Theory and Practice of Science and Technology
2025, VOL. 6, NO. 9, 136-140
DOI: 10.47297/taposatWSP2633-456927.20250609

A License Plate Recognition System for Highway Emergency Lane Monitoring Using YOLOv11-pose and PaddleOCR

Xiaonan Liang*, Shu Li, Chenyan Fan, Xinyue Zhang

School of Science, Xi'an Shiyou University, Xi'an, Shaanxi, 710065, China

ABSTRACT
Emergency lanes are critical for ensuring timely access of rescue vehicles during traffic congestion, significantly reducing accident fatalities when kept unobstructed. This study proposes a robust license plate recognition (LPR) system for monitoring illegal occupation of highway emergency lanes. The method integrates YOLOv11-pose for precise four-corner detection and PaddleOCR (PP-OCRv5) for high-speed character recognition, achieving an average recognition time of less than 10 milliseconds per plate. The models are unified via ONNX for efficient inference and evaluated on real-world surveillance videos from the Changshen Expressway. Experimental results demonstrate high accuracy and strong robustness, achieving 98.79% recognition accuracy and reliable performance across diverse scenarios, including new energy vehicle plates and long-distance captures. The system effectively supports automated traffic violation detection.

KEYWORDS
YOLO, PaddleOCR; License plate recognition; Highway emergency lane monitoring

1 Introduction

With the expansion of highway networks and the continuous growth of vehicle ownership, traffic accident rates have risen significantly. To enhance road safety and emergency response efficiency, emergency lanes are incorporated into highway designs to ensure unimpeded access for rescue vehicles during congestion. Studies show that timely medical intervention within 20 minutes of an accident can reduce mortality by over 50% [1]. However, illegal occupation of emergency lanes by drivers severely hampers rescue operations, increases the risk of fatalities, exacerbates traffic congestion, disrupts traffic order, and may lead to secondary accidents. To improve enforcement, surveillance systems are widely deployed on highways to record violations in different countries [2]. Traditional manual monitoring is costly and inefficient. Nevertheless, surveillance systems often suffer from degraded image quality under challenging conditions—such as adverse weather, variable lighting, and road disturbances—reducing the accuracy of license plate recognition (LPR) and speed estimation, thereby limiting traffic management effectiveness. To tackle these challenges, existing research primarily adopts the deep learning-based method, which aims to improve the accuracy of LPR and the reliability of speed estimation to enhance intelligent transportation systems.

Since the early 1990s, neural networks have been applied to character recognition. LeNet-5, one of the first deep convolutional neural networks (CNNs), was developed for handwritten digit recognition in postal services and laid the foundation for LPR [3]. Silva et al. proposed a novel CNN-based automated license plate recognition (ALPR) system that simultaneously detects and rectifies multiple distorted license plates in unconstrained scenarios [4]. Björklund et al. demonstrated that a CNN-based LPR system trained on synthetically generated images with controlled variance in key visual factors can outperform systems trained on real data, achieving superior performance across multiple real-world datasets [5]. With the advancement of deep learning, researchers have developed end-to-end LPR systems integrating detection, segmentation, and recognition. CNNs have since become central to both plate detection and character recognition. YOLO (You Only Look Once), known for its real-time detection speed, has also been adopted. Agarwal et al. proposed a YOLO World-based ALPR system that achieved high accuracy and real-time performance across multiple datasets, outperforming prior works and commercial systems [6]. Putluru et al. developed a privacy-aware, edge-deployable license plate recognition system using YOLOv8 and OCR that achieved over 95% mAP and 90% OCR accuracy on multi-source datasets, enabling real-time, compliant intelligent transportation applications [7].

Deep learning approaches, particularly CNN and YOLO-based models, offer superior accuracy and adaptability, enabling more reliable and scalable solutions for intelligent traffic monitoring. Therefore, this paper proposes a robust LPR system using YOLOv11-pose for precise detection with four-corner localization and PaddleOCR for accurate character recognition. The models are integrated and deployed via ONNX for efficient inference, and evaluated on real-world highway emergency lane surveillance videos, demonstrating high accuracy and reliability under challenging conditions.

* **Corresponding Author:** Xiaonan Liang, xiaonan_liang@139.com

2 Method

2.1 Experiment platform

The experiments in this study were conducted using a desktop system equipped with a 13th Gen Intel(R) Core(TM) i5-13400F CPU (2.50 GHz) and an NVIDIA GeForce RTX 3060 Ti GPU for model training and evaluation. The proposed YOLOv11-Pose and PaddleOCR framework was implemented in Python 3.11, leveraging the PyTorch 2.8.0 deep learning library. To enhance training efficiency and model performance, the YOLOv11-pose component was initialized with a pre-trained model, taking advantage of its strong generalization capabilities.

2.2 Dataset

The Chinese City Parking Dataset (CCPD) is a large-scale, highly diverse dataset of vehicle license plates, developed to mitigate limitations observed in existing datasets concerning scale and scene variation [8]. It comprises more than 250,000 unique vehicle images collected across varying time periods, weather conditions, and illumination settings. Each image is accompanied by comprehensive annotations, including the license plate number, bounding box coordinates, four corner points, as well as horizontal and vertical rotation angles. Additionally, CCPD offers specialized subsets—such as CCPD-base and CCPD-Challenge—to facilitate robust evaluation of license plate detection and recognition models under complex real-world scenarios. It is noteworthy that CCPD2019 contains license plates from conventional fuel vehicles, while CCPD2020 introduces green license plates assigned to new energy vehicles, which also exhibit an increased character count. The highway emergency lane video dataset comprises traffic surveillance footage collected from four monitoring locations along the Changshen Expressway in China. The 1280×720 resolution video data serves as the input to the YOLO-pose model integrated with PaddleOCR. The output of the model consists of the corresponding video frames automatically annotated with license plate information.

2.3 YOLO

YOLOv11 builds upon its predecessors by introducing significant enhancements that bolster its performance across a range of computer vision tasks, including object detection, instance segmentation, pose estimation, and oriented object detection [9]. Figure. 1 illustrates the overall network architecture of the original YOLOv11 model, which comprises three main components: the backbone, the beck and the head.

At the core of YOLOv11's architecture lies a sophisticated backbone designed for efficient feature extraction. This backbone incorporates the novel C3k2 block, a computationally efficient variant of the Cross Stage Partial (CSP) Bottleneck. By utilizing two smaller convolutions instead of a single large one, the C3k2 block reduces computational overhead while maintaining performance. Additionally, the backbone integrates the Spatial Pyramid Pooling-Fast (SPPF) block and the innovative Cross Stage Partial with Spatial Attention (C2PSA) block. The C2PSA block enhances spatial attention, enabling the model to focus more effectively on key regions within images, thereby improving detection accuracy for objects of varying sizes and positions.

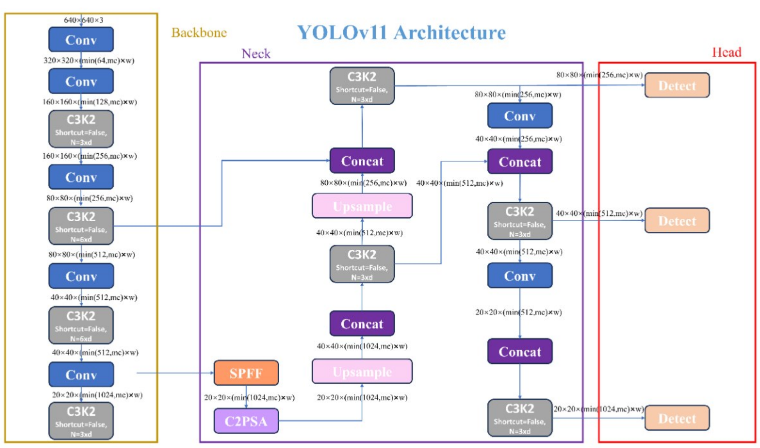

Figure. 1 YOLOv11 architecture and its three main components

The neck component of YOLOv11 is responsible for aggregating and enhancing features from different scales. It leverages the C3k2 block to streamline the feature aggregation process, resulting in faster and more efficient performance. This is further complemented by the C2PSA module, which emphasizes spatial attention, allowing the model to concentrate on critical areas within images. This attention mechanism is particularly beneficial for detecting smaller or partially occluded objects, a common challenge in object detection tasks.

The head of YOLOv11 is designed to generate precise predictions for object localization and classification, making it particularly suitable for license plate detection tasks. It employs multiple C3k2 blocks to efficiently process and refine feature maps at different depths. The flexibility of the C3k2 block is enhanced by customizable kernel sizes, which are particularly useful for extracting detailed features from images. Furthermore, the head incorporates CBS (Convolution-BatchNorm-SiLU) layers after the C3k2 blocks. These layers refine feature maps by extracting relevant features, stabilizing data flow through batch normalization, and introducing non-linearity via the Sigmoid Linear Unit (SiLU) activation function. The final convolutional layers and Detect layer consolidate predictions, outputting bounding box coordinates, objectness scores, and class scores for detected objects.

YOLOv11 also demonstrates remarkable versatility across different model sizes, catering to diverse application needs. The YOLOv11m variant attains higher mean Average Precision (mAP) on the COCO dataset compared to YOLOv8m, despite having 22% fewer parameters, highlighting enhanced computational efficiency without sacrificing detection accuracy [10]. These optimizations ensure that YOLOv11 is well-suited for deployment in resource-constrained environments.

2.4 PaddleOCR

PaddleOCR is an open-source Optical Character Recognition (OCR) toolkit used in this study for recognizing characters on Chinese license plates [11]. The task presents unique challenges due to the fixed layout of license plates, limited character variations, and potential environmental issues affecting image quality. PaddleOCR's PP-OCRv5 model is particularly well-suited for this application, offering a lightweight yet powerful solution.

PP-OCRv5 comprises four key stages: image preprocessing, text detection, text line orientation classification, and text recognition. Image preprocessing is crucial for addressing environmental factors such as lighting and angle distortions. Techniques like denoising and contrast enhancement are applied to improve image quality, while perspective correction helps normalize the license plate's orientation. Text detection utilizes the PP-HGNetV2 backbone network, which is sensitive to small, closely spaced characters typical of license plates. This advanced detector identifies the license plate region and the individual characters within it, even when spacing is minimal. Text line orientation classification ensures that the detected characters are correctly oriented for accurate recognition, which is particularly important given the fixed layout of license plates. The text recognition component employs an attention-based mechanism in its Narrowing Residual Text Recognition (NRTR) branch, adept at decoding characters despite potential noise or distortion in the image. NRTR demonstrates a significant advantage in enhancing model performance for complex text recognition scenarios by effectively handling multiple languages and diverse writing formats, such as the combination of Chinese characters and alphanumeric characters in license plates. This is vital for recognizing license plate characters that may be affected by environmental conditions or image degradation.

The PP-OCRv5 model's advantages for license plate character recognition lie in its ability to handle the specific challenges of this task. Its lightweight nature allows for efficient processing, while its high accuracy ensures reliable character recognition. By leveraging PaddleOCR's capabilities, this study aims to achieve accurate and efficient Chinese license plate character recognition.

3 Results and discussion

In the YOLOv11 model, pose estimation is employed to leverage the four corner bounding box annotations in the CCPD dataset. Training and validation results over 30 epochs, based on 263,359 training images and 71,006 validation images, demonstrate strong convergence and performance in license plate detection and pose estimation, despite initial instability observed in the validation metrics. All training losses, including pose loss, classification loss (cls_loss), and distribution focal loss (dfl_loss), decreased monotonically, indicating effective optimization, as shown in Figure. 2.

Training pose loss fell from 0.534 to 0.018 and classification loss from 0.569 to 0.244. Validation losses peaked early (pose loss: 1.359; classification loss: 6.309) but later declined to 0.018 and 0.225, respectively, showing improved generalization. Precision and recall for pose estimation exceeded 0.998 from epoch 7 onward, reflecting the model's high accuracy and reliability in localizing license plates. The mAP50 metric maintained a stable level above 0.994 from epoch 8 to 30, while mAP50-95 converged to approximately 0.994 after epoch 9, underscoring the model's precision across a range of IoU thresholds and its ability to produce well-calibrated bounding boxes. The model demonstrates highly effective learning behavior, achieving excellent convergence and strong generalization on the CCPD dataset.

Additionally, we continue to employ CCPD for model training and validation within the PaddleOCR framework. The dataset is partitioned into a training set containing 203,671 license plate images and a validation set with 33,657 images. A predefined dictionary encompassing all relevant Chinese characters, English letters and digits used in license plates was integrated to restrict the recognition scope, thereby significantly reducing the risk of erroneous character predictions. The

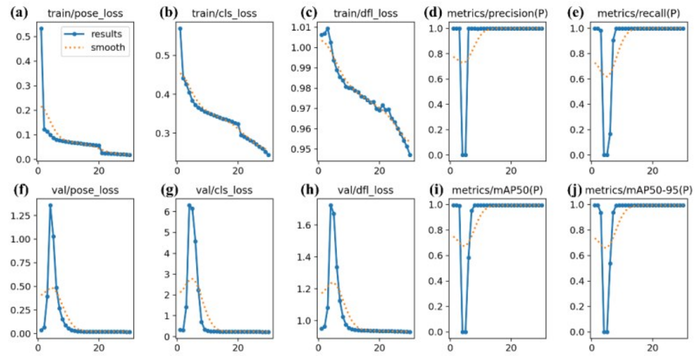

Figure. 2 Evaluation metrics for YOLO-pose on the training and validation sets

model was trained for 100 epochs, and the best-performing checkpoint achieved an average recognition accuracy of 98.79%, with a recognition time of less than 10 milliseconds.

Figure. 3 License plate recognition of expressway emergency lane monitoring

The trained YOLOv11-pose model for detection and PaddleOCR model for recognition were converted and integrated using the ONNX runtime to ensure efficient deployment. The combined pipeline was applied to highway surveillance video data for joint license plate detection and recognition via frame-by-frame processing. Representative results are illustrated in the accompanying Figure. 3(a) shows a new energy vehicle license plate, which is characterized by a green background and eight characters. Our model correctly localizes and recognizes the plate under this condition. Figure. 3(b) depicts a vehicle captured from a frontal view by an emergency lane camera. The system accurately identifies the license plate without performance degradation due to perspective variation, demonstrating high robustness. Figure. 3(c) illustrates a vehicle lane-changing into the emergency lane at a considerable distance, resulting in partial blur and information loss in the video. Despite these challenges, our model correctly detects and recognizes the plate while effectively avoiding false positives from non-emergency lane vehicles. Figure. 3(d) presents a car-following scenario common in emergency lanes, where narrow field-of-view and high speed often lead to occlusion and missed detections. Our model successfully detects and tracks multiple vehicles simultaneously through temporal frame analysis, substantially reducing omission errors. The experimental outcomes confirm that the proposed system achieves high accuracy and robustness under various challenging real-world conditions, making it suitable for automated LPR in dynamic highway environments.

4 Conclusion

This paper presents an effective license plate recognition system tailored for intelligent monitoring of emergency lane violations on highways. By combining YOLOv11-pose for accurate four-corner detection and PaddleOCR for high-precision character recognition, the proposed framework achieves robust performance under real-world challenges such as motion blur, partial occlusion, perspective distortion, and varying illumination. Training on the large-scale CCPD dataset ensures strong generalization, with the detection model achieving mAP50-95 of 0.994 and the recognition model reaching 98.79% accuracy. The integration via ONNX enables efficient deployment for real-time video analysis. Experimental results on highway surveillance footage confirm the system's capability to accurately detect and recognize license plates from both conventional and new energy vehicles, even in complex traffic scenarios. This approach significantly enhances the automation and reliability of traffic enforcement systems, offering a scalable solution for intelligent transportation management.

CRediT authorship contribution statement

Xiaonan Liang: Writing-original draft, Validation, Methodology, Investigation, Conceptualization, Project administration. Shu Li: Writing - review & editing, Supervision, Resources. Chenyan Fan: Validation, Methodology, Investigation. Xinyue Zhang: Writing-review & editing, Resources.

References

[1] Swaroop, V., Ahmed, N. (2023) Prehospital care in road traffic accident cases. Journal of Surgical Specialties and Rural Practice, 4: 147-152.

[2] Alobaidallah, A. M., Alqahtany, A., & Maniruzzaman, K. M. (2025). Safety Effectiveness of Automated Traffic Enforcement Systems: A Critical Analysis of Existing Challenges and Solutions. Future Transportation, 5(1), 25.

[3] Lecun, Y., Bottou, L., Bengio, Y., et al. (1998) Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86: 2278-2324.

[4] Silva, S.M., Jung, C.R. (2018) License plate detection and recognition in unconstrained scenarios. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), Computer Vision - ECCV 2018. Springer International Publishing, Cham. pp. 593-609.

[5] Björklund, T., Fiandrotti, A., Annarumma, M., et al. (2019) Robust license plate recognition using neural networks trained on synthetic images. Pattern Recognition, 93: 134-146.

[6] Agarwal, V., Bansal, G. (2024) Automatic number plate detection and recognition using YOLO world. Computers and Electrical Engineering, 120: 109646.

[7] Putluru, A.R., Pannuru, V.R., AshokKumar, C., et al. (2025) An optimized YOLO-based license plate recognition system with integrated privacy safeguards. In: 2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS). pp. 1053-1060.

[8] Xu, Z., Yang, W., Meng, A., et al. (2018) Towards end-to-end license plate detection and recognition: A large dataset and baseline. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (Eds.), Computer Vision - ECCV 2018. Springer International Publishing, Cham. pp. 261-277.

[9] Khanam, R., Hussain, M. (2024) YOLOv11: An overview of the key architectural enhancements. https://api. semanticscholar. org/CorpusID: 273532028.

[10] Ultralytics. (2024) Ultralytics YOLOv11. https://docs.ultralytics.com/models/yolo11/.

[11] Cui, C., Sun, T., Lin, M., Gao, T., et al. (2025). PaddleOCR 3.0 technical report. arXiv. https://arxiv.org/abs/2507.05595.