International Journal of Advances in Intelligent Informatics
Vol. 12, No. 2, May 2026, pp. 319-332

Enhancing anomaly lane detection accuracy using a feature cross attention in the U-Net architecture

Zahir Zainuddin [a,1], Muhammad Abdillah Rahmat [a,2,*], Elly Warni [a,3], A. Ais Prayogi [a,4]

[a] Departement of Informatics, Hasanuddin University, Makassar, Indonesia
[1] zahir@unhas.ac.id; [2] abdi.rahmat@unhas.ac.id; [3] elly@unhas.ac.id; [4] aisprayogi@unhas.ac.id
* corresponding author

ARTICLE INFO

Article history
Received December 11, 2025
Revised January 14, 2026
Accepted February 4, 2026
Available online May 31, 2025

Keywords
Autonomous driving
Anomaly detection
Semantic Segmentation
U-Net
Feature Cross Attention

ABSTRACT

Lane departure detection is a crucial task in advanced driver assistance systems (ADAS) and autonomous driving, aimed at reducing accidents caused by unintentional road deviation. This study proposes a modified U-Net architecture enhanced with Feature Cross Attention (FCA) to improve lane departure anomaly detection. The objective is to enhance spatial sensitivity and context awareness in segmentation, especially under challenging driving conditions such as occlusions, poor lighting, and distorted lane geometry. The materials used include the publicly available Comma 2k19 LD dataset, comprising 2,000 manually annotated frames extracted from highway driving scenarios. Each frame includes synchronized video and driving telemetry, offering diverse visual conditions. Preprocessing steps include resizing, normalization, and annotation conversion to binary masks. An anomaly is defined as a spatial deviation threshold between predicted and ground-truth lane boundaries. The proposed method incorporates FCA at the bottleneck and decoder levels of the U-Net architecture. Evaluation was performed using Intersection over Union (IoU), Pixel Accuracy, and threshold-based anomaly criteria. The model achieved 99.19% Pixel Accuracy and 98.47% IoU, outperforming the baseline U-Net (97.56% and 97.46%, respectively). Visual results showed improved detection of subtle lane shifts. A confusion matrix generated over 210 validation images demonstrated perfect classification of normal and anomalous cases. These results confirm that FCA integration enhances segmentation precision and anomaly sensitivity. The approach is suitable for real-time deployment in autonomous systems. Future research may focus on temporal integration, lightweight optimization for embedded devices, and the extension of the framework to multi-lane or urban traffic environments.

1. Introduction

Lane departure detection is an important aspect of both autonomous driving and advanced driver assistance systems (ADAS). It helps keep drivers from being distracted and crashing into other cars on the road [1]. As self-driving cars become more common, being able to reliably detect lane departures is becoming more important for keeping roads safe and making sure cars stay stable in real world driving situations [2]. The Enhanced Lane Departure Warning System (ELDWS) is a recent improvement in lane detection that uses advanced image processing and machine learning algorithms to improve detection accuracy under different road and lighting situations [3]. But there are still problems, especially with complicated road shapes, bad weather, and unclear lane markings, which means that detection systems need to be stronger [4].

Despite recent progress, lane departure detection remains challenging due to dynamic road conditions, adverse weather, and occlusions caused by surrounding vehicles or infrastructure [5], [6]. Environmental factors such as glare, reflections, shadows, and day-night transitions further degrade lane visibility and boundary localization [5], [6], [7], [8]. Various strategies, including sub-image feature extraction, enhanced lane verification, and refined edge recognition, have been proposed to address these issues [6], [8]. However, existing methods still struggle to achieve robust and reliable performance in real-world autonomous driving scenarios.

People have utilized traditional U-Net architectures a lot to find lane departures, but they have trouble with complicated situations like lane shifts due to construction or glare from the sun [9]. Recent research has suggested alterations to the U-Net architecture to address these constraints [10], [11], such as the implementation of federated learning in FedLANE to boost privacy and accuracy [12], and the incorporation of directional lane and edge attention maps to improve performance [13]. In addition, using U-Net and YOLO together for object identification has worked well in real-time applications [14]. These advancements indicate that integrating attention processes into the U-Net may improve detection reliability in difficult driving conditions.

One such interesting improvement is the addition of feature cross-attention techniques, which have been demonstrated to greatly improve both semantic segmentation and anomaly detection [15], [16], [17]. In segmentation tasks, feature cross-attention improves edge accuracy and feature fitting, leading to superior intersection over union and pixel accuracy [17]. In anomaly detection, strategies such as early halting and elastic regularization have surpassed traditional methods by tailoring pre-trained features to the target distribution [18]. Also, techniques like GeneralAD, which use Vision Transformers to detect anomalies, demonstrate their effectiveness across different fields [19]. These improvements show that attention-based methods can greatly boost the performance of models for both semantic segmentation and anomaly detection.

This study presents a modified U-Net architecture with feature cross-attention mechanisms designed to improve lane-departure anomaly detection in autonomous driving systems. The suggested design enhances lane recognition accuracy by overcoming the constraints of conventional U-Net models, especially under difficult driving conditions. The distinctive contribution of this study lies in the integration of Feature Cross Attention into a lightweight U-Net framework, specifically designed to enhance lane segmentation robustness under challenging environmental conditions such as glare, occlusion, and lane distortion. Rather than proposing a fundamentally new attention mechanism, this work focuses on adapting and positioning existing attention principles to improve anomaly-sensitive lane segmentation in practical autonomous driving scenarios. This paper makes three contributions: (1) we propose a modified U-Net architecture that includes feature cross attention mechanisms, (2) we compare this improved model to the baseline U-Net model, and (3) we show how well the new architecture works to improve lane departure detection in a variety of real-world situations.

This study adopts a task-driven formulation of anomaly detection rather than unsupervised or generative approaches that implicitly learn normal driving patterns. Lane departure anomalies are defined as interpretable spatial deviations derived from lane segmentation outputs that align with practical ADAS requirements such as explainability, robustness, and real-time performance. While this approach does not aim to replace learning-based anomaly modeling, it provides a controlled, deployment-oriented framework for detecting lane-departure events under challenging visual conditions.

To avoid ambiguity between problem formulations, this study explicitly distinguishes between two related but distinct tasks. The primary task addressed is lane segmentation, which aims to accurately delineate lane boundaries under challenging visual conditions. Anomaly detection in this work is treated as a secondary, decision-level process derived from the segmentation output, in which lane departure events are identified based on spatial deviation criteria. By structuring the problem in this manner, the proposed method focuses on improving segmentation robustness through Feature Cross Attention, while anomaly detection operates as an interpretable post-processing step rather than an independent learning objective.

To clearly guide the scope and evaluation of this study, the following research questions are formulated:

RQ1: To what extent does the integration of Feature Cross Attention improve lane boundary segmentation accuracy compared to the baseline U-Net under challenging visual conditions such as low illumination, occlusion, and lane distortion?

RQ2: How does the proposed FCA-enhanced U-Net affect the sensitivity of anomaly detection based on spatial lane deviation, particularly in detecting subtle lane departure events?

RQ3: What are the trade-offs between segmentation accuracy, anomaly detection robustness, and computational efficiency when incorporating Feature Cross Attention into a lightweight U -Net architecture?

The rest of this work is structured as follows: Section 2 discusses the methodology and architecture of the updated U-Net model. In Section 3, the results and discussion are shown, with a focus on the benefits of the suggested method. Finally, Section 4 wraps up the study and talks about where research on lane departure detection and self-driving cars should go in the future.

2. Method

This research presents a deep learning methodology for lane departure detection in autonomous driving systems, employing a modified U-Net architecture integrated with feature cross-attention methods. Cropping the image, downsizing it to 128x128 pixels, and normalizing it are all processes in preprocessing. Lane annotations are converted to binary masks, and interpolation is applied to smooth the lane boundaries. The improved U-Net architecture successfully captures spatial dependencies, while the attention methods focus on essential aspects to boost anomaly identification. We test and evaluate the model in real-world settings to assess its performance relative to traditional methods.

2.1. Data Collection

This study uses the Comma2k19-LD dataset, obtained from Kaggle, which consists of 2,000 manually annotated frames from 100 highway driving scenarios recorded at 20 Hz. Each scenario includes front-facing video frames and synchronized vehicle telemetry, such as steering angle, speed, and position. Although limited in size compared to large-scale datasets, Comma2k19-LD provides high-quality lane annotations and diverse real-world driving conditions, making it suitable for controlled evaluation of lane departure detection. A scene-level split strategy is employed to prevent temporal leakage and ensure fair performance assessment.

A subset of frames with clearly visible left and right lane markers is selected, and frames are sampled at regular intervals to reduce redundancy across scenes. Annotation data in XML format is converted into lane-point sequences for supervised training. Fig. 1 illustrates a representative preprocessed frame with annotated lane markings, reflecting the visual quality and annotation structure used in this work.

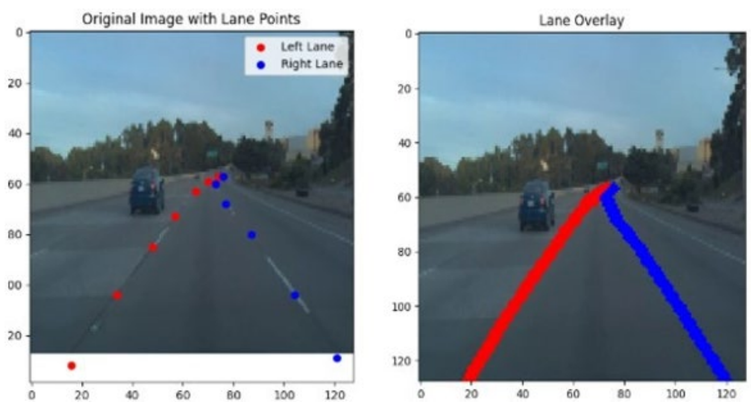

Fig. 1. Example image from the comma 2k19 LD dataset.

2.2. Data Preprocessing

The preprocessing stage transforms raw driving images into structured inputs for lane departure anomaly detection. Each frame is cropped to the driving corridor and resized to 128×128 pixels. Lane boundary annotations are converted into smooth curves and binary segmentation masks representing lane regions. Images and masks are stored in NumPy format to enable efficient batch loading and training. To enhance anomaly sensitivity, synthetic anomaly masks are generated during preprocessing to simulate realistic lane deviations, exposing the model to both normal and abnormal lane configurations and promoting stable convergence.

To ensure reliable training and evaluation, a scene-level splitting strategy is adopted. All frames from the same driving scenario are assigned exclusively to a single data partition, preventing temporal leakage between training, validation, and test sets. Specifically, 70% of scenes are used for training, 15% for validation, and 15% for testing. Anomaly injection is applied only during training, whereas the validation and test sets contain natural driving sequences to ensure an unbiased performance evaluation.

2.3. Model Development

This work proposes a modified U-Net architecture that integrates Feature Cross Attention (FCA) to improve anomaly detection in lane-departure scenarios. The architecture preserves spatial accuracy while incorporating global contextual information, enabling the model to distinguish normal lane following from abnormal deviations. The design follows a standard encoder–decoder structure with skip connections, augmented with attention modules at the bottleneck and decoder stages to enhance feature representation. As illustrated in Fig. 2, FCA at the bottleneck captures global semantic context, while decoder-level FCA refines spatial details during upsampling. This design balances global awareness and fine-grained localization, resulting in improved robustness to complex road geometries and visual degradation.

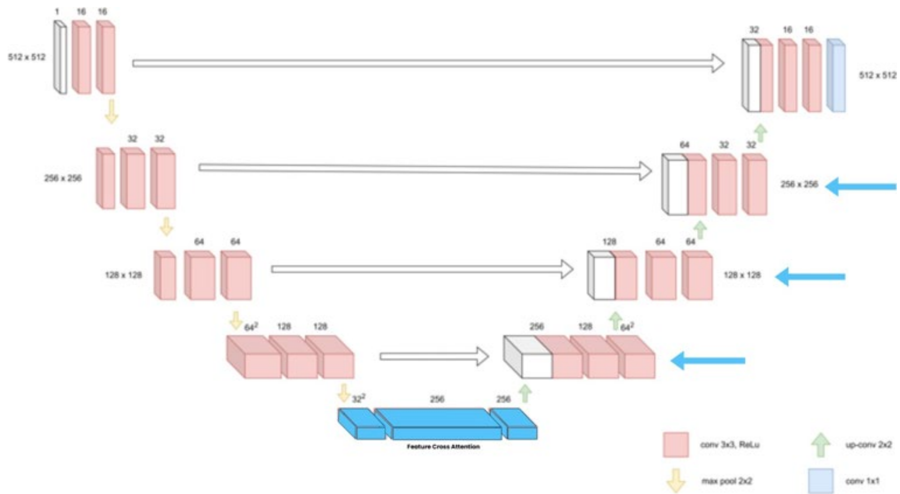

Fig. 2. The modified U-Net architecture with Feature Cross Attention for lane departure anomaly detection.

U-Net is widely adopted for image segmentation due to its symmetric encoder– decoder design and multi-scale feature extraction via skip connections [20], [21]. However, conventional skip connections may produce blurred boundaries. Recent studies have addressed this limitation through attention mechanisms and modified connections to better preserve low-level spatial information [21], [22]. Feature Cross Attention enhances semantic segmentation and anomaly sensitivity by enabling cross-channel interaction and spatial refinement [15]. Its effectiveness in improving edge accuracy and contextual awareness has been demonstrated in DeepLabv3+ [17] and semantic-aware anomaly detection frameworks [23], [24]. To address performance degradation in complex driving environments, attention-based extensions of U-Net, such as SE, SCSE, and EMA, have been proposed to suppress background noise and emphasize relevant spatial regions [14], [25], with demonstrated gains in rare-event detection and spatial change segmentation tasks [19], [26].

The proposed FCA does not introduce a novel attention mechanism but adapts established channel and spatial attention principles within the U-Net framework. The novelty lies in the placement and fusion of FCA at the bottleneck and decoder stages, which strengthen encoder–decoder feature interactions and improve lane boundary localization under challenging visual conditions.

Fig. 3 shows the Feature Cross Attention (FCA) module, which improves feature representation by integrating spatial attention (SA) and channel attention (CA) [17]. SA obtains low-level spatial information for boundary mapping, while CA obtains high-level semantic context via global pooling [27]. Their merger improves lane segmentation in challenging situations, such as obstructions or changes in lighting [28]. This modification, based on attention, improves border recognition in complicated road scenarios [29]. The combined attention also makes the model more robust by letting it adaptively focus on both spatial and semantic dimensions.

The fusion process starts by combining the two attention outputs, followed by convolution and activation layers. Channel attention gives a global context, while spatial attention improves pixel-level localization. This two-branch topology helps the network find lane changes more accurately by balancing spatial accuracy and knowledge of the environment. The model stays strong even as the road textures and weather change since both algorithms work together. This synergy enhances the model's ability to generalize across unseen road scenarios and ensures consistent detection performance.

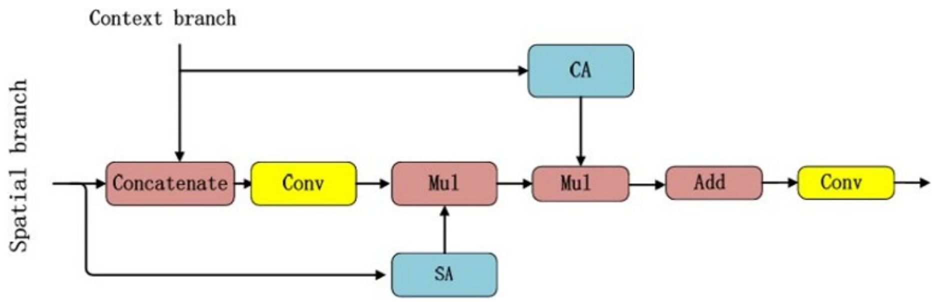

Fig. 3. Feature Cross Attention (FCA).

Let F ∈ ℝ^{C×H×W} denote the intermediate feature map extracted from the encoder-decoder pathway. The proposed Feature Cross Attention (FCA) module refines F by sequentially applying Channel Attention (CA) and Spatial Attention (SA) to enhance both semantic relevance and spatial localization.

Channel Attention aims to model inter-channel dependencies by emphasizing informative feature channels. Global average pooling is first applied to compress the spatial dimensions of F:

$$Z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_c\,(i,j) \tag{1}$$

where z ∈ ℝ^{C} represents the channel-wise descriptor. The attention weights are then computed using a lightweight gating mechanism:

$$s = \sigma\big(W_2\, \delta(W_1 z)\big) \tag{2}$$

where W_1 and W_2 denote learnable weight matrices, $\delta(\cdot)$ is the ReLU activation, and $\sigma(\cdot)$ is the sigmoid function. The channel-refined feature map is obtained as:

$$F_c^{CA} = s_c \cdot F_c \tag{3}$$

Spatial Attention highlights informative spatial regions by exploiting contextual cues across channels. The channel-refined feature map F^{CA} is aggregated along the channel dimension using average pooling

$$M(i,j) = \frac{1}{C} \sum_{c=1}^{C} F_c^{CA}\,(i,j) \tag{4}$$

The resulting spatial attention map is computed as:

$$A = \sigma(\mathrm{Conv}(M)) \tag{5}$$

where $\mathrm{Conv}(\cdot)$ denotes a convolution operation with a kernel size of 3×3. The final FCA output is obtained by spatially re-weighting the feature map:

$$F^{FCA}(i,j) = A(i,j) \cdot F^{CA}(i,j) \tag{6}$$

By combining Channel Attention and Spatial Attention within a unified FCA module, the network is able to emphasize semantically relevant feature channels while preserving fine-grained spatial details. This design strengthens encoder-decoder feature interaction, particularly at the bottleneck and decoder stages, enabling improved boundary refinement and anomaly-sensitive lane segmentation under challenging visual conditions.

2.4. Comparative Analysis with State-of-the-Art Methods

Recent advances in lane detection and attention-enhanced semantic segmentation have explored diverse architectural strategies to improve spatial consistency, contextual reasoning, and robustness under complex driving conditions. Structural-prior-based methods emphasize geometric modeling, while real-time approaches prioritize inference efficiency, and attention mechanisms enhance channel-wise and spatial feature refinement. Within this context, Table 1 summarizes representative state-of-the-art methods and positions the proposed U-Net with Feature Cross Attention relative to their key characteristics, strengths, and limitations.

Table 1. Comparison with representative state-of-the-art lane detection and attention-enhanced segmentation methods.

Method	Year	Backbone / Type	Relevant Contribution	Strengths	Limitations
SCNN: Spatial CNN [30]	2018	Spatial Convolution	Sequential spatial message passing to maintain lane continuity	Strong for long & curved lanes	Heavy and slow inference
LaneATT [31]	2020	ResNet based	Anchor based lane detection with attention for candidate prioritization	High accuracy, faster than SCNN	Requires anchor tuning and heuristic design
UltraFast Lane Detector [32]	2020	ResNet 18	Extremely fast row based lane classification	Real time (300+ FPS)	Less stable in extreme or noisy conditions
CondLaneNet [33]	2021	ResNet	Conditional convolution modeling inter lane relationships	Very high accuracy	Computationally expensive
ENet SAD [34]	2019	ENet	Self attention distillation for lightweight lane detection	Efficient and lightweight	Sensitivity to occlusion
DANet: Dual Attention Network [34]	2019	ResNet	Channel + spatial dual attention for segmentation	Strong global context reasoning	High computational cost
CBAM [35]	2018	Modular attention	Lightweight channel & spatial attention module	Easy to integrate; low overhead	Limited cross feature interaction
UNet++ + Attention Gate [36], [37]	2018	U-Net	Attention gates for suppressing irrelevant regions	Good segmentation accuracy	Heavier than standard U-Net
Proposed U-Net + Feature Cross Attention	**2025**	**U-Net (128×128)**	**Cross attention between the encoder and decoder for enhanced feature alignment**	**Lightweight, real-time friendly, anomaly sensitive**	**Requires adaptive thresholding**

Table 1 summarizes representative state-of-the-art lane detection and attention-enhanced segmentation methods, highlighting differences in architectural complexity, computational cost, and modeling assumptions. Structural-prior-based approaches emphasize geometric consistency, while real-time models prioritize inference speed. Attention-enhanced architectures improve contextual reasoning but often introduce additional computational overhead.

In contrast, the proposed U-Net with Feature Cross Attention strengthens encoder–decoder feature interaction using a lightweight cross-attention mechanism, achieving a favorable balance between segmentation robustness, anomaly sensitivity, and computational efficiency. While not intended to replace state-of-the-art lane extractors, the model is well-suited for real-time anomaly-aware lane monitoring, where efficiency and stability take precedence over complex geometric inference.

2.5. Training Configuration

The proposed model was trained using the Adam optimizer with an initial learning rate of 0.001 and a cosine annealing schedule over 10 epochs to promote stable convergence. A batch size of 4 was used, and the training objective combined Binary Cross-Entropy and Dice loss to balance pixel-wise accuracy and the structural consistency of lane boundaries. All models were initialized with Xavier uniform initialization and trained on an NVIDIA GPU. Training convergence was monitored through loss curves and validation accuracy, which indicated stable learning behavior without severe overfitting.

No geometric or photometric data augmentation was applied during training. Given the controlled nature of the dataset and the focus on preserving lane geometry, augmentation was intentionally avoided to prevent introducing unrealistic distortions. Instead, anomaly injection was applied during training to simulate lane deviation scenarios and improve robustness. Scene-level splitting was employed to reduce temporal bias and overfitting. Future work will explore geometry-preserving augmentation strategies, such as mild illumination variation or camera-aware transformations, to further enhance robustness.

2.6. Testing Model

The suggested model was tested on a separate test set that included lane-departure situations under varying road conditions, lighting, and obstructions. We used common segmentation metrics such as Intersection over Union (IoU) and Pixel Accuracy to assess performance. We also used a task-specific evaluation based on lane adherence and deviation tolerance [38], [39]. We compared the model's predictions with the ground-truth lane masks to demonstrate that it could accurately detect deviations from the annotated lane borders [40]. This assessment configuration ensures that the model's performance meets the demands of real-world driving variability and anomaly detection. This supports recent research that calls for task-aware evaluation procedures in autonomous driving systems [41].

Fig. 4 shows some of the model's results on test photos. The model effectively detected small drifts and misalignments that conventional U-Net topologies overlooked.

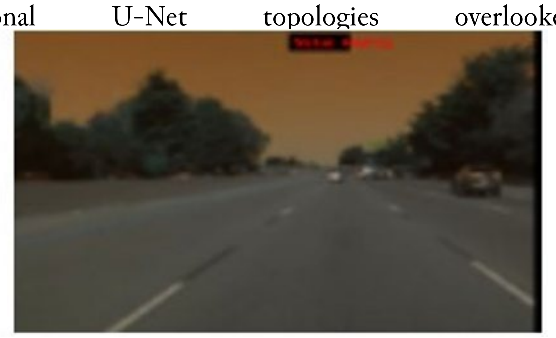

Fig. 4. Example outputs of the proposed model showing normal driving status (left) and lane departure anomaly detection (right).

These results show that adding Feature Cross Attention can help improve spatial awareness and segmentation quality in difficult driving situations in the real world. The left image represents a typical driving status that was accurately anticipated, whereas the right image indicates an anomaly that was discovered where the vehicle goes off the expected lane boundaries. These visual comparisons further prove that the model is strong enough to tell the difference between important lane departure situations in a variety of weather conditions.

2.7. Evaluation

We used conventional segmentation metrics such as Intersection over Union (IoU) and Pixel Accuracy to assess the performance of the proposed model. We did this on a test set that included different driving conditions, such as occlusions, curvy roadways, and low-light situations [42], [43], [44], [45], [46]. These metrics provide a comprehensive assessment of segmentation performance across diverse visual conditions.

Intersection over Union (IoU) measures the degree of overlap between the predicted lane mask and the ground truth, and is defined as.

$$IoU = \frac{TP}{TP+FP+FN} \tag{7}$$

where TP, FP, and FN denote the number of true positive, false positive, and false negative pixels, respectively. A higher IoU indicates more accurate segmentation with minimal false detections and omissions.

Pixel Accuracy evaluates the overall correctness of pixel-wise predictions, including both foreground and background, and is given by:

$$\text{Pixel } Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{8}$$

Where TN denotes true negatives, Pixel Accuracy reflects the proportion of correctly classified pixels across the image. Both Pixel Accuracy and Intersection over Union (IoU) are widely used evaluation metrics in semantic segmentation for autonomous driving applications [40], [43].

In this study, IoU is computed for the lane class only (foreground and IoU) to avoid background dominance and provide a more representative assessment of lane segmentation quality. Pixel Accuracy is reported as a complementary metric and interpreted alongside IoU, given the inherent imbalance between lane and background pixels.

Lane annotations are represented as binary masks within a predefined region of interest (ROI) covering the driving corridor. Mask thickness follows the original lane annotations and is preserved during resizing to 128×128. All evaluations are performed on the cropped ROI to ensure that reported metrics focus on lane-relevant regions rather than background-dominated areas. For anomaly detection, a spatial deviation threshold of τ=20 pixels is applied, and a predicted lane region is considered anomalous when the deviation from the ground truth exceeds this threshold:

$$\text{Anomaly } (x, y) = \begin{cases} 1, \text{if} ||P9(x, y) - G(x, y)|| > 20 \\ \quad 0, \text{otherwise} \end{cases} \tag{9}$$

where P (x, y) and G (x, y) stand for the projected and actual lane positions at pixel (x, y). The 20-pixel threshold was chosen based on real-world data to match the maximum lateral offset used during synthetic anomaly injection (±20 px), ensuring that the training anomalies and the assessment criteria were consistent.

Also, earlier research on lane departure warning systems shows that lane shifts in image space that are more than 15-20 pixels generally match real-world drifts that are big enough to require safety measures in ADAS. So, this threshold strikes a balance between sensitivity to real departures and avoiding false positives from noise, perspective distortion, or small model errors.

3. Results and Discussion

The proposed U-Net with Feature Cross Attention (U-Net FCA) was evaluated on a held-out test set and achieved a Pixel Accuracy of 99.19% and an Intersection over Union (IoU) of 98.47%, outperforming the baseline U-Net (97.56% Pixel Accuracy and 97.46% IoU). These results indicate that integrating Feature Cross Attention improves feature representation and spatial prediction accuracy, while the consistent decrease in validation loss suggests stable training behavior.

Despite these encouraging results, the relatively limited dataset size constrains the generalizability of the proposed model. Although scene-level splitting and anomaly injection enhance robustness within the evaluated scenarios, the reported performance may not fully reflect behavior across diverse geographic regions, camera configurations, or traffic conditions. In addition, the frame-based formulation does not explicitly model temporal continuity, which may lead to instability under rapidly changing driving conditions. Consequently, the evaluation focuses on per-frame segmentation accuracy and anomaly sensitivity rather than long-term temporal consistency.

The use of a fixed 20-pixel spatial deviation threshold enables effective identification of lane departure events and improves sensitivity to subtle lane shifts under occlusion and distorted lane geometry. Visual inspection confirms more stable lane boundary estimation compared to the baseline model, particularly in challenging conditions. These findings suggest that the proposed approach is suitable for real-time lane monitoring applications, while acknowledging that further validation on larger datasets and with temporal modeling is required for broader deployment.

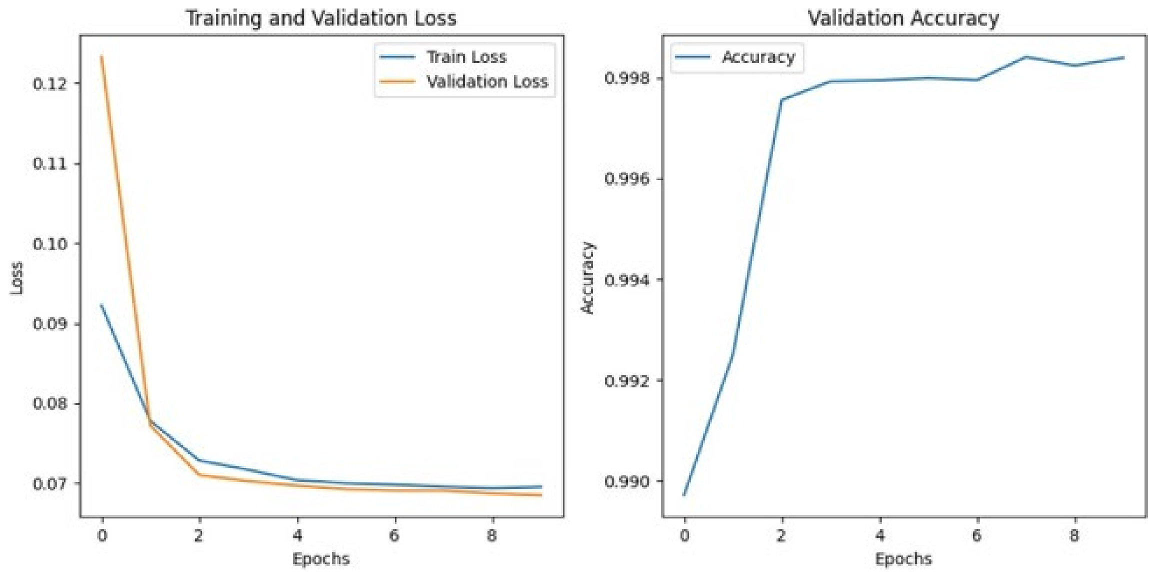

Fig. 5. Training and validation loss (left) and validation accuracy (right) curves of the proposed model.

Fig. 5 illustrates the training dynamics of the proposed model over 10 epochs. Both training and validation losses decrease consistently, while validation accuracy stabilizes at approximately 99%, indicating stable convergence without severe overfitting. The small gap between training and validation curves suggests that the attention-enhanced architecture generalizes reasonably well within the evaluated scenarios.

To further confirm anomaly identification, a confusion matrix was made utilizing 210 validation photos, which included both normal and abnormal situations. Fig. 6 shows that the proposed model correctly classified all 95 normal and 115 anomalous samples in the validation set under the defined experimental conditions. This result indicates that the segmentation -based anomaly criterion is effective for distinguishing lane departure events within the controlled evaluation setting. However, this outcome should be interpreted with caution, as the anomaly definition is based on a predefined spatial deviation threshold, and the dataset size remains limited. The classification accuracy shows that the integrated attention mechanism helps the model pick up on little changes and changes in context that are hard for regular segmentation algorithms to see.

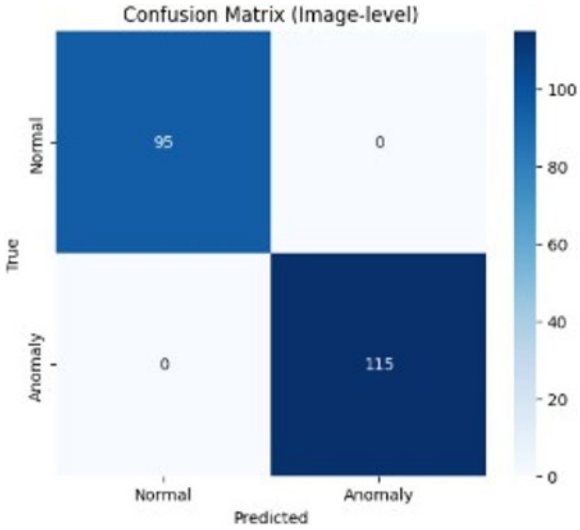

Fig. 6. Confusion matrix showing image-level classification results.

Table 2 shows that the proposed U-Net architecture with Feature Cross Attention outperforms the baseline model on both assessment metrics. It shows significant increases in both lane anomaly detection and segmentation accuracy. The model achieves a Pixel Accuracy of 99.19% and an IoU of 98.47%, which are better than the baseline U-Net's scores of 97.56% and 97.46%, respectively. These steady advances show that FCA is good at improving spatial and contextual feature representations, which helps the model keep lane structures intact and find small changes even when there are occlusions, noise, or geometric distortions. These kinds of advances are especially important for self-driving cars, since being able to accurately read lanes is directly related to safety and reliability.

Table 2. Performance Comparison Between Baseline U-Net and Proposed U-Net With Feature Cross Attention (FCA).

Model	Pixel Accuracy (%)	IoU (%)
Baseline U-Net	97.56	97.46
U-Net + FCA	**99.19**	**98.47**

The suggested method shows strong performance in segmentation and anomaly detection; however, a key limitation is its reliance in a fixed 20-pixel spatial deviation threshold to detect lane departures. This threshold performs well under the current dataset and experimental setup but is sensitive to camera resolution, field of view, and mounting configuration. In real-world scenarios, a constant pixel-based threshold may not consistently capture meaningful lateral deviations, potentially leading to false positives or missed anomalies. In practical deployments, classification performance is also expected to be influenced by factors such as camera calibration, perspective distortion, and scene dynamics, which are not fully captured in the current frame -based evaluation.

Although the proposed model demonstrates strong segmentation and anomaly detection performance under most evaluated conditions, several challenging scenarios remain. Performance degradation is observed under extremely low-light or nighttime conditions, strong glare or overexposure during sunrise or sunset, and road surfaces with worn, faded, or partially occluded lane markings, which introduce ambiguity in lane boundary estimation. Abrupt changes in road geometry, such as sharp curves or lane merges, further challenge the frame-based formulation due to the absence of temporal smoothing. These observations indicate that while the method is robust in typical highway scenarios, its performance may be limited in the presence of severe visual degradation or highly dynamic environments.

Anomaly detection is based on a fixed spatial deviation threshold τ = 20 pixels after resizing the frames to 128×128, consistent with the maximum offset used during synthetic anomaly injection (±20 px). Smaller thresholds increase sensitivity but may introduce false positives due to segmentation noise or perspective distortion, whereas larger thresholds reduce false alarms at the cost of missing subtle lane departures. The selected threshold represents a practical balance between sensitivity and robustness within the evaluated scenarios. Since pixel-based thresholds are influenced by camera calibration and perspective effects, future work will investigate adaptive or geometry-aware thresholding strategies to improve reliability.

The proposed U-Net with Feature Cross Attention also exhibits a favorable computational profile for real-time deployment, with approximately 39.23 million parameters and an estimated 9 .63 GFLOPs per 128×128 input. This efficiency supports deployment on both desktop GPUs and embedded platforms, reinforcing the method's practicality as a lightweight, deployment-oriented enhancement to attention-augmented U-Net architectures.

4. Conclusion

This study proposed a revised U-Net architecture enhanced with Feature Cross Attention (FCA) to detect anomalies in lane-departure situations. The model showed better segmentation performance than the baseline by using attention methods at both the bottleneck and decoder phases. It achieved 99.19% pixel accuracy and 98.47% IoU on a variety of driving circumstances. The inclusion of a spatial deviation threshold made it even easier to find unusual lane departures by separating small drifts from regular behavior. The attention-enhanced design improves spatial awareness, edge refinement, and generalization, making it perfect for real-time autonomous driving systems. This is backed up by both qualitative and quantitative data. Future endeavors may investigate lightweight modifications for embedded deployment and integration with temporal models to augment resilience in dynamic settings.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.

Funding statement. None of the authors has received any funding or grants from any institution or funding body for the research.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

[1] P. Shivale, N. Sonawane, S. Dharmale, T. Lokhandwala, and P. D. M. B. Wagh, "Advanced Driver Assistance System," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 3, pp. 1111–1113, Mar. 2023, doi: 10.22214/ijraset.2023.49547.

[2] N. B. Chetan, J. Gong, H. Zhou, D. Bi, J. Lan, and L. Qie, "An Overview of Recent Progress of Lane Detection for Autonomous Driving," in *2019 6th International Conference on Dependable Systems and Their Applications (DSA)*, IEEE, Jan. 2020, pp. 341–346. doi: 10.1109/DSA.2019.00052.

[3] V. S. S.D. and P. C.J., "Revolutionary Enhanced Lane Departure Detection Techniques for Autonomous Vehicle Safety using ADAS.," *Int. J. Electron. Commun. Eng.*, vol. 11, no. 9, pp. 22–35, Sep. 2024, doi: 10.14445/23488549/IJECE-V11I9P103.

[4] S. Waykole, N. Shiwakoti, and P. Stasinopoulos, "Review on Lane Detection and Tracking Algorithms of Advanced Driver Assistance System," *Sustainability*, vol. 13, no. 20, p. 11417, Oct. 2021, doi: 10.3390/su132011417.

[5] S. Huang, N. A. M. Zin, and M. H. I. Hamzah, "A Review of Deep Learning-Based Lane Detection Methods in Complex Environments," *Int. J. Basic Appl. Sci.*, vol. 14, no. 4, pp. 549–561, Aug. 2025, doi: 10.14419/wb7z2179.

[6] S. Sultana, B. Ahmed, M. Paul, M. R. Islam, and S. Ahmad, "Vision-Based Robust Lane Detection and Tracking in Challenging Conditions," *IEEE Access*, vol. 11, pp. 67938–67955, 2023, doi: 10.1109/ACCESS.2023.3292128.

[7] Y. Zhang, Z. Lu, X. Zhang, J.-H. Xue, and Q. Liao, "Deep Learning in Lane Marking Detection: A Survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 5976–5992, Jul. 2022, doi: 10.1109/TITS.2021.3070111.

[8] G. Krishna Kushwaha *et al.*, "Review on Computer Vision Techniques to enhance Road Lane Detection by using Open CV," *Int. J. Innov. Res. Adv. Eng.*, vol. 11, no. 02, pp. 102–109, Feb. 2024, doi: 10.26562/ijirae.2024.v1102.07.

[9] N. E. Brown *et al.*, "Real World Use Case Evaluation of Radar Retro-reflectors for Autonomous Vehicle Lane Detection Applications," in *SAE Technical Papers*, Apr. 2024, pp. 1–11. doi: 10.4271/2024-01-2042.

[10] H. Aisha S and S. K T, "Deep Learning Based Lane Detection for Auto Driving Vehicles," *Int. J. Sci. Res. Eng. Manag.*, vol. 09, no. 08, pp. 1–9, Aug. 2025, doi: 10.55041/IJSREM51949.

[11] R. Liu *et al.*, "BiAttentionNet: a dual-branch automatic driving image segmentation network integrating spatial and channel attention mechanisms," *Sci. Rep.*, vol. 15, no. 1, p. 13193, Apr. 2025, doi: 10.1038/s41598-025-95470-4.

[12] S. Santhiya, I. Johnraja Jebadurai, G. Jeba Leelipushpam Paulraj, P. Pavan Venkata Vamsi, M. Aravind Reddy, and P. Poulraju, "FedLANE: a federated U-Net architecture for lane detection," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 32, no. 3, p. 1621, Dec. 2023, doi: 10.11591/ijeecs.v32.i3.pp1621-1629.

[13] S.-H. Lee and S.-H. Lee, "U-Net-Based Learning Using Enhanced Lane Detection with Directional Lane Attention Maps for Various Driving Environments," *Mathematics*, vol. 12, no. 8, p. 1206, Apr. 2024, doi: 10.3390/math12081206.

[14] O. Hotkar, P. Radhakrishnan, A. Singh, N. Jhamnani, and R. V. Bidwe, "U-Net and YOLO: AIML Models for Lane and Object Detection in Real-Time," in *Proceedings of the 2023 Fifteenth International Conference on Contemporary Computing*, New York, NY, USA: ACM, Aug. 2023, pp. 467–473. doi: 10.1145/3607947.3608049.

[15] D. Liu, D. Zhang, L. Wang, and J. Wang, "Semantic segmentation of autonomous driving scenes based on multi-scale adaptive attention mechanism," *Front. Neurosci.*, vol. 17, p. 1291674, Oct. 2023, doi: 10.3389/fnins.2023.1291674.

[16] T.-K. Yin, L.-Y. Wu, and T.-P. Hong, "Axial Attention Inside a U-Net for Semantic Segmentation of 3D Sparse LiDAR Point Clouds," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, Jun. 2022, pp. 1543–1549. doi: 10.1109/IV51971.2022.9827257.

[17] H. Zeng, S. Peng, and D. Li, "Deeplabv3+ semantic segmentation model based on feature cross attention mechanism," *J. Phys. Conf. Ser.*, vol. 1678, no. 1, p. 012106, Nov. 2020, doi: 10.1088/1742-6596/1678/1/012106.

[18] T. Reiss, N. Cohen, L. Bergman, and Y. Hoshen, "PANDA: Adapting Pretrained Features for Anomaly Detection and Segmentation," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2021, pp. 2805–2813. doi: 10.1109/CVPR46437.2021.00283.

[19] L. P. J. Sträter, M. Salehi, E. Gavves, C. G. M. Snoek, and Y. M. Asano, "GeneralAD: Anomaly Detection Across Domains by Attending to Distorted Features," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 15095 LNCS, Springer Science and Business Media Deutschland GmbH, 2025, pp. 448–465. doi: 10.1007/978-3-031-72913-3_25.

[20] H. V. Patri, M. B. Priya, M. B. Mothukuri, D. M. Kumar, K. V. Ratna Prabha, and S. R. K. Gandham, "U-Net Advancements in Semantic Segmentation for Autonomous Vehicles," in *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, Mar. 2024, pp. 2292–2296. doi: 10.1109/ICACCS60874.2024.10717217.

[21] H. Zunair and A. Ben Hamza, "Sharp U-Net: Depthwise convolutional network for biomedical image segmentation," *Comput. Biol. Med.*, vol. 136, no. September, p. 104699, Sep. 2021, doi: 10.1016/j.compbiomed.2021.104699.

[22] J. Di, S. Ma, J. Lian, and G. Wang, "A U-Net Network Model for Medical Image Segmentation Based on Improved Skip Connections," in *2022 14th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, IEEE, Jan. 2022, pp. 298–302. doi: 10.1109/ICMTMA54903.2022.00064.

[23] J. Yang, Y. Shi, and Z. Qi, "Learning deep feature correspondence for unsupervised anomaly detection and segmentation," *Pattern Recognit.*, vol. 132, no. December, p. 108874, Dec. 2022, doi: 10.1016/j.patcog.2022.108874.

[24] W. Ma, Y. Li, S. Lan, W. Wang, W. Huang, and W. Zhu, "Semantic-aware normalizing flow with feature fusion for image anomaly detection," *Neurocomputing*, vol. 590, no. July, p. 127728, Jul. 2024, doi: 10.1016/j.neucom.2024.127728.

[25] J. Fang, X. Zhang, B. Yang, S. Chen, and B. Li, "An Attention-based U-Net Network for Anomaly Detection in Crowded Scenes," in *14th International Conference on Computer Research and Development (ICCRD)*, IEEE, Jan. 2022, pp. 202–206. doi: 10.1109/ICCRD54409.2022.9730481.

[26] K. Kalinaki, O. A. Malik, and D. T. Ching Lai, "FCD-AttResU-Net: An improved forest change detection in Sentinel-2 satellite images using attention residual U-Net," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 122, no. August, p. 103453, Aug. 2023, doi: 10.1016/j.jag.2023.103453.

[27] C. Wang and N. Aouf, "Fusion Attention Network for Autonomous Cars Semantic Segmentation," in *2022 IEEE Intelligent Vehicles Symposium (IV)*, IEEE, Jun. 2022, pp. 1525–1530. doi: 10.1109/IV51971.2022.9827377.

[28] Z. Guo, Y. Huang, H. Wei, C. Zhang, B. Zhao, and Z. Shao, "DALaneNet: A Dual Attention Instance Segmentation Network for Real-Time Lane Detection," *IEEE Sens. J.*, vol. 21, no. 19, pp. 21730–21739, Oct. 2021, doi: 10.1109/JSEN.2021.3100489.

[29] Y. Kao, S. Che, S. Zhou, S. Guo, X. Zhang, and W. Wang, "LHFFNet: A hybrid feature fusion method for lane detection," *Sci. Rep.*, vol. 14, no. 1, p. 16353, Jul. 2024, doi: 10.1038/s41598-024-66913-1.

[30] X. Pan, J. Shi, P. Luo, X. Wang, and X. Tang, "Spatial as Deep: Spatial CNN for Traffic Scene Understanding," *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, pp. 7276–7283, Apr. 2018, doi: 10.1609/aaai.v32i1.12301.

[31] L. Tabelini, R. Berriel, T. M. Paixao, C. Badue, A. F. De Souza, and T. Oliveira-Santos, "Keep your Eyes on the Lane: Real-time Attention-guided Lane Detection," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2021, pp. 294–302. doi: 10.1109/CVPR46437.2021.00036.

[32] Z. Qin, H. Wang, and X. Li, "Ultra Fast Structure-Aware Deep Lane Detection," in *Computer Vision -- ECCV 2020*, vol. 12369, A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, Eds., in Lecture Notes in Computer Science, vol. 12369. , Cham: Springer, 2020, pp. 276–291. doi: 10.1007/978-3-030-58586-0_17.

[33] L. Liu, X. Chen, S. Zhu, and P. Tan, "CondLaneNet: a Top-to-down Lane Detection Framework Based on Conditional Convolution," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2021, pp. 3753–3762. doi: 10.1109/ICCV48922.2021.00375.

[34] J. Fu *et al.*, "Dual Attention Network for Scene Segmentation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2019, pp. 3141–3149. doi: 10.1109/CVPR.2019.00326.

[35] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *Computer Vision -- ECCV 2018*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11211. , Cham: Springer, 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.

[36] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A Nested U-Net Architecture for Medical Image Segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, vol. 11045, in Lecture Notes in Computer Science, vol. 11045. , Springer, 2018, pp. 3–11. doi: 10.1007/978-3-030-00889-5_1.

[37] O. Oktay *et al.*, "Attention U-Net: Learning Where to Look for the Pancreas," *arXiv Prepr.*, pp. 1–10, 2018, [Online]. Available: https://arxiv.org/abs/1804.03999.

[38] P. Santhiya, I. J. Jebadurai, G. J. L. Paulraj, J. A, S. K. Karan, and E. Naveen.V, "Deep Vision: Lane Detection in ITS: A Deep Learning Segmentation Perspective," in *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*, IEEE, Jun. 2024, pp. 21–26. doi: 10.1109/ICICI62254.2024.00012.

[39] S. Swain and A. K. Tripathy, "Real-time lane detection for autonomous vehicles using YOLOV5 Segmentation Model," *Int. J. Sustain. Eng.*, vol. 17, no. 1, pp. 718–728, Dec. 2024, doi: 10.1080/19397038.2024.2400965.

[40] R. Yousri, M. A. Elattar, and M. S. Darweesh, "A Deep Learning-Based Benchmarking Framework for Lane Segmentation in the Complex and Dynamic Road Scenes," *IEEE Access*, vol. 9, pp. 117565–117580, 2021, doi: 10.1109/ACCESS.2021.3106377.

[41] T. Sato and Q. A. Chen, "Towards Driving-Oriented Metric for Lane Detection Models," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2022, pp. 17132–17141. doi: 10.1109/CVPR52688.2022.01664.

[42] S. Cakir, M. Gauß, K. Häppeler, Y. Ounajjar, F. Heinle, and R. Marchthaler, "Semantic Segmentation for Autonomous Driving: Model Evaluation, Dataset Generation, Perspective Comparison, and Real-Time Capability," no. July, pp. 1–8, Jul. 2022, [Online]. Available: http://arxiv.org/abs/2207.12939.

[43] J. Hyeon, M. Jeong, G. Shin, W.-C. Chern, V. K. Asari, and H. Kim, "Evaluating Road Segmentation Performance in Participatory Sensing: An Investigation into Alternative Metrics," in *Proceedings of the International Symposium on Automation and Robotics in Construction*, International Association for Automation and Robotics in Construction (IAARC), Jul. 2023, pp. 506–512. doi: 10.22260/ISARC2023/0068.

[44] B. Baheti, S. Innani, S. Gajre, and S. Talbar, "Semantic scene segmentation in unstructured environment with modified DeepLabV3+," *Pattern Recognit. Lett.*, vol. 138, no. October, pp. 223–229, Oct. 2020, doi: 10.1016/j.patrec.2020.07.029.

[45] J. Gu, M. Bellone, R. Sell, and A. Lind, "Object Segmentation for Autonomous Driving Using iseAuto Data," *Electronics*, vol. 11, no. 7, p. 1119, Apr. 2022, doi: 10.3390/electronics11071119.

[46] A. D. P. -, A. R. -, I. R. S. -, J. P. J. -, and S. K. -, "Computer Vision-Based Anomaly Detection in Industrial Components," *Int. J. Sci. Technol.*, vol. 16, no. 2, p. 13, Jun. 2025, doi: 10.71097/IJSAT.v16.i2.6154.