Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 8, Issue 1: 57-63, DOI: 10.25236/AJCIS.2025.080108

Lane detection based on boundary feature enhancement and information interaction

Zhenkun Hu[1,a,*], Yi Shen[1,b]

[1]*School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing, 400074, China*
[a]*2582809329@qq.com*, [b]*1639801731@qq.com*
[*]*Corresponding author*

Abstract: *Lane detection and drivable area segmentation are crucial for safe and efficient navigation on roads. To address the challenges of poor recognition under complex traffic conditions and limited generalization ability in existing lane detection methods, we propose an efficient and lightweight approach for lane and drivable area detection. By leveraging the concept of difference, we introduce a Differential Boundary-Aware Module (DBAM) that enhances lane boundary features and effectively captures the elongated nature of lane markings in complex scenarios. Additionally, we incorporate an Interactive Attention Network (IAN) to learn spatial dependencies between different task features, alleviating potential conflicts. Our method achieves competitive results on the BDD100K dataset, with lane detection Intersection over Union (IoU) reaching 31.3%, and drivable area mean IoU (mIoU) achieving 91.2%, while maintaining a processing speed of 130 FPS. The results demonstrate that the proposed method achieves excellent detection performance in complex scenes.*

Keywords: *Autonomous Driving, Lane Detection, Drivable Area Segmentation, Boundary-Aware, Interactive Attention*

1. Introduction

Lane detection technology is a key component of intelligent driving systems, crucial for accurate vehicle localization and reliable navigation. In autonomous driving tasks, multi-sensor fusion techniques, such as monocular, binocular, and LiDAR sensors, are commonly used in intelligent vehicles. However, camera-based methods stand out due to their low cost. Therefore, combining cameras with deep learning models has become a powerful solution. Additionally, high precision, real-time processing, and lightweight design are essential in autonomous driving tasks.

Object detection and image segmentation are fundamental challenges in computer vision, both aiming to identify regions of interest in images using different processing techniques. A series of pioneering works have been proposed for object detection, including CenterNet[1], Faster R-CNN[2], and the YOLO series [3, 4, 5, 6]. For segmentation, common networks such as FCN[7], SegFormer[8], and UNet [9] have been widely used. While these methods are effective, they are typically designed for single tasks and are not capable of handling multiple tasks simultaneously. In recent years, efforts have been made to implement more complex strategies for computer vision tasks[10, 12, 13, 14, 15, 16], with unified multitask models offering the advantage of simplifying algorithms and saving computational resources. However, in complex scenarios, lane detection remains a challenging task due to the small size and limited representation of lane markings in images. To address this, more robust networks are needed to capture fine-grained features while considering the distinct characteristics of different tasks.

To address this, we propose a lane detection algorithm suitable for complex traffic scenarios, which improves detection accuracy and robustness while maintaining inference speed. Inspired by the current state-of-the-art method, TwinLiteNet, our approach first introduces a Differential Boundary-Aware Network (DBAN) by leveraging the concept of differences, and embedding it into the backbone network to enhance boundary features. Next, we incorporate Interactive Attention to learn the spatial dependencies between lane markings and drivable area features, mitigating conflicts. By introducing a self-attention mechanism, we capture the spatial dependencies between any two positions within the same feature map and across different feature maps, allowing for the extraction of sufficient object-specific information. Finally, our method achieves competitive results on the BDD100K dataset, particularly in lane detection.

ISSN 2616-5775 Vol. 8, Issue 1: 57-63, DOI: 10.25236/AJCIS.2025.080108

2. Related Works

2.1. Lane Detection

Lane detection is a critical component of autonomous driving systems, enabling vehicles to stay within the road boundaries and follow the correct path. Accurate lane detection ensures safe and efficient navigation by identifying lane markings, which are essential for tasks such as trajectory planning, lane-keeping assistance, and adaptive cruise control. The task is particularly challenging due to various factors such as varying road conditions, lighting, occlusions, and the subtle appearance of lane markings in some environments.

Traditional lane detection methods primarily relied on edge detection, Hough transforms[17], and geometric models. These methods, while effective in well-structured environments, often struggle in complex real-world scenarios, where road markings may be partially occluded, faded, or missing. As a result, they typically have limited robustness to changes in weather, lighting, and road conditions.

LaneNet[18] is another popular deep learning-based method for lane detection, which uses a combination of a CNN-based feature extractor and a clustering algorithm to identify lane markings. LaneNet first uses a deep network to generate a binary mask indicating the presence of lane markings. Then, a clustering algorithm is applied to group the connected lane pixels into distinct lanes. This approach works well in scenarios where lane markings are relatively continuous and easy to separate. However, it may struggle when lane markings are broken or occluded.

An important advancement in lane detection has been the adoption of transformer-based models. Transformers, renowned for their capacity to model long-range dependencies, have demonstrated significant potential in capturing global spatial relationships between lane markings and road structures. By leveraging self-attention mechanisms, transformers can effectively capture contextual information from distant parts of the image, which is particularly beneficial for handling complex and curved lane geometries, as well as occlusions. In addition to improving detection accuracy, real-time performance is critical for autonomous driving systems. As a result, recent research has increasingly focused on developing lightweight models that strike a balance between high accuracy and computational efficiency. These models leverage various techniques such as network pruning, quantization, and the use of efficient backbone networks to achieve fast inference times while preserving robust lane detection performance. Architectures like MobileNetV2[19] and ShuffleNet[20] are frequently integrated into lane detection pipelines to ensure that the system can operate in real-time on resource-constrained embedded hardware.

2.2. Drivable Area Segmentation

Drivable area segmentation is a crucial task in autonomous driving systems, aiming to identify the regions of the road that are safe and feasible for a vehicle to navigate. This task plays a key role in path planning and navigation, ensuring the vehicle can safely and efficiently move through complex and dynamic environments. The goal is to accurately delineate the drivable areas from the surrounding obstacles, lane markings, and other irrelevant regions.

As the field progressed, the rise of deep learning has significantly advanced the performance of this task. Convolutional Neural Networks (CNNs) have become the backbone for segmentation models due to their ability to learn hierarchical features from raw pixel data. FCN[7] and U-Net[9] have been widely adopted for semantic segmentation tasks, including drivable area segmentation, where the goal is to classify each pixel as either drivable or non-drivable.

Recent advancements in drivable area segmentation have leveraged multitask learning, where a single network simultaneously addresses multiple tasks, such as lane detection, obstacle detection, and drivable area segmentation. This approach not only simplifies the overall system architecture but also enhances the synergy between tasks, leading to improved performance across all objectives. Notable examples of this approach include methods like YOLOP[12] and A-YOLOM[16], which integrate object detection and semantic segmentation into a unified framework, demonstrating the potential of multitask learning for improving both efficiency and accuracy in autonomous driving systems.

3. Proposed method

Figure 1 shows the BILane. The input image is fed into the backbone network, which is responsible for extracting features related to lane markings and drivable area regions. After passing through the

ISSN 2616-5775 Vol. 8, Issue 1: 57-63, DOI: 10.25236/AJCIS.2025.080108

backbone[11], a shared feature map is obtained. This shared feature map is subsequently processed by an interactive attention mechanism, which learns the spatial dependencies between lane markings and drivable area features, thereby mitigating conflicts between these two tasks. Finally, the processed features are forwarded to separate decoders, each dedicated to a specific task: lane detection and drivable area segmentation. Each decoder consists of upsampling layers and detection heads, which produce the final outputs: lane detection results and drivable area segmentation.

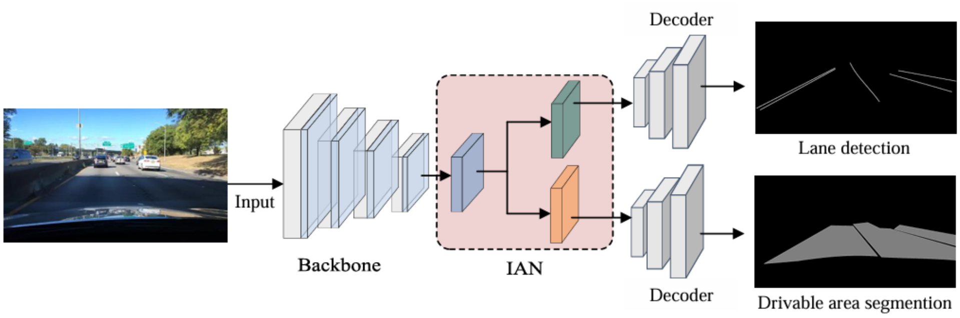


Figure 1: Network architecture.

3.1. Differential Boundary-Aware Network

In complex driving scenarios, lane markings may become blurred due to factors such as shadows, lighting variations, and adverse weather conditions, making it challenging to capture subtle lane information in scenes with weak appearance cues, such as severe occlusion, wear, or nighttime driving. Inspired by [21], we propose a Differential Boundary Awareness Module, which is embedded into the backbone network to perceive changes in lane boundaries. This module enhances the model's ability to adapt to challenging traffic environments by effectively capturing and emphasizing the variations in lanes.

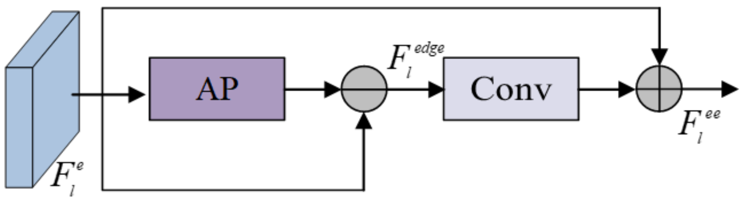


Figure 2: Differential boundary-aware network architecture.

The Differential Boundary-Aware Network is illustrated in Figure 2. Inspired by the differential approach, we propose a Differential Boundary Awareness Module designed to enhance lane boundary features. In this architecture, the feature map F_l^e is subtracted from the average-pooled feature layer to obtain a feature map F_l^{edge}, as described in equation (1). This is followed by a 1×1 convolution operation, and the result is then added back to the original feature map F_l^e, yielding an enhanced feature map with boundary information, as described in equation (2). This architecture enables the network to effectively extract boundary information for lane markings, allowing it to capture crucial features even in complex scenarios such as nighttime driving, shadows, and occlusions, thereby demonstrating significant practical value.

$$F_l^{\text{edge}} = F_l^e - AP(F_l^e) \tag{1}$$

$$F_l^{ee} = conv(F_l^{\text{edge}}) + F_l^e \tag{2}$$

3.2. Interactive Attention Network

To effectively extract sufficient lane features, the Interactive Attention Network (IAN) [22] introduced in this paper successfully reduces feature conflicts between the lane detection and drivable area segmentation tasks, as illustrated in Figure 3. The IAN is built upon a self-attention mechanism, where the internal process optimizes vector distances through dot-product operations. When two vectors are close in the feature space, it indicates their similarity. In this work, we leverage the self-attention mechanism to compute the correlations between different positions and establish spatial topological relationships, thereby enhancing the focus on key task-specific features. In the IAN network, the self-attention mechanism effectively enables each task to focus more on its specific objective while coordinating the competition between multiple tasks. This enhances the network's overall ability to extract relevant feature information.

As shown in Figure 3, the shared features $f^t \in \mathbb{R}^{H \times W \times C}$ is obtained from the backbone network. This improves the overall ability of the network architecture to extract feature information. First, average pooling is applied to f^t to reduce computation and obtain the feature map $f'^t \in \mathbb{R}^{H' \times W' \times C}$. Then, convolutional layers are applied to f'^t to encode and separately construct the lane detection feature map A_1 and the drivable area feature map A_2.

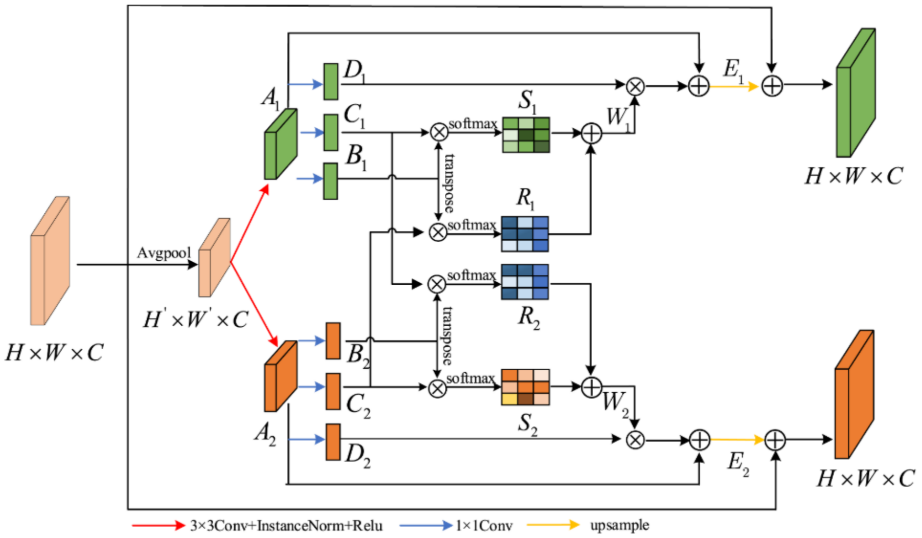


Figure 3: Interactive attention network.

A_k is then fed into a 1 × 1 convolution to build two feature maps B_k and C_k of size $\mathbb{R}^{C \times N}$,where $N = H' \times W'$. B_1^T and C_1^T as well as B_2^T and C_2^T perform matrix multiplication, where *T* denotes a transpose operation. The Softmax function is used to determine a task's spatial attention $\{S_1, S_2\} \in \mathbb{R}^{N \times N}$. The calculation method is shown in (3):

$$S_k^{ji} = \frac{\exp\left(B_k^i \cdot C_k^j\right)}{\sum_{l=1}^{N} \exp\left(B_k^l \cdot C_k^j\right)}, k \in \{1, 2\} \tag{3}$$

Where B_k^i and C_k^i are the i-th and j-th lines of B_k and C_k, respectively; s_k^{ji} is the value of the (*j,i*) position on s_k, which is the correlation between the i-th and j-th positions in the feature map; and B_1^T and C_2, as well as B_2^T and C_2, perform matrix multiplication, and calculate the spatial attention $\{R_1, R_2\} \in \mathbb{R}^{N \times N}$ between different tasks using the Softmax function.

$$\boldsymbol{W}_k = \eta_k \times \boldsymbol{S}_k + (1 - \eta_k) \times \boldsymbol{R}_k, k \in \{1, 2\} \tag{4}$$

Then, the two spatial attention maps are fused using a learnable parameter η, resulting in $\{W_1, W_2\} \in \mathbb{R}^{N \times N}$. The calculation method is shown in (4). Meanwhile, A_k is passed through a 1×1 convolutional layer to generate new feature maps $D_k \in \mathbb{R}^{C \times N}$. Matrix multiplication is then performed between D_1 and W_1^T, as well as between D_2 and W_2^T, and the outputs are reshaped into tensors of size $\mathbb{R}^{H' \times W' \times C}$.Then, element-wise summation is performed on the feature maps A_1 and A_2 to obtain the output

ISSN 2616-5775 Vol. 8, Issue 1: 57-63, DOI: 10.25236/AJCIS.2025.080108

$E_k \in \mathbb{R}^{H' \times W' \times C}$. The calculation method is shown in (5):

$$E_k^j = \sum_{i=1}^{N}\left(w_k^{ji} D_k^i\right) + A_k^j, k \in \{1,2\} \tag{5}$$

Finally, the feature maps E_1 and E_2 are upsampled to the same scale as the input feature f_t, and then fused with f_t to obtain two types of features for lane detection and drivable area segmentation.

4. Experimental Analysis

4.1. Dataset

The BDD100K dataset is used for training and evaluation of BILane, encompassing diverse weather conditions and terrains. It is divided into three subsets: a training set with 70,000 images, a validation set with 10,000 images, and a test set with 20,000 images. Since the test set lacks ground truth labels, we evaluate our method on the 10,000 images from the validation set. For consistency, we resize the images from the original resolution of $1280 \times 720 \times 3$ to $720 \times 640 \times 3$.

4.2. Evaluation Metrics and Experimental Environment

The evaluation metrics used in the experiments follow those in YOLOP [12]. For lane detection, we adopt Intersection over Union (IoU), while for drivable area segmentation, we use mean Intersection over Union (mIoU). In addition, the model inference speed is measured in terms of FPS (frames per second), while the model size is represented by the number of parameters.

All experiments were conducted on a system equipped with an RTX 3090 GPU and an Intel(R) Xeon(R) Platinum 8362 processor. The model training and evaluation were implemented using Python 3.8 and the PyTorch 1.8.0 framework. The training process utilized the Adam optimizer with a weight decay of 0.0005. The batch size was set to 8. The model was trained for a total of 100 epochs, taking approximately 80 hours to complete.

4.3. Lane Detection Result

Table 1 presents the results on the BDD100K dataset, where bold text indicates the best performance and underlined text represents the second-best. As shown, our proposed method achieves an IoU of 31.3%, surpassing A-YOLOM(n)[16], Sparse U-PDP[15], and TwinLiteNet[13] by 3.1%, 0.1%, and 0.8%, respectively. Additionally, BILane operates at 3.25 times the inference speed of A-YOLOM(n)[16]. Intelligent driver assistance systems require timely responses to lane changes, which are critical for vehicle safety. Compared to TwinLiteNet, our method strikes an optimal balance between IoU and FPS. These findings demonstrate the effectiveness of BILane in handling diverse driving scenarios.

Table 1: Comparison with other state-of-the-art methods on the BDD100K dataset.

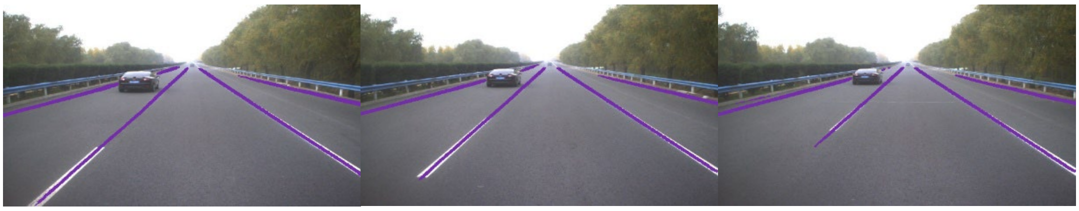
Method	Lane IoU(%)	Drivable Area mIoU(%)	Param.	FPS
YOLOP[12]	26.5	**91.6**	7.9M	49
IALaneNet (ResNet-18)[14]	30.4	90.6	17.1M	58
IALaneNet (ResNet-34) [14]	30.5	90.5	27.2M	40
A-YOLOM(n) [16]	28.2	90.5	4.4M	40
A-YOLOM(s) [16]	28.8	91.0	13.6M	40
Sparse U-PDP[15]	<u>31.2</u>	<u>91.5</u>	12.1M	29
TwinLiteNet[13]	30.5	91.0	**0.4M**	**185**
BILane(Ours)	**31.3**	91.2	<u>1.4M</u>	<u>130</u>

4.4. Drivable Area Segmentation Result

We present the drivable area segmentation results in Table 1, where it can be observed that BILane achieves competitive performance in terms of mIoU. Specifically, its mIoU surpasses A-YOLOM(n)[16] and TwinLiteNet[13] by 0.7% and 0.2%, respectively. We introduce the Interactive Attention Mechanism, which effectively mitigates the conflicts between lane detection and drivable area segmentation tasks. By learning the spatial dependencies between these tasks, the mechanism enhances the accuracy of drivable area detection, leading to more precise and robust segmentation results. Our method achieves a mIoU that is 0.3% lower than Sparse U-PDP [15]; however, it operates at 4.48 times the speed of Sparse U-PDP, demonstrating the effectiveness of our approach.

4.5. Real-World Road Testing

To validate the generalization ability of the proposed method, we conducted experiments on the Lane Detection dataset from Xi'an Jiaotong University, with the visualization results shown in Figure 4. As can be seen, our method effectively detects lane markings, demonstrating strong adaptability and robustness. This further proves the effectiveness and broad applicability of our proposed model in complex scenarios.



(a) *(b)* *(c)*

Figure 4: Actual road visualization.

5. Conclusions

In this paper, we propose the BILane, a method that leverages the Boundary-Aware Module to effectively capture boundary variations and enhance lane line features, making it well-suited for complex driving scenarios. This module enables the model to adapt to challenging environments, where lane markings may be ambiguous or occluded due to factors such as changes in lighting, weather conditions, or road degradation. Furthermore, we introduce the interactive attention mechanism, which effectively alleviates the conflicts between lane detection and drivable area segmentation tasks. By learning the spatial dependencies between these tasks, the mechanism significantly improves the accuracy of both lane detection and drivable area segmentation. The proposed method demonstrates robust performance in real-world, complex traffic scenarios, achieving an optimal balance between detection accuracy and computational efficiency.

References

[1] Duan K, Bai S, Xie L, et al. Centernet: Keypoint triplets for object detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2019: 6569-6578.
[2] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39(6): 1137-1149.
[3] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
[4] Wang C Y, Bochkovskiy A, Liao H Y M. Scaled-yolov4: Scaling cross stage partial network[C]// Proceedings of the IEEE/cvf conference on computer vision and pattern recognition. 2021: 13029-13038.
[5] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 7464-7475.
[6] Varghese R, Sambath M. YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024: 1-6.
[7] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.
[8] Xie E, Wang W, Yu Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [J]. Advances in neural information processing systems, 2021, 34: 12077-12090.
[9] Huang H, Lin L, Tong R, et al. Unet 3+: A full-scale connected unet for medical image

ISSN 2616-5775 Vol. 8, Issue 1: 57-63, DOI: 10.25236/AJCIS.2025.080108

segmentation[C]// ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2020: 1055-1059.

[10] Qian Y, Dolan J M, Yang M. DLT-Net: Joint detection of drivable areas, lane lines, and traffic objects [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(11): 4670-4679.

[11] Mehta S, Rastegari M, Caspi A, et al. Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation[C]//Proceedings of the european conference on computer vision (ECCV). 2018: 552-568.

[12] Wu D, Liao M W, Zhang W T, et al. Yolop: You only look once for panoptic driving perception[J]. Machine Intelligence Research, 2022, 19(6): 550-562.

[13] Che Q H, Nguyen D P, Pham M Q, et al. TwinLiteNet: An efficient and lightweight model for driveable area and lane segmentation in self-driving cars[C]//2023 International Conference on Multimedia Analysis and Pattern Recognition (MAPR). IEEE, 2023: 1-6.

[14] Tian W, Yu X, Hu H. Interactive attention learning on detection of lane and lane marking on the road by monocular camera image[J]. Sensors, 2023, 23(14): 6545.

[15] Wang H, Qiu M, Cai Y, et al. Sparse u-pdp: A unified multi-task framework for panoptic driving perception [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(10): 11308-11320.

[16] Wang J, Wu Q M J, Zhang N. You only look at once for real-time and generic multi-task[J]. IEEE Transactions on Vehicular Technology, 2024.

[17] Satzoda R K, Sathyanarayana S, Srikanthan T, et al. Hierarchical additive Hough transform for lane detection [J]. IEEE Embedded Systems Letters, 2010, 2(2): 23-26.

[18] Neven D, De Brabandere B, Georgoulis S, et al. Towards end-to-end lane detection: an instance segmentation approach[C]//2018 IEEE intelligent vehicles symposium (IV). IEEE, 2018: 286-291.

[19] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 4510-4520.

[20] Zhang X, Zhou X, Lin M, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 6848-6856.

[21] Zhang Z, Bao L, Xiang S, et al. B2CNet: A Progressive Change Boundary-to-Center Refinement Network for Multi-Temporal Remote Sensing Images Change Detection [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024.

[22] Ma S, Duan S, Hou Z, et al. Multi-object tracking algorithm based on interactive attention network and adaptive trajectory reconnection[J]. Expert Systems with Applications, 2024, 249: 123581.