Performance Evaluation of Road Marking Detection Models in Indian Conditions

Khushi Agrawal, Jafri Syed Mujtaba, Samarth Bankar

School of Engineering and Applied Science, Ahmedabad University

Email: { khushi.a2, jafri.h, samarth.b2}@ahduni.edu.in

Team: **CV Project**

April 17, 2025

Problem Statement

Ahmedabad University

Performance Evaluation of Road Marking Detection Models in Indian Conditions

This project aims to perform pixel-wise segmentation of Indian road scenes using deep learning models. The aim is to identify road elements such as lane markings, dividers, and road surfaces in drone or street-level images. This helps improve the understanding of the roads for autonomous driving and traffic analysis systems.

- Automate the pixel-wise classification of Indian road images into multiple classes.
- Handle real-world challenges like lighting variation, class imbalance, and faded markings.
- Evaluate and compare deep learning models to find the most accurate and efficient solution.

Instructor's Feedback

Ahmedabad University

- Suggested improving results by deeper training and better augmentation, and focusing on other datasets with Drone-imagery
- Explore different loss functions.
- Recommended class-wise performance evaluation to handle imbalances.
- Explore different evaluation metrics

Our Approach: Models Used

Ahmedabad University

We implemented and fine-tuned three models:

- **DeepLabV3+** – Encoder-decoder with ASPP; high pixel accuracy.
- **U-Net** – Lightweight with skip connections; faster and interpretable.
- **SegFormer** – Transformer-based; efficient with better contextual understanding.

All models were pretrained on ImageNet and fine-tuned using transfer learning.

Dataset and Preprocessing

Dataset: AU-Drone Dataset with Indian roads captured from UAVs.
Preprocessing:

- Images and masks resized to 640×360.
- Converted RGB masks to single-channel class labels.
- Applied data augmentation: flips, color jitter, affine transforms.
- Normalized images using ImageNet statistics.

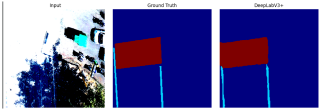
Results – Quantitative Metrics

Model	Pixel Accuracy	mIoU	F1 Score
U-Net	0.89	0.7502	0.8477
DeepLabV3+	0.98	0.65	–
SegFormer	0.97	0.64	–

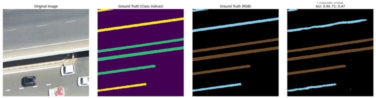
Table 1: Performance metrics across models after training

Observation: While U-Net shows high mIoU and F1, DeepLabV3+ gives the best pixel accuracy. SegFormer maintains competitive performance with lower training epochs.

Results – Qualitative Visuals



DeepLabv3+ Output



SegFormer Output

Visual Insight: DeepLabv3+ captures clearer road edges and markings. U-Net performs well but struggles slightly with fine boundaries. Both perform better than baseline segmentation.

Results – Evaluation Metrics

Table 2: Evaluation Metrics Comparison

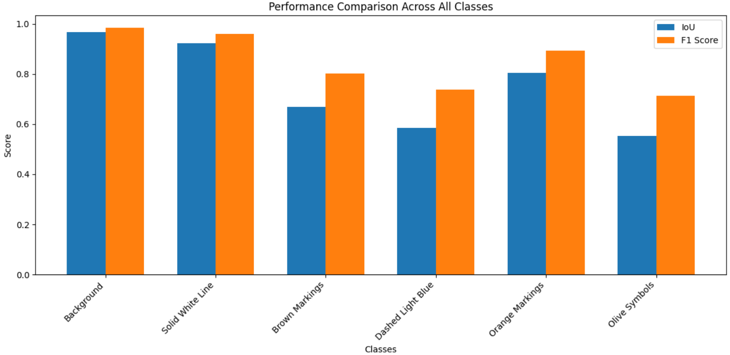
DeepLabV3+

Metric	Value
Pixel Accuracy	0.9864
Mean IoU	0.6589
Class 0 IoU (BG)	0.9849
Class 1 IoU (Marking)	0.7105
Class 2 IoU (Other)	0.0000
Class 3 IoU (Road)	0.9401

SegFormer

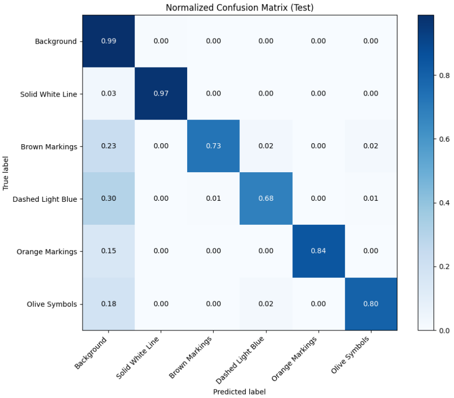
Metric	Value
Pixel Accuracy	0.9734
Mean IoU	0.6431
Class 0 IoU (BG)	0.9805
Class 1 IoU (Marking)	0.6923
Class 2 IoU (Other)	0.4437
Class 3 IoU (Road)	0.4569

Results – IoU and F1 Comparison

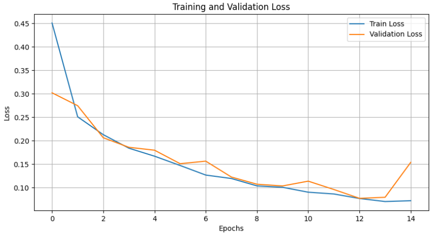


Comparison of IoU and F1 across all classes

Results – Confusion Matrix and Training History



Confusion Matrix



Training vs Validation Loss

Future Work

Ahmedabad University

- Train for more epochs to improve generalization
- Use focal loss to handle class imbalance
- Try additional models and transformer variants
- Add real-time segmentation and post-processing filters
- Explore deployment on edge/embedded devices for field use
- Expand dataset diversity (weather, road types, angles)

References

[1] Chen, L., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking Atrous Convolution for Semantic Image Segmentation*. arXiv:1706.05587.

[2] Long, J., Shelhamer, E., & Darrell, T. (2014). *Fully Convolutional Networks for Semantic Segmentation*. arXiv:1411.4038.

[3] Howard, A., Sandler, M., Chu, G., Chen, L.C., Tan, M., Wang, W., et al. (2019). *Searching for MobileNetV3*. arXiv:1905.02244.

[4] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J.M., & Luo, P. (2021). *SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers*. arXiv:2105.15203.

[5] Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. arXiv:1505.04597.

[6] Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2016). *Pyramid Scene Parsing Network*. arXiv:1612.01105.