Tree Species Identification and Crown Delineation Using Classification Models and Remote Sensing Data

Justin Mau

Montana State University

LRES 525 – Applied Remote Sensing

Introduction

In our forests, an important aspect of preserving, protecting, and managing them is knowing the distribution and health of a forest's trees. This knowledge is used by forest researchers to understand the composition of the ecosystem, by forest managers to know the health and status of the forest, and by environmentalists to monitor the spread of invasive species and the health of native species. However, it can be very difficult to identify and monitor every tree in a forest. To do so would require many workers, many hours, and lots of funding, especially since many forests are expansive, remote, and difficult to access. In order to expand the scope of researchers and forest managers, remote sensing data can bridge that gap, allowing them to get a comprehensive overview of the forest without having to constantly monitor every tree.

For this research, I look to test the effectiveness of deep learning models to identify individual trees using satellite imagery data. My hypothesis is that a general model can be used to effectively identify most individuals; however, I do not believe that satellite imagery data will be good enough to accurately identify species. To test this hypothesis, I used remote sensing NEON data to attempt to identify tree crowns. The satellite imagery and tree data come from the University of Notre Dame Environmental Research Center, in the southwest Ottawa National Forest, located between northern Wisconsin and Michigan's Upper Peninsula. The dominant species in this location include red and sugar maple, aspen, paper birch, balsam fir, and hemlock.

Methods

Using the NEON API in R, I downloaded 4 tiles of satellite imagery for the UNDE site that include most of the tree study plots. I also downloaded vegetation structure data for the UNDE site, which includes apparent tree individuals, their status and plot ID, stem size, health, etc. First, any trees that did not have a canopy position of either "Full sun" or "Partially shaded" were filtered out. From there, the geoNEON R package was used to extract coordinate locations for each apparent individual, which I then exported as a .csv file for each plot. The .csv files also contained each tree's species, canopy diameter, and uncertainty. These .csv files were then added to ArcGIS and converted to feature layers.

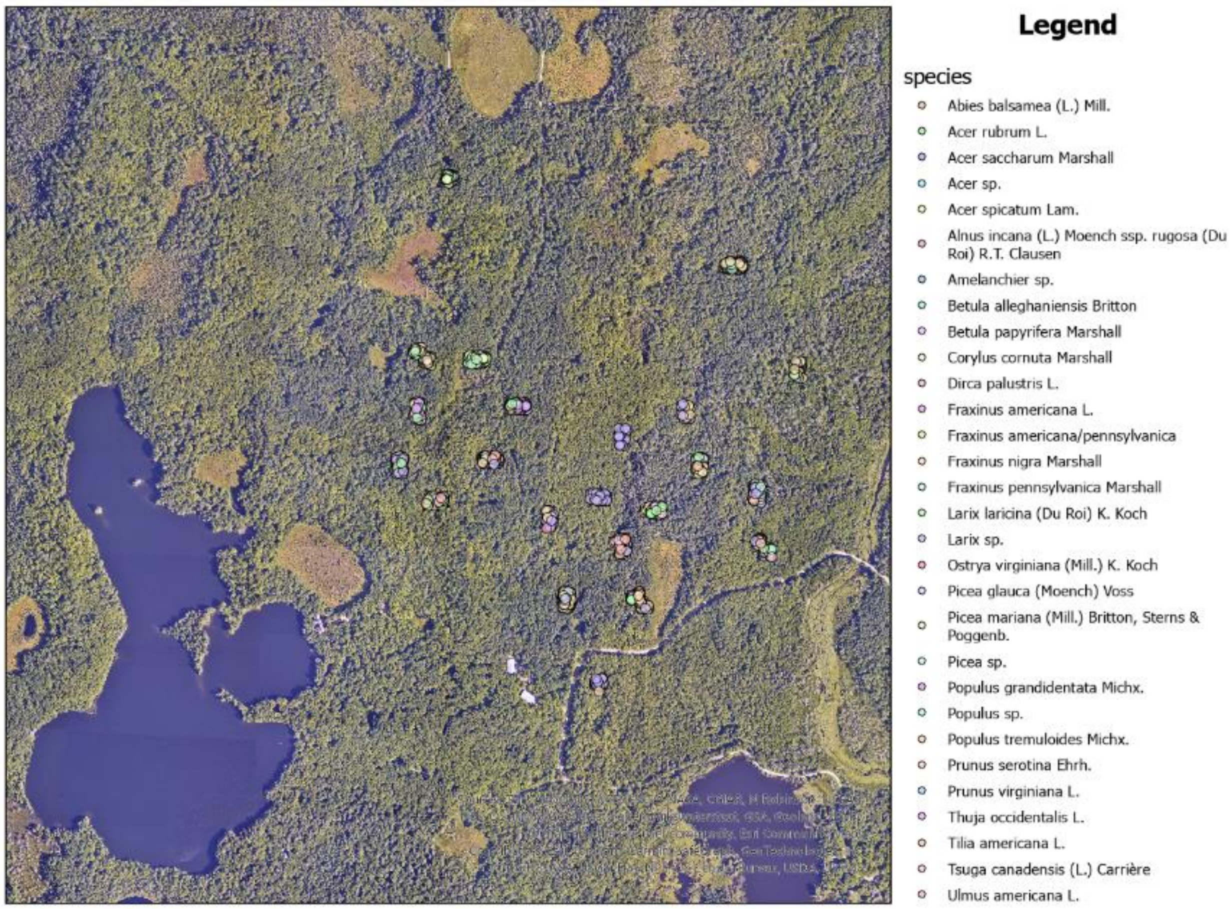

Figure 1: Map of identified trees

After loading in each apparent individual tree as a point feature, I then changed the symbology to change the color of each point based on the species' scientific name (Fig. 1). From here, I was able to find a study site that had a relatively high species diversity, as well as fairly clear satellite imagery (some imagery was mildly distorted near some study sites). Study site 065 was selected to test the deep learning model.

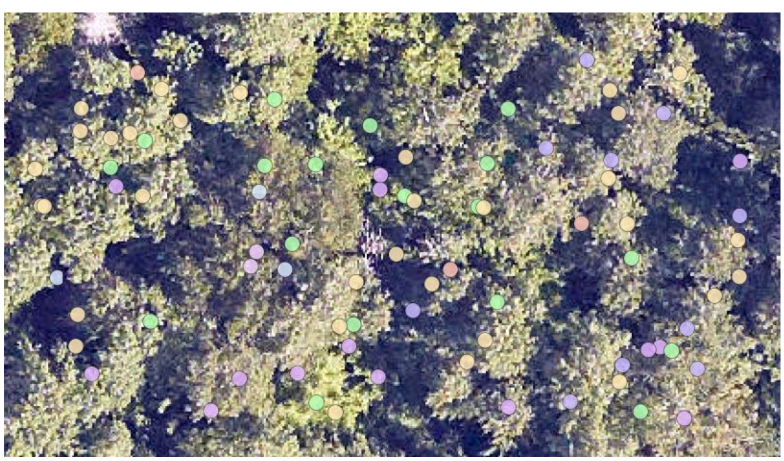

Figure 2: Plot 065, with labelled apparent tree individuals

To detect and delineate tree canopies, I used the DeepForest model, a Python package that contains pretrained models for tree identification. The satellite imagery was cropped to just the selected plot 065 (Fig. 2). This plot was then run through the DeepForest model using a relatively small window size of 250 pixels and 25% overlap.

Results

Unfortunately, testing the validity of the results is difficult. Each study site only has roughly 50 trees identified, far fewer than the number present at the site. Therefore, the only trees that can be confirmed as properly identified are those. While there is a canopy size field from the vegetation structure data, I could not manage to get the tree points to scale properly with that canopy size.

Figure 3: Results of DeepForest tree canopy identification, with points for each known tree

In the plot 065 region, a total of 57 known trees with mostly prevalent canopy positions were collected (Fig. 3). The DeepForest model detected a total of 60 tree canopies, with 26 of these canopies containing a known tree point. From a rough glance at the output image, most of these 60 do seem to be accurate. However, there are many instances where multiple known trees are included in one detected canopy. In fact, only 11 of the 26 detected canopies with known trees contain just one tree. Nine of the known trees were not included in any detected canopy at all. Of these, however, five were “partially shaded”, which may mean that their crown is not visible from the overhead view of the canopy. Another factor that may affect the placement of these known trees is the uncertainty for each tree. Some tree data has an associated uncertainty for its coordinates, so some trees may be more or less under crowns identified by the model.

Conclusions

Overall, while the ability of this model to predict trees is promising, as it is, its performance is not great. My hypothesis was partially true, but the model's performance is not good enough to say that it can accurately identify tree individuals. There are a number of factors limiting its performance in the current state. For one, these NEON sites only have a small subset of trees sampled at the base, making it difficult to ground-truth the model's predictions. If they had one plot that was fully sampled and had confirmed crown identification, then that could be annotated and used to better train the DeepForest model. As it is, with only a few trees sampled in each plot, there is not enough data there to accurately annotate training data for the model. While the prebuilt model certainly does a good job at delineating the most obvious tree crowns, it is not yet ready to predict tree species.

The next step for this research would be to extend this model, training it on annotated crown data, including species data. This way, the model could then attempt to predict the species of each identified tree. With a model capable of both crown delineation and species identification, forest management projects could expand in scope, allowing researchers and managers to gain a better understanding of the makeup of a forest ecosystem.

Code & Comments

See the following GitHub repository for code:

https://github.com/jdmau72/LRES-525---Tree-Crown-Delineation.git