rule configuration module and precompiling or preloading it through the rule loading module to improve the efficiency of subsequent data processing.

Rule-Based Validation and Pattern Matching: Finally, the system applies the rule-based validation and pattern matching modules to further process the data. The rule-based validation module ensures that the data meets the predefined format and dependency requirements, while the pattern matching module identifies and processes anomalies and specific patterns in the data, ensuring its accuracy and consistency.

In practical applications, users can modify the data processing flow of the lightweight rule engine based on the specific data scenarios they face. For example, the “Data Deduplication Module” can be omitted in scenarios where duplicate data rarely occurs, and the process can end after deduplication in cases where anomalies are rare. This flexible design allows the lightweight rule engine to be freely deployed in different smart terminal environments, meeting various personalized needs.

In summary, the lightweight rule engine successfully implements efficient data cleaning in power smart terminal environments through modular design and targeted optimization. The engine simplifies the operation process while retaining key functions, offering high flexibility and adaptability, and providing reliable data cleaning support for smart terminal databases.

3.3 Data Compression Technology for Smart Terminals Based on FVD-SDT

Smart terminals need to process large amounts of sensor data, log data and status information. Given the limited storage and computing resources of these devices, selecting the right data compression technology is critical to achieving high performance storage. The Swing Door Trending compression algorithm is a lossy compression technique based on linear fitting, primarily used to reduce redundancy in data storage. It achieves data compression by constructing parallelograms to select and retain key data points. As shown in Fig. 4, point A is the starting point of the first compression segment. At a vertical distance of E (compression deviation) from point A, two support points are set, and a parallelogram is constructed that includes points A and B. Since the parallelogram only includes points A and B, the line segment AB cannot form a compression segment, meaning the compression process is not yet complete. Next, the line segment AC is evaluated. It is found that the parallelogram constructed with a side length of twice the compression deviation (2E) and containing points A and C includes all data points between A and C. If the parallelogram constructed with points A and D does not include point C, then the line segment AC forms the first compression segment, saving points A and C and discarding point B. The next compression segment begins with point D, and the process continues until all data points are traversed.

The SDT algorithm has a low computational complexity, and the compression deviation can be controlled by the threshold value E, which influences the construction of the parallelogram by determining its side length. However, once the threshold value E is set, it cannot be changed throughout the compression process, leading to poor controllability. If the data collected by the smart terminal exhibits strong periodicity, this can result in wasted computational and storage resources. To address this, this paper introduces Flow Variability Degree (FVD) to dynamically adjust the compression deviation E, proposing a data compression algorithm based on FVD-SDT. The method of adjusting the SDT compression using FVD is as follows:

- (1) Extracting Data for a Time Period and Calculating the Deviation. Let δ_i be the deviation of the i-th data point, x_i the value of the data point, and μ the mean of the dataset.

$$\delta_i = |x_i - \mu| \quad (1)$$

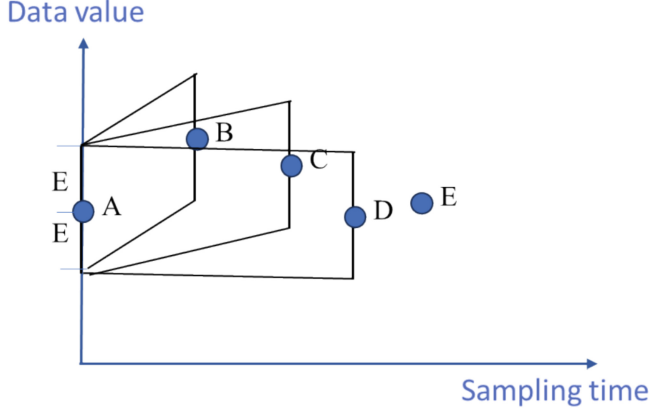


Fig. 4. Swing Door Trending compression algorithm

- (2) Calculating FVD: Let N be the total number of data points, and $\bar{\delta}$ be the mean deviation.

$$FVD = \frac{1}{N} \sum_{i=1}^N \left(\frac{\delta_i - \bar{\delta}}{\bar{\delta}} \right)^2 \quad (2)$$

- (3) Updating Compression Deviation E : Let E_0 be the base compression deviation, and α be an adjustment parameter that controls the sensitivity of E to changes in FVD.

$$E = E_0 * (1 + \alpha * FVD) \quad (3)$$

- (4) Applying the Updated E as the SDT Compression Deviation: The input data is divided into multiple segments, each containing a series of data points. Within each segment, a data point is chosen as the starting point, and a parallelogram is constructed using this point and another point. In each segment, the algorithm gradually checks whether a new parallelogram can include more data points. If a data point is found outside the current parallelogram, that point is saved, and a new compression segment begins.

The variability of data is measured by calculating the variance of the deviation of data points relative to the mean deviation. FVD reflects the extent to which individual data points deviate from average behavior within the dataset. A high FVD indicates significant variation in the data flow, requiring the compression algorithm to discard data more conservatively, thus reducing the compression deviation E . This ensures that fewer data points are discarded to capture important changes. Conversely, a low FVD indicates a relatively stable data stream with minimal variation. In this case, the compression deviation E can be increased, allowing more data points to be discarded, thereby improving the compression ratio without significantly affecting data integrity. The FVD-SDT algorithm optimizes data compression performance by combining the Swing Door Trending compression technique with an adaptive adjustment mechanism. By introducing Flow Variability Degree as a basis for dynamic adjustment, the algorithm can flexibly adjust

the compression deviation E , enhancing both compression efficiency and data quality. Compared to the original SDT algorithm, FVD-SDT performs better when handling different data characteristics, providing a more efficient and applicable solution for data compression in smart terminal environments.

4 Experimental Validation

4.1 Simulation Environment

To validate the practical application effectiveness of the proposed high-performance database storage solution for smart terminals, a series of simulation experiments were designed. Considering the background that smart terminals need to handle large volumes of sensor data, log data, and status information while operating with limited storage and computing resources, an embedded development platform was chosen as the simulation environment to mimic actual smart terminal operations.

To evaluate the performance of the FVD-SDT algorithm, experiments were designed based on three key simulation metrics: Compression Ratio (CR), Mean Compression Error (MCE), and Compression Time (CT).

CR: The compression ratio is one of the fundamental metrics for measuring the effectiveness of data compression, representing the ratio of the data size after compression to the data size before compression. A higher compression ratio indicates better compression performance.

$$CR = \frac{\text{Original Data Size}}{\text{Compressed Data Size}} \quad (4)$$

MCE: The mean compression error reflects the average deviation between the compressed and decompressed data points and the original data points, indicating the extent of information loss during compression. A smaller MCE suggests that the data compression has a minimal impact on the original data, ensuring higher data fidelity. Let x_i be the original data point, $\hat{x}_i$ be the decompressed data point, and N be the total number of data points.

$$MCE = \frac{1}{N} \sum_{i=1}^N |x_i - \hat{x}_i| \quad (5)$$

CT: Compression time refers to the time required to complete the data compression process, reflecting the computational efficiency of the compression algorithm. A shorter compression time indicates that the algorithm is more suitable for use in resource-constrained environments.

The FVD-SDT algorithm was compared with the original SDT algorithm and two other adaptive adjustment deviation (AAD) SDT algorithms. The first variant, referred to as AAD-SDT1 [11], adjusts the compression deviation E based on the comparison between the length of each compression interval and an interval length threshold (ILT). The second variant, referred to as AAD-SDT2 [12], adjusts the compression deviation E based on the comparison between the actual mean error of the compression and the expected mean error (EME). The dataset used for the experiments was an industrial sensor network dataset consisting of 15,000 data points. The initial compression deviation E was set to 5, with AAD-SDT1's ILT set to 20, and AAD-SDT2's EME set to 2.5.

4.2 Results of the Experiment

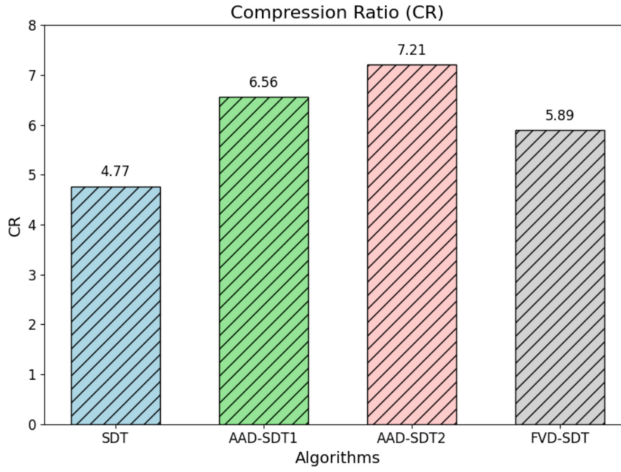


Fig. 5. Comparison of compression ratios

Figure 5 illustrates the compression ratio results of the four algorithms. It can be observed that the AAD-SDT2 algorithm exhibits the highest compression ratio, while the other algorithms also demonstrate strong performance in this metric. The FVD-SDT algorithm achieves a higher compression ratio than the original SDT algorithm due to its ability to increase compression efficiency in low flow variability scenarios.

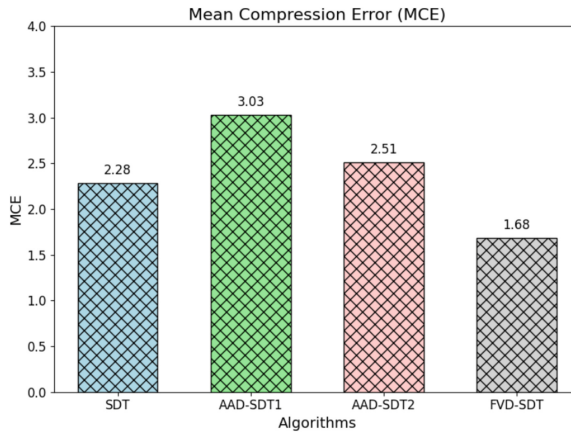


Fig. 6. Comparison of Mean Compression Error

Figure 6 shows the comparison of mean compression errors across the four algorithms. The FVD-SDT algorithm exhibits the lowest MCE, indicating that it retains

the highest data fidelity by losing the least amount of data during compression. This is because the FVD-SDT algorithm dynamically adjusts the compression deviation E , more accurately preserving the original characteristics of the data.

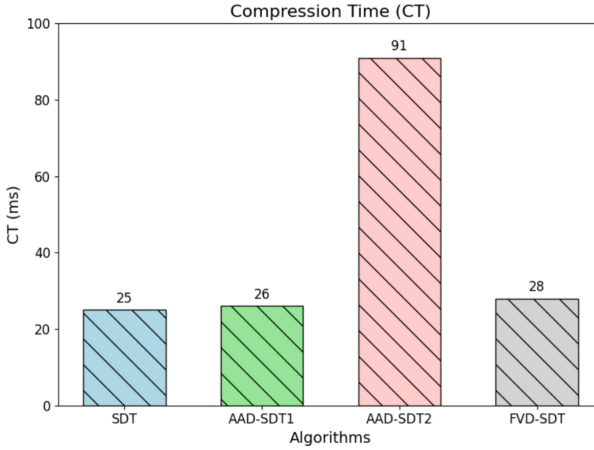


Fig. 7. Comparison of compression Time

Figure 7 presents the compression time results for the four algorithms. Although the AAD-SDT2 algorithm excels in compression ratio, its compression time is significantly higher than that of the other algorithms. The FVD-SDT algorithm, while maintaining a high compression ratio and low error, achieves a relatively short compression time, demonstrating its suitability for use in resource-constrained smart terminal environments.

By analyzing the experimental data, the FVD-SDT algorithm shows clear advantages in smart terminal data compression. In terms of compression ratio, the FVD-SDT algorithm improves by 23.5% compared to the SDT algorithm. In terms of mean compression error, the FVD-SDT algorithm reduces it by 26.3%, significantly enhancing data fidelity. Regarding compression time, the FVD-SDT algorithm only takes 3ms longer than the SDT algorithm, but its computational efficiency is far superior to the AAD-SDT2 algorithm. These results indicate that the FVD-SDT algorithm achieves a good balance between data quality and compression efficiency, making it an effective data compression technology suitable for smart terminal environments where resources are limited.

5 Conclusion

This paper addresses the application requirements of smart terminals in new power systems by proposing an efficient database storage and processing solution. We have designed a lightweight database framework tailored for smart terminals, effectively addressing the challenges of data storage and processing in resource-constrained environments. The lightweight rule engine-based data cleaning technology ensures data accuracy and reliability, meeting the demands of complex data environments within smart

terminals. Additionally, the proposed FVD-SDT data compression technology enhances storage efficiency while preserving data integrity. Experimental validation demonstrates that the proposed methods exhibit significant advantages in storage space optimization, data processing efficiency, and overall system performance, providing a solid technical foundation for the efficient operation of power smart terminals.

Acknowledgement. This Paper is Supported by Research and Development of Common Platform and Typical Application Products for Power Intelligent Terminal.

References

1. Pasricha, S.: Embedded systems education: experiences with application-driven pedagogy. *IEEE Embed. Syst. Lett.* **14**(4), 167–170 (2022)
2. Kang, W., Son, S.H., Stankovic, J.A.: Design, implementation, and evaluation of a QoS-aware real-time embedded database. *IEEE Trans. Comput.* **61**(1), 45–59 (2012)
3. Li, S., Kot, A.C.: Privacy protection of fingerprint database. *IEEE Signal Process. Lett.* **18**(2), 115–118 (2011)
4. Yafooz, W.M.S., Fahad, S.A., Abidin, S.Z.: FlexiDC: a flexible platform for database conversion. In: *Proceedings 2018 International Conference Smart Computing and Electronic Enterprise (ICSCEE)*, Shah Alam, Malaysia, pp. 1–7 (2018)
5. Wang, C., Mu, G., Cao, Y.: A method for cleaning power grid operation data based on spatiotemporal correlation constraints. *IEEE Access* **8**, 224741–224749 (2020)
6. Wang, Z., Li, S., Feng, J., Liang, Y.: A book recommendation algorithm based on data cleaning and association rules. In: *Proceedings 2021 IEEE 5th Adv. Information Technology, Electronic and Automation Control Conf. (IAEAC)*, Chongqing, China, pp. 2586–2590 (2021)
7. Wang, Q., Guo, Y., Yu, L., Chen, X., Li, P.: Deep Q-network-based feature selection for multisourced data cleaning. *IEEE Internet Things J.* **8**(21), 16153–16164 (2021)
8. Koc, B., Arnavut, Z., Kocak, H.: Lossless compression of dithered images. *IEEE Photonics J.* **5**(3), 6800508, Art. no. 6800508 (2013)
9. Zhang, X.: Lossy compression and iterative reconstruction for encrypted image. *IEEE Trans. Inf. Forensics Secur.* **6**(1), 53–58 (2011)
10. Wang, R.: SDT compression algorithm based on approximate query of sensor time series data. In: *Proceedings 2022 IEEE 5th International Conference Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China, pp. 666–670 (2022)
11. Zhang, W., Chen, X.: Research and application of improved process data compression algorithm-SDT. *Industrial Control Computer* **22**(8), 1–3 (2009)
12. Bristol, E.H.: Swinging door trending: adaptive trend recording. In: *Proceedings ISA 90th Int. Conf.*, pp. 749–753 (1990)



Quantum Key Distribution Strategy for Power Quantum Communication Networks Based on Trusted Relay Nodes

Donghai Huang^(✉), Xingnan Li, Jiajia Fu, Wenjuan Liang, and Zhongmiao Kang

Guangdong Power Dispatch and Control Center, Guangdong 510000, China
1334539052@qq.com

Abstract. Due to the theoretically unconditional security offered by quantum communication, it has broad application prospects in new power systems. However, the integration of quantum communication into these systems faces the challenge of reconciling the low quantum key generation rate with the high data transmission demands of power operations. This issue is particularly exacerbated by transmission distance limitations, necessitating the support of relay technology in the construction of quantum communication networks within power systems, which further amplifies the problem of low quantum key generation rates. To address these challenges, this paper models a quantum communication network architecture for power systems based on trusted relay nodes and designs a reinforcement learning-based quantum key distribution (QKD) routing algorithm. This algorithm aims to reduce network congestion and minimize routing hops by selecting the next node for the quantum key distribution task using reinforcement learning. Experimental results demonstrate that this algorithm can effectively improve the quantum key generation rate.

Keywords: Quantum Key Distribution · Trusted Relay Nodes · Routing

1 Introduction

Unlike traditional communication systems, which are grounded in classical physics, quantum communication leverages the unique properties of quantum mechanics to provide enhanced security and more efficient communication. As research in quantum mechanics advances, quantum communication technologies are becoming increasingly mature.

With the development of the energy internet and smart grids, the demands for information security and reliability in power systems are steadily increasing. Quantum communication offers a novel solution for the information security of new power systems. In smart grids, the dispatch center needs to maintain real-time communication with various power nodes, transmitting critical data such as power load and transmission line status. Through quantum communication technology, the dispatch center and power nodes can achieve absolutely secure communication, preventing data tampering or interception.

Additionally, quantum communication can be used for quantum authentication between devices, ensuring the authenticity of smart devices. New power systems include a large number of smart devices, such as smart meters and distributed generation equipment, which require identity authentication to prevent unauthorized device access. The application of quantum communication in new power systems has broad prospects, providing technical support for the construction of a more intelligent and reliable power network. As quantum communication technology continues to mature, it will play an increasingly important role in the future of new power systems.

Compared to traditional communication methods, quantum communication offers absolute security. It is based on theories such as quantum superposition, quantum entanglement, and the Heisenberg uncertainty principle. The most well-known form of quantum communication is QKD, which ensures that the keys generated during transmission are secure. A QKD channel typically includes both a quantum channel and a classical channel. The quantum channel is used to transmit quantum signals, which are quantum states encoded with classical information. The classical channel, on the other hand, is responsible for transmitting classical signals to synchronize and negotiate keys between the sender and receiver. If an eavesdropper attempts to intercept part of the quantum states from the quantum channel, these quantum states will not be used for key distribution because the receiver will not receive them. Furthermore, even if the eavesdropper tries to measure and replicate these quantum states before sending them to the receiver, the quantum no-cloning theorem ensures that the replicated quantum states will inevitably change, resulting in significant errors. Therefore, any potential eavesdropping on the QKD process can be detected.

This paper proposes a quantum key distribution strategy for power quantum communication networks based on trusted relay nodes. To address the current challenge where the quantum key generation rate fails to meet the encryption needs of power operations, we first model a power quantum communication network based on trusted relay nodes, considering the priority of power operations to schedule power tasks and prioritize high-priority tasks. We then design a reinforcement learning-based quantum communication routing strategy to reduce congestion caused by slow quantum key generation. The method proposed in this paper provides a comprehensive solution for applying quantum communication to new power systems.

2 Related Works

The fundamental principles of quantum mechanics ensure the security of quantum communication networks, characterized by theoretically “unconditional security” [1], and have become a focus of research in network security in recent years. Among these, quantum key distribution (QKD) technology, an important branch of quantum secure communication, has garnered widespread attention. The BB84 and MDI protocols are two classic QKD protocols. The BB84 protocol [2], proposed by Charles Bennett and Gilles Brassard in 1984, transmits keys using four possible photon polarization states. By leveraging the uncertainty principle of quantum mechanics, it ensures that any eavesdropping attempts will result in detectable errors, thus securing communication. The MDI protocol [3] is an improved quantum key distribution protocol where two communicating parties send photons to a third-party relay station for interference measurement. The

communicating parties establish a secure key based on the measurement results. Even if the relay station is controlled by an eavesdropper, the MDI protocol still ensures the security of the key, thereby enhancing overall communication security.

In recent years, some studies have focused on improvements to QKD protocols. Chen et al. [4] applied the “send or not send” twin-field (TF) QKD protocol over 509 km of ultra-low-loss fiber. Additionally, some research has focused on network-related studies in the field of quantum communication. Farahbakhsh et al. [5] proposed an opportunistic quantum routing algorithm, which, unlike traditional algorithms, allows requests to advance even when some resources on the path are not yet ready, thereby managing entanglement resources more efficiently. Zeng et al. [6] studied the design issues of multi-routing entanglement protocols in quantum networks, aiming to simultaneously maximize the number of quantum user pairs and their expected throughput.

Quantum communication technology has broad application prospects in power systems. Yu et al. [7] proposed an improved weighted fair queue (WFQ) algorithm, LD-WFQ, which estimates the expected delay of packets and prioritizes the transmission of packets nearing timeout. This approach effectively reduces the packet timeout rate, improves the effective utilization of quantum keys, and maintains a low packet timeout rate in high-priority queues. Wu et al. [8] proposed a quantum key distribution system based on wireless channels, aiming to address the limitations of traditional fiber-optic channel key distribution methods in terms of fiber resources and construction costs, particularly to meet the security enhancement needs of numerous mobile terminals.

3 Power Line Carrier and Wireless Dual-Mode Integrated Communication Method

In the field of quantum communication, trusted relay systems can enhance the transmission distance of quantum keys and improve their scalability. A trusted relay is a fully trusted relay node whose security is strictly ensured. To guarantee the absolute security of a trusted relay, it is typically necessary to deploy the relay in a guarded location or use physical isolation measures to defend against potential attacks. Trusted relays are responsible for storing and transmitting keys in a QKD network, utilizing their security to encrypt and decrypt keys, thereby ensuring the overall security of the network's keys. For example, as shown in Fig. 1, QKD node A and the trusted relay node generate quantum key pairs K_A and K_B respectively with the trusted relay node and QKD node B. Long-distance quantum key distribution in a QKD network with trusted relays involves the following steps:

1. The trusted relay node encrypts key K_A using key K_B , resulting in the ciphertext $K_A \oplus K_B$;
2. The trusted relay node sends the ciphertext $K_A \oplus K_B$ to QKD node B;
3. QKD node B decrypts the ciphertext $K_A \oplus K_B$ to obtain $K_B \oplus (K_A \oplus K_B)$, resulting in key K_A .

By placing several trusted relays, the communication distance between two QKD nodes can theoretically be extended indefinitely, provided that all trusted relay nodes along the line are secure.

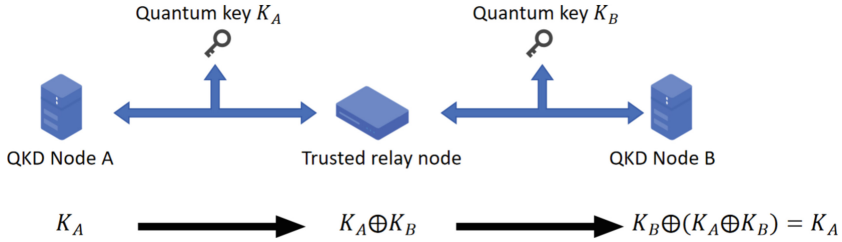


Fig. 1. Quantum Key Distribution Based on Trusted Relay Nodes

3.1 System Architecture

The architecture of a power system communication network based on quantum communication is as illustrated. The system mainly includes QKD nodes and trusted relay nodes. QKD nodes encompass smart meters, substation control devices, dispatch centers, and other components. In long-distance communication, trusted relay nodes are used to relay and re-encrypt keys, ensuring the security of the keys throughout the entire transmission chain.

Given the limited generation rate of quantum keys, a quantum key pool is maintained between QKD nodes, between QKD nodes and relay nodes, and between relay nodes. During quantum key distribution, nodes will consume keys stored in the quantum key pool. The primary function of the quantum key pool is to centrally manage and store keys generated through QKD technology, allowing quick access and usage by nodes in the network when needed, thus performing a “peak shaving and valley filling” role. To prevent resource wastage, a storage limit is set for each quantum key pool to avoid excessive generation of quantum keys. If the quantum key generation rate between QKD nodes m and n is k_{mn} , then the storage limit for the quantum key pool between QKD nodes m and n is set to uk_{mn} (Fig. 2).

The system network can be represented as a directed graph $G = (V, E)$, where V denotes the set of nodes, which include both QKD nodes and relay nodes, and E are links between nodes. Each QKD node may have multiple tasks arriving in a short period; therefore, a priority scheduling strategy can be used to manage these tasks. This strategy orders tasks based on their priority, with higher-priority tasks being processed first. This approach is simple and effective, suitable for scenarios requiring rapid response and the assurance of critical task completion.

Firstly, each task needs to be assigned a priority. Priority is determined based on factors such as the type of task and its urgency. For example, in a power system, operations such as fault recovery and safety protection have higher priorities.

Currently, system operations in power systems can be classified into four security zones [9]: Security Zone I, II, III, and IV. Security Zones I and II are collectively referred to as the Production Control Zone, while Security Zones III and IV are collectively referred to as the Management Information Zone. Security Zone I includes operations related to line relay protection, safety and stability systems, dispatch automation, and dispatch telephones. Security Zone II includes operations related to protection management information systems, safety and stability management information systems, and electric energy measurement and telemetry systems. Security Zone III includes operations

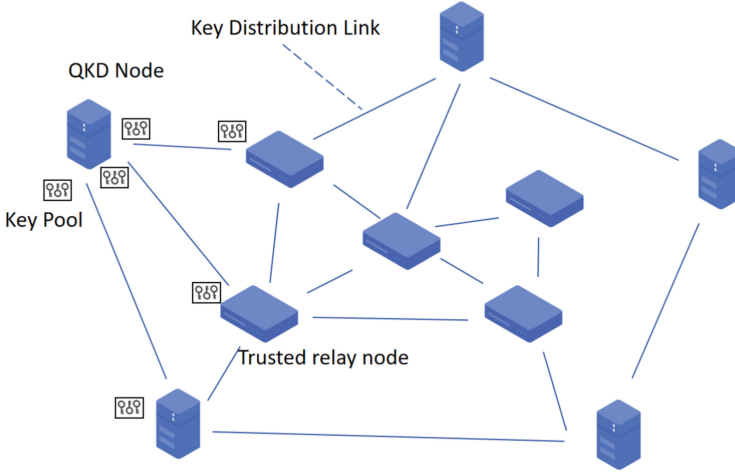


Fig. 2. System Architecture

related to dispatch management information systems, lightning positioning monitoring systems, substations, and video surveillance systems. Security Zone IV includes operations related to video conferencing systems, administrative telephones, the Internet, and disaster recovery backup systems.

In our proposed method, each QKD node maintains a separate task queue for each of the four security zones. When a task arrives at a QKD node, it is inserted only into the queue corresponding to its respective security zone. Within each queue, tasks are ordered in ascending order of delay requirements, with tasks requiring lower delays being placed at the front of the queue.

3.2 Routing Strategy Based on Reinforcement Learning

Quantum key transmission requires the involvement of multiple relay nodes. When the keys on the links in the transmission route are insufficient to meet the key consumption needs, it can lead to blockage in the key transmission process. To ensure smooth quantum key transmission, the routing selection for quantum key transmission should favor links with adequate key reserves or higher key generation rates.

To address the routing selection problem using reinforcement learning, we model it as a Markov Decision Process (MDP) and solve it using reinforcement learning algorithms. An MDP is a Markov process that includes rewards and decisions and can be represented by a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is the action space, $\mathcal{P}$ is the state transition function, $\mathcal{R}$ is the reward function, and γ is the discount factor. The probability of transitioning from state s_t to state s_{t+1} after taking action a_t is denoted by $P(s_{t+1}|s_t, a_t)$. The immediate reward obtained from taking action a_t in state s_t and transitioning to state s_{t+1} is denoted by $R(s_t, a_t, s_{t+1})$.

Let N denote the total number of nodes in the system, and divide the nodes into three sets: L_s, L_d , and L_n , representing the sets of source QKD nodes, destination QKD nodes, and relay nodes, respectively. Let N_s, N_d and N_n denote the numbers of nodes

in these sets. All nodes participate in the key transmission process, and whether the transmission process is blocked depends on the key reserves and key generation rates on each link. Let $D_t = \{D_{1,t}, D_{2,t}, \dots, D_{N,t}\}$ denote the total amount of keys to be transmitted by each node at time slot, and $B_t = \{B_{1,2,t}, B_{1,3,t}, \dots, B_{N-1,N,t}\}$ denote the remaining key pool size on each link at time slot. Let $q_t = \{q_{1,2}, q_{1,3}, \dots, q_{N-1,N}\}$ denote the key generation rate for each link, and $s_t = \{s_{1,2}, s_{1,3}, \dots, s_{N-1,N}\}$ denote the key consumption rate for each link. Source QKD nodes generate new quantum key transmission demands at each time slot, and the amount of keys required for transmission is related to the corresponding task data volume, which affects the routing choice for key transmission. Let $V_t = \{V_{1,t}, V_{2,t}, \dots, V_{N,t}\}$ denote the new quantum key transmission demand at time slot, assuming the data generation process follows a Poisson process. Thus, the joint state of the entire network at time slot is defined as $(D_t, B_t, V_t, q_t, s_t)$.

For each key transmission task, the key data to be transmitted to the next node may be distributed across different nodes, and the size and location of this data will also affect the choice of the next transmission node. Let P_t denote the size and location of the key data to be transmitted at time slot t . $P_{i,t}$ represents the amount of key data to be transmitted at node i at time slot t . Therefore, the state for each key transmission task can be represented as $S_t = (D_t, B_t, V_t, q_t, s_t, P_t)$.

The data transmission task selects actions based on the network state and the location and size of data packets, specifically choosing the link between the current router and the next-hop router. Not all links are available for selection; only those connected to the location of the current data packet can be chosen as valid actions. Thus, during the data transmission process, the set of valid actions changes with the movement of the data packet location.

The reward function is used to quantify the value of performing a given action in a given state. In quantum key transmission routing, the focus of the reward function design is on avoiding network congestion rather than simply seeking the shortest path. Therefore, the reward function should be designed to encourage congestion-avoidance behaviour while optimising path length without causing congestion. The reward function should penalise actions that could lead to congestion and inefficient actions (such as selecting congested links), while also considering path length to improve network efficiency and reduce resource consumption. The reward function $R(s_t, a_t, s_{t+1})$ should be set as follows:

$$R(s_t, a_t, s_{t+1}) = \begin{cases} r_c, & \text{when congestion occurs} \\ r_e, & \text{when } a_t \text{ is invalid action} \\ 0, & \text{when key transfer task completed} \\ -1, & \text{others} \end{cases} \quad (1)$$

In this context, r_c and r_e are large negative values used to penalize the occurrence of congestion and the selection of ineffective actions, respectively. The value -1 is used to record the number of hops that data packets make within the network.

A policy π is a mapping from states to actions. $V^\pi(s)$ represents the expected total return starting from state s and following policy π ; and $Q^\pi(s, a)$ represents the expected total return from taking action a in state s and then following policy π .

The Bellman equation provides a recursive representation of the value functions. For the state value function $V^\pi(s)$:

$$V^\pi(s) = \sum_{a \in A} \pi(a|s) \sum_{s' \in S} P(s'|s, a) [R(s, a, s') + \gamma V^\pi(s')] \quad (2)$$

For the action value function $Q^\pi(s, a)$:

$$Q^\pi(s, a) = \sum_{s' \in S} P(s'|s, a) [R(s, a, s') + \gamma \sum_{a' \in A} \pi(a'|s') Q^\pi(s', a')] \quad (3)$$

In reinforcement learning, the goal is to find an optimal policy π^* that provides the maximum expected return for each state. Correspondingly, there are optimal state value function $V^*(s)$ and optimal action value function $Q^*(s, a)$. These functions satisfy the following Bellman optimality equations:

$$V^*(s) = \max_{a \in A} \sum_{s' \in S} P(s'|s, a) [R(s, a, s') + \gamma V^*(s')] \quad (4)$$

$$Q^*(s, a) = \sum_{s' \in S} P(s'|s, a) \left[R(s, a, s') + \gamma \max_{a' \in A} Q^*(s', a') \right] \quad (5)$$

Methods such as policy iteration can be used to solve the Bellman equations and obtain the optimal policy π^* .

The Bellman optimality equation is used to update the policy:

$$\pi' = \operatorname{argmax}_{a \in A} \sum_{s' \in S} P(s'|s, a) [R(s, a, s') + \gamma V^\pi(s')] \quad (6)$$

Policy evaluation and policy improvement steps are repeated until the policy converges.

In this paper, Deep Q-Networks (DQN) are used as the model for deep reinforcement learning.

4 Experiments

In the experimental phase of this study, the experiments were conducted using PyTorch to set up the experimental environment. The hardware used included an NVIDIA GTX 3080 GPU, and the operating system was Ubuntu 16.04. The network topology configured for the experiments is illustrated in the figure, consisting of 11 QKD nodes and 7 trusted relay nodes. Quantum key distribution tasks generated by each QKD node in the experiment followed a Poisson process, with subsequent adjustments made to this parameter to evaluate the method's performance under different conditions. The parameter unit is MB/s. The quantum key generation rate on each link followed a normal distribution with parameters (1, 0.5), and the quantum key consumption rate on each link followed a normal distribution with parameters (1, 0.8). The discount factor for the reinforcement learning was set to 0.9. Each link's key pool could store a maximum of keys generated over 20 time units, with initial key pools set to full capacity.

The shortest path algorithm was chosen for comparison, with the blocking probability and quantum key throughput of the system used as reference indicators to validate the

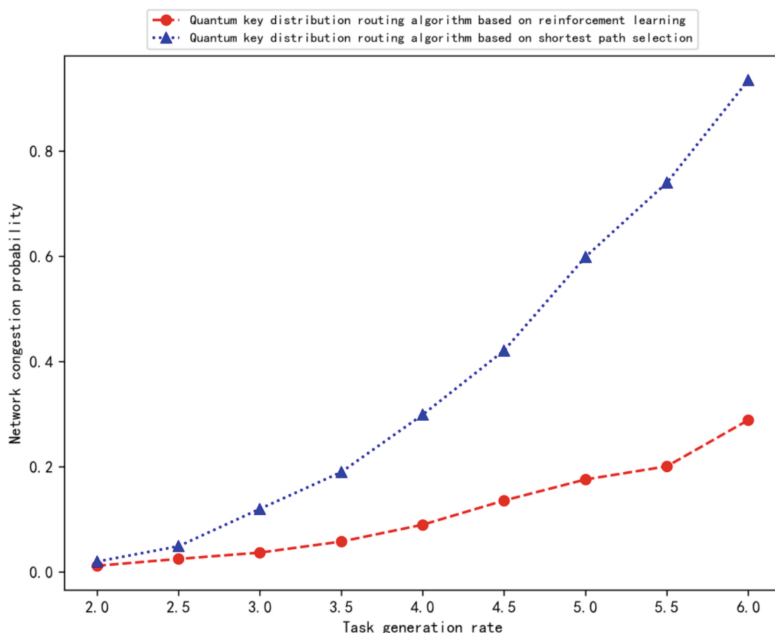


Fig. 3. Network blocking rate comparison

performance of the proposed method. The quantum key throughput here represents the average quantum key throughput for quantum key distribution tasks completed per unit time in the system. The unit for quantum key throughput is MB/s (Fig. 3).

We first compared the network blocking probabilities of the two algorithms under different task generation rates. Network blocking occurs when the keys generated on the chosen links cannot support the key consumption. It was observed that as the task generation rate increased, the probability of network blocking also increased for both algorithms. The slope of the curves generally exhibited an upward trend, indicating that network blocking worsens with increasing task generation rates. Compared to the shortest path routing algorithm, the proposed reinforcement learning-based routing algorithm significantly reduced the probability of network blocking and effectively delayed the worsening of network congestion (Fig. 4).

By comparing the system quantum key throughput under different task generation rates for the two algorithms, we found that the system quantum key throughput increased with the task generation rate for both algorithms. However, the slope of the curves decreased, due to the increased probability of network blocking with higher task generation rates. Compared to the shortest path routing algorithm, the proposed method showed a significant advantage in system quantum key throughput, with this advantage becoming more pronounced as the task generation rate increased.

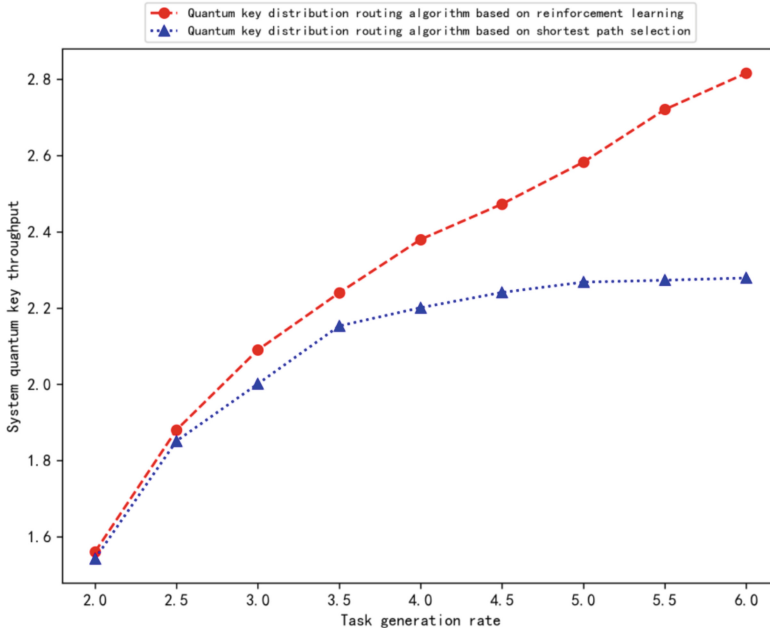


Fig. 4. System quantum key throughput comparison

5 Conclusion

The advancement of quantum mechanics poses new challenges for data security in power systems, necessitating the development of quantum communication networks to protect data security in power systems. Due to the extensive coverage area of power networks, the inclusion of relay nodes is required in the construction of power quantum communication networks, increasing the complexity of network topology and the quantum key distribution process. This paper first models the power quantum communication network based on trusted relay nodes and establishes a task scheduling strategy. It then proposes a reinforcement learning-based quantum key distribution routing algorithm that considers the overall network situation and the attributes of current quantum key distribution tasks when selecting routing nodes. Experiments demonstrate that this algorithm effectively reduces network blocking and improves quantum key distribution rates.

Acknowledgement. This Paper is Supported by Research on Key Technologies of High Security Protection of Basic Information Facilities for New Power System (GDKJXM20230420).

References

1. Gisin, N., Ribordy, G., Tittel, W., et al.: Quantum cryptography. *Rev. Mod. Phys.* **74**(1), 145 (2002)
2. Bennett, C.H.: Quantum cryptography: Public key distribution and coin tossing. In: *Proc of IEEE International Conference on Computers*. Institute of Electrical and Electronics Engineers (1984). <https://doi.org/10.1016/j.tcs.2011.08.039>

3. Lo, H.K., Curty, M., Qi, B.: Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **108**(13), 130503 (2012)
4. Chen, J.P., Zhang, C., Liu, Y., et al.: Sending-or-not-sending with independent lasers: secure twin-field quantum key distribution over 509 km. *Phys. Rev. Lett.* **124**(7), 070501 (2020)
5. Farahbakhsh, A., Feng, C.: Opportunistic routing in quantum networks. In: *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, pp. 490–499 (2022)
6. Zeng, Y., Zhang, J., Liu, J., et al.: Multi-entanglement routing design over quantum networks. In: *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. IEEE, pp. 510–519 (2022)
7. Yu, Z., Zhang, Y., Han, B., Chen, W., Wang, D.: Power services scheduling algorithm based on quantum communication. *Telecommunications Science* **35**(11), 108–116 (2019)
8. Wu, J., Bian, Y., Zhang, Q., Feng, B.: Wireless quantum power distribution system based on wireless. *Telecommunications Science* **36**(2), 68–75 (2020)
9. Li, W., Ma, L., Sun, X., et al.: Research on combination traffic forecasting method based on power grid IMS platform framework. *Electric Power Information and Communication Technol.* **16**(3), 43–49 (2018)



Research on Network Traffic Anomaly Detection Method Integrating Graph Topology Optimization and Random Restart Walk Algorithm

Wen Luo^(✉), Haotian Li, Kunxiang Niu, and Longfei Yao

Nanjing University of Science and Technology, Nanjing 210094, China
xlangyu@njjust.edu.cn

Abstract. Due to the high sparsity and complexity of network traffic, the detection effect of existing network traffic anomaly detection methods was still very limited, and it was difficult to effectively utilize the topological information and traffic characteristics in the network. To solve this problem, a sampling optimization method based on the graph neural network E-GraphSAGE that combines graph topology optimization and random walk restart algorithm was proposed, which effectively utilized network topology and traffic characteristics for network traffic anomaly detection. The graph topology optimization strategy improved the performance of the graph by transforming the graph structure, so that the RWR algorithm could be used for sampling more effectively, and then the sampling results were aggregated to obtain the final edge features for training and detection models. Experiments on three common network anomaly traffic datasets was conducted. The experimental results showed that the recall rate, precision, and F1 score of the proposed method on all datasets exceeded 99%. In particular, on the ISCXTor2016 dataset, not only the evaluation indicators had reached the most advanced level, but also the recall rate, precision, and F1 score of the proposed method had been improved by 19%, 17%, and 19% respectively compared with the most advanced graph neural network model. The experimental results prove the effectiveness and robustness of the proposed method for network traffic anomaly detection tasks.

Keywords: Graph neural network · network traffic anomaly detection · RWR algorithm · graph topology optimization

1 Introduction

With the rapid development of internet technology, network traffic data is experiencing explosive growth, with malicious traffic from attackers mixed in, posing significant threats to cyberspace security. Network traffic anomaly detection, as an essential part of network security, can effectively identify and prevent various network attacks by analyzing network traffic data [1–3]. However, the data from network intrusion behavior often exhibit high sparsity and complexity, making traditional methods based on predefined rules and feature engineering ineffective when handling large-scale traffic data.

The challenges in network traffic anomaly detection mainly stem from the nature of the data itself. First, network traffic data is highly sparse, with only a small fraction of the data related to malicious activity. This sparsity complicates the detection of anomalies since traditional methods struggle to effectively differentiate between normal and malicious traffic in such conditions [4, 5]. Second, the complex topology of the network introduces difficulties in feature extraction, as the interactions between nodes and traffic patterns are not easily captured using standard machine learning approaches. Lastly, the high-dimensional nature of network traffic data requires models capable of handling non-Euclidean data structures, such as graphs, which are more suitable for representing relationships in traffic networks. Researchers have started exploring the use of graph neural networks (GNN) for network traffic anomaly detection [6, 7]. GNNs model the relationships between nodes and edges in graph structures, enabling them to capture higher-order feature relationships such as network topology through message passing and feature aggregation within graph structures [8].

To address these challenges, we propose a sampling optimization method based on the graph neural network E-GraphSAGE model, which is particularly designed to handle non-Euclidean graph structures in network traffic data. The E-GraphSAGE model, through its inductive learning framework, can efficiently generate edge embeddings, allowing for the integration of both node features and topological information. By combining graph topology optimization and the random walk with restart (RWR) algorithm, our method enhances sampling quality and improves the model's ability to capture complex relationships in traffic data, leading to more accurate network traffic anomaly detection. The main contributions of this research are as follows:

- 1) Graph Topology Optimization Method: We designed a graph topology optimization technique to address the high sparsity and complexity of network traffic. This technique improves graph connectivity through structural modifications, enhancing the efficiency and effectiveness of the RWR algorithm for more uniform and comprehensive sampling.
- 2) RWR-Based Sampling Optimization Method: To overcome the challenges of sampling complex correlations in network traffic, we propose an optimization method based on the RWR algorithm, replacing E-GraphSAGE's random uniform sampling with importance sampling to improve sampling quality and achieve efficient capture and representation of network traffic data.
- 3) Integration Strategy: By combining graph topology optimization and the RWR algorithm, we propose an integration strategy tailored for network traffic anomaly detection. This strategy synergizes graph structure modifications with efficient sampling, enabling better handling of large-scale, complex network traffic data and enhancing model detection performance and robustness.

The structure of this paper is organized as follows: Sect. 1 reviews the research status of network traffic anomaly detection methods. Section 2 introduces the proposed network traffic anomaly detection method based on improved graph neural networks. Section 3 presents the construction of graph neural networks tailored for network traffic. Section 4 discusses the experiments conducted on datasets and analyzes the experimental results. Finally, Sect. 5 concludes the study.

2 Related Work

2.1 Traditional Neural Network-Based Methods

Traditional neural network methods have been widely applied in network traffic anomaly detection. These methods typically rely on feature engineering, where key features of network traffic are manually extracted and selected to build detection models.

For example, Sinha et al. [9] proposed a deep learning model combining the advantages of convolutional neural networks (CNN) and BiLSTM to simultaneously learn spatial and temporal features of network traffic data. Feng et al. [10] introduced a method that effectively detects encrypted hijacking traffic by only accessing anonymous, content-agnostic network traffic metadata from the network gateway. Wu et al. [11] proposed a hierarchical CNN + RNN neural network, LuNet, considering the spatio-temporal features in network traffic data. Mirsky et al. [12] utilized multiple autoencoder models and ensemble learning techniques to learn network traffic behavior patterns for detecting network intrusion events.

2.2 Graph Neural Network-Based Methods

Recently, graph neural networks have gained widespread attention in network traffic anomaly detection. GNNs capture high-order feature relationships through feature propagation and aggregation in graph structures, making them suitable for processing network traffic data.

Kipf et al. [6] first introduced the concept of Graph Convolutional Networks (GCN). Lin et al. [13] proposed a GNN-based method for P2P botnet detection. Cheng et al. [14] introduced an Alert-GCN framework that inputs network traffic structure graphs into a two-layer GCN for node classification. Zhang et al. [15] proposed a new network traffic anomaly detection method to address the nonlinear and high-dimensional characteristics of Industrial IoT data. Nguyen et al. [16] introduced a Traffic-Aware Self-Supervised Learning (TS-IDS) framework for network traffic anomaly detection.

Several GCN variants have also been developed. Veličković et al. [17] proposed Graph Attention Networks (GAT). Hamilton et al. [18] introduced GraphSAGE, a general inductive framework that effectively generates node embeddings for unseen data using node feature information. Wai Weng Lo et al. [19] extended GraphSAGE to the network traffic detection problem and proposed the E-GraphSAGE algorithm, which captures edge features and topological information of graphs for network traffic detection. Caville et al. [20] proposed Anomal-E, a self-supervised GNN-based method for network traffic anomaly detection. Liu et al. [21] introduced a GraphSAGE-based Dynamic Spatial-Temporal Graph Convolutional Network (DST-GraphSAGE) for traffic prediction. Chang et al. [22] improved the E-GraphSAGE and GAT architectures by integrating residual learning into GNNs to enhance performance in handling highly imbalanced data.

3 Network Traffic Anomaly Detection Method Based on Improved Graph Neural Networks

In network traffic anomaly detection, traditional detection methods rely heavily on feature engineering and predefined rules, which are inadequate for handling the high dimensionality, sparsity, and complexity of network traffic data. While GNNs have shown significant advantages in processing complex graph data, existing GNN models still face challenges in sampling and feature extraction from network traffic data. Therefore, this paper proposes a network traffic anomaly detection method that integrates graph topology optimization and the RWR algorithm [23], improving the E-GraphSAGE model. The graph topology optimization technique enhances the performance of network graphs, and the RWR algorithm is used for sampling, allowing for better capture of topological information and traffic characteristics in network traffic. The aim is to improve the effectiveness of the model for network traffic anomaly detection tasks.

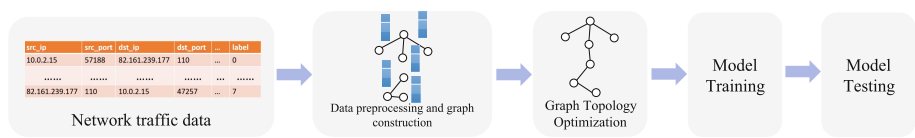


Fig. 1. Network traffic anomaly detection method flow

3.1 Graph Neural Network-Based Methods

The process of the network traffic anomaly detection method in this paper is shown in Fig. 1, which mainly includes the following steps:

- 1) Data Preprocessing and Graph Construction: Network traffic data is standardized through one-hot encoding and normalization, followed by modeling to construct a graph.
- 2) Graph Topology Optimization: A graph topology optimization technique is introduced to enhance the connectivity of the network graph, providing a strong foundation for subsequent model training.
- 3) Model Training: Edge embeddings are obtained by sampling and aggregating network traffic edges, followed by supervised learning on the training set.
- 4) Model Testing: The model’s effectiveness is tested on network traffic test sets, and performance is evaluated and compared.

The focus of this method is the model training process, as illustrated in Fig. 2.

3.2 Data Preprocessing for Network Traffic

In network traffic, many features are symbolic data, which need to be processed into numerical data through one-hot encoding, and then normalized to ensure that all features contribute equally to the analysis.

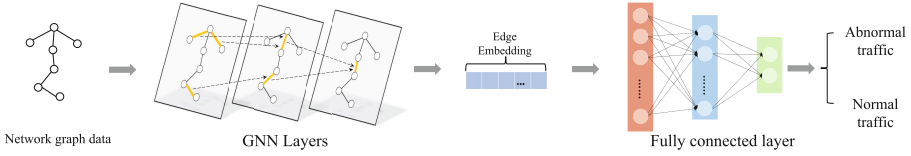


Fig. 2. Model training process

Common normalization methods include Min-Max Normalization and Standardization. In this paper, we employ Min-Max Normalization to scale the numerical data into a specific range, typically $[0, 1]$. The transformation formula for Min-Max Normalization can be expressed as Formula (1):

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

for each numerical feature, the minimum value $x_{\min}$ and the maximum value $x_{\max}$ are calculated. The original value x is then normalized to x' using the Min-Max Normalization formula mentioned earlier.

Network traffic data is primarily composed of three main components: source address, destination address, and various flow features. By using the source address and destination address as endpoints in a graph, we can naturally construct a graph representation of the network traffic. This method is not only intuitive but also efficient, as it leverages the inherent structure of the network traffic data.

3.3 Data Preprocessing for Network Traffic

Considering that the RWR algorithm requires performing random walks on a complete graph, and given the high sparsity of network traffic, the performance of the RWR algorithm can be somewhat limited when applied to sparse graphs. To address this limitation, this paper introduces a graph topology optimization technique that enhances graph connectivity by connecting nodes with the same IP address but different ports.

The process of connecting nodes with the same IP but different ports can be mathematically represented by Formula (2):

$$E' = \{e_{i,i+1} | (v_i, v_{i+1}), \forall i \in \{1, 2, \dots, n-1\}\} \cup \{e_{n,1} = (v_n, v_1)\} \quad (2)$$

where $\{v_1, v_2, \dots, v_n\}$ is a sequence containing n nodes. For each node v_i in the sequence ($i = 1, 2, \dots, n-1$), an edge $e_{i,i+1}$ is added between nodes v_i and v_{i+1} . Additionally, an edge is added between the last node v_n and the first node v_1 , forming a closed loop.

This method effectively enhances the connectivity of the graph without significantly increasing its complexity, providing a more uniform sampling basis for the RWR algorithm. The results, as shown in Table 1, clearly illustrate the change in graph connectivity before and after the optimization. This enhancement ensures that the RWR algorithm can operate more efficiently, leading to better performance in network traffic anomaly detection.

Table 1. Comparison of connectivity before and after

	Before optimization	After optimization
ISCXTor2016	982	1
UNSW-NB15	76050	7
ToN-IoT	920	6

3.4 Model Training and Testing

After applying the graph topology optimization to the network traffic graph, the next step involves sampling and aggregating the edges in the network traffic to obtain edge embeddings. These edge embeddings serve as the input for the supervised learning process, where the model is trained on the training dataset.

Once the model is trained, it is then tested on the network traffic test set to evaluate its effectiveness. The performance of the model is assessed through various metrics, and a comparative analysis is conducted to benchmark the model against other existing methods.

This process ensures that the model is capable of accurately detecting anomalies in network traffic, leveraging the enhanced connectivity from the graph topology optimization and the efficient sampling strategy provided by the RWR algorithm.

4 Graph Neural Network Construction for Network Traffic

4.1 Application of RWR Algorithm in Network Traffic

The RWR algorithm is an effective method for computing the weighted scores of connections between nodes in a constructed graph model. Its application in network traffic analysis helps the model better understand the relationships between nodes, thereby optimizing model performance.

Mathematically, the RWR algorithm at node v can be represented by the following iterative formula:

$$\mathbf{P}_v^{(t+1)} = (1 - c)\mathbf{A}\mathbf{P}_v^{(t)} + c\mathbf{r} \quad (3)$$

where $\mathbf{P}_v^{(t)}$ is the node visitation probability vector at time t , c is the restart probability, $\mathbf{r}$ is the restart vector, which is equal to $\mathbf{P}_v^{(0)}$ (typically initialized with 1 at the starting node position and 0 elsewhere), and $\mathbf{A}$ is the graph transition probability matrix, whose elements A_{ij} are defined as follows, where d_i is the degree of node i :

$$A_{ij} = \begin{cases} \frac{1}{d_i}, & \text{if } d_i > 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

As the number of iterations increases, $P_v^{(t)}$ converges to a stable probability distribution, reflecting the relative importance or similarity of nodes in the graph.

Through the introduction of the RWR algorithm, the model can more comprehensively cover the graph structure, efficiently capturing relationships between nodes, reducing sampling bias, improving feature representation quality, and enhancing model performance in large-scale graph data sampling processes.

4.2 E-GraphSAGE Algorithm

In the E-GraphSAGE algorithm[19], the sampling method samples a fixed-sized subset of neighboring nodes to propagate neighbor messages. The goal of neighborhood sampling is to minimize the size of the computational graph to reduce the algorithm's space and time complexity. The E-GraphSAGE algorithm performs graph convolution operations on the graph $G(V, E)$, with nodes represented as V and edges represented as E , and node features represented as x_v .

To specify the depth of the GNN, the K -hop neighborhood parameter is set, indicating the K -hop neighbor edge information that needs to be aggregated in each iteration. Another important aspect of the GNN is selecting a differentiable aggregation function AGG_k ($k \in \{1, 2, \dots, K\}$) for aggregating K -hop neighbor edge information. Various aggregation methods, including mean, pooling, or different types of neural networks, can be applied in E-GraphSAGE.

For each node, the E-GraphSAGE model iteratively aggregates K -hop depth adjacent edge information. In each iteration, a fixed number of neighboring nodes can be sampled to reduce the algorithm's space and time complexity. The edge embeddings between the sampled nodes and the source node are then used to generate the next layer of node embeddings. At the k -th layer ($k \in \{1, 2, \dots, K\}$), the aggregated node neighbor information $h_{N(v)}^k$ at node v can be represented by the following Formula (5):

$$h_{N(v)}^k = AGG_k(\{e_{uv}^{k-1}, \forall u \in \mathcal{N}(v), uv \in \mathcal{E}\}) \quad (5)$$

The edge embedding e_{uv}^{k-1} represents the edge embedding of the edge uv in the previous layer. These adjacent edge embeddings are aggregated into the embedding of node v at the k -th layer, and the K -hop neighborhood edge features of each graph node are aggregated based on the aggregation function. The aggregated representation from the sampled neighbors is then concatenated with the node representation h_v^{k-1} of itself. After that, the model weights W_k can be applied and the result can be passed to a nonlinear activation function σ (such as ReLU) to obtain the final node embedding h_v^k , as shown in Formula (6):

$$h_v^k = \sigma(W_k \cdot \text{CONCAT}(h_v^{k-1}, h_{N(v)}^k)) \quad (6)$$

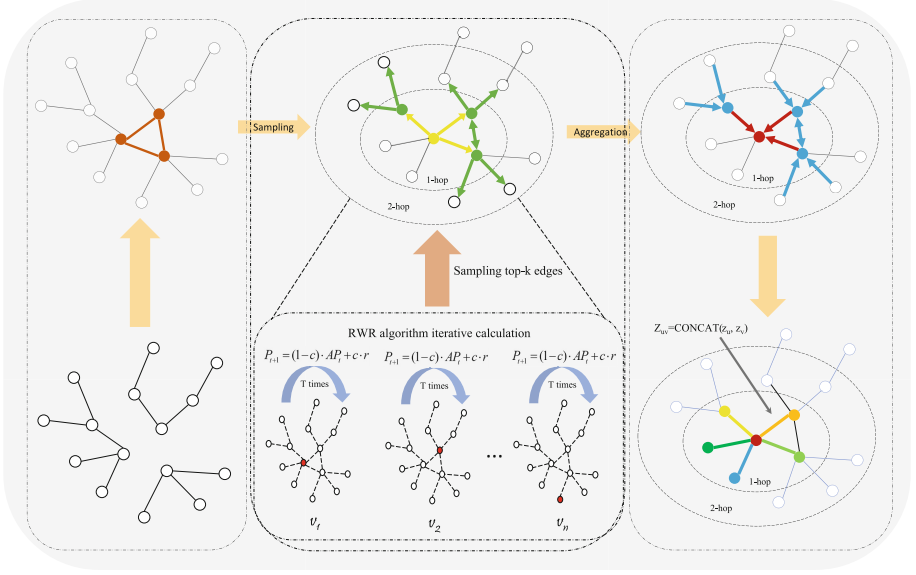


Fig. 3. Model framework

Algorithm 1 E-GraphSAGE Integrating Graph Topology Optimization and RWR Algorithm

input: Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$; Edge Features $\{e_{uv}, \forall uv \in \mathcal{E}\}$; Node Features $\mathbf{x}_v = \{1, \dots, 1\}$; depth K ; Weighted Matrix $\mathbf{W}^k, \forall k \in \{1, \dots, K\}$; Sample Number N ; Non-linear activation function σ ; Aggregation Function AGG_k ; The length of the restart vector $\mathbf{e}_v$ is $|\mathcal{V}|$ and $\mathbf{e}_v = \{1 \text{ if } u=v, 0 \text{ otherwise}, \forall u \in \mathcal{V}, \forall v \in \mathcal{V}\}$; Restart Probability c ; Transfer Matrix A ; Iteration Limit T ;

output: Edge Embedding $\mathbf{z}_{uv}, \forall uv \in \mathcal{E}$

```

1   $\mathbf{h}_v^0 \leftarrow \mathbf{x}_v, \forall v \in \mathcal{V}$ 
2  for  $v \in \mathcal{V}$  do
3      // RWR algorithm iteration process
4       $\mathbf{P}_v^{(0)} \leftarrow \mathbf{e}_v$ 
5       $t \leftarrow 0$ 
6      do
7           $\mathbf{P}_v^{(t+1)} \leftarrow (1-c) \mathbf{A} \mathbf{P}_v^{(t)} + c \mathbf{P}_v^{(0)}$ 
8           $t \leftarrow t+1$ 
9      while  $\mathbf{P}_v^{(t)}$  not converged or  $t \leq T$ 
10      $\mathcal{N}_{RWR}(v) \leftarrow \text{Select TOP-}N \text{ nodes from } \mathbf{P}_v^{(t)}$ 
11     for  $k \leftarrow 1$  to  $K$  do
12         for  $v \in \mathcal{V}$  do
13              $\mathbf{h}_{\mathcal{N}_{RWR}(v)}^k \leftarrow AGG_k(\{e_{uv}^{k-1}, \forall u \in \mathcal{N}_{RWR}(v), uv \in \mathcal{E}\})$ 
14              $\mathbf{h}_v^k \leftarrow \sigma(\mathbf{W}^k \cdot \text{CONCAT}(\mathbf{h}_v^{k-1}, \mathbf{h}_{\mathcal{N}_{RWR}(v)}^k))$ 
15      $\mathbf{z}_v = \mathbf{h}_v^K$ 
16     for  $uv \in \mathcal{E}$  do
17          $\mathbf{z}_{uv}^K \leftarrow \text{CONCAT}(\mathbf{z}_u^K, \mathbf{z}_v^K)$ 
  
```

The final node representation of node v can be expressed as $\mathbf{z}_v$, which describes the node embedding of the final layer, as shown in Formula (7).

$$\mathbf{z}_v = \mathbf{h}_v^k, \forall v \in V \quad (7)$$

Finally, to obtain the final edge embedding $\mathbf{z}_{uv}^k$, the concatenation of the edge embeddings of nodes u and v at the k -th layer is calculated, as shown in Formula (8). To perform edge classification, the edge embedding $\mathbf{z}_{uv}^k$ is sent to the softmax layer for classification.

$$\mathbf{z}_{uv}^k = \text{CONCAT}(\mathbf{z}_u^K, \mathbf{z}_v^K), uv \in \mathcal{E} \quad (8)$$

4.3 Integrating Graph Topology Optimization and RWR Algorithm in E-GraphSAGE

The proposed framework integrating graph topology optimization and RWR algorithm in E-GraphSAGE is illustrated in Fig. 3, with the complete algorithm process detailed in Algorithm 1.

5 Experiments and Analysis

5.1 Datasets

This paper conducts experiments on multiple public network anomaly traffic datasets, including ISCXTor2016 [24], UNSW-NB15 [25] and ToN-IoT [26]. The datasets are widely used for evaluating the performance of GNNs, and their main features are listed in Table 2.

Table 2. Basic information of the dataset

Datasets	Number of features	Quantity				
ISCXTor2016	24	NonTor	Browser	Email	Chat	Audio
		231466	4142	681	966	1872
		Video	P2P	File Transfer	VoIP	
		2156	2501	2016	5263	
UNSW-NB15	43	Normal	Analysis	Backdoor	Dos	Exploits
		2218761	2677	2329	16353	44525
		Fuzzers	Generic	Reconnaissance	Shellcode	Worms
		24246	215481	13987	1511	174
ToN-IoT	39	Normal	Backdoor	Scanning	Injection	Password
		300000	20000	20000	20000	20000
		XSS	DDOS	Ransomware	DOS	MITM
		20000	20000	20000	20000	1043

The ISCXTor2016 dataset includes Scenario A and Scenario B, which correspond to binary classification tasks and multi-classification tasks respectively. The binary classification task includes NonTor and Tor traffic, and the multi-classification task only

includes Tor traffic. The UNSW-NB15 and ToN-IoT datasets both contain binary classification and multi-classification tasks in one scenario. The training set, validation set, and test set are divided into 70%, 10%, and 20% ratios.

5.2 Experimental Setup

The experiment was conducted in an environment with Windows 11 operating system, AMD Ryzen 7 5800H CPU, 16G memory, and NVIDIA GeForce RTX 3060 GPU. Python 3.8.18 and PyTorch 2.1.1 were used to build, train, and evaluate the model. The same hyperparameters were used for the model proposed in this article and the compared model.

5.3 Evaluation Metrics

This experiment uses recall, precision, F1-score and AUC as evaluation indicators. The formula is as follows:

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

Among them, TP is the number of samples correctly classified as positive examples, TN is the number of samples correctly classified as negative examples, FN is the number of positive samples incorrectly classified as negative examples.

Because of the imbalance of the categories of the data set, simply calculating the weighted average of the above indicators in multi-classification tasks may cause the prediction performance of minority classes to be ignored. Therefore, the macro average recall rate (Macro Recall), macro average precision (Macro Precision), macro average F1-score, and macro average AUC evaluation indicators are introduced to comprehensively evaluate the performance of the model in the network traffic anomaly detection classification task. The above indicator formula is as follows:

$$\begin{aligned} Macro\ Recall &= \frac{1}{N} \sum_{i=1}^N Recall_i \\ Macro\ Precision &= \frac{1}{N} \sum_{i=1}^N Precision_i \\ Macro\ F1 - score &= 2 \times \frac{Precision_{Macro} \times Recall_{Macro}}{Precision_{Macro} + Recall_{Macro}} \end{aligned} \quad (12)$$

Among them, N is the number of categories, $Recall_i$ represents the Recall of the i -th category, and $Precision_i$ represents the Precision of the i -th category.

5.4 Result Analysis

In this section, we will conduct a detailed analysis of the experimental results on the ISCXTor2016, UNSW-NB15, and ToN-IoT datasets. The models evaluated include E-GraphSAGE [19], GAT [17], E-ResSAGE, E-ResGAT [22], the model using only the Graph Topology Optimization method (Based-GTO), the model using only the RWR method (Based-RWR), and the model that combines both Graph Topology Optimization and RWR methods (GTO-RWR). The binary classification results are shown in Table 3, while the multi-class classification results are presented in Table 4.

For the ISCXTor2016 dataset, our method in the binary classification task, the Recall, Precision, and F1-score reached 0.9908, 0.9909, and 0.9909, respectively, which are significantly higher than those achieved by other methods. In the multi-class classification task, the Recall, Precision, and F1-score achieved 0.9942, 0.9944, and 0.9942, respectively, again outperforming the other methods significantly. This demonstrates that the proposed method is highly effective in accurately identifying and classifying anomalous traffic. Additionally, the method performed exceptionally well in macro-average metrics, further showcasing its strength in addressing class imbalance issues, which are common in network anomaly detection datasets. These results highlight the model's ability to effectively identify minority classes without sacrificing performance on the more frequent classes, offering a balanced and comprehensive detection mechanism.

For the UNSW-NB15 dataset, the proposed method exhibited outstanding performance across all metrics in the binary classification task, with Recall and F1-score reaching 0.9946 and 0.9947, respectively, surpassing the other methods. In the multi-class classification task, the method particularly excelled in Recall, Precision, and F1-score, achieving 0.9923, 0.9917, and 0.9916, respectively, clearly outperforming the other methods. The method also continued to perform excellently in macro-average metrics.

For the ToN-IoT dataset, in the binary classification task, the Recall, Precision, and F1-score all reached 0.9986, slightly lower than E-ResSAGE and the RWR-based method but still higher than other methods. In the multi-class classification task, the method achieved Recall, Precision, and F1-score values of 0.9849, 0.9853, and 0.9847, respectively, slightly lower than E-ResGAT but still higher than other methods. The proposed method performed well in macro-average metrics across all categories. The slight decrease in performance compared to E-ResSAGE could be attributed to the unique characteristics of IoT traffic, which often involve more frequent and diverse interactions between devices. Despite these challenges, the proposed method's performance remained robust, showing its general applicability to different network traffic types.

Overall, across all datasets, The proposed method outperforms other models in most cases in both binary and multi-classification tasks. This highlights the advantage of integrating graph topology optimization and RWR-based sampling in enhancing the quality of edge embeddings and improving the model's ability to detect anomalies. Furthermore, the method's robustness in handling imbalanced data ensures its effectiveness in real-world network environments, where traffic distributions are rarely uniform. The combination of high precision, recall, and balanced macro-average metrics demonstrates that our method is not only accurate but also reliable in varied and complex network scenarios.

Table 3. Binary classification experimental results

Datasets	Methods	Recall	Precision	F1-score	AUC
ISCXTor2016	E-GraphSAGE	0.9783	0.9781	0.9779	0.99
	GAT	0.9672	0.9665	0.9665	0.99
	E-ResSAGE	0.9778	0.9779	0.9778	0.99
	E-ResGAT	0.9653	0.9646	0.9638	0.99
	Based-GTO	0.9792	0.9790	0.9789	0.99
	Based-RWR	0.9788	0.9803	0.9800	0.99
	GTO-RWR	0.9908	0.9909	0.9909	0.99
UNSW-NB15	E-GraphSAGE	0.9926	0.9934	0.9928	1
	GAT	0.9938	0.9948	0.9941	1
	E-ResSAGE	0.9910	0.9917	0.9912	1
	E-ResGAT	0.9944	0.9951	0.9946	1
	Based-GTO	0.9944	0.9951	0.9946	1
	Based-RWR	0.9930	0.9937	0.9932	1
	GTO-RWR	0.9946	0.9950	0.9947	1
ToN-IoT	E-GraphSAGE	0.9991	0.9992	0.9991	1
	GAT	0.9972	0.9974	0.9973	1
	E-ResSAGE	0.9995	0.9997	0.9996	1
	E-ResGAT	0.9980	0.9981	0.9980	1
	Based-GTO	0.9986	0.9986	0.9986	1
	Based-RWR	0.9997	0.9996	0.9996	1
	GTO-RWR	0.9986	0.9986	0.9986	1

To intuitively demonstrate the sampling quality of the model proposed in this paper, we use the ISCXTor2016 dataset as an example. The edge embeddings from the final layer of the model are dimensionally reduced using the t-SNE algorithm, mapping the high-dimensional data into a two-dimensional space for visualization. As shown in Fig. 4, the edge embeddings produced by our model exhibit higher class separability.

Since some classes in the ISCXTor2016 dataset are extremely sparse, this might result in less distinct distribution patterns of edge embeddings for certain classes. To address this issue, GAN-generated data was used to increase the sample density of these minority classes, leading to a more enriched and informative visualization. This aids in better identifying and understanding the differences between various classes. As shown in Fig. 5, the edge embeddings produced by our model form more distinct clusters, illustrating the improved clustering quality.

Table 4. Multi-classification experimental results

Datasets	Methods	Recall	Precision	F1-score	AUC	Macro Recall	Macro Precision	Macro F1-score	Macro AUC
ISCXTor2016	E-GraphSAGE	0.7985	0.8237	0.7927	0.98	0.7384	0.7998	0.7388	0.98
	GAT	0.7742	0.7718	0.7637	0.96	0.6308	0.6759	0.6352	0.94
	E-ResSAGE	0.7922	0.8112	0.7847	0.98	0.7469	0.7520	0.7236	0.98
	E-ResGAT	0.8051	0.8206	0.8038	0.96	0.6781	0.7226	0.6839	0.95
	Based-GTO	0.8335	0.8475	0.8282	0.98	0.7406	0.8011	0.7445	0.97
	Based-RWR	0.9905	0.9906	0.9905	1	0.9897	0.9893	0.9894	1
	GTO-RWR	0.9942	0.9944	0.9942	1	0.9940	0.9947	0.9943	1
UNSW-NB15	E-GraphSAGE	0.9838	0.9866	0.9843	1	0.6414	0.6124	0.6056	0.9
	GAT	0.9842	0.9848	0.9838	1	0.6314	0.5423	0.5614	0.99
	E-ResSAGE	0.9820	0.9793	0.9801	1	0.5615	0.5127	0.5224	0.9
	E-ResGAT	0.9872	0.9873	0.9868	1	0.5802	0.5651	0.5580	0.98
	Based-GTO	0.9903	0.9908	0.9901	1	0.6390	0.5945	0.5952	0.9
	Based-RWR	0.9840	0.9853	0.9840	1	0.6446	0.5814	0.5934	0.98
	GTO-RWR	0.9923	0.9917	0.9916	1	0.6803	0.6265	0.6340	0.96
ToN-IoT	E-GraphSAGE	0.9841	0.9852	0.9838	1	0.9284	0.9292	0.9267	1
	GAT	0.9960	0.9961	0.9960	1	0.9923	0.9920	0.9918	1
	E-ResSAGE	0.9838	0.9849	0.9834	1	0.9335	0.9301	0.9295	1
	E-ResGAT	0.9968	0.9969	0.9968	1	0.9934	0.9932	0.9930	1
	Based-GTO	0.9826	0.9833	0.9826	1	0.9498	0.9581	0.9520	1
	Based-RWR	0.9904	0.9910	0.9904	1	0.9660	0.9686	0.9662	1
	GTO-RWR	0.9849	0.9853	0.9847	1	0.9557	0.9608	0.9570	0.99

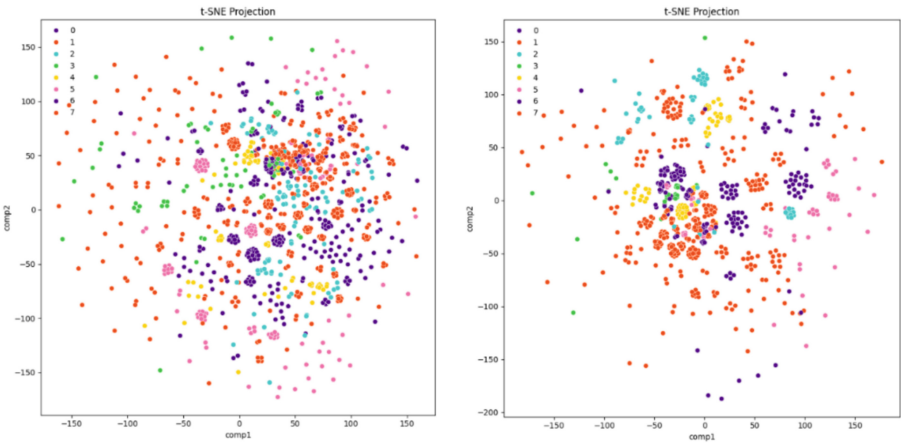


Fig. 4. The left side shows the edge embedding distribution of the original model, and the right side shows the edge embedding distribution of the model proposed in this paper.

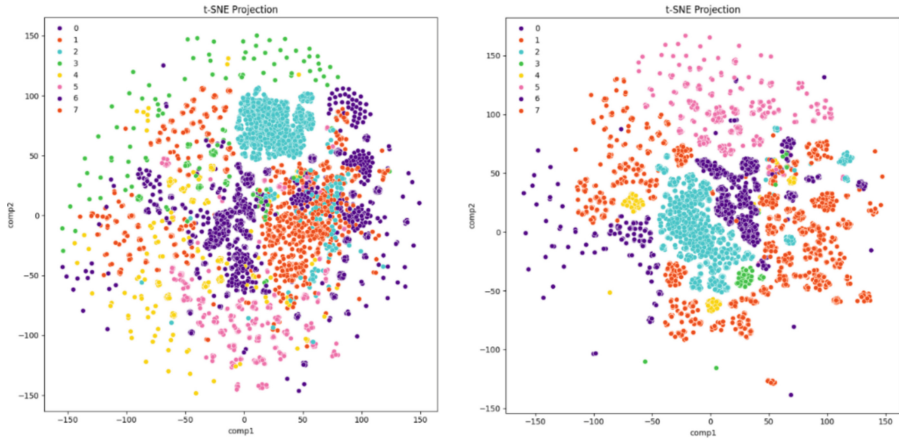


Fig. 5. The edge embedding distribution of the original model on the left and the edge embedding distribution of the proposed model on the right.

6 Conclusion

This study explores the performance enhancement of GNN-based models in network traffic anomaly detection tasks by integrating graph topology optimization and RWR algorithm sampling. The experimental results show that the proposed method improves sampling quality and significantly enhances model performance across different graph datasets. The method's emphasis on optimizing the sampling process highlights the importance of this step in improving GNN performance. By integrating graph structure modifications and data preprocessing strategies, the proposed method offers a new approach to graph data analysis and processing.

Although the study achieved positive results, it also has limitations, including the adaptability of the graph modification method to different types of graph data. Future research will explore more diverse graph modification strategies to accommodate a broader range of network structures and types. Overall, the proposed method and findings provide a new perspective for utilizing GNNs to process network traffic data, making important contributions to advancing research in network traffic anomaly detection methods.

References

1. Wang, Z., Li, J., Wang, Y., et al.: Optimal repair strategy against advanced persistent threats under time-varying networks. *IEEE Transactions on Information Forensics and Security* (2023)
2. Garcia-Teodoro, P., Diaz-Verdejo, J., Maciá-Fernández, G., et al.: Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security* **28**(1–2), 18–28 (2009)
3. Almaiah, M.A.: An efficient smart weighted and neighborhood-enabled load balancing scheme for constraint oriented networks. *International Journal of Advanced Computer Science and Applications* **11**(12) (2020)

4. Wang, W., Jian, S., Tan, Y., et al.: Representation learning-based network intrusion detection system by capturing explicit and implicit feature interactions. *Comput. Secur.* **112**, 102537 (2022)
5. Layeghy, S., Portmann, M.: On Generalisability of Machine Learning-Based Network Intrusion Detection Systems. arXiv preprint [arXiv:2205.04112](https://arxiv.org/abs/2205.04112) (2022)
6. Scarselli, F., Gori, M., Tsoi, A.C., et al.: The graph neural network model. *IEEE Trans. Neural Networks* **20**(1), 61–80 (2008)
7. Jiang, W.: Graph-based deep learning for communication networks: a survey. *Comput. Commun.* **185**, 40–54 (2022)
8. Wu, Z., Pan, S., Chen, F., et al.: A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems* **32**(1), 4–24 (2020)
9. Sinha, J., Manollas, M.: Efficient deep CNN-BiLSTM model for network intrusion detection. In: *Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Pattern Recognition*, pp. 223–231 (2020)
10. Feng, Y., Li, J., Sisodia, D.: CJ-Sniffer: measurement and content-agnostic detection of cryptotjacking traffic. In: *Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses*, pp. 482–494 (2022)
11. Wu, P., Guo, H.: LuNET: a deep neural network for network intrusion detection. In: *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, pp. 617–624 (2019)
12. Mirsky, Y., tshman, T., Elovici, Y., et al.: Kitsune: an Ensemble of Autoencoders for Online Network Intrusion Detection. arXiv preprint [arXiv:1802.09089](https://arxiv.org/abs/1802.09089) (2018)
13. 林宏刚,张运理,郭楠馨,陈麟.基于图神经网络的P2P僵尸网络检测方法[J].*工程科学与技术* **54**(2), 65–72 (2022)
14. Cheng, Q., Wu, C., Zhou, S.: Discovering attack scenarios via intrusion alert correlation using graph convolutional networks. *IEEE Commun. Lett.* **25**(5), 1564–1567 (2021)
15. Zhang, Y., Yang, C., Huang, K., et al.: Intrusion detection of industrial internet-of-things based on reconstructed graph neural networks. *IEEE Transactions on Network Science and Engineering* (2022)
16. Nguyen, H., Kashef, R.: TS-IDS: traffic-aware self-supervised learning for IoT network intrusion detection. *Knowl.-Based Syst.* **279**, 110966 (2023)
17. Veličković, P., Cucurull, G., Casanova, A., et al.: Graph Attention Networks. arXiv preprint [arXiv:1710.10903](https://arxiv.org/abs/1710.10903) (2017)
18. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. *Advances in Neural Information Processing Systems* **30** (2017)
19. Lo, W.W., Layeghy, S., Sarhan, M., et al.: E-graphsage: a graph neural network based intrusion detection system for iot. In: *NOMS 2022–2022 IEEE/IFIP Network Operations and Management Symposium*. IEEE, pp. 1–9 (2022)
20. Caville, E., Lo, W.W., Layeghy, S., et al.: Anomal-E: a self-supervised network intrusion detection system based on graph neural networks. *Knowl.-Based Syst.* **258**, 110030 (2022)
21. Liu, T., Jiang, A., Zhou, J., et al.: GraphSAGE-based dynamic spatial–temporal graph convolutional network for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems* (2023)
22. Chang, L., Branco, P.: Embedding residuals in graph-based solutions: the E-ResSAGE and E-ResGAT algorithms. a case study in intrusion detection. *Applied Intelligence*, pp. 1–16 (2024)
23. Page, L., Brin, S., Motwani, R., et al.: The PageRank citation ranking: Bringing order to the web. *Stanford infolab* (1999)
24. Lashkari, A.H., Gil, G.D., Mamun, M.S.I., et al.: Characterization of tor traffic using time based features. In: *International Conference on Information Systems Security and Privacy*. SciTePress **2**, 253–262 (2017)

25. Moustafa, N., Slay, J.: UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In: 2015 Military Communications and Information Systems Conference (MilCIS). IEEE, pp. 1–6 (2015)
26. Alsaedi, A., Moustafa, N., Tari, Z., et al.: TON_IoT telemetry dataset: a new generation dataset of IoT and IIoT for data-driven intrusion detection systems. *IEEE Access* **8**, 165130–165150 (2020)



An Evaluation Method for Cutted Generated Video Based on Eigenvectors

Ji Yu, Haotian Gao, and Jundi Ding^(✉)

1-School of Computer Science and Engineering, Nanjing University of Science and
Technology, Nanjing, China
{121106010763, gaohaotian, dingjundi2010}@njust.edu.cn

Abstract. With the development of deep learning, the artificial intelligence text generation video model Sora combines the characteristics of Diffusion models, Transformer architecture, and multimodal learning to generate high-quality video content. In this case, we propose an evaluation method for generated videos after cutting based on eigenvectors, which evaluates the AI generated videos after cutting and determines whether the cutted short videos meet the evaluation criteria. This algorithm uses linear algebra knowledge in mathematics to process the eigenvectors of images and calculates the similarity between images using relevant threshold. We calculate the minimum “orthogonal value” of the eigenvectors of the first frame of a generated short video and the last frame of another generated short video, and the accuracy of the evaluation method is 99% when the selected threshold is 0.75. This evaluation method has surpassed neural network algorithms in terms of code simplification and algorithm simplification, achieving good results.

Keywords: Eigenvectors · Image Clustering · Video Action Recognition

1 Introduction

1.1 Image Clustering Algorithm

Image clustering is an unsupervised learning task that aims to group images based on their similarity. Clustering algorithms usually use image features as input, such as color, texture, shape, etc., and group images based on the distance or similarity between these features. Common clustering algorithms include K-Means, DBSCAN, and image clustering algorithms based on deep learning.

Traditional Clustering Algorithms. K-Means clustering is a clustering algorithm that uses distance to measure similarity. The core idea is to divide the data into k clusters, so that the internal distance of each cluster is minimized and the external distance is maximized. The algorithm first randomly initializes the centroids, then divides the data points into the nearest clusters based on the

distance between the data points and the centroids, and finally recalculates the center point of each cluster. This process will be repeated until the centroid no longer changes. The K-Means algorithm is simple and efficient, with low time and space complexity, but it also has obvious limitations, such as the clustering results being dependent on the selection of initial clustering centers, being sensitive to abnormal sample points, and not being able to handle non-convex data.

DBSCAN is a density-based clustering method that divides data points into dense and sparse regions, grouping data points in dense areas together. DBSCAN clustering requires determining two parameters: the radius (eps), which indicates the neighborhood range around a given point P, and the minimum number of points (MinPts) required in the neighborhood around point P. If the number of points in the eps neighborhood of point P is greater than or equal to MinPts, P is termed a core point. If the number of points is less than MinPts but falls within the neighborhood of a core point, P is referred to as a boundary point. The specific steps for DBSCAN are: 1. Label all points as core points, boundary points, or noise points (neither core nor boundary points); 2. Remove noise points; 3. Add an edge between core points that are within eps distance of each other; 4. All connected core points form a cluster; 5. Include boundary points in the cluster of the core point if they fall within the eps neighborhood. Compared to the K-Means algorithm, the advantage of DBSCAN is that it does not require prior specification of the number of clusters, is insensitive to noise, and can handle clusters of arbitrary shapes. Its disadvantage is that the clustering results are highly dependent on parameter selection, and DBSCAN may struggle to select suitable parameters for clusters with varying density.

Neural Network Based Clustering Algorithm. Autoencoders achieve data clustering by extracting features and reconstructing input data. An autoencoder is a type of neural network model consisting of an encoder and a decoder. It trains the model by minimizing reconstruction errors and uses the hidden layers of the encoder as feature representations for clustering.

1.2 Video Action Recognition

Video action recognition is the core field of video understanding, although it mainly recognizes the actions of people in videos, it can also be applied to scene cutting. Video action recognition refers to the detection and classification of human movements contained in video sequences in both temporal and spatial dimensions. As a dynamic sequence of images sorted by time, videos have correlation relationships between image frames. Therefore, video action recognition not only needs to process spatial image information, but also needs to model the time series. By combining the spatiotemporal information in the video, we can better understand the dynamic changes in the video and more accurately cut the scene.

Traditional Video Action Recognition Methods. Traditional video action recognition methods include Dense Trajectory Algorithm (DT algorithm) and Improved Dense Trajectory Algorithm (IDT algorithm). The DT algorithm utilizes trajectory shape features, HOF, HOG, and MBH features, followed by encoding the features using the BoF method, and finally training an SVM classifier for action recognition. The basic framework of IDT algorithm is the same as DT algorithm, with the main improvements being the optimization of optical flow images, the improvement of feature regularization methods, and the improvement of feature encoding methods. By estimating camera motion to eliminate optical flow and trajectories on the background, SURF keypoint descriptors and dense optical flow are used for feature point matching between frames to eliminate or mitigate the impact of camera motion. Use L1 regularization to normalize the obtained features. Unlike the BoF encoding method used in DT algorithm, IDT algorithm adopts Fisher Vector encoding and then uses SVM for classification. The IDT algorithm, as the best traditional algorithm in action detection, has excellent performance and robustness, achieving an accuracy of 91.2% on the UCF50 dataset and 57.2% on the HMDB51 dataset.

Video Action Recognition Method Based on Deep Learning. The process of video action recognition based on deep learning can be divided into two main steps: feature extraction and action classification. In the feature extraction stage, we need to extract useful features from the video for subsequent classification tasks. In the feature extraction stage, the commonly used method is to use CNN to process video frames and extract spatiotemporal features. In the action classification stage, various deep learning models can be used to classify the extracted features. Through these methods and techniques, we can better utilize deep learning to achieve video action recognition, providing more accurate results for practical applications.

Our Method. This study proposes eigenvectors based evaluation method to assess the AI generated videos after cutting and determine whether the cutted short videos meet the evaluation criteria. This algorithm uses linear algebra knowledge in mathematics to process the eigenvectors of images, and uses correlation threshold to calculate the similarity between images. We propose the concept of “orthogonal value” to evaluate the minimum “orthogonal value” of the eigenvectors of the first frame of a generated short video and the last frame of another generated short video. With a selected threshold of 0.75, the accuracy of evaluation method is 99%, achieving good results.

2 Related Work

2.1 Image Clustering

Image clustering algorithms include traditional image clustering algorithms and deep learning based image clustering algorithms. Image clustering refers to classifying images in image datasets based on their features, so that images in the

same category have high similarity in features, while images in different categories have significant differences in features.

Traditional Image Clustering Algorithms. Traditional image clustering algorithms include K-MEANS algorithm, DBSCAN algorithm, and spectral clustering algorithm. Spectral clustering algorithm is a clustering method based on graph theory, with the basic steps: 1. Constructing a similarity matrix: For a given dataset, the similarity between them needs to be calculated first. Similarity calculation can be selected based on the characteristics of the corresponding dataset. Usually, methods such as Euclidean distance and cosine similarity can be used to calculate similarity and obtain a similarity matrix. 2. Constructing Laplacian matrix: Based on the similarity matrix, a Laplacian matrix can be constructed. The Laplacian matrix is obtained by regularizing the similarity matrix, and can be constructed in two ways: diagonal matrix and adjacency matrix. 3. Solving eigenvalues and eigenvectors: By performing eigenvalue decomposition on the Laplacian matrix, a series of eigenvalues and corresponding eigenvectors can be obtained. 4. Select eigenvectors: Based on the eigenvectors corresponding to the first K smallest eigenvalues, these eigenvectors form a low dimensional space, and the original dataset is projected into this low dimensional space. 5. Clustering: The K-MEANS algorithm can be used to perform clustering analysis on the projected dataset to obtain the final clustering results. The biggest advantage of spectral clustering algorithm is that it can effectively solve the clustering problem of high-dimensional data. The main disadvantage is that multiple parameters need to be selected during the implementation of the algorithm, and the selection of these parameters has a significant impact on the clustering results.

Deep Learning Based Image Clustering Algorithms. Deep learning based image clustering algorithms are methods of using deep learning models for clustering analysis. Deep learning clustering methods typically involve the following three steps: 1. Encoding data using deep learning models such as autoencoders, variational autoencoders, etc. 2. Cluster using encoded data. 3. Evaluate and visualize the clustering results. The core concepts of deep learning clustering include: 1. Encoder: a function that converts raw data into low dimensional representations. 2. Decoder: A function that converts low dimensional representations back to the original data. 3. Loss function: a function used to evaluate the performance of encoders and decoders. With the flourishing development of deep learning [3, 6, 10, 13], many researchers have turned their attention to deep unsupervised feature learning and clustering [2, 4, 9]. Jianlong Wu et al. [12] proposed a new clustering framework called Deep Comprehensive Correlation Mining (DCCM), which addressed the clustering problem by proposing triplet mutual information between features, elevating recently discovered instance level deep mutual information to a triplet structure, further facilitating the learning of more discriminative features. The advantage of this method is that it has achieved good performance, such as a clustering accuracy of 62.3% on

CIFAR-10, which is 10.1% higher than the state-of-the-art results. The disadvantage is that DCCM still lacks local robustness in data transformation, extraction of relationships between the same data in different network layers, and mutual relationships between other samples. Mathilde Caron et al. [1] proposed a new unsupervised method for end-to-end learning of convnets, which iteratively clustered the features generated by convnets using k-means and updated their weights by predicting the clusters as pseudo labels in the discriminative loss. The disadvantage is that the computational cost is high, and this algorithm requires clustering of all data points every iteration.

2.2 Video Action Recognition Method

Video action recognition methods include traditional video action recognition methods and deep learning based video action recognition methods. Video action recognition is an important research direction in the field of computer vision, which involves automatically analyzing the actions of people or objects in videos and classifying them into specific categories. Video action recognition technology achieves accurate classification and recognition of actions by analyzing the temporal information between image frames in the video.

Traditional Video Action Recognition Technology. Traditional video action recognition technology mainly includes three steps: feature extraction, feature encoding, and classifier construction. Feature extraction is the most critical step among them. Feature extraction involves extracting special relationships between video frames through techniques such as frame difference and optical flow, and converting them into a digital feature matrix. In the IDT method, Fisher Vecor was used to encode the features, and then SVM classifier was used to classify the encoded features.

Video Action Recognition Methods Based on Deep Learning. There are many video action recognition methods based on deep learning. Limin Wang et al. [11] proposed a temporal segmentation network framework consisting of a spatial stream convolutional neural network and a temporal stream convolutional neural network, aiming to use the visual information of the entire video for video level prediction. The advantage is that it solves the problem of large computational complexity in long videos, but the disadvantage is that it performs poorly when dealing with short-term actions. Boyuan Jiang et al. [5] proposed a simple and effective STM network that effectively encodes spatiotemporal and motion features. The advantage is that the effect is very good on many datasets, including Something, Kinetics, Jester, UCF101, and HMDB-51 all achieving SOTA, but the disadvantage is that they cannot compare to TSM method in terms of time performance. Yan Li et al. [7] proposed a Time Excitation and Aggregation (TEA) block, which included a Motion Excitation (ME) module and a Multi Time Aggregation (MTA) module, specifically designed to capture short-term and long-term temporal evolution. The advantage is that the ME module is used

to replace the traditional manual extraction of optical flow, which was input into a 2D convolutional network-based dual flow framework for action recognition, solving the problem of time-consuming and storage required for optical flow calculation. The disadvantage is that it requires abundant computing resources. Zhaoyang Liu et al. [8] proposed a new time modeling module called TEI module, which decoupled the task of time feature learning into channel level enhancement and local time interaction, and can achieve significant improvements over previous time modeling methods at lower computational costs. The disadvantage is that it requires abundant computing resources.

3 Our Method

3.1 The Main Idea of Our Method

We mainly use eigenvectors in linear algebra to process images, and then use correlation threshold to evaluate the different cutted short videos. Each evaluation compares the minimum “orthogonal value” of the eigenvectors of the first frame of one short video and the last frame of another short video.

The calculation of eigenvectors and eigenvalues has a wide range of applications in machine learning, signal processing, quantum mechanics, and other fields. Eigenvalues and eigenvectors are fundamental concepts in linear algebra. If matrix A is multiplied by a non-zero vector x , the result is:

$$Ax = cx \quad (1)$$

where c is a scalar, then c is the eigenvalue of A , and x is the corresponding eigenvector. It should be noted that the eigenvector is a non-zero vector, so the solved x value cannot be a zero vector.

The Physical Meaning of Eigenvectors and Eigenvalues. The physical significance of eigenvectors and eigenvalues mainly lies in their description of the properties of vector scaling and direction invariance in linear transformations. The size of the eigenvalue indicates the degree of stretching of the eigenvector in a certain direction. For an image, it can be converted into a matrix of grayscale values, and different linearly independent eigenvectors become basis vectors, forming a linear space.

3.2 The Implementation Steps of the Method in this Article

The implementation steps of this method in this article are image preprocessing, calculating the eigenvalues and eigenvectors of the image, and then using relevant threshold to evaluate the minimum “orthogonal value” of corresponding rows of eigenvectors of the first frame of one short video and the last frame of another short video.

Image Preprocessing. This article uses the transforms function to transform the image in the preprocessing steps, adjusting the image size to (256, 256) size. There are many image preprocessing operations that can be selected and combined according to specific application scenarios and model requirements, such as normalizing, flipping, rotating, and filtering images. This article simply uses the method of adjusting image size, which does not require a lot of computational complexity and time cost.

Processing the Eigenvalues and Eigenvectors of an Image. Processing the eigenvalues and eigenvectors of an image is a fundamental concept in linear algebra. This article generates a grayscale image from a color image (usually composed of three channels: red, green, and blue) by calculating the average value of each pixel across all channels. Subsequently, the grayscale values of the image are used to calculate the eigenvalues and eigenvectors. The eigenvalues are sorted in ascending order, and the eigenvectors are arranged correspondingly to the eigenvalues. Subsequently, calculate the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the first frame of one short video and the last frame of another short video.

Selection of Threshold. This article uses a threshold of 0.75 to evaluate the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the first frame of a short video and the last frame of another short video. The input of the program is two images, namely the first frame of a short video and the last frame of another short video. The eigenvalues are sorted in ascending order, and the eigenvectors are arranged correspondingly to the feature values. Then, the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the first frame of one short video and the last frame of another short video is calculated. The calculation method of the “orthogonal value” is as follows: 1. Calculate the cosine value between the corresponding rows of eigenvectors of two images to calculate the similarity value, with a threshold range of $[-1, 1]$. 2. According to the threshold range of cosine values, the calculation method for “orthogonal values” is:

$$\text{Orthogonal value} = 1 - |\text{cosine value}| \quad (2)$$

The reason for using this calculation formula is that the closer the cosine value between vectors is to 1, the more similar the vectors are. After using this formula, it means that the closer the “orthogonal value” between vectors is to 1, the less similar the two vectors are.

The reason for setting the threshold to 0.75 is that if it is lower than the threshold of 0.75, the evaluation result of this threshold is meaningless, because for short videos with different source materials, the minimum “orthogonal value” between the corresponding rows of eigenvectors in the first frame of one short video and the last frame of another short video is basically greater than 0.75. When the threshold is higher than 0.75, the evaluation accuracy of this threshold will gradually decrease.

Evaluation Methods and Criteria. We conduct experiments on different generated short videos, with a total of 33 materials. The number of combinations of different short videos between each pair is $33 * 32 = 1056$ combinations. Our evaluation method is: if the minimum “orthogonal value” between the corresponding rows of the eigenvectors in the first frame of one short video and the last frame of another short video is less than 0.75, then it can be determined that these two short videos belong to the same video sequence, have the same materials and scenes, and are not included in the accuracy. The accuracy of this evaluation method is 99.62%.

4 Dataset

The dataset we use is obtained from promotional videos on the Sora website, which contains 33 source materials. We manually cut the videos into 33 subfolders, each with a corresponding video frame sequence for experimentation and evaluation (Figs. 1 and 2).



Fig. 1. The first and last frames of a generated video in the dataset



Fig. 2. The first and last frames of another generated video in the dataset

5 Experimental Results and Evaluation

We arrange and combine 33 generated short videos, with 1056 different combinations of short videos between each pair. We select the threshold and conduct experiments to calculate the accuracy. The calculation method for accuracy is: if the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the first frame of one short video and the last frame of another short video is greater than or equal to 0.75, it is included in the accuracy (Fig. 3 and Table 1).



Fig. 3. Only display one frame and the next frame of the generated video

Table 1. The relationship between threshold selection and accuracy

Selection of threshold	0.75	0.8	0.85	0.9	0.95
Accuracy	99.62%	87.02%	13.92%	0%	0%

To demonstrate the effectiveness of our method, we conduct comparative experiments on the entire promotional video and calculate the minimum “orthogonal value” between the eigenvectors of the two frames before and after. The accuracy obtained is 77.78%. This indicates that our evaluation method is more effective and accurate (Table 2).

Table 2. Comparative experiment with a threshold of 0.75

Compared frames	two different generated short videos	the entire video
Accuracy	99.62%	77.78%

6 Limitations of Our Method

We propose an eigenvector based evaluation method for assessing cutted short videos from AI generated videos and determining whether the cutted videos meet the evaluation criteria. Our evaluation method is: if the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the first frame of one short video and the last frame of another short video is less than 0.75, it can be determined that these two generated short videos belong to the same video sequence, have the same materials and scenes. The limitation of our method is that it cannot detect all generated videos belonging to the same video sequence. According to experiments, the minimum “orthogonal value” between the corresponding rows of the eigenvectors of the generated videos of the same video sequence can be greater than 0.75. Our evaluation method can be used as a supplementary detection method, which is a limitation of our approach.

7 Conclusion

We propose an evaluation method for generated video after cutting based on eigenvectors, and its code implementation link is <https://github.com/freezeparadox/Image-Clustering-Algorithm>. We propose an eigenvector based evaluation method to assess the cutted short videos from AI generated videos and determine whether the cutted videos meet the evaluation criteria. This evaluation method has a good effect, with accuracy of 99% when the selected threshold is 0.75. Our method can be applied to evaluate whether the cutted short videos from AI generated videos are accurately cutted, which has practical significance.

Statement. Our method only uses open-source datasets and there are no conflicts of interest.

References

1. Caron, M., Bojanowski, P., Joulin, A., Douze, M.: Deep clustering for unsupervised learning of visual features. In: European Conference on Computer Vision (2018)
2. Chang, J., Wang, L., Meng, G., Xiang, S., Pan, C.: Deep adaptive image clustering (2017)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition (2015)
4. Ji, P., Zhang, T., Li, H., Salzmann, M., Reid, I.: Deep subspace clustering network (2017)
5. Jiang, B., Wang, M., Gan, W., Wu, W., Yan, J.: STM: spatiotemporal and motion encoding for action recognition (2019)
6. Li, X., Wu, J., Lin, Z., Liu, H., Zha, H.: Recurrent squeeze-and-excitation context aggregation net for single image deraining (2018)
7. Li, Y., Ji, B., Shi, X., Zhang, J., Kang, B., Wang, L.: Tea: temporal excitation and aggregation for action recognition (2020)
8. Liu, Z., et al.: Teinet: towards an efficient architecture for video recognition (2019)

9. Peng, X., Xiao, S., Feng, J., Yau, W., Yi, Z.: Deep subspace clustering with sparsity prior (2016)
10. Wang, F., et al.: Residual attention network for image classification (2017)
11. Wang, L., et al.: Temporal segment networks: towards good practices for deep action recognition (2016)
12. Wu, J., et al.: Deep comprehensive correlation mining for image clustering (2019)
13. Xie, X., Wu, J., Zhong, Z., Liu, G., Lin, Z.: Differentiable linearized ADMM (2019)



Joint Optimization of Task Scheduling and Resource Allocation for Hierarchical Federated Learning

Nan Ding¹, Tianshu Huang²(✉), and Liutao Zhao³

¹ Beijing Nickle Net Computer Technology Co., Ltd., Beijing 100192, China

² Guangxi Botanical Garden of Medicinal Plant, Beijing 530012, China
huangts@gxyyzwy.com

³ Beijing Academy of Science and Technology, Beijing 100094, China

Abstract. With the explosive growth of IoT devices and data volumes, traditional centralized machine learning methods can no longer meet the demands of modern application scenarios. Hierarchical Federated Learning (HFL), as an enhanced federated learning framework, builds a multi-level collaborative mechanism among devices at different layers. It ensures data security while maintaining high scalability, offering broad application prospects. However, existing HFL schemes still face numerous challenges in task scheduling and resource allocation. This paper proposes a joint optimization solution to address the task scheduling and resource allocation problems within the HFL framework. Firstly, we model the task scheduling and resource allocation problem in the HFL system as a constrained optimization problem, with the objective of minimizing global training time and energy consumption. Subsequently, we introduce a hybrid heuristic optimization algorithm to solve this problem. This hybrid algorithm leverages the strengths of both PSO and SA, enabling fast solutions while reducing the likelihood of getting trapped in local optima. Experimental results demonstrate that our algorithm can efficiently complete task scheduling and resource allocation in the hierarchical federated learning system with high quality.

Keywords: Hierarchical Federated Learning · Task Scheduling · Resource Allocation

1 Introduction

In today's data-driven world, Federated Learning (FL) has emerged as an innovative distributed machine learning paradigm, allowing multiple participants to collaboratively train a model without sharing raw data. This effectively addresses the issue of data silos while ensuring data privacy. The basic framework of FL involves a central server and multiple edge devices or clients, where the core principle is to keep the data “stationary”—meaning that model updates are aggregated centrally, rather than transferring the data itself. However, in practice, as the number of participants grows, traditional FL faces challenges such as low communication efficiency and increased network latency, particularly in large-scale cross-regional deployments.

Hierarchical Federated Learning (HFL) is an advanced distributed machine learning framework that introduces an intermediate layer (e.g., edge servers or regional centers) into the traditional client-server architecture, forming a multi-tiered communication and computational structure. The core concept of HFL is to offload part of the aggregation work to intermediate layers closer to the data source, thereby significantly reducing the communication load on the central server and improving the overall communication efficiency and system responsiveness. Additionally, HFL exhibits greater scalability, supporting a wider range of distributed learning environments, while enhancing system robustness to better handle challenges such as heterogeneous data distributions and varying computational resources. Therefore, HFL not only optimizes resource utilization but also facilitates cross-organizational collaborative learning while preserving data privacy.

In federated learning, task scheduling and resource allocation refer to how to efficiently assign computational tasks and resources in heterogeneous computing environments. Since the devices participating in FL differ significantly in terms of computational power, memory, and network bandwidth, intelligent task scheduling algorithms are required to dynamically allocate tasks and optimize resource usage to overcome the challenges posed by device heterogeneity. Proper scheduling and resource allocation can enhance overall learning efficiency and accelerate model convergence. Conversely, improper task scheduling and resource allocation can lead to a range of issues in FL, including low training efficiency, slow model convergence, increased communication costs, and resource wastage. Specifically, poor scheduling might overload certain clients while leaving others idle, negatively impacting the quality of the global model. Additionally, inefficient resource allocation can exacerbate network congestion, prolong training cycles, and potentially prevent some clients from participating due to insufficient resources, ultimately affecting the stability and reliability of the entire FL system.

Compared to single-layer FL, hierarchical federated learning faces even more complex challenges in task scheduling and resource allocation. In particular, coordination across multiple layers requires efficient task assignment and resource utilization mechanisms not only between clients and edge servers but also between edge servers and the central server to prevent any layer from becoming a bottleneck in the system. The increased complexity of the system means that more communication paths and potential data synchronization issues must be managed, especially during data transmission and model updates across different layers. The broader heterogeneity also demands that scheduling algorithms flexibly adapt to differences in hardware, network conditions, and computational capacities across layers, ensuring that the system maintains high performance and stability even in the face of network instability or limited computational resources.

In this study, we address the challenge of optimizing both task scheduling and resource allocation within the HFL system by framing it as a constrained optimization problem. We introduce a hybrid heuristic algorithm designed to tackle this issue, which devises strategies for both task scheduling and resource management in HFL. The experimental results indicate that our proposed approach excels in reducing latency and minimizing energy consumption.

2 Related Works

In recent years, conducting machine learning within wireless edge networks has become a focal area for researchers. In traditional machine learning practices, users' local data must be uploaded to cloud servers with substantial computational power for centralized model training. This system architecture typically consists of a central cloud server and multiple participating nodes responsible for collecting data. However, in such a scenario, the massive data upload not only consumes considerable energy but also increases communication latency. For participants handling sensitive data, there is an added risk of privacy breaches.

To address these challenges, researchers have focused on developing distributed machine learning solutions. Notably, Google first introduced the concept of Federated Learning (FL) in 2017 [1].

In the traditional two-tier federated learning framework, terminal devices interact directly with a central server, uploading locally trained model parameters for global model updates. However, in the hierarchical federated learning (HFL) architecture [2–4], the concept of edge computing is introduced, where edge servers act as intermediaries between terminal devices and the central server. The primary role of the edge server is to receive local model parameters from terminal devices and perform preliminary model aggregation at the edge level. The aggregated model parameters are then uploaded to the central server for further global model fusion and updates. This hierarchical approach significantly reduces the data transmission between terminal devices and the central server.

Federated learning still faces several challenges, including consistency [5, 6]. To improve the efficiency and resource utilization in FL, some studies have focused on optimizing task scheduling and resource allocation within federated learning frameworks.

For instance, in reference [7], researchers aimed to minimize training energy consumption under delay and bandwidth constraints while using federated learning to train machine learning models. They formulated an optimization problem to minimize training latency. Additionally, they compared training latency performance under different transmission strategies. In addressing the power control problem, FAROOQ et al. [8] proposed an innovative approach by transforming the originally non-convex optimization problem into a convex programming problem. Based on this transformation, they applied Nesterov's smoothing technique and the accelerated projected gradient method to solve the transformed convex problem.

Some works focused on the long-term energy consumption constraints of terminal devices, aiming to maximize the number of devices participating in training during each cycle. This study transformed the original problem using the Lyapunov optimization framework, utilizing only the current cycle's Channel State Information (CSI), ensuring that devices could meet long-term energy constraints while analyzing the upper bounds of long-term learning performance and energy consumption.

In the context of hierarchical federated learning, some works proposed an optimization strategy that minimizes the weighted sum of training latency and energy consumption by designing bandwidth allocation and device association strategies. An iterative

optimization method was employed to efficiently determine the optimal resource management strategy for a single edge server. Additionally, the study introduced a heuristic device association method, consisting of device migration and device swapping steps, which could achieve convergence within a few iterations.

3 Joint Optimization of Task Scheduling and Resource Allocation for Hierarchical Federated Learning

In a typical hierarchical federated learning (HFL) framework based on a cloud-edge-end architecture, there are three main entities: terminals, a cloud server and edge servers. Terminals collect data and maintain local datasets, with limited computational capabilities. They perform local training based on their datasets and send local model updates to the edge server for aggregation, while synchronizing the aggregated model from the edge server. Each edge server is associated with multiple terminals and acts as the aggregation center for those terminals, receiving their local model updates and executing edge aggregation.

In this paper, we focus on optimizing task scheduling and resource allocation in hierarchical federated learning to improve resource utilization in the federated learning system. The key concerns are federated learning efficiency and the energy consumption of the terminals. The efficiency of federated learning is measured by the total time required for global training. In the system considered in this paper, terminals establish connections with edge servers via wireless networks, while edge servers connect to the cloud server via optical fiber. Therefore, the global training time is primarily determined by the local training time at the terminals and the communication delay between terminals and edge servers, while the model aggregation and distribution delays are ignored.

For task scheduling, we relax the typical constraint in federated learning that data must remain local, allowing some terminals with limited computational resources to offload their computational tasks to edge server by uploading their data. The edge server aggregates the uploaded data from the associated terminals as a local dataset, performs local training, and participates in edge aggregation.

In terms of resource allocation, since multiple terminals are connected to edge servers via wireless networks, bandwidth allocation is a critical factor that determines the model update transmission delays from terminals to edge servers, thus affecting the overall efficiency of the federated learning process.

The hierarchical federated learning architecture in this paper is shown in Fig. 1, which includes M terminals and N edge servers. Terminal v_i maintains its own local dataset $\mathcal{D}_i$. $|\mathcal{D}_i|$ represents the sample size of the dataset. The association relationship between terminal v_i and edge server s_j is represented by $x_{i,j}$. Specifically, if $x_{i,j} = 1$, it means that s_j and v_i are associated, otherwise $x_{i,j} = 0$. $\mathcal{V}_j$ represents the set of terminals associated with edge server s_j . The computing resources of terminal v_i are defined as f_i , and the computing resources of edge server s_j are defined as q_j . The amount of data unloaded by terminal v_i to the associated edge server is o_{ij} . The delay of terminal v_i performing local training is

$$t_i^{local} = \frac{c \times E \times (\mathcal{D}_i - o_{ij})}{f_i}$$

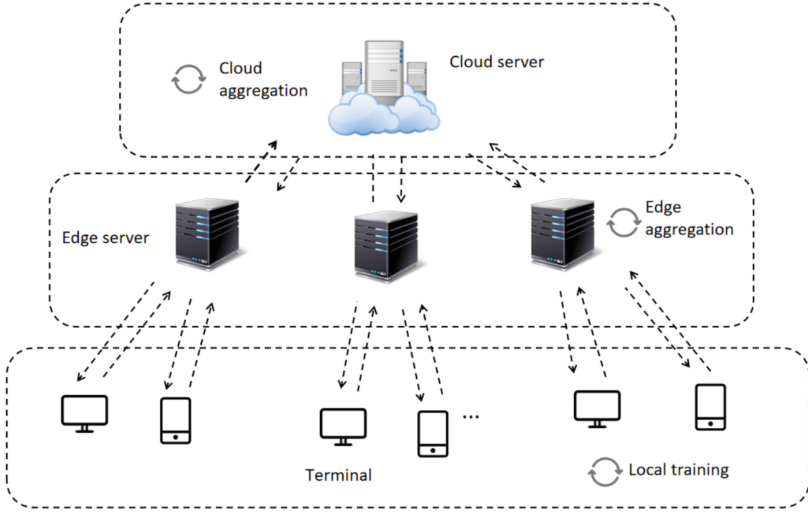


Fig. 1. System architecture diagram

The number of CPU cycles required to train a single sample is represented by E , which represents the number of local training rounds.

The local training latency of the edge server can be expressed as

$$t_j^{edge} = \frac{c \times E \times \sum_{i \in \mathcal{V}_j} o_{ij}}{q_j}$$

Under Shannon's theorem, the transmission rate between terminal v_i and the associated edge server is defined as

$$r_i = B_{ij} \log_2 \left(1 + \frac{hP_i}{\sigma} \right)$$

where B_{ij} represents the bandwidth of the communication link between terminal v_i and server s_j , h represents the channel gain, P_i represents the transmission power of terminal v_i , and σ represents the noise power. The local model update upload delay of terminal v_i is

$$t_i^{up} = \frac{|\mathbf{w}|}{r_i}$$

In the formula, $|\mathbf{w}|$ represents the size of the model parameters.

The energy consumption of terminal v_i performing local training and edge server s_j performing local training are defined as

$$E_i^{local} = t_i^{local} * P_i^v$$

$$E_j^{Edge} = t_j^{edge} * P_j^s$$

where P_i^v and P_j^s represents the energy consumption of terminal and edge servers performing floating-point calculations.

The data transmission energy consumption of terminal v_i is

$$E_i^{up} = t_i^{up} * P_i^{up}$$

where P_i^{up} represents the energy consumption of the terminal to transmit data through the wireless network.

The time required for the global training of federated learning is

$$T_{train} = \max(\max(t_i^{local} + t_i^{up}), t_j^{edge}), i \leq M, j \leq N$$

The energy consumption of global training performed by federated learning is

$$E_{train} = \sum_{i \leq M} E_i^{local} + E_i^{up} + \sum_{j \leq N} E_j^{Edge}$$

In this paper, task scheduling and resource allocation optimization in a hierarchical federated learning system is modeled as an optimization problem, which aims to achieve higher federated training efficiency and lower energy consumption by adjusting the amount of tasks offloaded from each terminal to the associated edge server. This problem can be formally expressed as

$$\min_S f(S) = \min_S (T_{train}(S) + V \cdot E_{train}(S))$$

$$s.t. \sum_{i \in \mathcal{V}_j} B_{ij} \leq B_j, j \leq N$$

$$0 \leq o_{ij} \leq \mathcal{D}_i, i \leq M$$

The solution to the optimization problem in this paper is defined as $S = \langle O, B \rangle$, where O is a vector representing the task scheduling strategy of each terminal, and B is a vector representing the bandwidth between each terminal and the associated edge server. V is a weight measurement.

The joint optimization problem of task scheduling strategy and bandwidth resources is an NP-hard problem. This paper proposes a hybrid algorithm. The core idea is to use PSO for global search and introduce SA's local search at the appropriate time to avoid falling into the local optimum and improve the accuracy of the global optimal solution.

In the PSO part of this algorithm, the particle swarm searches for the optimal solution by iteratively updating the position and speed in the global search space. After PSO searches to a certain stage, the global optimal solution or specific particles are selected for SA local optimization. The introduction of SA can help to jump out of the local optimal solution.

In the PSO-SA hybrid algorithm, the particle swarm is first initialized, a group of particles are randomly generated in the search space, the initial position and speed of the particle swarm are set, and the initial temperature $T(0)$ of the simulated annealing is set. Then PSO is executed, and the speed and position of PSO are updated during

this process, and the optimal solution is searched through group cooperation. The speed update formula is as follows

$$\mathbf{v}_i(t+1) = \omega \mathbf{v}_i(t) + c_1 r_1 (\mathbf{p}_i(t) - \mathbf{x}_i(t)) + c_2 r_2 (\mathbf{g}(t) - \mathbf{x}_i(t))$$

where ω is the inertia weight, which controls the influence of the current speed. c_1 and c_2 are learning factors. r_1 and r_2 are random numbers between $[0,1]$. $\mathbf{p}_i(t)$ is the current individual optimal position of particle i . $\mathbf{g}(t)$ is the global optimal position of the particle group.

The particle position update formula is:

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1)$$

Through iterative updates, particles gradually approach the global optimal solution. After several iterations (or when the global optimal solution is no longer significantly improved), SA local optimization is introduced for the optimal particle to jump out of the possible local optimal area. The objective function of simulated annealing is the same as that of PSO. After each local perturbation, the energy difference is calculated:

$$\Delta E = f(\mathbf{x}') - f(\mathbf{x})$$

where $\mathbf{x}'$ is a new solution generated by random perturbation.

If the new solution $\mathbf{x}'$ is better than the current solution (i.e. $\Delta E < 0$), it is directly accepted; otherwise, the worse solution is accepted with probability P :

$$P = \exp\left(-\frac{\Delta E}{T}\right)$$

where T is the current temperature, which decreases gradually with each iteration. The temperature T decreases according to the following formula:

$$T(t+1) = \alpha T(t)$$

where α is the temperature attenuation coefficient.

The algorithm concludes upon reaching the maximum number of iterations, or the temperature T drops to a sufficiently small value, or the improvement of the solution is less than the preset threshold.

4 Experiments

In the experimental setup, we configured a hierarchical federated learning (HFL) system comprising one cloud server, four edge servers, and multiple terminal devices. Each terminal was assigned different transmission power levels and computational capacities. The maximum transmission power for the terminal devices was set to 20 dBm, and the maximum computational power of each terminal was set to 2 GHz. The number of CPU cycles required to process the data depends on the structure of the terminal devices' chips, and the effective switching capacitance for each terminal device was set to 1×10^{-27} . The total bandwidth allocated for each edge server was 10 MHz.

During the federated learning process, the training model used was a convolutional neural network (CNN) with a total parameter count of 1×10^5 . The CIFAR-10 dataset was used for training, which is an image dataset comprising 10 classes. For benchmarking purposes, we used the algorithm from reference [2] as the comparison algorithm in our experiments (Fig. 2).

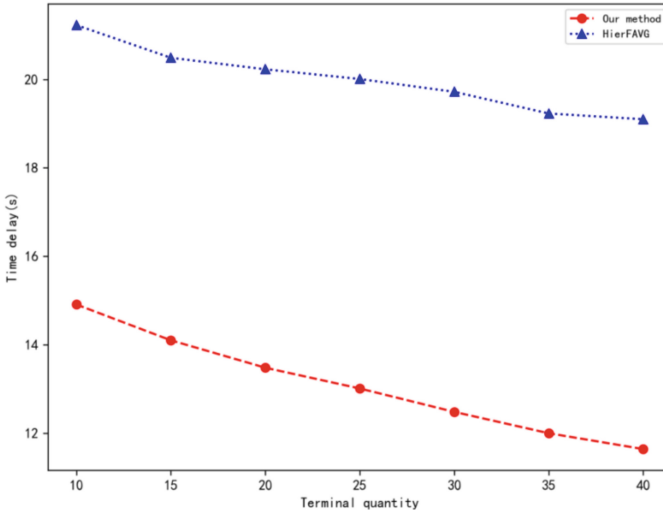


Fig. 2. Comparison of time delay in a round of global training

First, we compared the time taken for a single round of global aggregation between the two algorithms under different numbers of terminal devices. The results showed that our method has a significant advantage in terms of latency compared to HireFAVG. On the one hand, by offloading some computational tasks to the edge servers, our method fully utilizes the computational resources of the edge servers, accelerating the federated training process. On the other hand, the total time for global training in the HireFAVG method is affected by slower devices. We also observed that as the number of terminal devices increases, the reduction rate of the time required for a round of global aggregation in our method gradually slows down. This is because with an increasing number of terminals, the proportion of transmission time in the total global aggregation time increases, making it difficult to further reduce the overall training latency solely by decreasing the local training time (Fig. 3).

Next, we compared the energy consumption for executing a round of global training between the two algorithms under different numbers of terminal devices. Our observations showed that, as the number of terminals increases, the energy consumption for both algorithms exhibits an upward trend. This is primarily because, although the overall training task remains constant, introducing more terminals results in higher transmission energy consumption. Compared to HireFAVG, our method demonstrates a significant advantage in terms of energy efficiency. The modeled optimization problem effectively guides a more optimal task scheduling strategy and communication resource allocation.

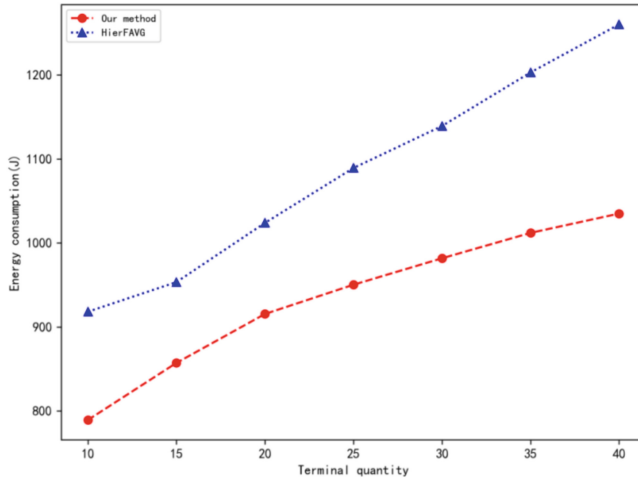


Fig. 3. Comparison of energy consumption in a round of global training

As the number of terminals increases, the energy consumption of our method approaches linear growth. This increase is mainly attributed to the additional communication energy required between the newly added terminals and the edge servers.

5 Conclusion

This paper proposes a joint optimization scheme for task scheduling and resource allocation in a hierarchical federated learning system. By offloading computational tasks from certain terminals to edge servers and allocating communication resources to terminal devices efficiently, the proposed approach improves federated learning efficiency and resource utilization. The task scheduling and resource allocation problem in the hierarchical federated learning system is modeled as a constrained optimization problem, and a hybrid heuristic optimization algorithm is designed for rapid problem-solving. Compared to the benchmark methods, our approach demonstrates significant advantages in both latency and energy consumption.

Acknowledgment. This study was supported by the grant of Guangxi Science and Technology Major Program (No. GuiKe AA23023035).

References

1. McMahan, H.B., Moore, E., Ramage, D., et al.: Communication-Efficient Learning of Deep Networks from Decentralized Data (2016). <https://doi.org/10.48550/arXiv.1602.05629>
2. Liu, L., Zhang, J., Song, S.H., et al.: Client-edge-cloud hierarchical federated learning. In: ICC 2020 - 2020 IEEE International Conference on Communications (ICC). IEEE (2020)

3. Saadat, H., Aboumadi, A., Mohamed, A., et al.: Hierarchical federated learning for collaborative IDS in IoT applications. In: 2021 10th Mediterranean Conference on Embedded Computing (MECO). IEEE, pp. 1–6 (2021)
4. Abdellatif, A.A., Mhaisen, N., Mohamed, A., et al.: Communication-efficient hierarchical federated learning for IoT heterogeneous systems with imbalanced data. *Futur. Gener. Comput. Syst.* **128**, 406–419 (2022)
5. Zhang, C., Xie, Y., Bai, H., et al.: A survey on federated learning. *Knowl.-Based Syst.* **216**, 106775 (2021)
6. Kairouz P, McMahan H B, Avenet B, et al. Advances and open problems in federated learning[J]. *Foundations and trends® in machine learning*, 2021, 14(1–2): 1–210
7. Farooq, M., Ngo, H.Q.: A low-complexity approach for max-min fairness in uplink cell-free massive MIMO. In: 2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring). IEEE, pp. 1–6 (2021)
8. Yang, Z., Chen, M., Saad, W., et al.: Energy efficient federated learning over wireless communication networks. *IEEE Trans. Wireless Commun.* **20**(3), 1935–1949 (2020)



A Smart Grid Planning Data Decision Scoring Model Based on Random Forest

Lijun Xiao^{1,2}, Xiangwei Meng¹(✉), Dezhi Han², and Dacheng He¹

¹ School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411100, China

ljxiaoxy@126.com, xiangwei_meng@126.com, wskdddn@163.com

² College of Information Engineering, Shanghai Maritime University, Shanghai 200000, China
dzhan@shmtu.edu.cn

Abstract. The changes in power load patterns and distribution locations pose significant challenges to the stability and long-term development of the smart grid system. Therefore, it is of great significance for power grid planning. The rationality of grid planning directly affects the stability and reliability of the smart grid system. This paper proposes a smart grid planning data decision scoring model based on random forest (RF-ETOPSIS). Firstly, a scoring mathematical model based on controller sorting was designed to score the grid planning data. Afterwards, a data decision-making method for grid planning based on adaptive random forest was proposed, which improves the accuracy of decision-making by constructing multiple decision trees. In the experimental phase, the proposed model shows 24.7% higher average performance than related models.

Keywords: Power Grid · Planning Data · Decision Scoring Model · Random Forest

1 Introduction

Smart grid integrated operation system is a comprehensive system integrating advanced information technology, communication technology, control technology and artificial intelligence algorithm [1, 2]. It monitors, analyzes and predicts the operating state of the power grid in real time, and automatically generates or assists the generation of power grid operation, maintenance and expansion decisions according to the operation objectives, constraints and optimization criteria of the power grid [3–5]. By adopting a centralized unified management mode, the integrated operation system can centrally monitor each decentralized power grid module, coordinate the allocation of resources among the modules, meet the power needs of users, ensure the stable operation of the power grid and the reliability of the power supply, reduce the system operating costs, and achieve the best economic benefits [6–8].

However, the dynamic nature of smart grid load, the uncertainty brought by various renewable resources and the variability of user behavior make its optimization problems particularly complicated, and the traditional constrained optimization methods based on

mathematics are often difficult to search for the optimal solution when dealing with such problems [9–12]. Therefore, the integrated scoring model can provide a more flexible and adaptable solution to help the integrated operating system make more accurate and efficient decisions in the dynamic and changing grid environment by evaluating the effects of different decisions.

The current comprehensive scoring models, such as fuzzy comprehensive evaluation method and TOPSIS method, have some applicability in some scenarios, but they have obvious limitations in dealing with the integrated operation of smart grid. Specifically, fuzzy comprehensive evaluation method can deal with uncertain information, it still has subjectivity in determining fuzzy membership degree and weight, and the calculation process is complex, which is difficult to adapt to the needs of real-time decision making [13–15]. The TOPSIS method assumes that all indicators are in the same direction, that is, the higher the better or the lower the better, but in a smart grid, some indicators may have an optimal range, rather than a simple extreme value. For example, the load rate of the grid is too high or too low, which can adversely affect the stability of the grid. These methods often ignore the feedback mechanism and time series changes within the power system, and fail to fully consider load fluctuations, resource uncertainties and user behavior changes, which makes it difficult to achieve accurate decision support in practical applications [16–18].

To deal with these issues, this paper proposes a smart grid planning data decision scoring model based on random forest (RF-ETOPSIS). The main contributions of this study are as follows:

- A mathematical model for smart grid planning decision scoring based on controller sorting is designed, and a terminal equipment data scoring model based on adaptive random forest is constructed.
- A multi-attribute decision-making method based on adaptive random forest is proposed, which can dynamically analyze the key performance indicators of the controller, including response time, reliability, energy consumption, etc.

The rest of this paper is as follows: Section 2 reviews the related work. Section 3 introduces the proposed decision scoring model. Section 4 shows the experimental results of the proposed model. Section 5 concludes this paper.

2 Related Work

In this section, we introduce the relevant research on random forest and smart grid decision system, respectively.

2.1 Random Forest

Random forest is a popular machine learning method for developing predictive models. Speiser et al. [19] proposed a technique to assess different variable selection in the setting of random forest classification, to determine the optimal method based on the application. Schoppa et al. [20] proposed a flood flow simulation model based on random forest. The results show that random forest has a strong competitiveness in middle and

low flood magnitudes simulation. Jackins et al. [21] proposed an intelligent prediction scheme for clinical diseases based on random forest, which classifies many diseases such as diabetes, heart disease, cancer, etc. to check whether patients are affected by the disease.

To accurately predict short-term power grid load, Fan et al. [22] proposed a hybrid model, that hybridizes the random forest and mean generating function. The authors used the time variable, random forest forecasting value, and forecasting value as input variables for the hybrid model. The experimental results indicate that the hybrid model has high accuracy in predicting short-term load. It is difficult for traditional methods to predict heart disease. Asadi et al. [23] proposed a prediction model combining the random forest and multi-objective particle swarm optimization. The model adopts an evolutionary multi-objective approach, with each training each tree having different samples and features. Chen et al. [24] proposed a risk assessment model based on random forest to realize early warning of security risks of large-scale group activities, and the classification accuracy of the assessment model reached 0.86.

Random forest is also widely used in economic forecasting scenarios. Yoon et al. [25] proposed a GDP growth prediction model based on random forest and gradient boosting model. The model uses the cross-validation process to choose the optimal hyper parameters. Liu et al. [26] proposed a lithium-ion battery classification framework based on random forest. The framework not only achieves reliable classification of electrode properties, but also effectively quantifies the manufacturing feature importance and correlations.

2.2 Smart Grid Decision System

Lu et al. [27] proposed a supplementary electricity price decision system based on reinforcement learning to minimize electricity payment and consumption dissatisfaction of grid end users. Integrating active distribution network functionality into distributed energy resources is a challenging problem. Therefore, Zadsar et al. [28] proposed an active distribution network operation decision framework based on two stage risk. The first stage is to consider the uncertainty of renewable resources, electricity prices, and customer loads. The second stage aims to keep the shed load at minimum level.

Islam et al. [29] proposed a centralized control strategy based on the Internet of Things, using differential evolution optimization algorithm to coordinate the distribution of electric vehicles and distributed energy sources. Jiang et al. [30] applied the heuristic intelligent neural decision support system to renewable energy systems, optimizing algorithm methods to improve local energy production and forecast energy.

3 Proposed Decision Scoring Model

In this section, we propose a load decision scoring model based on adaptive random forest, which integrates real-time data and dynamically adjusts decision parameters to achieve accurate evaluation of terminal load data. This model consists of four phases: 1) Data conversion phase: the terminal equipment data is preprocessed; 2) Weight calculation phase: using naive Bayes theory to determine the importance of controller attribute

parameters; 3) Utility evaluation phase: Utility evaluation of terminal equipment is carried out by multi-attribute decision-making method based on alternative scheme and compromise ranking to improve the stability of controller ranking; 4) Stability prediction phase: Adaptive random forest classifier is used to score the scores of smart grid terminal devices.

3.1 Data Conversion Phase

The choice of terminal devices can be defined as a multi-attribute decision problem with K terminal devices defined $T_E = \{T_{E1}, T_{E2}, \dots, T_{EK}\}$, where $K \geq 2$. Each terminal device has L properties $P = \{P_1, P_2, \dots, P_L\}$. At the same time, a performance matrix is constructed P_{KL} , where $K = \{1, 2, \dots, (X-1), \dots, X\}$, $L = \{1, 2, \dots, (Y-1), \dots, Y\}$:

$$P_{KL} = \begin{bmatrix} T_{E11} & T_{E12} & \cdots & T_{E1(Y-1)} & \cdots & T_{E1Y} \\ T_{E21} & T_{E22} & \cdots & T_{E2(Y-1)} & \cdots & T_{E2Y} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ T_{E(X-1)1} & T_{E(X-1)2} & \cdots & T_{E(X-1)(Y-1)} & \cdots & T_{E(X-1)Y} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ T_{EX1} & T_{EX2} & \cdots & T_{EX(Y-1)} & \cdots & T_{EXY} \end{bmatrix}_{X \times Y} \quad (1)$$

The sliding window average method is used to remove noise, and the outliers in the data set are detected and processed in real time. Different data sources may have different dimensions and numerical ranges, direct use of raw data may result in too much or too little weight for certain features, which can affect the final utility evaluation. Therefore, data need to be normalized to ensure that each feature's contribution to utility evaluation is fair. The further processed dimensionless performance matrix NP_{KL} is obtained by normalizing the performance matrix in the range $[0,10]$. The P_{KL} is normalized as follows:

$$NP_{KL} = \left(\frac{T_{EXY}}{IS_{T_{EXY}}} \right) * 10, \forall Y \text{ is a benefit attribute.} \quad (2)$$

$$NP_{KL} = \left(\frac{IS_{T_{EXY}}}{T_{EXY}} \right) * 10, \forall Y \text{ is a cost attribute.} \quad (3)$$

where, $IS_{T_{EXY}}$ is calculated in P_{KL} from the minimum value $\min_y T_E$ of the cost attribute and the maximum value $\max_y T_E$ of the benefit attribute.

3.2 Weight Calculation Phase

This section uses the data set to determine the weights of individual attributes, which influence the final result and the selection of alternatives. The traditional approach is to use feature selection methods in machine learning to obtain weights from data sets. These methods usually determine the weight based on the degree of influence of the independent variable on the dependent variable, and the more influential the feature, the

higher the weight [31]. However, in multi-attribute decision analysis problems, since each criterion may be treated as a dependent variable and no single criterion can be used as a benchmark for all other criteria, traditional feature selection methods may not be suitable. This is because criteria in a multi-attribute decision analysis problem are often interdependent, and the importance of each criterion to the decision may vary from situation to situation. In this study, we consider the interdependence of multiple response variables and criteria, and use naive Bayes method combined with real-time data to dynamically calculate important controller attribute parameters to make effective decisions.

First, the Tsallis-Havrda-Charvát entropy metric method is used to determine the initial weights of controller attribute parameters, as follows:

$$W_j^{THC} = \frac{e_j}{\sum_{j=1}^n e_j} \quad (4)$$

where, W_j^{THC} is the weight of the j -th controller property, and e_j is the entropy of the j -th controller property.

Then the projection distance B_{PM} between each reference value and the target value is calculated using the bidirectional projection metric method, as shown below:

$$B_{PM}(B_{pi}, T_{pi}) = \frac{||B_{pi}|| ||T_{pi}||}{||B_{pi}|| ||T_{pi}|| + |||B_{pi}|| - ||T_{pi}|| |(B_{pi} \cdot T_{pi})|} \quad (5)$$

where, $||B_{pi}||$, $||T_{pi}||$ are the Euclidean norms of the reference value and the target value, respectively.

The projection metric weights for each controller attribute are as follows:

$$W_j^{PM} = \frac{B_{PM}(B_{pi}, T_{pi})}{\sum_{j=1}^n B_{PM}(B_{pi}, T_{pi})} \quad (6)$$

Finally, use Bayesian formula to recalculate the weights of the criteria:

$$\mathcal{N}(\mathcal{T}_j/\mathbb{R}) = \frac{\mathcal{N}(\mathcal{T}_j)\mathcal{N}(\mathcal{T}_j/\mathbb{R})}{\sum_{j=1}^n \mathcal{N}(\mathcal{T}_j)\mathcal{N}(\mathcal{T}_j/\mathbb{R})} \quad (7)$$

where, $\mathcal{N}(\mathcal{T}_j) = \mathcal{N}_j$ represents the initial weight of the j -th attribute $\mathcal{T}_j$; $\mathbb{R}$ represents events, which obtain new weights for attributes through different weighting methods; $\mathcal{N}(\mathcal{T}_j/\mathbb{R})$ represents the recalculated attribute weights. Recalculate the attribute weights using the geometric mean method as input for the third stage, as shown below:

$$\mathcal{W}_j = \frac{w_j \omega_j}{\sum_{j=1}^m w_j \omega_j} \quad (8)$$

3.3 Utility Evaluation Phase

First, a multi-attribute decision method based on alternative and compromise ranking is used to evaluate the performance of terminal device controller parameters and capture the nonlinear relationship between them and ranking stability. Then, the utility value of each terminal device relative to best reference value (BRV) and worst reference value (WRV) can be obtained by calculation to distinguish its positive utility and negative utility. Then the utility function of each terminal device is established to comprehensively consider its positive utility relative to BRV and negative utility relative to WRV. Finally, a utility function matrix is obtained through utility function to achieve decision scoring of power grid planning data.

3.4 Stability Prediction Phase

Using UF_{KL} matrix as input, the adaptive random forest classifier is used to predict the stability of power grid planning, which mainly includes three parts: 1) The adaptive random forest classifier builds an adaptive decision tree on the data instances of terminal devices, 2) predicts the stability of each terminal device in real time, and 3) selects the best power grid planning by majority voting.

4 Experiment

4.1 Experimental Setup

The experiment uses Dell PowerEdge T150 server, equipped with Intel Xeon E-2378 processor (8-core 16 thread, 2.6 GHz clock speed, 16 M cache), supporting Turbo to 65 W and 3200 MT/s memory speed. Equipped with 32 GB ECC memory (4x8GB UDIMM) to ensure data accuracy. In terms of storage, three 2TB SATA 6 Gbps 7.2 K hard drives are used to support large amounts of data storage.

Table 1. Controller attribute parameter settings.

Classification	Property	Load	Signal strength	Sweep frequency
Reference value	3	50%	−70 dbm	1000 ms
Target value	10	80%	−30 dbm	100 ms

In the experiment, the reference value and target value of controller properties are set, as shown in Table 1.

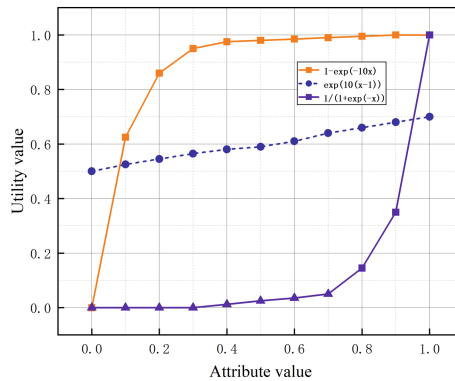
The benchmark value for controller load is set to 50%, which is the minimum available memory percentage, to avoid overload, and the target value is set to 80% to maintain stable system operation under ideal conditions while having sufficient margin to handle unexpected situations; The Bluetooth signal strength usually ranges from −70 dBm to −30 dBm; The benchmark value for scanning frequency is 100 ms, but the transmission effect may be insufficient. The target value is set at 1000 ms to improve the data update speed with a faster scanning frequency. Table 2 provides a detailed record of the attribute values of some controllers at a specific moment in time.

Table 2. Observation values of four controllers.

Number	Property	Load	Signal strength	Scanning frequency
1	9	75%	38 dbm	155 ms
2	8	40%	64 dbm	855 ms
3	8	54%	42 dbm	855 ms
4	10	69%	55 dbm	155 ms

4.2 Performance Comparison

Figure 1 shows a comparative analysis of increasing, decreasing, and linear functions. These functions can calculate the corresponding utility value based on the specific value of each indicator, providing a more accurate basis for decision scoring.

**Fig. 1.** Contrast of utility function.

To further verify the ranking stability of the proposed RF-ETOPSIS in dealing with controller availability, assume that the top two controllers in the smart grid system fail simultaneously. Figure 2 shows that the ranking stability of relevant decision model fluctuates significantly with the increase of the number of controllers, which is related to the maximum value of dependent controller attributes in the forward process. Once the maximum value is removed, the ranking results will change dramatically, resulting in unstable ranking stability. RF-ETOPSIS showed high ranking stability, and its fluctuation range remained near the level of 1. The stability of RF-ETOPSIS was improved by 24.7% compared to other model.

In addition, by comparing the average value of the first ranked controller attributes with the reference value, the experimental data obtained are shown in Fig. 3, which further indicates that the decision result of RF-ETOPSIS is reasonable.

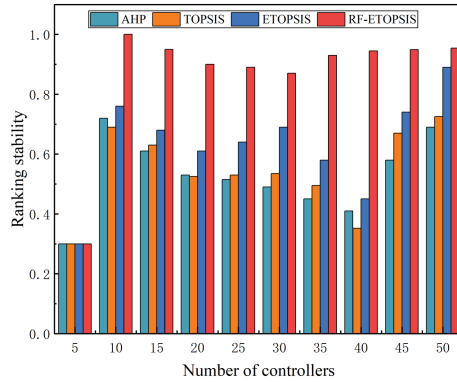


Fig. 2. First comparison of RF-ETOPSIS stability.

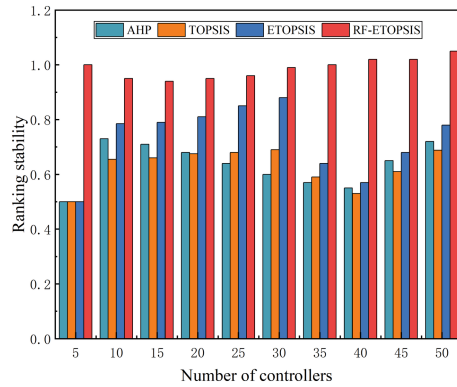


Fig. 3. Second comparison of RF-ETOPSIS stability.

5 Conclusion

In order to improve the stability of smart grid system, this paper proposes a smart grid planning data decision scoring model based on random forest. In the process of decision scoring, the adaptive random forest can comprehensively analyze performance indicators and provide optimal control decisions. The experimental results show that the performance of RF-ETOPSIS has improved by 24.7% compared to the relevant models.

Acknowledgements. This work was partially supported by the National Natural Science Foundation of China under Grants 62072170 and 61976087, the Science and Technology Project of the Department of Communications of Hunan Provincial under Grant 202101, the Key Research and Development Program of Hunan Province under Grant 2022GK2015, and the Hunan Provincial Natural Science Foundation of China under Grant 2021JJ30141.

References

1. Kakran, S., Chanana, S.: Smart operations of smart grids integrated with distributed generation: a review. *Renew. Sustain. Energy Rev.* **81**, 524–535 (2018)
2. Faheem, M., Shah, S.B.H., Butt, R.A., et al.: Smart grid communication and information technologies in the perspective of Industry 4.0: Opportunities and challenges. *Computer Science Review* **30**, 1–30 (2018)
3. Xie, K., Ouyang, Y., Wang, X., et al.: Deep adversarial tensor completion for accurate network traffic measurement. *IEEE/ACM Trans. Networking* **31**(5), 2101–2116 (2023)
4. Sarker, E., Halder, P., Seyedmahmoudian, M., et al.: Progress on the demand side management in smart grid and optimization approaches. *Int. J. Energy Res.* **45**(1), 36–64 (2021)
5. Rathor, S.K., Saxena, D.: Energy management system for smart grid: an overview and key issues. *Int. J. Energy Res.* **44**(6), 4067–4109 (2020)
6. Xiao, L., Han, D., Yang, C., et al.: TS-DP: an efficient data processing algorithm for distribution digital twin grid for industry 5.0. *IEEE Transactions on Consumer Electronics* (2023)
7. Kaur, D., Aujla, G.S., Kumar, N., et al.: Tensor-based big data management scheme for dimensionality reduction problem in smart grid systems: SDN perspective. *IEEE Trans. Knowl. Data Eng.* **30**(10), 1985–1998 (2018)
8. Liang, W., Xiao, J., Chen, Y., et al.: TMHD: twin-bridge scheduling of multi-heterogeneous dependent tasks for edge computing. *Futur. Gener. Comput. Syst.* **158**, 60–72 (2024)
9. Hafeez, G., Alimgeer, K.S., Khan, I.: Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid. *Appl. Energy* **269**, 114915 (2020)
10. Chung, H.M., Maharjan, S., Zhang, Y., et al.: Distributed deep reinforcement learning for intelligent load scheduling in residential smart grids. *IEEE Trans. Industr. Inf.* **17**(4), 2752–2763 (2020)
11. Zeeshan, M., Jamil, M.: Adaptive moth flame optimization based load shifting technique for demand side management in smart grid. *IETE J. Res.* **68**(1), 778–789 (2022)
12. Chen, X., Liang, W., Zhou, X., et al.: An efficient transmission algorithm for power grid data suitable for autonomous multi-robot systems. *Inf. Sci.* **572**, 543–557 (2021)
13. Wu, X., Hu, F.: Analysis of ecological carrying capacity using a fuzzy comprehensive evaluation method. *Ecol. Ind.* **113**, 106243 (2020)
14. Hu, Y., Wu, L., Pan, X., et al.: Comprehensive evaluation of cloud manufacturing service based on fuzzy theory. *Int. J. Fuzzy Syst.* **23**(6), 1755–1764 (2021)
15. Liao, M., Lu, Y., Li, X., et al.: An unsupervised image dehazing method using patch-line and fuzzy clustering-line priors. *IEEE Transactions on Fuzzy Systems* (2024)
16. Ciardiello, F., Genovese, A.: A comparison between TOPSIS and SAW methods. *Ann. Oper. Res.* **325**(2), 967–994 (2023)
17. Li, W., Dong, F., Ji, Z., et al.: Evaluation of provincial power supply reliability with high penetration of renewable energy based on combination weighting of game theory-TOPSIS method. *Sustainable Energy, Grids and Networks* **35**, 101092 (2023)
18. Liang, W., Li, Y., Xie, K., et al.: Spatial-temporal aware inductive graph neural network for C-ITS data recovery. *IEEE Trans. Intell. Transp. Syst.* **24**(8), 8431–8442 (2022)
19. Speiser, J.L., Miller, M.E., Tooze, J., et al.: A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst. Appl.* **134**, 93–101 (2019)
20. Schoppa, L., Disse, M., Bachmair, S.: Evaluating the performance of random forest for large-scale flood discharge simulation. *J. Hydrology* **590**, 125531 (2020)
21. Jackins, V., Vimal, S., Kaliappan, M., et al.: AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *J. Supercomput.* **77**(5), 5198–5219 (2021)

22. Fan, G.F., Zhang, L.Z., Yu, M., et al.: Applications of random forest in multivariable response surface for short-term load forecasting. *Int. J. Electr. Power Energy Syst.* **139**, 108073 (2022)
23. Asadi, S., Roshan, S.E., Kattan, M.W.: Random forest swarm optimization-based for heart diseases diagnosis. *J. Biomed. Inform.* **115**, 103690 (2021)
24. Chen, Y., Zheng, W., Li, W., et al.: Large group activity security risk assessment and risk early warning based on random forest algorithm. *Pattern Recogn. Lett.* **144**, 1–5 (2021)
25. Yoon, J.: Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach. *Comput. Econ.* **57**(1), 247–265 (2021)
26. Liu, K., Hu, X., Zhou, H., et al.: Feature analyses and modeling of lithium-ion battery manufacturing based on random forest classification. *IEEE/ASME Trans. Mechatron.* **26**(6), 2944–2955 (2021)
27. Lu, T., Chen, X., McElroy, M.B., et al.: A reinforcement learning-based decision system for electricity pricing plan selection by smart grid end users. *IEEE Transactions on Smart Grid* **12**(3), 2176–2187 (2020)
28. Zadsar, M., Sebtahmadi, S.S., Kazemi, M., et al.: Two stage risk based decision making for operation of smart grid by optimal dynamic multi-microgrid. *Int. J. Electr. Power Energy Syst.* **118**, 105791 (2020)
29. Islam, M.R., Lu, H., Hossain, M.J., et al.: An IoT-based decision support tool for improving the performance of smart grids connected with distributed energy sources and electric vehicles. *IEEE Trans. Ind. Appl.* **56**(4), 4552–4562 (2020)
30. Jiang, D., Zhu, W., Muthu, B.A., et al.: Importance of implementing smart renewable energy system using heuristic neural decision support system. *Sustainable Energy Technol. Assess.* **45**, 101185 (2021)
31. Zuo, Z., Zeng, Z., Su, W., et al.: Specification transformation method for functional program generation based on partition-recursion refinement rule. *Inf. Sci.* **633**, 613–632 (2023)



Feature Importance on Axial Compressive Strength Prediction of Concrete-Filled Steel Tubular Columns Using Random Forest Algorithm

Bo Su¹, Zhenhong Wu¹, Dongqi Jiang²✉ , Gang Bi³, Jiaxin Fan², and Kai Yan⁴

¹ China Harbor Engineering Company Ltd., Beijing 100027, China

² Nanjing University of Science and Technology, Nanjing 210094, China
jiangdongqi@njjust.edu.cn

³ Nanjing Tech University, Nanjing 211816, China

⁴ Suzhou Port and Shipping Business Development Center, Suzhou 215000, China

Abstract. Concrete-filled steel tubular (CFST) columns are widely used in the construction of civil engineering structures due to their excellent structural performance, which are able to utilize advantages of both concrete and steel materials. The axial compressive strength is considered to be one of the most critical parameters used in the design of CFST columns, and different formulae have been provided in codes and specifications to predict the design strength values. In this paper, 1581 CFST column concentric compressive load test data points are collected from a variety of scientific literature published in China, USA and Japan to establish an extensive database. Random forest (RF) algorithm is adopted to predict the axial compressive strength of CFST columns by using a randomly assigned test set from the database. Two types of features are selected as key input parameters in the development of RF model, including basic parameters and intermediate parameters. In particular, indirect parameters with physical meanings defined in Chinese design codes are used as input features. The comparison of the proposed RF model with the existing code-specified equations underscores the efficiency and accuracy of RF technique in the predication of axial compressive strength. Furthermore, higher prediction accuracy is achieved by machine learning techniques based on intermediate parameters than those based on original parameters.

Keywords: Axial Compression Capacity · Code-specified formulae · Machine Learning · Random Forest

1 Introduction

Concrete-filled steel tubular (CFST) column is a composite structural member formed by pouring concrete into the steel pipe, which has been widely used as key load carrying components in building structures due to the superior structural efficiency [1]. The confinement effect of the outer steel tube significantly enhances strength and deformability

of the concrete inside, and the local buckling induced failure of the steel tube itself can also be delayed with the help of the infilled concrete core. In particular, circular sections are more popular in the engineering practice as compared to other cross-section shapes of CSFT columns because of the superiority in the confinement effect and aesthetics [2], as shown in Fig. 1. The axial compressive strength is considered to be one of the most critical parameters used in the design of CFST columns. Different formulae have been proposed by scholars and also specified in codes to predict the design strength, but the format and accuracy vary from one equation to another.

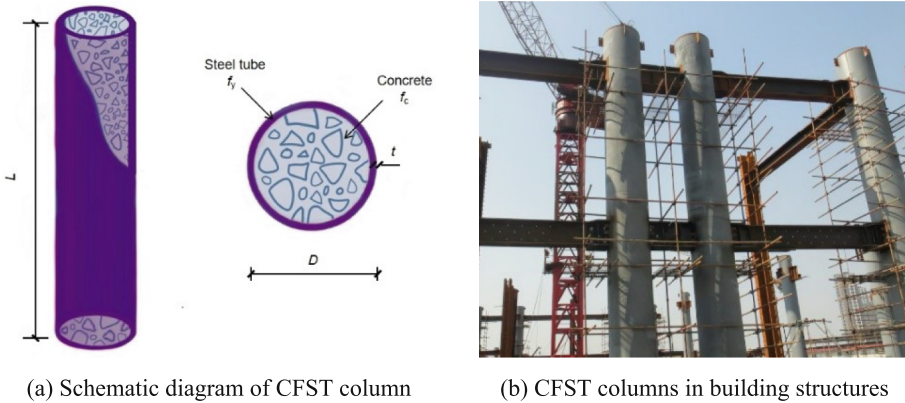


Fig. 1. CFST column and the application in building structures.

To achieve a quick and reliable prediction, this paper explores the application of the machine learning technique in determining the axial compressive strength value of CFST columns. As a multidisciplinary approach, machine learning techniques have been successfully applied in several research areas in the structural engineering. Mangalathu et al. [3] utilized artificial neural network (ANN) technique to identify the failure mode and predict the ultimate strength of reinforced concrete beams. Luo et al. [4] evaluated the effective stiffness of concrete columns with support vector machine (SVM) algorithm. Tran et al. [5] predicted the axial compression capacity of square concrete-filled steel tubular (SCFST) columns using the ANN technique. Jiang et al. [6, 7] and Xiao et al. [8] compared the efficiency and accuracy of several machine learning techniques in the identification of failure patterns of reinforced concrete-steel plate shear walls. Nguyen et al. [9] utilized the extreme gradient boosting (XGBoost) algorithm to predict the punching shear resistance of reinforced concrete interior slabs without shear reinforcement. Nikbin et al. [10] proposed an empirical formula for the prediction of concrete fracture energy based on the ANN technique. Nguyen et al. [11] developed a machine learning-based formulation to predict the shear capacity of squat flanged RC walls. In this paper, the random forest (RF) algorithm is used to predict CFST columns' axial compressive strength values, and the effect of feature selection on the prediction accuracy of RF models is further evaluated.

2 Experimental Database

Because it is quite difficult for one researcher to obtain sufficient experimental test data for machine learning algorithm training, a variety of scientific literatures published in China, Japan, U.S. and other countries are collected to assemble an extensive database consisting of 1581 concentric compressive test data points. The axial compressive strength of CFST columns are affected by geometric characteristics, material properties and other relevant parameters. Key basic parameters include the outer diameter of the circular cross section (D), the thickness of the steel tube (t), the length of the CFST column (L), the yield stress of the steel tube (f_y), and the concrete compressive strength (f_{ck}). Table 1 presents the minimum value, maximum value, mean value and standard deviation of aforementioned five basic parameters. In the established database for concentric loaded CFST samples, the outer diameter of the cross section (D) and the thickness of the steel tube (t) vary from 25.4 mm to 477.0 mm and 0.5 mm to 12.8 mm; the length of the CFST column (L) ranges from 150 mm to 4365.0 mm. The geometric dimensions are suitably used for low, medium and high-rise buildings. As for the material properties, the steel yield strength is up to 853.0 MPa and the concrete compressive strength is up to 145.0 MPa, which vary from the normal to the high strength of both steel and concrete materials.

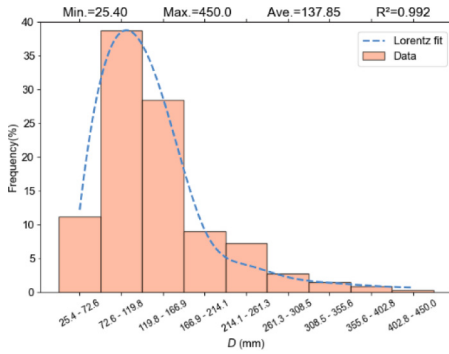
Table 1. Statistical information on basic parameters of CFST column database.

Parameters	D/mm	t/mm	L/mm	f_y/MPa	f_{ck}/MPa
Min	25.4	0.5	150	178.3	6.1
Max	477.0	12.8	4365.0	853.0	145.0
Mean	140.7	3.8	916.0	335.5	39.1
Stv	71.5	2.15	790.4	83.1	21.8

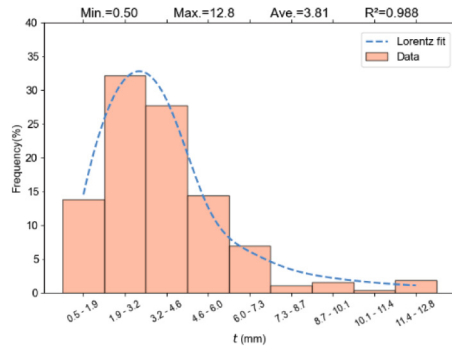
Figure 2 depicts statistical distributions of each considered variable together with the fitted distribution curve. It can be seen that the distribution of basic parameters is preferably described by two types of functions: Lorentz function (i.e. D , t , and L) and Lognormal function (i.e. f_y , and f_{ck}). Besides, several basic statistical features are also indicated in Fig. 2. In particular, the coefficient of determination R^2 is regarded as an evaluation criterion in the selection of the appropriate function to approximate distribution, which is expressed in Eq. (1). A more ideal approximation will be obtained when the value of R^2 is closer to 1.

$$R^2 = 1 - \frac{\sum_i (\hat{x}_i - x_i)^2}{\sum_i (x_i - \bar{x}_i)^2} \quad (1)$$

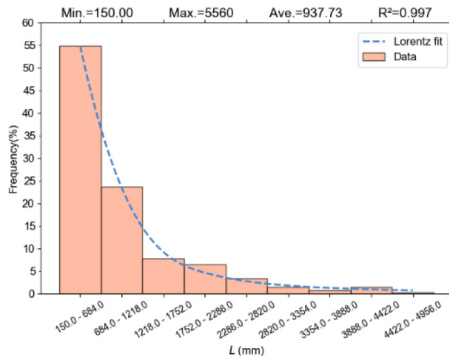
where $\hat{x}_i$ is the predicted value, x_i is the true value, and $\bar{x}_i$ is the mean value.



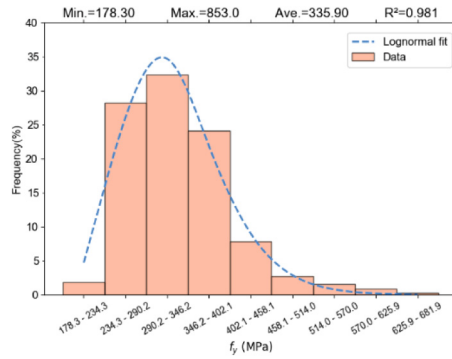
(a) Outer diameter, D (mm)



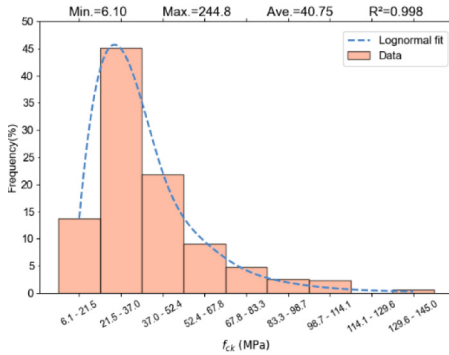
(b) Steel tube thickness, t (mm)



(c) Length, L (mm)



(d) Steel yield strength, f_y (MPa)



(e) Concrete compressive strength, f_{ck} (MPa)

Fig. 2. Frequency distribution diagram for basic parameters of CFST columns.

3 Machine Learning Technique

3.1 Random Forest (RF) Algorithm

As a powerful machine learning approach, RF generates an ensemble of decision trees and merges them together to achieve the prediction more accurately and stably, as shown in Fig. 3. By using a bootstrap sample of the training dataset, each decision tree is created independently, and then the average output value from individual trees in the forest is utilized to make the decision [6]. In the RF algorithm, extra randomness is brought in the selection of a subset of features. Thus, RF is able to search for the best feature when splitting a node. The RF model implementation generally provides a wide diversity and returns satisfactory predicted values.

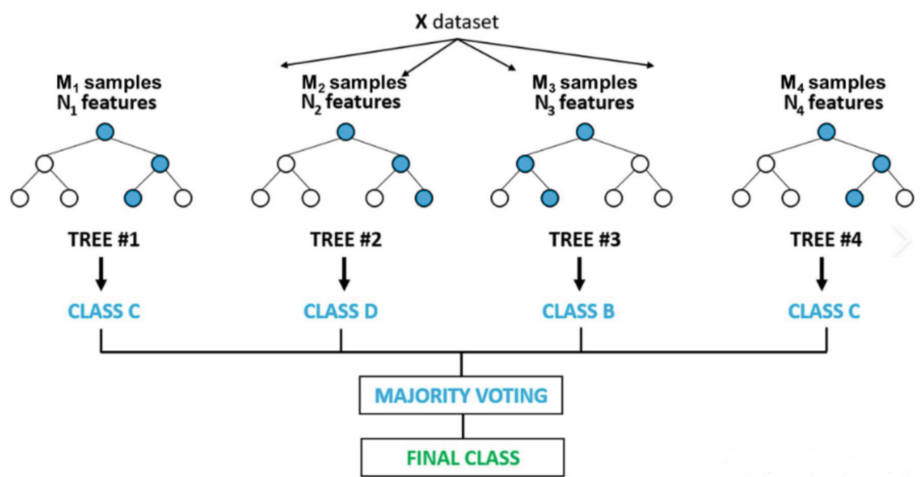


Fig. 3. Illustration of random forest algorithm [12].

3.2 Input Feature Selection

Appropriate feature selection causes a great influence on the prediction accuracy and computational cost of the machine learning technique. In terms of axial compressive strength prediction for CFST columns, two types of features can be selected as input variables in the development of RF model, including basic parameters and intermediate parameters. The basic parameters and the corresponding statistical properties have been illustrated in the previous section. As for the intermediate parameters, scholars and engineers can define them with specific physical meanings through the combination of basic parameters. Herein, we use indirect parameters specified in two Chinese design standards, CECS 254 (Technical specification for solid and hollow concrete-filled steel tubular structure) and CECS 28 (Technical specification for concrete-filled steel tubular structures), as input features. In the standard CECS 254, the design value of axial

compressive strength for the CFST column is determined by Eqs. (2)–(8):

$$N_0 = A_{sc}f_{sc} \quad (2)$$

$$f_{sc} = (1.212 + B\theta + C\theta^2)f_c \quad (3)$$

$$\alpha_{sc} = \frac{A_s}{A_c} \quad (4)$$

$$\theta = \alpha_{sc} \frac{f_y}{f_c} \quad (5)$$

$$N_u = \varphi N_0 \quad (6)$$

$$\varphi = \frac{1}{2\bar{\lambda}_{sc}^2} \left\{ \bar{\lambda}_{sc}^2 + (1 + 0.25\bar{\lambda}_{sc}) - \sqrt{\left[\bar{\lambda}_{sc}^2 + (1 + 0.25\bar{\lambda}_{sc}) \right]^2 - 4\bar{\lambda}_{sc}^2} \right\} \quad (7)$$

$$\bar{\lambda}_{sc} = \frac{\lambda_{sc}}{\pi} \sqrt{\frac{f_{sc}}{E_{sc}}} \approx 0.01\lambda_{sc}(0.001f_y + 0.781) \quad (8)$$

where A_{sc} is the CFST composite section area, f_{sc} is the CFST compressive strength, A_s and A_c are section areas of the steel tube and in-filled concrete, respectively, α_{sc} is the steel ratio, θ is the confinement coefficient, f_y is the steel yield strength, f_c is the concrete compressive strength, B and C are the shape coefficients on the confinement effect, φ is the stability coefficient, λ_{sc} is the slenderness ratio, $\bar{\lambda}_{sc}$ is the regular slenderness ratio, and E_{sc} is the CFST material modulus.

As for the standard CECS 28, Eqs. (9)–(13) are provided to calculate the design value of axial compressive strength for the CFST column:

$$N_u = \varphi_e \varphi_l N_0 \quad (9)$$

$$\begin{cases} N_0 = 0.9A_c f_c (1 + \alpha\theta), & \theta \leq \frac{1}{(\alpha-1)^2} \\ N_0 = 0.9A_c f_c (1 + \sqrt{\theta} + \theta), & \theta > \frac{1}{(\alpha-1)^2} \end{cases} \quad (10)$$

$$\theta = \frac{A_s f_y}{A_c f_c} \quad (11)$$

$$\varphi_e \varphi_l \leq \varphi_0 \quad (12)$$

$$\begin{cases} \varphi_l = 1, & \frac{L_e}{D} \leq 4 \\ \varphi_l = 1 - 0.115\sqrt{\left(\frac{L_e}{D} - 4\right)}, & \frac{L_e}{D} > 4 \end{cases} \quad (13)$$

where α is the factor related to the concrete strength grade, φ_e is the eccentricity coefficient, φ_l is the slenderness coefficient, and L_e is the effective column length.

Figure 4 shows parameter systems extracted from design equations provided by the standards CECS 254 and CECS 28. Variables marked with the blue box are selected as

intermediate parameters for RF model training. The CFST composite section area, A_{sc} , CFST compressive strength, f_{sc} , CFST material modulus, E_{sc} , and the slenderness ratio, λ_{sc} , are referred to CECS 254, while the concrete section area, A_c , slenderness coefficient, φ_l , concrete strength grade related coefficient, α , concrete compressive strength, f_{ck} , and confinement coefficient, θ , are defined in CECS 28.

Table 2 and Table 3 list the minimum values, maximum values, mean values and standard deviations of intermediate parameters referred to the standards CECS 254 and CECS 28, respectively. Figure 5 shows statistical distributions of intermediate parameters together with fitted distribution curves. Three main types of functions approximately describe the distribution of intermediate parameters: Lorentz function (i.e. A_{sc} , f_{sc} , E_{sc} , and θ), Gaussian function (i.e. λ_{sc}), and Lognormal function (i.e. A_c , φ_l , and f_{ck}). In addition, several statistical features are also indicated in Fig. 5.

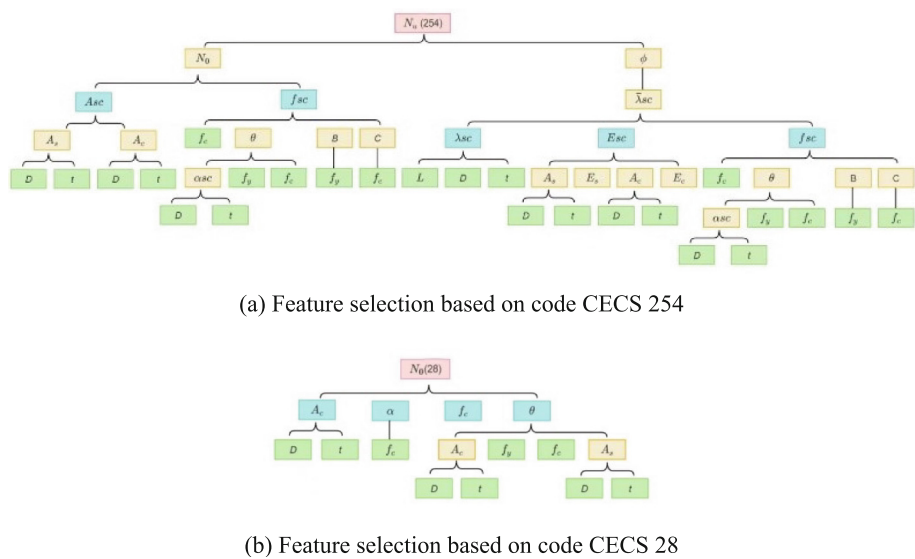


Fig. 4. Key features selection based on two Chinese design codes.

Table 2. Statistical information on code CECS 254-based intermediate parameters of CFST column database.

Parameters	A_{sc}/mm^2	f_{sc}/MPa	E_{sc}/MPa	λ_{sc}
Min	506.7	20.6	33806.6	0.03
Max	817128.2	364.1	158037.8	2.3
Mean	23304.6	94.5	71261.3	0.3
Stv	9160.9	85.9	77571.9	0.13

Table 3. Statistical information on code CECS 28-based intermediate parameters of CFST column database.

Parameters	A_c/mm^2	f_{ck}/MPa	α	θ	φ_l
Min	437.4	22.8	1.8	0.03	0.2
Max	786529.5	323.8	2.0	26.7	1
Mean					
	20798.7	93.5	1.9	1.6	0.9
Stv	42670.0	90.3	0.1	0.9	1

In the determination of the axial compressive strength of CFST columns, the correlations between different parameters should be carefully checked to avoid multicollinearity. Figure 6 shows the Pearson correlation coefficient matrix for all basic and intermediate parameters. The darker color with the coefficient closer to 1 reveals the higher correlation.

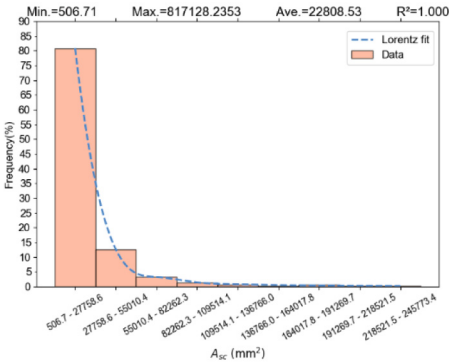
4 Results and Discussion

4.1 Accuracy Evaluation

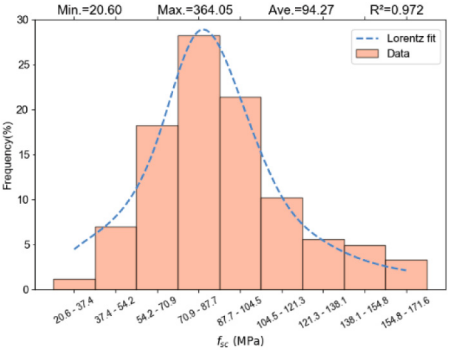
Figure 7 shows the comparison of axial compressive strength values determined by RF models and code formulae. Dash lines indicate different levels of errors in the figure. According to the calculated coefficient of determination R^2 , it is found that the proposed RF models show more accurate axial compressive strength prediction as compared to the existing code-specified equations. Furthermore, slightly higher prediction accuracy is achieved by RF algorithm techniques based on intermediate parameters than those based on original parameters.

4.2 Feature Importance

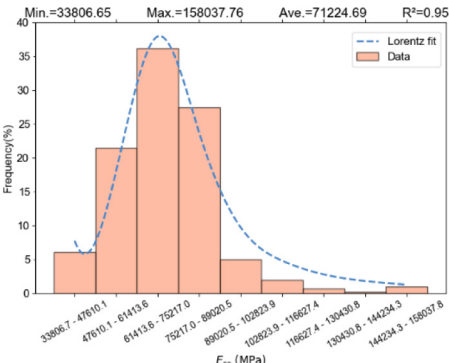
Shapley Additive exPlanations (SHAP) technique is further adopted herein to better explain the RF model and identify the importance of various input features. Figure 8 shows the SHAP value plot and Beeswarm plot for all basic and intermediate parameters. The range and distribution of input variables on the axial compressive strength prediction can be highlighted in the plots, which are colored according to their values from high (red) to low (blue). It can be seen that geometric dimensions are more important than material strength properties. In particular, the section area and outer diameter are two crucial parameters. Note that although the confinement coefficient is considered to be a key design variable in the design standards, the Beeswarm plot indicates its low level of importance.



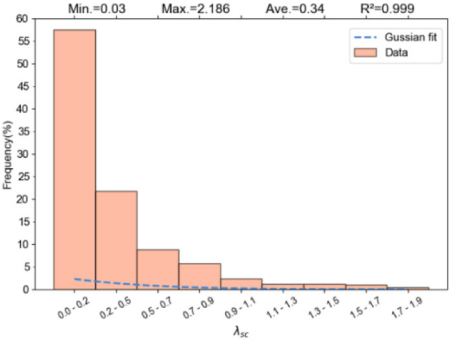
(a) Composite section area, A_{sc} (mm²)



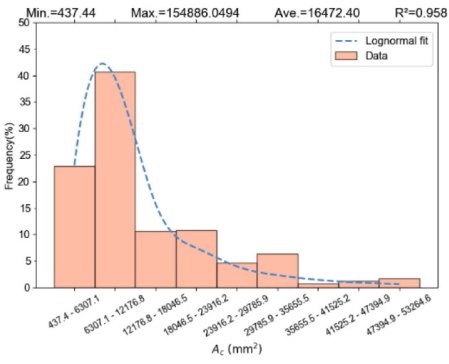
(b) CFST compressive strength, f_{sc} (MPa)



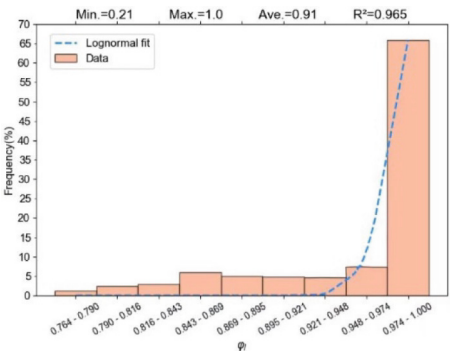
(c) CFST material modulus, E_{sc} (MPa)



(d) Slenderness ratio, λ_{sc}



(e) Concrete section area, A_c (mm²)



(f) Coefficient for slenderness ratio, φ_i (MPa)

Fig. 5. Frequency distribution diagram for intermediate parameters of CFST columns.

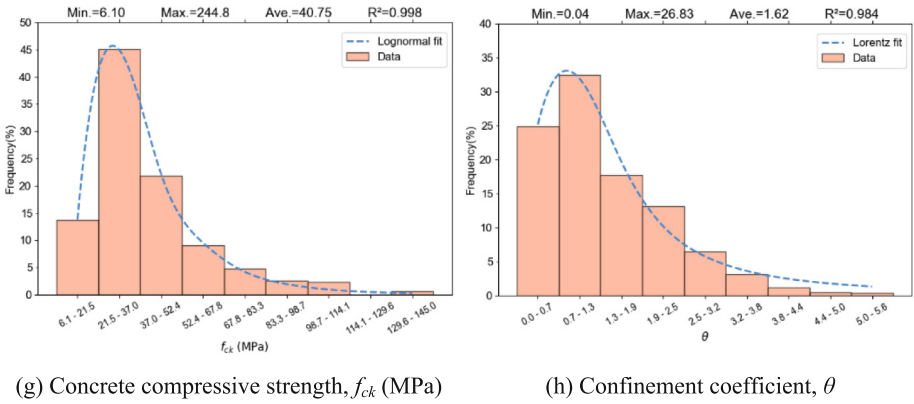


Fig. 5. (continued)

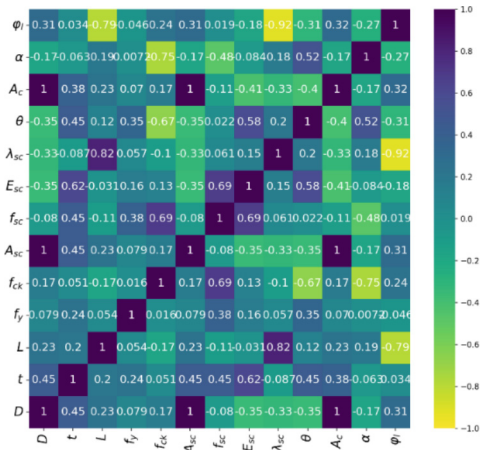


Fig. 6. Pearson correlation coefficient matrix.

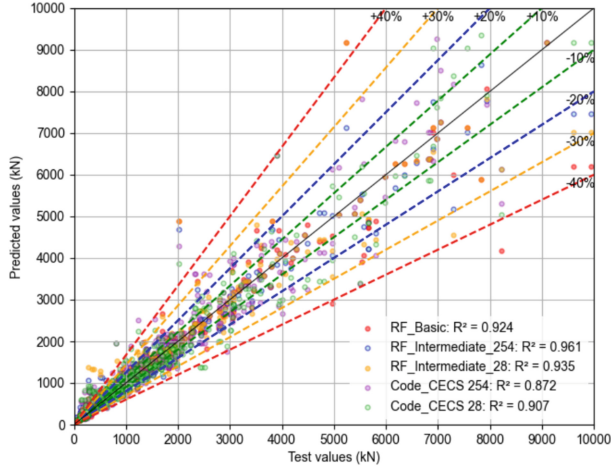
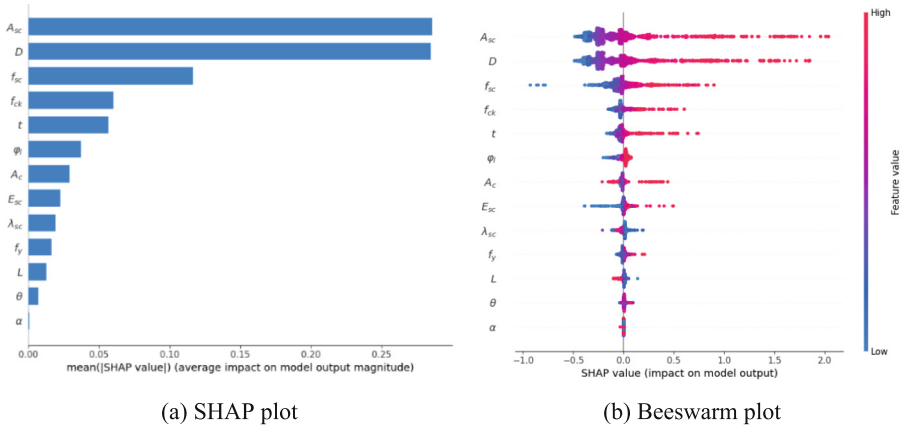


Fig. 7. Comparison between RF models and code formulae.



(a) SHAP plot

(b) Beeswarm plot

Fig. 8. Shapley Additive exPlanations (SHAP) summary plots

5 Conclusion

An extensive database consisting of 1581 CFST column axial compressive load test data points is established in this paper to develop appropriate RF models for predicting the design strength values. Both basic parameters and intermediate parameters are selected as key input features in the development of RF model. SHAP technique is further utilized to better explain the RF model and identify the importance of various input features. The comparison of the proposed RF model with the existing code-specified equations underscores the efficiency and accuracy of RF technique in the predication of axial compressive strength. In addition, higher prediction accuracy is achieved by machine learning techniques based on intermediate parameters than those based on original parameters.

Acknowledgement. This work is supported by National Natural Science Foundation of China (Grant No. 52478180, 51808292), Fundamental Research Funds for the Central Universities (Grant No. 30922010907), State Key Laboratory of Ocean Engineering (Shanghai Jiao Tong University) (Grant No. GKZD010089) and Natural Science Foundation of Jiangsu Province (Grant No. BK20180487).

References

1. Wang, X., Fan, F., Lai, J.: Strength behavior of circular concrete-filled steel tube stub columns under axial compression: a review. *Constr. Build. Mater.* **322**, 126144 (2022)
2. Chen, P., Wang, Y., Zhang, S.: Size effect prediction on axial compression strength of circular CFST columns. *J. Constr. Steel Res.* **172**, 106221 (2020)
3. Mangalathu, S., Hwang, S., Jeon, J.: Failure mode and effects analysis of RC members based on machine-learning-based shapley additive explanations (SHAP) approach. *Eng. Struct.* **219**, 110927 (2020)
4. Luo, H., Paal, S.: A locally weighted machine learning model for generalized prediction of drift capacity in seismic vulnerability assessments. *Computer-Aided Civil and Infrastructure Eng.* **34**(4), 935–950 (2019)
5. Tran, V., Thai, D., Kim, S.: Application of ANN in predicting ACC of SCFST column. *Compos. Struct.* **228**, 111332 (2019)
6. Jiang, D., Liu, S., Chen, T., Bi, G.: Machine learning-based failure mode identification of RCSPSW. In: *IABSE Congress: Resilient technologies for sustainable infrastructure*, Christchurch, New Zealand, 3–5 February 2021, pp. 1150–1157 (2021)
7. Jiang, D., Xiao, C., Chen, T., Zhang, Y.: Experimental study of high-strength concrete-steel plate composite shear walls. *Appl. Sci.* **9**, 2820 (2019)
8. Xiao, C., Tian, C., Chen, T., Jiang, D.: Compression-bending behavior of steel plate reinforced concrete shear walls with high axial compression ratio. In: *15th World conference on earthquake engineering*, Lisbon, Portugal, 24–28 September 2012, pp. 24–28 (2012)
9. Nguyen, H., Truong, G., Shin, M.: Development of extreme gradient boosting model for prediction of punching shear resistance of r/c interior slabs. *Eng. Struct.* **235**, 112067 (2021)
10. Nikbin, I., Rahimi, S., Allahyari, H.: A new empirical formula for prediction of fracture energy of concrete based on the artificial neural network. *Eng. Fract. Mech.* **186**, 466–482 (2017)
11. Nguyen, D., Tran, L., Ha, D., Nguyen, V., Lee, T.: A machine learning-based formulation for predicting shear capacity of squat flanged RC walls. *Structures* **29**, 1734–1747 (2021)
12. Innovation Yourself. <https://innovationyourself.com/random-forest-classification>. Accessed 10 Sep 2024



AEDs Treatment and Cognition Analysis of BECTS Through EEG Features

Zhezhaoh Zhang^{1,2}, Jiagui Chen³, Dinghan Hu^{1,2}(✉), Jiajia Fang⁴,
and Jiuwen Cao^{1,2}(✉)

¹ Machine Learning and I-health International Cooperation Base of Zhejiang Province, Hangzhou Dianzi University, Hangzhou 310018, China
{hdh, jwcao}@hdu.edu.cn

² Artificial Intelligence Institute, Hangzhou Dianzi University, Hangzhou 310018, Zhejiang, China

³ Hangzhou Zhichuang Information Technology Company Limited, Hangzhou 310050, Zhejiang, China

⁴ Department of Neurology, The Fourth Affiliated Hospital, Zhejiang University, Yiwu, Zhejiang Province, China

Abstract. Treatment with anti-epileptic drugs (AEDs) exerts a certain degree of ameliorative effect on brain activity and cognitive functioning. This study presents a comprehensive analysis of electroencephalogram (EEG) features during and after AEDs treatment for Benign Epilepsy of Childhood with Centrottemporal Spikes (BECTS) patients. Experimental results demonstrated an increase in α band activity during AEDs treatment and a decline in cognitive performance accompanied by a high kurtosis value. Following the treatment with AEDs, there was a reduction in α band activity and an improvement in cognitive performance, accompanied by a narrowing of the inter-individual differences. The results indicate that AEDs may enhance cognitive function by suppressing α band activity, which facilitating the guided use of AEDs and cognitive enhancement.

Keywords: Electroencephalogram · Anti-epileptic drugs · Cognitive function · Feature analysis

1 Introduction

Epilepsy exhibits close association with the cognitive development [1,2], especially Benign Epilepsy of Childhood with Centrottemporal Spikes (BECTS) [3].

This work was supported in part by the National Key Research and Development Program of China under Grants 2021YFE0100100, in part by Zhejiang Provincial Natural Science Foundation of China under Grant No. LZ24F030010, in part by the key projects of major health science and technology plan of Zhejiang Province (WKJ-ZJ-2129), and in part by the National Natural Science Foundation of China under Grant U1909209.

© The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025
H. Jin (Ed.): IAIC 2024, CCIS 2451, pp. 201–211, 2025.
https://doi.org/10.1007/978-981-96-7188-5_18

Patients may experience cognitive impairment and subsequently receive a diagnosis of epilepsy [4]. A significant number of frequent short-term epilepsy discharges (IEDS) are observed with high frequency during non-rapid eye movement sleep (NREM) in EEG of BECTS [3]. These epileptic discharges have been linked to a decline in cognitive function in these patients [5], such as memory problems [6]. The cognitive impairments may be linked to diminished functional connectivity and structure of the brain [7]. Anti-epileptic drugs (AEDs) represent the primary treatment option for BECTS [8,9]. For the early years of BECTS, AEDs contribute to the modulation of enhanced brain connectivity [10]. The impact of AEDs on sleep and cognitive function is complex and contingent upon the specific context. While some studies have demonstrated the potential benefits of AEDs in improving sleep quality [11], others have reported adverse effects [12]. Consequently, the scientific analysis of the EEG across different stages of AEDs treatment has the potential to inform subsequent treatment strategies and facilitate cognitive enhancement, thereby improving the orientation of clinical and therapeutic interventions.

Currently, the principal study on AEDs conditions is that of clinical outcomes. AEDs are effective in the management of epilepsy [13], and most of patients will ultimately require treatment with multiple antiepileptic drugs (AEDs) [14]. Nevitt et al. demonstrates that for individuals with focal onset seizures, carbamazepine and lamotrigine, and levetiracetam (LEV) exhibit the most favourable profile in terms of treatment seizure [15]. Brivaracetam (BRV) is also a type of AEDs as an adjunctive therapy for epileptic individuals [16]. Operto et al. observed that patients treated with LEV exhibited a notable enhancement in cognitive performance [17,18]. Li et al. demonstrated that the cognitive abilities of patients are below the level observed in typically developing children, as measured by the Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV) [19]. It can be reasonably deduced from the findings of Strzelczyk et al. that LEV potentially exert beneficial influences for cognitive functioning [20]. It is evident that there is a notable absence of literature examining the relationship between EEG and cognitive outcomes in patients undergoing treatment with AEDs. With the aim of attaining a profound comprehension of the AEDs influence on EEG, we carried out the subsequent study: (1) Group analysis: The different stages of AEDs treatment and cognition groups are analysed. Our findings indicate that while individuals within the α band who are undergoing AEDs treatment with poor cognition may exhibit heightened susceptibility, this could be evidence to suggest that they may benefit from treatment. (2) Individualized analysis: Individualized analyses were conducted within different stages of AEDs groups. The investigation confirmed that significant inter-individual differences were observed among individuals undergoing AEDs treatment, suggesting that the efficacy of AEDs varies among patients.

Figure 1 shows the experimental flowchart of this paper, the purpose of which is to analyze the effect of AEDs across various stages and levels of cognition for BECTS patients. It consists of three primary components: preprocessing, feature extraction, and analysis. Preprocessing involves several steps, including

filtering, channel selection and fragmentation. All features consist of statistical parts, time-domain elements, frequency-domain elements, and nonlinear parts are extracted. Statistical Analysis such as Mann-Whitney U test and linear correlation analysis are employed.

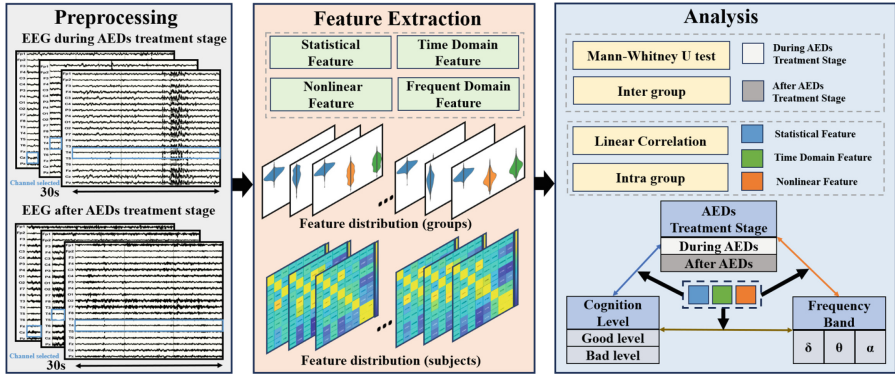


Fig. 1. The flowchart of experiment aims to analyze the effect of AEDs across various stages and levels of cognition for BECTS patients.

2 Methods and Results

2.1 Dataset

The EEG dataset was collected from the Fourth Affiliated Hospital of Zhejiang University. The individuals are diagnosed with BECTS, who were monitored for approximately eight hours NREM sleep. The dataset comprises EEG recording from two groups of patients: 10 patients undergoing AEDs treatment and 10 patients who have completed their AEDs treatment. The group undergoing treatment is denoted as Group 1, while the completion treatment is denoted as Group 2. The patients were administered a Wechsler Intelligence Test under the guidance of a clinical psychologist. The final test results yielded 4 indexes for each subject: Verbal Comprehension Index (VCI), Perceptual Reasoning Index (PRI), Working Memory Index (WMI), and Processing Speed Index (PSI). Besides, Full-Scale IQ (FSIQ) score is usually used to evaluate the overall cognition for tester. The sum of the 4 indexes represents the cognitive level. When FSIQ is greater than 85 and the sum of the 4 indexes is less than 400, it indicates poor cognition. Conversely, when FSIQ exceeds 85 and the sum of the 4 indexes surpasses 400, it means good cognition. Table 1 lists the dataset details, including information on age, sex, selected channel, FSIQ, VCI, PRI, WMI and PSI.

Table 1. Dataset Specifications

Group	Subject	Age	Sex	Channel	FSIQ	VCI	PRI	WMI	PSI
Group1	#1	8	M	C4	92	102	100	91	77
	#2	9	F	Cz	96	108	92	97	86
	#3	9	M	Cz	93	103	92	88	95
	#4	8	F	C4	98	108	104	88	83
	#5	8	M	C4	85	86	94	85	86
	#6	16	F	C4	99	100	122	88	74
	#7	7	F	T4	113	106	127	97	102
	#8	15	M	T4	126	120	130	109	116
	#9	7	M	C4	113	120	102	85	131
	#10	14	M	Cz	108	108	122	94	95
Group2	#1	14	F	T4	91	94	82	91	110
	#2	12	M	T4	99	100	116	88	83
	#3	15	F	C4	112	122	114	88	104
	#4	13	F	Cz	130	129	127	109	122
	#5	14	F	T4	102	103	104	94	102
	#6	12	M	Cz	129	122	125	109	128
	#7	13	M	Cz	103	122	96	103	89
	#8	10	M	Cz	108	137	100	100	83
	#9	14	M	Cz	113	126	120	91	92
	#10	9	M	T4	107	103	116	97	104

2.2 EEG Preprocessing and Feature Extraction

EEG signals are preprocessed according to the following steps: (1) Signal filtering: An effective from 0.5 to 70 Hz band-pass filter, and 50 Hz notch filter are employed to eliminate artifact interference. (2) Channel selection: It is determined that a lower standard deviation is deemed less susceptible to interference for channel. “Channel” in Table 1 refers to the channel selected by standard deviation for each EEG channel. (3) Data slicing: The selected EEG channel is segmented into 30-second intervals with the 50% overlap to facilitate subsequent feature extraction. The numbers of samples during and after AEDs treatment are 9113 and 13279, respectively.

From the preprocessed EEG data, we extracted 11 features for each segment: 3 features in statistics including mean (μ), variance (var) and median (median); 2 features in the time domain: namely skewness (ske) and kurtosis (kur); 3 relative power spectral densities (PSD) in the frequency domain: denoted as $p(\delta)$, $p(\theta)$ and $p(\alpha)$, and 3 nonlinear features: Hurst exponent (Hurst), approximate entropy (ApEn) and sample entropy (SampEn). The formulas for these features are shown in Table 2, where x represents discrete EEG sequence and N denotes the number of sample points.

Table 2. Features Description

Feature	Equation
μ	$\frac{1}{N} \sum_{i=1}^N x(i)$
Median	$\begin{cases} x(\frac{i+1}{2}) & \text{if } i \text{ is odd} \\ \frac{1}{2}(x(\frac{i}{2}) + x(\frac{i}{2} + 1)) & \text{if } i \text{ is even} \end{cases}$
Var	$\frac{\sum_{i=0}^N (x(i) - \mu)^2}{N-1}$
Ske	$\frac{\frac{1}{N} \sum_{i=1}^N (x(i) - \bar{x})^3}{[\frac{1}{N} \sum_{i=1}^N (x(i) - \bar{x})^2]^{\frac{3}{2}}}$
Kur	$\frac{\frac{1}{N} \sum_{i=1}^N (x(i) - \bar{x})^4}{[\frac{1}{N} \sum_{i=1}^N (x(i) - \bar{x})^2]^2}$
p_δ	$\frac{1}{ML} \sum_{i=1}^L \left \sum_{n=0}^{M-1} x_m^i(n) e^{-j\omega_\delta n} \right ^2$
p_θ	$\frac{1}{ML} \sum_{i=1}^L \left \sum_{n=0}^{M-1} x_m^i(n) e^{-j\omega_\theta n} \right ^2$
p_α	$\frac{1}{ML} \sum_{i=1}^L \left \sum_{n=0}^{M-1} x_m^i(n) e^{-j\omega_\alpha n} \right ^2$
Hurst	$\frac{\log(R/S)}{\log(T)}$
ApEn	$\Phi^{d_E-1}(r) - \Phi^{d_E}(r)$
SampEn	$\log \left(\frac{B_r^m}{B_r^{m+1}} \right)$

2.3 Analysis

Statistical Analysis. In order to assess the differences observed during and after AEDs treatment, the Mann-Whitney U test was utilized, considering the non-normal distribution and unequal variances of the EEG features. The results consistently yielded p to assess the differences observed during and after AEDs treatment-values below 0.001, indicating a statistically significant difference. Figure 2 illustrates the distribution of relative PSD ratio in δ , θ , and α bands, which indicates EEG α band activity is heightened during AEDs treatment. It follows that AEDs perform epilepsy treatment by suppressing α activity of the EEG. Following the treatment with AEDs, a significant decrease in EEG α activity is observed, accompanied by a substantial reduction in the incidence of spikes. These findings suggest that AEDs exert their therapeutic effect on epilepsy by suppressing the high-frequency activity observed in the EEG.

With a view to comprehend the effects of AEDs treatment, we conducted frequency band-specific analyses, building upon the results obtained in the previous study. The features extracted for each frequency band, aside from those in the frequency domain, were subsequently analyzed using the Mann-Whitney U test. The numerical results are shown in Table 3. Nearly all features showed significant differences within δ band, with the exception of mean ($p = 0.619$) and Hurst ($p = 0.048$). The results indicate that EEG activity undergoing and following AEDs treatment exhibits a high degree of similarity with respect to long-term memory in δ . In θ and α bands, only the mean and variance exhibited no significant differences between groups, with p -values less than 0.001 for the remaining features. There was a continued prevalence of α activity and large

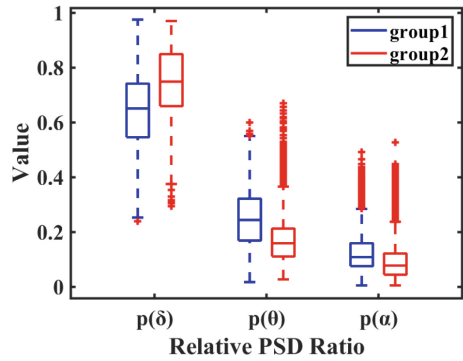


Fig. 2. Relative PSD Ratio distribution of δ , θ and α . Blue and red line represent distribution during and after AEDs treatment, respectively. (Color figure online)

Table 3. Results of Mann-Whitney U test

Feature	μ	Var	Median	Ske	Kur	Hurst	ApEn	SampEn
p_{δ}	0.619	< 0.001	< 0.001	< 0.001	< 0.001	0.048	< 0.001	< 0.001
p_{θ}	0.352	< 0.001	0.015	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
p_{α}	0.056	< 0.001	0.147	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

number of spikes in EEG, which could be exploited as a means of identifying an appropriate treatment strategy for patient during AEDs treatment stages.

Correlation Analysis. A linear correlation analysis was conducted to examine the relationship between different features intra each groups, during AEDs treatment, after AEDs treatment, respectively. The Pearson correlation coefficient, which ranges from -1 to $+1$, was employed, quantifying the correlation for each features. Values ranging from $+1$ to -1 indicate a gradual transition from positive to negative correlation. The majority of features exhibited a low correlation, suggesting minimal redundancy among them. Entropy-based features demonstrated a negative correlation with time domain features, particularly in θ and α bands during the treatment stages of AEDs. Entropy-based features reflect the complexity and irregularity of EEG, whereas time-domain features represent the waveform of EEG. Based on the aforementioned analysis, it can be inferred that EEG in α band is more active during AEDs treatment, resulting in enhanced complexity and the presence of numerous spikes. The irregularity of EEG is closely related to the dominant waveform.

Cognition Analysis. BECTS patients undergoing and after AEDs treatment are categorised into poor and good cognition groups based on WISC-IV scores. This section examines the relationship between kurtosis and cognitive levels within the same AEDs treatment group. Kurtosis can be employed to describe

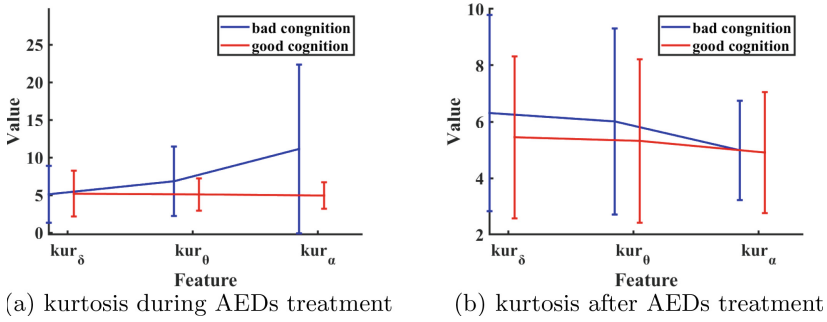


Fig. 3. Distribution and trend of kurtosis in δ , θ and α . The blue line represents poor cognition, while the red line represents the distribution good cognition. (Color figure online)

the signal peak sharpness [21] to reflect the waveform. As illustrated in Fig. 3, it is notable that during AEDs treatment, differences between those with good and poor cognition are observed in α band. Patients with poor cognition exhibited an average kurtosis of approximately 10, in contrast to approximately 5 for those with good cognition, indicating a significant gap of approximately 5 between the groups. Patients with poor cognition exhibit higher values, indicating that the peak in EEG, such as epileptic discharges, is more pronounced in patients with poor cognition. Furthermore, as frequency bands increase, kurtosis values for patients with poor cognitive abilities exhibit a marked upward trend, while those for patients with good cognitive abilities show a slight downward trend. This trend could be attributed to heightened high-frequency α band activity, such as epileptic discharges, which significantly impair cognition in patients with poor cognitive abilities. Similarly, after AEDs treatment stages, minimal disparity is observed between patients with poor and good cognitive abilities across all three frequency bands. It can be observed that AEDs treatment appears to suppresses these α activity, such as epileptic discharges, which has the effect of impacting patients cognition to some extent.

Individual Analysis. With the aim of examining the relevance of different patients, we conducted a study of the EEG features during and after AEDs treatment, utilising kurtosis and SampEn. Figure 4 shows the average feature values for each patient across three frequency bands, with individuals on the x-axis, features on the y-axis, and feature values represented by bar height. Figure 4 (a) illustrates the variations in kurtosis in α band among individuals during AEDs treatment, with most individuals having an average value around 5. However, case1, case4 and case5 exhibit considerably elevated bar heights in comparison to the others. Similarly, the average values of kurtosis in θ for these cases are slightly higher. In Fig. 4 (b), case1 and case4 stand out from the others regarding SampEn in α , despite the majority of individuals having average values of approximately 0.4 and 0.3. SampEn can be used to describe the complexity of

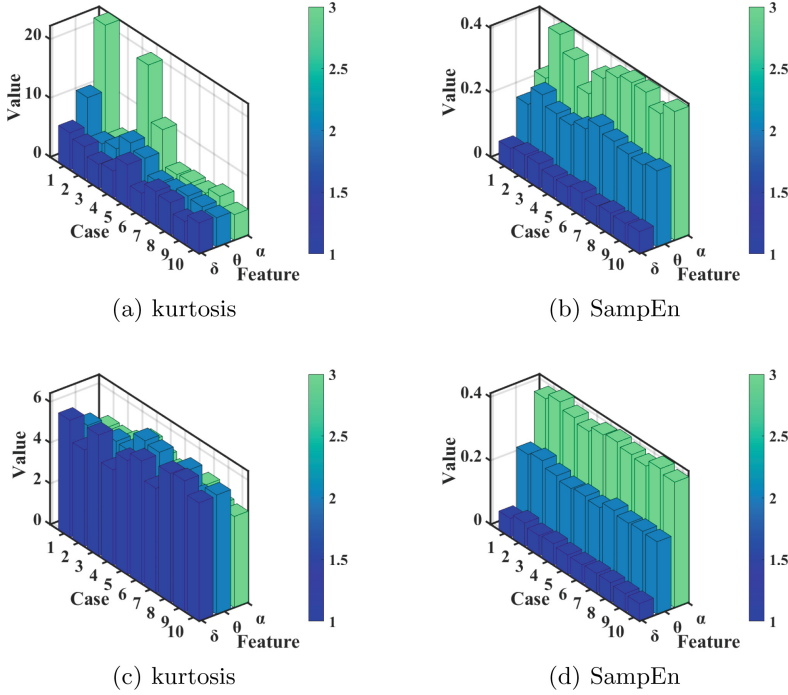


Fig. 4. Average value of features among individuals in δ , θ and α . (a) and (b) represent kurtosis and SampEn during AEDs treatment. (c) and (d) represent kurtosis and SampEn after AEDs treatment.

EEG [22]. The observed variability within the group may be attributed to differing levels of AEDs tolerance among patients, which in turn resulting in varying effects on α EEG activity. Following AEDs treatment, there is a reduction in the variability among individuals. Figure 4 (c) and (d) demonstrate that kurtosis value in three frequency bands is slightly uneven among patients, but the bar heights of SampEn are almost the same. This indicates that the patient's condition has improved, resulting in more stable and normal EEG brain activity, particularly in α band.

The probability density curves of SampEn is plotted as Fig. 5 to understand the impact of AEDs for individuals deeply. The probability density distribution of SampEn for θ is more dispersed, spanning a larger range. In case 7, the values range from 0.15 to 0.3 and the distribution is relatively sharp. For α , half of the individuals are centred between 0.25 and 0.5. Besides, a few are extremely dispersed. These observations demonstrate differences in θ and α among individuals during AEDs treatment. SampEn metric serves as an indicator of complexity. Therefore, when it exhibits higher complexity during AEDs treatment, it signifies heightened activity within the EEG [23]. Following treatment, there is a reduction in inter-subject variability. The probability density distributions of

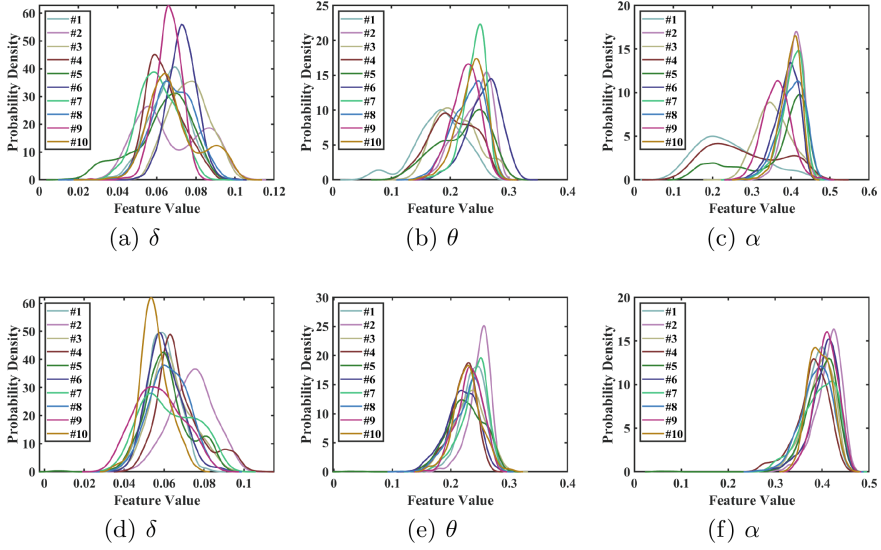


Fig. 5. Probability density distribution of SampEn in δ , θ and α . (a), (b) and (c) represent distribution during AEDs treatment. (d), (e) and (f) represent distribution after AEDs treatment.

SampEn for individuals following AEDs treatment exhibit a greater degree of concentration than those observed during AEDs treatment, particularly in α band. The distributions exhibit a tendency to range between 0.3 and 0.5.

3 Conclusions

Our study sought to explore the effects of AEDs on BECTS patients. All 20 patients with BECTS were kindly included in this study, consisting of 10 patients undergoing AEDs treatment and 10 patients who had completed AEDs treatment. A total of features were thoughtfully extracted for our analysis, including 3 statistical features, 2 time-domain features and 3 nonlinear features, which were utilized for both statistical and individual analyses. Furthermore, the patients were categorized into two groups according to their cognitive performance, assessing by the WISC-IV scores, with the objective of conducting a cognitive analysis. The findings of the experimental study yielded several significant insights: (1) EEG activity in α band was found to be more pronounced during the stages of AEDs treatment compared to post-treatment. This finding can provide a significant foundation for customizing the treatment related epilepsy. (2) The complex EEG activity observed during AEDs treatment exhibited a higher degree of inter-individual variability compared to that observed post-treatment. Therefore, it is important to develop an individualized treatment plan for each patient, taking into consideration their specific AEDs tolerance. (3) In the exploration of cognitive groupings, kurtosis has proven to be

a noteworthy indicator of visible differences between individuals with good and poor cognition. The presence of α band activity may result in alterations to the structure of the brain network, subsequently exerting an impact on the patient's cognitive abilities. This study may serve as a valuable reference for future AEDs treatment and provide insights into the cognitive changes associated with AEDs for BECTS patients. However, the type of epileptic syndrome categories analysis of AEDs effects was not conducted as the limited sample size. Future studies could explore the role of AEDs for cognitive development in greater detail.

Ethical Standards. This study has been approved by the Ethic Committee of the Forth Affiliated Hospital of Zhejiang University. All patients gave their informed consent prior to their inclusion in the study.

References

1. Chai, X., Xiao, Z., Zhao, Q., et al.: Cognitive impairment as a comorbidity of epilepsy in older adults: analysis of global and domain-specific cognition. *Epileptic Disord.* **25**(1), 65–73 (2023)
2. Sayed, N.M., Aldin, M., Ali, S.E., et al.: Cognitive functions and epilepsy-related characteristics in patients with generalized tonic-clonic epilepsy: a cross-sectional study. *Middle East Curr. Psychiatry* **30**(1), 15 (2023)
3. Dryżałowski, P., Józwiak, S., Franckiewicz, M., et al.: Benign epilepsy with centrotemporal spikes—current concepts of diagnosis and treatment. *neurologia i neurochirurgia polska* **52**(6), 677–689 (2018)
4. Hata, M., Satake, Y., Miyazaki, Y., et al.: Hidden cases of epilepsy in cognitive impairment clinics: exploring the use of a portable device for simplified electroencephalography testing. *Epilepsy Behav. Rep.* **27**, 100701 (2024)
5. Vannest, J., Tenney, J.R., Gelineau-Morel, R., et al.: Cognitive and behavioral outcomes in benign childhood epilepsy with centrotemporal spikes. *Epilepsy Behav.* **45**, 85–91 (2015)
6. Dharan, A.L., Bowden, S.C., Peterson, A., et al.: A cross-sectional investigation of cognition and epileptiform discharges in juvenile absence epilepsy. *Epilepsia* **64**(3), 742–753 (2023)
7. Sekimoto, M., Kato, M., Muramatsu, R., et al.: Cognitive dysfunction in drug-naïve late-onset temporal lobe epilepsy. *Epilepsy Behav.* **146**, 109356 (2023)
8. Zhang, X., Zeng, J., Gu, X., et al.: Relapse after drug withdrawal in patients with epilepsy after two years of seizure-free: a cohort study. *Neuropsychiatric Dis. Treat.* 85–95 (2023)
9. Tiwary, H., Bhavsar, A.: Anti-seizure medication classification using EEG signals via attention-based CNN. In: 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS), pp. 605–610. IEEE (2023)
10. Tsai, M.L., Wang, C.C., Yu-Der Wang, A., et al.: Antiseizure medications normalize electroencephalographic functional connectivity and power in children with benign epilepsy with centrotemporal spikes. *Pediatr. Neurol.* **156**, 41–50 (2024)
11. Liguori, C., Toledo, M., Kothare, S.: Effects of anti-seizure medications on sleep architecture and daytime sleepiness in patients with epilepsy: a literature review. *Sleep Med. Rev.* **60**, 101559 (2021)

12. Novak, A., Vizjak, K., Rakusa, M.: Cognitive impairment in people with epilepsy. *J. Clin. Med.* **11**(1), 267 (2022)
13. Cho, D., Yu, M.S., Shin, J., et al.: A computational clinical decision-supporting system to suggest effective anti-epileptic drugs for pediatric epilepsy patients based on deep learning models using patient's medical history. *BMC Med. Inform. Decis. Mak.* **24**(1), 149 (2024)
14. Rahim, F., Azizimalamiri, R., Sayyah, M., et al.: Experimental therapeutic strategies in epilepsies using anti-seizure medications. *J. Exp. Pharmacol.* 265–290 (2021)
15. Nevitt, S.J., Sudell, M., Cividini, S., et al.: Antiepileptic drug monotherapy for epilepsy: a network meta-analysis of individual participant data. *Cochrane Database Syst. Rev.* (4) (2022)
16. Ricci, L., Tombini, M., Savastano, E., et al.: Quantitative EEG analysis of brivaracetam in drug-resistant epilepsy: a pharmaco-EEG study. *Clin. Neurophysiol.* **163**, 152–159 (2024)
17. Operto, F.F., Pastorino, G., Mazza, R., et al.: Cognitive profile in bects treated with levetiracetam: a 2-year follow-up. *Epilepsy Behav.* **97**, 187–191 (2019)
18. Cao, L., Chen, Y., Lv, N., et al.: Clinical study of the effect of 5 kinds of antiepileptic drugs on the postictal state. *Epilepsy Behav.* **158**, 109897 (2024)
19. Li, Y., Sun, Y., Niu, K., et al.: The relationship between neuromagnetic activity and cognitive function in benign childhood epilepsy with centrotemporal spikes. *Epilepsy Behav.* **112**, 107363 (2020)
20. Strzelczyk, A., Schubert-Bast, S.: Psychobehavioural and cognitive adverse events of anti-seizure medications for the treatment of developmental and epileptic encephalopathies. *CNS Drugs* **36**(10), 1079–1111 (2022)
21. Mahajan, R., Morshed, B.I.: Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA. *IEEE J. Biomed. Health Inform.* **19**(1), 158–165 (2014)
22. Chen, W., Wang, Y., Ren, Y., et al.: An automated detection of epileptic seizures EEG using CNN classifier based on feature fusion with high accuracy. *BMC Med. Inform. Decis. Mak.* **23**(1), 96 (2023)
23. Ren, Z., Yue, M., Han, X., et al.: The potential of the lempel-ziv complexity of the EEG in diagnosing cognitive impairment in patients with temporal lobe epilepsy. *Epileptic Disord.* **25**(3), 331–342 (2023)



Light Source Optimization Design of Intelligent Grading Equipment for Flue-Cured Tobacco

Chen Pengbiao, Hou Kaihu^(✉), Gong Peiyao, and Bai Fan

Faculty of Mechanical and Electrical Engineering, Kunming University of Science and Technology, Kunming 650500, China
576521346@qq.com

Abstract. Aiming at the problem of low grading accuracy caused by illumination non-uniformity of light source in an intelligent grading equipment for flue-cured tobacco, an optimization solution to improve the non-uniformity was proposed. Firstly, a 54-point measurement method was proposed to measure the illumination non-uniformity of the light source, and the problem of illumination non-uniformity of the light source of the used equipment was analyzed. Secondly, the optimization solution of LED line light source and LED surface light source was proposed, and the variance, standard deviation and coefficient of variation of light source illumination were calculated, and the illumination uniformity was analyzed. Finally, the LightGBM model was used to verify the rationality of the optimization solution. The results showed that the illumination uniformity and image sample grading accuracy of the optimized light source were better than those of the old light source to meet the accuracy requirements of the actual acquisition of flue-cured tobacco.

Keywords: Light source · Optimization solution · Illumination uniformity · LightGBM

1 Introduction

As an important part of machine vision and image processing system, the intensity, uniformity, color rendering and stability of light source directly affect the quality of images [1–3]. In practical application, the unreasonable layout and structure of the light source and the uniformity distribution of the light source in the lighting equipment of the flue-cured tobacco intelligent grading equipment used by a domestic company lead to the obvious difference between the light and shade of the collected image samples and the shadow part, which greatly increases the difficulty of subsequent image processing and analysis, which leads to the low grading accuracy of the grading equipment.

In recent years, the field of secondary optical design of light sources had attracted the in-depth exploration and remarkable results of a large number of researchers because of its potential in improving the uniform illumination of light sources [4, 5]. Wu Fei [6] optimized the distance of LED light source by using Sparrow's law, determined the optimal distance, and constructed a triangular light source array. Qian [7] optimized

the design of an internal and external refractive curved optical lens with a distance-to-height ratio of 5.5 in the backlight system. Under the condition that the actual luminous area was $1.7 \text{ mm} \times 1.7 \text{ mm}$, the spot uniformity was 0.83. He et al. [8] introduced the fruit fly optimization algorithm to optimize the three-dimensional layout parameters of rectangular and circular LED arrays. Through the simulation verification of TracePro software, the illumination uniformity was 97.4% and 96.0% respectively. Zhang Shuai et al. [9] used a pyramid-shaped optical element to divide the light emitted by the LED into three parts: transmission, reflection and direct light, which improved the uniformity of light mixing and color mixing on the light output surface. Using the Taguchi method, a plant lighting surface light source with an illumination uniformity of 92.13% and a color mixing uniformity of 89.72% was obtained. Wang J [10] et al. established the illumination calculation model of the ring light source. Based on the theoretical model, the mathematical relationship between the distance between the detection plane and the LED plane and the illumination uniformity was obtained, and the optimal solution was obtained to obtain the best uniform illumination field. Wang Jiawen et al. [11] constructed the uniformity evaluation function, and used the simulated annealing algorithm to optimize the circular and 4×4 and 5×5 rectangular LED arrays, and obtained 97%, 96% and 95% illumination uniformity, respectively. Luo Xiaoyan [12] et al. used the bacterial colony algorithm to optimize the three-dimensional position parameters of the two LED arrays. TracePro was used to verify the three-dimensional optimization results of the light source, and the illumination uniformity of the LED array is 88.58% and 95.8%.

In this paper, 54-point measurement method was proposed to measure the illumination uniformity of the light source, and the problem of illumination non-uniformity of the light source of the used equipment was analyzed. Then, two optimization solutions of LED line light source and LED surface light source were proposed, and the variance, standard deviation and coefficient of variation of the illumination of the light source were calculated, and the illumination uniformity was analyzed. Finally, in order to verify the actual effect of the optimization solution and its influence on the grading accuracy, the LightGBM grading model was used to classify the collected images.

2 Used Equipment Light Source Layout Problem

The lighting system of a domestic company's flue-cured tobacco intelligent grading equipment was composed of four lamp tubes and supporting lampshade (as shown in Fig. 1 and Fig. 2). The bottom double lamp tubes directly illuminate the belt with a 30° angle light source, while the upper double lamp tubes illuminate the lampshade at the same angle and then reflect it to the belt, forming a unique lighting layout.

The image collected under the condition of the old light source was shown in Fig. 3. The light and dark distribution was significantly non-uniformity, and the upper right of the tobacco leaf was particularly bright. According to the national standard of flue-cured tobacco GB2635-1992 [13], color accuracy is the key to grading. This illumination non-uniformity may directly interfere with the color variable extraction of the machine vision system, which in turn affects the accuracy and reliability of the intelligent grading. And the tobacco leaf folding or wrinkled by the light source irradiation will leave a significant

shadow part, for the subsequent image processing to increase the difficulty. According to the national standard of flue-cured tobacco GB2635–1992, the length and width of tobacco leaves will affect the determination of position and grade, and the inaccurate shadow of image processing will affect the accuracy and reliability of intelligent grading.

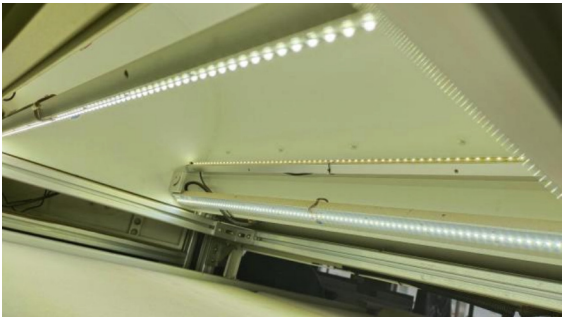


Fig. 1. The used equipment light source physical map.



Fig. 2. The used equipment light source structure diagram.

To further verify the problem of uneven illumination of the equipment, it is necessary to measure the illumination of the light source of the lighting system. The traditional illumination uniformity measurement method is the 13-point measurement method [14], which has the advantage of simple operation and can better reflect the illumination non-uniformity of the whole area. However, it also has some limitations, for example, it may ignore the problem of local illumination non-uniformity. In practical applications, the number and location of measurement points need to be adjusted according to the specific situation.

To accurately quantify the illumination uniformity of the light source, this paper proposed a more refined 54-point measurement method. Compared with the traditional 13-point method, this method adds 54 measurement points in the rectangular area (as shown in Fig. 4), fully covering the area to be measured, effectively reducing the omission of local illumination non-uniformity.



Fig. 3. The image of the used device light source.

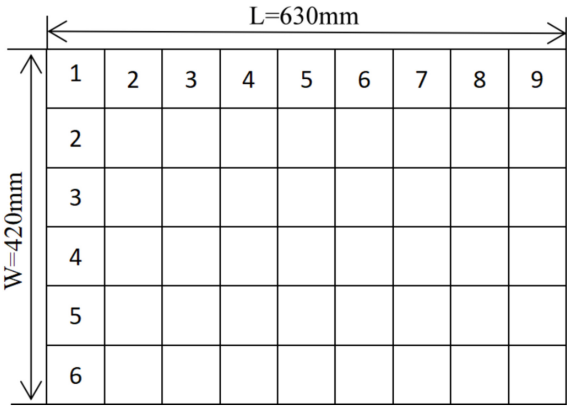


Fig. 4. Light source uniformity 54-point method measurement point diagram.

The light intensity data collected by the light source of the used equipment lighting system was shown in Table 1.

From the light intensity of 54 points in Table 1, the mean, variance, standard deviation and coefficient of variation were calculated:

- (1) The light intensity in the middle area (fourth row) was significantly weaker than that of the edge position;
- (2) The coefficient of variation = 0.21 > 0.1, indicated that there was a large fluctuation and uncertainty in the data, that is, the illumination non-uniformity of the light

Table 1. Measurement data of light intensity of the old light source at 54 points (unit: LUX).

	Illumination Intensity($\times 1000$)										Avg	Var	Sig
Illumination Intensity	35.70	36.10	37.30	40.60	33.60	46.70	45.30	44.40	42.10	40.20	22.34	4.73	
	56.30	62.50	62.60	65.00	66.40	77.00	78.00	76.20	71.00	68.33	58.01	7.62	
	61.40	65.00	68.30	69.00	68.70	62.00	60.50	59.70	56.60	63.47	20.04	4.48	
	45.10	48.90	50.40	51.30	51.40	50.20	48.70	47.30	44.20	48.61	6.79	2.61	
	48.00	51.20	52.80	54.40	54.20	62.10	63.40	60.40	54.90	55.71	26.86	5.18	
	70.80	75.00	74.60	76.10	72.50	76.20	72.70	71.30	66.40	72.84	9.73	3.12	
Avg	52.88	56.45	57.67	59.40	57.80	62.37	61.43	59.88	55.87	58.19			
Var	156.98	190.50	183.09	169.29	209.68	159.83	165.71	158.99	133.41		151.95		
Sig	12.53	13.80	13.53	13.01	14.48	12.64	12.87	12.61	11.55			12.33	
Coefficient of variation = 0.21													

source, and the inconsistent quality of the collected image will affect the reliability and grading effect of the subsequent image acquisition.

In summary, the lighting system of the equipment had the problem of illumination non-uniformity and needed to be optimized.

3 Light Source Optimization

3.1 Optimization Solution One

To solve the problems existing in the equipment, combined with the actual situation, the old LED light source was used in the selection of the light source. Considered the characteristics of the irradiated object (flue-cured tobacco), the motion state on the equipment and the box space of the lighting system, the layout and structure of the light source were changed, and the optimization solution was proposed: the line light source and the surface light source do the forward lighting layout. The physical diagram and structure diagram of the first solution were shown in Fig. 5 and Fig. 6. Four LED lamps were placed inside the light box in parallel. The camera and the light source were located on the same side of the target. The light source was irradiated on the conveyor belt. The reflected light enters the camera directly, and the distance between the light source and the irradiation plane was adjusted to make the illumination of the light source more uniform on the measured object.

The image collected by the optimization solution was shown in Fig. 7. Compared with the used equipment, the illumination effect was more uniform, and the illumination difference of the irradiated object part was almost invisible. The overall outline of the tobacco leaf image was clear, and there was no shadow at the fold, which lays a foundation for subsequent image processing.

The light intensity data collected by the optimized lighting system light source were shown in Table 2. It can be seen that:

The coefficient of variation of the light intensity of the optimization solution was 0.09, which was much lower than the light source illumination of the used equipment,



Fig. 5. Optimization solution one physical map



Fig. 6. Optimization solution one structure diagram.



Fig. 7. The image collected by the optimization solution one.

indicating that the light source stability of the optimization solution was significantly improved and the data fluctuation was greatly reduced.

Table 2. Light intensity 54 points measurement data of optimization solution one (unit: LUX).

	Illumination Intensity($\times 1000$)									Avg	Var	Sig
Illumination Intensity	176.90	184.70	191.90	192.20	188.70	176.80	168.40	156.00	145.70	175.70	266.09	16.31
	188.50	197.50	205.30	208.10	204.50	194.30	184.80	170.50	156.80	190.03	294.81	17.17
	187.70	198.90	208.00	212.30	210.90	202.80	193.40	179.40	164.50	195.32	253.21	15.91
	183.10	195.20	204.60	209.10	208.30	202.20	193.60	180.10	167.60	193.76	202.77	14.24
	175.80	186.50	196.70	201.80	202.10	197.00	189.10	177.60	164.50	187.90	169.82	13.03
	166.50	176.90	185.40	189.90	191.60	187.80	180.80	170.90	158.10	178.66	132.65	11.52
Avg	179.75	189.95	198.65	202.23	201.02	193.48	185.02	172.42	159.53	186.89		
Var	70.05	74.89	78.40	87.17	80.88	97.29	90.91	82.03	63.06		253.85	
Sig	8.37	8.65	8.85	9.34	8.99	9.86	9.53	9.06	7.94			15.93
Coefficient of variation = 0.09												

3.2 Optimization Solution Two

Optimization solution two: The surface light source was shown in Fig. 8, and its structure was shown in Fig. 9. LED lamp beads were densely distributed and projected to the conveyor belt in parallel to achieve more delicate light coverage and reduce spot and shadow.

The image collected by the second optimization solution (surface light source structure) was shown in Fig. 10: It also showed that there was no significant difference in the illumination of tobacco leaves in the picture except for a more uniform illumination effect.

The light intensity data collected by the optimized surface light source were shown in Table 3. It can be seen that:

The coefficient of variation of the light intensity of the optimization solution was reduced to 0.08, which was much lower than that of the old light source, indicating that the stability of the light source was significantly improved and the data fluctuation was greatly reduced.

4 Illumination Uniformity Data Analysis

To facilitate the comparison of the illumination uniformity of the light sources of the two lighting systems, it was necessary to standardize the data, so that the three sets of data were converted into data with a mean value close to 100 for comparison. The formula is as follows:

$$X'_n = X_n \cdot (100/\overline{x_n}) \tag{1}$$



Fig. 8. Optimization solution two physical map.



Fig. 9. Optimization solution two structure diagram.

Among them, X_n' is the transformed data, X_n is the original data, and $\bar{x}_n$ is the mean value of the n th set of data. The following data(Table 4, Table 5, Table 6) were obtained and their mean, variance, standard deviation and coefficient of variation were calculated. Under the same dimension, it can be seen from Table 4, Table 5 and Table 6 that the variances of the three groups were 448.48, 74.02 and 57.95 respectively. The standard deviations were: 21.18, 8.60, 7.61; the coefficient of variation were: 0.21, 0.09, 0.08.



Fig. 10. The image collected by the optimization solution two.

Table 3. Light intensity 54 points measurement data of optimization solution two (unit: LUX).

	Illumination Intensity($\times 1000$)									Avg	Var	Sig
Illumination Intensity	99.10	105.00	111.10	115.00	116.90	117.80	116.40	110.90	103.50	110.63	44.86	6.70
	117.40	124.90	131.40	133.10	126.40	125.60	123.80	122.30	114.60	124.39	35.02	5.92
	126.70	133.80	138.90	139.90	132.50	133.00	130.90	129.60	120.60	131.77	34.88	5.91
	126.90	133.40	138.60	140.10	133.20	133.70	132.10	126.70	121.10	131.76	35.98	6.00
	122.60	128.50	134.40	136.40	130.70	132.10	129.80	123.10	117.70	128.37	37.15	6.09
	114.00	119.40	126.40	128.90	125.80	125.30	121.70	116.30	109.60	120.82	42.11	6.49
Avg	117.78	124.17	130.13	132.23	127.58	127.92	125.78	121.48	114.52	124.62		
Var	109.87	117.49	108.91	89.37	36.83	38.27	38.17	47.05	47.21		90.10	
Sig	10.48	10.84	10.44	9.45	6.07	6.19	6.18	6.86	6.87			9.49
Coefficient of variation = 0.08												

By further comparing the variance, standard deviation and coefficient of variation of the two sets of data, it is significantly revealed that the dispersion of the old light source in the light intensity was the most significant, that was, the data fluctuates the most. Specifically, the illumination uniformity of the light source of the used equipment was very poor, and the light intensity had a large fluctuation; in contrast, the optimized light source showed better illumination uniformity, low data dispersion and high stability. Among them, the surface light source of the second optimization solution performed best.

Among them, Figs.11(a), (b) and (c) are the standardized data linear regression fitting plots of the original light source, optimization solution one and optimization solution two, respectively.

According to the linear regression of Fig. 11, the fitting equation of the old light source was: $y = 0.22x + 93.94$, the fitting equation of the first optimization solution was: $y = -0.22x + 105.45$, and the fitting equation of the second optimization solution

Table 4. Standardization data of the old light source.

	Illumination Intensity									Avg	Var	Sig
Illumination Intensity	61.33	62.02	64.08	69.75	57.72	80.23	77.83	76.28	72.33	69.06	65.93	8.12
	96.72	107.38	107.55	111.67	114.08	132.29	134.00	130.91	121.98	117.40	171.23	13.09
	105.49	111.67	117.34	118.54	118.03	106.52	103.94	102.56	97.24	109.04	59.13	7.69
	77.48	84.01	86.59	88.13	88.31	86.24	83.67	81.26	75.94	83.51	20.04	4.48
	82.46	87.96	90.71	93.46	93.12	106.69	108.92	103.77	94.32	95.71	79.27	8.90
	121.63	128.85	128.16	130.74	124.56	130.91	124.90	122.49	114.08	125.15	28.71	5.36
Avg	90.85	96.98	99.07	102.05	99.30	107.15	105.54	102.88	95.98	99.98		
Var	463.34	562.26	540.38	499.67	618.89	471.76	489.10	469.26	393.78		448.48	
Sig	21.53	23.71	23.25	22.35	24.88	21.72	22.12	21.66	19.84			21.18
Coefficient of variation = 0.21												

Table 5. Standardized data of optimization solution one.

	Illumination Intensity									Avg	Var	Sig
Illumination Intensity	95.53	99.74	103.63	103.79	101.90	95.47	90.94	84.24	78.68	94.88	77.59	8.81
	101.79	106.65	110.86	112.37	110.43	104.92	99.79	92.07	84.67	102.62	85.97	9.27
	101.36	107.41	112.32	114.64	113.89	109.51	104.44	96.88	88.83	105.47	73.84	8.59
	98.87	105.41	110.48	112.91	112.48	109.19	104.54	97.25	90.50	104.63	59.13	7.69
	94.93	100.71	106.22	108.97	109.13	106.38	102.11	95.90	88.83	101.47	49.52	7.04
	89.91	95.53	100.12	102.55	103.46	101.41	97.63	92.29	85.37	96.47	38.68	6.22
Avg	97.07	102.57	107.27	109.21	108.55	104.48	99.91	93.11	86.15	100.92		
Var	20.43	21.84	22.86	25.42	23.59	28.37	26.51	23.92	18.39		74.02	
Sig	4.52	4.67	4.78	5.04	4.86	5.33	5.15	4.89	4.29			8.60
Coefficient of variation = 0.09												

Table 6. Standardized data of optimization solution two.

	Illumination Intensity									Avg	Var	Sig
Illumination Intensity	79.48	84.21	89.10	92.23	93.75	94.48	93.35	88.94	83.01	88.73	28.85	5.37
	94.15	100.17	105.38	106.75	101.37	100.73	99.29	98.08	91.91	99.76	22.53	4.75
	101.61	107.31	111.40	112.20	106.27	106.67	104.98	103.94	96.72	105.68	22.43	4.74
	101.77	106.99	111.16	112.36	106.83	107.23	105.94	101.61	97.12	105.67	23.14	4.81
	98.33	103.06	107.79	109.39	104.82	105.94	104.10	98.73	94.40	102.95	23.89	4.89
	91.43	95.76	101.37	103.38	100.89	100.49	97.60	93.27	87.90	96.90	27.09	5.20
Avg	94.46	99.58	104.37	106.05	102.32	102.59	100.88	97.43	91.84	99.95		
Var	70.67	75.57	70.05	57.49	23.69	24.62	24.55	30.26	30.36		57.95	
Sig	8.41	8.69	8.37	7.58	4.87	4.96	4.96	5.50	5.51			7.61
Coefficient of variation = 0.08												

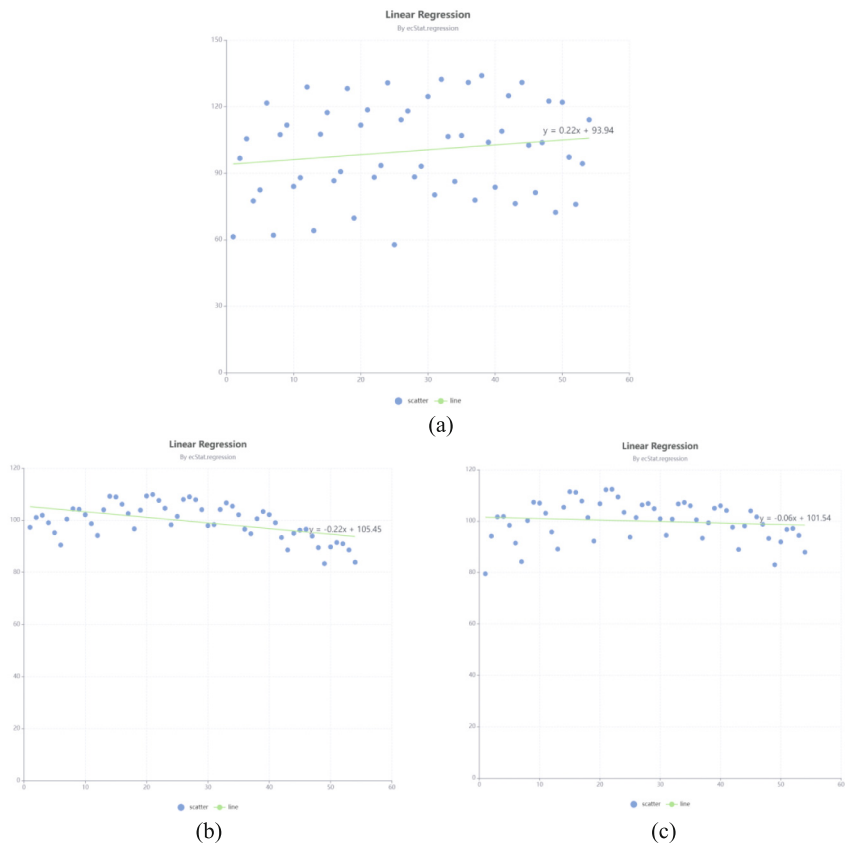


Fig. 11. Three groups of illuminance data linear regression fitting diagram.

was: $y = -0.06x + 101.54$. It can be seen that the fitting effect of the second optimization solution was the best. The mean absolute values of the residuals of the old light source and the optimization solution one data fitting were 17.57 and 4.69, respectively. The illumination data dispersion of the optimization solution one was better than that of the old light source.

In summary, for the problem of illumination non-uniformity of the light source of the used equipment, the two optimization solutions proposed had performed well, significantly improving the illumination uniformity of the light source, basically eliminating the shadow caused by the illumination of the light source, and laying a solid foundation for subsequent image acquisition and intelligent grading. Among them, the performance of the two-sided light source of the optimization solution was better.

5 LightGBM Model Grading Verification

To verify whether the image quality collected by the optimized light source is better, this study used OpenCV to extract 27 features including color, shape, etc.as shown in Table 7 from the tobacco leaf images collected by the three sets of light sources according to the national standard of GB2635–1992 flue-cured tobacco. To distinguish the B-channel in RGB and LAB images, the B-channel features in LAB images were distinguished by *.

Table 7. Extract image feature types and names.

Type	Names
Shape	Width and Height
Color	B_mean, G_mean, R_mean, B_var, G_var, R_var
	H_mean, S_mean, V_mean, H_var, S_var, V_var
	L_mean, A_mean, B_mean*, L_var, A_var, B_var*
Texture	Mean, asm, con, eng, idm, Auto_correlation, std

The train_test_split function in the scikit-learn library was used to divide the same batch of 310 flue-cured tobacco image samples collected by the three groups into training set and test set according to 8:2, and 248 training sets and 62 test sets were obtained. The features extracted from the samples collected by the three light source design solutions were input into the LightGBM grading model, and the grading accuracy and various evaluation parameters of each group of data were recorded. The experimental results were shown in Table 8, and the confusion matrix was shown in Fig. 12, Fig. 13 and Fig. 14.

From Table 8 we can see that:

- (1) The accuracy of the image samples collected by the old light source in the grading model was 0.65, and the accuracy of the optimized light source was 0.85 and 0.86, respectively, which are higher than the grading accuracy of the samples collected by the old light source, indicating that the image collected by the optimized light source can improve the accuracy of image grading and the quality of the image is better.
- (2) The Recall and F1-score of the two optimized light sources were greater than the values of the old light source, which further illustrates that the image samples collected by the optimized light source had better accuracy and integrity when the model recognizes the positive samples.

From Fig. 12, there were 17 samples with correct grading in C2F samples, 4 samples with C3F and 1 sample with C4F; among the C3F samples, there were 20 samples with correct grading, 3 samples with C2F, and 3 samples with C4F. There were 4 samples with correct grading in C4F samples, 0 samples with C2F, and 10 samples with C3F.

From Fig. 13, it can be seen that there were 20 samples with correct grading in C2F samples, 2 samples with C3F and 0 sample with C4F; among the C3F samples, there

Table 8. Evaluation parameters of LightGBM grading model discriminant samples.

Type	Precision	Recall	F1-score
The old light source	0.65	0.61	0.61
Solution 1	0.85	0.85	0.85
Solution 2	0.86	0.88	0.87

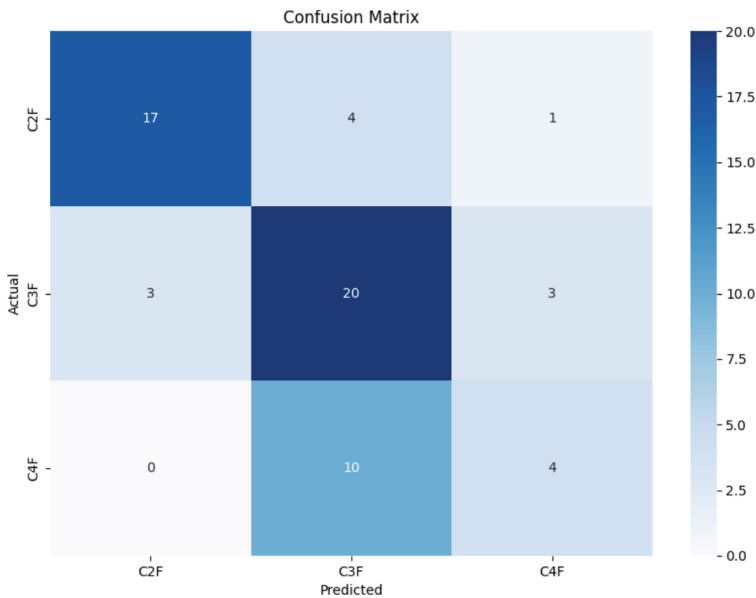


Fig. 12. The sample grading confusion matrix of the old light source.

were 22 samples with correct grading, 1 samples with C2F, and 3 samples with C4F. There were 11 samples with correct grading in C4F samples, 1 samples with C2F, and 2 samples with C3F.

From Fig. 14, it can be seen that there were 19 samples with correct grading in C2F samples, 2 samples with C3F and 1 sample with C4F; among the C3F samples, there were 22 samples with correct grading, 1 samples with C2F, and 3 samples with C4F. There were 13 samples with correct grading in C4F samples, 1 samples with C2F, and 0 samples with C3F.

Through three sets of confusion matrices, it can also be observed intuitively that the grading accuracy of the images collected by the optimized light source had been significantly improved. Two sets of optimized light sources effectively improved the quality of the image, making the features in the image more distinct and easy to identify, thereby improving the accuracy of image grading.

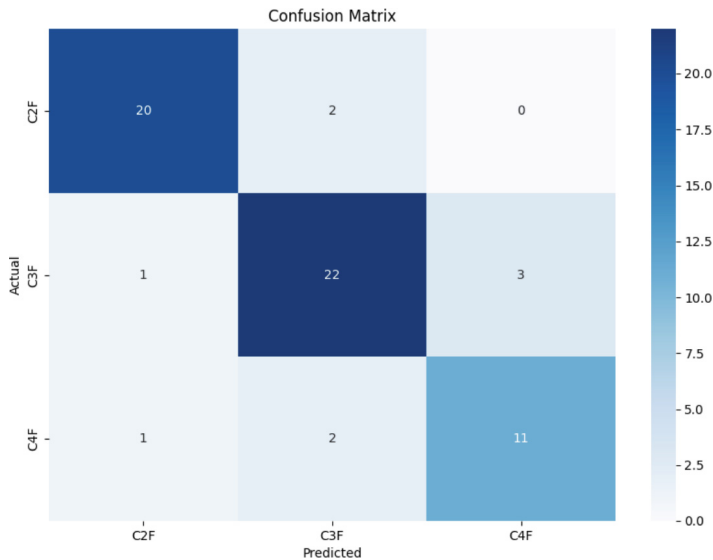


Fig. 13. The sample grading confusion matrix collected in solution one.

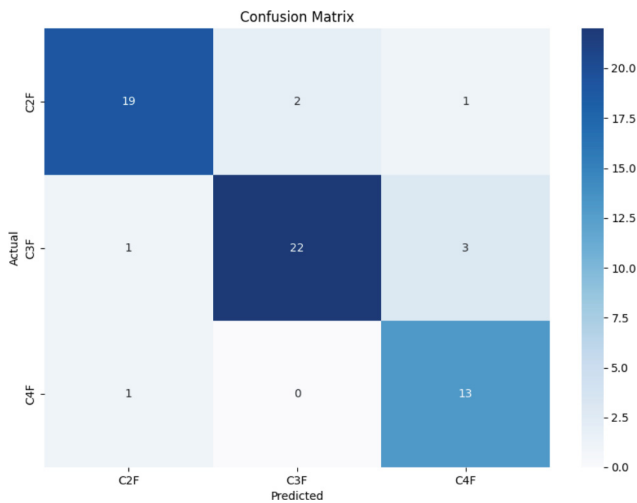


Fig. 14. The sample grading confusion matrix collected in solution two.

6 Discussions

In this paper, aiming at the problem of low grading accuracy caused by illumination non-uniformity of light source in a flue-cured tobacco intelligent grading equipment, two optimization solutions of LED line light source and LED surface light source were proposed. The illumination data of the three light sources are measured by 54-point measurement method, and the variance, standard deviation and coefficient of variation

of the illumination of the light source were calculated, and the illumination uniformity was analyzed. Finally, the LightGBM grading model was used to classify the images collected by the three groups of light sources to verify the rationality of the optimized light source. The experimental results showed that the variation coefficients of the illumination data of the light source of the first and second optimization solutions were 0.09 and 0.08, respectively, which are significantly improved compared with the 0.21 of the old light source. The grading accuracy of the optimized light source through the LightGBM model were 85% and 86%, respectively, which is 20% and 21% higher than that before optimization. As a result, the illumination uniformity of the optimized light source had been significantly improved, which greatly improved the image quality collected by the equipment, effectively eliminates the influence of non-uniformity brightness and shadow on subsequent image processing, and improves the problem of low grading accuracy of the used equipment due to illumination non-uniformity of the light source.

References

1. Xiang'an, Z., Mingyang, W., Yueming, W., et al. Effect of light source non-uniformity for imaging ellipsometry measurements. *Optics Communications* 545 (2023)
2. Su, Y.: Research on Fabric Color Difference Detection Method Based on Machine Vision. Xi'an University of Technology (2021)
3. Yang, R.: Experimental Study on Machine Vision Lighting Technology and Device. Zhongyuan University of Technology (2017)
4. Sugimoto, R., Maruyama, R., Tamada, Y., et al.: Contrast enhancement by oblique illumination microscopy with an LED array. *Optik* 183 (2019)
5. Huang, D., Feng, Q., Tang, T., et al.: Surface energy mapping optimization design method based on extended light source. *J. Optics* **39**(12), 298–305 (2019)
6. Fei, W., Guangda, Y., Fenglan, W., et al.: Light source design of finger vein recognition sensor. *Sensor and Microsystem* **36**(04), 74–77 (2017)
7. Keyuan, Q.: Design optimization of LED near-field optical model and direct-type backlight lens. *J. Opt.* **35**(05), 306–313 (2015)
8. He, G., Wen, S., Fu, M., et al.: Optimization design of three-dimensional LED light source array based on fruit fly optimization algorithm. *Lighting Journal* **39**(08), 1128–1135 (2018)
9. Zhang, S., Wen, S., Ma, B., et al.: Design of high non-uniformity LED surface light source for plant lighting. *Luminous J.* **39**(03), 403–413 (2018)
10. Wang, J., Wang, G., Li, M., et al.: Hand vein recognition based on PCET. *Optik* **127**(19) (2016)
11. Wang Jiawen, S., Zeping, Y.Z., et al.: The problem of illumination non-uniformity in LED array modularization. *Acta Photonica Sinica* **43**(08), 22–28 (2014)
12. Luo, X., Xie, Y.: Three-dimensional optimization design of imaging for vein recognition system. *Sensors and Microsystems* **41**(10), 92–95+99 (2022)

13. GB 2635–1992, Flue-cured tobacco
14. Yu, W., Wang, J., Wang, X., et al.: Laser processing hot-pressed light guide plate dot construction and optimization design. *Optical Technology* **46**(05), 582–586 (2020)



Multidimensional Attention Convolution Plugin for Ship Radiated Noise Classification

Zhongxiang Zheng^{1,2} and Peng Liu^{1,2} 

¹ Hainan Acoustics Laboratory Institute of Acoustics Chinese Academy of Sciences,
Haikou, China
liup@dsp.ac.cn

² Lingshui Marine Information Hainan Observation and Research Station,
LingShui, China

Abstract. In the passive recognition task of ship radiated noise, the efficiency of extracting relevant classification information from the noise directly affects the classification results. This study proposed a multidimensional attentional convolution (MAConv) module as a fundamental component for capturing decisive classification information. Drawing inspiration from dynamic convolutional methods such as CondConv, we perform multidimensional attentional integration in three dimensions: input channel, output channel, and spatial channel, which involves a linear combination of different static convolutional kernels. The weights of the combination are data-dependent, with different samples using different weights, so the network will focus on different input channels, output channels and spatial channels depending on the samples in classification. As a drop-in replacement of traditional convolutions, MAConv can be easily integrated into popular CNN architectures, and subsequent experimental results on the open-source dataset Deepship show that CNN networks using the MAConv plugin can significantly improve the performance of ship radiation noise classification.

Keywords: Deep Learning · Underwater acoustic target recognition · dynamic convolutional

1 Introduction

With the continuous advancement of modern military technology, research and application of ship target classification technology have become an indispensable and essential part in military equipment automation control, intelligence

This work was supported in part by Innovation Promotion Association, Chinese Academy of Sciences under Grant 2022022, in part by South China Sea Nova project of Hainan Province under Grant NHXXRCXM202340, and in part by the Scientific Research Foundation Project of Hainan Acoustics Laboratory under grant ZKNZ2024002.

reconnaissance and surveillance, missile attacks, and other fields. However, due to the diverse types and large quantities of ship targets, coupled with complex factors such as ocean environmental noise interference, accurately and rapidly identifying and classifying ship targets has always been a challenge troubling military technology personnel. Therefore, research on ship target classification technology based on deep learning holds significant practical significance and application value.

Deep learning methods can automatically learn target features, reduce human interference, and effectively address the problems faced by traditional target classification methods. Deep learning-based end-to-end recognition methods are now replacing traditional recognition methods as the focal point in the field of underwater target recognition and classification.

In the domain of underwater target recognition, convolutional neural networks, deep belief networks, and long short-term memory networks are primarily used for target identification. Reference [1] extracted the spectrum and Mel-frequency cepstral coefficients of ship radiation noise, used these extracted feature images as training sets, and employed convolutional neural networks for identification, effectively improving the accuracy of recognition. Reference [2] designed a deep convolutional neural network for feature extraction based on signal time-frequency characteristics, and utilized deep convolutional networks for target identification, significantly enhancing recognition accuracy. Reference [3] employed convolutional neural networks for signal feature recognition, followed by deep neural networks (DNN) for target identification, resulting in a significant improvement in classification effectiveness. Reference [4] directly utilized convolutional neural networks for target identification, using line spectra and modulation features extracted from signals as training sets, and employing convolutional neural networks for target recognition. References [5–9] utilized convolutional neural networks with features extracted from DEMON spectra, LOFAR spectrograms, and constant Q-transform techniques as training sets for ship classification. Spectrograms are more complex compared to images; if spectrograms are solely used as features for convolutional neural network classification, classification issues may arise. Reference [10] established an identification and classification model based on long short-term memory networks by extracting Mel-frequency cepstral coefficients and leveraging long short-term memory networks.

2 Related Work

With the continuous evolution of marine information technology, ship target identification has emerged as a pivotal research domain. Traditional classification methods face significant challenges in parameter selection when encountering a high number of “anomalies” within data, rendering them ineffective in addressing target classification issues under the complexities of the marine environment. Deep learning methodologies enable batch learning of data, continual adjustment of parameters based on batch data samples, insensitivity to outliers,

and higher recognition rates for intricate classification dilemmas. Kamal S et al. [11] introduced an underwater target identification approach based on deep belief networks, which extracts invariant features from signals through a multi-layered structure and demonstrates promising results through experimental validation.

Mcquay C et al. [12] employed a convolutional neural network combined with a support vector machine classification model to classify and identify the sounds of various marine mammals, achieving higher recognition rates compared to traditional Mel-frequency cepstral coefficients. Xu C et al. [13] utilized stacked autoencoders (SAE) to extract stable features from underwater target noise and proposed an underwater target classification method combining SAE with soft-max. However, the aforementioned classification methods rely on extensive training data samples, posing limitations for small-sample target classification issues in underwater acoustics.

Researchers have also delved into exploring solutions for small-sample recognition problems. Jin G et al. [14] utilized generative adversarial networks to extend samples of LOFAR spectrograms of raw signals and employed convolutional neural networks for classification and recognition, achieving superior recognition rates compared to methods without sample extension. Yang H et al. [15] introduced a competitive deep belief network that can process both labelled and unlabelled samples to extract valid classification information, reaching a classification accuracy of 90.89%.

It is evident that directly applying existing network structures from the computer field to underwater target classification issues proves challenging. Continuous exploration is needed to address classification problems with insufficient sample data in underwater acoustics.

3 Method

In this section, we begin by revisiting dynamic convolution through a general formula. Subsequently, we elucidate the formula of our Multidimensional Attention Convolution (MAConv), clarifying its characteristics and detailing its implementation.

A standard convolutional layer employs static convolutional kernels that are applied universally across all input samples. Networks using static convolution layers can only enhance their generalization by increasing depth and channel numbers, leading to significantly higher computational requirements and increased deployment complexity. In contrast, dynamic convolution utilizes distinct convolutional kernels for different inputs, parameterizing certain kernel aspects as linear combinations of expert knowledge [16]. This approach allows for an increase in model capacity by expanding the number of experts, which proves more efficient than enlarging kernel sizes. Moreover, expert knowledge requires only one combination, facilitating efficient reasoning while enhancing model capacity.

In general, dynamic convolution can be represented by the following formula:

$$Output(x) = \sigma((\alpha_1 W_1 + \alpha_2 W_2 + \cdots \alpha_n W_n) * x) \quad (1)$$

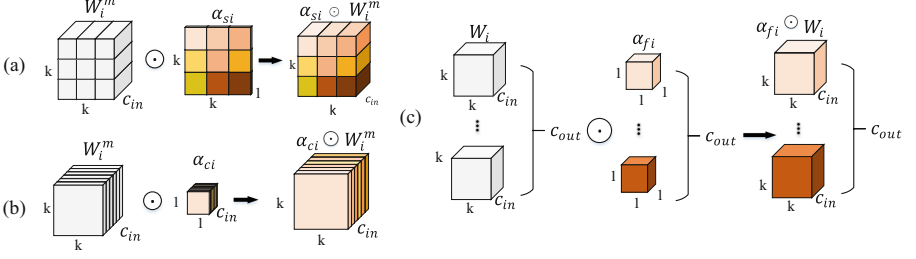


Fig. 1. Attention for convolution along three different dimensions in MAConv. (a) position multiplication operation along the spatial dimension, (b) channel multiplication operation along the input channel dimension, and (c) filter multiplication operation along the output channel dimension. The terminology will be explained in the Methods section.

where $\alpha_i = r_i(x)$ denotes sample-dependent weighted parameters, and W_n represents convolutional kernels with different parameters. Traditional convolution relies on enlarging kernel sizes and channel numbers to boost capacity, which escalates overall network computations. Dynamic convolution, on the other hand, involves computing weighted convolutional kernels for input samples through multiple experts before convolution operations. Crucially, each kernel is calculated once and applied to various locations. This implies that augmenting expert data volume enhances network capacity, resulting in minimal inference computational load: each additional parameter necessitates only a single multiplication and addition operation [17].

Dynamic convolution essentially forms a linear combination of multiple static convolutions, providing an equivalent capacity to n experts but with superior computational efficiency. The key aspect lies in the weighted parameters, which must exhibit data dependency; Otherwise, attention convolution will be equivalent to static convolution. How to design these data-dependent parameters is addressed through the following approach:

$$r(x) = (\text{Sigmoid}(\text{GAP}(x)R)) \quad (2)$$

Therefore, the derivation of the convolutional kernel parameters for a CondConv [18] layer is based on the aforementioned linear combination formula. The entire process can be summarized as follows: dependent on the input x , prior to the convolution operation, each expert's coefficient α_i is computed through a function $\text{routing}(r_i(x))$, and then, through linear combination, the final kernel of the CondConv layer is obtained. This kernel is subsequently convolved with the input x and activated.

For the input x , the initial step consists of global average pooling, followed by multiplication with the matrix R (right-multiplying the altered matrix maps the input dimensions to n experts, allowing for linear combinations afterwards). Finally, the weight coefficients on each dimension are mapped to the interval $[0,1]$

by sigmoid activation. Thus, different weight vectors are available depending on the input x , thus providing different kernels for the CondConv layer.

As shown in the Fig. 1, the MAConv convolution proposed in this paper features adaptable parameters across input channels, output channels, and spatial channels. Following the notations in Eq. (1), MAConv can be defined by:

$$\text{Output}(x) = \sigma((\alpha_{s1} \odot \alpha_{c1} \odot \alpha_{f1} \odot W_1 + \cdots \alpha_{sn} \odot \alpha_{cn} \odot \alpha_{fn} \odot W_n) * x) \quad (3)$$

In the provided text, the symbols $\alpha_{si} \in \mathbb{R}^{k \times k}$, $\alpha_{ci} \in \mathbb{R}^{c_{in}}$, and $\alpha_{fi} \in \mathbb{R}^{c_{out}}$ represent three newly introduced attention mechanisms. These attentions are computed across the spatial dimension, input channel dimension, and output channel dimension of the convolutional kernel space for W_i . The symbol $\odot$ denotes the multiplication operations across various dimensions of the kernel space. Specifically, α_{si} , α_{ci} , and α_{fi} are determined using a multi-head attention component $\pi_i(x)$.

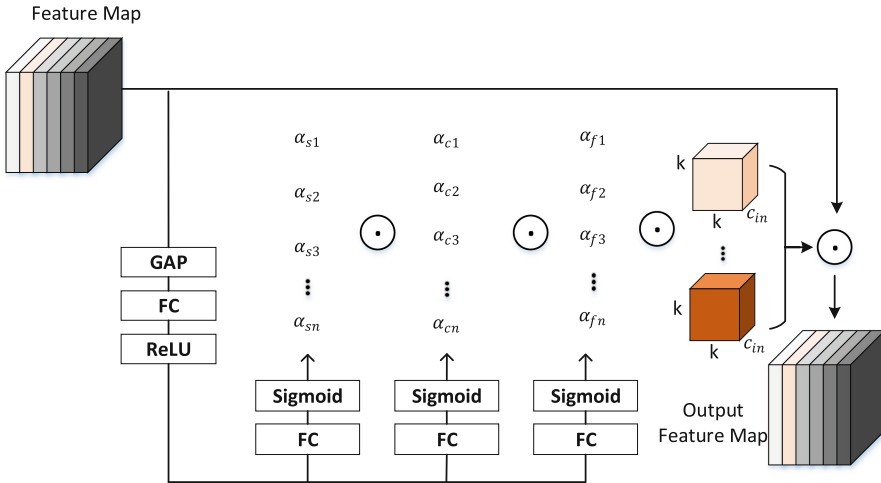


Fig. 2. Unlike CondConv, which computes a single attention scalar αw_i for a convolutional kernel W_i , MAConv utilises a novel multi-dimensional attention mechanism to compute three types of attention α_{si} , α_{ci} , and α_{fi} along all three dimensions of the kernel space in a parallel manner.

The passage outlines the procedure of multiplying the three types of attention by n convolutional kernels, depicted in Fig. 2. These attention types, addressing position, channel, and filter dimensions, are harmoniously integrated into the convolutional kernel w_i . This approach ensures the uniqueness of each convolution operation across all spatial positions, input channels, and filters of input x , guaranteeing effective capture of rich contextual cues. As a result, MAconv notably amplifies the feature extraction potential of fundamental CNN convolution operations.

The method for computing the weights $\pi_i(x)$ is similar to the calculation approach of an SE-type attention mechanism. An SE-typed attention module [19] with multiple heads, denoted as $\pi_i(x)$, is utilized for this computation, the structure of which is depicted in Fig. 2. Initially, the input x undergoes channel-wise Global Average Pooling (GAP) operations to compress it into a feature vector with a length of c_{in} . This feature vector then passes through a Fully Connected (FC) layer and three head branches. Following the FC layer, a Rectified Linear Unit (ReLU) activation function is applied. The FC layer reduces the dimensionality of the feature vector to a lower-dimensional space based on a reduction ratio r (set as $1/16$ in main experiments to avoid excessive model complexity).

For each of the three head branches, there is an FC layer with output sizes of $k \times k$, $c_{in} \times 1$, and $c_{out} \times 1$, along with a Softmax or Sigmoid function to generate normalized attentions αs_i , αc_i , and αf_i , respectively. Additionally, αs_i , αc_i , and αf_i are shared across all convolutional kernels.

In the experimental part, we will verify the effectiveness of MAConv as a plug-in in ship radiated noise identification.

4 Experiments

This section provides a detailed account of the dataset, implementation process, and evaluation criteria. Subsequently, thorough experiments on the Deepship dataset are conducted to assess the efficacy of the MAConv module as an integrated component within three distinct CNN architectures for neural networks (ResNet50 [20], DenseNet121 [21], MobileNetV2 [22]).

The study utilised the publicly accessible DeepShip dataset, which records the marine environment in real-world conditions during different seasons and sea states. In addition to vessel signals, the recorded signals include natural background noise, marine mammal noise and any other noise from human activities. The dataset consists of 609 ship radiated noise (SRN) recordings, categorised as tankers (240), tugs (69), passenger ships (191) and cargo ships (109), with audio sampled at 32 kHz.

We integrated 7 h and 52 min of background noise from the marine environment into the dataset and divided it into 5 different groups, as detailed in the article [23]. Initially, the dataset was partitioned into a training set and a test set, ensuring the absence of shared segments of identical radiated noise in both sets, thereby enhancing the authenticity of the recognition process.

Each recording was then split into 3-second samples with a frame shift of 1 s. In total, 137,369 training samples and 58,567 test samples were generated from the frame splitting. The training samples included 117,734 ship radiated noise and 19,635 ocean background noise, while the test samples included 50,434 ship radiated noise and 8,133 ocean background noise. For a more elaborate breakdown, please refer to the Data Set Segmentation available at <https://github.com/ZhuPengsen>.

Each sample has 96000 signal points, after windowing and pre-emphasis operations are performed on the signal, four features are extracted for each sample,

which are CQT transform, mel spectrograms from 1–100 HZ, mel spectrograms from 100–1000 HZ, and mel spectrograms from 1–16 KHZ. They are saved as images of total size 224×224 . Finally the mean and standard deviation of all sample images were normalised.

To assess the classification performance of the various methods, the experiments employed metrics such as recognition accuracy, macro precision, macro recall, and macro F1 score. These metrics were utilized consistently throughout the experiments for evaluation purposes.

The experimental setup involved a server configuration featuring an NVIDIA RTX3090 graphics card, which offers 24 GB of on-chip memory and up to 35.7 TFLOPS of single-precision floating-point computing power. This robust hardware configuration enabled efficient processing and computation for the experiments conducted.

The construction of the network structure was carried out using the PyTorch framework with a batch size of 32 for training and a total of 100 epochs, with model validation using the validation set at the end of each epoch. The highest accuracy epoch is selected as the final classification network, for which recall, precision, accuracy and F1 score are calculated.

The precision, recall, and F1 score are key metrics used to evaluate the performance of classification models. Here's a breakdown of each metric:

Precision: Precision measures the proportion of correctly predicted positive cases out of all predicted positive cases.

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

TP: True Positives (correctly predicted positive cases) FP: False Positives (negative cases incorrectly predicted as positive)

Recall: Recall, also known as sensitivity, calculates the proportion of correctly predicted positive cases out of all actual positive cases.

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

FN: False Negatives (positive cases incorrectly predicted as negative)

F1 Score: The F1 score is the harmonic mean of precision and recall, providing a balance between the two metrics.

$$F1 \text{ score} = \frac{(Precision \times Recall)}{(Precision + Recall)} \quad (6)$$

These metrics are crucial for understanding the model's predictive performance, accuracy, and ability to correctly identify positive instances. The F1

score is particularly useful in situations where you want to balance precision and recall [24].

By considering precision, recall, and the F1 score, you can gain a comprehensive understanding of how well your model is performing in classification tasks.

In a multi-category classification task, the three metrics of precision, recall, and F1 score will be replaced by macro-precision, macro-recall, and macro-F1 score, respectively. They are calculated by computing precision, recall, and F1 score separately for each individual category and then averaging across all categories. These metrics provide a comprehensive evaluation of the overall performance of a multiclassification model.

MAConv convolution is used to replace several traditional 3×3 convolution kernels. Since the convolutional sense field at the front end of the network is small and can only capture local information, replacing it with MAConv convolution does not improve the network much, but slows down the network operation, so we only replace several convolutions at the last part of the network, the specific replacement is as follows:

ResNet50: Used MAConv to replace the 3×3 convolutional kernels in all blocks of the last layer.

DenseNet121: Replaced the last 5 3×3 convolutional kernels with MAConv.

MobileNetV2: Substituted the last 5 3×3 convolutional kernels with MAConv.

By applying these modifications and evaluating the macro-precision, macro-recall, and macro-F1 score, we can assess how well the model performs across all classes in a multiclass classification scenario. This approach helps in understanding the overall effectiveness of the model modifications on different CNN architectures.

Table 1. Ablation Experiments for SFE and MCA Modules.

Method	Conv	MAConv	Accuracy	Precision	Recall	F1-score
ResNet50	✓		76.71%	78.14%	78.17%	0.7794
MAResNet		✓	76.97%	78.25%	78.31%	0.7811
DenseNet	✓		75.14%	76.84%	76.42%	0.7627
MADenseNet		✓	75.76%	77.80%	77.09%	0.7684
MobileNetV2	✓		74.65%	76.94%	76.10%	0.7569
MAMobileNetV2		✓	75.58%	77.49%	76.10%	0.7569

The experimental results are shown in Table 1, where all three networks show different degrees of improvement after replacing several conventional convolutional layers using the MAConv proposed in this paper. In terms of accuracy, the MResNet, MADenseNet, and MAMobileNetV2 networks using MAConv are improved by 0.26%, 0.62%, and 0.93%, respectively. It can be seen that the effect of using MAConv instead of the traditional static convolutional layer is minimal in the original ResNet50 network with higher accuracy and larger number of parameters. Presumably this is because the structure that has been residually connected in the ResNet50 network is already effective in capturing valid information in the classification, and continuing to increase the learnable parameters of the convolutional kernel has less of an impact on the results.

5 Conclusion

In conclusion, the utilization of the multidimensional attentional convolution (MAConv) module in the context of passive ship radiated noise recognition has demonstrated significant promise in enhancing classification outcomes. By leveraging multidimensional dynamic convolutional techniques, MAConv offers a robust framework for capturing crucial classification insights. The data-dependent nature of the weight combinations ensures adaptive attention allocation tailored to individual samples, thereby enhancing the network's ability to discern pertinent details across input, output, and spatial channels during classification tasks.

Experimental results show that the ResNet, DenseNet, and MobileNetV2 networks obtain 0.26%, 0.62%, and 0.93% improvement after replacing several conventional convolutional layers with MAConv, respectively. Future research endeavors could delve deeper into optimizing MAConv integration strategies and exploring its applicability across diverse classification domains.

References

1. Yang, H., Yang, X., Li, G.: Dual feature extraction system for ship-radiated noise and its application extension. *Ocean Eng.* **285**, 115352 (2023)
2. Lian, Z., Wu, T.: Feature extraction of underwater acoustic target signals using gammatone filterbank and subband instantaneous frequency. In: 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 944–949. IEEE (2022)
3. Xue, L., Zeng, X., Jin, A.: A novel deep-learning method with channel attention mechanism for underwater target recognition. *Sensors* **22**(15), 5492 (2022)
4. Szymak, P.: Recognition of underwater objects using deep learning in matlab. In: 2018 International Conference on Applied Mathematics and Computational Science (ICAMCS.NET), pp. 53–535. IEEE (2018)
5. Lu, J., Song, S., Hu, Z., Li, S.: Fundamental frequency detection of underwater acoustic target using demon spectrum and CNN network. In: 2020 3rd International Conference on Unmanned Systems (ICUS), pp. 778–784. IEEE (2020)

6. Cao, X., Togneri, R., Zhang, X., Yang, Yu.: Convolutional neural network with second-order pooling for underwater target classification. *IEEE Sens. J.* **19**(8), 3058–3066 (2019)
7. Wu, H., Song, Q., Jin, G.: Deep learning based framework for underwater acoustic signal recognition and classification. In: *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence, CSAI' 18*, pp. 385–388. ACM (2018)
8. Yue, H., Zhang, L., Wang, D., Wang, Y., Lu, Z.: The classification of underwater acoustic targets based on deep learning methods. In: *Proceedings of the 2017 2nd International Conference on Control, Automation and Artificial Intelligence (CAAI 2017)*, CAAI-17. Atlantis Press (2017)
9. Cinelli, L., Chaves, G., Lima, M.: Vessel classification through convolutional neural networks using passive sonar spectrogram images. In: *Anais de XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, SBrT2018*. Sociedade Brasileira de Telecomunicações (2018)
10. Zhang, W., Wu, Y., Wang, D., Wang, Y., Wang, Y., Zhang, L.: Underwater target feature extraction and classification based on gammatone filter and machine learning. In: *2018 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, pp. 42–47. IEEE (2018)
11. Kamal, S., Mohammed, S.K., Saseendran Pillai, P.R., Supriya, M.H.: Deep learning architectures for underwater target recognition. In: *2013 Ocean Electronics (SYMPOL)*, pp. 48–54. IEEE (2013)
12. McQuay, C., Sattar, F., Driessen, P.F.: Deep learning for hydrophone big data. In: *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, pp. 1–6. IEEE (2017)
13. Cao, X., Zhang, X., Yu, Y., Niu, L.: Deep learning-based recognition of underwater target. In: *2016 IEEE International Conference on Digital Signal Processing (DSP)*, pp. 89–93. IEEE (2016)
14. Jin, G., Liu, F., Wu, H., Song, Q.: Deep learning-based framework for expansion, recognition and classification of underwater acoustic signal. *J. Exp. Theor. Artif. Intell.* **32**(2), 205–218 (2019)
15. Yang, H., Shen, S., Yao, X., Sheng, M., Wang, C.: Competitive deep-belief networks for underwater acoustic target recognition. *Sensors* **18**(4), 952 (2018)
16. Zhao, Z., Dong, M.: Channel-spatial dynamic convolution: an exquisite omnidimensional dynamic convolution. In: *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)*, pp. 1707–1711. IEEE (2023)
17. Nguyen, H.D., Chamroukhi, F.: Practical and theoretical aspects of mixture-of-experts modeling: an overview. *Wiley Interdisc. Rev. Data Min. Knowl. Discov.* **8**(4), e1246 (2018)
18. Sharmila, V., Geetha, S.: Gastro intestinal disease classification using hierarchical spatio pyramid tranfonet with pittree fusion and efficient-condconv swishnet. *IEEE Access* **12**, 113972–113987 (2024)
19. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE (2018)
20. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2016)
21. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2017)

22. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C.: MobileNetV2: inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE (2018)
23. Zhu, P., Zhang, Y., Huang, Y., Zhao, C., Zhao, K., Zhou, F.: Underwater acoustic target recognition based on spectrum component analysis of ship radiated noise. *Appl. Acoust.* **211**, 109552 (2023)
24. Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., Li, S.: Breast cancer multi-classification from histopathological images with structured deep learning model. *Sci. Rep.* **7**(1), 4172 (2017)



Underwater Acoustic Target Detection Network Based on Comprehensive Feature Extraction

Denghui Sun^{1,2}, Kaiqiao Wang^{1,3} , Peng Liu^{1,3}  , and Chun Zhang^{1,3}

¹ Hainan Acoustics Laboratory, Institute of Acoustics, Chinese Academy of Sciences,
Haikou 570105, China

liup@dsp.ac.cn

² University of Chinese Academy of Sciences, Beijing 100049, China

³ Lingshui, Marine Information, Hainan Observation and Research Station,
Lingshui 572423, China

Abstract. Underwater acoustic target detection is a task that presents both challenges and application potential. With the successful application of deep learning technology in computer vision and signal processing, an increasing number of researchers are applying these techniques and concepts to the field of underwater acoustic target detection, resulting in significant progress. However, existing methods still have room for performance improvement, as they often overlook the unique characteristics of underwater tasks. To address the challenges of high noise, poor image quality, and low target recognition accuracy in forward-looking sonar (FLS) images, This paper introduces a network designed to detect underwater acoustic targets, which is built on the principle of comprehensive feature extraction. By combining the convolution operation with the self-attention mechanism, this paper designed an enhanced version of the Comprehensive Feature Extraction (CFE) module. The convolution branch is a core component of the network that operates on the input image by sliding the convolution kernel to efficiently capture local features such as edges, textures, corners, and other relevant patterns, while the self-attention mechanism is designed to capture global patterns. We integrate the CFE module into the backbone of the baseline Yolov10 network and conduct experiments on two FLS image datasets. Experimental results show that our proposed model exhibits excellent detection performance in object detection tasks and demonstrates strong competitiveness in the field, objectively demonstrating the advantages of the new model.

Keywords: Underwater acoustic target detection · Sonar target detection · Comprehensive feature extraction

This research was partially supported by the Innovation Promotion Association, Chinese Academy of Sciences (Grant No. 2022022), and by the South China Sea Nova project of Hainan Province (Grant No. NHXXRCXM202340), and in part by the Scientific Research Foundation Project of Hainan Acoustics Laboratory under grant ZKNZ2024001.

1 Introduction

Underwater target detection focuses on identifying targets in the underwater environment, with important applications in oceanography [6], diving navigation [8], and aquaculture [11]. Sound waves have less attenuation in water compared to light waves and electromagnetic waves, making sonar devices effective for this purpose. Sonar equipment creates sonar images by emitting sound waves and capturing echoes, where pixel intensity indicates echo strength. Strong echoes result in high pixel values, while weak ones produce low values. Therefore, the use of advanced image processing methods can help quickly extract valuable information from such images, obtain the location information of the target in a very short time, and reduce the false detection rate and missed detection rate, significantly improving the performance of the underwater detection system.

Early methods [1, 10, 17, 33] usually focus on manually selecting and extracting certain key features in sonar images, leading to challenges like low efficiency, high complexity, and poor generalization. With the continuous development of deep learning technology, especially the continuous progress in hardware acceleration and algorithm optimization, the field of sonar image target detection has broken through the limitations of traditional methods, making the target detection results in complex underwater environments more accurate, fast, and reliable. Among the various deep learning-based object detection approaches, one-stage algorithms have gained widespread adoption due to their efficiency and relatively simpler architecture. These algorithms process images in a single pass, eliminating the need for region proposals or multiple stages of computation, which is common in two-stage methods. The Single Shot Multibox Detector (SSD [14]) and You Only Look Once (YOLO [20]) are two of the most well-known one-stage object detection frameworks. On the other hand, two-stage algorithms like Faster R-CNN [25] and Mask R-CNN [9] offer significantly better accuracy, making them suitable for applications where detecting small or obscured objects is of paramount importance. While two-stage models excel in terms of accuracy, they are computationally more expensive and slower, which leads to obvious limitations in handling application scenarios that require high real-time performance, especially in environments that require fast response. Nonetheless, their ability to deliver higher-quality results makes them ideal for applications such as post-processing analysis of sonar images, where accuracy is prioritized over speed.

The enhancements in one-stage target detection algorithms for optical images have contributed to progress in one-stage sonar image target detection. For example, Neves et al. [19] proposed a novel method for detecting rotated targets in sonar images using YOLOv2 [30] for initial positioning and a rotation attention mechanism to capture angle details. Yulin et al. [35] created the TransNet model, leveraging YOLOv3 [21] and transfer learning to effectively extract global features from sonar images. Kong et al. [12] integrated YOLOv3 [21] with a Dual Path Network (DPN) to develop the YOLOv3-DPFIN model, which captures multi-scale features via DPN and employs dense connections for extracting global deep semantic features. Yu et al. [34] improved feature extraction for underwa-

ter objects by adding Transformer [26] self-attention and coordinated attention within the backbone and neck of YOLOv5 [3]. While current sonar image target recognition methods can extract and enhance feature information, they still encounter challenges such as high redundancy, low semantic content, and the disregard for feature contextual relationships. Furthermore, sonar images are frequently impacted by noise, blur, and occlusion, increasing the requirements for effective feature extraction.

To overcome the challenges mentioned earlier, this study introduces an underwater acoustic target detection framework that utilizes Comprehensive Feature Extraction (CFE). By incorporating the CFE module into the backbone network, the model significantly boosts its feature extraction ability, ultimately enhancing the precision of target detection. Considering that a traditional convolutional network often struggles to capture global features due to its local nature and limited receptive field, we introduce the self-attention mechanism embedded in the Transformer [26]. The introduction of this mechanism enables the model to better understand the complex spatial relationships in the image, thereby achieving significant performance improvements in a series of visual tasks, especially in application scenarios that need to process long-distance dependencies and global information. Therefore, we have designed a lightweight CFE module that integrates convolution and self-attention mechanisms in the backbone network, improving the feature expression ability of the backbone model and achieving higher accuracy in object detection. Experiments were carried out using two sonar image datasets, MDD and WHD, which encompass a range of scenarios and varying levels of difficulty. Through systematic experimental comparison, it is proved that the model has shown stronger performance in multiple key indicators, including higher detection accuracy, faster reasoning speed, and better robustness. The innovations are as follows:

- This paper proposes a CFE module that can capture comprehensive contextual features in sonar images, including local and global features, by integrating convolution and self-attention mechanisms.
- We integrated the CFE module into the backbone network of the baseline network Yolov10 and conducted experiments on two FLS image datasets, MDD and WHD. The results of the experiments demonstrated that integrating our module into the baseline network enhances its detection performance.

2 Related Work

Traditional sonar image target detection methods mainly rely on classical image processing techniques and statistical methods. These methods typically include edge detection [2, 18], shape analysis [4, 32], and template based matching [7, 16, 31]. Edge detection algorithms such as Canny [5] and Sobel [13] operators are widely used in sonar images to extract potential target contours. Shape analysis techniques identify different types of targets by calculating geometric features such as aspect ratio, area, etc. In addition, template-based methods

achieve target detection by searching for regions in sonar images that are similar to pre-defined templates. The limitations of these traditional methods are also obvious. For example, these methods are sensitive to noise and can be significantly influenced by environmental factors like water flow and clutter, resulting in reduced detection accuracy. Furthermore, traditional methods often require manual adjustment of parameters and have poor adaptability to complex scenarios, making it difficult to learn and optimize automatically.

As deep learning technology advances, sonar image target detection algorithms utilizing deep learning networks have increasingly gained attention in research. Deep learning frameworks such as YOLO series algorithms, SSD [14], and Faster R-CNN [25] have been widely used in sonar image processing. These models effectively overcome the shortcomings of manual feature extraction in traditional methods through end-to-end training, combined with a substantial volume of annotated data, it can automatically recognize targets in various complex backgrounds. Among them, the YOLO series of algorithms, as a representative of the significant progress made in the field of object detection in recent years, has continuously optimized and improved optical image processing, which has improved the accuracy and efficiency of image recognition and target detection, and also provided strong technical support for breakthroughs in sonar image target detection. For example, Neves et al. [19] introduced an innovative method for rotation target detection in sonar images, using YOLOv2 [30] to determine initial position data and employing a rotation attention mechanism to obtain angle information. Yulin et al. [35] introduced the TransNet model, which integrates YOLOv3 [21] with transfer learning to fully extract global features from sonar images. Kong et al. [12] combined YOLOv3 [21] with a dual path network (DPN) to create the YOLOv3-DPFIN model, which obtains multi-scale features through DPN and extracts global deep semantic features using dense connection operations. Zhang et al. [36] incorporates transfer learning to address the challenge of limited sonar image datasets and upgrades the feature extraction network in YOLOv5 [3], significantly improving the detection accuracy for small targets in sonar images. YOLOv10 represents the most recent model in the YOLO series, by eliminating the requirement for non-maximum suppression (NMS) and enhancing the network structure, YOLOv10 enhances both detection performance and efficiency. However, the complex underwater environment poses challenges, as forward-scanning sonar images usually have low resolution and more noise characteristics, which will have some impact on how well the YOLOv10 network performs in detection tasks. To this end, we conducted in-depth research and improvements on the original convolution and attention fusion module, optimizing it to more effectively align with the specific characteristics of sonar images.

Among the various object detection methods available, the YOLO framework distinguishes itself with an exceptional balance of speed and accuracy, delivering rapid and dependable object recognition in images. Despite the promising prospects of YOLO-based algorithms in sonar image target detection, YOLO-based algorithms are general target detection algorithms, and their architectures are designed specifically for optical images. To further improve their effectiveness

in detecting targets within sonar images, designing specialized structures that align with the unique characteristics of sonar images is crucial. Consequently, this paper introduces a CFE module tailored for this purpose simultaneously, which is designed to integrate local and global image information, aiming to achieve further optimization of YOLOv10 in sonar image target detection.

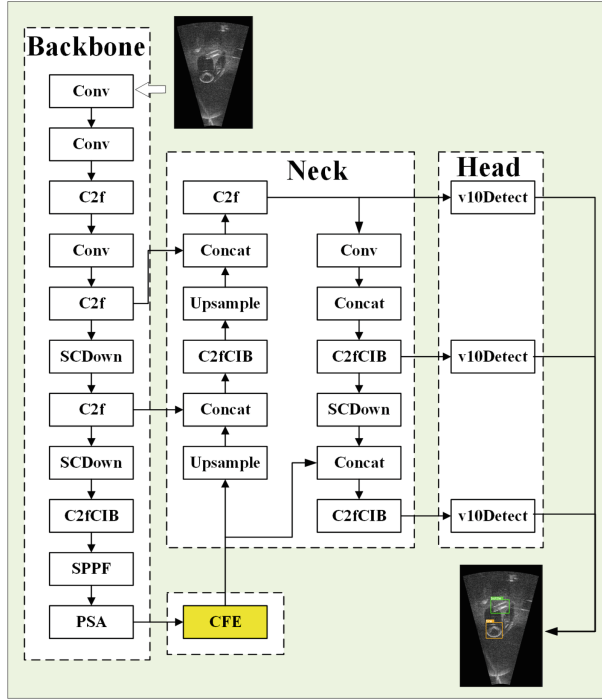


Fig. 1. The diagram illustration of the enhanced underwater acoustic target detection network.

3 Proposed Method

The general structure of the underwater acoustic target detection model is illustrated in Fig. 1 of this study. In this study, we employ the YOLOv10 model as the core architecture to improve the accuracy and efficiency of sonar image target detection. The YOLOv10 model consists of three core parts that complement each other: backbone, neck, and head. These three components work together and each has a specific function, making the entire detection process more accurate and efficient. The backbone consists of multiple convolutional layers and other deep learning modules, which can automatically identify different levels and details in the image. While the neck further refines these features for

improved processing. The head is responsible for outputting the category probability of the target and the coordinate information of the bounding box. As part of our innovation, we introduce the CFE module, which is integrated immediately after the backbone, as illustrated in Fig. 1. A comprehensive explanation of the CFE module will be provided in the subsequent sections.

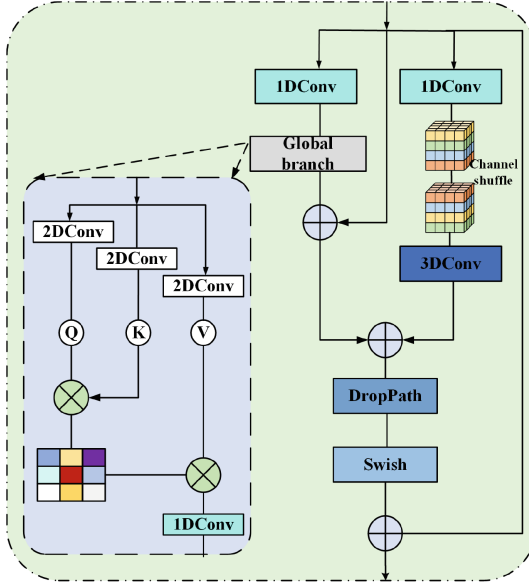


Fig. 2. The architecture of the proposed CFE module’s network.

Figure 2 illustrates the architecture of the CFE module, which mainly includes a local part and a global part. In the case of the local part, in the process of processing local features, improving the interactivity and information fusion between channels is the key to enhancing model performance. To this end, this study uses two effective techniques: one-dimensional convolution and channel shuffling operations. First, through the one-dimensional convolution operation, we can effectively adjust the channel dimension of the input feature to create better conditions for subsequent feature fusion. Secondly, channel shuffling allows information to flow more flexibly between different channels by rearranging the relationship between channels. The channel shuffle technique is an innovative method for improving the feature extraction capabilities of deep learning models when processing complex input data. First, during the channel shuffle process, the input tensor is divided into several groups along the channel axis, and the originally independent channels are assigned to different groups, providing more flexibility for subsequent convolution operations. Next, a depth-wise separable convolution operation is used within each group to randomize the

channels. Finally, each group of output tensors after the depthwise separable convolution is merged along the channel axis. This merge operation integrates the information from different groups to generate a new tensor that incorporates the features extracted from all groups. Next, we apply a 3D convolution operation for feature extraction. The Local part formula is as follows:

$$F_{conv} = W_{3 \times 3 \times 3}(CS(W_{1 \times 1}(Y))) \quad (1)$$

where F_{conv} refers to the output from the local part, $W_{1 \times 1}$ represents the weight matrix for the 1D convolution, $3 \times 3 \times 3$ represents the weight matrix for the 3D convolution, The operation CS represents the channel randomization, and Y corresponds to the input features of the local part.

In the global part, we first adjust the channel dimension through 1D convolution. Next, different features are extracted from the input data through three 2D convolution operations. The query(Q) guides the model to find relevant feature information in the input data. The key(K) helps the model determine the relevance of the input data to the current query, thereby improving the accuracy of information selection. The value(V) generates a weighted output through pairing. The Q, K, and V are generated through this three-layer convolution, yielding three tensors with dimensions of $H \times W \times C$. Then, we compute the attention mechanism for Q, K, and V, as described below:

$$F_{att} = W_{1 \times 1}Attention(Q, K, V) + Y \quad (2)$$

$$Attention(Q, K, V) = softmax(\frac{QK^T}{\alpha})V \quad (3)$$

where F_{att} is the output of the global branch, α serves as a trainable scaling factor, that adjusts the magnitude of the matrix multiplication between QK^T prior to the $softmax$ activation. This scaling parameter helps regulate the influence of the attention mechanism. The final calculation for the output of the CFE module is then expressed as follows:

$$F_{out} = Swish(DP(F_{att} + F_{conv})) + Y \quad (4)$$

$$Swish(x) = x\sigma(x\beta) \quad (5)$$

where DP is a DropPath regularization technique that prevents overfitting and improves the model's generalization ability by randomly discarding some paths during the training process, Swish is a nonlinear activation function that effectively reduces the gradient vanishing problem, allowing better gradient flow, making the training process more stable, and accelerating network learning and performance improvement, σ is the sigmoid activation function.

4 Experiments

4.1 Datasets

This paper utilized the MDD dataset [24] and the WHD dataset [29] to evaluate the proposed object detection method. The MDD dataset was collected by ARIS Explorer 3000 FLS (3.0 MHz frequency) from the Marine Systems Laboratory of Heriot-Watt University. It contains 1868 static images covering ten target categories: tires, cans, bottles, hooks, beverage cartons, shampoo bottles, propellers, upright bottles, chains, and valves. The WHD dataset was collected by Tritech 1200ik FLS in Weihai, China. It contains 4000 labeled images. There are eight categories used for detection tasks: tires, cubes, spheres, human bodies, cylinders, metal barrels, round cages, and square cages. Referring to previous studies [15], these two datasets were randomly partitioned into three distinct subsets to ensure proper model training, evaluation, and validation. Specifically, the datasets were divided in a ratio of 7:2:1, the training set accounts for 70% of the dataset, and the validation set and test set account for 20% and 10% of the dataset respectively, which can ensure that the model is fully trained, optimized, and evaluated at different stages. The training, validation, and test sets of MDD and WHD are shown in Table 1.

Table 1. THE DATASETS OF MDD AND WHD.

Set	MDD		WHD	
	Number of Images	Number of Labels	Number of Images	Number of Labels
Training	1345	1345	2880	5068
Validation	149	149	320	565
Test	374	374	800	1444
Total	1868	1868	4000	7077

4.2 Experiments Setup

In this study, all experiments were performed utilizing the Ultralytics YOLOv10 framework, leveraging the capabilities of an NVIDIA GeForce RTX 3090 GPU, with computations facilitated by PyTorch 2.0.1. For optimization, the system was configured to an automatic setting, and the training process employed the Stochastic Gradient Descent optimizer. The learning rate was initialized at $1e-2$, with a momentum value of 0.9, which was used to enhance convergence during model training. Set the batch size to 32, and the improved model was set for 500 epochs. Our input FLS images were resized to 640×640 for training, utilizing multi-GPU parallel processing. From input data to output prediction, all processing steps in the model training process are performed through a unified framework. By using box loss, class loss, and distribution focus loss, the

model can simultaneously optimize the target positioning accuracy, and category recognition accuracy, and increase the attention to difficult-to-classify samples, thereby improving the overall performance of the model in a variety of target detection tasks.

4.3 Metrics

We employ the Mean Average Precision (mAP) to measure the overall accuracy of the model in multi-class object detection, mAP comprehensively considers the prediction precision and recall of the model in each target category, providing a unified evaluation metric to help compare the performance of different methods in handling object detection tasks. mAP is defined as the average accuracy of all target categories n calculated based on a fixed intersection over union (IoU) threshold, and is an important evaluation metric in object detection. Specifically, mAP@0.5 refers to the average accuracy when the IoU threshold is set to 0.5, the IoU threshold of 0.5 means that when the overlap area between the predicted box and the true box exceeds 50%, the detection is considered correct. The equations for Average Precision and mean Average Precision are given below:

$$AP = \int_0^1 P(r)dr \quad (6)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (7)$$

where r is the integration variable, and $P(r)$, represents the PR curve, which is particularly useful for processing class-imbalanced datasets. The horizontal axis indicates the recall rate, which is the proportion of actual positive examples correctly identified by the detection model, while the vertical axis shows the precision rate, which is the proportion of predicted positive examples that are correctly identified as positive by the model. The area beneath this curve corresponds to the AP score, which quantifies the average precision of the algorithm for a single category. The following are their formulas:

$$Precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

where TP indicates the count of samples that are actually positive and correctly predicted as positive (true positives), FP indicates the count of samples that are actually negative but are mispredicted as positive (false positives), and FN indicates the count of samples that are actually positive but are mispredicted as negative (false negatives). Precision measures the proportion of samples that are actually positive among the samples predicted by the model to be positive, in other words, precision focuses on how many samples predicted to be positive are

correct. Recall measures the proportion of all samples that are actually positive that the model can successfully identify as positive, recall focuses on the proportion of all actual positive samples that are correctly predicted to be positive. By combining these indicators, we can more comprehensively and objectively evaluate the performance of the model in different tasks and adjust the optimization strategy for specific tasks.

4.4 Results and Comparisons

In order to show that our proposed model is highly competitive in the field of object detection, we compared it with leading object detection algorithms to verify its actual ability in object detection tasks, which helped us to more objectively demonstrate the advantages of the new model. This includes general object detection algorithms like Faster RCNN [22], RetinaNet [23], and YOLOv10 [27], as well as sonar image object detection algorithms MBSNN [28] and UFIDNet [15].

Table 2. Performance comparison of our proposed model and advanced target detection algorithms on MDD and WHD.

Datasets	Model	mAP ₅₀₋₉₅ (%)	mAP ₅₀ (%)	mAP ₇₅ (%)
MDD	Faster RCNN	60.2	87.6	72.8
	RetinaNet	60.9	82.6	78.3
	MBSNN	52.5	73.7	61.9
	UFIDNet	70.5	88.0	83.4
	YOLOv10	75.7	90.9	84.7
	Ours	77.0	93.6	87.6
WHD	Faster RCNN	37.8	82.3	24.0
	RetinaNet	37.2	78.4	27.1
	MBSNN	38.9	85.3	25.6
	UFIDNet	47.3	90.9	43.0
	YOLOv10	55.3	95.8	47.3
	Ours	56.5	96.3	48.8

The experimental results demonstrate that the proposed method effectively addresses the challenges presented by the two datasets, exhibiting strong robustness and generalizability in the target detection task. As shown in Table 2, compared to other algorithms, the results confirm the competitiveness of the proposed model in the field of target detection, highlighting its advantages in this task, specifically general detectors and those tailored for sonar images. This suggests that the model is not only effective but also versatile, performing well in processing sonar image data, while YOLOv10 also demonstrated good performance, achieving second-best results across all three metrics. Specifically, for

the MDD dataset, our model achieved a mAP_{50-95} score of 77.0%, which is 1.3% higher than YOLOv10. Additionally, the proposed model can perform well in multiple evaluation indicators, especially in the two key indicators of mAP_{50} and mAP_{75} , which are better than other traditional or advanced object detection methods, with scores of 93.6% and 87.6%, respectively, surpassing YOLOv10 by 2.7% and 2.9%.

Unlike MDD, which is a single-target detection task, WHD is a multi-target detection task and therefore more challenging. Overall, the three metrics for WHD are lower than their corresponding metrics for MDD. Nevertheless, our model still achieved the best performance on this dataset, attaining mAP_{50-95} , mAP_{50} , and mAP_{75} scores of 56.5%, 96.3%, and 48.8%, respectively. Since the only difference between our model and YOLOv10 is the CFE module, we attribute this improvement to the application of the CFE module. The CFE module enables the model to extract more comprehensive features from sonar images, improving the ability to represent features and resulting in increased detection precision.

5 Conclusion

In the present research, we designed a CFE module and integrated it into the YOLOv10 framework to construct an underwater acoustic target detection network. To thoroughly assess the performance and applicability of this network, we performed a set of experiments using two sonar image datasets, each presenting distinct levels of complexity and challenges. The findings show that the improved model greatly surpasses both existing sonar image detectors and general object detection models, demonstrating its adaptability in different environments. While acknowledging the progress made, we also recognize that further improvements are needed, particularly in optimizing the model's efficiency and speeding up its inference process. To tackle this challenge, upcoming research will aim to investigate methods for optimizing the model architecture to reduce computational burdens and enhance inference efficiency, thereby further increasing the practical applicability and operability of the proposed model.

References

1. Abu, A., Diamant, R.: A statistically-based method for the detection of underwater objects in sonar imagery. *IEEE Sens. J.* **19**(16), 6858–6871 (2019). <https://doi.org/10.1109/JSEN.2019.2912325>
2. Anitha, U., Malarkkan, S., Premalatha, J., Prince, P.G.K.: Study of object detection in sonar image using image segmentation and edge detection methods. *Indian J. Sci. Technol.* **9**(42) (2016)
3. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M., Liu, Y.J.: YOLOv5: you only look once v5: more. *arXiv preprint arXiv:2007.12092* (2020)
4. Brandes, T.S., Ballard, B., Ramakrishnan, S., Lockhart, E., Marchand, B., Rabenold, P.: Environmentally adaptive automated recognition of underwater mines with synthetic aperture sonar imagery. *J. Acoust. Soc. Am.* **150**(2), 851–863 (2021)

5. Canny, J.: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 679–698 (1986)
6. Chen, L., et al.: Underwater object detection using invert multi-class adaboost with deep learning. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2020)
7. Dobeck, G.J., Hyland, J.C., et al.: Automated detection and classification of sea mines in sonar imagery. In: *Detection and Remediation Technologies for Mines and Minelike Targets II*, vol. 3079, pp. 90–110. SPIE (1997)
8. Fan, B., Chen, W., Cong, Y., Tian, J.: Dual refinement underwater object detection network. In: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (eds.) *ECCV 2020*. LNCS, vol. 12365, pp. 275–291. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58565-5_17
9. He, K., Gkioxari, G., Dollár, P., Girshick, R.: Mask R-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–2969 (2017)
10. Hurtós, N., Palomeras, N., Carrera, A., Carreras, M.: Autonomous detection, following and mapping of an underwater chain using sonar. *Ocean Eng.* **130**, 336–350 (2017)
11. Jian, Z., Liang, Z., Li-nan, Z., Nan, L.: The design and research of intelligent search and rescue device based on sonar detection and marine battery. In: 2017 International Conference on Computer Network, Electronic and Automation (ICCNEA), pp. 383–387 (2017). <https://doi.org/10.1109/ICCNEA.2017.73>
12. Kong, W., et al.: YOLOv3-DPFIN: a dual-path feature fusion neural network for robust real-time sonar target detection. *IEEE Sens. J.* **20**(7), 3745–3756 (2019)
13. Lateef, R.A.R.: Expansion and implementation of a 3×3 sobel and prewitt edge detection filter to a 5×5 dimension filter. *Baghdad College of Economic Sciences University*, pp. 336–348 (2008)
14. Liu, W., et al.: SSD: single shot MultiBox detector. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9905, pp. 21–37. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_2
15. Long, H., Shen, L., Wang, Z., Chen, J.: Underwater forward-looking sonar images target detection via speckle reduction and scene prior. *IEEE Trans. Geosci. Remote Sens.* **61**, 1–13 (2023)
16. Lopera, O., Dupont, Y.: Automated target recognition with SAS: shadow and highlight-based classification. In: 2012 Oceans, pp. 1–5. IEEE (2012)
17. McKay, J., Monga, V., Raj, R.G.: Robust sonar ATR through bayesian pose-corrected sparse classification. *IEEE Trans. Geosci. Remote Sens.* **55**(10), 5563–5576 (2017). <https://doi.org/10.1109/TGRS.2017.2710040>
18. Mukherjee, K., Gupta, S., Ray, A., Phoha, S.: Symbolic analysis of sonar data for underwater target detection. *IEEE J. Oceanic Eng.* **36**(2), 219–230 (2011)
19. Neves, G., Ruiz, M., Fontinele, J., Oliveira, L.: Rotated object detection with forward-looking sonar in underwater applications. *Expert Syst. Appl.* **140**, 112870 (2020)
20. Redmon, J.: You only look once: unified, real-time object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016)
21. Redmon, J.: YOLOv3: an incremental improvement. *arXiv preprint arXiv:1804.02767* (2018)
22. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1137–1149 (2016)

23. Ross, T.Y., Dollár, G.: Focal loss for dense object detection. In: proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2980–2988 (2017)
24. Singh, D., Valdenegro-Toro, M.: The marine debris dataset for forward-looking sonar semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3741–3749 (2021)
25. Sun, X., Wu, P., Hoi, S.C.: Face detection using deep learning: an improved faster RCNN approach. *Neurocomputing* **299**, 42–50 (2018)
26. Vaswani, A.: Attention is all you need. In: *Advances in Neural Information Processing Systems* (2017)
27. Wang, A., et al.: YOLOv10: real-time end-to-end object detection. arXiv preprint [arXiv:2405.14458](https://arxiv.org/abs/2405.14458) (2024)
28. Wang, J., Feng, C., Wang, L., Li, G., He, B.: Detection of weak and small targets in forward-looking sonar image using multi-branch shuttle neural network. *IEEE Sens. J.* **22**(7), 6772–6783 (2022)
29. Wang, Z., Zhang, S., Huang, W., Guo, J., Zeng, L.: Sonar image target detection based on adaptive global feature enhancement network. *IEEE Sens. J.* **22**(2), 1509–1530 (2021)
30. Wang, Z., Xu, K., Wu, S., Liu, L., Liu, L., Wang, D.: Sparse-yolo: hardware/software co-design of an FPGA accelerator for yolov2. *IEEE Access* **8**, 116569–116585 (2020). <https://doi.org/10.1109/ACCESS.2020.3004198>
31. Williams, D.P.: Fast target detection in synthetic aperture sonar imagery: a new algorithm and large-scale performance analysis. *IEEE J. Oceanic Eng.* **40**(1), 71–92 (2014)
32. Williams, D.P.: The mondrian detection algorithm for sonar imagery. *IEEE Trans. Geosci. Remote Sens.* **56**(2), 1091–1102 (2017)
33. Williams, D.P.: The mondrian detection algorithm for sonar imagery. *IEEE Trans. Geosci. Remote Sens.* **56**(2), 1091–1102 (2018). <https://doi.org/10.1109/TGRS.2017.2758808>
34. Yu, Y., Zhao, J., Gong, Q., Huang, C., Zheng, G., Ma, J.: Real-time underwater maritime object detection in side-scan sonar images based on transformer-yolov5. *Remote Sens.* **13**(18), 3555 (2021)
35. Yulin, T., Jin, S., Bian, G., Zhang, Y.: Shipwreck target recognition in side-scan sonar images by improved yolov3 model based on transfer learning. *IEEE Access* **8**, 173450–173460 (2020). <https://doi.org/10.1109/ACCESS.2020.3024813>
36. Zhang, H., Tian, M., Shao, G., Cheng, J., Liu, J.: Target detection of forward-looking sonar image based on improved yolov5. *IEEE Access* **10**, 18023–18034 (2022)



Research on a Temperature Prediction Model for Meteorological Data Integrating XGBoost and MLP

Wenhua Xue^(✉), Yuxing Gao, Long Chen, and Enqi Lv

Xi'an Meteorological Bureau, No. 102-1 Weiyang Road, Xi'an, Shaanxi Province, China
396397223@qq.com

Abstract. This study develops an efficient temperature prediction model by integrating multiple machine learning techniques to enhance the accuracy of meteorological data analysis. Utilizing ten years of meteorological data encompassing features such as relative humidity, cloud cover, visibility, precipitation, wind speed, and sunshine duration, we performed comprehensive data preprocessing and feature selection. Initially, missing values were addressed, and wind direction was encoded. Important features were then identified using the mutual information method. Subsequently, data normalization was carried out through standardization, and temporal as well as lag features were extracted to bolster the model's predictive capabilities. For model construction, we selected XGBoost (Extreme Gradient Boosting) and MLP (Multilayer Perceptron) as base models and employed linear regression as the final predictor, thereby establishing a stacked regression framework. The dataset was divided into training and testing sets with an 8:2 ratio. The model achieved a Root Mean Squared Error (RMSE) of 1.34 and a Coefficient of Determination (R^2) of 0.98 on the training set, and an RMSE of 1.74 and R^2 of 0.97 on the testing set. Additionally, 5-fold cross-validation yielded an average RMSE of 1.92, demonstrating the model's stability and generalization performance.

Keywords: Temperature Prediction Model · Machine Learning · XGBoost · MLP · Feature Selection

1 Introduction

Temperature prediction is a crucial topic in meteorological research, significantly impacting various sectors such as agriculture, transportation, and energy. It also plays a vital role in disaster prevention, mitigation, and addressing climate change. Accurate temperature forecasting can help individuals better respond to extreme weather events, thereby reducing their negative impacts on the economy and human life [1]. In recent years, with the rapid development of machine learning technologies, an increasing number of researchers have begun applying these techniques to meteorological data analysis and temperature prediction to enhance prediction accuracy and stability [2].

In the field of temperature prediction, traditional statistical methods such as time series analysis and linear regression have been widely used [3, 4]. However, with the

advancement of machine learning, more studies are exploring the use of machine learning models for temperature forecasting. For example, models like Support Vector Machines (SVM), Random Forests (RF), and Long Short-Term Memory (LSTM) networks have demonstrated excellent performance in temperature prediction [5, 6]. XGBoost, an efficient gradient boosting decision tree algorithm, has gained widespread application in meteorological data analysis due to its advantages in handling nonlinear relationships and high-dimensional data [7]. On the other hand, Multilayer Perceptron (MLP), a classic neural network model, possesses strong nonlinear fitting capabilities, enabling it to capture complex temperature variation patterns [8]. Although single models have achieved certain success in temperature prediction, they often struggle to comprehensively capture the diversity and complexity of temperature changes [9]. Consequently, ensemble learning methods that integrate multiple models have gradually become a research hotspot. Ensemble learning can effectively improve prediction accuracy and stability by combining the outputs of multiple base models [10, 11].

In this study, we utilized meteorological data from the Jinghe Meteorological Station in the Xi'an region, spanning from January 2013 to July 2024. The dataset includes multiple features such as Average Relative Humidity (ARH), Minimum Relative Humidity (MRH), Average Total Cloud Cover (ATCC), Minimum Horizontal Visibility (MHV), Precipitation (P), Average Wind Speed (AWS), Maximum Daily Wind Direction (MDWD), Maximum Wind Speed (MWS), Extreme Wind Direction (EWD), Extreme Wind Speed (EWS), and Sunshine Duration (SD). These data were preprocessed and subjected to feature selection to ensure that the model could capture the key factors influencing temperature changes. Initially, missing values were handled, and wind direction was encoded, followed by the selection of important features using the mutual information method [12, 13]. Handling missing values is a crucial step in data preprocessing, as it effectively improves data integrity and model prediction performance. Wind direction encoding converts the categorical wind direction feature into numerical values, allowing the model to better understand and utilize this information. Subsequently, data normalization was performed using standardization, and temporal and lag features were extracted to enhance the model's predictive capabilities. Standardization eliminates the dimensional differences between different features, making model training more stable and efficient [14]. The extraction of temporal and lag features aims to capture the temporal dependencies and trend information of temperature changes, which are essential for improving prediction accuracy [15].

During the model construction phase, we selected XGBoost (Extreme Gradient Boosting) and MLP (Multilayer Perceptron) as base models and employed linear regression as the final predictor to build a stacked regression model. XGBoost is an ensemble learning algorithm based on gradient boosting, capable of handling high-dimensional data and nonlinear relationships [16]. MLP is a feedforward neural network that can capture complex feature relationships through multiple layers of nonlinear transformations [17]. The stacked regression model combines the predictions of multiple base models, thereby further enhancing prediction accuracy and stability [18]. Additionally, we validated the model's stability and generalization ability using 5-fold cross-validation [19].

This study not only theoretically explores the application of integrating multiple machine learning models in temperature prediction but also empirically demonstrates the effectiveness and reliability of this approach in practical applications. In the future, we plan to further optimize model parameters and explore additional feature selection and data preprocessing methods to further improve temperature prediction accuracy and stability. We also intend to apply this model to predict other meteorological variables, such as precipitation and wind speed, to verify its applicability and generalizability under different meteorological conditions.

Through this research, we aim to provide a new perspective and methodology for meteorological data analysis and temperature prediction, thereby advancing the development of meteorological science and offering robust support for related fields of research and application.

2 Data and Methods

2.1 Data Preprocessing

The data used in this study were obtained from the Jinghe Meteorological Station of the Xi'an Meteorological Information Center, spanning daily meteorological records from January 1, 2013, to July 1, 2024. The dataset includes multiple features: Average Relative Humidity (ARH), Minimum Relative Humidity (MRH), Average Total Cloud Cover (ATCC), Minimum Horizontal Visibility (MHV), Precipitation (P), Average Wind Speed (AWS), Maximum Daily Wind Direction (MDWD), Maximum Wind Speed (MWS), Extreme Wind Direction (EWD), Extreme Wind Speed (EWS), and Sunshine Duration (SD). The ten-year variations of these meteorological parameters are illustrated in Fig. 1, the frequency of data changes is shown in Fig. 2, and the statistical results are presented in Table 1.

From the information provided by Fig. 1, Fig. 2, and Table 1, it is evident that the average temperature exhibits a wide range, from -6.7°C to 35.7°C , with a mean of 15.53°C and a standard deviation of 9.77°C , indicating significant fluctuations in average temperature. Similarly, other meteorological factors also show considerable variability without any consistent patterns. Outliers in the dataset were replaced with NaN (Not a Number) and subsequently filled using the column-wise mean to handle missing values. To enable the model to process non-numeric data and select the most relevant features, several preprocessing steps were undertaken, including wind direction encoding, mutual information computation and feature selection, data normalization, extraction of temporal features, and generation of lag features. These steps provide a fundamental understanding of the data's basic characteristics and distribution, thereby facilitating further analysis and model development.

2.2 Model Construction

To identify the features that most significantly influence the prediction target, Average Temperature (AT), we first conducted a correlation analysis between AT and other meteorological factors [20], as illustrated in Fig. 3. From Fig. 3, it is evident that AT exhibits a

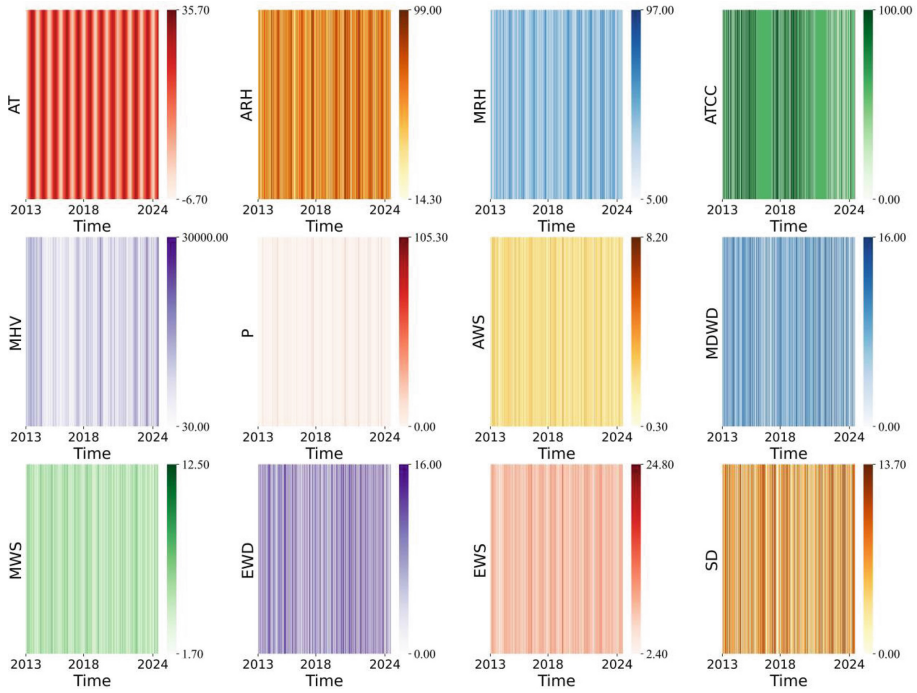


Fig. 1. Overview of Meteorological Data Changes from January 2013 to July 2024 at Jinghe Observation Station in Xi'an

positive correlation with Minimum Horizontal Visibility (MHV) and Sunshine Duration (SD), with a particularly strong correlation with SD, having a correlation coefficient of 0.33. Conversely, AT shows a negative correlation with Average Total Cloud Cover (ATCC) and Minimum Relative Humidity (MRH). However, the correlation results do not allow for an intuitive extraction of the features that have a substantial impact on AT.

To further explore the features that influence Average Temperature (AT), we calculated the mutual information values between each feature and the target variable, as shown in Eq. (1) [21, 22].

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (1)$$

where $p(x, y)$ is the joint probability distribution of X and Y , $p(x)$ is the marginal probability distribution of X , and $p(y)$ is the marginal probability distribution of Y .

We selected mutual information for feature selection due to its ability to capture both linear and nonlinear relationships between variables, providing a more comprehensive measure of feature relevance compared to traditional correlation coefficients. Additionally, mutual information does not assume any specific distribution of the data, making it suitable for the diverse range of meteorological variables considered in this study.

The mutual information results between AT and other meteorological factors are presented in Fig. 4. Mutual information was calculated using the `mutual_info_regression`

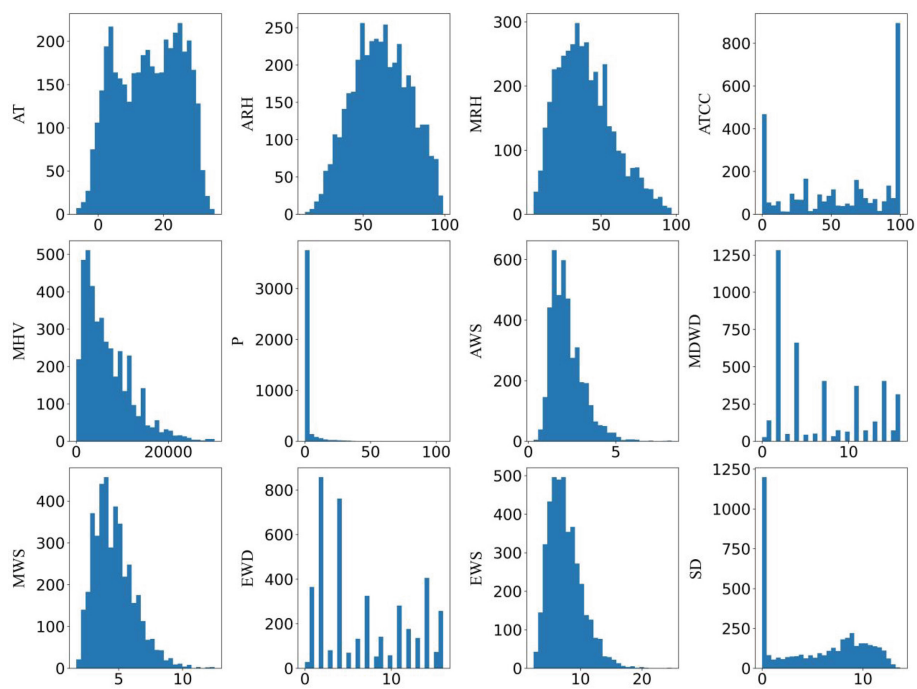


Fig. 2. Frequency Distribution of Meteorological Data from January 2013 to July 2024 at Jinghe Observation Station in Xi'an

Table 1. Meteorological Data Statistics from January 2013 to July 2024 at Jinghe Observation Station in Xi'an.

	AT	ARH	MRH	ATCC	MHV	P	AWS	MDWD	MWS	EWD	EWS	SD
count	4170	4172	4171	3258	4171	4191	4172	4199	4171	4199	4171	4158
mean	15.5	60.8	39.8	58.6	6814.1	1.6	2.2	6.9	4.7	6.9	7.5	5.4
std	9.8	17.5	18.6	36.8	5229.7	6.0	0.9	5.1	1.6	5.0	2.7	4.4
min	-6.7	14.3	5.0	0.0	30.0	0.0	0.3	0.0	1.7	0.0	2.4	0.0
25%	6.9	48.0	26.0	26.7	2716.5	0.0	1.5	2.0	3.5	2.0	5.5	0.0
50%	15.9	60.5	37.0	66.7	5480.0	0.0	2.0	4.0	4.4	5.0	7.1	6.0
75%	23.9	74.0	51.0	99.3	10000.0	0.0	2.7	11.0	5.6	11.0	9.1	9.2
max	35.7	99.0	97.0	100.0	30000.0	105.3	8.2	16.0	12.5	16.0	24.8	13.7

function from the Scikit-learn library, which estimates the dependency between each feature and the target variable without assuming any particular distribution. This method ensures that both linear and nonlinear dependencies are effectively captured. From Fig. 4, it can be observed that the mutual information values of Average Relative Humidity (ARH), Minimum Relative Humidity (MRH), Minimum Horizontal Visibility (MHV),

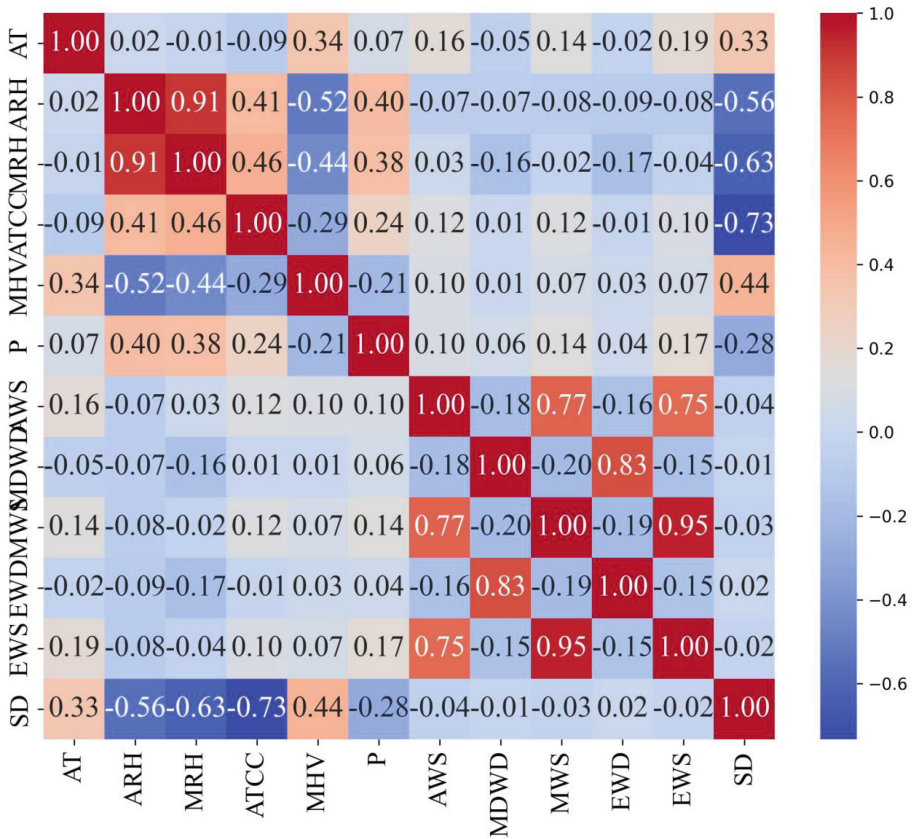


Fig. 3. Correlation Analysis between AT and Other Meteorological Elements

and Sunshine Duration (SD) exceed 0.1. This indicates that these factors share a substantial amount of information with AT, thereby being considered important features for predicting or analyzing AT. The threshold of 0.1 was selected based on preliminary experiments that demonstrated improved model performance when selecting features above this value. This cutoff ensures that only the most informative features are included, reducing the dimensionality of the dataset and mitigating the risk of overfitting. Specifically, the mutual information values are 0.1397 for ARH, 0.1064 for MRH, 0.1191 for MHV, and the highest is 0.1816 for SD. These results are consistent with the correlation analysis shown in Fig. 3. The relatively high mutual information values of these factors suggest a strong association with AT, making them key factors in forecasting temperature changes.

To ensure the accuracy of the model analysis, we standardized the selected features so that each feature has a mean of 0 and a standard deviation of 1, which helps eliminate the dimensional differences between different features [23]. When handling time series data, we considered the effects of seasonality and periodicity by extracting temporal features such as month, season, and whether the day is a working day. Additionally,

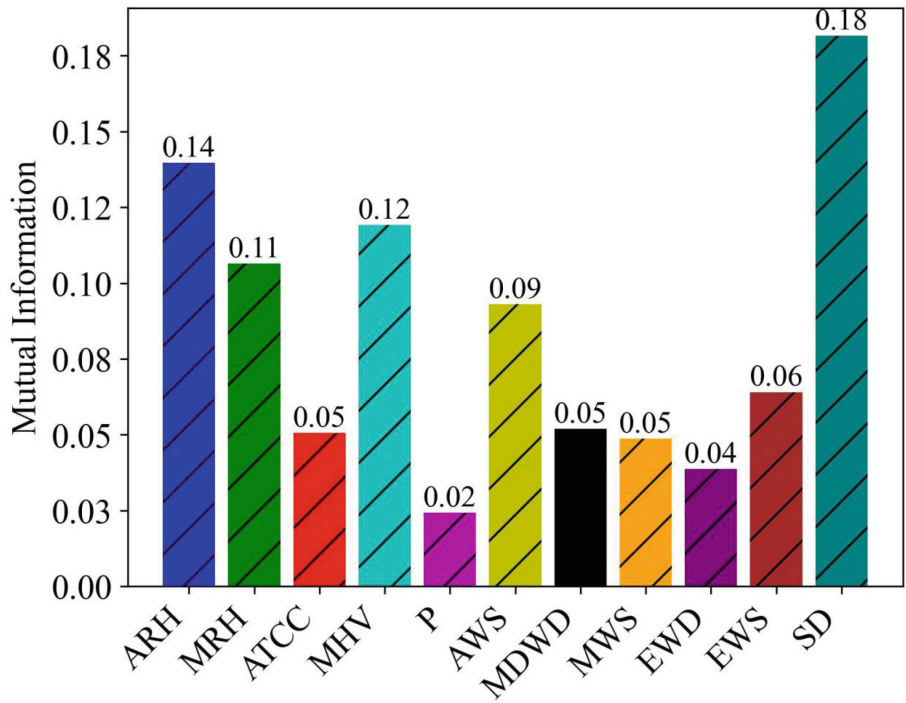


Fig. 4. Mutual Information Results between AT and Other Meteorological Elements

we constructed lag features, including the average temperatures of the previous 1 day, 2 days, and 3 days, to capture the temporal dependencies of temperature. To evaluate the model’s generalization capability, we split the data chronologically into training and testing sets with a ratio of 8:2, using data from January 1, 2013, to March 14, 2022, as the training set and data from March 15, 2022, to July 1, 2024, as the testing set, ensuring that the model was not exposed to the testing set data during the training process.

In this study, a stacking regressor method was employed to construct the prediction model. Stacking regression is an ensemble learning technique that enhances the model’s predictive performance by combining the predictions of multiple base models. We selected two base models: the XGBoost regression model and the Multilayer Perceptron (MLP) regression model, and used linear regression as the final meta-learner. XGBoost is a gradient boosting tree model characterized by its efficiency, flexibility, and accuracy. It optimizes the objective function by incrementally building decision trees, effectively handling missing values and preventing overfitting. The specific parameters for XGBoost were chosen based on a combination of grid search optimization and cross-validation to identify the optimal configuration. The parameters were set as follows: maximum depth (`max_depth`) = 5, learning rate (`learning_rate`) = 0.1, number of base learners (`n_estimators`) = 100, and random seed (`random_state`) = 42. These settings were selected to balance model complexity and computational efficiency, ensuring robust performance without overfitting.

MLP is a feedforward neural network capable of capturing nonlinear relationships in the data. We selected an MLP model with a single hidden layer consisting of 50 neurons and employed L2 regularization (alpha) to prevent overfitting. The hyperparameters for the MLP were determined through a systematic hyperparameter tuning process using grid search combined with cross-validation. Specifically, the hidden layer sizes (hidden_layer_sizes) were set to 50 based on preliminary experiments that indicated this configuration provided a good trade-off between model capacity and training time. The regularization parameter (alpha) was tuned to 0.01 to effectively mitigate overfitting, and the maximum number of iterations (max_iter) was set to 1000 to ensure convergence. The random seed (random_state) was fixed at 42 to guarantee reproducibility of the results.

The stacking regression model utilizes the predictions of the base models as inputs and employs a meta-learner for the final prediction. In this study, linear regression was chosen as the meta-learner. The linear regression model effectively and simply integrates the predictions of the base models, further enhancing predictive performance.

2.3 Model Evaluation and Validation

To evaluate the performance of the stacking regression model, we calculated the Root Mean Square Error (RMSE) and the coefficient of determination (R²) on both the training and testing sets, as shown in Eqs. (2) and (3) [24]. Additionally, we conducted 5-fold cross-validation and computed the mean and standard deviation of the cross-validated RMSE to assess the model's stability and generalization capability.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

where n is the number of samples, y_i is the actual value of the i th sample, $\hat{y}_i$ is the predicted value of the i th sample, and $\bar{y}$ is the mean of the actual values.

We calculated the RMSE and the R^2 of the stacking regression model on both the training and testing sets to evaluate the model's performance on known and unseen data. To assess the model's stability and generalization capability, we conducted 5-fold cross-validation. First, the dataset was divided into five mutually exclusive subsets. In each iteration, four subsets were used to train the model, and the remaining subset was used for validation [25]. This process was repeated five times, with each subset serving as the validation set once. For each partition, the stacking regression model was trained and the RMSE on the validation set was calculated. The mean and standard deviation of the cross-validated RMSE is expressed by Eqs. (4) and (5), respectively.

$$\text{Mean RMSECV} = \frac{1}{5} \sum_{k=1}^5 \text{RMSE}_k \quad (4)$$

$$\text{Std RMSECV} = \sqrt{\frac{1}{5} \sum_{k=1}^5 (\text{RMSE}_k - \text{Mean RMSECV})^2} \quad (5)$$

Through the aforementioned evaluation metrics and methods, we were able to comprehensively assess the performance of the stacking regression model. This ensures that the model demonstrates strong predictive capabilities on both the training and testing sets and maintains stable performance across different data partitions.

3 Results and Analysis

Figures 5 and 6 illustrate the comparison between the actual and predicted values of AT, as well as the fitting curves for the actual versus predicted AT values using the XGBoost-MLP ensemble model and the MLP single model, respectively. From these figures, it is evident that the model effectively captures the overall trend of the data. The R2 for the training and validation sets are 0.981 and 0.9689, respectively, while the RMSE values are 1.3091 and 1.6948, respectively. These metrics indicate that the model possesses excellent overall performance, lacks significant outliers, and does not exhibit overfitting on the training data.

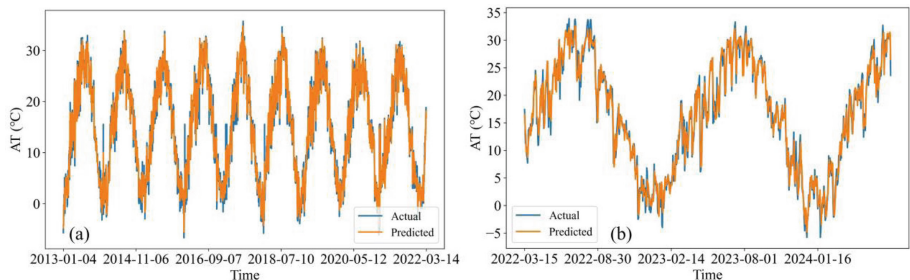


Fig. 5. Comparison of Actual and Predicted Values of AT (a) Training Set (b) Validation Set

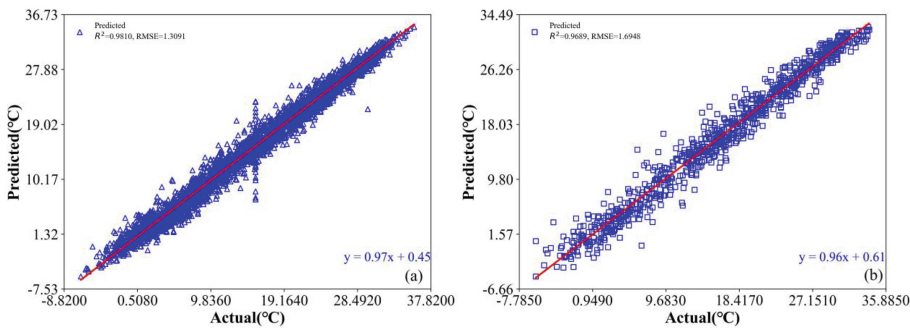


Fig. 6. Fitting Curves of Actual and Predicted Values of AT (a) Training Set (b) Validation Set

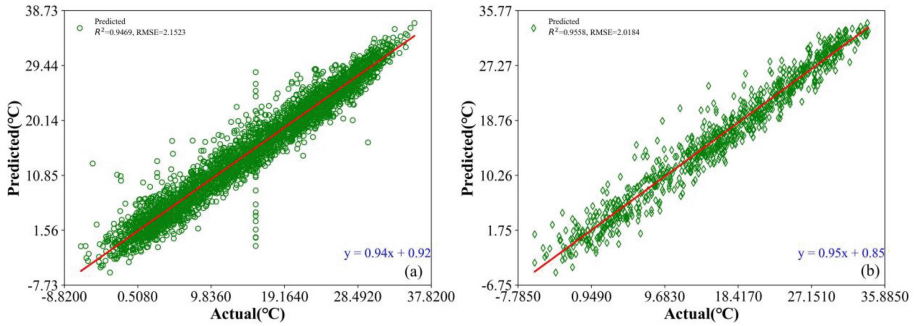


Fig. 7. Fitting Curves of Actual and Predicted Values of AT using MLP Single Model (a) Training Set (b) Validation Set

Figure 7 presents the fitting curves of the actual versus predicted AT values using the single MLP model. Combined with the comparison results between multiple models and the single model in Table 2, it is evident that, in the performance evaluation of the temperature prediction model, the stacking regression model integrating XGBoost and MLP exhibits exceptionally high accuracy on the training set, with an R^2 of 0.981 and an RMSE of 1.3091. On the validation set, the model similarly maintains high accuracy, achieving an R^2 of 0.9689 and an RMSE of 1.6948. Additionally, the mean RMSE obtained from 5-fold cross-validation is 1.917, further confirming the model's stability and generalization capability.

In contrast, the single MLP model, while attaining a respectable R^2 value of 0.9469 on the training set, exhibits a relatively higher RMSE of 2.1523. On the validation set, its R^2 decreases to 0.9558, and the RMSE increases to 2.0184. The mean RMSE from 5-fold cross-validation for the MLP model is 2.381, indicating that the MLP model has comparatively weaker generalization ability on new data. Therefore, the ensemble model not only performs superiorly on the training data but also demonstrates better predictive performance and stability on unseen data.

Table 2. Performance Evaluation of Predictive Models.

Model	Training set		Validation set		Mean CV RMSE
	R^2	RMSE	R^2	RMSE	
XGBoost + MLP	0.981	1.3091	0.9689	1.6948	1.917
MLP	0.9469	2.1523	0.9558	2.0184	2.381

4 Conclusion and Discussion

Using meteorological data from the Jinghe Observation Station provided by the Xi'an Meteorological Information Center for the period 2013–2024, this study integrated multiple meteorological factors, including relative humidity, cloud cover, visibility, precipitation, wind speed, and sunshine duration. Through data preprocessing, feature selection, mutual information methods, and standardization, an efficient temperature prediction model was constructed using a machine learning technique that fuses XGBoost and MLP. The model's performance was comprehensively evaluated using an 8:2 training-test split and 5-fold cross-validation. The main conclusions are as follows:

Data Preprocessing and Feature Selection: Data preprocessing involved handling missing values, encoding wind direction, and selecting key features using the mutual information method, ensuring that the model was trained on high-quality data. Feature selection identified meteorological elements closely related to temperature variations, such as sunshine duration (SD) and average relative humidity (ARH).

Construction and Evaluation of the Temperature Prediction Model: The stacked regression model combining XGBoost and MLP demonstrated excellent performance on the training set, achieving a Root Mean Square Error (RMSE) of 1.34 and a coefficient of determination (R^2) of 0.98. On the test set, the model maintained high accuracy with an RMSE of 1.74 and an R^2 of 0.97.

Validation of Model Stability and Generalization Capability: Through 5-fold cross-validation, the model achieved an average RMSE of 1.92, confirming that the model not only performs well on training data but also possesses strong generalization capability and stability.

Model Optimization and Future Applications: Although the current model exhibits high predictive accuracy, there is potential for further optimization. Future work will focus on adjusting model parameters, exploring additional feature selection methods, and applying the model to predict other meteorological variables.

This study not only provides a novel method for temperature prediction but also experimentally validates its effectiveness and reliability. This approach is expected to advance meteorological science and offer critical decision support for sectors such as agriculture, transportation, and energy. Additionally, it can facilitate the analysis of the spatiotemporal variations in climate quality levels under different meteorological conditions, providing a scientific basis for climate quality assessment and management.

Despite the model's good predictive performance on the current dataset, certain limitations exist. For example, the model's ability to predict extreme meteorological events requires further validation. To address this limitation, future work will explore the incorporation of specialized models designed to capture extreme events, such as extreme value theory models or advanced deep learning architectures with attention mechanisms. Additionally, integrating additional features that are indicative of extreme meteorological conditions, such as extreme precipitation and wind speed indicators, will be considered to enhance the model's capability in predicting these events. Future work will include further optimization of model parameters, exploration of more feature selection and data preprocessing methods, and application of the model to predict other meteorological variables to verify its applicability and generalizability under different meteorological conditions.

References

1. Zhang, Y., He, Q.X., Zeng, S.Y.: Research on global temperature prediction based on CNN-LSTM model. *Adv. Appl. Math.* **13**(1), 302–312 (2024)
2. Wu, W.C., Wei, F.Y., Wang, Y.Q., Zhu, E.D.: Machine learning-based regression analysis of the winter temperature anomalies in China and associated influencing factors. *Acta Meteor. Sin.* **81**(1), 163–174 (2023)
3. Zhao, C.B., Liu, D.X., Xie, X.P.: Research on seasonal temperature forecasting based on time series. *J. Anhui Jianzhu University* **30**(03), 83–89 (2022)
4. Zhou, Y.H., Tang, Q.Y.: Global warming analysis based on ARIMA multiple linear regression model. *Modeling and Simulation* **12**(2), 1257–1271 (2023)
5. Wang, Z.W., Zheng, Z.F., Chen, M.: Prediction of meteorological elements based on nonlinear support vector machine regression method. *J. Applied Meteorological Science* **23**(05), 562–570 (2012)
6. Tao, Y., Du, J.L.: Temperature prediction using long short term memory network based on random forest. *Computer Engineering and Design* **40**(3), 737–743 (2019)
7. Li, D.L., Xiao, Z.X., Xie, N.X.: A study on the adjusting spring and summer surface air temperature of ECMW model by a hybrid feature selection method in machine learning of Guangxi. *J. Chengdu University of Information Technol.* **38**(05), 602–609 (2023)
8. Pan, S.Y., Ding, X., Cui, Y.: Application of MLP neural network to prediction of heating indoor temperature. *Gas & Heat* **39**(07), 40–44 (2019)
9. Peng, D.Y., Zhao, S.L., Wu, Y.Y., Jiang, X.Y.: Global average temperature prediction based on VMD and LSTM. *Climate Change Research Letters* **13**(5), 1055–1063 (2024)
10. Ju, Q., Wu, J.Y., Wang, X.P.: Application of machine learning algorithms in multimodal integration of precipitation and temperature. *Water Resources Protection* **40**(03), 106–115 (2024)
11. Yu, X., Wang, Q.L.: Weather forecast based on ensemble learning. *Information & Computer* **33**(7), 81–83 (2021)
12. Liu, K., Wang, Y.F., Peng, Z.Q.: Machine learning for hyperspectral characteristic bands extraction in soil nutrient analysis. *Acta Optica Sinica* **45**(06), 01 (2024)
13. Lei, B.: An infrared image matching method based on separating information entropy. *Scientific and Technological Innovation* **18**, 89–92 (2024)
14. Tian, H.D., Zhang, M.Z., Chang, R.: A survey on large model training technologies. *ZTE Technology Journal* **30**(02), 21–28 (2024)
15. Zhao, Q., Li, S.S., Guo, Y.M.: Time-varying distributed lag non-linear model: using temperature-mortality association as an example. *J. Environ. Occup. Med.* **37**(1), 9–14 (2020)
16. Li, P.K., Chen, X.M., Lu, W.B.: Research on nonlinear relationship between subway built environment and travel distance of stations based on XGBOOST-SHAP. *Journal of Railway Science and Engineering* **21**(04), 1624–1633 (2024)
17. Zhao, Z.X., Chen, L.M., Yao, L.N.: Prediction of compressive strength of mechanical sand concrete based on BP neural network. *Industrial Control Computer* **37**(08), 57–58+60 (2024)
18. Liu, B., Ding, H.: A diversified stacked regression model for seasonal product demand forecasting. *Logistics Technology* **43**(06), 15–30 (2024)
19. Chen, C., Yu, C.B., Zuo, L.X.: Power system transient stability assessment based on powershap feature selection. *Thermal Power Generation* **53**(08), 143–151 (2024)
20. Liu, S.R., Li, P.Q., Chen, J.Y.: Allocation of hybrid energy storage capacity based on pearson correlation analysis and T-tFD algorithm under VMD decomposition. *Electric Power* **57**(07), 82–97 (2024)
21. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**(3), 379–423 (1948)

22. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley-Interscience (2006)
23. Kuhn, M.: Applied Predictive Modeling. Springer (2013)
24. Zhang, Y.R., Cheng, B.Y., Yang, B.: Quantitative analysis of nitric acid using Raman spectroscopy based on machine learning. *J. Detection & Control* **46**(04), 63–68 (2024)
25. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence. IJCAI, pp. 1137–1143 (1995)



A Review of Research Advances in Image Segmentation of Skin Lesions

Mengmeng He and Xiaoyan Zhang^(✉)

College of Computer Science and Technology, Xi'an 710054, Shannxi, China
1161880978@qq.com

Abstract. Skin cancer is a major global disease whose main threat lies in the risk of rapid spread and remote metastasis, with melanoma being the most lethal type. Automated and accurate segmentation of skin lesions with the help of computer-aided diagnostic systems is clinically significant, not only providing critical information to physicians, but also significantly improving the segmentation accuracy. In order to gain a deeper understanding of the related technology, this study summarizes the latest research progress in this field. Traditional image segmentation algorithms and their advantages and disadvantages are first introduced, followed by a discussion of deep learning-based approaches, focusing on classical structures such as convolutional neural networks (CNN) and U-Net. For the widely used U-Net and its variant models, they are subdivided into segmentation methods based on multi-scale features, contextual information, attention mechanisms, and dual codecs, and the technical improvements of each type of methods are analyzed. Finally, the publicly available datasets of skin lesion images are sorted out and the experimental results of some typical algorithms are compared, and the technical challenges and future development directions are discussed in light of the current situation.

Keywords: Convolutional Neural Network · Deep Learning · U-Net network · Skin Lesion Segmentation

1 Introduction

The occurrence of skin lesions is a common clinical medical problem, and its accurate and efficient recognition and segmentation is crucial for medical diagnosis and treatment. According to authorities such as the World Health Organization (WHO) and the American Cancer Society (ACS) [1, 2], melanoma is the deadliest type of skin cancer. More than 100,000 new cases of melanoma are expected in the United States in 2023, and about 7,000 deaths due to melanoma.

The key objective of skin lesion segmentation is to accurately identify and segment different types of lesions in skin images. Accurate segmentation of skin lesion regions is challenging due to the unique hairs, blood vessels, discoloration, and other interfering factors on the skin surface. Therefore, researchers have used various advanced image segmentation methods to make the lesion areas in the images clearer and clearer, which

helps doctors locate and identify the lesion areas more accurately and improves the credibility of clinical diagnosis. As shown in Fig. 1, some typical datasets with lesions occurring in the skin and examples of segmentation effects are demonstrated.

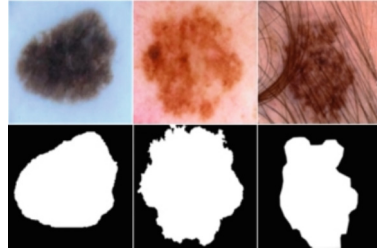


Fig. 1. Skin lesion image and example of segmentation effect

In recent years, deep learning has been widely used in the field of natural images, and its advantage lies in the fact that it can extract more representative feature information in images by increasing the number of network layers. Among them, U-Net and its various improved versions have become a hotspot of research, which draws on the idea of full convolutional network and introduces jump connections between the encoder and decoder, which helps to better preserve the spatial information of the image [3]. In addition, some novel networks based on attention mechanism and multi-scale feature fusion have been proposed, which are able to learn more robust feature representations from a large amount of image data through end-to-end learning, significantly improving the accuracy and efficiency of skin damage segmentation [4].

In this paper, we summarize the network models for skin lesion image segmentation based on deep learning in recent years, aiming to provide a deeper learning perspective and comprehensive understanding for further research. Through extensive research on literature, further deep learning-based approaches in the field of skin lesion segmentation are further subdivided into segmentation models based on multi-scale features, contextual information, attention mechanisms, and dual codecs based on lesion types, regional features, and preprocessing needs. In addition, the commonly used datasets and evaluation metrics in skin lesion segmentation are described in detail, and finally the experimental results of the network models of different methods on commonly used datasets. It is hoped that it can provide a more comprehensive reference and guidance for future research on skin lesion segmentation.

2 Traditional Skin Lesion Segmentation Based Methods

2.1 Clustering-Based Segmentation Methods

Clustering-based segmentation algorithm is an unsupervised learning algorithm whose basic idea is to consider the pixel points in an image as samples and classify them into different categories using clustering algorithms [5]. Commonly used clustering algorithms include K-Means, Mean Shift clustering algorithm, Spectral Clustering (Spectral

Clustering) and so on. Literature [6] embeds constraints into the feature space based on K-Means algorithm, groups image pixels with the same attributes, and merges the pixel sets into several super regions. This method is suitable for segmenting skin lesion regions with distinct color and texture properties. Literature [7] introduced a density-based boundary-driven framework to recognize skin lesions in dermoscopic images. The algorithm extends the regions only at the boundaries of the clusters, thus speeding up the processing without loss of precision or recall.

2.2 Region-Based Segmentation Method

The core idea of the region-based segmentation method is to aggregate pixels into different regions based on local features such as gray scale, color, texture, etc., and output these regions as segmentation results. The method makes full use of the local feature information of the image to produce segmented regions with internal consistency. Literature [8] used statistical region fusion method (SRM) for pre-segmentation of dermatoscopically acquired medical images of skin cancer and further segmentation by conversion to HSV color space and K-mean algorithm to improve the segmentation boundaries of foreground and background. Literature [9] showed that the region growing method performed well in extracting edge information and segmenting lesions individually from CT images, which is suitable for clinical CT images that require fast and accurate segmentation, and improves the efficiency of the procedure.

2.3 Threshold-Based Segmentation Method

The basic idea of threshold-based segmentation method is to classify pixels into two or more categories by comparing the gray value or color value of the pixels in the image with a pre-set threshold, which is computationally simple, fast running and easy to implement. Literature [10] utilizes an adaptive thresholding technique to segment the skin lesion region directly and uses a skew estimation technique to correct the asymmetry of the histogram to further improve the segmentation results. This method is suitable for skin lesion regions with asymmetric histogram distribution. Literature [11] used four different thresholding methods for skin lesions to perform automatic edge detection, which relies more on the selection of color histograms, however, the color of the lesion region of the image has a large difference, and there is a certain degree of difficulty in determining the threshold value. Therefore, this method is suitable for skin lesion regions with obvious color histogram distribution.

2.4 Edge Based Segmentation Methods

The basic idea of edge based segmentation methods is to recognize the edges between the object and the background in an image by edge detection. Commonly used edge detection algorithms include Canny and Sobel. Their main advantage is that they can process the details and edge information in the image more finely, but they are also easily affected by noise and image quality. Literature [12] used Canny edge detector to detect the edges of the skin lesion region and used median filtering to obtain segmentation

results after several iterations of processing. This method solves the problem of blurred melanoma edges by refining the edges through multiple iterations. Literature [13] uses Sobel convolution kernel instead of Gaussian convolution kernel to compute the image gradient size, which improves the anti-noise performance while simplifying the Canny edge detection process. Literature [14] increased to four Sobel operators in the gradient computation to improve the accuracy of gradient size computation and make the Canny edge detection with better anti-noise performance.

3 Deep Learning Based Skin Lesion Image Segmentation Method

Although traditional skin lesion image segmentation methods based on traditional methods have achieved certain results, due to the fact that the foreground and background of medical images are usually more complex, some of the accuracy is often lost during the segmentation process, resulting in low segmentation accuracy. In recent years, with the rapid development of deep learning technology, its segmentation accuracy has been significantly improved compared to traditional methods. Deep learning algorithms are able to learn from a large amount of training data and effectively extract image features to achieve higher accuracy in image segmentation tasks. These algorithms help to improve the accuracy, reliability and efficiency of image segmentation. By analyzing the models related to medical image segmentation based on deep learning in recent years, this chapter will provide a comprehensive summary in terms of both segmentation models based on convolutional neural networks and segmentation models based on U-Net and its variant structures.

3.1 Segmentation Model Based on Convolutional Neural Network

Convolutional Neural Network (CNN) is a common deep learning network [15], whose artificial neurons can respond to a portion of the surrounding units in the coverage area, and it performs well in processing large images. CNN acquires the spatial feature relationships of the image through the convolutional kernel, and uses the back propagation algorithm to adjust the convolutional kernel parameters, and ultimately generates a model that can effectively acquire useful information within the image. Typical deep learning neural network models include LeNet network [16], AlexNet network [17], VGG network [18], Inception network [19] and ResNet network [20].

Segmentation Model Based on FCN Structure

Nowadays, many deep learning algorithms still use the idea of CNN [21]. Fully Convolutional Network (FCN) [22] also inherits the idea of CNN, replaces the fully connected layer with a convolutional layer, and uses an inverse convolutional layer to up-sample the feature maps of the last convolutional layer, so that it can process an input image of arbitrary size and output pixel-level segmentation predictions. The network structure of FCN is shown in Fig. 2 is shown.

Literature [23] proposed a fully convolutional network with a dense pooling layer, which can ensure that as little image information as possible is lost during the forward propagation process, thus better capturing the detailed features of melanoma. Literature

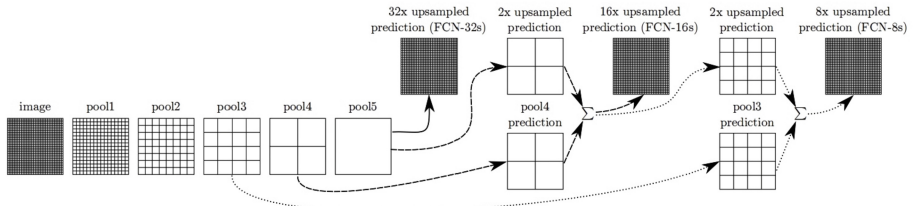


Fig. 2. FCN network structure

[24] used the strategy of hyperpixel method to refine the output results of FCN, the processing results are relatively fine edges in the case of ideal FCN output, but the effect is not good in the case of complex image structure.

Segmentation Model Based on RNN Structure

Recurrent Neural Network (RNN) is a deep learning network model [25], which is different from CNN in that it has a recurrent unit for processing sequence data, retaining the information from the previous moment and using it as the current input, thus being able to process sequence data of arbitrary length. However, due to the parameter sharing of the recurrent unit, RNNs are prone to the problems of gradient vanishing, gradient explosion and long-term dependence, while Long Short-term Memory (LSTM) [26] combines short-term and long-term memory through forgetting gates, which ameliorates these problems to some extent. Literature [27] proposes a new deep learning framework, FCSLSTM, that combines FCN for spatial models and LSTM for temporal models. This is the first DL framework that deeply integrates FCN and LSTM to achieve joint 4D segmentation that can be trained end-to-end.

Segmentation Model Based on GAN Structure

Generative Adversarial Network (GAN) is a deep learning model [28], similar to the network model of a two-player game, which mainly consists of a combination of a generative network (G) and a discriminative network (D). The structure of the GAN network is shown in Fig. 3. Literature [29] proposes a novel framework based on adversarial learning called Efficient-GAN (EGAN) that utilizes unsupervised generative networks to generate accurate lesion masks. The composite scaling path generator module of this framework utilizes a bottom-up squeezing excitation mechanism to better handle the multi-scale features of the skin lesion region. The top-down asymmetric laterally connected path and discriminator modules, on the other hand, can effectively differentiate between original and synthetic masks. Literature [30] proposed a deep convolutional neural network (CNN) trained using a structurally conditioned GAN network, a virtual staining framework that can rapidly convert in vivo RCM images into 3D microscopic images with H&E-like color contrast for the diagnosis of normal skin, basal cell carcinomas, and pigmented melanocytic nevi.

3.2 Segmentation Model Based on U-Net and Its Variant Architectures

U-Net [31] is a deep learning architecture for image segmentation, originally proposed by Ronneberger et al. in 2015. The model has shown great advantages in the field of

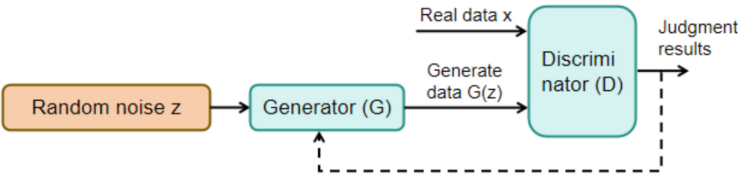


Fig. 3. GAN network structure

medical image segmentation and has become one of the most widely used models. Due to the better segmentation effect of this framework, a large number of subsequent researchers have improved on U-Net network many variant network models based on U-Net have appeared and shown powerful segmentation effect. Such as AttentionU-Net [32], U-Net++ [33], U-Net3 + [34], nnUNet [35], R2U-Net [36] and TransUNet [37]. This paper categorizes and summarizes the representative algorithms in recent years, which will be followed by a detailed description of the four aspects of the segmentation model, namely, based on multi-scale features, based on contextual information, based on the attention mechanism, and based on dual codecs.

Segmentation Model Based on Multi-scale Features

Multi-scale features are image features extracted at different scales, which can capture different levels of information from global to local and improve the model performance. Literature [38] proposed a new multi-scale residual encoding and decoding network (Ms RED) for skin lesion segmentation and the network structure is shown in Fig. 4. A multiscale residual coding fusion module(MsR-EFM) is used in the encoder, and a multiscale residual decoding fusion module(MsR-DFM) is applied in the decoder in order to adaptively fuse multiscale features. By proposing a multi-resolution, multi-channel feature fusion module(M2F2) to replace the traditional encoder and decoder networks with convolutional layers, better segmentation performance is obtained.

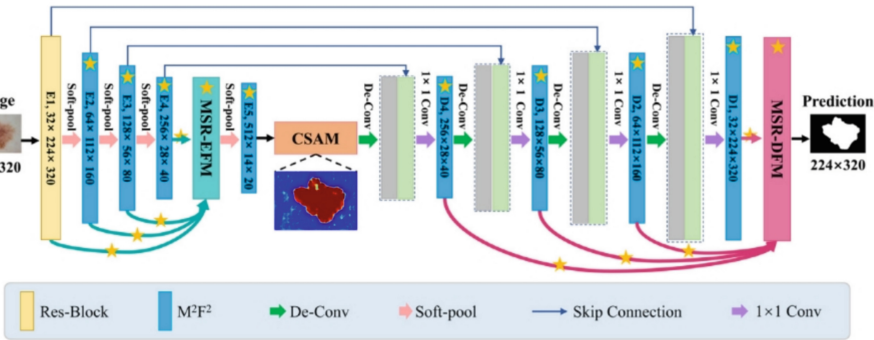


Fig. 4. Ms RED Network structure model

Literature [39] introduced a multiscale feature fusion network (MSF-Net) based on integrated attention convolutional neural network, which introduced a multiscale dilation

convolution module (MDC) and a multiscale feature fusion module (MFF) in the convolutional block to extract cross-scale contextual information and adaptively adjust the sensory field size of the feature maps, which helps to improve the segmentation accuracy at the edges of the lesion region. Literature [40] proposed a multi-scale convolutional neural network REDAUNet based on UNet3+. Its encoder consists of four feature extraction layers through two cross residual (CR) units, allowing the module to extract deep semantic information while avoiding the gradient vanishing problem. Literature [41] proposed an adaptive position-aware skin lesion segmentation model, SapFormer, which is suitable for segmentation of complex skin lesion regions by designing multiple hybrid transformers for multiscale feature encoding and using a transformer decoder for multiscale positional feature sensing of the encoded features to obtain fine-grained features of the skin lesion region and optimize the regional feature distribution.

Segmentation Model Based on Contextual Information

Literature [42] proposes a multi-scale context-guided network called MSCGnet, which enhances the model's ability to deal with complex boundaries by utilizing contextual information to design a context-based attention structure (Cas) during feature encoding, and selecting effective contextual features in the decoding path. To capture local and global contextual information, literature [43] proposed EIU-Net network as shown in Fig. 5, EIU-Net uses inverted residual block and efficient pyramid squeezed attention (EPSA) block as the main encoder at different stages and downsampling after the last encoder using irregular space pyramid pooling (ASPP) and soft pooling methods. The ability of the network to segment skin lesion boundaries and details is improved by effectively fusing the feature distributions in different encoders through the multilayer fusion (MLF) module.

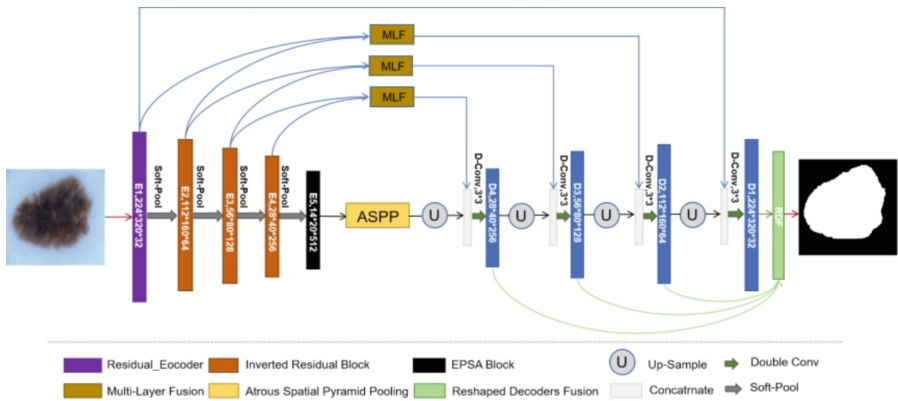


Fig. 5. EIU-Net Network structure model

Literature [44] proposed UNeXt, which uses the U-Net coder-decoder structure and the design of hopping connections. For skin lesions that need to be diagnosed by combining multiple imaging information, UNeXt utilizes the complementary information of different modalities to improve the segmentation accuracy, while possessing a strong

generalization ability. Literature [45] proposes a novel Neighborhood Context Refinement Network (NCRNet) containing a shared encoder and two different but closely related decoders, for skin lesion regions with complex boundaries, the Parallel Attention Decoder (PAD) effectively extracts and fuses multi-level local detail information and global semantic information to localize skin lesions of different sizes and shapes, improving segmentation accuracy.

Segmentation Model Based on Attention Mechanism

People will naturally choose to focus on the more important information or regions when observing things, inspired by this phenomenon, Attention Mechanisms have been introduced into neural network models, and the efficiency and accuracy of image segmentation can be greatly improved by introducing Attention Mechanisms into deep learning segmentation models. Literature [46] proposed a Gated Fusion Attention Network (GFANet) to accurately segment skin lesion images by designing two asymptotic relational decoders. Multi-level contextual features are fused using a contextual feature gated fusion decoder (CGFD), and the final output of the encoder's low-level features and shape streams are fused by GCF. The boundary segmentation accuracy of the lesion region is effectively improved to avoid over- or under-segmentation.

Literature [47] proposed a multiscale attention network (MSAU-Net) for skin lesion segmentation. By improving the typical U-Net, an attention mechanism is inserted at the bottleneck, which combines the BDC-LSTM structure to effectively capture the complex features of the lesion region and improve the segmentation accuracy. Literature [48] proposes a deep network for accurate skin lesion segmentation based on a U-shaped structure. The ACCAPA module can model the features of the lesion region by dynamically learning the channel information, contextual information, and global structural information, and GFF is used for different levels of semantic information interactions between the encoder and decoder layers. The segmentation efficiency is improved.

Segmentation Model Based on Dual Codecs

A novel network based on encoder-decoder architecture, Dual Encoder U-Net (DEU-Net), has been proposed in the literature [49]. For skin lesions with complex morphology and diverse textures, DEU-Net is able to effectively capture local details and global contextual information through a dual encoder branch consisting of a convolutional encoder and a transformer encoder to improve the accuracy and reliability of segmentation. Literature [50] proposed a dual branch feature extraction network (DFE-Net). Two encoders based on Transformer and Efficient Channel Attention (ECA) modules are used, one for extracting global features and the other for extracting local features. Meanwhile, the jump connection from the ECA feature extraction module to the decoder is designed to reduce the feature loss during decoding, which effectively recovers the lesion boundaries and details and improves the segmentation accuracy. Literature [51] proposed a new skin lesion segmentation network DBCGN. In the feature extraction stage, three-layer high-level semantic features using Transformer branches and CNN branches are used to construct cross-scale global contextual dependencies for the semantic features. For skin lesions with fuzzy boundaries and complex morphology, DBCGN effectively outlines

the boundaries of the target region through multi-scale feature fusion and inverse map inference module to improve the segmentation accuracy and reliability.

With the above classification methods, different variant models can focus on solving different segmentation difficulties. Multi-scale feature-based methods can effectively capture the global context and local details of the skin lesion region, thus enhancing the ability to recognize lesions at different scales. For example, by introducing pyramid structure or cavity convolution, it is possible to locate tiny lesions more accurately while retaining the overall morphological information to enhance the segmentation accuracy. Segmentation methods based on contextual information improve the robustness of the segmentation model to complex backgrounds by capturing the relationship between the lesion region and the surrounding tissues, e.g., by using RNNs or CRFs to enhance the contextual dependency. Segmentation offenses based on attention mechanisms can accurately localize lesion boundaries and improve the focusing ability of the model by dynamically assigning weights, highlighting important features of the lesion region, and suppressing irrelevant background interference. For example, the introduction of channel attention and spatial attention helps to solve the segmentation challenges of low contrast and boundary blurring. The segmentation method based on the dual codec structure solves the trade-off between detail preservation and global perceptual ability of a single codec model by simultaneously processing the detailed features and global semantic information of the lesion region. The method significantly improves the segmentation performance of complex morphological lesions, especially for irregular or multiple lesions.

4 Discussion and Analysis

4.1 Datasets

Datasets are indispensable in the task of skin lesion segmentation, not only as a basis for training, validation, and evaluation of deep learning algorithms, but also as a rich source of examples for researchers. By using the same dataset, researchers can fairly compare the performance of various algorithms under similar conditions to better process skin lesions and evaluate the performance of different algorithms.

Common skin lesion datasets include the ISIC2016 dataset [52], which contains 900 skin lesion images divided into 600 training set images and 300 test set images, which are segmented and labeled according to lesion type and image source. The ISIC2017 dataset [53] contains 2,000 training images, 150 validation images and 600 test images, all of which have corresponding Ground Truth (GT). ISIC2018 dataset [54] contains 2594 images of skin lesions, where the GT data of the mask images are generated by multiple techniques and reviewed and curated by specialized dermatologists. ISIC2019 dataset [55] covers 8 different types of skin lesions, with a size of about 13 GB, with image sizes ranging from 256x256 pixels to 6000x4500 pixels. The ISIC2020 dataset [56] includes 33,126 images of benign and malignant skin lesions from 2056 patients. Each image has a unique patient identifier, which contributes to a more accurate melanoma diagnosis and false-positive exclusion. The PH2 [57] dataset was provided by the Department of Dermatology, Pedro Hispano Hospital, Matosinhos, Portugal. The dataset contains 200 dermoscopic images and corresponding labeled images, all stored in PNG format. The

HAM10000 [58] dataset contains 10015 training images for the detection of pigmented skin lesions. This benchmark dataset can be used for machine learning diagnostics for comparison with human experts and covers a representative set of all important diagnostic categories in the field of pigmented lesions. The above mentioned datasets commonly used for image segmentation of skin lesions are summarized as shown in Table 1.

Table 1. Skin injury image dataset

Database	Amount	Resolution	Date
PH2	200	560*768	2013
ISIC2016	900	540*720	2016
ISIC2017	2000	540*772; 2000*3008; 4499*6748	2017
ISIC2018	2594	540*720	2018
ISIC2019	25331	540*720	2019
ISIC2020	33126	480*640; 5184*3456; 6000*6000	2020
HAM10000	10015	450*600	2018

4.2 Analysis of Typical Algorithm Results

Table 2 shows the evaluation metrics results of different approaches in deep learning based on U-Net and its variant structures on ISIC2017 dataset, ISIC2018 dataset and PH2 dataset, where the experimental results are taken from the related literature.

From the comparative experimental results in Table 2, it can be seen that on the ISIC2017 dataset, the multiscale feature-based segmentation method SapFormer [41] obtains high Iou and Dice values, and by designing multiple hybrid transformers to encode multiscale features in skin images, it obtains fine-grained features in skin lesion regions, reduces erroneous segmentation of non-skin lesion regions, and improves the efficiency of skin image segmentation by reducing the error segmentation of non-skin lesion regions. Literature [42] consists of four feature extraction layers through two cross residual (CR) units, which can recognize lesions of different sizes by integrating dense connectivity and ASPP as well as multiscale information fusion, thus obtaining better semantic consistency. Literature [48] designed a boundary-aware contextual neural network that can acquire multi-granularity edge information, and proposed a new interactive attention to bridge the complementary information between different tasks, thus effectively utilizing the boundary information to provide strong cues for better segmentation prediction, resulting in better edge and semantic integrity of segmentation results.

On the ISIC2018 dataset, the highest Iou values are obtained by the dual codec-based segmentation method DFE-Net [50], which extracts the information with two Transformer and Efficient Channel Attention (ECA) module-based coders to maximize the recovery of boundary and local features. On the PH2 dataset, both literature [49] and

Table 2. Experimental results based on U-Net and its variants of structural modeling methods

Methods	Model	ISIC2017		ISIC2018		PH ²	
		Dice	Iou	Dice	Iou	Dice	Iou
Multi-scale feature-based	Ms RED 错误!未找到引用源。	86.48	78.55	89.99	83.45	94.65	90.14
	MSF-Net 错误!未找到引用源。	-	-	93.89	88.75	-	-
	REDAUNet 错误!未找到引用源。	90.29	82.98	90.20	83.23	-	-
	SapFormer 错误!未找到引用源。	92.60	86.40	94.30	89.40	-	-
Contextual information	MSCGnet 错误!未找到引用源。	-	-	-	-	93.36	88.21
	EIU-Net 错误!未找到引用源。	85.50	77.10	90.20	83.60	91.60	84.70
	UNeXt 错误!未找到引用源。	-	-	-	81.70	-	-
	NCRNet 错误!未找到引用源。	86.55	78.62	-	-	-	-
Attention mechanism	GFANet 错误!未找到引用源。						
	ULFAC-Net 错误!未找到引用源。	86.33	78.10	90.81	83.52	-	-
Dual encoder-decoder	DEU-Net 错误!未找到引用源。	87.16	79.51	90.81	84.44	95.65	91.81
	DFU-Net 错误!未找到引用源。	-	-	-	89.68	-	93.79

literature [50] obtained high Iou values, but at the same time, the two-branch structure makes the model more complicated and increases the number of parameters. Ms RED [38] proposed a new multiscale residual coding and decoding network (Ms RED), which adopts the multiscale residual coding fusion module (MsR-EFM) in the encoder, and applies the multiscale residual decoding fusion module (MsR-DFM) in the decoder in order to adaptively fuse the multiscale features, and obtains a better segmentation performance.

5 Summary and Outlook

This paper briefly describes the significance of medical image segmentation in the context of dermatology, and firstly introduces the traditional segmentation methods based on clustering, region-based, thresholding and edge detection, and summarizes their advantages and disadvantages. Secondly, image segmentation methods based on deep learning are summarized, which are divided into segmentation models based on convolutional neural network and segmentation models based on U-Net and its variants. Finally, the

experimental results of typical algorithms on commonly used datasets are compared and analyzed. After finishing and analyzing, it is found that there are still some challenges in this field, and the future research direction can be carried out from these directions.

- (1) Insufficient data diversity. Many existing datasets cover a limited variety of skin lesions, while in reality there are a wide variety of lesions. Techniques such as generative adversarial networks can be utilized to synthesize more diverse skin lesion images to expand the dataset and increase the adaptability of the model to various lesions.
- (2) Labeling difficulties. Accurate labeling of large-scale skin lesion images is a complex task that requires the involvement of a medical professional. By using weakly supervised learning or semi-supervised learning, combined with techniques such as unsupervised learning and reinforcement learning, labels are automatically generated from unlabeled data to reduce the burden of manual labeling.
- (3) Blurred and mixed boundaries. The boundaries of skin lesions are usually blurred, which brings challenges to the accuracy and stability of segmentation algorithms. By introducing operations such as attention mechanism and morphology, the model can pay more attention to the boundary information of the lesion region in order to improve the model's perception of boundary information and segmentation effect.

In actual diagnosis, the use of automated segmentation technology for remote skin lesion screening can quickly give segmentation results, speeding up the diagnostic process and helping users to seek medical treatment in a timely manner.

References

1. World Health Organization. Skin Cancers. <https://www.who.int/news-room/fact-sheets/detail/skin-cancers>
2. American Cancer Society. Cancer Facts & Statistics. <https://www.cancer.org/research/cancer-facts-statistics.html>
3. Al-Masni, M.A., Al-Antari, M.A., Choi, M.T., et al.: Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine* **162**, 221–231 (2018)
4. Chen, P., Huang, S., Yue, Q.: Skin lesion segmentation using recurrent attentional convolutional networks. *IEEE Access* **10**, 94007–94018 (2022)
5. Wu, Q.P., Wu, C.M.: Fast robust kernel-based fuzzy C-means clustering segmentation. *J. Image and Graphics* **23**(12), 1838–1851 (2018)
6. Zhou, H., Chen, M., Zou, L., et al.: Spatially constrained segmentation of dermoscopy images. In: *Proceedings of the IEEE International Symposium on Biomedical Imaging: from Nano to Macro*. IEEE, pp. 800–803 (2008)
7. Mete, M., Kockara, S., Aydin, K.: Fast density-based lesion detection in dermoscopy images. *Comput. Med. Imaging Graph.* **35**(2), 128–136 (2011)
8. Tian, Q.F., Meng, R.S., Jiang, Z.G., et al.: Research on a comprehensive segmentation method for skin melanocyte tumor images. *Chinese Stereology and Image Analysis* **12**(3), 207–211 (2007)
9. Shu, T.H., Xu, B.G., Hu, Z.H.: Medical image segmentation based on region growing method. *Microcomputer Information* **18**, 284–285 (2008)

10. Sforza, G., Castellano, G., Arika, S.K., et al.: Using adaptive thresholding and skewness correction to detect gray areas in melanoma in situ images. *IEEE Trans. Instrum. Meas.* **61**(7), 1839–1847 (2012)
11. Celebi, M.E., Iyatomi, H., Schaefer, G., et al.: Border detection in dermoscopy images using statistical region merging. *Skin Research & Technology* **14**(3), 347–353 (2008)
12. Liu, Z., Zerubia, J.: Skin image illumination modeling and chromophore identification for melanoma diagnosis. *Phys. Med. Biol.* **60**(9), 3415–3431 (2015)
13. Wang, Y.X., Chen, J.M.: Iris edge detection algorithm based on adaptive Canny operator and multi-directional Sobel operator. *Computer and Digital Engineering* **48**(11), 2744–2749 (2020)
14. Dong, E., Zhao, Y., Yu, X., et al.: An improved NMS-based adaptive edge detection method and its FPGA implementation. *Journal of Sensors* **8**(6), 31–39 (2016)
15. Bouvrie, J.: Notes on Convolutional Neural Networks (2006)
16. Lecun, Y., Bottou, L., Bengio, Y., et al.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
17. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Adv. Neural. Inf. Process. Syst.* **25**(2), 1097–1105 (2012)
18. Simonvan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *Computer Science*, pp. 1–13 (2014)
19. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9 (2015)
20. He, K., Zhang, X., Ren, S., et al.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778 (2016)
21. Patel, S., Patel, R., Ganatra, N., et al.: Spatial feature fusion for biomedical image classification based on ensemble deep CNN and transfer learning. *International Journal of Advanced Computer Science and Applications* **13**(5) (2022)
22. Long, J., Shelhamer, E., Darrell, T., et al.: Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. USA: IEEE Press*, pp. 3431–3440 (2015)
23. Nasresfahani, E., Rafiei, S., Jafari, M.H., et al.: Dense fully convolutional network for skin lesion segmentation. *Computerized Medical Imaging and Graphics* **78** (2017)
24. Bozorgabar, B., Sedai, S., Roy, P.K., et al.: Skin lesion segmentation using deep convolution networks guided by local unsupervised learning. *Ibm Journal of Research & Development* **61**(4), 6:1–6:8 (2017)
25. Zaremba, W., Sutskever, I., Vinyals, O.: Recurrent Neural Network Regularization. *arXiv preprint [arXiv:1409.2329](https://arxiv.org/abs/1409.2329)* (2014)
26. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Computation* **9**(8), 1735–1780 (1997)
27. Gao, Y., Phillips, J.M., Zheng, Y., et al.: Fully convolutional structured LSTM networks for joint 4D medical image segmentation. In: *Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging. IEEE*, pp. 1104–1108 (2018)
28. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al.: Generative adversarial nets. *Advances in Neural Information Processing Systems*, pp. 1–9 (2014)
29. Innani, S., Dutande, P., Baid, U., et al.: Generative adversarial networks based skin lesion segmentation. *Sci. Rep.* **13**(1), 13467 (2023)
30. Li, J., Garfinkel, J., Zhang, X., et al.: Biopsy-free in vivo virtual histology of skin using deep learning. *Light: Science & Applications* **10**(1), 233 (2021)
31. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Germany, October 5–9*, pp. 234–241 (2015)

32. Oktay, O., Schlemper, J., Folgoc, L.L., et al.: Attention u-net: Learning where to Look for the Pancreas. arXiv preprint [arXiv:1804.03999](https://arxiv.org/abs/1804.03999) (2018)
33. Zhou, Z., Siddiquee, M.R., Tajbakhsh, N., et al.: Unet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Trans. Med. Imaging* **39**(6), 1856–1867 (2019)
34. Huang, H., Lin, L., Tong, R., et al.: UNet 3+: a full-scale connected unet for medical image segmentation. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1055–1059 (2020)
35. Isensee, F., Jaeger, P.F., Kohl, S.A., et al.: NnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
36. Alom, M.Z., Hasan, M., Yakopcic, C., et al.: Recurrent Residual Convolutional Neural Network based on u-net (r2u-net) for Medical Image Segmentation. arXiv preprint [arXiv:1802.06955](https://arxiv.org/abs/1802.06955) (2018)
37. Chen, J., Lu, Y., Yu, Q., et al.: Transunet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv preprint [arXiv:2102.04306](https://arxiv.org/abs/2102.04306) (2021)
38. Duwei, D., Caixia, D., Songhua, X., et al.: Ms RED: a novel multi-scale residual encoding and decoding network for skin lesion segmentation. *Med. Image Anal.* **75**, 102293 (2021)
39. Shao, D., Ren, L., Ma, L.: MSF-Net: a lightweight multi-scale feature fusion network for skin lesion segmentation. *Biomedicines* **11**(6), 1733 (2023)
40. Liu, L., Zhang, X., Li, Y., et al.: An improved multi-scale feature fusion for skin lesion segmentation. *Appl. Sci.* **13**(14), 8512 (2023)
41. Xin, C., Liu, Z., Ma, Y., et al.: Transformer guided self-adaptive network for multi-scale skin lesion image segmentation. *Comput. Biol. Med.* **169**, 107846 (2024)
42. Tang, Y., Fang, Z., Yuan, S., et al.: Imsegnet: Iterative multi-scale context-guided segmentation of skin lesion in dermoscopic images. *IEEE Access* **8**, 39700–39712 (2020)
43. Yu, Z., Yu, L., Zheng, W., et al.: EIU-Net: Enhanced feature extraction and improved skip connections in U-Net for skin lesion segmentation. *Computers in Biology and Medicine*, pp. 107081 (2023)
44. Valanarasu, J.M.J., Patel, V.M.: Unext: Mlp-based rapid medical image segmentation network. In: *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer Nature Switzerland, pp. 23–33 (2022)
45. Liu, Q., Wang, J., Zuo, M., et al.: NCRNet: neighborhood context refinement network for skin lesion segmentation. *Comput. Biol. Med.* **146**, 105545 (2022)
46. Qiu, S., Li, C., Feng, Y., et al.: GFANet: gated fusion attention network for skin lesion segmentation. *Comput. Biol. Med.* **155**, 106462 (2023)
47. Alahmadi, M.D.: Multiscale attention U-Net for skin lesion segmentation. *IEEE Access* **10**, 59145–59154 (2022)
48. Ma, Y., Wu, L., Gao, Y., et al.: ULFAC-Net: ultra-lightweight fully asymmetric convolutional network for skin lesion segmentation. *IEEE J. Biomed. Health Inform.* **27**(6), 2886–2897 (2023)
49. Karimi, A., Faez, K., Nazari, S.: DEU-Net: dual-encoder u-net for automated skin lesion segmentation. *IEEE Access* **11**, 134804–134821 (2023)
50. Fan, C., Yang, L., Lin, H., et al.: DFE-Net: dual-branch feature extraction network for Enhanced segmentation in skin lesion. *Biomed. Signal Process. Control* **81**, 104423 (2023)
51. Song, P., Li, J., Fan, H., et al.: DBCGN: dual branch cascade graph network for skin lesion segmentation. *Int. J. Mach. Learn. Cybern.* **14**(8), 2847–2865 (2023)
52. Gutman, D., Codella, N.C.F., Celebi, E., et al.: Skin Lesion Analysis Toward Melanoma Detection: a Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC). arXiv preprint [arXiv:1605.01397](https://arxiv.org/abs/1605.01397) (2016)

53. Codella, N.C.F., Gutman, D., Celebi, M.E., et al.: Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: Proceedings of the 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018). IEEE, pp. 168–172 (2018)
54. Codella, N., Rotemberg, V., Tschandl, P., et al.: Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (isic). arXiv preprint arXiv: 1902.03368 (2019)
55. Combalia, M., Codella, N.C.F., Rotemberg, V., et al.: Bcn20000: Dermoscopic Lesions in the Wild. arXiv preprint [arXiv:1908.02288](https://arxiv.org/abs/1908.02288) (2019)
56. International Skin Imaging Collaboration. Siim-isic 2020 challenge dataset. International Skin Imaging Collaboration (2020)
57. Mendonca, T., Ferreira, P.M., Marques, J.S., et al.: PH2-A dermoscopic image database for research and benchmarking. In: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, pp. 5437–5440 (2013)
58. Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Scientific data **5**(1), 1–9 (2018)

Author Index

B

Bao, Lin [II-201](#)
Bi, Gang [I-189](#)

C

Cao, Jiuwen [I-201](#)
Chang, Yuhao [I-109](#), [I-119](#)
Chen, Geng [II-93](#), [II-106](#)
Chen, Jiagui [I-201](#)
Chen, Long [I-252](#)
Chen, Zaixin [II-77](#), [II-132](#)
Cheng, Jinwei [I-61](#)
Chu, Jie [I-49](#)
Cui, Mingtao [I-109](#)

D

Ding, Jundi [I-158](#)
Ding, Nan [I-169](#)
Dong, Yifan [I-39](#)
Du, Peike [II-77](#)

F

Fan, Bai [I-212](#)
Fan, Jiaxin [I-189](#)
Fang, Jiajia [I-201](#)
Fang, Jinguo [I-119](#)
Feng, Shuyi [II-3](#)
Feng, Xiao [I-99](#), [I-109](#), [I-119](#)
Fu, Jiajia [I-3](#), [I-132](#)

G

Gan, Lan [II-18](#), [II-147](#)
Gao, Haotian [I-158](#)
Gao, Yuxing [I-252](#)

H

Han, Dezhi [I-179](#)
Han, Hao [II-181](#)
Hao, Jun [II-201](#)
Hao, Zhifei [I-99](#)

He, Dacheng [I-179](#), [II-121](#)
He, Mengmeng [I-265](#)
He, Wenchen [I-39](#), [I-89](#)
Hu, Dinghan [I-201](#)
Huang, Donghai [I-3](#), [I-132](#)
Huang, Jain [II-31](#)
Huang, Lvchao [I-119](#)
Huang, Tianshu [I-169](#)
Huo, Yingzi [II-43](#)

J

Jiang, Dongqi [I-189](#)
Jiang, Man [II-121](#)
Jiao, Ankang [II-181](#)

K

Kaihu, Hou [I-212](#)
Kang, Zhongmiao [I-3](#), [I-132](#)
Ke, Meiguo [II-43](#)

L

Li, Biqing [II-132](#)
Li, Haotian [I-142](#)
Li, Kefei [II-43](#)
Li, Wendi [II-181](#)
Li, Xingnan [I-3](#), [I-132](#)
Li, Yanbo [II-31](#)
Li, Yitao [I-49](#), [I-89](#)
Li, Zi [II-66](#)
Liang, Wenjuan [I-132](#)
Liu, Guangwei [I-99](#), [I-119](#)
Liu, Peng [I-228](#), [I-239](#)
Liu, Wanyu [I-61](#)
Liu, Wenbin [I-109](#)
Liu, Xingtong [I-12](#)
Liu, Yuelin [II-66](#)
Liu, Yuxin [II-43](#)
Lu, Chenxin [I-61](#), [I-76](#), [II-3](#)
Lu, Jiangang [I-3](#)
Luo, Wen [I-142](#), [II-54](#)

Lv, Dong I-39, I-89
Lv, Enqi I-252

M

Mao, Yang II-132
Meng, Xiangwei I-179

N

Ning, Shan I-49
Niu, Kunxiang I-142, II-54

P

Peiyao, Gong I-212
Peng, Yumeng II-201
Pengbiao, Chen I-212

Q

Qin, Jiawei I-89
Qiu, Yuxiang II-18, II-147

S

Shen, Ji I-76
Su, Bo I-189
Sui, Mingshuang I-39
Sun, Denghui I-239
Sun, Xiaoyan II-201
Sun, Yong I-109

W

Wan, Ming II-147
Wang, Fei II-161
Wang, Haiyang I-39, I-49
Wang, Kaiqiao I-239
Wang, Liqiang I-39, I-49, I-89
Wang, Weiming II-106
Wang, Yong I-23
Wang, Yu II-201
Wang, Zihan I-89
Wu, Kaijun II-77

Wu, Yanbing II-43
Wu, Zhenhong I-189

X

Xiao, Lijun I-179
Xu, Ke II-121
Xue, Wenhua I-252

Y

Yan, Hanming I-99
Yan, Kai I-189
Yan, Tihua II-77, II-132
Yang, Hang II-201
Yao, Longfei I-142
Yu, Ji I-158
Yuan, Jie I-76, II-3

Z

Zeng, Qingtian II-93, II-106
Zhang, Chun I-239
Zhang, Hang I-99
Zhang, Hongbin II-66
Zhang, Lei II-18
Zhang, Peiyao I-109
Zhang, Ronglin II-43
Zhang, Xiaoyan I-265
Zhang, Yiming II-93
Zhang, Zhen II-3
Zhang, Zhezhaoyao I-201
Zhao, Dongmei II-66
Zhao, Jingjing I-12
Zhao, Liutao I-169
Zheng, WenFa I-23
Zheng, Yang I-61, I-76, II-3
Zheng, Zhongxiang I-228
Zhou, Qirong I-49
Zhou, Shuhui I-61, I-76
Zhu, Honglong II-181
Zhu, Yuan II-18, II-147