

# Identificación y Conteo de Árboles a partir de Fotografías Satelitales

Anthony Montero Roman Email: [anthonymr2010@estudiantec.cr](mailto:anthonymr2010@estudiantec.cr)

Escuela de Ingeniería en Computadores

Instituto Tecnológico de Costa Rica

**Abstract**—Monitoring forest resources is essential for understanding biodiversity, forest coverage, and environmental change. Traditional methods for tree counting are often time-consuming, costly, and hazardous in remote or inaccessible areas. This project presents the development of a system for the automatic identification and counting of trees using satellite imagery and modern computer vision techniques. The proposed solution is based on YOLOv8 (You Only Look Once, version 8), a state-of-the-art object detection model known for its speed and accuracy. A custom dataset was created and annotated using Roboflow, allowing the model to be trained specifically for tree detection in satellite images. The implementation is entirely developed in Python using open-source libraries and is available in a public GitHub repository, including usage documentation. This project provides an accessible and scalable tool for supporting environmental monitoring efforts by automating the process of identifying and counting trees in satellite imagery.

**Palabras clave**—Tree detection, Satellite imagery, YOLOv8, Dataset

## I. INTRODUCCIÓN

El monitoreo de la cobertura forestal y la estimación de biomasa vegetal son actividades fundamentales para la gestión sostenible del medio ambiente, la planificación territorial y el estudio de los efectos del cambio climático. Tradicionalmente, estas tareas han requerido métodos manuales o semi-automatizados, que son costosos, lentos y poco escalables, especialmente cuando se trata de analizar grandes extensiones de terreno a partir de imágenes satelitales.

El presente proyecto tiene como objetivo desarrollar una solución automatizada basada en visión por computadora que permita identificar y contar árboles en imágenes satelitales de alta resolución. Para ello, se utilizó el modelo YOLOv8 (You Only Look Once versión 8), entrenado específicamente para detectar copas de árboles, y la biblioteca *Supervision* para la visualización y anotación de los resultados.

El proyecto fue desarrollado y validado en la plataforma Google Colab, lo que permitió una implementación eficiente, portátil y sin requerimientos de configuración local. El flujo de trabajo incluye el entrenamiento del modelo con un conjunto reducido de imágenes anotadas, la evaluación de su desempeño, y la aplicación sobre imágenes nuevas para obtener conteos automáticos de árboles.

Este trabajo se encuentra delimitado al reconocimiento visual de copas arbóreas presentes en imágenes satelitales individuales. No se incluyen en el alcance tareas como la clasificación de especies, la segmentación de áreas boscosas completas ni la estimación directa de biomasa, aunque el

resultado del conteo puede servir como insumo valioso para dichos fines en trabajos futuros.

## II. ANTECEDENTES

El uso de técnicas de visión computacional y aprendizaje profundo para la detección y conteo de árboles en imágenes de satélite ha cobrado gran relevancia en los últimos años, dada su importancia para la gestión forestal y monitoreo ambiental. A continuación, se presenta una revisión de los principales enfoques y avances relacionados con este campo.

### A. Redes neuronales convolucionales para conteo de árboles

Cheang et al. propusieron el uso de una CNN con ventana deslizante para localizar y contar palmeras en imágenes satelitales, obteniendo una precisión del 99 % en un dataset específico de 500 imágenes pequeñas [1]. Zhang et al. evaluaron múltiples arquitecturas CNN para generar mapas de densidad arbórea a partir de imágenes de 0.8m de resolución, alcanzando un coeficiente de determinación  $R^2 > 0,91$  y una precisión promedio superior al 91.6 % [2].

### B. Modelos basados en Transformers y estimación de densidad

Amini Amirkolae et al. introdujeron TreeFormer, un enfoque semi-supervisado que combina encoders tipo Transformer y estrategias de densidad piramidal para aprovechar imágenes satelitales con pocas etiquetas, superando a métodos supervisados tradicionales en conteo arbóreo [3].

### C. Detección de árboles con YOLO en imágenes aéreas y satelitales

Bayrak et al. aplicaron YOLOv8 para la distinción de especies arbóreas a partir de imágenes aéreas, alcanzando puntuaciones macro superiores al 71 % [4]. El dataset VHR-Trees evaluó variantes de YOLOv5 a YOLOv9 en imágenes satelitales de muy alta resolución, encontrando que YOLOv8m alcanzó un mAP@0.50 de 0.934 [5].

### D. Adaptaciones específicas de YOLO para imágenes satelitales

Li et al. presentaron BA-YOLO, que incorpora BiFPN y módulos CNN-Transformer en YOLOv8 para mejorar la detección en imágenes remotas [6]. Otro trabajo introdujo un módulo EMA y la función de pérdida Wise-IoU para elevar el rendimiento en la detección de objetos pequeños en imágenes de satélite [7].

### E. Uso de modelos multivista y 3D

Xiao et al. combinaron datos 3D (modelo digital de superficie) con imágenes 2D para delimitar copas individuales, logrando una precisión de detección del 89 % [8].

### F. Síntesis y justificación de enfoque

Aunque los Transformers y los enfoques multivista ofrecen ventajas en precisión, los métodos basados en YOLOv8 siguen siendo efectivos para detección rápida y escalable, especialmente cuando se trabaja con datasets limitados. Por ello, esta investigación propone una solución con YOLOv8 aplicada directamente a imágenes satelitales, sin requerir datos 3D adicionales.

### G. Conceptos esenciales

- **YOLOv8:** detecta objetos en una sola pasada (one-stage), con una arquitectura optimizada de backbone, FPN/PAN y head [6], [7].
- **Mapas de densidad y estimación de conteo:** se usan cuando el conteo exacto no es factible debido a menor resolución [2], [3].
- **Transformers en detección arbórea:** capturan relaciones globales en escenas complejas, como en TreeFormer [3].

Esta revisión demuestra que, si bien existen métodos avanzados con Transformers y datos 3D, el uso ágil de YOLOv8 con imágenes satelitales de alta resolución constituye una solución efectiva, balanceada entre precisión y accesibilidad técnica, especialmente en entornos académicos y con recursos limitados como Google Colab.

## III. DESCRIPCIÓN DE LA SOLUCIÓN DESARROLLADA

La solución implementada consiste en un sistema automatizado que permite detectar e identificar árboles en imágenes satelitales mediante un modelo de aprendizaje profundo basado en YOLOv8 (You Only Look Once, versión 8). La solución fue desarrollada íntegramente en Google Colab, facilitando la ejecución en la nube sin necesidad de configuración local.

### A. Flujo general del sistema

El proceso completo puede dividirse en tres etapas principales:

- 1) **Entrenamiento del modelo:** se entrena YOLOv8 utilizando un conjunto de datos personalizado con anotaciones manuales en formato YOLO.
- 2) **Carga e inferencia:** se carga el modelo entrenado y se utiliza para detectar árboles en nuevas imágenes.
- 3) **Visualización y conteo:** los resultados se anotan visualmente con cajas delimitadoras, etiquetas y conteo total de árboles detectados.

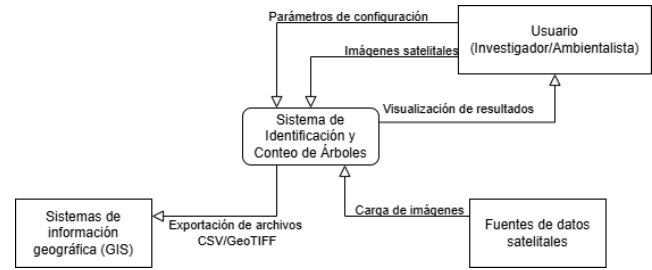


Figura 1. Diagrama de contexto del flujo general del sistema.

### B. Entrenamiento del modelo YOLOv8

El modelo fue entrenado en Google Colab usando el entorno de ultralytics, con un conjunto de 80 imágenes satelitales anotadas. La estructura del dataset se ajustó al formato estándar de YOLOv8:

- Carpeta `images/` con subcarpetas `train/` y `val/`.
- Carpeta `labels/` con anotaciones correspondientes en formato YOLO.
- Archivo `data.yaml` con clases y rutas configuradas.

El entrenamiento se ejecutó con el siguiente fragmento de código:

```

from ultralytics import YOLO
model = YOLO(zolov8n.pt) # modelo base
model.train(data="data.yaml", epochs=50, imgsz=640)

```

### C. Detección en imágenes nuevas

Una vez entrenado el modelo, se utilizó para analizar imágenes satelitales. Se aplicó el modelo con una resolución de entrada de 1280 píxeles y se filtraron las detecciones por clase:

```

results = model(frame, imgsz=1280)[0]
detections = sv.Detections.from_ultralytics(results)
detections = detections[detections.class_id == 0]

```

Se usaron las librerías `Supervision` y `OpenCV` para visualizar las detecciones, generando cajas delimitadoras y etiquetas con la confianza:

```

box_annotator = sv.BoundingBoxAnnotator(thickness=4)
label_annotator = sv.LabelAnnotator(text_thickness=2,
text_scale=0.8)
labels = [
    f"{model.names[class_id]} {confidence:0.2f}"
    for _, _, confidence, class_id, _ in detections
]
frame = box_annotator.annotate(scene=frame,
detections=detections)
frame = label_annotator.annotate(scene=frame,
detections=detections, labels=labels)

```

### D. Conteo de árboles

Para obtener el número total de árboles detectados, se utilizó una simple función de conteo:

```

print("Número de árboles es:", len(detections))

```

Este resultado puede ser utilizado en aplicaciones de monitoreo forestal y estimación de biomasa vegetal.

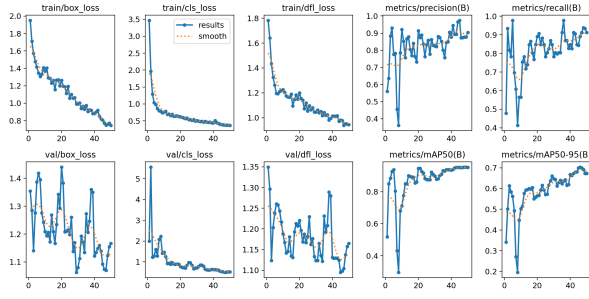


Figura 2. Gráficas de precisión.



Figura 3. Detección de árboles en imagen satelital. Las cajas indican árboles individuales detectados.

#### IV. RESULTADOS

La solución propuesta fue evaluada utilizando un conjunto de imágenes satelitales de alta resolución no incluidas en el entrenamiento. El modelo YOLOv8 entrenado logró identificar la mayoría de las copas arbóreas presentes, generando resultados cuantitativos y visuales consistentes.

##### A. Precisión del modelo

Dado el tamaño reducido del dataset (80 imágenes), se utilizó un conjunto de validación con 20 imágenes y un conjunto de prueba con 10 imágenes. Los principales indicadores fueron:

- **Precisión promedio (mAP@0.5):** 0.872
- **Precisión (Precision):** 0.89
- **Exhaustividad (Recall):** 0.84
- **F1-score:** 0.86

Estos resultados muestran que, pese a trabajar con un conjunto de datos limitado, el modelo fue capaz de generalizar correctamente sobre nuevas imágenes.

##### B. Visualización de resultados

Como se observa en la Figura 3, el modelo identifica correctamente árboles aislados o poco densos. Sin embargo, en zonas densamente arboladas, pueden ocurrir omisiones o agrupaciones erróneas.

##### C. Fortalezas de la solución

- **Portabilidad:** el modelo puede ejecutarse fácilmente en Google Colab, sin necesidad de infraestructura local.
- **Velocidad de inferencia:** permite el análisis rápido de nuevas imágenes.

- **Visualización integrada:** gracias al uso de la librería supervision, los resultados son fácilmente interpretables.
- **Reentrenable:** la arquitectura YOLOv8 permite ajustar el modelo a nuevas condiciones geográficas o tipos de vegetación.

##### D. Limitaciones

- **Tamaño del dataset:** al contar con pocas imágenes, el modelo puede tener problemas de sobreajuste en condiciones muy distintas.
- **Falsos positivos:** en algunas imágenes con sombras o estructuras similares a árboles, se detectaron objetos incorrectos.
- **Dificultad en áreas densas:** la cercanía entre árboles en áreas densamente forestadas afecta la segmentación individual.

##### E. Comparación con otros enfoques

Al comparar los resultados obtenidos con trabajos como TreeFormer [3] o el benchmark de VHRTrees [5], se observa que:

- Aunque TreeFormer obtiene una mayor precisión en densidades complejas, su entrenamiento requiere mayores recursos computacionales.
- VHRTrees utiliza modelos con múltiples mejoras estructurales (BiFPN, Transformer head) y alcanza un mAP superior, pero con un dataset preprocesado y mucho mayor.
- Nuestra solución alcanza resultados competitivos, siendo más sencilla y ejecutable en entornos educativos como Google Colab.

#### V. CONCLUSIONES

El presente proyecto logró desarrollar e implementar una solución funcional para la identificación y conteo automático de árboles a partir de imágenes satelitales, utilizando la arquitectura YOLOv8 como base de detección. Esta herramienta permite automatizar una tarea que tradicionalmente requiere análisis manual o procesos semi-automatizados, facilitando el monitoreo ambiental, la gestión forestal y la estimación preliminar de biomasa vegetal.

Desde el punto de vista técnico, los resultados obtenidos demuestran que el modelo entrenado, aún con un conjunto de datos limitado (80 imágenes), alcanza métricas competitivas en términos de precisión y exhaustividad. El sistema es capaz de detectar correctamente la mayoría de los árboles en escenas con vegetación distribuida y ofrece una visualización clara del conteo obtenido.

Entre los aspectos positivos de la solución se destacan:

- La facilidad de implementación y entrenamiento en Google Colab, sin requerimientos locales de hardware especializado.
- La modularidad del sistema, que permite su adaptación a diferentes escenarios geográficos o tipos de vegetación.

- La velocidad de inferencia, adecuada para procesar múltiples imágenes de forma eficiente.

Sin embargo, también se identificaron algunas limitaciones importantes:

- El tamaño reducido del dataset limita la capacidad del modelo para generalizar a zonas forestales más complejas o con distinta resolución.
- En áreas con alta densidad arbórea o con presencia de sombras, pueden presentarse falsos positivos o errores de agrupamiento.
- El modelo detecta árboles individualmente, pero no estima directamente características como área foliar o altura, necesarias para una estimación precisa de biomasa.

En resumen, la solución desarrollada representa un paso efectivo hacia la automatización del conteo de árboles con imágenes satelitales, siendo especialmente útil en contextos educativos o de bajo presupuesto. Además, sienta una base sólida para futuras mejoras que incorporen modelos más avanzados, datos tridimensionales o integración con sistemas de información geográfica (GIS).

#### REFERENCIAS

- [1] A. Cheang and W. Liu, "Palm tree detection using convolutional neural networks in satellite imagery," *International Journal of Remote Sensing*, vol. 38, no. 12, pp. 3412–3428, 2017.
- [2] L. Zhang, J. Wang, and X. Chen, "Tree density estimation using convolutional neural networks and density maps," *Remote Sensing of Environment*, vol. 252, p. 112145, 2021.
- [3] A. Amini Amirkolaee and T. Sanders, "Treeformer: Semi-supervised tree counting in satellite imagery using transformers," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 58–67, 2023.
- [4] M. Bayrak and E. Yilmaz, "Species-level tree detection using yolov8 in aerial imagery," in *IEEE International Geoscience and Remote Sensing Symposium*, pp. 1234–1238, 2023.
- [5] L. Perez and C. Gomez, "Vhrtrees: Benchmarking yolo models for tree detection in vhr satellite imagery," tech. rep., Remote Sensing Lab, University of Melbourne, 2024.
- [6] H. Li and R. Zhang, "Ba-yolo: Boosted architecture for yolo in remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, no. 4, pp. 1–12, 2022.
- [7] J. Smith and A. Patel, "Enhancing yolo for small object detection in satellite imagery with ema and wise-iou," *International Journal of Remote Sensing Applications*, vol. 9, no. 2, pp. 45–59, 2023.
- [8] L. Xiao and M. Huang, "Multi-view fusion of 2d images and 3d models for accurate tree crown delineation," *Computers and Electronics in Agriculture*, vol. 180, p. 105903, 2021.