

Tree object detection using airborne images and LiDAR point clouds

Alexandre Bry
École polytechnique, Geodan B.V.
alexandre.bry.21@polytechnique.org

2024-07-08

Abstract This is the abstract. It can be on multiple lines and containing **Markdown**.

Table of contents

1 Introduction	1
2 State-of-the-art	1
2.1 Datasets	2
2.1.a Requirements	2
2.1.b Existing tree datasets	2
2.1.c Things to talk about	3
2.2 Models	3
3 The dataset	3
4 The model	3
5 Results	3
6 Conclusion	3
Bibliography	3

1 Introduction

The goal of the internship was to study the possibility of combining LiDAR point clouds and aerial images in a deep learning model to perform instance segmentation of trees. The two types of data are indeed complementary, as point clouds capture the shape of the worlds, while images capture the colors. However, combining them into a format that allows a model to handle them simultaneously is not an easy task because they inherently have a very different spatial repartition and encoding.

The second major topic of the internship was to acquire a proper dataset matching all the criteria required for the project. Most of the datasets containing tree annotations only used either RGB images or LiDAR point clouds, but not both. Therefore, I had to create such a dataset by myself, using the openly available images and point clouds in the Netherlands, by annotating trees by hand to properly train and evaluate the methods.

2 State-of-the-art

2.1 Datasets

2.1.a Requirements

Before presenting the different promising datasets and the reasons why they were not fully usable for the project, let's enumerate the different conditions and requirements for the tree instance segmentation task:

- Multiple types of data:
 - Aerial RGB images
 - LiDAR point clouds (preferably aerial)
 - (Optional) Aerial infrared images
- Tree crown annotations or bounding boxes
- High-enough resolution:
 - For images, about 25 cm
 - For point clouds, about 10 cm

Here are the explanations for these requirements. As for the types of data, RGB images and point clouds are required to experiment on the ability of the model to combine the two very different kinds of information they hold. Having infrared data as well could be beneficial, but it was not necessary. Regarding tree annotations, it was necessary to have a way to spatially identify them individually, using crown contours or simply bounding boxes. Since the model outputs bounding boxes, any kind of other format could easily be transformed to bounding boxes. Finally, the resolution had to be high enough to identify individual trees and be able to really use the data. For the point clouds especially, the whole idea was to see if and how the topology of the trees could be learnt, using at least the trunks and even the biggest branches if possible. Therefore, even if they are not really comparable, this is the reason why the required resolution is more precise for the point clouds.

Unfortunately, none of the datasets that I found matched all these criteria. Furthermore, I didn't find any overlapping datasets that I could merge to create a dataset with all the required

types of data. In the next parts, I will go through the different kinds of datasets that exist, the reasons why they did not really fit for the project and the ideas I got when searching for a way to use them.

2.1.b Existing tree datasets

As explained above, there were quite a lot of requirements to fulfill to have a complete dataset usable for the task. This means that almost all the available datasets were missing something, as they were mainly focusing on using one kind of data and trying to make the most out of it, instead of trying to use all the types of data together.

The most comprehensive list of tree annotations datasets was published in A. Ouaknine, T. Kattenborn, E. Laliberté, and D. Rolnick [1]. Without enumerating all of them, there were multiple kinds of datasets that all have their own flaws regarding the requirements I was looking for.

Firstly, there are the forest inventories. T. Jucker *et al.* [2] is probably the most interesting one in this category, because it contains a lot of spatial information about almost 500K trees, with their locations, their crown radii and their heights. Therefore, everything needed to localize trees is in the dataset. However, I didn't manage to find RGB images or LiDAR point clouds of the areas where the trees are located, making it impossible to use these annotations to train tree detection.

Secondly, there are the RGB datasets. G. Reiersen *et al.* [3] and B. Weinstein [4] are two of them and the quality of their images are high. The only drawback of these datasets is obviously that they don't provide any kind of point cloud, which make them unsuitable for the task.

Thirdly, there are the LiDAR datasets, such as E. Kalinicheva, L. Landrieu, C. Mallet, and N. Chehata [5] and S. Puliti *et al.* [6]. Similarly to RGB datasets, they lack one of the data source for the task I worked on. But unlike them, they have the advantage that the missing data can be much easier to acquire from another source, since RGB aerial or satellite images are much more common than LiDAR point clouds. However, this solution was abandoned for two main reasons. First it is quite challenging to find the exact locations where the point clouds were acquired. Then, even when the location is known, it is often in the middle of a forest where the quality of satellite imagery is very low.

- OpenForest A. Ouaknine, T. Kattenborn, E. Laliberté, and D. Rolnick [1]: list of all the openly available forest monitoring datasets
- FoMo-Bench N. I. Bountos, A. Ouaknine, and D. Rolnick [7]: benchmark for forest monitoring
- NEON B. G. Weinstein, S. Marconi, S. Bohlman, A. Zare, and E. White [8]: LiDAR and RGB but quality was too low for me at the beginning
- MDAS J. Hu *et al.* [9]: RGB, hyperspectral and DSM, no LiDAR.

2.1.c Things to talk about

Difference between object detection and instance segmentation. The first one is easier to train because it only requires bounding boxes

Open public tree datasets in several municipalities in France and in the Netherlands: sometimes no position

2.2 Models

3 The dataset

4 The model

5 Results

6 Conclusion

Blablabla

Bibliography

- [1] A. Ouaknine, T. Kattenborn, E. Laliberté, and D. Rolnick, “OpenForest: A data catalogue for machine learning in forest monitoring.” 2023.
- [2] T. Jucker *et al.*, “Tallo: A global tree allometry and crown architecture database,” *Global Change Biology*, vol. 28, no. 17, pp. 5254–5268, 2022, doi: <https://doi.org/10.1111/gcb.16302>.
- [3] G. Reiersen *et al.*, “ReforesTree: A Dataset for Estimating Tropical Forest Carbon Stock with Deep Learning and Aerial Imagery.” [Online]. Available: <https://arxiv.org/abs/2201.11192>
- [4] B. Weinstein, “MillionTrees.” Accessed: Jul. 08, 2024. [Online]. Available: <https://milliontrees.idtrees.org/>
- [5] E. Kalinicheva, L. Landrieu, C. Mallet, and N. Chehata, “Multi-Layer Modeling of Dense Vegetation from Aerial LiDAR Scans.” [Online]. Available: <https://arxiv.org/abs/2204.11620>
- [6] S. Puliti *et al.*, “FOR-instance: a UAV laser scanning benchmark dataset for semantic and instance segmentation of individual trees.” [Online]. Available: <https://arxiv.org/abs/2309.01279>
- [7] N. I. Bountos, A. Ouaknine, and D. Rolnick, “FoMo-Bench: a multi-modal, multi-scale and multi-task Forest Monitoring Benchmark for remote sensing foundation models,” *arXiv preprint arXiv:2312.10114*,

2023, [Online]. Available: <https://arxiv.org/abs/2312.10114>

- [8] B. G. Weinstein, S. Marconi, S. Bohlman, A. Zare, and E. White, “Individual Tree-Crown Detection in RGB Imagery Using Semi-Supervised Deep Learning Neural Networks,” *Remote Sensing*, vol. 11, no. 11, 2019, doi: 10.3390/rs11111309.
- [9] J. Hu *et al.*, “MDAS: a new multi-modal benchmark dataset for remote sensing,” *Earth System Science Data*, vol. 15, no. 1, pp. 113–131, 2023, doi: 10.5194/essd-15-113-2023.