UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
INSTITUTO DE INFORMÁTICA (INF)
PROGRAMA DE PÓS–GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

WELINGTON GALVÃO RODRIGUES

Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas

GOIÂNIA
2025

UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO (TECA) PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TESES

E DISSERTAÇÕES NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo das Teses e Dissertações disponibilizado na BDTD/UFG é de responsabilidade exclusiva do autor. Ao encaminhar o produto final, o autor(a) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do material bibliográfico

[ ] Dissertação [X] Tese [ ] Outro*:_____

*No caso de mestrado/doutorado profissional, indique o formato do Trabalho de Conclusão de Curso, permitido no documento de área, correspondente ao programa de pós-graduação, orientado pela legislação vigente da CAPES.

Exemplos: Estudo de caso ou Revisão sistemática ou outros formatos.

2. Nome completo do autor

Welington Galvão Rodrigues

3. Título do trabalho

Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas

4. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador)

Concorda com a liberação total do documento [X] SIM [ ] NÃO[1]

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante:
a) consulta ao(à) autor(a) e ao(à) orientador(a);
b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo da tese ou dissertação.
O documento não será disponibilizado durante o período de embargo.
Casos de embargo:
- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro;
- Publicação da dissertação/tese em livro.

Obs. Este termo deverá ser assinado no SEI pelo orientador e pelo autor.

Documento assinado eletronicamente por **Fabrizzio Alphonsus Alves De Melo Nunes Soares** , **Professor do Magistério Superior**, em 26/11/2025, às 14:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020 .

Documento assinado eletronicamente por **Welington Galvão Rodrigues**, **Discente**, em 27/11/2025, às 14:39, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5761171** e o código CRC **732CEF5F**.

Referência: Processo nº 23070.041055/2025-16 SEI nº 5761171

WELINGTON GALVÃO RODRIGUES

Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas

Tese apresentada ao Programa de Pós–Graduação em Ciência da Computação, do Instituto de Informática (INF), da Universidade Federal de Goiás (UFG), como requisito para obtenção do título de Doutor em Ciência da Computação.
Área de Concentração: Ciência da Computação.
Linha de Pesquisa: Sistemas Inteligentes e Aplicações.

Orientador: Prof. Dr. Fabrizzio Alphonsus Alves de Melo Nunes Soares
Co-Orientador: Prof. Dr. Gabriel da Silva Vieira

GOIÂNIA
2025

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Rodrigues, Welington Galvão
Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas. [manuscrito] / Welington Galvão Rodrigues. - 2025.
CCXLVI, 246 f.: il.

Orientador: Prof. Fabrizzio Alphonsus Alves de Melo Nunes Soares; co-orientador Gabriel da Silva Vieira.
Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Informática (INF), Programa de Pós-Graduação em Ciência da Computação, Goiânia, 2025.
Bibliografia. Apêndice.
Inclui siglas, mapas, fotografias, tabelas, lista de figuras, lista de tabelas.

1. Inventário Florestal. 2. LiDAR. 3. Deep Learning. 4. Nuvens de Pontos 3D. 5. Eucalipto. I. Soares, Fabrizzio Alphonsus Alves de Melo Nunes, orient. II. Título.

CDU 004

Processo: 23070.041055/2025-16 Documento: 5662444

UNIVERSIDADE FEDERAL DE GOIÁS

INSTITUTO DE INFORMÁTICA

ATA DE DEFESA DE TESE

Ata nº **22** da sessão de Defesa de Tese de **Welington Galvão Rodrigues**, que confere o título de Doutor em Ciência da Computação, na área de concentração em Ciência da Computação.

Aos vinte e um dias do mês de outubro de dois mil e vinte e cinco, a partir das oito horas, via sistema de webconferência da RNP, realizou-se a sessão pública de Defesa de Tese intitulada **" Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas"**. Os trabalhos foram instalados pelo Orientador, Professor Doutor Fabrizzio Alphonsus Alves de Melo Nunes Soares (INF/UFG) com a participação dos demais membros da Banca Examinadora: Professor Doutor Gabriel da Silva Vieira (IF Goiano), coorientador; Professor Doutor Christian Dias Cabacinha (ICA/UFMG), membro titular externo; Professor Doutor Helio Pedrini (IC/Unicamp), membro titular externo; Professor Doutor Ronaldo Martins da Costa (INF/UFG), membro titular interno; Professora Doutora Deborah Silva Alves Fernandes (INF/UFG), membra titular interna. A realização da banca ocorreu por meio de videoconferência. Durante a arguição os membros da banca não fizeram sugestão de alteração do título do trabalho. A Banca Examinadora reuniu-se em sessão secreta a fim de concluir o julgamento da Tese, tendo sido o candidato **aprovado** pelos seus membros. Proclamados os resultados pelo Professor Doutor Fabrizzio Alphonsus Alves de Melo Nunes Soares, Presidente da Banca Examinadora, foram encerrados os trabalhos e, para constar, lavrou-se a presente ata que é assinada pelos Membros da Banca Examinadora, aos vinte e um dias do mês de outubro de dois mil e vinte e cinco.

TÍTULO SUGERIDO PELA BANCA

Documento assinado eletronicamente por **Ronaldo Martins Da Costa**, **Professor do Magistério Superior**, em 21/10/2025, às 13:09, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Deborah Silva Alves Fernandes**, **Professora do Magistério Superior**, em 21/10/2025, às 13:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Helio Pedrini**, **Usuário Externo**, em 21/10/2025, às 13:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Christian Dias Cabacinha**, **Usuário Externo**, em 21/10/2025, às 13:15, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Fabrizzio Alphonsus Alves De Melo Nunes Soares**, **Professor do Magistério Superior**, em 21/10/2025, às 13:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Welington Galvão Rodrigues**, **Discente**, em 21/10/2025, às 15:29, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5662444** e o código CRC **477AE751**.

Referência: Processo nº 23070.041055/2025-16 SEI nº 5662444

WELINGTON GALVÃO RODRIGUES

Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas

Tese defendida no Programa de Pós–Graduação do Instituto de Informática da Universidade Federal de Goiás como requisito parcial para obtenção do título de Doutor em Ciência da Computação, aprovada em 21 de Outubro de 2025, pela Banca Examinadora constituída pelos professores:

Prof. Dr. Fabrizzio Alphonsus Alves de Melo Nunes Soares
Instituto de Informática – UFG
Presidente da Banca

Prof. Dr. Gabriel da Silva Vieira
Instituto de Informática – UFG

Prof. Dr. Christian Dias Cabacinha
Instituto de Ciências Agrárias – UFMG

Prof. Dr. Hélio Pedrini
Instituto de Computação – UNICAMP

Prof. Dra. Deborah Silva Alves Fernandes
Instituto de Informática – UFG

Prof. Dr. Ronaldo Martins da Costa
Instituto de Informática – UFG

Welington Galvão Rodrigues

received a bachelor's degree in Information Systems from the Faculty of Imperatriz (FACIMP) in 2015. Completed a Master's degree in Computer Science in 2019 and is currently a professor at the State University of Tocantins (UNITINS).

Ao meu filho Benício, por me revelar uma nova maneira de compreender e enxergar a vida.

Acknowledgement

The last four years have been a period of intense transformation. Amid so many anxieties and challenges, life continued its course, bringing new friendships, farewells, and, above all, profound personal growth. It was during this time that I discovered the meaning of the most sublime love with the arrival of my son, who gave new meaning to my entire journey.

This thesis is the result of a journey that began long before, almost unpretentiously. I remember being in the computer lab at work when my colleague Felipe Rizzo approached me with a problem involving counting trees in an image. That challenge sparked a curiosity in me that led me to audit a course offered by Professor Jaques, in a partnership between UFG and FACIMP. That initial contact ignited a desire to delve deeper into the field. When I decided to pursue a graduate program in Computer Science at UFG, still at the master's level, I knew that moving to another state to study at the renowned Institute of Informatics would be an immense challenge.

And indeed, it was. There were moments when I considered giving up, and I might have if it weren't for my advisor, Professor Fabrizzio. A truly remarkable person, he welcomed and guided me with admirable patience and wisdom. I am deeply grateful for his guidance, which not only made this thesis possible but also introduced me to wonderful people in our research group. Among them is my dear friend Gabriel Vieira, whose inspiration and help were crucial during both my master's and doctoral studies.

After a period of uncertainties, marked by the pandemic and radical changes, I began my doctoral journey, feeling more resilient to face both academic and personal challenges. I am immensely grateful to have my partner, Liz, by my side, for all her understanding and for giving me the most precious gift I have ever received: our son Benício.

I thank my family, my foundation. To my mother, Maria, for all her unconditional support, love, and care. To my father, Raimundo, for all the support he has always provided. To my brother, Walison, for being the problem-solver and for accompanying me on some field studies. And to my sister, Weslaine, who, with her technical expertise in forestry engineering, contributed immensely to this work.

Above all, I thank God for the gift of life, for being the foundation of everything,

and for all the blessings bestowed upon my family.

Agradecimentos

Os últimos quatro anos foram um período de intensa transformação. Em meio a tantas angústias e desafios, a vida seguiu seu curso, trazendo novas amizades, despedidas e, acima de tudo, um profundo crescimento pessoal. Foi nesse período que conheci o significado do amor mais sublime com a chegada do meu filho, que ressignificou toda a minha jornada.

Esta tese é o resultado de uma trajetória que começou bem antes, de forma quase despretensiosa. Lembro-me de estar no laboratório de informática do trabalho quando meu colega Felipe Rizzo me procurou com um problema de contagem de árvores em uma imagem. Aquele desafio despertou em mim uma curiosidade que me levou a buscar, como aluno ouvinte, uma disciplina ofertada pelo professor Jaques, em uma parceria entre a UFG e a FACIMP. Esse primeiro contato acendeu o desejo de me aprofundar cada vez mais na área. Quando decidi ingressar na pós-graduação em Ciência da Computação na UFG, ainda no mestrado, sabia que mudar de estado para estudar no renomado Instituto de Informática seria um desafio imenso.

E, de fato, foi. Houve momentos em que pensei em desistir, e talvez o tivesse feito se não fosse pelo meu orientador, professor Fabrizzio. Uma pessoa verdadeiramente incrível, que me acolheu e guiou com uma paciência e sabedoria admiráveis. Sou extremamente grato por sua orientação, que não apenas tornou esta tese possível, mas também me apresentou a pessoas fantásticas em nosso grupo de pesquisa. Entre elas, meu caro amigo Gabriel Vieira, cuja inspiração e ajuda foram fundamentais tanto no mestrado quanto no doutorado.

Após um período de incertezas, marcado pela pandemia e por mudanças radicais, iniciei a jornada do doutorado, sentindo-me mais resiliente para enfrentar os desafios, tanto acadêmicos quanto pessoais. Sou imensamente grato por ter ao meu lado minha parceira, Liz, por toda a sua compreensão e por me dar o presente mais precioso que já recebi: nosso filho Benício.

Agradeço à minha família, meu pilar. À minha mãe, Maria, por todo o apoio, amor e carinho incondicionais. Ao meu pai, Raimundo, por todo o suporte que sempre me deu. Ao meu irmão, Walison, por ser o resolvedor de problemas e por me acompanhar em alguns estudos de campo. E à minha irmã, Weslaine, que com seu conhecimento técnico

em engenharia florestal, contribuiu imensamente para este trabalho.

Acima de tudo, agradeço a Deus pelo dom da vida, por ser a base de tudo e por todas as bênçãos concedidas à minha família.

"Study hard what interests you the most in the most undisciplined, irreverent and original manner possible."

Richard Feynman

Resumo

Rodrigues, Welington Galvão. **Estimativa de Diâmetro de Troncos de Eucalipto a Partir de Nuvens de Pontos LiDAR de Smartphone e Redes Neurais Profundas**. Goiânia, 2025. 246p. Tese de Doutorado. Instituto de Informática (INF), Universidade Federal de Goiás (UFG).

A medição precisa de parâmetros dendrométricos, particularmente o Diâmetro à Altura do Peito (DAP), é fundamental para o inventário florestal e o manejo sustentável. No entanto, os métodos de campo tradicionais são trabalhosos, demorados e propensos a erros. A integração de sensores LiDAR em *smartphones* oferece uma alternativa escalável e de baixo custo, mas requer métodos computacionais robustos para processar os dados 3D resultantes. Esta tese apresenta um *framework* para a medição automatizada do DAP em árvores de eucalipto, utilizando aprendizado profundo para reconstrução 3D e segmentação semântica. Desenvolvemos e validamos dois *pipelines* de processamento distintos: um que reconstrói modelos de árvores completos de 360 graus a partir de varreduras rápidas e parciais, usando uma arquitetura de completude de nuvem de pontos, e outro que processa varreduras circulares completas capturadas diretamente no campo. Ambos os fluxos de trabalho convergem para uma etapa de segmentação compartilhada, onde uma rede PointTransformer realiza uma segmentação de alta fidelidade para isolar com precisão a região do DAP do tronco da árvore. Um algoritmo calcula então o diâmetro a partir da nuvem de pontos segmentada. Nossos resultados demonstram que o *framework* proposto alcança alta precisão, com medições que não apresentam diferença estatisticamente significativa em relação aos métodos manuais com suta. O *pipeline* de varredura direta mostrou-se superior, atingindo um RMSE inferior a 0,55 cm em todas as classes de árvores. A metodologia proposta aumenta significativamente a eficiência operacional, reduzindo o tempo de coleta em campo de três a trinta vezes. O fluxo de trabalho validado, de alta precisão e baixo custo, avança a automação dos inventários florestais e promove práticas mais eficientes na silvicultura de precisão. Dois conjuntos de dados públicos com 880 varreduras anotadas são disponibilizados para futuras pesquisas.

Palavras–chave

Inventário Florestal, LiDAR Móvel, aprendizado profundo, Nuvens de Pontos 3D, Eucalipto.

Abstract

Rodrigues, Welington Galvão. **Estimating Eucalyptus Stem Diameter from Smartphone LiDAR Point Clouds and Deep Neural Networks**.
Goiânia, 2025. 246p. PhD. Thesis. Instituto de Informática (INF), Universidade Federal de Goiás (UFG).

Accurate measurement of dendrometric parameters, particularly the Diameter at Breast Height (DBH), is fundamental for forest inventory and sustainable management. However, traditional field methods are labor-intensive, time-consuming, and prone to error. The integration of LiDAR sensors into consumer-grade smartphones offers a scalable and cost-effective alternative, yet requires robust computational methods to process the resulting 3D point cloud data. This thesis presents a novel, end-to-end framework for the automated measurement of DBH in eucalyptus trees by leveraging deep learning for 3D reconstruction and semantic segmentation. We develop and validate two distinct processing pipelines: one that reconstructs complete 360-degree tree models from rapid, partial scans using a point cloud completion architecture, and another that processes full-circle scans captured directly in the field. Both workflows converge into a shared segmentation stage where a PointTransformer network performs high-fidelity semantic segmentation to precisely isolate the DBH region from the tree stem. An automated algorithm then calculates the diameter from the segmented point cloud. Our results demonstrate that the proposed framework achieves high precision, with measurements showing no statistically significant difference from manual caliper methods. The direct-scan pipeline proved superior, achieving a RMSE of less than 0.55 cm across all tree classes. Critically, the methodology yields a transformative improvement in operational efficiency, reducing in-field data collection time by three to thirty-fold. By validating a high-precision, low-cost workflow, this research provides a significant step toward the automation of forest inventories, enabling more efficient and data-driven practices in precision forestry. To foster future research, two novel datasets containing 880 annotated tree scans are also made publicly available.

Keywords

Forest Inventory, Mobile LiDAR, Deep Learning, 3D Point Clouds, Eucalyptus.

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

1-NNA 1-Nearest Neighbor Accuracy

3D Three-Dimensional

3D-CNN Three-Dimensional Convolutional Neural Network

AAEs Adversarial Auto-Encoders

ACC Accuracy

ADAM Adaptive Moment Estimation

AI Artificial Intelligence

ALS Airborne Laser Scanning

ANOVA Analysis of Variance

BEV Bird-Eye View

BIAS Percent Bias

CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CD Chamfer Distance

CHM Canopy Height Model

CNN Convolutional Neural Network

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico

COV Coverage

CPU Central Processing Unit

CSCN Cascade of Sparse Coding-based Networks

CV Computer Vision

DBH Diameter at Breast Height

DL Deep Learning

DNN Deep Neural Networks

DPM Diffusion Probabilistic Model

DSM Digital Surface Model

DT Decision Trees

DTM Digital Terrain Model

DU-DEAL Depth Up-sampling with Deep Edge-Aware Learning

EC Exclusion Criteria

ECC Edge-Conditioned Convolution

EMD Earth Mover's Distance

FAPEG Fundação de Amparo à Pesquisa do Estado de Goiás

GANs Generative Adversarial Networks

GEDI Global Ecosystem Dynamics Investigation

GDP Gross Domestic Product

GF Guided Filter

GNSS Global Navigation Satellite System

GPU Graphics Processing Unit

GRNN General Regression Neural Networks

GRU Gated Recurrent Unit

HR High-Resolution

IC Inclusion Criteria

IoU Intersection over Union

kNN k-Nearest Neighbors

LISTA Learned Iterative Shrinkage and Thresholding Algorithm

LiDAR Light Detection and Ranging

LR Low-Resolution

LSTM Long Short-Term Memory

LSTM-DA Long Short-Term Memory with Data Augmentation

LSTM-VA Vanilla Long Short-Term Memory

l-GANs Latent Generative Adversarial Networks

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

mAP mean Average Precision

MCC Matthews Correlation Coefficient

MCTI Ministério da Ciência, Tecnologia e Inovações

MSE Mean Squared Error

MSG-Net Multi-Scale Guidance Network

MMD Minimum Matching Distance

MLP Multilayer Perceptron

MLP-DAPHT Multilayer Perceptron with DBH and Total Height

MLP-PD Multilayer Perceptron for Direct Prediction

MLP-PR Multilayer Perceptron for Recursive Prediction

MRF Markov Random Field

MSPF Multi-Scale Progressive Fusion

mIoU mean Intersection over Union

nDSM Normalized Digital Surface Model

PDDSR Perceptual Depth Denoising and Super-Resolution

PCCN Point Cloud Convolutional Network

PGR Pillar-based Ground Removal

PReLU Parametric Rectified Linear Unit

PSNR Peak Signal-to-Noise Ratio

PVT Point Voxel Transformer

QC Quality Criteria

RAM Random Access Memory

Raw Generative Adversarial Networks r-GANs

RDN Residual Dense Network

RevNN Recursive Neural Networks

RF Random Forest

RGB Red Green Blue

RGB-D Red, Green, Blue - Depth

RMSE Root Mean Square Error

RNN Recurrent Neural Networks

ROC Receiver Operating Characteristic

RQ Research Question

SFA Structural Feature Aggregation

SGD Stochastic Gradient Descent

SLR Systematic Literature Review

SR Super-Resolution

SRCNN Super-Resolution Convolutional Neural Network

SRFBN Super-Resolution Feedback Network

SVM Support Vector Machines

SVR Support Vector Regression

TGV Total Generalized Variation

TLS Terrestrial Laser Scanner

ToF Time-of-Flight

VAEs Variational Auto-Encoders

WGANs Wasserstein Generative Adversarial Networks

CHAPTER 1

Introduction

Happy is the one who finds wisdom.

Proverbs 3:13

Accurate measurement of dendrometric parameters is a cornerstone of forest engineering and management, with stem diameter being the primary metric for estimating volume, basal area, and tree classification. With the advancement of remote sensing technologies, mobile Light Detection and Ranging (LiDAR) has emerged as a promising tool for large-scale collection of three-dimensional data in the field. In this context, this thesis investigates and develops machine learning techniques to automate the detection, segmentation, and extraction of initial diameters of eucalyptus trees from point clouds. The work begins with the analysis of a reconstructed point cloud to validate the extraction of data at key stem points, demonstrating the potential of this approach to optimize the collection of crucial information for forest management.

For the automated extraction of tree diameters, we systematically explored different approaches that compose a complete data processing pipeline. This workflow encompasses critical stages, ranging from the initial image processing, including depth image upscaling, point completion and point cloud generation, to data refinement through ground removal and the detection and segmentation of regions of interest. Ultimately, the process culminates in the calculation of eucalyptus tree diameters. All these stages were integrated into a single, cohesive pipeline. The performance of each stage was rigorously evaluated using specific assessment metrics, and comparative studies with similar works were conducted to validate our results. Thus, the methods developed in this thesis are capable of generating the point cloud from RGB-D images, filtering the ground, and subsequently detecting and segmenting regions of interest for the precise extraction of Diameter at Breast Height (DBH).

The workflow for point cloud analysis was implemented following a pipeline architecture, where each processing stage is encapsulated within an independent and modular software component. From point cloud generation to the final DBH computation, each functionality operates autonomously, facilitating the replacement and evaluation of differ-

ent algorithms for the same task (e.g., alternative ground removal methods). This modular approach was crucial for the research, as it simplified debugging within a complex Three-Dimensional (3D) data environment and enabled isolated performance analysis of each component. Additionally, this architecture allows the creation of lightweight processes capable of running on edge devices and facilitates the integration of the solution with larger digital forest management platforms.

Another innovation of this work lies in the robustness of the models to variations in terrain slope around the tree, a well-known challenge for measurement accuracy. In manual approaches, terrain inclination introduces a significant margin of error since defining the reference point for breast height measurement becomes ambiguous and operator-dependent. Conversely, photogrammetry-based techniques often suffer from inaccuracies in 3D reconstruction on uneven terrains due to difficulties in camera calibration and point correspondence. In contrast, the models developed in this thesis learn general structural and geometric features of the tree stem during training, enabling them to identify the tree base and compute DBH more consistently, mitigating errors caused by topography.

This thesis is organized into ten chapters. Chapter 1 introduces the study, detailing the motivation, research problem, hypotheses, and objectives of the investigation. Chapter 3 explores the prediction of eucalyptus diameters and volume using LSTM neural networks, establishing a baseline with dendrometric data. Chapter 2 presents a Systematic Literature Review on the application of Artificial Intelligence and LiDAR in forest inventories, contextualizing the state of the art. Chapter 4 focuses on improving input data quality through depth map super-resolution using a MobileNetV2-enhanced architecture. In Chapter 5, a comparative analysis of deep learning models for eucalyptus point cloud segmentation is conducted. Chapter 6 describes a pipeline for accurate diameter measurement using smartphone LiDAR, combining Faster R-CNN and PointNet. Chapter 7 addresses the challenge of incomplete data, evaluating point cloud completion methods. Chapter 8 investigates the generation of high-fidelity synthetic point clouds, comparing architectures ranging from GANs to Diffusion Models. Chapter 9 consolidates the previous stages by presenting and comparing two complete and integrated pipelines for automated diameter measurement: one for partial scans with reconstruction and another for full scans. Finally, the concluding remarks, limitations, and directions for future research are presented in Chapter 10. Additionally, Appendix A catalogs the contributions generated throughout the research, including papers, software, and new datasets.

1.1 Motivation and Significance

The Brazilian planted forest sector stands out as a strategic pillar of the national economy, driven by remarkable resilience and innovation capacity. Confidence in its

growth trajectory is evidenced by a robust investment portfolio, projecting contributions of R$105 billion in expansion by 2028. This strength is reflected in its global leadership, with international trade reaching US$12.7 billion annually and Brazil holding the position of the world's largest exporter of pulp (IBÁ, 2024).

The foundation of this economic powerhouse lies in its massive and technology-intensive operational scale. In 2023, the total area dedicated to tree plantations in Brazil surpassed, for the first time, the 10-million-hectare mark, representing a 3% increase compared to the previous year. These data were obtained through satellite imagery mapping conducted by *Canopy Remote Sensing Solutions*. The sector plants approximately 1.8 million trees every day, which are transformed into a wide and growing range of renewable-source products, such as books, notebooks, clothing, pulp, paper, packaging, toilet paper, diapers, wood panels, laminate flooring, and many others.

Also in 2023, the gross production value of the planted tree sector reached R$202.6 billion, with a compound annual growth rate (CAGR) of 3.2% over the past decade. The forest production chain accounted for 0.9% of Brazil's Gross Domestic Product (GDP) that year. Specifically, forest production represented 4.2% of the agricultural GDP and 4% of the manufacturing industry GDP.

From the perspective of domestic demand, forest-based products accounted for 1.9% of intermediate consumption and 0.5% of household final consumption, also encompassing gross capital formation. Regarding production and import structures, the participation of the planted tree production chain was 1% and 0.4%, respectively. These figures not only sustain essential production chains—from paper and packaging to advanced bioproducts—but also position Brazil as a central player in the transition toward a global bioeconomy, offering renewable alternatives to fossil-based products (IBÁ, 2024).

The magnitude and increasing complexity of this sector demand continuous advancement in technology and knowledge to sustain competitiveness and productivity. Eucalyptus, occupying 7.8 million hectares—or 76% of the total planted area—is the key species within this dynamic and, therefore, the main focus for technological innovation. Although the national average productivity of eucalyptus is 33.7 $m^3/ha/year$, the significant variation across different producing regions reveals a vast potential for optimization. As illustrated in Figure 1.1, eucalyptus consolidates itself as the central species driving this dynamic.

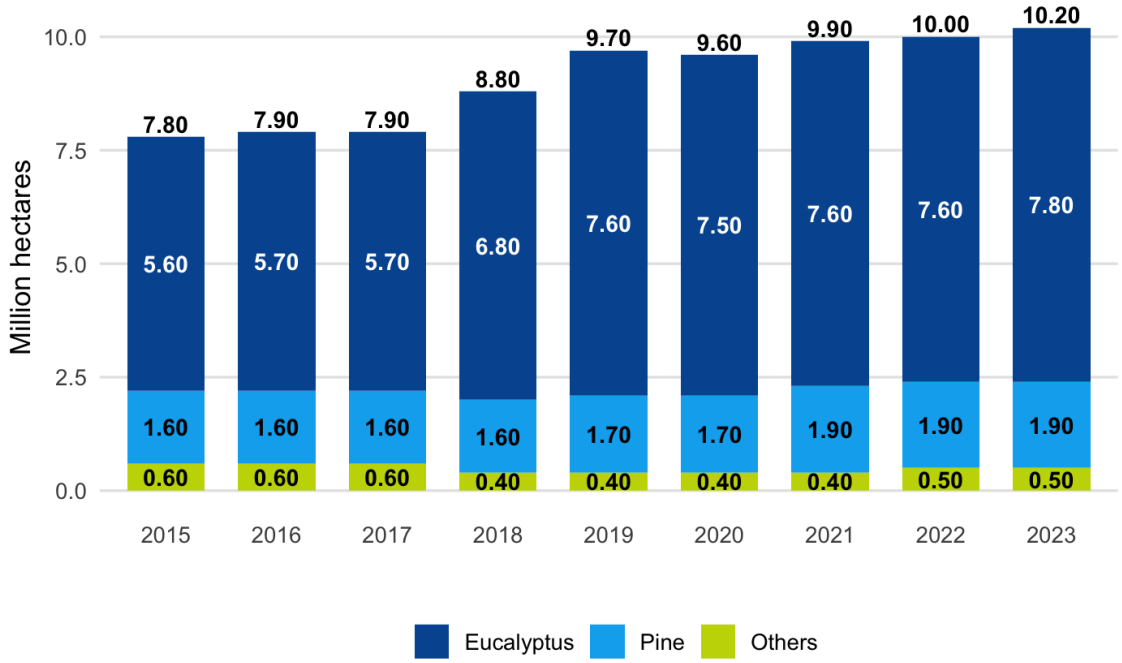


Figure 1.1: *History of the distribution of planted area in Brazil by species (2015–2023) in millions of hectares (IBÁ, 2024).*

The management and strategic planning of a biological asset of such magnitude fundamentally depend on a central technical tool: the *Forest Inventory*. Conceptually, the inventory is the instrument responsible for the "counting of trees" and the systematic collection of qualitative and quantitative information that characterizes the forest resources of a given area (CAMPOS; LEITE, 2006). As highlighted in the literature, it provides the diagnostic basis for the forest's productive and protective potential, serving as the informational foundation that supports critical management decisions, harvest planning, silvicultural treatments, and, crucially, the assessment of the feasibility of large-scale investment projects. In summary, the forest inventory is the primary data source for in-depth knowledge of the forest asset (CABACINHA, 2003).

However, the effectiveness of a forest inventory is intrinsically linked to its execution methods. The traditional approach, based on the observation and measurement of a limited number of sample units (plots), faces challenges in terms of cost, time, and labor intensity, especially when applied to large areas (CUNHA, 2004). Although statistically valid, this dependence on sampling imposes practical limitations on the granularity and frequency of data collection. Given the increasing complexity of the sector and the relentless pursuit of productivity gains, such limitations have become a bottleneck. The need for adaptive and precision-oriented management, capable of responding to local variations in soil, climate, and genetic material, demands a volume and quality of data that conventional inventory methods struggle to provide in an economically viable way.

Overcoming these sectoral challenges requires the modernization of forest in-

ventory practices, as traditional methods, costly, laborious, and with low spatiotemporal resolution, are insufficient to meet the demands of precision forestry. This creates a clear motivation to develop dendrometric data collection technologies that are automated, accurate, and scalable. Such methods are essential strategic tools to enhance forest management, drive productivity gains, and maximize the economic return of the sector.

In this context, remote sensing technologies emerge as promising solutions for the modernization of forest inventory, with special emphasis on LiDAR. This technology uses laser pulses to measure distances with extremely high precision, generating detailed three-dimensional models of the environment known as point clouds (DUBAYAH; DRAKE, 2000). Its application has transformed ecosystem mapping, enabling the extraction of parameters such as forest vertical structure, tree height, and wood volume estimation with a high degree of precision and spatial detail.

Within the forest inventory framework, the use of LiDAR provides disruptive advantages over manual methods. The ability to collect data without direct contact with the trees not only drastically accelerates the diameter measurement process but also enables access to difficult terrain and mitigates systematic errors associated with traditional instruments, such as calipers. Furthermore, the growing integration of LiDAR sensors into mobile platforms such as *smartphones* opens the perspective of transforming the inventory process from a sampling-based and sporadic activity into a large-scale, continuous and high-resolution data collection practice, more comprehensive, detailed and aligned with the demands of precision forest (LUETZENBURG *et al.*, 2021).

Diameter stands as a crucial variable in studies focused on volumetric quantification, biomass assessment, and growth analysis. As highlighted in the literature, diameter is the fundamental parameter for calculating volume, biomass, basal area, and for analyzing forest structure as expressed by the diameter distribution (FLORIANO, 2021). Therefore, accuracy in its measurement is not an end in itself, but rather the foundation for a comprehensive and reliable forest inventory.

Different strategies for performing dendrometric calculations using alternative technologies, such as imagery, have been investigated in the literature; however, they have revealed considerable discrepancies between estimated and actual diameters. The study by (MAULONI *et al.*, 2011) explored digital photogrammetry for dendrometric calculations, revealing a significant margin of error ranging from 3 to 7 centimeters. In a similar approach, (ASHLEY; ROGER, 1968) employed photogrammetry for diameter measurement, involving camera calibration and observing an increase in error with tree height. Although approaches such as that presented by (CLARK *et al.*, 1998) demonstrate promising results, it is important to note that such systems require customized calibration for each application, considering factors such as slope and lighting conditions. Despite encouraging findings, photogrammetric techniques have been scarcely employed in the

national context due to the effort required, particularly for precise diameter measurement. This performance gap in photogrammetry reinforces the need for direct 3D measurement technologies, such as LiDAR, which are inherently more robust to such sources of error.

1.2 Problems

The predominant dendrometric variables used for conducting forest inventories are tree height and diameter, whose measurements can be used to derive additional metrics, such as basal area and wood volume within a forest ecosystem (FREITAS; WICHERT, 1998). However, obtaining diameter measurements at different heights along the tree requires considerable time and resources (GAFFREY *et al.*, 2001).

The measurement of these forest parameters presents a series of specific scientific challenges that may arise due to the particular characteristics of eucalyptus trees and the environments in which they are cultivated. Some of the main issues include, for example:

- the irregular shape of eucalyptus trees, which may vary in form and size due to factors such as light competition, variable planting spacing, and soil conditions;
- the quality of the stem, which can be affected by pests, diseases, and competition. The presence of dead branches, knots, and irregularities along the stem can hinder the accurate determination of true diameter; and
- the sensitivity of volumetric parameters to measurement errors. During a forest inventory, small inaccuracies can significantly affect total volume estimates. (COUTO; BASTOS, 1987) demonstrated that an error of 1 meter in total height measurement corresponds to approximately 14% of the cylindrical volume of a tree, while an error of 1 cm in DBH results in an error of approximately 19% of the total tree volume.

Traditional methods, such as rigorous cubage, are simple and widely used to obtain precise measurements. However, a key disadvantage lies in the need to fell multiple selected trees. In addition to being a slow process, it requires immediate commercialization to avoid product loss (CABACINHA, 2003).

Determining tree volume is a fundamental task for sustainable forest management and resource assessment. . In this process, diameter stands out as the most important dendrometric variable. Its prominence derives from being easily accessible in the field while serving as the basis for a wide range of derived calculations, including not only volume but also biomass, basal area (which expresses stand density), and forest structure, as revealed by the diameter distribution. The standardized measurement of this parameter, typically obtained at 1.30 meters above ground level using a caliper or diameter tape,

assumes that the stem cross-section is approximately elliptical. However, this traditional approach, while seemingly simple, presents several challenges that can compromise the accuracy of estimates and, consequently, the quality of management decisions (CUNHA, 2004). Figure 1.2 illustrates the main difficulties encountered in the diameter measurement process.

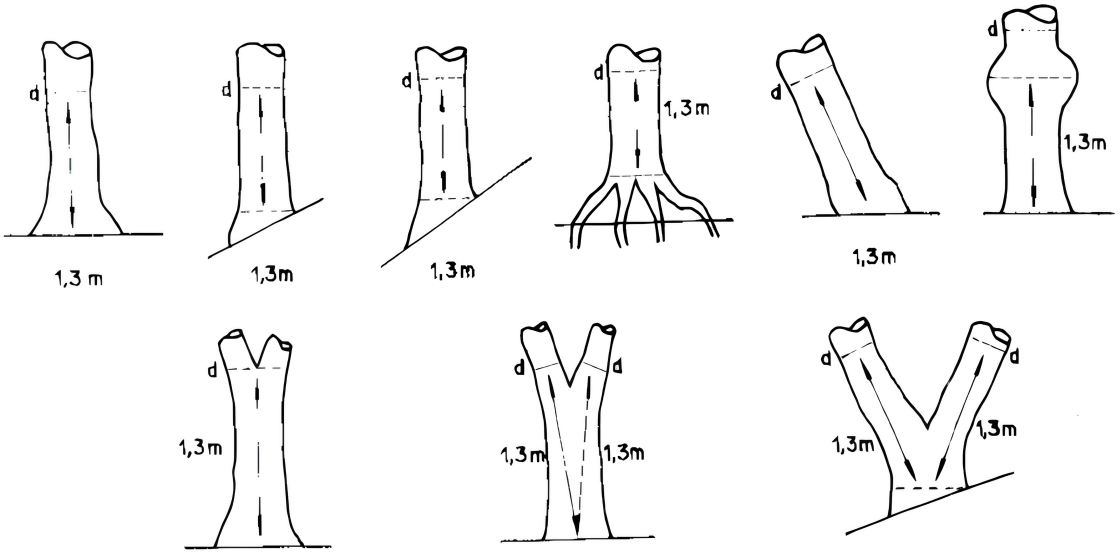


Figure 1.2: *Process of measuring the diameter in different terrains and tree stems.*

The slope of the terrain, for example, makes the definition of the starting point for measuring the DBH ambiguous, introducing a margin of error that depends on the operator. Similarly, irregularities in the stem, such as forks, calluses, or the presence of buttress roots, prevent measurement at a standardized point and can distort the values obtained with a caliper or diameter tape. These factors compromise not only the accuracy of individual estimates but also the consistency of data in large-scale inventories. To mitigate such sources of error and ensure the standardization of the collected data, this work adopted a specific measurement protocol. The following Figure 1.3 details the diameter collection process carried out in this thesis.

Figure 1.3: *On-site data collection process. (a, b) Measuring diameter with a forestry caliper. (c) Recording field data. (d) Tree marked at the standard DBH height. (e, f) Views of the eucalyptus stand environment. (g) Measuring tree height with a digital clinometer. (h, i) Acquiring 3D data with a smartphone-based LiDAR scanner.*

As depicted in Figure 1.3, each tree was first measured using conventional tools. The DBH, marked at 1.30 meters, was measured with both a forestry caliper

and a diameter tape to ensure accuracy. These ground-truth values, along with other observations, were manually logged in the field. Concurrently, a smartphone equipped with a LiDAR sensor was mounted on a tripod and used to perform a 360-degree scan of the tree stem. This dual-method approach allowed for a direct comparison between the established manual measurements and the data derived from the 3D point clouds, forming the basis for evaluating the accuracy and efficiency of the proposed automated methodology.

To overcome the challenges of accuracy, cost, and scalability inherent to traditional methods, this thesis proposes an end-to-end computational approach based on deep learning models and mobile LiDAR sensor data for the automatic extraction of eucalyptus diameters. The realization of this objective, however, requires addressing a chain of specific subproblems, each representing a critical stage in the data processing pipeline, from acquisition to final measurement:

- enhancing the resolution of point clouds generated from smartphone sensors through RGB-D image upscaling, aiming to achieve a denser and more accurate 3D representation of the scene;
- developing a method for the automatic removal of ground points from the point cloud, a fundamental step to isolate vegetation and reduce computational complexity for subsequent stages;
- designing a model capable of identifying and segmenting individual tree stems from the surrounding environment, accurately isolating the object of interest;
- addressing the issue of occlusions, which are common in field scans, by applying point completion techniques to reconstruct the complete stem shape from partial point clouds;
- improving the robustness and generalization capability of the models through the generation of synthetic eucalyptus point cloud data, mitigating the scarcity of annotated samples; and
- finally, developing an algorithm for the accurate computation of DBH from the segmented stem point cloud, ensuring that the method remains resilient under diverse field conditions, such as stem irregularities and variations in capture quality.

The sequential design and integrated resolution of these challenges form the methodological backbone of this thesis, aiming to significantly enhance the management and monitoring of forest production.

1.3 Research Questions

The research questions guiding this study can be summarized into five central axes, which aim to investigate topics ranging from technological aspects of data acquisition to advanced processing and analytical methodologies. These questions are outlined as follows:

RQ1: Is it feasible to use LiDAR sensors integrated into smartphones to capture point clouds with sufficient precision for the automated measurement of DBH in eucalyptus stands?

RQ2: How can deep learning techniques, such as depth map super-resolution and point cloud completion, be integrated into a unified pipeline to reconstruct complete and high-fidelity 3D models of trees from partial and low-resolution field scans?

RQ3: Which deep learning architectures, particularly point-based models, demonstrate superior effectiveness for the precise semantic segmentation of the DBH region in eucalyptus point clouds, considering datasets of varying quality?

RQ4: Can generative models be employed to synthesize structurally faithful eucalyptus point clouds, thereby enabling data augmentation and the robust training of deep learning models?

RQ5: What is the most effective method for calculating DBH from a segmented point cloud, and how does its accuracy compare to that of traditional forest inventory measurements?

1.4 Hypotheses

Based on the research questions outlined, five central hypotheses are formulated, each designed to provide specific answers to the investigated problems, as described below:

H1: Data acquisition using low-cost LiDAR sensors, such as those integrated into *smartphones*, produces point clouds with sufficient accuracy to enable the extraction of dendrometric parameters comparable to manual measurement methods.

H2: The integration of deep learning techniques within a unified pipeline enables the reconstruction of complete and high-fidelity 3D models from fast, partial, and low-resolution field scans.

H3: Point-based deep learning architectures demonstrate strong effectiveness for semantic segmentation tasks and can accurately isolate the DBH region within eucalyptus point clouds.

H4: It is feasible to employ advanced generative models to synthesize eucalyptus point clouds with high structural and perceptual fidelity, providing a valuable resource for data augmentation and enhancing the robustness of forest analysis models.

H5: DBH extraction through geometric fitting of a segmented point cloud section can yield measurements with high correlation and low mean error compared to manual reference measurements.

1.5 Aims

In this thesis, we investigate the development of a computational pipeline for the automatic extraction of dendrometric parameters of eucalyptus, with a focus on DBH from point clouds. In this context, the objective is to propose solutions based on deep learning and low-cost LiDAR sensor data for precision forestry, with attention to automated computational processes, high accuracy, and algorithms capable of generalizing to the diverse conditions encountered in the field.

More specifically, this study examines the following research themes within the fields of Forest Engineering and Forestry.

1. Develop and validate a method for the automated measurement of DBH in eucalyptus stands using smartphone LiDAR.
2. Quantify the gain in operational efficiency (reduction in data collection time) provided by the pipelines compared to conventional inventory practices.
3. Analyze the morphology of the stem and its variability under real field conditions.
4. Reconstruct information from partially occluded trees for a more complete inventory.
5. Statistically validate the accuracy of the proposed methods compared to manual field measurements (caliper).

Additionally, we investigate the following research topics related to the field of Computer Science:

1. Techniques for processing, filtering, and analyzing 3D point clouds.
2. Application of deep learning for the upscaling of depth images (RGB-D).
3. Semantic segmentation of point clouds, object detection, and 3D reconstruction (point completion).
4. Generation of synthetic 3D data to enhance model robustness and generalization.
5. Development of algorithms for fitting geometric shapes to point clouds.
6. Performance analysis of models considering accuracy, robustness, and computational efficiency.

7. Create and publicly release new annotated point cloud datasets of eucalyptus to promote reproducibility and future research in the field.

1.6 Methodology

The research methodology used in this work consisted of the following steps:

- Literature review with research on forest inventory, precision forestry, remote sensing with LiDAR, and the state-of-the-art in deep learning models for 3D data processing.
- Investigation of point cloud processing techniques applied to tree analysis and the results provided by computer vision approaches in 3D data.
- Research on public point cloud databases, such as *ShapeNet* (CHANG *et al.*, 2015a), *ModelNet* (WU *et al.*, 2015), *KITTI* (LIAO *et al.*, 2022), *nuScenes* (CAESAR *et al.*, 2020), among others, with the aim of identifying relevant data to support the research.
- Planning and execution of field data collection, using calipers and measuring tapes for manual DBH measurements, as well as LiDAR sensors integrated into smartphones, to build a proprietary dataset aligned with the specific needs of the study.
- Construction of the computational pipeline design, with attention to the implementation of automated and high-performance processes for extracting dendrometric parameters.
- Development of a software architecture model with code modularization and self-contained functionalities (e.g., preprocessing, segmentation, diameter extraction).
- Planning and implementation of deep learning algorithms for stem segmentation, reconstruction of partial point clouds (point completion), and generation of synthetic data.
- Formulation of experimental tests to evaluate the computational models presented in this research, using metrics of accuracy, robustness, and efficiency.
- Development of comparative studies with related works, considering both quantitative aspects (measurement errors) and qualitative aspects (robustness in different field scenarios).
- Organization and presentation of research results in the form of scientific articles for publication in journals and international conferences.

1.7 Contributions

This thesis presents an investigative study on the development of computational models for precision forestry, with a primary focus on automating forest inventory through a 3D data processing pipeline. The research addresses key aspects of dendrometric analysis, including point cloud acquisition and preprocessing, individual tree segmentation, partial data reconstruction, and robust diameter extraction. Each component contributed to the creation of algorithms designed to perform automated tasks, achieving reliable results with minimal human intervention and reduced dependence on field specialists. In alignment with the principles of Forestry 4.0, the integration of computational tools into dendrometric analysis is crucial for enhancing accuracy while reducing human effort, operational costs, and subjectivity (FENG; AUDY, 2020).

The presence of experts is fundamental for decision-making in forest management, but the scale of plantations in Brazil makes traditional inventory a logistical and financial challenge. Manual methods, in addition to being labor intensive, are susceptible to the subjectivity of the measurer, which can lead to inaccurate information and erroneous judgments, directly affecting volume calculations, where a 1 cm error in diameter can result in up to a 19% error in volume. For this reason, computational tools for dendrometric analysis are essential to reduce human labor and the operational costs of manual operations, while increasing accuracy through processes free of subjectivity.

The desired automation faced challenges inherent to working with 3D point cloud data, especially those from low-cost sensors. Field data acquisition can lead to variations in point density, tree occlusion, sensor noise, and topographic variations of the terrain. To address these issues, we modeled the programs to be robust to these sources of error. We specifically explored solutions for reconstructing trees with partial data using *point completion* techniques, a common problem in field scans where neighboring trees obstruct the view. Additionally, we developed algorithms for generating synthetic point clouds with realistic eucalyptus characteristics, aiming to enhance the robustness of learning models.

Another highlight of this work is the focus on building computational models with cohesive software components and efficient processes, suitable for the reality of precision forestry and the use of low-cost devices. The feasibility of models operating on smartphone data is crucial for environments with limited computational resources. In this regard, we measured results in terms of effectiveness (accuracy) and efficiency (processing time). The experiments show highly accurate results and involve the agile execution of the models. Furthermore, the generalization capability of the models is a critical feature that enables their use in different planting scenarios, with variations in spacing, age, and terrain conditions.

While similar works often focus on only one stage of the process (e.g., segmentation alone) or use high-cost sensors, our work presents a complete and integrated *pipeline*, from capture with mobile devices to the final extraction of diameter. Among the new functionalities, we highlight the application of *deep learning* models not only for segmentation but also for improving the quality of input data (*upscaling*) and reconstructing missing data (*point completion*). The generation of a synthetic dataset of eucalyptus point clouds for training neural networks is, to the best of our knowledge, one of the first discussions applied to this specific context.

Another contribution of this thesis is the scientific articles published with partial results of this investigative study. For each achieved result, we described the motivating context, modeled the proposals, specified the algorithms, prepared experimental tests, conducted data analysis, and discussed the observed results. In this way, we organized and presented the articles, disseminated our algorithms, prepared databases with eucalyptus point clouds, and implemented original software.

In summary, the main contributions of this thesis are:

- A complete computational workflow, from the acquisition of raw data with smartphones to the extraction of DBH, was statistically validated, with accuracy comparable to manual methods but with significantly higher efficiency.
- Reconstruction of partial tree point clouds using point completion techniques.
- Validation of a low-cost approach demonstrating that the use of consumer LiDAR sensors integrated into mobile devices is a viable and scalable alternative.
- Generation of a synthetic dataset of eucalyptus point clouds for model training.
- Modularized computer programs for the analysis of 3D data in the forestry context.
- Deep learning-based methods for processing point clouds, including *upscaling* of depth images, segmentation, and reconstruction.
- Computational models with high performance and accuracy, validated against traditional methods.
- Scientific production with the research results.
- Creation of new annotated datasets with eucalyptus scans, made publicly available, contributing to research and reproducibility in forestry environments.

1.8 Organization of the Thesis

This thesis is organized into ten chapters, with the first presenting the introduction and the tenth addressing the conclusion of the research work. In addition, the thesis includes appendices containing the original articles, software, and datasets developed by the authors. To provide an overview of each development chapter (Chapters 2 to 9) and the appendices, their main topics and respective contents are outlined below.

In Chapter 2, we present a Systematic Literature Review (SLR) on the applications of Artificial Intelligence (AI) and LiDAR in forest inventories, aiming to identify, evaluate, and interpret the results of primary studies at this intersection. Traditional forest inventory approaches are characterized by being costly, time-consuming, and sometimes invasive, resulting in inaccuracies in obtaining critical data such as tree volume and diameter. The SLR shows that the integration of AI and LiDAR data has proven to be an effective solution for automating these tasks and achieving results superior to conventional methods. We observed a growing trend in the application of deep learning, which demonstrates the ability to automatically extract complex features, surpassing traditional machine learning algorithms in various metrics. However, the review highlights that direct diameter (DBH) estimation from aerial scans is a minimally explored aspect, despite its fundamental importance for timber volume calculations. Even with the computational demands and training time requirements of deep learning models, the research underscores their efficiency and broad applicability for improving forest management.

In Chapter 3, we present the fundamentals and development of neural network models for predicting the diameter and volume of eucalyptus trees, focusing on the stages of data processing and initial training. A critical initial task is to address the limitations of traditional forest inventory methods, which are labor-intensive, time-consuming, costly, and often involve tree felling for precise measurements. We discuss how neural network algorithms can be employed to model the complex nonlinear relationships between tree measurements, enhancing accuracy and efficiency. We introduce the five neural network models developed: Vanilla Long-Short Term Memory (LSTM-VA), Long-Short Term Memory with Data Augmentation (LSTM-DA), Multilayer Perceptron for Recursive Prediction (MLP-PR), Multilayer Perceptron for Direct Prediction (MLP-PD), and Multilayer Perceptron with DBH and Total Height (MLP-DAPHT). The results demonstrated the potential of neural networks, particularly LSTMs, to revolutionize forest inventory by providing accurate and generalizable estimates, significantly benefiting forest management, decision-making, and the sustainability of the industry.

In Chapter 4, we present an approach for super-resolution of depth maps, capable of generating dense representations from sparse measurements and overcoming the limitations of low-cost sensors. Unlike traditional methods, which introduce noise and lose edge details, our solution uses an architecture based on MobileNetV2 integrated with the MSG-Net framework, combining multi-scale information from RGB images and depth maps. The network has two main branches, a color encoder with a pre-trained MobileNetV2 and a depth encoder with traditional convolutions, whose features are hierarchically fused for reconstruction. Evaluated on the Middlebury and MPI Sintel Depth datasets, the proposed approach demonstrated superior performance compared to the standard MSG-Net, achieving low RMSE values, reduced noise, and greater clarity,

while requiring only 2,000 training epochs, in contrast to over 500,000 for MSG-Net, yielding competitive results even against the state-of-the-art MSPF method.

In Chapter 5, we present an innovative methodology for the precise measurement of eucalyptus tree diameters, overcoming the limitations of traditional inventories, which are labor-intensive, time-consuming, and prone to errors. We propose a pipeline that integrates smartphone LiDAR and deep learning models, combining Faster R-CNN for tree detection in Bird-Eye View (BEV) projections with stem segmentation via PointNet, trained with a hybrid loss function (weighted Focal Loss and Dice Loss) to address class imbalance. A specific algorithm was developed to calculate the diameter from segmented regions in the 3D point cloud. The results showed high accuracy, with an RMSE below 1 cm, strong correlation with manual measurements ($r = 0.98$), and no statistically significant differences. In addition to reducing collection time and errors, the approach positions smartphone LiDAR as a cost-effective and efficient alternative for forest inventories, with potential for greater sustainability in management. A dataset containing field measurements and LiDAR point clouds was made publicly available to support future research.

In Chapter 6, we present the application of deep learning architectures for the segmentation of 3D point clouds of eucalyptus trees scanned by LiDAR, aiming for greater accuracy in forest inventories. We evaluated point-based models such as PointNet, PointNet++, DGCNN, PointMLP, and PointNeXt-S, as well as attention-based models like Point Transformer and Point Voxel Transformer (PVT), using a dataset of 55 clonal trees scanned with an iPhone 12 Pro Max. The results, measured by mIoU, showed that transformers (PVT and Point Transformer) achieved higher accuracy, while PointNeXt-S offered competitive performance with lower inference time, highlighting a trade-off between accuracy and efficiency. Despite the dataset's limitations, the models successfully segmented critical tree regions, reinforcing the potential of deep learning for structural measurements in forests and its feasibility on devices with lower computational power.

In Chapter 7, we address the challenge of precise 3D reconstruction of eucalyptus trees from point clouds in real forest environments, where occlusion, noise, and structural complexity compromise fundamental measurements. We evaluated seven state-of-the-art point cloud completion methods (FoldingNet, PCN, GRNet, SnowflakeNet, PoinTr, PointAttN, and ODGNet) using the new Euca3D dataset composed of terrestrial LiDAR scans. The models, pre-trained on ShapeNet, were adapted to the eucalyptus domain using a five-fold cross-validation protocol. Analysis using Chamfer Distance showed that transformer-based architectures, such as SnowflakeNet and PoinTr, performed best (43.08 and 69.16, respectively), preserving the global tree structure, though with limitations in reconstructing fine morphological details. The results also highlighted a performance drop compared to synthetic benchmarks, underscoring the challenges posed by real data

and the need for larger and more diverse datasets. In summary, the study establishes initial benchmarks and points to directions for domain-specific adaptations in forestry, essential for robust 3D reconstructions aimed at ecological monitoring and sustainable management.

In Chapter 8, we present a comparative study of five state-of-the-art generative architectures for synthesizing 3D point clouds of eucalyptus trees, a critical task for forest metrics such as height and DBH. To overcome the limitations of traditional acquisition methods, we introduce the Euca3D dataset, comprising 440 LiDAR scans collected in the field with an iPhone 14 Pro Max. We evaluated several model families, including l-GANs, PointFlow, a diffusion probabilistic model (DPM), and the autoregressive CanonicalVAE. The evaluation involved fine-tuning pre-trained ShapeNet models and assessing performance using metrics such as Chamfer Distance (CD) and Earth Mover's Distance (EMD). The results showed that the DPM achieved higher surface fidelity, while the CanonicalVAE better preserved structural integrity and perceptual quality, achieving performance nearly indistinguishable from real data. This study demonstrates that generating high-quality synthetic point clouds for complex natural shapes is feasible, providing a solid foundation for data augmentation and applications in forestry and computer vision, with greater objectivity and scalability than traditional methods.

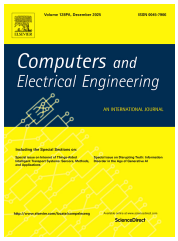
Chapter 9 represents the consolidation of the thesis, where the individual components developed in previous chapters are integrated into functional end-to-end pipelines. This chapter details and compares two experimental approaches for DBH measurement: the first reconstructs a complete 3D model from a partial and rapid scan, using depth super-resolution and point cloud completion techniques (SnowflakeNet); and the second directly processes a full 360° field scan. Both approaches converge on a shared workflow of ground removal and semantic segmentation with a PointTransformer to isolate the DBH region. Rigorous statistical validation demonstrates that both methods achieve accuracy statistically indistinguishable from manual measurements, with the direct scan method achieving an RMSE below 0.55 cm. More importantly, the chapter quantifies the transformative gain in operational efficiency, with up to a 30-fold reduction in data collection time, representing the main practical contribution of the work and validating the methodology's feasibility for forest inventory automation.

Finally, Chapter 10 presents the concluding remarks, synthesizing the results and contributions of the thesis. The chapter reaffirms the success of developing an automated, accurate, and low-cost pipeline for diameter measurement. Key findings, such as the statistical validation of accuracy and the significant efficiency gains, are discussed. Subsequently, the study's limitations are addressed, including the focus on a single species (Eucalyptus) within a plantation stand and the dataset size. Lastly, well-defined avenues for future research are outlined, including the expansion of the proposed

methodology to encompass additional species and native forest types, the estimation of further dendrometric parameters (e.g., height, volume, tortuosity), and the optimization of predictive models for deployment in field applications with real-time data processing, thereby advancing the concept of an accessible and efficient tool for precision forestry.

With the exception of Chapters 6, 7, and 9, whose manuscripts are still under review by journals, the remaining chapters have been published in conferences and journals, as presented below.

Chapter 2



Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Renato F. Bulcão-Neto, and Fabrizzio Soares. "Applications of artificial intelligence and LiDAR in forest inventories: A Systematic Literature Review." In Computers and Electrical Engineering 120 (2024) 109793, 2024. (RODRIGUES *et al.*, 2024b)

Chapter 3

Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, and Fabrizzio Soares. "Eucalyptus diameter and volume prediction with deep neural networks: A Long Short-Term Memory model approach." In Expert Systems With Applications 271 (2025) 126704, 2025. (RODRIGUES *et al.*, 2025)

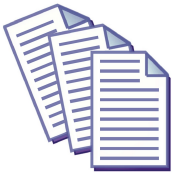
Chapter 4

Rodrigues, Welington G., Emilia A. Nogueira, Thamer H. Nascimento, Gabriel S. Vieira, Deborah S. A. Fernandes, and Fabrizzio Soares. "MobileNetV2-Enhanced Depth Map Super-Resolution Through Multi-scale Image Guidance." In 19th International Symposium, ISVC 2024 Lake Tahoe, NV, USA, October 21–23, 2024 Proceedings, Part II, Advances in Visual Computing, 2024. (RODRIGUES *et al.*, 2024a)

Chapter 5

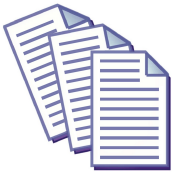
Rodrigues, Welington G., Gabriel Vieira, Afonso U. Fonseca, Juliana P. Felix, Deborah S. A. Fernandes, and Fabrizzio Soares. "3D Point Cloud Segmentation of LiDAR-Scanned Eucalyptus Trees: a Deep Learning Analysis." In 2025 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2025 (Waiting to be published in the proceedings)

Chapter 6



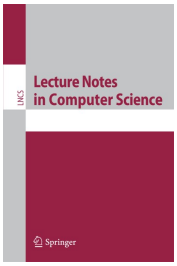
Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Helio Pedrini, Juliana P. Felix, and Fabrizzio Soares. "Accurate Measurement of Eucalyptus Tree Diameters Using Smartphone LiDAR and Deep Learning." (Under review).

Chapter 7



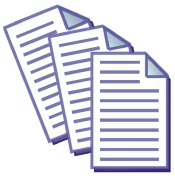
Rodrigues, Welington G., Christian D. Cabacinha, Fernandes, Deborah S. A. and Fabrizzio Soares. "3D Point Cloud Completion in Real-World Forestry: A Comparative Study for Eucalyptus Plantations. (Under review).

Chapter 8



Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Juliana P. Felix, Thamer H. Nascimento and Fabrizzio Soares. "Generating High-Fidelity Eucalyptus Point Clouds: A Comparative Study from GANs to Diffusion Models." In 19th International Symposium on Visual Computing (ISVC) 2025, Proceedings Part II, Advances in Visual Computing (Accepted for publication).

Chapter 9



Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Juliana P. Felix, Thamer H. Nascimento and Fabrizzio Soares. "3D Reconstruction and Segmentation of Eucalyptus Trees for Automated Diameter Measurement with Mobile LiDAR." (Under review).

Appendix A compiles a comprehensive listing of the original articles, developed software, and image datasets prepared by the authors throughout the research. This supplementary material aims to systematically document the outputs resulting from the work, serving both as a record of the scientific and technological contributions and as a resource to support reproducibility and the continuation of future studies.

During the research period, we explored various relevant lines of inquiry, including information retrieval and disparity in stereo images. We also conducted studies focused on human-computer interaction through smartwatches, as well as approaches to spell-checking on mobile devices. A detailed list of the results obtained is presented in Appendix A. In the same section, we present original software, accompanied by descrip-

tions and links to their respective repositories. These programs were developed to support the activities carried out in the research. Additionally, new datasets obtained from 3D scanners of eucalyptus trees are made available.

CHAPTER 2

Applications of Artificial Intelligence and LiDAR in Forest Inventories: A Systematic Review

> All things work for the good of those who love God.
>
> *Romans 8:28*

Forest inventory is a crucial tool for managing forest resources by providing quantitative and qualitative information about a particular region, much of which is collected manually in the field. Using devices such as Light Detection and Ranging (LiDAR) assists in collecting and analyzing various parameters of forest inventory. Adopting artificial intelligence (AI) techniques has sparked interest among forestry engineers seeking to work with forest LiDAR data. In this context, this study presents a Systematic Literature Review (SLR) to identify, evaluate, and interpret the results of primary studies related to the intersection between AI and Forestry Engineering. The automated search strategy retrieved 218 studies, of which 46 were selected after applying inclusion and exclusion criteria and quality assessment. After analyzing and synthesizing the data, the results showed that deep learning is becoming an increasing trend in recent research and that the direct estimation of tree diameter from aerial scans, although critical, has been minimally explored, highlighting an open field for future research.

Thus, this chapter presents a Systematic Literature Review on the integration of artificial intelligence and forestry engineering, with emphasis on the use of LiDAR data for forest inventory. The study highlights the growing adoption of deep learning techniques and identifies the limited exploration of tree diameter estimation from aerial scans as a promising research direction. This chapter was published in Computers and Electrical Engineering (RODRIGUES *et al.*, 2024b).

2.1 Introduction

In the field of forest engineering, the inventory is the basis for managing forest resources, whether native or planted forests (BRENA; NETO, 1997). Several qualitative and quantitative parameters can be obtained in the forest inventory, such as the volume, diameter, and height of trees, forest growth, canopy height, stem quality, and assortments. However, many of these measures are obtained in the field with manual equipment, which makes this activity time-consuming and costly. Furthermore, in some specific cases, tree felling may be necessary to obtain accuracy, leading to the prior disposal of the product.

In this way, it becomes essential to employ technologies that allow speeding up the process, such as LiDAR. In the field of remote sensing, LiDAR is a sensor capable of obtaining three-dimensional characteristics of forests on scales larger than those possible by traditional field methods (BOLTON *et al.*, 2018). LiDAR can provide aerial, terrestrial, and satellite data which can be used in the analysis of large forest areas (LEE *et al.*, 2018a). Because of that, the application of LiDAR in conjunction with AI has been the subject of research in various data collection tasks for forest inventory.

AI techniques and LiDAR data are often used to estimate and classify crucial variables in forest inventory such as height, diameter, and volume of trees. Very assertive results can be obtained which can be compared to traditional methods that use manual tools such as compasses, measuring tapes, diameter tapes, and clinometers. Since automatic approaches can handle a large volume of data, AI techniques are efficient to process and map entire areas of forests, plantations, and crop fields. In contrast, traditional methods work essentially in forest plots with statistical models to extrapolate information obtained from just a few samples of the total area.

2.1.1 LiDAR in the Forestry Sector

LiDAR is an active remote sensing sensor that emits laser signals to calculate distances based on the time delay of the returned laser pulses. They can record a dense array of distance return values assembled into the highly detailed 3D reconstruction of the surfaces. The data provided by LiDAR is a point cloud, which contains the three spatial coordinates of the intercepted points and additional characteristics such as reflectance, intensity or RGB values. These have several applications in Forestry, such as:

- Detailed surface modeling is employed to generate high-resolution digital elevation products for the planning and management of forests.
- LiDAR technology facilitates the acquisition of more precise information essential for accurate ecological and land use classification.

- LiDAR can map forests by measuring the vertical structures of the canopy and its density. These models aid in comprehending the intricate forest structures and producing precise forest inventories.
- Precision forestry, tailored to specific forest areas, can potentially enhance site productivity in terms of tree quality and overall yield.

LiDAR scanner scans can be conducted in various ways, including from satellites, aircraft, or the ground. Airborne Laser Scanning (ALS) provides a broad, comprehensive view of forests, enabling fast and efficient coverage of vast areas. On the other hand, the Terrestrial Laser Scanner (TLS) is ideal for obtaining precise, high-resolution details about the vertical structure of the forest, including tree height, canopy density, and terrain features. Figure 2.1 shows an example of the scanning process using the LiDAR sensor in data acquisition.

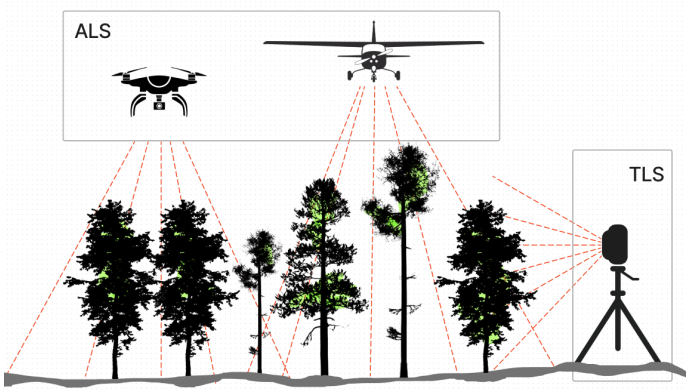


Figure 2.1: *Schematic representation of the ALS and TLS scanning methods.*

ALS

ALS works by emitting laser pulses from an airborne platform, such as an airplane, helicopter, and, in recent studies, drones (ALMEIDA *et al.*, 2020). These laser pulses travel to the Earth's surface, interacting with objects and terrain features. The time taken for the laser pulses to return to the sensor is recorded, allowing the calculation of distances between the sensor and the Earth's surface (BAZEZEW *et al.*, 2018).

During data collection, the ALS system emits laser pulses rapidly, often thousands of pulses per second. As the aircraft travels over the target area, the laser scanner records the time and intensity of each laser return. By measuring the time it takes for the laser pulse to return and knowing the speed of light, the ALS system can calculate the distance to the object or surface that reflects the laser pulse. The resulting georeferenced point cloud data allows for computation of a digital terrain model (DTM), digital surface model (DSM), canopy height model (CHM), or normalized DSM (nDSM).

The inherent limitation of ALS lies in the reduction of point density within the lower strata of forest vegetation due to significant reflection of the laser beam upon traversing the upper canopy (KORPELA *et al.*, 2012; KÜKENBRINK *et al.*, 2017). Consequently, methodologies prevalent in ALS applications tend to focus on the indirect estimation of understory coverage, often relying on statistical parameters derived from the distribution of reflections.

TLS

TLS, also known as ground-based LiDAR, is a remote sensing technique used to capture highly detailed three-dimensional information about objects and terrain features from a stationary position on the ground.

TLS works on a similar principle to ALS, but instead of being mounted on an airborne platform, the laser scanner is positioned on the ground. The TLS system emits laser pulses in a horizontal and/or vertical scanning pattern, covering the target area with laser beams. As the laser pulses interact with objects and surfaces, their reflections are recorded by the sensor.

In recent decades, several applications have been developed to use the TLS to replace conventional methods for measuring forest inventory attributes. The first step is, however, isolating every tree or stem in the point cloud. The attributes of each tree are then estimated, such as diameter at breast height (DBH) (SIMONSE *et al.*, 2003), tree height (SRINIVASAN *et al.*, 2015), biomass (HOLOPAINEN *et al.*, 2012), and volume (DASSOT *et al.*, 2012).

2.1.2 Review Goals and Paper Structure

This study presents a SLR that aims to identify, evaluate, and interpret results from primary studies in order to map and guide research at the intersection of AI and forest inventory using LiDAR. We organized the study into eight fundamental research questions (RQs) which include listing the contributions and types of AI algorithms, AI techniques/methods, forest parameters, type of LiDAR scanner used, and the achieved results in relation to conventional methods. To assist in the elaboration of the SLR, we used the *Parsif.al* tool[1] for planning and conducting the SLR protocol.

Based on the protocol, we performed an automatic search strategy in scientific databases and applied inclusion and exclusion criteria to identify relevant studies. We identified and analyzed 46 relevant studies by applying the protocol and classified them according to their objectives. We found that the most significant contributions were

[1] Avaliable in `https://parsif.al/`

model proposals for estimating and classifying various forestry parameters. Therefore, this document provides researchers with valuable guidance on how to apply AI models to forest inventory data obtained from LiDAR.

This review is organized as follows. Section 2.2 details the methodology used to plan and execute the SLR protocol; Section 2.3 presents a background about artificial intelligence; Section 2.4 reports the results obtained from the conduction of the SLR and analyzes the obtained data; Section 2.5 summarizes the SLR results; Section 2.6 discuses some limitations related to our SLR, and Section 2.7 presents final considerations and future studies.

2.2 Methodology

An SLR is a widely recognized and used method for identifying, evaluating, and interpreting relevant research about a specific question, area, or phenomenon of interest. An SLR is a widely recognized and used method for identifying, evaluating, and interpreting relevant research about a specific question, area, or phenomenon of interest (KEELE, 2007). The SLR is a secondary study whose objective is to survey research with similar scopes, critically evaluate them in their methodology, and combine them in a statistical analysis or meta-analysis when this is feasible.

Following the guidelines established by Keele (2007), the SLR methodology adopted in this study was structured into four main phases. First was the planning phase, where the studies' objectives and inclusion and exclusion criteria were defined. Next, the selection phase involved searching and screening the relevant articles according to the previously established criteria. Subsequently, relevant information from the selected studies was systematically extracted in the data extraction phase. Finally, the synthesis of results phase consisted of analyzing and aggregating the obtained data, enabling a comprehensive and coherent interpretation of the evidence found.

2.2.1 Planning

In the planning phase, some keywords related to forest inventory, artificial intelligence, and LiDAR, were selected to elaborate the search string engine. After defining the keywords and applying some pilot search tests, the search string was refined and the following string was defined:

(("forest inventory" OR "forest mensuration") AND ("machine learning" OR "deep learning" OR "neural network" OR "artificial intelligence") AND ("diameter" OR "volume" OR "height") AND ("lidar" OR "remote sensing")).

The search string was applied in the Scopus, EI Compendex, Web of Science, and IEEE Xplore scientific databases to answer the following research questions:

- **Q1**: What is the main contribution of AI techniques algorithms in the identified primary studies?
- **Q2**: What AI algorithms were used?
- **Q3**: What metrics were used to report the results?
- **Q4**: What AI approaches did the studies use?
- **Q5**: What types of forest inventory variables were obtained through the LIDAR scan?
- **Q6**: What LiDAR scanning methods were performed?
- **Q7**: In which type of forest was the study carried out?
- **Q8**: Are the datasets used publicly available?

2.2.2 Selection

Following the recommendations of Keele (2007), we defined selection criteria for primary studies, consisting of inclusion (IC) and exclusion (EC) criteria.

Inclusion criteria

Studies that met the following inclusion criteria were selected:

- **IC1**: The study proposes the use of AI and LiDAR

Exclusion criteria

Studies that met any of the 5 exclusion criteria were removed:

- **EC1**: The article does not deal with AI with LIDAR.
- **EC2**: The article is not a conference or journal publication.
- **EC3**: The article does not describe a primary study.
- **EC4**: The article was published more than 5 years ago (it must be between 2018 and 2022).
- **EC5**: The article is not in the English language.

Figure 2.2 presents the results obtained after applying the selection phase. Initially, 162 articles were obtained from Scoups database, 31 from EI Compendex, 17 from WoS, and 8 from IEEE Xplorer. From the 218 retrieved studies, 50 were removed because there was more than one entry from the same paper in the different databases. The remaining 168 articles were evaluated considering the inclusion and exclusion criteria and a total of 46 studies were obtained for further analysis.

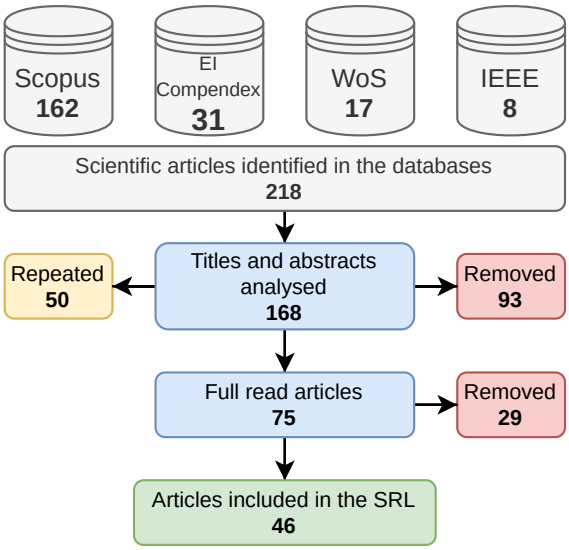


Figure 2.2: *Stages of study selection.*

Quality Criteria (QCs)

Assessing the methodological quality of a primary study involves evaluating the importance of individual studies considering the obtained results. To investigate whether differences in study quality can account for differences in study results, we evaluated the following criteria:

- **QC1**: Search type.
- **QC2**: Field application.
- **QC3**: Dataset availability.

The quality criteria were scored as follows:

- (QC1) According to the classification by Petersen *et al.* (2015), research classified as solution proposals may still have results that are insufficiently validated, thus their maturity level should be considered lower than that of validation and evaluation research. Studies whose results are based on validation methods but have not been applied in practice have less maturity than studies whose results are derived from evaluations conducted by industry professionals, also known as evaluation surveys. Therefore, we assigned scores for the different types of studies as follows: 0 points for solution proposals, 1 point for validation proposals, and 2 points for evaluation surveys.
- (QC2) Studies aimed at evaluating new methodologies and comparing them to traditional methods used in the field were deemed more significant because new approaches may yield superior results or require fewer resources while still being sufficiently adequate. Therefore, we assigned 2 points to studies that applied new

methodologies in the field, 1 point to studies that partially applied it, and 0 points to studies that did not present such novelty.

- (QC3) The availability of the dataset is important for the future reproduction of the experiment. Therefore, this criterion seeks to give greater importance to datasets that are available for download and thus replicate the study. Studies with available datasets received 1 point, and studies with unavailable datasets received 0 points.

Once the quality criteria are presented, the following final quality index formula presents a numerical quantification as a means of classifying the relevant primary studies.

$$S = 2\sum_{i=1}^{2} QC_i + QC_3, \tag{2-1}$$

where QC represents the point attributed to the quality criterion, i represents the index of the applied criterion, and S represents the score assigned to the study considering the quality criteria.

After presenting the process of defining and evaluating the quality criteria, studies that obtained a score lower than two were removed from the SLR. Table 2.1 presents the results of the primary studies, including the respective scores assigned after rigorous application of the quality criteria. Primary studies that did not meet the established quality criteria or obtained scores lower than 2 were excluded from this study.

Table 2.1: *List of studies selected after applying the quality criteria and attribution of grades.*

Score (S)	ID	Articles
2	S1	(ESTEBAN *et al.*, 2020a)
3	S2	(SILVA *et al.*, 2018)
4	S3-16	(ZHOU *et al.*, 2018; **??**; PEARSE *et al.*, 2019; LAMB *et al.*, 2018) (WITTKE *et al.*, 2019; KUKUNDA *et al.*, 2018; CORTE *et al.*, 2020a; QUEINNEC *et al.*, 2021) (SHENDRYK; GORROD, 2020; LEE *et al.*, 2018b; PARK *et al.*, 2021; OZKAN *et al.*, 2022) (SUN *et al.*, 2022; HAUGLIN *et al.*, 2018)
5	S17-29	(HUDAK *et al.*, 2020; MAHONEY *et al.*, 2018; HEINZEL; GINZLER, 2018; NGUYEN *et al.*, 2021) (CACCAMO *et al.*, 2018; OEHMCKE *et al.*, 2022; WINDRIM; BRYSON, 2020a; CAO *et al.*, 2019) (HOOVER *et al.*, 2018; HAO *et al.*, 2021; ZHANG *et al.*, 2019; OBATA *et al.*, 2021) (ALON *et al.*, 2019)
6	S30-38	(ASTOLA *et al.*, 2021; SHANG *et al.*, 2019; SANCHEZ-RUIZ *et al.*, 2019; TIJERÍN-TRIVIÑO *et al.*, 2022) (BEAUDOIN *et al.*, 2022; MEDINA; PAFFENROTH, 2021; BOLTON *et al.*, 2020; CASTILLA *et al.*, 2022) (CHIRICI *et al.*, 2020)
7	S39-45	(SHOOT *et al.*, 2021; NOVO-FERNÁNDEZ *et al.*, 2019; EHLERS *et al.*, 2022) (XIE *et al.*, 2020a; CHEN *et al.*, 2020a; MAYRA *et al.*, 2021) (PERSSON *et al.*, 2022)
8	S46	(AYREY; HAYES, 2018)

Data extraction

By performing a comprehensive reading of the 46 articles selected in our SLR, the following contents corresponding to the research questions for each of the articles:

1. the type of contribution of the studies;
2. the most evaluated algorithms;
3. the techniques/methods used;
4. the type of forest (native or planted).
5. the most worked variables;
6. the most reported metrics; and
7. the availability of the dataset.

A detailed categorization of the studies reviewed is presented in Table 2.2 for further consultation.

Table 2.2: *Identification of studies included in the SLR*

ID	Paper	Algorithm	Contribution	Scanning Task	A.I. Parameter	Forest Parameter	Metric	Dataset	Forest Type
S1	(ESTEBAN *et al.*, 2020a)	RF	Mdl	A	C	Vol	Acc	Private	N
S2	(SILVA *et al.*, 2018)	KNN	Mdl	A	R	Vol	Pre, RMSE	Private	P
S3	(ZHOU *et al.*, 2018)	DT, RF	Frm, Mdl, Met	A	C, S	Tc	Acc	Private	N
S4	(**??**)	DT, RF, SVM	Frm, Mdl	A	C	Ht	Acc	Private	N
S5	(PEARSE *et al.*, 2019)	RF	Mdl	A	R	Ht, Vol	R^2, RMSE	Private	P
S6	(LAMB *et al.*, 2018)	KNN	Mdl	A	R	Ht, Dia, Vol	RMSE	Private	P
S7	(WITTKE *et al.*, 2019)	RF	Mdl, Met	A	R	Ht, Bio, Dia, Vol	R^2, RMSE	Private	P
S8	(KUKUNDA *et al.*, 2018)	DT, RF, NN	Mdl, Met	A	C, R, S	Ht, Tc	ROC, Acc	Private	N
S9	(CORTE *et al.*, 2020a)	RF, NN, SVM	Met	A	C, R	Ht, Dia, Vol	R^2, RMSE	Private	P
S10	(QUEINNEC *et al.*, 2021)	KNN	Alg, Mdl, Met	A	R	Ht	RMSE	Private	N
S11	(SHENDRYK; GORROD, 2020)	RF	Mdl	A	R	Vol	R^2, RMSE	Private	N
S12	(LEE *et al.*, 2018b)	DT, RF, SVM	Mdl, Met	A	R	Ht	R^2, RMSE	Private	P
S13	(PARK *et al.*, 2021)	DL	Mdl, Met	A	C	Cla	ROC, Pre, Rec	Private	N
S14	(OZKAN *et al.*, 2022)	RF	Mdl, Met	A	R	Ht, Bio, Dia, Vol	R^2, RMSE	Private	N
S15	(SUN *et al.*, 2022)	RF, NN, SVM	Mdl	A	R	Dia	R^2, RMSE	Private	P
S16	(HAUGLIN *et al.*, 2018)	RF	Mdl	A	R	Vol	MAE, R^2, RMSE	Private	P
S17	(HUDAK *et al.*, 2020)	RF	Mdl	A	C, S	Bio	R^2, RMSE	Private	N
S18	(MAHONEY *et al.*, 2018)	KNN	Frm, Mdl	A	R	Ht	RMSE	Private	N
S19	(HEINZEL; GINZLER, 2018)	NN	Frm, Met	T	R, S	Dia, Vol	MAE	Private	N
S20	(NGUYEN *et al.*, 2021)	RF	Alg, Frm, Mdl, Met	T	C, S	Cla, Dia	F1-Score, Pre, Rec	Private	N
S21	(CACCAMO *et al.*, 2018)	RF	Mdl, Met	A	R	Vol	RMSE	Private	P
S22	(OEHMCKE *et al.*, 2022)	DL, RF	Mdl	A	R	Bio, Vol	MAPE, R^2, RMSE	Private	N
S23	(WINDRIM; BRYSON, 2020a)	CNN, DL, NN	Alg, Frm, Mdl, Met	A	C, O, R	Ht, Tc, Dia	Acc, F1, Pre, Rec, RMSE	Private	P
S24	(CAO *et al.*, 2019)	RF	Mdl	A	R	Bio	R^2, RMSE	Private	P
S25	(HOOVER *et al.*, 2018)	RF	Mdl	A	R	Ht, Bio	R^2, RMSE	Private	P
S26	(HAO *et al.*, 2021)	KNN, RF	Frm, Mdl	A	C, O, R	Dia	RMSE	Private	P
S27	(ZHANG *et al.*, 2019)	KNN, RF	Mdl	A	R	Dia	R^2, RMSE	Private	P
S28	(OBATA *et al.*, 2021)	RF	Mdl	A	R	Vol	RMSE	Public	P
S29	(ALON *et al.*, 2019)	CNN, DL, NN	Frm, Mdl, Met	A	O, S	Tc	Pre, Rec	Private	N
S30	(ASTOLA *et al.*, 2021)	DL, RF, NN	Mdl	A	R	Vol	RMSE	Public	N
S31	(SHANG *et al.*, 2019)	RF	Mdl	A	R	Vol	R^2, RMSE	Private	P
S32	(SANCHEZ-RUIZ *et al.*, 2019)	RF	Mdl	A	R	Vol	RMSE	Public	N
S33	(TIJERÍN-TRIVIÑO *et al.*, 2022)	KNN, RF	Mdl	A	C, R	Cla	Acc	Public	N
S34	(BEAUDOIN *et al.*, 2022)	KNN	Mdl	A	R	Bio	R^2, RMSE	Public	N
S35	(MEDINA; PAFFENROTH, 2021)	KNN, RF, NN	Mdl	A	C	Cla	F1-Score, Pre, Rec	Public	N
S36	(BOLTON *et al.*, 2020)	RF	Mdl	A	R	Ht, Vol	R^2, RMSE	Private	P
S37	(CASTILLA *et al.*, 2022)	KNN	Frm, Mdl	A	C, R	Bio, Cla, Vol	R^2, RMSE	Private	N
S38	(CHIRICI *et al.*, 2020)	KNN, RF	Mdl	A	R	Vol	R^2, RMSE	Public	P
S39	(SHOOT *et al.*, 2021)	KT, KNN, RF, SVM	Mdl, Met	A	C	Cla	Acc	Public	N
S40	(NOVO-FERNÁNDEZ *et al.*, 2019)	DT, KNN, RF	Mdl	A	R	Vol	R^2, RMSE	Public	P
S41	(EHLERS *et al.*, 2022)	RF	Mdl	A	R	Bio	R^2, RMSE	Public	N
S42	(XIE *et al.*, 2020a)	KNN, RF	Mdl	A	R	Vol	MAE, R^2, RMSE	Public	N
S43	(CHEN *et al.*, 2020a)	DL	Alg, Frm, Mdl	A	O, S	Dia	Acc	Private	P
S44	(MAYRA *et al.*, 2021)	CNN, DL, RF, NN, SVM	Frm, Mdl, Met	A	C, O, S	Cla, Tc	Acc, F1-Score	Public	N
S45	(PERSSON *et al.*, 2022)	CNN, DL	Mdl	A	C, R	Ht, Dia, Vol	R^2, RMSE	Private	P
S46	(AYREY; HAYES, 2018)	CNN, DL, RF, NN	Frm, Mdl	A	R, S	Bio, Tc, Vol	RMSE	Public	P

Models (Mdl), algorithm (Alg), framework (Frm), method (Met), aerial (A), terrestrial (T), classification (C), segmentation (S), regression (R), object detection (O), biomass (Bio), volume (Vol), diameter (Dia), species classification (Cla), tree count (Tc), height (Ht), determination coefficient (R^2), root mean square error (RMSE), accuracy (Acc), mean absolute percentage error (MAPE), receiver operating characteristic curve (ROC), native (N), planted (P).

2.3 Background on Artificial Intelligence

Before presenting the results of the review, it is essential to introduce some fundamental concepts related to artificial intelligence.

Artificial intelligence is a field of computer science that focuses on creating systems capable of performing tasks that typically require human intelligence, such as learning, reasoning, perception, problem solving, and language understanding (LECUN *et al.*, 2015).

AI refers to the ability of a digital computer or computer-controlled robot to perform cognitive tasks typically associated with intelligent beings. It has gained increasing relevance in several fields, with numerous applications that have generated significant advances in several sectors (ALBAHRI *et al.*, 2024).

Machine learning (ML) is a subfield of artificial intelligence that empowers computers to discern patterns and associations within data through the application of mathematical models. This process involves training algorithms on datasets that may be either labeled (annotated) or unlabeled (unannotated), enabling the algorithms to learn from experience and subsequently make predictions or decisions based on input data. ML techniques are typically categorized as either supervised or unsupervised, contingent upon whether the training involves labeled data (SACHDEVA *et al.*, 2023).

Recent advancements in computational power and the increased availability of memory in modern computer architectures have led to the emergence of an enhanced approach to traditional machine learning, known as Deep Learning (DL).

Deep learning represents a paradigm shift in how machines learn patterns from data, as it automatically extracts salient features for a given computational task rather than relying on human intuition. Deep learning methods are characterized by architectures composed of multiple layers that apply mathematical transformations to the data as it progresses through the network. These transformations are governed by sets of learnable parameters, which are adjusted using various learning and optimization algorithms. These computational layers and parameters collectively constitute models that can be trained for specific tasks by iteratively updating the parameters based on the model's performance on a set of training data (WATSON *et al.*, 2022).

ML and DL are two subfields within the broader domain of AI, in which ML focuses on identifying patterns and associations within data through mathematical models. In contrast, deep learning, or Deep Neural Networks (DNNs), involves architectures where "deep" refers to the number of hidden layers within the network. Various types of deep networks exist, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Recursive Neural Networks (RevNNs), and Long Short-Term Memory networks (LSTMs), among others (ALZUBAIDI *et al.*, 2021).

2.4 Results

Figure 2.3 illustrates the distribution of selected articles according to their respective year of publication. Notably, between 2018 and 2022, there was considerable fluctuation in publications about integrating LiDAR and AI technologies. Specifically 2018, 12 articles were identified, followed by a decline of 33% and 42% in 2019 and 2020, respectively. However, there was a resurgence in publications in 2021, matching the count observed in 2018. Conversely, in 2022, the number of publications decreased again, with only seven articles identified.

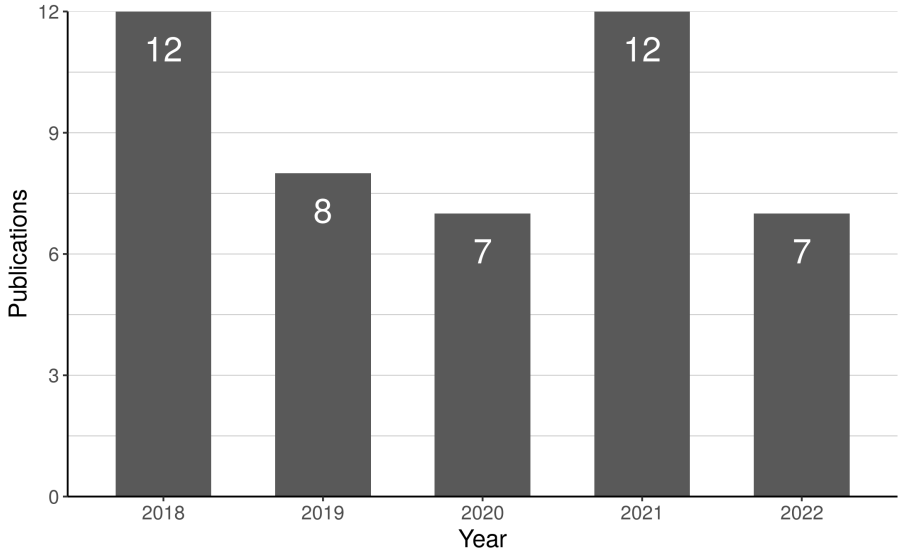


Figure 2.3: *Articles per year discussing AI and LiDAR according to our SLR.*

2.4.1 Data synthesis

This section offers a synthesis of data extracted from primary studies pertinent to addressing the research inquiries.

Main contributions (Q1)

The analyzed studies adopt multiple proposals for data processing, including recommendations for algorithms, new methods, models, and frameworks. Figure 2.4 illustrates the correlation between the number of publications and their primary contributions. Most articles presented AI models to handle data extracted by LiDAR sensors. Others presented methods and frameworks to support learning models for LiDAR approaches. On the other hand, new algorithms were presented in fewer papers.

Out of the 46 studies scrutinized, 44 (95%) delineated AI models tailored towards forest analysis, while 15 (32%) methods were devised to address LiDAR variables or explore novel strategies for noise reduction, facilitating the examination of innovative

approaches. Additionally, frameworks were prominently employed, evidenced by their usage in 12 articles (ZHOU *et al.*, 2018; BOUTSOUKIS *et al.*, 2019; LAMB *et al.*, 2018; MAHONEY *et al.*, 2018; HEINZEL; GINZLER, 2018; NGUYEN *et al.*, 2021; WINDRIM; BRYSON, 2020a; HAO *et al.*, 2021; ALON *et al.*, 2019; CASTILLA *et al.*, 2022; CHEN *et al.*, 2020a; MAYRA *et al.*, 2021; AYREY; HAYES, 2018). In contrast, only four algorithms were introduced for process automation and addressing specific challenges, as documented in (NGUYEN *et al.*, 2021; WINDRIM; BRYSON, 2020a; CHEN *et al.*, 2020a). Additionally, efforts were made to explore forest structures and cultivate a computational library, as proposed by Queinnec *et al.* (2021).

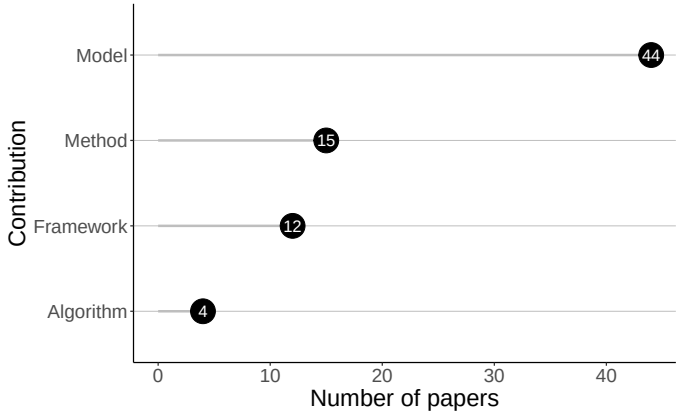


Figure 2.4: *Main contributions*

The multifaceted nature of the proposed contributions is most effectively visualized through a Venn diagram, as depicted in Figure 2.5. The inherent complexity of relationships in diagrams with more than three sets precludes circular representations. Consequently, this study employs ellipses to delineate the four distinct sets. The confluence of these sets, as represented by their intersection, encapsulates studies that concurrently employ all outlined approaches. Notably, two studies, Nguyen *et al.* (2021) and Windrim and Bryson (2020a), have been identified as incorporating an AI model, method, framework, and algorithms comprehensively.

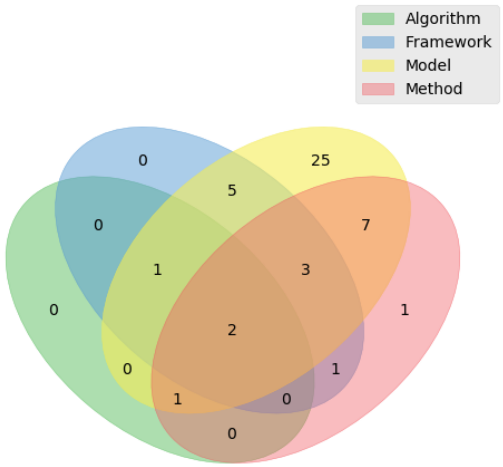


Figure 2.5: *Representation of the number of studies that apply more than one contribution simultaneously.*

AI algorithms (Q2)

This research question points out the AI algorithms applied in the studies. The aim is to discern and chart the trajectory of algorithmic trends that have emerged in recent scholarly contributions. These trends are categorized based on the algorithmic typologies employed within the articles. Given the multifaceted nature of algorithmic applications, it is not uncommon for a single study to be featured across multiple categories. For instance, the study by Lee *et al.* (2018b) encompasses the utilization of Random Forest (RF), Support Vector Machines (SVM), and Decision Trees (DT) within its methodology, thereby warranting its inclusion in three distinct algorithmic classifications.

Figure 2.6 delineates the prevalence of various algorithms within the corpus of selected studies. RF algorithm emerges as the preeminent choice, employed in 34 distinct studies. This predilection can be attributed to RF's efficacy in addressing classification and regression challenges. Moreover, its implementation simplicity starkly contrasts the intricacies associated with more complex algorithms. Implementing RF across diverse programming languages and libraries further cements its status as a versatile tool. In the realm of classification, RF's utility is exemplified in the categorization of tree species. Conversely, in regression scenarios, it is adept at estimating continuous variables, such as height, biomass, and volume.

The K-Nearest Neighbors (KNN) algorithm is the second most frequently utilized algorithm in the reviewed literature. Its implementation is widespread across numerous libraries and is widely used in forestry applications. A significant proportion of the studies leveraged KNN for the development of models aimed at variable estimation, as evidenced by (SUN *et al.*, 2022; HAO *et al.*, 2021; ZHANG *et al.*, 2019; NOVO-

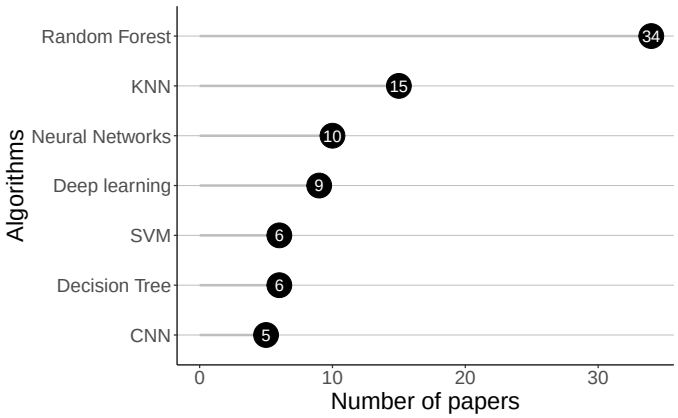


Figure 2.6: *AI algorithms used in selected studies.*

FERNÁNDEZ *et al.*, 2019; XIE *et al.*, 2020a). Additionally, KNN has been instrumental in classifying variables, a technique employed in studies such as (TIJERÍN-TRIVIÑO *et al.*, 2022; MEDINA; PAFFENROTH, 2021; CHIRICI *et al.*, 2020; SHOOT *et al.*, 2021).

The integration of DT and SVM was explored in Lee *et al.* (2018b) to estimate tree heights utilizing LiDAR data. This combination was also applied to address classification challenges in studies such as (BOUTSOUKIS *et al.*, 2019; LAMB *et al.*, 2018; SHOOT *et al.*, 2021). Neural networks (NN) and deep learning have similar applications. However, deep learning specifically refers to neural network architectures with multiple layers. Among the various deep learning architectures, CNNs have played a crucial role in research involving the conversion of point clouds into image data, enabling the application of advanced techniques for tasks such as classification, segmentation, and object detection. Table 2.3 presents the main algorithms used in different tasks and relevant studies related to these methods.

Table 2.3: *Algorithms adopted by task in the studies*

Tasks	Algorithms	Studies	Total
Regression	RF, KNN, NN, DL, SVM	S2, S5, S6, S7, S8, S9, S10, S11, S12, S14, S15, S16, S18, S19 S21, S22, S23, S24, S25, S26, S27, S28, S30, S31, S32, S33, S34 S36, S37, S38, S40, S41, S42, S45, S46	35
Classification	RF, KNN, NN, DT, DL SVM, CNN	S1, S3, S4, S8, S9, S13, S17, S20 S23, S26, S33, S35, S37, S39, S44, S45	16
Segmentation	CNN, NN, DL	S3, S8, S17, S19, S20, S29, S43, S44, S4	9
Object Detection	CNN, NN, DL	S23, S26, S29, S43, S44	5

Evaluated metrics (Q3)

In our investigation, we scrutinized the evaluation metrics utilized across the studies. We observed the absence of benchmarks tailored to specific tasks, such as species classification or estimating other variables. While many approaches encompassed

classical AI algorithms, others showcased more intricate neural networks. Consequently, the metrics presented herein solely reflect the prevalent evaluation measures employed. The distribution of these metrics is depicted in Figure 2.7.

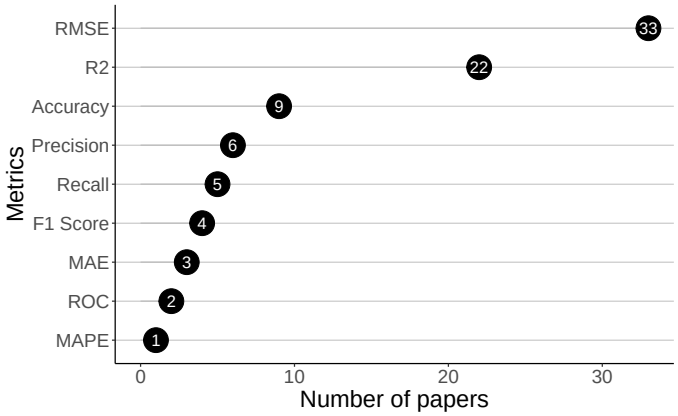


Figure 2.7: *Type of performance metrics used in the selected studies.*

Most algorithms, such as RF and KNN, were employed to develop models to estimate continuous variables. Within forestry, numerous statistical models were also utilized to evaluate continuous variables. Consequently, the most prevalent evaluation metrics included the Root Mean Square Error (RMSE) and the coefficient of determination R^2. While Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were also utilized for this purpose, they were typically presented alongside RMSE and R^2, as demonstrated in studies such as (HEINZEL; GINZLER, 2018; XIE *et al.*, 2020a; HAUGLIN *et al.*, 2018; OEHMCKE *et al.*, 2022).

Studies concentrating on tree detection and individual tree feature extraction have addressed various metrics for classification tasks. Accuracy emerged as the most frequently observed measure, followed by Precision, Recall, F1-Score, and the area under the Receiver Operating Characteristic (ROC) curve. These metrics were utilized to evaluate 3D modeling and NN, as evidenced in studies such as (NGUYEN *et al.*, 2021; WINDRIM; BRYSON, 2020a; MEDINA; PAFFENROTH, 2021; MAYRA *et al.*, 2021).

Types of AI approaches (Q4)

This question seeks to elucidate the prevalent techniques for estimating forest parameters, encompassing qualitative and quantitative variables. Noteworthy quantitative variables include height, volume, biomass, and the number of trees. The commonly employed techniques for estimation include classification, regression, segmentation, object detection, and clustering. It is pertinent to highlight that individual studies may employ multiple approaches to address the designated problem. However, our SLR uncovered no studies that simultaneously employed these techniques.

Figure 2.8 illustrates the principal techniques for estimating tree parameters. Most studies focus on resolving regression (35 papers, 76%) problems, aiming to estimate continuous variables. For instance, (OEHMCKE *et al.*, 2022; ASTOLA *et al.*, 2021; AYREY; HAYES, 2018; CHIRICI *et al.*, 2020) leveraged deep learning methodologies to estimate volume stock. Additionally, attributes such as biomass and crown size were estimated in studies such as (OEHMCKE *et al.*, 2022; ALON *et al.*, 2019; AYREY; HAYES, 2018).

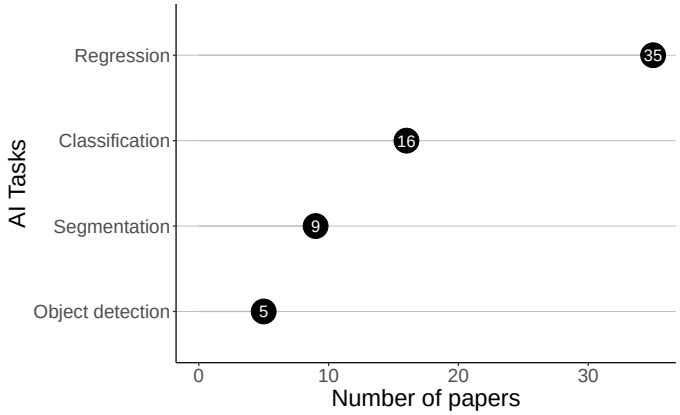


Figure 2.8: *Approaches used in studies.*

Studies addressing classification (16 papers, 34%) tasks were undertaken to estimate the number of trees within a defined window or to classify trees and forests within specific regions, as documented in (PARK *et al.*, 2021; NGUYEN *et al.*, 2021; TIJERÍN-TRIVIÑO *et al.*, 2022; MEDINA; PAFFENROTH, 2021; CASTILLA *et al.*, 2022; SHOOT *et al.*, 2021; MAYRA *et al.*, 2021). The utilization of LiDAR technology for tree counting was explored in studies such as (ZHOU *et al.*, 2018; KUKUNDA *et al.*, 2018; WINDRIM; BRYSON, 2020a; MAYRA *et al.*, 2021). Moreover, in Mayra *et al.* (2021), various approaches were employed to classify *Pinus* species, while regression techniques were applied to estimate tree attributes.

Studies that address object detection in their analyses often deal concomitantly with segmentation tasks for classifying regions or tree species. In the study by Windrim and Bryson (2020a), all of these approaches are worked on simultaneously.

Forest parameters (Q5)

This research question aims to identify the parameters, including height, diameter, volume, tree canopy size, and others, that were predominantly investigated. These parameters constitute vital variables collected during forest inventory processes. Given the interconnected nature of many characteristics, such as height, volume, and diameter of trees, several studies delve into multiple parameters simultaneously. Figure 2.9 elucidates the interrelationship between the parameters explored in the selected studies.

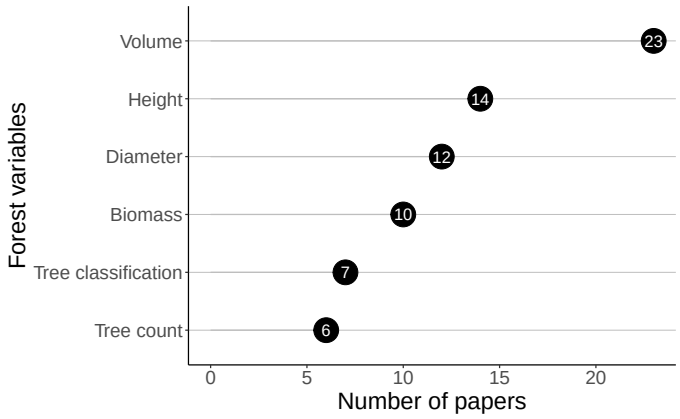


Figure 2.9: *Distribution of evaluated parameters.*

In the domain of forest inventory, the computation of the total and commercial volume of trees stands as a pivotal undertaking. This variable is essential to populations undergoing environmental licensing or management planning. Consequently, selecting the most suitable models for evaluating the study area constitutes a fundamental task in inventory management. This prominence of volume estimation is evidenced in Figure 2.9, where it emerges as the most extensively explored variable.

The calculation of tree height involves measuring the linear distance between the ground and the highest point of its canopy, a crucial variable for analyzing the tree's growth site and estimating parameters like height increment, volume, and basal area. Several studies, such as Boutsoukis *et al.* (2019) and Lamb *et al.* (2018), propose methods to approximate forest canopy height, while others, like Lamb *et al.* (2018) and Windrim and Bryson (2020a), assess commercial height estimation techniques. Height estimation often complements other variables to estimate additional characteristics, such as volume and basal area Wittke *et al.* (2019), Ozkan *et al.* (2022), Hudak *et al.* (2020), Bolton *et al.* (2020), height and volume Corte *et al.* (2020a), and volume and diameter Ozkan *et al.* (2022), Persson *et al.* (2022).

The diameter constitutes a dendrometric measurement that can be directly assessed and utilized to compute parameters such as cross-sectional area, basal area, form quotient, volume, and assortment estimates, enabling the derivation of multiple products. Thus, it is a fundamental characteristic of forest inventory, given its accessibility and high correlation with other variables. Recognizing its significance, Heinzel and Ginzler (2018), Nguyen *et al.* (2021), Chen *et al.* (2020a) evaluate the diameter prediction employing 3D models. Furthermore, studies such as Zhou *et al.* (2018), Sun *et al.* (2022), Hao *et al.* (2021) identify and estimate diameter. Additionally, several other studies investigate diameter in conjunction with previously described variables.

Forest biomass, a renewable organic material derived from forests, holds significant potential for energy production, serving as an alternative to finite energy sources such

as petroleum. Its utilization gains increasing relevance due to its effectiveness in mitigating greenhouse gas emissions. In this context, Hudak *et al.* (2020) introduces a carbon monitoring system to map aboveground biomass. DL techniques were also employed in Wittke *et al.* (2019), Ayrey and Hayes (2018) for forest biomass estimation. Furthermore, a variety of LiDAR and satellite image alternatives for assessing biomass potential are explored in studies such as Mahoney *et al.* (2018), Cao *et al.* (2019), Hoover *et al.* (2018), Beaudoin *et al.* (2022), Ehlers *et al.* (2022).

During the forest inventory process of native forests, alongside traditional procedures like measurement, a critical step involves identifying existing species. This phase is of immense importance as it enables the assessment of the study area, facilitates efficient forest mappings, and provides an in-depth analysis of forest structure and composition. In the studies retrieved through the SLR, this approach is frequently coupled with tree counting, necessitated by isolating individual trees for subsequent identification purposes. In Zhou *et al.* (2018), this approach is implemented by analyzing different wavelengths. Furthermore, Kukunda *et al.* (2018) and Shoot *et al.* (2021) compare different algorithms for species identification. In Medina and Paffenroth (2021), 3D modeling is utilized for species identification, while in Park *et al.* (2021), Alon *et al.* (2019), Mayra *et al.* (2021), deep learning techniques are employed for the same purpose.

Figure 2.10 illustrates the publications from the last five years, indicating a notable emphasis on volume measurement, primarily attributable to its practical applications and significance in forest inventory. However, in recent years, owing to advancements in AI techniques and the emergence of additional measures such as height and species identification, there has been an augmented focus on species identification and tree counting.

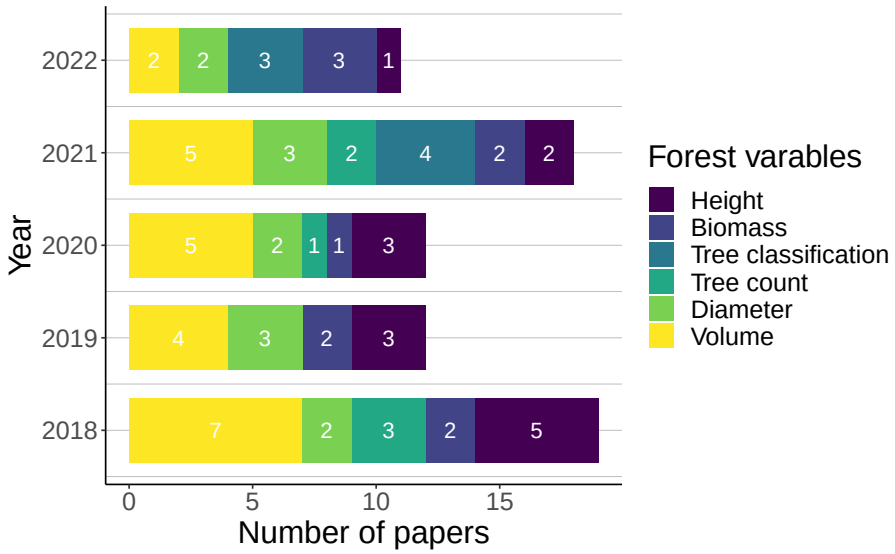


Figure 2.10: *Main forest parameters investigated in the period.*

Scanner methods (Q6)

This section examines the primary types of LiDAR utilized in the selected studies. Aerial scanning emerges as a prevalent approach, owing to its capacity to efficiently cover expansive areas.

Several studies integrate data from aerial scans with information from various scanning sources, including images, infrared data, and satellite datasets such as Landsat Esteban *et al.* (2020a), Hudak *et al.* (2020), Obata *et al.* (2021), Sanchez-Ruiz *et al.* (2019), Beaudoin *et al.* (2022), Bolton *et al.* (2020), Castilla *et al.* (2022), Chirici *et al.* (2020), Sentinel Bolton *et al.* (2018), Wittke *et al.* (2019), Astola *et al.* (2021), WorldView-2 Wittke *et al.* (2019), Kukunda *et al.* (2018), Ozkan *et al.* (2022), and the Global Ecosystem Dynamics Investigation (GEDI) (CAO *et al.*, 2019). These studies aim to assess the feasibility of utilizing multi-temporal multispectral satellite images (orthophotos) for forest inventories and compare their efficacy with data acquired via ALS.

Notably, the selected studies incorporating satellite data are conducted concurrently with data obtained from aerial scans (44 paper, 96%). Therefore, satellite scans were classified within the same category as aerial scans. Regarding the studies examined, we found that only two (2 papers, 4%) of them used a TLS technique (HEINZEL; GINZLER, 2018; NGUYEN *et al.*, 2021).

Figure 2.11 illustrates the proportion of studies using orthophoto and LiDAR images as primary data for comparison. Twenty studies (43%) used multispectral images to compare with data obtained by aerial scans, while 26 studies (57%) used airborne LiDAR data directly.

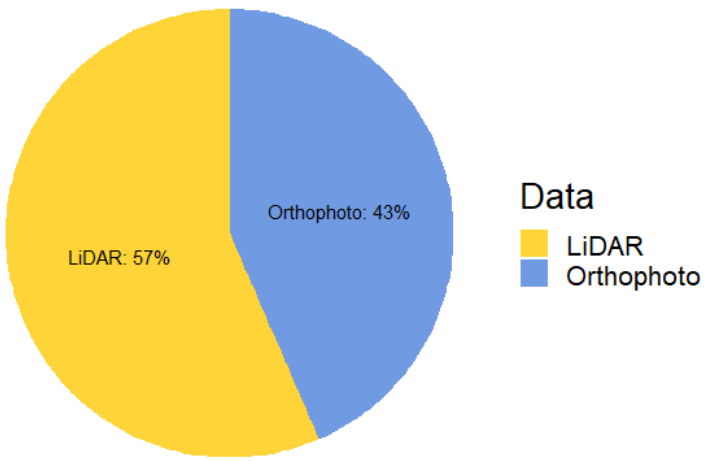


Figure 2.11: *Most common types of LiDAR scans.*

Forest type (Q7)

The SLR revealed a greater volume of research focusing on the amalgamation of AI and LiDAR technologies within native forests, accounting for 24 articles (52%), as opposed to 22 articles (48%) concentrating on planted forests.

Despite the inherent homogeneity of planted forests compared to their native counterparts, there is a notable apprehension regarding deploying computational models in areas with restricted accessibility. This hesitation stems from the necessity to rationalize the extensive preparatory efforts, data collection endeavors, and associated expenses intrinsic to adopting cutting-edge technologies.

Native forests, with their intrinsic unstructured complexity, pose significant obstacles that thwart the application of traditional research methodologies. The insights gleaned from our SLR highlight the benefits of employing AI and LiDAR technologies in settings marked by heterogeneous topography, a diverse array of tree species, and fluctuating climatic conditions.

In juxtaposition, the utilization of AI and LiDAR modalities in planted forests is primarily associated with prospects for augmenting the precision and economic efficiency of forest inventory protocols, as referenced in Sun *et al.* (2022), Hao *et al.* (2021), Zhang *et al.* (2019), Chen *et al.* (2020a).

Data availability (Q8)

Data availability is crucial to ensure reproducibility in a study, especially in fields such as machine learning, where new ideas disseminate rapidly. This study assessed the accessibility of data utilized in the selected works.

Public data made available in selected studies are generally presented in .laz format. This format represents a common standard for storing point cloud data obtained by LiDAR systems. It records information regarding the three-dimensional location of each point in the cloud and possible additional attributes, such as laser return intensity, terrain classification, and chromatic information. Including these additional attributes varies depending on the LiDAR equipment used in the data collection process. Adopting the .laz format is widespread due to its compatibility with various point cloud processing software, enabling practical analysis, visualization, and manipulation of data obtained through LiDAR.

Upon reviewing the 46 selected studies, it was observed that 28% (13) utilized public data, whereas 72% (33) relied on private data. Public data sources typically stemmed from government agencies, although certain companies also provided access to their datasets, which encompassed LiDAR and field survey data. However, most data remained inaccessible due to private entities' ownership or unavailability within particular

countries. Table 2.4 presents a list of the publicly available datasets, a link to access the datasets, and the respective primary studies.

Table 2.4: *List of Publicly Available LiDAR Datasets*

Article	Dataset	Access Link
(OBATA *et al.*, 2021)	FIA Datamart	<https://www.fs.usda.gov/research/products/dataandtools/tools/fia-datamart>
(ASTOLA *et al.*, 2021)	Sentinel-2	<https://sentiwiki.copernicus.eu/web/s2-applications>
(SANCHEZ-RUIZ *et al.*, 2019)	NFI3 (SNFI-3)	<https://www.miteco.gob.es/es/biodiversidad/servicios/banco-datos-naturaleza/informacion-disponible/ifn3.html>
	Landsat-5 TM and Landsat-7 ETM	<https://code.earthengine.google.com/>
	SIOSE	<https://www.siose.es/web/guest/descargar>
(TIJERÍN-TRIVIÑO *et al.*, 2022)	PNOA-LIDAR	<https://centrodedescargas.cnig.es/CentroDescargas/buscadorCatalogo.do?coT1amilia=02211#>
	SNFI-4	<https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/cuarto_inventario.html>
(BEAUDOIN *et al.*, 2022)	MVI Project[1]	<https://www.mdpi.com/2072-4292/14/5/1108>
	PALSAR	<https://www.eorc.jaxa.jp/ALOS/en/dataset/fnf/fnf_palsar20140116_e.htm>
	Landsat TM	<https://earthenginepartners.appspot.com/science-2013-global-forest>
(MEDINA; PAFFENROTH, 2021)	USGS Explorer	<https://earthexplorer.usgs.gov/>
(CHIRICI *et al.*, 2020)	Italian NFI	<https://www.inventarioforestale.org/it/>
	Landsat 5 TM	<https://code.earthengine.google.com/>
(SHOOT *et al.*, 2021)	FIA data	<https://apps.fs.usda.gov/fia/datamart/datamart_sqlite.html>
	G-LiHT Data	<https://glihtdata.gsfc.nasa.gov>
(NOVO-FERNÁNDEZ *et al.*, 2019)	SNFI-4	<https://www.miteco.gob.es/es/biodiversidad/temas/inventarios-nacionales/inventario-forestal-nacional/cuarto_inventario.html>
	CNIG	<http://centrodedescargas.cnig.es/CentroDescargas/buscadorCatalogo.do?codFamilia=LIDAR#>
(EHLERS *et al.*, 2022)	FIA data	<https://www.fs.usda.gov/research/programs/fia>
	NCSDD	<https://sdd.nc.gov/>
	Sentinel-1	<https://sentiwiki.copernicus.eu/web/sentinel-1>
	Landsat 8	<https://earthexplorer.usgs.gov/>
(XIE *et al.*, 2020a)	Landsat-8 OLI, Sentinel-1A	<https://code.earthengine.google.com/>
	SRTM	<http://srtm.csi.cgiar.org/srtmdata/>
	FROM-GLC30	<http://data.ess.tsinghua.edu.cn/fromglc2017v1.html>
(MAYRA *et al.*, 2021)	Personal Dataset[1]	<https://github.com/jaeeolma/tree-detection-evo>
(AYREY; HAYES, 2018)	G-LIHT Data	<https://glihtdata.gsfc.nasa.gov/> <https://github.com/Eayrey/3D-Convolutional-Neural-Networks-with-LiDAR>

[1] Available on request.

Among the evaluated studies, only Mayra *et al.* (2021) and Ayrey and Hayes (2018) provided the inventory data collected in the field, LiDAR data, and source code, which are available on GitHub.

2.5 Discussion

The primary studies evaluated in this SLR, which explored the intersection between AI and LiDAR in the context of forest inventories, provide significant insights into emerging trends in this area.

Our findings showed that the integration of AI in the forestry context is predominantly aimed at building models capable of estimating various forestry variables. Many

of the activities carried out in forests still depend on manual methods involving instruments and fieldwork. Although LiDAR is already widely used in several sectors of the forestry industry, it is essential to develop techniques and algorithms to analyze this data and extract relevant information for managing native or planted forests.

Traditional machine learning algorithms such as RF, SVM, DT, and others discussed in this study are widely employed compared to deep learning methods. Machine learning algorithms comprise the vast majority of techniques used, with RF alone representing 74% (34 papers) of the works evaluated, while deep learning algorithms represent 19% (9 papers).

However, our evidence indicates that deep learning techniques outperform machine learning algorithms due to the ability to automatically extract features from raw data, with features at higher levels of the hierarchy formed by the composition of lower-level features (LECUN *et al.*, 2015). Deep learning can solve more complex problems efficiently and quickly due to the more complex models used, which allows for massive parallelization (WEISS *et al.*, 2016).

A disadvantage of deep learning is that it generally takes longer to train. Other disadvantages include problems that can arise when using pre-trained models on small or significantly different datasets, optimization issues due to the complexity of the models, and hardware constraints.

Despite these limitations, deep learning has proven to be superior due to the shorter time required to estimate variables and better metrics, as evidenced in previous studies (PARK *et al.*, 2021; OEHMCKE *et al.*, 2022; ASTOLA *et al.*, 2021; CHEN *et al.*, 2020a). Among deep learning architectures, CNNs represented 10% (5 articles) of the studies evaluated. However, the results obtained in these studies were significantly better compared to machine learning algorithms, considering the metrics used (WINDRIM; BRYSON, 2020a; ALON *et al.*, 2019; MAYRA *et al.*, 2021; PERSSON *et al.*, 2022; AYREY; HAYES, 2018).

Although deep learning techniques do not constitute the majority, the results achieved by these algorithms have demonstrated efficiency and broad applicability. Considering the observed performances, it is possible to consider this class of algorithms as a trend in future studies involving the integration of AI and LiDAR.

Section 2.4.1 discusses the main forest variables evaluated in this RSL. The four most studied variables were volume, representing 50% (23 papers) of the studies, height with 30% (14 papers), diameter with 26% (12 papers), and biomass with 21% (10 papers). These forest variables are continuous and frequently estimated due to their significant economic value in planted forests and their crucial environmental value in native forests. The regression approach, corresponding to 74% (35 papers), was the most used to estimate continuous variables, and RMSE and R^2 were the metrics most used in this context.

Among the variables observed, the diameter corresponds to 26% of the studies. The diameter at breast height (DBH) is the essential element, as its measurement serves as a basis for other calculations, such as, for example, obtaining the sectional area (degree of occupation by wood in a given area) at chest height, an essential measure in calculating the volume of trees. Trees and stands. In the studies covered, DBH is a measurement observed indirectly regarding aerial scans (95% of studies). This occurs due to this type of scanning, as the density of the tree canopy makes it difficult for the pulses used by LiDAR to be emitted and returned.

Consequently, the diameters have a high margin of error compared to those measured in the field with other methods. In Heinzel and Ginzler (2018), terrestrial scanners showed superior results, with a 2 to 8% MAE. In the work of Nguyen *et al.* (2021), although the diameters of regions of defective branches are measured, DBH measurements are not estimated.

Our analysis revealed a need to develop automated DBH identification and quantification methods by applying AI algorithms to the point cloud in LiDAR datasets. The scarcity of research addressing this metric with a sufficiently minimal margin of error is particularly noteworthy, considering the pivotal role DBH plays in forestry assessments.

Tree species classification and tree counting represent 15% (7 papers) and 13% (6 papers) of the studies covered. In native forest scenarios, species classification is widely used to identify and categorize the main species of a given area, as discussed in previous studies (PARK *et al.*, 2021; NGUYEN *et al.*, 2021; TIJERÍN-TRIVIÑO *et al.*, 2022; MEDINA; PAFFENROTH, 2021; CASTILLA *et al.*, 2022; SHOOT *et al.*, 2021; MAYRA *et al.*, 2021).

Counting and classifying tree species using LiDAR are more complex tasks that require segmentation and object detection. Deep learning approaches, especially CNNs, have proven very effective, as discussed by Windrim and Bryson (2020a), Mayra *et al.* (2021).

ALS approaches were widely used (44 articles, 95%) due to the practicality of covering large areas with just one pass over the region. The number of variables obtained and the ease of exploring these LiDAR data allow us to derive several important information for the forest inventory. On the other hand, TLS scans were covered on a smaller scale (2 articles, 4%), as they depend on professionals in the field manipulating LiDAR scanners to collect data.

TLS requires more time and makes obtaining these variables more costly. Despite having denser and more accurate data, such as defect identification (for example, (NGUYEN *et al.*, 2021), with branch diameter errors more minor than 5 mm) and more accurate DBH analysis, the cost and collection time showed the main factors that limit adherence to this technique are identified.

Considering that 48% (22 articles) of the studies were carried out in planted forests, that is, forests planted by humans for commercial use, the LiDAR data obtained are private (33 articles, 72%) and are not publicly available for reproducibility. The public data (13 articles, 28%) available for access come from scans carried out in native forests (24 articles, 52%). However, some data from native forests were inaccessible and considered private. More details about which studies have available data are presented in 2.2.

We observed that research on AI techniques applied to forest inventory has produced excellent results in proposing new methods and applying established techniques in the field. This area of study is extensive and offers many opportunities for researchers interested in improving the automation of forest inventory processes, increasing productivity, and assisting forestry engineers and technicians in the responsible management of forest resources.

2.6 Research Scope and Limitations

While our study was conducted following the guidelines established by Keele (2007), our analysis has certain limitations, primarily related to our research methodology and data extraction process.

The present study faces potential threats to the external validity of the systematic review due to limitations in the search strategy employed. While multiple pilot searches and a review of prior studies informed the construction of the search string, it is possible that certain relevant studies may have been excluded due to the omission of key search terms or combinations. Additionally, the scope of the study was limited to four databases (SCOPUS, EI Compendex, Web of Science, and IEEE Xplore), which may have resulted in the omission of relevant studies indexed in other databases.

Despite this, primary study data was extracted regarding the type of contribution such as algorithms, AI techniques used, forest parameters and LiDAR scan; and comparison of results. To ensure even more accuracy in the extracted data, a protocol was developed to define the strategy and format of data extraction on the *Parsif.al* tool. During extraction, the consistency of relevant information was verified, and the extracted data met the research questions. A forestry science expert resolved any disagreements or discrepancies among the authors during the process.

2.7 Conclusions

This chapter presents a SLR on the application of artificial intelligence techniques to forest data obtained through LiDAR scanners, aiming to identify, evaluate, and

interpret primary studies at the intersection of AI and forestry engineering. The findings reveal that AI and LiDAR not only automate manual tasks but also outperform traditional methods, with deep learning emerging as a prominent trend for classification and prediction of forest parameters. However, despite the importance of tree diameter for volume estimation, it is often treated as a secondary measure, indicating a promising research gap. This review provides researchers with valuable insights into potential applications and considerations for implementing AI in forestry, highlighting opportunities for future studies that explore additional databases, parameters, and advanced AI techniques. Furthermore, the results reinforce the relevance of combining cutting-edge computational approaches with forestry practices to support sustainable management strategies. By outlining both achievements and open challenges, this SLR contributes to guiding future innovations in the field.

CHAPTER 3

Eucalyptus Diameter and Volume Prediction with Deep Neural Networks: A Long Short-Term Memory Model Approach

> Commit to the Lord whatever you do,
> and your plans will succeed.
>
> *Proverbs 16:3*

Effective management of forest resources is vital for forestry companies. Accurate information is essential for planning planted forests. Therefore, forest inventory is a necessary procedure for obtaining quantitative and qualitative information about a region. Total volume serves as a critical metric for evaluating the potential of a specific area. This study presents an approach for modeling and forecasting diameters and total volume in planted forests. Diameter data at different tree heights in three different forest sites were used, and five neural network models were developed to estimate diameters and calculate volumes of eucalyptus clones. The models underwent training and testing using cross-validation techniques and were compared to statistical models. The results showed that the models based on Long Short-Term Memory (LSTM) are superiors. The LSTM models generalized tree volumes from diverse forest sites with greater accuracy than statistical models, even without site-specific training data. The percentage differences between the cubed volumes at sites (I, II, and III) and the volumes obtained by LSTM were smaller (0.01%, 1.06%, and 0.42%) in comparison to statistical models (0.35%, 1.81%, and 0.17%). These results highlight the potential of LSTM networks to provide accurate, generalizable, and non-invasive solutions for forest volume estimation, offering significant benefits for forestry management and decision-making. The contents of this chapter were published in Expert Systems with Applications (RODRIGUES *et al.*, 2025).

3.1 Introduction

The Brazilian forestry industry is globally renowned for its high productivity, characterized by the shortest turnover time between planting and harvesting. In 2021, the total area of planted trees in Brazil increased by 8.3% compared to 2020, driven solely by the expansion of eucalyptus plantations. Data indicates that eucalyptus plantations now occupy 7.53 million hectares of the country's planted area (IBÁ, 2022). This expansion is primarily attributed to industries investing in eucalyptus as a raw material for their production processes.

In 2021, the Brazilian forestry sector experienced a growth of 7.5% compared to the previous year, exceeding the national Gross Domestic Product (GDP) growth and achieving a record gross revenue of R$244.6 billion, equivalent to 1.2% of the GDP.

Incorrect manual measurements, such as diameter, can compromise the accuracy of total tree volume estimates during forest inventory surveys. Couto and Bastos (1987) demonstrated that a one-meter error in measuring the tree's total height corresponds to approximately a 14% error in the cylindrical volume of the trees. Similarly, a 1 cm error in DBH, measured at a height of 1.30 meters above the ground, results in a 19% error in the total volume of the tree.

The total volume of planted wood in a given area is crucial for effective forest resource management. Therefore, achieving the highest possible precision in its quantification is of paramount importance. While there are methods to estimate the volume of standing trees, these approaches are generally less accurate than those obtained from felled trees (rigorous measurement).

According to Gomes (1957), rigorous measurement measures successive diameters along the tree stem. The author mentions that diametric measurements were carried out every 0.30, 1.30, 3.30, 5.30, and so on, every 2 meters; however, the distances between diametric measurements can be changed depending on the shape and rigor of the measurement process. However, the distances between these diametric measurements can be adjusted according to the shape and precision of the measurement process. Generally, this approach involves cutting down the sampled trees and measuring various diameters along the stem, which is laborious and differs from the routine measurements of DBH and the heights of the trees that make up the plot inventory.

Another disadvantage of rigorous measurement is the need to fall trees during the process of this activity in Brazil. In addition, it is a time-consuming process that requires immediate commercialization of the felled trees to avoid product loss (CABACINHA, 2003).

The stem of a tree resembles geometric solids of revolution. The base of the tree is similar to a hyperboloid, the intermediate part is similar to a paraboloid, and the end is

similar to a cone (GOMES, 1957; SCOLFORO *et al.*, 1997).

Due to the high variability in shape among individual trees, the relationship between diameters along the stem and taper may exhibit linear or nonlinear behavior. Therefore, algorithms such as neural networks have been employed in various applications, including diameter prediction, volume calculation, and other forest parameters.

Neural networks (NNs) have garnered significant attention and demonstrated remarkable achievements across various domains, including communication security, image processing, and associative memory (SONG *et al.*, 2024).

Studies in the literature have shown the efficiency of artificial NNs in modeling the intricate relationships between tree measurements, increasing precision and efficiency in inventory and forest management.

This work presents five neural network models for estimating diameters and calculating the volume of eucalyptus clones. The models Vanilla Long-Short Term Memory (LSTM-VA), Long-Short Term Memory with Data Augmentation (LSTM-DA), and Multilayer Perceptron for Recursive Prediction (MLP-PR) use recursive prediction to estimate diameters and perform volume calculation. On the other hand, the Multilayer Perceptron for Direct Prediction (MLP-PD) and Multilayer Perceptron with DBH and Total Height (MLP-DAPHT) models use direct prediction, considering only the DBH and total height to estimate tree volumes without the need to know the diameters previously.

This study introduces LSTM neural network models developed to generalize tree volumes effectively across diverse forest sites, achieving satisfactory results without incorporating data from these specific sites during the training phase. Forest sites represent distinct classes of productivity capacity, significantly influencing tree shape and, consequently, their volume. The hypothesis under examination is whether LSTM networks can accurately generalize tree volumes under these conditions. Therefore, models trained on one site can be applied to others without requiring tree felling for direct measurement

This work is divided as follows. Section 3.2 presents the literature review. Section 3.3 discusses the materials and methods used. Section 3.4 compares the results with traditional methods and the models' generalizability. Finally, Section 3.5 presents the final considerations.

3.2 Literature Review

Numerous scientific studies in the forestry sector aim to utilize neural networks to acquire estimates of critical parameters like height, volume, and diameter at various locations on the tree stem.

In the study by Xiao *et al.* (1998), multi-layer neural networks were utilized to estimate tree height using high-frequency data from interferometry instruments. Leduc

et al. (2001) employed neural networks to estimate the diameters of *Pinus* trees. Diamantopoulou (2005) proposed a neural network architecture for estimating diameters and calculating the volume of *Pinus* trees. These studies demonstrate that neural networks represent a promising alternative to traditional regression methods. Additionally, Huang *et al.* (2009) utilized a neural network to predict the diameter distribution of *Pinus* trees, while Silva *et al.* (2009) employed neural networks to adapt the Schumacher and Hall model for estimating the volume of eucalyptus.

Niska *et al.* (2010) conducted a study in which neural networks were employed for species prediction and volume calculation of trees from three different species using laser scans and aerial images. Özçelik *et al.* (2010) utilized neural networks to estimate the volume of tree stems belonging to four species. Diamantopoulou (2010) addressed the challenge of filling gaps in diameter measurements using data from standing trees.

Guo *et al.* (2011) demonstrated the effectiveness of neural networks in solving complex nonlinear problems, particularly in modeling diameter distributions. Soares *et al.* (2012) investigated recursive prediction methods for estimating diameters of eucalyptus clones, considering scenarios with and without prior knowledge of tree height. Silva *et al.* (2009) estimated the volumes of eucalyptus trees using various measurements, including DBH, tree height, and diameters at specific heights above ground level. Volumes were calculated for different height thresholds, such as up to 2m and 4m. Vieira *et al.* (2018) utilized neural network and fuzzy logic techniques to forecast the growth of eucalyptus forests, predicting diameter and height based on DBH measurements.

Azevedo *et al.* (2020) evaluate strategies for modeling multi-volumes of eucalyptus tree stems, emphasizing their importance for forest management and economic decision-making. Results indicate that while volume-specific models excel in accuracy, multi-volume models offer practical advantages. NNs outperformed regression models in estimating multi-volumes, demonstrating superior accuracy and practicality.

Ercanlı (2020) introduce deep learning as a novel application for predicting the relationship between individual tree height and diameter at breast height. Equivalence tests further validated the effectiveness of deep learning in independent datasets, highlighting its potential for forest management applications.

Diamantopoulou *et al.* (2023) investigated cost-efficient alternatives to traditional tree height measurement methods in mixed oak stands. They assessed six modeling techniques, including General Regression Neural Networks (GRNN) and Support Vector Regression (SVR), highlighting SVR as a reliable and cost-effective tool for forest management and sustainable resource planning.

Rocha *et al.* (2023) investigated machine learning models to estimate tree volume and biomass in the Southwestern Amazon, finding that NN delivered the best performance. The study highlights NN's potential to support sustainable forest management and

emphasizes artificial intelligence's role in addressing deforestation and land-use changes.

Unlike previous studies that primarily focused on applying neural networks to specific datasets or forest conditions, our work advances the field by incorporating a deep learning approach to analyze and compare the performance of multiple neural network models. A significant contribution of this study lies in testing these models across different forest sites, which represent distinct productivity classes, to evaluate their generalization capacity.

3.2.1 Novelty and Main Contributions

Traditionally, when performing a forest inventory, volume estimates such as Schumacher and Hall Log, or diameter estimators such as taper functions, are developed by the forester. Building these volume estimation models requires cutting trees from different DBH classes to prepare a dataset for modeling. Similarly, constructing a taper function requires cutting a larger number of trees.

It is important to highlight that cutting trees requires effort and cost and, once cut, the tree must be sold, as the wood degradation process begins. Furthermore, these models are directly related to the location and population where the trees were harvested and are based on geometric model assumptions, which cannot represent the majority of trees in the population.

In this work, we focus on the use of Neural Networks to estimate diameters and volumes without geometric dependence. However, the greatest benefit is provigind a more general model, using models trained to estimate the diameters and volumes of other locations, avoiding cutting down trees or creating new models.

Although we have not accounted for it, it is possible to see that there is a positive monetary impact, cost and effort reduction, once an accurate model is built. Therefore, this work is providing a reusable model for different locations and populations, contributing to greater gains for the forestry engineer in several aspects.

3.3 Materials and Methods

3.3.1 Recurrent Model for Diameter Prediction

Diameters can be viewed as a sequential data (SOARES *et al.*, 2013), thus sequential data refer to data ordered by time, where the sequence and timing of data points are crucial for understanding patterns and making predictions. In predicting eucalyptus diameters, the data can be considered a sequential data because it involves measurements taken at various heights along the tree's stem. Each diameter measurement corresponds

to a specific point on the tree's measurement, reflecting the tree's dimensional changes along its stem.

In this regard, direct and recursive prediction is among the models commonly used in time series forecasting (JI *et al.*, 2005).

The direct prediction approach constructs $M+1$ different models to predict values of a series.

$$\hat{d}_{i+m} = f_m(d_{i-1}, d_{i-2}, ..., d_{i-n}), \tag{3-1}$$

Let $m = 0, 1, ...M$, where M is defined as the maximum size of the prediction horizon. The variable d_i represents the diameters at height i, which form the regressor, where n is the size of the regressor.

Recursive prediction builds the model by initially performing one step forward.

$$\hat{d}_i = f(d_{i-1}, d_{i-2}, ..., d_{i-n+1}), \tag{3-2}$$

Unlike direct prediction, in recursive prediction, the next value to be predicted takes as input the result of itself in the previous step, as can be observed in Equation 3-3.

$$\hat{d}_{i+1} = f(\hat{d}_i, d_{i-1}, d_{i-2}, ..., d_{i-n}) \tag{3-3}$$

Direct prediction requires a specific model for each step and necessitates the correction of inputs. Errors occurring in one step cannot influence the estimation of subsequent steps. Conversely, recursive prediction estimates values using a single model, where inputs are replaced at each step. As a result, initial values can be handled differently, eliminating the need to establish a prediction horizon and avoiding biased model creation.

Soares *et al.* (2013) demonstrated that diameters at heights $d_{0.3}$, $d_{0.7}$, and $d_{1.3}$ exhibit strong autocorrelation. Models that use the recursive prediction method receive these same diameters as inputs and return the next diameter $\hat{d}$ at height $d_{2.0}$. As the prediction is recursive, the next inputs for the model will be d_{i-1}, d_{i-2}, and $\hat{d}$, and so on until the tree height is reached, generating a vector $\hat{p}$ with all estimated diameters.

3.3.2 Dataset

The total data used in this study consists of 1,819 trees from 48 plots that were felled and rigorously measured using the Smalian method. Rigorous measurement is the process of measuring successive diameters along the tree stem. These trees are from three forest sites owned by a private company, and the data was made available by Cabacinha (2003).

During the measurement process, diameters were measured at heights of 0.10, 0.30, 0.50, 0.70, 0.90, 1.10, 1.30, and 2.00*m* relative to the ground, and in the remaining stem at intervals of 1 meter until the height of the commercial diameter of 4 cm was reached. The measurement process can be observed in Figure 3.1, where a caliper was used to take the measurements.

Table 3.1: *Distribution of trees.*

Site	# Trees
I	585
II	619
III	615
Total	1,819

Site I has the highest production capacity, Site II has intermediate production capacity, and Site III has the lowest production capacity. Table 3.1 shows the distribution of trees in each site.

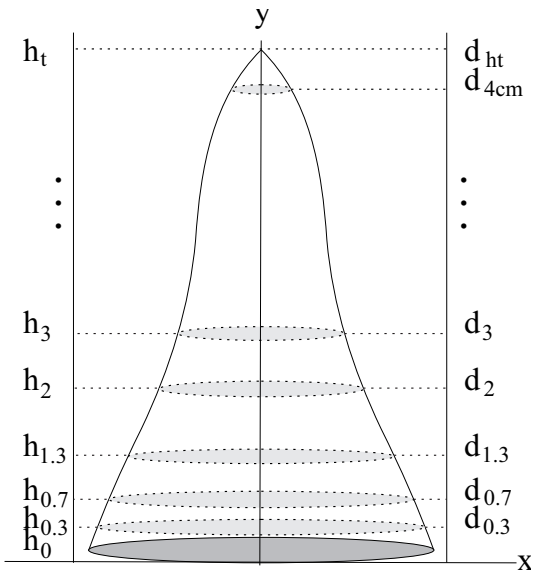


Figure 3.1: *Height and Diameter measuring approach.*

To construct and train neural network models, we selected 615 trees from Site III. This subset was chosen due to its lower tree growth capacity than other sites. The remaining sites will be utilized to validate the trained models.

Although the eucalyptus clones are from the same site, studies have shown that tree characteristics vary according to the DBH. Schneider *et al.* (1996) and Machado Pires and Calegario (2007) have previously reported this finding.

For a better understanding of the database used in this study, Table 9.1 presents a statistical summary.

Table 3.2: *Dataset statistics.*

Class	Limits		Qty.	Autocorr(d_i)		DBH					Height				
	Inf.	sup.		Min	Max	Average	Var[1]	DP[2]	Min	Max	Average	Var[1]	DP[2]	Min	Max
1	7.15	10.15	14	0.63	0.84	8.68	0.95	0.97	7.15	10.10	17.60	3.49	1.87	14.20	19.70
2	10.15	13.15	106	0.77	0.87	12.01	0.54	0.73	10.32	13.14	22.83	2.75	1.66	18.00	26.70
3	13.15	16.15	302	0.78	0.89	14.62	0.68	0.82	13.15	16.14	25.74	1.57	1.25	18.40	29.40
4	16.15	19.15	178	0.83	0.90	17.23	0.64	0.80	16.15	19.14	27.52	1.39	1.18	24.20	30.10
5	19.15	22.15	15	0.82	0.90	20.21	1.70	1.31	19.31	24.55	29.23	1.06	1.03	27.20	30.60

[1] Variance. [2] Standard Deviation.

Linear Pearson correlations were computed to examine the relationship between tree diameters and their neighboring diameters. Each tree was analyzed with displacements ranging from 1 to 3. These intervals were selected based on this study's observed strong correlations between tree diameters.

As depicted in Table 3.2, correlations for a 1-interval ranged from 0.63 to 0.90, while for a 3-interval, correlations ranged from 0.55 to 0.81.

Our statistical analysis has provided a clear picture. It indicates strong correlations among the first three measurements of randomly selected trees, ranging from 0.78 to 0.88. As the distance between diameters increases, correlations gradually decrease but remain above 0.68. This gradual decrease in correlations is a key aspect of our findings, illustrating the progressive nature of our analysis. Figure 3.2 further illustrates this relationship among the first 10 diameters of a representative tree.

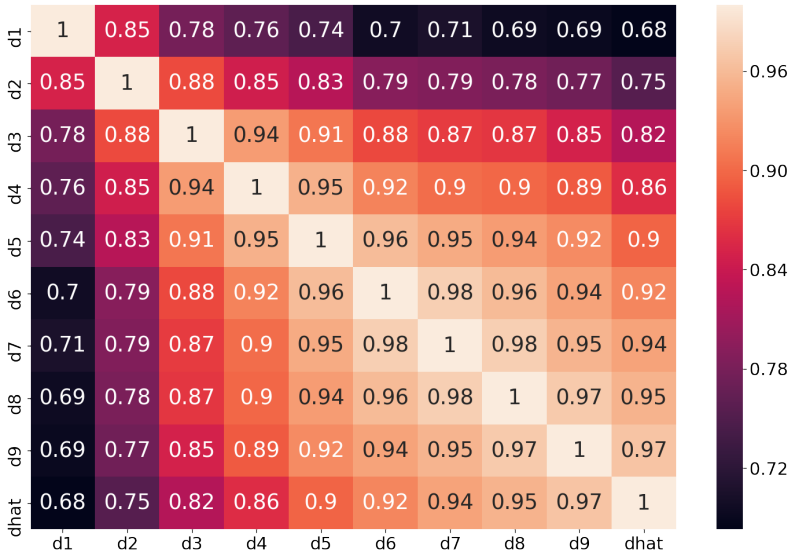


Figure 3.2: *Relationship between the first ten diameters of a tree.*

3.3.3 Normalization

The data was normalized in the range of [-1,1] according to Equation 3-4.

$$X' = a + \frac{(X - X_{min})(b - a)}{X_{max} - X_{min}}, \tag{3-4}$$

where X' is the normalized diameter value, a and b represent the normalization range, X_{min} and X_{max} represent the value of the smallest and largest diameter, respectively, of the dataset.

3.3.4 *Data Augmentation*

Data augmentation is a widely used technique for increasing the amount of data and helping with generalization capacity, reducing overfitting, and improving performance in problems with unbalanced data (SIMARD *et al.*, 2003; CIREŞAN *et al.*, 2010).

Data augmentation techniques play a crucial role in expanding the variability of training data by introducing synthetic variations or transformations into existing samples. Standard data augmentation methods include rotating, flipping, scaling, cropping, and adding noise to images. In the context of eucalyptus diameter and volume estimation, data augmentation was employed as a complementary strategy to address the inherent challenges of dataset imbalance and limited sample size. By generating synthetic variations in the input data, this technique enhances the model's exposure to a wider range of scenarios, thereby improving generalization. Unlike traditional feature engineering, data augmentation does not rely on predefined assumptions, ensuring flexibility and scalability in the predictive framework. In this work, new diameter values were generated for training the LSTM-DA model. The value of the new diameter was obtained through Equation 3-5,

$$d_{i_{new}} = e \cdot \frac{d_i}{c}, \tag{3-5}$$

where $d_{i_{new}}$ represents the new synthetic diameter generated, d_i is the value of the diameter read at the i-th position, e is the value of the Euler's constant, and the constant $c = 2.5$ was empirically determined through a series of experiments, aiming to calibrate the formula and adjust the model to the observed conditions, ensuring greater consistency and accuracy in the results. By applying this equation, the model generates synthetic diameter values that reflect plausible growth scenarios. This is particularly important in the context of limited data, as it allows the model to "see" a wider range of potential tree sizes and growth patterns, which can improve its generalization capabilities.

The process was repeated four times. The new diameter values generated using the data augmentation technique were added to the original dataset.

3.3.5 Modeling and Training of Neural Networks

The sections of a tree exhibit nonlinear relationships. Given neural networks' capability to handle nonlinear problems and their application in various forestry studies, five distinct neural network models were developed.

The methodology prioritizes a balance between accuracy and practicality by leveraging LSTMs networks and recursive prediction frameworks. While LSTMs are not inherently simple, they provide a structured approach to handling sequential data, such as tree diameters measured at various heights, without requiring extensive feature engineering. This makes them a pragmatic choice for forestry applications where domain-specific features may not always generalize across diverse forest sites.

The first model, LSTM-VA, employs a primary LSTM recurrent neural network without modifications. The LSTM-DA model is similar but incorporates Data Augmentation during training to enhance generalization. The MLP-PR model, on the other hand, performs diameter estimation using principles akin to LSTMs but without the complexities of recurrent neural network architectures. It utilizes feedforward neural network principles to predict subsequent diameters based on input diameters at different tree heights. Unlike LSTMs, MLP-PR focuses on sequential prediction tasks without the memory and feedback mechanisms inherent in LSTM models.

The described models receive as input the diameters at heights $d_{0.3}$, $d_{0.7}$, and $d_{1.3}$ meters and return as output the next diameter to be predicted every $1m$ until the height of the tree is reached.

The models employing direct prediction methodology, where tree volume calculation occurs directly from network inputs, were constructed using the classical feedforward architecture of neural networks. These models include the Multilayer Perceptron for Direct Prediction (MLP-PD), which takes diameters at heights $d_{0.3}$, $d_{0.7}$, and $d_{1.3}$) as input and outputs the total volume of the tree. Additionally, the Multilayer Perceptron with DBH and Total Height (MLP-DAPHT) receives the DBH and the total height of the tree as input, predicting the tree's total volume.

The topology of the models used in this work is presented in Table 3.3.

Table 3.3: *Structure of the built models.*

Model	N[1]. Layer Entry	N[1]. Layer Hidden	N[1]. Layer Output	Function of Activation	Variables input	Variables output
LSTM-VA	3	16	1	ReLU	$d_{0.3}$, $d_{0.7}$, and $d_{1.3}$	$\hat{d}$
LSTM-DA	3	32	1	ReLU	$d_{0.3}$, $d_{0.7}$, and $d_{1.3}$	$\hat{d}$
MLP-PR	3	20	1	Tanh	$d_{0.3}$, $d_{0.7}$, and $d_{1.3}$	$\hat{d}$
MLP-PD	3	25	1	Tanh	$d_{0.3}$, $d_{0.7}$, and $d_{1.3}$	v
MLP-DAPHT	2	25	1	Tanh	DBH and Total Height	v

[1] Number of neurons in the layer.

A systematic approach to parameter tuning entails the fine-tuning of hyperparameters, where critical parameters such as learning rate, batch size, number of epochs, and network architecture are iteratively adjusted following each training cycle. This study used techniques such as grid search and random search to identify the ideal configuration that maximizes the model's effectiveness (TUNÇEL; DURAN, 2023). At the end of training each model, the best one was defined as the one that presented the lowest Mean Squared Error (MSE) value.

The training process involved splitting the training and testing sets by 70% and 30%, respectively. Since the neural network training was conducted using data from Site III, the validation set and generalization capacity of the models were tested on Sites I and II.

For models that perform recursive analysis, input samples $[d_{i-2}, d_{i-1}, d_i]$ and output $[d_{i+1}]$ were generated according to Equation 3-3. The diameters used as input for the models consists of $d_{0.3}$, $d_{0.7}$, and $d_{1.3}$, and the network's response is the estimated diameter at height $d_{2.0}$. The next input set will consist of diameters $d_{0.7}$, $d_{1.3}$, and $d_{2.0}$ to estimate the diameter at height $d_{3.0}$. This process continues until the total height of the tree is reached.

We employed the L-BFGS algorithm to optimize the MLPs models (MORITZ *et al.*, 2016). This approach does not require parameters such as learning rate and momentum.

The LSTM networks were trained with Stochastic Gradient Descent (SGD) and the Adaptive Moment Estimation (ADAM) optimization algorithm with a learning rate a of 0.0001.

During the experiments, data augmentation analyses were performed on all models. However, the results obtained were quite different from what was expected, except for the LSTM architecture, which presented results closer to reality. Therefore, only the LSTM-DA model was evaluated using the data augmentation process.

Overfitting is one of the biggest problems in neural network training. It occurs when the data is overfitted, hindering the model's generalization ability. In addition to varying the training parameters, the early stop technique was used to try to minimize this problem.

The coefficient of determination R^2 is one of the metrics used to evaluate regression models. It indicates how much of the independent variable can be explained by the model. The Mean Absolute Percentage Error (MAPE) shows, in percentage terms, the average absolute difference between actual and predicted values. Table 3.4 shows the values of R^2 and MAPE that the models achieved during the training phase.

During the training phase, all models were configured to run for up to 1,000 epochs until convergence or until the Mean Squared Error (MSE) value either increased or

Table 3.4: *Analysis of metrics R^2 and MAPE.*

Metric	Class	Training					Test				
		LSTM		MLP			LSTM		MLP		
		VA	DA	PR	PD	DAPHT	VA	DA	PR	PD	DAPHT
R^2	1	0.86	0.89	0.82	0.59	0.59	0.90	0.90	0.89	0.61	0.63
	2	0.97	0.98	0.94	0.74	0.87	0.97	0.98	0.95	0.66	0.84
	3	0.98	0.99	0.96	0.75	0.85	0.98	0.99	0.96	0.71	0.89
	4	0.98	0.98	0.96	0.63	0.80	0.98	0.97	0.96	0.53	0.68
	5	0.99	0.99	0.96	0.84	0.18	0.98	0.99	0.97	0.43	0.46
MAPE	1	8.03	7.95	8.40	0.79	0.73	5.20	7.66	7.79	0.75	0.59
	2	3.81	2.98	5.04	0.92	0.70	3.51	3.11	4.48	0.97	0.70
	3	2.71	2.44	4.08	1.49	1.01	2.74	2.41	4.17	1.19	0.77
	4	2.74	2.95	3.75	1.93	1.32	2.68	3.14	3.90	1.91	1.54
	5	2.95	2.981	3.96	2.28	4.47	3.32	2.91	4.14	2.35	2.66

showed no improvement for more than five consecutive epochs, triggering the activation of the early stopping mechanism and halting the process.

For instance, the MLP-PR model in Class 1 achieved convergence in just 35 epochs, demonstrating its efficiency. Conversely, the MLP-DP model required 1,000 epochs to achieve acceptable error levels. The remaining models converged, on average, around 250 epochs.

Figure 3.3 illustrates the evolution of MSE across epochs for the various models, including LSTM-VA, LSTM-DA, MLP-PR, MLP-PD, and MLP-DAPHT. The graph highlights differing convergence rates and identifies the models achieving the best MSE performance.

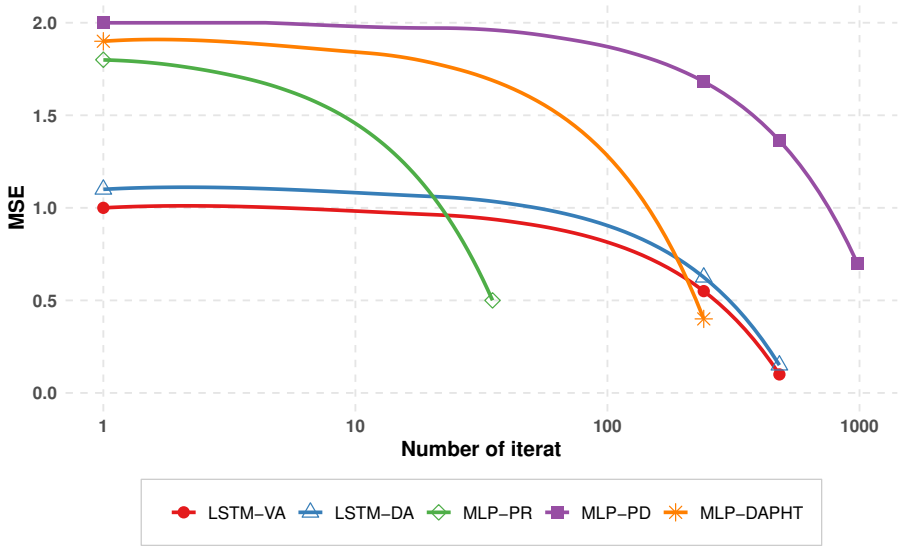


Figure 3.3: *Evolution of MSE over epochs for different deep learning models. The graph compares LSTM-VA, LSTM-DA, MLP-PR, MLP-PD, and MLP-DAPHT, highlighting the different convergence rates and the best MSE results.*

Figure 3.4 compares the training times of the models, offering insight into computational efficiency. The LSTM-VA model demonstrated a training time of 10.3

seconds, while MLP-PR completed training in just 6.8 seconds, underscoring its speed relative to other architectures. The computational resources utilized in this study are summarized in Table 3.5.

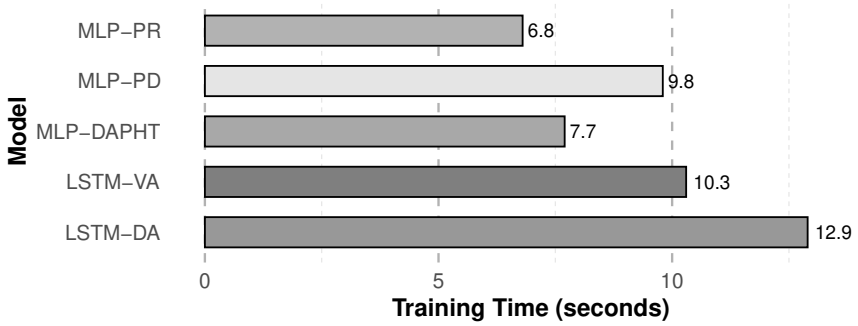


Figure 3.4: *Comparison of training time (in seconds) between different models.*

Table 3.5: *Computational resources and development environment used for model training and evaluation.*

Resources	**Description**
Processor	10th generation Intel i7 processor
Memory	8 GB RAM
Development Environment	
Python Version	Python 3.8
IDE	Spyder IDE, Jupyter Notebook
Libraries	Sklearn, Tensorflow

3.3.6 Methods

The methods used to analyze the model's performance consist of well-known and widely used metrics in statistical analyses.

Calculation of diameter errors

The metrics such as Root Mean Square Error ($RMSE_{\%}$) and linear correlation (r) were the metrics used to analyze the diameters estimated by the models.

The metrics used are represented by Equations 3-6 to 3-8, respectively.

$$RMSE = \sqrt{\frac{1}{n}\sum (y_i^a - y_i^e)^2} \tag{3-6}$$

$$RMSE_{\%} = \left(\frac{RMSE}{\bar{y}^a}\right) 100, \tag{3-7}$$

$$r = \frac{\sum_{i=1}^{n}(y_i^a - \bar{y}^a)(y_i^e - \bar{y}_i^e)}{\sqrt{\sum_{i=1}^{n}(y_i^a - \bar{y}^a)^2}\sqrt{\sum_{i=1}^{n}(y_i^e - \bar{y}^e)^2}}, \tag{3-8}$$

where n is the total of diameters estimated by the model, y^a is the vector of real diameters, y^e is the vector of estimated diameters, $\bar{y}^a$ and $\bar{y}^e$ are the average values of the real and estimated diameters, respectively.

Calculation of tree volume

The actual volume of all the trees cubed in the site and the volume calculated from the estimates made by the models were calculated using the Smalian formula, as shown in Equations 3-9 and 3-10

$$v = \sum^{n-1} i = 1\frac{gi + g_{i+1}}{2} + \frac{g_n l_n}{3}, \tag{3-9}$$

$$g_i = \frac{(d_i/2)^2\pi}{10.000}, \tag{3-10}$$

where v is the total volume of the tree, g_i is the basal area at the i-th position, l_i is the length of the section at the i-th position, g_n is the basal area of the cone of the tree, l_n is the length of the cone, and d_i is the diameter at the i-th position.

The respective volumes of the actual trees and the volumes estimated by the MLP were compared with the Schumacher and Hall (log) model. The Schumacher and Hall (log) model is a model that calculates the volume of trees based on the DBH and the total height (H_t), as described in Equation 3-11.

$$\ln V = \beta_0 + \beta_1 \ln DBH + \beta_2 \ln H_t + \varepsilon, \tag{3-11}$$

where v is the volume of the tree, DBH is the mean diameter at breast height, Ht is the total height of the tree, and the coefficients β_0, β_1, and β_2 are parameters to be adjusted.

The Schumacher and Hall (log) model utilized in this study is adjusted based on the total height of the tree and incorporates parameters β_0, β_1, and β_2 defined (CABACINHA, 2003).The models investigated in this research focus on estimating the tree volume up to a commercial diameter of 4 centimeters, rather than the entire tree volume. Therefore, this model serves solely as a benchmark for comparison with the models developed in this study.

Calculation of error estimation for volume

To analyze the accuracy of the volumes calculated by the models, the metrics of Mean Absolute Error ($MAE_{\%}$) Equation 3-12 (XIE *et al.*, 2020b), percent bias ($Bias_{\%}$)

Equation 3-13 (LEITE; ANDRADE, 2002), percent root mean square error ($RMSE_{\%}$) Equation 3-6, and linear correlation (r) Equation 3-8 (ÖZÇELIK *et al.*, 2010) were used.

$$MAE_{\%} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|y_i^a - y_i^e|}{y_i^a}\right)100, \tag{3-12}$$

$$Bias\% = \frac{\sum_{i=1}^{n} y_i^a - \sum_{i=1}^{n} y_i^e}{\sum_{n=1}^{i} y_i^a}100, \tag{3-13}$$

where y^a corresponds to the vector of cubed volumes, y^a is the vector of volumes calculated by the model, and n is the total number of reported volumes.

Calculation of percentage volume difference between models

The percentage differences between the total volume calculated by the Smalian method from the diameters obtained by rigorous scaling and the volume estimated by the models were calculated by Equation 3-14.

$$D_{\%} = \frac{|v_c - v_e|}{v_c}.100, \tag{3-14}$$

where $D_{\%}$ is the value of the percentage difference, v_c is the total cubed volume, and v_e is the total volume estimated by the model.

3.3.7 Statistical Tests

Given the multitude of experiments conducted, the results may appear closely aligned, prompting considerations about whether one model predominates over another or if the outcomes stem from random variability.

We employed Graybill's F-test at a significance level of 5% (GRAYBILL, 1976) to assess potential significant distinctions between the developed models and actual cubic values. This statistical method compares multiple distribution variants to ascertain substantial differences, typically among different groups or samples. The test computes the F statistic, representing the ratio of variance between groups to variance within groups. Graybill's F-test parallels Analysis of Variance (ANOVA) and applies to scenarios involving more than two groups, contingent upon data adhering to a normal distribution.

To verify our data's normality and the variances' homogeneity, we conducted hypothesis tests using the Lilliefors and Shapiro-Wilk tests. Results yielded p-values exceeding 0.05 at a significance level of $a = 0.05$, indicating that our data satisfied the assumptions of normality and homogeneity. Consequently, Graybill's F-test was deemed appropriate for analyzing the experimental outcomes.

All statistical analyses were executed utilizing the *forestmangr* library within RStudio.

3.4 Results and Discussion

3.4.1 Diameter Prediction

The LSTM-VA, LSTM-DA, and MLP-PR models recursively predict diameters until the tree height is reached. The metrics used to analyze the performance were the $RMSE_{\%}$ and the Linear Correlation (r).

Table 3.6: *Diameter Prediction Error by class.*

Class	Model	$RMSE_{\%}$				r	
		Min	Máx	Média	DP[1]	Min	Máx
1	**LSTM-VA**	3.36	7.97	5.44	1.53	0.97	0.99
	LSTM-DA	1.55	13.94	5.66	3.72	0.97	0.99
	MLP-PR	3.83	21.40	9.27	5.27	0.96	0.99
2	**LSTM-VA**	1.32	16.61	5.53	3.49	0.96	0.99
	LSTM-DA	2.67	19.14	6.36	2.71	0.95	0.99
	MLP-PR	2.17	18.76	6.91	4.22	0.96	0.99
3	**LSTM-VA**	1.24	15.29	5.37	2.57	0.97	0.99
	LSTM-DA	2.61	16.94	7.23	2.70	0.96	0.99
	MLP-PR	1.67	20.08	6.94	3.53	0.96	0.99
4	**LSTM-VA**	1.54	13.58	5.50	2.26	0.96	0.99
	LSTM-DA	2.13	18.29	6.62	3.41	0.97	0.99
	MLP-PR	1.51	16.84	6.58	3.16	0.98	0.99
5	**LSTM-VA**	2.64	7.55	4.86	1.57	0.97	0.99
	LSTM-DA	3.42	7.79	5.39	1.33	0.98	0.99
	MLP-PR	1.96	9.58	5.33	2.18	0.99	0.99

Table 3.6 shows the results of the error estimates of the predictions made by the models. Class 1 consists of dominated trees under the other classes and does not receive direct sunlight, hindering their development. They have the smallest dataset compared to the others.

For Class 1, the LSTM-VA model exhibited the lowest standard deviation among all models. Despite the limited training data, it reported an $RMSE_{\%}$ of 7.97, whereas the MLP-PR showed a higher error rate of 21.40%. The LSTM-DA model achieved an $RMSE_{\%}$ ranging from 1.55 to 13.94, demonstrating the benefits of data augmentation in mitigating the effects of sample scarcity on network performance. All models showed robust correlation coefficients ranging from 97% to 99%. Table 3.6 provides a detailed overview of the prediction errors estimated by the models. Class 1 comprises suppressed trees beneath other canopy layers, receiving limited direct sunlight, adversely affecting their growth. Consequently, it possesses the smallest dataset compared to different classes.

Classes 2, 3, and 4 consist of subordinate, intermediate, and co-dominant trees, respectively. Co-dominant trees receive direct light from both the canopy and lateral

sources, while subordinate trees grow between dominant and co-dominant ones. These classes encompass the most significant number of trees and exhibited strong performance during the training phase. Although these classes showed higher $RMSE_{\%}$ values, this can be attributed to the error propagation inherent in recursive prediction methods. Nevertheless, the models achieved favorable minimum $RMSE_{\%}$ values, low standard deviations, and robust correlations.

In contrast to Class 1, Class 5 comprises dominant trees that are taller and receive more sunlight, leading to more significant development and larger dimensions. Despite having fewer trees in the training dataset, Class 5 represents the tallest trees.

The tallest trees in Class 5 in forest inventories provide more diameter measurements, contributing more data during training. Consequently, the models estimate diameters more accurately due to the increased information available. The proposed models demonstrated promising results with minimal standard deviations, lowest maximum $RMSE_{\%}$ values, and maintained strong correlations ranging from 97% to 99%.

Figure 3.5 displays the actual and estimated diameters for each model in the best and worst cases of each class. The figure illustrates that better estimates are achieved for taller trees, whereas shorter trees tend to show poorer estimation accuracy. This discrepancy arises because taller trees provide more diameter measurements, offering richer training data for the neural networks. Notably, the LSTM-VA model displayed the best overall performance among all models, achieving an $RMSE_{\%}$ of 1.24. Conversely, the MLP-PR model exhibited the least favorable performance, recording an $RMSE_{\%}$ of 21.40.

Table 3.6 presents results where some RMSE values exceed 10%. However, it is important to emphasize that the average RMSE values across all classes remained below 10%. This indicates that, on average, all models delivered satisfactory results.

Regarding Figures 3.5(b) and 3.5(f), it is noticeable that there is overestimation in all diameters. However, this phenomenon was not consistently observed. However, this phenomenon was not consistently observed as a trend. While the highest RMSE values occurred when the models overestimated the actual diameters, during testing, the models alternated between underestimating and overestimating diameters, suggesting variability in diameter estimations. Therefore, there is no conclusive evidence indicating a systematic tendency of the models to overestimate the actual diameters consistently.

3.4.2 Comparison Between the Volume Calculated from the Actual Diameters, Diameters Estimated by the Models and the Volume Calculated Through Schumacher and Hall (log)

The actual volumes of the logged trees were calculated using the Smalian formula, Equation 3-9. The same method was employed to estimate the volumes from the estimated diameters for the same heights defined for the exact volume measurements for all the recursive models. The Schumacher and Hall (log) model was used as a benchmark, given that it is widely used in forestry.

According to Equation 3-11, the parameters used to calculate volume were adjusted by Cabacinha (2003). The equation adjusted to the data was: $\beta_0 = -10.79449927$, $\beta_1 = 1.942384069$, $\beta_2 = 1.226015699$, and $\varepsilon = 0$. With $R^2(\%) = 99.44$, residual standard error $S_{yx} = 0.011993$ m, and $S_{yx}\% = 5.81$.

After the diameters were estimated by the models in Table 3.6, the volume calculation can be estimated using the Smalian formula, except for MLP-PD and MLP-DAPHT models that perform direct estimation. The comparisons of real volumes, calculated by the models and by Schumacher and Hall (log), can be observed in Figure 3.6. The values presented correspond to the total estimated volume for each class, considering the sum of the training and testing sets.

Figure 3.6 shows that the total volume calculated using the Smalian method for diameters obtained through rigorous scaling, the total volume calculated using the Smalian method for diameters estimated by recursive models, the total volume directly calculated by the direct prediction models, and the volume estimated by Schumacher and Hall (log) are very close.

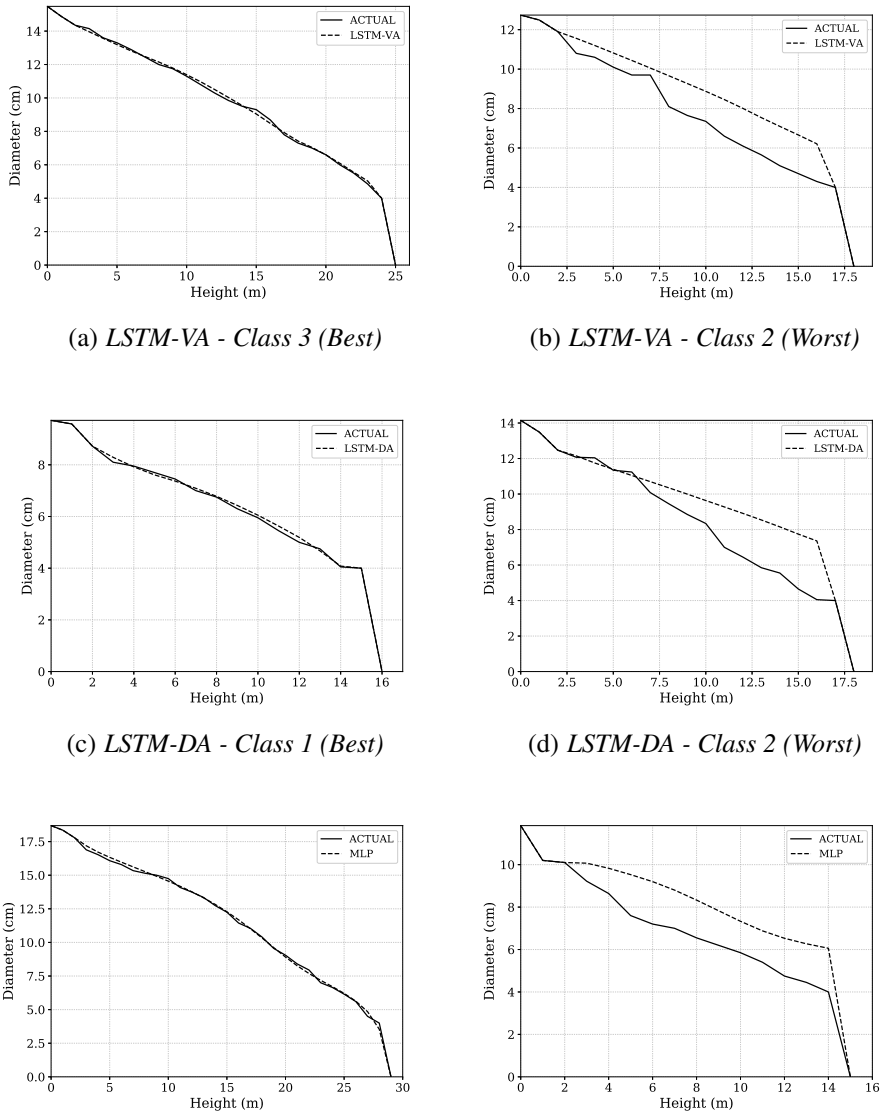
Analyzing Figure 3.6, variations in the percentage differences between the total volume obtained from rigorous volume measurement and the volumes calculated by the diameter estimation models are observed.

The LSTM-VA model presented a percentage difference of 0.42% in relation to the measured volumes, while the LSTM-DA and MLP-PR models showed 0.91% and 0.06%, respectively.

Among the models that directly estimate volume, MLP-PD showed a difference of 0.21%, while the MLP-DAPHT model had a difference of 0.13%. The Schumacher and Hall (log) model showed a variation of 0.17%.

Although the models differed, all presented percentage differences below 1%, with the LSTM-DA exhibiting the largest variation.

We found significant similarity by comparing the results of the models developed in this study to Soares' model Soares *et al.* (2011). Furthermore, utilizing advanced machine learning techniques and optimization enabled the achievement of results compa-



(a) *LSTM-VA - Class 3 (Best)*

(b) *LSTM-VA - Class 2 (Worst)*

(c) *LSTM-DA - Class 1 (Best)*

(d) *LSTM-DA - Class 2 (Worst)*

(e) *MLP-PR - Class 4 (Best)*

(f) *MLP-PR - Class 1 (Worst)*

Figure 3.5: *Prediction of MLP-PR diameters (best and worst case).*

rable to the best works in the literature. We calculated volume estimation errors for each model to determine the significance of variations and compared them with the results of (SOARES *et al.*, 2011). The outcomes are provided in Table 3.7.

The $MAE_{\%}$ among models employing the recursive prediction approach were higher than those using the direct prediction approach, especially for MLP-PR, which achieved a value of 11.54% in Class 1. However, the models demonstrated good results for the other larger classes.

The effect of data augmentation can be observed in the behavior of the LSTM-DA model, where the $Bias_{\%}$ showed the highest variations in larger classes. However, this effect was positive in smaller classes, with a lower tendency to overestimate the cubed values. The other models maintained a $Bias_{\%}$ below 1, which is positive since the

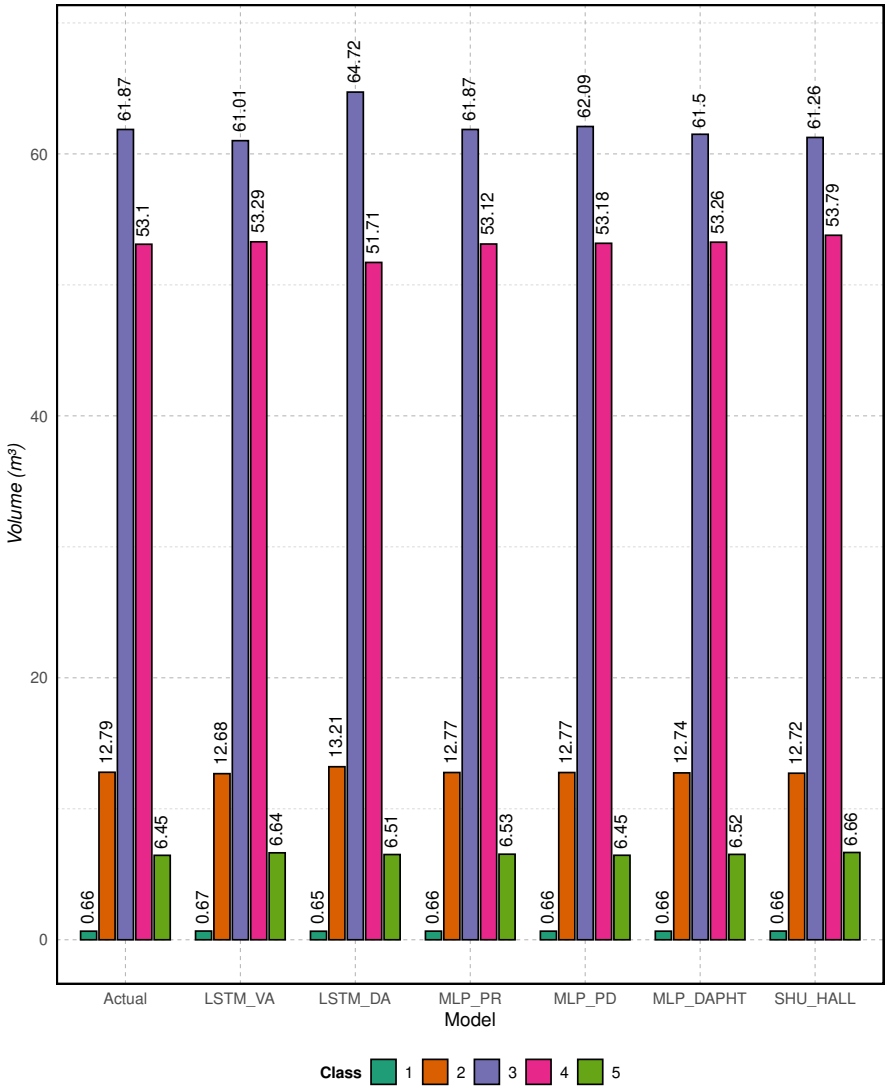


Figure 3.6: *Comparison between the total volume calculated (training + test) from the measured diameters, diameters estimated by the models and Schumacher and Hall (log) of Site III.*

Schumacher and Hall (log) model (appropriately parameterized) achieved 3.37 in Class 5.

The correlation coefficients for all studied models were extremely high, indicating a strong correlation between actual and estimated values. The $RMSE_{\%}$ values were very close in the direct prediction models, especially for MLP-DAPHT, which is very similar to the Schumacher and Hall (log) model and uses the same input variables. The MLP-PD model presented interesting results since it does not use the total tree height as a parameter.

However, as shown in Figure 3.6, the percentage difference between the actual total volume and the models was less than 1%. Nonetheless, although the volume was very close, it was observed that in some cases, the models attempt to compensate for

Table 3.7: *Statistics were conducted among the actual volumes, the trained models, and Schumacher and Hall's (log) method.*

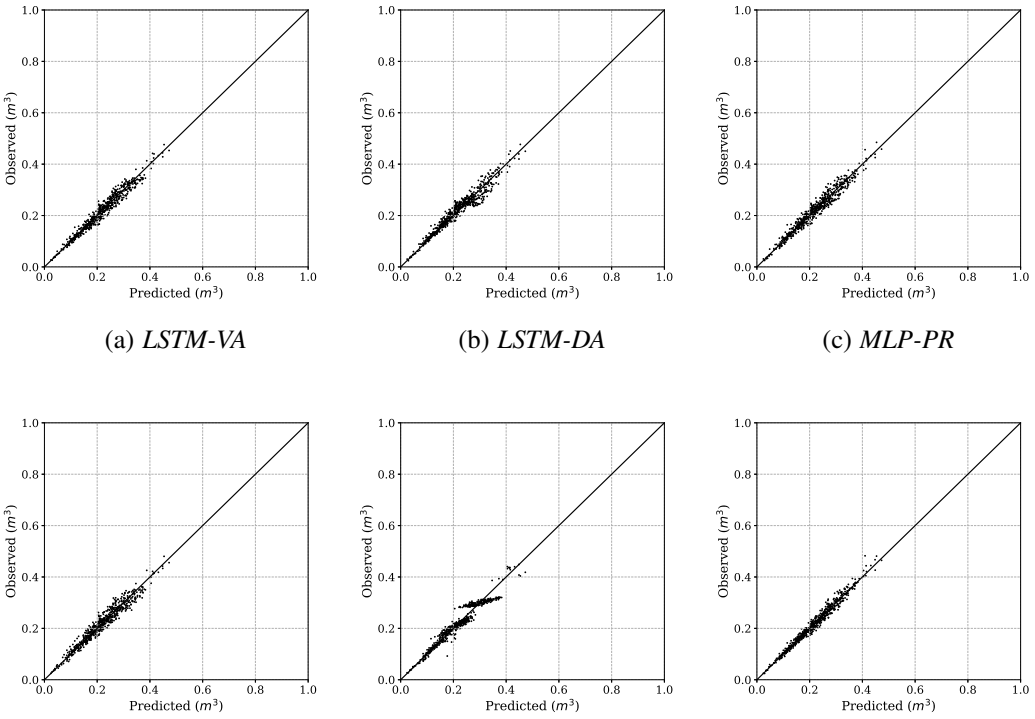
Class	Model	$MAE_{\%}$	$Bias_{\%}$	r	$RMSE_{\%}$
	LSTM-VA	6.57	1.52	0.95	9.05
	LSTM-DA	8.42	0.60	0.94	9.99
1	MLP-PR	11.54	0.44	0.88	15.32
	MLP-PD	8.57	0.16	0.92	11.39
	MLP-DAPHT	**5.08**	**0.05**	0.97	**6.29**
	(SOARES *et al.*, 2011)	10.40	0.07	**0.99**	9.88
	Sch. Hall (log)	6.16	1.15	0.96	8.29
	LSTM-VA	6.02	0.83	0.92	7.40
	LSTM-DA	6.14	3.29	0.94	6.14
2	MLP-PR	6.95	0.15	0.90	8.22
	MLP-PD	7.62	0.15	0.87	9.25
	MLP-DAPHT	5.08	**0.05**	0.97	6.29
	(SOARES *et al.*, 2011)	6.93	0.72	0.90	8.59
	Sch. Hall (log)	**4.75**	0.57	**0.95**	**5.53**
	LSTM-VA	6.01	1.54	0.89	7.57
	LSTM-DA	7.16	4.61	0.90	8.52
3	MLP-PR	6.49	**0.02**	0.87	7.94
	MLP-PD	6.77	0.37	0.86	8.20
	MLP-DAPHT	6.47	0.59	0.87	8.08
	(SOARES *et al.*, 2011)	6.90	0.76	0.86	8.78
	Sch. Hall (log)	**4.48**	0.98	**0.93**	**5.65**
	LSTM-VA	6.07	0.35	0.83	6.99
	LSTM-DA	6.86	2.62	0.81	6.86
4	MLP-PR	6.40	0.38	0.80	7.39
	MLP-PD	6.83	**0.13**	0.76	8.08
	MLP-DAPHT	8.03	0.29	0.87	9.51
	(SOARES *et al.*, 2011)	6.41	0.33	0.81	7.59
	Sch. Hall (log)	**4.25**	1.29	**0.91**	**5.22**
	LSTM-VA	5.07	2.99	0.94	7.70
	LSTM-DA	5.02	0.95	0.94	6.87
5	MLP-PR	4.78	1.34	0.94	6.45
	MLP-PD	**3.30**	**0.08**	**0.95**	**4.14**
	MLP-DAPHT	5.82	1.12	0.87	6.87
	(SOARES *et al.*, 2011)	4.86	3.25	0.87	5.67
	Sch. Hall (log)	4.62	3.37	0.94	6.15

erroneous diameter estimates by overestimating or underestimating the actual diameters, mainly in Classes 1 and 5, as evidenced by the $RMSE_{\%}$ values for both classes. Finally, the calculated volume values were very close to the actual values.

Comparing the tested models with the Schumacher and Hall (log) model highlights that the latter estimates the total volume of the tree, whereas the neural network-generated models calculate the volume up to the height of the commercial diameter. Consequently, these differences manifest in more sensitive measures such as $MAE_{\%}$ and $RMSE_{\%}$.

Figure 3.7 illustrates the dispersion between the cubed volume and the estimated volume for each model in all analyzed classes, considering both training and testing data.

After analyzing Figure 3.7(a), it is evident that the dispersion for the LSTM-VA model was significantly reduced. However, a higher dispersion was observed in the LSTM-DA model in Figure 3.7(b), which was expected due to the noise generated from



(a) *LSTM-VA* (b) *LSTM-DA* (c) *MLP-PR*

(d) *MLP-PD* (e) *MLP-DAPHT* (f) *Schumacher e Hall (log)*

Figure 3.7: *Volume scatter plot for the models.*

the data augmentation.

Although the dispersions between the recursive prediction model in Figure 3.7(c) and the direct prediction model in Figure 3.7(d) are quite similar, there is slightly greater dispersion in the MLP-PD model shown in Figure 3.7(d). However, neither of these dispersions showed significant differences in the calculated volumes.

In Figure 3.7(e), the MLP-DAPHT model shows large differences in relation to the Schumacher and Hall (log) model (3.7(f)). Although it presented interesting results in volume calculation, it showed a noisy dispersion where estimated volume values vary considerably and, at other times, are closer. This indicates a strong bias in the model and suggests that it may not be as effective for generalization purposes.

The results of the Graybill F-test are presented in Table 3.8. At a significance level of 5%, the models did not exhibit a significant $F(H_0)$ statistic, indicating that H_0 cannot be rejected. This suggests no significant differences exist between the values obtained by the models and those measured in the field. However, it is noteworthy that the MLP-PD model shows more significant variation than the other models, even within the equality criteria of the applied test.

Table 3.8: *Statistical differences evaluated by Graybill's F test.*

Models	$F(H_0)$
LSTM-VA	0.983
LSTM-DA	0.032
MLP-PR	0.953
MLP-PD	7.117
MLP-DAPHT	0.40
Sch. Hall (log)	0.034

3.4.3 Volume Residual Analysis

Figure 3.8 shows the residual volume plot for each of the models for all 615 trees at Site III (training + testing). The formal analysis of the distribution of residuals was evaluated using the Lilliefors normality test.

Figure 3.8(a) shows a random dispersion distributed around zero, indicating that the LSTM-VA model has good fitting ability. The normality test indicated a normal distribution with a p value of 0.059 for the residuals of this model.

Figure 3.8(b) shows the residual distribution of the models. The observed overfitting is caused by errors during the prediction of diameters resulting from data augmentation, which subsequently affects the volume calculation.

The MLP-PR model (Figure 3.8(c)) exhibited good fitting capacity, revealing a distribution similar to that observed in the LSTM-VA model but with a better fit regarding outliers. The model presented a normal distribution with a p-value of 0.08.

On the other hand, the MLP-PD and MLP-DAPHT models (Figures 3.8(d) and 3.8(e), respectively), which both estimate the final volume of trees, exhibited significant divergence in terms of residual distribution, highlighting the contrast in their performance.

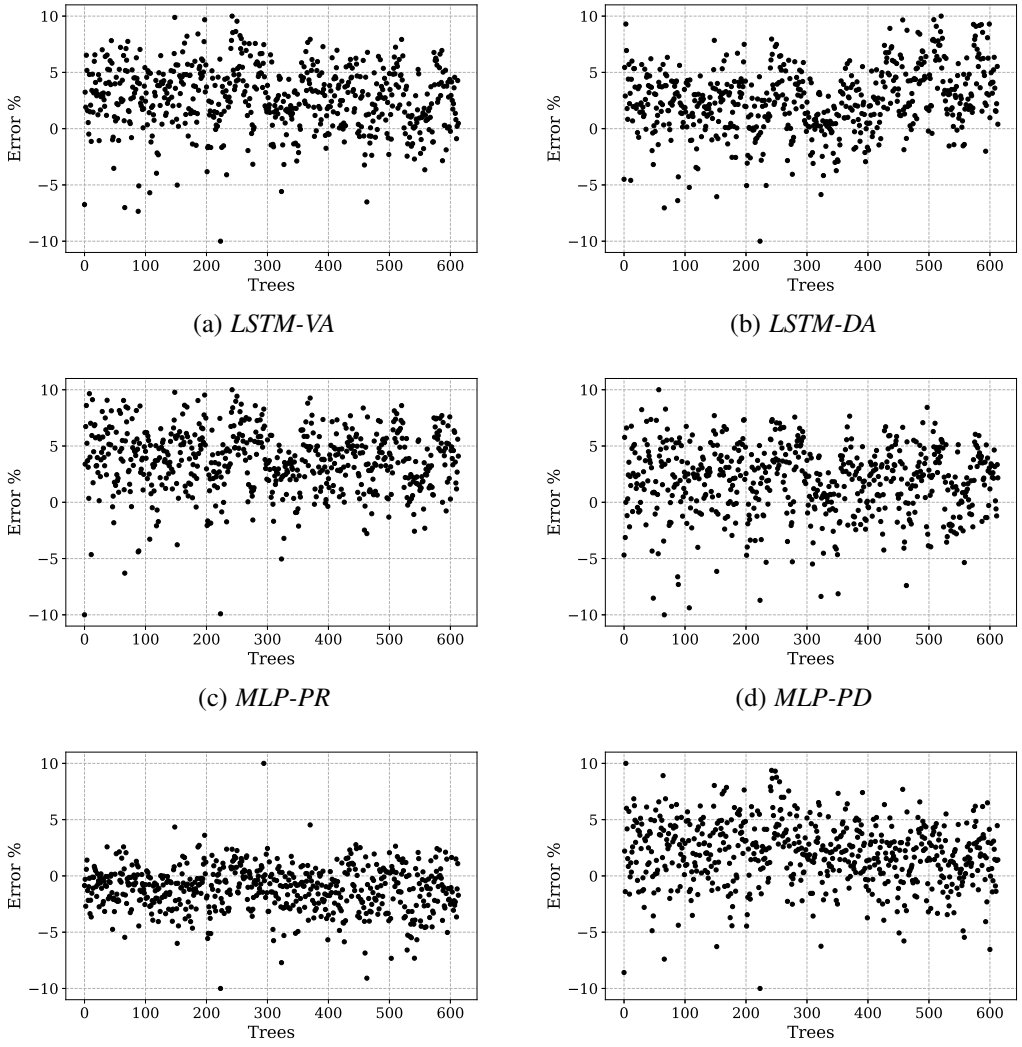
The MLP-PD model showed a random distribution around zero, demonstrating moderate fitting to the data even in the presence of outliers. The maximum relative errors varied around 20%, and the p-value was 0.058.

The MLP-DAPHT model presented a significant variation, indicating that it did not adapt well to the dataset. This model obtained outliers with values greater than 45%, which was also reflected in the normality test with a p-value of 0.008, proving that the null hypothesis should be rejected.

3.4.4 Comparison with other sites

Machine learning problems typically depend heavily on the quality and quantity of data. The ability of a network to handle unknown information with reasonable responses is of great importance.

Considering this, diameter prediction and volume calculation analyses were performed using trees from Sites I and II to verify the models' ability to work with data



(a) *LSTM-VA* (b) *LSTM-DA*

(c) *MLP-PR* (d) *MLP-PD*

(e) *MLP- DAPHT* (f) *Schumacher and Hall (log)*

Figure 3.8: *Residual distribution of the models.*

different from the training set.

Site I diameter prediction

The trees in Site I are slightly taller than those in Site III used in the experiments where the model is adapted to smaller trees, and the class division also presents different amplitudes. The results can be evaluated in Table 3.9.

Comparing the results presented in Table 3.9, it is possible to observe that all models showed logically close values and a strong correlation. As previously noted, the classes with more trees (Classes 2 and 3) presented higher values of $RMSE_{\%}$, and therefore, errors are propagated to estimate future diameters, resulting in higher errors.

The MLP-PR had the lowest performance and higher errors than the other models. As previously observed, the LSTM-VA and LSTM-DA models achieved very

Table 3.9: *Evaluation of models for estimating Site I diameters.*

Class	Model	RMSE%				*R*	
		Min	Max	Mean	SD	Min	Max
1	LSTM-VA	5.27	7.81	6.40	0.95	0.98	0.99
	LSTM-DA	1.98	7.64	4.56	2.59	0.98	0.99
	MLP-PR	2.52	8.35	5.24	2.42	0.98	0.99
2	LSTM-VA	2.67	8.40	5.38	1.97	0.98	0.99
	LSTM-DA	2.75	9.87	6.66	2.20	0.96	0.99
	MLP-PR	2.64	9.07	5.36	1.88	0.98	0.99
3	LSTM-VA	2.09	24.90	7.18	4.74	0.93	0.99
	LSTM-DA	2.41	17.47	5.82	2.39	0.95	0.99
	MLP-PR	2.44	29.55	10.23	6.18	0.91	0.99
4	LSTM-VA	2.64	12.84	6.85	2.23	0.96	0.99
	LSTM-DA	1.74	18.17	6.97	3.15	0.97	0.99
	MLP-PR	2.27	21.25	10.82	5.12	0.92	0.99
5	LSTM-VA	2.45	18.83	9.52	3.78	0.94	0.99
	LSTM-DA	2.77	19.33	9.52	3.74	0.95	0.99
	MLP-PR	2.58	16.40	7.66	2.93	0.95	0.99

Table 3.10: *Comparison between the total volumes calculated for Site I.*

Class	Actual	LSTM-VA	LSTM-DA	MLP-PR	MLP-PD	MLP-DAPHT	Sch. and Hall (log)
1	0.0869	0.0926	0.9833	0.0946	0.0735	0.1930	0.0821
2	1.4335	1.4820	1.5062	1.4742	0.8054	0.7394	1.3729
3	21.1229	20.9958	20.8304	20.0445	18.7600	23.5379	21.3479
4	66.2907	64.6813	62.5992	59.28	59.4458	69.6962	66.4334
5	55.0128	56.6890	55.2270	60.50	53.0777	78.4189	55.2160
Total	143.9468	143.9407	141.1461	140.9443	141.3772	172.5854	144.4523

close results. For smaller classes, LSTM-DA showed advantages with increased data.

Despite the high RMSE values presented by the models, the mean of the LSTM-VA and LSTM-DA models remained below 10%, demonstrating that their capacity for generalization in diameter prediction yielded good results.

Site II diameter prediction

The trees in Site II have characteristics such as class distribution and diametric proximity similar to the training dataset. Table 3.11 presents the analyses performed on Site II trees.

The behavior of the models regarding Site II displayed similarities with Classes 2, 3, and 4, where the highest values of RMSE% were observed across all models.

The LSTM models stood out in our analysis because of their consistent performance. In contrast, the MLP model, while still performing, exhibited the lowest performance among the models. The differences between actual and estimated diameters were more pronounced, as indicated by the RMSE values. However, it's worth noting that, except for Classes 2 and 3 for the LSTM-VA model, the average RMSE values were be-

Table 3.11: *Evaluation of models for estimating Site II diameters.*

Class	Model	RMSE%				R	
		Min	Max	Mean	SD	Min	Max
1	LSTM-VA	3.12	6.66	4.89	1.32	0.98	0.99
	LSTM-DA	4.00	7.68	5.37	1.39	0.97	0.99
	MLP-PR	3,15	7,22	5,15	1,84	0,96	0,99
2	LSTM-VA	4,61	22,83	12,57	4,24	0,92	0,99
	LSTM-DA	2,60	8,18	4,70	1,40	0,98	0,99
	MLP-PR	3,40	20,97	10,58	3,94	0,91	0,99
3	LSTM-VA	1,71	32,56	12,29	7,65	0,91	0,99
	LSTM-DA	2,07	18,79	6,26	2,82	0,96	0,99
	MLP-PR	2,25	37,68	17,77	8,33	0,90	0,99
4	LSTM-VA	2,61	17,20	6,98	2,76	0,96	0,99
	LSTM-DA	2,47	16,34	6,43	2,11	0,98	0,99
	MLP-PR	1,90	35,34	18,50	7,99	0,80	0,99
5	LSTM-VA	2,03	13,34	5,78	2,72	0,98	0,99
	LSTM-DA	2,58	10,75	6,43	2,11	0,98	0,99
	MLP-PR	4,92	17,15	10,46	3,20	0,95	0,99

low 10%. The LSTM-DA model consistently showed results below 10% for all observed classes, indicating good generalization ability.

These findings strongly suggest that, for diameter predictions as a sequential data, the recurrent architecture of LSTM models, which consistently outperformed other models, proved to be more effective, even when applied to a different data set.

Comparison and statistical analysis of volume calculation for Sites I and II

As mentioned, all trees were measured, and volume was calculated using the Smalian formula. Sites I and II were parameterized for the Schumacher and Hall (log) model with the established coefficients β, as was Site III (CABACINHA, 2003). The tree division into classes at the Site I resulted in 6 classes, with the sixth class having measurements outside the defined amplitude for class separation.

For Site I, the corresponding values for fitting the Schumacher and Hall (log) model were: $\beta_0 = -10.5904$, $\beta_1 = 1.8906$, $\beta_2 = 1.2173$, and $\varepsilon = 0$. The R^2 value was 99.61%, the residual standard error S_{yx} was 0.01189, and $S_{yx}\%$ was 5.28. For Site II, the corresponding values were: $\beta_0 = -10.6390$, $\beta_1 = 1.7762$, $\beta_2 = 1.3261$, and $\varepsilon = 0$. The R^2 value was 99.64%, the residual standard error S_{yx} was 0.01099, and $S_{yx}\%$ was 4.45. The results of comparing the volumes of all models are presented in Figure 3.9 and detailed in Table 3.10.

As this is a set of data not seen during the training phase of the models, some volumes are significantly different from the actual values. However, some models showed good approximation. The LSTM-VA model showed an excellent approximation of 0.01% in relation to the cubed volumes. The LSTM-DA model presented a percentage difference of 1.94%, and the MLP-PR model reached 2.08%, values that are much higher compared

to the training site. Among the direct prediction models, the MLP-PD and MLP-DAPHT showed percentage differences of 1.78% and 19.89%, respectively. The latter model is unfeasible for field use, considering that the maximum allowable error in the forestry sector is 10%.

Table 3.12: *Comparison between the total volumes calculated for Site II.*

Class	Actual	LSTM-VA	LSTM-DA	MLP-PR	MLP-PD	MLP-DAPHT	Sch. and Hall (log)
1	0.1587	0.1670	0.1621	0.1742	0.1736	0.1568	0.1583
2	6.8116	6.0744	6.6105	6.1240	6.4034	6.4564	6.9742
3	49.7618	47.7294	47.2327	45.3864	47.5062	50.4661	50.7882
4	69.4123	70.8067	66.5720	60.1668	66.4532	70.5456	70.6064
5	17.6997	17.5400	17.4133	18.4459	16.8859	19.8470	17.9236
Total	143.8441	142.3175	137.9906	130.2943	137.4223	147.4719	146.4507

The LSTM-VA model showed the best generalization ability among all the

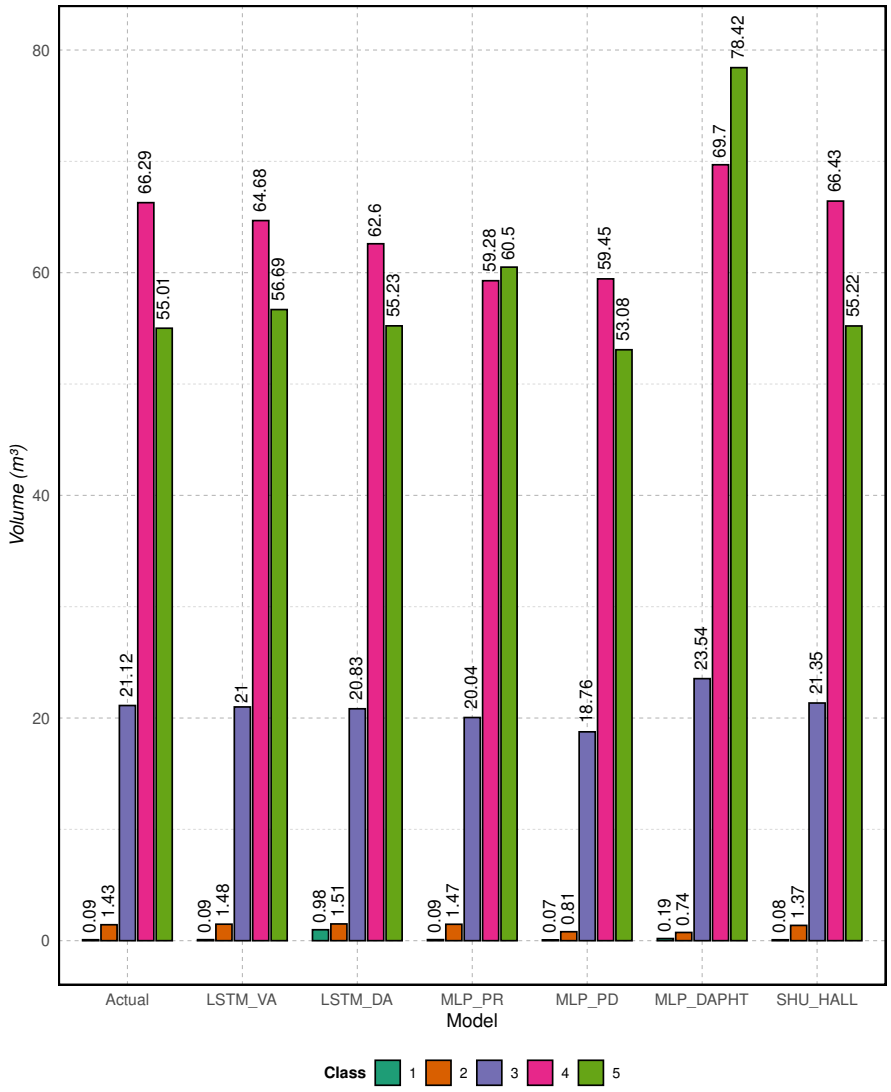


Figure 3.9: *Comparison of total volumes calculated for Site I.*

models studied. The Schumacher and Hall (log) model with parameters adjusted for Site I showed a difference of 0.35% compared to 0.01% for the neural network. The same process was performed to analyze the trees in Site II, and the results are presented in Figure 3.10 and detailed in Table 3.12.

The LSTM-VA model demonstrated the best generalization ability among all the studied models for Site II. Some of the models underestimated the actual values, showing a tendency to calculate lower volumes than expected, which led to higher percentage differences.

Nevertheless, the LSTM-VA model exhibited a percentage difference of 1.06%, the LSTM-DA model showed a difference of 4.06%, and the MLP-PR model achieved a significant difference of 9.41% compared to Sites I and III. The direct volume calculation models, MLP-PD and MLP-DAPHT, reached percentage differences of 4.46% and 2.52%,

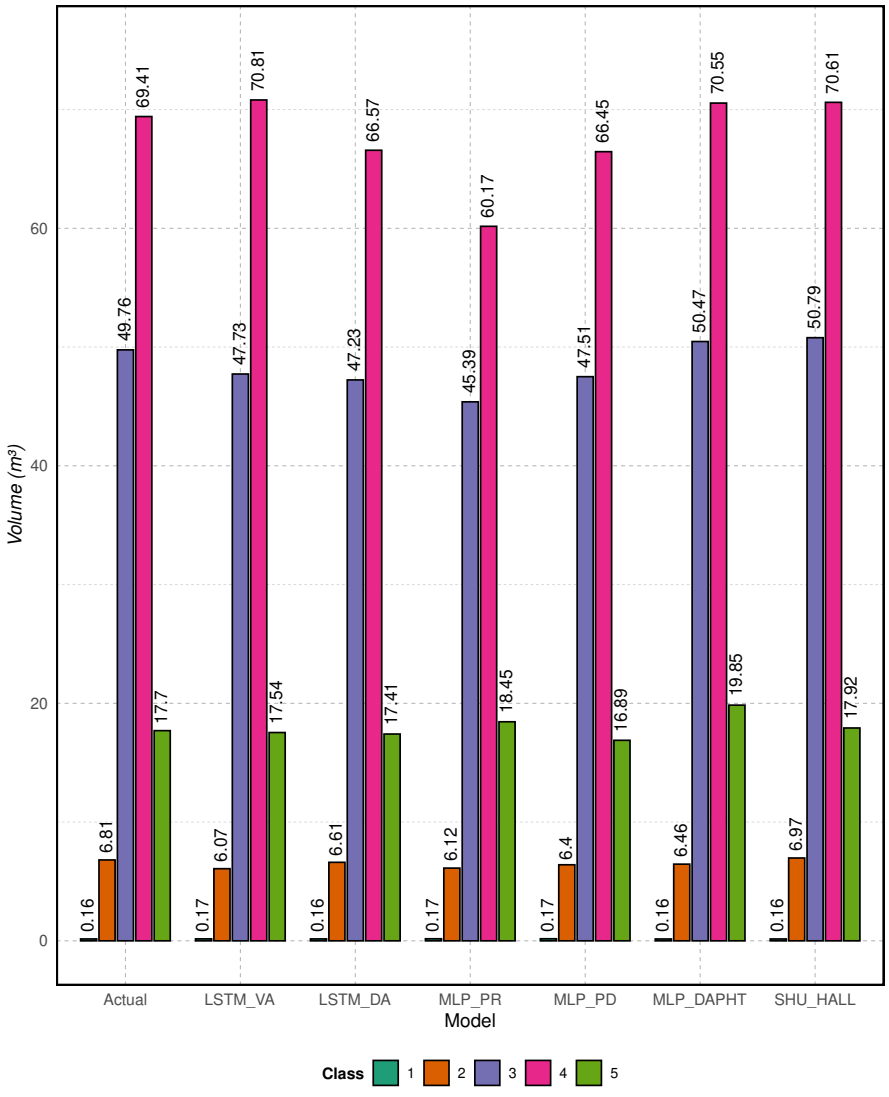


Figure 3.10: *Comparison of total volumes calculated for Site II.*

respectively. The Schumacher and Hall (log) dendrometric model, parameterized for Site II, achieved a percentage difference of 1.81%.

Therefore, among the models across all comparison sites, the LSTM-VA demonstrated robust generalization ability, achieving consistently good results.

Examining Table 3.13, it is evident that the models based on LSTMs achieved the lowest values for $MAE_{\%}$ and $\text{RMSE}_{\%}$, except the Schumacher and Hall (log) model. Conversely, the models that directly calculate the volume, MLP-PD and MLP-DAPHT, performed the worst across all evaluated metrics. This suggests that these models are unsuitable for analyzing sites for which they were not adequately trained. They tended to overestimate the cubed values, rendering them impractical for field use.

For Sites I and II, the Schumacher and Hall (log) model, previously adjusted for these sites, was used to compare the data from the rigorous measurement with a previously adjusted model. This comparison underscores the generalization power of the LSTM models.

3.5 Conclusions

This chapter analyzed diameter and volume estimation and their generalization across eucalyptus sites. The main challenges in applying deep learning models, particularly LSTMs, were related to data quantity and quality. Since these models require large, high-quality datasets for optimal training, data scarcity in forestry—where collection is costly and labor-intensive—remains a key limitation.

Direct prediction models yielded accurate volume estimates within the training site using only a few input variables but showed limited generalization to other sites (I and II), where they tended to underestimate actual values. In contrast, recursive volume estimation models proved more effective by capturing tree shape, enabling the assessment of taper and other features essential for estimating multiple forest products and economic value.

Our results confirmed the robustness of modern neural networks for sequential diameter analysis. LSTMs outperformed MLPs due to their ability to retain temporal information. Tests at Sites I and II, particularly with the LSTM-VA model, highlighted their potential for improving forest inventory practices, reducing sampling effort, and eliminating the need for destructive measurements.

Although deep learning often demands high computational resources, our models require only three numerical input variables to estimate successive diameters, with volume calculated externally. This allows LSTM models to run efficiently on CPUs, ensuring practical and cost-effective deployment without GPUs.

Table 3.13: *Statistic of volume estimation errors between the cubed volume and that estimated by the models for Sites I and II.*

		Site I				Site II			
Class	Model	$MAE_{\%}$	*BIAS*	*r*	$RMSE_{\%}$	$MAE_{\%}$	*BIAS*	*r*	$RMSE_{\%}$
	LSTM-VA	7.16	6.66	0.98	8.18	4.99	5.22	0.99	5.80
	LSTM-DA	6.58	7.24	0.99	8.92	4.20	2.14	0.99	3.88
1	MLP-PR	9.01	8.85	0.98	12.39	8.17	9.76	0.98	17.08
	MLP-PD	46.25	19.08	0.49	52.38	11.51	9.36	0.97	11.76
	MLP-DAPHT	245.00	121.00	0.81	208.00	6.12	5.65	0.98	6.12
	Sch. Hall (log)	7.60	5.53	0.98	8.03	2.87	0.24	0.99	3.05
	LSTM-VA	5.85	3.38	0.98	6.60	11.25	10.01	0.92	12.60
	LSTM-DA	6.29	5.08	0.89	7.38	5.39	2.95	0.94	6.59
2	MLP-PR	5.39	2.82	0.98	6.26	10.69	10.00	0.94	11.72
	MLP-PD	56.85	43.80	0.95	52.15	9.89	5.11	0.88	11.61
	MLP-DAPHT	66.67	48.31	0.97	57.33	8.14	5.21	0.96	9.04
	Sch. Hall (log)	6.43	4.22	0.92	8.00	4.29	2.38	0.97	5.26
	LSTM-VA	7.63	0.59	0.89	9.14	7.24	4.08	0.88	8.80
	LSTM-DA	6.89	1.38	0.91	8.33	6.93	5.08	0.92	8.73
3	MLP-PR	8.82	5.10	0.90	10.21	9.58	8.79	0.92	11.03
	MLP-PD	13.79	11.18	0.86	15.63	7.74	4.53	0.87	9.79
	MLP-DAPHT	14.45	11.42	0.93	14.65	8.50	1.41	0.95	9.64
	Sch. Hall (log)	5.20	1.06	0.94	6.55	4.42	2.06	0.96	5.23
	LSTM-VA	7.24	3.93	0.89	9.60	6.05	2.00	0.86	7.29
	LSTM-DA	8.52	7.07	0.98	10.01	6.49	4.14	0.87	8.18
4	MLP-PR	11.98	10.56	0.88	13.35	13.75	13.36	0.87	15.16
	MLP-PD	10.84	10.32	0.84	13.35	7.09	4.26	0.75	9.98
	MLP-DAPHT	11.14	5.13	0.90	12.29	8.84	1.63	0.90	10.34
	Sch. Hall (log)	4.38	0.20	0.94	5.20	3.82	1.72	0.95	4.65
	LSTM-VA	6.88	5.19	0.78	8.48	5.38	0.90	0.82	6.68
	LSTM-DA	7.23	5.60	0.79	8.87	5.56	1.61	0.89	7.03
5	MLP-PR	5.89	0.01	0.76	6.97	6.42	4.21	0.86	7.73
	MLP-PD	14.92	14.37	0.66	17.17	7.92	4.59	0.66	10.28
	MLP-DAPHT	33.11	31.70	0.28	33.49	15.38	12.13	0.27	17.07
	Sch. Hall (log)	4.00	0.36	0.89	4.92	3.76	1.26	0.93	4.46

3.5.1 Future Works

Future work will focus on refining models with new data to improve accuracy and generalization across eucalyptus sites. We also plan to explore Gated Recurrent Units (GRUs) and compare them with other architectures in terms of accuracy, efficiency, and generalization. Such models can enhance forest inventories by lowering operational costs, improving volume estimation accuracy, and supporting sustainable forest management and industry growth.

CHAPTER 4

MobileNetV2-Enhanced Depth Map Super-Resolution

> Science is not just a body of knowledge, but a way of thinking.
>
> *Carl Sagan*

Depth super-resolution is a process aimed at generating dense depth maps from sparse depth measurements, addressing the challenge of recovering detailed depth information and semantic context. Traditional methods in this domain have focused solely on enhancing sparse depth data. However, they often need to improve in reconstructing depth details and integrating semantic information, particularly in the absence of multimodal data. This study presents a novel approach to depth map super-resolution, leveraging a MobileNetV2-enhanced architecture that integrates multi-scale image guidance to reconstruct high-resolution (HR) depth maps from low-resolution (LR) counterparts. Our methodology employs a fusion of feature maps, utilizing the feature maps extracted from MobileNetV2, facilitating improved depth completion. We evaluate the proposed framework against established state-of-the-art methods using the publicly available Middlebury dataset to ensure a comprehensive performance assessment. The experimental results demonstrate significant improvements in both qualitative and quantitative metrics, underscoring the efficacy of our approach in enhancing depth map resolution. This chapter was published in International Symposium on Visual Computing (ISVC) (RODRIGUES *et al.*, 2024a).

4.1 Introduction

Depth maps, which provide a pixel-by-pixel representation of the distance between objects in a scene and the camera, are essential in many advanced technological applications such as autonomous navigation, 3D reconstruction, human-computer interaction, and virtual reality (ZHU *et al.*, 2024). The advent of low-cost depth cameras has

made depth information more accessible in everyday use. However, these affordable sensors typically offer limited resolution, posing significant challenges for their effective use. To fully utilize depth data, it is often necessary to perform upsampling, a process where a high-resolution depth map is reconstructed from a low-resolution depth map.

Depth map super-resolution (SR) is a complex and demanding task. The fine details in HR images are frequently lost or distorted in LR images due to their limited spatial resolution. Simply upsampling an LR image can result in blurred structures that should have sharp boundaries, degrading the quality and utility of the depth map.

Recent advancements in depth sensors have greatly enhanced their application in both 3D computer vision research and practical scenarios such as somatosensory interaction, autonomous vehicles, and 3D reconstruction. Despite these improvements, depth sensors still lag behind RGB cameras in terms of resolution, limited by cost and current hardware technologies. This limitation results in sparse point clouds that are inadequate for many downstream applications. Furthermore, LR depth maps often lack high-frequency information, leading to significant loss of detail. Addressing this, depth upsampling aims to recover HR depth maps from their LR counterparts, a task made challenging by the need to preserve fine structures and details.

Convolutional neural networks (CNNs) have shown remarkable performance in image perception tasks, leading to significant progress in depth map super-resolution (DONG *et al.*, 2015; LEDIG *et al.*, 2017; LIU *et al.*, 2022; PARK *et al.*, 2020; ZHANG *et al.*, 2018). Unlike RGB images, depth images capture the relative positions of objects in three-dimensional space, resulting in clear and sharp edges between different surfaces. Traditional CNN models, however, tend to interpolate transitions at these edges, introducing noise and blurring. Additionally, many existing networks fail to fully integrate the features of LR depth maps with corresponding guidance images, resulting in blurred depth pixels in the reconstructed HR maps. Therefore, achieving high-quality HR depth maps remains a formidable challenge.

Many neural networks have been developed to tackle this task, but they often suffer from poor utilization of heterogeneous data and a tendency to favor sparse depth inputs, which compromises their performance. As a result, effective depth completion and depth map super-resolution continue to be critical areas of research in the field of computer vision.

We propose a novel architecture derived from the MSG-Net (HUI *et al.*, 2016) framework, a convolutional neural network designed to upsample low-resolution depth maps using a corresponding high-resolution intensity image as guidance. In contrast to the original MSG-Net approach, which employs an early spectral decomposition to extract high-frequency intensity features from RGB and depth inputs, our architecture utilizes a MobileNetV2 as a feature extractor in the color encoder branch. Meanwhile, the depth

encoder branch performs straightforward convolutions. This design choice streamlines the feature extraction process, enhancing the model's efficiency and effectiveness.

4.2 Related Works

In the domain of depth map super-resolution, numerous methodologies have been devised to tackle the challenge of reconstructing high-resolution depth maps from low-resolution inputs. These approaches can be broadly classified into four primary categories: local methods, global methods, dictionary-based methods, and deep learning techniques.

4.2.1 Local methods

Local methods primarily utilize filtering techniques to perform upsampling by leveraging HR color images and local pixel relationships to do up-sampling for LR depth maps. Yang *et al.* (2007) proposed the use of joint bilateral filter (KOPF *et al.*, 2007) to weights the degree of smoothing in each depth patch based on the color similarity between the center pixel and its neighborhood. Liu *et al.* (2013) introduced upsampling weights designed using geodesic distances to improve the filtering process. Lu and Forsyth (2015) utilized image segmentation to develop a smoothing method that reconstructs depth structures within each segment, enhancing the clarity of depth boundaries. Choi and Jung (2014) proposed distinct up-sampling strategies for continuous and discontinuous regions in the depth map. For discontinuous areas, their depth-histogram-based method enhanced the sharpness of the recovered depth boundaries. These approaches often focus on maintaining the integrity of local features and reducing artifacts during upsampling.

4.2.2 Global methods

Global methods formulate depth upsampling as an optimization problem, incorporating constraints from associated intensity images. Diebel *et al*. introduced a Markov Random Field (MRF) formulation that combines a data term from the LR depth map with a smoothness term from the HR intensity image (DIEBEL; THRUN, 2005). Park *et al.* (2011) employed nonlocal means filtering, using intensity features as weights in depth regularization. Ferstl *et al.* (2013) proposed anisotropic operators to solve the optimization problem of depth upsampling. Xie *et al.* (2015) integrated self-similarity and HR edge map guidance into their MRF-based depth super-resolution method, resulting in improved performance. Li *et al.* (2016) proposed a cascaded global interpolation framework to recover HR depth maps. This cascading structure helps reduce texture-copy artifacts and over-smoothing around weak edges caused by color map guidance. These methods

aim to leverage the global structure and correlations between depth and intensity images to achieve more accurate HR depth maps. However, they are often computationally intensive and may struggle with real-time performance.

4.2.3 Dictionary methods

Dictionary-based methods exploit the relationship between paired LR and HR depth patches through sparse coding. Yang *et al.* (2010) sought the coefficients of this representation to generate HR outputs, demonstrating the potential of dictionary learning in depth map SR. Timofte *et al.* (2013) improved upon this by introducing anchored neighborhood regression, enhancing the robustness and accuracy of the sparse-coding approach. Ferstl *et al.* (2015) proposed learning a dictionary of edge priors for anisotropic guidance, further refining the depth map reconstruction process. Kwon *et al.* (2015) proposed a data-driven approach to generate HR depth maps using multi-dictionary sparse representation, effectively addressing the problem of over-smoothing in the recovered depth. These methods highlight the effectiveness of sparse representations in capturing the intricate details necessary for high-quality depth upsampling.

4.2.4 Deep learning-based methods

With the advent of deep learning, CNN-based methods have emerged as powerful tools for depth mapping. Dong *et al.* (2015) pioneered the use of an end-to-end super-resolution convolutional neural network (SRCNN) for image restoration, showcasing the potential of deep learning in this domain. Wang *et al.* (2015) developed a cascade of sparse coding-based networks (CSCN), which integrates modules from the learned iterative shrinkage and thresholding algorithm (LISTA) (GREGOR; LECUN, 2010). Riegler *et al.* (2016) advanced the generation of HR depth maps by incorporating variational optimization within the framework of deep convolutional neural networks. Guo *et al.* (2018) devised a methodology to learn residuals across different resolutions, thereby enhancing the interpolation accuracy of LR depth maps. Voynov *et al.* (2019) introduced perceptual metrics to constrain the network during HR depth map reconstruction, demonstrating through experiments that such quality measures, aligned with human perception, are more effective. These approaches, however, required careful patch reconstruction to ensure consistency in the HR output. Recent advancements include multi-scale networks and cascaded restoration networks that leverage HR intensity images and edge information to guide depth upsampling, resulting in significant improvements in boundary resolution and overall accuracy. Hui *et al.* (2016) proposed MSG-Net, a convolutional neural network that performs upsampling of low-resolution depth maps using a corresponding high-resolution intensity image as guidance. The network uses a multi-scale architec-

ture for progressive depth map upsampling, with guidance from high-resolution intensity features at each level. Xian *et al.* (2020) introduces a multi-scale progressive fusion network aimed at improving the resolution of low-quality depth maps. The authors propose a method that integrates features from high-resolution color images to enhance depth map quality. Their approach utilizes a step-wise fusion strategy to address issues related to boundary clarity and detail preservation.

4.3 Proposed Method

In this section, we introduce our novel architecture inspired by the MSG-Net framework.

4.3.1 Overview

Consider an LR depth map D_l, which represents a sampled version of its high-resolution counterpart D_h. Additionally, we have access to a corresponding high-resolution image I_h of the same scene. Our objective is to reconstruct D_h using both D_l and I_h, employing the high-resolution RGB image as a guide to enhance the quality and accuracy of the recovered depth map.

4.3.2 Network Architecture

As depicted in Figure 4.1, the architecture of the proposed method consists of two primary branches: the depth encoder branch and the color encoder branch. These encoder branches, with multi-level receptive fields, generate a set of deep hierarchical features. Assume the input resolution for these encoders is $H \times W$. For the depth encoder branch, a low-resolution depth map $D_{lH' \times W'}$ is initially upscaled to the specified resolution using bicubic interpolation, where H' and W' are determined according to a downsampling factor s (e.g., $s = 2, 4, 8$, or 16). Features are then extracted from this upscaled depth map through convolutional operations. Simultaneously, the corresponding high-resolution RGB image $I_{H \times W}$ is processed through the color encoder branch. The MobileNetV2 architecture is employed as a feature extractor, with the classification head removed, to process the high-resolution RGB image. The extracted features then guide the progressive upsampling of the low-resolution depth map at multiple levels, aiming to produce the desired high-resolution depth map.

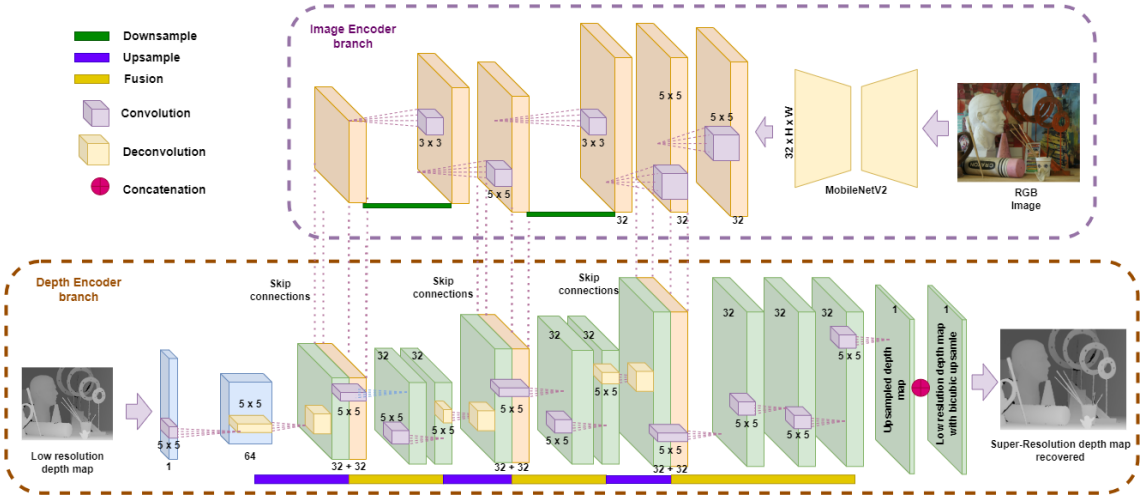


Figure 4.1: *The architecture of DepthMobileNet. Our architecture is inspired by MSG-Net. Similarly, we only show an upsampling CNN with upscaling factor 8. Here there are three levels of upsampling at multiple scales. Each level consists of an upsampling and a fusion stage.*

RGB Encoder Branch

Since high-resolution color images are easier to obtain and possess rich informational characteristics, we extract structural features from the color branch to guide the recovery of HR depth maps. We used a pretrained model of MobileNetV2 trained on the ImageNet dataset to avoid training the model from scratch and speed up the learning process (SANDLER *et al.*, 2018). The MobileNets architecture employs depth-wise separable convolutions to construct lightweight deep neural networks, enhancing computational efficiency (HOWARD *et al.*, 2017). MobileNets are compact size and high-speed computational capabilities make them well-suited for deep learning tasks on mobile platforms (MBUNGE *et al.*, 2021). Following the feature extraction from the preceding layers, a 1×1 convolution is applied to adjust the number of output channels to 32. Additionally, an upsampling layer is incorporated to enhance the feature maps to the desired spatial dimensions $H \times W$.

Depth Encoder Branch

The depth encoder branch is analogous to the color encoder branch. However, unlike the color encoder branch, we replace the MobileNetV2 block with a traditional convolutional. This modification is necessary because the depth map contains less channel information than the color image, and using a more complex feature extraction mechanism would likely lead to network overfitting. We first up-sample the input LR depth map by bicubic interpolation to match the resolution $H \times W$ of the color image. Then, we use two transitional convolutional layers with 5×5 convolution kernel to generate the input feature map. The convolutional layer is followed by an activation function for non-

linear mapping. We use parametric rectified linear unit (PReLU) (HE *et al.*, 2015) as the activation function σ. After this process, an inverse convolution (deconvolution layer) is performed, and PReLU is used as the activation function. A deconvolution layer, also known as a transposed convolution layer, is used in neural networks to increase the spatial resolution of an image or feature map.

Reconstruction

The augmented feature maps in the depth encoder branch are combined with the corresponding feature maps in the color encoder branch, which have the same resolution (as illustrated in Figure 4.1). Since the features in I_h may not be consistent with the depth structures in D_h, a post-fusion layer is introduced to learn a better coupling. An additional post-fusion layer is included to further enhance the fusion process before reconstruction.

4.4 Experiments

In this section, we evaluate the performance of our proposed network on publicly available datasets from both qualitative and quantitative perspectives. In implementing our method, we use the ADAM optimizer to adjust the network parameters, keeping the learning rate fixed at 1×10^{-5} throughout the training process. We trained using PyTorch on a PC with an AMD Ryzen 9x processor and a NVIDIA RTX 3070 GPU.

4.4.1 Datasets

We used two public datasets in this study: The Middlebury dataset (SCHARSTEIN; SZELISKI, 2002; SCHARSTEIN; PAL, 2007) contains high-quality depth maps and color maps, combined with the MPI Sintel depth dataset (BUTLER *et al.*, 2012) provides HR color maps and corresponding depth maps.

Following the methodology outlined in Hui *et al.* (2016), we selected 58 RGB-D images from the MPI Sintel Depth dataset and 34 RGB-D images from the Middlebury dataset. We used a total of 82 images for training and 10 images for validation. To prepare the data, we cropped the HR depth maps to a size of 128×128 pixels and performed sampling with a stride of 32 for scaling factors of 2, 4, 8, and 16. Each patch was augmented with a 90° rotation, resulting in approximately 700,000 training patches for each scale. The LR depth maps were generated by down-sampling the full-resolution input patches using bicubic interpolation with the specified scaling factors (2, 4, 8, and 16). We trained this architecture for 2e+3 epochs.

4.4.2 Evaluation

To effectively evaluate the performance of our proposed method, we conducted experiments comparing our architecture with bicubic up-sampling and several state-of-the-art methods. These include a local method (GF (HE *et al.*, 2012)), a global optimization method (TGV (MANDAL *et al.*, 2016)), CNN-based color map super-resolution methods (SRCNN (DONG *et al.*, 2015), RDN (ZHANG *et al.*, 2018), SRFBN (LI *et al.*, 2019)), and CNN-based depth map super-resolution methods (MSG-Net (HUI *et al.*, 2016), DU-DEAL (WANG *et al.*, 2020), DepthSR (GUO *et al.*, 2018), and PDDSR (VOYNOV *et al.*, 2019)). We performed both quantitative and qualitative analyses at four scales (2, 4, 8, and 16) on the datasets processed as described above. We used Root Mean Squared Error (RMSE) to evaluate the performance of our method compared to other state-of-the-art methods. Tables 4.1 and 4.2 present the numerical results of these experiments.

4.4.3 Results

We conducted experiments on the Middlebury dataset, utilizing publicly available codes for several methods: RDN (ZHANG *et al.*, 2018), TGV (MANDAL *et al.*, 2016), SRFBN (LI *et al.*, 2019), DepthSR (HUI *et al.*, 2016), DU-DEAL (WANG *et al.*, 2015), MSG-Net (HUI *et al.*, 2016), and MSPF (XIAN *et al.*, 2020). Additionally, we directly used the data from MSG-Net (HUI *et al.*, 2016) available on GitHub[1]. For inputs with a resolution of 1320 × 1080, we compared the results using four down-sampling scale factors.

The quantitative results are presented in Tables 4.1 and 4.2, which show that the RMSE values of our proposed method, based on six test datasets (Art, Books, Dolls, Laundry, Moebius, Reindeer), are generally lower than those of other methods. Most learning-based methods outperform traditional methods in SR results. Among CNN-based methods, those focused on depth map SR typically achieve better outcomes than those focused on RGB image SR. Figure 4.2 demonstrates that the model improves its reconstruction quality, measured by PSNR, rapidly in the first training epochs, gradually converging as the number of interactions increases.

Similarly to Hui *et al.* (2016), we calculated the time required to upscale different LR depth maps to their maximum resolution using MSG-Net and the proposed model. The results obtained are presented in Table 4.3. It can be seen that the computation time of MSG-Net is slightly lower than that of our model. This increase is due to the depth of

[1] https://github.com/twhui/MSG-Net/tree/master

our model and its enhanced feature extraction capacity, which raises the computational complexity, leading to a longer processing time during upscaling.

The visual comparison results are presented in Figure 4.3. As illustrated, our SR method significantly outperforms the standard MSG-Net and traditional methods, producing less noise. Additionally, compared to current learning-based methods, our approach provides clearer and more accurate SR results for depth maps. However, when compared to the MSPF (XIAN *et al.*, 2020) results, our method was slightly inferior to the state-of-the-art performance. Nevertheless, the training time of our model was significantly shorter than that of the MSG-Net. While MSG-Net required over 500,000 epochs, our model achieved better results with only 2,000 epochs of training.

Table 4.1: *Quantitative analysis results across four scales based on RMSE values.*

Method	Art				Books				Dolls			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Bibcubic	3.53	3.84	4.47	5.72	1.31	1.61	2.34	3.34	3.28	3.34	3.47	3.72
GF	2.75	3.91	5.32	8.36	1.36	1.76	2.1	3.36	1.23	2.48	3.97	4.86
TGV	3.03	3.78	7.08	11.59	1.29	1.61	2.15	3.05	1.63	1.96	2.62	4.08
RDN	2.61	3.82	5.87	-	1.46	2.01	3.08	-	1.25	1.7	2.22	-
SRFBN	1.99	3.02	3.58	-	0.54	1.22	1.51	-	1.04	1.81	2.06	-
MSG-Net	0.66	1.47	2.45	4.57	0.37	0.67	1.03	1.6	0.34	0.69	1.05	1.6
DU-REAL	0.62	1.15	2.15	4.32	0.34	0.57	1.01	1.54	0.31	0.65	0.98	1.42
DepthSR	0.53	1.2	2.22	3.91	0.31	0.6	0.89	1.51	0.32	0.62	0.85	1.48
MSPF	**0.51**	1.05	**2.08**	**3.87**	**0.3**	**0.55**	**0.83**	**1.47**	**0.29**	0.58	**0.8**	**1.35**
Ours	0.54	**1.03**	2.08	3.90	0.34	0.59	1.01	1.49	0.34	**0.55**	1.01	1.37

Table 4.2: *Quantitative analysis results across four scales based on RMSE values.*

Method	Laundry				Moebius				Reindeer			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Bibcubic	3.35	3.49	3.77	4.35	3.28	3.36	3.5	3.81	3.4	3.52	3.83	5.82
GF	2.26	2.67	3.84	5.23	1.92	2.29	3.85	5.22	2.27	2.89	3.98	5.85
TGV	2.15	2.51	3.82	6.41	1.21	1.65	2.13	2.73	2.41	2.71	3.79	7.27
RDN	2.53	3.22	4.65	-	1.22	1.61	2.39	-	3.32	2.93	4.41	-
SRFBN	1.67	2.13	2.28	-	1.12	1.43	1.52	-	1.63	2.07	2.15	-
MSG-Net	0.37	0.79	1.51	2.62	0.35	0.66	1.02	1.63	0.42	0.98	1.76	2.91
DU-REAL	0.35	0.76	1.49	2.56	0.34	0.62	0.97	1.54	0.39	0.95	1.61	2.53
DepthSR	0.34	0.78	1.32	2.26	0.32	0.59	0.92	1.51	0.39	0.96	1.57	2.47
MSPF	**0.32**	**0.71**	**1.21**	**2.15**	**0.29**	**0.56**	**0.85**	**1.42**	**0.35**	**0.82**	**1.45**	**2.21**
Ours	0.36	0.78	1.21	2.22	0.31	0.59	0.91	1.47	0.39	0.88	1.72	2.89

4.5 Conclusions

In this study, we propose a neural network inspired by MSG-Net for depth map super-resolution. Our method employs a pre-trained MobileNetV2 as a feature extractor for high-resolution color images. The features extracted by the color encoder branch are fused with those from the low-resolution depth encoder branch. During the reconstruction process, the high-resolution color image features are hierarchically

Table 4.3: *Upscaling time (sec).*

Model	2x	4x	8x	16x
MSG-Net	0.247	0.296	0.326	0.368
Ours	0.252	0.311	0.331	0.388

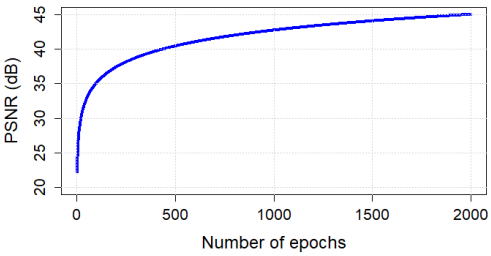


Figure 4.2: *Convergence curve.*

integrated with the low-resolution depth map features. Our experiments demonstrate that the proposed network achieves competitive performance compared to MSG-Net; however, the MSPF architecture proved more advantageous in the tests conducted. In the future, we aim to apply this architecture to additional datasets and enhance the network. Like other learning-based super-resolution networks, our method tends to produce biased results in regions with significant depth errors and data sparsity. Therefore, we plan to expand the dataset to enable the network to be used in a variety of real-world scenarios.

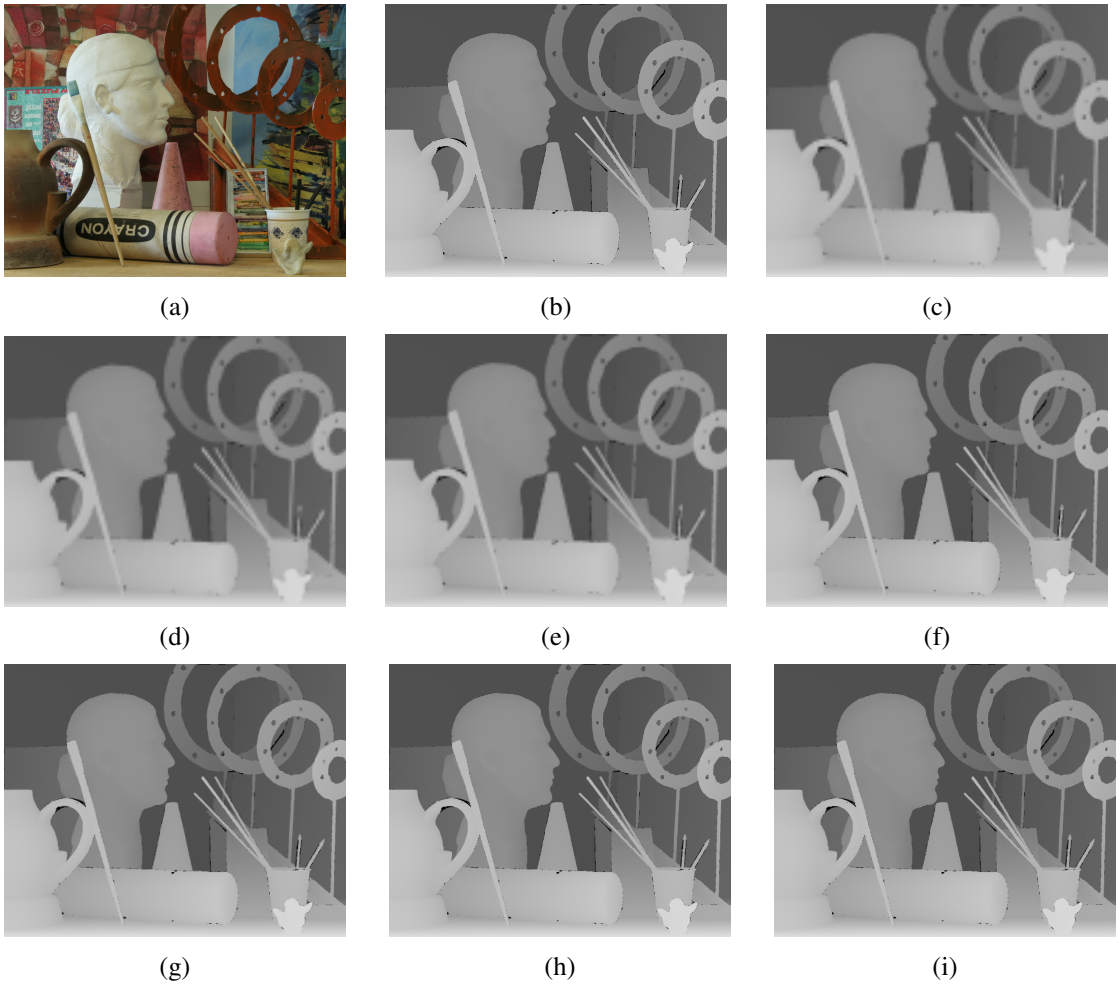


Figure 4.3: *Visual comparison of depth map super-resolution at 8× magnification on the Middlebury datasets. Panel (a) displays the color image. Panels (b) through (i) show the high-resolution depth map, as well as the results obtained using the Bicubic method, RDN (ZHANG* et al.*, 2018), TGV (MANDAL* et al.*, 2016), SRFBN (LI* et al.*, 2019), MSG-Net (HUI* et al.*, 2016), MSPF (XIAN* et al.*, 2020), and our proposed method.*

CHAPTER 5

3D Point Cloud Segmentation of LiDAR-Scanned Eucalyptus Trees: A Deep Learning Analysis

> Make requests, and they will be answered.
>
> *Luke 11:9*

Accurate segmentation of 3D point clouds is essential for advancing precision forestry, particularly in eucalyptus plantations, where precise tree measurements underpin resource management and yield estimation. This study evaluates some state-of-the-art deep learning architectures, including point-based, voxel-based, and attention-based models, for segmenting LiDAR-scanned eucalyptus trees. A dataset comprising 55 clonal eucalyptus trees was constructed using an iPhone 12 Pro Max LiDAR sensor, with diameter measurements taken at four key heights. The performance of PointNet, PointNet++, DGCNN, PointMLP, PointNeXt-S, Point Transformer, and Point Voxel Transformer (PVT) was assessed using the mean mIoU% metric. Results indicate that transformer-based models achieved superior segmentation accuracy, while classical point-based networks demonstrated competitive performance with lower computational costs. Despite dataset limitations, the models effectively delineated key tree structures, demonstrating their applicability to forestry tasks. This research highlights the potential of deep learning in forestry point cloud analysis and suggests avenues for enhancing segmentation performance through larger datasets and advanced augmentation techniques. This chapter has been accepted for publication in the proceedings of the Canadian Conference on Electrical and Computer Engineering (CCECE).

5.1 Introduction

The Brazilian forestry sector is globally recognized for its efficiency and rapid planting-to-harvesting cycles. In 2023, Brazil's commercial tree cultivation area surpassed 10 million hectares, with eucalyptus plantations accounting for 7.8 million hectares, or 76% of the total. Over the past decade, eucalyptus plantations have grown by 41%, underscoring their importance in Brazil's forestry economy (IBÁ, 2024).

Accurate measurement of tree structural attributes, such as height, diameter at breast height (DBH), and stem taper, is critical for forest inventories in eucalyptus plantations (FREITAS; WICHERT, 1998). These variables are essential for calculating key metrics like basal area, wood volume, and overall productivity. Traditionally, data collection in eucalyptus forestry is conducted manually within sampling areas, a process that is labor-intensive, time-consuming, and prone to human error (GAFFREY *et al.*, 2001). Errors in crucial variables, such as DBH and tree height, can substantially affect eucalyptus volume estimates. For instance, a one-meter error in tree height can result in a 14% miscalculation of cylindrical tree volume, while a mere 1 cm error in DBH can lead to a 19% discrepancy (COUTO; BASTOS, 1987). This level of inaccuracy not only affects the reliability of forest inventories but also leads to suboptimal resource allocation and economic inefficiencies in eucalyptus production.

The emergence of advanced technologies is crucial for generating accurate forest inventories, and LiDAR stands out as a key tool in this area. Light Detection and Ranging (LiDAR) utilizes laser pulses to measure distances and produce detailed three-dimensional data on objects and surfaces. This technology is widely applied in fields such as topographic mapping, environmental monitoring, remote sensing, and autonomous vehicles (DUBAYAH; DRAKE, 2000). LiDAR produces highly accurate point clouds, providing critical insights for forestry applications. It is widely used to create digital terrain models, identify trees, estimate wood volumes, and analyze forest structure (NGUYEN *et al.*, 2021).

This study explores the application of advanced deep learning architectures, ranging from traditional point-based and voxel-based models to modern transformer-based approaches, for segmenting point clouds generated from LiDAR scans. Through statistical analysis, we assessed the performance of architectures renowned for excelling on classic benchmarks like ShapeNetPart (YI *et al.*, 2016), ModelNet40 (WU *et al.*, 2015), S3DIS (ARMENI, 2017) and Semantickitti (BEHLEY *et al.*, 2019).

The primary contribution of our work lies in the detailed segmentation and analysis of the regions associated with the first four key measurements of standing eucalyptus trees. By addressing this specific and critical aspect of eucalyptus mapping, our study not only extends the application of point cloud segmentation techniques to forestry

but also demonstrates the potential of state-of-the-art neural networks in improving the accuracy of structural measurements essential for forest inventories and sustainable management.

The chapter proceeds as follows. Section 5.2 surveys the relevant literature. Section 5.3 elaborates on the proposed methodology, providing mathematical formulations and schematic representations of the workflow. Section 5.4 analyzes the experimental findings and discusses their implications. Concluding remarks are presented in Section 5.5.

5.2 Literature Review

The 3D point cloud (RUSU; COUSINS, 2011), an innovative primitive representation for objects, has become increasingly prevalent across various research fields due to its simplicity, flexibility, and powerful representational capabilities. This representation is widely used in object recognition (CHEN *et al.*, 2021) and reconstruction (GIRDHAR *et al.*, 2016).

5.2.1 Point Cloud Analyses

Point cloud analysis plays a pivotal role in 3D vision, primarily relying on two methodologies for extracting structural representations: point-based and voxel-based approaches. Point-based methods (QI *et al.*, 2017a; QI *et al.*, 2017b; WANG *et al.*, 2019; THOMAS *et al.*, 2019; QIAN *et al.*, 2022; MA *et al.*, 2022) directly process unordered points and can be classified into Set Abstraction-based (QI *et al.*, 2017a; QI *et al.*, 2017b; QIU *et al.*, 2021), Dynamic Graph-based (WANG *et al.*, 2019), and Attention-based (YU *et al.*, 2021; YU *et al.*, 2022; ZHAO *et al.*, 2021a; ZHANG *et al.*, 2022) approaches, each using different strategies to model point relationships. Voxel-based methods (MATURANA; SCHERER, 2015; KLOKOV; LEMPITSKY, 2017; RIEGLER *et al.*, 2017) partition 3D space into ordered voxels and use 3D convolutions for feature extraction, with SparseConvolution-based techniques (KLOKOV; LEMPITSKY, 2017) enabling efficient processing through sparse convolutions.

5.2.2 Point-based Models for Point Cloud Learning

In their seminal work, Qi *et al.* (2017a) introduced PointNet, a deep learning architecture that processes raw point clouds directly for 3D classification and segmentation, preserving geometric details and enhancing computational efficiency. Building on this, PointNet++ (QI *et al.*, 2017b) was developed as a hierarchical neural network that

captures local structures through recursive partitioning. Recent advancements have focused on improving point modeling with novel sampling algorithms (YANG *et al.*, 2019; LANG *et al.*, 2020; YANG *et al.*, 2020; NEZHADARYA *et al.*, 2020). PointMLP (MA *et al.*, 2022) simplifies point cloud analysis by using a residual MLP network, avoiding complex local feature extractors while maintaining high effectiveness. PointNeXt (QIAN *et al.*, 2022) modernizes PointNet++ with enhanced training strategies and scalable architectural modifications, demonstrating significant performance gains.

Several advanced methods enhance point cloud analysis by leveraging graph-based learning techniques. Dynamic Graph CNN (DGCNN) (WANG *et al.*, 2019) utilizes dynamic graph convolutions on kNN graphs to capture geometric relationships effectively. PointWeb (ZHAO *et al.*, 2021a) improves local neighborhood connectivity, while Edge-Conditioned Convolution (ECC) (SIMONOVSKY; KOMODAKIS, 2017) employs dynamic filters based on point cloud edges.

Advanced methods have been developed to apply continuous convolutions directly to 3D point sets, bypassing the need for quantization. PCCN (WANG *et al.*, 2018) uses MLPs for convolutional kernels, while PointConv (WU *et al.*, 2019) and KPConv (THOMAS *et al.*, 2019) base convolution weights on input coordinates. PointCNN (LI *et al.*, 2018) reorders unordered point clouds with specialized operators to facilitate convolution operations.

5.2.3 Attention-based Models

Transformer and self-attention models, initially transformative in natural language processing, have demonstrated successful adaptation to 2D image and 3D point cloud analysis through (HU *et al.*, 2019a; RAMACHANDRAN *et al.*, 2019), who focus on applying self-attention to local image patches. Vector self-attention was applied in Zhao *et al.* (2020), while a novel paradigm reinterpreting images as sequences of patches was introduced by Dosovitskiy *et al.* (2020). The field of 3D point cloud processing was advanced by Zhao *et al.* (2021a) through a transformer model optimized for kNN computation, while the Point-Voxel Transformer (PVT) proposed by Zhang *et al.* (2022) integrates voxel-based and point-based advantages for effective 3D analysis.

This study conducts a comparative analysis of state-of-the-art models, including PointNet (QI *et al.*, 2017a), PointNet++ (QI *et al.*, 2017b), DGCNN (WANG *et al.*, 2019), PointMLP (MA *et al.*, 2022), PointNeXt-S (QIAN *et al.*, 2022), Point Transformer (PointT) (ZHAO *et al.*, 2021a) and Point Voxel Transformers (PVT) (ZHANG *et al.*, 2022) to evaluate their performance in point cloud segmentation for eucalyptus forestry. The analysis highlights the strengths and weaknesses of each architecture, offering valuable insights into their applicability and potential for improving point cloud segmentation

tasks in forestry.

5.3 Materials and Methods

Problem definition: Given a point cloud $P = p_1, p_2, \ldots, p_n$ where each point p_i is represented by its 3D coordinates and potentially additional attributes, the goal is to partition into a set of disjoint subsets $S = S_1, S_2, \ldots, S_k$ such that:

1. $\bigcup_{i=1}^{k} S_i = P$ (Union of all subsets equals the original point cloud).
2. $S_i \cap S_j = \emptyset$ for all $i \neq j$ (Subsets are mutually exclusive).
3. Points within each subset S_i share similar properties or belong to the same object.

the segmentation process can be mathematically expressed as a function $f : P \rightarrow 1, 2, \ldots, k$ that assigns a label to each point, where k is the number of segments.

5.3.1 Dataset

The dataset comprises information collected from 55 clonal eucalyptus trees (*Eucalyptus sp.*) in June 2023. Diameter measurements were taken at four different heights for each tree: 0.30 m, 0.70 m, 1 m, and the DBH at 1.3 m from the ground, using a caliper and diameter tape.

Point clouds were acquired using a LiDAR sensor integrated into an iPhone 12 Pro Max (LUETZENBURG *et al.*, 2021). The scanner settings included a 2-meter range and a 5-millimeter resolution, resulting in a 3D reconstruction error of approximately ±5 mm. The acquisition process involved rotating the smartphone around each tree to capture all points within a 2-meter radius and an average height of 2 meters.

The annotation of each region corresponding to diameter measurements was inspired by the approach presented by ShapeNetPart (YI *et al.*, 2016). This method ensures that each data point is precisely and consistently annotated, enhancing the dataset's accuracy and reliability. The dataset underwent normalization for training, adhering to the methodology outlined by Qi *et al.* (2017a). Figure 5.1 illustrates the prepared and annotated dataset.

5.3.2 Training Details

We conducted an evaluation of five prominent and widely recognized architectures for tasks including classification, semantic segmentation, and partial segmentation. Each network was implemented in accordance with the specifications detailed in their respective original papers. Table 5.1 provides an overview of the evaluated networks along

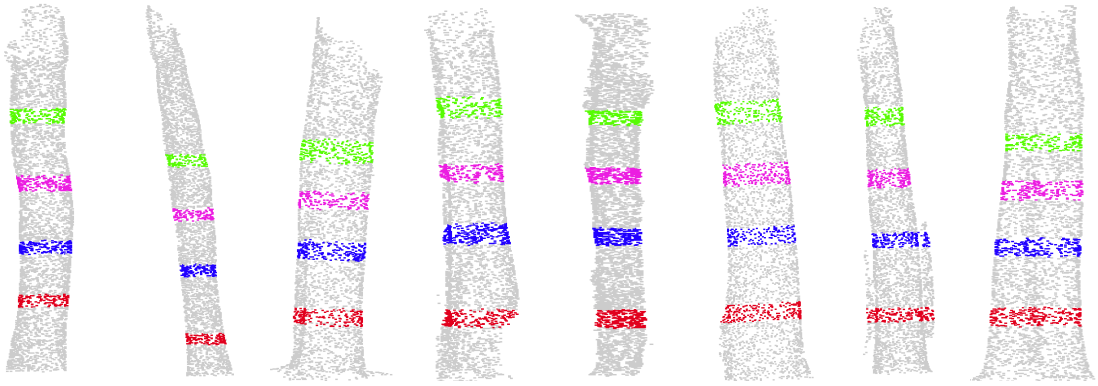


Figure 5.1: *Eucalyptus dataset with annotated tree classes: stem (gray), d30cm (red), d70cm (blue), d100cm (magenta), and dbh (green).*

with the hyperparameters used in our experiments. Certain parameters, such as batch size, were adjusted to accommodate the limited dataset size.

Table 5.1: *Training configurations of evaluated models for eucalyptus point cloud segmentation.*

Methods	θ[1]	lr[2]	Bs[3]	Epochs
PointNet (QI *et al.*, 2017a)	Adam	1×10^{-3}	8	2500
PointNet++ (QI *et al.*, 2017b)	Adam	1×10^{-3}	8*	2500
DGCN (WANG *et al.*, 2019)	SGD	$10^{-1} - 10^{-3}$	8	2500
PointMLP (MA *et al.*, 2022)	SGD	1×10^{-1}	8*	2500
PointNeXt-S (QIAN *et al.*, 2022)	AdamW	1×10^{-3}	2	2500
PointT (ZHAO *et al.*, 2021a)	SGD	5×10^{-2}	8	300
PVT (ZHANG *et al.*, 2022)	SGD	1×10^{-1}	8*	300

[1] Optimizer.
[2] Learning rate.
[3] Batch size.

We use the 5-fold cross-validation strategy to evaluate our model's performance. The dataset is divided into five subsets, and in each fold, four subsets (80%) are used for training while the remaining subset (20%) serves as validation. This process is repeated five times, ensuring that each subset is used for validation exactly once.

All networks were implemented using the PyTorch framework. The experiments were conducted on computing setup featuring an NVIDIA RTX 3090 GPU, 64GB of RAM, and an AMD Ryzen 9 processor.

5.3.3 Metrics

To compare models, we use the mean Intersection-over-Union (mIoU) metric on points, following the evaluation scheme of PointNet. While PointNet evaluates shapes with 16 parts and 50 categories as in ShapeNetPart, in this case, we have only 1 part

(referring to the tree) and 5 categories: *stem*, *d30cm*, *d70cm*, *d100cm*, and *dbh*. The IoU of a shape is calculated by averaging the IoUs of its different parts, and the IoU of a category is obtained by averaging the IoUs of all shapes within that category. The mIoU is then determined by averaging the IoUs of all testing shapes.

5.4 Results and Discussion

We trained the models according to their original specifications, making necessary adjustments such as modifying the number of epochs and batch size to accommodate the smaller size of our dataset. The resulting mIoU values are presented in Table 5.2.

Methods	mIoU	stem	d30cm	d70cm	d100cm	dbh
Point-based						
PointNet (QI *et al.*, 2017a)	0.553	0.605	0.545	0.530	0.525	0.560
PointNet++ (QI *et al.*, 2017b)	0.569	0.625	0.565	0.545	0.540	0.570
DGCNN (WANG *et al.*, 2019)	0.572	0.630	0.570	0.550	0.545	0.565
PointMLP (MA *et al.*, 2022)	0.601	0.655	0.595	0.580	0.575	0.600
PointNeXt-S (QIAN *et al.*, 2022)	0.612	**0.661**	0.605	0.590	0.594	0.610
Attention Point and Voxel-based						
PointT (ZHAO *et al.*, 2021a)	0.586	0.640	0.580	0.565	0.560	0.585
PVT (ZHANG *et al.*, 2022)	**0.623**	0.660	**0.612**	**0.606**	**0.614**	**0.623**

Table 5.2: *Performance comparison of point-based and attention-based methods for eucalyptus structure segmentation. Results show mIoU scores for different classes.*

The results obtained in this study are lower than those achieved by models trained on the ShapeNetPart dataset. Our experiments indicate that the reduced size of our dataset negatively impacted the mIoU metric. This suggests that the models, while effective, require larger datasets to achieve optimal performance metrics comparable to those seen with ShapeNetPart.

Despite dataset limitations, the models exhibited a remarkable ability to accurately segment significant point cloud regions. This indicates that even with a smaller dataset, the models can identify and delineate large areas relevant to the segmented regions. This performance underscores the robustness of the models, even when faced with less-than-ideal data volume.

To provide a deeper understanding of the segmentation quality, a qualitative analysis is presented in Figure 5.3. This analysis visually demonstrates the models' segmentation capabilities, highlighting their strengths and areas for potential improvement.

By examining these qualitative results, we gain valuable insights into how the models interpret and process the data, which can inform future adjustments and improve-

ments in the training process. Although Table 5.2 shows some close values, the visual analysis reveals that certain models have distinct advantages over others. Specifically, the Point Transformers and Point Voxel Transformers models demonstrated superior performance. Nevertheless, classical architectures such as DGCNN and PointNeXt-S also exhibited strong segmentation capabilities.

The inference analysis was conducted on complete point clouds, without the need for point reduction, ensuring that the models' performance was evaluated under realistic conditions. The inference times for each model are presented in Figure 5.2, providing a comprehensive overview of their efficiency in processing point cloud data.

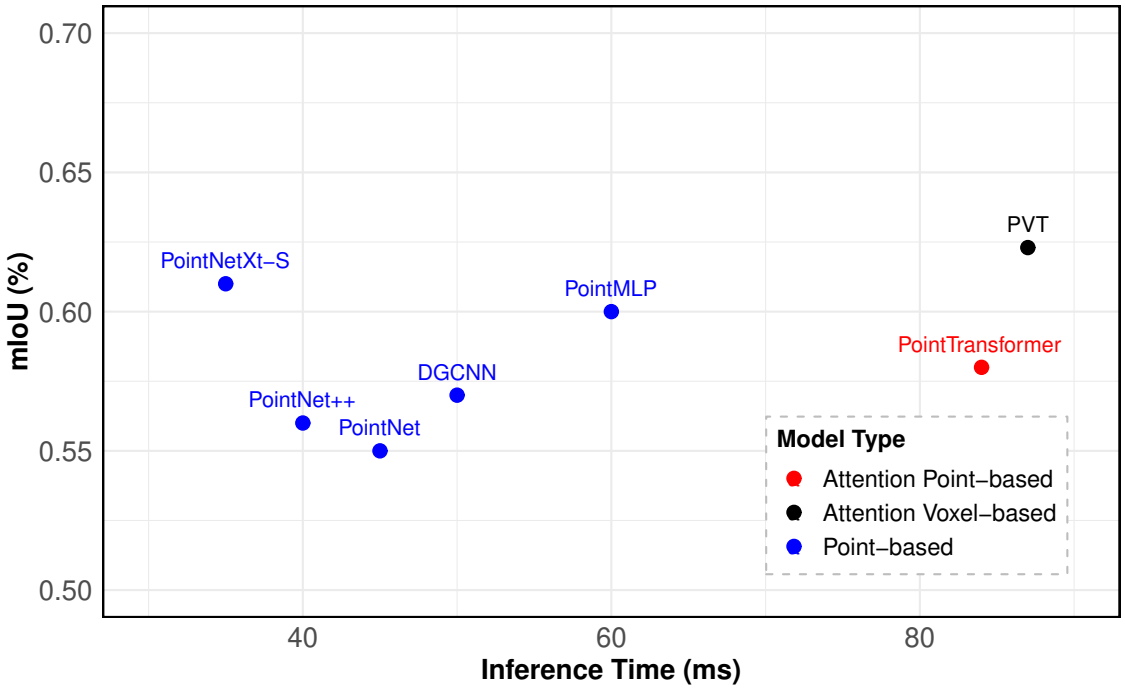
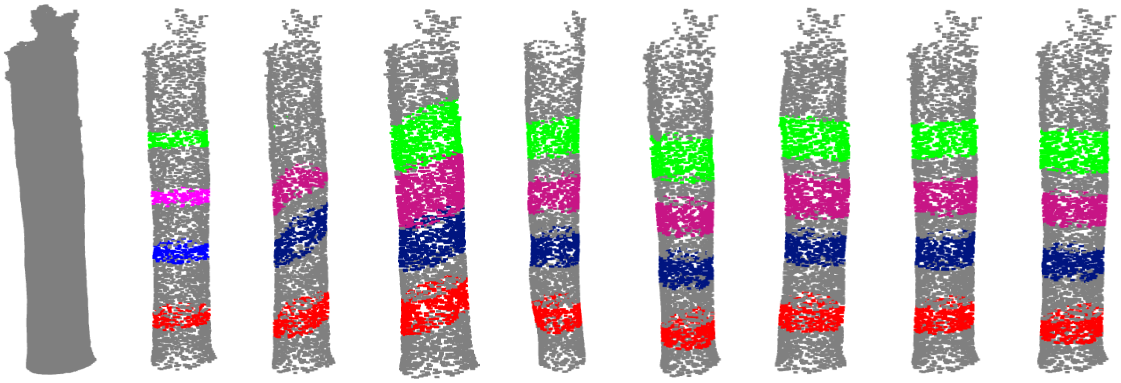


Figure 5.2: *Comparison of partial segmentation models for eucalyptus, highlighting the trade-off between mIoU performance and inference time: point-based models deliver faster processing speeds, while attention-based models achieve greater accuracy.*

The PointNeXt-S model demonstrated exceptional performance, achieving results comparable to those of transformer-based models while maintaining the shortest inference time among all evaluated models. This comprehensive analysis highlights the strengths and identifies potential areas for improvement in these models, providing valuable guidance for future research and development efforts in point cloud segmentation within forestry contexts.



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 5.3: *Visualization of segmented eucalyptus trees using different models. (a) Input point cloud, (b) Ground Truth, (c) PointNet, (d) PointNet++, (e) DGCNN, (f) PointNeXt-S, (g) PointMLP, (h) Point Transformer, and (i) PVT.*

5.5 Conclusions and Future Work

In this chapter, we present a comparative analysis of five distinct deep learning models for segmenting point clouds of eucalyptus trees. The performance, measured by mIoU, was similar across the models but lower compared to classical benchmarks, likely due to the smaller dataset. Despite this, the models demonstrated strong performance within these constraints.

The qualitative results show that the models are efficient at segmenting diameter regions, crucial for measuring variables in forest environments. This efficiency highlights their potential to provide valuable insights for forestry applications, even with limited data. Additionally, the models' processing times indicate they can run on computers with lower computational power, making them suitable for field deployment where high-performance resources may not be available.

Future work will focus on expanding the dataset to validate model performance further, implementing advanced data augmentation techniques to enhance robustness, and conducting a comprehensive statistical analysis to compare actual diameters with those derived from segmentation, offering deeper insights into the models' accuracy and reliability in practical applications.

CHAPTER 6

Accurate Measurement of Eucalyptus Tree Diameters Using Smartphone LiDAR and Deep Learning

> The greatest enemy of knowledge is the illusion of it.
>
> *Stephen Hawking*

Accurate measurement of tree structural characteristics, such as height and diameter, is crucial for forest inventories. Dendrometric variables like DBH, basal area, and volume are fundamental. However, accurately obtaining diameters traditionally requires significant resources due to manual measurement in extensive sampling areas. In this context, LiDAR sensors are a promising technology that uses laser pulses for three-dimensional mapping, contributing to more efficient and sustainable forest inventories. This study explores the use of smartphone-integrated LiDAR technology combined with deep learning models to efficiently measure the DBH of *Eucalyptus sp*. trees. Traditional methods for obtaining accurate measurements in forestry require significant manual labor and resources. Here, we propose an efficient pipeline for tree detection, segmentation, and diameter assessment using Faster R-CNN and PointNet applied to LiDAR-generated point clouds. Our results demonstrate that the proposed method can achieve high-precision diameter measurements with *RMSE* below 1 cm and no statistically significant differences compared to traditional methods. To support further research and validation, we provide a publicly available dataset comprising field measurements and corresponding LiDAR point clouds. These findings position smartphone LiDAR technology as a cost-effective alternative for forest inventory and commercial forestry operations. This chapter has been submitted to the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS).

6.1 Introduction

The Brazilian forestry industry is internationally recognized for its remarkable productivity, particularly noted for the reduced time interval between planting and harvesting. In 2023, the total area dedicated to tree planting in Brazil surpassed 10 million hectares for the first time, marking a 3% increase compared to the previous year. Among the various types of plantations, eucalyptus stands out, covering 7.8 million hectares, which accounts for 76% of the total planted area. This represents a 41% growth over the past ten years, underscoring the increasing significance of eucalyptus in the Brazilian forestry economy (IBÁ, 2024).

Precise quantification of tree structural attributes, including height, diameter, crown width, and taper, is imperative for adequate forest inventories in managed forests and commercial plantations. Consequently, it is crucial to obtain accurate measurements. Achieving precise measurements requires collecting data from representative forest samples, which facilitates a comprehensive understanding of population dynamics (SOARES *et al.*, 2011).

The forest inventory represents a meticulous process that encompasses collecting, analyzing, and interpreting data relating to a specific forest area. In areas of planted forests, it mainly focuses on spaces where trees have been strategically planted for commercial, industrial, or production purposes (HAKAMADA *et al.*, 2015). This procedure plays an essential role in monitoring the growth and development of planted trees, calculating the amount of biomass and wood available, predicting production yield, and making appropriate management decisions.

Essential metrics for forest inventories encompass structural aspects such as tree height, diameter at breast height, stem basal area and volume, and stem geometric properties such as taper. Traditionally, such characteristics are measured at the individual level, in broad sampling areas, through manual field procedures.

The predominant dendrometric variables for conducting forest inventories are height and diameter, whose information can be used to derive additional measurements, such as calculations of the basal area and volume of wood present in a forest ecosystem (FREITAS; WICHERT, 1998). However, obtaining diameter measurements at different heights in the tree requires considerable time and resources (GAFFREY *et al.*, 2001).

In this context, the advent of technologies that help generate accurate inventories is essential for the forestry sector. Among these technologies, LiDAR stands out. Light Detection and Ranging (LiDAR) is a sensor that measures distances and generates three-dimensional information about objects and surfaces using laser pulses. LiDAR is frequently used in various applications, including topographic mapping, environmental monitoring, remote sensing, vehicle autonomy, and other applications (DUBAYAH;

DRAKE, 2000).

These sensors measure depth by projecting a laser pulse onto a scene and calculating the time it takes for the pulse to return. In time-of-flight (ToF) cameras, a specialization of this technology, the laser pulse is divided into a dense array of thousands of points, measuring an entire scene in a single scan and achieving a resolution and frame rate comparable to small RGB cameras (ZEA; HANEBECK, 2022).

In forest mapping, LiDAR offers substantial insights into trees and forest areas. Its application is often aimed at generating digital terrain models, identifying trees, estimating wood volumes, and analyzing forest structure, among other applications (NGUYEN *et al.*, 2021).

This study introduces and validates a novel methodology for accurately measuring the DBH of *Eucalyptus sp.* trees using smartphone-integrated LiDAR sensors. Central to this methodology is a specialized deep learning pipeline, engineered to overcome the challenges of processing smartphone LiDAR data for precise forest inventory tasks. The main contributions of this research include:

- An Integrated 2D-3D Deep Learning Framework for Hierarchical Tree Analysis: We developed and implemented an innovative framework that synergistically integrates 2D and 3D deep learning architectures for robust, multi-stage tree analysis from point cloud data. This hierarchical approach involves:
  - *Efficient Tree Detection:* A Faster R-CNN model is applied to 2D Bird-Eye View (BEV) projections, derived from the 3D LiDAR point clouds, to achieve efficient and accurate detection of individual trees.
  - *Precise Stem Segmentation:* Following detection, a PointNet-based neural network performs fine-grained semantic segmentation directly on the 3D point clouds of the identified tree stems.
- Advanced Training Strategy for Robust Segmentation with Imbalanced Data: To address the significant class imbalance inherent in forestry point cloud datasets, our PointNet segmentation model is trained using a hybrid loss function. This function combines weighted Focal Loss and Dice Loss, to optimize segmentation overlap (Intersection over Union - IoU).
- A Novel Algorithm for Diameter Derivation from 3D Semantic Segments: We introduce a novel algorithm to derive DBH and other key tree diameters. This algorithm directly processes the 3D coordinates of points within the specifically segmented regions (e.g., the DBH band) to calculate robust and accurate diametric measurements.

This work is structured as follows. Section 6.2 describes the results found in the literature regarding applying LiDAR in the forestry sector. Section 6.3 addresses the

methodology adopted in this research. Section 6.4 covers computational resources and metrics. Section 6.5 presents the results obtained with the proposed method and compares them with field data. Section 6.6 discusses the results achieved and the feasibility of the application. Finally, Section 6.7 presents and discusses the conclusions derived from this study.

6.2 Literature Review

Accurately quantifying forest structural metrics from Airborne Laser Scanning (ALS) data is challenging due to factors like scan angles, low point density, and canopy obstruction. Therefore, methods have been developed to establish empirical relationships between LiDAR-derived variables, such as canopy height, and field measurements (HYYPPA *et al.*, 2001; MALTAMO *et al.*, 2007).

ALS approaches represent a significant portion of studies utilizing LiDAR in forestry. For instance, citelee2018machine adopted traditional machine learning approaches, such as decision trees and Support Vector Machines, for tree height estimation. In their study, Esteban *et al.* (2020b) based canopy height estimation on digital canopy height models (CHM) obtained from LiDAR data and texture analysis. Similarly, (SHOOT *et al.*, 2021) employed several machine learning algorithms to classify forest types based on LiDAR readings.

Determining the total and commercial volume of trees is essential for environmental licensing and forest management planning. Forest features can be extracted from LiDAR data obtained through aerial scans and satellite imagery. Height estimation, combined with other variables, can be used to derive characteristics such as volume and basal area (WITTKE *et al.*, 2019; HUDAK *et al.*, 2020; BOLTON *et al.*, 2020; OZKAN *et al.*, 2022). Different machine learning approaches have been explored for estimating height and volume (CORTE *et al.*, 2020b), while other studies have focused on predicting volume and diameter (PERSSON *et al.*, 2022; OZKAN *et al.*, 2022).

Forest biomass is an increasingly relevant energy source due to its effectiveness in reducing greenhouse gas emissions. A carbon monitoring system has been proposed to map above-ground biomass (HUDAK *et al.*, 2020). Deep learning approaches have been employed to estimate forest biomass (WITTKE *et al.*, 2019; AYREY; HAYES, 2018). Additionally, various studies have explored the use of LiDAR and satellite imagery to assess biomass potential (MAHONEY *et al.*, 2018; CAO *et al.*, 2019; HOOVER *et al.*, 2018; BEAUDOIN *et al.*, 2022; EHLERS *et al.*, 2022).

The Terrestrial Laser Scanner (TLS) can provide detailed information about specific trees or forest areas. In recent decades, several applications have been designed to explore the advantages of TLS, replacing conventional methods of measuring attributes

in forest inventories. Unlike ALS, the approach using TLS, for the most part, requires manual operation of the LiDAR equipment by a professional in the field.

Diameter is a fundamental dendrometric parameter essential for calculating various forest metrics, making it a key variable in forest inventories due to its accessibility and strong correlation with other attributes. Diameter estimation has been performed using the TLS approach (HEINZEL; GINZLER, 2018), while 3D models have been explored for diameter prediction (NGUYEN *et al.*, 2021; CHEN *et al.*, 2020a). Tree identification and diameter estimation have also been addressed in several studies (ZHOU *et al.*, 2017; SUN *et al.*, 2022; HAO *et al.*, 2020). The TLS approach is generally more efficient than ALS, as it overcomes limitations related to scan angles and canopy obstruction, which can reduce correlation with field-measured data (THIES; SPIECKER, 2004).

Neural networks have been widely used in forest inventory tasks, demonstrating their versatility in addressing various challenges. Noteworthy examples include the utilization of RetinaNet on ALS LiDAR data to perform tree detection and segmentation (ALON *et al.*, 2019). An end-to-end pipeline for tree diameter estimation, integrating semantic segmentation, odometry, and LiDAR mapping techniques, has also been proposed (CHEN *et al.*, 2020a). Additionally, Convolutional Neural Networks (CNNs) have been applied to classify species using hyperspectral images and LiDAR data (MAYRA *et al.*, 2021). CNNs have further been employed to interpret LiDAR data and estimate various forest attributes, illustrating the broad applicability of neural network-based approaches (AYREY; HAYES, 2018).

CNNs and Three-Dimensional Convolutional Neural Network (3D-CNN) have been applied to LiDAR point clouds for tree detection and segmentation (WINDRIM; BRYSON, 2019). Later, a method to detect, segment, and extract individual tree stem features using the PointNet Segmentation architecture on LiDAR data was introduced (WINDRIM; BRYSON, 2020b).

In this study, we employ various deep learning architectures to identify and segment eucalyptus trees in three-dimensional point clouds. Unlike previous studies, which primarily focus on using CNNs, 3D-CNNs, or traditional machine learning methods, our approach integrates advanced techniques such as Faster R-CNN for object detection and PointNet for segmentation. Additionally, we introduce a novel method to calculate the segmented DBH using a model trained with PointNet. This comprehensive pipeline not only enhances the accuracy of tree detection and segmentation but also provides precise DBH measurements, offering a robust and scalable solution for forest inventory and management.

6.3 Materials and Methods

6.3.1 Materials

This subsection outlines the materials employed in our experiments, including the study area and dataset.

Study Area

The study area corresponds to *Eucalyptus sp.* plantations located on a farm in the municipality of Açailândia, Maranhão, Brazil. The region is characterized by a hot and humid climate for approximately half of the year, with significant rainfall occurring mainly between November and May.

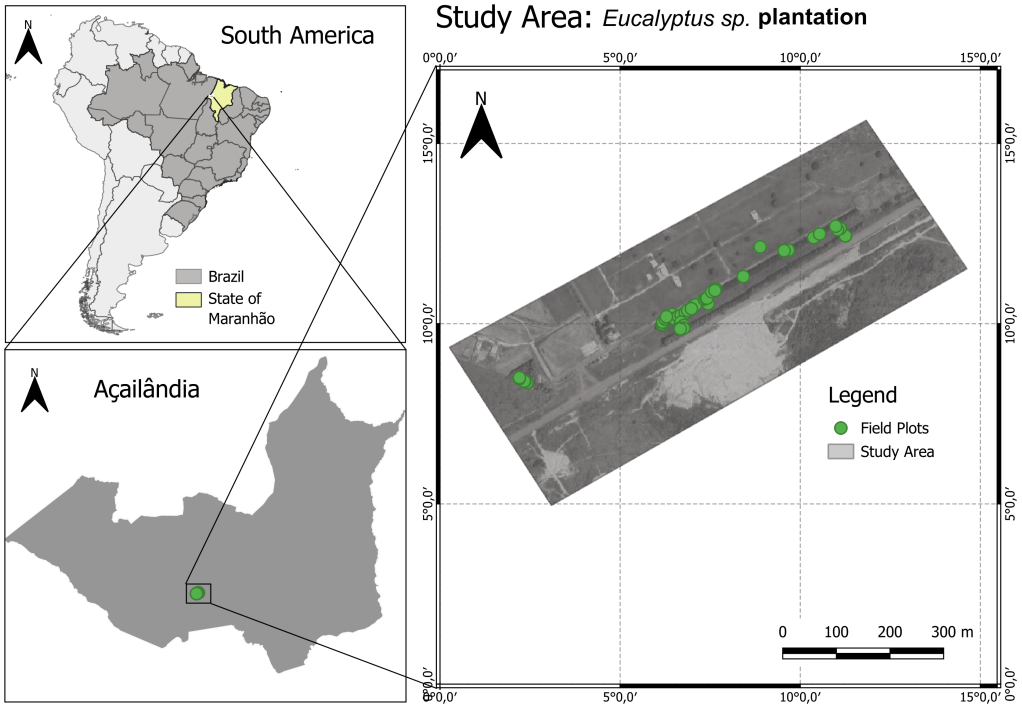


Figure 6.1: *Location of study area in Açailândia, Maranhão, Brazil. The green dotes indicate the location of the Eucaliptus sp. trees.*

Dataset

The dataset comprises dendrometric measurements and corresponding LiDAR point clouds for several eucalyptus trees, collected over approximately one month. To

support reproducibility and further research, the complete dataset is publicly available for download[1].

Measurements were collected from 55 eucalyptus trees of the clonal type of the genus Eucalyptus sp. in June 2023. All trees were georeferenced with GPS (Garmin 76csx), ensuring a location error of less than 10 meters. For each tree, diameter measurements were taken at four different heights: 0.30 m, 0.70 m, 1 m, and the DBH diameter at 1.3 m from the ground. Measurements were carried out using a caliper and diameter tape. Additionally, the total height of a random subsample (20%) of trees was measured.

The trees were categorized into four classes based on DBH, with an interval of 5 centimeters: dominated, intermediate, codominant, and dominant. Dominant trees (D) are the tallest and most developed, with crowns that extend above the general canopy, receiving both direct and lateral light. Codominant trees (C) are of average size within the stand, receiving direct light from above but limited lateral exposure. Intermediate trees (I) grow between dominant and codominant trees, receiving only partial light at the crown tip, leading to smaller crowns. Suppressed or dominated trees (S) are the smallest, receiving minimal to no direct light, resulting in significantly reduced crown development. The statistics are presented in Table 6.1.

Table 6.1: *Dataset statistics, including DBH limits, average, and variability.*

Class	Limits (DBH)		Avg[1]	Sd[2]	Range (ht)
	Inf.	Sup.			
Dominated	-	12	10.38	1.53	7.0 - 11.9
Intermediaries	12	17	14.98	1.37	12.8 – 16.9
Codominant	17	22	19.90	1.49	17.0 – 21.5
Dominants	22	27	24.43	1.64	22.1 – 27.0

[1] Average.
[2] Standard Deviation.

LiDAR point clouds were acquired using the LiDAR sensor integrated into the iPhone 12 Pro Max. The scanner settings, optimized for the data acquisition procedure of capturing points within a radius of up to 2 meters around each tree, included an effective operational range of 2 meters and a resolution of 5 millimeters, resulting in a 3D reconstruction error of approximately ±5 mm. The point cloud acquisition procedure involved rotating the smartphone around each tree to capture all corresponding points within a 2-meter radius from the stem and up to the tree's average height of 2 meters.

[1]Dataset available in Google Drive: <https://drive.google.com/drive/folders/1VibXd6UdRrArYJG9hoTHMUBvUz6AMCEp?usp=sharing>

Direct access to raw LiDAR depth data via general software libraries is often restricted by the device's operating system; therefore, frameworks like Apple's ARKit are typically used to process this information and generate depth estimates (APPLE, 2021). This framework is designed to process the information from the LiDAR scanner and the RGB camera, performing a fusion of this data to generate a depth estimate. Additionally, ARKit provides additional information, such as a real-time mesh reconstruction and a point cloud based on identified RGB features (ZEA; HANEBECK, 2022). The 3D Scanner App was utilized to acquire the point cloud data used in this study.

The main limitation of the LiDAR sensor used in this study is its restricted operating range, which covers areas with an average range of approximately 5 meters (LUETZENBURG *et al.*, 2021). Furthermore, this type of sensor does not have intensity, return, and classification data, which are characteristics of professional sensors such as those in the FARO Focus line. Consequently, the data resulting from a scan performed with this device was limited to the three-dimensional location of the point cloud (*xyz* coordinates).

6.3.2 Methods

This methodology was developed to identify individual trees in a LiDAR point cloud and estimate stem diameters at multiple heights. To provide a rigorous validation, the quantitative analysis herein focuses exclusively on the DBH, the most fundamental measurement in forest inventories. The developed pipeline is, however, capable of measuring all other annotated stem sections without modification.

Overview

In order to accomplish this goal, a multi-stage pipeline is implemented, consisting of individual tree detection, soil characterization and removal, and segmentation of points of interest on the tree stem. Following these procedures, the initial diameters of each tree can be inferred using a pre-trained model. Figure 6.2 illustrates an overview of the steps outlined in this study.

Tree Detection

In this work, identification is performed using a common approach for object detection in point clouds in which a 3-channel 2D image is generated through a panoramic projection known as BEV (i.e., projected onto the xy-plane) of the 3D point cloud (CHEN *et al.*, 2020b; WANG *et al.*, 2023). After that, a Faster R-CNN object detector delimits the region corresponding to the tree in the 2D projected image.

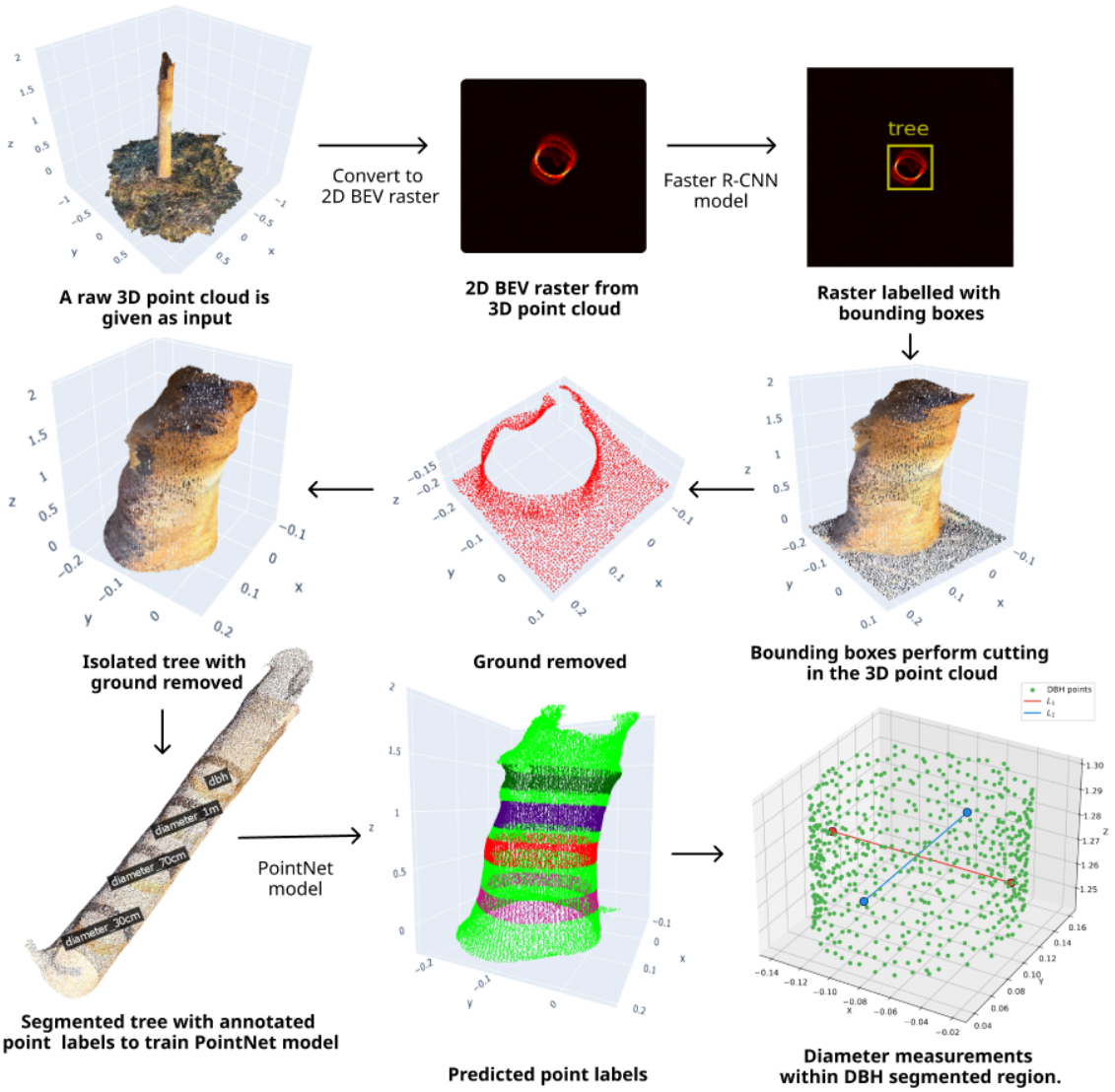


Figure 6.2: *Pipeline for obtaining tree diameter measurements.*

The BEV projection encodes the 3D information of a point cloud in a 2D representation. In this study, the scanned grid size was approximately $4 \times 4 \times 2$ meters. The resulting data is mapped to a 3-channel, producing a $256 \times 256 \times 3$ color image.

The strategy of projecting the point cloud into BEV and using a Faster R-CNN object detector proved effective due to the predominantly cylindrical structure of trees, which facilitates a clear representation in two dimensions. This approach enhances object detection efficiency by reducing variability among objects similar to trees. Once identified, individual trees are delineated by projecting the detected bounding box coordinates from the xy-plane back onto the 3D point cloud, effectively isolating points corresponding to each tree.

Ground Removal

Unlike specialized LiDAR ALS or TLS sensors that provide detailed information such as intensity, reflectance, and classification, the iPhone 12 Pro Max sensor is limited to *xyz* coordinates and RGB channels corresponding to each point. After identifying the tree using Faster R-CNN, soil removal is simplified as the initial delimitation isolates much of the soil. A Digital Terrain Model (DTM) is then constructed to remove the remaining soil points.

To construct the DTM, the point cloud is discretized into 20×20-centimeter bins within the *xy*-plane, preserving the minimum height (z coordinate) within each bin. A grid with 1×1-meter cells is then generated to cover the entire point cloud. The KD-Tree data structure (BENTLEY, 1975) is employed to find the four nearest points (using Euclidean distance in the *xy*-plane) from the previously stored subset of heights to the center of each grid cell. The average height for each *xy* location on the grid is calculated by weighting these four points according to their distance from the grid cell center. Once the DTM is estimated, ground points are removed by discarding all points within a specified height limit above the DTM. In our experiments, we used a height threshold of 10 cm above the ground.

Tree Detection Training and Inference

The three-dimensional point clouds obtained for each tree were converted into BEV images for the Faster R-CNN training process. Each image was annotated, and axis-aligned bounding boxes were created using the LabelStudio software (TKACHENKO *et al.*, 2022). The annotations were categorized into two classes: “ground”, representing the soil, and “tree”. The inclusion of the “ground” class served to reduce false-positive detections of trees. The BEV images, along with their corresponding bounding box labels, were then used to train the Faster R-CNN object detector (REN *et al.*, 2015).

The Faster R-CNN model was implemented using a ResNet-50 backend, previously trained on the MS COCO dataset (HE *et al.*, 2016). Hyperparameters were tuned empirically to optimize detection performance. We employed the Stochastic Gradient Descent (SGD) optimizer, arriving at a final learning rate of 0.001, a momentum of 0.9, and a mini-batch size of 8. The model was trained for 100 epochs to ensure convergence

In the inference step, the point cloud of a tree, including the ground, is transformed into a BEV image. Tree detection in the BEV projection is then performed, identifying it through bouding boxes and associating it with the corresponding class, using the previously trained Faster R-CNN model. The “ground” class is excluded, resulting in only the region that circumscribes the tree. This procedure significantly contributes to the

substantial removal of soil, reducing the computational load during the construction of the DTM.

The bounding box coordinates, initially in the local coordinates of the BEV projection, are later converted to local coordinates in the point cloud. Finally, the 2D bounding boxes are projected onto 3D cuboids, covering all points whose x and y coordinates are contained in the bounding boxes, thus isolating the representation three-dimensional tree.

Stem Segmentation

After isolating and detecting the tree, it is returned to a set of three-dimensional points for subsequent segmentation. In this context, a specific deep learning architecture was used for segmenting 3D point clouds, called PointNet (QI *et al.*, 2017a). This approach uses the direct representation of raw data points to train a neural network capable of inferring labels.

Training this network involved manually annotating point clouds corresponding to individual trees and reducing the number of points to a maximum size of 2048 points. This process was carried out using the CloudCompare software tool (GIRARDEAU-MONTAUT, 2016). For each tree, the points were labeled about the following classes: *stem, diameter-30cm, diameter-70cm, diameter-1m* and *dbh*. Figure 6.3 illustrates an example of a tree stem point cloud following this manual annotation process, showcasing the labeled segments used to train the PointNet model for identifying these key regions.

Stem Segmentation Training and Inference

The approach adopted in this work is based on the famous PointNet architecture, a neural network style designed for points. It is flexible enough to be used for classification and segmentation. This study is used to map points from individual trees to class labels mentioned above, being a segmentation task. The network generates a score for each of the five classes for all points.

A significant disparity is evident in the data, revealing a pronounced imbalance in which the class associated with the tree stem exhibits a substantially higher representation compared to other categories. Nevertheless, it is important to emphasize that the classes corresponding to regions previously segmented as diameters are more critical than the general stem class, as these segments will be directly employed in diameter measurements.

When using a standard loss function such as Cross Entropy, it inherently assumes that all classes are of equal importance during training. This implies that an error in classifying a segmented diameter is treated equivalently to an error in classifying the

Figure 6.3: *Manually annotated LiDAR point cloud of a tree stem segment, showing regions explicitly labeled for diameter measurements (DBH, 1m, 70cm, and 30cm). The remaining points on the stem segment constitute the 'stem' class, according to the annotation scheme used for training the deep learning model.*

stem. However, in this context, the model should prioritize the diametric categories over the stem, as the former are more critical for the intended diameter measurements.

To address the issue of class imbalance, an effective strategy is to adopt a weighted loss function. This approach adjusts the model's learning focus, emphasizing classes that are underrepresented or more challenging to classify, while reducing attention to classes that are easier to predict correctly.

This study employs a loss function based on Focal Loss (LIN *et al.*, 2017), an extension of the Cross Entropy loss designed specifically for unbalanced classification scenarios. Focal Loss works by directing the model's training efforts toward harder-to-classify examples, thereby mitigating the effects of class disparity and improving performance. Additionally, considering the segmentation nature of the task, Dice Loss is also incorporated to further enhance the model's performance, particularly in boosting the IoU (SUDRE *et al.*, 2017).

Although the network architecture can support point clouds with different numbers of points, computational constraints require the same number of points in the point clouds in a given batch. In this work, the point clouds had a fixed size of 2048 points. These points have been uniformly normalized to a unit sphere. This normalization pro-

cedure in a sphere ensures that all points are evenly distributed on the surface of a unit sphere, thus mitigating distortions in the analysis and facilitating an accurate representation of the data in a three-dimensional space (QI *et al.*, 2017a).

The PointNet architecture was trained from scratch. For training, we utilized the Adam optimizer, as it consistently yielded stable convergence in initial tests. Key hyperparameters were selected after preliminary experiments: the learning rate was set to 0.0001 with no weight decay, and the network was trained for 2500 epochs. During training, input point clouds were subsampled to a fixed size of 2048 points. For our hybrid loss function, the α parameter for the Focal Loss was set to 0.1 for the 'stem' class and 0.9 for the 'diameter' classes to counteract the severe class imbalance.

During inference, point reduction is unnecessary for tree segmentation so that the model can deal with a cloud with thousands of points. Therefore, after the tree is identified and the soil is removed, it is submitted to the model trained by the PointNet. Then, the classes of interest are segmented to measure the diameter.

Diameter Measurement

After the segmentation of classes, the next step involves measuring the DBH. This process begins by calculating the average z-axis value within the point cloud corresponding to the segmented region.

Once the average z-coordinate (average height) of the points within the segmented DBH region is determined, two linear measurements, L_1 and L_2, are computed to estimate the diameter. L_1 is defined as the 3D Euclidean distance between the specific pair of points within this segmented region that have the overall maximum and minimum y-coordinates, respectively. Similarly, L_2 is the 3D Euclidean distance between the pair of points within the region that possess the overall maximum and minimum x-coordinates.Figure 6.4 provides an illustration of the procedure used to measure the segmented region.

Given two points, $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$, the length of L_1 can be determined using the Euclidean distance formula:

$$L_1 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \tag{6-1}$$

For the second line L_2, defined by the points $Q_1(x_3, y_3, z_3)$ and $Q_2(x_4, y_4, z_4)$, the length is computed similarly:

$$L_2 = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2 + (z_4 - z_3)^2}. \tag{6-2}$$

Consequently, the lengths of L_1 and L_2 are determined. The diameter of the

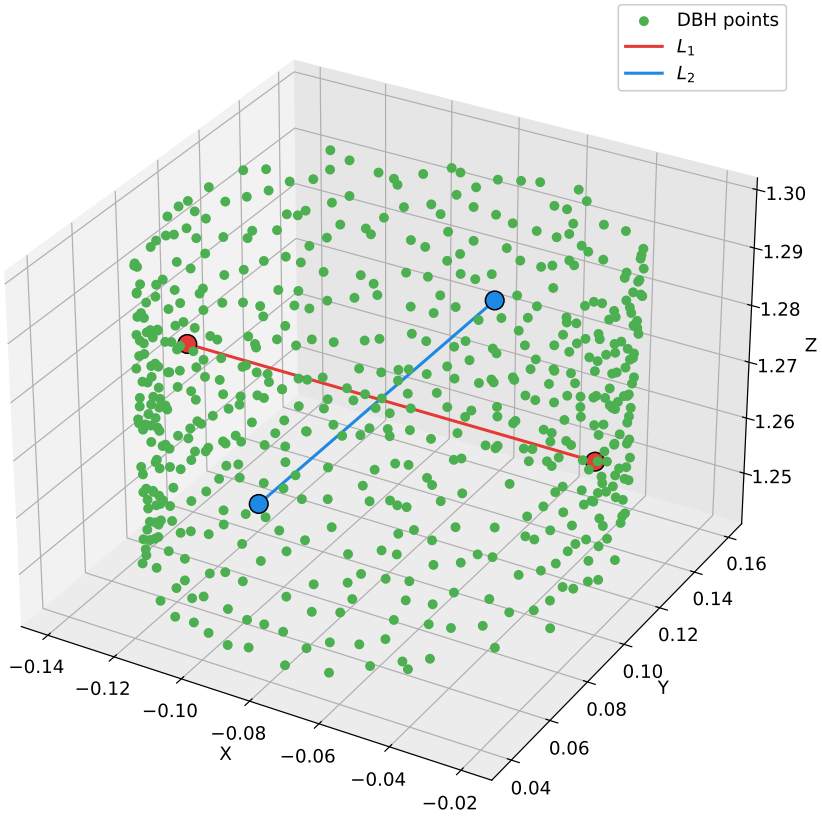


Figure 6.4: *Process for measuring the segmented region corresponding to the DBH.*

segmented region is then defined as the arithmetic mean of these two distances, as expressed in the following equation:

$$D = \frac{L_1 + L_2}{2}, \tag{6-3}$$

where D represents the estimated diameter. This computation provides the average diametric measure obtained along the DBH region. This computation provides an average diametric measure for the analyzed stem section. The algorithm detailing these steps is presented in Algorithm 1.

6.4 Experimental Setup

In training the R-CNN detector, the 55 point clouds corresponding to the scanned trees were converted into images. Each tree was annotated with bounding boxes, as described in Section 6.3. Individual trees were extracted and annotated at the point level for segmentation experiments, and bounding boxes for detection experiments.

The dataset was randomly divided into 80% for training and 20% for testing. Due to the limited number of samples in the dataset, cross-validation (k-fold) with $k = 3$

Algorithm 1 Estimate Diameter of Segmented Region

procedure CALCULATEDIAMETER(P_1, P_2, Q_1, Q_2)

▷ Calculate the length of L_1 using Euclidean distance

$\Delta x_1 \leftarrow x_2 - x_1$

$\Delta y_1 \leftarrow y_2 - y_1$

$\Delta z_1 \leftarrow z_2 - z_1$

$L_1 \leftarrow \sqrt{(\Delta x_1)^2 + (\Delta y_1)^2 + (\Delta z_1)^2}$

▷ Calculate the length of L_2 using Euclidean distance

$\Delta x_2 \leftarrow x_4 - x_3$

$\Delta y_2 \leftarrow y_4 - y_3$

$\Delta z_2 \leftarrow z_4 - z_3$

$L_2 \leftarrow \sqrt{(\Delta x_2)^2 + (\Delta y_2)^2 + (\Delta z_2)^2}$

▷ Calculate the estimated diameter D as the arithmetic mean of L_1 and L_2

$D \leftarrow \frac{L_1 + L_2}{2}$

return D

was employed to evaluate the detection and segmentation methods.

To ensure a fair and robust evaluation, the hyperparameters for all deep learning models were carefully tuned. The final selected parameters, determined through a series of preliminary experiments aimed at optimizing performance on a validation subset of the data, are detailed in the respective methodology sections below and summarized in Table 6.2.

The experiments were conducted on a 64-bit desktop equipped with an AMD Ryzen 9 5900X processor (12 cores and 24 threads), 32 GB of RAM, and an Nvidia GeForce RTX 3070 GPU. The neural network models for detection and segmentation were implemented in Python using the PyTorch framework.

6.4.1 Metrics

To assess the performance of tree detection, we employed fundamental metrics, including precision, recall, F1 score, and mean Average Precision (mAP). The mAP, recognized as a standard metric for evaluating object detection models, is widely acknowledged in the literature (ZHIQIANG; JUN, 2017). In this context, a correct detection is determined when the bounding boxes associated with an actual region and those predicted by the model exhibit a higher IoU overlap of 50% or more (FLETCHER *et al.*, 2018). In contrast, detection is classified as incorrect if this criterion is not achieved.

The metrics used to evaluate the accuracy of three-dimensional tree segmentation consisted of accuracy, the Matthews Correlation Coefficient (MCC) index (YAO; SHEP-

Table 6.2: *Selected Hyperparameters for the Deep Learning Models*

Model	Hyperparameter	Value
Faster R-CNN		
	Backend	ResNet-50
	Optimizer	SGD
	Learning Rate	0.001
	Momentum	0.9
	Batch Size	8
	Epochs	100
PointNet		
	Optimizer	Adam
	Learning Rate	0.0001
	Weight Decay	0
	Epochs	2500
	Input Points (Training)	2048
	Focal Loss α (Stem)	0.1
	Focal Loss α (Diameter)	0.9

PERD, 2020), and IoU for each specific class. Such metrics were computed by converting the point clouds noted as true and the predicted point clouds, followed by comparing them.

Root Mean Square Error (RMSE) is a method for evaluating accuracy in DBH prediction. It quantifies the average magnitude of errors between predicted and observed DBH values, thereby assessing the model's accuracy. Additionally, Mean Absolute Percentage Error (MAPE) is utilized to gauge the mean absolute percentage error generated by a model.

A statistical analysis uses the paired Student's t-test to evaluate the predictive performance. The Student t-test requires that the data set follows a normal distribution. Therefore, the Shapiro-Wilk test was applied to evaluate whether the distribution is similar to a normal distribution.

6.5 Results

6.5.1 Individual Tree Detection

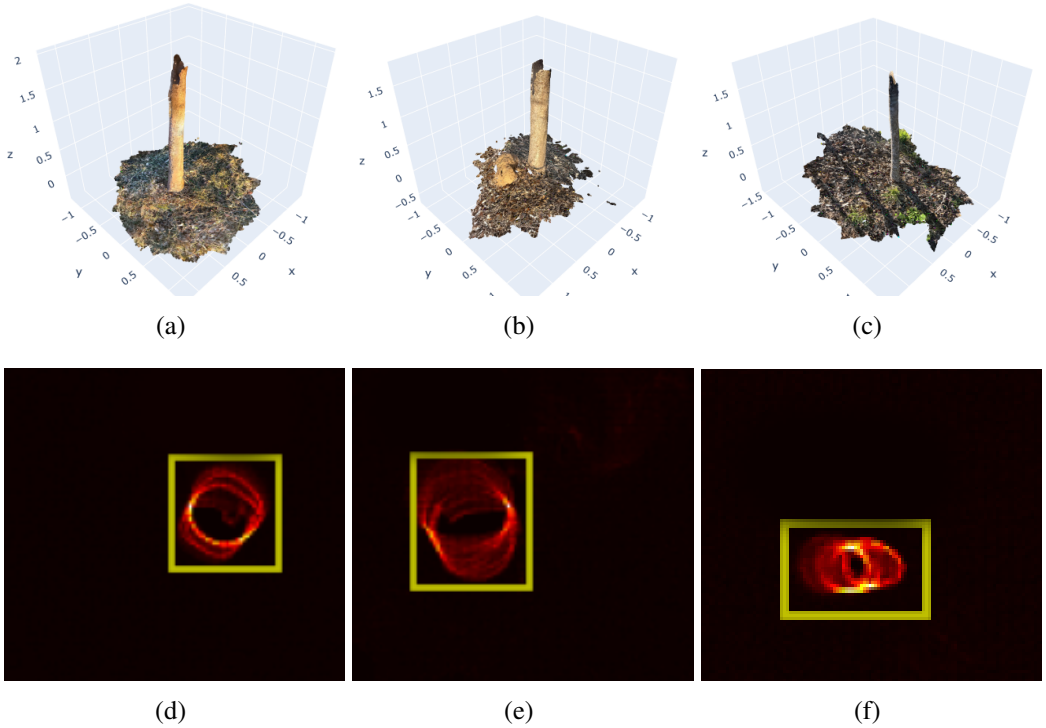
Table 6.3 presents the metrics related to the detection of individual trees achieved in this study. Throughout the experiments, R-CNN-based object detection methods exhibited a low incidence of false positives (e.g., bushes misidentified as trees). Notably, at confidence levels above 0.85, the precision reaches its maximum value, indicating the absence of false positives. This high performance can be attributed to the use of a pre-

trained ResNet50 model and the significant disparities between the classes (ground and tree) in the image dataset, which contribute to the maximization of metrics in each *k*-fold iteration.

Figure 6.5 qualitatively illustrates the detection results, showcasing the transformation from raw point clouds to the corresponding BEV images. Additionally, the figure displays the R-CNN bounding boxes and the contours derived from the points, along with the ground reference contours.

Table 6.3: *Individual tree detection cross-validation results on the validation set.*

k-fold	Precision	Recall	mAP (50)	mAP (50–95)
1	0.9941	1.0000	0.980	0.660
2	1.0000	1.0000	0.995	0.665
3	0.9971	1.0000	0.995	0.668
Avg.	0.9970	1.0000	0.990	0.664



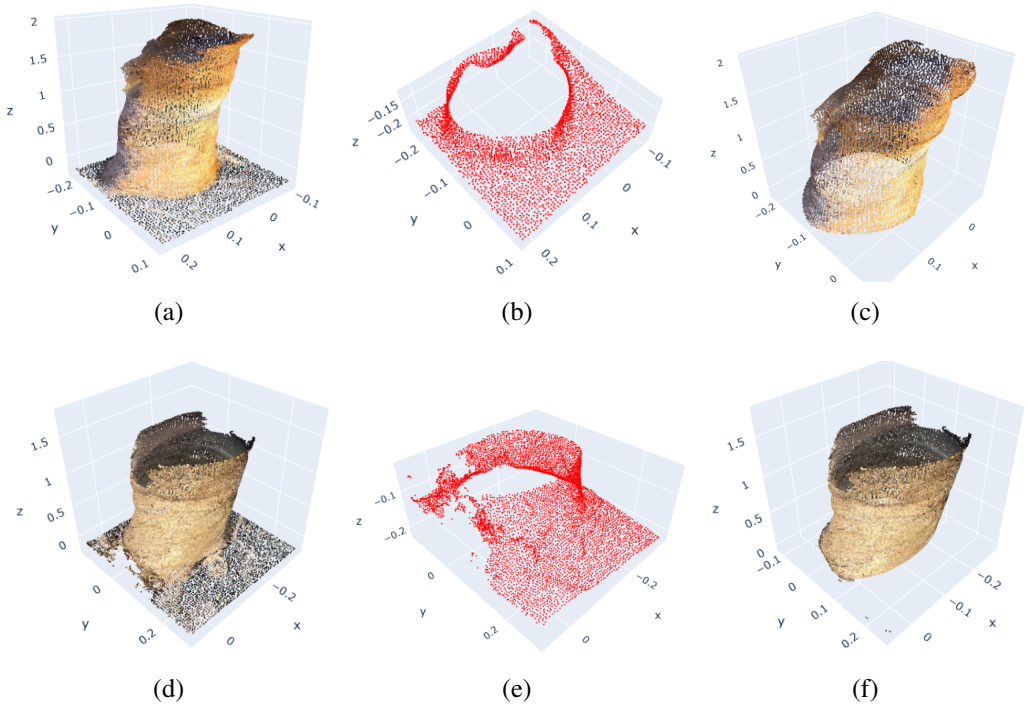
(a) (b) (c)

(d) (e) (f)

Figure 6.5: *Figures **(a)** to **(c)** depict the raw point cloud obtained using the iPhone 12 Pro Max's LiDAR sensor. Images **(d)** to **(f)** represent the same trees after the object detection process using the Faster R-CNN method, with the bounding boxes presented in their respective projections in BEV.*

6.5.2 Ground Removal

Figure 6.6 shows examples of DTMs extracted from point clouds and point clouds after the tree identification and delimitation process with ground points removed. The point clouds have terrain with different slopes from the ground, and, in each case, the shape of the DTM resembles the terrain.



(a) (b) (c)

(d) (e) (f)

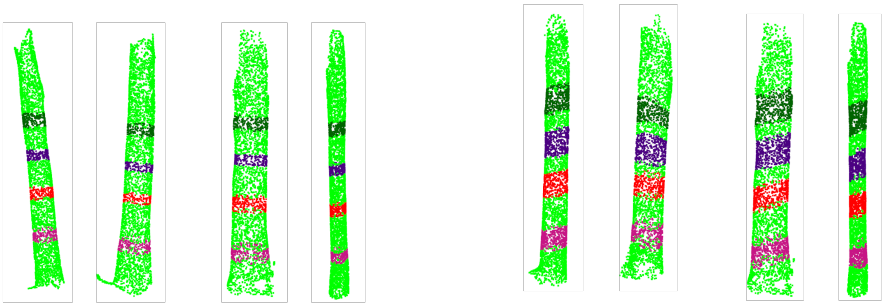
Figure 6.6: *Qualitative results of soil removal. The process illustrated above shows the tree in its already delimited point cloud, followed by the region created by the DTM and subsequently the point cloud with the isolated tree.*

6.5.3 Stem Segmentation

Results for stem segmentation were computed at the tree level. For each iteration of the cross-validation method, the average segmentation accuracy was calculated for each dataset. These values were then averaged across the three folds, resulting in accuracy = 0.33, MCC = 0.47, and IoU = 0.56.

Figure 6.7 shows qualitative examples of the segmentation, displaying the regions corresponding to the classes correctly identified by the proposed method. These visualizations demonstrate the model's ability to accurately distinguish different components of the tree while preserving the structural characteristics of the point clouds. Additionally, Figure 6.8 presents the confusion matrix that summarizes the classification per-

formance of the segmented points, highlighting the agreement between the actual classes and those assigned by the model.



(a) *Ground truth* (b) *Predicted*

Figure 6.7: *Visualization of the segmentation results of selected trees, with comparison with the ground truth. It can be seen that there is a slight divergence between the segmented regions **(a)** and the estimated regions **(b)**. Point density also varies between locations, with the predicted data set having a higher density.*

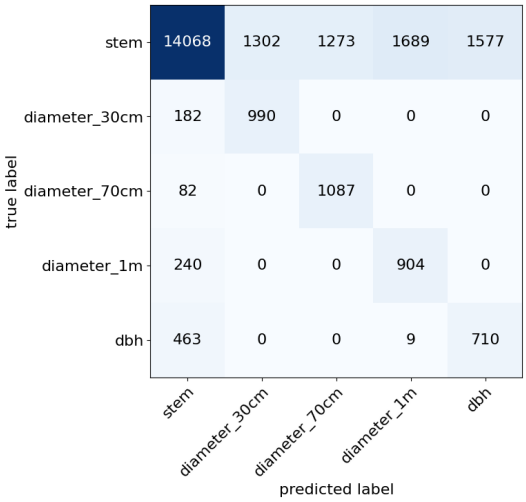


Figure 6.8: *Confusion matrix of regions estimated in the stem segmentation process.*

The analysis revealed that PointNet's performance in segmenting regions was low due to the high dispersion of the point cloud, combined with the considerable discrepancy in the diameter and stem regions. However, these results were sufficient to identify the corresponding regions.

6.5.4 Diameter Measurement

The method and results of the diameter measurements are presented in Table 6.4. The trees were categorized into four diameter classes based on their DBH, as described in Section 6.3. The Shapiro-Wilk test yielded a p-value of 0.59 for the DBH diameters measured by caliper and a p-value of 0.20 for the diameters estimated through segmentation using PointNet. These p-values indicate that both distributions follow a normal distribution, thus justifying the use of the Student's t-test for further analysis.

Table 6.4: *Results for diameters measured by the proposed method.*

Class	RMSE (cm)	MAPE (%)	DBH max error (cm)	t-Student (p-value)
Dominated	0.9	8.06	1.5	0.59
Intermediary	0.9	5.35	1.6	0.38
Codominant	0.7	2.49	0.7	0.13
Dominant	0.9	3.02	2.0	0.66

Upon examining the table, it is notable that the maximum error observed was 2 cm, specifically within the dominant trees category. However, RMSE remains below 1 cm, and MAPE values indicate that the disparities between observed and estimated values are consistently below 10%. Analyzing the significance of p values for each category reveals that, for all classes, these values exceed the significance level ($\alpha = 0.05$). Consequently, the null hypothesis cannot be rejected, suggesting the absence of significant differences between measurements conducted manually with a caliper and values estimated by the proposed method.

Figure 6.9 portrays a scatter plot encompassing the entire dataset, illustrating the relationship between actual and estimated diameters. Notably, a strong positive correlation is evident, as indicated by the high Pearson correlation coefficient (r) 0.98.

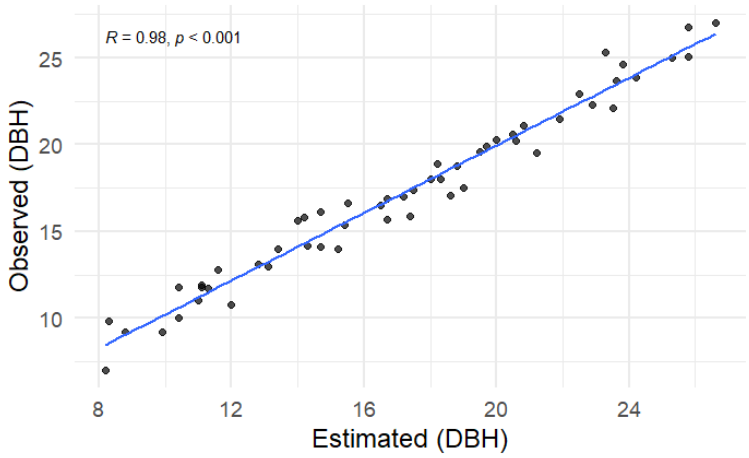


Figure 6.9: *Graph representing dispersion between estimated and observed diameters.*

6.6 Discussion

This section presents and discusses the main results obtained with the proposed method.

6.6.1 Tree Detection and Ground Removal

Approaches that employ TLS demonstrate greater promise in obtaining tree stem measurements compared to methods that use ALS. Although ALS-based methods attempt to derive stem information from the point cloud, the method adopted in this study offers substantial advantages in achieving more accurate measurements of the tree stem.

The approach utilizing object detection through Faster R-CNN in BEV images proved highly effective by streamlining the ground detection and removal process. This is particularly noteworthy, given that alternative soil removal methods often affect extensive tree areas, posing a potential risk to the accuracy of diameter measurements.

Another factor that improves precise tree identification is the individual scanning process employed for each tree, as opposed to a comprehensive scan. This method transforms the point cloud into BEV images, emphasizing a significant portion of the tree stem. Consequently, this enhances the efficacy of the object detector.

After delimiting the region with the object detector, where a considerable portion of the soil is removed, the remaining soil is easily excluded when creating the DTM. This process is carried out without compromising any significant measurable portion of the tree diameters. This approach proved more efficient than removing soil prior to converting the point cloud to BEV.

Although no instances of false positives (misidentification of background or bushes as trees) were observed, the inherent advantage of using a Faster R-CNN detector lies in the potential for improvement by simply including more examples in the training set. This procedure results in more efficient and accurate bounding boxes when performing the tree identification task.

6.6.2 Stem Segmentation

The outcomes suggest that the achieved metrics could have been more favorable. Specifically, the IoU falls notably below anticipated levels in the context of segmentation problems. Examination of Figure 6.8 reveals that various regions associated with the stem were incorrectly identified as diameter regions, directly impacting the assessed metrics. Among several hypotheses, the considerable disparity between the distribution of points related to the stem and those corresponding to the diameter sections stands out. However,

a qualitative analysis of Figure 6.7 reveals a remarkable similarity between the regions segmented by the PointNet model and those segmented manually.

Another relevant hypothesis is that the PointNet architecture analyzes the data cloud point by point, resulting in minor discrepancies that can be reflected in the final metrics. Additionally, it is essential to consider the limitation of the dataset, which exhibits restricted tree variations.

For future approaches to address this issue, it is possible to explore more recent point-to-point architectures, such as PointNet++ (QI *et al.*, 2017b), KPConv (THOMAS *et al.*, 2019), and GeomGCNN (SRIVASTAVA; SHARMA, 2021). Additionally, architectures based on voxelization, a technique that converts three-dimensional data into three-dimensional grid representations called voxels, can be considered. Examples of these architectures include VoxNet (MATURANA; SCHERER, 2015), 3D-FCN (NIE *et al.*, 2018), and 3D U-Net (ÇIÇEK *et al.*, 2016).

6.6.3 Diameter Measurement

During the data acquisition process, we aimed to represent a variety of DBH sizes to cover the most common classes in a eucalyptus stand. Upon analysis of the different diameter classes, it was observed that there are no statistically significant differences between the diameters measured using traditional methods and those estimated by the method proposed in this study.

Although the tree stem segmentation stage did not demonstrate high rates, the estimated diameters showed a high correction (r=0.98) with the measured diameters. In particular, the most significant divergence was identified in the dominant class, with a difference of 2 cm, while the smallest error was in the codominant class, with a difference of 0.7 cm. The MAPE errors were less than 10%, and the RMSE values were less than 1 cm, as shown in Table 6.4.

The approach of measuring the region's average height along the z-axis aimed to obtain a measurement closely aligned with the height determined by the caliper. Similarly, measuring two distances along the x and y axes and then calculating their average as the final diameter mirrors the manual measurement process with a caliper. Consequently, this method yielded closely comparable results. Moreover, this approach demonstrated superior efficiency, as LiDAR readings circumvent several challenges associated with direct tree measurement, particularly human errors in handling the tools.

6.7 Conclusions

This chapter introduced a novel approach for tree identification, segmentation, and diameter measurement using smartphone-integrated LiDAR sensors combined with deep learning models. The developed methodology successfully identified trees in 3D point clouds and performed segmentation to obtain highly accurate measurements of Eucalyptus tree diameters. The results demonstrated that the proposed method achieves precision comparable to traditional techniques, with RMSE values below 1 cm and no statistically significant differences in diameter measurements.

One of the most significant advantages of this approach is the substantial reduction in time required for data collection compared to traditional tape measure methods. By leveraging LiDAR technology, the need for manual measurements is eliminated, thereby minimizing the potential for human error. This efficiency is particularly valuable in large-scale forest inventories, where traditional methods are labor-intensive and prone to inconsistencies. The automated pipeline not only accelerates the measurement process but also ensures greater consistency and reliability in the data collected.

Furthermore, the dataset generated in this study, comprising field measurements and LiDAR point clouds, has been made publicly available. This resource provides a valuable foundation for future research, enabling other researchers to reproduce, validate, and improve upon the methods presented here. By fostering collaboration and innovation, this dataset has the potential to drive advancements in forest inventory techniques and expand the applicability of smartphone LiDAR technology in forestry science.

For future work, we intend to expand the dataset to include a broader range of tree species and environmental conditions, further enhancing the generalizability of the models. Additionally, we will explore advanced deep learning architectures and segmentation techniques to improve the accuracy of diameter measurements and extend the methodology to estimate other forest attributes, such as height and volume. The integration of RGB data from smartphone cameras may also offer complementary insights, further refining the segmentation process.

CHAPTER 7

3D Point Cloud Completion in Real-World Forestry: a Comparative Study for Eucalyptus Plantations

It is better to know than to imagine.

Leonardo da Vinci

Accurate 3D reconstruction of trees from point cloud data remains a significant challenge in precision forestry due to occlusion, noise, and structural complexity inherent in natural environments. This study presents a comprehensive evaluation of seven state-of-the-art point cloud completion methods (FoldingNet, PCN, GRNet, SnowflakeNet, PoinTr, PointAttN, and ODGNet) using the novel Euca3D dataset, comprising terrestrial LiDAR scans of plantation eucalyptus trees. Our quantitative analysis reveals transformer-based architectures, particularly SnowflakeNet and PoinTr, achieve superior performance (Chamfer Distance of 43.08 and 69.16 respectively) compared to traditional approaches, outperforming traditional approaches. However, all methods demonstrated reduced accuracy when compared to synthetic benchmarks. Qualitative assessments highlight persistent challenges in reconstructing fine-scale morphological features critical for forestry applications. The findings underscore the need for domain-specific adaptations to address the unique characteristics of forest point clouds. This work establishes foundational performance benchmarks and identifies key research directions, including the integration of multimodal data and biological priors, to advance robust 3D reconstruction for ecological monitoring and forest inventory management.

7.1 Introduction

Point clouds are an efficient data structure for representing three-dimensional (3D) objects as sets of point coordinates. With increasing interest in digital twins-virtual - replicas of real-world environments - point clouds are now widely acquired using Light

Detection and Ranging (LiDAR), depth cameras, photogrammetry, and structured light scanning. Despite their utility, raw point clouds from these sensors often exhibit sparsity and incompleteness (GEIGER *et al.*, 2013), severely limiting their use in applications such as autonomous driving (ZENG *et al.*, 2018; LI *et al.*, 2020), object detection (ZHOU *et al.*, 2022; SHI; RAJKUMAR, 2020), and 3D segmentation (ZHANG *et al.*, 2023; ZHAO *et al.*, 2021a). Consequently, reconstructing missing regions in incomplete point clouds is a critical task in 3D computer vision

However, point cloud completion is highly challenging. Accurate reconstruction requires both high-level semantic understanding of the object and low-level geometric reasoning about local point relationships. The problem is ill-posed, meaning a single incomplete input may correspond to multiple plausible outputs, complicating inference of missing (FEI *et al.*, 2022).

Although deep learning has demonstrated promising results in point cloud completion, existing approaches remain largely constrained to synthetic datasets, with few methods rigorously validated on real-world data. Real-world point clouds exhibit inherent challenges including non-uniform sparsity, sensor noise, and structural complexity, where missing regions result from occlusions, limited viewing angles, material properties, and environmental conditions (JMAA; CHAIEB, 2024). These issues are especially acute in natural environments like forests, where dense vegetation creates severe occlusion patterns and variable lighting conditions introduce additional noise and artifacts, significantly compromising data integrity (DASSOT *et al.*, 2011).

Eucalyptus plantations exemplify these challenges. These high-density monocultures consist of fast-growing trees forming closed, homogeneous canopies, typically planted in 3×2 m to 3×3 m spacings (PICOS *et al.*, 2020). Their simplified vertical architecture, evergreen foliage, and shallow root systems create a dense canopy that restricts sunlight penetration, resulting in sparse understory vegetation and reduced biodiversity (FAYAD *et al.*, 2021).

The precise measurement of the attributes of the eucalyptus tree, including height, diameter at breast height (DBH), and stem taper, is vital for forest inventories. These metrics underpin calculations of basal area, wood volume, and productivity. Traditional methods rely on manual sampling, which is labor intensive, time consuming, and error prone (GAFFREY *et al.*, 2001).

The structural complexity of eucalyptus plantations introduces significant challenges for LiDAR-based data acquisition. The high leaf density of eucalyptus canopies attenuates approximately 40% of laser pulses, severely limiting the detection of stems and lower-canopy structures. In mature plantations, where tree crowding is pronounced, individual tree identification rates decline to 70–80%, with height measurement errors exceeding 3 meters for suppressed trees adjacent to dominant individuals (OLIVEIRA *et*

al., 2012).

These combined factors - signal attenuation, occlusion effects, and sensor noise - underscore the need for robust point cloud completion methods to enable accurate forest inventory assessments

This study investigates state-of-the-art deep learning architectures for point cloud completion of eucalyptus trees captured via LiDAR scanning. Through comprehensive statistical and qualitative analyzes, we evaluate the performance of the leading completion networks, including FoldingNet (YANG *et al.*, 2018), PCN (YUAN *et al.*, 2018), GRNet (XIE *et al.*, 2020c), SnowflakeNet (XIANG *et al.*, 2021a), PoinTr (YU *et al.*, 2021), PointAttN (WANG *et al.*, 2024) and ODGNet (CAI *et al.*, 2024), in real-world scenarios of eucalyptus forest plantation. Our comparative assessment focuses on these models' ability to handle the complex challenges posed by natural forest environments, such as occlusion patterns, structural sparsity, and sensor noise inherent to LiDAR-based forestry surveys.

This work makes three key contributions to point cloud completion in forest applications.

- We present the first comparative analysis of state-of-the-art point completion networks in real-world eucalyptus forests, evaluating their robustness to natural challenges;
- We introduce and make publicly available for download Euca3D, a novel dataset containing paired partial-complete point clouds of plantation eucalyptus trees acquired through terrestrial LiDAR scanning;
- Our quantitative and qualitative results establish performance baselines for this critical agricultural use case.

The remainder of the chapter is structured as follows. Section 7.3 details the methodology used to simulate predation, the attributes of the pests considered in this study, and the preparation of a new image dataset containing defoliated soybean leaves. Section 7.4 presents the experimental results and a discussion considering models assertiveness and time performance of different convolutional neural networks. Finally, the work is concluded in Section 7.5.

7.2 Related Works

Point cloud completion methods primarily follow two approaches: traditional techniques that leverage geometric assumptions or shape retrieval from databases, and deep learning methods that directly learn shape priors from data for more flexible completion. Recently, transformer-based architectures have emerged as particularly effective

due to their ability to simultaneously model both local geometric details and global structural relationships, offering superior performance in reconstruction tasks.

7.2.1 Traditional Methods

Geometry-based point cloud completion methods use intrinsic geometric cues to fill missing data without external training (BERGER *et al.*, 2014; DAVIS *et al.*, 2002; SORKINE; COHEN-OR, 2004; NGUYEN *et al.*, 2016). Although effective for structured and moderately incomplete scans, these methods struggle with real-world data characterized by irregular or extensive gaps, limiting their applicability in complex environments.

Alignment-based shape completion methods match partial inputs with templates from databases, retrieving either full shapes (HAN; ZHU, 2008; LI *et al.*, 2015; NAN *et al.*, 2012; PAULY *et al.*, 2005; SHAO *et al.*, 2012), parts for assembly (KALOGERAKIS *et al.*, 2012; KIM *et al.*, 2013; MARTINOVIC; GOOL, 2013; SHEN *et al.*, 2012), or deformable models (FELZENSZWALB *et al.*, 2009; GUPTA *et al.*, 2015; LI *et al.*, 2010; ROCK *et al.*, 2015). Some approaches instead use geometric primitives (CHAUVE *et al.*, 2010; NAN *et al.*, 2010; SCHNABEL *et al.*, 2009; YIN *et al.*, 2014). However, these methods often require costly optimization in inference time and are sensitive to noise, which limits their practicality for real-time use.

7.2.2 Deep Learning-based Methods

The rise of deep learning techniques for 3D point cloud processing (HUANG *et al.*, 2020; WANG *et al.*, 2020; GU *et al.*, 2020; NIE *et al.*, 2020; HU *et al.*, 2019b; HU *et al.*, 2020) has significantly advanced point cloud completion research. A key challenge remains the generation of high-quality, complete shapes due to the discrete nature of point clouds. Notably, FoldingNet (YANG *et al.*, 2018), though not originally intended for this task, stands as one of the early influential works in the field.

Coarse-to-fine frameworks have gained attention in point cloud completion for their interpretability and controllable generation process, like GRNet (XIE *et al.*, 2020c). Methods such as PCN (YUAN *et al.*, 2018) adopt a two-stage approach, first generating a coarse, low-resolution point cloud, which is then refined and densified through a lifting module.

7.2.3 Transformers-based Methods

Originally introduced for sentence encoding in natural language processing by Vaswani *et al.* (2017), the Transformer architecture gained popularity in 2D computer

vision (CV) tasks (PARMAR *et al.*, 2018; DOSOVITSKIY *et al.*, 2020). Its success in 2D CV has since inspired its adoption in 3D point cloud research, with early works such as PoinTransformer (ZHAO *et al.*, 2021a), PCT (GUO *et al.*, 2021), and Pointformer (PAN *et al.*, 2021) leveraging Transformer-based frameworks for point cloud representation learning.

SnowflakeNet (XIANG *et al.*, 2021a) enhances point cloud completion by integrating hierarchical point deconvolution with cross-level attention. By modeling point generation as a structured, snowflake-like process guided by skip-transformer modules, it achieves superior accuracy over coarse-to-fine methods (e.g., PCN (YUAN *et al.*, 2018)), particularly in enforcing local geometric consistency. Nonetheless, its reliance on computationally intensive transformer refinement and synthetic data highlights the need for improved efficiency and generalization to real-world scenarios.

PoinTr (YU *et al.*, 2021) introduces a transformer-based framework that formulates point cloud completion as a set-to-set translation task. By integrating geometry-aware blocks and dynamic query mechanisms, the model enhances both detail preservation and generalization capability, achieving state-of-the-art performance on benchmarks such as PCN (YUAN *et al.*, 2018) and KITTI (GEIGER *et al.*, 2012).

PointAttN (WANG *et al.*, 2024) introduces a density-agnostic attention mechanism using GDP and SFA modules instead of kNN, enabling adaptive local-global feature learning. While showing promising results, its reliance on synthetic data and lack of real-world validation limit proven scalability and generalization

ODGNet (CAI *et al.*, 2024) proposes a refined deep learning architecture designed to address limitations commonly observed in two-stage point cloud completion frameworks. Unlike prior methods such as SnowflakeNet (XIANG *et al.*, 2021a) and SeedFormer (ZHOU *et al.*, 2022), which often generate oversimplified shapes, ODGNet employs a U-Net-based architecture with multi-level feature fusion and a learnable dictionary module that encodes orthogonal shape priors.

This study presents a comparative analysis of several state-of-the-art point cloud completion models—namely FoldingNet (YANG *et al.*, 2018), PCN (YUAN *et al.*, 2018), GRNet (XIE *et al.*, 2020c), SnowflakeNet (XIANG *et al.*, 2021a), PoinTr (YU *et al.*, 2021), PointAttN (WANG *et al.*, 2024) and ODGNet (CAI *et al.*, 2024)—within real-world scenarios involving eucalyptus forest plantations. The evaluation examines the respective strengths and limitations of each architecture, offering critical insights into their effectiveness and suitability for point cloud completion tasks in complex natural environments, with particular emphasis on applications in eucalyptus forest mapping and analysis.

7.3 Materials and Methods

Problem statement: Let X represent a set of 3D points captured from the visible surfaces of an object, acquired through a single scan or a sequence of scans using a 3D sensor. Let Y denote a denser set of 3D points that uniformly samples both the visible and occluded (unobserved) regions of the object's surface. The shape completion task is thus defined as the problem of predicting the complete set Y, given only the partial observations in X.

7.3.1 Dataset

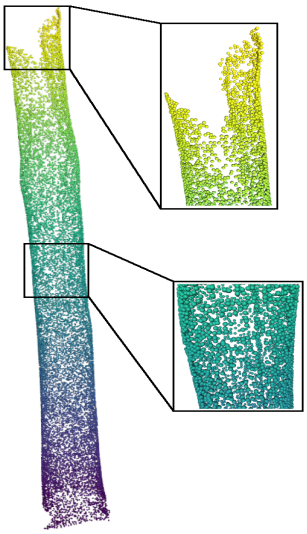
In this study, we introduce Euca3D, a novel benchmark dataset for point cloud completion tasks, focusing on real-world scans of eucalyptus trees[1]. The dataset is publicly available and comprises point cloud scans of 55 clonal eucalyptus trees (*Eucalyptus sp.*) , collected during a field campaign in June 2023 using the LiDAR sensor embedded in an iPhone 12 Pro Max. The sensor had a range of 2 meters and a spatial resolution of 5 mm. To maximize coverage of the tree structure, the device was manually rotated around each tree at a consistent height and distance of 2 meters. This procedure enabled the detailed capture of the stem and lower canopy under realistic field constraints.

To ensure compatibility and allow for direct comparison with existing methods, the Euca3D dataset is structured following the PCN methodology (YUAN *et al.*, 2018). In accordance with this format, all point clouds were uniformly resampled to 16384 points to ensure consistency across evaluations. A corresponding partial point cloud dataset was then created by employing a cross-slicing technique to randomly remove approximately 40% of the points from the original scans. Additionally, points corresponding to the ground surface were systematically removed to reduce noise and emphasize the tree structure. This entire process aims to replicate the constraints commonly encountered in real-world scanning scenarios, such as occlusions, limited accessibility, and interference from dense vegetation, which often hinder the complete circumferential capture of tree stems in natural forest environments.

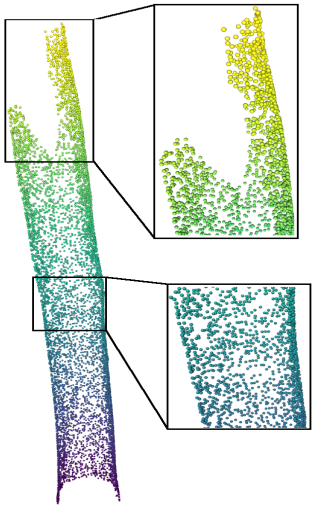
7.3.2 Evaluation Metrics

To assess the quality of the reconstructed point clouds, we compare each model's output against the corresponding ground truth using the Chamfer Distance (CD) metric (FAN *et al.*, 2017). This metric evaluates the average discrepancy between two point sets,

[1]The Euca3D dataset is available for download at: <https://drive.google.com/drive/folders/1544iUa3puaU0jkqjFcL_y2pGP29hLo4b?usp=sharing>



(a) *Complete point cloud captured through full circumferential scanning.*



(b) *Simulated partial point cloud generated by reducing point density to mimic real-world scanning constraints.*

Figure 7.1: *Visualization of a eucalyptus tree stem from the Euca3D dataset.*

$S_1, S_2 \in \mathbb{R}^3$, by measuring the proximity of each point in one set to its nearest neighbor in the other. In this study, we adopt both the L_1 and L_2 norms of the CD to provide a comprehensive evaluation of reconstruction accuracy. The L_1 variant emphasizes absolute differences and is more robust to outliers, while the L_2 version penalizes larger deviations more heavily, offering a sensitive measure of geometric precision. This dual-metric approach allows for a more nuanced comparison of model performance across varying reconstruction characteristics.

The Chamfer Distance is formally defined to L_1 as:

$$d_{CD}^{L1}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_1 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_1 \tag{7-1}$$

and L_2,

$$d_{CD}^{L2}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2 \tag{7-2}$$

This symmetric measure is widely used in 3D vision tasks, as it captures both precision and recall aspects of the reconstruction, ensuring that predicted points closely approximate the target surface while covering its full extent.

7.3.3 Training and Evaluation Protocol

To ensure the robustness and reliability of our results, we adopted a 5-fold cross-validation strategy for model evaluation. In this approach, the Euca3D dataset was partitioned into five equal subsets. For each of the five folds, four subsets (80%) were used for training the model, while the remaining subset (20%) was reserved for validation. This process was repeated until every subset had served as the validation set exactly once. The final performance metrics were then computed by averaging the results across all folds, a method that reduces variance and mitigates potential bias from any single data split.

Given the limited size of our Euca3D dataset, the training within each fold was conducted using a transfer learning approach. We employed a fine-tuning strategy that leveraged models pre-trained on the large-scale ShapeNet 3D object dataset (CHANG *et al.*, 2015a). For architectures where the original authors provided publicly available pre-trained weights, we used these as the starting point for our fine-tuning process. If pre-trained weights were not available, we first trained the models from scratch on ShapeNet. To preserve methodological consistency across all experiments, we strictly adhered to the implementation specifications and training configurations detailed in each model's respective publication.

To adapt these general-purpose models to the specific domain of eucalyptus trees, we implemented an encoder-focused fine-tuning technique with gradual unfreezing.

This method helps preserve the robust, low-level feature representations learned from ShapeNet while enabling the model to learn the specific, high-level features of our target data. The fine-tuning process involved sequentially unfreezing the encoder layers, starting from layer 6. This allowed the model to gradually adjust its weights for the Euca3D dataset, promoting effective knowledge transfer while mitigating the risk of catastrophic forgetting.

To ensure a fair and reproducible comparison, the fine-tuning process was standardized using a consistent set of hyperparameters for all evaluated models. These parameters, summarized in Table 7.1, were carefully selected to balance model adaptation to our Euca3D dataset with overall training stability. The configuration specifies the learning rate, choice of optimizer, regularization methods, and the gradual unfreeze strategy applied during training.

Table 7.1: *Fine-Tuning Hyperparameters for the Euca3D Dataset*

Hyperparameter	**Value**
Learning Rate	5e-5
Optimizers	Adam
Learning Rate Scheduler	Cosine Annealing
Batch Size	8
Epochs	250
Weight Decay	1e-4
Gradient Clipping	1.0
Freeze Strategy	Gradual Unfreeze (Encoder)

7.3.4 Environment Setting

All models were implemented using the PyTorch deep learning framework. Experimental evaluations were performed on a high-performance computing setup equipped with an NVIDIA RTX 3090 GPU, 64 GB of RAM, and an AMD Ryzen 9 processor.

7.4 Results Analysis and Discussion

This section presents the quantitative and qualitative analyses of the point cloud completion models discussed in the previous sections. The performance of each approach was evaluated using our eucalyptus dataset, following the experimental protocol and test configurations described in Section 7.3.

Table 7.2 summarizes the performance of the evaluated models on the point cloud completion task using the Euca3D dataset.

Analysis of Table 7.2 reveals that the models performed worse on the Euca3D dataset than on synthetic benchmarks like ShapeNet, highlighting the challenges of real-

Table 7.2: *Point cloud completion results on the Euca3D dataset are reported using the Chamfer Distance, computed with both L_1 and L_2 norms on point clouds. For readability, L_1 and L_2 values are scaled by 10^3 and 10^4, respectively. The best results in each metric are highlighted in bold.*

Methods	**CD-L_1**	**CD-L_2**
FoldingNet (YANG *et al.*, 2018)	167.02	1164.67
PCN (YUAN *et al.*, 2018)	110.02	883.04
GRNet (XIE *et al.*, 2020c)	72.27	226.20
SnowflakeNet (XIANG *et al.*, 2021a)	**43.08**	**84.61**
PoinTr (YU *et al.*, 2021)	69.16	236.50
PointAttN (WANG *et al.*, 2024)	98.15	456.5
OGDNet (CAI *et al.*, 2024)	72.27	228

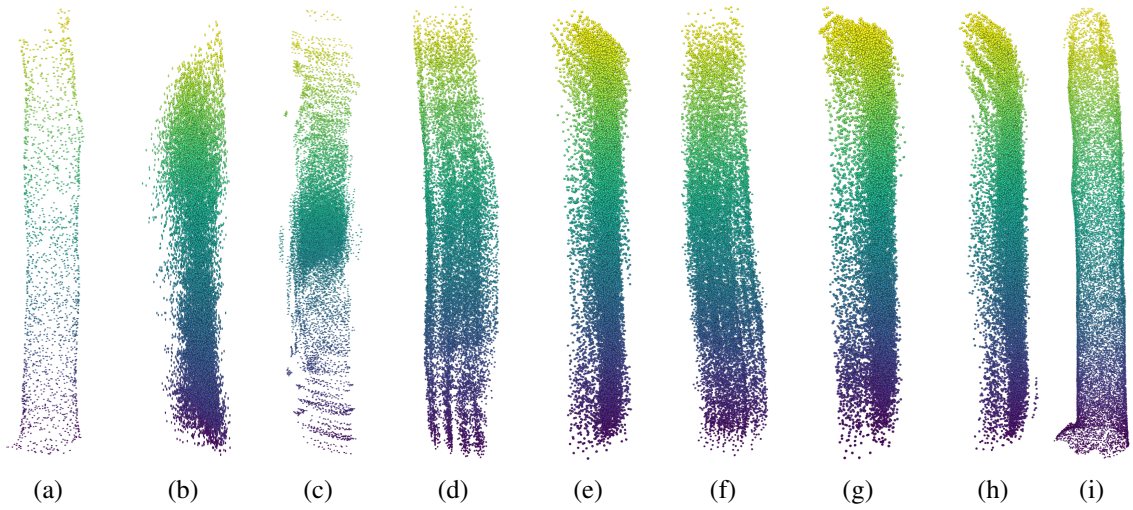


Figure 7.2: *Qualitative comparison of point cloud reconstruction results. From left to right: (a) partial input, (b) FoldingNet, (c) PCN, (d) GRNet, (e) SnowflakeNet, (f) PoinTr, (g) PointAttN, (h) ODGNet, and (i) ground truth.*

world data. This discrepancy becomes particularly evident in real-world conditions, such as those found in eucalyptus plantations. We propose two main hypotheses to explain this behavior: (i) the inherent differences between real-world point clouds and synthetic data generated from CAD models, particularly with respect to noise, occlusion, and irregular sampling; and (ii) the limited size and variability of our dataset, which comprises point clouds from only 55 eucalyptus trees.

Despite these limitations, some models demonstrated encouraging performance, suggesting that with larger and more diverse datasets, their results could improve significantly. Classical point-based architectures yielded comparatively lower accuracy, likely due to their limited capacity for geometric abstraction. Transformer-based models, while generally more robust, also struggled to generalize under the data-scarce regime, highlighting their dependence on large-scale training data. In particular, ODGNet (CAI *et al.*, 2024) and SnowflakeNet (XIANG *et al.*, 2021a) delivered more promising results,

indicating their relative effectiveness in handling incomplete and noisy input under real-world conditions. Figure 7.2 provides a visual comparison of the outputs generated by each model, highlighting differences in reconstruction quality and structural detail.

Beyond the quantitative metrics, a qualitative assessment based on the visual output of the models provides critical insights into their performance. Figure 7.2 presents a visual comparison of the reconstructions, which highlights significant differences in geometric coherence, structural integrity, and the ability to reproduce fine-scale details that are not fully captured by the CD metric. Overall, the results reveal that the networks struggled to accurately reconstruct dense and coherent point clouds from the partial inputs. While certain methods partially capture the vertical structure of the tree stem, many reconstructions fail to represent the expected cylindrical geometry, appearing sparse or disorganized. Our qualitative comparison shows that transformer-based models were more effective at preserving global structure. In particular, SnowflakeNet (XIANG *et al.*, 2021a) and PoinTr (YU *et al.*, 2021) demonstrate visually superior performance, managing to approximate the stem's overall shape more clearly than other methods.

Interestingly, although ODGNet (CAI *et al.*, 2024) was designed to enhance local geometric details through attention mechanisms, its output in this scenario appears overly smooth and less structurally defined compared to others. ODGNet, despite its focus on local details, produced overly smooth reconstructions, suggesting limited generalization to real-world forestry data. This discrepancy highlights the difficulty of generalizing from synthetic or controlled environments to the complexity of real-world forest data.

Figure 7.3 illustrates the coarse-to-fine reconstruction process used by point cloud completion models for eucalyptus tree stems. In this visualization, the red dots represent the partial input point cloud, which simulates incomplete real-world scans due to occlusions or limited sensor coverage. The blue dots correspond to the coarse output generated by the model, which captures the basic geometry of the stem and serves as the foundation for generating a denser point cloud in later stages.

These findings underscore the challenges point cloud completion models face when applied to natural, unstructured environments such as eucalyptus plantations. Factors such as occlusion, irregular geometry, and noise introduce significant obstacles. Furthermore, the limited size and diversity of our dataset may have constrained the learning capacity of the models, affecting their ability to generalize and accurately complete missing structures. These observations suggest that future improvements may require both architectural advances and the development of larger, more diverse real-world datasets.

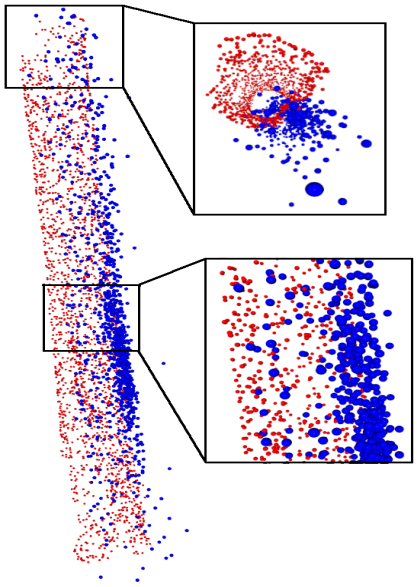


Figure 7.3: *Visualization of the coarse-level reconstruction within the point cloud completion pipeline. The top square represents a top-down view.*

7.5 Conclusions

This study evaluated state-of-the-art point cloud completion methods in the challenging context of real-world eucalyptus forestry. Our comparative analysis of seven deep learning models on the Euca3D dataset revealed significant performance gaps between synthetic benchmarks and real-world LiDAR scans, highlighting the difficulties posed by occlusion, noise, and structural sparsity in natural environments. Transformer-based models, particularly SnowflakeNet and PoinTr, demonstrated superior reconstruction capabilities by leveraging hierarchical attention mechanisms, though even these approaches struggled with fine-grained details like stem taper and branching patterns. These findings underscore the need for domain-specific adaptations in point cloud completion, particularly for precision forestry applications where accurate estimation of tree attributes (e.g., DBH, height, and volume) is critical for inventory management and productivity assessments.

In future work, we plan to expand and diversify the dataset by including a larger number of eucalyptus trees captured under varying conditions. By increasing the dataset's scale, we aim to improve the accuracy of training and evaluation in realistic forest environments. Additionally, we intend to integrate LiDAR data with complementary sensing technologies and optimize computational performance to support practical field deployment. We also plan to explore uncertainty quantification strategies to better assess the reliability of the reconstructed point clouds.

CHAPTER 8

Generating High-Fidelity Eucalyptus Point Clouds: A Comparative Study from GANs to Diffusion Models

I saw you.

John 1:48

The generation of high-fidelity 3D point clouds is a critical yet challenging task, particularly for complex natural objects. This paper presents a systematic comparative analysis of five state-of-the-art generative architectures, including *l*-GANs, PointFlow, a Diffusion Probabilistic Model (DPM), and the autoregressive CanonicalVAE, on the specific task of synthesizing eucalyptus tree point clouds. To facilitate this investigation, we introduce Euca3D, a novel benchmark dataset of 440 real-world eucalyptus tree scans. Through rigorous quantitative and qualitative evaluation, our findings reveal that DPM and CanonicalVAE emerge as the most effective models. While DPM demonstrates superior performance in surface fidelity as measured by the Chamfer Distance , CanonicalVAE excels at preserving global structural integrity and perceptual quality, achieving the best results on Earth Mover's Distance-based metrics. Visually, the point clouds generated by CanonicalVAE are of exceptional quality, nearly indistinguishable from the reference data. This study demonstrates that generating high-quality synthetic point cloud data of complex natural forms is achievable, establishing a strong foundation for data augmentation to enhance downstream tasks in forestry and computer vision.

8.1 Introduction

The representation of 3D data as point clouds has emerged as a fundamental choice for diverse applications in computer vision, making the ability to synthesize or reconstruct high-resolution, high-fidelity point clouds a critical area of research. Point clouds are recognized as an expressive and compact 3D representation, capable of

capturing fine geometric details with spatial efficiency. This benefits a wide range of synthesis tasks, such as reconstruction and super-resolution. Furthermore, they are an essential component in domains including robotics, medicine, augmented reality, and virtual reality, and serve as the standard acquisition format for devices like LiDAR, Kinect, and the iPhone's Face ID (LUETZENBURG *et al.*, 2021). The development of generative models for point clouds is also a powerful paradigm for unsupervised representation learning, enabling the characterization of data distributions and providing a foundation for downstream tasks such as shape completion, upsampling, and data augmentation.

Despite the recent success of deep learning models in discriminative point cloud tasks like classification and segmentation, the generative modeling of point clouds remains a significant challenge. Unlike images, which possess a regular grid structure, point clouds are characterized by irregular sampling patterns in 3D space and are inherently invariant to the permutation of their constituent points.

In forestry, high-density point clouds are essential for measuring tree metrics such as height and diameter at breast height (DBH) to calculate wood volume and productivity (GAFFREY *et al.*, 2001). However, current acquisition methods have significant trade-offs: Airborne Laser Scanning (ALS) is effective for large areas but its top-down view fails to capture structural details in lower canopies and stems (LOHANI; GHOSH, 2017), while Terrestrial Laser Scanning (TLS) provides high detail but is limited by high costs, occlusion, and the need for multiple scans (TELLING *et al.*, 2017). These challenges are particularly evident in high-density eucalyptus plantations, where fast-growing trees are planted in close spacings of 3×2 m to 3×3 m, forming dense canopies that exacerbate occlusion issues (PICOS *et al.*, 2020; FAYAD *et al.*, 2021). The limitations of existing technologies underscore the need for alternative approaches to point cloud generation.

In this regard, the use of data augmentation to expand point cloud datasets can substantially improve the performance of models for classification, segmentation, and object detection tasks (HIMMELSBACH *et al.*, 2009; FERNANDES *et al.*, 2021; UNAL *et al.*, 2021; QI *et al.*, 2017a; QI *et al.*, 2017b). It is envisioned that with high-fidelity synthetic data, the accuracy of these models in estimating tree diameters could potentially exceed that of conventional manual methods, such as using calipers and dendrometer tapes. This would offer a more objective, scalable, and less error-prone alternative to traditional field measurements.

This paper provides a systematic comparative analysis of five state-of-the-art architectures for 3D point cloud generation. We evaluate and contrast leading methods from different generative families, including l-GANs (ACHLIOPTAS *et al.*, 2018), PointFlow (YANG *et al.*, 2019), a Diffusion Probabilistic Model (DPM) (LUO; HU, 2021), and the

autoregressive CanonicalVAE (CHENG *et al.*, 2022). Through rigorous qualitative and quantitative benchmarking on the specific task of generating synthetic eucalyptus tree point clouds, our investigation reveals a nuanced performance landscape.

8.2 Related Works

The representation of 3D objects as point clouds is notable for its expressiveness and compactness in capturing fine geometric details. While deep learning models have achieved significant advances in discriminative tasks, the generation of high-resolution and high-fidelity point clouds remains a central challenge in the field.

8.2.1 Deep Learning on Point Clouds

Deep learning has been widely applied to enhance performance across various discriminative point cloud tasks, such as classification (QI *et al.*, 2017a; QI *et al.*, 2017b; QIAN *et al.*, 2022; YANG *et al.*, 2018) and segmentation (QI *et al.*, 2017a; QI *et al.*, 2017b; QIAN *et al.*, 2022), as well as critical point sampling (DOVRAT *et al.*, 2019). Furthermore, substantial progress has been made in point cloud synthesis, including auto-encoding (ACHLIOPTAS *et al.*, 2018; YANG *et al.*, 2018; GROUEIX *et al.*, 2018), single-view 3D reconstruction (FAN *et al.*, 2017; JIANG *et al.*, 2018; NGUYEN *et al.*, 2024; CHEN *et al.*, 2021; HUANG *et al.*, 2025), stereo reconstruction (CHEN *et al.*, 2019; YANG *et al.*, 2024), and point cloud completion (ZHANG *et al.*, 2021; HUANG *et al.*, 2020).

Gadelha *et al.* (2018) applied Variational Auto-Encoders (VAEs), while Zamorski *et al.* (2018) utilized **AEEs!** (**AEEs!**). Achlioptas *et al.* (2018) explored Generative Adversarial Networks (GANs) operating both on the raw data space and the latent space of a pre-trained auto-encoder. However, their fixed-matrix approach constrains generation to a fixed number of points and disregards permutation invariance, leading to reduced parameter efficiency.

8.2.2 3D Generative Models

Generative Adversarial Networks:

GANs have been extensively applied to point cloud generation, operating either directly on raw data (r-GANs) or within the latent space of auto-encoders (l-GANs) (ACHLIOPTAS *et al.*, 2018). Although l-GANs demonstrate greater stability and superior performance, GANs are generally susceptible to training instability and mode collapse. To

mitigate these limitations, variants such as Wasserstein GANs (WGANs) (ARJOVSKY *et al.*, 2017) have been explored.

Variational Auto-encoders:

VAEs are frameworks that learn to reconstruct their input through a latent bottleneck, thereby providing a low-dimensional representation (KINGMA; WELLING, 2013). PointFlow (YANG *et al.*, 2019) introduces a probabilistic framework that models a point cloud as a "distribution of distributions"—a distribution over shapes and a distribution over points given a shape—by employing continuous normalizing flows for both levels. Following PointFlow, other models such as DPF-Net (KLOKOV *et al.*, 2020) utilize affine coupling layers as normalizing flows to model the distribution of points. These models afford greater flexibility and enable the computation of exact likelihoods.

Autoregressive models:

Autoregressive approaches, such as PointGrow (SUN *et al.*, 2019), synthesize point clouds sequentially in a point-by-point fashion by modeling the discrete probability distribution of point positions conditioned on previously generated points. Although effective in capturing local dependencies, these methods face significant challenges when applied to the inherently unordered structure of point clouds. In particular, Transformer-based autoregressive models require a meaningful point ordering, and defining such a sequence without compromising spatial relationships or geometric coherence remains a fundamental limitation.

8.2.3 Transformers for Point Clouds

Transformers have demonstrated remarkable performance in generative tasks for text, images, and audio (VASWANI *et al.*, 2017). Recent studies have begun to explore their application to point clouds, primarily within the encoder architecture for discriminative tasks and feature extraction. Examples include employing self-attention layers for classification and segmentation (ZHAO *et al.*, 2021a), enhancing local geometric information (GUO *et al.*, 2021), and extracting local and global features for shape completion (ENGEL *et al.*, 2021; XIANG *et al.*, 2021b). Kim *et al.* (2021) applied latent code hierarchies for flexible subset modeling. To address the challenge of sequentialization, CanonicalVAE (CHENG *et al.*, 2022) decomposes point clouds into semantically meaningful compositions within a learned canonical space. This is achieved by using a canonical auto-encoder and self-supervised clustering to establish dense inter-instance correspondences.

8.2.4 Diffusion Models

Diffusion probabilistic models, inspired by non-equilibrium thermodynamics (HO *et al.*, 2020), have emerged as a promising approach for point cloud generation. In these methods, the reverse diffusion process is modeled as a Markov chain conditioned on a shape latent variable. This conditioning allows for greater shape diversity compared to models that only learn the marginal distribution of points; moreover, the prior distribution can be parameterized via normalizing flows. The work by Luo and Hu (2021), for instance, demonstrates competitive performance in generation and auto-encoding tasks. It also shows potential for unsupervised learning, yielding informative latent representations that are well-separated by category.

This study compares five architectures for point cloud generation: *l*-GAN (CD)(ACHLIOPTAS *et al.*, 2018), *l*-GAN (EMD) (ACHLIOPTAS *et al.*, 2018), PointFlow (YANG *et al.*, 2019), DPM(LUO; HU, 2021) and CanonicalVAE (CHENG *et al.*, 2022). We provide a critical analysis of their efficacy in generating synthetic point clouds of complex natural environments, with a specific focus on applications involving eucalyptus trees.

8.3 Material and Methods

8.3.1 Dataset

In this study, we introduce Euca3D, a novel benchmark dataset for point cloud completion tasks, focusing on real-world scans of eucalyptus trees[1]. The dataset is publicly available and comprises point cloud scans of 440 clonal eucalyptus trees (*Eucalyptus sp.*), collected during a field campaign in June-August 2025 using the LiDAR sensor embedded in an iPhone 14 Pro Max. The sensor had a range of 2 meters and a spatial resolution of 5 mm. To maximize coverage of the tree structure, the device was manually rotated around each tree at a consistent height and distance of 2 meters. This procedure enabled the detailed capture of the stem and lower canopy under realistic field constraints.

8.3.2 Evaluation metrics

Simmilarity metrics:

To measure the similarity between point clouds, researchers traditionally employ two primary metrics: the Chamfer Distance (CD) and the Earth Mover's Distance (EMD).

[1]The Euca3D dataset is available for download at: <https://drive.google.com/drive/folders/1544iUa3puaU0jkqjFcL_y2pGP29hLo4b?usp=sharing>

EMD solves a transportation problem that seeks to transform one point set into another by minimizing the distance between corresponding points. This metric requires that the point clouds have the same number of points (i.e., equal cardinality).

$$\mathrm{EMD}(X,Y) = \min_{\phi:X\to Y} \sum_{x\in X} \|x-\phi(x)\|_2$$

where X and Y are two point clouds of equal cardinality, and ϕ is a bijection between them.

The Chamfer Distance is defined as:

$$\mathrm{CD}(X,Y) = \sum_{x\in X} \min_{y\in Y} \|x-y\|_2^2 + \sum_{y\in Y} \min_{x\in X} \|y-x\|_2^2$$

CD is differentiable and compared to EMD more efficient to compute.

Evaluation Metrics for Generative Models:

To evaluate generative models, we first consider the three metrics introduced by Achlioptas *et al.* (2018):

Coverage (COV): COV measures the diversity of generated point clouds by computing the fraction of reference clouds matched by at least one generated sample. For each generated cloud $X \in S_g$, its nearest neighbor in the reference set S_r is identified, and COV is defined as:

$$\mathrm{COV}(S_g,S_r) = \frac{\left|\left\{\arg\min_{Y\in S_r} D(X,Y) \mid X \in S_g\right\}\right|}{|S_r|}$$

where $D(\cdot,\cdot)$ is either the CD or EMD.

Minimum Matching Distance (MMD): MMD evaluates the fidelity of generated samples by measuring the average distance from each reference cloud $Y \in S_r$ to its nearest generated cloud:

$$\mathrm{MMD}(S_g,S_r) = \frac{1}{|S_r|} \sum_{Y\in S_r} \min_{X\in S_g} D(X,Y)$$

where $D(\cdot,\cdot)$ can be either CD or EMD.

1-Nearest Neighbor Accuracy (1-NNA)

The 1-Nearest Neighbor Accuracy (1-NNA), proposed by Lopez-Paz and Oquab (2018), is a two-sample test used to evaluate whether two distributions are identical. Given sets S_g (generated) and S_r (reference), the 1-NNA is the leave-one-out accuracy of a 1-NN classifier defined as:

$$\text{1-NNA}(S_g, S_r) = \frac{\sum_{X \in S_g} \mathbb{I}[N_X \in S_g] + \sum_{Y \in S_r} \mathbb{I}[N_Y \in S_r]}{|S_g| + |S_r|},$$

where $\mathbb{I}[\cdot]$ is the indicator function and N_X denotes the nearest neighbor of X. If S_g and S_r come from the same distribution, the expected accuracy converges to 50%.

8.3.3 Training and Evaluation Protocol

To establish a robust and standardized evaluation framework, we partitioned the Euca3D dataset into three distinct, non-overlapping subsets: training, validation, and testing, following an 85%-5%-10% ratio, respectively. The training set (85% of the data) was used exclusively to learn the model parameters. The validation set (5%) served to monitor the training progress for early stopping, prevent overfitting, and perform hyperparameter tuning. Following the approaches adopted in the analyzed works, we use 2048 points to represent each shape during both training and testing, while the models can also sample an arbitrary number of points. Finally, the held-out test set (10%) was reserved for a final, unbiased evaluation of the model's generalization performance on entirely unseen data.

Given the limited size of our specialized Euca3D dataset, we adopted a transfer learning methodology. This approach involved a fine-tuning strategy that leveraged models pre-trained on three structurally complex and diverse classes from the large-scale ShapeNet dataset (CHANG *et al.*, 2015a): *airplane*, *chair*, and *car*. For architectures where the original authors provided publicly available pre-trained weights for these classes, we utilized them as our initialization point. If no such weights were available, we first trained the models from scratch on the specified ShapeNet classes. To ensure methodological consistency and reproducibility, we strictly adhered to the implementation details and training configurations described in each model's original publication.

To adapt these general-purpose models to the specific domain of eucalyptus trees, we employed a specialized fine-tuning strategy focused on the model's encoder using gradual unfreezing. This technique is designed to preserve the robust, low-level feature representations learned from ShapeNet while enabling the model to adapt to the specific high-level geometric features of our target data. Specifically, the process involved sequentially unfreezing the encoder layers, starting from the sixth layer. This approach allowed the model to progressively adjust its weights for the Euca3D dataset, promoting effective knowledge transfer while mitigating the risk of catastrophic forgetting.

8.3.4 Experimental Setup

We implemented all models in the PyTorch framework and conducted training and evaluation on a workstation equipped with an AMD Ryzen 9 CPU, 64 GB of RAM, and an NVIDIA RTX 3090 GPU.

8.4 Results and Discussion

This section presents a comprehensive evaluation of point cloud generation methods. We conduct a detailed comparative analysis by benchmarking five prominent state-of-the-art architectures in the specific task of generating eucalyptus tree point clouds, utilizing established evaluation metrics. The section concludes with an assessment of the reconstruction quality and representation learning capabilities of the models' auto-encoder components.

8.4.1 Quantitative analyses

The results of this comparative analysis are summarized in Table 8.1. For each metric, lower values indicate better performance, with the exception of Coverage, where higher values are desirable. The best-performing model for each metric is highlighted in bold. Note that MMD-CD and MMD-EMD scores have been scaled by factors of 10^3 and 10^2 respectively, for presentation.

Table 8.1: *Quantitative Generation Results*

Model	**MMD** (↓)		**COV** (%, ↑)		**1-NNA** (%, ↓)	
	CD	EMD	CD	EMD	CD	EMD
l-GAN (CD) (ACHLIOPTAS *et al.*, 2018)	19.12	21.15	17.5	18.83	99.59	99.52
l-GAN (EMD) (ACHLIOPTAS *et al.*, 2018)	18.24	20.01	19.55	18.62	95.13	98.60
PointFlow (YANG *et al.*, 2019)	10.45	15.48	41.89	**43.70**	83.91	86.33
DPM (LUO; HU, 2021)	**7.98**	12.31	44.94	44.48	**73.23**	70.81
CanonicalVAE (CHENG *et al.*, 2022)	8.91	**12.28**	**42.11**	43.29	**73.23**	70.51

↑ indicates the greater values the better, ↓ indicates the lower values the better.

The quantitative results, presented in Table 8.1, highlight a competitive landscape among the generative models, with DPM and CanonicalVAE emerging as the most effective architectures. Importantly, no single model consistently outperforms the others across all evaluation metrics.

The DPM model exhibits outstanding performance in terms of fidelity, as measured by the Chamfer Distance. It achieves the lowest score for MMD-CD (7.98) and also performs strongly in both coverage (COV-CD) and classifier-based accuracy (1-NNA-CD), where it ties for the best result.

In contrast, CanonicalVAE demonstrates a clear advantage in metrics that emphasize point-to-point correspondence and global structural coherence. It attains the best results for MMD-EMD (12.28) and 1-NNA-EMD (70.51), indicating that its generated shapes more closely preserve structural integrity and perceptual similarity to the reference data. Notably, PointFlow achieves the highest score in COV-EMD, underscoring its strength in capturing mode diversity when evaluated using EMD.

8.4.2 Qualitative Analysis

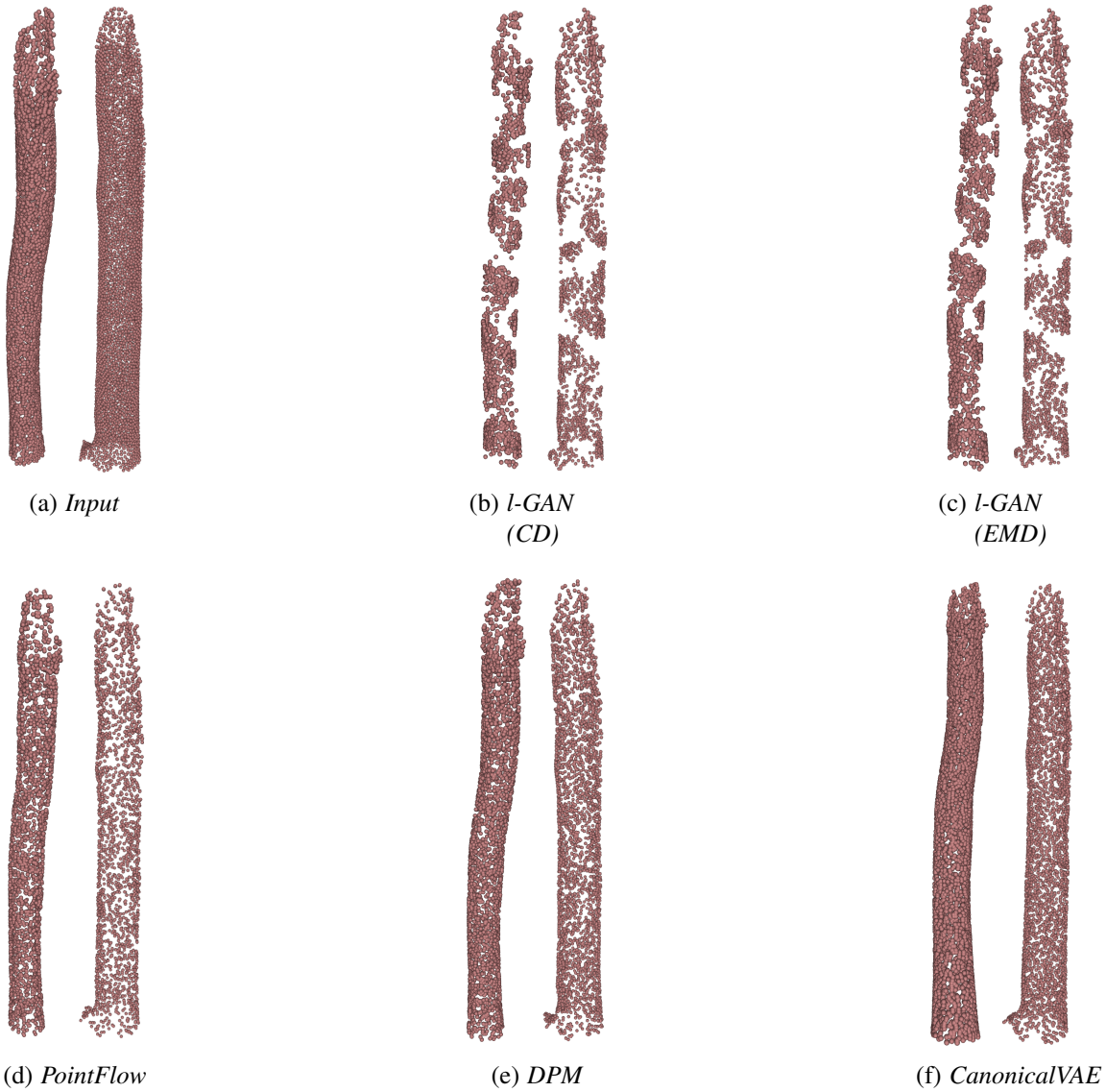
In addition to the quantitative metrics, we conducted a qualitative analysis to visually assess the fidelity, detail, and realism of the point clouds generated by each model. This visual inspection provides crucial insights into the practical quality of the synthetic data that numerical scores alone cannot fully capture. Figure 8.1 presents a side-by-side comparison, showing a reference eucalyptus tree from the test set alongside representative samples generated by the five evaluated models.

As depicted in Figure 8.1, the visual inspection reveals significant disparities in generation quality across the different architectures. The l-GAN models, under both CD (Fig. 1b) and EMD (Fig. 1c) metrics, produce the least coherent structures. These models struggle to maintain the object's continuity, resulting in fragmented point clouds with substantial structural gaps and significant noise, failing to represent a cohesive tree stem.

The PointFlow (Fig. 1d) and DPM (Fig. 1e) architectures demonstrate a marked improvement over the l-GANs, successfully generating a complete and recognizable stem structure. However, they still exhibit visible artifacts, such as surface irregularities, non-uniform point density, and a lack of the fine textural detail present in the input data.

In stark contrast, the CanonicalVAE model (Fig. 1f) produces an output of exceptional visual quality. The generated point cloud is visually almost indistinguishable from the input sample (Fig. 1a), successfully capturing not only the overall cylindrical structure but also the fine-grained surface details and subtle curvature of the stem. The resulting sample exhibits a high degree of realism and structural integrity, lacking the noise, gaps, or surface artifacts that affect the other methods.

Overall, this qualitative assessment strongly corroborates the quantitative findings presented in Table 8.1. The visual evidence confirms the superior capability of the CanonicalVAE model to synthesize complex, high-fidelity point clouds of natural objects like eucalyptus trees, establishing it as the most suitable architecture for this task among the methods evaluated.



(a) *Input* (b) *l-GAN (CD)* (c) *l-GAN (EMD)*

(d) *PointFlow* (e) *DPM* (f) *CanonicalVAE*

Figure 8.1: *Qualitative comparison of generated eucalyptus trees. While earlier models like l-GAN (b,c) exhibit significant noise and structural gaps, the CanonicalVAE (f) produces a point cloud with exceptional detail and geometric coherence, marking it as the top-performing method.*

8.5 Conclusions

This study presented a systematic comparison of five generative models for synthesizing 3D eucalyptus tree point clouds, identifying DPM and CanonicalVAE as the most effective architectures. Our analysis revealed a nuanced performance trade-off: DPM achieved the highest fidelity as measured by the Chamfer Distance, while CanonicalVAE excelled at preserving global structural integrity and perceptual similarity, earning the best scores on EMD-based metrics. This distinction was decisively confirmed by qualitative evaluation, where CanonicalVAE generated outputs of exceptional visual quality that were nearly indistinguishable from real-world scans, unlike the fragmented and noisy re-

sults from methods like *l*-GAN. These findings underscore that for applications in forestry, where geometric accuracy is paramount, the autoregressive approach of CanonicalVAE is superior. Ultimately, this work demonstrates that generating high-quality synthetic eucalyptus point cloud data is achievable through the best-performing models, paving the way for data augmentation and enhanced performance in downstream forestry applications.

CHAPTER 9

Mobile LiDAR-Based 3D Reconstruction and Segmentation for Automated Measurement of Eucalyptus Tree Diameter

"For where your treasure is, there your heart will be also."

Matthew 6:21

The integration of LiDAR sensors into consumer-grade smartphones presents a transformative opportunity for forest inventory, offering a cost-effective and scalable alternative to traditional methods. This study introduces a novel computational framework that leverages smartphone-based LiDAR and deep learning to accurately measure eucalyptus tree diameters. We developed and compared two distinct pipelines: one that reconstructs complete 360-degree point clouds from rapid, partial scans using a SnowflakeNet architecture, and another that processes full-circle scans captured directly in the field. Both workflows culminate in a shared segmentation stage where a PointTransformer network isolates the Diameter at Breast Height (DBH) region for automated measurement. Our results demonstrate that both methods achieve high accuracy, with no statistically significant difference compared to manual caliper measurements. The direct-scan approach proved superior, reducing the RMSE to less than 0.55 cm across all tree classes. Critically, our smartphone-based methods dramatically increase operational efficiency, reducing in-field data collection time by three to thirty-fold. By validating a high-precision, low-cost methodology, this work presents a significant step toward the automation of forest inventories. Furthermore, we publicly release two datasets of 880 annotated tree scans to foster future research and innovation in precision forestry. This chapter has been submitted to the IEEE Access Journal.

9.1 Introduction

The Brazilian planted forest sector is a strategic pillar of the economy, marked by resilience and innovation. Its trajectory is supported by investments projected to reach R$ 105 billion by 2028 and by global leadership, with U$ 12.7 billion in annual exports, positioning the country as the world's largest pulp exporter (IBÁ, 2024). The management and strategic planning of forest assets depend on the forest inventory, a tool that systematically gathers and analyzes data about the resources of a given area . However, the traditional method, based on limited sampling, has significant disadvantages, such as high costs, extensive time consumption, and intensive labor demands. These factors become particularly critical in large areas and in contexts requiring high-resolution data for precision forestry (HAKAMADA *et al.*, 2015).

In this context, remote sensing technologies are emerging as promising solutions for modernizing forest inventory. Among these, LiDAR (Light Detection and Ranging) stands out; it uses laser pulses to measure distances with extremely high precision, generating detailed three-dimensional models of the environment known as point clouds (DUBAYAH; DRAKE, 2000). In forest mapping, the application of LiDAR offers disruptive advantages, such as the ability to generate digital terrain models, identify individual trees, and estimate dendrometric variables like height and diameter with high accuracy and without direct contact with the target (NGUYEN *et al.*, 2021). The growing integration of LiDAR sensors into mobile platforms, such as smartphones, represents a new frontier for forest data collection. This innovation has the potential to transform forest inventory from a sporadic, sample-based process into a comprehensive, large-scale data collection effort aligned with the demands of precision forestry. The technology not only accelerates the measurement of essential variables, such as diameter at different stem heights, but also mitigates systematic errors associated with traditional instruments and enables surveys in difficult-to-access terrain (LUETZENBURG *et al.*, 2021; GAFFREY *et al.*, 2001).

Among the essential dendrometric variables, the diameter at breast height (DBH) is paramount, as even minor inaccuracies in its measurement can significantly compromise the final calculation of a tree's volume (COUTO; BASTOS, 1987). Despite its importance, a comprehensive review conducted by Rodrigues *et al.* (2024b) revealed that DBH is frequently the most underestimated variable in studies employing LiDAR technology. This highlights a critical gap between the technology's potential and its practical application, a challenge this study aims to address by developing a robust and accurate measurement framework.

This study introduces a novel computational framework for accurately measuring the DBH of eucalyptus trees using smartphone-integrated LiDAR. Our approach leverages

a multi-stage pipeline that begins with raw sensor data and employs deep learning for point cloud reconstruction, semantic segmentation, and ultimately, automated diameter calculation. The main contributions of this work are:

- A novel processing pipeline that integrates depth map upsampling and point cloud completion with deep learning-based semantic segmentation for robust tree analysis.
- The development and comparison of two distinct workflows: one for reconstructing complete models from rapid, partial scans, and another for processing full 360-degree scans.
- The creation and public release of two new datasets, *PartialEucaliptoTree* and *FullEucaliptoTree*, comprising 880 field-scanned samples to support future research and reproducibility.
- A comprehensive statistical validation demonstrating that the proposed method achieves accuracy comparable to traditional techniques while offering up to a thirty-fold improvement in data collection efficiency.

The remainder of this study is organized as follows. Section 9.2 reviews the relevant literature on LiDAR applications in forestry. Section 9.3 details the proposed methodology, while Section 9.4 describes the experimental setup and evaluation metrics. The results and discussions are presented and compared with field data in Section 9.5. Finally, Section 9.6 concludes the chapter and outlines future research directions.

9.2 Literature Review

Accurately quantifying forest structural metrics from Airborne Laser Scanning (ALS) data is challenging due to factors like scan angles, low point density, and canopy obstruction. Therefore, methods have been developed to establish empirical relationships between LiDAR-derived variables, such as canopy height, and field measurements (HYYPPA *et al.*, 2001; MALTAMO *et al.*, 2007).

ALS approaches represent a significant portion of studies utilizing LiDAR in forestry. For instance, Lee *et al.* (2018c) adopted traditional machine learning approaches, such as decision trees and Support Vector Machines, for tree height estimation. In their study, Esteban *et al.* (2020b) based canopy height estimation on digital canopy height models (CHM) obtained from LiDAR data and texture analysis. Similarly, Shoot *et al.* (2021) employed several machine learning algorithms to classify forest types based on LiDAR readings.

Determining the total and commercial volume of trees is essential for environmental licensing and forest management planning. Studies have utilized aerial scans and

satellite data to extract forest features from LiDAR information. Height estimation, combined with other variables, can be used to estimate characteristics such as volume and basal area (WITTKE *et al.*, 2019; HUDAK *et al.*, 2020; BOLTON *et al.*, 2020; OZKAN *et al.*, 2022). Corte *et al.* (2020b) explored different machine learning approaches for estimating height and volume. Additionally, Persson *et al.* (2022) and Ozkan *et al.* (2022) focused on estimating volume and diameter.

Forest biomass is an increasingly relevant energy source due to its effectiveness in reducing greenhouse gas emissions. Hudak *et al.* (2020) proposed a carbon monitoring system to map above-ground biomass. To estimate forest biomass, deep learning approaches were employed by Wittke *et al.* (2019) and Ayrey and Hayes (2018). Furthermore, studies by Mahoney *et al.* (2018), Cao *et al.* (2019), Hoover *et al.* (2018), Beaudoin *et al.* (2022), and Ehlers *et al.* (2022) explore various alternatives for using LiDAR and satellite images to measure biomass potential.

The Terrestrial Laser Scanner (TLS) can provide detailed information about specific trees or forest areas. In recent decades, several applications have been designed to explore the advantages of TLS, replacing conventional methods of measuring attributes in forest inventories. Unlike ALS, the approach using TLS, for the most part, requires manual operation of the LiDAR equipment by a professional in the field.

Diameter is a fundamental dendrometric parameter essential for calculating various forest metrics, making it crucial in forest inventory due to its accessibility and strong correlation with other variables. Recognizing the importance of the diameter, Heinzel and Ginzler (2018) performed diameter estimation using the TLS approach, while Nguyen *et al.* (2021) and Chen *et al.* (2020a) investigated diameter prediction using 3D models. Tree identification and diameter estimation are also addressed in studies by Zhou *et al.* (2017), Sun *et al.* (2022), and Hao *et al.* (2020). The TLS approach demonstrates greater efficiency than ALS, as it overcomes limitations associated with scan angles and canopy obstruction, factors that reduce correlation with data obtained through field measurements (THIES; SPIECKER, 2004).

Neural networks have been widely used in forest inventory tasks, demonstrating their versatility in addressing various challenges. Noteworthy examples include the study by Alon *et al.* (2019), which utilizes RetinaNet on ALS LiDAR data to perform tree detection and segmentation. Chen *et al.* (2020a) introduced an end-to-end pipeline for tree diameter estimation, integrating semantic segmentation, odometry, and LiDAR mapping techniques. Furthermore, Mayra *et al.* (2021) leveraged Convolutional Neural Networks (CNNs) to classify species utilizing hyperspectral images and LiDAR data. Similarly, Ayrey and Hayes (2018) employed CNNs to interpret LiDAR data and estimate various forest attributes, showcasing the broad applicability of neural network-based approaches in forest inventory tasks.

Early work explored tree detection and segmentation in LiDAR point clouds using CNN and 3D-CNN-based architectures, demonstrating the feasibility of deep learning for structural forest analysis (WINDRIM; BRYSON, 2019). Building on this direction, subsequent efforts advanced towards individual-tree characterization by incorporating PointNet-based segmentation for detecting stems and extracting dendrometric attributes directly from point clouds (WINDRIM; BRYSON, 2020b). These developments highlight a shift from general canopy-level detection toward finer, tree-centric geometric understanding.

While the reviewed studies demonstrate the significant potential of LiDAR in forestry, they predominantly rely on specialized and costly ALS or TLS systems. This study bridges a critical gap by introducing an end-to-end framework built upon accessible, consumer-grade smartphone LiDAR. Our work diverges from previous methods by developing a comprehensive pipeline that integrates advanced deep learning techniques at multiple stages: a MobileNetV2-based architecture for depth map super-resolution, a SnowflakeNet generative model for 3D point cloud completion from partial scans, and a PointTransformer network for high-fidelity semantic segmentation of the tree stem. By combining these modern architectures, we present and validate a novel, cost-effective, and highly efficient workflow for DBH measurement, directly addressing the practical limitations of both traditional inventory methods and specialized remote sensing systems.

9.3 Materials and Methods

This section details the materials and methods employed in this study. It begins by describing the study areas and the procedures for field data collection, followed by an explanation of the LiDAR data acquisition process. Subsequently, it outlines the multi-stage computational pipeline developed for tree identification, segmentation, and diameter estimation.

9.3.1 Materials

This subsection outlines the materials used in our experiments, including the study areas and the datasets.

Study area

The research was conducted in commercial *Eucalyptus sp.* plantations situated in two distinct regions of Brazil, as depicted in Figure 9.1. The first study site is located in the municipality of Açailândia, in the state of Maranhão, while the second is in the municipality of Alagoinhas, in the state of Bahia. Both regions are characterized by a hot

and humid tropical climate, with a pronounced rainy season that typically extends from November to May, providing suitable conditions for eucalyptus cultivation.

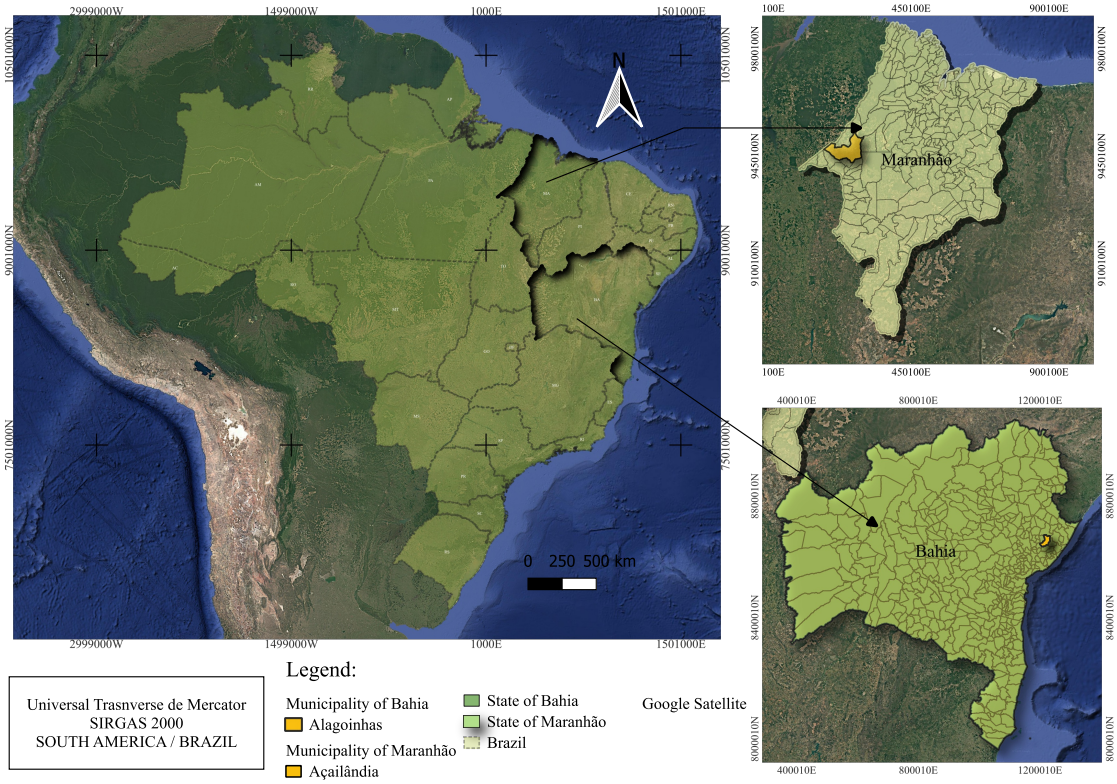


Figure 9.1: *Location of the study areas within Brazil. The maps highlight the municipalities of Açailândia in the state of Maranhão and Alagoinhas in the state of Bahia.*

Field data

The dataset comprises dendrometric measurements and corresponding LiDAR point clouds for 440 eucalyptus trees. To support reproducibility and further research, the complete dataset is publicly available for download[1].

Field sampling was conducted in two campaigns (June 2023 and June 2024), encompassing a total of 440 trees from a single clone of *Eucalyptus sp.* The geographic coordinates of each tree were recorded using a handheld GNSS receiver (Garmin 76CSx), which has an estimated positional accuracy of less than 10 meters.

For each tree, the DBH was measured at 1.30 m above the ground using either a forestry caliper (Signumat 600mm) or a diameter tape. Total height was directly measured for a random subsample of 20% of the trees. The heights of the remaining 80% were

[1]Dataset available at: <https://drive.google.com/drive/folders/1VibXd6UdRrArYJG9hoTHMUBvUz6AMCEp?usp=sharing>

then estimated using a hypsometric regression model calibrated with the subsample data. Concurrently with these manual measurements, both partial and complete LiDAR scans were performed on each tree for subsequent comparison.

The sampled trees were categorized into four social classes based on DBH, using 5 cm intervals: suppressed, intermediate, codominant, and dominant. Dominant (D) trees are the tallest, with crowns extending above the main canopy and receiving full overhead and lateral light. Codominant (C) trees form the main canopy and receive direct overhead light but limited lateral light. Intermediate (I) trees are shorter, with crowns that receive only partial overhead light. Suppressed (S) trees are the smallest and are overtopped by other trees, receiving little to no direct sunlight. The descriptive statistics for these classes are presented in Table 9.1.

Table 9.1: *Dataset statistics, including DBH limits, averages, and variability.*

Class	N° Trees	Limits		Avg[1]	Sd[2]	Range (ht)
		Inf.	Sup.			
Dominated	117	6.71	11.71	9.93	1.50	6.71–13.90
Intermediaries	131	11.71	16.65	14.79	1.68	13.05–19.90
Codominant	143	16.72	21.67	19.05	1.41	19.00–28.05
Dominant	117	21.72	27.45	24.46	1.57	22.80–28.44

[1] Average.
[2] Standard deviation.

LiDAR data

The LiDAR point clouds were acquired using the sensor integrated into an iPhone 14 Pro Max smartphone. Data acquisition was managed through Apple's ARKit framework, which processes and fuses information from the LiDAR sensor and the RGB cameras to generate depth estimates. In addition to the primary point cloud, ARKit provides auxiliary data, such as real-time three-dimensional mesh reconstructions and point clouds derived from visual (RGB) features (ZEA; HANEBECK, 2022).

Compared to professional-grade terrestrial laser scanners (e.g., FARO Focus), the embedded sensor has notable limitations. These include an effective operational range of approximately 5 meters and an inability to capture data on return intensity, number of returns, or automatic point classification. Consequently, the primary data obtained in this study are restricted to point clouds where each point is defined by its three-dimensional coordinates (x, y, z) and the corresponding RGB color values (LUETZENBURG *et al.*, 2021).

9.3.2 Methods

This methodology was developed to identify individual trees in a LiDAR point cloud and estimate stem diameters at DBH height. To provide a rigorous validation, the quantitative analysis herein focuses exclusively on the DBH, the most fundamental measurement in forest inventories. The developed pipeline is, however, capable of measuring all other annotated stem sections without modification.

Data collection protocols

Data acquisition in the field followed two distinct and complementary protocols, designed to generate the PartialEucaliptoTree and FullEucaliptoTree datasets, each serving specific purposes in the processing and analysis pipeline.

Protocol A: Partial raw data acquisition — This protocol focused on acquiring raw, unprocessed sensor data to generate the PartialEucaliptoTree dataset, providing full control over the subsequent reconstruction workflow. For this method, the smartphone was mounted on a static tripod to eliminate operator instability and ensure a fixed reference frame, which is essential for the quality and temporal alignment of the data. The *StrayScanner* application was used as a data logger, capturing and saving synchronized sequences of raw depth images (from the LiDAR sensor) and color images (from the RGB camera). The acquisition was set to a rate of 60 frames per second (fps) for a fixed interval of 2 seconds per tree, yielding approximately 120 pairs of RGB-D images for each sample. This procedure generated a dense frontal scan that served as the primary input for the 3D reconstruction and point cloud completion workflow, as detailed in Section 9.4.

Protocol B: Complete in-field scanning — This protocol was designed to generate the FullEucaliptoTree dataset by capturing a complete (360°), cohesive point cloud directly in the field, minimizing the need for complex post-processing. The *3D Scanner App* was used to perform real-time data fusion and point cloud registration. The procedure involved a manual scan where the operator walked around each tree in a continuous circular motion. The smartphone was maintained at a controlled distance of approximately 2 meters from the stem, a distance established as the optimal balance for operating within the LiDAR sensor's effective range while capturing the stem geometry without significant perspective distortion. The scan covered a vertical height of up to 2 meters to ensure a robust capture of the DBH section, located at 1.30 m. Capture settings were configured for maximum density with a 5 mm resolution, resulting in an estimated 3D reconstruction error of approximately ± 5 mm.

Ground removal

The raw point clouds generated by the smartphone's LiDAR sensor are limited to *xyz* coordinates and associated RGB color channels, lacking the detailed classification (e.g., ground, vegetation) often provided by professional-grade ALS or TLS systems. Therefore, a critical preprocessing step is to accurately segment the ground from the tree structure. For this task, we employed a Pillar-based Ground Removal (PGR) algorithm (ZENG *et al.*, 2024). The PGR method operates by first discretizing the point cloud's horizontal *xy*-plane into a grid of fixed-size cells. Each cell defines a vertical column, or pillar, that contains all points falling within its 2D boundaries. Within each pillar, the algorithm identifies the point with the lowest elevation (minimum z-coordinate), which is designated as a potential ground seed point. The collection of these seed points forms a sparse representation of the terrain. A Digital Terrain Model (DTM) is then generated by interpolating this sparse set to create a continuous ground surface. Finally, each point in the original cloud is classified as either a ground or non-ground point based on its vertical distance from the interpolated DTM. Points below a defined height threshold are classified as ground and subsequently removed. This process effectively isolates the points belonging to the tree, creating a clean point cloud for subsequent analysis. The precise removal of the ground is an essential prerequisite, ensuring that subsequent dendrometric measurements, such as DBH, height, and volume, are calculated exclusively from the tree's aerial biomass, thereby eliminating a significant source of potential error. The workflow for this process is illustrated in Figure 9.2.

Depth upsampling

For the PartialEucaliptoTree dataset, a depth super-resolution step was implemented to address the significant resolution disparity between the high-resolution RGB images (1920×1440 pixels) and the low-resolution depth maps (256×192 pixels). We utilized a MobileNetV2-Enhanced Depth Map Super-Resolution Through Multi-scale Image Guidance architecture, a computationally efficient convolutional neural network (CNN), to upsample the depth maps to match the resolution of their corresponding RGB images (RODRIGUES *et al.*, 2024a). The primary purpose of this process was to generate dense, spatially aligned RGB-D pairs for constructing high-fidelity point clouds. The quality of this initial 3D representation is a critical factor, as it directly influences the accuracy of all subsequent steps in the pipeline, including registration, completion, and the final DBH estimation.

Point cloud completion

A significant challenge with the PartialEucaliptoTree dataset is that its point clouds are geometrically incomplete, representing only a frontal view of each tree. This is due to a static data acquisition protocol that causes self-occlusion. To enable accurate structural analysis, it was necessary to reconstruct a full 360° model of the tree stem from this partial data. For this task, we implemented the SnowflakeNet architecture (XIANG *et al.*, 2021b), a generative model adept at reconstructing complex organic shapes through a hierarchical process that generates morphologically plausible point clouds. A SnowflakeNet model, pre-trained on the ShapeNet dataset, was fine-tuned using

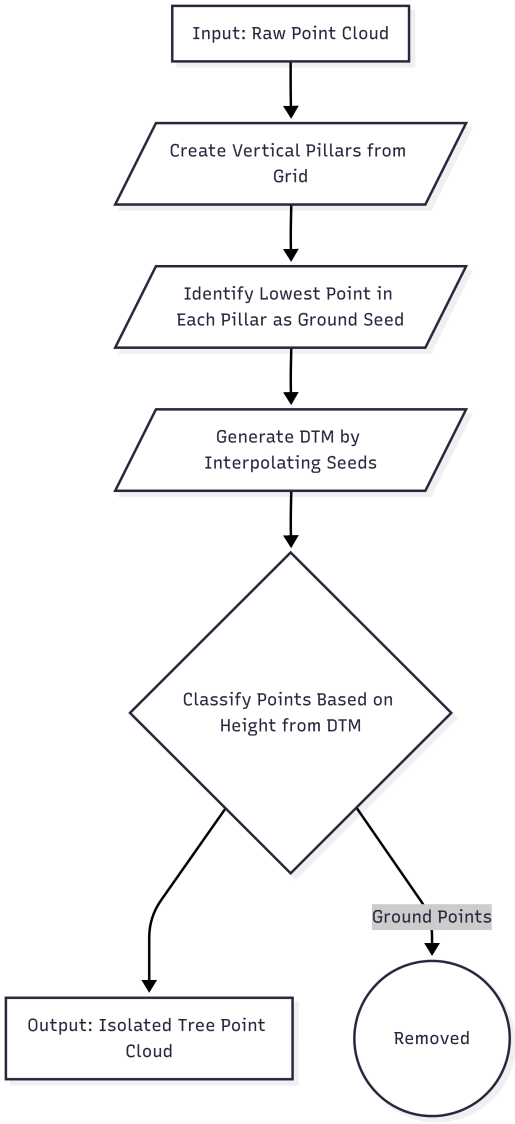


Figure 9.2: *Flowchart illustrating ground removal with the PGR algorithm. Points classified as inliers by PGR are used to build the DTM. Points labeled as outliers are preserved as candidate tree points for subsequent DBH segmentation.*

our PartialEucaliptoTree data to accurately transform rapid, partial scans into complete 3D representations.

Stem labelling and ground truth generation

To train and validate the segmentation models, a ground truth dataset with precise semantic annotations was created. This manual annotation process was applied to all tree point clouds from both the FullEucaliptoTree and PartialEucaliptoTree datasets using the open-source software CloudCompare[2].

Prior to labeling, each point cloud was downsampled to a uniform size of 4,096 points to ensure consistent input dimensions for the deep learning models. Each tree's points were then manually annotated into two semantic classes relevant to dendrometric analysis: stem, representing the main trunk, and DBH, a subset of stem points corresponding to the cross-section at 1.30 m above ground.

Diameter measurement

After the semantic segmentation of the point cloud, the challenge is to extract an accurate diametral measurement from the subset of three-dimensional coordinates representing the DBH region. For this purpose, an algorithm was developed that emulates the field measurement methodology with a caliper, which recommends obtaining two orthogonal measurements to compensate for the often-elliptical cross-section of the trunk.

As illustrated in Figure 9.3, the proposed method identifies the extreme points of the segmented point cloud along the x and y axes. The blue line (D_1) represents the three-dimensional Euclidean distance between the points with the minimum and maximum coordinates on the x-axis, while the red line (D_2) represents the distance between the points with the minimum and maximum coordinates on the y-axis. Subsequently, the arithmetic mean of these two orthogonal distances is calculated, resulting in a robust diameter estimate (D_e) that is less susceptible to local geometric variations of the stem.

Given two points, $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$, the length of D_1 can be determined using the Euclidean distance formula:

$$D_1 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \tag{9-1}$$

For the second line D_2, defined by the points $Q_1(x_3, y_3, z_3)$ and $Q_2(x_4, y_4, z_4)$, the length is computed similarly:

$$D_2 = \sqrt{(x_4 - x_3)^2 + (y_4 - y_3)^2 + (z_4 - z_3)^2}. \tag{9-2}$$

[2]CloudCompare homepage <https://www.cloudcompare.org/>

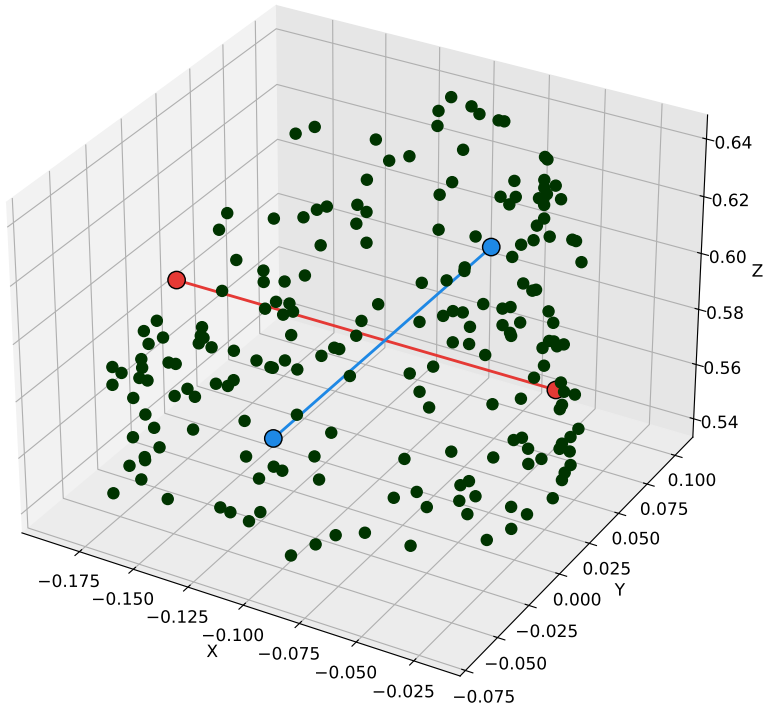


Figure 9.3: *Illustration of the algorithm for calculating the diameter from the DBH region's point cloud. The blue (D_1) and red (D_2) lines represent the two orthogonal measurements used to compute the final average diameter (D_e).*

Consequently, the lengths of D_1 and D_2 are determined. The diameter of the segmented region is then defined as the arithmetic mean of these two distances, as expressed in the following equation:

$$D_e = \frac{D_1 + D_2}{2}, \tag{9-3}$$

where D_e represents the estimated diameter. This computation provides the average diametric measure obtained along the DBH region. This computation provides an average diametric measure for the analyzed stem section. The algorithm detailing these steps is presented in Algorithm 2.

9.4 Experimental Setup

9.4.1 Experiment I - Processing Pipeline for Partial Scans

This processing pipeline was applied by adapting the FullEucaliptoTree dataset, where pairs of partial and complete point clouds were recalled in a manner consistent with

Algorithm 2 Estimate Diameter of Segmented Region

procedure CALCULATEDIAMETER(P_1, P_2, Q_1, Q_2)

▷ Calculate the length of D_1 using Euclidean distance

$\Delta x_1 \leftarrow x_2 - x_1$

$\Delta y_1 \leftarrow y_2 - y_1$

$\Delta z_1 \leftarrow z_2 - z_1$

$D_1 \leftarrow \sqrt{(\Delta x_1)^2 + (\Delta y_1)^2 + (\Delta z_1)^2}$

▷ Calculate the length of D_2 using Euclidean distance

$\Delta x_2 \leftarrow x_4 - x_3$

$\Delta y_2 \leftarrow y_4 - y_3$

$\Delta z_2 \leftarrow z_4 - z_3$

$D_2 \leftarrow \sqrt{(\Delta x_2)^2 + (\Delta y_2)^2 + (\Delta z_2)^2}$

▷ Calculate the estimated diameter D as the arithmetic mean of D_1 and D_2

$D_e \leftarrow \frac{D_1 + D_2}{2}$

return D_e

the strategy described in the PCN paper (YUAN *et al.*, 2019). The workflow was designed to reconstruct complete 3D tree models from raw sensor data and subsequently perform the final segmentation for DBH estimation.

First, the raw data, comprising low-resolution depth maps (256×192 pixels) and high-resolution RGB images (1920×1440 pixels), were processed. As detailed in Section 9.3.2, the depth maps were upsampled using a MobileNetV2 architecture to match the resolution of the corresponding RGB images. These aligned RGB-D pairs were subsequently converted into a time-ordered sequence of point clouds using the camera's intrinsic parameters, a process facilitated by the Open3D library. The reprojection of a pixel (u, v) with a corresponding depth value d to a 3D point (X, Y, Z) is performed using the camera's intrinsic matrix, K, as follows:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = d \cdot K^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{9-4}$$

where K is defined as:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \tag{9-5}$$

In this matrix, f_x and f_y represent the focal lengths in pixels, and (c_x, c_y) are the coordinates of the principal point (the optical center of the image). This mathematical

transformation converts each pixel from the RGB-D pair into a 3D coordinate, forming the point cloud.

Given the 60 fps acquisition rate over a two-second interval, approximately 120 individual point clouds were generated for each tree. The Multiway Registration method was employed to register and fuse this sequence into a single, dense, and unified frontal point cloud (CHOI *et al.*, 2015).

Since frontal scans do not capture the full circumference of the tree, a point cloud completion step was required. To this end, we employed the SnowflakeNet architecture (XIANG *et al.*, 2021b), initialized with a model pre-trained on the ShapeNet (CHANG *et al.*, 2015b) dataset and subsequently adapted to our dataset (see Section 9.3.2). This approach enabled the reconstruction of complete 360-degree eucalyptus stem representations from partial inputs.

Finally, the fully reconstructed point cloud underwent the ground removal procedure (Section 9.3.2). The resulting isolated tree point cloud was then processed by a PointTransformer (ZHAO *et al.*, 2021b) network for partil segmentation, which identified and separated the DBH region from the rest of the stem.

9.4.2 Experiment II - Processing Pipeline for Complete Scans

This pipeline was designed for the FullEucaliptoTree dataset, which contains complete, 360-degree point clouds that were reconstructed in real-time during the in-field scanning process using Apple's proprietary ARKit framework[3]. As the internal reconstruction algorithms of ARKit are not publicly documented, this method treats the generated point cloud as the direct input, simplifying the workflow.

The process begins with the complete point cloud provided by the scanning application. The same ground removal procedure described in Section 9.3.2 was applied to ensure that only the points corresponding to the tree were analyzed.

Subsequently, the isolated tree point cloud was passed to the PointTransformer network for semantic segmentation. This model identified and classified the points belonging to the stem and, more specifically, the region corresponding to the DBH, enabling the final dendrometric analysis. The PointTransformer model was trained on a combined dataset, including both the reconstructed point clouds from Pipeline A and the directly acquired point clouds from this pipeline, to ensure its robustness across different data sources.

A comprehensive overview of the distinct workflows for Experiment I and Experiment II is presented in Figure 9.4. This diagram illustrates how the two pipelines

[3] ARKit framework <https://developer.apple.com/documentation/arkit>

begin with different data inputs—raw partial scans versus complete point clouds—and converge into a shared, final processing stage for ground removal, segmentation, and measurement.

9.4.3 Stem Segmentation: Training and Evaluation Protocol

The partial segmentation of the tree stem was performed using a PointTransformer architecture. This network was chosen for its effectiveness in processing raw point cloud data, mapping each input point to a predefined class label (e.g., stem or DBH).

A preliminary analysis of the labeled dataset revealed a significant class imbalance, with the number of points belonging to the generic ‘stem‘ class far outweighing the points in the critical diameter-measurement regions. Standard loss functions, such as Cross-Entropy, would treat misclassifications equally, leading to a model biased towards the majority class. To address this, we implemented a hybrid weighted loss function combining Focal Loss and Dice Loss (LIN *et al.*, 2017; LI *et al.*, 2020). The Focal Loss component directs the model's focus toward more difficult-to-classify examples, while the Dice Loss directly optimizes the Intersection over Union (IoU) metric, improving the spatial accuracy of the segmentation.

For training, all point clouds were standardized to a uniform size of 2,048 points and normalized to fit within a unit sphere. The model was trained for 2,500 epochs using the Adam optimizer with a learning rate of 0.0001 and no weight decay. To counteract the class imbalance, the α parameter of the Focal Loss was set to 0.1 for the stem class and 0.9 for the DBH class, giving greater importance to the correct identification of diameter-related regions.

During the inference stage, the full-resolution point cloud (after ground removal) was provided to the trained model for segmentation. The dataset was randomly divided into training (80%) and testing (20%) subsets. A 3-fold cross-validation procedure was then applied within the training set to optimize the model and assess its generalization prior to the final evaluation on the independent test set.

9.4.4 Evaluation Metrics

The performance of the proposed pipelines was evaluated using a set of metrics tailored to each specific task: point cloud completion, semantic segmentation, and DBH estimation. To assess the quality of the 3D reconstruction in the point completion stage (Section 9.4.1), we employed the Chamfer Distance (CD), reporting both the L1 and L2 norms.

The CD measures the average squared distance between each point in one cloud and its nearest neighbor in the other, providing a robust measure of similarity

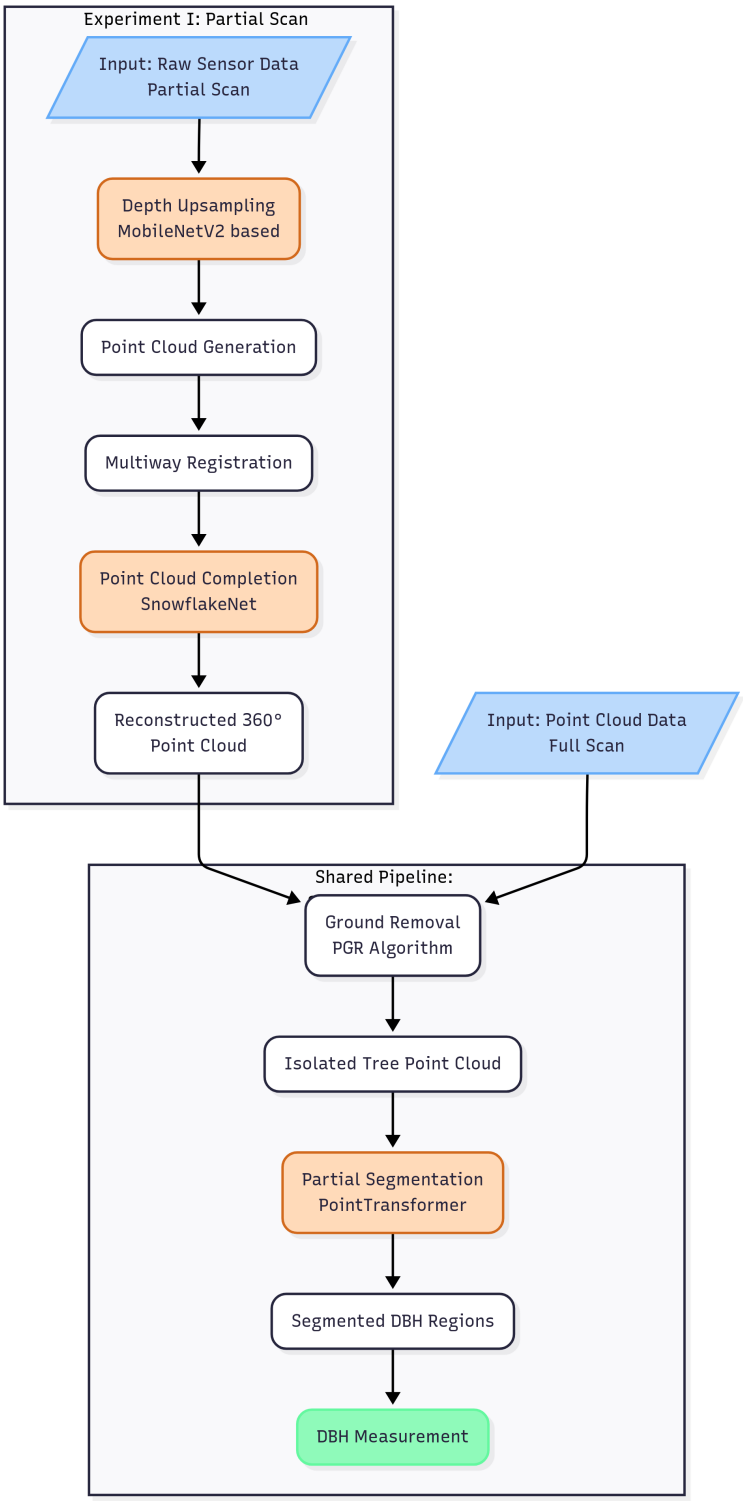


Figure 9.4: *Comparative workflow of the two experimental pipelines. Experiment I (left path) involves a multi-stage reconstruction process to generate a complete 3D model from a partial scan. Experiment II (right path) begins with a complete 360-degree point cloud. Both pipelines converge for the final shared steps of ground removal, segmentation, and measurement.*

between the reconstructed and ground-truth point clouds. The accuracy of the 3D semantic segmentation was evaluated using overall accuracy and the Mean Intersection over Union (mIoU) for each class. Overall accuracy provides a general measure of the proportion of correctly classified points, while the IoU for a given class is defined as the ratio of the intersection between the predicted and ground-truth points of that class to their union. The mIoU is the average of the IoU values across all classes, providing a balanced assessment of segmentation quality across classes. Together, these metrics offer a comprehensive evaluation of the model's ability to correctly classify each point.

The final accuracy of the DBH prediction was evaluated using the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE). These metrics quantify the average magnitude of the errors between the model's predictions and the field-measured values. To assess whether statistically significant differences existed between the measurement techniques, a stratified analysis was conducted for each DBH class. The normality of the diameter distributions for each method within these classes was evaluated using the Shapiro-Wilk test. As the assumption of normality was not met for all groups, the non-parametric Kruskal-Wallis test was employed to verify whether there were significant differences between the diameters measured by the traditional caliper method and those estimated by the methods proposed in Experiment I and Experiment II.

9.4.5 Computational Environment

All experiments were conducted on a 64-bit desktop workstation equipped with an AMD Ryzen 9 5900X processor, 64 GB of RAM, and an Nvidia GeForce RTX 3070 GPU. The deep learning models used throughout the processing pipelines were implemented in Python using the PyTorch framework.

9.5 Results and Discussion

This section presents the results of the experimental evaluation, following the sequential stages of the processing pipelines detailed in Section 9.4. We first present the performance of the individual tree detection from the initial raw point clouds. Next, we assess the quality of the 3D reconstruction from the point completion model for the PartialEucaliptoTree dataset. Subsequently, we evaluate the accuracy of the stem segmentation for both pipelines. The section culminates in a comprehensive statistical analysis of the DBH measurements, comparing the values estimated by our method against the ground-truth data collected in the field.

9.5.1 Ground Removal Performance

A foundational step common to both Experiments I and II was the accurate separation of the tree from the ground, as any residual ground points can introduce significant errors into subsequent measurements. The performance PGR algorithm was evaluated qualitatively, with a representative result presented in Figure 9.5.

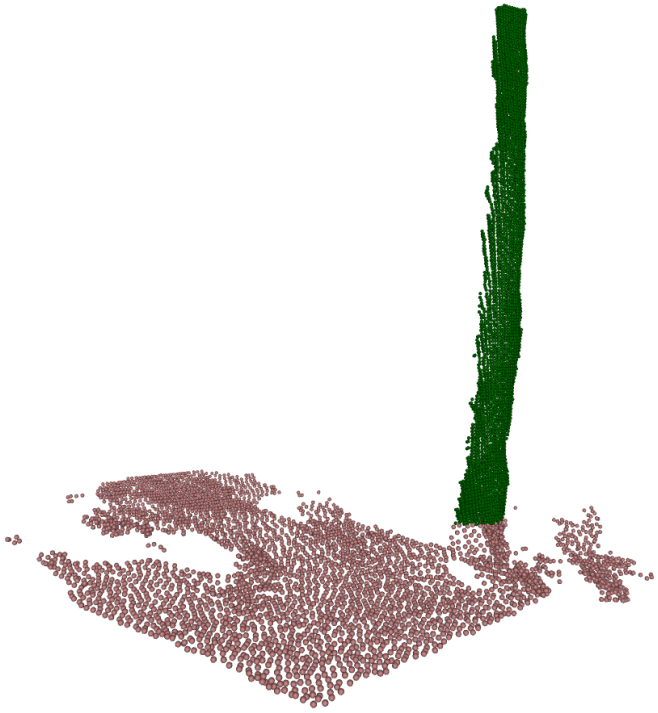
Figure 9.5(a) displays a raw point cloud, which includes both the target tree and the surrounding terrain. Following the application of the PGR algorithm, the result is a semantically segmented point cloud, as shown in Figure 9.5(b). Here, the algorithm has successfully differentiated the ground points (brown) from the tree structure (green). This correct classification is the critical outcome of this stage; it allows for the subsequent processing steps to proceed using only the points belonging to the tree. By ensuring that non-tree points are correctly identified and excluded from the subsequent analysis, the PGR algorithm prevents the propagation of errors and lays a clean foundation for accurate stem segmentation and reliable DBH estimation in both pipelines.

9.5.2 Depth Upsampling

A significant challenge in using consumer-grade smartphone sensors for 3D reconstruction is the substantial resolution mismatch between the high-resolution RGB camera and the low-resolution LiDAR sensor.

As illustrated in Figure 9.6 illustrates the resolution gap between RGB and depth images. While the high-resolution RGB image preserves fine structural details, the corresponding low-resolution depth map is sparse and lacks geometric precision. This mismatch leads to a low-density point cloud, limiting its applicability for tasks requiring accurate measurements, such as DBH estimation. To overcome this limitation, we applied a MobileNetV2-based super-resolution model to enhance the depth maps. The qualitative result of this process is shown in Figure 9.7. The upsampled depth map exhibits a dramatic increase in detail, successfully inferring high-frequency geometric features from the corresponding high-resolution RGB image. The model accurately preserves the macro-level structure, such as the curvature of the tree trunk, while generating plausible fine-grained details on the bark and ground texture. The primary benefit of this upsampling is the generation of a significantly denser and more geometrically accurate point cloud. This high-fidelity 3D representation is fundamental to the success of the entire pipeline, as it provides the necessary point density and structural detail for the subsequent ground removal, stem segmentation, and, ultimately, the reliable estimation of dendrometric parameters. Without this crucial step, the inherent sparsity of the raw sensor data would severely compromise the accuracy of any derived measurements.

(a) *Original pointcloud*



(b) *Segmented pointcloud*

Figure 9.5: *Qualitative results of the PGR algorithm. (a) The original raw point cloud, including the tree and the surrounding ground terrain. (b) The resulting point cloud after the ground points have been segmented and removed, effectively isolating the tree structure for further analysis.*

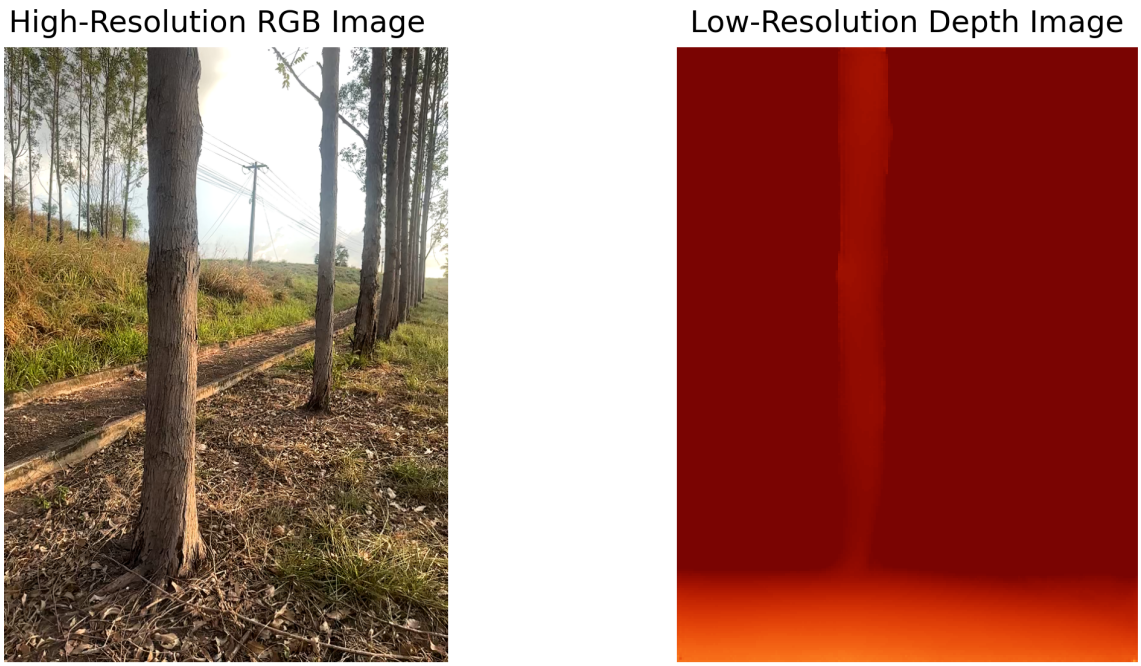
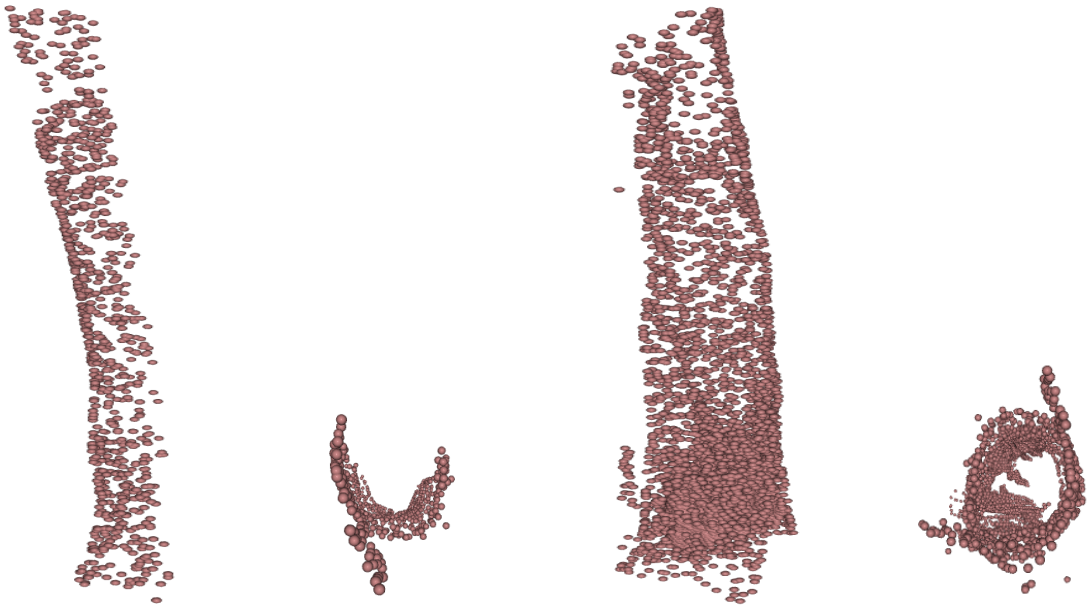


Figure 9.6: *Visualization of the input data for the depth upsampling process. The figure illustrates (left) the high-resolution RGB image and (right) the corresponding low-resolution depth map, visualized with a color map.*

Figure 9.7: *The high-resolution depth map generated by the MobileNetV2 super-resolution model, which exhibits significantly enhanced geometric detail.*

9.5.3 Point Completion Results

A critical stage in the pipeline for Experiment I was the reconstruction of a complete 360-degree tree model from a partial scan. As detailed in Section 9.3.2, this was accomplished using a fine-tuned SnowflakeNet architecture. The model's performance was evaluated both quantitatively and qualitatively to ensure the geometric integrity of the reconstructed point clouds. Quantitatively, the model's accuracy was assessed against a ground-truth dataset using the CD distance, achieving a CD-L1 of 24.81 and a CD-L2 of 4.98. These low error values indicate a high degree of fidelity between the generated outputs and the complete reference models, confirming that the network produced geometrically accurate reconstructions. The qualitative results, presented in Figure 9.8, demonstrate the model's effectiveness in generating a complete and geometrically plausible tree structure. The top row (Figures 9.8(a) and 9.8(b)) shows the sparse input point cloud from a partial scan. Notably, the top-down view reveals a significant gap in the data, representing the unscanned backside of the tree. The bottom row (Figures 9.8(c) and 9.8(d)) displays the output from the completion network. The model has successfully filled in the missing points, creating a dense and complete representation of the stem. The completed top-down view now shows a full, circular cross-section. This comprehensive 3D model is highly advantageous for subsequent processing, as a complete surface representation is critical for robustly segmenting the DBH and performing other geometric inferences on the Eucalyptus tree form.



(a) *Partial input (perspective view)*

(b) *Partial input (top view)*

(c) *Completed output (perspective view)*

(d) *Completed output (top view)*

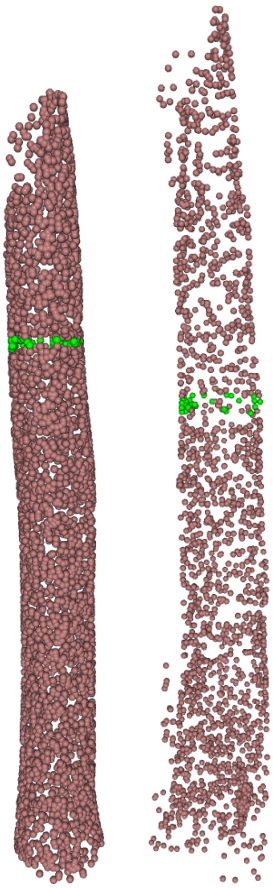
Figure 9.8: *Qualitative results of the point completion process. The top row (a, b) shows the partial point cloud from two viewpoints, while the bottom row (c, d) shows the corresponding completed point cloud generated by the model.*

9.5.4 Stem Segmentation Performance

Following the preprocessing stages, the PointTransformer network was tasked with the semantic segmentation of the tree stem to isolate the DBH region. The model demonstrated a high level of performance, achieving an overall accuracy of 97.9% and a mIoU of 79.3% across the test dataset. These strong quantitative metrics underscore the model's capability to effectively learn the distinguishing features of the DBH cross-section and accurately classify points at a fine-grained level.

The practical success of this segmentation across both experimental pipelines is illustrated in Figure 9.9. The figure visually contrasts the segmentation quality based on the input point cloud's origin. Figure 9.9(a) displays the result on a high-fidelity point cloud from the direct 360-degree scan (Experiment II). Here, the model precisely identifies a clean, coherent horizontal band corresponding to the DBH section. The segmentation is sharp and well-defined, reflecting the high quality of the input data. Conversely, Figure 9.9(b) shows the segmentation applied to a point cloud that underwent the full reconstruction pipeline, including point completion (Experiment I). While the model still successfully isolates the DBH region, the resulting band appears slightly less uniform, which is attributable to the inherent geometric variations of the reconstructed

point cloud. In both cases, the model's ability to create a distinct boundary with minimal misclassification is evident. This high degree of precision is a critical prerequisite for the final stage of the pipeline, as the reliability of the calculated diameter is directly contingent on the quality of this segmentation. The robust performance of the PointTransformer validates its selection for this task and confirms its effectiveness in enabling automated dendrometric analysis from point cloud data.



(a) (b)

Figure 9.9: *Qualitative results of DBH segmentation. (a) shows the segmentation on a point cloud from a direct 360-degree scan. (b) shows the segmentation on a point cloud reconstructed via the point completion pipeline.*

9.5.5 Diameter Measurement and Pipeline Comparison

The final and most critical evaluation of the proposed methodology lies in the precision of the DBH estimations. This stage serves as the ultimate test for both experimental pipelines, Experiment I (relying on reconstructed point clouds from partial scans) and Experiment II (using directly captured 360-degree point clouds). The performance of each was compared to traditional field caliper measurements, with the comprehensive results summarized in Table 9.2.

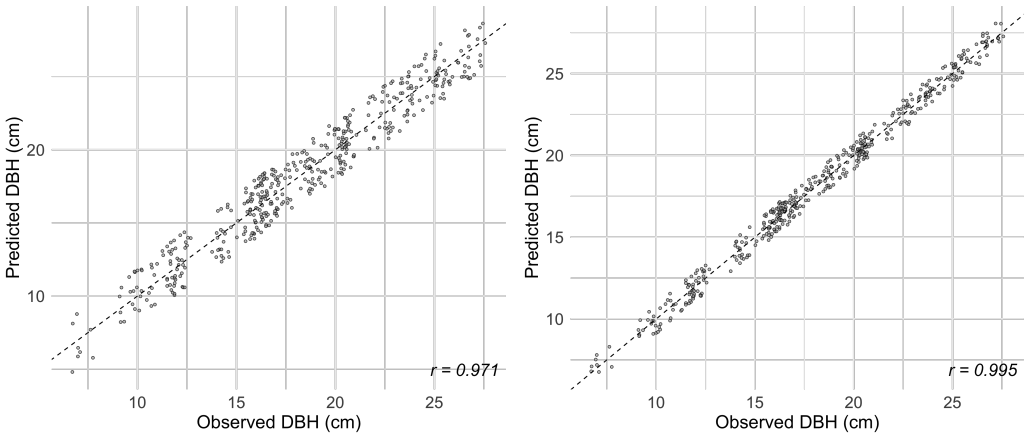
Table 9.2: *Comparison of DBH estimation performance for different tree classes in Experiments I and II.*

Class	Experiment I			Experiment II		
	RMSE (cm)	**MAPE (%)**	**DBH max (cm)**	**RMSE (cm)**	**MAPE (%)**	**DBH max (cm)**
Dominated	1.20	11.05	2.0	0.54	4.78	0.88
Intermediary	1.21	7.20	2.0	0.52	3.15	0.90
Codominant	1.12	5.21	1.98	0.48	2.17	0.90
Dominant	1.19	4.35	1.98	0.48	1.68	0.90

The Table 9.2 reveal a significant performance gap between the two pipelines. Experiment II, which utilized direct 360-degree scans, consistently outperformed the reconstruction-based approach of Experiment I, achieving substantially lower RMSE and maximum error values across all tree classes. For instance, in the "Dominant" class, the RMSE was more than halved, decreasing from 1.19 cm in Experiment I to just 0.48 cm in Experiment II, while the maximum error saw a similar reduction from 1.98 cm to 0.90 cm. This marked improvement stems directly from the superior fidelity of the input data in Experiment II, where the cohesive, high-density point clouds generated by the full-scan protocol allowed for a more precise segmentation and, consequently, a more accurate final measurement.

In contrast, the multi-stage reconstruction pipeline of Experiment I (which involves depth upsampling, Multiway Registration, and point completion) introduces multiple steps where minor errors can accumulate. Potential geometric artifacts from the upsampling network or slight misalignments during registration can propagate through the pipeline, resulting in a final point cloud that is a less perfect representation of the tree's cylindrical form. The higher quality of the input data in Experiment II allows the PointTransformer to perform a more precise segmentation of the DBH region, which in turn leads to more accurate and reliable final diameter measurements that are much closer to the ground-truth values obtained in the field.

The strong linear relationship between the field measurements and the estimates from both pipelines is visualized in the scatter plots in Figure 9.10. The results for Experiment I, shown in Figure 9.10(a), exhibit a clear positive correlation but with noticeable dispersion around the regression line. In stark contrast, the plot for the superior Experiment II pipeline, presented in Figure 9.10(b) reveals a much tighter and more precise correlation, with the data points aligning almost perfectly along the 1:1 line. This visual distinction is confirmed by a high Pearson correlation coefficient ($r = 0.99$) for Experiment II, underscoring the exceptional precision of the direct-scan method and demonstrating its readiness for operational deployment in forest inventories.



(a) (b)

Figure 9.10: *Scatter plots comparing observed (field-measured) DBH against predicted DBH for both experimental pipelines. (a) Results for Experiment I showing a positive correlation with noticeable dispersion. (b) Results for Experiment II demonstrating a significantly tighter correlation and higher precision, with data points closely aligned to the 1:1 line.*

9.5.6 Statistical Equivalence Analysis

To determine if the proposed methods produced results statistically equivalent to the traditional method, a comparative analysis was performed.

Normality testing

The Shapiro-Wilk test was applied to evaluate the normality of the data within each DBH class (Dominated, Intermediary, Codominant, and Dominant) for the three measurement groups: *Caliper*, Experiment I, and Experiment II. At a significance level of $\alpha = 0.05$, the null hypothesis (H_0) of normality was rejected in eleven of the twelve cases. Only the Experiment I group within the Codominant class exhibited a normal distribution ($p = 0.275$). Given the predominance of non-normal distributions, non-parametric tests were employed for subsequent analyses.

Comparative analysis

The Kruskal-Wallis test, a non-parametric alternative to ANOVA, was used to assess for statistical differences among the *Caliper*, Experiment I, and Experiment II measurements within each DBH class. The results are summarized in Table 9.3 and visualized in Figure 9.11.

Table 9.3: *Kruskal-Wallis Test Results by DBH Class.*

DBH Class	H Statistic (χ^2)	p-value
Dominated	2.2729	0.3210
Intermediary	0.1988	0.9054
Codominant	0.0032	0.9984
Dominant	0.4971	0.7799

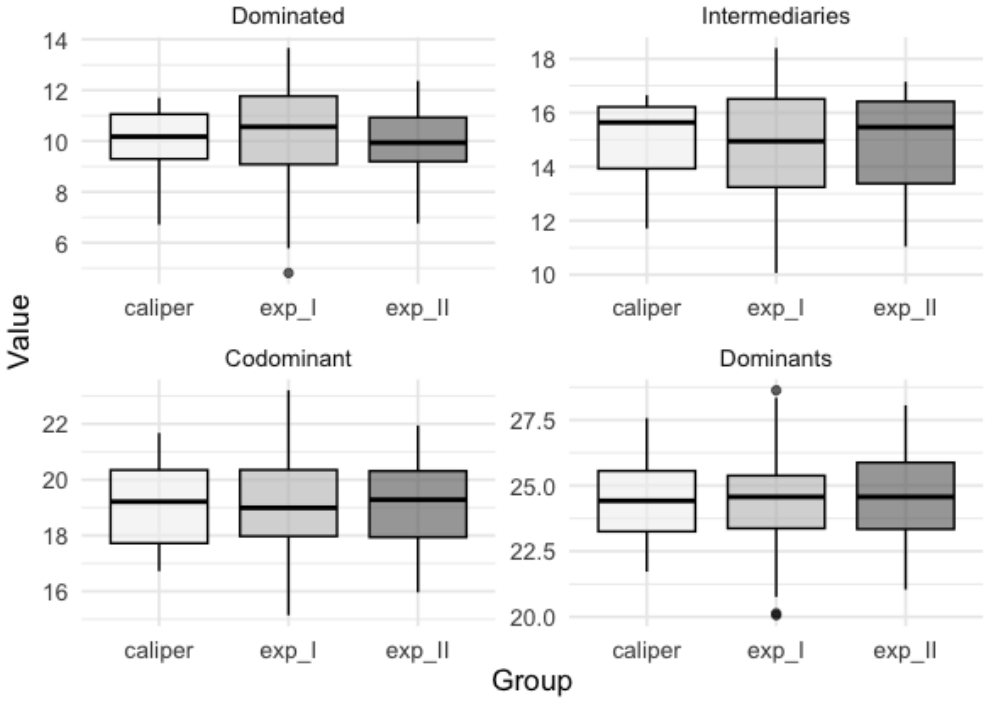


Figure 9.11: *Kruskal–Wallis Test by DAP Class. Comparison among Caliper, Experiment I, and Experiment II. No significant differences were observed, corroborating the Kruskal–Wallis and Dunn results.*

No statistically significant differences ($p > 0.05$) were detected among the three measurement groups for any DBH class. This indicates that the measurements produced by both Experiment I and Experiment II are statistically equivalent to those obtained with a traditional caliper. Dunn's post-hoc test with Bonferroni correction further confirmed this, showing no significant pairwise differences. These findings validate that the proposed LiDAR-based methods provide reliable estimates across the full range of tree diameters considered.

The decision to conduct a stratified analysis by DBH class, rather than a single global analysis, provides a more granular understanding of the model's performance. This approach makes it possible to determine whether the proposed methods exhibit varying levels of accuracy depending on the tree's diameter. A global analysis, while useful for an overall summary, would obscure these nuances, potentially masking whether

the methods are more or less precise for certain size classes (e.g., suppressed versus dominant trees). By confirming consistent performance across all diameter categories, this stratified validation strengthens the conclusion that the LiDAR-based pipelines are robust and reliable for operational use in diverse forest stands.

9.5.7 Collection Time Analysis

Beyond accuracy, the operational efficiency of a forest inventory method is a critical factor for its practical adoption. We analyzed the in-field data collection time per tree for the traditional caliper method against our two proposed smartphone-based LiDAR pipelines. The results, visualized in Figure 9.12 using a logarithmic scale to accommodate the wide range of values, reveal a dramatic improvement in efficiency. As shown in the boxplot, the manual caliper measurements are the most time-consuming, with a median collection time of approximately 75 seconds per tree and significant variability. In stark contrast, both LiDAR-based methods offer substantial time savings. Experiment II, which involves a full 360-degree scan, reduces the median collection time to roughly 22 seconds—a more than threefold improvement. The most significant leap in efficiency comes from Experiment I, where a quick, partial scan takes a median of just 2.5 seconds per tree. This represents a thirty-fold reduction in the time required for data acquisition in the field. When contextualized with the accuracy results from the previous subsection—which established that both pipelines produce measurements statistically indistinguishable from the manual method—this time saving becomes a compelling advantage. The ability to collect high-fidelity data at speeds 3 to 30 times faster than traditional methods, using a ubiquitous device like a smartphone, marks a significant step towards the automation of forest inventories. This efficiency gain not only reduces labor costs and mitigates potential human error but also enables a much higher sampling intensity, leading to more robust and representative datasets. Ultimately, this opens a clear path for more agile, data-driven forest management, providing a decisive operational advantage and paving the way for the future of precision forestry.

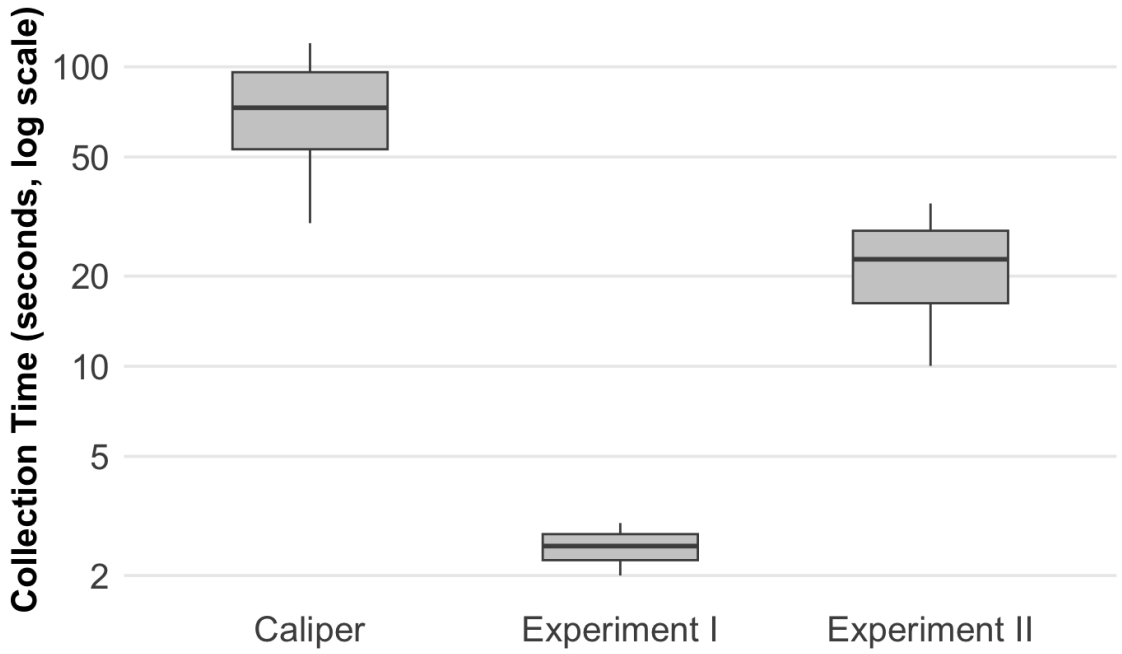


Figure 9.12: *Boxplot comparison of in-field data collection time per tree for the traditional caliper method, Experiment I (partial scan), and Experiment II (full scan). The y-axis is presented on a logarithmic scale to clearly visualize the significant time differences.*

9.6 Conclusions

This study successfully demonstrated the viability of a novel methodology for measuring eucalyptus tree diameters using smartphone-integrated LiDAR sensors and deep learning. We developed and validated two distinct processing pipelines: one reconstructing complete 3D models from rapid, partial scans, and another processing full 360-degree scans captured in the field. Our findings confirm that both approaches can produce DBH measurements with high accuracy, showing no statistically significant difference from traditional manual caliper measurements.

The primary contribution of this work lies in its significant operational advantages. The full-scan method reduced in-field data collection time by more than threefold, while the partial-scan method achieved a remarkable thirty-fold increase in efficiency. This substantial reduction in time and labor, combined with the use of ubiquitous consumer technology, has the potential to democratize and automate forest inventories. By minimizing reliance on specialized equipment and mitigating human error, this approach enables higher-density sampling, leading to more robust and representative datasets for data-driven forest management. While the direct-scan pipeline (Experiment II) yielded slightly lower error metrics due to higher initial data fidelity, the success of the reconstruction pipeline (Experiment I) is noteworthy. It proves that even extremely rapid, in-

complete scans can be computationally transformed into viable 3D models for accurate dendrometric analysis, opening possibilities for new, highly efficient data collection protocols.

Furthermore, the public release of our *PartialEucaliptoTree* and *FullEucaliptoTree* datasets provides a valuable resource to the research community for future validation and innovation. Limitations of this study include its focus on a single tree species within commercial plantation environments. Future work should therefore aim to expand the dataset to include a wider variety of species, age classes, and more complex forest structures to enhance the generalizability of the deep learning models. Subsequent research could also extend the methodology to estimate other critical forest attributes, such as tree height, stem volume, and taper. Finally, optimizing these computational models for on-device deployment could lead to a fully integrated, real-time measurement tool, marking a significant leap forward in the field of precision forestry.

CHAPTER 10

Final Considerations

> Constant doubt is the mother of progress.
>
> *Carl Friedrich Gauss*

This thesis reports the development, implementation, and rigorous validation of a complete computational pipeline for the automated measurement of dendrometric parameters in eucalyptus trees, with particular emphasis on the DBH. The central motivation was to address the operational, financial, and logistical limitations inherent in traditional forest inventories by offering a more efficient and scalable method. To this end, the study sought to promote a paradigm shift by leveraging the accessibility of consumer-grade technology, specifically LiDAR sensors integrated into smartphones, and the advanced pattern recognition capabilities of deep learning. The results demonstrate that the proposed approach constitutes a robust, accurate, and highly efficient alternative for precision forestry, yielding a workflow that is demonstrably faster and statistically equivalent to manual field methods.

The investigative framework of this research was systematically designed to address five core Research Questions (RQs) and their corresponding Hypotheses (Hs). The foundational inquiry, RQ1, examined the fundamental feasibility of using smartphone-integrated LiDAR sensors to capture point clouds with sufficient precision for the automated measurement of DBH. To test Hypothesis 1 (H1), we posited that these low-cost sensors could generate data comparable to manual measurement methods. The validation presented in Chapter 9 confirmed this hypothesis unequivocally. Results from the Kruskal–Wallis tests revealed no statistically significant difference ($p > 0.05$) between the diameters obtained using the proposed pipeline and those measured with manual calipers, thereby establishing the suitability of this low-cost technology for high-precision forestry applications.

The second major challenge, addressed by RQ2, was methodological in nature: determining how deep learning techniques, such as depth map super-resolution and point cloud completion, could be integrated into a unified pipeline to reconstruct high-fidelity

3D models from the partial and low-resolution field scans typical of low-cost sensors. This inquiry required a multi-stage investigation. First, in Chapter 4, we addressed data quality at the sensor level, demonstrating that a MobileNetV2-based architecture effectively performed depth map super-resolution, generating high-resolution depth images that consequently improved the accuracy and density of the subsequent point cloud reconstruction. Second, in Chapter 7, we tackled data incompleteness through a comparative analysis that identified transformer-based architectures, specifically SnowflakeNet, as the most effective approach for 3D point cloud completion in forestry data. The integration of these components into the Experiment I workflow (Chapter 9) validated Hypothesis 2 (H2), confirming that a unified pipeline can indeed reconstruct complete and high-fidelity 3D models from rapid, partial field acquisitions.

The core of the automated pipeline was centered on RQ3: identifying the superior effectiveness of deep learning architectures for the precise semantic segmentation of the DBH region. The comparative analysis presented in Chapter 5, followed by the implementation in the final pipeline (Chapter 9), validated Hypothesis 3 (H3). The study demonstrated that point-based architectures, particularly the PointTransformer network, exhibit strong effectiveness and can accurately isolate the DBH region. The high segmentation accuracy achieved in Chapter 9 (97.9% overall accuracy and 79.3% mIoU), even when processing complex and noisy real-world data, was pivotal to the success of the automated measurement algorithm.

Recognizing that the performance of deep learning models is often constrained by data scarcity, a persistent challenge in forestry research, RQ4 investigated the feasibility of employing generative models for data augmentation. The results presented in Chapter 8 provided strong support for Hypothesis 4 (H4). Through a systematic evaluation of state-of-the-art generative models, the study demonstrated that advanced autoregressive architectures, particularly CanonicalVAE, can synthesize eucalyptus point clouds with high structural and perceptual fidelity. The synthetic data produced were nearly indistinguishable from real-world scans, thus offering a validated resource for enhancing the robustness of future model training.

Finally, RQ5 addressed the concluding step: identifying the most effective method for calculating DBH from the segmented point cloud and comparing its performance against traditional measurements. The validation of the complete, integrated pipeline in Chapter 9 confirmed Hypothesis 5 (H5). The automated algorithm, which replicates the orthogonal measurement principle of manual calipers, produced results characterized by high correlation and low mean error. The direct-scan pipeline (Experiment II) achieved superior performance, with an RMSE of less than 0.55 cm across all tree classes. Notably, this statistical equivalence in accuracy was accompanied by transformative improvements in operational efficiency: our methods reduced in-field data collection

time by a factor of three (Experiment II) to thirty (Experiment I) compared with manual caliper measurements. This convergence of accuracy, cost-effectiveness, and efficiency fulfills the central objective of this thesis.

Limitations and Future Work

Despite the promising results, this study has several limitations. Our research focused on a single species (Eucalyptus sp.) within commercial plantation settings, and the models' generalizability to other species or more complex, heterogeneous native forests remains to be tested. The datasets developed, while novel and now publicly available, are modest in size compared to large-scale computer vision benchmarks. The methodology is also dependent on specific smartphone hardware equipped with LiDAR sensors. Future work should aim to expand the datasets to include a wider variety of species, age classes, and forest structures to enhance model robustness. The framework can also be extended to estimate other critical forest attributes, such as total height, stem volume, and taper. A significant next step would be the optimization of these computational models for on-device deployment, creating a fully integrated, real-time measurement tool that could revolutionize field data collection. In conclusion, this thesis successfully bridges the gap between consumer technology and precision forestry. By developing a cost-effective, accurate, and scalable solution for forest inventory, this work provides a tangible contribution toward the automation and democratization of sustainable forest management, paving the way for more agile, data-driven practices in the forestry sector.

Bibliography

ACHLIOPTAS, P.; DIAMANTI, O.; MITLIAGKAS, I.; GUIBAS, L. Learning representations and generative models for 3d point clouds. In: PMLR. *International conference on machine learning*. [S.l.], 2018. p. 40–49.

ALBAHRI, A.; KHALEEL, Y. L.; HABEEB, M. A.; ISMAEL, R. D.; HAMEED, Q. A.; DEVECI, M.; HOMOD, R. Z.; ALBAHRI, O.; ALAMOODI, A.; ALZUBAIDI, L. A systematic review of trustworthy artificial intelligence applications in natural disasters. *Computers and Electrical Engineering*, Elsevier, v. 118, p. 109409, 2024.

ALMEIDA, D. R. A. de; ZAMBRANO, A. M. A.; BROADBENT, E. N.; WENDT, A. L.; FOSTER, P.; WILKINSON, B. E.; SALK, C.; PAPA, D. d. A.; STARK, S. C.; VALBUENA, R. *et al.* Detecting successional changes in tropical forest structure using gatoreye drone-borne lidar. *Biotropica*, Wiley Online Library, v. 52, n. 6, p. 1155–1167, 2020.

ALON, A. S.; FESTIJO, E. D.; JUANICO, D. E. O. Tree detection using genus-specific retinanet from orthophoto for segmentation access of airborne lidar data. In: IEEE. *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*. [S.l.], 2019. p. 1–6.

ALZUBAIDI, L.; ZHANG, J.; HUMAIDI, A. J.; AL-DUJAILI, A.; DUAN, Y.; AL-SHAMMA, O.; SANTAMARÍA, J.; FADHEL, M. A.; AL-AMIDIE, M.; FARHAN, L. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, Springer, v. 8, p. 1–74, 2021.

APPLE. Arkit. 2021. Accessed on 15.03.2024. [Online] Available: https://developer.apple.com/videos/play/wwdc2021/10073.

ARJOVSKY, M.; CHINTALA, S.; BOTTOU, L. Wasserstein generative adversarial networks. In: PMLR. *International conference on machine learning*. [S.l.], 2017. p. 214–223.

ARMENI, I. Joint 2d-3d semantic data for indoor scene understanding. *arXiv preprint arXiv:1702.01105*, 2017.

ASHLEY, M.; ROGER, R. Tree heights and upper stem diameters. 1968.

ASTOLA, H.; SEITSONEN, L.; HALME, E.; MOLINIER, M.; LÖNNQVIST, A. Deep neural networks with transfer learning for forest variable estimation using sentinel-2 imagery in boreal forest. *Remote Sensing*, MDPI, v. 13, n. 12, p. 2392, 2021.

AYREY, E.; HAYES, D. J. The use of three-dimensional convolutional neural networks to interpret lidar for forest inventory. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 10, n. 4, p. 649, 2018.

AZEVEDO, G. B.; TOMIAZZI, H. V.; AZEVEDO, G. T. d. O. S.; TEODORO, L. P. R.; TEODORO, P. E.; SOUZA, M. T. P. de; BATISTA, T. S.; EUFRADE-JUNIOR, H. de J.; GUERRA, S. P. S. Multi-volume modeling of eucalyptus trees using regression and artificial neural networks. *PloS one*, Public Library of Science San Francisco, CA USA, v. 15, n. 9, p. e0238703, 2020.

BAZEZEW, M. N.; HUSSIN, Y. A.; KLOOSTERMAN, E. Integrating airborne lidar and terrestrial laser scanner forest parameters for accurate above-ground biomass/carbon estimation in ayer hitam tropical forest, malaysia. *International journal of applied earth observation and geoinformation*, Elsevier, v. 73, p. 638–652, 2018.

BEAUDOIN, A.; HALL, R. J.; CASTILLA, G.; FILIATRAULT, M.; VILLEMAIRE, P.; SKAKUN, R.; GUINDON, L. Improved k-nn mapping of forest attributes in northern canada using spaceborne l-band sar, multispectral and lidar data. *Remote Sensing*, MDPI, v. 14, n. 5, p. 1181, 2022.

BEHLEY, J.; GARBADE, M.; MILIOTO, A.; QUENZEL, J.; BEHNKE, S.; STACHNISS, C.; GALL, J. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 9297–9307.

BENTLEY, J. L. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, ACM New York, NY, USA, v. 18, n. 9, p. 509–517, 1975.

BERGER, M.; TAGLIASACCHI, A.; SEVERSKY, L. M.; ALLIEZ, P.; LEVINE, J. A.; SHARF, A.; SILVA, C. T. State of the art in surface reconstruction from point clouds. In: THE EUROGRAPHICS ASSOCIATION. *35th Annual Conference of the European Association for Computer Graphics, Eurographics 2014-State of the Art Reports*. [S.l.], 2014.

BOLTON, D. K.; TOMPALSKI, P.; COOPS, N. C.; WHITE, J. C.; WULDER, M. A.; HERMOSILLA, T.; QUEINNEC, M.; LUTHER, J. E.; LIER, O. R. van; FOURNIER, R. A. *et al.* Optimizing landsat time series length for regional mapping of lidar-derived forest structure. *Remote Sensing of Environment*, Elsevier, v. 239, p. 111645, 2020.

BOLTON, D. K.; WHITE, J. C.; WULDER, M. A.; COOPS, N. C.; HERMOSILLA, T.; YUAN, X. Updating stand-level forest inventories using airborne laser scanning and landsat time series data. *International journal of applied earth observation and geoinformation*, Elsevier, v. 66, p. 174–183, 2018.

BOUTSOUKIS, C.; MANAKOS, I.; HEURICH, M.; DELOPOULOS, A. Canopy height estimation from single multispectral 2d airborne imagery using texture analysis and machine learning in structurally rich temperate forests. *Remote Sensing*, v. 11, n. 23, 2019. Cited By 1. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85076556240&doi=10.3390%2frs11232853&partnerID=40&md5=6aec7429501efa8b3012f33323161991>.

BRENA, D.; NETO, S. P. Inventário florestal. *Curitiba.[Links]*, 1997.

BUTLER, D. J.; WULFF, J.; STANLEY, G. B.; BLACK, M. J. A naturalistic open source movie for optical flow evaluation. In: SPRINGER. *Computer Vision–ECCV 2012: 12th European*

Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VI 12. [S.l.], 2012. p. 611–625.

CABACINHA, C. D. *A method to perform forest inventories without upper-stem diameter measurements*. 116 p. PhD Thesis (PhD Thesis) — Universidade Federal de Lavras., 2003.

CACCAMO, G.; IQBAL, I.; OSBORN, J.; BI, H.; ARKLEY, K.; MELVILLE, G.; AURIK, D.; STONE, C. Comparing yield estimates derived from lidar and aerial photogrammetric point-cloud data with cut-to-length harvester data in a pinus radiata plantation in tasmania. *Australian Forestry*, Taylor & Francis, v. 81, n. 3, p. 131–141, 2018.

CAESAR, H.; BANKITI, V.; LANG, A. H.; VORA, S.; LIONG, V. E.; XU, Q.; KRISHNAN, A.; PAN, Y.; BALDAN, G.; BEIJBOM, O. nuscenes: A multimodal dataset for autonomous driving. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 11621–11631.

CAI, P.; SCOTT, D.; LI, X.; WANG, S. Orthogonal dictionary guided shape completion network for point cloud. In: *Proceedings of the AAAI conference on artificial intelligence*. [S.l.: s.n.], 2024. v. 38, n. 2, p. 864–872.

CAMPOS, J. C. C.; LEITE, H. G. *Mensuração florestal: perguntas e respostas*. [S.l.]: UFV, 2006.

CAO, L.; COOPS, N. C.; SUN, Y.; RUAN, H.; WANG, G.; DAI, J.; SHE, G. Estimating canopy structure and biomass in bamboo forests using airborne lidar data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 148, p. 114–129, 2019.

CASTILLA, G.; HALL, R. J.; SKAKUN, R.; FILIATRAULT, M.; BEAUDOIN, A.; GARTRELL, M.; SMITH, L.; GROENEWEGEN, K.; HOPKINSON, C.; SLUIJS, J. van der. The multisource vegetation inventory (mvi): A satellite-based forest inventory for the northwest territories taiga plains. *Remote Sensing*, MDPI, v. 14, n. 5, p. 1108, 2022.

CHANG, A. X.; FUNKHOUSER, T.; GUIBAS, L.; HANRAHAN, P.; HUANG, Q.; LI, Z.; SAVARESE, S.; SAVVA, M.; SONG, S.; SU, H.; XIAO, J.; YI, L.; YU, F. *ShapeNet: An Information-Rich 3D Model Repository*. [S.l.], 2015.

CHANG, A. X.; FUNKHOUSER, T.; GUIBAS, L.; HANRAHAN, P.; HUANG, Q.; LI, Z.; SAVARESE, S.; SAVVA, M.; SONG, S.; SU, H.; XIAO, J.; YI, L.; YU, F. *ShapeNet: An Information-Rich 3D Model Repository*. 2015. Available at: <https://arxiv.org/abs/1512.03012>.

CHAUVE, A.-L.; LABATUT, P.; PONS, J.-P. Robust piecewise-planar 3d reconstruction and completion from large-scale unstructured point data. In: IEEE. *2010 IEEE computer society conference on computer vision and pattern recognition*. [S.l.], 2010. p. 1261–1268.

CHEN, H.; ZUO, Y.; TONG, Y.; ZHU, L. 3d point cloud generation reconstruction from single image based on image retrieval. *Results in Optics*, Elsevier, v. 5, p. 100124, 2021.

CHEN, R.; HAN, S.; XU, J.; SU, H. Point-based multi-view stereo network. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 1538–1547.

CHEN, S.; YU, T.; LI, P. Mvt: Multi-view vision transformer for 3d object recognition. *arXiv preprint arXiv:2110.13083*, 2021.

CHEN, S. W.; NARDARI, G. V.; LEE, E. S.; QU, C.; LIU, X.; ROMERO, R. A. F.; KUMAR, V. Sloam: Semantic lidar odometry and mapping for forest inventory. *IEEE Robotics and Automation Letters*, IEEE, v. 5, n. 2, p. 612–619, 2020.

CHEN, Y.; TAI, L.; SUN, K.; LI, M. MonoPair: Monocular 3D object detection using pairwise spatial relationships. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2020. p. 12093–12102.

CHENG, A.-C.; LI, X.; LIU, S.; SUN, M.; YANG, M.-H. Autoregressive 3d shape generation via canonical mapping. In: SPRINGER. *European Conference on Computer Vision*. [S.l.], 2022. p. 89–104.

CHIRICI, G.; GIANNETTI, F.; MCROBERTS, R. E.; TRAVAGLINI, D.; PECCHI, M.; MASELLI, F.; CHIESI, M.; CORONA, P. Wall-to-wall spatial prediction of growing stock volume based on italian national forest inventory plots and remotely sensed data. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 84, p. 101959, 2020.

CHOI, O.; JUNG, S.-W. A consensus-driven approach for structure and texture aware depth map upsampling. *IEEE transactions on image processing*, IEEE, v. 23, n. 8, p. 3321–3335, 2014.

CHOI, S.; ZHOU, Q.-Y.; KOLTUN, V. Robust reconstruction of indoor scenes. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 5556–5565.

ÇIÇEK, Ö.; ABDULKADIR, A.; LIENKAMP, S. S.; BROX, T.; RONNEBERGER, O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: SPRINGER. *Medical Image Computing and Computer-Assisted Intervention: 19th International Conference*. Athens, Greece, 2016. p. 424–432.

CIREŞAN, D. C.; MEIER, U.; GAMBARDELLA, L. M.; SCHMIDHUBER, J. Deep, big, simple neural nets for handwritten digit recognition. *Neural computation*, MIT Press, v. 22, n. 12, p. 3207–3220, 2010.

CLARK, N.; WYNNE, R. H.; SCHMOLDT, D. L.; ARAMAN, P. A.; WINN, M. F. Use of a non-metric digital camera for tree stem evaluation. In: *Proceedings, 1998 ASPRS/RT Annual Convention.(pre-published version)*. [S.l.: s.n.], 1998.

CORTE, A.; SOUZA, D.; REX, F.; SANQUETTA, C.; MOHAN, M.; SILVA, C.; ZAMBRANO, A.; PRATA, G.; ALMEIDA, D. Alves de; TRAUTENMüLLER, J.; KLAUBERG, C.; MORAES, A. de; SANQUETTA, M.; WILKINSON, B.; BROADBENT, E. Forest inventory with high-density uav-lidar: Machine learning approaches for predicting individual tree attributes. *Computers and Electronics in Agriculture*, v. 179, 2020. Cited By 13. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85092727022&doi=10.1016%2fj.compag.2020.105815&partnerID=40&md5=969d49dd092d6c2b302df16bbae5da85>.

CORTE, A. P. D.; SOUZA, D. V.; REX, F. E.; SANQUETTA, C. R.; MOHAN, M.; SILVA, C. A.; ZAMBRANO, A. M. A.; PRATA, G.; ALMEIDA, D. R. A. de; TRAUTENMÜLLER, J. W. *et al.* Forest inventory with high-density uav-lidar: Machine learning approaches for predicting individual tree attributes. *Computers and Electronics in Agriculture*, Elsevier, v. 179, p. 105815, 2020.

COUTO, H. d.; BASTOS, N. L. M. Modelos de equações de volume e relações hipsométricas para plantações de eucalyptus no estado de são paulo. *IPEF, Piracicaba*, n. 37, p. 33–44, 1987.

CUNHA, U. S. D. *Dendrometria e inventário florestal*. [S.l.]: Manaus, Brasil, 2004.

DASSOT, M.; COLIN, A.; SANTENOISE, P.; FOURNIER, M.; CONSTANT, T. Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. *Computers and Electronics in Agriculture*, Elsevier, v. 89, p. 86–93, 2012.

DASSOT, M.; CONSTANT, T.; FOURNIER, M. The use of terrestrial lidar technology in forest science: application fields, benefits and challenges. *Annals of forest science*, Springer, v. 68, p. 959–974, 2011.

DAVIS, J.; MARSCHNER, S. R.; GARR, M.; LEVOY, M. Filling holes in complex surfaces using volumetric diffusion. In: IEEE. *Proceedings. First international symposium on 3d data processing visualization and transmission*. [S.l.], 2002. p. 428–441.

DIAMANTOPOULOU, M. J. Artificial neural networks as an alternative tool in pine bark volume estimation. *Computers and electronics in agriculture*, Elsevier, v. 48, n. 3, p. 235–244, 2005.

DIAMANTOPOULOU, M. J. Filling gaps in diameter measurements on standing tree boles in the urban forest of thessaloniki, greece. *Environmental Modelling & Software*, Elsevier, v. 25, n. 12, p. 1857–1865, 2010.

DIAMANTOPOULOU, M. J.; ÖZÇELIK, R.; KOPARAN, B.; ALKAN, O. Artificial intelligence as an alternative modelling strategy for reliable height-diameter predictions of mixed-oaks species. *Turkish Journal of Agriculture and Forestry*, v. 47, n. 2, p. 228–241, 2023.

DIEBEL, J.; THRUN, S. An application of markov random fields to range sensing. *Advances in neural information processing systems*, v. 18, 2005.

DONG, C.; LOY, C. C.; HE, K.; TANG, X. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 38, n. 2, p. 295–307, 2015.

DOSOVITSKIY, A.; BEYER, L.; KOLESNIKOV, A.; WEISSENBORN, D.; ZHAI, X.; UNTERTHINER, T.; DEHGHANI, M.; MINDERER, M.; HEIGOLD, G.; GELLY, S. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

DOVRAT, O.; LANG, I.; AVIDAN, S. *Learning to Sample*. 2019. Available at: <https://arxiv.org/abs/1812.01659>.

DUBAYAH, R. O.; DRAKE, J. B. Lidar remote sensing for forestry. *Journal of forestry*, Oxford University Press, v. 98, n. 6, p. 44–46, 2000.

EHLERS, D.; WANG, C.; COULSTON, J.; ZHANG, Y.; PAVELSKY, T.; FRANKENBERG, E.; WOODCOCK, C.; SONG, C. Mapping forest aboveground biomass using multisource remotely sensed data. *Remote Sensing*, MDPI, v. 14, n. 5, p. 1115, 2022.

ENGEL, N.; BELAGIANNIS, V.; DIETMAYER, K. Point transformer. *IEEE Access*, v. 9, p. 134826–134840, 2021.

ERCANLI, İ. Innovative deep learning artificial intelligence applications for predicting relationships between individual tree height and diameter at breast height. *Forest Ecosystems*, Springer, v. 7, p. 1–18, 2020.

ESTEBAN, J.; MCROBERTS, R.; FERNáNDEZ-LANDA, A.; TOMé, J.; MARCHAMALO, M. A model-based volume estimator that accounts for both land cover misclassification and model prediction uncertainty. *Remote Sensing*, v. 12, n. 20, p. 1–25, 2020. Cited By 1. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85092911299&doi=10.3390%2frs12203360&partnerID=40&md5=088f2235c3fe42466e54271f8a9ec523>.

ESTEBAN, J.; MCROBERTS, R. E.; FERNÁNDEZ-LANDA, A.; TOMÉ, J. L.; MARCHAMALO, M. A model-based volume estimator that accounts for both land cover misclassification and model prediction uncertainty. *Remote Sensing*, MDPI, v. 12, n. 20, p. 3360, 2020.

FAN, H.; SU, H.; GUIBAS, L. J. A point set generation network for 3d object reconstruction from a single image. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 605–613.

FAYAD, I.; BAGHDADI, N. N.; ALVARES, C. A.; STAPE, J. L.; BAILLY, J. S.; SCOLFORO, H. F.; ZRIBI, M.; MAIRE, G. L. Assessment of gedi's lidar data for the estimation of canopy heights and wood volume of eucalyptus plantations in brazil. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE, v. 14, p. 7095–7110, 2021.

FEI, B.; YANG, W.; CHEN, W.-M.; LI, Z.; LI, Y.; MA, T.; HU, X.; MA, L. Comprehensive review of deep learning-based 3d point cloud completion processing and analysis. *IEEE Transactions on Intelligent Transportation Systems*, IEEE, v. 23, n. 12, p. 22862–22883, 2022.

FELZENSZWALB, P. F.; GIRSHICK, R. B.; MCALLESTER, D.; RAMANAN, D. Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 32, n. 9, p. 1627–1645, 2009.

FENG, Y.; AUDY, J.-F. Forestry 4.0: a framework for the forest supply chain toward industry 4.0. *Gestão & Produção*, SciELO Brasil, v. 27, p. e5677, 2020.

FERNANDES, D.; SILVA, A.; NÉVOA, R.; SIMÕES, C.; GONZALEZ, D.; GUEVARA, M.; NOVAIS, P.; MONTEIRO, J.; MELO-PINTO, P. Point-cloud based 3d object detection and classification methods for self-driving applications: A survey and taxonomy. *Information Fusion*, Elsevier, v. 68, p. 161–191, 2021.

FERSTL, D.; REINBACHER, C.; RANFTL, R.; RÜTHER, M.; BISCHOF, H. Image guided depth upsampling using anisotropic total generalized variation. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2013. p. 993–1000.

FERSTL, D.; RUTHER, M.; BISCHOF, H. Variational depth superresolution using example-based edge representations. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2015. p. 513–521.

FLETCHER, S.; ISLAM, M. Z. *et al.* Comparing sets of patterns with the Jaccard index. *Australasian Journal of Information Systems*, Australian Computer Society, v. 22, 2018.

FLORIANO, E. *DENDROMETRIA*. [S.l.: s.n.], 2021. ISBN 978-65-00-28337-2.

FREITAS, A. G.; WICHERT, M. C. P. Comparação entre instrumentos tradicionais de medição de diâmetro e altura com o criterion 400. 1998.

GADELHA, M.; WANG, R.; MAJI, S. Multiresolution tree networks for 3d point cloud processing. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. [S.l.: s.n.], 2018. p. 103–118.

GAFFREY, D.; SLOBODA, B.; FABRIKA, M.; ŠMELKO, Š. Terrestrial single-image photogrammetry for measuring standing trees, as applied in the dobroč virgin forest. 2001.

GEIGER, A.; LENZ, P.; STILLER, C.; URTASUN, R. Vision meets robotics: The kitti dataset. *The international journal of robotics research*, Sage Publications Sage UK: London, England, v. 32, n. 11, p. 1231–1237, 2013.

GEIGER, A.; LENZ, P.; URTASUN, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2012.

GIRARDEAU-MONTAUT, D. CloudCompare. *France: EDF R&D Telecom ParisTech*, v. 11, p. 5, 2016.

GIRDHAR, R.; FOUHEY, D. F.; RODRIGUEZ, M.; GUPTA, A. Learning a predictable and generative vector representation for objects. In: SPRINGER. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, the Netherlands, October 11-14, 2016, Proceedings, Part VI 14*. [S.l.], 2016. p. 484–499.

GOMES, A. M. d. A. *Medição dos arvoredos*. [S.l.]: Livraria Sá da Costa, 1957.

GRAYBILL, F. A. *Theory and application of the linear model*. [S.l.: s.n.], 1976.

GREGOR, K.; LECUN, Y. Learning fast approximations of sparse coding. In: *Proceedings of the 27th international conference on international conference on machine learning*. [S.l.: s.n.], 2010. p. 399–406.

GROUEIX, T.; FISHER, M.; KIM, V. G.; RUSSELL, B. C.; AUBRY, M. A papier-mâché approach to learning 3d surface generation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 216–224.

GU, J.; MA, W.-C.; MANIVASAGAM, S.; ZENG, W.; WANG, Z.; XIONG, Y.; SU, H.; URTASUN, R. Weakly-supervised 3d shape completion in the wild. In: SPRINGER. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part V 16*. [S.l.], 2020. p. 283–299.

GUO, C.; LI, C.; GUO, J.; CONG, R.; FU, H.; HAN, P. Hierarchical features driven residual learning for depth map super-resolution. *IEEE Transactions on Image Processing*, IEEE, v. 28, n. 5, p. 2545–2557, 2018.

GUO, F.; GAO, G.; HUANG, J.; WANG, L.; WANG, D. Diameter distribution prediction of populus shelterbelts based on artificial neural network. In: IEEE. *Natural Computation (ICNC), 2011 Seventh International Conference on*. [S.l.], 2011. v. 2, p. 762–765.

GUO, M.-H.; CAI, J.-X.; LIU, Z.-N.; MU, T.-J.; MARTIN, R. R.; HU, S.-M. Pct: Point cloud transformer. *Computational visual media*, Springer, v. 7, n. 2, p. 187–199, 2021.

GUPTA, S.; ARBELÁEZ, P.; GIRSHICK, R.; MALIK, J. Aligning 3d models to rgb-d images of cluttered scenes. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 4731–4740.

HAKAMADA, R. E.; STAPE, J. L.; LEMOS, C. C. Z. d.; EMANUEL, A.; ALMEIDA, A.; SILVA, L. F. Uso do inventário florestal e da uniformidade entre árvores como ferramenta de monitoramento da qualidade silvicultural em plantios clonais de eucalipto. *Scientia Forestalis*, v. 43, n. 105, p. 27–39, 2015.

HAN, F.; ZHU, S.-C. Bottom-up/top-down image parsing with attribute grammar. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 31, n. 1, p. 59–73, 2008.

HAO, Y.; WIDAGDO, F.; LIU, X.; QUAN, Y.; DONG, L.; LI, F. Individual tree diameter estimation in small-scale forest inventory using uav laser scanning. *Remote Sensing*, v. 13, n. 1, p. 1–21, 2021. Cited By 5. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85098724568&doi=10.3390%2frs13010024&partnerID=40&md5=26bf51852e445459f36fef93612dd14d>.

HAO, Y.; WIDAGDO, F. R. A.; LIU, X.; QUAN, Y.; DONG, L.; LI, F. Individual tree diameter estimation in small-scale forest inventory using uav laser scanning. *Remote Sensing*, MDPI, v. 13, n. 1, p. 24, 2020.

HAUGLIN, M.; HANSEN, E.; SØRNGÅRD, E.; NÆSSET, E.; GOBAKKEN, T. Utilizing accurately positioned harvester data: modelling forest volume with airborne laser scanning. *Canadian Journal of Forest Research*, NRC Research Press, v. 48, n. 8, p. 913–922, 2018.

HE, K.; SUN, J.; TANG, X. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 35, n. 6, p. 1397–1409, 2012.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2015. p. 1026–1034.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2016. p. 770–778.

HEINZEL, J.; GINZLER, C. A single-tree processing framework using terrestrial laser scanning data for detecting forest regeneration. *Remote Sensing*, MDPI, v. 11, n. 1, p. 60, 2018.

HIMMELSBACH, M.; LUETTEL, T.; WUENSCHE, H.-J. Real-time object classification in 3d point clouds using point feature histograms. In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. [S.l.: s.n.], 2009. p. 994–1000.

HO, J.; JAIN, A.; ABBEEL, P. *Denoising Diffusion Probabilistic Models*. 2020. Available at: <https://arxiv.org/abs/2006.11239>.

HOLOPAINEN, M.; VASTARANTA, M.; KANKARE, V.; RÄTY, M.; VAAJA, M.; LIANG, X.; YU, X.; HYYPPÄ, J.; HYYPPÄ, H.; VIITALA, R. *et al.* Biomass estimation of individual trees using stem and crown diameter tls measurements. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, v. 38, p. 91–95, 2012.

HOOVER, C. M.; DUCEY, M. J.; COLTER, R. A.; YAMASAKI, M. Evaluation of alternative approaches for landscape-scale biomass estimation in a mixed-species northern forest. *Forest ecology and management*, Elsevier, v. 409, p. 552–563, 2018.

HOWARD, A. G.; ZHU, M.; CHEN, B.; KALENICHENKO, D.; WANG, W.; WEYAND, T.; ANDREETTO, M.; ADAM, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.

HU, H.; ZHANG, Z.; XIE, Z.; LIN, S. Local relation networks for image recognition. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 3464–3473.

HU, T.; HAN, Z.; SHRIVASTAVA, A.; ZWICKER, M. Render4completion: Synthesizing multi-view depth maps for 3d shape completion. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. [S.l.: s.n.], 2019. p. 0–0.

HU, T.; HAN, Z.; ZWICKER, M. 3d shape completion with multi-view consistent inference. In: *Proceedings of the AAAI conference on artificial intelligence*. [S.l.: s.n.], 2020. v. 34, n. 07, p. 10997–11004.

HUANG, J.; ZHAO, J.; GAO, G.; MENG, X.; GUAN, Y. Prediction of stand diameter distribution with artificial neural network. In: IEEE. *Natural Computation, 2009. ICNC'09. Fifth International Conference on*. [S.l.], 2009. v. 2, p. 79–82.

HUANG, Z.; BOSS, M.; VASISHTA, A.; REHG, J. M.; JAMPANI, V. Spar3d: Stable point-aware reconstruction of 3d objects from single images. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. [S.l.: s.n.], 2025. p. 16860–16870.

HUANG, Z.; YU, Y.; XU, J.; NI, F.; LE, X. Pf-net: Point fractal network for 3d point cloud completion. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 7662–7670.

HUDAK, A. T.; FEKETY, P. A.; KANE, V. R.; KENNEDY, R. E.; FILIPPELLI, S. K.; FALKOWSKI, M. J.; TINKHAM, W. T.; SMITH, A. M.; CROOKSTON, N. L.; DOMKE, G. M. *et al.* A carbon monitoring system for mapping regional, annual aboveground biomass across the northwestern usa. *Environmental Research Letters*, IOP Publishing, v. 15, n. 9, p. 095003, 2020.

HUI, T.-W.; LOY, C. C.; TANG, X. Depth map super-resolution by deep multi-scale guidance. In: SPRINGER. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*. [S.l.], 2016. p. 353–369.

HYYPPA, J.; KELLE, O.; LEHIKOINEN, M.; INKINEN, M. A segmentation-based method to retrieve stem volume estimates from 3-D tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 39, n. 5, p. 969–975, 2001.

IBÁ. *Relatório Anual 2024*. Brasília, DF, 2024.

IBÁ. *Relatório Anual IBÁ*. São Paulo: Indústria Brasileira de Árvores São Paulo, Brazil, 2024.

IBÁ, I. B. *Relatório anual 2022*. 2022. Available at: <https://iba.org/datafiles/publicacoes/rela\torios/relatorio-anual-iba2022-compactado.pdf>.

JI, Y.; HAO, J.; REYHANI, N.; LENDASSE, A. Direct and recursive prediction of time series using mutual information selection. In: SPRINGER. *International Work-Conference on Artificial Neural Networks*. [S.l.], 2005. p. 1010–1017.

JIANG, L.; SHI, S.; QI, X.; JIA, J. Gal: Geometric adversarial loss for single-view 3d-object reconstruction. In: *Proceedings of the European conference on computer vision (ECCV)*. [S.l.: s.n.], 2018. p. 802–816.

JMAA, A. B. B.; CHAIEB, F. Low-density 3d point cloud classification. In: IEEE. *2024 IEEE 36th International Conference on Tools with Artificial Intelligence (ICTAI)*. [S.l.], 2024. p. 617–623.

KALOGERAKIS, E.; CHAUDHURI, S.; KOLLER, D.; KOLTUN, V. A probabilistic model for component-based shape synthesis. *Acm Transactions on Graphics (TOG)*, ACM New York, NY, USA, v. 31, n. 4, p. 1–11, 2012.

KEELE, S. *Guidelines for performing systematic literature reviews in software engineering*. [S.l.], 2007.

KIM, J.; YOO, J.; LEE, J.; HONG, S. Setvae: Learning hierarchical composition for generative modeling of set-structured data. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2021. p. 15059–15068.

KIM, V. G.; LI, W.; MITRA, N. J.; CHAUDHURI, S.; DIVERDI, S.; FUNKHOUSER, T. Learning part-based templates from large collections of 3d shapes. *ACM Transactions on Graphics (TOG)*, Acm New York, NY, USA, v. 32, n. 4, p. 1–12, 2013.

KINGMA, D. P.; WELLING, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

KLOKOV, R.; BOYER, E.; VERBEEK, J. Discrete point flow networks for efficient point cloud generation. In: SPRINGER. *European Conference on Computer Vision*. [S.l.], 2020. p. 694–710.

KLOKOV, R.; LEMPITSKY, V. Escape from cells: Deep kd-networks for the recognition of 3d point cloud models. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2017. p. 863–872.

KOPF, J.; COHEN, M. F.; LISCHINSKI, D.; UYTTENDAELE, M. Joint bilateral upsampling. *ACM Transactions on Graphics (ToG)*, ACM New York, NY, USA, v. 26, n. 3, p. 96–es, 2007.

KORPELA, I.; HOVI, A.; MORSDORF, F. Understory trees in airborne lidar data—selective mapping due to transmission losses and echo-triggering mechanisms. *Remote Sensing of Environment*, Elsevier, v. 119, p. 92–104, 2012.

KÜKENBRINK, D.; SCHNEIDER, F. D.; LEITERER, R.; SCHAEPMAN, M. E.; MORSDORF, F. Quantification of hidden canopy volume of airborne laser scanning data using a voxel traversal algorithm. *Remote Sensing of Environment*, Elsevier, v. 194, p. 424–436, 2017.

KUKUNDA, C. B.; DUQUE-LAZO, J.; GONZÁLEZ-FERREIRO, E.; THADEN, H.; KLEINN, C. Ensemble classification of individual pinus crowns from multispectral satellite imagery and airborne lidar. *International journal of applied earth observation and geoinformation*, Elsevier, v. 65, p. 12–23, 2018.

KWON, H.; TAI, Y.-W.; LIN, S. Data-driven depth map refinement via multi-scale sparse representation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 159–167.

LAMB, S. M.; MACLEAN, D. A.; HENNIGAR, C. R.; PITT, D. G. Forecasting forest inventory using imputed tree lists for lidar grid cells and a tree-list growth model. *Forests*, MDPI, v. 9, n. 4, p. 167, 2018.

LANG, I.; MANOR, A.; AVIDAN, S. Samplenet: Differentiable point cloud sampling. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2020. p. 7578–7588.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

LEDIG, C.; THEIS, L.; HUSZÁR, F.; CABALLERO, J.; CUNNINGHAM, A.; ACOSTA, A.; AITKEN, A.; TEJANI, A.; TOTZ, J.; WANG, Z. *et al.* Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 4681–4690.

LEDUC, D. J.; MATNEY, T. G.; BELLI, K. L.; BALDWIN, V. C. Predicting diameter distributions of longleaf pine plantations: a comparison between artificial neural networks and

other accepted methodologies. *Res. Pap. SRS-25. Asheville, NC: US Department of Agriculture, Forest Service, Southern Research Station. 24p.*, v. 25, 2001.

LEE, J.; IM, J.; KIM, K.; QUACKENBUSH, L. Machine learning approaches for estimating forest stand height using plot-based observations and airborne lidar data. *Forests*, v. 9, n. 5, p. 268, May 2018. ISSN 1999-4907.

LEE, J.; IM, J.; KIM, K.; QUACKENBUSH, L. Machine learning approaches for estimating forest stand height using plot-based observations and airborne lidar data. *Forests*, v. 9, n. 5, 2018. Cited By 19. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047016941&doi=10.3390%2ff9050268&partnerID=40&md5=153b38dfee70e47de3a9b968943eb05e>.

LEE, J.; IM, J.; KIM, K.; QUACKENBUSH, L. J. Machine learning approaches for estimating forest stand height using plot-based observations and airborne lidar data. *Forests*, MDPI, v. 9, n. 5, p. 268, 2018.

LEITE, H. G.; ANDRADE, V. C. L. d. Um método para condução de inventários florestais sem o uso de equações volumétricas. SciELO Brasil, 2002.

LI, G.; LIU, L.; ZHENG, H.; MITRA, N. J. Analysis, reconstruction and manipulation using arterial snakes. *ACM Trans. Graph.*, v. 29, n. 6, p. 152, 2010.

LI, X.; SUN, X.; MENG, Y.; LIANG, J.; WU, F.; LI, J. *Dice Loss for Data-imbalanced NLP Tasks*. 2020. Available at: <https://arxiv.org/abs/1911.02855>.

LI, Y.; BU, R.; SUN, M.; WU, W.; DI, X.; CHEN, B. Pointcnn: Convolution on x-transformed points. *Advances in neural information processing systems*, v. 31, 2018.

LI, Y.; DAI, A.; GUIBAS, L.; NIESSNER, M. Database-assisted object retrieval for real-time 3d reconstruction. In: WILEY ONLINE LIBRARY. *Computer graphics forum*. [S.l.], 2015. v. 34, n. 2, p. 435–446.

LI, Y.; MA, L.; ZHONG, Z.; LIU, F.; CHAPMAN, M. A.; CAO, D.; LI, J. Deep learning for lidar point clouds in autonomous driving: A review. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, v. 32, n. 8, p. 3412–3432, 2020.

LI, Y.; MIN, D.; DO, M. N.; LU, J. Fast guided global interpolation for depth and motion. In: SPRINGER. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*. [S.l.], 2016. p. 717–733.

LI, Z.; YANG, J.; LIU, Z.; YANG, X.; JEON, G.; WU, W. Feedback network for image super-resolution. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 3867–3876.

LIAO, Y.; XIE, J.; GEIGER, A. *KITTI-360: A Novel Dataset and Benchmarks for Urban Scene Understanding in 2D and 3D*. 2022. Available at: <https://arxiv.org/abs/2109.13410>.

LIN, T.-Y.; GOYAL, P.; GIRSHICK, R.; HE, K.; DOLLÁR, P. Focal loss for dense object detection. In: *IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2017. p. 2980–2988.

LIU, M.-Y.; TUZEL, O.; TAGUCHI, Y. Joint geodesic upsampling of depth images. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2013. p. 169–176.

LIU, Q.; YUE, H.; LYU, Z.; WANG, W.; LIU, Z.; CHEN, W. Sehlnet: separate estimation of high-and low-frequency components for depth completion. In: IEEE. *2022 International Conference on Robotics and Automation (ICRA)*. [S.l.], 2022. p. 668–674.

LOHANI, B.; GHOSH, S. Airborne lidar technology: A review of data collection and processing systems. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, Springer, v. 87, n. 4, p. 567–579, 2017.

LOPEZ-PAZ, D.; OQUAB, M. *Revisiting Classifier Two-Sample Tests*. 2018. Available at: <https://arxiv.org/abs/1610.06545>.

LU, J.; FORSYTH, D. Sparse depth super resolution. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 2245–2253.

LUETZENBURG, G.; KROON, A.; BJØRK, A. A. Evaluation of the apple iphone 12 pro lidar for an application in geosciences. *Scientific reports*, Nature Publishing Group, v. 11, n. 1, p. 1–9, 2021.

LUO, S.; HU, W. *Diffusion Probabilistic Models for 3D Point Cloud Generation*. 2021. Available at: <https://arxiv.org/abs/2103.01458>.

MA, X.; QIN, C.; YOU, H.; RAN, H.; FU, Y. Rethinking network design and local geometry in point cloud: A simple residual mlp framework. *arXiv preprint arXiv:2202.07123*, 2022.

MACHADO PIRES, L.; CALEGARIO, N. Fitting linear and nonlinear stochastic models to describe longitudinal tree profiles. *Revista Árvore*, SciELO Brasil, v. 31, n. 5, p. 845–852, 2007.

MAHONEY, C.; HALL, R. J.; HOPKINSON, C.; FILIATRAULT, M.; BEAUDOIN, A.; CHEN, Q. A forest attribute mapping framework: A pilot study in a northern boreal forest, northwest territories, canada. *Remote sensing*, MDPI, v. 10, n. 9, p. 1338, 2018.

MALTAMO, M.; SUVANTO, A.; PACKALÉN, P. Comparison of basal area and stem frequency diameter distribution modelling using airborne laser scanner data and calibration estimation. *Forest Ecology and Management*, Elsevier, v. 247, n. 1-3, p. 26–34, 2007.

MANDAL, S.; BHAVSAR, A.; SAO, A. K. Depth map restoration from undersampled data. *IEEE Transactions on Image Processing*, IEEE, v. 26, n. 1, p. 119–134, 2016.

MARTINOVIC, A.; GOOL, L. V. Bayesian grammar learning for inverse procedural modeling. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2013. p. 201–208.

MATURANA, D.; SCHERER, S. Voxnet: A 3d convolutional neural network for real-time object recognition. In: IEEE. *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. [S.l.], 2015. p. 922–928.

MAULONI, J. A.; DISPERATI, A. A.; MACHADO, Á. M. L.; MITISHITA, E. A. Investigação fotogramétrica em imagens digitais para cálculos dendrométricos. *XV Simpósio Brasileiro de Sensoriamento Remoto*, 2011.

MAYRA, J.; KESKI-SAARI, S.; KIVINEN, S.; TANHUANPÄÄ, T.; HURSKAINEN, P.; KULLBERG, P.; POIKOLAINEN, L.; VIINIKKA, A.; TUOMINEN, S.; KUMPULA, T. *et al.* Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sensing of Environment*, Elsevier, v. 256, p. 112322, 2021.

MBUNGE, E.; SIMELANE, S.; FASHOTO, S. G.; AKINNUWESI, B.; METFULA, A. S. Application of deep learning and machine learning models to detect covid-19 face masks-a review. *Sustainable Operations and Computers*, Elsevier, v. 2, p. 235–245, 2021.

MEDINA, F. P.; PAFFENROTH, R. Machine learning in lidar 3d point clouds. *Advances in Data Science*, Springer, p. 113–133, 2021.

MOMBACH, J.; FONSECA, A. U.; NASCIMENTO THAMER H RODRIGUES, W. G. R.; GRESSLER, H.; ROSSI, F. D.; SOARES, F. A new approach to performing paper-based children's spelling tests on mobile devices. In: IEEE. *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. [S.l.], 2020. p. 4346–4351.

MORITZ, P.; NISHIHARA, R.; JORDAN, M. A linearly-convergent stochastic l-bfgs algorithm. In: *Artificial Intelligence and Statistics*. [S.l.: s.n.], 2016. p. 249–258.

NAN, L.; SHARF, A.; ZHANG, H.; COHEN-OR, D.; CHEN, B. Smartboxes for interactive urban reconstruction. In: *ACM Siggraph 2010 papers*. [S.l.: s.n.], 2010. p. 1–10.

NAN, L.; XIE, K.; SHARF, A. A search-classify approach for cluttered indoor scene understanding. *ACM Transactions on Graphics (TOG)*, Acm New York, NY, USA, v. 31, n. 6, p. 1–10, 2012.

NASCIMENTO, T. H.; FERREIRA, C. B.; RODRIGUES, W. G.; SOARES, F. Interaction with smartwatches using gesture recognition: a systematic literature review. In: IEEE. *2020 IEEE 44th annual computers, software, and applications conference (COMPSAC)*. [S.l.], 2020. p. 1661–1666.

NEZHADARYA, E.; TAGHAVI, E.; RAZANI, R.; LIU, B.; LUO, J. Adaptive hierarchical down-sampling for point cloud classification. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2020. p. 12956–12964.

NGUYEN, A.-D.; CHOI, S.; KIM, W.; KIM, J.; OH, H.; KANG, J.; LEE, S. Single-image 3-d reconstruction: Rethinking point cloud deformation. *IEEE Transactions on Neural Networks and Learning Systems*, v. 35, n. 5, p. 6613–6627, 2024.

NGUYEN, D. T.; HUA, B.-S.; TRAN, K.; PHAM, Q.-H.; YEUNG, S.-K. A field model for repairing 3d shapes. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2016. p. 5676–5684.

NGUYEN, V.-T.; CONSTANT, T.; COLIN, F. An innovative and automated method for characterizing wood defects on trunk surfaces using high-density 3d terrestrial lidar data. *Annals of Forest Science*, BioMed Central, v. 78, n. 2, p. 1–18, 2021.

NIE, D.; WANG, L.; ADELI, E.; LAO, C.; LIN, W.; SHEN, D. 3-D fully convolutional networks for multimodal isointense infant brain image segmentation. *IEEE Transactions on Cybernetics*, IEEE, v. 49, n. 3, p. 1123–1136, 2018.

NIE, Y.; LIN, Y.; HAN, X.; GUO, S.; CHANG, J.; CUI, S.; ZHANG, J. *et al.* Skeleton-bridged point completion: From global inference to local adjustment. *Advances in Neural Information Processing Systems*, v. 33, p. 16119–16130, 2020.

NISKA, H.; SKON, J.-P.; PACKALEN, P.; TOKOLA, T.; MALTAMO, M.; KOLEHMAINEN, M. Neural networks for the prediction of species-specific plot volumes using airborne laser scanning and aerial photographs. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 48, n. 3, p. 1076–1085, 2010.

NOVO-FERNÁNDEZ, A.; BARRIO-ANTA, M.; RECONDO, C.; CÁMARA-OBREGÓN, A.; LÓPEZ-SÁNCHEZ, C. A. Integration of national forest inventory and nationwide airborne laser scanning data to improve forest yield predictions in north-western spain. *Remote Sensing*, MDPI, v. 11, n. 14, p. 1693, 2019.

OBATA, S.; CIESZEWSKI, C. J.; III, R. C. L.; BETTINGER, P. Random forest regression model for estimation of the growing stock volumes in georgia, usa, using dense landsat time series and fia dataset. *Remote Sensing*, MDPI, v. 13, n. 2, p. 218, 2021.

OEHMCKE, S.; LI, L.; REVENGA, J. C.; NORD-LARSEN, T.; TREPEKLI, K.; GIESEKE, F.; IGEL, C. Deep learning based 3d point cloud regression for estimating forest biomass. In: *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*. [S.l.: s.n.], 2022. p. 1–4.

OLIVEIRA, L. T. d.; CARVALHO, L. M. T. d.; FERREIRA, M. Z.; OLIVEIRA, T. C. d. A.; JUNIOR, F. W. A. Application of lidar to forest inventory for tree count in stands of eucalyptus sp. *Cerne*, SciELO Brasil, v. 18, p. 175–184, 2012.

ÖZÇELIK, R.; DIAMANTOPOULOU, M. J.; BROOKS, J. R.; JR, H. V. W. Estimating tree bole volume using artificial neural network models for four species in turkey. *Journal of environmental management*, Elsevier, v. 91, n. 3, p. 742–753, 2010.

OZKAN, U. Y.; DEMIREL, T.; OZDEMIR, I.; SAGLAM, S.; MERT, A. Predicting forest stand attributes using the integration of airborne laser scanning and worldview-3 data in a mixed forest in turkey. *Advances in Space Research*, Elsevier, v. 69, n. 2, p. 1146–1158, 2022.

PAN, X.; XIA, Z.; SONG, S.; LI, L. E.; HUANG, G. 3d object detection with pointformer. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2021. p. 7463–7472.

PARK, J.; JOO, K.; HU, Z.; LIU, C.-K.; KWEON, I. S. Non-local spatial propagation network for depth completion. In: SPRINGER. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16*. [S.l.], 2020. p. 120–136.

PARK, J.; KIM, H.; TAI, Y.-W.; BROWN, M. S.; KWEON, I. High quality depth map upsampling for 3d-tof cameras. In: IEEE. *2011 International Conference on Computer Vision*. [S.l.], 2011. p. 1623–1630.

PARK, S.-H.; JUNG, H.-S.; LEE, S.; KIM, E.-S. Mapping forest vertical structure in sogwang-ri forest from full-waveform lidar point clouds using deep neural network. *Remote Sensing*, MDPI, v. 13, n. 18, p. 3736, 2021.

PARMAR, N.; VASWANI, A.; USZKOREIT, J.; KAISER, L.; SHAZEER, N.; KU, A.; TRAN, D. Image transformer. In: PMLR. *International conference on machine learning*. [S.l.], 2018. p. 4055–4064.

PAULY, M.; MITRA, N. J.; GIESEN, J.; GROSS, M. H.; GUIBAS, L. J. Example-based 3d scan completion. In: *Symposium on geometry processing*. [S.l.: s.n.], 2005. p. 23–32.

PEARSE, G. D.; WATT, M. S.; DASH, J. P.; STONE, C.; CACCAMO, G. Comparison of models describing forest inventory attributes using standard and voxel-based lidar predictors across a range of pulse densities. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 78, p. 341–351, 2019.

PERSSON, H. J.; OLOFSSON, K.; HOLMGREN, J. Two-phase forest inventory using very-high-resolution laser scanning. *Remote Sensing of Environment*, Elsevier, v. 271, p. 112909, 2022.

PETERSEN, K.; VAKKALANKA, S.; KUZNIARZ, L. Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, Elsevier, v. 64, p. 1–18, 2015.

PICOS, J.; BASTOS, G.; MÍGUEZ, D.; ALONSO, L.; ARMESTO, J. Individual tree detection in a eucalyptus plantation using unmanned aerial vehicle (uav)-lidar. *Remote Sensing*, MDPI, v. 12, n. 5, p. 885, 2020.

QI, C. R.; SU, H.; MO, K.; GUIBAS, L. J. Pointnet: Deep learning on point sets for 3d classification and segmentation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 652–660.

QI, C. R.; YI, L.; SU, H.; GUIBAS, L. J. PointNet++: Deep hierarchical feature learning on point sets in a metric space. *Advances in Neural Information Processing Systems*, v. 30, 2017.

QIAN, G.; LI, Y.; PENG, H.; MAI, J.; HAMMOUD, H.; ELHOSEINY, M.; GHANEM, B. Pointnext: Revisiting pointnet++ with improved training and scaling strategies. *Advances in neural information processing systems*, v. 35, p. 23192–23204, 2022.

QIU, S.; ANWAR, S.; BARNES, N. Dense-resolution network for point cloud classification and segmentation. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. [S.l.: s.n.], 2021. p. 3813–3822.

QUEINNEC, M.; TOMPALSKI, P.; BOLTON, D. K.; COOPS, N. C. Foster—an r package for forest structure extrapolation. *Plos one*, Public Library of Science San Francisco, CA USA, v. 16, n. 1, p. e0244846, 2021.

RAMACHANDRAN, P.; PARMAR, N.; VASWANI, A.; BELLO, I.; LEVSKAYA, A.; SHLENS, J. Stand-alone self-attention in vision models. *Advances in neural information processing systems*, v. 32, 2019.

REN, S.; HE, K.; GIRSHICK, R.; SUN, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, v. 28, 2015.

RIEGLER, G.; FERSTL, D.; RÜTHER, M.; BISCHOF, H. A deep primal-dual network for guided depth super-resolution. *arXiv preprint arXiv:1607.08569*, 2016.

RIEGLER, G.; ULUSOY, A. O.; GEIGER, A. Octnet: Learning deep 3d representations at high resolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 3577–3586.

ROCHA, S. J. S. S. d.; ROMERO, F. M. B.; TORRES, C. M. M. E.; JACOVINE, L. A. G.; RIBEIRO, S. C.; VILLANOVA, P. H.; SCHETTINI, B. L. S.; JUNIOR, V. T. M. d. M.; REIS, L. P.; RUFINO, M. P. M. X. *et al.* Machine learning: Volume and biomass estimates of commercial trees in the amazon forest. *Sustainability*, MDPI, v. 15, n. 12, p. 9452, 2023.

ROCK, J.; GUPTA, T.; THORSEN, J.; GWAK, J.; SHIN, D.; HOIEM, D. Completing 3d object shape from one depth image. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 2484–2493.

RODRIGUES, W. G.; NOGUEIRA, E. A.; NASCIMENTO, T. H.; VIEIRA, G. S.; FERNANDES, D. S.; SOARES, F. Mobilenetv2-enhanced depth map super-resolution through multi-scale image guidance. In: SPRINGER. *International Symposium on Visual Computing*. [S.l.], 2024. p. 347–358.

RODRIGUES, W. G.; VIEIRA, G. S.; CABACINHA, C. D.; BULCÃO-NETO, R. F.; SOARES, F. Applications of artificial intelligence and lidar in forest inventories: A systematic literature review. *Computers and Electrical Engineering*, Elsevier, v. 120, p. 109793, 2024.

RODRIGUES, W. G.; VIEIRA, G. S.; CABACINHA, C. D.; SOARES, F. Eucalyptus diameter and volume prediction with deep neural networks: A long short-term memory model approach. *Expert Systems with Applications*, Elsevier, v. 271, p. 126704, 2025.

RUSU, R. B.; COUSINS, S. 3d is here: Point cloud library (pcl). In: IEEE. *2011 IEEE international conference on robotics and automation*. [S.l.], 2011. p. 1–4.

SACHDEVA, M.; KUSHWAHA, A. K. S. *et al.* The power of deep learning for intelligent tumor classification systems: A review. *Computers and Electrical Engineering*, Elsevier, v. 106, p. 108586, 2023.

SANCHEZ-RUIZ, S.; MORENO-MARTINEZ, A.; IZQUIERDO-VERDIGUIER, E.; CHIESI, M.; MASELLI, F.; GILABERT, M. A. Growing stock volume from multi-temporal landsat imagery through google earth engine. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 83, p. 101913, 2019.

SANDLER, M.; HOWARD, A.; ZHU, M.; ZHMOGINOV, A.; CHEN, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 4510–4520.

SCHARSTEIN, D.; PAL, C. Learning conditional random fields for stereo. In: IEEE. *2007 IEEE conference on computer vision and pattern recognition*. [S.l.], 2007. p. 1–8.

SCHARSTEIN, D.; SZELISKI, R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International journal of computer vision*, Springer, v. 47, p. 7–42, 2002.

SCHNABEL, R.; DEGENER, P.; KLEIN, R. Completion and reconstruction with primitive shapes. In: WILEY ONLINE LIBRARY. *Computer graphics forum*. [S.l.], 2009. v. 28, n. 2, p. 503–512.

SCHNEIDER, P. R.; FINGER, C. A. G.; KLEIN, J. E. M.; TOTTI, J. A.; BAZZO, J. L. Forma de tronco e sortimentos de madeira de eucalyptus grandis maiden para o estado do rio grande do sul. *Ciência Florestal*, v. 6, n. 1, p. 79–88, 1996.

SCOLFORO, J.; MACHADO, S.; SILVA, S. O manejo da vegetação nativa através de cortes seletivos. *EMBRAPA-CNPF. Documentos*, In: CURSO DE MANEJO FLORESTAL SUSTENTAVEL, Curitiba. Tópicos em manejo florestal sustentável. Colombo: EMBRAPA-CNPF. p. 127-155., 1997.

SHANG, C.; TREITZ, P.; CASPERSEN, J.; JONES, T. Estimation of forest structural and compositional variables using als data and multi-seasonal satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 78, p. 360–371, 2019.

SHAO, T.; XU, W.; ZHOU, K.; WANG, J.; LI, D.; GUO, B. An interactive approach to semantic modeling of indoor scenes with an rgbd camera. *ACM Transactions on Graphics (TOG)*, ACM New York, NY, USA, v. 31, n. 6, p. 1–11, 2012.

SHEN, C.-H.; FU, H.; CHEN, K.; HU, S.-M. Structure recovery by part assembly. *ACM Transactions on Graphics (TOG)*, ACM New York, NY, USA, v. 31, n. 6, p. 1–11, 2012.

SHENDRYK, Y.; GORROD, E. Leveraging airborne lidar data and gradient boosting for mapping the density of different sized trees. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE, v. 14, p. 1572–1579, 2020.

SHI, W.; RAJKUMAR, R. Point-gnn: Graph neural network for 3d object detection in a point cloud. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 1711–1719.

SHOOT, C.; ANDERSEN, H.-E.; MOSKAL, L. M.; BABCOCK, C.; COOK, B. D.; MORTON, D. C. Classifying forest type in the national forest inventory context with airborne hyperspectral and lidar data. *Remote Sensing*, MDPI, v. 13, n. 10, p. 1863, 2021.

SILVA, C. A.; KLAUBERG, C.; HUDAK, A. T.; VIERLING, L. A.; LIESENBERG, V.; BERNETT, L. G.; SCHERAIBER, C. F.; SCHOENINGER, E. R. Estimating stand height and tree density in pinus taeda plantations using in-situ data, airborne lidar and k-nearest neighbor imputation. *Anais da Academia Brasileira de Ciências*, SciELO Brasil, v. 90, p. 295–309, 2018.

SILVA, M. L. Marques da; BINOTI, D. H. B.; GLERIANI, J. M.; LEITE, H. G. Adjustment of the schumacher and hall model and application of artificial neural networks to estimate volume of eucalypt trees. *Revista Árvore*, SciELO Brasil, v. 33, n. 6, p. 1133–1139, 2009.

SIMARD, P. Y.; STEINKRAUS, D.; PLATT, J. C. *et al.* Best practices for convolutional neural networks applied to visual document analysis. In: *Icdar*. [S.l.: s.n.], 2003. v. 3, n. 2003.

SIMONOVSKY, M.; KOMODAKIS, N. Dynamic edge-conditioned filters in convolutional neural networks on graphs. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 3693–3702.

SIMONSE, M.; ASCHOFF, T.; SPIECKER, H.; THIES, M. Automatic determination of forest inventory parameters using terrestrial laser scanning. In: CITESEER. *Proceedings of the scandlaser scientific workshop on airborne laser scanning of forests*. [S.l.], 2003. v. 2003, p. 252–258.

SOARES, F. A. A.; FLÔRES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction and volume calculation of eucalyptus trees using multilayer perceptron networks. *Computers and electronics in agriculture*, Elsevier, v. 78, n. 1, p. 19–27, 2011.

SOARES, F. A. A.; FLÔRES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction for calculating merchantable volume of eucalyptus clones without previous knowledge of total tree height using artificial neural networks. *Applied Soft Computing*, Elsevier, v. 12, n. 8, p. 2030–2039, 2012.

SOARES, F. A. A.; FLORES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction for calculating merchantable volume of eucalyptus clones using multilayer perceptron. *Neural Computing and Applications*, Springer, v. 22, n. 7-8, p. 1407–1418, 2013.

SOARES, F. A. A. M. N.; FLORES, E. L.; CABACINHA, C. D.; CARRIJO, G. A.; VEIGA, A. C. P. Recursive diameter prediction and volume calculation of eucalyptus trees using multilayer perceptron networks. *Computers and electronics in agriculture*, Elsevier, v. 78, n. 1, p. 19–27, 2011.

SONG, X.; PENG, Z.; SONG, S.; STOJANOVIC, V. Anti-disturbance state estimation for pdt-switched rdnns utilizing time-sampling and space-splitting measurements. *Communications in Nonlinear Science and Numerical Simulation*, Elsevier, v. 132, p. 107945, 2024.

SORKINE, O.; COHEN-OR, D. Least-squares meshes. In: IEEE. *Proceedings Shape Modeling Applications, 2004.* [S.l.], 2004. p. 191–199.

SRINIVASAN, S.; POPESCU, S. C.; ERIKSSON, M.; SHERIDAN, R. D.; KU, N.-W. Terrestrial laser scanning as an effective tool to retrieve tree level height, crown width, and stem diameter. *Remote Sensing*, MDPI, v. 7, n. 2, p. 1877–1896, 2015.

SRIVASTAVA, S.; SHARMA, G. Exploiting local geometry for feature and graph construction for better 3D point cloud processing with graph neural networks. In: IEEE. *IEEE International Conference on Robotics and Automation*. [S.l.], 2021. p. 12903–12909.

SUDRE, C. H.; LI, W.; VERCAUTEREN, T.; OURSELIN, S.; CARDOSO, M. J. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In: SPRINGER. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, and 7th International Workshop, Held in Conjunction with MICCAI 2017*. Québec City, QC, Canada, 2017. p. 240–248.

SUN, Y.; JIN, X.; PUKKALA, T.; LI, F. Predicting individual tree diameter of larch (larix olgensis) from uav-lidar data using six different algorithms. *Remote Sensing*, MDPI, v. 14, n. 5, p. 1125, 2022.

SUN, Y.; WANG, Y.; LIU, Z.; SIEGEL, J. E.; SARMA, S. E. *PointGrow: Autoregressively Learned Point Cloud Generation with Self-Attention*. 2019. Available at: <https://arxiv.org/abs/1810.05591>.

TELLING, J.; LYDA, A.; HARTZELL, P.; GLENNIE, C. Review of earth science research using terrestrial laser scanning. *Earth-Science Reviews*, Elsevier, v. 169, p. 35–68, 2017.

THIES, M.; SPIECKER, H. Evaluation and future prospects of terrestrial laser scanning for standardized forest inventories. *Forest*, v. 2, n. 2.2, p. 1, 2004.

THOMAS, H.; QI, C. R.; DESCHAUD, J.-E.; MARCOTEGUI, B.; GOULETTE, F.; GUIBAS, L. J. Kpconv: Flexible and deformable convolution for point clouds. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 6411–6420.

TIJERÍN-TRIVIÑO, J.; MORENO-FERNÁNDEZ, D.; ZAVALA, M. A.; ASTIGARRAGA, J.; GARCÍA, M. Identifying forest structural types along an aridity gradient in peninsular spain: Integrating low-density lidar, forest inventory, and aridity index. *Remote Sensing*, Multidisciplinary Digital Publishing Institute, v. 14, n. 1, p. 235, 2022.

TIMOFTE, R.; SMET, V. D.; GOOL, L. V. Anchored neighborhood regression for fast example-based super-resolution. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2013. p. 1920–1927.

TKACHENKO, M.; MALYUK, M.; HOLMANYUK, A.; LIUBIMOV, N. *Label Studio: Data labeling software*. 2022. Open source software available from https://github.com/heartexlabs/label-studio. Available at: <https://github.com/heartexlabs/label-studio>.

TUNÇEL, M.; DURAN, A. Effectiveness of grid and random approaches for a model parameter vector optimization. *Journal of Computational Science*, Elsevier, v. 67, p. 101960, 2023.

UNAL, O.; GOOL, L. V.; DAI, D. Improving point cloud semantic segmentation by learning 3d object detection. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. [S.l.: s.n.], 2021. p. 2950–2959.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. *Advances in neural information processing systems*, v. 30, 2017.

VIEIRA, G. C.; MENDONÇA, A. R. de; SILVA, G. F. da; ZANETTI, S. S.; SILVA, M. M. da; SANTOS, A. R. dos. Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. *Science of The Total Environment*, Elsevier, v. 619, p. 1473–1481, 2018.

VOYNOV, O.; ARTEMOV, A.; EGIAZARIAN, V.; NOTCHENKO, A.; BOBROVSKIKH, G.; BURNAEV, E.; ZORIN, D. Perceptual deep depth super-resolution. In: *Proceedings of the ieee/cvf international conference on computer vision*. [S.l.: s.n.], 2019. p. 5653–5663.

WANG, J.; CUI, Y.; GUO, D.; LI, J.; LIU, Q.; SHEN, C. Pointattn: You only need attention for point cloud completion. In: *Proceedings of the AAAI Conference on artificial intelligence*. [S.l.: s.n.], 2024. v. 38, n. 6, p. 5472–5480.

WANG, S.; SUO, S.; MA, W.-C.; POKROVSKY, A.; URTASUN, R. Deep parametric continuous convolutional neural networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 2589–2597.

WANG, S.; ZHAO, X.; XU, H.-M.; CHEN, Z.; YU, D.; CHANG, J.; YANG, Z.; ZHAO, F. Towards Domain Generalization for Multi-view 3D Object Detection in Bird-Eye-View. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2023. p. 13333–13342.

WANG, X.; JR, M. H. A.; LEE, G. H. Cascaded refinement network for point cloud completion. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 790–799.

WANG, Y.; SUN, Y.; LIU, Z.; SARMA, S. E.; BRONSTEIN, M. M.; SOLOMON, J. M. Dynamic graph cnn for learning on point clouds. *ACM Transactions on Graphics (tog)*, Acm New York, NY, USA, v. 38, n. 5, p. 1–12, 2019.

WANG, Z.; LIU, D.; YANG, J.; HAN, W.; HUANG, T. Deep networks for image super-resolution with sparse prior. In: *Proceedings of the IEEE international conference on computer vision*. [S.l.: s.n.], 2015. p. 370–378.

WANG, Z.; YE, X.; SUN, B.; YANG, J.; XU, R.; LI, H. Depth upsampling based on deep edge-aware learning. *Pattern Recognition*, Elsevier, v. 103, p. 107274, 2020.

WATSON, C.; COOPER, N.; PALACIO, D. N.; MORAN, K.; POSHYVANYK, D. A systematic literature review on the use of deep learning in software engineering research. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, ACM New York, NY, v. 31, n. 2, p. 1–58, 2022.

WEISS, K.; KHOSHGOFTAAR, T. M.; WANG, D. A survey of transfer learning. *Journal of Big data*, Springer, v. 3, p. 1–40, 2016.

WINDRIM, L.; BRYSON, M. Forest tree detection and segmentation using high resolution airborne LiDAR. In: IEEE. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. [S.l.], 2019. p. 3898–3904.

WINDRIM, L.; BRYSON, M. Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning. *Remote Sensing*, v. 12, n. 9, p. 1469, May 2020. ISSN 2072-4292.

WINDRIM, L.; BRYSON, M. Detection, segmentation, and model fitting of individual tree stems from airborne laser scanning of forests using deep learning. *Remote Sensing*, MDPI, v. 12, n. 9, p. 1469, 2020.

WITTKE, S.; YU, X.; KARJALAINEN, M.; HYYPPÄ, J.; PUTTONEN, E. Comparison of two-dimensional multitemporal sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 76, p. 167–178, 2019.

WU, W.; QI, Z.; FUXIN, L. Pointconv: Deep convolutional networks on 3d point clouds. In: *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 9621–9630.

WU, Z.; SONG, S.; KHOSLA, A.; YU, F.; ZHANG, L.; TANG, X.; XIAO, J. 3d shapenets: A deep representation for volumetric shapes. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 1912–1920.

XIAN, C.; QIAN, K.; ZHANG, Z.; WANG, C. C. Multi-scale progressive fusion learning for depth map super-resolution. *arXiv preprint arXiv:2011.11865*, 2020.

XIANG, P.; WEN, X.; LIU, Y.-S.; CAO, Y.-P.; WAN, P.; ZHENG, W.; HAN, Z. Snowflakenet: Point cloud completion by snowflake point deconvolution with skip-transformer. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2021. p. 5499–5509.

XIANG, P.; WEN, X.; LIU, Y.-S.; CAO, Y.-P.; WAN, P.; ZHENG, W.; HAN, Z. *SnowflakeNet: Point Cloud Completion by Snowflake Point Deconvolution with Skip-Transformer*. 2021. Available at: <https://arxiv.org/abs/2108.04444>.

XIAO, R.; CARANDE, R.; GHIGLIA, D. A neural network approach for tree height estimation using ifsar data. In: IEEE. *Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS'98. 1998 IEEE International*. [S.l.], 1998. v. 3, p. 1565–1567.

XIE, B.; CAO, C.; XU, M.; BASHIR, B.; SINGH, R. P.; HUANG, Z.; LIN, X. Regional forest volume estimation by expanding lidar samples using multi-sensor satellite data. *Remote Sensing*, v. 12, n. 3, p. 360, Jan 2020. ISSN 2072-4292.

XIE, B.; CAO, C.; XU, M.; BASHIR, B.; SINGH, R. P.; HUANG, Z.; LIN, X. Regional forest volume estimation by expanding lidar samples using multi-sensor satellite data. *Remote Sensing*, MDPI, v. 12, n. 3, p. 360, 2020.

XIE, H.; YAO, H.; ZHOU, S.; MAO, J.; ZHANG, S.; SUN, W. Grnet: Gridding residual network for dense point cloud completion. In: SPRINGER. *European conference on computer vision*. [S.l.], 2020. p. 365–381.

XIE, J.; FERIS, R. S.; SUN, M.-T. Edge-guided single depth image super resolution. *IEEE Transactions on Image Processing*, IEEE, v. 25, n. 1, p. 428–438, 2015.

YANG, G.; HUANG, X.; HAO, Z.; LIU, M.-Y.; BELONGIE, S.; HARIHARAN, B. Pointflow: 3d point cloud generation with continuous normalizing flows. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2019. p. 4541–4550.

YANG, J.; WRIGHT, J.; HUANG, T. S.; MA, Y. Image super-resolution via sparse representation. *IEEE transactions on image processing*, IEEE, v. 19, n. 11, p. 2861–2873, 2010.

YANG, J.; ZHANG, Q.; NI, B.; LI, L.; LIU, J.; ZHOU, M.; TIAN, Q. Modeling point clouds with self-attention and gumbel subset sampling. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2019. p. 3323–3332.

YANG, Q.; YANG, R.; DAVIS, J.; NISTÉR, D. Spatial-depth super resolution for range images. In: IEEE. *2007 IEEE conference on computer vision and pattern recognition*. [S.l.], 2007. p. 1–8.

YANG, R.; HE, Y.; LU, X.; ZHAO, Y.; LI, Y.; YANG, Y.; KONG, W.; LIU, F. 3d-based precise evaluation pipeline for maize ear rot using multi-view stereo reconstruction and point cloud semantic segmentation. *Computers and Electronics in Agriculture*, Elsevier, v. 216, p. 108512, 2024.

YANG, Y.; FENG, C.; SHEN, Y.; TIAN, D. Foldingnet: Point cloud auto-encoder via deep grid deformation. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 206–215.

YANG, Z.; SUN, Y.; LIU, S.; JIA, J. 3dssd: Point-based 3d single stage object detector. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 11040–11048.

YAO, J.; SHEPPERD, M. Assessing software defection prediction performance: Why using the Matthews correlation coefficient matters. In: *24th International Conference on Evaluation and Assessment in Software Engineering*. [S.l.: s.n.], 2020. p. 120–129.

YI, L.; KIM, V. G.; CEYLAN, D.; SHEN, I.-C.; YAN, M.; SU, H.; LU, C.; HUANG, Q.; SHEFFER, A.; GUIBAS, L. A scalable active framework for region annotation in 3d shape collections. *ACM Transactions on Graphics (ToG)*, ACM New York, NY, USA, v. 35, n. 6, p. 1–12, 2016.

YIN, K.; HUANG, H.; ZHANG, H.; GONG, M.; COHEN-OR, D.; CHEN, B. Morfit: interactive surface reconstruction from incomplete point clouds with curve-driven topology and geometry control. *ACM Trans. Graph.*, Citeseer, v. 33, n. 6, p. 202–1, 2014.

YU, X.; RAO, Y.; WANG, Z.; LIU, Z.; LU, J.; ZHOU, J. Pointr: Diverse point cloud completion with geometry-aware transformers. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2021. p. 12498–12507.

YU, X.; TANG, L.; RAO, Y.; HUANG, T.; ZHOU, J.; LU, J. Point-bert: Pre-training 3d point cloud transformers with masked point modeling. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2022. p. 19313–19322.

YUAN, W.; KHOT, T.; HELD, D.; MERTZ, C.; HEBERT, M. Pcn: Point completion network. In: IEEE. *2018 international conference on 3D vision (3DV)*. [S.l.], 2018. p. 728–737.

YUAN, W.; KHOT, T.; HELD, D.; MERTZ, C.; HEBERT, M. *PCN: Point Completion Network*. 2019. Available at: <https://arxiv.org/abs/1808.00671>.

ZAMORSKI, M.; ZIEBA, M.; NOWAK, R.; STOKOWIEC, W.; TRZCINSKI, T. Adversarial autoencoders for generating 3d point clouds. *arXiv preprint arXiv:1811.07605*, v. 2, n. 3, 2018.

ZEA, A.; HANEBECK, U. Modeling Spatial Uncertainty for the iPad Pro Depth Sensor. *Journal of Advances in Information Fusion*, v. 17, n. 2, 2022.

ZENG, P.; PRESTA, A.; REINIS, J.; BHARADIA, D.; QIU, H.; COSMAN, P. *Can We Remove the Ground? Obstacle-aware Point Cloud Compression for Remote Object Detection*. 2024. Available at: <https://arxiv.org/abs/2410.00582>.

ZENG, Y.; HU, Y.; LIU, S.; YE, J.; HAN, Y.; LI, X.; SUN, N. Rt3d: Real-time 3-d vehicle detection in lidar point cloud for autonomous driving. *IEEE Robotics and Automation Letters*, IEEE, v. 3, n. 4, p. 3434–3440, 2018.

ZHANG, C.; WAN, H.; SHEN, X.; WU, Z. Pvt: Point-voxel transformer for point cloud learning. *International Journal of Intelligent Systems*, Wiley Online Library, v. 37, n. 12, p. 11985–12008, 2022.

ZHANG, C.; WU, Z.; WU, X.; ZHAO, Z.; WANG, S. Few-shot 3d point cloud semantic segmentation via stratified class-specific attention based transformer network. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.: s.n.], 2023. v. 37, n. 3, p. 3410–3417.

ZHANG, X.; FENG, Y.; LI, S.; ZOU, C.; WAN, H.; ZHAO, X.; GUO, Y.; GAO, Y. View-guided point cloud completion. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2021. p. 15890–15899.

ZHANG, Y.; TIAN, Y.; KONG, Y.; ZHONG, B.; FU, Y. Residual dense network for image super-resolution. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2018. p. 2472–2481.

ZHANG, Z.; CAO, L.; MULVERHILL, C.; LIU, H.; PANG, Y.; LI, Z. Prediction of diameter distributions with multimodal models using lidar data in subtropical planted forests. *Forests*, MDPI, v. 10, n. 2, p. 125, 2019.

ZHAO, H.; JIA, J.; KOLTUN, V. Exploring self-attention for image recognition. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. [S.l.: s.n.], 2020. p. 10076–10085.

ZHAO, H.; JIANG, L.; JIA, J.; TORR, P. H.; KOLTUN, V. Point transformer. In: *Proceedings of the IEEE/CVF international conference on computer vision*. [S.l.: s.n.], 2021. p. 16259–16268.

ZHAO, H.; JIANG, L.; JIA, J.; TORR, P.; KOLTUN, V. *Point Transformer*. 2021. Available at: <https://arxiv.org/abs/2012.09164>.

ZHIQIANG, W.; JUN, L. A review of object detection based on convolutional neural network. In: IEEE. *36th Chinese Control Conference*. [S.l.], 2017. p. 11104–11109.

ZHOU, H.; CAO, Y.; CHU, W.; ZHU, J.; LU, T.; TAI, Y.; WANG, C. Seedformer: Patch seeds based point cloud completion with upsample transformer. In: SPRINGER. *European conference on computer vision*. [S.l.], 2022. p. 416–432.

ZHOU, T.; POPESCU, S.; LAWING, A.; ERIKSSON, M.; STRIMBU, B.; BüRKNER, P. Bayesian and classical machine learning methods: A comparison for tree species classification with lidar waveform signatures. *Remote Sensing*, v. 10, n. 1, 2018. Cited By 19. Available at: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85040699889&doi=10.3390%2frs10010039&partnerID=40&md5=a882bae04d7833eb9166a9455fd85216>.

ZHOU, T.; POPESCU, S. C.; LAWING, A. M.; ERIKSSON, M.; STRIMBU, B. M.; BÜRKNER, P. C. Bayesian and classical machine learning methods: a comparison for tree species classification with lidar waveform signatures. *Remote Sensing*, MDPI, v. 10, n. 1, p. 39, 2017.

ZHU, J.; GAO, C.; SUN, Q.; WANG, M.; DENG, Z. A survey of indoor 3d reconstruction based on rgb-d cameras. *IEEE Access*, IEEE, 2024.

APPENDIX A

Contributions

A.1 Original Papers, Software, and Datasets

The following are presented: the scientific manuscripts, original software, and datasets published in journals, international conferences, and online repositories during the course of this research.

A.1.1 Publications directly related to this thesis

1. Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, and Fabrizzio Soares. "Eucalyptus diameter and volume prediction with deep neural networks: A Long Short-Term Memory model approach." In Expert Systems With Applications 271 (2025) 126704, 2025. (RODRIGUES *et al.*, 2025) **(Qualis A1)**
2. Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Renato F. Bulcão-Neto, and Fabrizzio Soares. "Applications of artificial intelligence and LiDAR in forest inventories: A Systematic Literature Review." In Computers and Electrical Engineering 120 (2024) 109793, 2024. (RODRIGUES *et al.*, 2024b) **(Qualis A1)**
3. Rodrigues, Welington G., Emilia A. Nogueira, Thamer H. Nascimento, Gabriel S. Vieira, Deborah S. A. Fernandes, and Fabrizzio Soares. "MobileNetV2-Enhanced Depth Map Super-Resolution Through Multi-scale Image Guidance." In 19th International Symposium, ISVC 2024 Lake Tahoe, NV, USA, October 21–23, 2024 Proceedings, Part II, Advances in Visual Computing, 2024. (RODRIGUES *et al.*, 2024a) **(Qualis A4)**
4. Rodrigues, Welington G., Gabriel Vieira, Afonso U. Fonseca, Juliana P. Felix, Deborah S. A. Fernandes, and Fabrizzio Soares. "3D Point Cloud Segmentation of LiDAR-Scanned Eucalyptus Trees: a Deep Learning Analysis." In 2025 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2025. (Waiting to be published in the proceedings) **(Qualis A3)**

5. Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Helio Pedrini, Juliana P. Felix, and Fabrizzio Soares. "Accurate Measurement of Eucalyptus Tree Diameters Using Smartphone LiDAR and Deep Learning." In IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS), 2025 (Waiting for a response). **(Qualis A1)**
6. Rodrigues, Welington G., Christian D. Cabacinha, Fernandes, Deborah S. A. and Fabrizzio Soares. "3D Point Cloud Completion in Real-World Forestry: A Comparative Study for Eucalyptus Plantations."International Conference on Computer Vision Theory and Applications" (VISAPP), 2026 (Waiting for a response). **(Qualis A3)**
7. Rodrigues, Welington G., Gabriel S. Vieira, Christian D. Cabacinha, Juliana P. Felix, Thamer H. Nascimento and Fabrizzio Soares. "Generating High-Fidelity Eucalyptus Point Clouds: A Comparative Study from GANs to Diffusion Models." In 20th International Symposium on Visual Computing (ISVC) 2025, Proceedings Part II, Advances in Visual Computing (Accepted for publication). **(Qualis A4)**

A.1.2 Other publications as coauthor

1. Junior, Silvio V. M. Welington G. Rodrigues, Gabril S. Vieira and Fabrizzio Soares. "Comparative Study of Depth Anything Model V2 and LiDAR sensors for Depth Map Estimation in Forest Environment."In: 2025 IEEE 46th annual computers, software, and applications conference (COMPSAC) (Awaiting publication) **(Qualis A3)**
2. Junior, Silvio V. M. Welington G. Rodrigues, Gabril S. Vieira and Fabrizzio Soares. "Tree Diameter Estimation Using LiDAR-Equipped Smartphones for Forest Inventory Applications". In IEEE International Conference on Industrial Informatics, 2025 (INDIN) (Awaiting publication) **(Qualis A3)**
3. Claudio G. De L. Almeida, Emília Alves Nogueira, Juliana Felix and Afonso Ueslei Fonseca; Gabriel Vieira; Welington Galvão Rodrigues, Deborah S. A. Fernandes. "Application of Real-ESRGAN to Improve Resolution in Agricultural Satellite Images". In 2025 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (Awaiting publication) **(Qualis A3)**
4. Nascimento, Thamer Horbylon. Jamilly Santos, Marcos Alves Vieira, Deborah Fernandes, Juliana Felix, Welington Rodrigues, Fabrizzio Soares and Thamer Nascimento. "Low-Cost Immersive Molecule Visualization from Hand-Drawn Smartwatch Input for Virtual Reality Learning". In 20th International Symposium on Visual Computing (ISVC) 2025 (Awaiting publication) **(Qualis A4)**

5. Nascimento, Thamer Horbylon, Deborah Fernandes, Diego Siqueira, Gabriel Vieira, Gustavo Moreira, Leonardo Silva, and Fabrizzio Soares. "Interaction with smartwatches using gesture recognition: a systematic literature review". In: 2020 IEEE 44th annual computers, software, and applications conference (COMPSAC). IEEE, 2020. p. 1661-1666. (NASCIMENTO *et al.*, 2020) **(Qualis A2)**
6. Mombach, Jaline. Afonso U. Fonseca, Thamer H. Nascimento, Pinheiro, Gabriel S. Vieira,Henrique Gressler, Fábio D. Rossi, and Fabrizzio Soares. "A new approach to performing paper-based children's spelling tests on mobile devices." In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2020. p. 4346-4351 (MOMBACH *et al.*, 2020) **(Qualis A2)**

A.1.3 Original Software

1. *PointCloud Visualizer*

   Description:

   The software, developed as part of this PhD thesis is a specialized toolkit designed for advanced visualization and processing of 3D point cloud data, with a primary focus on forestry applications. It leverages mobile LiDAR technology to enable automated diameter measurement of eucalyptus trees through precise 3D reconstruction and segmentation techniques. The toolkit offers robust tools for handling large-scale LiDAR datasets, allowing users to visualize, analyze, and extract meaningful geometric features from point clouds.

   Github Repository:

   <https://github.com/wgalvao/pointcloud_visualizer>

A.1.4 New Datasets

1. *EucaliptTree55Raw*

   Description:

   The Eucalyptus Point Cloud Dataset is the initial dataset, consisting of point clouds from 55 eucalyptus trees, with each file corresponding to an individual tree. The point clouds were obtained using the LiDAR sensor of the iPhone 12 Pro Max. The dataset was prepared to include

the three spatial coordinates (x, y, z) and the RGB color channels for each point.

Google Drive Repository: <https://drive.google.com/drive/folders/1Jdw0WHJKxDI7mvyXYSXS3wSv2xbpQmlA?usp=drive_link>

2. *Euca3D dataset*

Description:

The Euca3D Dataset is a pre-processed database derived from the original point clouds. The ground was removed, and each point cloud was reduced to 4,096 points, ensuring an efficient and standardized representation of each tree. Additionally, the point clouds were manually segmented at heights corresponding to diameters at 30 cm, 70 cm, 1 meter, and 1.30 meters (DBH) above the ground. The dataset features an organized structure, with each segment associated with its respective labels, facilitating tasks such as analysis, training of machine learning models, and quantitative evaluations.

Github Repository: <https://drive.google.com/drive/folders/1RH1mHr1I9cqnUHYghCn-dbnCWCvBHsa7?usp=drive_link>

3. *Euca3D Point Completion*

Description:

The Euca3D Point Completion is a variation of the original dataset, in which the pre-processed and normalized point clouds were sagittally cropped, splitting each tree into two parts. Additionally, random point removal was performed to simulate noise and structural damage in the point cloud. As a result, the task associated with this dataset involves completing the degraded point cloud to reconstruct its original shape.

Google Drive Repository: <https://drive.google.com/drive/folders/1khAjKMLTJCvqDieL1SGYBisvSO71QNPb?usp=drive_link>

4. *Eucalitpo 3D dataset*

Description: The Eucalyptus 3D Dataset is the complete dataset, comprising 440 eucalyptus trees scanned using the iPhone 12 Pro Max and iPhone 14 Pro Max devices, both equipped with LiDAR technology. This database includes the 55 trees previously released in the initial dataset, supplemented by 385 new trees captured with the iPhone 14 Pro Max. Each tree is represented by an individual point cloud containing spatial coordinates (x, y, z) with a range of up to 2 meters, ensuring the structural and geometric fidelity of the scanned vegetation.

Google Drive Repository: <https://drive.google.com/drive/folders/1RH1mHr1I9cqnUHYghCn-dbnCWCvBHsa7?usp=drive_link>

APPENDIX B

Authorization for Reuse of Published Papers

All papers originating from the thesis research were duly cited, and the publishers granted the authors the right to use them in theses and/or dissertations, as follows.

19th International Symposium on Visual Computing - ISVC 2024

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Oct 31, 2025

This Agreement between WELINGTON GALVAO RODRIGUES ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	6139471010856
License date	Oct 31, 2025
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	MobileNetV2-Enhanced Depth Map Super-Resolution Through Multi-scale Image Guidance
Licensed Content Author	Welington G. Rodrigues, Emilia A. Nogueira, Thamer H. Nascimento et al
Licensed Content Date	Jan 1, 2025
Type of Use	Thesis/Dissertation
Requestor type	academic/university or research institute
Format	electronic
Portion	Text extract
Number of text extracts	10
Will you be translating?	no

ELSEVIER – Rights and Permissions

The following page contains a PDF file presenting the formal permission statements for the reuse of articles published by Elsevier, which were developed as an integral part of this doctoral thesis.

According to Elsevier's copyright policy, as cited in the attached documents, authors "retain the right to include their article in a thesis or dissertation, provided it is not published commercially." The policy further clarifies that "No permission is required, but the journal must be cited as the original source."

Based on this policy, the following articles published by Elsevier are included herein:

1. Article in *Computers & Electrical Engineering*

A formal permission statement is provided for the article published in the journal *Computers & Electrical Engineering*.

- **Article Title:** Applications of artificial intelligence and LiDAR in forest inventories: A Systematic Literature Review
- **Author's Declaration:** The corresponding author, Welington G. Rodrigues, declares that the aforementioned article is incorporated into this doctoral thesis "solely for academic and research purposes," ensuring proper citation of *Computers & Electrical Engineering* as the original source.
- **Original Reference:**

  RODRIGUES, W. G.; VIEIRA, G. S.; CABACINHA, C. D.; BULCÃO-NETO, R. F.; SOARES, F. Applications of artificial intelligence and LiDAR in forest inventories: A Systematic Literature Review. *Computers & Electrical Engineering*, v. 119, 2024. Elsevier Ltd. 2024.

- **DOI:** The article can be accessed at: <https://doi.org/10.1016/j.compeleceng.2024.109793>

2. Article in *Expert Systems with Applications*

A permission statement is also provided for the article published in the journal *Expert Systems with Applications*.

- **Article Title:** Eucalyptus Diameter and Volume Prediction with Deep Neural Networks: A Long Short-Term Memory Model Approach

- **Author's Declaration:** The corresponding author, Welington G. Rodrigues, reaffirms that the article is incorporated into this thesis exclusively for academic and research purposes. In accordance with the publisher's policy, the journal *Expert Systems with Applications* is cited as the original source of publication.
- **Original Reference:**

  RODRIGUES, Welington G.; VIEIRA, Gabriel S.; CABACINHA, Christian D.; SOARES, Fabrizzio. Eucalyptus diameter and volume prediction with deep neural networks: A Long Short-Term Memory model approach. *Expert Systems with Applications*, v. 211, p. 119585, 2025. DOI: 10.1016/j.eswa.2025.119585.

- **DOI:** The article can be accessed at: <https://doi.org/10.1016/j.eswa.2025.126704>

Copyright | Elsevier policy

elsevier.com/about/policies-and-standards/copyright

Authors' licensing and copyright options when publishing research articles with Elsevier

In order for Elsevier to publish and disseminate research articles, we need authors to grant us certain publishing rights, which are determined by the author's choice of publishing model and, in the case of open access articles, the author's choice of end user license.

Irrespective of the model under which an article is published, Elsevier is committed to protecting and defending authors' works and their reputation. We take allegations of infringement, plagiarism, ethical disputes, and fraud very seriously.

Elsevier offers two publishing routes for authors to choose from, which have different implications for authors and others:

1. Pay-to-publish, also known as open access, which usually requires a fee from the author, or their institution or funder, to publish. When publishing open access with Elsevier, the author retains their copyright in the article and grants Elsevier a license to publish it. Authors have a choice of end user license:

- If the author selects a Creative Commons Attribution (CC BY) license, end users are bound by this license. Each author can re-use their article themselves and *also* grant others (additional to Elsevier) a license to re-use their article for any purpose without permission from, or payment to, Elsevier.

- If the author selects a Creative Commons Attribution-NonCommercial-NoDerivatives (CC BY-NC-ND) license or a Creative Commons Attribution-NonCommercial (CC BY-NC) license, end users are bound by the license. Each author retains the right to re-use the article for their own commercial and non-commercial purposes, including creating derivative works, without permission from, or payment to, Elsevier.

2. Pay-to-read, where articles are free to publish, but usually require a subscription or payment to read. When publishing subscription the author typically transfers copyright to the publisher. However, each author has the right to re-use the article for certain personal and scholarly purposes, including the creation of some derivative works, without permission from, or payment to, Elsevier.

In all cases, any rights the author grants to Elsevier apply only to the published article. Elsevier does not require, nor ever request, rights to the underlying research itself. Authors (and institutions or employers) will continue to retain patent, trademark and other intellectual property rights relating to the research, as well as the right to use their research data freely and without restriction.

Comparing reuse rights

The below tables compare rights for authors and others under the publishing agreements for open access articles (OA) with CC BY, CC BY-NC and CC BY-NC-ND licenses and subscription articles.

Authors' rights in the article

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Receive proper attribution and credit for their published work	Yes	Yes	Yes	Yes

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Re-use their article in their own new works, without permission from Elsevier or payment to Elsevier, including by: • making copies of the article (or part of the article) to promote companies or products they own, whether or not such promotion is commercial; • including the article in a thesis or dissertation; • extending the article to a book, including the article in a subsequent compilation of their own work, or re-using portions, excerpts, and their own figures, tables and images from the article in their own new works (which in each case may be published with Elsevier or with a third party commercial or non-commercial publisher, at the author's discretion)	Yes	Yes	Yes	Yes[1]
Use and share their article for scholarly purposes, including: • for classroom teaching; • at conferences; • for non-commercial Massive Open Online Courses; • create translations of the article and authorize others to do so for non-commercial scholarly collaborations and sharing	Yes	Yes	Yes	Yes[2]
Publicly share the **preprint** anywhere at any time	Yes	Yes	Yes	Yes
Publicly share the **final published article** immediately on **non-commercial** sites e.g., institutional repositories, ensuring attribution	Yes	Yes	Yes	No

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Publicly share the **final published article** immediately on **commercial scholarly collaboration networks**, ensuring attribution, **for the purpose only of hosting** by those websites	Yes	Yes	Yes	No
Publicly share the **final published article** immediately on **commercial scholarly collaboration networks**, ensuring attribution, **for reuse** by those websites	Yes	No	No	No
Publicly share the **accepted manuscript** on **non-commercial sites** after an embargo period and attaching a CC BY-NC-ND end user license	Yes	Yes	Yes	Yes
Right to license others to exercise the above rights for commercial purposes	Yes	No	No	No
Hold copyright of their article	Yes	Yes	Yes	No

Notes:

- [1]The list is exhaustive in the case of articles published under the subscription model; additional uses/ sharing beyond those listed would require permissions to be requested from Elsevier.
- [2]In the case of articles published under the subscription model: Sharing the Accepted Manuscript for non-commercial Massive Open Online Courses is permitted after the embargo period and provided a CC BY-NC-ND end user license is attached. While the author can create translations of the article for non-commercial scholarly collaborations and sharing, they cannot authorize others to do so.

Authors' institutions' rights in the article

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Distribute copies for classroom teaching and internal training purposes	Yes	Yes	Yes	Yes
Include in (online) coursework and courseware programs for use within the institution	Yes	Yes	Yes	Yes
Include in non-commercial Massive Open Online Courses	Yes	Yes	Yes	No
Include in applications for grant funding	Yes	Yes	Yes	Yes
Post publicly as part of theses and dissertations, with DOI links to the formal publication	Yes	Yes	Yes	Yes

Elsevier's rights in the article

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Rights to publish, reproduce, display, distribute and otherwise use all or any part of the article, including tables, illustrations or other materials, in print, electronic and all other media (whether now known or later developed)	Yes	Yes	Yes	Yes
To prepare derivative works, in any form, in all languages, throughout the world, for the full term of copyright	Yes	Yes	Yes	Yes

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
The right to license others to exercise the above rights	Yes	Yes	Yes	Yes
Manage permissions for third party reuse and adaptation of the work for commercial purposes	No	Yes	Yes	Yes

Others' rights to use the article without permission or payment

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Distribute the article for **non-commercial** purposes	Yes	Yes	Yes	No
Distribute the article for **commercial** purposes	Yes	No	No	No
Include in a collective work (such as an anthology) for **non-commercial** purposes	Yes	Yes	Yes	No
Include in a collective work (such as an anthology) for **commercial** purposes	Yes	No	No	No
Create and distribute revised versions, adaptations or derivative works of or from an article (such as a translation) for **non-commercial** purposes	Yes	Yes	No	No
Create and distribute revised versions, adaptations or derivative works of or from an article (such as a translation) for **commercial** purposes	Yes	No	No	No

Empty table header	OA with CC BY	OA with CC BY-NC	OA with CC BY-NC-ND	Subscription
Text or data mine the article for **non-commercial** purposes	Yes	Yes	Yes[1]	No[2]
Text or data mine the article for **commercial** purposes	Yes	No	No	No

Notes:

- Users accessing content as part of a subscription agreement may have additional rights under that agreement. Users should contact their library or subscription management group for more information.
- Further permission may be required from the rights owner for any content within an article that is identified as belonging to a third party.
    - [1]In line with the terms of the end user license, adaptations cannot be shared with others.
    - [2]Where legal access is obtained by a user, that user is able to text or data mine subscription articles for non-commercial purposes without sharing any adaptation of the original content with others.

Publishing Agreements at Elsevier

Authors publishing with Elsevier sign a Publishing Agreement. This has several important functions, including to confirm the author has complied with publishing ethics policies and that the published article will be the final version of record. When publishing open access, the publishing agreement will set out the terms of the license the author grants to Elsevier to publish their article. In the case of publishing subscription, the agreement will transfer copyright to Elsevier and set out reuse rights to the work. Note: In some circumstances, authors may instead grant Elsevier (or the learned society on whose behalf we publish) an exclusive license to publish and disseminate their work. Please see links to sample Publishing Agreements below.

Additional information on how Elsevier manages commercial licenses for open access articles published under CC BY-NC-ND or

CC BY-NC licenses

Our experience, based on requests from authors globally across decades, is that the rights that authors themselves have, when publishing open access under a non-commercial Creative Commons license (outlined above), cover the vast majority of the situations in which authors wish to use their work. Authors who are unclear whether their intended use is covered by these rights, should submit their enquiry here to obtain further guidance.

Elsevier sometimes receives direct requests from third parties, such as pharmaceutical companies, to re-use and adapt the article. Elsevier maintains relationships with service providers (such as Rightslink and other reproduction rights organizations) to manage these requests. Elsevier incurs costs in undertaking these activities, and may charge third parties a fee when granting permissions for re-use or adaptation of the article.

Authors publishing under the CC BY-NC-ND or CC BY-NC licenses agree not to license any third party to reuse their articles or any part of their articles for commercial purposes. Elsevier has the exclusive right to license third parties to do this. This enables Elsevier to ensure compliance with applicable laws and regulations and control over appropriate commercial use by third parties.

Authors who are government employees

For US government employees, works created within the scope of their employment are considered to be public domain and Elsevier's publishing agreements do not require a transfer or license of rights for such works.

In the UK and certain commonwealth countries, a work created by a government employee is copyrightable, but the government may own the copyright (Crown copyright). Click here for information about UK government employees publishing open access.

Authors publishing in society-owned journals

Please note that the above terms may differ for society-owned journals. Please refer to the journal homepage and Guide for Authors, or contact the journal for further information.

Find out more

- Download a sample publishing agreement for articles financed by journal subscriptions in English, French and Spanish
- Download a sample publishing agreement for articles published open access with a commercial user license (CC BY) in English, French, and Spanish

- Download a sample publishing agreement for articles published open access with a non-commercial user license (CC BY-NC-ND) in English, French, and Spanish
- Download a sample publishing agreement for articles published open access with a non-commercial user license (CC BY-NC) in English, French, and Spanish
- For authors who wish to self-archive see our sharing guidelines
- See our author pages for further details about how to promote your article
- See our hosting page for additional information on hosting research published by Elsevier
- For use of Elsevier material not defined here please see our permissions page or visit the Permissions Support Center
- If an author has become aware of a possible plagiarism, fraud or infringement we recommend contacting their Elsevier publishing contact who can then liaise with our in-house legal department
- If you are publishing in a society or third party owned journal, they may have different publishing agreements. Please see the journal's Guide for Authors for journal specific copyright information