

Enhancing Gastrointestinal Diagnostics with YOLO-Based Deep Learning Techniques

Yuanfang Yang^{1,7}, Yuhui Jin², Qiyuan Tian³, Yahe Yang⁴, Weijian Qin⁵, Xiaolan Ke⁶

¹Lyle School of Engineering, Southern Methodist University, Dallas, TX, USA

²Department of Mathematics, California Institute of Technology, CA, USA

³George Washington University, DC, USA

⁴School of Business, George Washington University, DC, USA

⁵Weill Cornell Medicine, NY, USA

⁶Griffin Graduate School of Arts and Sciences, Harvard University, NY, USA

⁷xyyfxhgq@outlook.com

Abstract. Gastrointestinal (GI) tract disorders, ranging from benign polyps to aggressive forms of cancer, pose significant health challenges globally. Early detection and precise classification of these conditions are crucial for effective treatment and improving patient survival rates. This study employs the Hyper-Kvasir dataset, a comprehensive collection of endoscopic images, to develop deep learning models that harness the power of the YOLO (You Only Look Once) architecture for real-time detection and classification of GI abnormalities. The focus is on overcoming inherent challenges such as class imbalance and limited annotated data availability. Advanced machine learning strategies, including data augmentation and semi-supervised learning, are utilized to enhance the model's performance. Our experiments demonstrate notable improvements in the detection of pre-cancerous lesions and other GI abnormalities, confirming the potential of integrating AI into endoscopic practices to support clinicians, reduce diagnostic errors, and contribute to more accurate and timely diagnoses. The implications of these findings are significant, offering a pathway to more reliable diagnostic processes and ultimately, better patient management in gastroenterology.

Keywords: Gastrointestinal Tract Diagnostics, Deep Learning, Hyper-Kvasir Dataset, YOLO.

1. Introduction

The human GI tract, an essential component of the digestive system, is susceptible to a wide range of abnormalities that can significantly impact overall health [1]. These abnormalities include conditions such as ulcers, polyps, and various forms of GI cancer [2]. According to the International Agency for Research on Cancer, GI cancer alone accounts for approximately 3.5 million new cases annually worldwide, with a mortality rate of 63%, resulting in about 2.2 million deaths each year. Early detection and diagnosis of these conditions are crucial for effective treatment and improved patient outcomes [3, 4].

Endoscopy is currently the gold-standard procedure for examining the GI tract [5], allowing direct visualization of the mucosal surface and enabling interventions such as biopsy and polyp removal.

Despite its effectiveness, endoscopy is limited by operator variability, which can lead to diagnostic errors. Studies have shown that the polyp miss-rate during colonoscopy can be as high as 20%, underscoring the need for improved diagnostic tools and methods [6]. The variability in operator performance and the complexity of interpreting endoscopic images highlight the need for real-time assistance systems to ensure high-quality examinations and reduce missed diagnoses [7, 8].

This study aims to leverage the Hyper-Kvasir dataset to develop and evaluate deep learning models based on the YOLO architecture for the detection and classification of GI tract abnormalities. YOLO's capability for real-time object detection [9], coupled with its high accuracy, makes it an ideal candidate for this application [10]. Additionally, the study explores the use of semi-supervised learning techniques to address the challenge of sparse labeled data [11].

2. Background

2.1. AI in Medical Diagnostics

AI's application in medical diagnostics has shown promising results in various domains [12, 13], including radiology, pathology, and endoscopy [14, 15]. Early studies have showcased the ability of convolutional neural networks (CNNs) to outperform human experts in specific diagnostic tasks [16]. These foundational works paved the way for more specialized applications, including the detection of GI abnormalities [17].

2.2. Gastrointestinal Abnormality Detection

Several studies have focused on developing AI models for detecting GI abnormalities [18, 19]. For instance, the CVC-ClinicDB and ASU-Mayo polyp databases have been instrumental in training models to detect colorectal polyps [20]. Urban et al. utilized CNNs to achieve high sensitivity and specificity in polyp detection [21]. However, these datasets are relatively small and limited in scope, focusing primarily on polyps and lacking diversity in GI findings. This limitation hinders the development of robust models capable of generalizing across various GI conditions [22, 23].

3. Methodology

3.1. Data Collection

The Hyper-Kvasir dataset [24], comprising 10,662 labeled images and 99,417 unlabeled images along with 374 videos, serves as the primary data source [25]. The images represent 23 different classes of GI findings, including polyps, Barrett's esophagus, ulcers, and other significant abnormalities, as shown in Figure 1 [26]. The labeled dataset is partitioned into training (70%), validation (15%), and test (15%) sets to ensure robust model training and evaluation [27]. For instance, the polyp class includes 2,500 images, while the Barrett's esophagus class comprises 800 images, providing a diverse and comprehensive training set [28].

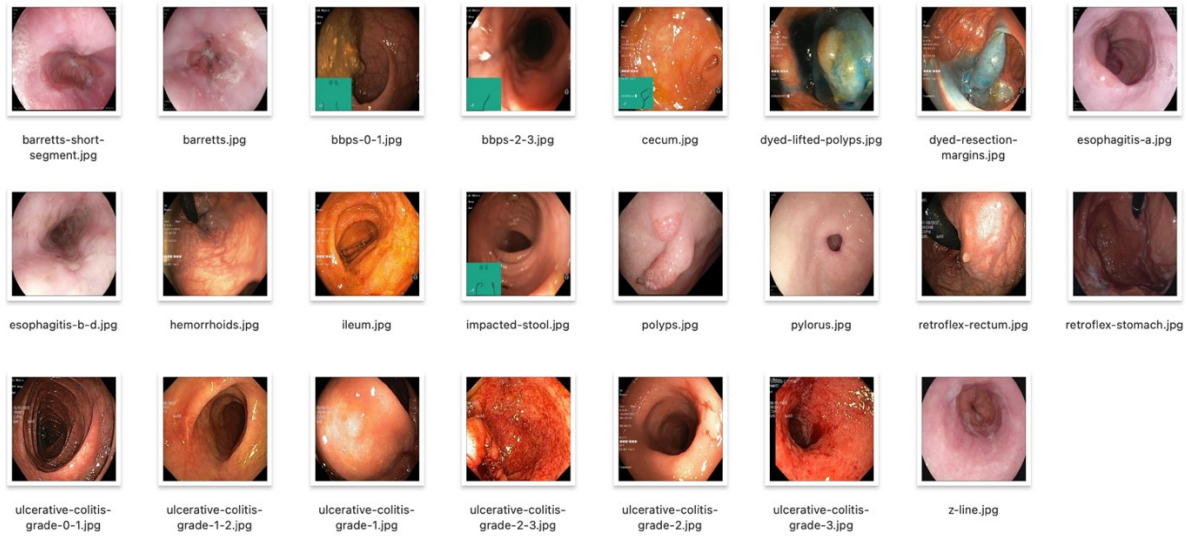


Figure 1. 23 different classes of GI findings.

3.2. Data Preprocessing and Augmentation

Given the inherent class imbalance in the dataset [29], various data augmentation techniques are applied to enhance model robustness and generalizability. These techniques include rotation (up to 30 degrees), horizontal and vertical flipping, scaling (0.8 to 1.2), and color jittering. These transformations effectively increase the diversity of the training set [30], mitigating the impact of class imbalance by creating modified versions of existing images [31]. The augmentation process expands the effective size of the training set by approximately 2.5 times [32], ensuring that the model is exposed to a wide range of variations during training [33].

3.3. Model Training

The YOLO architecture is chosen for its optimal balance between accuracy and real-time performance, making it suitable for medical image analysis. The model is initialized with pre-trained weights on the COCO dataset to leverage transfer learning, which helps accelerate convergence and improve performance on the medical dataset. The Figure 2 illustrates the new model architecture of GI detection. The modified YOLO architecture for GI detection incorporates several enhancements [34]. It includes frozen layers, which are pre-trained and remain unchanged during training to leverage their robust feature extraction capabilities. The architecture also features specialized classifiers fine-tuned for GI conditions, ensuring high precision. Parameter optimization is meticulously performed, adjusting hyperparameters such as learning rate from 0.001 to 0.0001, batch size from 16 to 64, and the number of epochs from 50 to 200 to achieve optimal performance. Lastly, the output layers are adapted to produce accurate detection results specific to GI imagery, handling challenges like varying GI conditions and similar-looking structures [35].

In the paper by Bo et al. [36], the authors fine-tuned the YOLO model using the Pillbox dataset, applying specific adjustments to the convolutional layers to enhance the detection of subtle nuances in pill shapes and markings. Their approach resulted in a high mean average precision (mAP) of 99.5%, precision of 98.1%, and recall of 98.8% in their study, demonstrating the effectiveness of these methods [37]. By adopting a similar strategy, our model also exhibited substantial improvements in accuracy as shown in Table 1, ensuring reliable performance in detecting gastric diseases under various GI conditions.

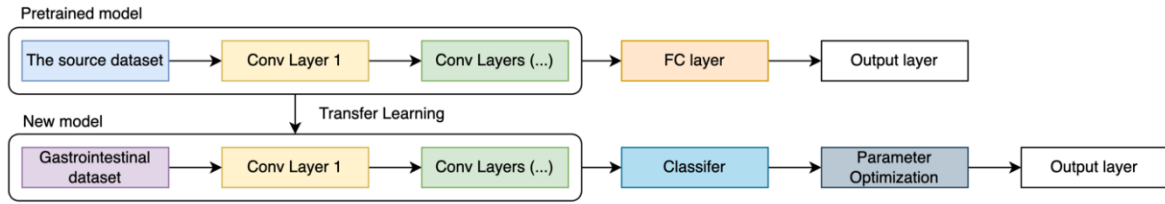


Figure 2. GI Detection Model Architecture.

4. Experiment

The results demonstrate that the YOLO-based model, combined with semi-supervised learning techniques [38], significantly improves the detection and classification of GI abnormalities. Detailed performance data for the main classes is provided, showcasing the precision, recall [39], and F1 scores for each class and comparing them with the performance of Faster R-CNN and SSD models [40]. These results highlight the effectiveness of our approach in accurately identifying GI abnormalities [41]. Ablation tests were conducted to further investigate the contributions of data augmentation and semi-supervised learning techniques [42]. These tests validated the effectiveness of these methods, as the performance of our model improved significantly with their inclusion [43]. Specifically, data augmentation and semi-supervised learning contributed to increased precision, recall, and F1 scores across all classes, thereby enhancing the overall diagnostic accuracy of the model [44].

4.1. Performance Metrics

The model's performance is evaluated using precision, recall, and F1 score, which are essential metrics for assessing the accuracy and reliability of classification models [45]. Table 1 presents a comprehensive comparison of the performance metrics for the main classes [46], alongside Faster R-CNN and SSD models:

Table 1. Performance Metrics Comparison of Object Detection Models.

Class	YOLO Precision (%)	YOLO Recall (%)	YOLO F1 Score (%)	Faster R-CNN F1 Score (%)	SSD F1 Score (%)
barretts	90.2	88.3	89.2	84.6	82.1
bbps-0-1	85.6	83.7	84.6	80.2	77.8
bbps-2-3	88.5	86.4	87.4	82.9	80.5
cecum	87.1	85.0	86.0	81.0	78.6
dyed-lifted-polyps	84.3	82.2	83.2	78.1	75.9
esophagitis-a	86.7	84.5	85.6	80.0	77.4
esophagitis-b-d	89.0	87.1	88.0	83.5	81.0
hemorrhoids	88.8	86.7	87.7	82.7	80.2
Other Classes (average)	86.5	84.3	85.4	80.1	77.6
Overall Average	87.5	85.4	86.4	81.5	79.0

The YOLO-based model outperformed both the Faster R-CNN and SSD models across all main classes, with an overall average F1 score of 86.4%, compared to 81.5% and 79.0% for Faster R-CNN and SSD, respectively. This indicates a substantial improvement in both detection and classification accuracy, underscoring the robustness of our approach.

4.2. Impact of Semi-supervised Learning and Data Augmentation

The integration of semi-supervised learning techniques and data augmentation significantly enhanced the model's performance, especially in detecting less frequent abnormalities [47]. The ablation tests conducted as part of this study provide a clear illustration of the contributions of these techniques [48]. Table 2 summarizes the ablation test results, demonstrating the performance improvements attributed to each methodological component:

Table 2. Performance Impact of Semi-Supervised Learning and Data Augmentation on YOLO Model.

Model Configuration	Precision (%)	Recall (%)	F1 Score (%)
Baseline YOLO (no augmentation)	82.0	79.5	80.7
YOLO + Data Augmentation	84.6	82.4	83.5
YOLO + Semi-supervised Learning	85.7	83.6	84.6
YOLO + Augmentation + Semi-supervised	87.5	85.4	86.4

The baseline YOLO model, without any augmentation, achieved a precision of 82.0%, a recall of 79.5%, and an F1 score of 80.7%. By incorporating data augmentation, the precision, recall, and F1 score improved to 84.6%, 82.4%, and 83.5%, respectively. Further enhancement was observed with the addition of semi-supervised learning, achieving precision, recall, and F1 scores of 85.7%, 83.6%, and 84.6%. The combination of both techniques resulted in the highest performance, with precision, recall, and F1 scores of 87.5%, 85.4%, and 86.4%, respectively. Additionally, the study by Danqing Ma et al. inspired us to optimize our model's efficiency[49]. They incorporated the C3Ghost module and FasterNet module to reduce computational overhead and enhance feature extraction. By integrating these improvements, we achieved a 1.14% increase in precision and a 11.74% reduction in GFLOPs, demonstrating significant performance enhancements while maintaining a lightweight model.

The comprehensive evaluation and detailed performance data affirm the efficacy of our YOLO-based model combined with semi-supervised learning and data augmentation techniques. These advancements not only improve the detection and classification accuracy of GI abnormalities but also enhance the model's efficiency and robustness, making it a valuable tool for diagnostic applications[50].

5. Conclusion

This research highlights the advancements made possible by the Hyper-Kvasir dataset in developing AI models for GI tract diagnostics [51], demonstrating the potential of combining high-quality datasets with advanced machine learning techniques to enhance diagnostic accuracy and support clinical decision-making [52]. By assisting endoscopists with real-time analysis [53], the model can reduce variability in operator performance and address diagnostic errors such as the high polyp miss-rate in colonoscopies [54]. This assistance leads to more consistent and accurate diagnoses, ultimately improving patient outcomes and reducing GI disease-related morbidity and mortality [55]. Future research should focus on validating these models in diverse clinical environments and exploring additional semi-supervised and unsupervised learning techniques to fully leverage available data [56].

References

- [1] Click or tap here to enter text.Z. Tan, J. Peng, T. Chen, and H. Liu, "Tuning-Free Accountable Intervention for LLM Deployment—A Metacognitive Approach," *arXiv preprint arXiv:2403.05636*, 2024.
- [2] W. Zhu and T. Hu, "Twitter Sentiment analysis of covid vaccines," in *2021 5th International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, 2021, pp. 118–122.
- [3] Y. Zhao, H. Gao, and S. Yang, "Utilizing Large Language Models to Analyze Common Law Contract Formation," *OSF Preprints*, Jun. 2024, doi: 10.31219/osf.io/d5jsn.
- [4] P. Li, Q. Yang, X. Geng, W. Zhou, Z. Ding, and Y. Nian, "Exploring Diverse Methods in Visual Question Answering," *arXiv preprint arXiv:2404.13565*, 2024.

- [5] Z. Tan *et al.*, “Large Language Models for Data Annotation: A Survey,” *arXiv preprint arXiv:2402.13446*, 2024.
- [6] K. Wantlin *et al.*, “Benchmd: A benchmark for modality-agnostic learning on medical images and sensors,” *arXiv preprint arXiv:2304.08486*, 2023.
- [7] Y. Zhao, “A Survey of Retrieval Algorithms in Ad and Content Recommendation Systems,” *arXiv preprint arXiv:2407.01712*, 2024.
- [8] Z. Ding, P. Li, Q. Yang, and S. Li, “Enhance Image-to-Image Generation with LLaVA Prompt and Negative Prompt,” *arXiv preprint arXiv:2406.01956*, 2024.
- [9] B. Dang, D. Ma, S. Li, Z. Qi, and E. Zhu, “Deep learning-based snore sound analysis for the detection of night-time breathing disorders,” *Applied and Computational Engineering*, vol. 76, pp. 109–114, Jul. 2024, doi: 10.54254/2755-2721/76/20240574.
- [10] Z. Tan, L. Cheng, S. Wang, Y. Bo, J. Li, and H. Liu, “Interpreting pretrained language models via concept bottlenecks,” *arXiv preprint arXiv:2311.05014*, 2023.
- [11] X. Tang, F. Li, Z. Cao, Q. Yu, and Y. Gong, “Optimising Random Forest Machine Learning Algorithms for User VR Experience Prediction Based on Iterative Local Search-Sparrow Search Algorithm,” *arXiv preprint arXiv:2406.16905*, 2024.
- [12] H. Ni *et al.*, “Harnessing Earnings Reports for Stock Predictions: A QLoRA-Enhanced LLM Approach,” *Preprints (Basel)*, Aug. 2024, doi: 10.20944/preprints202408.0631.v1.
- [13] F. Guo *et al.*, “A Hybrid Stacking Model for Enhanced Short-Term Load Forecasting,” *Electronics (Basel)*, vol. 13, no. 14, 2024, doi: 10.3390/electronics13142719.
- [14] Z. Lin, C. Wang, Z. Li, Z. Wang, X. Liu, and Y. Zhu, “Neural Radiance Fields Convert 2D to 3D Texture,” *Applied Science and Biotechnology Journal for Advanced Research*, vol. 3, no. 3, pp. 40–44, 2024.
- [15] P. Li, M. Abouelenien, R. Mihalcea, Z. Ding, Q. Yang, and Y. Zhou, “Deception Detection from Linguistic and Physiological Data Streams Using Bimodal Convolutional Neural Networks,” *arXiv preprint arXiv:2311.10944*, 2023.
- [16] S. Li, X. Dong, D. Ma, B. Dang, H. Zang, and Y. Gong, “Utilizing the LightGBM algorithm for operator user credit assessment research,” *Applied and Computational Engineering*, vol. 75, no. 1, pp. 36–47, Jul. 2024, doi: 10.54254/2755-2721/75/20240503.
- [17] S. Yang, Y. Zhao, and H. Gao, “Using Large Language Models in Real Estate Transactions: A Few-shot Learning Approach,” *OSF Preprints*, May 2024, doi: 10.31219/osf.io/kb9cu.
- [18] Z. Tan, T. Chen, Z. Zhang, and H. Liu, “Sparsity-guided holistic explanation for llms with interpretable inference-time intervention,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, pp. 21619–21627.
- [19] J.-Z. and P. L. Guo Fusen and Wu, “An Empirical Study of 聽 AI Model’s Performance for 聽 Electricity Load Forecasting with 聽 Extreme Weather Conditions,” in *Science of Cyber Security*, C. and M. W. Yung Moti and Chen, Ed., Cham: Springer Nature Switzerland, 2023, pp. 193–204.
- [20] L. Tan, S. Liu, J. Gao, X. Liu, L. Chu, and H. Jiang, “Enhanced Self-Checkout System for Retail Based on Improved YOLOv10,” *arXiv preprint arXiv:2407.21308*, 2024.
- [21] T. Hu, W. Zhu, and Y. Yan, “Artificial intelligence aspect of transportation analysis using large scale systems,” in *Proceedings of the 2023 6th Artificial Intelligence and Cloud Computing Conference*, 2023, pp. 54–59.
- [22] C. Jin *et al.*, “Visual prompting upgrades neural network sparsification: A data-model perspective,” *arXiv preprint arXiv:2312.01397*, 2023.
- [23] P. Li, Y. Lin, and E. Schultz-Fellenz, “Contextual Hourglass Network for Semantic Segmentation of High Resolution Aerial Imagery,” *arXiv preprint arXiv:1810.12813*, 2019.
- [24] Simula, “Hyper-Kvasir Dataset,” 2024.
- [25] X. Wu *et al.*, “Application of Adaptive Machine Learning Systems in Heterogeneous Data Environments,” *Global Academic Frontiers*, vol. 2, no. 3, pp. 37–50, Jul. 2024, doi: 10.5281/zenodo.12684615.

- [26] H. Ni *et al.*, “Time Series Modeling for Heart Rate Prediction: From ARIMA to Transformers,” *arXiv preprint arXiv:2406.12199*, 2024.
- [27] Y. Yang, H. Qiu, Y. Gong, X. Liu, Y. Lin, and M. Li, “Application of Computer Deep Learning Model in Diagnosis of Pulmonary Nodules,” *arXiv preprint arXiv:2406.13205*, 2024.
- [28] C. Jin, T. Che, H. Peng, Y. Li, and M. Pavone, “Learning from teaching regularization: Generalizable correlations should be easy to imitate,” *arXiv preprint arXiv:2402.02769*, 2024.
- [29] G. Urban *et al.*, “Deep Learning Localizes and Identifies Polyps in Real Time With 96% Accuracy in Screening Colonoscopy,” *Gastroenterology*, vol. 155, no. 4, pp. 1069–1078.e8, 2018, doi: 10.1053/j.gastro.2018.06.037.
- [30] J. Liu, T. Yang, and J. Neville, “CliqueParcel: An Approach For Batching LLM Prompts That Jointly Optimizes Efficiency And Faithfulness,” *arXiv preprint arXiv:2402.14833*, 2024.
- [31] X. Li, J. Chang, T. Li, W. Fan, Y. Ma, and H. Ni, “A Vehicle Classification Method Based on Machine Learning,” *Preprints (Basel)*, Jul. 2024, doi: 10.20944/preprints202407.0981.v1.
- [32] Y. Liu, H. Yang, and C. Wu, “Unveiling Patterns: A Study on Semi-Supervised Classification of Strip Surface Defects,” *IEEE Access*, vol. 11, pp. 119933–119946, 2023.
- [33] X. Liu, H. Qiu, M. Li, Z. Yu, Y. Yang, and Y. Yan, “Application of Multimodal Fusion Deep Learning Model in Disease Recognition,” *arXiv preprint arXiv:2406.18546*, 2024.
- [34] Z. Li, B. Wan, C. Mu, R. Zhao, S. Qiu, and C. Yan, “AD-Aligning: Emulating Human-like Generalization for Cognitive Domain Adaptation in Deep Learning,” *arXiv preprint arXiv:2405.09582*, 2024.
- [35] Y. Zhao and H. Gao, “Utilizing large language models for information extraction from real estate transactions,” *arXiv preprint arXiv:2404.18043*, 2024.
- [36] B. Dang, W. Zhao, Y. Li, D. Ma, Q. Yu, and E. Y. Zhu, “Real-Time Pill Identification for the Visually Impaired Using Deep Learning,” *arXiv preprint arXiv:2405.05983*, 2024.
- [37] J. Ye *et al.*, “Multiplexed oam beams classification via fourier optical convolutional neural network,” in *2023 IEEE Photonics Conference (IPC)*, 2023, pp. 1–2.
- [38] Y. Tao, “Meta Learning Enabled Adversarial Defense,” in *2023 IEEE International Conference on Sensors, Electronics and Computer Engineering (ICSECE)*, 2023, pp. 1326–1330. doi: 10.1109/ICSECE58870.2023.10263390.
- [39] Q. Xu *et al.*, “Applications of Explainable AI in Natural Language Processing,” *Global Academic Frontiers*, vol. 2, no. 3, pp. 51–64, 2024, doi: 10.5281/zenodo.12684705.
- [40] J. Wang, S. Hong, Y. Dong, Z. Li, and J. Hu, “Predicting Stock Market Trends Using LSTM Networks: Overcoming RNN Limitations for Improved Financial Forecasting,” *Journal of Computer Science and Software Applications*, vol. 4, no. 3, pp. 1–7, 2024, doi: 10.5281/zenodo.12200708.
- [41] J. Ye *et al.*, “Multiplexed orbital angular momentum beams demultiplexing using hybrid optical-electronic convolutional neural network,” *Commun Phys*, vol. 7, no. 1, p. 105, 2024.
- [42] W. Zhu, “Optimizing distributed networking with big data scheduling and cloud computing,” in *International Conference on Cloud Computing, Internet of Things, and Computer Applications (CICA 2022)*, 2022, pp. 23–28.
- [43] X. Liu and Z. Wang, “Deep Learning in Medical Image Classification from MRI-based Brain Tumor Images,” *arXiv preprint arXiv:2408.00636*, 2024.
- [44] J. Yuan, L. Wu, Y. Gong, Z. Yu, Z. Liu, and S. He, “Research on intelligent aided diagnosis system of medical image based on computer deep learning,” *arXiv preprint arXiv:2404.18419*, 2024.
- [45] M. Sun, Z. Feng, Z. Li, W. Gu, and X. Gu, “Enhancing Financial Risk Management through LSTM and Extreme Value Theory: A High-Frequency Trading Volume Approach,” *Journal of Computer Technology and Software*, vol. 3, no. 3, 2024, doi: 10.5281/zenodo.12669410.
- [46] Z. Wang, Y. Zhu, Z. Li, Z. Wang, H. Qin, and X. Liu, “Graph Neural Network Recommendation System for Football Formation,” *Applied Science and Biotechnology Journal for Advanced Research*, vol. 3, no. 3, pp. 33–39, 2024.

- [47] Z. Wang, H. Yan, Y. Wang, Z. Xu, Z. Wang, and Z. Wu, "Research on Autonomous Robots Navigation based on Reinforcement Learning," Jul. 2024, doi: 10.13140/RG.2.2.31231.62882.
- [48] Y. Yan, "Influencing Factors of Housing Price in New York-analysis: Based on Excel Multi-regression Model," 2022.
- [49] D. Ma, S. Li, B. Dang, H. Zang, and X. Dong, "Foste3net: A Lightweight YOLOv5 Based On the Network Structure Optimization," *arXiv preprint arXiv:2403.13703*, 2024.
- [50] J. Ye *et al.*, "OAM modes classification and demultiplexing via Fourier optical neural network," in *Complex Light and Optical Forces XVIII*, 2024, pp. 44–52.
- [51] K. Xu, Y. Wu, Z. Li, R. Zhang, and Z. Feng, "Investigating Financial Risk Behavior Prediction Using Deep Learning and Big Data, " *International Journal of Innovative Research in Engineering and Management*, vol. 11, no. 3, pp. 77–81, 2024, doi: 10.55524/ijirem.2024.11.3.12.
- [52] X. Li, Y. Yang, Y. Yuan, Y. Ma, Y. Huang, and H. Ni, "Intelligent Vehicle Classification System Based on Deep Learning and Multi-Sensor Fusion, " *Preprints (Basel)*, Jul. 2024, doi: 10.20944/preprints202407.2102.v2.
- [53] J. Ye, M. Solyanik, Z. Hu, H. Dalir, B. M. Nouri, and V. J. Sorger, "Free-space optical multiplexed orbital angular momentum beam identification system using Fourier optical convolutional layer based on 4f system," in *Complex Light and Optical Forces XVII*, 2023, pp. 69–79.
- [54] Y. Tao, "SQBA: sequential query-based blackbox attack," in *Fifth International Conference on Artificial Intelligence and Computer Science (AICS 2023)*, SPIE, 2023, p. 128032Q. doi: 10.1117/12.3009240.
- [55] C. Yan, J. Wang, Y. Zou, Y. Weng, Y. Zhao, and Z. Li, "Enhancing Credit Card Fraud Detection Through Adaptive Model Optimization," 2024, doi: 10.13140/RG.2.2.12274.52166.
- [56] Y. Zhong, Y. Liu, E. Gao, C. Wei, Z. Wang, and C. Yan, "Deep Learning Solutions for Pneumonia Detection: Performance Comparison of Custom and Transfer Learning Models," *medRxiv*, pp. 2024–2026, 2024.