UTRECHT UNIVERSITY

Department of Information and Computing Science

Artificial Intelligence Master Thesis

Optical Remote Sensing for Agroforestry Vegetation Segmentation & Biomass Estimation

Supervisor:
Ronald Poppe

Author:
Weronika Gajda

External Supervisor:
Xi Zhu

In cooperation with:
Rabobank Acorn

July 11, 2024

Abstract

Accurately estimating biomass in forestry systems is important for evaluating carbon offsets. Traditional methods of manual measurement are labor-intensive, costly, and use considerable resources. This research explores the potential of high-resolution optical remote sensing as a more cost-effective and easily available data source.

Current deep learning and machine learning techniques face challenges due to the limited availability of data and the complexity of labeling satellite imagery. Typically, these studies rely on high-resolution data (10m-30m), which may not be optimal for biomass monitoring, with very few exploring super-high-resolution satellite data.

Additionally, existing studies often employ either pixel-to-pixel segmentation or deep learning (DL) methodologies. Pixel-to-pixel approaches are limited as they fail to capture the surrounding of each datapoint. Meanwhile, DL methods are data-intensive, often impractical due to the scarcity of ground truth data and the laborious, sometimes inaccurate process of manual labeling.

This study proposes a novel vegetation segmentation technique that requires minimal labeling. We developed a texture-based machine learning segmentation approach which uses Local Binary Pattern and Random Forest, and which achieved up to 95% accuracy on 4-classes segmentation. We also showed that this approach outperforms UNET, as well as that incorporating outcomes of this segmentation in the feature set significantly improves biomass estimation.

Keywords: Remote sensing, vegetation segmentation, biomass estimation, random forest, texture features

Contents

Appendix

1. Introduction

Humanity faces a crucial moment in its fight against climate change, marked by various challenging aspects, including the increased levels of CO2 in the Earth's atmosphere [1]. This increase in atmospheric CO2, primarily attributed to the burning of fossil fuels and deforestation, is a significant driver of global warming and climate change [1]. In this situation, the amount of carbon stored in forests and other natural ecosystems becomes crucial for reducing the effects of climate change. These stocks act as natural carbon sinks, absorbing CO2 from the atmosphere and thereby helping to balance the global carbon budget. However, the effective management and preservation of these carbon stocks necessitate a reliable method of supervision and monitoring [2]. It is here that Artificial Intelligence (AI) can play a transformative role. By applying AI in conjunction with advanced remote sensing technologies, we can not only enhance our understanding of carbon stock dynamics but also improve the accuracy and efficiency of monitoring these vital resources. This thesis explores the innovative integration of AI with high-resolution optical imagery to develop robust, scalable, and cost-effective methodologies for carbon stock assessment, thereby contributing to the broader efforts of combating climate change and ensuring a sustainable future.

1.1 Carbon Stocks and Their Monitoring

Carbon stocks refer to the quantity of carbon held within various ecosystems, such as forests, soil, and oceans. In the context of climate change, these stocks play a crucial role [2]. Forests, for example, act as carbon sinks by absorbing carbon dioxide (CO2) from the atmosphere through photosynthesis and storing it in the form of biomass in trees and soil. This natural process of carbon sequestration is vital for regulating atmospheric CO2 lev-

els, one of the primary greenhouse gases contributing to global warming [1]. The maintenance and enhancement of these carbon stocks are therefore essential in mitigating the adverse effects of climate change, as they directly influence the global carbon balance.

Building upon the concept of carbon stocks, carbon offset projects are instrumental in increasing these stocks, thereby playing a key role in climate change mitigation. These projects, which include forestry initiatives, renewable energy developments, and energy efficiency improvements, are designed to either sequester additional CO2 or reduce emissions. By capturing or preventing the release of carbon emissions, carbon offset projects not only contribute to increasing carbon stocks but also provide a means for companies and individuals to balance out their own carbon footprints. This approach promotes a shift towards more sustainable and environmentally responsible practices in both business operations and daily life.

Forestry projects, in particular, are a cornerstone of carbon offset initiatives. They focus on either preserving existing forests or planting new ones, both of which are vital in sequestering carbon dioxide from the atmosphere [3]. These projects not only help in reducing the concentration of greenhouse gases, but also contribute to biodiversity conservation, water resource management, and the provision of livelihoods for local communities [3]. However, the success of forestry projects depends on accurate measurement and monitoring of biomass, which is where the challenge lies [3].

One of the types of forestry projects uses agroforestry - incorporating trees/shrubs into farming systems, and therefore allowing for the production of trees and crops or livestock from the same piece of land in order to obtain economic, ecological, environmental and social benefits as discussed by the authors [4]. This types of forests present an even more complex scenario for biomass prediction, as vegetation complexity makes monitoring and quantifying biomass in these systems particularly challenging. The variability in species, tree densities, and the interplay with agricultural crops complicates the data collection process, necessitating more sophisti-

cated and adaptable monitoring techniques [5] .

1.2 Carbon Measurements Methodologies

Traditional methods for measuring carbon stocks, particularly in forest ecosystems, have primarily relied on field-based techniques, among which allometry stands out as an important approach [6]. Allometry in ecological context refers to the study and measurement of the relative growth of a part of an organism in relation to the entire organism. In forestry, this concept is applied to understand and quantify the relationship between the size of a tree (usually measured as diameter at breast height, diameter at breast height (DBH), wood density, and height) and its biomass, which in turn indicates the amount of carbon stored [7].

The allometric approach involves developing equations that relate easily measured tree parameters, like DBH and height, to more difficult-to-measure attributes, such as total biomass. These allometric equations are derived from detailed measurements of a sample of trees, which typically involve cutting them down and weighing each component (trunks, branches, leaves, and roots) to calculate their total biomass [7]. Once these equations are established, they can be applied to standing trees, allowing for the estimation of biomass and carbon stocks without the need for destructive sampling. The development of these equations, however, requires extensive fieldwork and can be specific to tree species and ecological zones, thus necessitating the creation of different equations for different forest types.

1.2.1 Sampling and Allometry in Practice

In the practical implementation of carbon offset projects, field sampling plays a crucial role. Typically, a team of field researchers and foresters is deployed to a designated area, often demarcated into several plots, to conduct detailed tree measurements [8]. The number of plots can vary significantly depending on the size of the area and the project's objectives, but the aim is to ensure a representative sample of the entire area. In each plot, the team measures a specific number of trees, which is determined based on statisti-

cal sampling techniques to achieve an accurate representation of the forest's overall biomass [8].

The process involves recording the species, as well as measuring the DBH and, in some cases, the height of each selected tree. The number of trees measured in each plot can range from a few dozen to several hundred, again depending on the project's scope and the forest's density [8]. These measurements are then used to estimate the biomass for each sampled tree using allometric equations, as previously discussed.

Extrapolation is a key step in this process. The collected data from the sampled plots are extrapolated to estimate the total biomass of the larger area. This extrapolation is based on statistical methods that account for the variability in tree sizes, species, and densities across the entire project area. By using these representative samples and careful extrapolation, researchers can estimate the total carbon stock of the forest without needing to measure every tree individually. This method provides a balance between accuracy and efficiency, enabling the practical assessment of large forest areas for carbon offset projects.

One significant drawback of the standard sampling/allometry approach is the challenge in data acquisition. As mentioned previously, most allometric equations are based on tree species, height, and DBH, necessitating manual sampling. This typically involves a team of foresters traveling to selected areas to collect the data. This is hardly a scalable solution; it influences both the quantity and quality of the data samples collected. To simplify the process and reduce costs, measurements are often taken at the forest's edge or near roads. These locations, however, may not be representative of the entire forest plot, potentially leading to biased or less accurate assessments of the forest's biomass and carbon stock.

1.2.2 The Potential of Remote Sensing

Remote Sensing involves the acquisition of information about an object or area from a distance, typically from aircraft or satellites, making it a non-intrusive and efficient method for environmental monitoring. Since tradi-

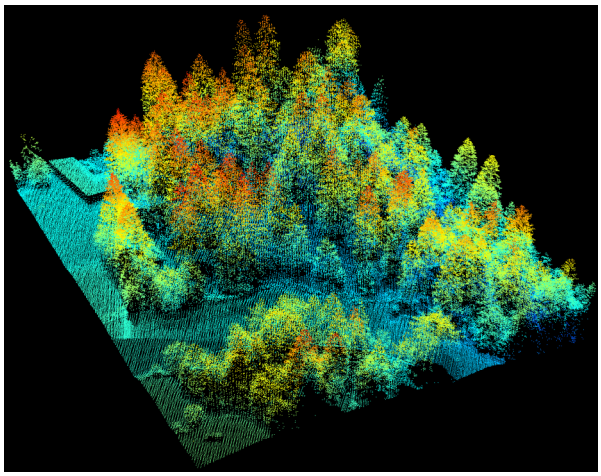


Figure 1.1: An example of LiDAR pointcloud, captured over a forest [14].

tionally biomass data collection has been a manual, labor-intensive process fraught with methodological inaccuracies and high costs, and there is a need for developing a less time-consuming and less prone to errors approach, Remote Sensing technology emerges as a promising alternative. A number of studies has proven successful estimation of biomass using this type of data [9] [10] [11].

1.2.2.1 LiDAR

Light Detection and Ranging (LiDAR) technology has become a cornerstone in the field of biomass assessment due to its ability to generate precise three-dimensional point clouds. A LiDAR point cloud is a collection of data points in space produced by a laser scanner. These points collectively form a detailed representation of the scanned area, capturing the contours and structures of the environment with high accuracy [12]. In the context of forest biomass, LiDAR's ability to measure the height and structure of trees is particularly valuable. Tree height is a critical parameter in biomass estimation as it correlates strongly with the overall mass of the tree. This direct relationship between tree height and biomass allows for accurate calculations of the carbon stock within a forested area [11] [13].

However, despite its precision, LiDAR technology comes with significant costs. The requirement for specialized equipment and aerial platforms, typically aircraft or drones, adds to the expense. Furthermore, operating

and interpreting LiDAR data requires skilled personnel, which further limits its accessibility and widespread use. These factors make LiDAR a less feasible option for frequent or large-scale biomass monitoring, especially in resource-limited settings.

1.2.2.2 Optical Data

High-resolution optical imagery, as provided by platforms like SkySat satellites, offers a more accessible alternative for biomass monitoring. The primary advantage of this type of data lies in its wide availability, comprehensive coverage, and relative affordability [15]; moreover, satellites nowadays capture a wide range of light spectrum - from which especially near-infrared band is helpful for vegetation monitoring [16].

Unlike LiDAR, high-resolution optical imagery is two-dimensional, providing detailed visual information about the surface but lacking direct data on the height and three-dimensional structure of vegetation. This absence of height information poses a challenge in accurately assessing biomass, as height is a key component in biomass estimation [17]. However, studies show the additional spectral information (e.g., the aforementioned NIR band) may be just as important in biomass monitoring [18].

Recent advancements in remote sensing technologies have significantly bolstered our capabilities in biomass estimation using high-resolution optical data. However, these methods often exhibit a degree of specialization that limits their broader applicability. For instance, algorithms developed for temperate forests might not perform as well in tropical or boreal ecosystems due to distinct structural and phenological differences. This specialization can hinder the scalability of these methods, making it challenging to apply them to larger and more diverse forest stands.

One promising approach to overcome this drawback is to focus on canopy segmentation methods. Knowledge about tree canopy size and shape, combined with the information about the species and allometry, may be sufficient to estimate the biomass: e.g, recent methodologies have shown promise in linking tree crown dimensions, potentially observable in two-dimensional

imagery, to biomass [19].

Another important indicator of biomass, photosynthetic activity, is possible to asses using optical data. Through the analysis of color variations and patterns in the imagery, it is possible to infer the health and productivity of vegetation, which are closely tied to biomass. These advancements suggest that high-resolution optical imagery could potentially match the accuracy of LiDAR (which does not indicate photosynthetic activity) in biomass prediction, while offering the benefits of lower costs and greater accessibility. This potential makes high-resolution optical data a promising avenue for efficient and cost-effective biomass monitoring, especially in extensive and frequent applications.

Choosing an appropriate AI method may also pose a challenge. Convolutional Neural Networks, especially models such as YOLO [20], or SAM [21], have gained a lot of popularity in recent years and seem to be an obvious choice for a segmentation (or instance segmentation) task. Unfortunately, these models have not been trained on satellite imagery, and show low performance when utilized for this type of data. There exist of course other models and architectures more suitable for the task, however these are not free from limitations: they are usually utilizing a specific dataset, such as SpaceNet [22] or specific type of satellite imagery - usually Landsat or Sentinel2 [23]. Moreover, the segmentation is usually performed on human infrastructure (roads, buildings), with very limited vegetation classes (often just 'vegetation').

A very significant limitation of satellite data is high variability between satellites, difficulty in labelling the data, and, therefore, lack of datasets suitable for training a Machine Learning model. Due to variability in vegetation between ecosystems, different sun angle in distinct places on Earth, and inconsistent satellite imagery properties in images captured by unassociated satellites, it is often the case that the algorithm would need to be retrained for each individual use. This is an obvious obstacle for using Deep Learning models which often require a significant number of learning examples.

On the other hand, using a simpler ML model usually relies on pixel-by-

pixel segmentation. In vegetation segmentation, this may be problematic, as single pixels of multiple vegetation classes may have the exact same properties. What is actually crucial for the task to succeed, is to consider the context (the surroundings) of the pixel. This approach is, however, uncommon and not sufficiently explored in research yet.

1.3 Project Goal & Research Question

The primary objective of this project is to develop a new biomass estimation methodology using optical imagery, which would surpass existing solutions in three key aspects: accuracy, versatility, and efficiency. It aims to deliver more precise biomass estimations across a diverse range of ecosystems and significantly reduce the dependency on extensive ground truth data, which is a common limitation in current approaches. The proposed process encompasses three distinct stages:

- Developing and Validating Vegetation Segmentation Mask: This initial step involves creating a vegetation segmentation mask using super high-resolution optical imagery (Pleiades Neo, 0.5m resolution). This classification mask is crucial for the subsequent stages.
- Biomass Estimation Using Vegetation Segmentation, with LiDAR-derived biomass as a label: this step involves using Machine Learning in order to estimate biomass values, with the use of remote sensing data and segmentation mask derived in the previous step.

The ultimate goal is for this methodology to serve as a baseline, which would be later fine-tuned in order to use cheaper, more accessible high-resolution imagery (such as Sentinel-2 with 10m resolution). This adaptation aims to apply the developed techniques to a wider range of satellite data sources.

Following this objective, the research questions that emerge are:

- What texture functions can be used in order to achieve at least 90% accuracy on the 4-classes canopy segmentation task? and
- Can the texture features in combination with random forest exceed the

performance of popular segmentation methodologies for small datasets?

- Does the use of canopy segmentation input increases the biomass estimation accuracy, in comparison to a traditional method using raw satellite imagery as input?

2. Literature Review

In this section, we provide a literature overview of existing research on biomass monitoring and species segmentation using remote sensing (RS). We begin with a concise description of Remote Sensing data sources (2.1) followed by describing Vegetation Indices (2.2), which are widely utilized in RS research. Next, we discuss the current methods of biomass monitoring using AI and RS data, along with their existing limitations (2.3). The third part of this chapter provides an overview of AI and RS in plant species segmentation (2.4). In both chapters 2 and 3, deep learning and non-deep learning methods are described.

2.1 Remote Sensing Imagery

In this section, we explore the various data sources utilized in remote sensing, highlighting their distinct advantages and applications. Understanding the diversity of these data sources is fundamental to comprehending the broader research domain of remote sensing.

2.1.1 Radar Data

Radar data, specifically Synthetic Aperture Radar (SAR) data, is important in remote sensing due to its capability to penetrate cloud cover and operate under all weather conditions. This feature makes SAR data useful for applications such as deforestation monitoring, land surface mapping, and disaster management. The ability of radar to provide consistent and reliable data regardless of environmental conditions is particularly beneficial for continuous and long-term monitoring projects.

Among the most commonly used data sources for radar data are Sentinel-1 and ALOS PALSAR.

Sentinel-1: Sentinel-1 is a constellation of two polar-orbiting satellites, Sentinel-1A and Sentinel-1B, launched by the European Space Agency (ESA). Sentinel-1A was launched in April 2014 and Sentinel-1B in April 2016. These satellites provide all-weather, day-and-night radar imagery with a revisit time of six days at the equator. The C-band radar instrument on Sentinel-1 operates at a wavelength of approximately 5.6 cm, which is suitable for various applications, including land and water monitoring, emergency response, and maritime surveillance. The high temporal frequency and spatial resolution of Sentinel-1 make it valuable for observing changes over time, such as monitoring deforestation or land subsidence [24].

ALOS PALSAR: The Advanced Land Observing Satellite (ALOS) with the Phased Array type L-band Synthetic Aperture Radar (PALSAR) was launched by the Japan Aerospace Exploration Agency (JAXA) in January 2006 and operated until May 2011. ALOS-2, its successor, was launched in May 2014. PALSAR operates in the L-band with a wavelength of about 23.6 cm, which allows for deeper penetration into vegetation and soil, making it particularly useful for forestry, agriculture, and geological applications. The longer wavelength of L-band radar is advantageous for monitoring biomass and forest structure, as it can penetrate the canopy and provide information about the underlying vegetation [25], [26].

However, these data sources have limitations when it comes to vegetation monitoring, particularly for biomass estimation:

- **Sentinel-1**: While Sentinel-1's C-band radar can provide detailed surface information, its shorter wavelength has limited penetration capabilities compared to L-band radar. This makes it less effective for capturing information about the structure of dense forests or estimating biomass. The C-band is more sensitive to surface scattering, which means it primarily captures information about the canopy surface rather than the volume of vegetation [24].
- **ALOS PALSAR**: Although ALOS PALSAR's L-band radar can penetrate deeper into vegetation and is better suited for monitoring forest structure and biomass, its spatial resolution is lower compared to

Sentinel-1. This can limit the level of detail and accuracy in biomass estimation, particularly in heterogeneous landscapes. Additionally, the availability of ALOS PALSAR data can be less frequent, and its operational timeline is discontinuous, with gaps between ALOS and ALOS-2 [25], [26].

In summary, while radar data from sources like Sentinel-1 and ALOS PALSAR is valuable for many remote sensing applications, they have inherent limitations for vegetation monitoring, especially for biomass estimation. These limitations necessitate the use of complementary data sources and advanced analytical methods to improve the accuracy and reliability of biomass assessments.

2.1.2 Optical Data

Optical data, derived from sensors capturing visible, near-infrared, and short-wave infrared light, is widely employed for its high spatial and spectral resolution. These characteristics make optical data essential for a range of applications, including agricultural monitoring, land cover classification, and urban planning. The detailed spectral information obtained from optical sensors allows for precise identification and analysis of various surface materials and conditions, enhancing the accuracy of remote sensing studies.

Among the various optical satellite data sources, several stand out due to their unique characteristics:

Sentinel-2: The Sentinel-2 mission, part of the European Space Agency's Copernicus program, consists of two satellites, Sentinel-2A and Sentinel-2B, launched in June 2015 and March 2017, respectively. These satellites provide high-resolution optical imagery in 13 spectral bands, ranging from visible to short-wave infrared. With a revisit time of five days at the equator, Sentinel-2 offers a spatial resolution of 10 meters for most bands, making it highly suitable for detailed land cover classification and vegetation monitoring. Sentinel-2 data is freely available, making it an accessible resource for researchers and practitioners [27].

Landsat: The Landsat program, a joint initiative of NASA and the US

Geological Survey, has been providing Earth observation data since 1972. The latest satellites, Landsat 8 (launched in 2013) and Landsat 9 (launched in 2021), capture imagery in 11 spectral bands, including visible, near-infrared, and thermal infrared wavelengths. Landsat offers a spatial resolution of 30 meters for most bands and a revisit time of 16 days. Landsat data is also freely available, which has made it a cornerstone for long-term environmental monitoring and change detection studies. [28], [29].

Pleiades: The Pleiades satellites, operated by Airbus Defence and Space, offer very high-resolution optical imagery with a spatial resolution of 0.5 meters for panchromatic and 2 meters for multispectral bands. The satellites capture data in four spectral bands (RGB and near-infrared). Pleiades provides high-detail imagery suitable for urban planning and precision agriculture, but access to this data comes at a higher cost compared to Sentinel-2 and Landsat [30].

Drone Imagery: Drones equipped with multispectral or hyperspectral sensors can capture high-resolution imagery with spatial resolutions down to centimeters. This level of detail is beneficial for small-scale studies, precision agriculture, and detailed vegetation analysis. Drones offer flexibility in data acquisition, allowing for targeted and timely data collection. However, their coverage area is limited compared to satellites, and flight operations can be constrained by regulations, weather conditions, and battery life [31].

Plane-Captured Imagery: Airborne platforms equipped with advanced sensors can cover larger areas than drones while providing high-resolution imagery. These platforms can carry a variety of sensors, including multispectral, hyperspectral, and LiDAR, offering comprehensive data for diverse applications. The main advantages of plane-captured imagery include high spatial resolution and the ability to cover large areas efficiently. However, the operational costs are higher, and data acquisition is less frequent compared to satellites.

When it comes to biomass monitoring, each data source has its strengths and weaknesses:

- **Sentinel-2 and Landsat**: Both provide free, high-quality imagery with

sufficient spatial and spectral resolution for large-scale biomass monitoring. The frequent revisit times and comprehensive spectral bands are advantageous for tracking vegetation changes and estimating biomass over large areas [27], [28], [29].

- **Pleiades**: The very high spatial resolution of Pleiades imagery allows for detailed biomass assessments at smaller scales. This can be useful for precision agriculture and urban forestry, though the higher cost and limited spectral bands may restrict its broader applicability [30].
- **Drone and Plane-Captured Imagery**: These platforms provide unparalleled spatial resolution and flexibility, making them ideal for detailed biomass studies in specific areas. The ability to deploy various sensors adds versatility. However, their limited coverage area and higher operational costs can be drawbacks for large-scale biomass monitoring [31].

In summary, the choice of data source for biomass monitoring depends on the scale of the study, budget, and specific requirements for spatial and temporal resolution. Combining multiple data sources can often provide the most comprehensive and accurate biomass assessments.

2.1.3 LiDAR Data

LiDAR (Light Detection and Ranging) data, which utilizes laser light to generate detailed three-dimensional maps of the Earth's surface, offers high accuracy and resolution. This technology is particularly useful for applications such as topographic mapping, forestry, autonomous vehicle navigation, and archaeological studies. By measuring the time it takes for laser pulses to return after hitting a target, LiDAR can produce precise elevation models and detailed structural information [32].

2.1.3.1 Applications and Advantages

LiDAR's ability to create high-resolution 3D models makes it invaluable for a variety of applications:

- **Topographic Mapping**: LiDAR is extensively used to create detailed Digital Elevation Models (DEMs) that accurately represent the terrain.
- **Forestry**: It can measure forest structure, including tree height, canopy density, and biomass, which are critical for forest management and carbon stock assessment [33].

2.1.3.2 Limitations

Despite its advantages, LiDAR data has some limitations:

- **Weather Sensitivity**: LiDAR's effectiveness can be compromised by adverse weather conditions such as heavy rain, fog, or snow, which can scatter or absorb the laser pulses.
- **Dense Vegetation**: While LiDAR can penetrate through gaps in the canopy to some extent, very dense vegetation can limit its ability to capture ground-level details accurately.
- **Operational Costs**: High operational costs are associated with LiDAR data acquisition, particularly from airborne platforms. The cost includes the equipment, flight operations, and data processing, which can be significant compared to other remote sensing methods [33], [32].

2.1.3.3 LiDAR for Biomass Monitoring

LiDAR is particularly well-suited for biomass monitoring due to its ability to measure forest structure in three dimensions. Here are some key points:

- **High Accuracy**: LiDAR can accurately estimate tree height, canopy volume, and biomass density, providing crucial data for forest carbon stock assessments and ecological studies.
- **Detailed Vegetation Structure**: Unlike optical sensors, LiDAR can penetrate the forest canopy and provide detailed information about the vertical structure of the forest, which is essential for understanding biomass distribution [32].

Challenges

- **Cost**: The high cost of LiDAR data acquisition can be a significant barrier for extensive biomass monitoring projects. The expenses associated with both drone-based and plane-based LiDAR need to be considered in the context of the project budget and scale [32].

2.2 Vegetation Indices

Vegetation Indices (VIs) are mathematical expressions that combine measured reflectance in multiple spectral bands of remote-sensing data to generate values useful in assessing plant growth, vigor, and various vegetation properties such as biomass and chlorophyll content [34]. Each VI requires specific band combinations as input. Bands are defined as specific ranges of wavelengths in the electromagnetic spectrum that satellite sensors detect, measured in nanometers (nm).

Band availability varies depending on the image source; for instance, super-high-resolution satellite imagery from Pleiades Neo offers 6 bands capturing different wavelength spectra: Deep Blue, Blue, Green, Red, Red Edge, and Near-infrared. In contrast, widely used Sentinel-2 provides 13 bands, including Coastal/Aerosol, Blue, Green, Red, Vegetation Red Edge 1, Vegetation Red Edge 2, Vegetation Red Edge 3, Near Infrared (NIR), Narrow Near Infrared (NNIR), Water Vapor, Shortwave Infrared 1 (SWIR1), Shortwave Infrared 2 (SWIR2), and Shortwave Infrared 3 (SWIR3). Notably, bands with the same names may not necessarily cover the same wavelength spectra (for a detailed band description, see A.1). Hence, it's essential to recognize that a Vegetation Index calculated with a specific formula may yield varying results with different data sources.

Commonly employed VIs include the Normalized Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Chlorophyll Index (CI), Enhanced Vegetation Index (EVI), and Leaf Area Index (LAI) [34].

The process of calculating VIs involves transforming bands in relation to

each other. For example, the formula for NDVI is:

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} \tag{2.1}$$

To apply this in practice, element-wise calculations are performed on Near Infrared and Red Bands. Consequently, the output of a Vegetation Index is a 2-dimensional matrix with the same height and width as the input bands, where each pixel's value carries information about the relative health and density of vegetation at its corresponding location on the Earth's surface.

In machine learning research, VIs are typically generated during preprocessing as part of feature engineering. They are often used as input variables to machine learning models, either instead of or in addition to raw satellite bands [35].

2.3 Biomass estimation

In this section, we explore how artificial intelligence (AI) methods are used to estimate biomass, which is important for monitoring ecosystems. We will start by looking at different ways to monitor biomass, including using deep learning and non-deep learning methods. Then, we will talk about the challenges of using AI for biomass prediction, like not having enough data or making models too complicated. Finally, we will discuss why it's important to divide vegetation into different parts when predicting biomass. This section aims to help us understand how AI can be used to estimate biomass and what challenges we might face.

2.3.1 Deep Learning Approaches

Multiple studies have used Deep Learning, especially Convolutional Neural Networks (CNN) [36], for biomass estimation. The advantages of using Deep Learning in this context include its ability to automatically learn hier-

archical features from raw data, which can be particularly beneficial when working with complex and high-dimensional remote sensing datasets. Additionally, Deep Learning models have shown promise in capturing intricate spatial patterns and relationships within imagery, thereby potentially improving the accuracy of biomass estimates. Moreover, Deep Learning approaches are often more flexible and adaptable to different types of data and can handle large volumes of information efficiently.

Deep Learning and optical remote sensing methodologies have been used for biomass estimation in various studies. One study developed a model using artificial neural networks (ANN) [37] and vegetation indices derived from remotely sensed data (Landsat5, 30m spatial resolution) to estimate Above-Ground Biomass (AGB) in the Brazilian Amazon [18]. The study however reported a quite big error (20% of biomass values). This is likely because of the resolution of the imagery used, and exceptionally high biomass values in Amazon Forest. Similarly, another study by [38] reported an R^2 value of 0.27 when predicting biomass using (Landsat8, 30m spatial resolution) data in China. These studies, utilizing only optical bands and simple vegetation indices as input features, underscore the limitations of existing Deep Learning approaches [38].

Other researchers use different Deep Learning architectures in order to lift the performance of the model in estimating biomass. Li et. al. [23] compared three non-parametric models, including ANN [37], random forests (RFs) [39], and quantile regression neural network (QRNN)[40], for AGB estimation in *Pinus Densata* forests in China. In this study, a high accuracy has been reported. It must be mentioned (and the authors themselves did so) that *Pinus Densata* forest is a very specific case of an ecosystem - it contains only one type of tree, and this species is quite remarkable as coniferous trees have distinctive tops and rarely intersect. Moreover, all the trees in the forest were reported to be of the same age, which is rarely the case in natural ecosystems, and even less likely to occur in carbon offset projects, especially agroforestry.

Poor generalization [41] of the algorithms used for biomass prediction

must be noted. A specific study by Pascarella et al. [42] developed a Regressive UNET for biomass estimation using Sentinel-2 data and ESA CCI BIOMASS Project [43] labels, which cover multiple species in Europe. Despite achieving what the authors deemed as very good results, the reported R^2 value of 0.6 +/- 0.1 indicates a significant error in terms of biomass estimation accuracy, particularly in the context of metric tons per hectare (MgT/ha). Moreover, the study highlights the challenge of generalizing Deep Learning models trained on European forests to other regions, as evidenced by a drop in accuracy when applied to Asian forests ($R^2 = 0.4$), despite being labeled as a "remarkable generalization" .

Recent research suggests that incorporating texture measures into the modeling process may offer a promising avenue for improving biomass prediction accuracy. For instance, another study [44] demonstrated that including texture features such as mean, variance, dissimilarity, entropy, contrast, correlation, and second moment significantly improved biomass prediction performance, achieving R^2 values of 0.63 using CNN [36] and 0.70 using CNN-LSTM [45] models.

Despite the theoretical advantages of Deep Learning methods for biomass estimation using remote sensing data, several limitations and shortcomings have been observed in published studies. One notable limitation is the requirement for a large amount of training data, which often leads researchers to utilize freely available datasets such as Sentinel-2 and Landsat. While these datasets offer broad coverage and accessibility, they come at the cost of spatial resolution, which may limit the ability to capture fine-scale vegetation patterns and heterogeneity. Furthermore, the labels used for training these Deep Learning models are often derived from datasets with relatively low resolution, such as the ESA CCI BIOMASS Project [43], which may introduce errors and inaccuracies into the training process [38].

The sub-field of super-high-resolution Remote Sensing remains relatively underexplored, primarily due to the high costs associated with acquiring this type of data. Notably, as of today (17/05/2024), Google Scholar does not report any study that utilizes Deep Learning with super-high resolu-

tion imagery for Forest Biomass estimation. While Sentinel-2 (10m spatial resolution) and Landsat 5/8 (30m spatial resolution) are suitable for many scenarios, biomass mapping for carbon offset purposes, particularly in agroforestry, necessitates imagery with much higher resolution, preferably under 1 meter. However, a significant challenge lies in the scarcity of input data, i.e., satellite imagery, and labels. There is a notable absence of widely accepted biomass datasets with resolutions of 1 meter or less. Therefore, developing algorithms for biomass estimation at such high resolutions poses a formidable challenge.

2.3.2 Non-Deep Learning Approaches

Non-Deep Learning methods for biomass estimation have also been extensively explored in the literature. These approaches encompass a diverse range of techniques, including traditional machine learning algorithms such as Random Forest [39], Support Vector Machines [46], and Decision Trees [47], as well as statistical and regression-based models [47]. Non-Deep Learning methods are often favored for their simplicity, interpretability, and computational efficiency, making them suitable for applications where computational resources are limited or where transparency and interpretability of the model are crucial [48]. Additionally, Non-Deep Learning approaches can be effective in situations where the dataset is relatively small or when feature engineering plays a significant role in model performance [49].

As in the case of Deep Learning-based studies, researchers here also use data of different resolutions to support their hypotheses. For example, one study compared biomass prediction using Random Forest [39], CNN [36], and CNN-LSTM [45], and found that using Random Forest allows for achieving comparable results to CNN and CNN-LSTM: for instance, RF with Landsat 8 can predict biomass with $R^2 = 0.73$ [44], while the CNN-LSTM and CNN obtained worse results ($R^2 = 0.72$ and $R^2 = 0.56$ respectively). The same study reports that the use of a geostationary satellite Gaofen-6 (2m resolution) and Sentinel-2 (10m resolution) gives comparable results.

In the realm of higher-resolution data, WorldView-2-based NDVI (spatial resolution of 2m) was used with Random Forest to estimate the biomass of wetlands [50]. An important contribution of this study is that the authors calculated NDVI multiple times, using different Near-Infrared bands available in WorldView-2 (Red Edge with a 705-745 nm wavelength, NIR with a 770-895 nm wavelength, and NIR2 with an 860-1040 nm wavelength). Random Forest trained on traditional NDVI, calculated using the NIR band, achieved much worse accuracy ($R^2 = 0.39$) than when using the Red Edge band instead of NIR in calculations ($R^2 = 0.79$).

This illustrates a critical issue with Remote Sensing-based Machine Learning: data sources differ not only in resolution but also in the specific wavelengths captured under the same band name. The NIR band from the WorldView 2 satellite is not the same as the NIR band from Sentinel-2 or Landsat 5, and therefore, developing a methodology with one specific type of data does not guarantee success if the same methodology is used with different types of data. For example, another study [51] used similarly prepared features and Random Forest, but with Landsat 5 data, achieving an accuracy of $R^2 = 0.22$.

Because of this, it must be underlined that algorithms relying solely on VIs are sensor-dependent. One of solutions to this issue is to use texture information, which is less sensor-dependent, as an additional input in these methodologies. For instance, Eckert et. al. [52] used Worldview (5m) imagery with texture information to estimate carbon biomass in Madagascar. By employing Step-wise Multiple Linear Regression [53], the authors achieved an $R^2 = 0.865$ using GLCM texture features. Remarkably, the study distinguished between non-degraded and degraded forests, estimating the biomass in each separately. As the authors note, "the more heterogeneous the forest canopy structure, the stronger the correlation with textural parameters." They also observed that the correlation between vegetation indices and biomass saturates at high biomass values [52].

2.3.3 Importance of Vegetation Segmentation in Biomass Prediction

Building upon the previous findings, another study highlights the importance of not only texture but also species detection in Above Ground Biomass (AGB) modeling. Zhang et al. [44] used Sentinel2 data and Support Vector Machine to first segment the imagery into 5 groups (Pine, Oak, Walnut, Other, and non-forest) and later used the segmentation results as input to the Random Forest algorithm, together with information about Vegetation Indices, texture, and SAR data. The reported accuracy of species segmentation is 80%, while R^2 of predicted biomass differs between species: it is equal to 0.61 in pine, 0.62 in oak, 0.78 in walnut, and 0.65 in forests with other dominant species. The authors of the study underline the importance of separate models for separate species, noting that because of AGB variations in different species, a model used for classifying all 3 species together was significantly overestimating low biomass values, and significantly underestimating high biomass values. They stress it is necessary to construct AGB models for different tree species.

A meta-analysis of studies on biomass modeling with Unmanned Aerial Systems corroborates these findings [54]. The authors emphasize that growth stage and plant structure, both of which are species-dependent, are likely the most influential factors affecting above-ground biomass (AGB) estimation accuracy as well as model robustness and transferability. Specifically, the analysis indicates that optical remote sensing is more susceptible to species variability due to its lack of structural information, which is otherwise provided by LiDAR data. In this case, information about the species can be used as a proxy for more detailed relationship between spectral information and biomass.

2.3.4 Challenges in AI-based Biomass Prediction: Summary

As described in previous sections, AI-based above-ground biomass (AGB) prediction using optical remote sensing faces several significant challenges. While Deep Learning (DL) methods are often robust and capable of achiev-

ing superior results compared to non-DL methods [55], their application in this domain is hindered by the absence of foundational models, a significant lack of data, and issues related to dataset labeling and the collection of ground-truth data. Unlike standard computer vision tasks, where target labels can be created relatively easily, preparing biomass heat maps (for image-to-image translation) or even parcel-level biomass estimations (where an image is regressed to a single biomass value) is a resource-intensive process [56].

The process of creating accurate biomass labels requires substantial effort and resources. A team of foresters must be deployed to the field to gather data, a task that is both costly and logistically challenging [57]. This is particularly true for tropical ecosystems, which are often remote and difficult to access due to poor transportation infrastructure and other logistical barriers [57]. Furthermore, the ground-truth data collected is only valid for a limited time after collection, as biomass can change rapidly due to factors like deforestation and vegetation growth. Therefore, it is crucial that optical imagery be captured shortly before or after ground measurements. This synchronization is difficult to achieve, especially considering that approximately 67% of the Earth's surface is covered by clouds at any given time [58], limiting the availability of cloud-free imagery.

This results in a very limited number of usable learning examples despite significant effort. It is estimated that it takes about three hours for an experienced team to measure a 0.1-hectare plot (1,000 square meters) [59] [60]. In publicly available datasets (e.g. [61], [43]), each such measurement typically corresponds to a single label. This demonstrates the considerable effort required to assemble a dataset of sufficient size for training deep learning models, which typically need large amounts of data. Consequently, despite the potential benefits of deep learning for biomass prediction, the practical constraints on data collection pose a significant barrier to its widespread application in this field.

Moreover, the variability in AGB across different species, ecoregions, and environmental conditions, along with fluctuations due to precipitation,

seasonality, and climate, complicates the use of a single methodology for biomass estimation across different satellite data, vegetation types, or regions. Remote sensing studies often produce results that are sensor-specific and region- or species-specific [51] [50]. This specificity underscores the need for methods that require minimal training examples, as these models will likely need retraining when applied to different ecosystems or locations. Consequently, a robust and adaptable biomass estimation approach should be capable of effectively utilizing a limited number of training samples to accommodate the diverse conditions encountered in various ecological settings.

Given these challenges, Machine Learning (ML) appears more suitable for biomass estimation as it generally requires fewer labeled examples [62]. Current research underscores the importance of texture information in this task, which can enhance the accuracy of ML models. Moreover, it would be beneficial to adopt a more step-wise approach, as the current error rates are too high for practical applications such as carbon markets. A step-wise method would also allow for more precise error tracing.

One promising solution is highlighted in the study by Zhang et al. [63], which proposes a two-step process: first performing species segmentation and then estimating biomass based on these segmented regions. This approach can potentially improve accuracy by addressing the specific characteristics of different vegetation types before estimating biomass.

Another notable factor highlighted in current studies is the subjectivity of scale. Despite successful attempts to map biomass using spectral data at resolutions ranging from 10 to 30 meters, the best results achieved still exhibit a mean absolute error (MAE) as high as 18.2 megatonnes of biomass per hectare [38] [42]. While this level of error may not be problematic at the national level, it could lead to significant miscalculations in carbon offset projects, particularly those involving agroforestry.

Furthermore, super-high-resolution imagery remains underexplored in the context of biomass estimation, paralleling the situation observed in Machine Learning studies. To date, only one study has utilized Pleiades Neo

for biomass predictions, yielding sub-optimal results [64]. Notably, this study did not employ advanced methodologies such as texture functions or a comprehensive set of vegetation indices. The limited number of studies in this area can be attributed to the high costs associated with acquiring high-resolution data. However, this research gap presents a significant opportunity for further investigation, as higher-resolution models typically offer increased accuracy compared to their low-resolution counterparts. Additionally, high-resolution models are better suited for applications requiring precise delineation, particularly those involving small parcels of land.

In summary, available research on biomass monitoring shows that an accurate estimation may be possible, but the uniformity of the species and high-resolution data are crucial for this task. The information about the species could possibly be delivered in the form of a segmentation mask, but current methodologies suffer from being ecosystem-specific, as well as lack the ability to use the main feature that seems to be logical to use - the texture of the image. This project aims to fill this gap and investigate the potential of using texture-derived segmentation masks for biomass estimation.

2.4 Segmentation

Building on the promising findings of Zhang et al. [44], further exploration of species segmentation and texture analysis holds considerable promise. Segmentation in Computer Vision refers to the process of dividing a digital image into multiple segments or regions to facilitate analysis and understanding of the image content. This technique plays a crucial role in various fields such as disease detection, crop management, robotics, and medical image analysis [65][66]. Image segmentation is particularly significant in clinical cardiology for anatomical segmentation using artificial intelligence-based computer vision, aiding in automation and novel applications within the medical field [67]. Deep learning models, like the U-Net architecture, have been widely utilized for image segmentation tasks, achieving high accuracy scores such as an IoU of 0.80 and a dice coefficient of 0.82, surpassing some state-of-the-art techniques [68]. Additionally, segmentation is essen-

tial in tasks like building segmentation from aerial images and LIDAR data, where deep learning frameworks like U-Net have been successfully applied, achieving competitive IoU scores for accurate building segmentation [69].

In this section, both deep-learning and non-deep learning segmentation methods for remote satellite imagery will be delineated and subsequently compared. Additionally, the pivotal roles of texture in enhancing segmentation accuracy and effectiveness will be underscored.

2.4.1 Deep Learning vs. Non-Deep Learning Approaches

In recent years, numerous studies have demonstrated significant advancements in satellite image segmentation ([70], [71], [72]). While these methods are highly accurate, they primarily focus on detecting human infrastructure. In contrast, using AI for vegetation segmentation presents a more formidable challenge. Notably, one of the most popular GitHub repositories on AI for remote sensing, [73], boasts 8,000 stars, but within its 40 topic categories and over 1,000 links, only 12 methods pertain to tree segmentation/classification. Of these, only two address closed-canopy tree segmentation/classification, and just one is based on non-hyperspectral data [74].

A logical choice would be to use one of the state-of-the-art (SOTA) models, such as Mask R-CNN [75], YOLO [20], or SAM [21]. However, these models were not pre-trained on satellite imagery, leading to sub-optimal performance on this type of data. While they perform reasonably well for segmenting human infrastructure (e.g., roofs), they struggle to accurately segment vegetation, especially different vegetation classes. For this research, SAM and YOLO were applied to various datasets containing vegetation, and their performance was found to be inadequate.

Due to the challenges associated with species segmentation, researchers have explored alternative methods, such as UNET [76], for this task. In one notable study, Kattenborn et al. [77] utilized UNET for canopy segmentation using drone data with a resolution of 3 cm. This study focused on distinguishing *Pinus radiata* and *Ulex europaeus*, achieving segmentation accuracies of 87% and 84%, respectively. The findings demonstrate that with

high-resolution data and a sufficient number of training examples, convolutional neural networks (CNNs) can perform the segmentation task effectively. However, the extrapolative capabilities of the algorithm remain uncertain, as it was not validated on the same tree species growing in different ecoregions.

The authors also highlighted the significant effort required to compile their dataset. In addition to UAV imagery, fieldwork was necessary to geotag trees and provide additional information for accurate delineation. Despite this effort, the entire dataset comprises only seven images, each capturing an areas of 150 by 150 meters. While this may be sufficient for 3 cm imagery, the scenario changes when considering satellite imagery. For instance, with the super high-resolution Pleiades Neo satellite imagery at 0.5 m resolution, one such tile would be only 300 by 300 pixels, which is insufficient to train a neural network without risking overfitting. This example underscores the substantial effort required to gather an adequate dataset for species segmentation tasks.

These findings are corroborated by Schiefer et al. [78], who employed CNNs for species segmentation in a German forest. Their study focused on distinguishing between nine classes. The authors collected data from thirteen plots, each measuring 100 by 100 meters. They acknowledged that accurate delineation of the classes required ground measurements such as tree height, geolocation, and diameter at breast height (DBH). Similar to the study by Kattenborn et al., the amount of training data they gathered is sufficient for training a neural network if the data resolution is 3 cm. However, if we aim to use satellite imagery with a resolution of 0.5 meters, this amount of data would be inadequate.

Fricker et al. [79] employed Convolutional Neural Networks (CNNs) along with 1m resolution hyperspectral data cubes from NEON Airborne Observation Platform (AOP) [80] to perform tree classification. This hyperspectral imagery, comprising 426 bands [81], is notably uncommon in remote sensing data, which typically features far fewer bands. The authors achieved an impressive average classification f-score of 0.83 for hy-

perspectral data and 0.59 for RGB data. This highlights the advantage of using high-dimensional hyperspectral data, as most super-high resolution imagery generally contains only four bands, while high-resolution (10m) data typically has twelve bands.

The study faced significant challenges in its preparation. LiDAR had to be flown over the entire area to create accurate labels for canopy segmentation. Additionally, a highly-specialized GPS system was necessary to align field samples with the trees visible in the LiDAR and optical data, as regular GPS systems lacked the required accuracy. The researchers also emphasized the critical need for optimal lighting and weather conditions during data capture, noting that even slight deviations in lighting could impair the algorithm's performance. Despite these efforts, only 544 trees were measured, underscoring the demanding nature of the study.

In order to mitigate the problem of data scarcity, Non-Deep Learning approaches are also popular in segmentation tasks. One study used Support Vector Machine to detect and map rubber plantation with Sentinel 2- and Landsat 8-derived Vegetation Indices [82]. The reported accuracy was 87% and 85%, respectively. Another study [83] used Vegetation Indices and Random Forest to address a similar task: rice and crop mapping. The accuracy, even though Sentinel 2 data was supplemented with Sentinel 1, reached 73%. This level of accuracy is not ideal. Moreover, Sentinel 2 is 10m resolution data and is very well suitable for national-level predictions, is more challenging to use plantations with small parcel sizes.

Another way to combat the lack of high resolution labelled datasets is to use unsupervised methods for such task. In the realm of super-high resolution data, SLIC segmentation [84] in combination with Random Forest was adapted to extract tree canopy [85]. The study reported high classification accuracy, with the algorithm being able to distinguish between single- and multi-tree canopies. There are however very big limitations to this method: it was developed and tested on areas with very sparse vegetation, growing on bare ground, which is not a very commonly occurring scenario, as in the majority of use cases the trees grow in between bushes, or are otherwise

surrounded with other vegetation. Moreover, the method is only suitable for classifying one type of tree only and is unable to differentiate between species. This is likely because SLIC segmentation bases on colour features [84], and different tree species can have very similarly looking canopies, especially in their green-leaf stage [86].

Closed-canopy forest segmentation is a particularly challenging task. Similar to biomass prediction, research suggests that texture analysis can be advantageous for this task as well. For example, Immitzer et al. [87] employed wavelet-based texture measures to perform superpixel segmentation, followed by using a random forest classifier to distinguish between two different tree species within the canopy. Notably, unlike the majority of studies, this research utilized high-resolution (5m) WorldView-2 data. The classification accuracy achieved was 0.74 for pine and 0.79 for spruce. Similar results were achieved by Wang et. al. [88], who used JSEG color segmentation [89] and later wavelet-based texture measures for forest delineation on orthophotos with 0.5m spatial resolution.

2.4.2 Challenges and Limitations of Current Approaches

In Section 2.2, it was emphasized that biomass prediction using Deep Learning methods is challenging due to the data-intensive nature of these algorithms. Gathering a sufficiently large and well-labeled dataset for successful training is very difficult. This challenge is also present in segmentation tasks. There are many widely available datasets for rooftop mapping, road mapping, or water body mapping. However, fewer datasets are available for land cover mapping. In these datasets, vegetation is typically labeled either as a single class, "vegetation," or divided into forest and non-forest vegetation. This granularity is insufficient for the objectives of this thesis, which focus on mapping different vegetation types.

Moreover, as with biomass estimation, available land-use datasets are often at a 10m resolution because they are based on Sentinel-2 data. One such popular dataset is the CORINE Land Cover dataset, which, as noted by [61], is not perfect and frequently contains mislabeled pixels. While this dataset

was sufficient to train a UNET architecture, issues such as mislabeling and unbalanced classes make the task very challenging.

It is also necessary to note that there is no universal definition of terms such as 'forest,' 'cropland,' or 'bushland'; different organizations and countries specify them differently. For example, the United Nations Framework Convention on Climate Change defines forests as "land having a minimum canopy cover of 10-30%, minimum tree height of 2-5 meters, and a minimum area of 0.1 hectares" [90]. On the other hand, the European Environment Agency defines a forest as "land spanning more than 0.5 hectares with trees higher than five meters and a canopy cover of more than 10%, or trees able to reach these thresholds in situ" [91]. This variation poses an additional challenge when using pre-trained models, as such discrepancies can render a method unusable.

In practice, this means it is nearly impossible to create a single labeled dataset that can be universally applied for training Machine Learning models. Labeling must be redone each time a new region is analyzed, leading to a scarcity of labeled data. Consequently, methods that require less training data are more valuable.

2.5 Summary

Artificial intelligence (AI) in remote sensing remains a relatively underexplored domain, particularly when it comes to leveraging super high-resolution satellite imagery. While many researchers utilize widely available data sources such as Landsat 5 (30m), Landsat 8 (30m), and Sentinel-2 (10m), there is significantly less focus on data with 5m resolution like WorldView-2, and almost no studies employ satellites like Pleiades Neo with 50cm resolution. This underutilization of high-resolution imagery presents both a challenge and an opportunity for advancing remote sensing applications.

Both biomass prediction and species segmentation are inherently difficult tasks, primarily due to the limited availability of datasets and the substantial effort required to prepare these datasets. Furthermore, the extrap-

olative capabilities of current algorithms are limited; methods that work well in one ecoregion or with one species often do not perform effectively with others. This lack of generalizability hinders broader application and adoption of these techniques across diverse environments.

These tasks are interlinked, as demonstrated by studies indicating that species segmentation can enhance biomass predictions, improving accuracy and error traceability. Traditionally, researchers have focused on vegetation indices for these tasks, but recent studies highlight the significant role of texture features in both biomass prediction and species segmentation. Incorporating texture analysis can provide more detailed and reliable results, underscoring the need for more comprehensive approaches.

The aim of this work is, therefore twofold: first, to develop a segmentation framework based on texture features and a machine learning (non-deep learning) model that can be trained with a limited amount of examples and does not require precise delineation of species. Second, to utilize the results of this segmentation framework to improve biomass prediction. This approach seeks to address the limitations of current methods by providing a more efficient and generalizable solution for remote sensing applications.

3. Data

Multiple data sources were used in this study, and the selection of final sites was based on the data availability - they cover locations where Acorn projects are located, and therefore the data was easy to obtain for these areas.

Originally, we wanted to focus on vegetation segmentation using high-resolution satellite imagery (Pleiades Neo). This part of the data is described in a section 3.1 of this chapter.

However, what we also wanted to verify is the importance of vegetation segmentation for biomass estimation. Because of external issues with data acquisition, we were not able to obtain LiDAR data, necessary for biomass calculation, for the initial project area. For this reason, another area was selected, for which optical and biomass data were available. The optical data for this region is described in section 3.2, whereas the biomass data is described in the section 3.3

3.1 Pleiades Neo - Kenya site

3.1.1 Project site description

Pleiades Neo data was captured over an area in Kenya, in Embu county (see figures 3.1 and 3.2). This type of imagery has 0.3 m spatial resolution and is delivered in 6 spectral bands: red, green, blue, red-edge, near-infrared, and deep blue. For vegetation monitoring, bands red, green, and near-infrared are most commonly used. See figure 3.5 for reference.

The image 1, located in the north-western part of the county, was captured on the 03.02.2024. The image 2, located more towards the centre of the county, was captured on the 28.01.2024

Figure 3.1: Red dots indicate the location of the project in Kenya

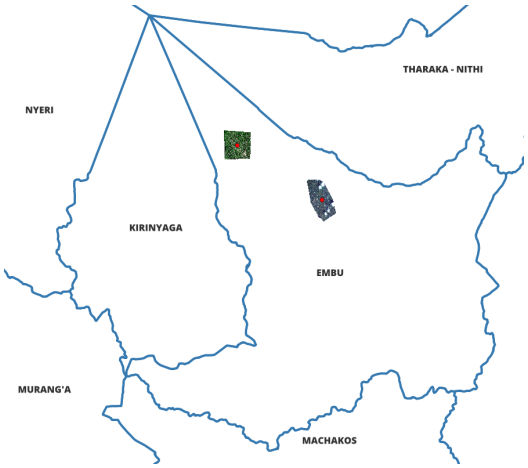


Figure 3.2: Red dots indicate the location of the project in Embu County

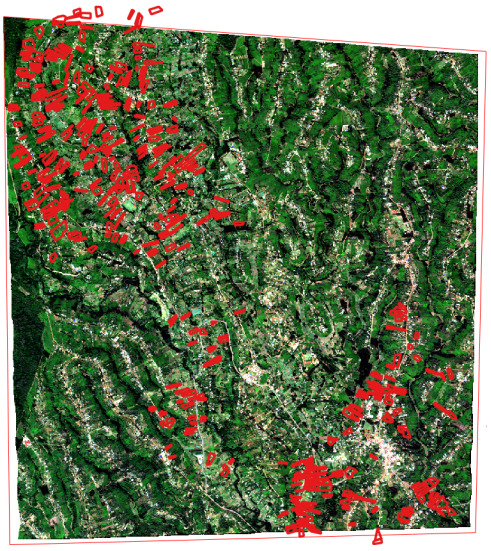


Figure 3.3: Pleiades Neo image of the Kenya project area - image 1

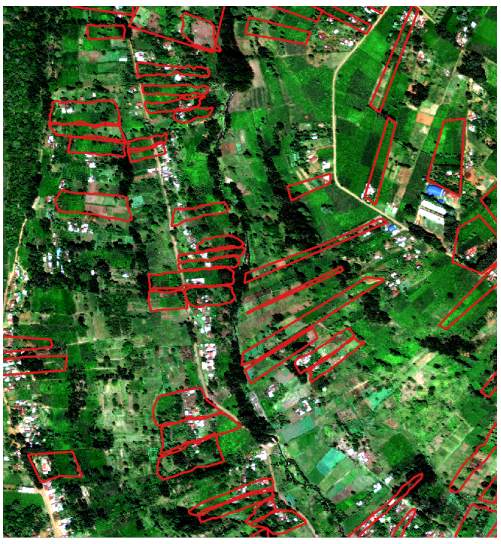


Figure 3.4: Zoom in into Pleiades image

Figure 3.5: Pleiades Neo: Kenya

3.1.2 Data Preprocessing

3.1.2.1 Vegetation Detection Methodology

The dataset underwent a cleaning process to ensure the exclusion of non-vegetative regions. This was accomplished by the computation of the Normalized Difference Vegetation Index (NDVI) [34], which is given by the equation:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \tag{3.1}$$

where NIR represents the near-infrared band, and Red denotes the red band of the electromagnetic spectrum. Regions exhibiting an NDVI value less than 0.5 were classified as non-vegetated and subsequently excluded from the analysis.

Additionally, the red-to-green ratio was calculated to further improve the vegetation detection process:

$$\text{Red to Green Ratio} = \frac{R}{G} \tag{3.2}$$

where R and G represent the reflectance values in the red and green bands, respectively. A threshold value of 1.0 was applied to this ratio, with values exceeding this threshold indicating the absence of vegetation.

The final vegetative areas were delineated by applying an OR operation to both of these masks; the non-vegetation mask can be defined as:

$$\text{non\_vegetation} = (\frac{R}{G} > 1.0) \text{ or } (\text{NDVI} < 0.5) \tag{3.3}$$

Consequently, the remaining dataset exclusively represented regions with a confirmed presence of vegetation (3.8).

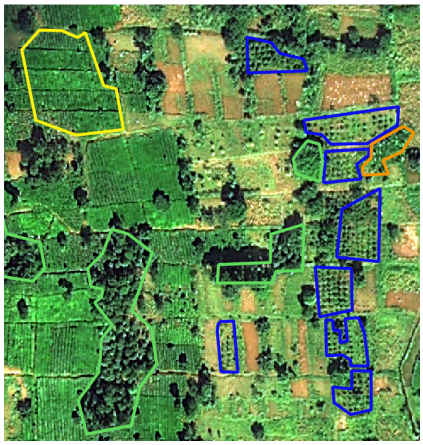


Figure 3.6: Sub-area captured in Pleiades Neo image. One can notice multiple regions not covered with vegetation.

Figure 3.7: The same sub-area with vegetation mask applied. Pink areas designate non-vegetated parts.

Figure 3.8: Vegetation detection

3.1.2.2 Shadow Detection

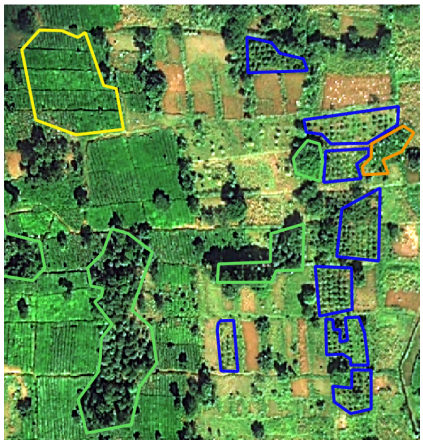
One of the hypotheses of this project suggests the influence of vegetation-casted shadows on correct classification methodology. In order to test it, we developed a simple algorithm for shadow detection, with the following formula:

$$\text{shadow} = ln(\text{red} + \text{green} + \text{blue}) < 7.3 \tag{3.4}$$

The 7.3 parameter was chosen experimentally and was guided by visual inspection of the image.

3.1.3 Data Annotation

Because this research utilizes a supervised learning approach, high-resolution imagery was manually labeled (using QGIS Desktop 3.30.2) into 4 distinc-



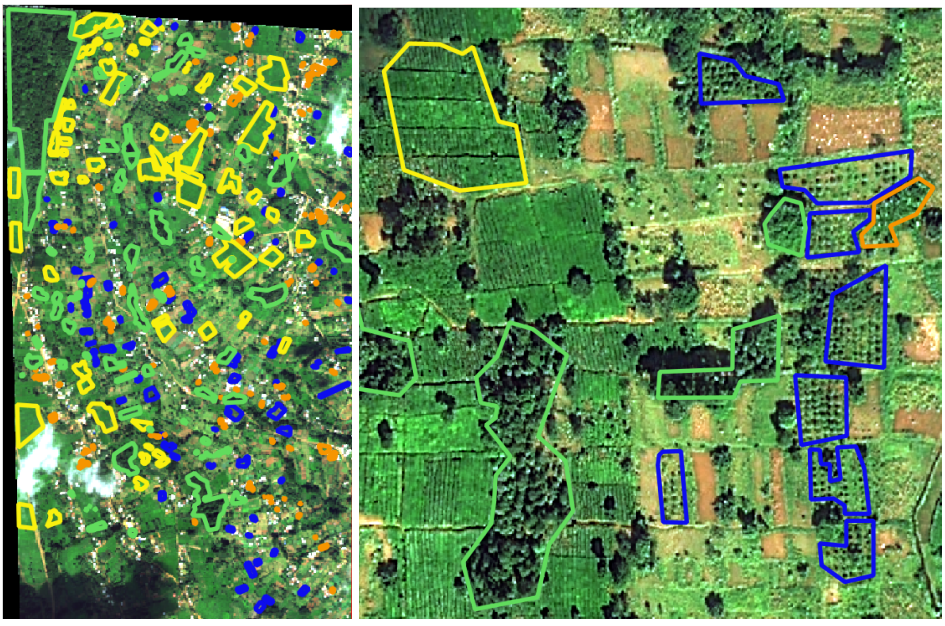
(a) Sub-area captured in Pleiades Neo image. One can notice multiple regions covered in shadow.

(b) The same sub-area with vegetation mask (pink) and shadow mask (dark blue) applied. Pink areas designate non-vegetated parts.

Figure 3.9: Shadow detection

tive classes of vegetation: broadleaves, palms, crops, and coffee trees. In total, an area of 13032 m^2 (1.30325 ha) was annotated, resulting in 52130 learning examples. The annotations were not complete, meaning that around 60% of the raster used were not annotated.

An example of these classes can be seen in Figure 3.10.



(a) Original image **(b)** Zoomed-in image

Figure 3.10: Areas selected as vegetation examples: green indicates broadleaves, yellow - crops, blue - coffee, and orange - palms.

3.2 Orthophoto - Colombia site

A high-resolution (3cm), plane-based orthophoto was captured over an area in Colombia, Risaralda Department (see Figure 3.11 and 3.12). This image consists of three bands (Red, Green, and Blue) and covers an area of 2281 hectares. See Figure 3.15 for reference.

Figure 3.11: Red dot indicates project location in Colombia

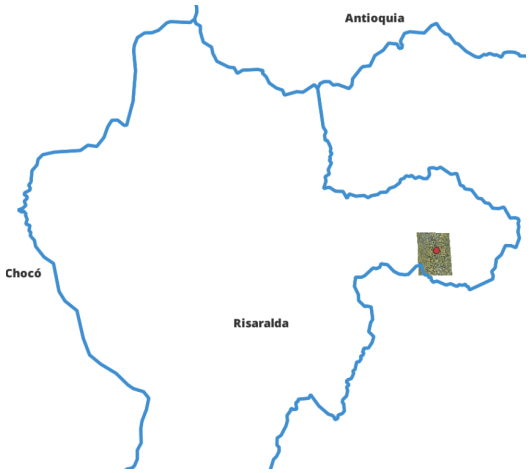


Figure 3.12: Red dot indicates the location of the project in Risaralda Department

3.2.1 Data Preprocessing

The orthophoto captured over the Colombia site consists of only three layers (Red, Green, and Blue), so NDVI and Vegetation Mask were not computed. Additionally, the orthophoto was captured around midday, minimizing the presence of shadows, and therefore the shadow masking step was skipped.

The original resolution of the orthophoto, at 3 cm, did not align with the 30 cm resolution of the Pleiades Neo imagery, which was used to develop the segmentation method. To harmonize these discrepancies and enable the reuse of the segmentation approach, the orthophoto was downsampled to a 30 cm resolution (see Figure 3.18). This adjustment was made to more closely match the Pleiades Neo imagery specifications. The downsampling

Figure 3.13: Orthophoto image over Colombia site

Figure 3.14: Zoom in into the orthophoto.

Figure 3.15: Orthophoto: Colombia

process was straightforward: every tenth pixel was sampled in both the width and height dimensions, utilizing NumPy for efficiency. This method ensured that the data remained consistent across different imaging sources.

Figure 3.16: Zoom-in into the orthophoto: before downsampling (3 cm resolution)

Figure 3.17: Zoom-in into the orthophoto: after downsampling (30 cm resolution)

Figure 3.18: Orthophoto Downsampling

3.2.2 Data Annotation

To segment the orthophoto, a part of this image had to be annotated, as it's properties are significantly different from Pleiades Neo image (e.g. only 3 bands). In the area, there were 6 distinctive vegetation classes that were used: coffee trees, palms, banana trees, broadleaves, coniferous trees, and grasslands. The details about annotations can be seen in a table 3.1

Segmentation Class	Area (m^2)	Amount of Pixels
Broadleaves	102470	$\approx$ 113855556
Coniferous	90433	$\approx$ 100480000
Grassland	67529	$\approx$ 75032222
Palms	78795	$\approx$ 87550000
Banana Trees	37130	$\approx$ 41255556
Coffee Trees	64673	$\approx$ 71858889

Table 3.1: Statistical data for different vegetation classes with calculated pixel amounts.

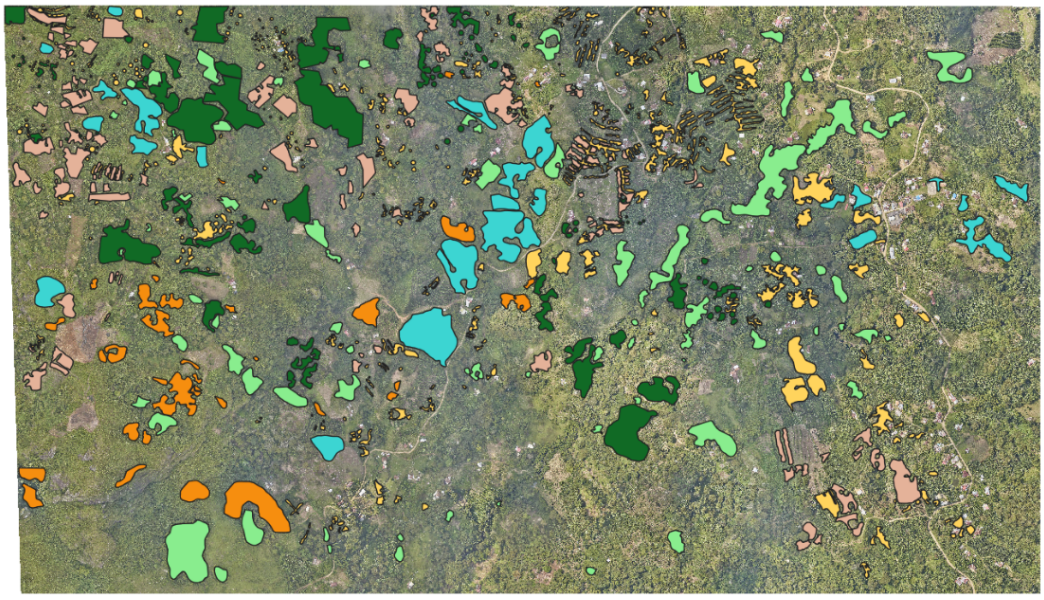


Figure 3.19: Zoom in into the annotated orthophoto.

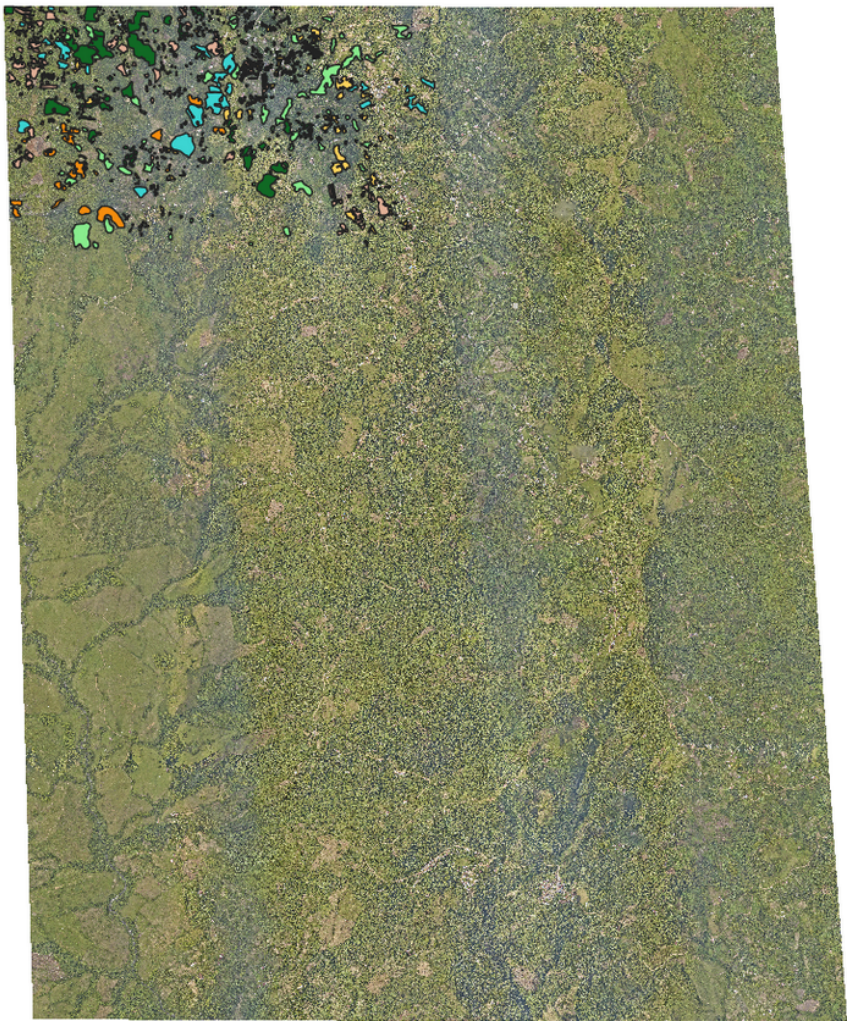


Figure 3.20: Annotated orthophoto image over Colombia site

3.3 Biomass data - Colombia site

The biomass ground truth data was derived from LiDAR using Acorn's proprietary methodology. This data, delivered at a 10 m resolution, covers the entire area of the orthophoto (2281 ha). It expresses biomass values in tonnes of biomass per hectare.

For mapping purposes, the raster was upsampled to a 30 cm resolution using bicubic interpolation [92] to match the orthophoto's resolution. An example comparison can be seen on the figure 3.23.

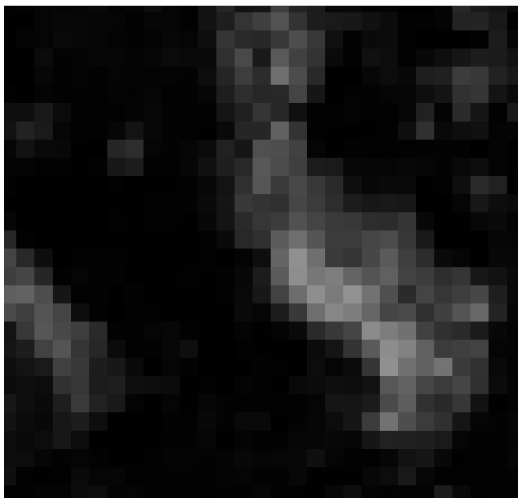


Figure 3.21: Zoom-in into the biomass layer: before upsampling (10 m resolution)

Figure 3.22: Zoom-in into the biomass layer: after upsampling (30 cm resolution)

Figure 3.23: Biomass layer upsampling

The original range of values of biomass was between 0 and 96 tonnes per hectare. After interpolation, the range of biomass changed to be between -11 and 104 (see Figure 3.24 for reference) in order to mitigate it, the values of interpolated biomass were clipped to the old range.

At the conclusion of the segmentation phase, we successfully generated 2,236 tiles, each measuring 256 by 256 pixels. Each tile included data for the red, green, and blue bands, alongside corresponding segmentation and biomass data. It is pertinent to note that these tiles represent only a subset of the entire area covered by the segmentation, not the entirety.

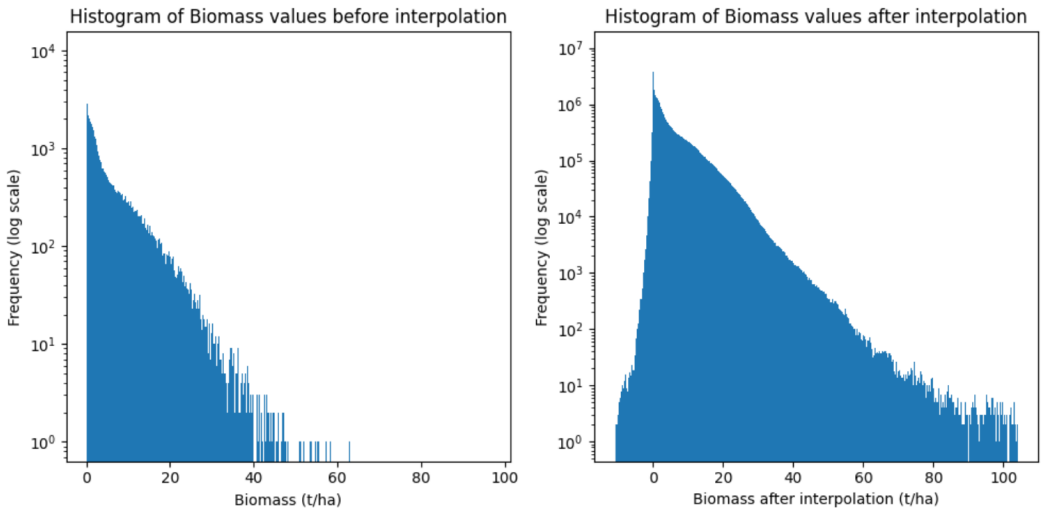


Figure 3.24: Comparison of biomass values distribution before and after bicubic interpolation

The tiles were subsequently loaded into a Python environment and transformed into a pixel-wise dataframe using Pandas. Initially, this dataframe comprised a total of 7,929,856 rows, with each row representing individual pixel information.

This raw dataframe underwent further processing to enhance data quality:

- Rows containing NaN values were removed.
- Pixels where all RGB values were either 255 or 0 were excluded from analysis.

These steps refined the dataset down to 7,249,180 entries. The highest recorded value of biomass within this cleaned dataset was 56.8125.

Following data cleaning, a log transformation was applied to normalize the distribution of biomass values [93]. This transformation adjusted the range of biomass values to between 0 and approximately 4.023. The effects of this normalization on the data distribution are illustrated in Figure 3.25.

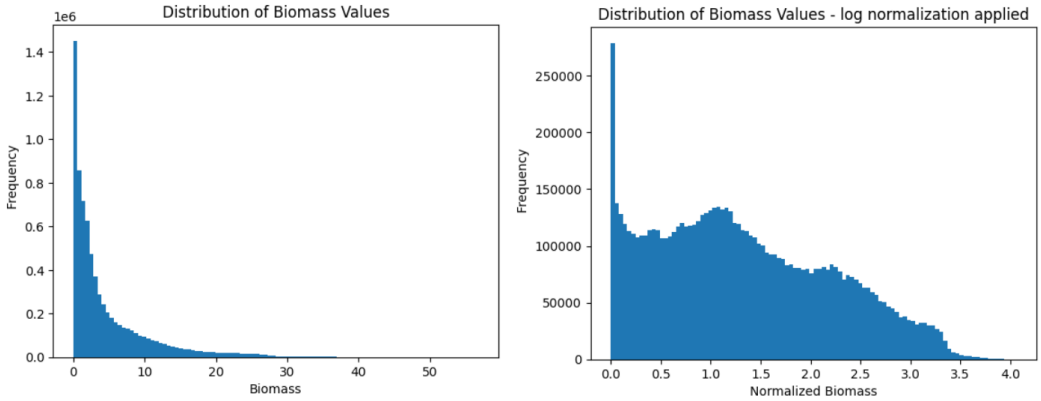


Figure 3.25: Comparison of values distribution before and after log normalization

4. Method

This section outlines the step-by-step methods used in this study, including data annotation and initial processing, species separation, and biomass estimation. Our approach utilizes high-resolution optical satellite images to improve current segmentation methods, as well as prove their utility in the biomass mapping task.

Because this research project consists of two parts, this chapter is divided into two sections - developing a methodology for species segmentation (section 4.1) and biomass estimation (section 4.2)

4.1 Species Segmentation

The species segmentation part of the project has a 6 step methodology. The aim of it is to classify high-resolution Pleiades Neo satellite imagery into different vegetation types. The six steps are:

- Data Annotation;
- Data points selection;
- Feature Preparation;
- Iterative feature selection using RF classification;
- N parameter selection;
- Random Forest hyper-parameters fine-tuning.

The aim of these steps is to obtain a set of features, as well as a trained RF algorithm, that can be used to prepare a 2D segmentation mask. The mask can later be used for statistics, as well as as an input for biomass calculations.

Figure 4.1 shows the performed steps for Random Forest Training.

Moreover, since the texture features selection, as well as data points def-

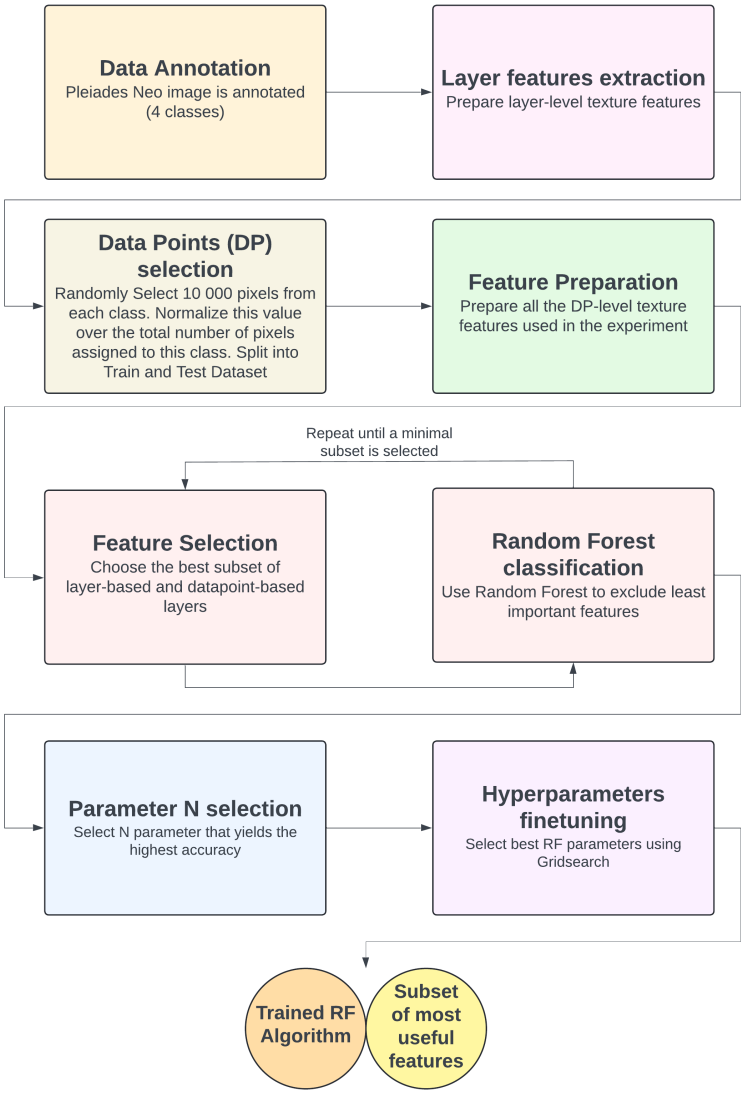


Figure 4.1: Visualization of an ML pipeline for Random Forest training

inition, are an important part of this research, this part of the pipeline (the first four blocks, or the first two rows of Figure 4.1) was visualized in more detail in the Figure 4.2.

In the following subsections, the steps from this Figure will be explained and described.

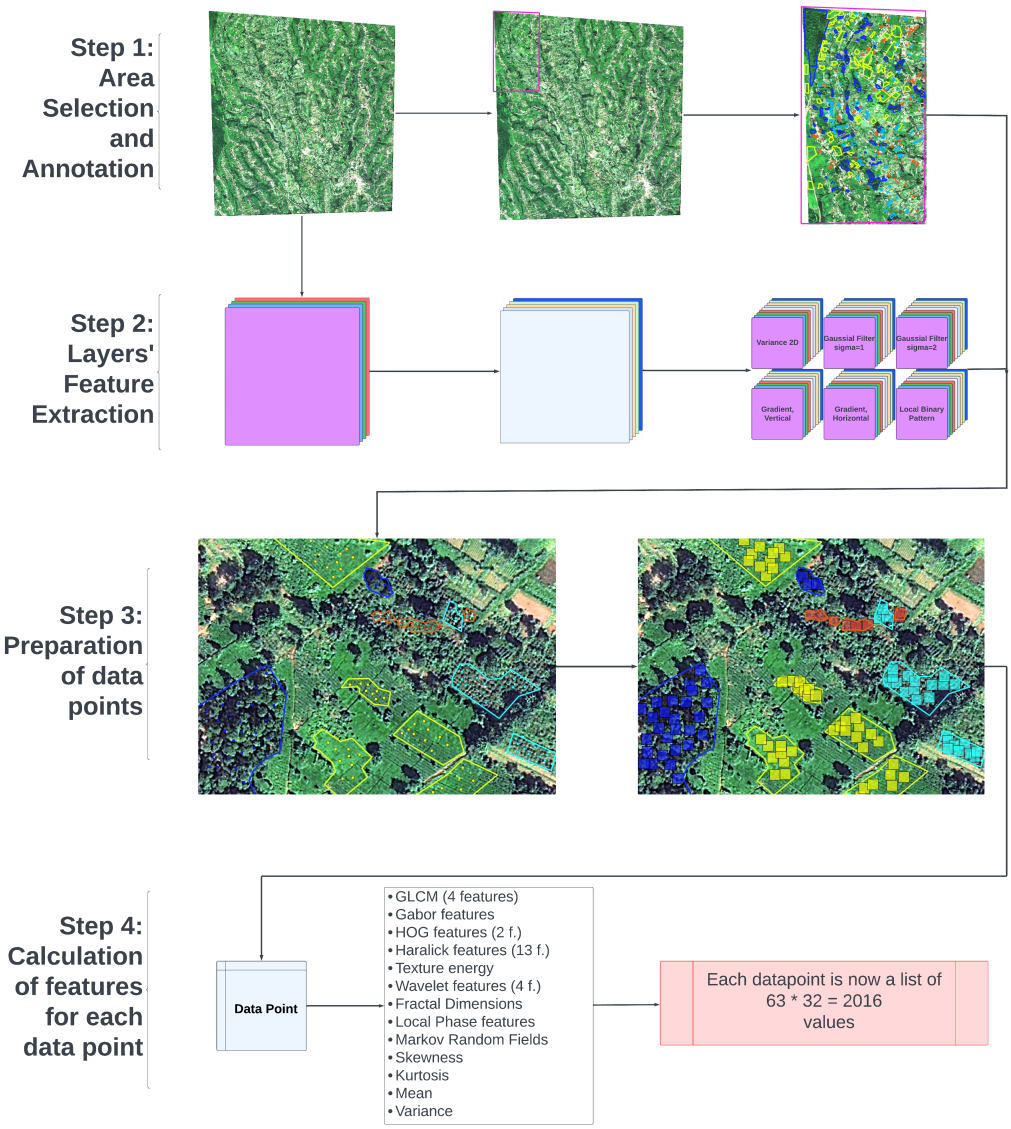


Figure 4.2: Processing & Feature Engineering pipeline, encompassing 4 different steps. The pipeline allows for effective computing of all texture features for one set of datapoints.

4.1.1 Step 1: Data Annotation

This step of the pipeline was described in the data chapter 3, in the section 3.1.3

4.1.2 Step 2: Layers processing and layer features extraction

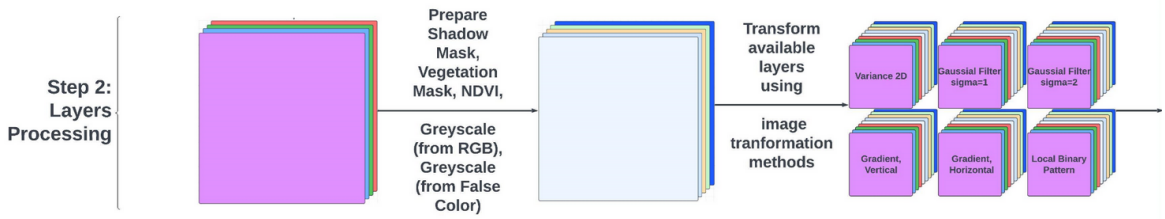


Figure 4.3: Step 2 of the pipeline: Layers Processing and Feature Extraction

This step was partially described in section 3. Here, the input raster was loaded and the bands were processed. In the training stage, the raster was clipped to a smaller area, closely bounding the annotated areas. As a result, 4 input bands (red, green, blue, and nir) were used to create Shadow Mask, Vegetation Mask, Greyscale 1 (prepared from an RGB image), and Greyscale 2 (prepared from a False Color image).

These were used as input to calculate additional layers: each of the 8 input layers was transformed using 6 different functions: Variance 2d function, Gaussian Filter (sigma=1)[94], Gaussian Filter (sigma=2), Vertical Gradient, Horizontal gradient, and Local Binary Pattern [95].

An example of these transformations can be seen in Figure 4.4, while the mathematical definition of these transformations is summarised in Figure B.1.

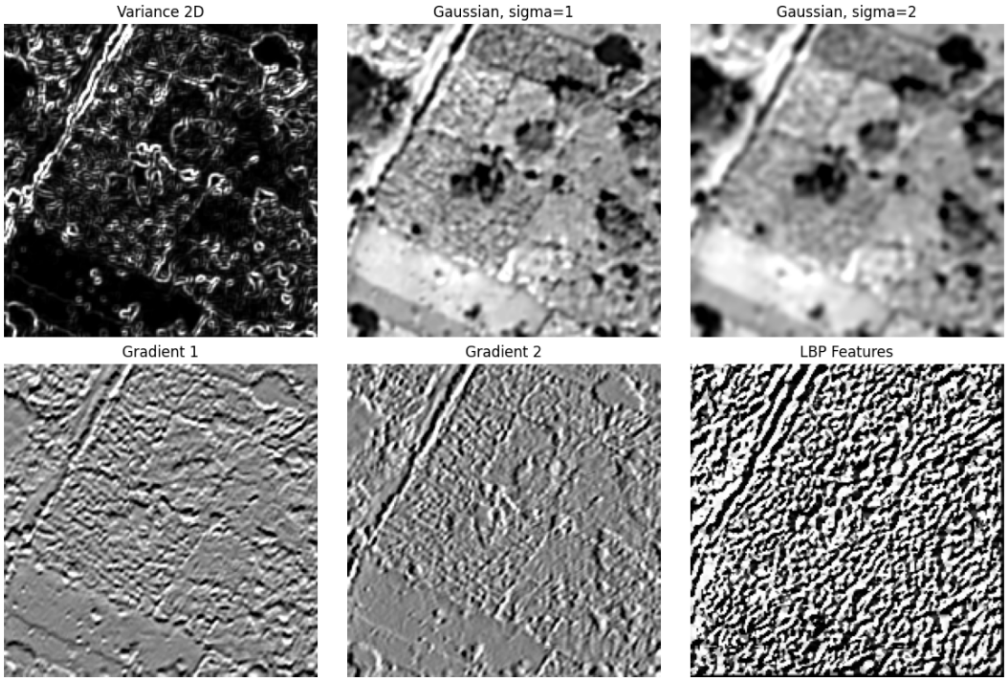


Figure 4.4: Step 2 of the pipeline: Layers Processing and Feature Extraction

4.1.3 Step 3: Preparation of data points

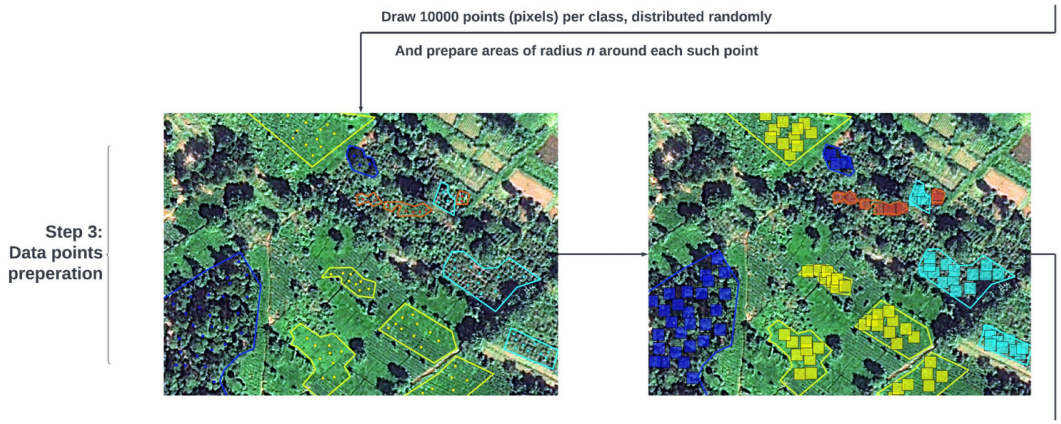


Figure 4.5: Step 3 of the pipeline: Data points preparation

In this step, 10000 pixels from each labeled class were randomly selected. This happened only once and these pixels were used in all further calculations and experiments, in order to obtain comparable results. Based on them, 4 different datasets were created, differing in the size of the window surrounding each pixel:

- DS-10 with a window size of 10 by 10 pixels surrounding the image
- DS-15 with a window size of 15 by 15 pixels surrounding the image
- DS-20 with a window size of 20 by 20 pixels surrounding the image
- DS-25 with a window size of 25 by 25 pixels surrounding the image.

An area of N by N pixels surrounding a given coordinate will be referred to as a **Data Point (DP)**

4.1.4 Step 4: Datapoint-level feature extraction

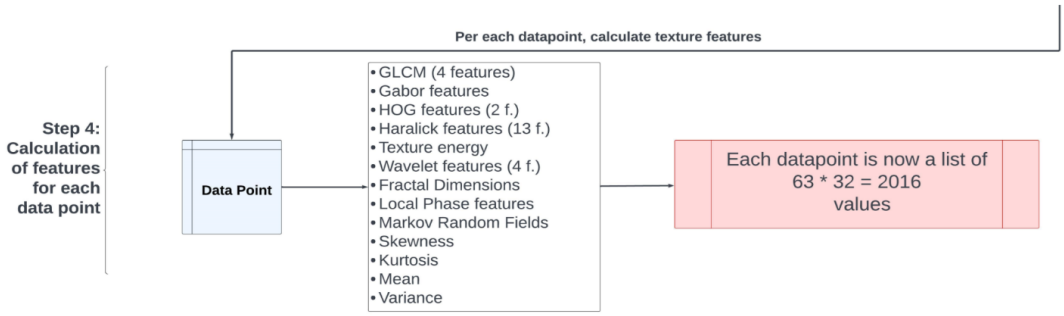


Figure 4.6: Step 4 of the pipeline: Data points level feature extraction

In Step 4, DP-level texture feature calculations were employed. This process includes calculating 13 different texture features for each data point of size n by n: GLCM, Gabor features, HOG features, Haralick features, texture energy, wavelet features, fractal dimensions, local phase features, Markov random fields, skewness, kurtosis, mean, and variance. The pre-selection of these features was made based on current literature and the ease of use. The description of these features can be found in Figures B.3, B.2, and B.4.

4.1.5 Feature Selection Loop

Feature selection is an essential step in data analysis, particularly when dealing with large and complex datasets. In our study, each DP consists of 2016 features - we start with 9 layers, each transformed 63 times, which gives us 567 layers; they are then transformed with texture features which together return 32 values, which gives us 2016 values per DP, in total.

Given that our classification task involves only 4 classes, it is reasonable to assume that all 2016 features would not be necessary; moreover, some features may be highly correlated. Another point to consider is that computing all features for segmentation purposes is impractical, as it takes an extensive amount of time.

To address these challenges, feature selection was performed to reduce the dimensionality of the dataset. This process helps in identifying and retaining only the most relevant features, thus making the analysis more efficient and manageable.

Two different approaches were employed to optimize feature selection for the task at hand, each with distinct focus points. Both approaches began with the full set of features and systematically excluded groups of features one at a time, performing classification using Random Forest algorithm, and evaluating the impact on accuracy.

The first approach focused on excluding texture features initially. It started with all possible layers and sequentially excluded texture features from each layer. After determining the essential texture features, it then proceeded to exclude entire layers.

The second approach took the opposite strategy by initially excluding entire layers. Once the relevant layers were identified, it then focused on excluding texture features based on data points (DP).

This difference in the sequence of feature exclusion defines the primary distinction between the two approaches, leading to varying impacts on classification accuracy and computational efficiency.

4.1.6 N parameter selection

The core assumption of this study is that texture features can be encoded into a single float value and subsequently used for species segmentation. This raises the natural question: how large should the surrounding area be to effectively capture texture? Intuitively, larger surrounding area provides more context, but can also introduce noise. To address this question, we investigated square areas of size N by N and examined the optimal value of N from the set 10, 15, 20, 25, 30, 35, 40. Considering the image is in 40cm resolution, these values correspond to 4, 6, 8, 10, 12, 14, and 16 meters.

After identifying the optimal subset of texture features to use, the data points are recalculated using different values of N. It is important to note that the pixel coordinates remain unchanged to maintain consistency and ensure comparability of results. Once the texture features are recalculated for various values of N, the accuracy of the Random Forest classifier is evaluated. The value of N that yields the highest accuracy is then established as the best parameter.

4.1.7 Hyperparameter Fine-Tuning

The final phase included hyperparameter fine-tuning of the Random Forest classifier. With the selected features and N parameter, the data was passed through a GridSearch algorithm using Cross Validation to determine the best hyperparameters.

Cross Validation [96] is a statistical method used to estimate the skill of machine learning models. It involves splitting the dataset into a set of training and test sets multiple times in different ways and ensuring that every observation from the original dataset has the chance of appearing in the training and test set. This provides a robust estimate of model performance. In this study, 10-fold cross validation was used. It is an important step to include in our project, as we operate on a relatively small dataset of features.

Grid Search [97] is a systematic way of working through multiple combinations of parameter tunes, cross-validating as it goes to determine which

tune gives the best performance. The algorithm exhaustively searches over the specified parameter values.

The range of hyperparameters, as well as their description, can be found in table 4.1.

Hyperparameter Name	Description	Values Considered
`n_estimators`	The number of trees in the forest.	[50, 100, 150, 200, 250, 300, 350]
`max_depth`	The maximum depth of the tree.	[None, 5, 10, 20, 30]
`min_samples_split`	The minimum number of samples required to split an internal node.	[2, 5, 7, 10]
`min_samples_leaf`	The minimum number of samples required to be at a leaf node.	[1, 2, 3, 4]
`bootstrap`	Whether bootstrap samples are used when building each tree.	[True, False]

Table 4.1: Hyperparameters considered for fine-tuning the Random Forest classifier.

These parameters were chosen based on their impact on the model's performance, and the search aimed to find the optimal combination to maximize the accuracy of the Random Forest classifier.

4.2 Biomass Estimation

Because of the lack of biomass data from the Kenya region, a high-resolution drone orthophoto from Colombia was employed to test the hypothesis about biomass estimation. The original orthophoto, with a resolution of 3 cm per pixel, was downsampled to 30 cm per pixel to match more closely the resolution of the original data used in the development of the segmentation methodology. It is important to note here that the orthophoto does not include a Near-Infrared layer, which was used for the initial development of the methodology.

To evaluate the hypothesis that segmentation improves biomass estimation, the orthophoto's annotations were used to retrain the Random Forest algorithm described in Section 4.1.5. The entire image was then segmented using the trained Random Forest model. The input raster data was divided into 256 by 256-pixel tiles, each of which was segmented, resulting in a 256x256x1 layer where each integer value represents a specific class assigned to each pixel. This segmentation layer was then combined with the original input channels, including Red, Green, and Blue channels, creating a dataset that integrates both spectral and segmentation information.

A biomass layer derived from LiDAR data was delivered to us with a 10m resolution. To enhance the resolution to match the input image resolution, the image was resized using the OpenCV Python library. Cubic interpolation [92], a more complex method compared to linear interpolation, was chosen for this task as it typically yields smoother results and a sharper output image. The biomass layer was then tiled into 256 by 256 tiles, and added to the according tiles of RGB(N) + segmentation layer, creating

Each tile was then flattened to create a dataframe. For setup 1, the data frame included columns for Red, Green, Blue, Segmentation, and Biomass values. For setup 2, the data frame included columns for Red, Green, Blue, and Biomass values. The resulting data frames comprised a total of 7,929,856 rows.

Next, the data was divided into three subsets to ensure an unbiased eval-

uation of the model's performance. The training set consisted of 70% of the data, while the testing and validation sets each contained 15% of the data. This division allowed for distinct data subsets for training, testing, and validation, facilitating a robust assessment of the model's accuracy and generalizability.

XGBoost [98], a gradient-boosting algorithm known for its efficiency and performance with structured data, was used to estimate biomass. The algorithm was chosen as it is relatively lightweight, as well as less prone to overfitting than other methods.

The inclusion of segmentation layers was hypothesized to improve the model's ability to capture spatial patterns and heterogeneity in the data, leading to more accurate biomass estimation. By comparing the performance of the model with and without the segmentation layers, this study aimed to determine the impact of segmentation on biomass prediction accuracy. The results of this analysis provided insights into the effectiveness of incorporating segmentation information in remote sensing applications for biomass estimation.

5. Results

This thesis aimed at finding answers to multiple different hypothesis. Both species segmentation and biomass estimation results will be discussed, focusing on the results of feature selection, N parameter exploration, species segmentation results using Random Forest algorithm, and biomass estimation error analysis.

5.1 Species Segmentation Results

5.1.1 Feature Selection

The methodology for feature selection utilized two distinct approaches: one focused on layer-based features and the other on DP-based features. Each approach identified a unique set of optimal features. Additionally, to assess the performance impact of omitting texture information, we incorporated two variants of a control approach into our analysis. This allowed us to evaluate whether a dataset lacking texture information could achieve comparable results.

Approach 1: Emphasis on Layer Features: the first approach prioritized layer features over texture features. The optimal number of features (N parameter) for this method was determined to be between 20 and 30. This approach achieved an accuracy of 92% in data point classification. The layers utilized in this approach included the primary color bands (red, green, blue, nir) and their corresponding local binary pattern (LBP) transformations (red_lbp, green_lbp, blue_lbp, nir_lbp). For texture features, only the mean and variance were selected. The time required for data preparation of a single tile (250x250) was relatively short at 9 seconds.

Approach 2: Emphasis on Texture Features: the second approach focused more on texture features, reducing the variety of layers used. Here,

the optimal number of features (N parameter) was determined to be between 20 and 30. This approach yielded a higher classification accuracy of 95%. The layers employed were limited to blue and near-infrared. A more extensive set of texture features was used, including 13 Haralick features and skewness. Consequently, the time required for data preparation of a single tile (250x250) was longer, amounting to 30 seconds.

The difference in time between these two approaches comes from the difference in computing layer-based texture features versus data-point-based texture features. Layer-based texture features can be computed once for the entire project area before any data points are selected. In contrast, data-point-based texture features must be calculated for each datapoint, involving each $n \times n$ sliding window. This fundamental difference in computational requirements contributes to the significant disparity in processing times between the two approaches. Given the substantial amount of data that needed to be segmented, we were willing to compromise accuracy slightly in favor of faster processing, leading to the selection of Approach 1 as the final method.

A comprehensive comparison of the two approaches and results of experiments not utilizing texture features can be found in Appendix C

5.1.2 *N* parameter exploration

N parameter indicates the size of the area, expressed in pixels, surrounding each pixel - the context of a pixel for which the texture is measured for. With the resolution of the image being 30 cm, the smallest N value considered is of size 3 by 3 meters (10 by 10 pixels).

The results indicated that larger *N* values corresponded to higher accuracy. Specifically, for a four-class classification task, the accuracies were as follows: 73%, 85%, 92%, 93.5%, 94%, 95%, and 96% for *N* values of 10, 15, 20, 25, 30, 35, and 40, respectively (see Figure 5.1).

However, upon visual inspection of the results, a downside of large window sizes became apparent, which was not captured by the accuracy metric. Since the samples were manually labeled, it was much easier to label

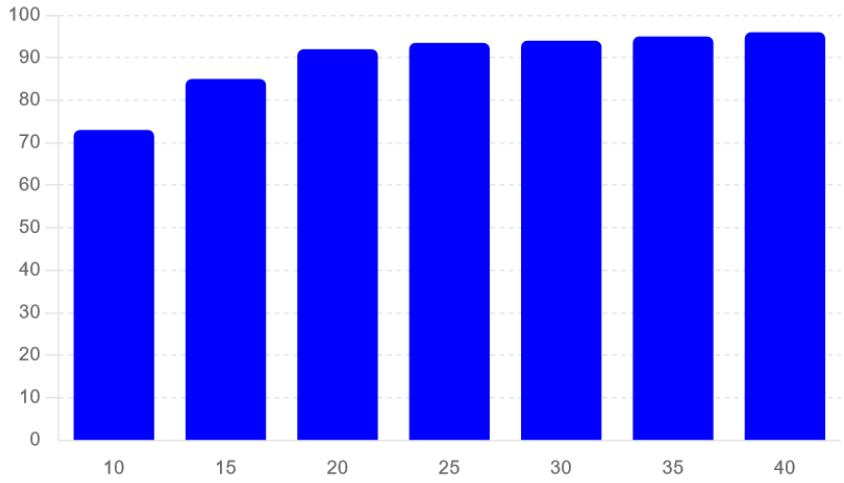


Figure 5.1: 4-class classification accuracy with different size of N parameter

Figure 5.2: Manual annotation of the broadleaves class (orange). It can be noted that the annotation is rough, not aligning with the actual tree boundary, which means that edge pixels of the class were not included in the training dataset

areas within a single vegetation type rather than at the edges where different types meet (see Figure 5.2). Consequently, the majority of labels were placed away from the borders of vegetation groups. The algorithm tended to confuse classes for pixels located at the boundaries between different vegetation clusters, an issue that worsened as N increased (see Figure 5.3).

Therefore, despite the higher accuracy associated with larger N values, we opted for $N = 20$. This value provided a good balance between high accuracy and minimizing misclassifications at the edges of vegetation clusters.

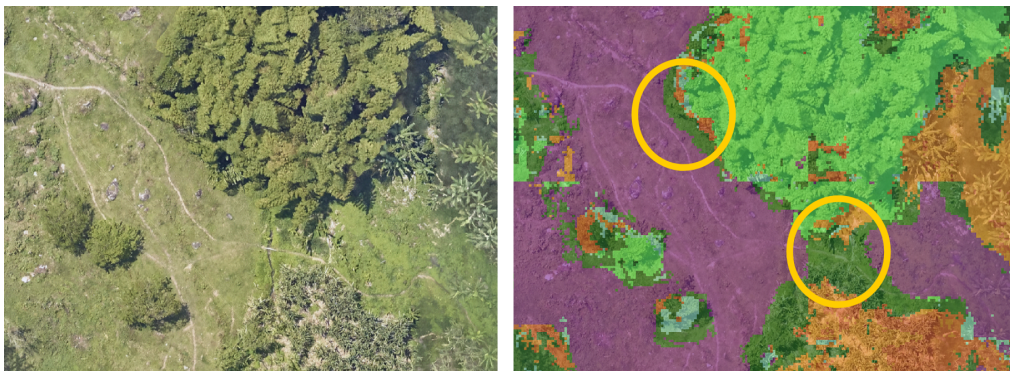


Figure 5.3: Resulting segmentation with edge pixels misclassification - dark green indicates broadleaved vegetation, not present in the area.

5.1.3 Random Forest hyperparameters finetuning

In our study, the optimization of the Random Forest model's performance was critical for enhancing the accuracy of vegetation segmentation. To identify the most effective set of hyperparameters, we employed a systematic approach using GridSearchCV, a method that iteratively explores a range of potential settings to determine the optimal combination for model training. This process focuses on adjusting several key parameters that influence the behavior and capability of the Random Forest algorithm. The parameters optimized included the number of trees in the forest (n_estimators), the maximum depth of each tree (max_depth), the minimum number of samples required to split an internal node (min_samples_split), the minimum number of samples required to be at a leaf node (min_samples_leaf), and whether bootstrap samples are used when building trees (bootstrap). The best hyperparameters selected through this exhaustive search, as listed above, are pivotal for ensuring that the model achieves a high level of predictive accuracy while avoiding overfitting, thereby making it robust in handling complex data structures typical in remote sensing datasets.

The best hyperparameters selected with GridSearch are:

- n_estimators: 300;
- max_depth: 20;
- min_sample_split: 2;
- min_sample_leaf:1
- bootstrap: true.

5.1.4 Pleiades Neo (Kenya) Species Segmentation Results

The classification results for the Pleiades Neo dataset are as follows: the accuracy of data point classification was 94% for the training set and 92% for the test set. The most successfully recognized class was Crops (98.45% accuracy)

The one-vs-all accuracy, f1, and f2 score for each class are detailed in

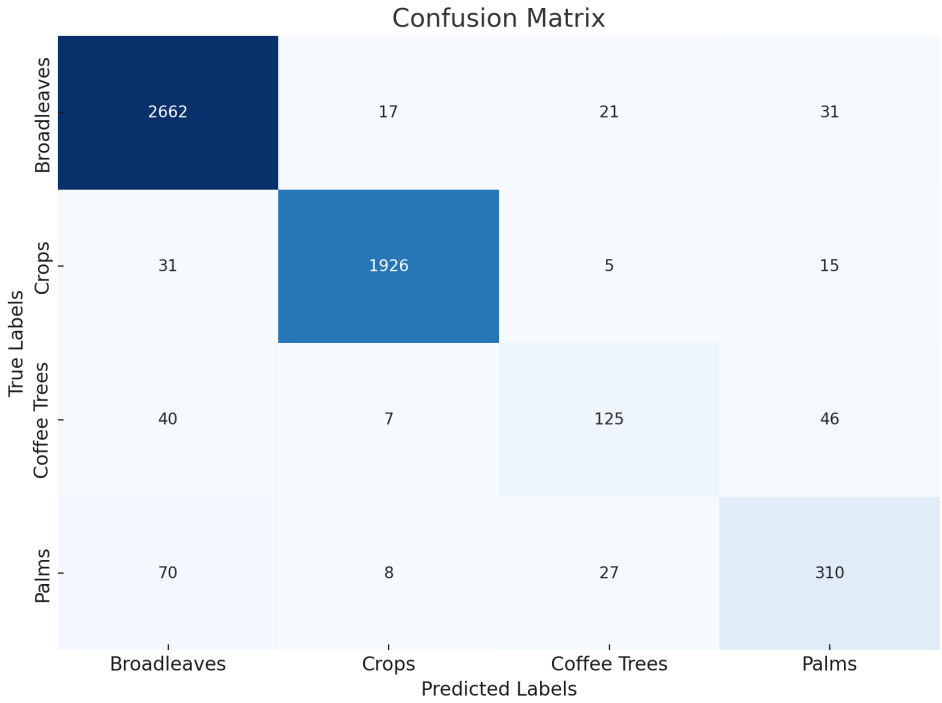


Figure 5.4: Confusion Matrix for Pleiades Neo Classification

the table 5.2. The classification performance metrics for the Pleiades Neo dataset indicate varying levels of effectiveness across different classes. The accuracy of the data point classification is notably high for all classes, with Broadleaves achieving 96.07%, Crops 98.45%, Coffee Trees 97.27%, and Palms 96.31%.

When examining the F1 scores, which balance precision and recall, Crops exhibit the highest performance at 0.9789, followed closely by Broadleaves at 0.9621. Palms have a moderate F1 score of 0.7589, while Coffee Trees have a significantly lower F1 score of 0.6313, suggesting challenges in correctly classifying this category.

The F2 scores, which emphasize recall more than precision, show a similar trend. Broadleaves and Crops maintain high F2 scores of 0.9696 and 0.9761, respectively. Palms have a slightly lower F2 score of 0.7517, and Coffee Trees have the lowest F2 score of 0.5952, further highlighting difficulties in capturing all relevant instances of Coffee Trees.

Overall, while the classification model performs well for Broadleaves and Crops, improvements are needed to enhance the classification of Coffee

Trees and Palms.

Class	Accuracy (%)	F1 Score	F2 Score
Broadleaves	96.07	0.9621	0.9696
Crops	98.45	0.9789	0.9761
Coffee Trees	97.27	0.6313	0.5952
Palms	96.31	0.7589	0.7517

Table 5.1: Classification Performance Metrics

The segmentation results for selected examples are illustrated in the figures 5.5,5.10,5.11,5.12.

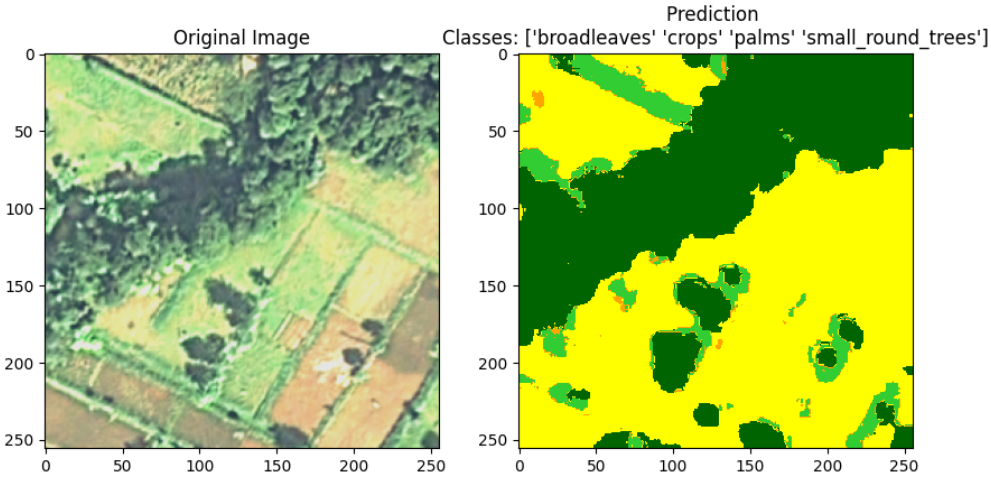


Figure 5.5: Segmentation Result 1

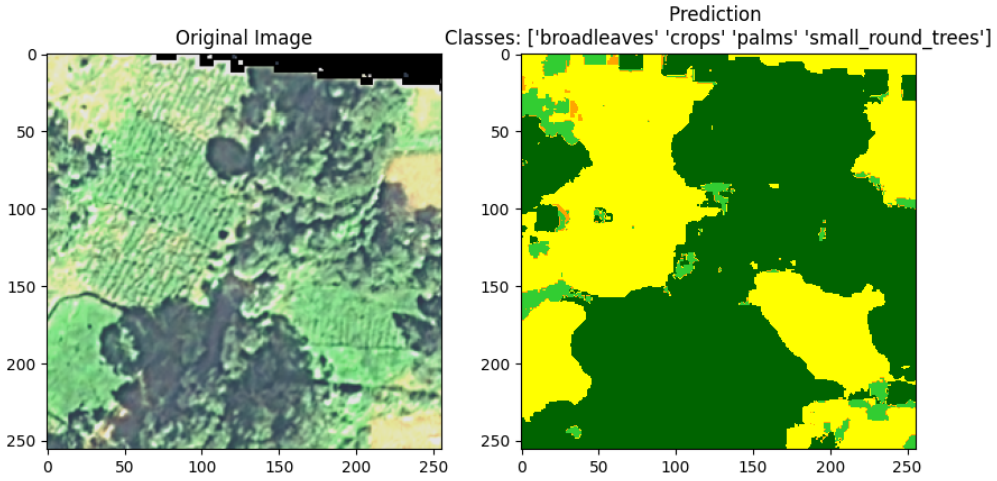


Figure 5.6: Segmentation Result 2

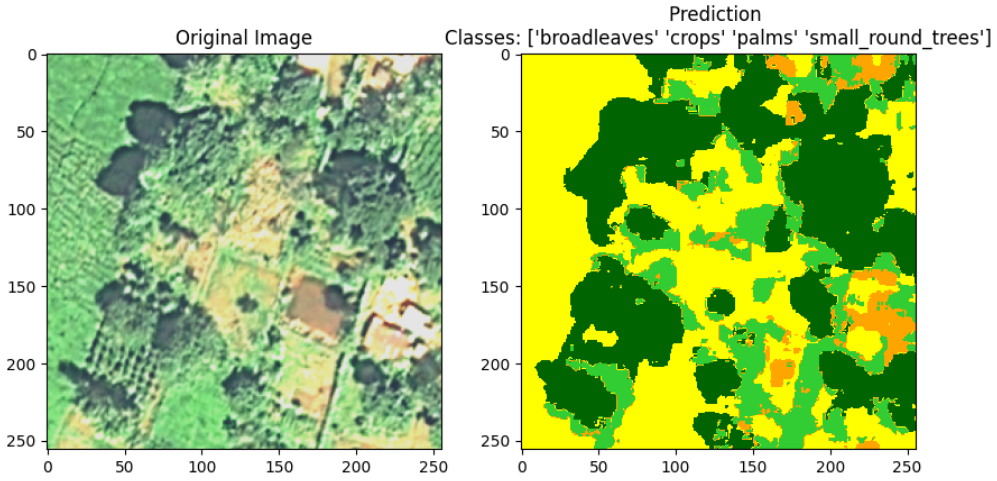


Figure 5.7: Segmentation Result 3

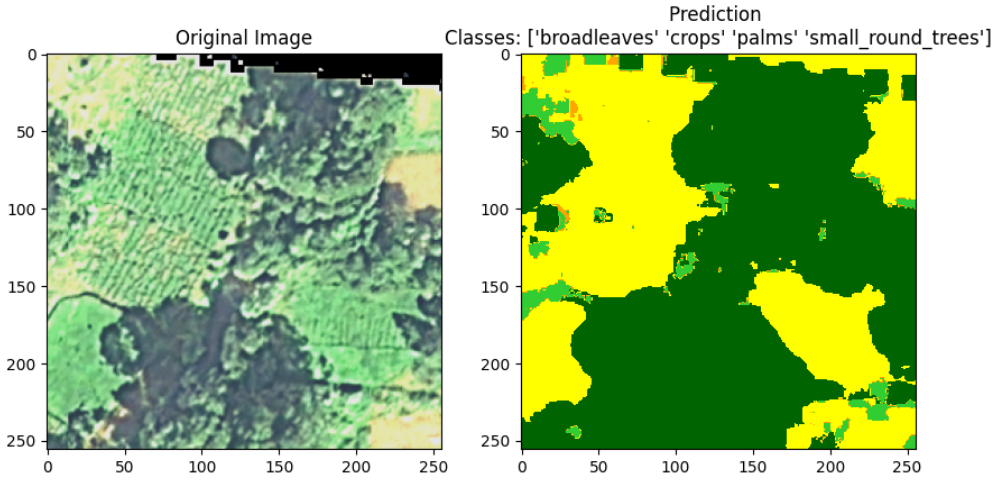


Figure 5.8: Segmentation Result 4

5.1.5 Orthophoto (Colombia) Species Segmentation Results

In this study, we aimed to test the hypothesis that species-specific segmentation can improve biomass estimation accuracy. Due to the absence of biomass data for the Kenya site, we focused our efforts on segmenting orthophotos from the Colombia region. Six species were selected for this analysis: broadleaves, coniferous trees, palms, banana trees, coffee, and grassland.

Despite the higher spatial resolution of the orthophotos (3 cm), the lack of a near-infrared band presented a challenge. To adhere to our previously established methodology, we calculated local binary patterns (LBP) for all available bands, as well as the red-to-green ratio. This resulted in eight layers (red, green, blue, red by green, and the LBPs of each) being used for the segmentation process. Mean and Variance were used as DP-based texture features, with N settled at 20.

Initially, when the images were used at their original resolution, the segmentation accuracy averaged 64%. To address this, the images were downsampled to a 30 cm resolution, in order to more closely follow the Pleiades Neo scenario. This significantly improved the segmentation accuracy to an average of 86%.

Class	**Accuracy (%)**	**F1 Score**	**F2 Score**
Broadleaved	69.89	0.7400	0.7148
Grassland	86.98	0.8621	0.8667
Coniferous	94.59	0.9387	0.9430
Coffee	89.07	0.8827	0.8875
Palms	84.84	0.8278	0.8401
Banana trees	84.14	0.8361	0.8393

Table 5.2: Classification Performance Metrics

The classification metrics indicate varying performance across different land cover types. The Coniferous class achieved the highest accuracy at 94.59%, with corresponding high F1 and F2 scores (0.9387 and 0.9430, respectively), indicating that it was the easiest for the model to detect with balanced precision and recall.

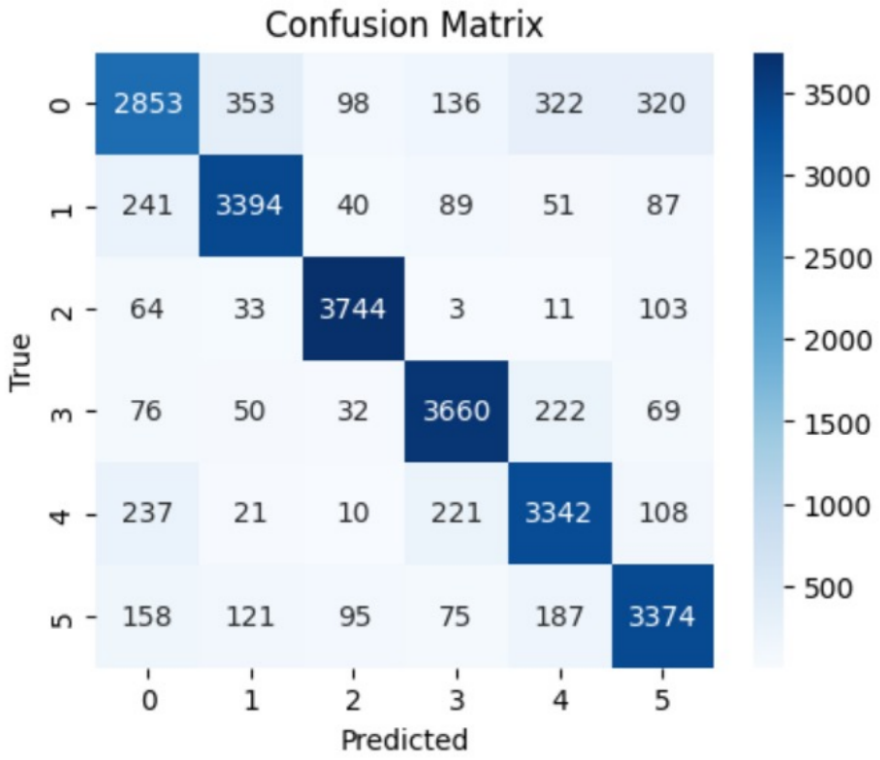


Figure 5.9: Confusion Matrix for Orthophoto Classification. Classes: 0 - Broadleaved; 1 - Grassland; 2 - Coniferous; 3 - Coffee; 4 - Palms; 5 - Banana Trees.

In contrast, the Broadleaves class had the lowest accuracy at 69.89%, along with the lowest F1 and F2 scores (0.7400 and 0.7148). This suggests that Broadleaves areas were the most challenging for the model to classify accurately, indicating significant room for improvement.

Grassland and Coffee classes performed well, with accuracies of 86.98% and 89.07%, respectively. The F1 and F2 scores for Grassland (0.8621 and 0.8667) and Coffee (0.8827 and 0.8875) indicate robust performance, though some enhancements could further improve recall, particularly for Grassland.

Palms and Banana trees classes had similar accuracies (84.84% and 84.14%) and comparable F1 and F2 scores. While these results are satisfactory, targeted improvements could help achieve a better balance in precision and recall for these classes. Overall, while the model shows strong performance for certain classes, others like Broadleaved require additional focus for improvement.

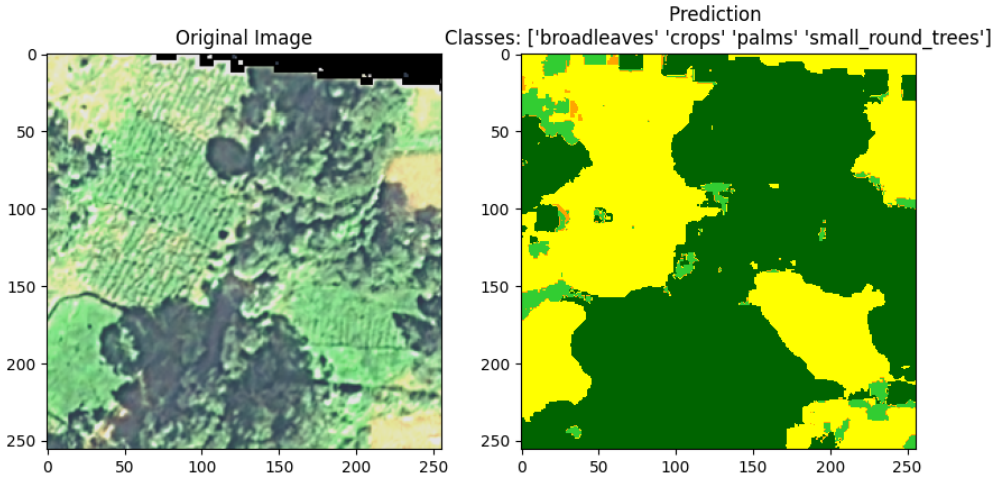


Figure 5.10: Segmentation Result 2

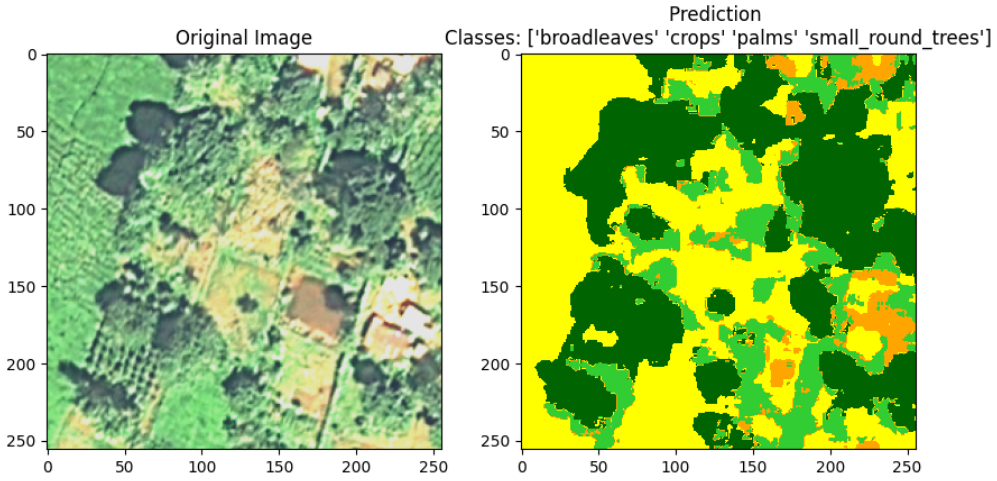


Figure 5.11: Segmentation Result 3

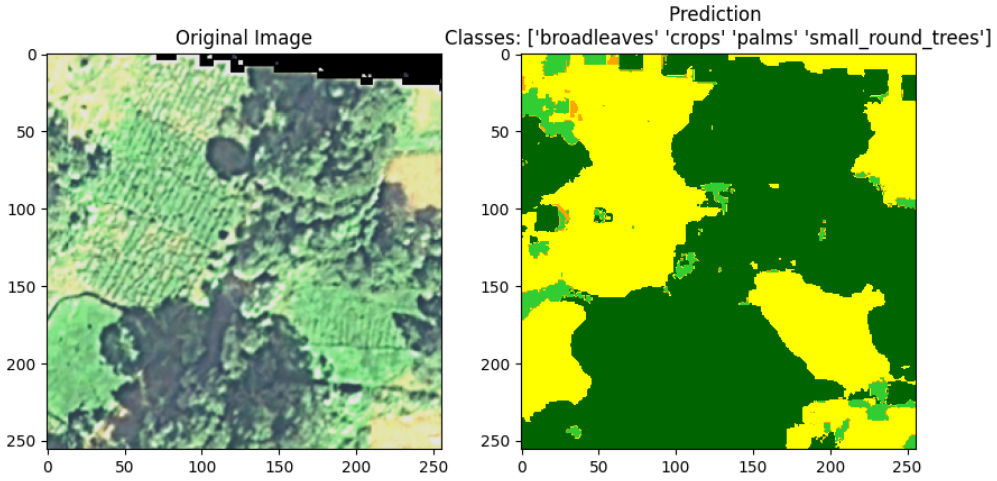


Figure 5.12: Segmentation Result 4

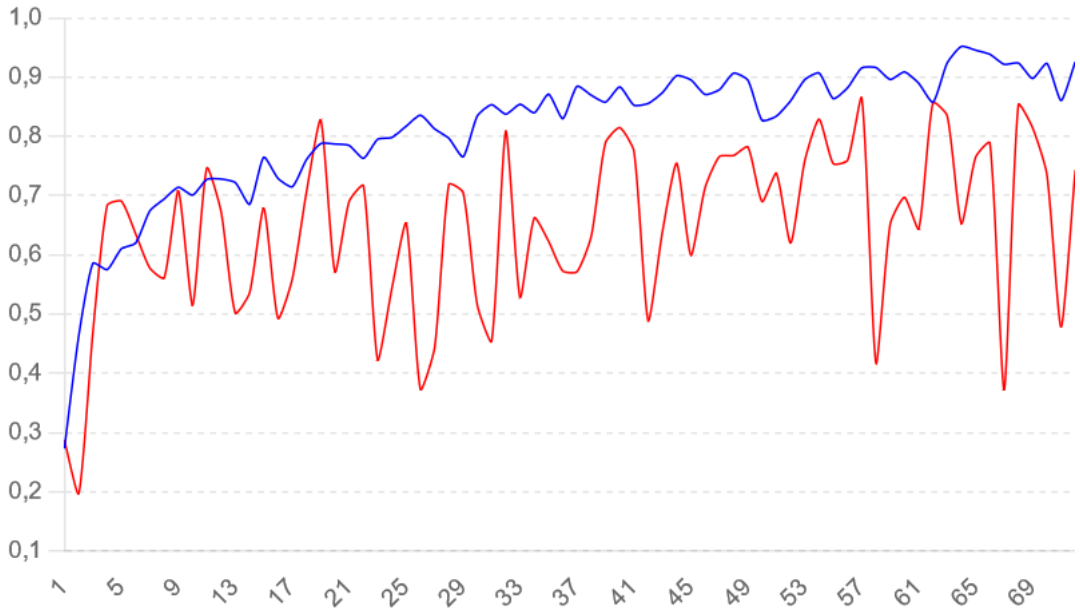


Figure 5.13: Training and testing accuracy over 73 epochs

5.1.6 Unet Species Segmentation

One of the assumptions we are testing is that texture-based segmentation combined with machine learning is superior to well-established deep learning segmentation methods, such as Unet, in scenarios where we have a limited amount of labels. This hypothesis was tested by running Unet on the entire segmented dataset that was available.

To challenge this hypothesis, the available manually annotated dataset was divided into training (70%), testing (15%), and validation (15%) sets. Data augmentation techniques, including random rotation and mirroring, were applied to the training set. The training was conducted with a batch size of 16 and utilized early stopping to prevent overfitting.

The results after 72 epochs indicated that the segmentation accuracy was 0.92 on the training set and 0.744 on the test set. However, upon inspecting the plot visualizing the test accuracy, we can conclude that the performance is highly unstable. This indicates overfitting of the algorithm (see figure 5.13).

5.2 Biomass Estimation

To test the hypothesis that information about vegetation class enhances biomass estimation, the orthophoto from Colombia was segmented using the algorithm outlined in subsection 5.1.5. For simplicity, the XGBoost algorithm was selected for biomass estimation, rather than more complex methods such as CNN. The experiment utilized a dataset of 200 000 data points, which were randomly selected from the image. The data points included features such as the values of the red, green, and blue bands, with an additional information of vegetation class assigned through the segmentation process.

The biomass estimation was conducted using XGBoost was employed for the estimation.

The findings, detailed in Table 5.3 and visualized on Figure 5.14, demonstrate that including the vegetation classification layer substantially improves biomass estimation. The feature importance analysis conducted using the XGBoost algorithm revealed that the segmentation layer was the most pivotal feature, with an importance score of 0.939 when included. Additionally, excluding this layer from the model resulted in a notable escalation in the error rate, increasing from 0.60 to 0.71.

Feature/MSE	Without Segmentation	With Segmentation
Red	0.4897	0.0216
Green	0.2901	0.0192
Blue	0.2202	0.0201
Segmentation	-	0.9391
MSE	0.7100	0.5960

Table 5.3: Comparison of Feature Importances and MSE With and Without Vegetation Class Feature

A detailed comparison of the distribution of actual versus predicted biomass values for each class is illustrated in Figure 5.16.

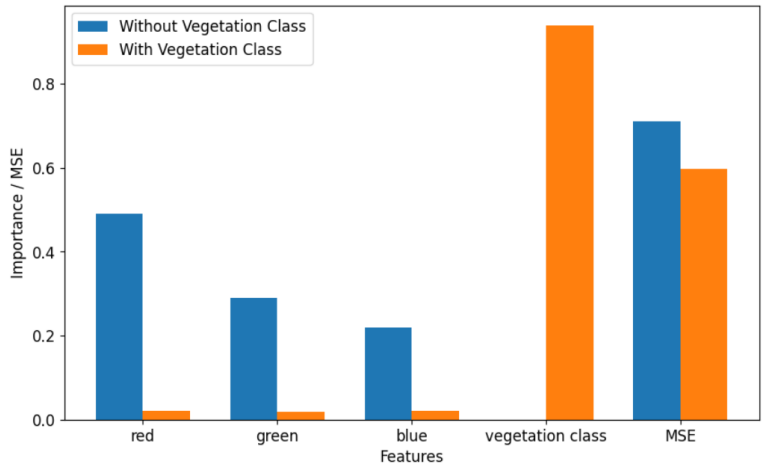


Figure 5.14: Comparison of Feature Importances and MSE With and Without Vegetation Class Feature

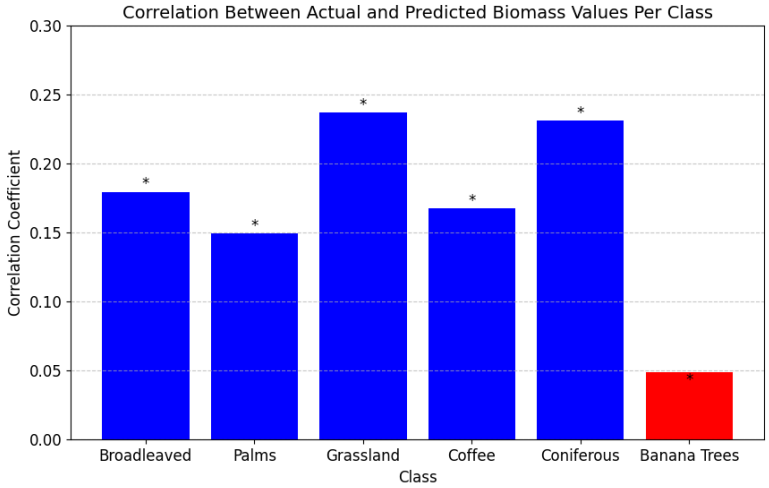


Figure 5.15: Correlation between classes and biomass values. Blue bars represent significant correlations, while red bars represent non-significant ones.

Interestingly, there is a weak but significant correlation between biomass values and four out of the five vegetation classes, with the coniferous class showing the highest correlation at 0.23. These findings are visualized in Figure 5.15

The observed low correlation and high error prompted us to investigate the effect of removing outliers from the dataset, suspecting potential misclassification of some pixels. We excluded pixels with biomass values below the 5th percentile or above the 95th percentile within each class's biomass distribution. The visualization of predictions after this modification is shown in Figure 5.17

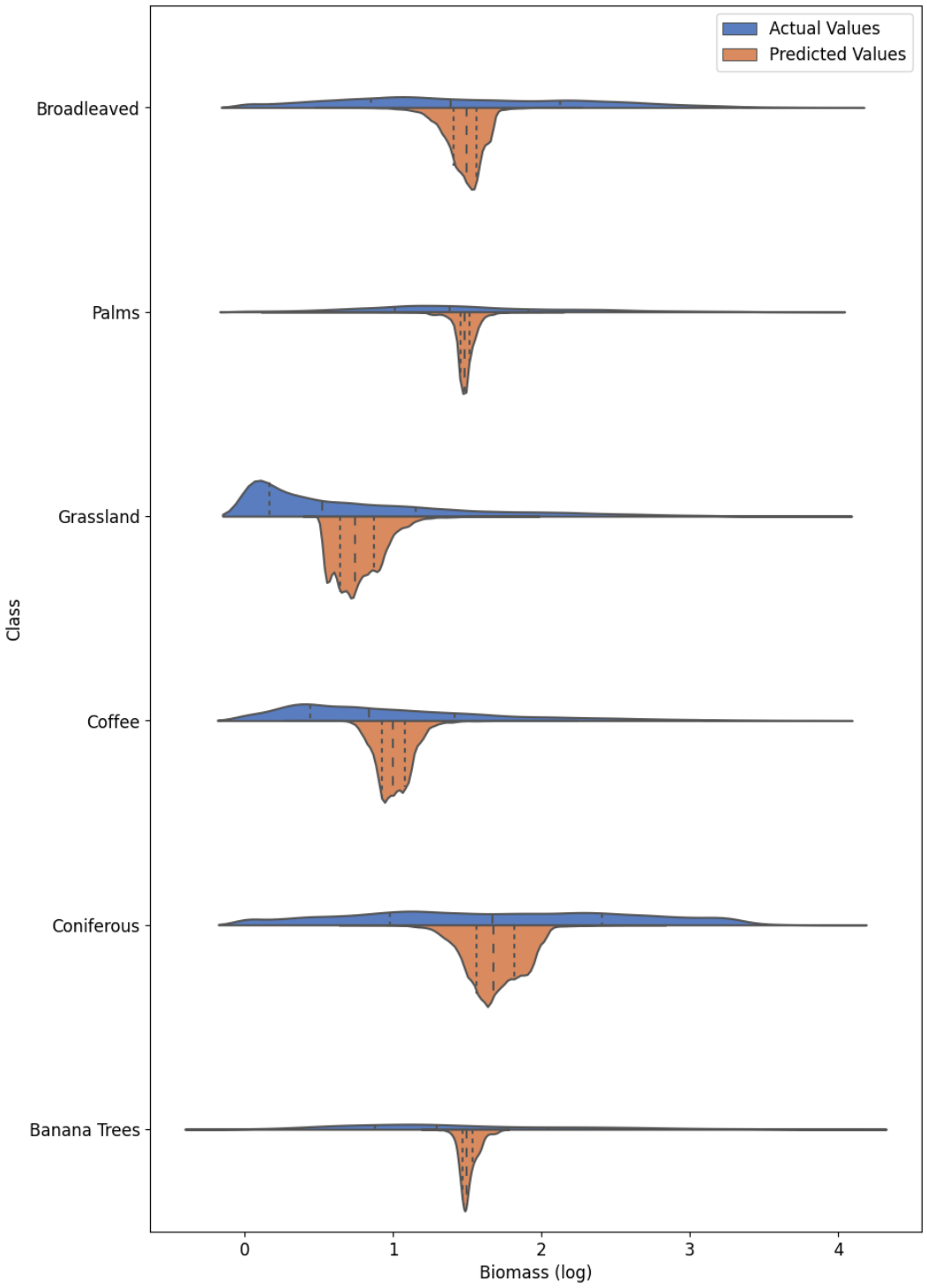


Figure 5.16: Visualization of actual vs predicted biomass value per class.

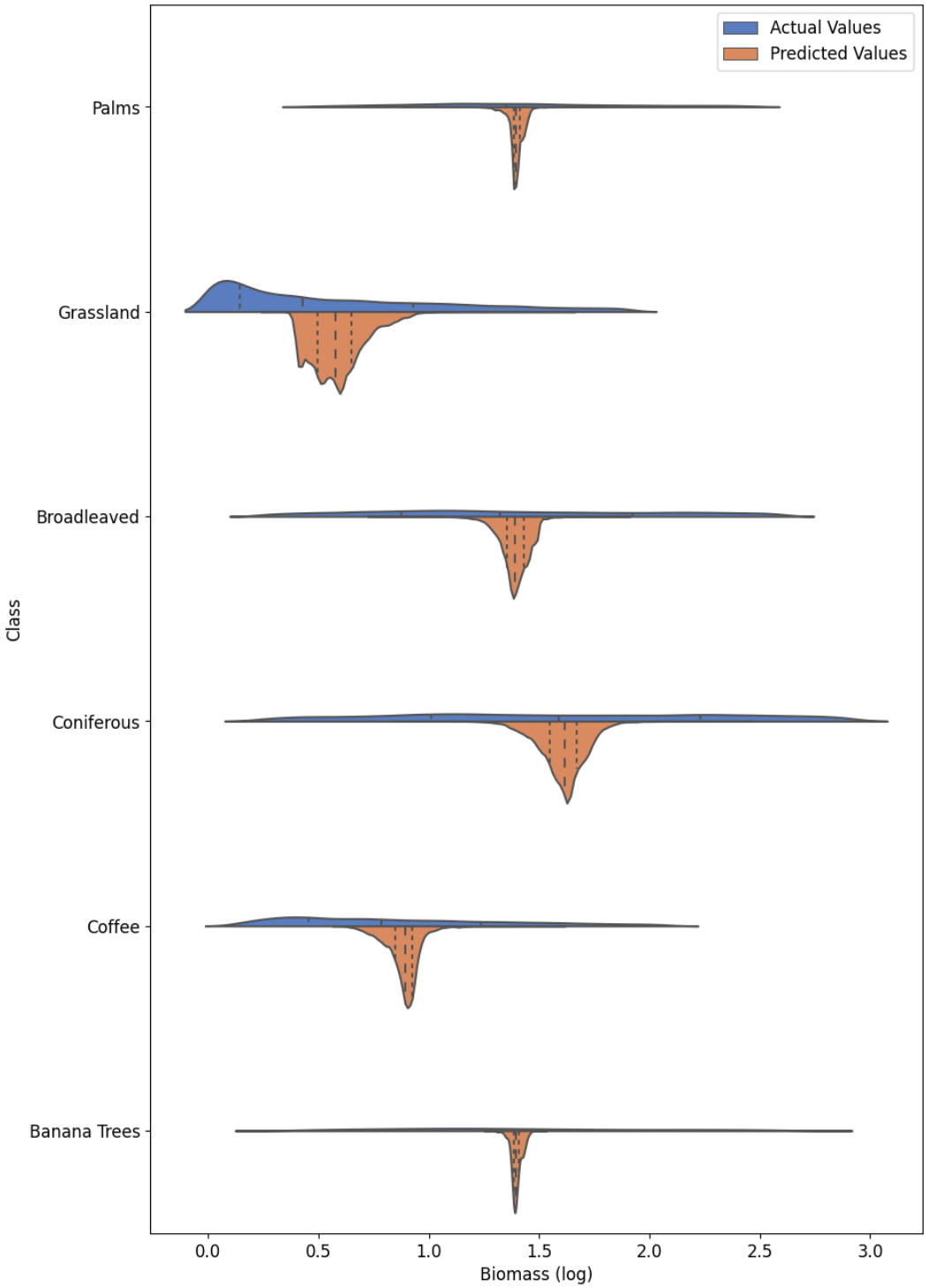


Figure 5.17: Visualization of actual vs predicted biomass value per class after excluding the outliers

6. Discussion

In this project, we aimed to validate the significance of texture features in vegetation segmentation, demonstrate their advantages over deep learning algorithms, and determine whether prior vegetation segmentation aids in biomass estimation. This chapter reviews the results obtained, explores their applicability in broader contexts, compares them with findings from the literature, and discusses the method's limitations.

We will structure this chapter into two sections. The first section will focus on the task of vegetation segmentation, while the second will concentrate on biomass estimation.

The discussion reflects on the research questions and objectives, as well as the methodology and results.

6.1 Vegetation Segmentation

In this project, the texture-based vegetation segmentation methodology was developed and later evaluated, on two different sites, data sources, and sets of classes. This allows us for a comprehensive and in-depth review of the accuracy and limitations of the method.

6.1.1 Texture Features

The core aspect of our methodology is the application of a sliding window technique combined with the calculation of texture features to develop a dataset. This dataset is subsequently employed to train a Random Forest model for vegetation segmentation tasks.

To facilitate this, several texture functions were computed (see **??**, **??**, **??**, **??**) and incorporated into an iterative training loop with Random Forest to identify the most essential subset of features.

The feature selection process yielded some interesting insights. Notably, despite the scarcity of studies incorporating texture features in remote sensing data, those that do typically rely heavily on the Gray Level Co-occurrence Matrix (GLCM) or a limited array of features, often concluding the superiority of GLCM features [53]. Contrary to these findings, our research indicates that GLCM features were among the least significant in our study. This may be because of the relative simplicity of GLCM. Additionally, the existing Python implementations of GLCM are significantly slower compared to other features utilized, which might have deterred their broader adoption in research.

In our study, we introduce a significant distinction in texture feature computation that is often overlooked in remote sensing literature. Texture features can be categorized into two types: layer-based features and window-based features. Layer-based features, such as gradients, are computed once for an entire large image, producing an output array of the same shape as the input array. On the other hand, window-based features, like Haralick features, need to be computed separately for each sliding window, returning a set of float values.

This distinction is particularly crucial in the context of remote sensing, especially with super-high-resolution imagery. These images are typically very large, and the ability to precompute layer-based features significantly enhances the speed and efficiency of data processing. In our research, we found that Local Binary Patterns (LBP) were the most influential texture features. This finding is consistent with previous successes of LBP in texture classification tasks [99], but it is noteworthy to observe its substantial potential in remote sensing applications.

The capability to efficiently compute and utilize potent texture features significantly enhances the accuracy and performance of vegetation segmentation and other remote sensing tasks. This assertion is corroborated by our research findings. We achieved over 90% segmentation accuracy in a 4-class classification task. In contrast, other studies employing Random Forest algorithms, such as those cited in [88], reported only 74% accuracy for pine

classification and 79% for spruce. These comparisons highlight the substantial improvements our methodologies offer over traditional approaches, emphasizing the effectiveness of our advanced feature utilization in handling complex remote sensing data.

6.1.2 Comparison to Traditional Deep Learning Approaches

Remote sensing analysis frequently has to deal with the scarcity of labeled data, compounded by the laborious and challenging nature of the labeling process. Additionally, methods developed for one ecoregion often do not seamlessly transfer to another, necessitating adaptable, easily retrainable approaches that require minimal training examples.

Given these constraints, deep learning methods, which typically demand large datasets, may not be ideally suited for such tasks. To validate this, we trained a UNet model on the same dataset used for our texture-based segmentation to compare the effectiveness of the two approaches.

As anticipated, the limited number of training examples led to overfitting in the UNet model. This outcome underscores the advantages of precomputing texture information. Theoretically, UNet's convolutional layers operate similarly to a manual sliding window approach, but they lack predetermined filters and require substantial training data to develop these effectively. Our manual approach, informed by domain-specific insights into desirable filter characteristics, circumvents this need.

While we hypothesize that UNet could potentially surpass our method under conditions of abundant data, the reality of our situation—where preparing the dataset alone consumed an entire workweek—highlights the impracticality of such data-intensive methods in our current context.

6.1.3 Evaluation & Limitations

Misclassifications on Class Edges

Our segmentation method exhibits several limitations that offer avenues for future exploration. A significant challenge arises in classifying pixels lo-

cated at the boundaries of different vegetation classes or where these classes overlap. The training accuracy metric, which assigns a single class value to each pixel, coupled with less-than-ideal annotations (it is easier to annotate homogeneous areas within a class than at transitional edges), leads to imperfect segmentation at these critical junctions.

There are several potential solutions to this issue. Most directly, incorporating more complex boundary areas into the training and testing datasets could improve the model's ability to accurately classify these regions.

However, this approach has limitations, primarily due to the feature extraction method employed. We hypothesize that the misclassification of boundary pixels stems from their feature values, derived from a 20x20 pixel window. This makes the feature context of a pixel at the edge of two classes similar to that of a nearby pixel that, while close to the border, belongs to a distinct class.

A promising direction for future research could involve adopting superpixel segmentation techniques, similar to those used in [88], to define context windows. This method would allow for the extraction of texture features within these dynamically shaped clusters rather than relying on fixed, square windows. Implementing such an approach would necessitate further investigation into its effectiveness and feasibility.

Overlapping classes

The algorithm demonstrates particular difficulty in areas where two vegetation classes coexist, as notably observed at the Colombia site where palms overlap with coffee trees. In these cases, the algorithm tends to prioritize palms, erroneously creating a buffer around this class rather than accurately classifying the overlapping pixels as coffee. This issue stems from the inherent behavior of the Random Forest model, which classifies based on the majority vote from its decision trees.

One potential solution to this problem involves delving deeper into the classifier's output. Instead of solely considering the most frequent class, extracting the second most popular class or even a distribution of votes

among several top classes could provide more nuanced insights. Leveraging ecological knowledge of the region, these refined predictions could be intelligently combined to distinguish more accurately between overlapping classes. This approach not only promises to enhance classification accuracy in complex mixed-class areas but also aligns closely with the ecological realities observed in field studies, potentially leading to significant improvements in the practical application of remote sensing data in vegetation analysis.

Comparison between Pleiades Neo and Orthophoto segmentation

Pleiades Neo and Orthophotos are distinct data sources that offer varying capabilities for vegetation segmentation. Pleiades Neo provides a 30 cm resolution along with a near-infrared band, whereas Orthophotos offers a higher resolution of 3 cm but lacks the near-infrared spectrum. This difference fundamentally affects their utility in segmentation tasks.

A key observation from our study was the lower segmentation accuracy with Orthophotos compared to Pleiades Neo. This discrepancy is likely due to the absence of the near-infrared band in Orthophotos, which plays a crucial role in vegetation monitoring and classification. Additionally, artificially downsampling the Orthophoto to 30 cm to match Pleiades Neo's resolution may have introduced further inaccuracies.

Notably, the accuracy of class segmentation varied significantly between the two data sources. For instance, in the Colombian site, the detection accuracy for broadleaves was only 69% with the Orthophoto, the lowest among all classes, while Pleiades Neo achieved a 96% accuracy for the same class. Despite broadleaves being a visually distinct class, this stark contrast in performance suggests underlying issues. We hypothesize two main factors contributing to this discrepancy: the diversity within the broadleaves category, which includes a variety of tree species, potentially confusing the algorithm; and the similarity in canopy texture between some broadleaves and other species, such as coffee trees, which can lead to misclassifications. An example of this is depicted in Figure 6.1, where parts of the broadleaf canopy mistakenly resemble the round-shaped canopy of coffee trees.

Figure 6.1: Example of broadleaves misclassification. Dark green indicates broadleaves, light blue - coffee trees.

To address these issues, integrating the use of multiple window sizes in the segmentation process could be beneficial. This approach would allow the capture of contextual variations at different scales, potentially improving the accuracy of class differentiation and overall segmentation performance.

Time of Processing

A significant challenge encountered in this research was the extensive time required to compute texture features, which notably impacted both the results and the pace of the project. For example, to perform segmentation on the Pleiades image using a comprehensive set of texture features, the algorithm required a staggering 36 hours to complete. Even when reducing the number of features, the process remained time-intensive, not achieving the speed efficiency typically seen with neural networks.

One potential solution to enhance processing speed is to leverage GPU acceleration. Due to time constraints, this was not implemented during the current project. However, future work could benefit greatly from this approach. By utilizing pre-computed local binary pattern (LBP) layers and restricting the feature set to mean and variance calculations, it becomes feasible to adapt the entire framework for GPU execution. Tools such as TensorFlow or PyTorch, which are designed for high-performance computational parallelism on GPUs, could dramatically reduce processing time. This adjustment would not only accelerate the computation but also bring the pro-

cessing time more in line with that of advanced neural network methods, enabling more efficient and scalable remote sensing analytics.

6.2 Biomass Estimation

Due to the unavailability of biomass data for the site in Kenya, the biomass estimation component of this project was exclusively carried out at the site in Colombia. The primary objective of this experiment was to determine whether incorporating a texture-derived segmentation layer improves biomass estimation compared to using only RGB channels.

6.2.1 Vegetation Segmentation Influence

Our study appears to confirm the hypothesis that vegetation segmentation positively influences biomass estimation. As detailed in 5.2, the inclusion of the segmentation layer significantly reduced the Mean Squared Error (MSE) in the XGBoost model. Additionally, feature importance analysis shows that the segmentation layer is heavily utilized, with a feature importance (FI) score of 0.93.

Despite the efforts, the estimation error remains considerably high. An MSE of 0.59 is unsatisfactory, indicating the need for a more thorough analysis of the process.

The analysis of predicted versus actual biomass values per class (5.16) reveals that XGBoost assigned a very narrow margin of values per class, often close to the median value. This pattern, combined with the high importance of the vegetation class, suggests that there may be another feature, not available in our dataset, that could further differentiate pixels belonging to the same class in a more nuanced way.

The findings about the low correlation between biomass values and the assigned class are quite surprising. One might expect a high correlation in some cases, such as grassland being expected to have a very low biomass value. However, it is surprising to see that each class of vegetation covers the entire range of biomass values. This indicates possible misclassifica-

Figure 6.2: An example of missclassification: broadleaved vegetation was classified as grassland

tion, as it is very unlikely that grassland would have high biomass values. Upon closer inspection of the confusion matrices and segmented images, it becomes evident that a common type of misclassification is assigning the grassland class to a specific species of broadleaves with a flat canopy structure and fine leaves (see Figure 6.2). Coincidentally, this type of vegetation has a very high biomass.

This analysis highlights the need for additional features or refinement in the classification model to better capture the nuances in vegetation types and their corresponding biomass values. By incorporating more detailed features or improving the classification accuracy, the model can provide more precise biomass predictions, particularly for classes that are currently being misclassified.

In an effort to identify and remove potentially misclassified data points, the dataset was modified by excluding pixels whose biomass values fell below the 5th percentile or above the 95th percentile for their respective classes. This approach resulted in a significantly higher mean squared error (MSE) of 0.36. Despite a better result, this method remains suboptimal as it does not allow for distinguishing between outliers caused by segmentation errors and those resulting from natural variations in biomass values.

6.2.2 Limitations

Our methodology faces several limitations. Firstly, the heavy reliance on segmentation means that any misclassification of pixels significantly impacts accuracy. This dependency is particularly highlighted in the interpretation of our results.

Secondly, our inability to test on Pleiades Neo data due to the lack of available biomass data for that region represents a significant shortfall. The absence of the near-infrared band in orthophoto data, which is crucial for vegetation assessment in many studies, limits the effectiveness of our analysis. The Normalized Difference Vegetation Index (NDVI) alone, which can correlate with biomass values up to 0.6 [100], could not be utilized. This raises an important question: would the importance of segmentation diminish if the near-infrared band were included in the analysis? Despite these challenges, we speculate that segmentation would still play a critical role. Studies support this, noting instances where the NDVI of grasslands can be comparable to that of forests [101], despite forests having significantly higher above-ground biomass. Thus, segmentation and near-infrared data could potentially complement each other in biomass estimation.

Thirdly, the necessity to artificially upsample the biomass data introduced additional errors. Notably, some values became negative after bicubic interpolation and required clipping to zero, while others significantly exceeded their original values. These alterations affected the distribution of biomass values, potentially degrading the quality of the ground truth data and hindering the algorithm's ability to learn accurately.

7. Conclusions

The increasing utilization of the carbon stock market emphasizes the importance of developing methodologies for rapid and cost-efficient estimation of above-ground biomass. Due to the limitations inherent in remote sensing data, it is imperative to focus on methodologies that require minimal data and are specifically developed for vegetation detection in remote-sensing scenarios, distinguishing them from classical computer vision techniques.

This study aimed to develop a lean segmentation methodology for analyzing vegetation through high-resolution remote sensing data and to assess whether such segmentation could improve biomass estimation. We devised a method that leverages texture features for species segmentation, effective even with minimal annotated datasets (demonstrated with as few as 40,000 pixels, equivalent to annotating a 200 by 200 pixel image). This method achieved an accuracy of 86% in a six-class classification using RGB orthophotos and up to 95% in a four-class classification using RGB plus near-infrared imagery from Pleiades Neo, exceeding our initial expectations. We validated our hypothesis that texture information, coupled with a Random Forest classifier, offers greater efficacy for this task with limited data compared to traditional segmentation approaches such as U-Net. Furthermore, the local binary pattern was identified as the most effective texture function, surpassing the performance of the commonly utilized GLCM in terms of both efficiency and speed.

The study also demonstrated that incorporating vegetation classification could improve biomass estimation. Although the results did not meet all expectations, they underscored the significant role of this additional feature, affirming our second hypothesis.

Future Work

The potential of this methodology merits further exploration. Future

studies should extend the testing of this segmentation approach to a wider array of vegetation classes and data sources, and assess its effectiveness across different resolutions and ecosystems; our evaluations were confined to a 30cm resolution. Additionally, re-engineering the segmentation method to utilize GPU processing could enhance computational speed. More rigorous research is also needed in biomass prediction, as this study had to rely on an artificially upsampled biomass dataset, introducing potential inaccuracies. Integrating both texture features and the near-infrared band for comprehensive biomass assessment should be considered in subsequent studies.

Conclusion

In summarizing this study, we acknowledge the advancements made in refining methodologies for vegetation segmentation and biomass estimation using remote sensing data. We managed to confirm and find satisfactory answers to all three of our hypotheses:

- What texture functions can be used in order to achieve at least 90% accuracy on the 4-classes canopy segmentation task?

  The usage of Local Binary Pattern allows for that.

- Can the texture features in combination with random forest exceed the performance of popular segmentation methodologies for small datasets?

  Yes. The method outperformed UNET, which exhibited overfitting.

- Does the use of canopy segmentation input increases the biomass estimation accuracy, in comparison to a traditional method using raw satellite imagery as input? **Yes. The MSE of biomass estimation was significantly lower when vegetation classes were used as a feature.**

The outcomes contribute to the methodological framework of remote sensing applications in ecological research and offer avenues for improving environmental monitoring techniques. The continuation of this research is expected to provide significant contributions to the field of environmental science.

Bibliography

[1] K. Sharma and D. Karamanev, *Investigating the historical correlation between atmospheric carbon dioxide concentration and global temperature change*, 2021. DOI: 10.32006/EEEP.2021.1.0516.

[2] Á. Enríquez-de-Salamanca, "Climate change mitigation in forestry: Paying for carbon stock or for sequestration?" *Atmosphere*, 2022. DOI: 10.3390/atmos13101611.

[3] C. Pan, A. Shrestha, J. L. Innes, *et al.*, "Key challenges and approaches to addressing barriers in forest carbon offset projects," *Journal of Forestry Research*, vol. 33, no. 4, pp. 1109–1122, 2022. DOI: 10.1007/s11676-022-01488-z.

[4] A. Wirz-Justice, M. P. Ivantysyn Michal Vaclavikova Dagmar, and D. R. Wood, "Agroforestry," 2023. DOI: 10.1016/b978-0-12-822974-3.00269-x.

[5] E. R. Lines, M. J. Allen, C. Cabo, *et al.*, *Ai applications in forest monitoring need remote sensing benchmark datasets*, 2022. DOI: 10.1109/BigData55660.2022.10020772.

[6] M. G. Ryan, "The enduring mystery of differences between eddy covariance and biometric measurements for ecosystem respiration and net carbon storage in forests.," *New Phytologist*, 2023. DOI: /10.1111/nph.19105.

[7] A. E. L. Stovall, A. G. Vorster, R. Anderson, and P. H. Evangelista, "Developing nondestructive species-specific tree allometry with terrestrial laser scanning," *Methods in Ecology and Evolution*, 2022. DOI: 10.1111/2041-210X.14027.

[8] M. Ciolfi, F. Chiocchini, G. Russo, L. Spaccino, M. Mattioni, and M. Lauteri, "Sampling strategies in a forest environment for the elaboration of isoscapes," 2018. DOI: 10.30441/SMART-ELAB.V12I0.203.

[9] G. Sun, K. J. Ranson, Z. Guo, Z. Zhang, P. Montesano, and D. Kimes, "Forest biomass mapping from lidar and radar synergies," *Remote Sensing of Environment*, vol. 115, no. 11, pp. 2906–2916, 2011, DESDynI VEG-3D Special Issue, ISSN: 0034-4257. DOI: 10.1016/j.rse.2011.03.021.

[10] W. Yao, P. Krzystek, and M. Heurich, "Tree species classification and estimation of stem volume and dbh based on single tree extraction by exploiting airborne full-waveform lidar data," *Remote Sensing of Environment*, vol. 123, pp. 368–380, 2012, ISSN: 0034-4257. DOI: 10.1016/j.rse.2012.03.027.

[11] G. Patenaude, R. Hill, R. Milne, D. Gaveau, B. Briggs, and T. Dawson, "Quantifying forest above ground carbon content using li-

dar remote sensing," *Remote Sensing of Environment*, vol. 93, no. 3, pp. 368–380, 2004. DOI: 10.1016/j.rse.2004.07.016.

[12] R. C. Shinde and S. S. Durbha, "Deep convolutional compressed sensing-based adaptive 3d reconstruction of sparse lidar data: A case study for forests," *Remote sensing*, 2023. DOI: 10.3390/rs15051394.

[13] S. Popescu, R. Wynne, and R. Nelson, "Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass," *Canadian Journal of Remote Sensing*, vol. 29, pp. 564–577, Oct. 2003. DOI: 10.5589/m03-027.

[14] *How we can better understand our forest ecosystems with laser scanning*, Accessed: 2024-05-25, 2022. [Online]. Available: http://resilience-blog.com/2022/07/06/how-we-can-better-understand-our-forest-ecosystems-with-laser-scanning/.

[15] I. Dowman, K. Jacobsen, G. Konecy, and R. Sandau, *High Resolution Optical Satellite Imagery*. 2012.

[16] B. Shrestha, J. Posom, P. Sirisomboon, and B. P. Shrestha, "Comprehensive assessment of biomass properties for energy usage using near-infrared spectroscopy and spectral multi-preprocessing techniques," *Energies*, 2023. DOI: 10.3390/en16145351.

[17] M. J. P. Sullivan, S. L. Lewis, S. L. Lewis, *et al.*, "Field methods for sampling tree height for tropical forest biomass estimation," *Methods in Ecology and Evolution*, 2018. DOI: 10.1111/2041-210X.12962.

[18] V. Simeonov, "Artificial intelligence tools and vegetation indices combined to estimate aboveground biomass in tropical forests," *Journal of Applied Remote Sensing*, 2023. DOI: 10.1117/1.jrs.17.024512.

[19] M. Dalponte, L. Frizzera, H. O. Ørka, T. Gobakken, E. Næsset, and D. Gianelle, "Predicting stem diameters and aboveground biomass of individual trees using remote sensing data," *Ecological Indicators*, vol. 85, pp. 367–376, 2018. DOI: 10.1016/j.ecolind.2017.10.066.

[20] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," pp. 779–788, 2016. DOI: 10.1109/CVPR.2016.91.

[21] A. Kirillov, E. Mintun, N. Ravi, *et al.*, "Segment anything," Apr. 2023. DOI: 10.48550/arXiv.2304.02643.

[22] A. V. Etten, D. Lindenbaum, and T. M. Bacastow, *Spacenet: A remote sensing dataset and challenge series.* 2018. DOI: 10.48550/arXiv.1807.01232.

[23] L. Li, B. Zhou, Y. Liu, *et al.*, "Reduction in uncertainty in forest aboveground biomass estimation using sentinel-2 images: A case study of pinus densata forests in shangri-la city, china," *Remote sensing*, 2023. DOI: 10.3390/rs15030559.

[24] E. S. Agency, *Sentinel-1*, https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1, Accessed: 2024-07-03, 2024.

[25] E. S. Agency, *Palsar*, https://earth.esa.int/eogateway/instruments/palsar, Accessed: 2024-07-03, 2024.

[26] eoPortal, *Alos-2 (advanced land observing satellite-2) / daichi-2*, https://www.eoportal.org/satellite-missions/alos-2, Accessed: 2024-07-03, 2023.

[27] European Space Agency, *Copernicus sentinel-2 collection 1 msi level-2a (l2a)*, https://sentinels.copernicus.eu/web/sentinel/sentinel-data-access/sentinel-products/sentinel-2-data-products/collection-1-level-2a, Accessed: 2024-07-03, 2021.

[28] U.S. Geological Survey, *Landsat project documents*, https://www.usgs.gov/landsat-missions/landsat-project-documents, Accessed: 2024-07-03, 2024.

[29] U.S. Geological Survey, *Landsat 8 data users handbook*, https://www.usgs.gov/landsat-missions/landsat-8-data-users-handbook, Accessed: 2024-07-03, 2024.

[30] Airbus Defence and Space, *Pleiades neo user guide*, https://wp-cdn.apollomapping.com/web_assets/user_uploads/2021/11/08103301/2021.10_PleiadesNeo_UserGuide-EarlyRelease_20211015.pdf, Accessed: 2024-07-03, 2021.

[31] E. Alvarez-Vanhard, T. Corpetti, and T. Houet, "Uav & satellite synergies for optical remote sensing applications: A literature review," *Science of Remote Sensing*, vol. 3, p. 100 019, 2021, ISSN: 2666-0172. DOI: https://doi.org/10.1016/j.srs.2021.100019.

[32] S. Dittmann, E. Thiessen, and E. Hartung, "Applicability of different non-invasive methods for tree mass estimation: A review," *Forest Ecology and Management*, vol. 398, pp. 208–215, 2017, ISSN: 0378-1127. DOI: https://doi.org/10.1016/j.foreco.2017.05.013.

[33] N. Takhtkeshha, G. Mandlburger, F. Remondino, and J. Hyyppä, "Multispectral light detection and ranging technology and applications: A review," *Sensors*, vol. 24, no. 5, 2024, ISSN: 1424-8220. DOI: 10.3390/s24051669.

[34] S. Goswami, J. Gamon, S. Vargas Zesati, and C. Tweedie, "Relationships of ndvi, biomass, and leaf area index (lai) for six key plant species in barrow, alaska," Mar. 2015. DOI: 10.7287/PEERJ.PREPRINTS.913V1.

[35] S. Marino and A. Alvino, "Vegetation indices data clustering for dynamic monitoring and classification of wheat yield crop traits," *Remote Sensing*, vol. 13, no. 4, 2021, ISSN: 2072-4292. DOI: 10.3390/rs13040541.

[36] S. Indolia, A. K. Goswami, S. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network- a deep learning approach," *Procedia Computer Science*, vol. 132, pp. 679–688, 2018,

International Conference on Computational Intelligence and Data Science, ISSN: 1877-0509. DOI: 10.1016/j.procs.2018.05.069.

[37] B. Macukow, "Neural networks – state of art, brief history, basic models and architecture," K. Saeed and W. Homenda, Eds., pp. 3–14, 2016. DOI: 10.1007/978-3-319-45378-1_1.

[38] C. Du, W. Fan, Y. Ma, H.-I. Jin, and Z. Zhen, "The effect of synergistic approaches of features and ensemble learning algorithms on aboveground biomass estimation of natural secondary forests based on als and landsat 8," *Sensors*, vol. 21, no. 17, p. 5974, 2021. DOI: 10.3390/s21175974.

[39] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.

[40] J. W. Taylor, "A quantile regression neural network approach to estimating the conditional density of multiperiod returns," *Journal of forecasting*, vol. 19, no. 4, pp. 299–311, 2000. DOI: 10.1002/1099-131X(200007)19:4<299::AID-FOR775>3.0.CO;2-V.

[41] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, and M. Data, "Practical machine learning tools and techniques," in *Data mining*, Elsevier Amsterdam, The Netherlands, vol. 2, 2005, pp. 403–413.

[42] A. E. Pascarella, G. Giacco, M. Rigiroli, S. Marrone, and C. Sansone, "Reuse: Regressive unet for carbon storage and above-ground biomass estimation," *Journal of Imaging*, vol. 9, no. 3, 2023, ISSN: 2313-433X. DOI: 10.3390/jimaging9030061.

[43] M. Santoro and O. Cartus, *Esa biomass climate change initiative (biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017 and 2018, v2*, Centre for Environmental Data Analysis, Mar. 17, 2021. DOI: 10.5285/84403d09cef3485883158f4df2989b0c.

[44] Y. W. Yali Zhang Ni Wang and M. Li, "A new strategy for improving the accuracy of forest aboveground biomass estimates in an alpine region based on multi-source remote sensing," *GIScience & Remote Sensing*, vol. 60, no. 1, p. 2 163 574, 2023. DOI: 10.1080/15481603.2022.2163574.

[45] R. Yang, S. K. Singh, M. Tavakkoli, *et al.*, "Cnn-lstm deep learning architecture for computer vision-based modal frequency detection," *Mechanical Systems and Signal Processing*, vol. 144, p. 106 885, 2020, ISSN: 0888-3270. DOI: https://doi.org/10.1016/j.ymssp.2020.106885.

[46] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995. DOI: 10.1007/BF00994018.

[47] X. Wu, V. Kumar, J. R. Quinlan, *et al.*, "Top 10 algorithms in data mining," *Knowledge and information systems*, vol. 14, no. 1, pp. 1–37, 2008. DOI: 10.1007/s10115-007-0114-2.

[48] S. Zafeiriou, M. Bronstein, T. S. Cohen, *et al.*, "Guest editorial: Non-euclidean machine learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. DOI: 10.1109/TPAMI.2021.3129857.

[49] D. Fonseka and C. Chrysoulas, "Data augmentation to improve the performance of a convolutional neural network on image classification," *Proceedings of the International Conference on Data Science and Advanced Analytics (DASA)*, 2020. DOI: 10.1109/DASA51403.2020.9317249.

[50] O. Mutanga, E. Adam, and M. A. Cho, "High density biomass estimation for wetland vegetation using worldview-2 imagery and random forest regression algorithm," *International Journal of Applied Earth Observation and Geoinformation*, vol. 18, pp. 399–406, 2012, ISSN: 1569-8432. DOI: https://doi.org/10.1016/j.jag.2012.03.012.

[51] H. Tamiminia, B. Salehi, M. Mahdianpari, C. M. Beier, L. Johnson, and D. B. Phoenix, "A comparison of decision tree-based models for forest above-ground biomass estimation using a combination of airborne lidar and landsat data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. V-3-2021, pp. 235–241, 2021. DOI: 10.5194/isprs-annals-V-3-2021-235-2021.

[52] S. Eckert, "Improved forest biomass and carbon estimations using texture measures from worldview-2 satellite data," *Remote Sensing*, vol. 4, no. 4, pp. 810–829, 2012, ISSN: 2072-4292. DOI: 10.3390/rs4040810.

[53] H. J. Breaux, "On stepwise multiple linear regression," 1967. [Online]. Available: https://api.semanticscholar.org/CorpusID:119012914.

[54] L. G. Poley and G. J. McDermid, "A systematic review of the factors influencing the estimation of vegetation aboveground biomass using unmanned aerial systems," *Remote Sensing*, vol. 12, no. 7, 2020, ISSN: 2072-4292. DOI: 10.3390/rs12071052.

[55] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021, ISSN: 1422-8890. DOI: 10.1007/s12525-021-00475-2.

[56] I. Veza, Irianto, H. Panchal, *et al.*, "Improved prediction accuracy of biomass heating value using proximate analysis with various ann training algorithms," *Results in engineering*, 2022. DOI: 10.1016/j.rineng.2022.100688.

[57] R. A. de Melo Reis, F. N. dos Santos, and L. Santos, *Forest Robot and Datasets for Biomass Collection*. 2019. DOI: 10.1007/978-3-030-35990-4_13.

[58] S. Gary, "Two-thirds of earth is covered in clouds," *ABC Science*, May 2015, Accessed: 2024-05-25. [Online]. Available: https://www.abc.net.au/science/articles/2015/05/12/4233502.htm.

[59] T. Woldemariam, "Ghg emission assessment guideline: Volume ii: Aboveground biomass field guide for baseline survey," Ministry of Agriculture, Addis Ababa, Ethiopia, Tech. Rep., Apr. 2015.

[60] D. Helman, A. Mussery, I. M. Lensky, and S. Leu, "Detecting changes in biomass productivity in different land management regimes in drylands using satellite-derived vegetation index," *Soil Use and Management*, 2013, Received March 2013; accepted after revision December 2013. DOI: 10.1111/sum.12099.

[61] European Environment Agency, *Corine land cover dataset*, Accessed: 2024-05-19, 2018. [Online]. Available: https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-corine.

[62] S. N. Ferdous, X. Li, K. Sahoo, and R. Bergman, "Analysis of biomass sustainability indicators from a machine learning perspective," *arXiv.org*, 2023. DOI: 10.48550/arXiv.2302.00828.

[63] X. Tian, J. Li, F. Zhang, H. Zhang, and M. Jiang, "Forest aboveground biomass estimation using multisource remote sensing data and deep learning algorithms: A case study over hangzhou area in china," *Remote Sensing*, vol. 16, no. 6, 2024, ISSN: 2072-4292. DOI: 10.3390/rs16061074.

[64] R. Sadono and E. Soraya, "Enhancing the estimation accuracy of above-ground carbon storage in eucalyptus urophylla plantation on timor island, indonesia, through higher spatial-resolution satellite imagery," *Journal of Degraded and Mining Lands Management*, vol. 11, no. 3, pp. 5623–5634, Apr. 2024. DOI: 10.15243/jdmlm.2024.113.5623.

[65] A. Kakumani, L. P. Sree, B. V. Kumar, S. K. Rao, M. Garrepally, and M. Chandrakanth, "Segmentation of cell nuclei in microscopy images using modified resunet," 2022. DOI: 10.1109/GCAT55367.2022.9971978.

[66] P. Kaur and A. M. Mishra, "Segmentation of deep learning models," in *Machine Learning for Edge Computing*, 1st ed., eBook ISBN 9781003143468, CRC Press, 2022, p. 12, ISBN: 9781003143468.

[67] A. C. Kwan, G. Salto, G. Salto, S. Cheng, S. Cheng, and D. Ouyang, "Artificial intelligence in computer vision: Cardiac mri and multimodality imaging segmentation," *Current Cardiovascular Risk Reports*, 2021. DOI: 10.1007/S12170-021-00678-4.

[68] Y. S. Bicakci and B. Sarica, "Attransunet: Semantic segmentation model for building segmentation from aerial image and laser data," *Nordic Machine Intelligence*, 2023. DOI: 10.5617/nmi.10039.

[69] Y. Wang, O. Sakhi, A. E. Ayadi, M. Hagen, and A. Estelle, "Computer vision: Deep dive into object segmentation approaches," 2020. DOI: 10.1145/3394486.3406710.

[70] X. Shi, Y.-C. Wang, Y. Li, and S. Q. Dou, "Remote sensing image segmentation based on hierarchical student's-t mixture model and spatial constrains with adaptive smoothing," *Remote sensing*, 2023. DOI: 10.3390/rs15030828.

[71] X. Pan, J. Xu, J. Zhao, and X. Li, "Hierarchical object-focused and grid-based deep unsupervised segmentation method for high-resolution

remote sensing images," *Remote sensing*, 2022. DOI: 10.3390/rs14225768.

[72] P. Fang, J. Yu, X. Zhang, and T.-P. Li, "Hi-resnet: A high-resolution remote sensing network for semantic segmentation," *arXiv.org*, 2023. DOI: 10.48550/arXiv.2305.12691.

[73] D. Robinson, *Techniques for using deep learning on satellite and aerial imagery*, https://github.com/satellite-image-deep-learning/techniques, 2024.

[74] R. Bienz, *Arborizer: A tool for tree data processing and visualization*, https://github.com/RaffiBienz/arborizer, 2024.

[75] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask r-cnn," Oct. 2017, pp. 2980–2988. DOI: 10.1109/ICCV.2017.322.

[76] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *arXiv preprint arXiv:1505.04597*, 2015. DOI: 10.48550/arXiv.1505.04597.

[77] T. Kattenborn, J. Eichel, and F. E. Fassnacht, "Convolutional neural networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution uav imagery," *Scientific Reports*, vol. 9, p. 17 656, 2019, Received 30 August 2019; Accepted 06 November 2019; Published 27 November 2019. DOI: 10.1038/s41598-019-53797-9.

[78] F. Schiefer, T. Kattenborn, A. Frick, *et al.*, "Mapping forest tree species in high resolution uav-based rgb-imagery by means of convolutional neural networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 170, pp. 205–215, Dec. 2020. DOI: 10.1016/j.isprsjprs.2020.10.015.

[79] G. A. Fricker, J. D. Ventura, J. A. Wolf, M. P. North, F. W. Davis, and J. Franklin, "A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery," *Remote Sensing*, vol. 11, no. 19, 2019, ISSN: 2072-4292. DOI: 10.3390/rs11192326.

[80] N. E. O. Network, *Neon airborne observation platform*, Accessed: 2024-05-26, 2024. [Online]. Available: https://www.neonscience.org/data-collection/airborne-remote-sensing#Data%20Products.

[81] N. E. O. Network, *Neon airborne observation platform*, Accessed: 2024-05-26, 2024. [Online]. Available: https://www.neonscience.org/resources/learning-hub/tutorials/hyper-spec-intro.

[82] N. Yusof, H. Shafri, and N. S. Shaharum, "The use of landsat-8 and sentinel-2 imageries in detecting and mapping rubber trees," *Journal of Rubber Research*, vol. 24, pp. 1–15, Jan. 2021. DOI: 10.1007/s42464-020-00078-0.

[83] D. Schulz, H. Yin, B. Tischbein, S. Verleysdonk, R. Adamou, and N. Kumar, "Land use mapping using sentinel-1 and sentinel-2 time series in a heterogeneous landscape in niger, sahel," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 178, pp. 97–111, 2021, ISSN: 0924-2716. DOI: 10.1016/j.isprsjprs.2021.06.005.

[84] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, 2012, pp. 2274–2282. DOI: 10.1109/TPAMI.2012.120.

[85] A. Adhikari, M. Kumar, S. Agrawal, and R. S, "An integrated object and machine learning approach for tree canopy extraction from uav datasets," *Journal of The Indian Society of Remote Sensing*, 2021. DOI: 10.1007/S12524-020-01240-2.

[86] M. Onishi and T. Ise, "Explainable identification and mapping of trees using uav rgb image and deep learning," *Scientific Reports*, vol. 11, p. 903, 2021, Received 16 June 2020; Accepted 07 December 2020; Published 13 January 2021. DOI: 10.1038/s41598-020-79653-9.

[87] M. Immitzer, K. Einzmann, S. Böck, *et al.*, "Erstellung von fichten- und kiefernanteilskarten auf basis von satellitendaten für bayern," Nov. 2015.

[88] Z. Wang and R. Boesch, "Color- and texture-based image segmentation for improved forest delineation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, pp. 3055–3062, Nov. 2007. DOI: 10.1109/TGRS.2007.896283.

[89] Y. Deng and B. S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 8, pp. 800–810, Aug. 2001. DOI: 10.1109/34.946985.

[90] The REDD Desk, *Forest*, Accessed: 2024-05-19, n.d. [Online]. Available: https://theredddesk.org/encyclopaedia/forest.

[91] European Environment Agency, *Forests and forestry*, Accessed: 2024-05-19, n.d. [Online]. Available: https://www.eea.europa.eu/en/topics/in-depth/forests-and-forestry.

[92] D. Abdullah, F. Fajriana, M. Maryana, *et al.*, "Application of interpolation image by using bi-cubic algorithm," *Journal of Physics: Conference Series*, vol. 1114, p. 012 066, Nov. 2018. DOI: 10.1088/1742-6596/1114/1/012066.

[93] R. M. West, "Best practice in statistics: The use of log transformation," *Ann Clin Biochem*, vol. 59, no. 3, pp. 162–165, May 2022. DOI: 10.1177/00045632211050531.

[94] J. P. D'Haeyer, "Gaussian filtering of images: A regularization approach," *Signal Processing*, vol. 18, no. 2, pp. 169–181, 1989, ISSN: 0165-1684. DOI: https://doi.org/10.1016/0165-1684(89)90048-0. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0165168489900480.

[95] A. Hadid, "The local binary pattern approach and its applications to face analysis," in *2008 First Workshops on Image Processing Theory, Tools and Applications*, 2008, pp. 1–9. DOI: 10.1109/IPTA.2008.4743795.

[96] A. Ng, "Preventing "overfitting" of cross-validation data," *Proceedings of the Fourteenth International Conference on Machine Learning*, Jan. 1998.

[97] P. Liashchynskyi and P. Liashchynskyi, *Grid search, random search, genetic algorithm: A big comparison for nas*, 2019. arXiv: 1912.06059 [cs.LG]. [Online]. Available: https://arxiv.org/abs/1912.06059.

[98] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '16, ACM, Aug. 2016. DOI: 10.1145/2939672.2939785. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939785.

[99] G. Kylberg and I.-M. Sintorn, "Evaluation of noise robustness for local binary pattern descriptors in texture classification," *EURASIP Journal on Image and Video Processing*, vol. 2013, no. 1, p. 17, 2013, ISSN: 1687-5281. DOI: 10.1186/1687-5281-2013-17.

[100] S. Devkota, R. A. Mandal, and A. Khadka, "Assessing the correlation between aboveground carbon stock, ndvi and tree species diversity (a study of kailali and kanchanpur district)," *International Journal of World Policy and Development Studies*, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:255926633.

[101] N.-H. Seong, D. Jung, J. Kim, and K.-S. Han, "Evaluation of ndvi estimation considering atmospheric and brdf correction through himawari-8/ahi," *Asia-Pacific Journal of Atmospheric Sciences*, vol. 56, 2020. DOI: 10.1007/s13143-019-00167-0.

A. Appendix A

A.1 Sentinel 2 and Pleiades Neo Spectral Bands

Band Name	Wavelength (nm)	Resolution (m)
Pleiades Neo		
Deep Blue	400-450	0.3
Blue	450-520	0.3
Green	530-590	0.3
Red	620-690	0.3
Red Edge	700-750	0.3
Near-infrared	770-880	0.3
Sentinel-2		
Coastal aerosol	443	60
Blue	490	10
Green	560	10
Red	665	10
Vegetation red edge	705	20
Vegetation red edge	740	20
Vegetation red edge	783	20
Near-infrared	842	10
Narrow NIR	865	20
Water vapour	945	60
SWIR - Cirrus	1375	60
SWIR	1610	20
SWIR	2190	20

Table A.1: Comparison of spectral bands and resolutions for Pleiades Neo and Sentinel-2 satellites.

B. Appendix B

Transformation	Mathematical Formula	Explanation
Variance 2D	$\mathrm{Var}(X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2$	Variance 2D measures the variability of pixel intensities in a 2D region, indicating texture or noise levels.
Gaussian Filter sigma=1	$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \sigma = 1$	Gaussian filtering smooths the image by averaging pixel values with a Gaussian kernel (sigma=1).
Gaussian Filter sigma=2	$G(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \sigma = 2$	Gaussian filtering smooths the image by averaging pixel values with a Gaussian kernel (sigma=2).
Gradient, Vertical	$\frac{\partial I}{\partial y}$	The vertical gradient detects changes in pixel intensities along the vertical direction, highlighting vertical edges.
Gradient, Horizontal	$\frac{\partial I}{\partial x}$	The horizontal gradient detects changes in pixel intensities along the horizontal direction, highlighting horizontal edges.
Local Binary Pattern	$LBP(x_c, y_c) = \sum_{p=0}^{P-1} s(i_p - i_c)2^p,$ where $s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$	Local Binary Pattern encodes texture by thresholding the neighborhood of each pixel and converting the result into a binary number.

Figure B.1: Description of the layer-based texture transformations.

Algorithm	Mathematical Formula	Explanation
Local Phase features	$PC = \frac{\sum A_n \cos(\phi_n - \bar{\phi})}{\sum A_n}$	Local phase features capture edge and symmetry information based on phase congruency.
Markov Random Fields	$P(I) = \frac{1}{Z} \exp\left(-\sum_c V_c(I_c)\right)$	Markov Random Fields model the spatial dependencies of pixels in an image.
Skewness	$S = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^3$	Skewness measures the asymmetry of the pixel intensity distribution in an image.
Kurtosis	$K = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$	Kurtosis measures the "tailedness" of the pixel intensity distribution in an image.

Figure B.2: Data Point level texture features used, excluding Haralick features (p1)

Algorithm	Mathematical Formula	Explanation
GLCM (4 features)	$GLCM(i,j) = \sum_{p,q}\{1 \text{ if } I(p,q) = i \text{ and } I(p+dx, q+dy) = j\}$	Gray Level Co-occurrence Matrix measures texture by considering the spatial relationship of pixels.
Gabor features	$g(x,y;\lambda,\theta,\psi,\sigma,\gamma) = \exp\left(-\frac{x'^2+\gamma^2y'^2}{2\sigma^2}\right)\cos\left(2\pi\frac{x'}{\lambda}+\psi\right)$	Gabor filters capture spatial frequency characteristics, useful in texture analysis.
HOG features (2)	$\text{HOG} = \text{histogram}(\nabla I, \text{bins})$	The HOG feature extraction involves computing gradients of the image and accumulating a histogram of gradient directions within localized regions.
Texture Energy	$E = \sum_{x,y} I(x,y)^2$	Texture energy is calculated by summing the squared intensities of each pixel in the image, emphasizing regions of higher intensity and texture variations.
Wavelet features (4)	$W = \text{DWT}(I)$	The DWT decomposes an image into a set of wavelet coefficients that represent different frequency bands at different spatial resolutions.
Fractal Dimensions	$D = \lim_{\epsilon\to 0}\frac{\log N(\epsilon)}{\log(1/\epsilon)}$	Fractal dimensions measure the complexity of an image by analyzing its pattern repetition at various scales.

Figure B.3: Data Point level texture features used, excluding Haralick features (p2)

Algorithm	Mathematical Formula	Explanation
Haralick features (13)		
Contrast	$\sum_{i,j} p(i,j)(i-j)^2$	Measures the amount of local variations in the image, highlighting contrast between pixels.
Correlation	$\sum_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}\right]$	Assesses how pixel pairs are correlated to each other across the whole image.
Energy	$\sum_{i,j} p(i,j)^2$	Sum of squared elements in the GLCM, reflecting textural uniformity of image.
Homogeneity	$\sum_{i,j} \frac{p(i,j)}{1+(i-j)^2}$	Measures closeness of the distribution of elements in the GLCM to the GLCM diagonal.
Entropy	$-\sum_{i,j} p(i,j) \log p(i,j)$	Measures the randomness or complexity in the image texture.
Variance	$\sum_{i,j} (i-\mu)^2 p(i,j)$	Measures the dispersion from the mean of the gray levels in the image.
Sum Average	$\sum_i i \sum_{i,j} p(i,j)$	Represents the average sum of gray levels within the window.
Sum Variance	$\sum_i (i-\mathrm{SA})^2 \sum_{i,j} p(i,j)$	Measures the variance of the sum average, indicating variations from the sum average.
Sum Entropy	$-\sum_i \sum_{i,j} p(i,j) \log \sum_{i,j} p(i,j)$	Represents entropy of the sum of gray levels, measuring randomness in sum values.
Difference Variance	$\mathrm{var}(i-j) = \sum_k k^2 \cdot p_{\mathrm{diff}}(k) - (\sum_k k \cdot p_{\mathrm{diff}}(k))^2$	Measures the variance in gray level differences between pixel pairs, providing insight into the contrast variation in the image.
Difference Entropy	$-\sum_k p_{\mathrm{diff}}(k) \log(p_{\mathrm{diff}}(k))$	Entropy of the difference in gray levels between pixel pairs, indicating the randomness in texture contrast.
Information Measures of Correlation (1)	$HXY - \frac{HXY1}{HX}$	Measures dependency of gray levels at the specified positions relative to each other.
Information Measures of Correlation (2)	$\sqrt{1-\exp -2(HXY2 - HXY)}$	Another measure of gray level dependencies, providing a different scale of correlation.
Maximal Correlation Coefficient	Typically computed using the second highest eigenvalue of the Q matrix derived from GLCM	Highest correlation coefficient between any two features.

Figure B.4: Data Point level texture features used: Haralick features

C. Appendix C

Feature	Approach 1	Approach 2	Basic Segmentation - with context	Basic Segmentation - no context
Focus	Layer features	Texture features	No texture features	No texture features
N parameter	20	20	20	1
Accuracy of data point classification	92%	95%	64%	35%
Layers used	red, green, blue, nir; red_lbp, green_lbp, blue_lbp, nir_lbp	blue, nir	red, green, blue, nir, rgb (greyscale)	red, green, blue, nir, rgb (greyscale)
Texture Features used	mean, variance	haralick (13 features); skewness	mean	pixel value
Time of data preparation (one tile, 250x250)	9s	30s	1s	0.1s

Table C.1: Comparison of Feature Selection Approaches