

Huaping Liu
Di Guo *Editors*

Proceedings of 2025 Chinese Intelligent Automation Conference

Volume III

Series Editors

Leopoldo Angrisani, *Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Napoli, Italy*

Marco Arteaga, *Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico*

Samarjit Chakraborty, *Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany*
Shanben Chen, *School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China*

Tan Kay Chen, *Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore*

Rüdiger Dillmann, *University of Karlsruhe (TH) IAIM, Karlsruhe, Germany*

Haibin Duan, *Beijing University of Aeronautics and Astronautics, Beijing, China*

Gianluigi Ferrari, *Dipartimento di Ingegneria dell'Informazione, Sede Scientifica Università degli Studi di Parma, Parma, Italy*

Manuel Ferre, *Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain*

Faryar Jabbari, *Department of Mechanical and Aerospace Engineering, University of California, Irvine, USA*

Limin Jia, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*

Janusz Kacprzyk, *Intelligent Systems Laboratory, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland*

Alaa Khamis, *Department of Mechatronics Engineering, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt*

Torsten Kroeger, *Intrinsic Innovation, Mountain View, USA*

Yong Li, *College of Electrical and Information Engineering, Hunan University, Changsha, China*

Qilian Liang, *Department of Electrical Engineering, University of Texas at Arlington, Arlington, USA*

Ferran Martín, *Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Spain*

Tan Cher Ming, *College of Engineering, Nanyang Technological University, Singapore, Singapore*

Wolfgang Minker, *Institute of Information Technology, University of Ulm, Ulm, Germany*

Pradeep Misra, *Department of Electrical Engineering, Wright State University, Dayton, USA*

Subhas Mukhopadhyay, *School of Engineering, Macquarie University, Sydney, NSW, Australia*

Cun-Zheng Ning, *Department of Electrical Engineering, Arizona State University, Tempe, AZ, USA*

Toyoaki Nishida, *Department of Intelligence Science and Technology, Kyoto University, Kyoto, Japan*

Luca Oneto, *Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genova, Genova, Italy*

Bijaya Ketan Panigrahi, *Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India*

Federica Pascucci, *Department di Ingegneria, Università degli Studi Roma Tre, Rome, Italy*

Yong Qin, *State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China*

Gan Woon Seng, *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore*

Joachim Speidel, *Institute of Telecommunications, University of Stuttgart, Stuttgart, Germany*

Germano Veiga, *FEUP Campus, INESC Porto, Porto, Portugal*

Haitao Wu, *Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China*

Walter Zamboni, *Department of Computer Engineering, Electrical Engineering and Applied Mathematics, DIEM—Università degli studi di Salerno, Fisciano, Italy*

Kay Chen Tan, *Department of Computing, Hong Kong Polytechnic University, Hong Kong, Hong Kong*

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering—quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

Huaping Liu · Di Guo
Editors

Proceedings of 2025 Chinese Intelligent Automation Conference

Volume III

Editors

Huaping Liu
Computer Science and Technology
Tsinghua University
Beijing, China

Di Guo
Beijing University of Posts
and Telecommunications
Beijing, China

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-981-95-4048-8

ISBN 978-981-95-4049-5 (eBook)

<https://doi.org/10.1007/978-981-95-4049-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Singapore Pte Ltd. 2026

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.

Contents

Integration of Modal Decomposition and Error Correction for Port Power Load Forecasting	1
<i>Ziru Niu, Qiang Wang, Yaixiang Yang, Hanrui Jiang, and Xibin Xiao</i>	
Active Fault-Tolerant Control for Uncertain Steer-by-Wire System Using Radial Basis Function Neural Network	10
<i>Qingwang Zeng, Yue Zhu, Ming Yu, and Zhening Liang</i>	
A Two-Stage Optimization Algorithm for the Multi-dropoff Flying Sidekick Traveling Salesman Problem	21
<i>Hongrui Ao, Jihui Zhang, and Shengzong Chen</i>	
Optimization of Vibration Suppression for a Humanoid Robot's Lower Limb	32
<i>Jie Yang, Jisong Ding, Yiqun Peng, and Tianyu Liu</i>	
Design and Implementation of an Intelligent Waste Sorting and Classification System Based on YOLOv5s	43
<i>Yulin Peng, Xiangpeng Liu, Hui Zhang, and Kang An</i>	
Missed Sample Exploration by Class-Agnostic Detector for Embodied Learning	55
<i>Panpan Wang and Di Guo</i>	
Node Classification and Link Prediction Methods Based on Hybrid Teacher Models	64
<i>Yuxin Chen, Hao Yan, Qingyou Dai, Diting Zhou, Haohao Fu, Maoyu Wang, and Zhenbo Cheng</i>	
Optimized On-Off Control of Artificial Climate Chambers via Physics-Neural Hybrid Modeling and DQN	76
<i>Han Zeng, Linli Zhou, Dayong Wang, and Caiguo Tang</i>	
Parking Path Planning Method Based on Improved Hybrid A* Algorithm	86
<i>Jiaming Li, Bin Lan, Lunfei Liang, Gang Yan, and Jiahao He</i>	
UAV Imagery Based Railroad Tunnel Facility Instance Segmentation Using Post-processing of Spatial Topological Relationships	93
<i>Tong Meng, Yong Qin, Fanteng Meng, Ninghai Qiu, Chongchong Yu, and Zhipeng Wang</i>	

Self-Organized Radar Signal Sorting via Multi-order DTOA Spectral Analysis and Parameter Entropy Dynamic Fusion	105
<i>Xun Huang, Dan Huang, Lu Ding, Chenggang Wang, and Lei Song</i>	
Fault-Tolerant Control for Dual Three-Phase PMSM Based on Multi-Objective Optimization and Multi-Subspace Coordination	116
<i>Wenhong Pang, Yang Gao, and Qingwei Chen</i>	
Three-Vector-Based Model Predictive Flux Control For SPMSM Using Ultralocal Model	128
<i>Rongbo Lai, Lu Liu, Shihong Ding, Keqi Mei, and Jinlin Sun</i>	
TB-DeepSeek-Agent: A Large Language Model for Tuberculosis Incidence Prediction with Web-Based Automatic Report	137
<i>Mingming Chen, Yu Lu, and Tenglong Li</i>	
Bayesian-Optimized Eye Movement Event Detection for Assisted Alzheimer's Diagnosis	147
<i>Xin Wang, Yujie Nie, Jia Zhao, Yahan Wang, Yibin Li, Xiangqing Xu, and Xin Ma</i>	
Enhanced MobileNetV3 for Lightweight Yi Script Recognition	157
<i>Haipeng Sun, Jian Sun, Han Yu, Bohan Guo, Mengyan Liu, and Hengrui Dai</i>	
Three-Dimensional Motion Estimation of Multibeam Imaging Sonar	167
<i>Yupei Huang, Peng Li, Zhengxing Wu, and Junzhi Yu</i>	
A High Throughput Deep Learning Platform for Edge Inference Based on Hardware/Software Codesign	175
<i>Yuwei Zhang, Fan Wu, and Chaonong Xu</i>	
Visual SLAM Based on Point Line Feature Fusion in Dynamic Environments	189
<i>Benlong Xiang, Lichuan Ning, and Yuanmin Xie</i>	
Cooperative Path Planning for Multi-UAVs Using MASAC with SEAttention and Prioritized Experience Replay	200
<i>Wei Xia, Boxiang Li, Jinliang Zhu, and Denghao Pang</i>	
Robust and Natural Bipedal Robot Locomotion Control in Complex Terrains Using Enhanced Adversarial Motion Priors	212
<i>Hanting Li, Kang An, and Yaqing Song</i>	

Experimental Investigation of Drag and Lift Forces on Objects Moving Through Granular Media with Varying Geometries and Speeds	222
<i>Chundong Liu, Xiaopu Hao, Zhiqiang Song, Qiang Zhang, Shuai Kang, and Longchuan Li</i>	
Structure Design and Output Characteristics of Hemispherical Tactile Sensor for 3D Force Detection	230
<i>Xiaotao Du, Ling Weng, Shichao Zuo, Bowen Cui, Xiaopeng Ji, and Lanyang Hao</i>	
Time-Optimal Trajectory Planning for Robotic Arms Based on VPPSO Algorithm	240
<i>Xiao Shi, Zhiyong He, and Yucai Zhou</i>	
Multi-spacecraft Orbital Game Combined with Potential Field Function	251
<i>Zihan Li and Shan Lu</i>	
Research on the Design and Control Method of a Controllable Deformable Wheeled Stair-Climbing Robot	260
<i>Dongsheng Shao, Lunfei Liang, Xiaoxiong Du, Jiansheng Zhou, Shuo Cao, and Yu Liu</i>	
Intelligent Stable Walking Based on Nonlinear Disturbance Observer	271
<i>Helin Wang</i>	
VLM-Based Dense Reward for Zero-Shot Robot Manipulation	278
<i>Kehao Shi, Zhenyi Xu, Yang Cao, Jun Huang, and Yu Kang</i>	
Reinforcement Learning-Based Autonomous Control Strategy for Snake Robots in Confined Terrains	287
<i>Cheng Wang, Yunhu Zhou, Yuhong Na, Ku Du, and Darong Huang</i>	
Optimization of YOLOv12 for Steel Surface Defect Detection Under Class Imbalance	297
<i>Hui Han, Yunyun Yang, Xiaoyuan Liu, Jiamin Wei, and Xinying Xu</i>	
Autonomous Landing Control Method for UAV Multistate Platform Based on Dynamic Weights PPO	307
<i>Yalin He, Bo Huang, Yinlong Yuan, and Yun Cheng</i>	
Real-Time Intrusion Object Detection in Low-Light Rail Surveillance Images Based on Triple Attention Mechanism	317
<i>Xuan Xu, Yi Wang, Yeting Gong, Yuhang Shi, Weibin Zhu, and Tangwen Yang</i>	

A Quantitative Calculation Scheme for Network Vulnerable Nodes Based on Dynamic Trust Evaluation	328
<i>Jian Liang, Pin Lv, Yangming Kang, and Yaxing Zhang</i>	
Big-Five Personality Trait Analysis in Videos via DLF-Based Multi-Scale Spatiotemporal Modeling	335
<i>ZhengHua Guo, Tao Ning, Jin Zhang, QiDong Hou, and JingSong Chen</i>	
Research Progress of Elderly Care and Rehabilitation Robots	345
<i>Yang Yang, Lunfei Liang, Yu Liu, and Houde Liu</i>	
Anti-swing Control of 5-DOF Marine Crane Based on ASMC-PD Mode Switching Method	354
<i>Qin Zhang, Gaoli Xu, and Xiong Hu</i>	
Co-simulation and Validation Methodology for Unmanned Tracked Vehicle	365
<i>Lunfei Liang, Huixin Huang, Geng Chen, Song Liu, and Yan Pan</i>	
6-DoF Object Pose Estimation via Improved Attitude Optimization Network	375
<i>Yue Wang, Kang An, and Hui Zhang</i>	
The Design and Kinematic Analysis of a Seven-DOF Electric Operation Robot	384
<i>Ding Ma, Haibo Du, Xunhong Sun, Wenwu Zhu, Runchen Zhao, and Yuchen Li</i>	
Design and Output Characteristics of Carbon Nanotube-Silicone Rubber Based Piezoresistive Tactile Sensors	394
<i>Bowen Cui, Ling Weng, Shichao Zuo, Xiaotao Du, Xiaopeng Ji, and Lanyang Hao</i>	
Multi-objective Path Planning for Two-Stage Mobile Robot Based on Improved ACO and Q-Learning	407
<i>Xianyong Ruan, Lichuan Ning, Bo Tao, and Yuanmin Xie</i>	
Adaptive Neural Bipartite Containment Control for Stochastic Non-strict Feedback Multi-agent Systems	417
<i>Hailin Tang and Tianping Zhang</i>	
Vibration-Assisted Dual-Helix Burrowing Robot for Enhanced Locomotion in Granular Media	425
<i>Guancheng Zhao, Longchuan Li, Zhiqiang Song, Shuqian He, and Shuai Kang</i>	

A Reinforcement Learning Control Method Based on Pre-training and Truncated PPO for Hand Rehabilitation	434
<i>Lu Liu, Lu Sun, Benyan Huo, Yanhong Liu, and Zan Zhang</i>	
Non-contact Heart Rate Estimation Based on U-Net Global Facial Video	444
<i>Cheng Chen, Yurun Ma, Jiaqi Li, and Yudan Liu</i>	
A Robust SLAM Backend Framework with Multi-landmark Constraints in Industrial Logistics Scenarios	457
<i>Yu Li, Bo Li, Yuanpeng Li, Jianyuan Ruan, Jianyun Hong, and Aihua Yu</i>	
Multi-scale Hierarchical Feature Attention Interaction for Fault Diagnosis of Rolling Bearings Under Noisy Conditions	469
<i>Hanwen Chang, Fun Zhou, and Jiechen Sun</i>	
Electromechanical Co-simulation Experiment System of Floating Offshore Wind Turbine	480
<i>Chuanjie Li and Zhaohua Liu</i>	
Author Index	489



Integration of Modal Decomposition and Error Correction for Port Power Load Forecasting

Ziru Niu¹, Qiang Wang¹, Yaxiang Yang¹, Hanrui Jiang², and Xibin Xiao¹(✉)

¹ School of Transportation and Logistics Engineering, Wuhan University of Technology, Wuhan 430063, China

wangqiang@whut.edu.cn, 1214278648@qq.com

² Central Southern China Electric Power Design Institute Co., LTD. of China Power Engineering Consulting Group, Wuhan 430071, China

Abstract. Accurate power load forecasting is critical for ensuring the stable operation of port microgrids. To address the inherent randomness and volatility of port power loads, this paper proposes a novel forecasting method that integrates optimized mode decomposition with error correction. The Pearson correlation coefficient is employed to evaluate the correlation between the load and influencing factors, such as production operations. Optimal parameters for the mode decomposition model are determined using the Sparrow Search Algorithm (SSA), which effectively decomposes the original load signal into distinct components that capture both the overall trend and local characteristics. Subsequently, a prediction and reconstruction model incorporating error correction is established. This model leverages forecasted error values to refine preliminary predictions, thereby significantly improving forecasting accuracy. Validation using real-world port operational data demonstrates the effectiveness of the proposed method across multiple time scales. Notably, the performance is particularly robust for 8-h-ahead and 24-h-ahead forecasts, achieving a Mean Absolute Percentage Error (MAPE) below 5% and forecasting accuracy as high as 95%.

Keywords: Port power load forecasting · Variational modal · Decomposition · Sparrow search algorithm · Error correction

1 Introduction

With the global emphasis on sustainable development, ports are encountering increasing challenges related to energy consumption. To mitigate these challenges, many ports have incorporated renewable energy sources such as wind and solar power into their operations. However, the intrinsic volatility and intermittency of these renewable sources pose significant threats to the stable operation of port microgrids. This issue is further compounded by operational uncertainties and the heterogeneous power demands of various port equipment. Given that relying solely on generation-side regulation is inadequate, the importance of demand-side management becomes increasingly pronounced. Consequently, precise forecasting of port power consumption emerges as a critical component for effective microgrid planning, dispatching, and optimization, thereby ensuring grid stability and facilitating the sustainable transformation of ports.

Load forecasting methods can generally be categorized into two main types: statistical methods and machine learning methods. Statistical methods, such as time series analysis [1] and regression analysis [2], are well-established but often depend significantly on the stability of input sequences. Reference [3] utilized the FARIMA model combined with a CS algorithm for parameter optimization, achieving high-accuracy short-term power load forecasting. Reference [4] enhanced linear regression by introducing fuzzy input variables, thereby mitigating the influence of weather-related errors and improving prediction accuracy.

Machine learning methods, such as neural networks [5] and support vector machines (SVM) [6], demonstrate superior capabilities in capturing both linear and nonlinear relationships within power load data. Reference [7] optimized the structure of neural networks by leveraging Elman and BP networks to construct forecasting models, thereby significantly enhancing prediction accuracy and computational efficiency. Reference [8] integrated multivariate phase space reconstruction with SVM to achieve precise load forecasting for integrated energy systems. To address the nonlinear and non-stationary characteristics of load data, Reference [9] combined empirical mode decomposition (EMD), convolutional neural networks (CNN), and long short-term memory (LSTM) networks for forecasting. However, EMD is prone to mode mixing, which adversely affects prediction performance. Reference [10] mitigated this issue by combining EMD with local mean decomposition (LN) and LSTM, yet boundary effects still constrained the overall accuracy. Reference [11] adopted variational mode decomposition (VMD) to preprocess load data, effectively avoiding mode mixing and boundary effects, and subsequently combined it with gated recurrent unit (GRU) networks for prediction, achieving high accuracy. Nevertheless, this method relies on empirical settings for critical VMD parameters, such as the penalty factor and decomposition mode number, which lack theoretical justification. This reliance may lead to over-decomposition or under-decomposition, thereby impacting VMD performance and compromising the accuracy and stability of load forecasting.

Port power loads are characterized by frequent fluctuations, which pose significant challenges to the stability of microgrid operations. Existing forecasting methods struggle to adequately address these challenges due to the complex and dynamic nature of port operational environments. Consequently, there is a critical need for approaches that are more closely aligned with the unique characteristics of port power loads. To this end, this study proposes a novel port power load forecasting method that integrates optimal modal decomposition with error correction mechanisms. The proposed method was validated through experiments utilizing real-world power load data collected from a domestic port. The results indicate that this method not only achieves higher prediction accuracy but also demonstrates superior reliability and stability compared to existing techniques.

2 Port Power Load Forecasting Method

2.1 Forecasting Framework

To address the challenges associated with decomposition parameter selection and model accuracy in port power load forecasting, this paper proposes an integrated forecasting method that combines optimal modal decomposition with error correction. The proposed

approach utilizes advanced modal decomposition techniques for refined data processing and integrates an error correction mechanism to improve prediction accuracy and reliability. This provides an effective solution for enhancing the performance of port power load forecasting. As depicted in Fig. 1, the framework comprises four key modules.

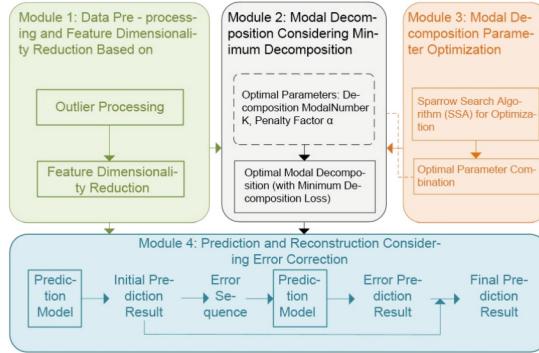


Fig. 1. Framework for Port Power Load Forecasting Integrating Optimal Modal Decomposition and Error Correction

Module 1: Raw load data is preprocessed to detect and correct outliers, resulting in refined load data. The Pearson Product-Moment Correlation Coefficient (PPMCC) is computed between the corrected load data and potential influencing factors to evaluate their relationships.

Module 2: This module establishes a Variational Mode Decomposition (VMD) parameter optimization model. By leveraging the Sparrow Search Algorithm (SSA) from Module 3, it determines the optimal penalty factor and decomposition mode number that minimize decomposition loss, ensuring high-quality modal decomposition.

Module 3: SSA is utilized to iteratively search for the optimal penalty factor and decomposition mode number that minimize decomposition loss. This process enhances the effectiveness and precision of the modal decomposition procedure.

Module 4: This module develops a predictive model to generate initial forecasts. Subsequently, the error sequence between the initial predictions and the corrected load data is analyzed. The Intrinsic Mode Functions (IMFs) of the error sequence are predicted and integrated into the correction of the initial forecasts, thereby producing final refined predictions with significantly improved accuracy.

2.2 Data Preprocessing and PPMCC

Prior to constructing the forecasting model, it is imperative to address anomalies through two critical steps: detection and correction, thereby ensuring the reliability of the data. Anomalies are identified when a data point deviates by more than 10% from its preceding point. This criterion can be formally expressed as:

$$\left| \frac{y_n - y_{n-1}}{y_{n-1}} \right| > 10\% \quad (1)$$

where y_n is the abnormal data; y_{n-1} is the data point immediately preceding the abnormal value.

When a data point is identified as an outlier, it is replaced using the following formula:

$$y'_n = y_{n-1} + \frac{y_n - y_{n-1}}{|y_n - y_{n-1}|} MC \quad (2)$$

$$MC = \frac{1}{n} \sum_{i=1}^n |y_i - y_{i-1}| \quad (3)$$

where y'_n is the replacement value. MC is the average rate of change across all data points.

To minimize the adverse effects of dimensional inconsistencies among input features on prediction accuracy, feature normalization is conducted using the following formula:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4)$$

where x_n is the normalized data; x is the original data; x_{\max} is the maximum value in the original dataset; x_{\min} is the minimum value in the original dataset.

Selecting appropriate influencing features as auxiliary inputs is critical for enhancing prediction accuracy. Excessive or irrelevant features can lead to increased model dimensionality and computational burden, which may degrade prediction performance. To address this issue, this paper utilizes the Pearson Product-Moment Correlation Coefficient (PPMCC) for feature selection and dimensionality reduction, thereby ensuring the efficiency and accuracy of the predictive model. The calculation formula is presented as follows:

$$\rho = \frac{Cov(X, Y)}{\sqrt{Var[X]Var[Y]}} \quad (5)$$

where ρ is the Pearson correlation coefficient, representing the correlation between X and Y; $Cov(X, Y)$ is the covariance between x and y; $Var[X]$ and $Var[Y]$ is the variance of x and y. The correlation criteria are shown in Table 1.

Table 1. Pearson Correlation Criterion

$ \rho $	correlation
0. 8–1. 0	very strong correlation
0. 6–0. 8	strong correlation
0. 4–0. 6	moderate correlation
0. 2–0. 4	weak correlation
0. 0–0. 2	very weak or negligible correlation

2.3 Modal Decomposition Considering Minimum Decomposition Loss

In the field of load forecasting, the selection of the number of decomposition modes and the penalty factor in Variational Mode Decomposition (VMD) [12] plays a critical role in ensuring the accuracy and effectiveness of the forecasting model.

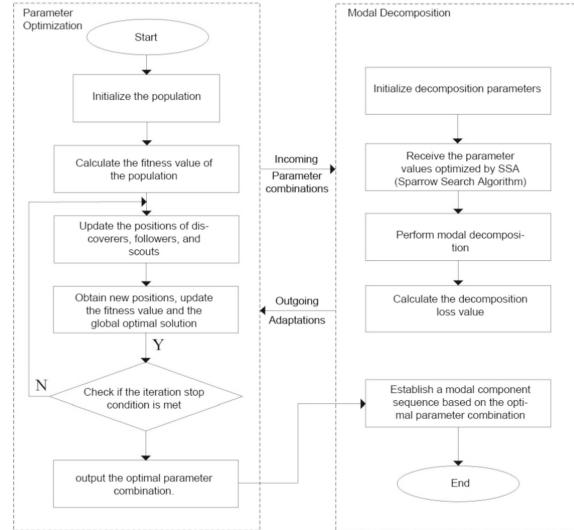


Fig. 2. Modal decomposition process considering the minimum decomposition loss

The setting of the penalty factor directly influences the bandwidth of the modal components, thereby affecting the decomposition performance. This paper proposes a modal decomposition method that minimizes decomposition loss by introducing the Sparrow Search Algorithm (SSA) [13] to intelligently optimize the decomposition parameters. The core concept of this method is to minimize decomposition loss, with the decomposition loss serving as the fitness function. The SSA is utilized to search for the optimal combination of the number of decomposition modes and the penalty factor. The specific steps are illustrated in Fig. 2.

2.4 Prediction and Reconstruction with Error Correction

Considering that the load Intrinsic Mode Functions (IMFs) consist of both smoothly varying low-frequency components and highly stochastic high-frequency components, this study proposes an integrated module that combines error correction with predictive reconstruction, thereby achieving more accurate forecasting results.

Traditional Gradient Boosting Decision Tree (GBDT) algorithms are susceptible to gradient estimation bias, thereby increasing the risk of overfitting. In contrast, CatBoost (Categorical Boosting) optimizes the training process via ordered boosting, effectively reducing the potential for overfitting while exhibiting superior generalization capabilities. Consequently, this study selects CatBoost as the core prediction model and integrates

it with a one-step error correction strategy to achieve efficient and precise forecasting. As illustrated in Fig. 3, this paper proposes an integrated module that combines prediction, reconstruction, and error correction. This module not only fully exploits CatBoost's exceptional generalization capability but also further enhances prediction accuracy through the error correction mechanism, providing a reliable and precise solution for load forecasting.

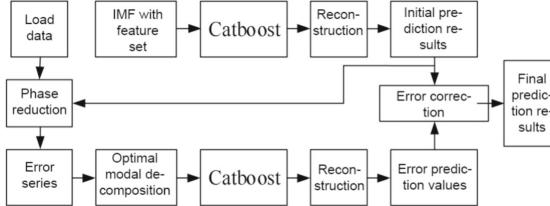


Fig. 3. Prediction and reconstruction process considering error correction

3 Experimental Analysis

3.1 Experimental Setup

This paper employs 181 days of power load and influencing factor data collected from a port between November 1, 2021, and April 30, 2022. The data, sampled at an interval of 1 h, comprises a total of 4,338 data points. The load data curve is presented in Fig. 4.

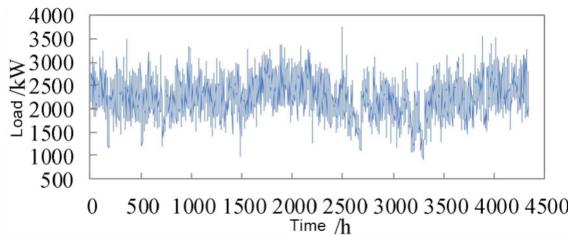


Fig. 4. Load data of a certain port

The dataset is divided into three subsets: an initial training set, an error training set, and a validation set, following an 8:1:1 ratio. The initial training set is utilized to train the primary forecasting model, the error training set is employed to train the error prediction model, and the validation set is used to assess the forecasting performance. The influencing factor features comprise ten variables, including temperature and the number of vessels in port, with detailed information provided in Table 2.

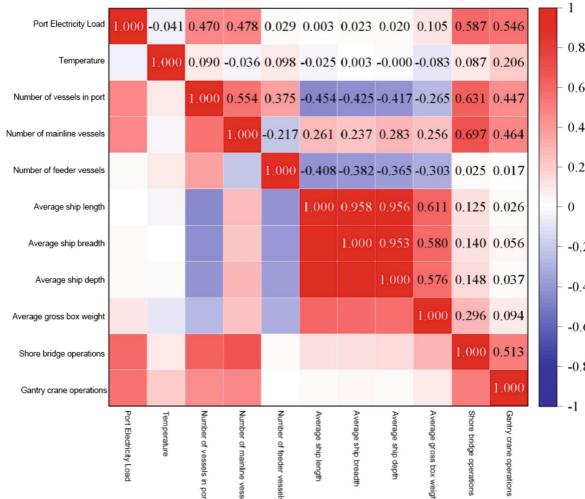
To evaluate the predictive performance of the model, the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are selected as evaluation metrics.

Table 2. Definitions of Factors Influencing Port Power Load

Factor	Definition
Temperature	Ambient air temperature
Number of vessels in port	Total count of vessels currently in port
Number of feeder vessels	Count of feeder vessels in port
Average vessel length	Mean length of vessels in port
Average vessel width	Mean width of vessels in port
Average vessel draft	Mean depth of vessels
Average container weight	Mean weight of loaded/unloaded containers per vessel
Quay crane throughput	Total container volume handled by all quay cranes
RTG throughput	Total container volume handled by all rubber - tired gantry cranes (RTGs)

3.2 Data Preprocessing and Feature Dimensionality Reduction

Based on Eqs. (1, 2, 3, and 4), anomalies in the original port power load data were detected and corrected. The Pearson Product-Moment Correlation Coefficient (PPMCC) between the port power load and each influencing factor was computed, with the results presented in Fig. 5. In the figure, darker colors signify stronger correlations, whereas lighter colors denote weaker correlations.

**Fig. 5.** PPMCC between the power load and the characteristics of various influencing factors

Based on this analysis, quay crane workload, rubber-tired gantry crane workload, mainline vessel count, and in-port vessel count are identified as the key input features for the forecasting model.

3.3 Forecasting Results

The data in Table 3 indicate that the error-corrected final prediction results substantially outperform the initial prediction results across all time scales of 8 h, 24 h, and 144 h in terms of prediction evaluation metrics. Specifically, for the 8 h time scale prediction, the final prediction results achieve the best performance with MAPE and RMSE values of only 4.2% and 90.9 kW, respectively. Similarly, for the 24 h and 144 h time scale predictions, the final prediction results show MAPE improvements of 1.9% and 1.3%, respectively, while the RMSE values are reduced by 47.0 kW and 34.5 kW, respectively. These findings clearly demonstrate the effectiveness of the error correction strategy, leading to a significant improvement in the prediction accuracy of port power load.

Table 3. Comparison of Evaluation Metrics for Final Forecasting Results

Decomposition Method	Model	8 h		24 h		144 h	
		MAPE%	RMSEkW	MAPE%	RMSEkW	MAPE%	RMSEkW
None	Initial Forecasts	6.2	133.3	6.8	145.8	6.7	192.8
With Adjustment	Final Forecast	4.2	90.9	4.9	98.8	5.4	158.3

4 Conclusions

This paper proposes a novel forecasting method that integrates optimal modal decomposition and error correction techniques for highly stochastic and volatile port power load data. By adopting minimum decomposition loss as the optimization objective and utilizing the Sparrow Search Algorithm (SSA) to determine decomposition parameters, the method achieves optimal modal decomposition of load data while eliminating subjective biases. Moreover, a unified prediction module combining forecasting, reconstruction, and error correction is designed to achieve precise load predictions. Experimental results indicate that the proposed method performs exceptionally well in port power load forecasting, accurately capturing both overall load trends and local variations while surpassing comparative methods in terms of prediction stability and performance. Despite its ability to integrate multiple algorithms for delivering high accuracy and robustness, the method has certain limitations, including a relatively complex model architecture, substantial computational demands, and longer runtime.

References

1. Gasparin, A., Lukovic, S., Alippi, C.: Deep learning for time series forecasting: the electric load case [J]. *CAAI Trans. Intell. Technol.* **7**(1), 1–25 (2022)
2. Aprillia, H., Huang, C.M., Yang, H.T.: Short term load forecasting using PSO – DWT – MLR at system and end – user levels [C]. In: 2019 8th International Congress on Advanced Applied Informatics (IIAI – AAI), pp. 462–467. IEEE (2019)
3. Wu, F., Cattani, C., Song, W., et al.: Fractional ARIMA with an improved cuckoo search optimization for the efficient short – term power load forecasting [J]. *Alex. Eng. J.* **59**(5), 3111–3118 (2020)
4. Petrusic, A., Janjic, A.: Fuzzy Multiple linear regression for the short term load forecast [C]. In: 2020 55th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Niš, Serbia, Sep 10–12, 2020. IEEE, pp. 102–105 (2020)
5. Zhao, J., Zhang, C., Chen, J.: A combined short – term power forecasting method incorporating time series decomposition [J]. *J. Chongqing Univ. Technol. (Nat. Sci.).* **35**(12), 180–188 (2021)
6. Zhang, A., Zhang, P., Feng, Y.: Short – term load forecasting for microgrids based on DA – SVM [J]. *COMPEL Int. J. Comput. Math. Electr. Electron. Eng.* **38**(1), 68–80 (2019)
7. Zheng, X., Ran, X., Cai, M.: Short – term load forecasting of power system based on neural network intelligent algorithm [J]. *IEEE Access.* (2020)
8. Liu, H., Tang, Y., Pu, Y., et al.: Short – term load forecasting of multi – energy in integrated energy system based on multi – variate phase space reconstruction and support vector regression mode [J]. *Electric Power Syst. Res.* **210**, 108066 (2022)
9. Xu, Y., Xiang, Y., Ma, T.: Short – term power load forecasting based on EMD – CNN – LSTM hybrid model [J]. *Journal of North China Electric Power University.* **49**(02), 81–89 (2022)
10. Wei, A., Mao, D., Han, W., et al.: Research on short – term power load forecasting based on EMD and LSTM network [J]. *J. Eng. Thermal Energy Power.* **35**(04), 203–209 (2020)
11. Huang, D., Tang, Z., Hu, A., et al.: Short – term power load forecasting method based on VMD – GRU [J]. *Internet Things Technol.* **11**(04), 67–70 (2021)
12. Zhang, L., Li, T., Gong, Y., et al.: Research on power load forecasting considering variational mode decomposition residuals [J]. *J. Chongqing Univ. Technol. (Nat. Sci.).* **36**(01), 165–170 (2022)
13. Xue, J., Shen B: A novel swarm intelligence.



Active Fault-Tolerant Control for Uncertain Steer-by-Wire System Using Radial Basis Function Neural Network

Qingwang Zeng^{1(✉)}, Yue Zhu², Ming Yu¹, and Zhenying Liang¹

¹ Hefei University of Technology, 230009 Hefei, China

2022010047@mail.hfut.edu.cn

² SWJTU-Leeds Joint School, Southwest Jiaotong University, 611756 Chengdu, China

Abstract. In this paper, a fault-tolerant control method is developed for steer-by-wire system with sensor fault and uncertainties. When a sensor fault occurs in the system, an adaptive sliding mode observer is employed for fault reconstruction. Based on the reconstructed results, a recursive non-singular terminal sliding mode control (RNTSMC) based on radial basis function neural network (RBFNN) is proposed to track the angle of the front wheel. The simulation results demonstrate that the proposed active fault-tolerant method effectively achieves fault-tolerant control of sensor fault with high accuracy. Moreover, compared to traditional PID controller, the RBFNN-RNTSMC exhibits superior tracking performance in terms of both accuracy and speed.

Keywords: Steer-by-wire system · fault-tolerant control · non-singular terminal sliding mode · radial basis function neural network

1 First Section

In recent years, steer-by-wire (SBW) system has gained increasing popularity among automotive researchers due to its ability to enhance steering efficiency and meet the demands of modern automobile intelligence and lightweight design [1]. The SBW system eliminates the mechanical linkage between the steering wheel and front wheels, using sensors to detect driver input and electronically transmitting control signals to actuators [2]. In comparison to conventional steering systems, the SBW system optimizes steering performance to enhance operational stability and ride comfort, thereby establishing a robust hardware foundation for autonomous steering in self-driving vehicles [3]. The use of electromechanical equipment increases the risk of system failure and hardware malfunction, requiring a strong emphasis on safety-related steering functionalities in vehicle development.

This work was supported by the National Natural Science Foundation of China under Grant 62173119 and the National College Student Innovation and Entrepreneurship Training Program under Grant 202310359020.

Fault-tolerant control is widely employed in critical domains such as aerospace and nuclear energy, serving as a crucial approach to enhance system security and reliability. There are two approaches which are commonly adopted to maintain system stability: passive fault-tolerant control (PFTC) and active fault-tolerant control (AFTC) [4–6]. However, when compared with PFTC, AFTC often exhibits superior control performance and offers more flexible design methodologies. Therefore, AFTC holds greater practical value in real applications. In [7], an adaptive fault observer is devised for a three-degree-of-freedom helicopter system which is capable of concurrently detecting actuator and sensor faults. After that, the controller is compensated online using the estimated values obtained from adaptive observer.

With the advancement of vehicle chassis control technology, fault-tolerant control has been introduced to enhance safety and stability. The approach in [8] employs analytical redundancy through multiple position estimation methods for motor and steering sensors to improve fault tolerance. In [9], a delta domain minimax model predictive control is used to ensure robust tracking despite actuator faults, system uncertainties, and disturbances, supporting fault-tolerant control in steer-by-wire systems.

In this paper, an AFTC method for the SBW systems with uncertainty and sensor fault is proposed. Firstly, a recursive non-singular terminal sliding mode control (RNTSMC) based on radial basis function neural network (RBFNN) is developed to track the front wheel angle under healthy condition. In case of sensor fault, an adaptive sliding mode observer (ASMO) is employed to reconstruct the fault, and an adaptive active fault tolerance control (AAFTC) is developed using the fault reconstruction results. The fault detection and isolation (FDI) mechanism is used to switch the control law in real time to achieve AFTC under sensor fault.

2 SBW System Modeling

2.1 Front Wheel Subsystem Modeling

The dynamic model of the front wheel subsystem of the SBW system is described as follows:

$$\ddot{\theta}_{fw} = \psi V_{in} + \Psi(\theta_{fw}, \dot{\theta}_{fw}) + d \quad (1)$$

$$\text{with } \psi = \frac{G_{p1}G_{p2}r_pK_1K_2}{(G_{p1}G_{p2})^2J_s + (G_{p2}r_p)^2M_p + r_p^2J_{fw}},$$

$$\begin{aligned} \Psi(\theta_{fw}, \dot{\theta}_{fw}) &= \frac{G_{p1}}{r_p} T_{v,s} \dot{\theta}_{fw} \\ &+ (T_{s,s} - T_{c,s}) e^{-\frac{a_i G_{p1} G_{p2}}{r_p} |\dot{\theta}_{fw}|} \text{sign}(\dot{\theta}_{fw}) \\ &+ T_{c,s} \text{sign}(\dot{\theta}_{fw}) + G_r T_{v,r} \dot{\theta}_{fw} \\ &+ (T_{s,r} - T_{c,r}) e^{-a_i G_{p2} |\dot{\theta}_{fw}|} \text{sign}(\dot{\theta}_{fw}) \\ &+ T_{c,r} \text{sign}(\dot{\theta}_{fw}) + \frac{1}{G_r} T_{v,fw} \dot{\theta}_{fw} \\ &+ (T_{s,fw} - T_{c,fw}) e^{-a_i |\dot{\theta}_{fw}|} \text{sign}(\dot{\theta}_{fw}) \\ &+ T_{c,fw} \text{sign}(\dot{\theta}_{fw}) \end{aligned}$$

where θ_{fw} is the front wheel steering angle, V_{in} is the vector of inputs, G_{p1} is the reduction ratio, G_{p2} is the ratio between rack displacement and front wheel angle, r_p is the radius of the pinion, K_1 is the voltage-to-current ratio, K_2 is the current-to-torque ratio, J_s is the rotor inertia, M_p is the mass of the rack, J_{fw} is the inertia of front wheel, $T_{v,i}$, $i \in \{s, p, fw\}$, is the viscous friction coefficient, $T_{s,i}$, $i \in \{s, p, fw\}$, is the Coulomb friction torque, $T_{c,i}$, $i \in \{s, p, fw\}$, is the static friction torque, $a_i \in (0, 1)$, $i \in \{s, p, fw\}$, is a constant describing the transition between $T_{s,i}$ and $T_{c,i}$, and $|d| < D$ is the disturbance from both internal and external sources and D is the upper bound of the disturbance.

3 Design of Fault Estimation Observer

3.1 Sensor Fault Reconstruction

The state-space model of the front wheel subsystem under sensor fault is described as follows:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) + \mathbf{M}\Omega(\mathbf{x}, t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{E}f(t) \end{cases} \quad (2)$$

where $\mathbf{x}(t) = [x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6] = [\theta_s \ \dot{\theta}_s \ \theta_p \ \dot{\theta}_p \ \theta_{fw} \ \dot{\theta}_{fw}]$, with θ_s is the steering motor rotational angle, θ_p is the rack displacement, $f(t)$ is the sensor fault vector, $\mathbf{A} =$

$$\left[\begin{array}{cccccc} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{K_1}{J_s} & -\frac{T_{v,s}}{J_s} & \frac{G_{p1}C_s}{r_p J_s} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{G_{p1}C_s}{r_p M_p} & 0 & -\frac{C_s G_{p1}^2}{M_p r_p^2} - \frac{G_{p2}^2 C_p}{M_p} - \frac{T_{v,r}}{M_p} & \frac{C_p G_{p2}}{M_p} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{C_p G_{p2}}{J_{fw}} & 0 & -\frac{C_p}{J_{fw}} - \frac{T_{v,fw}}{J_{fw}} & \end{array} \right], \quad C_s \text{ and } C_p \text{ are the shaft rigidity}$$

of steering motor and pinion, respectively, $\mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T$, $\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$,

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Omega(\mathbf{x}, t) = [0 \ g(x_1, s)/J_s \ 0 \ g(x_3, p)/M_p \ 0 \ g(x_5, fw)/J_{fw}]$$

$[g(\vartheta, i) = T_{\tau,i}\dot{\vartheta} + T_{c,i}\text{sign}(\dot{\vartheta}) + (T_{s,i} - T_{c,i})e^{-a_i|\dot{\vartheta}|}\text{sign}(\dot{\vartheta})]$, $\vartheta \in \{x_1, x_2, x_3\}$, $i \in \{s, p, fw\}$, and $\Omega(\mathbf{x}, t)$ satisfies the Lipschitz condition with respect to state x :

$$\|\Omega(\mathbf{x}_1, t) - \Omega(\mathbf{x}_2, t)\| \leq \beta \|\mathbf{x}_1 - \mathbf{x}_2\| \quad (3)$$

where β is the Lipschitz constant.

Then, a two-dimensional state variable $z(t) \in R^2$ is introduced:

$$z(t) = \int_0^t y(\tau) d\tau = \int_0^t [Cx(\tau) + f(\tau)] d\tau \quad (4)$$

Combining (2) and (4) leads to the augmented state-space model as follows:

$$\begin{cases} \dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}u(t) + \bar{M}_d\Omega(\bar{x}, t) + M_f f(t) \\ \bar{y}(t) = \bar{C}\bar{x}(t) \end{cases} \quad (5)$$

where $\bar{x} = \begin{bmatrix} z \\ x \end{bmatrix}$, $\bar{A} = \begin{bmatrix} 0 & C \\ 0 & A \end{bmatrix}$, $\bar{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}$, $\bar{C} = [I_2 \ 0]$, $\bar{M}_d = \begin{bmatrix} 0 \\ M \end{bmatrix}$, and $M_f = \begin{bmatrix} I_2 \\ 0 \end{bmatrix}$.

According to the principle of the Edwards-Spurgeon observer [10], the following ASMO can be designed:

$$\begin{cases} \dot{\hat{x}}(t) = \bar{A}\hat{x}(t) + \bar{B}u(t) + \bar{M}_d\Omega(\hat{x}, t) \\ \quad - L(\hat{y} - \bar{y}) + M_f v \\ \hat{y}(t) = \bar{C}\hat{x}(t) \end{cases} \quad (6)$$

where \hat{x} is the estimate of \bar{x} , \hat{y} is the estimate of \bar{y} , L is the linear gain matrix of ASMO, and v is a discontinuous sliding mode input term, which can be represented as follows:

$$v = \begin{cases} -\frac{\hat{\phi}He_y}{\|He_y\|}, \|He_y\| \neq 0 \\ 0, \quad \|He_y\| = 0 \end{cases} \quad (7)$$

where $e_y = \hat{y} - \bar{y}$ represents the estimation error of the output, and $\hat{\phi}$ is an adjustable gain parameter with the following adaptive laws:

$$\dot{\hat{\phi}} = \phi_1 \|He_y\|, \hat{\phi} \leq \phi^* \quad (8)$$

where $\phi_1 > 0$, $\phi^* > \delta$.

Then, by defining the state estimation error $e_x = \hat{x} - \bar{x}$ and combining (5) and (6), the following error system can be obtained:

$$\dot{e}_x = \dot{\hat{x}} - \dot{\bar{x}} = (\bar{A} - \bar{L}\bar{C})e_x + \bar{M}_d[\Omega(\hat{x}, t) - \Omega(\bar{x}, t)] + M_f v - M_f f \quad (9)$$

where $\Omega(\hat{x}, t) - \Omega(\bar{x}, t) \leq \bar{\beta}$.

3.2 Stability Analysis

Lemma 1. For any non-linear function $\varphi(\tilde{x}, t)$ that satisfies the Lipschitz condition, there exists a symmetric matrix \tilde{P} such that the function fulfills the following equation:

$$2\tilde{e}^T \tilde{P} [\varphi(\tilde{x}_1, t) - \varphi(\tilde{x}_2, t)] \leq \beta \tilde{e}^T \tilde{P} \tilde{P} \tilde{e} + \tilde{e}^T \tilde{e} \quad (10)$$

where $\tilde{e} = \tilde{x}_1 - \tilde{x}_2$.

Lemma 2. According to Schur complement lemma, matrix inequality $\Phi(x) < 0$, $Q(x) - \Theta(x)\Phi^{-1}(x)\Theta^T(x) < 0$, where $\Phi(x) = \Phi^T(x)$, $Q(x) = Q^T(x)$, can be equivalent to matrix inequality $\begin{bmatrix} Q(x) & \Theta(x) \\ \Theta^T(x) & \Phi(x) \end{bmatrix} < 0$.

Consider the Lyapunov function \hat{V} :

$$\hat{V} = e_x^T P e_x + \frac{1}{2} \bar{\omega} \phi^2 \quad (11)$$

where $\bar{\omega} > 0$, $\bar{\omega}\phi_1 > 2$, and $\tilde{\phi} = \phi^* - \hat{\phi} \geq 0$.

Taking the derivative of (11) yields:

$$\begin{aligned} \dot{\hat{V}} &= \dot{e}_x^T P e_x + e_x^T P \dot{e}_x - \bar{\omega}\phi_1 \tilde{\phi} \|He_y\| \\ &= e_x^T \Gamma e_x + 2e_x^T P M_f v + 2e_x^T P M_d \Omega(\hat{x}, t) \\ &\quad - \Omega(\bar{x}, t) - 2e_x^T P M_f f - \bar{\omega}\phi_1 \tilde{\phi} \|He_y\| \\ &\leq e_x^T (\Gamma + P \bar{\beta}^2 P + I) e_x - (\bar{\omega}\phi_1 - 2)\tilde{\phi} \|He_y\| \\ &\leq e_x^T (\Gamma + P \bar{\beta}^2 P + I) e_x \end{aligned} \quad (12)$$

where $\Gamma = (\bar{A} - L\bar{C})^T P + P(\bar{A} - L\bar{C})$. According to *Lemma 2*, when the following matrix inequality holds:

$$\begin{bmatrix} (\bar{A} - L\bar{C})^T P + P(\bar{A} - L\bar{C}) + I & P \\ P & -\frac{1}{\bar{\beta}^2} I \end{bmatrix} < 0 \quad (13)$$

where $\dot{\hat{V}} < 0$, which means that the state estimation error e_x is asymptotically stable, and \hat{x} gradually approaches \bar{x} . Then, combined with (7), $f \approx v$ can be obtained. Additionally, the utilization of v_{eq} instead of v in practical applications can mitigate buffeting effects induced by v .

$$v_{eq} = -\hat{\phi} \|He_y\| / (\|He_y\| + \xi) \quad (14)$$

where $\xi > 0$ and extremely small.

According to the above analysis, the reconstruction outcome of the fault vector f is denoted as $\tilde{f} \approx v_{eq}$, where $\tilde{f} = [0 \ 0 \ \tilde{f}_{fw}]^T$ and \tilde{f}_{fw} represents the reconstructed result of f_{fw} .

4 Design of SBW FTC Scheme

4.1 RBFNN-RNTSMC based FTC Design Scheme

The previous section presents the application of ASMO for reconstructing sensor fault in the front wheel angle sensor of the SBW system. Consequently, in case of a sensor fault, the actual front wheel angle of the system can be obtained through utilizing the reconstructed sensor fault.

The true front wheel angle $\tilde{\theta}_{fw}$ after \tilde{f}_{fw} compensation is defined as $\tilde{\theta}_{fw} = \theta_{fw} - \tilde{f}_{fw}$. The dynamic equation of the system under the fault condition is formulated as follows:

$$\ddot{\tilde{\theta}}_{fw} = \psi V_{in} + \Psi\left(\tilde{\theta}_{fw}, \dot{\tilde{\theta}}_{fw}\right) + d \quad (15)$$

The tracking error of the front wheel angle is defined as $e_f = \theta_{fwd} - \tilde{\theta}_{fw}$, where θ_{fwd} represents the reference signal for the front wheel angle.

Define the sliding mode surface as follows:

$$s = \sigma + \lambda \varsigma \quad (16)$$

where $\sigma = \dot{e}_f + k e_f + v |e|^{p/q} \operatorname{sign}(e_f)$, $\dot{\varsigma} = |\sigma|^\gamma \operatorname{sign}(\sigma)$, $k > 0$, $v > 0$, $0 < p < q$, $\lambda > 0$, $0 < \gamma < 1$, and $\varsigma(0) = -\lambda^{-1} \sigma(0)$.

Taking the derivative of (16) yields:

$$\dot{s} = \ddot{e}_f + k \dot{e}_f + v \frac{p}{q} |e_f|^{\frac{p}{q}-1} \dot{e}_f + \lambda \dot{\varsigma} \quad (17)$$

When disregarding disturbance and setting $\dot{s} = 0$, in conjunction with (1), the equivalent control term u_{eq} can be derived as follows:

$$U_{eq} = \psi^{-1} \left[\ddot{\theta}_{fw} - \Psi(\theta_{fw}, \dot{\theta}_{fw}) + k \dot{e}_f + v \frac{p}{q} |e_f|^{\frac{p}{q}-1} \dot{e}_f + \lambda \dot{\varsigma} \right] \quad (18)$$

To enhance the robustness of the system against uncertain factors and ensure convergence time, a switching control term U_{sw} is designed as follows:

$$U_{sw} = \psi^{-1} [\mu_1 s + \mu_2 |s|^\eta \operatorname{sign}(s)] \quad (19)$$

where $\mu_1 \in R^+$, $\mu_2 \in R^+$, and $\eta \in R^+$. μ_2 denotes the switch gain.

The control law of RNTSMC can be formulated as u_1 :

$$u_1 = U_{eq} + U_{sw} \\ = \psi^{-1} \left[\ddot{\theta}_{fw} - \Psi(\theta_{fw}, \dot{\theta}_{fw}) + k \dot{e}_f + v \frac{p}{q} |e_f|^{\frac{p}{q}-1} \dot{e}_f + \lambda \dot{\varsigma} + \mu_1 s + \mu_2 |s|^\eta \operatorname{sign}(s) \right] \quad (20)$$

In the presence of uncertain factors, RNTSMC has the theoretical capability to achieve convergence of system error to zero within a finite time. However, the selection of switching gain requires knowledge of the upper bound of the uncertainty. In practical applications, larger gains are typically opted for to enhance system robustness, but this also engenders chattering. In order to mitigate the chattering effect, an uncertainty estimator based on RBFNN is employed in conjunction with RNTSMC to approximate and compensate for uncertain factors, thereby suppressing chattering and enhancing the robustness of system.

The continuous smooth function $F_i(X) : R^m \rightarrow R$, $i = 1, 2, \dots, n$ can be expressed by RBFNN as follows:

$$F(X) = \mathbf{W}^T \mathbf{h} \quad (21)$$

where $X \in \Omega_X \subset R^m$ is the input vector, $\mathbf{W} \in R^p$ is the weight vector, where $p > 1$ is the node number of NN. $\mathbf{h} = \exp\left(\frac{\|X - c_i\|^2}{2b_i}\right)$, $i = 1, 2, \dots, n$ is the basis function, where $c_i = [c_{i,1} \ c_{i,2} \ \dots \ c_{i,n}]^T$ is the center vector of the i th hidden layer, and b_i is the width of the i th hidden layer. Then, the unknown smooth function can be expressed as follows:

$$F(X) = \mathbf{W}^*{}^T \mathbf{h} + \varepsilon \quad (22)$$

where \mathbf{W}^* is the ideal weight vector, ε denotes an approximation error. In general, \mathbf{W}^* and ε are bounded by the positive constants $\bar{\mathbf{W}}$ and $\bar{\varepsilon}$.

Finally, the RBFNN-RNTMC design of the system under the healthy state is denoted as u_{nor} :

$$\begin{aligned} u_{nor} &= u_1 - u_{rbf} \\ &= \psi^{-1} \left[\ddot{\theta}_{fwd} - \Psi(\theta_{fw}, \dot{\theta}_{fw}) + k \dot{e}_f \right. \\ &\quad + v \frac{p}{q} |e_f|^{\frac{p}{q}-1} \dot{e}_f + \dot{\varsigma} + \mu_1 s \\ &\quad \left. + \mu_2 |s|^\eta \text{sign}(s) - \bar{\mathbf{W}}^T \mathbf{h} \right] \end{aligned} \quad (23)$$

4.2 Stability Analysis

Consider the Lyapunov function as follows:

$$V = \frac{1}{2} s^2 - \frac{1}{2} \tilde{\mathbf{W}}^T \delta^{-1} \tilde{\mathbf{W}} \quad (24)$$

where $\tilde{\mathbf{W}} = \bar{\mathbf{W}} - \mathbf{W}^*$ is the estimation error of weight.

Taking the derivative of V and substituting (3), (9) and (10) into (11) yields:

$$\begin{aligned} \dot{V} &= ss - \tilde{\mathbf{W}}^T \delta^{-1} \tilde{\mathbf{W}} \\ &= s \left[\ddot{\theta}_{fwd} - \psi u_{nor} - \Psi(\theta_{fw}, \dot{\theta}_{fw}) \right. \\ &\quad \left. - (\mathbf{W}^*{}^T \mathbf{h} + \varepsilon) + k \dot{e}_f + v \frac{p}{q} |e_f|^{\frac{p}{q}-1} \dot{e}_f + \lambda \dot{\varsigma} \right] - s \tilde{\mathbf{W}}^T \mathbf{h} \\ &= s \left[-\mu_1 s - \mu_2 |s|^\eta \text{sign}(s) + \hat{\mathbf{W}}^T \mathbf{h} \right. \\ &\quad \left. - \mathbf{W}^*{}^T \mathbf{h} - \varepsilon \right] - s \tilde{\mathbf{W}}^T \mathbf{h} \\ &= -\mu_1 s^2 - \mu_2 |s|^{\eta+1} - s \varepsilon \\ &\leq -\mu_1 s^2 - \mu_2 |s|^{\eta+1} - \varepsilon |s| \\ &= -\mu_1 s^2 - |s|^{\eta+1} \left(\mu_2 + \frac{\varepsilon}{|s|^\eta} \right) \end{aligned} \quad (25)$$

According to (25), it can be concluded that $\dot{V} \leq 0$, which indicates the asymptotic stability of the close-loop system

5 Simulation Results

5.1 Simulation Environment

To validate the performance of presented FTC method, the simulation is carried out by MATLAB/Simulink platform and the parameter nominal value of each component in the SBW system is set as: $G_{p1} = 15$ rad, $r_p = 0.08$ m, $G_{p2} = 1.4$ rad/m, $J_s = 2.8e^{-4}$ kg \bullet m², $K_1 = 7$ A/V, $K_2 = 0.097$ Nm/A, $M_p = 3$ kg, $J_{fw} = 0.05$ kg \cdot m², $T_{v,s} = 2.06e^{-4}$ Nm/(rad/s), $T_{s,s} = 0.06$ Nm, $T_{c,s} = 0.059$ Nm, $T_{v,p} = 0.0124$ Nm/(rad/s), $T_{s,p} = 0.69$ Nm, $T_{c,p} = 0.05$ Nm, $T_{v,fw} = 25$ Nm/(rad/s), $T_{s,fw} = 0.7$ Nm, $T_{c,fw} = 0.6$ Nm, $C_s = 1$ Nm/rad, $C_p = 250$ Nm/rad, and $a_s = a_p = a_{fw} = 0.001$. The simulation time and sampling period are set as 30 s and 0.001 s, respectively.

The fault of the front wheel angle sensor is simulated by introducing additive fault f_{fw} into the encoder input module of the front wheel angle sensor. This paper examines two sensor fault conditions, namely additive fault (case 1) and intermittent fault (case 2), which are described as follows:

$$\text{Case 1 : } f_{fw} = \begin{cases} 0, & 0s < t < 15s \\ 50, & 15s \leq t \leq 30s \end{cases} \quad (26)$$

$$\text{Case 2 : } f_{fw} = \begin{cases} 33.33, & 2.5s < t < 3.5s \\ 41.53, & 18s < t < 19s \\ 39.84, & 27s < t < 28s \\ 0, & \text{else} \end{cases} \quad (27)$$

The control law u_{nor} and u_{fit} are characterized by the following parameters: $k = 160$, $v = 10$, $p = 7$, $q = 5$, $\mu_1 = 10$, $\mu_2 = 8$, $\eta = 0.62$, $\lambda = 50$, $\tilde{k} = 155$, $\tilde{v} = 14$, $\tilde{p} = 7$, $\tilde{q} = 5$, $\tilde{\mu}_1 = 10$, $\tilde{\mu}_2 = 9$, $\tilde{\eta} = 1.3$, $\tilde{\lambda} = 42$, $\delta = \tilde{\delta} = \text{diag}\{0.06, 0.06, 0.06, 0.06, 0.06\}$, and the number of neurons in the three layers of RBFNN-RNTSM is 2, 5, and 1, respectively. The linear gain matrix L and H in ASMO are defined as: $L = [26.32 \ 0.92 \ 97.65 \ -712.86 \ 26.48 \ 164.28]^T$, $H = [6.54 \ 0.78 \ 0.78 \ 6.42]$.

In addition, in order to verify the fault tolerance performance of the AFTC method based on RBFNN-RNTSMC, the PID controller is used for comparison.

5.2 Simulation Results

The simulation results under case 1 are depicted in Figs. 1, 2, and 3. Among them, Fig. 1 illustrates the fault reconstruction results, where \tilde{f}_{fw} represents the output of ASMO. As depicted in the figure, ASMO exhibits prompt completion of fault reconstruction when a fault occurs, with the reconstructed value being essentially consistent with the fault value. To validate the superiority method proposed in this article, PID controller is employed for simulation comparison. Figure 2 shows the front wheel angle tracking results achieved through AFTC, while Fig. 3 presents the front wheel angle tracking error under fault conditions.

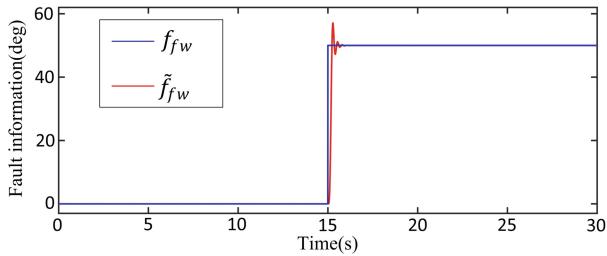


Fig. 1. Fault information of case 1.

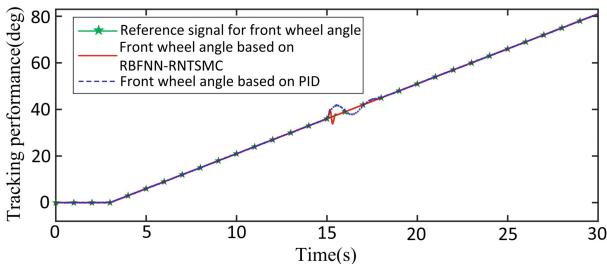


Fig. 2. Tracking preformance of case 1.

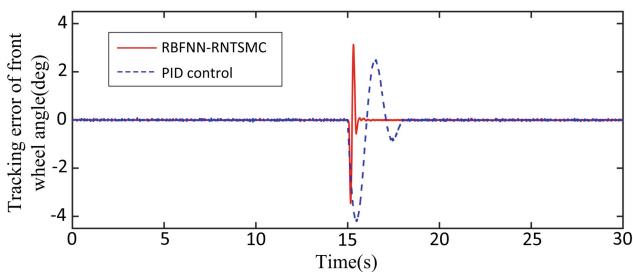


Fig. 3. Tracking error of case 1.

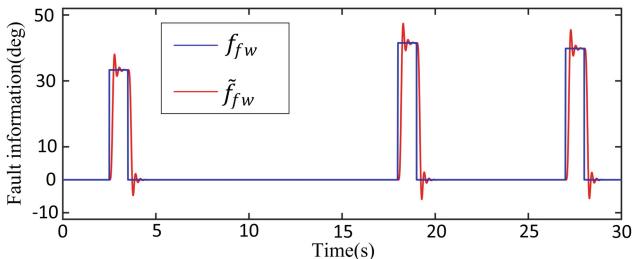


Fig. 4. Fault information of case 2.

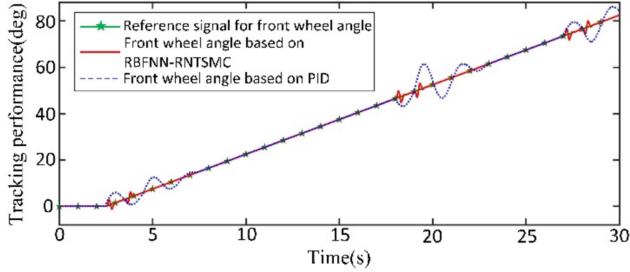


Fig. 5. Tracking preformation of case 2.

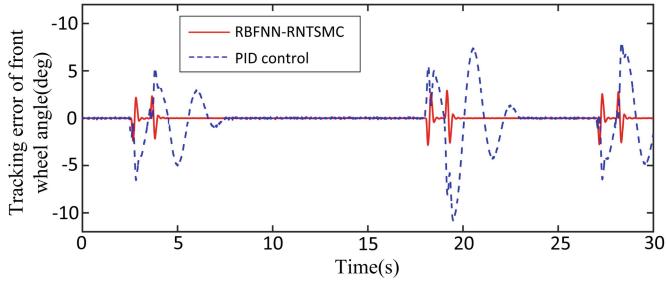


Fig. 6. Tracking error of case 2.

The simulation results for case 2 are presented in Figs. 4, 5, and 6. Figure 4 shows the fault reconstruction outcomes, demonstrating ASMO can rapidly and accurately complete fault reconstruction. Figs. 5 and 6 illustrate the front wheel angle tracking results and the corresponding tracking error, respectively.

Overall, the AAFTC method based on RBFNN- RNTSMC exhibits robust fault tolerance in addressing single additive sensor fault as well as intermittent sensor fault. At the same time, compared with the traditional PID based AFTC, the proposed controller shows better tracking performance in both accuracy and speed.

6 Conclusion

The article proposes an AAFTC method for SBW systems. Firstly, a dynamic model of the SBW system is constructed and a tracking control strategy based on RBFNN- RNTSMC is further proposed to achieve accurate tracking of front wheel steering angle under normal system conditions. In case of sensor fault, ASMO is introduced for sensor fault reconstruction and adaptive active fault-tolerant control laws are designed based on the reconstruction results. Meanwhile, real-time switching control laws are implemented guided by FDI results. Finally, through simulation analysis, the feasibility of this active fault-tolerant control method under sensor fault conditions is verified.

References

1. Perozzi, G., Rath, J.J., Sentouh, C., Floris, J., Popieul, J.C.: Lateral shared sliding mode control for lane keeping assist system in steer-by-wire vehicles: theory and experiments. *IEEE Trans. Intell. Veh.* **8**(4), 3073–3082 (2023)
2. Chen, L., Tang, L.: Yaw stability control for steer-by-wire vehicle based on radial basis network and terminal sliding mode theory. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **237**(8), 2036–2048 (2022)
3. Shi, Q., Wei, Y., Xie, D., Li, F., Song, K., He, L.: A model predictive control approach for angle tracking of steer-by-wire system with nonlinear transmission ratio. *Asian J. Control.* **25**(2), 1156–1166 (2023)
4. Zhang, Y., Jiang, J.: Bibliographical review on reconfigurable fault-tolerant control systems. *Annu. Rev. Control.* **32**(2), 229–252 (2008)
5. Merheb, A.R., Noura, H., Bateman, F.: Design of passive fault-tolerant controllers of a quadrotor based on sliding mode theory. *Int. J. Appl. Math. Comput. Sci.* **25**(3), 561–576 (2015)
6. Yang, P., Liu, Z., Li, D., Zhang, Z., Wang, Z.: Sliding mode predictive active fault-tolerant control method for discrete multi-fault system. *Int. J. Control. Autom. Syst.* **19**(3), 1228–1240 (2021)
7. Zeghlache, S., Benslimane, T., Bouguerra, A.: Active fault tolerant control based on interval type-2 fuzzy sliding mode controller and non linear adaptive observer for 3-DOF laboratory helicopter. *ISA Trans.* **70**, 280–303 (2017)
8. Mortazavizadeh, S.A., Ghader, A., Hajian, M., Ebrahimi, M.: A fault-tolerant steer-by-wire system based on angular position estimation. *Electr. Eng.* **103**(5), 2403–2414 (2021)
9. Huang, C., Naghdly, F., Du, D.: Observer-based fault-tolerant controller for uncertain steer-by-wire systems using the delta operator. *IEEE-ASME Trans. Mechatron.* **23**(6), 2587–2598 (2018)
10. Tan, C., Edwards, C.: Sliding mode observers for detection and reconstruction of sensor fault. *Automatica.* **38**(10), 1815–1821 (2002)



A Two-Stage Optimization Algorithm for the Multi-dropoff Flying Sidekick Traveling Salesman Problem

Hongrui Ao¹, Jihui Zhang^{1,2(✉)}, and Shengzong Chen³

¹ School of Automation, Qingdao University, Qingdao 266071, China
zhangjihui@qdu.edu.cn

² Shandong Key Laboratory of Industrial Control Technology, Qingdao 266071, China

³ School of Management, Beihang University, 17, Beijing 100191, China

Abstract. The drone-truck collaborative delivery system has emerged as a promising logistics paradigm with significant potential for improving efficiency. This study formulates a mixed integer programming model for the multi-dropoff flying sidekick traveling salesman problem (MD-FSTSP) to minimize total delivery time. We propose a two-stage optimization algorithm featuring: (1) a list-based adaptive temperature control strategy that improves upon traditional large neighborhood search by enhancing parameter attenuation in generating initial routes; and (2) a drone range-constrained segmentation strategy that decomposes global paths into feasible sub-paths with local search refinements. Experimental validation using TSPLIB benchmark data and parameter sensitivity analysis shows its superior solution accuracy and convergence speed compared to the existing methods.

Keywords: Multi-dropoff flying sidekick traveling salesman problem · Segmentation strategy · List-based adaptive large neighborhood search algorithm

1 Introduction

The rapid growth of e-commerce and logistics has made the efficient management of terminals a crucial factor for achieving a competitive advantage. Although conventional truck delivery has advantages in terms of payload and cost, it faces significant limitations [1]. However, drones' inherent limitations, such as limited endurance (less than 30 min) and payload capacity (less than 5 kg), prevent them from being used independently on a large scale. The drone-truck collaborative model [2, 3] effectively combines the freight capacity of trucks with the mobility of drones, creating an intelligent logistics network that is particularly well-suited to urban last-mile and remote-area deliveries.

In the seminal work [4], Murray and Chu [4] developed two models: the Flying Sidekick Traveling Salesman Problem (FSTSP) and the Parallel Drone Scheduling Traveling Salesman Problem (PDSTSP). The FSTSP involves a truck transporting and deploying a drone for the purpose of coordinated parcel delivery, whereas the PDSTSP utilizes drones operating autonomously from distribution centers, independently of trucks. To

optimize the FSTSP, the authors developed a two-phase heuristic approach: first optimizing truck routes, then incorporating drone-specific operational constraints. Agatz et al. [5] advanced the field by introducing the Traveling Salesman Problem with Drones (TSP-D), which allowed for multiple drones launches and recoveries at individual customer nodes, yielding substantial efficiency gains. These pioneering contributions have formed the theoretical basis for numerous subsequent investigations [6–9], catalyzing ongoing model refinements and practical applications.

However, existing research predominantly assumes a single-dropoff delivery mode, in which drones serve only one customer per flight. As drone technology advances and delivery demands intensify, the limitations of single-dropoff models have become increasingly apparent, prompting a shift in research toward multi-dropoff delivery systems. Gonzalez et al. [10] were among the first to take this approach to extend the FSTSP framework to multi-dropoff flights. Manyam et al. [11] subsequently formulated the problem as a mixed-integer linear program and used a branch-and-cut approach to achieve optimal solutions for 20-node instances within 6 h. Wang et al. [12] proposed a capacitated joint delivery model. Nevertheless, exact methods [11, 12] are computationally intractable as the scale of the problem increases, making heuristic approaches [13–19] more practical due to their superior efficiency and flexibility.

Recent heuristic developments have significantly advanced multi-dropoff optimization. Zhang et al. [13] adopted a decomposition strategy, solving the problem through a column generation-based heuristic. Liu et al. [14] developed hybrid heuristics for joint drone-truck routin. Luo et al. [15] integrated urban road network constraints into a novel two-stage algorithm. Liu et al. [16] implemented a simulated annealing algorithm enhanced with nearest-neighbor initialization and local search. Setyo et al. [17] contributed an adaptive large neighborhood search (ALNS) specifically tailored for multi-dropoff optimization. Despite these advancements, heuristic methods continue to face scalability limitations.

The MD-FSTSP combines the classical assignment problem with the traveling salesman problem (TSP). However, most existing studies solve it holistically rather than through phased optimization. Given the well-established solution frameworks for the TSP, this study uses high-quality TSP paths as initial solutions for the MD-FSTSP. We have developed a two-stage optimization algorithm that optimizes truck routing and drone delivery strategies sequentially, thereby improving overall distribution efficiency.

The main contributions of this paper are twofold: (1) a List-Based Adaptive Large Neighborhood Search (LBALNS) algorithm using a new list-managed cooling mechanism to control the decay of the temperature parameter; and (2) a segment-based perturbation strategy combing a local search operator to reduce the number of infeasible solutions and to accelerates convergence.

2 Problem Description

Given a set of customers each requiring a delivery service by the truck or drone. The truck and drone depart from the depot together and serve different customers separately. They return the depot synchronously after finishing all service. The drone may launch or land the truck several times at any customer node during the process. At meeting

points (some of the customer nodes), truck service is mandatory. The drone must serve its customers sequentially without violating its maximum capacity before meeting the truck. The objective is to make a service plan to minimize the makespan.

The main assumptions are listed as follows.

- (1) Service uniqueness: each customer is served exactly once, either by the truck or the drone, but not by both.
- (2) Truck capability: the truck has no capacity limit thus can serve all customers independently.
- (3) Meeting constraints: the meetings may only occur at customer nodes or the depot. Battery replacement and package loading are performed exclusively at these nodes. If either vehicle arrives earlier, it must wait for the other to synchronize operations.
- (4) Drone operations: no restrictions are imposed on drone launch or retrieval operations. The drone's battery is fully recharged prior to each launch.
- (5) Multi-customer sorties: provided sufficient battery capacity, the drone may serve multiple customers per sortie before meeting with the truck at a designated customer node.
- (6) Battery conservation: the drone consumes no battery power while being transported by the truck.
- (7) Distance metric: all inter-node distances are computed using the Euclidean metric.
- (8) Energy model: drone energy consumption follows a linear model, excluding auxiliary costs associated with takeoff, landing, or payload delivery.
- (9) Speed invariance: both vehicles maintain constant speeds during operation.

The problem can be formally modeled using a graph $G = (V, A)$, where V represents the node set comprising both the depot and customer locations; A denotes the edge set containing all arcs connecting any two nodes in V ; N constitutes the customer set with a total of n customers. Each customer must be exclusively served by either the truck or the drone. The collaborative delivery system operates under five distinct operational modes (see Fig. 1), with the drone independently servicing customers in all but the first mode. This requirement mandates strict compliance with the maximum range constraint specified in Eq. (1).

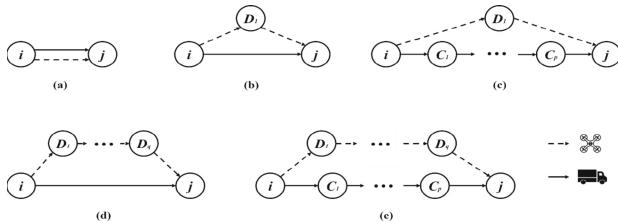


Fig. 1. The five basic operational modes for MD-FSTSP.

This problem can be formally defined as selecting an optimal subset of feasible joint delivery route segments from all possible combinations that satisfies two key requirements: Eq. (1) complete network node coverage, and Eq. (2) total delivery makespan

minimization. The computational method for determining each segment's delivery time is explicitly given in Eqs. (2), (3), and (4). The total service duration required for delivery task OP_k is denoted by t_k . T_k denotes the sequence of nodes visited by the truck in OP_k . D_k denotes the sequence of nodes serviced by the drone in OP_k . C represents the set of nodes visited by the truck's route excluding launch and retrieval nodes. $t(T_k)$ denotes the time required for truck delivery operations within the specified segment. $t(D_k)$ denotes the time required for drone delivery operations within the specified segment. $d_{i \rightarrow j}^T$ and $t_{i \rightarrow j}^T$ denote the travel distance and travel time, respectively, for the drone's movement from customer i to customer j . S_L and S_R denote the preparatory time required for package loading prior to drone launch and battery replacement following drone retrieval, respectively.

$$\begin{cases} d_{i \rightarrow D_1}^D + d_{D_1 \rightarrow j}^D \leq L_{\max}, & \text{if } q = 1 \\ d_{i \rightarrow D_1}^D + d_{D_1 \rightarrow j}^D + \sum_{i=1}^{q-1} d_{D_i \rightarrow D_{i+1}}^D \leq L_E, & \text{if } q > 1 \end{cases} \quad (1)$$

$$t_k = t(OP_k) = \begin{cases} \max\{t(T_k), t(D_k)\}, & \text{if } p \geq 1 \\ t(T_k), & \text{otherwise} \end{cases} \quad (2)$$

$$t(D_k) = \begin{cases} s_L + t_{i \rightarrow D_1}^D + \sum_{i=1}^{p-1} t_{D_i \rightarrow D_{i+1}}^D + t_{D_q \rightarrow j}^D + s_R, & \text{if } q > 1 \\ s_L + t_{i \rightarrow D_1}^D + t_{D_1 \rightarrow j}^D + s_R, & \text{otherwise} \end{cases} \quad (3)$$

$$t(T_k) = \begin{cases} s_L + t_{i \rightarrow C_1}^T + t_{C_1 \rightarrow j}^T + s_R, & \text{if } D \neq \emptyset \text{ and } p = 1 \\ s_L + t_{i \rightarrow C_1}^T + \sum_{i=1}^{p-1} t_{C_i \rightarrow C_{i+1}}^T + t_{C_p \rightarrow j}^T + s_R, & \text{if } D \neq \emptyset \text{ and } p > 1 \\ t_{i \rightarrow j}^T, & \text{otherwise} \end{cases} \quad (4)$$

Let $K = \{OP_1, \dots, OP_{|k|}\}$ be the set of $|k|$ feasible truck-drone collaborative delivery route segments; $K(v)$ the set of all feasible route segments containing nodes $v \in V$; $K^-(v)$ the set of feasible route segments with origin/launch nodes $v \in V$; $K^+(v)$ the set of feasible route segments with terminal/meeting nodes $v \in V$. For each node subset $S \subset V$, let $K^+(S) \subset K$ be the set of route segments whose terminal nodes are contained in S . $K^T(v)$ is the set of feasible route segments where nodes $v \in V$ is served by the truck. $K^D(v)$ denotes the set of feasible route segments where nodes $v \in V$ is served by the drone. The model additionally incorporates two binary decision variables, x_k and Y_v . Y_v is a binary indicator variable that denotes whether node v is selected as a launch node in at least one route segment. x_k is defined as a binary decision variable that equals 1 if the route segment is selected in OP_k , and 0 otherwise. The complete mathematical formulation is presented as follows.

$$\min M = \sum_{k \in K} t_k x_k \quad (5)$$

$$\sum_{k \in K(v)} x_k \geq 1, \quad \forall v \in N \quad (6)$$

$$\sum_{k \in K^D(v)} x_k \leq 1, \quad \forall v \in N \quad (7)$$

$$\sum_{k \in K^T(v)} x_k \leq 2 \left(1 - \sum_{k \in K^D(v)} x_k \right), \quad \forall v \in N \quad (8)$$

$$\sum_{k \in K^+(v)} x_k \leq nY_v, \quad \forall v \in V \quad (9)$$

$$\sum_{k \in K^+(v)} x_k = \sum_{k \in K^-(v)} x_k, \quad \forall v \in V \quad (10)$$

$$\sum_{k \in K^+(0)} x_k \geq 1 \quad (11)$$

$$\sum_{k \in K^+(S)} x_k \geq Y, \quad \forall S \subset N, v \in S \quad (12)$$

$$Y_0 = 1 \quad (13)$$

$$x_k \in \{0, 1\}, \quad \forall k \in K \quad (14)$$

$$Y_v \in \{0, 1\}, \quad \forall v \in V \quad (15)$$

The objective function (5) minimizes the total delivery time. Constraint (6) ensures service coverage for all customer nodes. Constraint (7) restricts drone visits to at most one occurrence per node in N . Constraint (8) ensures mutual exclusivity between drone and truck services for each customer node, prohibiting duplicate servicing. Constraint (9) establishes path continuity by permitting node v to serve as the origin of subsequent route segments if it terminates any existing path. Constraint (10) enforces path concatenation by requiring any selected path terminating at node v to be immediately succeeded by another path originating from v . Constraint (11) restricts the number of paths terminating at the depot node to exactly one. Constraint (12) maintains route connectivity while preventing revisiting of customer nodes by the truck. Constraint (13) designates node 0 as the mandatory starting point. Constraints (14) and (15) formally define the decision variables' domain and properties.

3 Solution Approach

3.1 Solution Representation

The dual-layer encoding scheme proposed by Gonzalez et al. [10] is adopted. The first encoding layer (AN) employs integer values to denote a TSP route. The second layer (AM) specifies service modes for customer nodes: 1 means service only by drones, 2 represents service only by trucks, and 0 describes the movement process of drones carried by trucks.

3.2 List-Based Adaptive Large Neighborhood Search

The algorithm proceeds as follows.

Step1: Initialization: Generate the initial temperature list and the initial solution S ; Set the initial parameters (the number of inner cycles N , the total number of outer cycles M , and the length of the temperature list L).

- Step2: Find the maximum temperature T_{\max} in the temperature list.
- Step3: Select a set of remove-repair operators from the removal operator set and the repair operator set with a probability, and use them to perturb the current solution S to obtain a new one S' which is accepted as the current solution according to the Metropolis criterion. The acceptance probability p is calculated by Eq. (16). If the accepted new solution is worse than the current one, calculate $t_i = -(f(S') - f(S))/\ln(r_i)$ simultaneously where r_i is a uniform random number in (0,1).
- Step4: check the inner loop termination condition. If not, return to Step3. Otherwise, replace the highest temperature in the thermometer with the average temperature $t_{avg} = \sum_{i=1}^c t_i/c$ of t_i in the current inner loop, and then update the probability of the operator being selected based on the number of uses and scores of each operator in this inner loop.
- Step5: check the out loop termination condition. If the condition is satisfied, end and output the historical optimal solution, otherwise, return to step2.

$$p = \begin{cases} 1, & \text{if } f(S') < f(S) \\ \exp(-(f(S') - f(S))/T), & \text{otherwise} \end{cases} \quad (16)$$

Different from ref. [18], the proposed LBALNS algorithm adopts an adaptive initialization strategy: instead of relying on the parameter p_0 , it generates a feasible initial solution and uses the objective function difference from remove-repair perturbations to derive initial temperature values, with the list size set to $3L$ and selects the top L highest temperatures to form the initial temperature list [19].

In the LBALNS, an adaptive operator selection mechanism is implemented through weighted roulette wheel selection to optimally balance solution quality and computational efficiency. All removal, repair operators and adaptive probability control strategy in this paper are the same as the improved ALNS [20].

3.3 Segmentation Strategy and Local Search Operators

Keeping the capacity of the drone (such as its maximum visiting range) in mind, customers with long distances are given priority allocation to the drone, and neighboring customer segments are added to form continuous service segments for the drone. Then, the second layer encoding of customer nodes within the service segment is adjusted through local perturbations. The specific steps are as follows.

- (1) Traverse each customer node not in the tabu list, assuming the currently selected customer is j , find its previous customer i and next customer k , and calculate the distance $savings = d_{ij} + d_{jk} - d_{ik}$ after assigning the current customer to the drone.
- (2) Sort and assign the customer j with the highest $savings$ value to the drone. During this process, customer i serves as the takeoff point for the drone, while k serves as the landing point. Then, the allocation scheme is calculated to determine if it meets the drone's endurance range. If it does, customer j is assigned to the drone, and then proceed to step (3), otherwise add customer j to the tabu list, and go to step (5).
- (3) To optimize the path of the drone, calculate the total objective function value after assigning takeoff or landing points to the drone separately. Under the range limit of

the drone, if the takeoff or landing node is not in the tabu list, add the node with smallest objective value to the customer list of drone services. At the same time, add one node before (after) the node as the takeoff (landing) point. This process terminates when either the endurance limit is reached or both adjacent nodes in the current segment are in the tabu list.

- (4) For the current service segment, local search operators are used to perturb the second layer encoding, only accepting better solutions until the upper limit of perturbation times is reached. After completing the perturbation, if the optimal solution obtained during the process is better than the original solution before segmentation, it is set as the current solution, end the optimization process of this service segment, and add the service segment (including takeoff and landing points) to the tabu list; otherwise, add customer j to the tabu list.
- (5) If the termination condition is satisfied (all client nodes exist in the tabu list or the maximal iteration G is reached), the algorithm will exit and output the global optimal solution. Otherwise, return (1).

The local search operators are specifically optimized and adjusted for the second-layer coding structure of MD-FSTSP. Based on the characteristics of the problem, after processing the path segmentation, the set of customer points within each sub-segment is served only by trucks or drones alone. Therefore, the solution space of each sub-segment can be modeled as a binary sequence (that is, a discrete code containing only 1 and 2, where 1 represents truck service and 2 represents drone service). Three classical neighborhood operation operators are adopted to perturb and optimize the current solution. The specific operation is shown in Fig. 2.

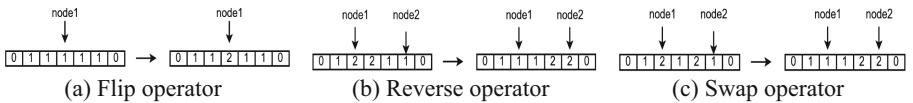


Fig. 2. Local search operators

4 Experiments

The experimental parameter settings are as follows: $N = 70$, $M = 150$, the speed of the truck is $v_1 = 1$, the speed of the drone is $v_2 = 1.5$. The time for loading packages and replacing batteries on the drone is 5. In each neighborhood search, the number of nodes removed is $[0.25num, num]$, where the upper limit parameter num for the number of removed nodes. The number of local perturbations K , and the maximal iterations in the segmentation G vary with the size of the instance, are shown in Table 1.

4.1 Sensitivity Analysis of the Length of the Temperature List

To analyze the sensitivity on the parameter L , 11 different parameter values (with a tolerance of 2 and a range of 20 to 40) are selected for testing. On four classic instances in the

Table 1. Partial parameter settings for LBALNS.

parameters	instance size			parameters	instance size		
	$n \leq 200$	$200 < n \leq 1000$	$1000 < n$		$n \leq 200$	$200 < n \leq 1000$	$1000 < n$
<i>num</i>	20	0.1n	0.1n	<i>K</i>	10n	10n	10n
<i>L</i>	20	30	25	<i>G</i>	100	100	100

standard test database TSPLIB, 20 experiments are conducted for each parameter value, and the average deviation rate error defined by Eq. (17) is recorded. The experimental results are shown in Fig. 3.

$$Error = \frac{\text{average} - \text{optimal}}{\text{optimal}} \times 100\% \quad (17)$$

Results indicate LBALNS's list length sensitivity scales with problem size, remaining negligible for small instances but substantial for large ones (see Fig. 4). While increased parameters enable thorough exploration through higher initial temperatures and gradual cooling, they impair final convergence precision and computational efficiency due to prolonged high-temperature phases accepting inferior solutions.

4.2 Performance Testing and Analysis of LBALNS Algorithm

To comprehensively evaluate the solving performance of LBALNS, experiments were conducted independently 20 times on classical TSP test instances. Experimental results (Table 2) demonstrate that LBALNS algorithm can quickly converge to the optimal solution when solving small and medium-scale TSP. For large-scale TSP, it can also converge to a satisfactory solution within an acceptable time. Successful processing of instances exceeding 5000 nodes confirms its scalability for large TSPs.

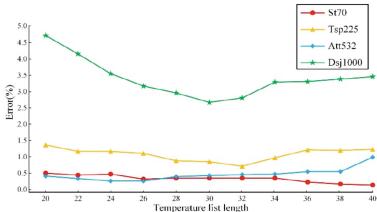


Fig. 3. The influence of parameter *L* on the solution effect.

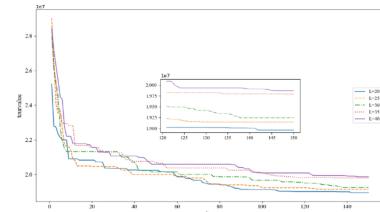


Fig. 4. The influence of parameter *L* on the temperature parameter.

Table 2. Results of TSP instances solved by LBALNS algorithm.

Instance	OPT	Best	Mean	Time(s)
St70	675	675	675.8	2.6
Eil101	629	629	632.7	3.7
Pr124	59030	59030	59030	4.9
KroA200	29368	29368	29420	11.2
Att532	87686	87455	87907	69.3
RL5915	565530	598090	604117	5506

4.3 Performance Test and Analysis of Two Stage Algorithm

Our experimental evaluation using TSPLIB instances (with drone range set to 10% of optimal TSP distances) demonstrated significant efficiency gains (Table 3) with an average 30% time reduction.

Table 3. Results of the algorithm used in this paper to solve MD-FSTSP case studies (20 runs).

Instance	Initial average delivery time of trucks	The shortest time for joint delivery	Time saving percentage	Average time for joint delivery	Time saving percentage	Running time(s)
Eil51	426.3	376.7	11.6%	379	11.1%	2.1
St70	675.8	542.4	19.7%	547	19.1%	2.9
Eil76	538	432.6	19.6%	441	18.0%	3.4
Eil101	632.7	467.3	26.1%	471	25.6%	4.2
KroA150	26198	16285	37.8%	16680	36.3%	8.1
A280	2619.6	1641	37.4%	1651	37.0%	24.9

5 Conclusion

In this paper, we present a two-stage optimization framework for the MD-FSTSP. The first stage introduces LBALNS, a novel list-based adaptive large neighborhood search algorithm that dynamically manages temperature parameters through real-time updates, enhancing both solution stability and computational efficiency while minimizing manual parameter tuning. The second stage develops a segmentation strategy that accounts for drone range limitations, decomposing large-scale problems into tractable sub-problems through intelligent task allocation and local search optimization. Experimental results demonstrate the algorithm's effectiveness in handling complex delivery scenarios with reduced computational complexity.

Acknowledgement. This research is supported by the Natural Science Foundation of China [grant numbers 61673228, 62072260], the Natural Science Foundation of Shandong Province [grant number ZR2020MF094] and the Science Technology Program of Qingdao [grant number 21-1-2-16-zhz].

References

1. Moshref-Javadi, M., Winkenbach, M.: Applications and research avenues for drone-based models in logistics: a classification and review. *Expert Syst. Appl.* **177**, 114854 (2021)
2. Chung, S.H., Sah, B., Lee, J.: Optimization for drone and drone-truck combined operations: a review of the state of the art and future directions. *Comput. Oper. Res.* **123**, 105004 (2020)
3. Ren, X., Huang, H., Yu, S.W., et al.: Review on vehicle-UAV combined delivery problem. *Control Decis.* **36**(10), 2313–2327 (2021)
4. Murray, C.C., Chu, A.G.: The flying sidekick traveling salesman problem: optimization of drone-assisted parcel delivery. *Transp. Res. Part C Emerg. Technol.* **54**, 86–109 (2015)
5. Agatz, N., Bouman, P., Schmidt, M.: Optimization approaches for the traveling salesman problem with drone. *Transp. Sci.* **52**(4), 965–981 (2018)
6. Jia, Z.H., Wang, S.G., Liu, C.: Vehicle and drones joint distribution model and optimization algorithm in multi-mode. *Control Decis.* **39**(07), 2125–2132 (2024)
7. Freitas, J.C., Penna, P.H.V., Toffolo, T.A.M.: Exact and heuristic approaches to truck–drone delivery problems. *EURO J. Transp. Logist.* **12**, 100094 (2023)
8. Yan, R., Chen, L.S., Zhu, X.N., et al.: Research on truck-drone collaborative routing problem with regional restrictions. *Chin. J. Manag. Sci.* **30**(05), 144–155 (2022)
9. Yurek, E.E., Ozmutlu, H.C.: A decomposition-based iterative optimization algorithm for traveling salesman problem with drone. *Transp. Res. Part C Emerg. Technol.* **91**, 249–262 (2018)
10. Gonzalez-R, P.L., Canca, D., Andrade-Pineda, J.L., et al.: Truck-drone team logistics: a heuristic approach to multi-drop route planning. *Transp. Res. Part C Emerg. Technol.* **114**, 657–680 (2020)
11. Manyam, S.G., Sundar, K., Casbeer, D.W.: Cooperative routing for an air–ground vehicle team—exact algorithm, transformation method, and heuristics. *IEEE Trans. Autom. Sci. Eng.* **17**(1), 537–547 (2019)
12. Wang, Z., Sheu, J.B.: Vehicle routing problem with drones. *Transp. Res. B Methodol.* **122**, 350–364 (2019)
13. Zhang, G., Zhu, N., Ma, S., et al.: Humanitarian relief network assessment using collaborative truck-and-drone system. *Transp. Res. Part E Logist. Transp. Rev.* **152**, 102417 (2021)
14. Liu, Y., Shi, J., Liu, Z., et al.: Two-layer routing for high-voltage powerline inspection by cooperated ground vehicle and drone. *Energies.* **12**(7), 1385 (2019)
15. Luo, H., Zhang, P., Wang, J., et al.: Traffic patrolling routing problem with drones in an urban road system. *Sensors.* **19**(23), 5164 (2019)
16. Liu, Y., Liu, Z., Shi, J., et al.: Two-echelon routing problem for parcel delivery by cooperated truck and drone. *IEEE Trans. Syst. Man Cybern. Syst.* **51**(12), 7450–7465 (2020)
17. Mara, S.T.W., Rifai, A.P., Sophya, B.M.: An adaptive large neighborhood search heuristic for the flying sidekick traveling salesman problem with multiple drops. *Expert Syst. Appl.* **205**, 117647 (2022)
18. Zhan, S., Lin, J., Zhang, Z., et al.: List-based simulated annealing algorithm for traveling salesman problem. *Comput. Intell. Neurosci.* **2016**(1), 1712630 (2016)

19. Chen, S.Z., Zhang, J.H., et al.: A simulated annealing algorithm with wave temperature control for the traveling salesman problem. *Control Decis.* **38**(04), 911–920 (2023)
20. Ao, H.R., Zhang, J.H., Chen, S.Z.: An improved adaptive large neighborhood search algorithm and its application to the traveling salesman problem. *Appl. Res. Comput.* **42**(6), 1713–1718 (2025)



Optimization of Vibration Suppression for a Humanoid Robot's Lower Limb

Jie Yang^{1(✉)}, Jisong Ding^{1,2}, Yiqun Peng¹, and Tianyu Liu¹

¹ Jianghuai Advance Technology Center, Hefei 230000, China

m13063499211@163.com

² Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230000, China

Abstract. Humanoid robots mimic human-like movements by emulating the human body's morphology, with a single lower limb typically configured with five or more degrees of freedom (DOF) to execute complex motions. However, vibration coupling induced by multi-motor coordination can lead to forced vibrations or even resonance in structural components, significantly increasing the complexity of vibration analysis. This study employs the finite element method (FEM) to construct a single lower limb model of a humanoid robot. Modal analysis is performed to extract the lower limb's natural frequencies and mode shapes, while harmonic response analysis evaluates amplitude variations at critical regions under different excitation frequencies. Further optimization via material selection is conducted on robotic components, targeting modifiable regions—including hip joints, thigh linkages, and calf structures—to mitigate vibrational behavior. The results provide actionable guidance for enhancing vibration suppression capabilities of lower limb, offering hardware-level support for high-speed motion stability and operational precision in complex environments.

Keywords: Humanoid robots · Vibration suppression · FEM · Modal analysis

1 Introduction

Humanoid robots serve as the core embodiment of embodied intelligence. Their anthropomorphic structural characteristics enable superior adaptability to human society compared to traditional industrial robots confined to fixed workstations, allowing them to thrive in unstructured environments with broader applicability. A stable body structure is critical for supporting complex motions such as running and jumping, during which robotic components cannot be treated as rigid bodies, as vibrations may occur. As application scenarios expand, rapid and precise manipulation will become essential skills for humanoid robots. Concurrently, the trend toward structural lightweighting and high load capacity further amplifies the risk of vibration [1].

Existing vibration analyses predominantly focus on industrial robots, where inertial effects from rapid motion are the primary cause of vibration [2]. Vibration suppression methods fall into two categories: passive and active control. Some studies passively suppress vibrations by adding constrained damping layers to structures to increase damping

ratios [3, 4], though this complicates motion control. For active control, researchers have explored trajectory optimization in motion planning [5, 6] and input shaping techniques to mitigate residual vibrations [7, 8]. However, these methods rely on idealized dynamic models with linear parameters, which may diverge from real-world complex structures, leading to reduced generalization efficacy.

This study focuses on humanoid robot which employs finite element simulations to conduct modal analysis on a complete lower-limb model, extracting its natural frequencies and mode shapes. By analyzing motors' current data, the actual output frequency during operation is identified as vibrational input for harmonic response analysis, evaluating lower-limb vibration characteristics. Finally, structural material modifications at multiple locations are tested to assess vibration responses, providing practical optimization recommendations for vibration suppression.

2 Modal Analysis

2.1 Model

The lower limb of the humanoid robot studied in this work comprises distinct thigh, calf, and foot segments, along with joint connectors at the hip. Five rotary motors provide five degrees of freedom (DOFs) to enable human-like walking. In the finite element model illustrated in Fig. 1, critical components such as the two joint connectors, thigh, calf, ankle, foot, three linkages, five motors, all bearings, and mass-significant parts are retained, while minor components (e.g., wires) are reasonably simplified. Materials include 7075 aluminum alloy for the joint connectors, thigh, calf, ankle, and foot; 40Cr steel for the linkages; and non-load-bearing structural components replaced with mass blocks to preserve weight and center of gravity.

For modal analysis, interactions between component contact surfaces are defined as initial frictionless contacts and bonded connections. The rotor and stator of each motor are modeled separately to accurately simulate their interactions. Boundary conditions constrain all six degrees of freedom at the top of the hip joint connector, as depicted in Fig. 2. A meshing strategy balancing computational efficiency and accuracy is adopted (Fig. 3). Modal analysis is performed based on this model.

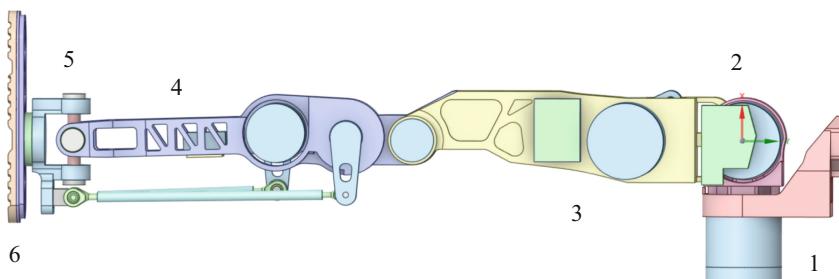


Fig. 1. The finite element model of a humanoid robot lower limb including six structural parts: 1. joint connect #1, 2. joint connect #2, 3. thigh, 4. calf, 5. ankle, 6. foot.

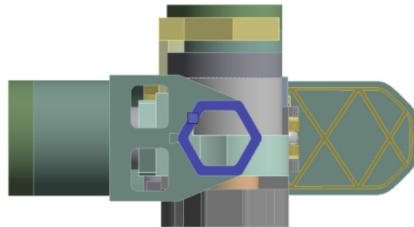


Fig. 2. Schematic Diagram of Constraint Application Locations in the Model.

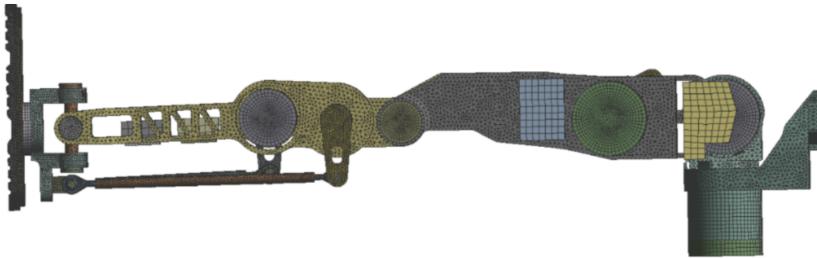


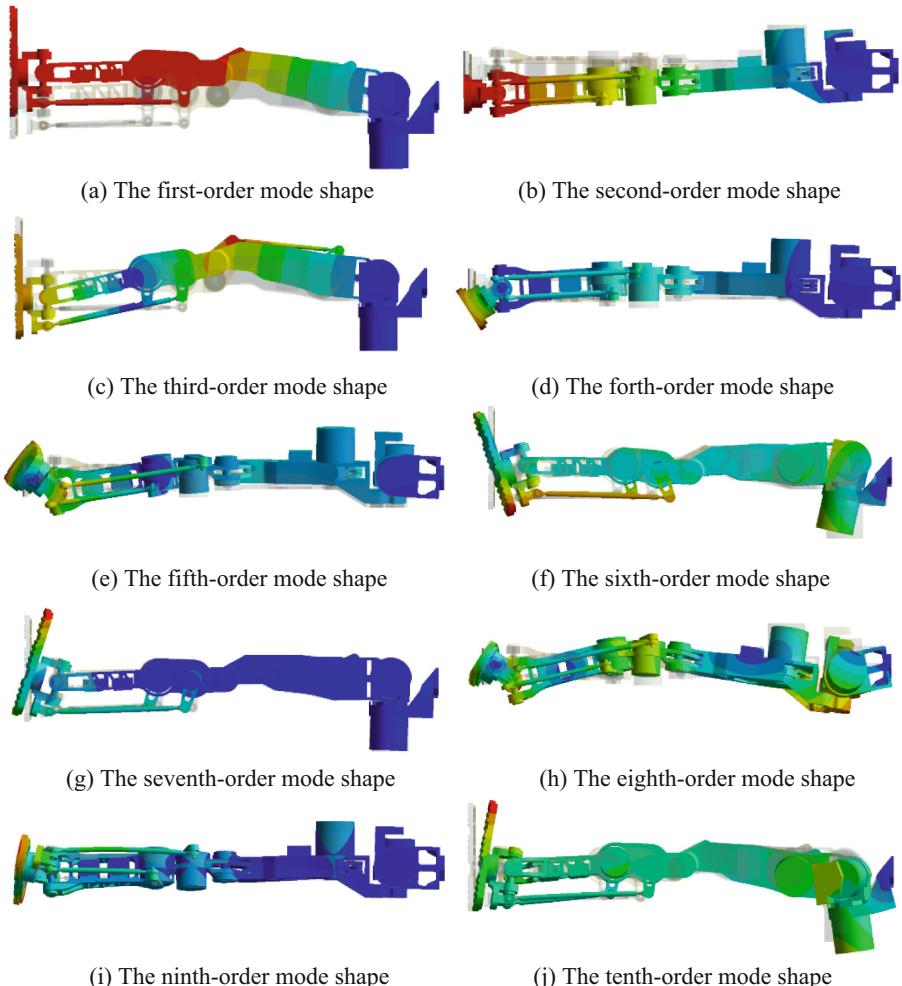
Fig. 3. Schematic Diagram of Model Mesh Generation.

2.2 Results

The modal analysis prioritized the first 10 natural frequencies (Table 1) based on their effective mass contributions across six degrees of freedom directions, with corresponding mode shapes depicted in Fig. 4. The analysis identified distinct modal behaviors: Mode 1 featured thigh rotation about Joint Connector #2, Mode 2 involved synchronized rotation of Joint Connector #2 and its substructure about Joint Connector #1, Mode 3 combined thigh rotation with pronounced calf rotation at the knee, and Mode 4 exhibited dominant ankle rotation. Higher-order modes (5–10) manifested coupled elastic deformations and mechanical interactions. These findings suggest a dual-path vibration suppression strategy: For mechanism-driven vibrations (Modes 1–4), optimizing motor control strategies (e.g., switching from position to force control) effectively reduces vibrational energy input without structural modifications. Conversely, elastic deformation-dominated vibrations (Modes 5–10) necessitate structural redesign to shift natural frequencies away from operational excitation bands, thereby suppressing resonance risks. This tiered methodology achieves targeted vibration mitigation while preserving mass-performance equilibrium in lower-limb systems.

Table 1. The natural frequencies of first 10 order vibration mode.

Mode	Frequency/Hz	Mode	Frequency/Hz
1	1.4	6	70.3
2	4.8	7	75.5
3	14.5	8	94.1
4	43.4	9	110.9
5	65.4	10	130.4

**Fig. 4.** Schematic diagrams of first ten order vibration mode of humanoid robot lower limb used in this paper.

3 Harmonic Response Analysis

3.1 Model

The lower-limb structure studied in this work integrates five electric motors. In the finite element model, the stator and rotor components of each motor are explicitly distinguished to enhance simulation accuracy. The operational torque data of these motors, derived from real-world measurements (Table 2), serve as vibrational excitation inputs for the analysis, as illustrated in Fig. 5. Given that all motors operate at frequencies below 15 Hz and are simultaneously energized with synchronized phase activation, the initial phase differences between them are negligible and thus excluded from the analysis.

Table 2. The operational torque data of all motors in use.

Motor number	Torque/N•mm	Motor number	Torque/N•mm
1	9000	4	540
2	4000	5	760
3	2000	—	—

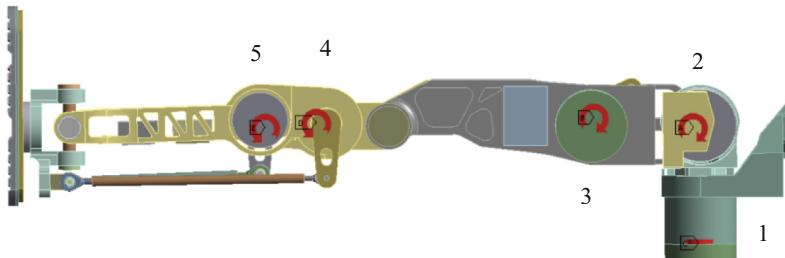


Fig. 5. Schematic diagrams of motors' location on the lower limb.

To evaluate the structural vibration characteristics, a harmonic response analysis spanning 0–20 Hz is conducted. This frequency range comprehensively covers the operational spectrum of the motors while allowing for resonance risk assessment within adjacent frequency bands. The analysis focuses on identifying steady-state vibrational responses (e.g., displacement and stress) under sinusoidal excitation, aiming to predict potential resonance points and optimize structural integrity to avoid fatigue or failure.

3.2 Results

To evaluate vibration characteristics, specific regions at the thigh, knee, calf, and ankle were selected as measurement points, with their average vibration amplitudes systematically analyzed. This approach effectively captures the dynamic responses of structural components within the 0–20 Hz frequency range, as vibrations in this spectrum primarily originate from mechanical interactions between subsystems rather than elastic deformations of individual structural elements. Fig. 6 illustrates the frequency-dependent amplitude variations of these regions under harmonic excitation. Key observations reveal distinct patterns: fore-aft vibrations exhibit peak amplitudes aligned with the first natural frequency (consistent with modal analysis results), where the thigh region demonstrates significantly lower amplitudes compared to the knee, calf, and ankle—the latter three showing nearly identical response magnitudes. In contrast, lateral vibrations display maximum amplitudes at the second natural frequency, with a linear amplitude-distance relationship observed along the limb, likely influenced by structural connectivity dynamics.

These vibration patterns strongly correlate with the system's kinematic chain configuration. The uniformity of fore-aft responses in distal regions suggests symmetric load transmission through parallel linkage mechanisms, while the progressive amplitude increase in lateral vibrations reflects cumulative compliance effects within serial joint connections. Building on these findings, optimization strategies can prioritize structural modifications and control adjustments to mitigate vibration. Potential measures include redesigning connection interfaces to disrupt resonance energy transfer pathways, implementing localized damping mechanisms at amplitude-concentrated areas (e.g., knee, calf, and ankle), and tuning joint stiffness ratios to decouple natural frequencies from operational excitation frequencies. Such amplitude-frequency mapping provides a quantitative foundation for both parametric control refinement and targeted structural enhancements to achieve vibration suppression.

3.3 Vibration Suppression Optimization

Focusing on four structural components—joint connector 1, joint connector 2, thigh, and calf—as optimization targets, material substitutions were implemented (replacing 7075 aluminum alloy with Q235 steel) to evaluate their influence on ankle vibration response amplitudes. This study aimed to identify critical structural contributors to overall vibration behavior and guide targeted optimization. Fig. 7 illustrates the ankle vibration responses in both fore-aft and lateral directions under these material modifications. Analysis reveals distinct directional dependencies: in the fore-aft direction, material changes in all components except the calf showed negligible impact, whereas increasing the calf's mass and stiffness reduced ankle vibration amplitudes to 28% of baseline values.

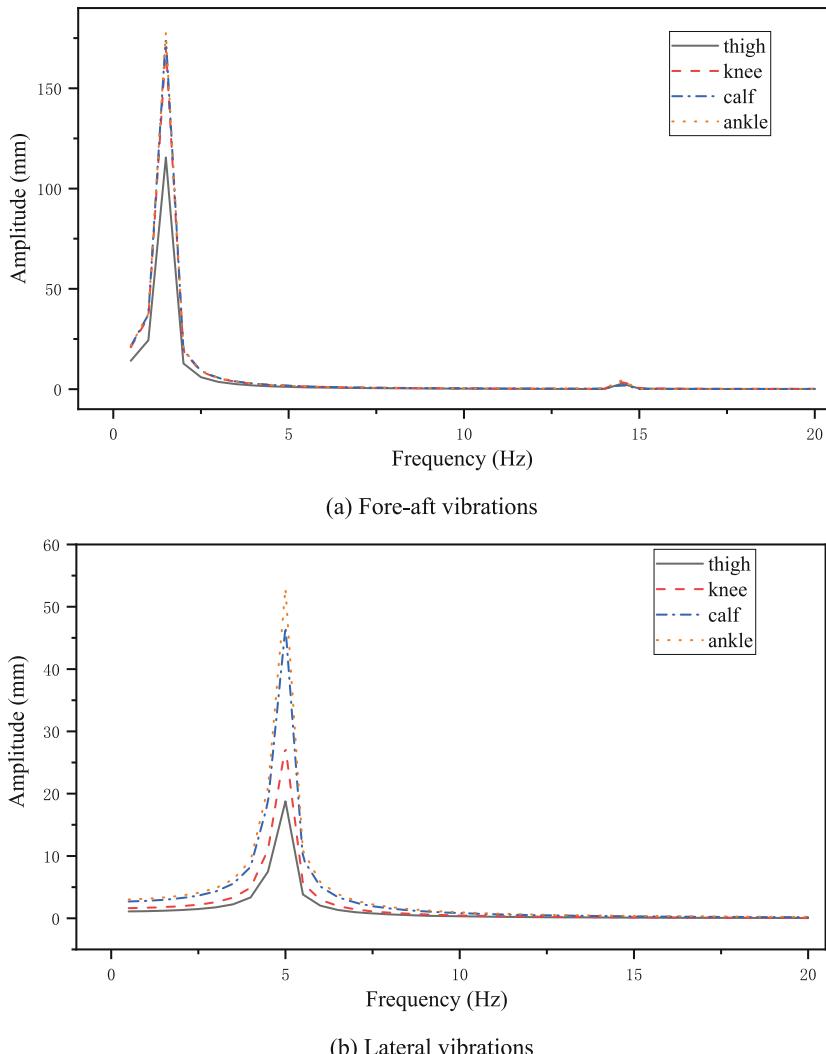


Fig. 6. The frequency-dependent amplitude variations of different regions under harmonic excitation.

Conversely, lateral vibration responses exhibited contrasting patterns. Enhancing the thigh's mass and stiffness significantly reduced ankle vibration amplitudes, while modifications to joint connector 1 had minimal effect. Notably, increasing the mass and stiffness of joint connector 2 or the calf amplified lateral vibrations dramatically, with calf modifications causing over fourfold amplitude increases compared to baseline. This divergence highlights the inadequacy of uniform material substitution for vibration suppression.

Further investigation through modal analysis and structural characterization clarified these phenomena. The lower limb exhibits multiple degrees of freedom in the fore-aft

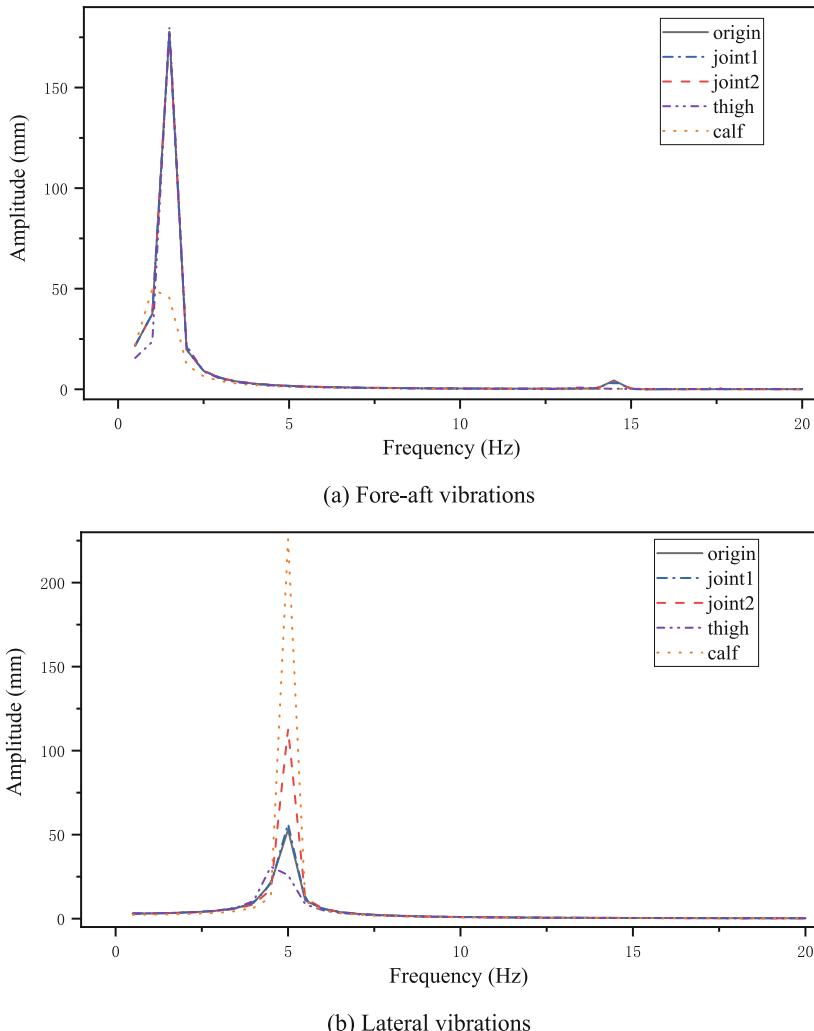


Fig. 7. The frequency-dependent amplitude variations of the ankle vibration responses under material modifications of different components.

direction but none in the lateral direction. Increased calf mass shifts the system's center of gravity downward, inadvertently amplifying resonance energy during lateral vibrations. Building on this understanding, subsequent optimization will focus on maintaining the calf's mass while selectively increasing stiffness, aiming to decouple mass-related resonance effects from stiffness-driven vibration control. This dual approach addresses both directional vibration mechanisms while mitigating unintended amplitude amplification.

A comparative study of overall vibration responses under three calf material configurations was conducted in the simulation model: the initial state (7075 aluminum

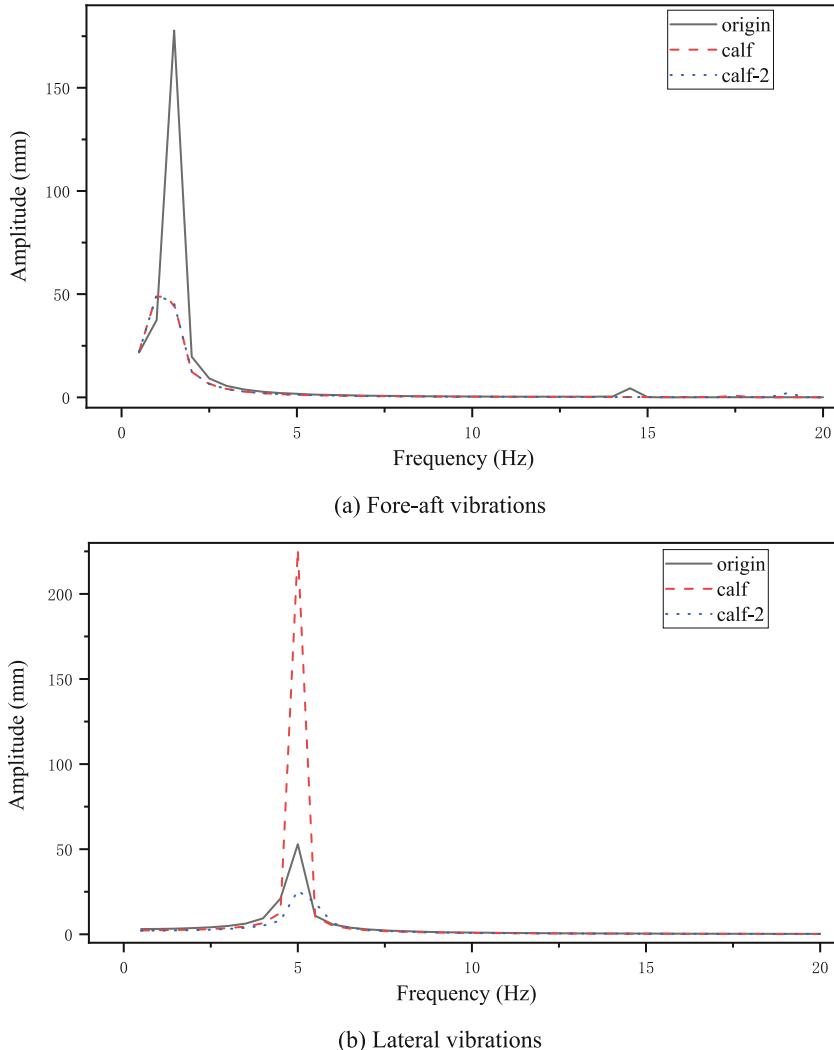


Fig. 8. The frequency-dependent amplitude variations of the ankle vibration responses under two approaches for the calf.

alloy), steel, and a hybrid material configuration combining steel's elastic modulus with aluminum alloy's density (stiffness-enhanced while maintaining original mass). The harmonic response analysis results, shown in Fig. 8, reveal distinct directional behaviors. For fore-aft vibrations, both modified material configurations exhibited nearly identical response amplitudes, significantly lower than the baseline aluminum alloy. In lateral vibrations, the hybrid material configuration (preserving initial density/mass) achieved the smallest amplitude, demonstrating effective vibration suppression. By strategically tailoring material properties in the simulation model, the redesigned calf structure simultaneously reduced vibration amplitudes in both fore-aft and lateral directions. These findings provide actionable insights for designing vibration-resistant structures, emphasizing the importance of decoupling stiffness and mass effects to address multidirectional dynamic challenges.

4 Conclusion

This study systematically investigates vibration characteristics and suppression strategies for humanoid robot lower limbs through finite element simulations. Modal analysis identifies the first 10 natural frequencies and their corresponding mode shapes, distinguishing between vibrations dominated by mechanical motion and those influenced by structural elastic deformation. Harmonic response analysis under realistic motor operating conditions quantifies vibration responses within the 0–20 Hz frequency range, revealing that fore-aft vibrations are primarily driven by low-frequency mechanical interactions, while lateral vibrations are significantly affected by structural compliance. To address vibration suppression, finite element simulations were employed to evaluate the impact of material modifications across four key structural components. Results demonstrated opposing trends in vibration responses for the calf structure between fore-aft and lateral directions. By incorporating idealized material properties (specifically, high specific stiffness) into the calf design through simulations, significant vibration reduction was achieved, providing actionable guidance for practical optimization. Subsequent research will focus on structural refinement using aluminum alloy materials to enhance stiffness while maintaining mass, thereby advancing vibration-resistant designs for robotic lower limbs.

References

1. Luo, N.: Research on Rigid-Elastic Coupling Dynamics and Vibration Suppression of Spacecraft (PhD thesis). Harbin Engineering University, Harbin (2023)
2. Routa, A., Dileep, M., Mahanta, G.B., et al.: Optimal time-jerk trajectory planning of 6 axis welding robot using TLBO method. Procedia Comput. Sci. **133**, 537–544 (2018)
3. John, C.D.: Finite element prediction of damping in structures with constrained viscoelastic layers. AIAA. **20**, 1284–1290 (1984)
4. Rikards, R.: Finite element analysis of vibration and damping of laminated composites. Compos. Struct. **24**, 193–204 (1993)
5. Yang, C., Jiang, Y., Na, J., et al.: Finite-time convergence adaptive fuzzy control for dual-arm robot with unknown kinematics and dynamics. IEEE Trans. Fuzzy Syst. **27**(3), 574–588 (2019)

6. Vijay, M., Jena, D.: Backstepping terminal sliding mode control of robot manipulator using radial basis functional neural networks. *Comput. Electr. Eng.*, S0045790617312004. <https://doi.org/10.1016/j.compeleceng.2017.11.007>
7. Singhose, W.: Command shaping for flexible systems: a review of the first 50 years. *Int. J. Precis. Eng. Manuf.* **10**(4), 153–168 (2009)
8. Du, Y., Wang, C., Lu, J., et al.: Vibration suppression using multi-impulse robust shaping method of zero vibration and derivative. *J. Sound Vib.* **440**, 277–290 (2019)



Design and Implementation of an Intelligent Waste Sorting and Classification System Based on YOLOv5s

Yulin Peng, Xiangpeng Liu^(✉), Hui Zhang, and Kang An

College of Information, Mechanical & Electrical Engineering, Shanghai Normal University, 100 Haisi Road, Shanghai 201418, China
xliu@shnu.edu.cn

Abstract. Rapid urbanization and improved living standards have led to increasing complexity and volume of household waste, making efficient sorting a critical challenge for urban management and environmental sustainability. In this paper, we propose an intelligent waste classification system that integrates visual recognition with mechanical sorting. The design incorporates a dual-conveyor zigzag mechanism for primary and secondary waste separation, a Raspberry Pi 4B deployed to perform object detection via the YOLOv5s model, and a dual-servo gimbal that directs waste into designated bins based on predictions from a PyTorch-trained model converted to TFLite. Utilizing the lightweight YOLOv5s model, the system achieves an mAP₅₀ of 92.4% along with high category-specific average precision (AP): 99.2% for kitchen waste, 92.9% for hazardous waste, 91.6% for recyclable waste, and 85.9% for other waste. Experimental results confirm the system's robust and real-time performance, highlighting its potential to improve automated urban waste management and promote sustainable environmental practices.

Keywords: intelligent waste classification · YOLOv5s · object detection

1 Introduction

Amid rapid urbanization and rising living standards, the diversity and volume of household waste have increased significantly, rendering waste classification a pressing challenge in urban governance and environmental sustainability. Manual sorting methods are generally inefficient and may pose safety risks [1]. With the increasing application of artificial intelligence across various domains, developing intelligent and automated approaches for waste classification has become an active area of research. The realization of intelligent waste classification necessitates the coordinated optimization of two essential stages: mechanical implementation for physical sorting and visual algorithm-based recognition for precise waste categorization.

In recent years, numerous investigations have examined the deployment of robotic arms and associated apparatuses in waste sorting [2–4], leading to measurable improvements in both operational efficiency and sorting accuracy. Robotic arms remain challenged by complex motion control, limited system stability, and high equipment costs in practical applications [5,6]. This paper presents a stable, cost-effective and well-designed conveyor system. It employs a two-stage separation structure, where precise speed control enables both primary and secondary waste separation, thus providing effective preprocessing for subsequent detection and classification.

In the field of visual inspection, Jangsamsi [7] employed local binary patterns (LBP) and a support vector machine (SVM) classifier to extract features from waste images, achieving an average accuracy of 80%. However, due to the limited dataset size, the method remains inadequate for meeting the demands of accurate and efficient waste classification in practical applications. Arebey et al. [8] proposed a method based on feature extraction using the gray-level co-occurrence matrix (GLCM). By optimizing the relevant parameters and integrating a multi-layer perceptron (MLP) with a k-nearest neighbors (KNN) classifier, the method enables efficient detection and classification. Satvikar [9] evaluated five models for waste classification: SVM, random forest (RF), extreme gradient boosting (XGBoost), KNN and convolutional neural network (CNN). The models achieved classification accuracies of 65%, 62%, 70%, 52% and 89%, respectively. Among these, the CNN model demonstrated the highest classification performance.

With the advancement of deep learning, CNN-based detection frameworks have become increasingly dominant. Two-stage detection algorithms, such as Faster R-CNN [10] and Mask R-CNN [11], improve detection accuracy through the integration of a region proposal network (RPN) and exhibit strong performance in waste classification tasks. Kulkarni et al. [12] evaluated a Faster R-CNN-based object detector using the Inception V2 backbone, achieving an accuracy of 84.2%. Two-stage detection algorithms are limited by computational complexity and processing speed, as their dependence on RPN fails to meet the demands of efficient real-time processing in practical applications [13,14].

In contrast, the one-stage YOLO detector series has attracted considerable attention due to its exceptional trade-off between speed and accuracy [15]: Kumar et al. [16] trained and evaluated the YOLOv3 algorithm on six categories of waste—cardboard, glass, metal, paper, plastic and organic waste—demonstrating promising results. However, the model is insufficient in addressing challenges such as category overlap, and its performance in complex scenarios remains suboptimal. Lin [17] proposed the YOLO-Green detection algorithm, derived from YOLOv4, which demonstrates effective detection performance while maintaining a relatively small model size. He et al. [18] introduced enhancements to the YOLOv3 model, which improves detection performance in waste collection and transportation monitoring. By incorporating depthwise separable convolutions and a ternary attention mechanism, the model not only accelerates detection speed but also enhances multi-object recognition capabilities. Glučina et al. [19] proposed a YOLOv4-based method for detecting recy-

clable packaging, focusing on plastic, glass and aluminum waste. However, the method has limitations in hyperparameter tuning and architectural optimization, which affect its ability to prevent overfitting and improve generalization performance.

With the ongoing evolution of the YOLO series, YOLOv5 has attracted considerable attention owing to its enhanced inference speed, superior detection accuracy and modular architecture. YOLOv5 includes several variants—YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. The lightweight design of YOLOv5s ensures real-time detection performance while maintaining high accuracy, making it especially suitable for environments with limited hardware resources. This paper employs YOLOv5s for waste detection and classification, utilizing a custom dataset for model training. This approach enables the automatic identification and categorization of four types of waste through modules such as real-time object detection and intelligent control.

2 Overall System Design

This paper presents an intelligent waste classification system based on visual recognition, which enables the automatic identification and sorting of waste, while effectively displaying classification results and supporting additional functionalities. The system consists of four waste bins, a microcontroller, a display screen, a dual-servo gimbal and a camera. Image recognition technology is employed to identify the waste captured by the camera. Upon successful recognition, the microcontroller transmits commands to the servos, which in turn drive the dual-servo gimbal to rotate and deposit the identified waste into the appropriate bin.

During operation, the waste is introduced into the system through the intake and is initially separated by the first conveyor mechanism. Subsequently, the waste is conveyed to the second conveyor system for secondary separation. The conveyors are arranged in a zigzag pattern, and speed variation is employed to segregate the waste based on the distances between individual pieces. The camera is positioned above the second conveyor belt to capture images. The system is controlled by a Raspberry Pi 4B, which processes the images before they are input into the classification model. The image classification model, implemented using the Pytorch framework, is trained on a PC. The resulting weight, label and configuration files are subsequently converted to TFLite format for deployment on a Raspberry Pi terminal. After processing by the classification model, a prediction is generated. The system then uses this prediction to control the servo motor, which rotates the bottom of the trash bin, allowing the waste to fall into it. The operation process is shown in Fig. 1.

3 Key Technologies

The waste classification system developed in this study is built around the Raspberry Pi 4B single-board computer, which serves as the primary control unit.

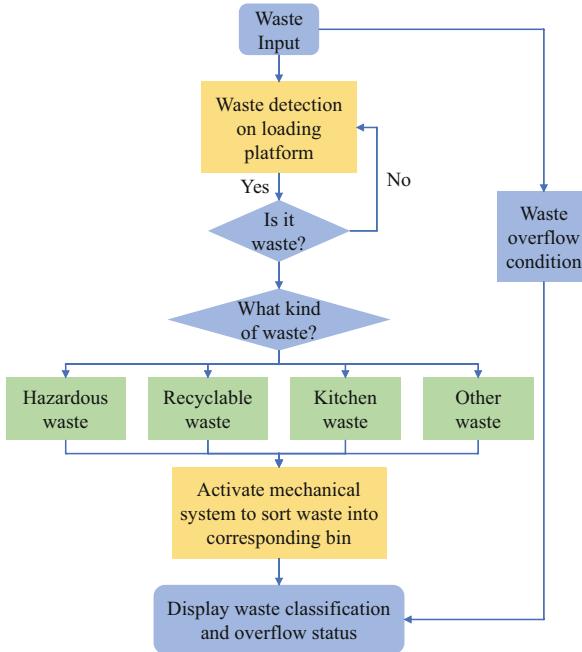


Fig. 1. Workflow of the intelligent waste sorting system.

The system integrates several key modules, including electronic control, display, fully loaded detection and visual recognition, ensuring stable, accurate and efficient operation. The Raspberry Pi precisely controls the rotation of the trash can and the operation of the conveyor belt, utilizing motor and servo adjustments to ensure that the waste is accurately delivered to the designated bin along a predefined path; The display module enables real-time visualization of waste classification results, playback of promotional videos and fully loaded bin alerts, thus providing intuitive feedback and supporting interactive engagement; The fully loaded detection module employs optical sensors to monitor the waste level inside the bin in real time, ensuring timely alerts and preventing overflow; The YOLOv5s algorithm is employed for real-time waste detection and classification in visual inspection, providing essential visual support for the subsequent sorting process.

3.1 Electronic Control Module

The Raspberry Pi activates the motor based on the results of image recognition to rotate the trash can. The conveyor belt is actuated by 42 stepper motors, with their rotation controlled by a Raspberry Pi via the DBD motor driver module. This configuration allows precise control of the conveyor belt, including acceleration and deceleration capabilities. The rotation mechanism of the bottom trash

bin is controlled by an MG996 servo, as depicted in Fig. 2. The Raspberry Pi interfaces with the PCA8695 through the I2C protocol to regulate the servo's rotational angle. The gear transmission system, connected to the central large gear, drives the rotation of the trash can, ensuring that the waste is accurately deposited into the designated bin. Coaxial coupling and bolt fastening are utilized in the gear system to secure the base plate to the four waste bins, thereby ensuring the system's overall stability and reliability.

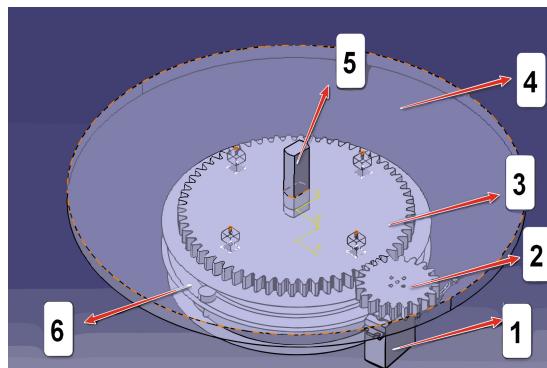


Fig. 2. View of the bottom rotary mechanism assembly. The labeled components are: (1) MG996 servo motor; (2) drive pinion; (3) central spur gear; (4) baseplate; (5) central shaft; and (6) bearing housing.

This design achieves intelligent control of the waste classification system. Precise motor and servo control directs waste along predefined paths into the corresponding bins, thereby improving sorting efficiency and accuracy.

3.2 Display Module

The display module is equipped with a high-definition screen that offers high resolution and enhanced visual clarity, enabling accurate display of waste classification results, instructional videos on waste sorting procedures, and fully loaded bin alerts.

The control program for the display module is developed using the PYQT framework to facilitate interaction with the system control unit. The system transmits real-time data to the display via the Raspberry Pi, ensuring the precise presentation of key information, such as waste classification results. The display is also capable of issuing alerts based on the system's status, such as when the trash reaches capacity. The display module functions not only as an output device but also as an interactive interface between the user and the waste classification system. The display module facilitates user interaction by presenting classification results, playing public service videos, and providing real-time feedback on system status.

3.3 Fully Loaded Detection Module

The Fully Loaded Detection Module employs an optoelectronic sensor to monitor the waste level in the bin in real time, based on the sensor's characteristics and variations in the electrical signal. When the waste level in the bin reaches the preset maximum threshold, the photodetector generates a corresponding electrical signal, which is transmitted to the control system. The Fully loaded status is then displayed, ensuring the timely emptying of the bin and preventing overflow or misclassification of waste.

Due to their non-contact detection capabilities, optoelectronic sensors exhibit high response speed and reliability, effectively reducing contamination and errors associated with direct contact with waste. The optoelectronic sensor was selected for its high sensitivity and adaptability to various waste environments, thus ensuring the accuracy and stability of the detection system under diverse operational conditions.

3.4 YOLOv5s Object Detection Algorithm

The YOLOv5 model, which was developed by the Ultralytics team [20], is a one-stage object detector. Among the four variants in the YOLOv5 family, it has the smallest depth and width, resulting in the fastest inference speed. Thus, it is particularly suitable for deployment on embedded devices. The network architecture comprises four key components: Input, Backbone, Neck, and Head. The Backbone employs a modified version of cross stage partial darknet-53 (CSP-Darknet53) as the principal feature extraction module. The Focus module is used to slice and rearrange the input image, converting the original $608 \times 608 \times 3$ image into a $304 \times 304 \times 32$ feature map. The CSP structure alleviates redundant gradient information by partitioning the gradient flow, thereby reducing the number of model parameters to 7.2 million. The Neck adopts a dual-pyramid architecture based on a feature pyramid network (FPN) and a path aggregation network (PAN), enabling effective multi-scale feature fusion via both upsampling and downsampling operations. The proposed design enhances the detection performance for small objects, such as cigarette butts and bottle caps, while improving the ability to distinguish stacked or overlapping objects. The head employs a Coupled Head architecture, where classification and regression tasks share feature channels. In conjunction with the AutoAnchor strategy for adaptive anchor box computation, this design maintains high localization accuracy amid the diverse shapes of waste objects. generalized intersection over union (GIoU) is employed in place of the conventional IoU in the loss function to improve the detection of multiple and occluded objects.

4 Experiments and Results

4.1 Dataset

A custom waste dataset was constructed for the experiment, comprising four categories: hazardous waste, recyclable waste, kitchen waste and other waste,

with a total of 6,100 images. Image annotation was conducted using LabelImg, with each annotation saved as a .txt file sharing the same name as its corresponding image. The dataset comprises a training set of 500 images and a test set of 5,600 images. Figure 3 presents the visualized statistical analysis of the dataset. (a) illustrates the distribution of object instances across different categories; (b) shows the number and dimensions of bounding boxes; (c) displays the spatial distribution of object centers, where the horizontal and vertical axes represent their coordinates; (d) depicts the distribution of object sizes, with the axes corresponding to width and height.

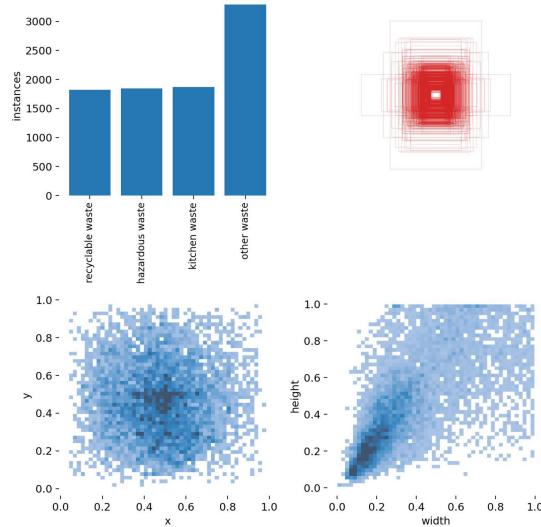


Fig. 3. Dataset visualization.

4.2 Evaluation Indicators

In this study, the performance of the detection model is quantitatively evaluated using precision, recall and AP.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

In which true positives (TP) are positive instances that are correctly classified, while true negatives (TN) are negative instances that are accurately identified. False positives (FP) refer to negative instances that are incorrectly predicted as

positive, and false negatives (FN) are positive instances that are misclassified as negative.

$$AP = \sum_n (R_n - R_{n-1})P_n \quad (3)$$

where R_n and R_{n-1} refer to the recall rates at the n-th and (n-1)-th thresholds, and P_n represents the precision at the n-th threshold.

The average precision in this study was calculated based on the commonly used IoU threshold of 0.5, according to the PASCAL VOC evaluation standard. A detection was considered a true positive only when the intersection over union (IoU) between the predicted bounding box and the ground truth box was greater than or equal to 0.5. Under this criterion, the AP for each category was computed, and the overall mean average precision at IoU 0.5 (mAP₅₀) was used as the primary performance metric.

In addition, the model's performance was evaluated across different confidence thresholds by plotting F1 curves, as shown in Fig. 4. The optimal confidence threshold was selected based on the point at which the F1 score reached its maximum; in this study, the best-performing threshold was determined to be 0.444.

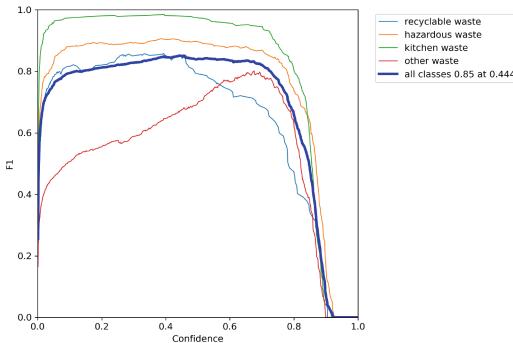


Fig. 4. F1-confidence curve.

4.3 Results Analysis and Discussion

As shown in Fig. 5, the trained YOLOv5 model achieved recall rates of 99% for kitchen waste and 96% for other waste, whereas the recall rates for recyclable and hazardous waste were 92% and 87%, respectively. The model exhibited a primary confusion rate of 7% for recyclable waste and 2% for hazardous waste. In addition, the false negative rate across all categories remained below 6%, reflecting a low incidence of missed detections. Most background misdetections were erroneously classified as other waste, accounting for 85% of background false

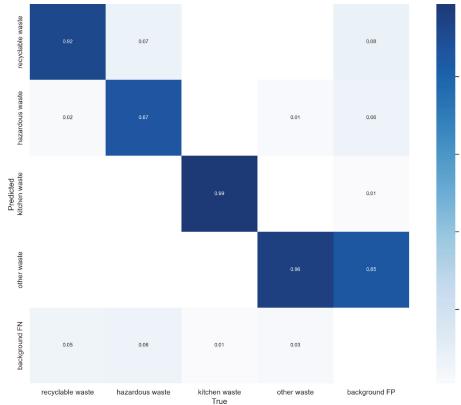


Fig. 5. Confusion matrix of a YOLOv5s model.

positives. This underscores the necessity of further enhancing the background suppression strategy.

Upon completion of model training, the test dataset was fed into the model, and the results are presented in Fig. 6. The blue bounding box in the image signifies recyclable waste, while the orange bounding box represents hazardous waste. The green bounding box indicates kitchen waste, and the red bounding box denotes other waste.

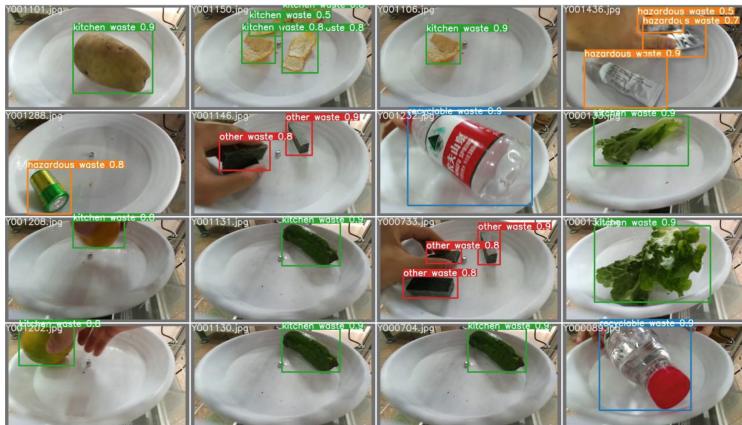


Fig. 6. Algorithm performance evaluation.

A comprehensive evaluation of the YOLOv5s-based waste classification model was conducted using multiple performance metrics, including precision and recall. Figure 7 displays the precisionrecall curve of the model, highlighting the trade-off between precision and recall at various confidence thresholds.

According to the $\text{IoU} = 0.5$ evaluation standard defined in Sect. 4.2, the model achieved an overall mAP_{50} of 92.4%. The AP for kitchen waste was the highest at 99.2%, maintaining high recall alongside strong precision. The AP for hazardous and recyclable waste was 92.9% and 91.6%, indicating robust and consistent performance. The other waste showed the lowest AP of 85.9%, with precision declining noticeably at high recall levels. Each curve preserved a precision above 70% within the recall range of 60% to 80%. This indicates the model's robustness across various threshold settings and supports flexible selection of confidence thresholds tailored to specific application scenarios during deployment.

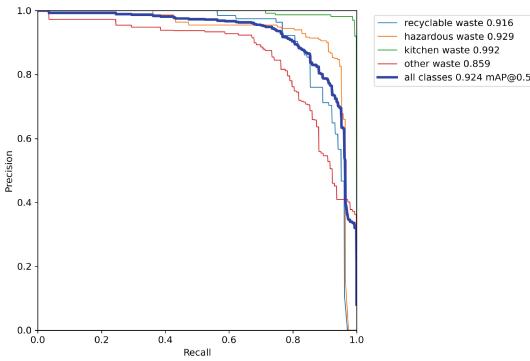


Fig. 7. Precision-recall curve.

5 Conclusion

This work proposes an intelligent waste classification system based on the lightweight YOLOv5s model, which performs automated waste recognition, sorting, and real-time information display. Experimental results show that the proposed system achieves an overall mAP_{50} of 92.4%. The category-wise AP is 99.2% for kitchen waste, 92.9% for hazardous waste, 91.6% for recyclable waste, and 85.9% for other waste. The model exhibits high recognition accuracy across different waste categories and demonstrates strong robustness. The system employs a zigzag configuration with dual conveyor belts to achieve both primary and secondary waste separation, ensuring clear and continuous image acquisition by the camera. A well-designed mechanical structure, integrated with image recognition, enables efficient coordination across all stages of waste processing, including conveying, recognition and sorting. In the future, the system's applicability and robustness in a broader range of real-world scenarios may be further improved by expanding the training dataset, refining the model architecture, and enhancing hardware response speed. Additionally, analysis reveals that a large proportion of background false positives are misclassified as other waste, likely due to overlapping features between this category and background textures.

such as floors or walls. To mitigate this issue, future research will explore the inclusion of explicitly labeled background samples, refinement of category definitions, and optimization of the loss function. These improvements aim to reduce misclassification and further enhance the model's robustness in complex real-world environments. These enhancements can offer more practical and valuable solutions for waste management and resource recycling.

References

1. Ji, T., Fang, H., Zhang, R., Yang, J., Fan, L., Li, J.: Automatic sorting of low-value recyclable waste: a comparative experimental study. *Clean Technol. Environ. Policy* **25**(3), 949–961 (2023). <https://doi.org/10.1007/s10098-022-02418-7>
2. Hegde, R.B., D'Mello, G., Swathi, N. C: Smart garbage sorting system: integrating robotic arm and machine learning for segregation. In: 2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), pp. 249–254. IEEE (2024). <https://doi.org/10.1109/discover62353.2024.10750572>
3. Lahoti, J., Sn, J., Krishna, M.V., Prasad, M., Bs, R., Mysore, N., Nayak, J.S.: Multi-class waste segregation using computer vision and robotic arm. *PeerJ Comput. Sci.* **10**, e1957 (2024). <https://doi.org/10.7717/peerj-cs.1957>
4. Thao, L.Q.: An automated waste management system using artificial intelligence and robotics. *J. Mater. Cycles Waste Manag.* **25**(6), 3791–3800 (2023). <https://doi.org/10.1007/s10163-023-01796-4>
5. Zhang, Y., Yu, Z., Liang, D., Liu, D., Wang, G., Ding, W.: Research on mechanical claw control method for UAV with robotic arm. In: 2024 4th International Symposium on Artificial Intelligence and Intelligent Manufacturing (AIIM), pp. 241–244. IEEE (2024). <https://doi.org/10.1109/aiim64537.2024.10934204>
6. Zhang, X., Xu, J., Fu, H., Hu, S.: Research on moving arm grasping based on computer vision. In: 2022 5th International Conference on Robotics, Control and Automation Engineering (RCAE), pp. 298–301. IEEE (2022). <https://doi.org/10.1109/rcae56054.2022.9995949>
7. Jangsamsi, K.: Conventional machine learning approach for waste classification. In: Proceedings of the 2023 6th Artificial Intelligence and Cloud Computing Conference, pp. 7–12 (2023). <https://doi.org/10.1145/3639592.3639594>
8. Arebey, M., Hannan, M.A., Begum, R.A., Basri, H.: Solid waste bin level detection using gray level co-occurrence matrix feature extraction approach. *J. Environ. Manag.* **104**, 9–18 (2012). <https://doi.org/10.1016/j.jenvman.2012.03.035>
9. Satvilkar, M.: Image based trash classification using machine learning algorithms for recyclability status. Ph.D. thesis, National College of Ireland, Dublin (2018)
10. Nie, Z., Duan, W., Li, X.: Domestic garbage recognition and detection based on Faster R-CNN. *J. Phys. Conf. Ser.* **1738**(1), 012089 (2021) <https://doi.org/10.1088/1742-6596/1738/1/012089>
11. Jaikumar, P., Vandaele, R., Ojha, V.: Transfer learning for instance segmentation of waste bottles using Mask R-CNN algorithm. In: International Conference on Intelligent Systems Design and Applications, pp. 140–149. Springer (2020) https://doi.org/10.1007/978-3-030-71187-0_13
12. Kulkarni, H.N., Raman, N.K.S.: Waste object detection and classification. CS230 Stanford (2019)

13. Zou, Z., Chen, K., Shi, Z., Guo, Y., Ye, J.: Object detection in 20 years: a survey. arXiv preprint [arXiv:1905.05055](https://arxiv.org/abs/1905.05055) (2019)
14. Du, L., Zhang, R., Wang, X.: Overview of two-stage object detection algorithms. *J. Phys. Conf. Ser.* **1544**(1), 012033 (2020). <https://doi.org/10.1088/1742-6596/1544/1/012033>
15. Ali, M.L., Zhang, Z.: The YOLO framework: a comprehensive review of evolution, applications, and benchmarks in object detection. *Comput.* **13**(12), 336 (2024). <https://doi.org/10.3390/computers13120336>
16. Kumar, S., Yadav, D., Gupta, H., Verma, O.P., Ansari, I.A., Ahn, C.W.: A novel yolov3 algorithm-based deep learning approach for waste segregation: Towards smart waste management. *Electronics* **10**(1), 14 (2020). <https://doi.org/10.3390/electronics10010014>
17. Lin, W.: Yolo-green: A real-time classification and object detection model optimized for waste management. In: 2021 IEEE International Conference on Big Data (Big Data), pp. 51–57. IEEE (2021). <https://doi.org/10.1109/bigdata52589.2021.9671821>
18. He, Y., Li, J., Chen, S., Xu, Y., Liang, J.: Waste collection and transportation supervision based on improved YOLOv3 model. *IEEE Access* **10**, 81836–81845 (2022). <https://doi.org/10.1109/access.2022.3195995>
19. Glučina, M., Baressi Šegota, S., Andelić, N., Car, Z.: Automated detection and classification of returnable packaging based on YOLOV4 algorithm. *Appl. Sci.* **12**, 11131 (2022). <https://doi.org/10.3390/app122111131>
20. Tzedakis, G., Tzamali, E., Spanakis, E.G., Antonakakis, M., Zervakis, M., Sakkalis, V.: Comparing YOLO-based detectors for Pedestrian and car detection in aerial static video: an evaluation of generalization capacity and performance. In: 2023 IEEE International Conference on Imaging Systems and Techniques (IST), pp. 1–5. IEEE (2023). <https://doi.org/10.1109/ist59124.2023.10355688>



Missed Sample Exploration by Class-Agnostic Detector for Embodied Learning

Panpan Wang and Di Guo^(✉)

Beijing University of Posts and Telecommunications, The School of Artificial Intelligence, Beijing, China
{wpanpan,di.guo}@bupt.edu.cn

Abstract. Offline trained object detectors perform well on public datasets but degrade in the real world due to distributional shifts. Embodied intelligence, by enabling sample collection in the target environment, offers a solution. A key challenge is determining which samples to collect. Current methods primarily focus on semantically uncertain data, overlooking the importance of missed detection samples. This paper addresses this by using a class-agnostic detector to identify potential objects within the agent’s egocentric view, integrating local and global information as rewards. Experiments show that the proposed method effectively captures informative detection samples for perception model training.

Keywords: Embodied active learning · data collection · class-agnostic detection · reinforcement learning

1 Introduction

Modern object detection models typically use offline supervised learning, training on static datasets. However, this requires high annotation costs. Active learning reduces these costs by selecting valuable samples, training better models with fewer samples. Current methods select samples from static datasets, which are collected offline and may have distributional biases, leading to performance degradation in real-world environments. Embodied intelligence can interact with the environment online and discover valuable samples through real-time feedback, known as embodied active learning. A key challenge is determining which samples to acquire. Chaplot et al. [2], Jing et al. [3] and Scarpellini et al. [8] used different module like semantic map increasement, KL distance of prediction and a top-down disagreement map as uncertrain rewards, respectively. These methods focus on uncertain samples but overlook missed detections.

From a human perspective, we initially perceive distant objects without categorizing them, classifying them only when closer. This latter stage can be achieved with a class-specific detector. Inspired by this, we use a class-agnostic

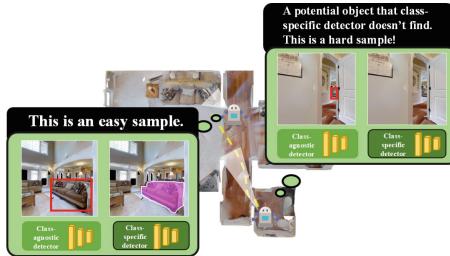


Fig. 1. Unrecognized objects of class-specific detector. We use a class-agnostic detector to assist class-specific detector for finding missed object thus hard sample.

detector to assist class-specific detectors in identifying unrecognized objects and collecting hard samples. Unlike typical class-specific detectors that detect and classify targets, class-agnostic detectors only localize potential “objects”, as shown in Fig. 1.

This paper presents a two-phase framework for embodied active learning, consisting of policy learning and detector training (Fig. 2). During policy learning, the agent uses global and local information from class-specific and class-agnostic detectors within a perception module to train a sampling policy in training scenes. In the detector training phase, the trained policy collects samples in test scenes to fine-tune the class-specific detector, improving its perception. Our key contributions are: 1) A reward system integrating global map and local egocentric perspectives to train the sampling policy. 2) A class-agnostic detector to identify missed detections from a local perspective. 3) Experimental results demonstrating effective collection of missed detection samples, leading to strong performance on the Gibson-tiny dataset.

2 Related Work

2.1 Active Learning for Object Detection

Active learning selects a small and informative subset of data in unlabeled dataset for model training. The methods primarily fall into two categories based on uncertainty [4, 9] and diversity [6, 11]. Active learning in object detection like [1, 13, 14]. However, active learning typically samples on static datasets.

2.2 Embodied Learning

Embodied learning involves robots learning through environmental interactions to complete tasks. This includes instruction following, visual exploration, and room arrangement. Our task, embodied active learning, focuses on informative sample collection for perception model enhancement. Chaplot et al. [2] used reinforcement learning (RL) to train agents for data collection, rewarding increases

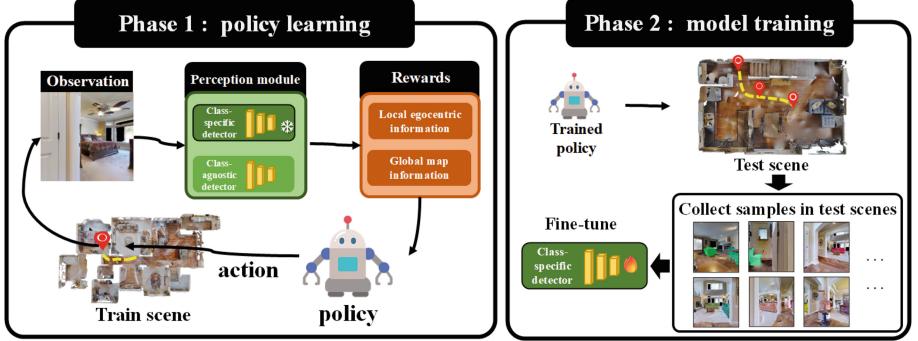


Fig. 2. The framework of embodied active learning. It contains two phases: the first phase involves training an exploration agent in train scenes. For RGB required per step, two detectors in the perception module predict separately to obtain local ego-view and global map information. These two signals serve as rewards to train the exploration policy. During this phase, the parameters of class-specific detector are fixed. In the second phase, the trained exploration policy explores in the test scenes, and collect samples to fine-tune the class-specific detector.

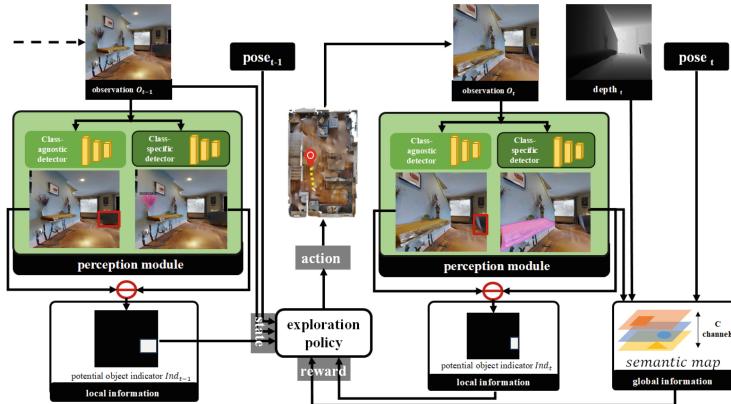


Fig. 3. The framework of exploration policy. At each step, the policy's input comprises observations, potential object indicators, and agent pose. The potential object indicator, derived from the perception module, highlights objects detected by the class-agnostic detector but missed by the class-specific detector. Reward based on local and global information returned at next step.

in semantic map accuracy. Jing et al. [3] further constructed a 3D semantic distribution map. Zhao et al. [15] also employed RL to identify inconsistent 3D bounding box predictions. Scarpellini et al. [8] created a top-down disagreement map for rewards. Wang et al. [12] considered both consistency and inconsistency in predictions for rewards. Nilsson et al. [7] used RL for local navigation and

rewarded perception model metric enhancements. These methods primarily focus on low-confidence or inconsistent detections, neglecting missed detections.

3 Problem Formulation

In embodied active learning, an agent explores to collect samples D_{sample} covering K target classes $C = C_1, \dots, C_k$ to train a class-specific detection model M_{spec} . The goal is to minimize D_{train} size while maximizing M_{spec} performance. Our two-phase approach involves training a sampling policy π in training environments S_{train} in phase one. In phase two, π collects samples in test environments S_{test} to build D_{train} for M_{spec} , enhancing its performance in S_{test} . The agent's exploration is a Markov decision process (MDP) defined by $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$. At timestep t , the agent receives state s_t , takes action a_t , and the environment provides reward r_t and transitions to s_{t+1} . The state transition probability is $P = \mathbb{P}[s_{t+1}|s_t, a_t]$. In phase one within S_{train} , the agent learns to explore. At timestep t , it gets an egocentric observation O_t , processed by M_{spec} for semantic predictions B_{sem} and by M_{agnos} for objectness prediction B_{obj} . An egocentric potential object indicator ind_t is derived from B_{sem} and B_{obj} . The state $s_t = O_t, ind_t, Pose_t$ inputs into π , which selects action a_t . Then the environment returns s_{t+1} . In phase two, π explores S_{test} . At each time t , the agent receives observation o_t . At maximum timestep T , we get $D_{sample} = o_1, \dots, o_T$. After removing samples without target objects, we derive $D_{train} = \{o \in D_{sample}; |Class(o) \in C\}$. M_{spec} is fine-tuned on D_{train} . A randomly collected test set D_{test} in S_{test} evaluates it.

4 Methods

This paper focuses on training an embodied agent via reinforcement learning to explore and gather challenging samples for a class-specific detector in novel environments. Our approach employs a perception module with class-specific and class-agnostic detectors to build global and local information, guiding the active exploration agent, as shown in phase 1 of Fig. 2.

4.1 Global and Local Information from Perception Module

Figure 3 illustrates the agent's perception module, comprising a class-specific detector (target model for finetuning) and a class-agnostic detector (assistant for environment exploration). Local information is derived from the prediction discrepancy between the two detectors. Upon acquiring an image, both detectors infer semantic B_{sem} and object B_{obj} predictions, where b_j^{sem} and b_i^{obj} are bounding boxes. A set Ind identifies potential objects missed by the class-specific detector based on IoU thresholds:

$$Ind = b_i^{obj} \mid \max_{b_j^{sem} \in B_{sem}} IoU(b_i^{obj}, b_j^{sem}) < \tau_{IoU} \quad (1)$$

A binary mask $ind(B_{sem}, B_{obj})$ marks these potential objects. Global information is provided by a semantic map. The agent receives RGB, depth, and pose data, using the class-specific detector to project object segmentations into a point cloud, which is then max-pooled across height into a two-dimensional map. This map, with C channels for K target categories, is updated with last one iteratively to retain the maximum value at each step.

4.2 Exploration Policy

We train a policy by reinforcement learning to determine the agent’s actions in the environment. It can be seen in Fig. 3, at timestep $t - 1$, collecting the current RGB observation O_{t-1} , the agent acquires the potential object indicator ind_{t-1} from the prediction of perception module. The agent’s policy network then takes as input the RGB observation O_{t-1} , the potential object indicator ind_{t-1} and the agent’s $pose_{t-1}$. Then it select an action from $[left, right, forward]$. We design a reward that incorporate both global and local information. The ind_t is taken as the local reward, which is defined as the number of potential objects detected by the agent:

$$reward_{t-1,local} = \lambda_{local} |Ind_t| \quad (2)$$

λ_{local} is the coefficient of local reward. Additionally, to promote the exploration of more objects in the global level, we adopt a global reward from semantic map M^{sem} , namely the increase number of the semantic map voxels:

$$reward_{t-1,global} = \lambda_{global} (M_t^{sem} - M_{t-1}^{sem}) \quad (3)$$

λ_{global} is the coefficient of global reward. Our overall reward is sum of them:

$$Reward_{t-1} = reward_{t-1,local} + reward_{t-1,global} \quad (4)$$

5 Experiment

5.1 Policy Training

Experiments are conducted on the Gibson-tiny dataset with 24 train and 4 test scenes using the Habitat simulator. The policy is trained via Proximal Policy Optimization (PPO)[10] for 10 billion frames, taking 3–4 days on 3 NVIDIA 4090 GPUs. Maximum episode length is 500. The policy includes a ResNet18 visual encoder, a convolution layer, and two linear layers for action distribution and value prediction. A fourth convolution channel of visual encoder with random initialization is added for potential object indicators. PPO epoch is set to 4, with an entropy coefficient of 0.001 and a value loss coefficient of 0.5. The Adam optimizer is used with a learning rate of $2.5e - 5$. Global map reward and local egocentric reward coefficients are $2.5e - 3$ and 1, respectively. The IoU threshold τ_{IoU} for potential object indicators is 0.5.

5.2 Model Finetuning

Our class-agnostic detector uses a pretrained Faster-RCNN with a ResNet101 backbone with weights from [5]. The class-specific detector employs a pretrained COCO Mask-RCNN with a ResNet50 backbone, focusing on 5 categories: couch, chair, toilet, plant, and bed. Both detectors have a prediction score threshold of 0.5. The class-specific detector is fine-tuned for 10 epochs with a batch size of 8. Ground truth samples are provided by the environment simulator. We collect samples for 5 episodes in each environment, removing samples without target objects to create the final valid training dataset D_{train} for fine-tuning the class-specific detector. We then evaluate it on a pre-collected test set D_{test} , acquiring 1000 samples of the target object in each test environment.

5.3 Baselines

We compare with the following baselines:

- **Random:** Random policy selects actions randomly at each time step.
- **Forward:** Forward policy is a heuristic policy that goes forward continuously and rotates when collision occurs.
- **Semantic Curiosity:** Semantic Curiosity policy from [2] are reproduced by us. We train the policy on the Gibson-tiny dataset. The reward is the incremental of semantic map:

$$reward_{t-1} = \lambda_{global}(M_t^{sem} - M_{t-1}^{sem}) \quad (5)$$

5.4 Results

We compared fine-tuned class-agnostic detectors of all methods using the *map50* metric on D_{test} , as shown in Table 1. Results indicate that our fine-tuned model, compared to the pre-trained model, shows the most significant improvement in accuracy among all baselines, with a 16% increase in *map50*. Notably, Semantic Curiosity policy's underperformed compared to the random policy, suggesting that exploration based solely on the global semantic map is ineffective for discovering difficult samples. Our method, leveraging the class-agnostic detector's guidance from an egocentric perspective, effectively identifies challenging samples. Furthermore, even potential objects from the class-agnostic detector not belonging to the target categories serve as an exploration signal, leading to the discovery of more objects. About the number of valid samples collected by different policies. While the random policy collected the most samples, the fine-tuned model's *map50* was not the highest. Our policy, collecting 30% fewer samples than the random policy, achieved better fine-tuning results, demonstrating the higher quality of our D_{train} .

To further investigate the quality of D_{train} collected by policies of baselines, we evaluated the performance of the pre-trained model on these D_{train} using *mar10* and *map50* (Table 2). A positive correlation between *mar10* and the

Table 1. The *map50* of pre-trained and finetuned class-specific models from various policies, and corresponding valid sample number of D_{train}

Method	<i>Map50</i>	Number of valid samples
Pre-trained	23.9%	—
Random	36.2%	6776
Forward	21.85%	3592
Semantic Curiosity	33%	4648
Ours	40.5%	4712

Table 2. The performance of pre-trained model M_{spec} on D_{train} from different policies

Policy	<i>Mar10</i>	<i>Map50</i>
Random	41.4%	49.6%
Semantic Curiosity	43.8%	42.3%
Ours	36.8%	41.5%

fine-tuned model’s *map50* improvement (Table 1) suggests that missed samples significantly enhance the class-specific detector. Notably, on our policy’s D_{train} , the pre-trained model had the lowest *mar10* (38.1%), but fine-tuning on this dataset increased *map50* from 23.9% to 40.5%. Since *mar10* reflects the recall of the top 10 prediction boxes per image, this indicates our method’s D_{train} had the most missed detections, highlighting the pre-trained model’s lowest recall rate on this dataset.

We evaluate the impact of policy input. We discard pose input in our policy for ablation. It can be seen in Table 3, the pose input contributed to the policy. Otherwise, we compare the impact of the local reward coefficient λ_{local} in [1, 0.1, 0.01, 0.001] in Table 4. And we set the best learning rate 1.

Finally, we visualized both detectors’ predictions on our method’s D_{train} (Fig. 4). The visualization reveals that our method successfully collects missed samples, including occluded ones (a-d), that the class-specific detector fails to recognize. Furthermore, we visualized the detector’s performance on D_{test} before and after finetuning, with selected samples shown in Fig. 5. The enhanced detector, after applying our method, can detect objects previously indiscernible in the scene.

Table 3. Ablation of pose in policy input

Policy input	<i>Map50</i>	Number of valid samples
pose+RGB+ind	40.5%	4712
RGB+ind	39.55%	5312

Table 4. Impact of local reward coefficient λ_{local}

λ_{local}	Map50	Number of valid samples
1	40.5%	4712
0.1	34.6%	5872
0.01	39.5%	6352

**Fig. 4.** Prediction on samples collected by our method. The first and second rows are from class-agnostic and class-specific detector, respectively.**Fig. 5.** Predictions of class-specific detector in test set before and after finetune.

6 Conclusion

This paper presents a policy integrating global and local rewards, emphasizing the use of a class-agnostic detector for local guidance in uncovering challenging samples for class-specific detectors in embodied settings. Results indicate that our method effectively identifies difficult, often missed, samples, thereby improving the agent’s performance.

References

1. Agarwal, L., Arora, H., Anand, S., Arora, C.: Contextual diversity for active learning. In: Computer Vision-ECCV: 16th European Conference, Glasgow, UK, 23–28 Aug 2020, pp. 137–153. Proceedings, Part XVI 16. Springer (2020)

2. Chaplot, D.S., Jiang, H., Gupta, S., Gupta, A.: Semantic curiosity for active visual learning. In: Computer Vision-ECCV,: 16th European Conference, Glasgow, UK, 23–28 Aug 2020, Proceedings, Part VI 16, pp. 309–326. Springer (2020)
3. Jing, Y., Kong, T.: Learning to explore informative trajectories and samples for embodied perception. In: 2023 IEEE International Conference on Robotics and Automationpp. 6050–6056. IEEE (2023)
4. Lewis, D.D., Catlett, J.: Heterogeneous uncertainty sampling for supervised learning. In: Machine Learning Proceedings, pp. 148–156. Elsevier (1994)
5. Maaz, M., Rasheed, H., Khan, S., Khan, F.S., Anwer, R.M., Yang, M.H.: Class-agnostic object detection with multi-modal transformer. In: European Conference on Computer Vision, pp. 512–531. Springer (2022)
6. Nguyen, H.T., Smeulders, A.C.: Active learning using pre-clustering. In: Proceedings of the Twenty-First International Conference on Machine Learning, p. 79 (2004)
7. Nilsson, D., Pirinen, A., Gärtner, E., Sminchisescu, C.: Embodied learning for lifelong visual perception, *arXiv preprint arXiv:2112.14084* (2021)
8. Scarpellini, C., Rosa, S., Morerio, P., Natale, L., Del Bue, A.: Look around and learn: self-training object detection by exploration. In: European Conference on Computer Vision, pp. 72–88. Springer (2024)
9. Scheffer, T., Decomain, C., Wrobel, S.: Active hidden Markov models for information extraction. In: International Symposium on Intelligent Data Analysis, pp. 309–318. Springer (2001)
10. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017)
11. Van den Bergh, M., Boix, X., Roig, G., De Capitani, B., Van Gool, L.: Seeds: superpixels extracted via energy-driven sampling. In: Computer Vision–ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, 7–13 Oct 2012, Proceedings, Part VII 12, pp. 13–26. Springer (2012)
12. Wang, J., Liu, X., Zhao, D., Dai, B., Liu, H.: Self-supervised embodied learning for semantic segmentation. In: 2023 IEEE International Conference on Development and Learning, pp. 383–390 (2023)
13. Wu, J., Chen, J., Huang, D.: Entropy-based active learning for object detection with progressive diversity constraint. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9397–9406 (2022)
14. Yu, K., Albro, S., DeSalvo, G., Kothawade, S., Rashwan, A., Tavakkol, S., Battmanghelich, K., Yin, X.: Two-step active learning for instance segmentation with uncertainty and diversity sampling. *arXiv preprint arXiv:2309.16139* (2023)
15. Zhao, Q., Zhang, L., Wu, L., Qiao, H., Liu, Z.: A real 3d embodied dataset for robotic active visual learning. IEEE Robot. Autom. Lett. **7**(3), 6646–6652 (2022)



Node Classification and Link Prediction Methods Based on Hybrid Teacher Models

Yuxin Chen^{1,2}, Hao Yan^{1,2}, Qingyou Dai⁴, Diting Zhou³, Haohao Fu^{1,2}, Maoyu Wang^{1,2}, and Zhenbo Cheng^{1,2(✉)}

¹ College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

czb@zjut.edu.cn

² Zhejiang Key Laboratory of Visual Information Intelligent Processing, Zhejiang University of Technology, Hangzhou 310023, China

³ Library, Zhejiang University of Technology, Hangzhou 310023, China

⁴ Guangdong Institute of Special Equipment Inspection and Research, Guangdong 528251, China

Abstract. The knowledge question-answering system leverages knowledge graphs for retrieval and inference, enabling it to handle complex queries. However, knowledge graphs in specialized technical fields are often incomplete, with limited coverage, and fail to provide sufficient support for question-answering systems. To address these issues, this paper delves into relevant knowledge graph technologies and proposes a node classification and link prediction method based on a hybrid teacher model. This method consists of multiple teacher models, integrating the strengths of different teacher models' knowledge. Based on this foundation, tasks for node classification and link prediction have been developed. In addition to analyzing each module's overall results, experiments on citation network datasets and similar datasets verify that the model performs better in terms of network classification and link prediction tasks across many datasets.

Keywords: Knowledge graph · Node classification · Link prediction · Question-answering system

1 Introduction

With the rapid development of generative large language models, knowledge question-answering systems now combine these models with knowledge graphs for response generation. However, traditional knowledge graph construction methods, both semi-automated and fully automated [1, 2], suffer from incompleteness. Knowledge graph completion methods tackle this problem, and the majority of approaches rely on graph neural networks. However, many use single models, which fail to fully capture complex structural relationships and semantic features, limiting performance in certain tasks.

To improve generalization, recent studies have used multi-teacher graph neural networks, integrating heterogeneous knowledge from teacher models to enhance the target model's learning. For example, Asif et al. [3] proposed a multi-teacher ensemble framework, and Shen et al. [4] designed a knowledge fusion framework. However, these methods often assume equal contributions from all teachers and lack dynamic weight allocation, potentially introducing low-quality information.

To address this, we propose a knowledge graph node classification and link prediction method based on a hybrid teacher model. The model integrates multiple teacher models' knowledge through a teacher weight allocation module and a hybrid model alignment module, overcoming the limitations of single models and capturing the complexity and diversity of knowledge graphs. In other words during the training process, capturing the diversity and complexity of the knowledge graph in a more comprehensive manner.

2 Method

This paper proposes a knowledge graph node classification and link prediction method based on a hybrid teacher model (HTM-KG) for tasks in knowledge graphs. Figure 1 shows that there are three primary parts to the model in its structure: the hybrid model alignment module, the teacher weight allocation module, and the multi-teacher model module. Together, these modules enable the hybrid model to effectively absorb knowledge from multiple teacher models, leading to improved performance in knowledge graph classification and prediction tasks.

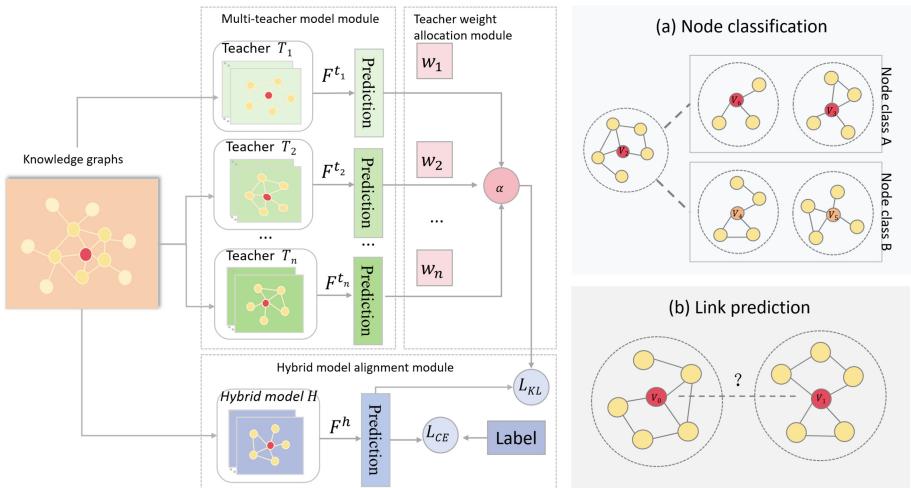


Fig. 1. Overall framework of the knowledge graph classification and prediction model based on the hybrid teacher model.

2.1 Multi-teacher Model Module

In this method, an ensemble strategy using multiple teacher models is adopted. The core idea is to enable the hybrid model to learn diverse knowledge from different teacher models. Three different types of graph neural networks, denoted as T1, T2, and T3, are selected as teacher models. In the node classification task, teacher models with varying layers are chosen, as they capture different semantic information from shallow to deep layers. This design allows the hybrid model to gradually learn representations from local features to global features, thereby enhancing classification performance.

Finally, these teacher models calculate the cross-entropy loss using the input data and true labels, indicating how accurately the model predicts the current data. The calculation formula is as follows:

$$L_{T_i} = \text{CrossEntropy}(P_t(X), Y) = - \sum_{i=1}^N Y_i \log(P_t(X)_i) \quad (1)$$

Where, given an input X , the predicted probability of the i -th class by the teacher model T_i is denoted as $P_t(X)_i$, whereas Y_i represents the ground truth label for class i .

2.2 Teacher Weight Allocation Module

Since different teacher models focus on different regions and learn different knowledge, the hybrid model is able to acquire implicit knowledge and deep semantics from multiple teacher models. To this end, this module introduces an attention mechanism to dynamically allocate weights to different teacher models, ensuring that each teacher has independent discriminative power in the hybrid feature space, and quantifying the differences between the hybrid model and each teacher. The prediction result of the teacher model is represented as $P_t(X)_i$, and the prediction result of the hybrid model is represented as $P_h(X)$. The degree of match between the hybrid model and the teacher models is evaluated by calculating the similarity between their prediction results. The calculation formula for this similarity is as follows:

$$\text{sim}(P_h(X), P_t(X)_i) = \frac{P_t(X)_i \bullet P_h(X)}{\|P_h(X)\| \|P_t(X)_i\|} \quad (2)$$

Where, $\|P_t(X)_i\|$ and $\|P_h(X)\|$ are the norms of the prediction results of the teacher and hybrid models, respectively, representing their similarity in the feature space.

After calculating the similarity, the similarity results are processed using $\log_{softmax}$ and averaged to obtain the final dynamic weight α_{T_k} for each teacher model. This approach effectively adjusts the contribution of each teacher model to the hybrid model, ensuring a reasonable weight allocation. The specific formula is as follows:

$$\alpha_{T_k} = \text{Mean}(\log_{softmax}(\text{sim}(P_t(X)_i, P_h(X)))) \quad (3)$$

2.3 Hybrid Model Alignment Module

To measure the difference between the hybrid model and the teacher models, The hybrid alignment model incorporates the Kullback-Leibler (KL) divergence. The overall loss for the hybrid model is computed as a combination of cross-entropy loss and KL divergence. First, during the training of the hybrid model, cross-entropy loss is used to quantify the difference between the hybrid model's output and the real labels, ensuring that the output of the hybrid model aligns with the true labels. The specific calculation formula is as follows:

$$L_{CE} = \text{CrossEntropy}(P_h(X), Y) - \sum_{i=1}^N Y_i \log(P_h(X)_i) \quad (4)$$

Where, $P_h(X)$ is the predicted probability of the i -th class by the hybrid model H on the input data X , and Y_i is the value of the i -th class in the true labels. By optimizing the cross-entropy loss, the hybrid model will try to make its output as close as possible to the true labels.

Subsequently, the KL divergence is applied to quantify the distance between the probability distributions produced by the hybrid model and the teacher models, thus guiding the hybrid model to learn from the teachers. The alignment loss function is formulated as:

$$L_{align} = L_{KL} \left(\sigma \left(\frac{P_h(X)}{\tau} \right) \sigma, \sum_{i=1}^N \alpha_{T_k} \bullet \sigma \left(\frac{P_{T_k}(X)}{\tau} \right) \sigma \right) \quad (5)$$

Where, σ is the softmax function that converts model outputs into probabilities, and τ is a temperature parameter that modulates the output distribution's smoothness to enhance alignment learning.

The final total loss is calculated as follows:

$$L_{total} = \lambda L_{CE} + (1 - \lambda) L_{align} \quad (6)$$

Where, λ is a hyperparameter that adjusts the relative contributions of the cross-entropy loss and the alignment loss.

By jointly optimizing the above loss functions, the hybrid model can gradually learn and absorb knowledge from multiple teacher models while maintaining consistency with the true labels. This ultimately enhances the reasoning ability and performance of the hybrid model on specific tasks.

3 Experiments

3.1 Node Classification

Datasets. The experiment uses three commonly used citation network datasets for graph neural networks: Cora [5], CiteSeer [5], and PubMed [6], as well as two domain-specific graph datasets, A-Photo [7] and Physics [7]. The relevant statistics of the datasets are shown in Table 1.

Table 1. Dataset statistics information

Dataset	Node	Edge	Feature	Type
Cora	2708	5429	1433	7
CiteSeer	3312	4723	3703	6
PubMed	19717	44338	500	3
A-Photo	7487	119043	745	8
Physics	34493	247962	8415	5

Experimental Settings. The experiment was conducted on a computer equipped with an NVIDIA RTX A6000 GPU, using the DGL and PyTorch frameworks. For comparison, the Adam optimizer was used for 200 training epochs, with a batch size of 64 for both teacher and hybrid models. The learning rate was set to 0.001, weight decay to 0.005, and the temperature parameter to 5. To ensure a fair comparison, three experiments were performed with different seed values 0, 1, 2, and the average performance was calculated.

Evaluation Metrics. To evaluate the performance of HTM-KG, Micro-F1 was used as the metric for the node classification task. Micro-F1 is a variant of the F1 score, used to assess the performance of multi-class or multi-label classification models. It combines precision and recall, providing a balanced performance measure through their harmonic mean. The specific formula is as follows:

$$\text{Micro-F1} = 2 \times \frac{\text{precision} + \text{recall}}{\text{precision} \times \text{recall}} \quad (7)$$

Comparative Experimental Analysis. As shown in Table 2, HTM-KG outperforms all baseline methods, except for Graph-BERT, in Micro-F1 scores across five public datasets, demonstrating its advantages in node classification tasks. HTM-KG uses the optimal multi-teacher combination from the ablation experiment, selecting GAT and SAGE as teacher models, and GCN as the hybrid model.

Compared to single-teacher models such as LSP, CPF, GLNN, and methods by Jing et al., HTM-KG shows improvements across all five datasets. Single-teacher models have limited ability to capture both local structure and global semantic features, while, HTM-KG leverages the complementary strengths of multiple teacher models to achieve more comprehensive feature representations. Although MSKD and E2GNN, as multi-teacher knowledge transfer methods, perform well on some datasets, HTM-KG improves by 1.42% and 1.31% on CiteSeer and A-Photo, respectively. This is due to its dynamic weighting mechanism and the joint optimization of cross-entropy and KL divergence, which effectively enhance knowledge transfer and model robustness.

Graph-BERT, as the current strong baseline, outperforms all traditional GNN and multi-teacher methods across five datasets, showing its superior per-

formance relying on large-scale pre-trained models. However, Graph-BERT consumes significant computational resources and is less efficient than HTM-KG in resource-constrained scenarios. HTM-KG achieves results close to Graph-BERT without significantly increasing computational burden, making it particularly suitable for applications requiring computational efficiency or interpretability.

In summary, HTM-KG significantly increases node classification task performance. Through the collaboration of multiple teacher models, it improves the model’s ability to capture complex graph structures, thus boosting its generalization capability.

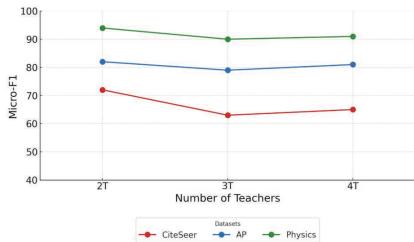
Table 2. Micro-F1 results of HTM-KG method and baseline methods on five datasets

Model	Dataset				
	Cora	CiteSeer	PubMed	A-Photo	Physics
LSP [8]	80.02	68.32	78.14	78.15	92.43
CPF [9]	80.59	68.79	78.36	79.23	92.56
GLNN [10]	75.82	66.80	72.34	84.56	91.78
Jing et al. [11]	78.69	68.52	73.87	83.95	92.14
MSKD [5]	81.62	69.68	77.50	85.23	92.68
E2GNN [12]	82.36	70.31	78.43	84.78	93.05
Graph-BERT [13]	84.3	71.2	79.3	94.5	94.9
HTM-KG	82.17	71.73	78.92	86.09	93.68

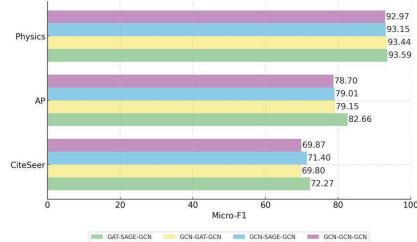
Ablation Experimental Analysis. To comprehensively evaluate the performance of HTM-KG in node classification, this chapter designs two ablation experiments and performs sensitivity analysis on the hyperparameters τ and λ , using Micro-F1 as the evaluation metric. The experiments use three datasets: Citeseer, A-Photo, and Physics, representing academic, social, and interdisciplinary graph data, respectively, with a focus on HTM-KG’s performance on these datasets. To study the impact of the number of teacher models on performance, the number of teachers was gradually increased from 2 to 4 (using GCN, GAT, and SAGE, with APPNP added). The results, shown in Fig. 2, reveal that the model performs best with 2 teachers, and adding more teachers leads to negative effects due to redundant information interfering with the training process. However, on the smaller Citeseer dataset, the model with 4 teachers performs better than the one with 3. Therefore, the optimal configuration for HTM-KG was set to 2 teachers and 3 layers.

After determining the optimal number of teachers, the impact of different teacher model combinations on performance was further explored. The results show that the hybrid model performs best when GAT and SAGE are used as teacher models, significantly outperforming variants that include GCN as a

teacher. This finding demonstrates that combining heterogeneous teacher models such as GAT and SAGE enables the hybrid model to learn more diverse features, revealing the positive impact of teacher model heterogeneity on performance and emphasizing the importance of diversity and complementarity when designing models.



(a) Ablation study on the number of teacher models



(b) Ablation study on different combinations of teacher models

Fig. 2. Ablation study on the number of teacher models and different combinations of teacher models.

3.2 Link Prediction

Datasets. This experiment uses two datasets: FB15k-237[14] and WN18RR[15], with their statistical data shown in Table 3.

Table 3. Dataset statistics

Dataset	Entity	Relation	Training	Validation	Test
FB15k-237	14541	237	272115	17535	20466
WN18RR	40943	11	86835	3034	3134

Experimental Settings. The method is built on the PyTorch framework and implemented on a GPU-equipped computer with an NVIDIA RTX A6000. A grid search over the learning rate range of $\{3 \times 10^{-5}, 5 \times 10^{-5}\}$ is performed, and the AdamW optimizer with a learning rate of 0.01 is used, combined with a linear learning rate decay strategy. The batch size for both the teacher and hybrid models is set to 64, with 200 epochs.

Comparative Experimental Analysis. Table 4 presents the performance of HTM-KG compared to eight baseline methods on the WN18RR and FB15k-237 datasets. The results show that HTM-KG outperforms most baseline methods on all metrics in the WN18RR dataset, and also performs well on the FB15k-237 dataset, demonstrating its adaptability in complex knowledge graphs.

On WN18RR, HTM-KG improves by 26.9% over traditional embedding methods and by 25.3% over pre-trained language model-based methods, showing its advantage in capturing complex semantics and graph structures. Although the gap with PD-MGF is small (2.4%), HTM-KG shows more stability in Hits@k. On the FB15k-237 dataset, HTM-KG performs best in MRR, especially compared to pre-trained models. However, it slightly lags behind embedding-based methods in Hits@3 and Hits@10. HTM-KG still demonstrates strong generalization ability, with performance comparable to other advanced models, especially in handling complex relationships.

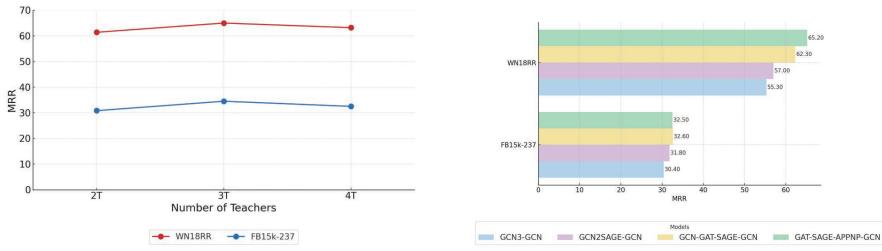
SimKGC, as a strong baseline, outperforms traditional methods and HTM-KG on several datasets, particularly achieving a higher MRR on WN18RR. However, SimKGC mainly relies on contrastive learning and struggles to fully capture complex graph structures and semantic features. In contrast, HTM-KG, through multi-teacher model fusion, better adapts to complex graph data and achieves results close to SimKGC without significantly increasing computational cost, making it particularly suitable for multi-relational and complex semantic scenarios.

Table 4. Overall performance comparison of the HTM-KG method in the link prediction task.

Model	Dataset							
	WN18RR				FB15k-237			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
TransE [16]	24.3	4.3	44.1	53.2	27.9	19.8	37.6	44.1
ConvE [15]	43.0	40.0	44.0	52.0	32.5	23.7	35.6	50.1
RotatE [17]	47.6	42.8	49.2	57.1	33.8	24.1	37.5	53.3
MTL-KGC [18]	33.1	20.3	38.3	59.7	26.7	17.2	29.8	45.8
KG-BERT [19]	41.1	31.6	43.5	55.3	26.9	18.8	28.3	42.3
KGLM [20]	46.7	33.0	53.8	74.1	28.9	20.0	31.4	46.8
PD-MGF [21]	62.8	52.9	69.5	80.4	31.3	23.3	33.9	48.6
SimKGC [22]	66.6	58.7	71.7	80.0	33.6	24.9	36.2	51.1
HTM-KG	63.5	54.7	70.3	80.9	32.4	23.9	35.4	50.7

Ablation Experimental Analysis. To comprehensively evaluate the performance advantages of HTM-KG, this chapter designs two ablation experiments

based on the link prediction task and analyzes the impact of the number of teachers on model performance using the MRR metric. The experiments were conducted on the WN18RR and FB15k-237 datasets, with the number of teachers gradually increased from 2 to 4 (including GCN, GAT, SAGE, and APPNP as teacher models). Figure 3a displays the results that show that when there are 3 teachers on both datasetsthe model performs best. However, when the number of teachers increases to 4, the MRR slightly decreases. This may be due to the introduction of redundant or irrelevant information from too many teachers, which interferes with the training of the student model. Considering both the efficiency and effectiveness of knowledge transfer, the final configuration sets the number of teachers to 3 and the number of layers to 9.



(a) Ablation study on the number of teacher models

(b) Ablation study on different combinations of teacher models

Fig. 3. Ablation study on the number of teacher models and different combinations of teacher models.

Based on the previous finding that the model performs best with 3 teachers, this study further investigates the impact of different graph neural network combinations on HTM-KG’s performance through ablation experiments, quantitatively analyzing the contribution of different GNN structures to the hybrid model. The experiments were conducted on the FB15k-237 and WN18RR datasets, and the results shown in Fig. 3b indicate that the GAT-SAGE-APPNP-GCN combination (with GAT, SAGE, and APPNP as teacher models) performs best, significantly outperforming other combinations. Further analysis reveals that the superiority of GAT-SAGE-APPNP-GCN stems from the high heterogeneity of its teacher models: GAT’s attention mechanism excels at capturing node relationship weights, SAGE’s sampling and aggregation strategy enhances the model’s adaptability to large graphs, and APPNP’s personalized propagation mechanism optimizes information transfer efficiency. This diverse combination allows the hybrid model to more comprehensively learn graph structures and semantic features. When designing model architectures, heterogeneous teacher combinations should be prioritized to enhance adaptability and robustness.

Building on this optimal teacher selection, further ablation experiments explore the impact of the temperature parameter τ and the weight parameter λ

between L_{CE} and L_{align} in the hybrid model alignment loss function on training performance. The experiments, conducted on the WN18RR and FB15k-237 datasets with different τ and λ values, are shown in Fig. 4. The results indicate that the choice of temperature and λ hyperparameters has a significant impact on HTM-KG's link prediction performance, particularly across different datasets. Properly setting the temperature and λ can significantly enhance the model's prediction ability, especially when handling complex graph structures.

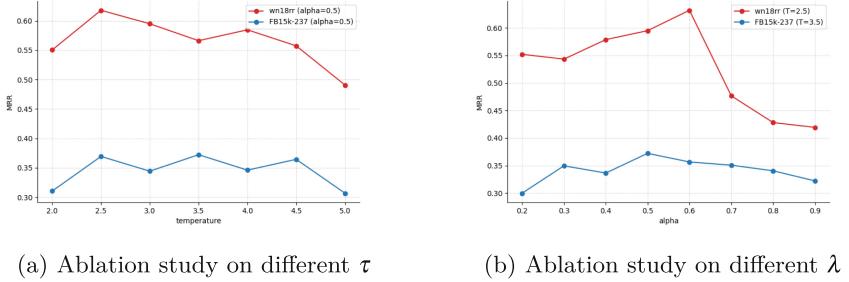


Fig. 4. Ablation study on different τ and λ

4 Conclusion

This paper proposes a knowledge graph node classification and link prediction method based on a hybrid teacher model. First, the relevant background and definitions of knowledge graph node classification and link prediction are introduced, emphasizing their importance in the field of knowledge question answering. Next, the multi-teacher model ensemble architecture is described in detail, including the overall framework, the functions of each module, and the methods used to assess performance in multi-task learning scenarios. The classification and prediction model based on the hybrid teacher model is specifically introduced, covering the design and implementation of the multi-teacher model module, the teacher weight allocation module, and the hybrid model alignment module. The paper focuses on how the model achieves knowledge fusion and alignment through alignment loss and cross-entropy loss. Ultimately the method's efficacy is proved by experiments using node categorization and link prediction. The model's performance gains across multiple datasets are assessed, and the impact of individual modules on overall effectiveness is systematically examined. The experimental results show that this method has significant advantages in both tasks, especially excelling in handling complex relationships.

References

1. Zhang, B., Soh, H.: Extract, define, canonicalize: an llm-based framework for knowledge graph construction. arXiv preprint [arXiv:2404.03868](https://arxiv.org/abs/2404.03868) (2024)

2. Chen, H., Shen, X., Lv, Q., et al.: Sac-kg: exploiting large language models as skilled automatic constructors for domain knowledge graphs. arXiv preprint [arXiv:2410.02811](https://arxiv.org/abs/2410.02811) (2024)
3. Asif, U., Tang, J., Harrer, S.: Ensemble Knowledge Distillation for Learning Improved and Efficient Networks, pp. 953–960. ECAI. IOS Press (2020)
4. Shen, C., Wang, X., Song, J., et al.: Amalgamating knowledge towards comprehensive classification. In: Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA, vol. 33(01), pp. 3068–3075 (2019)
5. Zhang, C., Liu, J., Dang, K., et al.: Multi-scale distillation from multiple graph neural networks. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36(4), pp. 4337–4344 (2022)
6. Namata, G., London, B., Getoor, L., et al.: Query-driven active surveying for collective classification. In: 10th international Workshop on Mining and Learning with Graphs, vol. 8 (2012)
7. Shchur, O., Mumme, M., Bojchevski, A., et al.: Pitfalls of graph neural network evaluation. arXiv preprint [arXiv:1811.05868](https://arxiv.org/abs/1811.05868) (2018)
8. Yang, Y., Qiu, J., Song, M., et al.: Distilling knowledge from graph convolutional networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Seattle, Washington, USA, pp. 7074–7083 (2020)
9. Yang, C., Liu, J., Shi, C.: Extract the knowledge of Zhang S, Liu Y, Sun Y, et al. Graph-less neural networks: teaching old mlps new tricks via distillation. arXiv preprint [arXiv:2110.08727](https://arxiv.org/abs/2110.08727) (2021)
10. Zhang, S., Liu, Y., Sun, Y., et al.: Graph-less neural networks: teaching old MLPS new tricks via distillation. arXiv preprint [arXiv:2110.08727](https://arxiv.org/abs/2110.08727) (2021)
11. Jing, Y., Yang, Y., Wang, X., et al.: Amalgamating knowledge from heterogeneous graph neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, Tennessee, USA, pp. 15709–15718 (2021)
12. Zhang, X., Zha, D., Tan, Q.: E2GNN: efficient graph neural network ensembles for semi-supervised classification. arXiv preprint [arXiv:2405.03401](https://arxiv.org/abs/2405.03401) (2024)
13. Zhang, J., Zhang, H., Xia, C., et al.: Graph-bert: only attention is needed for learning graph representations. arXiv preprint [arXiv:2001.05140](https://arxiv.org/abs/2001.05140) (2020)
14. Toutanova, K., Chen, D.: Observed versus latent features for knowledge base and text inference. In: Proceedings of the 3rd Workshop on Continuous Vector Space Models and Their Compositionality, pp. 57–66 (2015)
15. Dettmers, T., Minervini, P., Stenetorp, P., et al.: Convolutional 2D knowledge graph embeddings. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32(1) (2018)
16. Bordes, A., Usunier, N., Garcia-Duran, A., et al.: Translating embeddings for modeling multi-relational data. *Adv. Neural Inform. Process. Syst.* **26** (2023)
17. Sun, Z., Deng, Z.H., Nie, J.Y., et al.: Rotate: knowledge graph embedding by relational rotation in complex space. arXiv preprint [arXiv:1902.10197](https://arxiv.org/abs/1902.10197) (2019)
18. Ma, T., Lin, X., Song, B., et al.: KG-MTL: knowledge graph enhanced multi-task learning for molecular interaction. *IEEE Trans. Knowl. Data Eng.* **35**(7), 7068–7081 (2022)
19. Yao, L., Mao, C., Luo, Y.: KG-BERT: BERT for knowledge graph completion. arXiv preprint [arXiv:1909.03193](https://arxiv.org/abs/1909.03193) (2019)
20. Youn, J., Tagkopoulos, I.: KGLM: integrating knowledge graph structure in language models for link prediction. arXiv preprint [arXiv:2211.02744](https://arxiv.org/abs/2211.02744) (2022)

21. Fan, C., Chen, Y., Xue, J., et al.: Progressive distillation based on masked generation feature method for knowledge graph completion. In: Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, California, USA, vol. 38(8), pp. 8380–8388 (2024)
22. Wang, L., Zhao, W., Wei, Z., et al.: Simkgc: simple contrastive knowledge graph completion with pre-trained language models. arXiv preprint [arXiv:2203.02167](https://arxiv.org/abs/2203.02167) (2022)



Optimized On-Off Control of Artificial Climate Chambers via Physics-Neural Hybrid Modeling and DQN

Han Zeng^{1,2}, Linli Zhou^{1,2(✉)}, Dayong Wang³, and Caiguo Tang^{1,2}

¹ Intelligent Agriculture Engineering Laboratory of Anhui Province, Institute of Intelligent Machines, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, People's Republic of China

linlizhou@iim.ac.cn

² University of Science and Technology of China, Hefei 230026, People's Republic of China

³ School of Information and Artificial Intelligence, Anhui Agricultural University, Hefei 230036, People's Republic of China

Abstract. To address the issues of energy consumption and environmental instability caused by frequent switching of on-off control devices in artificial climate chambers, this paper proposes an intelligent control framework that integrates physical constraints with deep reinforcement learning. A Dual-Path Prediction Network (DPP) is constructed to build a high-precision simulation environment for the climate chamber, achieving comprehensive environmental prediction with a modeling accuracy of $R^2 = 0.9979$ and $MSE = 0.0016$. Based on this, a discrete control strategy is designed using a Deep Q-Network (DQN), with a composite reward function that incorporates penalties for device switching. Experimental results show that, compared to traditional threshold-based control, the proposed method reduces the switching frequency of the heating plate and compressor by 78.4% and 72.5%, respectively, while maintaining temperature accuracy within ± 0.5 °C. This study provides an intelligent control paradigm for on-off control systems in climate chambers, offering practical value in reducing operational costs and enhancing environmental stability.

Keywords: On-off control · Deep Reinforcement Learning · Artificial climate chamber

1 Introduction

The artificial climate chamber is widely used in modern biological cultivation due to its cost-effectiveness and ease of implementation, particularly with its simple on-off control design [1]. As shown in Fig. 1, the chamber used in this study features a sealed vertical cabinet with transparent front doors, a bottom ventilation outlet, and an integrated control panel displaying temperature, humidity,

CO concentration, and light intensity. Internally, it includes heating plates, a compressor, a humidifier, a CO regulator, fans, and LED lights—all controlled through basic on/off switching.

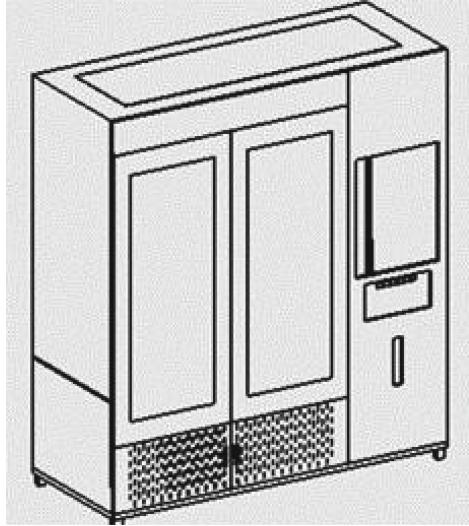


Fig. 1. Climate chamber structure.

Currently, threshold control algorithms are commonly used, switching devices based on preset upper and lower limits [2]. While simple, this method has two major drawbacks: (1) frequent switching increases energy use and mechanical wear [3]; (2) nonlinear interactions between multiple devices make single-variable control ineffective, especially for coordinated temperature-humidity regulation [4].

Deep Reinforcement Learning (DRL) has recently shown great potential in complex control systems [5–7]. Unlike traditional PID or MPC methods [9], DRL learns control strategies through interaction and multi-objective optimization. It has achieved up to 37% energy savings in HVAC applications [8]. However, applying DRL to on-off control in climate chambers faces two key challenges: (1) real-device training is risky and time-consuming, requiring a reliable simulation environment; (2) discrete actions and delayed responses make learning inefficient.

To solve these issues, we propose a DRL-based intelligent control framework using a Deep Q-Network (DQN) [10,11]. First, we construct a physics-constrained neural network simulator (Fig. 2) using historical data, achieving high prediction accuracy ($R^2 = 0.9994$, see Table 1) to support training. Second, a composite reward function is designed to balance temperature tracking accuracy and device switching cost, enabling efficient and stable control.

2 Climate Chamber Simulation Environment Modeling

2.1 Dual Path Prediction Network

In artificial climate chambers, purely physics-based models often struggle to accurately capture real environmental changes triggered by device switching. To address this, we propose the DPP that integrates physical constraints with neural network capabilities to predict the next environmental state based on control device inputs.

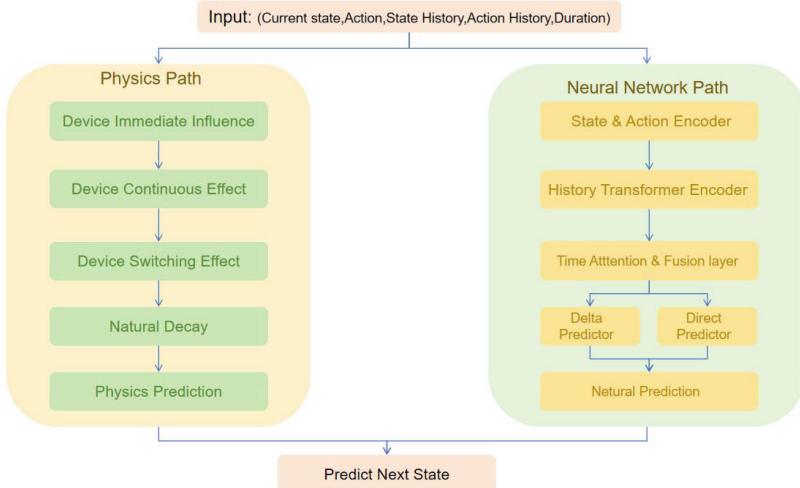


Fig. 2. Dual path prediction network structure.

As illustrated in Fig. 2, the model adopts a dual-path architecture:(1)The physics-guided path encodes domain knowledge and physical rules, modeling immediate influence,continuous effect, device switching effect, and natural decay.(2)The neural network path learns complex patterns from historical data, consisting of state/action encoders, a sequence transformer, temporal attention and fusion layers, and both incremental and direct predictors.

The model takes as input the current state, current action, historical states and actions, and device duration. The outputs from both paths are fused using a confidence-weighted mechanism to generate the final prediction of the next environmental state. This design ensures both high predictive accuracy and physical consistency.

The DPP model can be formally defined as:

$$\mathbf{s}_{t+1} = f_{\text{DPP}}(\mathbf{s}_t, \mathbf{a}_t, \mathbf{S}_{t-h:t-1}, \mathbf{A}_{t-h:t-1}, \mathbf{d}_t) \quad (1)$$

where $\mathbf{s}_t \in \mathbb{R}^{n_s}$ is the current state vector consisting of n_s environmental variables, $\mathbf{a}_t \in \mathbb{R}^{n_a}$ is the action vector representing n_a on-off control devices, $\mathbf{S}_{t-h:t-1}$ and $\mathbf{A}_{t-h:t-1}$ denote the historical state and action sequences of length h , and \mathbf{d}_t is a vector indicating the activation durations of each device.

2.2 Training of the DPP

The training data was collected from an artificial climate chamber operating continuously over a period of three months. The system sampled data every 5 s, synchronously recording four types of environmental variables—temperature, humidity, CO concentration, and light intensity—along with the on-off states of control devices. After applying sliding-window normalization and outlier removal, a time-series dataset containing 1.19 million samples was constructed.

Model training was conducted using a simulated annealing strategy, with an initial learning rate of 0.005 and a cosine annealing scheduler. The loss function consists of a prediction loss and a consistency loss:

$$\mathcal{L} = \mathcal{L}_{\text{pred}}(\mathbf{s}_{t+1}, \mathbf{s}_{t+1}^{\text{true}}) + \omega \cdot \mathcal{L}_{\text{cons}}(\mathbf{s}_{t+1}^{\text{phy}}, \mathbf{s}_{t+1}^{\text{neural}}) \quad (2)$$

where ω is a hyperparameter weighting the consistency term, and $\mathcal{L}_{\text{cons}}$ denotes the mean squared error between the outputs of the physical path and the neural path. The training loss curve is shown in Fig. 3.

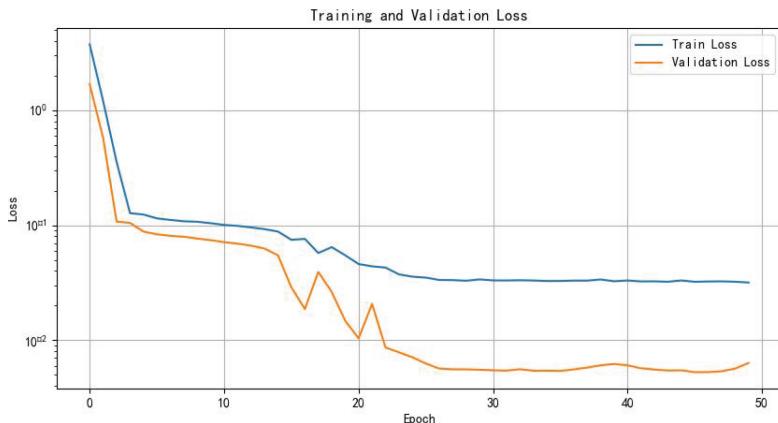


Fig. 3. loss curve for DDP training.

We used data from a full day of operation of the artificial climate chamber as the test set. The model's predictive performance on key environmental factors was evaluated using MSE and R^2 . The experimental results, as shown in Table 1, each environment variable predicts very well. The overall model performance metrics show an MSE of 0.0016 and an average R^2 of 0.9979. The results indicate that the DPP model can accurately simulate the dynamic characteristics of the climate chamber, providing a reliable environment for subsequent deep reinforcement learning strategy training.

Table 1. Test set prediction accuracy metrics for environmental factors

Factor	MSE	R^2
Temperature	0.001597	0.9994
Humidity	0.087929	0.9974
CO ₂	0.915838	0.9983
Illumination	0.000066	0.9965
Overall	0.251358	0.9979

3 Deep Reinforcement Learning

3.1 Designing for Deep Reinforcement Learning

To achieve precise temperature control and optimize the frequency of device switching in the climate chamber environment, we adopt the DQN approach [10, 11]. This method enables the agent to learn a state-to-action mapping policy through continuous interaction with the environment, thereby optimizing both environmental objectives and device operation costs.

The temperature regulation task is formalized as a Markov Decision Process (MDP) defined by a 5-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where:

\mathcal{S} : the state space, consisting of the current environmental state of the climate chamber and the status of control devices;

\mathcal{A} : the action space, representing the on/off control actions for each device;

\mathcal{P} : the state transition function, simulated by our custom-designed climate chamber environment model, reflecting the dynamic evolution of temperature in response to specific control actions;

\mathcal{R} : the reward function, representing the combined effect of control actions on temperature stability and the frequency of device switching;

γ : the discount factor, which balances immediate rewards and long-term control strategies.

Based on the DQN framework, a feedforward neural network is used to approximate the action-value function $Q(s, a; \theta)$. The network architecture consists of two hidden layers with ReLU activations, containing 128 and 64 neurons respectively. The input is the current environmental state, and the output is the estimated Q-values for each possible action. The agent selects the optimal action by maximizing the Q-value, guiding the control of the devices.

To improve training stability and sample efficiency, we incorporate the Experience Replay mechanism [12] and a Target Network. At each training step, a mini-batch of transition samples is drawn from the replay buffer to update the network parameters. The target Q-value is calculated using the target network with fixed parameters:

$$y_t = r_t + \gamma \cdot \max_{a'} Q(s_{t+1}, a'; \theta^-) \quad (3)$$

The current network is then optimized by minimizing the following loss function:

$$\mathcal{L}(\theta) = \mathbb{E} \left[(y_t - Q(s_t, a_t; \theta))^2 \right] \quad (4)$$

Additionally, an ϵ -greedy strategy is adopted to balance exploration and exploitation. The value of ϵ gradually decays from an initial value to a stable minimum during training, helping the agent converge to the optimal policy.

To simultaneously achieve stable temperature control and minimized switching of devices, we design the following composite reward function:

$$r_t = -\alpha \cdot |T_t - T_{\text{target}}| - \beta \cdot \left(\mathbb{I}_{[a_t^{\text{heater}} \neq a_{t-1}^{\text{heater}}]} + \mathbb{I}_{[a_t^{\text{cooler}} \neq a_{t-1}^{\text{cooler}}]} \right) \quad (5)$$

where:

T_t is the current temperature, and T_{target} is the target temperature;
 a_t^{heat} and a_t^{cool} represent the on/off states of the heater and cooler, respectively;

$\mathbb{I}_{[\cdot]}$ is an indicator function that equals 1 when a device changes its switching state;

α, β are adjustable weighting coefficients used to balance temperature stability and control costs.

This reward function encourages the agent to maintain the temperature close to the target while penalizing frequent switching of devices, thereby improving both the stability and energy efficiency of the control strategy.

3.2 Training Results of Deep Reinforcement Learning

The training results are shown in Fig. 4, illustrating the evolution of four key indicators during the reinforcement learning process: DQN Score, Device Switching Count, Average Temperature Error, and Exploration Rate.

In the DQN Score plot, we observe significant negative fluctuations during the early training phase. Subsequently, the model rapidly improves, stabilizing after approximately 50 training episodes and maintaining a score close to zero. This indicates that the agent has learned an effective control policy.

The Device Switching Count plot displays the switching frequency of the heating (blue line) and cooling (orange line) devices. In the initial phase—particularly within the first 100 training episodes—the heating device exhibits large fluctuations, peaking at around 110 switches. As training progresses, although the switching count of the heater remains relatively volatile, the overall trend becomes more stable, mostly fluctuating between 20 and 60. In contrast, the switching frequency of the cooling device remains more consistent, staying around 40 switches.

The Average Temperature Error plot reflects the accuracy of the model's temperature control. In the early training stage, the error peaks at around 16 °C, then rapidly decreases. Notably, after around 600 episodes, the error stabilizes below 1 °C, indicating that the model has achieved a high degree of temperature control precision.

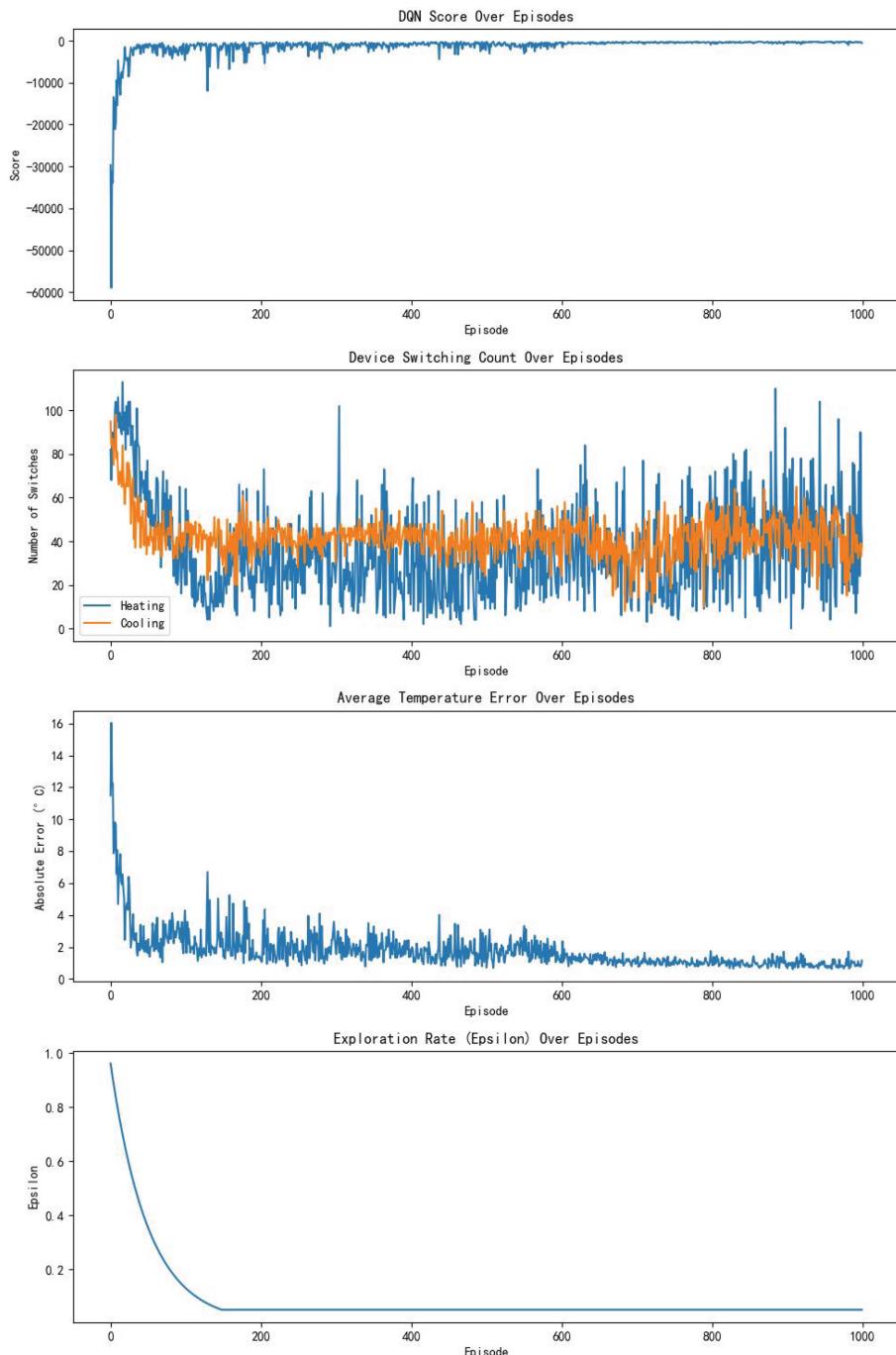


Fig. 4. DQN training results.

The Exploration Rate (ε) plot shows the model's transition from exploration to exploitation. Initially, ε is close to 1, representing high exploratory behavior. As training progresses, ε quickly decays and stabilizes at around 0.05 after approximately 200 episodes, suggesting that the model has shifted from random exploration to exploiting the learned policy.

These training curves demonstrate that our reinforcement learning model has successfully learned a temperature control policy. After sufficient training, it is capable of maintaining a low temperature error and reasonable device switching frequency while achieving a stable reward signal. This training pattern aligns with the typical convergence behavior in deep reinforcement learning, where early instability and high exploration gradually give way to stable performance.

4 Experiments and Results

To assess the effectiveness of the reinforcement learning strategy for climate box temperature control, we conducted experiments in a simulated environment. The primary performance metrics were heater/cooler switching frequency and temperature stability. The target temperature was set to 23 °C over 600 time steps, with the first 60 steps initialized using real-world data.

Figure 5a and b compare the DQN-based control and traditional threshold control. Each figure includes three subplots: temperature (top), heater status (middle), and cooler status (bottom). The red solid line shows actual temperature; the dashed line indicates the target. Square waves represent device on/off states (1 for on, 0 for off).

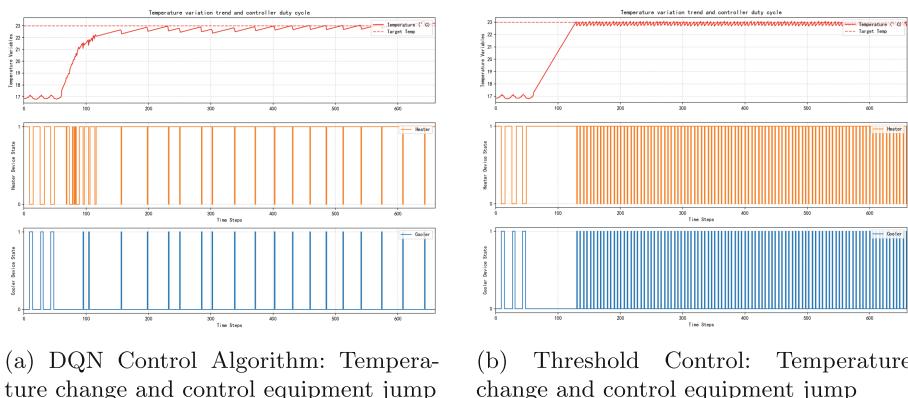


Fig. 5. Comparison of two control algorithms.

Temperature Control: The DQN strategy (Fig. 5a) stabilizes near 23 °C within 100 steps, with fluctuations contained within ± 0.5 °C. In contrast, threshold control (Fig. 5b) reaches the target later and shows persistent sawtooth-like oscillations, indicating unstable regulation.

Device Behavior: With DQN, the heater is activated more frequently early on, then enters a low-frequency mode, while the cooler operates minimally and both devices avoid simultaneous operation. Threshold control, however, causes high-frequency switching and alternating activation—exacerbating the “ping-pong effect” and increasing wear and energy use.

As shown in Table 2, when the target temperature is set to 23 °C, the temperature deviations at the 600th time step for the threshold-based and DQN-based control algorithms are 0.19 °C and 0.23 °C, respectively.

In terms of control device switching frequency, under the DQN control strategy, the total number of on-off switches for the heater and cooler were 56 and 44 times, respectively. In contrast, under the threshold-based control strategy, both the heater and cooler were switched on and off 204 times.

The DQN strategy reduced the number of switches for both the heating and cooling devices by more than 70%, significantly lowering the control system’s operational complexity and reducing hardware stress.

Table 2. Control performance comparison

Method	Temp@600 (°C)	Heater switches	Cooler switches
Threshold Control	22.81	204	204
DQN Control	23.23	44 (↓78.4%)	56 (↓72.5%)

5 Conclusion

This paper proposes a dual-path climate chamber simulation environment that integrates physical principles with neural networks, combined with a DQN-based intelligent control framework. The approach effectively addresses the issue of frequent on-off switching of devices in traditional threshold-based control systems for on/off-type climate chambers.

By employing a composite reward mechanism that penalizes device switching, the system maintains temperature stability within ± 0.5 °C while reducing the switching frequency of heating and cooling devices by more than 70%, significantly mitigating the “ping-pong effect.”

The main innovations are as follows: (1) Construction of a simulation environment that fuses physical constraints with neural networks, effectively narrowing the sim-to-real gap; (2) Introduction of a device switching penalty term, leading to a composite reward function that jointly optimizes temperature tracking accuracy and control cost; (3) Validation of the effectiveness of deep reinforcement learning in multi-device discrete control scenarios.

Future work will extend to multivariable coordinated control.

Acknowledgments and Disclosure of Funding. This work was supported by the Anhui Provincial Key Research and Development Project (2023n06020028).

References

1. Sánchez, C., Dessì, P., Duffy, M., et al.: OpenTCC: an open source low-cost temperature-control chamber. *HardwareX* **7**, e00099 (2020). <https://doi.org/10.1016/j.ohx.2020.e00099>
2. Tang, L., Zhao, J.: Switched threshold-based fault detection for switched nonlinear systems with its application to Chua's circuit system. In: *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 66, no. 2, pp. 733–741 (2019).<https://ieeexplore.ieee.org/document/8335806/>
3. Ma, K., Yuan, C., Yang, J., Liu, Z., et al.: Switched control strategies of aggregated commercial HVAC systems for demand response in smart grids. *Energies* (2017).<https://doi.org/10.3390/en10070953>
4. Behrooz, F., Mariun, N., et al.: Review of control techniques for HVAC systems—nonlinearity approaches based on fuzzy cognitive maps. *Energies* **11**(3), 495 (2018).<https://doi.org/10.3390/en11030495>
5. Bucci, M.A., Semeraro, O., Allauzen, A., et al.: Control of chaotic systems by deep reinforcement learning. *Proc. R. Soc. A* **475** (2231), 20190351 (2019).<https://doi.org/10.1098/rspa.2019.0351>
6. Wang, L., He, X., Luo, D.: Deep reinforcement learning for greenhouse climate control. In: *2020 IEEE International Conference on Knowledge Graph (ICKG)*. IEEE (2020).<https://doi.org/10.1109/ICBK50248.2020.00073>
7. Morcego, B., et al.: Reinforcement learning versus model predictive control on greenhouse climate control. *Comput. Electron. Agric.* **215**, 108372 (2023).<https://doi.org/10.1016/j.compag.2023.108372>
8. Wang, H., et al.: Energy optimization for HVAC systems in multi-VAV open offices: a deep reinforcement learning approach. *Appl. Energy* **356**, 122354 (2024).<https://doi.org/10.1016/j.apenergy.2023.122354>
9. Mayne, D.Q.: Model predictive control: Recent developments and future promise. *Automatica* **50**(12), 2967–2986 (2014). <https://doi.org/10.1016/j.automatica.2014.10.128>
10. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Playing atari with deep reinforcement learning (2013).<https://arxiv.org/pdf/1312.05602.pdf>
11. Mnih, V., Kavukcuoglu, K., Silver, D. et al.: Human-level control through deep reinforcement learning. *Nature* **518**, 529–533 (2015). <https://doi.org/10.1038/nature14236>
12. Lin, L.-J.: Self-improving reactive agents based on reinforcement learning, planning and teaching. *Mach. Learn.* **8**, 293–321 (1992).<https://doi.org/10.1007/BF00992699>



Parking Path Planning Method Based on Improved Hybrid A* Algorithm

Jiaming Li, Bin Lan, Lunfei Liang, Gang Yan^(✉), and Jiahao He

JiangHuai Advanced Technology Center, HeFei 230000, China
yangang1010@126.com

Abstract. To solve the problems of inefficient expansion and time-consuming planning of the traditional hybrid A* algorithm in parking scenarios, an improved Hybrid A* path planning method with dynamic adjustment of sampling arc length is proposed. When the algorithm is executed, the number of collision-free child nodes of the parent node is used to represent the distribution of environmental obstacles, and the sampling step length of the subsequent node expansion is corrected in real-time based on this. This not only improves the node expansion efficiency of the vehicle in the open scene outside the garage but also takes into account the obstacle avoidance requirements of the complex environment near the garage. ROS simulation experiments show that the improved hybrid A* algorithm has similar parking planning path lengths, but has obvious advantages in search time, expansion nodes, and running time, which can effectively solve the path planning problem in the parking field.

Keywords: Hybrid A* Algorithm · Path Planning · Autonomous Parking

1 Introduction

With the steady progress of economic construction, the number of cars in China has increased yearly. However, due to the high threshold of parking technology and insufficient construction of urban parking lots, “difficult parking” has become a difficult problem that needs to be solved urgently in the process of urbanization. Autonomous parking technology can realize automatic parking and one-click parking of smart cars without human intervention, which can effectively alleviate traffic congestion and parking difficulties [1].

Path planning is an important part of the automatic parking system. The mainstream parking path planning methods are divided into search-based algorithms, sampling-based algorithms, and optimization-based algorithms [2, 3]. The sampling-based algorithm constructs a feasible road network through free space sampling and then uses the Dijkstra algorithm [4] to complete the optimal path search, but the algorithm consistency is poor and it is difficult to deal with narrow channels [5, 6]. The optimization-based algorithm transforms the path planning problem into an optimal control problem and obtains the unmanned vehicle trajectory by solving the optimal value of the objective function that satisfies the constraints. The trajectory solved by the algorithm has high quality, but it

is difficult to handle obstacles equivalently and the optimization process depends on good initial value input [7, 8]. The search-based algorithm is divided into two steps: environmental map construction and path search, mainly including the A* algorithm and the Dijkstra algorithm, which can realize the shortest path search of the grid map, but the path has many turns, which is not conducive to tracking control [9, 10].

The Hybrid A* algorithm [11] integrates the incomplete constraints of the unmanned vehicle on the basis of the traditional A* algorithm and can plan a continuous feasible path that meets the initial and final posture conditions [12]. Some scholars have made a lot of improvements to the Hybrid A* algorithm [13–16]. Qin [17] et al. incorporated the idea of jump point search into the Hybrid A* algorithm, reducing the search time of U-turns; SHENG [18] et al. integrated the traditional hybrid A* algorithm and the STC algorithm to solve the driving planning problem with unstructured narrow channels, but the path planning efficiency is low; LI [19] et al. completed the variable curvature planning of the Hybrid A* algorithm based on the road curvature, but there are too many search nodes.

To solve the problems of the traditional Hybrid A* algorithm, such as too many search nodes, low computational efficiency, and time-consuming collision detection, this paper proposes an adaptive Hybrid A* algorithm that dynamically determines the sampling time according to the number of non-collision child nodes of the parent node. Based on the ROS operating system, simulation tests are carried out in parallel parking scenarios to verify the effectiveness of this method.

2 Problem Description

2.1 Vehicle Kinematics Model

During parking operations, the driving speed of the smart car is lower than 5 km/h. Ignore the irregular parts on the surface of the car body and simplify it into a rigid rectangle. Based on the Ackerman steering principle [20], the vehicle kinematic model of the low-speed state is established as shown in Fig. 1. The equivalent rectangle is composed of the left front vertex A, the right front vertex B, the left rear vertex C, and the right rear vertex D. The center coordinates of the rear axle of the vehicle are (x_r, y_r) , the center coordinates of the front axle are (x_f, y_f) , the wheelbase of the vehicle is l , the front overhang length is l_f , the rear overhang length is l_r , the body width is W , the vehicle heading angle is θ , the equivalent front wheel is located at the center of the front axle, the front wheel turning angle is represented by δ , and the center speed of the rear axle of the vehicle is v_r . The vehicle state quantities are the center position and heading angle of the rear axle, and the control quantities are the center speed of the rear axle and the front wheel turning angle. During the movement of the vehicle, the center of the rear axle has no displacement and rotation relative to the body, so the kinematic equation is established with the center of the rear axle of the vehicle as the reference point:

$$\begin{bmatrix} \dot{x}_r \\ \dot{y}_r \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta \\ \sin \theta \\ \frac{\tan \delta}{l} \end{bmatrix} v_r \quad (1)$$

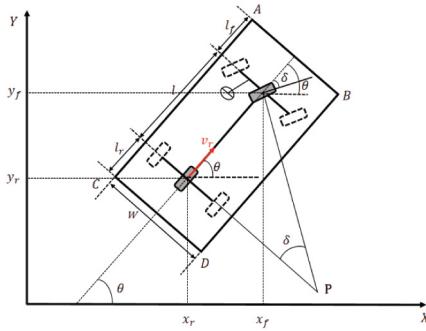


Fig. 1. Vehicle kinematics model

2.2 Traditional Hybrid A* Algorithm

The A* algorithm combines the advantages of the Dijkstra algorithm and the best search algorithm and can complete the search for the optimal path at a relatively fast speed. It is widely used in global path planning. $f(n)$ represents the total cost between the starting point and the current node, $g(n)$ represents the actual cost between the starting point and the current node, and $h(n)$ represents the estimated cost between the current node and the target point. The calculation formula for the path node cost function is:

$$f(n) = g(n) + h(n) \quad (2)$$

To alleviate the problems of the A* algorithm, such as many path turns and failure to meet the vehicle's non-holonomic constraints, the Hybrid A* algorithm improves the search dimension, node expansion method, actual cost, and heuristic function based on the algorithm, making it easier to plan a smooth path with complete probability that meets the vehicle's non-holonomic constraints:

1. Search dimension: The Hybrid A* algorithm introduces the node heading angle, which increases the search dimension from two dimensions to three dimensions (x, y, θ) .
2. Node expansion method (Fig. 2): The Hybrid A* algorithm considers the minimum turning radius of the vehicle and generates successor nodes in the left front, front, right front, left rear, rear, and right rear directions based on the parent node posture. The child nodes can be located at any position in the grid.
3. Actual node cost: The Hybrid A* algorithm introduces corresponding penalty factors for reversing and turning operations during node expansion to ensure the generation of a smoother path.
4. Node heuristic cost: The Hybrid A* algorithm takes into account the non-holonomic constraints of the vehicle when calculating the heuristic cost of the successor node and the target node. $h_{\text{obstacles}}$ represents the shortest path between nodes that considers the distribution of obstacles in the surrounding environment but ignores the non-holonomic constraints of the vehicle. With the endpoint as the origin, the reverse

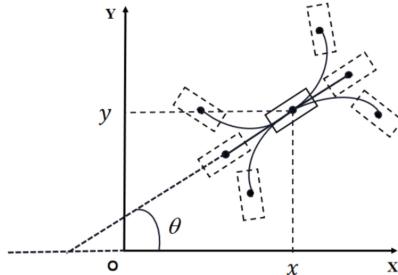


Fig. 2. Hybrid A* algorithm node expansion method. The parent node position and orientation are represented as (x, y, θ) , and six successor nodes are generated considering the minimum turning radius of the vehicle.

Dijkstra algorithm can be used to obtain the value $h_{\text{obstacles}}$ of each grid in the grid map. $h_{\text{non-holonomic}}$ is the heuristic cost that considers the non-holonomic constraints of the vehicle but ignores environmental obstacles. The Dubins curve or Reeds-Shepp curve is used to generate the path between the start and end points, and the curve length is the heuristic cost.

$$h(n) = \max(h_{\text{obstacles}}, h_{\text{non-holonomic}}) \quad (3)$$

3 Improved Hybrid A* Algorithm

The Hybrid A* algorithm can generate the shortest path connecting the starting and ending positions that satisfy the non-holonomic constraints of the vehicle. However, when the algorithm is running, the six successor nodes of the parent node need to perform collision detection with all obstacles, which is computationally intensive and time-consuming. At the same time, the use of a fixed step size to generate successor child nodes in an open environment with sparse obstacles slows down the algorithm expansion efficiency. Therefore, an adaptive sampling step size is introduced when the successor nodes of the Hybrid A* algorithm are expanded to improve the efficiency and real-time performance of path planning.

The parent node generates six successor nodes in the expansion process. Each child node needs to perform collision detection. Since the six child nodes have different postures, the number of collision-free successor nodes reflects the distribution of obstacles of the parent node in the local environment (Fig. 3). When the node is expanded to generate successor nodes, the number of safe child nodes of its parent node is first obtained, and the number of nodes is used as the independent variable of the adaptive sampling arc length. The sampling arc length is obtained according to the following formula, and then six successor nodes are generated.

$$l_{\text{arc_sample}} = \left(0.06 * (6 - n_{\text{safe}})^2 - 0.6 * (6 - n_{\text{safe}}) + 2\right) * l_{\text{arc_fixed}} \quad (4)$$

Among them, n_{safe} represents the number of collision-free child nodes of the parent node, $l_{\text{arc_fixed}}$ represents the fixed sampling arc length, and the simulation scene takes 0.125 m, which $l_{\text{arc_sample}}$ represents the adaptive sampling arc length. It can be inferred that the more collision-free child nodes there are, the more spacious the environment is, and the larger the adaptive sampling arc length is.

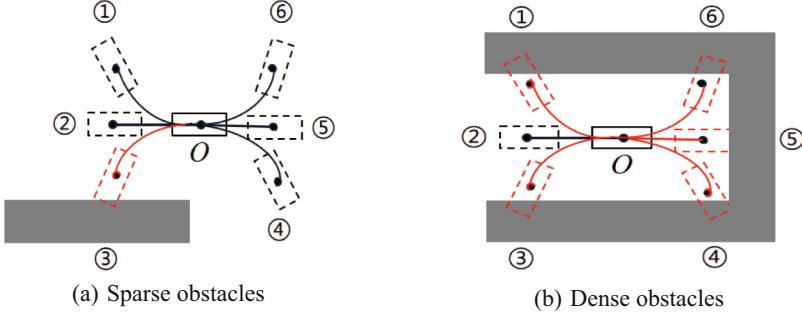


Fig. 3. Schematic diagram of successor node collision. Nodes ①–⑥ represent the six successor nodes with node o as the parent node, and the successor nodes that collide with the gray obstacle area are represented in red. a) Sparse obstacle scenario: There are five non-collision successor nodes. b) Dense obstacle scenario: There is only one non-collision successor node.

4 Simulation Analysis

Based on the ROS operating system, a parallel parking simulation scene is built: The vehicle length of the path planning experiment is 4.6 m, the width is 1.8 m, the wheelbase is 2.7 m, the front overhang is 0.92 m, the rear overhang is 0.98 m, the wheelbase is 1.6 m, and the minimum turning radius is 4.56 m. The parking garage is 6 m long and 2.4 m wide. The same starting and ending postures are set in the Rviz platform. The traditional Hybrid A* algorithm and the improved Hybrid A* algorithm are used to obtain the parking path visualization results shown in the figure below. The black line represents the garage outline, the blue box with a green arrow represents the starting posture of the smart car, the blue box with a red arrow represents the end posture, the red dot is the parking path point, and the orange box is the vehicle posture corresponding to the path point. From the Rviz visualization results, it can be seen that both the traditional Hybrid A* algorithm and the improved Hybrid A* algorithm can successfully plan collision-free parking paths, and the path shapes are similar. However, the traditional Hybrid A* algorithm uses a fixed sampling arc length, and the path points are evenly spaced. Due to the use of adaptive sampling arc length, the improved Hybrid A* algorithm has a large spacing between path points in the open area, while the path points near the garage edge are densely distributed. That is, the improved algorithm takes into account the obstacle avoidance requirements of planning while improving the node expansion efficiency. Table 1 shows that the parking path planned by the improved Hybrid A* algorithm is only 0.84% higher than that of the traditional Hybrid A* algorithm, but the

number of expanded nodes has been reduced from 7231 to 1725, and the algorithm time consumption has been reduced by about 34.3%. The parking planning efficiency and algorithm real-time performance have been greatly improved (Fig. 4).

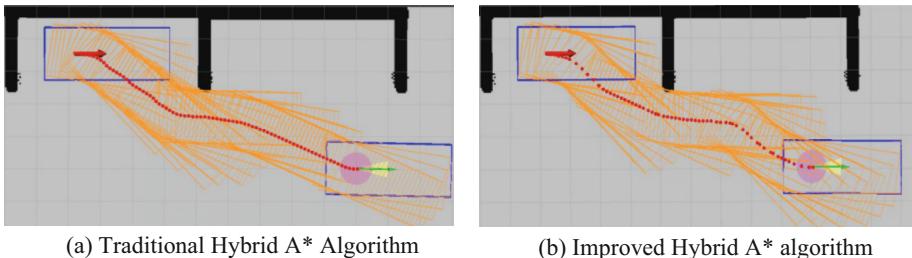


Fig. 4. Simulation comparison before and after Hybrid A* Algorithm Improvement

Table 1. Parameter comparison before and after Hybrid A* Algorithm Improvement

	Path length (m)	Running time (ms)	Number of nodes expanded
Traditional Hybrid A* algorithm	9.5	10.5	7231
Improved Hybrid A* algorithm	9.58	6.9	1725
Contrast	Up 0.84%	Down 34.3%	Down 76.1%

5 Conclusion

Aiming at the problems of low node expansion efficiency and time-consuming path calculation in the parking scene of the hybrid A* algorithm, an improved hybrid A* path planning method based on adaptive sampling arc length is proposed. The sampling arc length of the successor node is dynamically adjusted by the number of collision-free child nodes of the parent node, which can not only improve the node expansion efficiency in open scenes but also take into account the obstacle avoidance requirements. Through the parallel parking simulation experiment under the ROS system, the significant advantages of the improved hybrid A* algorithm in expansion efficiency and search time are verified.

References

1. Banzhaf, H., Nienhüser, D., Knoop, S., et al.: The future of parking: a survey on automated valet parking with an outlook on high density parking. In: 2017 IEEE Intelligent Vehicles Symposium (IV), pp. 1827–1834. IEEE (2017)

2. Jang, C., Kim, C., Lee, S., et al.: Re-plannable automated parking system with a standalone around view monitor for narrow parking lots. *IEEE Trans. Intell. Transp. Syst.* **21**(2), 777–790 (2019)
3. Paden, B., Čáp, M., Yong, S.Z., et al.: A survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Trans. Intell. Veh.* **1**(1), 33–55 (2016)
4. Dijkstra, E.W.: A note on two problems in connexion with graphs. In: Edsger Wybe Dijkstra: His Life, Work, and Legacy, pp. 287–290. ACM, New York (2022)
5. Kang, J.G., Lim, D.W., Choi, Y.S., et al.: Improved RRT-connect algorithm based on triangular inequality for robot path planning. *Sensors.* **21**(2), 333 (2021)
6. Klemm, S., Oberländer, J., Hermann, A., et al.: RRT*-Connect: Faster, asymptotically optimal motion planning. In: 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 1670–1677. IEEE (2015)
7. Zips, P., Böck, M., Kugi, A.: Optimisation based path planning for car parking in narrow environments. *Robot. Auton. Syst.* **79**, 1–11 (2016)
8. Zhang, X., Liniger, A., Sakai, A., et al.: Autonomous parking using optimization-based collision avoidance. In: 2018 IEEE Conference on Decision and Control (CDC), pp. 4327–4332. IEEE (2018)
9. Hart, P.E., Nilsson, N.J., Raphael, B.: A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans. Syst. Sci. Cybern.* **4**(2), 100–107 (1968)
10. Liu, X., Deng, R., Wang, J., et al.: COStar: a D-star lite-based dynamic search algorithm for codon optimization. *J. Theor. Biol.* **344**, 19–30 (2014)
11. Dolgov, D., Thrun, S., Montemerlo, M., et al.: Path planning for autonomous vehicles in unknown semi-structured environments. *Int. J. Robot. Res.* **29**(5), 485–501 (2010)
12. Chen, J., Zhang, X., Xu, L.: Path planning algorithm based on hybrid A* and adaptive ant colony optimization. In: 2022 18th International Conference on Computational Intelligence and Security (CIS), pp. 43–48. IEEE (2022)
13. Tu, K., Yang, S., Zhang, H., et al.: Hybrid A* based motion planning for autonomous vehicles in unstructured environment. In: 2019 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–4. IEEE (2019)
14. Li, B., Acarman, T., Zhang, Y., et al.: Tractor-trailer vehicle trajectory planning in narrow environments with a progressively constrained optimal control approach. *IEEE Trans. Intell. Veh.* **5**(3), 414–425 (2019)
15. Wang, M., Shi, L., Li, Y.: Multi-path-planning in complex intersection based on improved hybrid A* and minimum jerk optimization. In: 2023 China Automation Congress (CAC), pp. 3093–3098. IEEE (2023)
16. Zhao, L., Mao, R., Bai, Y.: Local path planning for unmanned surface vehicles based on hybrid A* and B-spline. In: 2022 IEEE International Conference on Unmanned Systems (ICUS), pp. 81–85. IEEE (2022)
17. Qin, Z., Chen, X., Hu, M., et al.: A novel path planning methodology for automated valet parking based on directional graph search and geometry curve. *Robot. Auton. Syst.* **132**, 103606 (2020)
18. Sheng, W., Li, B., Zhong, X.: Autonomous parking trajectory planning with tiny passages: a combination of multistage hybrid A-star algorithm and numerical optimal control. *IEEE Access.* **9**, 102801–102810 (2021)
19. Li, C., Yu, D., Lu, W.: Variable-curvature Hybrid A* search for AMR path planning in limited space. In: Proceedings of the 2021 3rd International Symposium on Robotics & Intelligent Manufacturing Technology (ISRIMT). IEEE (2021)
20. Ren, X., Cai, Z.: Kinematics model of unmanned driving vehicle. In: 2010 8th World Congress on Intelligent Control and Automation, pp. 5910–5914. IEEE (2010)



UAV Imagery Based Railroad Tunnel Facility Instance Segmentation Using Post-processing of Spatial Topological Relationships

Tong Meng^{1,2} , Yong Qin^{1,3} , Fanteng Meng^{1,2} , Ninghai Qiu^{1,2} , Chongchong Yu⁴ , and Zhipeng Wang¹

¹ The State Key Laboratory of Advanced Rail Autonomous Operation, Beijing Jiaotong University, Beijing 100044, China
`{24120793,yqin,ftmeng,qiuninghai,zpwang}@bjtu.edu.cn`

² School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

³ Key Laboratory of Railway Industry of Proactive Safety and Risk Control, Beijing Jiaotong University, Beijing 100044, China

⁴ School of Computer and Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China

Abstract. Railroad tunnels, as important infrastructures where the line crosses complex mountains, require frequent and focused monitoring, especially at the tunnel entrance facilities connecting the mountains and slopes. Unmanned aerial vehicle (UAV)-based railroad tunnel entrance inspection has great potential to be an effective solution due to its highly maneuverable and wide aerial views. Unfortunately, current convolution neural network (CNN)-based methods struggle to accurately segment the railroad tunnel entrance facilities. This study presents a coarse-to-fine post-processing method that integrates the image processing algorithm and spatial topological representations between different facilities. First, outlier removal based on clustering is designed for coarse processing the segmentation results. Second, segmentation consistency verification and boundary refinement guided by structural adjacency relations progressively execute for fine optimizing and achieving the results. Finally, experimental results conducted on complex UAV railroad tunnel entrance dataset demonstrate that the proposed method exhibits strong generalization capability across different types of tunnels entrance and achieves stable and accurate structural segmentation.

Keywords: Railroad Tunnel · UAV Imagery · Segmentation Post-Processing · Spatial Topological Relationships

1 Introduction

By the end of 2024, the total operating length of China's railroad network has exceeded 160,000 km, including 46,000 km of high-speed rail, ranking first in the world [1]. The number of railroad tunnels has surpassed 21,000, with a total tunnel length of over 20,000 km [2]. Tunnel entrances, which connect tunnel structures with natural slopes,

are critical components affecting train operation safety. For example, mudslides [3] or rockfalls [4] encroach on tunnel portals and are not promptly detected, they lead to train derailments, resulting in casualties and significant economic losses. Currently, tunnel entrance monitoring primarily relies on fixed surveillance cameras along the tracks, which suffer from blind spots in field of view. This makes it difficult to monitor upper structures such as tunnel tops, vents, and side slopes effectively (shown in Fig. 1). Fortunately, UAVs (unmanned aerial vehicles), which are highly maneuverable and offer wide aerial views, have been widely adopted for railroad infrastructure inspection. Using UAVs equipped with high-definition optical cameras to capture tunnel entrance areas from a broader perspective presents a more efficient alternative to traditional monitoring methods.



Fig. 1. Category label names and their corresponding areas

However, UAV-based imagery faces challenges including large data volumes, significant viewpoint variations, complex lighting conditions, frequent occlusions, and noisy backgrounds. These factors call for the development of automated algorithms capable of accurately segmenting key tunnel entrance facilities from aerial images to reduce manual effort. While end-to-end convolutional neural networks have gained widespread attention for their powerful feature extraction capabilities, models like the Fully Convolutional Network (FCN) [5], Mask R-CNN [6], and the Segment Anything Model (SAM) [7] have shown good performance in general segmentation tasks, but still face key challenges in identifying facilities within railroad tunnel entrance scenes: complex background interference, heterogeneous instances share similar features and blurry boundaries between adjacent structures.

To address these challenges, this paper leverages the advanced YOLOv11-segment [8] model as the base framework for initial segmentation of tunnel entrance components. The main contributions of this paper are summarized as follows:

- A deep model post-processing framework is proposed and it can be seamlessly integrated with any segmentation-related deep learning model.
- A clustering approach is introduced to remove outlier noise in tunnel entrance facility segmentation.

- spatial continuity constraint and spatial adjacency constraint models are constructed by leveraging the inherent topological relationships among tunnel components shaped during engineering construction.

2 Related Work

2.1 UAV-Based Railroad Infrastructure Detection

In recent years, unmanned aerial vehicles (UAVs) have been widely applied to railroad infrastructure inspection. For track structure detection, methods such as NL-LinkNet-SSR [9], DenseTrackNet [10], and UNet-ResNet hybrids [11] have shown effectiveness in extracting centerlines and track boundaries by integrating attention modules, edge-aware blocks, and residual feature enhancement. In bridge and tunnel inspection tasks, semantic segmentation on 2D images [12] and 3D point clouds [13] has been employed to identify internal tunnel structures, bridge piers, and portals. Meanwhile, for environmental risk identification, models like YOLARC [14], YOLORS [15], and UYOLO [16] have been developed for segmenting foreign objects and small-scale hazards in UAV imagery, leveraging multi-scale perception and risk-level inference. However, these approaches consider optimizing feature extraction methods to improve the efficiency of railroad risk hazard identification, but never consider using the established spatial topological relationship constraints of railroad scenarios to ensure the rationality of identification.

2.2 End-to-End Deep Convolution Segmentation Model

Classical semantic segmentation models such as Fully Convolutional Networks (FCN) [6] and U-Net [17], as well as instance-level frameworks like Mask R-CNN [6], PointRend [18], and SOLOv2 [19], have significantly advanced pixel-wise understanding in complex scenes. These models are capable of identifying fine-grained regions and have been widely applied in domains such as medical imaging, natural scene parsing, and general object recognition. Moreover, the advent of large-scale pretrained zero-shot segmentation models, such as Segment Anything Model (SAM) [7] and SEEM [20], further improved generalization performance across unseen categories and domains. However, they suffer from handling of small-scale ambiguous structures, boundary inconsistency, and lacking domain-specific priors required to segment infrastructure with structural logic. These limitations motivate the development of task-specific post-processing frameworks that integrate spatial semantics and engineering priors to support high-precision railroad tunnel entrance facility segmentation.

3 Methodology

3.1 Algorithm Framework

Although the YOLOv11 segment model demonstrates strong performance in structural object detection and segmentation tasks, its output is still subject to challenges such as false positives, small-object noise, and blurred boundaries. To further enhance the structural completeness and geometric consistency of instance segmentation results,

this study proposes a post-processing pipeline guided by structural semantics and spatial constraints, progressively optimizing the raw model predictions. It consists of the following four key steps:

1. Binary mask extraction: Based on the YOLOv11-segment predicted category labels, binary masks are separately extracted for each structure type to facilitate independent processing.
2. Outlier removal: Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering is applied to identify and eliminate outlier instances that may result from texture noise or misdetections of small objects.
3. Vent false positive removal: A direction vector is constructed between the geometric centers of the track area and top. The angular deviation between this vector and each vent center is calculated. If it exceeds a defined threshold, the vent instance is removed.
4. Boundary refinement: By utilizing the adjacency between structures like the track area, slope, and external structure, the method smooths and merges boundaries to enhance spatial continuity.

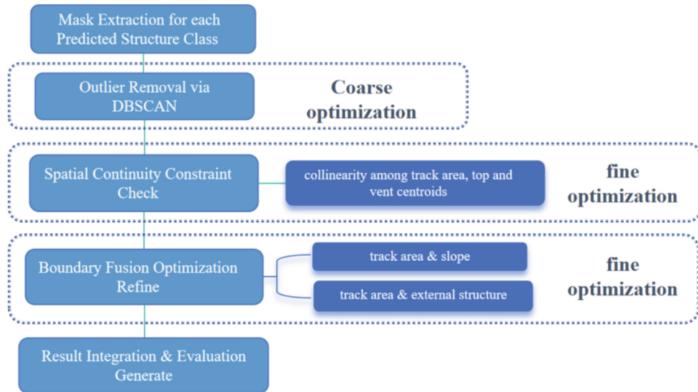


Fig. 2. Data post-processing flowchart

This pipeline fully incorporates structural geometric relationships, semantic category priors, and traditional image processing operations, enabling the transformation from coarse initial segmentation to be refined structural outputs. The overall framework is illustrated in Fig. 2.

3.2 YOLOv11-Segment Model Output Description and Category-Wise Mask Extraction

In this study, we adopt YOLOv11-segment as the base instance segmentation model to automatically identify structural components at railroad tunnel entrances. As shown in Fig. 3, YOLOv11-segment performs structure-level instance segmentation on the input image, outputting the category label, confidence score, bounding box, and pixel-level

mask for each detected instance. We extract binary masks for each predicted structure class. Specifically, instances belonging to the same class are grouped and converted into a unified binary mask, in which only the target structure is preserved while all other classes are excluded. All category-wise masks are then recombined into a full-scene instance map with clearly separated structure types.

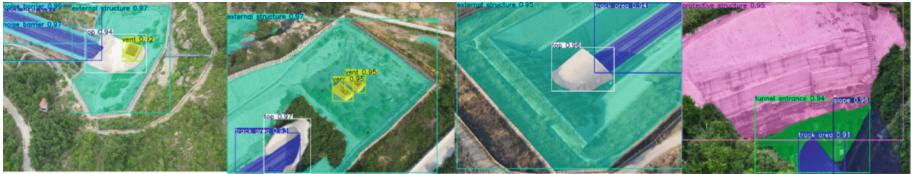


Fig. 3. Segmentation performance of YOLOv11-segment

3.3 Outlier Removal via DBSCAN Clustering

During the instance segmentation process, the model may produce small-sized targets, which are often noise or false detections that can interfere with subsequent structural analysis. To remove such outliers, we adopt a DBSCAN-based clustering approach to automatically identify and eliminate spatially isolated structural instances.

The core idea of this method is to leverage the density-based spatial distribution of object instances to distinguish between main structural regions and spatially isolated outlier points. The pseudocode of the process is shown in Algorithm 1. This method significantly improves segmentation quality by removing noisy outputs, especially around the periphery of tunnel entrances, as illustrated in Fig. 4.



Fig. 4. Illustration of mask results before and after outlier removal

Algorithm 1

Algorithm1 DBSCAN_Filter. Filtering spatial outliers based on DBSCAN

```

1 Input: The binary segmentation mask of a structure, mask
2 Output: The filtered mask after removing outliers, filtered_mask
3 contours ← findContours(mask)
4 for each contour in contours do
5     M ← moments(contour)
6     if M['m00'] ≠ 0 then
7         cx ← M['m10'] / M['m00']
8         cy ← M['m01'] / M['m00']
9         centroids.append([cx, cy])
10    end if
11 end for
12 labels ← DBSCAN(centroids, eps = 25, min_samples = 3)
13 for each i in range(0, len(labels)) do
14    if labels[i] ≠ -1 then
15        drawContour(filtered_mask, contours[i], color = 255, fill = True)
16    end if
17 end for
18 return filtered_mask

```

3.4 Segmentation Consistency Validation Based on Spatial Continuity Constraints

In the instance segmentation task, the vent is prone to misdetection due to its small size, simple geometry, and frequent occurrence in visually complex backgrounds such as vegetation. As shown in Fig. 5, background objects with similar colors or shapes may be mistakenly recognized as vents, even when the confidence score is relatively high.

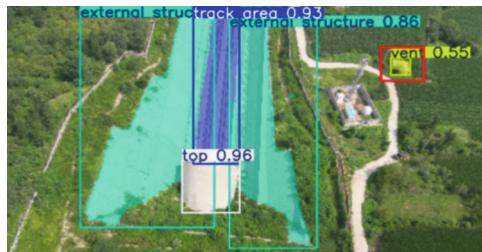


Fig. 5. Vent misclassification (the region within the red box shows a falsely detected structure)

To address this issue, we propose a vent misdetection removal strategy based on geometric spatial constraints. The core assumption is that, in real tunnel scenes, a valid vent structure is generally aligned along the extended axis between the centers of the track area and top. Hence, the geometric centers of the three components should lie approximately along a straight line. If we treat the line connecting the centers of the track area and top as the primary structural axis (indicated by the black arrow in the

figure), then the vent center is expected to be positioned near this axis. This strategy uses an angular threshold (set to 30° in this study) to eliminate spatially inconsistent targets. This is illustrated in Fig. 6(a) and (b).

The specific procedure is as follows (The pseudocode of the process is shown in Algorithm 2): First, the geometric centers of the track area, top, and vent are calculated. A direction vector is constructed from the track area center to the top center, which defines the primary structural axis. Then, the angle between this axis and the line connecting to each vent center (shown as orange arrows in figures) is computed. If the angle is smaller than the predefined threshold (30°), the vent is considered to follow a reasonable structural alignment and is retained. Otherwise, it is marked as a spatial anomaly and removed as a false detection.

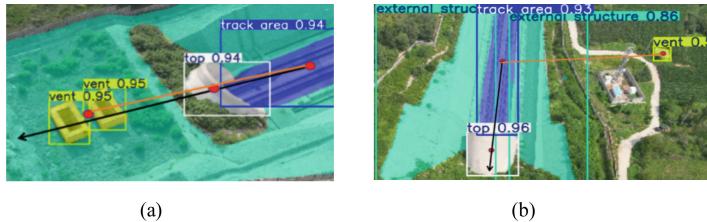


Fig. 6. Illustration based on the spatial continuity constraints. (a) Example of correct classification. (b) Example of fault classification

Algorithm 2.

Algorithm2 Vent_Filter. False positive removal of vents

```

1  Input: Binary masks of track area, top, and vent structures
2  Output: Filtered mask of vent structures after removing spatial outliers
3  track_center ← computeCentroid(track_mask)
4  top_center ← computeCentroid(top_mask)
5  vent_centers ← computeCentroids(vent_mask)
6  main_axis ← vector(track_center, top_center)
7  for each vent_center in vent_centers do
8      vent_vector ← vector(track_center, vent_center)
9      angle ← computeAngle(main_axis, vent_vector)
10     if angle < 30 degrees then
11         keep vent_center
12     else
13         remove vent_center from mask
14     end if
15  end for
16  return filtered_vent_mask

```

3.5 Boundary Refinement Based on Spatial Adjacency Constraints

During the instance segmentation process, the contact boundaries between adjacent structures often exhibit jagged edges or discontinuities, which negatively affect subsequent structural analysis and geometric modeling. To address this issue, we propose a boundary refinement method based on spatial adjacency constraints.

The core assumption behind this method is that, in realistic tunnel entrance scenes, the adjacent boundaries between the track area and external structure, as well as between the track area and slope, should ideally form smooth straight lines.

The detailed process is as follows (The pseudocode of the process is shown in Algorithm 3):

1. Contact region detection: Overlapping areas between the two masks are identified. These regions generally indicate physical contact zones between neighboring structures.
2. Line fitting: If overlapping regions are detected, the shared boundary is approximated by a straight line.

Through this optimization process, the boundaries between adjacent structures are effectively refined, eliminating zigzag artifacts and fractures.

Algorithm 3.

Algorithm3 Boundary_Fusion. Optimizing adjacent structure boundaries

```

1 Input: Binary masks of two adjacent structures, mask_A and mask_B
2 Output: Updated masks, new_mask_A and new_mask_B
3 contours_A ← findContours(mask_A)
4 contours_B ← findContours(mask_B)
5 intersection ← bitwise_and(mask_A, mask_B)
6 if sum(intersection) > 0 then
7     for each contour_A in contours_A do
8         for each contour_B in contours_B do
9             epsilon ← 0.02 * arcLength(contour_A, True)
10            approx ← approxPolyDP(contour_A, epsilon, True)
11            if length(approx) = 2 then
12                pt1 ← approx[0]
13                pt2 ← approx[1]
14                drawLine(mask_A, pt1, pt2, color = 255, thickness = 2)
15                drawLine(mask_B, pt1, pt2, color = 255, thickness = 2)
16            end if
17        end for
18    end for
19 end if
20 return mask_A, mask_B

```

4 Experiments

4.1 Dataset and Experimental Settings

In this study, we constructed a tunnel entrance dataset that covers a variety of complex scenarios, including mountainous regions, hilly terrain, and bridge–tunnel transitions. The dataset comprises a total of 340 images of tunnel entrances from diverse terrain conditions, such as high-speed railroad portals and conventional railroad tunnels in mountainous areas. The images were collected across different geographical conditions to ensure diversity and representativeness. In addition, all images were manually annotated using the LabelMe tool. The annotation categories include nine common tunnel entrance structures (as shown in Fig. 1): track area, noise barrier, top, external structure, slope, protective structure, passive network, vent, and tunnel entrance. These categories cover the most typical structural components in tunnel entrance environments. Examples of annotated images and structure types are shown in Fig. 7. Each annotated image was verified through multiple rounds of manual review to ensure accurate and unambiguous structural labeling. The category distribution across the dataset is relatively balanced, which contributes to the diversity and generalization capability of the model during training.

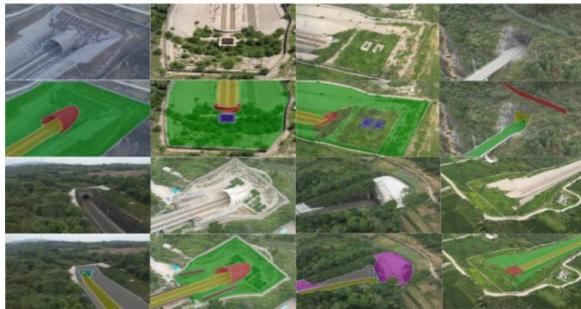


Fig. 7. Typical categories of tunnel entrances and their labeled images

The experimental process follows the modular post-processing framework described in Subsect. 3.1. In terms of evaluation metrics, this study adopts the following commonly used indicators to comprehensively assess the model’s performance:

- Intersection over Union (IoU): Measures the overlap between the predicted mask and ground truth mask for each category, quantifying segmentation accuracy at the instance level.
- Mean Intersection over Union (mIoU): Computes the average IoU across all structure categories, providing a holistic evaluation of the model’s multi-class segmentation performance.

4.2 Outlier Noise Removal

By applying this method, small and meaningless targets can be effectively removed, ensuring that the final mask output contains only valid structural information. A comparison of the segmentation results before and after outlier removal is presented in Table 1. As shown in the table, the mIoU improved for most structure categories after applying the outlier removal process, except for the noise barrier classe. Notably, the protective structure and external structure categories showed a significant reduction in false detections, leading to better segmentation accuracy overall.

Table 1. Comparison of mIoU before and after outlier removal

classes	mIoU (Before)	mIoU (After)	Improvement
track area	93.05%	93.60%	+0.55%
noise barrier	80.12%	79.98%	-0.14%
top	89.32%	89.76%	+0.44%
external structure	89.31%	89.59%	+0.28%
slope	81.44%	81.48%	+0.04%
protective structure	54.81%	57.20%	+2.39%
passive network	81.89%	81.89%	+0.00%
vent	51.69%	51.72%	+0.03%
tunnel entrance	82.43%	82.88%	+0.45%

4.3 Spatial Consistency Validation

In the ablation study, we compare the model’s performance on vent structures with and without applying the spatial validation strategy. The experimental results show that this method significantly improves segmentation accuracy across multiple structural categories. Particularly in scenarios with cluttered tunnel entrance backgrounds, the proposed strategy effectively prevents incorrect merging of vent instances and enhances the model’s overall performance. A detailed comparison is provided in Table 2.

Table 2. Comparison of mIoU before and after vent misclassification elimination

Experimental Setting	Mean IoU (mIoU)
Without vent misclassification elimination	51.72%
With vent misclassification elimination	59.95%

After applying the vent false detection elimination strategy, the mIoU score for vent structures improved by 8.23%, demonstrating the strategy’s effectiveness in reducing false positives and improving the model’s ability to distinguish and segment fine-grained structures.

4.4 Adjacency Aware Boundary Optimization

To evaluate the effectiveness of this strategy, comparative experiments were conducted on three typical structural categories. The results before and after boundary optimization are shown in Table 3.

Table 3. Comparison of mIoU before and after adjacency-aware boundary optimization

Class	mIoU (Before)	mIoU (After)	Improvement
track area	93.60%	93.65%	+0.05%
external structure	89.59%	90.50%	+0.91%
slope	81.48%	85.54%	+4.06%

The results indicate that this method improves segmentation accuracy across all tested categories. In particular, the slope class showed the most significant gain, with an mIoU increase of over 4%, demonstrating the method's effectiveness in handling irregular, low-contrast boundary regions. For the external structure, clearer boundary definition contributed to improved structural completeness. Although the track area already achieved high segmentation accuracy prior to optimization, smoothing the boundary still yielded minor gains in stability.

5 Conclusion

This paper proposes a post-processing method for segmentation of railroad tunnel entrances. By incorporating structural semantics and spatial topological constraints, the proposed approach significantly improves the segmentation performance. This method includes a series of coarse-to-fine optimization strategies: outlier removal via clustering, spatial consistency validation based on positional constraints, and adjacent-aware boundary refinement. Experimental results demonstrate that the proposed method enhances structural-level segmentation quality without compromising the baseline model's efficiency. Moreover, the method exhibits strong adaptability across various tunnel entrance scenes, including high speed railroad, mountainous, and bridge-tunnel transition zones, showing promising potential for engineering deployment.

Acknowledgements. This work was supported in part by Autonomous Research Project of State Key Laboratory of Advanced Rail Autonomous Operation RAO2023ZZ003 and in part by National Key R&D Program of China under Grant 2022YFB4300600.

References

1. Zhang, Z.: China's railway leads the world as railway mileage surpasses 160,000 kilometers. Baijiahao (Baidu News) (2024). [Online]. Available: <https://baijiahao.baidu.com/s?id=1810685835588287153>. Accessed 20 Sep 2024

2. China State Railway Group Co., Ltd.: China Railway Statistical Yearbook 2023. China Railway Press, Beijing (2024)
3. Liu, S.: Debris flow causing derailment of train D2809. China News Network (2022). [Online]. Available: <https://www.chinanews.com.cn/sh/2022/06-04/9774304.shtml>. Accessed 04 Jun 2022
4. The State Council of the People's Republic of China: Report on Gaoyangzhai tunnel collapse in Yichang-Wanzhou Railway. gov.cn (2007). [Online]. Available: https://www.gov.cn/jrzq/2007-11/25/content_814936.htm. Accessed 25 Nov 2007
5. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440 (2015)
6. 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)
7. Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the European Conference on Computer Vision, pp. 801–818 (2018)
8. Jocher, G., Chaurasia, A., Chang, L.: YOLO by ultralytics (2024). <https://github.com/ultralytics/ultralytics>. Accessed 11 Apr 2024
9. Wang, J., Liu, Y., Guo, H.: Railway tracks extraction from high resolution unmanned aerial vehicle images using improved NL-LinkNet network. Drones **8**(11), 611 (2024)
10. Chen, L., Li, B., Sun, K.: DenseTrackNet: dense attention network for railway track detection. In: Proceedings of the IEEE International Conference on Intelligent Transportation Systems, pp. 902–908 (2023)
11. Diakogiannis, F. I., Waldner, F., Caccetta, P., Wu, C.: ResUNet-a: a deep learning framework for semantic segmentation of remotely sensed data. ISPRS J. Photogrammetry Remote Sens. **162**, 94–114 (2020)
12. Kim, H., Narazaki, Y., Spencer Jr., B.F.: Automated bridge component recognition using close-range images from unmanned aerial vehicles. Eng. Struct. **274**, 115184 (2023)
13. Grandio, J., Riveiro, B., Soilán, M., Arias, P.: Point cloud semantic segmentation of complex railway environments using deep learning. Autom. Constr. **141**, 104425 (2022)
14. Wu, Y., Meng, F., Qin, Y., et al.: UAV imagery based potential safety hazard evaluation for high-speed railroad using real-time instance segmentation. Adv. Eng. Inform. **55**, 101819 (2023)
15. Wu, Y., Zhao, Z., Chen, P., et al.: Hybrid learning architecture for high-speed railroad scene parsing and potential safety hazard evaluation of UAV images. Measurement **239**, 115504 (2025)
16. Manole, A., Dioşan, L.-S.: UOLO: A multitask U-Net YOLO hybrid model for railway scene understanding. IEEE Trans. Intell. Veh. **10**(3), 2015–2026 (2025)
17. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Proceedings of the Medical Image Computing and Computer-Assisted Intervention, pp. 234–241 (2015)
18. Kirillov, A., Wu, Y., He, K., Girshick, R.: PointRend: image segmentation as rendering. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9799–9808 (2020)
19. Wang, X., Zhang, R., Kong, T., et al.: SOLOv2: dynamic and fast instance segmentation. In: Proceedings of the Neural Information Processing Systems (2020)
20. Zou, X., Xie, Z., Wu, Y., et al.: SEEM: segment everything everywhere all at once. arXiv Preprint arXiv:2304.06718 (2023)



Self-Organized Radar Signal Sorting via Multi-order DTOA Spectral Analysis and Parameter Entropy Dynamic Fusion

Xun Huang¹, Dan Huang², Lu Ding¹(✉), Chenggang Wang³, and Lei Song³

¹ School of Electrical Engineering, Guangxi University, Nanning 530004, China
2312391026@st.gxu.edu.cn, dinglu@gxu.edu.cn

² School of Aeronautics and Astronautics, Shanghai Jiao Tong University, Shanghai 200240, China
huangdan@sjtu.edu.cn

³ School of Automation and Sensing, Shanghai Jiao Tong University, Shanghai 200240, China
{cgwang-auv, songlei_24}@sjtu.edu.cn

Abstract. Modern radar signal sorting faces critical challenges in complex electromagnetic environments characterized by severe parameter overlapping and dynamic modulation patterns. Traditional methods relying on single-domain parameter analysis (e.g., RF, PW, DOA) exhibit limited robustness under such conditions, while existing PRI-based approaches struggle with multi-modulation adaptability and computational efficiency. This paper proposes a self-organized radar signal sorting framework integrating multi-order DTOA spectral analysis and parameter entropy dynamic fusion. First, frequency-domain features are extracted from TOA sequences through multi-order DTOA spectral analysis, where threshold-controlled DTOA frequency matrices are transformed into latent PRI features via softmax-based correlation mapping. Second, an adaptive weighting mechanism dynamically fuses multi-domain features by evaluating feature entropy and standard deviation, generating a weighted similarity matrix that captures inter-pulse correlations. Third, manifold learning optimizes the high-dimensional similarity matrix into a low-dimensional discriminative space for clustering. Experiments on five simulated radar pulse descriptor datasets demonstrate the framework's superiority over state-of-the-art methods, achieving average Rand Index (RI) of 0.960 and Adjusted Rand Index (ARI) of 0.902. The proposed method resolves parameter overlap, uneven pulse density, and PRI modulation diversity in modern radar signal sorting for congested environments.

Keywords: Radar Signal Sorting · Features Dynamic Fusion · Clustering

1 Introduction

Modern radar systems face significant challenges in signal sorting within congested electromagnetic environments, where traditional parameter-based methods fail due to severe parameter overlap and complex modulation patterns. As a key metric reflecting

The work is supported by the Innovation Project of Guangxi Graduate Education (ZX01030030225017).

temporal patterns, Pulse Repetition Interval (PRI) enables critical pulse discrimination, yet existing PRI extraction methods struggle with adaptive feature fusion and dynamic weighting in complex scenarios [1].

Current research predominantly focuses on histogram-based PRI extraction. While SDIF variants [2, 3] and pseudo peak recognition [4] improve periodic PRI identification, they exhibit limited adaptability to diverse modulation types. High computational overheads or substantial prior knowledge requirements hinder alternative methods based on time-frequency transforms [5, 6] and deep learning [7, 8]. Though clustering algorithms eliminate prior information dependence [9, 10], they inadequately leverage PRI-temporal correlations.

To address these limitations, we propose a self-organized framework integrating multi-order DTOA spectral analysis with entropy-driven feature fusion. Our key contributions are:

- (1) Frequency-domain PRI embedding: Combining multi-order DTOA matrices with histogram analysis to generate discriminative representations for various PRI modulations.
- (2) Adaptive similarity weighting: Dynamically prioritizing features through entropy-standard deviation metrics, validated by nonlinear manifold visualization.
- (3) Experimentally validated robustness: Demonstrating state-of-the-art clustering accuracy on simulated datasets under complex electromagnetic conditions.

Experimental results confirm our framework's superior accuracy and interpretability compared to unsupervised baselines, providing an effective solution for real-time radar signal sorting.

2 Feature Extraction of Pulse Repetition Interval

2.1 Frequency Domain Feature Embedding

Radar PRI features, representing temporal modulation patterns, are extracted from interleaved TOA sequences. The DTOA between the i -th and j -th pulses (where $j > i$) is defined as $TOA_j - TOA_i$. As Fig. 1 illustrates, PRI constitutes a subset of DTOA values containing genuine PRI components. Effective frequency-domain embedding of DTOA features critically enables robust PRI extraction [11].

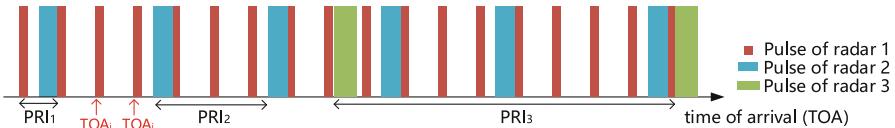


Fig. 1. Schematic diagram of radar pulse TOA sequence.

Radar PRI components and harmonics exhibit distinct spectral peaks in DTOA distributions [4] (Fig. 2 (a)). We extract frequency-domain signatures by calculating the DTOA histogram (frequency matrix). To prevent PRI feature dilution from higher-order

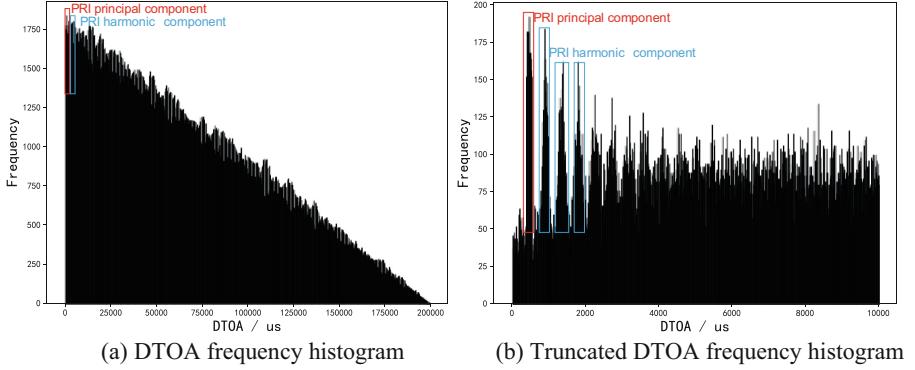


Fig. 2. Multi order DTOA spectrum.

TOA differences, statistical bounds are dynamically constrained via threshold truncation (Fig. 2 (b)).

For the TOA sequence $\mathbf{T} = \{t_1, t_2, t_3, \dots, t_n\}$, the interval frequency count $f_{i,j}$ ($0 \leq i, j \leq n$) is obtained by statistically binning pairwise time differences $t_i - t_j$ in a frequency histogram. This $f_{i,j}$ represents the histogram bin value plotted on the y-axis of the DTOA spectrum, corresponding to the time difference interval $t_i - t_j$. When $i = j$ or the corresponding DTOA exceeds the predefined threshold, $f_{i,j}$ is set to 0. Thus, the resulting DTOA frequency matrix \mathbf{F} is defined as:

$$\mathbf{F}(i, j) = \begin{cases} f_{i,j}, & i \neq j \\ 0, & i = j \end{cases} \quad (1)$$

This matrix \mathbf{F} serves as the primary frequency-domain feature.

2.2 Feature Hidden Space Mapping

Constrained DTOA statistical ranges cause non-zero elements in matrix \mathbf{F} to cluster diagonally, hindering cross-instance similarity assessments. Our solution transforms each row vector of \mathbf{F} via softmax normalization to derive the time correlation matrix:

$$\mathbf{R}(i, j) = \frac{\exp(\mathbf{F}(i, j))}{\sum_{k=1}^n \exp(\mathbf{F}(i, k))} \quad (2)$$

The DTOA occurrence frequencies are converted into probability values through softmax normalization, where higher values indicate greater temporal similarity between instances. To integrate PRI features with other parameters, we derive one-dimensional temporal embeddings via matrix dot-product compression. For instance i , its temporal feature is computed as:

$$z_i = \alpha \sum_{j=1}^n \mathbf{F}^T(i, j) \bullet \mathbf{R}(j, i) \quad (3)$$

Where α serves as a scaling factor to normalize the magnitude of latent temporal features.

3 Weight Allocation of Feature Similarity Matrix

3.1 Feature Similarity Matrix

To enhance extraction of local relationships within radar pulse feature domains, Gaussian kernel functions quantify pairwise pulse similarity. These values constitute correlation features, with each row of the resultant similarity matrix representing a pulse's correlation feature vector.

For the eigenvalues $\mathbf{V} = \{v_1, v_2, v_3, \dots, v_n\}$ of the k -th domain, the numerical differences are first standardized to eliminate the influence of data dimensionality or range, making the differences between different datasets comparable. For any $i, j \in \{1, 2, \dots, n\}$, define the elements of the original distance matrix:

$$\mathbf{M}(i, j) = \frac{1}{2} \left(\frac{v_i - v_j}{T} \right)^2 \quad (4)$$

Where T is the tolerance, defined as $T = v_{\max} - v_{\min}$.

The matrix \mathbf{M} is mapped nonlinearly element by element, and finally the elements are normalized to yield a similar matrix for the k -th field:

$$S_k(i, j) = \frac{\mathbf{E}(i, j) - \min(\mathbf{E}(i, j))}{\max(\mathbf{E}(i, j)) - \min(\mathbf{E}(i, j))} \quad (5)$$

Where $\mathbf{E}(i, j) = e^{-\mathbf{M}(i, j)}$ is a non-linear mapping of $\mathbf{M}(i, j)$.

3.2 Characteristic Entropy

Entropy has some influence on assessing the significance of features [2, 12]. For normalized eigenvalues $\mathbf{X}_i = \{x_1, x_2, x_3, \dots, x_n\}$ in the i -th domain, we derive their kernel density function f to mitigate noise effects. Bin partitioning via the Freedman-Diaconis rule [13, 14] yields intervals $[a_j, b_j]$, for which probability density integrals are computed:

$$p_j = \int_{a_j}^{b_j} \hat{f}_i(x) dx \quad (6)$$

The parameters a_j and b_j are defined as $a_j = \frac{j-1}{n_{bins}} \cdot x_{\max}$ and $b_j = \frac{j}{n_{bins}} \cdot x_{\max}$, respectively, where n_{bins} denotes the number of bins calculated via the Freedman-Diaconis rule. The function $\hat{f}_i(x)$ represents the Gaussian kernel density estimation function.

To ensure that the total probability over all bins sums to unity, the following normalization procedure was applied:

$$p_j^{\text{norm}} = \frac{p_j}{\sum_{k=1}^{n_{bins}} p_k} \quad (7)$$

The entropy corresponding to feature \mathbf{X}_i is defined as:

$$E_i = - \sum_{j=1}^{n_{bins}} p_j^{\text{norm}} \cdot \log(p_j^{\text{norm}} + \epsilon) \quad (8)$$

where $\epsilon = 10^{-10}$ is a numerical stability term added to avoid undefined logarithmic operations when $p_j^{\text{norm}} = 0$.

3.3 Weight Allocation

Radar signal sorting requires multi-perspective analysis [15] due to complex parameter dynamics causing feature discriminability variations. We evaluate feature importance through entropy and bin frequency standard deviation.

Let d denote the number of feature domains. For the i -th feature domain, the initial attention weight is computed as:

$$W_i = \frac{E_i}{Std_i + \epsilon} \quad (9)$$

For the i -th feature domain, E_i denotes its entropy and Std_i represents the standard deviation of bin frequencies (as defined in Sect. 3.2). The weight allocation follows two principles:

1. High Std_i (indicating dynamic parameter variations or environmental interference) reduces weight priority.
2. High E_i (implying severe parameter overlap) similarly necessitates weight suppression.

The penalty terms are formalized as:

$$StdPenalty_j = \begin{cases} 0.4, & \text{if } Std_i > 3.0 \\ 0.7, & \text{if } 2.0 < Std_i \leq 3.0 \\ 1.0, & \text{otherwise} \end{cases} \quad (10)$$

$$EntropyPenalty_i = \begin{cases} \left(\frac{E_i}{2.5}\right)^2, & \text{if } E_i < 2.5 \\ 1.0, & \text{otherwise} \end{cases} \quad (11)$$

After adjusting the initial weights W_i using the two penalty coefficients, normalization is performed to obtain the final attention weight for feature domain i :

$$w_i = \frac{W_i \cdot StdPenalty_j \cdot EntropyPenalty_i}{\sum_{j=1}^d W_j \cdot StdPenalty_j \cdot EntropyPenalty_j} \quad (12)$$

The overall feature similarity matrix is derived by computing the weighted summation of similarity matrices across all feature domains:

$$\mathbf{S} = \sum_{j=1}^d w_j \mathbf{S}_j \quad (13)$$

The row vectors of the similarity matrix serve as correlation features for radar pulses, with diagonal elements normalized to 1. To address high dimensionality, we employ t-SNE [16] for nonlinear dimensionality reduction (2D projection) followed by DBSCAN clustering on the low-dimensional embeddings.

4 Experiments

4.1 Datasets

We generated five datasets through numerical simulation, wherein the distribution patterns of PW, PRI, and CF parameters are determined by the modulation method. When the signal sources exhibit no prior spatial preference, the DOA angle parameter can be regarded as equally probable within the 0° – 360° range, satisfying the core assumption

Table 1. Radar pulse description word simulation data set details.

Dataset	Number of Radars	Parameter Variation Types	Dataset Key Characteristics	DOA/PW/RF Feature Space Visualization
1	2	Fixed, Jittered, Staggered (pulse-group)	Large parameter variation range; Close radar locations	
2	3	Fixed, Jittered, Hopping, Staggered (pulse-group)	Severe parameter overlap; Similar parameter variation patterns	
3	5	Fixed, Jittered, Hopping, Staggered (pulse-group)	Large parameter variation range; Severe parameter overlap; Close radar locations; Uneven data density	
4	6	Fixed, Jittered, Hopping, Staggered (pulse-group)	Large parameter variation range; Large number of emitters	
5	15	Fixed, Jittered, Staggered	Severe parameter overlap; Large number of emitters	

of a uniform distribution. Consequently, DOA was modeled as uniformly distributed in the numerical simulations. Regarding specific parameter ranges, the parameters for Datasets 1–3 originate from real-world measurements, while those for Datasets 4–5 are sourced from references [17, 18]. The details of the datasets are summarized in Table 1.

4.2 Validity Verification Experiment

We jointly visualized the Direction of Arrival (DOA) and the extracted latent PRI features, with the results illustrated in Fig. 3. As demonstrated by the results, the proposed feature extraction method resulted in similar distributions of identical PRI modulation patterns within the latent feature space, while distinct PRI modulation patterns exhibited clear separability.

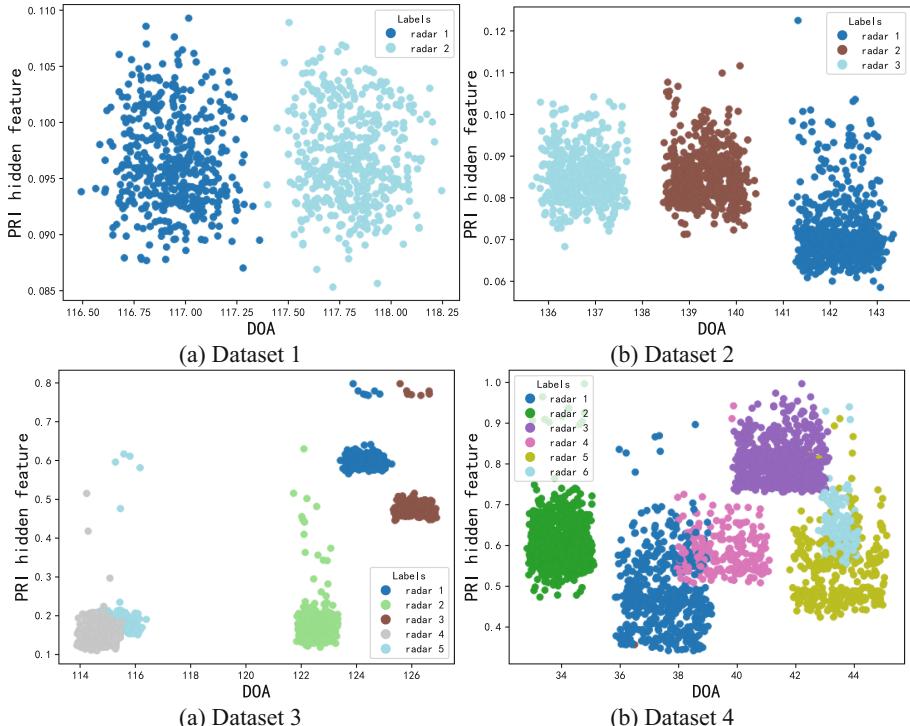
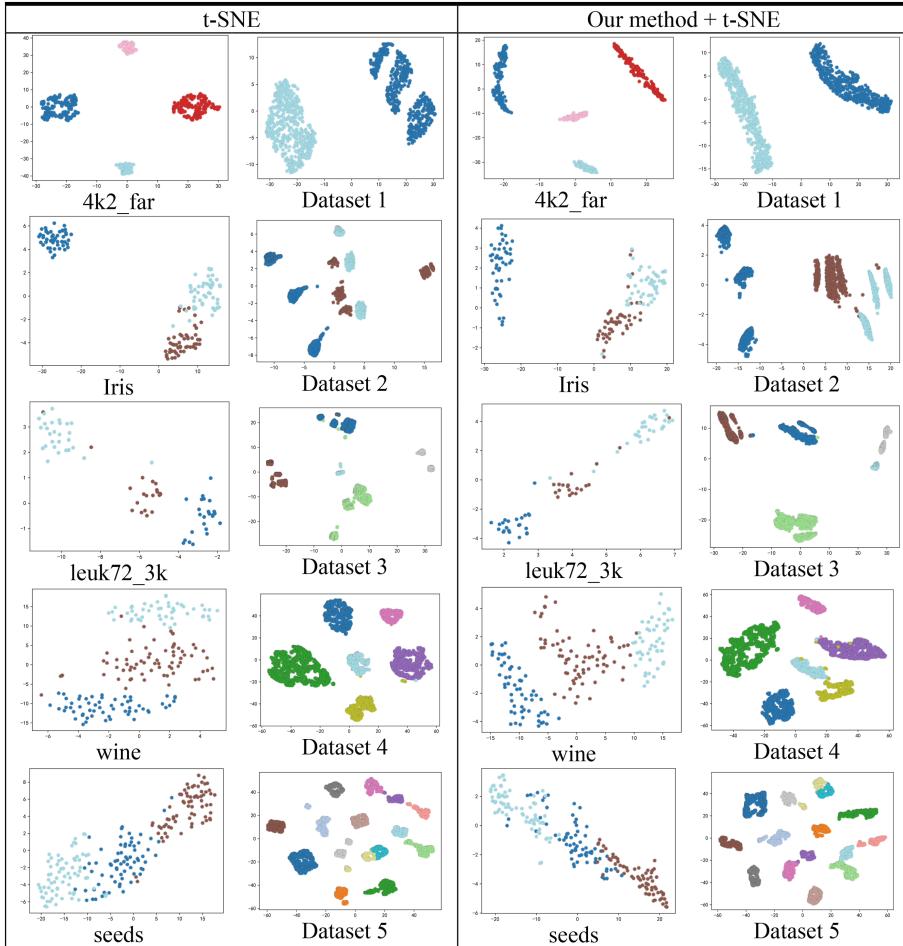


Fig. 3. 2D Visualization of PRI Hidden Features and DOA Features.

To validate the effectiveness of the proposed feature fusion method on heterogeneous data, experiments were conducted on five simulated radar pulse descriptor word datasets and five widely-used standard clustering datasets (4k2_far, Iris, leuk72_3k, wine and seeds). First, the pulse descriptor word vectors, defined as $\mathbf{PDW} = \{PW, DOA, RF, PRI\}$ (where PRI denotes the one-dimensional latent feature obtained via Eq. (3)), were transformed into correlation-aware feature vectors using the proposed method. These

vectors were then projected into a two-dimensional plane via t-SNE. The visualization results were compared against those generated by applying t-SNE directly to the raw data, with the comparative outcomes summarized in Table 2.

Table 2. Ablation experiment results.



As observed from the visualization results, the proposed method exhibited no significant difference in performance on standard clustering datasets. However, when applied to simulated radar pulse descriptor word datasets, our method effectively reduced the intra-class distance.

4.3 Comparative Experiment

The proposed method (with DBSCAN) was benchmarked against four state-of-the-art approaches: DSets-DBSCAN [19], improved-GMM [20], DE-DPC [12], and improved-SC [21]. Quantitative evaluation using Rand Index (RI) and Adjusted Rand Index (ARI) demonstrated superior performance (average RI = 0.960, ARI = 0.902), with full comparisons visualized in Fig. 4 and Table 3.

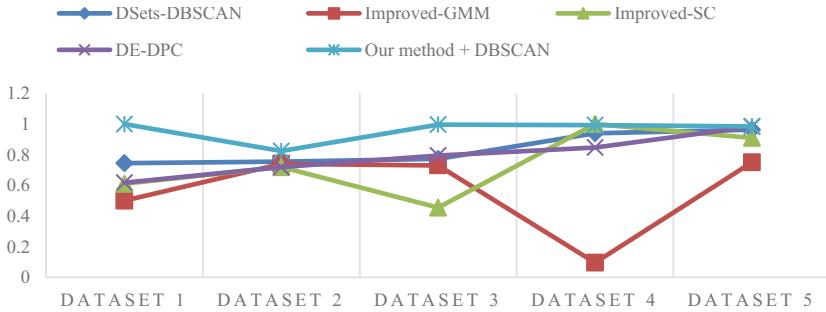


Fig. 4. RI Scores of Different Methods Across Various Datasets

Table 3. Results of algorithm comparison.

Methods	DSets-DBSCAN	Improved-GMM	DE-DPC	Improved-SC	Our method + DBSCAN
Average RI	0.835	0.618	0.793	0.739	0.960
Average ARI	0.493	0.144	0.478	0.510	0.902

RI and ARI metrics approaching 1 indicate higher agreement between clustering results and ground truth. As evidenced by the experimental results, the proposed method demonstrates superior sorting performance over baseline methods on simulated radar pulse descriptor word datasets, while maintaining robust clustering effectiveness even when handling complex and diverse radar pulse signals.

5 Conclusion

This paper proposes a self-organized radar signal sorting framework that integrates multi-order DTOA spectral analysis and entropy-driven dynamic feature fusion. By extracting PRI features through threshold-controlled DTOA frequency matrices and aligning them with PDW parameters via adaptive similarity weighting, the method achieves robust clustering with an average Rand Index (RI) of 0.960 and Adjusted Rand Index

(ARI) of 0.902, significantly outperforming existing approaches. Validated on five simulated radar datasets with parameter overlap, uneven pulse density, and PRI modulation diversity, the framework demonstrates enhanced discriminative power through manifold learning-optimized nonlinear embeddings, offering prior-free adaptability in complex electromagnetic environments. However, our method relies on statistical histograms, so parameters such as the histogram's statistical range and bin size significantly affect algorithm performance.

References

1. Sui, J.P., Liu, Z., Liu, L., Li, X.: Progress in radar emitter signal deinterleaving. *J. Radars.* **11**(03), 418–433 (2022)
2. Zhang, C.J., Qing, S., Deng, Z.A., Liu, Y.C.: Signal deinterleaving algorithm for dwell and switch radar. *J. Syst. Eng. Electron.* **46**(06), 1925–1934 (2024)
3. Zhang, C.J., Liu, Y.C., Si, W.J.: Synthetic algorithm for deinterleaving radar signals in a complex environment. *IET Radar Sonar Navig.* **14**(12), 1918–1928 (2020)
4. Dong, H.X., Ge, Y.Z., Zhou, R., Wang, H.Y.: An improved sorting algorithm for periodic PRI signals based on congruence transform. *Symmetry.* **16**(4), 398 (2024)
5. Fan, F.H., Shi, Y.C., Qin, L.L.: A method of fast detecting sequence for stable pulse repetition interval. *J. Electron. Inf. Technol.* **45**(4), 1293–1302 (2023)
6. Xie, M., Zhao, C., Zhao, Y.J., Hu, D.X., Wang, Z.W.: A novel method for deinterleaving radar signals: first-order difference curve based on sorted TOA difference sequence. *IET Signal Process.* **17**(1), e12162 (2023)
7. Xiang, H.R., Shen, F.R., Zhao, J.: Deep ToA Mask-based recursive radar pulse deinterleaving. *IEEE Trans. Aerosp. Navig. Electron.* **59**(2), 989–1006 (2022)
8. Lin, J., Juliano, J., Erdogan, A., George, K.: Pulse separation using time-frequency mask and machine learning. In: Paul, R. (eds.) CCWC 2021, IEEE, pp. 0067–0072. <https://doi.org/10.1109/CCWC51732.2021.9376161>
9. Wang, X., Chen, X., Zhou, Y.P., Chen, Y., Xiao, B.S., Wang, H.X.: A radar signal sorting algorithm based on improved DBSCAN algorithm. *J. Air Force Eng.* **22**(3), 47–54 (2021)
10. Su, Y.H., Chen, Z., Gong, L.F., Xu, X., Yao, Y.F.: An improved adaptive radar signal sorting algorithm based on DBSCAN by a novel CVI. *IEEE Access.* **12**, 43139–43154 (2024)
11. Xie, M., Zhao, C., Hu, X.D., Liu, C.C.: Signal sorting method for jagger PRI radar based on first-order difference curve of DTOA and correlation analysis. *Acta Electron. Sin.* **51**(11), 3271–3281 (2023)
12. Yu, Z.C., Xia, H.P.: Radar signal sorting based on density entropy and density peak clustering. *Electron. Opt. Control.* **29**(7), 53–56 (2022)
13. Freedman, D., Diaconis, P.: On the histogram as a density estimator: L₂ theory. *Z. Wahrscheinlichkeitstheor. verw. Geb.* **57**(4), 453–476 (1981)
14. Wang, H.Y., Zhu, M.T., Fan, R.Z., Yan, L.: Parametric model-based deinterleaving of radar signals with non-ideal observations via maximum likelihood solution. *IET Radar Sonar Navig.* **16**(8), 1253–1268 (2022)
15. Wang, J.D., Wang, X., Zhou, Y.P., Chen, Y.: Radar emitter classification based on a multi-perspective collaborative clustering method and radar characteristic spectrum. *Int. J. Aerosp. Eng.* **2022**(1), 1297735 (2022)
16. Lang, P., Fu, X.J., Cui, Z.D., Feng, C., Chang, J.Y.: Subspace decomposition based adaptive density peak clustering for radar signals sorting. *IEEE Signal Process. Lett.* **29**, 424–428 (2021)

17. Zhang, Y.X., Guo, W.P., Kang, K., Yao, Y.L., Zhang, L.K., Zhang, W.: Radar signal sorting method based on data field combined PRI transform and clustering. *J. Syst. Eng. Electron.* **41**(7), 1509–1515 (2019)
18. Ma, Z.F., Zhang, Y., Dong, J., Fu, X.J.: Adaptive radar signal sorting based on improved fuzzy ART. *J. Beijing Inst. Technol.* **44**(9), 990–996 (2024)
19. Liu, L.T., Wang, L.L., Li, P., Chen, T.: Radar signal sorting algorithm for DSets-DBSCAN without parameter clustering. *J. Natl. Univ. Def. Technol.* **44**(4), 158–163 (2022)
20. Wang, L., Zhang, Z.Y., Zeng, W.G., Cao, S.L., Zhang, T.H.: An improved GMM clustering based on data field and decision graph. *J. Syst. Eng. Electron.* **44**(9), 2743–2751 (2022)
21. Wang, Y.L., Yang, Y.M.: A radar signal sorting algorithm based on improved spectral clustering. *Telecommun. Eng.* **63**(9), 1348–1354 (2023)



Fault-Tolerant Control for Dual Three-Phase PMSM Based on Multi-Objective Optimization and Multi-Subspace Coordination

WenHong Pang, Yang Gao^(✉), and Qingwei Chen

Nanjing University of Science and Technology, Nanjing 210094, China
gaoyang_zdh@njust.edu.cn

Abstract. Multiphase permanent magnet synchronous motors, renowned for their high power density, low torque ripple, and exceptional fault tolerance, have emerged as a key research focus in aerospace propulsion systems. To address the single-phase open-circuit fault in dual three-phase permanent magnet synchronous motors (DT-PMSM), this paper proposes a multi-objective optimization-based fault-tolerant field-oriented control strategy via multi-subspace coordination, aiming to enhance system reliability. First, a mathematical model of the motor under both normal operation and single-phase open-circuit fault conditions is established using the space vector decoupling method. Then, by coordinating current distribution between the d - q and z_1 - z_2 - z_3 subspaces with the objectives of minimizing phase current amplitude and optimizing stator copper loss, high-precision control under fault conditions is achieved. Simulation results demonstrate that, compared to traditional current hysteresis fault-tolerant control, the proposed strategy reduces torque ripple by 75.26% and copper loss by 2.67% while maintaining output torque accuracy, significantly enhancing the system's fault-tolerant performance and energy efficiency.

Keywords: Dual three-phase PMSM · Extreme space environment · Fault-tolerant control · Field-oriented control · Subspace coordination

1 Introduction

Recent aerospace applications require propulsion systems with enhanced operational dependability, stimulating substantial progress in multiphase motor drive technology development. Compared with traditional three-phase systems, multiphase motors have become an ideal choice for high-reliability applications—such as aerospace propulsion, electric vehicles, and marine power systems [1–6]—due to their superior fault tolerance, lower torque ripple, and higher power density.

Supported by National Natural Science Foundation of China (Grant No. U24B6014 and No. 62203222), and the Fundamental Research Funds for the Central Universities (No. 30925010526).

In aerospace applications, intense radiation and thermal cycling effects significantly increase the probability of winding open-circuit faults. Therefore, robust fault-tolerant control strategies must be employed to ensure continuous and stable system operation. For multi-three-phase motors, the simplest fault-tolerant strategy is to directly disconnect the entire set of faulty phase windings. However, this leads to a reduction in overall power. In references [7–9], the fault-tolerant characteristics of a six-phase permanent magnet motor were analyzed under the vector control framework, and a method to achieve fault-tolerant operation by adjusting the reference current was proposed. However, its dynamic performance is poor, and the parameter adjustment is complex, limiting its practical application. Due to its fast response, direct torque control has been introduced into the fault-tolerant systems of multi-phase motors [10–13]. However, the hysteresis controller exacerbates torque fluctuations, especially since harmonic suppression becomes more challenging after a fault.

Fault-tolerant field-oriented control (FT-FOC) can reconstruct effective torque and excitation currents in the remaining phases of multiphase motors under open-circuit faults, achieving smooth torque output and reduced ripple [14–16]. However, this strategy increases current amplitude in healthy phases and causes additional resistive losses from zero-sequence currents without contributing to torque, leading to reduced efficiency and overheating risks.

To address the contradiction between fault tolerance and system efficiency in FT-FOC strategies, this paper proposes a multi-subspace coordinated fault-tolerant field-oriented control (MSFT-FOC) strategy for single-phase open-circuit faults in DT-PMSM for aerospace applications. The strategy maintains a constant torque output in the α - β subspace while optimizing copper loss in the zero-sequence subspace and compensates for the missing magnetomotive force through improved current reconstruction. Simulations show the proposed method significantly reduces torque ripple and copper loss compared to hysteresis-based approaches, without hardware redundancy, proving valuable for weight-constrained spacecraft propulsion.

2 Modeling Healthy and Faulty Motors

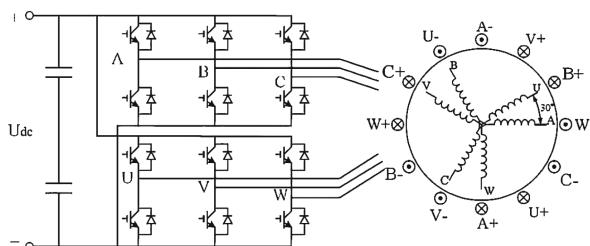


Fig. 1. DT-PMSM Drive System.

2.1 Mathematical Model of Six-Phase System

Figure 1 shows the DT-PMSM drive structure with two 30° -offset three-phase windings (A-B-C and U-V-W) sharing a non-isolated neutral point. Each set of windings is controlled by its respective three-phase inverter bridge circuit, and both drive circuits share a common DC bus.

Neglecting magnetic saturation, leakage inductance, and iron losses with sinusoidally distributed windings, the DT-PMSM model equations the six-phase stationary coordinate system are:

$$\mathbf{U}_s = \mathbf{R}_s \mathbf{I}_s + \frac{d}{dt} \boldsymbol{\psi}_s \quad (1)$$

$$\boldsymbol{\psi}_s = \mathbf{L}_s \mathbf{I}_s + \boldsymbol{\gamma}_s \boldsymbol{\psi}_{fd} \quad (2)$$

$$T_e = p_n \left(\frac{1}{2} \mathbf{i}_s^T \frac{\partial \mathbf{L}_s}{\partial \theta} \mathbf{i}_s + \mathbf{i}_s^T \frac{\partial \boldsymbol{\gamma}_s}{\partial \theta} \boldsymbol{\psi}_{fd} \right) \quad (3)$$

$$\frac{d\omega_m}{dt} = \frac{1}{J} (T_e - B\omega_m - T_L) \quad (4)$$

where $\theta = p_n \theta_m$ is the electrical angle.

The variables used in the aforementioned equations are defined in Table 1.

Table 1. Variable definitions for DT-PMSM model

Symbol	Unit	Description	Symbol	Unit	Description
\mathbf{U}_s	V	Stator voltage vector	T_e	N · m	Electromagnetic torque
\mathbf{R}_s	Ω	Stator resistance matrix	p_n	—	Pole pair number
\mathbf{I}_s	A	Stator current vector	θ_m	rad	Mechanical angle
$\boldsymbol{\psi}_s$	Wb	Stator flux linkage vector	J	$\text{kg} \cdot \text{m}^2$	Moment of inertia
\mathbf{L}_s	H	Stator inductance matrix	B	$\text{N} \cdot \text{m} \cdot \text{s}/\text{rad}$	Damping coefficient
$\boldsymbol{\psi}_{fd}$	Wb	PM flux amplitude	T_L	N · m	Load torque
$\boldsymbol{\gamma}_s$	—	Flux coefficient matrix	ω_m	rad/s	Mechanical angular velocity

2.2 Vector Space Decoupling Mathematical Model of DT-PMSM

The DT-PMSM's nonlinear strong coupling complicates analysis and control, but can be resolved by vector space decoupling (VSD) [17], which can decouple the mathematical model into three orthogonal sub-planes: α - β , z_1 - z_2 , and $o1$ - $o2$.

$$T_{6s} = \frac{1}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (5)$$

By transforming the matrix equation (5), the variables are mapped into three stationary subspaces. Among these, only the torque plane α - β participates in electromechanical energy conversion and needs to be transformed into the d - q coordinate system through rotation to simplify the model. In contrast, the harmonic plane $z1$ - $z2$ and the zero-sequence plane $o1$ - $o2$ do not participate in energy conversion and thus do not require rotation. The resulting rotation transformation matrix is P_6 :

$$P_6 = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & I_4 \end{bmatrix} \quad (6)$$

where I_4 is the 4×4 identity matrix.

The final form of the transformation matrix is as follows:

$$T_{6r} = P_6 T_{6s} \quad (7)$$

By applying transformation matrix (7) to (1)–(3), the system is decoupled, yielding the VSD model equations: voltage equation (8), flux linkage equation (9), and electromagnetic torque equation (10).

$$U_{vsd} = T_{6r} U_s = T_{6r} R_s I_s + T_{6r} \frac{d}{dt} \psi_s \quad (8)$$

$$\psi_{vsd} = T_{6r} \psi_s = (T_{6r} L_s T_{6r}^{-1})(T_{6r} I_s) + T_{6r} \gamma_s \psi_{fd} \quad (9)$$

$$T_{vsd} = 3p_n [(L_d - L_q)i_d i_q + i_q \psi_{fd}] \quad (10)$$

2.3 VSD-Based Mathematical Modeling of DT-PMSM Under Single-Phase Open-Circuit Fault

For fault-tolerant control design of DT-PMSM, an accurate mathematical model of the faulty motor must be established. For convenience, the variables of the fault motor are still represented by the symbols used during normal operation. The voltage and flux equations under W-phase open-circuit fault are as follows:

$$u_{5s} = R_{5s} i_{5s} + \frac{d\psi_{5s}}{dt} \quad (11)$$

$$\psi_{5s} = L_{5s}(\theta) i_{5s} + \gamma_{5s}(\theta) \psi_{fd} \quad (12)$$

In this context, the magnetic flux coefficient matrix $\gamma_{5s}(\theta)$ is obtained by removing the sixth row from matrix $\gamma_s(\theta)$. The inductance coefficient matrix $L_{5s}(\theta)$ and the resistance coefficient matrix R_{5s} are obtained by removing the sixth row and the sixth column from matrices R_s and $L_s(\theta)$, respectively.

$$T_{5s} = \frac{1}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{1}{5} & \frac{\sqrt{3}}{2} - \frac{1}{5} & -\frac{\sqrt{3}}{2} - \frac{1}{5} & \frac{1}{2} - \frac{1}{5} & \frac{1}{2} - \frac{1}{5} \\ 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ -\frac{2}{3} & -\frac{2}{3} - \frac{\sqrt{3}}{3} & -\frac{2}{3} + \frac{\sqrt{3}}{3} & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (13)$$

Through VSD-based decoupling, the single-phase open-circuit DT-PMSM model is separated into two orthogonal coordinate systems: the α - β frame and the z_1 - z_2 - z_3 frame [18]. The corresponding stationary transformation matrix for the faulted motor is shown in (13).

The rotational transformation matrix P_5 , defined in (14), maps the α - β subspace to the rotating d - q frame. In contrast, the z_1 - z_2 - z_3 subspace remains untransformed since it is not involved in the energy conversion process between electrical and mechanical domains.

$$P_5 = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & I_3 \end{bmatrix} \quad (14)$$

The final form of the transformation matrix is as follows:

$$T_{5r} = P_5 T_{5s} \quad (15)$$

Similar to normal motor modeling, applying (15) to (11) and (12) yields the d - q axis stator voltage equations:

$$U_{dq} = R_{dq} I_{dq} + L_{dq} \frac{dI_{dq}}{dt} + \left(\frac{dL_{dq}}{dt} - \Omega L_{dq} \right) I_{dq} + \omega \psi_{fd} N(\theta) \quad (16)$$

where $N(\theta) = \begin{bmatrix} -0.2 \sin 2\theta \\ 0.8 - 0.2 \cos 2\theta \end{bmatrix}$.

The electromagnetic torque is derived as:

$$T_e = 3p_n [(3L_d - 3L_q)i_d i_q + i_q \psi_{fd}] \quad (17)$$

3 Design of Fault-Tolerant Control Strategy

This section proposes an MSFT-FOC strategy based on multi-objective optimization for DT-PMSM single-phase faults. The strategy optimizes the distribution of the remaining phase currents to maintain a circular rotating magnetic field while minimizing current amplitude and copper loss. The d - q subspace retains conventional control, while the z_1 - z_2 - z_3 subspace introduces closed-loop control to achieve fault-tolerant current tracking, balancing torque output and efficiency. Additionally, a current hysteresis control method based on flux conservation is introduced for comparative studies in subsequent chapters.

3.1 Multi-Objective Optimized Post-Fault Current Reconstruction

During operation, the motor establishes a circular rotating magnetic field to maintain smooth torque output. The phase currents in a healthy DT-PMSM are expressed as:

$$\begin{cases} i_A = I_m \cos \theta, i_B = I_m \cos(\theta - 120^\circ), i_C = I_m \cos(\theta + 120^\circ) \\ i_U = I_m \cos(\theta - 30^\circ), i_V = I_m \cos(\theta - 150^\circ), i_W = I_m \cos(\theta + 90^\circ) \end{cases} \quad (18)$$

where I_m represents the current amplitude and θ represents the electrical angle of phase A current.

The expression for the synthesized magnetomotive force in a healthy DT-PMSM during normal operation is:

$$F = \frac{3}{2}NI_m \cos(\theta - \phi) = \frac{3}{4}NI_m(e^{j\theta}e^{-j\phi} + e^{-j\theta}e^{j\phi}) \quad (19)$$

Under open-phase fault of winding W, to preserve the original rotating magnetomotive force characteristics, the remaining five-phase currents should be regulated to meet:

$$3I_m e^{j\theta} = i_A + i_B e^{j120^\circ} + i_C e^{-j120^\circ} + i_U e^{j30^\circ} + i_V e^{j150^\circ} \quad (20)$$

Express the remaining phase currents in the following form:

$$i_K = a_K I_m \cos \theta + b_K I_m \sin \theta = a_K i_\alpha + b_K i_\beta \quad (21)$$

where $K = \{A, B, C, U, V\}$, and a_K, b_K are coefficients to be determined.

Substituting (21) into (20) and separating real and imaginary parts yields:

$$\begin{cases} a_A - 0.5a_B - 0.5a_C + \sqrt{3}a_U/2 - \sqrt{3}a_V/2 = 3 \\ b_A - 0.5b_B - 0.5b_C + \sqrt{3}b_U/2 - \sqrt{3}b_V/2 = 0 \\ \sqrt{3}a_B/2 - \sqrt{3}a_C/2 + 0.5a_U - 0.5a_V = 0 \\ \sqrt{3}b_B/2 - \sqrt{3}b_C/2 + 0.5b_U - 0.5b_V = 3 \end{cases} \quad (22)$$

Since the motor's neutral point is not isolated, the sum of the remaining five phase currents must be zero, resulting in the following constraint condition:

$$\begin{cases} a_A + a_B + a_C + a_U + a_V = 0 \\ b_A + b_B + b_C + b_U + b_V = 0 \end{cases} \quad (23)$$

For the concurrent optimization of both phase current amplitude minimization and stator copper loss reduction, the multi-objective optimization problem is formulated as:

$$\begin{cases} \min \mathbf{F}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x})]^T \\ \text{subject to (22) -- (23)} \end{cases} \quad (24)$$

where $f_1(\mathbf{x}) = \max_{k \in K}(a_k^2 + b_k^2)$ is the maximum squared phase current, and $f_2(\mathbf{x}) = \sum_{k \in K}(a_k^2 + b_k^2)$ represents the copper-loss-related term.

The optimal parameter solution is obtained through multi-objective optimization using the NSGA-II algorithm, yielding:

$$\begin{cases} \mathbf{a}_K = [0.9284, -0.3009, -0.4209, 0.8845, -1.0909]^T \\ \mathbf{b}_K = [-0.7879, 1.3993, -1.5327, 0.8773, 0.0444]^T \end{cases} \quad (25)$$

The expressions for the remaining five-phase currents are as follows:

$$\begin{cases} i_A = 0.9284I_m \cos \theta - 0.7879I_m \sin \theta \approx 1.22I_m \cos(\theta + 40.3^\circ) \\ i_B = -0.3009I_m \cos \theta + 1.3993I_m \sin \theta \approx 1.43I_m \cos(\theta - 77.8^\circ) \\ i_C = -0.4209I_m \cos \theta - 1.5327I_m \sin \theta \approx 1.59I_m \cos(\theta + 74.6^\circ) \\ i_U = 0.8845I_m \cos \theta + 0.8773I_m \sin \theta \approx 1.25I_m \cos(\theta - 44.8^\circ) \\ i_V = -1.0909I_m \cos \theta + 0.0444I_m \sin \theta \approx 1.09I_m \cos(\theta - 0.1^\circ) \end{cases} \quad (26)$$

3.2 Fault-Tolerant Control Strategy Based on Current Hysteresis

The hysteresis current control features a simple structure and does not rely on mathematical models. After determining the optimal fault-tolerant current reference values for the remaining stator windings, the system compares the commanded current with the measured phase current. The hysteresis controller generates control signals based on the magnitude and direction of the current error to drive the inverter state switching. By confining the current deviation within the hysteresis bandwidth, the system achieves fast and accurate current tracking. As shown in Fig. 2, the DT-PMSM employs a dual closed-loop fault-tolerant control scheme: the inner loop implements hysteresis current control, while the outer loop performs PI speed regulation.

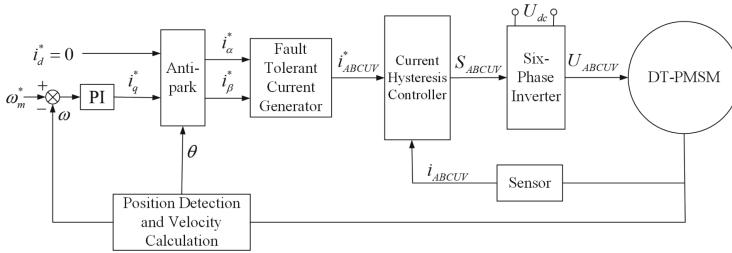


Fig. 2. DT-PMSM Current Hysteresis Fault-Tolerant Control System.

3.3 Fault-Tolerant Control Strategy Based on MSFT-FOC

The MSFT-FOC strategy achieves fault-tolerant control through coordinated operation of two decoupled subspaces: torque generation control in the d-q subspace and harmonic suppression control in the z1-z2-z3 zero-sequence subspace for fault compensation. The overall control architecture is illustrated in Fig. 3.

In the z1-z2-z3 subspace, there is no coupling between axes, allowing independent design of current controllers; whereas strong coupling exists between the d-q axes, requiring a feedforward compensator $M(\theta)$ for decoupling control. The system introduces decoupling quantities u_m and u_n , whose relationship with the d-q axis voltages is given by:

$$U_{mn} = M^{-1}(\theta)U_{dq} = [u_m \ u_n]^T \quad (27)$$

where $M(\theta) = \begin{bmatrix} 0.8 + 0.2 \cos 2\theta & -0.2 \sin 2\theta \\ -0.2 \sin 2\theta & 0.8 - 0.2 \cos 2\theta \end{bmatrix}$.

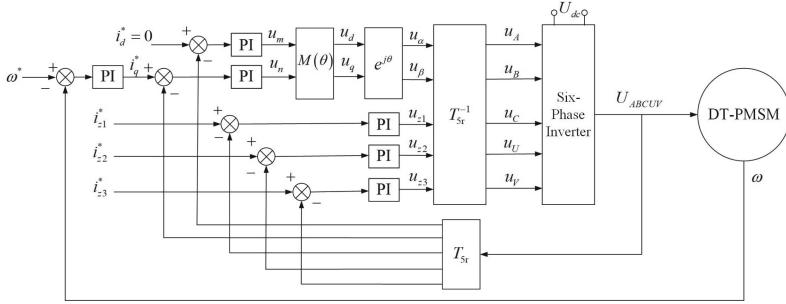


Fig. 3. DT-PMSM Fault-Tolerant Control System Based on MSFT-FOC.

Equation (26) presents the optimal phase current expressions that achieve both stator current amplitude minimization and copper loss optimization. Through transformation by T_{5r} , the required current injection magnitudes for i_{z1} , i_{z2} , and i_{z3} can be obtained as follows:

$$\begin{cases} i_{z1}^* = -0.141i_\alpha^* - 0.481i_\beta^* = 0.141i_d^* \sin \theta - 0.481i_q^* \cos \theta \\ i_{z2}^* = -0.138i_\alpha^* - 0.052i_\beta^* = 0.138i_d^* \sin \theta - 0.052i_q^* \cos \theta \\ i_{z3}^* = 0 \end{cases} \quad (28)$$

4 Simulation and Results

4.1 Motor Simulation Setup and Parameter Configuration

The parameters of the simulation motor are shown in Table 2. The simulation conditions are as follows: DC bus voltage $U_{dc} = 300$ V, speed setpoint is set to 500 rpm, and load torque is set to 50 N · m, with the motor starting under load.

Table 2. Simulation motor parameters

Parameters	R (Ω)	L_d (mH)	L_q (mH)	p	Ψ_m (Wb)	J ($\text{kg} \cdot \text{m}^2$)
Value	1.4	7.9	7.9	3	0.08	0.015

4.2 Motor Fault Operation Without Fault-Tolerant Control

Fig. 4 shows the results under single-phase open-circuit condition without fault-tolerant control. At the instant of disconnection, the current surges abruptly, causing severe fluctuations in both rotational speed and torque. During fault operation, the remaining phase currents exhibit significant distortion with increased amplitudes, and their higher harmonics exacerbate vibration, noise, and operational instability. The non-sinusoidal waveforms also induce additional losses and reduce motor lifespan. Persistent speed and torque fluctuations lead to substantial performance degradation. Therefore, investigating fault-tolerant control strategies for motors under failure conditions is crucial.

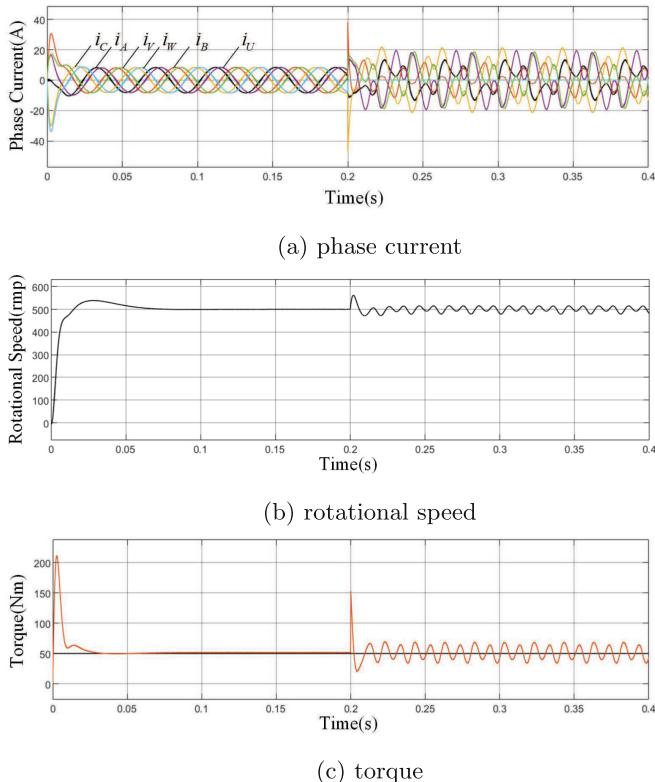
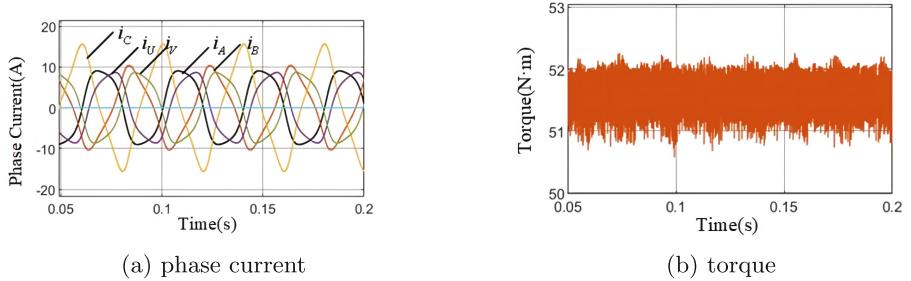
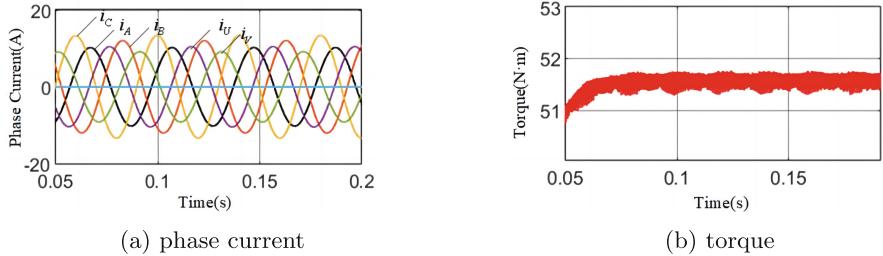


Fig. 4. Operating results graph of faulty motor without fault-tolerant control.

4.3 Motor Fault Operation with Fault-Tolerant Control

A comparative analysis of fault-tolerant control performance is shown in Fig. 5 and Fig. 6 for current hysteresis control and MSFT-FOC, respectively. Table 3

**Fig. 5.** Current Hysteresis Fault-Tolerant Control.**Fig. 6.** Fault-Tolerant Control System Based on MSFT-FOC.

displays some performance indicators of the two control strategies during motor fault operation.

Under normal operating conditions of the DT-PMSM, the phase current amplitude is 8.5A with an ideal sinusoidal waveform. After an open-phase fault occurs, both fault-tolerant control strategies result in an increase in phase current amplitude. However, the MSFT-FOC strategy demonstrates superior performance: its maximum phase current amplitude is 13.4A, which is 14.4% lower than the 15.65A of the current hysteresis control, indicating that the multi-subspace collaborative optimization effectively improves current distribution balance. According to the current waveform comparison in Fig. 5a and Fig. 6a, the hysteresis control exhibits significant distortion, while the MSFT-FOC control maintains an ideal sinusoidal current. Copper losses decrease from 401.9W to 391.2W, a reduction of 2.67%, achieving higher energy efficiency. Meanwhile, torque ripple drops sharply from 1.94% to 0.48%, a reduction of 75.26%, significantly improving control precision, which is particularly critical for high-precision applications such as aerospace. Overall, the MSFT-FOC strat-

Table 3. Performance metrics of different control strategies

Performance metric	Current Hysteresis	MSFT-FOC
Maximum Amplitude of Phase Current (A)	15.65	13.4
Copper loss (W)	401.9	391.2
Torque Ripple (%)	1.94	0.48

egy demonstrates clear advantages in current stress, energy efficiency, and torque quality.

5 Conclusion

This paper proposes a FT-FOC strategy based on multi-objective optimization and multi-subspace coordination for single-phase open-circuit faults in DT-PMSM. By establishing a unified motor model through space vector decoupling, the strategy maintains torque output in the d - q subspace while optimizing copper loss in the z_1 - z_2 - z_3 subspace, achieving dual objectives of minimizing phase current amplitude and reducing power loss.

Simulation results show that, compared with conventional hysteresis control, the proposed method reduces torque ripple and copper loss without additional hardware. Its computational load—primarily linear matrix operations and PI controllers—is easily handled by industrial DSPs, ensuring real-time feasibility while improving reliability and efficiency.

References

1. Jiang, X., Huang, W., Li, J., et al.: Overview of high reliability fault-tolerant permanent magnet motor and its drive system for aerospace applications. *Micromotors* **53**, 108–118 (2020). <https://doi.org/10.15934/j.cnki.micromotors.2020.06.021>
2. Wang, L., Aleksandrov, S., Tang, Y., et al.: Fault-tolerant electric drive and space-phasor modulation of flux-switching permanent magnet machine for aerospace application. *IET Electr. Power Appl.* **11**, 1416 (2017)
3. Hasoun, M., Afia, A.E., Khafallah, M.: Performance comparison of two SVPWM strategies based vector space decomposition controlled dual three-phase PMSM for electric ship propulsion. In: 7th International Renewable and Sustainable Energy Conference, pp. 1–6. IEEE, Agadir, Morocco (2019)
4. Abdel Aziz, G., Amin, M., El-Sousy, F.F.M., et al.: MRAS-based super-twisting sliding-mode observer with DTC of six-phase induction motor for ship propulsion application. *IEEE Trans. Ind. Appl.* **57**, 6646–6658 (2021)
5. Gaber, M., El-Banna, S.H., Hamad, M.S., et al.: Studying the effect of using multi-phases switched reluctance motor to reduce the torque ripple for ship propulsion system. In: 2020 IEEE PES/IAS PowerAfrica. pp. 1–5. IEEE, Nairobi, Kenya (2020)
6. Kui, Y., Gui, D.: A vector control method for 15-phase motor of marine electric propulsion system. *J. Shanghai Ship Transp. Res. Inst.* **45**, 15–22 (2022)
7. Munim, W.N.W.A., et al.: A unified analysis of the fault tolerance capability in six-phase induction motor drives. *IEEE Trans. Power Electron.* **32**, 7824–7836 (2016)
8. Ye, D., et al.: Study on steady-state errors for asymmetrical six-phase permanent magnet synchronous machine fault-tolerant predictive current control. *IEEE Trans. Power Electron.* **35**, 640–651 (2019)
9. Eldeeb, H.M., et al.: Pre-and postfault current control of dual three-phase reluctance synchronous drives. *IEEE Trans. Ind. Electron.* **67**, 3361–3373 (2019)
10. Zhou, Y., Chen, G.: Predictive DTC strategy with fault-tolerant function for six-phase and three-phase PMSM series-connected drive system. *IEEE Trans. Ind. Electron.* **65**, 9101–9112 (2017)

11. Freire, N.M.A., Cardoso, A.J.M.: A fault-tolerant direct controlled PMSG drive for wind energy conversion systems. *IEEE Trans. Ind. Electron.* **61**, 821–834 (2013)
12. Jiang, B.: Three-Dimensional Space Vector Model Predictive Control of Open-Winding Permanent Magnet Synchronous Motor. Ph.D. dissertation, Yanshan University (2023). <https://doi.org/10.27440/d.cnki.gysdu.2023.002061>
13. Li, K., et al.: A novel fault-tolerant control technique for an inverter with hysteresis current. *Comput. Sci. Eng.* **21**, 92–102 (2018)
14. Zhu, X., et al.: Passive fault-tolerant scheme of a 2×3 -phase SPMSM driven by mono-inverter based on field oriented control. *IEEE Trans. Ind. Informat.* **20**(4), 7040–7052 (2024)
15. Prieto, I.G., et al.: Field-oriented control of multiphase drives with passive fault tolerance. *IEEE Trans. Ind. Electron.* **67**(9), 7228–7238 (2019)
16. Zhang, L., et al.: Research on fault-tolerant field-oriented control of a five-phase permanent magnet motor based on disturbance adaption. *Energies* **15**(9), 3436 (2022)
17. Zhao, Y.F., Lipo, T.A.: Space vector PWM control of dual three-phase induction machine using vector space decomposition. *IEEE Trans. Ind. Appl.* **31**(5), 1100 (1995)
18. Zhao, Y.F., Lipo, T.A.: Modeling and control of a multi-phase induction machine with structural unbalance, Part I—machine modeling and multi-dimensional current regulation. *IEEE Trans. Energy Convers.* **11**(3), 570–577 (1996)



Three-Vector-Based Model Predictive Flux Control For SPMSM Using Ultralocal Model

Rongbo Lai, Lu Liu, Shihong Ding^(✉), Keqi Mei, and Jinlin Sun

School of Electrical and Information Engineering, Jiangsu University, Zhenjiang,
Jiangsu 212013, China
2222307080@stmail.ujs.edu.cn, {liulu2021,dsh,mkq,jsun}@ujs.edu.cn

Abstract. This article proposes a three-vector-based model predictive flux control(TV-MPFC) using an ultralocal model. The conventional mathematical model of PMSM is replaced by an ultralocal model, and an extended state observer (ESO) is designed to estimate the unknown dynamics in the model, thereby eliminating disturbances caused by unmodeled dynamics. The control objective of the inner loop is shifted from current vectors to stator flux vectors, and three voltage vectors are applied within one period to achieve zero dq-axis flux tracking error. Simulations and experiments demonstrate that the proposed model predictive control achieves excellent current and speed regulation.

Keywords: Model predictive flux control · ultralocal model · three voltage vectors

1 Introduction

Over the past few years, surface-mounted permanent magnet synchronous motor (SPMSM) have become increasingly popular in various industries, including electric vehicles, industrial automation, and aerospace. This is largely due to their exceptional efficiency, precision, and reliability [1]. However, precise control of SPMSM remains challenging due to issues like system parameter variations, external load disturbances, and model uncertainties. Conventional control methods, such as Field-Oriented Control (FOC), have limitations in handling complex conditions and high-precision requirements [2].

In order to overcome these challenges, Model predictive control (MPC) as an innovative control strategy is gradually attracting interest from researchers [3]. MPC predicts future system behavior and optimizes control inputs to ensure stable and precise motor control [4]. Three-vector-based model predictive control (TV-MPC), a novel MPC approach, applies three voltage vectors each control period to more flexibly shape the desired voltage vector, reducing current ripple and boosting control accuracy. For instance, [5] introduces a three-vector-based model predictive torque control that effectively cuts current distortion and

torque ripple. Reference [6] presents a three-vector model predictive control that simplifies voltage vector selection using Direct Torque Control principles.

However, conventional MPC relies on accurate mathematical models, which are challenging to derive, especially under varying parameters and external disturbances [7]. Ultralocal models, as a data-driven approach, reflect the system's current state in real-time and adapt to changes through online parameter estimation. Integrating ultralocal models with MPC enhances the control system's robustness and adaptability, thereby mitigating reliance on precise mathematical models [8].

The contributions of this article are as follows: First, an ultra-local model is adopted to replace the PMSM flux linkage mathematical model, thereby eliminating unmodeled dynamics including voltage dead-time effects. Second, by establishing the analytical relationship between reference voltage and flux linkage, the proposed method directly selects based voltage vectors, which circumvents the conventional three-vector-based MPC approach requiring exhaustive computation of cost functions across all voltage vector combinations. Finally, the derived ultra-local PMSM model can be used for the calculation of voltage duty cycle, thereby reducing MPC computational load. The simulation and experimental results confirm that this approach enables precise motor control.

2 Machine Equation of SPMSM And ESO

2.1 Machine Equation of SPMSM

The discretised mathematical model of the stator flux linkage for SPMSM is presented as follows:

$$\begin{cases} \psi_d(k) = L_d i_d(k) + \psi_f \\ \psi_q(k) = L_q i_q(k) \end{cases} \quad (1)$$

$$\begin{cases} \psi_d(k+1) = T_s [U_d(k) - R_s i_d(k) + \omega_e \psi_q(k)] + \psi_d(k) \\ \psi_q(k+1) = T_s [U_q(k) - R_s i_q(k) - \omega_e \psi_d(k)] + \psi_q(k) \end{cases} \quad (2)$$

where $U_d(k)$, $U_q(k)$, $i_d(k)$, $i_q(k)$, $L_d(k)$, $L_q(k)$, $\psi_d(k)$ and $\psi_q(k)$ are the d-axis and q-axis components of the input voltage, stator current, stator inductance and stator flux linkage at the constant k , respectively. R_s is the stator resistance. ψ_f is the permanent magnet flux. T_s is the interval of a control period. Equation (2) presents an idealised mathematical model. However, the actual SPMSM model includes components like inverter dead-time effects, which are difficult to obtain the precise mathematical models. This article uses $\mathbf{f}(k) = \begin{bmatrix} f_d(k) \\ f_q(k) \end{bmatrix}$ to denote these unmodelled dynamics, and Eq. (2) is thus reformulated as follows:

$$\begin{cases} \psi_d(k+1) = T_s [f_d(k) + U_d(k) - R_s i_d(k) + \omega_e \psi_q(k)] + \psi_d(k) \\ \psi_q(k+1) = T_s [f_d(k) + U_q(k) - R_s i_q(k) - \omega_e \psi_d(k)] + \psi_q(k) \end{cases} \quad (3)$$

2.2 ESO Based on Ultralocal Model

Express (3) in the form of a first-order discretized ultralocal model:

$$\frac{\mathbf{y}(k+1) - \mathbf{y}(k)}{T_s} = \mathbf{F}(k) + \alpha \mathbf{u}(k) \quad (4)$$

where $\mathbf{y}(k) = \psi(k) = \begin{bmatrix} \psi_d(k) \\ \psi_q(k) \end{bmatrix}$ is control output variables, $\mathbf{u}(k) = \begin{bmatrix} U_d(k) \\ U_q(k) \end{bmatrix}$ is control input variables, $\mathbf{F}(k) = \begin{bmatrix} -R_s i_d(k) + \omega_e \psi_q(k) + f_d(k) \\ -R_s i_q(k) - \omega_e \psi_d(k) + f_q(k) \end{bmatrix}$ is unknown quantity and $\alpha = 1$ is control gain.

Based on Eq. (4), the following second-order linear ESO is constructed.

$$\begin{cases} \mathbf{e}_\psi(k) = \hat{\psi}(k) - \psi(k) \\ \hat{\psi}(k+1) = \hat{\psi}(k) - \beta_1 \mathbf{e}_\psi(k) + T_s (\hat{\mathbf{F}}(k) + \alpha \mathbf{u}(k)) \\ \hat{\mathbf{F}}(k+1) = \hat{\mathbf{F}}(k) - \beta_2 \mathbf{e}_\psi(k) \end{cases} \quad (5)$$

where $\hat{\psi}(k)$ and $\hat{\mathbf{F}}(k)$ are estimates of $\psi(k)$ and $\mathbf{F}(k)$, respectively. $\beta_1 = 2\omega_0 T_s$ and $\beta_2 = \omega_0^2 T_s$ are the gains of observer. ω_0 is the bandwidth of the ESO. The stability of ESO has been demonstrated in [9] and for the sake of brevity this article will not repeat the discussion on the stability of ESO.

3 Principle of the Proposed MPC

3.1 Selection of Action Vectors

Conventional TV-MPC requires evaluating the control effects of different voltage vector combinations, incurring a substantial computational burden. To avoid the enormous computation from enumerating these combinations, this article leverages the flux linkage deadbeat principle to calculate the reference voltage, enabling direct selection of the two adjacent voltage vectors and the voltage zero vector in the reference voltage vector's sector. The reference flux can be provided by (6):

$$\begin{cases} \psi_d^{ref} = \psi_f \\ \psi_q^{ref} = \frac{T_e^{ref} L_q}{1.5 n_p \psi_f} \end{cases} \quad (6)$$

where the reference torque T_e is derived from the PI controller within the speed loop. And n_p is the number of pole pairs of the SPMSM. By combining Eqs. (4), (5) and (6), the reference voltage can be obtained.

$$\begin{cases} U_d^{ref} = \frac{\Delta \psi_d}{\alpha T_s} - \frac{\hat{F}_d(k+1)}{\alpha} \\ U_q^{ref} = \frac{\Delta \psi_q}{\alpha T_s} - \frac{\hat{F}_q(k+1)}{\alpha} \end{cases} \quad (7)$$

where the flux linkage errors in the d-axis and q-axis are denoted as $\Delta\psi_d = \psi_d^{ref} - \hat{\psi}_d(k+1)$ and $\Delta\psi_q = \psi_q^{ref} - \hat{\psi}_q(k+1)$, respectively. It should be noted that the ESO eliminates the additional computational burden associated with delay compensation. The angle of the reference voltage vector can be obtained through Eq. (8).

$$\theta^{ref} = \arctan\left(\frac{U_q^{ref}}{U_d^{ref}}\right) + \theta_r \quad (8)$$

where θ_r is the electrical angle of the rotor. As an example, consider the reference voltage vector in Sector I. Figure 1 demonstrates that the required target vectors in this sector are the adjacent non-zero vectors u_1 and u_2 , combined with the zero vectors u_0 and u_7 . Table 1 presents a detailed overview of the target vector combinations for the reference voltage in different sectors.

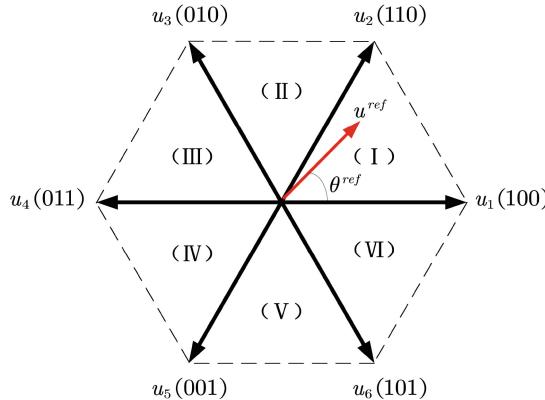


Fig. 1. Voltage space vector diagram.

Table 1. Target vectors lookup table

Sector	The target vectors
I ($0 \leq \theta^{ref} < \frac{\pi}{3}$)	u_1, u_2, u_0 and u_7
II ($\frac{\pi}{3} \leq \theta^{ref} < \frac{2\pi}{3}$)	u_2, u_3, u_0 and u_7
III ($\frac{2\pi}{3} \leq \theta^{ref} < \pi$)	u_3, u_4, u_0 and u_7
IV ($\pi \leq \theta^{ref} < \frac{4\pi}{3}$)	u_4, u_5, u_0 and u_7
V ($\pi \leq \theta^{ref} < \frac{4\pi}{3}$)	u_5, u_6, u_0 and u_7
VI ($\frac{5\pi}{3} \leq \theta^{ref} < 2\pi$)	u_1, u_6, u_0 and u_7

3.2 Vector Acting Time Calculation

To achieve zero tracking error in the dq-axis flux, equation set (9) can be obtained.

$$\begin{cases} \Delta\psi_d = \sum_{i=0}^2 T_i (F_d(k+1) + \alpha U_{id}) \\ \Delta\psi_q = \sum_{i=0}^2 T_i (F_q(k+1) + \alpha U_{iq}) \end{cases} \quad (9)$$

where T_i respectively represent the acting time of the three target voltage vectors. U_{id} and U_{iq} are the d-axis and q-axis components of the three target voltage vectors, respectively. Solving equation set (9) yields the acting time of each voltage vector within a period.

$$\begin{cases} T_1 = \frac{\Delta\psi_d(k_{2q} - k_{0q}) + \Delta\psi_q(k_{0d} - k_{2d}) + T_s(k_{0q}k_{2d} - k_{2q}k_{0d})}{A} \\ T_2 = \frac{\Delta\psi_d(k_{0d} - k_{1q}) + \Delta\psi_q(k_{1d} - k_{0d}) + T_s(k_{1q}k_{0d} - k_{1d}k_{0q})}{A} \\ T_0 = T_s - T_1 - T_2 \end{cases} \quad (10)$$

where,

$$\begin{cases} k_{0d} = \hat{F}_d(k+1) \\ k_{0q} = \hat{F}_q(k+1) \\ k_{1d} = \hat{F}_d(k+1) + \alpha U_{1d} \\ k_{1q} = \hat{F}_q(k+1) + \alpha U_{1q} \\ k_{2d} = \hat{F}_d(k+1) + \alpha U_{2d} \\ k_{2q} = \hat{F}_q(k+1) + \alpha U_{2q} \\ A = (k_{0q}k_{2d} + k_{1q}k_{0d} + k_{2q}k_{1d} - k_{1q}k_{2d} - k_{2q}k_{0d} - k_{0q}k_{1d}) \end{cases} \quad (11)$$

The block diagram of the proposed TV-MPFC is illustrated in Fig. 2.

4 Simulation and Experimental Results

4.1 Simulation Results

MATLAB simulation models for the conventional one-vector flux model predictive control (Method 1), conventional FOC (Method 2), and the proposed model predictive control (Method 3) have been developed to highlight the advantages of the proposed strategy. To guarantee impartiality, the proportional coefficients of the PI controllers in the speed loops for all three methods are uniformly set to 0.1, and the integral parameters are all set to 2.5. The proportional coefficient

of the PI controller in the current loop for Method 2 is 15, and the integral parameter is 514.87. Additionally, the controllers for all three methods operate at an execution frequency of 20 kHz. The bandwidth of the ESO for Method 3 is selected to be 5000. Table 2 lists the SPMSM parameters used for algorithm verification.

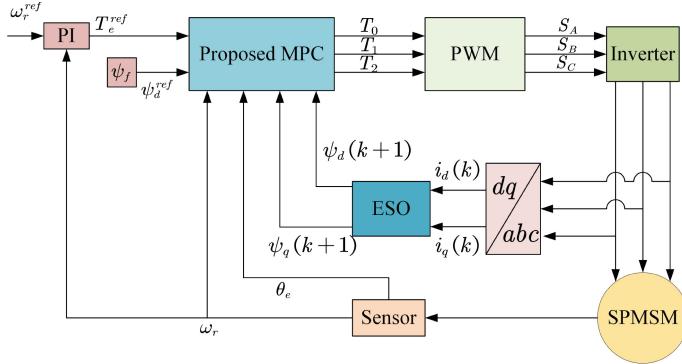


Fig. 2. Control block diagram of the proposed model predictive control.

Table 2. Parameters of the SPMSM.

Parameter	Value
Rated power (kW)	1.5
Rated speed (rpm)	1500
Number of poles pairs	4
Stator resistance (Ω)	1.5
Stator inductance (mH)	4.37
Rotor magnet flux linkage (Wb)	0.142
Rotational inertia (kg m^2)	0.00194

Figure 3 demonstrates significant variations in total harmonic distortion (THD) characteristics across the three approaches. Method 1 demonstrates the highest THD of 15.43%, whereas Method 2 displays a marked improvement with 3.41% THD. The optimal performance is attained by Method 3, achieving a THD value as low as 2.25%. The results demonstrate that method 3 surpasses the other methods in current control performance, exhibiting the most effective regulation of current with the least harmonic distortion and the highest stability. As illustrated in Fig. 4, the speed response curves for the three methods during acceleration from standstill to 1000 rpm reveal distinct performance differences.

It is evident that since Method 1 and Method 3 have faster current regulation, both of them are able to reach a reference speed more quickly than Method 2. Although method 1 has the same speed tracking capability as method 3 but Fig. 5 demonstrates that the torque ripple of method 3 is significantly reduced compared to method 1 and is marginally lower than that of method 2. This shows that the proposed method minimises the torque tracking error.

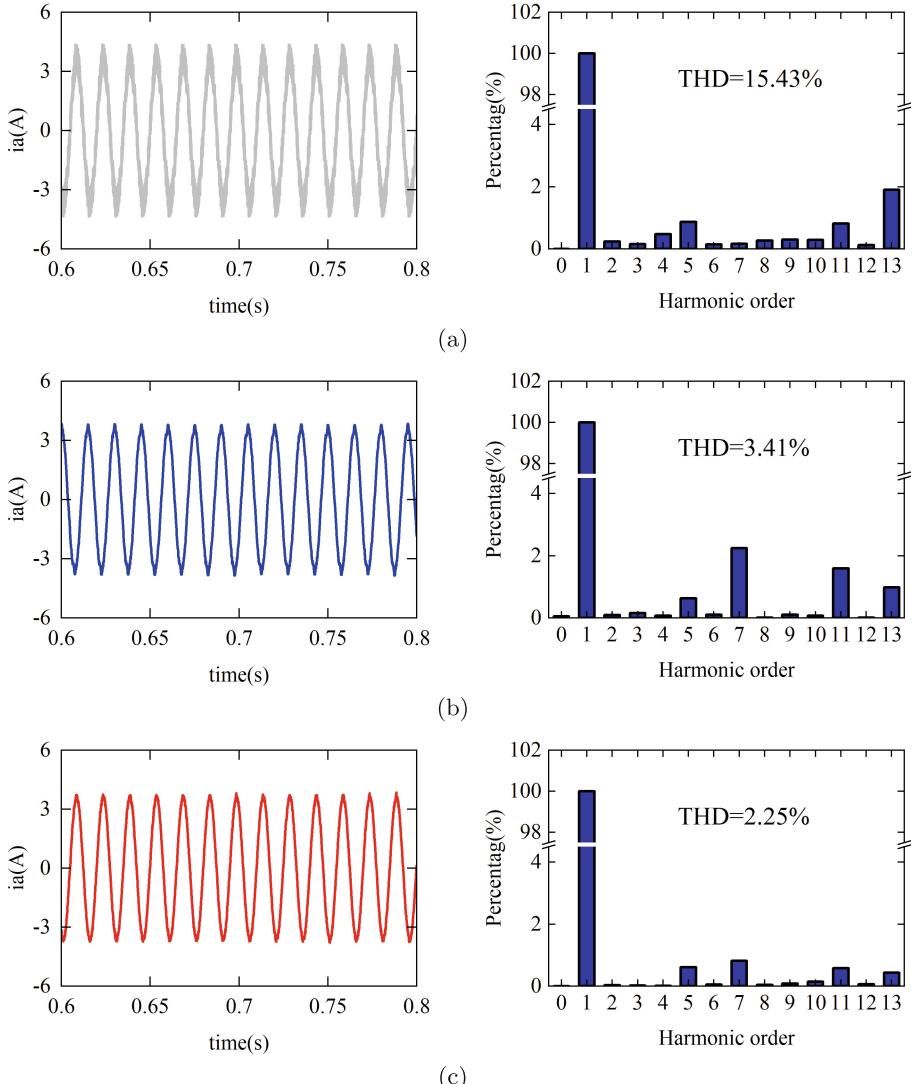


Fig. 3. A-phase current response simulation curve and corresponding FFT analysis.
(a) Method 1. (b) Method 2. (c) Method 3.

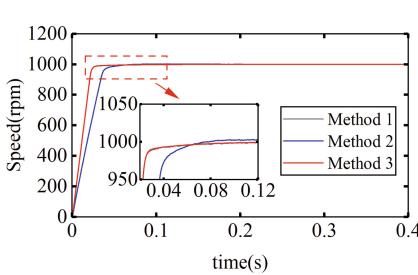


Fig. 4. Speed response simulation curve.

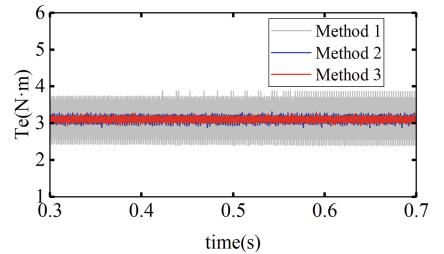


Fig. 5. Torque response simulation curve.

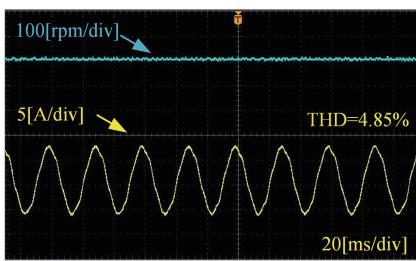


Fig. 6. The steady-state experimental results of the FOC.

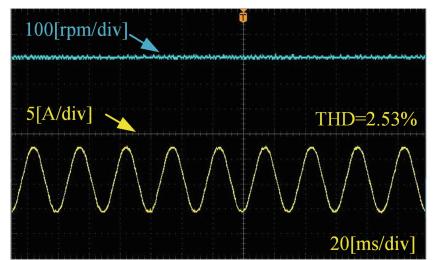


Fig. 7. The steady-state experimental results of the proposed method.

4.2 Experimental Results

Figures 6 and Fig. 7 show the comparison of the speed characteristic curves and A-phase current waveforms between the FOC and the proposed method in steady state operation conditions, respectively. The experimental results indicate that when the speed stabilizes at 500 rpm, the conventional FOC has a THD of 4.85%, whereas the proposed method achieves a substantially decreased THD of 2.53%. The above comparative analysis shows that the control method proposed in this article has better harmonic suppression capability and speed stability under steady state operating conditions, which verifies its effectiveness in SPMSM control optimization.

The proposed TV-MPFC theoretically requires only the inductance parameter, demonstrating stronger parametric robustness compared to conventional MPC. In terms of controller structural complexity, conventional PI controllers require tuning four independent parameters, while TV-MPFC has only the ESO bandwidth as its sole tunable parameter, simplifying parameter configuration. Based on the deadbeat control principle, the TV-MPFC algorithm achieves precise single-cycle tracking of reference currents when the DC bus voltage is sufficient, thus exhibiting superior current tracking capability over PI controllers.

5 Conclusion

The conventional MPC suffers from issues such as inaccurate mathematical modeling and high computational complexity. The TV-MPFC based on the ultralocal model proposed in this article effectively addresses these challenges. By employing an ultralocal model, the proposed method mitigates the impact of unmodeled dynamics. Furthermore, direct selection of the voltage vector based on flux reference eliminates the computationally intensive enumeration process inherent to conventional MPC. Finally, simulations and experiments are conducted to validate the effectiveness of the proposed TV-MPFC.

References

1. Buja, G.S., Kazmierkowski, M.P.: Direct torque control of PWM inverter-fed AC motors—a survey. *IEEE Trans. Industr. Electron.* **51**(4), 744–757 (2004). <https://doi.org/10.1109/TIE.2004.831717>
2. Casadei, D., Profumo, F., Serra, G., Tani, A.: FOC and DTC: two viable schemes for induction motors torque control. *IEEE Trans. Power Electron.* **17**(5), 779–787 (2002). <https://doi.org/10.1109/TPEL.2002.802183>
3. Cortes, P., Kazmierkowski, M.P., Kennel, R.M., Quevedo, D.E., Rodriguez, J.: Predictive control in power electronics and drives. *IEEE Trans. Industr. Electron.* **55**(12), 4312–4324 (2008). <https://doi.org/10.1109/TIE.2008.2007480>
4. Wang, F., Li, S., Mei, X., Xie, W., Rodríguez, J., Kennel, R.M.: Model-based predictive direct control strategies for electrical drives: an experimental evaluation of PTC and PCC methods. *IEEE Trans. Industr. Inf.* **11**(3), 671–681 (2015). <https://doi.org/10.1109/TII.2015.2423154>
5. Chen, L., Xu, H., Sun, X., Cai, Y.: Three-vector-based model predictive torque control for a permanent magnet synchronous motor of EVs. *IEEE Trans. Transport. Electrification* **7**(3), 1454–1465 (2021). <https://doi.org/10.1109/TTE.2021.3053256>
6. Li, X., Xue, Z., Zhang, L., Hua, W.: A low-complexity three-vector-based model predictive torque control for SPMSM. *IEEE Trans. Power Electron.* **36**(11), 13002–13012 (2021). <https://doi.org/10.1109/TPEL.2021.3079147>
7. Zhang, X., Zhang, L., Zhang, Y.: Model predictive current control for PMSM drives with parameter robustness improvement. *IEEE Trans. Power Electron.* **34**(2), 1645–1657 (2019). <https://doi.org/10.1109/TPEL.2018.2835835>
8. Zhou, Y., Li, H., Yao, H.: Model-free control of surface mounted PMSM drive system. In: 2016 IEEE International Conference on Industrial Technology (ICIT), Taipei, Taiwan, pp. 175–180 (2016). <https://doi.org/10.1109/ICIT.2016.7474746>
9. Zhang, Y., Jin, J., Huang, L.: Model-free predictive current control of PMSM drives based on extended state observer using ultralocal model. *IEEE Trans. Industr. Electron.* **68**(2), 993–1003 (2021). <https://doi.org/10.1109/TIE.2020.2970660>



TB-DeepSeek-Agent: A Large Language Model for Tuberculosis Incidence Prediction with Web-Based Automatic Report

Mingming Chen, Yu Lu, and Tenglong Li^(✉)

Xi'an Jiaotong-Liverpool University, Suzhou 215123, China

Tenglong.Li@xjtlu.edu.cn

Abstract. Tuberculosis (TB) remains a global health crisis, necessitating accurate prediction models to guide resource allocation and interventions. Traditional time-series and machine learning models often lack flexibility and precision in capturing complex TB trends. This study introduces TB-DeepSeek-Agent, a novel framework integrating a large language model (DeepSeek) with an Agent-based architecture to enhance TB incidence forecasting. The model leverages historical data, memory mechanisms, and reflective thinking to refine predictions dynamically. Evaluated on 135,802 TB cases from five Chinese cities (2011–2021), TB-DeepSeek-Agent achieved superior performance, with a 100% acceptance ratio (5% error margin), the lowest deviation ratio (0.0129), and RMSE (30.5876), outperforming traditional models (e.g., ARIMA, LSTM) and standalone LLM-based approaches. A web-based application was developed to automate simulations and generate actionable reports, streamlining TB research and policymaking. This work demonstrates the potential of Agent-enhanced AI models in public health analytics, offering a cost-effective solution for TB trend prediction and resource optimization.

Keywords: Tuberculosis Prediction · Large Language Model · Agent Reflective Mechanism · Public health analytics

1 Introduction

Tuberculosis (TB) persists as a global health priority, ranking as the second leading infectious cause of death worldwide [1–3]. Effective prevention and control strategies rely on timely and accurate TB data to inform policy and track epidemiological trends. Countries such as Nigeria have established surveillance systems to collect comprehensive data [4–7], yet the logistical complexity and high costs of large-scale screening impede data acquisition [8, 9]. Predictive modeling offers a viable alternative, leveraging historical data to enable cost-efficient forecasting and timely interventions such as active case-finding. This approach enhances resource optimization and targeted public health strategies, curbing transmission and mitigating outbreak risks [10]. Accurate prediction is critical: underestimation may lead to insufficient resource allocation, while overestimation risks economic strain. Thus, precise forecasting is essential for effective

intervention design and understanding TB's full impact [11]. Traditional time-series models often lack flexibility to capture the complex, evolving dynamics of TB data [12–14]. Consequently, machine learning (ML) and artificial intelligence (AI) methods are increasingly explored to improve trend predictions [15]. Recent studies indicate that hybrid (e.g., CNN-LSTM) and ensemble models (e.g., XGBoost) show superior performance for TB incidence forecasting [16, 17].

Despite their potential, many ML and AI models still face challenges related to robustness, generalizability, and accuracy, often struggling to meet the diverse quality standards required by different health authorities [18, 19]. Additionally, the inconsistent temporal dynamics and limited feature availability in TB data complicate the training and performance of AI models. Leveraging the architecture of Agent, we can potentially predict TB cases with higher accuracy and thereby insight the TB situation correctly. This paper presents the development of an advance Agent-based forecasting model incorporating large language model DeepSeek for TB prediction (LLMTBP), coupled with an integrated application designed to generate accurate TB simulation data for research applications. The Agent-based mechanism can effectively integrate the predictive values generated by LLMTBP and historical TB records, leveraging memory to uncover deeper temporal patterns in TB trends. This paper presents the TB-DeepSeek-Agent model, combining Agent-based forecasting with DeepSeek for accurate TB prediction, and an integrated web-based application for generating simulation data for research. The principal contributions of this study are:

- We have developed a TB-DeepSeek-Agent computational model that predicts TB infection data more precisely.
- An interactive website has been designed to integrate our model, allowing researchers to perform TB data analysis accurately and conveniently.

2 Method

2.1 Data

Using the dataset sourced from Jiangsu Provincial Center for Disease Control and Prevention, we collected daily TB infection cases from Changzhou, Suzhou, Nanjing, Zhenjiang and Wuxi within Jiangsu Province, spanning from 2011 to 2021. After excluding the missing data, there are totally 135802 TB cases. For the purpose of this study, the data has been divided into two main sets: the training dataset and the testing dataset. The training dataset consists TB infection data from the years 2011 to 2017, which is used to train our predictive model. The testing dataset contains the data from 2018 to 2021, which serves to evaluate the performance of our model.

2.2 TB-DeepSeek-Agent Architecture

As illustrated in Fig. 1. The conceptual architecture of TB-DeepSeek-Agent, our TB-DeepSeek-Agent model consist of three core components: Large Language Model for TB Prediction (LLMTBP), the memory and the reflection flow (For the other three Agent-based models, each incorporates its own original approach to TB prediction while

integrating memory and reflection flow within the same Agent framework as Fig. 1. The conceptual architecture of TB-DeepSeek-Agent). The TB data could be input into LLMTBP and undergoes preprocessing to ensure the dataset is cleaned where missing values are deleted. These preprocessing steps are essential for maintaining the integrity and reliability of the output generated by DeepSeek. It employs multi-scale spatiotemporal attention mechanisms to dynamically identify and prioritize critical TB features across diverse time horizons, thereby enhancing predictive decision-making [20, 21]. The following two systematically designed prompts were developed to generate summaries regarding TB-case distribution:

- Daily TB-case Summary: Records daily TB cases from 2018 to 2021, along with a daily summary.
- Monthly TB-case Summary: Captures monthly TB usage over the same period, summarized monthly.

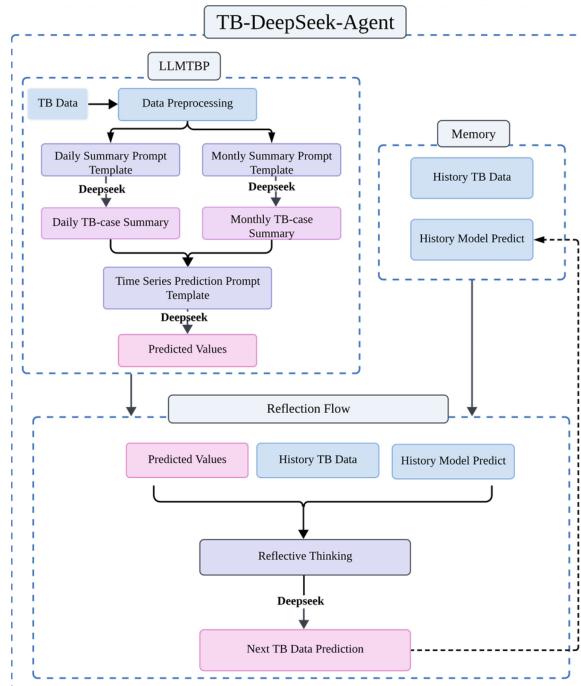


Fig. 1. The conceptual architecture of TB-DeepSeek-Agent

Equations 1 and 2 describe the methods for summarizing daily and monthly TB cases, respectively. (See parameter descriptions in Table 1)

$$RN_{c,td} = D\text{Summarize}(P_{c,td}) \quad (1)$$

$$RN_{c,tm} = M\text{Summarize}(P_{c,tm}) \quad (2)$$

These summaries are then combined for designed time intervals to provide further inputs based on the provided dataset to engineer the final time series prompts. After that, utilizing the time series prompt template, LLMTBP computes the predicted values defined by Eq. 3. (See parameter descriptions in Table 1)

$$PN_{c,ty} = \Delta Predict(RN_{c,td}, RN_{c,tm}) \quad (3)$$

At this stage, the previously outlined illustration represents the complete structure of the LLMTBP model. In contrast, our proposed TB-DeepSeek-Agent model introduces significant advancements by enhancing the initial predictions generated by LLMTBP. Specifically, in the TB-DeepSeek-Agent model, the prediction process begins by calling LLMTBP to generate an initial set of predicted values for the TB cases. While LLMTBP serves as a strong foundation for the initial prediction, the TB-DeepSeek-Agent model incorporates additional mechanisms that refine and improve these predictions by leveraging past information and reflective thinking of DeepSeek. The first key addition is the memory component, which stores both historical TB data and the TB-DeepSeek-Agent's own previous predictions. This memory mechanism is essential, as it allows the model to not only consider the current epidemiological trends but also to recall and learn from past predictions and data. By integrating this historical information, the model is able to capture long-term trends and evolving patterns in TB incidence, which is crucial for improving the accuracy and reliability of predictions over time. Once the memory component has been called by Agent, this information is fed into the reflection flow. The reflection mechanism dynamically adjusts predictions by analyzing historical prediction errors from the LLMTBP model. Specifically, it calculates the 95% confidence interval (95%CI) of these errors over a sliding time window. Based on the upper and lower bounds of this interval, the system automatically generates a dynamic correction factor. This correction is applied to the raw LLMTBP output, enabling the TB-DeepSeek-Agent model to compensate for systematic biases and temporal uncertainties, ultimately producing optimized and statistically robust predictions. By considering the broader epidemiological context, the reflective flow enhances the model's ability to account for subtle shifts in the data that might otherwise be overlooked. It allows the system to dynamically adjust its predictions based on patterns observed over time, facilitating a more responsive and context-aware approach to TB prediction. The strength of the DeepSeek reflection lies in its capacity to continuously refine the model's understanding of complex, nonlinear trends, ensuring that the predictions are not static but instead reflect the most up-to-date and nuanced epidemiological insights. (See parameter descriptions in Table 1)

$$DPN_{c,ty} = \Delta DPredict(M_{d,ty}, PN_{c,ty}) \quad (4)$$

Additionally, the final predicted TB cases, as described above, are simultaneously provided as responses to the memory for its update.

Through the integration of LLMTBP, the memory and the reflective thinking, TB-DeepSeek-Agent provides more refined control over the data, ensuring that predicted TB cases are accurate. This refinement improves the accuracy of TB case predictions compared to the original LLMTBP model.

Table 1. Definitions of Parameters Used in Mathematical Formulas

Parameter	Description
RN_{c,t_d} :	Real number of TB cases of city c on day t_d .
RN_{c,t_m} :	Real number of TB cases of city c in month t_m .
P_{c,t_d} :	An individual being diagnosed as a TB case in city c on day t_d .
P_{c,t_m} :	An individual being diagnosed as a TB case in city c on day t_m .
DSummarize():	Function utilizes the LLMTBP model to summarize daily TB cases.
MSummarize():	Function utilizes the LLMTBP model to summarize monthly TB cases.
PN_{c,t_y} :	Predicted number of TB cases of city c for the year t_y using the LLMTBP model.
$\Delta Predict()$:	The prediction operation using the LLMTBP model.
M_{d,t_y} :	List of history TB Data d and history model predict of year t_y in the memory.
$\Delta DPredict()$:	The prediction operation using the DeepSeek model after applying reflective thinking.
DPN_{c,t_y} :	TB predicted number of city c for the year t_y

3 Result

To assess the robustness and reliability of the TB-DeepSeek-Agent predictive model, we conducted a comparative analysis of the forecasting accuracy of eight algorithms for the number of TB infections in five cities from 2018 to 2021. In evaluating the performance of these TB infection forecasting algorithms, we utilized four key metrics: acceptance ratio, deviation ratio, and Root Mean Square Error (RMSE).

The comprehensive performance comparison of the TB prediction models for the five cities presented in Table 2. Forecast results for 5 cities in Jiangsu from 2018 to 2021 highlights significant variations among the algorithms. Among traditional models, LR demonstrates relatively low prediction accuracy, with an acceptance ratio of 30% and considerable deviation from actual values evidenced by a high deviation ratio (0.0822) and RMSE (208.3093). In contrast, the ARIMA model shows a notable performance advantage, with an acceptance ratio of 85% and a reduced RMSE (86.81) compared to LR.

Nevertheless, the real breakthrough in performance comes with the introduction of the Agent-based mechanism, which substantially improves model accuracy. Specifically, combining the Agent with LR dramatically improves its acceptance ratio from 30% to 85%, with the RMSE dropping from 208.3093 to 82.9696. Though the performance difference between ARIMA and ARIMA-Agent is not significant, the higher acceptance ratio (from 85% to 90%) illustrates the Agent-based model's ability to adapt quickly to new patterns. A more noticeable improvement is seen with LSTM when adding the Agent mechanism: the acceptance ratio increases from 65% to 100%, the deviation ratio drops from 0.0407 to 0.0157, and RMSE reduces from 100.7956 to 50.2668. This

demonstrates the effectiveness of integrating the Agent framework into deep learning models for accurate TB infection case prediction. Among all the models compared, TB-DeepSeek-Agent yielded the most accurate predictions, with the highest acceptance ratio (100%), the lowest deviation ratio (0.0129), and the lowest RMSE (30.5876).

Table 2. Forecast results for 5 cities in Jiangsu from 2018 to 2021

Algorithms	Acceptance Ratio	Deviation Ratio	RMSE
LR	30.00%(6/20)	0.0822	208.3093
LR-Agent	85.00%(17/20)	0.0277	82.9696
LSTM	65.00%(13/20)	0.0407	100.7956
LSTM-Agent	100.00%(20/20)	0.0157	50.2668
ARIMA	85.00%(17/20)	0.0271	86.8159
ARIMA-Agent	90.00%(18/20)	0.0205	75.3356
LLMTBP	35.00%(7/20)	0.0678	178.6214
TB-DeepSeek-Agent	100.00%(20/20)	0.0129	30.5876

Note: The four evaluation metrics used for model comparisons are explained as follows: Acceptance ratio is anchored in the requirement for predictions to maintain a 5% margin of error relative to true values, as this threshold is widely recognized as the standard for high-precision health surveillance analytics [22]; Deviation Ratio refers to the average absolute deviation of predicted values from the true values, normalized by the true values, which provides a measure of relative predictive accuracy; RMSE measures the magnitude of the prediction errors by calculating the square root of the average squared differences between the predicted and observed values.

We further examined the models' annual predictions for each city from 2018 to 2021 (Appendix Table 1). The TB-DeepSeek-Agent model consistently exhibited outstanding stability and accuracy, significantly surpassing the performance of all other models. For instance, it achieved perfect prediction accuracy in Zhenjiang (2020), with predicted and actual values both at 928 cases (ratio of predicted to true value: 1.0000). In Changzhou (2018), the model exhibited minimal deviation, predicting 1578 cases against the actual 1579 (ratio: 0.9994), demonstrating its ability to capture subtle epidemiological changes. In Nanjing (2021), the model predicted 2715 cases versus the actual 2712, showing only a 3-case difference (ratio: 1.0011), confirming its reliability in TB trend forecasting. The TB-DeepSeek-Agent model also demonstrated superior performance in complex epidemiological scenarios. In Changzhou (2019), it accurately predicted the upward trend, forecasting an increase from 1578 to 1610 cases (actual increase: 1579 to 1590, ratio: 1.0126), while other models failed to capture this trend. Similarly, in Nanjing's sustained growth period (2019–2020), TB-DeepSeek-Agent's predictions (2471 and 2522 cases) closely matched actual values (2485 and 2537), whereas other models miscalculated the trend. In Wuxi, where the downward trend persisted from 2018 to 2021, TB-DeepSeek-Agent consistently captured this reduction with minimal deviations (2255 to 2205 predicted versus 2205 to 2197 actual), outperforming ARIMA, which erroneously projected a 6.8% increase in 2021.

Finally, we developed an advanced web-based platform integrated with the TB-DeepSeek-Agent model for simulating TB data. This software is designed to transform TB research by providing a robust environment for comprehensive simulation experiments. In addition to automating TB simulations, the system generates detailed analytical reports, transforming complex datasets into actionable insights. By leveraging DeepSeek's advanced analytical capabilities, the platform produces well-structured reports that researchers can directly use, streamlining their workflow and enhancing the efficiency of their studies. This innovative tool offers researchers precision, depth, and convenience, making it an invaluable asset in the ongoing fight against TB (See Fig. 2).

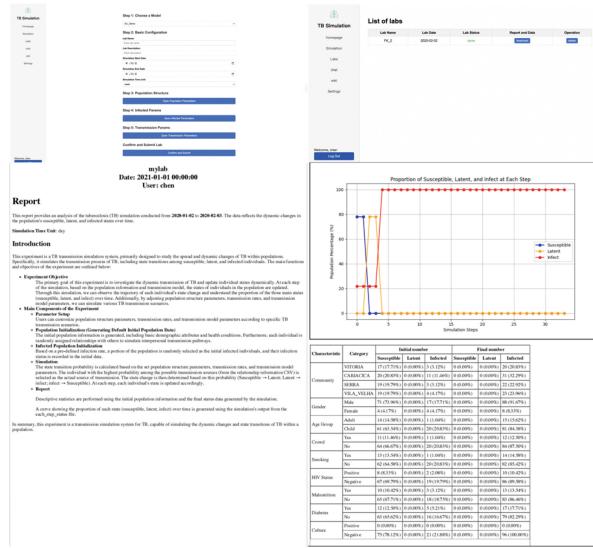


Fig. 2. The Website and Report of Our Application

4 Conclusion

This study demonstrates the efficacy of TB-DeepSeek-Agent, a novel framework integrating Agent architecture, reflective workflow with memory, and a large language model. Evaluated on TB infection records from five Jiangsu Province cities, the model significantly enhances prediction accuracy. Comparative analysis against seven baseline models revealed substantial error reduction, particularly outperforming the Large Language Model for TB Prediction (LLMTBP). The TB-DeepSeek-Agent architecture shows strong potential for improving TB forecasting, offering critical advantages for optimizing policy development and resource allocation in data collection and control programs. However, this study acknowledges limitations including potential data biases from missing value handling in historical records and unmeasured external confounding factors (e.g., healthcare disruptions during the COVID-19 pandemic). Future work

should explore its generalizability across diverse regions and extend its application to broader prediction tasks.

Appendix

Appendix Table 1. Forecasting results for 5 cities in Jiangsu from 2018 to 2021

		2018	2019	2020	2021
Changzhou	True	1579	1590	1350	1241
	LR	1724(1.0918)	1762(1.1082)	1539(1.1400)	1365(1.0999)
	LR-Agent	1589(1.0063)	1652(1.0390)	1429(1.0585)	1240(0.9992)
	LSTM	1706(1.0804)	1731(1.0887)	1395(1.0333)	1319(1.0629)
	LSTM-Agent	1597(1.0114)	1586(0.9975)	1402(1.0385)	1222(0.9847)
	ARIMA	1489(0.9430)	1612(1.0138)	1366(1.0119)	1216(0.9799)
	ARIMA-Agent	1608(1.0184)	1598(1.0050)	1363(1.0096)	1262(1.0169)
	LLMTBP	1664(1.0538)	1566(0.9849)	1474(1.0919)	1336(1.0766)
	TB-DeepSeek-Agent	1578(0.9994)	1610(1.0126)	1405(1.0407)	1210(0.9750)
Nanjing	True	2480	2485	2537	2712
	LR	2758(1.1121)	2715(1.0926)	2471(0.9740)	2601(0.9591)
	LR-Agent	2584(1.0419)	2523(1.0153)	2453(0.9669)	2875(1.0601)
	LSTM	2519(1.0157)	2544(1.0237)	2443(0.9629)	2767(1.0203)
	LSTM-Agent	2536(1.0226)	2466(0.9924)	2498(0.9846)	2761(1.0181)
	ARIMA	2466(0.9944)	2378(0.9569)	2454(0.9673)	2874(1.0597)
	ARIMA-Agent	2520(1.0161)	2475(0.9960)	2604(1.0264)	2903(1.0704)
	LLMTBP	2290(0.9234)	2243(0.9026)	2245(0.8849)	2339(0.8625)
	TB-DeepSeek-Agent	2517(1.0149)	2471(0.9944)	2522(0.9941)	2715(1.0011)
Suzhou	True	4095	3597	3120	3292
	LR	4283(1.0459)	4028(1.1198)	3479(1.1151)	3185(0.9675)
	LR-Agent	4220(1.0305)	3528(0.9808)	3203(1.0266)	3456(1.0498)
	LSTM	4337(1.0591)	3734(1.0381)	3210(1.0288)	3238(0.9836)
	LSTM-Agent	4183(1.0215)	3498(0.9725)	3218(1.0314)	3287(0.9985)
	ARIMA	3967(0.9687)	3388(0.9419)	3175(1.0176)	3254(0.9885)
	ARIMA-Agent	4297(1.0493)	3632(1.0097)	3117(0.9990)	3288(0.9988)
	LLMTBP	3891(0.9502)	3857(1.0723)	3448(1.1051)	3144(0.9550)
	TB-DeepSeek-Agent	4141(1.0112)	3586(0.9969)	3125(1.0016)	3237(0.9833)
Wuxi	True	2411	2396	2205	2197
	LR	2719(1.1277)	2566(1.0710)	2447(1.1098)	2450(1.1152)
	LR-Agent	2472(1.0253)	2378(0.9925)	2340(1.0612)	2252(1.0250)

(continued)

Appendix Table 1. (continued)

		2018	2019	2020	2021
	LSTM	2446(1.0145)	2368(0.9883)	2383(1.0807)	2300(1.0469)
	LSTM-Agent	2392(0.9921)	2385(0.9954)	2309(1.0472)	2210(1.0059)
	ARIMA	2414(1.0012)	2330(0.9725)	2229(1.0109)	2356(1.0724)
	ARIMA-Agent	2453(1.0174)	2351(0.9812)	2350(1.0658)	2212(1.0068)
	LLMTBP	2322(0.9631)	2333(0.9737)	2279(1.0336)	2230(1.0150)
	TB-DeepSeek-Agent	2377(0.9859)	2366(0.9875)	2255(1.0227)	2205(1.0036)
Zhenjiang	True	1234	1179	928	949
	LR	1207(0.9781)	1250(1.0602)	973(1.0485)	1010(1.0643)
	LR-Agent	1279(1.0365)	1181(1.0017)	936(1.0086)	956(1.0074)
	LSTM	1250(1.0130)	1258(1.0670)	980(1.0560)	968(1.0200)
	LSTM-Agent	1266(1.0259)	1185(1.0051)	929(1.0011)	946(0.9968)
	ARIMA	1214(0.9838)	1150(0.9754)	932(1.0043)	927(0.9768)
	ARIMA-Agent	1257(1.0186)	1216(1.0314)	931(1.0032)	967(1.0190)
	LLMTBP	1122(0.9092)	1089(0.9237)	981(1.0571)	870(0.9168)
	TB-DeepSeek-Agent	1227(0.9943)	1226(1.0399)	928(1.0000)	930(0.9800)

References

1. World Health Organization: Global Tuberculosis Report 2024. World Health Organization, Geneva (2024)
2. Dheda, K., Perumal, T., Moultrie, H., Perumal, R., Esmail, A., Scott, A.J., et al.: The intersecting pandemics of tuberculosis and COVID-19: population-level and patient-level impact, clinical presentation, and corrective interventions. *Lancet Respir. Med.* **10**(6), 603–622 (2022)
3. World Health Organization: Tuberculosis response recovering from pandemic but accelerated efforts needed to meet new targets. <https://www.who.int/news-room/detail/07-11-2023-tuberculosis-response-recovering-from-pandemic-but%2D%2Daccelerated-efforts-needed-to-meet-new-targets>, last accessed 2025/02/07
4. Kusimo, O.C., Ugwu, C.I., Aduh, U., Okoro, C.A.: Implementing TB surveillance in Nigeria: best practices, challenges and lessons learnt. *J. Tuberc. Res.* **8**(4), 199–208 (2020)
5. Puyat, J.H., Brode, S.K., Shulha, H., Romanowski, K., Menzies, D., Benedetti, A., et al.: Predicting risk of tuberculosis (TB) disease in people who migrate to a low-TB incidence country: development and validation of a multivariable, dynamic risk-prediction model using health administrative data. *Clin. Infect. Dis.* **ciae561** (2024)
6. Engel, N., Ochodo, E.A., Karanja, P.W., Schmidt, B.M., Janssen, R., Steingart, K.R., et al.: Rapid molecular tests for tuberculosis and tuberculosis drug resistance: a qualitative evidence synthesis of recipient and provider views. *Cochrane Database Syst. Rev.* **4** (2022)
7. Gabrielian, A., Engle, E., Harris, M., Wollenberg, K., Juarez-Espinosa, O., Glogowski, A., et al.: TB DEPOT (Data Exploration Portal): A multi-domain tuberculosis data analysis resource. *PLoS One.* **14**(5), e0217410 (2019)

8. Cole, B.: Essential components of a public health tuberculosis prevention, control, and elimination program: recommendations of the Advisory Council for the Elimination of Tuberculosis and the National Tuberculosis Controllers Association. *MMWR Recomm. Rep.* **69** (2020)
9. Sarkar, M.: Tuberculosis infection prevention and control. *Indian J. Tuberc.* (2024)
10. Gharamaleki, O.G., Colijn, C., Sekirov, I., Johnston, J.C., Sobkowiak, B.: Early prediction of Mycobacterium tuberculosis transmission clusters using supervised learning models. *Sci. Rep.* **14**(1), 27652 (2024)
11. Siawaya, J.F.D., Bapela, N.B., Ronacher, K., Veenstra, H., Kidd, M., Gie, R., et al.: Immune parameters as markers of tuberculosis extent of disease and early prediction of anti-tuberculosis chemotherapy response. *J. Infect.* **56**(5), 340–347 (2008)
12. Maipan-Uku, J.Y., Cavus, N.: Forecasting tuberculosis incidence: a review of time series and machine learning models for prediction and eradication strategies. *Int. J. Environ. Health Res.*, 1–16 (2024)
13. Aryee, G., Kwarteng, E., Essuman, R., Agyei, A.N., Kudzawu, S., Djagbletey, R., et al.: Estimating the incidence of tuberculosis cases reported at a tertiary hospital in Ghana: a time series model approach. *BMC Public Health.* **18**, 1–8 (2018)
14. Liao, Z., Zhang, X., Zhang, Y., Peng, D.: Seasonality and trend forecasting of tuberculosis incidence in Chongqing, China. *Interdiscip. Sci. Comput. Life Sci.* **11**, 77–85 (2019)
15. Frauenfeld, L., Nann, D., Sulyok, Z., Feng, Y.S., Sulyok, M.: Forecasting tuberculosis using diabetes-related Google Trends data. *Pathog. Global Health.* **114**(5), 236–241 (2020)
16. Abdualgalil, B., Abraham, S., Ismael, W.M., George, D.: Modeling and forecasting tuberculosis cases using machine learning and deep learning approaches: a comparative study. In: International Conference on Data Management, Analytics & Innovation, pp. 157–171. Springer, Singapore (2022)
17. Hamna Mariyam, K.B., Jose, S.A., Jirawattanapanit, A., Mathew, K.: A comprehensive study on tuberculosis prediction models: integrating machine learning into epidemiological analysis. *J. Theor. Biol.* **597**, 111988 (2025)
18. Al Sailawi, A.S.A., Kangavari, M.R.: Utilizing AI for extracting insights on post WHO's COVID-19 vaccination declaration from X (Twitter) social network. *AIMS Public Health.* **11**(2), 349–378 (2024)
19. Awasthi, R., Nagori, A., Nasri, B.: Global Generalisability of AI-Driven COVID-19 Vaccination Policies: A Cross-Sectional Observational Study. *medRxiv* 2023.01 (2023)
20. Guo, D., et al.: Deepseek-r1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. *arXiv preprint arXiv:2501.12948* (2025)
21. Liu, A., et al.: Deepseek-v2: A Strong, Economical, and Efficient Mixture-of-Experts Language Model. *arXiv preprint arXiv:2405.04434* (2024).
22. Wang, Y., et al.: Temporal trends analysis of tuberculosis morbidity in mainland China from 1997 to 2025 using a new SARIMA-NARNNX hybrid model. *BMJ Open.* **9**(7), e024409 (2019)



Bayesian-Optimized Eye Movement Event Detection for Assisted Alzheimer's Diagnosis

Xin Wang^{1,2}, Yujie Nie^{1,2}, Jia Zhao^{1,2}, Yahan Wang³, Yibin Li^{1,2},
Xiangqing Xu³, and Xin Ma^{1,2(✉)}

¹ School of Control Science and Engineering, Shandong University, Jinan 250061, China

² Engineering Research Center of Intelligent Unmanned System, Ministry of Education, Beijing, China
maxin@sdu.edu.cn

³ Department of Neurology, Affiliated Hospital of Shandong University of Traditional Chinese Medicine, Jinan 16369, China

Abstract. Alzheimer's disease (AD) is a progressive neurodegenerative disorder requiring reliable early diagnostic tools. We propose PSOs-Net, a Bayesian-optimized 1DCNN-BLSTM model for automated eye movement detection, alongside the novel AD-VSEMD dataset capturing visual search behaviors in AD patients and controls. The model achieves state-of-the-art performance, with event-level Kappa scores of 0.972 (fixations), 0.968 (saccades), and 0.786 (post-saccadic oscillations) on Lund2013, surpassing existing methods. Results show AD-specific impairments, including prolonged fixations, increased saccadic activity, and altered search patterns linked to cognitive decline. This demonstrates that automated eye movement analysis offers a sensitive, non-invasive AD assessment, with our Bayesian-optimized framework enhancing neurodegenerative disease detection and progression monitoring through eye-tracking and machine learning.

Keywords: Alzheimer's disease · Eye movement event detection · Bayesian optimization · Deep learning

1 Introduction

Neurodegenerative diseases represent a critical global health challenge amid rapid population aging, with Alzheimer's disease (AD) accounting for 60–70% of dementia cases worldwide. AD presents as progressive cognitive impairment (CI), behavioral abnormalities, and functional decline [1]. WHO projects 152 million CI cases by 2050, creating substantial socioeconomic burdens for health-care systems.

The disease involves progressive cognitive decline, particularly in visual search tasks, due to degeneration of neural circuits governing visual attention

and spatial cognition. Core pathology includes abnormal β -amyloid deposition and tau aggregation, causing synaptic dysfunction [2,3]. These changes originate in medial temporal lobe structures (hippocampus/entorhinal cortex), then spread to parietal and frontal visual-attention regions [4]. Posterior parietal damage (inferior/superior lobules) especially impairs target identification in complex scenes [5].

Eye movement analysis has emerged as a promising non-invasive biomarker for neurodegenerative diseases. Studies using classic paradigms (fixation, smooth pursuit, saccades) reveal distinct abnormalities in AD patients, including increased fixation counts and prolonged durations during visual tasks, which correlate with cognitive decline [6,7]. These behavioral patterns reflect impairments in visual processing pathways, particularly in attention and executive function.

Current eye movement detection relies on threshold-based algorithms requiring subjective manual tuning. To address this, GazeNet [8] introduced a hybrid CNN-GRU architecture for end-to-end classification, later enhanced with attention mechanisms [9] and U-Net structures [10] for multi-scale feature extraction. However, these 2D CNN-based approaches may overemphasize spatial features at the expense of temporal dependencies in eye-tracking data, while GRUs underperform LSTMs in capturing long-range temporal patterns - potentially limiting dynamic movement detection.

To overcome these limitations, we propose a novel AD diagnostic framework using optimized eye movement analysis. First, we built AD-VSEMD, a specialized dataset capturing AD visual search behaviors. We then developed a 1DCNN-BLSTM model, optimized via Bayesian hyperparameter tuning, and trained/tested it on public benchmarks for event detection. The model was applied to AD-VSEMD for automated labeling, ensuring consistency with established methods. Finally, we extracted discriminative features from annotated events and used statistical analysis to identify AD-specific abnormalities, aiming to create a reliable diagnostic tool.

2 Methods

2.1 Data Collection and Preprocessing

Unlike naturalistic visual search, our controlled laboratory paradigm (Fig. 1) used simplified arrow arrays to study AD-related eye movements. Participants searched for a uniquely oriented target arrow among distractors in two matrices: Grid-4 (4×4) and Grid-6 (6×6), with target direction (left/right/up/down) and position randomized uniformly across 8 repetitions per condition. Rest intervals between Grid-4/Grid-6 blocks minimized fatigue.

Visual stimuli were displayed on a 23.8" monitor (1920×1080) driven by an ROG gaming laptop (Intel i9-14900HX, NVIDIA RTX 4090). Eye movements were recorded at 250 Hz using a Tobii Pro Fusion 250 eye tracker, with five-point calibration performed for each participant to ensure data quality. The study protocol was approved by the Ethics Committee of Shandong University

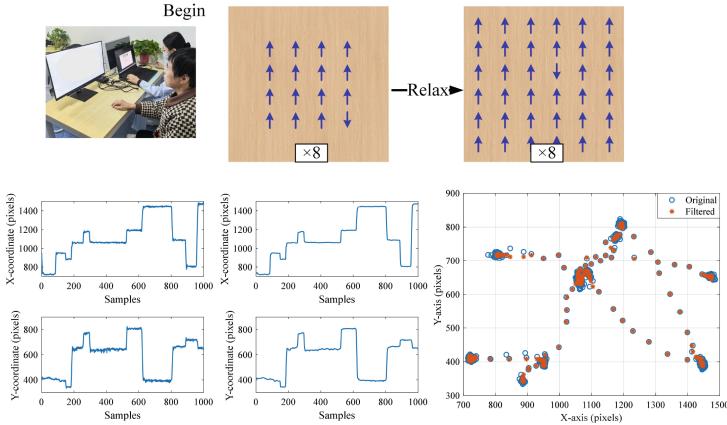


Fig. 1. Data collection process and examples.

of Traditional Chinese Medicine Affiliated Hospital. All participants provided written informed consent following full disclosure of procedures, with strict confidentiality maintained for all personal data.

The study was conducted June–December 2024 at Shandong University of Traditional Chinese Medicine Affiliated Hospital (East Campus). Participants were excluded if < 50 years old or with psychological/vascular disorders. Four groups were recruited: healthy young controls (YHC, $n=39$, MoCA ≥ 26), healthy elderly controls (HC, $n=38$, MoCA ≥ 26), mild cognitive impairment (MCI, $n = 20$, $20 \leq \text{MoCA} \leq 25$), and AD patients ($n=22$, MoCA ≤ 19). The mean ages of the MCI&AD, HC, and YHC groups were 65, 60, and 26 years, respectively, with male-to-female ratios of 15:27, 13:26, and 25:13.

Raw eye movement data were preprocessed to ensure quality, excluding trials with $> 50\%$ missing data or fewer than 101 sample points. Missing values were linearly interpolated, followed by noise reduction using a moving median filter (window size = 11, ~ 44 ms at 250 Hz). The filtering effects on waveform characteristics are illustrated in Fig. 1.

2.2 Bayesian Optimized Eye Movement Event Detection Model

Figure 2 shows the architecture of our PSOs-Net algorithm, combining a 1DCNN-BLSTM hybrid model (PyTorch implementation) with Bayesian optimization for parallel detection of three eye movement events: fixations, saccades, and post-saccadic oscillations (PSOs). The figure details the hyperparameter configurations and their search space boundaries. We performed 50 optimization iterations using Gaussian process-based Bayesian optimization (via Scikit-Optimize) to identify the optimal hyperparameter combination.

The model architecture comprises: A 1D CNN module with multiple convolutional layers sharing identical kernel sizes but varying neuron counts per

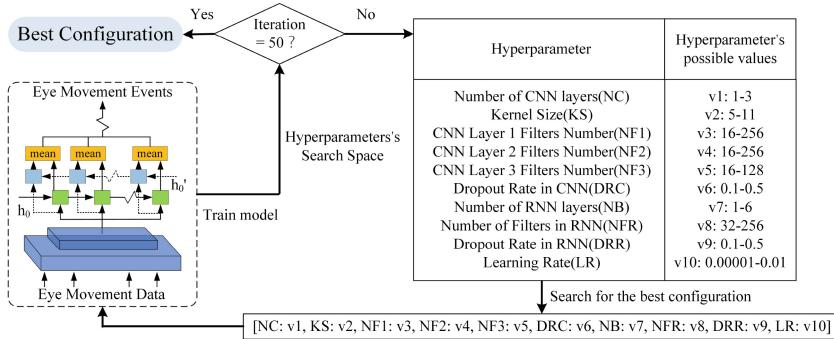


Fig. 2. Overall structure of PSOs-Net.

layer, enabling effective local feature extraction; A BLSTM module with uniform neuron counts across layers to capture temporal dependencies, employing mean aggregation; A final fully connected layer that generates the eye movement event detections.

A composite objective function was designed based on dual-scale metrics:

$$f(x) = \min_{x \in X} \|\kappa e, \kappa s\|_2 \quad (1)$$

The optimization framework uses κe (event-level Kappa) and κs (sample-level Kappa) as dual objectives within hyperparameter space X . The L2-norm formulation ensures balanced improvement of both metrics, preventing unilateral dominance while requiring synchronous convergence for global optimization.

2.3 Extraction of Eye Movement Features from Patients

Eye movement events in the AD-VSEMD dataset were automatically annotated using PSOs-Net. The annotated data underwent preprocessing: Saccade events with amplitudes 1° and subsequent PSOs events (if present) were reclassified as fixations to merge adjacent fixations fragmented by minor saccade errors; PSOs events not preceded by saccade events were relabeled as fixation events to rectify algorithm-induced misannotations; Fixation events lasting 60 ms were excluded to ensure data robustness.

Table 1 summarizes the extracted features, including four behavioral and nine eye movement metrics. Data were collected using two arrow-matrix configurations (Grid-4 and Grid-6), each with eight randomized target trials. For consistency, trials within each configuration were aggregated via averaging (not summation) to derive composite metrics.

3 Experimental Setup and Results Analysis

We trained and tested PSOs-Net on the Lund2013 dataset (500 Hz) [12], employing early stopping (validation every 25 steps, 150-check patience, ~ 3.5 epochs)

Table 1. Disease related quantitative indicators and their definitions.

Category	Metrics	Definition
Behavior	<i>Acc</i>	Percentage of successful trials relative to total attempts
	<i>N_{roi}</i>	Mean number of gaze entries into areas of interest per trial
	<i>T</i>	Mean response latency per completed trial
	<i>Eff</i>	Accuracy to mean reaction time ratio
Eye Movement	<i>P</i>	Mean pupil size across fixations per trial
	<i>N_{sac}</i>	Average number of saccades per trial
	<i>V</i>	Mean angular speed of saccades (degrees/sec)
	<i>N_{pso}</i>	Mean number of PSOs per trial
	<i>F_{pso}</i>	Frequency of PSOs following saccadic movements
	<i>T_{fix}</i>	Mean duration of fixations per trial
	<i>FR</i>	Percentage of trial time spent in fixation
	<i>A_{search}</i>	Mean scanned area relative to total display area
	<i>K</i>	Standardized difference between saccade amplitude and fixation duration scores [11]

to prevent overfitting. Due to the small validation set (~ 1 min) and overfitting risks [8], we optimized using the mean of validation and test set objective functions. The model used weighted cross-entropy loss (fixations=0.1443, saccades=0.8955, PSOs=0.9602) with RMSprop (batch=100, epochs=50). Inputs were processed in 100-sample windows, augmented via Gaussian noise (0–0.4°) and channel swapping. Two labeling errors in Lund2013 were corrected: UH27_img_vy_labelled_RA.mat (37 samples, 8676–8748 ms) smooth pursuit → fixation, and UH29_img_Europe_labelled_RA.mat (5 samples, 2868–2876 ms) pre-saccadic PSOs → fixation.

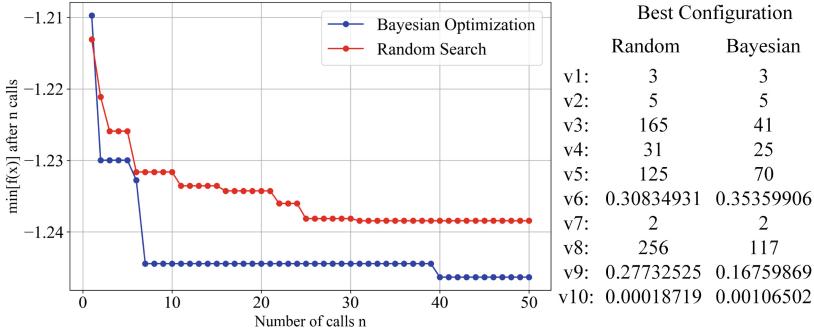
Eye movement events exhibit spatiotemporal invariance, so our feature extraction focuses on relative kinematic measures: velocity (v_x, v_y), acceleration (a_x, a_y), their norms ($\|\mathbf{v}\|, \|\mathbf{a}\|$), and binocular disparity (d_x, d_y). These form an 8D feature vector \mathbf{x} , capturing essential motion dynamics. For cross-dataset validation, we developed Equivalent Feature Scale Mapping (EFSM) to handle 250 Hz GazeCom data without resampling. EFSM scales coordinate offsets by 0.5 to match 500 Hz feature scales while preserving original signal characteristics.

To address class imbalance, we use Zemblys et al.’s revised κ metric [13], computed as $\kappa = (p_o - p_e)/(1 - p_e)$. This penalizes prediction bias through p_e (chance agreement), ensuring rigorous evaluation under imbalance by weighting performance according to class prevalence.

3.1 Comparisons with State-of-the-Art Methods

Figure 3 compares Bayesian optimization (BO) and random search convergence during hyperparameter tuning. BO rapidly approaches near-optimal performance (−1.414) early, reaching the best value (−1.246) by iteration 40, while random search converges slower to a suboptimal value (−1.238). This demonstrates BO’s superior efficiency in high-dimensional hyperparameter space exploration.

Both methods converge to similar architectures: three convolutional layers (kernel size 5) and two recurrent layers, with a U-shaped neuron distribution

**Fig. 3.** Hyperparameter optimization curve.

across convolutional layers. This design may benefit PSOs detection – deeper initial/final layers enhance spatiotemporal feature extraction, while the narrower middle layer filters information to prevent overfitting. The results demonstrate the 1DCNN-BLSTM architecture’s adaptability for eye movement analysis.

Table 2. Experimental results on the Lund2013 test dataset.

Algorithm	Event level			Sample level		
	FIX	SAC	PSOs	FIX	SAC	PSOs
IRF [14]	0.893	0.885	0.607	0.865	0.839	0.555
gazeNet [8]	0.934	0.934	0.744	0.873	0.898	0.699
GazeUNet [10]	0.962	0.954	0.775	0.892	0.908	0.727
PSOs-Net (random search)	0.962	0.961	0.770	0.895	0.912	0.742
Bayesian-optimized PSOs-Net	0.972	0.968	0.786	0.898	0.913	0.745

Table 2 shows PSOs-Net’s performance on the Lund2013 test set. It outperforms all methods, achieving event-level κ of 0.972 (fixations), 0.968 (saccades), and 0.786 (PSOs), with sample-level κ of 0.898, 0.913, and 0.745. Compared to GazeUNet, PSOs-Net improves PSO detection by 1.1% (event-level) and 1.8% (sample-level). Table 3 evaluates model generalization on the 250 Hz GazeCom dataset (no PSOs). PSOs-Net shows strong generalization, with EFSM processing boosting performance without resampling. While its sample-level κ for fixations is slightly lower than the random-search variant, the Bayesian-optimized PSOs-Net outperforms in all other metrics, confirming superior adaptability.

3.2 Abnormal Visual Search Behavior in AD Patients

The Mann-Whitney U test ($\alpha = 0.05$) assessed group differences in eye movement features, robust to non-normal distributions and small samples. While

Table 3. Experimental results on the GazeCom dataset.

Algorithm	Event level		Sample level	
	FIX	SAC	FIX	SAC
IRF [14]	0.796	0.900	0.486	0.776
gazeNet [8]	0.919	0.892	0.853	0.795
GazeUNet [10]	0.911	0.895	0.861	0.814
PSOs-Net (random search)	0.903	0.898	0.864	0.812
Bayesian-optimized PSOs-Net	0.915	0.900	0.863	0.803
Bayesian-optimized PSOs-Net (EFSM)	0.947	0.940	0.847	0.841

Table 4. Comparison of eye movement metrics across different groups.

Task	Metrics	YHC vs. HC	HC vs. MCI	MCI vs. AD	HC vs. MCI&AD
Grid-4	<i>Acc</i>	0.58	0.44	0.21	0.071
	<i>N_{roi}</i>	2.9e-08	6.0e-03	0.19	2.7e-06
	<i>T</i>	3.0e-13	9.0e-05	0.20	1.9e-07
	<i>Eff</i>	4.7e-13	8.5e-05	0.20	2.5e-07
	<i>P</i>	3.0e-04	0.021	0.25	0.020
	<i>N_{sac}</i>	2.5e-10	3.0e-04	0.026	1.8e-07
	<i>V</i>	0.098	0.82	0.49	0.84
	<i>N_{pso}</i>	6.5e-08	5.0e-04	0.043	1.55e-07
	<i>FR</i>	0.022	0.79	0.20	0.34
	<i>T_{fix}</i>	4.5e-11	7.4e-05	0.057	2.4e-08
	<i>FP</i>	0.61	0.99	0.59	0.72
	<i>A_{search}</i>	1.6e-09	5.7e-03	0.55	4.2e-05
	<i>K</i>	1.1e-05	0.96	0.085	0.36
Grid-6	<i>Acc</i>	0.11	0.24	0.72	0.26
	<i>N_{roi}</i>	7.0e-06	0.070	0.16	7.0e-04
	<i>T</i>	4.4e-12	7.1e-03	0.10	4.4e-05
	<i>Eff</i>	3.2e-12	5.0e-03	0.20	4.4e-05
	<i>P</i>	1.6e-03	0.025	0.29	0.032
	<i>N_{sac}</i>	4.2e-10	0.035	0.07	3.0e-04
	<i>V</i>	0.70	0.48	0.95	0.44
	<i>N_{pso}</i>	7.9e-08	0.047	0.034	2.0e-04
	<i>FR</i>	0.010	0.85	0.11	0.23
	<i>T_{fix}</i>	1.1e-10	3.5e-03	0.16	3.1e-05
	<i>FP</i>	0.076	0.34	0.72	0.42
	<i>A_{search}</i>	1.1e-11	0.19	0.20	9.2e-03
	<i>K</i>	3.5e-06	0.018	0.52	2.0e-03

behavioral metrics (Acc , N_{roi} , T , Eff) showed no significant differences between MCI and AD ($p > 0.05$), significant differences emerged in N_{roi} , T , and Eff when comparing: YHC vs HC, HC vs MCI, and HC vs MCI&AD ($p < 0.05$). Participants achieved high Acc in Grid-4, confirming appropriate task difficulty. YHC typically found targets in two scans versus HC's four, suggesting age-related visual neglect risk. AD patients showed significantly higher rescanning than HC ($p < 0.01$), indicating attentional and working memory deficits.

Eye movement features enhanced discrimination, with N_{sac} ($p = 0.026$) and N_{pso} ($p = 0.043$) differing significantly between MCI and AD. Pupil diameter progressively declined from YHC to HC to AD ($p < 0.05$), indicating cognitive load and autonomic dysfunction. AD patients showed: Prolonged T_{fix} , Increased N_{sac} and N_{pso} and Expanded A_{search} ($p < 0.05$), reflecting inefficient visual processing and attentional dysregulation. Notably, N_{sac} and N_{pso} remained significant between MCI and AD despite comparable reaction times ($p > 0.05$), revealing disease-specific saccadic impairment. In contrast, FR lost significance, suggesting T_{fix} differences may reflect generalized slowing rather than disease-specific effects.

Task difficulty modulated effects: in Grid-6, previously significant differences in N_{roi} , A_{search} (HC vs. MCI), and N_{sac} (MCI vs. AD) disappeared, likely due to training effects from Grid-4. However, feature K newly distinguished HC from both MCI and MCI&AD ($p < 0.05$), indicating a shift toward prolonged fixations and short saccades – a maladaptive strategy that expanded search area while reducing efficiency.

4 Conclusion and Discussion

This study introduces PSOs-Net, a novel Bayesian-optimized framework for eye movement event detection to aid Alzheimer's disease (AD) diagnosis. Combining a 1DCNN-BLSTM architecture with Bayesian hyperparameter tuning, our model outperforms state-of-the-art methods in detecting fixations, saccades, and post-saccadic oscillations (PSOs), achieving high event-level and sample-level Kappa scores. The proposed Equivalent Feature Scale Mapping (EFSM) further enhances generalization across datasets with varying sampling frequencies. Applied to the AD-VSEMD dataset, PSOs-Net identifies AD-specific visual processing abnormalities—prolonged fixations, increased saccade counts, and altered search strategies—correlating with cognitive decline and attentional deficits.

Despite these advancements, several limitations should be acknowledged. While PSOs-Net achieves high accuracy in detecting fixations and saccades, its performance for PSOs remains comparatively lower, suggesting the need for architectural refinements or improved feature representation. Additionally, the dataset—though carefully curated—is limited in size, particularly for MCI and AD groups; expanding its diversity and scale would enhance robustness and enable finer subgroup analyses. Furthermore, while the controlled laboratory paradigm ensures standardization, it may not fully replicate naturalistic visual search behaviors, potentially limiting real-world generalizability.

Table 5. Median (interquartile range) of eye movement metrics across groups.

Metric	YHC	HC	MCI	AD
Grid_4_Acc	1.0 (1.0, 1.0)	1.0 (1.0, 1.0)	1.0 (0.94, 1.0)	1.0 (0.75, 1.0)
N_{roi}	1.6 (1.3, 2.1)	3.0 (2.4, 4.3)	4.6 (3.4, 8.8)	7.0 (5.8, 8.5)
T	1.6 (1.4, 1.8)	3.4 (2.8, 4.4)	5.6 (4.3, 8.2)	7.9 (4.3, 13)
Eff	0.61 (0.54, 0.72)	0.28 (0.23, 0.36)	0.16 (0.12, 0.23)	0.13 (0.057, 0.20)
P	3.1 (2.9, 3.3)	2.8 (2.5, 3.0)	2.4 (2.2, 2.8)	2.5 (2.3, 3.0)
N_{sac}	4.6 (3.8, 5.4)	9.5 (6.6, 12)	18 (12, 22)	25 (14, 33)
V	1.0e2 (98, 1.2e2)	1.1e2 (1.0e2, 1.3e2)	1.2e2 (93, 1.3e2)	1.2e2 (1.0e2, 1.3e2)
N_{pso}	1.9 (1.4, 2.4)	3.5 (2.5, 5.0)	7.8 (4.4, 9.1)	9.3 (5.4, 16)
FR	0.44 (0.39, 0.49)	0.39 (0.34, 0.45)	0.41 (0.36, 0.43)	0.42 (0.37, 0.47)
T_{fix}	1.4 (1.2, 1.7)	2.6 (2.0, 3.8)	5.3 (3.7, 6.6)	6.7 (4.1, 13.2)
FR	0.87 (0.86, 0.88)	0.87 (0.84, 0.89)	0.87 (0.85, 0.89)	0.88 (0.85, 0.89)
A_{search}	0.029 (0.021, 0.037)	0.073 (0.052, 0.093)	0.13 (0.072, 0.17)	0.13 (0.085, 0.20)
K	-0.17 (-0.23, -0.076)	-0.024 (-0.082, 0.034)	-0.0056 (-0.076, 0.016)	0.015 (-0.053, 0.048)
Grid_6_Acc	1.0 (1.0, 1.0)	1.0 (1.0, 1.0)	1.0 (0.94, 1.0)	1.0 (1.0, 1.0)
N_{roi}	2.0 (1.5, 2.4)	2.9 (2.1, 4.1)	4.0 (2.6, 7.4)	5.5 (3.7, 7.2)
T	1.8 (1.5, 2.3)	4.2 (3.6, 5.1)	5.5 (4.5, 7.2)	6.9 (5.4, 18)
Eff	0.55 (0.44, 0.65)	0.23 (0.18, 0.28)	0.17 (0.11, 0.22)	0.14 (0.05, 0.18)
P	3.1 (2.8, 3.3)	2.8 (2.5, 3.0)	2.4 (2.2, 2.9)	2.5 (2.4, 3.0)
N_{sac}	5.4 (4.3, 6.8)	13 (11, 16)	17 (12, 27)	24 (17, 41)
V	1.3e2 (1.1e2, 1.4e2)	1.3e2 (1.1e2, 1.5e2)	1.2e2 (1.1e2, 1.5e2)	1.3e2 (1.1e2, 1.5e2)
N_{pso}	2.3 (1.7, 2.9)	4.9 (3.8, 6.5)	6.3 (4.6, 12)	9.6 (7.2, 18)
FR	0.43 (0.38, 0.51)	0.39 (0.36, 0.42)	0.38 (0.35, 0.45)	0.43 (0.38, 0.50)
T_{fix}	1.5 (1.3, 2.0)	3.3 (2.3, 4.0)	4.4 (3.4, 6.8)	6.4 (3.9, 14)
FR	0.86 (0.84, 0.88)	0.85 (0.82, 0.87)	0.86 (0.84, 0.87)	0.85 (0.82, 0.88)
A_{search}	0.067 (0.051, 0.091)	0.19 (0.14, 0.26)	0.23 (0.16, 0.34)	0.28 (0.21, 0.36)
K	-0.18 (-0.24, -0.10)	-0.040 (-0.080, 0.010)	0.0067 (-0.020, 0.030)	0.023 (-0.024, 0.037)

These findings highlight the promise of eye movement analysis as a non-invasive AD biomarker, especially when enhanced by machine learning. Key features like saccade count and fixation patterns effectively differentiate healthy aging, MCI, and AD, demonstrating strong discriminative power. Future studies should adopt longitudinal designs to evaluate their prognostic value in tracking disease progression, while multimodal integration (e.g., neuroimaging, genetics) could reveal deeper neural mechanisms. This work advances eye-tracking's potential for objective AD diagnosis, paving the way for data-driven diagnostic tools in neurodegenerative disease management.

Acknowledgement. This work was jointly supported by National Key Research and Development Program Project under Grant 2023YFB4706104. Key R&D Program of Shandong Province (Major Science and Technology Innovation Project) under Grant 2024CXGC010603. Fundamental Research Funds for the Central Universities under Grant 2022JC011.

References

1. Dubois, B., Villain, N., Frisoni, G.B., et al.: Clinical diagnosis of Alzheimer's disease: recommendations of the International Working Group. *Lancet Neurol.* **20**(6), 484–496 (2021)
2. Zhang, J., Zhang, Y., Wang, J., et al.: Recent advances in Alzheimer's disease: mechanisms, clinical trials and new drug development strategies. *Signal Transduct. Target. Ther.* **9**(1), 211 (2024)
3. Braak, H., Braak, E.: Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **82**(4), 239–259 (1991)
4. Migliaccio, R., Cacciamani, F.: The temporal lobe in typical and atypical Alzheimer disease. *Handb. Clin. Neurol.* **187**, 449–466 (2022)
5. Husain, M., Nachev, P.: Space and the parietal cortex. *Trends Cogn. Sci.* **11**(1), 30–36 (2007)
6. Akkoyun, M., Koçoğlu, K., Eraslan Boz, H., et al.: Visual search for real-world scenes in patients with Alzheimer's disease and amnestic mild cognitive impairment. *Brain Behav.* **14**(6), e3567 (2024)
7. Li, D., Butala, A.A., Moro-Velazquez, L., et al.: Automating the analysis of eye movement for different neurodegenerative disorders. *Comput. Biol. Med.* **170**, 107951 (2024)
8. Zemblys, R., Niehorster, D.C., Holmqvist, K.: gazeNet: end-to-end eye-movement event detection with deep neural networks. *Behav. Res. Methods* **51**, 840–864 (2019)
9. Li, X., Zhang, H., Wu, S., et al.: Attention model based RNN for automated eye-movement event detection. In: Proceedings of the IEEE International Conference on Intelligent Autonomous Systems, pp. 33–38 (2021)
10. Zheng, Y., Yu, Z., Fu, H., et al.: Characterising eye movement events with multi-scale spatio-temporal awareness. *IEEE Signal Process. Lett.* (2024)
11. Krejtz, K., Duchowski, A., Krejtz, I., et al.: Discerning ambient/focal attention with coefficient K. *ACM Trans. Appl. Percept.* **13**(3), 1–20 (2016)
12. Larsson, L., Nyström, M., Stridh, M.: Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. *IEEE Trans. Biomed. Eng.* **60**(9), 2484–2493 (2013)
13. Zemblys, R., Niehorster, D.C., Holmqvist, K.: Evaluating three approaches to binary event-level agreement scoring. A reply to Friedman. *Behav. Res. Methods* **53**(1), 325–334 (2021)
14. Zemblys, R., Niehorster, D.C., Komogortsev, O., et al.: Using machine learning to detect events in eye-tracking data. *Behav. Res. Methods* **50**, 160–181 (2018)



Enhanced MobileNetV3 for Lightweight Yi Script Recognition

Haipeng Sun¹, Jian Sun^{2(✉)}, Han Yu^{3(✉)}, Bohan Guo⁴, Mengyan Liu², and Hengrui Dai²

¹ School of Chinese Ethnic Minority Languages and Literatures, Minzu University of China, Beijing 100081, China

² Computer Science and Engineering College, Dalian Minzu University, Liaoning 116600, China
m_sunjian@163.com

³ School of Foreign Languages, China University of Petroleum, Beijing 100000, China

15142100483@163.com

⁴ Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia

Abstract. As one of the six classical scripts in the world, the Yi script holds immense cultural significance, and its recognition plays a pivotal role in preserving and digitizing the cultural heritage of ethnic minorities. This study presents the first dedicated Yi character sample library for the Daliangshan region of Sichuan Province, addressing the challenges posed by limited data availability and inadequate algorithmic adaptability in Yi script recognition. Moreover, we propose a lightweight recognition model built on an enhanced MobileNetV3 architecture. By optimizing the networks depth and width, the model significantly boosts feature extraction capabilities while maintaining a compact structure. Experimental results reveal that the model achieves a recognition accuracy of 94.6% on Yi character datasets. This study provides an effective technical solution for digitizing ancient literature and enabling the intelligent processing of ethnic languages and scripts, offering valuable practical insights for advancing informatization in minority regions.

Keywords: Deep learning · Yi character recognition · MobileNetV3 · cultural heritage digitization

1 Introduction

OCR technology plays a crucial role in the digital era, enabling text recognition in languages such as Chinese, English, Korean, Japanese, and Arabic, with notable advancements in accuracy and processing power [1–3]. In China, research on ethnic minority scripts like Tibetan, Uighur, Kazakh, and Mongolian has made significant progress, resulting in a relatively mature technical system. However, the digitization of Yi, an important minority language, remains

underdeveloped. The Yi script, one of the six major ancient writing systems, is culturally and historically significant, primarily used in Sichuan, Yunnan, and Guizhou. Despite its importance, Yi character digitization relies mainly on manual methods, limiting informatization and posing challenges to the preservation of Yi heritage [4–6].

The rapid advancement of deep learning offers a promising solution for Yi character recognition. AI-based OCR systems leveraging convolutional neural networks, recurrent neural networks, and transformers autonomously extract image features, enabling accurate recognition even in complex backgrounds [7, 8]. Deep learning overcomes the challenges posed by the intricate glyphs and stroke structures of the Yi script by learning symbol patterns from large labeled datasets, improving recognition accuracy [9]. This approach aids the informatization of Yi, supporting the digital preservation and transmission of its cultural heritage. The research lays the foundation for Yi character processing, promotes minority language development, and enriches Chinese cultural heritage, providing empirical support for the global digitization of multi-ethnic languages. This paper develops a Yi script sample library and a recognition method using MobileNetV3 to enhance recognition efficiency and accuracy in digital applications.

- This study constructs a Yi script sample library consisting of recognition datasets, which are augmented with morphological transformations, such as noise addition, background insertion, translation, and rotation, to enhance diversity.
- The research focuses on Yi character recognition using the MobileNetV3 network, a lightweight model that reduces computational requirements and accelerates image feature extraction. The model achieves high recognition speed by optimizing components such as feature extraction, fusion, network depth, and loss function while maintaining accuracy.
- Experimental results demonstrate that the final model achieves 94.6% accuracy on the recognition dataset, showcasing its strong potential for Yi character recognition applications.

2 Related Works

AI-based OCR systems, powered by deep learning, offer an effective solution for Yi character recognition by leveraging CNNs, RNNs, and transformers for accurate feature extraction in complex backgrounds [10, 11]. Deep learning overcomes challenges in Yi scripts intricate glyphs and stroke structures, improving accuracy. This approach supports Yi informatization and digital preservation, advancing minority script recognition and global digitization efforts. This paper uses MobileNetV3 to enhance recognition efficiency and accuracy.

Traditional OCR methods, such as template matching and feature extraction, struggle with complex backgrounds and font variations, leading to low accuracy. Techniques like HOG and SURF with models like SVM and K-nearest neighbors

still face challenges with distorted text and complex scenes [12–15]. These methods are insufficient for recognizing minority scripts, requiring more advanced approaches.

Deep learning methods like CNNs and CRNNs capture local features and temporal relationships, improving recognition for noisy or distorted text. Innovations such as Wang et al.’s [16] curved text recognizer. Despite this, deep learning’s application to Yi recognition remains limited, as traditional methods still struggle with the complexity of Yi characters. This paper aims to develop an efficient deep learning model for Yi recognition, advancing digitization and informing future minority language research [17].

3 Dataset Construction and Methodology Improvement

3.1 Processing and Construction of Yi Character Recognition Images

Preprocessing of Yi Recognized Images. The recognition dataset in this paper is composed of an automatically generated and manually labeled real dataset, including synthetic Yi images and semi-automatically labeled Yi images. The dataset is created by segmenting the original images to obtain the authentic dataset, as shown in Fig. 1.

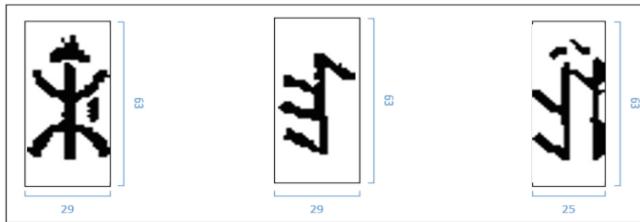


Fig. 1. Segmented images of Yi characters.

As shown in Fig. 1, Yi character images often have inconsistent aspect ratios, with edges overlapping character boundaries, complicating model training. To address this, the images are normalized, improving adaptability to varying sizes and ratios. The dataset includes both automatically generated and manually labeled real data, comprising synthetic and semi-automatically labeled Yi images. The normalization process, shown in Fig. 2, reduces size variations, enhancing the models generalization and accuracy.

Construction of Yi Language Recognition Dataset. The recognition dataset consists of two components: the synthetic dataset, generated from annotated images, and the semi-automated labeled dataset, based on model-detected cropped Yi characters and manual corrections. Experimental analysis shows that around 200 samples per character are needed to meet accuracy requirements. An example of the synthetic dataset is shown in Fig. 3.

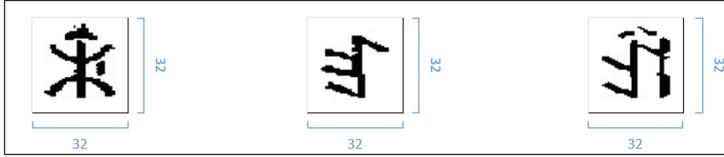


Fig. 2. Recognizing the result after normalizing the image.



Fig. 3. Synthetic datasets.

As shown in Fig. 4, this paper creates a synthetic dataset by combining 1165 Yi characters with scene images. The UTF-8 encoded corpus uses white backgrounds to match scanned Yi books. Characters are embedded using a text style migration technique, diversifying the dataset. Data augmentation (noise, blurring, and size modifications) enhances generalization. The dataset generates 200,000 images, with each character appearing 200 times, improving recognition accuracy and robustness.

3.2 Yi Character Recognition Method Based on MobileNetv3

Deep Separable Convolution. As shown in Fig. 5, depthwise separable convolution, a key technique in MobileNet, reduces computational load by decomposing standard convolution into depthwise and pointwise convolutions. Depthwise convolution applies a separate kernel to each input channel, reducing redundant computations, while pointwise convolution uses a 11 kernel to fuse features and adjust output channels. This decomposition reduces model parameters and maintains performance, making it ideal for resource-constrained environments like mobile devices and embedded systems.

Depthwise separable convolution, a key technique in MobileNet, splits standard convolution into depthwise and pointwise convolutions. In depthwise con-



Fig. 4. Yi character corpus dictionary.

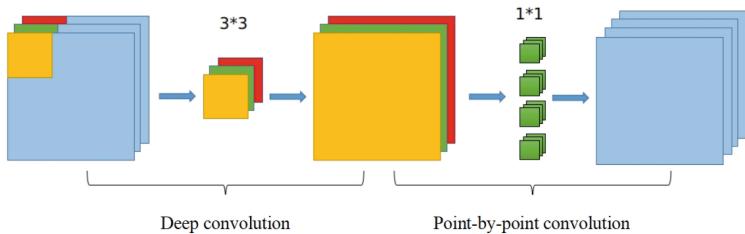


Fig. 5. Deep separable convolution.

volution, each input channel is convolved independently with its corresponding kernel, as shown in Fig. 6.

The feature map from the previous step has the same number of channels as the input, limiting capacity. To address this, a 11 convolution kernel increases the channel count. As shown in Fig. 7, combining depthwise and pointwise convolutions reduces model parameters while maintaining accuracy. This design improves efficiency, reduces computational costs, and shortens training time, offering a cost-effective deployment solution.

Proposed Module. The SE channel attention mechanism improves network representation in three steps: global average pooling compresses spatial information into a channel descriptor, two fully connected layers adjust the channel count, and the excitation-generated weight vector recalibrates the input feature map. Using 11 convolutions and the H-Sigmoid function, the SE module enhances efficiency and feature representation without adding significant complexity (Fig. 8).

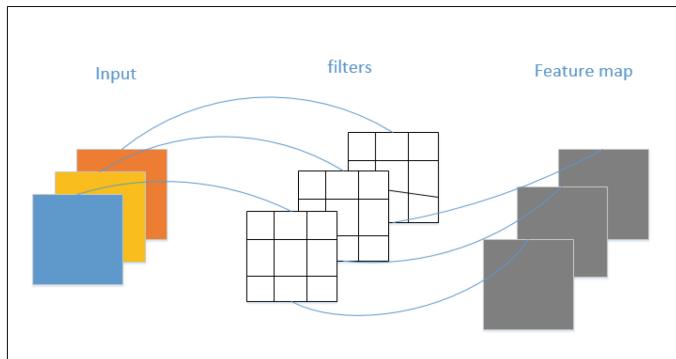


Fig. 6. Deep convolution

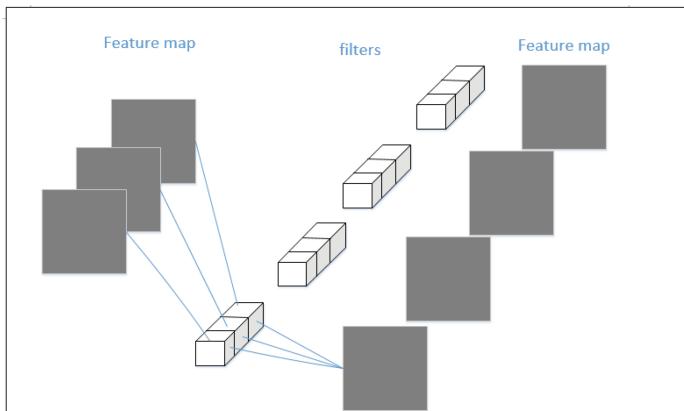


Fig. 7. Point-by-point convolution.

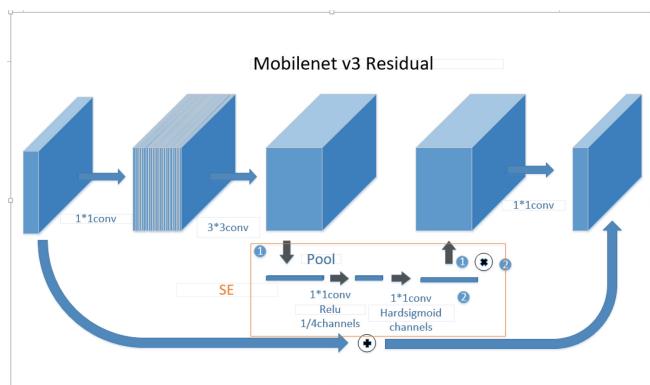


Fig. 8. Proposed module.

4 Experimental Results and Analysis

4.1 Evaluation Indicators for Yi Character Recognition

This study uses quantitative analysis to evaluate the performance of the Yi recognition model, focusing on two key metrics: recognition accuracy (ACC) and training time (H). Recognition accuracy is the primary evaluation metric, and its calculation formula is as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FN} + \text{FP}} \quad (1)$$

$$H = \text{Training time} \quad (2)$$

4.2 Results and Analysis

This study uses a high-performance computing platform with an NVIDIA RTX 3090Ti graphics card (24 GB memory), essential for deep learning training and inference. The software environment runs on Ubuntu 18.04, with the PaddlePaddle deep learning framework selected for its strong performance and stability in computer vision tasks, as shown in Table 1.

Table 1. Overview of the experimental environment

Experiment environment	Version information
Operating system	Ubuntu
Programming language	Python 3.8.0
Development tool	Pycharm
Graphics card	GTX 3090Ti
Memory	24G
Paddle	2.4.0

This study uses a systematic parameter configuration to ensure reliability and comparability. The Adam optimizer (momentum=0.9) combines momentum with an adaptive learning rate for better convergence. The learning rate follows an exponential decay, starting at 0.005 to balance fast convergence and fine-tuning. The batch size is set to 256 for optimal memory capacity and gradient stability, with training lasting 500 epochs to ensure full convergence.

This study evaluates four neural networks for Yi character recognition, as shown in Table 2. MobileNetV3 outperforms the standard CNN baseline, achieving 94.6% accuracy with the lowest time consumption. Compared to ResNet-18, MobileNetV3 reduces parameters by 65% while retaining key features, thanks to its depthwise separable convolution and lightweight attention mechanism. NRTR, benefiting from the transformer architecture, tends to overfit on medium-sized datasets. The Yi character set, consisting of about 1,000 canonical characters with well-defined stroke structures, favors lightweight networks. Deep networks risk attenuating low-level features and overfitting, leading to unnecessary

computational overhead. Experimental results highlight that lightweight networks strike the best balance between accuracy and efficiency for tasks with distinct structural features, making them ideal for minority character recognition. For systems with limited character sets and strong regularity, lightweight architectures should be prioritized over deep networks.

Table 2. Experimental results under the dataset made in this paper, showing the recognition accuracy and training time for different models

Model	ACC	H (h)
CNN	0.696	40
MobileNetV3	0.946	30
ResNet-18	0.702	45
NRTR	0.901	32

As shown in Table 3, MobileNetV3 outperforms the first two models in accuracy, with training time similar to MobileNetV1 but surpassing its performance. MobileNetV2 matches MobileNetV3 in training time but has lower accuracy, while MobileNetV1 requires more training time with the lowest accuracy. After training and testing, all models consume similar time, making MobileNetV3 the optimal choice, achieving 94.6% accuracy and meeting Yi character recognition requirements.

Table 3. Experimental results of different MobileNet under the same dataset

Model	ACC	H (h)
CNN	0.696	40
MobileNetV1	0.896	33
MobileNetV2	0.921	35
MobileNetV3	0.946	35

The proposed Yi character recognition model enables end-to-end processing, converting canonical images into digitized text (TXT format) through pre-processing, feature extraction, and recognition modules, as shown in Fig. 9. Experiments show the model effectively handles challenges such as noise in ancient book images and character variants, achieving 94.6% accuracy. This model supports ancient book digitization and corpus construction, addressing key bottlenecks in Yi character digitization and offering a solution for preserving endangered languages and scripts through the image-to-text processing paradigm.



Fig. 9. Yi character image recognition result.

5 Summary and Outlook

This study addresses two key challenges in Yi character recognition: dataset construction and algorithm optimization. We propose an "algorithmic pre-labeling + expert verification" approach to create the first professional dataset of 1,165 Yi characters, combining automated and manual labeling for high quality and efficiency. The improved MobileNetV3 model achieves 94.6% recognition accuracy, surpassing traditional methods, while its lightweight design ensures real-world applicability.

Future developments should focus on expanding multi-dialect samples for better generalization, exploring semi-supervised learning to reduce reliance on labeled data, and creating a cross-script transfer learning framework. These advancements will enhance Yi character recognition and support the digital preservation of minority languages and cultural heritage.

References

- Batra, P., Phalnikar, N., Kurmi, D., et al.: OCR-MRD: performance analysis of different optical character recognition engines for medical report digitization. *Int. J. Inf. Technol.* **16**(1), 447–455 (2024). <https://doi.org/10.1007/s41870-023-01610-2>
- Avyodri, R., Lukas, S., Tjahyadi, H.: Optical character recognition (OCR) for text recognition and its post-processing method: a literature review. In: Proceedings of the International Conference on Technology Innovation and its Applications (ICTIIA). IEEE, pp. 1–6 (2022). <https://doi.org/10.1109/ICTIIA54654.2022.9935961>
- Anakpluek, N., Pasanta, W., Chantharasukha, L., et al.: Improved tesseract optical character recognition performance on Thai document datasets. *Big Data Res.* **39**, 100508 (2025). <https://doi.org/10.1016/j.bdr.2025.100508>
- Shang, W., Chen, L.L., Xiang, W.: Yi printed character recognition based on deep learning. *Procedia Comput. Sci.* **242**, 584–591 (2024). <https://doi.org/10.1016/j.procs.2024.08.111>
- Sun, H., Ding, X., Sun, J., et al.: A method for detecting and recognizing Yi character based on deep learning. *Comput. Mater. Continua* **78**(2), 2721. <https://doi.org/10.32604/cmc.2024.046449>

6. Zhuang, H., Lin, Z., Yang, Y., et al.: An analytic formulation of convolutional neural network learning for pattern recognition. *Inf. Sci.* **686**, 121317 (2025). <https://doi.org/10.1016/j.ins.2024.121317>
7. Zhao, X., Wang, L., Zhang, Y., et al.: A review of convolutional neural networks in computer vision. *Artif. Intell. Rev.* **57**(4), 99 (2024). <https://doi.org/10.1007/s10462-024-10721-6>
8. Sun, H., Ding, X., Li, Z., et al.: Linguistic-visual based multimodal Yi character recognition. *Sci. Rep.* **15**(1), 11874 (2025). <https://doi.org/10.1038/s41598-025-96397-6>
9. Duan, J., Xiong, J., Li, Y., et al.: Deep learning based multimodal biomedical data fusion: An overview and comparative review. *Inf. Fusion*, 102536 (2024). <https://doi.org/10.1016/j.inffus.2024.102536>
10. Kasem, M. S. E., Mahmoud, M., Kang, H. S.: Advancements and challenges in Arabic optical character recognition: A comprehensive survey. arXiv preprint [arXiv:2312.11812](https://arxiv.org/abs/2312.11812) (2023). <https://doi.org/10.48550/arXiv.2312.11812>
11. Nigam, S., Verma, S., Nagabhushan, P.: Document analysis and recognition: A survey. *Authorea Preprints* (2021). <https://doi.org/10.36227/techrxiv.22336435.v1>
12. Abdulazeez, A.: A review on utilizing machine learning classification algorithms for skin cancer. *J. Appl. Sci. Technol. Trends* **5**(2), 60–71 (2024). <https://doi.org/10.38094/jastt52191>
13. Sun, J., Zhang, J., Liu, L., et al.: Output consensus control of multi-agent systems with switching networks and incomplete leader measurement. *IEEE Trans. Autom. Sci. Eng.* **21**(4), 6643–6652 (2023)
14. Sun, J., Lei, H., Zhang, J., et al.: Adaptive consensus control of multiagent systems with an unstable high-dimensional leader and switching topologies. *IEEE Trans. Ind. Inf.* **20**(9), 10946–10953 (2024)
15. Sun, J., Zhang, J., Liu, L., et al.: Event-triggered consensus control of linear multiagent systems under intermittent communication. *J. Franklin Inst.* **361**(5), 106650 (2024)
16. Shi, B., Wang, X., Lyu, P., et al.: Robust scene text recognition with automatic rectification. In: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 4168–4176 (2016). <https://doi.org/10.1109/CVPR.2016.452>
17. Liu, X., Han, X., Chen, S., et al.: Ancient Yi script handwriting sample repository. *Sci. Data* **11**(1), 1183 (2024). <https://doi.org/10.1038/s41597-024-03918-5>



Three-Dimensional Motion Estimation of Multibeam Imaging Sonar

Yupei Huang^{1,2}, Peng Li^{1,2}, Zhengxing Wu^{1,2(✉)}, and Junzhi Yu³

¹ Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

² University of Chinese Academy of Sciences, Beijing 100049, China

zhengxing.wu@ia.ac.cn

³ College of Engineering, Peking University, Beijing 100871, China

Abstract. To address the key challenges in 3D motion estimation of sonar sensors, this paper proposes a novel sonar odometry approach. The proposed method establishes a complete 3D motion estimation pipeline for sonar, incorporating two-view acoustic bundle adjustment, 3D-2D acoustic Perspective-n-Point (PnP) solving, and global bundle adjustment optimization. Without relying on planar assumptions or pre-defined artificial acoustic landmarks, the proposed method introduces a graph-based acoustic structure-from-motion technique to overcome the failure of analytical closed-form solutions caused by the nonlinear projection model. Additionally, a motion direction prior is employed to effectively resolve the symmetric dual-pose ambiguity resulting from elevation angle uncertainty in sonar measurements. The integration of global bundle adjustment allows the system to fully utilize historical observations, significantly improving pose estimation accuracy. The effectiveness of the proposed method are validated on datasets generated in a self-developed simulation environment, covering both 2D planar scenes and randomly generated 3D environments.

Keywords: 3D motion estimation · sonar odometry · graph optimization

1 Introduction

Underwater robots enable autonomous operations in complex underwater environments, supporting tasks like search and rescue, equipment inspection, seabed mapping, and resource exploration [1]. Autonomous localization through environmental perception is vital for these applications. Commonly used sensors include optical cameras and acoustic imaging sonars. Cameras offer rich visual information at low cost but are hindered in underwater settings by poor lighting, scattering, turbidity, and suspended particles, limiting vision-based state estimation [2,3]. Conversely, imaging sonars, which utilize acoustic waves, remain effective in low-visibility and turbid waters, making them more suitable for underwater perception and localization [4,5].

Multibeam imaging sonar generates two-dimensional (2D) acoustic images by emitting multiple acoustic beams and capturing their echoes, providing range and azimuth information of objects within its field of view. These images enable pose estimation of the sonar sensor [6, 7]. However, the nonlinear nature of the sonar projection model, unlike traditional optical models, precludes the use of linear analytical methods like matrix decomposition for motion estimation [8]. Additionally, the loss of elevation angle information during sonar imaging introduces symmetric dual-pose ambiguity in sonar epipolar geometry, further complicating accurate motion estimation.

Some existing methods aim to estimate three-dimensional (3D) motion from two-dimensional acoustic images but often assume planar feature landmark distribution or depend on known artificial targets [9–11], significantly limiting their applicability. While some approaches extend to 3D scenes [12, 13], they typically estimate 6-degree-of-freedom (6DOF) poses using additional motion sensors such as Doppler velocity logs (DVLs) and inertial measurement units (IMUs). However, these methods fail to address the symmetric dual-pose ambiguity caused by elevation angle uncertainty in sonar measurements. This highlights the pressing need for effective methods to estimate 3D sonar motion.

To address the above challenges, this paper proposes a novel method for 3D motion estimation of sonar sensors. The proposed approach comprises a complete sonar odometry pipeline, including 2D-2D initialization, 3D-2D acoustic Perspective-n-Point (PnP), and global bundle adjustment (BA) optimization. The main contributions of this article can be summarized as follows:

- A complete sonar odometry pipeline is proposed for estimating the three-dimensional pose of a multibeam imaging sonar device during motion. The proposed method does not rely on planar assumptions or require pre-known artificial acoustic landmarks.
- A graph optimization-based acoustic structure-from-motion method is introduced, which addresses the failure of analytical closed-form solutions caused by the nonlinear projection model of sonar. In addition, a motion direction prior is employed to resolve the symmetric dual-pose ambiguity arising from the elevation angle uncertainty in sonar measurements.
- A global bundle adjustment optimization is applied, which fully leverages historical measurement information during the motion process to enhance the accuracy of sonar pose estimation and improve robustness against noise.
- The effectiveness of the proposed method is validated on datasets collected in a self-developed simulation environment, covering both planar scenes and randomly generated three-dimensional scenes.

The remainder of this paper is organized as follows. Section 2 discusses the sonar projection model, epipolar geometry, and the symmetric dual-pose ambiguity. Section 3 describes the proposed complete sonar odometry pipeline, followed by results and analysis conducted in simulation in Sect. 4. Finally, Sect. 5 concludes the article.

2 Preliminaries

Transformation matrices $\mathbf{T} \in \mathbb{SE}(3)$ are used to represent the 3D poses, consisting of a rotation matrix $\mathbf{R} \in \mathbb{SO}(3)$ and a translation vector $\mathbf{t} \in \mathbb{R}^3$. The transformation ${}^{n_i}\mathbf{T}_{n_j}$ denotes the pose relation between the coordinate system $\{N\}$ at frame j and the coordinate system $\{M\}$ at frame i .

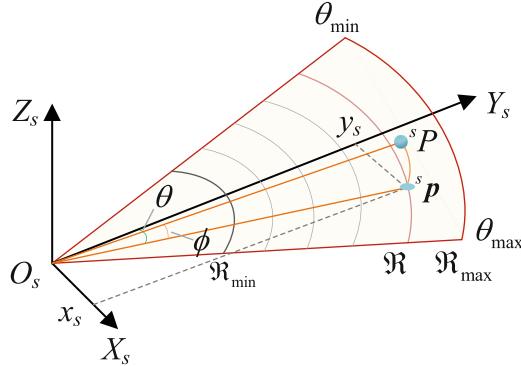


Fig. 1. The projection model of multibeam imaging sonar.

As illustrated in Fig. 1, the echo data with the same range \mathfrak{R} and horizontal azimuth angle θ are displayed on the central imaging plane of sonar. In this process, 3D points are mapped onto a two-dimensional image along an arc centered at the sonar's origin, leading to the loss of pitch angle information ϕ . In Fig. 1, the homogeneous coordinates of a 3D point ${}^s\mathbf{P}$ can be expressed as:

$${}^s\mathbf{P} = \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix} = \begin{bmatrix} \mathfrak{R} \cos \phi \sin \theta \\ \mathfrak{R} \cos \phi \cos \theta \\ \mathfrak{R} \sin \phi \\ 1 \end{bmatrix} \quad (1)$$

where \mathfrak{R} is the Euclidean distance from the 3D point to the origin of sonar coordinate frame, θ is the horizontal azimuth angle, and ϕ is the pitch angle.

Then the homogeneous coordinates of the 2D image projection ${}^s\mathbf{p}$ corresponding to the 3D point ${}^s\mathbf{P}$ can be obtained as:

$${}^s\mathbf{p} = \Pi_s({}^s\mathbf{P}) = \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} = \mathbf{Q} {}^s\mathbf{P} = \begin{bmatrix} \cos^{-1} \phi & 0 & 0 & 0 \\ 0 & \cos^{-1} \phi & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix} \quad (2)$$

where $\Pi_s(\cdot)$ represents the sonar projection function, which can be expressed by the transformation matrix \mathbf{Q} . The matrix \mathbf{Q} reflects the nonlinear nature of the sonar projection model.

The nonlinear sonar projection model, as expressed in (2), causes the essential or homography matrix to depend on the sonar motion \mathbf{R} and \mathbf{t} through complex trigonometric functions. Currently, no closed-form decomposition exists to extract motion parameters from this model [8]. Additionally, the lack of *elevation* angle information introduces the dual-pose ambiguity problem, where two symmetric pose solutions exist for the second sonar frame relative to the first, mirrored across the sonar’s X - Y imaging plane.

3 Sonar Motion Estimation Pipeline

Two-view Acoustic Bundle Adjustment. To complete the initialization of sonar motion estimation, we employ two-view acoustic bundle adjustment (ABA) to estimate the relative pose between the first two frames by optimizing a nonlinear least squares (NLS) objective based on 2D feature reprojection errors, simultaneously recovering the relative sonar motion and 3D landmark positions. We assume that the feature correspondence has already been obtained from feature extraction and matching algorithms such as SuperPoint and SuperGlue.

Figure 2a shows the factor graph for two-view ABA optimization, comprising sonar poses at two viewpoints ${}^w\mathbf{T}_{s_0}$ and ${}^w\mathbf{T}_{s_1}$ and k observed 3D acoustic landmarks ${}^w\mathbf{P}_1, \dots, {}^w\mathbf{P}_k$. The first sonar pose is fixed as the reference frame W (identity matrix), with all subsequent poses and landmarks represented relative to it. Thus, only ${}^w\mathbf{T}_{s_1}$ and the landmarks are optimized. Therefore, the complete formulation of the two-view ABA optimization can be expressed as:

$$\begin{aligned} \mathcal{S} &= \{{}^w\mathbf{T}_{s_1}, {}^w\mathbf{P}_1, \dots, {}^w\mathbf{P}_k\} \\ \mathcal{S}^* &= \underset{\mathcal{S}}{\operatorname{argmin}} \left\{ \sum_j^k \|{}^{s_0}\mathbf{p}_j - \Pi_s({}^w\mathbf{P}_j)\|^2 \right. \\ &\quad \left. + \sum_j^k \|{}^{s_1}\mathbf{p}_j - \Pi_s({}^w\mathbf{T}_{s_2}^{-1} \cdot {}^w\mathbf{P}_j)\|^2 \right\} \end{aligned} \quad (3)$$

where $\Pi_s(\cdot)$ denotes the sonar projection function. The problem (3) can be treated as a nonlinear least squares optimization and solved using the Levenberg-Marquardt (LM) method [14]. Since there are $6+3k$ variables and $4k$ constraint equations, each sonar image should provide at least 6 acoustic feature points to fully constrain this optimization problem.

Elimination of Dual-Pose Ambiguity. We propose to utilize the simple prior information about the Z -axis motion direction to eliminate the dual-pose ambiguity. The Z -axis motion direction prior can be easily obtained from common motion sensors such as IMU or pressure sensors. Specifically, we need to ensure that the sign of the Z -axis component in the translation vector of the second sonar pose, ${}^w\mathbf{T}_{s_1}$, is consistent with known motion direction along the Z -axis. This enables accurate initialization through two-view ABA optimization, successfully recovering both the sonar pose and 3D acoustic landmarks.

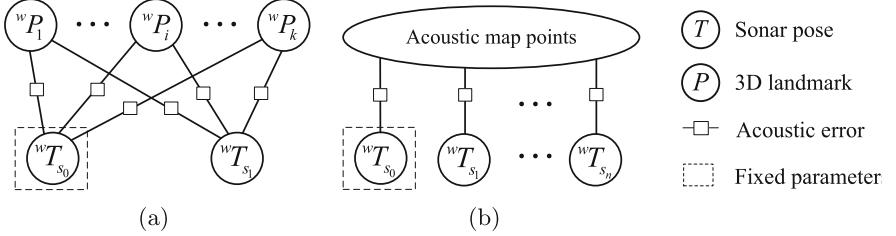


Fig. 2. Illustration of the factor graph used in sonar motion estimation. (a) Two-view acoustic bundle adjustment. (b) Global bundle adjustment of Sonar.

3D-2D PnP of Sonar. After two-view ABA optimization and resolving pose ambiguity, the sonar poses and 3D landmarks from the first two frames are recovered. For subsequent frames, sonar motion is estimated similarly to the 3D-2D PnP method: based on the feature correspondences, the sonar pose at new frame is optimized by minimizing 3D-2D reprojection errors.

Global Bundle Adjustment of Sonar. To enhance sonar motion accuracy, a global bundle adjustment (GBA) is applied using feature measurements from all frames to refine sonar poses and 3D landmarks, as shown in Fig. 2b. After previous steps, we have $n+1$ sonar poses ${}^wT_{s_0}, {}^wT_{s_1}, \dots, {}^wT_{s_n}$ and m 3D landmarks ${}^wP_1, \dots, {}^wP_m$, with ${}^wT_{s_0}$ fixed as the identity. The global BA optimization is then formulated as:

$$\begin{aligned} \mathcal{X}_s &= \{{}^wT_{s_0}, {}^wT_{s_1}, \dots, {}^wT_{s_n}; {}^wP_1, \dots, {}^wP_m\} \\ \mathcal{X}_s^* &= \underset{\mathcal{X}_s}{\operatorname{argmin}} \sum_i^n \sum_{j \in \mathcal{N}_{s_i}} \|{}^{s_i}p_j - \Pi_s({}^wT_{s_i} \cdot {}^wP_j)\|^2 \end{aligned} \quad (4)$$

where $\Pi_s(\cdot)$ is the sonar projection function, and \mathcal{N}_{s_i} denotes the number of 3D points observed by the i -th sonar pose.

4 Results and Analyses

To validate the effectiveness of the proposed sonar motion estimation method, evaluations and analyses were conducted in a simulated environment. In the simulation tests, the sonar's azimuth range was set to $[-65^\circ, 65^\circ]$ and the elevation range to $[-10^\circ, 10^\circ]$. We conducted evaluations in two types of scenarios to demonstrate the applicability of the proposed method. Specifically, in the random 3D scene, 20 three-dimensional points were randomly generated to serve as landmarks for sonar observations, as shown in Fig. 3a. In the planar scene, a planar board with 4×5 uniformly distributed features was generated. In both scenarios, multiple sonar poses were set and simulated sonar observation images were produced at these poses, as illustrated in Fig. 3b. To account for noise, zero-mean Gaussian noise with a standard deviation of 5 mm was added to the

simulated sonar images. This is because the independent Gaussian noise model of acoustic feature uncertainty will be transformed into a Rayleigh distribution for sonar distance measurement, which is consistent with the speckle noise model of sonar imaging systems.

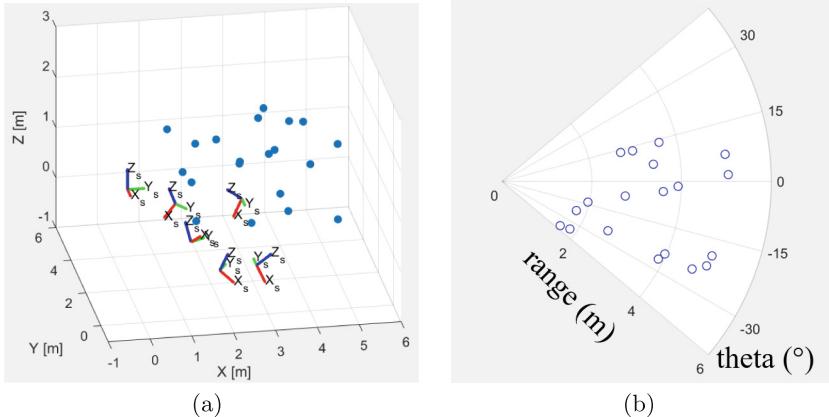


Fig. 3. The self-developed simulation environment and simulated acoustic 2D images. (a) The sonar poses and randomly generated 3D landmarks in simulation. (b) Acoustic image generated by projecting 3D landmarks onto a specific sonar pose.

Table 1. RMSE ATE (in meters) of Sonar Motion Estimation Results

Sequence	FLS-PnP		Ours without GBA		Ours with GBA	
	Average	Median	Average	Median	Average	Median
3D Scene	x	x	0.0865	0.0814	0.0127	0.0132
Planar Scene	0.0708	0.0717	0.0816	0.0758	0.0179	0.0173

We selected the FLS-PnP method [11] as a baseline for comparison, which relies on geometric priors from surrounding structures. In the planar scene tests, we provided it with 2D planar feature information to enable its operation. To simulate manufacturing errors or degradation in real-world conditions, noise was added to the planar board information supplied to the FLS-PnP method. The estimated sonar motion was compared against the ground truth, and the root mean square error (RMSE) of the absolute trajectory error (ATE) was reported. The data for each scene was tested 7 times, and the mean and median results are presented in Table 1.

As shown in Table 1, the FLS-PnP method failed in the three-dimensional scene due to its reliance on planar feature distribution and the absence of prior

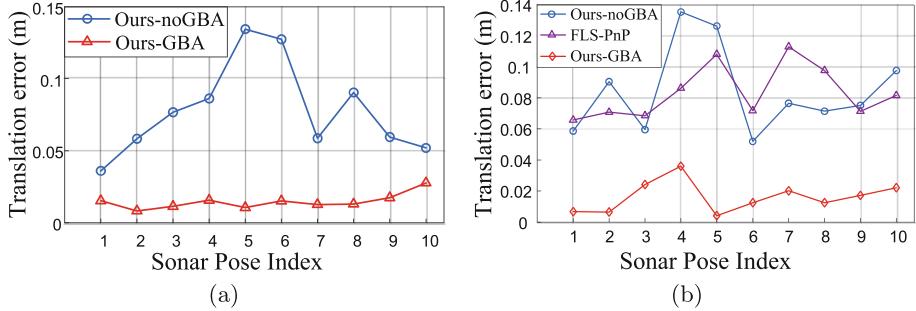


Fig. 4. The translation error curves at different sonar poses. (a) The error curves of the proposed method in 3D scenes. (b) The error curves of the proposed method and comparison method FLS-PnP in planar scenes.

structural information. In contrast, the proposed method was able to successfully estimate the sonar poses using the simulated acoustic images. Furthermore, benefiting from the proposed Global Bundle Adjustment (GBA) module, which optimizes the sonar poses using all valid observations throughout the motion sequence, the RMSE ATE error was significantly reduced. In the planar scene, the FLS-PnP method was able to run successfully when provided with structural information of the planar features. However, due to its strong reliance on prior information, the accuracy of its pose estimation is limited when the quality of planar feature structural information deteriorates and noise occurs. In this scenario, the initial results produced by our method through the two-view acoustic bundle adjustment and 3D-2D PnP processes were comparable to those of FLS-PnP. Moreover, after further optimization using the GBA module, the accuracy of the sonar pose estimation was significantly improved. The translation error curves shown in Fig. 4 further support the above conclusions. Our method can successfully achieve sonar motion estimation in both random 3D and planar scenes, and the proposed global bundling adjustment module can significantly improve pose estimation accuracy, demonstrating the effectiveness of the proposed method.

5 Conclusion

This paper presents a novel 3D motion estimation method for multibeam imaging sonar and develops a complete sonar odometry system. Without relying on artificial landmarks or planar scene assumptions, the proposed approach integrates graph-based optimization with motion priors to effectively address challenges caused by the nonlinear projection model and symmetric dual-pose ambiguity. The introduction of global bundle adjustment module further enhances the system's robustness and accuracy. Simulation results demonstrate that the proposed method achieves stable and reliable sonar pose estimation in both 3D and 2D scenes, providing valuable guidance for applications such as underwater robot

localization and navigation. Future work will focus on real-world underwater experiments to further validate the proposed method.

Acknowledgment. This work was supported in part by the National Natural Science Foundation of China under Grant 62433021, in part by Beijing Natural Science Foundation (JQ24026), in part by Youth Innovation Promotion Association CAS (2023039), and in part by the Hebei Natural Science Foundation under Grant F2024203130.

References

1. Zhang, B., Ji, D., Liu, S., Zhu, X., Xu, W.: Autonomous underwater vehicle navigation: a review. *Ocean Eng.* **273**(113861) (2023)
2. Song, J., Li, W., Zhu, X.: Acoustic-VINS: tightly coupled acoustic-visual-inertial navigation system for autonomous underwater vehicles. *IEEE Robot. Autom. Lett.* **9**(2), 1620–1627 (2024)
3. Joshi, B., et al.: Experimental comparison of open source visual-inertial-based state estimation algorithms in the underwater domain. In: Proceeding IEEE/RSJ International Conference Intelligent Robotics System (IROS), Macau, China, pp. 7227–7233 (2019)
4. Rahman, S., Quattrini Li, A., Rekleitis, I.: SVIn2: a multi-sensor fusion-based underwater SLAM system. *Int. J. Robot. Res.* **41**(11–12), 1022–1042 (2022)
5. Xu, S., Zhang, K., Hong, Z., Liu, Y., Wang, S.: DISO: direct imaging sonar odometry. In: Proceeding IEEE International Conference Robotics Automation (ICRA), Yokohama, Japan, pp. 8573–8579 (2024)
6. Wang, Y., Ji, Y., Tsuchiya, H., Ota, J., Asama, H., Yamashita, A.: Acoustic-n-point for solving 2D forward looking sonar pose estimation. *IEEE Robot. Autom. Lett.* **9**(2), 1652–1659 (2024)
7. Huang, T.A., Kaess, M.: Towards acoustic structure from motion for imaging sonar. In: Proceeding IEEE/RSJ International Conference Intelligent Robotics System (IROS), Hamburg, Germany, pp. 758–765 (2015)
8. Negahdaripour, S.: On 3-D motion estimation from feature tracks in 2-D FS sonar video. *IEEE Trans. Robot.* **29**(4), 1016–1030 (2013)
9. Sekkati, H., Negahdaripour, S.: 3-D motion estimation for positioning from 2-D acoustic video imagery. In: Proceeding Girona, Spain, Jun, Iberian conference on Pattern Recognition and Image Analysis, pp. 80–88 (2007)
10. Negahdaripour, S.: Analyzing epipolar geometry of 2-D forward-scan sonar stereo for matching and 3-D reconstruction. In: Proceeding Charleston, SC, USA, Oct, IEEE OCEANS, pp. 1–10 (2018)
11. Yang, D., He, B., Zhu, M., Liu, J.: An extrinsic calibration method with closed-form solution for underwater opti-acoustic imaging system. *IEEE Trans. Instrum. Meas.* **69**(9), 6828–6842 (2020)
12. Westman, E., Kaess, M.: Degeneracy-aware imaging sonar simultaneous localization and mapping. *IEEE J. Oceanic Eng.* **45**(4), 1280–1294 (2020)
13. Westman, E., Hinduja, A., Kaess, M.: Feature-based SLAM for imaging sonar with under-constrained landmarks. In: Proceeding IEEE International Conference Robotics Automation (ICRA), Brisbane, QLD, pp. 3629–3636 (2018)
14. Kümmel, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: G2o: a general framework for graph optimization. In: Proceeding IEEE International Conference Robotics Automation (ICRA), Shanghai, China, pp. 3607–3613 (2011)



A High Throughput Deep Learning Platform for Edge Inference Based on Hardware/Software Codesign

Yuwei Zhang, Fan Wu, and Chaonong Xu^(✉)

China University of Petroleum, Beijing 102249, China
xuchaonong@cup.edu.cn

Abstract. This paper proposes a deep learning platform for edge inference based on hardware/software codesign. On the hardware side, to achieve high inference throughput within the constraints of edge device, we introduce a star-style chip cluster architecture, which utilizes multiple NPUs connecting by PCIe bridge to expand deep learning computational power, and uses a CPU as the execution platform for task scheduling strategies. On the software side, hybrid parallelism scheduling strategy is designed based on data parallelism scheme on the whole and pipeline parallelism for each NPU task, so as to achieve high inference throughput. To maximize the efficiency of pipeline parallelism, we utilize the method of balancing pipeline segments. Take the most common object detection task for example, according to its processing flow and time characteristics, both pre-processing segment and post-processing segment are migrated onto CPU, and the number of threads for executing them are adjusted to balance with the intelligent processing segment. Compared with the traditional inference based on a single NPU, the system achieves the throughput acceleration ratio of 2.6X on COCO dataset, and its peak computing energy ratio is increased up to 2.5X. Last but not least, we also summarize some basic design principles based on observed experiment results, which presents useful insights for further performance improvement.

Keywords: Edge intelligence · Throughput · Deep learning · Hybrid parallelism

1 Introduction

As intelligent application scenarios continue to expand and the demand for real-time data processing grows, edge intelligence has become an important direction for artificial intelligence (AI) research.

In edge scenarios, due to various limitations such as size, power supply, and real-time processing requirements, edge intelligence often faces the challenge of insufficient computational power, greatly affecting its application in some industries. For example, in the intelligent application of high-speed drones, drones

need to take high-frequency images and quickly recognize targets in the images. If a drone shoots at a rate of 10 Frames Per Second (FPS), the time for processing each image cannot exceed 0.1 s. This is a requirement that current typical edge intelligence computing platforms like Jetson Nano (maximum computational power is 0.47 Tera Operations Per Second (TOPS)) [1], Jetson TX2 (maximum computational power is 1.3 TOPS) [2], and Jetson Xavier NX (maximum computational power is 21 TOPS) [3] are unable to meet. In this context, augmenting computational power by scaling up Graphics Processing Unit (GPU) resources is a conventional approach to overcome these challenges. NVIDIA's DGX-1 [4] exemplifies this strategy, representing a high-performance GPU cluster system that significantly amplifies computational power by aggregating multiple GPUs. It is designed and calibrated for environments where high-performance computing is paramount. However, its large size and high energy consumption (equipped with four 1600 W power supply units [5]) clearly distinguish it is unsuitable for edge intelligence scenarios.

Inspired by the architecture of DGX-1 and all kinds of constraints presented in edge computing, we propose a high throughput solution for edge inference systems based on hardware/software codesign. On the hardware side, considering the constraints of low power consumption and small size, we expand deep learning capabilities by bridging multiple Neural Processing Unit (NPU) chips with a Central Processing Unit (CPU), thereby forming the so-called star-style chip cluster architecture. On the software side, we enhance inference throughput by running application-oriented task scheduling algorithm so as to fully utilize the computational power of star-style chip cluster.

Specifically, in our design, we have expanded the computational power for Deep Learning (DL) by bridging four MLU220s via PCIe 3.0, and using FT-2000/4 CPU as the execution platform for task scheduling strategies. This design allows up to four NPUs to work in parallel, which directly increases computational power, and also significantly optimizes power efficiency and compactness by facilitating the shared utilization of non-NPU resources. On the software side, to meet the demand for high throughput inference, the scheduling strategy is designed based on a hybrid parallelism scheme, where data parallelism inference is adopted on the whole and pipeline parallelism is for each NPU task. Specifically, taking the most common object detection task as an example, four NPUs use same intelligent model and their data sources are assigned by CPU, so they can perform object detection independently. At the same time, we partitioned the object detection task into three parts, using pipeline parallelism to accelerate task inference. Hybrid parallelism scheme is developed both application-oriented and hardware-oriented, which lays the foundation for high throughput.

To evaluate the performance of deep learning platform based on hardware-software codesign proposed by us, we ran YOLOv5s model on the platform for object detection on COCO dataset. The experimental results show that compared with the traditional inference based on a single NPU, the throughput of our scheme is greatly related to the number of NPUs adopted. In our platform, our scheme can achieve up to a 2.6X throughput acceleration ratio, and

its Peak Computing Energy Ratio (PCER) is increased up to 2.5X. Last but not least, based on the observed experimental results, we also summarize some basic design principles for the co-designed platform, which lays theoretical foundation for further improving of our platform.

The main contributions of this paper are summarized as follows:

- We propose a high throughput deep learning platform for edge inference based on hardware/software codesign. On the hardware side, we propose the star-style chip cluster architecture, and on the software side, we propose hybrid parallelism which can fully utilize the computing capabilities of hardware platform.
- We evaluated the efficiency of the star-style chip cluster architecture using an object detection application. It can achieve up to 2.6X throughput acceleration ratio on COCO dataset using YOLOv5s model compared to the classic manner of single NPU, and its peak computing energy ratio is increased up to 2.5X.
- Based on our experimental results, we have summarized some basic design principles for the chip cluster architecture, which presents useful hints for further performance improvement.

The remainder of this paper is as follows. In Sect. 2, we discuss current issues in edge intelligence device and introduce our high throughput systematic method, where the architecture of star-style chip cluster and hybrid parallelism acceleration algorithm are focused on. We also introduce a platform realization based on MLU220 NPU and FT2000/4 CPU. Sect. 3 presents the performance evaluations and further discusses the design principle of the co-designed platform. The last section concludes the paper.

2 Problem Analysis And Solution Proposal

2.1 Problem Analysis

Edge intelligence seamlessly integrates the benefits of edge computing, namely low latency and high bandwidth efficiency, while embracing the sophisticated capabilities of artificial intelligence in data analytics and intelligent decision-making. As highlighted in [6], this integration enables intelligent applications, such as facial recognition systems and voice-enabled assistants, to perform efficient local processing during inference tasks. However, a substantial proportion of current edge intelligence platforms are constrained in their ability to handle high-throughput applications. These platforms often exhibit a limited configuration, typically comprising a single CPU coupled with a solitary NPU, characterized by inadequate computational power.

Taking the object detection process of a single image as an example, Fig. 1 presents the intelligent image processing flowchart on these platforms, which delineates three key stages: pre-processing, inference, and post-processing.

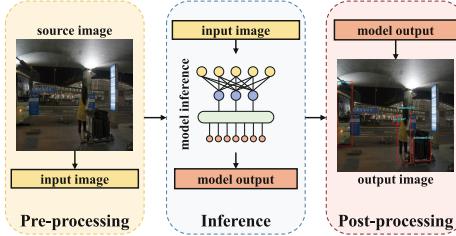


Fig. 1. Three steps for processing a single image.

The workflow begins with CPU performing initial pre-processing on the original image, which is then forwarded to NPU for the inference stage. Upon completion of this stage, CPU takes over again to carry out post-processing, culminating in the final output. However, this traditional setup often falls short in scenarios that demand high throughput. For example, in a heterogeneous inference architecture that integrates a FT2000/4 CPU with a MLU220 NPU which has 16TOPS@INT8 computational power, employing the YOLOv5 algorithm to process images from COCO dataset results in a throughput of about 22 FPS. This throughput is notably insufficient for practical applications which have larger image size and higher throughput requirements.

In light of these limitations, we need to consider how to design a hardware architecture that meets the needs of such applications. Further, under the architecture, how to design a scheduling algorithm for intelligent interference that works in concert with the hardware to maximize system throughput, in other words, a deep learning platform for high throughput based on hardware/software codesign, becomes the main research content of this paper.

2.2 Platform Proposal

(1) Hardware Architecture Design

To achieve high throughput, we refer to the hardware architecture of NVIDIA DGX-1, which is now the most essential platform for AI research. To fully utilize the capability of the hardware architecture, acceleration approaches based on data parallelism and pipeline parallelism, which have been widely used in SOTA inference software, such as FasterTransformer by Nvidia [7], DeepSpeed by Microsoft [8], and ORCA by FriendlAI [9], can be adopted.

Hinted by the design of DGX-1 and taking the size, power consumption, computation requirement and NPU chip status into consideration, we present a simplified hardware architecture for edge inference, which is shown in Fig. 2. In this figure, multiple NPUs (4 NPUs in Fig. 2) are connected through a PCIe3.0 switch, which is widely supported in current NPUs. Additionally, a CPU is also connected to this switch, thereby forming the so-called star-style chip cluster architecture. Besides, there is a Field-Programmable Gate Array (FPGA) chip to deal with some extra processing requirements.

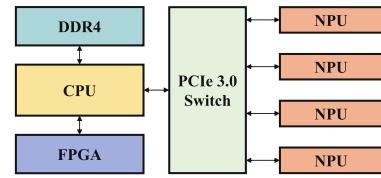


Fig. 2. The star-shaped chip cluster architecture.

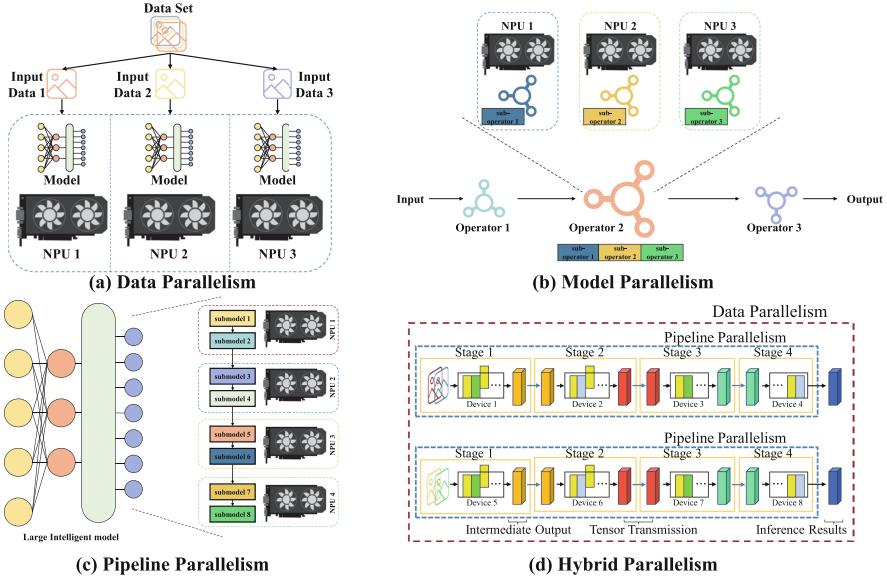


Fig. 3. Common inference acceleration techniques.

In this architecture, NPUs play the central role in executing deep learning tasks, focusing on high-intensity parallel computing tasks, particularly in the inference of deep learning models. CPU undertakes key functions such as platform-level task scheduling and data transfer, ensuring the efficiency and coordination of system operation. DDR4 memory, serving as the primary storage medium of the system, ensures the continuity and real-time processing of data with its high-speed access capabilities and ample storage capacity.

(2) Inference Middleware Architecture

The key point of inference middleware should be supporting the intelligent models of all sizes, while achieving high throughput. Considering the hardware architecture of multiple NPUs, the middleware should support data parallelism, model parallelism, pipeline parallelism, and hybrid parallelism, all aimed at fully leveraging the potential of the proposed star-style chip cluster architecture.

Data Parallelism: For intelligent models of a general scale that can be processed entirely by a single NPU, data parallelism method illustrated by Fig. 3(a) is a feasible way to achieve high throughput. In this approach, CPU is responsible for loading the intelligent model into every NPU, feeding input data to NPUs for parallel processing, and combining the results, while NPUs execute the same intelligent processing program.

Model Parallelism: For some intelligent models where the size of their large operators can exceed NPU memory capacity, the model cannot be executed on the traditional single-CPU-single-NPU architecture. Therefore, as a solution,

model parallelism illustrated by Fig. 3(b) can be adopted to solve the problem. Every time a large operator (Operator 2 in Fig. 3(b)) is encountered, CPU will partition it into multiple parallelized sub-operators, load them into NPUs, then combine the results from sub-operators, and progress to the next operator. Obviously, in this way, intelligent models with large operators can be supported.

Pipeline Parallelism: For some large intelligent models where the size is large enough, we can utilize pipeline parallelism to deal with them. As illustrated in Fig. 3(c), the large model is partitioned into several sub-models (8 sub-models in Fig. 3(c)). A NPU will execute the inference of at least one sub-models, and they form an execution pipeline until last NPU outputs the final result. CPU is only responsible for loading sub-models for NPUs, acquiring input data and feeding them into first NPU, and getting the final result from last NPU. Obviously, in this manner, the theoretical maximal acceleration ratio is the number of NPUs adopted.

Hybrid Parallelism: We can employ a parallelism scheme which combines the advantages of data parallelism, model parallelism, and pipeline parallelism, allowing for flexible adjustment of the degree of parallelism during the inference process. As shown in Fig. 3(d), there are two groups, each of which consists of four NPUs, and data parallelism can be adopted by two groups, and four NPUs in the same group can use pipeline parallelism to accelerate inference[10].

Through the aforementioned parallelization strategies, the middleware architecture demonstrates high flexibility and scalability, and can be configured and optimized according to the needs of different models and the status of hardware resources.

A clever design of the middleware is the load balancing in pipeline parallelism, particularly within application scenarios demanding high throughput. We all know that only when the execution time of every pipeline stage is equal can the throughput be maximized. According to conventional thinking, partitioning the model evenly among all NPUs is the optimal way to enhance throughput. In fact, CPU can also be used as a pipeline stage, and since its utilization ratio can be adjusted by multi-thread technology, the greater its computational power, the more flexible room it will provide for partitioning the model into sub-models. In our middleware, we also make use of the multi-thread technology of CPU to adjust model partition to enhance throughputs when employing pipeline parallelism.

(3) A Realistic Design Case

Based on the design guideline of star-style chip cluster, we design a real computation platform for high throughput as follows.

- Hardware Component

To ensure that the proposed platform achieves the requisite high-performance computing, energy-efficient operation, and elevated security and reliability



Fig. 4. The actual realization of star-style chip cluster architecture.

within edge intelligence inference contexts, we have chosen MLU220 from Cambriicon Corporation as NPU of star-style chip cluster architecture. This NPU is endowed with a compact form factor and remarkably low energy consumption, at merely 8.25 W [11]. FT-2000/4 has been selected as CPU for our platform. Its design is tailored to meet the demands for energy efficiency and seamless integration with embedded systems. FT-2000/4 CPU, featuring four FTC663 cores, offers robust computational capabilities while maintaining a typical operational power consumption of 10 W [12]. This combination of power efficiency and performance makes FT-2000/4 an ideal choice for the computational foundation of our platform.

We have expanded the computational power for deep learning by bridging four MLU220s via a PCIe 3.0 switch. The platform's architectural composition, as illustrated in Fig. 4, has been designed to address some specific demands of various edge inference scenarios. Through heat simulation evaluation, we customize a heat radiator according to its heat distribution, and use two fans of small size to satisfy the heat dissipation requirement. This design facilitates minimally impacting the increase in energy consumption. This approach not only amplifies the system's ability to process complex deep learning tasks but also ensures that the platform remains energy-efficient and scalable for various edge intelligence applications.

Moreover, the ability of FPGA to perform parallel processing operations at the hardware level can significantly reduce latency and increase throughput for certain tasks, such as signal processing, data pre-processing, and custom protocol handling [13]. Additionally, its quick response times make it an ideal candidate for augmenting the computational power of the platform without compromising energy efficiency. Considering the factor of power consumption and computation requirement, we incorporate EF2L15LG100B FPGA chip from Anlogic corporation into the platform [14].

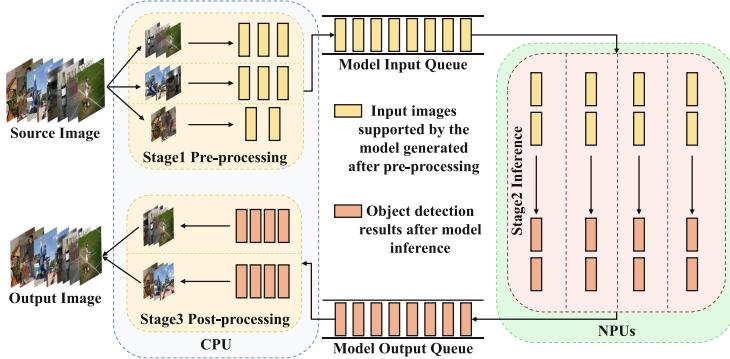


Fig. 5. Architecture of inference software.

- Software for High-throughput Object Detection

In the context of high-throughput requirements in edge intelligence computing, we select the widely applied task of object detection and optimize the system through a hybrid parallelism strategy, aiming to achieve a higher inference throughput. This strategy integrates the technical advantages of data parallelism and pipeline parallelism, reasonably scheduling pre-processing, inference, and post-processing stages according to the characteristics of the object detection task. As depicted in Fig. 5, hybrid parallel scheduling strategy assigns pre-processing and post-processing stages of object detection to CPU, while the inference stage is executed on NPU, utilizing queue structures for data buffering and transition.

These three stages together form the three main segments of pipeline parallelism, specifically, the inference stage employs a data parallel approach, enabling NPUs to load the same model and process different subsets of data in parallel.

Specifically, during pre-processing stage of the object detection task, we adopt different image capture methods according to different application scenarios. After acquisition, the images undergo a series of transformations through OpenCV functions to align the image data with the model's input specifications. Subsequently, the pre-processed images are incorporated into the model input queue, awaiting their deployment in the subsequent stage of the processing pipeline.

Within the inference stage, the utilization of MLU220 NPU necessitates the invocation of Cambricon Neuware Runtime Library (CNRT), which facilitates high-efficiency offline model inference. Within the pipeline framework, we extract input images from the model input queue, subject them to object detection via MLU220 NPU, and subsequently forward the resulting detections to the model output queue for further processing stages.

In the post-processing stage, we employ relevant functions from OpenCV to perform visualization operations on the original images by the detection results obtained from the model output queue. Ultimately, the processed images are

either locally stored or directed towards subsequent operations, depending on the integration with downstream applications.

During the inference stage, the implementation of data parallelism strategy enables multiple NPUs to concurrently perform model inference for object detection, thereby accelerating the inference process to meet the requirements for high throughput.

From a holistic software perspective, achieving enhanced throughput necessitates addressing the critical issue of balancing segment processing times within pipeline parallelism. Fortunately, our proposed platform offers adjustable parameters, such as the number of threads allocated for image processing on CPU and the number of utilized NPU. Considering each processing step as an independent task, pre-processing and post-processing can be executed in parallel using multiple threads. Referring to Fig. 5 as an example, we can adjust CPU threads dedicated to pre-processing and post-processing tasks to 3 and 2, respectively. Concurrently, the inference stage utilizes four NPUs to execute the task of object detection.

Assuming the processing of a single image, the time required for pre-processing is denoted as T_{pre} , the time for the inference segment as T_{inf} , and the time for post-processing as T_{post} . Let N_{pre} and N_{post} represent the number of threads used for pre-processing and post-processing, respectively, and N_{inf} denote the number of NPUs utilized in the inference segment. When processing a large volume of images and addressing the demand for high throughput, the average processing time per segment for pre-processing, inference, and post-processing becomes T_{pre}/N_{pre} , T_{inf}/N_{inf} , and T_{post}/N_{post} , respectively. To attain the highest throughput in hybrid parallelism, it is necessary to equalize these average segment processing times. That means:

$$\frac{T_{pre}}{N_{pre}} = \frac{T_{inf}}{N_{inf}} = \frac{T_{post}}{N_{post}} \quad (1)$$

However, since the time required to process a single image in each segment is relatively fixed, the problem of balancing the pipeline segment processing times is transformed into the optimization of N_{pre} , N_{inf} , and N_{post} . Therefore, to achieve the highest throughput under hybrid parallelism strategy, it is necessary to determine the optimal number of threads for pre-processing and post-processing, taking into account the number of NPUs utilized, in order to balance the processing times across all pipeline stages.

3 Evaluauon

In this section, we illustrate the performance of our platform, highlighting its efficiency in managing high-throughput computational tasks. The evaluation of performance is anchored by the metric of FPS, derived from our object detection software. FPS is defined as:

$$FPS = \frac{n}{T_{out}(n-1) - T_{in}(0)} \quad (2)$$

where n represents the total number of source images, $T_{in}(i)$ signifies the instant when image i is input to the program, $T_{out}(i)$ denotes the moment when image i finishes processing.

Table 1. Processing time of every stage in pipeline

Symbol	Processing time (ms)
T_{pre}	30ms
T_{inf}	47ms
T_{post}	6ms

3.1 Experimental Setting

Our experimental framework is designed to evaluate the inference acceleration ratio of our proposed heterogeneous AI chip cluster system. The software environment is constructed on Ubuntu 18.04 and utilizes C++ for programming. We use CNRT for inference execution on MLU220 and OpenCV for image pre-processing and post-processing, which means that the pre-processing and post-processing must be executed on FT2000/4 CPU. YOLOv5s model is chosen for object detection due to its renowned accuracy and efficiency. This model was trained on RTX3090 GPU in advance using COCO dataset, and then was quantized using INT8 so as to deploy on MLU220.

3.2 Experimental Procedure

In accordance with (1), the equilibrium of processing times across the various pipeline stages is vital for throughput when pipeline parallelism is adopted. To utilize the chip cluster to accelerate inference, we must know the elementary execution time for pre-processing, inference, and post-processing, so as to provide guides for adjusting computational power for each pipeline stage to achieve the pipeline stage equilibrium.

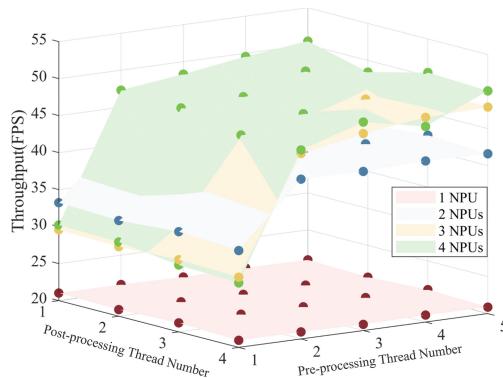
(1) Measuring Processing Time for Each Stage

In this study, we assess the processing time of handling a single image. We configure the thread number to be 1 for both pre-processing and post-processing stages, utilizing only one NPU to run YOLOv5s model. We randomly choose an image from COCO dataset, and then convert its resolution to 826×826 . We placed timers at the beginning and end of each pipeline segment, with a precision of microseconds, to record the times we needed.

Table 1 reveals that the inference executed on NPU is the most time-consuming stage, while the post-processing on CPU is the smallest. The findings indicate that the processing times of the three pipeline stages have great gaps, which is an undesirable for high throughput.

Table 2. Anticipated optimal thread number

Test scenario	Pre-processing thread number anticipated	Post-processing thread number anticipated
1 NPU	0.64 (1)	0.13 (1)
2 NPUs	1.28 (2)	0.26 (1)
3 NPUs	1.91 (2)	0.38 (1)
4 NPUs	2.55 (3)	0.51 (1)

**Fig. 6.** Throughput under hybrid parallelism.

(2) Throughput of Chip Cluster Platform

To attain high throughput, it is necessary to regulate the matching of segment processing times. Since pre-processing and post-processing stages are executed on CPU, and YOLOv5s is executed on NPU, they can be accelerated by multithreading and data parallelism, respectively. Further, if the time of pre-processing and post-processing are modelled as being linear with the thread number, and the inference time of YOLOv5s is also linear with the number of NPU adopted under data parallelism, the optimal thread number for pre-processing and post-processing can be computed manually, which is presented in Table 2. Note that the number in brackets is the optimal thread number since it should be an integer.

To attain enhanced stability in our experimental results, we repeated the object detection task 200 times using the identical image sourced from COCO dataset. It seems that verifying the conclusions presented by Table 2 appears to be a straightforward process: one would simply need to alter the thread number for pre-processing and post-processing, record FPS, and identify the optimal thread number corresponding to the maximum FPS. However, a critical issue arises during experimentation: as the number of NPUs increases, we can anticipate that the system throughput capacity will also increase, and thus only more

Table 3. Throughput performance

Test scenario	Peak throughput (FPS)
1 NPU	22.01
2 NPUs	40.22
3 NPUs	51.96
4 NPUs	58.23

images should be pre-processed and then feed to NPUs. In other words, the optimal pre-processing thread number will also increase. However, for any CPU, once its computing load is saturated, increasing thread number will have no benefits. Figure 6 presents the actual experiment results. We can find that for the case of 2 NPUs and 3 NPUs, the optimal thread numbers of pre-processing and post-processing are consistent with our anticipations in Table 2. However, for the case of 4 NPUs, system throughput almost remains constant when pre-processing thread number increases from 2 to 5, whatever post-processing thread number is. To explain the fact, we check the utilization ratio of CPU by using HTOP command, which could display the utilization ratio of CPU in real-time. We found that when the thread number of pre-processing is 1, 2 and 3, CPU utilization ratio is 46%, 92% and 100%, respectively, which shows that CPU is fully saturated when number of pre-processing is larger than 2. Therefore, continuing to increase the thread number of pre-processing will have no help for boosting throughput.

Using the optimal thread number we have obtained above, we test the optimal throughput performance. We use a diverse set of 400 images from COCO dataset, and the experiment results are presented in Table 3.

Compared with the traditional inference based on a single NPU, the system achieves a throughput acceleration ratio of 2.6X, and besides, due to CPU reaching peak resource utilization, the system's overall throughput fails to achieve the theoretically expected linear increase. For our platform equipped with only one FT2000/4 CPU, the optimal number of MLU220 seems to be 3.

(3) Peak Computing Energy Ratio

In this section, we conducted an evaluation of PCER to gain some insights into constructing a star-style chip cluster.

We have great difficulties in measuring peak computing/energy ratio, since we have no idea about which test program can accurately reflect PCER. A feasible way is to use a program which can exert the computing performance of NPU to the fullest extent possible, and we believe its computing/energy ratio is near to the peak value.

Nowadays, all NPUs are designed to maximize the computing performance of convolution since it is the most important operator and vital to the performance evaluation of NPU. After trying all convolution operators with distinct super-parameters, we find that the convolution operator with NCHW being 4, 256, 103,

Table 4. Peak computing/energy ratio

Test Scenario	PCER (TOPS/W)
1 NPU	0.57
2 NPUs	1.0
3 NPUs	1.43
4 NPUs	0.98

53 and the kernel size being 44 can achieve the highest computing throughput. So, we use the operator after INT8 quantization as the test program for PCER.

For computing PCER, we compute TOPS for the convolution operator manually, count the Average Throughput (AT) of executing the operator, and record the Average Power (AP) consumption of our entire platform using a power meter. The computing/energy ratio is defined as:

$$PCER = TOPS \times AT \div AP \quad (3)$$

Table 4 presents the results. We can find that a maximum 2.5X improvement on PCER is achieved when three MLU220 NPUs are utilized. This is consistent with the scenario when optimal throughput is achieved. So we think if higher throughput or PCER is intended, a higher performance CPU or multiple CPUs will be used instead of current use of one FT2000/4 CPU.

4 Conclusion

This paper introduces a high-throughput deep learning platform for edge intelligence, designed through a hardware/software codesign approach. The platform leverages PCIe bridging technology to expand the computational power of NPUs in low-power manner, while employing hybrid parallel task scheduling strategies in the software to meet the high-throughput demands of edge intelligence scenarios. Our platform was tested using YOLOv5s model for object detection on COCO dataset, and the experimental results demonstrate that, compared to the use of a single NPU, our platform achieves a throughput acceleration ratio of 2.6X and a peak energy efficiency ratio of 2.5X. In addition, when designing a star-shaped chip cluster architecture, it is crucial to pay close attention to the allocation of hardware resources. The issue of load balancing must be carefully addressed to prevent certain stages from becoming bottlenecks in the system.

Acknowledgments. This study is supported by the National Key R&D Program of China (Grant No. 2022YFB4501600).

References

1. NVIDIA Jetson Nano. <https://www.nvidia.cn/autonomous-machines/embedded-systems/jetson-nano/product-development/>
2. NVIDIA Jetson TX2. <https://www.nvidia.cn/autonomous-machines/embedded-systems/jetson-tx2/>
3. NVIDIA Xavier NX. <https://www.nvidia.cn/autonomous-machines/embedded-systems/jetson-xavier-nx/>
4. Kim, H., Ryu, J., Lee, J.: TCCL: Discovering Better Communication Paths for PCIe GPU Clusters. In: Proceedings of 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '24), vol. 3, pp. 999–1015. ACM, New York, NY, USA (2024)
5. NVIDIA DGX-1. <https://www.nvidia.com/en-gb/data-center/dgx-systems/dgx-1/>
6. Xu, D., et al.: Edge intelligence: Empowering intelligence to the edge of network. Proc. IEEE **109**(11), 1778–1837 (2021)
7. NVIDIA’s FasterTransformer. <https://github.com/NVIDIA/DeepLearningExamples/tree/master/FasterTransformer>
8. Rasley, J., Rajbhandari, S., Ruwase, O., He, Y.: DeepSpeed: System optimizations enable training deep learning models with over 100 billion parameters. In: Proceedings of 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '20), pp. 3505–3506. ACM, New York, NY, USA (2020)
9. Yu, G.-I., Jeong, J.S., Kim, G.-W., Kim, S., Chun, B.-G.: Orca: A distributed serving system for transformer-based generative models. In: 16th USENIX Symposium Operating Systems Design and Implementation (OSDI '22), pp. 521–538. Carlsbad, CA, USA (2022)
10. Xu, C., Kong, W., Liu, M., Zhang, M., Li, C., Gong, L.: Hybrid parallel inference for large model on heterogeneous clusters for high throughput. In: 2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS), pp. 549–554. Ocean Flower Island, China (2023)
11. Cambricon’s MLU220. <https://www.cambricon.com/index.php?m=content&c=index&a=lists&catid=57>
12. Phytiun’s FT-2000/4. <https://www.phytium.com.cn/homepage/production/6/>
13. Lielāmars, E., Cvetkovs, A., Novickis, R., Ozols, K.: Infrared image pre-processing and IR/RGB registration with FPGA implementation. Electronics **12**(4), 882 (2023)
14. Anlogic’s EF2L15LG100B. <https://www.datasheets.com/zh-cn/part-details/ef2l15lg100b-anlogic-424558258>



Visual SLAM Based on Point Line Feature Fusion in Dynamic Environments

Benlong Xiang¹, Lichuan Ning^{2(✉)}, and Yuanmin Xie²

¹ Wuhan University of Science and Technology, Wuhan 430081, Hubei, China

² Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering,
Wuhan 430081, Hubei, China
ning.lichuan@wust.edu.cn

Abstract. Aiming at the problems of feature matching error and insufficiency of static features in dynamic environments, we proposed a dynamic point and line SLAM system based on point and line feature fusion and lightweight improved instance segmentation network. Three-dimensional line equations for line features were fitted by principal component analysis and spatial validation model are constructed to correct depth outliers. Secondly, the YOLOv8-seg backbone network is reconstructed by PP-LCNet and integrated a lightweight attention mechanism to improve the network inference speed while maintaining the detection accuracy. Thirdly, we integrated line segment midpoint and endpoint projection error, combined with the Gaussian residual model to construct a dynamic probability prediction mechanism for point and line features. Validated by the TUM RGB-D dataset, this system outperforms the benchmark algorithms such as ORB-SLAM3 in terms of absolute trajectory error and relative position error, in which the maximum reduction of ATE reaches 94.35%, which fully verifies the robustness of the proposed method and the enhancement of the positioning accuracy in the dynamic environment.

Keywords: Dynamic visual SLAM · Point line feature fusion · Spatial validation model · Semantic segmentation

1 Introduction

As the core technology of autonomous navigation, simultaneous localization and mapping (SLAM) provides intelligent systems with the capabilities of environment perception by constructing environment maps and estimating carrier positions in real time. Recent years, SLAM technology has made breakthroughs in both theoretical research and engineering applications, and its applications have covered many fields such as intelligent driving and service robot navigation [1]. Visual sensors occupy an important position in SLAM research by virtue of their cost advantage, rich texture information. Visual-based SLAM solutions have formed two major technology routes, the direct method represented by SVO and DSO [2], and the feature-point-based optimization method represented by ORB series [3, 4].

Although existing SLAM frameworks show superior performance in static environments, their theoretical foundation based on the assumption of static environments leads to limitations in scenarios with dynamic interferences. To address the above problems, researchers have proposed strategies based on semantic segmentation, such as DynPL-SLAM [5] and DS-SLAM [6], to improve system robustness by detecting and rejecting dynamic region features. The instance segmentation network has high computational complexity and is difficult to meet real-time constraints. Therefore, designing lightweight semantic segment models for fast detection of moving targets have become the key break-through directions of dynamic visual SLAM.

The application of visual SLAM algorithms fusing multi-feature information in structured environments has received academic attention, especially in indoor scenes rich in line and surface geometric features. Typical systems such as ORB-LINE-SLAM [7] and DynPL-SLAM optimize the pose by constructing a reprojection error model for line features. However, the approach direct derivation of 3D line equations from the depth values of the endpoints ignores whether the middle part of the line satisfies the constraints of the fitted line equation, which leads to the spatial inconsistency between the constructed geometrical maps and the real texture of the scene.

To solve the above problems, this paper proposes a visual SLAM system based on point-line feature fusion in dynamic environments using the yolov8seg instance segmentation network. The main contributions include:

- (1) Constructing a spatial validation model for line features, combining the PCA algorithm to fit the line equations in the camera coordinate system, and correcting or compensating for the depth of pixel points on the line that have too large difference between the true value and the estimated value, or that do not have a depth value.
- (2) A lightweight improved yolov8seg instance segmentation algorithm is proposed in the front-end of the system, which uses a neural network PP-LCNet to reconstruct the yolov8seg backbone network layer, and performs motion attribute prediction of the point and line features based on Gaussian model.
- (3) The proposed lightweight improved Yolov8Seg segmentation network is trained on the COCO2017 dataset, and the robustness and accuracy of the proposed SLAM system for point and line feature fusion in dynamic environments are verified on TUM RGBD.

2 Related Work

2.1 Visual SLAM Based on Point and Line Features

Recent years, breakthrough progress has been made in visual SLAM systems based on point and line feature fusion. Line features provide geometric constraints in low-texture scenarios, while point features maintain high-precision matching in high-texture regions, and the two synergize to enhance the system's environmental adaptability. Zou et al. [8] constructed a joint optimization model of point and line, fused linear reprojection error and point feature error to construct a hybrid cost function, and combined Gauss-Newton optimization with chi-square test to achieve robust pose solution. Si et al. [9] designed an edge-constrained line detector, which improves the line features' integrity and the anti-jamming ability by suppressing gradient noise.

Although the current study effectively improves the localization robustness of low-texture scenes through point-line feature fusion, the high computational complexity and matching uncertainty of line features still constitute the core constraints on the real-time performance and accuracy of the system. How to construct a lightweight line feature representation model and design a fast match algorithm with geometric invariance is still a key scientific problem in this field.

2.2 Visual SLAM in Dynamic Environment

Traditional visual SLAM systems rejection dynamic feature based on static environment assumptions, which are prone to feature degradation triggering localization drift in scenes occupied by dynamic targets, and academics carry out a series of studies on dynamic scenes [10, 11]. Gong et al. [12] design an RGB-D SLAM system based on YOLO-Fastest, constructing effective geometric constraints by enhancing the density of the static point. Although the above methods suppress dynamic interference through the semantic-geometric fusion framework, they still face the risk of degradation due to insufficient constraints after feature rejection. Yuan et al. [13] propose a Bayesian dynamic feature tracking framework based on polar constraints to achieve decoupling of dynamic and static features, to improve the accuracy of the system.

However, the existing methods have not yet effectively solved the problem of geometric constraint breakage caused by dynamic occlusion and endpoint disappearance of line features. The current dynamic SLAM research still needs to break through the problem of feature continuity preservation in highly dynamic scenarios, and achieve engineering through lightweight network design and edge computing deployment.

3 Method

The system uses RGBD depth camera to capture image data, in the front-end tracking thread, three sub-threads are turned on to extract point and line features and object masks in parallel (see Fig. 1). Point features are extracted using raw FAST corner points, while line features are extracted using the ED lines algorithm. The extracted line features are fused and culled, and subsequently the inputs of the line features are fed into a spatial validation model to compensate and correct for missing depth values or anomalies at certain pixel points of the line features. Finally, a Gaussian model is used to calculate the dynamic probabilities to predict the motion properties of point and line features, and the dynamic features are eliminated for fine bit pose estimation.

3.1 Spatial Validation Model for Line Features

Since the line feature detection algorithm relies on image edge information, in structured scenes (e.g., edges of objects such as desktops), the coordinates of the endpoints of the extracted lines tend to exhibit spatial inconsistencies, and their pixel positions may be localized to the adjacent background regions (e.g., walls or floors). This geometric misalignment leads to deviations in the depth values of the endpoints, affecting the accuracy of the system's pose estimation and geometric representation of 3D line features. To solve

this problem, a spatial inspection model is proposed, which uses the PCA algorithm to fit the 3D line equations after restoring the line's pixel points to the camera coordinate system. Assuming that the point set of the line feature is $X = [x_1, x_2, \dots, x_n]$, the data are first decentred using Eq. 1.

$$X = X - \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

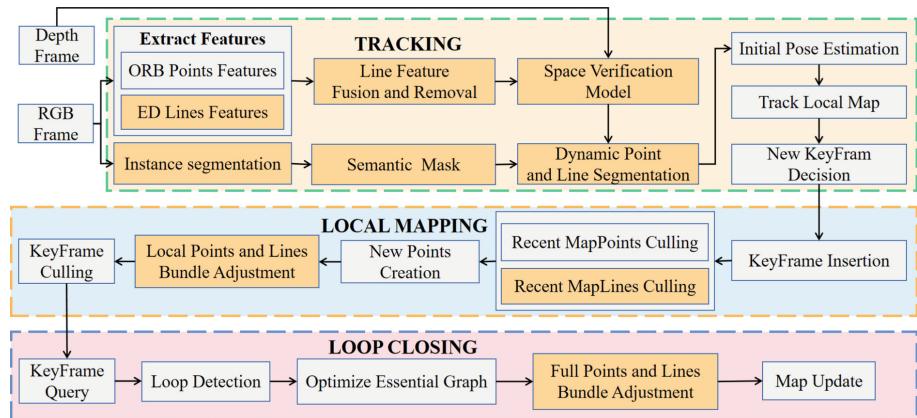


Fig. 1. Overall system framework diagram. Grey white represents the original ORBSLAM3 module, and light yellow represents the newly added or modified module in this article.

The covariance matrix is then calculated using Eq. 2. The eigenvalue decomposition of the covariance matrix is performed to obtain eigenmatrix, which is arranged in columns according to the eigenvalues from the largest to the smallest, where the eigenvector corresponding to the largest eigenvalue is the direction vector of the line.

$$C = \frac{1}{n} XX^T \quad (2)$$

Since the initial fitted line point set may contain more outliers, we use the initial estimated line equation to calculate the distance from each point to the line in the point set, and remove those outliers with too large a distance and then use the PCA algorithm to fit again to obtain the line equation as shown in Eq. 3.

$$\frac{x - x_0}{m} = \frac{y - y_0}{p} = \frac{z - z_0}{q} \quad (3)$$

where $[m, p, q]^T$ is the direction vector of the line and $[x_0, y_0, z_0]^T$ is the midpoint of the line, and the average distance σ to the line of all the interior points in the fitting process is obtained, and further the normalized coordinates of a pixel coordinate point on the

image under the camera coordinate system can be obtained from Eq. 4.

$$\begin{pmatrix} X' \\ Y' \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \stackrel{\text{def}}{=} K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (4)$$

where $P' = [X', Y', 1]^T$ is the normalised coordinate, multiply it by the depth value Z to get the coordinate of the pixel in the camera coordinate system, combined with the line Eq. 3 can be solved to get the fitted depth value of pixels in the depth map without depth value Z , the specific calculation formula is shown in Eq. 5.

$$Z = \frac{qx_0 - mz_0}{qX' - m} = \frac{qy_0 - pz_0}{qY' - p} \quad (5)$$

The pixel points with a difference of less than 3σ between the depth value on the depth map and the depth value fitted by the 3D linear equation are considered to have accurate depth values. Otherwise, replaced by using the depth value of the last adjacent pixel point, plus the linear increment of the depth value of the two adjacent pixel points on the fitted three-dimensional line equation.

An experimental validation is carried out in the TUM RGB-D dataset fr1/room sequence. The depth anomaly in the reconstruction results that have not been validated by spatial validation models leads to some map lines showing disordered distribution in 3D space, which seriously deviates from the geometrical constraints and texture features of the real scene (see Fig. 2(a)). After the optimization of the model, the map lines significantly converge to reasonable spatial locations, and highly consistent with the scene geometry (see Fig. 2(b)).

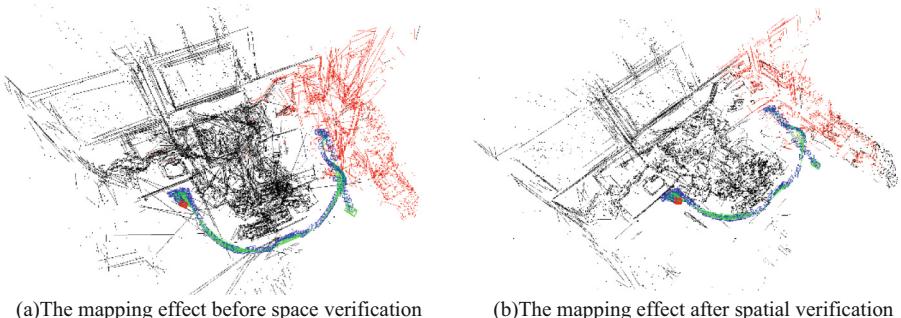


Fig. 2. The mapping effect before and after line feature space verification

3.2 Lightweight Improvement of Yolov8seg Based on PP-LCNet

Semantic information provides very important priori information when dealing with dynamic objects. Although yolov8n-seg is the most lightweight model, the inference speed is very fast but still can't meet the real-time requirements of SLAM system, need to improve its lightweight. we based on PP-LCNet reconstruction of the backbone

network layer, combined with the MLCA attention mechanism, the original instance segmentation network for lightweight improvement (see Fig. 3).

Specifically, 7-layer depth-separable convolutional DWConv is used in the early layer of the network for 4 times down-sampling to extract multiscale features, and the feature maps at P3, P4 levels are input into the Neck network layer for multiscale feature fusion. The activation function adopts H-Swish instead of ReLU, as shown in Eq. 6, which enhances the model representation. A large 5x5 convolution kernel is used in the deeper layers of the model to enhance the feeling field, but it will increase the amount of computation, balancing speed and accuracy.

$$H - \text{Swish}[x] = x \frac{\text{ReLU}6(x + 3)}{6} \quad (6)$$

Since light-weighting optimizes performance by reducing the number of parameters, but may lead to a decrease in feature expressiveness, it is necessary to add an attention mechanism to construct the important information. Using the MLCA attention mechanism performs local average pooling on the feature maps input to the Neck network layer, and processes the feature maps in two directions, one is to process the rescaled feature maps using a 1×1 convolutional network, and the other is to perform global average pooling on feature maps, and then perform a post-processing inverse pooling operation. Finally, the attention weights can be obtained by summing the two and then performing the inverse pooling operation.

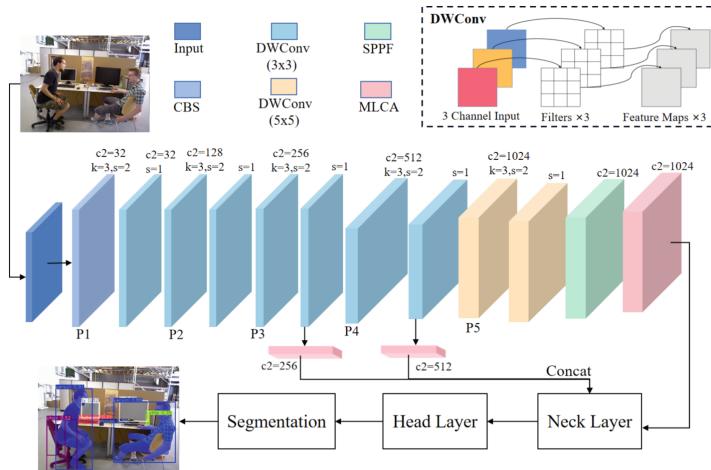


Fig. 3. Lightweight improved instance segmentation network framework diagram

3.3 Prediction of Point Line Feature Motion Attributes

Static feature points outside the mask of the moving target are screened from the previous frame, feature matching is constructed by optical flow tracking, and the pose between adjacent frames is solved by the essential matrix. The feature points inside the mask are projected to the current frame, and a Gaussian probability model is constructed based on

the reprojection errors to quantify the dynamic probability. For line features, the distance from the endpoint to the matching line is used to build a residual model, and combined with the midpoint distance between the two lines (see Eq. 7).

$$\begin{cases} \varepsilon_i^P = |p'_i - \pi(T_{cw}P_i)| \\ \varepsilon_i^L = |p_i^m - \pi(T_{cw}P_i^m)| + \frac{\pi(T_{cw}P_i^{s,e})^T l_i}{\sqrt{l_1^2 + l_2^2}} \end{cases} \quad (7)$$

where ε_i^P and ε_i^L represent the reprojection error of the point and line feature, respectively, p'_i is the pixel coordinate of the matched point, P_i is the 3D coordinate of the point in the world coordinate system, and l_i is the normal vector of the line matched to the current frame. We assume that the reprojection errors of the point and line features obey a Gaussian distribution with zero mean, Δ_p^2 and Δ_l^2 variance, respectively, and Δ_p and Δ_l denote the mean of the reprojection errors of the static point and line features in the initial pose estimation, respectively. The formula for calculating the dynamic probability of each point and line feature is shown in Eq. 8.

$$\begin{cases} \varepsilon_i^P = |p'_i - \pi(T_{cw}P_i)| \\ \varepsilon_i^L = |p_i^m - \pi(T_{cw}P_i^m)| + \frac{\pi(T_{cw}P_i^{s,e})^T l_i}{\sqrt{l_1^2 + l_2^2}} \end{cases} \quad (8)$$

where x_i represents the reprojection error of the i th point or line feature, i.e., ε_i^P or ε_i^L as calculated above, and $P(x_i)$ denotes the dynamic probability. The records with dynamic probability greater than 0.85 are regarded as high dynamic features, those with dynamic probability greater than 0.6 less than 0.85 are regarded as low dynamic features, and those less than 0.6 are regarded as static features.

$$Class(P(x_i)) \begin{cases} \text{High Dynamic, } P(x_i) > 0.85 \\ \text{Low Dynamic } 0.60 < P(x_i) < 0.85 \\ \text{Static } P(x_i) < 0.60 \end{cases} \quad (9)$$

4 Experiment

The experimental platform is configured with Intel Core i5-12490F processor and 32 GB RAM, and equipped with Ubuntu 20.04 operating system. Among them, the lightweight improved YOLOv8-seg instance segmentation network is trained on NVIDIA RTX 4060 GPUs, while the SLAM runtime phase and the object mask inference tasks are executed on the CPU side, aiming to adapt to the future engineering deployment requirements on resource-constrained embedded devices.

4.1 YOLOv8seg Lightweight Experiment

To verify the instance segmentation performance of proposed lightweight improved YOLOv8seg, we trained it on the COCO2017 dataset. The training batch size was 16 and a total of 400 rounds were trained. The experimental results show that the model improved by PP-LCNet can correctly segment the object mask. However, the contour of the segmented object mask is not fine enough. After being strengthened by the MLCA attention mechanism, the contour can be made finer (see Fig. 4).

4.2 Line Feature Experiment

In order to verify the effectiveness of the proposed line feature method, four static scene sequences (fr1/desk, fr1/xyz, fr2/desk, and fr2/xyz) in the TUM RGB-D dataset are selected to carry out comparative experiments. Compared with the classical point-line fusion SLAM frameworks ORB-LINE-SLAM and ORB-SLAM3, the algorithm in this paper achieves the highest optimization enhancement of 31.63% in absolute trajectory error (ATE) (see Table 1), which indicates that the global constraints of the line features effectively inhibit the trajectory cumulative drift. Meanwhile, the relative position error (RPE) is also improved by up to 18.57% (see Table 2), reflecting that the line features have a certain optimization effect on the local motion estimation.

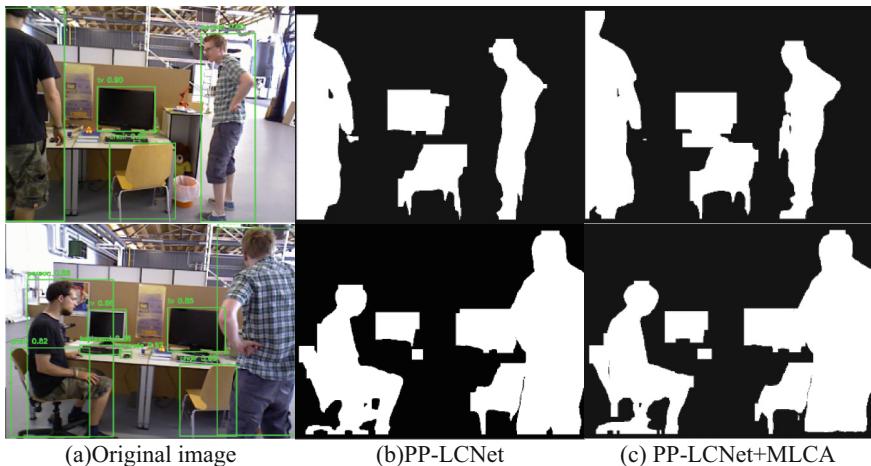


Fig. 4. The effect of YOLOv8seg after lightweight improvement

Table 1. Result of absolute trajectory error (ATE).

	ORB-SLAM3		ORB-LINE-SLAM		Ours		Improved	
	RMSE	STD	RMSE	STD	RMSE	STD	RMSE	STD
fr1/xyz	0.0101	0.0052	0.0102	0.0054	0.0090	0.0048	10.89%	7.69%
fr2/desk	0.0098	0.0044	0.0306	0.0166	0.0067	0.0027	31.63%	38.64%
fr2/xyz	0.0035	0.0017	0.0044	0.0021	0.0033	0.0015	5.71%	11.76%

Note: The algorithm data with the best results in the table is highlighted in bold

Table 2. Results of relative pose error (RPE).

	ORB-SLAM3		ORB-LINE-SLAM		Ours		Improved	
	RMSE	STD	RMSE	STD	RMSE	STD	RMSE	STD
fr1/xyz	0.0096	0.0051	0.0101	0.0056	0.0085	0.0046	11.46%	9.80%
fr2/desk	0.0070	0.0042	0.0075	0.0043	0.0057	0.0035	18.57%	16.67%
fr2/xyz	0.0053	0.0030	0.0054	0.0030	0.0054	0.0032	-1.89%	-6.67%

Note: The algorithm data with the best results in the table is highlighted in bold

4.3 Dynamic SLAM Experiment

To verify the performance of the proposed method in dynamic scenarios, six typical sequences in the TUM RGB-D dataset are selected for experiments. Compared with dynamic SLAM algorithm Dyna-SLAM, the proposed method achieved up to 94.31% improvement in absolute trajectory error, and up to 91.45% improvement in low-dynamic scenes (see Table 3). The relative pose error is improved to some extent (see Table 4), indicating that the dynamic feature suppression mechanism proposed effectively eliminates the interference of moving objects on pose estimation, thus improving the positioning accuracy of the system in dynamic environments.

Table 3. Result of absolute trajectory error (ATE).

	ORB-SLAM3		Dyna-SLAM		Ours		Improved	
	RMSE	STD	RMSE	STD	RMSE	STD	RMSE	STD
fr3/w/half	0.1967	0.0769	0.0296	0.0157	0.0220	0.0123	88.82%	84.01%
fr3/w/rpy	0.1405	0.0674	0.0354	0.0191	0.0345	0.0193	75.44%	71.36%
fr3/w/static	0.0175	0.0094	0.0068	0.0032	0.0137	0.0075	21.71%	20.21%
fr3/w/xyz	0.2567	0.1125	0.0164	0.0086	0.0146	0.0071	94.31%	93.69%
fr3/s/half	0.0623	0.0314	0.0186	0.0086	0.0194	0.0105	68.86%	66.56%
fr3/s/xyz	0.0117	0.0058	0.0127	0.0060	0.0010	0.0057	91.45%	1.72%

Note: The algorithm data with the best results in the table is highlighted in bold

Table 4. Results of relative pose error (RPE).

	ORB-SLAM3		Dyna-SLAM		Ours		Improved	
	RMSE	STD	RMSE	STD	RMSE	STD	RMSE	STD
fr3/w/half	0.0233	0.0165	0.0284	0.0149	0.0144	0.0087	38.20%	47.27%
fr3/w/rpy	0.0125	0.0092	0.0448	0.0262	0.0147	0.0135	-17.60%	-46.74%

(continued)

Table 4. (*continued*)

	ORB-SLAM3		Dyna-SLAM		Ours		Improved	
	RMSE	STD	RMSE	STD	RMSE	STD	RMSE	STD
fr3/w/static	0.0063	0.0019	0.0089	0.0044	0.0051	0.0022	19.05%	-15.79%
fr3/w/xyz	0.0323	0.0260	0.0217	0.0119	0.0121	0.0070	62.54%	73.08%
fr3/s/half	0.0096	0.0051	0.0239	0.0120	0.0115	0.0065	-19.79%	-27.45%
fr3/s/xyz	0.0136	0.0086	0.0317	0.0159	0.0098	0.0052	27.94%	39.53%

Note: The algorithm data with the best results in the table is highlighted in bold

4.4 Real-Time Analysis

Real-time performance is one of the most important indicators for evaluating SLAM performance. We analyzed the time of each module of algorithm, including the instance segmentation module (SEG), point and line feature extraction module (PLE), spatial validation module (SCM), and the entire tracking thread module (ALL). The results show that proposed algorithm takes an average of 38ms to process an image, which meets the real-time requirements of SLAM system (see Table 5).

Table 5. Time comparison of each module(ms).

	SEG	PLE	SCM	ALL
fr2/xyz	/	12.86	2.96	26.72
fr3/w/half	27.45	13.75	2.42	39.23
fr3/s/xyz	28.03	13.24	3.07	38.64

In addition, we also conduct time analysis with classic dynamic SLAM algorithms (DS-SLAM [6]) and state-of-the-art dynamic SLAM algorithms (Strong-SLAM [10]), as well as state-of-the-art dynamic point-line SLAM algorithms (DPL-SLAM [11], DynPL-SLAM [5]). Experimental results show that proposed algorithm is better than traditional dynamic SLAM, and also has competitive advantages in real-time performance compared with the state-of-the-art dynamic SLAM (see Table 6).

Table 6. Time comparison of different algorithms(ms).

Algorithm	DS-SLAM	Strong-SLAM	DPL-SLAM	DynPL-SLAM	Ours
Time	67.07	34.06	36.3	156	38.85

5 Conclusion

This article proposed a SLAM system based on point-line feature fusion and lightweight improved Yolov8-seg network. A spatial validation model is designed, and the depth outliers are corrected by fitting 3D linear equations through the PCA algorithm. By reconstructing the backbone network of Yolov8-seg using PP-LCNet lightweight network and integrating MLCA attention mechanism, real-time semantic segmentation was achieved, and combined with a Gaussian residual model to achieve dynamic probabilistic prediction. And experiments on the TUM dynamic dataset show that its ATE is improved 94.31%, and the RPE is improved by 62.54%.

References

1. Wang, X., Fan, X., Shi, P., et al.: An overview of key SLAM technologies for underwater scenes. *Remote Sens.* **15**(10), 2496 (2023)
2. Engel, J., Koltun, V., Cremers, D.: Direct sparse odometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(3), 611–625 (2017)
3. Mur-Artal, R., Tardós, J.D.: ORB-SLAM2: An open-source slam system for monocular, stereo, and RGBD cameras. *IEEE Trans. Rob.* **33**(5), 1255–1262 (2017)
4. Campos, C., Elvira, R., Rodríguez, J.J.G., et al.: ORB-SLAM3: an accurate open-source library for visual, visual–inertial, and multimap slam. *IEEE Trans. Rob.* **37**(6), 1874–1890 (2021)
5. Zhang, B., Dong, Y., Zhao, Y., et al.: DynPL-SLAM: a robust stereo visual slam system for dynamic scenes using points and lines. *IEEE Trans. Intell. Veh.* (2024)
6. Yu, C., Liu, Z., Liu, X.J., et al.: DS-SLAM: a semantic visual SLAM towards dynamic environments. In: 31th IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1168–1174. IEEE, Madrid (2018)
7. Alamanos, I., Tzafestas, C.: ORB-LINE-SLAM: an open-source stereo visual SLAM system with point and line features. *Author Preprints* (2023)
8. Zou, Y., Eldemiry, A., Li, Y., et al.: Robust RGB-D SLAM using point and line features for low textured scene. *Sensors* **20**(17), 4984 (2020)
9. Si, H., Yu, H., Chen, K., et al.: Point-line visual inertial odometry with optimized line feature processing. *IEEE Trans. Instrum. Meas.* **73**, 5027113 (2024)
10. Huang, W., Zou, C., Yun, J., et al.: Strong-SLAM: real-time RGB-D visual SLAM in dynamic environments based on StrongSORT. *Meas. Sci. Technol.* **35**(12), 126309 (2024)
11. Lin, Z., Zhang, Q., Tian, Z., et al.: DPL-SLAM: enhancing dynamic point-line SLAM through dense semantic methods. *IEEE Sens. J.* **9**(24), 14596–14607 (2024)
12. Gong, C., Sun, Y., Zou, C., et al.: Real-time visual SLAM based YOLO-fastest for dynamic scenes. *Meas. Sci. Technol.* **35**(5), 056305 (2024)
13. Yuan, C., Xu, Y., Zhou, Q.: PLDS-SLAM: point and line features SLAM in dynamic environment. *Remote Sens.* **15**(7), 1893 (2023)



Cooperative Path Planning for Multi-UAVs Using MASAC with SEAttention and Prioritized Experience Replay

Wei Xia, Boxiang Li, Jinliang Zhu, and Denghao Pang^(✉)

School of Internet, Anhui University, Hefei 230039, People's Republic of China
pangdenghao@ahu.edu.cn

Abstract. The advancement in unmanned aerial vehicle (UAV) technology has increasingly drawn the attention of researchers to cooperative operations among multi-UAVs, with path planning standing out as a critical challenge. This paper proposes a method leveraging the multi-agent soft actor-critic algorithm enhanced by squeeze-and-excitation attention and prioritized experience replay (MASAC-SEPR) to address the heterogeneous multi-UAVs path planning problem. The model is formulated as a partially observable Markov decision process (POMDP), integrating an advanced algorithmic framework that incorporates the Squeeze-and-Excitation Attention (SEAttention) module, actor and critic networks. Additionally, it employs a prioritized experience replay memory mechanism to ensure efficient training. Simulation results demonstrate that this model surpasses traditional approaches in terms of performance, convergence speed, and adaptability to complex environments. It facilitates efficient and accurate cooperative path planning for heterogeneous multi-UAVs, providing robust algorithmic support for practical applications.

Keywords: cooperative path planning · multi-UAVs · deep reinforcement learning · SEAttention · prioritized experience replay

1 Introduction

Unmanned aerial vehicles (UAVs), initially developed as military reconnaissance tools in the early 20th century [1, 2], have since evolved into indispensable assets across various industries. Enabled by advancements in micro-electro-mechanical systems (MEMS), communication technologies, and navigation systems, UAVs have become more compact, cost-effective, and capable, thereby significantly expanding their civilian applications [3, 4]. Cooperative UAVs operations, which capitalize on information sharing and enhanced operational efficiency, are now extensively utilized in logistics, inspection, security, and intelligent manufacturing sectors. Nevertheless, path planning for cooperative UAV missions remains

a formidable challenge due to the complexity and dynamism of operational environments, as well as the increasing size of UAV swarms [5].

Traditional path planning algorithms, such as Dijkstra's and A*, encounter significant challenges in managing exponentially increasing computational loads in complex scenarios [6]. To address the limitations of classical algorithms and mitigate computational complexity, heuristic algorithms, including improved genetic algorithms (GA) [7], particle swarm optimization (PSO) [8], and ant colony optimization (ACO) [9], have been employed to enhance efficiency and adaptability. However, in dynamic environments, existing heuristic algorithms still face difficulties in rapidly adjusting paths to effectively respond to new situations [10].

In recent years, deep reinforcement learning (DRL) has been increasingly utilized for dynamic path planning in UAVs [11, 12]. By integrating the perception capabilities of deep learning with the decision-making mechanisms of reinforcement learning, DRL enables UAVs to adaptively adjust their trajectories through interactions with the environment [13]. Algorithms such as double deep Q-network (DDQN) [14], proximal policy optimization (PPO) [15], and deep deterministic policy gradient (DDPG) [16] have been proposed to enhance the performance of path planning. However, the majority of existing studies predominantly focus on single-UAV scenarios and fall short in addressing multi-UAVs cooperation, particularly in complex and dynamically changing environments.

To address these challenges, the multi-agent soft actor-critic (MASAC) algorithm has been proposed for cooperative path planning of multi-UAVs [17]. MASAC models each UAV as an independent agent that learns cooperative strategies by interacting with both other agents and the environment, thereby enhancing mission execution efficiency. Despite its promising performance, MASAC encounters limitations stemming from complex inter-agent interactions and insufficient experience utilization, which lead to slow convergence, particularly in complex tasks and environments.

In this paper, we develop a novel multi-agent soft actor-critic with squeeze-and-excitation attention and prioritized experience replay (MASAC-SEPR) mechanism for achieving efficient multi-UAVs path planning. The key contributions of this work are summarized as follows: (1) A novel MASAC-SEPR mechanism is introduced to tackle the path planning challenge in heterogeneous multi-UAVs. (2) The Squeeze-and-Excitation Attention (SEAttention) module is incorporated into the framework, allowing the system to automatically emphasize critical environmental features while suppressing irrelevant information, thus significantly reducing computational complexity. (3) A prioritized experience replay memory is designed, where experiences derived from the interaction between agents and the environment are assigned priorities to facilitate more effective learning.

2 Problem Formulation

2.1 Basic Decision Model

Partially observable Markov decision process (POMDP) is a framework for decision-making under uncertainty, where the environment's state is only partially observable. It extends the standard Markov decision process (MDP). The fundamental decision model for UAVs is formulated as a POMDP, represented by a 7-tuple $S, A, T, R, \Omega, O, \gamma$, where S denotes the set of all possible UAV states s_n , with N being the total number of UAVs; A represents the set of all possible UAV actions a_n ; the probability of transitioning from state s to state s' upon taking action a is indicated by the state transition probability $T(s'|s, a)$; the immediate reward obtained when executing action a in state s is specified by the reward function $R(s, a)$; the observation set Ω encompasses all possible UAV observations z_n ; the observation probability $O(z|s, a)$ quantifies the probability of observing z given that action a is taken in state s ; and $\gamma \in [0, 1]$ signifies the discount factor for future rewards. Within this framework, UAVs can make informed decisions based on partially observable information, interact with the environment via various learning strategies, and thereby achieve collaborative path planning in complex scenarios.

Assumption 1. The dynamics of UAV motion are governed by a set of kinematic equations, which describe the evolution of the system over discrete time intervals. Specifically, for the n -th UAV, the equations of motion from time t to $t + 1$ are defined as:

$$\begin{cases} x_n(t+1) = x_n(t) + v_n(t) \cos \psi_n(t) \Delta t, \\ y_n(t+1) = y_n(t) + v_n(t) \sin \psi_n(t) \Delta t, \\ \psi_n(t+1) = \psi_n(t) + \omega_n(t) \Delta t, \\ v_n(t+1) = v_n(t) + u_n(t) \Delta t, \end{cases} \quad (1)$$

where all variables, namely position (x_n, y_n) , heading angle ψ_n , velocity v_n , angular velocity ω_n , and acceleration u_n are bounded within their respective operational limits. These variables are updated iteratively over discrete time steps Δt . By incorporating these equations into the path-planning framework, the algorithm ensures compliance with physical constraints, thereby improving both the precision and practicality of the generated trajectories.

Assumption 2. The presence of obstacles is represented as standard circular regions with a radius denoted by R_g . A collision is deemed to have occurred only when the Euclidean metric separating the UAV from obstacles is reduced to below the threshold R_g . Consequently, the condition for avoiding collision can be expressed as:

$$\sqrt{(x_n(t) - x_g)^2 + (y_n(t) - y_g)^2} \geq R_g, \quad (2)$$

where $(x_n(t), y_n(t))$ and (x_g, y_g) correspond to the instantaneous position of the UAV and the center of the obstacle, respectively.

2.2 Network Update

For the n -th agent, the adjustment of its network parameters is driven by the policy loss functions corresponding to its actor network, critic network, and entropy network. The structural relationships among the Actor, Critic, and entropy networks are as depicted in the “Network Structure” of Fig. 1, and their network update strategy adheres to the update strategy of the traditional Actor-Critic network [17]. Through the computation of these loss functions, the agent is able to efficiently update and refine its network architecture, thereby enhancing its task execution efficiency and decision-making capabilities. Specifically, the actor network generates optimal policies, the critic network evaluates the value of these policies, and the entropy network introduces an entropy regularization term to maintain the exploratory nature of the policies. By synthesizing the loss functions from all three components, the agent can dynamically fine-tune its network parameters, enabling superior performance in complex and dynamic environments.

3 MASAC-SEPR Algorithm

This work presents a new methodology aimed at addressing the deficiencies of traditional path-planning approaches in dynamic scenarios and multi-agent collaboration, thereby guaranteeing the robust and secure performance of collaborative aerial missions.

As illustrated in Fig. 1, the MASAC-SEPR algorithm module operates through a structured workflow.

The actor network receives observation inputs z from N agents (UAVs), generates corresponding actions a , and interacts with the environment to receive rewards R and the subsequent state S' . These experiences, along with the current state S and action set A , are stored in a prioritized experience replay memory. Subsequently, mini-batches of experiences B are sampled from this memory and fed into the respective network structures to compute the associated value estimates and loss functions. The network update module then refines the network architecture by propagating gradients and updating the prioritized experience replay memory on priority weights.

3.1 POMDP Model

In the realm of multi-UAVs cooperative path planning, where environments are inherently complex and information is partially observable, the POMDP model establishes a theoretical framework grounded in a septuple structure. This framework provides a rigorous mathematical foundation for the development of algorithms and decision-making processes.

(1) The POMDP model defines the state set S , action set A , and observation set Ω . For the n -th UAV, the state information is denoted as $s_n = [z_1, z_2, \dots, z_n]$, encapsulating both its own and other UAVs’ observation state information.

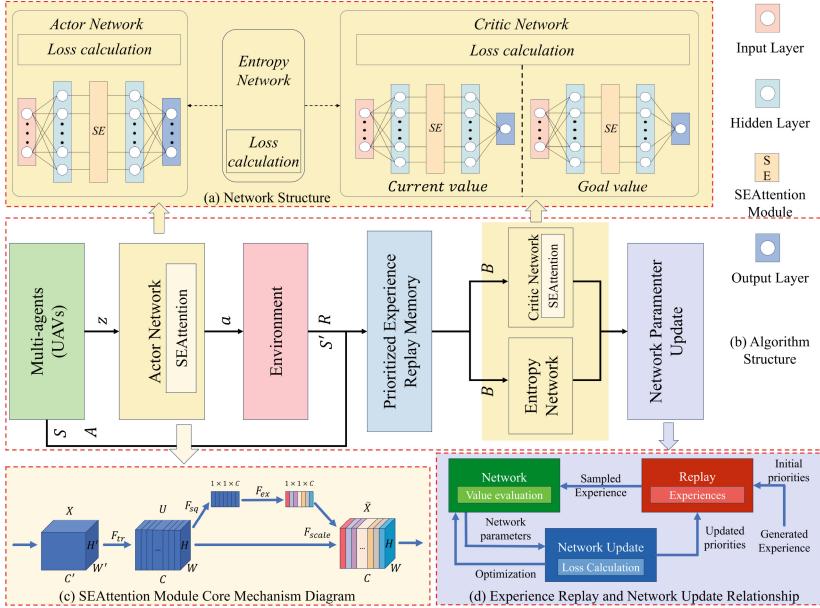


Fig. 1. Algorithm module structure.

Given N UAVs, the collective state reflects the interactions within the system. The action of the n -th UAV is represented as a continuous vector $a_n = [u_n, w_n]^T$, where u_n signifies acceleration, and w_n represents angular velocity.

The observation information for the n -th UAV, define $z_n = [x_n, y_n, v_n, \varpi_n]$. Here, $[x_n, y_n, v_n, \varpi_n]$ denote the flight state quantities of the n -th UAV, while ϖ_n represents its own observation state information. For the leader UAV, $\varpi_n = [x_o, y_o, FLAG]$, where (x_o, y_o) indicates the position of the target point, and $FLAG$ serves as an obstacle flag. Specifically, if the distance between the UAV and any obstacle exceeds twice the obstacle's radius, the system assumes no nearby obstacles ($FLAG = 0$); otherwise, $FLAG = 1$. For follower UAVs, $\varpi_n = [x_l, y_l, v_l]$, where (x_l, y_l) corresponds to the leader UAV's position, and v_l denotes the velocity.

(2) To ensure effective coordination among multi-UAVs, a well-defined reward mechanism is essential. The total reward set R is structured as:

$$R = [R_l, R_{1,f}, R_{2,f}, \dots, R_{N-1,f}], \quad (3)$$

comprising the reward R_l of the leader UAV and the rewards $R_{j,f}$ of $N - 1$ followers, where $j = 1, 2, \dots, N - 1$ indexes the followers. For the leader UAV, the reward function R_l integrates four components: a target-reaching reward r_o , an obstacle avoidance reward r_g , a speed cooperation reward r_s , and a formation distance reward r_d . These components are combined into the total reward

function as follows:

$$R_l = \omega_{1,l} r_o + \omega_{2,l} r_g + \omega_{3,l} r_s + \omega_{4,l} r_d, \quad (4)$$

where $\omega_{1,l}, \omega_{2,l}, \omega_{3,l}, \omega_{4,l}$ are the weights for the leader's reward components. For the $N - 1$ followers, the reward function $R_{j,f}$ for each j is formulated as:

$$R_{j,f} = \omega_{1,f} r_{j,s} + \omega_{2,f} r_{j,d}, \quad (5)$$

where $r_{j,s}$ and $r_{j,d}$ represent the speed cooperation reward and formation distance reward for the j -th follower UAV, respectively. The weights $\omega_{1,f}$ and $\omega_{2,f}$ adjust their relative importance.

Definition 1. The objective reward r_o directs the leader UAV to the target point efficiently, ensuring rapid and precise mission completion. It is defined as:

$$r_o = \begin{cases} 1000, & L_o \leq D_1 \\ -0.002 \times L_o, & L_o > D_1 \end{cases}, \quad (6)$$

where L_o denotes the distance between the leader UAV and the designated target point. If L_o drops below a threshold D_1 , a positive reward of 1000 is given. Otherwise, a penalty proportional to the distance is applied to optimize the trajectory and approach the target.

Definition 2. The obstacle avoidance reward r_g promotes safe navigation by discouraging UAVs from approaching obstacles or colliding. It is quantified as:

$$r_g = \begin{cases} 0, & L > 2R_g \\ -5, & R_g \leq L \leq 2R_g \\ -500, & L < R_g \end{cases}, \quad (7)$$

where L represents the distance between the leader UAV and an obstacle, and R_g the obstacle's radius. If $L > 2R_g$, the UAV is safe ($r_g = 0$). For $R_g \leq L \leq 2R_g$, a moderate penalty of -5 applies. If $L < R_g$ (collision), a severe penalty of -500 is enforced.

Definition 3. The cooperative reward r_s and the formation distance reward r_d ensure the stability and cohesion of the UAV formation. Specifically, r_s promotes synchronization by quantifying velocity coordination.

$$r_{j,s} = \begin{cases} 1, & \text{if } |v_l - v_j| < 1 \\ 0, & \text{if } |v_l - v_j| \geq 1 \end{cases} \quad (8)$$

where $r_{j,s}$ represents the velocity cooperation reward between the leader UAV and the j -th follower UAV, v_l is the leader's velocity and v_j is the j -th follower's velocity. The total velocity cooperation reward r_s is calculated as:

$$r_s = \sum_{j=1}^{N-1} r_{j,s}, \quad (9)$$

A reward of $r_{j,s} = 1$ is given if the absolute speed difference $|v_l - v_j|$ is less than 1 m/s, indicating successful velocity synchronization. Otherwise, $r_{j,s} = 0$.

Definition 4. The formation distance reward r_d is defined as follows:

$$r_{j,d} = \begin{cases} 0, & \text{if } L_{j,\min} \leq L_{jl} \leq L_{j,\max} \\ -0.002 \times L_{jl}, & \text{otherwise} \end{cases} \quad (10)$$

where L_{jl} denotes the distance between the leader and the j -th follower, while $L_{j,\min}$ and $L_{j,\max}$ represent the lower and upper bounds of the desired formation range. The total formation distance reward r_d is computed as:

$$r_d = \sum_{j=1}^{N-1} r_{j,d}, \quad (11)$$

When L_{jl} is within $[L_{j,\min}, L_{j,\max}]$, the j -th follower maintains its position with no penalty ($r_{j,d} = 0$). If L_{jl} falls outside this range, a penalty $r_{j,d} = -0.002 \times L_{jl}$ applies, discouraging deviations from the desired formation.

3.2 SEAttention Module

In complex environments characterized by abundant and diverse information, the SEAttention module is capable of automatically focusing on critical features, such as obstacles and target points, thereby preventing resource wastage. This module not only integrates seamlessly with the network but also adaptively adjusts weights across varying scenarios. As a result, it accurately extracts key information, enhances feature representation, and consequently improves the accuracy and robustness of decision-making processes.

The core mechanism of the SEAttention module is illustrated in Fig. 1. Initially, feature representations associated with the UAV flight environment are extracted, i.e., the input feature map X ($X \in \mathbb{R}^{H' \times W' \times C'}$). Convolutional operations (denoted as Ftr) are then applied, where each layer of the input feature map X undergoes a convolution to produce output feature sub-maps u_c , forming the feature map U ($U \in \mathbb{R}^{H \times W \times C}$), which consists of height H , width W , and the number of channels C . Subsequently, adaptive average pooling (F_{sq}) is utilized to perform compression operations, generating a channel description vector for each channel. This process effectively compresses global spatial information into channel-wise information, accurately capturing the corresponding global distribution while reshaping the feature map from $[H, W, C]$ to $[1, 1, C]$.

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j), \quad (12)$$

where z_c and u_c denote the globally average-pooled feature map and the feature map of the c -th channel, respectively.

After the compression process, the operation enters the excitation stage F_{ex} , where two linear transformations F_1 and F_2 are employed to eliminate redundant

information. Subsequently, ReLU and Sigmoid activation functions are utilized to investigate the intricate relationships among features, thereby generating channel attention weights. Ultimately, the generated channel attention weights are fused with the original input features to produce the final output of the SE module \tilde{X} . This process accomplishes feature re-calibration F_{scale} , thereby enhancing the capability to extract features from complex scenarios.

$$\tilde{X}_c = F_{scale}(u_c, s_c) = s_c \times u_c, \quad (13)$$

where $\tilde{X} = [\tilde{X}_1, \tilde{X}_2, \dots, \tilde{X}_c, \dots, \tilde{X}_C]^T$, \tilde{X}_c represents the re-calibrated feature of the c -th feature sub-map, and s_c represents the weight of the c -th feature sub-map.

3.3 Prioritized Experience Replay Memory

Prioritized experience replay memory is similar to the paper [18]. It assigns priorities to experiences from environment interactions, enabling more efficient learning by focusing on key experiences. This improves data utilization and enhances model performance.

First, parameter settings and initialization are carried out. The prioritized experience replay memory operates based on three essential parameters: memory capacity, batch size, and priority threshold. Two queues, each with a maximum length equal to the memory capacity, are initialized for storing recorded experiences in quadruples (state, action, reward, next state) and the other for storing corresponding priorities. Proper configuration of these parameters ensures an optimal balance between the storage and usage efficiency of experiences.

Secondly, following the storage of experiences, when new state transition experiences are generated, the maximum priority in the stored priority queue is retrieved (if the queue is empty, the maximum priority is initialized to 1.0). Subsequently, the new experience is appended to the stored state transition experience queue, and its corresponding maximum priority is added to the stored priority queue. This storage mechanism ensures that newly generated experiences are promptly recorded and each experience is assigned an initial priority.

Thirdly, experience sampling is conducted. The stored priority queue is processed to compute the sampling probabilities for each experience based on the priority values in the array. A batch of experiences is then randomly sampled from the experience queue according to the specified number of samples. Based on the indices, the corresponding experience samples are retrieved, and their state, action, reward, and next state information are extracted. Through priority-based sampling, critical experiences are selected for learning, thereby enhancing the efficiency of training data utilization.

Finally, priority updates are performed. The new priority of each experience sample in the stored priority queue is determined by the TD error δ , using the formula:

$$\delta = r + \gamma \hat{Q}(S', \pi_\theta(S'')) - Q(S, A), \quad (14)$$

where r represents the current reward, $\hat{Q}(S', \pi_\theta(S'))$ and $Q(S, A)$ represent the value outputs by the critic network based on the next state S' and the current state S with corresponding action strategy A , respectively. The new priority is then calculated as $p_i = |\delta| + \epsilon$ (ϵ is a extremely tiny positive value).

In Fig. 1, the Replay module stores, receives experiences with initial priorities, and samples them for value evaluation by the Network. Simultaneously, the Network Update module receives network parameters and action values from the Network, performs optimization through loss calculation, and updates priorities accordingly.

4 Simulation

4.1 Initialization

The UAV motion is modeled as a 2D plane for simplicity. The simulation environment has a resolution of 700×600 pixels, representing a real-world area of 7000×6000 m. The frame rate is kept constant at 60 FPS to ensure a realistic dynamic environment. After each episode, the simulation resets, and the initial states of all UAVs (position, velocity, heading angle, target points) and obstacle locations are randomly generated to reflect real-world variability.

The heading angle change range for the leader UAV is defined as 0 to 2π radians, with velocity from 100 to 150 m/s, angular velocity from -0.5 to 0.5 rad/s, and acceleration from -3 to 3 m/s. For follower UAVs, the heading angle is also 0 to 2π radians, velocity spans 100 to 250 m/s, angular velocity ranges from -1.0 to 1.0 rad/s, and acceleration varies from -6 to 6 m/s. These parameters, configured within realistic limits as in Algorithm 1, ensure physical constraints are considered during training.

4.2 Simulations Results and Analysis

During the multi-UAVs cooperative training process, the initial phase exhibits substantial fluctuations and predominantly negative rewards for followers, which suggests instability in the early stages of the system. Both leaders and total rewards also undergo significant fluctuations with marked positive and negative variations. As training progresses, although fluctuations persist, an upward trend emerges overall. After approximately 200 episodes, follower rewards stabilize near zero, leader rewards stabilize around 1000, and total rewards stabilize around 400. These results confirm the effectiveness of the algorithm, as the multi-UAVs system successfully identifies the optimal strategy. Visualization (Fig. 2d) illustrates the leader (blue) avoiding obstacles and coordinating a path in formation with followers (green) to reach the target.

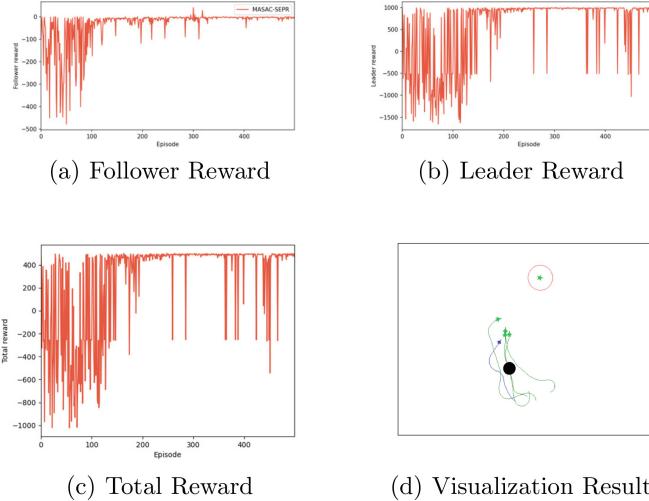
To evaluate the performance of the MASAC-SEPR algorithm in heterogeneous multi-UAVs collaborative path planning tasks, a series of experiments were conducted, comparing it with random strategies, MADDPG, and MASAC

Algorithm 1 MASAC-SEPR Training Algorithm

```

1: Initialize actor, critic, entropy network and prioritized experience replay memory
   D
2: for episode = 1 to MaxEpisode do
3:   Initialize the state  $s$ 
4:   for step = 1 to MaxStep do
5:     for each UAV  $n$  do
6:       According to the state  $s$ , input the observation information  $z_n$  into the
      actor network
7:     end for
8:     Perform Ornstein-Uhlenbeck noise action on the actor network to obtain the
      action  $a_n = [u_n, \omega_n]$ 
9:     Obtain the action set  $A$ , execute step function, and obtain the reward  $r$ ,
      next state  $s'$ 
10:    Store the experience  $(s, a, r, s')$  in  $D$ 
11:    if The sample size in  $D$  is sufficient then
12:      Sample from  $D$  to obtain a batch of experiences
13:      Calculate the action value, TD error, and actor, critic and entropy net-
      work loss function
14:      Update the network parameters and priority
15:    end if
16:    Update the state to  $s'$ 
17:  end for
18: end for

```

**Fig. 2.** Reward curves and visualization results.

algorithms. The experimental results are presented as follows (see Table 1 for details):

Table 1. Performance comparison of MASAC-SEPR algorithm vs. other strategies

Algorithm	Task completion rate (%)	Formation maintenance rate (%)	Flight time(s)	Flight Trajectory(m)	Energy Consumption
Random Strategy	2.2	0.00	934.67	459.28	943.66
MADDPG	81.0	0.67	330.52	196.32	319.48
MASAC	90.0	35.56	249.91	142.47	283.61
MASAC-SEPR	90.0	41.27	187.01	100.34	232.95

In heterogeneous multi-UAV collaborative path planning experiments, the MASAC-SEPR algorithm shows superior performance. It achieves a 90.0% task completion rate, 41.27% formation maintenance rate, 187.01 flight time units, 100.34 trajectory length units, and 232.95 energy consumption units. These results stem from its key mechanisms: the SE mechanism helps UAVs capture features in complex environments, prioritized experience replay accelerates learning, and cooperation rewards ensure stability. Overall, MASAC-SEPR outperforms other algorithms in this domain.

5 Conclusion

In this paper, we propose a heterogeneous multi-UAVs cooperative path planning algorithm based on MASAC. By incorporating SEAttention and prioritized experience replay mechanisms, the proposed algorithm addresses the limitations of conventional approaches in complex environments, thereby substantially improving both planning efficiency and accuracy. Simulation results confirm its superior convergence speed and adaptability. In future work, we will optimize the algorithm by incorporating factors such as adverse weather conditions and interfering signals, which will further bolster the system's robustness and applicability.

References

1. Shen, C., Li, L., Wu, Y., Liu, D.: Research on the capability of the US manned/unmanned autonomous collaborative operations. *Tactical Missile Technol.* **1**, 16–21 (2018)
2. Li, D., Li, M.: Research advance and application prospect of unmanned aerial vehicle remote sensing system. *Geomat. Inf. Sci. Wuhan Univ.* **39**(5), 505–514 (2014)
3. Li, W., Zhao, W., Du, Y.: Large-scale wind turbine blade operational condition monitoring based on UAV and improved YOLOv5 deep learning model. *Mech. Syst. Signal Process* **226**, 112386 (2025)
4. Lee, J.S., Jeong, S.H., Park, G., Kim, Y., Tutumluer, E., Kim, S.Y.: Geotechnical application of unmanned aerial vehicle (UAV) for estimation of ground settlement after filling and compaction. *Transp. Geotech.* **101517** (2025)
5. Gautam, S., Adhikari, B.R., Pudasaini, U., Lamsal, P.: Application of unmanned aerial vehicle (UAVs) for disaster risk reduction and management in Nepal. In: *Communication, Science, Technology, and Innovation in Disaster Risk Management: Recent Trends and Approaches in South Asia*, pp. 151–165. Cham, Springer Nature Switzerland (2025)

6. Dewangan, R.K., Shukla, A., Godfrey, W.W.: Three dimensional path planning using Grey wolf optimizer for UAVs. *Appl. Intell.* **49**(6), 2201–2217 (2019). <https://doi.org/10.1007/s10489-018-1384-y>
7. Yuan, J., Liu, Z., Lian, Y., Chen, L., An, Q., Wang, L., Ma, B.: Global optimization of UAV area coverage path planning based on good point set and genetic algorithm. *Aerospace* **9**(2), 86 (2022)
8. Sonny, A., Yeduri, S.R., Cenkeramaddi, L.R.: Autonomous UAV path planning using modified PSO for UAV-assisted wireless networks. *IEEE Access* **11**, 70353–70367 (2023)
9. Li, B., Qi, X., Yu, B., Liu, L.: Trajectory planning for UAV based on improved ACO algorithm. *IEEE Access* **8**, 2995–3006 (2019)
10. Zhu, D., Yang, S.X.: Bio-inspired neural network-based optimal path planning for UUVs under the effect of ocean currents. *IEEE Trans. Intell. Veh.* **7**(2), 231–239 (2021)
11. Wang, Y., Gao, Z., Zhang, J., Cao, X., Zheng, D., Gao, Y., Ng, D.W.K., Di Renzo, M.: Trajectory design for UAV-based internet of things data collection: A deep reinforcement learning approach. *IEEE Internet Things J.* **9**(5), 3899–3912 (2021)
12. Oubbat, O.S., Atiquzzaman, M., Baz, A., Alhakami, H., Ben-Othman, J.: Dispatch of UAVs for urban vehicular networks: A deep reinforcement learning approach. *IEEE Trans. Veh. Technol.* **70**(12), 13174–13189 (2021)
13. Xin, B., He, C.: DRL-based improvement for autonomous UAV motion path planning in unknown environments. In: 2022 7th International Conference on Control and Robotics Engineering (ICCRE), pp. 102–105. IEEE (2022)
14. Chu, Z., Wang, F., Lei, T., Luo, C.: Path planning based on deep reinforcement learning for autonomous underwater vehicles under ocean current disturbance. *IEEE Trans. Intell. Veh.* **8**(1), 108–120 (2022)
15. Qi, C., Wu, C., Lei, L., Li, X., Cong, P.: UAV path planning based on the improved PPO algorithm. In: 2022 Asia Conference on Advanced Robotics, Automation, and Control Engineering (ARACE), pp. 193–199 (2022)
16. Lan, W., Jin, X., Chang, X., Wang, T., Zhou, H., Tian, W., Zhou, L.: Path planning for underwater gliders in time-varying ocean current using deep reinforcement learning. *Ocean Eng.* **262**, 112226 (2022)
17. Haarnoja, T., Zhou, A., Abbeel, P., Levine, S.: Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International Conference on Machine Learning, pp. 1861–1870. PMLR (2018)
18. Yamani, H., Xing, Y., Ong, L.V.C., MacDonald, B.A., Williams, H.: Reward prediction error prioritisation in experience replay: The RPE-PER method. ArXiv Preprint ArXiv. **2501**, 18093 (2025)



Robust and Natural Bipedal Robot Locomotion Control in Complex Terrains Using Enhanced Adversarial Motion Priors

Hanting Li, Kang An^(✉), and Yaqing Song

Shanghai Normal University, 100 Haisi Road, Shanghai 201400, P.R. China
{ankang,yqsong}@shnu.edu.cn

Abstract. Stable and natural locomotion of bipedal robots in complex terrains provides the foundation for their application in real-world scenarios such as rescue and exploration. Robots trained using traditional deep reinforcement learning methods, for instance Proximal Policy Optimization (PPO), often exhibit high robustness owing to the use of standard network structures and conventional reward functions; however, the naturalness of the gaits generated within complex terrain scenarios is typically suboptimal. To this end, this paper proposes an enhanced PPO-Adversarial Motion Priors (AMP) framework specifically tailored for complex terrains. This framework enhances the system’s learning capability by refining the network structures of the PPO policy and value networks, as well as the AMP discriminator. It integrates terrain height detectors and contact force sensors, thereby expanding the observation space to augment the system’s environmental perception capabilities. Concurrently, this paper designs a novel reward function and modifies the self-collision condition from a traditional penalty term to an episode termination condition, aiming to encourage safer behaviors. During training, PPO-Clip and KL divergence constraints are incorporated to prevent excessive magnitudes in policy optimization steps. Experimental results based on the Isaac Lab simulation platform demonstrate that a model employing a deep network structure (six layers, approx. 2.79 million parameters) exhibits exceptional walking robustness and highly natural gaits closely approximating human locomotion across various complex terrains, even when the task reward comprises solely a survival reward with a weight accounting for merely 0.2. In contrast, robots trained using traditional PPO alone typically display only distinctly mechanical gaits.

Keywords: Bipedal Robot · Complex Terrain · Adversarial Motion Priors · Natural Gait · Robust Control

1 Introduction

The bipedal walking capability evolved by humans allows us to traverse diverse complex terrains with flexibility and ease, serving as a crucial source of inspiration for research in bipedal robotics. Compared to wheeled and tracked robots [1, 2], bipedal robots can adapt to scenarios like stairs, rugged ground, and obstacles [3] that are challenging for

traditional locomotion methods. However, achieving stable, natural, and autonomous walking control in complex terrains [4] remains one of the major challenges in the field of robotics.

Traditional bipedal control methods like the Zero-Moment Point (ZMP) criterion can ensure robot stability [5] but require precise environmental modeling and possess limited adaptability. With the rapid development of deep reinforcement learning techniques [6], researchers have begun exploring more robust and adaptive learning-based control methods. Among these, PPO [7] has quickly emerged as one of the mainstream methods in robot control due to its advantages in training stability and sample efficiency. However, standard PPO implementations often employ traditional reward function [8] designs that prioritize walking stability while neglecting motion aesthetics and naturalness. Consequently, bipedal robots trained in complex terrains still exhibit unnatural gaits. Natural gaits [9] have been shown not only to be aesthetically pleasing but also to yield improvements in energy efficiency and environmental adaptability. Achieving both robustness [4] and gait naturalness simultaneously [10] for bipedal robots is thus a critical direction for current research.

To address this issue, Peng et al. [11] learn natural gaits from reference motion data. Escontrela et al. [8] applied AMP to quadruped control, successfully achieving low-energy natural gaits by learning style priors from German Shepherd motion capture data, exhibiting appropriate gait transitions at different speeds. Peng et al. [3] further extended the AMP method to enable a quadruped robot to learn bipedal walking, using a teacher-student learning framework to allow the quadruped to walk stably on its hind legs with a natural gait style.

Despite the significant progress made by the AMP method in enhancing robot gait naturalness, existing research predominantly focuses on applications involving flat terrains [10] and quadrupedal robots; its application to bipedal robots in complex, unstructured terrains remains limited. Furthermore, current AMP frameworks typically employ simple network structures [12] and state representations, which struggle to process the high-dimensional perceptual information inherent in complex terrains.

Addressing the aforementioned challenges, this paper introduces an enhanced PPO-AMP framework aimed at achieving robust and natural walking control for bipedal robots in complex terrains. This framework incorporates three main innovations building upon the original PPO-AMP structure:

1. Significantly expanding the robot's environmental perception capabilities by integrating terrain height detectors and contact force sensors, establishing a foundation for adapting to complex terrains.
2. Optimizing the network architectures of PPO and AMP by increasing network depth and scale, thereby enhancing the system's capacity to process high-dimensional multi-modal inputs.
3. Redesigning the reward function and training strategy, changing the self-collision condition from a traditional penalty term to an episode termination condition, and incorporating PPO-Clip and KL divergence constraints to prevent excessively large update steps.

Experimental results demonstrate that the proposed enhanced PPO-AMP framework enables bipedal robots to achieve locomotion that is both robust and natural across a variety of complex terrains.

2 Enhanced Adversarial Motion Prior Control Framework

To overcome the dilemma where robustness and gait naturalness often seem mutually exclusive for bipedal robots in complex and variable terrains, we have refined the PPO-AMP based control framework. As shown in Fig. 1, this enhanced framework introduces innovative improvements and integrations in key aspects such as state representation, network architecture, reward mechanism, and training strategy, systematically enhancing the robot's capabilities in environmental perception, decision-making learning, and motion generation.

The enhanced framework architecture is illustrated in Fig. 1. The system fuses proprioceptive data from basic sensors with exteroceptive information from a terrain height detector and contact force sensors. The AMP module, trained using reference motion data, generates style rewards. These are combined with task-specific rewards derived from task objectives to guide the PPO training of the policy and value networks. Training stability is enhanced using KL divergence constraints and PPO-Clip.

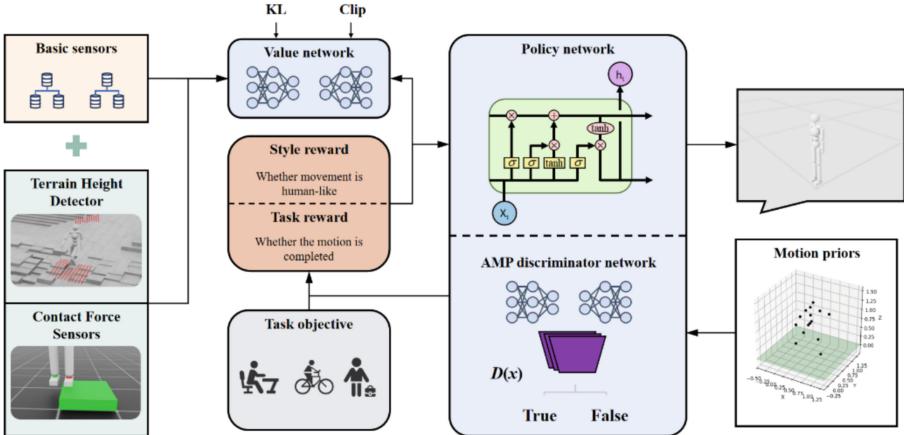


Fig. 1. Enhanced PPO-AMP control framework architecture.

2.1 System Framework and Network Modeling

We model the robot's continuous control problem as a Markov Decision Process, defined by a state space S , action space A , transition probability $P(s'|s, a)$, reward function $R(s, a, s')$, and a discount factor γ . At each discrete timestep t , the agent observes its current state $s_t \in S$, which comprises both proprioceptive information and processed environmental sensor data; the fusion of these two information types constitutes the

state input. Subsequently, the agent samples an action $a_t \in A$ based on its learned policy $\pi(a_t | s_t)$, where the action represents the target positions for the joints (the output of the policy network). After one forward pass, the physics simulation environment calculates the next state s_{t+1} based on the dynamics model and returns a reward signal $r_t = R(s_t, a_t, s_{t+1})$.

Our primary objective is to learn an optimal policy π^* that maximizes the expected long-term cumulative discounted reward starting from any initial state:

$$\pi^* = \operatorname{argmax}_{\pi} E_{\tau \sim \pi} \left[\sum_{t=0}^{\infty} \gamma^t r_t \right] \quad (1)$$

where τ represents a sequence of state-action pairs generated by policy π .

To ensure the generated gait accomplishes the terrain traversal task guided by the task reward r^{task} , the system is trained using the PPO algorithm via deep reinforcement learning. The training employs an Actor-Critic architecture; the former is the policy network outputting the action probability distribution, and the latter is the value network estimating the state value. Their combination allows the robot to learn task completion through trial and error. Additionally, to render the gait more human-like and natural, we utilize an AMP module. Its adversarial generative network distinguishes whether the current action output corresponds to real motion or simulated motion, providing a style reward r^{style} accordingly. The system maximizes r^{style} , thereby striving to mimic reference motion capture data as closely as possible. The synergy of these three networks enables the robot's walking to concurrently achieve both robustness and naturalness.

2.2 Enhanced State Representation via Multi-sensor Fusion

A key reason for the suboptimal performance of traditional bipedal control methods in complex terrains lies in their limited environmental perception capabilities. Relying solely on basic proprioceptive inputs like joint angles, velocities, and trunk pose/velocity forces the robot to react passively to environmental changes rather than actively adapt. We integrate a terrain height detector that scans the ground beneath the robot and contact force sensors at key joints. These sensors expand the state representation from 81 to 299 dimensions, enabling anticipatory terrain adaptation.

The robot's basic proprioceptive state is represented by an 81-dimensional vector. After integrating the processed information from the terrain height detector and contact sensors, the input state vector s_t for the policy and value networks expands to 299 dimensions, substantially improving their perceptual and predictive capabilities regarding complex terrains. Since the AMP discriminator focuses solely on the motion's style, its input remains the 81-dimensional proprioceptive state information.

2.3 Systematic Optimization of Deep Network Structure and Innovative Reward Mechanism

Experiments showed that a six-layer network with 2.79 million parameters achieved optimal performance, while a minimal three-layer network [512, 256, 128] with 840,000 parameters maintained basic walking ability.

Reward design and training strategies in traditional reinforcement learning often focus on singular objectives (e.g., walking stability or speed), neglecting the crucial aspect of gait naturalness. Furthermore, overly complex reward function designs not only increase the difficulty of manual parameter tuning but may also introduce conflicting learning objectives, hindering policy convergence. To address these issues, we proposed two key improvements:

First, we redesigned the reward mechanism, introducing an innovative modification specifically to the handling of self-collisions. Unlike traditional methods that treat self-collision as a penalty term with a negative weight, we define it as a condition that triggers immediate termination of the current episode. Experimental results indicate that this approach more effectively guides the agent to learn, plan, and execute collision-free action sequences, resulting in more coordinated and natural gaits. Concurrently, we discovered that for models employing deep networks and rich state representations, the style reward provided by AMP inherently contains substantial information conducive to accomplishing the walking task. Based on this insight, we adjusted the weighting between w_{task} and w_{style} , significantly reducing the weight of the task reward. In some scenarios, robust walking on complex terrains was achieved relying almost entirely on the style reward. This not only drastically simplifies reward design but also enhances the system's fidelity in mimicking the style of reference motions.

Second, to ensure training process stability, we adopted a training strategy combining an improved PPO-Clip mechanism with KL divergence constraints. PPO-Clip constrains the magnitude of policy updates by limiting the probability ratio between new and old policies. Its objective function is formulated as:

$$L^{CLIP}(\theta) = \hat{E}_t \left[\min \left(r_t(\theta) \hat{A}_t, \text{clip}(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon) \hat{A}_t \right) \right] \quad (2)$$

Where r_t is the probability ratio of new and old policies, \hat{A}_t is the advantage function estimate, and ε is the clipping coefficient. Additionally, we introduced a KL divergence constraint term into the objective function:

$$L^{KL}(\theta) = L^{CLIP}(\theta) - \beta \bullet KL[\pi_{\theta_{old}}(\bullet|s_t), \pi_\theta(\bullet|s_t)] \quad (3)$$

This dual-constraint mechanism effectively prevents drastic oscillations and overfitting during training, facilitating smoother convergence of the policy towards the optimal solution. The training process follows the standard PPO-AMP framework with our modifications for self-collision handling and reward weighting.

3 Experimental Validation and Analysis

To comprehensively evaluate the effectiveness of the proposed enhanced PPO-AMP framework for bipedal robot locomotion control in complex terrains, we conducted a series of systematic experimental studies within the Isaac Lab simulation environment.

3.1 Experimental Setup and Evaluation Methodology

As shown in Fig. 2, experiments used the 28-DoF Humanoid28 model on procedurally generated terrains including staircases (0.05–0.23 m steps), ramps (0–40% slopes), rugged landscapes, and random obstacles.

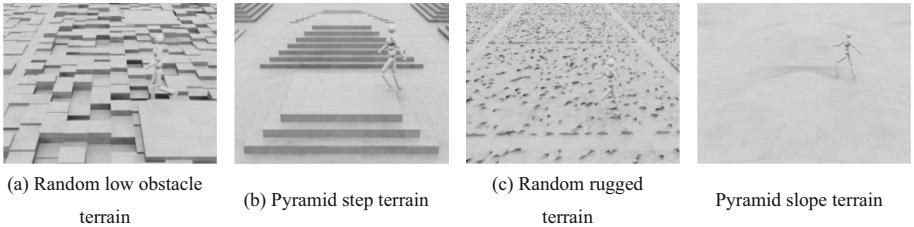


Fig. 2. The trained Humanoid28 robot achieves natural and robust locomotion across various complex terrains.

Performance assessment combined qualitative observation and quantitative analysis. The primary evaluation criterion was qualitative observation, focusing on:

- Gait stability: assessed by trunk sway magnitude, foot slippage occurrences, etc.
- Motion fluency and coordination.
- Adaptive gait adjustments to terrain features: such as leg lift height during obstacle negotiation, posture control on slopes, and foot placement selection on uneven surfaces.

Simultaneously, we recorded and analyzed key training metrics from selected experimental runs, including the evolution curves of discriminator loss, policy loss, and value loss against the total training timesteps, to aid in judging the dynamic characteristics and convergence quality of different configurations.

3.2 Empirical Investigation of Key Design Elements' Impact on Performance

Initially, the difference between self-collision handling strategies was evaluated. In Experiment 1, the conventional approach of penalizing self-collisions with negative rewards was employed, utilizing a three-layer network comprising 1.58 million parameters and training with a standard episode length limit. Observations revealed that the robot frequently exhibited non-functional hesitation or unnatural limb avoidance maneuvers during locomotion, failing to achieve smooth and stable walking. This finding corroborates the perspective from existing literature [13] that penalty terms can potentially guide the policy towards suboptimal local minima or induce unintended behaviors. In contrast, Experiment 2 strictly defined self-collision events as episode termination conditions. Compared to Experiment 1, the robot's walking stability exhibited marked improvement, demonstrating more natural and fluid fundamental locomotion capabilities.

Subsequently, we investigated the influence of network architecture and scale on learning walking strategies for complex terrains. When the network was configured with three layers and the parameter count was reduced to 840,000, this setup maintained fundamental walking capabilities and demonstrated higher computational efficiency, even with substantially fewer total parameters compared to the baseline experiments. However, upon further compressing the network parameters to approximately 570,000, a drastic degradation in walking performance was observed, with the robot virtually losing effective locomotion capability. This set of comparative experiments indicated

that below a certain critical network capacity, the model fails to effectively represent and learn the required perception-to-action mapping within complex environments.

When confronted with diverse and dynamically changing complex terrains, our optimally configured enhanced PPO-AMP framework (featuring a six-layer network architecture with approximately 2.79 million parameters, emphasizing style rewards supplemented by minimal necessary task guidance, and employing self-collision as a termination criterion) demonstrated comprehensively superior performance compared to baseline methods and other tested configurations. As shown in Fig. 4, despite having a shorter training record owing to the early stopping technique, the optimal network (light grey curve) clearly demonstrates faster convergence and superior performance. It maintains a distinct advantage over Experiment 5, which exhibited strong performance in the previous comparison. Under these conditions, the system was capable of generating gaits that were both stable and natural, exhibiting precise perception of and adaptation to terrain features. For instance, when encountering terrain with continuous irregular height variations, the robot could preemptively adjust leg lift height and foot placement based on detected terrain characteristics; during ascent or descent on slopes, it automatically adjusted its trunk posture to maintain balance. These complex behaviors attest to the superior capability of our proposed enhanced framework in processing high-dimensional sensory input and generating coordinated control signals.

In Experiment 6, which utilized the optimal network architecture, intentionally increasing the weight of the survival reward substantially within the total reward function caused the robot’s policy to converge to a behavior that completely suppressed forward locomotion; instead, it maintained balance in place to maximize cumulative survival reward. The anomalous nature of the corresponding experimental metrics is also evident in Table 2. This outcome corroborates that excessive emphasis on conventional task rewards can override the style guidance introduced by AMP, which implicitly encodes motion intent, leading the policy to deviate from high-level task objectives.

Figure 3 presents the training curves for Experiments 1–5, displaying (from left to right, top to bottom): mean episode length, discriminator loss, policy loss, and value loss versus training timesteps. The curves are color-coded as follows: dark grey (Exp. 1), light blue (Exp. 2), dark blue (Exp. 3), orange (Exp. 4), and magenta (Exp. 5), corresponding to the configurations detailed in Table 1 (Fig. 4).

Table 1. Experimental Parameters

Exp. No.	Network Layers	Network Params	Self-Collision Termination	Survival Timestep Limit
1	2 layers	1.58 M	False	Standard
2	2 layers	2.28 M	True	Standard
3	3 layers	0.84 M	True	Standard
4	3 layers	0.57 M	True	Standard
5	3 layers	0.84 M	True	Extended
6	6 layers	2.79 M	True	Extended

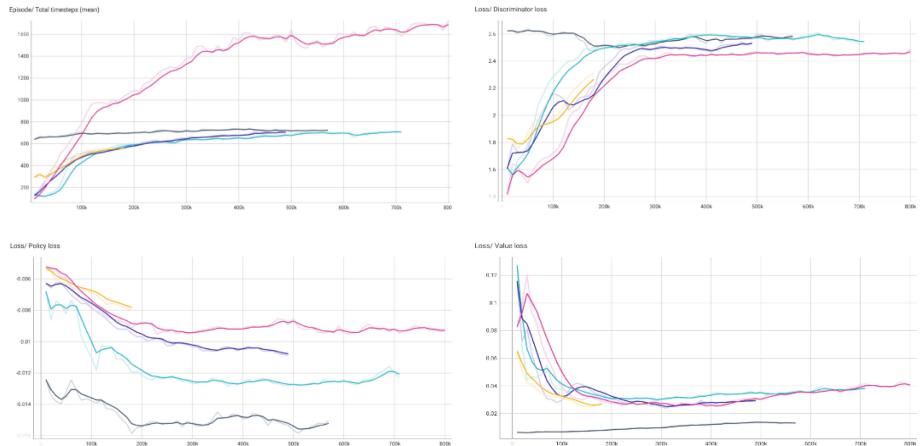


Fig. 3. Training performance metrics for Experiments 1–5.

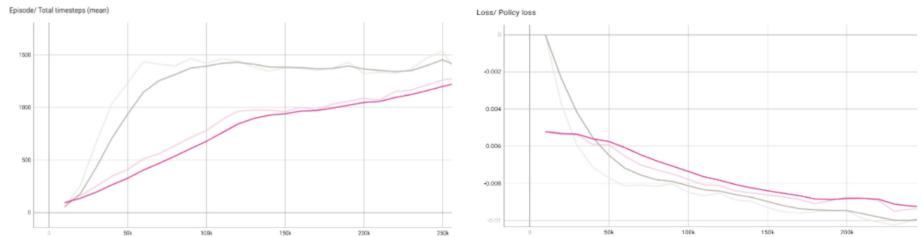


Fig. 4. Performance comparison of optimal network configuration.

Table 2. Average Survival Timesteps per Iteration

Exp. Metrics	Exp. No.	Start Value	End Value	Δ Value	Δ %
Survival Timesteps	1	639.8049	694.9182	55.1134	9%
	2	139.3844	575.4074	436.023	313%
	3	124.0024	559.7116	435.7092	351%
	4	292.2104	559.8938	267.6834	92%
	5	93.667	995.5627	901.8958	963%
	6	13.1568	52.4153	39.2585	298%
Discriminator Loss	1	2.6237	2.5124	-0.1114	-4%
	2	1.6157	2.4727	0.8571	53%
	3	1.6065	2.1542	0.5477	34%
	4	1.8286	2.2665	0.4378	24%
	5	1.4155	2.1443	0.7288	51%

(continued)

Table 2. (*continued*)

Exp. Metrics	Exp. No.	Start Value	End Value	Δ Value	Δ %
	6	1.4944	2.283	0.7885	53%
Policy Loss	1	-0.0124	-0.0152	-0.0028	23%
	2	-0.0068	-0.0114	-0.0047	69%
	3	-0.0063	-0.0092	-0.0029	47%
	4	-0.0053	-0.0078	-0.0025	47%
	5	-0.0052	-0.0088	-0.0036	69%
	6	0.0043	-0.0084	-0.0127	-297%
Value Loss	1	0.0064	0.0077	0.0013	21%
	2	0.1273	0.0324	-0.0949	-75%
	3	0.1158	0.0354	-0.0803	-69%
	4	0.0652	0.0267	-0.0385	-59%
	5	0.0825	0.0304	-0.0522	-63%
	6	0.2024	0.0152	-0.1872	-92%

4 Conclusion

This paper proposed and validated an enhanced AMP control framework designed to improve the walking robustness and naturalness of bipedal robots in complex, unstructured terrains. By integrating external environmental sensing systems, the robot's perceptual capabilities were significantly enhanced. Optimizing and employing deep neural network structures improved the policy network's ability to process high-dimensional multi-modal inputs and learn complex control strategies. Modifying the self-collision handling to be an episode termination condition, recognizing the dominant role of the AMP style reward, and incorporating improved PPO-Clip and KL divergence constraints simplified reward design and fostered the generation of safer, more natural gaits. Extensive experiments conducted in the high-performance Isaac Lab simulation platform across various challenging terrains demonstrated that the proposed framework, particularly the version utilizing deep networks, achieves significant improvements in task success rate, stability, gait naturalness, and generalization capability to unknown terrains compared to traditional AMP methods and approaches using only PPO.

References

1. Zhang, T., Mo, H.: Reinforcement learning for robot research: a comprehensive review and open issues[J]. Int. J. Adv. Robot. Syst. **18**(3), 17298814211007305 (2021)
2. Bakhsh, Q., Hasnan, K., Ahmed, A.: Comparative study between wheeled and tracked mobility system for Mobile robot[J]. Appl. Mech. Mater. **393**, 538–543 (2013)

3. Peng, T., Bao, L., Humphreys, J., et al.: Learning bipedal walking on a quadruped robot via adversarial motion priors[C]. In: Annual Conference Towards Autonomous Robotic Systems, pp. 118–129. Springer, Cham (2024)
4. Bao, L., Humphreys, J., Peng, T., et al.: Deep Reinforcement Learning for Bipedal Locomotion: A Brief Survey[J]. arXiv preprint arXiv:2404.17070 (2024)
5. Li, C., Li, M., Tao, C.: A parallel heterogeneous policy deep reinforcement learning algorithm for bipedal walking motion design[J]. *Front. Neurorobot.* **17**, 1205775 (2023)
6. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Playing Atari with Deep Reinforcement Learning[J]. arXiv preprint arXiv:1312.5602 (2013)
7. Schulman, J., Wolski, F., Dhariwal, P., et al.: Proximal Policy Optimization Algorithms[J]. arXiv preprint arXiv:1707.06347 (2017)
8. Escontrela, A., Peng, X.B., Yu, W., et al.: Adversarial motion priors make good substitutes for complex reward functions[C]. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 25–32. IEEE (2022)
9. Mikolajczyk, T., Mikołajewska, E., Al-Shuka, H.F.N., et al.: Recent advances in bipedal walking robots: review of gait, drive, sensors and control systems[J]. *Sensors*. **22**(12), 4440 (2022)
10. Wang, Y., Jiang, Z., Chen, J.: Amp in the Wild: Learning Robust, Agile, Natural Legged Locomotion Skills[J]. arXiv preprint arXiv:2304.10888, 6 (2023)
11. Peng, X.B., Ma, Z., Abbeel, P., et al.: Amp: adversarial motion priors for stylized physics-based character control[J]. *ACM Trans. Graph. (ToG)*. **40**(4), 1–20 (2021)
12. L’Erario, G., Hanover, D., Romero, A., et al.: Learning to walk and fly with adversarial motion priors[C]. In: 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 10370–10377. IEEE (2024)
13. Dulac-Arnold, G., Levine, N., Mankowitz, D.J., et al.: Challenges of real-world reinforcement learning: definitions, benchmarks and analysis[J]. *Mach. Learn.* **110**(9), 2419–2468 (2021)



Experimental Investigation of Drag and Lift Forces on Objects Moving Through Granular Media with Varying Geometries and Speeds

Chundong Liu¹, Xiaopu Hao¹, Zhiqiang Song², Qiang Zhang¹(✉), Shuai Kang³, and Longchuan Li²

¹ College of Mechanical Engineering, Hebei University of Architecture, Hebei Zhangjiakou 075000, China
qiangzhang1021@163.com

² College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China

³ College of Mechanical and Electrical Engineering, Beijing University of Chemical Technology, Beijing 100029, China

Abstract. Understanding the mechanics of object motion through granular media is essential for applications in terrain navigation, agricultural engineering, and planetary exploration. This study presents a comprehensive experimental investigation into the drag and lift forces acting on rigid bodies of different geometries—including flat plates and wedge-shaped objects—moving at controlled speeds and burial depths through a dry granular bed. A custom-designed, high-precision towing apparatus equipped with a six-axis force sensor was used to collect real-time force data. Results demonstrate strong coupling between drag/lift forces and object geometry, speed, and immersion depth. Flat plates exhibited periodic force fluctuations linked to granular rearrangements, while wedge models induced sharp force spikes and varied lift characteristics depending on front-edge inclination. Notably, increasing towing speed led to nonlinear changes in resistance, suggesting non-negligible inertial effects even in low-velocity regimes. These findings provide experimental insights into the dynamics of granular drag and inform the design of optimized devices for granular locomotion and interaction.

Keywords: Granular media · drag force · lift force

1 Introduction

Granular materials are ubiquitous in nature and industry, appearing in forms such as sand, soil, snow, gravel, grains, sugar, salt, coal, ores, pharmaceuticals, and chemical products [1]. These materials consist of macroscopic particles typically larger than 100 μm. Due to their widespread presence and relevance, granular materials are among the most important classes of matter on Earth. Their behavior plays a critical role in numerous scientific and engineering domains, including geotechnics, mining, chemical processing, pharmaceuticals, biology, agriculture, and environmental science [2].

Despite mature theories describing individual particle mechanics—such as elasticity and plasticity—and widely adopted numerical approaches like finite element methods [3], it remains extremely challenging to develop a unified model for the collective behavior of granular flows. Granular media exhibit complex properties that lie between those of solid and fluid phases, showing non-equilibrium, dissipative dynamics [4]. The development of a comprehensive theoretical framework for granular materials is recognized as one of the most significant open problems in modern science [5].

One important research scenario involves studying the motion of intruding bodies—such as tillage tools, robotic platforms, or planetary landers—through granular substrates. This process involves a rich interplay of mechanical resistance, energy dissipation, and local particle rearrangements. Understanding these mechanisms is vital for applications ranging from agriculture and construction to disaster recovery and extraterrestrial exploration. In this study, we conduct a detailed investigation of how object geometry, velocity, and immersion depth affect the drag and lift forces encountered during motion through granular media. Using a combination of controlled experiments, data analysis, and physical modeling, we aim to develop a predictive framework for drag dynamics, supporting both theoretical insights and practical design optimization [6].

Recent advances in granular physics have yielded various theoretical models and experimental strategies to better characterize force transmission during intrusion [7]. Building on these foundations, this work proposes a combined analytical and experimental approach to systematically study the influence of geometric and dynamic parameters on object-granular interactions.

2 Methods

2.1 Experimental Setup

The experimental system was custom-built for high-precision drag force measurements in granular media. It consists of a modular aluminum alloy frame, servo motor drive system, ball screw actuator, linear guide rails, and a six-axis force sensor. The test chamber is enclosed by transparent acrylic panels and filled to a depth of 0.5 m with irregular polyethylene granules (average diameter ~ 6 mm), mimicking natural coarse-grain environments.

The apparatus includes four primary modules: a load-bearing structural frame; an X-axis motion drive system for precise linear movement; a mounting platform for rigidly attaching test objects; and an integrated control system for synchronized motion execution and data acquisition [8]. During operation, the six-axis force sensor continuously measures forces acting on the object, enabling real-time analysis of drag and lift. The platform has a maximum traction capacity of 25,000 N and lift range of 3,000 N. The use of 6061 aluminum profiles and reinforced joints ensures high structural integrity and stiffness. Motion control is achieved via a servo motor coupled to a ball screw for precise, repeatable linear movement (Fig. 1).

2.2 Test Objects

Two sets of test bodies were developed to examine the effects of geometry on mechanical response: flat plates and wedge-shaped blocks (Fig. 2).

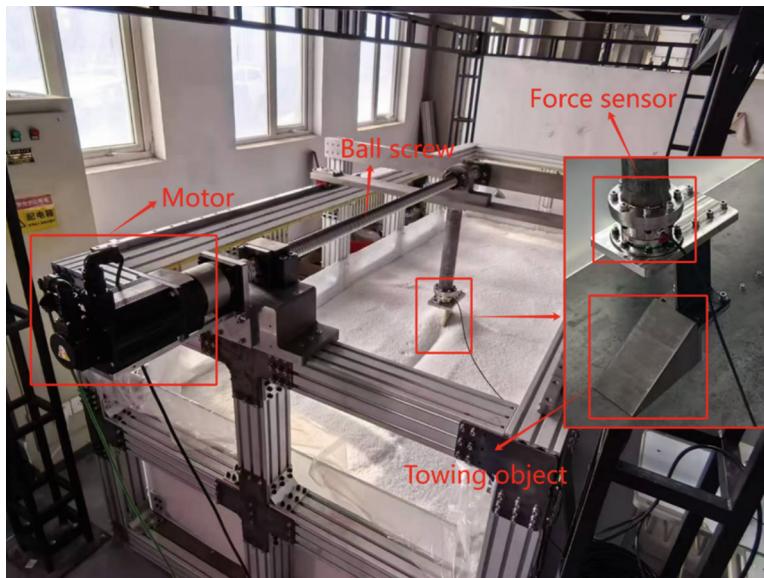


Fig. 1. Schematic of the towing experimental setup.

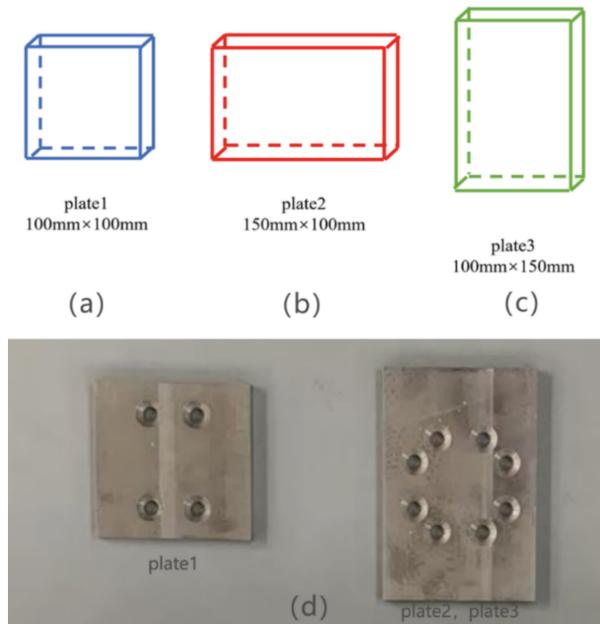


Fig. 2. Photographs of the flat plate models.

Flat Plates: This group includes three flat models, all with constant thickness of 10 mm. They differ in planar dimensions and orientation:

P1: Square plate ($100 \text{ mm} \times 100 \text{ mm}$)

P2: Rectangular plate ($150 \text{ mm} \times 100 \text{ mm}$)

P3: Rotated rectangular plate ($100 \text{ mm} \times 150 \text{ mm}$)

These configurations allow comparative analysis of symmetry and directional influence on drag behavior. A photograph of the models (Fig. d) clearly shows their shapes and relative sizes. The objective is to understand how planar dimensions and alignment affect resistance and lift under varying speeds and depths, providing empirical guidance for optimized structural design in granular environments [9].

Wedge Models: These were used to investigate how varying front-edge slopes affect hydrodynamic-like performance in granular media. All wedges share a base width of 0.1 m, length of 0.2 m, and height of 0.1 m. The inclination angle is adjusted by varying the leading edge height:

W1: $h = H/16 = 6.25 \text{ mm}$, consequently, $\alpha = 3.58^\circ$

W2: $h = H/8 = 12.5 \text{ mm}$, consequently, $\alpha = 7.13^\circ$

W3: $h = 3H/16 = 18.75 \text{ mm}$, consequently, $\alpha = 10.62^\circ$

These configurations enable systematic investigation of wedge slope effects on drag, lift, and flow-induced pressure fields. Figure (d) shows the physical models. Insights from these tests have implications for aquatic vehicle design, biomimetic propulsion systems, and sediment interaction analysis [10] (Fig. 3).

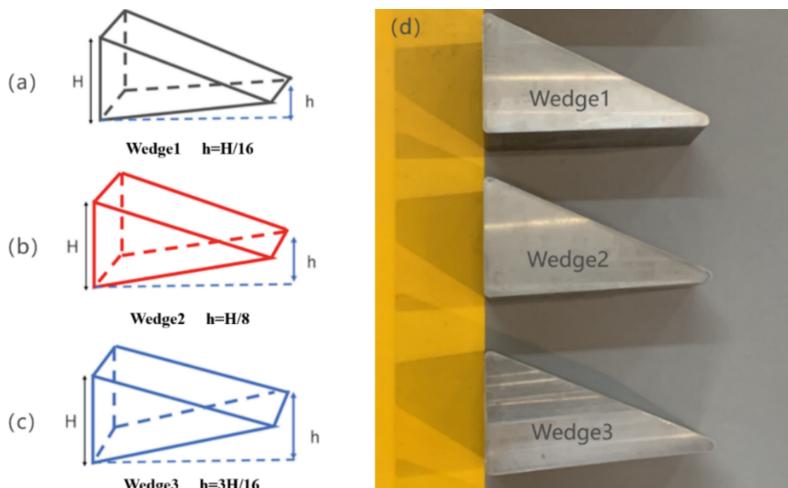


Fig. 3. Photographs of the wedge-shaped models.

2.3 Experimental Design

Each object was rigidly attached to the center of the platform using a high-stiffness H-beam linkage connected directly to the force sensor, minimizing extraneous torque or

force transmission error. The servo motor drove the objects at constant velocity along a linear guide track through the granular medium.

The two main experimental parameters were:

A1) Velocity: 0.01 m/s to 0.09 m/s in 0.01 m/s increments (9 total levels)

A2) Burial Depth: 0.1 m and 0.2 m

This range captures both quasi-static and slow dynamic regimes, facilitating analysis of depth-speed interactions on drag and lift.

3 Results and Discussion

3.1 Flat Plate Behavior

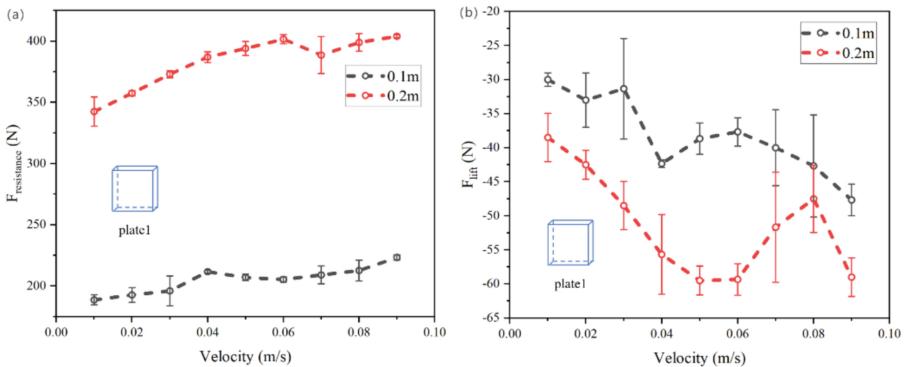


Fig. 4. Effect of towing speed on the drag and lift forces experienced by plate model P1.

Figure 4(a) shows drag force on P1 versus velocity at two burial depths. Drag increased with velocity in a nonlinear fashion, and deeper burial (0.2 m) resulted in significantly higher resistance. At higher speeds, a plateauing effect was observed, likely due to localized fluidization or particle rearrangement, which softened the force buildup. The presence of error bars indicates consistent, repeatable measurements.

Figure 4(b) presents lift force data for the same configuration. At 0.2 m depth, lift became more negative (stronger downward force), reaching a minimum around -60 N near 0.05 m/s. This suggests enhanced suction effects or more intense flow disturbance. Shallower depth yielded smaller, more stable lift values. A 3D diagram of P1 is provided for geometric reference.

Figure 5(a) compares drag between P2 and P3. P3 consistently exhibited higher drag, especially above 0.06 m/s, likely due to its taller profile in the motion direction. Drag increased steadily with speed for both models, but P2 showed irregular fluctuations in high-speed regimes, possibly due to wake instability or vortex shedding.

Figure 5(b) shows lift forces for P2 and P3. P2 experienced more negative lift, with values nearing -70 N at high speeds, while P3 had smaller, more stable lift. This difference likely results from P2's wider horizontal profile promoting greater pressure

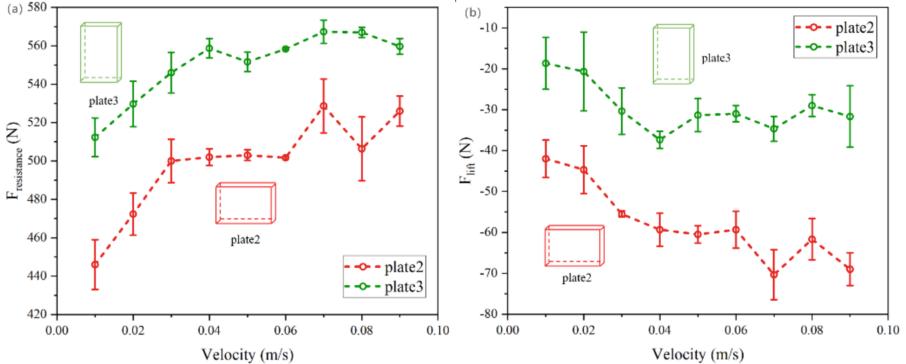


Fig. 5. Comparison of drag and lift forces on plate models P2 and P3 under identical conditions.

differential across the surface. Diagrams illustrate the distinct geometries and reinforce their aerodynamic interpretations.

3.2 Wedge Model Behavior

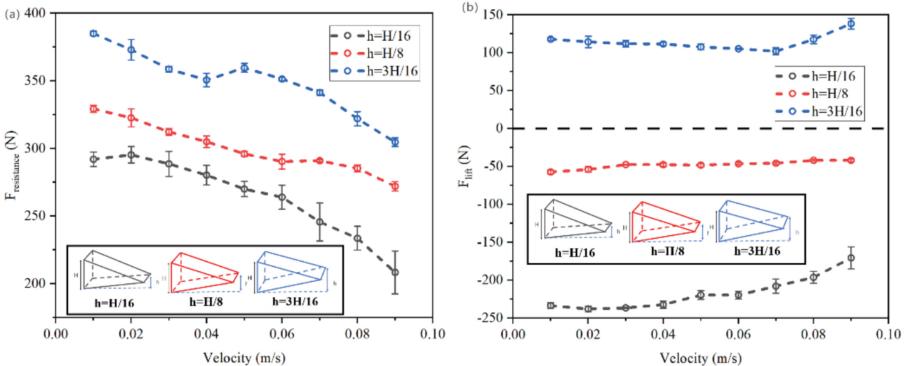


Fig. 6. Comparison of drag and lift forces on wedge models W1, W2, and W3 under identical conditions.

Figure 6(a) illustrates drag versus velocity for three wedge angles. All configurations showed decreasing drag with increasing speed, likely due to boundary layer reattachment or delayed separation. However, greater wedge angles resulted in higher drag across all speeds. At >0.06 m/s, drag ranged from ~ 210 N (W1) to ~ 310 N (W3), indicating the influence of frontal area and flow disturbance.

Figure 6(b) shows both drag and lift responses. W3 produced strong positive lift (up to 130 N), indicating upward flow pressure due to steep inclination. W1 generated large negative lift (~ -240 N), likely due to downward pressure gradients and concentrated upper surface flow. W2 yielded moderate, nearly stable lift near -50 N. This comparison

suggests a performance trade-off: steep wedges improve lift but increase drag, while shallow wedges reduce drag at the cost of stronger suction forces.

4 Conclusion

The movement of bodies through granular substrates involves complex, nonlinear interactions among body geometry, velocity, and media structure [11]. Using controlled experiments on flat plates and wedge-shaped bodies, this study demonstrated how geometry significantly influences drag and lift forces. Flat plates induced periodic force fluctuations associated with local particle avalanching and reconstruction, while wedges triggered abrupt force surges and granular ejection due to sharp leading edges [12].

Additionally, both drag and lift responses exhibited nonlinear trends with velocity, highlighting the importance of inertial effects even at low speeds. These effects must be accounted for when modeling drag forces in granular environments [13].

Overall, these findings provide valuable empirical data for developing predictive drag models and designing efficient structures for motion in granular media. Future work will extend to particle shape variation, wet granular systems, and multimodal particle distributions to further refine understanding of granular resistance mechanisms.

Acknowledgements. This work was supported in part by the Fund of Robot Technology Used for Special Environment Key Laboratory of Sichuan Province, China (Grant No. 23kftk05), by the Fund of the State Key Laboratory of Robotics at Shenyang Institute of Automation, China (Grant No. 2024-O16, 2024-Z08), by the Fundamental Research Funds for the Central Universities, China (Grant No. buctrcc202215), by the National Natural Science Foundation of China (No. Grant 62273340), by LiaoNing Revitalization Talents Program, China (Grant No. XLYC2403108), and by Key Research and Development Projects of Liaoning Province, China (Grant No. 2024JH2/10240045).

References

1. Liu, J., Wang, Z., Tian, Y., Li, L., Ma, S.: Drag force modeling with induced surface deformation in granular media. *Powder Technol.* **455**, 120757 (2025)
2. Ding, Y., Gravish, N., Goldman, D.I.: Drag induced lift in granular media. *Phys. Rev. Lett.* **106**(2), 028001 (2011)
3. Kumar, S., Dhiman, M., Reddy, K.A.: Magnus effect in granular media. *Phys. Rev. E* **99**(1), 012902 (2019)
4. Li, C., Zhang, T., Goldman, D.I.: A terradynamics of legged locomotion on granular media. *Science* **339**(6126), 1408–1412 (2013)
5. Guillard, F., Forterre, Y., Pouliquen, O.: Lift forces in granular media. *Phys. Fluids* **26**(4) (2014)
6. Ma, X., Wang, G., Liu, K.: Design and optimization of a multimode amphibious robot with propeller-leg. *IEEE Trans. Robot.* **38**(6), 3807–3820 (2022)
7. Liu, J., Tian, Y., Wang, Z., Li, L., Ma, S.: Modeling the impact of terrain surface deformation on drag force using discrete element method and empirical formulation. *Appl. Math. Model.* (2024)

8. Maladen, R.D., Umbanhower, P.B., Ding, Y., Masse, A., Goldman, D.I.: Granular lift forces predict vertical motion of a sand-swimming robot. In: 2011 IEEE International Conference on Robotics and Automation, pp. 1398–1403 (2011)
9. Patino-Ramirez, F., O'Sullivan, C.: Optimal tip shape for minimum drag and lift during horizontal penetration in granular media. *Acta Geotech.* **19**(1), 19–38 (2024)
10. Percier, B., Manneville, S., McElwaine, J.N., Morris, S.W., Taberlet, N.: Lift and drag forces on an inclined plow moving over a granular surface. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **84**(5 Pt 1), 051302 (2011)
11. Li, L., Zhao, C., He, S., Qi, Q., Kang, S., Ma, S.: Enhancing undulation of soft robots in granular media: a numerical and experimental study on the effect of anisotropic scales. *Biomimetic Intell. Rob.* **4**(2), 100158 (2024)
12. Potiguar, F.: Lift force on an asymmetrical obstacle immersed in a dilute granular flow. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **84**, 061302 (2011)
13. Potiguar, F., Ding, Y.: Lift and drag in intruders moving through hydrostatic granular media at high speeds. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **88**, 012204 (2013)



Structure Design and Output Characteristics of Hemispherical Tactile Sensor for 3D Force Detection

Xiaotao Du , Ling Weng (✉), Shichao Zuo , Bowen Cui , Xiaopeng Ji , and Lanyang Hao

State Key Laboratory of Intelligent Power Distribution Equipment and System, Hebei University of Technology, Tianjin 300401, People's Republic of China

llweng@hebut.edu.cn

Abstract. As an important part of wearable devices, flexible tactile sensors have attracted extensive attention from society and researchers. However, the existing flexible tactile sensors have problems such as small force measurement range and single detection direction, so it is very important to study large-range multi-direction force tactile sensors. Designed with large load range, high sensitivity which can realize the three dimensional force test of flexible tactile sensor. The sensor is composed of a silicone elastomer shell, an embedded NdFeB silicone elastomer magnetic cylinder and a tunnel reluctance element array at the bottom. The experimental results show that the tactile sensor can recognize the normal force in the range of 0–50 N and the tangential force in the range of 0–10 N, and the three-dimensional force direction can be judged. The sensor also has good dynamic performance, in the normal force loading and unloading experiments, the response time and recovery time are 42 ms and 40 ms, respectively, with good response and recovery ability. the sensitivity of normal force in the range of 0–30 N and 30–50 N is 11.51 mV/N and 3.03 mV/N, and the tangential force is 15.82 mV/N. This paper lays a foundation for further research on tactile sensors that can detect spatial triaxial forces.

Keywords: three-dimensional force · self-decoupling · tactile sensor · large range · high sensitivity

1 Introduction

Humans can easily identify textures or process objects of different shapes and sizes with the touch of their fingers. In recent years, the design of flexible manipulators and bionics based tactile sensors with the same capability as human hands has become a research hotspot. After years of research, scientists have made a lot of progress in the field of robot operation [1–3], but due to the lack of tactile sensing devices that can be compared with human hands, there are still challenges in the realization of contact force and adaptive grasping. The tactile sensing device simulates human skin and enables the robot to obtain various information such as temperature, shape and hardness of

the external environment through contact, and feedback the processed information, so that the robot can dynamically adjust the next operation behavior [4]. Human skin is mainly composed of four types of mechanoreceptors [5, 6], which can sense external information with ultra-high resolution and extremely low geometric distance, and can instantly obtain a variety of information feedback to the brain for information processing [7], and can easily distinguish between normal and shear forces at the same time [8]. Through the skin's powerful perception system, humans can dynamically adjust grasping posture, grasping force and other information according to the information obtained by touch to achieve accurate and reliable grasping, recognition and other functions [9]. With the development of tactile sensing theory and new magnetic materials, it is imperative to develop magnetic tactile sensors with high sensitivity, low power consumption and miniaturization. Compared with capacitive [10–12], piezoresistive [13–16], piezoelectric [17–20] and other tactile sensors based on other principles, magnetic tactile sensors have the advantages of high sensitivity, low hysteresis, low power consumption, and easy realization of three-dimensional detection and remote detection.

At present, flexible tactile sensors are mainly used for force or Angle measurement. The magnetic tactile sensor described in this paper is a new type of skin-like tactile sensor. The main detection part is made of NdFeB magnetic particles mixed with silicone elastomer, which combines the excellent permeability of NdFeB particles and the high flexibility of silicone elastomer, and has good thermal stability. The flexible tactile sensor based on NdFeB particles has simple manufacturing process, low cost, high sensitivity, large measurement range, rapid dynamic response, and can measure normal force, tangential force and dynamic force. The sensor can be installed on the mechanical hand to grasp the object, and can be used in the stiffness detection, contour and texture recognition of the object to be grasped.

2 Sensor Structure Design

As the largest sensory organ of the human body, skin plays a very important role in human interaction with the outside world, including sensing force, temperature, humidity, etc. A large number of mechanoreceptors are distributed in human hands, such as Meckel disc, Messner body, Pacini body, etc. These receptors cooperate with each other so that human fingers can identify pressure, temperature, humidity, etc. At the same time, fingers have highly flexible joints and muscles, which can passively accept external stimuli and actively explore the surface of objects to obtain more tactile information. Inspired by the sensory function of the finger skin, a hemispherical tactile sensing unit is designed to realize triaxial force measurement, as shown in Fig. 1. The unit consists of four parts, namely, an outer silicone elastomer shell (It is made of a half sphere with a radius of 7.5 mm and a cylinder with the same radius and a thickness of 2 mm), NdFeB silicone elastomer magnetic cylinder embedded in silicone elastomer, four TMR2003(Tunnel Magneto Resistance) components (size $3 \times 3 \times 1.45$ mm, package form SOT23-5) and the underlying PCB board. The magnetic cylinder, with a radius of 2 mm and a height of 7.5 mm, was filled in the hole at the top of the silicone elastomer shell after axial magnetization in the vertical downward direction. The projection position on the printed circuit board was the center of the four TMR components. TMR element for TMR2003

element, at the bottom of the packaging format for SOT23-5, the magnetic sensitive for three pin pointing in the direction of two pins. The susceptibility directions of the four TMR2003 elements remain consistent with respect to the central cylinder, so under the same of force can remain relatively consistent output. The outer hemispherical silicone elastomer shell is responsible for sensing the external stimulus. When the hemispherical shell deforms under the action of external stress, the inner NdFeB silicone elastomer magnetic cylinder deforms accordingly. The surrounding magnetic field change diagram is shown in Fig. 2. Through the tunnel reluctance effect, the corresponding voltage output is generated to complete the detection of the external stress. The output of the TMR element is the amount of voltage change.

The output of the TMR2003 component is divided into four regions, as shown in Fig. 3 [21]. When the magnetic field changes between the two regions, the TMR2003 component will produce a step output, according to which the direction arrangement of the TMR component can be designed to realize the self-decoupling of the three-axis force.

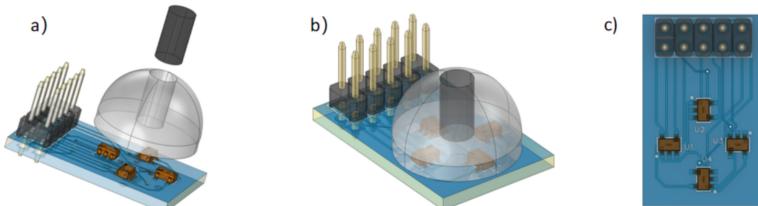


Fig. 1. (a) electronic skin structure disassembly diagram (b) sensor overall diagram (c) PCB

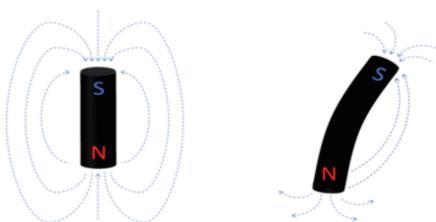


Fig. 2. Changes of the surrounding magnetic field when the NdFeB silicone elastomer magnetic cylinder is bent

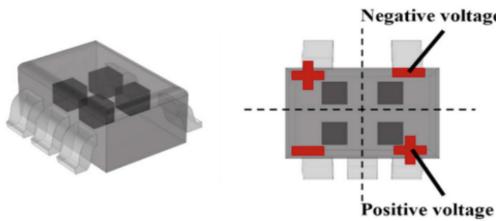


Fig. 3. TMR2003 component output positive and negative areas

3 Preparation of Silicone Elastomer Shell and Magnetic Cylinder

Figure 4 shows the preparation process of the silicone elastomer used in the tactile sensor and the NdFeB silicone elastomer magnetic cylinder embedded in it. The silicone elastomers used in the tactile sensor are Smooth-ON Eciflex 00-30 Series Liquid A (5% polysiloxane +95% Tovenki silicone oil) and Liquid B (5% hydrogen-containing polysiloxane +95% end vinyl silicone oil)1:1 mixing resulted in a silicone elastoplast prepolymer. The inner NdFeB silicone elastomer magnetic cylinder is mixed by NdFeB particles (particle size 150 μm , produced by Sinnord) and silica gel prepolymer according to the mass ratio of 3:1. The mold is cured at room temperature for 120 min, and the NdFeB silicone elastomer magnetic cylinder is obtained after the release.

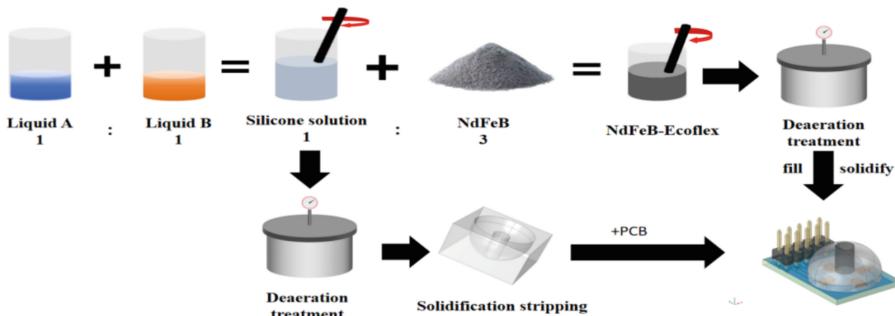


Fig. 4. Preparation process of tactile sensor

On the basis of using the same materials and the same preparation methods, the mixed mass ratio of different NdFeB powder and silicone elastomer prepolymer was studied. Set the mass ratio of NdFeB and silicone elastomer prepolymer to 1:1, 2:1, 3:1, 4:1, 5:1 for preparation. After completion, the NdFeB silicone elastomer magnetic cylinder is axially magnetized, and its magnetic field strength and hardness are tested after the completion of magnetization. The study found that when the mass ratio was 1:1, the magnetic cylinder flexibility of the prepared NdFeB silicone elastomer magnetic cylinder was the best, but the magnetic field intensity was too small; When the mass ratio is 5:1, the magnetic field intensity is the largest, but under this ratio, the magnetic cylinder has great hardness, is difficult to bend, and is easy to break under stress. Finally, 3:1 is determined as the best ratio.

When determining the length of the internal NdFeB silicone elastomer magnetic cylinder, COMSOL simulation software is used to explore the length of different NdFeB silicone elastomer magnetic cylinder. Different NdFeB silicone elastomer magnetic cylinder with lengths of 1 mm, 2 mm, 3 mm, 4 mm, 5 mm, 6 mm, 7 mm, 8 mm and 9 mm were prepared in groups, and axial magnetization was carried out after preparation. The position change of different NdFeB silicone elastomer magnetic cylinder is that the upper surface is fixed on the top of the hemispherical shell, and the different length of NdFeB silicone elastomer magnetic cylinder only affects its depth in the hemispherical shell, and then the magnetic field intensity generated on the surface of the TMR2003 component is respectively detected in the COMSOL simulation software. The detection method is to measure its Y-axis direction component at the height of the TMR2003 element, that is, the magnetic field strength component in the direction of the red solid line in Fig. 5(a). The test results are shown in Fig. 5(b), where the horizontal coordinate is the length of the red line in Fig. 5(a), and the vertical coordinate is the magnetic field strength. As can be seen from the figure: The maximum magnetic field intensity generated by the 1 mm–9 mm NdFeB silicone elastomer magnetic cylinder at the height of the TMR2003 element increases with the increase of the length of the NdFeB silicone elastomer magnetic cylinder, and when the length is 9 mm, the bottom of the cylinder contacts with the height of the TMR element. When selecting this length, the magnetic field generated by the magnetic cylinder at the height of the TMR2003 component is distorted, and the pressure sensing function cannot be completed, so 8 mm is selected as the best length.

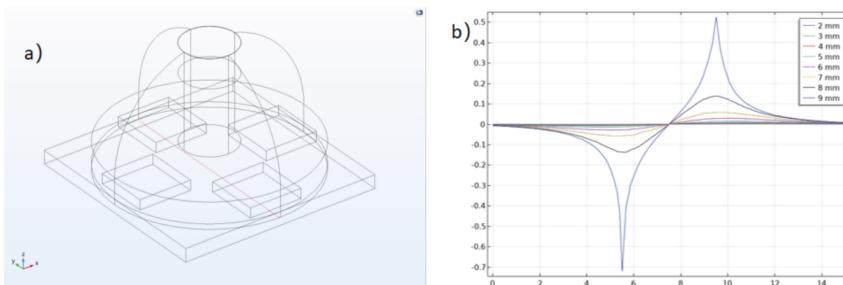


Fig. 5. (a) Direction of magnetic field intensity detected in simulation (b) magnetic field intensity generated by magnetic cylinders of different heights

4 Sensor Output Characteristic Test

4.1 Normal Force Test

Build the sensor output characteristic test platform, as shown in Fig. 6. The sensor is assembled according to the hemispherical tactile sensor pattern; The experimental platform is composed of a press contact with a diameter of 8 mm, a digital display push-pull meter, a DC power supply, a DH-8303 dynamic data acquisition card, and

a computer. The hemispherical touch sensor is fixed on the base of the digital display push tension meter, the circular press contact is fixed on the contact end of the push tension meter, and the rotation axis of the digital display push tension meter is adjusted so that the contact point of the push tension meter is directly opposite the top of the hemispherical touch sensor; The DC regulated power supply is used to provide power support for the TMR2003 component. The + port of the DC power supply is connected to the outgoing VCC, the - port of the DC power supply is connected to the outgoing GND port, and the + and - ports of one acquisition channel of the data acquisition card are respectively connected to the differential voltage V+ and V- ports of the same TMR2003 component. The output voltage signal after acquisition is displayed on the computer.

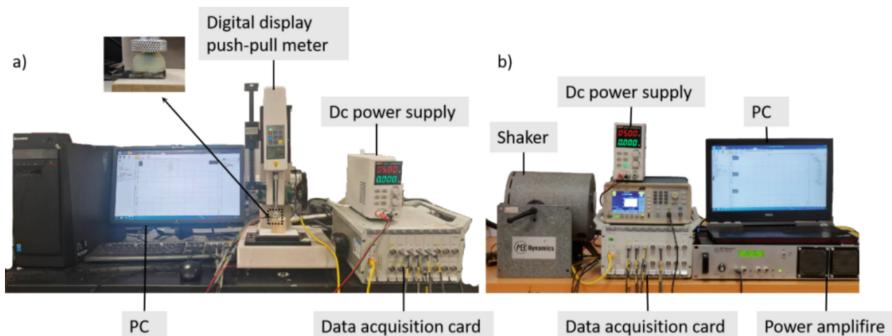


Fig. 6. (a) static force test platform (b) dynamic force test platform

Since the TMR2003 element array at the bottom adopts a symmetrical arrangement, when applying vertical downward pressure to the top of the tactile sensor, the output data of the four TMR2003 elements are basically the same when measuring the normal force. Therefore, in order to obtain more intuitive results, The output of a TMR2003 component was used to characterize the detection ability of the tactile sensor for normal force, and the corresponding relationship between normal force and output voltage was obtained.

With one of the four TMR2003 components at the bottom as the detection component, a press contact with a diameter of 8 mm was placed on the hemispherical touch sensor to conduct a press experiment to verify the relationship between the external stress accepted by the hemispherical touch sensor and the output voltage, and to test its force measurement range and output voltage signal range.

Uniform and vertical downward pressure is generated by the digital display push-pull meter, and 0–50 N normal force is applied to the top of the hemispherical tactile sensor with the step length of 5 N. The experiment is conducted in this way, and the output voltage is recorded successively, and the voltage waveform diagram is drawn as shown in Fig. 7. By observing the output voltage waveform, we can see that the output voltage of the first half increases with the increase of pressure, while the output voltage of the second half changes little with the increase of pressure. The reason is that when the pressure is initially increased, the magnetic field intensity on the height plane

of TMR2003 is small. When the hemispherically shaped tactile sensor is deformed by external pressure, the magnetic field intensity on the TMR2003 surface changes greatly. With the increase of displacement, the magnetic field intensity gradually increases. The magnetic field intensity produces saturation, and the magnetic field generated by the magnetic cylinder saturates the TMR2003 element, so the output voltage changes little. It can be seen that the normal force test range of the hemispherical tactile sensor is 0–50 N, and it has good linearity at 0–30 N.

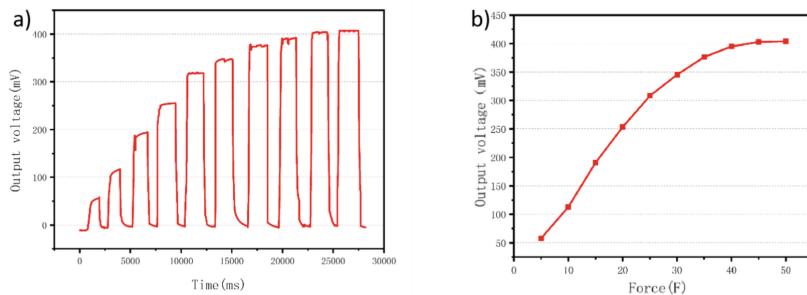


Fig. 7. (a) Output voltage of the tactile sensor under normal force (b) output voltage line diagram

4.2 Response Time Test

The tactile sensor was fixed on the experimental bench, and the input square wave signal was used to generate the output voltage waveform. The experimental results are shown in Fig. 8. The waveform obtained was analyzed, and the time between the signal generation and the signal peak was recorded as the response time, and the time from the peak to 0 point was recorded as the recovery time. It can be seen that the tactile sensor responds quickly and can return to 0 point after the pressure is removed. The response speed is fast, the recovery ability is strong, and the dynamic force performance is good.

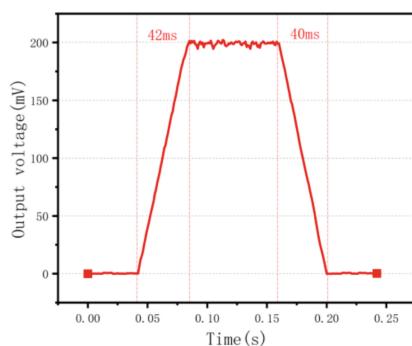


Fig. 8. Response time test

4.3 Dynamic Force Test

The normal sinusoidal pressure is applied to the tactile sensor. The experimental equipment is: signal generator (AFG2021-SC), power amplifier (MB500VI), and stimulator (MODAL50) combined to apply dynamic normal force. The data acquisition card is connected to PC to collect and process the output signal, use DC regulated power supply for power supply.

The tactile sensor is fixed on the experimental bench, and a 5 Hz sine wave signal is output by a signal generator. After processing by a power amplifier, the force in the form of a sine wave whose output frequency of the driver is 5 Hz is applied to the normal direction of the tactile sensor. The output obtained is shown in Fig. 9. In 1000 cycles of experiments, the output voltage error of the tactile sensor is small and there is no obvious change, which shows that the dynamic performance of the sensor is good and can be used for the measurement of dynamic force.

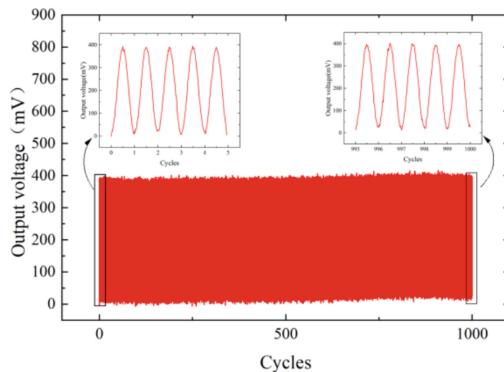


Fig. 9. Voltage output waveform of tactile sensor under dynamic force

4.4 Tangential Force Test

Take the array composed of four TMR2003 components as the detection part, replace the pressing contacts with a round contact with a diameter of 2 mm, glue the hemispherical touch sensor to the side of the test bench, make the digital display push pull meter press direction to the side of the hemispherical touch sensor, and apply tangential force to the hemispherical touch sensor to verify the output of the four TMR2003 components. The purpose of this experiment is to test the tangential force of the hemispherical tactile sensor and verify the output rule of the TMR2003 component array.

Uniform, vertically downward pressure generated by the digital display push-pull meter is applied to the side of the hemispherical tactile sensor as a tangential force, with the direction pointing in the direction of U1, as shown in Fig. 10 (a); A tangential force of 0–10 N was applied with a step length of 1 N, and the output voltage values of each TMR2003 component were recorded sequentially, and the output voltage line diagram was drawn, as shown in Fig. 10 (b). By analyzing its output, it can be seen that the step

output of TMR2003 component will be generated when the magnetic field changes in the magnetic sensitive direction, just like the output of U2 and U4. After the step, the output voltage basically remains unchanged; The output voltage of U1 increases with the increase of tangential force, and that of U3 decreases with the increase of tangential force. It can be seen that within the range of 1–10 N tangential force, the tangential force direction can be judged and the tangential force magnitude can be identified through the output of the TMR component array. The direction of the external pressure can be judged by the positive or negative output value of the TMR element to realize the sensor self-decoupling function.

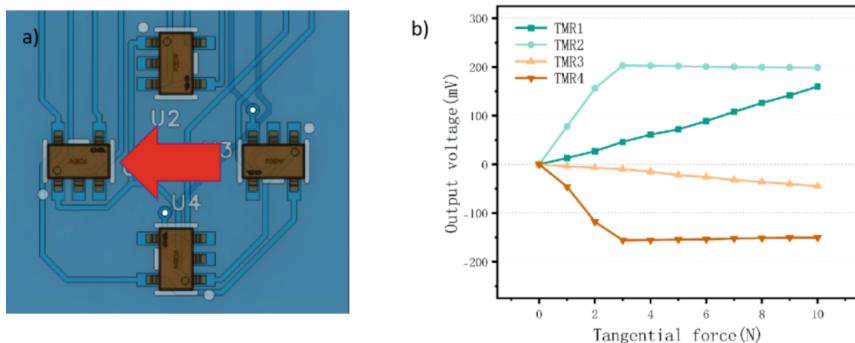


Fig. 10. (a) direction of tangential force (b) Output voltage of TMR array under tangential force

5 Conclusion

The hemispherical tactile sensor for three-dimensional force detection described in this paper is composed of NdFeB silicone elastomer magnetic cylinder, silicone elastomer shell, TMR2003 component array and printed circuit board, which can identify the direction and size of three-dimensional force. Using the radius of 2 mm, length of 7.5 mm NdFeB silicone elastomer magnetic cylinder as the sensing part, has the right size of the magnetic and has good flexibility. The use of silicone elastomer shell has good resilience in the range of force measurement, does not occur plastic deformation, has good service life and repeatability. The axial magnetized NdFeB silicone elastomer magnetic cylinder deforms after being subjected to pressure, causing the magnetic field intensity around the four TMR2003 components to change, and the output voltage waveform is obtained. According to the relationship between the input pressure and the output voltage, the recognition range of normal force and tangential force is 0–50 N and 0–10 N, the sensitivity of normal force in the range of 0–30 N and 30–50 N is 11.51 mV/N and 3.03 mV/N, and the tangential force is 15.82 mV/N, which can complete the three-dimensional force recognition function.

References

1. Billard, A., Kragic, D.: Trends and challenges in robot manipulation[J]. *Science*. **364**(6446), eaat8414 (2019)
2. Bartolozzi, C., Natale, L., Nori, F., et al.: Robots with a sense of touch[J]. *Nat. Mater.* **15**(9), 921–925 (2016)
3. Suzuki, H., Wood, R.J.: Origami-inspired miniature manipulator for teleoperated micro-surgery. *Nat. Mach. Intell.* **2**, 437–446 (2022)
4. Chi, C., Sun, X.G., Xue, N., et al.: Recent progress in technologies for tactile sensors[J]. *Sensors*. **18**(4), 948 (2018)
5. Abraira, V.E., Ginty, D.D.: The sensory neurons of touch[J]. *Neuron*. **79**(4), 618–639 (2013)
6. Johnson, K.: The roles and functions of cutaneous mechanoreceptors[J]. *Curr. Opin. Neurobiol.* **11**(4), 455–461 (2001)
7. Maksimovic, S., Nakatani, M., Baba, Y., et al.: Epidermal Merkel cells are mechanosensory cells that tune mammalian touch receptors[J]. *Nature*. **509**(7502), 617–621 (2014)
8. Wheat, H.E., Salo, L.M., Goodwin, A.W.: Cutaneous afferents from the Monkeys fingers: responses to tangential and normal forces[J]. *J. Neurophysiol.* **103**(2), 950–961 (2010)
9. Phillips, J.R., Johansson, R.S., Johnson, K.O.: Representation of braille characters in human nerve fibres[J]. *Exp. Brain Res.* **81**, 457–472 (1990)
10. Sotgiu, E., Aguiam, D.E., Calaza, C., et al.: Surface texture detection with a new sub-mm resolution flexible tactile capacitive sensor array for multimodal artificial finger[J]. *J. Microelectromech. Syst.* **29**(5), 629–636 (2020)
11. Sheng, L., Sicheng, C., Lei, Y.: Bimodal capacitive tactile sensor assisted by shield effect of triboelectric nanogenerator[J]. *Nano Energy*. **118**, 108946 (2023)
12. Zhang, Y., Lin, Z., Huang, X., et al.: Highly sensitive capacitive pressure sensor with elastic metallized sponge[J]. *Smart Mater. Struct.* **28**(10), 105023 (2019)
13. Ding, L., Xuan, S.H., Pei, L., et al.: Stress and magnetic field bimode detection sensors based on flexible CI/CNTs-PDMS sponges [J]. *ACS Appl. Mater. Interfaces*. **10**(36), 30774–30784 (2018)
14. Gupta, N., Adepu, V., Tathacharya, M.: Piezoresistive pressure sensor based on conjugated polymer framework for pedometer and smart tactile glove applications [J]. *Sensors Actuators A Phys.* **350**, 114139 (2023)
15. Li, H., Fu, Y., Guo, J., et al.: Impedance cell sensing technology from single-frequency to multifrequency: a review[J]. *IEEE Sensors J.* **24**(4), 4150–4160 (2024)
16. Yuan, Z., Yang, L., Yunna, S., Yanxin, Z., Guifu, D.: Recent advances in resistive sensor technology for tactile perception: a review[J]. *IEEE Sensors J.* **22**(16), 1 (2022)
17. Nassar, H., Khandelwal, G., Chirila, R.: Fully 3D printed piezoelectric pressure sensor for dynamic tactile sensing[J]. *Addit. Manuf.* **71**, 103601 (2023)
18. Yingxuan, Z., Feng, J., Xiaoyong, W., Dan, W., Wang, Y.: A piezoelectric tactile sensor for tissue stiffness detection with arbitrary contact angle[J]. *Sensors (Basel, Switzerland)*. **20**(22), E6607 (2020)
19. Zhang, Y., Ju, F., Wei, X., et al.: A piezoelectric tactile sensor for tissue stiffness detection with arbitrary contact angle[J]. *Sensors*. **20**(22), 6607 (2020)
20. Wang, J., Jiang, J., Zhang, et al.: Energy-efficient, fully flexible, high-performance tactile sensor based on piezotronic effect: piezoelectric signal amplified with organic field-effect transistors [J]. *Nano Energy*. **76**, 105050 (2020)
21. Zhang, H., Weng, L., Lin, G., et al.: Design of flexible tactile electronic skin array based on magnetic thin films and its position-sensing applications [J]. *IEEE Sensors J.* **24**(16), 25541–25549 (2024)



Time-Optimal Trajectory Planning for Robotic Arms Based on VPPSO Algorithm

Xiao Shi¹, Zhiyong He^{1(✉)}, and Yucai Zhou²

¹ College of Mechanical and Vehicle Engineering, Changsha University of Science & Technology, Changsha 410114, China
hzy73@sina.com

² School of Electrical and Information Engineering, Changsha University of Science & Technology, Changsha 410114, China

Abstract. To improve time-efficient trajectory planning for robotic arms and address limitations like premature convergence in standard PSO, this paper presents a robot trajectory planning method based on Velocity Pausing Particle Swarm Optimization is proposed. The method integrates a dynamic velocity pause strategy and adaptive parameter adjustment to balance global search with local refinement. We apply this algorithm to the IRB-2600 robot, building a trajectory model based on hybrid cubic-quintic polynomial interpolation under kinematic constraints. Results indicate that VPPSO significantly enhances convergence speed and optimization accuracy. Compared with traditional planning and classical PSO, VPPSO shortens trajectory duration by 47.74% and 18.55%, respectively, without compromising feasibility.

Keywords: Time-optimal · Trajectory planning · VPPSO · Polynomial interpolation

1 Introduction

Trajectory planning is the process of determining a robot's end-effector path based on task requirements, using forward and inverse kinematics to compute displacement, velocity, and acceleration over time [1]. With the robot's kinematic and dynamic parameters as input, joint motion variables can be derived. As a core aspect of robotic control, trajectory planning affects execution accuracy, motion smoothness, energy efficiency, and mechanical wear in industrial applications. An optimal trajectory should not only ensure task precision but also optimize time utilization and energy efficiency while maintaining system stability. Current research focuses on time-optimal, jerk-optimal, and energy-optimal planning methods [2, 3], among which time-optimal planning—aimed at minimizing motion time under constraints—is the most mature and widely studied.

Xu et al. [4] addressed issues of discontinuity and excessive jerk in cubic and quintic polynomial interpolation by proposing a 3-5-3 joint-space interpolation method,

enhancing planning precision and simplifying implementation in segmented trajectories. Kennedy et al. [5] introduced PSO, which gained popularity in trajectory planning for its simple structure, few parameters, and strong robustness. However, conventional PSO suffers from limited adaptability due to fixed parameters. To improve performance, Feng Bin et al. [6] used linearly decreasing inertia weight, while Peng Siwen et al. [7] introduced trigonometric-function-based inertia and dynamic learning factors, reducing trajectory optimization time by up to 45%. Shami et al. [8] proposed a VPPSO variant that introduces a velocity pause mechanism and dual-population structure, where the second population updates based on the global best, thus enhancing global exploration in high-dimensional constrained problems.

In this paper, a hybrid interpolation method combining cubic and quintic polynomials is used as the trajectory model, and a VPPSO algorithm is proposed for time-optimal planning. Compared with traditional PSO and conventional methods, the proposed VPPSO achieves faster convergence and higher accuracy in trajectory optimization.

2 Robot Modeling

The IRB-2600 is a six-axis industrial robot developed by ABB, with all joints designed as rotary actuators. To capture its kinematic structure, we use the Modified Denavit-Hartenberg (MD-H) method, it's an improvement over the classic D-H approach—which offers a more systematic way of assigning coordinate frames. For each joint, four parameters are defined: the link length (a), twist angle (α), joint angle (θ), and link offset (d). Together, these parameters describe how each link is positioned relative to its neighbors. Figs. 1 [9] and 2 illustrate the parameter setup and the robot's structural layout, while Table 1 lists the specific MD-H values used in the model. This formulation provides the foundation for computing both the robot's kinematics.

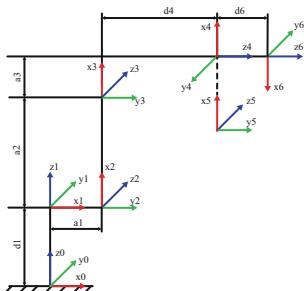


Fig. 1. D-H Parameter description

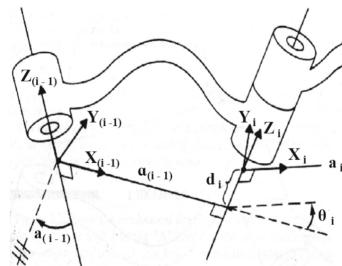


Fig. 2. Simplified diagram of robot

Table 1. Robot D-H parameters

Joint	$\theta/(\circ)$	$a/(\circ)$	a/mm	d/mm	Joint Limits/ (\circ)
1	θ_1	0	0	445	[-180,180]
2	θ_2-90	-90	150	0	[-95,110]
3	θ_3	0	700	0	[-180,75]
4	θ_4	-90	115	795	[-400,400]
5	θ_5	90	0	0	[-120,120]
6	θ_6-180	-90	0	85	[-400,400]

3 Robot Trajectory Planning

The robot motion trajectory is constructed using a hybrid cubic polynomial interpolation. By solving the interpolation functions simultaneously, the general system of equations for the 3-5-3 segmented polynomial interpolation is obtained, as shown in Eq. 1.

$$\begin{cases} \theta_{m1}(t) = \alpha_{m10} + \alpha_{m11}t_1 + \alpha_{m12}t_1^2 + \alpha_{m13}t_1^3 \\ \theta_{m2}(t) = \alpha_{m20} + \alpha_{m21}t_2 + \alpha_{m22}t_2^2 + \alpha_{m23}t_2^3 + \alpha_{m24}t_2^4 + \alpha_{m25}t_2^5 \\ \theta_{m3}(t) = \alpha_{m30} + \alpha_{m31}t_3 + \alpha_{m32}t_3^2 + \alpha_{m33}t_3^3 \end{cases} \quad (1)$$

$\theta_{m1}(t), \theta_{m2}(t), \theta_{m3}(t)$ represent the three-segment polynomial interpolation functions for joint m. t represents the motion time. Since the 3-5-3 polynomial interpolation does not possess the properties of traditional optimization methods, its time optimization can only be achieved through intelligent optimization methods [10, 11].

$$A = \begin{bmatrix} B & C & D \\ 0 & 0 & E \\ F & 0 & G \end{bmatrix} \quad (2)$$

In Eq. 2:

$$B = \begin{bmatrix} t_1^3 & t_1^2 & t_1 & 1 & 0 \\ 3t_1^2 & 2t_1 & 1 & 0 & 0 \\ 6t_1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & t_2^5 \\ 0 & 0 & 0 & 0 & 5t_2^4 \\ 0 & 0 & 0 & 0 & 20t_2^3 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ t_2^4 & t_2^3 & t_2^2 & t_2 & 1 \\ 4t_2^3 & 3t_2^2 & 2t_2 & 1 & 0 \\ 12t_2^2 & 6t_2 & 2 & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} 0 & t_3^3 & t_3^2 & t_3 & 1 \\ 0 & 3t_3^2 & 2t_3 & 1 & 0 \\ 0 & 6t_3 & 2 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$b = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ x_{i3} \ 0 \ 0 \ x_{i0} \ 0 \ 0 \ x_{i2} \ x_{i1}]^T \quad (3)$$

$$a = A^{-1} \cdot b = [a_{m13} \dots a_{m10} \ a_{m25} \dots a_{m20} \ a_{m33} \dots a_{m30}]^T \quad (4)$$

To ensure smooth and efficient motion, appropriate boundary conditions must be imposed. The solution process is outlined in Eqs. 2, 3, and 4, where Eq. 3 defines the four path points, and Eq. 4 expresses the coefficient vector for trajectory planning.

4 Time-Optimal Trajectory Planning Based on the Velocity Pausing Particle Swarm Optimization Algorithm

The Particle Swarm Optimization (PSO) algorithm is inspired by bird flocking behavior during foraging [12]. In PSO, each particle explores the search space to find optimal solutions, recording its best position (individual best) over iterations. These best positions are shared within the swarm to guide the search toward the global optimum. Each particle is characterized by two attributes: velocity, indicating its movement toward the optimum, and position, representing its location in the search space. The velocity and position update equations for the i -th particle are defined as follows:

$$\begin{cases} v_{id}(t+1) = w \times v_{id}(t) + c_1 \times r_1 \times (p_{id}(t) - x_{id}(t)) + c_2 \times r_2 \times (p_{gd}(t) - x_{id}(t)) \\ x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \end{cases} \quad (5)$$

In the equation, d represents the dimension, t is the current iteration number, w is the inertia weight, c_1 and c_2 are the individual and social learning factors, respectively, r_1 and r_2 are uniformly distributed random numbers between $[0, 1]$. p_{id} represents the individual best fitness, p_{gd} represents the global best fitness, and x and v are the position and velocity components of the particle in the d -dimensional space.

The VPPSO introduces an improved strategy known as the velocity pause mechanism. particles are not required to update their velocity at each iteration. This approach preserves the particle's inertia to some extent. The mathematical expression for the velocity pause can be represented as follows:

$$v_{id}(t+1) = \begin{cases} v_{id}(t) & \text{if } r < a \\ v_{id}(t)^{r_3 \times a(t)} + c_1 \times r_4 \times (p_{id}(t) - x_{id}(t)) + c_2 \times r_5 \times (p_{gd}(t) - x_{id}(t)) & \text{otherwise} \end{cases} \quad (6)$$

$$a(t) = \exp\left(-((b \times t)/T)^b\right) \quad (7)$$

in Eq. 7, a is the velocity pause parameter, $a(t)$ is an exponential function that changes with each iteration, b is a constant, and T is the maximum number of iterations. When the value of a is bigger than 1, the particle's velocity will tend to remain constant. When a falls within the range $[0, 1]$, the particle's velocity will be updated according to the above equation.

In the PSO, the learning factors c_1 and c_2 are typically set to fixed values, which can lead the algorithm to become stuck in local optima. To improve the algorithm's performance, this paper introduces a strategy of dynamically adjusting the learning factors. In the early stages of the search, a larger individual learning factor c_1 helps enhance the particle's exploration ability, preventing premature convergence to a local optimum. In the later stages, a larger social learning factor c_2 accelerates the particle's convergence toward the global optimum solution [13]. Therefore, this paper adopts nonlinear varying learning factors to replace the traditional fixed values.

Shami et al. [8] stated in their article that as a gradually decreases from 1 to 0.3, the performance of VPPSO improves. However, when a is further reduced to below 0.3, the performance begins to decline. This indicates that the value of a significantly affects the performance of VPPSO, and there exists an optimal range for a . In practical applications, the selection of the velocity pause parameter a should be adjusted according to the specific problem. Testing the low-dimensional complex functions F20 and F23 from the CEC testing function benchmark suite, the expression is shown in the Table 2. Among the table, Fun represents the function formula, dim represents the function dimension, bound represents the domain, and Fmin represents the global minimum value of the function.

Table 2. Test function

Fun	Dim	Bound	Fmin
$F_6(\mathbf{x}) = \sum_{i=1}^n x_i + 0.5 ^2$	30	$[-100, 100]^n$	0
$F_{10}(\mathbf{x}) = -20 \cdot \exp\left(-0.2 \cdot \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}\right)$ $- \exp\left(\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i)\right) + 20 + e$	30	$[-32, 32]^n$	0
$F_{20}(\mathbf{x}) = - \sum_{i=1}^4 c_i \cdot \exp\left(- \sum_{j=1}^6 a_{ij} \cdot (x_j - p_{ij})^2\right)$	6	$[0, 1]^n$	-3.32
$F_{23}(\mathbf{x}) = - \sum_{i=1}^{10} \left(\ \mathbf{x} - \mathbf{a}_i\ ^2 + c_i\right)^{-1}$	4	$[0, 10]^n$	-10

The velocity pause coefficient a in Eq. 7 varies from 1 to 0, as shown in the Fig. 3. When $a = 1$, the algorithm updates velocity and position according to the standard PSO method. For F20, the best global optimum is achieved when $a = 0.1$; for F23, the best result is obtained when $a = 0.3$.

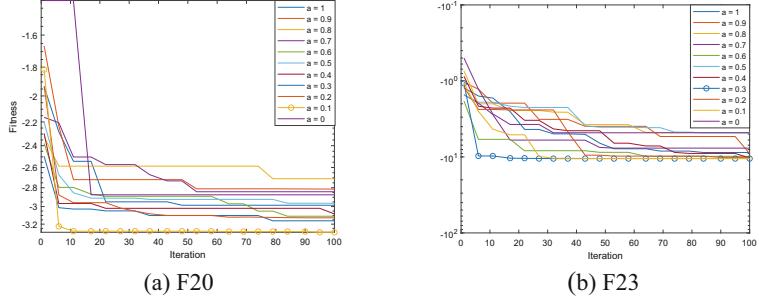


Fig. 3. The influence of different a on the algorithm

Shi et al. [14] first introduced the concept of inertia weight in PSO, noting that a larger weight promotes global exploration, while a smaller one favors local exploitation. Building on this, we propose an adaptive inertia weight strategy, enabling particles to dynamically adjust inertia based on the search stage and optimization needs.

To enhance its application in robotic trajectory planning, this paper also improves the update formulas for the learning factors and inertia weight. The equations are as follows:

$$\begin{cases} c_1 = c_{1 \max} - (c_{1 \max} - c_{1 \min}) \times (t/T)^{1.5} \\ c_2 = c_{2 \min} + (c_{2 \max} - c_{2 \min}) \times (t/T)^{1.5} \end{cases} \quad (8)$$

$$a(t) = \exp(-(2.5 \times t/T)^2 \times 2.5) \quad (9)$$

The velocity and position updates of the first group of particles in VPPSO are shown in Eqs. 10 and 11:

$$v_{id}(t+1) = \begin{cases} v_{id}(t) & \text{if } r < \alpha \\ v_{id}(t+1) \text{ as in (9)} & \text{otherwise} \end{cases} \quad (10)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (11)$$

Each particle in the second group updates its position as follows:

$$x_{id}(t+1) = \begin{cases} p_{gd}(t) + r_6 \times a(t) \times |p_{gd}(t)|^{a(t)} & \text{if } r_7 < 0.5 \\ p_{gd}(t) - r_8 \times a(t) \times |p_{gd}(t)|^{a(t)} & \text{otherwise} \end{cases} \quad (12)$$

In VPPSO, the optimization function is defined, and all particles are initialized with velocity and position values. The swarm is divided into two groups with distinct update strategies. The first group follows an improved PSO update mechanism, incorporating a modified velocity term and the velocity pause strategy. Their velocity and position are synchronously updated via Eq. 6. The second group uses a simplified strategy, updating positions based solely on the global best solution. The fitness function shown in Eq. 13, evaluates the performance of each trajectory, where m denotes the m -th joint.

$$f(t) = \min(t_{m1} + t_{m2} + t_{m3}) \quad (13)$$

To optimize the robotic arm's running time using VPPSO, initial and final velocities and accelerations must be zero to ensure smooth polynomial interpolation. Thus, the constraint in Eq. 14 is imposed. The optimization process flowchart is shown in Fig. 4.

$$\begin{cases} \max\{|v_{id}|\} \leq v_{id \max} \\ \max\{|a_{id}|\} \leq a_{id \max} \end{cases} \quad (14)$$

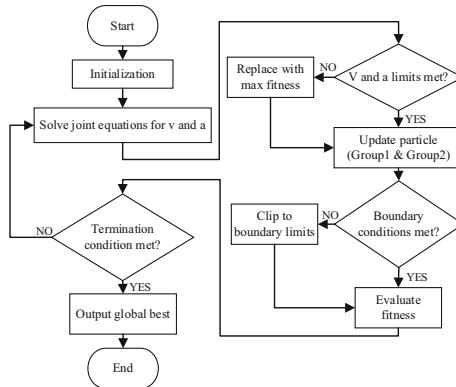


Fig. 4. Flow chart of the VPPSO

Thirty independent runs of VPPSO, PSO, IPSO, and CPSO were performed on CEC2017 benchmark functions F6, F10, F20, and F23 to explore global minima, as summarized in Table 3. Best represents the best result, Mean represents the average, and Std is the standard deviation.

Table 3. Results of function optimization

Fun		VPPSO	PSO	IPSO	CPSO
F6	Best	3.0611×10^{-5}	5.3338	12.886	76.681
	Mean	2.2354×10^{-2}	21.377	107.75	123.48
	Std	2.1994×10^{-2}	10.711	106.38	23.776
F10	Best	4.4409×10^{-16}	2.0418	4.4415	6.7659
	Mean	4.4409×10^{-16}	3.0535	5.9204	7.5338
	Std	0	0.58149	1.0519	0.40212
F20	Best	-3.3220	-3.3220	-3.3220	-3.1878
	Mean	-3.2320	-3.2582	-3.2782	-3.0017
	Std	9.0312×10^{-2}	6.6845×10^{-2}	5.8586×10^{-2}	8.7945×10^{-2}
F23	Best	-10.536	-10.536	-10.536	-9.5431
	Mean	-9.9397	-7.5778	-8.8551	-4.8691

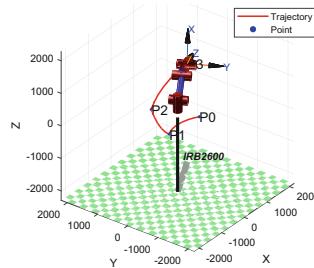
(continued)

Table 3. (*continued*)

Fun		VPPSO	PSO	IPSO	CPSO
Std		2.0201	3.6536	3.1335	1.6903

5 Simulation Experiment Analysis

Based on the parameters in Table 1, a robot simulation model was built using MATLAB Robotics Toolbox. The hybrid interpolated motion trajectory was generated via polynomial interpolation. Simulation results are shown in Fig. 5.

**Fig. 5.** Robot model and trajectory

To verify VPPSO's time optimization in robot trajectory planning, simulation experiments comparing PSO and VPPSO were conducted using the IRB-2600 robot and MATLAB Robotics Toolbox. The robot's end effector sequentially passes through pre-set waypoints: $P_0 = [965, 0, -435]$, $P_1 = [154, 389, -723]$, $P_2 = [564, 1291, -478]$, $P_f = [1397, 733, 728]$. Inverse kinematics converts these Cartesian coordinates into corresponding joint angles, listed in Table 4.

In VPPSO, the first group population size is 25, and the second group is 50. The maximum iterations are 200, with a dimension of 3. The individual and group learning factors c_1 and c_2 range between 0.5 and 2. Joint velocity limits follow the IRB-2600 manual: [150, 160, 170, 320, 400, 460]. Acceleration limits, unspecified in the manual, are set equal to velocity limits. The parameter a is set to 0.5.

Table 4. Joint angle values(rad)

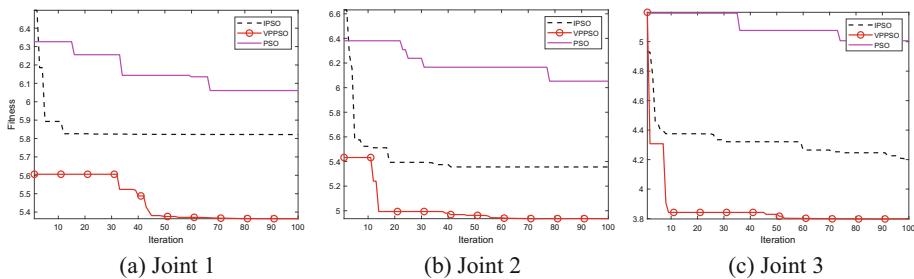
Joint	P0	P1	P2	Pf
1	0	1.1940	1.1589	0.4832
2	0	0.5198	0.4232	-0.6626
3	0	0.0042	-1.0552	-0.6036
4	0	0.0440	-0.0038	-0.0126

(continued)

Table 4. (continued)

Joint	P0	P1	P2	Pf
5	0	-0.0201	-0.1210	-0.0797
6	0	0.0139	0.1035	-0.2001

Implement the algorithm based on the designed flowchart and run simulations to obtain the fitness curves of the objective functions for each joint, as shown in Fig. 6. By comparing the trajectory planning results of VPPSO and IPSO, PSO on the same path points, the advantages of VPPSO in improving optimization efficiency and reducing runtime are evident. Figure 6 shows that VPPSO achieves significantly faster convergence and better optimization of the robot arm's joint interpolation times. The optimized runtimes for each joint's three-segment polynomial trajectory are summarized in Table 5.

**Fig. 6.** Joint fitness variation curves**Table 5.** Time optimized by PSO and VPPSO (s)

Joint	T1(PSO)	T2(PSO)	T3(PSO)	T1(VPPSO)	T2(VPPSO)	T3(VPPSO)
1	1.9409	2.6490	1.9783	2.0564	1.9603	1.3440
2	1.0994	2.4477	2.7978	1.1862	1.7938	1.9564
3	0.1000	2.1454	3.1092	0.1730	2.2582	1.3668
4	1.1652	0.6251	0.3503	0.3295	0.3647	0.1000
5	0.1866	0.5964	1.0478	0.1314	0.3893	0.2803
6	0.2281	0.8111	0.6874	0.1020	0.6571	0.7403

As shown in Fig. 7, the time-optimized trajectory simulation of the robotic arm was performed in MATLAB. Where q represents the change in robot joint angle, v represents the velocity, and a represents the change in acceleration. The times were used to generate the end-effector's motion trajectory through the preset points, producing optimized curves of angular displacement, velocity, and acceleration for each joint over time.

The maximum joint running time using VPPSO algorithm in Table 5 is 2.0564 s, 2.2582 s, and 1.9564 s, while the time using PSO algorithm is 1.9409 s, 2.6490 s, and 3.1092s. The total time for the three segments is 6.2710 s and 7.6991 s respectively, a decrease of 18.55% compared to the previous values. Overall, this is equivalent to a 5.729 s reduction in general planning time, approximately 47.74% reduction in time.

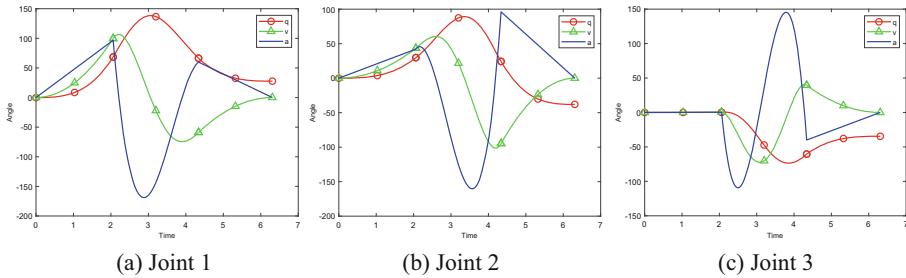


Fig. 7. Joint displacement, velocity and acceleration curves

Due to space limitations, this paper only presents the iterative curves of the first three joints, including the changes in angle, velocity, and acceleration, as well as the comparison between VPPSO and PSO.

6 Conclusion

This paper establishes forward and inverse kinematic models of the IRB-2600 robot and performs time optimization using the velocity pausing particle swarm optimization algorithm. Simulations were conducted with MATLAB's Robotics Toolbox. The key findings are:

1. Using the improved D-H method, a complete robot model was built and validated via simulation, confirming the arm's structural rationality and motion capability.
2. A 3–5–3 combined polynomial interpolation was applied in joint space for trajectory planning. VPPSO achieved time-optimal planning while ensuring smooth motion and satisfying constraints, demonstrating practicality.
3. Compared to PSO, VPPSO introduces a dual-population collaborative iteration, an improved velocity update, nonlinear inertia weights, and dynamic learning factors, better suiting robotic trajectory planning. Simulations show VPPSO reduces motion time by 18.55% versus PSO and 47.74% versus the initial plan, significantly improving operational efficiency.

References

1. Li, W., Deng, Z., Ge, J., et al.: Research progress on robot joint space trajectory planning[J]. Mach. Des. Manuf. **51**(10), 15–23 (2022)

2. Wurll, C., Henrich, D.: Point-to-point and multi-goal path planning for industrial robots[J]. *J. Robot. Syst.* **18**(8), 445–461 (2001)
3. He, C., Zhu, R., Huang, T., et al.: Research progress on trajectory planning and optimization of industrial robots[J]. *Mod. Manuf. Eng.* **514**(07), 150–159 (2023)
4. Xiang-Rong, X., Xiao-Gang, W., Feng, Q.: Trajectory planning of robot manipulators by using spline function approach[C]. In: Proceedings of the 3rd World Congress on Intelligent Control and Automation, No. 00EX393, vol. 2, pp. 1215–1219. IEEE (2000)
5. Kennedy, J., Eberhart, R.: Particle swarm optimization[C]. In: Proceedings of ICNN'95-International Conference on Neural Networks, vol. 4, pp. 1942–1948. IEEE (1995)
6. Feng, B., Liu, F., Zheng, L.: Optimal motion trajectory planning of robot joint space based on particle swarm optimization[J]. *Modular Mach. Tool Autom. Manuf. Technique.* **05**, 1–4 (2022)
7. Peng, S., Liang, K., Pan, L., et al.: Time optimal trajectory planning of manipulator based on improved particle swarm optimization algorithm[J]. *Modular Mach. Tool Autom. Manuf. Technique.* **06**, 31–34 (2023)
8. Shami, T.M., Mirjalili, S., Al-Eryani, Y., et al.: Velocity pausing particle swarm optimization: a novel variant for global optimization[J]. *Neural Comput. Applic.* **35**(12), 9193–9223 (2023)
9. Edition, T., Craig, J.J.: *Introduction to Robotics*[J], 3rd edn. China Machine Press, Beijing (2005)
10. Li, X., Cui, F., Luo, H., et al.: Kinematics analysis and simulation of six degrees of freedom industrial robots[J]. *Manuf. Autom.* **44**(7), 7–10 (2022)
11. Fu, R., Ju, H.: Time-optimal trajectory planning algorithm for manipulator based on PSO[J]. *Inf. Control.* **40**(6), 802–808 (2011)
12. Saramago, S.F.P., Junior, V.S.: Optimal trajectory planning of robot manipulators in the presence of moving obstacles[J]. *Mech. Mach. Theory.* **35**(8), 1079–1094 (2000)
13. Ma, B., Wu, Z.: Research on supply chain network equilibrium model based on improved particle swarm optimization[J]. *Oper. Res. Manage. Sci.* **29**(2), 122 (2020)
14. Shi, Y., Eberhart, R.: A modified particle swarm optimizer[C]. 1998 IEEE international conference on evolutionary computation proceedings. In: IEEE World Congress on Computational Intelligence, Cat. No. 98TH8360. IEEE, pp. 69–73 (1998)



Multi-spacecraft Orbital Game Combined with Potential Field Function

Zihan Li^(✉) and Shan Lu^(✉)

Shanghai Aerospace Control Technology Institute, Shanghai 201109, China
1421605066@qq.com

Abstract. The space orbital game involves the control problems of the spacecraft in the complex space environment. This paper addresses the fuel optimization problem of the escaper in multi-spacecraft orbital games and proposes an adaptive control strategy based on potential field functions. By establishing a dynamic game model of the chaser-escaper, the chaser is modeled as a repulsive source, and the repulsive field generated by the potential field function is used to determine the escaper's escape direction. Different from the fixed acceleration strategy, this method designs an adaptive adjustment mechanism based on the relative motion state to dynamically optimize the escaper's acceleration magnitude. Numerical simulation results show that within the specified confrontation time threshold, this strategy can ensure that the escaper maintains a safe distance and is not captured, while reducing fuel consumption by 13.2% compared to the full acceleration escape scheme.

Keywords: Orbital game · Potential function · Avoidance strategy

1 Introduction

As global space activities accelerate and competition for low-Earth orbit (LEO) resources grows, orbital space is becoming highly congested. Space debris, defunct spacecraft and adversarial spacecraft with maneuvering capability pose serious threats to our in-orbit assets, which highlight the importance of space security protection. In view of non-adversarial threats such as space debris, it is necessary to implement orbital maneuver prediction and collision avoidance based on perceived information. For adversarial targets with maneuvering abilities, safe evasive strategies are required.

Artificial Potential Field method (APF), a classical path planning method, has been widely used in robot, unmanned aerial vehicles (UAV) and other fields. Literature [1] improved the potential function based on gradient potential energy, effectively enhancing the robot's mobility in complex mountainous environments. Literature [2] integrated model predictive control with the artificial potential function to design trajectory planning algorithms for multiple drones in complex scenarios. Literature [3] constructed the potential field function with a continuously differentiable logarithmic function, eliminating local minima traps. Literature [4] focused on synchronizing path parameters and combined potential functions with obstacle constraints to correct the robot's control law.

Literature [5] utilized repulsive functions to reconstruct the guidance rate for trajectory tracking of autonomous surface vessels. Literature [6, 7] also proposed solutions to the local optimum problem of the APF in various application scenarios.

The spacecraft orbital game problem involves one or more spacecraft achieving specific mission objectives, such as pursuit-evasion, interception, and rendezvous, under the constraints of orbital dynamics through active control strategies. In the context of spacecraft evasion, literature [8] improved the potential field function using the Frenet coordinate system, effectively preventing trajectory deviation during evasion. Literature [9] introduced a correction term to enhance the artificial potential function guidance method, enabling the tracking spacecraft to avoid dynamic obstacles while approaching the target spacecraft. Literature [10] designed a control law based on geometric potential fields for space obstacle avoidance in scenarios with incomplete information. Although these methods have addressed the issue of the artificial potential function falling into local minima, which can cause path planning failure or oscillation, they have primarily focused on static obstacle avoidance tasks, with no applications in dynamic scenarios.

In this paper, APF method is introduced into the field of multi-spacecraft orbit game, a two-on-one scenario is constructed, and an escape spacecraft control strategy based on adaptive artificial potential field is designed to enhance the on-orbit safety protection capability.

2 Basic Principles

Coordinate System

Geocentric inertial coordinate system: With the center of the earth as the origin, the fundamentals are flat equatorial plane, z-axis point to the Earth's North Pole along the Earth's axis of rotation, x-axis and y-axis are vertical and located in the equatorial plane, x-axis points to the equinox, three axes constitute the right hand orthogonal Coordinate system (oxyz, Si).

The second orbit coordinate system, also known as the VVLH coordinate system (Vehicle Velocity Local Horizontal): Taking the spacecraft's center of mass as the origin, oz points to the center of the earth along the spacecraft's center of mass, oy along the negative normal of the spacecraft's orbit, ox along the spacecraft's track, and three axes form a right-handed orthogonal coordinate system (oxyz, So) (Fig. 1).

Equation of Motion

Assuming that the reference orbit is a circular orbit and the earth's higher-order perturbation term is ignored, the relative motion model of space vehicle based on Clohessy-Wiltshire equation (C-W equation) is obtained by Taylor expansion and simplified:

$$\begin{cases} \ddot{x} - 2\omega\dot{z} = a_x \\ \ddot{y} + \omega^2y = a_y \\ \ddot{z} - 3\omega^2z + 2\omega\dot{x} = a_z \end{cases} \quad (1)$$

where: $\{(x, y, z \in \mathbb{R}^3)\}$ is the three-dimensional position of the spacecraft in the orbital coordinate system of the target spacecraft, $\{\dot{x}, \dot{y}, \dot{z}\}$ is the velocity vector of the spacecraft

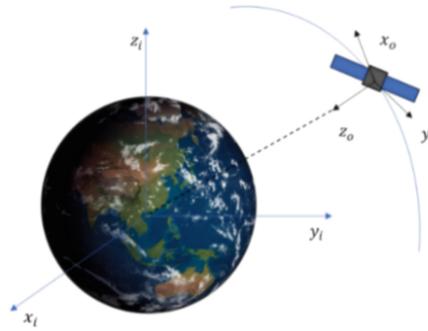


Fig. 1. Inertial coordinate system and orbit coordinate system

in the orbital coordinate system of the target spacecraft, $\omega = \sqrt{\mu/R^3}$ is the orbital angular velocity, $\mu = 3.986 \times 10^{14} \text{m}^3/\text{s}^2$ and a_i is the axis component of the control acceleration. In the case of circular orbit or near-circular orbit, the relative distance between spacecraft is small, the calculation accuracy can meet most of the dynamic requirements.

The relative state quantity of the spacecraft is $X = \{x, y, z, \dot{x}, \dot{y}, \dot{z}\}^T$, the control acceleration is $U = \{a_x, a_y, a_z\}^T$, then the spacecraft relative motion state space expression is:

$$\dot{X} = AX + BU \quad (2)$$

where: A is the state matrix, B is the input matrix.

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 2\omega \\ 0 -\omega^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3\omega^2 - 2\omega & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The exact discrete solution of the Clohessy-Wiltshire equation can be obtained as follows:

$$X^{(k+1)} = e^{At}X^{(k)} + \left(\int_0^t e^{A(t-\tau)} B d\tau \right) u^{(k)} \quad (4)$$

3 Space Orbital Game Control Strategy

3.1 Strategy Formulation

Take a virtual spacecraft in the orbit of 400km as the origin of the orbital coordinate system, and then in the opposite direction, the initial state information of the escaper and chasers are given: $X_E = [r_E, v_E]^T$, $X_{Ci} = [r_{Ci}, v_{Ci}]^T$, $i = 1, 2$.

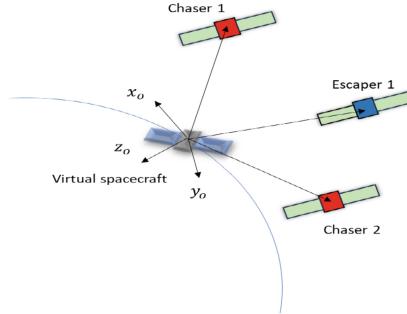


Fig. 2. Relative states

The following is the design of the control strategy of the escaper and the chaser: design the sight pointing control scheme of the chaser; For the escaper, the chaser is regarded as the repulsive force point the acceleration direction of the escaper is taken according to the force direction of the potential function. In addition, the acceleration grade is divided according to the relative distance between the chaser and the escaper to optimize the fuel consumption in the escape process (Fig. 2).

Chaser's Strategy Design

Firstly, the relative position between the chaser and the escaper is calculated in the orbital coordinate system of the virtual spacecraft:

$$\mathbf{r}_{rel,i} = \mathbf{r}_E - \mathbf{r}_{Ci} \quad (5)$$

where: $\mathbf{r}_{rel,i}$ is the relative position vector of the chaser pointing at the escaper, \mathbf{r}_E is the current position of the escaper, and \mathbf{r}_{Ci} is the current position of the chaser.

Then, calculate the chaser acceleration along the line of sight:

$$\mathbf{a}_{Ci} = \frac{\mathbf{r}_{rel,i}}{\|\mathbf{r}_{rel,i}\|} a_{\max} c \quad (6)$$

where: $a_{\max} c$ is the maximum acceleration of the chaser.

In order to prevent acceleration spillover, the calculated acceleration is limited to obtain the final acceleration of the chaser:

$$\mathbf{a}_{Ci} = \frac{\mathbf{a}_{Ci}}{\|\mathbf{a}_{Ci}\|} a_{\max} c \quad (7)$$

Escaper's Strategy Design

Firstly, taking the resultant force direction of the potential field function as its acceleration direction, the dynamic repulsive force field is calculated as follows [11]:

$$\mathbf{F}_{req} = \sum_{i=1}^2 \left[k_{req}(\rho_i) \left(\frac{1}{\rho} - \frac{1}{\rho_o(\rho_i)} \right) \frac{\mathbf{r}_E - \mathbf{r}_{Ci}}{\rho_i^3} \right] \quad (8)$$

where k_{req} : is the repulsion constant; ρ_i is the distance between the chaser and the escaper; ρ_o is the influence radius of the chaser; r_E is the position vector of the escaper; r_{Ci} is the position vector of the chasers;

The acceleration of the escape vehicle is obtained:

$$a_E = \frac{F_{req}}{\|F_{req}\|} a_{\max E} \quad (9)$$

where: $a_{\max E}$ is the maximum acceleration of the escape device.

In addition, different warning areas are set for the escaper. When the proximity distance between any chaser and the escaper reaches the set conditions, the escaper adopts different acceleration to maneuver:

$$\begin{aligned} 0 &\leq \|r_{rel,i}\| \leq r_1, \quad a_{E,current} = a_{\max E} \\ r_1 &\leq \|r_{rel,i}\| \leq r_2, \quad a_{E,current} = 0.8a_{\max E} \\ r_2 &\leq \|r_{rel,i}\|, \quad a_{E,current} = 0.6a_{\max E} \end{aligned} \quad (10)$$

The acceleration of the escape vehicle is obtained:

$$a_E = \frac{F_{req}}{\|F_{req}\|} a_{E,current} \quad (11)$$

Other Conditions

Approximate fuel consumption of each spacecraft in terms of cumulative velocity increment:

$$\Delta V = \sum \|a_k\| \cdot \Delta t \quad (12)$$

The acquisition decision condition is that the relative distance between any chaser and the escaper meets:

$$\|r_{rel,i}\| \leq r_{capture} \quad (13)$$

The capture is successful, $r_{capture}$ is the default capture radius.

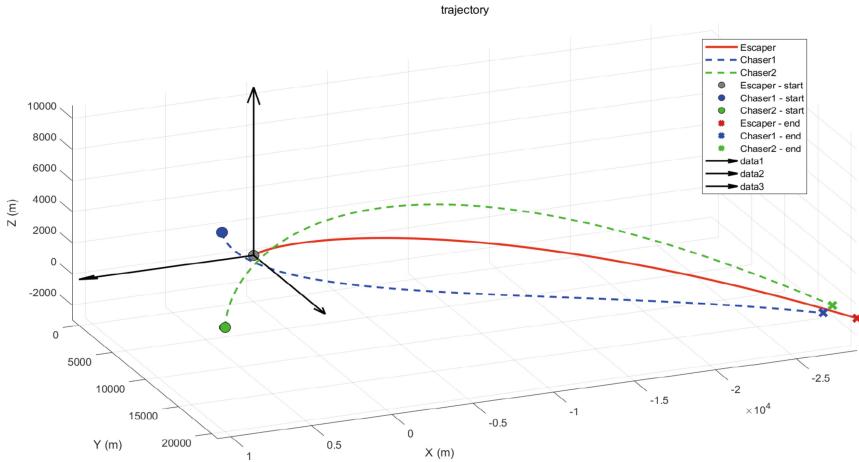
3.2 Simulation

The upper limit of the acceleration of the chaser is set at 0.03 m/s^2 , the acceleration of the escaper is 0.02 m/s^2 , the repulsion coefficient is $k_{req} = 5$, the influence range of the basic repulsion is $\rho_0 = 10000 \text{ m}$, the acquisition radius of the chaser is 100 m, and the reference orbit altitude is 400 km. The initial simulation conditions are set as follows (modified from literature [12]):

The four-order Runge-Kutta method is used for integration, the step length is 1 s, and the simulation time is 1500 s (Table 1).

Table 1. Initial conditions

Parameter	Escaper	Chaser 1	Chaser 2
Position (m)	(0, 0, 0)	(3300, 3500, 3200)	(3200, 3600, -3000)
Velocity (m/s)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)

**Fig. 3.** 3D game trajectory

Firstly, a two-on-one game with fixed acceleration is performed, and the three-dimensional game trajectory is obtained as follows (Fig. 3):

In the following Figures, we take the escaper as the reference to illustrate the proximity of the chasers.

Figures 4 and 5 shows the relative distance evolution curve of the two chasers relative to the escaper during the multi-spacecraft orbit game. Initially (0–600 s), the relative distance between the chaser and the escaper is rapidly shortened by active maneuvering, and the minimum relative distance reaches the extreme point in the game period. In the middle of the game (600–1000 s), the two sides are in the fierce confrontation period of the offensive and defensive strategy game, and the relative distance tends to be stable. At the end of the game (after 1000 s), the chaser leverages its acceleration vector adjustment advantage to gradually penetrate the escaper's defense, achieving quadratic convergence of the relative distance. Figures 6 and 7 show the relative velocity change curve of the chasers. Although the chasers can increase the relative velocity amplitude in the fast approach stage, its velocity vector direction adjustment capability limits this increase. To maintain offense-defense balance in the middle of the game, chasers reduce velocity vector offset, leading to a significant decline in relative velocity amplitude. Figure 8 gives comparative propellant consumption data. Under full acceleration control, the cumulative fuel consumption of the escape vehicle is 30 m/s, while the chaser 1 and chaser 2 reach 45 m/s.

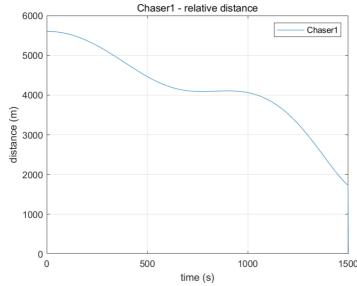


Fig. 4. Chaser1 - relative distance

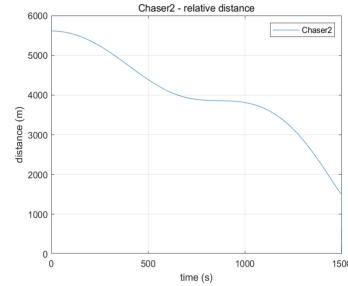


Fig. 5. Chaser2 - relative distance

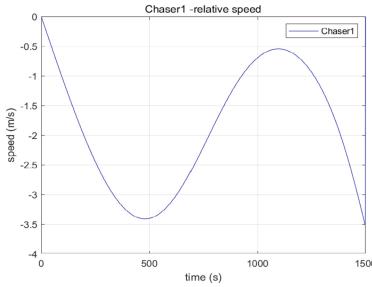


Fig. 6. Chaser1 - relative speed

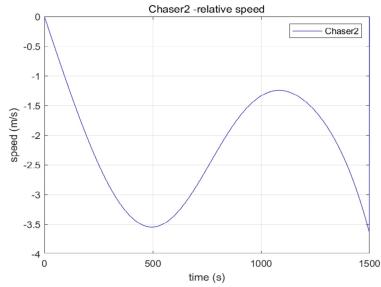


Fig. 7. Chaser2 - relative speed

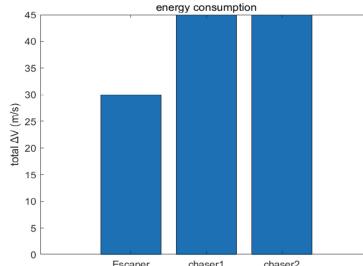


Fig. 8. Energy consumption

Then set the warning area of the escaler ($r_1=3000$ m, $r_2=5000$ m). The escaler's maneuvering acceleration is dynamically adjusted with the real-time relative distance of the chaser to form a hierarchical response mechanism.

Compared with the fixed acceleration scenario, in the dynamic acceleration strategy, the relative distance between the chaser and the escaler in the early stage of the game (0–700 s), presents a rapid convergence characteristic, enters the intense game stage at 700 s–1000 s, and then rapidly approaches the escaler again (Figs. 9 and 10). The amplitude of the velocity change curve in Fig. 11 and Fig. 12 is the same. The ratio increases and the relative velocity vector also enters the adjustment period. Due to the setting of the warning area, the acceleration of the escaler responds in stages as the

relative distance of the chaser decreases, and the maximum acceleration is kept away when the chaser is close (Fig. 13). The fuel consumption of the escaper is 26.03 m/s, and the chaser 1 and chaser 2 are 45 m/s (Fig. 14). During the specified simulation time, fuel consumption is reduced by 13.2% compared to the full acceleration scenario.

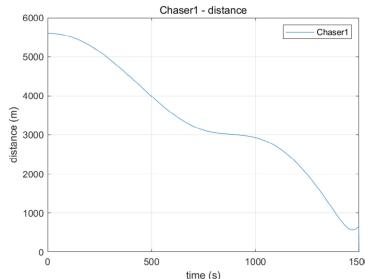


Fig. 9. Chaser1 - relative distance

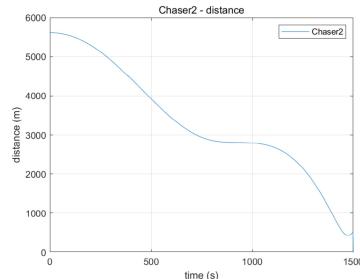


Fig. 10. Chaser2 - relative distance

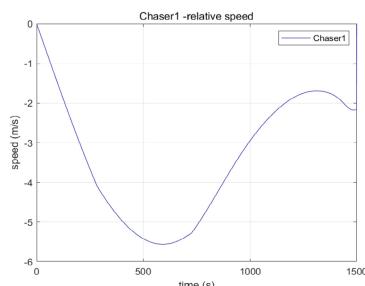


Fig. 11. Chaser1 - relative speed

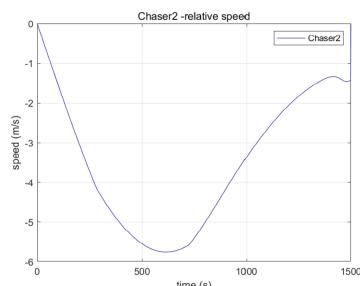


Fig. 12. Chaser2 - relative speed

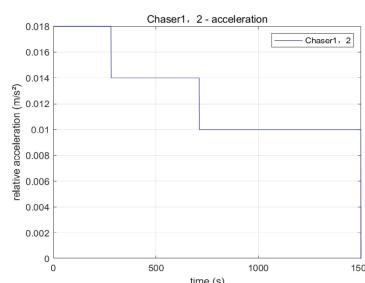


Fig. 13. Chaser's relative acceleration

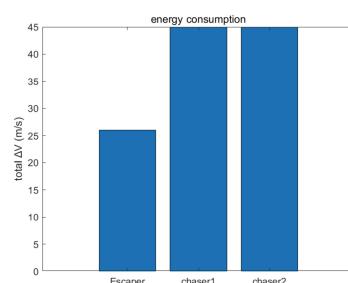


Fig. 14. Energy consumption

4 Conclusion

This paper proposes a layered control architecture of multi-spacecraft orbit game based on potential field function. The direction of the repulsive force is taken as the escape direction of the escaper. By designing a range-adaptive acceleration adjustment mechanism, and the escape maneuvering strategy is optimized. Within the specified simulation time, the escaper is not captured and the fuel consumption of the escaper in the dynamic acceleration game scenario is reduced by 13.2% compared with the escape process maintaining full acceleration.

References

1. Chen, R., Zhu, Y., Yang, L., He, J., Tang, Y.: Path planning method for quadruped robot in complex mountain environment. *J. Chin. Inertial Technol.* (12), 1250–1257, 1262 (2024)
2. Xian, B., Song, N.: A multiple UAVs path planning method based on model predictive control and improved artificial potential field. *Control Decis.* **07**, 2133–2141 (2024)
3. Fan, S., Qi, Q., Lu, K., et al.: Autonomous collision avoidance technique of cruise missiles based on modified artificial potential method. *Trans. Beijing Instit. Technol.* **38**(08), 828–834 (2018)
4. Su, Z., Liu, D., Sun, D., et al.: Path parameter consensus-based formation and obstacle avoidance control of autonomous underwater vehicles. *Chin. J. Intell. Sci. Technol.* **4**(04), 533–541 (2022)
5. Li, W., Zhang, Y.-C., Ning, J., et al.: Collision avoidance of under-actuated unmanned surface vehicles formation with improved artificial potential field method. *Control Decis.*, **40**(01), 252–260 (2025)
6. Zhang, G.Q., Han, J., Li, J.Q., et al.: APF-based intelligent navigation approach for USV in presence of mixed potential directions: guidance and control design. *Ocean Eng.* **260**, 111972 (2022)
7. Mohanan, M.G., Salgoanka, R.A.: A survey of robotic motion planning in dynamic environments. *Robot. Auton. Syst.* **100**, 171–185 (2018)
8. Liu, B.Y., Ye, X.B., Fang, S.L., et al.: Autonomous planning of on-orbit evasion path based on Frenet and improved artificial potential field. *J Beijing Univ. Aeron. Astron.* **47**(4), 731–741 (2021). (in Chinese)
9. Peng, G., Jianjun, L.: Adaptive artificial potential function guidance for dynamic obstacle avoidance of spacecraft. *Chin. Space Sci. Technol.* **32**(05), 1–8 (2012)
10. Hao, Z., Zhaojun, D.: Spacecraft cluster obstacle avoidance control adapted to the specific structure of obstacle. *Aerospace Control.* **40**(06), 23–29 (2022)
11. Guangjian, T., Jiyang, D., Jin, Y., Ning, W.: Multi-UAV trajectory planning based on adaptive segmented potential field method. *J. Syst. Simul.* **11**, 2368–2376 (2022)
12. Zhu, H.: Optimal control of spacecraft orbital pursuit-evasion based on differential game. National University of Defense Technology (2017)



Research on the Design and Control Method of a Controllable Deformable Wheeled Stair-Climbing Robot

Dongsheng Shao¹, Lunfei Liang^{2(✉)}, Xiaoxiong Du¹, Jiansheng Zhou¹, Shuo Cao¹,
and Yu Liu²

¹ Jianghuai Advance Technology Center, Hefei 230000, China

² School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China

lianglunfei@hotmail.com

Abstract. Aiming at the difficulties of conventional mobile robots in passing through step-like obstacles, a design method for a controllable deformable wheeled stair-climbing robot is proposed. A wheel structure with a variable diameter is designed, which consists of six crank-slider mechanisms and three fan-shaped wheel blades arranged in a circular array. The contraction and expansion of the wheel are driven by controlling multiple sets of motors. According to the feature recognition of obstacles, the shape and diameter of the wheel are adaptively adjusted to achieve the climbing over step-like obstacles. In this thesis, the modeling and simulation, prototype design, control algorithm, and experimental verification of the stair-climbing robot are studied.

Keywords: Controllable variable-diameter wheel · Stair-climbing robot · Adaptive recognition

1 Introduction

The motion execution device of a mobile robot plays a decisive role in its passability when performing tasks on complex road conditions. Due to the complex terrain environment in which it is used, the mobile robot is required to be light in weight, small in size, and highly adaptable, so as to complete different work tasks. The kinematic research of wheels shows that the larger the diameter, the more conducive it is to improving the passability on the road surface. However, a large wheel diameter will lead to an increase in the overall vehicle volume and weight. In recent years, there have also been mobile robots using non-single-wheel train designs such as legged, tracked, and planetary wheels [1–6]. However, these non-single-wheel train drives also bring problems such as increased volume, complex control, and limited dexterity.

Currently, mobile robots generally use fixed-diameter wheels. On the basis of limited volume, the movement ability of mobile robots using fixed-diameter wheels in obstacle environments with height differences such as stairs is greatly restricted and has no absolute advantage [7–9]. In summary, a method for a controllable deformable wheel

to climb up and down stairs controlled by double pull rods is proposed. It consists of multiple evenly divided wheels and is controlled by pull rods on different rotating shafts, which can achieve opening angles in different directions and changes in wheel diameter to adapt to various step heights and widths for obstacle crossing. Based on theoretical analysis, prototype design, and experimental verification, it brings new ideas for the design of intelligent stair-climbing wheelchairs and obstacle-crossing mobile robots.

2 Structural Design of the Controllable Deformable Wheel

Aiming at the difficulties of conventional mobile robots in passing through step-like obstacles, for the circular driving execution mechanism, the folding and stretching in both the axial and radial directions can be considered. However, due to the requirements of the operation task, the wheel structure should adapt to various environments as much as possible [10, 11]. Therefore, only the radial folding and stretching of the designed wheel are analyzed.

2.1 Principle Analysis of the Controllable Deformable Wheel Mechanism

The radial folding and stretching of the designed deployable wheel can be regarded as its radial telescopic motion. Therefore, the crank-slider mechanism is considered to realize this kind of motion form. Since the single wheel blade of the wheel needs to change its angle, a double-hinged point is used to control its opening and closing angle, so as to realize the transformation of the wheel fan blade (Fig. 1).

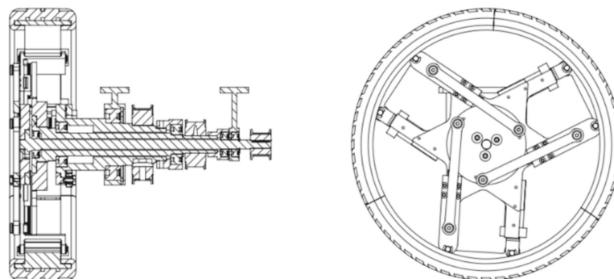


Fig. 1. Three-dimensional model of the controllable deformation wheel

According to the schematic diagram of the mechanism, the structure is decomposed as follows: The rotation of the wheel is jointly driven by Motor 1, Motor 2, and Motor 3. The rotation of Motor 1 and Motor 2 drives the crank to drive the linear slider to adjust the angle of the wheel fan (Fig. 2).

The schematic diagram of the mechanism of a single fan-shaped wheel in the initial state is shown in the figure (Fig. 3)

In the initial state, the center of the single wheel coincides with the center of the drive shaft. The distance between the guide rails is equal to l_2 , the hole distance of the wheel piece. The distance from l_2 to the center of the wheel fan is equal to l_1 . Analyzing

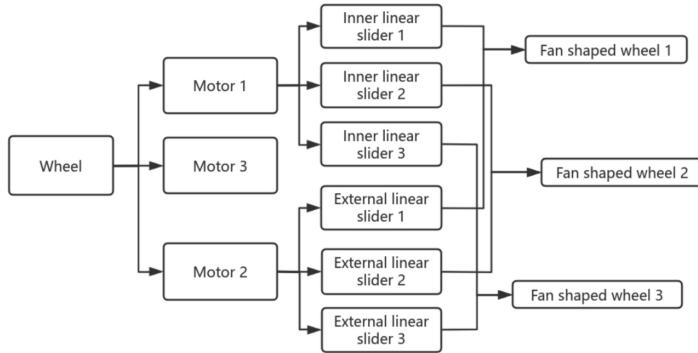


Fig. 2. Adjustment method of the fan-shaped wheel

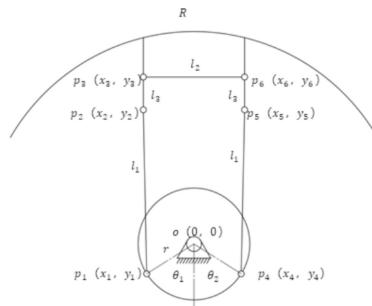


Fig. 3. Schematic diagram of the initial structure of a single fan-shaped wheel

the structure of a single wheel piece, the connecting rod l_4 moves along the linear guide rail. The hinge point of the connecting rods l_1 and l_3 moves along the linear guide rail. Taking the center of the wheel diameter as the origin of the coordinate system, the initial angle of the connecting rod is θ_1 .

$$\theta_1 = \cos^{-1} \left\{ r * \left[l_2 - 2 * \cos \left(\tan^{-1} \frac{2 * (R - l_5 - l_3)}{l_2 - 2 * r} \right) \right] * l_1 \right\} \quad (1)$$

In the initial state, $\theta_1 = \theta_2$.

Schematic diagram after the motor adjusts the movement of the wheel fan (Fig. 4)

Motor 1 rotates clockwise by φ_1 at the initial angle θ_1 , and Motor 3 rotates counter-clockwise by φ_2 at the initial angle. Then the coordinates of point $p_1(x_1, y_1)$, $p_4(x_4, y_4)$ are:

$$p_1(-r * \cos(\theta_1 + \varphi_1), -r * \sin(\theta_1 + \varphi_1)) \quad (2)$$

$$p_4(r * \cos(\theta_1 + \varphi_1), -r * \sin(\theta_1 + \varphi_1)) \quad (3)$$

The connecting rod l_4 moves along the linear guide rail. The hinge point of the connecting rods l_1 and l_3 moves along the linear guide rail. Then the coordinates of points $p_2(x_2,$

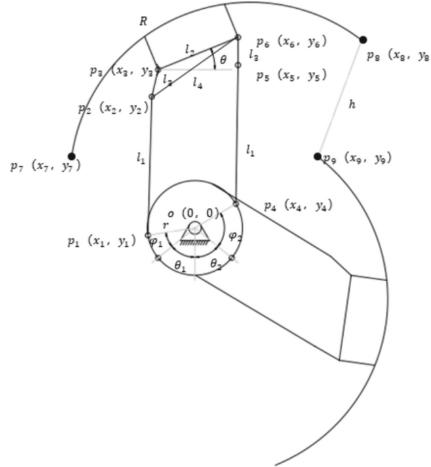


Fig. 4. Schematic diagram after the movement of the fan-shaped wheel piece

p_2) and $p_5(x_5, y_5)$, $p_6(x_6, y_6)$ are:

$$p_5 \left(\frac{l_2}{2}, \sqrt{l_1^2 - \left(\frac{l_2 - 2 * r \cos(\theta_2 + \varphi_2)}{2} \right)^2} - r * \sin(\theta_2 + \varphi_2) \right) \quad (4)$$

$$p_6 \left(\frac{l_2}{2}, \sqrt{l_1^2 - \left(\frac{l_2 - 2 * r \cos(\theta_2 + \varphi_2)}{2} \right)^2} - r * \sin(\theta_2 + \varphi_2) + l_3 \right) \quad (5)$$

$$p_2 \left(-\frac{l_2}{2}, \sqrt{l_1^2 - \left(\frac{l_2 - 2 * r \cos(\theta_1 + \varphi_1)}{2} \right)^2} - r * \sin(\theta_1 + \varphi_1) \right) \quad (6)$$

The distance l_4 from point p_2 to point p_6 is

$$l_4 = \sqrt{(x_6 - x_2)^2 + (y_6 - y_2)^2} \quad (7)$$

The coordinates of point $p_3(x_3, y_3)$ are

$$x_3 = \left(-\frac{l_2}{2} + l_3 * \cos\left(\tan^{-1} \frac{y_6 - y_2}{x_6 - x_2}\right) - \cos^{-1}\left(\frac{l_3^2 + l_4^2 - l_2^2}{2 * l_3 * l_4}\right) \right) \quad (8)$$

$$y_3 = \left(y_2 + l_3 * \sin\left(\tan^{-1} \frac{y_6 - y_2}{x_6 - x_2}\right) - \cos^{-1}\left(\frac{l_3^2 + l_4^2 - l_2^2}{2 * l_3 * l_4}\right) \right) \quad (9)$$

The rotation angle θ of the wheel fan is

$$\theta = \tan^{-1} \frac{y_6 - y_3}{x_6 - x_3} \quad (10)$$

The initial position coordinates of the end point are $p_7\left(-\sqrt{3}R, \frac{R}{2}\right)$, $p_8\left(\sqrt{3}R, \frac{R}{2}\right)$. After the transformation of the wheel fan, it can be regarded as formed by translation from the initial state and then rotation by θ . Then the coordinates of p_7 after the transformation are

$$x_7 = \left(-\sqrt{3}R - x_6\right) * \cos \theta - \left(\frac{R}{2} - y_6\right) * \sin \theta + x_6 \quad (11)$$

$$y = \left(-\sqrt{3}R - x_6\right) * \sin \theta - \left(\frac{R}{2} - y_6\right) * \cos \theta + y_6 \quad (12)$$

The coordinates of p_8 are

$$x_8 = \left(\sqrt{3}R - x_6\right) * \cos \theta - \left(\frac{R}{2} - y_6\right) * \sin \theta + x_6 \quad (13)$$

$$y_8 = \left(\sqrt{3}R - x_6\right) * \sin \theta + \left(\frac{R}{2} - y_6\right) * \cos \theta + y_6 \quad (14)$$

The wheel is composed of three fan-shaped wheel segments, that is, a single wheel fan is arrayed around the rotating shaft. The coordinates of the end points of adjacent wheel fans

$$x_9 = x_8 * \cos \frac{2\pi}{3} - y_8 * \sin \frac{2\pi}{3} \quad (15)$$

$$y_9 = x_8 * \sin \frac{2\pi}{3} - y_8 * \cos \frac{2\pi}{3} \quad (16)$$

Openings of different wheel fans

$$h = \sqrt{(x_7 - x_9)^2 + (y_7 - y_9)^2} \quad (17)$$

Through the calculation of the motion principle of the controllable deformation wheel mechanism, the rotation angle θ of the fan-shaped wheel segment and the opening h of adjacent segments can be calculated from the rotation angle of the motor, providing a theoretical basis for wheel control.

2.2 Simulation Analysis of Wheel Obstacle Negotiation

When walking on a flat ground, the wheel takes the shape of an initial complete circle, which can maintain high mobility. When the mobile trolley crosses an obstacle, an obstacle - crossing resistance will be generated at the interface where the wheel contacts the obstacle. Assuming that the obstacle does not move on the ground during obstacle negotiation, the wheel cannot climb over it due to the limitation of its diameter. The obstacle parameters are uploaded to the controller through the sensor to determine whether it can successfully climb over. During climbing, the shape of the fan - wheel is adjusted according to the obstacle parameters and the traveling direction. The following figure shows the simulation process taking stair - climbing as an example (Fig. 5).

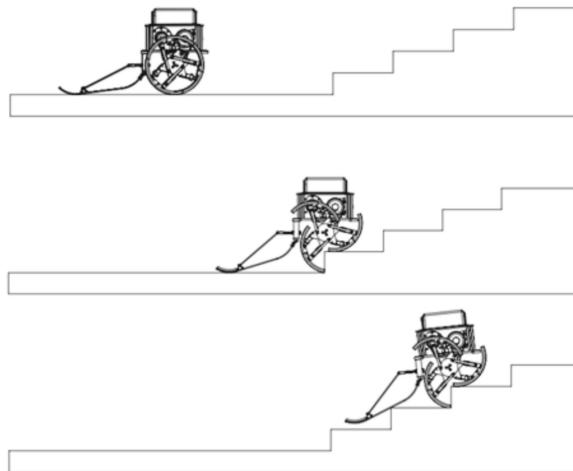


Fig. 5. Process of the mobile robot climbing stair - like obstacles

3 System Control and Test Scheme

3.1 System Hardware Construction

The stair - climbing robot adjusts the rotation of the motor based on the depth vision recognition of the camera and the real - time angle change data collected by the IMU. Taking the left wheel as an example, the diameter change of the left wheel is achieved by the simultaneous rotation of diameter - changing motor 1 and diameter - changing motor 2. The forward and backward movement of the left wheel is achieved by the simultaneous rotation of diameter - changing motor 1, diameter - changing motor 2 and drive motor 1. The hardware structure diagram of the whole machine is shown in Fig. 6:

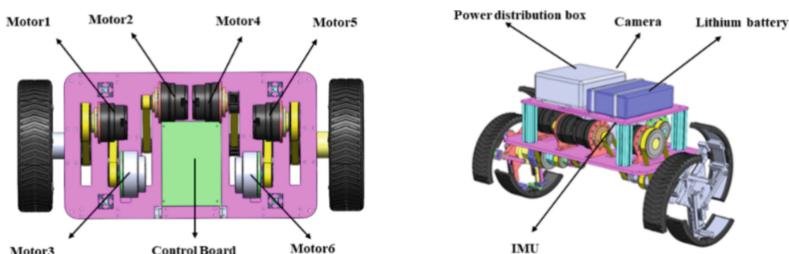


Fig. 6. Hardware structure diagram of the robot

The camera visually collects the edges of the stairs and obstacles [12]. The main control board will control the rotation of the diameter - changing motor and the drive motor according to the collected terrain conditions. The IUM always collects the angle between the robot platform and the horizontal ground and feeds it back to the main

control board for real - time correction. Design of the control algorithm for going up and down stairs

During the process of climbing stairs, the robot needs to use the sensor detection system to output the distance S between the robot and the stairs, the depth D of the stairs, and the width W of the stairs. Based on this, it determines whether to climb the stairs and the unfolding posture of the deformable wheels, and adjusts the direction of its tail drag to ensure that the robot can climb the stairs smoothly. The specific implementation process is shown in Fig. 7 below

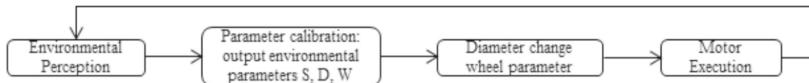


Fig. 7. Flowchart of the overall process

After obtaining the three - dimensional point cloud data through the depth camera, the width and depth of the stairs are extracted through plane fitting and edge detection. Since there will be errors and fluctuations in the point cloud data obtained based on the point cloud sensor, there will be a large deviation in the obtained stair width and depth, which makes it difficult for the stair - climbing robot to climb the stairs smoothly. This system corrects the errors by compensating the robot's motion distortion through the IMU, assists in positioning, and improves the measurement accuracy. First, construct the IMU pose matrix. Integrate the angular velocity $\omega(t)$ and acceleration $a(t)$ measured by the IMU to obtain the pose matrix (18):

$$T(t) = \begin{bmatrix} R(t) & P(t) \\ 0 & 1 \end{bmatrix} \quad (18)$$

Angle integration of the rotation part $R(t)$ (19):

$$R(t) = R(t_0) \cdot \text{Exp}\left(\int_{t_0}^t \omega(t) dt\right) \quad (19)$$

$\text{Exp}(\cdot)$ is the Lie algebra exponential mapping (20):

$$\text{Exp}(\omega) = I + \frac{\sin \|\omega\|}{\|\omega\|} [\omega] + \frac{1 - \cos \|\omega\|}{\|\omega\|^2} [\omega]^2 \quad (20)$$

Translation part (21):

$$P(t) = \int_0^t \left(\int_0^\tau a(s) ds \right) d\tau + v_0 t + P_0 \quad (21)$$

Where v_0 and P_0 are the initial velocity and position.

For point cloud distortion correction, for each original point $P_{\text{raw}} = (x, y, z, t)$, according to its acquisition time t , apply the corresponding pose matrix $T(t)$ for correction (22):

$$P_{\text{corrected}} = T(t)^{-1} \cdot P_{\text{raw}} \quad (22)$$

Map all point cloud data to a unified reference coordinate system (the camera coordinate system at the initial moment) to eliminate the geometric distortion caused by motion. The following calculation is the corrected coordinates. Then, fuse the multi - sensor data through filtering technology, and output the distance S from the robot to the stairs and the stair dimensions D, W.

For the stair width W and depth D, extract the point cloud data plane after obtaining the point cloud data set; obtain the data of multiple stair facades, and calculate the predicted stair width and depth based on the stair facade data. Obtain the stair facade set based on the point cloud data. The stair facade set includes the data of multiple stair facades; perform plane correction on the stair facades; obtain the average value of the distances; obtain the predicted stair width and depth according to all the average distance values. Adjust the normal vector of the plane equation corresponding to the stair facade to the horizontal to correct the stair facade to a vertical stair surface. The plane equation corresponding to the plane - corrected stair facade is shown in Eq. (23)

$$a_i x + b_i y + c_i z + d_i = 0 \quad (23)$$

Among them, $n(a_i, b_i, c_i)$ represents the normal vector of the plane equation corresponding to the i -th stepped facade, $p(x, y, z)$ represents the centroid coordinates, and d_i represents the distance from the i -th stepped facade to the origin;

The plane equation corresponding to the stepped facade after plane correction is shown in the following Eq. (24):

$$a_i x + b_i y + \cos \alpha_i \times d_i = 0 \quad (24)$$

Among them, α represents the included angle between the i -th stepped facade and the Z axis.

The vertical facade of the adjacent step with the shortest distance from the stair-climbing robot is defined as the first plane, and the vertical facade of the adjacent step with a larger distance from the stair-climbing robot is defined as the second plane. Obtain the average distance, calculate the first distance and the second distance. The first distance is the distance from the centroid of the point cloud corresponding to the second plane to the first plane, and the second distance is the distance from the centroid of the point cloud corresponding to the first plane to the second plane; then obtain the average distance.

Then the calculation formulas for the average distance are shown in Eqs. (25) and (26):

$$W = \frac{1}{n} \sum_{i=1}^n \left(\frac{\|n_2 \cdot p_1 + d_2\|}{\|n_2\|} + \frac{\|n_1 \cdot p_2 + d_1\|}{\|n_1\|} \right) \quad (25)$$

$$D = \frac{1}{n} \sum_{i=1}^n d_i \quad (26)$$

Among them, $n(a, b, c)$ is based on the normal vector of the plane equation and the centroid coordinates $p(x, y, z)$. Predict the stair width based on all the average distances.

In the initial state, the robot first needs to decide whether to go upstairs, downstairs or detour and make a motion plan. Upstairs mode: Place the tail drag at the rear to reduce the load on the front wheels and improve the climbing stability. Downstairs mode: Switch the tail drag to the front to avoid the risk of overturning caused by the backward tilting of the center of gravity. If it is determined that it cannot pass, then detour (Fig. 8).

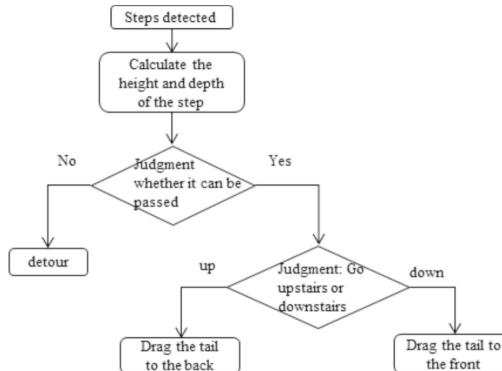


Fig. 8. Robot tail drag management

Wheel deformation control is the key to enabling the robot to climb stairs. In this paper, through inverse kinematics calculation, according to the stair size information (step depth D and step height W), and at the same time, design the forward calculation formula according to the wheel configuration to solve the required radius r and the rotation angle θ of the wheel rim of the deformable wheel. The specific calculation formulas are: as follows (27) and (28):

$$\theta = \arctan\left(\frac{W}{2r_0}\right) \quad (27)$$

$$r = r_0 + \frac{D}{2 \tan \theta} \quad (28)$$

Among them: r_0 is the initial radius of the variable-diameter wheel.

Based on θ and r, the operating angles (slider distances) of motors No. 1 and No. 2 for controlling the left wheel can be calculated. Similarly, the operating angles (slider distances) of motors No. 4 and No. 5 for the right wheel can also be calculated.

At the same time, obtain the distance S of the robot relative to the stairs and transmit it to the trajectory planning algorithm. When the robot is not in contact with the stairs, motors No. 1, 2, and 3 run at the same speed, and motors No. 4, 5, and 6 run at the same speed to achieve forward, backward, and differential operations.

According to the inverse kinematics model: Calculate the target radius r of the wheel and the rim rotation angle θ based on the step parameters (D, W). Here, r_0 is the initial radius of the wheel. Map (r, θ) to the slider strokes of the left and right wheel drive motors (for example, the slider stroke of motor No. 1 is 45 mm, 2 and that of motor No. 21 mm), and achieve dynamic adjustment of the wheel diameter through closed-loop control (Fig. 9).

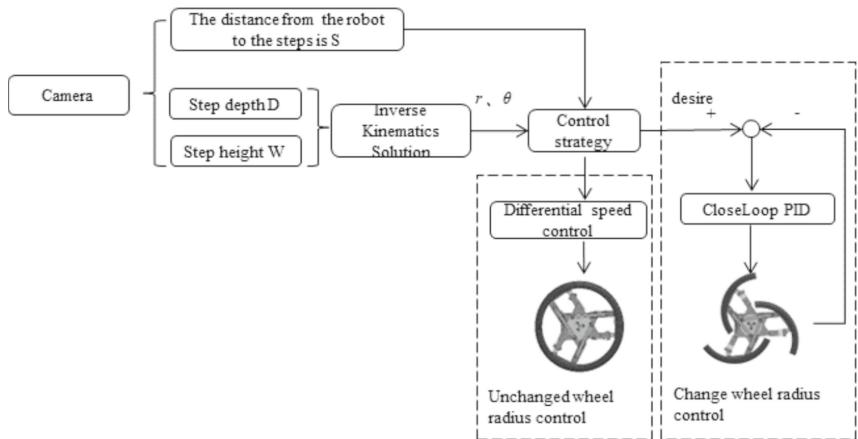


Fig. 9. Schematic diagram of stair-climbing control

4 Experimental Research

Keep consistent with the previous simulation environment. Realize the scenario of flat ground - going upstairs - flat ground - going downstairs - flat ground. Configure a physical diagram with the width and height dimensions of the steps. The height of a single stair step is 10 cm, the depth of a single step is 30 cm, and the width is 100 cm . There are a total of 4 steps. Then elaborate on the specific experimental steps (Fig. 10).

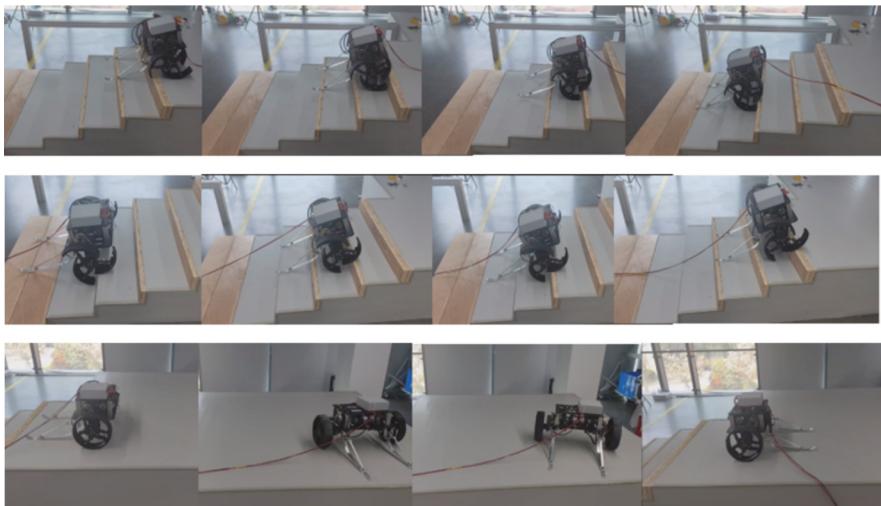


Fig. 10. Movement of the controllable deformable wheeled robot in different scenarios

Through experiments, aiming at the difficulties of conventional mobile robots in passing through step-like obstacles, the robot is modeled and simulated, the prototype is

designed, and the control algorithm method is verified. On a flat road surface, the robot moves with circular wheels to ensure its high mobility in motion. After encountering step-like obstacles, it identifies their characteristics and adaptively adjusts the wheel shape and diameter to climb over step-like obstacles.

5 Conclusions

The controllable deformable wheeled structure proposed in this paper has a compact structure and a small turning radius. In the initial state, it can maintain the high mobility inherent in wheeled structures. When encountering obstacles, it can adaptively adjust the wheel shape and diameter according to their characteristics to climb over step-like obstacles. The design and control methods proposed in this paper have been verified by experiments to be able to adaptively climb over step-like obstacles, bringing new ideas to the design of stair-climbing wheelchair robots, obstacle-crossing load-carrying robots, and special reconnaissance mobile robots.

References

1. Qin, H., He, L.P., Xing, L.Q., et al.: Kinematics analysis and simulation verification for the chassis of mobile robots based on an obstacle-crossing index[C]. In: Global Reliability and Prognostics and Health Management, pp. 1–6. Piscataway, IEEE (2020)
2. Chen, W.H., Lin, H.S., Lin, Y.M., et al.: TurboQuad: a novel leg wheel transformable robot with smooth and fast behavioral transitions[J]. *IEEE Trans. Robot.* **33**(5), 1025–1040 (2017)
3. Guo, J., Shi, J., Zhu, W., et al.: Approach to autonomous stair climbing for tracked robot[C]. In: 2017 IEEE International Conference on Unmanned Systems(ICUS), pp. 182–186. IEEE (2017)
4. Yuan, J., Paisley, R., Song, Y., et al.: Virtual realization of automatic stair-climbing motion by leg-wheeled hybrid mobile robot[C]. In: IEEE, International Conference on Robotics and Biomimetics, pp. 1352–1357 (2010)
5. Zhang, N., Shen, L.B., Li, Y.Z., et al.: Adaptive deformable wheeled mobile mechanism [J]. *Mech. Sci. Technol.* **1**(1), 1–9 (2020)
6. Liu, Y., Liu, G.: Track-stair interaction analysis and online tipover prediction for a self-reconfigurable tracked mobile robot climbing stairs[J]. *IEEE Trans. Mechatron.* **14**(5), 528–538 (2009)
7. Quaglia, G., Franco, W., Oderio, R.: A mechanical concept for a stair climbing wheelchair[C]. In: IEEE International Conference on Robotics and Biomimetics, pp. 800–805 (2009)
8. Quaglia, G., Franco, W., Nisi, M.: Kinematic analysis of an electric stair-climbing wheelchair[J]. *Ingenieria Y Universidad.* **21**(1), 27–48 (2017)
9. Chen, S.R., Xu, M., Wang, Y.P., et al.: Control and simulation of wheeled robot climbing steps [J]. *Design.* **24**(1) (2007)
10. Pan, L.H., Kuo, C.N., Huang, C.Y., et al.: The claw-wheel transformable hybrid robot with reliable stair climbing and high maneuverability[C]. In: IEEE International Conference on Automation Science and Engineering, pp. 233–238. Piscataway, IEEE (2016)
11. Kim, Y., Kim, J., Kim, H.S., et al.: Curved-spoke tri-wheel mechanism for fast stair-climbing[J]. *IEEE Access.* **7**, 173766–173773 (2019)
12. Jin, T.M.: Research on Motion Target Depth Information Extraction Method Based on Binocular Vision [D]. University of Electronic Science and Technology China (2017)



Intelligent Stable Walking Based on Nonlinear Disturbance Observer

Helin Wang^(✉)

Department of Electrical and Electronic Engineering, Shanghai Institute of Technology,
Shanghai, China
Whl7228@126.com

Abstract. Flexible and robust walking is the most basic guarantee for humanoid robots to complete various tasks, which requires complex mechanical control systems, environmental perception, and action planning capabilities. However, robots are inevitably disturbed in actual walking environments, which can affect the stability of walking and even disrupt the robot's periodic motion, causing it to fall. This paper is concerned with bipedal walking based on nonlinear disturbance observer. By constructing dynamic equations and utilizing Lyapunov stability theory to design observers, real-time estimation and compensation of composite disturbances, such as sudden changes in ground reaction forces, joint friction, can be achieved. We adjust the gain rate by designing the gain matrix of the observer. Finally, the effectiveness of the mentioned method is illustrated by simulations.

Keywords: Biped Walking · Nonlinear Disturbance Observer · Globally Stable

1 Introduction

Humanoid robot is a complex multi rigid body system, described as multiple sets of complex high-order nonlinear equations, with strong coupling between multiple degrees of freedom. This not only places high demands on the performance of gait planning algorithms for achieving stable walking, but also on the performance of their balance motion control. The algorithms usually consist of a set of computational tools to solve nothing more than a specific optimization problem. The humanoid robot research could provide a better test-bed in the study of artificial intelligence. For instance, we can study how the current state-of-the-art AI paradigms should be integrated to achieve simple tasks. We can also try to develop novel framework of artificial intelligence to overcome the challenge posed by the humanoid robots research.

Despite the enormous potential of humanoid robots, achieving complex tasks and dealing with highly complex robot dynamics remains challenging. These tasks require robots to safely collaborate with humans in dynamic environments and operate in unstructured environments. In order to quickly acquire motor and cognitive skills, humanoid robots can use human data for learning. With the help of human knowledge, humanoid robots can achieve rapid embodied intelligence, thereby accelerating skill acquisition. Perception algorithms can detect, classify, and segment various objects

in real-time. The model-based approach achieves agile and reliable motion and operation through prediction and reactive control. Deep learning strategies have demonstrated good control effects on robot hardware through exploration and imitation.

The environmental adaptability of humanoid robots is the core advantage compared to wheeled and tracked robots, and their design requires a comprehensive combination of motion control, perception ability, and energy efficiency technologies. In terms of terrain adaptability, robots can walk stably in unstructured terrains such as forests, slopes, and stairs through dynamic gait adjustment and whole machine parameter identification technology. Foot contact force control can adapt to changes in ground softness and hardness. Passive dynamic design, such as elastic joints, enhances the buffering ability against irregular ground and reduces the risk of instability caused by impact forces. In terms of multi-modal motion switching, some humanoid robots integrate a wheel foot hybrid motion mode to switch wheel movements on flat ground to improve efficiency, and restore foot movements when encountering obstacles.

In the perception system, the observer serves as a key state estimation tool, mainly used for real-time perception and compensation of system dynamic characteristics. Through high-precision state estimation and disturbance compensation, it significantly improves the stability and dynamic response capability in complex terrains, providing a reliable technical foundation for scenarios such as home services and industrial inspections. When designing a state feedback control law based on a linearized model, it is necessary to estimate the system state (such as centroid position, velocity, etc.) in real time. For example, humanoid robots use unscented kalman filter (UKF) to fuse multi-source data such as IMU and force sensors, thereby achieving accurate estimation of robot posture and motion state. In reinforcement learning frameworks, recurrent neural networks (RNN) can act as disturbance observers, compensating for model uncertainty by remembering historical state information and improving the robustness of strategies in partially observable environments. The observer needs to handle heterogeneous data from multiple sensors, such as IMU, force sensors, and cameras. For example, humanoid robots generate continuous and consistent state estimation outputs by integrating inertial measurement and joint encoder information through UKF. For body posture estimation, humanoid robots construct a discrete stochastic system model using six dimensional force sensors and accelerometers, and combine Kalman filtering algorithm to achieve real-time updates of pitch/roll angles.

The design of a humanoid robot observer mainly requires a systematic combination of model derivation, sensor fusion, and real-time optimization. Firstly, it is necessary to determine the estimated state variables, such as centroid position, velocity, joint angle, and external disturbances. Then selecting available sensors (IMU, encoder, force sensor, etc.) is based on the data frequency and noise characteristics. To realize model simplification and linearization, the impact range of model errors, such as friction parameter deviations, and external disturbances, such as sudden changes in ground reaction forces, will be qualified. There are currently several types of linear observers. The Kalman filter is suitable for linear Gaussian noise systems and requires the design of state equations and observation equations. The Luenberger observer estimates convergence speed through pole placement optimization. The extended Kalman filter processes weakly non-linear systems through local linearization. Unscented kalman filter is utilized with Sigma

point propagation to solve strongly nonlinear problems such as bipedal pose estimation. Sliding mode observer (SMO) suppresses model uncertainty through discontinuous control laws and is suitable for disturbance estimation. Disturbance Observer (DOB) aims at modeling disturbances as equivalent inputs and estimating external disturbances in real-time.

The rest of the paper is organized as follows. Section 2 tells how to construct a nominal model based on system dynamics equations as the benchmark for observer design, which injects external and internal disturbances into the simulation, and models them uniformly through equivalent disturbance expressions. Section 3 presents the method of designing a nonlinear observer structure based on the nominal model and analyzing the stability of the observer through Lyapunov analysis. Section 4 presents the proposed observer by testing the ability to maintain control performance under parameter perturbations in the testing model. Section 4 describes simulations of the observers with quantifying the suppression effect of external disturbances after sudden load. Finally, Sect. 5 provides conclusions with future work.

2 Modeling

At present, the dynamic model of humanoid robots mainly refers to the mechanical arm model, which is regarded as a multi-link mechanical arm model with no fixed base but in contact with the ground for free fall. Constructing robot dynamics equations is based on Lagrange equation or Newton Euler method during the single support phase and the double support phase, respectively, including joint drive, center of mass motion, and external disturbance terms. The dynamic model established based on Lagrange equation during the single-support phase is,

$$D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) = \Gamma + d \quad (1)$$

where $D(q)$ is the inertia matrix of robot, $C(q, \dot{q})$ is the centrifugal force and Coriolis force matrix, $G(q)$ is the gravity matrix, Γ is the joint driving force torque, and d is the external disturbances. The expression for d can be calculated from (1),

$$d = D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) - \Gamma \quad (2)$$

The basic idea of designing an observer is to estimate the difference between the output and the actual output and correct the estimated value. Therefore, the observer is designed as,

$$\begin{aligned} \hat{d} &= \ell(q, \dot{q})(d - \hat{d}) \\ &= \ell(q, \dot{q})\left(D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) - \Gamma\right) - \ell(q, \dot{q})\hat{d} \end{aligned} \quad (3)$$

Assuming that the change in dynamic characteristic interference relative to the observer is slow. It can be expressed mathematically as, $\dot{d} = 0$. Let the observation error be,

$$\xi = d - \hat{d} \quad (4)$$

Derive the observation error,

$$\dot{\xi} = -\hat{d} = -\ell(q, \dot{q})(d - \hat{d}) = -\ell(q, \dot{q})\xi \quad (5)$$

List the constraint equations that observation errors satisfy as follows,

$$\dot{\xi} + \ell(q, \dot{q})\xi = 0 \quad (6)$$

3 Controller and Observer Design

The objective in this section is to find

The interference observer provides a model and data fusion based interference suppression scheme, which is combined with PD control and adaptive control to form a composite control strategy. Online estimation of external disturbances, such as load changes and environmental disturbances and internal disturbances, such as model errors and parameter drift, acting on the robot through dynamic information input and actual output of the system. Based on the nominal model of the robotic system, the deviation between the actual output and the model output is equivalent to a “virtual interference”, and the estimated value of the interference is extracted through an observer algorithm.

Define auxiliary parameter variable,

$$\varpi = \hat{d} - \mathfrak{N}(q, \dot{q}) \quad (7)$$

where $\mathfrak{N}(q, \dot{q})$ is the pending design function.

Let $\ell(q, \dot{q})D(q)\ddot{q}$ be given by the nonlinear equation below,

$$\ell(q, \dot{q})D(q)\ddot{q} = \frac{d\mathfrak{N}(q, \dot{q})}{dt} = \left[\begin{array}{cc} \frac{\partial \mathfrak{N}(q, \dot{q})}{\partial q} & \frac{\partial \mathfrak{N}(q, \dot{q})}{\partial \dot{q}} \end{array} \right] \begin{bmatrix} \dot{q} \\ \ddot{q} \end{bmatrix} \quad (8)$$

From (7) and (8), it can be concluded that,

$$\begin{aligned} \dot{\varpi} &= \dot{\hat{d}} - \frac{d\mathfrak{N}(q, \dot{q})}{dt} \\ &= \dot{\hat{d}} - \ell(q, \dot{q})D(q)\ddot{q} \end{aligned} \quad (9)$$

Substitute (3) into (9),

$$\begin{aligned} \dot{\varpi} &= \dot{\hat{d}} - \ell(q, \dot{q})D(q)\ddot{q} \\ &= \ell(q, \dot{q}) \left(D(q)\ddot{q} + C(q, \dot{q})\dot{q} + G(q) - \Gamma \right) - \ell(q, \dot{q})\hat{d} - \ell(q, \dot{q})D(q)\ddot{q} \\ &= \ell(q, \dot{q}) (C(q, \dot{q})\dot{q} + G(q) - \Gamma) - \ell(q, \dot{q})\hat{d} \end{aligned} \quad (10)$$

Then substitute (7) into (10),

$$\begin{aligned} \dot{\varpi} &= \ell(q, \dot{q}) (C(q, \dot{q})\dot{q} + G(q) - \Gamma) - \ell(q, \dot{q})\hat{d} \\ &= \ell(q, \dot{q}) (C(q, \dot{q})\dot{q} + G(q) - \Gamma) - \ell(q, \dot{q}) (\varpi + \mathfrak{N}(q, \dot{q})) \\ &= \ell(q, \dot{q}) (C(q, \dot{q})\dot{q} + G(q) - \Gamma - \mathfrak{N}(q, \dot{q})) - \ell(q, \dot{q})\varpi \end{aligned} \quad (11)$$

Therefore, the nonlinear observer is designed as,

$$\begin{cases} \dot{\varpi} = \ell(q, \dot{q})(C(q, \dot{q})\dot{q} + G(q) - \Gamma - \mathfrak{R}(q, \dot{q})) - \ell(q, \dot{q})\varpi \\ \hat{d} = \varpi + \mathfrak{R}(q, \dot{q}) \end{cases} \quad (12)$$

where $\ell(q, \dot{q})$ can be calculated from (8), hence, we can obtain,

$$\dot{\xi} = -\hat{d} = -\dot{\varpi} - \mathfrak{R}(q, \dot{q}) \quad (13)$$

Substitute (8) and (12) into (13),

$$\begin{aligned} \dot{\xi} &= -\dot{\varpi} - \mathfrak{R}(q, \dot{q}) \\ &= \ell(q, \dot{q})\varpi - \ell(q, \dot{q})(C(q, \dot{q})\dot{q} + G(q) - \Gamma - \mathfrak{R}(q, \dot{q})) - \mathfrak{R}(q, \dot{q}) \\ &= \ell(q, \dot{q})(\varpi + \mathfrak{R}(q, \dot{q})) - \ell(q, \dot{q})(D(q, \dot{q})\dot{q} + C(q, \dot{q})\dot{q} + G(q) - \Gamma) \end{aligned} \quad (14)$$

In order to further calculate its derivative, we substitute (2) and (12) into (14),

$$\dot{\xi} = \ell(q, \dot{q})\hat{d} - \ell(q, \dot{q})d = -\ell(q, \dot{q})\xi$$

Therefore, the observation error equation is obtained,

$$\dot{\xi} + \ell(q, \dot{q})\xi = 0 \quad (15)$$

4 Simulation Results

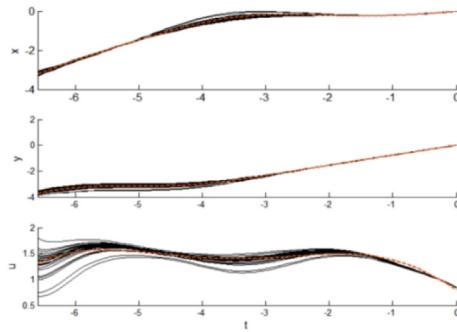


Fig. 1. Compared simulations against reference trajectory.

The simulation of interference observer needs to be based on model accuracy and interference equivalence, and ensure its effectiveness in actual systems through parameter optimization and closed-loop verification. Verify its robustness and practicality in complex environments through simulation. Figure 1 shows the compared simulations against reference trajectory. Figure 2 shows the polynomial approximation of trajectory, which describes the degree of fitting between the robot joint motion trajectory and the reference trajectory. Figure 3 shows polynomial approximation of lqr gains. And convex hull approximation of the reachable set is described in Fig. 4.

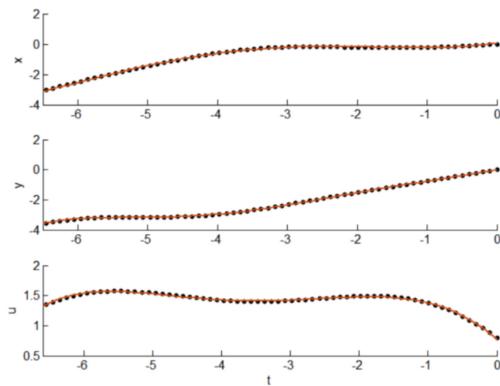


Fig. 2. Polynomial approximation of trajectory

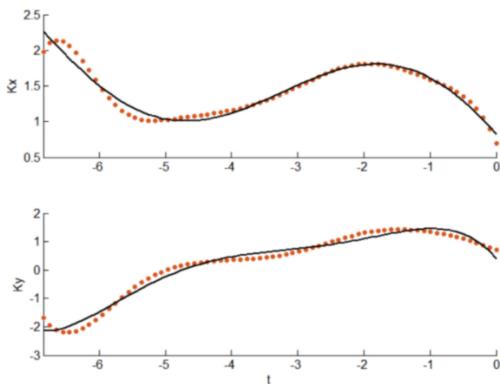


Fig. 3. Polynomial approximation of lqr gains

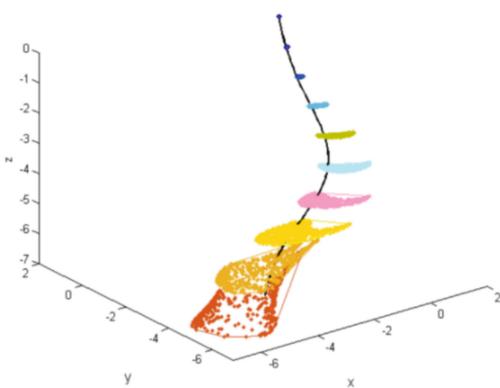


Fig. 4. Convex Hull approximation of the reachable set

5 Conclusions

In response to environmental adaptation issues, future development of bipedal robots will mainly target complex environmental conditions, such as outer space environments, complex ground or terrain after natural disasters, such as earthquakes, volcanic eruption and so on. This will lead to targeted and stable structural design for bipedal robots. Regarding the issue of autonomous control, when robots perform tasks in complex environmental conditions, they should have the ability to autonomously avoid obstacles and successfully complete tasks without human interference based on the job tasks and environmental perception results. The organic combination of robotics technology and biomimetic technology to design efficient autonomous control and decision-making control strategies is a top priority for future research. The ultimate goal of human research on bipedal robots is to enable robots to better serve humans and achieve human-machine integration. By combining robotics with biological science, intelligent controllers are designed using neural signals, providing a new research approach for robot intelligence.

Funding. This work is supported by National Natural Science Foundation (NNSF) of China under Grant (No. 61573260; No. 62073245), Artificial Intelligence Promotes Research Paradigm Reform and Empowers Discipline Leap Plan Project-AIZX-12(1021ZK250021012-A06).

References

1. Radosavovic, I., Zhang, B., Shi, B., Rajasegaran, J., Kamat, S., Sreenath, K.: Humanoid Locomotion as Next Token Prediction. arXiv preprint arXiv:2402.19469 (2024)
2. Radosavovic, I., Xiao, T., Zhang, B., Darrell, T., Malik, J.: Real-world humanoid locomotion with reinforcement Learning. *Sci. Robot.* **9**(89) (2024)
3. Ji, Y., et al.: Hierarchical reinforcement learning for precise soccer shooting skills using a quadrupedal robot. In: 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1479–1486 (2022)
4. Heydarnia, O., Dadashzadeh, B., et al.: Discrete sliding mode control to stabilize running of a biped robot with compliant kneeled legs. *Autom. Control. Comput. Sci.* **51**(5), 347–356 (2017)
5. Caron, S., Escande, A., Lanari, L.: Capturability-based pattern generation for walking with variable height. *IEEE Trans. Robot.* **36**(2), 517–536 (2019)
6. Li, Z., et al.: Reinforcement learning for robust parameterized locomotion control of bipedal robots. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 2811–2817 (2021)
7. Savin, S.: ZMP-Based Trajectory Generation for Bipedal Robots Using Quadratic Programming (2020)
8. Duan, H., Pandit, B., Gadde, M.S., van Marum, B.J.. Learning Vision-Based Bipedal Locomotion for Challenging Terrain. arXiv preprint arXiv: 2309.14594, (2023)
9. Kirchner, E.A., Bütefür, J.: Towards bidirectional and coadaptive robotic exoskeletons for neuromotor rehabilitation and assisted daily living: a review. *Curr. Robot. Rep.* **3**(2), 21–32 (2022)
10. Park, S., Oh, J.: Real-time continuous ZMP pattern generation of a humanoid robot using an analytic method based on capture point. *Adv. Robot.* **33**(1), 33–48 (2019)
11. Tran, P., Jeong, S., Herrin, K.R., et al.: A review: hand exoskeleton systems, clinical rehabilitation practices, and future prospects. *IEEE Trans. Med. Robot. Bionics.* **2021** (2021). <https://doi.org/10.1109/TMRB>



VLM-Based Dense Reward for Zero-Shot Robot Manipulation

Kehao Shi¹, Zhenyi Xu^{1,2,3}, Yang Cao^{1,3(✉)}, Jun Huang², and Yu Kang^{1,3(✉)}

¹ Department of Automation, Anhui Province Key Laboratory of Intelligent Low-Carbon Information Technology and Equipment, University of Science and Technology of China, Hefei 230026, China

I.Ekeland@princeton.edu

² Anhui Province Key Laboratory of Digital Twin Technology in Metallurgical Industry, Anhui University of Technology, Ma'anshan, Anhui 243032, China

³ Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230088, China

Abstract. Intelligent robots have been widely used in many fields, which can help humans greatly improve work efficiency. However, mainstream methods tend to use imitation learning or reinforcement learning, assuming there are expert demonstrations or priors to train a generalist, which is costly and time-consuming. Benefiting from large models, we propose DenseRL for zero-shot robot manipulation. Through the combination of pretrained vision-language models, DenseRL can provide multivariate zero-shot visual rewards. More importantly, DenseRL only uses frozen pretrained models without any finetuning. Experiments on Meta-World show that DenseRL can achieve the SOTA performance among zero-shot baselines, in which the average success rate is 16.7% higher than the best baseline.

Keywords: reward shaping · vision-language model · robot manipulation

1 Introduction

With the development of artificial intelligence and automation, the demand for robots that can work in different fields is increasing, which could help humans improve efficiency. Mainstream methods usually focus on collecting high-quality expert demonstrations to assist agents in learning professional skills. However, with the maturity of computer technology, generalists that execute under zero-shot or few-shot are becoming a new trend.

There are currently two mainstream methods realizing robot manipulation. First, imitation learning (IL) helps the agent obtain supervision signals from demonstrations and convert robot action prediction into supervised training

Y. Cao and Y. Kang.

[1,2], or learn rewards from demonstrations to match the action distribution between the agent and the expert [3,4], which reduces the requirements for experts' domain knowledge. However, it needs humans to remotely collect enough high-quality observation-action pairs in advance, which is very time-consuming and some even takes months to complete data preparation, such as RT-1 [5], Palm-e [6] and Octo [7]. In addition, for some complex tasks, the quantity and quality of demonstrations vary greatly, which would seriously affect the quality of skill learning. Therefore, an alternative approach is to learn the robot's skills through experience accumulation and interaction with the environment, i.e., reinforcement learning (RL) [8]. However, the formulation of rewards is a classic challenge for RL. Traditional methods use the reward elaborately designed by experts with domain knowledge to optimize the policy network [9,10], which requires a lot of time and effort. Moreover, hyperparameters in handcrafted rewards also need many adjustments under priori knowledge about the environment and task, which requires experts to own sufficient domain knowledge.

To address the above issues, this paper aims to utilize open-source pretrained vision-language models (VLMs) to achieve zero-shot robot manipulation. The key insight of this paper is that VLMs that have been pretrained on Internet-scale datasets have the rich background knowledge to generalize to unseen domains, which is useful for zero-shot skill learning. Therefore, this paper introduces pretrained VLMs to design rewards, which are used to alleviate the cost of designing manual rewards for different tasks and reduce the reliance on priors from the task and environment.

Specifically, we propose DenseRL, which combines VLMs pretrained on Internet-scale datasets. Through the complementarity of background knowledge among different VLMs, it could transform general knowledge into domain knowledge of robot manipulation without finetuning. Specifically, DenseRL would utilize VLMs to provide multivariate implicit rewards to the policy network based on the trajectory of the agent and the description of the task, without any robot-related priori knowledge. The contributions of this paper are as follows:

1. We propose DenseRL, a zero-shot dense reward for language-conditioned RL without collecting task-specific demonstrations or finetuning pretrained models.
2. DenseRL leverages the combination of different VLMs to achieve cooperation across background knowledge and realize zero-shot robot manipulation.
3. Experiments on MetaWorld show that DenseRL can achieve the SOTA performance among zero-shot baselines.

2 Related Work

Shaping effective rewards for the agent is a classic challenge of RL. Traditional works use rule-based rewards defined by experts while it's time-consuming. Benefiting from large models pretrained on Internet-scale datasets, some works choose

to use large language models (LLMs) to replace experts for generating rewards for agents. For example, Xie et al. [11] use GPT-4 to code dense reward functions for different tasks. Chu et al. [12] use LLMs to provide implicit reward functions to determine whether the agent trajectory is good or not. Kwon et al. [15] use LLMs to transform observations of the agent into binary rewards. Bai et al. [16] collect enough data and then finetune LLMs to provide shaped rewards. Although they can provide well-shaped rewards for agents, it requires task-specific priors or finetuning LLMs. Another popular paradigm is to use CLIP-based models for providing similarity rewards. Sontakke et al. [13] propose RoboCLIP to provide trajectory-level sparse rewards for agents. In addition, RoboCLIP focuses more on proving that IL can be achieved by providing one demonstration video for different tasks. Fu et al. [17] finetune the reward model through target states and task-specific data to provide well-shaped rewards. Unlike them, DenseRL does not specifically collect demonstrations to finetune VLMs or use task-specific priors. Instead, it directly provides dense rewards through different open-source pretrained VLMs to replace handcrafted ground-truth rewards.

3 Methodology

3.1 Preliminaries

We define the problem as a partially observable Markov decision process (POMDP) consisting of a 7-tuple $(O, S, A, R, \tau, \gamma, L)$, where O represents the observation space and S represents the state space. A represents the action space and R represents the reward function. $\gamma \in [0, 1]$ represents the discount factor and L represents the text description of the task, which remains unchanged throughout the training. $\tau(s_{t+1}|s_t, a_t)$ represents the probability of the environment state transition after taking the action $a_t \in A$. After transmitting to the next state, DenseRL calculates the dense reward $r^t = R(o_t)$ at time step t based on the observation $o_t \in O$. The goal is to learn a policy π that maximizes the expected cumulative reward, which can be written as, $E[\sum_{t=0}^H \gamma^t r^t]$.

3.2 Overview

DenseRL could provide zero-shot reward for the training of the agent based on different VLMs. To be specific, as can be seen from Fig. 1, at each timestep, DenseRL could measure the task progress based on current observations and task goals as well as storing the current observation into the buffer. Thus, at the end of each episode, DenseRL could provide trajectory-level similarity to control the movement of the agent. The above similarity scores would be used as the dense reward for the training of the agent.

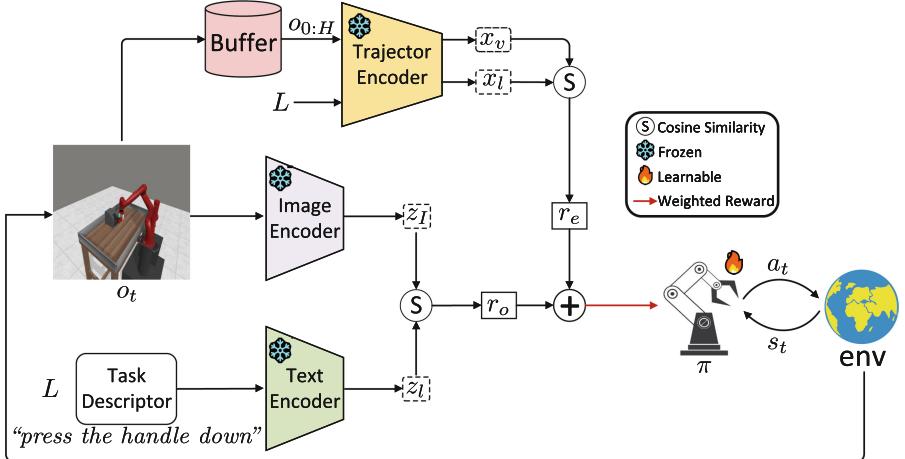


Fig. 1. DenseRL pipeline. It takes the pixel observation o_t and the task instruction L , and computes cosine similarity as the dense reward r_o at the timestep t . At the end of each episode, DenseRL provides trajectory-level sparse reward r_e in terms of robot trajectories recorded in the buffer. The weighted sum of r_o and r_e would replace the ground-truth reward to reduce the cost of designing handcrafted rewards.

3.3 Reward Generation

Unlike traditional methods that offer priori knowledge for IL or RL, DenseRL could provide zero-shot dense rewards based on pixel observations during training the policy π , which are used to replace the ground-truth rewards of environmental feedback and would reduce the cost of designing manual rewards. Specifically, at each time step, DenseRL calculates cosine similarity between the pixel observation and task instruction as dense rewards for the agent. At the same time, the pixel observation of each time step is recorded in the buffer. At the end of each episode, DenseRL calculates trajectory-level reward based on the historical observations recorded in the buffer. The final reward is a combination of the two rewards mentioned above.

Formally, given the task instruction L , at each time step t , the agent would receive the observation o_t , and the image encoder is used to encode o_t into the embedding space, which can be expressed as $z_I = \phi_i(o_t)$. Similarly, the text description L is encoded into the same embedding space through the text encoder, i.e., $z_L = \phi_T(L)$. In this paper, ϕ_i and ϕ_T are instantiated by LIV [18]. Thus, the dense reward at time step t can be calculated as:

$$r_o = \frac{z_I \cdot z_L}{\|z_I\|_2 \times \|z_L\|_2}. \quad (1)$$

Moreover, after each episode, the buffer maintained in the system will store the historical observation sequence $o_{0:H}$. Then, DenseRL encodes $o_{0:H}$ and the instruction L through trajectory-level encoders θ_i and θ_T . Next, we compute the

trajectory-level features x_v and x_l , which will be used to compute the trajectory-level sparse reward r_e . Note that r_e can only be obtained after each interaction, and equals 0 in other cases. The above process can be expressed as:

$$x_v = \theta_i(o_{0:H}), \quad (2)$$

$$x_l = \theta_T(L), \quad (3)$$

$$r_e = \begin{cases} \frac{x_v \cdot x_l}{\|x_v\|_2 \times \|x_l\|_2} & \text{at the end of the episode,} \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where θ_i and θ_T are instantiated by S3D [19]. Finally, we train the policy network in terms of the above rewards, which can be written as, $r^t = \lambda_1 * r_o + \lambda_2 * r_e$, where λ_1 and λ_2 are weights of r_o and r_e . In this paper, we set $\lambda_1 = 2$ and $\lambda_2 = 1$.

3.4 Action Execution

Once the reward r^t is computed, the agent would act as a_t based on s_t , i.e., $a_t \sim \pi(\cdot | s_t)$. In this paper, we train the policy π with PPO [21] and use zero-shot rewards to replace the ground-truth reward.

4 Experiment

4.1 Implement Details

DenseRL is implemented under PyTorch and Python 3.8. The CPU is a Intel(R) Core(TM) i7-10700K CPU @ 3.80GHz and the GPU is a NVIDIA GeForce RTX 3090 with 24GB memory. We choose 6 different tasks from MetaWorld [20] to clarify effectiveness of DenseRL and the learning rate is set to 0.0001. We run each task across 2 seeds and report the average success rates.

4.2 Comparison

We first report the success rates of different tasks on MetaWorld to verify the effectiveness of DenseRL, and related results are summarized in Table 1. We choose different zero-shot reward models as baselines. Notably, for the fifth line in the table, we use a binary reward for PPO, in which 0 is given when the task is not successful while 1 is given when the task is successful. It can be seen that in the absence of expert demonstrations, DenseRL can perform better than the baseline methods on most tasks and achieve the SOTA performance.

In addition, it can be seen that baselines perform poorly on most tasks because only one VLM is difficult to effectively adapt to new domains and visual rewards provided may have certain ambiguity, which is the reason previous works choose to collect task-specific demonstrations to finetune pretrained

Table 1. Comparison on MetaWorld. Average success rates are reported.

Task	VLM-RMs[14]	LIV[18]	R3M[22]	PPO[21]	DenseRL(Ours)
Coffee-button	0	0	0	0	50
Drawer-close	100	100	100	100	100
Drawer-open	0	0	0	0	0
Hammer	0	0	0	0	10
Handle-press	100	100	0	0	100
Window-open	0	0	0	20	40
Average	33.3	33.3	16.7	20	50

VLMs. Moreover, it's found that all methods do not work in `drawer-open`. It's believed that compared with coarse-grained tasks such as `drawer-close` and `handle-press`, opening a drawer requires more fine-grained operations, which means the robotic arm must accurately grasp the handle of the drawer and then pull it. However, it is difficult for existing pretrained models to provide such fine-grained and well-shaped rewards, which makes it difficult to align the rewards and leads to failures. It's expected that larger VLMs in this field can be raised for reward shaping.

4.3 Ablation Study

We then carry out ablation study on the dense reward r_o and trajectory-level reward r_e , which are summarized in Table 2. It can be seen that when the dense reward r_o generated by ϕ_i and ϕ_T is removed, the model performs worse than original DenseRL, and most tasks cannot be completed. In addition, when the trajectory-level reward r_e is removed, the task success rate also decreases. However, it can be seen that the dense reward r_o has a greater impact on the success rates than that of r_e , which illustrates that providing dense rewards is more important to the agent during skill learning.

Table 2. Ablation study on different rewards. We report the average success rates and “w/o” means without.

Task	w/o r_o	w/o r_e	DenseRL(Ours)
Coffee-button	0	20	50
Drawer-close	100	100	100
Drawer-open	0	0	0
Hammer	0	0	10
Handle-press	50	100	100
Window-open	0	0	40

4.4 Visualization

In Fig. 2, we visualize trajectories of the agent under different tasks. As can be seen from the figure, DenseRL can successfully achieve zero-shot robotic manipulation under Fig. 2(a) and (b) when only the text description of the task is given. Some failures are also shown in Fig. 2(c–f).

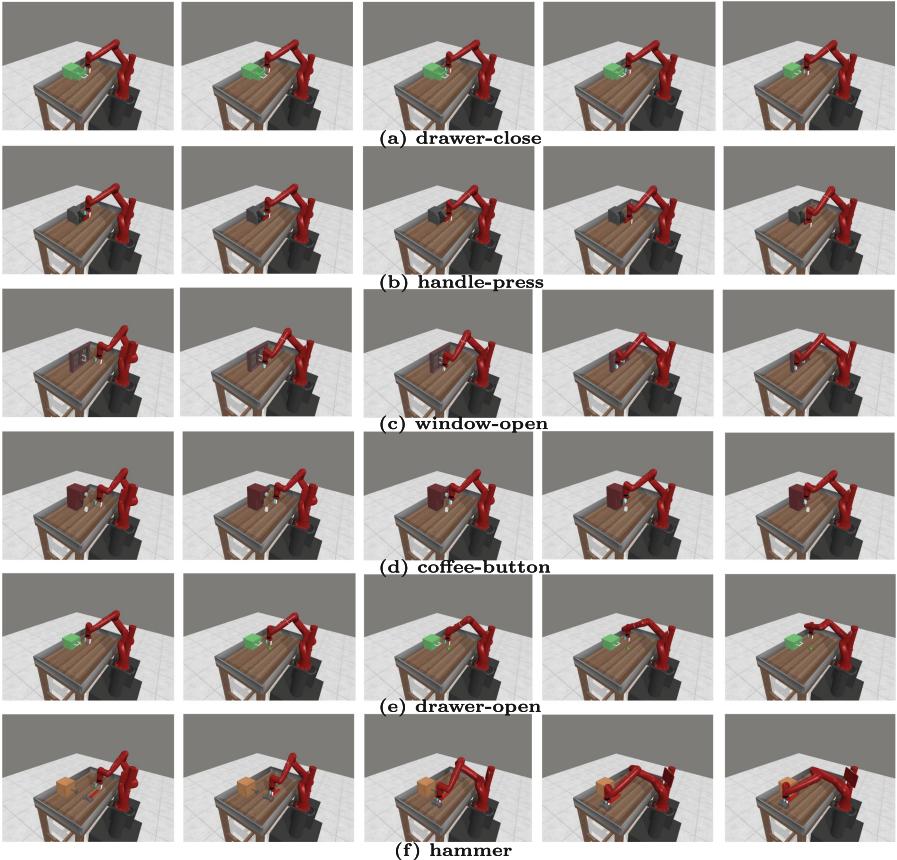


Fig. 2. Visualization. The first column is the start frame and each row changes over time from left to right. Moreover, (a) and (b) are successful while the rest are not successful.

5 Conclusion

In this paper, we propose DenseRL for zero-shot robot manipulation, in which the key insight is that pretrained VLMs can provide effective rewards for robots

without task-specific priors. Also, it's proved that dense rewards are important during skill learning. Experiments on different tasks show that DenseRL outperforms the zero-shot baselines, which could contribute to zero-shot robot manipulation and a stronger baseline.

Limitations and future work: Although DenseRL could realize zero-shot robot manipulation, all experiments are carried out in simulated environments. Therefore, for sim-to-real, we still need to consider more factors, such as the influence of the friction coefficient in the real environment and the appearance difference of objects in simulated and real environments.

References

1. Torabi, F., Warnell, G., Stone, P.: Behavioral Cloning From Observation (2018). arXiv preprint [arXiv:1805.01954](https://arxiv.org/abs/1805.01954)
2. Sasaki, F., Yamashina, R.: Behavioral cloning from noisy demonstrations. In: International Conference on Learning Representations (2020)
3. Chang, W.D., Fujimoto, S., Meger, D., et al.: Imitation Learning from Observation Through Optimal Transport (2023). arXiv preprint [arXiv:2310.01632](https://arxiv.org/abs/2310.01632)
4. Haldar S, Mathur V, Yarats D, et al. Watch and match: supercharging imitation with regularized optimal transport. In: Conference on Robot Learning, pp. 32–43. PMLR (2023)
5. Brohan, A., Brown, N., Carbajal, J., et al.: RT-1: Robotics Transformer For Real-World Control at Scale (2022). arXiv preprint [arXiv:2212.06817](https://arxiv.org/abs/2212.06817)
6. Driess, D., Xia, F., Sajjadi, M.S.M., et al.: Palm-e: An Embodied Multimodal Language Model (2023)
7. Team, O.M., Ghosh, D., Walke, H., et al.: Octo: An Open-Source Generalist Robot Policy (2024). arXiv preprint [arXiv:2405.12213](https://arxiv.org/abs/2405.12213)
8. Kalashnikov, D., Irpan, A., Pastor, P., et al.: Scalable deep reinforcement learning for vision-based robotic manipulation. In: Conference on Robot Learning , pp. 651–673. PMLR (2018)
9. Amodei, D., Olah, C., Steinhardt, J., et al.: Concrete Problems in AI Safety (2016). arXiv preprint [arXiv:1606.06565](https://arxiv.org/abs/1606.06565)
10. Shou, Z., Di, X.: Reward design for driver repositioning using multi-agent reinforcement learning. Transp Res Part C: Emerg Technol **119**, 102738 (2020)
11. Xie, T., Zhao, S., Wu, C.H., et al.: Text2reward: automated dense reward function generation for reinforcement learning. In: International Conference on Learning Representations (ICLR) (07/05/2024-11/05/2024, Vienna, Austria) (2024)
12. Chu, K., Zhao, X., Weber, C., et al.: Accelerating Reinforcement Learning of Robotic Manipulations via Feedback From Large Language Models (2023). arXiv preprint [arXiv:2311.02379](https://arxiv.org/abs/2311.02379)
13. Sontakke, S., Zhang, J., Arnold, S., et al.: Roboclip: one demonstration is enough to learn robot policies. Adv. Neural. Inf. Process. Syst. **36**, 55681–55693 (2023)
14. Rocamonde, J., Montesinos, V., Nava, E., et al.: Vision-Language Models are Zero-Shot Reward Models for Reinforcement Learning (2023). arXiv preprint [arXiv:2310.12921](https://arxiv.org/abs/2310.12921)
15. Kwon, M., Xie, S.M., Bullard, K., et al.: Reward Design with Language Models (2023). arXiv preprint [arXiv:2303.00001](https://arxiv.org/abs/2303.00001)

16. Bai, Y., Jones, A., Ndousse, K., et al.: Training a Helpful and Harmless Assistant with Reinforcement Learning From Human Feedback (2022). arXiv preprint [arXiv:2204.05862](https://arxiv.org/abs/2204.05862)
17. Fu Y, Zhang H, Wu D, et al. Furl: Visual-Language Models as Fuzzy Rewards for Reinforcement Learning (2024). arXiv preprint [arXiv:2406.00645](https://arxiv.org/abs/2406.00645)
18. Ma, Y.J., Kumar, V., Zhang, A., et al.: LIV: Language-image representations and rewards for robotic control. In: International Conference on Machine Learning, pp. 23301–23320. PMLR (2023)
19. Xie, S., Sun, C., Huang, J., et al.: Rethinking spatiotemporal feature learning: speed-accuracy trade-offs in video classification. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 305–321 (2018)
20. Yu, T., Quillen, D., He, Z., et al.: Meta-world: a benchmark and evaluation for multi-task and meta reinforcement learning. In: Conference on Robot Learning, pp. 1094–1100. PMLR (2020)
21. Schulman, J., Wolski, F., Dhariwal, P., et al.: Proximal Policy Optimization Algorithms (2017). arXiv preprint [arXiv:1707.06347](https://arxiv.org/abs/1707.06347)
22. Nair, S., Rajeswaran, A., Kumar, V., et al.: R3M: A Universal Visual Representation for Robot Manipulation (2022). arXiv preprint [arXiv:2203.12601](https://arxiv.org/abs/2203.12601)



Reinforcement Learning-Based Autonomous Control Strategy for Snake Robots in Confined Terrains

Cheng Wang^{1(✉)}, Yunhu Zhou^{1,2}, Yuhong Na¹, Ku Du¹, and Darong Huang¹

¹ School of Artificial Intelligence, Anhui University, Hefei 230000, China
wangcheng2018666@outlook.com, yunhuzhou@ahu.edu.cn

² R&D Center, Efort Intelligent Equipment Co., Ltd, Wuhu, China

Abstract. Snake robots exhibit significant potential in complex terrain exploration due to their exceptional flexibility and environmental adaptability. However, conventional vision-based navigation approaches demonstrate substantial limitations in low-light or completely dark environments (e.g., underground pipelines, disaster ruins), requiring reliance on high-power illumination systems or costly infrared/radar sensors, which lead to substantial increases in system costs and energy consumption. Furthermore, multi-sensor integration in confined spaces introduces structural complexity and reliability challenges. This paper proposes a vision-independent autonomous control method for serpentine robots based on reinforcement learning. The approach establishes a simulation environment incorporating walls, obstacles, and low-clearance passages, while implementing multi-faceted reward functions and optimized training strategies to achieve vision-free navigation in complex terrains. Leveraging the Proximal Policy Optimization (PPO) algorithm with joint state and positional information, the proposed method enables visual-input-free autonomous navigation through low-clearance scenarios. Experimental results demonstrate enhanced locomotion stability and improved navigational efficacy in confined spaces within simulated environments. This work provides a novel solution for vision-independent robotic control systems, particularly addressing challenges in unstructured subterranean environments.

Keywords: Snake robot · Reinforcement Learning · Vision-free Control

1 Introduction

Snake robots have garnered significant attention for their exceptional flexibility and environmental adaptability, demonstrating vast potential in complex terrain navigation, disaster rescue operations, and minimally invasive surgical procedures [1, 2]. However, conventional control methodologies predominantly rely on predefined motion patterns and computationally intensive path planning algorithms, which often exhibit limited adaptability and operational inefficiency in dynamically changing or unknown environments.

In recent years, reinforcement learning (RL) has emerged as a promising machine learning paradigm for robotic control, enabling autonomous policy optimization through

environmental interactions [3, 4]. Nevertheless, the effective application of RL to vision-independent autonomous control of serpentine robots — particularly in achieving reliable navigation through confined spaces with low vertical clearance — continues to face critical challenges. These include inefficient algorithmic convergence, prohibitive training durations, and insufficient generalization capabilities across diverse environmental configurations.

2 Related Work

The field of serpentine robot control has long grappled with the trade-off between environmental adaptability and computational efficiency. Current approaches fall into three categories:

2.1 CPGs Simulates the Biological Rhythm Movement of the Snake

Biomechanics-inspired conventional control: Early studies employed Central Pattern Generators (CPGs) to mimic biological rhythmic locomotion in snakes. For instance, [5–7] generated serpentine gaits via coupled oscillators, yet required manual presetting of frequency and phase parameters, causing failure modes when encountering unexpected obstacles. While introduced an impedance control-based joint coordination strategy to enhance terrain adaptability, its rigid motion patterns frequently induced joint lockup in narrow passages [8].

2.2 Vision-Dependent Deep Reinforcement Learning

With advancements in deep reinforcement learning (DRL), [9–11] pioneered the application of Deep Q-Networks (DQN) to serpentine robot navigation, achieving obstacle avoidance through RGB image inputs, [12]. However, this method exhibited critical limitations, including visual data processing latency reaching 200 ms and policy failures under illumination variations [13]. Further integrated Long Short-Term Memory (LSTM) networks with the Advantage Actor-Critic (A3C) algorithm to predict motion trajectories from sequential video frames [14–16]. While demonstrating improved temporal coherence, the model’s 1.2 M parameter count hindered real-time deployment on resource-constrained embedded platforms [17].

2.3 Non-visual Sensor Fusion Methods

Recent efforts have focused on minimizing visual dependencies, combined inertial measurement unit (IMU) data with joint encoder feedback, implementing a Twin Delayed DDPG (TD3) algorithm for indoor obstacle avoidance [18]. Despite eliminating vision sensors, this approach relied on a simplistic 9-dimensional joint state space that omitted obstacle position modeling, resulting in a passage success rate below 60% in cluttered environments, proposed a tactile feedback-driven DDPG framework utilizing pressure sensor arrays for collision detection [19, 20]. While innovative, the method suffered from

frequent policy oscillations due to unmodeled tactile signal noise and actuator response delays.

In response to the aforementioned challenges, this paper proposes a reinforcement learning-based control method for snake-like robots. The vision-independent approach enables efficient and stable motion control in complex environments through autonomous learning, while enhancing traversal capability in low and narrow spaces by utilizing joint state and positional sensor information. The main contributions include:

Constructing a simulation scene featuring walls, cubic obstacles, and low channels to model the snake-like robot's structure, joint configuration, and dynamic parameters, thereby establishing a realistic robotic simulation model.

Defining the state space and action space of the snake-like robot. A diversified reward function is designed by comprehensively considering multiple factors: movement speed of the snake head toward the target direction, alignment between motion direction and target orientation, distance variation to the target point, absolute distance to target, collision detection, and joint action amplitude with smoothness constraints.

Implementing a multilayer perceptron (MLP) architecture to construct policy and value networks. Reinforcement learning algorithms are applied for multi-round simulation training, enabling the robot to autonomously acquire optimized parameter sets for decision-making in various scenarios. This achieves optimal motion control for target-reaching tasks.

Organizing experimental data and validation results to demonstrate the method's effectiveness in complex environment navigation and confined space traversal.

3 Method

3.1 Simulation Environment and Robotic Modeling

A simulation environment containing walls, cubic obstacles, and low-clearance channels was constructed on the MuJoCo platform. Key features include: The low-clearance channel height was set and limited to 1.2 times the diameter of the robot to mimic the confined space of the real world. Cubic obstacles are randomly distributed on both sides of the channel, ensuring different obstacle layouts during training. A 1.0 m high side wall was set as a boundary constraint to prevent the robot from deviating from the mission area. Three directional light sources were configured to enhance the visualization of motion trajectory and collision detection. A snake-like robot is modeled, which consists of 12 joints in series, each controlled by an actuator.

State Space and Action Space

The state space (40-dimensional) encompasses joint angles (12 rad), angular velocities (12 rad/s), snake head coordinates (x, y), target distance (m), target direction (unit vector), positions of the four nearest obstacles (8 dimensions), and low-clearance passage boundaries (4 dimensions). The action space comprises 12 continuous torque commands normalized to $[-1, 1]$, linearly scaled to an actual torque range of $[-7, 7]$ (N/m).

Multiple Reward Function Design

The reward function integrates motion efficiency, obstacle avoidance and energy consumption control, and the specific formula is as follows:

$$R_t = R_{\text{forward}} + R_{\text{angle}} + R_{\text{distance}} + R_{\text{target}} + R_{\text{collision}} + R_{\text{cost}} \quad (1)$$

The forward reward R_{forward} was given based on the speed of the snake's head moving toward the target direction. Angle reward R_{angle} gives rewards based on the alignment between the motion direction of the reward snake head and the target direction. Distance change reward R_{distance} gives reward or punishment based on the change of distance between the snake head and the target point. Target distance penalty R_{target} punishes the snake based on the distance between the snake head and the target point. The goal achievement reward $R_{\text{collision}}$ punishes the snake head by judging that it has successfully reached the goal point. Control cost penalty R_{cost} gives a penalty based on the amplitude and variation of joint actions.

The above formula of R_{forward} is as follows:

$$R_{\text{forward}} = w_{\text{forward}} \cdot v_{\text{goal}} \quad (2)$$

where $v_{\text{goal}}, w_{\text{forward}}$ is the component and weight coefficient of snakehead velocity in the target direction;

The above formula of R_{angle} is as follows:

$$R_{\text{angle}} = \cos(\theta) \quad (3)$$

where θ is the Angle between the motion direction of the snake head and the target direction.

The above formula of R_{distance} is as follows:

$$R_{\text{distance}} = \begin{cases} 1.0 & \Delta d > 0 \\ -0.5 & \Delta d < 0 \\ 0 & \Delta d = 0 \end{cases} \quad (4)$$

where Δd represents the change of the distance from the snake to the target, so as to motivate the snake to keep moving and prevent it from getting stuck.

The above formula of R_{target} is as follows:

$$R_{\text{target}}(d) = \begin{cases} 0 & \text{if } d \leq d_{\text{threshold}} \\ -\min\left(\frac{d-d_{\text{threshold}}}{d_{\text{threshold}}}, 1\right) \cdot w & \text{if } d > d_{\text{threshold}} \end{cases} \quad (5)$$

where d represents the distance between the current snake head and the target point, and $d_{\text{threshold}}$ is the distance threshold.

The above formula of $R_{\text{collision}}$ is as follows:

$$R_{\text{collision}} = \begin{cases} -C_{\text{collision}} & \text{If there is a collision} \\ 0 & \text{Otherwise} \end{cases} \quad (6)$$

$C_{\text{collision}}$ is a constant.

R_{cost} is specifically the sum of squares of all joint actions multiplied by the weight coefficients.

3.2 Network Structure and Training Strategy

The PPO algorithm is employed, with a two-layer MLP policy network (256 neurons per layer and activation) and a value network. The objective function incorporates policy clipping and entropy regularization to balance exploration and exploitation.

The input layer has a 40-dimensional state vector. The hidden layer contains 2 fully connected layers (256 neurons) and activation. The output layer is the mean (μ) and variance (σ) of the 12-dimensional Gaussian distribution, and the action range is constrained by the tanh function.

The snake-like robot interacts with the environment, collects state s_t , action a_t , and reward R_t data, and stores them in a replay buffer. The reinforcement learning algorithm utilizes this data to compute the clipped objective function and value loss function. The policy network parameters θ are updated by minimizing the clipped objective function $L^{CLIP}(\theta)$, enabling the robot to select better actions for maximizing cumulative rewards. Simultaneously, the value network parameters ϕ are updated by minimizing the value loss function $L^{VF}(\phi)$, ensuring accurate state-value estimation and providing effective guidance for the policy network.

The clipped objective function is defined as follows:

$$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)} \quad (7)$$

$$L^{CLIP}(\theta) = \mathbb{E}_t [\min(r_t(\theta)\hat{A}_t, clip(r_t(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}_t)] \quad (8)$$

The value loss function is as follows:

$$L^{VF}(\phi) = \mathbb{E}_t \left[\left(V_\phi(s_t) - V_t^{target} \right)^2 \right] \quad (9)$$

The total loss function is as follows:

$$L(\theta, \phi) = L^{CLIP}(\theta) - c_1 L^{VF}(\phi) + c_2 S[\pi_\theta](s_t) \quad (10)$$

E_t denotes the expectation over all sampled timesteps t , averaging across the batch data to ensure stable and consistent policy updates. E_t denotes the expectation over all sampled timesteps t , averaging across the batch data to ensure stable and consistent policy updates. \hat{A}_t is the advantage function computed via Generalized Advantage Estimation (GAE), quantifying the relative benefit of action a_t in state s_t compared to the average behavior. Clip restricts the policy ratio $r_t(\theta)$ to the range $[1 - \epsilon, 1 + \epsilon]$, guiding policy updates toward favorable directions while leveraging \hat{A}_t for optimization. ϵ is a small constant controlling the clipping range. $\pi_\theta(a_t | s_t)$ is the probability density function of the current policy network selecting action a_t in state s_t . $\pi_{\theta_{old}}(a_t | s_t)$ is the probability density function of the old policy selecting action a_t in state s_t . $V_\phi(s_t)$ is the estimated value function for state s_t . V_t^{target} denotes the target value at timestep t , used for training the value network. $S[\pi_\theta](s_t)$ represents the entropy of policy π_θ in state s_t , measuring the policy's stochasticity. $r_t(\theta)$ is the policy ratio during the t training iteration. c_1, c_2 are hyperparameters.

The snake-like robot requires smooth motion in complex environments to avoid abrupt action changes caused by excessive policy updates. The clipped objective function ensures stable and continuous motion by constraining the magnitude of policy updates.

We introduce the following improvements over the standard PPO. The initial value of the dynamic clipping coefficient ϵ is 0.2, and it linearly decays to 0.1 with the training progress, balancing exploration and utilization. The entropy regularization term strategy entropy $S[\pi\theta]$ weighting coefficient $c_2 = 0.01$ to prevent premature convergence of the strategy, generalized Dominance Estimation (GAE) discount factor $\gamma = 0.99$, smoothing parameter $\lambda = 0.95$. In each round, 2048 steps of interaction data were collected(s_t, a_t, r_t, s_{t+1}). Through Critic network estimate \hat{A}_t . The trimmed objective function is maximized through policy update. The value update minimizes the mean square error.

4 Implementation

The experiment is based on the simulation environment built by MuJoCo physics engine, including cube obstacles and low channels. The snake-like robot consists of 12 torque-controlled joints, and the initial joint angles and angular velocities are normalized to $[-1, 1]$. PPO algorithm was used for training, and both the policy network and the value network were two-layer MLP. All experiments were performed on a workstation equipped with NVIDIA RTX 4060 GPU with 3 million training cycles.

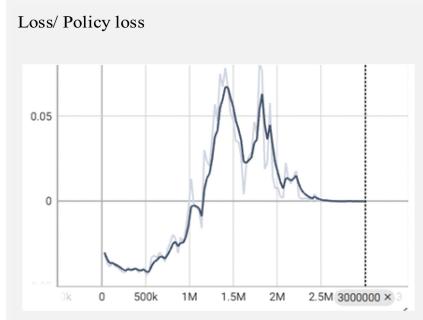


Fig. 1. Policy loss as a function of the number of training steps.

As shown in Fig. 1, the policy loss exhibits a fluctuating decline during the early training phase (0-500 k steps), indicating rapid learning of basic locomotion patterns through exploration. Between 1 M and 1.5 M steps, the loss shows significant oscillations (peaking at 0.07) due to increased environmental complexity (e.g., low-clearance passages), corresponding to the critical stage where the robot learns obstacle avoidance and posture adjustment. Additionally, the robot must learn to lower its head to traverse low-clearance passages while balancing forward speed with head pitch angle. When training exceeds 2 M steps, the loss stabilizes below 0.02, demonstrating the effectiveness of adaptive policy clipping and entropy regularization in enhancing training stability.

As shown in Fig. 2, in the initial exploration phase (0-500 k steps), the average reward gradually rises from an initial value of -0.2 to -0.1 , reflecting the agent's rapid

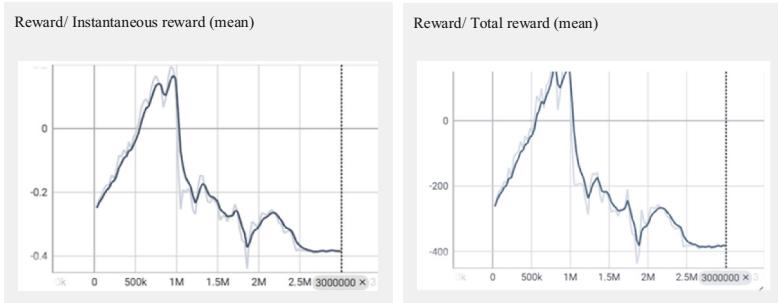


Fig. 2. Average Immediate Reward During Training.

learning of basic obstacle avoidance behaviors under a random policy. From 200 k to 1 M steps, the average reward climbs sharply from -0.1 to a peak of 0.1 (at 1 M steps), corresponding to the robot learning an efficient forward locomotion strategy and directional alignment. Between 1 M and 3 M steps, the average reward drops to -0.4 after the peak due to the introduction of low-clearance passage scenarios, requiring the robot to adjust its posture (lowering its head) to navigate through narrow spaces, leading to increased collision penalties and energy consumption. Ultimately, the reward stabilizes around -0.3 , indicating the policy achieves dynamic equilibrium under multi-objective constraints.

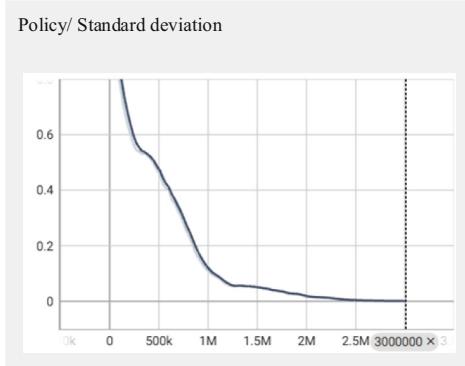


Fig. 3. Trend of output standard deviation of policy network.

As shown in Fig. 3, the policy network output changes with the number of training steps. The standard deviation is highly explored in the early stage, with an initial value as high as 0.8, indicating that the policy network actively performs high-variance actions, including random joint torques, to extensively explore the action space in the early stage of training. The initial fluctuation of the corresponding reward curve is between -400 and 0 , which verifies the high collision risk and low movement efficiency caused by high exploration, and the convergence is accelerated at 500 k-1.5 M steps. The standard deviation decreases rapidly to 0.4, which reflects the gradual convergence of the policy

network to the efficient action mode, indicating stable progress. It gradually stabilizes in the later stage of the 1.5 M–3 M step, and the standard deviation eventually tends to 0, which indicates that the policy network is almost deterministic.

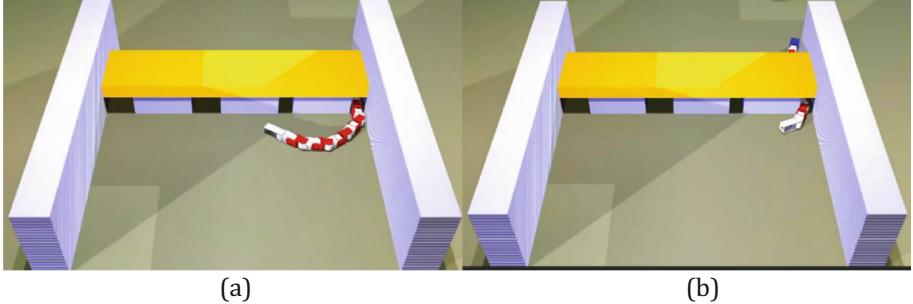


Fig. 4. Action breakdown diagram for the snake-like robot’s low channel traversal task.

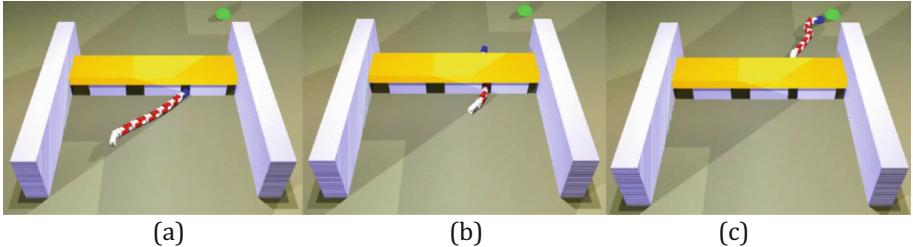


Fig. 5. The snake-like robot passes low through the mission and reaches the goal.

In Fig. 4, the yellow platform corresponds to the height limit of the low channel in the simulation scene, and the robot needs to adjust its attitude to pass through. The white vertical structures represent fixed obstacles or channel boundaries in the simulation environment. In Fig. 5, the green points represent the target points in the navigation task, which is consistent with the orientation alignment behavior driven by the reward function in the paper. In picture a, the robot depressions its body through joint Angle changes, triggering the direction alignment term in the reward function to ensure that the motion direction is aligned with the target. In picture b, the collision penalty term forces the policy network to optimize the joint torque allocation to reduce the collision risk when low channels are crossed. In picture c, the distance reward term guides the robot to reach the goal with the minimum action amplitude.

In this work, we compared this method with the previous CPG methods based on reinforcement learning, and proved its effectiveness in the navigation task in low and narrow passages. Table 1 shows the comparison results of the success rates of reinforcement learning and CPG in passing through low and narrow passages and reaching the target point. It can be seen that, compared with the ordinary CPG method, reinforcement

learning, after continuously pre-learning and obtaining experience, has a higher success rate in passing through low and narrow passages.

Table 1. A comparative experiment between CPG and reinforcement learning.

Method	Success Rate (%)	
	Low-clearance passages	Without passages
CPG	45 ± 3.2	3.984
RL(Ours)	72 ± 4.1	3.979

As shown in Table 2, the RL policy demonstrates superior adaptability by autonomously adjusting joint torques to lower the head posture and navigate narrow spaces, whereas the CPG controller frequently encounters collisions due to its rigid gait patterns. In addition, RL combines the reward function to balance the movement efficiency and shorten the path length.

Table 2. A comparison of performance Between CPG and RL in low-clearance passages.

Method	Path length	Collisions/Episode
CPG	3.8 ± 0.4	2.1 ± 0.5
RL(Ours)	2.1 ± 0.3	0.8 ± 0.2

5 Conclusions

Through the vision-free reinforcement learning framework, this method provides an efficient and low-cost solution for autonomous navigation of snake-like robots in low light or narrow Spaces (such as underground pipelines, disaster ruins), which significantly reduces the hardware complexity and energy consumption caused by traditional vision sensors or radar systems, and has direct application potential in industrial inspection, emergency rescue and other fields. Future research can further combine Domain Randomization technology to solve the problem of dynamic differences in the transfer from simulation to real world, explore the generalized control strategy of dynamic obstacles and variable terrain height, and introduce a multi-agent cooperation framework to improve the cooperation efficiency of multi-robots in complex tasks. It promotes the practical deployment of this method in real complex scenarios.

References

1. Seetohul, J., Shafiee, M.: Snake robots for surgical applications: a review. *Robotics*. **11**(3), 57 (2022)

2. Qu, J.Z., Qu, W.Z., Li, L., Jia, Y.: Reinforcement learning based multi-layer Bayesian control for snake robots in cluttered scenes. In: 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4702–4708. IEEE (2023)
3. Qiu, K., Zhang, H., Lv, Y., Wang, Y., Zhou, C., Xiong, R.: Reinforcement learning of serpentine locomotion for a Snake robot. In: 2021 IEEE International Conference on Real-Time Computing and Robotics (RCAR), pp. 468–473. IEEE (2021)
4. Song, Y.L., Chiu, W.Y., Manoharan, S.H.: Mixed compositional pattern-producing network-NeuroEvolution of augmenting topologies method for the locomotion control of a snake-like modular robot. In: 2023 SICE International Symposium on Control Systems (SICE ISCS), pp. 93–100. IEEE (2023)
5. Tamura, H., Kamegawa, T.: Parameter search of a CPG network using a genetic algorithm for a snake robot with tactile sensors moving on a soft floor. *Front. Robot. AI.* **10**, 1138019 (2023)
6. Wang, Z., Gao, Q., Zhao, H.: CPG-inspired locomotion control for a snake robot basing on nonlinear oscillators. *J. Intell. Robot. Syst.* **85**, 209–227 (2017)
7. Liu, X., Gasoto, R., Jiang, Z., Onal, C., Fu, J.: Learning to locomote with artificial neural-network and cpg-based control in a soft snake robot. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 7758–7765 (2020)
8. Nor, N.M., Ma, S.: CPG-based locomotion control of a snake-like robot for obstacle avoidance. In: 2014 IEEE International Conference on Robotics and Automation (ICRA), pp. 347–352 (2014)
9. Khan, S.N., Mahmood, T., Ullah, S.I., Ali, K., Ullah, A.: Motion planning for a snake robot using double deep q-learning. In: 2021 International Conference on Artificial Intelligence (ICAI), pp. 264–270. IEEE (2021)
10. Zhang, D., Ju, R., Cao, Z.: Reinforcement learning-based motion control for snake robots in complex environments. *Robotica.* **42**(4), 947–961 (2024)
11. Qi, X., Chen, D., Li, Z., Tan, X.: Back-stepping experience replay with application to model-free reinforcement learning for a soft snake robot. *IEEE Robot. Autom. Lett.* **9**, 7517–7524 (2024)
12. Ouyang, S., Wei, W.: Flexible adaptive control of snake-like robot based on LSTM and gait. *J. Phys. Conf. Ser.* **1487**(1), 012049 (2020)
13. Wang, L., et al.: Snake optimizer lstm-based uwb positioning method for unmanned crane. *PLoS One.* **18**(11) (2023)
14. Bai, Y., Xing, F., Wang, H.B., Liu, X.Y.: Research of localization method of coal mine snake detecting robot. *IAENG Int. J. Comput. Sci.* **51**(7) (2024)
15. Jiang, S., et al.: Hierarchical RL-guided large-scale navigation of a Snake robot. In: 2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM), pp. 1347–1352. IEEE (2024)
16. Hong, T., et al.: Adaptive control for a 3D snake-like robot based on mutual supervised reinforcement learning. In: 2022 International Conference on Advanced Robotics and Mechatronics (ICARM), pp. 144–150. IEEE (2022)
17. Liu, J., Tong, Y., Liu, J.: Review of snake robots in constrained environment. *Robot. Auton. Syst.* **141**, 103785 (2021)
18. Lyu, P.: Robot path planning algorithm with improved DDPG algorithm. *Int. J. Interact. Des. Manuf. (IJIDeM).* **19**(2), 1123–1133 (2025)
19. Liu, X., Gasoto, R., Jiang, Z., Onal, C., Fu, J.: Learning to locomote with artificial neural-network and cpg-based control in a soft snake robot. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 7758–7765. IEEE (2020)
20. Liu, X., Onal, C.D., Fu, J.: Reinforcement learning of CPG-regulated locomotion controller for a soft snake robot. *IEEE Trans. Robot.* **39**(5), 3382–3401 (2023)



Optimization of YOLOv12 for Steel Surface Defect Detection Under Class Imbalance

Hui Han , Yunyun Yang , Xiaoyuan Liu, Jiamin Wei , and Xinying Xu

College of Electrical and Power Engineering, Taiyuan University of Technology,
Taiyuan 030024, China
liuxiaoyuanz@126.com

Abstract. Steel surface defect detection plays a critical role in industrial quality control, where rare but high-risk defects must be accurately identified under imbalanced data conditions. As a subtask of object detection, defect detection also suffers from long-tailed class distributions, which hinder the learning of minority-class features. To address this issue, we propose a series of training-level enhancements to the YOLOv12 algorithm, including a Minority-Class Re-Learning Strategy and a Stability-Aware Loss Function. The former reinforces feature learning for underrepresented classes via a multi-stage fine-tuning and fusion process, while the latter improves localization robustness by penalizing uncertain predictions. Experimental results on the GC10-DET dataset demonstrate a 3.6% improvement in mAP@50 and 2.4% in mAP@50–95, confirming that our approach significantly improves detection performance for minority classes without compromising accuracy on majority classes, providing a practical and effective solution for industrial defect detection.

Keywords: Steel Surface Defects · Class Imbalance · YOLOv12 · Industrial Quality Inspection · Object Detection

1 Introduction

Object detection has achieved remarkable progress in recent years, driven by advances in deep learning and the availability of large-scale annotated datasets. However, a persistent challenge in practical applications is class imbalance. In many real-world scenarios, object categories follow long-tailed distributions, where a few classes dominate the sample space while others are severely underrepresented [1]. This imbalance often causes detection models to overfit majority classes and underperforms on minority classes. A representative example is the COCO dataset [2], which exhibits significant class imbalance. As shown in Fig. 1(a), class 0 (“person”) contains over 250,000 instances—several orders of magnitude more than many other classes. This distributional skew leads to model bias, making it difficult to learn robust representations for rare categories. Figure 1(b) illustrates this issue via the precision–recall curves of YOLOv12x, where substantial performance disparities across classes are observed.

This problem is equally critical in industrial applications, especially in surface defect detection. In such tasks, rare but high-risk defects must be reliably identified to ensure

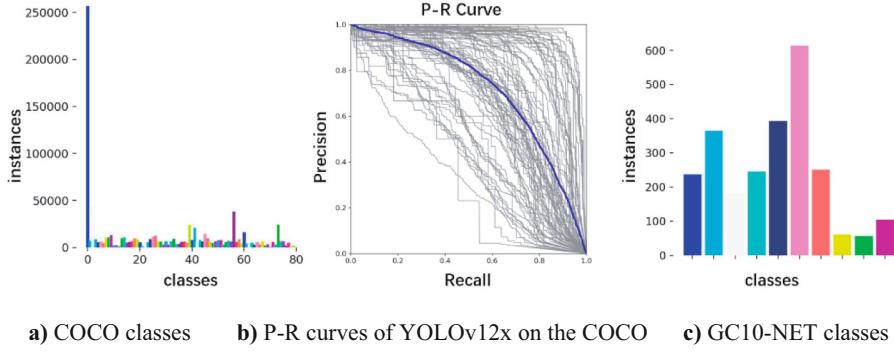


Fig. 1. Class imbalance in the COCO dataset

manufacturing quality and safety. However, models trained on imbalanced industrial datasets often fail to generalize to these minority defect types. As shown in Fig. 1(c), the GC10-DET dataset used in this study also demonstrates a highly skewed category distribution [3], highlighting that class imbalance remains a pervasive issue not only in public benchmarks but also in domain-specific detection scenarios.

To address this challenge, we focus on steel surface defect detection and adopt GC10-DET as the experimental benchmark. To improve model performance under such imbalance, we introduce two training-level enhancements: (1) a minority-class re-learning strategy that improves representation for minority classes, and (2) a stability-aware loss function that penalizes uncertain and dispersed localization distributions. We choose YOLOv12 as the base model for its enhanced attention-based feature extraction (via A2C2f blocks), which is particularly suited for capturing localized surface patterns. Moreover, its open structure facilitates integration of re-learning stages and distributional supervision. Experimental results demonstrate that the proposed method significantly improves the detection of minority-class defects while maintaining overall accuracy.

2 Optimization of YOLOv12 Target Detection Algorithm

We build our method upon the YOLOv12 architecture, a state-of-the-art object detector known for its balance of accuracy and efficiency [4]. As shown in Fig. 2, the original YOLOv12 framework consists of three main stages: feature extraction, feature fusion, and object detection. Two key modules are embedded within this architecture:

C3k2, a lightweight residual module based on the Cross Stage Partial (CSP) structure, enhances feature reuse while reducing computational overhead [5].

A2C2f, an attention-augmented fusion module, integrates multi-scale information with spatial awareness, improving the model's ability to discriminate fine-grained features under complex surface textures [6].

Despite these design strengths, YOLOv12—like other detectors—suffers from performance degradation in class-imbalanced scenarios, particularly in industrial surface defect detection, where critical but rare defect types (e.g., rolled pit, crease) are significantly underrepresented.

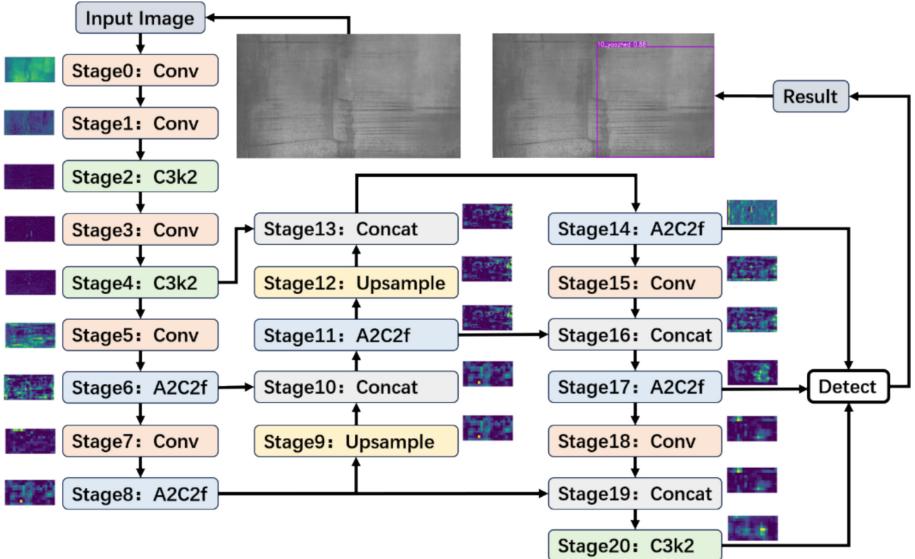


Fig. 2. Overall structural flowchart of the YOLOv12 object detection algorithm

In such settings, detection algorithms struggle to achieve uniform feature learning across defect categories. While data-level techniques like oversampling or augmentation are often used, they are insufficient to address the root cause of imbalance. To this end, we propose training-level optimizations centered on two complementary components: a Minority-Class Re-Learning Strategy, aimed at enhancing feature representations for rare defect types, and a Stability-Aware Loss Function, designed to suppress uncertain and scattered localization predictions. Together, these two components form a novel and theoretically motivated training framework that improves generalization for rare classes and enhances localization stability without introducing inference-time overhead.

2.1 Minority Class Re-learning Strategy

Class imbalance is a key obstacle in defect detection, as minority classes often exhibit weak feature activation and poor discriminability [1]. To address this, we propose a Minority-Class Re-Learning Strategy that enhances the model's ability to capture rare defect patterns. As shown in Fig. 3, feature coverage varies across classes: green and blue represent majority and minority classes respectively, with larger areas indicating stronger feature representation. White regions denote feature voids—under-activated areas resulting from insufficient training exposure.

Our strategy begins with a standard pre-training phase on the original imbalanced dataset, producing a baseline model denoted as $weight_{base}$ optimized predominantly for frequent defect types. To mitigate this, we perform single-class fine-tuning, where the model is individually trained on samples from each class. This process produces class-specific weights $\{weight_0, weight_1, \dots, weight_{n-1}\}$, which specialize in learning features

unique to their respective defect types. This targeted adaptation activates underutilized feature regions and improves class separability.

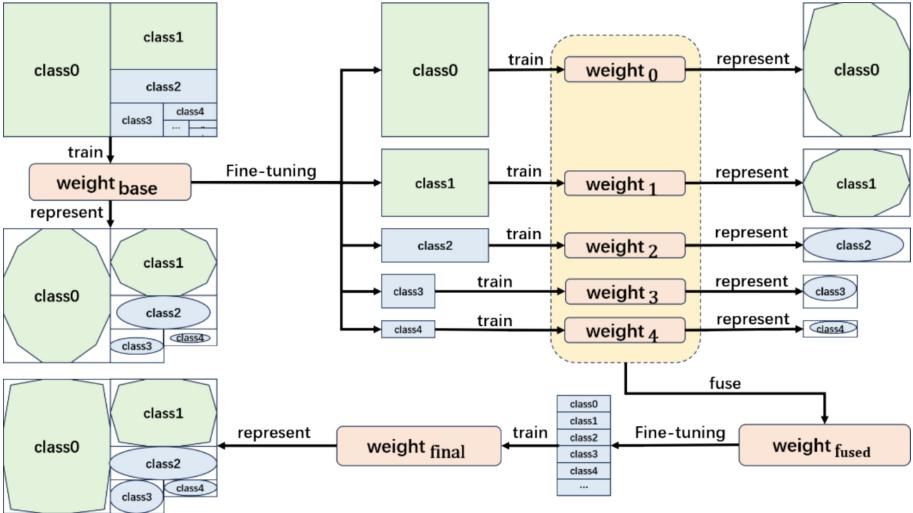


Fig. 3. Schematic of the Minority Class Re-Learning Strategy

To consolidate the improved representations, we introduce a class-aware weighted fusion mechanism (Fig. 4), which integrates the class-specific weights into a unified model.

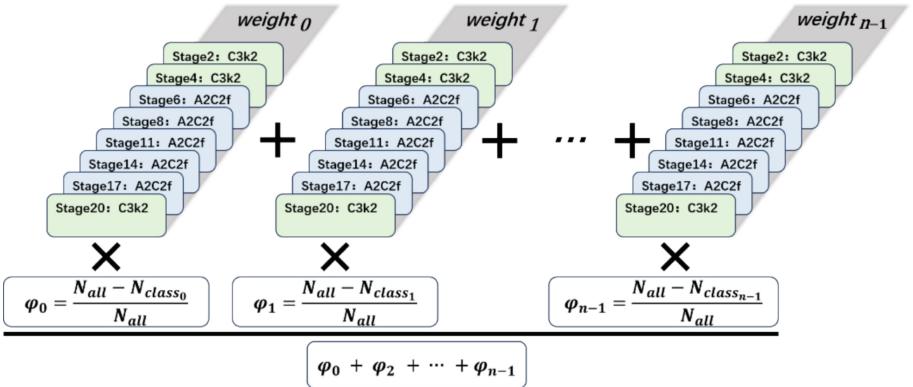


Fig. 4. Composition of $weight_{add}$

Rather than averaging all weights equally, we compute a weighted sum using coefficients φ_i derived from the inverse of class frequency:

$$\varphi_i = \frac{N_{all} - N_{class_i}}{N_{all}} \quad (1)$$

where N_{all} is the total number of training samples and N_{class_i} denotes the sample count for class i . This weighting scheme assigns greater influence to minority defect classes, ensuring their features are emphasized during fusion. The fusion primarily targets the C3k2 and A2C2f modules, which govern spatial encoding and attention mechanisms. Other modules are inherited directly from the baseline model to preserve generalization capability. The final fused model ($weight_{fused}$) is shown in Fig. 5.

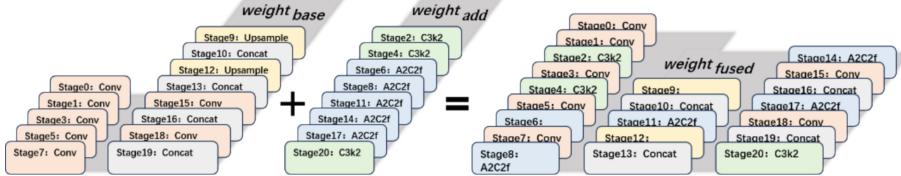


Fig. 5. Composition of $weight_{fused}$

In the final stage, the fused model is re-trained on a re-balanced version of the dataset, in which sample distributions are adjusted to emphasize minority categories. This stage focuses on semantic refinement under a balanced learning regime, enabling the model to better generalize across all defect types. Compared to directly training on $weight_{base}$, our re-learning strategy achieves significantly improved detection performance for rare defects, while maintaining accuracy on majority categories.

Compared to existing approaches such as Focal Loss [7] and Class-Balanced Loss [8], our proposed strategies differ fundamentally in their operating principle and application scope. While Focal and CB Loss modify the classification loss by adjusting sample weights based on difficulty or class frequency, our Minority-Class Re-Learning Strategy acts at the model level: it explicitly reshapes the parameter space via sequential fine-tuning and weight fusion, ensuring that rare classes receive dedicated optimization without compromising majority-class performance. This weight-space intervention mechanism is orthogonal to traditional probability-space reweighting.

2.2 Enhanced Loss Function with Stability-Aware Penalty

This section introduces a stability-aware enhancement to the YOLOv12 loss function. The original formulation comprises box regression loss, classification loss, and discrete feature loss (DFL). The DFL term guides the model to learn discrete probability distributions for the distances from the predicted center to each box boundary (d_t, d_b, d_l, d_r), ensuring the distribution's expectation aligns with the ground truth [7]. Ideally, these distributions are sharp and concentrated. However, during training, various suboptimal distribution patterns may arise, undermining localization accuracy.

As illustrated in Fig. 6, we categorize predicted distributions into four types:

Type I: Delta-like, with probability mass at a single bin—highly reliable.

Type II: Concentrated over two adjacent bins—still reliable.

Type III: Flat or noisy, indicating low confidence.

Type IV: Multimodal with distant peaks, suggesting semantic ambiguity.

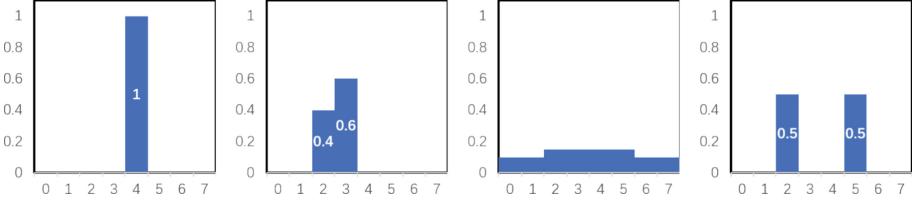


Fig. 6. Types of Distributions for Bounding Box Parameters

To mitigate the negative effects of unreliable distributions, we introduce an enhanced DFL loss that penalizes dispersion and enforces distributional concentration. Let the predicted distribution be $p = (p_0, p_1, \dots, p_N)$. We compute its variance as $\text{Var}(p)$. To discourage flat or dispersed distributions, we introduce an inverse-variance penalty:

$$L_{\text{var}} = \alpha \times \frac{1}{\text{Var}(p) + \varepsilon} \quad (2)$$

where α is a hyperparameter controlling the penalty strength, and ε is a small constant for numerical stability. This term encourages the network to produce high-variance (i.e., low-variance, sharply peaked) distributions. To further suppress Type IV distributions, we introduce a constraint mask M , which is activated only when the distribution is concentrated in one or two adjacent bins. Specifically,

$$M(p) = \begin{cases} 1, & \text{if } \exists i \text{ s.t. } p_i > \delta \text{ and } (p_{i-1} > \delta \text{ or } p_{i+1} > \delta) \\ & \text{or } p_i > \theta \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Here, θ is a threshold for single-bin sharp peaks, and δ for adjacent-bin concentration. The final enhanced DFL loss is then formulated as:

$$L_{\text{DFL-enhanced}} = L_{\text{DFL}} + \lambda_{\text{var}} \times M(p) \times L_{\text{var}} \quad (4)$$

Finally, the enhanced loss function is

$$\text{loss} = \lambda_1 \times \text{loss}_{\text{BOX}} + \lambda_2 \times \text{loss}_{\text{CLS}} + \lambda_3 \times \text{loss}_{\text{DFL-enhanced}} \quad (5)$$

This design enforces the concentration of probability mass around adjacent locations while penalizing unstable or fragmented distributions, thus improving the localization robustness of the model, especially under challenging conditions.

The proposed Stability-Aware Loss extends the Distribution Focal Loss (DFL) framework not by simply adding a penalty term, but by addressing a key overlooked issue: under class imbalance, the predicted discrete distributions over box offsets become noisy, fragmented, and multi-modal, especially for minority classes. To counteract this, we introduce a variance-aware regularization term that penalizes high-uncertainty predictions and encourages unimodal, compact distributions. This formulation is inspired by the statistical intuition that reliable localization should produce sharp, narrow distributions. The additional adjacency-constrained mask further restricts the mass to lie within ± 1 of the most probable offset, enforcing shape consistency and improving convergence.

3 Experimental Results and Analysis

Experiments were conducted on a high-performance platform running openEuler 22.03, equipped with dual Intel Xeon Silver 4316 CPUs and two NVIDIA RTX 4090 GPUs (24 GB VRAM each), enabling efficient training with high-resolution images.

To evaluate the effectiveness of our proposed method, we conduct comprehensive experiments on the GC10-DET dataset, comparing it with several representative approaches for addressing class imbalance. The dataset comprises 2,306 images and 3,570 annotated instances across 10 defect types: punching hole, welding line, crescent gap, water spot, oil spot, silk spot, inclusion, rolled pit, crease, and waist folding. It was randomly split into 70% training and 30% validation.

Table 1. Comparison with Mainstream Imbalance Handling Methods on GC10-DET

Methods	mAP@50	mAP@50–95	Params(M)	GFLOPs
Faster R-CNN [9]	0.589	0.296	28.48	941.2
SSD [10]	0.612	0.327	24.81	61.7
YOLOv3-tiny [11]	0.587	0.294	12.13	18.9
YOLOv6n [12]	0.639	0.315	4.23	11.8
YOLOv8n [13]	0.633	0.327	3.01	8.1
YOLOv10n [14]	0.600	0.332	2.71	8.2
YOLOv11n [5]	0.642	0.331	2.58	6.3
RT-DETR-R18 [15]	0.6474	0.314	19.88	57.0
SDS-YOLOv8n [16]	0.645	0.332	2.85	7.4
EMA-YOLO [17]	0.644	0.330	2.92	12.3
DCDF-YOLO [18]	0.678	0.345	2.15	7.3
YOLOv12n [4]	0.674	0.339	2.56	6.3
Ours	0.710	0.363	2.56	6.3

As shown in Table 1, our method achieves the highest detection performance across both mAP@50 and mAP@50–95 metrics. Notably, despite its superior accuracy, our model also maintains low parameter count and computational cost, which highlights its suitability for deployment in industrial scenarios with limited resources. While our approach introduces an additional fine-tuning stage during training, it is important to note that the inference-time model remains structurally identical to the baseline, incurring no extra parameters or runtime costs. The increased training cost is moderate and acceptable in industrial settings where inference speed is the primary concern.

This performance advantage stems from two synergistic components: the minority-class re-learning strategy and the stability-aware loss. Following the approach described in Sect. 2, we first train YOLOv12n on the imbalanced dataset to obtain the baseline weights ($weight_{base}$). Then, we perform single-class fine-tuning to derive a set of specialized weights $\{weight_0, weight_1, \dots, weight_9\}$, each optimized for one defect category.

Table 2. Comparison of mAP@50 across different training weights

<i>class_id</i>	0	1	2	3	4	5	6	7	8	9	all
instances	98	156	87	111	177	271	100	27	18	42	1087
<i>weight_{base}</i>	0.937	0.874	0.948	0.804	0.733	0.595	0.378	0.174	0.523	0.778	0.674
<i>weight₀</i>	0.995	0.27	0.882	0.79	0.619	0.486	0.263	0.167	0.213	0.572	0.526
<i>weight₁</i>	0.889	0.986	0.771	0.713	0.637	0.488	0.233	0.152	0.447	0.673	0.6
<i>weight₂</i>	0.933	0.877	0.982	0.8	0.706	0.592	0.367	0.17	0.433	0.789	0.665
<i>weight₃</i>	0.951	0.793	0.887	0.864	0.613	0.422	0.246	0.197	0.338	0.554	0.587
<i>weight₄</i>	0.934	0.399	0.666	0	0.794	0.305	0.174	0.37	0.409	0	0.405
<i>weight₅</i>	0.759	0.377	0.581	0.467	0.514	0.712	0.073	0.063	0	0	0.355
<i>weight₆</i>	0.953	0.822	0.929	0.727	0.655	0.527	0.541	0.198	0.459	0.708	0.652
<i>weight₇</i>	0.902	0.481	0.894	0.715	0.645	0.479	0.27	0.458	0.243	0.462	0.555
<i>weight₈</i>	0.933	0.771	0.925	0.741	0.673	0.551	0.291	0.078	0.881	0.525	0.637
<i>weight₉</i>	0.964	0.678	0.906	0.721	0.553	0.482	0.274	0.116	0.386	0.861	0.594
<i>weight_{final}</i>	0.937	0.851	0.947	0.808	0.725	0.601	0.351	0.518	0.561	0.803	0.71

Table 3. Comparison of mAP@50–95 across different training weights

<i>class_id</i>	0	1	2	3	4	5	6	7	8	9	all
<i>weight_{base}</i>	0.555	0.339	0.617	0.45	0.313	0.266	0.125	0.104	0.246	0.373	0.339
<i>weight₀</i>	0.646	0.13	0.549	0.463	0.262	0.217	0.085	0.098	0.06	0.317	0.282
<i>weight₁</i>	0.542	0.469	0.489	0.429	0.262	0.194	0.08	0.066	0.177	0.381	0.309
<i>weight₂</i>	0.549	0.336	0.664	0.452	0.315	0.263	0.113	0.102	0.199	0.382	0.338
<i>weight₃</i>	0.467	0.288	0.572	0.52	0.265	0.188	0.093	0.09	0.157	0.264	0.29
<i>weight₄</i>	0.574	0.179	0.458	0	0.346	0.174	0.062	0.23	0.105	0	0.212
<i>weight₅</i>	0.355	0.127	0.391	0.24	0.211	0.306	0.026	0.046	0	0	0.17
<i>weight₆</i>	0.515	0.321	0.63	0.379	0.281	0.228	0.189	0.082	0.218	0.366	0.321
<i>weight₇</i>	0.412	0.223	0.613	0.403	0.257	0.205	0.092	0.143	0.12	0.202	0.267
<i>weight₈</i>	0.47	0.226	0.609	0.385	0.271	0.221	0.109	0.057	0.403	0.298	0.305
<i>weight₉</i>	0.501	0.213	0.579	0.397	0.225	0.193	0.07	0.05	0.142	0.447	0.281
<i>weight_{final}</i>	0.551	0.317	0.637	0.507	0.353	0.256	0.115	0.26	0.222	0.408	0.363

These are integrated using a class-aware weighted fusion mechanism, and the resulting model is re-trained on a re-balanced dataset to produce the final weights (*weight_{final}*).

Performance at each stage is summarized in Tables 2 and 3. The class-specific models exhibit strong performance on individual categories but suffer from reduced generalization due to the absence of multi-class context. In contrast, the fused and re-trained

model effectively combines class-specific knowledge, achieving a 3.6% improvement in mAP@50 and 2.4% in mAP@50–95 over the baseline.

Unlike prior class-imbalance methods that adjust sample weights during loss computation, our approach directly intervenes in the model parameter space, enabling targeted optimization for minority classes. Although additional training stages are required, the final model remains structurally identical to the baseline and introduces no extra inference cost. Moreover, the re-learning and fusion process is architecture-agnostic and can be applied to other lightweight detectors with minimal adjustment.

4 Conclusion

This paper proposes a novel optimization framework for YOLOv12 to address the class imbalance problem in steel surface defect detection. By introducing a minority-class re-learning strategy and a stability-aware loss function, the approach enhances detection accuracy for rare defect types while maintaining high performance on majority classes.

Comparative experiments on the GC10-DET dataset demonstrate that our method outperforms mainstream class-imbalance remedies (e.g., Focal Loss, Class-Balanced Loss) in both mAP and localization precision, while preserving the lowest parameter count and inference complexity. Despite incorporating additional fine-tuning stages during training, the model introduces no overhead at inference, making it practical for industrial deployment.

The two proposed modules operate in orthogonal spaces—weight space and distribution space—and together form a generalizable solution for small-sample object detection. Future work will explore adapting these strategies to more complex industrial datasets, and integrating them with transformer-based detection backbones to further improve generalization under low-resource conditions.

Acknowledgments. This work was supported by the Shanxi Natural Science Foundation under Grant 202303021221002, and by the Open Project Fund of the State Key Laboratory of Virtual Reality Technology and Systems at Beihang University under Grant VRLAB2023A06.

References

1. Ghosh, K., Bellinger, C., Corizzo, R., et al.: The class imbalance problem in deep learning. *Mach. Learn.* **113**(7), 4845–4901 (2024)
2. Lin, T.Y., Maire, M., Belongie, S., et al.: Microsoft coco: common objects in context. In: Computer vision–ECCV 2014: 13th European Conference, pp. 740–755. Springer International Publishing (2014)
3. Lv, X., Duan, F., Jiang, J., et al.: Deep metallic surface defect detection: the new benchmark and detection network. *Sensors.* **20**(6), 1562 (2020)
4. Tian, Y., Ye, Q., Doermann, D.: Yolov12: attention-centric real-time object detectors. arXiv preprint arXiv:2502.12524 (2025)
5. Khanam, R., Hussain, M.: Yolov11: an overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725 (2024).

6. Li, R., Gao, Y., Zhang, R.: An improved YOLOv5-based small target detection method for UAV aerial image. In: Chinese Intelligent Automation Conference, pp. 298–312. Springer Nature Singapore, Singapore (2023)
7. Li, X., et al.: Generalized focal loss: towards efficient representation learning for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(3), 3139–3153 (2022)
8. Cui, Y., et al.: Class-balanced loss based on effective number of samples. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9268–9277 (2019)
9. Girshick, R.: Fast r-cnn. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448 (2015)
10. Liu, W., Anguelov, D., Erhan, D., et al.: Ssd: single shot multibox detector. Computer vision–ECCV 2016. In: 14th European Conference, pp. 21–37. Springer International Publishing (2016)
11. Redmon, J., Farhadi, A.: Yolov3: an incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
12. Li, C., et al.: YOLOv6: a single-stage object detection framework for industrial applications. arXiv preprint arXiv:2209.02976 (2022)
13. Varghese, R., Sambath, M.: Yolov8: a novel object detection algorithm with enhanced performance and robustness. In: 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICCS), pp. 1–6. IEEE (2024)
14. Wang, A., et al.: Yolov10: real-time end-to-end object detection. *Adv. Neural Inf. Proces. Syst.* **37**, 107984–108011 (2024)
15. Zhao, Y., et al.: Detrs beat yolos on real-time object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 16965–16974 (2024)
16. Guo, S., et al.: SDS-YOLOv8n: a lightweight detection method for flames and smoke. *J. Phys. Conf. Ser.* **2858**(1), 012020 (2024)
17. Xu, D., et al.: EMA-YOLO: a novel target-detection algorithm for immature yellow peach based on YOLOv8. *Sensors* **24**(12), 3783 (2024)
18. Li, C., et al.: An efficient lightweight detection model for steel surface defects with dynamic deformable head. *Eng. Res. Express.* **7**(1), 015282 (2025)



Autonomous Landing Control Method for UAV Multistate Platform Based on Dynamic Weights PPO

Yalin He, Bo Huang, Yinlong Yuan^(✉), and Yun Cheng

Nantong University, Nantong, China
yuanyl@ntu.edu.cn

Abstract. In recent years, UAVs have been widely used in a variety of fields such as firefighting, emergency communications, emergency lighting and surveillance. However, how to ensure the stable landing of UAVs on mobile platforms is still an issue of concern. Therefore, for the multimodal stability control challenges of autonomous landing of UAVs in complex dynamic environments, a cooperative control algorithm based on dynamic weight Proximal Policy Optimization (PPO) is proposed, the core of which is to reconstruct the positional error reward function through a dynamic weight adjustment mechanism and to establish a reinforcement learning model containing initial state constraints. The method addresses the conflict between trajectory tracking accuracy and attitude adjustment efficiency through the dynamic reward weight allocation strategy for stationary and moving dual-state landing platform scenarios. Simulation results show that the method can adapt to different platform motion noise conditions, providing a feasible technical path for robust autonomous landing control of UAVs in complex dynamic environments, which is of methodological reference value for multi-objective co-optimization research in the field of motion control of intelligent bodies.

Keywords: Proximal Policy Optimization · dynamic weighted positional rewards · autonomous landing control

1 Introduction

With the development of autonomous control technology, UAVs have gained wide application in various industries such as delivery [1] and agriculture [2]. Landing is an important stage for UAVs to perform tasks, and quadcopter UAVs usually first positioned to land the platform and then controlled to achieve a stable landing [3].

[4] investigated an autonomous precision UAV landing method utilizing a gimbaled multi-payload camera, where the image information acquired by the on-board camera is then used to control the UAV to stabilize the landing, and the sensors are dynamically switched to adapt to different distances and environments, which improves the detection range and adaptability. However, its image requirements are high, and once the image is

not recognized for a long time, the landing cannot be stabilized [5]. investigates a multi-stage visual guidance strategy based on deep reinforcement learning, which optimizes the accuracy of the basic flight movements of UAVs, fills in the shortcomings of the traditional GPS-based or sensor fusion methods. However, this method relies heavily on markers and does not take into account factors such as light influence and marker breakage, and the starting point, trajectory and target are fixed in the experiment, which is not adaptable enough in the actual task [6]. investigated decoupling visual detection from reinforcement learning control and using vision in combination with TD3 algorithm to realize autonomous landing, which can improve the scalability of the system, but the landing angle is not good considering the three-dimensional relative position, and it is easy to lose the target when the distance between the UAV and the platform is less than 3 m [7]. proposed a layered control architecture, where the high-level uses DDPG to learn to regulate the low-level PID parameters, realizing system self-adaptation to the environment and autonomous landing of the UAV [8, 9]. investigated a dual-module framework to improve the efficiency of error correction, a high degree of vision-IMU fusion to achieve lost recapture markers, respectively. They are more robust. However, it relies heavily on the UAV status information, and the computational cost is high [10]. proposed a new reinforcement learning framework, ROS-RL, which relies on the generalized interface and has high value for engineering applications, and is controlled with reinforcement learning in the horizontal direction and a fixed function in the vertical direction, but the dynamics coupling is neglected, and the reward function does not take into account the stability of the attitude angle.

The common points of these researches on UAV landing control aspects are: high dependence on prediction, conflict between trajectory tracking accuracy and attitude adjustment efficiency, and bad landing angle. But the landing of UAV is a highly nonlinear change process, and the demand of reward function is different for different relative states of the landing platform, and the prediction results have a large deviation, which makes it difficult to realize the stable landing of multi-scenarios.

Aiming at the above problems, this study proposes an autonomous landing control method based on dynamic weight PPO (Proximal Policy Optimization) for UAV multi-state platforms, which accomplishes the task of autonomous landing of UAVs to stationary and moving bi-state platforms. The position error reward function is reconstructed through the dynamic weight adjustment mechanism, and a reinforcement learning model containing initial state constraints is established to solve the conflict between trajectory tracking accuracy and attitude adjustment efficiency. Experimental results on the UE4 simulation platform show that the autonomous landing control method for UAV multi-state platforms using dynamic weight PPO can adapt to different platform motion noise conditions and exhibit high robustness.

2 Autonomous Landing Control of UAV Multistate Platform Based on Dynamic Weight PPO

2.1 Modeling of the Autonomous Landing Problem of UAV

The dynamics model of the UAV and mobile platform in this study adopts the default dynamics model in AirSim: Newton's classical mechanics model, aerodynamics model, and quadrotor UAV dynamics model. Autonomous UAV landing requires consideration of UAV state information $S(\Delta x, \Delta y, \Delta z, v_x, v_y, v_z)$, such as the UAV distance displacement ($\Delta x, \Delta y, \Delta z$), velocity state $a(v_x, v_y, v_z)$.

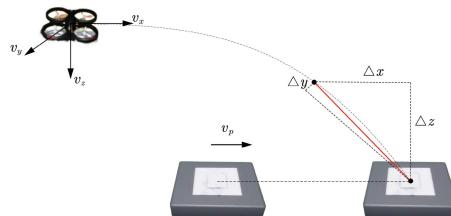


Fig. 1. State of autonomous landing motion of drone.

The motion state of UAV autonomous landing is shown in Fig. 1. To simulate more realistically, the initial speed of UAV is randomly determined over a range of speeds.

The landing platform also has different speeds. If the landing platform is moving, it will move at a relatively low speed to ensure that the UAV can realize the stable and autonomous landing. Even if the landing of the UAV is implemented in some extreme environments, the selected landing platform will not move too fast.

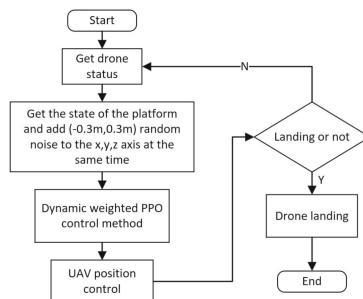


Fig. 2. Flowchart for autonomous drone landing.

In order to improve the robustness of the method in this study, positioning noise is added, and after obtaining the platform's positioning, positioning noise of $(-0.3 m, 0.3 m)$ is randomly added to its x, y, z axis. The specific process of autonomous landing of UAV in this study is shown in Fig. 2.

2.2 Dynamic Weighted Position Error Reward Function

UAV needs to have dynamic position error rewards at different relative positions, to design a position error reward function (1) based on a dynamic weight adjustment mechanism, which has higher flexibility compared to fixed position rewards.

$$R = w_1 \times H_r + w_2 \times V_r \quad (1)$$

In the above equation, H_r is the horizontal reward, V_r is the vertical reward, and w_1 and w_2 are the assigned weights of the horizontal and vertical rewards respectively, i.e., the horizontal and vertical reward coefficients.

$$H_r = \begin{cases} (H_{err_{last}} - H_{err_{now}}) \times 500, & H_{err_{last}} \neq 0 \\ 0, & H_{err_{last}} = 0 \end{cases} \quad (2)$$

$$V_r = \begin{cases} (V_{err_{last}} - V_{err_{now}}) \times 500, & V_{err_{last}} \neq 0 \\ 0, & V_{err_{last}} = 0 \end{cases} \quad (3)$$

Specifically, the UAV is firstly controlled to make adjustments in the horizontal direction to reduce the horizontal error H_{err} , and then control the vertical direction. Eq. (4) defines the horizontal reward coefficient, and Eq. (5) defines the relationship between the horizontal and vertical rewards:

$$w_1 = \begin{cases} 0.9, & H_{err} > 2 \\ 0.8, & H_{err} < 2 \text{ and } H_{err} > 1.5 \\ 0.7, & H_{err} < 1.5 \text{ and } H_{err} > 1 \\ 0.4, & H_{err} < 1 \text{ and } H_{err} > 0.8 \\ 0.3, & others \end{cases} \quad (4)$$

$$w_2 = 1 - w_1 \quad (5)$$

In order to prevent the UAV from repeatedly bouncing up and down and excessive movements during landing, the height rise penalty of Eq. (6) is introduced to ensure the stability in the vertical direction during landing:

$$V_r^- = \begin{cases} 50, & V_{err_{last}} < V_{err_{now}} \\ 0, & others \end{cases} \quad (6)$$

V_{err} is the vertical direction error. When the horizontal error H_{err} is larger, the horizontal reward coefficient w_1 is larger, which promotes the UAV to adjust horizontally. When the horizontal error is smaller, the vertical reward coefficient w_2 is larger, which promotes the UAV to land. When the horizontal error meets the landing requirement $w_2 > w_1$, the UAV focuses more on the vertical direction of the landing.

2.3 Control Based on PPO Algorithm

2.3.1 Network Structure of the PPO Algorithm

The PPO algorithm is iteratively updated with a small number of samples during training, and training is stable. For autonomous landings that may produce excessive mutations, it can directly limit the magnitude of policy updates and ensure that the UAV can adapt to multi-state landing platforms. The specific network structure is Fig. 3.

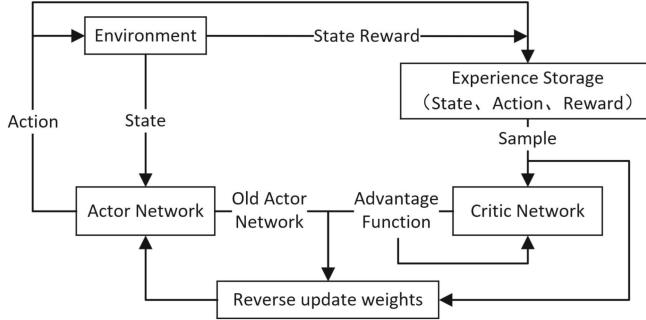


Fig. 3. Schematic diagram of PPO network structure.

2.3.2 Network Updates for PPOs

The PPO algorithm improves the stability and efficiency of strategy updating by estimating the advantage function more accurately by accumulating rewards across multiple time steps. In UAV landing, PPO needs to start calculating the cumulative rewards of each state and the advantage of the action after completing a trajectory. For each network update, specifically: after obtaining the initial state of the UAV and the shutdown platform, calculate the horizontal distance d_h and vertical distance d_v :

$$d_h = \|\Delta p_{xy}\| \quad (7)$$

$$d_v = \|\Delta p_z\| \quad (8)$$

If the landing is successful, i.e., the relative position distance $|\Delta p| = |p_d - p_p| < 0.5$, updating the distance reward (9):

$$r \leftarrow r + 100 \quad (9)$$

If a collision is detected, the UAV receives a distance penalty (10):

$$r \leftarrow r - 200d_h \quad (10)$$

For the remaining cases, the distance reward is Eq. (11):

$$r_{dist} = w_h(d_h^{t-1} - d_h) + w_v(d_v^{t-1} - d_v) \quad (11)$$

2.3.3 Model Training and Parameter Optimization Process

The pseudo-code of the reinforcement learning algorithm for UAV dynamic platform landing is shown in Table 1.

Table 1. Reinforcement learning algorithm pseudo-code for dynamic platform landing of UAVs.**Algorithm 1** Reinforcement Learning for Dynamic Drone Landing

```

1:   Initialize AirSim environment, PPO model
2:   Define observation space (relative position + platform velocity)
3:   Define action space (3D velocity control)
4:   while Training not completed do
5:     Reset environment
6:     Randomize drone position near platform
7:     for step=1 to MAX_STEPS do
8:       Get observation: drone_pos-platform_pos
9:       Predict action  $a_t = (v_x, v_y, v_z)$  using model
10:      Execute action in simulation
11:      Calculate reward:
        • Horizontal / vertical distance improvement
        • Landing success (+100) / collision penalty
12:      if collision or successful landing then
13:        Terminate episode
14:      end if
15:    end for
16:  end while

```

3 Experimental Results and Analysis

3.1 Experimental Setup

**Fig. 4.** Schematic diagram of simulation environment.

The method proposed in this study is trained in UE4 simulation environment and controlled based on AirSim. Figure 4 is a schematic diagram of the relative positions of the UAV and the landing platform in the simulation environment.

**Fig. 5.** Autonomous landing process in simulation environment.

The autonomous landing process in the simulation environment is shown in Fig. 5. The position parameter information of the UAV and platform is obtained directly from AirSim. The limit on the initial speed of the drone is: $v_x, v_y, v_z \in [-0.5, 0.5]$, the speed will also be reduced during landing. Three groups of comparison experiments were done with the speed of the landing platform as 0 m/s, 0.1 m/s, 0.2 m/s. We set some of the hyperparameters of the PPO as follows (Table 2):

Table 2. PPO partial hyperparameters and related data.

Symbol	Meaning	Value
learning_rate	The rate of learning	3×10^{-4}
n_steps	The number of steps collected in each training round	1024
batch_size	The number of samples used for each gradient update during training	64
gamma	Discount Factor	0.995
total_timesteps	Training step length	30000

3.2 Experimental Results and Analysis

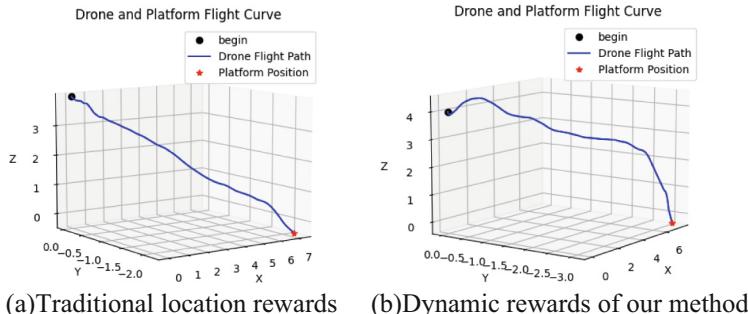


Fig. 6. Landing trajectory when the speed of the landing platform is 0 m/s.

Firstly, the speed of the stopping platform is set to 0 m/s, and the landing trajectory of the UAV is shown in Figs. 6(a), PPO landing trajectory based on the traditional positional reward, and Fig. 6(b) PPO landing trajectory based on the present method.

As can be seen in Fig. 6, the landing trajectory of the PPO rewarded by the traditional position is approximated to a straight line, although this results in the shortest trajectory of the UAV movement, but the over-pursuit of the shortest distance leads to the neglect of the angle of the UAV landing and the horizontal errors. In contrast, the landing trajectory of our method is similar to a parabola, when the horizontal error is too large, the UAV altitude decreases very little, and focuses more on adjusting the horizontal error, i.e., the previous segment of the trajectory in Fig. 6(b); when the horizontal error is small, the

UAV begins to focus on landing, almost perpendicular to the landing platform, reducing damage to the UAV and the landing platform.

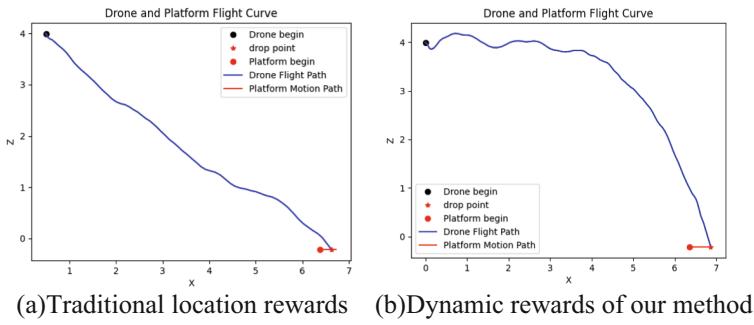


Fig. 7. Landing trajectory when the speed of the landing platform is 0.1 m/s .

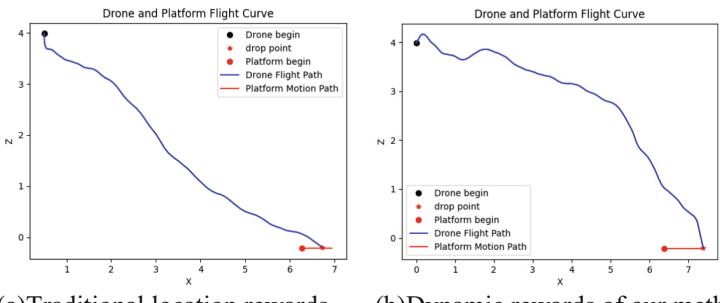


Fig. 8. Landing trajectory when the speed of the landing platform is 0.2 m/s .

In order to further verify the feasibility of the method proposed in this study, a comparison experiment was set up when the moving speed of the landing platform was 0.1 m/s and 0.2 m/s , and the results are shown in Figs. 7 and 8.

In Figs. 7 and 8 above, the red line is the motion trajectory of the landing platform, and the blue curve is the UAV landing trajectory, and it can be found that using the method proposed in this study for landing, the results are both focusing on horizontal position adjustment in the previous segment of the trajectory, and more focusing on the vertical direction in the latter stage, which is in line with the expectation of the dynamic weighting reward function design. However, our method focuses on horizontal position adjustment in the first phase, resulting in a longer landing distance compared to the PPO based on traditional positional reward. For autonomous UAV landing, the stability and flexibility of autonomous UAV landing is worse if only the reward component in horizontal or vertical direction is removed, so the fixed weight reward is selected to compare with the dynamic reward of this method, which directly reflects the effectiveness and flexibility of our method.

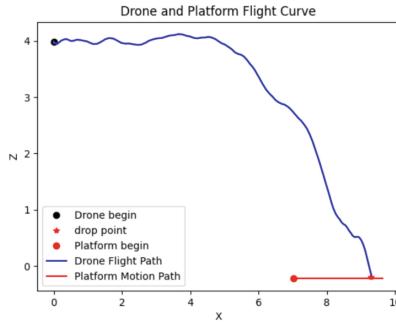


Fig. 9. UAV landing trajectory at 0.5 m/s speed of the landing platform.

In special cases, the landing platform will move faster, for this situation, the simulation experiment under the speed of 0.5 m/s of the landing platform is carried out, and it is found that the landing control method proposed in this study can still realize a better landing effect. The specific landing trajectory is shown in Fig. 9. To a certain extent, it can be seen that the method of this study solves the problem of conflict between trajectory tracking accuracy and attitude adjustment efficiency.

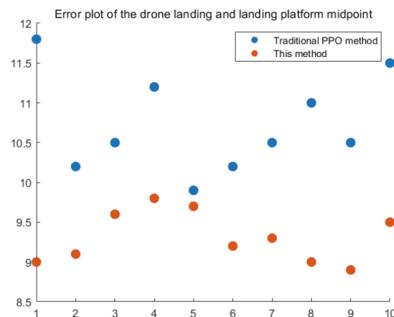


Fig. 10. Error plot of the drone landing and landing platform midpoint.

By analyzing Fig. 10, it is relatively intuitive to see that the landing error of the method proposed in this study is smaller compared to the PPO based on the traditional position reward and the landing using the method proposed in this study, specifically, the mean value of the landing error of the PPO based on the traditional position reward is 10.73 cm , and the landing error of the method proposed in this study is 9.31 cm .

4 Conclusion

Aiming at the multimodal stability control difficulties of autonomous landing of UAVs in complex dynamic environments, a dynamic weight PPO-based autonomous landing control method for UAVs with multimodal platforms is proposed. By reconfiguring the position error reward function through the dynamic weight adjustment mechanism,

combined with a reinforcement learning model containing initial state constraints, the autonomous stable landing of UAVs for stationary and moving dual-state platforms is realized, and at the same time, the conflict between the trajectory tracking accuracy and the attitude adjustment efficiency is solved. The experimental results show that with this research method, the UAV is able to adapt to different platform motion noise conditions during autonomous landing, and the autonomous landing realized by this research method is more stable and with less error than the traditional position rewarded PPO landing, which provides a more reliable reference for the robust autonomous landing control of UAVs in complex dynamic environments.

References

1. Aurambout, J.P., Gkoumas, K., Ciuffo, B.: Last mile delivery by drones: an estimation of viable market potential and access to citizens across European cities. *Eur. Transp. Res. Rev.* **11**(1), 1–21 (2019)
2. Chmaj, G., Selvaraj, H.: Distributed processing applications for UAV/drones: a survey. In: *Progress in Systems Engineering: Proceedings of the Twenty-Third International Conference on Systems Engineering*, pp. 449–454. Springer International Publishing (2015)
3. Xia, K., Shin, M., Chung, W., Kim, M., Lee, S., Son, H.: Landing a quadrotor UAV on a moving platform with sway motion using robust control. *Control. Eng. Pract.* **128**, 105288 (2022)
4. Springer, J., Guðmundsson, G.P., Kyas, M.: A precision drone landing system using visual and IR fiducial markers and a multi-payload camera. *arxiv preprint arxiv:2403.03806* (2024)
5. Chang, C.C., et al.: Accuracy improvement of autonomous straight take-off, flying forward, and landing of a drone with deep reinforcement learning. *Int. J. Comput. Intell. Syst.* **13**(1), 914–919 (2020)
6. Wang, C., et al.: Integrated learning-based framework for autonomous quadrotor UAV landing on a collaborative moving UGV. *IEEE Trans. Veh. Technol.* **73**, 16092–16107 (2024)
7. Wu, L., et al.: Deep reinforcement learning with corrective feedback for autonomous uav landing on a mobile platform. *Drones*. **6**(9), 238 (2022)
8. Saryazdi, S., et al.: Using reinforcement learning and error models for drone precise landing. *ACM Trans. Internet Technol.* **24**(3), 1–30 (2024)
9. Mu, L., et al.: Autonomous landing guidance for quad-UAVs based on visual image and altitude estimation. *Drones*. **9**(1), 57 (2025)
10. Jiang, Z., Song, G.: A deep reinforcement learning strategy for UAV autonomous landing on a platform. In: *2022 International Conference on Computing, Robotics and System Sciences (ICRSS)*, pp. 104–109. IEEE (2022)
11. Schulman, J., et al.: Proximal policy optimization algorithms *arxiv preprint arxiv:1707.06347* (2017)
12. Kalidas, A.P., et al.: Deep reinforcement learning for vision-based navigation of UAVs in avoiding stationary and mobile obstacles. *Drones*. **7**(4), 245 (2023)
13. Truong, N.Q., et al.: SlimDeblurGAN-based motion deblurring and marker detection for autonomous drone landing. *Sensors*. **20**(14), 3918 (2020)
14. Wu, Z., et al.: Autonomous UAV landing system based on visual navigation. In: *2019 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1–6. IEEE (2019)
15. Hodge, V.J., Hawkins, R., Alexander, R.: Deep reinforcement learning for drone navigation using sensor data. *Neural Comput. Appl.* **33**(6), 2015–2033 (2021)



Real-Time Intrusion Object Detection in Low-Light Rail Surveillance Images Based on Triple Attention Mechanism

Xuan Xu, Yi Wang, Yeting Gong, Yuhang Shi, Weibin Zhu,
and Tangwen Yang^(✉)

School of Computer Science and Technology, Beijing Jiaotong University, Beijing
100044, China
twyang@bjtu.edu.cn

Abstract. Railway transportation is important modern infrastructure, and video surveillance systems are frequently used to ensure safety and efficiency. However, intrusion object detection from surveillance images faces significant challenges under low-light condition, which is common during nighttime or within tunnel environment. These low-illumination conditions often result in elevated noise levels, diminished contrast ratios, and loss of critical visual details, thereby impairing the performance of conventional detection algorithms. To address these issues, an enhanced YOLOv8-based algorithm model that incorporates a triple attention mechanism is proposed to improve the accuracy and efficiency of intrusion object detection in low-light environments. We first introduce a comprehensive railway-specific dataset that includes self-collected railway surveillance images and relevant classes from the EXDark dataset. The triple attention mechanism that synergistically combines channel, spatial and temporal attention for improved feature representation is then presented. The experimental results show that the proposed model demonstrates real-time processing capabilities, meeting the stringent operational response requirements of railway applications.

Keywords: rail perimeter surveillance · low-light image · object detection · triple attention mechanism

1 Introduction

Railway transportation, as critical infrastructure, faces significant monitoring challenges under low-light conditions such as nighttime and tunnel environments. These conditions lead to severe image degradation including increased noise, reduced contrast, and loss of visual details, which substantially impair the performance of conventional computer vision methods [1, 2].

This work is supported in part by the National Key Research and Development Program of China under grant 2022YFC3005202.

To address the specific requirements of railway surveillance, we propose an innovative solution. The system must meet three core demands: high detection accuracy, real-time processing capability, and strong robustness against sudden illumination changes [3]. Our main contributions include: (1) a novel triple-attention mechanism combining length, width, and spatial attention to enhance feature extraction [4]; and (2) the creation of a specialized railway low-light dataset to support comprehensive model training.

This paper is organized as follows: Sect. 2 reviews relevant literature in railway surveillance and low-light image analysis. Section 3 details our proposed methodology, followed by experimental validation in Sect. 4. Finally, Sect. 5 concludes with key findings and future research directions.

2 Related Work

Low-light object detection is a challenging yet vital task in computer vision, with applications in night surveillance, autonomous driving, and security systems. Poor lighting conditions degrade image quality, causing blurred edges and loss of color information, which reduces detection accuracy.

Traditional methods rely on image enhancement techniques like histogram equalization and adaptive filtering [5], but these often distort original features or introduce noise, harming detection performance [6].

Deep learning-based approaches, such as Faster R-CNN [7], SSD [8], and YOLO [9], have improved detection accuracy by learning robust image features. However, their performance remains limited in low-light scenarios due to noise and blur.

Recent advances leverage attention mechanisms to enhance CNNs by mimicking human visual focus. For instance, WSA-YOLO [10] boosts YOLOv7's low-light detection through adaptive enhancement, while AMFLW-YOLO [11] employs coordinate attention to improve multi-scale target detection in remote sensing.

Despite progress, challenges remain, such as optimizing attention mechanisms for low-light conditions, reducing computational costs, and integrating multi-modal data (e.g., infrared, radar) [12]. Future work will address these issues to advance low-light object detection.

3 Method

3.1 Triple Attention Mechanism

The Triple Attention mechanism captures cross-dimension interactions between channel (C), height (H), and width (W) dimensions of an input tensor $\chi \in \mathbb{R}^{C \times H \times W}$. It consists of three branches that compute attention weights through rotation operations and residual transformations [13]. The input tensor is represented as:

$$\chi \in \mathbb{R}^{C \times H \times W} \quad (1)$$

The Channel-Height Attention mechanism integrates channel and height features through width-dimension rotation of the input feature map. Under low-light conditions, vertically oriented dark objects (e.g., nighttime pedestrians, streetlights) are particularly susceptible to background noise interference. This mechanism enhances vertical attention to effectively improve feature representation for linear structures like human silhouettes and signal poles, while capturing illumination gradient variations to prevent suppression of dark-region features.

For horizontal structural features (e.g., distant vehicle lights, railway track extensions, lane markings), the Channel-Width Attention mechanism employs height-dimension rotation to assign attention weights to channel combinations across horizontal structures. This approach demonstrates superior capability in extracting weakly illuminated horizontal features by enhancing both contour perception and feature consistency within horizontal bands, thereby significantly improving edge target detection performance. Rotation to reorient feature dimensions:

$$\hat{\chi}_i = \mathbb{R}_i(\chi), \quad i \in \{1, 2, 3\} \quad (2)$$

In low-light images, the contrast is poor, and the differences between the background and the target are not significant. Local noise is prone to causing false detections (such as watermarks and noise points), and uneven lighting can lead to partial loss of the target. Spatial attention can adaptively focus on local regions (such as high-response objects), suppress irrelevant backgrounds, integrate contextual spatial information, and enhance the recognition ability of incomplete targets. It can effectively focus on the “real target areas” and suppress background pseudo-activation areas. The Triple Attention is illustrated in Fig. 1 and in Algorithm 1.

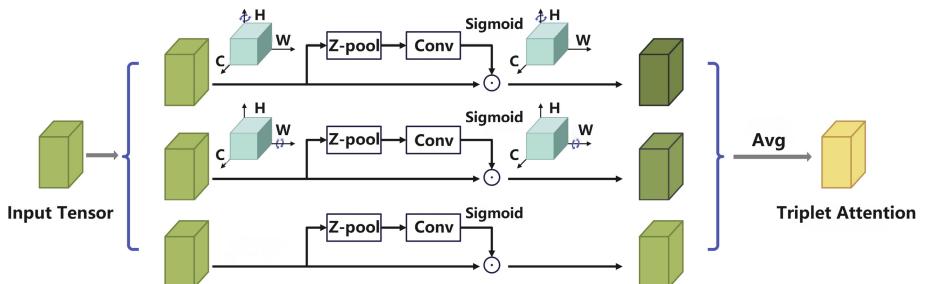


Fig. 1. Triple attention mechanism.

To compute attention weights for each directional branch, the rotated feature map is processed by a Z-pooling operation, which concatenates the max-pooled and average-pooled features along the channel dimension. Such a combination enhances robustness to noise and weak signals, reducing the risk of information loss that might occur if only a single pooling operation were used:

$$\hat{\chi}_i^* = \text{Z-pool}(\hat{\chi}_i) = \left[\max_c(\hat{\chi}_i), \text{avg}_c(\hat{\chi}_i) \right], \quad i \in \{1, 2, 3\} \quad (3)$$

A convolutional layer $\psi_i \in \mathbb{R}^{k \times k}$ is then applied to extract local spatial context, followed by a sigmoid activation function to generate the attention map:

$$\omega_i = \sigma(\psi_i(\hat{\chi}_i^*)), \quad \omega_i \in \mathbb{R}^{1 \times \cdot \times \cdot} \quad (4)$$

The attention map is then element-wise multiplied with the rotated feature map and subsequently rotated back to the original orientation:

$$\bar{y}_i = \mathbb{R}_i^{-1}(\hat{\chi}_i \odot \omega_i), \quad \bar{y}_i \in \mathbb{R}^{C \times H \times W} \quad (5)$$

The final output of the Triplet Attention module is obtained by averaging the outputs from all three branches:

$$y = \frac{1}{3} \sum_{i=1}^3 \bar{y}_i \quad (6)$$

This fusion mechanism effectively integrates cross-dimensional attention information, enabling the network to enhance weak feature representation, suppress noisy backgrounds, and improve target recognition under low-light conditions.

Algorithm 1 Triple Attention Mechanism (Rotation-based)

Require: Input tensor $\chi \in \mathbb{R}^{C \times H \times W}$
Ensure: Output tensor $y \in \mathbb{R}^{C \times H \times W}$

- 1: **Branch 1: Channel-Width Interaction**
- 2: $\hat{\chi}_1 \leftarrow \text{Rotate}(\chi, 90^\circ \text{ along } H)$
- 3: $\hat{\chi}_1^* \leftarrow \text{Z-pool}(\hat{\chi}_1) \{ \in \mathbb{R}^{2 \times H \times C} \}$
- 4: $\omega_1 \leftarrow \sigma(\psi_1(\hat{\chi}_1^*)) \{ \in \mathbb{R}^{1 \times H \times C} \}$
- 5: $\bar{y}_1 \leftarrow \text{RotateBack}(\hat{\chi}_1 \odot \omega_1)$
- 6: **Branch 2: Channel-Height Interaction**
- 7: $\hat{\chi}_2 \leftarrow \text{Rotate}(\chi, 90^\circ \text{ along } W)$
- 8: $\hat{\chi}_2^* \leftarrow \text{Z-pool}(\hat{\chi}_2) \{ \in \mathbb{R}^{2 \times C \times W} \}$
- 9: $\omega_2 \leftarrow \sigma(\psi_2(\hat{\chi}_2^*)) \{ \in \mathbb{R}^{1 \times C \times W} \}$
- 10: $\bar{y}_2 \leftarrow \text{RotateBack}(\hat{\chi}_2 \odot \omega_2)$
- 11: **Branch 3: Spatial (H-W) Interaction**
- 12: $\hat{\chi}_3 \leftarrow \text{Z-pool}(\chi) \{ \in \mathbb{R}^{2 \times H \times W} \}$
- 13: $\omega_3 \leftarrow \sigma(\psi_3(\hat{\chi}_3)) \{ \in \mathbb{R}^{1 \times H \times W} \}$
- 14: $y_3 \leftarrow \chi \odot \omega_3$
- 15: **Aggregate Outputs**
- 16: $y \leftarrow \frac{1}{3} (\bar{y}_1 + \bar{y}_2 + y_3)$
- 17: **return** y

3.2 Integration of Triple Attention Mechanism with YOLOv8

The Triple Attention mechanism is inserted at the end of the backbone network, specifically after the last feature extraction module (C2f). This placement is chosen because (1) the feature maps at this stage contain rich semantic information after multiple convolution and fusion layers, allowing the attention module to further enhance important features; (2) the output of the Triple Attention module directly connects to the Spatial Pyramid Pooling - Fast (SPPF) module, enabling better multi-scale feature fusion and improving detection performance for objects of various sizes [14]. In YOLOv8s backbone, the Triple Attention module is inserted at the final stage (P5/32 layer), right after the SPPF module. It takes the SPPF output, applies the attention mechanism, and outputs enhanced features for the detection head. This setup is referred to as T-attention1 (T-1), as shown in Fig. 2.

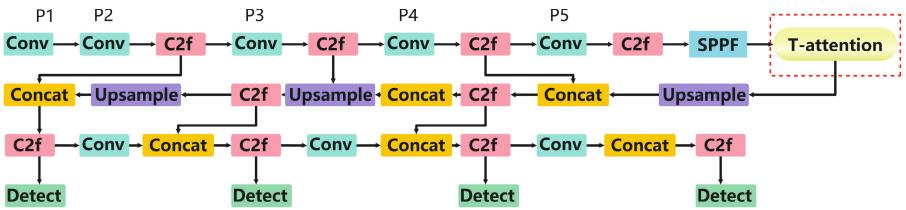


Fig. 2. Triple attention mechanism in the backbone of YOLOv8.

4 Experiments

To validate the effectiveness of the Triplet Attention mechanism, we designed the following experiments: A comparative study combining four commonly used detection models with four mainstream attention mechanisms to analyze their performance differences. We conducted these experiments by stacking 1 to 5 attention modules, aiming to demonstrate that the three-layer (triplet) attention structure achieves the best performance. In the ablation study, each attention mechanism was individually removed to evaluate its specific contribution within the overall system. Additionally, we assessed the models real-time detection capability.

4.1 Dataset

To construct a high-quality railway monitoring dataset, we collected a large number of images from surveillance cameras at multiple railway stations, focusing on low-light conditions from dusk to late night. By capturing images under various times and weather conditions, we observed the impact of lighting and

environmental changes on image quality, providing rich samples for model training. Ultimately, we gathered a dataset containing thousands of images, covering diverse scenes, object types, and variations in lighting and environmental conditions. The dataset includes point shots captured under various situations.

To better align with railway monitoring applications, we enhanced our self-collected dataset by incorporating the EXDark dataset [15], which is designed for low-light object detection and image enhancement. The expanded dataset, as shown in Fig. 3, better simulates real railway monitoring environments, providing diverse and comprehensive samples for model training. The combined dataset used for training and testing our model includes 7,864 images for the training set, 1,680 images for the validation set, and 1,663 images for the test set. Our dataset comprises 15 recognizable object categories.

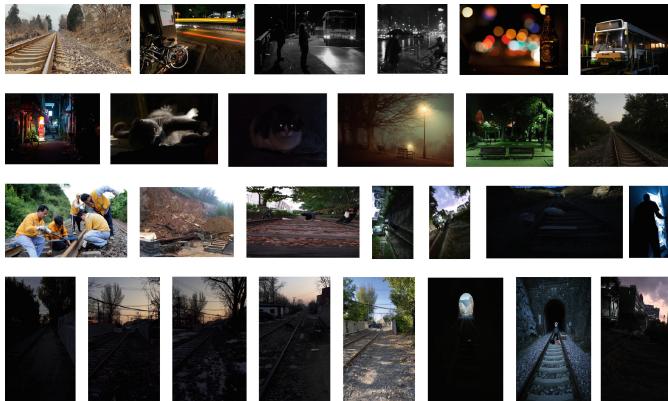


Fig. 3. Dark light image set.

4.2 Performance Comparison of Attention Mechanisms

This study presents a comprehensive evaluation of attention mechanisms in object detection systems, employing YOLOv8, YOLOv5, and Faster R-CNN as baseline architectures while investigating four prominent attention modules: Squeeze-and-Excitation (SE) attention that performs channel-wise feature recalibration through adaptive channel importance weighting; Coordinate Attention (CA) that effectively encodes both spatial coordinate information and inter-channel dependencies for enhanced feature representation; Convolutional Block Attention Module (CBAM) that sequentially combines channel and spatial attention mechanisms for dual-stage feature refinement; and Global Attention Mechanism (GAM) that establishes long-range dependencies through dimensionality reduction and feature reconstruction.

The evaluation framework incorporates four key metrics: bounding box precision (Box P) for localization accuracy, recall (R) for detection completeness,

mean Average Precision at IoU threshold 0.5 (mAP50), and mean Average Precision across IoU thresholds from 0.5 to 0.95 (mAP50-95) for comprehensive performance assessment. All models were trained for 100 epochs under identical experimental configurations to ensure fair comparison, with mAP50-95 serving as the primary indicator of detection robustness across varying localization requirements.

Table 1 details the comparative analysis. Our T-attention mechanism notably enhances detection performance, particularly in the mAP50-95 metric, which is indicative of improved localization accuracy under more stringent conditions. For YOLOv8, T-attention achieves an mAP50 of 0.713, surpassing other mechanisms like CBAM (0.685), SE (0.686), and GAM (0.698). This pattern is consistent across different models, validating the T-attention mechanism's effectiveness in boosting detection accuracy.

Table 1. Performance metrics of different models with various attention mechanisms

Model	Attention	Box (P)	R	mAP50	mAP50-95
YOLOv8	Baseline	0.715	0.637	0.686	0.403
	T-attention	0.767	0.659	0.713	0.421
	SE	0.726	0.613	0.686	0.391
	CA	0.722	0.632	0.678	0.397
	CBAM	0.765	0.611	0.685	0.401
	GAM	0.792	0.596	0.698	0.410
YOLOv5	Baseline	0.688	0.519	0.581	0.319
	T-attention	0.682	0.525	0.613	0.342
	SE	0.689	0.527	0.587	0.320
	CA	0.667	0.529	0.586	0.323
	CBAM	0.678	0.515	0.585	0.323
	GAM	0.695	0.532	0.590	0.330
Faster R-CNN	Baseline	0.633	0.478	0.536	0.294
	T-attention	0.628	0.483	0.566	0.315
	SE	0.634	0.485	0.542	0.295
	CA	0.615	0.488	0.541	0.298
	CBAM	0.625	0.474	0.540	0.298
	GAM	0.639	0.49	0.552	0.299

4.3 Individual Branch Contributions

To validate the advantages of T-attention, we conducted two experiments: (1) verifying the optimal branch configuration and (2) studying individual branch contributions.

Branch Configuration Analysis: Performance improves from single-branch to triple-branch design. A single H-C branch yields 0.584 mAP50, while adding a W-C branch boosts it to 0.682. The full T-attention (0.713 mAP50) integrates Channel-Height, Channel-Width, and Spatial branches. Beyond three branches, performance drops (quadruple: 0.688, quintuple: 0.685), confirming the three-branch design maximizes benefits without redundancy.

Individual Branch Contributions: The ablation study shows distinct values of each branch: W-C branch achieves 0.632 mAP50 by capturing horizontal features, H-C branch focuses on vertical structures (0.584), and spatial-only attention reaches 0.698 mAP50 through local feature enhancement. The complete T-attention module (0.713 mAP50) demonstrates effective synergy among these branches.

Table 2 shows that T-attention’s gains come from: (1) a balanced three-branch architecture avoiding redundancy, and (2) integrating complementary mechanisms capturing horizontal, vertical, and spatial features. The spatial branch contributes the most individually, while channel-directional branches provide specialized enhancements to optimize feature representation.

Table 2. Individual branch contributions

Configuration	Box (P)	R	mAP50	mAP50-95
Single (W-C)	0.698	0.565	0.632	0.352
Single (H-C)	0.683	0.535	0.584	0.333
Spatial only	0.777	0.616	0.698	0.406
Double (H-C + W-C)	0.777	0.616	0.682	0.406
T-attention	0.767	0.659	0.713	0.421
Quadruple	0.789	0.610	0.688	0.408
Quintuple	0.790	0.611	0.685	0.406

4.4 Insertion Position Optimization

As presented in Table 3, the performance of the T-attention module integrated into YOLOv8 is evaluated under two distinct placements: (1) after the SPPF layer of the backbone (T_attention1) and (2) after the detection heads (T_attention2). The experimental results demonstrate that T_attention1 achieves superior overall performance, attaining an mAP50 of 0.713. This improvement suggests that inserting the attention mechanism after the SPPF layer enhances feature refinement prior to detection. In contrast, T_attention2 exhibits higher precision but suffers from reduced recall and mAP50, indicating that its placement after the detection heads provides limited global feature enhancement.

Table 3. Optimal insertion position

Model	Box (P)	R	mAP50	mAP50-95
YOLOv8	0.715	0.637	0.686	0.403
YOLOv8 + T_attention1 (after SPPF)	0.767	0.659	0.713	0.421
YOLOv8 + T_attention2 (after heads)	0.792	0.622	0.698	0.410

4.5 Low-Light Detection Performance

To evaluate the robustness of our model in low-light conditions, we conducted extensive experiments on detecting diverse objects including vehicles (both nearby and distant), pedestrians, and static objects like boxes and umbrellas. These scenarios represent common challenges for vision systems in real-world applications such as nighttime surveillance and autonomous driving.

As quantitatively demonstrated in Table 4, our model achieves superior detection accuracy compared to state-of-the-art baselines. The performance gap is particularly significant for challenging cases: dark vehicle detection (0.90 mAP vs. YOLOv5’s 0.47 and YOLOv8’s 0.72) and distant vehicle identification (0.82 mAP). Qualitative results in Fig. 4 visually confirm these findings, showing reliable detection across varying illumination levels and object scales.

The system maintains practical efficiency with 95–110 FPS (7–10 ms latency) on an NVIDIA RTX 4090 GPU, demonstrating both real-time capability and

Table 4. Detection performance comparison in challenging scenarios

Model	Dark vehicle	Distant vehicle	Person	Box	Umbrella
YOLOv5	0.47	—	0.34	—	—
YOLOv8	0.72	—	0.54	—	—
Our	0.90	0.82	0.70	0.90	0.73

**Fig. 4.** Detection examples under low-light conditions.

robustness without frame drops critical requirements for production environments.

5 Conclusion

This paper presents an improved YOLOv8 model for low-light railway intrusion detection, incorporating a triple attention mechanism. By combining channel, spatial, and temporal attention, the model enhances feature representation, focuses on critical details, suppresses noise, and adapts to challenging lighting conditions. A dataset is established including self-collected low-light railway images and classes from the EXDark dataset, covering diverse scenes and conditions to provide a solid foundation for model training. The experimental results show that the model outperforms the baseline YOLOv8 and other attention mechanisms in respect to precision, recall, and mean average precision.

References

1. Cao, Z., et al.: Railway intrusion detection based on machine vision: a survey, challenges, and perspectives. *IEEE Trans. Intell. Transp. Syst.* **25**(7), 6427–6448 (2024). <https://doi.org/10.1109/TITS.2024.3412170>
2. Cao, Z., Qin, Y., Xie, Z., Liu, Q., Zhang, E., Zhiyu, W., Zujun, Y.: An effective railway intrusion detection method using dynamic intrusion region and lightweight neural network. *Measurement* **191**, 110564 (2022). <https://doi.org/10.1016/j.measurement.2021.110564>. ISSN 0263-2241
3. Xu, L., Philip Chen, C.L.: Comparison and combination of activation functions in broad learning system. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 3537–3542. Toronto, ON, Canada (2020). <https://doi.org/10.1109/SMC42975.2020.9282871>
4. Huang, H., Zhao, G., Bo, Y., Junyan, Yu., Liang, L.D., Yang, Y., Kai, O.: Railway intrusion detection based on refined spatial and temporal features for UAV surveillance scene. *Measurement* **211**, 112602 (2023). <https://doi.org/10.1016/j.measurement.2023.112602>. ISSN 0263-2241
5. Zhang, Z., Zhang, W., Zhang, C., Yang, S.: Exploring image enhancement for salient object detection in low light images. arXiv preprint [arXiv:2007.16124](https://arxiv.org/abs/2007.16124) (2020). <https://arxiv.org/abs/2007.16124>
6. Liu, X., Wang, Y., Liu, M.: Exploring low-light object detection techniques. arXiv preprint [arXiv:2107.14382](https://arxiv.org/abs/2107.14382) (2021). <https://arxiv.org/abs/2107.14382>
7. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS 2015), pp. 91–99. Montreal, Canada (2015). <https://arxiv.org/abs/1506.01497>
8. Liu, W., Anguelov, D., Erhan, D., et al.: SSD: single shot multibox detector. In: European Conference on Computer Vision (ECCV), pp. 21–37. Amsterdam, Netherlands (2016). <https://arxiv.org/abs/1512.02325>
9. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), pp. 779–788. Las Vegas, NV, USA (2016). <https://arxiv.org/abs/1506.02640>

10. Hui, Y., Wang, J., Li, B.: WSA-YOLO: weak-supervised and adaptive object detection in the low-light environment for YOLOv7. *IEEE Trans. Instrum. Meas.* **73**(99), 1–12 (2024). <https://doi.org/10.1109/TIM.2024.3350120>
11. Peng, J., Zhang, Y., Wang, S.: AMFLW-YOLO: a lightweight network for remote sensing image detection based on attention mechanism and multiscale feature fusion. *IEEE Trans. Geosci. Remote Sens.* **61**, 1–13 (2023)
12. Zhang, Y., et al.: A systematic literature review on object detection using near infrared and thermal images. *Neurocomputing* **518**, 250–262. <https://www.sciencedirect.com/science/article/abs/pii/S092523122300927X> (2023)
13. Liu, X., Wang, Y., Liu, M.: Rotate to attend: convolutional triplet attention module. arXiv preprint [arXiv:2010.03045](https://arxiv.org/abs/2010.03045) (2020). <https://arxiv.org/abs/2010.03045>
14. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI)* **37**(9), 1904–1916 (2015). <https://doi.org/10.1109/TPAMI.2015.2380471>
15. Loh, Y.H., et al.: Getting to know low-light images with the exclusively dark dataset. In: *Computer Vision and Image Understanding* (2020). <https://doi.org/10.1016/j.cviu.2020.103045>



A Quantitative Calculation Scheme for Network Vulnerable Nodes Based on Dynamic Trust Evaluation

Jian Liang^(✉), Pin Lv, Yangming Kang, and Yaxing Zhang

Institute of Automation, Chinese Academy of Sciences, Beijing 100190,
People's Republic of China
jian.liang@ia.ac.cn

Abstract. With the continuous increase in the scale and complexity of the current network environment, many challenges faced by network security issues have gradually become prominent. How to accurately identify the vulnerable nodes in a complex network is a crucial topic for ensuring the security and stable operation of the entire network. This paper proposes a calculation scheme for the vulnerability of nodes in a complex network based on dynamic trust evaluation, and provides suggestions for the maintenance of complex networks according to this scheme. This scheme first dynamically evaluates the trust relationships between network nodes. Then, by considering factors such as the behavioral and identity characteristics of nodes, historical interaction information, and credit records, a comprehensive dynamic trust evaluation model is constructed. On this basis, vulnerability indicators for network nodes can be formulated to quantitatively calculate the vulnerability degree of all nodes in the network. Finally, an example is used to prove that this method can effectively identify the vulnerable nodes in the network. Compared with the traditional qualitative analysis method, it can analyze the characteristics of complex networks more quickly, accurately, and quantitatively, providing strong support for topics related to network security.

Keywords: Quantitative calculation · network vulnerable · trust evaluation · security computation

1 Introduction

With the rapid development of information technology, the network has become an indispensable infrastructure in modern society. From individuals' daily lives to enterprises' business operations and even to a country's key sectors such as finance, energy, and transportation, all highly rely on the network to achieve information transmission, resource sharing, and business collaboration. The scale of the network is constantly expanding, and the number of connected devices is growing exponentially. A variety of new network

This work is supported by the Strategic Priority Research Program of Chinese Academy of Sciences. Grant No. XDA27050201.

technologies and applications are emerging in an endless stream, such as the Internet of Things (IoT) [1], cloud computing [2], big data [3], and 5G communication [4]. The emergence of these new technologies has made the network more powerful in function and more extensive in application, but at the same time, it has also greatly increased the complexity and dynamics of the network.

Today's networks are facing various security threats from both inside and outside, such as malicious attacks, software vulnerabilities, and mis-operations [5]. Frequent malicious attacks: The openness and complexity of the network provide more opportunities for attackers. Common network attack methods include Distributed Denial-of-Service (DDoS) attacks [6], malware infections, man-in-the-middle attacks, and zero-day vulnerability exploitation. Frequent interactions among different types of devices, systems, and applications have increased the difficulty of security management. Meanwhile, the dynamic changes in network topology and the diversity of communication protocols have also provided more opportunities for attackers [7]. These threats may cause some nodes in the network to malfunction or be maliciously exploited, thereby affecting the normal operation of the entire network. Network vulnerable node vulnerability is an indicator to measure the degree of vulnerability of a node in the network [8].

Vulnerable nodes refer to those nodes in the network that are easily attacked or prone to malfunction, and once they fail, they may have a relatively large negative impact on the network. However, in the actual network environment, the states and behaviors of nodes are dynamically changing, and simply relying on static features is difficult to accurately reflect the degree of node vulnerability. In recent years, trust evaluation technology has been widely used in the field of network security [9]. By evaluating the trust relationships between nodes, a more comprehensive understanding of the credibility and reliability of nodes in the network can be obtained. Dynamic trust evaluation refers to the real-time and dynamic evaluation of the trust relationships between nodes during the network operation [10]. Different from traditional static trust evaluation, dynamic trust evaluation not only considers the initial trust value of nodes but also continuously updates the trust value as the nodes' behaviors change, time passes, and the network environment changes. Dynamic trust evaluation can continuously update trust levels based on real-time behaviors, historical performance, and interaction situations, making it more flexible and accurate than static evaluation and better adapting to the dynamic changes of networks. Common dynamic trust evaluation models usually comprehensively consider the following factors [11]:

1. **Node behavior characteristics:** including the communication frequency of nodes, the accuracy of data transmission, and response time. For example, a node that frequently sends error packets or has a long response delay may have its trust value reduced.
2. **Historical interaction information:** Records the past interaction history between nodes and other nodes, such as the number of successful or failed communications and the frequency of cooperation. A good historical interaction record helps to increase the trust value of a node.
3. **Network environment changes:** Factors such as changes in network topology, the addition of new nodes or the removal of old nodes, and fluctuations in network traffic will all affect the trust relationships between nodes. For example, when there are

significant changes in the network topology, the trust relationships between nodes may need to be re-evaluated.

Common dynamic trust evaluation models include models based on Bayesian networks, models based on fuzzy logic, and models based on machine learning. Taking the model based on Bayesian networks as an example [12], it calculates the trust degree of nodes through probabilistic reasoning. The dynamic trust evaluation method has the following three advantages in the evaluation of network vulnerable nodes [13]. For example, when the network topology changes or new security threats emerge, the dynamic trust evaluation method can quickly adapt and adjust the evaluation results [14, 15].

To this end, this paper selects the method of using dynamic trust degree for node evaluation. Then, based on the trust evaluation results of nodes, a comprehensive analysis of the topological structure of the entire network is carried out to quantitatively calculate the vulnerability of each node. This is used to analyze the vulnerability of the entire network and enhance the understanding of this complex network.

2 Construction of the Dynamic Trust Model

In today's complex and ever-changing network environment, the behaviors and states of nodes are constantly changing. Traditional static trust models struggle to adapt to this dynamic nature and cannot accurately reflect the credibility of nodes at different times. However, constructing a dynamic trust model for nodes in the network is of great significance. It can evaluate the trust status of nodes in real-time, promptly detect potential security threats, and enhance the security and reliability of the network. In this section, an approach of calculate dynamic trust value of nodes in network is proposed.

1. **Node behavior characteristics:** Communication behavior: This includes communication frequency, communication time distribution, packet size, and format. For example, if a node frequently sends a large number of abnormal packets in a short period, it may be engaging in malicious behavior, and its trust value should be correspondingly reduced. Response behavior: This refers to the response time and response accuracy of a node to requests. If a node often responds with a delay or gives incorrect responses, it indicates low reliability, and its trust value will be affected.
2. **Historical interaction information:** Interaction frequency: The more interactions a node has with other nodes, the richer the accumulated information, and the more accurate the trust evaluation of it will be. For example, if two nodes have had long-term and stable cooperation, their mutual trustworthiness will be relatively high. Interaction results: The success or failure of each interaction matters. Multiple successful interactions will increase a node's trust value, while frequent interaction failures will decrease it.
3. **Network environment factors:** Topological structure changes: Alterations in the network topological structure, such as the addition or removal of nodes and the interruption or restoration of links, will affect the trust relationships between nodes. For example, when a critical node exits the network, the trust evaluation of the nodes connected to it may need to be re-adjusted. Network traffic fluctuations: A sudden increase or decrease in network traffic may imply abnormal situations. If a node exhibits abnormal behavior during peak traffic periods, its trust value may decline.

The core idea of the PageRank algorithm in the zero-trust architecture involves two points. Firstly, if a node is linked to by many other nodes, it is considered more important, resulting in a higher PageRank value. Secondly, if a node is linked to by nodes with high PageRank values, its own PageRank value will also increase. Therefore, in the zero-trust architecture, the PageRank algorithm can be used not only for determining the importance of nodes but also for establishing trust relationships and managing trust during interactions. By analyzing and evaluating the link relationships between nodes, it becomes possible to identify nodes with higher trust and trustworthiness, thus establishing a trust-based communication network. The PageRank value can serve as an essential metric to quantify and rank the trustworthiness of nodes within the zero-trust architecture, enabling complex networks to make more informed decisions.

3 Calculation of Node Vulnerability in Complex Networks

To address the “Dead Ends” and “Spider Traps” problems in zero-trust environments, we propose two mechanisms: the Zero Trust Random Jump mechanism and the Zero Trust Damping Factor, denoted as d .

The importance of nodes in a zero-trust architecture can be defined by incorporating a smoothing term. Consider an arbitrary zero-trust directed graph with n nodes and assume the existence of a first-order zero-trust Markov chain based on a random wandering model with a transfer matrix M . This zero-trust Markov chain may not possess a smooth distribution. We also consider another completely random wandering zero-trust model with a transfer matrix where all elements are set to $1/n$, indicating an equal probability of transfer from any node to any other node. By linearly combining these two transfer matrices, we can form a new transfer matrix, resulting in a new zero-trust Markov chain that guarantees a smooth distribution. This smooth distribution satisfies the following relation:

$$\begin{aligned} P &= \left(dM + \frac{1-d}{n} E \right) P \\ &= dMP + \frac{1-d}{n} 1 \end{aligned} \quad (1)$$

Here, d ($0 \leq d \leq 1$) represents the coefficient, referred to as the zero-trust damping factor, E is the matrix consisting of all 1's, P is an n -dimensional vector, and 1 is an n -dimensional vector with all components equal to 1. P serves as a general representation of node importance in a zero-trust directed graph. Linear combination generates comprehensive indicators by weighted fusion of multi-dimensional information (such as node trustworthiness, path cost, historical access frequency, etc.), providing more flexible guidance for search.

In the equation, the first term represents the probability of visiting each node based on the transfer matrix M when the state distribution is smooth, and the second term represents the probability of visiting each node completely at random. When d is close to 1, the random tour is primarily conducted according to the transfer matrix M , while d approaching 0 indicates that the random tour mainly visits each node with equal probability at random.

The formula for calculating the importance of each node in a zero-trust architecture can be derived from the above equation as follows:

$$PR(v_i) = d \left(\sum_{v_j \in M(v_i)} \frac{PR(v_j)}{L(v_j)} \right) + \frac{1-d}{n}, i = 1, 2, \dots, n \quad (2)$$

where N denotes the total number of nodes in the zero-trust topological network, d represents the zero-trust damping factor (typically set as 0.85 in a zero-trust environment), $M(v_i)$ represents the set of nodes pointing to node v_i and $L(v_i)$ represents the out-degree of node v_i .

The second term in the equation is referred to as the smoothing term, ensuring that the importance of all nodes is non-zero. This term possesses the following properties:

$$PR(v_i) > 0, i = 1, 2, \dots, n \quad (3)$$

$$\sum_{i=1}^n PR(v_i) = 1 \quad (4)$$

In the context of a complex networks communication network under a zero-trust architecture, we can represent the network as a zero-trust directed graph, where the nodes correspond to the complex network communication topology nodes, and the directed edges represent the communication connections between the vehicle nodes. The algorithm follows the following computational procedure:

1. Initialization: Assign each node an initial importance value, which can be uniformly set to an initial value.
2. Iterative computation: Update the importance of nodes through iterations until convergence is achieved. In each iteration, the importance of nodes under the zero-trust architecture is updated based on the general formula for importance calculation.
3. Convergence check: Monitor the changes in node importance values. When the importance values of nodes no longer change significantly, the algorithm is considered to have converged, and the computation can be stopped.

4 The Example

In this section, an example is given to verify the theory presented in this article. This example give the difference of the influence of common point and key node. From the result, it is obvious that the key node is more important than other nodes, and it could make larger influence to the whole network.

In the Fig. 1, there exist a network with 100 nodes, the connected index is 0.8, for the first, there occurs a malicious cyber-attack in a random node. Then the malicious cyber-attack is transmitted through the lines to the connected nodes of the first attacked node. The index β shows the proportion of abnormal nodes of all nodes. The probability of harmful transmission is 0.1. In the Fig. 2. In subfigure (c), the harmful information first spread from Node 88th to the critical nodes. Therefore, after the second spread of the harmful information, the total number of affected nodes reached 17. In contrast, in subfigure (d), the harmful information from node 88th only spread among ordinary nodes.

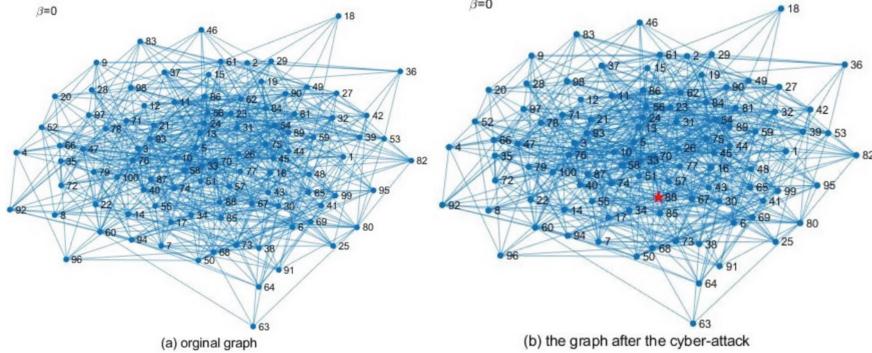


Fig. 1. The complex network. (In the subfigure (a), it is the original graph; the subfigure (b) shows there is a cyber attack to a random node.)

So, after the second spread of the harmful information, only 5 nodes were ultimately affected. By comparing the two figures, the impact of critical nodes and ordinary nodes on the security of the entire network can be clearly observed.

Because of the index β shows the proportion of abnormal nodes of all nodes. It is obvious that the values of β in subfigure (c) is larger than that in subfigure (d).

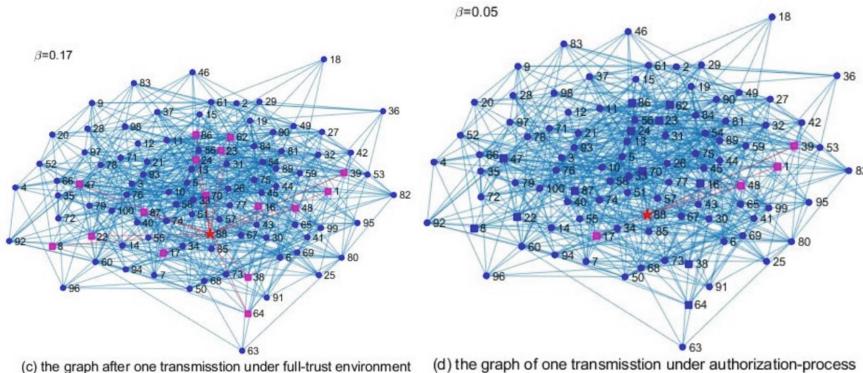


Fig. 2. The different influence of common node and key node.

After conducting 20 simulation experiments on random topologies with 100 nodes each, it was found that under the full-trust environment, the average value of β is 0.18, while under the zero-trust environment, the average value of β is 0.05.

5 The Conclusion

This paper presents a methodology for computing the vulnerability of network nodes grounded in dynamic trust evaluation. By establishing a dynamic trust model, the proposed approach comprehensively incorporates factors such as the communication behaviors of nodes, historical interaction data, and alterations in the network environment to

conduct real-time assessments of the trust relationships among nodes. Subsequently, it calculates the vulnerability indices of the nodes. The experimental findings demonstrate that this method exhibits a high degree of accuracy and adaptability in pinpointing vulnerable network nodes. It can effectively address the dynamic changes within the network environment and offers a robust means for network security management and optimization. In future research endeavors, it is advisable to further optimize the parameter settings and computational algorithms of the dynamic trust evaluation model to enhance computational efficiency. Additionally, consideration should be given to applying this method in actual large-scale network environments to validate its efficacy and feasibility in real-world scenarios.

References

1. Zhu, B., Wang, X.: Hypergraph-aided task-resource matching for maximizing value of task completion in collaborative IoT systems. *IEEE Trans. Mob. Comput.* **23**, 12247–12261 (2024)
2. Zhang, Z., Zhou, F., Wang, Y.: Multi-user privacy-preserving image search in distributed cloud computing. *Comput. Electr. Eng.* **118** (2024)
3. Al-Jumaili, A.H.A., et al.: Big data analytics using cloud computing based frameworks for power management systems: status, constraints, and future recommendations. *Sensors.* **23**(6) (2023)
4. Mao, C., et al.: Deep learning in physical layer communications: evolution and prospects in 5G and 6G networks. *IET Commun.* **17**(16) (2023)
5. Maene, P., Götzfried, J., De Clercq, R., Müller, T., Freiling, F., Verbauwheide, I.: Hardware-based trusted computing architectures for isolation and attestation. *IEEE Trans. Comput.* **67**(3), 361–374 (2017)
6. Kim, Y., Hakak, S., Ghorbani, A.: Detecting distributed denial-of-service (DDoS) attacks that generate false authentications on electric vehicle (EV) charging infrastructure. *Comput. Secur.* **144**, 20 (2024)
7. Wang, Z., et al.: Graph convolutional networks and deep reinforcement learning for intelligent edge routing in IoT environment. *Comput. Commun.* **232**, 108050 (2025)
8. Amelina, N., et al.: Approximate consensus in stochastic networks with application to load balancing. *IEEE Trans. Inf. Theory.* **61**(4), 1739–1752 (2015)
9. Lin, P., et al.: Dynamic default mode network across different brain states. *Sci. Rep.* **7**, 46088 (2017)
10. Tomás Olivier, et al.: Rule-mediated connectivity in social-ecological-technological systems: a comparative network analysis of reservoir operation rules in Coyote Valley dam (United States) and Ameghino dam (Argentina). *J. Environ. Manag.* **374** (2025)
11. Liu, C., et al.: An adaptive trust evaluation model for detecting abnormal nodes in underwater acoustic sensor networks. *Sensors.* **24**(9), 16 (2024)
12. Huang Ping, Corman Francesco: Probabilistic modeling of train operations for uncertainty quantification: a context-aware Bayesian network approach. *IEEE Trans. Intell. Transp. Syst.* (2024)
13. Wan, Y., Wen, G., Yu, X., Kurths, J.: Distributed event-based resilient secondary control for AC microgrids: a trust-reputation approach. *IEEE Trans. Smart Grid.* **15**(2), 2116–2128 (2024)
14. Jiang, R., et al.: An electronic medical record access control model based on intuitionistic fuzzy trust. *Inf. Sci.* **658**(000), 19 (2024)
15. Wang, Z., Yu, X., Xue, Y.J.L.: Research on medical security system based on zero trust. *Sensors.* **23**(7) (2023)



Big-Five Personality Trait Analysis in Videos via DLF-Based Multi-Scale Spatiotemporal Modeling

ZhengHua Guo^{1,2}, Tao Ning^{1,2(✉)}, Jin Zhang^{1,2}, QiDong Hou^{1,2}, and JingSong Chen^{1,2}

¹ College of Computer Science and Technology, Dalian Minzu University, Dalian 116000, China

daliannt@126.com

² State Ethnic Affairs Commission Key Laboratory of Big Data Applied Technology, Dalian Minzu University, Dalian 116000, China

Abstract. This study proposes the DLF Multi-Scale Network (DLF) for recognizing Big Five personality traits from multimodal (visual/audio) video features, addressing limitations in feature stability and cross-modal correlation inherent in traditional methods. The model integrates a Discrete Wavelet Transform-CNN (DWTC) module that decomposes video frames into multi-scale components (e.g., facial expressions, body movements) to extract robust visual representations, alongside an MLP for processing audio spectral data; a Learnable Multi-head Self-Attention (LMSA) mechanism then aligns cross-modal features to resolve spatiotemporal inconsistencies before decision-level fusion. Evaluated on the First Impressions V1 dataset, DLF-MSN achieves 0.9177 average accuracy, outperforming baselines. This work advances behavior-personality mapping with applications in mental health assessment and personalized recommendation systems.

Keywords: Multimodal fusion · Big Five personality · Deep learning · Behavioral analysis

1 Summary

Short video platforms have emerged as a globally significant social medium (with over 3 billion monthly active users and 95 min of daily average usage), generating massive user behavior data that provides a new, ecologically valid source for psychological research. Traditional personality assessment methods (e.g., the Big Five model) are limited by subjective reporting biases, while AI technologies can more objectively capture personality traits by analyzing multimodal data from videos, including micro-expressions and vocal characteristics. Theoretically, this study deepens our understanding of personality expression mechanisms in digital environments. Practically, it can be applied to mental health screening, personalized recommendation optimization, and the development of digital therapeutics to enhance the accuracy of remote psychological assessments.

2 Related Work

2.1 Research Status of Deep Learning-Based Personality Trait Recognition

The interdisciplinary integration of Computer Science and Psychology initially emerged through audio-based personality trait recognition research. Mohammad et al. [1] proposed an APP method using prosodic features, achieving high prediction accuracy particularly for Extraversion and Conscientiousness. Su et al. [2] further advanced this approach by employing wavelet transforms to decompose speech signals followed by CNN classification. While audio-based methods demonstrated promise, visual modality research progressed concurrently. Biel et al. [3] analyzed video sequences frame-by-frame using SVM classifiers, identifying Extraversion as the most saliently recognized trait from visual cues. Expanding on this, Dhall et al. [4] combined manually extracted multi-level features (including facial action units) with kernel partial least squares regression (KPLSR), followed by CNN modeling to improve trait recognition from facial images.

Most computational personality recognition studies have historically relied on unimodal data (e.g., audio-only or visual-only approaches). A significant advancement occurred in 2016 when the ChaLearn First Impressions challenge established a benchmark for multimodal video-based personality analysis. This shift prompted several innovative fusion methodologies: Zhang et al. [5] proposed a Dual-Branch Residual (DBR) framework that effectively combined audiovisual features through late fusion. Building on this, Panagiotis et al. [6] developed a more comprehensive system integrating three specialized modules: (1) CNNs for spectral speech features, (2) ResNet-50 for visual processing, and (3) LSTMs for modeling emotional context. The most recent advancement by Zhao et al. [7] introduced a hybrid framework leveraging pretrained VGGish (audio) and VGG-Face (visual) models for feature extraction, augmented with Bi-LSTMs and Transformers to capture long-term temporal dependencies, followed by decision-level fusion with linear regression for final trait prediction. Despite these methodological innovations, current approaches face two critical limitations: First, they often fail to adequately capture subtle but psychologically meaningful motion details in dynamic scenes (e.g., micro-gestures or rapid facial expressions). Second, their attention mechanisms remain oversimplified, struggling to model the complex, nonlinear temporal relationships between multimodal features across extended sequences. These shortcomings ultimately lead to suboptimal information integration, constraining prediction accuracy.

Therefore, the contributions of this paper will be summarised in the following directions:

In video key-frame feature extraction, traditional methods are prone to losing structural information when merging small-scale feature maps. An improved approach combines multi-scale wavelet fusion (Discrete Wavelet Transform, DWT) with Convolutional Neural Networks (CNNs) to separately process high-frequency and low-frequency features and fuse them back to the original scale, enabling more detailed capture of local and global image details.

After extracting feature sequences, temporal order and content relevance between features of different frames pose challenges for traditional methods. To address this limitation, the improved Learnable Multi-head Self-Attention (LMSA) leverages attention mechanisms to analyze inter-feature dependencies across time steps, thereby enhancing the integration of sequential information.

The paper is structured as follows: Chap. 3 introduces the proposed DLF multi-scale network; Chap. 4 presents experimental configurations; Chap. 5 visualizes results; Chap. 6 concludes the work;

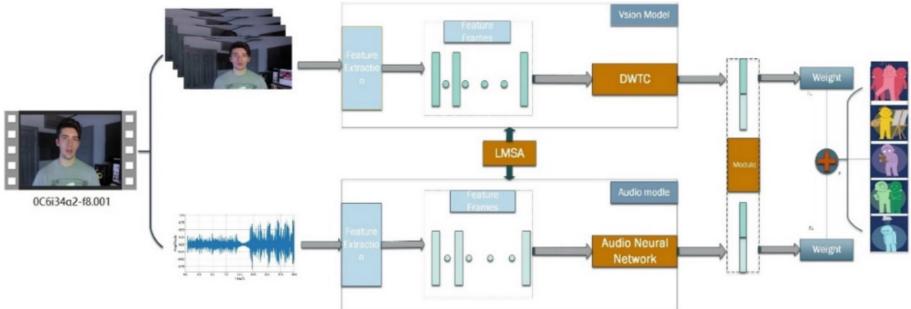


Fig. 1. Framework diagram of the DLF multi-scale, multi-modal personality recognition model

To achieve multi-scale recognition of personality traits in videos, this paper proposes a DLF multi-scale network for multimodal personality recognition in videos, processing dual modalities (audio signals and visual frames) to capture complementary features across scales. As illustrated in Fig. 1, the model integrates audio-visual data through a cross-modal fusion strategy, enabling comprehensive character analysis and accurate trait recognition by synthesizing multi-perspective information.

3 DLF Multi-Scale Network Framework

3.1 Visual Feature Extraction at DWTC

The Discrete Wavelet Transform (DWT) [8] effectively decomposes images into multi-scale features, capturing intricate details and structural information, which is particularly advantageous for processing complex textures and edge-rich images. When integrated with Convolutional Neural Networks (CNNs), DWT facilitates the extraction of higher-level semantic information from these multi-scale features, thereby enhancing model performance and robustness. The operational workflow involves initial image size reduction via average pooling followed by upsampling, concurrent extraction and fusion of global and local features, channel-wise concatenation of the fused features with the upsampled map, further feature extraction and transformation through a convolutional layer, enhancement of significant channel features via a channel attention module, and final feature representation generation using pooling and fully connected layers, as depicted in Fig. 2.

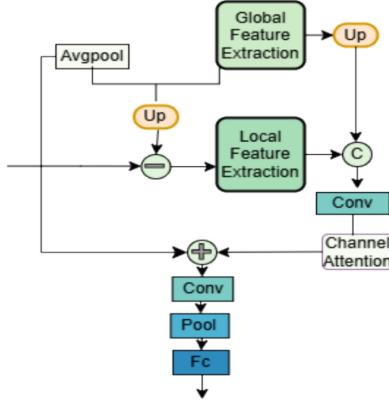


Fig. 2. DWTC model diagram

The Discrete Wavelet Transform (DWT) decomposes an image I into multi-scale features to capture detail and structural information, expressed as:

$$I = \sum_j \sum_k c_{j,k} \psi_{j,k} + \sum_j \sum_k d_{j,k} \phi_{j,k} \quad (1)$$

Where, $c_{j,k}$ and $d_{j,k}$ represent the low-frequency and high-frequency coefficients respectively, $\psi_{j,k}$ and $\phi_{j,k}$ denote the wavelet basis functions; the low-frequency and high-frequency features are concatenated along the channel dimension and restored to the original image size via the Inverse Discrete Wavelet Transform (IDWT).

$$I' = \sum_j \sum_k c_{j,k} \psi_{j,k} + \sum_j \sum_k d_{j,k} \phi_{j,k} \quad (2)$$

CNN extract high-level semantic information from images through convolutional layers and pooling layers. The convolution operation is defined as:

$$F = W * I + b \quad (3)$$

Where F denotes the feature map after convolution, W represents the convolution kernel, I is the input image, and b indicates the bias term.

3.2 LMSA Feature Sequence Alignment

LMSA enhances model performance in processing sequences with temporal and contextual dependencies by capturing temporal dependencies. The workflow is as follows: the input embeddings are transformed into queries, keys, and values through three linear layers, then partitioned into multiple heads for independent attention computation. Within each head, attention scores are computed via the dot product of queries and keys; a local window mask is applied to restrict the attention scope, followed by softmax normalization. The normalized scores weight and sum the values to produce outputs. The outputs from all heads are concatenated and passed through a final linear layer, yielding aligned feature sequences.

3.3 Auditory Feature Extraction

Auditory feature extraction transforms input audio into compact yet discriminative representations. A three-layer Multilayer Perceptron (MLP) processes the 26-dimensional input audio features: the first fully-connected (FC) layer maps features to 512 dimensions, followed by Batch Normalization and ReLU activation; the second FC layer reduces dimensionality to 256, applying Batch Normalization, ReLU activation, and Dropout for regularization; the third FC layer outputs a 128-dimensional feature vector ($h_3 \in \mathbb{R}^{128}$). This refined audio representation enhances multimodal fusion for subsequent model integration.

3.4 Integration Method

Decision fusion enhances predictive performance by integrating video and audio information. Video and audio data are processed through independent visual and audio models to extract features:

$$F_v = \text{VisionModel}(V; \theta_v) \quad (4)$$

$$F_a = \text{AudioModel}(A; \theta_a) \quad (5)$$

Subsequently, the video features F_v and audio features F_a are concatenated along the feature dimension to form an extended vector F. The concatenated vector F is processed through one or more neural network layers comprising linear transformation, ReLU activation, and Dropout regularization, ultimately generating predictions via a sigmoid function:

$$F' = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot F + b_1) + b_2) \quad (6)$$

This fusion approach enhances prediction accuracy and robustness by integrating information from diverse modalities and leveraging their complementarity, thereby providing greater flexibility for model design and optimization.

4 Experimental Setup

This section introduces relevant datasets, categorizes baselines, and elaborates on implementation details including feature preprocessing, parameter configuration, and evaluation metrics.

4.1 Dataset

This study employs the First Impressions V1 dataset [9], comprising 10,000 video clips (~15 s each) sourced from 3,000 diverse YouTube HD videos featuring English-speaking subjects facing the camera. Partitioned into training/validation/test sets (3:1:1 ratio), the dataset spans varied demographics (gender, age, nationality, ethnicity). Personality traits (Five-Factor Model: Openness, Agreeableness, Extraversion, Neuroticism, Conscientiousness) are annotated via Amazon Mechanical Turk (AMT) on a 0-to-1 continuous scale. Sample visualizations are provided in Fig. 3.



Fig. 3. Data set display diagram

4.2 Data Preprocessing

During the data preprocessing stage, for audio, FFmpeg [10] was used to extract the audio track from the video, configured with stereo channels, a sampling rate of 44,100 Hz, and a quality of 1,411 kbps. Subsequently, 26 key audio features were extracted using openSMILE [11], with each audio file size approximately 2.57 MB. For video, OpenCV was used for equidistant sampling to extract 20 key frames from a 15-s video (sampling rate of 450 frames per second). These frames were resized to 224×224 pixels to accommodate the facial feature recognition model, optimising facial feature extraction while maintaining information integrity. These steps laid the foundation for multimodal personality analysis.

4.3 Implementation Details

During feature preprocessing, recent benchmark studies indicate that most models demonstrate comparable performance with minimal differences when utilizing facial regions versus full frames. This suggests that either approach exerts relatively limited influence on model efficacy. While background and spatial context may provide personality-related information, they also risk introducing extraneous noise that compromises model accuracy. In data processing, preprocessed multimodal data (video frame images and audio data) are loaded. For model architecture, the Discrete Wavelet Transform-based Convolutional network (DWTC) serves as the visual feature extractor, which is subsequently integrated with an audio model for multimodal fusion. Training parameters are configured as follows: the L1 Loss function is employed for optimization, the AdamW optimizer updates parameters, and a cosine annealing scheduler with warm restarts adjusts the learning rate. During training, loss and accuracy metrics are periodically logged; after each epoch, weights are saved and evaluated on the validation set.

4.4 Rating Indicators

The coefficient of determination (R^2), which measures the explanatory power of the model for the data. Accuracy, Measures the accuracy of the model's predictions.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad Accuracy = \frac{1}{n} \sum_{i=1}^n$$
(7)

5 Experimental Visualisation Analysis

In this section, experiments are conducted on the ChaLearn First Impression V1 dataset to evaluate the performance of the Deep Learning-based Contextual (DLF) multi-scale network model. The effectiveness of the DLF multi-scale network is validated by comparing its results with baseline methods; ablation study outcomes are further presented to interpret model components.

5.1 Comparison of Different Ablation Experiments

Table 1: Audio modality personality recognition demonstrates MLP extracting deep audio features for prediction. In MLP + LMSA, features extracted by MLP are fed into LMSA for serialized feature alignment, serving as input to a subsequent linear regression layer for prediction. Method performance ranking: MLP + LMSA > MLP > LMSA. MLP + LMSA achieves optimal performance on audio personality recognition tasks, with an average Big Five personality prediction score of 0.8909. Although MLP yields a higher average accuracy ($R^2 = 0.8993$), MLP + LMSA overall surpasses other common methods. The superior performance of MLP + LMSA over standalone MLP and LMSA indicates complementarity between the two components.

Table 2: Video modality personality recognition employs Discrete Wavelet Transform (DWTC) for video feature processing. In DWTC+LMSA, features extracted by DWTC are fed into LMSA for serialized feature alignment, serving as input to a subsequent linear regression layer for prediction. Performance ranking: DWTC+LMSA > SWINT+LMSA > DWTC > LMSA. Key observations: DWTC+LMSA achieves optimal performance on the personality recognition task with an average Big Five prediction score of 0.9089 and $R^2 = 0.9071$, outperforming other common methods. However, SWINT+LMSA demonstrates higher metrics on Neuroticism (N), indicating stronger explanatory power for this trait. The superior performance of DWTC+LMSA over standalone DWTC and LMSA suggests complementarity between the components.

Table 1. Experimental results of ablation study on the audio modality of the First Impression dataset.

	E	A	C	N	O	Avg	R ²
MLP	0.8897	0.8939	0.8935	0.8897	0.8899	0.8911	0.8713
LMSA	0.8872	0.8714	0.8908	0.8817	0.8806	0.8825	0.8814
MLP + LMSA	0.8907	0.8902	0.8908	0.8989	0.9021	0.8909	0.8993

Table 2. Experimental results of ablation studies on the visual modality of the first impression dataset

	E	A	C	N	O	Avg	R ²
DWTC	0.8912	0.8891	0.8942	0.8916	0.8991	0.8917	0.8831
LMSA	0.8772	0.8814	0.8958	0.8867	0.8816	0.8845	0.8921
SWINT+LMSA	0.8881	0.8856	0.8986	0.9091	0.8883	0.8940	0.8931
DWTC+LMSA	0.9017	0.9092	0.9138	0.9089	0.9011	0.9069	0.9021

5.2 Multimodal Fusion

Table 3 shows the results of audio-video modality fusion. The bold entries indicate the high efficiency of the proposed model. However, due to the limited expressive power of the model in terms of openness (O), the accuracy value is relatively low.

Table 3. Experimental results of multimodal fusion ablation research at the decision layer based on the first impression dataset

	E	A	C	N	O	Avg	R ²
DWTC	0.8912	0.8819	0.8945	0.8807	0.8971	0.8891	0.8991
LMSA	0.8772	0.8814	0.8958	0.8867	0.8816	0.8845	0.8905
SWINT+LMSA	0.8889	0.8956	0.8986	0.9061	0.9183	0.9045	0.8913
DWTC+LMSA	0.9189	0.9198	0.9160	0.9182	0.9159	0.9177	0.9043

5.3 Performance Comparison

Table 3 presents the performance comparison of the DLF multi-scale network against baseline models on the ChaLearn First Impressions dataset. The DLF model outperformed existing literature for Extraversion (E), Agreeableness (A), and Openness (O) traits, though exhibiting marginally lower scores than some baselines in Conscientiousness (C) and Neuroticism (N). Crucially, it achieved a significant improvement

in the total average score (Avg). This enhanced performance is attributed to two factors: (1) The model's multi-scale feature extraction captures fine-grained structural details and leverages CNNs to derive higher-level semantics; (2) The Long-range Multi-head Self-Attention (LMSA) mechanism models global contextual dependencies across video sequences, enabling more comprehensive personality feature extraction. Table 4 summarizes these comparative results, with best performance indicated in bold

Table 4. Comparison of DLF with various baselines

	E	A	C	N	O	Avg
Gor bova et al. [12]	0.878	0.877	0.894	0.875	0.883	0.881
Daniel Helm et al [13]	0.8883	0.8986	0.8886	0.8856	0.8916	0.8905
Yağmur GüçlüTÜRK [14]	0.9107	0.9102	0.9138	0.9089	0.9111	0.9109
Zhao, X [7].	0.9175	0.9177	0.9163	0.9150	0.9167	0.9167
DLF(our)	0.9189	0.9198	0.9160	0.9182	0.9159	0.9177

6 Conclusion

This paper proposes a DLF multi-scale network model architecture for analyzing personality traits of individuals in videos. The model extracts visual and audio features from subjects, employs Discrete Wavelet Transform (DWT) to capture global and local features, and combines Convolutional Neural Networks (CNN) to extract deep semantic information. Simultaneously, it utilizes a Long-range Multi-head Self-Attention (LMSA) mechanism to achieve feature sequence alignment and information integration, with fusion implemented at the decision layer. Extensive experiments conducted on the First Impressions V1 dataset demonstrate favorable performance of the proposed model.

Acknowledgments. This work is partially supported by Research Project on Economic and Social Development of Liaoning Province(2024lsljdybkt-017), Liaoning Province Science and Technology Joint Plan (2024JH2/102600113),Fundamental Research Funds for the Central Universities(0919-140249) and Research project of China Federation of logistics and procurement (Grant No.:2025CSLKT3-307).

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Mohammadi, G., Vinciarelli, A.: Automatic personality perception: prediction of trait attribution based on prosodic features [extended abstract]. In: 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), pp. 484–490. IEEE (2015)

2. Su, M.-H., et al.: Personality trait perception from speech signals using multiresolution analysis and convolutional neural networks. In: 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 1532–1536. IEEE (2017)
3. Biel, J.-I., Teijeiro-Mosquera, L., Gatica-Perez, D.: Facetube: predicting personality from facial expressions of emotion in online conversational video. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, pp. 53–56. ACM (2012)
4. Dhall, A., Hoey, J.: First impressions: predicting user personality from twitter profile images. In: Proceedings of the 7th International Workshop on Human Behavior Understanding (HBU 2016), pp. 148–158. Springer (2016)
5. Zhang, C.L., Zhang, H., Wei, X.S., Wu, J.: Deep bimodal regression for apparent personality analysis. In: Hua, G., Jégou, H. (eds.) Computer Vision – ECCV 2016 Workshops (Lecture Notes in Computer Science), vol. 9915, pp. 311–324. Springer (2016)
6. Tzirakis, P., Trigeorgis, G., Nicolaou, M.A., Schuller, B.W., Zafeiriou, S.: End-to-end multimodal emotion recognition using deep neural networks. *IEEE J. Selected Top. Signal Process.* **11**(8), 1301–1309 (2017)
7. Zhao, X., et al.: Integrating audio and visual modalities for multimodal personality trait recognition via hybrid deep learning. *Front. Neurosci.* **16**, 1107284 (2023)
8. Chen, G., et al.: Bracketing image restoration and enhancement with high-low frequency decomposition. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 6097–6107. IEEE (2024)
9. Ponce-López, V., et al.: ChaLearn LAP 2016: first round challenge on first impressions—dataset and results. In: Computer Vision – ECCV 2016 Workshops (Lecture Notes in Computer Science), vol. 9914, pp. 400–418. Springer (2016)
10. Huang, R., et al.: An algorithm for correcting video-audio synchronization based on SyncNet. In: 2021 IEEE International Conference on Engineering, Technology & Education (TALE), pp. 1–4. IEEE (2021)
11. Eyben, F., Schuller, B.: openSMILE: the Munich open-source large-scale multimedia feature extractor. *ACM SIGMultimedia Records.* **6**(1), 4–13 (2015)
12. Gorbova, J., et al.: Automated screening of job candidate based on multimodal video processing. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2017). <https://doi.org/10.1109/CVPRW.2017.214>
13. Helm, D., Kampel, M.: Single-modal video analysis of personality traits using low-level visual features. In: 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA). IEEE (2020)
14. Gültürk, Y., Gülü, U., van Gerven, M.A.J., van Lier, R.: Deep impression: Audiovisual deep residual networks for multimodal apparent personality trait recognition. *arXiv:1609.05119*. <https://arxiv.org/abs/1609.05119> (2016)



Research Progress of Elderly Care and Rehabilitation Robots

Yang Yang¹ , Lunfei Liang² , Yu Liu² , and Houde Liu³

¹ Capital Medical University, No. 10 You'anmen Wai Street, Beijing 100069, China

² Harbin Institute of Technology, No. 92, Xidazhi Street, Harbin 150001, China

³ Tsinghua Shenzhen International Graduate School, Taoyuan Street, Shenzhen 518055, China

liu.hd@sz.tsinghua.edu.cn

Abstract. With the intensifying trend of population aging, the rehabilitation and care of elderly individuals have become a key focus of societal attention, revealing the limitations of traditional eldercare models. As an important innovation in intelligent eldercare, elderly rehabilitation robots play a vital role in improving the quality of life for older adults and alleviating the burden on caregivers. This paper aims to review the application technologies of elderly rehabilitation robots, analyze their technical characteristics, and summarize their use in daily care, medical rehabilitation, and emotional companionship. Additionally, it discusses the current challenges faced by these robots, with the goal of providing a reference for future research and innovation in the field of elderly rehabilitation robotics.

Keywords: Rehabilitation · Robots · Deep Learning · Elderly Care

1 Introduction

According to the National Bureau of Statistics, by the end of 2023, China had 297 million people aged 60 and over, accounting for 21.1% of the population, and this is projected to exceed one-third within 20 years [1]. From 2021 to 2023, 46.54 million elderly individuals were disabled, underscoring the growing need for long-term care. However, China's eldercare system faces serious challenges, including caregiver shortages and rising costs, making traditional models increasingly inadequate [2].

In this context, elder rehabilitation robots have emerged as a key innovation in intelligent eldercare. These robots provide motor rehabilitation and daily assistance, while also supporting emotional well-being through intelligent interaction [3]. Advances in artificial intelligence (AI), internet of things, and robotics have significantly enhanced their design, intelligence, and functionality, making them valuable tools for elderly care and rehabilitation [4].

This paper aims to systematically review the major application technologies currently adopted in elder rehabilitation robots, analyze their use in daily care, medical rehabilitation, and emotional companionship, and explore future development trends and challenges. The findings are intended to offer references for future research and application in this domain.

2 Technological Applications of Elderly Rehabilitation Robots

2.1 Human–Robot Interaction Technologies

2.1.1 Physical and Haptic Interaction Technologies

Physical and haptic interaction technologies in elder rehabilitation robots establish direct physical contact between humans and robots using force sensors, tactile sensors, and actuators. These technologies focus on capturing the user’s movement state, haptic feedback, and force variations to provide precise support for rehabilitation training and assistive movements. Dragusanu et al. [5] developed a modular wrist and arm exoskeleton device, where each drive unit is equipped with position and force sensors. By combining position control and force feedback, the device can coordinate with natural human movements. Sun et al. [6] designed an exoskeleton that fits closely to the user’s waist and thighs using soft straps and fastening mechanisms. It incorporates cable force feedback and end-of-cable force sensors to realize a body-perception-based tactile interaction system.

These technologies offer real-time, intuitive interaction, making them suitable for elderly users with limited mobility or speech impairments. However, challenges remain in sensor accuracy, signal latency, and noise interference in complex environments. Moreover, purely physical feedback cannot convey emotional information, limiting its role in emotional companionship.

2.1.2 Voice and Visual Interaction Technologies

Voice and visual interaction technologies are critical for enabling elder rehabilitation robots to perceive and understand user behavior, state, and needs. These technologies integrate speech recognition, natural language processing, speech synthesis, and computer vision, enabling the robot to understand verbal commands, recognize faces, capture body movements, and detect emotional states.

In voice interaction, deep neural network-based speech recognition and language modeling have advanced significantly, enabling rehabilitation robots to support smooth, conversational dialogue. Chiu et al. [7] developed an interactive scenario-based training robot called “TrainBo”, which communicates with elderly dementia patients via voice during cognitive training.

Visual interaction uses cameras, Red Green Blue–Depth sensors, or infrared imaging to perform face recognition, motion capture, posture estimation, and attention detection. Gong et al. [8] proposed a rehabilitation robotic walker based on multimodal fusion and human-computer interaction control, which integrates video data from a depth-sensing camera with other sensor inputs to detect interaction information. This system provides walking assistance and weight support for patients during gait rehabilitation.

Studies suggest that combining visual and voice interaction enhances training effectiveness in social rehabilitation robots [9]. One research [10] integrated neural networks and rule-based reasoning to enable sustained conversational interaction: the robot guides exercises via voice prompts, visual interfaces, and mechanical gestures while using depth cameras to monitor movements and offer real-time feedback. In tests with 15 stroke patients, the system achieved an average F1 score of 0.81. By integrating voice

and visual data, robots establish an “observe–understand–respond” closed loop: they confirm verbal responses visually or explain visual anomalies via speech [11]. With the advent of cross-modal fusion perception technologies, voice–vision interaction is poised to play an even larger role in eldercare robotics.

2.1.3 Social and Emotional Interaction Technologies

Social and emotional interaction technologies leverage affective computing and deep learning methods to perceive and interpret users’ emotional intent, enhancing companionship and psychological comfort. These systems boost personalized responsiveness and encourage robots to function as social companions, significantly improving emotional support for the elderly. Du et al. [12] developed a social assistance robot that integrates speech emotion recognition, semantic sentiment analysis, and topic-based emotion detection. Using Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models, the robot analyzes vocal content and language to infer emotion, which then drives changes in facial expression and calming tone, creating an emotional feedback loop. Meng et al. [13] implemented an emotion-aware interactive robot designed to enhance user experience: it can intuitively sense emotional states and respond accordingly. This type of system is particularly suitable for companions for older adults living alone, offering comfort, health nudges, and lifestyle advice.

2.2 Driving Mechanisms

2.2.1 Rigid Actuation

Rigid actuators typically employ electric motors coupled with gearboxes or reducers to provide stable and high-torque power output for exoskeletons. These actuators are easy to control and enable precise motion control with high bandwidth, making them the most common actuation method in multi-joint rehabilitation robots [14]. In Dragusaru et al.’s [5] study, brushless direct current motors serve as the core driving elements, amplifying torque through high-precision planetary gear sets and worm gear mechanisms to provide adequate motion support. Hydraulic actuators, another form of rigid drive, are often used in high-power scenarios and can generate substantial output forces, suitable for assisting movements that require high loads. Su et al. [15] proposed an efficient hydraulic actuation design that utilizes hydraulic actuators as the primary power source for exoskeletons, effectively improving system efficiency while reducing energy loss caused by fluid leakage and internal friction.

Rigid actuators offer high control precision and fast response, making them well-suited for tasks requiring accurate motion. However, they are generally heavy and less compliant to the human body, leading to discomfort and potential safety risks. Furthermore, their complex mechanical structure and high maintenance costs make them less ideal for daily use.

2.2.2 Flexible Actuation

Flexible actuation aims to achieve safe, comfortable, and coordinated physical interaction between humans and robots. Unlike rigid actuators that focus on high torque,

flexible actuators prioritize compliance, shock absorption, and lightweight structure. Common implementations include series elastic actuators (SEAs), variable stiffness actuators (VSAs), Bowden cable transmission, and pneumatic actuators.

SEAs and VSAs introduce compliance via elastic components, effectively reducing impact forces and output impedance during physical contact, thus improving interaction safety and comfort [16]. One study [17] designed an SEA-based elbow joint device, where springs enhanced compliance while maintaining a compact form. Bowden cables transmit power remotely via internal steel wires and outer sheaths, allowing heavy motors to be placed away from the joints. Li et al. [18] designed a back-mounted Bowden system that improved transmission efficiency to 0.78 and reduced walking load for older users.

Pneumatic actuators provide lightweight and compliant motion using compressed air. A study [19] proposed an innovative pneumatic exoskeleton joint design using pneumatic artificial muscles as air tanks and employing cylinders to store and recycle energy to generate torque. By regulating cylinder pressure, the system achieves high response speed and precise torque control under varying conditions.

While flexible actuators offer superior comfort and long-term wearability, they typically provide lower torque, slower response, and require advanced control algorithms due to their nonlinear behavior.

2.2.3 Comparison of Driving Mechanisms

Rigid and flexible drives each have distinct advantages and limitations in elderly care and rehabilitation robots. Table 1 compares the characteristics of rigid and flexible drives across five aspects.

Table 1. Comparison of Characteristics between Rigid Drive and Flexible Drive

Characteristic	Rigid Drive	Flexible Drive
Torque/Bandwidth	High/High	Low/Low
Response Speed	Fast	Slow
Weight and Portability	Heavy, poor portability	Light, good portability
Safety and Comfort	Low, rigid structure	High, compliant to human movement
Control Difficulty	Mature algorithms	High nonlinearity, requires compensation

As shown by the comparison, rigid drives are suitable for scenarios requiring high-precision and high-stability control, such as hospital training systems or industrial collaborative robots. Their control strategies are mature, with fast response but sacrificed comfort. Flexible drives are better suited for rehabilitation and assistive devices that require long-term human contact, emphasizing safety and compliance, ideal for elderly or early rehabilitation patients, though their control design is more complex.

2.3 Control Strategies

2.3.1 Assist-as-Needed Control

Assist-as-needed (AAN) control enables the robot to dynamically adjust its support based on the user's voluntary effort, thereby encouraging active participation in rehabilitation. Common strategies include force-feedback control, impedance control, and intention recognition-based control.

Force-feedback control involves real-time monitoring of contact forces between the exoskeleton and the user. The control system adjusts the output force based on user needs to ensure smooth and natural movement. Wen et al. [20] proposed a novel control method based on a complete exoskeleton dynamics model, which can effectively estimate and regulate human-robot interaction forces.

Impedance control focuses on adjusting the contact impedance between the exoskeleton and the user to enhance movement coordination. This approach allows patients to control the exoskeleton through natural gait patterns. Murray et al. [21] proposed a lower-limb exoskeleton control method based on impedance control, which adapts the assistive force according to the patient's gait, thereby avoiding rigid motion patterns.

Intention recognition-based control uses signals such as EMG and EEG to detect the user's motor intent and convert it into control commands. It offers high personalization and is well-suited for active rehabilitation but requires advanced algorithms and signal processing. Kueper et al. [22] developed an exoskeleton control system that detects motor intentions using EEG signals combined with resonance training, enabling mirror rehabilitation of limbs. Experimental results demonstrated its ability to predict and trigger assistive actions.

2.3.2 Adaptive Control

Adaptive control enables elderly care and rehabilitation robots to dynamically adjust interaction strategies in real time based on the user's motor ability, rehabilitation progress, and emotional state, thereby improving training outcomes and user experience. Wang et al. [23] proposed a method that extracts gait motion trajectories using motion capture systems and integrates them into a gait planning framework. A control system based on adaptive iterative learning was simulated, demonstrating high-precision tracking performance. Studies have shown that intelligent rehabilitation robots can optimize training strategies through reinforcement learning to accommodate individual recovery processes [24]. This adaptability significantly enhances the effectiveness of training and aligns the robot's assistance with the actual needs of elderly users.

3 Applications of Elderly Rehabilitation Robots

3.1 Daily Care Assistance

Daily care assistance robots are designed to help older adults perform basic activities in daily life, with key functionalities including mobility support, life assistance, and safety monitoring.

Mobility assistance is one of the core functions, especially for elderly people with limited mobility. These robots provide walking support and enhance lower limb strength to reduce fall risks. Wei et al. [25] developed a modular lower-limb exoskeleton robot that accurately estimates the user's applied force and real-time balance status to offer safe walking support.

In terms of life care, daily care assistance robots also play an important role. Pei et al. [26] designed and implemented a binocular vision-based intelligent meal delivery robot system capable of autonomously navigating within homes or care institutions and delivering meals to elderly users on time, reducing difficulties caused by limited mobility. This type of safety monitoring robot can continuously monitor the elderly's movement and physiological states and respond promptly to emergencies.

3.2 Medical Rehabilitation Support

Medical rehabilitation support robots mainly target elderly individuals' motor rehabilitation, cognitive intervention, and chronic disease management.

In motor rehabilitation, these robots are widely used for gait training and limb function recovery in patients with stroke, spinal cord injury, Parkinson's disease, and other conditions. Liu et al. [27] proposed an upper-limb rehabilitation robot designed to enhance the naturalness of training and increase patient engagement. Experiments demonstrated that this robot effectively improves patients' range of motion and muscle strength recovery speed.

Medical rehabilitation support robots are also applied to neurodegenerative diseases such as Alzheimer's and mild cognitive impairment. Pandey et al. [28] introduced the Pepper robot developed by SoftBank Robotics, which provides Alzheimer's patients with natural and emotionally rich interactions, helping to alleviate social isolation to some extent.

For chronic disease management, wearable medical robots combined with AI provide personalized health management, physiological monitoring, and remote medical support for conditions such as hypertension, diabetes, and cardiovascular diseases [29].

3.3 Emotional Interactive Companion Robot

Emotional companion robots serve elderly living alone or with disabilities by providing reminders, emotional support, and social interaction via voice, facial recognition, and personalized learning. The Japanese biomimetic robot "Paro" uses tactile sensors, facial expressions, and voice interaction to improve dementia patients' conditions and reduce psychiatric drug use [30]. Another study integrated a large language model-based conversational agent into a Mini robot, enabling smooth, natural dialogue with users [31]. However, this field faces ethical controversies, such as the potential risks of emotional dependency, privacy and data security concerns, and the balance between human care and machine substitution [32]. Therefore, stricter regulations and policy support are urgently needed.

4 Challenges Faced by Elderly Care Rehabilitation Robots

4.1 Technical Limitations

Elderly care and rehabilitation robots involve complex hardware and software systems, and their technological maturity remains limited. When performing tasks like walking assistance or rehabilitation training, these robots may fail due to sensor errors, unstable algorithms, or mechanical faults, posing safety risks. A study [33] assessing privacy risks of care robots for the elderly pointed out potential software vulnerabilities that could be exploited by hackers, leading to privacy breaches. Moreover, most current elderly care robots are still in experimental stages, with product stability and reliability needing improvement. How to achieve wider adoption remains a key research focus [34].

4.2 High Costs

Developing an elderly care and rehabilitation robot involves multiple costly processes such as mechanical design, AI algorithm development, and testing optimization. Although studies show that in the long term these robots can reduce labor costs and shorten rehabilitation periods, the high upfront cost remains a major barrier to widespread use [35]. For example, ReWalk's exoskeleton robot is priced between \$70,000 and \$80,000, and Ekso Bionics is priced around \$80,000, making them unaffordable for most consumers. Future efforts may draw on the automotive industry's experience by using automated assembly lines to improve production efficiency and reduce costs, thus promoting more effective dissemination of elderly care robots.

4.3 Societal Perceptions

The promotion of elderly care robots is also influenced by social attitudes. Many elderly people have low acceptance of smart devices; research shows they tend to rely more on human caregivers and may feel a threat to their self-worth when robots replace caregiving functions [36]. Some groups misunderstand elderly care robots, believing they cannot provide sufficient humane care and worry that robots will replace caregiving jobs, impacting employment [37]. Also, limited understanding of these robots' functions leads to misconceptions and undervaluation. Improving public acceptance requires coordinated efforts from governments, companies, and social groups through education and outreach.

5 Conclusion

This paper explores the application technologies and scenarios of elderly care and rehabilitation robots, highlighting the roles of human–robot interaction, actuation technologies, and control strategies. It also reviews their current use in daily care, medical rehabilitation, and emotional support. Despite their great potential to improve the quality of life for older adults, widespread adoption is still limited by several challenges. Future research should focus on enhancing intelligence, reducing costs, and promoting policy support and public awareness to enable broader implementation.

References

1. National Bureau of Statistics: Statistical communiqué of the People's Republic of China on the 2023 national economic and social development. *China Statistics*. **3**, 4–21 (2024)
2. National Health Commission of the People's Republic of China: Statistical communiqué on China's health development in 2023. *Chinese Rural Health*. **16**(9), 6–14 (2024)
3. Luo, S., Meng, Q., Yu, H.: Research and application of rehabilitation robotics technology in China. *Chin. J. Rehabil. Med.* **38**(12), 1762–1768 (2023)
4. Cesário, L.C., Barbosa, P., Miguel, P.A.C., Mendes, G.H.S.: Service robots in caring for older adults: uncovering the current conceptual and intellectual structures and future research agenda. *Arch. Gerontol. Geriatr.* **131**, 105755 (2025)
5. Dragusanu, M., Iqbal, M.Z., Baldi, T.L., Prattichizzo, D., Malvezzi, M.: Design, development, and control of a hand/wrist exoskeleton for rehabilitation and training. *IEEE Trans. Robot.* **38**(3), 1472–1488 (2022)
6. Sun, L., Deng, A., Wang, H., Zhou, Y., Song, Y.: A soft exoskeleton for hip extension and flexion assistance based on reinforcement learning control. *Sci. Rep.* **15**(1), 5435 (2025)
7. Chiu Fung, K., Mow, W. H.: TrainBo: an interactive robot-assisted scenario training system for older adults with dementia, <https://arxiv.org/pdf/2505.12080.pdf>. Last accessed 2025/06/14
8. Gong, T., et al.: Multimodal fusion and human–robot interaction control of an intelligent robot. *Front. Bioeng. Biotechnol.* **11**, 1310247 (2023)
9. Garcia, S., Gomez-Donoso, F., Cazorla, M.: Enhancing human–robot interaction: development of multimodal robotic assistant for user emotion recognition. *Appl. Sci.* **14**(24), 11914 (2024)
10. Hun Lee, M., Siewiorek, D.P., Smailagic, A., Bernardino, A., Bermúdez i Badia, S.: Design, development, and evaluation of an interactive personalized social robot to monitor and coach post-stroke rehabilitation exercises. *User Model. User-Adapt. Interact.* **33**(2), 545–569 (2023)
11. Fahn, C.-S., et al.: Image and speech recognition technology in the development of an elderly care robot: practical issues review and improvement strategies. *Healthcare*. **10**(11), 2252 (2022)
12. Du, Y., Zhang, K., Trovato, G.: Composite emotion recognition and feedback of social assistive robot for elderly people. In: Degen, H., Ntoa, S. (eds.) 4th International Conference, AI-HCI 2023, Held as Part of the 25th HCI International Conference, HCII 2023, pp. 220–231. Springer, Cham (2023)
13. Meng, Z., et al.: PERCY: Personal emotional robotic conversational system, <https://arxiv.org/pdf/2503.16473.pdf>. Last accessed 2025/06/14.
14. Masengo, G., Zhang, X., Dong, R., Alhassan, A.B., Hamza, K., Mudaheranwa, E.: Lower limb exoskeleton robot and its cooperative control: a review, trends, and challenges for future research. *Front. Neurorobot.* **16**, 913748 (2023)
15. Su, Q., Pei, Z., Tang, Z.: Analysis of dynamic characteristics of hydraulic driven system for weight bearing lower limb exoskeleton. *J. Phys. Conf. Ser.* **2365**(1), 12011 (2022)
16. Bhujel, S., Hasan, S.K.: A comparative study of end-effector and exoskeleton type rehabilitation robots in human upper extremity rehabilitation. *Hum.-Intell. Syst. Integr.* **5**(1), 11–42 (2023)
17. Jenks, B., Levan, H., Stefanovic, F.: OpenSEA: a 3D printed planetary gear series elastic actuator for a compliant elbow joint exoskeleton. *Front. Robot. AI.* **12**, 1528266 (2025)
18. Li, X., Liu, J., Li, W., Huang, Y., Zhan, G.: Force transmission analysis and optimization of bowden cable on body in a flexible exoskeleton. *Appl. Bionics Biomech.* **2022**, 5552166 (2022)
19. Mišković, L., Dežman, M., Petrič, T.: Pneumatic exoskeleton joint with a self-supporting air tank and stiffness modulation: design, modeling, and experimental evaluation. *IEEE/ASME Trans. Mech.* **29**(5), 3415–3426 (2024)

20. Küçüktabak, E.B., et al.: Haptic transparency and interaction force control for a lower limb exoskeleton. *IEEE Trans. Robot.* **40**, 1842–1859 (2024)
21. Murray, S.A., Ha, K.H., Hartigan, C., Goldfarb, M.: An assistive control approach for a lower-limb exoskeleton to facilitate recovery of walking following stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **23**(3), 441–449 (2015)
22. Kueper, N., Kim, S.K., Kirchner, E.A.: Avoidance of specific calibration sessions in motor intention recognition for exoskeleton-supported rehabilitation through transfer learning on EEG data. *Sci. Rep.* **14**(1), 16690 (2024)
23. Wang, X., et al.: A multistage hemiplegic lower-limb rehabilitation robot: design and gait trajectory planning. *Sensors.* **24**(7), 2310 (2024)
24. Su, D., Hu, Z., Wu, J., Shang, P., Luo, Z.: Review of adaptive control for stroke lower limb exoskeleton rehabilitation robot based on motion intention recognition. *Front. Neurorobot.* **17**, 1186175 (2023)
25. Wei, X., Hu, Z.: Research on control and design of the lower limb exoskeleton robot based on the center of gravity stability constraints. *Mech. Eng. Technol.* **9**(2), 35–52 (2020)
26. Pei, H., Qin, H.: Design of intelligent meal delivery robot system based on binocular vision. *Appl. Electron. Tech.* **47**(1), 96–99 (2021)
27. Liu, G., Zeng, L., Meng, Q., Xu, X., Meng, Q., Yu, H.: Six-degree-of-freedom upper limb rehabilitation robot based on tight-coupled dynamic interactive control: design and implementation. *Robot. Auton. Syst.* **185**, 104895 (2025)
28. Pandey, A.K., Gelin, R.: A mass-produced sociable humanoid robot: pepper: the first machine of its kind. *IEEE Robot. Autom. Magaz.* **25**(3), 40–48 (2018)
29. Zhang, X., Zhang, F., Li, M., Wang, J.: Applications and challenges of intelligent robots in primary chronic disease management. *Chin. Gen. Pract.* **28**(1), 7–12 (2025)
30. Kang, H.S., Makimoto, K., Konno, R., Koh, I.S.: Review of outcome measures in PARO robot intervention studies for dementia care. *Geriatr. Nurs.* **41**(3), 207–214 (2020)
31. Esteban-Lozano, I., Castro-González, Á., Martínez, P.: Using a LLM-based conversational agent in the social robot mini. In: Degen, H., Ntoa, S. (eds.) 5th International Conference, AI-HCI 2024, Held as Part of the 26th HCI International Conference, HCII 2024, pp. 15–26. Springer, Cham (2024)
32. Miao, F.: The caring capacity and ethical boundaries of elderly care robots. *Med. Philos.* **45**(6), 31–36 (2024)
33. Grabler, R., Koeszegi, S.T.: Privacy beyond data: assessment and mitigation of privacy risks in robotic technology for elderly care. *ACM Trans. Human-Robot Interaction.* **14**(1), 1–23 (2025)
34. Silvera-Tawil, D.: Robotics in healthcare: a survey. *SN Comput. Sci.* **5**(1), 189 (2024)
35. Cano-de-la-Cuerda, R., et al.: Economic cost of rehabilitation with robotic and virtual reality systems in people with neurological disorders: a systematic review. *J. Clin. Med.* **13**(6), 1531 (2024)
36. Zhou, Y., Yao, Y., Ye, X.: A meta-synthesis of community-dwelling older adults' perceptions and expectations of care robots. *J. Nursing (China).* **30**(9), 43–48 (2023)
37. Su, Q., Wang, Y.: Ethical dilemmas and countermeasures in the application of artificial intelligence in clinical nursing practice. *Med. Philos.* **44**(24), 51–55 (2023)



Anti-swing Control of 5-DOF Marine Crane Based on ASMC-PD Mode Switching Method

Qin Zhang, Gaoli Xu, and Xiong Hu^(✉)

College of Logistics Engineering, Shanghai Maritime University, Shanghai 201306, China
1390606178@qq.com

Abstract. To address the control issues of a five-degree-of-freedom marine crane with parameter uncertainties, external disturbances, and unmodeled disturbances, an adaptive sliding mode and PD mode switching controller (ASMC-PD) is proposed. Firstly, a five-degree-of-freedom marine crane model is established, and a dynamic analysis method for the crane under sea conditions is proposed, which takes into account the effects of wave excitation and gravity. Secondly, the uncertainties in system parameters and unmodeled disturbances are estimated and compensated through adaptive laws, and then a composite control law is obtained by combining sliding mode control and PD mode switching to enhance the anti-interference ability of the marine crane. The stability of the proposed control system is strictly proved using the Lyapunov method. Finally, through simulation experiments, the control effects of this method, PID, MPC and undiminished pendulum are verified. The simulation results show that, compared with the no anti-sway control, the lateral and longitudinal sway angles of ASMC-PD are reduced by more than 50%. Compared with PID control and MPC control, especially in dynamic sea conditions of 4 to 6 levels, the pendulum reduction efficiency of ASMC-PD is very significant, confirming the feasibility and effectiveness of the method.

Keywords: five degrees of freedom control · marine crane model · adaptive sliding mode and PD mode switching controller

1 Introduction

The complex and changeable Marine environment is a great challenge for Marine cranes [1]. The interaction of interference factors such as wave impact, sea breeze and ocean current instability greatly increases the difficulty of crane control, which not only causes structural vibration, increases system uncertainty, but also affects its positioning accuracy and motion stability.

Marine crane system is a typical underactuated system, and its control input dimension is less than the number of generalized degrees of freedom of the system [2, 3]. In practical applications, the main objectives of the Marine crane system are two: first, the large arm can accurately and quickly reach the target position to ensure the transportation efficiency; The second is to suppress the swing of the payload during transportation to

keep it within an acceptable range for safety reasons. In addition, the nonlinear characteristics of Marine cranes make the mathematical model of the system complicated and difficult to accurately describe and analyze [4, 5].

In order to solve the above problems, many researchers have proposed a large number of control schemes, which are mainly divided into open-loop control technology and closed-loop control technology. Open-loop controller does not need swing Angle feedback, and can accurately suppress the swing of the hook and the payload. However, this method requires high precision of system parameters, and the control effect is seriously affected by interference. By estimating the system variables, the closed-loop control scheme can obtain the desired positioning, reduce the swing of the payload, and increase the stability of the system. The closed-loop control schemes widely used in Marine crane systems mainly include proportional integral differential control (PID) [6], fuzzy control [7], adaptive control [8–10] and sliding mode control [11–13]. However, facing the complex and changeable Marine environment, the performance of most control schemes is sensitive to changes in model parameters and perturbations. In the above control scheme, sliding mode control has great advantages in the controller design of Marine crane system, that is, it has good robustness to disturbance and model defects. However, this sliding mode control method still has some disadvantages such as buffering. In order to suppress chattering, the concept of reaching law is usually used to solve the jitter interference problem. The control method of disturbance suppression is also an important field of sliding mode control. Eun et al. [14] designed a discrete sliding mode observer to realize the observation of interference at the control input, so as to effectively compensate the interference and relatively reduce the switching gain. However, the switching gain reflects the system's resistance to uncertainty and interference to some extent. Reducing switching gain may make the system less robust to external disturbances and model uncertainties, resulting in the system being more likely to deviate from the ideal sliding mode motion when disturbed. Therefore, an adaptive gain controller is used to solve this problem.

The main problems of the existing crane control methods are as follows: (1) the existing crane model is not accurate, most of the methods only analyze the two-dimensional model, ignore the movement characteristics of the ship in the three-dimensional space, which is inconsistent with the actual situation, resulting in the control strategy cannot fully consider the impact of the ship's rolling, pitching and heave movements on the crane; (2) Most control methods require accurate model parameters, which are difficult to meet in practical applications. In addition, there are often some dynamic characteristics in the system that are difficult to model accurately, and these unmodeled dynamics may have a negative impact on the control performance of the system. The main work of this paper includes: (1) a five-degree-of-freedom dynamic model of Marine crane is established considering ship roll and pitch motion under the action of sea waves. (2) In order to provide the control efficiency and robustness of the system, an adaptive sliding mode and PD mode switching controller is designed. External interference is suppressed by sliding mode control and PD mode switching, and the uncertainty of the system is estimated and compensated in real time by introducing adaptive gain, which improves the security and reliability of the system.

2 Dynamic Model of Crane

In order to facilitate the Marine operation of Marine cranes, most of the cranes are placed on the side of the deck, so it is affected by the ocean rolling and pitching disturbance, as well as the hull size, the slewing angle of the rotating table, the luffing angle, the length of the rope and other factors, so that its inertia space makes a similar arc-shaped movement.

The five-degree of freedom Marine crane system considering ship roll and pitch disturbance is shown in Fig. 1. Coordinate system $o_E - xyz$ represents the inertial coordinate system, $o - x_s y_s z_s$ represents the ship coordinate system, and $o - x_c y_c z_c$ represents the crane coordinate system. The Angle between B' and x_c axis of the projection of the big arm B on $ox_c y_c$ is defined as the slewing angle φ_1 , and the Angle between the projection B' and the big arm itself is the luffing angle φ_2 . The plane where the Angle is located is called C , the transverse Angle between the payload p and its projection on the plane C is θ_1 , and the longitudinal Angle between the projection B' of the payload and the vertical line from the end point of the big arm to the $ox_c y_c$ plane is θ_2 , so that θ_1 and θ_2 can represent the swing of the payload. The parameters and variables of Marine cranes are defined in Table 1.

Table 1. Parameters and variables of Marine crane system

ID	System parameters/variables	symbol	unit
1	Slewing angle	φ_1	rad
2	Luffing angle	φ_2	rad
3	Lateral swing Angle	θ_1	rad
4	Longitudinal swing Angle	θ_2	rad
5	Rope length	l	m
6	Arm length	l_b	m
7	Turning moment	f_M	N
8	Pitching moment	f_N	N
9	Rope length tension	f_y	N

The hull coordinate system describes the attitude of the hull in space, so it changes with the change of the hull attitude, so it is necessary to introduce the coordinate transformation matrix. The rotation Angle of the ship around axis x_s is set as the roll Angle α' and the rotation Angle around axis y_s is set as the pitch Angle β' . In this paper, the rotation matrix of the ship coordinate system relative to the inertial coordinate system is obtained:

$$R_1 = R_x(\alpha')R_y(\beta')$$

$$\begin{aligned}
 R_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix} \\
 R_y &= \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix} \\
 R_z &= \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}
 \end{aligned} \tag{1}$$

In Formula (1), $R_i(i = x, y, z)$ is the rotation matrix in each direction of the three-dimensional space, and α, β, γ is the Angle of rotation around each axis.

Rotation matrix of the crane coordinate system with respect to the inertial coordinate system:

$$R_2 = R_x R_y R_z \tag{2}$$

Given the position change between the Marine crane and the ship, this paper gives the position vector of the origin of the Marine crane coordinate system o in the inertial coordinate system with the following function form:

$$p_2 = p_1 + R_1 p'_2 \tag{3}$$

In Formula (3): $p_1 = [0 \ 0 \ 0]^T$ represents the origin of the inertial coordinate system. $p_2 = [0 \ 2 \ 1]^T$ represents the position of the crane coordinate origin in the inertial coordinate system. p'_2 is the point in the crane coordinate system.

The position of payload p in the $o - x_c y_c z_c$ crane coordinate system is expressed as:

$$\begin{aligned}
 x_p &= l(\cos \varphi_1 \cos \theta_2 \sin \theta_1 - \sin \varphi_1 \sin \theta_2) + l_b \cos \varphi_2 \cos \varphi_1 \\
 y_p &= l(\sin \varphi_1 \cos \theta_2 \sin \theta_1 - \cos \varphi_1 \sin \theta_2) + l_b \cos \varphi_2 \sin \varphi_1 \\
 z_p &= -l \cos \theta_2 \sin \theta_1
 \end{aligned} \tag{4}$$

The formula for calculating the speed of the payload relative to $o - x_c y_c z_c$ crane can be obtained:

$$v_p = (\dot{x}_p \ \dot{y}_p \ \dot{z}_p)^T \tag{5}$$

In summary, the total kinetic energy T of the Marine crane system is the sum of the boom kinetic energy T_1 and the payload kinetic energy T_2 :

$$T = T_1 + T_2$$

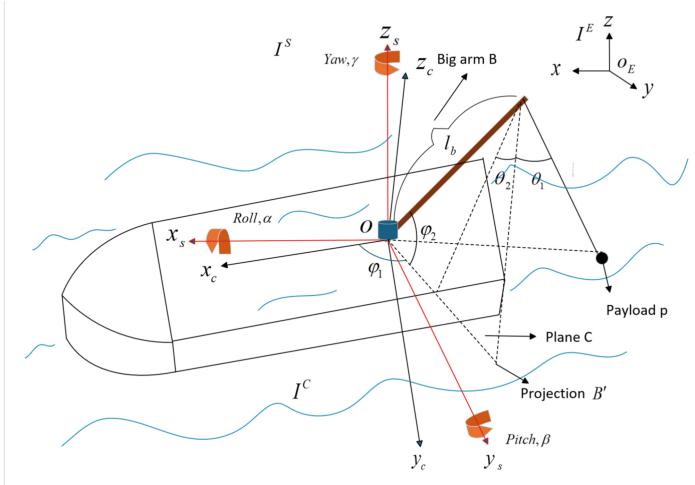


Fig. 1. Marine crane system model

$$\begin{aligned} T_1 &= \frac{1}{6}ml_b^2\dot{\varphi}_2^2 + \frac{1}{6}ml_b^2\dot{\varphi}_1^2\cos^2\varphi_2 \\ T_2 &= \frac{1}{2}m_p(\dot{x}_p^2 + \dot{y}_p^2 + \dot{z}_p^2) \end{aligned} \quad (6)$$

According to the Lagrange dynamics equation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial T}{\partial q} + \frac{\partial V_q}{\partial q} = F \quad (7)$$

In Formula (7), $q = [\varphi_1 \varphi_2 \ l \ \theta_1 \ \theta_2] \in R^5$ represents the generalized coordinate vector of the system, and $F = [f_M \ f_N \ f_y \ 0 \ 0]$ represents the generalized input force of the system, V_q represents the potential energy term.

The established dynamics equation of the five-degree-of-freedom Marine crane system is converted into the Lagrange standard state matrix description form:

$$\begin{aligned} M(q)\ddot{q} + a &= F \\ M(q) &= \begin{bmatrix} M_{11} & M_{12} & M_{13} & M_{14} & M_{15} \\ M_{21} & M_{22} & M_{23} & M_{24} & M_{25} \\ M_{31} & M_{32} & M_{33} & M_{34} & M_{35} \\ M_{41} & M_{42} & M_{43} & M_{44} & M_{45} \\ M_{51} & M_{52} & M_{53} & M_{54} & M_{55} \end{bmatrix} \\ a &= [a_1 \ a_2 \ a_3 \ a_4 \ a_5]^T \end{aligned} \quad (8)$$

In Formula (8), $M(q)$ is the inertia matrix of the system, and is a symmetric positive definite matrix, a is the wave excitation and gravity terms, see the appendix for specific definitions, and \ddot{q} is the acceleration vector.

3 Controller Design

3.1 Controller Design Based on Sliding Mode Variable Structure

Marine crane is a typical underactuated system, the 5° of freedom control is more complicated, this paper adopts sliding mode variable structure controller design. The sliding mode controller uses three separate sliding modes to control the slewing angle, the luffing angle, and the rope length. The defined error e is:

$$e = q - q' = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \varphi_1 - \varphi'_1 \\ \varphi_2 - \varphi'_2 \\ l - l' \\ 0 \\ 0 \end{pmatrix} \quad (9)$$

In Formula (9), φ'_1 represents the expected swing Angle of the lateral swing Angle, φ'_2 represents the expected swing Angle of the longitudinal swing Angle, and l' represents the expected rope length.

Through the error variable, the design sliding die surface s is:

$$s = \dot{e} + \sigma e \quad (10)$$

In Formula (10), $\sigma = [\sigma_1 \sigma_2 \sigma_3 \sigma_4 \sigma_5]^T$ is a constant diagonal gain matrix.

The derivative of the sliding mode surface obtained from Eq. (8) is:

$$\begin{aligned} \dot{s} &= \ddot{e} + \sigma \dot{e} = \ddot{q} - \ddot{q}' + \sigma \dot{e} \\ &= M^{-1}F - M^{-1}a - \ddot{q}' + \sigma \dot{e} \end{aligned} \quad (11)$$

For the five-degree-of-freedom Marine crane control system, because the model takes into account the change of slewing angle, variation of luffing angle and rope length, different control modes are set to determine the final control input $F(t)$. At this time, there are two modes.

- Mode 1, the final control input is a linear combination of equivalent control and switching control:

$$F(t) = k_1 F_a + k_2 F_1(t) + k_3 F_2(t)$$

$$\begin{cases} F_a = k'_p dq \\ F_1(t) = -k_s s \\ F_2(t) = -\zeta sign(s) \end{cases} \quad (12)$$

In Formula (12), k_1, k_2, k_3 is the parameter value, F_a is the equivalent control, k'_p is the proportional gain, and dq is the combination of the expected acceleration and velocity error. $F_1(t)$ is the linear approach term, k_s is the sliding mode control gain, this form

is essentially linear feedback control based on sliding mode surface s . $F_2(t)$ is a sign function switch, ζ is a correction parameter to the face value of the sliding mode, and $sign(s)$ is a sign function. The introduction of symbol function makes the control input have completely different symbols on both sides of the sliding mode plane. When $s > 0$, $sign(s) = 1$, the control input $F_2(t)$ is negative; When $s < 0$, $sign(s) = -1$, control input $F_2(t)$ is positive. This positive and negative switching characteristic can produce a strong driving force that drives the system state to move rapidly to the sliding mode surface, ensuring that the system can quickly enter the sliding mode.

- Mode 2, the final control input adopts PD control mode:

$$F(t) = -k_p e - k_d \dot{e} \quad (13)$$

In Formula (13), k_p is the proportional gain and k_d is the differential gain.

With the combination of the two control methods, in some uncertain links, PD control can adjust the control quantity in real time according to the error, make up for the lack of adaptability of sliding mode control to parameter changes under specific circumstances, and make the system maintain good performance under different working conditions. The rapid response characteristic of sliding mode control combined with the fine adjustment ability of PD control can make the control system play the advantages of different control methods in different stages and optimize the overall control effect.

3.2 Stability Analysis

Although the sliding mode control is robust to parameter variation and external interference, it is usually based on fixed robust performance design, which is difficult to adapt to the uncertain time-varying and complex characteristics. Adaptive gain can estimate and compensate the uncertainty of the system in real time, enhance the adaptability of the system to time varying, unknown interference and parameter perturbation, and improve the robustness of the system. Adaptive control is a control method that can modify its own characteristics to adapt to changes in the dynamic characteristics of objects and perturbations. Adaptive gain control can dynamically adjust the switching gain of sliding mode control to realize on-line compensation of system uncertainty.

The adaptive law of time-varying gain ζ is designed:

$$\zeta = \gamma |s| (\gamma > 0) \quad (14)$$

Assuming the system uncertainty (such as unmodeled dynamics, perturbations) is Δ , and $|\Delta| \leq D$ is satisfied, then the dynamic equation can be simplified to:

$$\dot{s} = -k_s s - \zeta sign(s) + \Delta \quad (15)$$

Define the Lyapunov function as:

$$V = \frac{1}{2} s^2 + \frac{1}{2\gamma} (\zeta - \zeta')^2 \quad (16)$$

In Formula (16), ζ' is the estimate of ζ , $\zeta' \geq D + \eta$ is the upper bound of the ideal gain ($\eta > 0$ is the design parameter).

Take the derivative of V :

$$\dot{V} = s\dot{s} + \frac{1}{\gamma}(\zeta - \zeta')\dot{\zeta} \quad (17)$$

Substituting formula (15) gives:

$$\dot{V} = s(-k_s s - \zeta \text{sign}(s) + \Delta) + (\zeta - \zeta')|s| \quad (18)$$

Using $|\Delta| \leq D$, there are:

$$\Delta s \leq |\Delta| |s| \leq D |s| \quad (19)$$

As a result:

$$\begin{aligned} \Delta s + (\zeta - \zeta')|s| &\leq D|s| + (\zeta - \zeta')|s| \\ &= (\zeta - \zeta' + D)|s| \end{aligned} \quad (20)$$

Since $\zeta' \geq D + \eta$, substituting yields:

$$\zeta - \zeta' + D \leq \zeta - \eta \quad (21)$$

To sum up:

$$\dot{V} \leq -k_s s^2 - \eta |s| \quad (22)$$

According to Lyapunov's theorem, \dot{V} is negative definite, the system is globally asymptotically stable at the equilibrium point $s = 0$, and the adaptive gain ζ is dynamically adjusted by $\dot{\zeta} = \gamma |s|$ to ensure $\zeta \geq \zeta' = D + \eta$, thus overcome uncertainty D and ensuring that the system state converges to the sliding mode plane $s = 0$ in a finite time.

4 Simulation Results

This section first verifies the effectiveness of adaptive sliding mode and PD mode switching controller, builds Simulink simulation model, sets different wind wave levels, different wave direction angles, different draft depths, different speed, initial slewing angle, luffing angle and rope length, simulates the rolling and pitching motion of different wave levels, and sets the relevant control parameters of the controller. The effectiveness of the method is verified by the simulation results, the swing amplitude, the response reality and the swing reduction efficiency are analyzed, and the feasibility and superiority of the adaptive sliding mode and PD mode switching controller are analyzed qualitatively and quantitatively.

Firstly, the wind wave-ship motion simulation software based on PM spectrum is used to simulate the roll and pitch motion of the ship under sea conditions of grades 4 to 6. Thus, the free swing variation of the effective load without pendulum reduction under sea conditions of grades 4 to 6 is obtained.

In order to further quantitatively describe the dynamic performance of the system and verify the superiority of the adaptive sliding mode and PD mode switching controller, the pendulum reduction efficiency is defined as the evaluation index. The expression is: as follows:

$$\eta = \frac{\theta_{\max} - \dot{\theta}}{\theta_{\max}} \quad (23)$$

In Formula (23), θ_{\max} is the maximum peak Angle of swing Angle swing without undiminished pendulum, $\dot{\theta}$ are the maximum peak angles of the swing Angle after subtracting the pendulum by adding the control method.

The initial parameter of the Marine crane is set as 70° slewing angle φ_1 , 0° luffing angle φ_2 , and 1 m rope length l . The system parameters in the simulation environment are shown as follows:

$$m_1 = 73.4, m_p = 30, l_b = 2.4$$

Consider two scenarios for simulation:

Case 1: In a static environment

The static environment is the test of anti-interference ability of Marine crane when it is not working. When it is subjected to roll and pitch interference, the simulation results are shown in Fig. 2. The undiminished pendulum control and MPC control show irregular and large oscillations. The ASMC-PD control converges within 10 s and is faster and more stable than the PID control. Undiminished pendulum control shows an irregular swing angle change and has no convergence trend, making the system difficult to stabilize. PID control and MPC control have large initial fluctuations and long adjustment times when the rope length changes, while ASMC-PD control has small fluctuations and fast convergence of the expected position.

Case 2: In a dynamic environment

The dynamic environment is the anti-interference ability test of Marine crane operation. At this time, it is disturbed by rolling and pitching. The input parameter slewing angle φ_1 is -50° , the luffing angle φ_2 is -30° , and the rope length l is 2 m. The simulation results are shown in Fig. 3. The swing amplitude of PID control is even larger than that of undiminished reduction control and MPC control, and its control effect is the worst. The swing amplitude of the undiminished pendulum control reaches 63.77° at the beginning, while the ASMC-PD control also fluctuates, but its change is relatively stable and the swing amplitude is minimal. Undiminished pendulum control shows an irregular swing angle change and it is unable to stabilize at the desired position. ASMC-PD control has smaller swing fluctuations and smaller steady-state errors than PID control and MPC control.

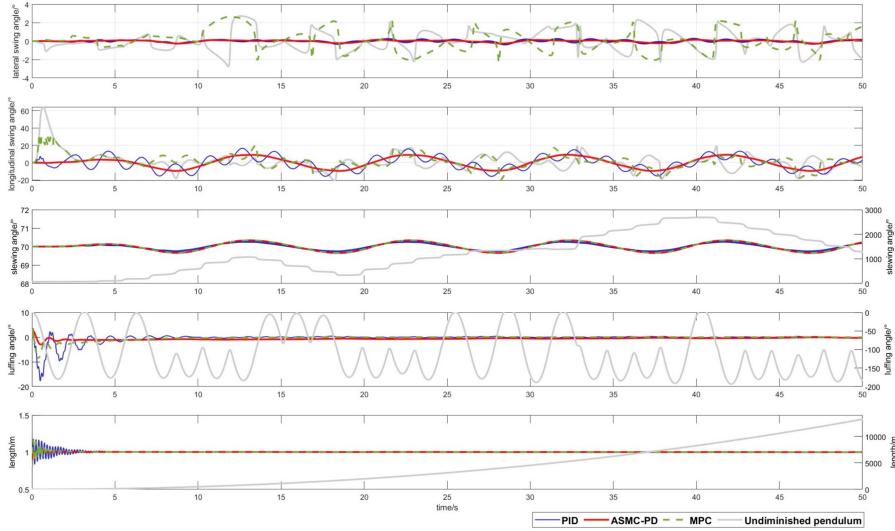


Fig. 2. Static environment comparison

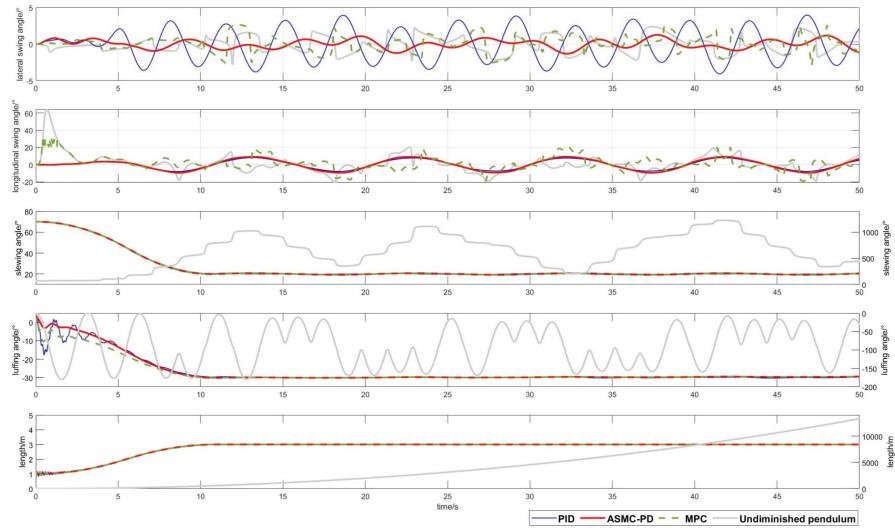


Fig. 3. Dynamic environment comparison

5 Conclusion

In this paper, an adaptive sliding mode and PD mode switching controller is proposed for a 5-DOF Marine crane under the influence of ship roll and pitch motion to eliminate buffeting and realize swing Angle reduction control by adaptive gain and switching mode. Among them, the slewing angle, luffing angle and rope length can reach and

stabilize the desired position in a limited time, and the swing Angle of the payload can also be suppressed. Detailed controller design steps and strict stability analysis are given. Finally, the static and dynamic simulation results under 4–6 sea state interference show that the stability and robustness of the controller are better than the PID control and MPC control.

References

1. Qian, Y., Hu, D., Chen, Y., Fang, Y., Hu, Y.: Adaptive neural network-based tracking control of Underactuated offshore ship-to-ship crane systems subject to unknown wave motions disturbances. *IEEE Trans Syst Man Cybern Syst.* **52**(6), 3626–3637 (2022)
2. Sun, N., Wu, Y., Chen, H., Fang, Y.: Antiswing cargo transportation of Underactuated tower crane systems by a nonlinear controller embedded with an integral term. *IEEE Trans. Automat. Sci. Eng.* **16**(3), 1387–1398 (2019)
3. Yang, T., Sun, N., Chen, H., Fang, Y.: Swing suppression and accurate positioning control for underactuated offshore crane systems suffering from disturbances. *IEEE/CAA J. Autom. Sinica.* **7**(3), 892–900 (2020)
4. Li, X., Dong, L., Sun, C.: Data-based optimal tracking of autonomous nonlinear switching systems. *IEEE/CAA J. Autom. Sin.* **8**, 227–238 (2021)
5. Sun, N., Fang, Y., Chen, H., Wu, Y., Lu, B.: Nonlinear Antiswing control of offshore cranes with unknown parameters and persistent ship-induced perturbations: theoretical Design and hardware experiments. *IEEE Trans. Ind. Electron.* **65**(3), 2629–2641 (2018)
6. Bandong, S., Kirana, R.C., Nazaruddin, Y.Y., Joelianto, E.: Optimal gantry crane PID controller based on LQR with prescribed degree of stability by means of GA, PSO, and SA. In: 2022 7th International Conference on Electric Vehicular Technology (ICEVT). IEEE, Bali (2022)
7. Zhao, Y., Gao, H.: Fuzzy-model-based control of an overhead crane with input delay and actuator saturation. *IEEE Trans. Fuzzy Syst.* **20**(1), 181–186 (2012)
8. Lu, B., Fang, Y.: Gain-adapting coupling control for a class of underactuated mechanical systems. *Automatica.* **125**, 109461 (2021)
9. Chen, H., Fang, Y., Sun, N.: An adaptive tracking control method with swing suppression for 4-DOF tower crane systems. *Mech. Syst. Signal Process.* **123**, 426–442 (2019)
10. Sun, N., Yang, T., Chen, H., Fang, Y., Qian, Y.: Adaptive anti-swing and positioning control for 4-DOF rotary cranes subject to uncertain/unknown parameters with hardware experiments. *IEEE Trans Syst Man Cybern Syst.* **49**(7), 1309–1321 (2019)
11. Uppal, A.A., Azam, M.R., Iqbal, J.: Sliding mode control in dynamic systems. *Electronics.* **12**(13), 2970 (2023)
12. Wang, S., Tao, L., Chen, Q., Na, J., Ren, X.: USDE-based sliding mode control for servo mechanisms with unknown system dynamics. *IEEE/ASME Trans. Mechatron.* **25**(2), 1056–1066 (2020)
13. Ding, S., Park, J.H., Chen, C.-C.: Second-order sliding mode controller design with output constraint. *Automatica.* **112**, 108704 (2020)
14. Eun, Y., Kim, J.H., Kim, K., Cho, D.I.: Discrete-time variable structure controller with a decoupled disturbance compensator and its application to a CNC servomechanism. *IEEE Trans. Control Syst. Technol.* **7**(4), 414–423 (1999)



Co-simulation and Validation Methodology for Unmanned Tracked Vehicle

Lunfei Liang^{1,2,3,4}, Huixin Huang^{2,3,4(✉)}, Geng Chen^{2,3,4}, Song Liu^{2,3,4},
and Yan Pan^{2,3,4}

¹ School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China

² Jianghuai Advance Technology Center, Hefei 230001, China

hhx15256092802@163.com

³ Anhui Provincial Key Laboratory of Humanoid Robots, Heifei 230088, China

⁴ Anhui Provincial Industry Innovation Center of HumanoidRobot, Heifei 230088, China

Abstract. The perception-planning performance of Unmanned Tracked Vehicle (UTV) serves as the foundation for executing complex and diverse missions. However, the intricate operational environment induces the invalid characterization of measured data. To address these challenges, this study proposes a systematic methodology structured as follows: First, a high-fidelity real-time co-simulation framework is established by integrating Unreal Engine 5 (UE5) virtual modeling with the Vortex dynamics engine. This framework enables closed-loop co-simulation of UTV-environment interactions, ROS-based algorithm-in-the-loop(AIL) control, and dynamic modeling. Second, Distributed interactive simulation communication is achieved via Data Distribution Service (DDS) middleware. Finally, experimental validation confirms the effectiveness of the high-fidelity simulation system in generating realistic data.

Keywords: Co-simulation · Unmanned Tracked Vehicle(UTV) · Perception-planning Testing

1 Introduction

With the rapid development of intelligent manufacturing and unmanned system technology, Unmanned Tracked Vehicle (UTV) have demonstrated exceptional terrain adaptability, load-bearing capacity, obstacle-crossing capability, and resilience in harsh environments. These attributes make them critical in national strategic domains: emergency rescue, geological hazard mitigation, and modern agriculture. During complex mission execution, the perception and planning capabilities of UTVs are pivotal for achieving operational flexibility and autonomy. By integrating real-time sensor data and geospatial information, UTVs dynamically optimize path planning based on mission objectives, environmental constraints, and safety considerations. Thus, robust perception-planning systems endow UTVs with the intelligence to navigate unstructured environments.

As UTVs are increasingly deployed in diverse scenarios, their operational efficacy hinges on algorithmic adaptability to dynamic interactions and decision-making reliability. However, physical field testing faces limitations due to high costs, safety risks,

and difficulties in replicating extreme scenarios, hindering efficient algorithm iteration. In this context, simulation technologies—leveraging digital twins of virtual environments and physical laws—have emerged as indispensable tools for validating algorithm robustness and exploring performance boundaries. Recent years have witnessed robotics research teams developing tailored simulation frameworks [1] and leveraging platforms like Gazebo (developed by the Open Source Robotics Foundation for the DARPA Robotics Challenge) [2], Webots [3], CoppeliaSim [4], and RobotDK [5] for performance evaluation. Notable advancements include: Yoon et al. integrated Robot Operating System (ROS)-based sensors with Gazebo’s physics engine to model robotic systems [6]. Yan et al. investigated multi-agent system (MAS) formation control under time-delayed discrete-time conditions via ROS-Gazebo co-simulation [7]. Kumar et al. proposed trajectory planning studies using Webots-MATLAB co-simulation [8]. Aslan et al. further validated the consistency between Webots-based 3D simulations and real-world testing [9].

While existing simulation platforms have proven valuable, the advent of Unreal Engine has exposed limitations in conventional simulators, particularly in environmental rendering and visualization. AirSim, leveraging Unreal Engine’s high-fidelity physics simulation and multimodal sensor emulation, has become a mainstream tool for unmanned system research. For instance, Shimada et al. utilized AirSim’s monocular camera-generated depth maps for drone obstacle avoidance analysis [10], while Zhu et al. developed swarm simulation models to evaluate Unmanned Aerial Vehicle reconnaissance strategies [11]. However, existing studies predominantly model instantaneous spatial states, failing to capture time-varying dynamics induced by multivariate uncertainties. Furthermore, traditional control models struggle to represent nonlinear, high-uncertainty environments as robotic perception data scales.

To address critical challenges in UTV perception-planning evaluation: (1) insufficient accuracy and scope in complex environment modeling; (2) absence of UTV-environment interaction simulation. The paper proposes two innovations:

1. Distributed real-time co-simulation: A Unreal Engine 5 (UE5)-dynamics-ROS framework integrated with Data Distribution Service (DDS) middleware for synchronized simulation.
2. Algorithm-in-the-loop (AIL) and hardware-in-the-loop (HIL) testing: Implementation of perception-planning algorithms for AIL/HIL validation.

These methodologies advance interpretability and credibility in UTV intelligence assessment, offering foundational theories and technical frameworks to enhance China’s competitiveness in autonomous systems.

2 High-Fidelity Co-simulation for UTV

To overcome limitations in UTV task data scarcity, multivariate environmental interference, and high real-world data acquisition costs, this study proposes: (1) Physics-based virtual environment: Development of a high-fidelity virtual testing environment with UTV dynamic models and embedded control algorithms. (2) Distributed data communication: Implementation of efficient DDS-based protocols for real-time data exchange

between virtual environments, controllers, and algorithms. (3) Closed-loop environment-control-feedback simulation: Integration of ROS-based control algorithms, perception-planning modules, and dynamic models to establish a UTV-environment-control feedback loop. The architecture is illustrated in Fig. 1.

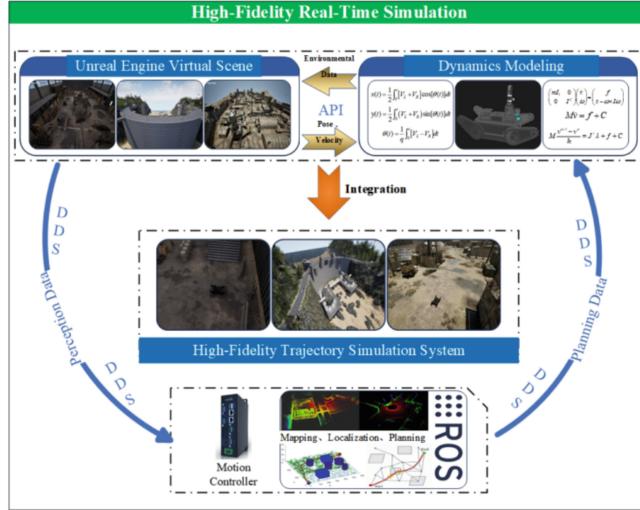


Fig. 1. High-Fidelity Real-Time Simulation Framework.

2.1 High-Fidelity Simulation Modeling

A simulation testing environment is constructed using Unreal Engine 5 (UE5) to achieve high-fidelity dynamics interactions between the UTV and its surroundings. Integration of the virtual environment with the dynamics model is accomplished through the Vortex dynamics engine, which ensures precise physics-based simulations. The perception-planning algorithms for the UTV are developed on the ROS, enabling real-time environmental awareness and decision-making. Communication between the UTV's onboard systems and algorithms is facilitated via DDS middleware, ensuring stable and efficient data exchange. The resulting high-fidelity real-time simulation system is illustrated in Fig. 2.

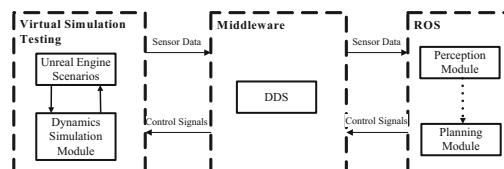


Fig. 2. High-Fidelity Real-Time Simulation System Architecture.

Leveraging UE5's advanced rendering capabilities, diverse virtual testing environments—including a small-scale indoor training environment, large-scale dam environment, and complex military camp environment—are constructed to evaluate UTV perception and planning performance under varying conditions. These environments are depicted in Fig. 3.



Fig. 3. UE5-Based Virtual Testing Environments.

The UTV mechanical structure is modeled in SolidWorks, with topology optimization and surface detailing performed in 3DS Max to meet engineering standards and real-time rendering requirements. A multibody dynamics model is developed using the Vortex physics engine, incorporating constraints, collision parameters, and terrain mechanics to balance computational accuracy and real-time performance. Bidirectional data communication between UE5 and Vortex is achieved via API interfaces, enabling real-time synchronization of terrain data, environmental parameters, and dynamics feedback. The integrated UTV dynamics model and virtual environment framework are detailed in Fig. 4.

Environmental and pose data are acquired from UE5 and Vortex, with a focus on simulating the UTV's LiDAR sensor. LiDAR pulses are emulated using ray-casting techniques, where collisions with surfaces provide direct distance measurements. Key LiDAR parameters are summarized in Table 1.

The virtual LiDAR implementation in UE5 is visualized in Fig. 5.

2.2 Real-Time Communication and Control

Distributed real-time co-simulation demands robust, low-latency communication between components. The communication is enabled by Data Distribution Service (DDS) middleware, which is founded on a data-centric publish-subscribe (DCPS) model. DDS provides configurable Quality of Service (QoS) policies (e.g., reliability, deadlines) enabling deterministic data exchange and automatic discovery. This fundamentally addresses the stringent requirements for synchronized, scalable interaction within the closed-loop UTV simulation and validation framework.

A DDS-based publish/subscriber architecture is adopted to enable distributed real-time communication, emphasizing data-centricity and robust QoS policies. This framework supports large-scale data exchange between components, integrating ROS with the high-fidelity simulation system. DDS communication modules—comprising discovery

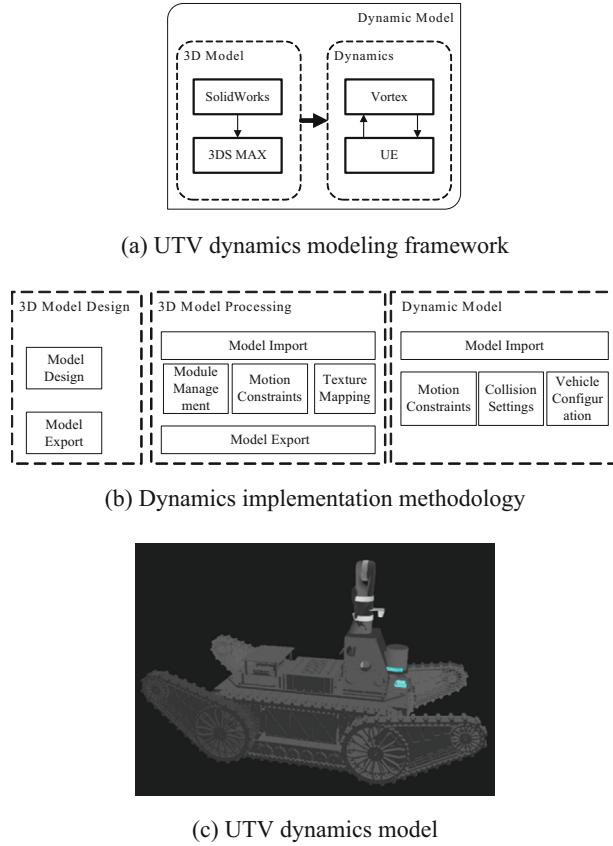


Fig. 4. UTV Dynamics Model Development Framework.

Table 1. LiDAR Specifications.

Parameter	Value
Number of channels	32
Measurement range	0.05–200 m
Horizontal FOV(Field of View)	360°
Horizontal resolution	0.2°
Vertical FOV	40° (−25° to 15°)
Transmission frequency	10 Hz

servers, publishers, subscribers, data handlers, and logging units—are embedded into the simulation system via plugins, while a standalone DDS node is implemented in ROS. The architecture is illustrated in Fig. 6.

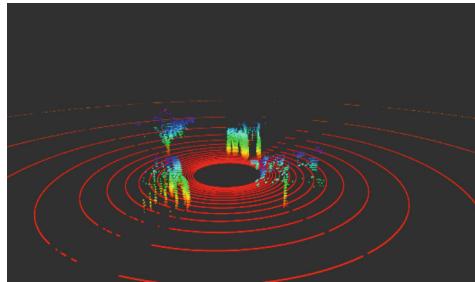


Fig. 5. Virtual LiDAR Sensor Simulation.

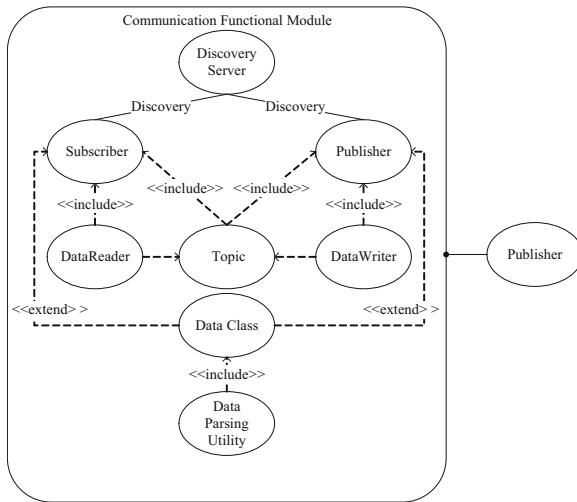


Fig. 6. DDS Communication Module Architecture.

With the simulation framework established, ROS-based control algorithms are deployed for UTV perception and planning. Real-time data interaction between the simulation system and UTV controllers enables Algorithm-in-the-Loop (AIL) and Hardware-in-the-Loop (HIL) testing.

3 Experimental Validation

3.1 Test Environment Configuration

The hardware setup for the high-fidelity simulation system, UTV motion controller, and perception-planning modules is summarized in Table 2.

A local area network (LAN) ensures real-time communication between components. System readiness, including DDS node activation and client-server synchronization.

Table 2. Hardware Configuration

Component	Specifications	OS/Platform
High-Fidelity Simulation	CPU: Intel i9-14900K; GPU: RTX 4090; 128GB DDR5	Windows 11
UTV Motion Controller	CPU: Intel i7-1165G7; GPU: RTX 2060; 32GB DDR4	Ubuntu 20.04 (ROS)
Perception/Planning Module	CPU: Intel i7-13700F; GPU: RTX 3060 Ti; 32GB DDR5	Ubuntu 20.04
Network Switch	Ruijie RG-NBS3100	–

3.2 Simulation Results Analysis

To validate the simulation system's realism, real-world data were collected from Hefei Zhongan Chuanggu Phase II (Fig. 7), a complex urban-terrain site with diverse obstacles. The virtual replica retained key topological features (e.g., building layouts, vegetation) to ensure comparability.

First, This study collected real-world LiDAR and RTK data, then processed them using the Automatic LiDAR Odometry and Mapping(A-LOAM) algorithm to achieve localization.



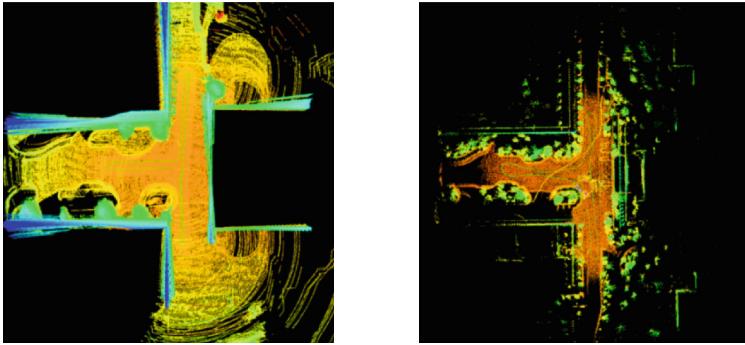
Fig. 7. Real-World Environment: Zhongan Chuanggu Phase II.

A virtual replica of this environment was constructed in the simulation system (Fig. 8).



Fig. 8. Virtual Environment Reconstruction.

Second, The simulation server executes control commands, generates synthetic LiDAR data, and runs A-LOAM to produce localization results. The point cloud datasets from both the high-fidelity real-time simulation system and the real-world environment under identical scenarios were processed using the A-LOAM algorithm, with their operational results demonstrated in Fig. 9. Trajectory data, sensor outputs, and system logs are stored for post-processing.



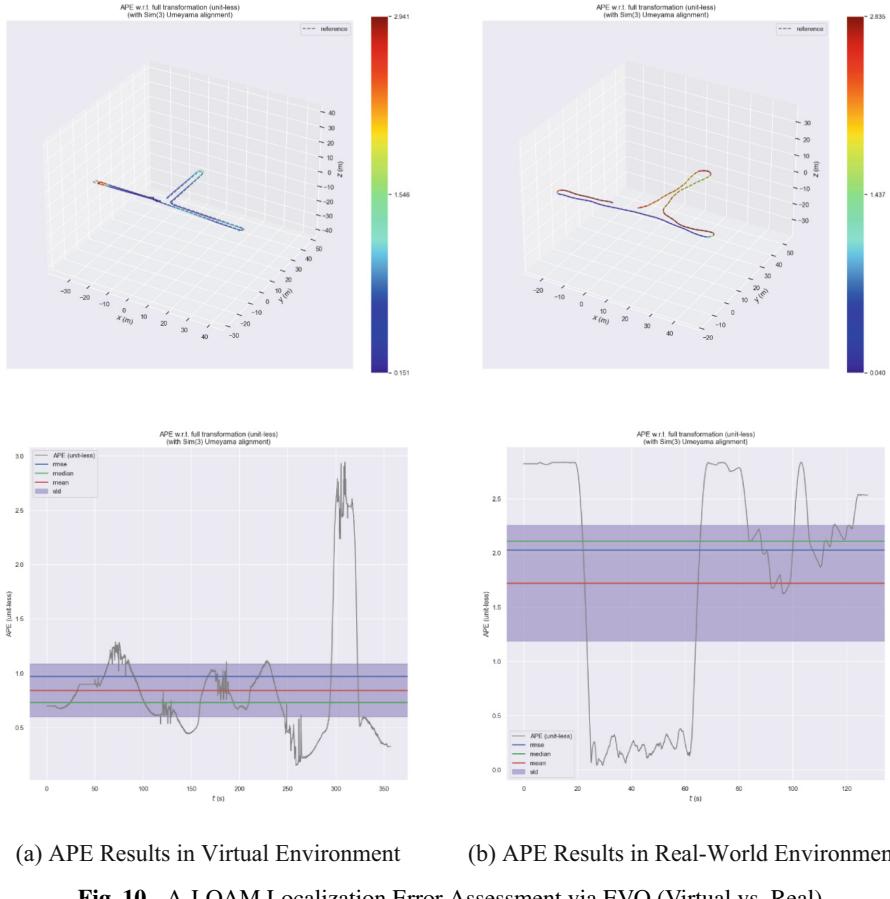
(a) Results under high fidelity real-time simulation system (b) Results under real environment

Fig. 9. A-LOAM Localization Results (Virtual vs. Real).

Third, the Absolute Pose Error (APE) was evaluated using the EVO tool to compare the high-fidelity simulation system with real-world conditions. APE quantifies the direct deviation between estimated poses and ground truth, providing an intuitive measure of algorithmic accuracy and global trajectory consistency. The results are presented in Fig. 10. Key metrics include:

APE results in virtual environment: Maximum error = 2.94 m, Minimum error = 0.15 m, Standard deviation = 0.48, Root Mean Square Error (RMSE) = 0.973.

APE results in real-world environment: Maximum error = 2.83 m, Minimum error = 0.04 m, Standard deviation = 1.06, Root Mean Square Error (RMSE) = 2.025.



(a) APE Results in Virtual Environment

(b) APE Results in Real-World Environment

Fig. 10. A-LOAM Localization Error Assessment via EVO (Virtual vs. Real).

To validate the high-fidelity simulation system's effectiveness, this study employed the A-LOAM algorithm to process point cloud data from both real and simulated environments, with APE assessed using EVO. Analysis of the ATE results revealed that while the maximum error in the virtual environment (2.94 m) was slightly higher than that in the real environment (2.83 m), it encompassed the extreme error observed in reality (3.7%), aligning with a conservative engineering approach. Crucially, the standard deviation (0.48) was significantly lower than that of the real environment (1.06), and the RMSE (0.973) was also lower than the real environment's 2.025. These findings demonstrate that the virtual environment not only exhibits more stable localization errors but also achieves a high degree of consistency with the real environment in terms of overall error levels. Consequently, the simulation system demonstrates highly comparable error metrics to the real-world environment, confirming its effectiveness in facilitating algorithm verification and scenario simulation.

4 Conclusions and Future Work

This study proposes a high-fidelity co-simulation framework integrating UE5, multi-body dynamics (Vortex), and ROS to address challenges in UTV perception-planning evaluation. Key contributions include:

Dynamic Coupling: High-fidelity modeling of UTV-environment interactions via DDS-based real-time communication.

Virtual-Real Consistency Verification: By applying the A-LOAM algorithm to process synchronized point cloud data from both simulated and real-world environments under identical conditions, using the EVO toolkit demonstrated a maximum localization error deviation of 3.7% between virtual and real trajectories, conclusively proving the simulation system's reliability in replicating real-world.

The methodology enhances testing efficiency and accuracy, providing a cost-effective solution for emergency response applications. Future work will integrate deep learning for dynamic environment prediction and expand the platform into a digital twin for lifecycle performance monitoring.

References

1. Krotkov, E., et al.: The DARPA robotics challenge finals: results and perspectives. *J. Field Robot.* **34**(2), 229–240 (2017)
2. Koenig, N., Howard, A.: Design and use paradigms for gazebo, an open source multi robot simulator. In: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), vol. 3, pp. 2149–2154. IEEE (2004)
3. Michel, O., Webots, T.M.: Professional mobile robot simulation. *J. Adv. Robot. Syst.* **1**(1), 3942 (2004)
4. Rohmer, E., Singh, S.P.N., Freese, M.: V-REP: a versatile and scalable robot simulation framework. *IEEE/RSJ Int. Conf. Intell. Robots Syst.* **2013**, 1321–1326 (2013)
5. Montaqim, A.: Offline prog ramming software for industrial robots from RobotDK offers hundreds of virtual industrial robots from top robotics companies. *Robot. Autom. News.* (2015)
6. Yoon, D., et al.: Research on the robot platform construction method based on the simulation. In: 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), pp. 413–416 (2019)
7. Yan, Z.W., et al.: Event-triggered formation control for time delayed discrete time multi-agent system applied to multi-UAV formation flying. *J. Franklin Inst.* **360**(5), 3677–3699 (2023)
8. Kumar, S., Parhi, D.R.: Multi target trajectory planning and control technique for autonomous navigation of multiple robots. *ISA Trans.*, 15–26 (2023)
9. Aslan, E., Arserim, M.A.: Development of push-recovery control system for humanoid robots using deep reinforcement learning. *Ain Shams Eng. J.*, 102167 (2023)
10. Shimada, T., et al.: Pix2Pixbased monocular depth estimation for drones with optical flow on AirSim. *Sensors.* **22**(6), 2097 (2022)
11. Zhu, B.S., et al.: Simulation and analysis of drone swarm reconnaissance and search based on AirSim platform. *J. Comput. Appl.* **41**(S1), 196–201 (2021)



6-DoF Object Pose Estimation via Improved Attitude Optimization Network

Yue Wang, Kang An^(✉), and Hui Zhang^(✉)

These authors are with the College of Information, Mechanical and Electrical Engineering,
Shanghai Normal University, Shanghai, China
1000549399@mail.shnu.edu.cn, {ankang, huiz}@shnu.edu.cn

Abstract. Achieving accurate and efficient object pose estimation is a highly valuable research topic today. However, static object pose estimation is no longer sufficient for most applications, and achieving accurate pose estimation in dynamic, unstructured scenes remains a significant challenge. Most existing methods are limited in real-world scenarios due to reliance on 3D models, closed-category detection, environmental noise, and densely labeled views. In addition, template-based methods require a large number of views for retrieval, and the refinement process is insufficiently sensitive to rotational errors. We propose an object pose detection method that uses sparse 2D images and a multi-scale attention mechanism to improve pose accuracy, with a pose refiner to correct rotational errors through dense correspondences. Additionally, we performed real-time 6-DoF pose estimation in a real-world environment, and the algorithm outperforms Gen6D. Compared to Gen6D, our method shows improvements of 4.63% and 1.74% on the LINEMOD dataset under ADD and Proj-2D metrics, respectively. On LINEMOD-Oc, the improvement is 0.15%. In terms of the median rotation angle, the model achieves 24.53°, which represents a reduction of 20.32° compared to Gen6D.

Keywords: Dynamic Scene Pose Estimation · Multi-scale Attention Mechanism · Dense Correspondence

1 Introduction

In recent years, dexterous grasping technology has been widely applied across various fields, including industrial manufacturing [1], social services [2, 3], precision medicine [4, 5], aerospace [6, 7], and more. However, detecting the pose of objects not seen during training remains a challenging task due to the lack of prior knowledge and the potential for unseen variations in object appearance and environmental conditions. Describing an object's state, particularly its six degrees of freedom (6DoF) pose, is essential for enabling fine-grained interactions in complex environments. Traditional instance-level [8–12] or category-level [13, 14] pose estimators have limitations in handling diverse objects, as they are designed for specific instances or categories. These methods, which rely on CAD models or well-defined canonical spaces for categories, restrict their ability to generalize to unseen instances or categories during testing. In summary, our main contributions are as follows:

- We propose a fusion method of multi-scale attention and zero convolution for 6-DoF pose estimation, aimed at segmenting feature regions.
- We designed a pose refinement network that effectively corrects errors in pose estimation.
- We conducted an in-depth study from the perspectives of pose translation error and rotation error, the results show that our method achieves significant improvements in both aspects.

2 Related Work

CosyPose [9], proposed by Facebook AI Research, has a key limitation: it relies on specific objects or categories during training and requires a pre-existing CAD model for pose estimation. Acquiring accurate CAD models is challenging due to the time and resources required for manual modeling, and model discrepancies can lead to pose estimation errors. Some methods, like LatentFusion [19], can generalize to new objects but still depend on high-quality target 3D models or depth maps. LatentFusion combines neural networks with traditional 3D pose estimation techniques, using RGB images and depth maps to better handle 3D shape. However, like CosyPose, it cannot estimate poses for unknown objects without corresponding models or representations. OnePose [20] and OnePose++ [21] use sparse point clouds for pose estimation but need dense views. We propose a method that first estimates a rough pose and refines it using a network inspired by Trabelsi [22]. The challenge of refining the relative pose is addressed with correspondence-based and direct pose regression methods, where pixel-level correspondences are established to solve the fundamental matrix [23]. For example, DeepVO [24] combines CNNs and RNNs for visual odometry, analyzing relative pose via optical flow. PVNet [11] is robust to occlusion and complex backgrounds, but its accuracy can drop with uneven keypoint distribution or weak texture. Direct pose regression methods include SA6D [25], which uses self-attention to capture long-range dependencies for improved accuracy in complex environments, and Gen6D [26], which refines pose estimates using dense reference views and 3D neural networks. POPE [27] enhances pose estimation through feature extraction and matching. Inspired by these works, we propose a correspondence-based method with a coarse-to-fine paradigm.

3 Proposed Method

3.1 Network Architecture

Our method consists of two main components: a pose proposal module and a pose refinement module. The pose proposal module outputs object classification and initial pose estimation from the RGB input. The pose refinement module employs a renderer and an iterative refiner based on multi-scale attention. The renderer initializes the object crop using the initial pose and sparse viewpoint images. During refinement, query image features are back-projected into a 3D volume, and the 6 closest reference images are selected, with their features projected and concatenated. Inaccuracies in the initial pose cause inconsistencies, which are addressed by a 3D CNN to regress translation residuals. For rotational residuals, the reference image with the highest similarity is used to establish dense correspondences and regress the rotational residual (Fig. 1).

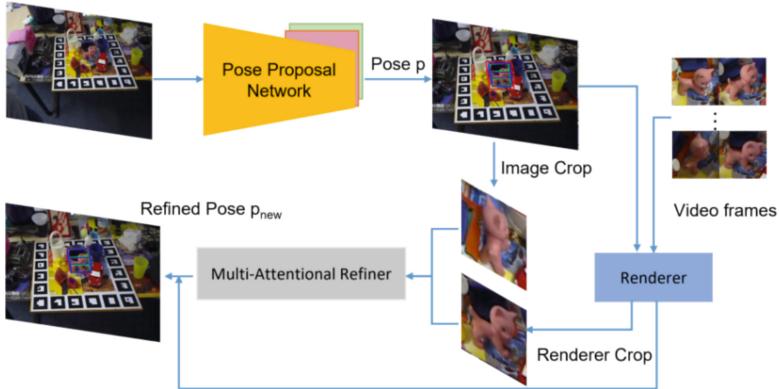


Fig. 1. The pose estimation network estimates a rough pose (red box) and refines it (blue box) by regressing residuals, with gaps corrected using multi-scale attention and dense feature analysis.

In target detection, large-scale objects are easier to detect due to richer features, while small-scale objects are more challenging. To enhance multi-scale detection, we reduce the down sampling rate and use hybrid null convolution. The first layer uses a void rate of 2 to expand the input area, the second layer uses a void rate of 4 to enlarge the receptive field, and a 1×1 standard convolution kernel refines it to a fine-grained localized receptive field, preserving edge information. For distinguishing similar objects in complex environments, the model calculates similarity scores and selects the part with the highest score. As shown in Fig. 2, the purple object, with less occlusion, is selected for segmentation due to its higher similarity to the template.



Fig. 2. The blue box in the first image indicates the region of interest segmented by the attention mechanism, the blue box in the second image shows the correct object pose, and the red box represents the initial pose from rough template matching. After refinement in the third image, the red box is much closer to the correct pose.

3.2 Attitude Optimization Network

We obtain an initial rough object pose by combining the translation estimated by the object detector and the rotation estimated by the viewpoint selector. This initial pose is then refined using a volume-based 3D pose refinement module. During feature extraction,

we apply various normalization layers, with instance normalization used to process the input data and ensure consistency across different feature layers. The outputs from different feature layers are then integrated through a feature fusion operation to create multi-scale feature representations. The feature extraction process can be expressed as follows:

$$F_i = \text{Conv}(F_i - 1) + \text{Norm}(F_i - 1) \quad (1)$$

where F_i represents the output feature map of the i_{th} layer. Using the pose optimization network, we compute the object's rotation quaternion, translation vector, and scale parameters, representing the object's pose information. The predicted values for rotation, translation, and scale are given by:

$$r = \frac{FC1(x)}{\|FC1(x)\|}, \quad t = \text{ReLU}(FC2(x)), \quad s = \text{sigmoid}(FC3(x)) \quad (2)$$

where r , t and s are the predicted values for rotation, translation, and scale, respectively, and x is the object feature. The formula introduces activation functions such as ReLU and sigmoid, and normalizes the rotation quaternion to ensure its validity. This approach helps prevent overfitting, allowing the network to learn smoother and more robust parameters for rotation, translation, and scaling.

By establishing dense correspondences between the best matched views, we then estimate the relative pose of the camera [28]. It is worth noting that our method can accurately recover the relative rotation. However, the predicted translation is up-to-scale, similar to other relative pose estimators [26, 29]. This limitation arises from unwanted rotational interference caused by background clutter. Two-view image matching often fails on complex targets due to noisy matches. By finding the highest similarity score in the matrix, we retrieve corresponding images from two views. This process can be extended to $M \times K$ for multiple prompt images. We introduce local descriptors to capture key features such as edges, corners, and textures, which enhance the global representations. We select matches with confidence scores exceeding the threshold and count them among the total matches n . This criterion is defined as:

$$\text{Criteria} = \frac{1}{n} \sum_{i=1}^N \alpha_i \cdot \beta(b_i \geq \sigma) \quad (3)$$

Where b_i denotes the confidence score of the i_{th} match, and β is the indicator function that returns 1 if its argument is true, and 0 otherwise. By introducing α , we can weight the confidence of each match, giving higher weight to matches with higher confidence in the final score. The proposal with the largest criteria score among the Top-K proposals is selected as the best matched pair, providing a more reliable estimation of the object pose. The visualization process is shown in Fig. 3.

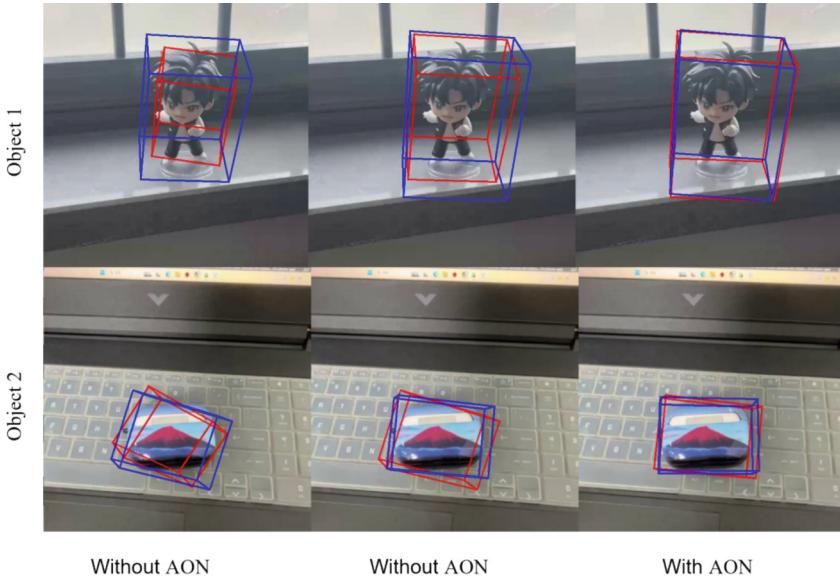


Fig. 3. The blue box represents the correct object pose, and the red box represents the estimated pose. Without the AON (Attitude Optimization Network), there is a significant rotational error. After correction by the AON, the estimated pose is much closer to the correct one.

4 Result Analysis

The refiner is trained on images from the LINEMOD, GenMOP, ShapeNet, and Google Scanning Objects datasets, and evaluated on LINEMOD and GenMOP. The LINEMOD dataset includes 1,578 images of 13 weakly textured objects, while GenMOP consists of 10 objects (e.g., “scissor” and “chair”) captured in two video sequences under varying conditions, with each sequence containing 200 images.

The region proposal module combines 2D image features with geometric features from the induced point cloud model to improve robustness against clutter and occlusion, compared to Gen6D. The region of interest (ROI) identifies the target object’s location. While Gen6D predicts the ROI of novel objects, it struggles with occlusion, relying only on visual similarity between reference and test images (Table 1).

Table 1. evaluation of add-0.1d on linemod-Occ datasets against category-agnostic baselines.

Ref. Num	Ref. Num	eggbox	duck	glue	ape	Avg.
Gen6D [26]	20	0.02	0.07	0.03	0.12	0.06
SA6D [25]	20	0.08	0.09	0.04	0.10	0.07
LF [19]	20	0.01	0.00	0.18	0.45	0.16
Ours	20	0.11	0.18	0.24	0.32	0.21

We compare the proposed multimodal feature fusion method with others on the LINEMOD dataset using ADD(S) and Proj-2D metrics. Despite weak textures in most objects, the cavity convolution improves global feature learning (Table 2).

Table 2. Test in linemod, most objects have weak textures. the multi-scale attention mechanism enhances global features, helping the model learn pose by referencing the reference image.

Metrics	Method	Cat	Duck	Lamp	Driller	Avg.
Add	Brachmann [30]	42.70	30.20	67.00	61.90	50.45
	OnePose [20]	45.81	33.99	85.32	71.26	59.09
	Gen6D [26]	60.68	41.97	89.83	67.39	64.96
	Ours	64.57	43.60	91.12	79.09	69.59
Metrics	Method	Cat	Duck	Lamp	Driller	Avg
Proj-2D	Brachmann [30]	42.70	30.20	67.00	61.90	50.45
	OnePose [20]	75.64	71.17	87.52	75.02	77.33
	Gen6D [26]	95.81	80.85	91.17	73.34	85.29
	Ours	95.19	88.03	93.04	71.89	87.03

We conducted tests on the LINEMOD dataset, evaluating the average median error and accuracy at 30°and 15°, which demonstrate the model’s advantage in rotation correction (Table 3).

Table 3. Test on the LINEMOD dataset, measuring the median error and accuracy at 30°and 15°averages. We also reported the accuracy at 30°broken down by category.

Dataset	Method	All Categories			Per Category(Med.Err↓)			
		Med.Err(↓)	Acc30(↑)	Acc15(↑)	Eggbox	Can	Iron	Camera
LINEMOD	Gen6D	44.85	0.36	0.09	31.78	30.40	30.09	35.97
	LoFTR	33.03	0.56	0.32	16.88	17.58	17.90	22.55
	Ours	24.53	0.67	0.39	12.14	17.92	15.23	19.03

5 Limitations and Future Work

While the proposed method exhibits strong performance in both accuracy and real-time capability, certain limitations remain. Specifically, the framework can be sensitive to severe occlusions and extreme lighting conditions, which may degrade feature quality and introduce pose ambiguity. Furthermore, although the integration of multi-scale attention and zero convolution enhances feature discrimination, the model’s generalization ability in highly cluttered or previously unseen environments remains limited without domain adaptation. In future work, we aim to improve the system’s robustness by

incorporating more advanced cross-modal attention mechanisms and uncertainty-aware modeling. In addition, extending the framework to handle deformable and articulated objects is a promising direction to further broaden its applicability.

6 Conclusion

In this paper, we presented a novel 6-DoF object pose estimation framework designed to address the challenges of dynamic and unstructured environments. By leveraging sparse 2D image inputs and integrating a multi-scale attention mechanism with zero convolution, our approach enhances the segmentation of discriminative feature regions. Furthermore, a dedicated pose refinement module is introduced to compensate for rotational inaccuracies by establishing dense and geometrically consistent cross-view feature correspondences, thereby improving the precision and robustness of pose estimation. Extensive experiments demonstrate that our method significantly outperforms existing approaches in both translation and rotation accuracy on benchmark datasets. In addition, real-world experiments confirm the framework's effectiveness in achieving accurate and real-time 6-DoF pose estimation. These results highlight the practical value and strong generalization ability of our approach, making it a promising solution for pose estimation in complex, real-world scenarios.

References

1. Kong, S., et al.: Review of application prospect of deep reinforcement learning in intelligent manufacturing. *Compt. Eng. Appl.* **57**(02), 49–59 (2021)
2. Garcia-Haro, J.M., Oña, E.D., Hernandez-Vicen, J., Martinez, S., Balaguer, C.: Service robots in catering applications: a review and future challenges. *Electronics* **10**, 47 (2021)
3. Huang, S., ZhenshengWu, Z.R., Liu, H., Gui, Y.: Review of research on intelligent power inspection robots. *Electr. Measur. Instrument.* **57**(02), 26–38 (2020). <https://doi.org/10.19753/j.issn1001-1390.2020.002.005>
4. Yan, L., Wang, Q., Li, H., Li, D., Xia, J.: Review of continue surgical robot actuated by SMA. *J. Mech. Eng.* **57**(11), 138–152 (2021)
5. Taylor, R.H.: Medical robotics and computer-integrated surgery. In: 2008 32nd Annual IEEE International Computer Software and Applications Conference, vol. 2008, pp. 1–1, Turku. <https://doi.org/10.1109/COMPSAC.2008.234>
6. Liu, H., Liu, D., Jiang, Z.: Space manipulator technology: review and prospect. *Acta Aeronautica et Astronautica Sinica* **42**(1), 524164–524164 (2021)
7. Sun, J., Sun, Z., Xin, P., Liu, B., Wei, Q., Yan, C.: Review on development of legged robots for deep space landing exploration. *China Mech. Eng.* **32**(15), 1765–1775 (2021)
8. Kehl, W., Manhardt, F., Tombari, F., Ilic, S., Navab, N.: Ssd-6d: making rgb-based 3d detection and 6d pose estimation great again. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1521–1529 (2017)
9. Labb  , Y., Carpentier, J., Aubry, M., Sivic, J.: Cospose: consistent multi-view multi-object 6d pose estimation. In: European Conference on Computer Vision, pp. 574–591. Springer International Publishing, Cham (2020)
10. Li, Z., Wang, G., Ji, X.: Cdpn: coordinates-based disentangled pose network for real-time rgb-based 6-dof object pose estimation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 7678–7687 (2019)

11. Peng, S., Liu, Y., Huang, Q., Zhou, X., Bao, H.: Pvnet: pixel-wise voting network for 6dof pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4561–4570 (2019)
12. Tekin, B., Sinha, S.N., Fua, P.: Real-time seamless single shot 6d object pose prediction. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 292–301 (2018)
13. Ahmadyan, A., Zhang, L., Ablavatski, A., Wei, J., Grundmann, M.: Objectron: a large scale dataset of object centric videos in the wild with pose annotations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7822–7831 (2021)
14. He, W., Sridhar, S., Huang, J., Valentin, J., Song, S., Guibas, L.J.: Normalized object coordinate space for category-level 6d object pose and size estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2642–2651 (2019)
15. Kehl, W., Milletari, F., Tombari, F., Ilic, S., Navab, N.: Deep learning of local RGB-D patches for 3D object detection and 6D pose estimation. In: European Conference on Computer Vision, pp. 205–220 (2016)
16. Wohlhart, P., Lepetit, V.: Learning descriptors for object recognition and 3D pose estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3109–3118 (2015)
17. Tejani, A., Tang, D., Kouskouridas, R., Kim, T.K.: Latent-class hough forests for 3d object detection and pose estimation. In: European Conference on Computer Vision, pp. 462–477. Springer International Publishing, Cham (2014)
18. Liu, Y., et al.: Regression-based three-dimensional pose estimation for textureless objects. *IEEE Trans. Multimedia*. **21**(11), 2776–2789 (2019)
19. Park, K., Mousavian, A., Xiang, Y., Fox, D.: Latentfusion: end-to-end differentiable reconstruction and rendering for unseen object pose estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10710–10719 (2020)
20. Sun, J., et al.: Onepose: one-shot object pose estimation without cad models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022)
21. He, X., Sun, J., Wang, Y., Huang, D., Bao, H., Zhou, X.: Onepose++: Keypoint-free one-shot object pose estimation without CAD models. *Adv. Neural Inf. Proces. Syst.* **35**, 35103–35115 (2022)
22. Trabelsi, A., et al.: A pose proposal and refinement network for better 6D object pose estimation. In: Workshop on Applications of Computer Vision. IEEE (2021)
23. Trabelsi, A., Chaabane, M., Blanchard, N., Beveridge, R.: A pose proposal and refinement network for better 6d object pose estimation. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 2382–2391 (2021)
24. Wang, S., Clark, R., Wen, H., Trigoni, N.: DeepVO: towards end-to-end visual odometry with deep recurrent convolutional Neural networks. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 2043–2050, Singapore (2017). <https://doi.org/10.1109/ICRA.2017.7989236>
25. Gao, N., Ngo, V.A., Ziesche, H., Neumann, G.: Sa6d: self-adaptive few-shot 6d pose estimator for novel and occluded objects. In: 7th Annual Conference on Robot Learning (2023)
26. Liu, Y., et al.: Gen6d: generalizable model-free 6-dof object pose estimation from rgb images. In: European Conference on Computer Vision. Springer Nature Switzerland, Cham (2022)
27. Fan, Z., et al.: POPE: 6-DoF Promptable pose estimation of any object, in any scene, with one reference. In: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 7771–7781 (2023)
28. Nist, D.: An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(6), 756–770 (2004)

29. Goodwin, W., Vaze, S., Havoutis, I., Posner, I.: Zero-shot category-level object pose estimation. In: Computer Vision ECCV 2022: 17th European Conference, Tel Aviv, pp. 516–532. Springer (2022)
30. Brachmann, E., Michel, F., Krull, A., Yang, M.Y., Gumhold, S.: Uncertainty-driven 6d pose estimation of objects and scenes from a single rgb image. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3364–3372 (2016)



The Design and Kinematic Analysis of a Seven-DOF Electric Operation Robot

Ding Ma¹, Haibo Du^{1(✉)}, Xunhong Sun¹, Wenwu Zhu^{1,2}, Runchen Zhao¹, and Yuchen Li¹

¹ School of Electrical Engineering and Automation, Hefei University of Technology, Hefei 230009, People's Republic of China
haibo.du@hfut.edu.cn

² Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing 210096, People's Republic of China

Abstract. In order to enhance the intelligence of substation power cabinet inspection, this paper investigates a 7-DOF hybrid emergency robot integrating a lift mechanism and a six-axis robotic arm. The robot's structure, control system, and kinematic model are presented, with an emphasis on the improved Denavit-Hartenberg (D-H) method used to establish its forward and inverse kinematics. Simulations in MATLAB validate the kinematic model and demonstrate the advantages of the 7-DOF configuration through workspace comparisons. Motion planning based on real operation scenarios further verifies task feasibility. The results provide a theoretical foundation for intelligent control and improve the efficiency, safety, and adaptability of substation inspection tasks.

Keywords: substation operation · hybrid robot · D-H method · kinematic analysis

1 Introduction

With the advancement of technology and increasing demands, mobile robots are increasingly used in industrial tasks such as transportation, inspection, and operations [1–4]. In substations, stable operation and fault handling of power cabinets are crucial to ensure continuous and reliable power supply [5,6]. Existing substation robots, typically equipped with mobile platforms and cameras, can perform autonomous inspections [7]. However, most are limited to monitoring functions and cannot perform physical operations on electrical cabinets. This paper introduces a 7-DOF hybrid robot system capable of performing cabinet operations, addressing the challenges of mounting a manipulator on a mobile base as opposed to fixed-base industrial setups [8].

To realize integrated inspection and operation in substations, this paper proposes a hybrid robot combining a mobile chassis, lift mechanism, and six-axis

arm. Its kinematic model, based on an improved D-H method, is verified through simulation to support accurate motion control and task execution.

2 Structural Design of the Seven-DOF Hybrid Robot

The mechanical structure and prototype of emergency operation hybrid robot are shown in Fig. 1. The robot consists of three main parts: an AGV mobile chassis, a lifting mechanism, and a six-DOF robotic arm.



Fig. 1. The electric operation hybrid robot

2.1 Mobile Chassis Design for the Hybrid Robot

The hybrid robot adopts a six-wheel layout of two driving wheels and four driven wheels, as shown in Fig. 2, and automatically adjusts the motion Angle through differential driving.

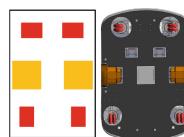
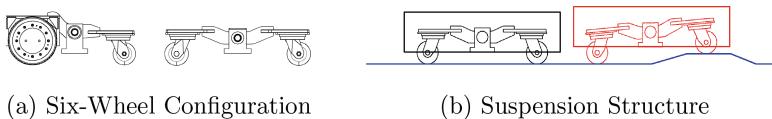


Fig. 2. Six-wheel configuration

The suspension system includes three pairs: two for the drive and passive wheels, and one for the dual passive wheels as shown in Fig. 3-a. Wheels are mounted on pivot axes to conform to uneven terrain and stabilize the chassis. As shown in Fig. 3-b.



(a) Six-Wheel Configuration

(b) Suspension Structure

Fig. 3. Chassis layout and suspension mechanism of the wheel system.

2.2 Upper Structure Design

Considering the special operational requirements at the bottom and top of electrical cabinets in substations within the power system environment, the composite robot studied in this paper, compared to traditional inspection robots, innovatively incorporates a lift mechanism and a six-axis robotic arm. This design meets the seven-degree-of-freedom operational requirement with the chassis remaining stationary during inspection tasks. The lower part features a servo electric cylinder with a 300 mm stroke, which achieves stable and safe position control of the lift mechanism through a single low-voltage DC servo motor and limit switches. The upper part uses a custom-developed six-axis collaborative robotic arm with a maximum payload of 5 kg and a maximum reach of 912 mm, which offers high sensitivity and fast response, meeting the needs for precise positioning and inspection of electrical cabinets.

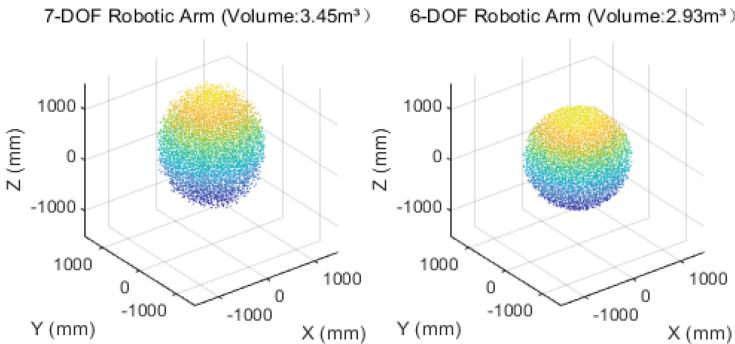


Fig. 4. 7-DOF vs 6-DOF robotic arm workspace comparison

2.3 Hybrid Robot Control Architecture

In order to realize the emergency operation function, the composite robot must have the ability of environment recognition, decision planning and environment adaptability. Therefore, the control architecture design of the composite robot is divided into three layers: perception layer, decision layer and control layer, as shown in Fig. 5. The decision control layer takes controller RK3588 and stm32 as the decision core, and realizes the environment identification, autonomous navigation, action planning and communication coordination through the decision control layer [9]. As the core of the robot action control layer, the robot drive control module is responsible for controlling the action to realize the specific action of the actuator. The sensing layer is mainly composed of the robot sensor module and its communication bus module, the core of which is the information processing computer, which processes the information of all sensor modules through the communication bus [10].

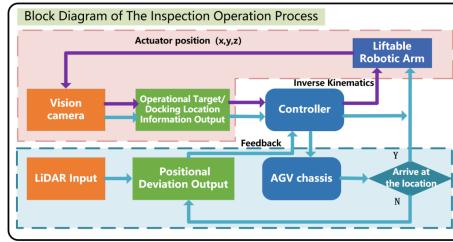


Fig. 5. Hybrid robot control architecture

2.4 Decomposition of Operation Tasks

The emergency operation hybrid robot can be modularly divided into three parts: the AGV chassis, the lifting mechanism, and the robotic arm. The AGV chassis is referred to as the lower structure, while the lifting mechanism and robotic arm constitute the upper structure. In a power inspection environment, electrical cabinets are typically located along the inspection route and equipped with information panels, monitoring lights, and operation buttons.

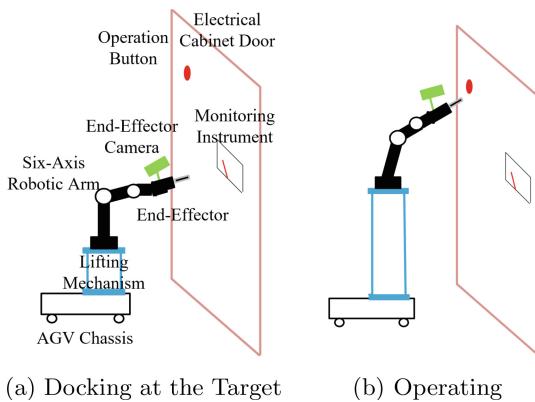


Fig. 6. Inspection and operation process of the hybrid robot

In autonomous control mode, the hybrid robot inspects cabinets along a predefined route. Upon reaching a target point, the robot first monitors the cabinet's status indicators, as shown in Fig. 6-a. After monitoring is completed, the robot waits for an operation command. For instance, in the operation of the red button on the cabinet shown in Fig. 6-b, the lifting mechanism raises the robotic arm to the appropriate position. Then, the vision camera identifies and positions the button, and the robotic arm's end-effector presses the button. This process involves solving the inverse kinematics for the specified target positions of each joint.

3 Kinematic Analysis and Simulation of the Seven-DOF Hybrid Robot

3.1 Establishment of the Robot's D-H Model

Based on the previous analysis of the seven-DOF hybrid robot's motion structure, the chassis remains stationary during operations on the electrical cabinet. Therefore, the kinematic analysis of the robot can be simplified to focus solely on the upper structure. Using the improved Denavit-Hartenberg (D-H) method, the kinematic model of the robot is established, as illustrated in Fig. 7.

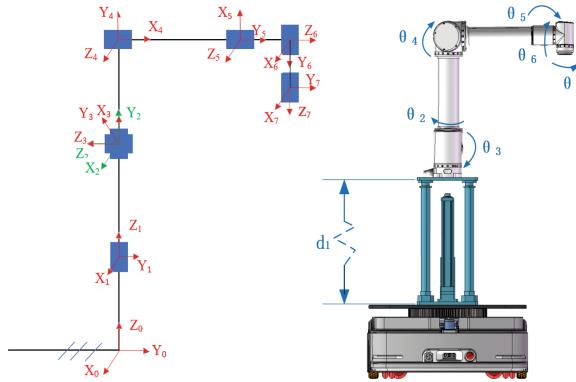


Fig. 7. Kinematic model of the hybrid robot.

Based on the joint coordinate systems established in the kinematic diagram, the Denavit-Hartenberg parameter table of each joint of the robot superstructure is written as shown in Table 1, where d_1 represents the stroke range of the robot's lifting mechanism, $d_1 \in [0, 300 \text{ mm}]$.

Table 1. D-H parameters of the hybrid robot

Joint i	Twist angle α_{i-1}	Link length a_{i-1}	Offset d_i	Joint angle θ_i
1	0	0	d_1	0
2	0	0	112.3	θ_2
3	90	0	85.6	θ_3
4	0	425	0	θ_4
5	0	401	0	θ_5
6	90	0	86	θ_6
7	-90	0	72.5	θ_7

3.2 Forward Kinematic Analysis

The Denavit-Hartenberg (D-H) parameter table of the robot essentially defines the relative pose relationships of each joint in three-dimensional space [11]. By taking the coordinate system $X_0 - Y_0 - Z_0$ as the base coordinate system, the coordinate transformation matrix from the end-effector coordinate system to the base coordinate system can be obtained. This is achieved through the transformation matrices of adjacent link coordinates:

$${}^{i-1}_iT = R_X(\alpha_{i-1})D_X(a_{i-1})R_Z(\theta_i)D_Z(d_i). \quad (1)$$

The general expression for ${}^{i-1}_iT$ is:

$${}^{i-1}_iT = \begin{bmatrix} c\theta_i & -s\theta_i & 0 & a_{i-1} \\ s\theta_i c\alpha_{i-1} & c\theta_i c\alpha_{i-1} & -s\alpha_{i-1} & -s\alpha_{i-1} d_i \\ s\theta_i s\alpha_{i-1} & c\theta_i s\alpha_{i-1} & c\alpha_{i-1} & c\alpha_{i-1} d_i \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

By substituting the parameters of each joint from the Denavit-Hartenberg parameter table into the expression, the coordinate transformation matrix between adjacent joints and its inverse matrix can be obtained, and by performing the consecutive multiplication of ${}_1^0T, {}_2^1T, {}_3^2T, {}_4^3T, {}_5^4T, {}_6^5T, {}_7^6T$, the homogeneous transformation matrix ${}_7^0T$ of the robot end-effector relative to its base coordinate system can be obtained:

$${}_7^0T = {}_1^0T {}_2^1T {}_3^2T {}_4^3T {}_5^4T {}_6^5T {}_7^6T, \quad (3)$$

where $c\theta$ is the shorthand for $\cos \theta$, and $s\theta$ is the shorthand for $\sin \theta$. The above equation forms the forward kinematics equation of the intelligent inspection hybrid robot.

3.3 Inverse Kinematics Analysis

This subsection focuses on solving the inverse kinematics equations for the hybrid robot. Assuming that the desired end-effector pose relative to the base coordinate system is given as , it is expressed as T_{end} . Similarly, according to the forward kinematics equations, it can be determined that:

$${}_E^B T = {}_7^0T = {}_1^0T {}_2^1T {}_3^2T {}_4^3T {}_5^4T {}_6^5T {}_7^6T. \quad (4)$$

In this case, the homogeneous transformation matrices ${}_j^iT, {}_j^iT^{-1}(i = 0..6, j = 1..7)$ for any two joints are derived from equation (3), and θ_i is solved using the method of separation of variables.

In practical tasks, d_1 is fixed to specific values $\{0, 100, 200, 300\}$ based on the z-axis height of the operation point on the electrical cabinet. Therefore, the inverse solution process focuses solely on the six-degree-of-freedom manipulator, and the solution for $\theta_i(i = 2..7)$ is as follows:

The process for sequentially solving $\theta_2, \theta_6, \theta_7$ based on the corresponding element equality relationships in ${}_E^2T$ is as follows. The homogeneous transformation matrices ${}_j^iT$ and inverse matrices ${}_j^iT^{-1}$, derived from the forward kinematics equations, ${}_E^2T$ can be expressed as follows:

$${}_E^2T = {}_3^2T {}_4^3T {}_5^4T {}_6^5T {}_7^6T = {}_2^0T^{-1} {}_E^BT. \quad (5)$$

The elements (2, 3) and (2, 4) in the two matrices correspond to each other, as do the elements (2, 1) and (2, 2), resulting in:

$$\begin{cases} a_y c\theta_2 - a_x s\theta_2 = -c\theta_6 \\ p_y c\theta_2 - p_x s\theta_2 = -d_3 - d_7 c\theta_6 \end{cases}, \quad \begin{cases} n_y c\theta_2 - n_x s\theta_2 = -c\theta_7 - s\theta_6 \\ o_y c\theta_2 - o_x s\theta_2 = s\theta_6 s\theta_7 \end{cases}. \quad (6)$$

The system of equations is solved by combining them, yielding the following solution:

$$\begin{aligned} \theta_2 &= \arctan \left(\frac{d_3}{\pm \sqrt{((d_7 a_x - p_x)^2 + (p_y - d_7 a_y)^2) - d_3^2}} \right) \\ &\quad + \arctan \left(\frac{p_y - d_7 a_y}{d_7 a_x - p_x} \right), \end{aligned} \quad (7)$$

$$\theta_6 = \arccos(a_y c\theta_2 - a_x s\theta_2), \theta_7 = \arctan \frac{o_y c\theta_2 - o_x s\theta_2}{n_x s\theta_2 - n_y c\theta_2}. \quad (8)$$

The process of sequentially solving for θ_3, θ_4 using the corresponding elements from is as follows. Similar to part a, ${}_5^2T$ can be written as:

$${}_5^2T = {}_3^2T {}_4^3T {}_5^4T = {}_2^0T^{-1} {}_E^BT {}_E^6T^{-1} {}_6^5T^{-1}. \quad (9)$$

Equating the elements (1, 3) and (3, 4) in both matrices gives (X, Y are formulas that contain parameters $c\theta_2, c\theta_7, s\theta_2, s\theta_7, d_{1,2,6,7}, o_{x,y,z}, a_{x,y,z}, n_{x,y,z}$, and due to length, they will not be elaborated here):

$$\begin{cases} a_3 c\theta_3 + a_4 c\theta_{34} = X \\ a_3 s\theta_3 + a_4 s\theta_{34} = Y \end{cases} \quad (10)$$

By squaring and adding the two equations, we obtain the solution:

$$\theta_4 = \pm \arccos \frac{X^2 + Y^2 - a_3^2 - a_4^2}{2a_3 a_4}, \theta_3 = \arctan \frac{a_4 X s\theta_4 + Y(a_3 + a_4 c\theta_4)}{X(a_3 + a_4 c\theta_4) - a_4 Y s\theta_4}. \quad (11)$$

The process of sequentially solving for θ_5 using the corresponding elements from is as follows. ${}_E^4T$ can be written as:

$${}_E^4T = {}_5^4T {}_6^5T {}_E^6T = {}_4^3T^{-1} {}_3^2T^{-1} {}_2^1T^{-1} {}_E^BT. \quad (12)$$

Equating the elements (1, 3) and (2, 3) in both matrices gives:

$$\begin{cases} a_x c\theta_{34} + a_y s\theta_2 c\theta_{34} + a_z s\theta_{34} = -c\theta_5 s\theta_6 \\ -a_x c\theta_2 s\theta_{34} - a_y s\theta_2 s\theta_{34} + a_z c\theta_{34} = -s\theta_5 s\theta_6 \end{cases}. \quad (13)$$

After performing the joint calculations, we obtain the solution:

$$\theta_5 = \arctan \frac{-a_x c\theta_2 s\theta_{34} - a_y s\theta_2 s\theta_{34} + a_z c\theta_{34}}{a_x c\theta_{34} + a_y s\theta_2 c\theta_{34} + a_z s\theta_{34}}. \quad (14)$$

At this point, $\theta_i (i = 2 \dots 7)$ have been solved, and the inverse kinematics solution for the composite robot is complete.

3.4 Forward Kinematics Verification

The forward kinematics verification method compares the results obtained from the homogeneous transformation matrix calculated using the D-H modeling method with those obtained from the MATLAB Robotics Toolbox. By using the MATLAB robot teaching interface (shown in Fig. 8), the displacement of each joint can be adjusted [12].

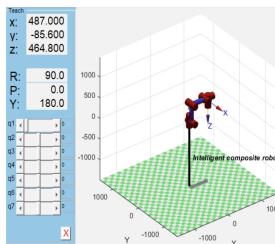


Fig. 8. MATLAB robot teaching interface

As shown in Table 2, tests with three joint variable sets yielded highly consistent results with negligible differences, confirming the forward kinematics model's accuracy.

Table 2. Verification of forward kinematics analysis results

Joint variables [$d_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6, \theta_7$]	Calculated Position(mm)	MATLAB teaching Interface Position(mm)
$[0, \frac{\pi}{4}, -\frac{\pi}{4}, \frac{\pi}{3}, \frac{\pi}{6}, \frac{\pi}{6}, \frac{\pi}{2}]$	(595.6772, 525.8858, 533.0210)	(595.677, 525.886, 533.021)
$[150, 0, -\frac{\pi}{4}, 0, \frac{\pi}{3}, -\frac{\pi}{6}, 0]$	(683.3903, -121.8500, 240.8816)	(683.390, -121.850, 240.882)
$[300, 0, -\frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{3}, 0, 0]$	(632.3, 85.6, 1066.9)	(632.276, -85.600, 1066.86)

3.5 Inverse Kinematics Verification

This section verifies the correctness of the inverse kinematics analysis for the hybrid robot through motion planning. Based on the linear trajectory planning function in the MATLAB Robotics Toolbox, joint trajectories for the power cabinet button operation are generated using the derived inverse kinematics equations. The accuracy of the inverse solution is validated by the motion planning results. The motion is divided into three stages: visual inspection, button operation, and resetting. The start point is set as $P_0 = [0.401, -0.158, 0.451]^T$, the inspection point as $P_i = [0.577, -0.085, 0.700]^T$, and the operation point as $P_e = [0.783, -0.080, 0.757]^T$. The joint displacement curves with respect to step size are shown in Fig. 9, illustrating joint movements at every 100 steps.

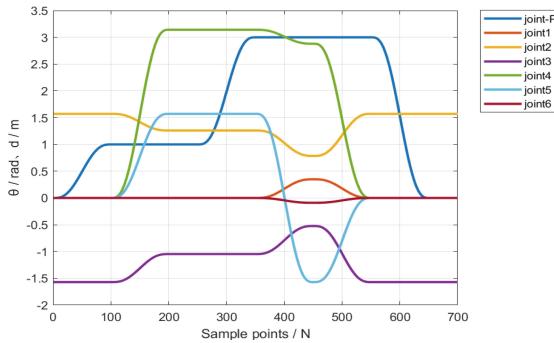


Fig. 9. Joint displacement vs. step size.

In summary, using the inverse kinematics solution for the hybrid robot presented in Section D helps the MATLAB motion planner generate displacement curves for each joint during the motion process, confirming that the inverse kinematics equations solved in this paper are applicable for the robot during the operation phase. The actual motion effects of each joint as the step size changes are shown in Fig. 10.

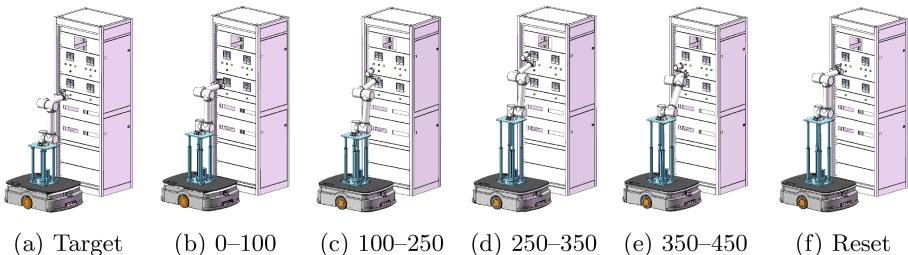


Fig. 10. Hybrid robot emergency operation process

4 Conclusion

Aiming at the working environment and tasks of power cabinet operation in substations, this paper focuses on a composite robot as the research subject. The structure and control system of the robot are first designed, the spatial advantages of a seven-degree-of-freedom robot are analyzed, and the tasks of the entire emergency operation process are described, including the design, control, and planning of the AGV chassis, lifting mechanism, and six-axis robotic arm. Subsequently, based on an improved D-H parameter method, the forward and inverse kinematics models of the robot are derived. Their correctness is verified through MATLAB calculations, and the entire operation process is visualized, providing a theoretical foundation for the coordinated control of the chassis and the seven-degree-of-freedom composite robot.

References

1. Guo, R., Li, B., Sun, Y.: A patrol robot for electric power substation. In: 2009 International Conference on Mechatronics and Automation, pp. 55–59 (2009)
2. Pereira, G.A.S., Pimentel, B.S., Chaimowicz, L., Campos, M.F.M.: Coordination of multiple mobile robots in an object carrying task using implicit communication. In: Proceedings 2002 IEEE International Conference on Robotics and Automation, pp. 281–286 (2002)
3. Gregori, S., Tur, M., Nadal, E.: Parametric model for the simulation of the railway catenary system static equilibrium problem. *Finite Elem. Anal. Design* **115**, 21–32 (2016)
4. Roa, M.A., Berenson, D., Huang, W.: Mobile manipulation: toward smart manufacturing. *IEEE Robot. Autom. Mag.* **22**, 14–15 (2015)
5. Allan, J., Beaudry, J.: Robotic systems applied to power substations-a state-of-the-art survey. In: International Conference on Applied Robotics for the Power Industry, pp. 1–6 (2014)
6. Menendez, O., Auat Cheein, F.A., Perez, M., Kouro, S.: Robotics in power systems: enabling a more reliable and safe grid. *IEEE Ind. Electron. Mag.* **11**, 22–34 (2017)
7. Lu, S., Zhang, Y., Su, J.: Mobile robot for power substation inspection: a survey. *IEEE/CAA J. Autom. Sin.* 830–847 (2017)
8. Fisher, M., Cardoso, R., Collins, E.: An overview of verification and validation challenges for inspection robots. *Robotics* **10**(67), 1–27 (2021)
9. Arduengo, M., Torras, C., Sentis, L.: Robust and adaptive door operation with a mobile robot. *Intell. Serv. Robot.* 409–425 (2021)
10. Ott, C., Bäuml, B., Borst, C.: Autonomous opening of a door with a mobile manipulator: a case study. *IFAC Proc.* 349–354 (2007)
11. Zhu, D., Zhang, C., Wu, W.: Fundamentals of Robotic Kinematics. Mechanical Industry Press, Beijing, China (2020)
12. Barakat, A.N., Gouda, K.A., Bozed, K.A.: Kinematics analysis and simulation of a robotic arm using MATLAB. In: 2016 4th International Conference on Control Engineering and Information Technology (CEIT), pp. 1–5 (2016)



Design and Output Characteristics of Carbon Nanotube-Silicone Rubber Based Piezoresistive Tactile Sensors

Bowen Cui , Ling Weng, Shichao Zuo, Xiaotao Du, Xiaopeng Ji, and Lanyang Hao

State Key Laboratory of Intelligent Power Distribution Equipment and System, Hebei University of Technology, Tianjin 300401, People's Republic of China

llweng@hebut.edu.cn

Abstract. Inspired by the tactile perception mechanism of human fingertips, an array flexible piezoresistive tactile sensor composed of carbon nanotube particles and silicone rubber was designed and fabricated, which has both strong flexibility and high sensitivity. The sensor comprises a pressure-sensitive conductive film, an array flexible printed circuit board, an organosilicon elastic substrate and a post-processing circuit. The flexible tactile sensor circuit board is embedded in the grooves on the silicone elastic substrate, and the pressure-sensitive conductive film is covered on the flexible tactile sensor circuit board. The flexible printed circuit board is an array of 3×3 comb sensing units. The pressure-sensitive conductive film consists of a carbon film and nine hemispherical conductive contacts distributed on the carbon film. Each comb sensing unit corresponds to a hemispherical conductive contact. The post-processing circuit converts pressure into an electrical signal by using the principle of voltage division. The force measurement range of the sensor is 0.1 N to 10 N. The pressure and output voltage basically change in a piecewise linear manner. As the pressure increases, the rate of voltage change decreases. This sensor can effectively detect the magnitude and direction of three-dimensional forces.

Keywords: Hemispherical carbon film · tactile sensor · three-dimensional force detection · High sensitivity · dynamic response characteristics

1 Introduction

Touch, as one of the most important senses in the human body, can sense information such as temperature, pressure, shape, position and texture of the outside world, and feed it back to the human body to achieve fine operation and interaction with the environment. Tactile sensors, by simulating the tactile function of living organisms, endow robots with the ability of tactile perception, becoming a key component for robots and wearable devices to achieve human-machine interaction and having broad application prospects. Common tactile sensors include piezoelectric [1–5], capacitive [6–9], piezoresistive [10–14], electromagnetic [15, 16], etc. Researchers have developed and optimized various types of tactile sensors to achieve highly sensitive and multi-functional tactile sensing.

Among them, the piezoresistive tactile sensor uses activated carbon material as the sensing unit. It features a flexible structural design, a wide range of material options, and high signal recognition accuracy. Gao et al. [17] proposed a flexible multifunctional piezoresistive tactile sensor based on microchannel-constrained MXene by designing electrode layers with micro-nano structures. By taking advantage of the microchannel confinement effect of the fingerprint-structured channels, MXenes can be well confined in the grooves and form a three-dimensional stacking structure. This endows MXenes with a larger deformation space and more sensitive fretting ability. The distance between adjacent interlayers of different MXenes will decrease to varying degrees with the change of external force, thereby resulting in the change of corresponding resistance, which can achieve the function of flexible tactile sensing. Pu Nie et al. proposed a high-performance piezoresistive electronic skin with bionic hierarchical microstructure and microcracks [18]. Inspired by the hierarchical microstructure on the surface of banana leaves, a PDMS substrate with a mountain pattern of secondary and tertiary ridges arranged at the micrometer scale was prepared. A thin layer of nano-silver wire was deposited on the surface of the microstructure PDMS as the sensitive electrode of piezoresistive electronic skin. Under the action of pressure, microcracks were formed in the nano-silver wire layer and the resistance changed. So as to achieve the sensing function. Longlong Chen et al. proposed a simple high-performance piezoresistive tactile sensor based on polymer membrane microcracks [19]. These cracks are formed by the self-stretching of the metal layer sputtered on the polydimethylsiloxane film and embedded in the cracks. The sensitive material is Ti₃C₂-MXenes. The sensor features high sensitivity, high elasticity and high resolution. These excellent characteristics prove that it has broad application prospects in the fields of touch screens and human activity monitoring.

Zhu Y [20] developed a tactile sensor for three-dimensional force detection and sliding recognition. Its sandwich structure sensor includes an upper PDMS bump, a conductive rubber layer located in the middle, and an electrode layer at the lower level. Although the sensor with this sandwich structure can measure force, due to its design of the contact layer as a convex block, it cannot quickly rebound to the initial state after deformation, and the recovery ability of the sensor is weak.

To solve the problems of small pressure detection range and weak recovery ability after stress detection of piezoresistive sensors, this paper innovatively proposes a hemispherical piezoresistive tactile sensor based on carbon nanotube-silicone rubber composite materials. It is made of a hemispherical carbon film composed of upper carbon nanotube particles mixed with silicone rubber, a middle PCB flexible circuit board and a bottom silicone protective layer. The hemispherical film made by compounding carbon nanotube particles with silicone rubber is simple to prepare and can achieve the detection of three-dimensional forces. Meanwhile, it has good flexibility and strong force rebound ability, which can effectively solve the problem of long response time of ordinary piezoresistive tactile sensors. The force measurement range of the sensor is 0.1 N to 10 N. The pressure and output voltage basically change in a piecewise linear manner, and as the pressure increases, the rate of voltage change decreases. When the pressure is between 0.1 and 0.5 N, the sensitivity is 8.75 MΩ/N; when the pressure is between 0.5 and 1 N, the sensitivity is 02 MΩ/N; when the pressure is between 1 and 10 N, the voltage output is nonlinear.

2 The Structural Design of the Hemispherical Piezoresistive Tactile Sensor

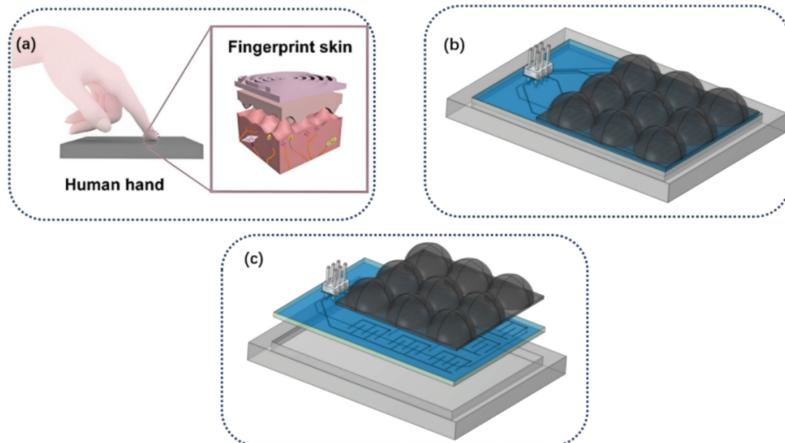


Fig. 1. Schematic diagram of the sensor (a) Structure of human fingertips (b) Schematic diagram of the sensor unit structure (c) Disassembly diagram of the sensor unit structure

The structure diagram of the human fingertip is shown in Fig. 1(a). The surface is distributed with a rich array of sensing units. The human body can sense external information by directly contacting objects and feed it back to the nerve tissue in the skin. Fig. 1(b) is a schematic diagram of the structure of the hemispherical pressure sensing unit, and Fig. 1(c) is a disassembly diagram of the hemispherical sensor. It is composed of a 3×3 array hemispherical contact film, a 3×3 comb-shaped PCB board and a silicone protective shell. The hemispherical contact film made of activated carbon can efficiently convert the external compressive stress into the deformation of the contact surface and conduct the force to the PCB circuit board. The PCB circuit board adopts a two-layer structure. The upper layer is an electrocomb-shaped sensing unit led out from the leftmost input, which corresponds one-to-one with the array hemispherical carbon film. Each of its comb-shaped sensing units has a via hole punched in the upper right corner to connect to the lower layer of the PCB circuit board, and then the lower layer is connected to the output terminal, thus forming a complete sensing unit. What is rather unique is that this circuit board remains in an open circuit state when it is not subjected to external pressure. A voltage signal is generated only when it comes into contact with the upper hemispherical carbon film under external pressure. At the same time, the circuit board returns to the open circuit state after the pressure disappears, which will not affect the next test. The unique comb-shaped sensing unit on the circuit board can expand the measurement range and reduce the measurement blind zone. Through the collaborative design of materials - structure - circuit, this sensing system realizes the conversion from compressive stress to electrical signal output and completes the tactile perception.

2.1 The Preparation Method of Hemispherical Contact Films

The preparation method of the hemispherical carbon film contact film is shown in Fig. 2. Firstly, mix the silicone rubber agent A and agent B in a 1:1 ratio to form a prepolymer. Then, mix the silicone rubber prepolymer with carbon nanotube particles (with a particle size of 10 μm , produced by the Chengdu Institute of Organic Chemistry, Chinese Academy of Sciences) in a mass ratio of 15: Uniformly mix to form a carbon nanotube-silicone rubber prepolymer. To improve the influence of temperature on the sensor, add PI crystals in a ratio of 15:2 to the carbon nanotube-silicone rubber prepolymer and then place it in a mixer. Due to the small particle size of carbon nanotubes, in order to prevent them from not being evenly distributed with the silicone rubber prepolymer, they should be stirred in a mixer for at least 15 min or more. After the stirring was completed, the prepolymer was placed in a vacuum environment at room temperature for deaeration treatment. The density of bubbles on the surface of the prepolymer was observed. After repeated pressurization and depressurization several times, it could be observed that the bubbles were significantly reduced. After taking it out, it was poured into the mold and left to stand for about 8 h. After curing, it was taken out, forming a 3×3 array hemispherical contact film with a diameter of 15 mm.

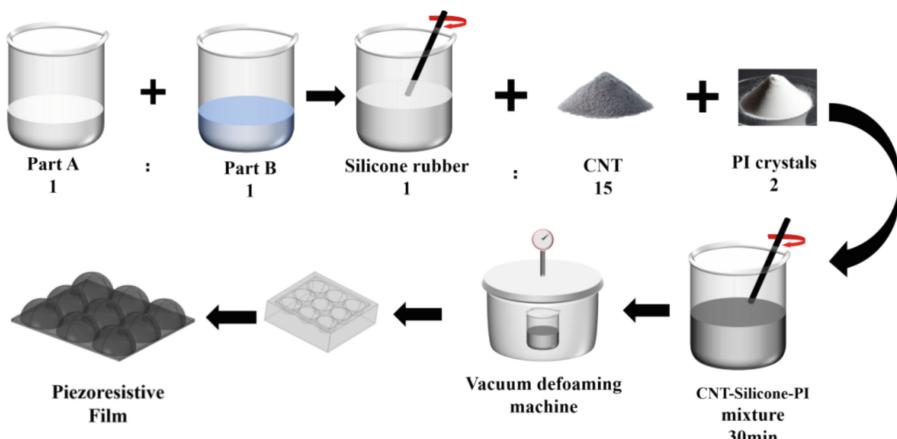


Fig. 2. Preparation process of hemispherical carbon film

The carbon films of different ratios were tested using the ohm range of a multimeter and a Young's modulus tester. The results are shown in Fig. 3. The silicone rubber prepolymer and carbon nanotube particles were divided into four groups in the ratios of 1:5, 1:10, 1:15, and 1:20. As can be seen from Fig. 3, as the amount of added carbon nanotube particles gradually increased, the conductivity of the carbon film became better and better, and the Young's modulus also gradually increased. In order to ensure that the carbon film has good flexibility and electrical conductivity, a carbon film with a ratio of 1:15 is used as the contact film on the surface of the sensing unit.

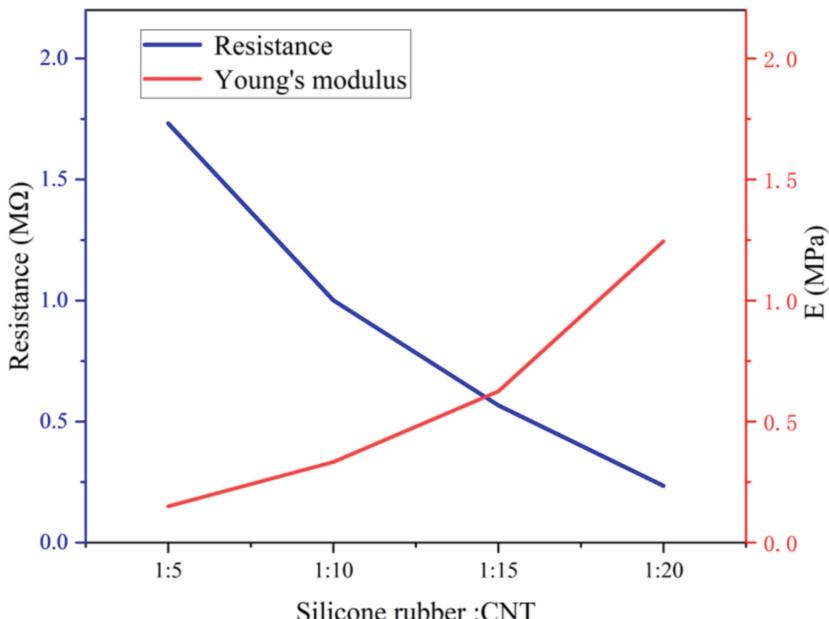


Fig. 3. Shows the corresponding resistances and Young's modulus of carbon nanotubes at different ratios

2.2 Post-processing Circuit and Comb Array Design of Hemispherical Carbon Film Sensor

When collecting data, since the data acquisition card can only collect voltage signals, while the piezoresistive sensor causes changes in resistance values and more intuitively the current signal, a post-processing circuit is designed for the piezoresistive sensor to supply power to the sensor and detect the data. In the design of the post-processing circuit, a resistor is connected in series in the circuit board by using the principle of series voltage division of resistors. By using the method of comparative experiments, 1 M ohms, 100 K ohms and 10 K ohms were serialized into the circuit, and the experiments were conducted on the data acquisition card. The output result graph is shown in Fig. 4.

After the experiment, it was found that the electrical signal change of the sensor with a 1- Ω resistance value inserted in series was more obvious. In order to observe the change of the electrical signal more intuitively, a 1- Ω resistor was adopted as the resistor of the post-processing circuit.

The sensing circuit board is designed as an electric comb-shaped sensing unit. The external information detected by a single comb sensor unit is limited and it cannot identify objects with large areas and high complexity. Therefore, it is designed as a 3×3 array as shown in Fig. 1(c). Each comb sensing unit is divided into external comb teeth and internal comb teeth. The external comb teeth are rectangular, with three evenly distributed comb teeth pointing downward. The inner comb teeth are facing upwards, and the two comb teeth are distributed in the gap between the upper comb teeth. Each of

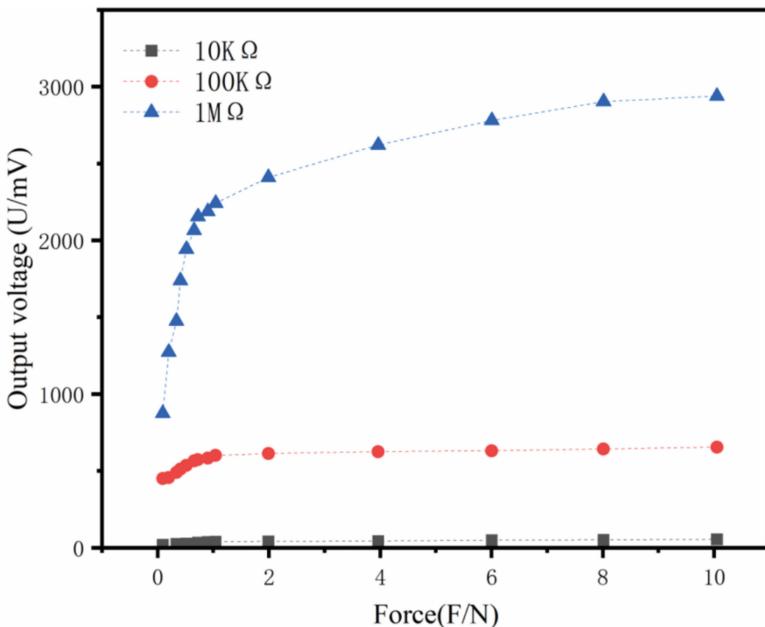


Fig. 4. Shows the output of the corresponding sensor acquisition card under different resistances

the three internal comb teeth in each row is led out through a single lead and connected to an input voltage interface.

3 Performance Analysis of Hemispherical Sensing Arrays

3.1 Static Output Characteristics of Hemispherical Sensing Units

According to the designed connection circuit, connect the power supply and output signal terminals to the left side of the post-processing circuit, and connect the sensor circuit board to the right side of the post-processing circuit. After the connection is completed, test the single sensing unit. The experimental platform is set up as shown in Fig. 5(a), including a high-precision electronic balance, a digital push-pull force gauge, a DC power supply, a data acquisition card and a display. Among them, the output part of the sensor is placed on the digital display push-pull force gauge to form a simple static force application device. Compared with the ordinary digital force gauge, it can apply a much smaller force and control the step size. By reading the gram m on the balance display screen and using $F = mg$ (the acceleration due to gravity is $g = 9.8 \text{ m/s}^2$), the static force acting on the antennae sensor at this time can be obtained. By adjusting the digital push-pull force gauge downward to apply a greater pressure, different pressures correspond to different static forces of varying magnitudes, and the output voltage value of the sensor is recorded.

By adjusting the digital push-pull force gauge downward, the reading m on the display screen is read to calculate the static pressure F. Each pressure F corresponds to an output

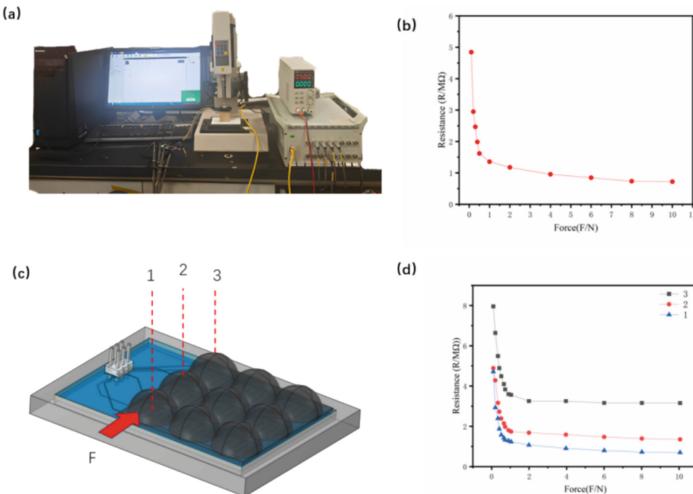


Fig. 5. Static characteristic test results: (a) Static characteristic test platform; (b) Curve of the relationship between the output voltage and pressure of different pressure sensors applied; (c) Schematic diagram of the tangential force of the sensor; (d) Output voltage of each sensor unit when the tangential force is applied

voltage U . The voltage signal is input to the PC terminal through the data acquisition card to read the voltage value. Take the sensor unit in the upper left corner as an example. Figure 5(b) shows the output voltages corresponding to different static forces. It can be seen that when the minimum detectable force of this sensor is 0.1 N, the output voltage reaches saturation and no longer changes when the force is 10 N. Therefore, the force measurement range is 0.1 N to 10 N. By observing the corresponding relationship graph, the pressure and output voltage basically change in a piecewise linear manner, and as the pressure increases, the rate of voltage change decreases. When the pressure is between 0.1 and 0.5 N, the sensitivity is $8.75 \text{ M}\Omega/\text{N}$; when the pressure is between 0.5 and 1 N, the sensitivity is $2 \text{ M}\Omega/\text{N}$; when the pressure is between 1 and 10 N, the voltage output is nonlinear.

To ensure the reliability of the experimental data, 20 repeated stress tests were conducted, and the maximum deviation of the multiple experimental data did not exceed 3.86%. To calibrate the linearity of the output voltage of the sensor, R^2 is introduced for characterization, and its calculation formula is: $R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$

Among them, SS_{res} is the sum of residual squares, that is, the sum of the squares of the differences between the measured actual voltage values and the predicted values, and SS_{tot} is the sum of total squares, that is, the sum of the squares of the differences between the actual voltage values measured in twenty experiments and the twenty mean values

According to the requirements, when the pressure is between 0.1 and 0.5 N, substituting SS_{res} and SS_{tot} , it is calculated that the output voltage R^2 is 0.95 when the pressure is between 0.1 and 0.5 N, and the output voltage R^2 is 0.89 when the pressure is

between 0.5 and 1 N. Experiments and calculations have verified that the sensor shows good linearity within the above-mentioned range.

When a tangential force is applied to any sensing unit, the voltage variation of adjacent sensing units and the voltage of the sensing unit farthest apart are detected. The force diagram is shown in Fig. 5(c), and the output result diagram is shown in Fig. 5(d). Because the hemispherical carbon film is designed to be closely adjacent to each other, the force will be transmitted to the surrounding sensing units. It can be known from Fig. 5(d) that the voltage amplitude of the detection unit is the highest, followed by the adjacent units, and the voltage amplitude of the sensing unit farthest apart is the lowest. This property can be used to determine the direction of the force acting on the sensing unit.

The hysteresis test of the flexible array piezoresistive sensor is shown in Fig. 6. It can be seen from the figure that the maximum difference between pressing and releasing is 130 mV, and the hysteresis rate is 4.37%, which is a normal error. This verifies that the output voltage of the sensor array is consistent. The cause of this hysteresis rate is generally the material of the hemispherical sensing unit and the sensor base. Subsequently, the packaging material can be optimized and a packaging structure with fewer contact points can be set, which can effectively reduce the hysteresis rate of the sensor.

In conclusion, to ensure the service life and output accuracy of this sensor, the pressure measurement range of this sensor is 0.1 to 10 N.

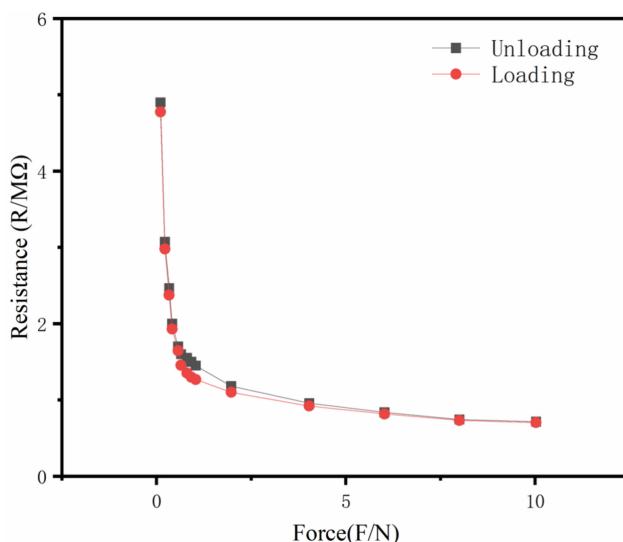


Fig. 6. Shows the hysteresis performance test loop

3.2 Dynamic Output Characteristics

Any flexible array piezoresistive sensing unit was selected for dynamic testing. The experimental platform was built as shown in Fig. 7(a), including a dynamic force generation device, a DC power supply, a data acquisition device and a PC display screen, etc. The dynamic force generation device consists of three parts, namely the signal generator (AFG2021-SC), the power amplifier (MB500VI), and the exciter (MODAL50). The flexible tactile sensing unit was placed on the platform for testing. The signal generator provided a 6 Hz square wave signal, which generated a dynamically changing force through the power amplifier and exciter and acted on the sensing unit. The response time is the period from when the sensor receives the pressure signal to when the output voltage reaches its peak. The recovery time, also known as the rebound time, is the time it takes for the output voltage to return to its initial state after the pressure is removed. As a hemispherical sensing unit, the shorter the response time, the stronger the perception ability of irregular shapes, the faster and more reliable the recognition of external signals, the shorter the recovery time, the stronger the recognition ability of continuous multiple stimuli, and the data will not be lost due to the failure to return to the initial state

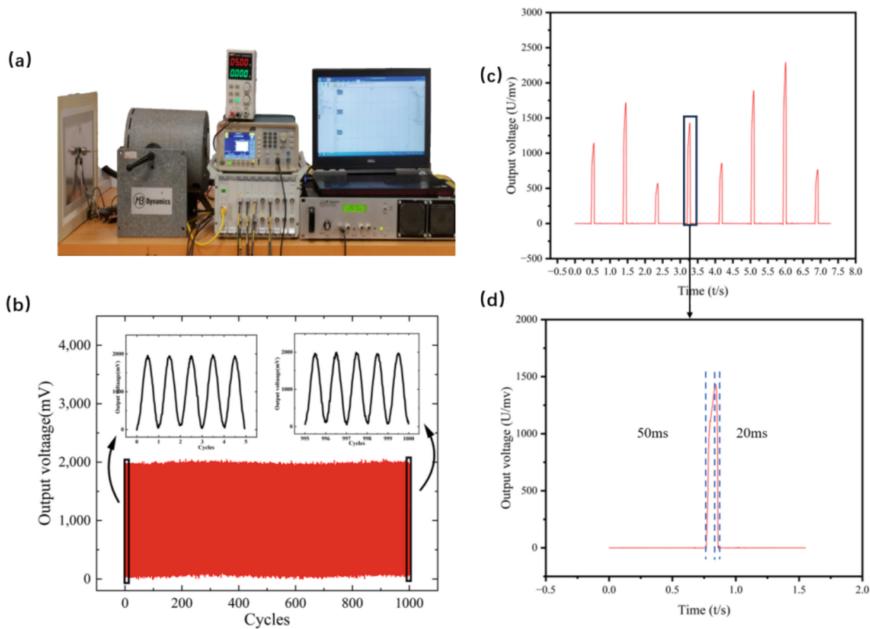


Fig. 7. Dynamic characteristic test diagram: (a) Dynamic test platform; (b) Output voltage diagram of 1000 cycles of continuous application of 1 N pressure; (c) Intermittent pressing output waveform; (d) Enlarged view of a randomly selected pressing waveform

A dynamic pressure of 1 N was applied to the hemispherical sensing unit for 1000 cycles. The output waveform is shown in Fig. 7(b). The waveforms at 10–15 s,

300–305 s, and 995–1000s were randomly selected for amplification and analysis. The output waveforms of the dynamic force at the initial measurement stage, the middle section, and the end section were observed. It was found that the waveforms output by the hemispherical sensing unit were basically consistent, and no significant deviation occurred. The hemispherical sensing unit of the sensor has stable dynamic characteristics and a long service life.

The hemispherical sensing unit was fixed on the experimental platform, and the output signal waveform was dynamically tested by manually and quickly pressing. The signal waveform of its output voltage is shown in Fig. 7(c). The overall waveform is stable, and the response is consistent as a pointed top wave under rapid pressing of different pressures, with the peak value being the measured value corresponding to the pressure. A single waveform was randomly amplified as shown in Fig. 7(d). It was observed that the response time was 50 ms and the recovery time was 20 ms. Both before and after the response returned to the equilibrium zero point, indicating that the hemispherical tactile sensor has a fast response speed, strong recovery ability, and high dynamic signal recognition accuracy.

3.3 Temperature Resistance Performance Test of the Sensor

In actual industrial production, the ambient temperature may constantly change, and the core component of this piezoresistive tactile sensor unit is a hemispherical sensing unit made of silicone rubber - carbon nanotubes. The change in temperature will cause the resistivity of materials to change, which in turn leads to thermal stress, affects contact and alters the properties of the materials. Therefore, in order to ensure the reliability of the hemispherical sensing unit, it is very necessary to analyze the influence of environmental temperature changes on the force measurement sensitivity of the tactile sensing unit. The influence of temperature on traditional piezoresistive sensors is significant. To improve this defect, PI crystals are added during the sensor preparation to reduce its coefficient of thermal expansion, thereby achieving the purpose of improving the influence of temperature.

In the experiment, the hemispherical sensing unit was selected as the object to be measured. The room temperature measured in the experiment was 23 °C. By adjusting the numerical control device to change the temperature of the heating plate, the value of the thermocouple temperature sensor close to the sensing unit gradually increased with a step size of 9 °C. As shown in Fig. 8(a), it is the temperature relationship curve between the heating plate and the thermocouple sensor. From the experimental results, it can be known that when the heating plate rises to the maximum 150 °C, The maximum temperature of the sensing unit is 77 °C, and at the same time, a forward pressure of 0-10 N is applied to the sensing unit by using the force transmission rod of the digital push-pull force gauge. As shown in Fig. 8(b), it is the variation curve of the force measurement sensitivity of the sensing unit with temperature. It can be seen that within the range of 23–59 °C, the change in the force sensitivity of the tactile sensor array does not exceed 6%, demonstrating good sensitivity stability (Table 1).

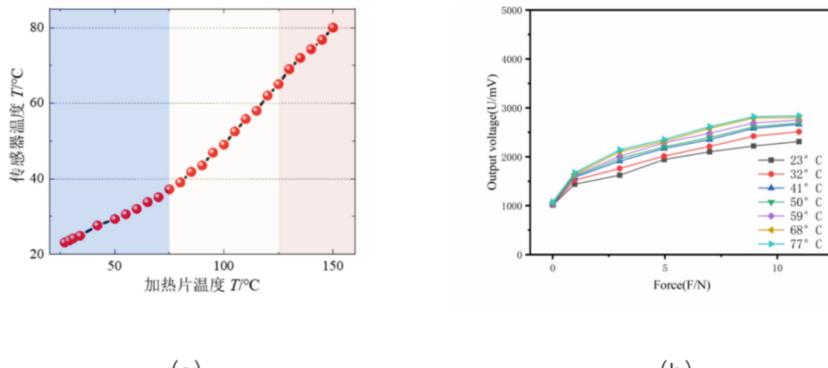


Fig. 8. (a) Temperature relationship curve between the sensor and the heating plate; (b) Output voltage of the hemispherical sensing unit under different temperatures and pressures

Table 1. Comparison of the performance of different sensors

Sensor models	Principle of sensing	Force measurement range	Force measurement sensitivity	Ref.
	piezoelectric effect	0–10 N	228.2 mV/N	[21]
	electromagnetic induction	0–15 N	56.3 mV/N	[22]
	Magnetostriction	0–3 N	126 mV/N	[23]
	Piezoresistive sensor	0–10 N	2 V/N	This work

4 Summary

Inspired by the force acting on human fingertips, a hemispherical piezoresistive sensor array that can be used for three-dimensional force direction and magnitude detection was designed and fabricated. The hemispherical contact film of silicone rubber - carbon nanotubes was prepared, and multiple sets of control experiments were carried out to

optimize the design of the carbon film. The sensor can identify a minimum pressure of 0.1 N, with a sensitivity of $8.75 \text{ M}\Omega/\text{N}$, the output voltage remains stable after 1000 presses, and has a long service life. The carbon film was prepared into a 3×3 hemispherical shape with a diameter of 15 mm. In the future, we will consider how to add a temperature measurement module to achieve multi-functional identification of pressure and temperature.

References

1. Payo, I., Sánchez, M., Rodríguez, D., Juárez, S.: Voltage drift compensation in charge amplifiers for DC measurements: application to piezoelectric paint sensors. *Measurement*. **201**, 111640 (2022)
2. Luo, J., et al.: Highly sensitive, wide-pressure and low-frequency characterized pressure sensor based on piezoresistive-piezoelectric coupling effects in porous wood. *Carbohydr. Polym.* **315**, 120983 (2023)
3. Ando, S., et al.: Analysis of respiratory properties using the low-frequency piezoelectric sensor in patients undergoing intravenous sedation: a prospective observational study. *Trends Anaesthet. Crit. Care.* **56**, 101357 (2024)
4. Park, H., Gbadam, G.S., Niu, S., Ryu, H., Lee, J.H.: Manufacturing strategies for highly sensitive and self-powered piezoelectric and triboelectric tactile sensors. *Int. J. Extreme Manuf.* **7**(1), 1–22 (2025)
5. Carluccio, V., et al.: Fabrication, characterization, and signal processing optimization of flexible and wearable piezoelectric tactile sensors. *IEEE Sensors J.* **23**(10), 10959–10969 (2023)
6. Alshawabkeh, M., et al.: Highly stretchable additively manufactured capacitive proximity and tactile sensors for soft robotic systems. *IEEE Trans. Instrum. Meas.* **72**, 7502210 (2023)
7. Xuan, W., et al.: Interdigital capacitor sensor-based cable health monitoring. *IEEE Trans. Ind. Electron.* **70**(7), 7301–7309 (2023)
8. Narayanan, P., et al.: A novel capacitance-to-time converter for differential-type capacitive sensor insensitive to offset mismatch and parasitic capacitance. *IEEE Sens. Lett.* **7**(10) (2023)
9. Tsao, C.T., Lu, M.S.C.: Development of a flexible capacitive tactile-proximity sensor array with CMOS integration for enhanced sensitivity. *IEEE Sensors J.* **24**(24), 40541–40548 (2024)
10. Wang, S., et al.: Integrated vertical force transfer structure for high-performance MEMS-based array piezoresistive tactile sensor. *Surf. Coat. Technol.* **497**, 131785 (2025)
11. Wu, R., et al.: Research on temperature drift mechanism and compensation method of silicon piezoresistive pressure sensors. *AIP Adv.* **13**(3), 035323 (2023)
12. Zhang, Z., et al.: Fully sprayed MXene-based high-performance flexible piezoresistive sensor for image recognition. *Nano Mater. Sci.* **6**(1), 77–85 (2024)
13. Wu, Y., et al.: Convex microarrays-based liquid metal soft Piezoresistive stress sensor with high sensitivity and large measurement range. *IEEE Sensors J.* **23**(9), 9176–9182 (2023)
14. Tan, M., Liu, W.: Research on flexible Piezoresistive tactile sensor based on silver nanowire. *Chin. J. Sens. Actuators.* **37**(1), 22–28 (2024)
15. Man, J., Chen, G., Chen, J.: Recent Progress of biomimetic tactile sensing technology based on magnetic sensors. *Biosensors*. **12**(11), 1054 (2022)
16. Ren, M., et al.: Flexible tactile sensors with self-assembled cilia based on magnetoelectric composites. *ACS Appl. Mater. Interfaces.* **17**(4), 6936–6947 (2025)
17. Gao, Y., et al.: Microchannel-confined MXene based flexible piezoresistive multifunctional micro-force sensor. *Adv. Funct. Mater.* **30**(11), 190–198 (2020)

18. H Song, T Bhattacharjee, SS Srinivasa, Sensing shear forces during food manipulation: resolving the trade-off between range and sensitivity, 2019 International Conference on Robotics and Automation, ICRA, IEEE, 2019,37(30):8367–8373.
19. Fang, H.S., et al.: Graspnet-lbillion: a large-scale benchmark forGeneral object grasping. In: Proceedings of the EEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11444–11453 (2020)
20. Zhu, Y., et al.: A flexible three-dimensional force sensor based on PI piezoresistive film. *J. Mater. Sci. Mater. Electron.* **29**, 19830–19839 (2018)
21. Chen, X., et al.: Scalable imprinting of flexible multiplexed sensor arrays with distributed piezoelectricity-enhanced micropillars for dynamic tactile sensing. *Adv. Mater. Technol.* **5**(7), 2000046 (2020)
22. Xie, S., et al.: High sensitivity and wide range soft magnetic tactile sensor based on electromagnetic induction. *IEEE Sensors J.* **21**(3), 2757–2766 (2021)
23. Weng, L., et al.: Magnetostrictive tactile sensor array for force and stiffness detection. *J. Magn. Magn. Mater.* **513**(0), 167068 (2020)



Multi-objective Path Planning for Two-Stage Mobile Robot Based on Improved ACO and Q-Learning

Xianyong Ruan^{1,2}, Lichuan Ning^{1,3,4(✉)}, Bo Tao^{1,2,3,4}, and Yuanmin Xie^{1,3,4}

¹ Key Laboratory of Metallurgical Equipment and Control, Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China
ninglichuan@wust.edu.cn

² Center for Biomechanics and Intelligent Measurement and Control, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China

³ Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, Hubei, China

⁴ National Experimental Teaching Center for Machinery, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China

Abstract. To meet the practical demands of mobile robots in multi-object handling and multi-point inspection tasks, a two-stage fusion algorithm is proposed, combining an improved Ant Colony Optimization (ACO) algorithm with an enhanced Q-Learning approach for efficient multi-objective path planning. In the first stage, the ACO algorithm is enhanced with an optimized distance heuristic function and an elite ant strategy to refine the pheromone update mechanism, thus determining an optimal visiting sequence of targets. In the second stage, an improved Q-Learning algorithm plans the paths between adjacent targets by redefining the state and action spaces, refining the reward function, and accelerating convergence through a combination of the Boltzmann strategy and an adaptive ϵ -greedy strategy. Furthermore, a path optimization strategy is introduced to enhance the overall path quality. Simulation experiments conducted in three different scenarios demonstrate that the proposed improvements significantly enhance the performance of Q-Learning, and the two-stage fusion algorithm proves to be both effective and feasible for solving multi-objective path planning problems in mobile robotics applications.

Keywords: ACO algorithm · Q-learning algorithm · Reinforcement learning · Mobile robot · Multi-objective path planning

1 Introduction

The advancement of robotics and artificial intelligence has enabled robots to perform increasingly complex tasks [1]. Path planning, as a core technology in robotics, plays a vital role in applications such as autonomous driving [2]. Existing research focuses on single-objective path planning, while real-world scenarios often require robots to arrive at multiple objectives sequentially to accomplish tasks, highlighting the need for multi-objective path planning research.

Path planning aims to find a safe route to a designated goal. Since its inception in the 1950s, the field has seen significant progress. Current methods are generally classified into four categories: graph search-based, sampling-based, intelligent algorithm-based, and learning-based approaches. Graph search-based methods, such as A* [3], convert the problem into a graph search but struggle with continuous spaces. Sampling-based methods like PRM [4] and RRT [5] are effective in high-dimensional spaces but often yield suboptimal paths that require smoothing. Intelligent algorithm-based methods, including ACO [6], and PSO [7], are suitable for multi-objective tasks but can be sensitive to parameter tuning, risking convergence to local optima. Learning-based methods, such as Q-Learning [8–10], and DQN [11], adapt well to dynamic environments but require extensive training, which may increase planning time.

While many studies have addressed single-objective path planning, practical applications often require multi-objective path planning. This involves navigating multiple target points, where the goal is to determine the shortest and safest path that visits all endpoints. Yu et al [12] improved the Hungarian algorithm and RRT algorithm for multi-objective path planning of surface vessels and introduced the ocean current constraint function in the RRT algorithm, which improved the effectiveness of the algorithm. Xiang et al [13] improved the A* algorithm by means of improving the evaluation function, and the improved A* algorithm was used for multi-objective planning by combining the improved A* algorithm with the greedy algorithm. Yu et al [14] treated the multi-objective path planning problem as a TSP problem, calculated the multi-objective access sequence using the gray wolf optimization algorithm, and used the improved D* Lite algorithm to achieve path planning between adjacent objectives. Zhang et al [15] improved the A* algorithm to plan the optimal paths between objectives to form the distance search matrix, and used the improved ACO algorithm to determine the optimal objective access sequence, thus achieving multi goal path planning.

Research on reinforcement learning for multi-objective path planning remains limited. To address this gap, a two-stage algorithm is proposed, integrating an improved ACO algorithm with an enhanced Q-Learning approach for efficient multi-objective path planning.

2 Problem Description and Preparation

2.1 Problem Description

Multi-objective path planning aims at designing a trajectory for an intelligent robot to visit all specified goals. In practice, the mobile robot starts from the starting position and needs to avoid obstacles to reach all the goal points in turn and return to the starting point. To reduce energy consumption, the path length should be shortened as much as possible to make the overall path as concise as possible. Multi-objective path planning consists of two main aspects: one is to determine the access order of multiple goal points; the other is to solve the optimal path between neighboring goal points according to the planned goal list. Therefore, multi-objective path planning can be regarded as a TSP and single-objective path planning.

At this stage, it is more efficient to determine the optimal access sequence before calculating individual paths. For N goals, computing all pairwise shortest paths requires

$N(N + 1)/2$ evaluations, while determining the sequence first reduces this to just $N + 1$ paths. This significantly lowers computational cost and planning time. Therefore, this study adopts the strategy of planning the access sequence first, followed by path planning between adjacent goals.

2.2 Map Processing

Path planning plays a fundamental role in fields such as autonomous driving, robotics, and logistics, and it relies heavily on the construction of accurate environmental maps. In practical applications, real-world environments must be transformed into simulation maps suitable for algorithmic processing. Among common map types—raster maps, vector maps, and free space maps—raster maps are often preferred due to their simplicity and intuitive representation of obstacles and navigable areas. Techniques such as YOLO can be applied to process real-world images into usable simulation maps.

In this study, a raster map is employed for experimental validation, as shown in Fig. 1(a). The map has a size of 10×10 with a grid resolution of 1. Black grids denote obstacles, while white grids indicate navigable areas. Unlike traditional methods that perform path planning directly on raster maps, this work introduces a preprocessing step in which small, adjacent obstacles are merged into larger obstacle regions. The boundaries of these merged regions are then extracted. To enhance safety during navigation, four safety points are generated for each obstacle by extending a predefined safe distance outward from its vertices, as shown in Fig. 1(b). These safety points are subsequently used to guide the path planning process.

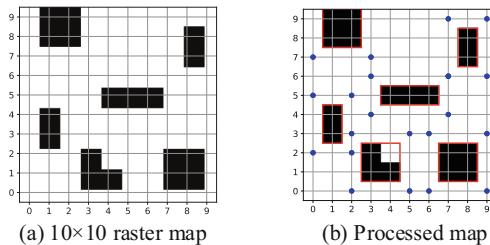


Fig. 1. Raster map processing

3 Multi-objective Path Planning Algorithm Description

3.1 Improvement of the ACO Algorithm

3.1.1 Optimize the Distance Heuristic Function

In traditional ACO algorithm for solving the TSP, the distance heuristic typically considers only the Euclidean distance between two nodes. However, in real-world environments, obstacles often exist between targets, and robots must avoid them, resulting in actual travel distances that are longer than straight-line distances. To address this, an

obstacle-related factor is introduced into the distance heuristic function in this study. This enhancement enables a more accurate representation of the effective distance between targets in the presence of obstacles, thereby contributing to shorter overall path lengths. The improved distance heuristic function is defined as follows:

$$\eta_{ij} = d_{ij} \cdot \left(1 + e^{-\frac{1}{obs_{ij}}} \right) \quad (1)$$

where d_{ij} is the Euclidean distance between target i and target j , and obs_{ij} is the number of obstacles crossed by the line connecting target i and target j .

3.1.2 Improve the Pheromone Updating Mechanism

In the ACO algorithm, ants attract other ants by releasing pheromones during the exploration of the path, and more ants will gather to the path with shorter length, the higher the concentration of pheromones on that path. To improve the iteration efficiency of the algorithm and speed up the convergence speed, the elite ant strategy is introduced. After all the ants have found the optimal path, the path lengths are counted and sorted by size, with the median as the boundary, and the ants whose path lengths are smaller than the median are regarded as elite ants. Reward the elite ants and punish the other ants. This will cause the ant colony to concentrate towards the optimal path faster. The new pheromone increment is shown as follows:

$$\Delta \tau_{ij} = \frac{Q}{l_k} \left(1 + a_1 \cdot \frac{l_{mid} - l_k}{l_{mid}} + a_2 \cdot \frac{l_{min} - l_k}{l_{min}} \right) \quad (2)$$

where l_k is the path length traveled by the k th ant, l_{mid} is the median path length of the round, l_{min} is the shortest path length of the round, and $a_1, a_2 \in (0, 1)$ is the reward-punishment factor, which indicates the degree of reward for the elite ants and also reflects the degree of punishment for the non-elite ants.

3.2 Improvement of Q-Learning Algorithm

3.2.1 Redefine the State Space and Action Space

In traditional Q-Learning, states are defined as grid points, resulting in a state space of size $n \times n$ for an $n \times n$ map. With four-neighborhood exploration, the action space is 4, leading to a Q-table size of $4 \times n \times n$. As the map grows, the Q-table expands rapidly, making path planning increasingly difficult.

To address this, the state and action spaces are redefined in this study. Safety points are used as states, and obstacles are reprocessed by identifying the nearest one intersecting the line from the current state to the goal. The four safety points at the diagonal corners of the selected obstacle are defined as actions. This approach significantly reduces the Q-table size and accelerates path planning.

3.2.2 Optimize the Reward Function

The reward function in reinforcement learning is a key feedback mechanism that influences learning efficiency and behavioral strategies. Traditional Q-learning algorithms

typically use sparse rewards or a single reward function, leading to a lack of reward continuity. To address this, a new continuous reward function is proposed to improve search efficiency.

The new reward function consists of five components: (1) distances from the current state to the goal, current state to the next state, and next state to the goal, to encourage favorable actions; (2) penalties for collisions between the current state and obstacles; (3) penalty for collision between next state and obstacle; (4) a significant reward for reaching the endpoint to motivate completion; and (5) evaluation of the distance between the current and next state relative to the goal. The total reward is calculated as a weighted sum of these components, as detailed in the formula below.

$$r_1 = \text{dis}(s_{\text{current}}, s_{\text{goal}}) - \text{dis}(s_{\text{current}}, s_{\text{next}}) - \text{dis}(s_{\text{next}}, s_{\text{goal}}) \quad (3)$$

$$r_2 = \begin{cases} -\text{dis}(s_{\text{current}}, s_{\text{goal}}) & \text{if } \text{path}(s_{\text{current}}, s_{\text{goal}}) \text{ collision} \\ 0 & \text{others} \end{cases} \quad (4)$$

$$r_3 = \begin{cases} -\text{dis}(s_{\text{next}}, s_{\text{goal}}) & \text{if } \text{path}(s_{\text{next}}, s_{\text{goal}}) \text{ collision} \\ 0 & \text{others} \end{cases} \quad (5)$$

$$r_4 = \begin{cases} 10 & \text{if } s_{\text{next}} == s_{\text{goal}} \\ 0 & \text{others} \end{cases} \quad (6)$$

$$r_5 = \text{dis}(s_{\text{current}}, s_{\text{goal}}) - \text{dis}(s_{\text{current}}, s_{\text{next}}) \quad (7)$$

$$r = b_1 \cdot r_1 + b_2 \cdot r_2 + b_3 \cdot r_3 + b_4 \cdot r_4 + b_5 \cdot r_5 \quad (8)$$

Where s_{current} , s_{next} and s_{goal} are the current state, next state and goal state respectively, $\text{dis}()$ is the Euclidean distance between different states, $\text{path}()$ collision indicates whether the connecting line of different states crosses the obstacle or not, $b_1, b_2, b_3, b_4, b_5 \in (0, 1)$, and $b_1 + b_2 + b_3 + b_4 + b_5 = 1$.

3.2.3 Improved Action Selection Strategy

Traditional algorithms often adopt the ε -greedy strategy for action selection. However, setting ε too small reduces exploration and risks convergence to a local optimum, while a large ε hinders convergence. To address this, an adaptive ε -greedy strategy is introduced, where ε decreases progressively with the number of iterations, balancing exploration and exploitation, and accelerating convergence. The adaptive ε formulation is provided below.

$$\varepsilon_k = \varepsilon_{\text{initial}} + \sin\left(\frac{\pi}{2} \cdot \frac{n_k - n_0}{N}\right)(\varepsilon_{\text{final}} - \varepsilon_{\text{initial}}) \quad (9)$$

where ε_k is the k th iteration, $\varepsilon_{\text{initial}}$, $\varepsilon_{\text{final}}$ the initial and final ε , respectively, n_k is the current iteration round, n_0 is the iteration number threshold, and N is the total number of iterations.

4 Experiments and Analysis

This section presents three sets of simulation experiments to evaluate the effectiveness and superiority of the proposed fusion algorithm. All experiments were conducted on a platform equipped with Python 3.10 and an Intel i5-12600KF processor.

4.1 Validation of the Effectiveness of the Improved Q-Learning Algorithm

To evaluate the effectiveness and generalization of the improved Q-Learning algorithm, three sets of experiments were conducted on maps of varying scales: 20×20 , 30×30 , and 40×40 . Results show that the improved algorithm successfully generated collision-free paths across all maps, with path lengths of 28.740, 42.996, and 60.545, respectively. The planned paths are illustrated in Fig. 2, while Fig. 3 presents the reward-iteration curves, indicating convergence after approximately 150 iterations. These results validate the effectiveness of the improved reward function and action selection strategy, and demonstrate the algorithm's robust path planning and generalization capabilities.

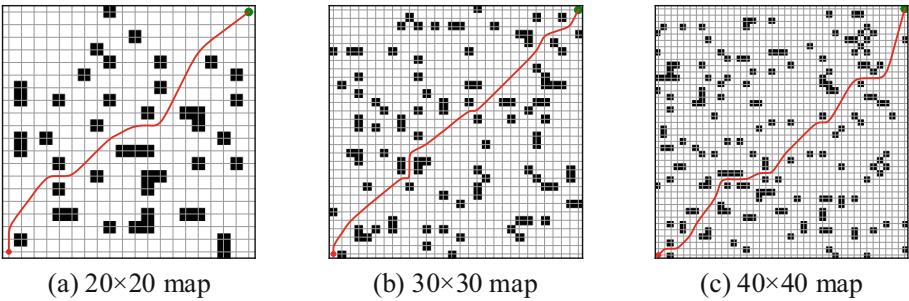


Fig. 2. Optimal path of improved Q-Learning algorithm on different scale maps

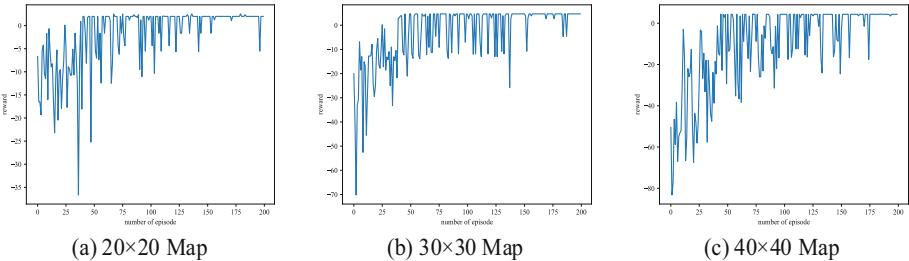


Fig. 3. Reward-iteration relationship on maps of different scales

4.2 Superiority Verification of the Improved Q-Learning Algorithm

To validate the effectiveness of the improved Q-Learning algorithm, comparative experiments were conducted in a 30×30 map environment with 20% obstacle coverage.

Algorithms from [8] (Algorithm 1), [9] (Algorithm 2), and [10] (Algorithm 3) were used as benchmarks. Each algorithm was run 20 times to ensure experimental reliability. Results indicate that the improved Q-Learning algorithm outperforms the benchmarks across multiple key metrics. Detailed performance data are presented in Table 1.

Table 1. Results of different algorithms

Targets	Data Type	Algorithm 1	Algorithm 2	Algorithm 3	Ours
Path Length	Best	58	58	46.498	47.878
	Mean	58.2	58.4	47.980	47.878
Inflection Points	Best	19	21	21	13
	Mean	28.1	27.8	25.35	13
Planning Time(s)	Best	14.071	10.709	70.474	8.114
	Mean	15.239	12.024	76.592	8.455

In terms of path length, the improved algorithm significantly outperforms Algorithm 1 and Algorithm 2, primarily due to the refined design of the action space. It also achieves the fewest number of inflection points, indicating smoother and more practical paths. Regarding planning time, Algorithm 3 requires the initialization of a Q-table using the ACO algorithm to construct the initial path, resulting in longer computation times. In contrast, the improved Q-Learning algorithm enhances both the state and action spaces, thereby increasing search efficiency and reducing computation time. These results not only validate the effectiveness of the algorithm in path optimization but also demonstrate its high efficiency in handling large-scale, complex environments. Figure 4 shows the optimal paths planned by different algorithms, and the superiority of the improved Q-Learning algorithm in terms of path quality can be clearly seen through visual comparison. This experimental result further validates the effectiveness of the improved algorithm.

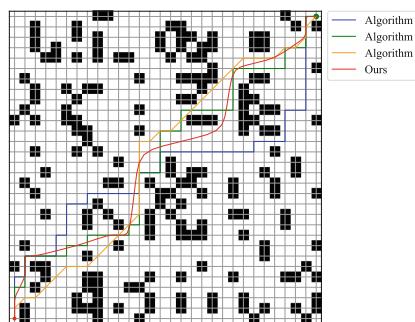


Fig. 4. Optimal paths of different algorithms on the 30×30 map

4.3 Experimental Validation of Multi-Objective Path Planning

To further verify the performance advantage of the fusion algorithm in complex environments, 12 targets are randomly set on an 80×80 simulation map to comprehensively evaluate the performance of the fusion algorithm in multi-target tasks. The evaluation indexes include the path length, the number of path turning points and the total turning angle to ensure the objectivity and comparability of the experimental results.

The experimental results show that the proposed fusion algorithm exhibits good performance in multi-objective path planning tasks. The comparison experimental data with Algorithm 1 in [14], Algorithm 2 in [15], and Algorithm 3 (ACO + QL) are shown in Table 2. Figure 5 shows the visualization results of the relevant data.

Table 2. Experimental results of multi-objective path planning simulation

Targets	Data Type	Algorithm 1	Algorithm 2	Algorithm 3	Ours
Path Length	Best	403.84	362.50	374.01	364.53
	Mean	404.57	362.79	375.18	366.48
Inflection Points	Best	107	104	111	90
	Mean	111.2	106.4	117.1	91.5
Turning Angle	Best	3243.9	3600	5400	3071.2
	Mean	3686.5	3757.5	5899.5	3112.6

In general, compared to Algorithm 1 and Algorithm 3, the shortest path lengths planned by the fusion algorithm are reduced by about 10% and 2.4% on average, with no significant difference from Algorithm 2. In addition, the overall optimization of path quality is also a highlight of the fusion algorithm. The number of path inflection points planned by the fusion algorithm is reduced by 6.6%, 12.6%, and 38.5% compared to Algorithm 1, Algorithm 2, and Algorithm 3, respectively, and the reduction in the number of inflection points means that the paths are more concise and continuous, which reduces unnecessary steering and complexity. Meanwhile, the proposed fusion algorithm also performs well in the optimization of path turning angles, with a significant reduction in the total turning angles of the paths compared to the other three algorithms, making the paths smoother and easier to execute and control. This result not only improves the practical executability and efficiency of the path, but also promises to reduce energy and time costs in practical applications. From the consideration of path safety, all positions of the path planned by the fusion algorithm maintain a certain safety distance from the obstacles, and the paths planned by other algorithms have overlapping edges and corners with the obstacles, which indicates that the paths planned by the fusion algorithm are safer. Figure 6 illustrates the optimal path for each algorithm.

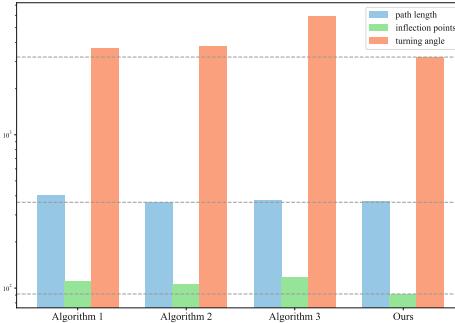


Fig. 5. Comparison of results of different algorithms

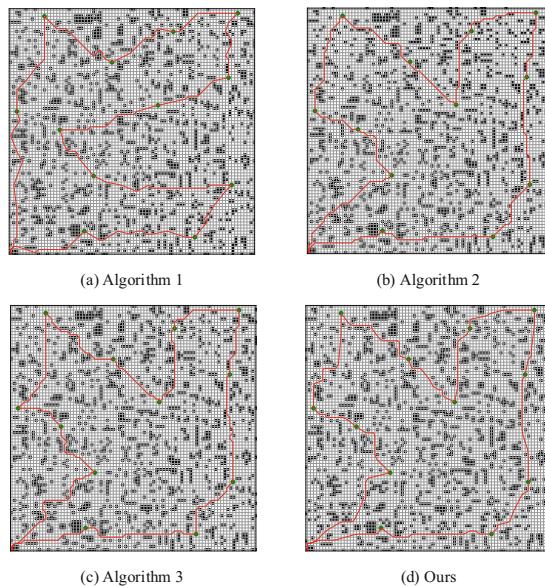


Fig. 6. Optimal paths of various algorithms in multi-objective path planning simulations

5 Conclusion

To address multi-objective path planning for intelligent mobile robots, a fusion algorithm combining improved ACO and Q-Learning is proposed. The problem is divided into two stages: solving the TSP for goal sequence and single-objective path planning between targets. ACO is enhanced with an optimized distance heuristic and elite ant strategy to improve pheromone updates and speed up convergence. Q-Learning is improved by redefining the state-action space and using a continuity-based reward function and action selection strategy, leading to more efficient search. Experiments show that the proposed method achieves better planning time and path quality than traditional approaches.

Future work will explore cooperative mechanisms and task allocation strategies in multi-robot systems to improve efficiency and flexibility. Additionally, efforts will focus

on enhancing learning efficiency and adaptability to dynamic environments, enabling real-time obstacle avoidance and task execution.

Acknowledgement. Hubei Science and Technology Innovation Talent Program for Science and Technology Service Talents Special (2024DJC073).

References

1. Licardo, J.T., Domjan, M., Orehovački, T.: Intelligent robotics—a systematic review of emerging technologies and trends. *Electronics*. **13**(3), 542 (2024)
2. Reda, M., Onsy, A., Haikal, A.Y., Ghanbari, A.: Path planning algorithms in the autonomous driving system: a comprehensive review. *Robot. Auton. Syst.* **174**, 104630 (2024)
3. Liu, N., Hu, Z., Wei, M., Guo, P., Zhang, S., Zhang, A.: Improved A* algorithm incorporating RRT* thought: a path planning algorithm for AGV in digitalised workshops. *Comput. Oper. Res.*, 106993 (2025)
4. Xu, Z., Deng, D., Shimada, K.: Autonomous UAV exploration of dynamic environments via incremental sampling and probabilistic roadmap. *IEEE Robot. Autom. Lett.* **6**(2), 2729–2736 (2021)
5. Li, Y., Wei, W., Gao, Y., Wang, D., Fan, Z.: PQ-RRT*: an improved path planning algorithm for mobile robots. *Expert Syst. Appl.* **152**, 113425 (2020)
6. Liu, C., et al.: An improved heuristic mechanism ant colony optimization algorithm for solving path planning. *Knowl.-Based Syst.* **271**, 110540 (2023)
7. Zhao, J., Deng, C., Yu, H., Fei, H., Li, D.: Path planning of unmanned vehicles based on adaptive particle swarm optimization algorithm. *Comput. Commun.* **216**, 112–129 (2024)
8. Zhou, Q., Lian, Y., Wu, J., Zhu, M., Wang, H., Cao, J.: An optimized Q-learning algorithm for mobile robot local path planning. *Knowl.-Based Syst.* **286**, 111400 (2024)
9. Zhong, Y., Wang, Y.: Cross-regional path planning based on improved Q-learning with dynamic exploration factor and heuristic reward value. *Expert Syst. Appl.* **260**, 125388 (2025)
10. Fang, W., Liao, Z., Bai, Y.: Improved ACO algorithm fused with improved Q-learning algorithm for Bessel curve global path planning of search and rescue robots. *Robot. Auton. Syst.* **182**, 104822 (2024)
11. Yang, X., Shi, Y., Liu, W., Ye, H., Zhong, W., Xiang, Z.: Global path planning algorithm based on double DQN for multi-tasks amphibious unmanned surface vehicle. *Ocean Eng.* **266**, 112809 (2022)
12. Yu, J., Chen, Z., Zhao, Z., Yao, P., Xu, J.: A traversal multi-target path planning method for multi-unmanned surface vessels in space-varying ocean current. *Ocean Eng.* **278**, 114423 (2023)
13. Dan, X., et al.: Combined improved A* and greedy algorithm for path planning of multi-objective mobile robot. *Sci. Rep.* **12**(1) (2022)
14. Yu, J., et al.: A hybrid multi-target path planning algorithm for unmanned cruise ship in an unknown obstacle environment. *Sensors*. **22**(7), 2429 (2022)
15. Zhang, B., Jin, W., Gao, X., Chen, W.: A multi-goal global dynamic path planning method for indoor mobile robot. In: 2021 3rd International Symposium on Robotics & Intelligent Manufacturing Technology (ISRIMT), pp. 97–103. IEEE (2021)



Adaptive Neural Bipartite Containment Control for Stochastic Non-strict Feedback Multi-agent Systems

Hailin Tang^{1,2} and Tianping Zhang^{1,2(✉)}

¹ College of Information Engineering, Yangzhou University, Yangzhou 225127, China

² School of Mathematical Sciences, Yangzhou University, Yangzhou 225002, China
tpzhang@yzu.edu.cn

Abstract. A novel adaptive compensation bipartite containment control strategy based on improved dynamic surface control (DSC) method and radial basis function neural networks (RBFNNs) is designed for non-linear stochastic multi-agent systems (SMASs). A definition for implementing containment control in probability is proposed. A high-gain observer is employed to reconstruct the unmeasurable system state, and the non-strict feedback term in the unknown function is reduced by a new processing method. A second-order command filter (CF) is constructed to eliminate the influence of filtering error on tracking performance and simplify the complexity of the controller.

Keywords: Stochastic multi-agent systems · Bipartite containment control · Dynamic surface control · Command filter

1 Introduction

In the context of complex and dynamically changing control objects, traditional single control system often struggle to meet practical requirements. Consequently, multi-agent systems (MASs) have rapidly emerged as a prominent research focus due to their exceptional fault tolerance and robust collaborative capabilities. In [1], the DSC method was introduced to prevent computational overload in backstepping. In [2], the CF technology was introduced to overcome the shortcomings of backstepping and DSC and a consensus tracking control strategy was designed. A fuzzy stochastic consensus controller was designed in [3]. When there is more than one leader in the MASs, consensus control evolves into containment control [4,5]. References [6,7] studied the containment control problem of simple SMASs. Tang et al. further simplified the design process by CF and DSC, and proposed a new neural adaptive output feedback containment control algorithm in [8].

However, in the actual situation, interactions among agents involve both collaboration and competition. In [9], Altafini first proposed a bipartite consensus

protocol for symbolic networks. In [10, 11], the bipartite consensus problem of different MASs was studied. In [12, 13] the bipartite containment control problem of MASs was explored. References [14–16] extended the bipartite consensus control problem to SMASs, and designed an adaptive bipartite controller using RBFNNs and fuzzy logic systems respectively. For non-strict feedback SMASs, a new adaptive bipartite containment controller based on CF and RBFNNs was proposed in [17]. At present, although SMASs has been widely studied and has produced fruitful results, each design method has some limitations, so we will propose an improved DSC method to overcome these problems and simplify the controller design. In addition, we do not just extend deterministic systems theory to stochastic systems, but propose a definition that satisfies containment control in probability.

2 Problem Preliminaries and Description

2.1 Graph Theory

The information exchange between N followers and M leaders is modeled by a signed directed graph $\mathcal{G} = (\mathcal{D}, \mathcal{B}, \mathcal{A})$, where $\mathcal{D} = \{1, \dots, N, N+1, \dots, N+M\}$ represents all agents (vertices), $\mathcal{B} \subset \mathcal{D} \times \mathcal{D}$ denotes the edge set, with $(j, i) \in \mathcal{B}$ indicating that agent i receives information from agent j . In particular, all leaders can only send information but cannot receive it. $\mathcal{A} = [a_{i,j}] \in R^{(N+M) \times (N+M)}$ is an adjacency matrix, if agent i can receive information from the agent j , i.e., $(i, j) \in \mathcal{B}, a_{i,j} \neq 0$. Otherwise, $a_{i,j} = 0$. In addition, $a_{i,i} = 0$. The follower and leader sets are $F = \{1, \dots, N\}$ and $L = \{N+1, \dots, N+M\}$. $\mathbb{N}_i = \{j | (j, i) \in \mathcal{B}\}$ is all the neighbors of the agent i ($i = 1, \dots, N$), and $\mathbb{N}_i^F = F \cap \mathbb{N}_i$, $\mathbb{N}_i^L = L \cap \mathbb{N}_i$. $\boldsymbol{\iota} = [\iota_{i,j}] = d - \mathcal{A} \in R^{(N+M) \times (N+M)}$ is the Laplacian matrix and $\boldsymbol{\iota}$ can be described as

$$\boldsymbol{\iota} = \begin{bmatrix} \iota_1 & \iota_2 \\ 0_{M \times N} & 0_{M \times M} \end{bmatrix}, \quad (1)$$

where $d = \text{diag}(d_1, \dots, d_N, d_{N+1}, \dots, d_{N+M})$ is the degree matrix and $d_i = \sum_{j \in \mathcal{D}} |a_{i,j}|$, $\iota_1 \in R^{N \times N}$, $\iota_2 \in R^{N \times M}$. If the vertex set \mathcal{D} can be divided into two subsets \mathcal{D}_1 and \mathcal{D}_2 that satisfy $\mathcal{D}_1 \cup \mathcal{D}_2 = \mathcal{D}, \mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$. When $\{i, j\} \in \mathcal{D}_o$ ($o \in \{1, 2\}$), they are in a cooperative relationship and $a_{i,j} = 1$. When $i \in \mathcal{D}_o, j \in \mathcal{D}_{3-o}$, they are in a competitive relationship and $a_{i,j} = -1$. Then, \mathcal{G} is said to be structurally balanced.

Assumption 1. [1] \mathcal{G} is structurally balanced, and every follower receives instruction from at least one leader.

Lemma 1. [17] Under Assumption 1, $\Lambda \boldsymbol{\iota}_1 \Lambda$ is an M -matrix, every element of $-\Lambda \boldsymbol{\iota}_1^{-1} \Lambda \boldsymbol{\iota}_2$ is nonnegative and each row sum of $-\Lambda \boldsymbol{\iota}_1^{-1} \Lambda \boldsymbol{\iota}_2$ equals 1, where $\Lambda = \text{diag}(\Lambda_1, \dots, \Lambda_i), i = 1, \dots, N, \Lambda_i = 1$ when $i \in \mathcal{D}_1$, otherwise, $\Lambda_i = -1$.

2.2 Problem Description

Consider a class of SMASs with M leaders and N followers, described by

$$\begin{cases} dx_{i,j} = [x_{i,j+1} + f_{i,j}(\bar{x}_i)]dt + g_{i,j}^T(\bar{x}_i)dw, \\ \quad i = 1, 2, \dots, N, j = 1, \dots, n-1, \\ dx_{i,n} = [u_i + f_{i,n}(\bar{x}_i)]dt + g_{i,n}^T(\bar{x}_i)dw, \\ \quad y_i = x_{i,1}, \end{cases} \quad (2)$$

where $\bar{x}_i = [x_{i,1}, \dots, x_{i,n}]^T \in R^n$ denotes the state vector, $f_{i,j}$, $g_{i,j}$ are smooth unknown functions, y_i , u_i represent the output and input, respectively, and only $x_{i,1}$ is measurable.

Control objective: All follower outputs converge to the desired convex hull, with all signals remaining semi-globally uniformly ultimately bounded (SGUUB) in probability.

Consider a general stochastic system as follows:

$$d\bar{h} = f(\bar{h})dt + g^T(\bar{h})dw, \quad (3)$$

where w is an r -dimensional Winer process.

$\forall V(\bar{h}) \in C^{1,2}$, we get

$$dV(\bar{h}) = LV(\bar{h})dt + \frac{\partial V(\bar{h})}{\partial \bar{h}^T} g^T dw, \quad (4)$$

$$LV(\bar{h}) = \frac{\partial V(\bar{h})}{\partial t} + \frac{\partial V(\bar{h})}{\partial \bar{h}^T} f + \frac{1}{2} \text{tr}[g^T \frac{\partial^2 V(\bar{h})}{\partial \bar{h}^T \partial \bar{h}} g]. \quad (5)$$

Definition 1. The convex hull $Co(Y_L)$ is formed by the leader outputs $Y_L = [y_{L,N+1}(t), \dots, y_{L,N+M}(t)]^T$, where $Co(Y_L) = \left\{ \sum_{j=N+1}^{N+M} v_{i,j} y_{L,j}(t) \mid v_{i,j} \geq 0, \sum_{j=N+1}^{N+M} v_{i,j} = 1 \right\}$, $i = 1, \dots, N$. If containment control is to be achieved in probability, then there is a set of coefficients $v_{i,j} \geq 0$, $\sum_{j=N+1}^{N+M} v_{i,j} = 1$, $\delta_1 > 0$ and $\delta_2 \in (0, 0.5)$ satisfying

$$\inf_{0 \leq t < \infty} P \left\{ \left| y_i(t) - \sum_{j=N+1}^{N+M} v_{i,j} y_{L,j}(t) \right| \leq \delta_1 \right\} \geq 1 - \delta_2. \quad (6)$$

Assumption 2. [8] There exist Lipschitz constants $F_i > 0$, $i = 1, \dots, N$ such that $|f_{i,j}(\bar{x}_i) - f_{i,j}(\hat{x}_i)| \leq F_i \sum_{j=1}^n |x_{i,j} - \hat{x}_{i,j}|$ hold, where $\bar{x}_i = [\hat{x}_{i,1}, \dots, \hat{x}_{i,n}]^T$.

3 Adaptive Bipartite Output Feedback Containment Controller

Firstly, the high-gain observer is constructed as follows:

$$\begin{cases} \dot{\hat{x}}_{i,j} = \hat{x}_{i,j+1} + k_{i,j} l_i^j (y_i - \hat{x}_{i,1}), j = 1, \dots, n-1, \\ \dot{\hat{x}}_{i,n} = u_i + k_{i,n} l_i^n (y_i - \hat{x}_{i,1}), \end{cases} \quad (7)$$

where $l_i > 1$ and $k_i(s) = s^n + k_{i,1}l_i s^{n-1} + \dots + k_{i,n-1}l_i^{n-1}s + k_{i,n}l_i^n$ is a Hurwitz polynomial.

Define the observation error as $E_{i,j} = \frac{x_{i,j} - \hat{x}_{i,j}}{l_i^{j-1}}$. According to (4) and (5), we get

$$d\bar{E}_{i,n} = [l_i B_i \bar{E}_{i,n} + \sum_{j=1}^n \frac{1}{l_i^{j-1}} \mathfrak{R}_{i,j} f_{i,j}] dt + \bar{g}_{i,n}^T dw, \quad (8)$$

where $\bar{E}_{i,n} = \begin{bmatrix} E_{i,1} \\ \vdots \\ E_{i,n} \end{bmatrix}$, $\mathfrak{R}_{i,j} = \begin{bmatrix} 0, \dots, 0, 1, \underbrace{0, \dots, 0}_{n-j} \end{bmatrix}^T$, $\bar{g}_{i,n} = [g_{i,1}, \dots, g_{i,n}]$, $B_i = \begin{bmatrix} -k_{i,1} \\ \vdots \\ -k_{i,n} \end{bmatrix}$, and there exists sole positive definite matrix $P_i > 0$ satisfying $B_i^T P_i + P_i B_i = -Q_i$.

Secondly, define the error transform as follows:

$$\begin{cases} s_{i,1} = \sum_{j \in \mathbb{N}_i^F} |a_{i,j}| [y_i - \text{sgn}(a_{i,j}) y_j] + \sum_{j \in \mathbb{N}_i^L} |a_{i,j}| [y_i - \text{sgn}(a_{i,j}) y_{L,j}], \\ s_{i,j} = \hat{x}_{i,j} - \omega_{i,j}, i = 1, \dots, N, j = 2, \dots, n, \end{cases} \quad (9)$$

where $\omega_{i,j}$ represent the output of the following CF

$$\begin{cases} \dot{\varsigma}_{i,j} = \phi_i \varsigma_{i,j+1}, \\ \dot{\varsigma}_{i,j+1} = -2\phi_i \Theta_i \varsigma_{i,j+1} - \phi_i (\varsigma_{i,j} - \alpha_{i,j-1}), \end{cases} \quad (10)$$

where $\phi_i > 0$, $\Theta_i \in (0, 1]$, $\omega_{i,j} = \varsigma_{i,j}$, $\omega_{i,j}(0) = \alpha_{i,j}(0)$, $\varsigma_{i,j+1}(0) = 0$, and $\alpha_{i,j-1}$ are the input of second-order CF.

According to (1) and (14), we have

$$\underline{s}_{N,1} = \Lambda \iota_1 \Lambda [Y - (-\Lambda \iota_1^{-1} \Lambda \iota_2 Y_L)], \quad (11)$$

where $\underline{s}_{N,1} = [s_{1,1}, \dots, s_{N,1}]^T$ and $Y = [y_1, \dots, y_N]^T$. From Assumption 1 and Lemma 1, let $-\Lambda \iota_1^{-1} \Lambda \iota_2 = [\iota_{i,j}] \in R^{N \times M}$, thus

$$\underline{s}_{N,1} = \Lambda \iota_1 \Lambda [Y - Y_d], \quad (12)$$

where $Y_d = -\Lambda \iota_1^{-1} \Lambda \iota_2 Y_L = \left[\sum_{j=1}^M \iota_{1,j} y_{L,N+j}(t), \dots, \sum_{j=1}^M \iota_{N,j} y_{L,N+j}(t) \right]^T$, and

$$\iota_{i,j} \geq 0, \sum_{j=1}^M \iota_{i,j} = 1, i = 1, \dots, N.$$

Then, design the following compensation signals

$$\begin{cases} \dot{\lambda}_{i,1} = -c_{i,1}\lambda_{i,1} + d_i(\lambda_{i,2} + \omega_{i,2} - \alpha_{i,1}), i = 1, \dots, N, \\ \dot{\lambda}_{i,j} = -c_{i,j}\lambda_{i,j} - \alpha_{i,j} + \lambda_{i,j+1} + \omega_{i,j+1} + k_{i,j}l_i^j(y_i - \hat{x}_{i,1}), \\ \dot{\lambda}_{i,n} = -c_{i,n}\lambda_{i,n} + k_{i,n}l_i^n(y_i - \hat{x}_{i,1}), c_{i,j} > 0, j = 1, \dots, n, \end{cases} \quad (13)$$

and reconstruct the coordinate transformation equations $z_{i,j} = s_{i,j} - \lambda_{i,j}$.

Finally, some Lyapunov functions for stability analysis are constructed

$$V_{i,0} = \bar{E}_{i,n}^T P_i \bar{E}_{i,n}, V_{z_{i,1}} = \frac{1}{4} z_{i,1}^4, V_{z_{i,j}} = \frac{1}{2} z_{i,j}^2, j = 2, \dots, n. \quad (14)$$

Step 1. Since $V_{z_{i,1}} = \frac{1}{4} z_{i,1}^4$ and $z_{i,1} = s_{i,1} - \lambda_{i,1}$, we have

$$\begin{aligned} LV_{z_{i,1}} &= z_{i,1}^3 \{ d_i(\alpha_{i,1} + f_{i,1} + z_{i,2} + l_i E_{i,2}) - \sum_{j \in \mathbb{N}_i^F} a_{i,j}(\hat{x}_{j,2} + l_j E_{j,2} + f_{j,1}) \\ &\quad - \sum_{j \in \mathbb{N}_i^L} a_{i,j} \dot{y}_{L,j} + c_{i,1} \lambda_{i,1} \} + \frac{3}{2} g_{i,1}^T z_{i,1}^2 g_{i,1} \\ &\leq z_{i,1}^3 (c_{i,1} \lambda_{i,1} + d_i \alpha_{i,1}) + z_{i,1}^3 H_i(X_i) + \frac{1}{4} z_{i,2}^4 + \frac{1}{2} \|\bar{E}_{i,n}\|^2 + \frac{1}{4}. \end{aligned} \quad (15)$$

where $H_i(X_i) = d_i f_{i,1} - \sum_{j \in \mathbb{N}_i^L} a_{i,j} \dot{y}_{L,j} + d_i^2 l_i^2 z_{i,1}^3 - \sum_{j \in \mathbb{N}_i^F} a_{i,j} (f_{j,1} + \hat{x}_{j,2}) + \frac{9}{4} \|g_{i,1}\|^4 z_{i,1}$
 $+ \frac{3}{4} d_i^{\frac{4}{3}} z_{i,1} - l_j^2 \left(\sum_{j \in \mathbb{N}_i^F} a_{i,j} \right)^2 z_{i,1}^3$ and
 $X_i = [z_{i,1}, \bar{x}_{i,n}^T, x_{j_1^{F_i}, 1}, \dots, x_{j_1^{F_i}, n}, \dots, x_{j_{k_i}^{F_i}, 1}, \dots, x_{j_{k_i}^{F_i}, n}, \dots, \bar{y}_{L, l_1^{L_i}}, \dots, \dot{y}_{L, l_{m_i}^{L_i}}, \hat{x}_{j_1^{L_i}, 2}, \dots, \hat{x}_{j_{k_i}^{L_i}, 2}]^T, \mathbb{N}_i^L = \{l_1^{L_i}, \dots, l_{m_i}^{L_i}\}, \mathbb{N}_i^F = \{j_1^{F_i}, \dots, j_{k_i}^{F_i}\}$. Using RBFNNs $\Im_i^{*T} \varphi(X_i)$ and error $\varepsilon_i(X_i)$ to process $H_i(X_i)$, we get

$$\begin{aligned} LV_{Z_{i,1}} &\leq z_{i,1}^3 (c_{i,1} \lambda_{i,1} + d_i \alpha_{i,1}) + \frac{3}{4} z_{i,1}^4 + \frac{1}{4} z_{i,2}^4 + \frac{1}{2} \|\bar{E}_{i,n}\|^2 + \frac{1}{4} \varepsilon_i^4(X_i) + \frac{a_{i0}^2}{2} + \frac{1}{4} \\ &\quad + \frac{z_{i,1}^6 \theta_{i,0} \|\varphi_i(X'_i)\|^2}{2a_{i0}^2} + \frac{z_{i,1}^6 \theta_{i,0}}{2a_{i0}^2} + \frac{a_{i0}^2 [\varphi_i(X_i) - \varphi_i(X'_i)]^2}{2}, \end{aligned} \quad (16)$$

where $\theta_{i,0} = \|\Im_i^*\|^2$, $X'_i = [z_{i,1}, x_{i,1}, \dot{y}_{L,j}]^T$ and $a_{i0} > 0$.

Therefore, we construct the following $\alpha_{i,1}$ and $\hat{\theta}_{i,0}$:

$$\alpha_{i,1} = -\frac{c_{i,1} s_{i,1}}{d_i} - \frac{z_{i,1}^3 \hat{\theta}_{i,0} [\|\varphi_i(X'_i)\|^2 + 1]}{2a_{i0}^2 d_i}, \quad (17)$$

$$\hat{\theta}_{i,0} = \gamma_{i0} \left(-\sigma_{i0} \hat{\theta}_{i,0} + \frac{z_{i,1}^6 [\|\varphi_i(X'_i)\|^2 + 1]}{2a_{i0}^2} \right), \gamma_{i0}, \sigma_{i0} > 0. \quad (18)$$

Step j ($2 \leq j \leq n-1$): According to (9), (13) and (14), we have

$$LV_{z_{i,j}} \leq z_{i,j} (\alpha_{i,j} + c_{i,j} \lambda_{i,j} + z_{i,j+1} - \dot{\omega}_{i,j}). \quad (19)$$

Thus,

$$\alpha_{i,j} = -c_{i,j}s_{i,j} + \dot{\omega}_{i,j}. \quad (20)$$

Step n: Design the following controller u_i

$$u_i = -c_{i,n}s_{i,n} + \dot{\omega}_{i,n}. \quad (21)$$

4 Main Results

Let

$$\Omega = \left\{ \left[\bar{E}_{N,n}^T, \bar{z}_{N,n}^T, \hat{\theta}_{N0}, \bar{\lambda}_{N,n}^T \right]^T : V \leq p \right\}, p > 0, \quad (22)$$

$$V = \sum_{i=1}^N \left(V_{i,0} + V_{z_{i,1}} + \sum_{j=2}^n V_{z_{i,j}} + \frac{1}{2} \sum_{j=1}^n \lambda_{i,j}^2 + \frac{1}{2\gamma_{i0}} \tilde{\theta}_{i0}^2 \right). \quad (23)$$

Theorem 1. Consider the SMASs consisting of system (2) subject to Assumptions 1 and 2, the virtual control laws (17) and (20), updating law (18) and controller (21). For bounded initial conditions satisfying $V(0) \leq p/n_0$ with $n_0 \geq 10$, if there exist positive design parameters $\mu_0, c_{i,j}, l_i, \gamma_{i0}, \sigma_{i0}$ satisfy (24), then all signals are SGUUB in probability.

$$\begin{cases} c_{i,1} \geq \frac{\mu_0}{4} + \frac{3}{4}, \\ c_{i,j} \geq \frac{\mu_0}{2} + 1, \\ l_i \lambda_{\min}(Q_i) \geq \frac{3}{2} + 2nF_i \|P_i\| + \mu_0 \lambda_{\max}(P_i), \\ \mu_0 \leq \min \{ \gamma_{10}\sigma_{10}, \dots, \gamma_{N0}\sigma_{N0} \}. \end{cases} \quad (24)$$

5 Simulation

Consider a class of non-strict feedback SMASs formed by 2 leaders and 4 followers, and the topology is shown in Fig. 1.

$$\begin{cases} dx_{i,1} = [x_{i,2} + 0.5\sin(3x_{i,1}x_{i,2})]dt + 0.7x_{i,1}\sin(x_{i,2}^2)dw, \\ dx_{i,2} = [u_i - 0.3\cos(x_{i,2})e^{-x_{i,1}^4}]dt + 0.5x_{i,1}x_{i,2}dw, \\ y_i = x_{i,1}, i = 1, \dots, 4, \\ y_{L,1}(t) = 0.4\sin(0.6t) + 0.9, y_{L,2}(t) = 0.5e^{-0.8t} - 0.7. \end{cases} \quad (25)$$

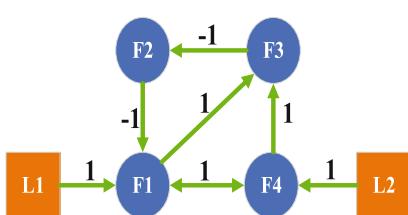


Fig. 1. Topological graph.

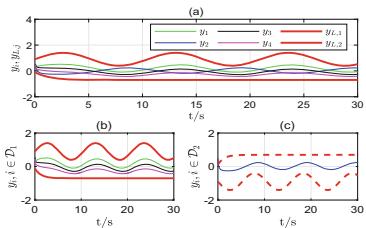


Fig. 2. Output signals y_i and $y_{L,j}$.

The simulation curves are depicted in Figs. 2, 3 and 4. Figure 2 shows the output signals of all agents, where it can be seen from (a) that the output signals of all followers are in the region formed by leaders, while (b) and (c) reflect the cooperative and antagonistic relationships between different agents. Figures 3 and 4 show the simulation results of the control signals and the transformation errors, respectively.

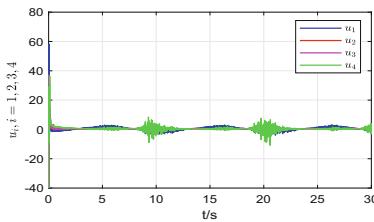


Fig. 3. Control signals u_i .

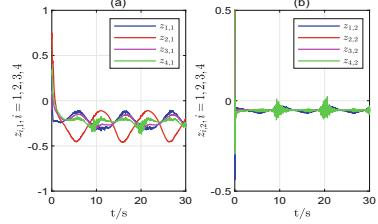


Fig. 4. Transformation errors $z_{i,1}$ and $z_{i,2}$.

6 Conclusion

A novel adaptive neural bipartite containment controller is proposed by combining the DSC method and the CF technique. First, we reconstruct the unmeasurable system state is using a high-gain observer, then introduce compensation signals to offset the filtering errors present in the conventional design, and simplify the controller design by constructing new coordinate transformations. For the unknown continuous functions generated in the design process, we approximate them by RBFNNs, but we deal with it by addition and subtraction mechanism instead of directly enlarging and reducing the selection requirement of basis function. In addition, we add all compensation signals to the final Lyapunov function to prove stability, thereby reducing some assumptions. Finally, the effectiveness of the control strategy is further verified by the experimental results.

Acknowledgements. This work was partially supported by the National Natural Science Foundation of China (62473329) and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX24_3708).

References

1. Zhang, T.P., Lin, M.F., Xia, X.N., Yi., Y.: Adaptive cooperative dynamic surface control of non-strict feedback multi-agent systems with input dead-zones and actuator failures. *Neurocomputing* **442**, 48–63 (2021).<https://doi.org/10.1016/j.neucom.2021.02.039>
2. Shen, Q.K., Shi, P.: Distributed command filtered backstepping consensus tracking control of nonlinear multiple-agent systems in strict-feedback form. *Automatica* **53**, 120–124 (2015).<https://doi.org/10.1016/j.automatica.2014.12.046>

3. Ren, C.E., Chen, L., Chen, C.L.P.: Adaptive fuzzy leader-following consensus control for stochastic multiagent systems with heterogeneous nonlinear dynamics. *IEEE Trans. Fuzzy Syst.* **25**(1), 181–190 (2017). <https://doi.org/10.1109/TFUZZ.2016.2554151>
4. Wang, W., Tong, S.C.: Adaptive fuzzy containment control of nonlinear strict-feedback systems with full state constraints. *IEEE Trans. Fuzzy Syst.* **27**(10), 2024–2038 (2019). <https://doi.org/10.1109/TFUZZ.2019.2893301>
5. Shang, K.F., Zhang, T.P., Zhang, E.Z.: Adaptive neural containment control of nonstrict-feedback multi-agent systems with unmodeled dynamics. *Int. J. Adapt. Control Signal Process.* **36**(8), 1879–1908 (2022). <https://doi.org/10.1002/acs.3426>
6. Li, K., Hua, C.C., Guan, X.P.: Distributed containment control for nonlinear stochastic multiagent systems. *IEEE Trans. Cybern.* **51**(6), 3361–3370 (2019). <https://doi.org/10.1109/TCYB.2019.2894502>
7. Yang, Y., Xi, X.R., Miao, S.T., Wu, J.R.: Event-triggered output feedback containment control for a class of stochastic nonlinear multi-agent systems. *Appl. Math. Comput.* **418**, 126817 (2022). <https://doi.org/10.1016/j.amc.2021.126817>
8. Tang, H.L., Zhang, T.P., Xia, M.Z.: Distributed adaptive finite-time output feedback containment control for nonstrict-feedback stochastic multi-agent systems via command filters. *Neurocomputing* **597**, 128021 (2024). <https://doi.org/10.1016/j.neucom.2024.128021>
9. Altafini, C.: Consensus problems on networks with antagonistic interactions. *IEEE Trans. Autom. Control* **58**(4), 935–946 (2013). <https://doi.org/10.1109/TAC.2012.2224251>
10. Wang, H., Yu, W.W., Wen, G.H., Chen, G.R.: Finite-time bipartite consensus for multi-agent systems on directed signed networks. *IEEE Trans. Circuits Syst. I* **65**(12), 4336–4348 (2018). <https://doi.org/10.1109/TCSI.2018.2838087>
11. Yang, H.J., Ye, D.: Adaptive fixed-time bipartite tracking consensus control for unknown nonlinear multi-agent systems: An information classification mechanism. *Inform. Sci.* **459**, 238–254 (2018). <https://doi.org/10.1016/j.ins.2018.04.016>
12. Chen, L.L., Shi, L., Qiu, G., Shao, J.L., Cheng, Y.H.: Bipartite containment tracking over switching signed networks. *Inform. Sci.* **612**, 959–973 (2022). <https://doi.org/10.1016/j.ins.2022.08.095>
13. Meng, D.Y.: Bipartite containment tracking of signed networks. *Automatica* **79**, 282–289 (2017). <https://doi.org/10.1016/j.automatica.2017.01.044>
14. Ren, C.E., Zhang, J.A., Guan, Y.: Prescribed performance bipartite consensus control for stochastic nonlinear multiagent systems under event-triggered strategy. *IEEE Trans. Cybern.* **53**(1), 468–482 (2023). <https://doi.org/10.1109/TCYB.2021.3119066>
15. Liang, H.J., Guo, X.Y., Pan, Y.N., Huang, T.W.: Event-triggered fuzzy bipartite tracking control for network systems based on distributed reduced-order observers. *IEEE Trans. Fuzzy Syst.* **29**(6), 1601–1614 (2021). <https://doi.org/10.1109/TFUZZ.2020.2982618>
16. Wu, Y., Pan, Y.N., Chen, M., Li, H.Y.: Quantized adaptive finite-time bipartite NN tracking control for stochastic multiagent systems. *IEEE Trans. Cybern.* **51**(6), 2870–2881 (2021). <https://doi.org/10.1109/TCYB.2020.3008020>
17. Guo, X.Y., Ma, H., Liang, H.J., Zhang, H.G.: Command-filter-based fixed-time bipartite containment control for a class of stochastic multiagent systems. *IEEE Trans. Syst. Man Cybern. Syst.* **25**(6), 3519–3529 (2022). <https://doi.org/10.1109/TSMC.2021.3072650>



Vibration-Assisted Dual-Helix Burrowing Robot for Enhanced Locomotion in Granular Media

Guancheng Zhao¹, Longchuan Li¹, Zhiqiang Song¹, Shuqian He¹,
and Shuai Kang^{2(✉)}

¹ College of Information Science and Technology, Beijing University of Chemical Technology,
Beijing 100029, China

² College of Mechanical and Electrical Engineering, Beijing University of Chemical
Technology, Beijing 100029, China
kangshuai@buct.edu.cn

Abstract. Locomotion in granular media poses significant challenges due to high resistance, heterogeneous interactions, and lift forces arising from pressure gradients. This paper presents a novel burrowing robot that combines dual-helix propulsion with a high-frequency vibration mechanism to improve mobility in such environments. The robot is equipped with a wedge-shaped head and lateral fins to mitigate lift forces and enhance vertical stability. A series of comparative experiments under varying configurations—vibration on/off and fin presence/absence—were conducted to evaluate the effects of each component. Results show that high-frequency vibration significantly reduces resistance by locally fluidizing the medium, while the lateral fins effectively suppress lift-induced drift. The integrated system demonstrated improved locomotion efficiency, consistent forward progression, and enhanced robustness against jamming. This work provides a practical and scalable approach for robotic locomotion in unstructured particulate terrains.

Keywords: Granular media locomotion · Burrowing robot · Dual-helix propulsion · High-frequency vibration

1 Introduction

With the rapid advancement of robotic technologies, mobile robots are increasingly deployed in complex environments [1]. Among these, granular media—such as sand, soil, and powders—pose unique challenges due to their complex mechanical behaviors, including high friction, cohesion, and non-Newtonian flow characteristics [2]. In such environments, conventional wheeled or tracked robots often suffer from significant resistance and low mobility efficiency, primarily because the propulsion force is strongly affected by the medium's density, particle shape, and size distribution. Designing locomotion strategies that effectively cope with these challenges has thus become a central research focus in recent years [3–5].

G. Zhao and L. Li—Contributed equally to this work.

To enhance performance in granular media, researchers have drawn inspiration from biological systems. For instance, sandfish lizards swim through sand by propagating body undulations. Inspired by this, Maladen et al. developed one of the first granular locomotion robots in 2011, utilizing sinusoidal body waves and anisotropic friction to enable self-propulsion [6]. In 2019, Liu and Ortiz independently developed robots using peristaltic locomotion—achieved through body undulation or periodic inflation—to traverse granular media [7, 8]. Drotman et al. later introduced a soft burrowing robot inspired by bristle worms, which employed asymmetrical structures to generate directional friction and forward motion using pneumatic actuation [9].

In parallel, limb-based systems have also shown promise. Chopra et al. (2023) designed a turtle-inspired robot using underactuated limbs for asymmetric propulsion and obstacle detection in granular terrain [10]. Similarly, Zhang et al. (2024) developed a mole-inspired robot for planetary soil exploration using biomimetic sweeping and thrusting motions [11]. Other designs draw from plant roots and use tip extension as a growth mechanism—Sadeghi and Biswas, for example, proposed additive-manufactured or magnetically actuated slender robots for low-disturbance burrowing [12, 13].

Overall, locomotion in granular media is commonly achieved through four principal strategies: (1) dual-anchor mechanisms using anisotropic structures and alternating anchoring-extension via pneumatic actuators; (2) peristaltic motion employing segmented chambers that cyclically expand and contract to generate traveling waves [14, 15]; (3) limb-based propulsion using flexible limbs with stiffness modulation for effective stroke-recovery cycles; and (4) undulatory motion where snake-like body waves continuously propagate along the robot's axis to generate forward thrust [16–18].

Beyond biomimetic approaches, helical propulsion has demonstrated significant potential in granular environments due to its continuous motion and relatively low disturbance. Several recent designs employ rotating helical drills for vertical or horizontal burrowing. For instance, Tang (2023) and Bagheri (2023) proposed robots equipped with rotating heads to achieve stable penetration. Additionally, flagella-like helical structures have been shown to provide directional movement through particulate media [19, 20].

However, challenges remain. Traditional helical systems often suffer from excessive resistance in dense granular media, leading to reduced efficiency. Moreover, the inhomogeneity of granular material can induce lift forces that cause trajectory deviation or unintended surfacing [21–23]. Optimizing structural design to improve both propulsion efficiency and directional stability remains an open research problem.

To address these issues, this paper presents a dual-helix burrowing robot enhanced with high-frequency vibration. The robot uses counter-rotating helical rods to generate balanced thrust, improving directional control. A high-frequency oscillator is integrated into the head to locally fluidize the surrounding particles, thereby reducing resistance [24–27]. Additionally, a wedge-shaped head and lateral fins are incorporated to redistribute pressure and counter lift forces. The proposed system is experimentally validated under various configurations to demonstrate its effectiveness in improving locomotion performance in granular media [28].

2 Mechanical Design and Locomotion Mechanism

2.1 Design Principles

As shown in Fig. 1, The robot employs two symmetrically arranged helical rods for propulsion. To ensure compactness and drive efficiency, the motors are mounted horizontally, transmitting torque through an internal gear-belt mechanism. This setup simplifies the actuation system while maintaining synchronized rotation.

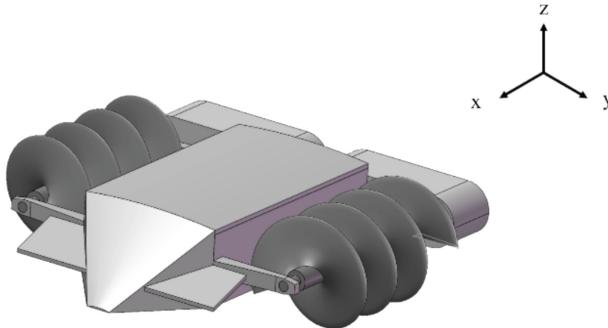


Fig. 1. Overall schematic of the robot.

Given the dense and frictional nature of granular media, effective resistance reduction is essential. The following principles were adopted:

- High-frequency internal vibration: An eccentric motor rotating about the Y-axis is embedded in the robot's head, producing vertical (Z-axis) vibrations. These vibrations locally disturb the surrounding particles, inducing a fluid-like state that reduces inter-particle friction and resistance.
- Bio-inspired deformation: The robot's front end is designed with a wedge-shaped profile inspired by sand-burrowing animals. This reduces frontal resistance and helps direct particles laterally. Additionally, lateral fins were added to suppress lift forces and maintain depth during locomotion. These structural adaptations also help minimize the robot's cross-sectional area and improve subsurface stability. The detailed design of the robot body is illustrated in Fig. 2.

2.2 Drive Structure and Locomotion Mechanism

The dual-helix system provides two main advantages over single-helix designs: (1) it cancels out angular momentum, preventing unwanted body rotation and improving directional stability, and (2) it allows for simple control—forward locomotion is achieved when the helices rotate at equal speeds in opposite directions.

The helices are driven by internal motors via a gear-belt transmission system enclosed within a protective shell to prevent particle ingress. This configuration helps maintain the center of mass within the robot body, enhancing mechanical balance.

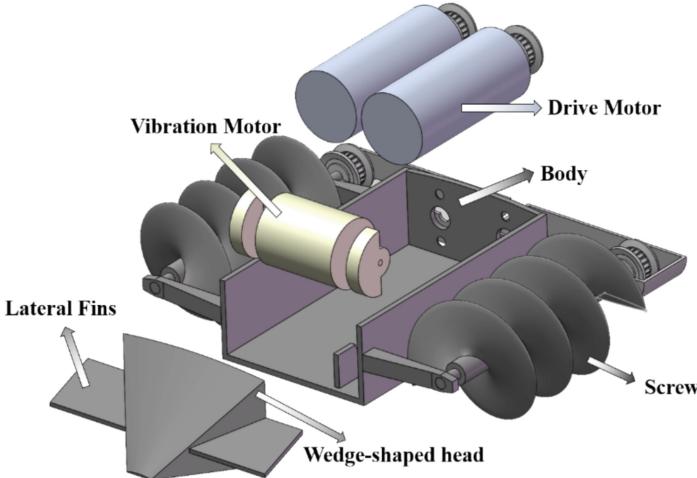


Fig. 2. Exploded view of the robot.

Compared to single-helix systems, the dual-helix configuration also provides greater surface contact, distributing thrust more evenly and reducing local pressure peaks. To ensure efficient propulsion, the helix pitch was empirically optimized. Excessively large pitches result in steep helix angles that increase radial displacement of particles, reducing thrust efficiency. Conversely, very small pitches concentrate force vertically, potentially destabilizing horizontal locomotion [29]. The selected pitch balances axial thrust, minimizes slip, and maintains efficient particle interaction.

3 Results and Analysis

3.1 Experimental Setup

A testbed measuring 650 mm × 450 mm × 400 mm was constructed and filled with polyethylene (PE) granules to a depth of 302.1 mm to simulate granular media. The particles had a diameter of approximately 4 mm, a bulk density of 0.53 g/cm³, and a material density of 0.90 g/cm³. Before each trial, the robot was manually embedded to a predefined depth. The motor speed was fixed at 22 rpm. Because internal locomotion was difficult to observe directly, a vertical indicator plate was attached to the rear to track displacement, as shown in Fig. 3.

3.2 Gait Analysis

Robot performance was evaluated under four configurations:

- VWS (Vibration + Wing + Screw)
- WS (Wing + Screw)
- VS (Vibration + Screw)
- S (Screw only)

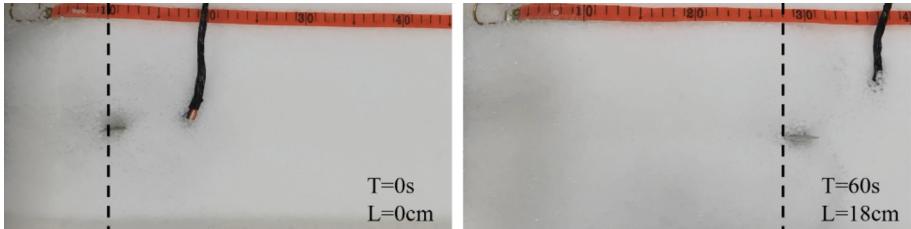


Fig. 3. Experimental setup and data acquisition method.

Each configuration was tested five times, with each trial lasting 60 s. Representative locomotion patterns are shown in Fig. 4.

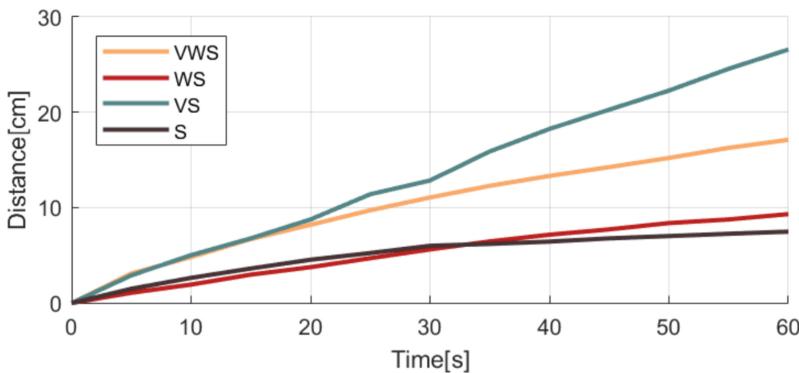


Fig. 4. Comparative results of experimental performance under different conditions.

In all cases, the robot was capable of locomotion, but with significant performance differences. In WS and S, propulsion diminished over time due to particle accumulation at the front and rear. This led to increased resistance and partial disengagement of the helical rods. In contrast, under VWS and VS conditions, high-frequency vibration improved particle distribution and prevented buildup. The combined effect of vibration and helical propulsion produced superior forward movement, as shown in Fig. 5. Statistical results with means and standard deviations are presented in Fig. 6.

3.3 Lift Mitigation

In WS and S tests, particle buildup prevented forward locomotion, masking any lift-suppressing effect of the fins. In VS, the robot gradually surfaced despite improved propulsion, indicating that the wedge-shaped head helped but was insufficient to fully counteract lift. Only in the VWS condition did the robot remain fully submerged, confirming that the addition of fins effectively mitigated lift.

To further validate the fin effect, we embedded the robot 2 cm below the surface in both VWS and VS modes, operated it for 30 s, and measured vertical displacement. As

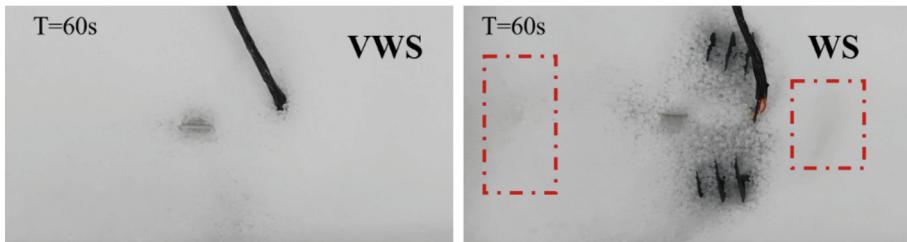


Fig. 5. Particle accumulation patterns under two different gait conditions.

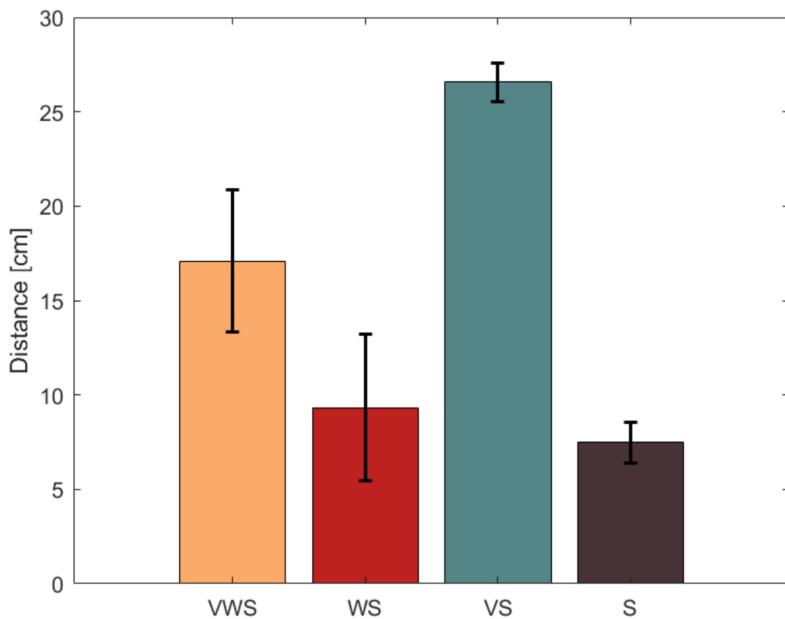


Fig. 6. Comparative analysis of locomotion performance in experiments.

shown in Fig. 7, the fins significantly reduced upward movement, confirming their role in enhancing vertical stability.

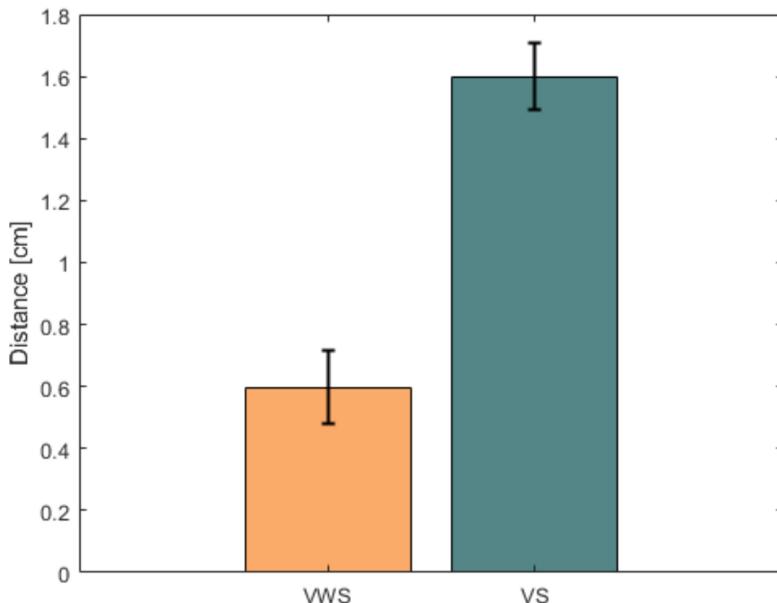


Fig. 7. Comparison of vertical displacement under two robot gait conditions.

4 Concluding Remarks

This study introduces a hybrid locomotion strategy that combines high-frequency vibration with dual-helix propulsion to enhance mobility in granular media. Unlike traditional burrowing robots that passively resist resistance, the proposed design actively modifies its environment to reduce drag, while the wedge-shaped head and lateral fins help counteract lift forces.

Experimental results demonstrate that vibration significantly improves propulsion efficiency, and that the fins contribute to maintaining depth. While these results are promising, challenges remain. The energy efficiency of the vibration system needs further improvement, and performance in more complex environments such as wet or cohesive soils requires additional investigation.

Future work will focus on optimizing vibration parameters for different media, studying the frequency-dependent behavior of various granular materials, and exploring alternative low-power actuation methods to reduce energy consumption while preserving fluidization effects.

References

1. Li, L., Asano, F., Tokuda, I.: High-speed sliding locomotion generation on slippery surface of an indirectly controlled robot with viscoelastic body. *IEEE Robot. Autom. Lett.* **4**(3), 2950–2957 (2019)
2. Miyamoto, H., et al.: Golith migration and sorting on asteroid Itokawa. *Science*. **316**(5827), 1011–1014 (2007)

3. Martinez, A., et al.: Bio-inspired geotechnical engineering: principles, current work, opportunities and challenges. *Géotechnique*. **72**(8), 687–705 (2022)
4. Hosoi, A.E., Goldman, D.I.: Beneath our feet: strategies for locomotion in granular media. *Annu. Rev. Fluid Mech.* **47**(1), 431–453 (2015)
5. Li, C., Lewis, K.: The need for and feasibility of alternative ground robots to traverse sandy and rocky extraterrestrial terrain. *Adv. Intell. Syst.* **5**(3), 2100195 (2023)
6. Maladen, R.D., Ding, Y., Umbanhowar, P.B., Goldman, D.I.: Undulatory swimming in sand: experimental and simulation studies of a robotic sandfish. *Int. J. Robot. Res.* **30**(7), 793–805 (2011)
7. Liu, B., Ozkan-Aydin, Y., Goldman, D.I., Hammond, F.L.: Kirigami skin improves soft earthworm robot anchoring and locomotion under cohesive soil. In: 2019 2nd IEEE International Conference on Soft Robotics (RoboSoft), pp. 828–833. IEEE (2019)
8. Ortiz, D., Gravish, N., Tolley, M.T.: Soft robot actuation strategies for locomotion in granular substrates. *IEEE Robot. Autom. Lett.* **4**(3), 2630–2636 (2019)
9. Drotman, D., Chopra, S., Gravish, N., Tolley, M.T.: Anisotropic forces for a worm-inspired digging robot. In: 2022 IEEE 5th International Conference on Soft Robotics (RoboSoft), pp. 261–266. IEEE (2022)
10. Chopra, S., Vasile, D., Jadhav, S., Tolley, M.T., Gravish, N.: Toward robotic sensing and swimming in granular environments using underactuated appendages. *Adv. Intell. Syst.* **5**(8), 2200404 (2023)
11. Zhang, T., et al.: Mole-inspired robot burrowing with forelimbs for planetary soil exploration. *Adv. Intell. Syst.* **6**(6), 2300392 (2024)
12. Sadeghi, A., Mondini, A., Mazzolai, B.: Toward self-growing soft robots inspired by plant roots and based on additive manufacturing technologies. *Soft Robot.* **4**(3), 211–223 (2017)
13. Biswas, A., Huynh, T., Desai, B., Moss, M., Kudrolli, A.: Dynamics of magnetoelastic robots in water-saturated granular beds. *Phys. Rev. Fluids.* **8**(9), 094304 (2023)
14. Liu, H., Guo, D., Cangelosi, A.: Embodied intelligence: a synergy of morphology, action, perception and learning. *ACM Comput. Surv.* **57**, 1–36 (2025)
15. Liu, H., Liu, X., Huang, K., Guo, D.: Embodied Intelligence. In: *Embodied Multi-Agent Systems*, pp. 3–48. Springer, Singapore (2025)
16. Li, L., Zhao, C., He, S., Qi, Q., Kang, S., Ma, S.: Enhancing undulation of soft robots in granular media: a numerical and experimental study on the effect of anisotropic scales. *Biomimetic Intell. Robot.* **4**(2), 100158 (2024)
17. Li, L., et al.: Embodying rather than encoding: towards developing a source-filter theory for undulation gait generation. *Biomimetic Intell. Robot.* **4**(3), 100173 (2024)
18. Li, L., et al.: Generation of efficient rectilinear gait based on dynamic morphological computation and its theoretical analysis. *IEEE Robot. Autom. Lett.* **6**(2), 841–848 (2021)
19. Tang, Y., Zhong, Y., Tao, J.: Bio-inspired rotational penetration and horizontal self-burrowing soft robot. *Acta Geotech.* **19**(3), 1345–1363 (2024)
20. Bagheri, H., et al.: A bio-inspired helically driven self-burrowing robot. *Acta Geotech.* **19**(3), 1435–1448 (2024)
21. Ding, Y., Gravish, N., Goldman, D.I.: Drag induced lift in granular media. *Phys. Rev. Lett.* **106**(2), 028001 (2011)
22. Maladen, R.D., Umbanhowar, P.B., Ding, Y., Masse, A., Goldman, D.I.: Granular lift forces predict vertical motion of a sand-swimming robot. In: 2011 IEEE International Conference on Robotics and Automation, pp. 1398–1403. IEEE (2011)
23. Liu, J., Wang, Z., Tian, Y., Li, L., Ma, S.: Drag force modelling with induced surface deformation in granular media. *Powder Technol.*, 120757 (2025)
24. Jeger, S.L., Wüest, V., Toumieh, C., Floreano, D.: Adaptive morphing of wing and tail for stable, resilient, and energy-efficient flight of avian-informed drones. arXiv preprint arXiv:2403.08598 (2024)

25. Nygaard, T.F., Martin, C.P., Torresen, J., Glette, K., Howard, D.: Real-world embodied AI through a morphologically adaptive quadruped robot. *Nat. Mach. Intell.* **3**(5), 410–419 (2021)
26. Zheng, Y., Asano, F., Yan, C., Li, L., Tokuda, I.T.: Tensegrity-based legged robot generates passive walking, skipping, and crawling gaits in accordance with environment. *IEEE/ASME Trans. Mechatron.* (2025)
27. Jiaxin, L., Tian, Y., Wang, Z., Li, L., Ma, S.: Modeling the impact of terrain surface deformation on drag force using discrete element method and empirical formulation. *Appl. Math. Model.* **136**, 115636 (2024)
28. Guillard, F., Forterre, Y., Pouliquen, O.: Lift forces in granular media. *Phys. Fluids.* **26**(4) (2014)
29. Karaseva, S., Belyakov, V., Makarov, V., Malahov, D.: Analysis of efficiency of full-submerged Archimedes screws of rotary-screw propulsion units of snow and swamp-going amphibious vehicles. In: VEHITS, pp. 438–444 (2022)



A Reinforcement Learning Control Method Based on Pre-training and Truncated PPO for Hand Rehabilitation

Lu Liu, Lu Sun, Benyan Huo^(✉), Yanhong Liu, and Zan Zhang

Zhengzhou University, Zhengzhou, China

sunlu@gs.zzu.edu.cn, {huoby,liuyh,zanzan}@zzu.edu.cn

Abstract. Patients who have had a stroke commonly develop upper limb motor dysfunction, particularly hand dysfunction, which significantly reduces their ability to perform Activities of Daily Living(ADL). As a neuromodulation technique, functional electrical stimulation (FES) offers a novel approach to enhancing hand-motor function. However, sophisticated hand motor control, customized stimulation program creation, and intelligent rehabilitation training are all lacking in current FES systems. This work proposes a hybrid reinforcement learning approach that uses pre-training and truncated proximal policy optimization (Truncated PPO) to build a FES based hand rehabilitation control system. A dataset representing the state-action correlation between hand joint motion and muscle stimulation is created by combining 24-channel flexible electrodes with high-precision Leap Motion sensors. The system's ability to regulate hand movements with high precision and smoothness on a virtual hand model is demonstrated by experimental findings, which also confirm the efficacy of the proposed approach by greatly enhancing training convergence speed and state similarity.

Keywords: Stroke rehabilitation · Functional electrical stimulation · Reinforcement learning · Truncated proximal policy optimization · Hand motion control

1 Introduction

Stroke commonly results in upper limb motor impairment, particularly hand dysfunction, severely impacting patients' daily activities [1]. While conventional rehabilitation methods offer some benefits [2–4], they often face practical limitations [5, 6]. Functional Electrical Stimulation (FES) presents a promising alternative, utilizing electrical currents to activate muscles and restore functional movement [7]. Recent advancements in FES include muscle synergy-based control [8, 9], feedback mechanisms [10], and Iterative Learning Control (ILC) [11, 12], aiming to enhance rehabilitation outcomes.

Despite these advances, achieving precise control over the complex dexterity of the human hand remains a significant challenge for FES systems [13]. Developing personalized stimulation strategies that adapt to individual neuromuscular conditions and recovery progress is crucial but difficult [14]. Furthermore, existing control methods often lack the adaptability required for dynamic, task-oriented rehabilitation, struggling to manage the high degrees of freedom involved in fine hand movements and adapt to changes in patient capabilities [15–17].

To overcome these limitations, this paper proposes a novel hand rehabilitation control system integrating Reinforcement Learning (RL) with FES. A hybrid RL approach is introduced, combining supervised pre-training with Truncated Proximal Policy Optimization (Truncated PPO). This method leverages expert demonstration data for initial policy learning and RL for subsequent refinement and adaptation. Key contributions include: (1) modeling muscle stimulation using high-density electrode arrays for enhanced precision; (2) a two-stage training strategy optimizing policy efficiency and effectiveness; and (3) a multi-component reward function designed to balance control accuracy, motion smoothness, and efficient exploration of stimulation patterns. The system acquires data using a 24-channel flexible electrode array and a Leap Motion sensor.

The remainder of this paper details the system design and data acquisition (Sect. 2), the RL control framework (Sect. 3), experimental setup and results (Sect. 4), and concludes with a discussion and outlook (Sect. 5).

2 System Design and Dataset Construction

2.1 System Architecture and Data Acquisition

A rehabilitation training system combining reinforcement learning (RL) and functional electrical stimulation (FES) was developed for finger rehabilitation in stroke patients (Fig. 1). The system integrates a Leap Motion sensor for angle measurement, a modular electrical stimulator, a wearable 24-channel electrode array, and a portable power supply.

2.2 Data Set Construction

A neurostimulation dataset was constructed for training the PPO model. Data collection involved synchronizing 17-dimensional hand joint motion data from Leap Motion (120 Hz) with stimulation signals from a 24-channel flexible electrode array (5 mm spacing), activating four electrodes per trial across 14 patterns. Both joint angles and pulse width signals were normalized to [0,1]. The dataset, stored in HDF5 format, comprises:

- State space: the neural state is represented by a floating point vector of length 17.
- Action space: the stimulus signal is represented as a floating-point vector of length 24.

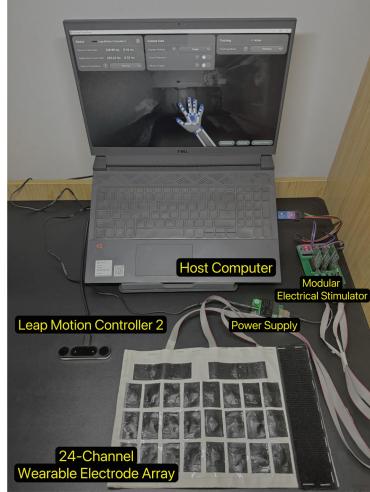


Fig. 1. System architecture of the hand rehabilitation training system based on functional electrical stimulation.

- Metadata: contains the maximum pulse width value, activation electrode ID, and combination ID as experimental reference data.

This dataset provides the state-action mapping necessary for the RL model.

2.3 Characteristics of Joint Angle Distribution

Figure 2 displays the radargram distribution of hand joint angles. Each axis represents a normalized [0,1] joint angle measurement, with colored polygons showing different hand posture samples. Thumb and index finger joints (upper region) exhibit greater variability essential for fine motor tasks, while wrist angles (lower-right region) remain stable, providing a consistent reference frame. This distribution reflects natural hand biomechanics where distal joints enable dexterity and proximal joints provide stability. The diverse motion patterns captured in this dataset are valuable for training the model to discern hand posture transitions, directly informing the design of state representations and reward functions in the reinforcement learning approach.

3 Reinforcement Learning Control Scheme Based on Pre-training and Truncated PPO

A hybrid reinforcement learning approach is proposed, combining pre-training with truncated proximal policy optimization to control hand movements. This approach aims to accurately guide the hand towards target states while maintaining motion smoothness and efficiently exploring electrode combinations.

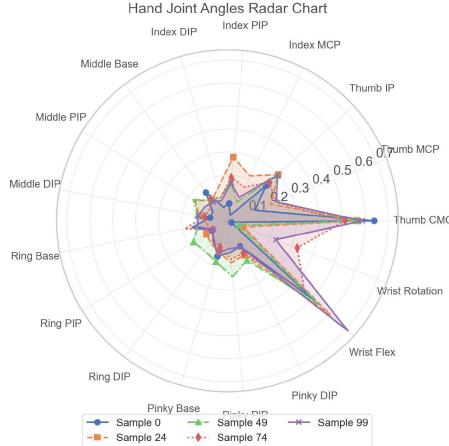


Fig. 2. Radargram distribution of hand joint angles showing the variation patterns across different samples.

3.1 Markov Decision Process Formulation

The hand control problem is formulated as a Markov Decision Process (MDP) with quintuple $M = (S, A, P, r, \gamma)$, where S represents the state space, A the action space, $P(s'|s, a)$ the state transition probability, $r(s, a)$ the reward function, and $\gamma \in [0, 1]$ the discount factor. The agent's objective is to maximize cumulative discounted rewards:

$$J(\pi) = E_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \right] \quad (1)$$

3.2 Two-Stage Training Strategy

A two-stage training method is employed to enhance policy learning efficiency:

3.2.1 Pre-training Phase

The initial policy is established through supervised learning using collected state-action pairs, minimizing the mean squared error:

$$\mathcal{L}_{\text{pretrain}}(\theta) = \frac{1}{N} \sum_{i=1}^N \|\pi_{\theta}(s_i) - a_i\|^2 \quad (2)$$

where s_i and a_i represent state and expert action, and π_{θ} denotes the policy network with parameters θ . Data augmentation techniques including Gaussian noise and state interpolation enhance generalization capability.

3.2.2 Truncated Proximal Policy Optimization

The pre-trained policy is further refined using Truncated PPO, which introduces a truncation mechanism to limit policy update magnitude, ensuring training stability while enabling exploration. The PPO objective function is:

$$\mathcal{L}^{\text{CLIP}}(\theta) = \hat{\mathbb{E}}_t[\min(\rho_t(\theta)\hat{A}_t, \text{clip}(\rho_t(\theta), 1 - \epsilon, 1 + \epsilon)\hat{A}_t)] \quad (3)$$

with $\rho_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{\text{old}}}(a_t|s_t)}$ representing the policy probability ratio, \hat{A}_t the estimated advantage, and $\epsilon = 0.2$ the clipping parameter. The complete loss function includes value function and entropy terms:

$$\mathcal{L}(\theta) = \mathcal{L}^{\text{CLIP}}(\theta) - c_1 \mathcal{L}^{\text{VF}}(\theta) + c_2 S \pi_\theta \quad (4)$$

where $c_1 = 0.5$ and $c_2 = 0.01$ are weighting coefficients.

3.3 Reward Function Design

A comprehensive integrated reward function with five components was designed to guide the learning process:

$$R_{\text{total}} = w_{\text{sim}} \cdot R_{\text{sim}} + w_{\text{explore}} \cdot R_{\text{explore}} + w_{\text{smooth}} \cdot R_{\text{smooth}} + R_{\text{success}} + R_{\text{time}} \quad (5)$$

The similarity reward evaluates proximity between current and target states:

$$R_{\text{sim}} = e^{-5 \cdot \text{mean}(|s_{\text{current}} - s_{\text{target}}|)} \quad (6)$$

The exploration reward uses the Upper Confidence Bound method:

$$R_{\text{explore}} = \sqrt{\frac{\ln(N_{\text{total}} + 1)}{N_{\text{combo}} + 1}} \quad (7)$$

The smoothness reward promotes continuous state transitions:

$$R_{\text{smooth}} = e^{-5 \cdot \text{mean}(|s_t - s_{t-1}|)} \quad (8)$$

The success and time rewards are defined as:

$$R_{\text{success}} = \begin{cases} 2.0, & \text{if } R_{\text{sim}} > 0.9 \\ 0.0, & \text{otherwise} \end{cases} \quad (9)$$

$$R_{\text{time}} = -0.001 \cdot \min(steps, 1000) \quad (10)$$

To balance exploration and exploitation, a dynamic weight adjustment strategy is implemented:

$$w_{\text{sim}} = \min(0.5 + \frac{\text{episode}}{\text{max_episodes} \cdot 0.5}, 0.7) \quad (11)$$

$$w_{explore} = \max(0.4 - \frac{episode}{max_episodes \cdot 0.5}, 0.2) \quad (12)$$

with $w_{smooth} = 0.1$ remaining constant. This approach sets the exploration reward weight high at the beginning of training to promote thorough exploration, then gradually increases the state similarity reward weight to shift the model from exploration to exploitation as training progresses.

Additionally, an adaptive tolerance mechanism was implemented that progressively raises the model's accuracy requirements during training, encouraging continuous improvement. This multi-component, dynamically adjustable reward function design guides the model to learn precise electrode stimulation parameter tuning, providing stroke patients with a more effective rehabilitation training experience.

The complete training algorithm integrating the above reward function design is outlined in Algorithm 1, which summarizes our two-stage approach combining supervised pre-training with reinforcement learning optimization.

Algorithm 1. Two-Stage Hand Control with Pretraining and PPO

```

1: Input: Dataset  $\mathcal{D}$  with state-action pairs
2: Output: Trained policy  $\pi_\theta$ 
3: procedure PRETRAINING( $\mathcal{D}$ )
4:   Initialize encoder-decoder model
5:   Train model using MSE loss with data augmentation
6:   return pretrained weights
7: end procedure
8: procedure PPOTRAINING(pretrained weights)
9:   Initialize policy with pretrained weights
10:  for episode = 1 to  $N_{episodes}$  do
11:    Sample target state, update reward weights
12:    Execute policy, collect transitions
13:    Update policy using clipped objective when buffer full
14:  end for
15:  return optimized policy
16: end procedure

```

4 Experimental Setup and Result Analysis

Systematic experiments were conducted to validate the effectiveness of the proposed pre-training and truncated PPO-based control algorithm for hand rehabilitation. The experiments evaluate training convergence, control accuracy, motion smoothness, and exploration efficiency.

4.1 Experimental Setup

The virtual hand model was implemented with a 17-dimensional state space representing joint positions and angles, and a 24-dimensional action space representing electrode stimulation parameters. All values were normalized to [0,1] for training and reward computation.

Training followed the two-stage approach described in Algorithm 1. The pre-training model used an encoder-decoder architecture with two linear layers ($17 \rightarrow 128 \rightarrow 128$) in the encoder and two linear layers ($128 \rightarrow 128 \rightarrow 24$) in the decoder. The PPO policy network maintained the encoder structure with additional layers for action mean, action standard deviation, and value function output. The key hyperparameters used in the training process are summarized in (Table 1).

Table 1. Key Hyperparameters for Training

Parameter	Value	Parameter	Value
State dimension	17	Learning rate (PPO)	3×10^{-4}
Action dimension	24	Learning rate (Pre)	1×10^{-3}
Hidden layer size	128	Discount factor (γ)	0.99
PPO epochs	4	Clip parameter (ϵ)	0.2
Batch size (PPO)	64	Batch size (Pre)	32
Max episodes	2000	Buffer size	2000

Data collection utilized a custom 24-channel FES system and Leap Motion Controller (120 Hz), with a flexible electrode array (5mm pitch) attached to the subject's forearm. All participants provided informed consent under an approved ethical protocol.

4.2 Results and Analysis

4.2.1 Training Performance

Figure 3(a) shows the pre-training model achieved stable convergence after approximately 500 rounds, with MSE loss decreasing from 0.5 to below 0.1. This effective state-action mapping provided a solid foundation for subsequent PPO training, significantly accelerating the overall process. The reward value increased rapidly during the first 500 rounds before entering a smooth optimization phase.

Figure 3(b) demonstrates how state similarity improved throughout training, rapidly increasing from 0.5 to 0.8 in the first 500 rounds, then gradually stabilizing above 0.9 (success threshold). The decreasing volatility, particularly in later training stages, indicates improved reliability and consistency.

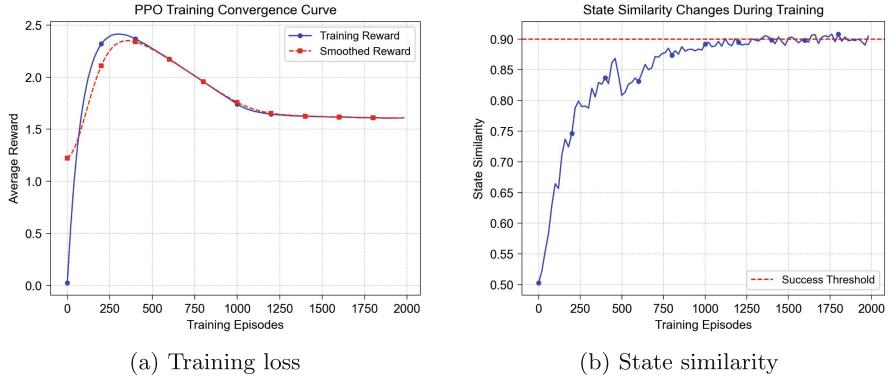


Fig. 3. Model training performance metrics

4.2.2 Reward Function Dynamics

Figure 4(a) illustrates the dynamic adjustment of reward component weights, with smoothness weight constant at 0.1, exploration weight decreasing from 0.4 to 0.2, and similarity weight increasing from 0.5 to 0.7. This strategy effectively balanced exploration and exploitation, focusing on electrode combination exploration in early stages and precise control later.

Figure 4(b) shows the pre-training MSE loss rapidly decreasing from 0.5 to below 0.1 around 200 rounds, eventually reaching 0.05. This confirms successful acquisition of the state-action mapping and demonstrates the effectiveness of data augmentation in preventing overfitting.

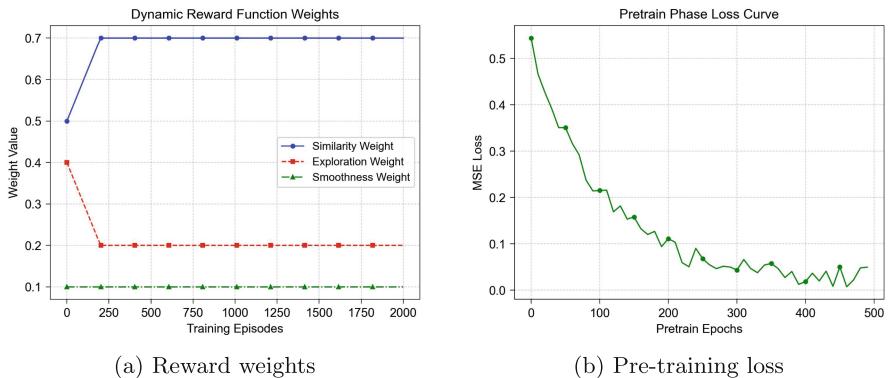


Fig. 4. Training dynamics

Table 2. Key Performance Metrics Across Training Phases

Metric	Early	Middle	Late	Change
Smoothness score	0.65	0.75	0.83	+27.7%
Combination entropy	2.3	3.1	3.8	+65.2%
Success rate	45%	78%	93%	+106.7%
Task completion steps	150	110	85	-43.3%
State similarity	0.65	0.82	0.94	+44.6%

Table 2 summarizes key performance indicators across training phases. The combination distribution entropy increased from 2.3 to 3.8, indicating the algorithm’s growing understanding of electrode-muscle response relationships. Movement smoothness improved by 27.7%, which is crucial for reducing muscle fatigue in FES applications.

The success rate (defined by state similarity > 0.9) increased from 45% to 93%, while average task completion steps decreased by 43.3%. These metrics demonstrate the method’s ability to learn precise hand control strategies essential for FES rehabilitation applications.

These results validate the effectiveness of the proposed system, though limitations remain. The approach was only tested in virtual environments, not accounting for muscle fatigue and adaptive changes with prolonged use. Additionally, pre-training depends on high-quality expert data, and the PPO method has relatively high computational complexity.

5 Discussion and Outlook

This work proposed a hybrid reinforcement learning approach that uses pre-training and truncated proximal policy optimization (Truncated PPO) to build an FES-based hand rehabilitation control system.

However, several limitations exist. The system was validated solely in a virtual environment, potentially differing from real-world FES applications. Factors like muscle fatigue and long-term adaptive changes were not considered. Furthermore, the pre-training phase relies on high-quality expert data, and the PPO algorithm entails considerable computational complexity.

Future research will focus on validating the proposed method on a physical FES system to assess real-world efficacy. Future work includes developing online adaptation algorithms to address individual variability and muscle fatigue, enhancing long-term usability. Investigating more efficient network architectures, training techniques, and model compression strategies is also crucial to reduce computational demands, facilitating deployment on resource-constrained embedded systems.

References

1. Tang, Q., et al.: Research trends and hotspots of post-stroke upper limb dysfunction: a bibliometric and visualization analysis. *Front. Neurol.* **15**, 1449729 (2024)
2. Bütefisch, C., Hummelsheim, H., Denzler, P., Mauritz, K.H.: Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand. *J. Neurol. Sci.* **130**(1), 59–68 (1995)
3. Charles, J., Gordon, A.M.: Development of hand-arm bimanual intensive training (habit) for improving bimanual coordination in children with hemiplegic cerebral palsy. *Dev. Med. Child Neurol.* **48**(11), 931–936 (2006)
4. Comelia, C.L., Stebbins, G.T., Brown-Toms, N., Goetz, C.G.: Physical therapy and parkinson's disease: a controlled clinical trial. *Neurology* **44**(3_part_1), 376–376 (1994)
5. Nam, C., et al.: Home-based self-help telerehabilitation of the upper limb assisted by an electromyography-driven wrist/hand exoneuromusculoskeleton after stroke. *J. Neuroeng. Rehabil.* **18**, 1–18 (2021)
6. Catanese, L., Tarsia, J., Fisher, M.: Acute ischemic stroke therapy overview. *Circ. Res.* **120**(3), 541–558 (2017)
7. Oujamaa, L., Relave, I., Froger, J., Mottet, D., Pelissier, J.Y.: Rehabilitation of arm function after stroke. Literature review. *Ann. Phys. Rehabil. Med.* **52**(3), 269–293 (2009)
8. d'Avella, A., Portone, A., Fernandez, L., Lacquaniti, F.: Control of fast-reaching movements by muscle synergy combinations. *J. Neurosci.* **26**(30), 7791–7810 (2006)
9. Niu, C.M., et al.: Synergy-based fes for post-stroke rehabilitation of upper-limb motor functions. *IEEE Trans. Neural Syst. Rehabil. Eng.* **27**(2), 256–264 (2019)
10. Razavian, R.S., Ghannadi, B., Mehrabi, N., Charlet, M., McPhee, J.: Feedback control of functional electrical stimulation for 2-D arm reaching movements. *IEEE Trans. Neural Syst. Rehabil. Eng.* **26**(10), 2033–2043 (2018)
11. Wannawas, N., Diaz-Pintado, C., Narayan, J., Faisal, A.A.: Controlling fes of arm movements using physics-informed reinforcement learning via co-kriging adjustment. In: 2024 IEEE International Conference on Robotics and Automation (ICRA), pp. 5555–5560. IEEE (2024)
12. Freeman, C.T., Hughes, A.M., Burridge, J.H., Chappell, P., Lewin, P., Rogers, E.: Iterative learning control of fes applied to the upper extremity for rehabilitation. *Control. Eng. Pract.* **17**(3), 368–381 (2009)
13. Peckham, P.H., Knutson, J.S.: Functional electrical stimulation for neuromuscular applications. *Annu. Rev. Biomed. Eng.* **7**(1), 327–360 (2005)
14. Sheng, Y., et al.: Upper limb motor function quantification in post-stroke rehabilitation using muscle synergy space model. *IEEE Trans. Biomed. Eng.* **69**(10), 3119–3130 (2022)
15. Zhou, J., Freeman, C.T., Holderbaum, W.: Multiple model iterative learning control for fes-based stroke rehabilitation. In: 2023 American Control Conference (ACC), pp. 2147–2152. IEEE (2023)
16. Meadmore, K., et al.: Electrical stimulation and iterative learning control for functional recovery in the upper limb post-stroke. In: 2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR), pp. 1–5. IEEE (2013)
17. Dou, H., Tan, K.K., Lee, T.H., Zhou, Z.: Iterative learning feedback control of human limbs via functional electrical stimulation. *Control. Eng. Pract.* **7**(3), 315–325 (1999)



Non-contact Heart Rate Estimation Based on U-Net Global Facial Video

Cheng Chen¹, Yurun Ma^{1(✉)}, Jiaqi Li², and Yudan Liu¹

¹ School of Electrical Engineering and Information Engineering, Lanzhou University of Technology, 730050 Lanzhou, China

mayr@lut.edu.cn

² Gansu Changfeng Electronic Technology Co., Ltd., 730070 Lanzhou, China

Abstract. Remote Photo-plethysmography (rPPG) is a method to obtain blood volume changes based on the principle of diffuse reflection of natural light on human skin. It can obtain physiological information such as human photoplethysmography signal and heart rate non-invasively and without contact. This research proposes a non-contact heart rate estimation method based on U-Net global facial segmentation to solve the problem that existing rPPG detection methods are easily affected by ambient light and individual movement, and fixed facial areas of interest can easily lead to inaccurate heart rate estimation. This method extracts full facial skin from the video through the U-Net segmentation network, then performs skin pixel reorganization, and extracts the green channel information of the reorganized skin to estimate individual heart rate. To verify the effectiveness and accuracy of this method, testing, and ablation experiments were conducted on the UBFC dataset. The average absolute error, root mean square error, and Pearson correlation coefficient were 4.52 bpm, 7.04 bpm, and 0.93 respectively.

Keywords: Remote Photo-plethysmography · Heart Rate · segmentation

1 Introduction

Heart rate (HR) is one of the important vital signs, and accurate estimation of HR is crucial for health management, disease diagnosis, and clinical research [1, 2]. In contrast to traditional contact HR measurement methods, which are based on electrocardiography and photoplethysmography, non-contact HR estimation methods are based on remote Photo-plethysmography (rPPG), which analyzes facial video information to estimate the heart rate without wearing contact sensors [3, 4]. As a result, its applications are becoming more and more widespread in the security field, atrial fibrillation detection [5], driver condition assessment [6], and critical care monitoring [7]. In recent years, non-contact heart rate estimation has become a research hotspot in the discipline of Instrumentation and Measurement (I&M) [8, 9].

More and more rPPG methods for estimating HR and other physiological signals directly from videos via deep learning approaches have been developed. Qiu et al. [10] automatically extracted ROIs and combined spatial and temporal filtering techniques

and CNNs to achieve end-to-end non-contact heart rate estimation; Chen et al. [11] proposed DeepPhys, which is capable of recovering HR and respiratory rate information from RGB or infrared video to recover HR and respiration rate information, and also supports spatial-temporal visualization of physiological information in the video; Liu et al. [12] used Multi-task Time Offset Convolutional Attention Network (MTTS-CAN) for end-to-end non-contact HR estimation. Their shortcoming, however, is that the end-to-end approach is susceptible to complex scenarios, such as individual movements and changes in illumination conditions, which may interfere with feature extraction during model learning, making it difficult to exclude features that are not related to rPPG. This not only leads to performance degradation during training but also usually requires a large number of training samples to achieve reliable quality assessment performance due to the large number of deep network parameters.

Although researchers have proposed several effective non-contact HR measurement methods, these methods still face several limitations. The main issues include (1) manual or semi-automatic ROI selection that relies on a priori knowledge is prone to inaccurate HR estimation due to missing facial information, (2) a high pixel percentage of non-skin regions increases the noise in the non-contact photoplethysmography signals, and (3) the end-to-end network may extract information that is not related to the rPPG, which raises the complexity of the computational process. Therefore, accurately extracting rPPG information from the video to estimate HR remains a challenge in the measurement and monitoring (I&M) discipline.

The main contributions of this paper are (1) the detection of rPPG signals is highly dependent on the accurate acquisition of images of the facial skin region. A three-stage processing method is proposed: firstly, the temporal image of facial skin is extracted, followed by skin pixel reorganization, and finally, features are extracted to estimate the HR; (2) in the facial skin extraction stage, a U-Net network is designed and applied to perform a global facial skin adaptive segmentation, which preserves the physiological information that contains only the facial skin; (3) by removing black pixels in the segmented image and reorganizing the remaining skin pixels are reorganized to construct a new image, which not only improves the computational efficiency but also ensures that only the physiological information of facial skin is preserved. Such image-processing steps effectively lay the foundation for subsequent feature extraction and HR estimation.

2 Methodology

In this thesis, we propose a three-stage approach that aims to reduce the loss of physiological information, reduce the noise in physiological information, and avoid the interference of irrelevant physiological information. As shown in Fig. 1, the basic process framework of rPPG extraction and HR estimation in this paper includes U-Net segmentation of images [13], pixel removal of non-skin regions and image reconstruction, and signal filtering and baseline drift removal. Specifically, in the first stage, facial skin regions are accurately extracted by the U-Net image segmentation network, thereby segmenting a time-series image containing only facial skin from a complex background. In the second stage, non-skin region pixels will be removed to realize the reconstruction of the time-series image. In the third stage, HR_{pre} , the predicted value of non-contact pulse

HR, is obtained by extracting the features of the green channel of the image and combining them with signal filtering, motion artifact removal processing, and peak detection. Meanwhile, HR_{gt} , the HR of the labeled data obtained by peak detection, is used as a reference.

2.1 U-Net Extraction of Facial Skin

The structure of the U-Net network is shown in Fig. 2. The segmentation principle of U-Net is to utilize CNN to perform in-depth feature extraction on the image and accurately segment the image by its unique symmetric U-shaped structure. Specifically, it includes the following key steps: (1) Feature extraction: in the contraction path, a series of convolutional layers and pooling layers are used to extract features from the input image. Each convolutional layer usually includes a convolution operation, an activation function (ReLU), and a batch normalization. The convolution operation is denoted as:

$$F_{i,j}^{(l)} = \phi \left(\sum_{m,n} K_{m,n}^{(l)} \cdot X_{i+m,j+n}^{(l-1)} + b^{(l)} \right) \quad (1)$$

Where $F_{i,j}^{(l)}$ denotes the feature map output of layer l at position (i, j), $K_{m,n}^{(l)}$ is the weight of the convolution kernel, $X_{i+m,j+n}^{(l-1)}$ is the input of the previous layer, $K_{m,n}^{(l)}$ is the bias term, and ϕ is the ReLU activation function. (2) Spatial information compression: after each convolutional block, a maximum pooling operation is used to reduce the spatial dimensions of the feature map and increase the sensory field to capture a wider range of contextual information. The pooling operation is denoted as:

$$P_{i,j}^{(l)} = \max_{a,b \in W} X_{i+a,j+b}^{(l)} \quad (2)$$

Where $P_{i,j}^{(l)}$ denotes the pooled feature map, W is the pooling window and $X_{i+a,j+b}^{(l)}$ is the input feature value in the window. (3) Feature map upsampling and merging: in the expansion path, the spatial dimension of the feature map is gradually increased using transpose convolution, while the feature maps of the corresponding layers in the contraction path are merged with the current layer's feature map through jump connections to recover image details. The transpose convolution operation is denoted as:

$$U_{i,j}^{(l)} = \sum_{m,n} K_{m,n}^{(l)} \cdot F_{i-m,j-n}^{(l-1)} \quad (3)$$

Where, $U_{i,j}^{(l)}$ denotes the feature map after upsampling, $K_{m,n}^{(l)}$ is the weight of the transposed convolution kernel, and $F_{i-m,j-n}^{(l-1)}$ is the feature map of the previous layer. (4) Pixel-level classification: finally, the multi-channel feature maps are converted into probability maps corresponding to the classification of each pixel using a 1×1 convolution. This step can be viewed as classifying each pixel, and the convolution operation of the output layer is denoted as:

$$S_{i,j}^{(l)} = \sigma \left(\sum_{m,n} K_{m,n}^{(l)} \cdot U_{i+m,j+n}^{(l-1)} + b^{(l)} \right) \quad (4)$$

Where, $S_{i,j}^{(l)}$ denotes the final segmentation map, the classification probability of each pixel at position (i,j), σ is the ReLU activation function, $K_{m,n}^{(l)}$ is the weight of position (m,n) in the convolution kernel of the lth layer, $U_{i+m,j+n}^{(l-1)}$ is the value of the input feature map (i.e., the output of the previous layer) of the lth layer at position (i + m, j + n), and $b^{(l)}$ is the bias term of the lth layer.

The detailed parameter settings of the U-Net network are visible in Fig. 2. The input data is the original facial image while the annotation by using the labelme tool contains detailed information about the facial skin region, hair, and background. This study aims to accurately extract the facial skin regions from images containing complex backgrounds using precise image segmentation techniques, while effectively excluding the influence of hair regions. This process aims to minimize the loss of physiological information and noise interference while mitigating the negative impact of facial motion on signal quality.

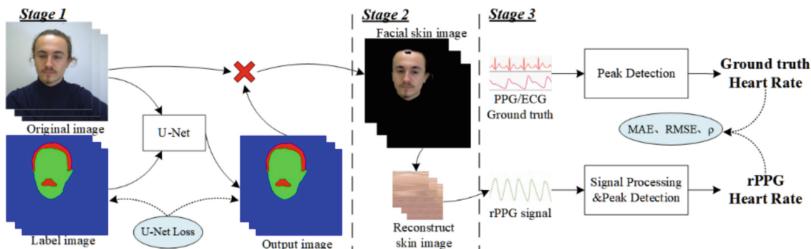


Fig. 1. General framework of the proposed two-stage approach. Stage 1: Facial skin region images are extracted by the U-Net network. Stage 2: Removal of non-skin region pixels and feature extraction. Stage 3: Non-contact heart rate measurement by signal processing and peak detection.

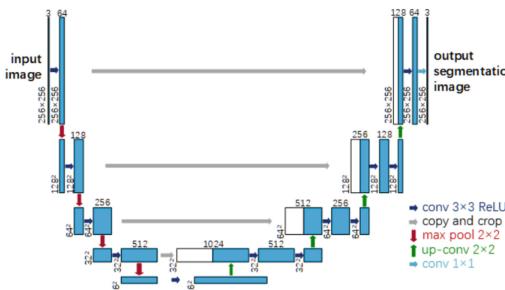


Fig. 2. U-Net structure diagram.

Then, by training a large number of U-Net on the images labeled with skin regions in the dataset, this study successfully generates segmented images that highly match the skin labeling map. Compared to methods using face detection or skin color space thresholding [14, 15], the latter often fails to provide comprehensive coverage when dealing with facial images and is prone to missing key physiological information. In

addition, the U-Net network can recognize color-treated hair, which is difficult to do with traditional skin color thresholding methods, especially when distinguishing yellow-dyed hair from skin color. Thus, U-Net provides us with significant advantages in terms of accuracy and functionality.

The loss of U-Net uses the BCE loss function, denoted as:

$$L(O_{img}, L_{img}) = -\omega * (L_{img} * \ln(O_{img}) + (1 - L_{img}) * \ln(1 - O_{img})) \quad (5)$$

Where ω is the weight value, generally 1, O_{img} is the original image data (Original image), which is the model prediction value, L_{img} is the image data labeled with facial skin, hair, and background using labelme (Label image), which is the label value, and the size between O_{img} and L_{img} is kept consistent. Eq. (5) is for a single sample when a batch has N samples, then the loss function of U-Net:

$$L_{U-Net} = \frac{1}{N} \sum L \quad (6)$$

After training, the complete facial region image can be segmented using the U-Net network for further non-contact HR estimation.

2.2 Image Reconstruction of Facial Regions

The underlying principle of non-contact pulse signaling technology is that regular heart-beat activity leads to a periodic flow of blood through the blood vessels, which in turn leads to a regular flow of blood under the skin. Particularly on the face, where capillaries are richly distributed, these changes in blood flow cause subtle and regular changes in the color of the facial skin. These color changes are difficult to capture with the naked eye and are therefore detected with a sensor digital camera. The basic principle of this measurement technique can be seen in Fig. 3.

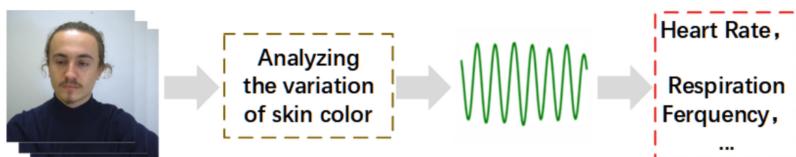


Fig. 3. Basic Principle Diagram of Non-Contact Pulse Detection and Heart Rate Estimation.

First, the facial image is processed using the U-Net network to generate a mask and apply it to the original image as a way of segmenting the facial region in the video and setting the pixel values of the non-facial background to zero. Further, a pixel reorganization operation is performed on the U-Net processed image to remove the black pixels in the image and automatically rearrange the remaining pixels to form a time series of facial skin reconstruction maps with consistent dimensions, this process is shown in Fig. 4.

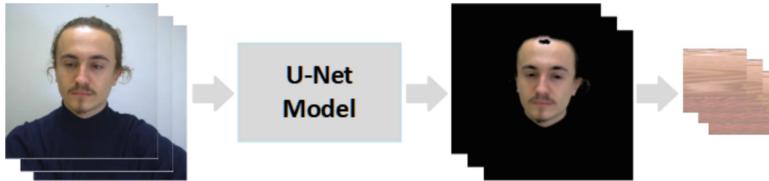


Fig. 4. Skin reconfiguration diagram.

2.3 Heart Rate Estimation

With the reconstructed time-series images, we were able to extract the preliminary signals of the non-contact pulse signals. Literature studies have shown that the green channel of the image is rich in physiologic signals. Although the R, G, and B channels all reflect the volumetric tracing map signal, the signal in the G channel is the most significant. This phenomenon is consistent with the fact that green light is more readily absorbed by hemoglobin compared to red light and that green light can detect vascular structures deeper into the skin compared to blue light. Preliminary tests also confirmed the superiority of the green channel in extracting photoplethysmography signals, and therefore, in the framework of this study, the green channel was chosen as the primary color channel for HR information extraction. This choice is based on the biological and optical properties of green light, which is more effective in extracting HR information from facial videos. After obtaining the skin reorganization image of each video frame, we calculate the mean value of its green channel, denoted as G_{mean} . the variation in heart rate is obtained by analyzing the G_{mean} curve. Where G_{mean} is defined as follows:

$$G_{mean} = \frac{\sum_{i=1}^{H*W} G_i}{H*W} \quad (7)$$

Where G_i is the pixel value of each pixel of the green channel of the image, where i is the position of each pixel, H is the height of the image, W is the width of the image, $H*W$ is the number of pixels of the image, and G_{mean} is the sum of all pixels of the green channel divided by the number of pixels of the image.

To obtain accurate HR variation, median filtering is applied to the G_{mean} curve in this paper. The formula for median filtering is shown in Eq. (8):

$$y_i = Med\{f_{i-v}, \dots, f_i, \dots, f_{i+v}\} \quad i \in N, v = \frac{m-1}{2} \quad (8)$$

A one-dimensional sequence f_1, f_2, \dots, f_n , takes the length of the window m (m is an odd number), its median filtering, that is, from the input sequence of the successive removal of m numbers, in the size of the m number of sorting, take the number of its serial number for the center of the number of that number as the filtered output.

In addition, the variation of the non-contact photoplethysmography signal is also caused by slight head movements, therefore, in this thesis, polynomial fitting is used to remove the baseline drift of the G_{mean} curve caused by respiration and slight head

movements. The equation for removing baseline drift:

$$P(t) = O(t) - B(t) \quad (9)$$

Where $P(t)$ is the processed data after removing baseline drift, $O(t)$ is the original data, and $B(t)$ is the baseline, which is obtained by polynomial fitting. Specifically, the baseline $B(t)$ is an nth-order polynomial that can be expressed as:

$$B(t) = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \cdots + a_2 \cdot t^2 + a_1 \cdot t + a_0 \quad (10)$$

Here t denotes the time data and a_n through a_0 are the coefficients obtained from the polynomial fit. The polynomial is of order 10, so $n = 10$. The baseline $B(t)$ is subtracted from the original data $O(t)$ and the resulting $P(t)$ is the signal with the baseline drift removed. Once the signal has been processed it is possible to calculate the heart rate of the non-contact pulse signal, using peak detection to estimate the non-contact pulse heart rate, and also to estimate the peak of the wrist pulse signal that was simultaneously acquired in the dataset.

3 Experiments and Analysis of Results

First, the publicly available UBFC dataset is introduced; subsequently, the evaluation metrics and detailed experimental setup are described; finally, the research results are presented through experimental results and analysis. By comparing with other non-contact pulse heart rate estimation methods, this study validates the efficiency and accuracy of our proposed method for non-contact ‘photoplethysmography signal detection and analysis.

3.1 Datasets

UBFC-rPPG dataset: the UBFC-rPPG dataset [16] contains 54 visible videos captured by the Logitech C920 HD Pro camera in ideal and HCI scenarios; 49 subjects participated in the data collection process. The frame rate and resolution of the videos were 30 fps and 640×480 , respectively. All videos were collected indoors with variable lighting conditions. A pulse oximeter (CMS50E) was used to record PPG waveforms and HR during video collection.

3.2 Evaluation Indicators

To evaluate the performance of different methods, we used the following widely adopted metrics in non-contact HR measurements: mean absolute error (MAE), root mean square error (RMSE), and Pearson correlation (Pearson correlation, ρ). Lower values of MAE and RMSE indicate better estimation, i.e., the estimated HR is closer to the true value. The ρ higher the value, the stronger their linear correlation.

Mean absolute error of heart rate HR_{mae}

$$HR_{mae} = \frac{1}{n} \sum_{i=1}^n |\hat{h}_i - h_i| \quad (11)$$

Root mean square error of heart rate HR_{rmse}

$$HR_{rmse} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{h}_i - h_i)^2} \quad (12)$$

Pearson correlation coefficient

$$\rho = \frac{\sum_{i=1}^n (h_i - \bar{h})(\hat{h}_i - \bar{\hat{h}})}{\sqrt{\sum_{i=1}^n (h_i - \bar{h})^2} \sqrt{\sum_{i=1}^n (\hat{h}_i - \bar{\hat{h}})^2}} \quad (13)$$

3.3 Experimental Setup

The experiment is roughly divided into two parts, the first part: pre-processing the dataset before training, to reduce the amount of computation and to make the U-Net network input match, the image of 640×480 resolution is adjusted to the image of 256×256 resolution, and then the image is fed into the U-Net network for the segmentation training of the facial skin region of the human face. After a large number of facial skin region segmentation training, the best model of the U-Net network is saved. Part II: The skin segmented image is subjected to black pixel removal and reconstructed image, and the pixel average of the green channel of the reconstructed sequence image is calculated. The rPPG signal can be obtained after a series of signal processing. the training epoch of the U-Net is set to 200, and the batch size is 20. The training process adopts the Adam optimization algorithm with a learning rate of 10^{-4} .

All the experiments were run on a machine with a Windows 10 operating system using an NVIDIA GeForce RTX 3090 graphics card. The processor is Intel I9 13900KF and U-Net is implemented using Pytorch deep learning framework.

3.4 Experimental Results and Analysis

U-Net Effect. Through a large number of training U-Net network can be well segmented facial skin region image, segmentation accuracy reached 99%, segmentation effect graph is shown in Fig. 5.

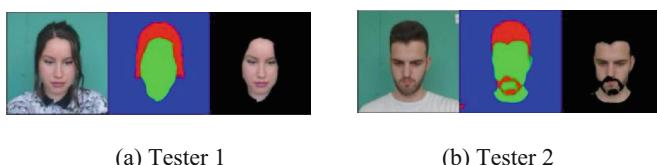


Fig. 5. Effect of U-Net segmentation of UBFC dataset (left: original image, middle: U-Net output, right: skin segmentation).

HR Extraction Results. Figure 6 shows the rPPG waveform estimated by our proposed method of labeling the PPG signal. Where the PPG signal is used as a reference standard for rPPG, the PPG signal was recorded using a pulse oximeter (CMS50E) during video acquisition of the UBFC dataset.

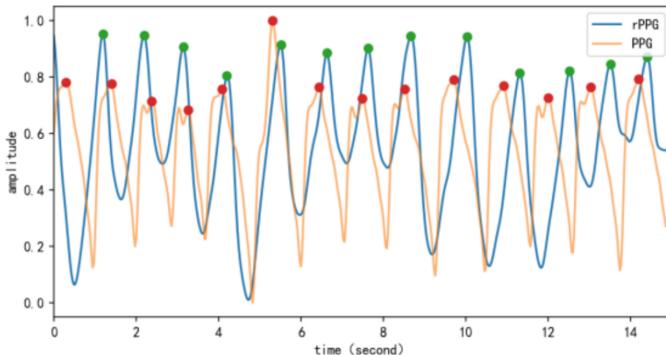


Fig. 6. Finger pulse signal (yellow) and non-contact pulse signal (blue).

We further evaluated the validity of our proposed method by comparing it with other non-contact HR measurement methods. The following methods were selected for comparison: traditional GREEN, CHROM, POS, ICA methods, and deep learning methods: SynRhythm, 3DCNN, CVD, Meta-rPPG. All methods were trained and tested on the same dataset. The results are summarized in Table 1.

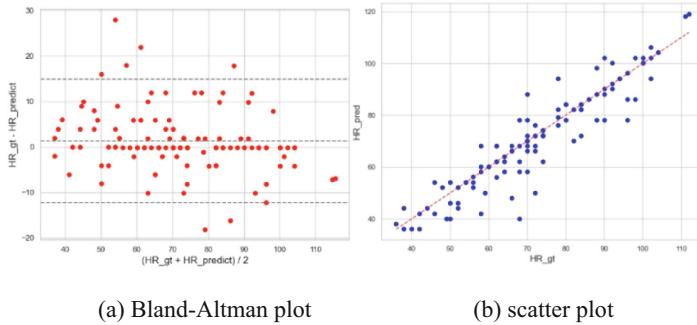
From the experimental results, it is observed that the method proposed in this paper can extract the non-contact photoplethysmography signals from the facial skin region in a targeted manner, and effectively overcome the noise-induced bias by enhancing the correlation between neighboring frames. Overall, the method performs well on the dataset with MAE and RMSE of 4.52 bpm and 7.04 bpm, respectively, which is an improvement compared to the traditional method. The Pearson correlation coefficient reaches 0.93, indicating that the method reaches the performance level of some current end-to-end network models; this not only demonstrates the robustness of the model and its excellent generalization ability but also reduces the computational complexity.

Validation of HR Extraction Validity. Figure 7 shows the Bland-Altman and scatter plot results for the dataset, where the non-contact heart rate estimate is denoted as $\text{HR}_{\text{predict}}$ and the finger pulse signal HR estimate is denoted as HR_{gt} . As shown in the figure, on the UBFC dataset, the distributions are centered on a mean deviation of -0.5 beats per minute (bpm), with a 95% consistency limit in the range of -7.8 to 6.5 bpm, and the predicted HR has a very good correlation with the labeled HR in the range of 50 to 120 bpm. The experimental results show that the proposed model can meet this criterion.

To deeply analyze the effectiveness of the proposed method in this research, we further visualized the HR estimation results on the UBFC dataset, demonstrating the confusion matrix as shown in Fig. 8. Given that the confusion matrix is mainly applied

Table 1. Comparison of HR estimation results based on UBFC

Method	MAE↓(bpm)	RMSE↓(bpm)	p↑
3DCNN [17]	8.35	10.45	0.76
CVD [18]	7.85	9.05	0.79
Meta-rPPG [19]	5.97	7.42	0.53
GREEN [20]	8.33	10.88	0.48
ICA [21]	5.63	8.53	0.71
CHROM [22]	6.69	8.82	0.82
POS [23]	7.80	11.44	0.62
SynRhythm [24]	5.59	6.82	0.72
Ours	4.52	7.04	0.93

**Fig. 7.** Plot of the difference between HR_gt and HR_predict on the UBFC dataset.

to analyze the results of the classification task, we categorized the HR into 8 intervals ranging from 40 to 120 bpm, each of which is 10 bpm in length. As demonstrated in Fig. 8, the predicted HR falls within the same interval as the actual HR in the vast majority of cases. The method demonstrates good performance at both lower heart rates (≤ 60 bpm) and higher heart rates (≥ 100 bpm), which further confirms the high estimation accuracy of our method over different HR intervals.

Ablation Experiment. To further demonstrate the effectiveness of removing the black pixels, thus the ablation experiment was set up to extract the HR of the non-contact pulse from the video, which was divided into three stages: i) extracting the skin region from the image by U-Net, ii) removing the black pixels from the extracted image of the skin region and reconstructing the image, and iii) extracting the eigenvalues of the green channel from the reconstructed image for further signal processing. The experimental setup is as follows:

Experiment 1: ① → G_mean → HR

Experiment 2: G_mean → ③ → HR

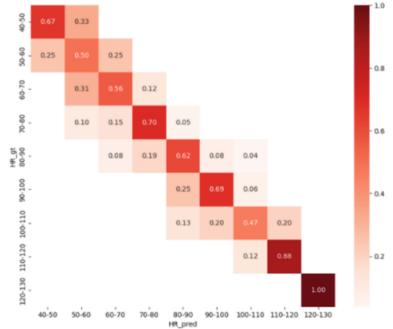


Fig. 8. Confusion matrix plot for HR estimation on the UBFC dataset. x- and y-axes denote HRpred and HRgt, respectively.

Experiment 3: ① → G_mean → ③ → HR

Experiment 4: ① → ② → G_mean → ③ → HR

Experiment 1 was coded as A, experiment 2 was coded as B, experiment 3 was coded as C, and experiment 4 was coded as D; the results of the ablation experiments are shown in Fig. 9.

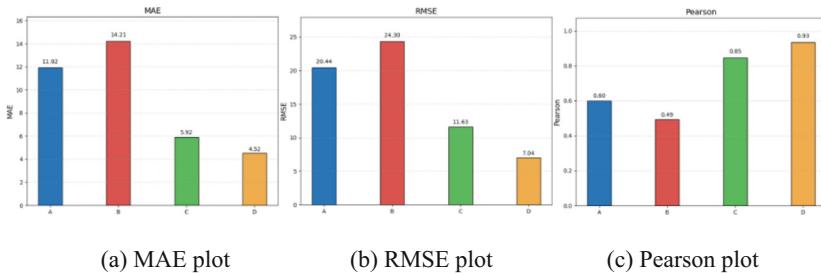


Fig. 9. Comparison chart of evaluation indicators for ablation experiments.

From the results of the ablation experiments, Experiment 4 has the smallest values of MAE and RMSE, indicating that the estimation is good, i.e., the closer the estimated HR is to the true value. The highest value of experiment 4 indicates the stronger their linear correlation. Thus, removing black pixels and performing image reconstruction can improve the effectiveness of the method in this paper.

4 Conclusion

In this study, we propose a three-stage approach aimed at achieving accurate non-contact HR estimation from facial videos. In the first stage, the image of the facial skin region of the face is extracted using the U-Net image segmentation network, as a means of generating an image that both enhances and recovers the rPPG signal and ensures that the

physiological signal information within the facial skin region is preserved. In the second stage, the noise in the non-contact photoplethysmography signal is further reduced by removing pixels in the non-skin region. In the third stage, the signal is preprocessed by applying methods such as median filtering and baseline drift removal, which in turn estimates the HR, a process that performs well in terms of saving computational resources and improving computational efficiency compared to the end-to-end deep network model. To further validate the effectiveness of the black pixel removal strategy, we conducted corresponding ablation experiments, and the experimental results confirmed that the removal of black pixels effectively improved the accuracy and robustness of the experiments. Thus, we provide a new scheme for non-contact HR estimation methods.

References

1. Metra, M., Teerlink, R.J.: Heart failure. *Lancet*. **390**(10106), 1981–1995 (2017)
2. Rachow, T., Berger, S., Israel, A.K., et al.: Relationship between electrodermal activity and heart rate in patients with acute schizophrenia. *Schizophr. Res.* **117**(2–3), 365 (2010)
3. Niu, X., Shan, S., Han, H., et al.: Rhythmnnet: end-to-end heart rate estimation from face via spatial-temporal representation. *IEEE Trans. Image Process.* **29**, 2409–2423 (2019)
4. Hu, M., Qian, F., Guo, D., et al.: ETA-rPPGNet: effective time-domain attention network for remote heart rate measurement. *IEEE Trans. Instrum. Meas.* **70**, 1–12 (2021)
5. Yan, B.P., Lai, W.H.S., Chan, C.K.Y., et al.: High-throughput, contact-free detection of atrial fibrillation from video with deep learning. *JAMA Cardiol.* **5**(1), 105–107 (2020)
6. Magdalena Nowara, E., Marks, T.K., Mansour, H., et al.: SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near-infrared. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1272–1281 (2018)
7. Aarts, L.A.M., Jeanne, V., Cleary, J.P., et al.: Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—a pilot study. *Early Hum. Dev.* **89**(12), 943–948 (2013)
8. Song, R., Zhang, S., Li, C., et al.: Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks. *IEEE Trans. Instrum. Meas.* **69**(10), 7411–7421 (2020)
9. Chen, X., Cheng, J., Song, R., et al.: Video-based heart rate measurement: recent advances and future prospects. *IEEE Trans. Instrum. Meas.* **68**(10), 3600–3615 (2018)
10. Qiu, Y., Liu, Y., Arteaga-Falconi, J., et al.: EVM-CNN: real-time contactless heart rate estimation from facial video. *IEEE Trans. Multimed.* **21**(7), 1778–1787 (2018)
11. Chen, W., McDuff, D.: Deepphys: video-based physiological measurement using convolutional attention networks. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 349–365 (2018)
12. Liu, X., Fromm, J., Patel, S., et al.: Multi-task temporal shift attention networks for on-device contactless vitals measurement. *Adv. Neural Inf. Proces. Syst.* **33**, 19400–19411 (2020)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention—MICCAI*. In: 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, pp. 234–241. Springer, Cham (2015)
14. Wang, W., Wei, Z., Yuan, J., et al.: Non-contact heart rate estimation based on singular spectrum component reconstruction using low-rank matrix and autocorrelation. *PLoS One*. **17**(12), e0275544 (2022)
15. Lugaresi, C., Tang, J., Nash, H., et al.: Mediapipe: a framework for building perception pipelines. arXiv preprint arXiv:1906.08172 (2019)

16. Bobbia, S., Macwan, R., Benerezeth, Y., et al.: Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recogn. Lett.* **124**, 82–90 (2019)
17. Bousefsaf, F., Pruski, A., Maaoui, C.: 3D convolutional neural networks for remote pulse rate measurement and mapping from facial video. *Appl. Sci.* **9**(20), 4364 (2019)
18. Niu, X., Yu, Z., Han, H., et al.: Video-based remote physiological measurement via cross-verified feature disentangling. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part II 16, pp. 295–310. Springer, Cham (2020)
19. Gideon, J., Stent, S.: The way to my heart is through contrastive learning: remote photoplethysmography from unlabelled video. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3995–4004 (2021)
20. Verkruyse, W., Svaasand, L.O., Nelson, J.S.: Remote plethysmographic imaging using ambient light. *Opt. Express.* **16**(26), 21434–21445 (2008)
21. Poh, M.Z., McDuff, D.J., Picard, R.W.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express.* **18**(10), 10762–10774 (2010)
22. De Haan, G., Jeanne, V.: Robust pulse rate from chrominance-based rPPG. *IEEE Trans. Biomed. Eng.* **60**(10), 2878–2886 (2013)
23. Wang, W., Den Brinker, A.C., Stuijk, S., et al.: Algorithmic principles of remote PPG. *IEEE Trans. Biomed. Eng.* **64**(7), 1479–1491 (2016)
24. Niu, X., Han, H., Shan, S., et al.: Synrhythm: learning a deep heart rate estimator from general to specific. In: 2018 24th International Conference on Pattern Recognition (ICPR), pp. 3580–3585. IEEE (2018)



A Robust SLAM Backend Framework with Multi-landmark Constraints in Industrial Logistics Scenarios

Yu Li¹, Bo Li^{1(✉)}, Yuanpeng Li¹, Jianyuan Ruan², Jianyun Hong³,
and Aihua Yu¹

¹ Zhejiang University of Science and Technology, Hangzhou 310023, China
120052@zust.edu.cn

² The Hong Kong Polytechnic University, Hong Kong, Hong Kong
jianyuan.ruan@connect.polyu.hk

³ Ningbo Sunny Bar Automation Co., Ltd., Yuyao, China

Abstract. The localization drift caused by the degeneracy problem and dynamic environments is a major challenge in LiDAR SLAM. This paper proposes a robust 2D LiDAR SLAM framework for industrial logistic scenarios, combining multiple landmark constraints to address these two issues. Three types of landmarks are integrated into the existing method, namely, retro-reflective column, planar reflector, and Data Matrix (DM) code. Customized methods are devised to extract the information of the first two types of landmarks through the intensity value of LiDAR measurements, while the DM codes are read by the on-board camera. The special constraint equations corresponding to these landmarks are proposed and incorporated into the existing backend framework. The experimental results show that both mapping and localization accuracy are significantly improved in degeneration scenarios and dynamic environments.

Keywords: LiDAR · SLAM · backend · Landmark

1 Introduction

Precise mapping and accurate localization are fundamental for autonomous navigation systems. SLAM [1] technology, which enables simultaneous environment mapping and real-time localization, provides reliable navigation support and has been widely adopted in industrial automation, warehouse logistics, and robotics. Representative systems base on 2D LiDAR like Cartographer [2] employ front-end scan registration for local map construction and backend graph optimization with loop closure detection to mitigate cumulative errors, thereby improving mapping quality and localization accuracy. However, practical deployments in factory environments face challenges from repetitive structures (degenerated scenarios) and frequent movements of personnel and equipment (dynamic environments). These factors induce drift and mismatches during mapping and localization processes, severely degrading system robustness and precision

Our approach is designed to complement the conventional LiDAR SLAM with additional constraints, integrating the corresponding constraint models to effectively improve the robustness and accuracy of the SLAM and localization system in degenerated and dynamic scenarios. For the issue of the dynamic environment, a feasible solution is to incorporate the fixed and invariant landmarks into the LiDAR measurements and utilize them to correct the drift. As to the degeneration problem, the DM code is adopted to identify the precise movement of the robot in the degenerated environment where the SLAM systems are more prone to failure when only LiDAR is employed.

The main contributions of this research can be summarized in three aspects:

- Customized methods are proposed to extract the information of two types of landmarks, reflective columns and planar reflectors, through the intensity value of LiDAR measurements.
- A map model based on the fusion of LiDAR and visual information of DM codes, point-shaped landmarks, line-shaped landmarks, and occupancy grid maps is proposed.
- A backend framework is proposed by designing constraint equations corresponding to the three types of landmarks and subsequently integrating them into the existing SLAM system.

2 Relate Work

Cartographer is an efficient 2D LiDAR SLAM system combining front-end real-time correlative Scan Matching (CSM) [3] and back-end efficient Sparse Pose Adjustment (SPA) [4] to achieve real-time mapping and optimization. CSM uses multi-resolution search to register scans to submaps, offering faster, more global matching than ICP [8]. SPA formulates a sparse pose-graph [5] optimization to reduce cumulative error. Its real-time performance, accuracy, and ease of integration have made it indispensable in research and industry. Cartographer also employs loop-closure detection to correct long-term drift. However it matching current scans to historical submaps and adding constraints, still fail in degraded or dynamic scenes, causing mismatches and drift.

In autonomous navigation, localization must supply precise global poses for planning. Monte Carlo Localization (MCL) [6] uses particle filters and resampling but converges slowly in feature-sparse or dynamic settings and suffers from spurious data associations. Cartographer's localization mirrors its mapping: the front end preserves local map accuracy, and the back end uses loop closures to correct drift [3,5]. However, in degraded and dynamic environments, loop closure constraints may fail to be established or may be erroneous, resulting in significant localization errors and drift.

Overall, existing localization methods perform suboptimal in complex environments equently [10], many studies have begun to explore the incorporation of additional information to enhance localization accuracy and system robustness. In recent years, the SLAM field has made significant progress in the study of

landmarks [11]. Some studies have attempted to use naturally occurring geometric structures (such as corners and door frames) as landmarks, but these features are often unstable in degraded and dynamic environments, which limits their performance. Therefore, integrating artificially designed landmarks such as high-reflectance retro-reflective columns and planar reflectors, which offer the advantage of unique encoding, has emerged as an effective complementary approach. In this paper, multiple landmark constraints are introduced into Cartographer, including retro-reflective columns, planar reflectors, and DM codes. Retro-reflective columns and planar reflectors possess strong geometric and semantic information and typically appear as high-reflectance point clusters in LIDAR data, making them easy to detect; whereas DM codes, with their unique identification and high robustness against interference [12], can provide precise pose information. The additional constraints provided by these landmarks effectively compensate for the deficiencies of LIDAR data in degraded and dynamic environments, further enhancing the robustness of Cartographer.

3 Methodology

3.1 Basic Backend Optimization Framework

The proposed backend framework is based on the backend module of Cartographer. To provide a more comprehensive introduction to the new methods proposed in this study, a concise overview of the Cartographer backend is given in this subsection.

Local scan matching registers the scan with the current submap and employs non-linear least squares to obtain the relative pose measurements between adjacent scans. Let the robot's pose at frame i be ε_i , and let the relative measurement between frames j and ε_j obtained from scan matching be ε_{ij} with covariance Σ_{ij} . The residual for the local matching constraint can be expressed as:

$$E_{local} = \sum_{(i,j) \in C_{local}} \|\varepsilon_{ij} - f(\varepsilon_i, \varepsilon_j)\|_{\Sigma_{ij}}^2 \quad (1)$$

where $f(\varepsilon_i, \varepsilon_j)$ denotes the predicted relative pose derived from i and j .

When a loop is detected, the system adds additional loop closure constraints to the pose graph to correct for cumulative errors. The loop closure measurement is denoted as ε_{ij}^{loop} with covariance Σ_{ij}^{loop} . the residual can be expressed as:

$$E_{loop} = \sum_{(i,j) \in C_{loop}} \|\varepsilon_{ij}^{loop} - f(\varepsilon_i, \varepsilon_j)\|_{\Sigma_{ij}^{loop}}^2. \quad (2)$$

Together, the local matching and loop closure constraints constitute the optimization objective of the Cartographer backend:

$$\min_{\varepsilon} \{E_{pose} = E_{local} + E_{loop}\} \quad (3)$$

where $\varepsilon = \{\varepsilon_1, \varepsilon_2, \varepsilon_3 \dots \varepsilon_n\}$ represents all robot poses, and the problem is solved efficiently using Sparse Pose Adjustment (SPA).

An additional landmark constraint is introduced in the Cartographer back end, with an idea similar to that of loop closure constraints: when the robot re-detects a stable and distinctive landmark in the environment, its observation is added as an extra constraint in the optimization problem.

$$e_i^L = Z_{i_1 i_2}^L - h(i_1, i_2) \quad (4)$$

$$E_{landmark} = \sum_{i \in C_{landmark}} \|e_i^L\|^2 \quad (5)$$

Here, $L = \{l_1, l_2, l_3 \dots l_m\}$ denotes all landmark constraints. After introducing a weighting factor λ , the overall optimization problem is updated as:

$$\min_{\varepsilon, L} \{E_{pose} = E_{local} + E_{loop} + \lambda_{landmark} E_{landmark}\}. \quad (6)$$

3.2 Extraction of Three Types of Landmarks and the Enhanced Backend Framework

Retro-Reflective Column. Retro-reflective columns appear in LiDAR scans as clusters of points with reflectance values significantly higher than that of the surrounding environment, and are distributed in the 2D plane in a circular or partial arc pattern as shown in Fig. 1(a). Exploiting this characteristic, points with reflectance above a threshold are filtered from the scan data, and then clustered to form candidate regions corresponding to individual retro-reflective columns. Subsequently, a least squares circle fitting method is employed to extract the circle center, which is used as the pose of the landmark. Let each scan point $p_i(x_i, y_i)$ be with reflectance R_i , and set a threshold τ_r to filter out the set of high-reflectance points:

$$P = \{p_i | p_i > \tau_r, d(i, j) < \tau_d\} \quad (7)$$

where $d_{(i,j)}$ is the Euclidean distance between p_i and p_j and τ_d denotes maximum intracluster distance threshold.

Subsequently, a least squares circle fitting method is employed to extract the circle center, which is used as the pose of the landmark. Assume the candidate cluster contains high-reflectance points $\{x_i, y_i\}_{i=1}^N$, and the standard equation of a circle is:

$$(x - x_c)^2 + (y - y_c)^2 = r^2 \quad (8)$$

where (x_c, y_c) is the center of the circle and r is its radius. Let $C = x_c^2 + y_c^2 - r^2$. For each point (x_i, y_i) , Expanding and rearranging the equation into a linear form yield:

$$2x_i x_c + 2y_i y_c + C = x_i^2 + y_i^2. \quad (9)$$

Then the objective function can be expressed as:

$$\min_{(x_c, y_c, C)} = \sum_{i=1}^N \|2x_i x_c + 2y_i y_c + C - (x_i^2 + y_i^2)\|^2 \quad (10)$$

$$A = \begin{bmatrix} 2x_1 & 2y_1 & 1 \\ 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots \\ 2x_N & 2y_N & 1 \end{bmatrix} \quad d = \begin{bmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ \vdots \\ x_N^2 + y_N^2 \end{bmatrix} \quad \beta = \begin{bmatrix} x_c \\ y_c \\ C \end{bmatrix} \quad (11)$$

$$\beta = (A^T A)^{-1} A^T d. \quad (12)$$

The pose of the retro-reflective column landmark is represented by its center (x_c, y_c) . the corresponding residual is:

$$E_{\text{column}} = \sum_{i \in l_{\text{column}}} \|Z_{i_1 i_2}^{\text{column}} - h(i_1, i_2)\|^2. \quad (13)$$

The backend optimization equation after adding the retro-reflective column landmark constraint becomes:

$$\min_{\varepsilon, l} \{E_{\text{pose}} = E_{\text{local}} + E_{\text{loop}} + \lambda_{\text{column}} E_{\text{column}}\}. \quad (14)$$

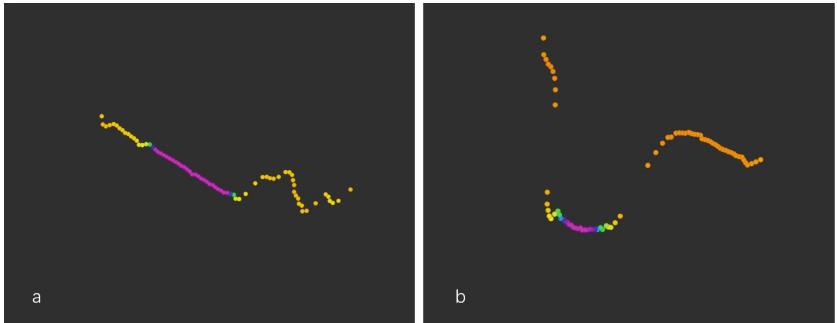


Fig. 1. a: The planar reflector appears as a line in the laser scan. b: The retro-reflective column appears as an arc in the laser scan.

Planar Reflector Landmarks. Planar reflectors also appear as high-reflectivity regions, but their point clusters in LiDAR data present a linear distribution. See Fig. 1a. Similarly, high-reflectivity points are first selected, and clustering is then used to separate candidate regions for planar reflectors.

The set of points in a candidate cluster for a planar reflector be denoted by P , and the general expression of a line is: $Ax + By = d$. The normal vector of the line is $N = (A, B)$, and d represents the intercept of the line.

The objective function is:

$$\min_{(N, d)} = \sum_{i=1}^N \|N \bullet P_i - d\|^2 \quad (15)$$

let

$$X = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & 1 \end{bmatrix} \quad I = \begin{bmatrix} N^T \\ -d \end{bmatrix}. \quad (16)$$

The objective function becomes: $\min(N, d) = \|XI\|^2$. Direct solving may yield the trivial solution, so an eigenvalue decomposition method is often used to determine the line parameters. The covariance matrix is $\Sigma = X^T X$. By performing eigenvalue decomposition on Σ , the eigenvector corresponding to the smallest eigenvalue represents the line parameters $[N \ -d]^T$.

Let segment center and segment normal direction be expressed as:

$$\begin{cases} x_c = \frac{1}{N} \sum_{i=1}^N x_i \\ y_c = \frac{1}{N} \sum_{i=1}^N y_i \\ \theta = \tan^{-1}(-A/B) \end{cases} \quad (17)$$

The pose of the planar reflector is $(x_c \ y_c \ \theta)$, and the corresponding residual is:

$$E_{board} = \sum_{i \in l_{board}} \|Z_{i_1 i_2}^{board} - h(i_1, i_2)\|^2 \quad (18)$$

The backend optimization equation after incorporating the planar reflector landmark constraint becomes:

$$\min_{\varepsilon, l} \{E_{pose} = E_{local} + E_{loop} + \lambda_{board} E_{board}\}. \quad (19)$$

DM Code. The recognition of DM codes is a relatively mature technology in the field of computer vision, and open-source libraries such as OpenCV [13] and libdmtx [14] can be used to obtain the DM code ID and its pose relative to the camera. Assume that the detected DM code center's pixel coordinate is (u_c, v_c) , the top-left point (u_1, v_1) and the top-right point (u_2, v_2) , the depth value is d , and the camera intrinsic matrix K is:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (20)$$

In the camera coordinate system, the three-dimensional coordinates of the DM code center are:

$$\begin{cases} x_c = \frac{u_c - c_x}{d} \\ y_c = \frac{v_c - c_y}{d} \\ z_c = d \end{cases} \quad (21)$$

Horizontal angle: $\theta = \tan^{-1} \left(\frac{v_2 - v_1}{u_2 - u_1} \right)$. The pose of the DM code $(x_c \ y_c \ \theta)$, and the corresponding residual is:

$$E_{DM} = \sum_{i \in l_{DM}} \|Z_{i_1 i_2}^{DM} - h(i_1, i_2)\|^2 \quad (22)$$

The backend optimization equation after adding the DM code landmark constraint becomes:

$$\min_{\varepsilon, l} \{E_{pose} = E_{local} + E_{loop} + \lambda_{board} E_{board}\}. \quad (23)$$

By integrating the above three types of landmark constraints into the graph optimization back-end of Cartographer, the final optimized objective function is:

$$\min_{\varepsilon, L} \{E_{pose} = E_{local} + E_{loop} + \lambda_{pole} E_{pole} + \lambda_{board} E_{board} + \lambda_{DM} E_{DM}\}. \quad (24)$$

4 Experimental Results

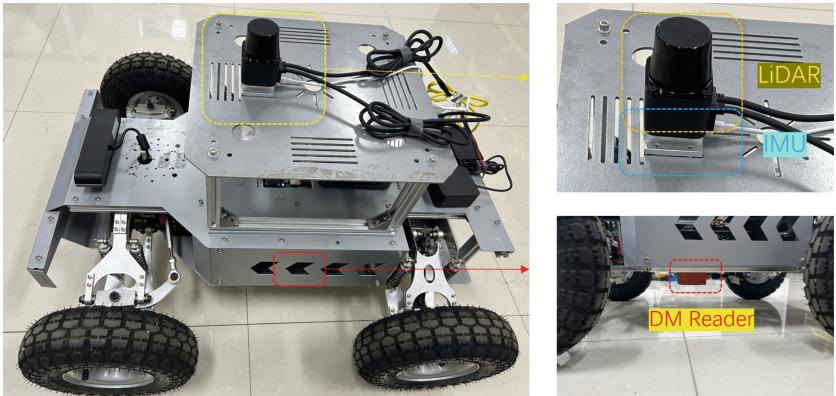


Fig. 2. Experimental platform.

Two experiments are deployed to evaluate our method's robustness and accuracy in degenerated and dynamic scenarios. The first experiment takes place in four long straight indoor corridors (degenerated environment); the second is performed in a busy workshop with a lot of moving objects. An Ackermann-chassis vehicle (see Fig. 2), equipped with 2D LiDAR, IMU, and code reader, serves as the SLAM platform.

We pre-deployed three types of landmarks (see Fig. 3)—retro-reflective column, planar reflector, and DM code—at the experiment site and set up experiments to compare the performance of the traditional Cartographer method with that of the system enhanced by landmark constraints in complex environments.



Fig. 3. Three types of landmarks.

Experiment-Degenerated Environments. This experiment was conducted in a long corridor environment. As shown in Fig. 4, the distinctive geometric features in the environment are sparse, and repetitive structures are prominent, which is a typical degenerated scenario. The experiment compared the mapping results with and without landmarks.

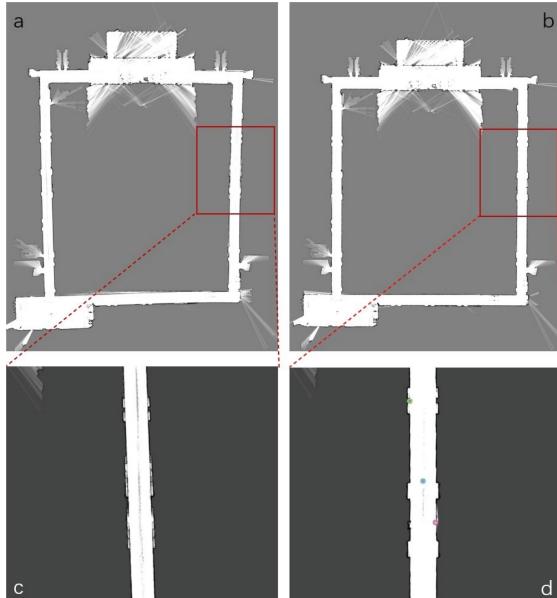


Fig. 4. Mapping results in the degenerated-environment Experiment. (a) Mapping results of the original cartographer. (b) Mapping results of our method. (c) Detailed mapping results of cartographer. (d) Detailed mapping results of our method.

Figure 4(a) shows the mapping result without landmarks, while Fig. 4(b) shows the result with landmarks added. Figure 4(c) and (d) present local enlarged comparisons of the degenerated scene. It can be clearly seen that Cartographer performs better in the degenerated environment after adding land-

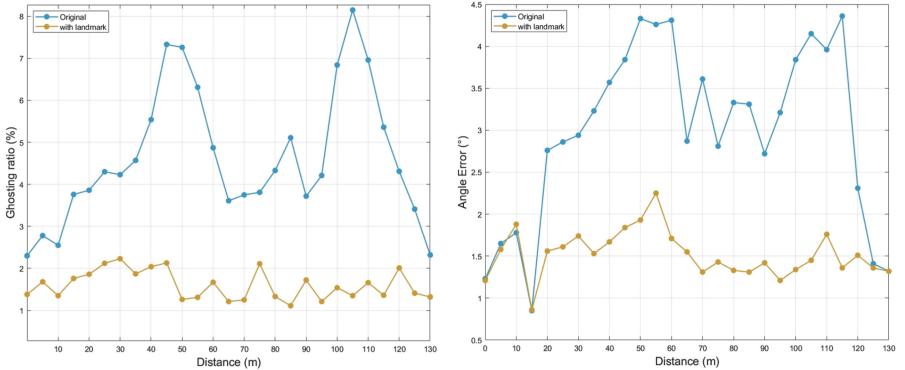


Fig. 5. The ghosting ratio and angular error of the SLAM system with and without landmark constraints intervals are 5-meter.

marks (Fig. 4(d)). Since it is difficult to obtain ground-truth trajectories in indoor environments, we take advantage of the parallel structure of the corridor: it is horizontally straight and, according to actual measurements, the layout of the test area is a standard rectangle. The angle difference between the current position and the standard rectangle is used as an approximation of SLAM angular error. The mapping quality is quantified by the ghosting ratio, calculated as the ratio of the ghosting area to the local map area. A lower ghosting ratio indicates superior mapping performance. The total length of the experiment is approximately 130 m. The angular error and ghosting ratio were calculated every 5 m.

The results are shown in Fig. 5. For the traditional Cartographer, the average angular error was 2.99° , and the average shadow ghosting ratio was 4.62%. After integrating landmark constraints, the average angular error dropped to 1.52° , and the ghosting ratio decreased to 1.61%.

The experiment site for the second experiment was selected in a typical factory environment, encompassing an area of approximately 10,000 square meters. In this scenario, frequent movement of personnel forms a typical dynamic environment. The localization performance was compared in this scenario.

As shown in Fig. 6, Fig. 6(b) represents a local part of the entire factory. Figure 6(c) shows the localization results without landmarks, with obvious map distortion. Figure 6(d) presents the localization results with landmarks added, showing no obvious map distortion and faster pose correction. Because Cartographer's map is base on occupancy grid map, the quality of localization can be quantitatively assessed by comparing the overlap ratio between the submap during localization and the corresponding submap in the original map; the higher the overlap ratio, the better the localization accuracy. The total route of the experiment was approximately 580 m, and the submap overlap ratio was calcu-

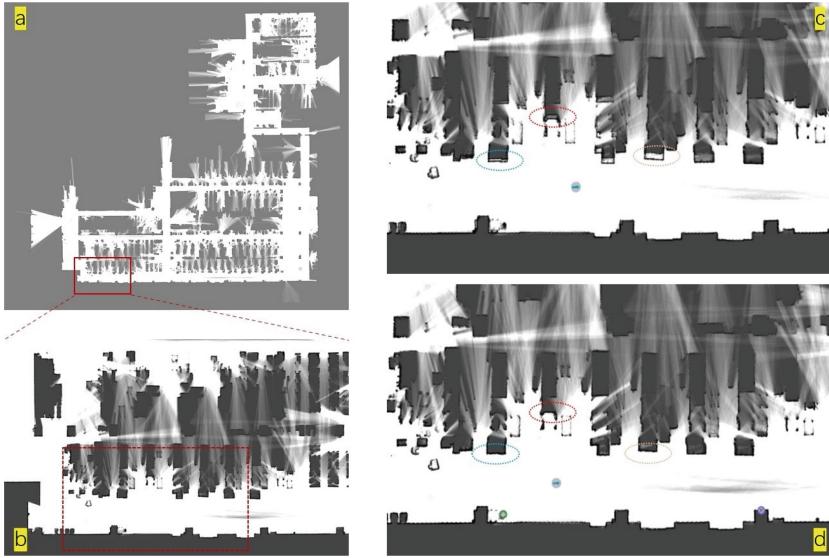


Fig. 6. Localization results in a dynamic environment. (a) Mapping results in the dynamic environment. (b) Local enlarged view of the mapping result. (c) Localization result of our method. (d) Localization result of the original cartographer.

lated every 20 m during localization; the experimental results are shown in Fig. 7. The localization accuracy increased from 85.25% to 96.36%.

The above experimental results are summarized in Table 1 .Overall, by deploying retro-reflective columns, planar reflectors, and Data Matrix codes at the experimental site and conducting controlled experiments, we have validated the effectiveness of the multi-modal landmark constraint method in enhancing the mapping and localization performance of the Cartographer system, thereby providing robust support for the application of autonomous navigation systems in degenerated and complex dynamic environments

Table 1. Experimental results

Metric	Cartographer	Cartographer with landmarks
Angular error (\circ)	2.99	1.52
Local ghosting ratio (%)	4.62	1.6
Localization accuracy (%)	85.25	96.36

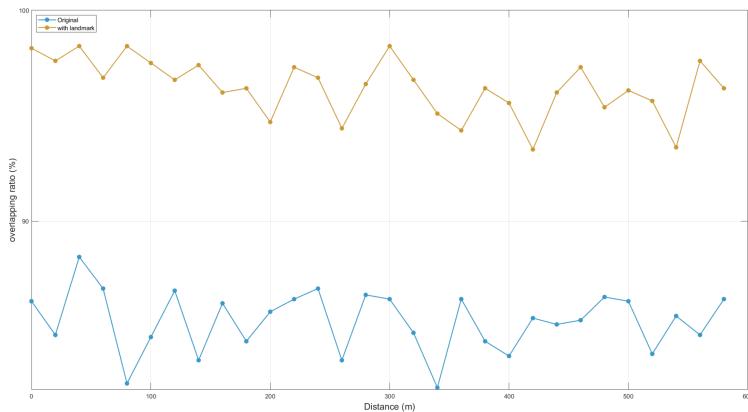


Fig. 7. A test position was selected at every 20-meter interval, and the overlap ratio between the current submap and the corresponding submap in the original map was calculated for both cases with and without landmarks.

5 Conclusion

This paper addresses the issue of insufficient mapping and localization stability of 2D LiDAR SLAM in degenerate and dynamic industrial environments by proposing an improved method based on multimodal landmark constraints. By incorporating three types of landmarks (retro-reflective column, planar reflector, and DM code) into the Cartographer system, additional pose information is integrated into the back-end graph optimization, compensating for the shortcomings of traditional 2D LiDAR SLAM in degenerate and dynamic scenarios. Experimental results demonstrate that the proposed method offers significant advantages in reducing accumulated errors, enhancing mapping accuracy, and improving localization robustness.

Acknowledgment. This work is supported by the Fundamental Research Funds Project of Zhejiang University of Science & Technology (Grant No. 2025QN011), Graduate Research Innovation Fund of Zhejiang University of Science & Technology (Grant No. F464101Q01).

References

1. Simultaneous Localization and Mapping (SLAM): Part I & II” IEEE Robotics & Automation Magazine (2006)
2. Hess, W., Kohler, D., Rapp, H., Andor, D.: Real-time loop closure in 2D LIDAR SLAM. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 1271–1278. Stockholm, Sweden (2016)
3. Olson, E.B.: Real-time correlative scan matching. In: IEEE International Conference on Robotics and Automation, pp. 4387–4393. Kobe, Japan (2009). <https://doi.org/10.1109/ROBOT.2009.5152375>

4. Konolige, K., Grisetti, G., Kümmerle, R., Burgard, W., Limketkai, B., Vincent, R.: Efficient sparse pose adjustment for 2D mapping. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 22–29. Taipei, Taiwan (2010). <https://doi.org/10.1109/IROS.2010.5649043>
5. Grisetti, G., Kümmerle, R., Stachniss, C., Burgard, W.: A tutorial on graph-based SLAM. *IEEE Intell. Transp. Syst. Mag.* **2**(4), 31–43 (2010). <https://doi.org/10.1109/MITS.2010.939925>
6. Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics. MIT Press (2005)
7. Mur-Artal, R., Tardós, J.D.: ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. In: IEEE Transactions on Robotics (2017)
8. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 239–256 (1992). <https://doi.org/10.1109/34.121791>
9. Saputra, M.R.U., et al.: Dynamic Object-Aware LiDAR SLAM for Autonomous Navigation. *IEEE ICRA* (2020)
10. Biswas, J., et al.: Towards Persistent Coarse-Grained Localization in Changing Environments. *IEEE ICRA* (2018)
11. Weingarten, J., et al.: 3D SLAM Using Planar Segments for Indoor/Outdoor Mobile Robots. *IEEE/RSJ IROS* (2006)
12. Olson, E., et al.: AprilTag: A Robust and Fiducial Marker System Using the Hamming Code. *IEEE/RSJ IROS* (2011)
13. Bradski, G.: The OpenCV library. *Dr. Dobb's J. Softw. Tools* **25**(11), 120–123 (2000)
14. Rahman, M.B.M.A., et al.: Robust data matrix code recognition using libdmtx in industrial environments. *IEEE Access* (2020). <https://doi.org/10.1109/ACCESS.2020.2992355>



Multi-scale Hierarchical Feature Attention Interaction for Fault Diagnosis of Rolling Bearings Under Noisy Conditions

Hanwen Chang, Funa Zhou^(✉), and Jiechen Sun

Shanghai Maritime University, Shanghai 201306, China
zhoufn2002@163.com

Abstract. Fault diagnosis of rolling bearings under strong noise interference faces challenges such as the masking of fault size features and difficulty in discrimination. Traditional diagnostic methods often neglect effective interactions across different scales and lack cross-layer feature fusion strategies, resulting in limited diagnostic performance under noise interference. To address these challenges, this paper proposes a Multi-Scale Hierarchical Feature Attention Interaction (MSH-FIA) method to enhance the model's ability to identify multi-scale fault features and improve noise robustness. The method first employs a multi-granularity feature collaboration fusion module to efficiently extract and integrate features of different granularities. A dual-path residual weighting strategy is introduced to strengthen the model's focus on critical channels and spatial regions. Lastly, a cross-scale attention interaction mechanism is used to enable cross-layer attention guidance between shallow and deep feature maps, enhancing the saliency of key features. Experimental results under various noise levels demonstrate that the proposed model maintains robust perception and discrimination capabilities for fault features of different sizes, even under strong noise interference.

Keywords: fault diagnosis · Convolutional neural network · Multi-scale CNN · Attention mechanisms · hierarchical feature learning

1 Introduction

Rolling bearings are vital components in rotating machinery, and their health directly affects operational safety and efficiency. In practical monitoring, vibration signals are often contaminated by multiple sources of noise. These noise factors can significantly obscure critical fault features. Existing fault diagnosis models may learn spurious features associated with the noise, leading to misclassifications and severely impacting the accuracy of fault identification [1].

The introduction of deep learning has significantly improved the accuracy and robustness of fault diagnosis using vibration signals, with Convolutional Neural Networks (CNNs) standing out for their powerful feature extraction capabilities. Diagnosing faults of the same type but different sizes under noise interference poses a significant challenge. Specifically, In the early stages of faults, the vibration signals are weak and highly

overlapped with noise, making them difficult to detect [2]. As fault size increases, the features grow nonlinearly, and noise further complicates this relationship, making it difficult for the model to establish an accurate correspondence between fault size and signal features [3]. Traditional fault diagnosis methods primarily focus on fault type classification and lack feature modeling mechanisms that account for size differences, resulting in insufficient recognition accuracy for faults of the same type but different sizes under strong noise conditions.

To address the above issues, researchers have introduced multi-scale learning mechanisms to enhance the model's ability to perceive faults of different sizes. Existing multi-scale feature extraction networks typically use multi-branch parallel structures with convolution kernels of different sizes to capture multi-scale features. However, this architecture lacks interaction mechanisms between different scales, features of smaller faults are filtered out as noise in coarser scale branches, leading to the loss of critical diagnostic information [4, 5]. Additionally, some methods introduce attention mechanisms to enhance feature selectivity and suppress noise. But they lack cross-scale interactions between different layers. Attention is usually applied independently at each network layer, without an effective cross-layer feature fusion mechanism, resulting in inconsistent noise suppression across scales. Consequently, fault features at important scales may be weakened during cross-layer propagation [6, 7].

This paper addresses the limitations of existing methods in quantitative diagnosis of fault sizes in rolling bearings under noisy environments. It focuses on solving the problems of insufficient multi-scale feature interaction and inconsistency in cross-level noise suppression strategies. A novel Multi-Scale Hierarchical Feature Attention Interaction (MSHFIA) method is proposed, which significantly enhances both the accuracy and robustness of fault size diagnosis. The main contributions of this work are as follows:

- (1) A cross-scale attention interaction method is proposed to establish cross-layer feature propagation pathways. It builds inter-layer feature correlations via cross-scale attention mechanisms, leveraging multi-scale fault feature dependencies to separate noise while ensuring consistent noise suppression across network depths.
- (2) A multi-receptive-field guided attention mechanism is introduced, which models channel features through multi-scale convolutional paths and dynamically adjusts channel weights based on multi-scale responses, achieving selective feature transmission during cross-layer fusion to strengthen critical feature expression.
- (3) A multi-granularity feature coordination strategy is designed. Progressive dilated convolutional groups are used to construct adaptive receptive fields, achieving efficient fusion of local fine-grained and global feature information.

2 Related Work

2.1 Noise Fault Diagnosis Based on Attention Mechanism

Vibration signals are often contaminated by noise, obscuring fault features and making it difficult for models to focus on critical regions. Attention mechanisms, due to their ability to select and enhance key features, improving accuracy and robustness. Wang et al. [8] proposed a CNN model with an attention mechanism to suppress noise by weighting feature maps. However, this method focuses only on the channel dimension,

neglecting spatial noise distribution, leading to poor performance under complex noise conditions. Wang et al. [6] later introduced a dual attention mechanism that combines both channel and spatial attention, cascading these modules to enhance the model's responsiveness to critical features. However, this sequential connection increases model complexity and still fails to effectively model channel and spatial information together. To address these limitations, Ma et al. [9] proposed a parallel residual attention mechanism, which computes channel and spatial attention in parallel and incorporates residual connections to couple feature enhancement with attention mask learning. This approach efficiently integrates both types of attention while reducing module complexity. However, it still lacks cross-scale and cross-level interactions and consistency constraints across different layers. For quantitative fault size diagnosis under noisy conditions, critical size-discriminative information tends to weaken or disappear during cross-layer propagation. Existing methods lack effective cross-scale attention mechanisms, severely limiting the model's ability to accurately identify fault sizes.

2.2 Cross-Scale Feature Fusion Strategies for Fault Diagnosis

In vibration signal-based fault identification, shallow convolutional kernels capture local time-frequency features, deep features capture global semantics through nonlinear transformations. To fully leverage multi-scale fault features and enhance model robustness in noisy environments, establishing effective interactions between shallow and deep features is a key challenge. Various cross-scale feature fusion strategies have been proposed. Huang et al. [10] introduced a Feature Pyramid Network (FPN), which uses top-down information pathways to fuse multi-level features by upsampling deep semantic features and concatenating them with shallow structural features. However, the unidirectional fusion strategy results in significant information loss during feature propagation. To address this, Chen et al. [11] proposed a Bidirectional Feature Pyramid Network (BiFPN) that adds bottom-up path augmentation modules to FPN, constructing multi-path information flows for iterative propagation and integration between upper and lower layer features, effectively improving fused feature representation. While this multi-path structure enhances feature interaction, its static fusion connections fail to dynamically adjust based on feature importance. To better focus on critical features, Guo et al. [12] developed a Spatial Pyramid Attention Network (SPANet), incorporating attention mechanisms into cross-scale fusion to highlight fault regions while suppressing noise. While existing cross-scale fusion methods improve shallow-deep feature interactions, they neglect cross-scale correlations and complementary information.

3 Methodology

This section introduces the Multi-Scale Hierarchical Feature Interaction Attention Convolutional Neural Network (MSHFIA-CNN). The network uses a progressive multi-granularity feature coordination strategy to dynamically fuse local and global features, addressing the limitations of traditional methods that process each scale independently. Additionally, MSHFIA-CNN employs a cross-scale interactive attention mechanism, establishing bidirectional attention pathways across layers. This mechanism enhances

feature representation by leveraging cross-scale correlations, suppresses noise, and prevents feature degradation during cross-layer propagation in noisy environments.

3.1 Cross-Scale Attention Interaction for Noise Suppression

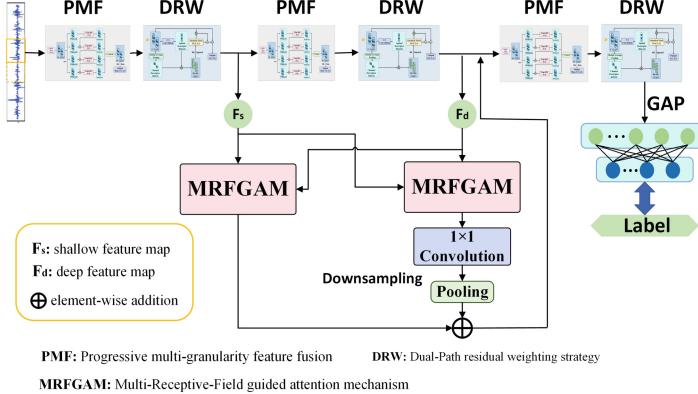


Fig. 1. The structure of MSHFIA-CNN.

Existing attention mechanisms emphasize multi-granularity features of different-sized faults in channel and spatial dimensions, but typically operate within individual network layers, overlooking the complementary nature of features across depths.

This paper proposes a cross-scale attention interaction mechanism that establishes bidirectional guidance between shallow and deep feature maps, facilitating semantic complementarity and information synergy. To improve scale adaptability in channel weight computation and enhance channel representation during cross-layer attention interaction, we design a Multi-Receptive Field Guided Attention Mechanism (MRFGAM) as the channel weight generation approach. As shown in Fig. 1, the MSHFIA-CNN architecture consists of Progressive Multi-granularity Fusion (PMF) layers and Dual-path Residual Weighting (DRW) layers. The shallow output feature map is denoted as $F_s \in \mathbb{R}^{C_s \times L_s}$, while the deep feature map is $F_d \in \mathbb{R}^{C_d \times L_d}$. The shallow feature map F_s is fed into MRFGAM to obtain cross-scale channel attention weights $w_{s-d} \in \mathbb{R}^{C_s \times 1}$, whose dimension matches the channel number of deep feature map F_d . These attention weights are then applied to F_d , enabling shallow-level information to guide deep feature enhancement:

$$\tilde{F}_d = w_{s-d} \otimes F_d \quad (1)$$

Similarly, the deep feature map F_d is processed through MRFGAM to generate cross-scale channel attention weights $w_{d-s} \in \mathbb{R}^{C_d \times 1}$, which are applied to guide the weighting of shallow feature maps.

The weighted shallow feature maps are downsampled and fused with the weighted deep feature maps to construct mutually enhanced cross-level representations. This

cross-scale attention interaction mechanism enables the model to integrate complementary feature representations across different receptive field hierarchies, thereby strengthening information exchange between shallow and deep features.

3.2 Multi-receptive-Field Guided Attention Mechanism

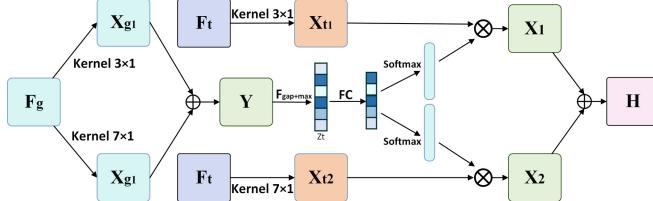


Fig. 2. Structure of the Multi-Receptive Field Guided Attention Mechanism

During cross-layer feature interaction, shallow layers capture local details, while deep layers focus on global semantics. Failure to account for receptive field size during channel weight generation can lead to feature selection bias, undermining the accuracy and stability of critical feature guidance. To address this issue, we propose the MRFGAM.

The structure of MRFGAM is shown in Fig. 2, takes two hierarchical feature maps as inputs: a guidance feature map $F_g \in \mathbb{R}^{C_g \times L_g}$ and a target feature map $F_t \in \mathbb{R}^{C_t \times L_t}$. The guidance feature map F_g is first processed through two parallel convolutional branches with different kernel sizes to extract channel response features under varying receptive fields. The outputs X_{g1} and X_{g2} from the two branches are fused through element-wise summation to form intermediate features $Y = X_{g1} \oplus X_{g2}$. Global average pooling and max pooling are applied along the channel dimension to generate the channel descriptor vector $z_t \in \mathbb{R}^{C_t \times 1}$. The vector z_t is then processed to produce multi-scale attention weights $\omega \in \mathbb{R}^{C_t \times 1}$:

$$w = \text{Softmax}(\text{FC} \cdot \text{ReLU}(\text{FC} \cdot z_t)) \quad (3)$$

The final attention weights ω are applied to the outputs X_{t1} and X_{t2} from both scale-specific branches, enabling channel-wise weighted modulation of the cross-scale feature maps under different receptive fields:

$$\tilde{X}_1 = w \otimes X_{t1}, \tilde{X}_2 = w \otimes X_{t2} \quad (4)$$

The channel-enhanced feature representations are then fused through element-wise summation to produce the final output feature map $H = \tilde{X}_1 \oplus \tilde{X}_2$.

The MRFGAM facilitates saliency guidance between hierarchical feature maps at different levels, enabling more selective transmission of features during cross-level fusion.

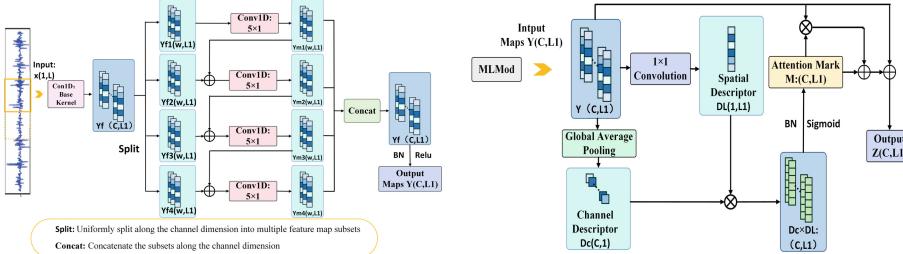


Fig. 3. Progressive Multi-granularity Feature Fusion Strategy.

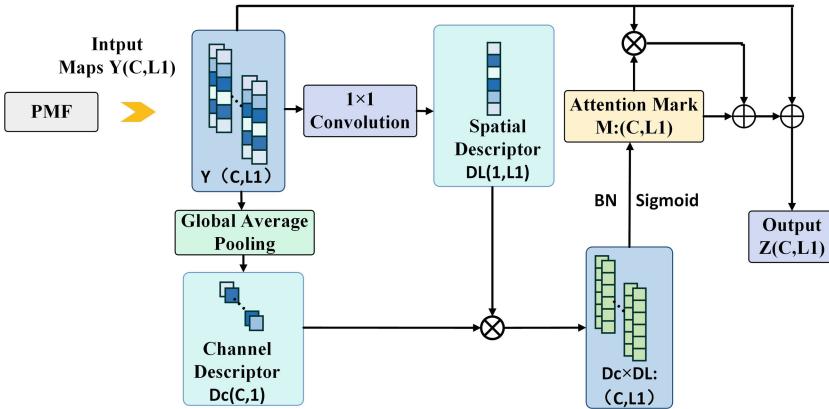


Fig. 4. The structure of Dual-Path Residual Weighting Strategy.

3.3 Multi-granularity Feature Fusion and Dual-Path Residual Weighting Strategy Under Noise Interference

Given that noise interference can easily obscure the characteristic distribution differences corresponding to various fault sizes, the capability to capture their intrinsic distinctions becomes crucial. This necessitates not only extracting multi-scale fault features but also enhancing interactions between features of different granularities, thereby achieving a compact yet discriminative feature fusion process.

The progressive multi-granularity features fusion architecture designed in the network backbone is illustrated in Fig. 3. The specific implementation steps are as follows: First, the input 1D vibration signal $x \in \mathbb{R}^{L \times 1}$ is processed through a convolutional layer to obtain the feature map $Y_f = [y_1^f, y_2^f, \dots, y_C^f]^T \in \mathbb{R}^{C \times L_1}$ with C channels. The feature map Y_f is then uniformly divided into s subsets, denoted as Y_f^i where $i \in \{1, 2, \dots, s\}$. Each subset $Y_f^i \in \mathbb{R}^{w \times L}$ represents a feature scale, with $w = \frac{C}{s}$ being the number of channels per subset. The first subset Y_f^1 is processed directly by its corresponding 1D convolutional layer $K_1(\bullet)$, while each subsequent subset Y_f^i is first summed with the output of the previous convolutional layer $K_{i-1}(\bullet)$ before being fed into the current

convolutional layer $K_i(\bullet)$. The output Y_m^i of this operation is expressed as:

$$Y_m^i = \begin{cases} K_i(Y_f^i), & i = 1 \\ K_i(Y_f^i + Y_m^i), & 2 \leq i \leq s \end{cases} \quad (5)$$

Finally, the subsets are concatenated along the channel dimension, resulting in a new feature map $Y_m = [Y_m^1, Y_m^2, \dots, Y_m^s]^T \in \mathbb{R}^{C \times L_1}$, which is then processed by batch normalization (BN) and a ReLU activation layer.

Due to noise interference, the output feature map Y after multi-granularity fusion typically contains redundant information. To address this issue, we designed a dual-path residual weighting strategy for multi-granularity features, which integrates both channel attention and spatial attention mechanisms. The detailed architecture is shown in Fig. 4. The multi-scale feature map $Y = [y_1, y_2, \dots, y_c]^T \in \mathbb{R}^{C \times L_1}$. Along the channel dimension, Global Average Pooling (GAP) is applied to Y to compress each channel's features, yielding the channel descriptor $D_C = [d_C^1, d_C^2, \dots, d_C^C]^T \in \mathbb{R}^{C \times 1}$. On the spatial dimension, a 1×1 convolutional layer is applied to transform Y into a single-channel feature map, generating the spatial descriptor $D_L = [d_L^1, d_L^2, \dots, d_L^{L_1}]^T \in \mathbb{R}^{1 \times L_1}$.

Then, matrix multiplication is used to combine the channel descriptor and spatial descriptor, thereby merging channel attention and spatial attention. After that, the attention mask $M \in \mathbb{R}^{C \times L_1}$ is obtained through the BN layer and the sigmoid activation function. The attention mask M is a weighted version of the multi-scale feature map Y , and because its shape is the same as that of Y and contains the original attention information, the output can directly learn the attention mask through residual connections, further enriching the feature information. Finally, the output of RAMod is:

$$Z = M \otimes Y \oplus M \oplus Y \quad (6)$$

4 Experimental Validation

4.1 Dataset Description

This study employs the Case Western Reserve University (CWRU) rolling bearing dataset [13], focusing on vibration signals from the drive-end bearing under zero-load conditions with a sampling frequency of 12 kHz. To evaluate the model's performance in handling multi-scale data, three fault types are investigated: inner race faults, outer race faults, and rolling element faults. Two fault diameters (0.007 in. and 0.014 in.) were selected, resulting in seven health conditions including normal state.

For each fault condition, 102,400 data points were utilized. Sample segmentation was performed using a sliding window approach with a window length of 1,024 data points and a stride of 512 data points. This processing generated 200 samples per health state, resulting in a complete experimental dataset comprising 1,400 samples. The dataset was split into training and test sets in a 7:3 ratio.

4.2 MSHFIA-CNN Model Structure and Parameter Settings

The main architecture of the MSHFIA-CNN model is shown in Table 1.

Table 1. MSHFIA-CNN model architecture parameters

Layer	Base-Kernel Layer			Convolutional Layer		
	Type	Kernel	Stride	Scale	Width	(C,L)
0	Input	—	—	—	—	(1,1024)
1	PMF	16	2	4	16	(64,512)
2	DRW	—	—	—	—	(64,512)
3	PMF	7	2	4	32	(128,512)
4	DRW	—	—	—	—	(128,512)
5	PMF	3	2	4	64	(256,256)
6	DRW	—	—	—	—	(256,256)
7	GAP	—	—	—	—	(256,1)
8	Softmax	—	—	—	—	ClassNum

The core model consists of three layers of Progressive Multi-Granularity Feature Fusion (PMF), three layers of Dual-Path Residual Weighting (DRW), one Global Average Pooling (GAP) layer, and one SoftMax layer, totaling eight layers. The hyperparameters for PMF include the kernel size and stride in the convolution layers, as well as the number and width of branches in the progressive convolution layers. The input signal has a size of (1,1024). When performing cross-scale attention interaction at different levels, the hyperparameters for downsampling should ensure that the size of shallow feature maps, after attention weighting, matches that of the deeper feature maps. The network optimizer uses the Adam algorithm with a batch size of 32.

4.3 Results and Analysis

To assess the classification performance of MSHFIA-CNN for fault size identification under noisy conditions, Gaussian noise with varying amplitudes was added to the original data from the Case Western Reserve University bearing dataset. The added noise has a mean of zero, and its variance is determined based on the preset signal-to-noise ratio (SNR). The magnitude of the SNR value characterizes the intensity of the noise interference. The SNR is defined as:

$$\text{SNR}_{db} = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \quad (7)$$

Four noise levels (-8 dB, -4 dB, -2 dB, 0 dB) were used for evaluation. For comparison, five state-of-the-art methods were also tested. The fault classification accuracy of the six models under varying noise interference levels is presented in Table 2.

Table 2. Accuracy of different methods under different SNR(%)

Models	SNR(dB)			
	-8	-4	-2	0
WDCNN	69.19	78.90	89.62	94.18
ResNet	72.38	79.27	89.91	94.27
DenseNet	74.05	79.42	91.28	95.84
BiLSTM	75.81	80.81	88.18	95.36
SPANet	76.19	83.72	91.54	97.43
MSHFIACNN	84.86	89.68	93.63	97.32

Experimental results show that as noise intensity increases, the fault recognition accuracy of all models decreases, highlighting the negative impact of noise on feature extraction and model performance. Under low noise conditions, the first four models perform similarly, while SPANet and the proposed method deliver better results due to their cross-scale feature fusion capabilities. As noise intensity rises, the advantage of the proposed method becomes more pronounced. Under extreme -8 dB noise, while other methods experience significant accuracy degradation, the proposed method maintains high recognition precision. This robustness is attributed to its multi-granularity feature extraction and cross-layer attention interaction, effectively modeling key features across fault sizes and improving discrimination of subtle differences.

To further analyze performance under strong noise, confusion matrices of all six models at -8 dB are shown in Fig. 5. Results indicate that under noise interference, the first five mainstream methods frequently misclassify different-sized samples of the same fault type. This occurs because strong noise weakens the expressive power of critical scale features, obscuring fine-grained differences and impairing size-level discrimination. In contrast, the proposed model excels in differentiating fault sizes under noise, thanks to its cross-scale feature attention interaction mechanism. By strengthening inter-layer dependencies and leveraging cross-scale correlations to highlight critical information, it significantly enhances classification robustness and accuracy.

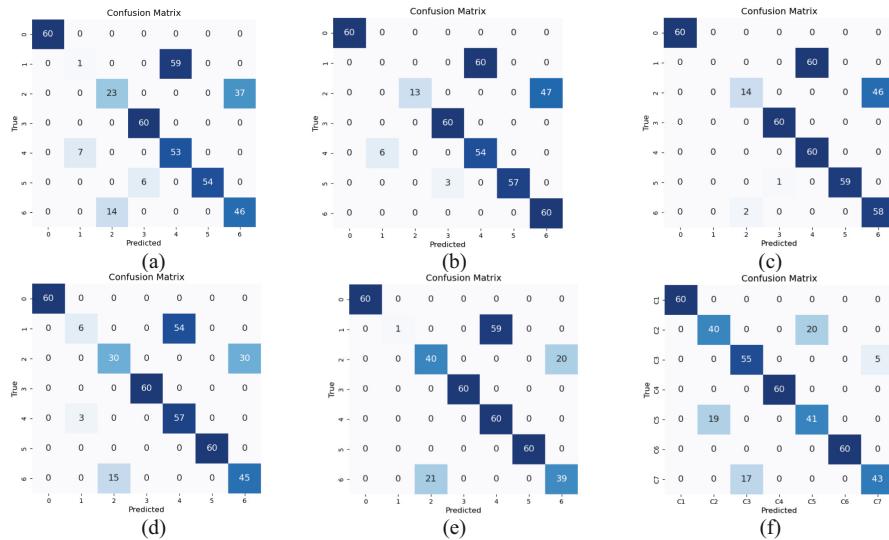


Fig. 5. Confusion matrix of six models for rolling bearing fault diagnosis (SNR = -8 dB). (a) WDCNN; (b) ResNet; (c) DenseNet; (d) BiLSTM; (e) SPANet; (f) MSHFIA-CNN.

5 Conclusion

This paper introduces a Multi-Scale Hierarchical Feature Interaction Attention method for rolling bearing fault diagnosis, aimed at improving the model's ability to perceive and differentiate multi-scale fault features under strong noise interference. The approach uses a progressive multi-granularity feature fusion strategy to combine local and global information hierarchically, enriching feature representation. A dual-path residual weighting strategy is employed to focus on critical channels and spatial regions while maintaining a lightweight network, enhancing sensitivity to subtle fault size differences. The cross-scale attention interaction mechanism enables bidirectional complementarity between shallow and deep semantic features, separating noise through fault feature correlations to boost noise robustness. Experimental results show that MSHFIA-CNN outperforms other methods in terms of robustness and accuracy under various noise levels, proving its effectiveness for industrial applications in complex environments.

References

- Yang, C., Qiao, Z., Zhu, R., et al.: An intelligent fault diagnosis method enhanced by noise injection for machinery. *IEEE Trans. Instrum. Meas.* **72**, 1–11 (2023)
- Chen, Z., Tian, S., Shi, X., et al.: Multiscale shared learning for fault diagnosis of rotating machinery in transportation infrastructures. *IEEE Trans. Ind. Inform.* **19**(1), 447–458 (2022)
- Wei, H., Sun, C., He, J., et al.: Multilayer adaptive denoising network for incipient fault diagnosis of gear system. In: 2024 International Conference on Sensing, Measurement & Data Analytics in the Era of Artificial Intelligence (ICSMD), pp. 1–8. IEEE (2024)

4. Han, S., Sun, S., Zhao, Z., et al.: Deep residual multiscale convolutional neural network with attention mechanism for bearing fault diagnosis under strong noise environment. *IEEE Sensors J.* **24**(6), 9073–9081 (2024)
5. Multiscale Kernel Based Convolutional Neural Networks for Gear Fault Diagnosis Under Variable Operation Conditions
6. Wang, H., Liu, Z., Peng, D., et al.: Understanding and learning discriminant features based on multiattention 1DCNN for wheelset bearing fault diagnosis. *IEEE Trans. Ind. Inform.* **16**(9), 5735–5745 (2019)
7. Shen, Q., Zhang, Z.: Fault diagnosis method for bearing based on attention mechanism and multi-scale convolutional neural network. *IEEE Access.* **12**, 12940–12952 (2024)
8. Wang, D., Li, Y., Jia, L., et al.: Novel three-stage feature fusion method of multimodal data for bearing fault diagnosis. *IEEE Trans. Instrum. Meas.* **70**, 1–10 (2021)
9. Ma, X.N., Qi, L., Zhao, M.: Bearing fault diagnosis based on attention mechanism and deep residual network. In: 2021 International Conference on Computer Engineering and Artificial Intelligence (ICCEAI), pp. 286–290. IEEE (2021)
10. Huang, S., Li, X., Jiang, Z., et al.: Hyper feature fusion pyramid network for object detection. In: 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 1–6. IEEE (2018)
11. Chen, J., Mai, H.S., Luo, L., et al.: Effective feature fusion network in BIFPN for small object detection. In: 2021 IEEE International Conference on Image Processing (ICIP), pp. 699–703. IEEE (2021)
12. Guo, J., Ma, X., Sansom, A., et al.: Spanet: Spatial pyramid attention network for enhanced image recognition. In: 2020 IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6. IEEE (2020)
13. Smith, W.A., Randall, R.B.: Rolling element bearing diagnostics using the Case Western Reserve University data: a benchmark study. *Mech. Syst. Signal Process.* **64**, 100–131 (2015)



Electromechanical Co-simulation Experiment System of Floating Offshore Wind Turbine

Chuanjie Li and Zhaohua Liu^(✉)

Hunan University of Science and Technology, Xiangtan 411100, China
jeic1@mail.hnust.edu.cn, zhaohua.liu@hnust.edu.cn

Abstract. To enhance the precision and efficiency of designing and simulating floating wind turbine systems, this paper presents the design and implementation of an electromechanical co-simulation experimental system for floating offshore wind turbines. The system employs OpenFAST software to create a detailed mechanical model of the floating wind turbine, simulating its dynamic behavior within the complex ocean environment. Subsequently, it utilizes Simulink software to construct the electrical model of the doubly-fed induction generator and its corresponding controller, achieving the co-simulation of the mechanical and electrical models. Considering the aerodynamic force, hydrodynamic force, dynamic response of mechanical structure, working condition of generator and control strategy of electrical system, the high-fidelity simulation of wind power system is realized. In order to enhance the intuitiveness and interactivity of the simulation results, the ParaView software is used to realize the 3D visualization of the simulation results. The results show that the electromechanical co-simulation can be closer to the actual situation, and the visualization platform provides a more intuitive reflection of wind turbine movements.

Keywords: Floating Offshore Wind Power · Electromechanical Co-simulation · Doubly-Fed Induction Generator

1 Introduction

China's Action Plan for Carbon Peak by 2030 emphasizes promoting green transformations in production and lifestyle, advancing carbon peak actions orderly, and basing economic growth on efficient resource utilization and green-low carbon development [1]. As a clean renewable energy, wind power, especially floating offshore wind power, has gained significant attention for its unique advantages in deep-sea applications, overcoming seabed topography/geology limitations to exploit strong wind resources efficiently with minimal marine environmental impact.

However, floating offshore wind power faces critical challenges: complex deep-sea environments such as currents, ice, waves, impose extreme loads on structures, while maintenance accessibility is limited. The system's strong electromechanical coupling requires interdisciplinary knowledge of mechanics, electricity, and control, complicating design and optimization. Developing a visual electromechanical co-simulation platform becomes crucial to model detailed operational behaviors under diverse conditions,

evaluate performance, and provide technical validation for sustainable clean energy development [2].

Current research on onshore wind turbines is mature, but offshore simulation platforms have two key limitations: 1) unbalanced focus on mechanical dynamics over grid-connected control strategies; 2) oversimplified mechanical models neglecting structural complexity and environmental load impacts, leading to simulation-actual deviations. Addressing these, a comprehensive platform must integrate floating platform/turbine mechanical dynamics with generator-converter-based power conversion processes to enable dynamic interactions, bridging the gap between electromechanical coupling theory and engineering practice.

2 Floating Offshore Wind Turbine Dynamics Model

Different from the traditional onshore wind power and the offshore installed fixed wind power, the environment of the floating offshore wind power system is complex and changeable, and it is often affected by the external environment such as wind, wave and ice. Therefore, these external environmental factors must be considered when modeling.

2.1 Aerodynamic Model of Floating Offshore Wind Turbine

It is defined C_P as the wind energy utilization coefficient, which C_P represents the ability of the wind turbine to extract wind energy from the airflow. In the design and control of the fan, it is generally considered as λ a function of the tip speed ratio and the pitch Angle β , and its specific relationship can be approximated by the following equation.

$$\begin{cases} C_p(\lambda, \beta) = c_1(c_2/\lambda_i - c_3\beta - c_4)e^{-c_5/\lambda_i} + c_6\lambda \\ \frac{1}{\lambda_i} = \frac{1}{\lambda+0.08\beta} - \frac{0.035}{\beta^3+1} \end{cases} \quad (1)$$

Where λ is the tip speed ratio; β is pitch Angle; The coefficients $c_1\dots c_6$ are constant and determined by the type of fan blade.

For different pitch angles β , the characteristic $C_P\sim\lambda$ curves are different. When β getting larger, the curve will obviously drop. When pitch Angle and wind speed are fixed, the blade tip speed ratio can be changed by adjusting the fan speed, so as to maximize the wind energy utilization coefficient of the fan.

2.2 Hydrodynamic Model of Floating Wind Turbine

Compared with the land wind turbine, the floating offshore wind turbine has significant differences in structural design and dynamic characteristics. The main reason is that it is built on a floating platform, which allows the wind turbine to have a certain activity space on the sea surface, so as to adapt to the dynamic changes of waves, tides and currents. The movement degrees of freedom of floating offshore wind turbines include the following six types: Surge, Sway, Heave, Pitch, Roll, and Yaw [4].

Floating platforms suffer from the dynamic effects of wave forces in the ocean, causing platform fluctuations. These wave loads are mainly derived from the coupling

effect of radiation stress, diffraction effect and hydrostatic pressure, which together constitute the wave loads borne by the platform. The expression is:

$$F_i = F_i^{\text{rad}} + F_i^{\text{waves}} + F_i^{\text{Hydrostatic}} \quad (2)$$

Where the subscript i denotes a degree of freedom; F_i Denotes the dynamic load received in the direction of this degree of freedom;

Radiation load F_i^{rad} describes the reaction force of wave radiation generated by the movement of the platform on the platform without the influence of incident waves. Diffraction load reflects the force effect on the platform caused by the incident wave scattered by the floating platform. F_i^{waves} In view of the random characteristics of natural waves, the wave amplitude spectrum is usually used to characterize this random process. $F_i^{\text{Hydrostatic}}$ is the total load of linear hydrostatic action on the floating platform.

2.3 Wind Turbine Simulation Software

This paper uses the OpenFAST simulation software to simulate the mechanical response of wind turbines.

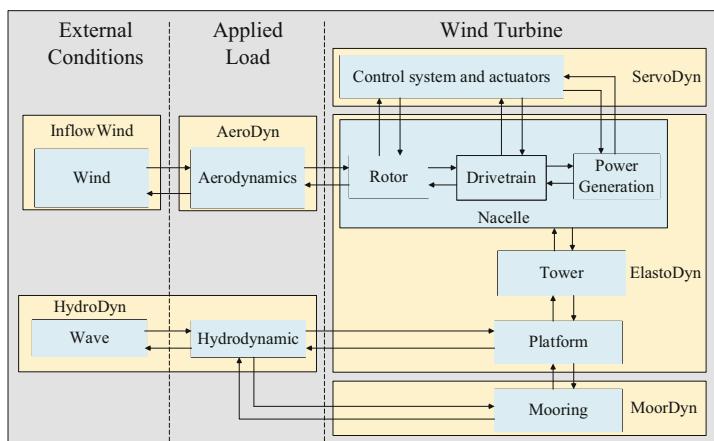


Fig. 1. Schematic diagram of the coupling of each OpenFAST module

OpenFAST software, developed by the United States National Renewable Energy Laboratory [5], is designed for wind turbine simulation and calculation. It encompasses a range of wind turbine types and models. These include structural dynamics calculation files, finite element blade structure dynamics files, aerodynamic calculation files, inflow wind files, electrical and control drive dynamics files, hydrodynamic calculation files, floating platform dynamics files, mooring system files, and ice load files. The relationships between the modules in each file are illustrated in Fig. 1:

3 Control Strategy of DFIG

The doubly-fed motor's alternating current (AC) excitation on the rotor allows for the adjustment of not only the amplitude but also the phase and frequency. This capability enables speed adjustment through changes in the excitation current frequency and motor power angle adjustment by altering the phase, thereby controlling both reactive and active power [6].

3.1 Grid-Side Converter Control Strategy

The grid-side converter adopts the directional vector control strategy based on the grid voltage, which includes the closed-loop regulation of current and voltage to control the energy of the grid. This control strategy allows the converter to switch between rectifier and inverter states according to the system demand, thus regulating the dc-side voltage and controlling the reactive power [8]. The control objective is mainly to keep the voltage on the DC bus stable, and only this voltage is stable, the machine side converter can work properly.

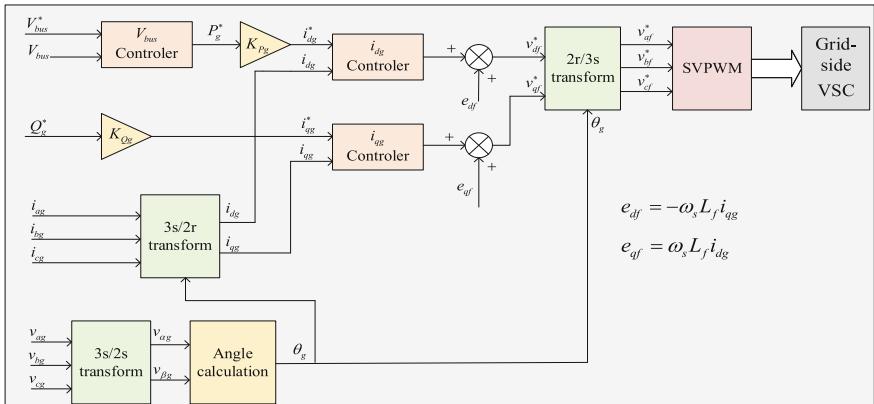


Fig. 2. Block diagram of vector control for grid-side converter based on grid voltage orientation

Figure 2 shows the grid voltage vector control block diagram of the grid-side converter. To achieve control, multiple voltage and current sensors are needed, and the quantities to be detected are: three-phase voltage v_{ag} , v_{bg} , v_{cg} on the grid side, three-phase current i_{ag} , i_{bg} , i_{cg} , and bus voltage V_{bus} on the DC side. The three-phase voltage and current collected by the sensor are transformed by Clark (3s/2s) and Parker (2s/2r) coordinates, and the voltage Angle of the grid θ_g is calculated by Clark transformation of the three-phase voltage of the grid.

The control includes two control loops, the voltage outer loop and the current inner loop, the voltage outer loop is mainly to control the voltage of the DC bus, the current regulator uses PI control, and the PI controller parameters can be converted from the current components i_{dg} and i_{qg} of the dq axis to the reference values u_{df}^* and u_{qf}^* of the

voltage components of the dq axis according to the mathematical model of the grid-side converter. The coupling terms e_{df} and e_{qf} in the model equation are given in the figure according to this formula. By transforming the dq rotating coordinate system, the d-axis is oriented to the grid voltage space vector. According to the grid-side mathematical model, adjusting i_{dg} can control the active power P_g , and adjusting i_{qg} can control the reactive power Q_g . After decoupling the control loop and then performing the Park inverse transformation (2r/3s) back to the three-phase stationary coordinate system, the space vector pulse width modulation (SVPWM) algorithm is used to control the grid-side converter.

3.2 Rotor-Side Converter Control Strategy

The Rotor-side converter employs stator flux-oriented vector control technology to achieve decoupled active and reactive power regulation at the stator terminal. The electromagnetic torque reference is determined through real-time generator speed measurement, enabling speed control via torque adjustment. Maximum power point tracking (MPPT) is implemented by optimizing the wind turbine's tip-speed ratio, thereby enhancing wind energy conversion efficiency [9].

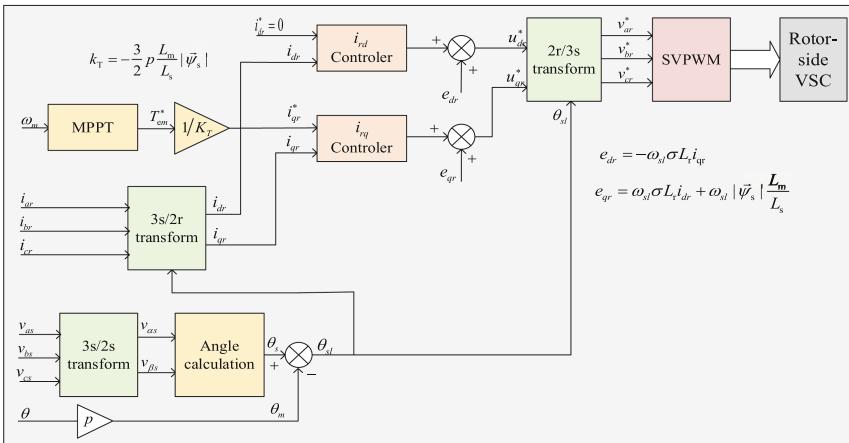


Fig. 3. Block diagram of the machine-side converter based on stator flux linkage directional vector control

As shown in Fig. 3, the control block diagram of the rotor-side converter, ω_m is the measured generator speed. The desired electromagnetic torque T_{em}^* is obtained through the MPPT control algorithm, and multiplying by the torque-component transformation gain $1/K_T$ from electromagnetic torque to rotor current gives the reference value i_{qr}^* of the rotor current q-axis component. Adopting the control strategy $i_{dr}^* = 0$ for the d-axis current, the actual values of the rotor dq-axis currents are obtained via measurement and coordinate transformation. e_{dr} and e_{qr} are compensation terms. Adding these terms per their mathematical model yields the expected values u_{dr}^* and u_{qr}^* of the rotor dq-axis voltage components, achieving decoupling. Then, via inverse Park transformation

($2r/3s$), it returns to the three-phase stationary coordinate system. θ_{sl} is the slip angle; θ_s is stator current electrical angle; θ_m is the rotor mechanical angle.

The MPPT control is under the rated wind speed, given the torque to control the generator speed, so as to adjust the impeller speed to change the tip speed ratio, the tip speed ratio is controlled at a certain value can make the maximum wind energy utilization coefficient [10, 11].

3.3 Joint Simulation Experiment

According to the fan model established by OpenFAST and the electrical and control models established by OpenFAST, the joint simulation model of the two was constructed. Coupling of each software diagram as shown in Fig. 4, the dotted box represents the software or module. TurbSim software generates turbulent wind, HydroDyn module generates wave force, OpenFAST calls aerodynamics module AeroDyn to calculate aerodynamic load of wind turbine blade generated by turbulent wind data, and receives aerodynamic force data from AeroDyn [12]. Based on these data and the structural parameters of the wind turbine, the structural motion calculation of the whole machine is carried out. At the same time, it will also couple the influence from the wave force, the wave first affects the floating platform and the mooring system, and then is transmitted to the wind turbine tower and other structures.

The wind turbine simulation model involves multiple time scales, covering multiple physical processes. Due to the differences in their intrinsic characteristics, these processes have different time constants and have different frequency effects on the system. In view of the relatively slow dynamic response of multi-body dynamics in wind turbines, the main characteristic frequency is usually below 20 Hz, and the transient process of the electrical part may contain high frequency components [13]. Therefore, in the electromechanical co-simulation, it is necessary to set the solution step reasonably according to the time scale of the electrical model. Here, the step size is set to 5e-5 to meet the requirements of electromagnetic transient simulation. Solver selection of Simulink ode4 solver, because its in the joint has good numerical stability [14].

The out the file after extracting the required data in MATLAB can be saved as a CSV file for read ParaView, extraction under the condition of 3 fan here six degrees of freedom movement information, wind speed and wave height in the software for the visualization, as shown in Fig. 5, clearly showing the Wind turbine, the movement of information, Graph in the vertical bar with moves to the right, the simulation time animation fans are synchronized, convenient read the observation data of each time node information, so then with the visualization process of the Wind turbine.

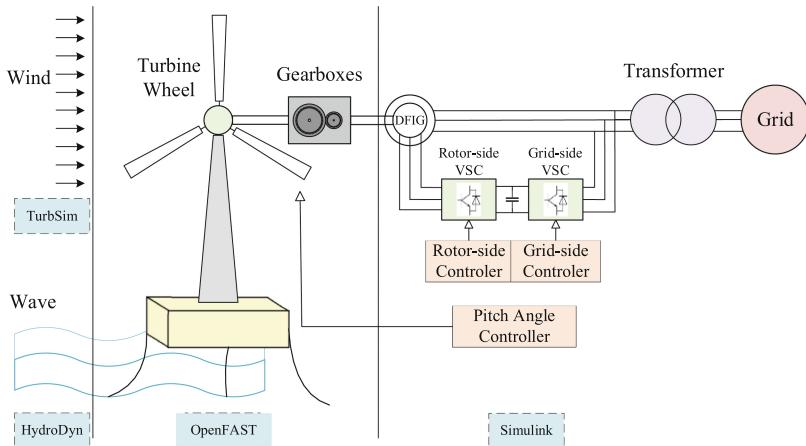


Fig. 4. Schematic diagram of the co-simulation model framework

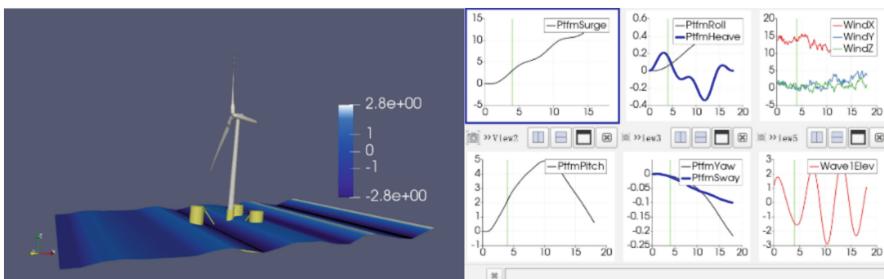


Fig. 5. Visualization and movement data of Wind turbine

4 Conclusion

This paper comprehensively integrates the mechanical dynamics of floating platforms and wind turbines with the electrical conversion processes of generators and converters, establishing dynamic interactions between mechanical and electrical models. Using OpenFAST, a high-fidelity mechanical model of floating wind turbines is developed to simulate platform dynamics under complex marine environments. Subsequently, a DFIG electrical model and controller are built in Simulink, enabling co-simulation of mechanical and electrical subsystems. This co-simulation incorporates aerodynamic, hydrodynamic, structural dynamic responses, generator operation, and control strategies for high-fidelity system modeling. By aligning simulation results with real-world offshore scenarios, the approach significantly enhances prediction accuracy and practical applicability of wind power system simulations. In addition, in order to improve the user experience and facilitate data analysis, the ParaView visualization platform is used to visually display the simulation results. Through the graphical interface, users can easily view the operation status, performance indicators and other data of the wind turbine.

References

1. Liu, X., Han, X., Qin, M.: China Offshore Oil Gas. **36**(02), 233–242 (2024)
2. Wang, F., Liu, Y.: In to the Deep Sea for wind power generation. Sci. Chin. (11), 26–28 (2023)
3. Meng, Q., Sun, H., Zhong, Z., et al.: Development and experiment of visualization electromechanical co-simulation platform for floating offshore wind power. Exp. Technol. Manag. **40**(10), 119–127+158 (2023)
4. Jonkman, J.M., Buhl, M.L.: FAST User's Guide. National Renewable Energy Laboratory, Golden (2005)
5. Jonkman, J.M., Butterfield, S., Musial, W., et al.: Definition of a 5-MW Reference Wind Turbine for Offshore System Development. National Renewable Energy Lab.(NREL), Golden (2009)
6. Liu, S.J., Zhuang, S.X., Shu, X.D.: Power control of doubly-fed wind turbine based on stator voltage direction. Large Motor Technol. (04), 8–41 (2010)
7. Miao, F.: Research on Influence of Electromechanical Coupling and Active Load Control of Doubly-Fed Wind Turbine. Beijing Jiaotong University, Beijing (2017)
8. Tian, D., Liu, F., Fang, J., et al.: Research on gas-mechanical-electric coupling modeling and simulation of wind turbine transmission chain. Wind Energy. (12), 72–78 (2021)
9. Singh, M., Muljadi, E., Jonkman, J., et al.: Simulation for Wind Turbine Generators—With FAST and MATLAB-Simulink Modules. National Renewable Energy Lab. (NREL), Golden (2014)
10. Gonzalo Abad, Awad, Lopez, et al. Modeling and Control of Doubly-Fed Induction Motor in Wind Power Generation. China Machine Press, Beijing 2014.
11. Ling, Y.: Modeling, Simulation and Control of Doubly-Fed Wind Power System. China Machine Press, Beijing (2017)
12. Xiong, H.: Research on Modeling and Control of Large-Scale Wind Turbine System Based on FAST Software. Chongqing University, Chongqing (2014)
13. Zhou, Y.: Research on Dynamic Characteristics of MW Stage Doubly-Fed Asynchronous Wind Turbine. Chongqing University, Chongqing (2016)
14. Bai, Y., Liu, H.: Simulation analysis of running state of floating fan based on OPENFAST. Metrol. Test. Technol. **51**(04), 49–52 (2024)

Author Index

A

- An, Kang 43, 212, 375
Ao, Hongrui 21

C

- Cao, Shuo 260
Cao, Yang 278
Chang, Hanwen 469
Chen, Cheng 444
Chen, Geng 365
Chen, JingSong 335
Chen, Mingming 137
Chen, Qingwei 116
Chen, Shengzong 21
Chen, Yuxin 64
Cheng, Yun 307
Cheng, Zhenbo 64
Cui, Bowen 230, 394

D

- Dai, Hengrui 157
Dai, Qingyou 64
Ding, Jisong 32
Ding, Lu 105
Ding, Shihong 128
Du, Haibo 384
Du, Ku 287
Du, Xiaotao 230, 394
Du, Xiaoxiong 260

F

- Fu, Haohao 64

G

- Gao, Yang 116
Gong, Yeting 317
Guo, Bohan 157
Guo, Di 55
Guo, ZhengHua 335

H

- Han, Hui 297
Hao, Xiaopu 222
He, Jiahao 86
Hao, Lanyang 230, 394
He, Shuqian 425
He, Yalin 307
He, Zhiyong 240
Hong, Jianyun 457
Hou, QiDong 335
Hu, Xiong 354
Huang, Bo 307
Huang, Dan 105
Huang, Darong 287
Huang, Huixin 365
Huang, Jun 278
Huang, Xun 105
Huang, Yupei 167
Huo, Benyan 434

J

- Ji, Xiaopeng 230, 394
Jiang, Hanrui 1

K

- Kang, Shuai 222, 425
Kang, Yangming 328
Kang, Yu 278

L

- Lai, Rongbo 128
Lan, Bin 86
Li, Bo 457
Li, Boxiang 200
Li, Chuanjie 480
Li, Hanting 212
Li, Jiaming 86
Li, Jiaqi 444
Li, Longchuan 222, 425
Li, Peng 167
Li, Tenglong 137

Li, Yibin 147
 Li, Yu 457
 Li, Yuanpeng 457
 Li, Yuchen 384
 Li, Zihan 251
 Liang, Jian 328
 Liang, Lunfei 86, 260, 345, 365
 Liang, Zhening 10
 Liu, Chundong 222
 Liu, Houde 345
 Liu, Lu 128, 434
 Liu, Mengyan 157
 Liu, Song 365
 Liu, Tianyu 32
 Liu, Xiangpeng 43
 Liu, Xiaoyuan 297
 Liu, Yanhong 434
 Liu, Yu 260, 345
 Liu, Yudan 444
 Liu, Zhaohua 480
 Lu, Shan 251
 Lu, Yu 137
 Lv, Pin 328

M
 Ma, Ding 384
 Ma, Xin 147
 Ma, Yurun 444
 Mei, Keqi 128
 Meng, Fanteng 93
 Meng, Tong 93

N
 Na, Yuhong 287
 Nie, Yujie 147
 Ning, Lichuan 189, 407
 Ning, Tao 335
 Niu, Ziru 1

P
 Pan, Yan 365
 Pang, Denghao 200
 Pang, Wenhong 116
 Peng, Yiqun 32
 Peng, Yulin 43

Q
 Qin, Yong 93
 Qiu, Ninghai 93

R
 Ruan, Jianyuan 457
 Ruan, Xianyong 407

S
 Shao, Dongsheng 260
 Shi, Kehao 278
 Shi, Xiao 240
 Shi, Yuhang 317
 Song, Lei 105
 Song, Yaqing 212
 Song, Zhiqiang 222, 425
 Sun, Haipeng 157
 Sun, Jian 157
 Sun, Jiechen 469
 Sun, Jinlin 128
 Sun, Lu 434
 Sun, Xunhong 384

T
 Tang, Caiguo 76
 Tang, Hailin 417
 Tao, Bo 407

W
 Wang, Cheng 287
 Wang, Chenggang 105
 Wang, Dayong 76
 Wang, Helin 271
 Wang, Maoyu 64
 Wang, Panpan 55
 Wang, Qiang 1
 Wang, Xin 147
 Wang, Yahan 147
 Wang, Yi 317
 Wang, Yue 375
 Wang, Zhipeng 93
 Wei, Jiamin 297
 Weng, Ling 230, 394
 Wu, Fan 175
 Wu, Zhengxing 167

X
 Xia, Wei 200
 Xiang, Benlong 189
 Xiao, Xibin 1
 Xie, Yuanmin 189, 407
 Xu, Chaonong 175
 Xu, Gaoli 354

- Xu, Xiangqing 147
Xu, Xinying 297
Xu, Xuan 317
Xu, Zhenyi 278
- Y**
Yan, Gang 86
Yan, Hao 64
Yang, Jie 32
Yang, Tangwen 317
Yang, Yang 345
Yang, Yaoxiang 1
Yang, Yunyun 297
Yu, Aihua 457
Yu, Chongchong 93
Yu, Han 157
Yu, Junzhi 167
Yu, Ming 10
Yuan, Yinlong 307
- Z**
Zeng, Han 76
Zeng, Qingwang 10
- Zhang, Hui 43, 375
Zhang, Jihui 21
Zhang, Jin 335
Zhang, Qiang 222
Zhang, Qin 354
Zhang, Tianping 417
Zhang, Yaxing 328
Zhang, Yuwei 175
Zhang, Zan 434
Zhao, Guancheng 425
Zhao, Jia 147
Zhao, Runchen 384
Zhou, Diting 64
Zhou, Funa 469
Zhou, Jiansheng 260
Zhou, Linli 76
Zhou, Yucai 240
Zhou, Yunhu 287
Zhu, Jinliang 200
Zhu, Weibin 317
Zhu, Wenwu 384
Zhu, Yue 10
Zuo, Shichao 230, 394