A Comparative Study of Multi-Model Lane Detection Methods Based on a Unified Evaluation Framework

Shaopu Zou[a]
School of Computer and Artificial Intelligence, Beijing Technology and Business University, Beijing, China

Keywords: Lane Detection, SCNN, PINet, LaneATT, CULane.

Abstract: Lane detection, as a crucial task in autonomous driving systems, faces the dual challenges of robustness and accuracy in complex road environments. This study conducts a comparative analysis of three representative deep learning models—Spatial Convolutional Neural Network (SCNN), Point Instance Network (PINet), and LaneATT. The models are reproduced and evaluated under consistent input settings using the unified CUHK Lane Dataset (CULane) and a standardized evaluation tool. Both quantitative metrics and visualized results are utilized to assess each model's detection performance across diverse driving scenarios. Experimental results demonstrate that LaneATT achieves the best overall performance, particularly exhibiting strong robustness in challenging conditions such as nighttime and shadowed environments. PINet excels in curved lane detection, while SCNN maintains stable outputs in standard road settings. This study establishes a unified evaluation framework for horizontal comparisons of lane detection models, providing a systematic basis for performance assessment under standardized conditions. The proposed framework contributes to the advancement of algorithmic benchmarking and offers methodological guidance for subsequent research on model optimization and real-world deployment.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

1 INTRODUCTION

With the rapid development of autonomous driving technology, lane detection has become a crucial component of the vehicle perception system and has attracted increasing attention (Singal et al., 2023). Accurate lane detection not only helps vehicles maintain correct trajectories on the road but also plays a vital role in Lane Keeping Assist Systems (LKAS) and Advanced Driver Assistance Systems (ADAS) (Waykole et al., 2021; Tian et al., 2021). However, complex road environments—such as varying lighting conditions, occlusions, and curved roads—still pose significant challenges to reliable lane detection (Sultana et al., 2023).

Traditional lane detection approaches primarily rely on image processing techniques such as edge detection and Hough transforms. While effective under ideal conditions, these methods are vulnerable to disturbances in complex environments, often resulting in reduced detection accuracy (Huang & Liu, 2021). In recent years, the emergence of deep learning has introduced significant breakthroughs in this field (Zakaria et al., 2023). Models based on Convolutional Neural Networks (CNNs) can automatically learn visual features from data, thus enhancing both robustness and accuracy in lane detection tasks (Kortli et al., 2022).

Among the various deep learning-based lane detection methods, the Spatial Convolutional Neural Network (SCNN) introduces spatial convolution modules to enable information propagation across feature maps, thereby improving the structural modeling of lane lines (Pan et al., 2018). The Point Instance Network (PINet) treats lane detection as a keypoint estimation and instance segmentation task, transforming it into a clustering problem of point instances, which enhances performance in detecting complex lane geometries (Ko et al., 2021). The LaneATT combines anchor-based mechanisms with attention modules, directly regressing lane parameters to achieve efficient lane detection (Tabelini et al., 2021).

Despite the strong performance demonstrated by these models in their respective studies, differences in

[a] https://orcid.org/0009-0007-6134-0737

Zou, S.
A Comparative Study of Multi-Model Lane Detection Methods Based on a Unified Evaluation Framework.
DOI: 10.5220/0014322600004718
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 2nd International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2025)*, pages 123-129
ISBN: 978-989-758-792-4
Proceedings Copyright © 2025 by SCITEPRESS – Science and Technology Publications, Lda.

experimental settings and evaluation criteria make it difficult to conduct fair and direct comparisons under the same conditions. To address this issue, this paper reproduces and evaluates SCNN, PINet, and LaneATT within a unified experimental environment. All models are tested using the same dataset–CUHK Lane Dataset (CULane) (Pan et al., 2018) – and assessed through consistent evaluation tools and metrics to facilitate a systematic comparison of their detection performance across various driving scenarios.

The remainder of this paper is organized as follows:

Chapter 2 provides a detailed introduction to the dataset used and the technical principles and architectures of the three lane detection models.

Chapter 3 elaborates on the experimental setup and evaluation metrics and presents a comparative performance analysis and visualizations across different scenarios.

Chapter 4 summarizes the research findings, discusses the limitations of current methods, and outlines potential directions for future work.

2 DATASET AND METHODS

2.1 Dataset

This study adopts the CULane dataset, a large-scale open dataset specifically designed for lane detection tasks, provided by the Multimedia Laboratory of The Chinese University of Hong Kong. The CULane dataset covers a wide range of real-world driving scenarios, aiming to provide a comprehensive and challenging evaluation environment for lane detection models.

The dataset consists of approximately 133,235 images, including around 88,880 images for training, 9,675 for validation, and 34,680 for testing. All images are extracted from real urban driving videos, ensuring strong practical relevance for real-world applications.

Each image in the dataset is manually annotated with lane markings using cubic spline curves. Even when lane lines are occluded or not clearly visible, annotations are inferred based on contextual information (as shown in Figure 1). This annotation method accurately captures the position and shape of lane lines and supports a variety of downstream tasks such as regression, detection, and segmentation.

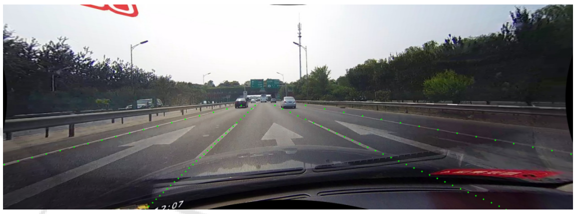

Figure 1. Example of point-wise lane annotation (Picture credit: Original)

The test set of the CULane dataset is divided into nine representative scenarios (some of which are illustrated in Figure 2): normal, crowded, dazzle light, shadow, no lane, arrow, curve, crossroad, and night. Each subset targets specific environmental challenges, allowing for a thorough evaluation of model generalization under diverse driving conditions.

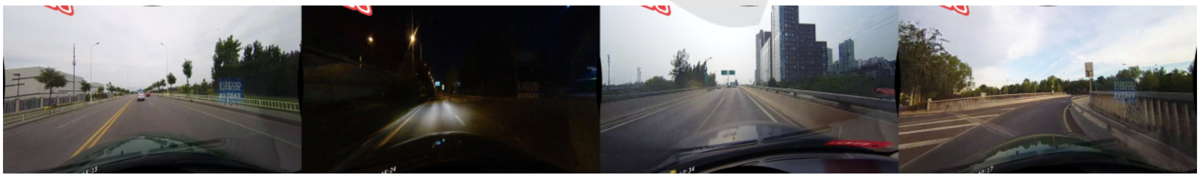

Figure 2. Examples of different road scenarios (Picture credit: Original)

Images in the CULane dataset have a resolution of 1640 × 590 pixels, preserving fine-grained visual details that facilitate accurate recognition of distant or thin lane markings.

In this study, all models are evaluated using the official CULane data split, with identical training and testing configurations. Inference is conducted at a fixed input resolution, and a unified evaluation tool is used to assess performance. The focus of this work is on model inference and evaluation on the standard test set without any retraining or fine-tuning, ensuring fairness and reproducibility in the comparison.

Given its large scale, diverse scenarios, and precise annotations, the CULane dataset serves as an ideal benchmark for lane detection experiments and provides a solid foundation for comparing algorithm performance in complex real-world environments.

2.2 Methods

To comprehensively compare the performance of different lane detection methods in complex road scenarios, this study selects three representative deep learning models−SCNN, PINet, and LaneATT−for experimental evaluation. These models differ in network architecture and detection strategies, reflecting the strengths and limitations of current mainstream approaches from various perspectives.

2.2.1 SCNN Model

The SCNN model introduces spatial convolution units into a traditional convolutional neural network, based on an adapted VGG16 backbone, enabling serialized information propagation along the spatial dimensions of the feature maps. Specifically, SCNN propagates features in horizontal and vertical directions, effectively modeling the elongated and continuous structure of lane lines.

The overall architecture of SCNN consists of a backbone network, spatial convolution modules, and a prediction head. The input image is first processed through a series of convolutional layers to extract base features. Then, spatial message passing is performed iteratively in four directions−up, down, left, and right−to enhance the representation of lane continuity in the feature maps. Finally, fully connected layers predict lane point positions along each scanline to produce the final detection results.

This approach maintains computational efficiency while significantly improving detection accuracy in complex environments, making it particularly suitable for extracting continuous lanes under occlusions or road wear.

2.2.2 PINet Model

The PINet model adopts an instance segmentation approach, transforming lane detection into a point-wise instance prediction problem. It uses a stacked hourglass network to extract multi-scale features and then predicts each pixel's lane instance affiliation and its spatial offset.

The architecture consists of three main components: a feature encoder, an instance embedding branch, and an offset regression branch. The encoder, typically a stacked Hourglass network, extracts hierarchical features from the input image. The embedding branch generates a low-dimensional vector for each pixel, allowing pixels belonging to the same lane to be clustered. The offset branch predicts each pixel's displacement relative to the centerline of its corresponding lane. A clustering algorithm (e.g., Mean Shift) is then used to group lane points into complete lane instances.

This method is highly adaptable to varying numbers and shapes of lane lines, showing strong robustness in dense or sharply curved road conditions.

2.2.3 LaneATT Model

The LaneATT model employs an anchor-based regression strategy combined with attention mechanisms, discarding traditional pixel-wise segmentation in favor of directly regressing lane parameters such as quadratic coefficients and endpoints. LaneATT typically utilizes ResNet-34 as its backbone and applies a Transformer encoder to capture global contextual features. Anchors are placed at predefined positions for efficient lane parameter regression.

Its architecture includes a feature extraction backbone (typically ResNet), a Transformer encoder, and a lane regression head. The backbone encodes the input image into intermediate features, which are then globally modeled by the Transformer module to capture long-range spatial dependencies. Finally, the model performs lane regression at anchor locations to produce fitted lane curves.

LaneATT achieves high detection accuracy while significantly improving inference speed, making it well-suited for real-time or large-scale deployment scenarios.

2.2.4 Unified Experimental Setup

To ensure fairness in model comparisons, all experiments in this study adopt the official training and testing splits of the CULane dataset. Input images are resized to a unified resolution across all models. Inference is conducted under the same hardware environment to ensure comparability in speed and resource consumption. Results are evaluated using consistent metrics and identical evaluation scripts.

To provide a clear overview of the differences in model architecture and detection strategy, the key characteristics of the three models are summarized in Table 1. This study focuses solely on the inference performance of the models without any retraining or modification of the original implementations.

Table 1. Comparison of the architectures of the three lane detection models

Model	Backbone	Core Module	Output Type	Characteristics
SCNN	VGG16	Spatial convolution unit	Scanline-wise point prediction	Strong continuity, suitable for occluded or worn lane markings
PINet	Hourglass	Instance embedding + offset regression	Clustering of points into lanes	Flexible in handling variable number and shape of lane lines
LaneATT	ResNet	Transformer encoder + anchor-based regression	Curve parameter regression	Fast inference, suitable for real-time applications

3 EXPERIMENTS

3.1 Evaluation Metrics

To comprehensively evaluate the performance of each lane detection model, this study adopts the official evaluation metrics provided by the CULane benchmark. The evaluation tool matches predicted lane lines with ground truth annotations based on the spatial overlap and calculate Precision, Recall, and F-measure as the core performance indicators.

3.1.1 Matching Method

The evaluation tool first fits both predicted and annotated lane lines using spline interpolation and renders their respective width masks in the image space. It then calculates the Intersection over Union (IoU) between the two lines, as illustrated in Equation (1).

$$\mathrm{IoU} = \frac{|\mathrm{Predict} \cap \mathrm{GT}|}{|\mathrm{Predict} \cup \mathrm{GT}|} \tag{1}$$

A predicted lane line is considered a successful match if its IoU with a ground truth lane line exceeds a threshold of 0.5. To ensure optimal matching in multi-lane scenarios, the tool applies the Hungarian Algorithm to compute one-to-one pairings based on the IoU similarity matrix. This process yields global statistics for True Positives (TP), False Positives (FP), and False Negatives (FN) across the entire image.

3.1.2 Metric Definitions

Based on the above matching process, the final evaluation metrics are defined as follows:

Precision: The proportion of correctly predicted lane lines out of all lane line predictions made by the model.

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

Recall: The proportion of ground truth lane lines that are successfully detected by the model.

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{3}$$

F-measure: The harmonic mean of Precision and Recall, providing a balanced assessment of both accuracy and completeness in lane detection.

$$\mathrm{F1} = \frac{2 \times \mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{4}$$

3.2 Experimental Results and Analysis

To systematically compare the lane detection performance of SCNN, PINet, and LaneATT under various driving scenarios, this study evaluates each model on the nine representative scenes defined in the CULane dataset. For each scenario, Precision, Recall, and F1-measure are calculated. Among them, F1-measure is considered the primary performance metric in this study. Table 2 presents the comparative results of the three models across all scenarios.

Table 2. F1-measure, Precision, and Recall comparison of three models across different CULane scenarios

Scenario	SCNN F1	PINet F1	LaneATT F1	SCNN Precision	PINet Precision	LaneATT Precision	SCNN Recall	PINet Recall	LaneATT Recall
Normal	0.9049	0.8985	0.9218	0.9070	0.9235	0.9384	0.9029	0.8749	0.9059
Crowd	0.6803	0.7184	0.7500	0.6922	0.8157	0.8100	0.6688	0.6418	0.6983
Hlight	0.6332	0.6458	0.6669	0.6482	0.7767	0.7455	0.6189	0.5527	0.6033
Shadow	0.6352	0.6683	0.7795	0.6340	0.8223	0.8216	0.6364	0.5629	0.7416

No line	0.4351	0.4769	0.4936	0.4538	0.7399	0.6611	0.4179	0.3518	0.3939
Arrow	0.8457	0.8350	0.8840	0.8556	0.9069	0.9226	0.8360	0.7738	0.8484
Curve	0.6159	0.6361	0.6767	0.6649	0.7679	0.7950	0.5736	0.5429	0.5890
Cross	0	0	0	0	0	0	-1	-1	-1
Night	0.6559	0.6605	0.7055	0.6590	0.8301	0.7996	0.6528	0.5484	0.6311

As shown in Table 2, the three models exhibit varying detection performance across different driving scenarios. Overall, LaneATT consistently achieves the highest F1-measure across all scenarios, demonstrating superior robustness, particularly in challenging conditions such as crowded, shadowed, and nighttime environments. In the Normal scenario, all three models perform well, with LaneATT slightly outperforming SCNN and PINet.

In more complex settings−Crowd, Shadow, and Night−LaneATT shows a substantial lead in F1 score compared to the other two models, highlighting its strong adaptability. While performance drops under extreme conditions such as Hlight (dazzling light) and No line (no visible lane markings), LaneATT still maintains a performance edge.

Notably, in the Curve scenario, both PINet and LaneATT demonstrate good adaptability, likely due to their ability to handle non-linear lane structures. In the Arrow scenario, SCNN and LaneATT achieve higher detection accuracy, with LaneATT attaining an F1-measure of 0.884.

It is important to note that although the CULane dataset includes the Cross (intersection) scenario, this subset does not provide ground truth lane annotations. As such, the evaluation tool is only used to test model robustness in this scene rather than actual detection performance. Consequently, all models score zero in this scenario−not due to poor performance, but due to the lack of ground truth, and this subset is excluded from the core performance comparison.

In summary, the overall model performance on the CULane test set can be ranked as: LaneATT > PINet > SCNN. LaneATT leverages anchor-based regression and Transformer-driven global modeling to achieve superior detection accuracy and robustness across diverse scenarios. PINet, with its point-instance detection strategy, excels in handling curved or lane-dense environments. SCNN, while slightly lower in overall accuracy, demonstrates stable continuity in lane detection, particularly under occlusion or degradation.

3.3 Visualization of Experimental Results

To further compare the performance of the three models across different scenarios, representative test image samples were selected, and the prediction results of SCNN, PINet, and LaneATT were visualized on the same images. Figure 3 presents a comparison between the predicted lane lines and the ground truth, where green lines represent the annotated lanes and red lines indicate the predicted outputs from each model.

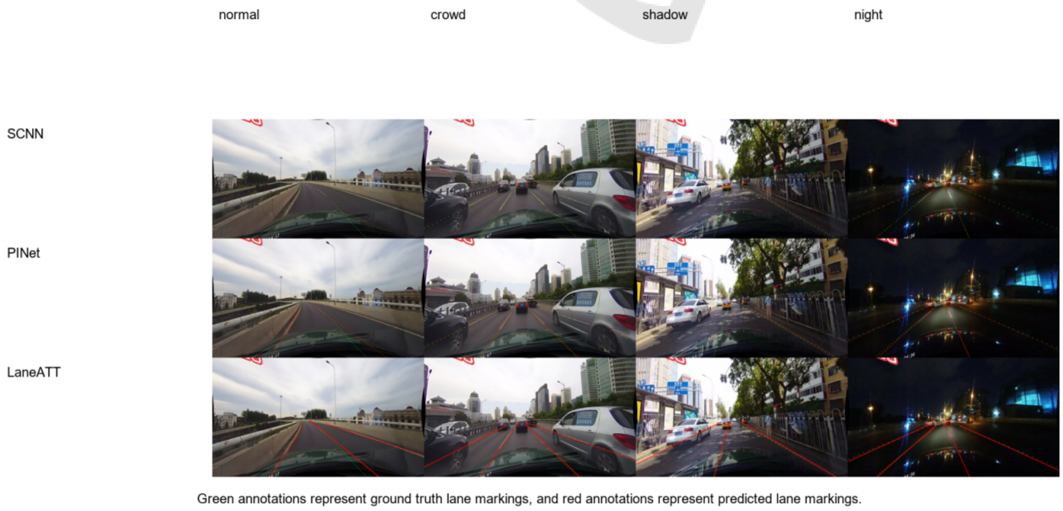

Figure 3. Visualization comparison of lane detection by the three models in typical scenarios (Picture credit: Original)

As illustrated in Figure 3, the performance of different models varies significantly under complex scenarios. In the Normal scenario, all three models are able to accurately detect lane lines, with predictions closely matching the ground truth. LaneATT, in particular, demonstrates smoother fitting at lane curvature points, reflecting superior detail recovery capabilities. In the Crowd scenario, all three models successfully detect the primary lane lines with minimal prediction error, showcasing good robustness.

In contrast, in the Shadow scenario, where lighting conditions change drastically, the models show noticeable differences. SCNN exhibits significant deviations and broken lines in its predictions, leading to reduced accuracy. PINet detects only two lane lines, but they align well with the ground truth. LaneATT successfully identifies all lane lines with predictions almost fully overlapping the annotations, demonstrating the best overall performance in this setting.

Under the Night scenario, both LaneATT and PINet maintain high detection accuracy, whereas SCNN shows missed detections under low-light conditions, failing to identify the rightmost lane line and exhibiting a notable performance drop.

In conclusion, the visual results further support the quantitative findings presented in Section 3.2. LaneATT demonstrates stronger robustness and generalization in complex scenarios, with more stable and accurate predictions. PINet maintains high localization accuracy in curved or partially occluded environments. SCNN, while stable in scenarios with clear lane continuity, exhibits limited performance under strong environmental interference.

4 CONCLUSIONS

This study focused on the task of lane detection by selecting three representative deep learning models– SCNN, PINet, and LaneATT – for systematic reproduction and performance comparison under a unified dataset (CULane) and evaluation framework. By standardizing the input-output settings, evaluation metrics, and visualization analysis, the aim was to explore the detection effectiveness of these models under various driving scenarios and provide an empirical foundation for future research.

Experimental results reveal significant differences in overall performance and detailed behavior among the three models. LaneATT achieved the highest F1 scores across all scenarios, demonstrating superior robustness and generalization capabilities, particularly in complex environments such as nighttime, crowded traffic, and variable lighting conditions. PINet performed well in handling curved roads and lane-dense scenes, making it suitable for recognizing structurally complex lane patterns. While SCNN maintained stable detection in standard scenarios with good lane continuity, its performance declined under more challenging conditions. The visual analyses further confirmed these quantitative findings, showcasing the prediction differences on specific test images.

Despite the comprehensive comparative analysis conducted in this study, some limitations remain. First, the evaluation focused solely on the inference stage without including the full training process. Second, only the CULane dataset was used, lacking cross-dataset generalization analysis. Third, practical deployment factors such as detection speed and resource consumption were not addressed.

Future research can be extended in several directions: further optimizing model architectures to improve adaptability in complex scenes; expanding evaluation to include diverse urban environments and varying weather conditions; incorporating lightweight network designs to enhance inference efficiency and promote real-world deployment in autonomous driving systems; and exploring multi-task learning approaches to integrate lane detection with other perception tasks.

This study holds practical relevance and reference value. On the one hand, reproducing and comparing typical models within a unified evaluation framework, clarifies the applicability and strengths of current mainstream lane detection methods under different scenarios, providing a basis for industrial model selection. On the other hand, the standardized comparison procedure and multi-perspective visualization analysis proposed in this work serve as an experimental paradigm and evaluation reference for future model improvements and academic studies.

REFERENCES

Singal, G., Singhal, H., Kushwaha, R., Veeramsetty, V., Badal, T., & Lamba, S. (2023). RoadWay: lane detection for autonomous driving vehicles via deep learning. Multimedia Tools and Applications, 82(4), 4965-4978.

Waykole, S., Shiwakoti, N., & Stasinopoulos, P. (2021). Review on Lane Detection and Tracking Algorithms of Advanced Driver Assistance System. Sustainability, 13(20), 11417.

Tian, J., Liu, S., Zhong, X., & Zeng, J. (2021). LSD-based adaptive lane detection and tracking for ADAS in structured road environment. Soft Computing, 25(7), 5709-5722.

Sultana, S., Ahmed, B., Paul, M., Islam, M. R., & Ahmad, S. (2023). Vision-based robust lane detection and tracking in challenging conditions. IEEE Access, 11, 67938-67955.

Huang, Q., & Liu, J. (2021). Practical limitations of lane detection algorithm based on Hough transform in challenging scenarios. International Journal of Advanced Robotic Systems, 18(2).

Zakaria, N. J., Shapiai, M. I., Abd Ghani, R., Yassin, M. N. M., Ibrahim, M. Z., & Wahid, N. (2023). Lane detection in autonomous vehicles: A systematic review. IEEE access, 11, 3729-3765.

Kortli, Y., Gabsi, S., Voon, L. F. L. Y., Jridi, M., Merzougui, M., & Atri, M. (2022). Deep embedded hybrid CNN‑LSTM network for lane detection on NVIDIA Jetson Xavier NX. Knowledge-based systems, 240, 107941.

Pan, X., Shi, J., Luo, P., Wang, X., & Tang, X. (2018). Spatial as Deep: Spatial CNN for Traffic Scene Understanding. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1). https://doi.org/10.1609/aaai.v32i1.12301

Ko, Y., Lee, Y., Azam, S., Munir, F., Jeon, M., & Pedrycz, W. (2021). Key points estimation and point instance segmentation approach for lane detection. IEEE Transactions on Intelligent Transportation Systems, 23(7), 8949-8958.

Tabelini, L., Berriel, R., Paixao, T. M., Badue, C., De Souza, A. F., & Oliveira-Santos, T. (2021). Keep your eyes on the lane: Real-time attention-guided lane detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 294-302).