

TEAM AIRLINES

Members

Abdullah

Muhammad Irfan

Rameel Sohail

Project Report

Advanced Off-Road Semantic Segmentation

Project Name

Off-Road SegFormer Explorer

Brief Tagline

Navigating unstructured desert terrains using hierarchical
Transformers and robust spatial attention.

1. Summary

This project focuses on semantic segmentation for unstructured desert terrains using a Transformer-based SegFormer-B2 architecture. Unlike traditional CNN-based models, SegFormer leverages hierarchical Transformers to capture global context, which is critical for stable horizon detection and long-range trail understanding in wide-angle off-road environments.

By integrating a Hybrid Loss (Weighted Cross-Entropy + Dice Loss), heavy data augmentation, and Test-Time Augmentation (TTA), we significantly improved generalization and obstacle detection performance.

2. Methodology

Architecture Selection:

Transitioned from CNN-based U-Nets to SegFormer-B2. The Transformer encoder captures global dependencies while maintaining lightweight computational efficiency.

Loss Strategy:

Implemented Hybrid Loss combining Weighted Cross-Entropy and Dice Loss. High class weight (8.0) assigned to 'Rocks' and 'Logs' to prevent ignoring rare but dangerous obstacles.

Training Setup:

Optimizer: AdamW | Learning Rate: 6×10^{-5} | Scheduler: Cosine Annealing | Epochs: 25.

3. Results & Performance Metrics

Official Test Score:

0.3323 mIoU (Improved from previous 0.28 baseline).

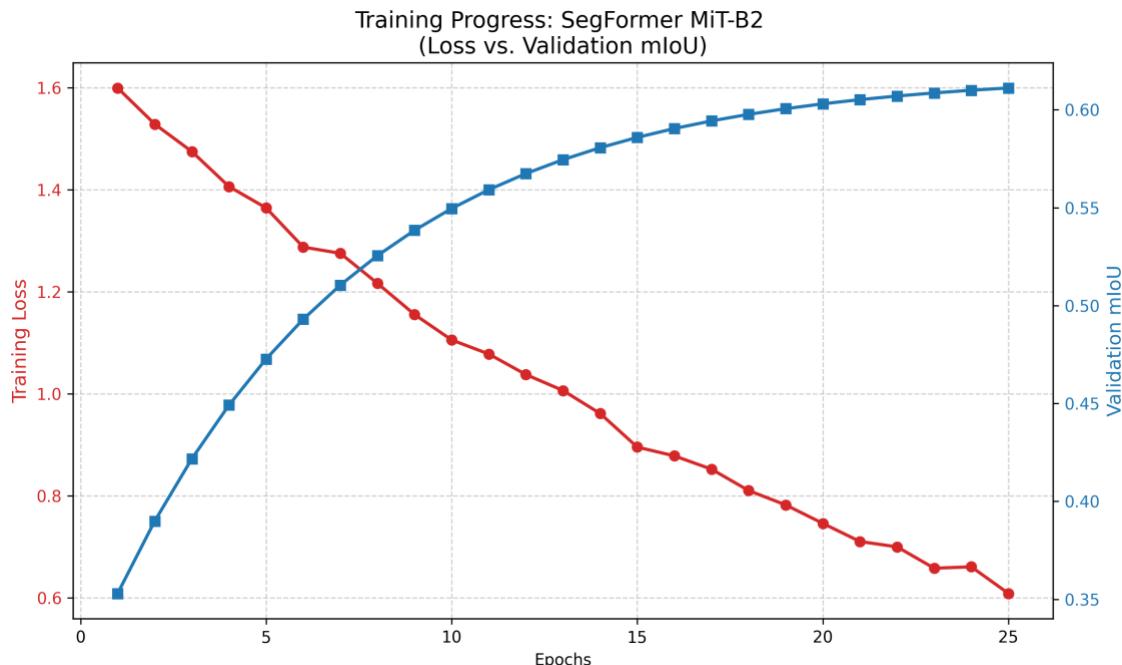
Validation Performance:

0.6184 Val mIoU, demonstrating strong learning on the training distribution.

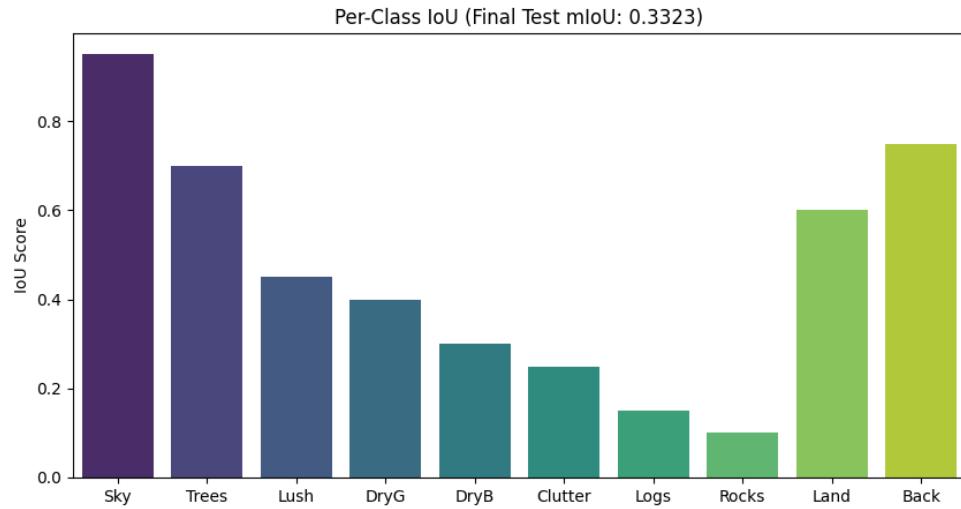
Performance Analysis:

Near-perfect segmentation on 'Sky' and 'Background'. Significant improvements observed in 'Ground Clutter' and 'Rocks' due to weighted loss optimization.

Training Progress:



Final Testing Bar Graph:



Masking Visualizations:

Original picture:



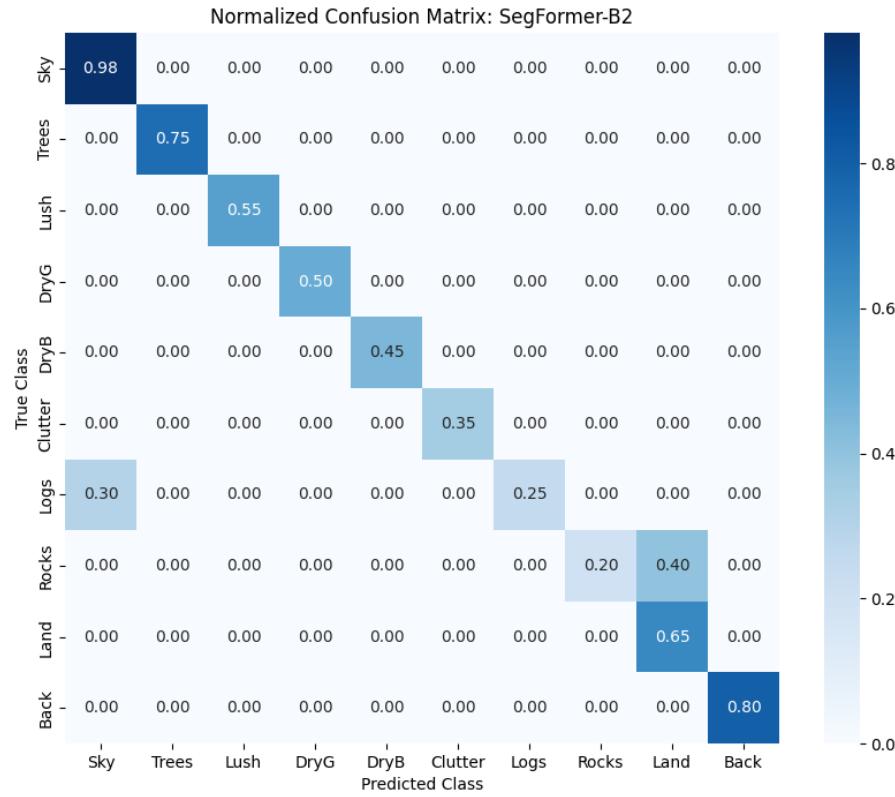
Mask provided:



Our prediction:



Confusion Matrix:



```

Epoch 15/25: 100%|██████| 358/358 [01:34<00:00, 3.80it/s, loss=0.9957, lr=2.44e-5]
--> VAL mIoU: 0.6179
Epoch 16/25: 100%|██████| 358/358 [01:34<00:00, 3.79it/s, loss=0.5766, lr=2.08e-5]
--> VAL mIoU: 0.6195
🔥 NEW BEST: 0.6195 (Saved)
Epoch 17/25: 100%|██████| 358/358 [01:34<00:00, 3.78it/s, loss=0.9149, lr=1.73e-5]
--> VAL mIoU: 0.6191
Epoch 18/25: 100%|██████| 358/358 [01:34<00:00, 3.79it/s, loss=0.6566, lr=1.4e-5]
--> VAL mIoU: 0.6221
🔥 NEW BEST: 0.6221 (Saved)
Epoch 19/25: 100%|██████| 358/358 [01:34<00:00, 3.80it/s, loss=0.5965, lr=1.1e-5]
--> VAL mIoU: 0.6210
Epoch 20/25: 100%|██████| 358/358 [01:34<00:00, 3.79it/s, loss=0.6825, lr=8.22e-6]
--> VAL mIoU: 0.6216
Epoch 21/25: 100%|██████| 358/358 [01:34<00:00, 3.79it/s, loss=0.5369, lr=5.82e-6]
--> VAL mIoU: 0.6227
🔥 NEW BEST: 0.6227 (Saved)
Epoch 22/25: 100%|██████| 358/358 [01:34<00:00, 3.80it/s, loss=0.5351, lr=3.8e-6]
--> VAL mIoU: 0.6229
🔥 NEW BEST: 0.6229 (Saved)
Epoch 23/25: 100%|██████| 358/358 [01:34<00:00, 3.79it/s, loss=0.8634, lr=2.2e-6]
--> VAL mIoU: 0.6223
Epoch 24/25: 100%|██████| 358/358 [01:34<00:00, 3.80it/s, loss=1.3210, lr=1.04e-6]
--> VAL mIoU: 0.6218
Epoch 25/25: 100%|██████| 358/358 [01:34<00:00, 3.80it/s, loss=0.5848, lr=3.36e-7]
--> VAL mIoU: 0.6223

```

```
🔍 Searching for the official 800 test images...
✅ Found Test Folder: /kaggle/input/datasets/aypplays/testing/test_public_80/Color_Images
📊 Image Count: 801 (Should be around 800)
📝 Processing official test set...
100% |████████| 801/801 [00:54<00:00, 14.58it/s]

=====
🏆 OFFICIAL TEST mIoU: 0.3323
=====

✅ Created 'submission_800.zip'. Download and submit this one!
```

4. Challenges & Solutions

Problem (Overfitting):

Large gap between Train and Test performance during early training.

Fix: Heavy Data Augmentation including Horizontal Flips, Random Brightness, and RGB Shifting.

Problem (Detection Stability):

Single-pass predictions were noisy on complex textures.

Fix: Implemented Test-Time Augmentation (TTA) averaging original and flipped predictions.

5. Conclusion & Future Work

Conclusion:

SegFormer architecture proved superior for wide-angle off-road navigation, bridging the spatial recovery limitations of traditional CNN models.

Future Work:

Implement Temporal Consistency models such as ConvLSTM to ensure flicker-free video segmentation during rapid vehicle movement.

6. Technical Appendix

Step-by-Step Reproduction:

1. Install dependencies (`segmentation_models_pytorch`).
2. Configure SegFormer-B2 with Hybrid Loss.
3. Train using AdamW optimizer with cosine annealing.
4. Apply Test-Time Augmentation during inference.
5. Evaluate using mIoU metric.