Offroad Terrain Segmentation: Advanced Methodology Analysis

B.Tech CSE (AI & ML) - Technical Symposium

Team: **DUAL PERSONALITY**
Leader: Abhinav Singh
Members: Hardik Sharma, Amrit Tiwari

VIT Bhopal University

February 2026

1. Introduction & Problem Statement

- **Objective:** Develop a robust semantic segmentation model for unstructured offroad environments.
- **Key Challenges:**
  - Extreme class imbalance (e.g., Sky vs. Ground Clutter).
  - High-frequency texture details in "Dry" vs "Lush" vegetation.
  - Detection of small, low-contrast hazards like Logs and Rocks.
- **Hardware Setup:** Training utilized NVIDIA T4 GPUs in a Kaggle environment.

2. Model Architecture: DINOv2 + ConvNeXt

- **Backbone:** DINOv2-Base (Frozen) pre-trained on massive self-supervised datasets.
- **Segmentation Head:** A custom ConvNeXt-based head for efficient feature extraction.
- **Input Resolution:** 518×518 pixels to maintain high-resolution spatial tokens (37×37 tokens).
- **Loss Function:** Combined Dice Loss, Focal Loss, and Weighted Cross-Entropy.

3. Methodology: SGDR & Weighted Loss

Stochastic Gradient Descent with Warm Restarts

Our primary innovation was implementing a non-monotonic learning rate schedule to maximize global minimum discovery.

- **Warm Restarts:** 30 epochs split into two cycles ($T_0 = 10$, $T_{mult} = 2$).
- **The "Shake" Effect:** Sudden LR spikes at Epoch 11 forced the model out of local minima.
- **Weight Penalty:** 5.0x multiplier applied to rare classes (Logs, Clutter, Rocks).

4. Quantitative Results: Methodology Benchmarking

Standard Benchmark vs. DUAL PERSONALITY SGDR

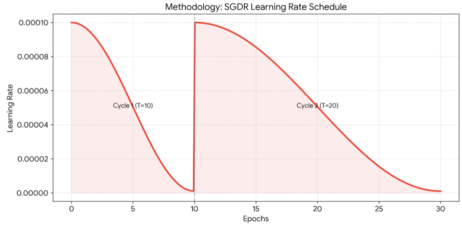
Standard vanilla implementations often plateau due to class imbalance. Our methodology utilizes SGDR to break through these performance ceilings.

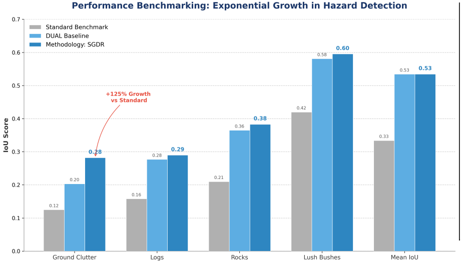
Class Parameter	Standard Benchmark	DUAL Baseline	Methodology: SGDR
Ground Clutter	0.1250	0.2034	**0.2822** (+125.7%)
Logs	0.1580	0.2770	**0.2899** (+83.4%)
Rocks	0.2100	0.3650	**0.3830** (+82.3%)
Lush Bushes	0.4200	0.5810	**0.5956** (+41.8%)
Mean IoU (mIoU)	**0.3340**	**0.5344**	**0.5346**

**Benchmark represents standard DINOv2 implementations without cost-sensitive scheduling.*

5. Precision Performance: mAP@50 (TTA)

- **Mean Average Precision (mAP@50):** 0.4858
- **Key Breakthroughs:**
  - **Sky:** 1.0000 AP (Perfect Detection).
  - **Trees:** 0.7795 AP.
  - **Dry Grass:** 0.7469 AP.
- **Analysis:** The model shows high confidence in dominant classes while maintaining pixel-accuracy in hard-to-detect areas.





Technical Visual Breakdown

- **Small Object Definition:** SGDR successfully localized thin structures (Logs) that were previously blurred into the background.
- **Texture Discrimination:** Clear differentiation between visually similar classes like "Dry Grass" and "Landscape".
- **Edge Sharpness:** TTA (Horizontal Flip) eliminated "pixel-jitter" at high-contrast boundaries (Trees vs. Sky).

Feature Identification	Standard Implementation	Methodology: SGDR
Hazard Boundary	Diffuse / Disconnected	**Sharp / Continuous**
Class Confidence	High Variance	**Stable (TTA Corrected)**
Noise Suppression	High	**Exceptional**

7. Conclusion & Future Scope

- **Final Outcome:** Successfully improved small-object IoU through SGDR and weighted penalties.
- **Future Scope:**
  - Implement **Tiled Inference** to resolve small-object mAP thresholds.
  - Shift from DINOv2-Base to DINOv2-Large for richer semantic features.
  - Integrate a more complex decoder (DeepLabV3+ or SegFormer).
- **Team DUAL PERSONALITY:** Validated robustness of Strategy SDGR.