Dissertationes Forestales 343

Detecting individual dead trees using airborne laser scanning

Einari Heinaro

Department of Forest Sciences
Faculty of Agriculture and Forestry
University of Helsinki

Academic dissertation
To be presented for public discussion with the permission of the Faculty of Agriculture and Forestry of the University of Helsinki, in Auditorium B2, B-building (Viikki Campus, Latokartanonkaari 7, Helsinki) on the 24th of November 2023 at 12 pm.

Title of dissertation: Detecting individual dead tree using airborne laser scanning
Author: Einari Heinaro

Dissertationes Forestales 343
https://doi.org/10.14214/df.343

Thesis Supervisors:
Doctor Topi Tahnuanpää
Department of Geographical and Historical Studies, University of Eastern Finland, Finland
Department of Forest Sciences, University of Helsinki, Finland

Professor Markus Holopainen
Department of Forest Sciences, University of Helsinki, Finland

Pre-examiners:
Associate professor Eduardo Maeda
Department of Geosciences and Georaphy, University of Helsinki, Finland

Docent Eva Lindberg
Department of Forest Resource Management, Swedish University of Agricultural Sciences, Sweden

Opponent:
Professor Terje Gobakken
Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Norway

ISSN 1795-7389 (online)
ISBN 978-951-651-776-9 (pdf)

Publishers:
Finnish Society of Forest Science
Faculty of Agriculture and Forestry of the University of Helsinki
School of Forest Sciences of the University of Eastern Finland

Editorial Office:
Finnish Society of Forest Science
Viikinkaari 6, 00790 Helsinki
http://www.dissertationesforestales.fi

Heinaro E. (2023). Detecting individual dead trees using airborne laser scanning. Dissertationes Forestales 343. 55 p. https://doi.org/10.14214/df.343

ABSTRACT

The global crises – climate change and biodiversity loss – have created a need for precise and wide-scale information of forests. Airborne laser scanning (ALS) provides a means for collecting such information, as it enables mapping large areas efficiently with a resolution sufficient for object-level information extraction. Deadwood is an important component of the forest environment, as it stores carbon and provides a habitat for a wide variety of species. Mapping deadwood provides information about the valuable areas regarding biodiversity, which can be used in, e.g., conservation and restoration planning. The aim of this thesis was to develop automated methodology for detecting individual fallen and standing dead trees from ALS data.

Studies **I** and **II** presented a line detection based method for detecting fallen trees and evaluated its performance on a moderate-density ALS dataset (point density approx. 15 points/m^2) and a high point density unmanned aerial vehicle borne laser scanning (ULS) dataset (point density approx. 285 points/m^2). In addition, the studies inspected the dataset, methodology, and forest structure related factors affecting the performance of the method. The studies found that the length and diameter of fallen trees significantly impact their detection probability, and that the majority of large fallen trees can be identified from ALS data automatically. Furthermore, study **I** found that the amount and type of undergrowth and ground vegetation, as well as the size of surrounding living trees determine how accurately fallen trees can be mapped from ALS data. Moreover, study **II** found that increasing the point density of the laser scanning dataset does not automatically improve the performance of fallen tree detection, unless the methodology is adjusted to consider the increase in noise and detail in the point cloud.

Study **III** inspected the feasibility of high-density discrete return ULS data for mapping individual standing dead trees. The individual tree detection method developed in the study was based on a three-step process consisting of individual tree segmentation, feature extraction, and machine learning based classification. The study found that, while some of the large standing dead trees could be identified from the ULS dataset, basing detection on discrete return data and the geometrical properties of trees did not suffice for acquiring applicable deadwood information. Thus, spectral information acquired with multispectral laser scanners or aerial imaging, or full-waveform laser scanning is necessary for detecting individual standing dead trees with a sufficient accuracy.

The findings of this thesis contribute to the existing deadwood detection methodology and improve the understanding of factors to take into account when utilizing ALS for detecting dead trees at a single-tree-level. Although remotely sensed deadwood mapping is still far from a resolved topic, these contributions are a step towards operationalizing remotely sensed biodiversity monitoring.

Keywords: deadwood, biodiversity monitoring, LiDAR, point cloud processing

PREFACE

My journey in the wonderful world of forest sciences began in the end of 2017 when I was seeking for a topic for my Master's thesis. I ended up contacting Markus Holopainen, as I knew his research group was working with LiDAR data – something I was interested in. After some discussion, Markus introduced me to Topi Tanhuanpää, who, at the time, was working with the monitoring of urban trees. Topi provided me with a topic related to LiDAR-based urban tree mapping, which I happily accepted. Little did I know, this marked the beginning of my journey as a researcher.

Once finishing my Master's thesis, Markus asked me whether I would like to continue my academic journey as a doctoral student. He had some funding from the Beetles LIFE project, which aimed at managing the habitats of several endangered beetle species, and he was searching for someone to produce the promised deadwood maps for the project. This role would be suitable for a doctoral student, as in addition to paying a salary, it would allow collecting LiDAR and field data that could be used in further studies. As I had considered pursuing a PhD, I accepted Markus' offer. As a result, my research focus shifted from urban trees to something I knew very little about – dead wood.

Now, several years into my doctoral studies, I have become very familiar with this key ecological indicator and improved my rather limited knowledge of forests in general. I have spent countless hours searching for dead trees in LiDAR point clouds, inspected the various forms of dead trees in the forest, and learned that their decay class is determined based on their penetrability (see, e.g., Heinaro et al. 2021). My journey amidst dead trees is finally reaching its culmination point once I defend this thesis in November. The journey has had its ups and downs, but most importantly, it has involved a large number of people without whom finishing this thesis would not have been possible.

Firstly, I want to thank Markus for accepting the young engineering student to his research group and introducing me to the customs of forest sciences. Markus has provided me with valuable support throughout my doctoral studies and, in addition to pursuing my academic goals, encouraged me to continue chasing my goals in orienteering. Secondly, I am forever grateful for my supervisor Topi who has guided me through the somewhat mysterious customs of the academic world, helped me plan my studies, provided constructive feedback, and sometimes simply been there for me when reaching the finish line has felt too big a challenge. Thirdly, I want to express my gratitude to all the people that have been involved in my studies. This includes my co-authors Topi, Markus, Mikko Vastaranta, Tuomas Yrttimaa, Antero Kukko, Teemu Hakala, and Teppo Mattsson who have helped me improve the manuscripts, assisted in data analysis, and collected data for my studies. It also includes Ninni Mikkonen and Johanna Roiha whose studies I had a privilege to be a part of, and Osmo Suominen and Aleksi Ritakallio with whom I spent several weeks in Eastern Finland collecting field data for my studies. Finally, it includes my pre-examiners Eva Lindberg and Eduardo Maeda who provided valuable comments that helped to improve this thesis.

I have had the privilege to work in a research group that, while taking science seriously, appreciates the other aspects of life as well. I have very much enjoyed being a part of a group that shares the same active lifestyle as I do. A morning run with Ville and Otto is perhaps the best way to start a workday, whereas going climbing with Otto, Topi, and Ninni is likely the best way to end it. I have also truly enjoyed our daily lunches with Jiri, Mohammad, Mikko

N, Otto, Markus, Topi, and two Villes. The culinary experiences offered by the restaurants around the campus would not have been the same without such good company.

Conducting research would not be possible without funding. I want to thank the Doctoral Programme in Sustainable use of Renewable Natural Resources of the University of Helsinki for offering me a salaried position as a doctoral researcher. The position allowed me to fully focus on research instead of spending a large portion of my time applying for funding. I also want to thank the EU's LIFE programme for enabling me to start my doctoral studies and to collect the majority of the datasets used in my studies.

The orienteering club Helsingin Suunnistajat has been an important part of my life for more than twenty years. I have found many of my best friends from the club and shared some of the most memorable moments with my clubmates. The club community has been an important support group that has helped me get through the hardships in the life of a doctoral student. I am very greatful for that.

Finally, I want to thank my parents Pauliina and Heikki for being the biggest supporters of me and my siblings Asko and Ellen. I am very lucky to have been brought up in a supportive environment and to still be able to share the ups and downs of my life with my parents. Having a mum who knows how to navigate the academic world and a dad who shares similar interests related to computing and artificial intelligence has also been pretty helpful during my doctoral studies.

Helsinki, September 2023

Einari Heinaro

LIST OF ORIGINAL ARTICLES

This thesis is based on findings presented in the following articles, referred to by the Roman Numerals I-III.

I Heinaro E, Tanhuanpää T, Yrttimaa T, Holopainen M, Vastaranta M (2021) Airborne laser scanning reveals large tree trunks on forest floor. Forest Ecol Manag 491: 119225. https://doi.org/10.1016/j.foreco.2021.119225

II Heinaro E, Tanhuanpää T, Vastaranta M, Yrttimaa T, Kukko A, Hakala T, Mattsson T, Holopainen M (2023) Evaluating Factors Impacting Fallen Tree Detection from Airborne Laser Scanning Point Clouds. Remote Sens 15(2), article number 382. https://doi.org/10.3390/rs15020382

III Heinaro E, Tanhuanpää T, Kukko A, Hakala T, Holopainen M (2023) Comparing field measurements and annotations as training data for UAV-based detection of standing dead trees in coniferous forest. Int J Remote Sens 44(17). https://doi.org/10.1080/01431161.2023.2248561

AUTHOR'S CONTRIBUTION

I) Heinaro planned the study design together with his supervisors, collected field reference data together with his colleagues, procured the airborne laser scanning dataset, developed the automatic point cloud processing method for detecting fallen trees, conducted all the analyses, wrote the first draft of the manuscript, and modified it based on suggestions given by co-authors.

II) Heinaro planned the study design together with his supervisors, collected field reference data together with his colleagues, developed the automatic point cloud processing method for detecting fallen trees, conducted all the analyses, wrote the first draft of the manuscript, and modified it based on suggestions given by co-authors.

III) Heinaro planned the study design together with his supervisors, collected field reference data together with his colleagues, developed the pipeline for detecting standing dead trees, conducted all the analyses, wrote the first draft of the manuscript, and modified it based on suggestions given by co-authors.

TABLE OF CONTENTS

ABBREVIATIONS

2D	two-dimensional
3D	three-dimensional
ALS	airborne laser scanning
ANOVA	analysis of variance
CHM	canopy height model
CWD	coarse woody debris
DBH	diameter at the breast height
DTM	digital terrain model
FN	false negative
FP	false positive
GNSS	global navigation satellite system
IMU	inertial measurement unit
ITD	individual tree detection and delineation
ML	machine learning
RANSAC	random sample consensus
RGB	red, green, and blue
RS	remote sensing
RTK	real-time kinematic
TLS	terrestrial laser scanning
TN	true negative
TP	true positive
UAV	unmanned aerial vehicle
ULS	UAV-borne laser scanning

1 INTRODUCTION

1.1 Background

Forests are a vital component of planet Earth. They are a home for the majority of species on our planet, provide food and other resources, and significantly contribute to the mitigation of climate change (FAO 2022). A long-sighted and sustainable use of forests and their resources is crucial for ensuring that forests continue to provide these ecosystem services in the future. Forest monitoring provides information that allows utilizing forests in a sustainable manner.

Many countries inventory their forests to gain knowledge of national forest resources. For example, the Natural Forest Resources Institute Finland (Luonnonvarakeskus, LUKE) produces annual estimates of the resources and sequestered carbon in Finnish forests as well as information about forest health and biodiversity (see Luonnonvarakeskus). These inventories include sampling-based field campaigns that are used for estimating forest variables at a regional and national level. Remote sensing can be utilized as additional information in this process, as attributes measured from remotely sensed data can be compared against the field measurements, which allows estimating the resources of areas not covered by the field campaigns more accurately (see, e.g., Næsset 2002). Still, the accuracy of the information collected using this modeling-based approach is rather coarse and only suffices for large-scale decision-making. This is especially true for information regarding rare and scattered resources, as the presence of these resources in the field-sampled data is not sufficient for precise modeling.

The developments in remote sensing technology and need for more precise forest information have driven a shift from area-level forest inventories (Næsset 1997b, 1997a, 2002) to estimating attributes of individual trees (Hyyppä and Inkinen 1999). Laser scanning has proven useful for estimating such attributes, as it provides detailed three-dimensional (3D) information of the target. To estimate attributes of individual trees, trees need to first be identified and segmented from the remote sensing dataset. This process is called individual tree detection and delineation (ITD), and a multitude of algorithms exist for performing this task from different types of remotely sensed data (see, e.g., Ke and Quackenbush 2011; Kaartinen et al. 2012; Wang et al. 2016). Thus far, the majority of the applications utilizing ITD have focused on living trees.

Biodiversity loss is one of the current global challenges. The European Commission has listed the degradation of biodiversity as a major threat to our planet and released a strategy for reversing this phenomenon (European Commission Directorate General for Environment 2021). This strategy involves nature protection and restoration actions that must be focused on areas with high ecological significance. Thus, there is a need for detailed information regarding biodiversity hotspots. Biodiversity is a complex concept that can be inspected at a genetic, species, or ecosystem level (Gaston and Spicer 2004). As a result, directly observing and measuring biodiversity is challenging. Thus, in practice, large-scale biodiversity monitoring is based on indirect observations of biodiversity, i.e. biodiversity indicators, which can be identified from remotely sensed data. Deadwood is one of the most significant of such indicators in forests (Lassauce et al. 2011).

1.2 Deadwood mapping

Deadwood or coarse woody debris (CWD) is a general term used for describing several types of decaying trees. It includes fallen and standing dead trees, stumps, and downed branches (Harmon et al. 1986). Deadwood has an important role in the forest environment due to multiple reasons. Perhaps the most notable reasons include its ecological value and its role as a carbon storage. Firstly, deadwood significantly contributes to the biodiversity in forests. A large variety of species rely on deadwood, as they use it as a habitat and feed on it (Stokland et al. 2012). The majority of these species are insects and fungi. In addition, a number of species indirectly depend on deadwood, as they feed on these insects and fungi (Siitonen 2001; Jonsson et al. 2005; Lassauce et al. 2011). Secondly, deadwood contributes to approximately 8% of the carbon in forests and is thus a significant carbon storage (Pan et al. 2011). The rate at which dead trees decompose and release carbon depend on various factors, including climate, fungi, and insects (Boddy et al. 2008; Bradford et al. 2014; Seibold et al. 2021). Understanding the dynamics of deadwood contributes to the understanding of the global carbon cycle.

Locational information of deadwood is important during the era of biodiversity loss and climate change. Due to its key role in maintaining biodiversity, deadwood can be used as a surrogate measure for estimating the ecological value of forests. Thus, deadwood mapping provides useful information that can be used for, e.g., conservation and restoration planning (see, e.g., Metsähallitus 2018). Conventional sampling-based forest inventory methods (Ståhl et al. 2001; Ståhl et al. 2010; Ducey et al. 2012) are rather inefficient and only provide coarse estimates of deadwood quantities in larger areas. This issue is highlighted by the scattered nature of deadwood occurrence, which places specific requirements for the sampling scheme (Kangas and Maltamo 2006) and limits the usability of deadwood information collected as a part of a larger field campaign, such as the Finnish National Forest Inventory (Luonnonvarakeskus).

Remote sensing (RS) methods, such as aerial imaging and airborne laser scanning (ALS), have been used for addressing the issues related to sampling-based deadwood inventory due to their ability to efficiently monitor large areas on a continuous scale. Modeling-based approaches have used RS-derived features for estimating the quantity of deadwood (Bater et al. 2007; Pasher and King 2009) and for guiding the placement of deadwood field inventory plots (Pesonen et al. 2010a; Pesonen et al. 2010b). Furthermore, deadwood presence has been estimated via canopy gaps identified from RS data (Miura and Jones 2010; Tanhuanpää et al. 2015). These approaches rely on proxy measures that are only weakly related to deadwood presence and thus the accuracy of deadwood information derived using such approaches is rather limited. To overcome this challenge, a growing number of studies have shifted from indirect observations to directly detecting deadwood. The resolution of current remote sensing datasets is not sufficient for detecting smaller deadwood components, such as stumps or downed branches, and thus the direct deadwood mapping approaches have mainly focused on detecting fallen and standing dead trees at a single tree level. Both aerial imaging and ALS have been utilized separately and in combination (Polewski 2017; Briechle et al. 2021; Huo et al. 2023). Especially with fallen trees, ALS has proven useful due to its unique ability to acquire information from below the canopy. In general, different types of approaches are

required for detecting fallen and standing dead trees. These approaches are discussed in more detail in section 1.5.

1.3 Airborne laser scanning

Airborne laser scanning (ALS) is an active remote sensing method in which a scanner, mounted on an aerial vehicle, sends laser pulses to a target, and observes returning signals reflected from the target. On a coarse level, a laser scanner consists of three components: The scanner, which sends and receives laser pulses, a global navigation satellite system (GNSS) receiver, which measures the location of the scanner, and an inertial measurement unit (IMU), which measures the orientation of the scanner (Wehr and Lohr 1999). These components allow converting the time or phase differences of transmitted and returning pulses into 3D coordinates. Transmitting a large number of pulses from different positions allows capturing a dense 3D point cloud of the target.

Laser scanners can be divided into two types based on how they record the returning signal (Baltsavias 1999; Lim et al. 2003). A single laser pulse sent to the target will likely reflect from multiple objects located at different distances from the scanner. As a result, the returning signal will be a continuous waveform containing multiple intensity peaks. These peaks represent reflections from an object. Discrete return scanners use a cut-off intensity value and only record peaks with an intensity above this value. This allows storing the acquired data efficiently, but some information is lost as a downside. In contrast, full-waveform scanners store the returning signal in full. This requires storing large amounts of data, but the advantage is that no information is lost. Certain properties of the full waveform, such as the widths of the peaks can be used for distinguishing different types of objects from each other (see, e.g., Heinzel and Koch 2011; Mücke et al. 2013).

Laser scanners emit laser pulses to different directions using a scanning mechanism based on distinct types of moving mirrors. Due to the scanning mechanism, the distribution of points in the acquired point cloud is never exactly uniform, but rather contains patterns whose shape depends on the scanning mechanism used. These patterns might complicate the detection of certain objects, especially if the pattern resembles the shape of the object of interest. There are several frequently used scanning mechanisms. Most scanning mechanisms, such as oscillating and rotating mirrors generate somewhat linear patterns in the point cloud, whereas the Palmer and wedge prism scanners form circular scan patterns (Wehr and Lohr 1999).

The most frequently used metric for comparing laser scanning datasets is the point density. The point density describes the number of laser returns per unit area and thus depicts the level of detail in the point cloud. The point density of an ALS dataset depends on a number of factors, including the flying altitude and speed during data acquisition, flight line overlap, and the pulse repetition frequency of the scanner (Baltsavias 1999; Wehr and Lohr 1999). In the forest environment, sparse point clouds (point density less than 1 point per m^2) are mostly suitable for observing large-scale phenomena, whereas object-level observations require point densities of at least several points per m^2 (see, e.g., Vauhkonen et al. 2008; Reitberger et al. 2009; Jakubowski et al. 2013a). The point densities of current large-scale laser scanning campaigns are reaching levels sufficient for object-level observations. For example, the point

density of the new nation-wide laser scanning dataset in Finland is 5 points per m^2 (National Land Survey of Finland).

The potential of ALS for monitoring forests is a widely studied topic. ALS has been used for forest inventory (see, e.g., Hyyppä et al. 2008), change detection (Gobakken and Næsset 2004; Yu et al. 2006; Vastaranta et al. 2012), biodiversity assessment (Goetz et al. 2007; Clawges et al. 2008; Kim et al. 2009; Müller and Brandl 2009), and various other forest-related applications. Generally, ALS-based forest monitoring can be divided into two approaches: the area-based approach (Næsset 2002), and the single-tree approach (Hyyppä and Inkinen 1999). The former is based on observing the relationship between field-measured forest attributes and ALS-derived metrics, whereas the latter is based on directly measuring attributes of individual trees from the point cloud.

Conventionally, the platform on which the laser scanner is mounted has been an airplane or helicopter. However, recent developments in laser scanning technology have enabled mounting scanners on smaller platforms, such as unmanned aerial vehicles (UAVs). The main advantages of UAV-borne laser scanning (ULS) include its agility and ability to fly at low altitudes. Due to the latter advantage, ULS enables acquiring datasets with significantly higher point densities than those acquired using a larger platform. This is especially useful when the goal is to identify individual objects from the dataset. In the forest environment, UAVs can even be flown under the canopy (Hyyppä et al. 2020), which allows shifting from the bird's-eye view to the eye level view. Thus far, forest-related studies regarding ULS have been rather scarce and have mainly focused on living trees. These studies have shown that ULS can be used for estimating the locations, heights, diameters at breast height (DBHs), and crown widths of individual trees relatively accurately (Jaakkola et al. 2010; Wallace et al. 2012; Chisholm et al. 2013; Wallace et al. 2014; Jaakkola et al. 2017).

1.4 Object-level information extraction in point clouds

Object-level information extraction can be divided into two categories based on whether the aim is to classify the whole scene or whether only objects of specific types are of interest. Segmentation methods (**Figure 1**a), such as semantic segmentation and instance segmentation (see, e.g., Landrieu and Simonovsky 2018; Wang et al. 2019; Xie et al. 2020), fall within the former category. When working with a point cloud, they aim to assign a label to each point, thus essentially identifying all objects in the point cloud. In contrast, object detection (Zhou and Tuzel 2017; Lang et al. 2019) focuses on identifying specific objects of interest, leaving all other objects as background (**Figure 1**b). Both categories fall within the field of computer vision, but require different kinds of approaches due to the fundamental difference in their objective.

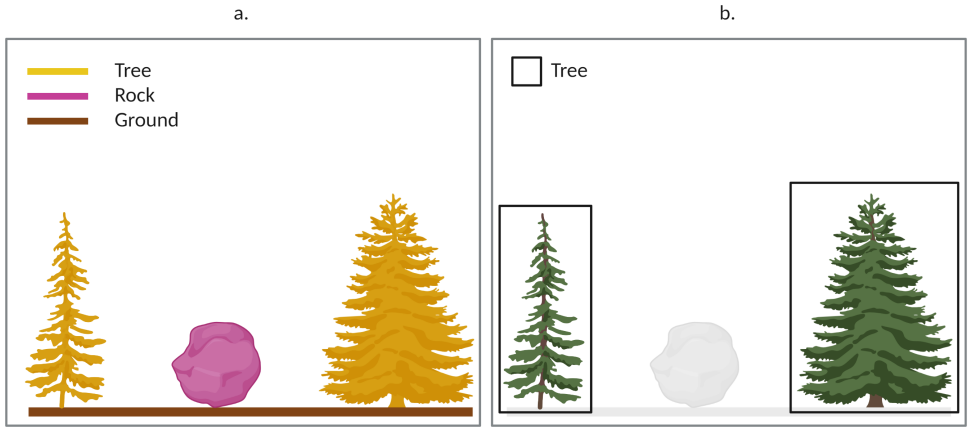

Figure 1 Semantic segmentation (a) versus object detection (b) in a single scene. Segmentation aims at labeling the whole scene, whereas object detection aims at detecting specific objects of interest. In this case, the objects of interest are trees. Please note that the provided example is rather simplistic and real-world segmentation and object detection tasks are much more complex. Figure created with BioRender.com.

Computer vision has conventionally focused on raster data (images) and raster data are still very widely in use. As a result, raster-based computer vision is a highly developed and ever-growing field. In contrast, point cloud based computer vision is a relatively new field, which heavily borrows from the more developed raster-based approaches. Two fundamental differences between these two types of data often prevent directly applying raster-developed methodology on point clouds. Firstly, as opposed to point clouds, rasters are a structured data type, which in many cases simplifies processing and the definition of certain properties. For example, defining the neighborhood in rasters is easy, as each raster cell has the same number of neighboring cells within a specific distance. In contrast, the density of a point cloud is often non-uniform, meaning that different points will have different numbers of points within a fixed distance around them. Secondly, point clouds are a 3D data type, whereas rasters are two-dimensional (2D) by nature. Even with these differences, raster-based approaches have proven useful for point clouds. The simplest way in which raster-based approaches can be used for point clouds is by converting the point cloud into one or more raster layers (e.g., Blanchard et al. 2011; Jakubowski et al. 2013b). The advantage of this approach is that the raster-based methods can be applied as is. However, such an approach induces a loss of information, as rasterizing the point cloud reduces the spatial resolution, and most importantly, loses one spatial dimension. To overcome the latter issue, point clouds can be converted into 3D rasters in which one cell – a voxel – is a cuboid taking a unit volume in space (see, e.g., Xu et al. 2021). Another option for using methods originally developed for rasters is by modifying them to handle unstructured data. Some methods, such as the Hough transform (Hough 1962; Duda and Hart 1972) naturally extend to unstructured data, whereas others require re-defining raster-defined concepts, such as neighborhood (Filin and Pfeifer 2005) and morphology (Calderon and Boubekeur 2014).

Segmentation methods rely on the assumption that elements comprising a single object have similar properties. When thinking of spatial data, spatial similarity lies at the core of this assumption. In other words, spatial segmentation methods assume that an object consists of elements that are located close to each other. Hence, neighborhood is perhaps the most fundamental concept of segmentation methods. For a specific unit of data (e.g., a point cloud point), the neighborhood defines the other units of data that potentially belong to the same object due to their spatial proximity. Other similarity measures or specific criteria are used for determining whether the neighboring units of data in fact belong to the same object. For example, region growing methods are based on an iterative neighborhood search (Gonzalez and Woods 2008). Region growing starts with a set of seed points representing individual segments. The method inspects the neighborhood of each seed point and adds units of data to the same segment as the neighboring seed point if they fulfil certain criteria. These criteria might be based on the similarity of some properties or geometric relationships (see, e.g., Vincent and Soille 1991). This process is repeated iteratively until a stopping condition is reached. Another type of segmentation approach – clustering – aims at grouping units of data to individual segments by maximizing the similarity of units of data within each segment and, in some cases, also minimizing the similarity between different segments (Jianbo and Malik 2000). Perhaps the most well-known clustering method, k-means clustering (e.g., MacQueen and Neyman 1967), starts by randomly assigning each unit of data into a cluster/segment and computing the centroid of each cluster. The cluster assignment is then revised by assigning each unit of data into the cluster with the most similar centroid. The algorithm repeats this process iteratively until the cluster assignment stabilizes.

As opposed to segmentation, object detection methods do not aim at classifying the whole scene, but rather aim at identifying specific objects from a dataset. This requires different types of approaches than the ones used for segmentation. The core of object detection methods is identifying objects based on their geometric and/or other properties. In point clouds, these other properties could include, for example, intensity-related features. For objects with simple shapes, parametric shape detection is a viable option. Parametric shape detectors, such as random sample consensus (RANSAC) (Fischler and Bolles 1981) and Hough transform (Hough 1962; Duda and Hart 1972) aim to detect simple shapes, such as lines in an image or point cloud. These approaches assume a parametric form for the shape to be detected and find the parameter values that best fit the data. Template matching can be used for more complex objects that cannot easily be represented in parametric form. As the name suggests, template matching uses a template that resembles the object of interest and compares different parts of the dataset against this template (Brunelli 2009). If a part of the dataset matches the template, an object is detected at that specific location. At its simplest, the template used is manually crafted and used for matching as is. The more recent approach is, however, using machine learning (ML). This approach includes crafting a set of features (multiple templates) and using samples of the object of interest for teaching the ML model which of these features are useful for detecting the object (Zou et al. 2023).

During the last decade, the developments in ML and especially deep learning have revolutionized computer vision (e.g., Voulodimos et al. 2018; Zou et al. 2023). Both segmentation and object detection have seen a wide variety of new algorithms. This has been mainly due to the increase in computational power, which has allowed approaching these tasks in a different manner. The main advantage of ML is that it removes the need for carefully crafted rules that solve a specific task. In ML, the model takes as input a set of

features and learns which of these features are most relevant for solving the given task. Deep learning takes a step further and, in addition to learning which features are relevant, learns the features themselves. This allows a more generic approach to object-level information extraction, as a unique algorithm is not required for each specific task. The disadvantage of ML is, however, that it requires training data and computational power. These issues are highlighted with deep learning models, which might include millions or even billions of parameters (see, e.g., Zaidi et al. 2022). Training such models require huge amounts of data and computational resources, which are often not available. Luckily, many of the state of the art ML models are openly available in online repositories (see, e.g., Hugging Face). These models can be used as is or fine-tuned to solve a new task for which only limited training data is available. The latter approach is called transfer learning (Weiss et al. 2016). The idea behind transfer learning is that state of the art ML models, such as deep convolutional neural networks, consist of a large stack of layers, each of which learns to detect specific types of features in a hierarchical manner. The first layers learn to detect simple features, such as lines or curves, which they then feed forward to the next layers. The next layers learn to combine these simple features into more complicated ones, and finally, the last layers learn to detect whole objects. Since the lower layers learn to detect features that are not specific to the task at hand, these layers can be utilized for another detection task, and only the last layers need to be retrained using new training data.

1.5 Detecting individual dead trees from ALS data

1.5.1 Single tree level information extraction in forests

Forests are complex and dynamic environments, which makes object-level information extraction challenging. Trees come in many shapes and sizes, making them a complicated object to detect, as object-level information extraction requires that objects belonging to the same class are somewhat similar. Furthermore, trees have overlapping crowns and may sometimes be occluded by crowns of larger trees, which generates problems in distinguishing trees from each other. Other objects, such as boulders and shrubs add to these challenges. This is especially evident with fallen trees that may not be visible in ALS data due to lush ground vegetation or abundant undergrowth. Stems are the single most distinguishable part of trees, as their shape is approximately cylindrical. However, the problem is that the stems of standing trees are vertical and thus they might not be visible in ALS data collected from above the canopy.

ITD is a widely studied topic, as it serves as the basis for most single tree level forest inventory applications. Such applications include tree species classification (Dalponte et al. 2014; Amiri et al. 2019), forest mensuration (Hyyppä and Inkinen 1999; Popescu et al. 2003) ,and standing dead tree detection (Yao et al. 2012; Casas et al. 2016). ALS-based ITD can roughly be divided into two main categories depending on whether individual trees are extracted from a canopy height model (CHM) derived from the point cloud (see, e.g., Hyyppä et al. 2001; Persson et al. 2002; Dalponte and Coomes 2016) or from the point cloud directly (Li et al. 2012; Lu et al. 2014). The CHM-based approach starts by detecting treetops as local maxima in a rasterized CHM. Individual trees are then segmented from the CHM using a segmentation algorithm, such as watershed segmentation (see, e.g., Vincent and Soille 1991),

which uses the treetop locations as seed points. Similarly, point cloud based approaches first identify tree locations as, for example, the highest point cloud points within their neighborhood. Individual tree segments are then delineated around these locations using a rule-based region growing algorithm. ITD algorithms may also combine CHM and point cloud approaches by first detecting treetops from a CHM and then using these as seed points for a point cloud based region growing method (Morsdorf et al. 2003; Reitberger et al. 2009). Forest characteristics vary significantly between different locations and even within a single area and thus there is no single ITD method that is guaranteed to work. Moreover, different ITD methods have different strengths and weaknesses. Generally, ITD is a compromise between over- and undersegmentation. For example, CHM-based ITD methods are generally less prone to oversegmentation (i.e., splitting single trees into multiple segments), but as a downside, are poor at detecting trees below the top-most canopy layer. In contrast, point cloud based approaches are better at detecting small trees, but tend to generate too many tree segments (Wang et al. 2016). Thus, the best choice for the ITD algorithm depends on the objective.

1.5.2 Fallen tree detection

Studies addressing ALS-based fallen tree detection have based the detection on the most distinguishable part of trees – the stem. Stems are cylindrical objects that can be detected based on their shape. Since fallen trees mostly lie on the forest floor, their stems are horizontal, which increases the chances of them being visible in ALS data. However, the challenge of fallen tree detection is that standing trees hinder the visibility to the ground and thus fallen trees located under dense canopy might not be visible in data collected from above the canopy. Laser scanners are able to penetrate canopy, but canopy limits the number of laser returns from the ground layer. Another challenge with fallen tree detection are other near-ground objects and ground vegetation, which might mistakenly be identified as fallen trees and which obstruct visibility to fallen trees.

Laser scanning based fallen tree detection methods can roughly be divided into segmentation-based approaches, template matching methods and shape detection methods. Segmentation-based methods (Blanchard et al. 2011; Mücke et al. 2013) create raster layers from several laser scanning metrics, perform segmentation on these layers, and identify segments resembling fallen trees based on their geometry. Template matching methods (Lindberg et al. 2013; Nyström et al. 2014; Polewski et al. 2015b, 2018) utilize 2D or 3D templates for detecting fallen trees from either the point cloud directly or from a rasterized height model generated from the point cloud. Thus far, the only method utilizing shape detection was introduced by Yrttimaa et al. (2019), who first filtered a terrestrial laser scanning (TLS) point cloud by RANSAC-based cylinder fitting and used a segmentation-based procedure for detecting fallen trees from the filtered point cloud.

There are two main goals in fallen tree detection. The first one is to detect as large a proportion of fallen trees as possible. The latter is making as few false detections as possible. These two goals often contradict each other, as developing fallen tree detection methods includes a number of decisions that affect the sensitivity of the method. Such decisions could be, for example, how linear must a segment be, or how well does a subset of the point cloud have to match the template for it to be identified as a fallen tree. Increasing the sensitivity of the method will generally result in a larger number of true detections but also a larger number

of false detections and vice versa. To address this challenge, different forms of filtering have been used for reducing the number of false detections while keeping the number of true detections relatively constant. These filtering procedures can be applied before or after the actual fallen tree detection phase. The filters applied before the actual detection phase aim at removing as many laser points not originating from fallen trees as possible. The simplest such filter is the height-based filter, which limits fallen tree search to a specific height range close to the ground, as this is where most fallen trees are located. Other pre-detection filters aim at removing points not originating fallen trees based on their properties, such as return type (Mücke et al. 2013; Nyström et al. 2014), echo width (Mücke et al. 2013), or geometric characteristics of the point neighborhood (Polewski et al. 2015b; Yrttimaa et al. 2019). In contrast, the filters applied after detection aim at removing false detections by comparing them to known instances of fallen trees. ML can be used for automating this procedure (Nyström et al. 2014; Polewski et al. 2015b).

1.5.3 Methods for detecting standing dead trees

Most studies attempting to detect standing dead trees from laser scanning data have based detection on a three-step process consisting of ITD, feature extraction, and classification (Yao et al. 2012; Polewski et al. 2015c; Casas et al. 2016; Amiri et al. 2019). The ITD step aims at extracting as accurate an individual tree segmentation as possible. The feature extraction step aims at describing the geometric and intensity-related properties of each segment. Often a large number of features are extracted, and some sort of feature selection, such as sequential feature selection (Amiri et al. 2019) or a genetic algorithm (Polewski et al. 2015c) is used for finding the most useful feature set. The classification step uses the extracted features for determining whether a tree segment represents a standing dead tree. Some ML method, such as logistic regression (Amiri et al. 2019) or a support vector machine (Yao et al. 2012) is trained using positive and negative samples and used as the classifier.

The successful segmentation of individual trees is crucial for standing dead tree detection methods based on the aforementioned three-step process due to several reasons. Firstly, if the ITD method undersegments the point cloud, some standing dead trees are left undetected, whereas oversegmentation increases the probability for false detections. Secondly, features calculated for inaccurately segmented trees might be distorted, which will hinder the learning process of the ML classifiers and complicate the classification of tree segments. Accurately detecting and segmenting trees in all size groups is a challenging task, as small trees are often occluded by larger ones (Wing et al. 2015; Casas et al. 2016). Thus, when aiming to detect standing dead trees, it might make sense to focus on the larger trees that can be distinguished with relative ease. Another option is to develop methodology that does not rely on ITD. The only such standing dead tree detection method presented thus far is the one by Miltiadou et al. (2018; 2020), who created a dead tree probability layer from voxelized ALS data. In addition, Wing et al. (2015) aimed at mitigating the impact of ITD by performing intensity-based point-level filtering for extracting all points originating from standing dead trees. These points were then input into a standard CHM-based segmentation procedure for extracting individual standing dead tree segments.

1.6 Objectives of the thesis

Existing research has highlighted the benefits of ALS in remotely sensed deadwood mapping. The ongoing development in laser scanning technology has enabled acquiring nation-wide laser scanning datasets with point densities sufficient for object-level mapping. However, collecting precise and utilizable deadwood information using ALS is not yet possible at an operational level. This is due to limitations in the accuracy and scalability of existing methodology and datasets. To enable operational deadwood mapping in the future, there is a need for more information regarding the requirements for the methodology and datasets used for deadwood mapping. Furthermore, it is important to gain insights on the various factors affecting the accuracy of ALS-based deadwood mapping. This thesis aims to address these gaps in knowledge.

The main objective of this thesis was to develop single tree level point cloud-based deadwood detection methodology. Furthermore, the objective was to gain an understanding of the various aspects affecting deadwood detection, including the properties of dead trees themselves, the characteristics of the forest surrounding them, and the properties of the datasets used for detecting dead trees. The thesis consists of three studies, each aiming to partly fulfil these objectives. **Figure 2** presents a summary of the thesis and its substudies.

Study **I** presents a method for detecting individual fallen trees from ALS point clouds. The study inspects whether fallen trees can be distinguished from point clouds based on their linear shape. Furthermore, the study inspects how the characteristics of fallen trees and their surrounding vegetation structure impacts the accuracy of fallen tree detection. The study aims to answer the following questions:

1. Can individual fallen trees be mapped from point clouds?
2. Which fallen tree characteristics impact their detection and how?
3. How does the surrounding vegetation structure influence the detection of fallen trees?

Study **II** extends on the methodology created in study **I** and inspects the differences in performance when the fallen tree detection method is applied on a moderate density ALS dataset and a high point density ULS dataset. The study aims to find out how the different properties of laser scanning datasets affect the performance of the line detection-based fallen tree detection method. Furthermore, the study performs a sensitivity analysis on the parameters of the fallen tree detection method to gain a deeper understanding of how the method should be tuned in different situations. The study aims to answer the following questions:

1. How do the properties of the laser scanning dataset impact the performance of fallen tree detection?
2. How can line detection-based fallen tree detection be adjusted for different datasets?
3. Can machine learning-based filters be used for improving the performance of fallen tree detection?

Study **III** shifts the focus from fallen trees to standing dead trees and presents a pipeline for detecting standing dead trees from laser scanning data. The study aims to find out what types of standing dead trees can be identified from point clouds. Furthermore, the study inspects the effects of using samples manually labeled from UAV images (i.e., annotated data) instead of field reference data as training data for standing dead tree classification. The study aims to answer the following questions:

1. How accurately can standing dead trees of different sizes be identified?
2. Does annotated training data bias standing dead tree detection?
3. Is it beneficial to focus on detecting large dead trees?

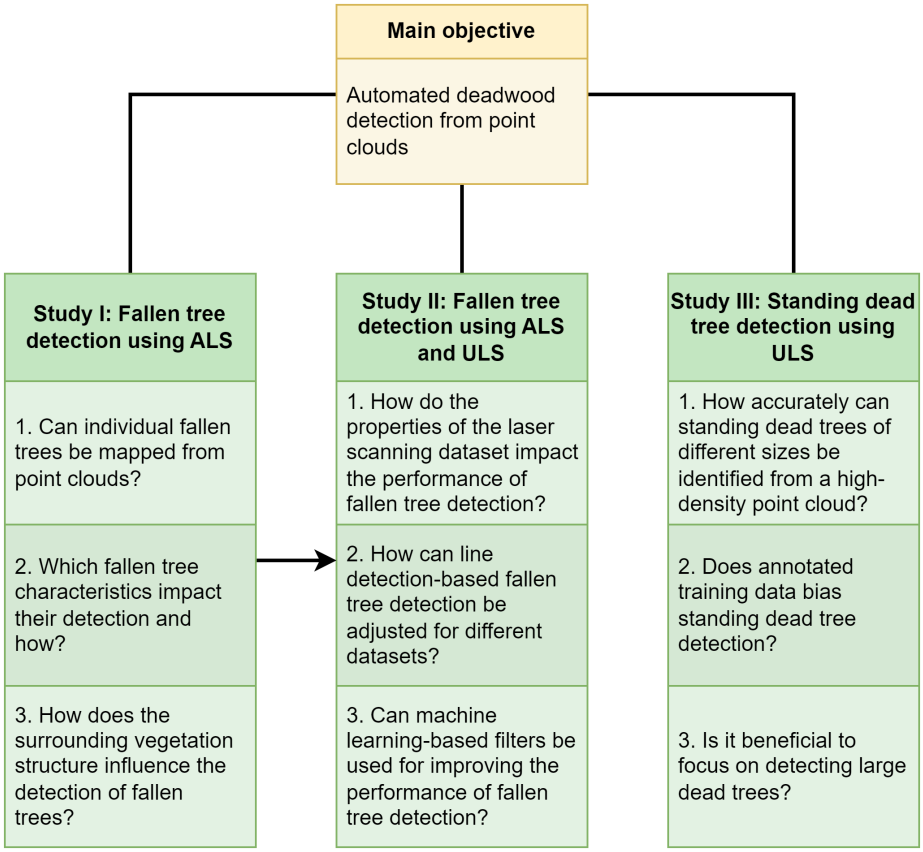

Figure 2 A summary of the thesis and its substudies. The figure presents the main objective of the thesis and the questions each substudy aims to answer. The arrow between studies I and II demonstrates that study II extends study I.

2 MATERIAL

2.1 Study site

All studies presented in this thesis were conducted in an approximately 16 km^2 area in the region of Kainuu in Finland (**Figure 3**). The study site can be characterized as typical boreal forest in which the main tree species are Norway spruce (*Picea abies* (L.) H. Karst.), Scots pine (*Pinus sylvestris* L.), silver birch (*Betula pendula* Roth), and downy birch (*Betula pubescens* Ehrh.). Some aspens (*Populus tremula* L.) could also be found within the site. The western and eastern parts of the study site were within Hiidenportti national park and Teeri-Lososuo mire conservation area, respectively. These parts mainly consisted of old-growth forest. In contrast, the part between these two conservation areas was state-owned managed forest with varying forest maturity. The topography of the site varied from flat to steep with elevations between 190 and 250 meters above sea level.

2.2 Airborne laser scanning datasets

2.2.1 Moderate-density airborne laser scanning data (studies I-II)

The moderate-density ALS dataset (solid black line in **Figure 3**) used in studies **I** and **II** was collected from within the whole study site using a Riegl VQ1560i laser scanner (RIEGL Laser Measurement Systems GmbH, Horn, Austria). The scanner collected data using two channels that scan the target in straight lines tilted 28 degrees against each other. As a result, the point distribution of the dataset was rather uniform with few patterns resulting from the scanning mechanism. The data were collected on May 17th, 2019 in leaf-off conditions to maximize the visibility of below-canopy objects. The data were acquired with five parallel flight lines with a 30% overlap and a single flight line perpendicular to the other lines. The point density of the dataset was around 15 points/m^2, although the density varied depending on flight line coverage.

2.2.2 High-density unmanned aerial vehicle -borne laser scanning data (studies II-III)

The high-density ULS dataset (black dashed lines in **Figure 3**) used in studies **II** and **III** was collected from five subsites (total area 2.4 km^2) within the study site using a Riegl miniVUX-1DL laser scanner. The scanner collected data using a wedge prism scanner, resulting in a circular scan patterns in the dataset. The data were collected in the beginning of June 2020 in mostly leaf-off conditions. Six to nine flight lines with 30% overlap were used for each subsite, depending on the shape and size of the subsite. The point density of the dataset was approximately 285 points/m^2, although the density varied depending on flight line coverage.

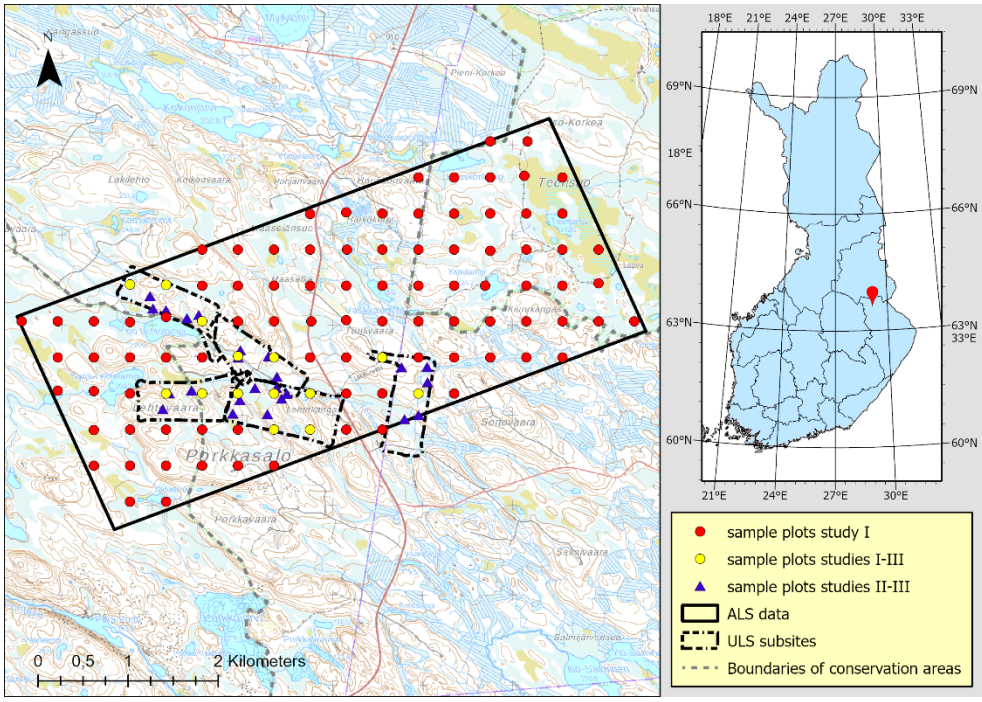

Figure 3 The study site. Background maps: National Land Survey of Finland.

2.3 Field reference data

The field reference data were collected in two campaigns. The data in both campaigns were collected from circular sample plots with a 9 m radius. Campaign 1 was executed between July and September 2019 and was designed for capturing the full variety of forest conditions within the study site. The campaign consisted of measuring 103 sample plots (red and yellow circles in **Figure 3**) distributed evenly across the study site. Campaign 2 aimed to supplement the field reference data within the ULS subsites. 23 sample plots (blue triangles in **Figure 3**) were placed within the ULS subsites and measured in November 2020. These sample plots were placed in locations with known abundance of deadwood to ensure that the field data would contain a sufficient number of dead tree observations.

The same measurement process was used in both campaigns. The locations of fallen trees and standing dead trees within the sample plots were measured using a Trimble R2 (Trimble Inc., Sunnyvale, California, USA) real-time kinematic (RTK) GNSS receiver. The locations of fallen trees were measured from both the top and bottom-end of the tree, whereas the locations of standing dead trees were measured by placing the receiver as close as possible to the tree stem. Fallen trees were measured in full even if they were only partly within the sample plots. Apart from the GNSS measurements taken at one sample plot, a fixed solution was acquired for all measurements and the positional accuracy of the measurement itself was within 30 centimeters. However, the placement of the GNSS device caused some further

positioning errors, as, for example, the receiver had to be placed on one side of a standing dead tree. Taking these factors into account, the positioning accuracy was still within one meter. The diameters of both fallen and standing dead trees were measured using steel calipers. Only fallen trees with a DBH over 100 mm and standing dead trees with a DBH over 45 mm were measured. However, the DBH could not be determined from some of the fallen trees, and in this case, the diameter was measured at the bottom-end of the tree. Lengths of fallen trees were determined based on the GNSS measurements, whereas lengths of standing dead trees were determined using a Vertex 5 measurement device (Haglöf Sweden AB, Långsele, Sweden). The species of both standing and fallen trees were determined based on visual inspection, whereas the decay state was measured on a scale of 1 (least decayed) to 5 (most decayed) using the guidelines of the Finnish national forest inventory (Metsäntutkimuslaitos 2009).

In addition to measuring dead trees, all living trees with a DBH over 45 mm located within the sample plots were mapped and measured to characterize the forest within the sample plots. The locations of living trees were determined based on the direction and distance from the sample plot center. The location of the sample plot center was measured using an RTK GNSS receiver. The DBH and species of each living tree were recorded. Furthermore, a representative sample of the living trees was selected from each sample plot and the heights of these sample trees were measured using a Vertex 5 measurement device. These heights were then used for deriving the heights of the other trees using the recorded DBHs and Näslund's (1936) function. Moreover, the undergrowth at each sample plot was determined by counting the species-specific occurrences of trees with a DBH less than 45 mm and a height over 1.3 m located within 5.4 m from the sample plot center. The information about living trees was used for describing the vegetation structure at each sample plot. More specifically, the measurements of living trees and undergrowth were aggregated at the sample plot level to acquire the general as well as the species-specific mean DBH, mean height, tree count, amount of undergrowth, basal area, and volume. Furthermore, the canopy cover of each sample plot was estimated from ALS data using the method presented in Polewski et al. (2015b).

Study **I** used the 103 sample plots measured in campaign 1, whereas studies **II** and **III** utilized all sample plots located within the ULS subsites. These consisted of the 23 sample plots measured in campaign 2 and 14 sample plots measured in campaign 1. **Table 1** presents a summary of the field-measured fallen trees in studies **I** and **II** and the field-measured standing dead trees in study **III**.

Table 1 A summary of the field-measured dead trees used in each study. Note that for studies I and II, the table presents a summary of fallen trees, whereas for study III, the table presents a summary of standing dead trees.

Study	I	II	III
Tree type	fallen trees	fallen trees	standing dead trees
Sample plots	103	37	37
Number of trees	273	197	75
Height (m)			
min	1.5	1.6	1.5
mean	10.7	11.9	11.2
max	23.1	28.8	27.1
standard deviation	4.9	5.1	5.7
Diameter (mm)			
min	100	100	92
mean	176.3	193	194
max	450	450	660
standard deviation	60.4	72.9	93
Decay class			
min	1	1	1
median	2	1	1
mode	1	1	1
max	5	5	3

2.4 Annotated data (study III)

Study **III** utilized an annotated standing dead tree dataset as training data. The dataset was collected by digitizing the crowns of all visually identifiable standing dead trees within several bounded areas in red, green, and blue (RGB) aerial images (i.e., virtual sample plots). The aerial images were collected in July 2019 using a UAV. The annotated dataset was collected by another research group and covered parts of the ULS subsites. Crown polygons located within the reference sample plots (see section 2.3) were excluded from the dataset. In total, the annotated dataset used in study **III** consisted of 148 standing dead trees. **Table 2** presents a summary of the annotated trees, whereas **Figure 4** shows their locations.

Table 2 A summary of the annotated standing dead trees. The tree heights were determined from ULS data.

Number of trees	148
Height (m)	
min	6.2
mean	18.1
max	28.6
standard deviation	4.5

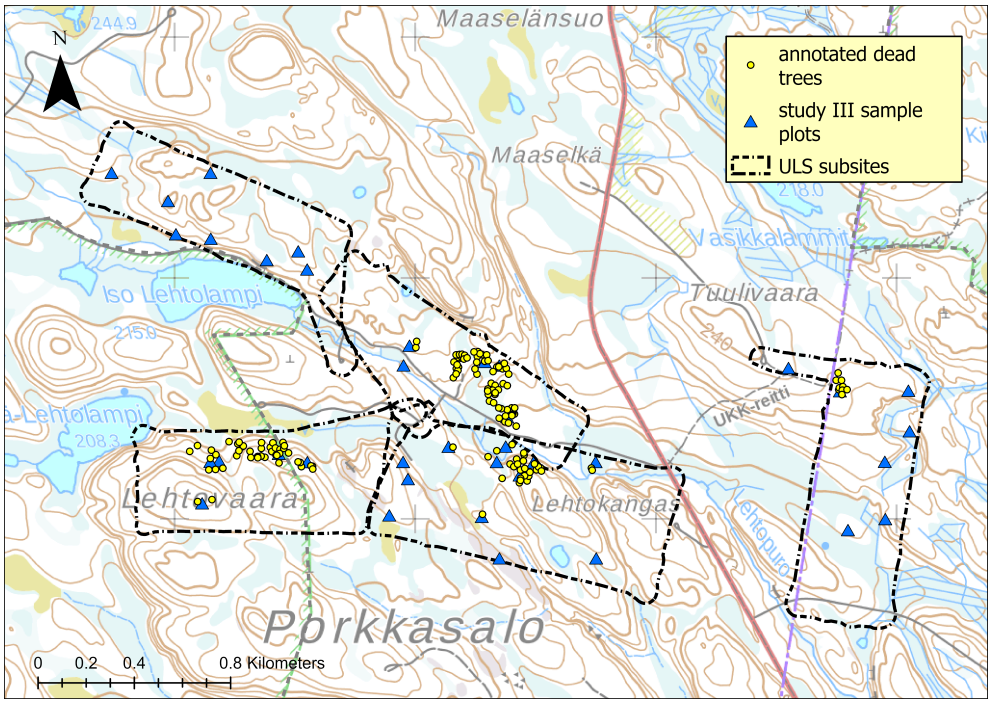

Figure 4 The locations of the annotated standing dead trees. Background map: National Land Survey of Finland.

3 METHODOLOGICAL OVERVIEW

3.1 Pre-processing of laser scanning point clouds

The deadwood detection methods used in the studies required height-normalized laser scanning point clouds. Thus, as a first step, both the ALS and ULS dataset were height-normalized by subtracting the corresponding digital terrain model (DTM) height from the z-coordinate of each point cloud point. As a result, the z-coordinates of the points represented heights above ground. The DTMs used for height-normalization were generated from the points classified as ground. In study **I**, the ALS dataset was ground-classified using a slightly modified version of Axelsson (2000) in the Terrascan software (Terrasolid Ltd., Espoo, Finland). In study **II**, the ALS and ULS datasets were ground-classified with the method by Zhang et al. (2003). In study **III**, the ULS dataset was ground-classified with the method by Zhang et al. (2016). The ground classification in studies **II** and **III** was performed using the functions in the lidR package (Roussel et al. 2020; Roussel and Auty 2022) developed for the R programming language.

3.2 The method for detecting fallen trees (studies I-II)

A method for detecting individual fallen trees in laser scanning point clouds was developed in study **I** and further inspected in study **II**. The method consisted of the following steps:

1. Filtering the height-normalized point cloud based on height and projecting the remaining points onto the xy-plane (step 3 in **Figure 5**).
2. Filtering the projected points based on the shapes of point groups (connected components) and an ML-based classifier (step 4a in **Figure 5**).
3. Placing a rectangular grid (cell size 20 × 20 m) on top of the projected point cloud and detecting line segments within each grid cell separately using iterative Hough transform (step 4b in **Figure 5**).
4. Merging the line segments detected in neighboring grid cells using distance-, angle-, and overlap-based criteria (step 4b in **Figure 5**).
5. Extracting fallen tree segments from the point cloud using a region growing algorithm that assigned nearby points to each detected line segment (step 5 in **Figure 5**).
6. Classifying the fallen tree segments as true or false fallen tree segments using a ML-based classifier (step 6 in **Figure 5**).

Each of these steps is discussed in more detail in the following subsections. In study **I**, the fallen tree detection methodology was applied on the ALS dataset, whereas in study **II**, the methodology was applied on both the ALS and the ULS dataset to inspect how the performance of the method differs between the two datasets. The fallen tree detection method was developed using the MATLAB programming language and is available in GitHub (see Heinaro 2021).

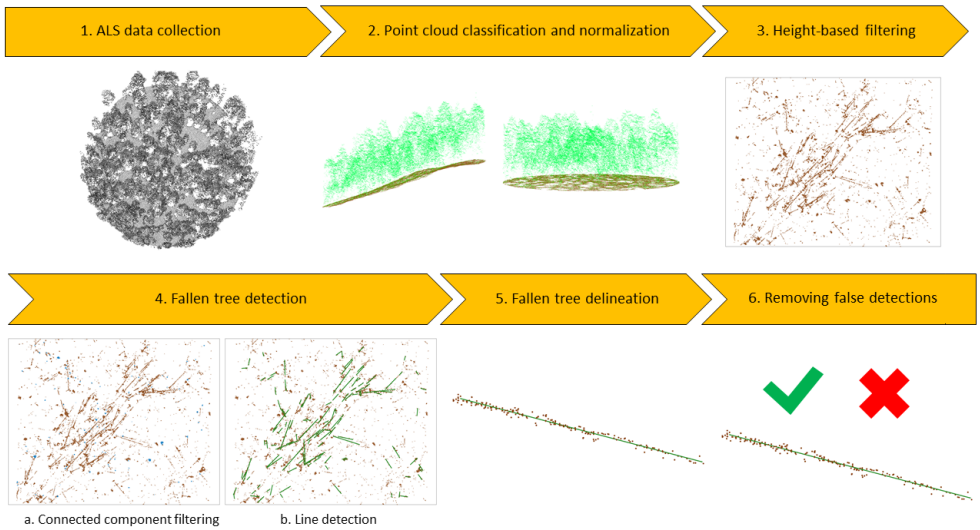

Figure 5 The fallen tree detection method.

3.2.1 Height-based filtering and point projection

As fallen trees mostly reside close to the ground, the height-normalized point cloud was first filtered based on height (i.e., only points falling within a specific height interval were retained for further analysis). This seemingly simple step is rather crucial for the performance of fallen tree detection, as it determines how large a fraction of fallen trees are visible and how much noise is present in the filtered point cloud. A suitable height interval should retain most of the fallen trees while filtering out as large a number of other point cloud points as possible. In both studies, the upper threshold for the height interval was set to 1 m, as this was empirically found to be a suitable interval. Ideally, the lower height threshold would be set to exactly 0 m, as this would retain all points close to ground. However, due to points originating from ground vegetation and errors in ground classification, setting the lower height threshold to exactly 0 m results in a very noisy point cloud from which the fallen trees cannot be identified. Thus, in practice, the lower height threshold must be set to a value slightly larger than 0 m. In study **I**, the lower height threshold was set to 0.2 m based on trial and error, whereas study **II** experimented with different threshold values. After height-based filtering, the points falling within the specified height interval were projected onto the xy-plane to simplify further analysis.

3.2.2 Connected component analysis

The height-filtered and projected point cloud contained points originating from fallen trees, but also a large number of points originating from other near-ground objects, such as undergrowth. Furthermore, even after height-based filtering, the point cloud contained

uniform point patches originating from errors in ground classification. To reduce the number of false fallen tree detections, the points not originating from fallen trees needed to be removed. Connected component analysis was used for this task. First, the point cloud was converted into a binary image with a cell size of 0.2 m. Each pixel in the image was assigned the value 1 if it contained at least one point cloud point. Otherwise, the pixel was labeled as a 0. Next, the binary image was segmented into connected components, which consisted of all pixels with a value 1 that were connected to each other through other pixels labeled as 1. 8-neighborhood was used for determining adjacency between pixels. Then, several shape descriptors describing geometrical properties were calculated for each connected component. Finally, the connected components were classified as originating or not originating from fallen trees based on the shape descriptors. Only point cloud points located within connected components originating from fallen trees were retained for further analysis. A shallow neural network consisting of one hidden layer was used as the classifier. The classifier was trained on manually labeled connected components that were extracted from the ALS dataset. The manually labeled components were extracted from outside the field-measured sample plots to ensure that the same components would not be used for training and testing.

3.2.3 Line detection using iterative Hough transform

Due to the linear shape of fallen trees, they can be detected as line segments in a point cloud. Thus, the core of the fallen tree detection method was a line detection algorithm. The projected and filtered point cloud was first split into 20 × 20 m rectangular regions. Then, line detection based on iterative Hough transform was applied to each region separately. Hough transform (Hough 1962; Duda and Hart 1972) searches for lines in a point cloud by transforming the coordinates of each point into all possible parameters of lines passing through the point. The outputs of Hough transform are the parameters of the line that passes through the largest number of points in the point cloud. After applying Hough transform, the infinite output line was cut into a finite line segment by searching for large gaps between the points falling on the line. The line segment was placed in the position of the largest continuous group of such points. All points located within 0.5 m from the line segment were then removed, and Hough transform was applied repeatedly until a stopping criterion was reached. Hence the name iterative Hough transform. The stopping criterion was based on the minimum number of points falling on a detected line. In study **I**, the threshold value was four points, whereas several different threshold values were tested in study **II**.

3.2.4 Merging line segments

The output of the iterative Hough transform algorithm was a set of line segments detected in each of the 20 × 20 m rectangular regions. As some fallen trees crossed the boundaries of these regions, it was possible that several line segments in neighboring regions represented different parts of the same fallen tree. Thus, line segments likely representing the same tree were merged using a merging scheme based on distance, angle, and overlap thresholds.

3.2.5 Extracting point cloud representations of fallen tree segments

A region growing algorithm was applied for extracting the point cloud representations of fallen trees. The line segments were used as the skeleton around which the segments were grown. First, all points within 0.5 meters from a line segment were assigned to the segment. Then, all points located within 0.2 meters of any point already assigned to the segment were added to the segment. This step was repeated until no further points were added.

3.2.6 Classifying segments as true or false detections

As a last step, the point cloud representations of fallen trees were input into a convolutional neural network (CNN) that determined whether each segment truly represented a fallen tree or not. The ones classified as false detections were removed. A pretrained AlexNet (Krizhevsky et al. 2012) was used as the base model and its weights were tuned using transfer learning. The training data consisted of manually labeled segments extracted from areas outside the sample plots.

3.3 The method for detecting standing dead trees (study III)

A method for detecting standing dead trees was developed in study **III**. The method consisted of five steps:

1. Extracting individual tree segments from the point cloud (step 3 in **Figure 6**)
2. Calculating features for the individual tree segments (step 4 in **Figure 6**).
3. Matching individual tree segments with known locations of dead trees to generate training data (step 5 in **Figure 6**).
4. Training binary classifiers aiming to detect standing dead trees (step 6 in **Figure 6**).
5. Classifying individual tree segments as dead or non-dead segments using the classifiers (step 7 in **Figure 6**).

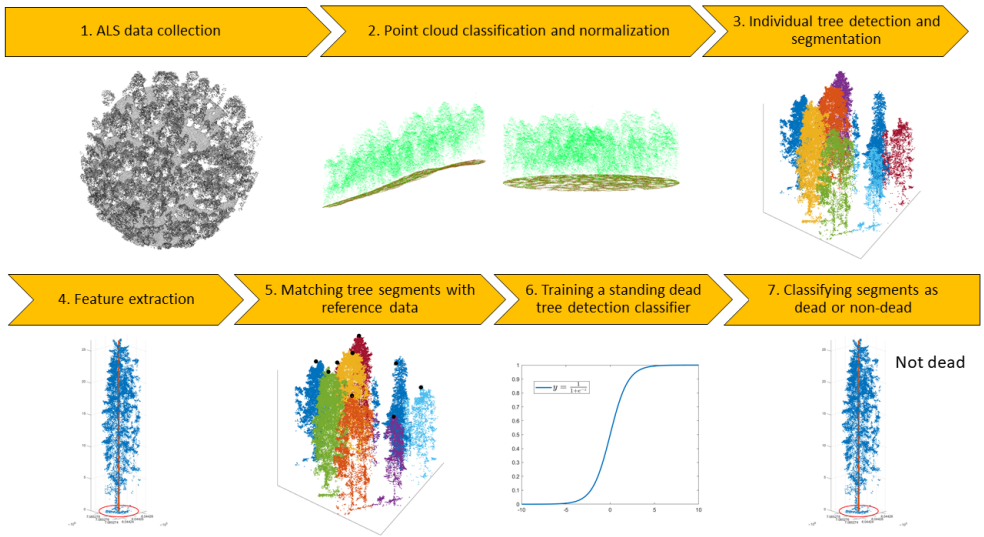

Figure 6 The standing dead tree detection process.

Step 1 was executed using R's lidR package (Roussel et al. 2020; Roussel and Auty 2022), whereas steps 2 and 3 were implemented in MATLAB. The last two steps were performed using the Python programming language and the scikit-learn machine learning library (Pedregosa et al. 2011).

As a first step, individual tree segments were extracted from the height-normalized ULS point cloud using a canopy height model (CHM) based algorithm. First, a CHM with a resolution of 0.5 m was generated using the method by Khosravipour et al. (2014). The method filters the point cloud using multiple height thresholds and generates a CHM using each subset of points. These CHMs are then combined to acquire the final CHM. Next, treetops were detected from the CHM with a method that uses a variable-sized search window whose size is based on the heights of surrounding point cloud points (Popescu and Wynne 2004). Finally, tree segments were delineated using the detected treetops as seed points. A region growing algorithm developed by Dalponte and Coomes (2016) was used for this task.

After extracting individual tree segments, various features were calculated for each segment. Most features described the geometry of the segment as a whole as well as the geometry of the estimated tree crown of the segment. Furthermore, several features described the intensity and point type (single, first, intermediate, last) distributions of the points belonging to the segment.

As a third step, annotated and field-measured standing dead trees (see sections 2.3 and 2.4) were matched with the extracted individual tree segments based on their location. 148 and 75 individual tree segments were matched with annotated and field-measured data, respectively, and labeled as standing dead trees. As binary classifiers require training samples from both classes, all non-matched segments located within the field reference sample plots were labeled as non-dead segments. Furthermore, 592 segments located in proximity of the

annotated dead trees, but not matched with them, were selected randomly and labeled as non-dead segments to be used together with the annotated standing dead trees. As a result, there were two datasets: one consisting of the 148 annotation-based dead tree segments and 592 non-dead segments (referred to as the annotation-based dataset), and another consisting of all segments located within the sample plots and labeled as dead or non-dead (referred to as the field-based dataset). Note that we labeled the segments as dead or non-dead instead of dead and living, as this described the segments not matched with known locations of dead trees more accurately.

Correctly segmenting small trees from the point cloud is a challenging task due to small trees often being occluded by larger trees surrounding them. Thus, the individual tree segments extracted for small trees might not accurately represent the true dimensions of the tree. Using such non-representative segments as training data might hinder the performance of standing dead tree classification. To inspect this effect, the annotation-based dataset and field-based dataset were further divided into three overlapping subsets using three segment height thresholds: 0, 7, and 14 meters. Thus, the total number of training sets was six. Each training set was used for training a classifier aiming to classify individual tree segments as dead or non-dead segments. Each classifier consisted of a feature normalization step, an analysis of variance (ANOVA) based feature selection step, and a logistic regression step. Due to the large number of possibly redundant features, L1 regularization was used within the logistic regression model. The number of features to be selected during the ANOVA based feature selection step, and the strength of L1 regularization were hyperparameters optimized using cross-validation. The area under the precision-recall curve was used as the optimization metric, as it places emphasis on both minimizing the number of false detections and maximizing the number of true detections.

3.4 Performance assessment

3.4.1 The performance of fallen tree detection (studies I-II)

The performance of the fallen tree detection method used in studies **I** and **II** was evaluated against the fallen trees on the field-measured sample plots (see section 2.3). Fallen tree segments detected by the method were first matched with the field-measured fallen trees. Manual matching was used in study **I**, whereas automatic matching based on distances and angles was used in study **II**. Detected fallen tree segments that were matched with a field-measured tree were labeled as true positives (TPs), whereas non-matched segments were labeled as false positives (FPs). Field-measured trees that were not matched with a segment were labeled as false negatives (FNs). The total numbers of TPs, FPs, and FNs were used for calculating the main evaluation metrics – precision and recall (equations 1 and 2). Precision represents the fraction of true detections from all detections, whereas recall represents the fraction of trees that were detected.

$$precision = \frac{\#TPs}{\#TPs + \#FPs} \tag{1}$$

$$recall = \frac{\#TPs}{\#TPs + \#FNs} \quad (2)$$

In study **I**, the general precision and recall of the ALS-based fallen tree detection method were calculated using data from all of the 103 sample plots measured during campaign 1. Furthermore, to inspect the performance of the method for different types of fallen trees, recall was determined for trees of different lengths, diameters, decay classes, and species. Note that tree characteristic specific precisions could not be calculated, as determining the length, diameter, decay class or species of false positives was not possible.

In addition to evaluating the performance of the fallen tree detection method on different types of fallen trees, study **I** inspected how the surrounding vegetation structure affects the detection of fallen trees. Logistic regression models were fitted using sample plot specific vegetation characteristics as independent variables, and sample plot specific precisions and recalls as dependent variables. The coefficients of the fitted models were then used for determining the relationship between each vegetation characteristic and the performance of fallen tree detection. Moreover, the difference in performance was compared between plots located in old-growth forests and plots located in managed forests.

Similarly to study **I**, study **II** evaluated the general precisions and recalls as well as tree characteristic specific recalls. The evaluation was applied separately for ALS- and ULS-based fallen tree detection. The performance metrics were calculated using the 37 sample plots located within the ULS subsites.

3.4.2 Sensitivity analysis (study II)

The fallen tree detection method used in studies **I** and **II** included several modifiable parameters. In study **I**, the values of these parameters were selected based on knowledge gained during method development. In study **II**, a sensitivity analysis was carried out to evaluate the impact of these parameters more comprehensively and to reveal the pros and cons of the method. To limit the computational and analytical complexity, four parameters were selected for sensitivity analysis. **Table 3** presents a description of the parameters. The sensitivity analysis was based on applying the fallen tree detection multiple times - each time changing the parameter values - and evaluating the performance against the field-measured data. Sobol's method (Sobol' 1990, 1993) was used for inspecting the contribution of each parameter to the total variance in performance.

Table 3 A description of the parameters included in sensitivity analysis.

Parameter	Description	Tested values
HR	The height range from which fallen trees are searched (see section 3.2.1)	Lower limit: 0.1, 0.2 and 0.3 m Upper limit: 1 m In total, 3 different height ranges.
CCC	Whether to use connected component classification or not (see section 3.2.2).	0 – connected component classification is not used 1 – connected component classification is used
MNP	The minimum number of point cloud points that must fall on the same line for a line to be detected (see section 3.2.3).	3, 6, 9,…, 30 points. 10 different values in total.
FTR	Whether to use the convolutional neural network that aims to remove false fallen tree detections (see section 3.2.6)	0 – false tree removal is not used 1 – false tree removal is used

3.4.3 The performance of standing dead tree detection (study III)

In study **III**, the performance of the standing dead tree classifiers was assessed using 100 repeats of stratified 5-fold cross-validation on the field-measured data. First, a standing dead tree classifier was trained using the annotated data. Then, at each iteration of cross-validation, $4/5^{ths}$ of the field-measured data was used for training a classifier. The remaining $1/5^{th}$ of field data was used for evaluating the performances of the field data trained classifier and the classifier trained on annotated data. Finally, the results of each iteration were averaged to get the final performance metrics. This process was repeated for all height-thresholded subsets of training data. The evaluated performance metrics were precision (equation 1), recall (equation 2), and Cohen's kappa (equation 3-5).

$$Cohen's\ kappa = \frac{p_o - p_e}{1 - p_e}, \tag{3}$$

where p_o, the relative observed agreement among raters is

$$p_o = \frac{\#TPs + \#TNs}{N} \tag{4}$$

and p_e, the probability of chance agreement is

$$p_e = \frac{\#TPs + \#FPs}{N} \times \frac{\#TPs + \#FNs}{N} + \frac{\#TNs + \#FNs}{N} \times \frac{\#TNs + \#FPs}{N} \quad (5)$$

In the equations, *#TNs* denotes the number of true negatives (i.e., non-dead segments correctly classified as non-dead segments), and $N = \#TPs + \#FPs + \#TNs + \#FNs$.

All three performance metrics were calculated for trees of different heights, whereas only recalls were calculated for trees of different diameters, decay classes, and species, as determining these characteristics from the individual tree segments would have been challenging.

4 RESULTS AND DISCUSSION

4.1 Summary of the results in the original articles

4.1.1 The performance of fallen tree detection (studies I-II)

The fallen tree detection method found 30% of all field-measured fallen trees in study **I**. In addition, the method generated a significant amount of false detections, as only 31% of all detections could be linked with a field-measured tree. The probability of detection (recall) of a fallen tree was highly dependent on the size of the tree. For example, the method detected 75% of all fallen trees over 20 m long (**Figure 7**a) and 73% of all fallen trees with a diameter larger than 250 mm (**Figure 7**b).

In addition to tree size, the species of a fallen tree impacted its detection probability. It seemed that deciduous trees were less likely to be detected than coniferous trees (**Figure 7**d). One reason for this phenomenon could be the differences in the vegetation structure in areas dominated by coniferous trees versus areas dominated by deciduous trees. Study **I** found that the amount of undergrowth was generally smaller in coniferous-dominant areas, allowing more favorable conditions for detecting fallen trees. A further reason could be the branch structure of coniferous versus deciduous trees. Nyström et al. (2014) reported similar results regarding the detection of fallen coniferous and deciduous trees and suggested that one reason for the high detection rate of pines could be their scarce branch structure, as the branches do not prevent laser pulses from hitting the trunk of a fallen tree.

Generally, the tree decay process seemed to reduce the probability of a fallen tree being detected (**Figure 7**c). A likely explanation for this is that as a fallen tree decomposes, its trunk loses its form and settles closer to the ground. Eventually a decomposed tree is covered by moss and other ground vegetation, and becomes very difficult to distinguish from the terrain. Note, however, that the recall of trees in decay class 5 was actually higher than that of less decayed trees. The reason for this anomaly could be that the reference trees in decay class 5 were larger than the trees in other decay classes.

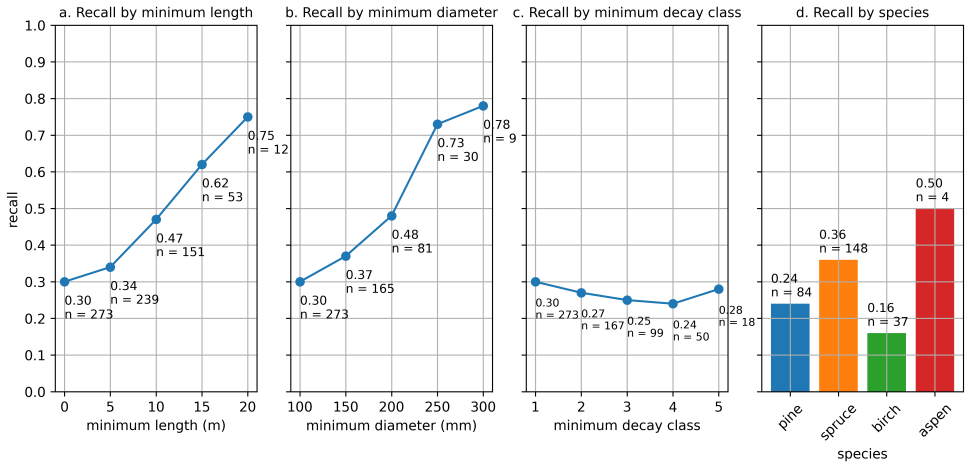

Figure 7 Recall by fallen tree type. Plots a, b, and c show the change in recall when the reference data was filtered using various length, diameter, and decay class thresholds, whereas plot d shows the recall by tree species.

The performance of the fallen tree detection method varied depending on the type of forest in which it was applied. Inspecting the relationship between different vegetation characteristics and performance revealed that the structure of vegetation significantly impacted performance. Firstly, the size (height, and diameter) of surrounding living trees was positively correlated with both precision and recall (**Table 4**), indicating that the method achieved a higher performance in areas with large trees. Secondly, the amount of undergrowth had a notable negative correlation with precision (**Table 4**), indicating that abundant undergrowth hindered the performance of the method. Thirdly, the performance of the fallen tree detection method was significantly higher on sample plots located in old-growth forest as opposed to those located in managed forest. The precision and recall for old-growth sample plots were 0.36 and 0.34, respectively. In contrast, the same metrics for sample plots located in managed forest were 0.02 and 0.04. The reference trees measured on sample plots located in managed forest were, on average, smaller than those located in old-growth forest, which partly explains this phenomenon.

Table 4 Correlation between different vegetation characteristics and the precision and recall of the fallen tree detection method.

Vegetation characteristic	Correlation with precision	Correlation with recall
Living tree height	0.452 ($p < 0.05$)	0.128 ($p < 0.05$)
Living tree DBH	0.474 ($p < 0.05$)	0.164 ($p < 0.05$)
Amount of undergrowth	-0.374 ($p < 0.05$)	-0.004 ($p > 0.05$)

The sensitivity analysis carried out in study **II** revealed that several dataset- and methodology-related factors impacted the performance of fallen tree detection. Firstly, the recalls of ULS-based fallen tree detection were generally higher than those of ALS-based fallen tree detection when no ML-based filters (CCC, FTR, see **Table 3**) were used. This implies that increasing the point density increases the proportion of fallen trees that can be detected. However, the tradeoff was that the higher point density resulted in a significant increase in the number of false detections. This issue was amplified by the circular scan pattern of the ULS dataset, which generated linear groups of points in the point cloud. The minimum number of points parameter (MNP, see **Table 3**) could be adjusted for balancing the tradeoff between precision and recall, but the false positive issue remained even with large values of MNP. Another factor affecting both precision and recall was the height range from which fallen trees were searched (HR, see **Table 3**). Using a higher value for the lower limit of HR increased precision, but decreased recall. Lastly, ML-based filters were able to reduce the number of false detections while also reducing the number of true detections. These filters worked reasonably well for the ALS dataset on which they were originally trained, but their performance on the ULS dataset was poor.

4.1.2 The performance of standing dead tree detection (study III)

The general performance of the standing dead tree detection method developed in study **III** was rather poor. This was, for a large part, due to challenges with individual tree segmentation, which resulted in a large fraction of the segments not representing individual trees. Based on kappa scores (**Figure 8**), the performance of the classifiers was close to random guessing for small trees, but the performance improved as tree height increased. This was especially evident with the classifiers trained using annotated data. For the largest tree class (trees taller than 14 meters), the best classifier trained on annotated data was able to achieve a cross-validated precision, recall, and kappa score of 0.23, 0.48, and 0.17, respectively. The corresponding values for the best classifier trained on field data were 0.17, 0.37, and 0.08.

Filtering training data based on height did not improve the performance of the standing dead tree classifiers. For annotated data, performance decreased while for field data, performance remained similar. This implies that, for annotated data, the reduction of the number of training samples had a more significant negative effect than the increase in the assumed representativeness of the segments.

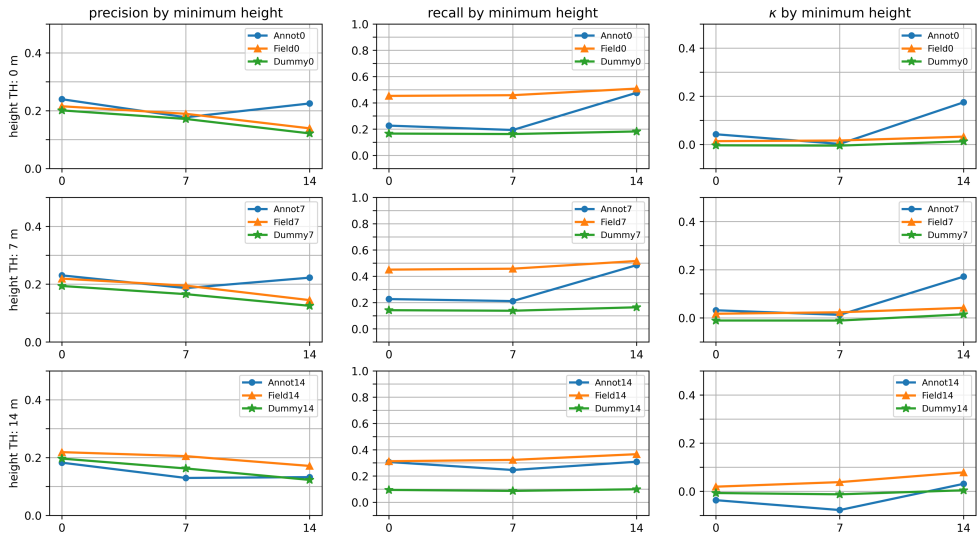

Figure 8 The performance of the standing dead tree classifiers on trees taller than 0, 7, and 14 meters. Each row of plots shows the results for classifiers trained using segments with a height above a specific threshold. For example, the plots on the third row show results for the classifiers trained with segments taller than 14 meters. AnnotX and FieldX denote the classifiers trained using segments taller than X meters in the annotated and field datasets, respectively. DummyX denotes a classifier making random predictions based on the proportion of dead and non-dead segments taller than X meters in the field dataset. κ denotes Cohen's kappa.

4.2 Major findings of the thesis

4.2.1 The majority of large fallen trees can be identified from airborne laser scanning point clouds based on their linear shape (study I)

Study **I** found that the length and diameter of fallen trees were the most significant characteristics affecting their detection probability. The detection rate (recall) of large trees was high, whereas the majority of small trees were left undetected. This result is likely due to large fallen trees standing out from the forest floor, as they are not as likely to be occluded by other near-ground objects as their smaller counterparts. In the point cloud, this means that large trees generate a sufficient number of laser returns for them to be detected. Similar results were reported by Mücke et al. (2013), Nyström et al. (2014), and Polewski et al. (2015b). This phenomenon is rather fortunate from an ecological viewpoint, as large trees have the most ecological significance (Andersson and Hytteborn 1991; Bader et al. 1995). It also raises a question on whether ALS-based fallen tree detection should focus on large trees, considering that detection methods always include a tradeoff between precision and recall. Focusing detection on large trees would allow decreasing the sensitivity of the method, which would reduce the number of false detections.

4.2.2 The structure of the surrounding vegetation impacts the detection of fallen trees (study I)

The performance of the fallen tree detection method was highest on sample plots with large living trees and scarce undergrowth. Two main reasons likely explain this result. Firstly, living tree size is an indicator of fallen tree size, as fallen trees have at some point been part of the living forest. Thus, in general, fallen trees are larger in areas where living trees are large. Large trees, in turn, are more likely visible in the point cloud. Secondly, as fallen trees reside close to the ground, other near-ground objects greatly affect their detection. Dense undergrowth and lush ground vegetation occlude fallen trees thus preventing their detection. Furthermore, line detection is rather sensitive to noise and thus other near-ground objects will generate false fallen tree detections. Similar challenges arising from undergrowth and ground vegetation were identified by Lindberg et al. (2013) and Mücke et al. (2013). In our study, these challenges were especially evident with the sample plots located in managed forests, on which the performance of the fallen tree detection method was very poor. The managed forests in the study area had abundant undergrowth and significant amounts of logging residue (branches and treetops), which resulted in the near-ground portion of the point cloud being full of points not originating from fallen trees (i.e., noise). The fallen tree detection method was not able to separate these points from the fallen tree points, which resulted in the method missing most of the fallen trees while generating a very high number of false detections. One possible solution for addressing the method's inability to separate points originating from fallen trees from other near-ground points would be to use full-waveform laser scanning data. Such data was already used for fallen tree detection by Mücke et al. (2013), who found that the waveform properties of points originating from fallen trees differ from those originating from other near-ground objects.

4.2.3 Fallen tree detection methodology should be adjusted to the properties of the laser scanning dataset available (study II)

Study **II** found that the performance of the fallen tree detection method was sensitive to the dataset used. The method was originally generated for the ALS dataset, and clearly favored this dataset compared to the denser ULS data. The main issue was the method's sensitivity to noise. Hough transform detects a line at any location where a sufficient number of points fall on the same line. It does not consider whether these points form an elongated object by themselves or are rather a part of a non-elongated point patch (see, e.g., Grompone von Gioi et al. 2008). Thus, any sufficiently large uniform point patches will result in line detections regardless of their shape. Hence, large point patches not originating from fallen trees should be removed before applying Hough transform on the point cloud. Already with the sparser ALS dataset, Hough transform's inability to distinguish elongated point patches from non-elongated ones was an issue. This issue was highlighted when the method was applied on the denser ULS data.

The ML-based filters applied before and after the iterative Hough transform step aimed at addressing the method's sensitivity issue and reducing the number of false detections. These filters were trained on samples extracted from the ALS data and seemed to work relatively well on this dataset. However, the filters seemed to lose their discriminative power

when applied on the ULS dataset, as in addition to reducing false detections, they significantly reduced the number of true detections as well. This implies that the two laser scanning datasets were too different to share common ML filters and highlights the importance of using dataset-specific models when utilizing ML.

Most forest inventory-related methods utilizing laser scanning data start by normalizing the heights of the point cloud. This allows a natural way of inspecting the point cloud, as e.g., the heights of different objects are easily measurable and comparable. Normalizing point heights requires extracting the ground from the point cloud (i.e., determining which point cloud points represent ground returns). When standing trees are of interest, the accuracy of ground extraction is not that big a concern, as minor inaccuracies in ground extraction do not significantly distort the properties of the trees. However, when inspecting fallen trees, an accurate ground extraction becomes essential, as it should be able to distinguish near-ground objects from the ground layer. The same observation was made by Mücke et al. (2013) and Polewski et al. (2015b), who highlighted the importance of the ground extraction phase and noted that the sensitivity of the ground extraction algorithm should be adjusted to balance two contradicting factors. Firstly, the algorithm should be sensitive enough to be able to separate near-ground objects from the terrain, as otherwise the objects will be left undetected. Secondly, the algorithm should not be too sensitive, as this might lead to small variations in the terrain being mistakenly identified as fallen trees. Finding a suitable balance between these factors is rather challenging, as the small undulations in the terrain might be of the magnitude as fallen trees. In practice, finding a suitable ground extraction algorithm is based on trial and error, as the suitability of a specific algorithm depends on various properties of the dataset and the area of interest.

Our line detection based fallen tree detection method required that the ground layer be removed in full before detecting fallen trees, as otherwise the ground points would have generated a very high number of false detections. As ground extraction is never perfect, the method used a cutoff height slightly above zero to ensure that the majority of ground points would be removed even if they were incorrectly labeled. Similar cutoff heights were used by earlier studies (Lindberg et al. 2013; Polewski et al. 2015b; Yrttimaa et al. 2019). The value of this cutoff height must be carefully considered, as setting it too high will result in fallen trees being removed from the point cloud, whereas setting it too low will result in a noisy point cloud. Study **II** found that the best cutoff heights were 0.1 m for ALS data and 0.2 for ULS data.

4.2.4 Discrete return ULS data is not sufficient for mapping standing dead trees (study III)

The results of study **III** revealed that discrete return high density laser scanning data does not suffice as the only dataset for standing dead tree detection. The three-step detection process consisting of ITD, feature extraction, and classification relied heavily on the geometric shapes of trees, which introduced several issues. Trees, including living and dead ones, come in a wide variety of shapes and sizes. Furthermore, trees may intertwine and grow beneath each other, which makes a forest a complex environment. This makes extracting individual trees from point clouds a challenging task. This issue is highlighted with standing dead trees. ITD methods have mostly been developed for living trees whose crowns are more homogeneous compared to those of dead trees. Furthermore, the crowns of standing dead trees are often scarce or non-existent. As a result, standing dead trees might not generate a

sufficient number of laser returns for being clearly visible in the point cloud. Moreover, dead crowns do not compete for light and thus living crowns can grow over and around them. Our standing dead tree detection method was based on the assumption that the majority of individual trees can be accurately delineated from the point cloud. This assumption turned out not to be true. As a result, a significant fraction of standing dead trees were not identified in the first place, and the ones that were identified, were not accurately delineated. The latter issue resulted in distorted segments whose geometric features did not represent the true geometric properties of the trees in the forest. Using such segments as training data prevented the ML-based classifiers from learning to distinguish standing dead trees from living ones, and complicated the classification of unseen samples. Combining the ULS data with a spectral dataset (Polewski et al. 2015a; Briechle et al. 2020) or using full-waveform laser scanning (Yao et al. 2012; Amiri et al. 2019; Miltiadou et al. 2020) would have provided additional information that could have helped in correctly segmenting the trees and identifying dead trees.

4.2.5 Annotating dead trees from aerial images provides an efficient means for collecting training data for dead tree detection (study III)

ML-based computer vision methods require that the samples used for training capture the full variety of the objects of interests. When standing dead trees are the objects of interest, a relatively large dataset is needed for fulfilling this requirement due to the complexity of these objects. Annotating dead trees from aerial images provides a means for collecting such a dataset efficiently, but it also introduces a bias, as the bird's eye perspective and inability to collect information from below the canopy only allows identifying standing dead trees that are relatively large and have a crown. The results of study **III** showed, however, that this limitation in the variability of training samples does not actually impact performance, as, even with its canopy penetration capabilities, laser scanning does not allow identifying a wider variety of standing dead trees. Thus, using annotated samples is a viable option, which could be further utilized for eliminating the need for an error-prone rule-based ITD phase. Collecting a training dataset significantly larger than the one used in study **III** would allow shifting to state of the art object detection approaches, which detect objects of interest from an unsegmented point cloud (see Zhou and Tuzel 2017; Lang et al. 2019).

4.3 Constraints and future perspectives

4.3.1 Technological and methodological constraints

This thesis aimed at developing laser scanning based deadwood detection methodology, keeping in mind the operational applicability of these methods. This meant that factors, such as efficiency, large-scale applicability, and generalizability were considered when making methodology-related decisions. It would likely have been possible to improve the performance of the methods by adding complex rules and fine-tuning the parameters of the methods for each sample plot separately, but this would have hindered the operational applicability of these methods and distorted the results. Furthermore, operational applicability was considered in the datasets used, as the methods solely utilized discrete

return laser scanning data, which is the type of laser scanning data most widely available. Unfortunately, it turned out that such data was not sufficient for standing dead tree detection, and spectral information (Polewski et al. 2015a; Amiri et al. 2019) or full-waveform data (Yao et al. 2012) would have been a requirement for a method that could be utilized in practice.

The datasets used in the studies were collected within a two-year time period for practical reasons. Within this time period, new dead trees might have appeared on the study site. As a result, there might have been some differences between the dead trees visible in the RS datasets and the collected reference data. Furthermore, fallen and standing dead tree detection have contradicting requirements regarding the optimal time for data acquisition. Fallen tree detection benefits from a maximized visibility to the ground and thus the laser scanning data should be collected in leaf-off conditions. In contrast, standing dead trees can be more easily identified from living ones in leaf-on conditions, as dead trees do not grow leaves. This is especially true for deciduous trees that died recently, as their only notable difference to living trees might be their inability to grow leaves. The ALS dataset, which was only used for fallen tree detection, was collected during spring 2019, before trees had grown leaves. In contrast, the ULS dataset was collected during the early summer of 2020, when leaf growth was in progress. This data acquisition time was neither optimal for fallen tree detection nor optimal for standing dead tree detection, but aimed at being a suitable compromise for both tasks. Coniferous trees dominated the majority of the sample plots and thus the impact of this compromise was relatively small. However, it is likely that the detection of fallen trees was slightly impaired, as the ground vegetation and deciduous undergrowth were in a more lush state.

The field reference data was collected using a standardized scheme that aimed at acquiring consistent data throughout the study area. However, since the collection of field data involved several people, and some of the measured attributes were not easily quantifiable, there might have been some inconsistencies in the field data. As an example, determining the decay state was based on how easily a knife was able to penetrate the wood, which largely depended on how much pressure was applied on the knife. In addition, the state at which a decaying tree is considered a part of ground is somewhat subjective, and thus similar trees might have been measured at some sample plots, whereas on others they might have been left unmeasured. These slight inconsistencies in the field reference data might have had minor impacts on the results.

The non-dead training segments in the annotation-based dataset used in study **III** were selected as a random set of segments located in the proximity of the annotated dead trees, but not matched with them. Selecting the non-dead samples in such a way likely introduced some errors in the results, as it was possible that some of the non-dead samples actually represented standing dead trees. The reason for using this non-optimal selection approach was that the annotated dataset was collected and provided by another research group, and we did not have access to the UAV imagery from which the standing dead trees were annotated. However, as the annotated data were collected by digitizing the crowns of all visually identifiable dead trees within virtual sample plots, it is safe to assume that the number of erroneous samples in the non-dead class was rather small.

4.3.2 Applicability of the methods and future insights

The methods presented in this thesis were developed for and tested in boreal forests. Although the methods were tested on a study site with varying forest characteristics, the suitability of the methods for other forest biomes cannot be guaranteed. Even within the boreal forest zone, applying the developed methodology will likely require some parameter tuning to optimize the performance for the area of interest. Nevertheless, the main principles of the presented methods should be generalizable for most types of forest. Although the methods include some tunable parameters, their development did not involve plot-level calibrations or complex study-specific rules that limit their generalizability. While this might impair the performance of the methods, it does not restrict their operational applicability.

Future research concerning ALS-based fallen tree detection could address two topics. Firstly, one key question highlighted by the findings of this thesis is how to reduce the number of false detections. The fallen tree detection method presented in this thesis aimed at addressing this question by applying ML-based filters before and after the actual fallen tree detection phase. However, the performance of these filters left room for improvement. The performance could potentially be improved by using intensity- or waveform-related features in addition to geometric features. Secondly, fallen tree detection itself only provides information about the location and number of fallen trees. Estimating further properties, such as the volume of the detected trees is not a trivial task, but would be highly beneficial for many applications of deadwood data.

This thesis highlighted the difficulty of ALS-based standing dead tree detection. There are several ways in which future research could advance this field and overcome the difficulties. Firstly, it would be beneficial to shift to object detection approaches that eliminate the need for the error-prone ITD phase. Secondly, acknowledging that such approaches require large amounts of training data, further research could focus on optimizing ITD for dead tree detection. Thirdly, more research is needed on how full-waveform and multispectral laser scanning can be utilized for detecting standing dead trees.

5 CONCLUSIONS

During the recent decades, remote sensing has achieved an increasingly important role in forest monitoring due to its ability to efficiently acquire information at large scales. Together with aerial and satellite imagery, ALS has been one of the widely utilized RS methods, as it provides detailed 3D information of the target. ALS-based forest inventory methods have largely focused on measuring forest attributes relevant for forestry, whereas ecological applications of ALS have not received as much focus. The global crises – climate change and biodiversity loss – have introduced an increasing need for accurate and up-to-date information related to forest ecology and biodiversity hotspots.

The goal of this thesis was to contribute to the methodology used for collecting accurate and timely information regarding the ecological status of forests. The thesis focused on developing laser scanning based deadwood detection methodology. Furthermore, the thesis provided insight on dataset-, methodology-, and forest structure-related factors that affect the

accuracy of deadwood mapping using laser scanning. The thesis aimed at approaching the topic with an operational mindset. Thus, the methodological decisions favored efficiency and large-scale applicability. Furthermore, the methods were evaluated against ground truth, thus providing unbiased estimates of their operational accuracy.

Studies **I** and **II** confirmed that current ALS devices are capable of collecting information with a level of detail sufficient for mapping individual fallen trees. The studies found that, while small fallen trees are often not visible in ALS data, the majority of large fallen trees can be identified based on their linear shape. However, undergrowth and other near-ground objects generate false detections, the number of which depends on the sensitivity of the detection methodology used. The aforementioned results suggest that ALS-based fallen tree mapping should focus on the detection of large trees, as aiming to detect trees of all sizes requires overly sensitive detection methodology that is prone to false detections. This would ensure that deadwood information collected using ALS is, although incomplete, as correct as possible, which would allow using such information in decision-making. This conclusion is supported by the fact that large dead trees have the most ecological significance.

Study **III** found that discrete return ALS data does not suffice for mapping individual standing dead trees, even if the data has a high point density. The bird's eye perspective limits the information about the vertical structure of trees, which complicates the identification of individual trees and prevents observing their true geometrical properties. This is especially true for dead trees, whose crowns are often scarce or non-existent. Spectral information would help in distinguishing individual trees, but the problem is that imagery collected from the bird's eye perspective only captures the top-most canopy layer. Multi-spectral and full-waveform laser scanning are viable options, as they provide detailed information about the reflectance of targets while retaining the ability to capture information from below the top-most canopy layer. These laser scanning methods have already been successfully utilized for tree species classification and the detection of dead trees (Yao et al. 2012; Mücke et al. 2013; Amiri et al. 2019). As these methods become more common, we may see a rapid development in individual tree level applications of ALS data.

The detection methods used in this thesis were a combination of rule-based approaches and machine learning. The advantage of rule-based approaches is that it is easy to incorporate domain knowledge into these approaches. However, the disadvantage is that rule-based approaches involve a number of human-made decisions and a set of parameters that need to be tuned. These decisions and parameters greatly affect the performance of the whole method and often need to be adjusted when applying the method to different areas. Furthermore, a single poorly made decision or poorly tuned parameter may result in the performance of the whole method being drastically reduced. For example, the main reason for the poor performance of the standing dead tree detection method was that the rule-based ITD was not able to delineate individual trees accurately enough. Shifting to fully ML-based object detection approaches would largely reduce the need for manually crafted rules that are highly sensitive to changes in the characteristics of the forest. The field of computer vision has undergone a paradigm shift and almost all state of the art object detection approaches are fully based on ML and deep learning. This paradigm shift has already revolutionized some fields, such as autonomous driving, and can be expected to land into forest-related applications as well.

ML-based object detection approaches require large amounts of training data. Collecting sufficient amounts of training data using conventional field inventory methods is not feasible,

and more efficient approaches are required. Annotating training samples from RS datasets improves the efficiency of training data collection significantly. The efficiency of annotation can be further improved by introducing semi-automatic annotation procedures in which, for example, the annotator identifies a part belonging to an object of interest and an automatic segmentation procedure determines the other parts belonging to the object.

This thesis investigated ALS-based deadwood mapping at an individual tree level. The thesis contributed to the understanding of the applicability of ALS for deadwood mapping and provided insight on the challenges related to such a mapping approach. The thesis showed that individual fallen trees can be identified from discrete return ALS data, whereas standing dead trees require different types of datasets. The results of this thesis showed that ALS-based deadwood mapping is far from a fully resolved topic. However, further developments in laser scanning technology and deadwood detection methodology could enable mapping deadwood with an accuracy sufficient for operational purposes.

REFERENCES

Amiri N, Krzystek P, Heurich M, Skidmore A (2019) Classification of tree species as well as standing dead trees using triple wavelength als in a temperate forest. Remote Sens. 11(22): 2614. https://doi.org/10.3390/rs11222614.

Andersson LI, Hytteborn H (1991) Bryophytes and decaying wood– a comparison between managed and natural forest. Ecography 14(2): 121-130. https://doi.org/https://doi.org/10.1111/j.1600-0587.1991.tb00642.x.

Axelsson P (2000) Dem generation from laser scanner data using adaptive tin models. Int. Arch. Photogramm. Remote Sens. 33(4): 110-117.

Bader P, Jansson S, Jonsson BG (1995) Wood-inhabiting fungi and substratum decline in selectively logged boreal spruce forests. Biol. Conserv. 72(3): 355-362. https://doi.org/https://doi.org/10.1016/0006-3207(94)00029-P.

Baltsavias EP (1999) Airborne laser scanning: Basic relations and formulas. ISPRS J. Photogramm. Remote Sens. 54(2): 199-214. https://doi.org/https://doi.org/10.1016/S0924-2716(99)00015-5.

Bater C, Coops N, Gergel S, Goodwin N (2007) Towards the estimation of tree structural class in northwest coastal forests using lidar remote sensing. ISPRS Workshop on Laser Scanning 2007 and Silvilaser 2007, Espoo, Finland, 12-14 September 2007.

Blanchard SD, Jakubowski MK, Kelly M (2011) Object-based image analysis of downed logs in disturbed forested landscapes using lidar. Remote Sens. 3(11): 2420-2439. https://doi.org/https://doi.org/10.3390/rs3112420.

Boddy L, Frankland JC, West Pv (2008) Ecology of saprotrophic basidiomycetes. Elsevier.

Bradford MA, Warren Ii RJ, Baldrian P, Crowther TW, Maynard DS, Oldfield EE, Wieder WR, Wood SA, King JR (2014) Climate fails to predict wood decomposition at regional scales. Nat. Clim. Change 4(7): 625-630. https://doi.org/10.1038/nclimate2251.

Briechle S, Krzystek P, Vosselman G (2020) Classification of tree species and standing dead trees by fusing uav-based lidar data and multispectral imagery in the 3d deep neural network pointnet++. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. V-2-2020: 203-210. https://doi.org/10.5194/isprs-annals-V-2-2020-203-2020.

Briechle S, Krzystek P, Vosselman G (2021) Silvi-net – a dual-cnn approach for combined classification of tree species and standing dead trees from remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 98: 102292. https://doi.org/https://doi.org/10.1016/j.jag.2020.102292.

Brunelli R (2009) Template matching techniques in computer vision : Theory and practice. Wiley.

Calderon S, Boubekeur T (2014) Point morphology. ACM Trans. Graph. 33(4): Article 45. https://doi.org/10.1145/2601097.2601130.

Casas Á, García M, Siegel RB, Koltunov A, Ramírez C, Ustin S (2016) Burned forest characterization at single-tree level with airborne laser scanning for assessing wildlife habitat. Remote Sens. Environ. 175: 231-241. https://doi.org/https://doi.org/10.1016/j.rse.2015.12.044.

Chisholm RA, Cui J, Lum SKY, Chen BM (2013) Uav lidar for below-canopy forest surveys. J. Unmanned Veh. Syst. 01(01): 61-68. https://doi.org/https://doi.org/10.1139/juvs-2013-0017.

Clawges R, Vierling K, Vierling L, Rowell E (2008) The use of airborne lidar to assess avian species diversity, density, and occurrence in a pine/aspen forest. Remote Sens. Environ. 112(5): 2064-2073. https://doi.org/https://doi.org/10.1016/j.rse.2007.08.023.

Dalponte M, Ørka HO, Ene LT, Gobakken T, Næsset E (2014) Tree crown delineation and tree species classification in boreal forests using hyperspectral and als data. Remote Sens. Environ. 140(2014): 306-317. https://doi.org/https://doi.org/10.1016/j.rse.2013.09.006.

Dalponte M, Coomes DA (2016) Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods Ecol. Evol. 7(10): 1236-1245. https://doi.org/https://doi.org/10.1111/2041-210X.12575.

Ducey MJ, Williams MS, Gove JH, Roberge S, Kenning RS (2012) Distance-limited perpendicular distance sampling for coarse woody debris: Theory and field results. Forestry 86(1): 119-128. https://doi.org/https://doi.org/10.1093/forestry/cps059.

Duda R, Hart P (1972) Use of the hough transformation to detect lines and curves in pictures. Commun. ACM 15(1): 11-15. https://doi.org/https://doi.org/10.1145/361237.361242.

European Commission Directorate General for Environment (2021) Eu biodiversity strategy for 2030 : Bringing nature back into our lives. Publications Office of the European Union. https://doi.org/doi/10.2779/677548

FAO (2022) The state of the world's forests 2022. Forest pathways for green recovery and building inclusive, resilient and sustainable economies. https://doi.org/https://doi.org/10.4060/cb9360en

Filin S, Pfeifer N (2005) Neighborhood systems for airborne laser data. Photogramm. Eng. Remote Sens. 71(6): 743-755. https://doi.org/10.14358/PERS.71.6.743

Fischler MA, Bolles RC (1981) Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24(6): 381–395. https://doi.org/10.1145/358669.358692.

Gaston KJ, Spicer JI (2004) Biodiversity : An introduction (2nd ed.). Blackwell Publishing.

Gobakken T, Næsset E (2004) Effects of forest growth on laser derived canopy metrics. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 36(W2): 224-227.

Goetz S, Steinberg D, Dubayah R, Blair B (2007) Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sens. Environ. 108(3): 254-263. https://doi.org/https://doi.org/10.1016/j.rse.2006.11.016.

Gonzalez RC, Woods RE (2008) Digital image processing (3rd ed.). Prentice Hall.

Grompone von Gioi R, Jakubowicz J, Morel J-M, Randall G (2008) On straight line segment detection. J. Math. Imaging Vis. 32(3): 313-347. https://doi.org/10.1007/s10851-008-0102-5.

Harmon ME, Franklin JF, Swanson FJ, Sollins P, Gregory SV, Lattin JD, Anderson NH, Cline SP, Aumen NG, Sedell JR, Lienkaemper GW, Cromack K, Cummins KW. (1986). Ecology of coarse woody debris in temperate ecosystems. In: MacFadyen A & Ford ED (eds) Adv. Ecol. Res. Academic Press, pp. 133-302.

Heinaro E (2021) Fallen_tree_detection. GitHub repository. https://github.com/Eikka12/Fallen_tree_detection.git

Heinzel J, Koch B (2011) Exploring full-waveform lidar parameters for tree species classification. Int. J. Appl. Earth Obs. Geoinf. 13(1): 152-160. https://doi.org/https://doi.org/10.1016/j.jag.2010.09.010.

Hough PVC (1962). Method and means for recognizing complex patterns (United States Patent No. 3069654). https://www.osti.gov/biblio/4746348

Hugging Face. Hugging face - the ai community building the future. Retrieved 28.3.2023 from https://huggingface.co/

Huo L, Strengbom J, Lundmark T, Westerfelt P, Lindberg E (2023) Estimating the conservation value of boreal forests using airborne laser scanning. Ecol. Indicators 147: 109946. https://doi.org/https://doi.org/10.1016/j.ecolind.2023.109946.

Hyyppä E, Hyyppä J, Hakala T, Kukko A, Wulder MA, White JC, Pyörälä J, Yu X, Wang Y, Virtanen J-P, Pohjavirta O, Liang X, Holopainen M, Kaartinen H (2020) Under-canopy uav laser scanning for accurate forest field measurements. ISPRS J. Photogramm. Remote Sens. 164: 41-60. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.03.021.

Hyyppä J, Inkinen M (1999) Detecting and estimating attributes for single trees using laser scanner. Photogramm. J. Finland 16: 27-42.

Hyyppä J, Kelle O, Lehikoinen M, Inkinen M (2001) A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. IEEE Trans. Geosci. Remote Sens. 39(5): 969-975. https://doi.org/10.1109/36.921414.

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

Jaakkola A, Hyyppä J, Kukko A, Yu X, Kaartinen H, Lehtomäki M, Lin Y (2010) A low-cost multi-sensoral mobile mapping system and its feasibility for tree measurements. ISPRS J. Photogramm. Remote Sens. 65(6): 514-522. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2010.08.002.

Jaakkola A, Hyyppä J, Yu X, Kukko A, Kaartinen H, Liang X, Hyyppä H, Wang Y (2017) Autonomous collection of forest field reference—the outlook and a first step with uav laser scanning. Remote Sens. 9(8): 785. https://doi.org/https://doi.org/10.3390/rs9080785.

Jakubowski MK, Guo Q, Kelly M (2013a) Tradeoffs between lidar pulse density and forest measurement accuracy. Remote Sens. Environ. 130: 245-253. https://doi.org/https://doi.org/10.1016/j.rse.2012.11.024.

Jakubowski MK, Li W, Guo Q, Kelly M (2013b) Delineating individual trees from lidar data: A comparison of vector- and raster-based segmentation approaches. Remote Sens. 5(9): 4163-4186. https://doi.org/10.3390/rs5094163.

Jianbo S, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 22(8): 888-905. https://doi.org/10.1109/34.868688.

Jonsson BG, Kruys N, Ranius T (2005) Ecology of species living on dead wood – lessons for dead wood management. Silva Fenn. 39(2): 289-309. https://doi.org/doi:10.14214/sf.390.

Kaartinen H, Hyyppä J, Yu X, Vastaranta M, Hyyppä H, Kukko A, Holopainen M, Heipke C, Hirschmugl M, Morsdorf F, Næsset E, Pitkänen J, Popescu S, Solberg S, Wolf BM, Wu J-C (2012) An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sens. 4(4): 950-974. https://doi.org/https://doi.org/10.3390/rs4040950.

Kangas A, Maltamo M (2006) Forest inventory : Methodology and applications. Springer.

Ke Y, Quackenbush LJ (2011) A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. Int. J. Remote Sens. 32(17): 4725-4747. https://doi.org/10.1080/01431161.2010.494184.

Khosravipour A, Skidmore A, Isenburg M, Hussin Y (2014) Generating pit-free canopy height models from airborne lidar. Photogramm. Eng. Remote Sens. 80: 863-872. https://doi.org/10.14358/PERS.80.9.863.

Kim Y, Yang Z, Cohen WB, Pflugmacher D, Lauver CL, Vankat JL (2009) Distinguishing between live and dead standing tree biomass on the north rim of grand canyon national

park, USA using small-footprint lidar data. Remote Sens. Environ. 113(11): 2499-2510. https://doi.org/https://doi.org/10.1016/j.rse.2009.07.010.

Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 25.

Landrieu L, Simonovsky M (2018) Large-scale point cloud semantic segmentation with superpoint graphs. arXiv 1711.09869. [Preprint]. http://arxiv.org/abs/1711.09869. Accessed 15 April 2023.

Lang AH, Vora S, Caesar H, Zhou L, Yang J, Beijbom O (2019) Pointpillars: Fast encoders for object detection from point clouds. arXiv 1812.05784. [Preprint]. http://arxiv.org/abs/1812.05784. Accessed 15 April 2023.

Lassauce A, Paillet Y, Jactel H, Bouget C (2011) Deadwood as a surrogate for forest biodiversity: Meta-analysis of correlations between deadwood volume and species richness of saproxylic organisms. Ecol. Indicators 11(5): 1027-1039. https://doi.org/https://doi.org/10.1016/j.ecolind.2011.02.004.

Li W, Guo Q, Jakubowski M, Kelly M (2012) A new method for segmenting individual trees from the lidar point cloud. Photogramm. Eng. Remote Sens. 78: 75-84. https://doi.org/10.14358/PERS.78.1.75.

Lim K, Treitz P, Wulder M, St-Onge B, Flood M (2003) Lidar remote sensing of forest structure. Prog. Phys. Geogr. 27(1): 88-106. https://doi.org/10.1191/0309133303pp360ra.

Lindberg E, Hollaus M, Mücke W, Fransson JES, Pfeifer N (2013) Detection of lying tree stems from airborne laser scanning data using a line template matching algorithm. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. II-5/W2: 169-174. https://doi.org/https://doi.org/10.5194/isprsannals-ii-5-w2-169-2013.

Lu X, Guo Q, Li W, Flanagan J (2014) A bottom-up approach to segment individual deciduous trees using leaf-off lidar point cloud data. ISPRS J. Photogramm. Remote Sens. 94: 1-12. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2014.03.014.

Luonnonvarakeskus. Valtakunnan metsien inventointi (vmi). Retrieved 20.3.2023 from https://www.luke.fi/fi/seurannat/valtakunnan-metsien-inventointi-vmi

MacQueen JB, Neyman J (1967) Some methods for classification and analysis of multivariate observations. In: Lecam L & Neyman J (eds) Proceedings of 5th Berkeley symposium on mathematical statistics and probability, Berkeley and Los Angeles, CA, USA, 21 June - 18 July 1965 and 27 December 1965 - 7 January 1966. pp 281-297.

Metsähallitus. (2018). Beetles life. Retrieved 20.3.2023 from https://www.metsa.fi/en/project/beetles-life-eng/

Metsäntutkimuslaitos. (2009). Vmi11. Retrieved 20.5.2019 from http://www.metla.fi/ohjelma/vmi/vmi11-maasto-ohje09-2p.pdf

Miltiadou M, Campbell NDF, Gonzalez Aracil S, Brown T, Grant MG (2018) Detection of dead standing eucalyptus camaldulensis without tree delineation for managing biodiversity in native australian forest. Int. J. Appl. Earth Obs. Geoinf. 67: 135-147. https://doi.org/https://doi.org/10.1016/j.jag.2018.01.008.

Miltiadou M, Agapiou A, Gonzalez Aracil S, Hadjimitsis DG (2020) Detecting dead standing eucalypt trees from voxelised full-waveform lidar using multi-scale 3d-windows for tackling height and size variations. Forests 11(2): 161. https://doi.org/https://doi.org/10.3390/f11020161.

Miura N, Jones SD (2010) Characterizing forest ecological structure using pulse types and heights of airborne laser scanning. Remote Sens. Environ. 114(5): 1069-1076. https://doi.org/https://doi.org/10.1016/j.rse.2009.12.017.

Morsdorf F, Meier EH, Allgöwer B, Nüesch D (2003) Clustering in airborne laser scanning raw data for segmentation of single trees. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 34(part 3): W13.

Mücke W, Deák B, Schroiff A, Hollaus M, Pfeifer N (2013) Detection of fallen trees in forested areas using small footprint airborne laser scanning data. Can. J. Remote Sens. 39(sup1): S32-S40. https://doi.org/https://doi.org/10.5589/m13-013.

Müller J, Brandl R (2009) Assessing biodiversity by remote sensing in mountainous terrain: The potential of lidar to predict forest beetle assemblages. J. Appl. Ecol. 46(4): 897-905. https://doi.org/https://doi.org/10.1111/j.1365-2664.2009.01677.x.

Næsset E (1997a) Estimating timber volume of forest stands using airborne laser scanner data. Remote Sens. Environ. 61(2): 246-253. https://doi.org/https://doi.org/10.1016/S0034-4257(97)00041-2.

Næsset E (1997b) Determination of mean tree height of forest stands using airborne laser scanner data. ISPRS J. Photogramm. Remote Sens. 52(2): 49-56. https://doi.org/https://doi.org/10.1016/S0924-2716(97)83000-6.

Næsset E (2002) Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ. 80(1): 88-99. https://doi.org/https://doi.org/10.1016/S0034-4257(01)00290-5.

National Land Survey of Finland. Laserkeilaus ja ilmakuvaus. Retrieved 27.4.2023 from https://www.maanmittauslaitos.fi/laserkeilaus-ja-ilmakuvaus

Nyström M, Holmgren J, Fransson JES, Olsson H (2014) Detection of windthrown trees using airborne laser scanning. Int. J. Appl. Earth Obs. Geoinf. 30: 21-29. https://doi.org/https://doi.org/10.1016/j.jag.2014.01.012.

Näslund M (1936) Skogsförsöksanstaltens gallringsförsök i tallskog. Meddelanden från Statens Skogsförsöksanstalt 29.

Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA, Phillips OL, Shvidenko A, Lewis SL, Canadell JG, Ciais P, Jackson RB, Pacala SW, McGuire AD, Piao S, Rautiainen A, Sitch S, Hayes D (2011) A large and persistent carbon sink in the world's forests. Science 333(6045): 988-993. https://doi.org/10.1126/science.1201609.

Pasher J, King DJ (2009) Mapping dead wood distribution in a temperate hardwood forest using high resolution airborne imagery. For. Ecol. Manag. 258(7): 1536-1548. https://doi.org/https://doi.org/10.1016/j.foreco.2009.07.009.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12: 2825-2830.

Persson A, Holmgren J, Soderman U (2002) Detecting and measuring individual trees using an airborne laser scanner. Photogramm. Eng. Remote Sens. 68(9): 925--932.

Pesonen A, Kangas A, Maltamo M, Packalén P (2010a) Effects of auxiliary data source and inventory unit size on the efficiency of sample-based coarse woody debris inventory. For. Ecol. Manag. 259(10): 1890-1899. https://doi.org/https://doi.org/10.1016/j.foreco.2010.02.001.

Pesonen A, Maltamo M, Kangas A (2010b) The comparison of airborne laser scanning-based probability layers as auxiliary information for assessing coarse woody debris. Int. J. Remote Sens. 31(5): 1245-1259. https://doi.org/10.1080/01431160903380607.

Polewski P, Yao W, Heurich M, Krzystek P, Stilla U (2015a) Active learning approach to detecting standing dead trees from als point clouds combined with aerial infrared imagery. 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, USA, 7-12 June 2015. pp 10-18. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7301378.

Polewski P, Yao W, Heurich M, Krzystek P, Stilla U (2015b) Detection of fallen trees in als point clouds using a normalized cut approach trained by simulation. ISPRS J. Photogramm. Remote Sens. 105: 252-271. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2015.01.010.

Polewski P, Yao W, Heurich M, Krzystek P, Stilla U (2015c) Free shape context descriptors optimized with genetic algorithm for the detection of dead tree trunks in als point clouds. In: Mallet C, Paparoditis N, Dowman I, Oude Elberink S, Raimond AM, Rottensteiner F, Yang M, Christophe S, Çöltekin A, & Brédif M (eds) Isprs Geospatial Week 2015, La Grande Motte, France, 28 September - 3 October 2015. pp 41-48. https://doi.org/10.5194/isprsannals-II-3-W5-41-2015.

Polewski P (2017) Reconstruction of standing and fallen single dead trees in forested areas from lidar data and aerial imagery. Technische Universität München.

Polewski P, Yao W, Heurich M, Krzystek P, Stilla U (2018) Learning a constrained conditional random field for enhanced segmentation of fallen trees in als point clouds. ISPRS J. Photogramm. Remote Sens. 140: 33-44. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2017.04.001.

Popescu SC, Wynne RH, Nelson RF (2003) Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Can. J. Remote Sens. 29(5): 564-577. https://doi.org/10.5589/m03-027.

Popescu SC, Wynne RH (2004) Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogramm. Eng. Remote Sens. 70: 589-604.

Reitberger J, Schnörr C, Krzystek P, Stilla U (2009) 3d segmentation of single trees exploiting full waveform lidar data. ISPRS J. Photogramm. Remote Sens. 64(6): 561-574. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2009.04.002.

Roussel J-R, Auty D, Coops NC, Tompalski P, Goodbody TRH, Meador AS, Bourdon J-F, de Boissieu F, Achim A (2020) Lidr: An r package for analysis of airborne laser scanning (als) data. Remote Sens. Environ. 251: 112061. https://doi.org/https://doi.org/10.1016/j.rse.2020.112061.

Roussel J-R, Auty D. (2022). Airborne lidar data manipulation and visualization for forestry applications. Retrieved 10.9.2022 from https://cran.r-project.org/package=lidR

Seibold S, Rammer W, Hothorn T, Seidl R, Ulyshen MD, Lorz J, Cadotte MW, Lindenmayer DB, Adhikari YP, Aragón R, Bae S, Baldrian P, Barimani Varandi H, Barlow J, Bässler C, Beauchêne J, Berenguer E, Bergamin RS, Birkemoe T, . . . Müller J (2021) The contribution of insects to global forest deadwood decomposition. Nature 597(7874): 77-81. https://doi.org/10.1038/s41586-021-03740-8.

Siitonen J (2001) Forest management, coarse woody debris and saproxylic organisms: Fennoscandian boreal forests as an example. Ecol. Bull.(49): 11-41. http://www.jstor.org/stable/20113262

Sobol' I (1990) On sensitivity estimation for nonlinear mathematical models. Matematicheskoe Modelirovanie 2(1): 112-118.

Sobol' I (1993) Sensitivity analysis for non-linear mathematical models. Math. Model. Comput. Exp. 1993(1): 407-414.

Stokland JN, Jonsson BG, Siitonen J (2012) Biodiversity in dead wood. Cambridge University Press.

Ståhl G, Ringvall A, Fridman J (2001) Assessment of coarse woody debris: A methodological overview. Ecol. Bull.(49): 57-70. http://www.jstor.org/stable/20113264

Ståhl G, Gove JH, Williams MS, Ducey MJ (2010) Critical length sampling: A method to estimate the volume of downed coarse woody debris. Eur. J. For. Res. 129(6): 993-1000. https://doi.org/https://doi.org/10.1007/s10342-010-0382-3.

Tanhuanpää T, Kankare V, Vastaranta M, Saarinen N, Holopainen M (2015) Monitoring downed coarse woody debris through appearance of canopy gaps in urban boreal forests with bitemporal als data. Urban For. Urban Green. 14(4): 835-843. https://doi.org/https://doi.org/10.1016/j.ufug.2015.08.005.

Vastaranta M, Korpela I, Uotila A, Hovi A, Holopainen M (2012) Mapping of snow-damaged trees based on bitemporal airborne lidar data. Eur. J. For. Res. 131(4): 1217-1228. https://doi.org/10.1007/s10342-011-0593-2.

Vauhkonen J, Tokola T, Maltamo M, Packalén P (2008) Effects of pulse density on predicting characteristics of individual trees of scandinavian commercial species using alpha shape metrics based on airborne laser scanning data. Can. J. Remote Sens. 34(sup2): S441-S459. https://doi.org/10.5589/m08-052.

Vincent L, Soille P (1991) Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13(6): 583-598. https://doi.org/10.1109/34.87344.

Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E (2018) Deep learning for computer vision: A brief review. Comput. Intell. Neurosci. 2018: 7068349. https://doi.org/10.1155/2018/7068349.

Wallace L, Lucieer A, Watson C, Turner D (2012) Development of a uav-lidar system with application to forest inventory. Remote Sens. 4(6): 1519-1543. https://doi.org/https://doi.org/10.3390/rs4061519.

Wallace L, Musk R, Lucieer A (2014) An assessment of the repeatability of automatic forest inventory metrics derived from uav-borne laser scanning data. IEEE Trans. Geosci. Remote Sens. 52(11): 7160-7169. https://doi.org/https://doi.org/10.1109/TGRS.2014.2308208.

Wang W, Yu R, Huang Q, Neumann U (2019) Sgpn: Similarity group proposal network for 3d point cloud instance segmentation. arXiv 1711.08588 [Preprint]. http://arxiv.org/abs/1711.08588. Accessed 23 March 2023.

Wang Y, Hyyppä J, Liang X, Kaartinen H, Yu X, Lindberg E, Holmgren J, Qin Y, Mallet C, Ferraz A, Torabzadeh H, Morsdorf F, Zhu L, Liu J, Alho P (2016) International benchmarking of the individual tree detection methods for modeling 3-d canopy structure for silviculture and forest ecology using airborne laser scanning. IEEE Trans. Geosci. Remote Sens. 54(9): 5011-5027. https://doi.org/10.1109/TGRS.2016.2543225.

Wehr A, Lohr U (1999) Airborne laser scanning—an introduction and overview. ISPRS J. Photogramm. Remote Sens. 54(2): 68-82. https://doi.org/https://doi.org/10.1016/S0924-2716(99)00011-8.

Weiss K, Khoshgoftaar TM, Wang D (2016) A survey of transfer learning. J. Big Data 3(1): 9. https://doi.org/10.1186/s40537-016-0043-6.

Wing BM, Ritchie MW, Boston K, Cohen WB, Olsen MJ (2015) Individual snag detection using neighborhood attribute filtered airborne lidar data. Remote Sens. Environ. 163: 165-179. https://doi.org/https://doi.org/10.1016/j.rse.2015.03.013.

Xie Y, Tian J, Zhu XX (2020) Linking points with labels in 3d: A review of point cloud semantic segmentation. IEEE Geosci. Remote Sens. 8(4): 38-59. https://doi.org/10.1109/MGRS.2019.2937630.

Xu Y, Tong X, Stilla U (2021) Voxel-based representation of 3d point clouds: Methods, applications, and its potential use in the construction industry. Autom. Constr. 126: 103675. https://doi.org/https://doi.org/10.1016/j.autcon.2021.103675.

Yao W, Krzystek P, Heurich M (2012) Identifying standing dead trees in forest areas based on 3d single tree detection from full waveform lidar data. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. I-7: 359-364. https://doi.org/10.5194/isprsannals-I-7-359-2012.

Yrttimaa T, Saarinen N, Luoma V, Tanhuanpää T, Kankare V, Liang X, Hyyppä J, Holopainen M, Vastaranta M (2019) Detecting and characterizing downed dead wood using terrestrial laser scanning. ISPRS J. Photogramm. Remote Sens. 151: 76-90. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2019.03.007.

Yu X, Hyyppä J, Kukko A, Maltamo M, Kaartinen H (2006) Change detection techniques for canopy height growth measurements using airborne laser scanner data. Photogramm. Eng. Remote Sens. 72(12): 1339-1348. https://doi.org/10.14358/PERS.72.12.1339.

Zaidi SSA, Ansari MS, Aslam A, Kanwal N, Asghar M, Lee B (2022) A survey of modern deep learning based object detection models. Digit. Signal Process. 126: 103514. https://doi.org/https://doi.org/10.1016/j.dsp.2022.103514.

Zhang K, Chen S-C, Whitman D, Shyu M-L, Yan J, Zhang C (2003) A progressive morphological filter for removing nonground measurements from airborne lidar data. IEEE Trans. Geosci. Remote Sens. 41(4): 872-882. https://doi.org/https://doi.org/10.1109/TGRS.2003.810682.

Zhang W, Qi J, Wan P, Wang H, Xie D, Wang X, Yan G (2016) An easy-to-use airborne lidar data filtering method based on cloth simulation. Remote Sens. 8(6): 501. https://www.mdpi.com/2072-4292/8/6/501

Zhou Y, Tuzel O (2017) Voxelnet: End-to-end learning for point cloud based 3d object detection. arXiv: 1711.06396 [Preprint]. https://doi.org/https://doi.org/10.48550/arXiv.1711.06396.

Zou Z, Chen K, Shi Z, Guo Y, Ye J (2023) Object detection in 20 years: A survey. Proc. IEEE 111(3): 257-276. https://doi.org/10.1109/JPROC.2023.3238524.