Assessing the Accuracy of Esri's Tree Crown Detection Model Utilizing UAV Imagery

Senior Inquiry
Augustana College Geography Department
Rock Island, Illinois

Lazlow Brian William Ziebel

December 13, 2024

Table of Contents

Table of Contents 2

Table of Figures 4

Table of Tables 5

Acknowledgements 6

Introduction 7

Literature Review 11

Technological Advancements in Drone Imagery and Remote Sensing 11

Applications and Implications for Forestry Management 12

Methodologies for Data Processing and Analysis of Tree Counts 13

Methods 15

Introduction 15

Research Design 15

Data Collection 16

Data Sources 17

Instruments 17

Ground Truthing 18

Data Analysis 18

Results 20

Ground Truthing 20

Ortho Photomosaic Production 21
Esri’s Deep Learning Package 24
Conclusion 29
References 32

Table of Figures

Figure 1. Three distinct layers in the forest's stratification referred to as the forest floor, understory, and canopy. Due to canopy minimizing resources to the lower layers, there is often competition. Cartography by Lazlow Ziebel 9

Figure 2. The study site boundary is defined by the purple rectangle. The main map depicts the location of the 35 images gathered to create an Ortho photomosaic. Cartography by Lazlow Ziebel 10

Figure 3. Canopy trees are depicted as dark grey diamonds and sub-canopy trees as light grey diamonds. The study boundary is defined by the purple border. 21

Figure 4 Ortho Photomosaics from Pix4DMapper, the September, October, and November Flights can be seen respectively. 22

Figure 5. Slight fragmentation in the production of the ortho photomosaic, likely due to the moving waters. A fountain can be seen. 23

Figure 6. True Positive detected tree crowns are symbolzied with a black star, totaling 67 detections. 26

Figure 7. True Positive detected tree crowns across the study site, totaling 42 individuals when compared with the ground-truthed trees, depicted by a black star. 28

Figure 8. Highlights all of the True Positive detected tree crowns across the study site, totaling 1 individual when compared with the ground-truthed trees, depicted by a black star. *The one individual identified was categorized as an understory tree, but was exposed after the leaves fell off the canopy trees. 29

Table of Tables

Table 1. Count of trees and confidence average were provided by the model. Number of true positives, false positives, and false negatives were all determined by referencing the ground-truthed trees with the detected tree polygons. 24

Table 2. Calculated performance metrics (Precision, Recall, Accuracy, and F1 Score) for different flight dates ranging from September-November 2024. 26

Acknowledgements

I would like to thank the Augustana College Geography Department for their guidance on this project, along with the Augustana College Environmental Department. I would specifically like to thank Dr. Chris Strunk for acting as my advisor and helping me design and complete my project. I would also like Dr. Jenny Arkle for inspiring me to find appreciation and a passion for the world of GIS and Cartography. Lastly, thanks to all my fellow Geography and Environmental peers for helping me develop myself as a writer through sessions of workshops.

Introduction

Forest management faces increasingly complex challenges due to climate change, changes in land use, and the spread of invasive pests (Choi and Park 2022). These factors contribute to altered forest ecosystems, making it difficult for managers to predict and respond to forest health, biodiversity, and productivity shifts. In this context, remote sensing technologies can play a crucial role in providing accurate, up-to-date information to support informed decision-making and effective management strategies (MacFarlane and Meyer 2005; Abbey 2022). Traditional forest management involves various field-based techniques, including surveying plant species, assessing tree density, evaluating forest health, and analyzing ecosystem dynamics. Decisions are then made based on these on-the-ground observations to address specific management goals, such as conservation, timber production, or biodiversity preservation.

Accurate information about forest composition, structure, volume, growth, and extent is essential for sustainable forest management (Ayres and Lombardero 2018). Over the past few decades, more attention has been brought to remote sensing as a tool for forestry management. Remote sensing is the acquisition of spatial information about the Earth's surface from an increased elevation (USGS). This includes using satellite imagery, UAVs or drones, and LiDAR applications (Onishi et al. 2022). Scholars are now starting to pair remote sensing with artificial intelligence (AI) for analysis (Onishi et al. 2022; Haq et al. 2020; Zhang et al. 2021). Machine Learning (ML) is a subfield of artificial intelligence that primarily relies on statistical algorithms. There is a subset of machine learning that uses artificial neural networks to process and analyze information called Deep Learning. Neural networks are composed of computational nodes that are layered within deep learning algorithms (Buchelt et al. 2024).

Classification and analysis within different areas of the forest are a necessity for proper management practices. These approaches target both forested study sites and urban study sites. Within urban study sites, deep learning can be utilized to create a map highlighting individual tree classification, reinforcing that deep learning is an effective way to classify urban tree species (Zhang et al. 2021). Similar approaches have been taken within dense forests combining high-resolution UAV images with deep learning to create a classification of forest inventory, proving that deep learning can be a very effective tool for gathering tree count information (Haq et al. 2020). One major distinction between an urban site and a densely forested site is that densely forested sites have increased resource competition among other trees (Trogisch et al. 2021). This creates three distinct layers in the forest's stratification referred to as the forest floor, understory or sub-canopy, and canopy (Hartmann et al. 2018). These three layers are consistently competing for resources (Figure. 1).

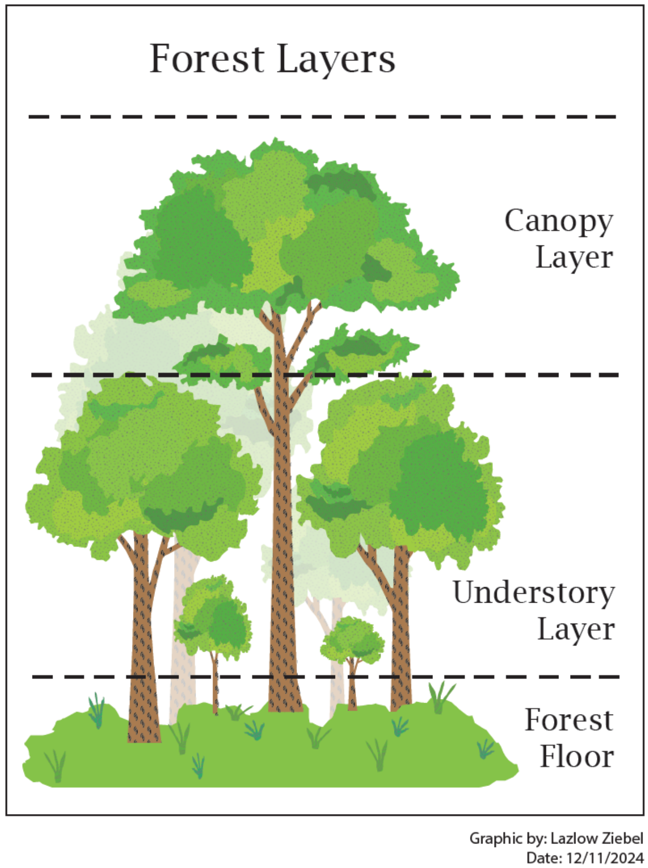

Figure 1. Three distinct layers in the forest's stratification referred to as the forest floor, understory, and canopy. Due to canopy minimizing resources to the lower layers, there is often competition. Cartography by Lazlow Ziebel

Although these are examples of where different machine-learning approaches have been utilized for forestry image classification, there is no universal deep-learning package that combines high-resolution UAV data with deep-learning methodologies. Therefore, the focus of this research is on applying a prebuilt deep-learning package trained for object classification of a densely forested area on Augustana College's campus (Figure 2). This study site was chosen because it includes multiple land uses; artificial slough, dense forest canopy, impervious surface (sidewalk, driveways, etc.), and varying topography. The forest of this study site is classified as a temperate deciduous forest (meaning that most of the trees lose their leaves during winter). It was also chosen because of its generally manageable size, sitting at 5 acres. This was taken into consideration, as a portion of this study involves physically ground-truthing the remote-sensed data.

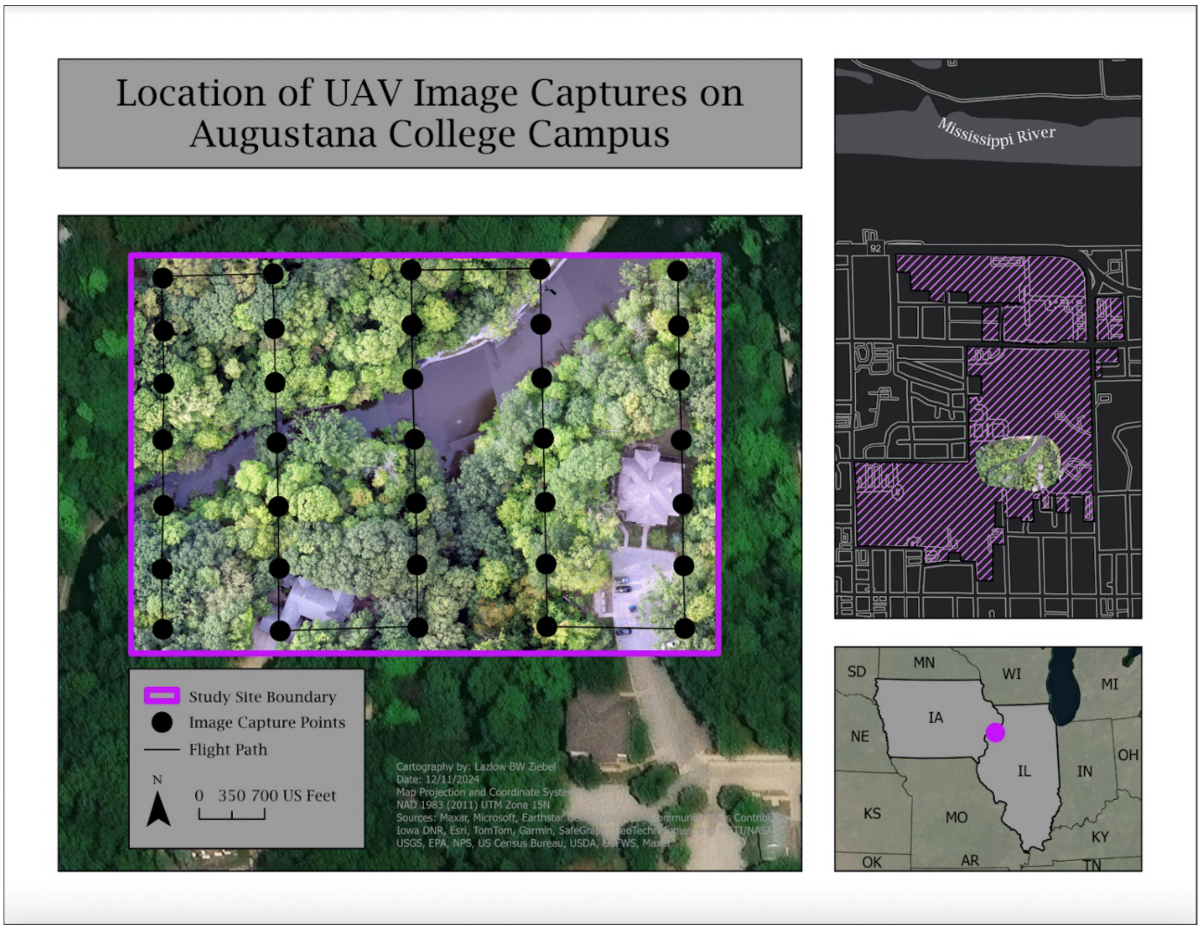

Figure 2. The study site boundary is defined by the purple rectangle. The main map depicts the location of the 35 images gathered to create an Ortho photomosaic. Cartography by Lazlow Ziebel

To better guide the development of drone remote sensing for sustainable forestry, this project focuses on using UAV Imagery in combination with Machine Learning to generate a Tree count without having to endure the long process of physical ID. This research aims to help fill in the knowledge gap around tree identification using artificial intelligence and deep learning. Due to the lack of publicly accessible Tree ID and Tree count software, this research aims to examine the accuracy of Esri's prebuilt Tree Detection model and determine if it can be applied to large-area study sites.

Literature Review

Technological Advancements in Drone Imagery and Remote Sensing

Remote sensing is the acquisition of spatial information about the Earth's surface from an increased elevation. Specifically, it is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance (USGS). Remote sensing is often used for change analysis, managing resources, and agricultural monitoring. Aerial footage is continually improving as technology advances. One method of remote sensing involves collecting data via satellites. Satellite data, while once a key tool in forest monitoring and resource management, is increasingly being phased out in favor of more reliable and accurate technologies. This shift is primarily due to the significant variability inherent in satellite data, which can result from factors such as atmospheric conditions, sensor limitations, and the timing of satellite passes over specific areas (Basu et al. 2015; Liang and Delahaye 2019).

While satellite images are good, a better resolution can be seen using unmanned aerial vehicles (UAVs) or drones. This allows for quick and precise monitoring of changes in forest damage at the spatial and temporal scales (Won and Park 2022). RGB (Red, Green, and Blue) imagery obtained from UAVs presents clear benefits for accurately gathering data that can be used for analysis. This is owing to its cost-effectiveness, user-friendly operation, and impressive image resolution, which can reach decimeter or even centimeter-level accuracy. With this data, image classification and object-based detection analysis can be done. Image classification is the process of assigning pixels to different classes based on their spectral values, often done using algorithms like Maximum Likelihood to extract thematic classes from images for generating statistical data (Merry et al. 2023). Object-based detection is very similar, but instead of pixels, whole objects are defined.

In contrast to alternative datasets, these qualities render UAV-based RGB data ideal for large-scale surveys of urban tree species. From this RGB imagery, image classification and object-based detection can be analyzed (Wu et al. 2024). Though usually more accurate, object-based detection is time-consuming. For an individual class to be identified, hundreds to thousands of RGB images are needed for an accurate classification.

Applications and Implications for Forestry Management

Within the forestry field, scholars are currently researching UAV usage for tracking wild forest fires (Wing et al. 2014), surveying land use in forests (Koh and Wich 2012), and intensive timber management (Wang et al. 2014). When tracking forest fires, UAVs allow for an “up-close examination” without putting any lives at risk. It also allows for real-time updates that can help confusion about fire intensity, fire severity, and burn severity (Lentile et al. 2006). Within forestry land use, ML is being used to create Land use/land cover (LULC) maps that can provide farmers and landowners with precise data on how their land is being used (Detsikas et al., 2024;). All these topics have one major similarity— large study sites. This is the reason why UAV Imagery is expanding very rapidly for urban planning and forestry management (Tang and Shao 2015).

Although the literature is emerging on species identification using RGB imagery, most of these studies are taking place in countries other than the US (examples include Japan, China, and New Zealand) and in addition to RGB are using LiDAR and satellite imagery (Onishi et al. 2022; Yi et al. 2016; Zhang et al. 2021). Because machine learning uses the data entered to create a model, most of these produced models do not correlate well to other international study sites. This is one major limitation of identifying tree crowns using ML.

Methodologies for Data Processing and Analysis of Tree Counts

One method that is becoming more popular at the intersection of Remote Sensing and Forestry management is machine learning (ML). ML is a subfield of artificial intelligence that primarily relies on statistical algorithms. It often requires substantial human intervention to extract relevant features and organize the data appropriately before a model can be trained. Machine learning techniques are broadly categorized into supervised and unsupervised learning, each offering unique advantages depending on the problem at hand. Supervised learning, for example, is used when labeled training data is available and can help create more accurate predictions. In contrast, unsupervised learning can discover hidden patterns in data without prior labels, which is particularly useful when dealing with large and complex datasets typical in forestry applications (Buchelt et al., 2024).

Due to machine learning's ability to quickly analyze large datasets of UAV imagery, it is advantageous when compared to ground-truthing, or physically classifying different data. When observing tree crowns, it is especially advantageous. Zhong et al. (2022) used a form of ML to create a map portraying the identification of individual tree species in a forested section of Northern China (Zhong et al. 2022). This study applied UAV and LiDAR data to a support vector machine (a supervised machine learning algorithm) to determine the total accuracy of the model. This study found that machine learning achieved a total accuracy of 84.62% in segmenting individual tree crowns. The success of this method underscores the growing importance of remote sensing technology in forestry, as it provides a reliable, non-invasive means of monitoring and managing forest ecosystems. This has enabled the rapid and accurate identification of tree species, making remote sensing technology one of the key methods for forest inventory.

Similar studies have also been applied to urban environments using machine learning to identify planted tree species (Zhang et al. 2021). Zhang et al. (2021) leveraged a multimethod study to determine which forms of AI yielded the greatest accuracies. This resulted in the findings that the classification accuracy of CNNs (a supervised form of ML) was much higher than other forms of AI. Though, it is important to note that because the studies were done in urban areas where there is a less complex background (ie. less tree canopy), the model would generally yield higher accuracy. Onishi and Ise (2021) constructed a machine vision system for tree identification and mapping using RGB imagery taken by a UAV and used a convolutional neural network (a form of machine learning) to obtain a tree count of the study site, and ultimately a map showing tree density. By hand annotating different tree crowns, and feeding them into the model, the overall accuracy of the model was able to reach 90%.

Methods

Introduction

The objective of this research project is to determine how accurate ESRI's deep learning model is in generating a tree count. Utilizing Augustana College's campus as a study site, UAV RGB data will be analyzed to determine the accuracy. The structure of this section is broken up into six sections: Research Design, Data Collection, Data Sources, Instruments, Ground Truthing, and finally Data Analysis. Each section provides a thorough overview of the methods and standards employed to attain the results.

Research Design

This research project's methods were inspired largely by the methods of Weinstein et al., (2019) which concludes that remote sensing can overcome a lack of labeled training data by using a semi-supervised network to successfully identify tree crowns. The basis of this training model data is derived from LIDAR, however, this will require training a new model for each geographic area while using both RGB and LIDAR training data. Using hand-generated outlines of tree crowns, Esri's available model was able to achieve an increased recall of 79% with a 60% total accuracy score. The model was trained on over 37 unique geographic sites across the United States, encompassing over 150 million trees (Weinstein et al. 2019).

The design for data collection was inspired and modeled after Onishi and Ise, (2021). The goal of their study was to create a CNN that could recognize tree crowns and, in turn, identify the species from seven different classes. I collected data through multiple drone flights under various leafing conditions, allowing for a more diverse study set. From this drone data, I produced an orthophotomosaic, enabling the user to import the data into ArcGIS Pro.

Data Collection

Remote sensing data was collected throughout multiple flight campaigns. Flights were conducted around 5:00 pm, spanning three different leafing conditions: September 16, 2024, which is the prime full leaf on green conditions, October 28, 2024, the onset of fall leaves, and November 25, 2024, the peak of fall leaf offset season, resulting in bare conditions. Using a UAV DJI Phantom 4.v2 (DJI, Shenzhen China), RGB spectral information can be collected using the onboard 1/2.3 CMOS sensor. A pre-planned mission was developed using the software Pix4DCapture, allowing for the mission to be synced with the UAV and flown asynchronously. Before the flight, a home point was established in a clearing to minimize disturbance among the canopy. The flight parameters were set as follows: the flight would take place at 325 Feet, the angle of the camera was 90 degrees, front overlap was set to 80%, side overlap to 70%, and picture trigger mode was set to fast mode (meaning the UAV does not stop to take an image, but instead flies through waypoints as it takes an image), and drone speed was set to normal. No ground-control points (GCPs) were needed as the Phantom 4v.2 has RTK positioning and ensures a geolocated image-capturing process. The ground sampling distance was 1.05 inches/pixel resulting in a 4:48 minute flight consisting of 35 images. From these 35 images taken by the UAV, an orthophotomosaic and digital surface model (DSM) were produced using the Pix4DMapper software.

Data Sources

All of the UAV data used in this research was sampled by myself around the same time, based on the same conditions. The average sampling day's forecast consisted of partly cloudy to cloudy conditions to ensure minimal shadows (this is also the reasoning behind flying at 5:00 pm), and a wind speed of 3-9 mph NW to minimize strong gusts pulling the UAV off course.

The study site was selected carefully for many reasons, one of which was land use. The site selected had a good variety of land use while still allowing for the forested area to be the dominant group. Within the study site, different factors would test the deep learning package's classifications: water, concrete, buildings, trees, etc. This would represent the grey area between urban areas and naturally forested areas. With how Illinois has urbanized over time, this study site is an appropriate representation of possible similarities observed in the Midwest. Another reason this study site was selected is that it allows for a large count of trees to be processed at once, within a smaller study site.

Instruments

As noted above, the instrument most heavily relied on is the UAV DJI Phantom 4.v2 (DJI, Shenzhen China). This UAV suits this research project well because it is very user-friendly, cost-efficient, and can produce a high amount of resolution, resulting in centimeter-level accuracy. One advantageous feature of the Phantom 4 is that the captured images are already georeferenced and geolocated when brought into Pix4DMapper. This saves countless hours and avoids using the "georeference wizard" tool in ArcGIS Pro. This UAV is also appropriate because no multispectral data is needed, meaning only RGB images are necessary for analysis.

Ground Truthing

To determine the accuracy of the DL, ground truthing or field validation refers to the process of verifying data or assumptions through direct observation or measurement in the real world. This study entails walking through the 5-acre site and recording whether each tree was a canopy tree or an understory tree. This hands-on approach allows for the collection of accurate, on-the-ground data. By cataloging the trees in this manner, a precise tree count can be established, which can be used to compare against and validate the findings produced by the deep learning package. The goal is to ensure that the model's predictions are correct and reliable by aligning them with the direct field observations. The ground truthing was carried out on November 24, 2024, taking 4.5 hours for one person to cover the entire study site. During the ground truthing, only non-saplings (trees with a diameter greater than 5 inches) were recorded. This was done utilizing ArcGIS Field Maps, and adding point data to the map when a canopy or sub-canopy tree was observed. One limitation to this form of ground-truthing is that the accuracy is limited to the GPS of a cell phone (accuracy or 5-10 ft).

Data Analysis

After creating the ortho photomosaic and DSM, these rasters were imported into ArcGIS Pro version 3.3.1 to begin the analysis. To examine the tree crowns, a prebuilt ESRI Deep Learning Package (DLPK) was downloaded and installed (ArcGIS). The Tree Detection DLPK is based on DeepForest and has been trained on data from the National Ecological Observatory Network (ESRI Analytics). After the package is installed, it can be accessed under the

Geoprocessing Pane, Deep Learning, and then by clicking *Detect Objects using Deep Learning*. After the model has run, an output polygon layer is added to the map that details a tree count, tree height, and confidence level.

Once the output layer is generated, comprehensive analysis can be performed to evaluate True Positives (TP), False Positives (FP), and False Negatives (FN), when the model's predictions are compared to the ground-truthed data. True Positives (TP) represent the number of instances where the model correctly identified a tree that was actually present in the ground-truthing. False Positives (FP) occur when the model incorrectly identifies an object as a tree that was not actually present, meaning the model made an error. False Negatives (FN) occur when the model misses a tree that is present in the ground-truthing data, failing to detect it. Once these values are calculated, they can be used to determine several key evaluation metrics: Accuracy is calculated as the proportion of correct predictions (TP) out of all predictions made (TP + FP + FN), providing a measure of overall model correctness. Precision is the proportion of correct positive predictions (TP) out of all predicted positives (TP + FP), reflecting how accurate the model's predictions are when it detects a tree. Recall is the proportion of correctly detected trees (TP) out of all actual trees in the ground truth (TP + FN), measuring the model's ability to identify all true trees. Finally, the F1 Score is the mean of Precision and Recall, offering a balanced metric when both false positives and false negatives are important. These metrics provide a comprehensive evaluation of the model's performance, helping to assess its accuracy, reliability, and ability to detect features on the study site.

Results

Ground Truthing

Since UAV imagery is shot from an overhead perspective, it can only access the tree crowns that are at the top of the canopy, distorting the overall count to exclude sub-canopy trees. This is critical when understanding the difference between a tree count developed in the field naturally, vs. Remote sensing. The predominant tree species in the forest were identified as Oaks (*Quercus*), Catalpa (*Catalpa speciosa*), and Hackberry (*Celtis occidentalis*). This process proved to be much more time-intensive compared to the 5-minute flight and subsequent GIS analysis used by the deep learning model to determine the tree count. The ground-truthing resulted in a total tree count of 227 canopy trees and 118 understory trees (Figure 3.)

Figure 3. Canopy trees are depicted as dark grey diamonds and sub-canopy trees as light grey diamonds. The study boundary is defined by the purple border.

When comparing the results of ground-truthing to the model, only the canopy tree count was assessed. The main reason for gathering a count of sub-canopy trees is to determine the potential limitation of excluding data from the study site. The ground-truthing could be further improved using a handheld Global Navigation Satellite System unit (GNSS).

Ortho Photomosaic Production

Three distinct ortho photomosaics were generated for each flight, with each mosaic corresponding to a different flight conducted during separate seasons, as organized in the introduction. As a result, significant variations in the natural elements—such as vegetation, tree cover, and overall landscape features—are evident across the mosaics. These seasonal differences reflect the natural changes in plant growth, foliage density, and overall environmental conditions that occur throughout the year. Such variations provide a dynamic view of the study area at different times, highlighting the impact of seasonal shifts on the landscape (Figure 4).

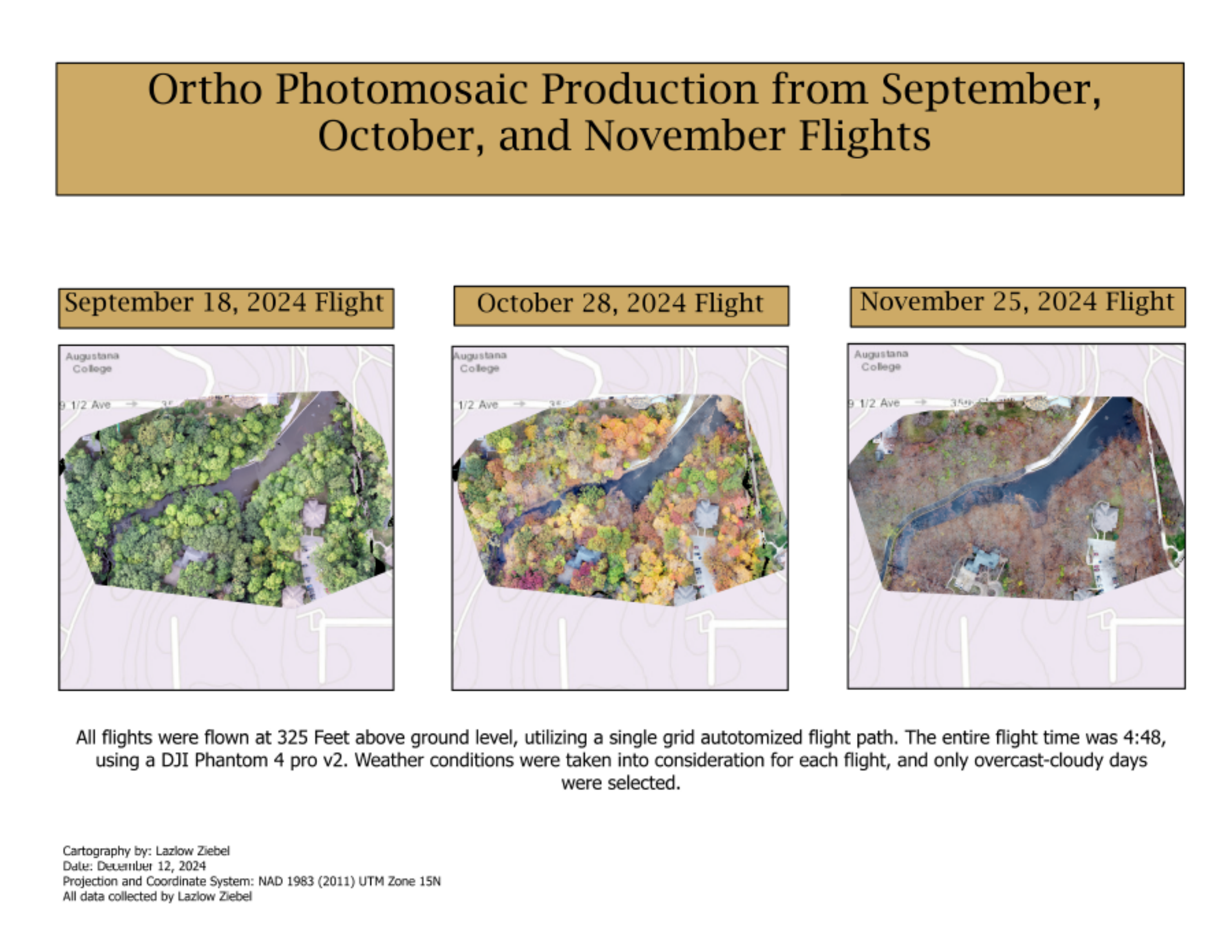

Figure 4 Ortho Photomosaics from Pix4DMapper, the September, October, and November Flights can be seen respectively.

The resolution of all three ortho photomosaics produced is exceptional due to the Ground sampling Distance of 1.05 inches/pixel. There were some points of contention when generating the ortho photomosaics in Pix4DMapper, resulting in areas of the slough being generated with fragmentation. When examining the September 18, 2024 mosaic, there is a slight fragmentation in the slough to the northeast. This is due to a lack of overlap between the 35 images used to generate the mosaic. Similar fragmentation occurred on the October 28, and November 25, 2024 flights due likely to the moving water, as there is a fountain in that area (Figure 5).

Figure 5. Slight fragmentation in the production of the ortho photomosaic, likely due to the moving waters. A fountain can be seen.

Esri's Deep Learning Package

The deep-learning package produced a differing number of tree crowns detected depending on the flight, resulting in 178, 111, and 1 tree(s) detected for the September, October, and November flights respectively. This data was then referenced with the ground-truthing point data; the detection numbers, confidence average, and number of True Positives (Table 1).

	1. September 18, 2024 Flight	2. October 28, 2024 Flight	3. November 25, 2024 Flight
Count of Trees Detected by DLPK	178	111	1
Confidence Average of Trees Detected	38.65%	31.70%	30.84%
Number of True Positives (correct detections)	67	42	N/A
Number of False Positives (incorrect detections)	111	69	N/A
Number of False Negatives (missed detections)	160	185	N/A

Table 1. Count of trees and confidence average were provided by the model. Number of true positives, false positives, and false negatives were all determined by referencing the ground-truthed trees with the detected tree polygons.

When looking at the September flight, the model did well in recognizing individual trees, considering the flight was during a time of peak leaf matter and density. The southern half of the study site (separated by the slough) had the highest tree density with a blend of individual trees within the canopy. Although Esri's Tree Detection deep learning model identified 178 tree crowns, only 67 were determined through ground truthing to be true positives (Table 1). From this, a precision, recall, accuracy, and F1 score of 37.6%, 29.52%, 19.82%, and 33.07% respectively were calculated (Table 2).

A precision score of 37.6% suggests that the model is misclassifying a significant portion of the predicted positives as false positives (incorrect predictions). Specifically, for every 100 positive predictions made by the model, only about 38 are correct. Given that there are 111 false positives, this suggests a high false positive rate, indicating that the model might be overly confident in predicting positive outcomes. A recall of 29.52% suggests that the model is missing a significant proportion of the actual positive cases (about 70% of them). Given that the false negative rate is 160, this indicates a relatively high number of false negatives, meaning the model is incorrectly classifying many actual positive instances as negative. A score of 19.82% indicates that the model is performing very poorly overall. It is correct less than 20% of the time, which means the vast majority of predictions are incorrect. When looking at the F1 Score of 33.07%, the model is not performing exceptionally well, it is somewhat balancing precision and recall. The F1 score is typically a good measure when there is a need to balance both false positives and false negatives. This indicates that both precision and recall are below optimal levels (Table 2).

	1. September 18, 2024 Flight	2. October 28, 2024 Flight	3. November 25, 2024 Flight
Precision	37.60%	37.84%	N/A
Recall	29.52%	20.79%	N/A
Accuracy	19.82%	14.19%	N/A
F1	33.07	21.35%	N/A

Table 2. Calculated performance metrics (Precision, Recall, Accuracy, and F1 Score) for different flight dates ranging from September-November 2024.

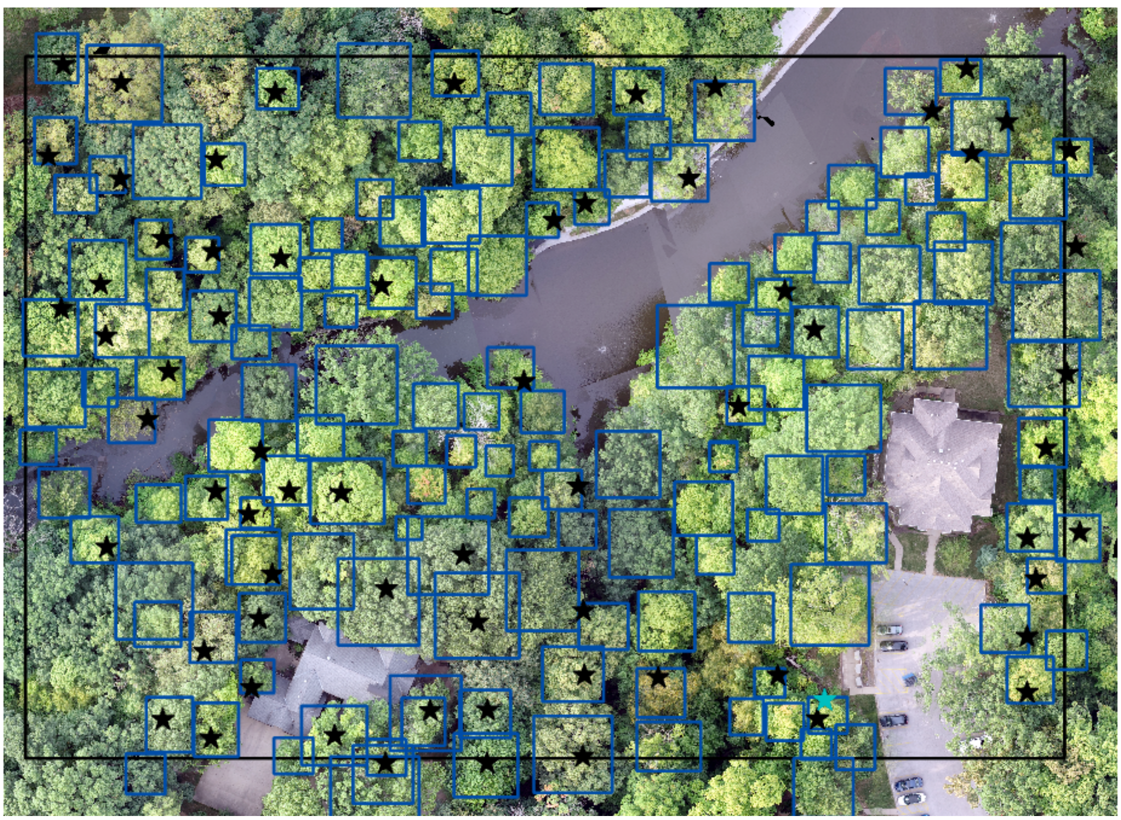

Figure 6. True Positive detected tree crowns are symbolzied with a black star, totaling 67 detections.

The October flight resulted in a precision, recall, accuracy, and F1 score of 37.84%, 20.79%, 14.19%, and 21.35% respectively. This data was calculated using Esri's Tree Detection deep learning model, identifying 111 tree crowns—only 42 of which were determined through ground truthing to be true positives (Figure 7). A precision score of 37.84% suggests that the model is not very reliable when it predicts positive outcomes. With 69 False positives detected, this reinforces the concept that the model is misclassifying many instances as positive when they are actually negative. A recall of 20.79% is very low, which suggests that the model is missing a large portion of the true positive instances. A score of 14.19% indicates that the model is incorrect 85.81% of the time. This is very low and suggests that the model is failing to correctly predict the class in almost all cases. An F1 score of 21.35% indicates that the model's overall performance is low, combining both precision and recall into a single metric (Table 2). This means that the model is not effectively identifying positive cases (low recall) and is also making many incorrect positive predictions (low precision).

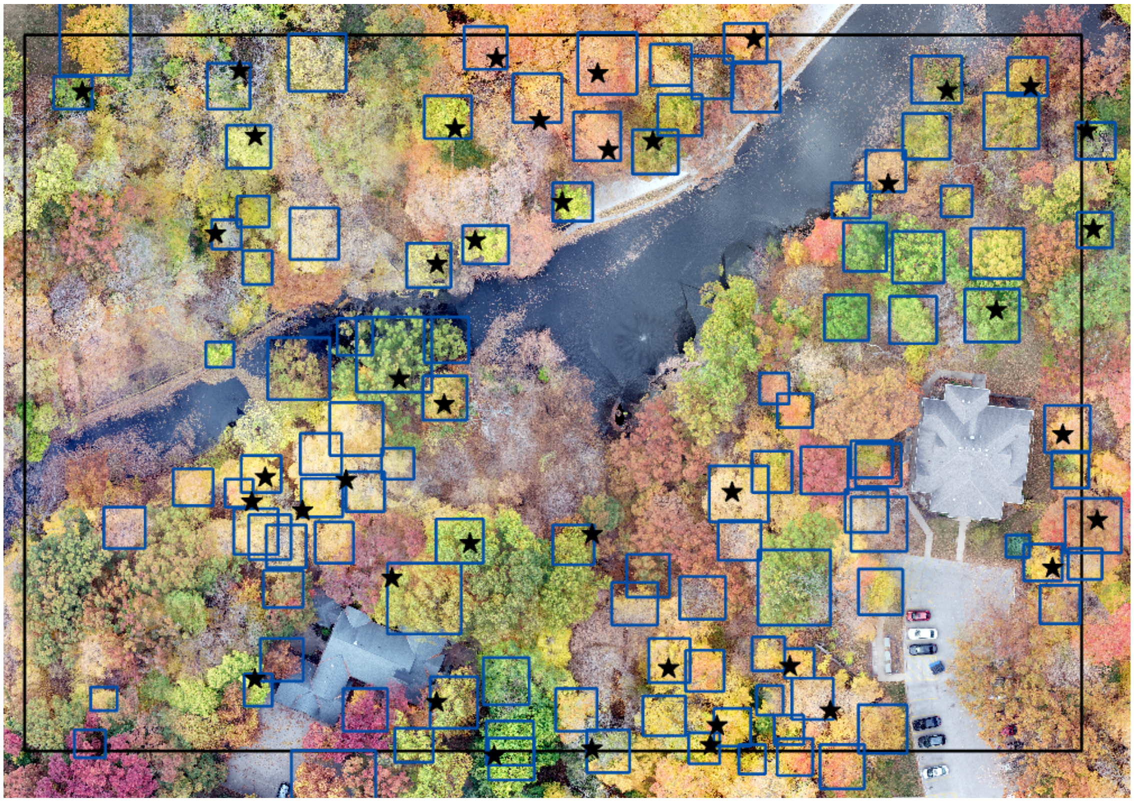

Figure 7. True Positive detected tree crowns across the study site, totaling 42 individuals when compared with the ground-truthed trees, depicted by a black star.

The last November flight data was inconclusive. The model struggled most with recognizing bare trees with little to no leaf matter, as depicted in Flight 3 of Table. 1. This makes sense, considering that the model was trained on full-bloom trees during high-intensity leafing conditions. The only tree detected was a conifer that was classified as an understory tree (Figure 8).

Due to the canopy trees mainly consisting of broad-leaved deciduous species losing their leaves in the winter, the conifer was exposed to sunlight, thus being captured via the UAV. Due to this being the only tree crown detection when the model was run, no calculations can be made to address the Precision, Recall, Accuracy, or F1.

*Figure 8. Highlights all of the True Positive detected tree crowns across the study site, totaling 1 individual when compared with the ground-truthed trees, depicted by a black star. *The one individual identified was categorized as an understory tree, but was exposed after the leaves fell off the canopy trees.*

Conclusion

This study aimed to evaluate the efficacy of a prebuilt, publicly available deep learning (DL) model, the Tree Detection Deep Learning Package (DLPK) by Esri, for classifying and detecting tree crowns in a temperate deciduous forest located on the Augustana College campus.

The study utilized UAV imagery collected during different leafing conditions and compared the results to ground-truthed data in order to assess the model's accuracy and precision. Although deep learning and other forms of artificial intelligence have shown promising signs within forestry management, the results from this study suggest that the model is currently not reliable enough for precise tree detection in complex, dense forest environments.

From the ground truthing, it was found that there was a total of 227 canopy trees and 118 understory trees present in the study area. The deep learning model’s performance varied significantly across the different flights, with precision, recall, accuracy, and F1 scores all indicating that the model did a suboptimal job at correctly identifying tree crowns. In the September flight, which captured the forest during peak leafing conditions, the model achieved a precision of 37.6% and recall of 29.5%, resulting in an overall low F1 score of 33.07%. These results reflect a high rate of false positives and false negatives, revealing that the model failed to detect a significant number of actual trees (trees present in ground truthing). The October flight exhibited similar suboptimal performance, with even lower recall and accuracy, suggesting that the model’s effectiveness diminishes as leaf cover decreases. The November flight was especially problematic, as the model failed to detect trees in the sparse conditions–one major limitation of the model.

Esri’s default deep learning package appears to be best suited for urban environments or open, low-density forests with well-defined tree crowns, rather than densely packed temperate deciduous forests, where the competition for light and space leads to obscured or overlapping tree canopies. Although the model boasts a 66% Precision Score, and 79% Recall, when tested on a real study site, these numbers were significantly lower. The model’s suboptimal performance in the leaf-off condition underscores the need for more robust training datasets that account for varying seasonal conditions. Forestry managers could develop a more accurate model by utilizing more hand-annotated tree crowns.

For future research, it is recommended that further refinements such as hand annotated polygons of individual tree crowns that include a broader range of seasonal conditions and forest structures be supplemented into the model. Due to time constraints, hand annotated tree crowns were not able to be imported into the model. While this study highlights the limitations of the current deep learning approach for dense forested areas, it also demonstrates the potential of UAV-based remote sensing combined with AI to enhance forest management practices. With continuing advancements in artificial intelligence and remote sensing, machine learning can become a powerful tool for forestry applications, supporting more efficient, accurate, and scalable management strategies. One significant limitation of the model, regardless of the accuracy, is determining how many understory trees are within the study site. Even with a 100% accurate model, understory trees cannot be detected utilizing UAV imagery and must be confirmed using ground truthing. As these models advance, there is an opportunity to apply these models to larger study sites, drastically reducing the time needed to ground truth.

References

Abbey, Tim. 2022. "Urban Tree Pests." Accessed May 14, 2024. https://extension.psu.edu/urban-tree-pests.

ArcGIS, dir. 2023. "Deep Learning in ArcGIS Pro Start to Finish." https://www.youtube.com/watch?v=JVizDjlN73c.

Ayres, Matthew P., and María J. Lombardero. 2018. "Forest Pests and Their Management in the Anthropocene." *Canadian Journal of Forest Research* 48 (3): 292–301. https://doi.org/10.1139/cjfr-2017-0033.

Choi, Won Il, and Young-Seuk Park. 2022. "Management of Forest Pests and Diseases." *Forests* 13 (11): 1765. https://doi.org/10.3390/f13111765.

Esri. 2024. "Classify Objects Using Deep Learning (Image Analyst)—ArcGIS Pro | Documentation." Accessed May 14, 2024. https://pro.arcgis.com/en/pro-app/latest/tool-reference/image-analyst/classify-objects-using-deep-learning.htm.

Haq, M. A., Rahaman, G., Baral, P., and Ghosh, A. 2021. Deep learning based supervised image classification using UAV images for forest areas classification. *Journal of the Indian Society of Remote Sensing* 49(3): 601–606. https://doi.org/10.1007/s12524-020-01231-3.

Koh, Lian Pin, and Serge A. Wich. 2012. "Dawn of Drone Ecology: Low-Cost Autonomous Aerial Vehicles for Conservation." *Tropical Conservation Science* 5 (2): 121–32. https://doi.org/10.1177/194008291200500202.

Liang, Man, and Daniel Delahaye. 2019. "Drone Fleet Deployment Strategy for Large Scale Agriculture and Forestry Surveying." In *2019 IEEE Intelligent Transportation Systems*

Conference (ITSC), 4495–4500. Auckland, New Zealand: IEEE. https://doi.org/10.1109/ITSC.2019.8917235.

MacFarlane, David W., and Shawna Patterson Meyer. 2005. "Characteristics and Distribution of Potential Ash Tree Hosts for Emerald Ash Borer." *Forest Ecology and Management* 213 (1–3): 15–24. https://doi.org/10.1016/j.foreco.2005.03.013.

Onishi, Masanori, and Takeshi Ise. 2021. "Explainable Identification and Mapping of Trees Using UAV RGB Image and Deep Learning." *Scientific Reports* 11 (1): 903. https://doi.org/10.1038/s41598-020-79653-9.

Onishi, Masanori, Shuntaro Watanabe, Tadashi Nakashima, and Takeshi Ise. 2022. "Practicality and Robustness of Tree Species Identification Using UAV RGB Image and Deep Learning in Temperate Forest in Japan." *Remote Sensing* 14 (7): 1710. https://doi.org/10.3390/rs14071710.

Tang, Lina, and Guofan Shao. 2015. "Drone Remote Sensing for Forestry Research and Practices." *Journal of Forestry Research* 26 (4): 791–97. https://doi.org/10.1007/s11676-015-0088-y.

Wang, Yangyang, Guangxin Bai, Guofan Shao, and Yukun Cao. 2014. "An Analysis of Potential Investment Returns and Their Determinants of Poplar Plantations in State-Owned Forest Enterprises of China." *New Forests* 45 (2): 251–64. https://doi.org/10.1007/s11056-014-9406-z.

Wing, M.G., Jonathan D. Burnett, and J. Sessions. 2014. "Remote Sensing and Unmanned Aerial System Technology for Monitoring and Quantifying Forest Fire Impacts." *International*

Journal of Remote Sensing Applications 4 (1): 18. https://doi.org/10.14355/ijrsa.2014.0401.02.

Wu, Jingru, Qixia Man, Xinming Yang, Pinliang Dong, Xiaotong Ma, Chunhui Liu, and Changyin Han. 2024. "Fine Classification of Urban Tree Species Based on UAV-Based RGB Imagery and LiDAR Data." *Forests* 15 (2): 390. https://doi.org/10.3390/f15020390.

Yi, Piyuan, Peng Tong, and Yingjun Zhao. 2016. "A Decision Tree Classification Method Combining Intensity and RGB Value for LiDAR Data." In *Geo-Informatics in Resource Management and Sustainable Ecosystem*, edited by Fuling Bian and Yichun Xie, 569:177–89. *Communications in Computer and Information Science*. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-49155-3_17.

Zhang, Chen, Kai Xia, Hailin Feng, Yinhui Yang, and Xiaochen Du. 2021. "Tree Species Classification Using Deep Learning and RGB Optical Images Obtained by an Unmanned Aerial Vehicle." *Journal of Forestry Research* 32 (5): 1879–88. https://doi.org/10.1007/s11676-020-01245-0.

Zhong, Hao, Wenshu Lin, Haoran Liu, Nan Ma, Kangkang Liu, Rongzhen Cao, Tiantian Wang, and Zhengzhao Ren. 2022. "Identification of Tree Species Based on the Fusion of UAV Hyperspectral Image and LiDAR Data in a Coniferous and Broad-Leaved Mixed Forest in Northeast China." *Frontiers in Plant Science* 13 (September). https://doi.org/10.3389/fpls.2022.964769.