Assessing and Predicting the Accuracy of TLS2trees Using Machine Learning Models

WQQR8

10207 words

Thesis submitted for consideration towards a degree of
MSc Environmental Modelling,
Dept of Geography,
UCL (University College London)

Supervisor: Prof. Mathias Disney

August, 2023

ABSTRACT

This study aims to evaluate the performance of the TLS2trees software in tree segmentation tasks, employing a variety of metrics such as Percentage Agreement (PA), Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Mean Squared Error (MSE), and Normalized Hausdorff Distance (NHD). A dataset comprising 417 refined tree samples was analyzed both numerically and visually.

Three machine learning models—Random Forest, Support Vector Machine, and Multi-Layer Perceptron—were employed to predict the accuracy of tree segmentations. Despite rigorous hyperparameter tuning, all models demonstrated limited predictive accuracy. The study acknowledges its limitations, including a sparse dataset and the absence of standardized data collection methodologies.

The research culminates in a set of recommendations for future work, such as the exploration of additional metrics, the employment of more sophisticated machine learning models, and the development of a real-time quality assessment feedback loop in the TLS2trees system. This study serves as a comprehensive evaluation and offers a roadmap for enhancing the efficiency and accuracy of machine learning-based tree segmentation algorithms, with potential applications in environmental monitoring and conservation.

ACKNOWLEDGEMENTS

I would like to extend my deepest gratitude to everyone who has been a part of this journey and contributed to the success of this study.

First and foremost, I owe a profound debt of thanks to my supervisor, Prof. Mathias Disney, for his invaluable guidance, constant encouragement, and meticulous attention to detail. His expertise and insights have been indispensable in shaping this research. I am also extremely grateful to Dr. Phil Wilkes for his unwavering support, invaluable suggestions, and for always pushing me to strive for excellence.

I would also like to express my heartfelt gratitude to my family, whose endless love and support have always been my strength. Your belief in me has made all the difference.

To my friends, thank you for your camaraderie, understanding, and the countless moments of levity that lightened the load. You made this journey both memorable and enjoyable.

Last but not least, I would like to thank all the faculty members, staff, and my fellow researchers who have contributed in numerous ways to this work. Your collective wisdom and cooperation have enriched this study and me, professionally and personally.

Thank you all for your contributions, direct or indirect, to this research. Your support has been the cornerstone of this endeavor.

LIST OF CONTENT

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Research background

Laser scanners have introduced a novel dimension to the area of surveying by enabling advanced data capture capabilities (Neumann et al., 2004). One of its branches, terrestrial laser scanning, has undergone significant advancements during the early 2000s, transitioning from a subject of research and development to a commercially available geo-data technology (Lemmens, 2011). This technique has significantly revolutionized forestry research. It was firstly introduced for fundamental forest measurement (Calders et al., 2020). Recently, notable advancements of sensor and algorithm provide detailed, three-dimensional data about forest structures, enabling the detailed examination of individual trees and the whole forest ecosystem (Disney, 2018).

The precise measurement of individual trees serves as the basis for research and management of forests, as it plays a significant role in elucidating ecological problems and processes, such as quantifying forest biomass, carbon storage, and habitat structure, among other things (Xu et al., 2021).

Especially, TLS provides more possibilities in determining Above Ground Biomass (AGB), which represents the carbon stored within terrestrial ecosystems. Traditional methodologies evaluate AGB using plot-level procedures, typically including areas of 1 hectare or larger, relying on inventory measurements combined with allometry. However, the operationalization of TLS in this domain has faced various obstacles. A notable obstacle encountered when working with TLS data involves the extraction of individual trees from the intricate and multi-scale point clouds. The extraction of individual trees requires substantial computation resources and advanced algorithms capable of discerning the intricate and interconnected patterns found within forest ecosystems. Traditionally, this work has been performed manually, utilizing interactive point cloud editing tools, which results in a time-consuming process (Chave et al., 2014).

In response to this formidable task, Dr. Phil Wilkes devised an open-source software called "TLS2trees," an automated pipeline specifically designed to extract individual trees from TLS data at a large scale. This software uses deep learning (DL) algorithms to discern distinctive tree shapes and structures, thereby efficiently segregating these entities from the 3D point cloud that represents lots of trees (Wilkes et al., 2022).

The application of this technology has great potential in forestry. If successful, TLS2trees can significantly enhance the efficiency and scalability of forest structural analysis by streamlining the process of extracting individual trees. However, like any novel technology, it is crucial to assess its accuracy. The basic purpose of this dissertation is to evaluate the accuracy of trees extracted by TLS2trees, providing a much-needed assessment of this

promising approach.

Traditionally, the evaluation of segmentation results necessitates a comparison with manually segmented datasets. However, the manual process is not only time-consuming and labor-intensive but also introduces variability due to differences in human judgment. Recognizing these limitations, we aimed to develop a more efficient and reproducible evaluation method. Specifically, our research focuses on utilizing tree model parameters extracted from the TLS2trees segmentation to predict the accuracy of the point cloud data, thereby circumventing the need for manual comparison and reducing subjectivity in the evaluation process.

Objective

- Make an assessment metric to evaluate the accuracy of trees extracted by the software TLS2trees
- To predict the accuracy of trees extracted by the software TLS2trees from tree model data by applying machine learning algorithms.

LITERATURE REVIEW

Laser Scanning

Laser scanning has emerged as a pivotal technique across several disciplines, acclaimed for its exceptional accuracy, speed, and adaptability. The fundamental basis of its origin lies in the concept of projecting a laser beam towards a designated objective and measuring the time it takes for the beam to return, usually referred to as the "time-of-flight." This methodology enables precise determination of distances (Pfeifer and Briese, 2007).

The emergence of diverse scanner types has been driven by the need to cater to different applications. For example, terrestrial, aerial, and handheld scanners, each possesses distinct capabilities that are specifically designed to cater to particular fields. As laser scanners capture their environment, they generate dense datasets known as 'point clouds'. These collections of points, representing the scanned environment, are essential in various analytical tasks.

When a laser scanner records its surroundings, it produces a substantial collection of data points. The unprocessed point cloud data often contains outliers or noise, which commonly arise from phenomena like as reflections, refractions, or faults in the scanning process. The elimination of noise is of the utmost significance in order to provide a clean and accurate representation of the scanned environment (Besl and H.D. McKay, 1992). After undergoing processing, the data has the capability to be presented in many visual representations. The

inherent flexibility of this technology has facilitated the development of 3D models (Remondino and El-Hakim, 2006), topographical contour maps (Baltsavias, 1999), and various other essential visual representations.

Laser scanning has found significant applications across diverse fields. In cultural heritage, it assists in detailed documentation of historical sites (Yastikli, 2007; Montuori et al., 2014). The technology has been instrumental in urban growth tracking (Pieraccini et al., 2006; (Erener, Sarp and Karaca, 2020); Zhu et al., 2020; Stilla and Xu, 2023) and in monitoring the health of building structures (Park et al., 2007). Laser scanning's precision aids in landslide surveillance (Prokop and Panholzer, 2009; Xu et al., 2023) and has revolutionized forest structure assessments, enabling accurate evaluations of various forest metrics (Sumnall, Hill and Hinsley, 2022; Wallace et al., 2016).

Technological advancements have further expanded laser scanning's potential. The integration of Artificial Intelligence (AI) has resulted in notable enhancements in the efficiency and precision of data processing (Hodgetts, 2013). In addition, the integration of laser scanners with dynamic platforms like drones has significantly enhanced data collecting capabilities, leading to the emergence of mobile laser scanning (Campana, 2017). Nevertheless, the field isn't devoid of challenges. Concerns such as the management of large datasets, the need for real-time processing, and the substantial expenses related to high-quality laser scanners are currently of great importance (Dai et al., 2017).

Terrestrial laser scanning

Terrestrial Laser Scanning (TLS), a subset of the broader laser scanning technology, has emerged as a powerful tool for capturing high-resolution, three-dimensional spatial data of terrestrial environments. Unlike airborne or mobile laser scanning systems, TLS devices are typically ground-based, stationary instruments that measure the distance between the scanner and the target object by emitting laser beams and then analyzing the reflected signals (Brasington, Vericat and Rychkov, 2012).

The origins of TLS technology can be traced back to the latter half of the 20th century, coinciding with the advancements made in laser-based measurement technologies. The advent of Terrestrial Laser Scanning (TLS) can be attributed to the convergence of advancements in laser technology, computer capabilities, and the demand for high-resolution spatial data. Early TLS systems were primarily used in industrial applications, such as manufacturing and reverse engineering. (Boehler, Heinz and Marbs, 2002) As the technology matured, the early 2000s saw a broader adoption of TLS in various fields. Its ability to capture detailed 3D data quickly made it invaluable for professionals in archaeology, civil engineering, and geology (Heritage and Hetherington, 2006). Over the span of the past decade, there have been notable advancements in the field of TLS technology. Contemporary systems include enhanced scanning rates, improved precision, and expanded range. The incorporation of TLS with many

technologies, including photogrammetry, has emerged as a significant trend (Gordon and Lichti, 2007).
A conventional TLS system consists of three main components: a laser source, a scanner, and a detector (Lichti and Jamtsho, 2006). The scanner emits laser pulses directed towards a designated target, and afterwards measures the duration it takes for the pulse to return to the scanner after being reflected off the target. The utilisation of time-of-flight measurement, in conjunction with the established speed of light, enables the computation of the distance between the scanning device and the designated object. Contemporary TLS systems possess the capability to acquire an extensive number of data points within a single second, reaching the magnitude of millions. Consequently, this enables the generation of very intricate point clouds that accurately depict the scanned surroundings (Shan and Toth, 2018).

One of the main benefits of TLS is in its ability to rapidly acquire geographical data with high resolution. Not only does this practise result in a reduction of fieldwork duration, but it also yields datasets that exhibit a higher level of comprehensiveness compared to conventional surveying techniques (Lugo et al., 2022). Nevertheless, like every technological advancement, TLS also possesses certain constraints. Line-of-sight restrictions mean that any object or terrain occluded from the scanner's view won't be captured (Beraldin et al., 2005). Furthermore, the precision of TLS may be influenced by atmospheric circumstances, the reflectance properties of the object being scanned, and the distance between the scanner and the target (Lichti, 2007).

In the future, the potential applications and efficiencies of TLS are expected to extend further, because continuous developments in sensor technology, data processing algorithms, and integration with other systems such as photogrammetry or unmanned aerial vehicles (UAVs).

TLS in forestry study

Terrestrial Laser Scanning (TLS) has emerged as a transformative tool in forestry studies, offering a paradigm shift in how forest ecosystems are quantified and analyzed. This section reviews the literature on the applications of TLS in forestry.

Traditional methods to assess tree height, diameter, and crown measurements are characterised by their labour-intensive nature and susceptibility to human error. TLS offers a non-invasive method for accurately capturing these data. Multiple studies have provided evidence about the effectiveness of TLS in estimating various structural characteristics such as tree height, crown width, and stem diameter (Calders et al., 2014).

The estimation of forest biomass plays a critical role in carbon accounting and the comprehension of forest health. The utilisation of Terrestrial Laser Scanning (TLS) has been implemented to generate comprehensive three-dimensional (3D) representations of trees, facilitating a more precise estimation of biomass compared to traditional methodologies

(Wilkes et al., 2017).

Additionally, TLS has the capability to detect tiny changes in the configuration of trees, so serving as a potential indicator of health-related concerns or reactions to external factors impacting the environment. This feature renders it a helpful instrument for the surveillance of tree health and identification of initial indications of diseases or infestations caused by pests (Liang et al., 2016).

While airborne laser scanning struggles with capturing the forest understory due to canopy occlusion, TLS demonstrates notable proficiency in this particular area. Airborne laser scanning has challenges in effectively collecting the forest understory as a result of canopy occlusion; however, TLS demonstrates notable proficiency in this particular area. The acquisition of detailed information regarding understory structure is crucial in order to comprehend forest ecology and effectively simulate fire behaviour. This can be accomplished through the utilisation of TLS technology (Wallace et al., 2012).

In summary, TLS presents unprecedented advantages in the field of forestry research, encompassing comprehensive structural analysis and health monitoring capabilities. The capacity to acquire high-resolution 3D data on forest stands offers researchers a thorough comprehension of forest ecosystems, hence assisting in the enhancement of forest management and conservation efforts.

Single tree segmentation

Single tree segmentation, the extraction of individual trees from a forest stand in remotely sensed data, commonly referred to as, has emerged as a prominent topic of interest within the field of forestry study. The present research provides an in-depth analysis of the methodology, applications, and issues pertaining to the segmentation of individual trees.

The initial efforts in the segmentation of individual trees predominantly relied on the utilisation of aerial images. The prevailing practise was manual interpretation, wherein analysts would utilise stereoscopic imaging to identify specific tree crowns (Wulder, Niemann and Goodenough, 2000). Although the traditional method can already provide a visual representation of forest stands, it had several drawbacks. The process was labour-intensive and time-consuming, frequently necessitating careful and detailed demarcation of individual tree canopies, particularly in large regions. The utilisation of a manual approach introduced subjectivity, as various analysts may interpret the same image differently, thus leading to inconsistencies. Aerial images primarily captured spatial and spectral data, but they lack the comprehensive vertical forest structural information provided by contemporary methods such as LiDAR. In dense forests, the presence of taller tree crowns may obstruct the view of shorter trees or understory vegetation, hence introducing the possibility of observational errors. Moreover, the images' quality may be affected due to atmospheric conditions, the presence of

shadows, and variations caused by seasonal changes. The cost of regularly updating aerial imagery and the challenges of automating the segmentation process further limited its efficiency and adaptability in the face of rapid technological advancements (Wulder et al., 2012).

Single tree segmentation was revolutionised by the introduction of LiDAR. Even in deep forests, the high-resolution vertical data enabled the identification of specific tree crowns (Hyyppä et al., 2008). Furthermore, a variety of advanced algorithms have been developed for single tree segmentation. Some of the most commonly used approaches in image segmentation include region-growing algorithms, watershed segmentation, and clustering techniques. In recent times, there has been a notable emergence of machine learning and deep learning methodologies, which have demonstrated considerable potential (Hans Ole Ørka et al., 2013).

While there have been significant advancements, challenges persist. Despite notable progress, persistent obstacles remain. The presence of overlapping crowns, dense understory vegetation, and diverse tree shapes can introduce complexities in the process of segmentation. Moreover, it is important to note that various forest types and structure may necessitate distinct methodologies (Vauhkonen et al., 2011). In the future, the integration of multiple sources of data, such as the combination of LiDAR and hyperspectral images, is regarded as a highly promising direction. Furthermore, the increasing utilisation of deep learning techniques, which possess the capability to autonomously acquire features from extensive datasets, has the potential to enhance the accuracy of single tree segmentation (Mäyrä et al., 2021).

TLS2trees

TLS2trees is a free open-source software developed by Dr. Phil Wilkes, which adopts deep learning algorithms to identify tree properties from plenty of TLS point clouds (Wilkes et al., 2022a). The workflow integrates adapted iterations of previously published procedures alongside new techniques. Its processing methodologies are horizontally scalable, which is accomplished by the implementation of a tile-based workflow. The codes are available at: https://github.com/philwilkes/TLS2trees.

The process that generates a series of individual tree point clouds S from the raw point cloud P can be concluded as follows: pre-processing -> semantic segmentation -> instance segmentation.

The specific processing starts by taking scans for the raw point clouds. The scans are adjusted based on their associated rotation matrices, then projected onto a 10x10 meter grid. This grid is divided into smaller units, or tiles, for processing (Martin-Ducup et al., 2021). The issue in TLS data that closer objects have more details than far ones (Burt, Disney and Calders, 2018) is also mitigated by applying Voxel Centre Nearest Neighbour method of PDAL (Bell, Chambers and Butler, 2020).

Then, the points are labelled into homogeneous groups, such as ground, leaf, wood, etc. Previously, points were classified alongside or after instance segmentation, but this can lead to issues, especially with leaf presence. In this version, a modified method called FSCT semantic segmentation method is used first to classify points into four main categories: ground, woody, leaf, or coarse woody debris (Krisanski et al., 2021).

The next step, instance segmentation, which is the most creative part of TLS2trees, is crucial in identifying and segmenting individual trees from the provided data. Initially, the software focuses on the woody components, essentially the structural elements of the trees. Once these foundational structures are identified, the software then associates the leaf components, ensuring that foliage corresponds correctly to its respective tree. Several advanced methods are integrated to achieve the target. Firstly, the process adopts a nuanced approach, employing a "convex hull" for each cluster (Wilkes et al., 2022a). This enveloping technique outlines the space each cluster occupies, preserving critical spatial information. By utilizing the k-nearest neighbour method, the software identifies vertices of these hulls that are close to other neighbouring clusters, emphasizing spatial relations. The weight or importance of connections between clusters is determined by the Euclidean distance between their edges, focusing on genuine spatial connectivity (Tao et al., 2015). Then, to facilitate this accurate pairing and manage the intricate nature of the dataset, TLS2trees utilizes Dijkstra's shortest path method (Dijkstra, 1959). This algorithm aids in discerning and mapping the connections between the tree components, ensuring each segment is correctly associated with its respective tree. Moreover, the software also introduces the RANSAC cylinder fitting method, which further refines the process by distinguishing tree bases from other elements, such as leaves (Ester et al., 1996). Accurate identification of tree stems, irrespective of their size, is crucial in order to avoid errors such as the merger of smaller stems with bigger ones.

QSM

Quantitative Structure Models (QSM) have become an integral tool in forestry, providing detailed 3D representations of individual trees. It encapsulates a comprehensive set of parameters to represent individual trees in detail. These parameters include tree height, Diameter at Breast Height (DBH), volume, branch geometry (encompassing branch length, diameter, and angle), crown dimensions (width, depth, and surface area), Leaf Area Index (LAI), tortuosity, branching density, internode length, and biomass (Hackenberg et al., 2015).

Among them, DBH measures the tree trunk's diameter approximately 1.3 meters off the ground. LAI represents the one-sided leaf tissue area per ground unit. Tortuosity gauges a branch's deviation from straightness, often indicative of growth patterns or environmental responses. Branching Density calculates the number of branches per unit volume, providing insights into the tree's growth patterns and health. Internode Length means the distance between successive branching points or nodes (Hackenberg et al., 2015).

The QSM emerged in response to the need for a detailed understanding of tree architecture and volume estimation, offering a more comprehensive perspective than traditional forest metrics (Raumonen et al., 2013). A primary application of QSM is in biomass estimation, where its detailed structural data facilitates more accurate biomass calculations compared to conventional allometric methods (Calders et al., 2014). Furthermore, QSMs have been instrumental in monitoring tree growth and assessing health, aiding in the early detection of diseases or structural anomalies (Åkerblom et al., 2017). However, the approach is not without challenges. Accurate segmentation, especially in dense canopies, and the computational demands of generating intricate models remain significant concerns (Côté et al., 2012).

Looking ahead, with the advancements in LiDAR technology and computational techniques, the potential and applications of QSMs are set to expand. The integration of QSMs with other data sources and the development of refined algorithms for swifter and more precise model generation are areas of ongoing research (Hackenberg et al., 2015).

Machine Learning in Forestry Study

Integrating machine learning into forestry studies has introduced a transformative approach that augments traditional methodologies. Emerging in tandem with advancements in computational capabilities and data availability, machine learning initially provided solutions for simple pattern recognition tasks but has since evolved to tackle more complex challenges central to contemporary forestry research (Jain, Duin and Mao, 2000, Reichstein et al., 2019).

Machine learning's applicability spans a broad range of forestry-related tasks. Deep learning algorithms, for example, have found utility in forest health monitoring, capable of rapidly identifying diseases and environmental stressors (Liang et al., 2016). Alongside this, machine learning has advanced the precision of biomass estimation, contributing to our understanding of forest ecology and carbon cycles (Mascaro et al., 2014). Furthermore, machine learning algorithms, particularly those capable of handling high-dimensional remote sensing data, have proven effective for identifying tree species, thereby informing biodiversity conservation efforts (Rina et al., 2023, He et al., 2016).

Despite these advancements, challenges persist. The effectiveness of machine learning models hinges on the quality and volume of training data. Poor data quality can result in model overfitting and reduce the generalizability of the model to new or unseen data (Danks, Ray and Shmueli, 2023). Additionally, interpreting complex models, especially when used to make policy decisions, remains an area of active research (King, Tomz and Wittenberg, 2000).

Looking ahead, the fusion of machine learning with emerging technologies such as the Internet of Things (IoT) and unmanned aerial vehicles (UAVs) is promising (Srikrishnan Divakaran, 2023). Such integration could facilitate real-time monitoring and adaptive forest

management. Moreover, advancements in machine learning techniques like transfer learning and ensemble methods are anticipated to further enhance model robustness and accuracy (Sorrentino et al., 2023). In summary, machine learning's application in forestry studies presents groundbreaking opportunities for research and practical management, underscoring its importance in future academic and industrial initiatives.

METHODOLOGY

Data Acquisition and Preparation

The datasets used in this study were directly acquired from other research. Those datasets were primarily captured from terrestrial laser scanning in two forests, wytham woods (Calders et al., 2022) and Malaysian tropical forest (Beyene et al., 2020). To ensure high-resolution and accurate scans, all scanning was conducted by using RIEGL VZ-400 scanner (RIEGL Laser Measurement Systems GmbH, Horn, Austria). The scanner's technical superiority, combined with its proven track record in similar studies, made it the optimal choice for our research.

The scanning methodology for each plot varied, yet adhered to systematic approaches to guarantee coverage and data quality. For many plots, a regular grid formation was followed (Wilkes et al. 2017). This methodology ensures a structured and systematic capture of data across the entire plot area. The star formation aims to cover diverse angles and perspectives, especially crucial for dense or intricate forest configurations. Importantly, at every designated scan position, two individual scans were undertaken. To facilitate precision and coherence between these scans, manually placed reflectors served a pivotal role. Positioned at strategic intervals, these reflectors functioned as tie points, guiding the co-registration of scans. Their placement ensured the data's spatial accuracy and provided a consistent reference throughout the scanning process.

Post-acquisition, the raw scan data underwent preliminary processing with the RiSCAN Pro software, which ranged from versions 2.1 to 2.9. This crucial step involved co-registering the individual scans to a unified coordinate system, tailored for each plot. A set of 4 x 4 transformation matrices was exported for each scan when registration was finished.

The primary focus of this study was to evaluate the accuracy of the TLS2trees software in individual tree delineation. In order to robustly evaluate the capabilities of the TLS2trees a comparative analysis was needed between the TLS2trees segmented data and a correct data. Recognizing the inherent challenges in obtaining 'true' individual tree data directly from the natural environment, particularly considering potential overlaps between trees, it was assumed that manually segmented tree point clouds would serve as a gold standard, although it may be different under different people's operation. Thus, the first dataset was derived through the

application of TLS2trees, producing point clouds for individual trees. In contrast, the second dataset was meticulously prepared through manual segmentation, crafted to act as a benchmark against which the automated results could be assessed.

The methodology for crafting the manually segmented dataset drew from established literature. In the WYT plots, an approach was adapted where all trees within a central 1.4 ha plot were segmented, derived from a larger 6 ha scanned area, thereby mitigating potential edge effects observed in other plots (Calders et al., 2018). The manual segmentation was facilitated using tools like RiSCAN Pro or CloudCompare v2.X for MLA01. In contrast, data from WYT were initially segmented with treeseg (Burt et al., 2019) and subsequently underwent manual refinements, such as the elimination of overlapping crowns. During this process, each tree was verified by a seasoned operator to minimize commission and omission errors.

Post-segmentation, tree point cloud underwent semantic segmentation into leaf and wood points using the TLSeparation Python package (Vicari, n.d.), except for WYT data, which was captured in leaf-off conditions. Subsequent modeling of tree structural traits was executed using the TreeQSM, focusing only on the point cloud classified as wood.

Accuracy Assessment Metrics

The core of this assessment approach is the juxtaposition of tree segmentations derived from TLS2trees against those meticulously curated through manual segmentation. The fundamental challenge lies in reliably comparing three-dimensional point clouds. Normally, there are five predominant metrics utilized for assessing the congruence between two three-dimensional point clouds: voxel accuracy, IoU, DSC, MSE, and Normalized Hausdorff distance (Taha and Hanbury, 2015). However, directly deploying these metrics on three-dimensional images requires an intricate and massive calculation after voxelization, which, although precise, is computationally intensive and time-consuming.

To circumvent these computational challenges, our methodology adopts a preliminary two-dimensional approach, aiming to offer a balance between computational efficiency and representational accuracy (Wang et al., 2023). This method is anchored in a process called "tri-view projection". By strategically discarding one of the three spatial dimensions of each point in the point cloud, we can obtain two-dimensional projections of the dataset. For instance, consider a point A with coordinates (x, y, z). By omitting the z-coordinate, the projection of the point on the xy-plane is given by the coordinates (x, y). By systematically discarding each of the three dimensions in turn, we can derive the point cloud's two-dimensional projections from three distinct perspectives. This method, termed the "tri-view method", strives to maximize the retention of the inherent three-dimensional characteristics of the point cloud in a two-dimensional format.

Upon completion of these projections, each two-dimensional representation is then pixelized.

Following this, we embark on computing the desired metrics: pixel accuracy, IoU, DSC, MSE, and Normalized Hausdorff distance for each pixelized image. The forthcoming sections delve into the detailed definitions and computational methodologies associated with each of these metrics. And the python script used to calculate those metrics is available at: https://github.com/ucfaune/GEOG0105.git

Pixel Accuracy

Pixel accuracy, often simply referred to as "accuracy," is a fundamental metric in the realm of remote sensing and image classification. Historically rooted in image analysis, its application has expanded with the advent of point cloud data, especially with the proliferation of LiDAR and photogrammetric techniques. Accuracy provides a straightforward measure of how well a classification or segmentation algorithm performs by comparing its results to a ground truth or reference data (Congalton, 1991).

$$PA = \frac{Number\ of\ correctly\ classified\ pixel}{Total\ number\ of\ pixel} \qquad \textbf{Eqn. 1}$$

This equation gives a ratio of the correctly classified pixels to the total pixels, providing a quick overview of the classification's overall correctness (Foody, 2002).

Intersection Over Union

Intersection over Union (IoU), also known as the Jaccard Index, is a metric that has gained prominence in the fields of image segmentation and computer vision. Originally utilized for binary classification problems, its application has expanded to evaluate the accuracy of object detection and segmentation tasks, particularly in the context of remote sensing and point cloud data (Jaccard, 1912).

$$IoU = \frac{\cap(A,B)}{\cup(A,B)} \qquad \textbf{Eqn. 2}$$

Where:
A represents the predicted segmentation and B is the ground truth. The metric provides a ratio of the overlapping area between the predicted and actual segmentations to the total area covered by both segmentations combined (Everingham et al., 2009).

Dice Similarity Coefficient

The Dice Similarity Coefficient (DSC), also known as the Sørensen–Dice coefficient, has its origins in ecology, where it was used to measure the similarity between two samples. Over

time, its application has permeated into the fields of image analysis and computer vision, particularly for binary and multi-class segmentation tasks. The DSC provides a robust measure of the spatial overlap between two sets, making it especially valuable for evaluating the accuracy of segmentation algorithms (Dice, 1945).

$$DSC = \frac{2\cap(A,B)}{|A|+|B|} \qquad \textbf{\textit{Eqn. 3}}$$

Where:
A represents the predicted segmentation and B is the ground truth. The coefficient yields a value between 0 (no overlap) and 1 (perfect overlap), offering a normalized measure of the similarity between two segmentations (Zou et al., 2004).

Mean Squared Error

The Mean Squared Error (MSE) is a widely recognized metric in the fields of statistics and signal processing. It quantifies the average squared differences between estimated and actual values, providing a measure of estimation accuracy. In image analysis, MSE has been extensively used to assess the quality of reconstructed or compressed images in comparison to their original versions (Gonzalez, Woods and Masters, 2009).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(A_i - B_i)^2 \qquad \textbf{\textit{Eqn. 4}}$$

Where:
A_i and B_i are the actual and estimated values, respectively, and n is the number of data points. The MSE provides a non-negative value, with a lower MSE indicating a closer fit to the actual data (Kubat, 1999).

Normalized Hausdorff Distance

The Hausdorff Distance (HD) is a metric that gauges the extent to which each point of a set lies near some point of another set and vice versa. Initially defined for subsets of a metric space, it has found widespread applications in image processing, computer vision, and shape matching. Particularly in medical imaging, the Hausdorff Distance is a popular metric for evaluating the accuracy of boundary-based segmentation methods, since it gives a measure of the greatest of all the closest distances between any two points in two sets (Huttenlocher, Klanderman and Rucklidge, 1993).

$$h(A,B) = Max_{a\in A}Max_{b\in B}d(a,b) \qquad \textbf{\textit{Eqn. 5}}$$

$$HD(A,B) = Max\big(h(A,B), h(B,A)\big) \qquad \textbf{\textit{Eqn. 6}}$$

Where:
A and B are non-empty subsets of a metric space, and $d(a, b)$ denotes the usual distance between the points a and b. Essentially, $h(A, B)$ calculates the furthest distance from a point in set A to the closest point in set B. The Hausdorff Distance itself is the maximum of these distances when considering both directions (from A to B and from B to A). The smaller the Hausdorff Distance, the closer the two sets are to each other in terms of shape and location (Dubuisson and Jain, 1994).

The Normalized Hausdorff Distance (NHD) is a variant of the traditional Hausdorff Distance (HD) and is particularly beneficial when comparing shapes or sets of different sizes or when operating in spaces where the absolute distances are less meaningful without normalization. By incorporating normalization, the measure becomes scale-invariant, making it a suitable metric for many computer vision and image processing tasks, especially where the objects of interest might vary in size or the scale of the image is not consistent.
The Normalized Hausdorff Distance can be defined as:

$$D(X) = Max_{x,x' \in X} d(x, x') \qquad \textbf{\textit{Eqn. 7}}$$

$$NHD(A, B) = \frac{HD(A,B)}{Max(D(A),D(B))} \qquad \textbf{\textit{Eqn. 8}}$$

Where:
$D(X)$ denotes the maximum intra-set distance for set X, which is essentially the diameter of the set. By dividing the Hausdorff Distance by the maximum of the diameters of the two sets being compared, we obtain a measure that varies between 0 (for perfectly matching sets) and 1 (for sets that are completely disjoint relative to their size).
The normalization ensures that the value remains scale-invariant and provides a more consistent understanding of the relative distance between two sets across different scenarios or datasets (Huttenlocher, Klanderman and Rucklidge, 1993).

Remove noise in datasets

Z-test was applied for noise removal in the dataset. Firstly, the mean and standard deviation were calculated as these metrics will be crucial for the Z-test calculations. Next, proceed to the actual Z-test where the Z-score for each data point x_i in the dataset is calculated using the formula:

$$Z = \frac{x_i - \mu}{\sigma} \qquad \textbf{\textit{Eqn. 9}}$$

Where:
μ is the mean and σ is the standard deviation
In this study, threshold used was 3. All x_i with an absolute Z-score over 3 will be removed.

Machine Learning Predictions and Assessment

Given the time and labor-intensive nature of manual segmentation, as well as the inherent variability introduced by different operators, there emerges a need for a more systematic and reproducible approach. In light of these challenges, an innovative solution was sought: to predict the accuracy of the point clouds using the tree model parameters derived from the TLS2trees segmentation.

Within the QSM parameters, there exist instances of inconsistent or anomalous data. These anomalies, either due to individual parameter values being out of realistic bounds or due to improbable relationships between multiple parameters, could potentially indicate the low accuracy of the associated tree point cloud. For humans, identifying such inconsistencies, especially given the complex relationships between different QSM parameters, is an intricate endeavor. Machine learning, however, with its adeptness at deciphering intricate patterns and relationships in vast data sets, presents a viable solution to this challenge.

In the endeavour to predict accuracy metrics based on QSM parameters, a foray into the domain of machine learning was undertaken, with a focus on multiple models celebrated for their efficacy in regression challenges. Given the intricate relationships inherent in the QSM data, a diverse array of algorithms was selected: the Random Forest model, the Support Vector Machines (SVM), and the Multilayer Perceptron. Beyond these standalone strategies, their strengths were amalgamated into a holistic ensemble model, aiming to encapsulate the nuanced intricacies of the dataset (Rodriguez-Galiano et al., 2015). In the ensuing sections, an exploration into each of these methods will be provided, shedding light on their distinct features and the accompanying optimization methodologies. All codes are available at: https://github.com/ucfaune/GEOG0105.git

Dataset preparation

The dataset was first loaded into a Pandas DataFrame. The data contained multiple features and targets. The features excluded columns such as 'sample', 'PA', 'IoU', 'DSC', 'MSE', 'NHD', which appeared to be either identifiers or target variables. The target variables were 'PA', 'IoU', 'DSC', 'MSE', and 'NHD'. Then, The dataset was split into a training set and a testing set using a 70-30 ratio. Specifically, 70% of the data was used for training, and the remaining 30% was used for testing. The data was randomly shuffled using a fixed random state to ensure reproducibility.

Random Forest model

The Random Forest algorithm stands as a robust ensemble learning method that constructs multiple decision trees for training (Sagi and Rokach, 2018). By outputting the mode of the classes for classification tasks or the mean prediction for regression tasks, it aims to improve

the model's performance. One of its defining features is the use of bagging, which employs bootstrapped datasets to train each individual tree (Gislason, Benediktsson and Sveinsson, 2006). This not only enhances the model's variance but also makes it more robust. Moreover, Random Forest incorporates feature randomness at each split in the decision tree, considering only a random subset of the features. This approach prevents overfitting, adding to the model's robustness (Ho, 1998).

Beyond these inherent features, Random Forest offers high interpretability. Although it may not be as straightforward as individual decision trees, it does provide valuable insights into feature importance (Bastani, Kim and Bastani, 2018). Furthermore, the model is exceptionally versatile, capable of being used for both classification and regression tasks. It is highly adaptive to different kinds of data, managing missing values effectively and offering robustness even when dealing with unbalanced datasets (Cutler, Cutler and Stevens, 2012). Another advantage is its parallelizability; the individual trees are independent of one another, making it easier to parallelize the training process (Wright and Ziegler, 2017).

In this study, scikit-learn library was used to generate the random forest model. Before training the model, a grid search was conducted to determine the optimal hyperparameters for the Random Forest Regressor, considering a range of values for the number of estimators (n_estimators), maximum number of features (max_features), and maximum depth (max_depth). A 5-fold cross-validation was employed during this grid search, using the negative mean squared error as the evaluation metric. After identifying the optimal hyperparameters, a Random Forest Regressor model was trained on the training dataset using all available CPU cores for computational efficiency (n_jobs = -1). Subsequently, another round of 5-fold cross-validation was executed on the training data to estimate the model's performance. Finally, the trained model was applied to the test set to make predictions for each target variable ('PA', 'IoU', 'DSC', 'MSE', 'NHD'), and its performance was evaluated using the mean squared error (MSE) metric. Scatter plots comparing the actual and predicted values were also created for visual assessment of the model's efficacy.

Support Vector Machine

The Support Vector Machine (SVM) algorithm is a supervised learning technique with widespread applications in both classification and regression tasks (Wang, 2005). It works by constructing hyperplanes in a multi-dimensional space to differentiate between different classes or to predict a continuous outcome (Wan Ling Wong et al., 2013). One of its key attributes is the utilization of the kernel trick. This technique enables SVM to tackle non-linear problems by implicitly mapping input features into a higher-dimensional space, allowing for more complex decision boundaries without directly computing the transformation (Park, 2008).

In terms of regularization, SVM offers a robust framework to avoid overfitting. By optimizing a regularization parameter, SVM balances the trade-off between margin maximization and

classification error. This feature makes the algorithm particularly effective in high-dimensional spaces where overfitting is a common concern (Cao and Tay, 2003). Though it may lack the interpretability found in tree-based models like Decision Trees or Random Forests, SVMs often excel in terms of accuracy, particularly in situations where the classes are not easily separable. One of the advantages of using SVM is its effectiveness in handling both imbalanced and small sample size datasets (Blagus and Lusa, 2013). However, a point of consideration is its computational intensity, particularly when dealing with large datasets or employing non-linear kernels, although various optimization methods can mitigate this to some extent (Tsang, Kwok and Hk, 2005). Despite these challenges, the SVM algorithm remains a strong choice for many machine learning applications, owing to its flexibility, robustness, and ability to produce high-accuracy models.

In this study, the scikit-learn library was utilized to develop a Support Vector Machine (SVM) model wrapped in a MultiOutputRegressor for multivariate regression tasks. Initially, a grid search was performed to pinpoint the best hyperparameters for the underlying SVM model. This involved evaluating different regularization parameter C through a 5-fold cross-validation scheme. The performance metric used during grid search was negative mean squared error. Upon acquiring the optimal hyperparameters, an SVM model was trained on the training dataset. Unlike the Random Forest Regressor, the SVM algorithm does not offer parallel computation via the n_{jobs} parameter; therefore, all computations were performed on a single core. To gauge the model's generalized performance, the training dataset was subjected to another round of 5-fold cross-validation. The trained SVM model was subsequently employed to make predictions on the test set across multiple target variables ('PA', 'IoU', 'DSC', 'MSE', 'NHD'). Model efficacy was assessed quantitatively using the mean squared error (MSE) for each output and qualitatively through scatter plots that compared the actual and predicted values for each target variable. These plots served as a visual metric for evaluating the model's predictive accuracy.

Multilayer Perceptron

The MultiLayer Perceptron (MLP) is a popular neural network architecture commonly used in supervised learning scenarios such as classification and regression (Murtagh, 1991). Structurally, an MLP is composed of an input layer, one or more hidden layers, and an output layer, each containing interconnected neurons. One of the key features of MLP is its ability to model complex, non-linear relationships in data through the application of activation functions like ReLU, sigmoid, or tanh at each neuron (Nwankpa et al., 2018). This capability allows MLPs to excel in tasks that are not easily solvable by linear models, making them highly versatile in handling a range of machine learning problems.

Despite their flexibility and high accuracy, MLPs have some challenges to consider. They are computationally expensive and can be sensitive to the choice of hyperparameters, requiring careful tuning especially when dealing with high-dimensional data (Tekaslan and Melike Nikbay, 2022). Regularization techniques such as dropout and L2 regularization can help in

mitigating overfitting, but unlike more interpretable models like Decision Trees, MLPs are often considered "black boxes," making them less interpretable (Zhang et al., 2021). Nevertheless, their robustness and adaptability make them a widely used tool in diverse domains, from natural language processing to image recognition.

In this study, the PyTorch library was utilized to construct a MLP model for regression tasks. Prior to training, a grid search was performed to identify the optimal hyperparameters for the MLP, taking into account different learning rates (0.01, 0.001) and hidden layer dimensions (16, 32). A 5-fold cross-validation was employed during this grid search, with the Mean Squared Error (MSE) serving as the evaluation metric. After pinpointing the optimal hyperparameters, the MLP model was trained on the entire training set for 50 epochs. The model architecture comprised an input layer, one hidden layer with batch normalization and ReLU activation, and an output layer with sigmoid activation. The Adam optimizer was used for training, and the model was deployed on a GPU for computational efficiency. Subsequently, another round of 5-fold cross-validation was executed on the training dataset to estimate the model's performance. Finally, the trained model was applied to the test set to make predictions for each of the target variables ('PA', 'IoU', 'DSC', 'MSE', 'NHD'). The model's performance was evaluated using the Mean Squared Error (MSE) metric. Scatter plots were also generated to visually compare the actual and predicted values for each target variable, offering further insight into the model's efficacy.

RESULTS

Accuracy Assessment Metrics

In total, we collected 466 samples for our study. After eliminating noise and outliers, we were left with a refined dataset of 417 samples.

Table 1 Statistical properties of the metrics for the whole dataset

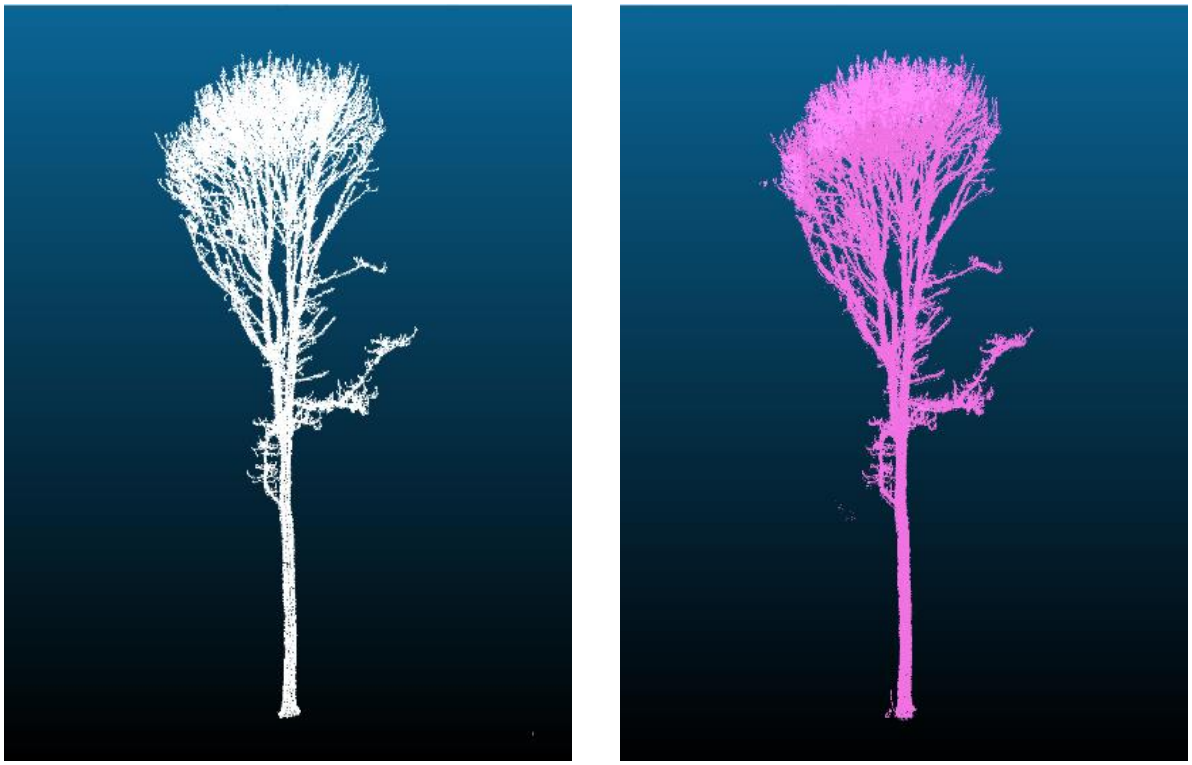
	PA	IoU	DSC	MSE	NHD
Maximum	1	1	1	0.371882	0.255464
Minimum	0.628118	0.302509	0.462065	0	0.001925
Median	0.895833	0.816764	0.89867	0.104167	0.03637
Average	0.880321	0.785773	0.870508	0.119679	0.056561
Standard Deviation	0.08292	0.146108	0.101199	0.08292	0.052529

The comprehensive details of these samples are provided in the accompanying table. Remarkably, the tree labeled as T321 stood out among all the samples. It exhibited a perfect score across multiple metrics: with a PA (Percentage Agreement), IoU (Intersection over Union), and DSC (Dice Similarity Coefficient) all equal to 1. Additionally, its MSE (Mean Squared Error) was observed to be zero, making it a unique specimen in our dataset.

To gain further insights, we randomly selected a subset of samples for a more in-depth analysis. The metrics for these samples are listed in the table below. Additionally, we utilized the CloudCompare software to visualize the point cloud images of these selected samples. The visualization results are displayed in the following figures.

Table 2 Metrics for specific tree samples

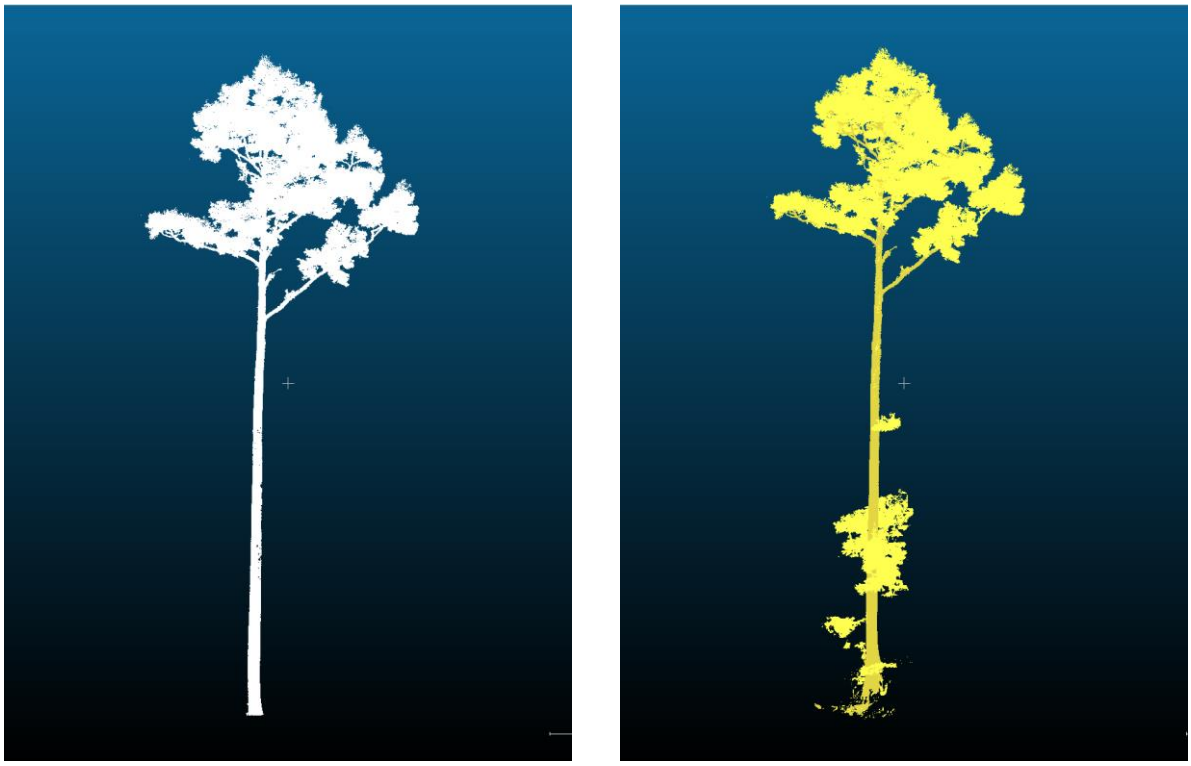
	PA	IoU	DSC	MSE	NHD
T321	1	1	1	0	0.027410
T1	0.944356	0.833276	0.905439	0.055644	0.016850
T38	0.891636	0.761324	0.860539	0.108364	0.072228
T417	0.952102	0.0.898263	0.946373	0.047898	0.002186



(a)Manually segmented T321 (b)TLS2trees segmented T321

Figure 1 Front view visualization of T321

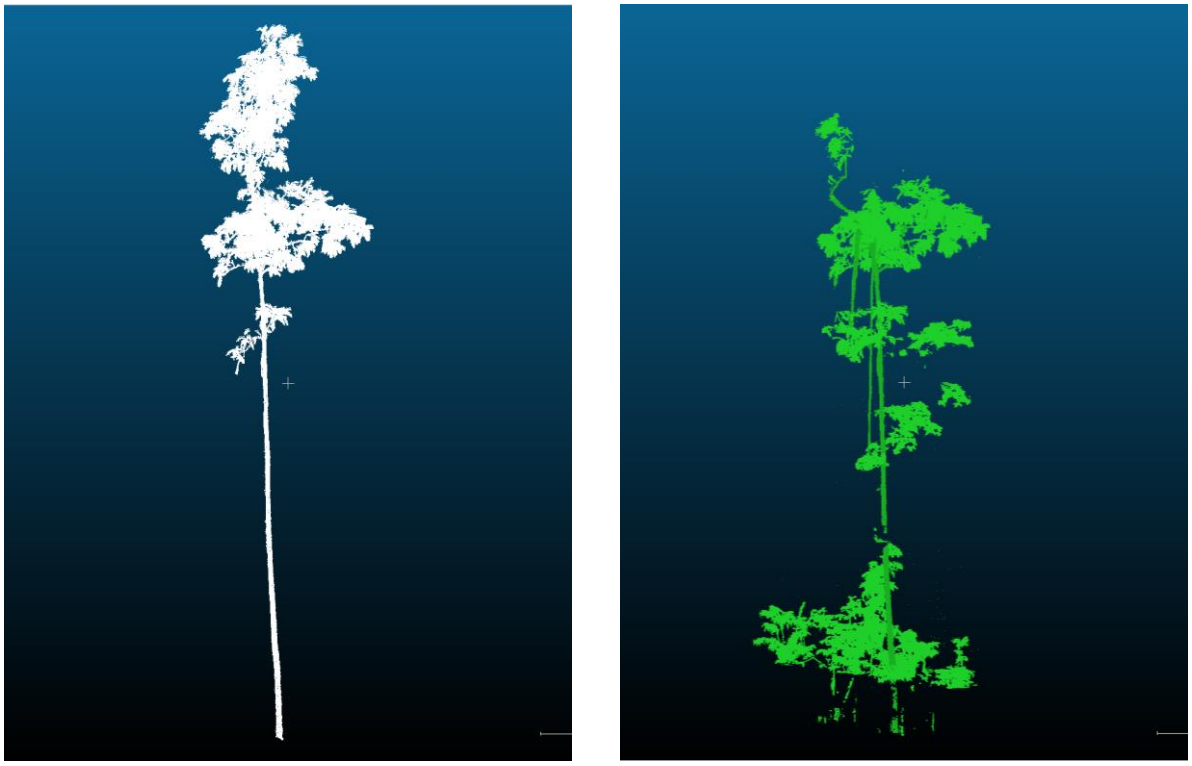
As can be seen, the two figures are almost identical, with only minor differences observed at the base of the trunk.



(a)Manually segmented T1 (b)TLS2trees segmented T1

Figure 2 Front view visualization of T1

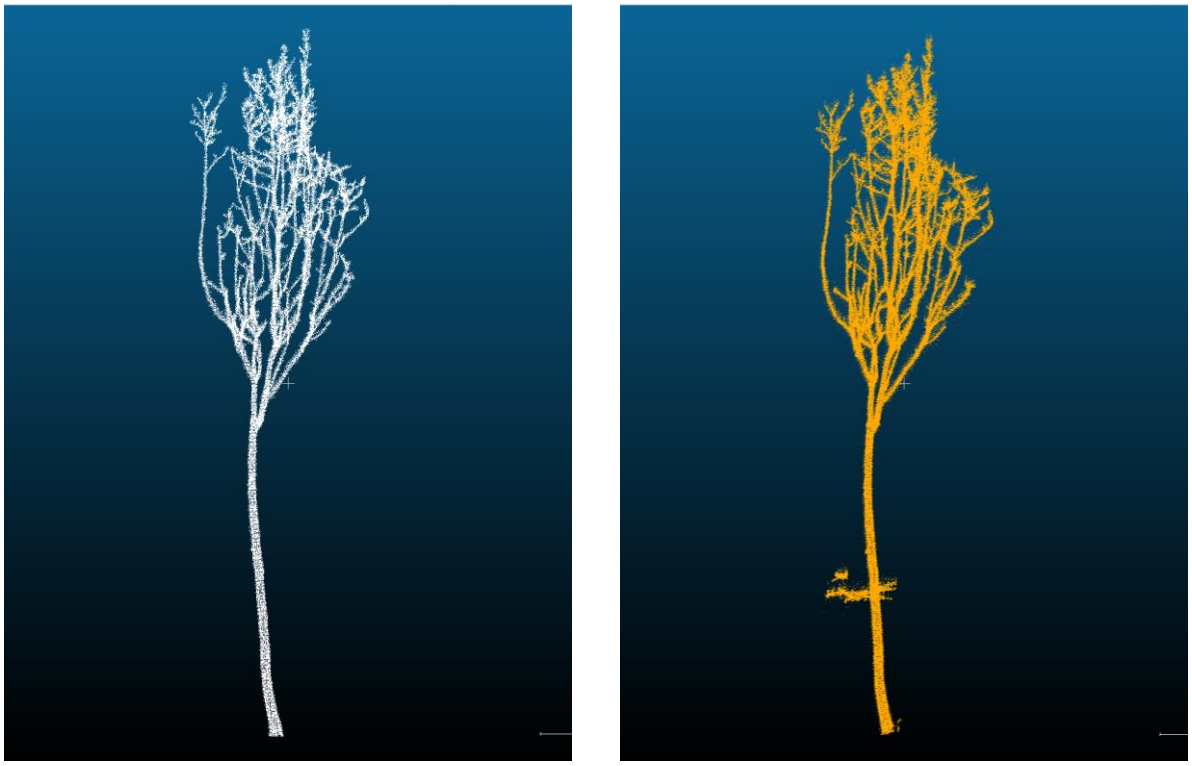
The similarity between the two trees is strikingly high, making it evident that they represent the same tree. The upper halves are virtually identical. However, in the case of the TLS2trees processed tree, the lower portion exhibits some leaf-like structures that are likely to belong to other trees.



(a)Manually segmented T38 (b)TLS2trees segmented T38

Figure 3 Front view visualization of T38

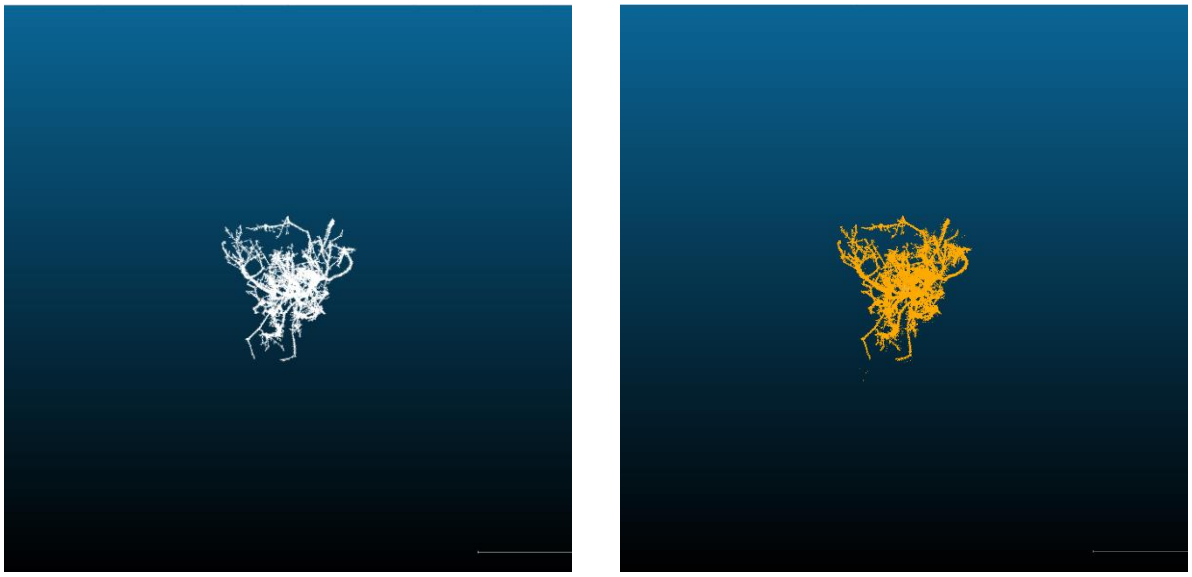
Upon initial inspection, it is challenging to ascertain that the two figures represent the same tree. But a closer look at the lower canopy (the middle section of the image) reveals that they do indeed correspond to the same tree. When using the manually segmented point cloud as the benchmark, the TLS2trees image shows notable discrepancies. Specifically, it lacks the upper canopy that is present in the standard image, and the bottom section includes numerous parts that actually belong to other trees. Additionally, parallel structures resembling trunks appear adjacent to the main trunk in the TLS2trees image.



(a)Manually segmented T417 (b)TLS2trees segmented T417

Figure 4 Front view visualization of T38

The two images are nearly identical; however, when using the manually segmented point cloud as a standard for comparison, the image from TLS2trees contains an extra element in the middle of the trunk that does not belong to this specific tree.



(a)Manually segmented T417 (b)TLS2trees segmented T417

Figure 5 Top view visualization of T38

When viewed from the top, minor discrepancies can also be observed between the two images; the lines in the TLS2trees image appear to be coarser.

Random Forest Model

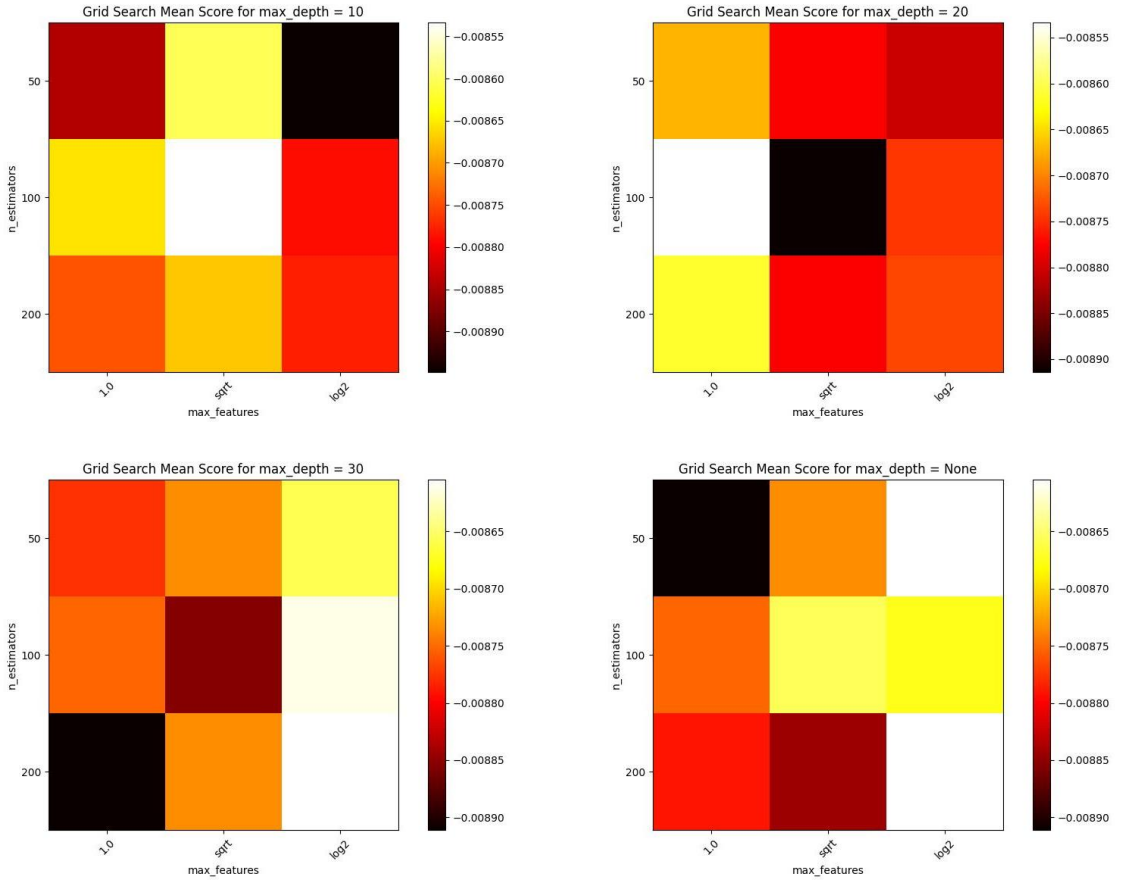


Figure 6 Hyperparameter tuning performance for random forest model

We employed GridSearchCV to fine-tune the hyperparameters of the Random Forest model, specifically focusing on the number of estimators (n_estimators), the maximum number of features (max_features), and the maximum depth (max_depth). Four figures are presented to show the model's performance at varying maximum depths of 0, 10, 20, and 30. In these figures, each cell corresponds to a unique combination of hyperparameters. The horizontal axis represents the maximum number of features, categorized as 'auto,' 'sqrt,' and 'log2.' The vertical axis indicates the number of estimators, listed as 50, 100, and 200. The color gradient, ranging from dark to light, signifies an increase in model performance; black indicates the poorest performance while white represents the best.

Based on the color bars, it is apparent that the performance variations across different combinations of hyperparameters are relatively minor. The best-performing set of hyperparameters was found to be (max_depth: 10, max_features: 'sqrt', n_estimators: 200).

The following are the figures indicating the performance of Random Forest Model (RF), in terms of predicting the accuracy of TLS2trees segmented trees based on their tree model data.

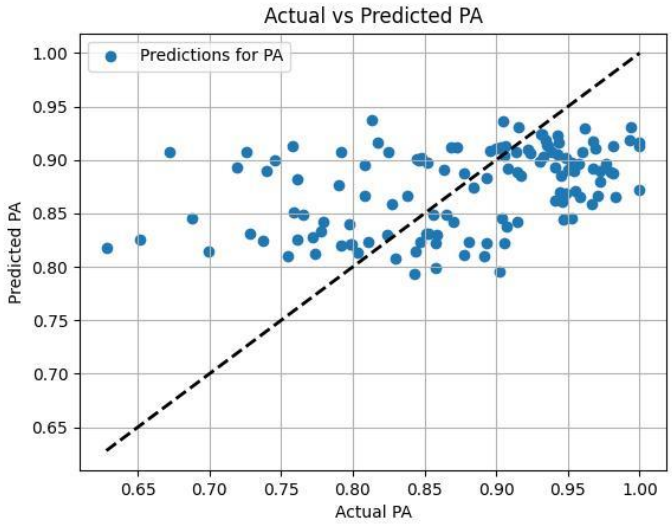


Figure 7 PA prediction by random forest

For the metric of Pixel Accuracy (PA), the Mean Squared Error is measured at 0.01. The dashed line in the corresponding graph represents an ideal scenario where predicted and actual PA would match perfectly. Unfortunately, the sparse distribution of points along this line suggests a lack of predictive accuracy. Points largely populate the y-axis range between 0.8 and 0.95, while stretching along the x-axis from about 0.65 to 1. This indicates that while the model can predict a general range for PA, it fails to closely align with the tree's model characteristics.

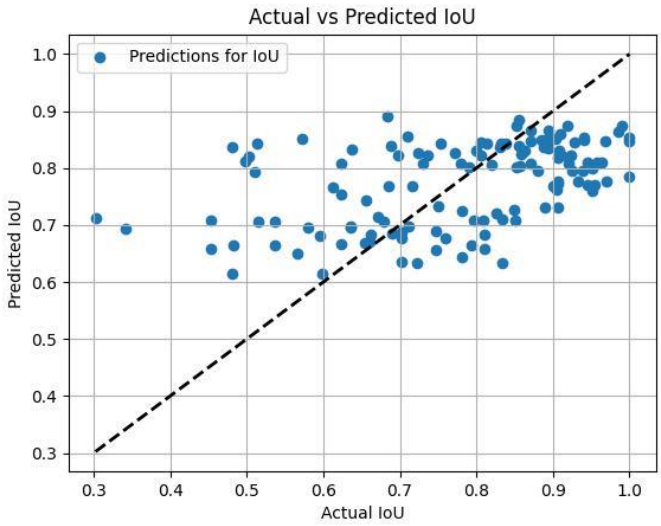


Figure 8 IoU prediction by random forest

In the case of Intersection over Union (IoU), the Mean Squared Error mirrors that of PA but comes in slightly higher at 0.02. Much like the prior metric, the sparseness of points on the ideal dashed line highlights the model's suboptimal predictive power. Most points on the graph hover on the y-axis between 0.6 and 0.9, but their x-axis distribution is wider, ranging from approximately 0.3 to 1. This underlines the model's weak performance and poor association with the tree attributes.

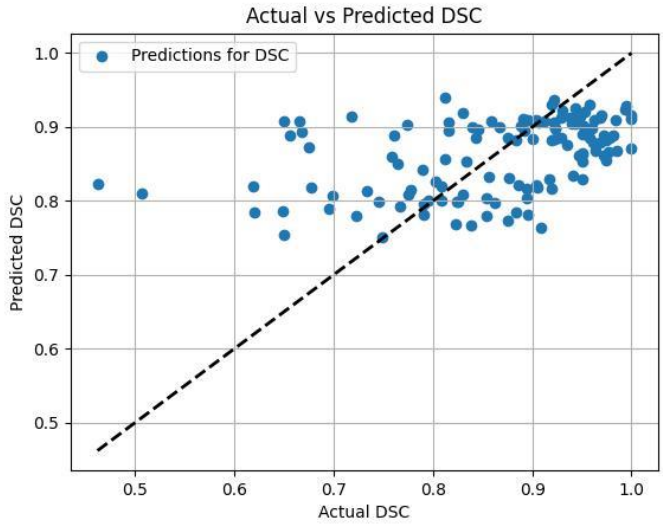


Figure 9 DSC prediction by random forest

Turning our attention to the Dice Similarity Coefficient (DSC), we observe a Mean Squared Error of 0.01. A comparable trend emerges: points scarcely align with the ideal dashed line, signalling low predictive accuracy. Here, the y-axis values are predominantly sandwiched between 0.75 and 0.95, while their x-axis counterparts vary between around 0.5 to 1. Again, this reveals that the model is only capable of generating approximate predictions that show minimal correlation with the properties of the trees.

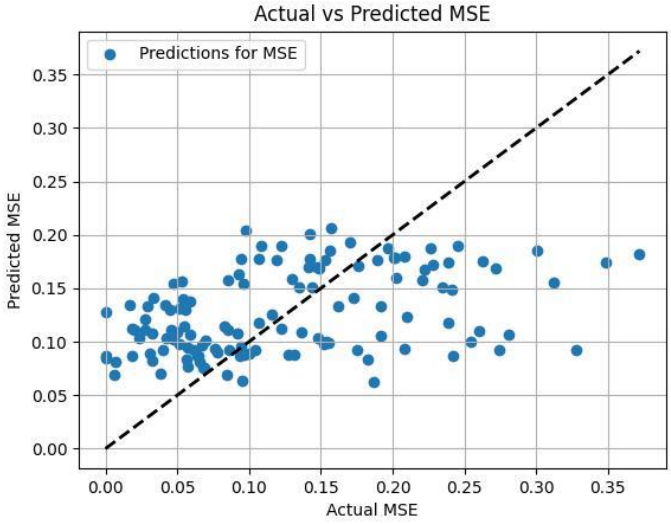


Figure 10 MSE prediction by random forest

When assessing Mean Squared Error (MSE) as a metric, the Mean Squared Error value of model is 0.01. The graph also showcases a shortage of points along the optimal dashed line. The points primarily cluster in the y-axis zone of 0.05 to 0.2 and span the x-axis from close to 0 up to 0.35. This illustrates that, much like the other metrics, the model can sketch a rough estimate but misses the mark in tying predictions closely to tree attributes.

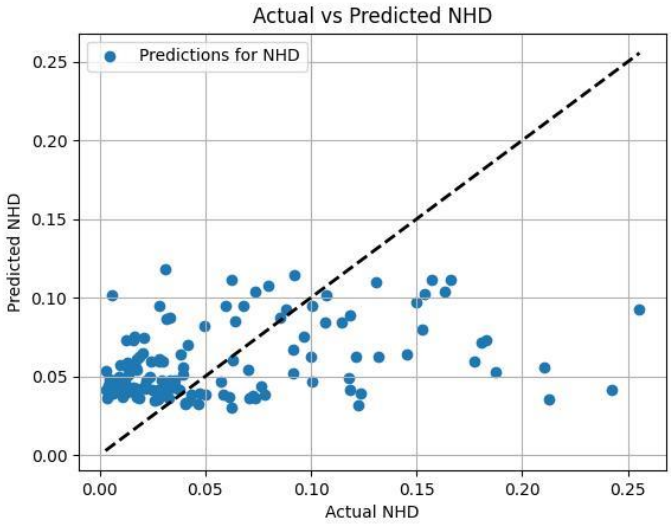


Figure 11 NHD prediction by random forest

Similarly, for the metric of Normalized Hausdorff Distance (NHD), the sparsity of points on the dashed line persists, indicating a limited predictive capability. Here, the y-values are confined between 0.04 and 0.1, while the x-values are sprinkled from roughly 0 to 0.25. This suggests that although the model can identify a generic range, it struggles to align its predictions with the innate attributes of the trees.

Support Vector Machine

The following are the figures indicating the performance of Support Vector Machine (SVM), in terms of predicting the accuracy of TLS2trees segmented trees based on their tree model data.

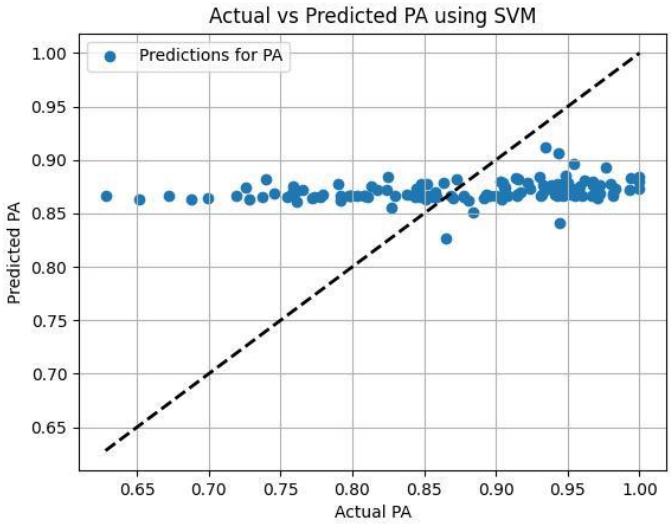


Figure 12 PA prediction by SVM

Similarly, for Pixel Accuracy (PA), we find a model test MSE of 0.01. The graph features a

dashed line to symbolize a perfect match between predicted and actual PA values. Regrettably, the actual data points rarely align with this ideal representation. Along the y-axis, points tend to cluster between 0.85 and 0.9, whereas their x-axis counterparts span from about 0.65 to 1. This suggests that the model has a knack for defining a general range but doesn't do well in fine-tuning the prediction to match the attributes of individual trees.

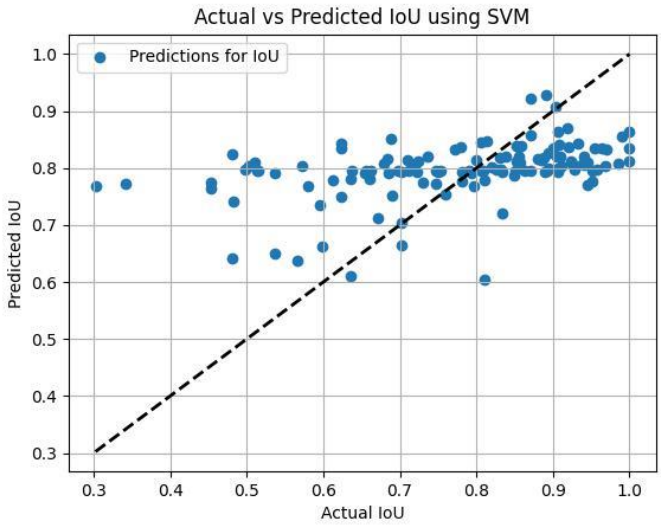


Figure 13 IoU prediction by SVM

Next, we analyze Intersection over Union (IoU), revealing an MSE of 0.02. Again, the graph's dashed line represents an ideal match between prediction and reality. The sparseness of points along this line serves as a glaring indicator of the model's imprecision. Most data points are positioned between 0.6 and 0.9 on the y-axis, and between 0.3 and 1 on the x-axis, once more indicating a lackluster correlation with the actual tree properties.

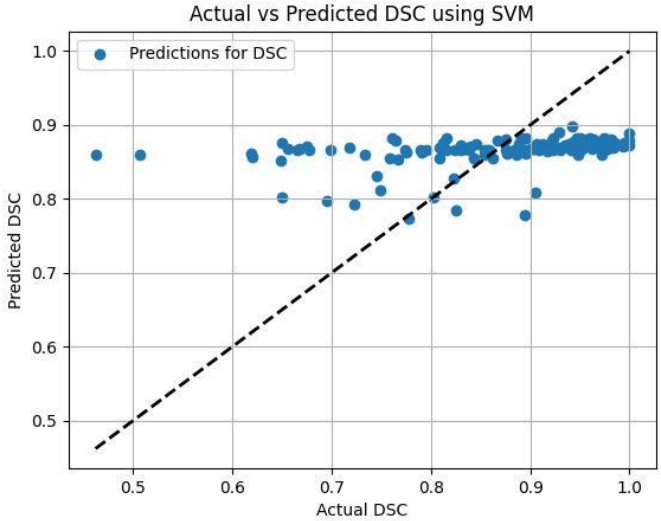


Figure 14 DSC prediction by SVM

For the Dice Similarity Coefficient (DSC), the trend persists, and the MSE measures at 0.01. The scarcity of points along the dashed 'ideal match' line continues to flag concerns about predictive accuracy. On the y-axis, values tend to hover between 0.8 and 0.9, and along the x-axis, they extend from around 0.5 to 1. While the model seems capable of generating ballpark figures, it doesn't successfully capture the nuances inherent to individual trees.

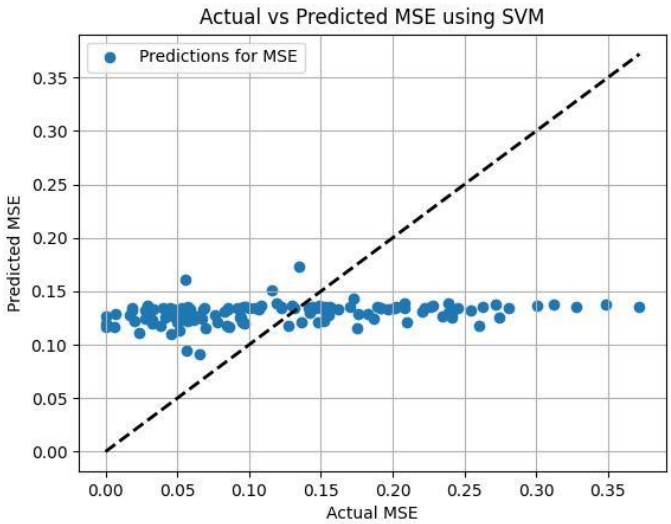


Figure 15 MSE prediction by SVM

Focusing on Mean Squared Error (MSE), we observe an model test MSE value of 0.01. Points again are sparse along the dashed line representing perfect prediction. Most points are confined to the y-axis range of 0.10 to 0.15 and stretch along the x-axis from nearly 0 to 0.35. This points to a model capable of rough estimations but failing to correlate these with specific tree characteristics.

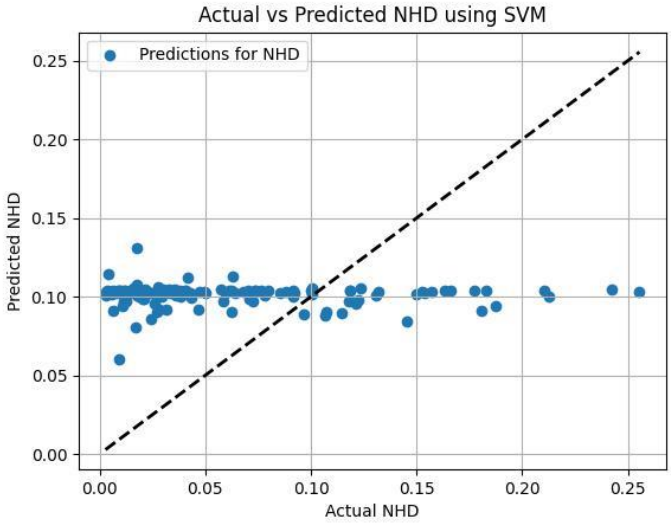


Figure 16 NHD prediction by SVM

Finally, for Normalized Hausdorff Distance (NHD), the story is much the same. Data points scarcely fall along the dashed line representing ideal predictions, signifying limited accuracy. On the y-axis, they mainly fall between 0.07 and 0.1, while their x-axis distribution ranges from about 0 to 0.25. This suggests that the model does manage to identify a general predictive range, yet falls short in aligning closely with the tree's actual properties.

Multilayer Perceptron

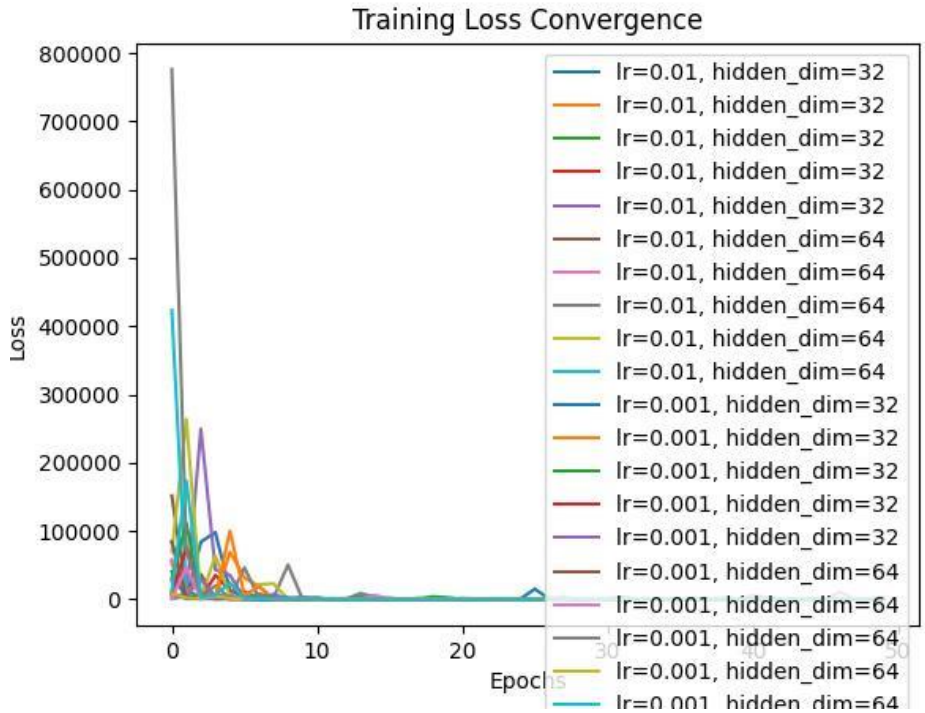


Figure 17 Convergence for the Multi-Layer Perceptron with different hyperparameter

The convergence behaviour of different groups during hyperparameter tuning for the Multi-Layer Perceptron (MLP) is observed in the figure above. Within fewer than ten iterations, rapid convergence is achieved by nearly all groups.

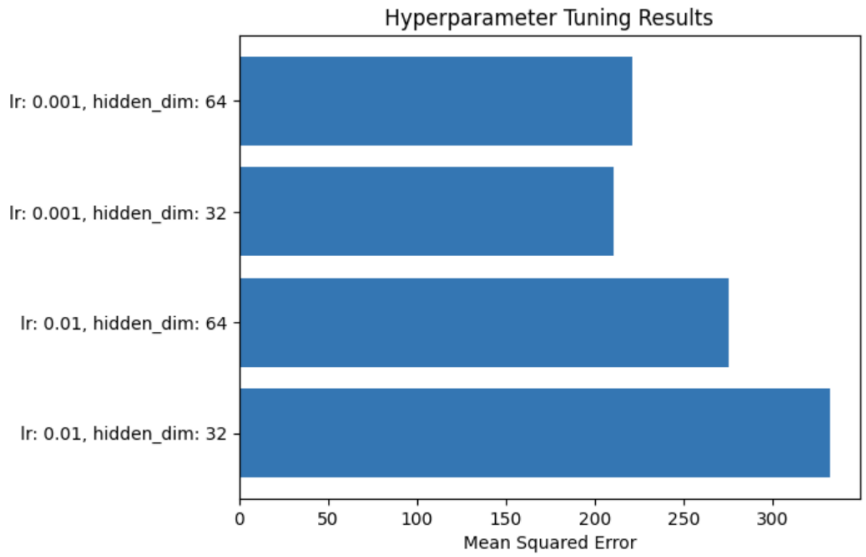


Figure 18 Hyperparameter Tuning Results

In this bar chart, four well-performing groups are identified, each representing a different combination of neural unit count and learning rate for the MLP's performance. Among these, the group with a learning rate of 0.001 and 32 neural units achieved the lowest mean squared error, standing at 210.81. Therefore, the MLP performs best with this specific set of hyperparameters.

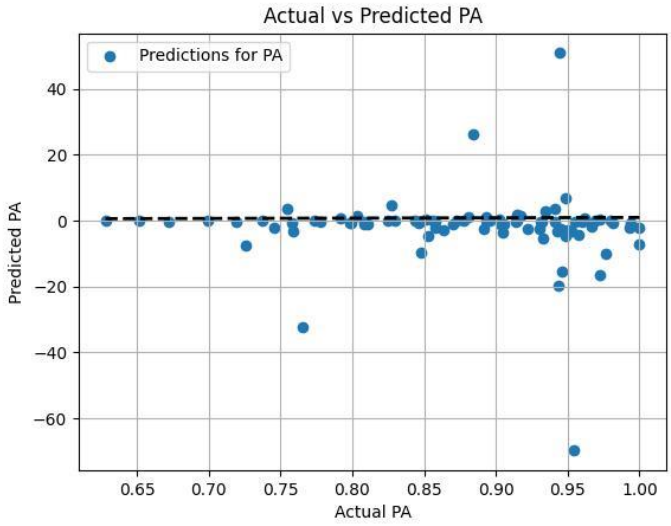


Figure 19 PA prediction by MLP

The dashed line in the graph signifies an ideal match between predicted and actual PA values. While most data points cluster around this line, indicating the model can approximate within a reasonable range, some outliers exceed the theoretical PA bounds of 0-1.

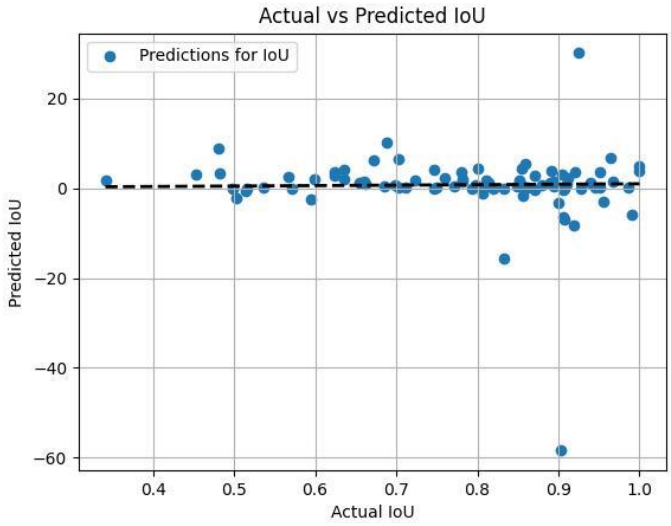


Figure 20 IoU prediction by MLP

The chart's dashed line serves as an equality benchmark between forecasted and actual IoU values. Although the majority of points hover near this line, suggesting the model offers a reasonable approximation, there are exceptions that surpass IoU's theoretical limits of 0-1.

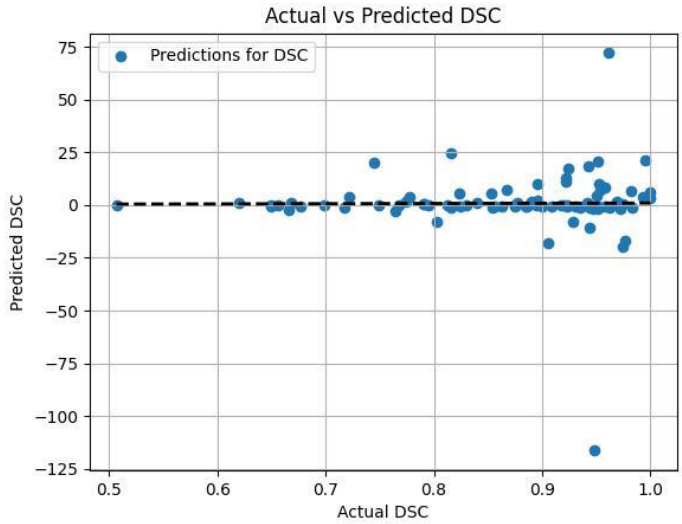


Figure 21 DSC prediction by MLP

In the graph, the dashed line denotes a perfect alignment between estimated and actual DSC. Many of the points appear near this line, providing evidence that the model's predictions are generally accurate. However, certain anomalies stretch beyond the 0-1 theoretical scope for DSC.

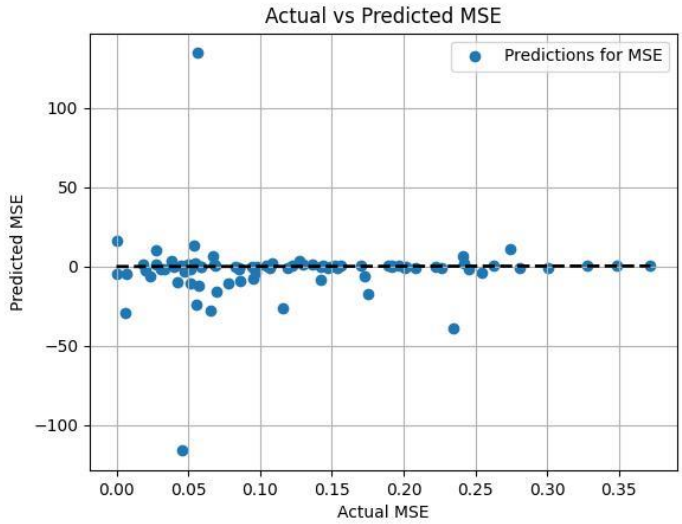


Figure 22 MSE prediction by MLP

The graph's dashed line marks where predicted MSE perfectly matches actual MSE. Most points are situated close to this line, which implies the model can generate a fairly accurate predictive range.

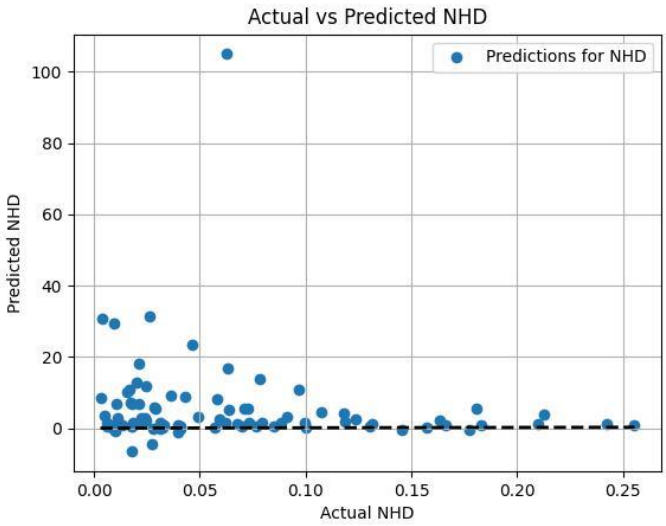


Figure 23 NHD prediction by MLP

The dashed line on the chart equates predicted NHD values with their actual counterparts. While a large number of points are adjacent to this line, indicating the model's predictive accuracy, some points anomalously go beyond NHD's theoretical 0-1 range.

DISCUSSION

Accuracy Assessment Metrics

In this study, a total of 466 samples were collected and subsequently refined to 417 samples after the elimination of noise and outliers. The statistical metrics presented indicate a fairly robust dataset with good values for PA (Percentage Agreement), IoU (Intersection over Union), DSC (Dice Similarity Coefficient), and relatively low values for MSE (Mean Squared Error) and NHD (Normalized Hausdorff Distance).

The tree labeled as T321 serves as a standout example, exhibiting perfect scores across multiple metrics. This suggests that for this particular sample, the techniques employed for capturing and processing the data yielded results of exceptional quality. However, it's crucial to recognize that this sample might be an outlier and should not overshadow the performance or issues present in the remaining dataset.

The CloudCompare visualizations add another layer to our analysis. The high degree of similarity between different trees suggests that the metrics are effective at capturing the salient features of each sample. However, the TLS2trees processed images do introduce some discrepancies, especially in the lower canopy and trunk structures. These inconsistencies could arise from various factors such as segmentation errors, or they could represent natural variations that the metrics failed to capture.
The metrics employed in this study serve as essential tools for evaluating the performance of the TLS2trees software in tree segmentation tasks. They offer a multidimensional framework to scrutinize how well the software aligns with manual or ground truth segmentation. By comparing the segmented point cloud images alongside these metrics, it becomes evident that these numerical values are not mere abstractions but indeed capture the quality of tree segmentation.

However, the study did face limitations, particularly in the form of MSE and NHD values, which had higher deviations. It might be beneficial to further refine the selection criteria for noise and outlier removal. Additionally, a more in-depth look into why certain samples like T321 outperform others could provide valuable insights into optimizing data collection and preprocessing steps.

Machine Learning Prediction

In this study, we employed various machine learning models to predict the accuracy of segmented trees generated by the TLS2trees software. Each model was tuned and evaluated across multiple metrics, including Pixel Accuracy (PA), Intersection over Union (IoU), Dice

Similarity Coefficient (DSC), Mean Squared Error (MSE), and Normalized Hausdorff Distance (NHD). This rich array of metrics provides a comprehensive assessment of the software's tree segmentation capabilities, supplementing our visual comparisons.

Random Forest showed consistent performance across various hyperparameter combinations, which is indicative of the model's robustness. However, its inability to closely approximate real values of PA, IoU, and DSC raises concerns. As pointed out by (Breiman, 2001), Random Forest is generally good for complex classification problems, but our results suggest it may not be ideal for this specific predictive task. On the other hand, although GridSearchCV was used for hyperparameter tuning in Random Forest, further fine-tuning might yield better results. Some hyperparameters like min_samples_split or min_samples_leaf were not investigated. Additionally, other optimization techniques such as Randomized Search CV or Bayesian Optimization could be explored.

In contrast, Support Vector Machine (SVM) showed similar performance shortfalls, most notably in its prediction of IoU and DSC. SVM was originally developed for two-group classification problems, and perhaps this fundamental architecture is less applicable to the continuous range of metrics we are considering here (Cortes and Vapnik, 1995).

Among all the models, Multi-Layer Perceptron (MLP) exhibited rapid convergence during hyperparameter tuning and showed promising results, especially when it came to predicting PA, IoU, and DSC. The specific architecture of MLP, well-suited for capturing complex relationships in the data as highlighted, seems to offer the most promise in this particular application (Maheswari and Gomathi, 2023).

However, it should be noted that all models struggled to achieve a high level of predictive accuracy across all metrics. This could be due to the inherent complexity and variability in tree structures, making it challenging for any model to offer precise predictions. Considering the complexity of tree structures, more sophisticated models like Convolutional Neural Networks (CNNs) or even Graph Neural Networks (GNNs) could be employed, as they have shown success in capturing complex relationships in other domains (Bronstein et al., 2017).

Future work

In subsequent research, there is room for employing a broader range of metrics to more comprehensively assess the performance of tree segmentation algorithms like TLS2trees, for example, Global Consistency Error, Volumetric Similarity etc. (Taha and Hanbury, 2015).

Moreover, we can delve into more detailed information about each tree model by enhancing semantic segmentation techniques. For instance, points that are part of the tree trunk but do not belong to the tree could be identified as "foreign points." Such an approach would allow us to calculate the ratio of foreign points in different sections of the tree. This would not only refine the existing tree model but also provide nuanced insights into the effectiveness of the segmentation algorithm in isolating individual trees from a point cloud containing multiple

entities.

One of the most pressing limitations of this study is the sparse dataset employed, which fails to capture the complexity inherent in tree models and their corresponding segmentation accuracy. This inadequacy affects the generalizability and robustness of the machine learning models developed. Compounding this issue is the inconsistency in data formats and parameters across various datasets encountered during initial data collection, which complicates the generation of a unified, comprehensive training set. To address these challenges, we strongly recommend that research institutions collaborate to standardize the methodologies for point cloud data collection and processing. Such standardization would simplify the integration of datasets and enable the creation of a more extensive and representative training set. Enhanced resource sharing could further augment this process, facilitating more in-depth and conclusive studies. Focusing on these directions could significantly improve the future landscape of machine learning-based tree segmentation algorithms.

Another potential reason for the suboptimal performance of the machine learning models in this study could be the inherently complex nature of the regression problem at hand. There may be an opportunity to improve the predictive accuracy of the machine learning models by reframing the task as a classification problem instead of a regression problem. However, transitioning to a classification-based approach would require the inclusion of categorical labels in the training set, representing different gradations of tree segmentation quality. These labels would likely need to be generated manually at the outset. One possible method could be to have multiple evaluators assign quality ratings to a given sample and then average those ratings. A sufficiently large dataset generated through this process could then serve as the basis for training a more effective classification model. By redefining the task in terms of classification and harnessing the power of collective human judgment for initial labeling, we may be able to construct machine learning models with significantly improved predictive capabilities. This approach could offer a promising avenue for future research, potentially overcoming some of the limitations inherent in treating the problem as a complex regression task.

Once a model with satisfactory predictive accuracy is developed, integrating it into the TLS2trees system could form a valuable feedback loop for future improvements. The enhanced model could serve as an evaluative tool within TLS2trees, enabling the system to automatically assess the quality of its tree segmentation in real-time. This self-evaluative mechanism could trigger adjustments or retraining of the system based on identified deficiencies or room for improvement. Over time, this feedback loop could help TLS2trees to become increasingly self-reliant and adaptive, learning from its previous iterations and progressively refining its performance. The prospect of a self-optimizing TLS2trees system, enhanced by a machine learning model with high predictive accuracy, offers a promising direction for future research and application. Such an integrated system would not only boost the efficiency and accuracy of tree segmentation but also contribute to the broader aim of automating and improving environmental monitoring technologies.

CONCLUSION

This comprehensive study delved into the performance of tree segmentation algorithms, particularly focusing on TLS2trees. Utilizing a rich array of metrics—Pixel Accuracy (PA), Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Mean Squared Error (MSE), and Normalized Hausdorff Distance (NHD)—the research provided a multi-faceted evaluation of the algorithm's capabilities.

Machine learning models like Random Forest, Support Vector Machine, and Multi-Layer Perceptron were employed to predict the accuracy of tree segmentations. Although these models showed promise in terms of parameter tuning and rapid convergence, they fell short in achieving high predictive accuracy across all metrics. This indicates room for improvement and possibly a shift in methodology, perhaps by employing more sophisticated models like Convolutional Neural Networks or Graph Neural Networks, or by reframing the task as a classification problem rather than a regression task.

The limitations of the study were openly acknowledged, particularly concerning the sparse dataset and the need for standardized data collection procedures. The discussion section offered insightful recommendations for future work, ranging from exploring additional metrics and semantic segmentation techniques to creating a feedback loop for real-time quality assessment in the TLS2trees system.

Ultimately, this study serves as both a rigorous evaluation of current tree segmentation algorithms and a roadmap for future research. It highlights the complexities involved in automating environmental monitoring technologies and sets the stage for collaborative efforts aimed at data standardization and methodological refinement. The findings and recommendations of this study could be instrumental in driving the field of machine learning-based tree segmentation algorithms toward higher levels of accuracy and efficiency, thereby contributing to broader environmental monitoring and conservation efforts.

AUTO-CRITIQUE

When I initially embarked on this research journey in the realm of terrestrial laser scanning (TLS), I had little knowledge or experience in the field. My decision to explore this topic was primarily driven by curiosity and a profound admiration for Prof. Mathias Disney, who later became my supervisor. After diving into numerous scholarly articles and engaging in enriching conversations with Prof. Disney, I began to grasp the inherent complexities and challenges associated with this line of research.

Starting from square one, I had to familiarize myself with not only the theoretical concepts but also the software and tools integral to the study, such as TLS2trees. I soon realized that the learning curve was steep, and time became a pressing concern. The datasets available for this research were vast, and each came in a unique format, demanding substantial time for pre-processing. The sheer volume of the files nearly overwhelmed my computer's storage capacity, and the computational power required for running and debugging the code was not insignificant.

Although I managed to produce results, I acknowledge that they are far from ideal. Due to my initial lack of expertise and the time constraints, the available training set was notably limited in size. This shortcoming impacted the generalizability and robustness of the machine learning models developed, hindering what could have been more promising outcomes. I believe that with a more extensive and diverse dataset, the performance of the models could have been significantly enhanced.

In retrospect, I see room for improvement and potential avenues for future research. Given the opportunity, I intend to revisit this study, incorporating a broader range of datasets to provide a more comprehensive and robust analysis. Despite the limitations and challenges faced, this experience has been an invaluable learning process, offering both professional growth and personal insights into the complexities of terrestrial laser scanning and machine learning methodologies.

LIST OF REFERENCES

Åkerblom, M., Raumonen, P., Mäkipää, R. and Kaasalainen, M. (2017). Automatic tree species recognition with quantitative structure models. *Remote Sensing of Environment*, 191, pp.1–12. doi:https://doi.org/10.1016/j.rse.2016.12.002.

Baltsavias, E.P. (1999). Airborne laser scanning: basic relations and formulas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3), pp.199–214. doi:https://doi.org/10.1016/s0924-2716(99)00015-5.

Bastani, O., Kim, C. and Bastani, H. (2018). Interpretability via Model Extraction. *arXiv:1706.09773 [cs, stat]*. [online] Available at: https://arxiv.org/abs/1706.09773.

Bell, A., Chambers, B. and Butler, H. (2020). *PDAL - Point Data Abstraction Library*. [online] pdal.io. Available at: https://pdal.io/en/latest/ [Accessed 26 Aug. 2023].

Beraldin, J.A., Picard, M., El-Hakim, S.F., Godin, G., Valzano, V. and Bandiera, A. (2005). Combining 3D technologies for cultural heritage interpretation and entertainment. *Proceedings of SPIE*. doi:https://doi.org/10.1117/12.594226.

Besl, P. and H.D. McKay (1992). *A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell*. [online] ResearchGate. Available at: https://www.researchgate.net/publication/3191994_A_method_for_registration_of_3-D_shapes_IEEE_Trans_Pattern_Anal_Mach_Intell.

Beyene, S.M., Hussin, Y.A., Kloosterman, H.E. and Ismail, M.H. (2020). Forest Inventory and Aboveground Biomass Estimation with Terrestrial LiDAR in the Tropical Forest of Malaysia. *Canadian Journal of Remote Sensing*, 46(2), pp.130–145. doi:https://doi.org/10.1080/07038992.2020.1759036.

Blagus, R. and Lusa, L. (2013). SMOTE for high-dimensional class-imbalanced data. *BMC Bioinformatics*, 14(1). doi:https://doi.org/10.1186/1471-2105-14-106.

Boehler, W., Heinz, G. and Marbs, A. (2002). *THE POTENTIAL OF NON-CONTACT CLOSE RANGE LASER SCANNERS FOR CULTURAL HERITAGE RECORDING*. [online] Available at: https://www.isprs.org/PROCEEDINGS/XXXIV/5-C7/pdf/2001-11-wb01.pdf [Accessed 26 Aug. 2023].

Brasington, J., Vericat, D. and Rychkov, I. (2012). Modeling river bed morphology, roughness, and surface sedimentology using high resolution terrestrial laser scanning. *Water Resources Research*, 48(11). doi:https://doi.org/10.1029/2012wr012223.

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), pp.5–32. doi:https://doi.org/10.1023/a:1010933404324.

Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A. and Vandergheynst, P. (2017). Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34(4), pp.18–42. doi:https://doi.org/10.1109/msp.2017.2693418.

Burt, A., Disney, M. and Calders, K. (2018). Extracting individual trees from lidar point clouds using *treeseg*. *Methods in Ecology and Evolution*. doi:https://doi.org/10.1111/2041-210x.13121.

Calders, K., Adams, J., Armston, J., Bartholomeus, H., Bauwens, S., Bentley, L.P., Chave, J., Danson, F.M., Demol, M., Disney, M., Gaulton, R., Moorthy, S.M.K., Levick, S.R., Saarinen, N., Schaaf, C., Stovall, A., Terryn, L., Wilkes, P. and Verbeeck, H. (2020). Terrestrial laser scanning in forest ecology: Expanding the horizon. *Remote Sensing of Environment*, [online] 251, p.112102. doi:https://doi.org/10.1016/j.rse.2020.112102.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J. and Kaasalainen, M. (2014). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. *Methods in Ecology and Evolution*, 6(2), pp.198–208. doi:https://doi.org/10.1111/2041-210x.12301.

Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y. and Lewis, P. (2018). Realistic Forest Stand Reconstruction from Terrestrial

LiDAR for Radiative Transfer Modelling. *Remote Sensing*, 10(6), p.933. doi:https://doi.org/10.3390/rs10060933.

Calders, K., Verbeeck, H., Burt, A., Origo, N., Nightingale, J., Malhi, Y., Wilkes, P., Raumonen, P., Bunce, R.G.H. and Disney, M. (2022). Laser scanning reveals potential underestimation of biomass carbon in temperate forest. *Ecological Solutions and Evidence*, 3(4). doi:https://doi.org/10.1002/2688-8319.12197.

Campana, S. (2017). Drones in Archaeology. State-of-the-art and Future Perspectives. *Archaeological Prospection*, 24(4), pp.275–296. doi:https://doi.org/10.1002/arp.1569.

Cao, L.J. and Tay, F.E.H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. *IEEE Transactions on Neural Networks*, 14(6), pp.1506–1518. doi:https://doi.org/10.1109/tnn.2003.820556.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E. and Pélissier, R. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, 20(10), pp.3177–3190. doi:https://doi.org/10.1111/gcb.12629.

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), pp.35–46. doi:https://doi.org/10.1016/0034-4257(91)90048-b.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, [online] 20(3), pp.273–297. doi:https://doi.org/10.1007/bf00994018.

Côté, J.-F., Fournier, R.A., Frazer, G.W. and Olaf Niemann, K. (2012). A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. *Agricultural and Forest Meteorology*, 166-167, pp.72–85. doi:https://doi.org/10.1016/j.agrformet.2012.06.007.

Cutler, A., Cutler, D.R. and Stevens, J.R. (2012). Random Forests. *Ensemble Machine Learning*, [online] pp.157–175. doi:https://doi.org/10.1007/978-1-4419-9326-7_5.

Dai, A., Nießner, M., Zollhöfer, M., Izadi, S. and Theobalt, C. (2017). BundleFusion. *ACM Transactions on Graphics*, 36(4), p.1. doi:https://doi.org/10.1145/3072959.3054739.

Danks, N.P., Ray, S. and Shmueli, G. (2023). The Composite Overfit Analysis Framework: Assessing the Out-of-Sample Generalizability of Construct-Based Models Using Predictive Deviance, Deviance Trees, and Unstable Paths. *Management Science*. doi:https://doi.org/10.1287/mnsc.2023.4705.

Dice, L.R. (1945). Measures of the Amount of Ecologic Association Between Species. *Ecology*, 26(3), pp.297–302. doi:https://doi.org/10.2307/1932409.

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), pp.269–271. doi:https://doi.org/10.1007/bf01386390.

Disney, M. (2018). Terrestrial LiDAR: a three-dimensional revolution in how we look at trees. *New Phytologist*, 222(4), pp.1736–1741. doi:https://doi.org/10.1111/nph.15517.

Dubuisson, M.-P. . and Jain, A.K. (1994). *A modified Hausdorff distance for object matching*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICPR.1994.576361.

Erener, A., Sarp, G. and Karaca, M.I. (2020). An approach to urban building height and floor estimation by using LiDAR data. *Arabian Journal of Geosciences*, 13(19). doi:https://doi.org/10.1007/s12517-020-06006-1.

Ester, M., Kriegel, H.-P., Sander, J. and Xu, X. (1996). *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*. [online] Available at: https://file.biolab.si/papers/1996-DBSCAN-KDD.pdf [Accessed 18 Apr. 2023].

Everingham, M., Van Gool, L., Williams, C.K.I., Winn, J. and Zisserman, A. (2009). The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision*, 88(2), pp.303–338. doi:https://doi.org/10.1007/s11263-009-0275-4.

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), pp.185–201. doi:https://doi.org/10.1016/s0034-4257(01)00295-4.

Gislason, P.O., Benediktsson, J.A. and Sveinsson, J.R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters*, [online] 27(4), pp.294–300. doi:https://doi.org/10.1016/j.patrec.2005.08.011.

Gonzalez, R.C., Woods, R.E. and Masters, B.R. (2009). Digital Image Processing, Third Edition. *Journal of Biomedical Optics*, [online] 14(2), p.029901. doi:https://doi.org/10.1117/1.3115362.

Gordon, S.J. and Lichti, D.D. (2007). Modeling Terrestrial Laser Scanner Data for Precise Structural Deformation Measurement. *Journal of Surveying Engineering*, 133(2), pp.72–80. doi:https://doi.org/10.1061/(asce)0733-9453(2007)133:2(72).

Hackenberg, J., Spiecker, H., Calders, K., Disney, M. and Raumonen, P. (2015). SimpleTree —An Efficient Open Source Tool to Build Tree Models from TLS Clouds. *Forests*, 6(12), pp.4245–4294. doi:https://doi.org/10.3390/f6114245.

Hans Ole Ørka, Dalponte, M., Terje Gobakken, Næsset, E. and Liviu Theodor Ene (2013). Characterizing forest species composition using multiple remote sensing data sources and inventory approaches. *Scandinavian Journal of Forest Research*, 28(7), pp.677–688. doi:https://doi.org/10.1080/02827581.2013.793386.

He, K., Zhang, X., Ren, S. and Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.770–778. doi:https://doi.org/10.1109/cvpr.2016.90.

Heritage, G. and Hetherington, D. (2006). Towards a protocol for laser scanning in fluvial geomorphology. *Earth Surface Processes and Landforms*, 32(1), pp.66–74. doi:https://doi.org/10.1002/esp.1375.

Ho, T.K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), pp.832–844. doi:https://doi.org/10.1109/34.709601.

Hodgetts, D. (2013). Laser scanning and digital outcrop geology in the petroleum industry: A review. *Marine and Petroleum Geology*, 46, pp.335–354. doi:https://doi.org/10.1016/j.marpetgeo.2013.02.014.

Huttenlocher, D.P., Klanderman, G.A. and Rucklidge, W.J. (1993). Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, [online] 15(9), pp.850–863. doi:https://doi.org/10.1109/34.232073.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X. and Maltamo, M. (2008). Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, 29(5), pp.1339–1366. doi:https://doi.org/10.1080/01431160701736489.

Jaccard, P. (1912). THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE.1. *New Phytologist*, [online] 11(2), pp.37–50. doi:https://doi.org/10.1111/j.1469-8137.1912.tb05611.x.

Jain, A.K., Duin, P.W. and Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), pp.4–37. doi:https://doi.org/10.1109/34.824819.

King, G., Tomz, M. and Wittenberg, J. (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science*, 44(2), p.347. doi:https://doi.org/10.2307/2669316.

Krisanski, S., Taskhiri, M.S., Gonzalez Aracil, S., Herries, D. and Turner, P. (2021). Sensor Agnostic Semantic Segmentation of Structurally Diverse and Complex Forest Point Clouds Using Deep Learning. *Remote Sensing*, 13(8), p.1413. doi:https://doi.org/10.3390/rs13081413.

Kubat, M. (1999). Neural networks: a comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7. *The Knowledge Engineering Review*, 13(4), pp.409–412. doi:https://doi.org/10.1017/s0269888998214044.

Lemmens, M. (2011). Terrestrial Laser Scanning. *Geo-information*, 5, pp.101–121. doi:https://doi.org/10.1007/978-94-007-1667-4_6.

Liang, X., Kankare, V., Hyyppä, J., Wang, Y., Kukko, A., Haggrén, H., Yu, X., Kaartinen, H., Jaakkola, A., Guan, F., Holopainen, M. and Vastaranta, M. (2016). Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, pp.63–77. doi:https://doi.org/10.1016/j.isprsjprs.2016.01.006.

Lichti, D.D. (2007). Error modelling, calibration and analysis of an AM–CW terrestrial laser scanner system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(5), pp.307–324. doi:https://doi.org/10.1016/j.isprsjprs.2006.10.004.

Lichti, D.D. and Jamtsho, S. (2006). Angular resolution of terrestrial laser scanners. *The Photogrammetric Record*, 21(114), pp.141–160. doi:https://doi.org/10.1111/j.1477-9730.2006.00367.x.

Lugo, G., Li, R., Chauhan, R., Wang, Z., Tiwary, P., Pandey, U., Patel, A., Rombough, S., Schatz, R. and Cheng, I. (2022). LiSurveying: A high-resolution TLS-LiDAR benchmark. *Computers & Graphics*, 107, pp.116–130. doi:https://doi.org/10.1016/j.cag.2022.07.010.

Maheswari, K.B. and Gomathi, S. (2023). *Analyzing the Performance of Diverse Deep Learning Architectures for Weather Prediction*. [online] IEEE Xplore. doi:https://doi.org/10.1109/ICIRCA57980.2023.10220887.

Martin-Ducup, O., Mofack, G., Wang, D., Raumonen, P., Ploton, P., Sonké, B., Barbier, N., Couteron, P. and Pélissier, R. (2021). Evaluation of automated pipelines for tree and plot metric estimation from TLS data in tropical forest areas. *Annals of Botany*, [online] 128(6), pp.753–766. doi:https://doi.org/10.1093/aob/mcab051.

Mascaro, J., Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Martin, R.E., Anderson, C., Higgins, M. and Chadwick, K.D. (2014). A Tale of Two 'Forests': Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. *PLoS ONE*, 9(1), p.e85993. doi:https://doi.org/10.1371/journal.pone.0085993.

Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T. and Vihervaara, P. (2021). Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sensing of Environment*, 256, p.112322. doi:https://doi.org/10.1016/j.rse.2021.112322.

Montuori, A., Luzi, G., Stramondo, S., Casula, G., Bignami, C., Bonali, E., Bianchi, M.G. and Crosetto, M. (2014). *Combined use of ground-based systems for Cultural Heritage conservation monitoring*. [online] IEEE Xplore. doi:https://doi.org/10.1109/IGARSS.2014.6947384.

Murtagh, F. (1991). Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5-6), pp.183–197. doi:https://doi.org/10.1016/0925-2312(91)90023-5.

Neumann, B., Mettenleiter, M., Hildebrandt, A., Fröhlich, C. and Abmayr, T. (2004). *TERRESTRIAL LASER SCANNING - NEW PERSPECTIVES IN 3D SURVEYING*. [online] Available at: https://www.prip.tuwien.ac.at/legacy/oldwww.prip.tuwien.ac.at/cvch07/download/download/lectures/FROEHLICH/index.pdf.

Nwankpa, C., Ijomah, W., Gachagan, A. and Marshall, S. (2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*. [online] arXiv.org. Available at: https://arxiv.org/abs/1811.03378.

Park, H.S., Lee, H.M., Adeli, H. and Lee, I. (2007). A New Approach for Health Monitoring of Structures: Terrestrial Laser Scanning. *Computer-Aided Civil and Infrastructure Engineering*, 22(1), pp.19–30. doi:https://doi.org/10.1111/j.1467-8667.2006.00466.x.

Park, Y. (2008). *Dynamic Task Scheduling onto Heterogeneous Machines Using Support Vector Machine*. [online] etd.auburn.edu. Available at: http://etd.auburn.edu/handle/10415/1047 [Accessed 30 Aug. 2023].

Pfeifer, N. and Briese, C. (2007). Laser scanning – principles and applications. *GeoSiberia 2007 - International Exhibition and Scientific Congress*. doi:https://doi.org/10.3997/2214-4609.201403279.

Pieraccini, M., Noferini, L., Mecatti, D., Atzeni, C., Teza, G., Galgaro, A. and Zaltron, N. (2006). Integration of Radar Interferometry and Laser Scanning for Remote Monitoring of an Urban Site Built on a Sliding Slope. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9), pp.2335–2342. doi:https://doi.org/10.1109/tgrs.2006.873574.

Prokop, A. and Panholzer, H. (2009). Assessing the capability of terrestrial laser scanning for monitoring slow moving landslides. *Natural Hazards and Earth System Sciences*, 9(6), pp.1921–1928. doi:https://doi.org/10.5194/nhess-9-1921-2009.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M. and Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sensing*, 5(2), pp.491–520. doi:https://doi.org/10.3390/rs5020491.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N. and Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), pp.195–204. doi:https://doi.org/10.1038/s41586-019-0912-1.

Remondino, F. and El-Hakim, S. (2006). Image-based 3D Modelling: A Review. *The Photogrammetric Record*, 21(115), pp.269–291. doi:https://doi.org/10.1111/j.1477-9730.2006.00383.x.

Rina, S., Ying, H., Shan, Y., Du, W., Liu, Y., Li, R. and Deng, D. (2023). Application of Machine Learning to Tree Species Classification Using Active and Passive Remote Sensing:

A Case Study of the Duraer Forestry Zone. *Remote Sensing*, [online] 15(10), p.2596. doi:https://doi.org/10.3390/rs15102596.

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M. and Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, [online] 71, pp.804–818. doi:https://doi.org/10.1016/j.oregeorev.2015.01.001.

Sagi, O. and Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4). doi:https://doi.org/10.1002/widm.1249.

Shan, J. and Toth, C.K. (2018). *Topographic Laser Ranging and Scanning: Principles and Processing, Second Edition*. [online] *Google Books*. CRC Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=dGpQDwAAQBAJ&oi=fnd&pg=PP1&dq=(Shan [Accessed 26 Aug. 2023].

Sorrentino, S., Manetti, F., Bresci, A., Vernuccio, F., Chiara Ceconello, Ghislanzoni, S., Bongarzone, I., Vanna, R., Cerullo, G. and Polli, D. (2023). Deep ensemble learning and transfer learning methods for classification of senescent cells from nonlinear optical microscopy images. *Frontiers in Chemistry*, [online] 11. doi:https://doi.org/10.3389/fchem.2023.1213981.

Srikrishnan Divakaran (2023). An Algorithmic Framework for Fusing Images from Satellites, Unmanned Aerial Vehicles (UAV), and Farm Internet of Things (IoT) Sensors. pp.75–88. doi:https://doi.org/10.1007/978-981-99-0577-5_4.

Stilla, U. and Xu, Y. (2023). Change detection of urban objects using 3D point clouds: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 197, pp.228–255. doi:https://doi.org/10.1016/j.isprsjprs.2023.01.010.

Sumnall, M.J., Hill, R.A. and Hinsley, S.A. (2022). Towards Forest Condition Assessment: Evaluating Small-Footprint Full-Waveform Airborne Laser Scanning Data for Deriving

Forest Structural and Compositional Metrics. *Remote Sensing*, 14(20), p.5081. doi:https://doi.org/10.3390/rs14205081.

Taha, A.A. and Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, [online] 15. doi:https://doi.org/10.1186/s12880-015-0068-x.

Tao, S., Wu, F., Guo, Q., Wang, Y., Li, W., Xue, B., Hu, X., Li, P., Tian, D., Li, C., Yang, H., Li, Y., Xu, G. and Fang, J. (2015). Segmenting tree crowns from terrestrial and mobile LiDAR data by exploring ecological theories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110, pp.66–76. doi:https://doi.org/10.1016/j.isprsjprs.2015.10.007.

Tekaslan, H.E. and Melike Nikbay (2022). A Multi-fidelity Prediction with Convolutional Neural Networks Using High-Dimensional Data. *AIAA AVIATION 2022 Forum*. doi:https://doi.org/10.2514/6.2022-3719.

Tsang, I., Kwok, J. and Hk, P. (2005). Core Vector Machines: Fast SVM Training on Very Large Data Sets Pak-Ming Cheung. *Journal of Machine Learning Research*, [online] 6, pp.363–392. Available at: https://www.jmlr.org/papers/volume6/tsang05a/tsang05a.pdf [Accessed 30 Aug. 2023].

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkanen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalen, P., Gobakken, T., Koch, B., Naesset, E., Tokola, T. and Maltamo, M. (2011). Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, 85(1), pp.27–40. doi:https://doi.org/10.1093/forestry/cpr051.

Vicari, M.B. (n.d.). *tlseparation: Performs the wood/leaf separation from 3D point clouds generated using Terrestrial LiDAR Scanners.* [online] PyPI. Available at: https://pypi.org/project/tlseparation/ [Accessed 31 Aug. 2023].

Wallace, L., Lucieer, A., Malenovský, Z., Turner, D. and Vopěnka, P. (2016). Assessment of Forest Structure Using Two UAV Techniques: A Comparison of Airborne Laser Scanning

and Structure from Motion (SfM) Point Clouds. *Forests*, 7(12), p.62. doi:https://doi.org/10.3390/f7030062.

Wallace, L., Lucieer, A., Watson, C. and Turner, D. (2012). Development of a UAV-LiDAR System with Application to Forest Inventory. *Remote Sensing*, 4(6), pp.1519–1543. doi:https://doi.org/10.3390/rs4061519.

Wan Ling Wong, Li, J., LI, X., Lamoureux, E., Cheung, C. and Wong, T. (2013). Novel and Traditional Biomarkers of Diabetic Retinopathy Severity: Multi-category Classifications Modeling. *Investigative Ophthalmology & Visual Science*, [online] 54(15), pp.1533–1533. Available at: https://iovs.arvojournals.org/article.aspx?articleid=2146141 [Accessed 30 Aug. 2023].

Wang, L. (2005). *Support Vector Machines: Theory and Applications*. *Studies in Fuzziness and Soft Computing*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi:https://doi.org/10.1007/b95439.

Wang, Y., Zuo, Y., Du, Z., Song, X., Luo, T., Hong, X. and Wu, J. (2023). MInet: A Novel Network Model for Point Cloud Processing by Integrating Multi-Modal Information. *Sensors*, [online] 23(14), pp.6327–6327. doi:https://doi.org/10.3390/s23146327.

Wilkes, P., Disney, M., Armston, J., Bartholomeus, H., Bentley, L., Brede, B., Burt, A., Calders, K., Chavana-Bryant, C., Clewley, D., Duncanson, L., Forbes, B., Krisanski, S., Malhi, Y., Moffat, D., Origo, N., Shenkin, A. and Yang, W. (2022a). TLS2trees: a scalable tree segmentation pipeline for TLS data. *Methods in Ecology and Evolution*. doi:https://doi.org/10.1101/2022.12.07.518693.

Wilkes, P., Disney, M., Armston, J., Harm Bartholomeus, Lisa Patrick Bentley, Brede, B., Burt, A., Calders, K., Chavana-Bryant, C., Clewley, D., Duncanson, L., Forbes, B., Krisanski, S., Malhi, Y., Moffat, D., Niall Origo, Shenkin, A. and Yang, W. (2022b). *TLS2trees*: a scalable tree segmentation pipeline for TLS data. *bioRxiv (Cold Spring Harbor Laboratory)*. doi:https://doi.org/10.1101/2022.12.07.518693.

Wilkes, P., Lau, A., Disney, M., Calders, K., Burt, A., Gonzalez de Tanago, J., Bartholomeus, H., Brede, B. and Herold, M. (2017). Data acquisition considerations for Terrestrial Laser Scanning of forest plots. *Remote Sensing of Environment*, 196, pp.140–153. doi:https://doi.org/10.1016/j.rse.2017.04.030.

Wright, M.N. and Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77(1). doi:https://doi.org/10.18637/jss.v077.i01.

Wulder, M., Niemann, K.O. and Goodenough, D.G. (2000). Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. *Remote Sensing of Environment*, [online] 73(1), pp.103–114. doi:https://doi.org/10.1016/S0034-4257(00)00101-2.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W. and Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, [online] 121, pp.196–209. doi:https://doi.org/10.1016/j.rse.2012.02.001.

www.northdata.com. (n.d.). *RIEGL Laser Measurement Systems GmbH, Horn, Austria*. [online] Available at: https://www.northdata.com/RIEGL+Laser+Measurement+Systems+GmbH [Accessed 31 Aug. 2023].

Xu, D., Wang, H., Xu, W., Luan, Z. and Xu, X. (2021). LiDAR Applications to Estimate Forest Biomass at Individual Tree Scale: Opportunities, Challenges and Future Perspectives. *Forests*, 12(5), p.550. doi:https://doi.org/10.3390/f12050550.

Xu, Q., Zhao, B., Dai, K., Dong, X., Li, W., Zhu, X., Yang, Y., Xiao, X., Wang, X., Huang, J., Lu, H., Deng, B. and Ge, D. (2023). Remote sensing for landslide investigations: A progress report from China. *Engineering Geology*, [online] 321, p.107156. doi:https://doi.org/10.1016/j.enggeo.2023.107156.

Yastikli, N. (2007). Documentation of cultural heritage using digital photogrammetry and laser scanning. *Journal of Cultural Heritage*, 8(4), pp.423–427. doi:https://doi.org/10.1016/j.culher.2007.06.003.

Zhang, Y., Tino, P., Leonardis, A. and Tang, K. (2021). A Survey on Neural Network Interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5(5), pp.726–742. doi:https://doi.org/10.1109/tetci.2021.3100641.

Zhu, J., J. Gehrung, Huang, R., Borgmann, B., Sun, Z., L. Hoegner, Hebel, M., Xu, Y. and Uwe Stilla (2020). TUM-MLS-2016: An Annotated Mobile LiDAR Dataset of the TUM City Campus for Semantic Point Cloud Interpretation in Urban Areas. *Remote Sensing*, 12(11), pp.1875–1875. doi:https://doi.org/10.3390/rs12111875.

Zou, K.H., Warfield, S.K., Bharatha, A., Tempany, C.M.C., Kaus, M.R., Haker, S.J., Wells, W.M., Jolesz, F.A. and Kikinis, R. (2004). Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index. *Academic radiology*, [online] 11(2), pp.178–189. doi:https://doi.org/10.1016/S1076-6332(03)00671-8.