



2025 | **16-20**
GIJÓN | **JUNIO**

9º CONGRESO FORESTAL ESPAÑOL

9CFE-1992

Actas del Noveno Congreso Forestal Español
Edita: **Sociedad Española de Ciencias Forestales. 2025.**
ISBN: **978-84-941695-7-1**

Organiza





Aplicabilidad de Red Neuronal Convolutiva para la segmentación de troncos a partir de datos láser obtenidos mediante escáner láser móvil en tiempo real

MARTÍN CORTÉS, C. (1,2), PORKHUN, O. (3), GARCÍA PASCUAL, B. (2,4), RUANO SASTRE, A. (3), ZHOU, X. (2), ACUNA, M. (4), MOLA YUDEGO, B. (2), PICHÍ, G. (3,5), GONZÁLEZ OLABARRIA, J.R. (1).

(1) Unidad mixta de investigación (JRU) CTFC - AGROTECNIO, Ctra de St. Llorenç de Morunys, Km 2, 25280, Solsona, España.

(2) Universidad de Finlandia Oriental (University of Eastern Finland, UEF), PO Box 111, 80101 Joensuu, Finlandia.

(3) Consorcio Centro de Ciencia y Tecnología Forestal de Cataluña (CTFC), Ctra de St. Llorenç de Morunys, Km 2, 25280, Solsona, España.

(4) Centro de Recursos Naturales de Finlandia (Luke), Yliopistokatu 6, 80100 Joensuu, Finlandia.

(5) Consejo Nacional de Investigación de Italia, Instituto por la bioeconomía (CNR), Via Madonna del Piano 10, 50019 Sesto Fiorentino, Italia.

Resumen

La planificación de operaciones forestales se ha basado tradicionalmente en inventarios forestales, los cuales requieren de mediciones manuales de campo, un proceso que suele ser costoso tanto monetariamente, como en tiempo de trabajo. El avance de tecnologías de sensorización remota, ha demostrado que es posible mejorar tanto la precisión como la eficiencia alrededor de la planificación de operaciones forestales, reduciendo los costes y la mano de obra necesaria. Sin embargo, la detección y segmentación de troncos generalmente ha requerido un post-procesado exhaustivo de los datos LiDAR, lo que limita su aplicabilidad en contextos donde se necesita información en tiempo real.

El presente trabajo estudia la viabilidad de la detección y segmentación de troncos de árboles en tiempo real, a partir de datos LiDAR obtenidos mediante escáner láser móvil (MLS) en tiempo casi real. Para ello, se analiza por separado cada fotograma capturado por los sensores LiDAR, definido como un registro de datos correspondiente a un escaneo individual del sensor. Estos se transforman en proyecciones 2D, donde una Red Neuronal Convolutiva se encarga de identificar y segmentar los troncos. De este modo, se elimina la necesidad de analizar las nubes de puntos y permite una respuesta rápida en entornos operativos. Una detección y segmentación precisa de troncos en tiempo real podría ser la base de la automatización y robotización de las operaciones forestales.

Palabras clave

LiDAR, operaciones forestales, maquinaria forestal, precisión forestal, automatización de procesos.



1. Introducción

Durante las últimas tres décadas, el mundo se ha visto inmerso en una revolución tecnológica que ha supuesto el impulso hacia las Industrias 4.0 y 5.0. La Industria 4.0 se caracteriza por la integración de sistemas ciberfísicos, Internet de las Cosas (IoT), macrodatos y la inteligencia artificial, teniendo como objetivo crear entornos inteligentes que permitan el procesamiento de datos en tiempo real para una toma de decisiones más eficiente (GHOBAKHLOO, 2020). Así mismo, la Industria 5.0 se ha centrado en complementarla, poniendo la investigación y la innovación al servicio de una transición hacia una producción sostenible, resiliente y centrada en el ser humano (GÜRDÜR BROO et al., 2022; JAVAID et al., 2022; LENG et al., 2022). Este nuevo enfoque ofrece oportunidades para el sector forestal, tales como la analítica predictiva, la automatización y la gestión de precisión, que podrían transformar las operaciones forestales tradicionales en prácticas inteligentes, eficaces y sostenibles, impulsando al sector hacia un marco conceptual denominado “Silvicultura 5.0” (HOLZINGER et al., 2024). Pero la implementación y adopción de nuevas tecnologías plantea desafíos de naturaleza tanto técnica como socioeconómica (CHESBROUGH, 2010).

No obstante, a pesar de los retos asociados con la adopción de nuevas tecnologías, los operadores forestales han comenzado a utilizar tecnologías avanzadas para mejorar los resultados de la gestión forestal, lo que se ha denominado “silvicultura de precisión” (CLOUDHRY & O’KELLY, 2018). En el ámbito de las operaciones forestales, el uso de tecnologías avanzadas a menudo busca una mayor automatización de los procesos, como es el caso de los aprovechamientos forestales, donde se hace un uso intensivo de maquinaria (JELAVIC et al., 2021; LINDROOS et al., 2017; VISSER & OBI, 2021). Estas operaciones han sido identificadas como un campo lógico donde se prevé probar máquinas autónomas en el futuro cercano (CLOUDHRY & O’KELLY, 2018).

Aunque el interés por una mayor automatización de las operaciones de aprovechamiento forestal ha sido claramente manifestado, persisten ciertas limitaciones. Lograr esta transición requiere una localización, navegación e identificación precisa de las características del entorno, aspectos fundamentales para los sistemas móviles (ROSSMANN et al., 2009). Estas características del entorno pueden comprenderse a través de sensores que permitan la toma de decisiones en tiempo real. Entre los sensores más comúnmente utilizados para percibir el entorno y medir distancias a objetos, destaca el sensor LiDAR (acrónimo del inglés, *Light Detection and Ranging*) (CAMPBELL et al., 2018). Este tipo de sensor genera una nube de puntos 3D que representa con precisión el bosque (MURTIYOSO et al., 2024), informando sobre la distribución espacial de los elementos que lo constituyen, lo cual brinda soporte técnico para la investigación y gestión de recursos forestales (BONAN, 2008; LIU et al., 2025).

La toma de decisiones en la gestión de recursos forestales depende en gran medida de información precisa recopilada mediante procesos de inventario (LIANG et al.,



2016). Actualmente, muchos gestores forestales, silvicultores e investigadores están acostumbrados a trabajar con estimaciones generales de volumen comercializable por rodal, número de árboles por hectárea y otras variables calculadas a nivel del rodal (KEEFE et al., 2022). Sin embargo, no centrarse en las características individuales de los árboles, limita la capacidad de tomar decisiones silvícolas ajustadas al potencial específico de cada árbol (ACHIM et al., 2022). Un ejemplo de cómo los datos a nivel de árbol permiten una gestión más precisa, es el tronzado óptimo del tronco, que puede mejorar la productividad y rendimiento de la gestión forestal (PRENDES et al., 2023), 2023). Por lo tanto, la detección individual de árboles mediante sensores LiDAR ha emergido como una herramienta prometedora para medir bosques a una escala ecológicamente significativa y útil para los gestores forestales (JERONIMO et al., 2018).

Los escáneres láser terrestres han ganado popularidad en los inventarios dendrométricos. Estos dispositivos permiten medir atributos como el diámetro a lo largo del tronco, la altura total del árbol, la curvatura y el volumen de este (PRENDES PÉREZ, 2023). Entre las tecnologías disponibles, el escáner láser móvil (MLS) destaca como especialmente ventajoso, ya que supera los desafíos de movilidad y tiempo de medición asociados al escáner láser terrestre (TLS), ya que este es un sistema fijo que no cuenta con capacidad de escanear en movimiento, haciendo que la adquisición de datos sea más eficiente y práctica (LIANG et al., 2014). Por esta razón, el MLS es la tecnología de escaneo láser principal en el diseño de sistemas móviles automatizados.

En el caso de la automatización del aprovechamiento forestal mediante métodos de corte a medida (CTL), está surgiendo una nueva línea de investigación centrada en la aplicación de escáneres láser. Pero el uso de sensores láser en procesadoras forestales para realizar CTL plantea dos desafíos clave que deben abordarse para manejar macrodatos y proporcionar soluciones en tiempo real: la localización y mapeo simultáneos (SLAM) y la detección de árboles con extracción de atributos. Si bien varios estudios han abordado la localización y mapeo simultáneos (FAITLI et al., 2024; JELAVIC et al., 2021, 2022; MIETTINEN et al., 2007; NEVALAINEN et al., 2020; ÖHMAN et al., 2008), pocos se han centrado en la detección de árboles con extracción de atributos (FAITLI et al., 2024; JELAVIC et al., 2021, 2022; JUTILA et al., 2007; MIETTINEN et al., 2010; NEVALAINEN et al., 2020; ÖHMAN et al., 2008; ROSSMANN et al., 2009; SAGAR et al., 2024; SIHVO et al., 2018).

En la literatura, los métodos presentados para detectar troncos de árboles en nubes de puntos (PC) en entornos forestales incluyen estrategias como el ajuste robusto de cilindros, el voto basado en vóxeles utilizando operaciones morfológicas para resaltar troncos, el filtrado local de puntos basado en características geométricas, el emparejamiento de segmentos de línea para modelar troncos verticales y la aplicación de redes neuronales convolucionales 3D (3D CNN) para la detección de troncos en datos de nubes de puntos (NEVALAINEN et al., 2020). Algunos de estos métodos parecen ser capaces de construir mapas locales de árboles en tiempo real o con postprocesado dentro de un rango limitado de aproximadamente 10 a 15 metros (NEVALAINEN et al., 2020; PIERZCHAŁA et al., 2018; PROUDMAN et al., 2022; ZHANG et al., 2019). Aunque ya se han presentado



métodos para detectar troncos de árboles en tiempo real en entornos forestales, todos estos métodos analizan nubes de puntos con el objetivo de clasificar puntos como troncos de árboles. Sin embargo, la principal limitación de capturar información en tiempo real con sensores LiDAR radica en el manejo de los grandes volúmenes de datos que generan.

Según nuestro conocimiento, la detección de objetos en entornos forestales utilizando nubes de puntos se realiza habitualmente mediante nubes de puntos 3D o convirtiéndolas en vóxeles. Estos enfoques son efectivos, pero también son computacionalmente intensivos y requieren recursos sustanciales, especialmente para aplicaciones en tiempo real. Un tercer método, que implica proyectar la nube de puntos y su información en un plano 2D como imágenes, no es comúnmente utilizado en entornos forestales. Sin embargo, este enfoque reduce significativamente la carga computacional, haciéndolo más adecuado para aplicaciones en tiempo real (JHALDIYAL & CHAUDHARY, 2023). Este enfoque de proyección 2D resulta particularmente ventajoso cuando se combina con datos LiDAR multicanal, como intensidad, infrarrojo y reflectividad, que mejoran la precisión de la clasificación al proporcionar representaciones 2D detalladas del entorno. Además, las redes neuronales convolucionales (CNN), que se encuentran en una fase más desarrollada para datos 2D que para datos 3D, pueden aprovecharse eficazmente para procesar estas proyecciones, mejorando el rendimiento de detección y segmentación. Uno de los modelos CNN más destacados para realizar segmentaciones semánticas e instancias en tiempo real es Yolo V8, un modelo de visión por computadora diseñado para tareas de detección de objetos en tiempo real (JOCHER et al., 2023).

2. Objetivos

Este estudio tiene como objetivo desarrollar una prueba de concepto y un flujo de trabajo para la detección en tiempo casi real de troncos de árboles, usando un sistema LiDAR montado en una procesadora forestal. Este flujo comienza con la proyección esférica de los múltiples canales del LiDAR, convirtiendo los datos de nubes de puntos 3D en proyecciones 2D. Este enfoque también incluye la implementación y evaluación de un modelo de aprendizaje profundo capaz de detectar troncos de árboles casi en tiempo real a partir de datos sin procesar capturados por el escáner. Específicamente, generamos imágenes RGBA que integran la información multicanal del LiDAR, incluida la intensidad, infrarrojo, reflectividad y rango y poder entrenar la red convolucional. El modelo YoloV8, optimizado para la detección de objetos en tiempo real, se emplea para realizar segmentación basada en instancias de troncos de árboles, permitiendo la identificación y localización de troncos individuales dentro de la escena.

3. Metodología

El flujo de trabajo seguido en este estudio y explicado en las siguientes secciones se muestra en la figura 1.

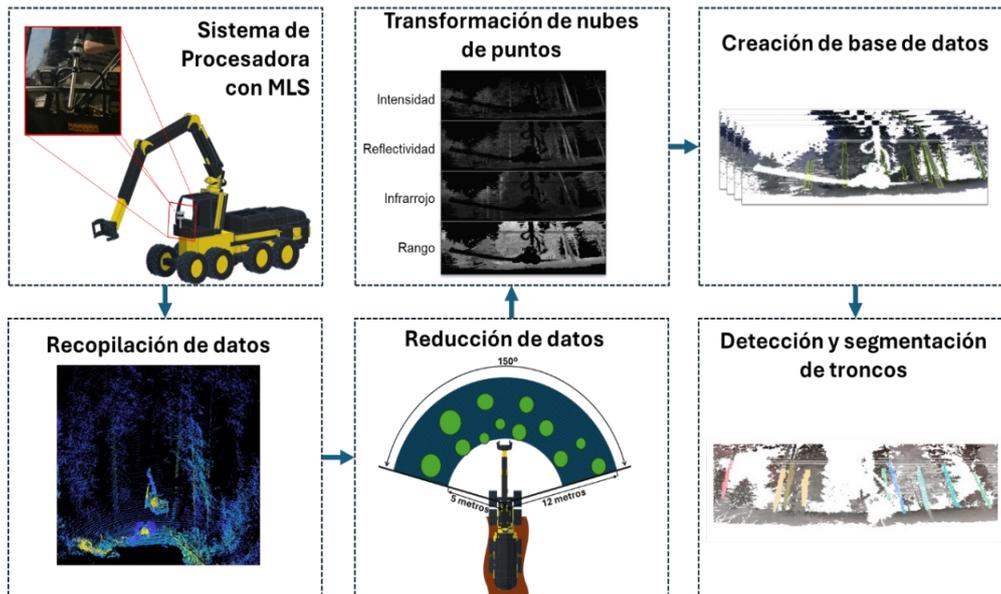


Figura 1. Flujo general del estudio.

Sistema de procesadora con escáner láser móvil

Se incorporó un sistema de escáner láser móvil Kudan Backpack Dev Kit (Kudan USA LLC., EE.UU), en una procesadora forestal Ponsse Scorpion (PONSSE, FINLANDIA). El escáner se montó en la parte delantera de la cabina de la procesadora para optimizar su campo de visión y garantizar la captura de datos durante las operaciones forestales. Esta posición se seleccionó para proporcionar una perspectiva similar a la del operador.

El Kudan Backpack Dev Kit estaba equipado con un sensor LiDAR Ouster OS0-128 Rev 7 (OUSTER, EE. UU.). La configuración de hardware incluía un procesador Intel Core i7 de 13.^a generación con 64 GB de memoria DDR-RAM y 2 TB de almacenamiento SSD. El set de sistemas de software consiste en Ubuntu 22.04 como sistema operativo y la versión ROS Noetic Full Desktop como middleware. El sistema utilizaba el controlador Ouster ROS para interactuar con aplicaciones basadas en ROS y Kudan LiDAR SLAM para ROS para la localización y el mapeo simultáneos. Kudan ROS Launcher era la interfaz gráfica de usuario principal para grabar y controlar el sistema LiDAR.

Sitio de estudio y recopilación de datos

La recolección de datos se llevó a cabo el 5 de junio de 2024 en un rodal ubicado en Carelia del Norte, aproximadamente a 25 km al oeste de Joensuu, en las coordenadas 62°35'25.72"N y 29°03'29.4"E. Durante el proceso de recolección de datos, la procesadora llevó a cabo la fase de corta final de una pequeña área de estudio de 0,5 hectáreas de *Pinus sylvestris*.

La recolección de datos LiDAR se realizó utilizando el software Kudan LiDAR SLAM, que permitió la adquisición y el registro en tiempo real de los datos generados por el sensor MLS. El sensor se configuró con una resolución vertical de



128 canales, una resolución horizontal de 1024 puntos y una velocidad de rotación de 20 Hz. Esta configuración fue elegida para maximizar el rendimiento del sensor, permitiendo la captura densa y precisa de información que le rodea.

Los datos capturados fueron almacenados en archivos de formato ROS bag (rosvbag), un formato estándar ampliamente utilizado en aplicaciones de robótica y sistemas basados en el Sistema Operativo Robótico (ROS, acrónimo del inglés, “Robot Operating System”). Este formato permite la grabación y reproducción de mensajes intercambiados, incluyendo aquellos que contienen nubes de puntos capturados por el sensor LiDAR. Los archivos rosvbag almacenaban información esencial sobre las coordenadas 3D (X, Y, Z) de los puntos, junto con datos de los canales de intensidad, reflectividad, rango e inarrojo, generando una nube de puntos por fotograma (comunmente conocido como frame), un fotograma se refiere al conjunto de datos obtenidos en una única captura generada por el escáner.

El proceso de recolección de datos se llevó a cabo durante la operación de una procesadora forestal que cortó árboles a lo largo del área de estudio. Durante esta operación, el sistema MLS capturó y almacenó los archivos rosvbag de forma secuencial, facilitando la gestión de los datos y posibilitando su posterior procesamiento y análisis. En total, se realizaron 3 registros de datos en la misma ubicación de la operación.

Reducción de datos y transformación de la nube de puntos

El proceso de reducción de datos se centró en reducir las nubes de puntos capturadas en cada fotograma a través de dos criterios específicos. En primer lugar, se filtraron los puntos en función de su distancia horizontal (distancia radial en el plano XY desde el sensor MLS), conservando solo aquellos puntos dentro de un rango horizontal de 5 a 12 metros. Aunque el brazo de una procesadora CTL puede alcanzar hasta 10 metros (PONSSE, s. f.), este rango horizontal se eligió para este estudio con objeto de permitir la detección temprana de árboles, y así detectarlos antes del momento de corte. La distancia mínima de 5 metros se definió para eliminar el ruido en la nube de puntos causado por la propia procesadora. En segundo lugar, se excluyeron los puntos fuera de un campo de visión de 150° y así replicar el rango visual operativo de un operador humano. En el flujo general (Figura 1) se ilustra los criterios de filtrado aplicados para una mejor comprensión.

Los objetivos principales de esta reducción de datos fueron dos: reducir el coste computacional asociado con el análisis de datos, reduciendo el volumen de las nubes de puntos, y enfocarse en el área operativa de la procesadora, reflejando tanto el alcance efectivo del brazo como el campo visual del operador.

Una vez filtradas las nubes de puntos de cada fotograma se han proyectado en planos 2D. Esta transformación se lleva a cabo para facilitar la detección de troncos, ya que proyectar la información de la nube de puntos en un plano 2D ha



demostrado ser una de las estrategias más viables para aplicaciones en tiempo real con datos LiDAR. Esto se debe a su menor costo computacional en comparación con los métodos de detección basados en puntos o vóxeles. Además, los métodos basados en proyección aprovechan la solidez de las redes neuronales convolucionales 2D (JHALDIYAL & CHAUDHARY, 2023).

El procedimiento implica proyectar la nube de puntos 3D en planos 2D y, posteriormente, identificar las etiquetas de clase ejecutando las proyecciones a través de CNN 2D preexistentes. Una vez que se etiqueta cada punto, las proyecciones 2D se vuelven a proyectar en la nube de puntos respectiva, lo que garantiza que cada punto 3D conserve su información de clase.

Existen dos métodos populares para proyectar nubes de puntos 3D en 2D: proyección esférica y vista aérea. En este estudio, se seleccionó la proyección esférica ya que esta representa una perspectiva similar a la del operador humano y permite la identificación precisa de los troncos de los árboles en cada fotograma. De manera similar al enfoque utilizado en SqueezeSeg (WU et al., 2018), se realiza una proyección esférica que proporciona una representación densa y estructurada en forma de una cuadrícula 2D. Esta transformación se basa en la conversión de coordenadas cartesianas (x, y, z) a coordenadas esféricas.

En este estudio, el sensor Ouster OS0, configurado según lo descrito previamente, habría producido proyecciones 2D de las nubes de puntos originales con una resolución de 1024x128 píxeles, pero, tras aplicar la reducción de datos, se han generado imágenes con una resolución de 428x128 píxeles. Cada punto incorpora diversas propiedades, como intensidad, reflectividad, infrarrojo, rango y las coordenadas x, y, z, lo que da lugar a múltiples proyecciones por cada fotograma, donde cada canal representa una de estas características, generando un total de 7 proyecciones.

Esta representación estructurada y densa se asemeja a una imagen convencional, lo que facilita la aplicación de técnicas de procesamiento de imágenes y aprendizaje profundo. Los procesos de reducción y transformación de datos se implementaron utilizando MATLAB (THE MATHWORKS INC., 2024) en una computadora portátil AORUS 16X (modelo ASG-63ESC65SH) (GIGABYTE, TAIWÁN) equipada con un procesador Intel Core i9 de 14.^a generación, una GPU NVIDIA RTX 4070 y 32 GB de RAM.

Creación del conjunto de datos y etiquetado

Se seleccionaron aleatoriamente un total de 253 proyecciones esféricas de los tres escaneos para crear el conjunto de datos. Una vez obtenidas las proyecciones, el proceso de creación del conjunto de datos se dividió en cuatro subfases: (1) Selección de canales; (2) Etiquetado de reflectividad; (3) Generación de imágenes RGBA ;(4) Transferencia de datos etiquetados. Para proporcionar una descripción general clara de todo el proceso, la figura 2 ilustra el flujo de trabajo de

preparación y anotación del conjunto de datos.

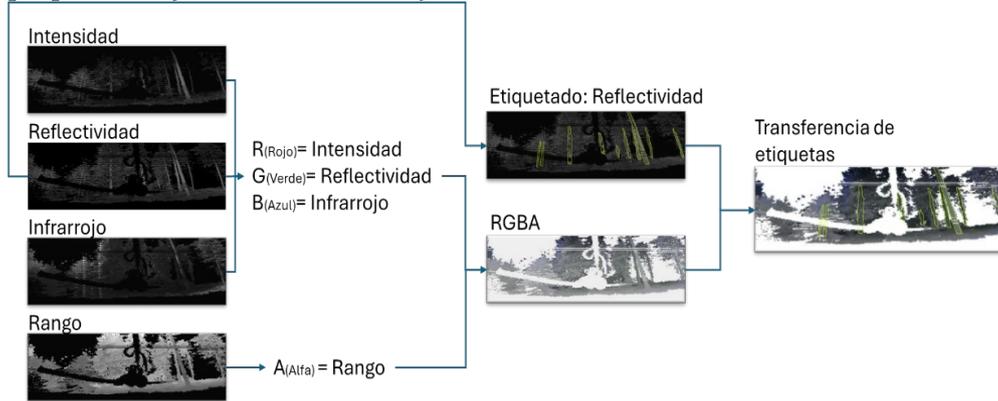


Figura 2. Flujo de trabajo de preparación y anotación del conjunto de datos.

Selección de canales

En este estudio, se ha propuesto maximizar la información extraída de cada punto. Para lograrlo, se han seleccionado cuatro canales de proyección: Intensidad, reflectividad, infrarrojo y rango. Se eligió el canal de rango en lugar de las coordenadas X, Y y Z, ya que integra de manera eficaz la información espacial de los tres, esto se debe a que el rango se calcula como la distancia euclidiana desde el sensor al punto en el espacio ($\text{Rango} = \sqrt{x^2 + y^2 + z^2}$).

Etiquetado de reflectividad

Para el proceso de etiquetado, se ha decidido etiquetar solo las proyecciones de reflectividad, esto se debe a que proporciona una medida de la capacidad de un objeto para reflejar la luz, ofreciendo información adicional más allá de la geometría y mejorando la precisión de la segmentación. A diferencia de la intensidad, la reflectividad es independiente del rango y el ángulo de incidencia, lo que proporciona una medida más consistente e interpretable de las características (VISWANATH et al., 2024). Esto permite una clara diferenciación de los troncos de los árboles, por lo que solo se etiquetaron las proyecciones de reflectividad.

El etiquetado de las proyecciones se llevó a cabo utilizando la plataforma Roboflow (DWYER et al., 2024) con la herramienta de etiquetado asistido “Smart Polygon” que utiliza la tecnología de SAM2 (RAVI et al., 2024), a su vez se realizaron correcciones manuales según fuera necesario. El uso de esta herramienta de etiquetado asistido redujo significativamente el tiempo necesario para el etiquetado de instancias.

Creación de imágenes RGBa

Las cuatro proyecciones de intensidad, reflectividad, infrarrojo y rango se combinaron para crear una imagen multicanal en formato RGBa. Cada canal codifica información específica de la siguiente manera:



- R (rojo): codifica los valores de intensidad, que representan la fuerza del pulso láser devuelto.
- G (verde): codifica la reflectividad, lo que proporciona información sobre las propiedades del material de la superficie.
- B (azul): codifica las mediciones de luz ambiental, que capturan la iluminación de fondo.
- A (alfa): codifica los datos de alcance, que indican la distancia de cada punto desde el sensor LiDAR.

Todas las anotaciones de las proyecciones de reflectividad se transfirieron a las imágenes RGBA correspondientes, lo que garantiza la coherencia y la alineación de la información etiquetada en todos los canales.

Utilización de Yolov8 para detección y segmentación por instancias de troncos

Este estudio se centra en la detección y segmentación de troncos de árboles mediante el uso de la red neuronal convolucional (CNN) YOLOv8. Este es un modelo de visión artificial diseñado para tareas de detección de objetos en tiempo real. Este modelo se basa en los avances de sus predecesores de la serie YOLO, ofreciendo mayor precisión y velocidad. Como arquitectura de última generación, YOLOv8 es particularmente adecuado para tareas que requieren alta eficiencia y precisión en entornos complejos (JOCHER et al., 2023).

Para abordar la tarea de segmentación de instancias de troncos, se ha seleccionado el modelo de aprendizaje profundo YOLOv8s-seg debido a su optimización para el procesamiento en tiempo real. Este modelo procesa imágenes en una sola pasada, lo que es fundamental para analizar grandes volúmenes de datos obtenidos a través del escaneo láser móvil. Con su enfoque de segmentación de una sola etapa, YOLOv8 reduce significativamente los tiempos de procesamiento al realizar la detección y la clasificación simultáneamente. En el contexto de la detección y segmentación de troncos a partir de datos proyecciones LiDAR, esta velocidad permite un procesamiento eficiente de grandes áreas forestales sin necesidad de calibraciones adicionales ni adaptaciones a condiciones de terreno variables. Además, aprovecha las proyecciones 2D compactas derivadas de nubes de puntos, logrando un equilibrio óptimo entre precisión y eficiencia computacional.

Arquitectura de YOLOv8-seg

El modelo YOLOv8s-Seg es una arquitectura robusta diseñada para tareas de segmentación de instancias, que permite la detección simultánea de objetos y la generación precisa de máscaras para instancias individuales. La estructura, ilustrada en la figura 3, se basa en las metodologías descritas en Fast Segment Anything (ZHAO et al., 2023) y YOLACT Real-time Instance Segmentation (BOLYA et al., 2019).

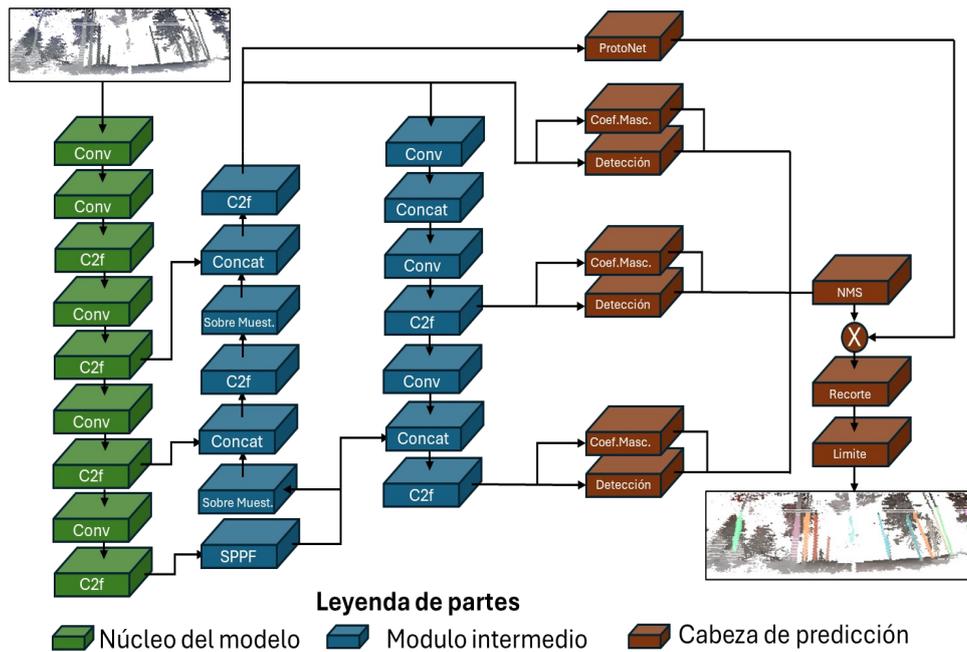


Figura 3. Descripción general de la arquitectura YOLOv8s-Seg.

El núcleo del modelo (en inglés *backbone*), está diseñado para extraer representaciones jerárquicas de las imágenes de entrada, esto se debe a que utiliza capas convolucionales para capturar características. La red del núcleo emplea bloques CBS (Convolución, Normalización por Lotes y Activación SiLU) y C2f, que mejoran la eficiencia y refinan las características de manera progresiva. Este diseño permite capturar información de múltiples resoluciones para tareas de detección y segmentación precisas; El Módulo Intermedio (en inglés *neck*), conecta el núcleo con la cabeza de predicción. Este módulo integra características a diferentes escalas utilizando arquitecturas como FPN o PAN, mejorando la representación espacial y contextual de los objetos detectados; La cabeza de predicción (en inglés *head*), se divide en dos ramas: detección y segmentación. La rama de detección es responsable de asignar las cajas delimitadoras y las clases de los objetos, mientras que la rama de segmentación genera máscaras de instancias utilizando la red ProtoNet, un enfoque que combina prototipos compartidos con coeficientes específicos para cada objeto detectado.

La segmentación se realiza a través de una combinación lineal ponderada de prototipos, seguida de una binarización mediante umbrales para separar objetos del fondo. El modelo incorpora postprocesamiento, como Supresión de No Máximos y filtrado por confianza (NMS), para refinar los resultados. La variante del modelo preentrenado que fue elegida, YOLOv8s-seg, utiliza el conjunto de datos COCO (LIN et al., 2014) e integra funciones de pérdida optimizadas específicamente para tareas de detección y segmentación. Además, ofrece configuraciones ajustables que permiten equilibrar de manera óptima la precisión, la velocidad y el tamaño.

Debido a su diseño eficiente y su capacidad de segmentación basada en prototipos, YOLOv8-Seg resulta particularmente adecuado para aplicaciones que demandan



alta precisión y rapidez en los resultados, como la segmentación de objetos en entornos complejos o en escenarios de procesamiento en tiempo real.

Métricas de evaluación

Se utilizan varias métricas clave para evaluar los resultados: precisión, *recall* (comúnmente conocido como ratio de verdaderos positivos), media de la Precisión Promedio (mAP) y la puntuación F1. La precisión se define como la proporción entre los objetos detectados correctamente y el total de objetos detectados. El *recall* mide la proporción de objetos clasificados correctamente respecto al total de objetos reales en el conjunto de datos.

La precisión y el *recall* se calculan en función del Índice de Intersección sobre Unión (IoU, por sus siglas en inglés), que se define como el área de intersección entre la región predicha y la región real, dividido por el área de su unión. Este índice evalúa qué tan bien las predicciones del modelo se alinean con los límites reales de los objetos en las imágenes.

Para una evaluación más integral del rendimiento general del modelo, se utiliza la media de la Precisión Promedio (mAP). Esta métrica se calcula como la media de los valores de Precisión Promedio para diferentes clases (en este caso, para una sola clase: tronco de árbol). La Precisión Promedio se define como el área bajo la curva de Precisión-Recall y proporciona una medida cuantitativa de la capacidad del modelo para detectar objetos de una clase específica con precisión. Específicamente, mAP@50 mide la precisión promedio con un umbral mínimo de IoU del 50%, mientras que mAP@50-95 calcula la precisión promedio en varios umbrales de IoU que van del 50% al 95%, con un incremento del 5%. Esto permite evaluar la precisión del modelo en diferentes niveles de IoU.

Para equilibrar la evaluación entre precisión y *recall*, también se emplea la puntuación F1, que es la media armónica entre estas dos métricas, proporcionando una visión más completa del compromiso entre ambas.

Flujo de datos

Para entrenar el modelo YOLOv8-seg, se utilizó un conjunto de datos compuesto por 253 imágenes RGBA. Estas imágenes se convierten en tensores y se normalizan considerando los valores medios y las desviaciones estándar de cada canal. Las etiquetas de las imágenes en los conjuntos de datos de entrenamiento y validación se proporcionan en el formato YOLOv8, representadas por las coordenadas de los vértices de los polígonos que describen las máscaras de los objetos en las imágenes.

La función de pérdida del modelo YOLOv8s-seg incluye: pérdida por clasificación, pérdida por predicción de cajas delimitadoras y pérdida derivada de la comparación entre la máscara predicha y la etiqueta real en la imagen. El conjunto de datos para el entrenamiento del modelo se divide en tres subconjuntos: entrenamiento, validación y prueba, en una proporción de 70/20/10. Durante el entrenamiento, el modelo minimiza la función de pérdida utilizando las imágenes

y sus máscaras correspondientes del conjunto de entrenamiento. Una vez finalizado el entrenamiento, el modelo se evalúa en el conjunto de prueba utilizando las métricas especificadas.

Este proceso se llevó a cabo utilizando la plataforma online Google Colab (GOOGLE, 2024) como plataforma para el entrenamiento y evaluación del modelo. Se empleó una configuración basada en Google Compute Engine, con un entorno de ejecución Python 3 (VAN ROSSUM & DRAKE, 2009), una CPU, 12,7 GB de memoria RAM y 107,7 GB de espacio en disco.

4. Resultados

Para evaluar la viabilidad de detectar y segmentar troncos individuales de árboles utilizando datos de sensores láser móviles y proyecciones esféricas 2D, ha realizado una evaluación incluyendo pruebas con el modelo pre entrenados YOLOv8s-seg, analizando su desempeño en tareas de detección de cajas delimitadoras y de máscaras.

Los resultados de este experimento se resumen en la **tabla 1**. Para ilustrar el rendimiento de segmentación, se seleccionó aleatoriamente un conjunto de imágenes del subconjunto de prueba (Figura 4).

Tabla 1. Resultados de métricas de evaluación.

YOLOv8s-seg		
Detección de cajas delimitadoras	Detección de máscaras	
Precisión	0,879	0,869
Recall	0,852	0,839
mAP50	0,914	0,899
mAP50-95	0,603	0,472

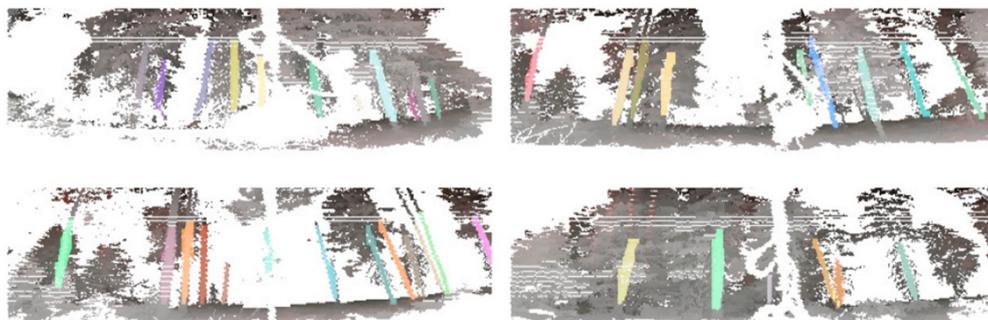


Figura 4. Identificación de troncos en imágenes del subconjunto prueba.

Para ilustrar mejor la dinámica de entrenamiento y el rendimiento del modelo, se presentan las curvas de pérdida de entrenamiento para las predicciones de cajas delimitadoras, clases y máscaras, así como las métricas de precisión, *recall*, mAP50 y mAP50-95 a lo largo de las épocas. Estas visualizaciones, incluidas en la Figura 5, muestran la convergencia de las funciones de pérdida y la estabilidad de las métricas de rendimiento durante el entrenamiento.

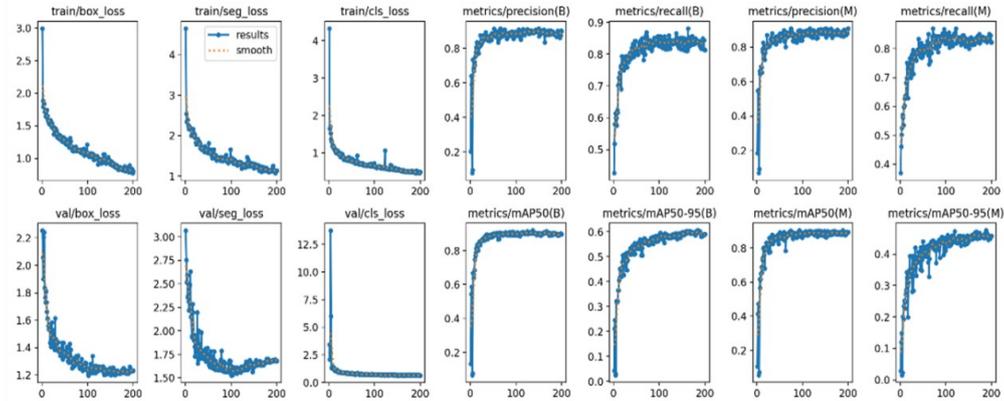


Figura 5. Evolución de las pérdidas y métricas durante la fase de entrenamiento.

El tiempo de inferencia para una sola imagen utilizando el modelo YOLOv8s-seg fue de 0,24 segundos.

El análisis de las curvas F1-Confianza, Precisión-Confianza y *Recall*-Confianza reveló que el mayor F1-score para todas las clases es 0.87 con un umbral de confianza de 0.481 (Figura 6). Para la inferencia, se seleccionó un umbral de confianza de 0.5, ya que está cerca del F1-score óptimo y proporciona un equilibrio práctico entre Precisión y *Recall*. Esta decisión asegura que se minimicen los falsos positivos y negativos, lo que hace que el modelo sea confiable para tareas de segmentación de troncos en escenarios reales.

Seleccionar un umbral de confianza de 0.5 simplifica el despliegue al evitar la necesidad de ajustes adicionales, manteniendo al mismo tiempo predicciones de alta calidad. La robustez de este umbral es particularmente importante en aplicaciones que requieren toma de decisiones en tiempo real.

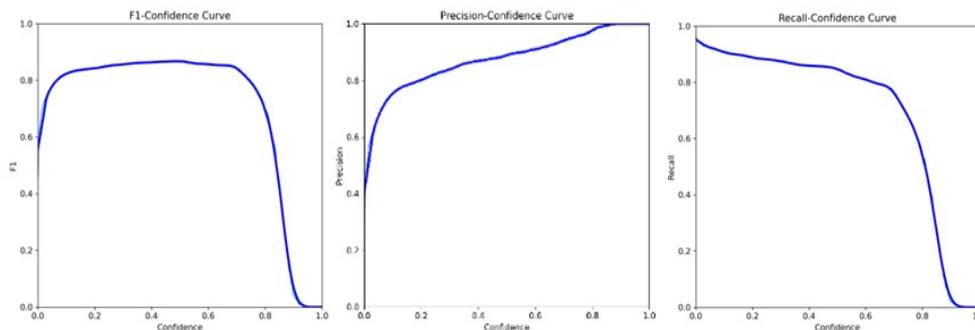


Figura 6. Curvas de confianza: F1, Precisión y Recall.

5. Discusión

Nuestra metodología se validó en campo durante un aprovechamiento final, corta



a hecho de pino silvestre en Finlandia. Durante el aprovechamiento, el sensor LiDAR recopiló datos mientras el operador efectuaba la utilizando una procesadora.

Los resultados demuestran que la detección de troncos de árboles basada en datos LiDAR y proyecciones esféricas 2D ofrece ventajas significativas frente a otros métodos, dada la eficiencia computacional de nuestro método, lo que aumenta la viabilidad de su implementación en tiempo real. Esto es así porque las redes CNN avanzadas están más optimizadas para su uso en imágenes 2D que en datos 3D. Además, emplear datos 2D permite reducir el tiempo de etiquetado de datos requerido para entrenar un modelo competente, lo que minimiza la intervención humana. Estas ventajas posicionan a esta metodología como una alternativa prometedora para aplicaciones prácticas en la gestión forestal de precisión.

En cuanto al modelo, se seleccionó la variante YOLOv8s-seg debido a su arquitectura compacta, ideal para implementaciones en dispositivos con recursos computacionales limitados (ZHAO et al., 2024), como GPUs de baja potencia. Su arquitectura optimizada permite procesar datos rápidamente, lo que resulta adecuado para imágenes de baja resolución, como proyecciones 2D derivadas de datos LiDAR (con configuraciones de resolución horizontal entre 512 y 2048 píxeles, y resoluciones verticales entre 16 y 128 píxeles). No obstante, resulta relevante estudiar variantes más pequeñas, como YOLOv8n (nano), ya que ofrece tiempos de inferencia inferiores a 0,24 segundos. Esto resulta vital para su uso en ciertas aplicaciones, como vehículos autónomos. Además, se aprovecharía el máximo potencial del sensor, que genera hasta 20 fotogramas por segundo. Sin embargo, en operaciones forestales se prioriza la precisión sobre la velocidad. Esto se debe a que las máquinas forestales pesadas, como las procesadoras, operan a velocidades reducidas en entornos forestales.

A pesar de su tamaño reducido, YOLOv8s-seg mantiene un equilibrio adecuado entre precisión y velocidad, garantizando resultados confiables en tareas de segmentación y detección. En este estudio, YOLOv8s-seg logró una precisión que confirma su capacidad para identificar troncos correctamente y minimizar falsos positivos. El *recall* alcanzado indica que el modelo detectó la mayoría de los troncos presentes. Además, los valores altos de mAP50 confirman el alto rendimiento del modelo con un umbral de IoU del 50%. Sin embargo, los valores más bajos de mAP50-95 revelan dificultades para mantener un rendimiento consistente con umbrales de IoU más estrictos. El análisis de las curvas F1-Confianza, Precisión-Confianza y *Recall*-Confianza determinó que el puntaje F1 más alto se alcanzó con un umbral de confianza de 0,481. No obstante, se seleccionó un umbral de confianza de 0,5 para las inferencias, ya que ofrece un equilibrio práctico entre precisión y *recall* cercano al óptimo. Esta elección minimiza falsos positivos y negativos, al tiempo que simplifica la implementación y garantiza predicciones consistentes y de alta calidad. A pesar de lo alentador de los resultados, también hay que tener en cuenta que esta prueba de concepto se basó en un conjunto de datos reducido (253 imágenes), y que la a la homogeneidad del escenario de estudio puede haber conducido a un sobreajuste del modelo al entorno.



Respecto al tiempo de cálculo y la aplicabilidad en tiempo real, el método presentado en este estudio logra un tiempo de inferencia del modelo segmentación de 0,24 segundos, permitiendo procesar hasta 4 fotogramas por segundo. Este rendimiento es suficiente para realizar segmentación por instancias en tiempo casi real en operaciones forestales, donde las procesadoras operan a velocidades reducidas. Sin embargo, este tiempo de inferencia se calculó utilizando recursos de Google Colab (GOOGLE, 2024), lo que podría diferir de las condiciones reales. Una implementación óptima requeriría integrar el procesamiento de datos LiDAR y la inferencia del modelo en un sistema propio, como un ordenador capaz de procesar los datos a la vez que se recolectan. Aunque no se evaluó explícitamente en este estudio, los requerimientos temporales de proyección de datos LiDAR, investigaciones previas han demostrado que las proyecciones esféricas pueden realizarse en tiempo real (WU et al., 2018). Además de esto, la resolución ampliada, tanto vertical como horizontal (128 píxeles y 2048-1024 píxeles, respectivamente), empleada en las proyecciones 2D, ofrece ventajas significativas para la detección precisa de troncos y la extracción de variables dendrométricas a nivel de árbol, como el diámetro a diferentes alturas, la curvatura y la inclinación. A pesar de ello, la caracterización de estas variables a partir de fotogramas presenta un desafío en torno a la estimación de variables, ya que cada fotograma únicamente refleja la parte del tronco visible desde la perspectiva del sensor.

Por último, resultaría relevante investigar si es posible estimar variables dendrométricas con precisión utilizando únicamente fotogramas individuales y determinar si esta información es suficiente para caracterizar el árbol en su totalidad. Además, es imprescindible analizar cómo los errores asociados al haz del láser del sensor Ouster OS0 afectan la precisión en la estimación de dichas variables (FAITLI et al., 2024).

6. Conclusiones

Este estudio presenta una metodología para la detección de troncos de árboles, desde una procesadora forestal, utilizando datos LiDAR, a través de transformaciones a proyecciones esféricas 2D de los mismos y un modelo CNN YOLOv8. Incluye una descripción exhaustiva de los procesos llevados a cabo, desde el manejo de los datos robóticos generados por el sensor LiDAR hasta la detección efectiva de troncos de árboles. Este enfoque tiene como objetivo abordar los desafíos de la automatización forestal proporcionando herramientas más eficientes y precisas para ayudar en las operaciones de aprovechamiento en entornos forestales dinámicos. Una característica clave de esta metodología es el uso de imágenes RGBA, que maximizan el potencial del sensor LiDAR al integrar información clave como la intensidad, la reflectividad, el alcance y el ambiente, lo que facilita un análisis más completo y robusto.

A pesar de los resultados positivos, este trabajo debe considerarse una prueba de concepto debido al tamaño limitado del conjunto de datos, que consta de 246 imágenes. El trabajo futuro planea abordar esta limitación mediante la recopilación de un conjunto de datos más grande y diverso, probando la metodología en diferentes especies, tipos de bosques y condiciones operativas para



evaluar su aplicabilidad y solidez en diversos escenarios. Esta limitación actualmente restringe la generalización de los resultados a otros entornos y condiciones. Sin embargo, los hallazgos sientan las bases para futuras investigaciones y desarrollos tecnológicos en este campo. Entre las recomendaciones para trabajos futuros se encuentra el desarrollo de algoritmos que integren todos los procesos descritos y permitan su implementación en tiempo real, con la velocidad suficiente para asistir al operador en la cabina. Adicionalmente, existe la necesidad de ampliar y diversificar el conjunto de datos utilizado para entrenar los modelos, lo que facilitaría la extrapolación de resultados a otras especies, tipos de bosque y condiciones operacionales. También se sugiere una mayor exploración de las características dendrométricas que se pueden obtener en tiempo casi real, explorando su potencial para implementar algoritmos de optimización de tronzado. Además, la creación o utilización de conjuntos de datos LiDAR etiquetados, tanto en 2D como en 3D, es esencial para mejorar la robustez y aplicabilidad de los modelos desarrollados.

Este estudio contribuye a avanzar en la automatización de procesos en el aprovechamiento forestal, donde la identificación precisa de árboles puede ayudar al operador con información relevante para optimizar las operaciones y mejorar la trazabilidad de los productos forestales. Los resultados abren nuevas oportunidades para la aplicación de tecnologías digitales en la gestión forestal de precisión, lo que representa un paso adelante en la modernización y mejora de la eficiencia en este sector.

7. Agradecimientos

El estudio forma parte del programa de doctorado financiado por la ayuda financiada con el apoyo del programa predoctoral AGAUR-FI ayudas (2024 FI-1 00896) Joan Oró de la Secretaría de Universidades e Investigación del Departamento de Investigación y Universidades de la Generalitat de Cataluña y del Fondo Social Europeo Plus. Asimismo, esta investigación ha recibido financiación del programa de investigación e innovación Horizon Europe de la Unión Europea bajo el acuerdo de subvención n.º 101082051 (Sintetic).

8. Bibliografía

ACHIM, A.; MOREAU, G.; COOPS, N. C.; AXELSON, J. N.; BARRETTE, J.; BÉDARD, S.; BYRNE, K. E.; CASPERSEN, J.; DICK, A. R.; D'ORANGEVILLE, L.; DROLET, G.; ESKELSON, B. N. I.; FILIPESCU, C. N.; FLAMAND-HUBERT, M.; GOODBODY, T. R. H.; GRIESS, V. C.; HAGERMAN, S. M.; KEYS, K.; LAFLEUR, B.; ... WHITE, J. C.; 2022. The changing culture of silviculture. *Forestry: An International Journal of Forest Research*, 95(2), 143-152.

BOLYA, D.; ZHOU, C.; XIAO, F.; LEE, Y. J.; 2019. YOLACT: Real-Time Instance Segmentation. 9157-9166.

BONAN, G. B.; 2008. Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, 320(5882), 1444-1449.



CAMPBELL, S.; O'MAHONY, N.; KRPALCOVA, L.; RIORDAN, D.; WALSH, J.; MURPHY, A.; RYAN, C.; 2018. Sensor Technology in Autonomous Vehicles: A review. 2018 29th Irish Signals and Systems Conference (ISSC), 1-4.

CHESBROUGH, H.; 2010. Business Model Innovation: Opportunities and Barriers. Long Range Planning, 43(2), 354-363.

CLOUDHRY, H.; O'KELLY, G.; 2018. Precision forestry: A revolution in the woods. McKinsey&Company.

DWYER, B.; NELSON, J.; HANSEN, T.; 2024. Roboflow (Version 1.0) [Software].

FAITLI, T.; HYYPPÄ, E.; HYYTI, H.; HAKALA, T.; KAARTINEN, H.; KUKKO, A.; MUHOJOKI, J.; HYYPPÄ, J.; 2024. Integration of a Mobile Laser Scanning System with a Forest Harvester for Accurate Localization and Tree Stem Measurements. Remote Sensing, 16(17), Article 17.

GHOBAKHLOO, M.; 2020. Industry 4.0, digitization, and opportunities for sustainability. Journal of Cleaner Production, 252, 119869.

GOOGLE.; 2024. Google Colab [Software]. Google.

GÜRDÜR BROO, D.; KAYNAK, O.; SAIT, S. M.; 2022. Rethinking engineering education at the age of industry 5.0. Journal of Industrial Information Integration, 25, 100311.

HOLZINGER, A.; SCHWEIER, J.; GOLLOB, C.; NOTHDURFT, A.; HASENAUER, H.; KIRISITS, T.; HÄGGSTRÖM, C.; VISSER, R.; CAVALLI, R.; SPINELLI, R.; STAMPFER, K.; 2024. From Industry 5.0 to Forestry 5.0: Bridging the gap with Human-Centered Artificial Intelligence. Current Forestry Reports, 10(6), 442-455.

JAVOID, M.; HALEEM, A.; SINGH, R. P.; SUMAN, R.; GONZALEZ, E. S.; 2022. Understanding the adoption of Industry 4.0 technologies in improving environmental sustainability. Sustainable Operations and Computers, 3, 203-217.

JELAVIC, E.; JUD, D.; EGLI, P.; HUTTER, M.; 2021. Towards Autonomous Robotic Precision Harvesting: Mapping, Localization, Planning and Control for a Legged Tree Harvester (arXiv:2104.10110). arXiv.

JELAVIC, E.; KAPGEN, T.; KERSCHER, S.; JUD, D.; HUTTER, M.; 2022, mayo 23. Harveri: A Small (Semi-)Autonomous Precision Tree Harvester. Innovation in Forestry Robotics: Research and Industry Adoption, ICRA 2022 IFRRRIA Workshop.



JERONIMO, S. M. A.; KANE, V. R.; CHURCHILL, D. J.; MCGAUGHEY, R. J.; FRANKLIN, J. F.; 2018. Applying LiDAR Individual Tree Detection to Management of Structurally Diverse Forest Landscapes. *Journal of Forestry*, 116(4), 336-346.

JHALDIYAL, A.; CHAUDHARY, N.; 2023. Semantic segmentation of 3D LiDAR data using deep learning: a review of projection-based methods. *Applied Intelligence*, 53(6), 6844-6855.

JOCHER, G.; QIU, J.; CHAURASIA, A.; 2023. Ultralytics YOLO (Versión 8.0.0).

JUTILA, J.; KANNAS, K.; VISALA, A.; 2007. Tree Measurement in Forest by 2D Laser Scanning. 2007 International Symposium on Computational Intelligence in Robotics and Automation, 491-496.

KEEFE, R. F.; ZIMBELMAN, E. G.; PICCHI, G.; 2022. Use of Individual Tree and Product Level Data to Improve Operational Forestry. *Current Forestry Reports*, 8(2), 148-165.

LENG, J.; SHA, W.; WANG, B.; ZHENG, P.; ZHUANG, C.; LIU, Q.; WUEST, T.; MOURTZIS, D.; WANG, L.; 2022. Industry 5.0: Prospect and retrospect. *Journal of Manufacturing Systems*, 65, 279-295.

LIANG, X.; HYYPPÄ, J.; KUKKO, A.; KAARTINEN, H.; JAAKKOLA, A.; YU, X.; 2014. The Use of a Mobile Laser Scanning System for Mapping Large Forest Plots. *IEEE Geoscience and Remote Sensing Letters*, 11(9), 1504-1508. *IEEE Geoscience and Remote Sensing Letters*.

LIANG, X.; KANKARE, V.; HYYPPÄ, J.; WANG, Y.; KUKKO, A.; HAGGRÉN, H.; YU, X.; KAARTINEN, H.; JAAKKOLA, A.; GUAN, F.; HOLOPAINEN, M.; VASTARANTA, M.; 2016. Terrestrial laser scanning in forest inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 63-77.

LIN, T.-Y.; MAIRE, M.; BELONGIE, S.; HAYS, J.; PERONA, P.; RAMANAN, D.; DOLLÁR, P.; ZITNICK, C. L.; 2014. Microsoft COCO: Common Objects in Context. En D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 740-755). Springer International Publishing.

LINDROOS, O.; LA HERA, P.; HÄGGSTRÖM, C.; 2017. Drivers of Advances in Mechanized Timber Harvesting – a Selective Review of Technological Innovation. *Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering*, 38(2), 243-258.



LIU, H.; XU, G.; LIU, B.; LI, Y.; YANG, S.; TANG, J.; PAN, K.; XING, Y.; 2025. A real time LiDAR-Visual-Inertial object level semantic SLAM for forest environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 219, 71-90.

MIETTINEN, M.; KULOVESI, J.; KALMARI, J.; VISALA, A.; 2010. New Measurement Concept for Forest Harvester Head. En A. Howard, K. Iagnemma, & A. Kelly (Eds.), *Field and Service Robotics* (pp. 35-44). Springer.

MIETTINEN, M.; OHMAN, M.; VISALA, A.; FORSMAN, P.; 2007. Simultaneous Localization and Mapping for Forest Harvesters. *Proceedings 2007 IEEE International Conference on Robotics and Automation*, 517-522.

MURTIYOSO, A.; CABO, C.; SINGH, A.; OBAYA, D. P.; CHERLET, W.; STODDART, J.; FOL, C. R.; BELOIU SCHWENKE, M.; REHUSH, N.; STEREŃCZAK, K.; CALDERS, K.; GRIESS, V. C.; MOKROŠ, M.; 2024. A Review of Software Solutions to Process Ground-based Point Clouds in Forest Applications. *Current Forestry Reports*, 10(6), 401-419.

NEVALAINEN, P.; LI, Q.; MELKAS, T.; RIEKKI, K.; WESTERLUND, T.; HEIKKONEN, J.; 2020. Navigation and Mapping in Forest Environment Using Sparse Point Clouds. *Remote Sensing*, 12(24), Article 24.

ÖHMAN, M.; MIETTINEN, M.; KANNAS, K.; JUTILA, J.; VISALA, A.; FORSMAN, P.; 2008. Tree Measurement and Simultaneous Localization and Mapping System for Forest Harvesters. En C. Laugier & R. Siegwart (Eds.), *Field and Service Robotics: Results of the 6th International Conference* (pp. 369-378). Springer.

PIERZCHAŁA, M.; GIGUÈRE, P.; ASTRUP, R.; 2018. Mapping forests using an unmanned ground vehicle with 3D LiDAR and graph-SLAM. *Computers and Electronics in Agriculture*, 145, 217-225.

PONSSE.; s. f. The cut-to-length method. Ponsse. Recuperado 17 de enero de 2025, de <https://www.ponsse.com/cut-to-length>

PRENDES, C.; ACUNA, M.; CANGA, E.; ORDOÑEZ, C.; CABO, C.; 2023. Optimal bucking of stems from terrestrial laser scanning data to maximize forest value. *Scandinavian Journal of Forest Research*, 38(3), 174-188.

PRENDES PÉREZ, C.; 2023. Algorithms for precision forestry inventories based on terrestrial point clouds, Tesis Doctoral, Universidad de Oviedo.

PROUDMAN, A.; RAMEZANI, M.; DIGUMARTI, S. T.; CHEBROLU, N.; FALLON, M.; 2022. Towards real-time forest inventory using handheld LiDAR. *Robotics and*



Autonomous Systems, 157, 104240.

RAVI, N.; GABEUR, V.; HU, Y.-T.; HU, R.; RYALI, C.; MA, T.; KHEDR, H.; RÄDLE, R.; ROLLAND, C.; GUSTAFSON, L.; MINTUN, E.; PAN, J.; ALWALA, K. V.; CARION, N.; WU, C.-Y.; GIRSHICK, R.; DOLLÁR, P.; FEICHTENHOFER, C.; 2024. SAM 2: Segment Anything in Images and Videos (arXiv:2408.00714). arXiv.

ROSSMANN, J.; SCHLUSE, M.; SCHLETTE, C.; BUECKEN, A.; KRAHWINKLER, P.; EMDE, M.; 2009. Realization of a highly accurate mobile robot system for multi purpose precision forestry applications. 2009 International Conference on Advanced Robotics, 1-6.

SAGAR, A.; KÄRHÄ, K.; EINOLA, K.; KOIVUSALO, A.; 2024. Assessing the Potential of Onboard LiDAR-Based Application to Detect the Quality of Tree Stems in Cut-to-Length (CTL) Harvesting Operations. Forests, 15(5), Article 5.

SIHVO, S.; VIRJONEN, P.; NEVALAINEN, P.; HEIKKONEN, J.; 2018. Tree Detection around Forest Harvester Based on Onboard LiDAR Measurements. 2018 Baltic Geodetic Congress (BGC Geomatics), 364-367.

THE MATHWORKS INC.; 2024. MATLAB R2024b [Software]. The MathWorks Inc.

VAN ROSSUM, G.; DRAKE, F. L.; 2009. Python 3 Reference Manual. CreateSpace.

VISSER, R.; OBI, O. F.; 2021. Automation and Robotics in Forest Harvesting Operations: Identifying Near-Term Opportunities. Croatian Journal of Forest Engineering: Journal for Theory and Application of Forestry Engineering, 42(1), 13-24.

VISWANATH, K.; JIANG, P.; SARIPALLI, S.; 2024. Reflectivity Is All You Need!: Advancing LiDAR Semantic Segmentation (arXiv:2403.13188). arXiv.

WU, B.; WAN, A.; YUE, X.; KEUTZER, K.; 2018. SqueezeSeg: Convolutional Neural Nets with Recurrent CRF for Real-Time Road-Object Segmentation from 3D LiDAR Point Cloud. 2018 IEEE International Conference on Robotics and Automation (ICRA), 1887-1893.

ZHANG, W.; WAN, P.; WANG, T.; CAI, S.; CHEN, Y.; JIN, X.; YAN, G.; 2019. A Novel Approach for the Detection of Standing Tree Stems from Plot-Level Terrestrial Laser Scanning Data. Remote Sensing, 11(2), Article 2.

ZHAO, M.; LIU, X.; WANG, K.; LIU, Z.; DONG, Q.; WANG, P.; SU, Y.; 2024. Welding Seam Tracking and Inspection Robot Based on Improved YOLOv8s-Seg Model.



Sensors, 24(14), Article 14.

ZHAO, X.; DING, W.; AN, Y.; DU, Y.; YU, T.; LI, M.; TANG, M.; WANG, J.; 2023. Fast Segment Anything (arXiv:2306.12156).