Deep Learning Enables Country-wide Individual Tree Canopy Segmentation, Counting and Height Estimation

Sizhuo Li, Martin Brandt, Philippe Ciais, Ankit Kariryaa, Rasmus Fensholt

University of Copenhagen, University Paris Saclay

PRESENTED AT:

INTRODUCTION

European forests have a wide range of social-ecological functions, such as storing carbon, preventing natural hazards, and providing food and shelters [1,2]. Monitoring the status of forests not only deepens our understanding of climate change and ecosystems but also helps guide the formulation of ecological protection policies. Remote sensing-based analyses of forests are typically limited to forest cover, and most of our knowledge of forests comes from forest inventories, where tree density, canopy cover, species, height, carbon stock, and other indicators are recorded. The inventories are conventionally established by manually collecting in-situ measurements, which can be time-consuming, labor-intensive, and difficult to scale up [3,4,5]. Forest cover also ignores trees outside forests, which can constitute a considerable woody resource and provide a variety of ecosystem services.

Recently, Brandt et al. [6] have shown that advances in computer vision and remarkably high resolution (0.5 m) satellite imagery enable the mapping of individual trees and shrubs in the Sahara Desert. While questions arise about whether similar approaches can be applied to dense forest trees. For most European countries, high-resolution aerial photos and LiDAR data are readily and publicly available and are often updated at an annual frequency, yet the feasibility of individual tree counting, crown segmentation, and height estimation in European forests has never been proven.

Here we introduce a deep learning-based 3-task framework which solves the individual tree counting, crown segmentation and height estimation problems from aerial photos and demonstrate the scalability of the methods by applying it for Denmark. We further demonstrate the robustness and transferability of the networks by adapting it to Southern Finland.

For more details regarding our global-scale individual tree mapping studies, visit www.treesoutsideforests.com!

DATA

The whole dataset contains 4-band (red, green, blue, near-infrared) 0.2 m resolution aerial photos in Denmark taken during the summer of 2018 and 0.4 m resolution canopy height model derived from LiDAR data collected in 5 years.

Two datasets were constructed for model training. For the individual tree counting and crown segmentation tasks, 19k trees in 84 sampling plots across Denmark were manually delineated for supervision, see Figure 1. For the height estimation task, 6.3k ha of aerial photos and corresponding canopy height data were collected and further split into a training (5k ha) and a testing (1.3k ha) set. We selected only the regions where LiDAR data was collected during the same period as the aerial photos (approximately 1/5 of Denmark) to ensure high consistency of the dataset.

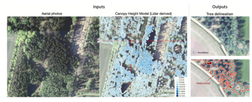


Figure 1: Datasets including the aerial photos, canopy height model, tree delineations as well as density maps.

METHODS

The framework contains 2 models: model 1 solves the individual tree counting and crown segmentation tasks jointly; model 2 solves the height estimation task independently. The models share a similar deep learning architecture adapted from the well-known UNet [7], see Figure 1.

Tree counting is modeled as a density counting problem where each tree object is modeled as a Gaussian kernel that sums to 1. Accordingly, the total tree count per plot equals to the integration of the corresponding density map [8]. Model 1 takes as input 4-band aerial photos, a computed NDVI band and canopy height maps, and generates tree segmentation masks as well as density maps. We use the Tversky loss [9] for the segmentation task, and a tree boundary mask [6] is introduced to weigh the losses and helps improve the separation of the touching crowns. We use mean squared error for the density counting branch.

The height estimation model takes as input 4-band aerial photos and a computed NDVI band, and generates pixel-wise canopy height maps. We construct the training dataset by balancing different forest types (coniferous, broadleaved, and non-forest) and varying height ranges to ensure equivalent focuses on different landscapes. We use smooth L1 loss for model training.

Transfer learning is used for adapting the model trained in Denmark to Southern Finland. Specifically, we train the counting and crown segmentation model using images and tree delineations in Denmark and then finetune the pretrained model with the original Denmark data as well as some additional training data (3k tree delineations) in Southern Finland.

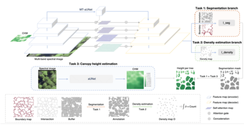


Figure 2: Overall framework containing 3 tasks: 1. segmentation, 2. density estimation and 3. pixelwise height estimation. Task 1 and 2 are solved jointly with a multi-tasking network (MT-uUNet). Task 3 is solved separately using a single task network (uUNet). Per-tree height map can be obtained by combining results from task 1 and 3.

RESULTS

We applied the proposed 3-task framework for Demark and noticed consistently high performance in diverse regions, and a total of 3.2M individual trees were detected, see Figure 3.

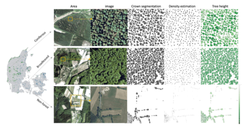


Figure 3: Products of the proposed framework, with examples in 3 major forest types (coniferous, broadleaved, and non-forest).

We validated the individual tree counting and crown segmentation performance on 3k manually delineated tree polygons that were never seen by the model during training. The model achieved robust performance in counting trees in different landscapes, reflected by fitting slopes close to 1. While a slight underestimation was observed in the predicted canopy area, reflected by fitting slopes from 1.1 to 1.4, see Figure 4.

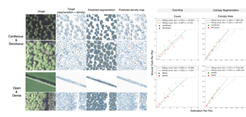


Figure 4: Validation of the counting and crown segmentation tasks using the testing data.

We validated the height estimation performance on the testing images (1.3k ha), randomly sampled from the whole dataset. We noted visually high agreement between the per-pixel height estimation and the reference LiDAR height. The individual tree height estimation obtained by combing the per-pixel height estimation with the crown segmentation results also strongly agrees with the references according to the heatmap, see Figure 5.

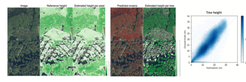


Figure 5: Validation of the per-pixel and per-tree height estimation using the testing data.

We validated the finetuned model in 3 sampling regions in Southern Finland and observed high performance in both counting and individual crown segmentation tasks, see Figure 6.

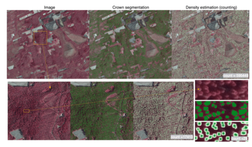
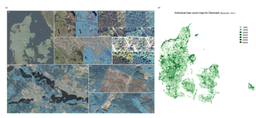


Figure 6: individual tree counting and crown segmentation applied in Southern Finland.

DISCUSSIONS

The proposed framework solves the individual tree counting, crown segmentation and height estimation problems in a scalable and transferable way, which could strongly support the traditional labor-intensive forest management studies. Additionally, the high efficiency of such intelligent models enables fast and precise country-wide tree inventory at an annual frequency. Yet a few limitations exist for the proposed methods. Firstly, the performance of the tree segmentation model is upper bounded by human vision, as the training is supervised by manually created annotations that might contain bias or mistakes. Secondly, the model tends to underestimate individual crown area since we aim at separating single trees. Thirdly, the model tends to underestimate the height of tall trees due to insufficient supervision of particularly tall trees. To summarize, our study verified the feasibility of country-scale individual tree inventory and paved the way for potential ecological analysis such as biomass estimation at individual tree level.



ABSTRACT

Forests have a wide range of social-ecological functions, such as storing carbon, preventing natural hazards, and providing food and shelters. Monitoring the status of forests not only deepens our understanding of climate change and ecosystems, but also helps guide the formulation of ecological protection policies. Remote sensing based analyses of forests are typically limited to forest cover, and most of our knowledge of forests comes from forest inventories, where tree density, canopy cover, species, height, carbon stock and other indicators are recorded. The inventories are conventionally established by manually collecting in-situ measurements, which can be time-consuming, labor-intensive and difficult to scale up. Here we present an automatic and scalable tree inventory pipeline based on publicly available aerial images from Denmark and deep neural networks, enabling individual-tree-level canopy segmentation, counting, and height estimation within different kinds of forests. The canopy segmentation and counting tasks are solved in a multi-tasking manner, where a convolutional neural network is trained to jointly predict a segmentation mask and a density map which sums up to the total tree count for a given image. Another network trained with LiDAR-derived height maps estimates pixel-wise canopy height from aerial photos, which, when combined subsequently with the canopy segmentation masks, allows for per-tree height mapping. The multi-tasking network achieves a segmentation mean IoU of 0.676, and a counting MAE of 1.6 trees/hectare when evaluated on 3k manually delineated trees of 3869 m2 (~0.84). Compared with independent LiDAR reference heights, the height estimation model achieves a pre-pixel mean absolute error (MAE) of 2.6 m on the testing set and a per-tree MAE of 3.0 m when assigning tree height with the maximum height estimate within each predicted canopy. The models perform robustly over diverse landscapes including dense forests (coniferous and broad-leaved), open fields, and urban areas. We further verify the scalability of the framework by detecting 312 million individual trees across Denmark.

REFERENCES

[1] Krieger, Douglas J. Economic value of forest ecosystem services: a review. The Wilderness Society. (2001).

[2] Fuhrer, Erwin. Forest functions, ecosystem stability and management. Forest Ecology and management-132). p29-38. Elsevier. (2000).

[3] Forest Europe (2020). State of Europe's forests. (2020).

[4] Hyyppä, J. et al. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. International Journal of Remote Sensing 29. (2008).

[5] Designing and Conducting a Forest Inventory - case: 9th National Forest Inventory of Finland.

[6] Brandt, M. et al. An unexpectedly large count of trees in the West African Sahara and Sahel. Nature 587. (2020).

[7] Ronneberger, Olaf and Fischer, et al. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. (234-241). (2015).

[8] Zhang, Cong and Li et al. Cross-Scene Crowd Counting via Deep Convolutional Neural Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2015).

[9]. Tversky. Features of similarity. Psychological Review. 84(4):327 – 352. 1977.