Theory and Practice of Science and Technology
2025, VOL. 6, NO. 8, 231-234
DOI: 10.47297/taposatWSP2633-456943.20250608

Research on Model Lightweighting in Road Lane Detection

Zhihao Wang

Digital Technology and Engineering College, Ningbo University of Finance and Economics, Computer Science and Technology, Ningbo, Zhejiang, 315000, China

ABSTRACT

With the rapid development of intelligent transportation systems, autonomous driving technologies have imposed higher requirements on the real-time performance and accuracy of road detection. However, traditional deep learning models often suffer from high computational complexity and large storage demands in road detection tasks, limiting their application on in-vehicle and edge computing devices. This paper focuses on the application of model lightweighting techniques in road detection, aiming to significantly reduce the computational complexity and storage requirements while maintaining high detection accuracy through model structure optimization, pruning, and quantization.This study first reviews the current status of model lightweighting technologies in road detection both domestically and internationally, analyzing the advantages of lightweight model structures such as MobileNet and the development of pruning and quantization techniques. Subsequently, a lightweight road detection model architecture based on an improved MobileNet is proposed. By optimizing network layer structures, adjusting convolution kernel sizes and numbers, and reducing redundant computations, the model dynamically adjusts computational load during runtime through sparse networks and dynamic pruning structures. Effective evaluation metrics and optimization algorithms are also designed to address the critical balance between accuracy and lightweighting.In the experimental section, comparative experiments were conducted on the YOLOv11 model using existing lightweighting techniques. Results show that the lightweight model retains high detection accuracy while significantly improving inference speed and reducing model size, making it more adaptable to various road scenarios and hardware conditions. The findings not only provide an efficient model solution for road detection tasks but also offer insights for applying lightweighting techniques to other computer vision tasks.

KEYWORDS

Model lightweighting; Road detection; Intelligent transportation; Real-time performance; Computational complexity

1 Introduction

With the rapid advancement of intelligent transportation systems, autonomous driving technology faces both significant challenges and opportunities. Road lane detection, as a critical component, has become increasingly important. In real-world road detection scenarios—whether for real-time traffic monitoring by intelligent transportation systems or rapid road condition recognition by autonomous vehicles—high detection efficiency is essential.

However, traditional complex deep learning models often require high-performance computing devices such as high-end GPUs, which not only increases deployment costs but also limits their application on resource-constrained embedded devices and mobile terminals, hindering large-scale deployment in devices like smart traffic cameras and in-vehicle systems. Moreover, road detection scenarios are highly diverse, including urban streets, highways, and rural roads, each with varying hardware and environmental conditions, posing severe challenges to the flexibility and generalization of detection models.

In this context, model lightweighting technology has emerged and developed rapidly. Lightweighting aims to reduce model parameters and computational load while maintaining acceptable accuracy, thereby improving inference speed, lowering hardware costs, and enabling efficient operation on low-cost hardware across diverse scenarios. Internationally, significant progress has been made: Google's MobileNet series uses depthwise separable convolutions; Facebook's ShuffleNet employs channel shuffling; NVIDIA's structured pruning techniques; and IBM's low-bit quantization algorithms have all shown strong performance in image classification and road scene segmentation. Domestically, Tsinghua University's attention-based lightweight road detection model and Huawei's adaptive pruning algorithm have also played key roles in practical applications, advancing the intelligent transportation industry.

Despite these achievements, challenges remain in road lane detection lightweighting. For instance, avoiding significant accuracy loss due to excessive compression or simplification while finding the optimal balance between accuracy and lightweighting remains a key issue. Additionally, designing lightweight architectures tailored to road detection tasks and ensuring stable and accurate performance across different road scenarios, weather conditions, and lighting environments are critical concerns.

Based on the above research status and challenges, this study proposes the topic "Research on Model Lightweighting in Road Lane Detection." Through in-depth exploration of model structure optimization, pruning, quantization, and other lightweighting techniques, we aim to develop lightweight architectures suitable for road detection tasks, addressing critical issues like balancing accuracy and lightweighting. We hope to provide an efficient, low-cost, and highly

generalizable solution for road lane detection in intelligent transportation systems, further advancing autonomous driving and intelligent transportation.

2 YOLOv11 model

YOLOv11 is a significant version in the YOLO series for object detection, inheriting the core idea of the YOLO family, which treats the object detection task as a single regression problem, directly mapping image pixels to bounding-box coordinates and class probabilities. This approach avoids the complex region-proposal extraction and multiple network forward passes required by traditional object detection methods, thereby achieving efficient real-time detection. Building on the strengths of its predecessors, YOLOv11 further refines model architecture and algorithmic performance, enhancing both detection accuracy and speed. Its network architecture consists of three main components: a backbone network, a neck network, and a head network. The backbone network is responsible for extracting image features and typically employs Darknet-53 or other efficient architectures such as MobileNet and ResNet. The output of the backbone network is multi-scale feature maps used for subsequent object detection. The neck network serves to fuse feature maps of different scales, enhancing feature representation; it commonly adopts structures such as the Feature Pyramid Network (FPN) or Path Aggregation Network (PANet), which use upsampling and fusion operations to improve detection of small objects. The head network is responsible for performing target predictions on the fused feature maps, containing multiple convolutional layers and prediction layers, each of which is tasked with predicting targets at a specific scale, outputting bounding-box coordinates (x, y, w, h), confidence scores, and class probabilities. Regarding the detection mechanism, the YOLOv11 pipeline can be summarized as feature extraction, feature fusion, target prediction, and non-maximum suppression (NMS). The input image first passes through the backbone network for feature extraction, generating multi-scale feature maps. The neck network fuses these feature maps to enhance their expressive power. The head network then performs target prediction on the fused feature maps; each prediction layer is responsible for predicting targets at a specific scale, outputting bounding-box coordinates, confidence scores, and class probabilities.

To eliminate duplicate detections, YOLOv11 employs the non-maximum suppression (NMS) algorithm during the prediction phase, which compares the confidence scores of bounding boxes, retains the box with the highest confidence, and removes others whose overlap exceeds a threshold. The loss function of YOLOv11 integrates bounding-box localization loss, confidence loss, and class probability loss. Localization loss measures the discrepancy between predicted and ground-truth bounding boxes and is typically implemented using mean squared error (MSE). Confidence loss measures the difference between the predicted confidence score and the ground-truth label and is usually implemented using binary cross-entropy loss. Class probability loss measures the difference between predicted class probabilities and the true classes and is implemented using categorical cross-entropy loss.

Through this composite loss function, YOLOv11 simultaneously optimizes bounding-box localization accuracy and improves target classification precision. Compared with earlier versions, YOLOv11 incorporates several improvements that further enhance detection accuracy and speed. Its refined backbone employs a more efficient architecture capable of extracting richer feature information while reducing computational load. The introduction of a multi-scale detection mechanism enables YOLOv11 to better handle objects of varying sizes, improving the detection of small objects. The improved loss function better aligns with practical detection requirements, further enhancing model robustness. Additionally, by optimizing the sizes and quantities of anchor boxes, YOLOv11 increases the model's adaptability to targets of different shapes.

3 Improved YOLOv11 network model

3.1 Introduction of improved mobilenet

MobileNet is a lightweight deep neural network architecture specifically designed for mobile and embedded devices, and its core innovation lies in the use of depthwise separable convolution. This convolution method decomposes the traditional standard convolution into two steps: a depthwise convolution and a pointwise convolution, drastically reducing both computational cost and parameter count. In the depthwise convolution, each input channel is convolved independently, eliminating cross-channel computation and significantly lowering computational complexity, while the pointwise convolution applies 1×1 kernels to linearly combine the outputs of the depthwise convolution, restoring inter-channel correlations. This design not only preserves the model's feature-extraction capability but also greatly improves computational efficiency, enabling MobileNet to run efficiently on low-power devices.

3.2 Overall lightweighting based on mobilenet

Building upon YOLOv11, the introduction of an improved MobileNet structure aims to optimize the model's overall performance through lightweight techniques. Specifically, the improved MobileNet is incorporated into YOLOv11 in the following ways: first, MobileNet is employed as the backbone network in place of YOLOv11's original traditional backbone. Its lightweight characteristics significantly reduce computational cost and parameter count while extracting image features, thereby providing a more efficient foundational feature representation for subsequent detection tasks. Second, within the neck network, the design philosophy of MobileNet is leveraged to further refine the feature-fusion mechanism. By adjusting the architectures of the Feature Pyramid Network (FPN) or Path Aggregation Network (PANet) to align with MobileNet's lightweight properties, the model fuses multi-scale features more effectively while lowering computational complexity, thus enhancing the detection of small objects. Finally, in the head network, the prediction layers are optimized to suit the lightweight feature representations. Through careful adjustment of the number and structure of convolutional layers, the model maintains accurate prediction of object bounding boxes and class probabilities despite the reduced computation.

The integration of the improved MobileNet into YOLOv11 not only demonstrates significant optimization potential in theory but also delivers outstanding performance in practice. Experimental results show that, with the improved MobileNet, YOLOv11 achieves a marked increase in inference speed, a substantial reduction in computational and storage requirements, and only a slight decrease—sometimes even no decrease—in detection accuracy under certain scenarios compared with the original model. This lightweight enhancement makes YOLOv11 better suited for deployment on resource-constrained devices such as mobile phones and smart cameras, meeting the demands of real-time object detection and providing a more efficient and practical solution for intelligent transportation, autonomous driving, and video surveillance applications.4 Experimental Results and Analysis

4 Experimental results and analysis

4.1 Experimental environment

The experiments were conducted on the Windows 11 operating system with an NVIDIA GeForce RTX 3060 GPU, 16 GB of RAM, and an Intel Core i7-12700H CPU. Model training was carried out using PyCharm as the deep-learning framework, CUDA 12.8 for GPU acceleration, and the Python programming language.

4.2 Opendatalab dataset

To validate the performance of the improved algorithm, this experiment adopts the OpenDataLab dataset as the evaluation benchmark. Within the OpenDataLab collection, the lane-line image dataset covers a variety of challenging scenarios, exemplified by the CurveLanes subset. CurveLanes is a newly introduced baseline for lane detection, comprising 150 K lane images specifically suited to difficult cases involving curved roads and multiple lanes. The images were captured in real urban and highway environments across several cities in China, making it the largest lane-detection dataset to date and establishing a more demanding benchmark for the community. For every image, all lanes are manually annotated with natural cubic splines. The images are carefully curated so that most contain at least one curved lane. The dataset further presents more demanding situations—such as S-curves, Y-shaped lanes, nighttime scenes, and multi-lane configurations—rendering it ideal for performance assessment under complex conditions.

4.3 Results

To comprehensively evaluate the performance of the improved model, this study adopts mean Average Precision at an IoU threshold of 0.5 (mAP50/%), parameter count (Params), Giga Floating-point Operations (GFLOPs), frames per second (FPS), and model weight file size (MB) as evaluation metrics. A Multi-Objective Optimization (MOO) framework is further employed to integrate these indicators for an overall assessment. Among them, mAP50/% balances precision and recall across all classes to reflect detection accuracy; its calculation is given in formula (1), where denotes the average precision of the i-th class.

Experimental comparisons reveal that the lightweight YOLOv11n reduces parameters from 2.62 M to 2.1 M (-20 %), lowers GFLOPs from 6.7 G to 4.2 G (-37 %), and shrinks the weight file to only 4.1 MB (-24 %). While preserving an mAP50 of approximately 37.8 %, FPS on the Jetson platform rises from 111 to 158 (+42 %), confirming the effectiveness of the MobileNetV4 backbone combined with pruning and quantization. This offers an efficient, low-cost real-time road-lane detection solution for edge deployment.

$$\mathrm{mAP50} = \frac{1}{\mathrm{m}} \sum_{i=1}^{m} AP_i \tag{1}$$

Table 1 Comparative experimental results

Model	*Params* /10^6	*GFLOPs* /10^9	*mAP*50/%	*FPS*	*Weight Size* /*MB*
*YOLOv*11*n*	2.62*M*	6.7*G*	39.5	111	5.4*MB*
*Lightweight YOLOv*11*n*	2.1*M*	4.2*G*	37.8	158	4.1*MB*

5 Conclusion

In this study, we address the problem of road-lane detection for intelligent transportation and autonomous-driving applications by proposing a model-lightweighting approach. By integrating MobileNet-based lightweight techniques into the YOLOv11 architecture, we simultaneously enhance detection efficiency and accuracy while substantially reducing computational complexity and storage requirements. Experimental results demonstrate that the lightweight model retains high detection precision and satisfies real-time constraints—both of which are essential for autonomous-driving and intelligent-transportation systems. Moreover, the lightweight model maintains stable and accurate performance across diverse road scenarios, weather conditions, and lighting environments, exhibiting strong generalization. Reduced hardware demands enable deployment on resource-constrained devices such as embedded systems and mobile terminals, facilitating large-scale adoption and lowering hardware costs.

The novelty of this work lies in applying MobileNet's depthwise-separable convolution to road-lane detection, achieving effective model lightweighting through carefully designed network structures and parameter optimization without sacrificing accuracy. Additionally, we explored emerging architectures—such as sparse networks and dynamically prunable networks—offering new avenues for further computational-cost reduction.

Although promising results have been achieved, several directions merit future investigation. More efficient lightweight strategies can be pursued to further shrink model size and computational load. Broader scenario validation—spanning varied traffic environments and complex road conditions—should be conducted. Refining the loss function and optimization algorithms to boost detection accuracy and robustness represents another key objective. Finally, adapting the lightweight model to diverse hardware platforms to fully leverage its advantages warrants continued attention.

Overall, this study delivers an effective lightweight solution for road-lane detection and constitutes a valuable step toward advancing intelligent transportation and autonomous-driving technologies. Future research is expected to further optimize model architectures and algorithmic performance, ultimately enabling even more efficient and accurate road-lane detection.

References

[1] An Mengjun, Wu Chaoming, Deng Chengzhi. MEC-YOLOv11n: Detection algorithm for small floating objects on water surfaces [J/OL]. Electronic Measurement Technology. https://link.cnki.net/urlid/11.2175.TN.20250513.1703.003.

[2] Yu Qingcang, Qiu Rui, Fu Linjie, et al. Lightweight YOLOv7-VSS model for segment appearance inspection of orange slices [J]. Modern Electronics Technique, 2025, 48(10): 85-91.

[3] Lai Chaofan, Hua Qiang, Mu Jingyue, Zhang Bo. Lightweight road surveillance detection model based on network structure [J/OL]. Electronic Measurement Technology. https://link.cnki.net/urlid/11.2175.TN.20250514.1613.004.

[4] Li Xuefang. Machine-learning-based computer network image-recognition system [J]. Information Technology and Informatization, 2022(8): 206-209.

[5] Dong Wenxuan, Liang Hongtao, Liu Guozhu, Hu Qiang, Yu Xu. Survey on deep convolutional applications in object detection algorithms[J]. Computer Science and Exploration, 2021, 15(9): 1591-1604.

[6] Xie Fu, Zhu Dingju. Survey of deep-learning object-detection methods [J]. Computer Systems & Applications, 2022, 31(2): 1-12. http://www.c-s-a.org.cn/1003-3254/8303.html.

[7] Li Kecen, Wang Xiaogiang, Lin Hao, Li Leixiao, Yang Yanyan, Meng Chuang, Gao Jing. Survey of single-stage small-object detection methods in deep learning [J]. *Computer Science and Exploration*, 2023, 17(2): 171-188.

[8] Li Jun, Ding Binbin, Shi Weijuan, Yang Lin. Improved YOLOv11 for UAV aerial-image detection [J/OL]. Electronic Measurement Technology. https://link.cnki.net/urlid/11.2175.TN.20250522.1718.014.

[9] Fang Hongsu, Chang Cheng, Shi Xinyu, Xiong Runlian, He Mentao. Research on the application of YOLO object-detection algorithms in autonomous-driving scenarios [J]. Intelligent and Connected Vehicles, 2022, 4(3): 45-54.