

Detección de palmas de aceite adultas en imágenes aéreas y ortoimágenes (RGB) usando modelos de aprendizaje profundo

Diego Alejandro García Cárdenas¹, Carlos Ricardo Bojacá¹

¹Área de Geomática, autor para correspondencia: dagarcia@cenipalma.org Corporación Centro de Investigación en Palma de Aceite, Cenipalma

Introducción

El cultivo de palma de aceite desempeña un papel fundamental en la economía colombiana, representando una fuente significativa de ingresos y empleo; según la Federación Nacional de Cultivadores de Palma de Aceite (Fedepalma), este sector genera más de 170.000 empleos directos e indirectos, contribuyendo al desarrollo social y económico de comunidades enteras (Fedepalma, 2023). Colombia se ha consolidado como el cuarto productor mundial de aceite de palma y el primero en América Latina, con una producción que supera las 1,7 millones de toneladas anuales (Cenipalma, 2022). Sin embargo, el cultivo de palma de aceite también enfrenta desafíos ambientales. Las extensas plantaciones requieren un manejo sostenible del cultivo. En este contexto, la tecnología juega un papel crucial para optimizar la producción y minimizar el impacto ambiental. Los avances recientes en la tecnología de vehículos aéreos no tripulados (UAV), combinados con las crecientes capacidades de la visión por computadora, han revolucionado el monitoreo agrícola (Bouguettaya *et al.*, 2022; Boursianis *et al.*, 2022; Wakchaure *et al.*, 2023). En este trabajo, se exploró el uso de modelos de aprendizaje profundo "you only look once" (YOLO), específicamente YOLOv8x y YOLO11x, para la detección y el conteo de palmas de aceite adultas en imágenes aéreas y ortoimágenes. Se entrenaron y compararon dos modelos, YOLOv8x y YOLO11, evaluando su rendimiento mediante métricas estándar como Precisión, Recall, la Precisión Media Promedio con un umbral de 0,5 (mAP:0,5) y la Precisión Media Promedio con umbrales de 0,5 a 0,95 (mAP:0,5:0,95). El modelo YOLOv8x fue validado en escenarios reales de plantaciones, lo que permitió analizar su capacidad de generalización.

Objetivo

Evaluar el rendimiento de modelos de aprendizaje profundo "you only look once" (YOLO) para la detección y conteo de palmas de aceite adultas en imágenes aéreas y ortoimágenes (RGB) obtenidas a partir de drones.

Metodología

Un total de 536 imágenes aéreas fueron seleccionadas de las zonas palmeras de Colombia (Central, Oriental y Norte). Las dimensiones de las imágenes comprendieron medianas de 5472 píxeles de ancho por 3648 píxeles de alto. Para el etiquetado, se definieron las clases Árbol, Cobertura, Palma anómala, Palma sana y Suelo. Al conjunto de datos se le aplicó la técnica de aumento de datos que consiste en ampliar la cantidad de imágenes y etiquetas del subconjunto de entrenamiento (Figura 1). Se aplicaron aumentos de: rotación de 90° (sentido horario, sentido antihorario, invertido), rotación entre -10° y +10°, corte transversal: ±10° horizontal, ±10° vertical, saturación entre -25 % y +25 %, brillo entre -20 % y +20 %, exposición entre -10 % y +10 %, desenfoque hasta 2,5 píxeles y ruido hasta 1,5 % de píxeles. También se agregaron aumentos de cajas delimitadoras como recortes (acercamiento mínimo 0 %, acercamiento máximo 20 %), rotación entre -20° y +20°, brillo entre -15 % y +15 %, exposición entre -10 % y +10 % y desenfoque hasta 2,5 px (Figura 2).

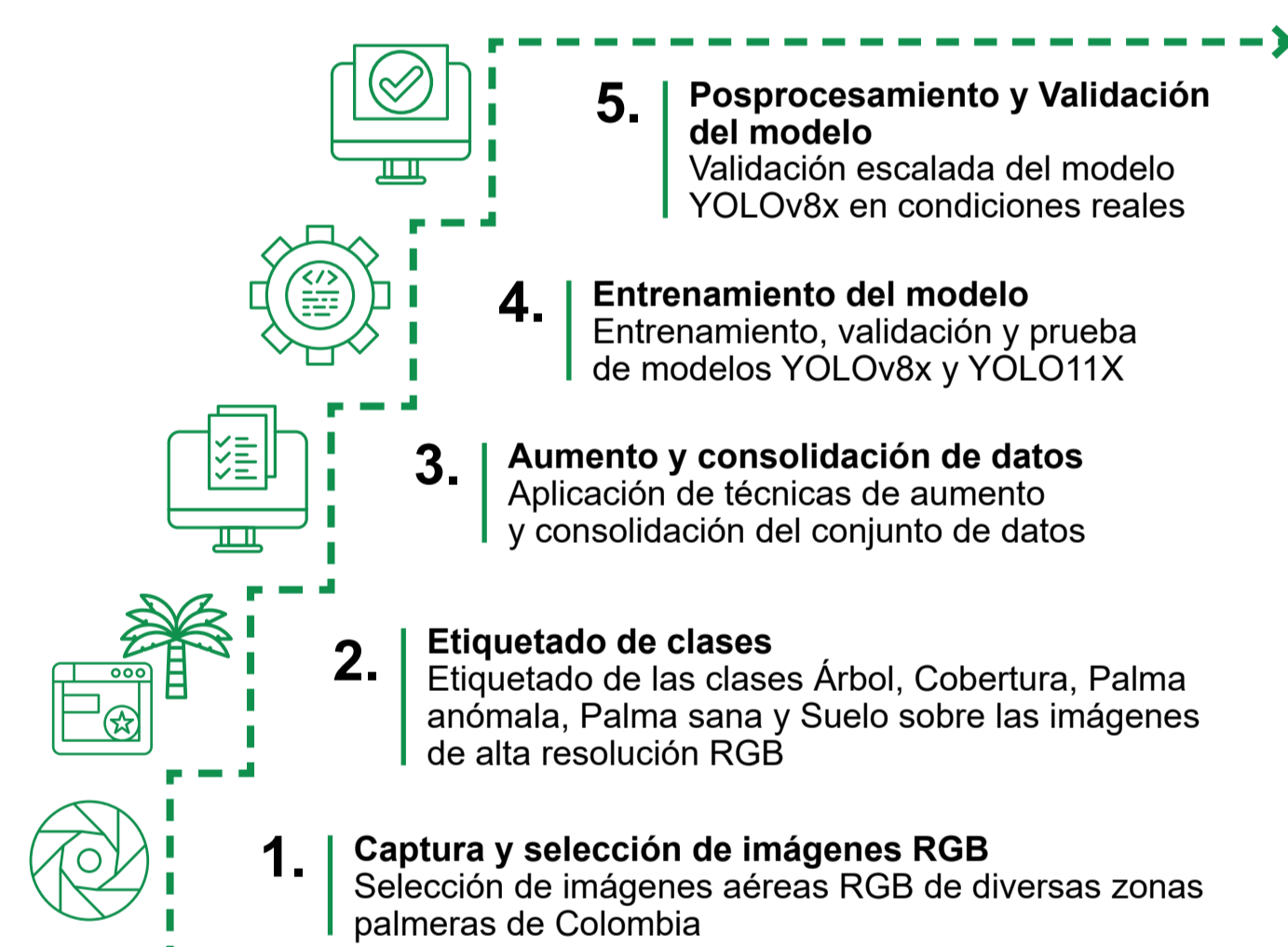


Figura 1. Desarrollo metodológico.

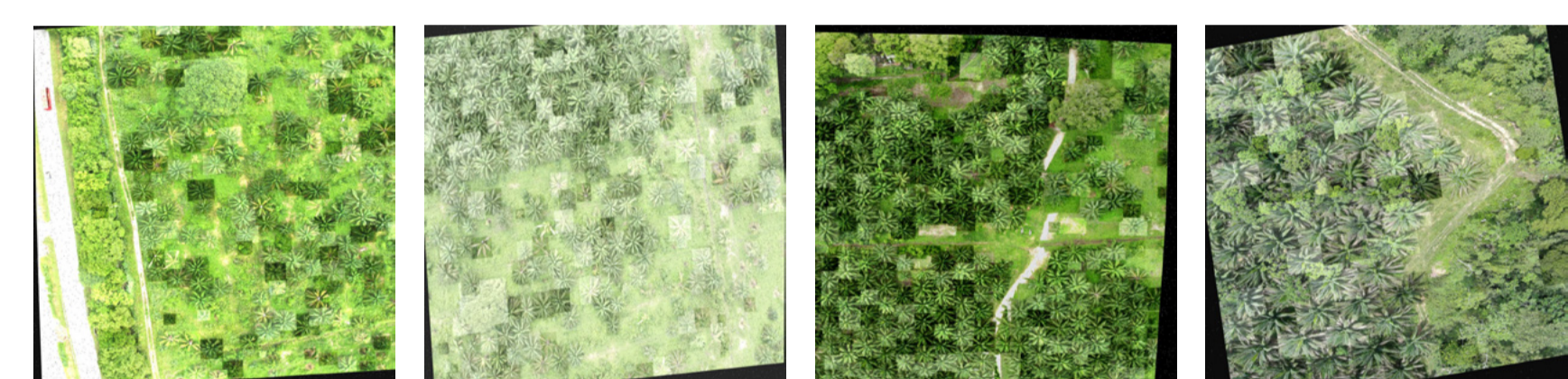


Figura 2. Colección de imágenes aumentadas del subconjunto de entrenamiento.

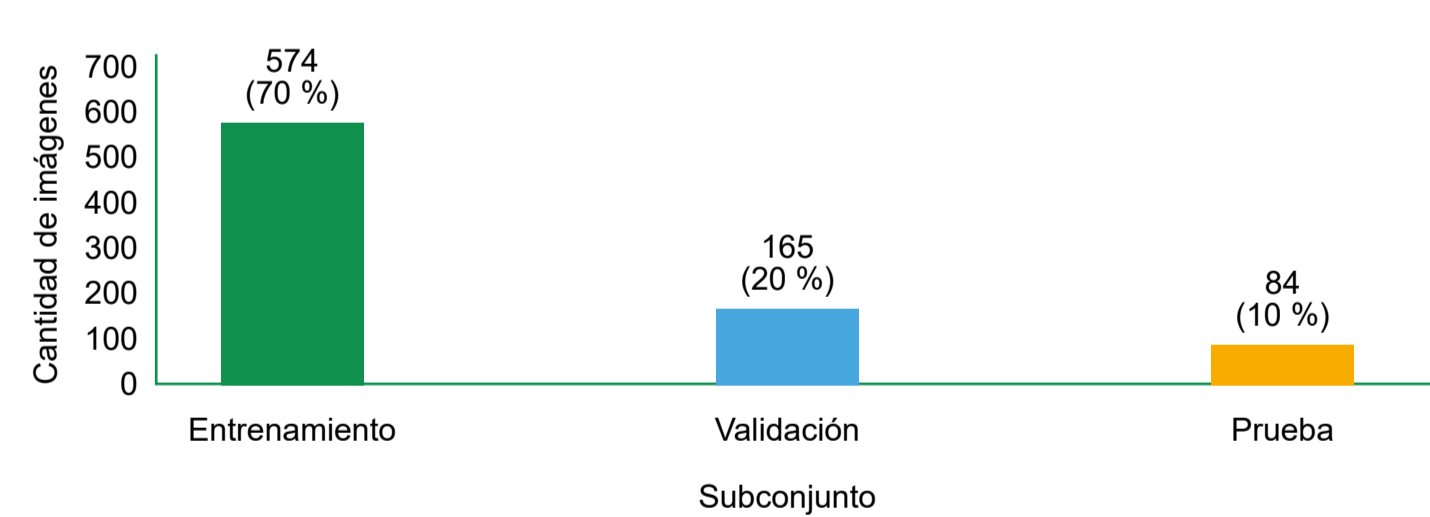


Figura 3. Cantidad de imágenes asignadas a las fases de entrenamiento, validación y prueba.

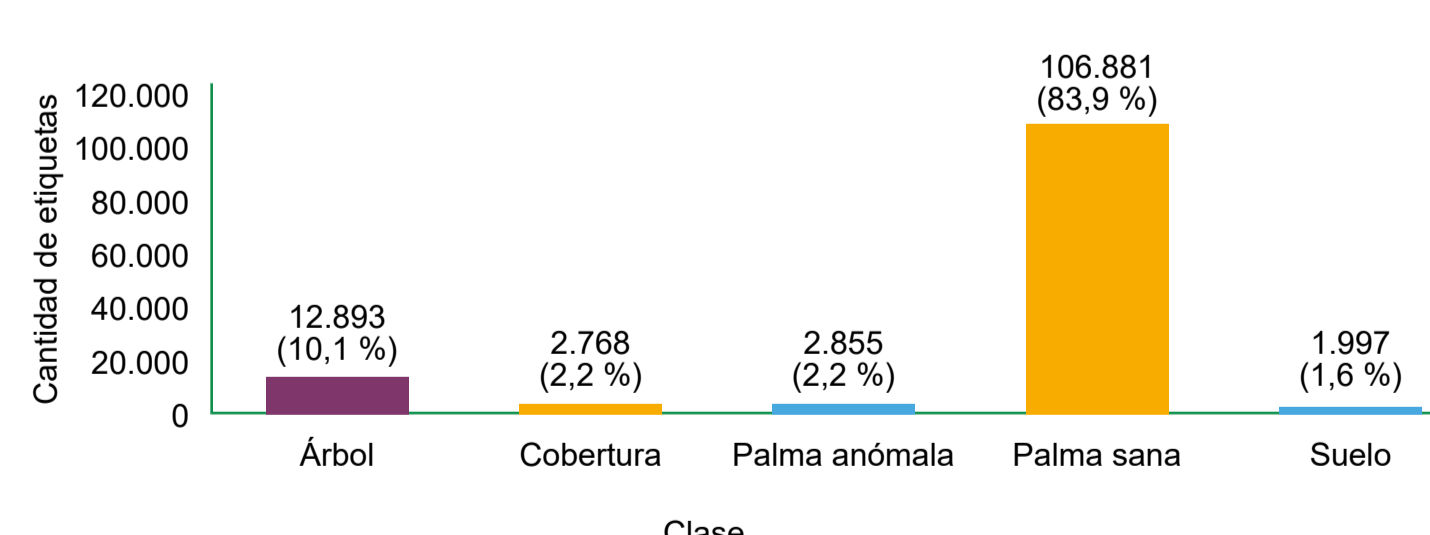


Figura 4. Cantidad de etiquetas del conjunto de datos.

Resultados

Entrenamiento y validación de modelos YOLO

Ambos modelos exhibieron la misma precisión al clasificar la clase Palma sana, es decir, del total de detecciones realizadas por cada modelo, el 93,90 % fueron clasificadas correctamente como verdaderos positivos (TP). Un valor tan elevado sugiere una baja tasa de falsos positivos (FP). La ligera superioridad en Recall del modelo YOLOv8x indica una menor propensión a omitir objetos reales. Adicionalmente, la consistencia de YOLOv8x en mAP:0,5 y su mejor indicador en mAP:0,5:0,95 sugieren que las detecciones de YOLOv8x no solo son clasificadas correctamente, sino que también presentan una mayor precisión en la localización de los objetos (Tabla 1).

Tabla 1. Métricas de los modelos YOLOv8x y YOLO11x para la clase Palma sana.

Modelo	Clase	Precisión	Recall	mAP:0,5	mAP:0,5:0,95
Yolov8x	Palma sana	93,9 %	95,4 %	96,9 %	76,8 %
Yolo11x	Palma sana	93,9 %	94,8 %	96,8 %	74,6 %

Posprocesamiento y validación del modelo YOLOv8x en condiciones reales

Se implementó una estrategia de posprocesamiento para refinar las predicciones del modelo YOLOv8x aplicado a ortoimágenes en condiciones reales. Las detecciones cuyos perímetros se situaban fuera de los límites definidos (inferior = $Q1 - 1,5 * IQR$, superior = $Q3 + 1,5 * IQR$) y que presentaron un umbral de confianza por debajo de 0,7 se consideraron detecciones erróneas. Para la validación se realizaron predicciones sobre 24 ortoimágenes correspondientes a 10 lotes de la Zona Oriental, 10 lotes de la Zona Central y 4 lotes de la Zona Norte. En este caso, la precisión mide, de todas las detecciones que hizo el modelo, qué proporción fueron realmente palmas (ecuación 1); el recall (sensibilidad o tasa de recuperación) mide de todas las palmas reales, qué proporción se detectó (ecuación 2).

Ecuación 1.

$$Precisión = \frac{TP}{TP + FP}$$

Ecuación 2.

$$Recall = \frac{TP}{TP + FN}$$

Donde, TP se refiere a los verdaderos positivos, FP a falsos positivos y FN a falsos negativos.

Los resultados revelan un rendimiento consistentemente alto del modelo a través de los diferentes lotes. La precisión promedio se situó en 96,1 % (mediana: 95,8 %) con una baja desviación estándar (1,2 %), lo que sugiere una estabilidad en la capacidad del modelo para generar predicciones correctas. Paralelamente, el recall promedio alcanzó el 98,2 % (mediana: 98,7 %) con una variabilidad similar (desviación estándar: 1,2 %), indicando una robusta habilidad para identificar la mayoría de los casos positivos (Figura 5).

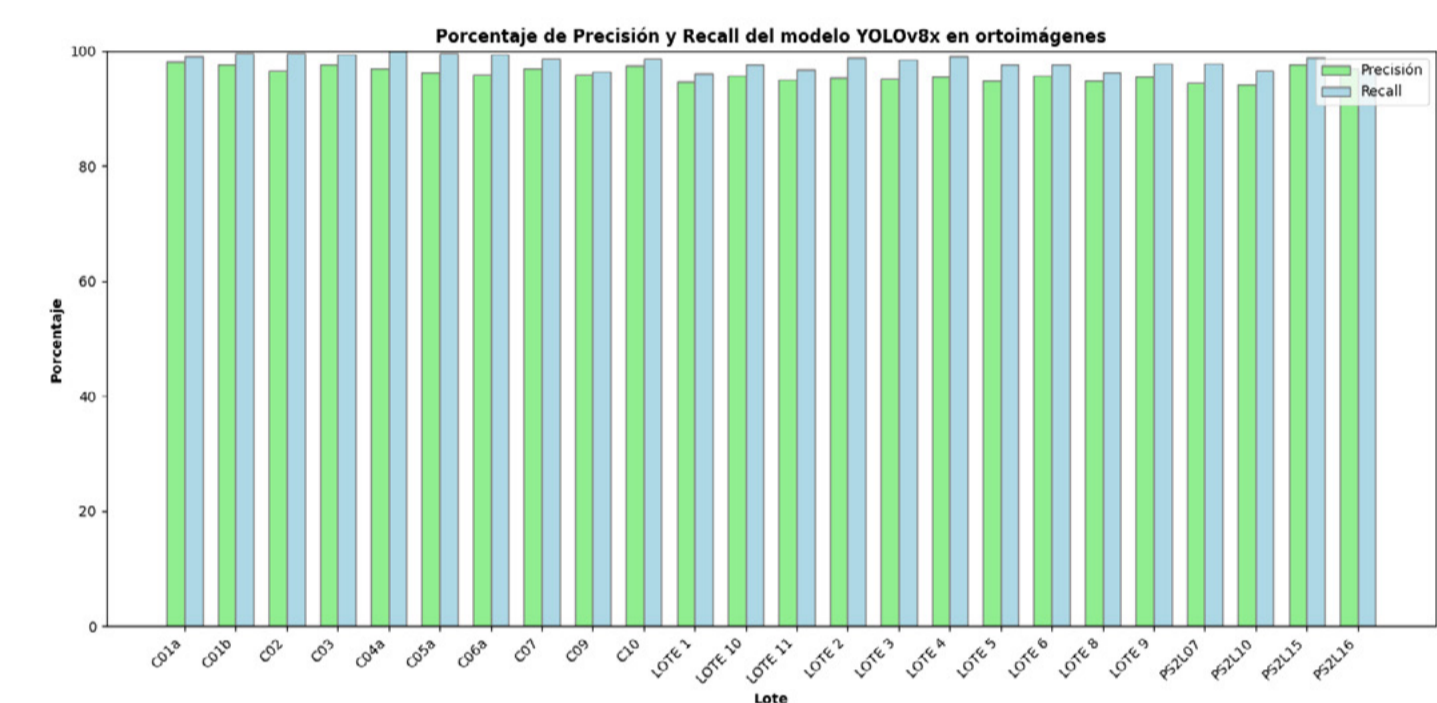


Figura 5. Porcentajes de precisión y recall del modelo YOLOv8x en ortoimágenes.

Inferencias modelo YOLOv8x

La Figura 6 presenta una colección de imágenes con dimensiones similares a las utilizadas durante el entrenamiento, en las que se muestran las detecciones de palma (cajas azules) realizadas mediante el modelo YOLOv8x. La Figura 7, por su parte, muestra las detecciones realizadas sobre una ortoimagen de un lote en la Zona Central, con dimensiones de 29115 x 10659 píxeles y una cobertura aproximada de 12 hectáreas.

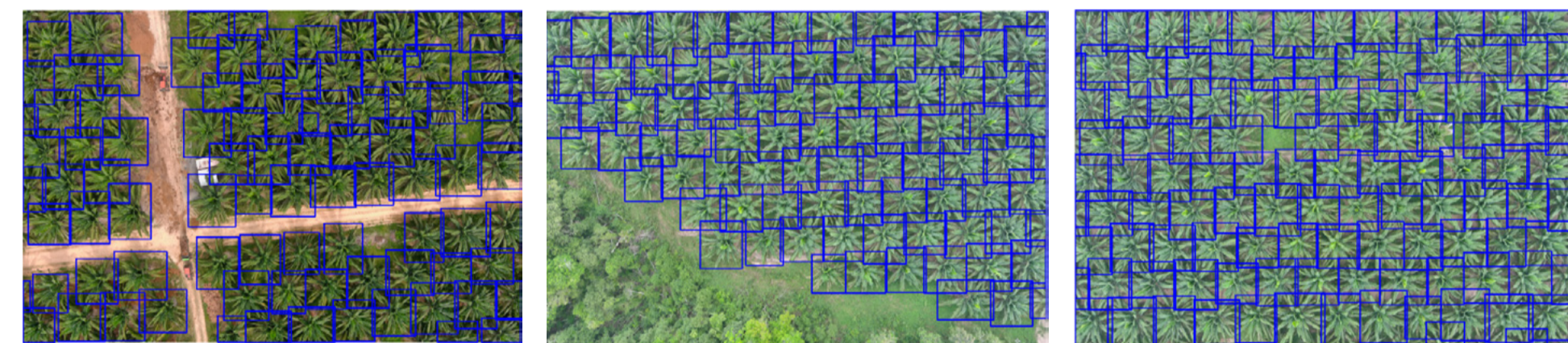


Figura 6. Ejemplos de inferencias con el modelo YOLOv8x en imágenes similares a las utilizadas en el entrenamiento.

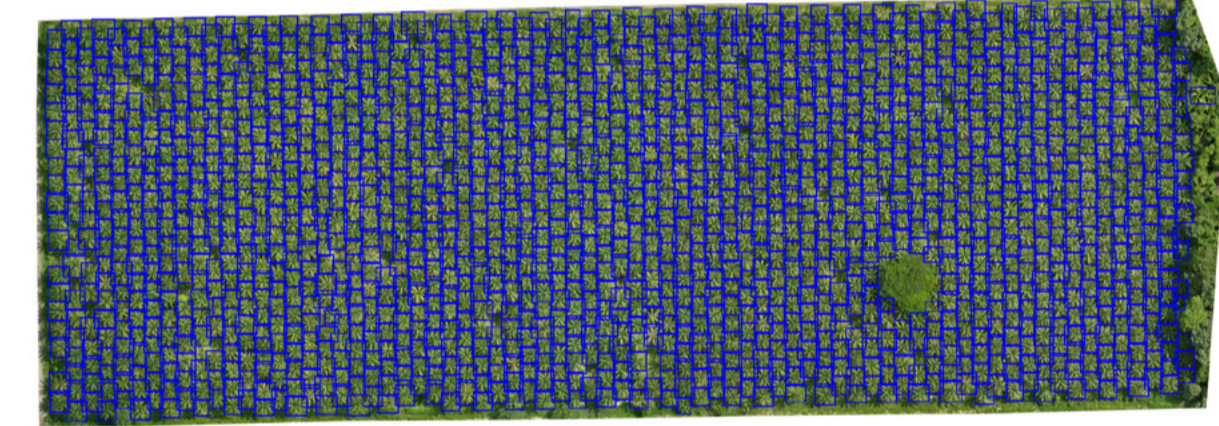


Figura 7. Ejemplos de inferencias con el modelo YOLOv8x en ortoimagen para un lote completo.

Conclusiones

La calidad del conjunto de datos es crucial para el rendimiento de los modelos de detección de objetos. La mejora en los datos conduce a incrementos significativos en precisión, recall y mAP. Resulta de gran importancia aumentar el número de clases para disminuir la confusión de los modelos con el fondo de las imágenes.

YOLOv8x se destacó como un modelo ligeramente superior en comparación con YOLO11x. Su mejor desempeño en las métricas de Recall y, especialmente, en mAP:0,5:0,95 indicaron una mayor capacidad para detectar la mayoría de los objetos presentes en las imágenes, con una localización más precisa.

La precisión del modelo YOLOv8x, tras el posprocesamiento en las ortoimágenes, demostró un rendimiento notablemente alto y consistente. La estrategia de filtrado implementada resultó efectiva para conservar una alta proporción de detecciones que coinciden con la realidad observada sobre el terreno, lo que sugiere una mejora en la fiabilidad de las predicciones del modelo en aplicaciones prácticas.

Referencias bibliográficas

- Bouguettaya, A., Zarzour, H., Kechida, A., Taberkit, A.M., 2022. Deep learning techniques to classify agricultural crops through UAV imagery: a review. *Neural Comput. & Applic.* 34, 9511–9536.
- Boursianis, A.D., Papadopoulou, M.S., Diamantoulakis, P., Liopa-Tsakalidi, A., Barouchas, P., Salahas, G., Karagiannidis, G., Wan, S., Goudos, S.K., 2022. Internet of things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: a comprehensive review. *Internet of Things* 18, 100187. Cenipalma. (2022).
- Fedepalma. (2022). Anuario Estadístico. Federación Nacional de Cultivadores de Palma de Aceite.
- Fedepalma. (2023). El sector palmero en cifras. Federación Nacional de Cultivadores de Palma de Aceite.
- Ultralytics. (s.f.). YOLO Models. Recuperado de: <https://docs.ultralytics.com/models/>
- Wakchaure, M., Patle, B., Mahindrakar, A., 2023. Application of AI techniques and robotics in agriculture: a review. *Artificial Intelligence in the Life Sciences* 3, 100057.

Agradecimientos

Los autores agradecen al Fondo de Fomento Palmero, administrado por Fedepalma, por la financiación de este estudio.

La información contenida en este póster es responsabilidad de los autores

CON EL APOYO DEL FONDO DE FOMENTO PALMERO