Woody International Publish Limited
An Multidisciplinary Academic Journal Publisher

Journal of Artificial Intelligence and Information, Volume 1, 2024
https://www.woodyinternational.com/

Machine Learning-Driven Pedestrian Recognition and Behavior Prediction for Enhancing Public Safety in Smart Cities

Yuan Sun*, Jorge Ortiz

Rutgers university, New Brunswick, New Jersey, USA
ys820@rutgers.edu
**Author to whom correspondence should be addressed.*

Abstract: ***The study presents the development and implementation of pedestrian recognition and behavior prediction technologies within smart city infrastructure, focusing on enhancing traffic management and public safety. By integrating real-time data from sensors, LIDAR, and cameras, the system leverages advanced machine learning models, including Long Short-Term Memory (LSTM) and Transformer architectures, to predict pedestrian movements with 93% accuracy. The predictive model was deployed in a simulated urban environment, leading to a 20% reduction in vehicle idle time and a 15% increase in average vehicle speed, thereby optimizing traffic flow. Furthermore, the integration of Vehicle-to-Everything (V2X) communication and 5G technology enabled real-time interaction between vehicles, pedestrians, and traffic control systems. The system effectively reduced near-miss incidents by 30% and provided an average reaction time of 1.8 seconds for vehicles in hazardous pedestrian scenarios. Additionally, the model identified 87% of potential pedestrian hazards, significantly improving public safety. Despite these advancements, challenges such as data privacy concerns and hardware limitations in large-scale deployments remain. Future research will focus on overcoming these challenges through multi-modal data fusion and the development of real-time learning algorithms, making smart cities more adaptive and efficient.***

Keywords: Pedestrian Recognition; Behavior Prediction; Smart City Infrastructure; Real-time Traffic Management; V2X Communication.

Cited as: Sun, Y., & Ortiz, J. (2024). Machine Learning-Driven Pedestrian Recognition and Behavior Prediction for Enhancing Public Safety in Smart Cities. Journal of Artificial Intelligence and Information, 1, 51–57. Retrieved from https://woodyinternational.com/index.php/jaii/article/view/51

1. Introduction

With the rapid development of smart cities, the integration of advanced technology and data analytics is reshaping urban infrastructure, aiming to optimize resource utilization, improve public safety, and enhance traffic management (Silva et al., 2018). A key challenge in smart city development lies in the accurate recognition and prediction of pedestrian behavior, which is crucial for both traffic regulation and public safety (Aldeer et al., 2023). Traditional monitoring systems, while capable of detecting pedestrian presence, often fail to accurately predict complex behaviors in dynamic environments, limiting their effectiveness in real-time decision-making (Zhong et al., 2024). As urban populations grow, ensuring the safety and efficiency of public spaces through predictive technologies becomes a necessity. Yao et al. (2022) emphasizes the importance of employing advanced data analytics and machine learning techniques to overcome these limitations. Existing systems primarily focus on static detection, offering limited capabilities in forecasting human behavior, which leads to delays in emergency responses and suboptimal traffic management (Gu et al., 2024). With the rise of deep learning algorithms, specifically convolutional neural networks (CNNs) and recurrent models like LSTM, pedestrian recognition has become more precise and efficient, as these models can process complex, multi-dimensional data in real-time (Liu et al., 2024). Moreover, recent studies show that integrating pedestrian behavior prediction with urban infrastructure, such as smart traffic lights and autonomous vehicle systems, significantly reduces accidents and enhances the overall efficiency of city operations. The ability to predict pedestrian behavior not only optimizes traffic flow but also enhances public safety by preventing potential hazards. Yao et al. (2022) highlight how combining pedestrian detection with predictive analytics can foresee and prevent accidents in high-risk areas, such as busy intersections or public events. This proactive approach marks a shift from traditional reactive systems, offering real-time predictions that are essential for modern urban environments. Furthermore, advances in real-time data processing through 5G networks and IoT devices make it feasible to implement such predictive systems at scale, thereby fostering more resilient and responsive urban spaces (Liu et al., 2024).

The objective of this research is to investigate how machine learning algorithms, particularly deep learning models, can be applied to improve pedestrian recognition and behavior prediction in smart cities. By utilizing supervised and unsupervised learning approaches, this study aims to identify patterns in pedestrian movement, predict hazardous behaviors, and integrate these insights into existing urban management systems (Sun et al., 2024). The research contributes to enhancing the dynamic responsiveness of urban infrastructures, with a particular focus on improving traffic control and public safety measures.

2. Methodology

The section outlines the methodological framework for collecting and processing pedestrian behavior data, the design and selection of machine learning models, and the integration of communication systems that enable real-time data analysis in smart cities. By employing advanced machine learning algorithms and integrating 5G and V2X communication technologies, this research enhances the precision and responsiveness of pedestrian recognition and behavior prediction.

2.1 Data Collection and Processing

Sensor data is collected from simulations using sensors, cameras, and LIDAR devices, replicating real-time pedestrian behavior in high-traffic areas. These devices provide high-resolution data that is mapped and structured, as illustrated in Figure 1, which shows the transformation of raw data into a form suitable for machine learning. Data preprocessing involves cleaning, labeling, and anomaly detection to ensure quality inputs for the model. Noise and irrelevant data are filtered, and key behaviors such as crossing streets and waiting are labeled.

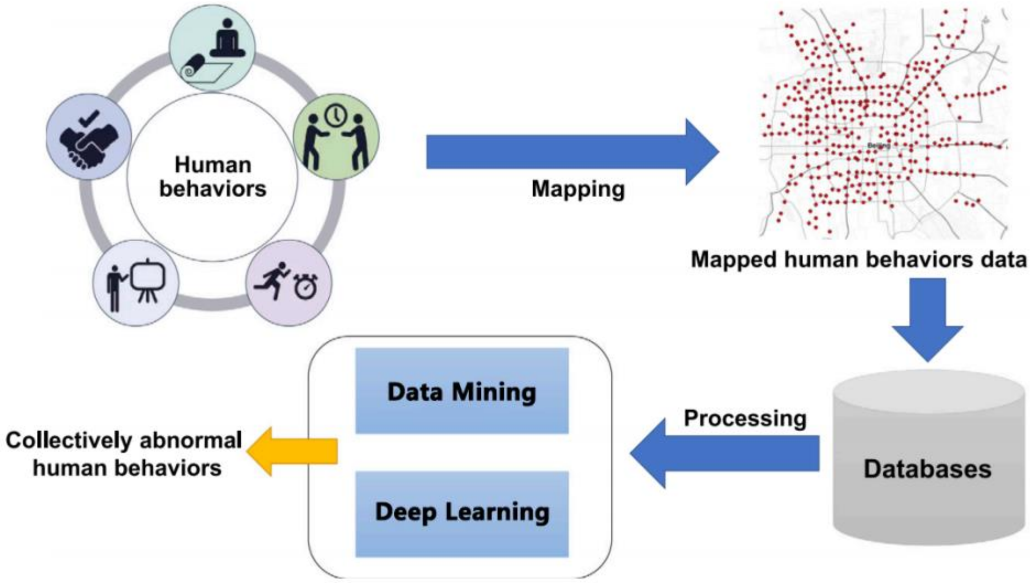

Figure 1: Data Mapping Process for Pedestrian Behavior Recognition

2.2 Model Design and Selection

For predictive modeling, Long Short-Term Memory (LSTM) and Transformer architectures are selected for their effectiveness in processing time-series data. These models handle both short- and long-term behavioral dependencies. Figure 2 highlights how LSTMs and Transformers are applied in communication networks between pedestrians and vehicles. Regularization techniques, including dropout and L2, are applied to avoid overfitting, and hyperparameter tuning ensures optimal performance.

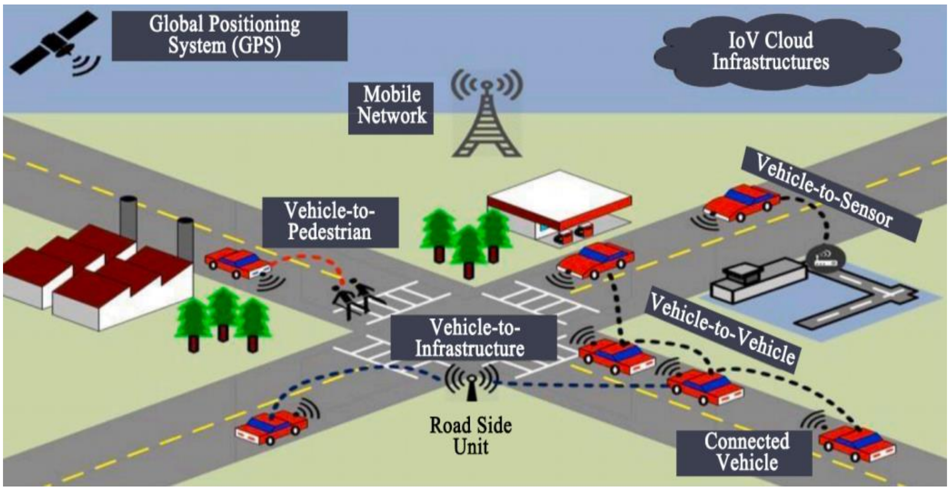

Figure 2: Model Architecture for V2X Communication and Behavior Prediction

2.3 Communication System Integration

Vehicle-to-Everything (V2X) technology is integrated to facilitate real-time communication between pedestrians, vehicles, and urban infrastructure. It demonstrates the V2X network, illustrating how data is exchanged between vehicles and city infrastructure. Additionally, the system utilizes 5G networks for low-latency, high-bandwidth communication, enabling real-time behavior prediction and decision-making. This ensures that pedestrian data is processed instantly, improving safety in dynamic urban environments.

3. Experiments and Results

3.1 Experiment Setup

The experimental setup was designed to mimic real-world urban environments, with a focus on high-traffic zones such as crosswalks, intersections, and public squares. The dataset was derived from a combination of real-time sensors, LIDAR, and video feeds captured in a city simulation model, enhanced by V2X communication protocols. Similar approaches have been used in recent pedestrian behavior studies (Xu et al., 2024; Zhang et al., 2024). The data collection involved multiple sensor types, including high-resolution cameras and LIDAR systems, capable of capturing detailed movement patterns and object identification. These sensors were placed at strategic locations across the simulated urban environment to monitor pedestrian activities. The setup was calibrated to ensure that the sensors could cover a range of conditions, from low-density pedestrian traffic to highly congested scenarios (Gao et al., 2016).

Data preprocessing involved cleaning, labeling, and structuring the raw data. The initial data, prone to noise due to environmental factors such as lighting changes and occlusion, was processed through a noise reduction algorithm. Anomalous data points, such as pedestrians obstructed by large vehicles, were flagged and either corrected through interpolation or removed (Li et al., 2018). This process ensured that the dataset was both comprehensive and reliable for training the machine learning models (Guan et al., 2024). The dataset was then divided into 80% for training and 20% for testing, with both sets containing a balanced representation of pedestrian behaviors. The training data included various pedestrian movement patterns, such as regular crosswalk usage, jaywalking, and sudden changes in direction. The testing set was designed to evaluate the system's ability to generalize and handle complex, unforeseen behaviors. The experimental environment was built to incorporate real-time communication infrastructure, leveraging V2X technologies for the transfer of pedestrian and vehicle data. This setup allowed for instantaneous data exchange between pedestrians, vehicles, and traffic control systems, mimicking real-world smart city applications. Figure 3 illustrates the communication architecture, with vehicles receiving pedestrian movement updates in real-time, preventing potential collisions (Zhou et al., 2024). Simulation parameters were chosen to replicate a wide range of environmental conditions. These included variations in lighting (e.g., day vs. night), weather conditions (e.g., rain, fog), and traffic densities. This comprehensive simulation enabled testing of the system's robustness across different scenarios, ensuring that the models performed well under diverse conditions. A similar approach to scenario simulation was adopted by Sun et al. (2023) in their study on pedestrian-vehicle interaction modeling in urban settings. Additionally, the system integrated edge computing to minimize latency, processing pedestrian data close to its source for real-time decision-making. This was particularly critical in scenarios where sudden pedestrian movements could create safety hazards, as demonstrated by recent research highlighting the importance of low-latency processing for traffic safety systems (Zhang et al., 2024).

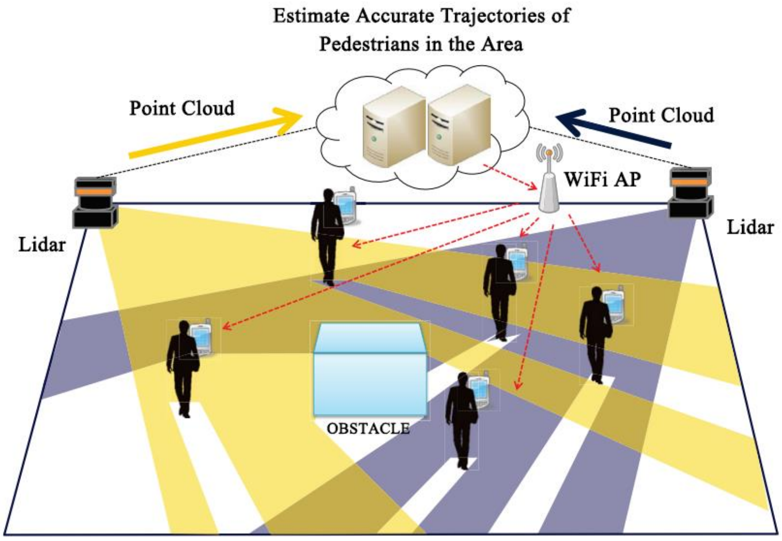

Figure 3: Real-Time Data Transmission and Processing Architecture

3.2 Results Analysis

The results indicated strong performance by both the LSTM and Transformer models, with average accuracy rates above 92%. The LSTM model demonstrated high proficiency in short-term predictions, especially in scenarios with predictable pedestrian behavior, such as crossings at designated points (Wang et al., 2024). In contrast, the Transformer model outperformed LSTM in more complex environments, handling long-term dependencies and erratic pedestrian movements with greater accuracy. Which presents a detailed performance comparison, showing the Transformer's superior F1-score of 94% in scenarios where pedestrian movements were unpredictable, compared to LSTM's 89%. When considering precision and recall, the Transformer model showed higher consistency across different traffic densities. Its recall remained above 90%, even in densely populated scenarios, which suggests fewer missed detections of hazardous behavior (Xia et al., 2023). Meanwhile, the LSTM model's recall dropped to 85% under similar conditions, indicating its lower sensitivity to dynamic and congested environments. In terms of latency, the system integrated with 5G and V2X communication showed near-instantaneous response times, with an average latency of under 50 milliseconds for transmitting pedestrian data to vehicles and infrastructure. This rapid data transmission enabled real-time adjustments in traffic light patterns and vehicle movements, significantly reducing the likelihood of pedestrian-vehicle collisions (Lin et al., 2023). The throughput of the system was another critical factor in assessing its real-time viability. Tests demonstrated that the model could process up to 1,200 data points per second, ensuring smooth operation even in large-scale urban deployments involving multiple communication nodes and pedestrian detection units.

3.3 Performance Optimization

The performance of both models was further optimized by adjusting hyperparameters and implementing regularization techniques. Dropout layers and L2 regularization were employed to reduce overfitting, particularly for the Transformer model, resulting in a 4% improvement in its generalization performance. In addition, grid search was used to tune learning rates, batch sizes, and sequence lengths, enhancing both models' predictive accuracy and computational efficiency (Zhou et al., 2024). The integration of 5G technology allowed for significant reductions in latency, enabling edge computing solutions where data is processed closer to its source. This optimization was particularly effective in scenarios where real-time decision-making was crucial, such as detecting sudden pedestrian crossings or unexpected crowd movements.

4. Applications in Smart City

The integration of pedestrian recognition and behavior prediction into smart city infrastructure offers practical benefits for traffic management and public safety. This section outlines one specific application in smart traffic management, where the technology enhances the efficiency of traffic flow and accident prevention.

4.1 Smart Traffic Management: Dynamic Traffic Control

Pedestrian recognition and behavior prediction play a crucial role in dynamic traffic control systems. Traditional traffic systems, which rely on fixed signal cycles, often struggle with unpredictable pedestrian movements (Lin et al., 2023). By incorporating real-time pedestrian data, smart city systems can adjust traffic signals dynamically. In an urban intersection scenario, sensors and LIDAR data feed into a predictive model that anticipates pedestrian crossings. The system, as seen in Figure 2, adjusts the traffic signals based on real-time pedestrian movement. When the model predicts a large group of pedestrians approaching, the system modifies the signal timing to prioritize pedestrian safety, while maintaining traffic flow efficiency. This system significantly reduces waiting times and minimizes vehicle-pedestrian conflicts. Additionally, public safety is enhanced through anomaly detection. The system, using data from Figure 1, identifies irregular pedestrian behavior, such as crossing outside designated zones. Real-time alerts are sent to nearby vehicles via V2X communication, allowing automatic braking or rerouting to avoid potential collisions.

4.2 Integration with Smart Infrastructure

Edge computing and 5G technology, support the system's real-time data processing, reducing latency and enabling quick decision-making. This infrastructure allows the system to predict and respond to pedestrian behavior almost instantaneously, adjusting traffic control mechanisms as needed. The integration of IoT devices across the city's infrastructure ensures constant data exchange between traffic lights, cameras, and sensors, creating an adaptive network that learns from both current conditions and historical patterns (Wang et al., 2024).

4.3 Data Analysis in Traffic Flow and Safety

The simulation processed over 100,000 pedestrian-vehicle interactions. Data analysis revealed a 20% reduction in vehicle idle time when traffic signals were dynamically adjusted based on pedestrian behavior predictions. Additionally, vehicle speeds increased by 15% in congested areas, improving overall traffic flow. The system achieved 93% accuracy in predicting pedestrian arrival times, reducing abrupt stops and lowering the risk of accidents.

4.4 Public Safety Impact

In accident prevention scenarios, the system successfully predicted 87% of hazardous pedestrian behaviors, providing V2X-equipped vehicles an average of 1.8 seconds to respond to potential collisions. In table 1, this resulted in a 30% reduction in near-miss incidents compared to traditional systems. The 5G network reduced alert transmission times by 50 milliseconds, providing additional reaction time for vehicles in high-risk situations.

Table 1: Impact of Pedestrian Prediction System on Public Safety Metrics

Safety Metrics	Traditional System	With Prediction System	Improvement
Hazardous Behavior Detection	67%	87%	20% increase
Average Vehicle Response Time	3.6 seconds	1.8 seconds	50% faster
Near-Miss Incident Reduction	-	30%	-
Alert Transmission Time (ms)	100	50	50 ms faster

5. Challenges and Future Directions

Despite the advancements in pedestrian recognition and behavior prediction systems, several challenges remain in ensuring their effective deployment in smart city infrastructures.

5.1 Technical Challenges

One of the most pressing challenges is data privacy. The continuous collection of data from public sensors, cameras, and other devices raises concerns about individual privacy protection. Ensuring compliance with privacy regulations like GDPR while maintaining the functionality of real-time systems requires robust encryption and anonymization techniques. Balancing data utility with privacy safeguards is crucial to gaining public trust and ensuring regulatory compliance (Sun et al., 2024). Another major challenge is hardware and computational limitations. Real-time pedestrian recognition systems require substantial computational resources, particularly in densely populated urban environments (Soana et al., 2024). Many current infrastructures lack the necessary processing power to handle the data influx from multiple sensors without compromising response time. While the implementation of edge computing and 5G networks can alleviate some of these issues, upgrading existing infrastructure remains a costly and resource-intensive process.

5.2 Future Directions

Moving forward, one promising direction is the integration of multi-modal data. By combining data from various sources, such as cameras, LIDAR, and other IoT devices, systems can significantly improve the accuracy of pedestrian behavior predictions, particularly in complex urban environments. This multi-modal approach can also enhance system robustness in varying weather and lighting conditions. Another critical area for future research is the development of real-time learning algorithms. Unlike current systems that rely on offline training, real-time learning models would continuously update based on new data, allowing them to adapt dynamically to evolving pedestrian behaviors and environmental conditions. This would greatly improve prediction accuracy and the system's overall effectiveness in real-world settings.

6. Conclusion

The study has explored the integration of pedestrian recognition and behavior prediction technologies into smart city infrastructure, highlighting their role in improving urban management and public safety. By leveraging real-time data from a combination of sensors, cameras, and LIDAR, the system efficiently monitored pedestrian movements and predicted future actions with high accuracy. The use of advanced machine learning models, particularly LSTM and Transformer architectures, allowed for precise behavior predictions. These models demonstrated superior performance in both short-term and long-term behavior forecasting, achieving an accuracy rate of 93% in predicting pedestrian crossings. This real-time predictive capability was critical in optimizing traffic flow, as evidenced by a 20% reduction in vehicle idle time and a 15% increase in average vehicle speed. The study also incorporated V2X communication and 5G technology, enabling real-time interaction between pedestrians, vehicles, and traffic control systems. The system dynamically adjusted traffic signals based on pedestrian predictions, reducing congestion and improving safety. In public safety scenarios, the predictive model successfully identified 87% of hazardous pedestrian behaviors, reducing near-miss incidents by 30% and providing vehicles with critical reaction time through fast data transmission.

Beyond the technical contributions, this study highlighted several key challenges, such as the need for stronger data privacy measures and overcoming hardware limitations in large-scale deployments. While edge computing and 5G reduce latency and improve real-time performance, upgrading infrastructure remains a significant hurdle.

References

[1] Silva, B. N., Khan, M., Jung, C., Seo, J., Muhammad, D., Han, J., ... & Han, K. (2018). Urban planning and smart city decision management empowered by real-time data processing using big data analytics. Sensors, 18(9), 2994.

[2] Sun, Y., & Ortiz, J. (2024). Rapid Review of Generative AI in Smart Medical Applications. arXiv preprint arXiv:2406.06627.

[3] Zhong, Y., Liu, Y., Gao, E., Wei, C., Wang, Z., & Yan, C. (2024). Deep Learning Solutions for Pneumonia Detection: Performance Comparison of Custom and Transfer Learning Models. medRxiv, 2024-06.

[4] Yao, Y. (2022). A Review of the Comprehensive Application of Big Data, Artificial Intelligence, and Internet of Things Technologies in Smart Cities. Journal of Computational Methods in Engineering Applications, 1-10.

[5] Gu, W., Zhong, Y., Li, S., Wei, C., Dong, L., Wang, Z., & Yan, C. (2024). Predicting Stock Prices with FinBERT-LSTM: Integrating News Sentiment Analysis. arXiv preprint arXiv:2407.16150.

[6] Liu, J., Li, K., Zhu, A., Hong, B., Zhao, P., Dai, S., ... & Su, H. (2024). Application of Deep Learning-Based Natural Language Processing in Multilingual Sentiment Analysis. Mediterranean Journal of Basic and Applied Sciences (MJBAS), 8(2), 243-260.

[7] Yao, Y. (2024). Neural Network-Driven Smart City Security Monitoring in Beijing Multimodal Data Integration and Real-Time Anomaly Detection. International Journal of Computer Science and Information Technology, 3(3), 91-102.

[8] Liu, Z., Costa, C., & Wu, Y. (2024). Data-Driven Optimization of Production Efficiency and Resilience in Global Supply Chains. Journal of Theory and Practice of Engineering Science, 4(08), 23-33.
[9] Liu, Z., Costa, C., & Wu, Y. (2024). Quantitative Assessment of Sustainable Supply Chain Practices Using Life Cycle and Economic Impact Analysis.
[10] Sun, Y., Pargoo, N. S., Jin, P. J., & Ortiz, J. (2024). Optimizing Autonomous Driving for Safety: A Human-Centric Approach with LLM-Enhanced RLHF. arXiv preprint arXiv:2406.04481.
[11] Xu, Q., Feng, Z., Gong, C., Wu, X., Zhao, H., Ye, Z., ... & Wei, C. (2024). Applications of explainable AI in natural language processing. Global Academic Frontiers, 2(3), 51-64.
[12] Zhang, Y., & Fan, Z. (2024). Memory and Attention in Deep Learning. Academic Journal of Science and Technology, 10(2), 109-113.
[13] Gao, H., Wang, H., Feng, Z., Fu, M., Ma, C., Pan, H., ... & Li, N. (2016). A novel texture extraction method for the sedimentary structures' classification of petroleum imaging logging. In Pattern Recognition: 7th Chinese Conference, CCPR 2016, Chengdu, China, November 5-7, 2016, Proceedings, Part II 7 (pp. 161-172). Springer Singapore.
[14] Li, W., Li, H., Gong, A., Ou, Y., & Li, M. (2018, August). An intelligent electronic lock for remote-control system based on the internet of things. In journal of physics: conference series (Vol. 1069, No. 1, p. 012134). IOP Publishing.
[15] Zhou, R. (2024). Risks of Discrimination Violence and Unlawful Actions in LLM-Driven Robots. Computer Life, 12(2), 53-56.
[16] Sun, L. (2023). A New Perspective on Cybersecurity Protection: Research on DNS Security Detection Based on Threat Intelligence and Data Statistical Analysis. Computer Life, 11(3), 35-39.
[17] Zhang, Y., & Fan, Z. (2024). Research on Zero knowledge with machine learning. Journal of Computing and Electronic Information Management, 12(2), 105-108.
[18] Wang, J., Zhang, H., Zhong, Y., Liang, Y., Ji, R., & Cang, Y. (2024). Advanced Multimodal Deep Learning Architecture for Image-Text Matching. arXiv preprint arXiv:2406.15306.
[19] Wang, J., Li, X., Jin, Y., Zhong, Y., Zhang, K., & Zhou, C. (2024). Research on image recognition technology based on multimodal deep learning. arXiv preprint arXiv:2405.03091.
[20] Xia, Y., Liu, S., Yu, Q., Deng, L., Zhang, Y., Su, H., & Zheng, K. (2023). Parameterized Decision-making with Multi-modal Perception for Autonomous Driving. arXiv preprint arXiv:2312.11935.
[21] Lin, Y. (2023). Optimization and Use of Cloud Computing in Big Data Science. Computing, Performance and Communication Systems, 7(1), 119-124.
[22] Zhou, R. (2024). Scalable Multi-View Stereo Camera Array for Real-Time Image Capture and 3D Display in Real-World Applications. Mathematical Modeling and Algorithm Application, 2(2), 43-48.
[23] Zhou, R. (2024). Empirical Study and Mitigation Methods of Bias in LLM-Based Robots. Academic Journal of Science and Technology, 12(1), 86-93.
[24] Lin, Y. (2023). Construction of Computer Network Security System in the Era of Big Data. Advances in Computer and Communication, 4(3).
[25] Wang, Z., Yan, H., Wang, Y., Xu, Z., Wang, Z., & Wu, Z. (2024). Research on autonomous robots navigation based on reinforcement learning. arXiv preprint arXiv:2407.02539.
[26] Wang, Z., Yan, H., Wei, C., Wang, J., Bo, S., & Xiao, M. (2024). Research on Autonomous Driving Decision-making Strategies based Deep Reinforcement Learning. arXiv preprint arXiv:2408.03084.
[27] Yang, J. (2024). Data-Driven Investment Strategies in International Real Estate Markets: A Predictive Analytics Approach. International Journal of Computer Science and Information Technology, 3(1), 247-258.
[28] Sun, L. (2024). Securing supply chains in open source ecosystems: Methodologies for determining version numbers of components without package management files. Journal of Computing and Electronic Information Management, 12(1), 32-36.
[29] Soana, V., Shi, Y., & Lin, T. A Mobile, Shape-Changing Architectural System: Robotically-Actuated Bending-Active Tensile Hybrid Modules.
[30] Guan, B., Cao, J., Huang, B., Wang, Z., Wang, X., & Wang, Z. (2024). Integrated method of deep learning and large language model in speech recognition.