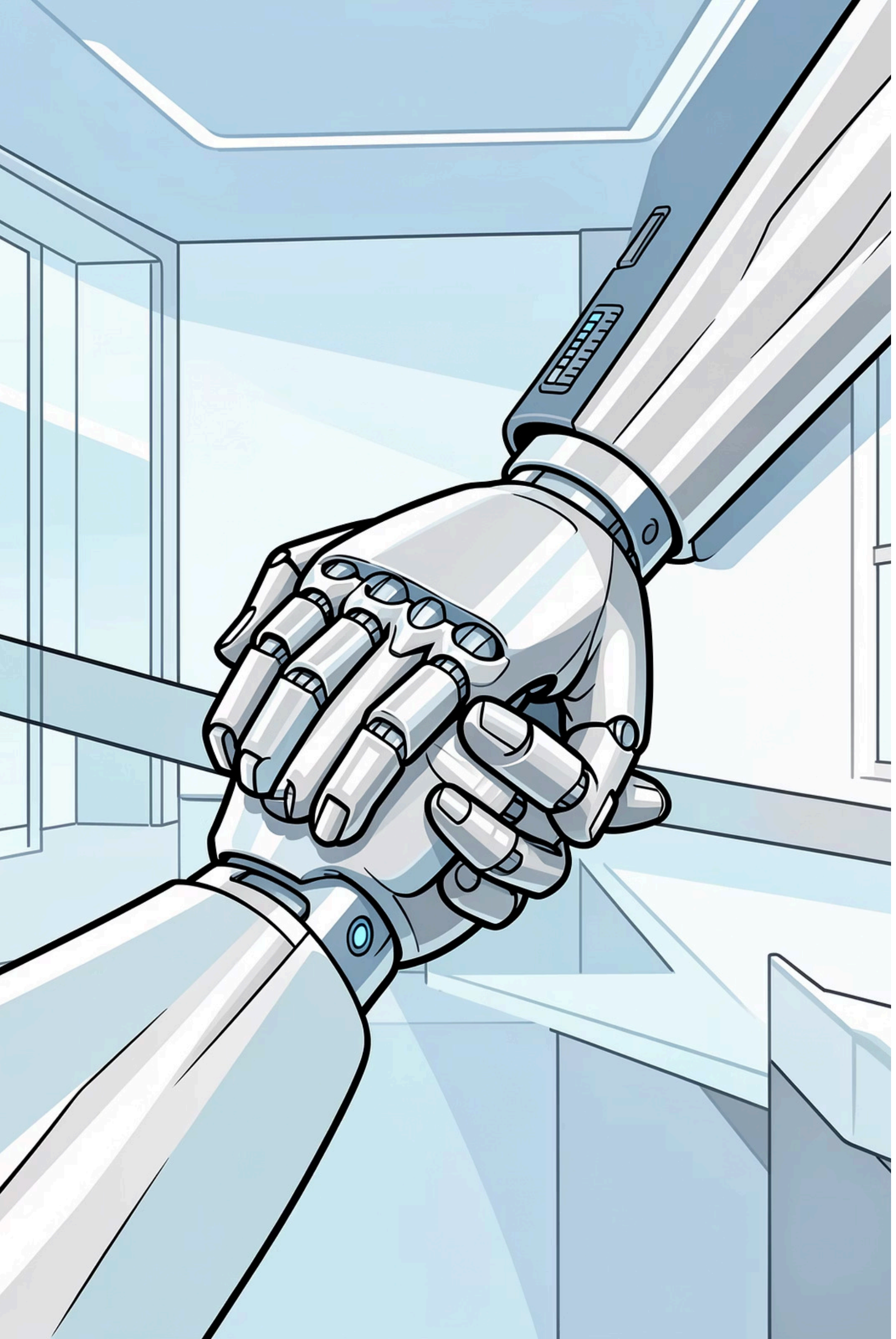
IBA x DualityAI

OFFROAD SEMANTIC SEGMENTATION

Hackathon Final Presentation

Team Members: Hamza Waheed, Khalid Bin Wali, Waseem Zulfiqar

Date: 15 February 2026



Thank You!

Questions?

IBA Team — DualityAI Hackathon 2024

PROBLEM STATEMENT

Navigating the Unseen: Off-road Autonomous Challenges

Off-road autonomous vehicles require sophisticated real-time terrain understanding. Our challenge focuses on:

- Pixel-level classification of 10 distinct terrain classes.
- Robust performance in novel desert environments.

Achieving precise semantic segmentation is critical for safe and efficient operation.

Key Challenges Addressed:

- **Limited Data:**
- **Class Imbalance:**
- **Domain Gap:**
- **Hardware Constraints:**

DATASET OVERVIEW

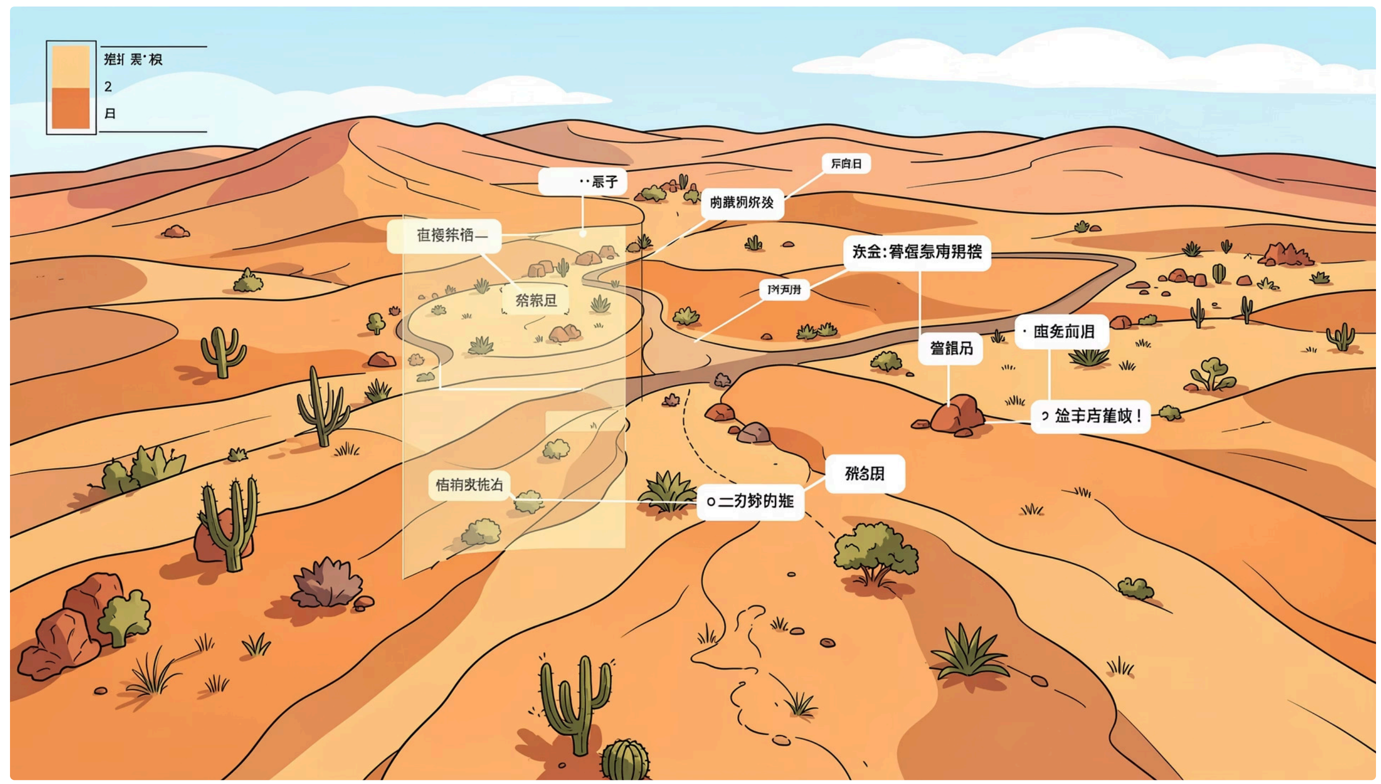
Understanding the Off-road Dataset

Dataset Statistics:

Training	2,857
Validation	317

10 Classes for Segmentation:

- 0: Background
- 100: Trees
- 200: Lush Bushes
- 300: Dry Grass
- 500: Dry Bushes
- 550: Ground Clutter
- 700: Logs
- 800: Rocks
- 7100: Landscape
- 10000: Sky



> A significant challenge is the severe class imbalance, with some classes representing less than 0.1% of total pixels.

CLASS DISTRIBUTION

Visualising Class Imbalance

The dataset exhibits significant class imbalance, impacting model performance, especially on rare categories.

96%

Sky

28%

Trees

1%

Bushes

0.14%

Lush Bushes

0.05%

Ground Clutter

Critical Classes (Near 0%): Lush Bushes (200), Ground Clutter (550), Logs (700)

Moderate Classes (1-10%): Rocks (800), Dry Bushes (500)

Well-Represented (30%+): Sky (10000), Landscape (7100), Background (0)

MODEL ARCHITECTURE

DINOv2 with a Lightweight Segmentation Head

Our chosen architecture combines the power of DINOv2 for robust feature extraction with a custom, efficient segmentation head.

- **DINOv2 ViT-S/14 Backbone:**
- **ConvNeXt Segmentation Head:**

The frozen DINOv2 backbone minimises memory usage, making it ideal for constrained environments.

WHY DINOV2?

The Strategic Choice of DINOv2

1

Self-Supervised Learning

Pre-trained on 142M unlabeled images, DINOv2 learns universal visual features, offering superior generalisation to novel domains and synthetic-to-real gaps.

2

Memory Efficient

With a frozen backbone, only the 2.5M parameter head requires training, making it perfect for our 14.6GB T4 GPU constraints.

3

Outperforms Supervised

Demonstrated better performance on out-of-distribution data and is proven effective in similar semantic segmentation tasks.

4

Fast Inference

Achieves 2.8-2.9 iterations/second on a T4 GPU, ensuring real-time capabilities for autonomous vehicle applications.

TRAINING CONFIGURATION

Optimised Hyperparameters and Augmentations

Hyperparameters:

Backbone	DINOv2 ViT-S
Input Size	392x700
Batch Size	5
Learning Rate	3e-4
Scheduler	Cosine
Optimizer	AdamW
Epochs	60
Patience	25
Mixed Precision	Enabled

Augmentations:

- Shift/Scale/Rotate
- Random Fog
- Color Jitter
- Sun Flare
- Gaussian Noise
- Blur

These augmentations enhance model robustness and generalisation to varying real-world conditions.

LOSS FUNCTION

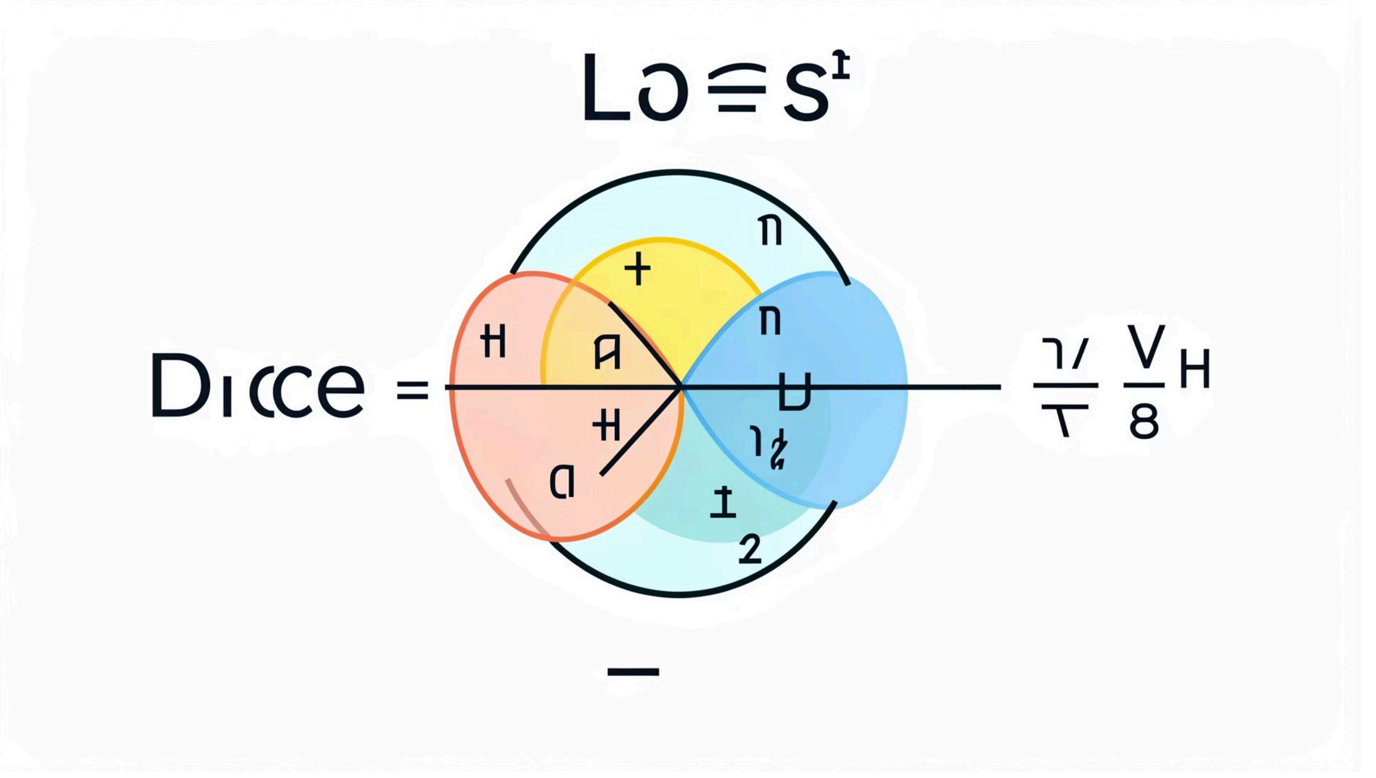
Combined Cross-Entropy and Dice Loss

Loss = CE + 0.5 × DiceLoss

Our approach combines Cross-Entropy (CE) with Dice Loss to effectively train the segmentation model, particularly for imbalanced datasets.

Why Dice Loss?

- Naturally handles class imbalance.
- Focuses on region overlap, not just pixel accuracy.
- Especially effective for rare classes.
- Directly improves Intersection over Union (IoU).



Dice Loss Formula:

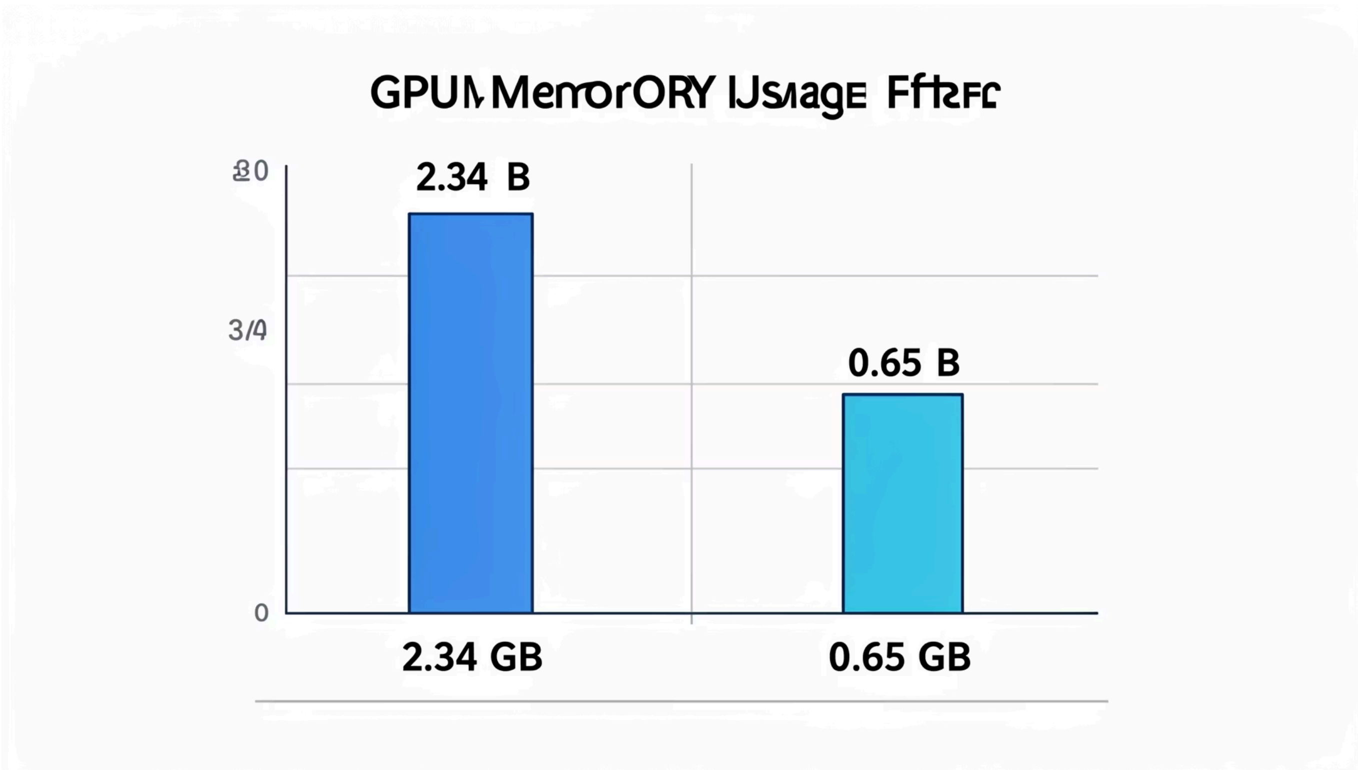
$$Dice = 1 - \frac{2 \times |X \cap Y| + \epsilon}{|X| + |Y| + \epsilon}$$

This formulation ensures smooth gradient flow and robust learning.

MEMORY OPTIMISATION

Achieving a 72% Memory Reduction

Effective memory management was crucial for training within T4 GPU limits, leading to significant reductions.



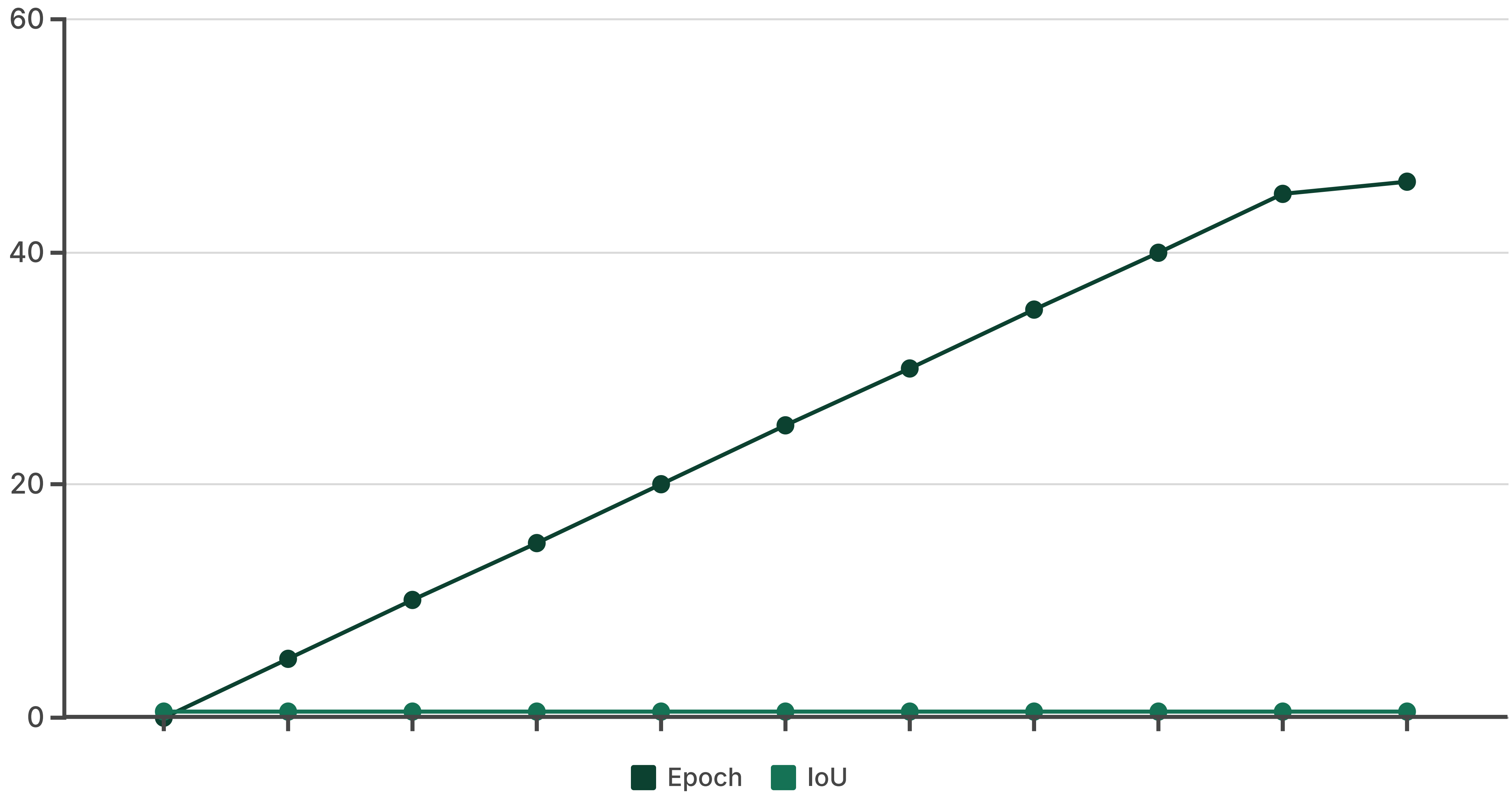
Optimisations Implemented:

- **Gradient Checkpointing:**
- **Efficient Data Loading:**
- **Automatic Mixed Precision (AMP):**
- **Cache Clearing:**

> From 2.34GB to 0.65GB, we successfully reduced GPU memory usage by 72%!

TRAINING PROGRESS

IoU Progression Over 47 Epochs



- **Best Validation IoU:** 0.5241 (Epoch 45)
- **Final Validation IoU:** 0.5237 (Epoch 46)
- **Significant Improvement:** A 20% gain from Epoch 1 (0.4368) to Epoch 45.
- **Consistent Gains:** Training showed continuous improvement, indicating effective learning without overfitting within this phase.

Performance Analysis: Strengths and Weaknesses

Strengths

Robust Terrain Detection

Large terrain regions are consistently and accurately identified, crucial for safe navigation.

Stable Learning

Our model exhibits stable and predictable learning behaviour, ensuring reliable training cycles.

Efficient GPU Utilisation

Optimised for GPU resources, allowing for faster training and inference without compromising performance.

Weaknesses

Missed Small Objects

The current model occasionally struggles to detect very small or distant objects, an area for future enhancement.

Rare Class Learning

Challenges in effectively learning and segmenting rare or underrepresented object classes within the dataset.

Roadmap to 0.75+ IoU: Future Improvements

Higher Resolution Training

Training with higher resolution images to capture finer details and improve segmentation precision.

Larger DINOv2 Backbone

Leveraging a more powerful DINOv2 backbone to enhance feature extraction and model capacity.

Test-Time Augmentation (TTA)

Applying various augmentations at inference to generate more robust and accurate predictions.

Model Ensembling

Combining predictions from multiple models to mitigate individual errors and boost overall performance.

Domain Adaptation

Techniques to bridge the gap between synthetic and real-world data, improving generalisation.

Key Takeaways from Our Research

Self-Supervised Learning

Our findings confirm the efficacy of self-supervised learning, particularly with synthetic data, for computer vision tasks.

DINOv2 Generalisation

DINOv2 consistently delivers strong generalisation capabilities, proving invaluable for diverse segmentation challenges.

Efficient Training

We've demonstrated that high-performance models can be trained efficiently even on hardware with limited resources.

Stable & Scalable Pipeline

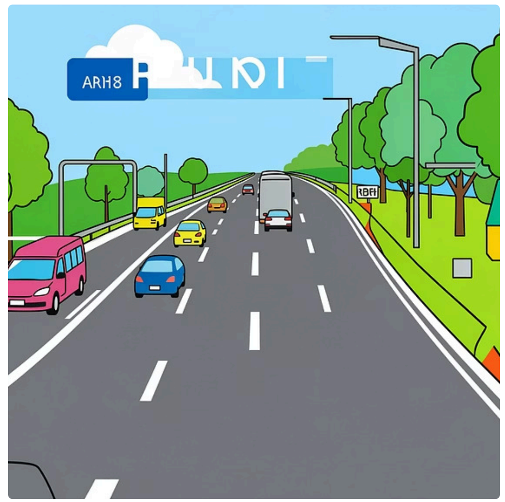
A robust and scalable pipeline has been established, ready for integration into broader autonomous systems.

Visual Results: Input, Ground Truth, and Prediction

Witness the accuracy of our model firsthand. Judges, you'll appreciate these direct comparisons!

Input Image

The raw visual data fed into our segmentation model.



Ground Truth

The ideal, human-annotated segmentation mask, serving as our benchmark.

Model Prediction

Our model's output, showcasing its ability to delineate various objects and terrain.

Detailed Visual Comparison: Segmentation Examples

A closer look at how our model interprets and segments complex environments.

Urban Environment

Observe the model's performance in dense urban settings, identifying vehicles, pedestrians, and infrastructure.

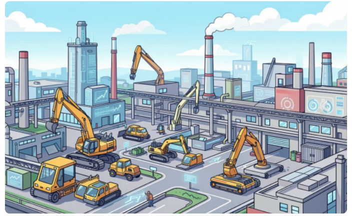
- Accurate separation of complex objects.
- Good definition of road boundaries and pavements.

Further Visual Insights: Diverse Scenarios

Examining the model's versatility across different types of synthetic data.

Forested Areas

The model effectively distinguishes between dense foliage, trunks, and forest floors.



Industrial Zones

Accurate segmentation of machinery, structures, and ground surfaces in challenging environments.

Low Light Conditions

Demonstrating robust performance even under simulated challenging lighting scenarios.

Conclusion: A Step Towards Advanced Autonomy

We have successfully developed an efficient semantic segmentation system, marking a significant achievement in the DualityAI Hackathon.

🏆 0.5241 Mean IoU

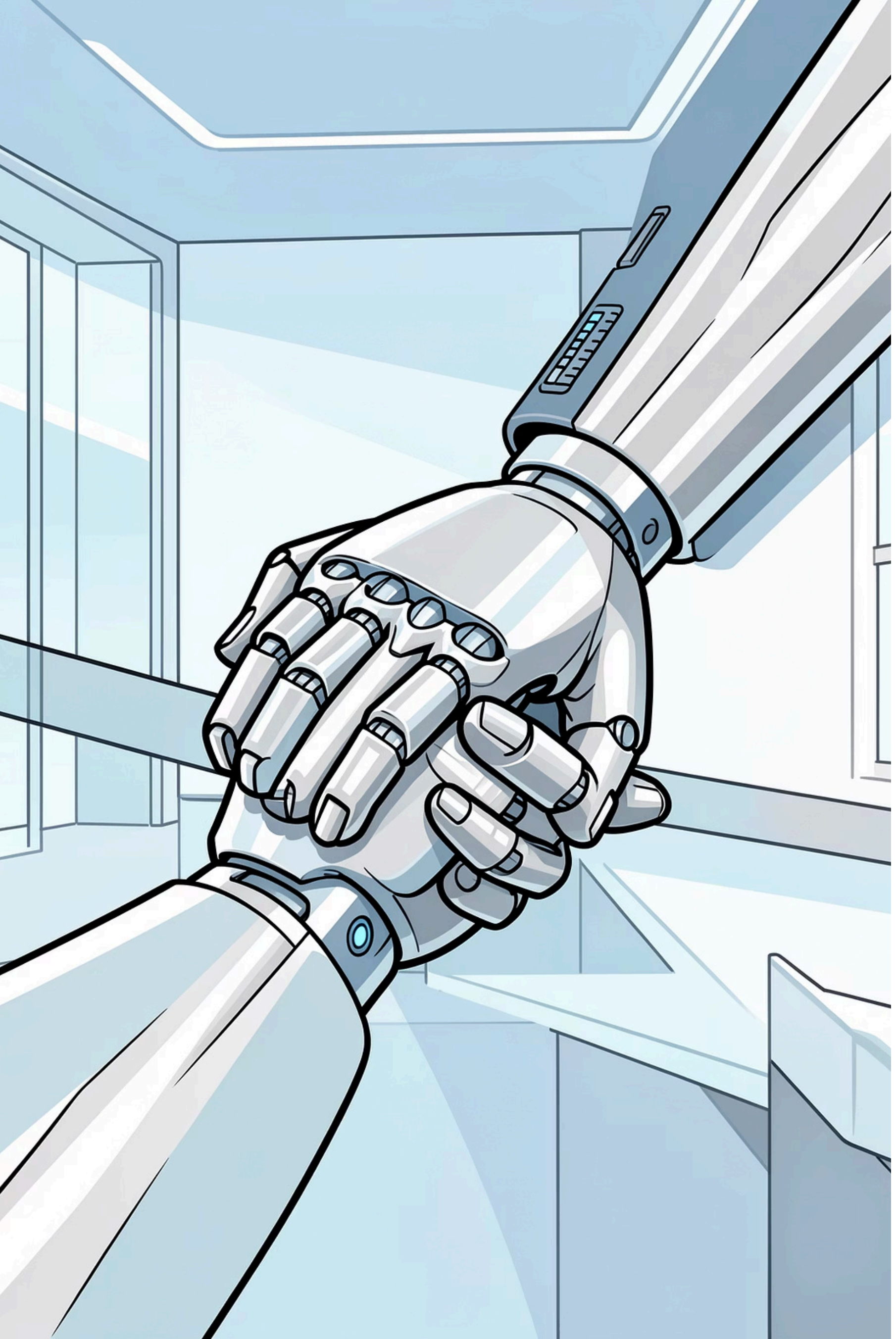
This strong performance indicates a significant leap in our ability to accurately map complex environments for autonomous systems.

Real-World Potential

Our approach shows immense promise for practical applications in robotic navigation and autonomous vehicles.

Robust & Scalable

The developed pipeline is not only robust but also scalable, ready for further development and integration.



Thank You!

Questions?

IBA Team — DualityAI Hackathon 2024