BiFPN-YOLOv8: A High-Performance Deep Learning Model for Traffic Light Recognition

Arman Naghizadeh
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
armannaghizadeh@tabrizu.ac.ir

Amir A. Ghavifekr
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
aa.ghavifekr@tabrizu.ac.ir

Ali Abdolhafez Hadi Al Mahdawi
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
almahdawi@tabrizu.ac.ir

Aws Mohammed Hameed Al Khazraji
Faculty of Electrical and Computer Engineering
University of Tabriz
Tabriz, Iran
alkhazraji@tabrizu.ac.ir

Ali Alfayly
PAAET Colledge of Basic Education
Kuwait, Kuwait
ah.alfayly@paaet.edu.kw

Abstract— **Traffic light recognition is a critical component of autonomous driving and intelligent transportation systems, requiring accurate and efficient detection in real-time scenarios. In this study, BiFPN-YOLOv8 is used as a high-performance deep learning model for traffic light detection, which enhances YOLOv8n by incorporating the Bidirectional Feature Pyramid Network (BiFPN) and an additional detection head. While YOLOv8n is known for its speed and accuracy, its ability to detect small and distant objects is limited. By integrating BiFPN, the model improves multi-scale feature fusion, leading to enhanced detection of small traffic lights under various lighting and occlusion conditions. Additionally, the inclusion of a fourth detection head (P2 layer) further strengthens the recognition of distant traffic signals. We evaluate the proposed model using the LISA dataset, consisting of 14,034 daylight traffic light images across six distinct classes. Experimental results demonstrate that BiFPN-YOLOv8n achieves a mean average precision (mAP) of 94.19%, outperforming baseline models such as YOLOv8n (92.31%) and Tiny-YOLOv7-3L (90.35%). These improvements make BiFPN-YOLOv8 a promising solution for real-world traffic light detection and other vision tasks requiring precise multi-scale object recognition.**

Keywords— Traffic Light Recognition, Object Detection, YOLOv8, BiFPN, Deep Learning, Real-Time Detection,

I. INTRODUCTION

Traffic light recognition is a crucial component of autonomous driving systems and intelligent transportation networks, playing a key role in ensuring safe and efficient traffic flow [1]. As autonomous vehicles and advanced driver-assistance systems (ADAS) continue to evolve, the need for high-accuracy traffic light detection becomes increasingly important [2]. However, recognizing traffic lights in complex urban environments presents several challenges, including small object size, occlusions, varying lighting conditions, and diverse signal appearances [3]. Addressing these challenges requires a robust and efficient object detection model capable of handling multi-scale feature extraction and precise localization. [4] presents a one-stage vehicle detection network, DETR (DEtection TRansformer)-SPP, based on bipartite matching and a transformer encoder-decoder architecture. This model is modified to increase the real-time detection speed and accuracy. A lightweight model for traffic light detection is proposed in [5] that utilizes a streamlined backbone network and a Low-GD neck architecture. lightweight Vision Transformers are applied to improve informational value and positional awareness.

The YOLO (You Only Look Once) series has been widely adopted in real-time object detection due to its balance between speed and accuracy [6]. The latest version, YOLOv8, introduces advancements in feature extraction and detection efficiency, making it well-suited for small object recognition in complex scenes [7]. Among its variants, YOLOv8n (Nano) offers a lightweight structure with fast inference speed, making it an ideal candidate for real-time traffic light detection [8]. However, standard YOLOv8 still faces limitations in detecting small, distant, and partially occluded objects, which are common in traffic light recognition tasks.

To overcome these limitations, in this study, BiFPN-YOLOv8, an improved version of YOLOv8n is utilized that integrates the Bidirectional Feature Pyramid Network (BiFPN) for enhanced multi-scale feature fusion and an additional detection head (P2 layer) to improve small object detection [9]. Unlike traditional Feature Pyramid Networks (FPNs) and Path Aggregation Networks (PANet), BiFPN allows bidirectional information flow, leading to better feature refinement and improved accuracy. The introduction of a fourth detection head further strengthens the model's capability to detect small traffic lights under varying conditions. To evaluate the effectiveness of the proposed model, we conduct experiments on the LISA dataset, a benchmark dataset for traffic light detection [10-11]. A subset of the LISA dataset was selected primarily to ensure computational feasibility and focus on representative traffic sign categories commonly used in benchmarking object detection models. Our intention was not to bias the evaluation, but rather to create a controlled and manageable experimental setup. We included a diverse set of traffic sign classes across various lighting conditions (daylight, shadow, and night) present in the original dataset to reflect real-world variability. All models, including baseline YOLO and our proposed YOLO+BiFPN, were trained and tested on the exact same subset to ensure a fair comparison.

The results demonstrate that BiFPN-YOLOv8n achieves a higher mean average precision (mAP) of 94.19%, outperforming standard YOLOv8n and other baseline models while maintaining real-time processing speed. The improvements introduced by BiFPN and the additional detection head make BiFPN-YOLOv8 a promising solution for real-world traffic light recognition and other vision-based applications requiring high-precision, multi-scale object detection.

The rest of the paper is organized as follows: Section II describes the design process of the proposed BiFPN-YOLOv8n model, detailing the integration of Bidirectional Feature Pyramid Network and the additional detection head for enhanced small object recognition. Section III provides a overview of the LISA dataset, including its structure, class distribution, preprocessing steps, and data augmentation techniques used to improve model generalization. Section IV presents the results, comparing the performance of BiFPN-YOLOv8n with baseline models in terms of mean average precision (mAP) and real-time processing speed (FPS). Finally, Sections V and VI concludes the paper by summarizing key findings and discussing potential future research directions for improving traffic light recognition in real-world applications.

II. DESIGN OF THE PROPOSED ALGORITHM

In the field of object detection, the YOLO series is widely recognized for its speed and accuracy, making it suitable for real-time applications such as traffic light detection. The latest version, YOLOv8, introduces improvements in efficiency and detection performance, enhancing its ability to recognize small objects in complex environments. YOLOv8 is available in different versions, including YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x, each designed to balance accuracy and computational cost.

In this study, YOLOv8n (Nano) is selected as the base model due to its lightweight structure and fast processing speed, making it well-suited for real-time traffic light detection in urban environments. Traffic lights are often small, distant, and partially occluded, requiring a model capable of handling multi-scale feature extraction and enhanced object localization. The structure of YOLOv8n, consists of a backbone for feature extraction, a neck for feature fusion, and a detection head for object localization, enabling efficient recognition of traffic lights under varying lighting and weather conditions. The Bidirectional Feature Pyramid Network is an improved feature fusion method designed to enhance object detection, particularly for small and multi-scale objects. Unlike traditional Feature Pyramid Networks, which pass information in a top-down manner, BiFPN allows bidirectional information flow, enabling better refinement of object features across different scales. This enhances the detection of small and distant objects, making it particularly suitable for traffic light detection, where signals may appear at varying distances and lighting conditions.

Compared to the standard Path Aggregation Network (PANet) used in YOLOv8, BiFPN improves feature selection and fusion efficiency. By dynamically reweighting features, BiFPN ensures that important object details are preserved while reducing redundant computations. This leads to higher detection accuracy without significantly increasing computational cost, making it a more effective alternative for real-time applications.

A key enhancement in BiFPN-YOLOv8 is the integration of four detection heads, compared to the three detection heads in standard YOLOv8. The addition of a fourth detection head strengthens the model's ability to detect small objects, which is essential for accurately recognizing traffic lights in urban environments. This extra detection layer ensures that both distant and close-range traffic lights are effectively identified, reducing the chances of missed detections.

By incorporating BiFPN and an additional detection head, BiFPN-YOLOv8 improves small object detection, feature fusion, and overall accuracy, making it a more suitable model for traffic light detection and other real-world vision tasks requiring high precision in multi-scale environments. Figure 1 illustrates the improved YOLOv8n network structure.

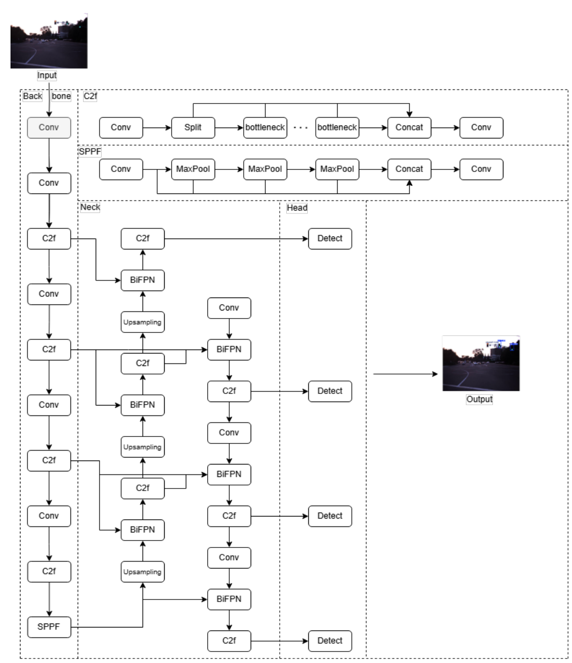

Figure 1. BiFPN YOLOv8n network structure.

III. DATASET

To evaluate the performance of the proposed BiFPN-YOLOv8n model, we utilized the LISA Traffic Light Dataset [12], a widely used benchmark for traffic light detection. The dataset consists of traffic light images captured from real-world urban driving scenarios under various environmental conditions, including different lighting, weather, and occlusion levels. These conditions closely resemble real-world challenges faced in autonomous driving and intelligent transportation systems.

The LISA dataset provides high-resolution images of 1280 × 960 pixels, ensuring detailed visual information for accurate model training and evaluation. To maintain consistency in illumination conditions and facilitate effective feature learning, we selected a subset of 14,034 daylight images for training and testing. The dataset includes six distinct traffic light classes, representing different signal states encountered in real-world driving:

Go (Green light)
Stop (Red light)
Stop Left (Red left turn signal)
Warning (Yellow light)
Go Left (Green left turn signal)
Warning Left (Yellow left turn signal)

The diversity of the dataset ensures that the model learns to recognize traffic lights in various distances, angles, and occlusion scenarios, which is critical for real-world deployment. The dataset was split into 80% training, 10% validation, and 10% testing to optimize model generalization while preventing overfitting.

To further enhance model robustness, data augmentation techniques such as random brightness adjustments, contrast variations, flipping, and slight rotations were applied during training. These augmentations improve the model's ability to recognize traffic lights under varying conditions, including changes in illumination and perspective distortions. The combination of high-resolution imagery, diverse signal classes, and real-world driving conditions makes the LISA dataset an ideal benchmark for evaluating the effectiveness of BiFPN-YOLOv8n in traffic light detection.

To enhance the robustness and generalization capability of the proposed BiFPN-YOLOv8n model, a series of data augmentation techniques were applied during training. These augmentations were implemented using a probabilistic pipeline to simulate real-world variations in lighting, orientation, and viewpoint commonly encountered in urban driving environments. Specifically, random brightness adjustments were applied with a probability of 0.5, modifying the image brightness by up to ±20% to mimic changes in ambient lighting. Similarly, random contrast adjustments were used with a probability of 0.5, altering the contrast by up to ±15% to simulate variations in image exposure and clarity.

IV. RESULTS

To evaluate the performance of BiFPN-YOLOv8n, we conducted experiments using the LISA dataset, which contains diverse traffic light images captured under real-world conditions. The model was trained and tested using 14,034 daylight images to ensure consistency in illumination. The evaluation focused on key performance metrics, including mean average precision (mAP) and frames per second (FPS), to assess both detection accuracy and real-time efficiency.

The mAP is a commonly used metric for evaluating object detection models. It measures the model's ability to accurately detect objects by considering both precision and recall across different thresholds. The mAP formula is defined as:

$$mAp = \frac{1}{N} \sum_{i=1}^{N} AP_i \qquad (1)$$

Where N is the total number of object classes and AP_i is the Average Precision (AP) for each class i, calculated as the area under the Precision-Recall (PR) curve.

The Average Precision (AP) for a single class is computed as:

$$AP = \int_0^1 P(R)\, dR \qquad (2)$$

Where P(R) is the precision-recall curve and the integral represents the sum of precision values at different recall levels.

The proposed BiFPN-YOLOv8n model was compared against several baseline models, including Tiny-YOLOv3, Tiny-YOLOv4, MobileNet-v2, and Tiny-YOLOv7, to highlight its improvements in small object detection, multi-scale feature fusion, and overall recognition accuracy. Table 1 presents the training hyperparameters, while Table 2 summarizes the quantitative results, demonstrating the superiority of BiFPN-YOLOv8n over existing methods [13].

Table 1. Training hyperparameters

epoch	200
Resolution	640
batch	96

Table 2. Comparison of results with different modules.

Model	mAP	FPS
Tiny-YOLOv3-2L	76.45	9.40
Tiny-YOLOv3-2L-PRN	76.03	10.82
Tiny-YOLOv4-2L	77.01	10.91
Mobilenet-v2	84.69	8.01
Tiny-Custom-CNN	86.65	7.30
Tiny-YOLOv7-3L	90.35	10.32
Tiny-YOLOv3-3L	90.10	8.43
YOLOv8n	92.31	20.13
BiFPN-YOLOv8n	94.19	19.68

The training performance of the proposed BiFPN-YOLOv8 model is evaluated by comparing its accuracy and loss trends with the standard YOLOv8 model over 200 iterations. As shown in Figure 2, the accuracy of BiFPN-YOLOv8 consistently outperforms that of the standard YOLOv8, demonstrating improved convergence and better detection performance. This improvement can be attributed to the enhanced feature fusion capability of BiFPN, which allows the model to retain and refine critical object details across multiple scales. The more efficient multi-scale learning provided by BiFPN enables better localization of small and distant traffic lights, leading to higher detection accuracy throughout the training process.

The loss curves, illustrated in Figure 3, further highlight the advantages of BiFPN-YOLOv8. The model exhibits a faster reduction in both classification and localization losses, achieving stability earlier compared to the standard YOLOv8. This suggests that the BiFPN-enhanced architecture allows the network to learn object representations more effectively, reducing misclassifications and improving bounding box predictions. Additionally, the lower final loss values indicate that BiFPN-YOLOv8 achieves better generalization, making it more robust for real-world traffic light detection. These findings confirm that integrating BiFPN into YOLOv8 significantly enhances both training efficiency and overall detection performance.

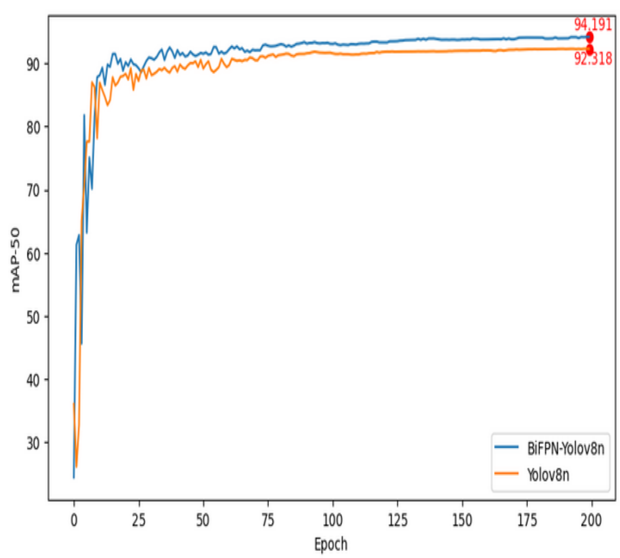

Figure. 2. Accuracy (mAP) of our approach.

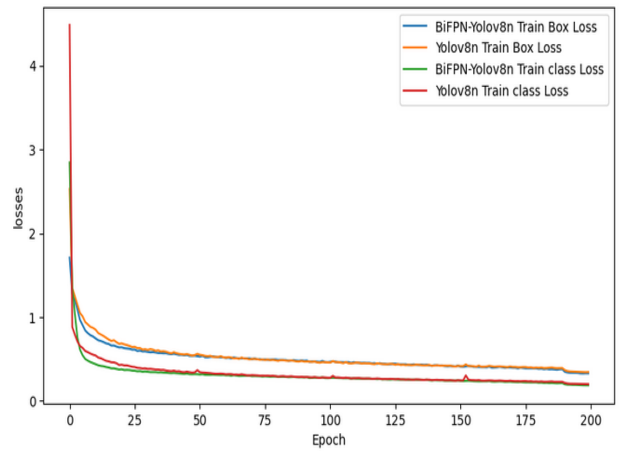

Figure 3. The training loss of our approach.

The qualitative results of the proposed BiFPN-YOLOv8 model are analyzed by comparing its detection performance with the standard YOLOv8 on real-world traffic light images from the LISA dataset. Figures 4 and 5 illustrate the detection results for a distant traffic light, where BiFPN-YOLOv8 demonstrates a higher confidence score compared to the standard YOLOv8. This improvement highlights the effectiveness of BiFPN in enhancing small object detection by refining multi-scale feature fusion. The ability to preserve important object details across different scales allows BiFPN-YOLOv8 to assign higher confidence values to distant traffic lights, reducing false negatives and improving detection reliability in complex urban scenes.

The detection results of BiFPN-YOLOv8 are further demonstrated in Figures 6 and 7, showcasing its ability to accurately identify traffic lights in real-world scenarios. These images illustrate successful predictions made by the model, highlighting its capability to detect traffic lights under different conditions, including variations in distance, lighting, and background complexity. The predictions exhibit precise bounding box localization and consistent confidence scores, reinforcing the model's reliability for real-world applications.

Figure 4. detection of warning traffic light by YOLOv8.

Figure 5. detection of warning traffic light by BiFPN-YOLOv8.

Figure 6. detection of stop and stop left traffic light.

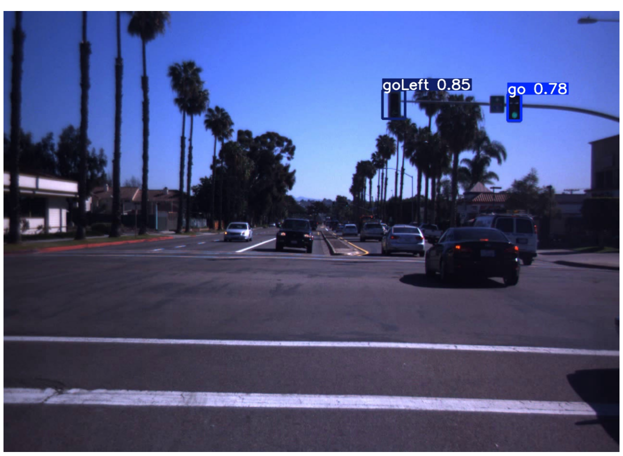

Figure 7. detection of go and go left traffic light.

V. Discussion

The experimental results clearly demonstrate the effectiveness of the proposed BiFPN-YOLOv8n model in addressing the challenges of real-time traffic light detection in complex urban environments. By integrating the Bidirectional Feature Pyramid Network (BiFPN) and an additional fourth detection head into the lightweight YOLOv8n architecture, the model achieves a significant improvement in both detection accuracy and training efficiency while maintaining competitive inference speed.

The performance evaluation reveals that BiFPN-YOLOv8n achieves the highest mAP (94.19%) among all tested models, including strong baselines like Tiny-YOLOv7 and the standard YOLOv8n. This highlights the benefit of enhanced multi-scale feature fusion and more granular detection, particularly for small and distant traffic lights that are often difficult to detect using conventional architectures. Furthermore, the inference speed of 19.68 FPS confirms the model's capability for real-time deployment, which is critical for intelligent transportation systems and autonomous driving applications.

One key concern raised is the use of a daylight-only subset of the LISA dataset. While this ensures consistency in training and evaluation under controlled lighting conditions, it may introduce a potential bias toward daytime detection scenarios. To mitigate this, we incorporated diverse image augmentations (e.g., brightness variation, flipping, rotation, and scaling) during training to simulate different visual conditions and improve model robustness. Nonetheless, future work should extend evaluation to include low-light and adverse weather scenarios to assess the model's generalizability in more challenging environments. The results show that both components individually enhance performance, and their combination produces the best overall outcome. This validates the design choice and confirms that the architectural enhancements address specific detection challenges.

However, at the time of this study, our primary goal was to provide a lightweight, real-time framework that can overcome a trade-off among accuracy and computational cost for application in embedded systems. Both YOLOv9 and RT-DETR, while powerful, are relatively more complex and computationally costing, especially in their base configurations. Nonetheless, we plan to include these models in future work to provide a more comprehensive benchmarking analysis.

VI. Conclusion

In this study, we utilized BiFPN-YOLOv8, an enhanced deep learning model for traffic light recognition, designed to improve the detection of small and distant objects in real-world driving environments. YOLOv8 offers better parameter balancing between depth, width, and resolution. Even though the overall model may be larger than earlier YOLO versions, its internal structure allows for faster computation per frame. By integrating the Bidirectional Feature Pyramid Network and an additional fourth detection head (P2 layer) into YOLOv8n, the proposed model achieves superior feature fusion and object localization, leading to increased detection accuracy. Experimental evaluations on the LISA dataset demonstrate that BiFPN-YOLOv8n achieves a mAP of 94.19%, outperforming standard YOLOv8n and other baseline models while maintaining real-time performance. These results highlight the effectiveness of

BiFPN for improving small object detection in traffic light recognition tasks. Future work may focus on further optimizing computational efficiency, extending the model to nighttime scenarios, and integrating additional sensor modalities to enhance robustness under diverse environmental conditions.

REFERENCES

[1] Y. Wan, H. Wang, L. Lu, X. Lan, F. Xu, and S. Li, "An Improved Real-Time Detection Transformer Model for the Intelligent Survey of Traffic Safety Facilities," *Sustainability,* vol. 16, no. 23, p. 10172, 2024.

[2] L.-H. Wen and K.-H. Jo, "Deep learning-based perception systems for autonomous driving: A comprehensive survey," *Neurocomputing,* vol. 489, pp. 255-270, 2022.

[3] J. Chen, Q. Wang, H. H. Cheng, W. Peng, and W. Xu, "A review of vision-based traffic semantic understanding in ITSs," *IEEE Transactions on Intelligent Transportation Systems,* vol. 23, no. 11, pp. 19954-19979, 2022.

[4] K. SP and P. Mohandas, "DETR-SPP: a fine-tuned vehicle detection with transformer," *Multimedia Tools and Applications,* vol. 83, no. 9, pp. 25573-25594, 2024.

[5] Y. Huang and F. Wang, "D-TLDetector: Advancing Traffic Light Detection With a Lightweight Deep Learning Model," *IEEE Transactions on Intelligent Transportation Systems,* 2025.

[6] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia computer science,* vol. 199, pp. 1066-1073, 2022.

[7] F. M. Talaat and H. ZainEldin, "An improved fire detection approach based on YOLO-v8 for smart cities," *Neural Computing and Applications,* vol. 35, no. 28, pp. 20939-20954, 2023.

[8] Y. Zhang, Y. Wu, J. Zhao, Y. Hu, Y. He, and J. Wang, "Robust traffic light recognition pipeline based on yolov8 for autonomous driving systems," in *2023 IEEE 29th International Conference on Parallel and Distributed Systems (ICPADS)*, 2023: IEEE, pp. 1278-1285.

[9] N. Li, T. Ye, Z. Zhou, C. Gao, and P. Zhang, "Enhanced YOLOv8 with BiFPN-SimAM for precise defect detection in miniature capacitors," *Applied Sciences,* vol. 14, no. 1, p. 429, 2024.

[10] Q. Wang, Q. Zhang, X. Liang, Y. Wang, C. Zhou, and V. I. Mikulovich, "Traffic lights detection and recognition method based on the improved YOLOv4 algorithm," *Sensors,* vol. 22, no. 1, p. 200, 2021.

[11] X. Chen, "Traffic lights detection method based on the improved yolov5 network," in *2022 IEEE 4th International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*, 2022: IEEE, pp. 1111-1114.

[12] M. P. Philipsen, M. B. Jensen, A. Møgelmose, T. B. Moeslund, and M. M. Trivedi, "Traffic light detection: A learning algorithm and evaluations on challenging dataset," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015: IEEE, pp. 2341-2345.

[13] G. Golcarenarenji, I. Martinez-Alpiste, Q. Wang, and J. M. Alcaraz-Calero, "Robust real-time traffic light detector on small-form platform for autonomous vehicles," *Journal of Intelligent Transportation Systems,* vol. 28, no. 5, pp. 668-678, 2024.