

Pérez Pablos, S.; Pagni Stanek, T.; Merino De Miguel, S. Desarrollo de un algoritmo para la detección, identificación y posicionamiento automático de árboles mediante deep learning y métodos topográficos

Desarrollo de un algoritmo para la detección, identificación y posicionamiento automático de árboles mediante deep learning y métodos topográficos

Pérez Pablos, Sergio ¹ Pagni Stanek, Tomás ¹ Merino De Miguel, Silvia ¹

¹ ETSI Montes, Forestal y del Medio Natural (Universidad Politécnica de Madrid), C/ José Antonio Novais, 10, 28040 Madrid

ORCID: Pagni Stanek [0009-0000-2527-1657](https://orcid.org/0009-0000-2527-1657) Merino De Miguel [0000-0002-4764-5311](https://orcid.org/0000-0002-4764-5311)

Correspondencia: sergiopzablos@gmail.com tomas.pagni@upm.es silvia.merino@upm.es

RESUMEN

Se presenta el desarrollo de un sistema automático para la detección, identificación y posicionamiento planimétrico de árboles individuales a partir de imágenes panorámicas terrestres georreferenciadas, sin empleo directo de sensores LiDAR en el cálculo posicional. La metodología integra un modelo de detección basado en Deep Learning (YOLOv8) con un algoritmo de seguimiento multiobjeto (Deep SORT) y un procedimiento topográfico de intersección directa simple para estimar coordenadas planimétricas. El dataset fue generado a partir de inventarios de vegetación en carreteras de la Comunidad de Madrid, utilizando imágenes equirectangulares capturadas mediante sistema Mobile Mapping. Se entrenó el modelo con imágenes etiquetadas manualmente, estructuradas en subconjuntos de entrenamiento, validación y prueba. El ajuste de hiperparámetros se realizó mediante optimización automática. Las detecciones obtenidas se asociaron entre imágenes consecutivas y se transformaron en lecturas angulares a partir de la geometría de la proyección panorámica y la orientación del sensor. La posición final de cada árbol se estimó mediante intersección de visuales desde estaciones GNSS conocidas. La validación se realizó comparando los resultados con la nube de puntos LiDAR del mismo sistema. Los resultados muestran que la combinación de detección 2D y métodos topográficos permite obtener posiciones planimétricas coherentes, con errores compatibles con inventarios forestales de carácter operativo, reduciendo significativamente la intervención manual.


Palabras clave: YOLOv8, inventario forestal, intersección directa, posicionamiento, visión por computador


Fecha de recepción: 13 febrero 2026 · Fecha de aceptación: 18 febrero 2026

Desarrollo de un algoritmo para la detección, identificación y posicionamiento automático de árboles mediante deep learning y métodos topográficos

Pérez Pablos, Sergio ⁽¹⁾, Pugini Stanek, Tomás ⁽¹⁾, Merino De Miguel, Silvia ⁽¹⁾

⁽¹⁾ ETSI Montes, Forestal y del Medio Natural (Universidad Politécnica de Madrid), 10 Madrid.

sergiopzablos@gmail.com ;  0009-0000-2527-1657, tomas.pugini@upm.es.

 0000-0002-4764-5311, silvia.merino@upm.es.

Resumen: Se presenta el desarrollo de un sistema automático para la detección, identificación y posicionamiento planimétrico de árboles individuales a partir de imágenes panorámicas terrestres georreferenciadas, sin empleo directo de sensores LiDAR en el cálculo posicional. La metodología integra un modelo de detección basado en *Deep Learning* (YOLOv8) con un algoritmo de seguimiento multiobjeto (Deep SORT) y un procedimiento topográfico de intersección directa simple para estimar coordenadas planimétricas. El dataset fue generado a partir de inventarios de vegetación en carreteras de la Comunidad de Madrid, utilizando imágenes equirectangulares capturadas mediante sistema *Mobile Mapping*. Se entrenó el modelo con imágenes etiquetadas manualmente, estructuradas en subconjuntos de entrenamiento, validación y prueba. El ajuste de hiperparámetros se realizó mediante optimización automática. Las detecciones obtenidas se asociaron entre imágenes consecutivas y se transformaron en lecturas angulares a partir de la geometría de la proyección panorámica y la orientación del sensor. La posición final de cada árbol se estimó mediante intersección de visuales desde estaciones GNSS conocidas. La validación se realizó comparando los resultados con la nube de puntos LiDAR del mismo sistema. Los resultados muestran que la combinación de detección 2D y métodos topográficos permite obtener posiciones planimétricas coherentes, con errores compatibles con inventarios forestales de carácter operativo, reduciendo significativamente la intervención manual.

Palabras clave: YOLOv8, inventario forestal, intersección directa, posicionamiento, visión por computador.

Development of an algorithm for automatic tree detection, identification and positioning using deep learning and topographic methods

Abstract: *This study presents the development of an automated system for individual tree detection, identification and planimetric positioning using georeferenced terrestrial panoramic imagery without direct LiDAR-based positioning. The methodology integrates a Deep Learning detection model (YOLOv8), a multi-object tracking algorithm (Deep SORT), and a topographic direct intersection method to estimate tree coordinates. The dataset was generated from roadside vegetation inventories in the Comunidad de Madrid (Spain), using equirectangular panoramic images acquired with a Mobile Mapping system. The model was trained with manually annotated images divided into training, validation and test subsets. Hyperparameter tuning was performed automatically to optimize performance. Detected objects were associated across multiple images and transformed into angular observations using panoramic projection geometry and sensor orientation data. Final tree positions were computed through the intersection of sight lines from GNSS-known stations. Validation was conducted against LiDAR point cloud data from the same acquisition system. Results indicate that combining 2D object detection with classical topographic methods enables consistent planimetric positioning suitable for operational forest inventories, significantly reducing manual intervention.*

Keywords: YOLOv8, forest inventory, direct intersection, positioning, computer vision.

1. INTRODUCCIÓN

La automatización del inventario de arbolado en entornos lineales constituye actualmente una interesante

línea de innovación en el ámbito de la geomática aplicada. La disponibilidad de sistemas *Mobile Mapping* equipados con cámaras panorámicas, sensores GNSS/IMU y escáneres LiDAR ha permitido generar

grandes volúmenes de datos geoespaciales de alta resolución en tiempos reducidos (Wu *et al.*, 2025). No obstante, el procesamiento y explotación eficiente de dichos datos para su transformación en información útil, continúa siendo un desafío técnico significativo.

En paralelo, el desarrollo reciente de arquitecturas de *Deep Learning* para detección de objetos ha demostrado una elevada capacidad para identificar automáticamente elementos discretos en imágenes complejas. En este sentido, los modelos de la familia YOLO (*You Only Look Once*) permiten realizar detección en tiempo real con un equilibrio adecuado entre precisión y coste computacional (Ali *et al.*, 2018), lo que los convierte en herramientas especialmente adecuadas para aplicaciones operativas.

En inventarios de arbolado viario, la detección automática de individuos constituye únicamente una parte del problema. Resulta igualmente necesario asociar las observaciones repetidas de un mismo pie en sucesivas imágenes, y estimar su posición espacial con precisión suficiente para su integración en bases de datos cartográficas (Ke *et al.*, 2024). De manera tradicional, el posicionamiento de árboles se ha venido realizando mediante explotación directa de la nube LiDAR (Mehtab, 2022). Sin embargo, ello implica un procesamiento costoso y no siempre necesario cuando el objetivo es exclusivamente planimétrico.

El presente trabajo propone un enfoque metodológico híbrido que combina: (i) detección 2D basada en *Deep Learning* sobre imágenes panorámicas equi-rectangulares, (ii) seguimiento multi-objeto para consolidación de identificadores y (iii) aplicación de un método topográfico clásico de intersección directa simple para el cálculo de coordenadas planimétricas. El sistema evita el uso directo del LiDAR en la estimación posicional, empleándolo únicamente como referencia independiente para validación.

El objetivo principal es evaluar la viabilidad técnica y la coherencia geométrica de este enfoque integrado para la realización de inventarios de arbolado en entornos antropizados con carácter operativo.

2. MATERIAL Y MÉTODOS

2.1. Área de estudio y datos de partida

El área de estudio se localiza en dos zonas de la Comunidad de Madrid (CAM) denominadas Zona Oeste y Zona Norte, cuya ubicación aparece en la Figura 1. La superficie total estudiada es de, aproximadamente, 15 km². En términos lineales, se han muestreado 260 km de tramos de carretera en la Zona Oeste, y 250 km en la Zona Norte, lo que hace un total de 510 km.

Las dimensiones y características del arbolado que aparece junto a estas carreteras permiten analizar la eficacia de la detección en distintos entornos y para distintas especies. Esto es debido al marcado gradiente altitudinal y la variedad de usos del suelo que quedan recogidos en la muestra de carreteras recorridas.

A modo de resumen, se pueden encontrar las siguientes formaciones arbóreas: Pinar de *Pinus pinea*, Pinar de *P.*

sylvestris, formaciones de *Quercus ilex* y de *Q. pyrenaica*, matorral disperso de *Juniperus oxycedrus* y *J. communis*, y vegetación colonizadora espontánea (principalmente *Salix sp.*, *Populus sp.* y *Fraxinus sp.*). Además, en los núcleos urbanos, destacan las especies más utilizadas en jardinería: *Platanus sp.*, *Morus sp.*, *Cedrus sp.*, y *Prunus sp.*, entre otros.

Los datos fueron capturados mediante un sistema *Mobile Mapping* IP-S3 de Topcon equipado con: (i) 6 cámaras panorámicas de 30 megapíxeles, (ii) un receptor GNSS y un sistema inercial (IMU), y (iii) sensor LiDAR de 32 haces y una velocidad de 700 000 pulsos por segundo. Las imágenes equi-rectangulares georreferenciadas constituyeron la base del procesamiento. Por otro lado, la nube de puntos LiDAR se empleó exclusivamente para validación.

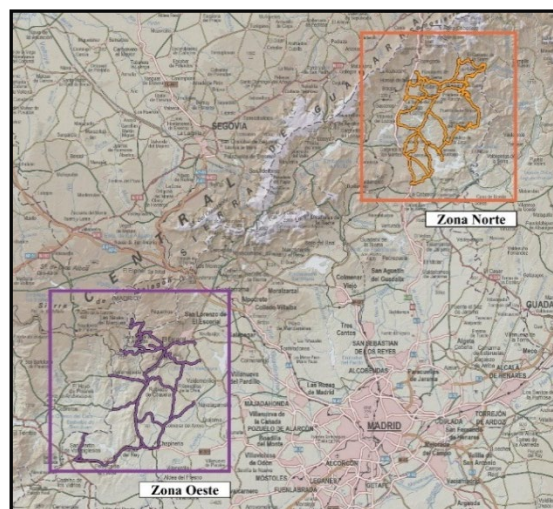


Figura 1 Localización del área de estudio en la CAM.

2.2. Métodos

La metodología seguida se estructura en cuatro etapas: (i) creación de un *dataset* para entrenamiento del modelo, (ii) entrenamiento del modelo de detección y seguimiento, (iii) extracción de parámetros de posición y mediciones, y (iv) validación de resultados.

2.2.1. Construcción del *dataset*

La primera fase consistió en la creación de un *dataset* formado por imágenes y etiquetas asociadas a cada árbol que apareciera en ellas. Para ello, se seleccionaron 921 imágenes representativas con alta variabilidad de especies, morfologías de copa, tamaño aparente y condiciones de iluminación.



Figura 2 Ejemplo de imagen con árboles identificados.

Las imágenes fueron etiquetadas manualmente utilizando *bounding boxes* con clase única (“árbol”), tal y como muestra la Figura 2. A continuación, el *dataset* se dividió en: (i) entrenamiento, (ii) validación, y (iii) prueba ($\approx 1\%$).

2.2.2. Entrenamiento, detección y seguimiento

Se empleó YOLOv8 como arquitectura de detección por su eficiencia en detección en una sola etapa y su capacidad para manejar objetos de escala variable. El entrenamiento de este algoritmo se realizó mediante optimizador *Adam* utilizando el *dataset* anteriormente construido y un ajuste automático de hiper-parámetros (*learning rate*, *batch size* y *momentum*). El rendimiento del modelo se evaluó mediante tres métricas: mAP (*mean Average Precision*), *precisión* (precisión) y *recall* (exhaustividad). El análisis de convergencia permitió seleccionar el modelo con mejor equilibrio entre sobreajuste y capacidad de generalización.

Las detecciones individuales por imagen generan múltiples observaciones del mismo árbol en capturas consecutivas. Dichas detecciones individuales fueron asociadas entre imágenes mediante un algoritmo *Deep SORT*, que combina: (i) predicción de estado mediante filtro de Kalman (predicción espacial), (ii) asociación basada en IoU (*Intersection over Union*), y (iii) métrica de similitud visual derivada de características profundas. Este procedimiento asigna un identificador único persistente a cada árbol individual detectado, permitiendo agrupar todas las visuales correspondientes antes del cálculo geométrico de la posición.

2.2.3. Posicionamiento mediante intersección directa

Cada detección proporciona una posición horizontal en la imagen panorámica. Para cada detección se calcula el azimut de la observación a partir de: (i) las coordenadas GNSS de la imagen, (ii) la orientación del sensor, y (iii) la posición horizontal del centroide de la *bounding box* en la imagen panorámica. Dado que la proyección equirectangular establece una relación lineal entre coordenada horizontal en píxeles y ángulo azimutal, se calculó el ángulo de observación θ_i para cada estación *i* mediante la ecuación 1.

$$\theta_i = \theta_{\text{sensor}} + \Delta\theta_{\text{pixel}} \quad (\text{Eq. 1})$$

donde θ_{sensor} corresponde al azimut del vehículo y $\Delta\theta_{\text{pixel}}$ se deriva de la posición relativa del *bounding box* respecto al centro de la imagen.

Con al menos dos visuales independientes desde dos estaciones A y B cualesquiera, la posición planimétrica del árbol V se obtiene por el método de intersección directa simple. Para ello es necesario conocer las coordenadas planimétricas X, Y de las estaciones A y B (dos posicionamientos del sensor derivadas de las coordenadas GNSS), y los azimuts de las alineaciones AV y BV, calculadas de acuerdo con la Eq. 1.

Como normalmente existen más de dos observaciones por árbol, se pudieron generar múltiples soluciones parciales, lo que permitió evaluar la precisión geométrica en función de variables como el ángulo de intersección, seleccionar aquellas configuraciones más robustas o

descartar aquellas configuraciones más próximas a la colinealidad.

2.2.4. Validación

La validación se realizó comparando las coordenadas planimétricas obtenidas a partir del procesado de las imágenes equi-rectangulares, con las posiciones medidas sobre la nube LiDAR que había sido captada por el mismo sistema (*Mobile Mapping IP-S3*) y de manera simultánea, eliminando así la variabilidad GNSS externa y evaluando exclusivamente el error metodológico.

La nube de puntos LiDAR, previamente georreferenciada, se utilizó para identificar manualmente posiciones de referencia para un subconjunto de árboles. De esta manera, se compararon las coordenadas planimétricas estimadas mediante el método propuesto con dichas posiciones de control, obteniendo métricas de error absoluto y análisis de dispersión.

3. RESULTADOS

El modelo YOLOv8 mostró una convergencia estable durante el entrenamiento, alcanzando valores elevados de mAP y manteniendo una adecuada relación entre precisión y exhaustividad. El *dataset* de entrenamiento confirmó la capacidad del modelo para detectar árboles bajo condiciones muy variables de iluminación, fondo y solapamiento parcial entre árboles.

El módulo de seguimiento mediante *Deep SORT* redujo significativamente las detecciones redundantes y permitió consolidar trayectorias coherentes para cada individuo, mejorando la coherencia espacial de los azimuts. La tasa de asignación incorrecta fue baja y se concentró principalmente en zonas con alta densidad de arbolado y copas solapadas.

Desde el punto de vista geométrico, la calidad del posicionamiento estuvo directamente relacionada con el ángulo de intersección entre visuales. Configuraciones próximas a 90° generaron soluciones estables con menor sensibilidad a errores angulares, mientras que ángulos pequeños y grandes incrementaron la incertidumbre planimétrica. Además, también influyen en la calidad del posicionamiento tanto la distancia entre posiciones del sensor utilizadas, como la confianza de detección.

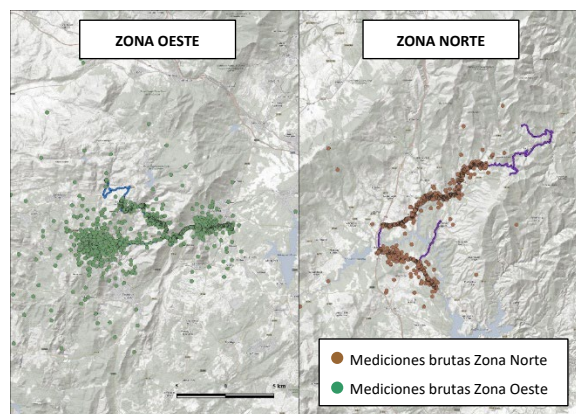


Figura 3 Distribución de todas las localizaciones (incluidos falsos positivos y errores de medición).

La Figura 3 muestra la distribución de los posicionamientos iniciales, es decir, incluidos los falsos positivos y los errores de medición. La figura 4 muestra aquellos posicionamientos derivados de la metodología expuesta y que contaron con incertidumbres aceptables.

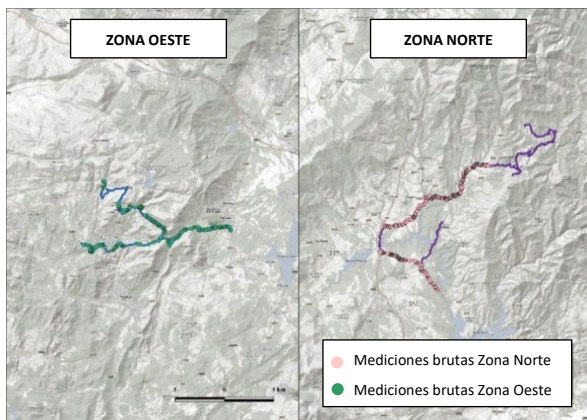


Figura 4 Distribución de localizaciones (solo mediciones con valores de incertidumbre aceptables).

La comparación con LiDAR evidenció que el error planimétrico se mantuvo dentro de rangos compatibles con inventarios de arbolado operativos en entornos lineales. Las Tablas 1 y 2 muestran el error por rangos, y las frecuencias absoluta y relativa para la Zona Oeste y Norte, respectivamente. En total se utilizaron 286 árboles para la validación de las posiciones en la Zona Oeste y 275 para la Norte. La dispersión observada responde a: (i) los errores angulares derivados de la discretización en píxeles, (ii) la incertidumbre en la determinación del ángulo azimutal del sensor, y (iii) la aproximación puntual del árbol mediante *bounding box* 2D.

En el caso de la Zona Oeste (Tabla 1), la mayoría de las observaciones (57%) presentaron errores inferiores a 3 m. Por otro lado, el mayor porcentaje se encontró en el intervalo 2-3 m. Sin embargo, un 24% de las mediciones exhibieron errores superiores a 4 m. En términos globales, el error medio fue de 2.8 m, con una desviación estándar global de 7.5 m. Los resultados para la Zona Norte (Tabla 2) mostraron un 60% de los árboles posicionados con un error inferior a 3 m. Sin embargo, un 22% de los pies fueron posicionados con más de 5 m de error. En términos globales, el error medio fue de 4.2 m, con una desviación estándar global de 5.3 m.

Tabla 1. Distribución del error en la Zona Oeste.

Zona	Rango error (m)	Frecuencia absoluta	Frecuencia relativa (%)
Oeste	0 – 1	14	5
	1 – 2	57	20
	2 – 3	92	32
	3 – 4	55	19
	4 – 5	32	11
	5 – 6	14	5
	> 6	22	8
	Total:	286	100%

Tabla 2. Distribución del error en la Zona Norte.

Zona	Rango error (m)	Frecuencia absoluta	Frecuencia relativa (%)
Norte	0 – 1	33	12
	1 – 2	81	29
	2 – 3	52	19
	3 – 4	35	13
	4 – 5	13	5
	> 5	61	22
	Total:	275	100%

En conjunto, los resultados demuestran que la combinación de detección automática y métodos topográficos clásicos constituye una alternativa técnicamente consistente para la generación de cartografía puntual de arbolado viario, con reducción sustancial del tiempo de procesamiento respecto a metodologías basadas exclusivamente en explotación de datos LiDAR.

4. CONCLUSIONES

Se ha desarrollado un sistema automatizado que integra: (i) detección 2D basada en *Deep Learning*, (ii) seguimiento multi-objeto, e (iii) intersección directa simple topográfica clásica. La metodología permite estimar posiciones planimétricas de árboles individuales utilizando exclusivamente imágenes panorámicas georreferenciadas para el cálculo, empleando LiDAR únicamente para validación. El enfoque demuestra que la combinación de inteligencia artificial y topografía clásica constituye una alternativa viable y eficiente para inventarios lineales de arbolado, reduciendo tiempos de procesamiento y dependencia de medición manual. Futuras mejoras podrían incluir ajuste robusto por mínimos cuadrados y estimación altimétrica.

5. AGRADECIMIENTOS

Los autores agradecen a RELIEVE MDC la cesión de los datos y herramientas empleados en este estudio.

6. REFERENCIAS

- Ali, W., Abdelkarim, S., Zidan, M., Zahran, M., & El Sallab, A. (2018). Yolo3d: End-to-end real-time 3d oriented object bounding box detection from lidar point cloud. In: Proceedings of the European conference on computer vision (ECCV) workshops.
- Ke, H., Li, H., Wang, B., Tang, Q., Lee, Y. H., & Yang, C. F. (2024). Integrations of Labellmg, You Only Look Once (YOLO), and Open Source Computer Vision Library (OpenCV) for Chicken Open Mouth Detection. *Sensors & Materials*, 36.
- Mehtab, S. (2022). Deep neural networks for road scene perception in autonomous vehicles using LiDARs and vision sensors (Doctoral dissertation, PhD Thesis, Auckland University of Technology, New Zealand).
- Wu, Y., Zhong, S., Ma, Y., Zhang, Y., & Liu, M. (2025). Application of SLAM-based mobile laser scanning in forest inventory: Methods, progress, challenges, and perspectives. *Forests*, 16(6), 920.