DETECTING WILDFIRES ON UAVS WITH REAL-TIME SEGMENTATION TRAINED BY LARGER TEACHER MODELS

Julius Pesonen[1] Teemu Hakala[1] Väinö Karjalainen[1] Niko Koivumäki[1] Lauri Markelin[1] Anna-Maria Raita-Hakola[2] Juha Suomalainen[1] Ilkka Pölönen[2] Eija Honkavaara[1]

[1]Finnish Geospatial Research Institute, National Land Survey of Finland [2]Faculty of Information Technology, University of Jyväskylä

Motivation

- Fire control flights are required in many areas to prevent major wildfires
- UAVs could bring huge efficiency improvements
- UAVs have challenges in areas of weak telecommunication
- Lack of sufficient training data is prominent – especially for segmentation

Figure 1. Examples from the used bounding box labelled wildfire smoke datasets. The ones with a 4:3 aspect ratio are from the AI For Mankind datasets and the rest are from the UAV dataset.

The approach in a nutshell

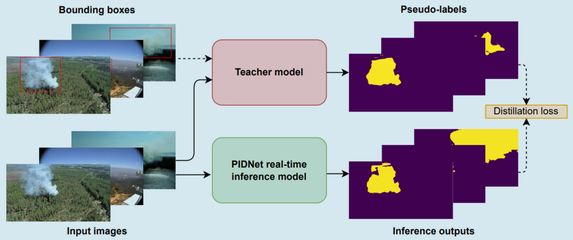

Figure 2. Inference model training scheme. The teacher model is either trained using bounding box supervision method such as BoxSnake [2] or guided with the bounding boxes like SAM [1].

- A wildfire smoke specific distillation of SAM (Segment Anything Model) [1] with bounding box annotated data
- Inference with a lightweight PIDNet segmentation model [2]

Data and training

- Trained using a combination of two bounding box annotated datasets. [3]
- All models trained for 50 epochs, optimized with AdamW [4] on a learning rate of 1e-3 using model free geometric and color augmentations.

Results

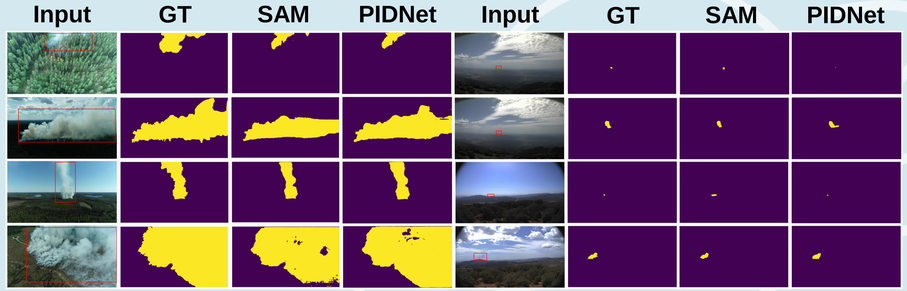

Table 1. Random examples from the manually annotated test set. From left to right in two columns: Input and a manually created bounding box, manual ground truth mask (GT), SAM-generated pseudo label, and the corresponding PIDNet-S prediction.

Teacher	Student	mIoU
ResNet-50-RCNN-FPN	-	0.560
SAM*	-	0.636
	Distilled models	
ResNet-50-RCNN-FPN	PIDNet-S	0.602
SAM	PIDNet-S	0.594

Table 2. Test set metrics of the different models. *The SAM test metrics are computed from the bounding box guided masks.

- The model also captured the smoke in a live test at a prescribed burning event from up to 9.7 km away – video found behind the QR code
- 25 fps inference with NVIDIA Jetson Orin NX at full HD resolution

Conclusions

- Showed that wildfire smoke can be segmented in real-time with on-board computational resources
- Implies that such tasks where typically bounding box detection is used could be changed to segmentation at low cost
- However, results are limited to the specific domain and the practical tests are also limited to a single geographical location
- Next steps would be to improve the data distribution and to test the method's generalization to other tasks

References

[1] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer White9 head, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 4015–4026, 2023.

[2] Jiacong Xu, Zixiang Xiong, and Shankar P Bhattacharyya. PIDNet: A real-time semantic segmentation network inspired by PID controllers. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 19529–19539, 2023.

[3] AI for Mankind. Open wildfire smoke datasets. https:// github.com/aiformankind/wildfire-smokedataset, 2020.

[4] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In International Conference on Learning Representations, 2019.

[5] Rui Yang, Lin Song, Yixiao Ge, and Xiu Li. BoxSnake: Polygonal instance segmentation with box supervision. 2023 IEEE/CVF International Conference on Computer Vision (ICCV), pages 766–776, 2023

NLS FINNISH GEOSPATIAL RESEARCH INSTITUTE FGI

JYVÄSKYLÄN YLIOPISTO UNIVERSITY OF JYVÄSKYLÄ

FireMan