



Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen
International Center for
Information Technology

DLRV Project proposal

Semantic Segmentation of aerial images of forest scenery

*Malika Navaratna, Urvashi Negi, Zain Ul Haq, Simon
Deussen*

Supervised by

Prof. Sebastian Houben

November 2021

1 Introduction

Deforestation is becoming an increasingly difficult problem, and it has adverse effects on the bio-diversity, global warming and more. Some of the factors that contribute to Deforestation is agricultural practices, raw materials for construction and household items, land clearing for development projects. [1]. Attempts are being made to reforest the environment and scientists and engineers are looking for better ways to this. For this reforestation efforts it is necessary to estimate the amount of deforested areas and for efficient planting of seeds it is important to find the land that is suitable for seeding. Meaning to find areas that are free of large trees, roads, logs that have already been cut down etc. Semantic segmentation of forest imagery is one solution to this problem.

The objective of the project is to perform semantic segmentation using Deep learning techniques. The project will use labelled data of forest scenery captured from an aerial view (most commonly by unmanned aerial vehicles). An Example of the type of data that will be used for the project is given in the Figure1.



Figure 1: Forest imagery

The expected final result is to have a pixel-wise segmentation of the forest into the following categories

- Empty forest area

- Trees
- Shrubbery
- Logs(cut down stacks of trees)
- Roads(dirt and tarmac)

1.1 Models and tools

- U-Net is the classical semantic segmentation architecture, builds on top of a fully convolutional network, but prevents the loss of resolution in the resulting masks.[2]
- Mask RNN another important architecture combining Faster RCNN and a Fully convolutional network for segmentation [3]
- SegNet is using a Encoder-Decoder architeture [4]
- Swin transformer is another architecture that was proved to be successfully. The paper [5] has shared the code of their implementation, in which they performed segmentation of forest scenery captured by satellites.

1.2 Data sets

Forest Segmentation Contains 5108 Images (256x256) and mask for forest segmentation

Satellite terrain A bit more complicated, contains huge amount of satellite images in different bands. We have to check if they are usable in our context (drone images) Contain 10 classes including road, track, trees, crops...

UAVID dataset Drone dataset, labels contain Road Tree Vegetation BUT its more urban areas

1.3 Potential problems

Computational capacity Working with images is always computational expensive. We might be able to use our team member Simon's high performance GPU, but this is not setup yet. Even then, there is the problem with accessing this async.

Dataset combination Combination of dataset to get enough data and enough labels is probably the hardest part.

References

- [1] A. Kamaris and F. X. Prenafeta-Boldú, “A review of the use of convolutional neural networks in agriculture,” *The Journal of Agricultural Science*, vol. 156, no. 3, p. 312–322, 2018.
- [2] M. Umar, L. B. Saheer, and J. Zarrin, “Forest terrain identification using semantic segmentation on uav images,” 2021.
- [3] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask r-cnn,” 2018.
- [4] V. Badrinarayanan, A. Kendall, and R. Cipolla, “Segnet: A deep convolutional encoder-decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 2481–2495, Dec 2017.
- [5] E. Guérin, K. Oechslin, C. Wolf, and B. Martinez, “Satellite image semantic segmentation,” 2021.