Deforestation Detection in Ukraine via Semantic Segmentation: Comparison of U-Net, GSAM, and AdaptFormer Approaches

B.-Y. V. Dekhtiar[1]

[1]National University of Kyiv-Mohyla Academy, Kyiv, Ukraine

yarema.dekhtiar@gmail.com

This study addresses the critical issue of deforestation in Ukraine through automated analysis of satellite imagery. Deforestation detection is formulated as a semantic segmentation task, wherein each pixel in a satellite image is classified as either deforested (label 1) or non-deforested (label 0), encompassing forests, fields, urban areas, and other land covers not identified as deforested.

The dataset utilized comprises Sentinel-2 satellite imagery from the Kharkiv region of Ukraine, annotated for deforestation regions [1]. To enhance model accuracy and generalization, the original satellite images were preprocessed by cropping into smaller 224×224 pixel tiles. This approach increases the effective dataset size and improves the detection of small deforested patches, which are often challenging to identify in larger frames.

A baseline was established using a U-Net architecture with a ResNet-50 backbone, inspired by methodologies from previous research on forest change detection in Ukrainian ecosystems [2]. This model achieved an F1 score of 59.31% on the validation set and 52.65% on the held-out test set.

Subsequently, the performance of the Segment Anything Model (SAM), a recently proposed vision foundation model, was investigated. Given SAM's requirement for a fixed input size of 1024×1024 pixels, the GSAM approach [3] was employed to allow for variable image sizes and efficient fine-tuning. This model achieved F1 scores of 58.57% on validation and 53.53% on the test set. While the validation performance was slightly lower than U-Net, GSAM demonstrated improved generalization on the test set, likely due to SAM's superior pretraining.

In parallel, AdaptFormer [4], a method designed to adapt transformer-based vision models for downstream tasks with limited labeled data, was utilized to fine-tune SAM. AdaptFormer incorporates lightweight adapters into frozen vision transformers, enabling efficient fine-tuning. This combined approach achieved 65.62% F1 on the validation set and 58.20% on the test set, significantly outperforming previous models.

Table 1: F1 Scores (%) for Validation and Test Sets

Model	Validation F1	Test F1
U-Net (ResNet-50)	59.31	52.65
GSAM	58.57	53.53
AdaptFormer	**65.62**	**58.20**

The GSAM codebase [5] was adapted for training and evaluation of the GSAM and AdaptFormer models, facilitating custom dataset loading, patch extraction, and evaluation procedures. The baseline U-Net model was implemented separately. The results underscore the promise of transformer-based models for environmental monitoring tasks, particularly when enhanced with adapter tuning strategies like AdaptFormer.

In conclusion, this study provides a comparative evaluation of multiple deep learning architectures for deforestation detection in Ukraine. The findings indicate that employing SAM, either through GSAM or fine-tuned with AdaptFormer, enhances detection performance compared to existing approaches, highlighting the potential of advanced transformer-based models in environmental monitoring.

Acknowledgements

The author expresses gratitude to supervisors Vladyslav Khramtsov and Halyna Kriukova for their guidance, to Quantum for providing computational resources, to Yur-Liubomysl Dekhtiar for supplying code utilized in training baseline models, and to Mykhailo Yushchuk for his helpful suggestions.

1. Isaienkov K. Deforestation in Ukraine Dataset. Kaggle, 2025. https://www.kaggle.com/datasets/isaienkov/deforestation-in-ukraine
2. Isaienkov K., Yushchuk M., Khramtsov V., Seliverstov O. Deep Learning for Regular Change Detection in Ukrainian Forest Ecosystem With Sentinel-2. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 2021, 14, 364–376. https://doi.org/10.1109/JSTARS.2020.3034186
3. Wuyue Y., et al. Generalized SAM: Efficient Fine-Tuning of SAM for Variable Input Image Sizes. arXiv preprint arXiv:2408.12406, 2024.
4. Chen S., et al. AdaptFormer: Adapting Vision Transformers for Scalable Visual Recognition. arXiv preprint arXiv:2205.13535, 2022.
5. Usagisukisuki. G-SAM: Generalized SAM Official Implementation. GitHub Repository. https://github.com/usagisukisuki/G-SAM