# Extending the Domain of Wildfire Smoke Segmentation with Diffusion-based Image Mixing

*Julius Pesonen[1,2]*

*[1]Finnish Geospatial Research Institute, National Land Survey of Finland* *[2]Department of Computer Science, Aalto University*

## Motivation

- Wildfire smoke segmentation enables UAV-based wildfire detection in large areas [1]
- However, the data for wildfire detection is difficult to find, and the few available data sources are specific to few geographical locations
- The performance of models trained with the data suffers greatly when applied in new locations – highlighted by the almost 50% decrease in Jaccard index in *Figure 1.*

**Training domain data – mIoU: 0.594**

**Target domain data – mIoU: 0.307**

*Figure 1. Examples and Jaccard index results (mean intersection over union or mIoU) of the training data and target domain test data with the baseline model.*

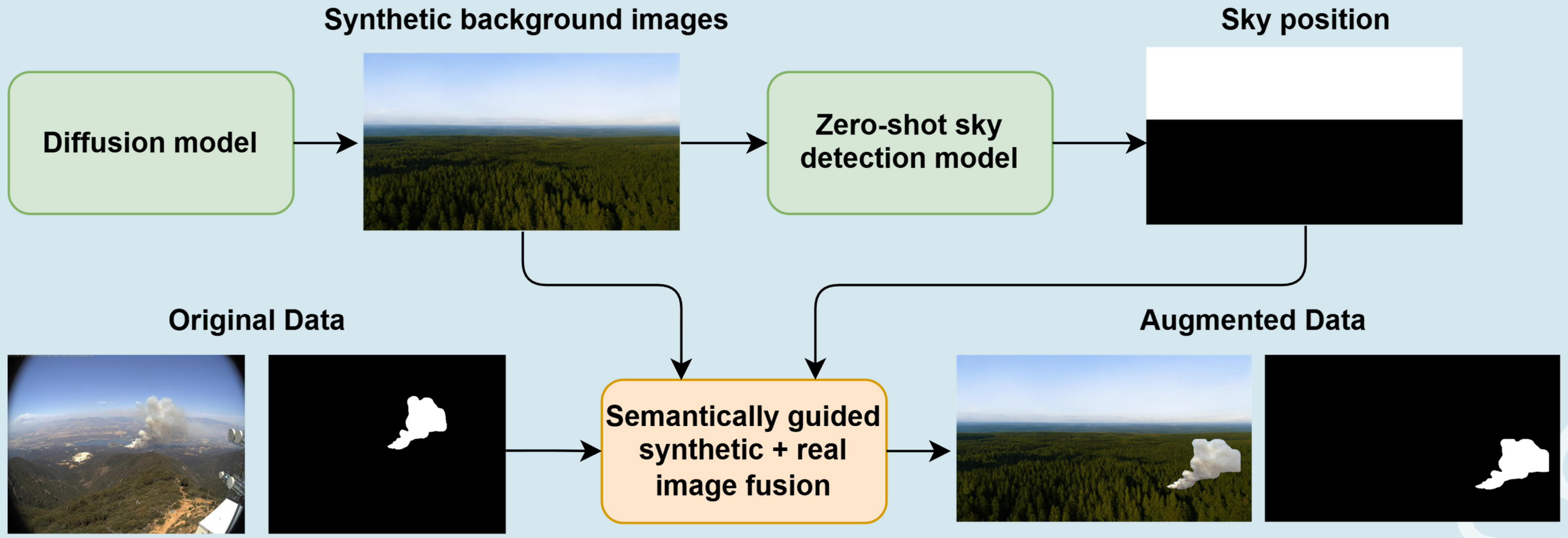


*Figure 2. The augmentation method process. The diffusion and sky segmentation model are used as is, without any retraining or fine-tuning. The image fusion is done using randomized copy-pasting with random scaling, that is adjusted based on the position of the smoke between the bottom of the image and the horizon. The sky position information is also used to prevent the smoke from being pasted completely on the sky, as such features should be detected as clouds.*

## Results

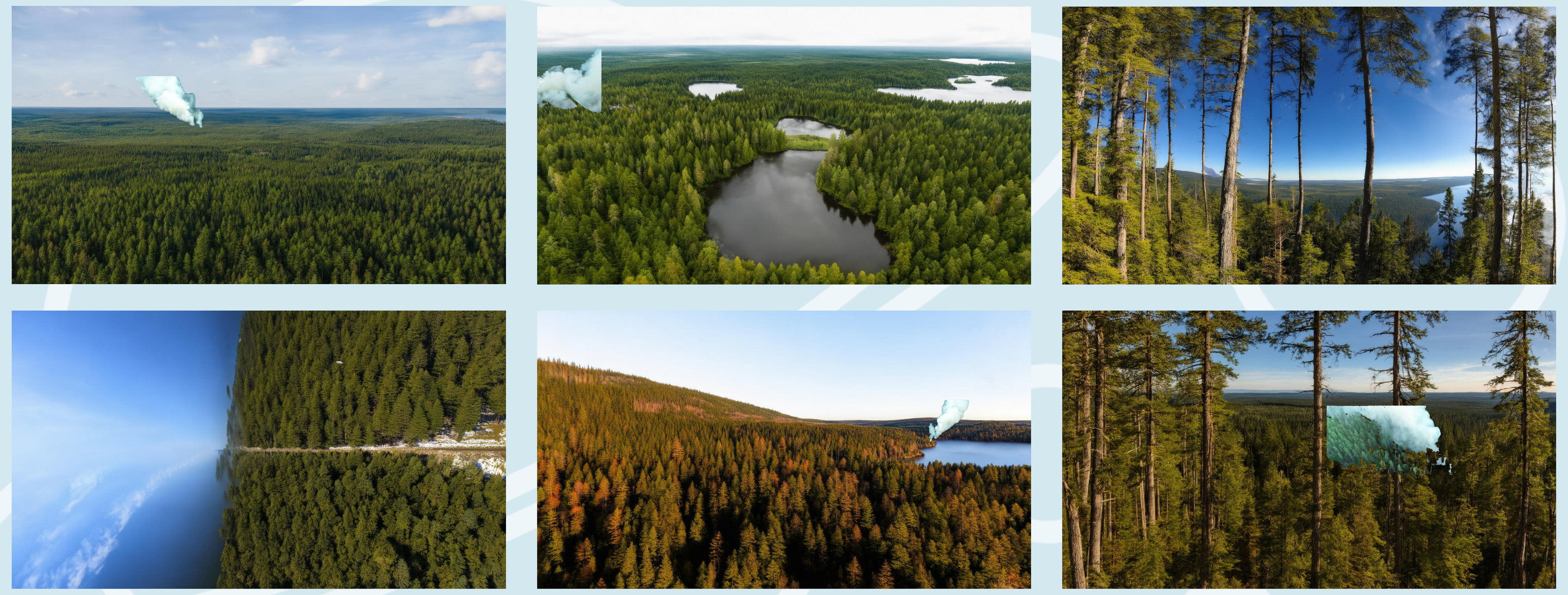

*Figure 3. Examples of the augmented smoke images.*

## Methods

- Fusing the original smoke data with diffusion model generated backgrounds creates new data which better represents the target domain
- The method is enhanced with guidance from a zero-shot object detection model to avoid generating problematic samples with the copy-pasting
- Results in realistic training samples where the original repositioned label masks can be used for supervision

- Stable Diffusion 3.5 [2] as the background image generator with varied prompting to generate the background images

```
"Boreal forest view {with some lakes, with some small
hills} captured from drone at a 150 meter altitude {in
the summer, in the spring, in the autumn} {with a
clear sky, with a cloudy sky, with a few clouds in the
sky} and a view towards the horizon {at daytime, in
the morning, in the evening, at night}."
```

- OmDet-Turbo-Tiny [3] as the zero-shot sky detector

## Discussion

- Provides a way to train models for new domains without real target domain specific data
- As seen in some of the augmented examples, the resulting images occasionally contain odd features
- The features are caused by either odd background generation results, failed sky labelling, or poor original smoke labels
- Testing with downstream models is in progress

## References

[1] J. Pesonen, T. Hakala, V. Karjalainen, N. Koivumäki, L. Markelin, A.-M. Raita-Hakola, J. Suomalainen, I. Pölönen, and E. Honkavaara, "Detecting wildfires on UAVs with real-time segmentation trained by larger teacher models," in Proceedings of the IEEE/CVF winter conference on applications of computer vision, 2025.

[2] P. Esser, S. Kulal, A. Blattmann, R. Entezari, J. Muller, H. Saini, ¨
Y. Levi, D. Lorenz, A. Sauer, F. Boesel, et al., "Scaling rectified flow transformers for high-resolution image synthesis," in Forty-first International Conference on Machine Learning, 2024.

[3] T. Zhao, P. Liu, X. He, L. Zhang, and K. Lee, "Real-time transformer based open-vocabulary detection with efficient fusion head," arXiv preprint arXiv:2403.06892, 2024.

NLS
FINNISH GEOSPATIAL RESEARCH INSTITUTE
FGI

Aalto University
School of Science

FireMan