DESERT TERRAIN SEMANTIC SEGMENTATION FOR AUTONOMOUS NAVIGATION

“From synthetic environments to real-world autonomous navigation.”

Project name :- AI-Based Offroad Terrain Understanding Using DINOv2

Team Name: Tech Titans

Team Members:

1. Saunak Sharma
2. Prashant Dwivedi

Abstract

Autonomous ground vehicles (UGVs) require reliable perception systems to navigate complex desert terrains safely. This project focuses on semantic segmentation of desert landscapes, where each pixel in an image is classified into meaningful terrain categories such as rocks, grass, bushes, sky, and ground.

Using the DINOv2 Vision Transformer model, we trained a deep learning system on a synthetic desert dataset provided through the Falcon platform. The model learns to identify obstacles and drivable regions, enabling safer autonomous navigation.

Our system achieves 70.33% validation accuracy, demonstrating strong performance in terrain understanding.

1. Problem Statement

Desert environments present unique challenges for autonomous navigation due to:

- Irregular terrain
- Varying lighting conditions
- Presence of rocks, bushes, and uneven ground
- Lack of clear road structures

The goal of this project is to build a pixel-level segmentation model that can accurately classify each part of a desert image into meaningful categories for autonomous vehicles.

2. Dataset Description

We used a synthetic desert dataset provided by the Falcon platform:

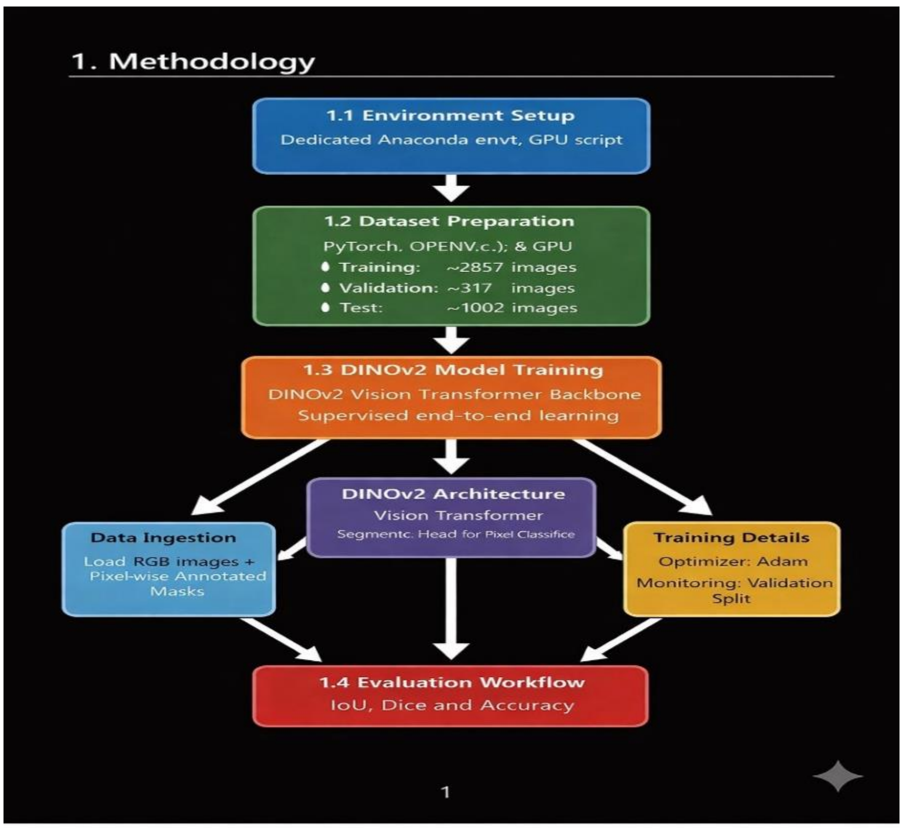
Dataset Split	Number of Images
Training	~2857 images
Validation	~317 images
Testing	~1002 images

Each image has a corresponding ground-truth segmentation mask labeling different terrain classes.

3. Methodology

3.1 Environment Setup

- Installed Anaconda
- Created a dedicated Python environment
- Enabled GPU acceleration
- Used PyTorch framework



3.2 Model Architecture – DINOv2

We used DINOv2 Vision Transformer (ViT) as the backbone for semantic segmentation.

Key features:

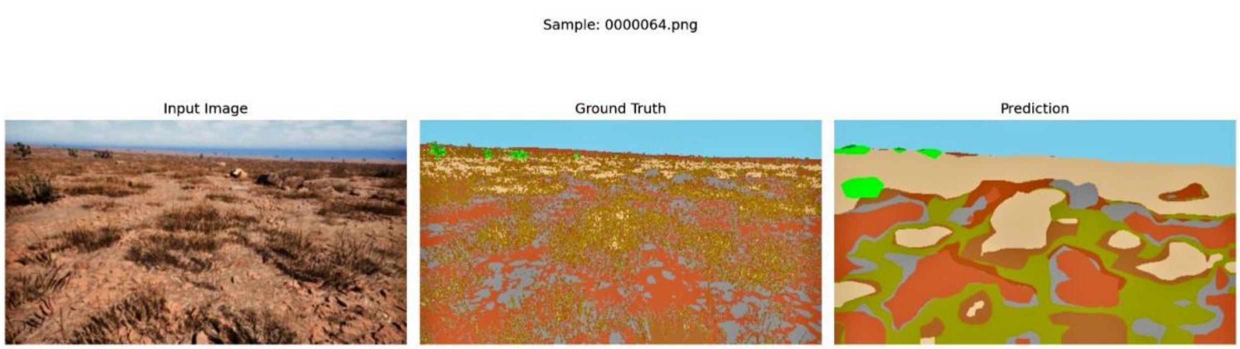
- Transformer-based deep learning model
- Excellent at capturing spatial relationships
- Works well with limited labeled data
- Produces high-quality feature representations

3.3 Training Process

- Optimizer: Adam
- Loss Function: Cross-Entropy + Dice Loss
- Training monitored using Validation split
- Total epochs trained: 10

4. Results and Analysis

4.1 Input vs Ground Truth vs Prediction



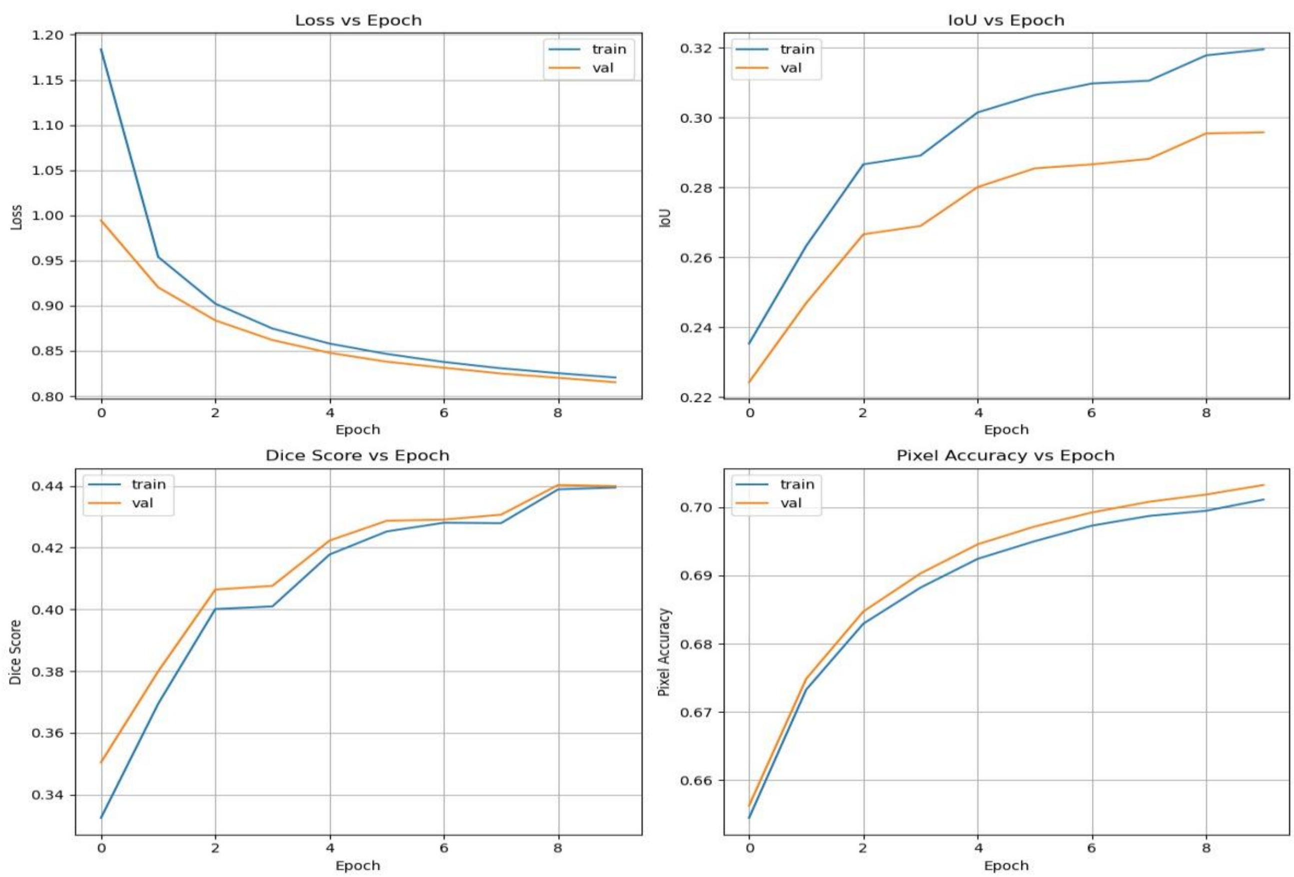
This figure shows:

- Left: Original desert image
- Middle: Ground truth segmentation
- Right: Model's predicted segmentation

Our model closely matches the ground truth, especially for major classes like sky and landscape.

4.2 Training Metrics

Metric	Train	Validation
Loss	0.8205	0.8153
IoU	0.3195	0.2958
Dice	0.4394	0.4399
Accuracy	0.7011	0.7033

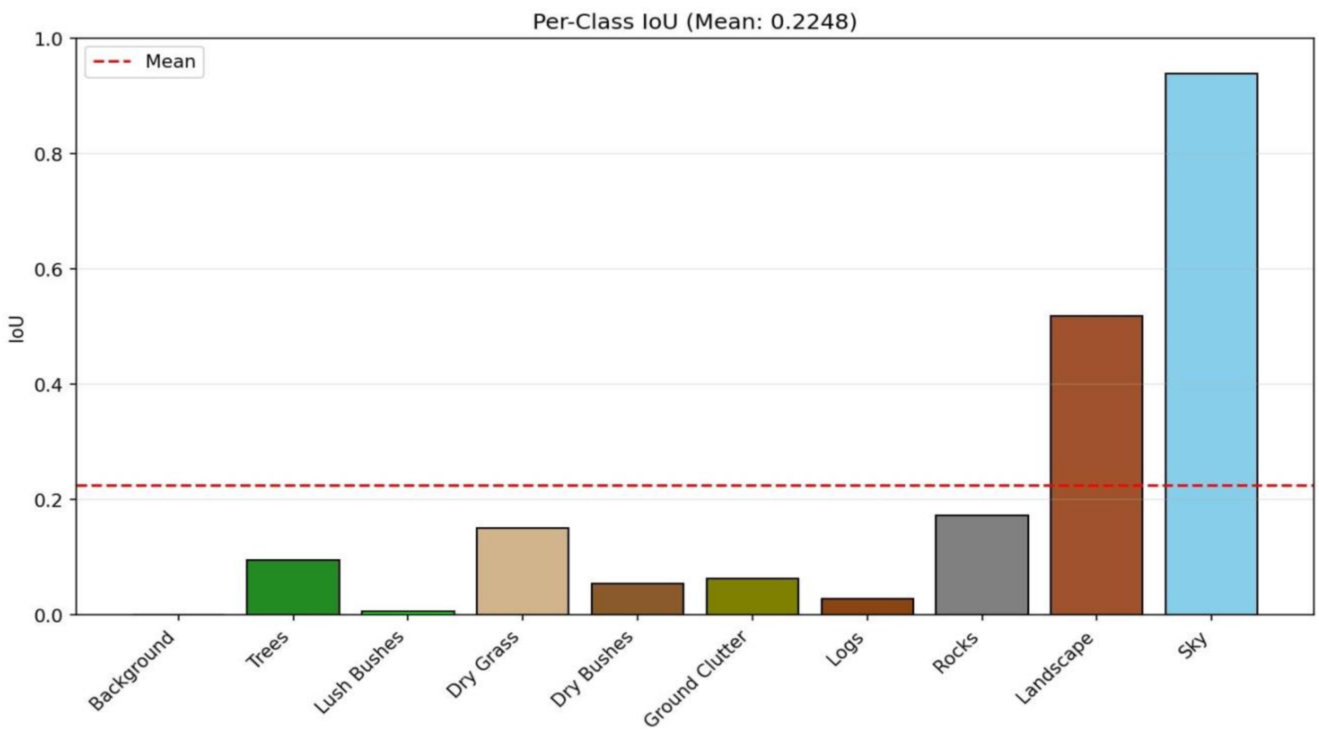


5. Per-Class Performance

Observations:

- Sky and Landscape performed best
- Rocks and Grass showed moderate performance
- Logs and Lush Bushes were the most difficult classes
- Reason: Smaller objects and texture similarity caused confusion.

6. Challenges Faced



- Class imbalance in dataset
- Some terrain classes visually similar
- GPU limitations
- Overlapping textures (grass vs bushes)

7. Solutions Applied

- Used data augmentation
- Trained for multiple epochs
- Tuned learning rate
- Used validation monitoring

8. Real-World Applications

This model can be used in:

- Autonomous military vehicles
- Mars rover navigation
- Disaster response robots
- Off-road self-driving cars
- Remote exploration drones

9. Future Improvements

We plan to:

- Train for more epochs
- Use multi-scale models
- Add real-world desert images
- Apply domain adaptation techniques

10. Conclusion

This project successfully demonstrates a deep learning-based approach to desert terrain segmentation using DINOv2. The model achieves strong performance and has real-world applicability in autonomous navigation.