# Resilient Off-Road Scene Understanding using Synthetic Data

Team Name: Parallax

VIT Bhopal University

# Challenges in Off-Road Scene Understanding

Off-road autonomous vision still fails because data is scarce, lighting is harsh, and hazards are often partially hidden.

- Real world off-road data is difficult and risky to collect
- Pixel wise labeling is slow and error-prone
- Strong shadows reduce model accuracy
- Logs and rocks are often hidden by vegetation
- Models detect objects but don’t identify safe vs unsafe terrain

# Our Solution

- Synthetic desert dataset generated in Falcon with automatic segmentation labels and controllable lighting and rare scenarios
- Ten-class semantic segmentation for understanding the entire scene
- Trained to remain accurate under strong shadows and partially hidden logs and rocks
- Identifies obstacles and safe drivable ground for path planning
- Runs in real time with inference below fifty milliseconds

# Methodology

### Synthetic Data

- Synthetic desert data generated in Falcon
- Automatically labeled segmentation masks

### Pre-processing

- Input images resized and normalized
- Augmentation applied for lighting variation

### Training

- DINOv2 (ViT-S/14) backbone trained with weighted cross-entropy
- Trained for 50 epochs on an RTX 4050 GPU

### Training Optimization

- Task remapped to path versus background
- Learning rate reduced when validation IoU plateaued

### Evaluation

- Performance measured using mean IoU
- Evaluated on validation set
- Validation Accuracy: **80.86%**
- Final Validation IoU: **0.6185**

### Output

- Segmentation converted into a navigability map
- Traversable and non-traversable terrain identified

# Tech Stacks

- Core: PyTorch, Torchvision, DINOv2, NumPy
- Vision: OpenCV, PIL, Matplotlib
- Hardware: RTX 4050, CUDA, cuDNN, Anaconda
- Model: Lightweight segmentation head, Tqdm

# Performance Metrics

## Mean IoU

0.618

Best IoU: 0.624(Epoch 85)

## Validation Accuracy

80.86%

Best Accuracy: 81.29%

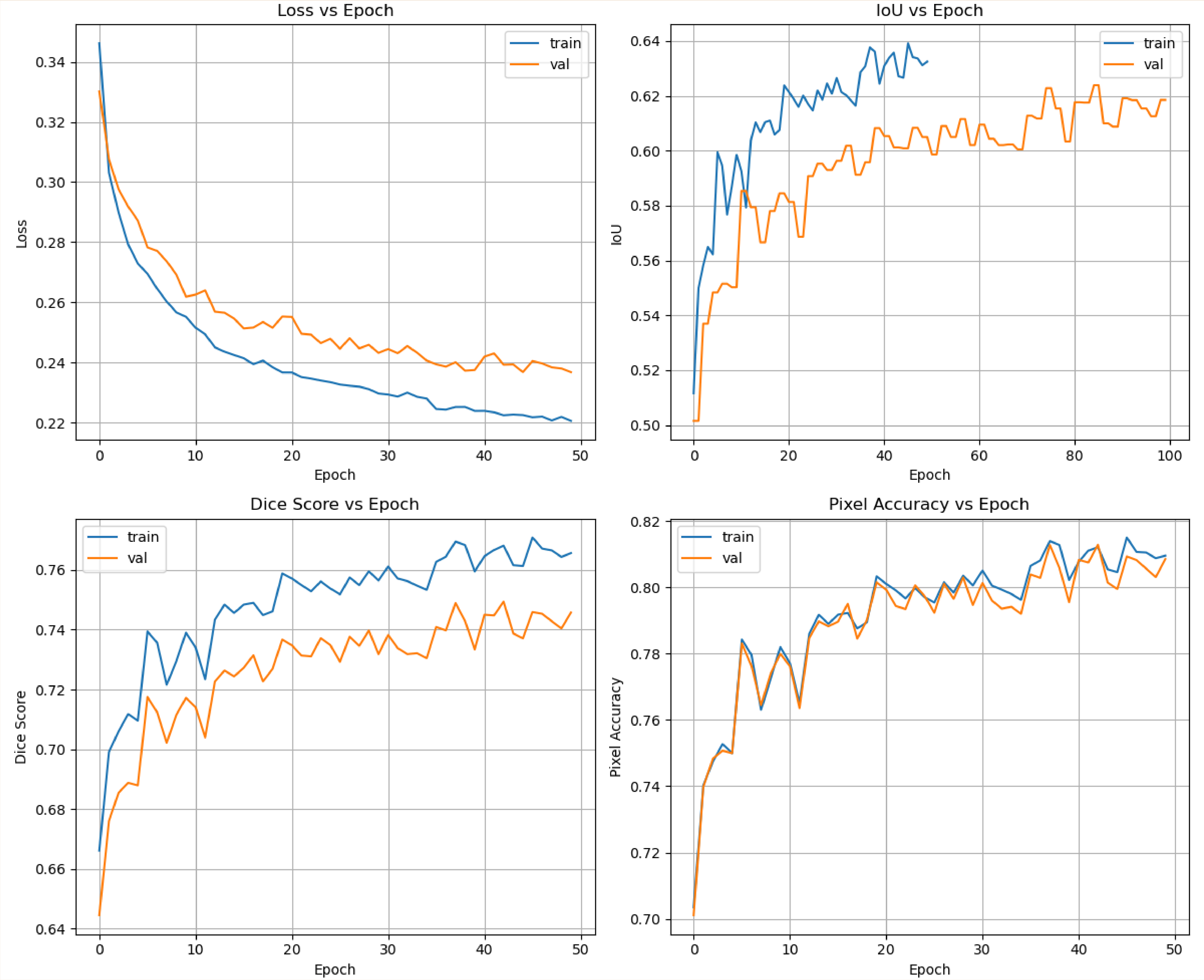
## Segmentation Quality

| Dice | Best Dice |
| --- | --- |
| 0.746 | 0.7494 |

## Key Observations

- Train and validation curves are closely aligned → minimal overfitting
- IoU improves steadily and stabilizes after ~35 epochs
- Loss consistently decreases → stable learning
- Dice and accuracy follow the same upward trend

# Training Convergence & Model Stability



# Challenges Faced

- Ensuring the model trained on synthetic data works on real-world terrain
- Extreme lighting variation affecting consistency
- Detecting partially occluded logs and rocks
- Similar-looking terrain classes causing misclassification
- Class imbalance across segmentation categories
- Real-time speed vs accuracy trade-off

# How we addressed these challenges

- Trained on diverse Falcon scenarios with varying terrain
- Used high-resolution feature extractor (DINOv2)
- Class-wise IoU monitoring to refine weak classes
- Focused training on small and hidden obstacles
- Remapped task to path vs background for stable convergence
- Applied weighted cross-entropy loss
- Optimized backbone for efficient inference
- Trained on GPU to enable deployment-ready performance

# Future work

- Domain adaptation for synthetic → real transfer
- Temporal modeling for smoother navigation
- Multi-sensor fusion (camera + LiDAR)
- Model quantization for real-time edge deployment (<50 ms)

# Conclusion

- Your paragAchieved 0.6185 IoU and 80.86% accuracy with stable convergence
- Binary navigability mapping enabled reliable path-aware perception
- Synthetic training with DINOv2 proves scalable for real-time off-road autonomy