Interreg North-West Europe Co-funded by the European Union
FlashFloodBreaker

Toward Occlusion-Robust River Segmentation for Flash Flood Forecasting: A Comparative Study of CNN-Based Architectures on UAV RGB Imagery

Hasitha Adikari[1,2*] Christian O'Leary[2] Joe Harrington[1] Conor Lynch[2]
[1]Sustainable Infrastructure Research & Innovation Group, Munster Technological University, Cork, Ireland
[2]Nimbus Research Centre, Munster Technological University, Cork, Ireland
*Corresponding Author: adikari.adikari@mymtu.ie

Introduction

- Flash floods require rapid and accurate mapping of river water extent.
- UAV imagery provides high resolution but suffers from canopy occlusion, shadows, reflections, and complex riverbanks.
- Deep learning segmentation helps, but most studies ignore occlusion in small rivers.

Contributions

- UAV dataset with 416 annotated images under canopy occlusion.
- Benchmark of 10 CNN architectures with 40 training variations.
- Hybrid BCE+Dice loss and conditional learning rate scheduling.
- Explainability with Grad-CAM and Grad-CAM++.
- Direct link to flash flood forecasting needs.

Study Area

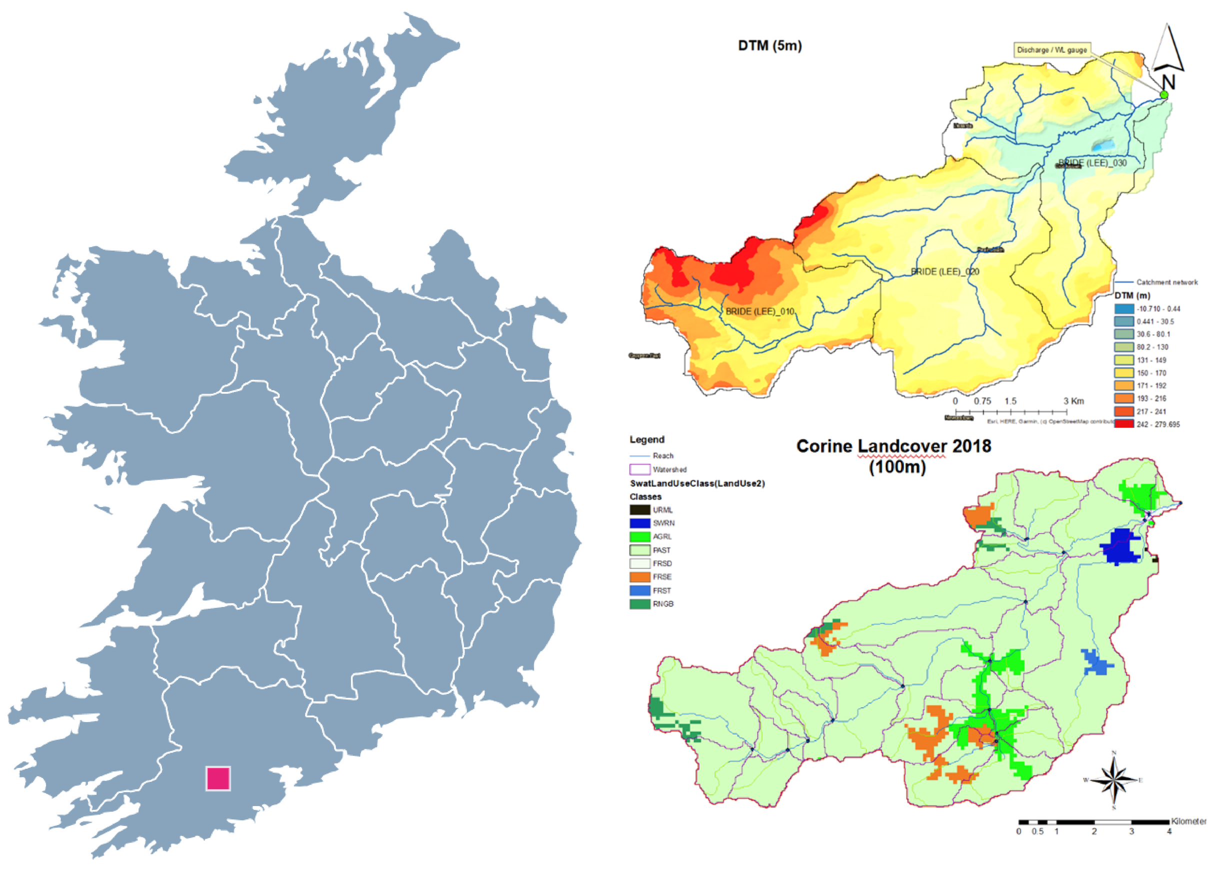
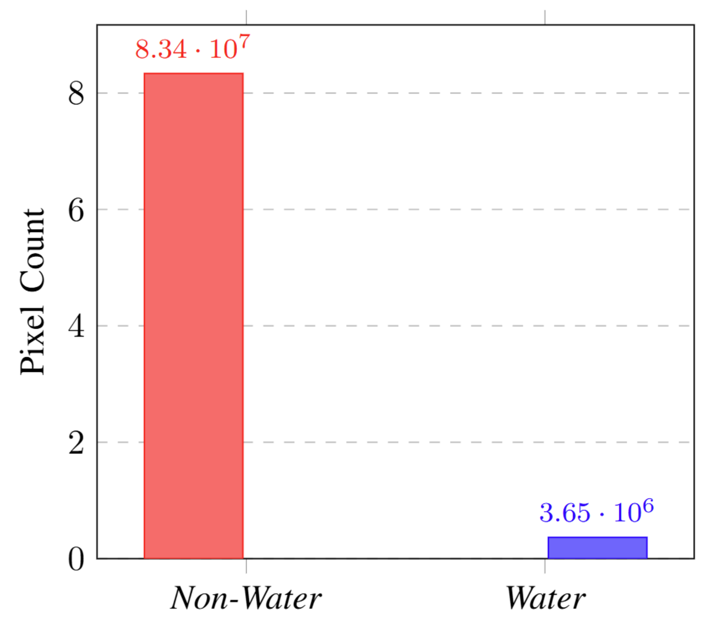


Figure 1: Study area: Bride River catchment, County Cork, Ireland.

- 70.21 km^2 in the Crookstown area in the valley of the Bride River basin.
- Recurrent flash flooding events, such as the storms of 2012 and 2015.
- Steep topography and narrow valleys that accelerate run-off.

Dataset

- 4 Drone surveys between March and June 2025.
- 2232 images ($\sim 20GB$), total $\sim 8km^2$ covered.
- 416 RGB UAV images (5280×3956 → resized 512×512).
- Class imbalance: $\sim$ 96% non-water vs $\sim$ 4% water (Pixel class distribution).



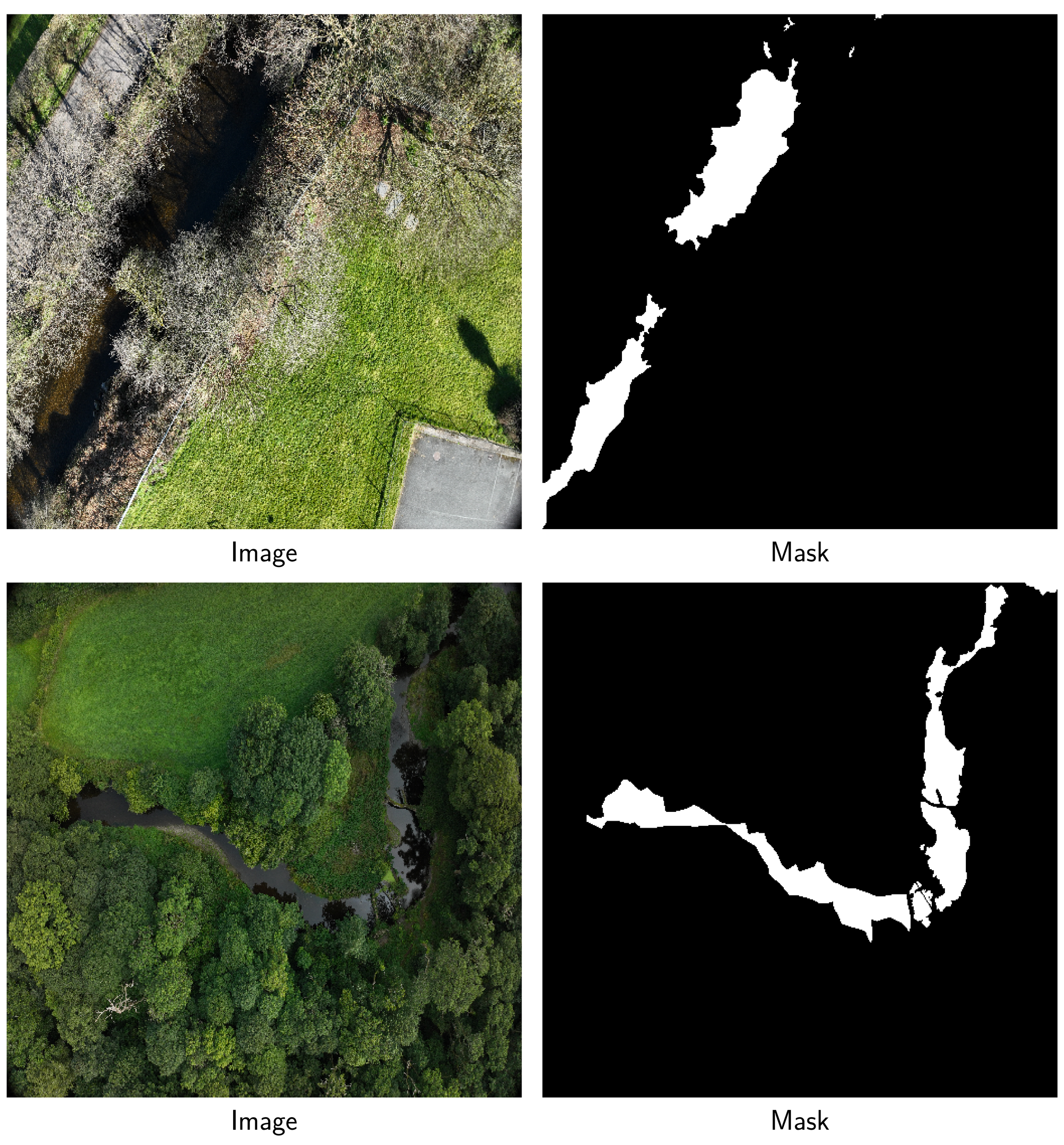


Figure 2: Sample image-masks pairs from the dataset

Methodology

- Workflow:
  - Preprocess and augment 512x512 tiles.
  - Train CNN models with BCE+Dice, Adam or AdamW.
  - Conditional LR scheduler (ReduceLROnPlateau).
  - Evaluate segmentation accuracy and efficiency.
  - Interpretability and explainability.
- 10 architectures:
  - UNET, UNET++, ResUNET++, UNET-FFC, UNET-VGG16, UNET-ResNet50, UNET-MobileNetV2, SegNet, SegNet-VGG16, DeepLabV3+.
- 40 variations: Adam vs AdamW combined with CLRS on vs off.

Augmentation

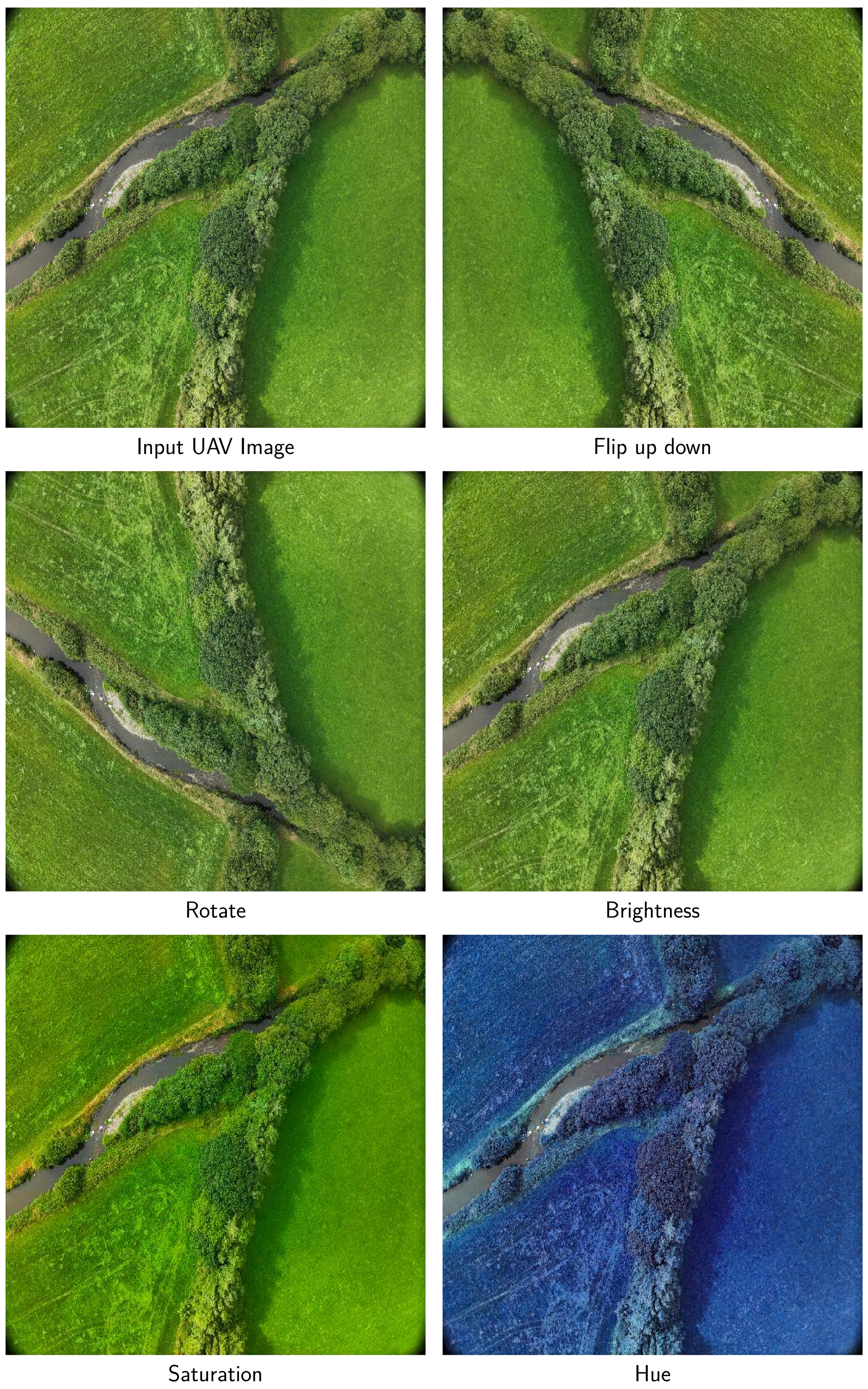


Figure 3: Set of random augmentation methods applied to the original image through the dataset loading pipeline

Explainability

DeepLabV3+ predictions with interpretability visualisations (Grad-CAM and Grad-CAM++). Red (low) to Blue (high) indicates the pixel-level importance.

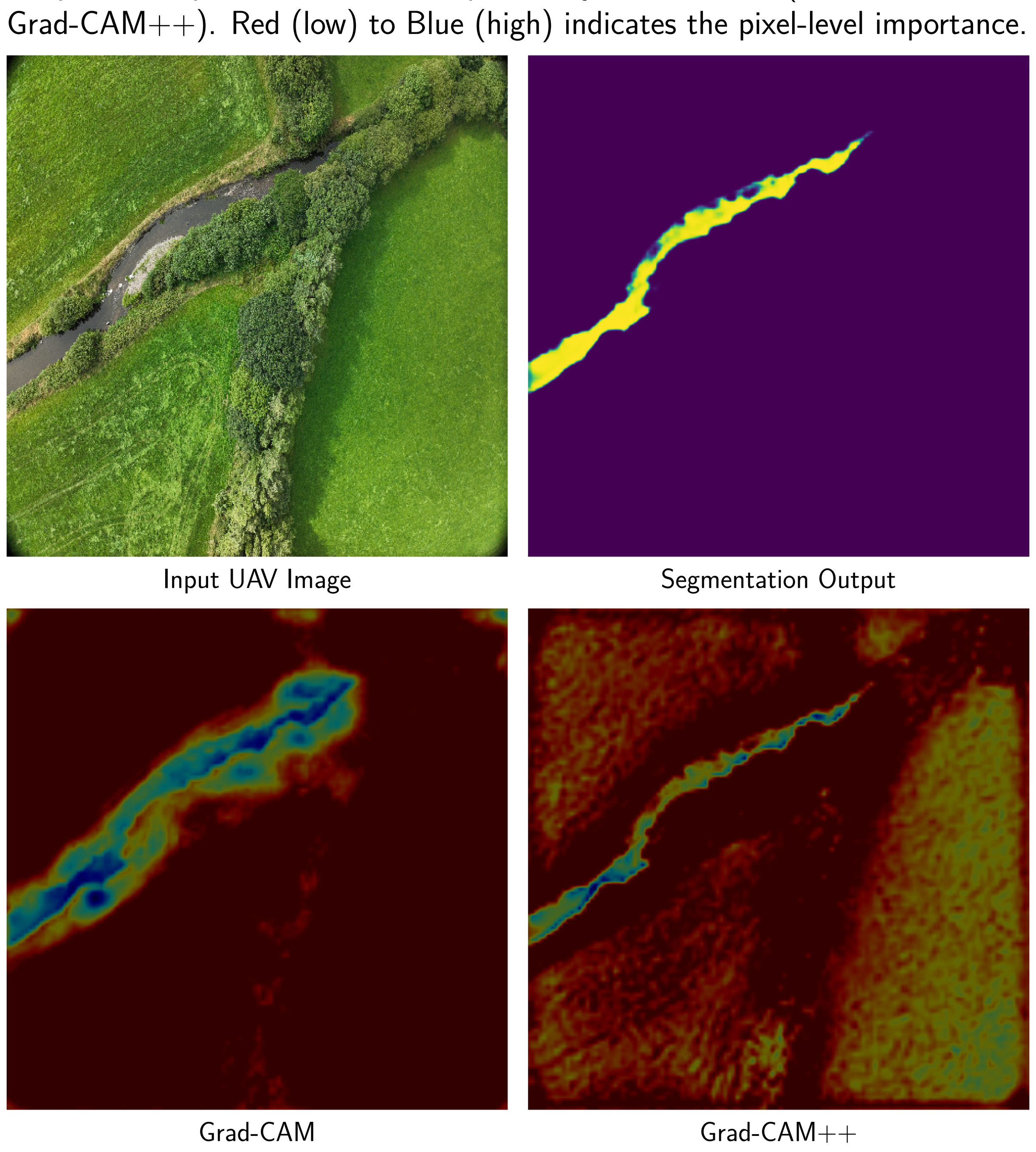


Figure 4: DeepLabV3+ Grad-CAM and Grad-CAM++ visualisations

Results: Quantitative Performance

Model	Dice	Recall	Time (ms)
DeepLabV3+	0.945	1.000	113
UNET++	0.948	0.927	186
UNET-ResNet50	0.947	0.878	172
UNET-MobileNetV2	0.933	1.000	100
SegNet-VGG16	0.940	1.000	116

Perfect recall and sub-120 ms inference are crucial for forecasting.

Computational Efficiency

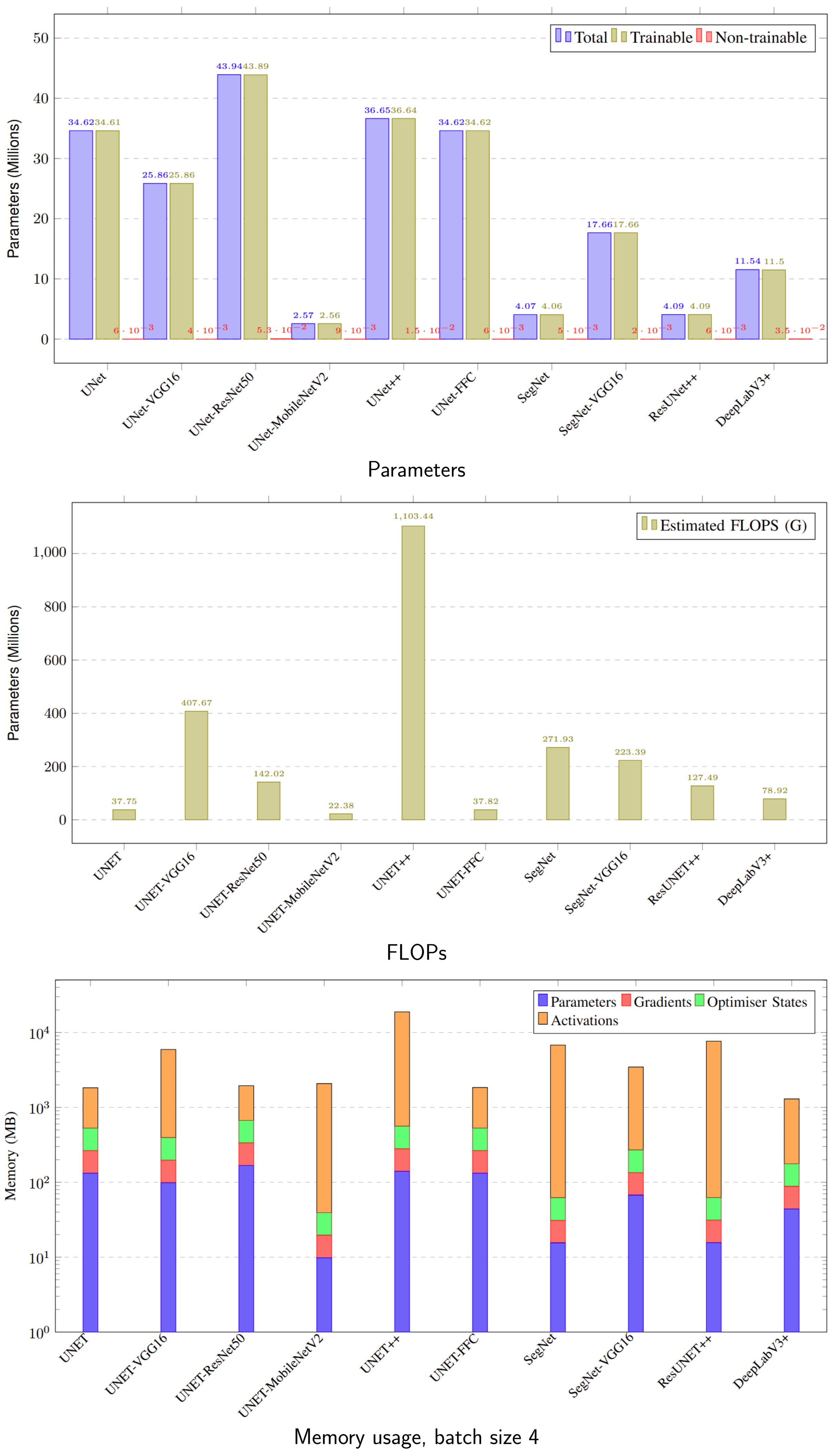


Figure 5: Computational efficiency and resource trade-offs among the selected models

Strengths and Weaknesses

DeepLabV3+	High Dice, recall 1.0, 113 ms	slight over-segmentation
UNET++	Highest Dice	slow, heavy compute
ResUNET++	Strong IoU	recall collapse risky for forecasting
UNET-MobileNetV2	Fastest and recall 1.0	precision slightly lower
SegNet-VGG16	Recall 1.0	moderate efficiency

Implications for Flash Flood Forecasting

- Recall as safety driver: do not miss water.
- Inference time for timeliness in response.
- Boundary accuracy supports hydraulic models.
- Efficiency enables field or edge deployment.

Takeaway: DeepLabV3+ is the most balanced for real-time operations. UNET-MobileNetV2 is a solid, lightweight option.

Conclusions and Links

- Occlusion-robust UAV dataset and benchmark.
- Practical guidance for real-world forecasting.
- Open data and code for reproducibility.

Dataset: https://doi.org/10.5281/zenodo.17236026
Code: https://doi.org/10.5281/zenodo.17236026

This research was supported by the EU Interreg North-West Europe FlashFloodBreaker project (Grant Number: NWE0200167). The authors also thank the continuous support of the Nimbus Research Centre and the Sustainable Infrastructure Research & Innovation Group (SIRIG) at Munster Technological University