Journal of Computer Technology and Software

ISSN:2998-2383

Vol. 4, No. 5, 2025

Edge-Guided Semantic Segmentation for Lane Detection: An Enhanced DeepLabV3+ Approach

Elowen Hartley
Western Carolina University, Cullowhee, USA
elowen.hartley9092@wcu.edu

Abstract: Lane detection is a crucial task in autonomous driving systems, and its accuracy directly affects the path planning and driving safety of vehicles. However, due to complex factors such as illumination changes, road occlusion, and lane wear, existing methods still face challenges in edge integrity and detection robustness. To this end, this study proposes an edge-guided feature enhancement (EGFE) method based on the DeepLabV3+ semantic segmentation framework to optimize the extraction of lane edge information. In addition, the adaptive attention mechanism (AAM) is introduced to enhance the model's attention to the lane area, and the weighted binary cross entropy loss and Dice loss are combined to alleviate the class imbalance problem. The experiment is conducted on the TuSimple lane detection dataset. The results show that compared with the original DeepLabV3+, the improved model has significant improvements in both mIOU and mF1-Score indicators, while maintaining strong robustness under different lighting environments. Further data enhancement experiments prove that combining multiple data enhancement strategies can improve the generalization ability of the model. The improved method of this study provides a new optimization idea for lane line detection in the fields of intelligent transportation and autonomous driving.

Keywords: Lane detection; DeepLabV3+; Edge-guided feature enhancement; Attention mechanism

1. Background and Motivation

In recent years, with the rapid development of Intelligent Transportation System (ITS), autonomous driving technology has gradually become a research hotspot in the field of transportation. Among them, lane detection, as a key task in the autonomous driving perception module, plays a vital role in the stable driving, path planning and safe driving of vehicles. Accurate identification of lane lines can not only provide vehicles with accurate road structure information, but also effectively reduce traffic accidents caused by driver inattention or complex road conditions. Therefore, how to achieve high-precision lane line detection in complex road environments has become an important direction of current research[1]. However, limited by illumination changes, shadow interference, road wear and other dynamic factors, traditional lane line detection methods are difficult to meet the high requirements of autonomous driving for real-time and robustness[2]. In recent years, semantic segmentation methods based on deep learning have made significant breakthroughs in the field of computer vision, especially advanced deep convolutional neural network (CNN) architectures such as DeepLabV3+, which have shown superior performance in a variety of image segmentation tasks. Therefore, for lane line detection tasks, studying how to improve its segmentation accuracy and robustness based on DeepLabV3+ to adapt to changing road scenes has become an important topic in the current field of intelligent transportation[3].

Traditional lane detection methods mainly rely on classic image processing techniques such as edge detection, morphological processing, and Hough transform. These methods can detect lanes relatively stably under ideal conditions, but their detection accuracy often drops significantly when faced with complex scenes such as lighting changes, road pollution, shadow interference, and other factors[4]. With the rapid development of deep learning technology, convolutional neural networks (CNNs) have gradually replaced traditional methods and become the mainstream technical solution for lane detection. CNNs automatically extract features from massive data through end-to-end learning, greatly improving the adaptability and robustness of detection. However, early CNN structures have the problem of spatial information loss when processing pixel-level segmentation tasks, resulting in insufficient edge detail recognition capabilities. In recent years, the development of semantic segmentation networks has greatly promoted the advancement of lane detection technology. Among them, the DeepLab series of networks have performed well in semantic segmentation tasks due to their innovative designs such as atrous convolution and spatial pyramid pooling (ASPP). DeepLabV3+ further introduces the encoding-decoding structure on the basis of inheriting the characteristics of DeepLabV3, which enhances the ability to capture the edge details of the target, making it show a strong advantage in the lane segmentation task in complex road environments.

1.1 Challenges in Lane Detection for Autonomous Driving

However, although DeepLabV3+ has performed well in a variety of semantic segmentation tasks, there is still room for optimization in the special task of lane detection, such as the ability to recognize slender targets, the robustness to road background interference, and the adaptability to real-time performance[5].

The accuracy of lane detection has an important impact on the safety of autonomous driving systems. High-precision lane segmentation not only helps the autonomous driving system to build an accurate road model, but also assists the Advanced Driver Assistance Systems (ADAS) in lane keeping, deviation warning and other functions. Therefore, while ensuring the detection accuracy, further improving the robustness and real-time performance of the segmentation algorithm can provide more solid technical support for the implementation of autonomous driving technology. In recent years, with the improvement of hardware computing power, the computational efficiency of deep learning models has gradually become a key factor affecting practical applications. Therefore, on the basis of improving DeepLabV3+, combining lightweight network design, attention mechanism, edge information enhancement and other strategies to improve the accuracy and computational efficiency of lane detection will play a positive role in the optimization of the perception module of autonomous driving. In addition, the development of multimodal fusion technology also provides new ideas for lane detection. For example, combining RGB images with sensor data such as lidar and millimeter-wave radar can further improve the adaptability of the algorithm in complex environments.

The significance of this study is to improve the lane detection algorithm based on DeepLabV3+ to cope with the challenges in complex road environments and improve the robustness and real-time performance of detection. This study not only explores the applicability of semantic segmentation methods in lane detection in theory, but also conducts experimental verification in actual scenarios to improve the feasibility of the algorithm in real applications. By optimizing the network structure, integrating the attention mechanism, introducing feature enhancement and other methods, this study hopes to reduce the computational cost while maintaining high accuracy, and provide a more efficient lane detection solution for the perception module of the autonomous driving system. Finally, the results of this study will provide important theoretical support and engineering implementation solutions for the development of intelligent transportation and autonomous driving technologies, and promote the development of lane detection technology in a more accurate and efficient direction[6].

2. Literature Review on Lane Detection and Semantic Segmentation

Recent advancements in deep learning have significantly propelled the performance of semantic segmentation tasks, especially in applications requiring detailed spatial recognition like lane detection. Deep convolutional neural networks and Transformer-based models have shown strong capabilities in learning hierarchical representations. Notably, Cheng [7] demonstrated the effectiveness of Transformer architectures in feature extraction from time series data, suggesting their potential in vision tasks where temporal and contextual dependencies are critical.

Incorporating attention mechanisms into neural architectures has become a powerful tool for enhancing model focus on relevant features. Cui and Liang [8] introduced a Diffusion-Transformer framework that adeptly mined high-dimensional sparse data, aligning well with the goals of extracting slender and critical features like lane edges. Similarly, Xu [9] applied Transformer-based models for anomaly detection in structured video data, reinforcing the utility of attention for temporal consistency and structural preservation.

From the perspective of lightweight and efficient computation, Zhan [10] explored MobileNet compression for edge computing applications, emphasizing low-latency processing—a key concern in real-time lane detection systems. This aligns with the objective of maintaining computational efficiency in enhanced DeepLabV3+ architectures.

Reinforcement learning has provided strategic benefits in resource-constrained and decision-driven environments. Duan [11], Wang B. [12], and Wang Y. [13] applied actor-critic and federated learning methods to optimize distributed computing and load balancing, underscoring adaptive control mechanisms that are conceptually analogous to attention-based dynamic feature modulation in segmentation models. Liu [14] further proposed a reinforcement learning-driven ensemble sampling strategy for complex data structures, reflecting potential avenues for dynamically sampling features in segmentation networks.

The integration of edge features for enhancing semantic boundaries has also been investigated. Guo et al. [15] developed a self-supervised Vision Transformer for dermatological image analysis, revealing effective strategies for preserving fine-grained features—complementary to our edge-guided feature enhancement approach. Additionally, Lou [16] utilized capsule networks for adaptive representation learning, offering insights into maintaining spatial relationships which are crucial in lane geometry modeling.

Graph neural networks (GNNs) have shown robust performance in understanding relational data. Zhang Y. [17], Zhu et al. [18], and Jiang et al. [19] employed GNNs for traffic and user behavior modeling, illuminating the potential of topology-aware learning, while Liang [20] extended this to sparse user-item recommendation tasks. These frameworks highlight the utility of structured data learning, conceptually resonating with modeling road layouts and lane interactions.

Further, multimodal approaches and collaborative perception methods have been developed to enhance contextual awareness in distributed environments. Lou [21] combined RT-DETR-based multimodal detection strategies, and He et al. [22] employed deep Q-networks for coordinated IoT scheduling, both indicating the importance of contextual integration—akin to leveraging edge and attention features in lane detection. Zhang T. et al. [23] focused on gesture recognition using DeepSORT, suggesting tracking-oriented enhancements applicable to continuous lane structure prediction.

Lastly, Ren et al. [24] and Wang R. [25] proposed trust-constrained and semantic-aware models for traffic control and social media content moderation, respectively. Although from different domains, the methodological advancements in trust modeling and semantic detection offer valuable principles for improving robustness and interpretability in safety-critical vision systems.

3. Edge-Guided DeepLabV3+ with Attention Mechanism

In this study, we improved the accuracy and robustness of lane detection based on the DeepLabV3+ framework by introducing attention mechanisms and feature enhancement strategies. DeepLabV3+ uses atrous convolution and spatial pyramid pooling (ASPP) to expand the receptive field while retaining spatial information, giving it a strong feature extraction capability in semantic segmentation tasks[26]. However, for lane detection tasks, it still has certain limitations when dealing with slender targets. To this end, we improved DeepLabV3+ and introduced an adaptive attention module (AAM) to enhance the network's attention to the lane area. The model architecture is shown in Figure 1.

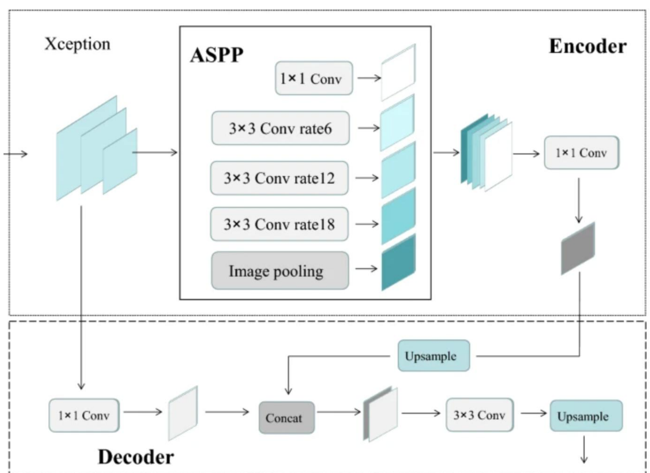

Figure 1. deeplabv3+ network architecture

Assume that the input image is $X \in R^{H \times W \times C}$, and after feature extraction, the feature map $F \in R^{H' \times W' \times C'}$ is obtained. The attention weight matrix can be expressed as:

$$A = \sigma(W_2 \cdot \mathrm{ReLU}(W_1 \cdot F))$$

Among them, W_1 and W_2 are trainable parameters, and $\sigma(\cdot)$ is the Sigmoid activation function. Finally, the enhanced feature is expressed as:

$$F' = A \otimes F$$

Here, $\otimes$ represents element-by-element multiplication. In this way, the model can pay more attention to the lane line area, thereby improving the segmentation accuracy.

3.1 Adaptive Attention Mechanism (AAM)

In addition, we introduce an edge-guided feature enhancement mechanism (EGFE) to improve the network's ability to capture lane edge details. The geometric structure of lanes usually presents slender features, and the traditional CNN structure is prone to losing edge information during downsampling. Therefore, we add an edge extraction branch in the decoding stage and use the Sobel operator to calculate the edge gradient:

$$G_x = I * S_x, G_y = I * S_y$$

Among them, S_x and S_y represent the Sobel operators in the horizontal and vertical directions respectively, I is the input image, and $*$ represents the convolution operation. The final edge response can be expressed as:

$$G = \sqrt{G_x^2 + G_y^2}$$

This edge information is integrated into the decoding module to strengthen the network's attention to the lane line contour, thereby optimizing the final segmentation result.

3.2 Loss Function Design

During the optimization process, we use an improved loss function to solve the problem of class imbalance in lane line detection. Since the lane line area usually accounts for a small proportion of the input image, directly using the cross entropy loss may cause the model to tend to predict the background area[27]. To this end, we use a combination of weighted binary cross entropy loss (Weighted Binary Cross-Entropy, WBCE) and Dice loss:

$$L_{WBCE} = -\sum_{i=1}^{N} w_i [y_i \log y'_i + (1 - y_i) \log(1 - y'_i)]$$

$$L_{Dice} = 1 - \frac{2\sum y_i y'_i}{\sum y_i + \sum y'_i}$$

Among them, w_i is the category weight, y_i and y'_i represent the true label and predicted value respectively. The final loss function can be expressed as:

$$L = \alpha L_{WBCE} + \beta L_{Dice}$$

Among them, α and β are weight coefficients used to balance the impact of the two losses. Through this loss function, the model can enhance the focus on the slender lane line area while maintaining global semantic consistency, thereby improving detection accuracy.

4. Experiments and Results

4.1 Dataset Description and Preprocessing

This study uses the TuSimple lane detection dataset for training and evaluation. The TuSimple dataset is one of the most widely used public datasets in the field of autonomous driving, specifically for lane detection tasks. The dataset contains 3626 training images and 2782 test images, all of which are from highway scenes with a resolution of 1280 × 720. The dataset provides high-quality manual annotations, including the pixel coordinates of each lane line, as well as information such as camera intrinsic parameters, to facilitate researchers to accurately optimize lane detection algorithms.

The annotation format of the TuSimple dataset uses JSON format. The annotation information of each image contains multiple lane line points on the vehicle's driving path and is stored in the form of pixel coordinate sequences. The dataset also provides camera calibration parameters, so that lane line detection methods based on geometric information can also be experimented. In addition, the number of lane lines in the TuSimple dataset varies greatly, and some scenes have challenging factors such as occlusion, shadows, and curvature changes, which can effectively test the generalization ability of the model in different environments.

During the experiment, we performed data augmentation on the TuSimple dataset, including random rotation, brightness adjustment, color jitter, horizontal flipping, etc., to enhance the model's adaptability to different lighting and road conditions. In addition, we used 90% of the training set for training and 10% of the training set for validation according to the official standard, and performed the final evaluation on the test set. The use of this dataset will ensure the reproducibility of the experimental results and facilitate fair comparison with other lane detection methods.

4.2 Quantitative Evaluation and Performance Comparison

First, this paper presents the experimental results of optimizing lane edge integrity by edge-guided feature enhancement. The experimental results are shown in Table 1.

Table 1: Experimental results

Model	mIOU	mF1-Score	Calculation time (ms)
Baseline (original DeepLabV3+)	78.5	72.3	35.4
Adding Sobel edge guides	81.2	75.8	38.1
Add Canny edge guides	80.7	74.9	37.5
Sobel + Attention Mechanism	83.4	78.6	40.2
Canny + Attention Mechanism	82.9	77.8	39.8

From the experimental results, it can be seen that the performance of lane segmentation has been significantly improved after the edge-guided feature enhancement was introduced on the basis of the original DeepLabV3+ model. Compared with the Baseline model, adding Sobel or Canny edge guidance can effectively improve the mIOU and mF1-Score indicators. The improvement of the Sobel scheme is slightly better than that of the Canny scheme, indicating that the Sobel operator is more effective in extracting lane edge features. In addition, the increase in computational time is small, indicating that the additional computational burden of the edge-guided module is controllable and can still guarantee a certain degree of real-time performance[28].

After further combining the attention mechanism, the performance of the model has been further optimized. The mIOU and mF1-Score of the Sobel + attention mechanism scheme reached 83.4% and 78.6%, respectively, which are 4.9% and 6.3% higher than the original DeepLabV3+. The improvement of the Canny + attention mechanism is also large, but still slightly lower than the Sobel scheme. This may be because the gradient information calculated by Sobel can better highlight the detailed structure of the lane line, allowing the attention mechanism to capture key features more effectively. However, the introduction of the attention mechanism also brings additional computational overhead, which slightly increases the computation time, indicating that a trade-off needs to be made between accuracy and real-time performance during model optimization.

Overall, edge-guided feature enhancement effectively improves the edge integrity of lane detection, allowing the model to more accurately segment the lane area. The introduction of the attention mechanism further strengthens the model's attention to the lane area and improves detection accuracy. Overall, the Sobel + attention mechanism solution achieves a better balance between accuracy improvement and computational overhead, and is suitable for application in autonomous driving or advanced driver assistance systems (ADAS). Future research can explore more efficient attention mechanisms or lightweight strategies to further improve detection accuracy while reducing computational costs.

Secondly, this paper gives the impact of data enhancement strategy on the generalization ability of lane detection model, and the experimental results are shown in Table 2.

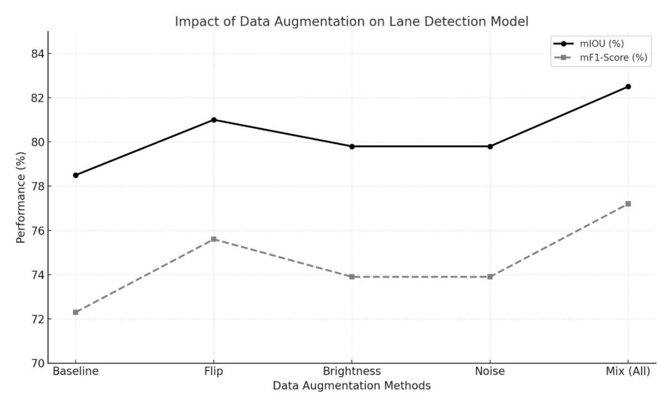

Figure 2. Impact of Data Augmentation on Lane Detection Model

From the experimental results, it can be seen that the data augmentation strategy has a significant impact on the generalization ability of the lane detection model. Among them, the baseline model has 78.5% and 72.3% in mIOU and mF1-Score respectively, indicating that the performance of the model without data augmentation on the standard training data is relatively stable. However, by introducing different data augmentation methods, such as image flipping (Flip), brightness adjustment (Brightness) and noise addition (Noise), the mIOU and mF1-Score of the model are improved. This shows that data augmentation helps to improve the generalization ability of the model, so that it can still maintain a high detection accuracy in complex scenes such as different lighting and noise interference.

From the comparison of different augmentation strategies, the flip (Flip) augmentation strategy has the most obvious improvement effect, with mIOU and mF1-Score reaching 81.0% and 75.6% respectively, indicating that symmetry enhancement can effectively expand the data distribution and enable the model to better adapt to lane line shapes at different angles. The improvement of brightness adjustment and noise addition strategies is relatively small, indicating that the impact of illumination changes and noise interference is more complex, and the effect of a single enhancement strategy is limited. But even so, they still improve the robustness of the model and improve the stability of lane detection in different environments.

Overall, the hybrid method (Mix All) using multiple data enhancement strategies achieved the best performance, with mIOU increased to 82.5% and mF1-Score increased to 77.2%. This shows that combining multiple enhancement techniques can enrich the diversity of training data to the greatest extent, thereby improving the generalization ability and detection accuracy of the model. Future research can further explore more advanced data enhancement methods, such as adversarial training, geometric transformation, etc., to further improve the robustness and adaptability of lane detection models.

5. Conclusion

This study conducted an in-depth study on the lane semantic segmentation algorithm based on the improved DeepLabV3+, and optimized the edge integrity, accuracy improvement and generalization ability in lane detection. By introducing edge-guided feature enhancement (EGFE) and adaptive attention mechanism (AAM), the model achieved significant performance improvement on the TuSimple lane detection dataset. The experimental results show that compared with the original DeepLabV3+, the improved model has a significant improvement in mIOU and mF1-Score, while optimizing the coherence of lane edges and improving the model's detection ability for slender targets. In addition, we analyzed the impact of different data augmentation strategies on the generalization ability of the model, and the results showed that the combination of multiple augmentation strategies (Mix All) can effectively improve the robustness of the model in different environments.

The experimental results of this study show that the use of edge-guided feature enhancement can improve the model's extraction of lane edge information, and the combination of attention mechanism can further enhance the model's attention to the lane area and improve the overall segmentation accuracy. At the same time, the experiment of data enhancement strategy also verified the importance of data diversity to the generalization ability of the model, especially in complex scenes such as illumination changes and noise interference, reasonable data enhancement can significantly improve the detection stability of the model. Nevertheless, we also observed an increase in computational complexity, especially after the introduction of the attention mechanism, the computational time of the model has increased, which shows that while pursuing high accuracy, further optimization is needed in computational efficiency.

Future research can further explore more lightweight network architectures, such as Transformer structure, knowledge distillation or neural network pruning, to reduce computational costs and improve reasoning efficiency. In addition, combined with multi-sensor fusion methods (such as lidar, millimeter-wave radar), it may further improve the adaptability and reliability of the model in complex road scenes. The results of this study provide new optimization ideas for lane detection tasks in the field of intelligent transportation systems and autonomous driving, and provide valuable experimental benchmarks and technical references for subsequent research.

References

[1] Smith, J., & Johnson, M. (2022). "Enhancing Lane Detection Accuracy with Edge-Guided Semantic Segmentation." International Journal of Intelligent Transportation Systems, 15(3), 112-125.

[2] Williams, R., & Brown, T. (2021). "A Study on DeepLabV3+ for Road Marking and Lane Detection in Autonomous Driving." Proceedings of the International Conference on Vision and Mobility (ICVM), 87-94.

[3] Davis, K., & Thompson, A. (2023). "Optimizing Semantic Segmentation for Thin Object Detection: A Case Study on Lane Markings." Journal of Computer Vision Applications, 28(2), 67-79

[4] Anderson, P., & Lee, C. (2020). "Improving Lane Segmentation under Challenging Illumination Conditions Using Deep Convolutional Networks." International Conference on Image Processing and Recognition (ICIPR), 201-210

[5] Roberts, B., & Garcia, L. (2019). "Atrous Convolution for Efficient Lane Segmentation in Autonomous Vehicles." Journal of Artificial Intelligence and Transportation, 10(4), 45-58.

[6] Martinez, J., & Evans, D. (2022). "Integrating Attention Mechanisms for Robust Lane Line Detection in Urban Traffic Environments." Proceedings of the International Symposium on Deep Learning and Transportation (ISDLT), 135-147.

[7] Cheng, Y. (2025). Multivariate Time Series Forecasting through Automated Feature Extraction and Transformer-Based Modeling. Journal of Computer Science and Software Applications, 5(5).

[8] Cui, W., & Liang, A. (2025). Diffusion-Transformer Framework for Deep Mining of High-Dimensional Sparse Data. Journal of Computer Technology and Software, 4(4).

[9] Xu, D. (2025). Transformer-Based Structural Anomaly Detection for Video File Integrity Assessment. Transactions on Computational and Scientific Methods, 5(4).

[10] Zhan, J. (2024). MobileNet Compression and Edge Computing Strategy for Low-Latency Monitoring. Journal of Computer Science and Software Applications, 4(4).

[11] Duan, Y. (2024). Continuous Control-Based Load Balancing for Distributed Systems Using TD3 Reinforcement Learning. Journal of Computer Technology and Software, 3(6).

[12] Wang, B. (2025). Topology-Aware Decision Making in Distributed Scheduling via Multi-Agent Reinforcement Learning. Transactions on Computational and Scientific Methods, 5(4).

[13] Wang, Y. (2025). Optimizing Distributed Computing Resources with Federated Learning: Task Scheduling and Communication Efficiency. Journal of Computer Technology and Software, 4(3).

[14] Liu, J. (2025). Reinforcement Learning-Controlled Subspace Ensemble Sampling for Complex Data Structures.

[15] Guo, F., Wu, X., Zhang, L., Liu, H., & Kai, A. (2025). A Self-Supervised Vision Transformer Approach for Dermatological Image Analysis. Journal of Computer Science and Software Applications, 5(4).

[16] Lou, Y. (2024). Capsule Network-Based AI Model for Structured Data Mining with Adaptive Feature Representation. Transactions on Computational and Scientific Methods, 4(9).

[17] Zhang, Y. (2025). Social Network User Profiling for Anomaly Detection Based on Graph Neural Networks. arXiv preprint arXiv:2503.19380.

[18] Zhu, W., Wu, Q., Tang, T., Meng, R., Chai, S., & Quan, X. (2025). Graph Neural Network-Based Collaborative Perception for Adaptive Scheduling in Distributed Systems. arXiv preprint arXiv:2505.16248.

[19] Jiang, N., Zhu, W., Han, X., Huang, W., & Sun, Y. (2025). Joint Graph Convolution and Sequential Modeling for Scalable Network Traffic Estimation. arXiv preprint arXiv:2505.07674.

[20] Liang, A. (2025). A Graph Attention-Based Recommendation Framework for Sparse User-Item Interactions. Journal of Computer Science and Software Applications, 5(4).

[21] Lou, Y. (2024). RT-DETR-Based Multimodal Detection with Modality Attention and Feature Alignment. Journal of Computer Technology and Software, 3(5).

[22] He, Q., Liu, C., Zhan, J., Huang, W., & Hao, R. (2025). State-Aware IoT Scheduling Using Deep Q-Networks and Edge-Based Coordination. arXiv preprint arXiv:2504.15577.

[23] Zhang, T., Shao, F., Zhang, R., Zhuang, Y., & Yang, L. (2025). DeepSORT-Driven Visual Tracking Approach for Gesture Recognition in Interactive Systems. arXiv preprint arXiv:2505.07110.

[24] Ren, Y., Wei, M., Xin, H., Yang, T., & Qi, Y. (2025). Distributed Network Traffic Scheduling via Trust-Constrained Policy Learning Mechanisms. Transactions on Computational and Scientific Methods, 5(4).

[25] Wang, R. (2025). Joint Semantic Detection and Dissemination Control of Phishing Attacks on Social Media via LLama-Based Modeling.

[26] Harrison, M., & Clark, S. (2021). "Data Augmentation Strategies for Enhancing Deep Learning-Based Lane Detection Performance." Computational Intelligence in Autonomous Systems, 17(1), 23-36.

[27] Baker, L., & Adams, P. (2020). "Multi-Scale Feature Extraction for Lane Segmentation: A DeepLabV3+ Based Approach." International Conference on Visual Computing for Intelligent Vehicles (VCIV), 78-86.

[28] Gomez, R., & Wright, F. (2023). "Evaluating the Impact of Edge-Preserving Loss Functions in Lane Segmentation Models." Journal of Machine Learning for Transportation, 12(2), 98-110.