Performance Evaluation of Road Marking Detection Models in Indian Conditions

Khushi Agrawal[1], Jafri Syed Mujtaba[2] and Samarth Bankar[3]
Ahmedabad University, Ahmedabad, Gujarat 380009, India
[1]khushi.a2@ahduni.edu.in, [2]jafri.h@ahduni.edu.in, [3]samarth.b2@ahduni.edu.in

***Abstract*—Road marker detection is essential for infrastructure monitoring, traffic control, and autonomous driving. Large-scale road marker identification can now be effectively done thanks to the growing availability of UAV (Unmanned Aerial Vehicle) footage. The purpose of this work is to use the AU-Drone dataset to assess the performance of many Indian road marking identification technologies, mostly segmentation-based techniques. Examining current publically accessible models, determining their efficacy, and applying transfer learning to modify them for Indian road conditions are all part of the project. Using ground-truth data, an assessment system based on Python will be created to benchmark the performance of the model. The finished framework will be open source, offering a strong instrument for further study and real-world uses in autonomous navigation and urban planning.**

***Index Terms*—Road Marking Detection, UAV Imagery, Semantic Segmentation, Deep Learning, Transfer Learning, Indian Road Infrastructure, Autonomous Navigation, Model Evaluation Framework**

I. Introduction

Road markers are crucial visual aids that improve road safety and help autonomous cars and human drivers navigate their lanes. Ground-level cameras and LiDAR are frequently used in traditional road marker detection techniques, however their scalability and affordability are limited. A viable substitute is aerial photography, which offers high-resolution aerial views of highways that can be processed by deep learning-based segmentation models.

The evaluation of current segmentation algorithms for the identification of road markers in the Indian setting is the main objective of this study. Indian roads offer particular difficulties due to their irregular signage, different lane lines, and environmental elements, including fading paint and occlusion. The search for publically accessible models, the examination of their designs, and the comparison of their results using the AU-Drone dataset will be the first steps in the study process. Transfer learning will be used if necessary to modify these models for Indian road conditions. A Python-based system will be created that incorporates important performance indicators, including precision, recall, IoU (Intersection over Union), and F1 score, to guarantee objective evaluation. To confirm the prediction of the model, a manually labeled ground truth dataset will also be created. An open source framework that makes it easier to detect road markings will be the end product, which will help with applications like traffic control, smart city planning, and autonomous car navigation.

II. Related Works

The extraction of road markings is a pivotal task in high-definition mapping for autonomous driving and transportation management. Given the challenges posed by variable illumination, occlusions, and the small size of many lane marking features, recent research has explored various deep learning approaches to improve segmentation accuracy. Several studies have pursued model architectures that explicitly address the spatial and structural complexity of road markings. For example, Chen et al. [1] developed an Attentive Capsule Feature Pyramid Network (ACapsFPN) that integrates capsule networks with a feature pyramid structure and attention mechanisms. Unlike traditional convolutional neural networks (CNNs) that use scalar activations, capsule networks preserve orientation and pose information, thereby enhancing the detection of irregular patterns in road scenes. While this approach improves robustness to geometric variations and occlusions, it also introduces challenges in controlling false positives in highly cluttered urban environments.

In another line of research, Azimi et al. [2] introduced an architecture that combines a Fully Convolutional Neural Network (FCNN) with Discrete Wavelet Transform (DWT) for lane marking segmentation. The incorporation of DWT allows the model to retain high-frequency details that are critical for segmenting thin and faded markings, while a cost-sensitive loss function mitigates the impact of severe class imbalance inherent in aerial imagery. This method is particularly adept at handling small objects, although it remains sensitive to drastic illumination changes and complex shadowing effects.

More recent work by Zhang et al. [3] has focused on benchmarking a wide range of semantic segmentation models—both CNN-based and transformer-based—using transfer learning. Their comparative analysis, which includes models such as U-Net, DeepLabV3+, and SegFormer, demonstrates that transformer architectures excel in capturing long-range dependencies and contextual information, yielding superior performance metrics such as mean Intersection over Union (mIoU) and F1-score. However, the increased computational overhead of transformer models poses practical challenges for real-time applications.

Complementing these approaches, Chen et al. [4] proposed a Dense Feature Pyramid Network (DFPN) model tailored for instance segmentation of road markings from Mobile

Laser Scanning (MLS) point clouds. By concatenating shallow and deep features across multiple pyramid levels, the DFPN model is able to capture both fine-grained spatial details and high-level semantic information. The integration of focal loss in the segmentation branch further improves performance on hard-to-classify samples. Despite these advantages, the DFPN approach may experience difficulties with boundary delineation and misclassification when instances overlap or when the foreground is sparsely represented. Collectively,

these studies underscore the evolution of deep learning techniques for road marking segmentation. The progression from CNNs with enhanced spatial representation to transformer-based models and dense pyramid structures reflects the ongoing effort to balance accuracy, computational efficiency, and robustness. Despite significant progress, challenges such as handling occlusions, reducing false positives, and efficiently processing high-resolution data remain. Our work builds upon these insights by aiming to integrate complementary strategies to further improve the segmentation accuracy and operational efficiency in complex real-world environments.

REFERENCES

[1] H. Guan, X. Lei, Y. Yu, H. Zhao, D. Peng, J. M. Junior, and J. Li, "Road marking extraction in uav imagery using attentive capsule feature pyramid network," *International Journal of Applied Earth Observation and Geoinformation*, vol. 107, p. 102677, 2022.

[2] S. M. Azimi, P. Fischer, M. Körner, and P. Reinartz, "Aerial lanenet: Lane-marking semantic segmentation in aerial imagery using wavelet-enhanced cost-sensitive symmetric fully convolutional neural networks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2920–2938, 2018.

[3] W. Liu, S. Qiao, K. Gao, H. He, M. A. Chapman, L. Xu, and J. Li, "Advancements in road lane mapping: Comparative fine-tuning analysis of deep learning-based semantic segmentation methods using aerial imagery," *arXiv preprint arXiv:2410.05717*, 2024.

[4] S. Chen, Z. Zhang, R. Zhong, L. Zhang, H. Ma, and L. Liu, "A dense feature pyramid network-based deep learning model for road marking instance segmentation using mls point clouds," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 1, pp. 784–800, 2020.