

Version 4 (V4) – Final Strategy Outline

Goal:

Build a robust, interpretable, and spatially accurate off-road semantic segmentation model.

1. Problem Re-Definition

Previous versions showed that pixel accuracy and loss engineering alone cannot solve semantic confusion.

V4 reframes the task as hierarchical semantic understanding, prioritizing spatial coherence over fine-grained noise.

2. Dataset Strategy

Classes are regrouped into five super-classes:

Sky, Vegetation, Terrain, Objects, Background.

This reduces label noise, improves class balance, and stabilizes learning.

3. Model Architecture

Encoder: ResNet50 (pretrained)

Context Module: DeepLabV3+ ASPP

Decoder: Attention-UNet style refinement

This captures global context while preserving fine spatial details.

4. Loss Design

Total Loss =

$0.5 \times \text{Weighted Categorical Cross-Entropy}$

$0.3 \times \text{Dice Loss}$

$0.2 \times \text{Boundary Loss}$

Each component addresses imbalance, region consistency, and boundary precision.

5. Training Strategy

Low learning rate, pretrained backbone, early stopping, and best-model checkpointing ensure stable convergence without overfitting noisy labels.

6. Evaluation Protocol

Avoid pixel accuracy.

Use Mean IoU, per-class IoU, and qualitative visual inspection.

7. Expected Outcomes

Target Mean IoU: 0.45–0.60

Improved boundary clarity and reduced class collapse.

8. Key Insight

V4 works because data, architecture, loss, and evaluation are aligned.

This is principled ML engineering.