



Programa de Doctorado en Recursos y Tecnologías Agrarias,
Agroambientales y Alimentarias

Universidad Miguel Hernández de Elche

– 2025 –

Tesis doctoral

**Teledetección con vehículos aéreos no
tripulados y técnicas de aprendizaje
profundo para la gestión agrícola.
Aplicación a dos casos de estudio en la
provincia de Alicante**

Ana María Codes Alcaraz

Directora: Dra. Carmen Rocamora Osorio

Codirectora: Dra. Herminia Puerto Molina



La presente Tesis Doctoral, titulada “**Teledetección con vehículos aéreos no tripulados y técnicas de aprendizaje profundo para la gestión agrícola. Aplicación a dos casos de estudio en la provincia de Alicante**”, se presenta bajo la modalidad de **tesis por compendio** de las siguientes **publicaciones**:

Publicación 1

Codes-Alcaraz, A. M., Puerto, H., & Rocamora, C. (2024). Image Recognition for Floating Waste Monitoring in a Traditional Surface Irrigation System. *Water*, 16(18), 2680. <https://doi.org/10.3390/w16182680>

Publicación 2

Codes-Alcaraz, A. M., Furnitto, N., Sottosanti, G., Failla, S., Puerto, H., Rocamora-Osorio, C., Freire-García, P., & Ramírez-Cuesta, J. M. (2025). Automatic Grape Cluster Detection Combining YOLO Model and Remote Sensing Imagery. *Remote Sensing*, 17(2), 243. <https://doi.org/10.3390/rs17020243>

Otras publicaciones incluidas en la Tesis Doctoral:

Ana María Codes Alcaraz, Herminia Puerto Molina, Carmen Rocamora Osorio, Nicola Furnitto, Sabina Failla, Juan Miguel Ramírez-Cuesta. (2024). Deciphering the role of UAV flight height and sensor inclination in the prediction of table grape production through computer vision techniques. In Proceedings of the IEEE International Workshop on Metrology for Agriculture and Forestry, MetroAgriFor 2024, Padua, Italy, 29–31 October 2024; pp. 19-23.

ISBN: 979-8-3503-5543-7



La Dra. D^a. Carmen Rocamora Osorio, directora, y la Dra. D^a Herminia Puerto Molina, codirectora, de la tesis doctoral titulada **“Teledetección con vehículos aéreos no tripulados y técnicas de aprendizaje profundo para la gestión agrícola. Aplicación a dos casos de estudio en la provincia de Alicante”**,

INFORMAN:

Que **D^a. Ana-María Codes Alcaraz** ha realizado bajo nuestra supervisión el trabajo titulado **“Teledetección con vehículos aéreos no tripulados y técnicas de aprendizaje profundo para la gestión agrícola. Aplicación a dos casos de estudio en la provincia de Alicante”** conforme a los términos y condiciones definidos en su Plan de Investigación y de acuerdo al Código de Buenas Prácticas de la Universidad Miguel Hernández de Elche, cumpliendo los objetivos previstos de forma satisfactoria para su defensa pública como tesis doctoral.

Lo que firmamos para que surta los efectos oportunos, en Orihuela a 27 de enero de 2025

Directora de la tesis

Dra. D^a Carmen Rocamora Osorio

Codirectora de la tesis

Dra. D^a Herminia Puerto Molina

La **Dra. D^a. Juana Fernández López**, Catedrática de Universidad y Coordinadora del Programa de Doctorado en Recursos y Tecnologías Agrarias, Agroambientales y Alimentarias (ReTos-AAA) de la Universidad Miguel Hernández de Elche (UMH),

CERTIFICA:

Que la Tesis Doctoral titulada **“Teledetección con vehículos aéreos no tripulados y técnicas de aprendizaje profundo para la gestión agrícola. Aplicación a dos casos de estudio en la provincia de Alicante”** de la que es autora la graduada en en Ingeniería Agroalimentaria y Agroambiental, y Máster Universitario en Ingeniería Agronómica **D^a Ana - María Codes Alcaraz**, ha sido realizada bajo la dirección de la **Dra. D^a. Carmen Rocamora Osorio** y la codirección de la **Dra. D^a. Herminia Puerto Molina**, considero que la Tesis es conforme, en cuanto a forma y contenido, a los requerimientos del Programa de Doctorado ReTos-AAA, siendo por tanto apta para su exposición y defensa pública.

Y para que conste a los efectos oportunos firmo el presente certificado en Orihuela a 27 de enero de 2025.

Dra. D^{ña}. Juana Fernández López

Coordinadora del Programa Doctorado ReTos-AAA

*Coordinadora: Dra. Juana Fernández López
Dpto. Tecnología Agroalimentaria, Escuela Politécnica Superior de Orihuela, Ctra.
Beniel Km 3,2 Orihuela-03312, Alicante, España
j.fernandez@umh.es Tlf: +34 966749784*

Indicios de calidad de las publicaciones

Publicación 1

Codes-Alcaraz, A. M., Puerto, H., & Rocamora, C. (2024). Image Recognition for Floating Waste Monitoring in a Traditional Surface Irrigation System. *Water*, 16(18), 2680. <https://doi.org/10.3390/w16182680>

Publication date: September 20th 2024.

ISSN: 2073-4441

Rank: JCR - Q2 (Water Resources)

CiteScore - Q1 (Water Science and Technology)

Impact factor (2023): 3.0 - Impact factor (5 years): 3.3

Article access statistics (MDPI web): 975 views as of January 27th 2025.

Mention as Most Notable Article (September–November 2024) *Water*

(<https://www.mdpi.com/message/74066bea9cdbb95b80208f1c880aa9b1/32e01814584890b33b9ec2cdfeaa337a>)

Publicación 2

Codes-Alcaraz, A. M., Furnitto, N., Sottosanti, G., Failla, S., Puerto, H., Rocamora-Osorio, C., Freire-García, P., & Ramírez-Cuesta, J. M. (2025). Automatic Grape Cluster Detection Combining YOLO Model and Remote Sensing Imagery. *Remote Sensing*, 17(2), 243. <https://doi.org/10.3390/rs17020243>

Publication date: January 11th 2025.

ISSN: 2072-4292

Rank: JCR - Q1 (Geosciences, Multidisciplinary)

CiteScore - Q1 (General Earth and Planetary Sciences)

Impact factor (2023): 4.2 - Impact factor (5 years): 4.9

Article access statistics (MDPI web): 506 views as of January 27th 2025.

Agradecimientos

Llegar hasta aquí no habría sido posible sin la ayuda, el apoyo y la compañía de muchas personas que me han acompañado durante este camino.

Primero, quiero agradecer enormemente a mi directora, Carmen Rocamora, y a mi codirectora, Herminia Puerto, por confiar en mí desde el primer momento, guiarme con mucha paciencia y por su constante apoyo en tantas horas dedicadas a este proyecto. Gracias por creer en mí.

A mis enormes profesores, José Antonio Flores y Salvador Castillo, gracias por todo lo que me habéis enseñado, por vuestra dedicación y por inspirarme a dar siempre un paso más allá, por tener siempre tiempo para mí y ayudarme en todo lo que habéis podido. Habéis sido y seréis un pilar fundamental en mi vida.

A Nestor por haberme ayudado y guiado cuando acababa de empezar el doctorado fuiste un gran apoyo y el motivo de querer seguir adelante.

A Juan Miguel gracias por acogerme en Catania, por haber confiado en mí, por haberme ayudado tanto y por ser mi gran apoyo cuando estaba lejos de casa.

A mis compañeros y amigos de la universidad, Joaquín y Rafa, no sé qué habría hecho sin vosotros durante todos estos años. Gracias por las risas, las charlas interminables, las largas tardes y noches de estudio juntos y por estar ahí en las buenas y en las malas.

A mis amigas, Puri, María y Lucía gracias por vuestro apoyo incondicional siempre, por escucharme cuando lo necesitaba y por sacarme siempre una sonrisa. Teneros cerca ha sido un regalo.

Por último, pero lo más importante, quiero dar las gracias a mis padres y a mi hermano. Vuestro amor, vuestra confianza y vuestro apoyo han sido mi mayor motor durante todo este tiempo. Esto no habría sido posible sin vosotros.

Gracias a todos por haber hecho que este camino, con sus altibajos, haya valido tanto la pena.

Índice

Índice	i
Índice de tablas.....	iii
Índice de figuras	iii
Resumen	iv
Abstract	v
1 Introducción.....	1
1.1 Antecedentes	1
1.2 Impactos ambientales de los residuos plásticos en el sistema de riego tradicional de la Vega Baja del Segura.....	3
1.2.1 Sistema de riego tradicional en la Vega Baja del Segura.....	4
1.3 Gestión de precisión del viñedo en la zona del Vinalopó	6
1.3.1 Uva embolsada del Vinalopó	9
1.4 Contribución a los Objetivos de Desarrollo Sostenible	10
2 Objetivos	13
3 Material y Métodos	15
3.1 Áreas de estudio	15
3.2 Adquisición de imágenes.....	16
3.3 Configuración de vuelo para la detección de flotantes.....	16
3.4 Imágenes satelitales.....	18
3.5 Procesamiento y entrenamiento de modelos.....	20
3.5.1 Validación del modelo de detección de residuos flotantes y de racimos	23
4 Resultados y discusión.....	24
4.1 Resultados de la clasificación y cuantificación de residuos flotantes	24
4.2 Resultados de conteo de racimos embolsados	26
4.2.1 Configuración del vuelo	26
4.2.2 Comparación de las versiones del modelo YOLO.....	26
4.2.3 Comparación de la estimación de cosecha con la producción real.	28
5 Discusión.....	31
5.1.1 Modelo de detección de flotantes	31
5.1.2 Modelo de detección de racimos	32
6 Conclusion / Conclusión	35
7 Referencias	39
8 Anexos	48



Índice de tablas

Tabla 1. Identificación, posición y descripción de los puntos de muestreo de flotantes en el sistema hidráulico de la Vega Baja.	16
Tabla 2. Índices de vegetación y fórmulas de cálculo ¡Error! Marcador no definido.	
Tabla 3. Ecuaciones de correlación, coeficiente de determinación (R^2) y error cuadrático medio (RMSE) con las distintas configuraciones de vuelo.....	26

Índice de figuras

Figura 1. Acumulación de residuos flotantes en una acequia de Orihuela	5
Figura 2. Embolsado de la Uva embolsada del Vinalopó	10
Figura 3. Objetivos de Desarrollo Sostenible	11
Figura 4. Localización de las áreas de estudio en la provincia de Alicante	15
Figura 5. UAV Autel EVO pro 2.....	16
Figura 6. Puntos de acumulación de residuos.....	17
Figura 7. Imagen del UAV adquirida a 17 metros y con un ángulo de 45°.....	18
Figura 8. Marco conceptual de la metodología de detección de objetos	20
Figura 9. Matriz de confusión.....	22
Figura 10. Gráficas de precisión (a) mAP@0,5, (b) mAP@0,5 a 0,95, (c) precisión....	24
Figura 11. Curva de precisión-recuperación de los objetos detectados.	25
Figura 12. Detección automática de los residuos flotantes	25
Figura 13. Detección automática de los racimos de uva embolsados	27
Figura 14. Comparación entre el número de racimos embolsados estimados por YOLOv7x y las mediciones reales observadas en las imágenes del UAV. ...	28
Figura 15. Resultados de estimación de cosecha de los modelos.	29
Figura 16. Mapa de producción según el número de racimos detectados por el modelo	30

Resumen

Esta tesis doctoral explora dos aplicaciones innovadoras de la teledetección con vehículos aéreos no tripulados (*unmanned aerial vehicles*, UAV) y técnicas de aprendizaje profundo.

El primer estudio se centra en el sistema de riego tradicional de la Vega Baja del Segura (España), donde se acumulan grandes cantidades de residuos flotantes que afectan al funcionamiento del riego y generan graves problemas ambientales. Se creó un conjunto de datos con 477 imágenes de elementos plásticos flotantes en diferentes entornos y se utilizó para entrenar un algoritmo basado en YOLOv5s para detectar y clasificar residuos flotantes a partir de imágenes de UAV. El modelo entrenado logró una precisión media promedio (mAP) del 96,9% y la velocidad de detección fue de 81,7 ms. Se comparó el recuento automático de objetos flotantes con su recuento manual en fotografías aéreas de distintos puntos de acumulación. El recuento automático redujo significativamente el tiempo de análisis y ambos métodos mostraron una buena concordancia, confirmando que las botellas de agua fueron el tipo de residuo flotante más abundante (95%). Este enfoque proporciona una herramienta confiable para monitorizar la acumulación de residuos, ayudando a las autoridades de la cuenca en la evaluación de medidas correctivas para mitigar la contaminación plástica.

El segundo estudio se centra en la estimación del rendimiento en un viñedo situado en el área de producción de uva de mesa embolsadas del Vinalopó. Se entrenó un algoritmo basado en YOLOv7x con imágenes de UAV para estimar el número de racimos de uva. Este modelo predijo con precisión el número de racimos de uva, con un valor R^2 de 0,64 y un error cuadrático medio (RMSE) de 0,78 racimos vid^{-1} . Adicionalmente, se utilizaron datos de los satélites Sentinel-2 y Planetscope para calcular distintos índices de vegetación: Índice de Vegetación de Diferencia Normalizada (NDVI), Índice de Vegetación Ajustado al Suelo (SAVI), Índice de Vegetación Mejorado (EVI), Índice de Vegetación de Diferencia Normalizada en Verde (GNDVI), Índice de Vegetación Ajustado al Suelo Modificado (MSAVI), Índice Vegetal Triangular Modificado (MTVI), Índice de Diferencia Normalizada del Agua (NDWI), Índice de Vegetación Optimizado Ajustado al Suelo (OSAVI), e Índice de Relación Simple (SR). Los valores de los índices de vegetación mostraron una correlación limitada con los parámetros de rendimiento ($R^2 < 0,23$), ya que estos índices reflejan principalmente el vigor vegetativo en lugar de atributos específicos de rendimiento. Sin embargo, la integración de datos multispectrales satelitales, imágenes de UAV de alta resolución, y modelos de detección avanzados ofrece un marco sólido para mejorar la gestión de los viñedos y promover prácticas agrícolas más sostenibles y eficientes.

En conjunto, estos estudios destacan el potencial transformador de los vehículos aéreos no tripulados y la inteligencia artificial para abordar desafíos relevantes en la gestión ambiental y agrícola.

Palabras clave: YOLO; plástico; residuos flotantes; acumulación de residuos; ríos; detección de objetos; Sentinel-2; Planetscope; detección de racimos embolsados; estimación de cosecha.

Abstract

This PhD thesis explores two innovative applications of remote sensing with unmanned aerial vehicles (UAVs) and deep learning techniques.

The first study focuses on the traditional irrigation system of the Vega Baja del Segura (Spain), where large amounts of floating waste accumulate affecting irrigation performance and generating serious environmental problems. A dataset was created with 477 images of floating plastic items in different environments and was used for training an algorithm based on YOLOv5s to detect and classify floating debris from UAV images. The trained model achieved a mean average precision (mAP) of 96.9% and the detection speed was 81.7 ms. Overhead photographs were taken with an unmanned aerial vehicle at strategic points of the river and channels, and its automatic count of floating objects was compared with their manual count. The automatic count significantly reduced analysis time and both methods showed good agreement, confirming that water bottles were the most abundant (95%) type of floating waste. This approach provides a reliable tool to monitor debris accumulation, supporting basin authorities in the evaluation of corrective measures to mitigate plastic pollution.

The second study focuses on yield estimation in a vineyard located in the Vinalopó bagged table grape production area. A YOLOv7x-based algorithm was trained using UAV imagery to estimate the number of grape clusters. This model accurately predicted the number of grape clusters, with an R^2 value of 0.64 and a root mean square error (RMSE) of 0.78 bunches vine⁻¹. Additionally, data from Sentinel-2 and PlanteScope satellites were used to calculate different vegetation indices: Normalized Difference Vegetation Index (NDVI), Soil Adjusted Vegetation Index (SAVI), Enhanced Vegetation Index (EVI), Green Normalized Difference Vegetation Index (GNDVI), Modified Soil Adjusted Vegetation Index (MSAVI), Modified Triangular Vegetation Index (MTVI), Normalized Difference Water Index (NDWI), Optimized Soil Adjusted Vegetation Index (OSAVI), and Simple Ratio Index (SR). The vegetation indices values showed limited correlation with yield parameters ($R^2 < 0.23$), as these indices mainly reflect vegetative vigor rather than specific yield attributes. However, the integration of multispectral satellite data, high-resolution UAV imagery, and advanced sensing models offers a robust framework to improve vineyard management and promote more sustainable and efficient agricultural practices.

Collectively, these studies highlight the transformative potential of UAVs and AI to address significant challenges in environmental and agricultural management

Keywords: YOLO; plastic; waste; waste accumulation; rivers; object detection; sentinel-2; planetscope; unmanned aerial vehicle; bagged cluster detection; yield estimation

1 Introducción

1.1 Antecedentes

La agricultura alicantina enfrenta grandes desafíos como son su escasez estructural de agua, la salinidad del agua de riego en ciertas zonas, su dependencia de trasvases de otras cuencas y una marcada presión urbanística. A pesar de estas dificultades, la agricultura de la provincia se apoya con fuerza en el regadío (Sánchez-Rubio, 2010) y en la calidad diferenciada de sus productos amparados por Denominaciones de Origen Protegidas.

La cantidad de vestigios de obras hidráulicas de todas las épocas hallados en la provincia enfatiza la importancia del regadío en su agricultura, caracterizada por un clima árido en su mayor parte y semiárido en la franja costera, y unas precipitaciones muy irregulares (Puerto y Rocamora, 2019). En las últimas décadas estos regadíos han evolucionado dando lugar a instalaciones modernizadas a lo largo de toda la provincia, que conviven con una amplia zona dedicada al cultivo hortícola donde se mantiene en uso el sistema de riego tradicional de la Vega Baja del Segura. Este regadío tradicional cuenta con un ancestral conjunto de obras hidráulicas mediante las cuales se deriva agua del río para riego. Los azudes son las obras de toma de agua del río hacia las acequias que distribuyen el agua hacia las tierras de regadío a través de una extensa, compleja y jerarquizada red de canales de riego. Posteriormente, los azarbes recogen el agua de drenaje que, a su vez, se utiliza para regar cuando se alcanza una altura suficiente sobre los campos aguas abajo, aunque la escasa pendiente de la llanura final del río Segura, inferior al 1 por 10.000, dificulta en gran medida el drenaje natural del terreno (Trapote Jaume et al., 2015).

Gracias a avances técnicos como el riego localizado y la modernización de infraestructuras, la agricultura de regadío se ha vuelto más eficiente, rentable y competitiva, especialmente en cultivos como cítricos, hortalizas y la uva de mesa embolsada del Vinalopó, un producto emblemático con Denominación de Origen Protegida (DOP) que combina tradición y tecnología para competir en mercados nacionales e internacionales. Además, las condiciones térmicas de la provincia permiten cosechas en momentos estratégicos, como el invierno, lo que le otorga una ventaja competitiva en el mercado global (Sánchez-Rubio, 2010).

A escala mundial, la agricultura ha experimentado una profunda transformación en las últimas décadas, impulsada por la incorporación de tecnologías avanzadas, especialmente dentro del ámbito de la Agricultura de Precisión (AP) (Bhakta et al., 2019). Este enfoque innovador ha surgido como respuesta a la creciente necesidad de implementar prácticas agrícolas más sostenibles y eficientes frente a los desafíos económicos y ambientales actuales (Kendall et al., 2022). La AP tiene como objetivo principal optimizar la gestión de los cultivos mediante el uso eficiente de los recursos agrícolas y la minimización de impactos negativos en el medio ambiente (Soumitra, 2024). Para ello, emplea herramientas tecnológicas como la teledetección, los Sistemas de Información Geográfica (SIG), el análisis de datos masivos (Big Data), la Inteligencia Artificial (IA) y el Internet de las Cosas (IoT). Estas tecnologías permiten recopilar información detallada sobre las condiciones del suelo y los cultivos, proporcionando a los agricultores datos precisos para tomar decisiones fundamentadas y en tiempo real (Liu et al., 2021; Leroux et al., 2018; Guo et al., 2018; Wolfert et al., 2017).

En los últimos años, los vehículos aéreos no tripulados (*unmanned aerial vehicles*, UAV) han emergido como herramientas clave en la agricultura de precisión. Los avances en áreas como la robótica, los sistemas de control y la visión artificial han permitido que los UAV sean más accesibles y eficaces, convirtiéndose en una opción destacada para diversas aplicaciones agrícolas (Krishna, 2018; Marinello et al., 2016; Veroustraete 2015; Daponte et al., 2019; Gutiérrez et al 2021; Apolo-Apolo et al., 2020; Ramírez-Cuesta et al., 2023). Estos dispositivos destacan por su versatilidad y capacidad para realizar tareas de teledetección de forma rápida y precisa, incluso bajo condiciones climáticas adversas. Además, ofrecen resoluciones espaciales y temporales superiores a las que pueden proporcionar la mayoría de los sistemas satelitales (Manfreda et al., 2018), lo que facilita la identificación de detalles y características que, de otro modo, pasarían desapercibidos en las imágenes de satélite (Payares et al., 2023). No obstante, el rendimiento de estos sistemas depende de numerosos factores, como el peso de la aeronave, su capacidad de carga útil, dimensiones, características de la batería, autonomía de vuelo, y condiciones ambientales. Además, el costo de implementación de los UAV en la agricultura incluye no solo la adquisición de la plataforma y equipos complementarios (como escáneres láser y cámaras multispectrales o hiperespectrales), sino también la formación y contratación de personal cualificado para el manejo del equipo y el procesamiento de los datos obtenidos (Toscano et al., 2024).

En un contexto de creciente presión sobre los recursos naturales y cambios en los patrones climáticos, la provincia de Alicante se enfrenta a desafíos significativos en los sectores ambiental y agrícola. Estos desafíos son particularmente visibles en áreas como la Vega Baja del Segura, donde el impacto de los residuos plásticos sobre el sistema de riego tradicional supone importantes retos a nivel ecológico, y el valle del Vinalopó, donde la uva de mesa embolsada representa el paradigma de una producción de alto valor añadido cuya sostenibilidad económica, ambiental y social depende de la optimización de la gestión de recursos. Esta zona, conocida por su rica tradición agrícola, ofrece un escenario único para estudiar y abordar estos problemas mediante la implementación de tecnologías emergentes como la teledetección y la inteligencia artificial (IA).

A lo largo de esta introducción se pondrá de manifiesto la creciente importancia de la inteligencia artificial (IA) en la detección de objetos, ya sean residuos plásticos o racimos de uva. Por ello se ha considerado adecuado abordar en este punto conceptos como las redes neuronales, los algoritmos de detección y el aprendizaje profundo.

La IA proporciona herramientas, como las redes neuronales, que permiten crear modelos capaces de identificar y mejorar la detección de objetos a partir de imágenes. Las redes neuronales son modelos computacionales inspirados en el funcionamiento del cerebro humano, diseñados para procesar datos y aprender patrones a través de múltiples capas de nodos interconectados (Nielsen, 2015). Estas estructuras utilizan algoritmos de aprendizaje automático, como el aprendizaje supervisado o no supervisado, para resolver problemas complejos en áreas como el reconocimiento de imágenes o la predicción de series temporales (LeCun et al., 2015). Cada nodo simula el comportamiento de una neurona, activándose en función de los estímulos recibidos y ajustando sus pesos durante el entrenamiento para optimizar el rendimiento del modelo (Nielsen, 2015). Un tipo particular de red neuronal, las redes neuronales convolucionales (CNN), son especialmente eficaces en tareas de visión por computadora (Nielsen, 2015). Utilizan filtros para detectar patrones en las imágenes, como bordes, texturas y formas y extraen automáticamente las características más relevantes de los datos.

Las redes neuronales, como potentes herramientas para el aprendizaje de patrones, han dado lugar a una variedad de algoritmos de detección de objetos. Entre ellos, destacan los algoritmos de una y dos etapas.

Los algoritmos de una etapa realizan la detección de objetos en una sola pasada sobre la imagen, identificando y clasificando objetos simultáneamente. Esto los hace extremadamente rápidos, pero pueden tener dificultades en la detección de objetos pequeños o en imágenes densas. Estos algoritmos suelen utilizar una red neuronal para predecir directamente las posiciones y categorías de todos los cuadros delimitadores posibles en cada imagen (Wang et al., 2024). Este enfoque incluye la serie You Only Look Once (YOLO), que se beneficia de la velocidad y el rendimiento en tiempo real (Redmon et al., 2016). YOLO es un algoritmo de detección de objetos muy popular que ha evolucionado a través de diversas versiones. Cada nueva versión de YOLO busca mejorar el rendimiento en términos de velocidad y precisión. Desde su versión inicial, YOLO ha experimentado múltiples iteraciones, como YOLOv2, YOLOv3, YOLOv4 y YOLOv5, entre otras. Estas versiones incorporan mejoras arquitectónicas, técnicas de entrenamiento y estrategias de aumento de datos para lograr mejores resultados en una amplia gama de tareas de detección de objetos.

Por otro lado, los algoritmos de dos etapas, como Faster R-CNN, siguen un enfoque más gradual. Primero generan propuestas de regiones donde es probable que haya un objeto y luego clasifican esas regiones. Este método suele ofrecer una mayor precisión, pero a costa de una mayor complejidad computacional y menor velocidad.

1.2 Impactos ambientales de los residuos plásticos en el sistema de riego tradicional de la Vega Baja del Segura

La presencia de desechos plásticos en agua dulce es un problema ambiental cada vez mayor. Este fenómeno implica un riesgo ecológico ya que afecta a la fauna y la flora acuática (Andrade et al., 2019; Inobeme et al., 2024). Además, los desechos de gran tamaño pueden obstruir el flujo natural del agua aumentando el riesgo de inundaciones en áreas urbanas, afectando la infraestructura local y causando graves pérdidas económicas. Los desechos plásticos se acumulan en las riberas de los ríos y los azudes, las compuertas y los canales de riego actúan como zonas de acumulación de macroplásticos. Un problema importante es la contaminación estética y la pérdida de valor recreativo de estos cuerpos de agua. Los desechos visibles bloquean la percepción de los ríos y lagos, reduciendo su atractivo para actividades recreativas como la natación, la pesca y el turismo, lo que puede tener impactos económicos adversos en las comunidades locales (Bellasi et al., 2020; Van Emmerik and Schwarz, 2019; Van Emmerik et al., 2022).

En la actualidad existe gran preocupación por la acumulación de residuos plásticos en mares y océanos. Varios estudios han establecido que la principal fuente de contaminación plástica marina es de origen terrestre, debido a la basura, el uso de bolsas de plástico y la eliminación de desechos sólidos (Derraik, 2022), y los ríos son el vehículo que conecta las superficies terrestres con el mar o los océanos al arrastrar los desechos que se acumulan en ellos con el agua que transportan hacia el mar. Dado que existen evidencias de que no todo el plástico que entra en un río se libera al océano (Van Emmerik et al., 2022), existe un creciente interés en caracterizar el flujo de residuos macroplásticos en los ríos para identificar y diseñar medidas para reducir la contaminación plástica global (Hurley et al., 2023).

Los sistemas de riego superficial en llanuras fluviales donde perviven sistemas de riego tradicional con canales abiertos extienden la red de flujos de agua, que actúan como sumideros de residuos, aumentando la probabilidad de que los desechos terrestres lleguen a los océanos. En estos sistemas de riego superficial, se ha estudiado el efecto de los restos leñosos flotantes en canales y embalses porque los desechos bloquean el flujo de agua y aumentan la altura de referencia en las estructuras de control del riego (Vaughn et al., 2021), y se han utilizado sistemas automatizados de eliminación de basura para proteger compuertas automáticas en proyectos de automatización de canales (Stringam et al., 2016), pero apenas hay menciones a la aparición de residuos plásticos flotantes en canales de riego en la literatura científica.

1.2.1 Sistema de riego tradicional en la Vega Baja del Segura

La comarca de la Vega Baja del Segura se encuentra situada al sur de la provincia de Alicante y abarca una extensión de unos 957 km². Linda al norte con las comarcas del Medio y del Bajo Vinalopó, al este con el mar Mediterráneo y al suroeste con la Región de Murcia. La comarca agrupa a 27 municipios (GVA, 2015), de los cuales 22 están dentro del perímetro del regadío tradicional del río Segura.

La Vega Baja se caracteriza por un complejo sistema de riego tradicional del que ya existe constancia desde el siglo XI, que se caracteriza por una red de acequias que derivan agua del río Segura y una red paralela de canales de drenaje o azarbes que recogen las aguas sobrantes de regadío y las reutilizan en parcelas de cota inferior. A lo largo de la Vega Baja existen ocho azudes o presas en el río que elevan el nivel del agua permitiendo que ésta se desvíe por las acequias (Melgarejo et al., 2013).

La gestión del reparto de agua en el sistema de regadío tradicional corre a cargo de Juzgados Privativos de Aguas, Sindicatos de Riego y Comunidades de Regantes, que tienen entre sus responsabilidades el mantenimiento de la red de acueductos para el correcto funcionamiento de los mismos (Melgarejo et al., 2013).

La Vega Baja enfrenta problemas crecientes debido a la acumulación de residuos flotantes que se produce en ciertos puntos del cauce del río y de los canales, causando problemas en la distribución del riego en acequias y canales de drenaje (Rocamora et al., 2021). Estos desechos, no solo afectan la distribución del agua, sino que también representan un riesgo para la sostenibilidad ambiental y económica de la región. Además, estos residuos impactan negativamente en la calidad del agua y en la funcionalidad de las infraestructuras, incrementando los costos operativos para los agricultores y generando conflictos con las autoridades de cuenca. Los residuos flotantes que se encuentran en el río y en la red de riego tienen un origen variado, predominando los residuos vegetales del propio río, como las cañas, así como los residuos de actividades domésticas, agrícolas e industriales, entre los que predominan los residuos plásticos (Rocamora et al., 2021).



Figura 1. Acumulación de residuos flotantes en una acequia de Orihuela. Fuente: Elaboración propia.

Las autoridades locales han planteado la implementación de campañas de sensibilización dirigidas a diferentes grupos de la población como una estrategia para reducir la llegada de residuos plásticos al río Segura. Sin embargo, para maximizar la efectividad de estas campañas, resulta fundamental identificar con precisión el sector de la población responsable de generar estos residuos. Analizar y determinar el tipo de plásticos presentes en los canales de riego podría ofrecer pistas sobre su origen: si estos desechos provienen mayoritariamente de actividades agrícolas, las campañas deberían enfocarse en los agricultores, mientras que, si tienen un origen más general, deberían orientarse hacia la población en su conjunto.

Se han propuesto diferentes enfoques para cuantificar y caracterizar residuos flotantes. Entre ellos destacan el desarrollo de software diseñado para analizar grandes concentraciones de desechos en el océano Pacífico (De Vries et al., 2021), metodologías basadas en aprendizaje profundo para identificar macroplásticos flotantes en imágenes capturadas en ríos de Yakarta (Van Lieshout et al., 2020), y el uso de algoritmos como APLASTIC-Q, que emplea redes neuronales convolucionales para clasificar y cuantificar diversos tipos de residuos, incluyendo botellas, poliestireno, cartones, zapatos y bolsas, en playas y sistemas fluviales de Camboya (Wolf et al., 2020).

En investigaciones más recientes, la detección de plásticos en cuerpos de agua ha evolucionado gracias al uso de modelos avanzados como YOLO. Por ejemplo, se han comparado diferentes configuraciones de YOLOv5 para identificar botellas y cartones en imágenes capturadas mediante vehículos de superficie no tripulados (Junzhe et al., 2023; Li et al., 2023). Los avances en aprendizaje profundo han impulsado el desarrollo de nuevas técnicas de detección de objetos, destacándose las redes neuronales profundas como herramientas clave. Los enfoques tradicionales de dos etapas, como las redes basadas en regiones (R-CNN) (He et al., 2015; Ren et al., 2016; Cai and Vasconcelos, 2018), han sido complementados y, en muchos casos, superados por métodos de una sola etapa como los algoritmos YOLO, que ofrecen una

detección más rápida y eficiente de desechos flotantes (Junzhe et al., 2023; Li et al., 2023; Tian et al., 2023). La introducción de las R-CNN por Girshick et al. (2014) marcó un hito en el desarrollo de métodos de detección de objetos basados en aprendizaje profundo. Aunque métodos como R-CNN han demostrado una alta precisión, se perciben como lentos (Inbar et al., 2023). A diferencia de los métodos de dos etapas, YOLO no necesita generar regiones candidatas; en cambio, divide la imagen en cuadrículas, cada una con cuadros delimitadores para detectar objetos en un solo paso, agilizando así el proceso (Junzhe et al., 2023). Además, en un estudio sobre detección de basura en el agua, Faster R-CNN mostró un rendimiento insatisfactorio debido a su pérdida de precisión en comparación con YOLOv5 (Junzhe et al., 2023).

A pesar de la creciente cantidad de trabajos sobre detección de imágenes para la cuantificación de basura en el medio ambiente, la mayoría de los artículos detectan residuos aislados (Junzhe et al., 2023), o residuos flotantes mediante navegación de superficie (Lin tal., 2021), existe menos evidencia sobre la detección de residuos acumulados en barreras y estructuras de contención de redes de riego.

La implementación de tecnologías avanzadas como los vehículos aéreos no tripulados (UAV) y los algoritmos de detección basados en inteligencia artificial ha surgido como una solución prometedora para abordar estos desafíos. En la Vega Baja, estas herramientas pueden proporcionar una forma eficiente y precisa de identificar, cuantificar y caracterizar los residuos flotantes, optimizando su gestión y reduciendo el impacto ambiental. Además, estas tecnologías permiten obtener una visión más completa y dinámica de los problemas que enfrenta la comarca, facilitando la toma de decisiones informadas.

1.3 Gestión de precisión del viñedo en la zona del Vinalopó

El uso de la inteligencia artificial en el sector vitícola todavía se encuentra en una fase incipiente, aunque presenta un gran potencial para transformar numerosos procesos clave en este sector. Su implementación permite optimizar significativamente actividades que tradicionalmente han dependido de métodos manuales. En los últimos años, el mundo del vino ha tenido que superar grandes retos, como la falta de trabajadores especializados y el aumento de los costes laborales. Estos problemas han dificultado mantener la productividad, mejorar la calidad del vino y recoger la uva a tiempo (Pérez-Zavala et al., 2018). Además, las tareas manuales suelen ser laboriosas, consumen mucho tiempo y están influenciadas por decisiones subjetivas de los trabajadores. Estas limitaciones han fomentado el desarrollo de tecnologías avanzadas, como sensores de última generación y el uso de UAV en combinación con inteligencia artificial. Las tecnologías digitales se están implementando en viticultura en aplicaciones tan diversas como evaluación de las propiedades del suelo, análisis del crecimiento, estado nutricional, detección de plagas y enfermedades, evaluación del estado hídrico de la vid (Tardáguila et al. 2021; Ramírez et al., 2023).

En los últimos años, el uso de sensores avanzados y la incorporación de tecnologías innovadoras en la viticultura de precisión han generado un notable incremento en la cantidad de datos obtenidos por hectárea, proporcionando a los viticultores una base de información mucho más rica y detallada (Sozzi et al., 2021). Por ejemplo, Gutiérrez et al. (2018) emplearon imágenes térmicas para evaluar el estado hídrico de la vid, mientras que en un estudio posterior, Gutiérrez et al. (2021) utilizaron UAV para detectar condiciones tempranas de estrés hídrico y correlacionarlas con el

índice NDVI. Asimismo, Ramírez-Cuesta et al. (2023) exploraron el uso de drones equipados con cámaras multiespectrales y térmicas para evaluar la eficiencia del uso del agua en cultivos leñosos, incluso en escenarios donde la vegetación no cubre completamente el suelo. Los métodos de obtención de datos empleados habitualmente en este campo pueden agruparse en tres categorías principales: imágenes satelitales (di Gennaro et al., 2019), imágenes capturadas mediante plataformas aéreas, como los UAV (Sepúlveda et al. 2016), e imágenes recogidas desde plataformas terrestres (González-Caballero, 2012; Diago et al., 2016; Palacios et al, 2020; Wang et al., 2024; Íñiguez et al., 2024). Debido a que los viñedos suelen ocupar áreas relativamente pequeñas, las plataformas aéreas y terrestres tienden a ser las más relevantes para evaluar diversos parámetros dentro de la viticultura de precisión (Squeri et al., 2021). A pesar de esto, las imágenes satelitales continúan ofreciendo un valor significativo, especialmente para el análisis a gran escala (Cogato et al., 2020). Para evaluar y comparar estas fuentes de datos en la gestión de viñedos es fundamental analizar factores clave como la resolución espacial, el número de bandas espectrales disponibles y los costes asociados. Las imágenes satelitales multiespectrales gratuitas, como las ofrecidas por el Programa Copernicus de la Unión Europea a través de los satélites Sentinel-2A y 2B, ofrecen una resolución espacial de 10 metros por píxel y una mayor variedad de bandas espectrales en comparación con las imágenes típicas obtenidas por UAV. También existen otras plataformas comerciales que ofrecen imágenes satelitales como PlanetScope SuperDove, una nueva generación de satélites CubeSat DOVE capaces de captar imágenes diarias de toda la superficie terrestre con una resolución de píxel de 3 metros. Sin embargo, el costo y la frecuencia de adquisición son aspectos críticos. Las imágenes capturadas por UAV tienen un costo por vuelo, lo que puede limitar la cantidad de datos obtenidos durante una temporada de cultivo. Por otro lado, los datos satelitales de Copernicus están disponibles de forma gratuita cada cinco días, lo que ofrece una ventaja significativa en términos de resolución temporal. En general, mientras que los UAV destacan por su alta resolución espacial y su capacidad para capturar detalles finos, las imágenes satelitales gratuitas sobresalen por su accesibilidad y frecuencia de actualización, ofreciendo así enfoques complementarios para la gestión de viñedos.

Una aplicación interesante de las tecnologías digitales es la estimación de la producción. Se trata de una información importante para la gestión adecuada de una explotación, pero implica un trabajo laborioso que requiere tiempo y mano de obra con experiencia o capacitación específica (Payne et al., 2013). Ballesteros et al. (2015) estimaron la producción de vid mediante datos como el índice de área foliar (LAI), la cobertura verde del dosel y el volumen del dosel, tomados desde sensores remotos montados en UAV. Muchos estudios se han centrado en la estimación de la producción mediante la detección de frutos en los árboles, incluidas manzanas, cerezas, uvas, aguacates, mangos, maracuyá y cítricos, entre otros (Robson et al., 2017; Shi et al., 2020; Apolo-Apolo et al., 2020; Wang et al., 2021; Íñiguez et al., 2024). Aunque la detección de frutos basada en el aprendizaje profundo ha tenido éxito, aún pueden ocurrir detecciones y localizaciones inexactas debido a la complejidad de los entornos agrícolas, las fluctuaciones de la iluminación y la imprevisibilidad, como el viento o las vibraciones mecánicas durante el proceso de cosecha real (Xiong et al., 2018; Jiménez et al., 2000).

La comparación entre distintos algoritmos de detección de imágenes se revela como una herramienta fundamental para obtener mejores resultados. Varios estudios

han comparado diferentes versiones de YOLO, como Santos et al. (2020), que compararon las redes YOLOv2, YOLOv3 y Mask R-CNN aplicadas a la detección de uvas, o Liu et al. (2020), que emplearon YOLOv3 para la detección de tomates utilizando cuadros delimitadores circulares en lugar de rectangulares tradicionales. Además, basándose en YOLOv5, Dong et al. (2004) realizaron la detección de lichis y calcularon las coordenadas de píxeles de los puntos de cosecha. Actualmente, la versión en la se han obtenido mejores resultados en detección de frutos es YOLO v7, que se ha aplicado para detectar la madurez de la uva en una variedad de uva blanca, aunque ya existen versiones posteriores como YOLO v10 y v11 (Ghahremani et al., 2025).

La detección de frutos se complementa con el análisis de índices de vegetación para obtener una visión más completa y cuantitativa del estado de los cultivos. Los índices de vegetación son herramientas fundamentales en el campo de la teledetección y el análisis de imágenes satelitales. Estos índices se calculan a partir de la combinación de diferentes bandas espectrales de la radiación electromagnética reflejada por la vegetación. Al comparar la reflectancia en distintas longitudes de onda, los índices de vegetación permiten obtener una estimación cuantitativa de la cantidad y calidad de la biomasa vegetal presente en una determinada área. Gracias a esta capacidad, los índices de vegetación se utilizan en una amplia variedad de aplicaciones, como el monitoreo de la salud de los cultivos, la detección de sequías, la evaluación de la deforestación y el estudio de los cambios en el tiempo de la cobertura terrestre.

Ballesteros et al. (2020) estimaron la producción de uva mediante una combinación de datos multiespectrales y geométricos del viñedo obtenidos desde un UAV. Las imágenes obtenidas con UAV, combinadas con índices como el índice de vegetación de diferencia normalizada (NDVI) (Rouse et al., 1974), se han utilizado para diferenciar entre vides vivas y secas (Gavrilović et al., 2024). Sin embargo, aunque algunos estudios han empleado índices de vegetación para estimar el rendimiento (Athanasίου et al., 2018), estas aproximaciones no incluyen la detección directa de los frutos (Sun et al., 2017). Los resultados muestran que índices como el NDVI enfrentan limitaciones significativas en la predicción del rendimiento, con variaciones en su efectividad según la etapa de desarrollo de las vides y valores del coeficiente de determinación R^2 inconsistentes que están influenciados por factores como el clima y los datos de referencia disponibles. Alternativas como el índice de vegetación mejorado (EVI) y el índice de vegetación ajustado al suelo (SAVI) también han mostrado resultados inconsistentes debido a problemas relacionados con la cobertura de nubes, las condiciones atmosféricas y dificultades en la calibración (Guan et al., 2019; Labus et al., 2002; Vannoppen and Gobin 2021; Karkauskaite et al., 2017; Jaafar et al., 2015; Ballesteros et al., 2020). Todas estas limitaciones ponen de manifiesto la necesidad de utilizar imágenes obtenidas por UAV para la detección directa de racimos de uvas y analizar posibles correlaciones entre estas estimaciones y los índices de vegetación. Este enfoque podría ofrecer una perspectiva más precisa y novedosa en la estimación del rendimiento vitícola.

Aunque las tecnologías actuales han progresado, aún no se ha logrado un consenso sobre su capacidad para proporcionar conteos de racimos completamente precisos. Los UAV, gracias a su alta resolución espacial, destacan por su eficacia para capturar detalles como los racimos de uvas individuales. En contraste, las imágenes satelitales, a pesar de su menor resolución, ofrecen ventajas por su amplia cobertura, lo que es valioso para monitorear la salud general de la vegetación y las tendencias de productividad en grandes extensiones. La integración de ambas herramientas podría ser

la clave para superar las limitaciones actuales, desarrollando sistemas más precisos y efectivos para la gestión vitivinícola.

1.3.1 Uva embolsada del Vinalopó

En la provincia de Alicante, la agricultura desempeña un papel crucial, no solo como motor económico, sino también como parte integral de su identidad cultural. La agricultura alicantina destaca por cultivos propios de la región como los cítricos o las hortalizas, otros más específicos como la granada y la higuera y, especialmente, la uva embolsada del Vinalopó, un producto con denominación de origen protegida.

La uva de mesa embolsada del Vinalopó es un producto agrícola emblemático de la provincia de Alicante, reconocido por su calidad excepcional y por un método de cultivo único que le otorga características distintivas. Cultivada en el valle del Vinalopó, esta uva de mesa goza de la protección de la DOP, que garantiza que el producto cumpla con estrictos estándares de calidad y se cultive exclusivamente en los municipios de Agost, Aspe, Hondón de las Nieves, Hondón de los Frailes, Monforte del Cid, Novelda y La Romana, según el Consejo Regulador de la DOP Uva de Mesa Embolsada del Vinalopó (CRDOP, 2024).

El clima mediterráneo del valle, caracterizado por inviernos suaves y veranos cálidos, en combinación con un suelo fértil y técnicas agrícolas tradicionales, contribuye significativamente a la calidad de este producto. Sin embargo, lo que verdaderamente distingue a la Uva Embolsada del Vinalopó de otras variedades de uva de mesa es su método de embolsado manual, una práctica artesanal que ha perdurado durante más de un siglo.

El embolsado de la uva es un proceso manual que comienza alrededor del mes de julio, cuando los racimos alcanzan un tamaño adecuado, pero aún están en desarrollo. Cada racimo se cubre cuidadosamente con una bolsa de papel biodegradable, diseñada específicamente para proteger y mejorar las uvas. Este método tiene múltiples ventajas. Por un lado, la bolsa protege los racimos de agentes externos como insectos, plagas, polvo y tratamientos químicos. Por otro lado, crea un microclima que favorece una maduración uniforme, protegiendo las uvas de condiciones climáticas adversas como el granizo, las altas temperaturas o las quemaduras solares. Además, este proceso mejora las características organolépticas de la fruta, otorgándole una piel más fina, un color uniforme y un sabor inconfundible.



Figura 2. Embolsado de la Uva embolsada del Vinalopó. Fuente: Agroinformación.

Sin embargo, el embolsado es un proceso laborioso y costoso, ya que requiere una alta dedicación y mano de obra especializada, lo que incrementa el valor final del producto.

Para garantizar la calidad y autenticidad de la Uva Embolsada del Vinalopó, la DOP establece un riguroso sistema de inspección y control en cada etapa del proceso, desde el cultivo hasta la comercialización. Las inspecciones del proceso de embolsado en la Uva Embolsada del Vinalopó son tareas que requieren una gran inversión de tiempo y mano de obra. Estas inspecciones son esenciales para garantizar que cada racimo cumpla con los estándares establecidos, ya que el embolsado es una práctica distintiva que asegura la calidad del producto. Sin embargo, debido a la naturaleza manual y laboriosa de esta tarea, se están explorando alternativas tecnológicas que permitan optimizar el proceso. Estas tecnologías no solo agilizarían la verificación del cumplimiento del embolsado, sino que, principalmente, permitirían estimar de manera precisa la producción de cada parcela y analizar la variabilidad de rendimiento dentro de la misma. Esta capacidad es fundamental para una agricultura de precisión, ya que proporciona a los agricultores información detallada y espacial para optimizar las prácticas de manejo, mejorar la eficiencia en el uso de recursos y maximizar la sostenibilidad del cultivo.

1.4 Contribución a los Objetivos de Desarrollo Sostenible

La agricultura desempeña un papel crucial en la consecución de los Objetivos de Desarrollo Sostenible (ODS) establecidos por las Naciones Unidas, que buscan abordar los desafíos globales relacionados con la pobreza, el hambre, la salud, el cambio climático y la sostenibilidad ambiental. En particular, la Agricultura de Precisión, junto con tecnologías avanzadas como los vehículos aéreos no tripulados y la inteligencia artificial, se alinea con múltiples metas de los ODS, proporcionando soluciones innovadoras para transformar prácticas tradicionales en sistemas agrícolas y de gestión de recursos más sostenibles y eficientes.



Figura 3. Objetivos de Desarrollo Sostenible. Fuente: ONU 2018.

En el contexto del ODS 2, “Hambre Cero”, la AP contribuye directamente a garantizar sistemas agrícolas que aumenten la productividad y la sostenibilidad. Mediante el uso de tecnologías como la teledetección, los SIG y el análisis de datos ayudan a usar de forma más eficiente recursos importantes como el agua, los fertilizantes y los pesticidas. Esto no solo mejora la producción, sino que también reduce el impacto en el medio ambiente. Por ejemplo, los UAV pueden detectar rápidamente problemas como plagas, enfermedades o falta de agua en los cultivos, permitiendo acciones más precisas y evitando desperdicios.

Además, los UAV están directamente relacionados con el ODS 13, “Acción por el Clima”, ya que proporcionan herramientas esenciales para monitorear y gestionar los impactos del cambio climático en los sistemas agrícolas. Los datos generados por estos dispositivos permiten evaluar y minimizar los efectos de fenómenos climáticos extremos, como sequías e inundaciones. La capacidad de los UAV para recopilar datos detallados también apoya estrategias de adaptación climática, fundamentales para la sostenibilidad a largo plazo.

En lo que respecta a la gestión de residuos, la acumulación de plásticos en los sistemas de riego representa un reto significativo para el cumplimiento de los ODS 6, “Agua Limpia y Saneamiento”, y 14, “Vida Submarina”. La presencia de plásticos flotantes afecta la calidad del agua y la biodiversidad acuática, además de perjudicar las infraestructuras agrícolas. Las tecnologías avanzadas, como los UAV combinados con IA, en esta tesis doctoral han demostrado ser eficaces para detectar, clasificar y cuantificar residuos plásticos en entornos acuáticos. Estas innovaciones permiten implementar políticas más efectivas para prevenir la contaminación, mejorando la gestión del agua y la sostenibilidad de las actividades humanas.

La acumulación de flotantes en los cauces representa una amenaza latente para nuestras infraestructuras y ecosistemas. Las recientes inundaciones en Valencia ponen de manifiesto la magnitud de esta amenaza y las consecuencias de no abordar este problema. Además de las infraestructuras de riego, puentes, alcantarillados y embalses se encuentran en riesgo. Para hacer frente a esta situación, es necesario invertir en tecnologías de detección temprana y remoción de flotantes, así como promover

prácticas sostenibles que reduzcan la generación de residuos y fomenten la conservación de los ecosistemas acuáticos. Al reducir la cantidad de residuos se contribuye a mitigar los efectos devastadores de avenidas. De esta manera, la metodología contribuye a la consecución del ODS 9, promoviendo infraestructuras resilientes, y del ODS 11, construyendo ciudades y comunidades más seguras y sostenibles.

La integración de estas tecnologías también promueve el ODS 12, “Producción y Consumo Responsables”, al fomentar el uso eficiente de recursos y la reducción de desperdicios en la agricultura. Además, el desarrollo y la adopción de estas herramientas impulsan la innovación tecnológica y el fortalecimiento de infraestructuras modernas, en línea con el ODS 9, “Industria, Innovación e Infraestructura”.

Esta tesis doctoral explora el uso de UAV e inteligencia artificial en dos áreas clave: la gestión de residuos plásticos en sistemas de riego y la agricultura de precisión aplicada en viticultura. Al abordar estos desafíos desde una perspectiva tecnológica y sostenible, esta investigación contribuye al cumplimiento de los ODS relevantes, promoviendo un futuro donde la agricultura sea más eficiente y los sistemas ambientales estén mejor protegidos frente a los impactos de la actividad humana.

2 Objetivos

El objetivo general de esta tesis es desarrollar soluciones tecnológicas basadas en vehículos aéreos no tripulados (UAV) e inteligencia artificial (IA) para abordar desafíos en la gestión de residuos plásticos en sistemas de riego y la agricultura de precisión. Estas áreas de investigación se centran en optimizar procesos agrícolas, mejorando tanto la productividad como la sostenibilidad en el contexto de la provincia de Alicante.

En el ámbito de la gestión de residuos plásticos en sistemas de riego, el objetivo general es proponer una metodología estandarizada y eficiente para la detección, conteo y caracterización de residuos flotantes acumulados en los sistemas de riego tradicionales, como los de la Vega Baja del Segura. Para ello se definen los siguientes objetivos específicos:

- Diseñar y entrenar un algoritmo basado en YOLOv5s para detectar automáticamente los diferentes tipos de residuos flotantes mediante análisis de imágenes.
- Evaluar el desempeño del modelo comparando sus resultados con imágenes previas que han sido contadas y clasificadas manualmente.
- Aplicar el modelo a fotografías aéreas capturadas con un vehículo aéreo no tripulado (UAV) en puntos estratégicos del río Segura y los canales de riego de la Vega Baja.
- Eliminar la subjetividad inherente al proceso manual de cuantificación y caracterización de residuos flotantes para obtener una descripción precisa de la magnitud del problema.

En el ámbito de agricultura de precisión en viticultura se propone evaluar el uso de inteligencia artificial y datos de teledetección para estimar el número de racimos de uva embolsados, mediante el análisis de imágenes obtenidas por UAV y satélites. Esto incluye la detección precisa de racimos de uvas embolsadas, la estimación del rendimiento de las vides individuales, y finalmente, de las parcelas. Asimismo, se investigará la correlación entre las estimaciones de rendimiento y los índices de vegetación determinados a partir de las imágenes satelitales. Los objetivos específicos que se plantean para cumplir este segundo objetivo son:

- Determinar la configuración óptima de vuelo de UAV para la adquisición de imágenes.
- Determinar la versión de YOLO más adecuada para la detección de racimos.
- Estimar el número de racimos de uva protegidos con bolsas de papel utilizando el modelo YOLOv7x en combinación con imágenes capturadas por UAV en un viñedo.
- Evaluar el potencial de diferentes índices de vegetación derivados de imágenes de los satélites Sentinel-2 y PlanetScope para estimar el número de racimos de uva.
- Explorar la novedad de incorporar inteligencia artificial y el uso de imágenes oblicuas para mejorar el proceso de detección de racimos de uva.

Este trabajo busca aportar soluciones innovadoras y prácticas a dos aspectos de

la actividad agrícola de la provincia que se pueden beneficiar de las técnicas de aprendizaje profundo, apoyándose en el avance tecnológico para mejorar la eficiencia en la gestión agrícola y ambiental, al tiempo que promueve un uso más sostenible de los recursos disponibles en el sector.

Los dos casos de estudio presentados son aplicaciones muy distintas de la teledetección con UAV y las técnicas de aprendizaje profundo, puesto que en un caso se usan imágenes cenitales y el objetivo es la detección y cuantificación de distintas clases de residuos, y en el otro caso, la aportación principal es la detección y conteo de objetos de una sola clase y su posicionamiento a lo largo de la espaldera y en el espacio de la parcela de cultivo, a fin de explorar posible relación con el vigor de las cepas.

3 Material y Métodos

3.1 Áreas de estudio

El presente estudio se desarrolló en dos ubicaciones de la provincia de Alicante, seleccionadas en función de sus características y de los objetivos específicos de la investigación.

La primera ubicación se sitúa en el sistema de riego tradicional de la Vega Baja del Segura localizado en el extremo sur de la provincia de Alicante (Comunidad Valenciana, España), este sistema hidráulico se extiende a lo largo del río Segura en su tramo final, distribuyendo agua mediante una compleja red de acequias y azarbes. La región es conocida por su agricultura intensiva y su sistema de riego histórico, que incluye azudes, acequias y otros elementos diseñados para la gestión eficiente del agua.

Se identificaron puntos específicos de acumulación de residuos en base a estudios previos (Rocamora et al., 2021) y nuevas observaciones realizadas durante el periodo de estudio. Estos puntos incluyeron tanto los puntos de muestreo determinados en los estudios previos como barreras de retención recién instaladas.

La segunda ubicación se situó en el municipio de Agost (Alicante) en una parcela de 1,03 hectáreas de *Vitis vinifera* L. cv. Aledo. El marco de plantación fue de 1,55 m x 2,50 m con un total de 2742 plantas, una producción registrada en 2022 de 12000 kg. Según el dato proporcionado por la cooperativa local, el peso medio de racimo fue de 598 g. Este viñedo está inscrito en la DOP Uva de Mesa Embolsada del Vinalopó.

Las condiciones climáticas mediterráneas de la zona, registradas en una estación meteorológica cercana, incluyeron una temperatura media anual de 17,86 °C, una humedad relativa del 63,81%, y una precipitación anual promedio de 338 mm. El riego por goteo, con una dotación estacional de 974 m³, aseguró que no se presentara estrés hídrico en el cultivo.

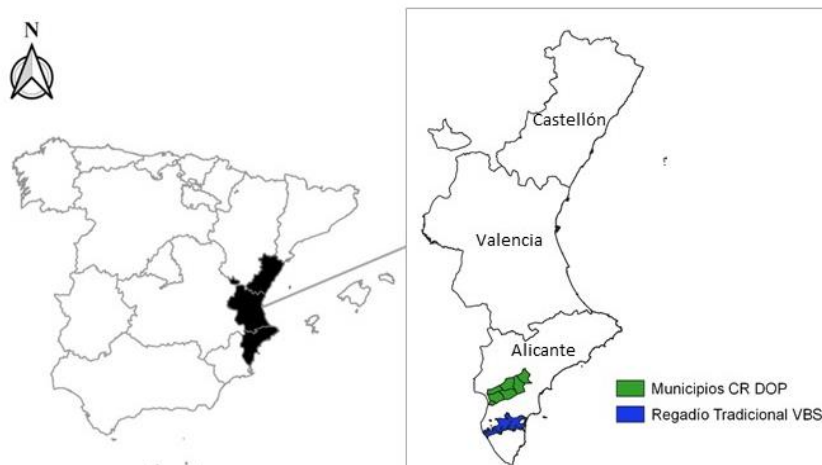


Figura 4. Localización de las áreas de estudio en la provincia de Alicante. Fuente: Elaboración propia

3.2 Adquisición de imágenes

Para la adquisición de imágenes de ambos estudios se utilizó el UAV Autel EVO pro 2 (UUC Technology Co., Ltd., Hong Kong), equipado con una cámara RGB visible con sensor CMOS de 1" (Sony Group Corporation, Tokio), apertura ajustable F2.8-F11 capaz de alcanzar una resolución de 20 megapíxeles y grabación de vídeo en 4K /30FPS.



Figura 5. UAV Autel EVO pro 2. Fuente: Elaboración propia

3.3 Configuración de vuelo para la detección de flotantes

Los vuelos se realizaron a alturas de 5, 10 y 20 metros, con rutas programadas para cubrir puntos de acumulación de residuos previamente identificados. Las imágenes se capturaron en vista cenital, asegurando una cobertura adecuada de las zonas de interés. Para definir los lugares de adquisición de imágenes, se mantuvieron también los mismos sitios de acumulación de basura a lo largo del río y la red de canales de riego descritos en Rocamora et al. (2021) añadiendo algunos nuevos sitios:

Tabla 1. Identificación, posición y descripción de los puntos de muestreo de flotantes en el sistema hidráulico de la Vega Baja.

Punto	Coordenadas	Descripción
ID 1	38.054452, -1.000317	Azud de las Norias
ID 2	38.080600, -0.966808	Barrera en el Molino de la Ciudad *
ID 2.a	38.085007, -0.952547	Barrera en la entrada de Orihuela
ID 3	38.087101, -0.823544	Barrera en Almoradí
ID 4	38.087319, -0.822053	Azud de Alfeitamí

ID 5	38.095986, -0.678104	Esclusas en la entrada de la Comunidad de Regantes de Riegos de Levante
ID 6	38.088012, -0.932084	Acequia de Callosa, sifón en el cruce con Acequia Vieja de Almoradí
ID 7	38.163356, -0.752248	Azarbe de Cebadas, sifón en el cruce con las acequias del Convenio
ID 8	38.110647, -0.650253	Pantalla flotante de Guardamar **

* En el año 2022 se añadió un nuevo punto de muestreo debido a una nueva barrera de retención en la entrada a la ciudad de Orihuela.

** En el año 2022 entraron en funcionamiento barreras de retención de residuos en las desembocaduras de los azarbes que vierten cerca del punto ID 8. Las nuevas barreras también se consideraron como puntos de evaluación adicionales para el presente estudio, y se numeran como ID8.a, ID8.b, ID8.c, ID8.d e ID8.e.

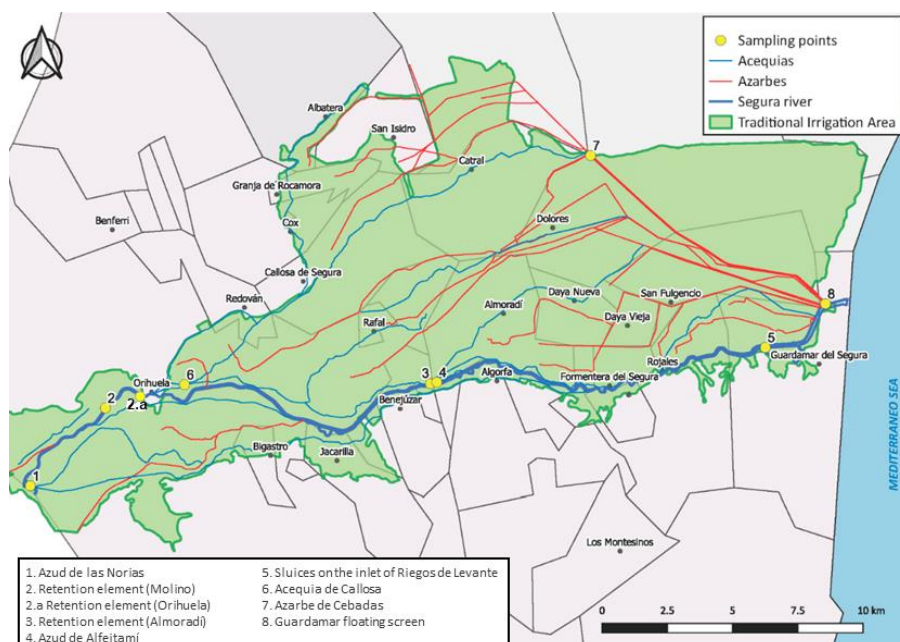


Figura 6. Puntos de acumulación de residuos. Fuente: Elaboración propia a partir de Rocamora et al., 2021

En total, se realizaron tres vuelos en cada punto durante las fechas seleccionadas (28 de julio, 8 de septiembre y 31 de octubre de 2022), obteniendo un total de 61 imágenes.

Para determinar la configuración de vuelo idónea que proporcionara una mejor precisión en la detección, se llevó a cabo un estudio previo (Codes et al., 2024) en el que se comparó dos configuraciones de vuelo de UAV: una a 10 m de altura con un ángulo de inclinación del sensor de 35° y otra a 17 m de altura con un ángulo de inclinación del sensor de 45°.

Los vuelos se llevaron a cabo siguiendo la línea de plantación. Para evitar la superposición excesiva de racimos, se aseguró un solape del 10% entre imágenes. Las imágenes obtenidas tuvieron una resolución de 3840 x 2160 píxeles, con un tamaño de píxel

aproximado de 1,1 cm.

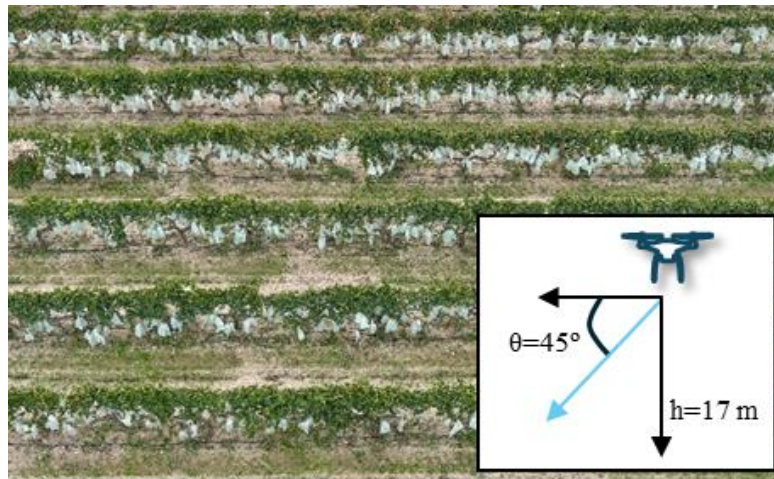


Figura 7. Imagen del UAV adquirida a 17 metros y con un ángulo de 45° . Fuente; Elaboración propia.

3.4 Imágenes satelitales

Para el cálculo de los índices de vegetación correspondientes a la parcela experimental, se emplearon imágenes satelitales de Sentinel-2 y PlanetScope.

Durante el periodo de estudio se descargaron 36 imágenes libres de nubes del portal Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) y del Copernicus Dataspace Ecosystem (<https://dataspace.copernicus.eu/>) a partir de octubre de 2023. De estas imágenes se utilizaron las bandas azul (490 nm), verde (560 nm), rojo (665 nm) y NIR (842 nm), todas ellas con resoluciones de píxel de 10 m.

Correspondientes al mismo periodo se descargaron 100 imágenes de alta resolución (tamaño de píxel de 3 m) de portal Planet Explorer (<https://www.planet.com/explorer/>). En este caso se seleccionaron las bandas borde rojo (705 nm), rojo (665 nm), verde (565 nm), verde I (531 nm), amarillo (610 nm), azul (490 nm), azul marino (442 nm) y NIR (865 nm).

Para el procesamiento de las imágenes satelitales se utilizó ArcGIS 8.3, un software de sistema de información geográfica (SIG) desarrollado por el *Environmental Systems Research Institute* (ESRI), California. Este software fue empleado tanto para extraer las bandas de Sentinel-2 y PlanetScope para cada píxel y calcular los índices de vegetación, como para geolocalizar el número de racimos por vid obtenido del conteo automático y relacionarlo con los cálculos anteriores.

En el estudio de detección de racimos embolsados se evaluaron correlaciones entre el número de racimos detectados y 11 índices de vegetación calculados a partir de imágenes Sentinel-2 y PlanetScope. Los índices de vegetación calculados se muestran en la tabla 2.

Tabla 2. Índices de vegetación y fórmulas de cálculo

Índices de vegetación	Ecuación	Referencia
Índice de Vegetación de Diferencia Normalizada (NDVI)	$\frac{(NIR - RED)}{(NIR + RED)}$	Rouse et al., 1974
Índice de Vegetación Ajustado al Suelo (SAVI)	$\frac{(1 + 0.5) * (NIR - RED)}{NIR + RED + 0.5}$	Huete et al., 1988
Índice de Vegetación Mejorado (EVI)	$\frac{2.5 * (NIR - RED)}{(NIR + 6 * RED - 7.5 * BLUE + 1)}$	Huete et al., 1997
Índice de Vegetación de Diferencia Normalizada en Verde (GNDVI)	$\frac{NIR - GREEN}{NIR + GREEN}$	Gitelson et al., 1996
Índice de Vegetación Ajustado al Suelo Modificado (MSAVI)	$0.5 * (2(NIR + 1) - \sqrt{((2 * NIR + 1)^2 - 8(NIR - RED))})$	Qi et al., 1994
Índice Vegetal Triangular Modificado (MTVI)	$\frac{1.5[1.2(NIR - GREEN) - 2.5(RED - GREEN)]}{\sqrt{(2NIR + 1)^2 - (6NIR - 5\sqrt{RED})} - 0.5}$	Haboudane et al., 2004
Índice de Diferencia Normalizada del Agua (NDWI)	$\frac{GREEN - NIR}{GREEN + NIR}$	McFeeters, 1996
Índice de Vegetación Optimizado Ajustado al Suelo (OSAVI)	$\frac{(NIR - RED)}{NIR + RED + 0.16}$	Rondeaux et al., 1996
Índice de Relación Simple (SR)	$\frac{NIR}{RED}$	Jordan, 1969

Los datos se integraron en mapas generados en ArcGIS para representar la distribución espacial de los racimos en la parcela.

Para ello durante el entrenamiento del algoritmo, se identificó y segmentó cada vid de la parcela, y las imágenes resultantes se utilizaron como entrada para el análisis del algoritmo. El algoritmo detectó y contó los racimos de uva presentes en cada vid. Luego, utilizando una cuadrícula de puntos creada en ArcGIS, se integraron los datos del racimo de cada vid, generando un mapa que muestra la distribución de las vides dentro de la parcela, junto con el recuento de racimos asociado. Este método permite una estimación precisa del número de vides y racimos de uva en cada píxel de la imagen. Para calcular los índices de vegetación, se extrajeron los valores de las bandas espectrales de Sentinel-2 y PlanetScope en los píxeles seleccionados en ArcGIS, y los datos se transfirieron a una hoja de cálculo de Excel, donde se calcularon los índices utilizando las fórmulas referidas en la tabla 2.

3.5 Procesamiento y entrenamiento de modelos

En la detección de residuos se recolectaron 477 imágenes que incluyeron datos de estudios previos (Rocamora et al., 2021) y nuevas fotografías. Estas imágenes fueron etiquetadas utilizando Makesense, una herramienta gratuita de etiquetado de fotografías en línea que no requiere ninguna instalación (<https://www.makesense.ai/index.html>), asignando cinco clases: botellas de agua, botellas de detergente, botellas de lejía, bidones agrícolas y pelotas. Estas clases corresponden a los objetos más frecuentes en el estudio previo. Las imágenes se dividieron en un 90% para entrenamiento y un 10% para validación, siendo posteriormente procesadas en formato YOLO. En este estudio se empleó YOLOv5 (Tian et al., 2023), una red de detección de objetos de una sola etapa basada en regiones. Este modelo aborda la detección de objetos como un problema de regresión, lo que facilita un procesamiento rápido. Su arquitectura se divide en tres componentes principales: columna vertebral, cuello y cabeza.

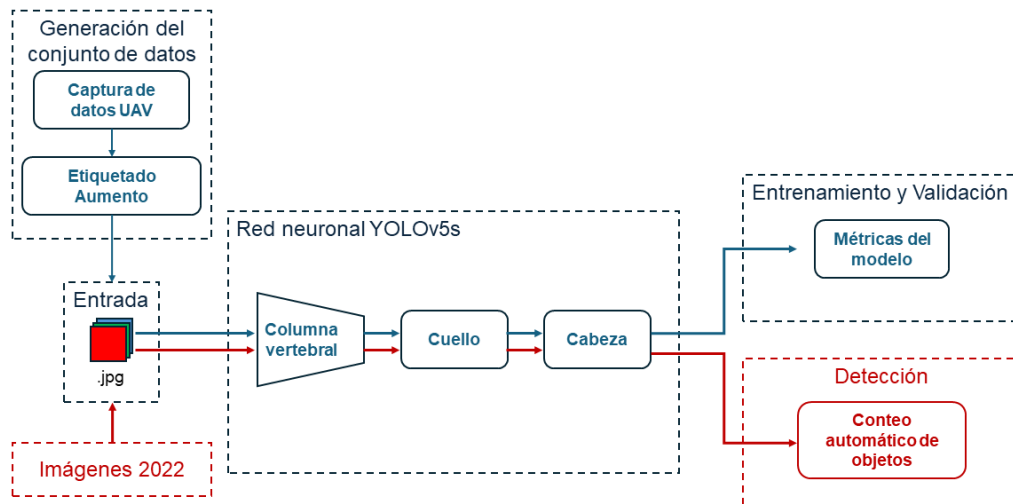


Figura 8: Marco conceptual de la metodología de detección de objetos. Fuente: Elaboración propia.

El módulo de columna vertebral se encarga de extraer características de la imagen de entrada utilizando elementos como Focus, Bottle-neckCSP (una red parcial interetapa) y SPP (piramidal de agrupación espacial), transmitiéndolas al siguiente módulo. Por su parte, el módulo de cuello genera una pirámide de características basada en PANet (*Path Aggregation Network*), lo que mejora la detección de objetos a múltiples escalas al combinar características espaciales de bajo nivel con semánticas de alto nivel de manera bidireccional. Finalmente, el módulo de cabeza produce las cajas delimitadoras, indicando la categoría, las coordenadas y la confianza asociada mediante el uso de anclajes aplicados a los mapas de características multiescala generados por el cuello. El modelo YOLOv5 cuenta con cuatro variantes principales: YOLOv5s, YOLOv5m, YOLOv5l y YOLOv5x. Además, el módulo SPP juega un papel importante al expandir el campo receptivo de la red y capturar características en distintas escalas (Redmon et al., 2016; Yao et al., 2021).

Por otro lado, en la detección de racimos de uva se etiquetaron 60 imágenes aéreas utilizando también la aplicación de Makesense para etiquetar cada una de las bolsas de cada racimo de uva, generando un total de 400 etiquetas por imagen. Estas fueron ajustadas también al formato YOLO. Cuando se realizó el análisis de estas imágenes (2024) existían versiones posteriores a YOLOv5. Para determinar la versión más adecuada se compararon las versiones YOLOv7-E6, YOLOv7x y YOLOv8 (Codes et al., 2024). Lo que distingue a YOLOv7 de versiones anteriores es su notable precisión en la detección, lograda gracias a una serie de innovaciones. Estas incluyen la optimización en la eficiencia de las capas, técnicas avanzadas de escalado de modelos, estrategias de reparametrización y la incorporación de una cabeza auxiliar con un enfoque de supervisión de lo general a lo específico. Estas mejoras combinadas aportan avances significativos en rendimiento, eficiencia y adaptabilidad, posicionando a YOLOv7 como una solución más efectiva para múltiples aplicaciones prácticas (Badeka et al., 2023). Entre las versiones desarrolladas dentro de YOLOv7 se encuentra YOLOv7x, que ha demostrado en diversos estudios obtener los mejores resultados (Badeka et al., 2023).

El entrenamiento de ambos modelos, detección de flotantes y detección de uva embolsada, se llevó a cabo en una la plataforma de cómputo en la nube Google Colaboratory (<https://colab.research.google.com/>), que permite el uso de Unidades de Procesamiento Gráfico (GPU) y Unidades de Procesamiento Tensorial (TPU) sin costo adicional. Este entorno facilitó la implementación de los algoritmos y su optimización mediante un ajuste de parámetros diferentes en cada estudio.

En el caso de la detección de flotantes en cauces en la Vega Baja, el modelo YOLOv5 se entrenó con imágenes de 1024 píxeles y 126 épocas. El tamaño de la imagen se refiere al número de veces que se divide una imagen durante su análisis, un aumento en este número significa que la imagen se analiza con mayor precisión. El número de épocas es un hiperparámetro que define el número de veces que funcionará el algoritmo de aprendizaje en todo el conjunto de datos de entrenamiento. Una época significa que cada muestra en el conjunto de datos de entrenamiento ha tenido la oportunidad de actualizar los parámetros del modelo interno (Brownlee, 2024).

En el conteo de racimos embolsados en viñedos, YOLOv7x se entrenó con imágenes de 640 píxeles y en este caso se necesitó 250 épocas para el entrenamiento del algoritmo. En ambos casos, el número de épocas se elige basándose en el análisis de las gráficas de precisión obtenidas durante el entrenamiento. Cuando estas gráficas muestran que la precisión del modelo deja de mejorar significativamente con el incremento de épocas, o incluso empieza a decrecer debido al sobreajuste, se detiene el entrenamiento. Este enfoque asegura que se logre un equilibrio entre un buen rendimiento del modelo y la eficiencia del entrenamiento.

La herramienta utilizada en el campo del aprendizaje automático y la estadística para evaluar el rendimiento de un modelo de detección como Yolo es la matriz de confusión (figura 9). Es una tabla que compara los resultados predichos por un modelo con los resultados reales, lo que permite evaluar la precisión y eficacia del modelo.

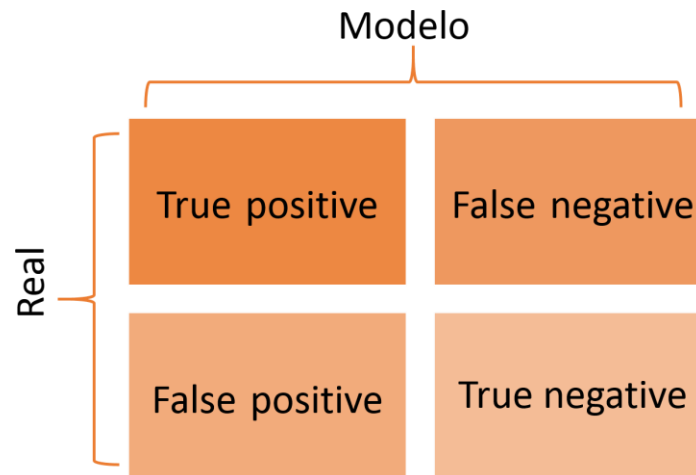


Figura 9. Matriz de confusión. Fuente: Elaboración propia

En la matriz de confusión el eje vertical representa las clases reales de los datos y el eje horizontal representa las clases predichas por el modelo. Los casos *True Positive* (TP) son aquellos en los que el modelo predijo correctamente la clase positiva y los casos *True Negative* (TN) son aquellos en los que el modelo predijo correctamente la clase negativa. De esta manera, la diagonal principal de la matriz indica el número de predicciones correctas, y los valores fuera de la diagonal representan los errores de clasificación.

Los casos *False Positive* (FP), también conocido como error tipo I, ocurren cuando el modelo predice como positiva una clase que realmente es negativa. Los casos *False Negative* (FN), también conocido como error tipo II, ocurre cuando el modelo predice como negativa una clase que realmente es positiva.

Para la validación de ambos modelos se utilizaron métricas estándar de detección de objetos como son la precisión (P), *recall* o ratio de verdaderos positivos (R), puntuación F1 y precisión promedio (Average Precision, AP) (Olson y Delen, 2008).

P es la proporción de muestras positivas verdaderas entre todas las muestras positivas predichas por el modelo (Ec. 1).

$$P = \frac{TP}{TP + FP} \cdot 100\% \quad (1)$$

La proporción de muestras positivas verdaderas predichas por el modelo con respecto a todas las muestras positivas verdaderas se considera mediante el término R (Ec. 2).

$$R = \frac{TP}{TP + FN} \cdot 100\% \quad (2)$$

Por lo general, existe una correlación inversa entre P y R. Para equilibrar los efectos de P y R y evaluar un modelo de manera global, se calculó AP a través del área bajo la curva P-R (Ec. 3), donde un valor más alto significa un mejor desempeño del modelo.

$$AP = \int_0^1 P(R) dR \quad (3)$$

La puntuación F1 (Ec. 4) es la media armónica de P y R, lo que proporciona una evaluación equilibrada del desempeño de un modelo al considerar tanto los falsos positivos como los falsos negativos (Ec. 4).

$$\text{puntuación } F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (4)$$

Una de las métricas más utilizadas para evaluar el entrenamiento es la Precisión Media Promedio (mAP), que mide la precisión promedio media del modelo considerando diferentes umbrales de confianza, donde mAP@0,5 representa el valor medio de mAP cuando el umbral IoU (Intersección sobre Unión) se establece en 0,5. Este valor, que varía de 0 a 1, refleja el rendimiento general del modelo a la hora de identificar correctamente los racimos de uva en la mayoría de las imágenes analizadas (Badeka et al., 2023).

3.5.1 Validación del modelo de detección de residuos flotantes y de racimos

Las predicciones del modelo desarrollado para el conteo de flotantes se contrastaron con los datos reales. Para ello, las 61 imágenes utilizadas en el modelo se analizaron con ayuda del software ImageJ 1.54g y se contó manualmente los objetos que aparecen en ellas. Los conteos automáticos del modelo y los conteos manuales con la asistencia de ImageJ se compararon utilizando el método de Bland y Altman (1986).

En el caso del viñedo para la validación del modelo se comparó el número de racimos y el rendimiento total estimado para cada una de las diferentes configuraciones de vuelo y versión de YOLO con el número de racimos identificados visualmente en cada imagen seleccionada. Para realizar una comparación exacta entre las diferentes versiones, se seleccionaron 50 imágenes: 25 imágenes con la configuración h= 10 m y $\theta = 35^\circ$; y 25 imágenes con la configuración h= 17 m y $\theta = 45^\circ$. Estas imágenes se recortaron para que coincidieran con la misma ubicación en ambas configuraciones de vuelo con el fin de determinar cuál era la más efectiva para el conteo de racimos.

Finalmente se comparó la estimación de la producción con la producción real de la parcela. La producción se estimó multiplicando el número de bolsas detectadas por el algoritmo por el peso medio del racimo facilitado por la cooperativa local, 598 g/racimo. El dato medido de producción real de la parcela fue proporcionado por la misma cooperativa.

4 Resultados y discusión

En ambos estudios se aplicaron algoritmos de detección basados en YOLO para resolver problemas específicos: identificar residuos flotantes en los cauces y detectar racimos de uvas embolsados en viñedos.

4.1 Resultados de la clasificación y cuantificación de residuos flotantes

En el caso de los residuos flotantes, al inicio del entrenamiento (Figura 10), se observa un incremento rápido, seguido de una estabilización progresiva alrededor de las 126 épocas, momento en el cual el algoritmo alcanza su rendimiento máximo en términos de $mAP@0,5$, logrando un valor de 0,96. En la Figura 8b se presenta el mAP en un rango de umbrales entre 0,5 y 0,95, donde se aprecia un crecimiento constante hasta aproximadamente 120 épocas, estabilizándose en un valor cercano a 0,73. Esto sugiere que el modelo alcanza su mejor desempeño para esta métrica en ese punto.

Por otro lado, la Figura 8c muestra la evolución de la precisión durante el entrenamiento. Al inicio, la precisión experimenta fluctuaciones significativas, pero estas disminuyen conforme avanza el proceso, alcanzando una estabilización en torno a 0,95 a partir de las 120 épocas, lo que indica que el modelo ha logrado su máxima precisión. Continuar el entrenamiento más allá de este punto no aportaría mejoras significativas en las métricas evaluadas.

En cuanto al tamaño de lote, se utilizó un valor de 2, ya que ofreció una adecuada matriz de confusión y evitar el uso de un tamaño mayor permitió optimizar el tiempo de cálculo.

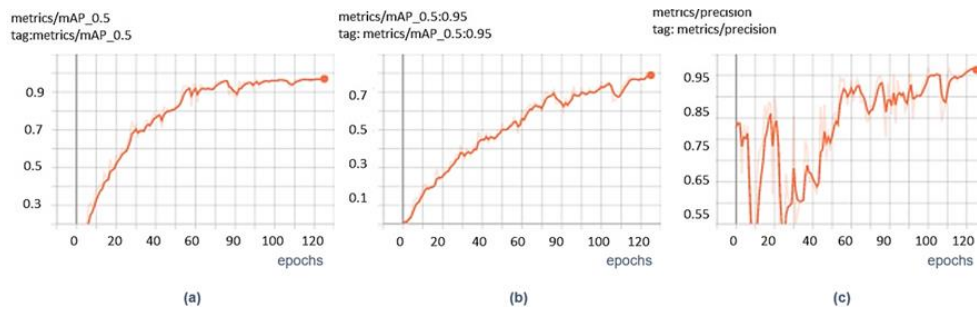


Figura 10. Gráficas de precisión (a) $mAP@0,5$, (b) $mAP@0,5$ a 0,95, (c) precisión.

El rendimiento del modelo se evaluó mediante precisión promedio media (mAP), precisión, *recall*, tiempo de prueba y puntuación F1. En este experimento, se utilizó un umbral de confianza de 0,5, obteniendo los siguientes resultados: un tiempo de entrenamiento de 3 horas y 7 minutos; una precisión del 92,04%; un recuerdo del 97,07%; un mAP del 96,90% y una puntuación F1 del 94,48%.

El objetivo principal de este trabajo fue identificar distintos residuos de manera rápida y precisa. Para ello, se consideraron como métricas clave el mAP (96,90%) y la velocidad de detección, que fue de 81,7 ms por imagen. Un mAP más alto indica un mejor rendimiento en la detección. Lo que demuestra al final una alta precisión y rapidez en la identificación de objetos, especialmente en botellas de lejía, bidones de uso agrícola y balones, con una precisión del 99,5%.

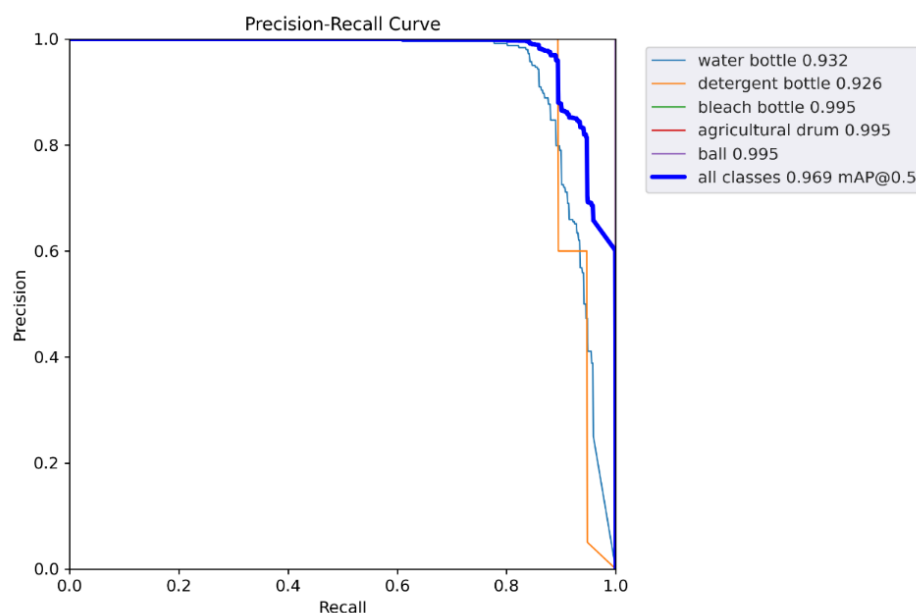


Figura 11. Curva de precisión-recuperación de los objetos detectados.



Figura 12. Detección automática de los residuos flotantes

En el conteo manual de los residuos flotantes se comprobó que los objetos más comunes fueron botellas de agua (96%), seguidas de botellas de detergente (2,07%) y lejía (1,12%), mientras que los menos frecuentes fueron bidones agrícolas (0,6%) y pelotas (0,3%). Este patrón se repitió en todos los puntos de muestreo y fechas.

El tiempo necesario para el recuento manual de 61 imágenes con la aplicación ImageJ fue de 7 horas y 34 minutos.

Al comparar las técnicas, el conteo automático siempre es igual o menor al manual en el caso de botellas de agua, mientras que en otras clases de objetos las diferencias son más variables. El análisis estadístico reveló un sesgo positivo para algunos objetos,

lo que indica una subestimación por parte del modelo en comparación con el conteo manual, especialmente en el caso de botellas apiladas.

4.2 Resultados de conteo de racimos embolsados

4.2.1 Configuración del vuelo

Para todas las versiones de YOLO evaluadas, el número estimado de racimos por vid fue más preciso utilizando la configuración de $h=10$ m y $\theta=35^\circ$ en comparación con la configuración de $h=17$ m y $\theta=45^\circ$. Más específicamente, la configuración de $h=10$ m y $\theta=35^\circ$ proporcionó términos de pendiente en el rango de 0,90-1,03, interceptos inferiores a 1,00, valores del coeficiente de determinación R^2 entre 0,79 y 0,84, y valores del error cuadrático medio RMSE inferiores a 1,17 racimos por vid. Por el contrario, la configuración de $h=17$ m y $\theta=45^\circ$ resultó en términos de pendiente más bajos, interceptos más altos, valores de R^2 más bajos y valores de RMSE más altos en comparación con la configuración de $h=10$ m y $\theta=35^\circ$. Estos resultados evidencian una mejor capacidad del modelo para identificar los racimos a bajas altitudes y ángulos de inclinación menores, probablemente debido a la mayor proximidad a los objetos, lo que facilita su detección precisa. Sin embargo, la identificación de racimos en vides situadas en el fondo de la imagen se ve más afectada por la obstrucción del dosel de vides cercanas al usar la configuración de $h=10$ m y $\theta=35^\circ$ que al utilizar la configuración de $h=17$ m y $\theta=45^\circ$.

Tabla 3. Ecuaciones de correlación, coeficiente de determinación (R^2) y error cuadrático medio (RMSE) con las distintas configuraciones de vuelo.

Versión de YOLO	Configuración de vuelo	Ecuación	R^2	RMSE (racimos vid ⁻¹)
v.7x	$h=10\text{m}$ $\theta=35^\circ$	$y = 1,03x + 0.29$	0,82	0,75
	$h=17\text{m}$ $\theta=45^\circ$	$y = 0,77x + 1.26$	0,64	0,78
v.7-E6	$h=10\text{m}$ $\theta=35^\circ$	$y = 0,90x - 0.32$	0,84	1,17
	$h=17\text{m}$ $\theta=45^\circ$	$y = 0,69x - 0.04$	0,71	2,36
v.8	$h=10\text{m}$ $\theta=35^\circ$	$y = 0,98x - 0.09$	0,79	0,62
	$h=17\text{m}$ $\theta=45^\circ$	$y = 0,56x + 2.10$	0,53	1,39

4.2.2 Comparación de las versiones del modelo YOLO

Con la configuración de vuelo a 17 m y ángulo de 45° , YOLOv7x tuvo un mejor desempeño que las otras versiones evaluadas (YOLOv7-E6 y YOLOv8), demostrando mayor estabilidad y menores valores de RMSE (Codes et al., 2024). En concreto, las tres versiones de YOLO mostraron valores de mAP similares, pero YOLOv7x presentó

los valores más bajos de RMSE (Tabla 3). El mejor desempeño de YOLOv7x también fue observado por Badeka et al. (2023) al evaluar el rendimiento total estimado a partir de las diferentes versiones de YOLO. Las estimaciones de rendimiento de la versión YOLOv7x no mostraron diferencias significativas con el rendimiento medido, mientras que se observó una subestimación significativa del rendimiento para las versiones YOLOv7-E6 y YOLOv8 con una configuración de altura de 17 m y un ángulo de inclinación del sensor de 45°. Dado que las imágenes utilizadas en el presente estudio fueron capturadas a 17 m, se eligió YOLOv7x como modelo de detección, ya que proporcionó resultados más fiables y consistentes con esta configuración de vuelo.

El modelo YOLOv7x aplicado a los viñedos logró un mAP@0,5 de 63%, con una precisión de 61%, lo que indica que, aunque es funcional, requiere mejoras para reducir falsos positivos y mejorar la localización precisa de los racimos. El *recall* alcanzó el 79%, lo que representa que el modelo detectó con éxito el 79% de los racimos de uva presentes en las imágenes. Finalmente, el puntaje F1 obtenido fue del 66%, indicando media armónica entre la precisión y el *recall*.

Los parámetros de entrenamiento variaron entre los estudios: para el modelo de residuos se usó un tamaño de *batch* de 2 y 126 épocas, mientras que el modelo para los racimos se entrenó con un tamaño de *batch* de 12 y 250 épocas. Esto refleja la adaptación de los hiperparámetros a las necesidades específicas de cada tarea.



Figura 13. Detección automática de los racimos de uva embolsados

En el caso de los racimos de uva, la correlación entre valores estimados y observados fue significativa, con un R^2 de 0,64. Sin embargo, se observaron errores debido a racimos parcialmente cubiertos o iluminación desfavorable, destacando la necesidad de mejoras en el preprocesamiento de las imágenes y en la representación de datos de entrenamiento.

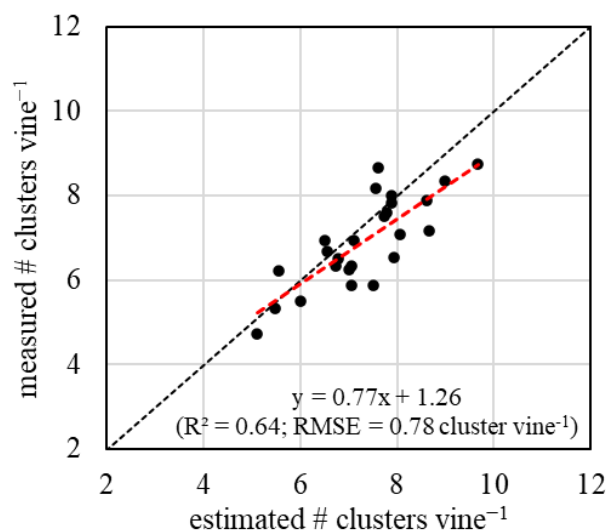


Figura 14. Comparación entre el número de racimos embolsados estimados por YOLOv7x y las mediciones reales observadas en las imágenes del UAV. La línea roja discontinua representa la línea de tendencia de la regresión, mientras que la línea negra discontinua indica la relación 1:1.

El estudio sobre la detección de la uva embolsada del Vinalopó incluyó un análisis de índices de vegetación obtenidos de los satélites Sentinel-2 y PlanetScope. La comparación entre el número total de racimos detectados por YOLOv7x y los índices de vegetación promedio obtenidos para cada imagen y sus valores acumulativos mostró que los índices de vegetación no predijeron con precisión el número de racimos. Esto se observó tanto con índices derivados de PlanetScope como con los calculados a partir de imágenes de Sentinel-2. En detalle, la falta de relación significativa se evidenció en los bajos valores de R^2 , con un máximo de 0,23 usando imágenes de Sentinel-2 (NDWI y GNDVI frente al número de racimos en vid el 6 de noviembre de 2022), y aún menores con imágenes de PlanetScope (máximo de 0,02 para NDVI el 25 de mayo de 2022). El uso de valores acumulativos durante el período evaluado tampoco mejoró los resultados, con valores de R^2 inferiores a 0,09 en la relación entre índices de vegetación y número de racimos en vid.

Esto subraya las limitaciones de estos índices para estimaciones específicas de rendimiento, ya que están diseñados para medir vigor y cobertura vegetal general más que para identificar detalles específicos como racimos.

4.2.3 Comparación de la estimación de cosecha con la producción real.

La producción de la parcela, según datos proporcionados por la cooperativa, fue de 12000 kg en la campaña de 2022.

La estimación de producción realizada a partir del número de bolsas detectadas por los modelos fue en todos los casos inferior a la producción medida (Figura 15).

Esta estimación a la baja de los modelos puede ser debida a la falta de detección de algunas bolsas que quedan tapadas por otras.

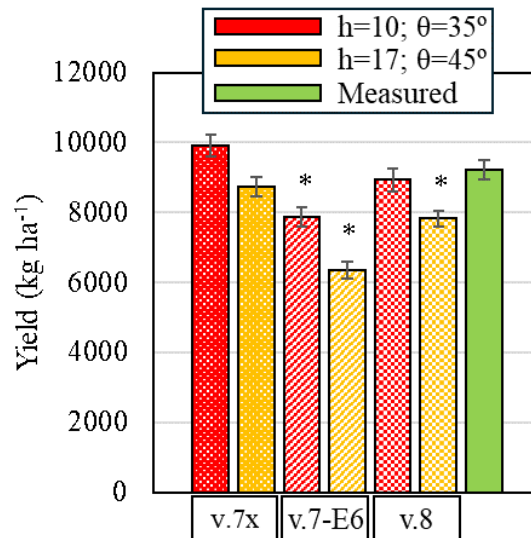


Figura 15. Resultados de estimación de cosecha de los modelos.

La figura 16 muestra un mapa de producción que refleja la variabilidad espacial de los racimos de uva detectados mediante el modelo YOLOv7x, aplicado a imágenes obtenidas por UAV. Este mapa permite visualizar cómo se distribuyen los racimos en la parcela, representando el número de racimos por unidad de longitud de la espaldera a través de un gradiente de colores. Estas variaciones espaciales pueden deberse a factores como el manejo agronómico (poda, fertilización, etc.), las condiciones del suelo, el estado fisiológico de las plantas o la disponibilidad de agua.

La variabilidad espacial evidenciada en el mapa es básica en la agricultura de precisión, ya que proporciona información detallada para la toma de decisiones. Por ejemplo, las zonas con baja producción podrían requerir ajustes en prácticas de manejo como la fertilización, el riego o el control fitosanitario, mientras que las áreas con alta densidad de racimos podrían ser manejadas para maximizar su potencial productivo. Además, este mapa facilita la estimación precisa de la cosecha total de la finca, lo que permite planificar eficientemente las actividades de recolección y logística. Asimismo, es útil para identificar zonas críticas donde haya problemas específicos como enfermedades, plagas o deficiencias nutricionales, para que puedan abordarse oportunamente.

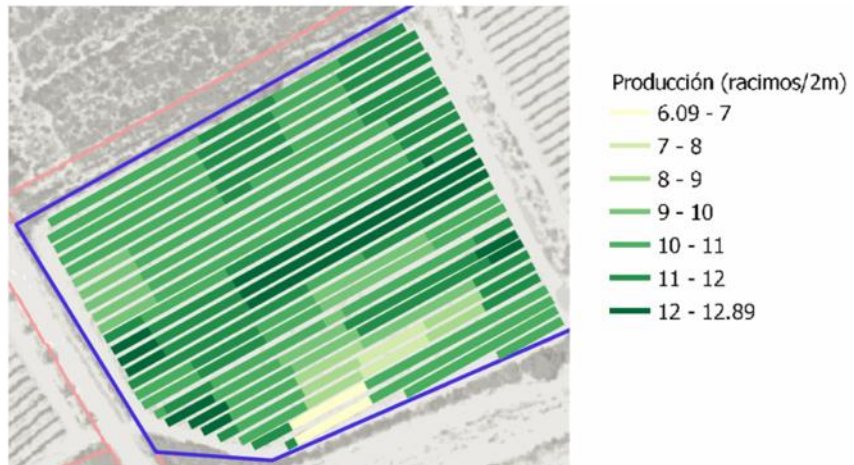


Figura 16. Mapa de producción según el número de racimos detectados por el modelo.

5 Discusión

5.1.1 Modelo de detección de flotantes

Debido a la reducida disponibilidad de modelos para detectar residuos plásticos cuando se acumulan en el medio acuático continental mediante imágenes de UAV, se introdujo este modelo para mejorar y facilitar la tarea de identificación y caracterización de residuos plásticos para agilizar las tareas de limpieza y poder aplicar medidas para reducir este problema de forma más rápida. Los resultados de este estudio muestran un excelente rendimiento en el tiempo de cómputo y la identificación de residuos. El modelo funcionó bien tanto para residuos aislados como para residuos agrupados, que se detectaron con precisión.

El algoritmo propuesto logró un $mAP@0,5$ del 96,9% y una velocidad de detección de 81,7 ms. En comparación con el de otros trabajos de investigación, como Zhang et al. (2023) con un $mAP@0,5$ del 98,1%, Li et al (2023) con un $mAP@0,5$ del 82,9% y Lin et al. (2021) con un $mAP@0,5$ del 76,12%, este enfoque proporcionó resultados sólidos tanto en precisión como en velocidad de detección. El algoritmo desarrollado en este trabajo puede clasificar diferentes tipos de botellas de plástico: botellas de agua, botellas de productos de higiene y limpieza, y latas de productos químicos utilizados en la agricultura. Esta distinción entre tipos de botellas de plástico representa una mejora con respecto al algoritmo de Wolf et al. (2020), ya que el tipo de botella puede indicar su origen. Sin embargo, este modelo sufre de detecciones falsas en el fondo de las imágenes, identificándolas como botellas de plástico. Las botellas de agua son las más difíciles de caracterizar, ya que se encuentran en gran número y se pueden ver superpuestas, lo que hace que la botella parezca de diferentes formas. Por el contrario, las botellas y bidones de lejía son más fáciles de identificar, ya que tienen una forma y un color característicos y no se encuentran en grandes cantidades.

Los resultados del recuento automático en tres fechas diferentes en sitios de acumulación de residuos flotantes obtenidos en este trabajo concuerdan con los reportados por Rocamora et al. (2021) en el mismo entorno, quienes encontraron que la clase de residuo más numerosa está constituida por botellas de agua, mientras que las botellas y bidones de productos fitosanitarios representaron solo el 3,8% de los elementos contabilizados en los puntos de muestreo. En el presente estudio, a lo largo de tres días de medición en ocho puntos de muestreo, se encontró que el 95% de los elementos flotantes eran botellas de agua de plástico (96% por recuento manual), y solo el 0,8% de los elementos eran envases agrícolas (0,6% por recuento manual).

Para comparar los resultados del algoritmo con los del recuento manual, es necesario destacar el tiempo empleado con las dos metodologías. El tiempo requerido para el recuento manual de 61 imágenes con la aplicación ImageJ fue de 7 h y 34 min. Por otro lado, el recuento automático de esas 61 imágenes duró 42 s. Por tanto, se observa una clara reducción del tiempo necesario para el análisis, incluso considerando el tiempo de formación, 3 h y 7 min, que solo se realizó una vez. A la reducción del tiempo de trabajo hay que añadir que el recuento automático no presenta problemas como subjetividad o errores por fatiga de la persona que realiza el recuento manual.

El algoritmo mostró una buena concordancia con el recuento manual salvo en el caso del recuento de pelotas, que se encuentran en cantidades muy pequeñas. Para las

botellas de agua, detergente y lejía, así como para los bidones agrícolas, el algoritmo mostró una buena concordancia con el recuento manual por un observador entrenado.

Con los datos obtenidos se ha podido monitorizar la presencia de residuos flotantes durante todo el periodo de estudio en cada uno de los puntos en los que se han tomado imágenes. Por tanto, con esta metodología se puede conseguir un seguimiento más continuo, dando resultados casi en tiempo real y siendo la única limitación la impuesta por los vuelos UAV para tomar las imágenes.

De los resultados de este estudio se desprende que los objetos presentes en mayor proporción no se pueden asociar exclusivamente a ningún sector de la población, ya que muchos de ellos pueden tener origen doméstico, como las botellas de productos de limpieza y la mayor parte de las botellas de agua. La presencia de este tipo de residuos puede deberse a la dispersión de las viviendas en la zona y a que determinados particulares vierten sus basuras al río o a la red de riego. Entre los residuos, hay una pequeña proporción de bidones agrícolas a pesar de la obligación de eliminarlos mediante un sistema de gestión integral de contenedores. Este recuento diferenciado ayuda a determinar el origen de estas acumulaciones en el río y así poder implementar medidas, como campañas de concienciación, para reducir esta problemática. La metodología propuesta en este trabajo puede ser una herramienta útil por su rapidez y objetividad en la evaluación de políticas públicas dirigidas a eliminar la contaminación plástica en las redes de riego tradicionales.

5.1.2 Modelo de detección de racimos

Los modelos de detección de objetos basados en YOLOv7x mostraron una precisión moderada en la detección de racimos de uvas, presentando valores aceptables en métricas como mAP y puntuación F1. Sin embargo, es necesario reducir los falsos positivos y mejorar la localización precisa de los racimos. Esto sugiere que, aunque el modelo detecta la mayoría de los racimos de uvas presentes en las imágenes, comete errores al detectar falsos positivos. En comparación con las versiones YOLOv5s y YOLOv3-spp, YOLOv7x demostró una mejor precisión (Wu et al, 2022); también en la detección de racimos YOLOv7x demostró una mayor consistencia que las versiones YOLOv7-E6 y YOLOv8 (Codes et al., 2024). La combinación de imágenes de UAV con algoritmos avanzados, como YOLOv7x, mejoró la detección en comparación con el uso exclusivo de datos satelitales, que carecen de resolución espacial para estimaciones de racimos de uva. De hecho, los tamaños de píxel de las imágenes satelitales consideradas (Sentinel-2 y PlanetScope), no solo tienen en cuenta la presencia de racimos de uvas, sino que incluyen todas las características incluidas dentro de los píxeles de 10 m y 3 m. En este sentido, la utilización de imágenes satelitales con una resolución espacial mayor, como las obtenidas de los satélites WorldView o incluso de UAV, podría permitir una identificación más precisa de los elementos de las imágenes. Además, el objetivo principal de los índices de vegetación era representar el vigor y la salud de la vegetación, en lugar de indicar directamente los parámetros de rendimiento.

A pesar del buen rendimiento general de YOLOv7x, su capacidad de detección está limitada por falsos positivos y dificultades para identificar racimos superpuestos o en condiciones de iluminación adversas. Una posible mejora sería el uso de técnicas de preprocesamiento de imágenes, como la normalización de la iluminación, y el aumento del conjunto de datos para el entrenamiento del algoritmo. Los resultados obtenidos en este estudio ponen de relieve la importancia de combinar diversas fuentes de datos de

teledetección para optimizar la detección de racimos de uvas en viñedos, contribuyendo así a la agricultura de precisión. La comparación entre las imágenes de UAV, Sentinel-2 y PlanetScope muestra que, aunque los UAV proporcionan una resolución espacial superior, los datos satelitales de Sentinel-2 y PlanetScope ofrecen ventajas complementarias en términos de estabilidad radiométrica y cobertura temporal.

En concreto, los UAV destacan por su capacidad para detectar detalles finos, como racimos de uvas, gracias a su alta resolución espacial, lo que resulta crucial para parcelas pequeñas y estudios detallados (Ortuani et al., 2024; Khaliq et al., 2019; Matese et al., 2015; Di Gennaro et al., 2022). Sin embargo, su menor estabilidad radiométrica y los costes asociados limitan su aplicación a escalas mayores o estudios prolongados. Por otro lado, Sentinel-2, con su mayor estabilidad radiométrica y frecuencia de revisita, es más adecuado para estudios a gran escala y a largo plazo, como lo demuestran investigaciones anteriores que destacan su eficiencia en el seguimiento del crecimiento y la salud de los cultivos en áreas más grandes (Farmonov et al., 2023; Lucero et al., 2024; Wang et al., 2024).

La inclusión de imágenes PlanetScope, con una resolución temporal alta y una resolución espacial intermedia, permitió capturar variaciones detalladas dentro de las parcelas. Sin embargo, se observó una mayor variabilidad en los valores de reflectancia, como señalan Manivasagam et al. (2021) y Huang et al. (2019). Los resultados superiores de PlanetScope se deben tanto a su mayor resolución geométrica, que se alinea mejor con las áreas de estudio, como a su capacidad para ajustar la pureza de píxeles, que mide el grado de homogeneidad con respecto al cultivo objetivo (Lucero et al., 2024). Esto sugiere que las imágenes PlanetScope podrían ser más útiles en estudios que requieren observaciones detalladas a escalas pequeñas, donde se necesitan ajustes radiométricos precisos (Lucero et al., 2024; Wang et al., 2024). En general, la combinación sinérgica de datos PlanetScope de alta resolución con datos Sentinel-2 de resolución media supera el rendimiento predictivo de cada conjunto de datos por separado, debido a la combinación de información espacial y espectral de ambas fuentes (Wang et al., 2024; Manivasagam et al., 2021; Huang et al., 2019; Zhou et al., 2021). Por lo tanto, se recomienda el uso de modelos multiescala para optimizar las predicciones.

En cuanto a la estimación del rendimiento, se observaron valores de correlación mayores en los datos de Sentinel-2 en comparación con PlanetScope. Aunque los datos de PlanetScope tuvieron una resolución temporal y espacial más alta, su resolución radiométrica fue más baja que la de Sentinel-2. La falta de dos bandas de borde rojo y SWIR adicionales en los datos de PlanetScope también puede ser un factor que influya en estos resultados. Sin embargo, la frecuencia de imágenes casi diaria de PlanetScope brinda oportunidades para mejorar la precisión de la predicción y promover la agricultura digital para el modelado de cultivos (Ziliani et al., 2021). Actualmente, las imágenes de Sentinel-2 se consideran las más adecuadas para la estimación del rendimiento a nivel de campo (Li et al., 2022). La mayor desviación estándar en PlanetScope sugiere que su mayor resolución espacial le permite capturar diferencias más finas en áreas con vegetación menos densa o mezclada con suelo desnudo. Por el contrario, Sentinel, con su resolución más baja pero más uniforme, ofrece resultados más consistentes para estudios a gran escala.

Los índices de vegetación utilizados, como NDVI, SAVI, EVI, GNDVI, MSAVI, MTVI, NDWI, OSAVI y SR, están diseñados para evaluar el vigor de la vegetación y la

cobertura general, pero no capturan detalles específicos. Los resultados de este estudio mostraron un aumento en los índices en otoño, posiblemente relacionado con el crecimiento de malezas en el suelo, en consonancia con otros estudios (Ortuani et al., 2024; Rouault et al., 2024a), donde este aumento también está relacionado con el consumo de agua del huerto dependiendo de las prácticas de riego y las condiciones pedoclimáticas (Rouault et al., 2024b; Elfarkh et al., 2023; Padilla-Díaz et al., 2018).

En estudios relacionados, se utilizaron imágenes terrestres para la detección de racimos de uva (Wang et al., 2024), mientras que otros utilizaron imágenes de UAV y NDVI para identificar vides vivas y secas (Gravilović et al., 2024). Algunos trabajos también compararon el desempeño de diferentes fuentes de imágenes satelitales, como WorldView, PlanetScope y Sentinel-2, concluyendo que las imágenes de alta resolución, como WorldView (es decir, <1 m), son esenciales para detectar características pequeñas (Rodenhizer et al., 2024). Esto podría explicar por qué ni las imágenes de Sentinel ni PlanetScope han mostrado ninguna relación con el número de racimos de uva en el presente estudio. Otros estudios se centraron en la predicción del rendimiento utilizando índices de vegetación, pero sin realizar la detección de la fruta (Sun et al., 2017). Otros autores han reportado esta limitación de los índices de vegetación (como NDVI) para derivar parámetros de rendimiento. En particular, Sun et al. (2017) y Guan et al. (2019) observaron que los índices de vegetación (por ejemplo, NDVI, EVI y SAVI) mostraron inconsistencia en la predicción del rendimiento dependiendo de factores como la cantidad de datos de referencia, el área de estudio, el clima, la cobertura de nubes, las condiciones atmosféricas y la calibración inadecuada.

Los resultados obtenidos demuestran que, si bien PlanetScope ofrece una resolución espacial más alta, los valores de reflectancia de Sentinel-2 tienen una mayor consistencia, lo que podría traducirse en una mejor capacidad para estudios a largo plazo o de áreas grandes. Por otro lado, PlanetScope, debido a su mayor variabilidad en bandas, podría ser más adecuado para aplicaciones que requieren una observación detallada de áreas más pequeñas, pero con una mayor necesidad de ajuste radiométrico.

Esta diferencia en bandas es crítica a la hora de seleccionar qué plataforma utilizar en función del objetivo del estudio. Sentinel-2 es ideal para una monitorización más homogénea y a largo plazo, mientras que PlanetScope es preferible cuando se requiere una resolución espacial muy fina y un análisis detallado de pequeñas parcelas agrícolas.

En resumen, este trabajo presenta un enfoque innovador para la detección automática de racimos de uva mediante la combinación del modelo YOLO y técnicas de teledetección. La capacidad de geoposicionar con precisión los racimos detectados permiten mapear la distribución espacial de la producción dentro de una parcela. Esta información es esencial para identificar la variabilidad intraparcularia, facilitando la implementación de prácticas de agricultura de precisión. Al comprender las diferencias en el rendimiento y las condiciones de crecimiento en distintas áreas del viñedo, los viticultores pueden ajustar de manera óptima las prácticas de manejo, como la aplicación de riego, fertilizantes y control de plagas, mejorando la eficiencia y sostenibilidad de la producción vitivinícola.

6 Conclusion / Conclusión

This work addresses two innovative applications of remote sensing and deep learning in precision agriculture: the detection of floating debris in traditional irrigation systems and the automatic identification of bagged grape clusters in vineyards. Both studies aim to integrate advanced technologies, such as UAVs and artificial intelligence models, to optimize processes and enhance agricultural sustainability.

For floating debris detection, an algorithm based on YOLOv5s was developed and trained to identify and classify different types of debris with remarkable accuracy (mAP@0.5 of 96.9%) and high processing speed (81.7 ms per image). By applying this system to aerial images obtained through UAVs, the analysis becomes faster and more objective, providing precise data to monitor and mitigate plastic pollution. Furthermore, the algorithm's ability to distinguish between types of debris offers valuable information about their origins, making it a useful tool for environmental evaluation and management.

In the case of bagged grape cluster detection, a comparison was conducted between UAV images and satellite data from Sentinel-2 and PlanetScope, using YOLOv7x models to evaluate performance. UAVs demonstrated a clear advantage due to their superior spatial resolution, essential for accurately identifying grape clusters. However, the study also revealed the limitations of traditional vegetation indices for this purpose, as they are designed to assess plant vigor rather than detect specific objects like bagged grape clusters. Additionally, factors such as vine pruning contribute to the lack of correlation, as pruning alters the plant's visible structure, which can confuse indices designed for general canopy assessments.

Among the YOLO model versions compared (YOLOv7x, YOLOv7-E6, and YOLOv8), the YOLOv7x version showed greater precision and consistency, establishing itself as the most suitable option for these specific tasks.

The combination of UAV images with satellite data and advanced deep learning models provides a robust approach to vineyard management, improving productivity and promoting sustainable agricultural practices. Both studies highlight the transformative potential of integrating remote sensing technology and artificial intelligence to address key challenges in agriculture. From reducing plastic pollution in irrigation systems to enhancing vineyard management efficiency, these solutions deliver precision, speed, and sustainability to agricultural operations.

The results also underscore the adaptability of YOLO-based models for automating complex tasks in various applications. While the floating debris detection model achieved outstanding performance, the grape cluster detection model demonstrated the need for further refinement, particularly in minimizing false positives and handling challenging environmental conditions.

The social contributions of these studies should be highlighted. In the case of the detection of plastic waste, the benefits include the reduction of plastics in the environment and microplastic pollution, the improvement of the landscape, the reduction of the impact on native fauna that inhabits rivers and nearby wetlands, as well as a better social perception of agriculture thanks to a cleaner environment.

On the other hand, in the case of grape cultivation, these studies contribute to reducing the need for inputs, improving the competitiveness of agricultural farms and optimising their economic results. This, in turn, encourages the continuity of work in the agricultural sector, helping to prevent the abandonment of rural areas.

Finally, the grape cluster detection study highlights the limitations of relying solely on vegetation indices to predict specific attributes such as cluster presence. It emphasizes the importance of integrating complementary technologies, such as LiDAR and high-resolution multispectral data. In the future, combining diverse data sources will not only improve object detection but also help address broader challenges in precision agriculture and environmental management.

Based on the findings of this study, the proposed future research directions are outlined as follows:

- To evaluate the impact of newer YOLO versions on model performance by comparing metrics.
- To use imaging beyond the RGB spectrum, such as thermal, multispectral, and hyperspectral images, to obtain additional information about the objects of interest. This is applicable to waste detection and fruit detection on trees.

In the field of floating debris detection:

- To incorporate new types of debris, such as plant residues, to expand the scope of the detection tool.
- To explore the identification of different types of plastics using hyperspectral imaging.

In the field of precision agriculture:

- To increase the number of studied plots to strengthen results to refine a harvest estimation model that can be generalized to the entire production area.
- To extend the methodology to other crops, such as figs and pomegranates, addressing specific challenges in automatic fruit detection. This approach should include parameters related to fruit maturity or quality.
- For fruit detection, to include parameters related to fruit maturity or quality.
- To incorporate LiDAR data to obtain three-dimensional information about the canopy and clusters, improving detection precision, and to evaluate the combination of LiDAR data, multispectral images, and advanced object detection models for automatic yield estimation in vineyards.

Conclusión

Este trabajo aborda dos aplicaciones innovadoras de la teledetección y el aprendizaje profundo en la agricultura : la detección de residuos flotantes en sistemas de riego tradicionales y la identificación automática de racimos de uva embolsados en viñedos. Ambos estudios tienen como objetivo integrar tecnologías avanzadas, como vehículos aéreos no tripulados y modelos de inteligencia artificial, para optimizar los procesos y mejorar la sostenibilidad agrícola.

Para la detección de residuos flotantes, se desarrolló y entrenó un algoritmo basado en YOLOv5s para identificar y clasificar diferentes tipos de residuos con una precisión notable ($mAP@0,5$ de 96,9%) y una alta velocidad de procesamiento (81,7 ms por imagen). Al aplicar este sistema a imágenes aéreas obtenidas a través de vehículos aéreos no tripulados, el análisis se vuelve más rápido y objetivo, proporcionando datos precisos para monitorizar y mitigar la contaminación plástica. Además, la capacidad del algoritmo para distinguir entre tipos de residuos ofrece información valiosa sobre sus orígenes, lo que lo convierte en una herramienta útil para la evaluación y gestión ambiental.

En el caso de la detección de racimos de uvas embolsados, se realizó una comparación entre imágenes de UAV y datos satelitales de Sentinel-2 y PlanetScope, utilizando modelos YOLOv7x para evaluar el rendimiento. Los UAV demostraron una clara ventaja debido a su resolución espacial superior, esencial para identificar con precisión los racimos de uvas. Sin embargo, el estudio también reveló las limitaciones de los índices de vegetación tradicionales para este propósito, ya que están diseñados para evaluar el vigor de la planta en lugar de detectar objetos específicos como racimos de uvas embolsados. Además, factores como la poda de la vid contribuyen a la falta de correlación, ya que la poda altera la estructura visible de la planta, lo que puede confundir los índices diseñados para evaluaciones generales del dosel.

Entre las versiones del modelo YOLO comparadas (YOLOv7x, YOLOv7-E6 y YOLOv8), la versión YOLOv7x mostró una mayor precisión y consistencia, estableciéndose como la opción más adecuada para estas tareas específicas.

La combinación de imágenes de UAV con datos satelitales y modelos avanzados de aprendizaje profundo proporciona un enfoque sólido para la gestión de viñedos, mejorando la productividad y promoviendo prácticas agrícolas sostenibles. Ambos estudios destacan el potencial transformador de la integración de la tecnología de teledetección y la inteligencia artificial para abordar los desafíos clave de la agricultura. Desde la reducción de la contaminación plástica en los sistemas de riego hasta la mejora de la eficiencia de la gestión de los viñedos, estas soluciones aportan precisión, velocidad y sostenibilidad a las operaciones agrícolas.

Los resultados también subrayan la adaptabilidad de los modelos basados en YOLO para automatizar tareas complejas en diversas aplicaciones. Si bien el modelo de detección de residuos flotantes logró un rendimiento sobresaliente, el modelo de detección de racimos de uvas demostró la necesidad de un mayor refinamiento, en particular para minimizar los falsos positivos y manejar condiciones ambientales desafiantes.

Es necesario destacar las contribuciones sociales de estos estudios. En el caso de la detección de residuos plásticos, los beneficios incluyen la reducción de plásticos en el medio ambiente y la contaminación por microplásticos, la mejora del paisaje, la

disminución del impacto en la fauna autóctona que habita en ríos y zonas húmedas cercanas, así como una mejor percepción social de la agricultura gracias a un entorno más limpio.

Por otro lado, en el caso del cultivo de la uva, estos estudios contribuyen a reducir la necesidad de insumos, mejorar la competitividad de las explotaciones agrarias y optimizar sus resultados económicos. Esto, a su vez, fomenta la continuidad del trabajo en el sector agrícola, ayudando a evitar el abandono de las zonas rurales.

Por último, el estudio de detección de racimos de uvas pone de manifiesto las limitaciones de confiar únicamente en los índices de vegetación para predecir atributos específicos como la presencia de racimos. Destaca la importancia de integrar tecnologías complementarias, como el LiDAR y los datos multiespectrales de alta resolución. En el futuro, la combinación de diversas fuentes de datos no solo mejorará la detección de objetos, sino que también ayudará a abordar desafíos más amplios en la agricultura de precisión y la gestión ambiental.

Tomando como base los resultados de este estudio, se proponen las siguientes líneas de investigación futuras:

- Evaluar el impacto de nuevas versiones de YOLO en el rendimiento del modelo a través de la comparación de métricas. Utilizar versiones más nuevas de YOLO, comparar sus métricas y explorar si mejoran el rendimiento del modelo.
- Utilizar imágenes más allá del espectro RGB, como imágenes térmicas, multiespectrales e hiperespectrales, para obtener información adicional sobre los objetos de interés. Esto es aplicable a la detección de desechos y detección de frutos en árboles.

En el ámbito de la detección de desechos flotantes:

- Incorporar nuevos tipos de desechos, como residuos vegetales, para ampliar el alcance de la herramienta de detección.
- Explorar la identificación de diferentes tipos de plásticos mediante imágenes hiperespectrales.

En el ámbito de la agricultura de precisión:

- Aumentar el número de parcelas estudiadas para fortalecer los resultados y refinar un modelo de estimación de cosecha que se pueda generalizar a toda el área de producción.
- Extender la metodología a otros cultivos, como higuera y granado, abordando desafíos específicos en la detección automática de frutos. Este enfoque puede incluir parámetros relacionados con la madurez o calidad de los frutos.
- Incorporar datos LiDAR para obtener información tridimensional sobre el dosel vegetal la copa y los racimos, mejorando la precisión de detección, y evaluar la combinación de datos LiDAR, imágenes multiespectrales y modelos avanzados de detección de objetos para la estimación automática del rendimiento en viñedos.

7 Referencias

- Anastasiou, E., Balafoutis, A., Darra, N., Psiroukis, V., Biniari, A., Xanthopoulos, G., & Fountas, S. (2018). Satellite and Proximal Sensing to Estimate the Yield and Quality of Table Grapes. *Agriculture*, 8(7), 94. <https://doi.org/10.3390/agriculture8070094>
- Andrade, M. C., Winemiller, K. O., Barbosa, P. S., Fortunati, A., Chelazzi, D., Cincinelli, A., & Giarrizzo, T. (2019). First account of plastic pollution impacting freshwater fishes in the Amazon: Ingestion of plastic debris by piranhas and other serrasalmids with diverse feeding habits. *Environmental Pollution*, 244, 766-773.
- Apolo-Apolo, O.E., Martínez-Guanter, J., Egea, G., Raja, P., Pérez-Ruiz, M. (2020). Deep learning techniques for estimation of the yield and size of citrus using a UAV. *European Journal of Agronomy*. 115, 126030.
- Badeka, E., Karapatzak, E., Karampatea, A., Bouloumpasi, E., Kalathas, I., Lytridis, C., Tziolas, E., Tsakalidou, V. N., & Kaburlasos, V. G. (2023). A deep learning approach for precision viticulture: Assessing grape maturity via YOLOv7. *Sensors*, 23, 8126. <https://doi.org/10.3390/s23198126>
- Ballesteros, R., Ortega, J. F., Hernández, D., & Moreno, M. Á. (2015). Characterization of *Vitis vinifera* L. canopy using unmanned aerial vehicle-based remote sensing and photogrammetry techniques. *American Journal of Enology and Viticulture*, 66(2), 120-129.
- Ballesteros, R.; Intrigliolo, D. S.; Ortega, J. F.; Ramírez-Cuesta, J. M.; Buesa, I.; Moreno, M. A. (2021). Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. *Precision Agriculture* 21, 1242-1262. <https://doi.org/10.1007/s11119-020-09717-3>.
- Bellasi, A., Binda, G., Pozzi, A., Galafassi, S., Volta, P., & Bettinetti, R. (2020) Microplastic contamination in freshwater environments: A review, focusing on interactions with sediments and benthic organisms. *Environments*, 7(4), 30. <https://doi.org/10.3390/environments7040030>
- Bhakta, I.; Phadikar, S.; Majumder, K. (2019). State-of-the-art technologies in precision agriculture: a systematic review. *J. Sci. Food Agric.* 99, 4878–4888. <https://doi.org/10.1002/jsfa.9693>
- Bland, J. M., & Altman, D. G. Statistical methods for assessing agreement between two methods of clinical measurement (1986). *The Lancet*, 327(8476), 307–310. [https://doi.org/10.1016/S0140-6736\(86\)90837-8](https://doi.org/10.1016/S0140-6736(86)90837-8)
- Brownlee, J. (2024) Difference between a batch and an epoch in a neural network. Retrieved July 19, 2024, from <https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/>
- Cai, Z., & Vasconcelos, N. Cascade R-CNN: Delving into high quality object detection (2018). In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/CVPR.2018.00644>
- Codes Alcaraz, A.M.; Puerto Molina, H.; Rocamora Osorio, C.; Furnitto, N.; Failla, S.; Ramírez-Cuesta, J.M. (2024). Deciphering the role of UAV flight height and sensor inclination in the prediction of table grape production through computer vision techniques. In *Proceedings of the IEEE International Workshop on Metrology for Agriculture and Forestry, MetroAgriFor 2024*, Padua, Italy, 29–31 October 2024; pp.

19-23.

- Cogato, A., Meggio, F., Collins, C., & Marinello, F. (2020). Medium-Resolution Multispectral Data from Sentinel-2 to Assess the Damage and the Recovery Time of Late Frost on Vineyards. *Remote Sensing*, 12(11), 1896. <https://doi.org/10.3390/rs12111896>
- CRDOP (2024). Consejo Regulador de la Denominación de Origen Protegida Uva de Mesa Embolsada del Vinalopó. Recuperado el 3 de octubre de 2024, de <https://uva-vinalopo.org>.
- Daponte, P.; Vito, L.; Glielmo, L.; Iannelli, L.; Liuzza, D.; Picariello, F.; Silano, G. (2019). A review on the use of drones for precision agriculture. *Terrestrial Environ. Sci.* 275.
- De Vries, R., Egger, M., Mani, T., & Lebreton, L. (2021). Quantifying floating plastic debris at sea using vessel-based optical data and artificial intelligence. *Remote Sensing*, 13(17). <https://doi.org/10.3390/rs13173401>
- Derraik, J. G. B. (2002). The pollution of the marine environment by plastic debris: A review. *Marine Pollution Bulletin*, 44(9), 842–852. [https://doi.org/10.1016/S0025-326X\(02\)00220-5](https://doi.org/10.1016/S0025-326X(02)00220-5)
- Di Gennaro, S. F., Dainelli, R., Palliotti, A., Toscano, P., & Matese, A. (2019). Sentinel-2 validation for spatial variability assessment in overhead trellis system viticulture versus UAV and agronomic data. *Remote Sensing*, 11(21), 2573.
- Dong, X.; Lan, J.; Zhu, Y. (2004). Method for identifying the collection position of lychee based on YOLOv5 and PSPNet. *Remote Sens.* 14.
- Diago, M. P., Fernandez-Novales, J., Fernandes, A. M., Melo-Pinto, P., & Tardaguila, J. (2016). Use of visible and short-wave near-infrared hyperspectral imaging to fingerprint anthocyanins in intact grape berries. *Journal of agricultural and food chemistry*, 64(40), 7658-7666.
- Elfarkh, J., Johansen, K., El Hajj, M. M., Almashharawi, S. K., & McCabe, M. F. (2023). Evapotranspiration, gross primary productivity and water use efficiency over a high-density olive orchard using ground and satellite based data. *Agricultural Water Management*, 287, 108423.
- Farmonov, N., Amankulova, K., Szatmári, J., Urinov, J., Narmanov, Z., Nosirov, J., & Mucsi, L. (2023). Combining PlanetScope and Sentinel-2 images with environmental data for improved wheat yield estimation. *International Journal of Digital Earth*, 16(1), 847-867.
- Gavrilović, M.; Jovanović, D.; Božović, P.; Benka, P.; Govedarica, M. (2024). Vineyard Zoning and Vine Detection Using Machine Learning in Unmanned Aerial Vehicle Imagery. *Remote Sens.* 16, 584. <https://doi.org/10.3390/rs16030584>
- General Directorate of Rural Development and Common Agricultural Policy. (2015) Specifications of the Protected Designation of Origin Vinalopó Bagged Table Grapes. Ciutat Administrativa 9 d'Octubre, Castán Tobeñas, Valencia.
- Ghahremani A, Adams SD, Norton M, Khoo SY, Kouzani AZ. (2025). Detecting Defects in Solar Panels Using the YOLO v10 and v11 Algorithms. *Electronics*. 14(2):344. <https://doi.org/10.3390/electronics14020344>
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation (2014). In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

<https://doi.org/10.1109/CVPR.2014.81>

- Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* 58, 289–298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7).
- González-Caballero, V., Sánchez, M.T., Fernández-Novales, J., López, M.I., Pérez-Marín, D. (2012). On-vine monitoring of grape ripening using near-infrared spectroscopy. *Food Anal Methods*, 5, pp. 1377-1385
- Guan, S.; Fukami, K.; Matsunaka, H.; Okami, M.; Tanaka, R.; Nakano, H.; Sakai, T.; Nakano, K.; Ohdan, H.; Takahashi, K. (2019). Assessing Correlation of High-Resolution NDVI with Fertilizer Application Level and Yield of Rice and Wheat Crops Using Small UAVs. *Remote Sens.* 11, 112.
- Guo, J.; Li, X.; Li, Z.; Hu, L.; Yang, G.; Zhao, C.; Fairbairn, D.; Watson, D.; Ge, M. (2018). Multi-GNSS precise point positioning for precision agriculture. *Precis. Agric.* 19, 895–911. <https://doi.org/10.1007/s11119-018-9563-8>
- Gutiérrez, S., Diago, M. P., Fernández-Novales, J., & Tardaguila, J. (2018). Vineyard water status assessment using on-the-go thermal imaging and machine learning. *PLoS One*, 13(2), e0192037.
- Gutiérrez-Gordillo, S., de la Gala González-Santiago, J., Trigo-Córdoba, E., Rubio-Casal, A. E., García-Tejero, I. F., & Egea, G. 2021. Monitoring of emerging water stress situations by thermal and vegetation indices in different almond cultivars. *Agronomy*, 11(7), 1419.
- GVA. (2015). Padrón Municipal de Habitantes. Explotación extensa. Año 2015. Conselleria de Economía Sostenible, Sectores Productivos, Comercio y Trabajo. En <http://www.ive.es/>
- Haboudane, D. (2004). Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* 90, 337–352. <https://doi.org/10.1016/j.rse.2003.12.013>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(9). <https://doi.org/10.1109/TPAMI.2015.2389824>
- Huang, J., Ma, H., Sedano, F., Lewis, P., Liang, S., Wu, Q., Su, W., Zhang, X., & Zhu, D. (2019). Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST–PROSAIL model. *European journal of agronomy*, 102, 1-13.
- Huete, A.R. (1988). A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* 25, 295–309. [https://doi.org/10.1016/0034-4257\(88\)90106-X](https://doi.org/10.1016/0034-4257(88)90106-X).
- Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* 59, 440–451. [https://doi.org/10.1016/S0034-4257\(96\)00112-5](https://doi.org/10.1016/S0034-4257(96)00112-5).
- Hurley, R.; Braaten, H.F.V.; Nizzetto, L.; Steindal, E.H.; Lin, Y.; Clayer, F., Van Emmerik, T., Buenaventura, N. T., Eidsvoll, D. P., Økelsrud, A., Norling, M., Adam, H. N., & Olsen, M. (2023). Measuring riverine macroplastic: Methods, harmonisation, and quality control. *Water Research*, 235, 119902. <https://doi.org/10.1016/j.watres.2023.119902>

- Inbar, O., Shahar, M., Gidron, J., Cohen, I., Menashe, O., & Avisar, D. (2023). Analyzing the secondary wastewater-treatment process using Faster R-CNN and YOLOv5 object detection algorithms. *Journal of Cleaner Production*, 416. <https://doi.org/10.1016/j.jclepro.2023.137913>
- Inobeme, A., Tadso, J., Adetuyi, B.O., Adetunji, C.O., Popoola, O.A., Olaitan, F.Y., Akinbo, O., Shahnawaz, M., Oyewole, O.A., Eniola, K.I.T., Yerima, M.B., Kolawole, O.M. (2024). Toxicity of Microplastics on the Flora and Fauna of the Freshwater Ecosystem: An Overview. In: Shahnawaz, M., Adetunji, C.O., Dar, M.A., Zhu, D. (eds) *Microplastic Pollution*. Springer, Singapore. https://doi.org/10.1007/978-981-99-8357-5_24
- Íñiguez R, Gutiérrez S, Poblete-Echeverría C, Hernández I, Barrio I, Tardáguila J. (2024). Deep learning modelling for non-invasive grape bunch detection under diverse occlusion conditions. *Computers And Electronics In Agriculture*. 226:109421. doi:10.1016/j.compag.2024.109421.
- Jaafar, H.H.; Zurayk, R.; King, C.; Ahmad, F.; Al-Outa, R. (2015). Impact of the Syrian Conflict on Irrigated Agriculture in the Orontes Basin. *Int. J. Water Resour. Dev.* 31, 436–449
- Jiménez, A.; Ceres, R.; Pons, J.L. (2000). A study of computer vision methods for locating fruit on trees. *Trans. ASAE* 43, 1911-1920.
- Jordan, C.F. (1969), Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50, 663–666. <https://doi.org/10.2307/1936256>.
- Junzhe, Z., Fuqiang, J., Yupeng, C., Weiyi, W., & Qing, W. (2023). A water surface garbage recognition method based on transfer learning and image enhancement. *Results in Engineering*, 19, 101340. <https://doi.org/10.1016/j.rineng.2023.101340>
- Karkauskaite, P.; Tagesson, T.; Fensholt, R. (2017). Evaluation of the Plant Phenology Index (PPI), NDVI and EVI for Start-of-Season Trend Analysis of the Northern Hemisphere Boreal Zone. *Remote Sens.* 9, 485.
- Kendall, H.; Clark, B.; Li, W.; Jin, S.; Jones, G.D.; Chen, J.; Taylor, J.; Li, Z.; Frewer, L.J. (2022). Precision agriculture technology adoption: a qualitative study of small-scale commercial “family farms” located in the North China Plain. *Precis. Agric.* 23, 319–351. <https://doi.org/10.1007/s11119-021-09839-2>
- Khalik, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chiaberge, M., & Gay, P. (2019). Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sensing*, 11(4), 436.
- Krishna, K.R. (2018). *Agricultural drones: a peaceful search*; CRC Press: Boca Raton, Florida, USA.
- Labus, M.P.; Nielsen, G.A.; Lawrence, R.L.; Engel, R.; Long, D.S. (2002). Wheat Yield Estimates Using Multi-Temporal NDVI Satellite Imagery. *Int. J. Remote Sens.* 23, 4169–4180
- LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. <https://doi.org/10.1038/nature14539>
- Leroux, C.; Jones, H.; Pichon, L.; Guillaume, S.; Lamour, J.; Taylor, J.; Naud, O.; Crestey, T.; Lablee, J.-L.; Tisseyre, B. (2018). GeoFIS: An open-source decision-support tool for precision agriculture data. *Agriculture* 8, 73. <https://doi.org/10.3390/agriculture8060073>

- Li, F., Miao, Y., Chen, X., Sun, Z., Stueve, K., & Yuan, F. (2022). In-season prediction of corn grain yield through PlanetScope and Sentinel-2 images. *Agronomy*, 12(12), 3176.
- Li, Y., Wang, R., Gao, D., & Liu, Z. (2023). A floating-waste-detection method for unmanned surface vehicle based on feature fusion and enhancement. *Journal of Marine Science and Engineering*, 11(12). <https://doi.org/10.3390/jmse11122234>
- Lin, F., Hou, T., Jin, Q., & You, A. (2021). Improved YOLO-based detection algorithm for floating debris in waterway. *Entropy*, 23(9). <https://doi.org/10.3390/e23091111>
- Liu, G.; Nouaze, J.T.; Kim, J.P. (2020). A robust algorithm for tomato detection based on YOLOv3. *Sensors* 20(7), 2145; <https://doi.org/10.3390/s20072145>
- Liu, J.; Xiang, J.; Jin, Y.; Liu, R.; Yan, J.; Wang, L. (2021). Boost precision agriculture with unmanned aerial vehicle remote sensing and edge intelligence: A survey. *Remote Sens.* 13, 4387. <https://doi.org/10.3390/rs13214387>
- Lucero, M. F., Hernández, C. M., Carcedo, A. J., Zajdband, A., Guillevic, P. C., Houborg, R., Hamilton, K., & Ciampitti, I. A. (2024). Enhancing alfalfa biomass prediction: An innovative framework using remote sensing data. *Remote Sensing*, 16(18), 3379.
- Manfreda, S., McCabe, M. F., Miller, P. E., Lucas, R., Pajuelo Madrigal, V., Mallinis, G., Ben Dor, E., Helman, D., Estes, L., Ciraolo, G., Müllerová, J., Tauro, F., De Lima, M. I., De Lima, J. L. M. P., Maltese, A., Frances, F., Caylor, K., Kohv, M., Perks, M., Ruiz-Perez, G., Su, Z., Vico, G., Toth, B. (2018). On the Use of Unmanned Aerial Systems for Environmental Monitoring. *Remote Sensing*, 10(4), 641. <https://doi.org/10.3390/rs10040641>
- Manivasagam, V. S., Sadeh, Y., Kaplan, G., Bonfil, D. J., & Rozenstein, O. (2021). Studying the feasibility of assimilating Sentinel-2 and PlanetScope imagery into the SAFY crop model to predict within-field wheat yield. *Remote Sensing*, 13(12), 2395.
- Marinello, F.; Pezzuolo, A.; Chiumenti, A.; Sartori, L. (2016). Technical analysis of unmanned aerial vehicles (drones) for agricultural applications. *Eng. Rural Dev.* 15, 870–875.
- Matese, A., Toscano, P., Di Gennaro, S. F., Genesio, L., Vaccari, F. P., Primicerio, J., Belli, C., Zaldei, A., Bianconi, R., & Gioli, B. (2015). Intercomparison of UAV, aircraft and satellite remote sensing platforms for precision viticulture. *Remote sensing*, 7(3), 2971-2990.
- McFeeters, S.K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. <https://doi.org/10.1080/01431169608948714>.
- Melgarejo J, Trapote A, Roca J. (2013). La infraestructura hidráulica y la gestión del agua en los regadíos tradicionales en la Vega Baja del Segura (Alicante). *Revista Pilquen. Sección Agronomía. Año XV. Nº 13*.
- Nielsen, M. A. (2015). *Neural Networks and Deep Learning*. Determination Press. <http://neuralnetworksanddeeplearning.com/index.html>
- Olson, D.L., Delen, D. (2008). Performance Evaluation for Predictive Modeling. In: *Advanced Data Mining Techniques*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-76917-0_9
- ONU - Organización de las Naciones Unidas (2018). *La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe*.

- https://repositorio.cepal.org/bitstream/handle/11362/40155/24/S1801141_es.pdf
- Ortuani, B., Mayer, A., Bianchi, D., Sona, G., Crema, A., Modina, D., D., Bolognini, M., Brancadoro, L., Boschetti, M., & Facchi, A. (2024). Effectiveness of Management Zones Delineated from UAV and Sentinel-2 Data for Precision Viticulture Applications. *Remote Sensing*, 16(4), 635.
- Padilla-Díaz, C. M., Rodriguez-Dominguez, C. M., Hernandez-Santana, V., Perez-Martin, A., Fernandes, R. D. M., Montero, A., García, J.M., Fernández, J. E. (2018). Water status, gas exchange and crop performance in a super high density olive orchard under deficit irrigation scheduled from leaf turgor measurements. *Agricultural Water Management*, 202, 241-252.
- Palacios, F., Bueno, G., Salido, J., Diago, M. P., Hernández, I., & Tardaguila, J. (2020). Automated grapevine flower detection and quantification method based on computer vision and deep learning from on-the-go imaging using a mobile sensing platform under field conditions. *Computers and Electronics in Agriculture*, 178, 105796.
- Payares, A.; Tarquis, A.H.; Peraló, R.; Canó, J.; Cámara, J.; Gómez del Campo, M. (2023). Multispectral and thermal sensors on board UAVs for heterogeneity in Merlot vineyard detection: Contribution to zoning maps. *Remote Sensing*, 15(16), 4024. <https://doi.org/10.3390/rs15164024>
- Payne, A.B.; Walsh, K.B.; Subedi, P.P.; Jarvis, D. (2013). Estimation of mango crop yield using image analysis – Segmentation method. *Comput. Electron. Agric.* 91, 57–64. <https://doi.org/10.1016/j.compag.2012.11.009>
- Pérez-Zavala, R.; Torres-Torriti, M.; Cheein, F.A.; Troni, G. (2018). A pattern recognition strategy for the visual detection of grape clusters in vineyards. *Comput. Electron. Agric.* 2018, 151, 136–149. <https://doi.org/10.1016/j.compag.2018.06.001>
- Puerto Molina H, Rocamora Osorio C. (2019). Nuevos regadíos en la provincia de Alicante. En *Canelobre: Revista del Instituto Alicantino de Cultura "Juan Gil-Albert"*, Nº 70: El agua en la provincia de Alicante. Territorio, patrimonio e innovación, págs. 228-239.
- Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. (1994). A Modified Soil Adjusted Vegetation Index. *Remote Sens. Environ.* 48, 119–126. [https://doi.org/10.1016/0034-4257\(94\)90134-1](https://doi.org/10.1016/0034-4257(94)90134-1).
- Ramírez-Cuesta, J.M., Intrigliolo, D.S., Lorite, I.J., Moreno, M.A., Vanella, D., (2023). Determining grapevine water use under different sustainable agronomic practices using METRIC-UAV surface energy balance model 281. <https://doi.org/10.1016/j.agwat.2023.108247>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 26 June–1 July 2016. <https://doi.org/10.1109/CVPR.2016.91>
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- Robson, A.; Rahman, M.M.; Muir, J. (2017). Using worldview satellite imagery to map yield in avocado (*Persea americana*): A case study in Bundaberg, Australia. *Remote Sens.* 9, 1223. <https://doi.org/10.3390/rs9121223>

- Rocamora, C., Puerto, H., Abadía, R., Brugarolas, M., Martínez-Carrasco, L., & Cordero, J. (2021). Floating debris in the Low Segura River Basin (Spain): Avoiding litter through the irrigation network. *Water*, 13(8), 1074. <https://doi.org/10.3390/w13081074>
- Rodenhizer, H., Yang, Y., Fiske, G., Potter, S., Windholz, T., Mullen, A., Watts, J.D., Rogers, B. M. (2024). A Comparison of Satellite Imagery Sources for Automated Detection of Retrogressive Thaw Slumps. *Remote Sensing*, 16(13), 2361.
- Rondeaux, G.; Steven, M.; Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107. [https://doi.org/10.1016/0034-4257\(95\)00186-7](https://doi.org/10.1016/0034-4257(95)00186-7).
- Rouault, P., Courault, D., Flamain, F., Pouget, G., Doussan, C., Lopez-Lozano, R., McCabe, M., Debolini, M. (2024). High-resolution satellite imagery to assess orchard characteristics impacting water use. *Agricultural Water Management*, 295, 108763.
- Rouault, P., Courault, D., Pouget, G., Flamain, F., Diop, P. K., Desfonds, V., Dousan, C., Chanzy, A., Debolini, M., McCabe, M., Lopez-Lozano, R. (2024). Phenological and biophysical Mediterranean orchard assessment using ground-based methods and Sentinel 2 data. *Remote Sensing*, 16(18), 3393.
- Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of the Third Earth Resources Technology Satellite—1 Symposium*, Washington, DC, USA, 10–14 December 1973; NASA SP-351.
- Sánchez-Rubio, C. J. (2010). La agricultura y el problema del agua en la provincia de Alicante. En: *A Vueltas Con la Agricultura: una actividad económica necesaria y marginada* / coord. por Segrelles J.A. ISBN 978-84-7784-565-2, págs. 105-130.
- Santos, T.T.; de Souza, L.L.; dos Santos, A.A.; Ávila, S. (2020). Detection, segmentation and tracking of grapes using deep neural networks and three-dimensional association. *Comput. Electron. Agric.* 170, 105247.
- Sepúlveda-Reyes, D., Ingram, B., Bardeen, M., Zúñiga, M., Ortega-Farías, S., & Poblete-Echeverría, C. (2016). Selecting Canopy Zones and Thresholding Approaches to Assess Grapevine Water Status by Using Aerial and Ground-Based Thermal Imaging. *Remote Sensing*, 8(10), 822. <https://doi.org/10.3390/rs8100822>
- Shi, R.; Li, T.; Yamaguchi, Y. (2020). An attribution-based pruning method for real-time mango detection with YOLO network. *Comput. Electron. Agric.* 169. <https://doi.org/10.1016/j.compag.2020.105214>
- Soumitra, N. (2024). A vision of precision agriculture: Balance between agricultural sustainability and environmental stewardship. *Agronomy Journal*, 116, 1126–1143. <https://doi.org/10.1002/agj2.21405>
- Sozzi, M., Cantalamessa, S., Cogato, A., Kayad, A., & Marinello, F. (2022). Automatic Bunch Detection in White Grape Varieties Using YOLOv3, YOLOv4, and YOLOv5 Deep Learning Algorithms. *Agronomy*, 12(2), 319. <https://doi.org/10.3390/agronomy12020319>
- Squeri, C., Poni, S., Di Gennaro, S. F., Matese, A., & Gatti, M. (2021). Comparison and Ground Truthing of Different Remote and Proximal Sensing Platforms to Characterize Variability in a Hedgerow-Trained Vineyard. *Remote Sensing*, 13(11), 2056. <https://doi.org/10.3390/rs13112056>
- Stringam, B. L., Gill, T., & Sauer, B. (2016). Integration of irrigation district personnel with

- canal automation projects. *Irrigation Science*, 34(1), 33–40. <https://doi.org/10.1007/s00271-015-0477-1>
- Sun, L., Gao, F., Anderson, M. C., Kustas, W. P., Alsina, M. M., Sanchez, L., Sams, B., McKee, L., Dulaney, W., White, W. A., Alfieri, J. G., Prueger, J. H., Melton, F., & Post, K. (2017). Daily Mapping of 30 m LAI and NDVI for Grape Yield Prediction in California Vineyards. *Remote Sensing*, 9(4), 317. <https://doi.org/10.3390/rs9040317>
- Tardáguila J, Diago MP, Balda P, et al. (2021). Smart applications and digital technologies in viticulture: A review. *Agronomy*. 11(8):1473. doi:10.3390/agronomy11081473
- Tian, Z., Huang, J., Yang, Y., & Nie, W. (2023). KCFS-YOLOv5: A high-precision detection method for object detection in aerial remote sensing images. *Applied Sciences*, 13(1). <https://doi.org/10.3390/app13010649>
- Toscano, F.; Fiorentino, C.; Capece, N. (2024). Unmanned Aerial Vehicle for Precision Agriculture: A Review. in *IEEE Access*, vol. 12, pp. 69188-69205, doi: 10.1109/ACCESS.2024.3401018
- Trapote Jaume, A., Roca Roca, J. F., & Melgarejo Moreno, J. (2015). Azudes y acueductos del sistema de riego tradicional de la Vega Baja del Segura (Alicante, España). *Investigaciones Geográficas*, (63). <https://doi.org/10.14198/ingeo2015.63.10>
- Van Emmerik, T., & Schwarz, A. (2020). Plastic debris in rivers. *WIREs Water*, 7(1), 1–24. <https://doi.org/10.1002/wat2.1398>
- Van Emmerik, T., Mellink, Y., Hauk, R., Waldschl ager, K., & Schreyers, L. (2022). Rivers as plastic reservoirs. *Frontiers in Water*, 3. <https://doi.org/10.3389/frwa.2021.786936>
- Van Lieshout, C., Van Oeveren, K., Van Emmerik, T., & Postma, E. (2020). Automated river plastic monitoring using deep learning and cameras. *Earth and Space Science*, 7(8). <https://doi.org/10.1029/2019EA000960>
- Vannoppen, A., & Gobin, A. (2021). Estimating Farm Wheat Yields from NDVI and Meteorological Data. *Agronomy*, 11(5), 946. <https://doi.org/10.3390/agronomy11050946>
- Vaughn, T., Crookston, B. M., & Pfister, M. (2021). Floating woody debris: Blocking sensitivity of labyrinth weirs in channel and reservoir applications. *Journal of Hydraulic Engineering*, 147(11). [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0001937](https://doi.org/10.1061/(ASCE)HY.1943-7900.0001937)
- Veroustraete, F. The rise of drones in agriculture. *EC Agricult*. 2015, 2, 325–327.
- Wang, W.; Shi, Y.; Liu, W.; Che, Z. (2024). An unstructured orchard grape detection method utilizing YOLOv5s. *Agriculture*14, 262. <https://doi.org/10.3390/agriculture14020262>
- Wang, Y.; Yi, Y.; Wang, X.-F.; Chen, J.; Li, X. (2021). Fig fruit recognition method based on YOLOv4 deep learning. In *Proceedings of the 2021 18th International Conference on Electrical Engineering, Electronics, Computer, Telecommunications and Information Technology: Smart Electrical System and Technology, ECTI-CON, Bangkok, Thailand, 24–26 March 2021*; pp. 303–306. <https://doi.org/10.1109/ECTI-CON51831.2021.9454904>
- Wang, Z., Wu, W., & Liu, H. (2024). Spatial estimation of soil organic carbon content utilizing PlanetScope, Sentinel-2, and Sentinel-1 data. *Remote Sensing*, 16(17), 3268.

- Wolf, M.; van den Berg, K.; Garaba, S.P.; Gnann, N.; Sattler, K.; Stahl, F.; Zielinski, O. (2020). Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q). *Environmental Research Letters*, 15(11). <https://doi.org/10.1088/1748-9326/abbd01>
- Wolfert, S.; Ge, L.; Verdouw, C.; Bogaardt, M.-J. (2017). Big data in smart farming-A review. *Agricultural Systems*, Vol. 153, 69-80. <https://doi.org/10.1016/j.agry.2017.01.023>
- Wu, D., Jiang, S., Zhao, E., Liu, Y., Zhu, H., Wang, W., & Wang, R. (2022). Detection of *Camellia oleifera* fruit in complex scenes by using YOLOv7 and data augmentation. *Applied sciences*, 12(22), 11318.
- Xiong, J.; Lin, R.; Liu, Z.; He, Z.; Tang, L.; Yang, Z.; Zou, X. (2018). The recognition of groups of lychee and the calculation of the collection point in a natural night-time environment. *Biosyst. Eng.* 166, 44–57.
- Yao, J., Qi, J., Zhang, J., Shao, H., Yang, J., & Li, X. (2021). A real-time detection algorithm for kiwifruit defects based on YOLOv5. *Electronics*, 10(14). <https://doi.org/10.3390/electronics10141711>
- Zhang, J.; Jiang, F.; Chen, Y.; Wu, W.; Wu, Q. (2023). A water surface garbage recognition method based on transfer learning and image enhancement. *Results Eng.* 19, 101340. <https://doi.org/10.1016/j.rineng.2023.101340>.
- Zhou, T., Geng, Y., Ji, C., Xu, X., Wang, H., Pan, J., Bumberger, J., Haase, D., & Lausch, A. (2021). Prediction of soil organic carbon and the C: N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. *Science of the Total Environment*, 755, 142661.
- Ziliani, M. G., Altaf, M. U., Aragon, B., Houborg, R., Franz, T. E., Lu, Y., Sheffield, J., Hoteit, I. & McCabe, M. F. (2022). Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model. *Agricultural and forest meteorology*, 313, 108736

ANEXOS

Article

Image Recognition for Floating Waste Monitoring in a Traditional Surface Irrigation System

Ana María Codes-Alcaraz ¹, Herminia Puerto ^{1,2}  and Carmen Rocamora ^{1,2,*} 

¹ CIAGRO-UMH, Centro de Investigación e Innovación Agroalimentaria y Agroambiental, Miguel Hernández University, 03312 Orihuela, Spain; acodes@umh.es (A.M.C.-A.); hpuerto@umh.es (H.P.)

² Escuela Politécnica Superior de Orihuela, Department of Engineering, Miguel Hernández University, 03312 Orihuela, Spain

* Correspondence: rocamora@umh.es

Abstract: In the traditional surface irrigation system of Vega Baja del Segura (Spain), large amounts of floating waste accumulate at certain points of the river, irrigation channels and drainage ditches, causing malfunctioning of the irrigation network and rising social problems related to the origins of waste. This work proposes a standardized and quick methodology to characterize the floating waste to detect changes in its amount and components. A dataset was created with 477 images of floating plastic items in different environments and was used for training an algorithm based on YOLOv5s. The mean Average Precision of the trained algorithm was 96.9%, and the detection speed was 81.7 ms. Overhead photographs were taken with an unmanned aerial vehicle at strategic points of the river and channels, and its automatic count of floating objects was compared with their manual count. Both methods showed good agreement, confirming that water bottles were the most abundant (95%) type of floating waste. The automatic count reduced the required time and eliminated human bias in image analysis of the floating waste. This procedure can be used to test the reach of corrective measures implemented by local authorities to prevent floating waste in the river.

Keywords: YOLO; Python; plastic; waste; waste accumulation; rivers



Citation: Codes-Alcaraz, A.M.; Puerto, H.; Rocamora, C. Image Recognition for Floating Waste Monitoring in a Traditional Surface Irrigation System. *Water* **2024**, *16*, 2680. <https://doi.org/10.3390/w16182680>

Academic Editor: William Frederick Ritter

Received: 25 July 2024

Revised: 12 September 2024

Accepted: 18 September 2024

Published: 20 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The presence of plastic waste in freshwater is an increasing environmental issue. Large debris can obstruct water flow, increasing flood risk in urban areas and affecting local infrastructure, causing economic losses. Plastic debris accumulates on riverbanks, potentially affecting tourism or real estate value. Dams, water inlets, sluices and canals can act as accumulation zones for macroplastics. A significant issue is the aesthetic pollution and loss of recreational value of these water bodies. Visible waste disrupts perception of rivers and lakes, reducing their appeal for recreational activities such as swimming, fishing and tourism, which can have an adverse economic impact on local communities [1–3].

Plastic waste in the aquatic ecosystem nowadays is a great concern due to its large amount in rivers and oceans [4] and its persistence in them [3]. Several studies have established that the main source of marine plastic pollution is land-based, due to littering, plastic bag usage and solid waste disposal [5], and rivers are the link connecting land surfaces to the sea or oceans. Since there is evidence that not all the plastic that enters a river is released into the ocean [3], there is growing interest in characterizing the flow of macroplastic residues in rivers to identify and design measures to reduce global plastic pollution [6].

Surface irrigation systems in traditionally irrigated river plains spread the water flow network, increasing the likelihood of land-based waste reaching the oceans. In these surface irrigation systems, the effect of floating woody debris in channels and reservoirs has been studied because debris blocks the water flow and increases reference head on flow control structures [7]. Automated trash racks removal systems have been used to protect automated

gates in canal automation projects [8], but there are no mentions of plastic floating waste occurrence in irrigation channels in the scientific literature.

In the Segura River and its traditional surface irrigation system of Vega Baja del Segura, in southeast Spain, where the traditional surface irrigation system dating from the 13th century is still in use, a large amount of floating waste accumulates at certain points, causing problems in the distribution of irrigation in channels and drainage ditches [9]. The floating debris that is found in the river and irrigation network has varied origins, predominantly plant waste from the river itself, such as reeds, as well as waste of anthropogenic origin [10]. The specific land-based sources of plastic waste in the Segura River include domestic activities, agricultural practices and industrial processes. Among these, plastic bottles of bottled water are the most significant contributor, comprising 74.2% of the waste identified. Containers of domestic origin, such as food containers and hygiene products, also play a significant role, alongside agricultural waste such as bottles and cans of phytosanitary products [9,10].

Besides impairment to the functioning of infrastructure, floating waste causes economic damage to farmers because they must face the cost of removal of floating waste, or the fines imposed by the river authority if they do not keep their channels clean and allow the discharge of floating material into the river [9].

Awareness campaigns aimed at different sectors of the population have been proposed by local authorities as an effective measure to prevent plastic waste from reaching the Segura River. However, for these campaigns to be effective, it is necessary to define the sector of the population to which they should be addressed. Knowing the type of plastic found in the irrigation infrastructure channels can disclose whether these wastes have a predominantly agricultural origin, and the campaigns should then be addressed to farmers, or are due to daily activities of the general population.

Quantifying and characterizing floating debris has been attempted using different approaches such as the development of software to quantify large amounts of waste in the Pacific ocean [11], a deep learning-based methodology to detect floating macroplastics on river surface images in Jakarta [12] and an algorithm based on convolutional neural networks (CNN), APLASTIC-Q, to classify (PLD-CNN) and quantify (PLQ-CNN) litter items in classes like water bottles, Styrofoam, canisters, cartons, bowls, shoes, polystyrene packaging, cups, textile and small or large carry bags in beaches and riverine systems in Cambodia [13]. Recent works have already used the detection of plastics in water with YOLO [14] and compare diverse configurations of YOLOv5 to detect bottles and cartons from images taken with unmanned surface vehicles [14,15].

The advances in deep-learning methods have boosted research in object detection, giving rise to a wide range of approaches based on deep neural networks. Traditional two-stage methods, such as region-based convolutional neural network (R-CNN) serial algorithms [16–18], are therefore distinguished from representative single-stage methods, such as YOLO serial algorithms [14,15,19].

The introduction of the convolutional neural network (R-CNN) in 2014 by Girshick et al. [20] marked a milestone in the development of deep learning-based object detection methods. Although methods like R-CNN have demonstrated high accuracy, they are perceived as slow [21]. Unlike two-stage methods, YOLO does not need to generate candidate regions; instead, it divides the image into grids, each with bounding boxes, to detect objects in a single step, thus streamlining the process [14]. In a study on garbage detection in water, Faster R-CNN showed unsatisfactory performance due to its accuracy loss compared to that of YOLOv5 [14]. Currently, YOLOv5 is the first widely adopted single-stage detector that outperforms two-stage detectors [15]. The better performance of YOLOv5 has been documented in challenging environments like the automatic identification of filaments, protozoa, spherical flocs and open flocs in microscopical images of wastewater treatment plants, in which YOLOv5 achieved an mAP of 67.0%, outperforming Faster R-CNN by 15% [21], or the identification of four biological categories in underwater conditions, where YOLOv5s, the lightest model with the fewest parameters, obtained an mAP of 84.9% [15].

According to some authors [14,15,21], YOLOv5s-based methodologies outperformed other deep-learning methods like Faster R-CNN in similar applications, particularly in challenging environments. YOLOv5s is noted for its higher mean Average Precision (mAP) and faster detection speed, making it more suitable for real-time applications when compared to two-stage detectors like R-CNN.

Despite the growing amount of work on image detection for litter quantification in the environment, most articles detect isolated waste where the objects appear against a smooth water surface [11,14], or detect floating waste by means of video cameras mounted on bridges [12] or surface vehicles [22]. No scientific works have dealt with detecting clustered floating waste, like the one that accumulates at booms and retaining structures of irrigation networks, using the YOLOv5 algorithm, perhaps because one of the main limitations of using YOLOv5 is its potential difficulty in accurately detecting small or overlapping objects in cluttered environments, which is common in floating waste scenarios.

The objective of this work is to propose a standardized and quick methodology to count and characterize the floating waste accumulated in the irrigation system to detect changes in its amount and composition. This objective will be attained by training an algorithm to automatically detect the different residues for their classification and quantification through image analysis using YOLOv5s, testing its performance on previous images counted manually and applying the model to overhead photographs taken with an unmanned aerial vehicle (UAV) at various strategic points of the Segura River and irrigation channels in the Vega Baja. The aim of the process is to eliminate the subjective component of manual quantification and characterization of floating waste to accurately describe the problem's magnitude.

While the study itself might not directly solve the problem of floating debris, its objective is to improve the detection and monitoring of such debris using UAVs and deep-learning techniques. This tool is devised to enhance data acquisition and analysis, improving the understanding of floating waste accumulation patterns, which could then inform targeted cleanup efforts and policymaking. The tool developed in this work will be at public disposal to impartially evaluate the effectiveness of the potential measures put in place by public authorities to solve the problem of floating debris accumulation in the river and the irrigation networks.

2. Materials and Methods

2.1. Description of the Study Area

The study took place in the Vega Baja region, which is located at the southern end of the province of Alicante, in the Valencian Community, Spain, along the system formed by the Segura River in its last section, and the irrigation and drainage channels that distribute water in the traditional irrigation area of Vega Baja (Figure 1 top).



Figure 1. Cont.

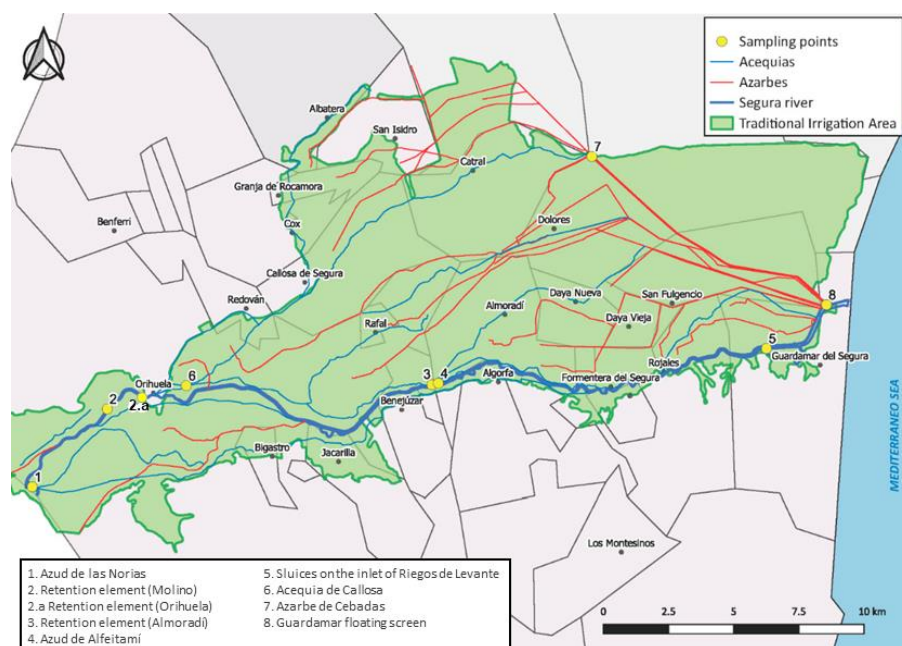


Figure 1. Location of the study area in the southern Alicante province (**top**) and location of the sampling sites along the river, irrigation channels (*acequias*) and drainage ditches (*azarbes*) (**bottom**).

The traditional irrigation of the Vega Baja del Segura has a unique set of hydraulic works that can be classified into two large groups: the dams that take the irrigation water directly from the river and the aqueducts that distribute the water to the traditional irrigated lands through an extensive and complex branched and hierarchical network of irrigation channels and drainage ditches [23]. A thorough description of the site can be found in Rocamora et al. [9].

2.2. Training of the Model

In previous work [9], the largest class of waste found in the Segura River and the surrounding irrigation network was made up of plastic bottles of bottled water (74.2%), containers of domestic origin like food containers, personal hygiene and cleaning products (14.9%), and containers of agricultural origin, bottles and cans of phytosanitary products (3.8%). Based on this information, we trained the algorithm to identify the following types of waste:

- Different-size plastic water bottles.
- Bleach bottles.
- Detergent bottles.
- Drums of chemical products from agricultural activity.
- Game balls.

2.2.1. Dataset

A carefully curated and diverse dataset was prepared to improve the detection accuracy of YOLOv5. A dataset of 477 images was created by gathering images from a previous study conducted in 2019, along with additional images of specific waste objects in the conditions of this study of the object classes we wanted our model to recognize. The selection criteria for these images focused on the types of plastic waste that were most commonly found in the Segura River and its irrigation network. These included water bottles, detergent bottles, bleach bottles, agricultural drums, and game balls. Of them, 430 images were labelled for training (90%) and 47 for model validation (10%). This dataset is partly available at the Rediumh repository (<https://hdl.handle.net/11000/32965>, accessed on 25 July 2024), which contains the images of our authorship.

2.2.2. Software

To process the images, several computer programs were used in sequence. The software used to label the images and carry out the model training was the following:

Makesense (<https://www.makesense.ai/index.html>, (accessed on 25 July 2024)): it is a free online photo-tagging tool that does not require any installation. Makesense was used to add all the labels of the different identified waste to a total of 477 images. After labelling, these images were downloaded in YOLO format.

YOLOv5 (https://pytorch.org/hub/ultralytics_yolov5/, (accessed on 25 July 2024)) [19]: a single-stage object detection network that is region-based. This model redefines object detection as a regression problem, allowing for high processing speed. Its architecture is made up of three main components: spine, neck and head. The backbone module extracts features from the input image using Focus, BottleneckCSP (inter-stage partial network) and SPP (spatial pyramid pooling) and transmits them to the neck module. The neck module generates a feature pyramid based on PANet (Path Aggregation Network), improving the ability to detect objects at different scales by bidirectionally fusing low-level spatial features with high-level semantic features. The head module produces the detection boxes, indicating the category, coordinates and confidence by applying anchor boxes to the multiscale feature maps generated by the neck module. There are four versions of the YOLOv5 model: YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. The SPP module mainly increases the receptive field of the network and captures features at different scales [24,25].

Google Colaboratory (<https://colab.research.google.com/>, (accessed on 25 July 2024)): it is a product of Google Research. It allows any user to write and execute arbitrary Python code in the browser. It is especially suitable for machine learning, data analysis and education tasks. From a more technical standpoint, Colaboratory is a zero-configuration Jupyter-hosted notebook service that provides access to computing resources such as GPUs.

Jupyter Notebook (<https://jupyter.org/>, (accessed on 25 July 2024)) docs are documents made up of code to be executed in Python and text that can include paragraphs, links, images, etc. Texts are often written using the Multi Markdown notation [26]. The documents contained the code and images that determined the model's behavior.

Python (<https://www.python.org/>, (accessed on 25 July 2024)) is a programming language created by Guido van Rossum in 1991 [27]. It is characterized by being high-level and interpreted, with special emphasis on code readability. It uses dynamic types and is a multi-paradigm language, which means that it supports object-oriented, structured, functional, etc programming. It is expandable, that is, it consists of a core to which modules and libraries can be added as required.

Sublime Text (<https://www.sublimetext.com/>, (accessed on 25 July 2024)) (© Sublime HQ Pty Ltd., Woollahra, Sydney) is a text editor for writing code in almost any file format. Sublime Text was used to create and edit the text documents necessary for training the YOLOv5 model. Specifically, it was used to define the number of classes, the names of these classes, and the paths to the training and validation folders.

2.2.3. Procedure

The first step was to label our images with Makesense software, where we used the following labels: water bottle, detergent bottle, bleach bottle, agricultural drum, and ball. Once these images were labelled, they were divided into two folders: one called "train", with the images that would be used to train the algorithm, and one called "val", with the images that would be used to validate the algorithm's training. Once the labelling was finished, we proceeded to extract these images in YOLO format to utilize them in Google Colab using YOLOv5.

The next step was to create a text document with Sublime Text, where we wrote the number of classes, the name of these classes and the address of our "train" and "val" folders.

YOLOv5 was downloaded in a Google Colaboratory file. The abovementioned "train" and "val" folders were added, as was the text document with the address of these folders. The image size, 1024 pixels, and the number of epochs, 126, were written in a new cell. These

parameters were chosen based on a balance between processing efficiency and accuracy. Image size refers to the number of times an image is divided when being analyzed: an increase in this number means that the image is analyzed more accurately. The number of epochs is a hyperparameter that defines the number of times the learning algorithm will work on the entire training data set. An epoch means that each sample in the training data set has had a chance to update the parameters of the internal model. At epoch 126, the model stopped improving and stabilized. The number of epochs was selected to ensure that the model had sufficient exposure to the training data to minimize error. The number of epochs is traditionally large, often hundreds or thousands, allowing the learning algorithm to run until model error has been sufficiently minimized [28]. Once the training cell was executed, a folder was created with validation and prediction of the results, as was a folder with the graphs analyzing the effectiveness of our trained algorithm.

2.3. Model Assessment

To verify the efficiency of the model, indicators such as precision, recall, mAP (mean Average Precision), detection speed and F1 score were used for evaluation in this study [29].

Precision (Equation (1)) is the proportion of true positive samples among all positive samples predicted by the model.

$$P = \frac{TP}{TP + FP} \cdot 100\% \quad (1)$$

where TP is the number of true positive cases, and FP is the number of false positive cases.

Recall (Equation (2)) indicates the proportion of true positive samples predicted by the model out of all true positive samples.

$$R = \frac{TP}{TP + FN} \cdot 100\% \quad (2)$$

where TP is as in Equation (1), and FN is the number of false negative cases.

Typically, there is a negative correlation between accuracy and recall, where one increases and the other decreases. To balance the effects of precision and recall and evaluate a model more comprehensively, AP (Average Precision) can be entered as a comprehensive evaluation index.

The value of AP (Equation (3)) is the area under the PR curve, a higher value meaning better model performance.

$$AP = \int_0^1 P(R) dR \quad (3)$$

The F1 score (Equation (4)) is the reconciled mean of precision and recall, considering both the precision (Equation (1)) and recall (Equation (2)) of the classification model.

$$F1score = (2 \cdot P \cdot R) / (P + R) \quad (4)$$

2.4. Image Acquisition

The images were obtained with an UAV, an Autel EVO 2 model quadrotor. The UAV specifications are as follows: weight: 1127 g, wheelbase: 397 mm, removable 7100 mAh 11.55 V batteries, maximum flight time: 40 min, maximum horizontal flight speed: 20 m/s, maximum flight distance: 25 km. Camera specifications: effective pixels: 48 MP, field of view: 79° and lens dimensions: 25.6 mm. The Autel EVO 2 quadrotor was chosen for its high-resolution camera, long flight time, and significant range. These features made it suitable for capturing detailed images over large areas, ensuring comprehensive coverage of the waste accumulation sites along the Segura River and its irrigation channels.

To define the image acquisition sites, we also kept the same garbage accumulation sites along the river and the irrigation channels network described in Rocamora et al. [9]. The most relevant accumulation sites were determined in a survey of the river layout along the Vega Baja, excluding those in which the regulations do not allow photographic flights

to be carried out because they are in an urban area. Regarding the accumulation sites in the channels, those indicated by the managers of the irrigation communities as most problematic were selected. Additional sites were included in this study since new retention booms were installed in 2022 by the Segura River Basin Authority (CHS) that were intended to capture floating waste more effectively.

Accumulation sites surveyed in this study are described in the following list:

Point 1 (ID1): Azud de las Norias,

Point 2 (ID2): Boom at the Town Mill,

In 2022, a new sampling point was added due to a new retention boom at the entrance to the city of Orihuela. This site is identified in the present work as ID2.a to maintain order in the direction of the river flow.

Point 2.a (ID2.a): Boom at the entrance of Orihuela,

Point 3 (ID3): Boom in Almoradí,

Point 4 (ID4): Azud de Alfeitamí,

Point 5 (ID5): Sluices on the inlet of the irrigation community of Riegos de Levante,

Point 6 (ID6): Acequia de Callosa, syphon at the crossing with Acequia Vieja de Almoradí,

Point 7 (ID7): Azarbe de Cebadas, syphon at the crossing with drainage channels del Convenio,

Point 8 (ID8): Guardamar floating screen.

In 2022, waste retention booms came into operation in the mouths of drainage channels discharging near point 8. The new booms were also considered as additional evaluation points for the present study, and are numbered as ID8.a, ID8.b, ID8.c, ID8.d and ID8.e.

These eight sites can be seen in Figure 1 (bottom) numbered along the river flow.

Flight planning was programmed for each of the different sampling points along with specific take-off and landing sites. In this study, three flights were carried out at each of the selected points along the river and channels on 28 July, 8 September and 31 October 2022. These flights were made at different heights: 5 m, 10 m and 20 m. The flights were conducted between 10:00 and 14:00 CET, under clear sky conditions, ensuring clear and consistent image capture across different sampling points. The UAV's camera took the aerial images in a zenithal view, and the number of photos at each point varied between one and nine, depending on the area extension of waste found in the river.

A total of 61 images, including all sampling dates and points, were analyzed using the algorithm trained in Section 2.2. The count results for each image were transferred to an Excel spreadsheet for further analysis.

Figure 2 depicts the conceptual framework of the methodology used to train and validate the model and to automatically detect the objects in a set of data acquired in the same environment during year 2022.

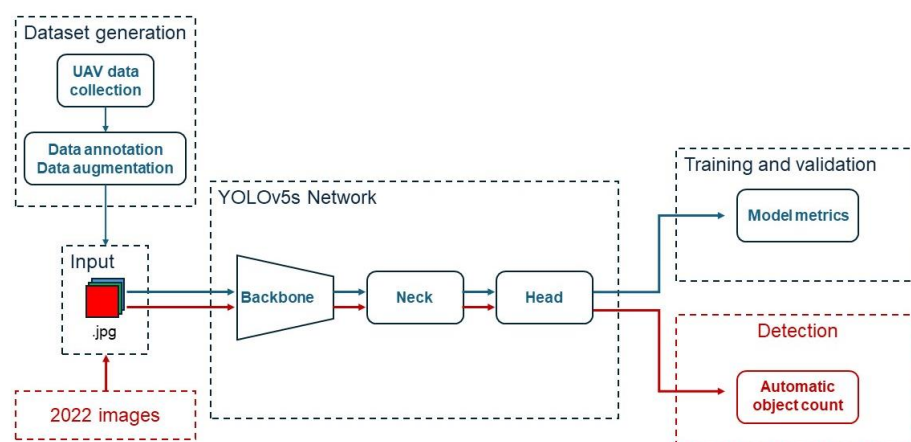


Figure 2. Conceptual framework of the object detection methodology. Blue color indicates the model training and validation, and red color indicates the process of object detection on the new images obtained in 2022.

2.5. Comparison of Automatic and Manual Count

To test the performance of the model, we manually marked and counted the objects appearing in the 61 images that were analyzed with the model with help of the ImageJ 1.54g application [30]. We recorded both the counting results by date and object class, and the time spent in the process.

The agreement between both methods was assessed using the methodology proposed by Bland and Altman [31], by studying the mean difference and constructing limits of agreement.

3. Results

3.1. Results of the Model Calibration and Performance Assessment

Model training was run for 126 epochs. The model's performance was validated using metrics like precision, recall, mean Average Precision (mAP) and the F1 score, as stated in Section 2.3. These metrics were continuously monitored over the 126 epochs to ensure that the model was learning effectively and improving its accuracy in detecting and classifying the different types of waste. A precision graph (Figure 3a) shows how the mAP@0.5 improved as training progressed. A rapid increase was observed at the beginning, followed by a gradual stabilization until around 126 epochs, where the value stabilized, indicating that the algorithm had reached its maximum performance in terms of mAP@0.5, achieving an mAP@0.5 of 0.96. Figure 3b shows the mAP over a range of thresholds from 0.5 to 0.95. A constant increase was seen until around 120 epochs, where the value stabilized at about 0.73, indicating that the algorithm had reached its maximum performance in this more stringent metric. Figure 3c illustrates the precision of the algorithm during training. Initially, the precision fluctuated significantly, but as training progressed, the fluctuations decreased, and the precision stabilized around 0.95 from 120 epochs onwards, suggesting that the algorithm had reached its maximum precision. This indicates that continuing training beyond this point did not provide significant additional improvements in these metrics. The batch size used was 2 because it presented a good confusion matrix, and choosing a higher batch size would have consumed more computing time.

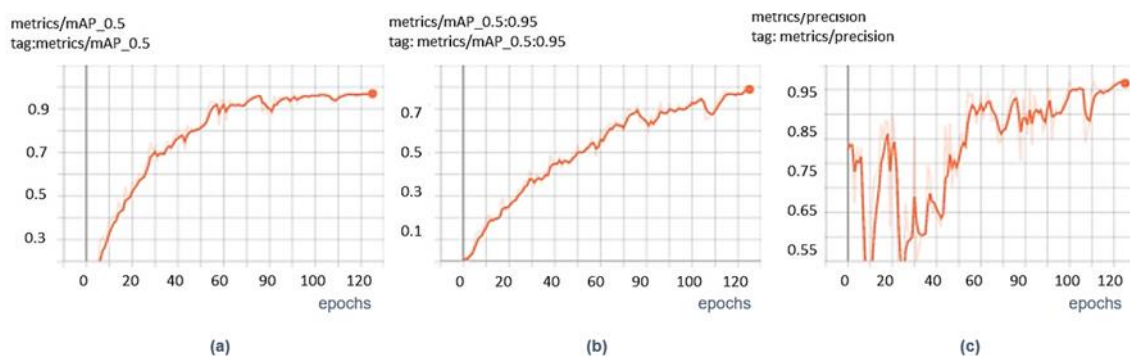


Figure 3. Precision graphs: (a) mAP@0.5, (b) mAP@0.5 to 0.95, (c) precision.

Model performance can be assessed by average precision, precision, recall, test time and F1 score. In this experiment, a confidence threshold of 0.5 was set, and precision, recall and F1 score were derived on this basis. The results obtained for this model were as follows: training time of 3 h 7 min; precision of 92.04%; recall of 97.07%; mAP of 96.90% and an F1 score of 94.48%.

The objective of this work was to identify the different residues quickly and accurately, for which the mAP values (96.90%) and the detection speed of 81.7 ms per image were used as the main evaluation metrics, where the higher the mAP value, the better the detection result.

The running training cell also provided the confusion matrix (Figure 4b) that allowed to visualize the performance of the supervised learning algorithm. Each column of the

matrix represents the number of predictions of each class, while each row represents the real class of each object, that is, it allows us to see what types of successes and errors our model had when going through the process of learning with the data [32]. In the present case, the probability that the detected objects were really those objects was high (89%), but there was an unusual fact with the background, which was confused with water bottles up to 85% of the time. This result might have been due to the fact that the analyzed photos were cropped to where only the waste layer could be seen, not the water in which they float, and at the bottom of this upper waste layer there were always more plastic water bottles.

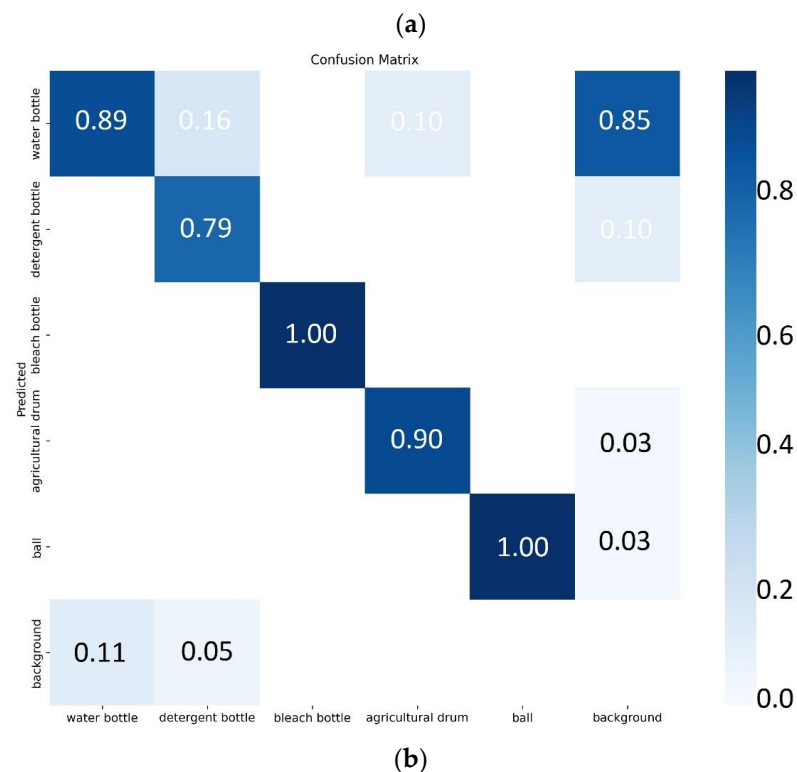
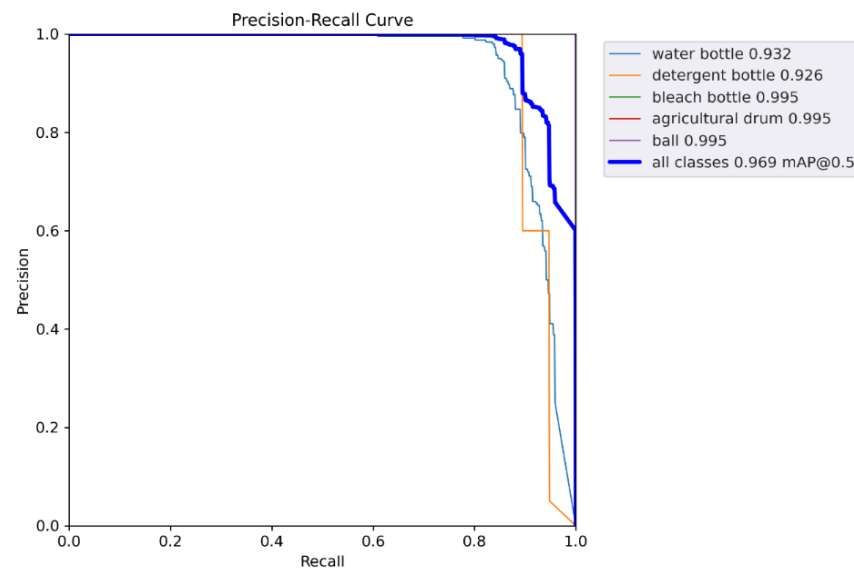


Figure 4. The model's precision–recall curve (a) and confusion matrix (b).

3.2. Automatic Count of Floating Waste

The algorithm was applied to the images captured throughout 2022 at the 11 sampling points defined in Section 2, Materials and Methods. A total of 61 images were analyzed in

the results cell, leading to the automatic counting of each one of them. The YOLO algorithm performed the identification and automatic count of residues on the analyzed images by means of bounding boxes (Figure 5).



Figure 5. Automatic characterization of waste at sampling point ID6 (**top**), and point ID8.b (**bottom**). Date 8 September 2022.

The results of the automatic count of floating objects are presented in Table 1. For each sampling point (in lines), the total of the detected objects is presented for each class (in columns). The results obtained by manual counting using the ImageJ application are also shown in Table 1 along with the automatic ones.

The most common objects were water bottles with a proportion of 96%, followed by detergent (2.07%) and bleach bottles (1.12%), and the least numerous were agricultural drums (0.6%) and balls (0.3%). This pattern of object classes' frequency distribution was found for all sampling points and on all dates. The number of elements fluctuated in some points, showing that there was no continuous accumulation. This was due to cleaning tasks that are carried out from time to time and the dynamics of the river and irrigation ditches.

Table 1. Count of objects detected with the algorithm and with the manual count.

Date	ID *	Automatic Counting with Algorithm					Manual Counting with ImageJ				
		Water Bottles	Agricultural Drum	Detergent Bottles	Bleach Bottles	Balls	Water Bottles	Agricultural Drum	Detergent Bottles	Bleach Bottles	Balls
28 July 2022	1	0					0				
	2	69		2		4	88				
	2.a	0					0				
	3	3					4				
	4	17					20				
	5	0					6				
	6	4					4				
	7	40	5	4	3		58	2	2	1	
	8	2		1	1		2				
8 September 2022	1	21					29		3		
	2	121		1	1		128	1	1	1	
	2.a	233	1	2	2	1	243	2	10	6	2
	3	154	2	5	2	1	173	3	5	6	2
	4	10		1	1		19			2	
	5	0					9				
	6	59				1	63		2		1
	7	37	2	5	2		49	4	2		
	8	90	1	3			96		2		
	8.a	236	1		2	1	329			5	3
	8.b	55					74		1		
	8.c	130		3			130	1	9		
	8.d	141	1	4	2		141	0	5	2	
	8.e	37					46	1	4		
31 October 2022	1	62	1	2	1		81	1		1	
	2	5	1	1	1	2	20				
	2.a	118	1			14	166	1			
	3	209		1	2	2	236		1	2	2
	4	0					0				
	5	199	6	8	4	1	218	4	9	4	1
	6	167	2	5	1		170	1	6	1	
	8	185		2	3		204		2	3	
	8.a	400	1	1	6		524	1	1	6	
8.d	8.e	153		2	8		174		2	8	
	8.d	58			1		72			1	

Note(s): * ID refers to the identification of sampling points as stated in Figure 1 of Section 2, Materials and Methods. 1: Azud de las Norias; 2: boom at the Town Mill; 2.a: boom at the entrance of Orihuela; 3: boom in Almoradí; 4: Azud de Alfeitami; 5: sluices on the inlet of the irrigation community of Riegos de Levante; 6: Acequia de Callosa, syphon at the crossing with Acequia Vieja de Almoradí; 7: Azarbe de Cebadas, syphon at the crossing with drainage channel Azarbe del Convenio; 8: Guardamar floating screen; 8.a, 8.b, 8.c, 8.d and 8.e, booms at channels discharging upstream of the Guardamar floating screen.

Comparing the number of objects obtained with each technique, the automatic count numbers were always equal or lower than the numbers obtained by the manual count in the case of water bottles. For other object classes, the differences between the two counts were more variable.

3.3. Comparison of Automatic Count with Manual Count

The time required for the manual counting of 61 images with the ImageJ application was 7 h and 34 min.

The comparison between the number of objects counted manually and the objects counted with the trained algorithm is shown in Figure 6. Each one of the scatter diagrams shows one of the objects classes. Each point shows the paired results of the manual count taken as reference in the X axes and the algorithm count in the Y axes. The difference between the two measurements can be noticed graphically in Figure 6, where the axes have been tailored to each object class depending on the counted quantities.

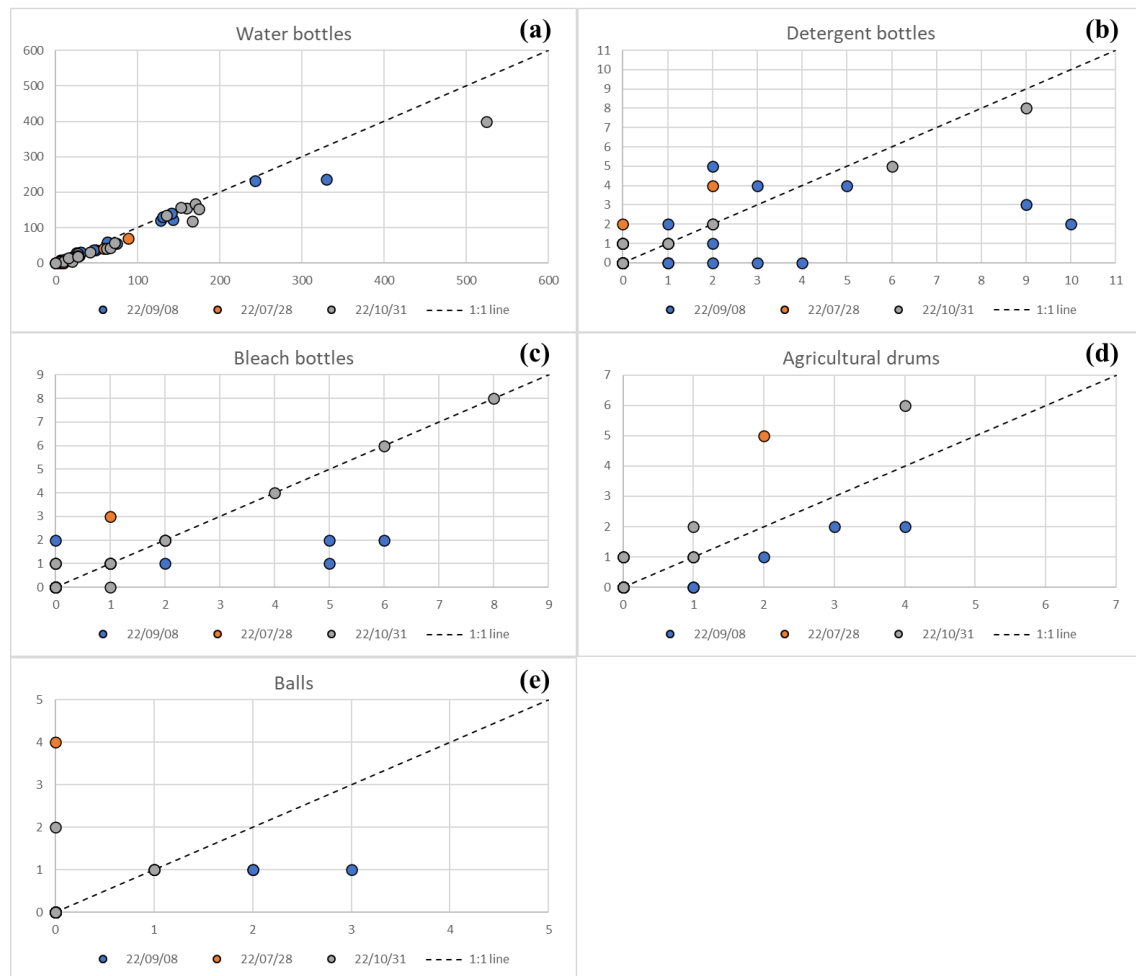


Figure 6. Comparison between manual (X axes) and automated (Y axes) count for water bottles (a), detergent bottles (b), bleach bottles (c), agricultural drums (d) and balls (e).

Agreement between object counts by the two methods, by hand and automatic, was analyzed, calculating bias and agreement ranges for each object class [30]. For each of the object classes, we found a bias of 11.4, and an agreement range from -32.25 to 55.05 units for water bottles (Figure 6a); a bias of 0.28 and an agreement range from -3.22 to 3.78 units for detergent bottles (Figure 6b); a bias of 0.12 and an agreement range from -1.92 to 2.16 units for bleach bottles (Figure 6c); a bias of -0.06 and an agreement range from -1.51 to 1.39 units for agricultural drums (Figure 6d); and a bias of -0.32 and an agreement range from -4.44 to 3.8 units for balls (Figure 6e).

For water bottles, detergent bottles and bleach bottles, the model tended to count fewer objects than the manual count did (positive bias); on the contrary, the model tended to count more agricultural drums and balls than the manual count did. In the case of water bottles, the bias seemed to change with count number, becoming higher when the count number was higher, probably due to the stacking of the bottles that could be appreciated in the manual count but not in the automatic one.

4. Discussion

Due to the reduced availability of models for detecting plastic waste when it is accumulated in the inland aquatic environment using UAV images, we introduced this model to improve and facilitate the task of identification and characterization of plastic waste to expedite cleaning tasks and to be able to apply measures for a faster way to reduce this problem. The results of this study show excellent performance in computing time and waste identification. The model worked well for both isolated residues and grouped residues, which were accurately detected.

Our proposed algorithm achieved an mAP50 of 96.9% and a detection speed of 81.7 ms. Compared to that of other research works, such as [14] with an mAP50 of 98.1%, [15] with an mAP50 of 82.9%, and [22] with an mAP50 of 76.12%, we can conclude that our approach provided strong results in both detection accuracy and speed.

The algorithm developed in this work could classify different types of plastic bottles: water bottles, bottles of hygiene and cleaning products, and cans of chemical products used in agriculture. This distinction among types of plastic bottles represents an enhancement over the algorithm of Wolf et al. [13], since the type of bottle can indicate its origin. However, our model suffers from false detections in the background of the images, identifying them as plastic bottles. Water bottles are the most difficult to characterize as they are found in large numbers and can be seen stacked together, making the bottle look like different shapes. On the contrary, bleach bottles and drums are easier to identify since they have a characteristic shape and color and are not found in large quantities.

The methodology used in this study, particularly the use of UAVs and YOLOv5 for object detection, could be adapted for monitoring other types of environmental issues, such as wildlife populations, vegetation health, or the spread of pollutants. The key advantage is the ability to cover large areas quickly and analyze the data with minimal human bias.

The results of the automatic count in three different dates on floating waste accumulation sites obtained in this work are in agreement with those reported by Rocamora et al. [9] in the same environment, who found that the most numerous class of waste is made up of water bottles, whereas bottles and cans of phytosanitary products represented only 3.8% of the elements counted at the sampling points. In our study, over three measurement days and more than eight measuring sites, we found that 95% of the floating items were plastic water bottles (96% by manual count), and only 0.8% of the items were agricultural drums (0.6% by manual count).

To compare the results of the algorithm with those of the manual count, it is necessary to highlight the time spent with the two methodologies. The time required for the manual counting of 61 images with the ImageJ application was 7 h and 34 min. On the other hand, the automatic counting of those 61 images lasted 42 s. Therefore, there is a clear reduction in the time required for the analysis, even considering the training time, 3 h and 7 min, which was only done once.

Besides the reduction in working time, we must add that the automatic count does not present issues such as problems of subjectivity or fatigue for the person carrying out the manual count.

The algorithm showed good agreement with the manual count except for the case of counting playing balls that are found in very small numbers. For water, detergent and bleach bottles, as well as for agricultural drums, the algorithm showed good agreement with the manual count by a trained observer.

With the data obtained, it was possible to monitor the presence of floating waste throughout the study period at each of the points where images were taken. Therefore, with this methodology, more continuous monitoring can be achieved, giving results almost in real time and the only limitation being imposed by UAV flights to take the images.

From the results of this study, the objects present in the greatest proportion cannot be exclusively associated with any sector of the population, since many of them can have domestic origin, such as cleaning product bottles and the larger part of water bottles. The presence of this type of waste may be due to the scattered homes in the area and the fact

that certain individuals dump their garbage into the river or into the irrigation network. Among the waste, there is a small proportion of agricultural drums despite the obligation to eliminate them through an integrated container management system. This differentiated count helps determine the origin of these accumulations in the river and therefore be able to implement measures, like awareness campaigns, to reduce this issue over time. The methodology that we propose in this work can be a useful tool due to its speed and objectivity in the evaluation of public policies aimed at eliminating plastic contamination in traditional irrigation networks.

The contribution of new knowledge of this work lies in the following features:

- (a) Application in a specific environmental context: This study focuses on the irrigation system of the Vega Baja region in Spain, a specific and understudied environment where floating plastic waste presents unique challenges. The complexity of this environment, including varying water levels, debris types, and narrow irrigation channels, necessitated the adaptation of existing models for effective waste identification.
- (b) Comprehensive dataset and longitudinal study: This work utilized a dataset that spans multiple years, capturing images in 2019 and 2022. This allowed us to evaluate the model's performance across different seasons and environmental conditions, providing insights into the model's robustness over time across accumulations varying in amount and composition.
- (c) Practical application and impact: Our study emphasizes the practical application of this methodology for a real-world situation. Local authorities in Vega Baja can integrate our findings into their waste reduction strategies, demonstrating the practical value of our work in preventing blockages and supporting public health initiatives.

Future research in our area could enhance the classes of objects detected, for instance, include Styrofoam trays, and drive more attention to the differentiation of stacked objects where waste accumulation takes place. Future enhancements might include testing the model in different locations to improve its robustness. Another potential enhancement could involve integrating other sensors (e.g., thermal or multispectral cameras) with the UAV to gather additional data that could improve detection accuracy in varied environmental conditions.

Author Contributions: Conceptualization, C.R. and H.P.; methodology, C.R. and H.P.; software, A.M.C.-A.; validation, A.M.C.-A.; resources and visualization, data curation, formal analysis, writing—original draft preparation, A.M.C.-A.; writing—review and editing, H.P. and C.R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by two research grants from Miguel Hernández University: 53ND0007RP Caracterización de residuos flotantes en ríos y canales mediante inteligencia artificial, and 79SH0010RP Análisis de imágenes en agricultura de precisión para cultivos de interés local.

Data Availability Statement: The dataset generated to train the YOLOv5s model in this article is deposited in the Universidad Miguel Hernandez of Elche repository at the handle <https://hdl.handle.net/11000/32965> (accessed on 18 September 2024).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Bellasi, A.; Binda, G.; Pozzi, A.; Galafassi, S.; Volta, P.; Bettinetti, R. Microplastic Contamination in Freshwater Environments: A Review, Focusing on Interactions with Sediments and Benthic Organisms. *Environments* **2020**, *7*, 30. [CrossRef]
2. van Emmerik, T.; Schwarz, A. Plastic debris in rivers. *WIREs Water* **2020**, *7*, e1398. [CrossRef]
3. van Emmerik, T.; Mellink, Y.; Hauk, R.; Waldschläger, K.; Schreyers, L. Rivers as Plastic Reservoirs. *Front. Water* **2022**, *3*, 786936. [CrossRef]

4. Hafeez, S.; Wong, M.S.; Abbas, S.; Kwok, C.Y.T.; Nichol, J.; Lee, K.H.; Tang, D.; Pun, L. Detection and Monitoring of Marine Pollution Using Remote Sensing Technologies. In *Monitoring of Marine Pollution*; IntechOpen: London, UK, 2019. [CrossRef]
5. Derraik, J.G.B. The pollution of the marine environment by plastic debris: A review. *Mar. Pollut. Bull.* **2002**, *44*, 842–852. [CrossRef]
6. Hurley, R.; Braaten, H.F.V.; Nizzetto, L.; Steindal, E.H.; Lin, Y.; Clayer, F.; van Emmerik, T.; Buenaventura, N.T.; Eidsvoll, D.P.; Økelsrud, A.; et al. Measuring riverine macroplastic: Methods, harmonisation, and quality control. *Water Res.* **2023**, *235*, 119902. [CrossRef]
7. Vaughn, T.; Crookston, B.M.; Pfister, M. Floating Woody Debris: Blocking Sensitivity of Labyrinth Weirs in Channel and Reservoir Applications. *J. Hydraul. Eng.* **2021**, *147*, 06021016. [CrossRef]
8. Stringam, B.L.; Gill, T.; Sauer, B. Integration of irrigation district personnel with canal automation projects. *Irrig. Sci.* **2016**, *34*, 33–40. [CrossRef]
9. Rocamora, C.; Puerto, H.; Abadía, R.; Brugarolas, M.; Martínez-Carrasco, L.; Cordero, J. Floating Debris in the Low Segura River Basin (Spain): Avoiding Litter through the Irrigation Network. *Water* **2021**, *13*, 1074. [CrossRef]
10. Abadía, R.; Brugarolas, M.; Rocamora, C.; Martínez-Carrasco, L.; Puerto, H.; Cordero, J. Causes, consequences and solutions to the problem of floating solid waste in the Segura River and its irrigation channels, in the district of Vega Baja (Alicante, Spain). In Proceedings of the 5th International Congress on Water, Waste and Energy Management (WWEM-19), Paris, France, 22–24 July 2019; p. 178.
11. de Vries, R.; Egger, M.; Mani, T.; Lebreton, L. Quantifying floating plastic debris at sea using vessel-based optical data and artificial intelligence. *Remote Sens.* **2021**, *13*, 3401. [CrossRef]
12. van Lieshout, C.; van Oeveren, K.; van Emmerik, T.; Postma, E. Automated River Plastic Monitoring Using Deep Learning and Cameras. *Earth Space Sci.* **2020**, *7*, e2019EA000960. [CrossRef]
13. Wolf, M.; van den Berg, K.; Garaba, S.P.; Gnann, N.; Sattler, K.; Stahl, F.; Zielinski, O. Machine learning for aquatic plastic litter detection, classification and quantification (APLastic-Q). *Environ. Res. Lett.* **2020**, *15*, 114042. [CrossRef]
14. Zhang, J.; Jiang, F.; Chen, Y.; Wu, W.; Wu, Q. A water surface garbage recognition method based on transfer learning and image enhancement. *Results Eng.* **2023**, *19*, 101340. [CrossRef]
15. Li, Y.; Wang, R.; Gao, D.; Liu, Z. A Floating-Waste-Detection Method for Unmanned Surface Vehicle Based on Feature Fusion and Enhancement. *J. Mar. Sci. Eng.* **2023**, *11*, 2234. [CrossRef]
16. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [CrossRef]
17. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]
18. Cai, Z.; Vasconcelos, N. Cascade R-CNN: Delving into High Quality Object Detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [CrossRef]
19. Tian, Z.; Huang, J.; Yang, Y.; Nie, W. KCFS-YOLOv5: A High-Precision Detection Method for Object Detection in Aerial Remote Sensing Images. *Appl. Sci.* **2023**, *13*, 649. [CrossRef]
20. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014. [CrossRef]
21. Inbar, O.; Shahar, M.; Gidron, J.; Cohen, I.; Menashe, O.; Avisar, D. Analyzing the secondary wastewater-treatment process using Faster R-CNN and YOLOv5 object detection algorithms. *J. Clean. Prod.* **2023**, *416*, 137913. [CrossRef]
22. Lin, F.; Hou, T.; Jin, Q.; You, A. Improved Yolo based detection algorithm for floating debris in waterway. *Entropy* **2021**, *23*, 1111. [CrossRef]
23. Trapote Jaume, A.; Roca Roca, J.F.; Melgarejo Moreno, J. Azudes y acueductos del sistema de riego tradicional de la Vega Baja del Segura (Alicante, España). *Investig. Geográficas* **2015**, *63*, 143–160. [CrossRef]
24. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016. [CrossRef]
25. Yao, J.; Qi, J.; Zhang, J.; Shao, H.; Yang, J.; Li, X. A real-time detection algorithm for kiwifruit defects based on yolov5. *Electronics* **2021**, *10*, 1711. [CrossRef]
26. Penney, F.T. MultiMarkdown User’s Guide. Version 6.6.0. Available online: <https://fletcherpenney.net/multimarkdown/> (accessed on 15 March 2024).
27. Van Rossum, G.; Drake, F.L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, USA, 2009.
28. Brownlee, J. Difference Between a Batch and an Epoch in a Neural Network. Available online: <https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/> (accessed on 19 July 2024).
29. Olson, D.L.; Delen, D. Performance Evaluation for Predictive Modeling. In *Advanced Data Mining Techniques*; Springer: Berlin/Heidelberg, Germany, 2008. [CrossRef]
30. US National Institutes of Health. Image J. Available online: <https://imagej.net/ij/index.html> (accessed on 19 July 2024).

31. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [[CrossRef](#)]
32. Arce, J.B. La Matriz de Confusión y Sus Métricas. Available online: <https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/> (accessed on 15 November 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Automatic Grape Cluster Detection Combining YOLO Model and Remote Sensing Imagery

Ana María Codes-Alcaraz ¹, Nicola Furnitto ², Giuseppe Sottosanti ², Sabina Failla ², Herminia Puerto ¹, Carmen Rocamora-Osorio ¹, Pedro Freire-García ³ and Juan Miguel Ramírez-Cuesta ^{2,*}

¹ Centro de Investigación e Innovación Agroalimentaria y Agroambiental (CIAGRO-UMH), Miguel Hernández University, 03312 Orihuela, Spain; acodes@umh.es (A.M.C.-A.); hpuerto@umh.es (H.P.); rocamora@umh.es (C.R.-O.)

² Department of Agriculture, Food and Environment (Di3A), University of Catania, 95123 Catania, Italy; nicola.furnitto@phd.unict.it (N.F.); giuseppe.sottosanti@unict.it (G.S.); sfailla@unict.it (S.F.)

³ Centro de Investigaciones sobre Desertificación (CIDE), CSIC-UV-GVA, Carretera CV 315, km 10.7, 46113 Valencia, Spain; pedrefre@gmail.com

* Correspondence: jm.ramirezcueta@unict.it

Abstract: Precision agriculture has recently experienced significant advancements through the use of technologies such as unmanned aerial vehicles (UAVs) and satellite imagery, enabling more efficient and precise agricultural management. Yield estimation from these technologies is essential for optimizing resource allocation, improving harvest logistics, and supporting decision-making for sustainable vineyard management. This study aimed to evaluate grape cluster numbers estimated by using YOLOv7x in combination with images obtained by UAVs from a vineyard. Additionally, the capability of several vegetation indices calculated from Sentinel-2 and PlanetScope satellites to estimate grape clusters was evaluated. The results showed that the application of the YOLOv7x model to RGB images acquired from UAVs was able to accurately predict grape cluster numbers (R^2 value and RMSE value of 0.64 and 0.78 clusters vine⁻¹). On the contrary, vegetation indexes derived from Sentinel-2 and PlanetScope satellites were found not able to predict grape cluster numbers (R^2 lower than 0.23), probably due to the fact that these indexes are more related to vegetation vigor, which is not always related to yield parameters (e.g., cluster number). This study suggests that the combination of high-resolution UAV images, multispectral satellite images, and advanced detection models like YOLOv7x can significantly improve the accuracy of vineyard management, resulting in more efficient and sustainable agriculture.

Keywords: object detection; sentinel-2; planetscope; unmanned aerial vehicle; yield estimation



Academic Editors: Wen Zhuo, Shibo Fang, Yi Xie and Xiaochuang Yao

Received: 6 December 2024

Revised: 8 January 2025

Accepted: 9 January 2025

Published: 11 January 2025

Citation: Codes-Alcaraz, A.M.; Furnitto, N.; Sottosanti, G.; Failla, S.; Puerto, H.; Rocamora-Osorio, C.; Freire-García, P.; Ramírez-Cuesta, J.M. Automatic Grape Cluster Detection Combining YOLO Model and Remote Sensing Imagery. *Remote Sens.* **2025**, *17*, 243. <https://doi.org/10.3390/rs17020243>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the current context of agriculture, which is in constant evolution, precision agriculture (PA) has emerged thanks to the strategic combination of intelligent sensors and advanced data analysis to guide agricultural decisions [1]. The purpose of PA is to optimize agricultural practices and management, increase resource efficiency, and reduce environmental impact [2]. This is achieved through the precise application of agricultural inputs only where and when they are needed, thus complying with environmental regulations and promoting sustainable agricultural practices. Additionally, PA addresses climate change through mitigation and adaptation strategies and proactively confronts barriers and challenges in its implementation [3]. This change in agriculture is driven by technologies such as remote sensing, geographic information systems (GISs), Global Positioning Systems (GPSs), big data analysis, Internet of Things (IoT), and artificial intelligence (AI) [4–7].

One of the technologies that has gained significant attention in recent years and has grown in the market is the use of UAVs for digital precision agriculture. Advances in control systems, robotics, electronics, and artificial intelligence have led to the development of sophisticated agricultural unmanned aerial vehicles (UAVs). UAVs offer advantages such as versatility, rapid and accurate remote sensing capabilities, and high-quality imaging at affordable prices. Research trends and patents show growing interest in UAVs for agricultural applications [8]. The use of advanced technologies, such as unmanned aerial vehicles (UAV) and computer vision systems, facilitates the collection of detailed data on soil and crop conditions. This data, analyzed by AI algorithms, allows for informed decisions on crop management, thus optimizing the productivity and sustainability of agricultural operations [9].

Various studies have indicated that UAVs stand out as one of the most effective technologies in precision agriculture, currently in widespread use and deployment [10–14]. These devices are extremely versatile and can perform remote sensing tasks quickly and accurately, even under adverse weather conditions, offering spatial and temporal resolutions that most satellite systems cannot achieve [15]. This high spatial resolution allows the identification of fine details and features that are often imperceptible in satellite images [16]. However, the performance and applicability of these systems depend on multiple factors, such as the weight of the aircraft, payload capacity, size, battery characteristics, flight autonomy under various conditions, acquisition cost, environmental conditions, and specific configurations. From an economic perspective, implementing UAVs in agriculture requires investments in purchasing the platform, additional equipment (such as laser scanners and multispectral or hyperspectral thermal cameras), and hiring a trained pilot both to handle the flight and to process the collected data [8].

The use of artificial intelligence in the viticulture sector is still in its early stages of development. Many processes in viticulture can be significantly improved by the use of artificial intelligence. However, its application has the potential to considerably optimize many processes within viticulture. In recent years, the wine industry has faced serious challenges due to the shortage of qualified labor and rising labor costs, which have affected productivity, wine quality, and timely grape harvesting [17]. Additionally, manual agricultural tasks are time-consuming and subject to the subjective decisions of workers. These problems have incentivized the development of new technologies, such as advanced sensors, including those from UAVs, along with artificial intelligence.

A critical aspect of vineyard management is estimating grape cluster numbers and yield, which is essential for the efficient planning of resources and labor. Traditionally, this task is labor-intensive and time-consuming, involving trained personnel to conduct visual counting on the plot. This has led to the exploration of remote sensing technologies, like the use of UAVs and satellites, offering faster and automated alternatives to manual counting. However, there is still no consensus on whether these technologies are fully capable of providing accurate grape cluster counts. UAVs offer a higher spatial resolution, making them suitable for identifying fine details such as grape clusters, whereas satellite images, with their lower spatial resolution but broader coverage, may contribute useful data on vegetation health and productivity trends over larger areas.

In recent years, the proliferation of modern sensors and the application of new technologies in precision viticulture have resulted in a significant increase in data generation per hectare. This has provided viticulturists with a wealth of information [18]. Data acquisition techniques commonly seen in viticulture can be classified into three groups: satellite imagery, data acquired using aerial platforms (most frequently UAVs), and data collected from terrestrial platforms. Given that vineyards typically cover relatively small areas, the latter two categories play a more important role in evaluating various parameters in

precision viticulture [19]. However, it is worth noting that satellite imagery still has its utility in this field [20].

To compare these data sources for vineyard management applications, it is essential to consider key variables such as ground sampling distance (or spatial resolution), acquired bands, and cost. Free multispectral satellite imagery, such as that provided by the European Union's Copernicus Program (through the Sentinel-2A and 2B satellites), starts at 10 m pixel⁻¹ and has more bands than typical UAV imagery. Cost is a relevant variable not only for a cost-benefit analysis in the context of user adoption but is also closely related to temporal resolution. The cost of each UAV image limits the number of images throughout the season, while free satellite data is available every five days. Generally speaking, and considering the cost, UAV images have an advantage in spatial resolution, while free Copernicus satellite images have an advantage in temporal resolution.

On the other hand, one of the most frequently raised issues today is the estimation of production on a farm, which is important for the proper management of a farm but involves laborious work that requires time and labor with specific experience or training [21]. Many studies have focused on production estimation by detecting fruits on the trees, including apples, cherries, grapes, avocados, mangoes, and passion fruits, among others [22–26]. Although fruit detection based on deep learning has been successful, inaccurate detection and localization can still occur due to the complexity of agricultural environments, lighting fluctuations, and unpredictability, such as wind or mechanical vibrations during the actual harvesting process [27,28].

In previous studies, grape cluster detection has commonly relied on ground-based imagery [26,29], while UAV imagery combined with NDVI has been used to differentiate live from dry vines [30]. Other authors have used vegetation indices to predict yield, though without direct fruit detection [31]. They revealed that vegetation indices like NDVI face challenges in yield prediction, with performance varying across growth stages and showing inconsistent R² values due to factors like climate and reference data. Alternatives like EVI and SAVI also produced contradictory results, further affected by cloud cover, atmospheric conditions, and calibration issues [24,32–36]. This highlights the need to perform grape cluster detection using UAV imagery and to investigate whether there is a relationship between these estimates and vegetation indices, providing a novel perspective for yield estimation.

Single-stage object detection algorithms typically use a neural network to directly predict the positions and categories of all possible bounding boxes in each image [29]. This approach includes the You Only Look Once (YOLO) series [37], which benefits from speed and real-time performance. Various studies have compared different versions of YOLO, such as [38], who compared YOLOv2, YOLOv3, and Mask R-CNN networks in grape detection research, or [39], who employed YOLOv3 for tomato detection using circular rather than traditional rectangular bounding boxes. Moreover, based on YOLOv5, ref. [40] performed lychee detection and calculated pixel coordinates of harvesting points. Currently, the latest version of YOLO is v7, which has been applied for detecting grape maturity in a white grape variety and outperforms other architectures both in precision and accuracy.

For all these reasons, the objectives of this research work were (i) to evaluate the number of grape clusters protected with paper bags estimated by using YOLOv7x in combination with images obtained by UAV from a vineyard and (ii) to assess the potentiality of different vegetation indexes obtained from Sentinel-2 and PlanetScope satellite images to estimate grape cluster number. Novelties of the present study include the use of artificial intelligence and oblique images, which could enhance the grape cluster detection process.

2. Materials and Methods

2.1. Study Site

The study plot (1.03 ha) is a vineyard (cv. Aledo) located in the municipality of Agost (Alicante, Spain; $38^{\circ}25'5.46''\text{N}$; $0^{\circ}38'22.56''\text{W}$) (Figure 1). The planting frame is $1.55\text{ m} \times 2.50\text{ m}$ with a total of 2742 plants, a registered production in 2022 of 12,000 kg, and an average cluster weight of 598 g (average data obtained by the cooperative after weighing 15 clusters). The Bagged Table Grape of Vinalopó used in this study is a unique fruit in the world due to its distinctive ripening process. Unlike other table grapes, the Vinalopó grape grows protected under a paper bag for at least 60 days, preserving each grape in the cluster until it reaches the consumer. Thanks to this particular cultivation method, the Vinalopó table grape possesses exceptional physical and gastronomic characteristics, making it the only bagged grape awarded a Protected Designation of Origin. Additionally, the bag shields the grapes from weather conditions, allowing them to develop much thinner skin, as they do not need to defend against the effects of rain, sun, or wind. This same protection also filters sun exposure, resulting in uniform color across the grapes in each cluster [41].



Figure 1. Study site location at Alicante, Spain ($38^{\circ}25'5.46''\text{N}$; $0^{\circ}38'22.56''\text{W}$); an example of RGB image acquired with the UAV at a height (h) of 17 m and a sensor tilt (θ) of 45° (green rectangle).

The meteorological conditions registered in the nearest weather station managed by the Sistema de Información Agroclimática para el Regadío (SIAR; <http://riegos.ivia.es/red-siar>, last accessed 5 December 2024), located 1 km far from the experimental plot (Agost weather station, Alicante, Spain; $38^{\circ}23'12''\text{N}$; $0^{\circ}45'28''\text{W}$) were typical Mediterranean with an average air temperature, relative humidity, wind speed, solar radiation and precipitation of 17.86°C , 63.81%, 1.28 m s^{-1} , $16.02\text{ MJ m}^{-2}\text{ day}^{-1}$, and 338 mm, respectively.

The irrigation system used in this plot was drip irrigation, which is known for its high-frequency water application. This system delivers water directly to the root zone, optimizing its use and minimizing evaporation and deep percolation losses. The seasonal water dose applied was 974 m³. The irrigation quantity was managed by the farmer in order to ensure no water stress was observed in the crop.

In addition, it is essential to consider the specific needs of the crop, as well as the efficiency of the soil, which is influenced by its composition. This composition affects the retention and use of water by the plants. Lastly, a water balance is used to compare the water supplied by rainfall with the amount needed by the crop, highlighting the importance of irrigation during months of water deficit.

The soil in the study area was of a limestone nature. Its lithological origin comes from terrace gravel deposits derived from continental sediments, marine sediments, molasses, limestone, and marl with sandstones and limestones. All the soils in the valley have a high calcium carbonate content and low organic matter [42].

2.2. UAV Flight Campaign

The aerial images were taken at 10:00 AM (local time) on 27 October 2022, using a UAV Autel EVO 2 pro (UUC Technology Co., Ltd., Hong Kong) equipped with a visible RGB camera with a 1" CMOS sensor (Sony Group Corporation, Tokyo, Japan), adjustable F2.8–F11 aperture capable of 20 megapixels resolution and video recording in 4 K/30 FPS. The UAV flight was conducted following a path parallel to the crop lines. The flight height (h) was 17 m, and the sensor tilt (θ) was 45°. To prevent overcounting of grape clusters, the images were captured with a 10% overlapping. The onboard camera obtained these images with a resolution of 3840 × 2160 pixels, resulting in a pixel size of about 1.1 cm. The total flight time was 6.5 min.

2.3. YOLO Model

2.3.1. Training Pre-Processing

To generate the dataset, the RGB images acquired from the UAV with the flight configuration shown in Figure 1 were used. These images were labeled using the Makesense application (<https://www.makesense.ai/index.html>, accessed on 25 November 2024), a free online photo-tagging tool). As output, grape clusters were annotated with rectangles (Figure 2). A total of 60 images were labeled; these aerial images used for the test were obtained from several flights carried out on different farms with the same flight configuration, where each of the images had 400 labels, out of which 90% were used for training the model and 10% for validation [43,44]. These images were annotated in YOLO format. These images were then resized to 640 × 640 pixels since YOLOv7x is a pre-trained algorithm with an image size of 640 pixels.

The RGB images were subsequently processed using the object detection algorithm known as YOLO. This algorithm is a state-of-the-art single-stage object detection model that approaches detection as a regression problem [37]. This model can detect multiple objects in a single image by dividing the image into a grid. Initially, each grid cell predicts the probability of a class. Then, the model determines bounding box coordinates and calculates confidence scores for each. As a single-stage detector, YOLO models achieve high computational speed [37]. YOLOv7 [45] is an advanced iteration of the widely used YOLO object detection algorithm, known for its single-pass architecture, which enhances efficiency over other object detection methods.

What sets YOLOv7 apart from earlier versions is its superior detection accuracy, which has been achieved through various innovations. These include optimized layer efficiency, effective model scaling techniques, re-parameterization strategies, and the addition of an

auxiliary head with a coarse-to-fine supervision method. Together, these enhancements result in significant performance, efficiency, and adaptability improvements, making YOLOv7 a more effective solution for many practical applications [45]. Within YOLOv7, there are new versions, such as YOLOv7x, which have been shown in various studies to achieve the best results [45]. Specifically, ref. [46] compared two UAV flight configurations, one at 10 m height with a 35° sensor tilt angle and another at 17 m height with a 45° sensor tilt angle, to assess which one provided better detection accuracy. This study concluded that, at 17 m, YOLOv7x performed better than the other evaluated versions (YOLOv7-E6 and YOLOv8), demonstrating greater stability and lower RMSE under this condition. Specifically, the three YOLO versions showed similar mean Average Precisions (mAP) (0.81–0.85), with YOLOv7x presenting the lowest RMSE values (0.78 clusters vine⁻¹ for YOLOv7x versus 1.39–2.36 clusters vine⁻¹ for YOLOv7-E6 and YOLOv8). The better performance of YOLOv7x was also observed by [46] when evaluating the total yield estimated from the different YOLO versions. The yield estimates from the YOLOv7x version did not differ from the measured yield, whereas a significant yield underestimation was obtained for the YOLOv7-E6 and YOLOv8 versions at a configuration of 17 m height and 45° sensor tilt angle. Given that the images used in the present study were captured at 17 m, YOLOv7x was chosen as the detection model, as it provided more reliable and consistent results with this flight configuration. The entire process of training the algorithm was carried out on Google Colaboratory, a free platform provided by Google. Google Colaboratory allows users to write and run Python v3.10 code directly from their browser, utilizing resources like Graphics Processing Units and Tensor Processing Units at no additional cost. This makes it an ideal tool for developing and testing machine learning models without the need for personal infrastructure. The training process can be observed in Figure 3.



Figure 2. Example of cluster detection using YOLOv7x with RGB images acquired from UAV. Numbers in the bounding boxes represent the model's confidence in detecting the objects, expressed as a value between 0 and 1.

2.3.2. Performance Metrics

To verify the effectiveness of YOLOv7x model, the accuracy (P), recall (R), Average Precision (AP), detection speed, and F1 score were used in this study.

Specifically, P is the ratio of true positive samples among all positive samples predicted by the model (Equation (1)). The proportion of true positive samples predicted by the model to all true positive samples is considered by the term R (Equation (2)). There is usually an inverse correlation between P and R . To balance the effects of P and R and evaluate a model more holistically, AP was calculated through the area under the PR curve (Equation (3)), where a higher value means a better model performance. The $F1_{score}$ (Equation (4)) is the

harmonic mean of P and R , providing a balanced assessment of a model's performance while considering both false positives and false negatives (Equation (4)).

$$P = \frac{TP}{TP + FP} 100(\%) \quad (1)$$

$$R = \frac{TP}{TP + FN} 100(\%) \quad (2)$$

$$AP = \int_0^1 P(R) dR \quad (3)$$

$$F1_{score} = \frac{2 \times P \times R}{P + R} \quad (4)$$

where TP , FP , and FN are the numbers of true positive cases, false positive cases, and false negative cases, respectively [25]. One of the most commonly used metrics for evaluating training is the mAP, which measures the average precision of the model considering different confidence thresholds, with the IoU (Intersection over Union) threshold set at 0.5 in this case. This value, which varies from 0 to 1, reflects the overall performance of the model in correctly identifying grape clusters in most of the analyzed images [45]. In addition, a dataset of 25 images of individual vines was selected to validate the number of grape clusters estimated using the YOLOv7x model. The validation procedure consisted of comparing the estimated value with the grape cluster number obtained in the field by visual counting. The images were carefully selected throughout the entire plot to capture the existing variability in the growing conditions. This strategic selection of images allows for a more representative analysis of the model's effectiveness.

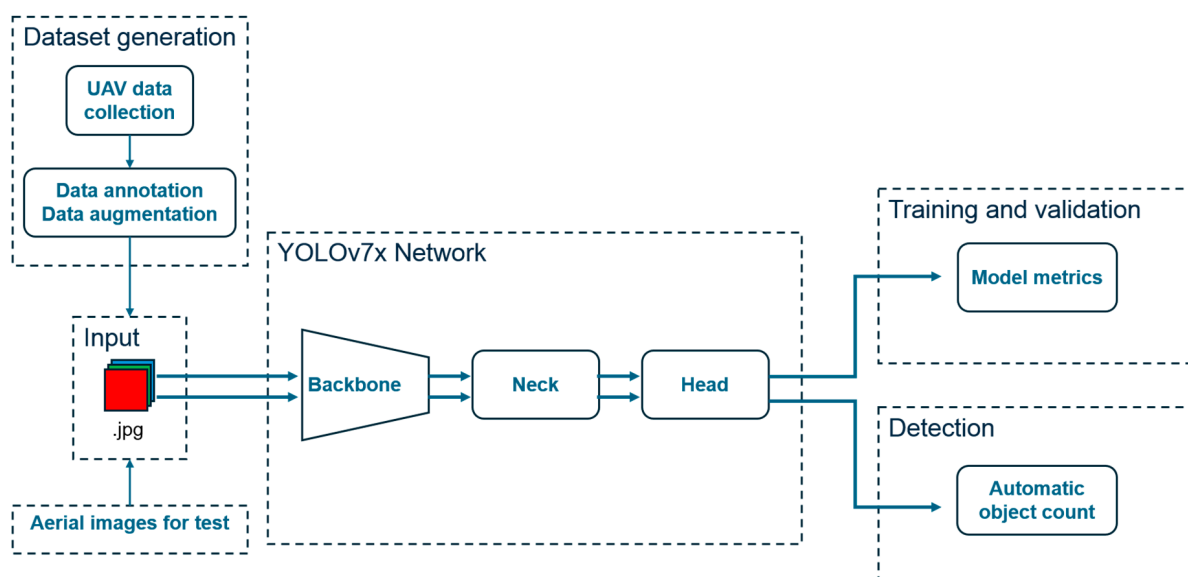


Figure 3. Conceptual framework of the object detection methodology.

2.4. Remote Sensing Data

This subsection describes the satellite products utilized for deriving the vegetation indexes and the methodology followed for performing the comparison between such indexes and the grape clusters number derived from the YOLOv7x model.

2.4.1. Sentinel-2 Imagery

Satellites Sentinel-2A and B were used, which have a multispectral instrument on board able to capture imagery across 13 spectral bands with a combined constellation

revisit of five days and a wavelength-dependent spatial resolution from 10 to 60 m [47]. In this study, we used the bands with a spatial resolution of 10 m, which are Band 2—Blue with a Central Wavelength of 490 nm, Band 3—Green 560 nm, Band 4—Red 665 nm, and Band 8—NIR 842 nm [48]. During the study period, 36 cloud-free images were downloaded from the Copernicus Open Access Hub website (<https://scihub.copernicus.eu/>; accessed on 10 July 2024) between February and November 2022. Specifically, Level-2A (L2A) products were considered since they provide images of the bottom of atmosphere reflectance covering the visible and NIR spectral range.

2.4.2. PlanetScope Imagery

PlanetScope SuperDove, a new generation of DOVE CubeSat, was used in this study. The PlanetScope SuperDove instrument, able to image the entire land surface of the Earth every day, has eight spectral bands covering the red edge (705 nm), red (665 nm), green (565 nm), green I (531 nm), yellow (610 nm), blue (490 nm), coastal blue (442 nm), and NIR (865 nm) spectral regions and a pixel size of 3 m [49]. For the period February to November 2022, 100 images of PlanetScope were downloaded from the Planet Explorer website (<https://www.planet.com/explorer/>, accessed on 10 July 2024). Specifically, the PlanetScope Ortho Analytic 8B SR product (level-3B) was used, which offers an orthorectified, scaled Surface Reflectance 8-band image product suitable for analytic applications [50].

2.4.3. Data Processing

ArcGIS 8.3, a geographic information system (GIS) software developed by the Environmental Systems Research Institute (ESRI), was utilized in this study. This software provides a framework for developing customized applications based on existing components [51]. It was employed to extract the Sentinel-2 and PlanetScope bands for each pixel. Initially, the Fishnet tool in ArcGIS was used to map all the vines within the plot, taking into account the planting layout. Subsequently, only the pixels that were fully covered by vines, with no empty spaces, were selected, excluding those on the edge of the plot. This pixel identification was later used to determine the number of grape clusters in each pixel and to associate this information with the Sentinel-2 and PlanetScope bands, enabling a correlation between the two datasets.

Additionally, a point shapefile with the location of each vine was generated, where the number of clusters per vine estimated by the YOLOv7x model was included. To establish the relationship between the estimated cluster number and the different vegetation indexes calculated from the Sentinel-2 and PlanetScope images, the average of the entire study area was considered. The comparative analysis was conducted both for individual dates and for the cumulative value over time.

Based on previous research on yield estimation [52–55], 11 widely used vegetation indices were calculated (Table 1), including the Normalized Difference Vegetation Index (NDVI) [56]; Green Normalized Difference Vegetation Index (GNDVI) [57]; Soil-Adjusted Vegetation Index (SAVI) [58]; Modified Soil-Adjusted Vegetation Index (MSAVI) [59]; Modified Triangular Vegetation Index (MTVI) [60]; Enhanced Vegetation Index (EVI) [61]; Optimized Soil Adjusted Vegetation Index (OSAVI) [62]; Simple Ratio (SR) [63]; and Normalized Difference Water Index (NDWI) [64].

All the UAV images of the plot analyzed by the algorithm were cropped by vines in order to later locate them in the corresponding pixels. This was performed to later correlate them with the vegetation indices obtained from Sentinel-2 and PlanetScope. Figure 2 shows an example of detection from the training. The bounding boxes, generated using the YOLOv7x model with a confidence threshold of 0.25, are overlaid on the image to show the detected objects.

Table 1. Vegetation indexes were calculated in the present study. GREEN, RED, and NIR are the spectral bands of Sentinel-2 and PlanetScope satellites corresponding to the green, red, and near-infrared spectral regions (wavelengths of 560, 665, and 842 nm, respectively, for Sentinel 2, and 565, 665, and 865 nm, respectively, for PlanetScope).

Vegetation Indices	Equation	Reference
Normalized Difference Vegetation Index (NDVI)	$\frac{(NIR - RED)}{(NIR + RED)}$	[56]
Soil Adjusted Vegetation Index (SAVI)	$\frac{(1 + 0.5)(NIR - RED)}{NIR + RED + 0.5}$	[58]
Enhanced Vegetation Index (EVI)	$\frac{2.5(NIR - RED)}{(NIR + 6RED - 7.5BLUE + 1)}$	[61]
Green Normalized Difference Vegetation Index (GNDVI)	$\frac{NIR - GREEN}{NIR + GREEN}$	[57]
Modified soil adjusted vegetation index (MSAVI)	$0.5 \left(2(NIR + 1) - \sqrt{(2NIR + 1)^2 - 8(NIR - RED)} \right)$	[59]
Modified triangular vegetation index (MTVI)	$\frac{1.5[1.2(NIR - GREEN) - 2.5(RED - GREEN)]}{\sqrt{(2NIR + 1)^2 - (6NIR - 5\sqrt{RED})} - 0.5}$	[60]
Normalized Difference Water Index (NDWI)	$\frac{GREEN - NIR}{GREEN + NIR}$	[64]
Optimized Soil-Adjusted Vegetation Index (OSAVI)	$\frac{(NIR - RED)}{NIR + \frac{RED}{NIR} + 0.16}$	[62]
Simple Ratio Index (SR)	$\frac{NIR}{RED}$	[63]

To evaluate the performance of the YOLOv7x model in detecting grape clusters, the most common object detection metrics were used: mAP, Precision, Recall, and F1 Score. During the training of the algorithm, each vine within the plot was identified and segmented, and the resulting images were used as input for the algorithm analysis. The algorithm detected and counted the grape clusters of each vine. Then, using a point grid created in ArcGIS, the cluster data for each vine was integrated, generating a map that showed the distribution of vines within the plot, along with the associated cluster number. This method allowed for a precise estimation of the number of vines and grape clusters in each pixel of the image. To calculate the vegetation indices, the spectral band reflectance values from Sentinel-2 and PlanetScope were extracted at the selected pixels in ArcGIS, and the data were transferred to an Excel spreadsheet, where the indices were calculated using the equations shown in Table 1.

3. Results

3.1. Model Training and Cluster Detection and Validation

As shown in Figure 4, after experimenting with combinations of batch size and epochs, it was determined that 250 epochs (i.e., number of times the learning algorithm will work on the entire training data set) with a batch size (i.e., number of times an image is divided when being analyzed) of 12 were appropriate for training the YOLOv7x model. In this case, mAP_0.5 reaches a value of 0.63 at epoch 249. Starting from epoch 249, it can be observed that the graph forms a straight line, indicating that the training will no longer improve.

The precision achieved was 61%, meaning that 61% of the objects detected by the model were indeed grape clusters (true positives). The recall reached 79%, indicating that the model successfully detected 79% of the grape clusters present in the images. Finally, the F1 Score obtained was 66%, which represents the harmonic mean between precision and recall.

The estimated number of clusters per vine varied between 5 and 10, being slightly higher than the measured in the field (4–9 clusters vine^{−1}). More specifically, the correlation between estimated and observed values was statistically significant (Figure 5), with an R² value and RMSE value of 0.64 and 0.78 clusters vine^{−1}, respectively, and had a slope term equal to 0.77 with an intercept value of 1.26 (Table 2).

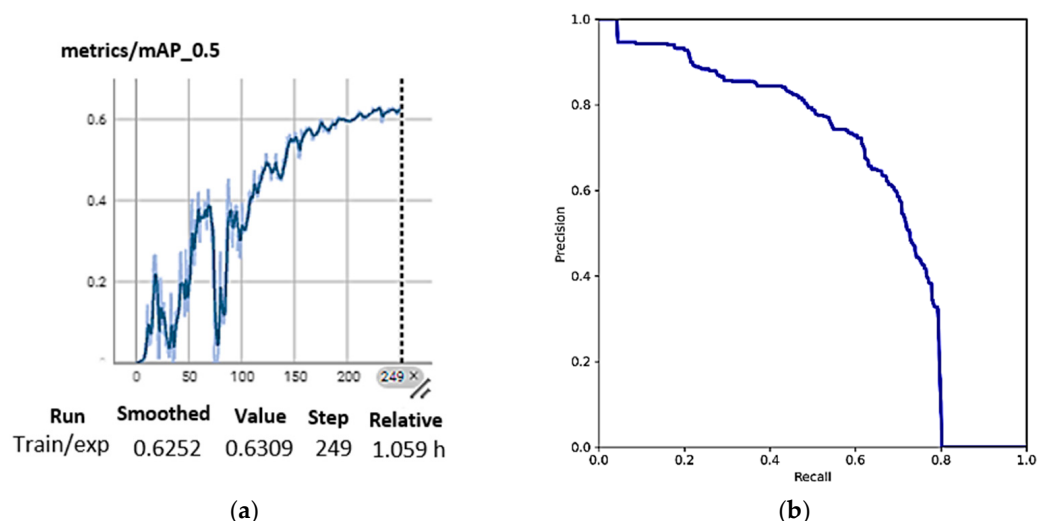


Figure 4. (a) Mean Average Precision at an IoU threshold of 0.5, reflecting the model's detection accuracy, and (b) Precision-Recall Graph indicating the model's effectiveness in detecting objects.

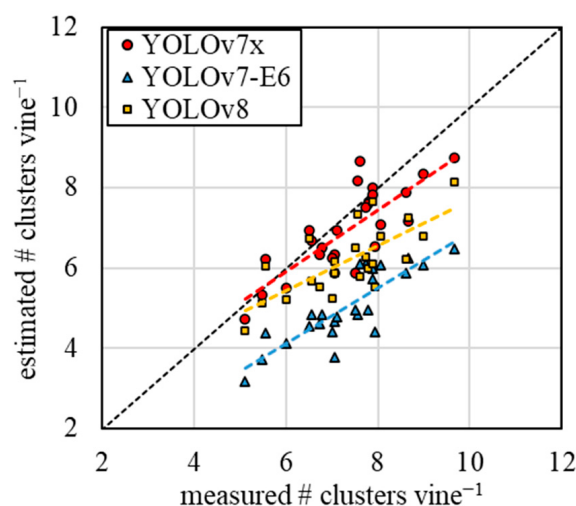


Figure 5. Comparison between the number of grape clusters per vine estimated using the YOLOv7x model and the actual measurements observed from the UAV images. The red dashed line represents the trend line regression, while the black dashed line indicates the 1:1 relationship. For comparison purposes, the trend lines obtained by using YOLOv7-E6 (blue dashed line) and YOLOv8 versions (yellow dashed line) have been included (adapted from [46]).

Table 2. Correlation equations, coefficient of determination (R^2), and Root Mean Square Error (RMSE) for the different YOLO versions (adapted from [46]).

YOLO Version	Equation	R^2	RMSE (Cluster Vine ⁻¹)
YOLOv7x	$y = 0.77x + 1.26$	0.64	0.78
YOLOv7-E6	$y = 0.69x - 0.04$	0.71	2.36
YOLOv8	$y = 0.56x + 2.10$	0.53	1.39

3.2. Comparison of Reflectance in the VIS-NIR Spectral Range Obtained from PlanetScope and Sentinel-2 Satellites

From the comparison between the reflectance values of the Blue, Green, Red, and NIR bands obtained from Sentinel-2 and PlanetScope satellites (Figure 6), a general over-estimation of Sentinel-2 in comparison with PlanetScope was observed. Specifically, the

correlation among both satellites showed a slope term of 0.89 and an intercept value of 0.11. The relation was highly significant, with an R^2 value of 0.76 and an RMSE value of 0.11.

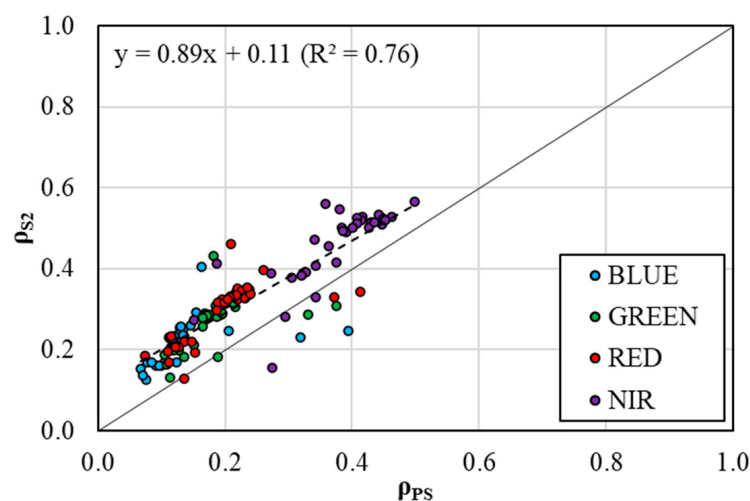


Figure 6. Comparison of the reflectance values from PlanetScope (ρ_{PS}) and Sentinel-2 (ρ_{S2}) satellites obtained in the visible (blue, green, and red) and near-infrared (NIR) spectral bands. The black dashed line represents the general trend line regression, while the black continuous line indicates the 1:1 relationship.

It was observed that the vegetation indices derived from PlanetScope and Sentinel-2 data demonstrated almost identical and consistent temporal patterns. An example of the NDVI obtained from Sentinel 2 and PlanetScope images is shown in Figure 7. All vegetation indices showed their lowest values (in absolute terms) at the beginning of the vegetative period. Vegetation indices values began to increase steadily after a few weeks, varying from Day of Year (DOY) 55–99 depending on the satellite used, indicating the start of vegetative stages (e.g., leaf development) and significant growth. The vine growth increased until reaching DOYs 130–190, corresponding to the highest growth values. From that point, the vine started to fruit, coinciding with when vegetation indices began to decrease at DOYs 205–249. A peculiarity in this case is that values rise again between DOYs 280 and 330, which corresponds to the fact that weeds have not been removed from the field. However, in absolute terms, the values calculated from Sentinel-2 and PlanetScope responded differently in the function of the evaluated vegetation index.

The major discrepancies among satellites occurred in terms of NDVI, GNDVI, NDWI, and SR (Figure 8). More specifically, NDVI derived from PlanetScope shows an average range of values from 0.11 to 0.51, while Sentinel-2 ranges between 0.09 and 0.33. GNDVI values in PlanetScope ranged from 0.19 to 0.51 during the crop cycle, whereas the values derived from Sentinel-2 varied from 0.14 to 0.34. NDWI values from PlanetScope and Sentinel-2 ranged from -0.51 to -0.19 and from -0.34 to -0.14 , respectively. In the case of SR, the values derived from PlanetScope imagery during the evaluated time period ranged from 1.25 to 3.16, whereas the values obtained from Sentinel-2 varied from 1.19 to 1.98.

Lower differences among the values provided by both satellites were observed in terms of SAVI and OSAVI. Thus, SAVI values ranged from 0.10 to 0.36 when using PlanetScope images, whereas the values ranged from 0.07–0.27 when using Sentinel-2 data. Similar values were obtained for OSAVI, with values ranging from 0.09 to 0.38 and from 0.07 to 0.26 when considering PlanetScope and Sentinel-2, respectively.

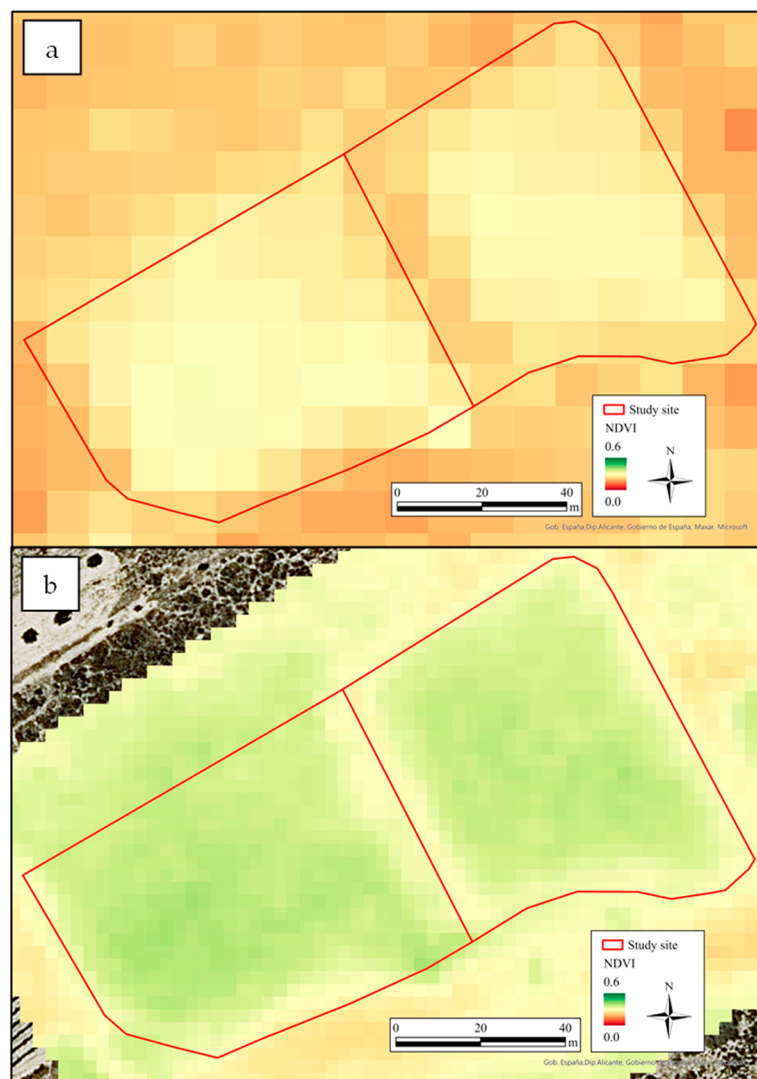


Figure 7. Example of NDVI calculated from Sentinel2 (a) and PlanetScope (b) images acquired on 29 June 2022.

The values calculated from PlanetScope and Sentinel-2 did not show significant differences in terms of EVI, MTVI, and MSAVI, except for the last dates of the cycle, where the differences became greater. Thus, EVI, MTVI, and MSAVI from both satellites varied from 0.09 to 0.39, from 0.00 to 0.54, and from 0.57 to 0.84, respectively.

3.3. Detection of Clusters with YOLOv7x and Comparison with Vegetation Indices

From the comparison between the total cluster number derived from YOLOv7x and the average vegetation indexes in the field obtained for each single image and for the cumulative value, it can be observed that vegetation indexes were not able to accurately predict the cluster numbers. These results occurred both when using PlanetScope-derived vegetation indexes and when using those vegetation indexes calculated from Sentinel-2 imagery (Tables 3–5). More specifically, the lack of a significant relationship was evidenced by the low R^2 values obtained, which reached a maximum of 0.23 when using Sentinel-2 imagery (NDWI and GNDVI versus cluster vine⁻¹ on 6 November 2022; Table 3), being even lower when PlanetScope imagery was used (maximum of 0.02 obtained for NDVI versus cluster vine⁻¹ on 25 May 2022; Table 4). The consideration of the cumulative values during the evaluated period did not improve the obtained results, with the R^2 values of the relationships vegetation indexes versus clusters vine⁻¹ lower than 0.09 (Table 5).

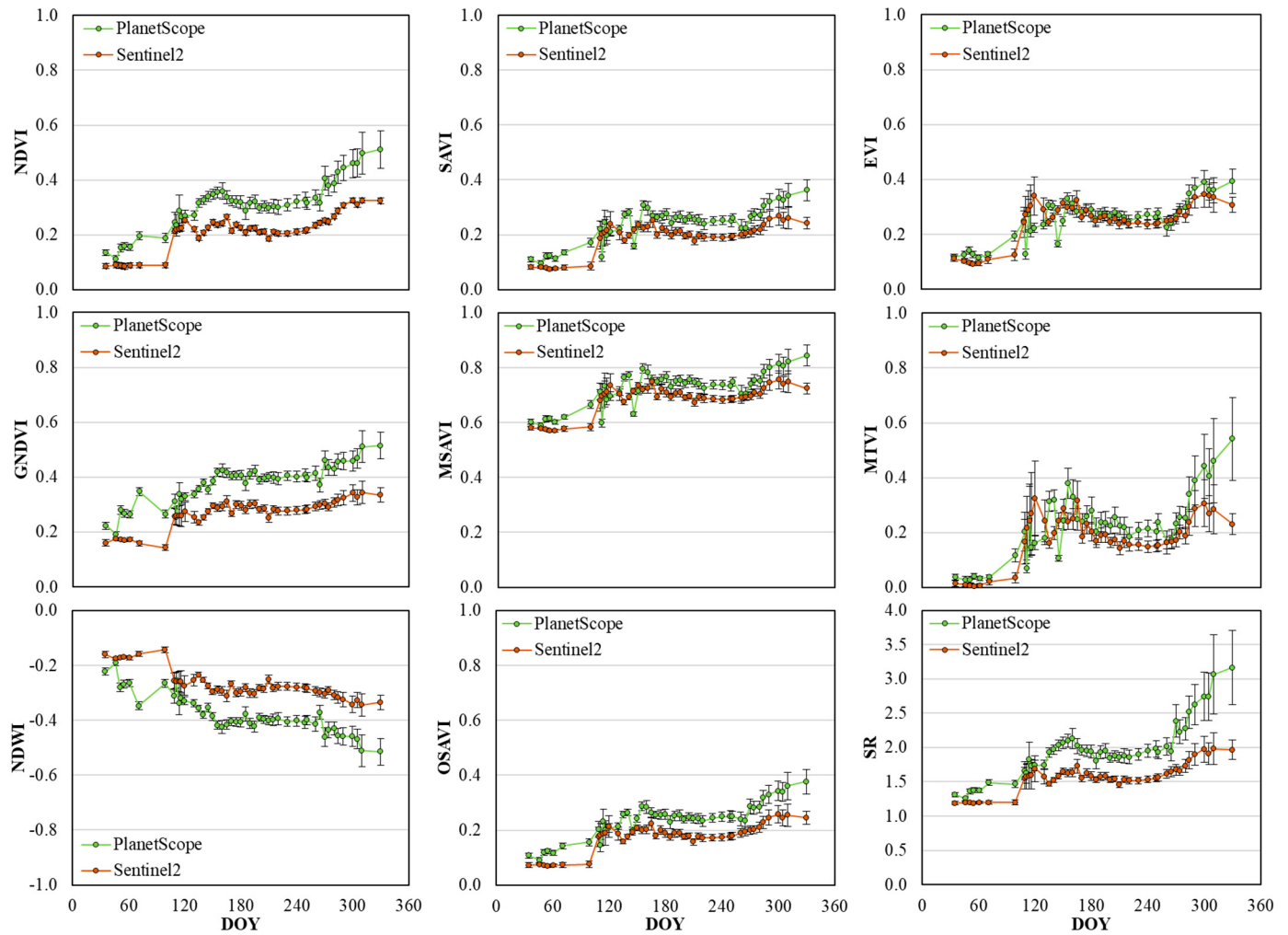


Figure 8. Temporal evolution of the different vegetation indexes calculated from PlanetScope and Sentinel-2 satellites (DOY: Day Of Year).

Table 3. R^2 values obtained for the relationships between the different vegetation indexes derived from the Sentinel-2 satellite and the number of clusters for each individual date. For the meaning of the indexes, the reader is referred to the main text. The bold values evidence the maximum obtained R^2 value.

Data	NDVI	SAVI	EVI	GNDVI	MSAVI	MTVI	NDWI	OSAVI	SR
4 February	0.007	0.003	2×10^{-4}	4×10^{-6}	0.001	0.020	4×10^{-6}	0.006	0.007
14 February	0.013	0.013	0.005	0.021	0.008	0.014	0.021	0.023	0.031
24 February	9×10^{-4}	0.005	0.007	0.062	0.007	0.031	0.062	0.003	0.001
1 March	0.004	0.010	0.006	0.055	0.013	7×10^{-4}	0.055	0.007	0.004
10 April	0.090	0.096	0.096	0.052	0.098	0.107	0.052	0.093	0.089
20 April	0.110	0.114	0.116	0.108	0.115	0.109	0.108	0.112	0.105
25 April	0.118	0.122	0.124	0.106	0.123	0.119	0.106	0.120	0.113
30 April	0.125	0.129	0.128	0.120	0.129	0.121	0.120	0.127	0.120
10 May	0.002	0.002	0.003	0.001	0.002	0.004	0.001	0.002	0.002
15 May	0.039	0.043	0.059	0.008	0.045	0.063	0.008	0.041	0.040
20 May	0.011	0.015	0.027	0.020	0.016	0.014	0.020	0.013	0.011
25 May	0.009	0.024	0.021	0.028	0.029	0.011	0.028	0.015	0.009
30 May	0.005	0.008	0.015	0.010	0.009	0.011	0.010	0.006	0.005
4 June	3×10^{-4}	4×10^{-4}	0.002	1×10^{-4}	4×10^{-4}	0.001	1×10^{-4}	4×10^{-4}	3×10^{-4}
9 June	0.007	0.008	0.007	0.007	0.008	0.007	0.007	0.008	0.007
14 June	0.004	0.004	0.004	0.005	0.004	0.002	0.005	0.004	0.003
19 June	0.023	0.021	0.018	0.003	0.020	0.026	0.003	0.022	0.022

Table 3. Cont.

Data	NDVI	SAVI	EVI	GNDVI	MSAVI	MTVI	NDWI	OSAVI	SR
24 June	0.032	0.032	0.042	0.024	0.032	0.030	0.024	0.032	0.030
29 June	3×10^{-5}	2×10^{-4}	9×10^{-4}	0.010	2×10^{-4}	4×10^{-4}	0.010	8×10^{-5}	7×10^{-5}
4 July	7×10^{-6}	2×10^{-4}	0.003	0.006	3×10^{-4}	0.005	0.006	5×10^{-5}	2×10^{-5}
9 July	0.042	0.040	0.055	0.025	0.039	0.043	0.025	0.041	0.042
14 July	0.001	3×10^{-4}	2×10^{-4}	0.001	1×10^{-4}	8×10^{-8}	5×10^{-4}	7×10^{-4}	8×10^{-4}
19 July	0.038	0.033	0.043	0.020	0.031	0.035	0.020	0.036	0.036
24 July	0.038	0.033	0.022	0.059	0.031	0.017	0.059	0.036	0.037
29 July	0.003	0.001	0.004	0.005	0.001	0.010	0.005	0.002	0.003
3 August	4×10^{-4}	2×10^{-5}	8×10^{-4}	0.001	2×10^{-4}	0.001	0.001	9×10^{-5}	3×10^{-4}
8 August	0.003	0.001	0.005	0.001	0.001	0.004	8×10^{-4}	0.002	0.003
13 August	0.091	0.092	0.094	0.062	0.093	0.056	0.062	0.092	0.092
28 August	3×10^{-4}	0.003	0.001	0.011	0.005	7×10^{-4}	0.011	0.001	4×10^{-4}
7 September	0.006	0.002	0.011	0.022	0.001	0.020	0.022	0.004	0.006
22 September	0.023	0.053	0.050	0.040	0.061	0.040	0.040	0.038	0.023
27 September	0.006	0.001	2×10^{-4}	0.002	4×10^{-4}	7×10^{-4}	0.002	0.004	0.005
2 October	8×10^{-4}	0.005	0.007	2×10^{-4}	0.007	0.010	2×10^{-4}	0.002	6×10^{-4}
7 October	2×10^{-4}	0.002	0.003	0.001	0.003	0.001	0.001	8×10^{-5}	2×10^{-4}
17 October	0.010	0.022	0.026	0.029	0.026	0.022	0.029	0.015	0.010
27 October	0.044	0.064	0.068	0.058	0.068	0.067	0.058	0.052	0.044
1 November	0.114	0.132	0.133	0.147	0.135	0.123	0.147	0.122	0.112
6 November	0.114	0.220	0.211	0.234	0.216	0.198	0.234	0.223	0.217
26 November	0.013	0.027	0.019	0.047	0.032	0.023	0.047	0.019	0.015

Table 4. R^2 values obtained for the relationships between the different vegetation indexes derived from the PlanetScope satellite and the number of clusters for each individual date. For the meaning of the indexes, the reader is referred to the main text. The bold value evidences the maximum obtained R^2 value.

Date	NDVI	SAVI	EVI	GNDVI	MSAVI	MTVI	NDWI	OSAVI	SR
4 February	0.002	0.001	0.001	3×10^{-4}	0.001	0.002	3×10^{-4}	0.002	0.002
15 February	0.001	0.002	0.001	0.001	0.002	0.003	0.001	0.001	0.001
20 February	0.002	0.002	0.001	0.003	0.002	4×10^{-4}	0.003	0.002	0.002
24 February	0.003	2×10^{-4}	1×10^{-4}	1×10^{-4}	3×10^{-6}	0.003	1×10^{-4}	0.001	0.002
2 March	0.002	0.005	0.005	0.001	0.006	0.004	0.001	0.003	0.002
12 March	0.007	0.010	0.009	0.007	0.010	0.006	0.007	0.008	0.007
9 April	0.005	0.005	0.005	0.002	0.004	0.006	0.002	0.005	0.005
19 April	0.008	0.009	0.009	0.010	0.009	0.008	0.010	0.009	0.008
21 April	0.006	0.002	0.002	0.003	0.002	0.004	0.003	0.004	0.006
24 April	0.007	0.007	0.007	0.007	0.008	0.007	0.007	0.007	0.006
26 April	0.010	0.011	0.010	0.011	0.011	0.010	0.011	0.010	0.010
10 May	3×10^{-6}	3×10^{-4}	0.001	0.003	4×10^{-4}	2×10^{-5}	0.003	7×10^{-5}	6×10^{-6}
15 May	0.005	0.006	0.007	0.006	0.007	0.006	0.006	0.006	0.005
20 May	0.006	0.006	0.009	0.001	0.006	0.007	0.001	0.006	0.006
25 May	0.022	0.019	0.018	0.016	0.018	0.018	0.016	0.021	0.021
4 June	0.004	0.003	0.003	0.002	0.002	0.003	0.002	0.003	0.004
9 June	7×10^{-7}	2×10^{-5}	9×10^{-6}	9×10^{-5}	4×10^{-5}	7×10^{-5}	9×10^{-5}	3×10^{-6}	5×10^{-8}
14 June	0.003	0.002	0.002	0.003	0.002	0.001	0.003	0.002	0.002
19 June	1×10^{-5}	8×10^{-6}	2×10^{-8}	2×10^{-7}	2×10^{-5}	2×10^{-5}	2×10^{-7}	1×10^{-7}	5×10^{-6}
29 June	0.002	0.001	0.001	3×10^{-4}	0.001	0.001	3×10^{-4}	0.001	0.001
4 July	0.019	0.018	0.017	0.020	0.017	0.015	0.020	0.019	0.019
9 July	0.006	0.005	0.004	0.005	0.005	0.004	0.005	0.006	0.006
14 July	0.010	0.007	0.008	0.007	0.007	0.007	0.007	0.009	0.009
19 July	0.004	0.003	0.003	0.003	0.003	0.002	0.003	0.004	0.004
24 July	0.007	0.005	0.006	0.006	0.005	0.004	0.006	0.006	0.007
29 July	0.003	0.001	0.001	4×10^{-4}	0.001	0.001	4×10^{-4}	0.002	0.003
3 August	0.010	0.007	0.006	0.008	0.006	0.006	0.008	0.009	0.009
8 August	0.001	4×10^{-4}	4×10^{-4}	0.001	3×10^{-4}	2×10^{-4}	0.001	0.001	0.001

Table 4. Cont.

Date	NDVI	SAVI	EVI	GNDVI	MSAVI	MTVI	NDWI	OSAVI	SR
18 August	0.007	0.004	0.004	0.003	0.003	0.003	0.003	0.006	0.006
28 August	0.010	0.007	0.010	0.009	0.007	0.006	0.009	0.009	0.009
6 September	0.003	0.001	0.002	0.001	0.001	0.001	0.001	0.002	0.003
8 September	0.010	0.006	0.006	0.006	0.005	0.005	0.006	0.008	0.009
17 September	0.021	0.015	0.016	0.008	0.013	0.016	0.008	0.020	0.020
22 September	0.008	0.001	0.001	0.004	0.001	0.001	0.004	0.004	0.007
27 September	0.008	0.005	0.005	0.006	0.004	0.005	0.006	0.007	0.007
1 October	0.010	0.006	0.005	0.009	0.004	0.004	0.009	0.008	0.009
7 October	0.001	2×10^{-4}	4×10^{-4}	0.001	4×10^{-4}	4×10^{-4}	0.001	2×10^{-5}	2×10^{-4}
11 October	2×10^{-4}	6×10^{-5}	3×10^{-5}	2×10^{-5}	2×10^{-4}	1×10^{-4}	2×10^{-5}	2×10^{-5}	1×10^{-4}
17 October	0.017	0.018	0.006	0.006	0.018	0.008	0.006	0.018	0.014
27 October	1×10^{-4}	0.001	0.001	2×10^{-5}	0.001	0.001	2×10^{-5}	2×10^{-5}	0.014
1 November	0.001	7×10^{-6}	1×10^{-4}	4×10^{-4}	8×10^{-5}	2×10^{-4}	4×10^{-4}	2×10^{-4}	2×10^{-4}
6 November	2×10^{-6}	0.001	0.001	6×10^{-5}	0.001	0.001	6×10^{-5}	1×10^{-4}	1×10^{-5}
26 November	0.008	0.005	0.002	0.007	0.004	0.002	0.007	0.007	0.005

Table 5. R^2 values obtained for the relationships between the different vegetation indexes derived from Sentinel-2 and PlanetScope satellites and the number of clusters for the cumulative period of the study. For the meaning of the indexes, the reader is referred to the main text.

Platform	NDVI	SAVI	EVI	GNDVI	MSAVI	MTVI	NDWI	OSAVI	SR
Sentinel-2	0.059	0.073	0.082	0.076	0.078	0.083	0.076	0.065	0.066
PlanetScope	0.001	2×10^{-5}	2×10^{-6}	3×10^{-4}	2×10^{-6}	6×10^{-7}	3×10^{-4}	0.001	0.001

4. Discussion

Object detection models based on YOLOv7x showed moderate accuracy in grape cluster detection, with metrics such as mAP and F1 score reaching acceptable values. However, it is necessary to reduce false positives and improve the precise localization of clusters. This suggests that, although the model is able to detect most of the grape clusters present in the images, it also makes errors by detecting incorrect objects. Compared to other YOLO versions, such as YOLOv5s and YOLOv3-spp, YOLOv7x demonstrated better accuracy [65]. For example, ref. [46] compared the precision and recall metrics of YOLOv7x, YOLOv7-E6, and YOLOv8, observing that YOLOv7x demonstrated greater consistency under different UAV flight configurations, with RMSE values of 0.75 and 0.78 grape clusters per vine. The combination of UAV images with advanced algorithms, such as YOLOv7x, significantly improved detection compared to the exclusive use of satellite data, which could lack enough spatial resolution to perform grape cluster estimations. In fact, the pixel sizes of the satellite images considered (Sentinel-2 and PlanetScope), do not account only for the presence of grape clusters but include all features included within the pixels of 10 m and 3 m. In this sense, the utilization of satellite images with a higher spatial resolution, such as those obtained from WorldView satellites or even from UAVs, could allow for a more accurate identification of the elements of the images. Additionally, the main purpose of vegetation indices was to represent vegetation vigor and health, rather than directly indicating yield parameters.

Despite the overall good performance of YOLOv7x, its detection capacity is limited by false positives and difficulties in identifying clusters partially covered or under adverse lighting conditions. A possible improvement would be the use of image preprocessing techniques, such as lighting normalization, and increasing the dataset for algorithm training.

The results obtained in this study highlight the importance of combining various remote sensing data sources to optimize grape cluster detection in vineyards, thus contribut-

ing to precision agriculture. The comparison between UAV, Sentinel-2, and PlanetScope images shows that, although UAVs provide superior spatial resolution, the satellite data from Sentinel-2 and PlanetScope offer complementary advantages in terms of radiometric stability and temporal coverage.

Specifically, UAVs excel in their ability to detect fine details, such as grape clusters, thanks to their high spatial resolution, which is crucial for small plots and detailed studies [66–69]. However, their lower radiometric stability and associated costs limit their application to larger scales or prolonged studies. On the other hand, Sentinel-2, with its greater radiometric stability and revisit frequency, is more suitable for large-scale and long-term studies, as evidenced by previous research highlighting its efficiency in monitoring crop growth and health over larger areas [52,70,71].

The inclusion of PlanetScope images, which have high temporal resolution and intermediate spatial resolution, allowed for capturing detailed variations within plots. However, higher variability in reflectance values was observed, as noted in [72,73]. PlanetScope's superior results are due to both its higher geometric resolution, which better aligns with the study areas, and its ability to adjust pixel purity, which measures the degree of homogeneity with respect to the target crop [70]. This suggests that PlanetScope images could be more useful in studies requiring detailed observations at small scales, where precise radiometric adjustments are needed [70,71]. Overall, the synergistic combination of high-resolution PlanetScope data with medium-resolution Sentinel-2 data outperforms the predictive performance of each dataset separately, due to the combination of spatial and spectral information from both sources [71–74]. Thus, the use of multi-scale models is recommended to optimize predictions.

Regarding yield estimation, higher correlation values were observed in Sentinel-2 data compared to PlanetScope, which is consistent with other studies comparing both datasets [63]. Although PlanetScope data had higher temporal and spatial resolution, it had lower radiometric resolution than Sentinel-2 data. The lack of two additional red-edge and SWIR bands in PlanetScope data may also be a factor influencing these results [63]. Nevertheless, PlanetScope's near-daily image frequency provides opportunities to improve prediction accuracy and promote digital agriculture for crop modeling [75]. Currently, Sentinel-2 images are considered the most suitable for field-level yield estimation [76]. The higher standard deviation in PlanetScope suggests that its higher spatial resolution allows it to capture finer differences in areas with less dense vegetation or mixed with bare soil. Conversely, Sentinel, with its lower but more uniform resolution, offers more consistent results for large-scale studies.

The vegetation indices used, such as NDVI, SAVI, EVI, GNDVI, MSAVI, MTVI, NDWI, OSAVI, and SR; are designed to assess vegetation vigor and overall cover but do not capture specific details like fruit clusters, which often have spectral characteristics similar to surrounding foliage or are partially obscured by the canopy.

This is because these indices are designed to assess overall vegetation cover but cannot accurately differentiate grape clusters from surrounding vegetation, which often share similar spectral characteristics or are partially covered by the canopy. Our results showed an increase in indices in autumn, possibly related to weed growth on the ground, consistent with other studies [66,77], where this increase is also related to orchard water consumption depending on both irrigation practices and pedoclimatic conditions [78–80].

In related studies, ground-based imagery is commonly used for grape bunch detection [29], while others used UAV imagery and NDVI to identify live and dry vines [30]. Some works also compared the performance of different satellite image sources, such as WorldView, PlanetScope, and Sentinel-2, concluding that high-resolution imagery, such as WorldView (i.e., <1 m), is essential for detecting small features [81]. This could explain why

neither Sentinel nor PlanetScope imagery has shown any relationship with the grape cluster number in the present study. Other studies focused on yield prediction using vegetation indices but without performing detection on the fruit [31]. This limitation of the vegetation indexes (such as NDVI) for deriving yield parameters has also been reported by other authors. In particular, refs. [31,32] observed that vegetation indexes (e.g., NDVI, EVI, and SAVI) showed inconsistency in yield prediction depending on factors such as the amount of reference data, study area, climate, cloud cover, atmospheric conditions, and inadequate calibration affect the accuracy of these measurements.

For future research, the use of LiDAR data is proposed, as it could provide three-dimensional information about the canopy and clusters, improving detection accuracy [74]. The combination of LiDAR data with multispectral images and advanced object detection models could open new opportunities for automatic yield estimation in vineyards.

Finally, although this study focused on grape cluster detection, it would be interesting to apply this approach to other fruit crops facing similar challenges in terms of automatic fruit detection.

These results demonstrate that although PlanetScope offers higher spatial resolution, Sentinel-2's reflectance values show greater consistency, which could translate into better capability for temporal or large-area studies. On the other hand, PlanetScope, due to its higher variability in bands, might be more suitable for applications requiring detailed observation of smaller areas but with a greater need for radiometric adjustment.

This difference in bands is critical when selecting which platform to use based on the study's objective. Sentinel-2 is ideal for more homogeneous and long-term monitoring, while PlanetScope is preferable when very fine spatial resolution and detailed analysis of small agricultural plots are required.

5. Conclusions

This study provides a detailed comparison between the use of UAV images and satellite images from Sentinel-2 and PlanetScope for the automatic detection of grape clusters in vineyards using the YOLOv7x model. The results of training the YOLOv7x model demonstrated moderate accuracy in grape cluster detection, with acceptable performance in metrics such as precision and recall, but with room for improvement in reducing false positives. It is concluded that vegetation indices alone are not sufficient to accurately predict the presence of clusters, as they are designed to assess vegetative vigor and biomass rather than detect specific objects such as grape clusters. However, it was observed that PlanetScope, despite its high spatial resolution, did not achieve as stable a correlation with grape cluster detection data compared to Sentinel-2.

It is observed that classical vegetation indices, such as NDVI, SAVI, EVI, GNDVI, MSAVI, MTVI, NDWI, OSAVI, and SR from both Sentinel-2 and PlanetScope, did not show a significant correlation with grape cluster detection. These indices, commonly used to assess plant vigor and health, do not effectively differentiate the clusters due to their spectral similarities with the surrounding foliage or their partial occlusion under the plant canopy.

In terms of data sources comparison, images obtained from UAVs offer a significant advantage in terms of spatial resolution, allowing for a more detailed identification of grape clusters. The detailed comparison of reflectance bands and vegetation indices between Sentinel-2 and PlanetScope reveals that, although PlanetScope offers a higher spatial resolution of 3 m, Sentinel-2 (10 m) shows greater radiometric stability and temporal consistency. This makes it more suitable for vegetation studies over large areas and prolonged temporal analyses, while PlanetScope could be more useful in studies where fine spatial resolution is of greater importance.

The integration of UAV images with advanced deep learning models such as YOLOv7x significantly enhances the detection of grape clusters by leveraging the superior spatial resolution of UAV data. This combination enables more accurate identification of clusters, reducing occlusion issues and increasing the precision of detection models. UAVs offer flexibility in data collection, allowing for timely and targeted surveys that can be adapted to specific vineyard conditions. This synergy between UAV imagery and deep learning provides a robust framework for addressing the limitations of satellite data and vegetation indices, ultimately supporting more precise, data-driven decision-making in vineyard management.

Finally, this study suggests that the most effective approach for detecting grape clusters in vineyards combines the use of UAV images to capture specific details, along with multispectral satellite data such as Sentinel-2 and PlanetScope, complemented by advanced object detection models like YOLOv7x. This approach allows for optimizing productivity and improving sustainability in vineyard management, providing more precise and effective tools for data-driven decision-making.

Author Contributions: Conceptualization, A.M.C.-A. and J.M.R.-C.; methodology, A.M.C.-A., N.F., G.S. and P.F.-G.; validation, A.M.C.-A., N.F. and G.S.; formal analysis, A.M.C.-A., N.F., G.S., S.F., H.P., C.R.-O., P.F.-G. and J.M.R.-C.; data curation, A.M.C.-A., N.F., G.S., P.F.-G. and J.M.R.-C.; writing—original draft preparation, A.M.C.-A. and J.M.R.-C.; writing—review and editing, P.F.-G., S.F., H.P. and C.R.-O.; supervision, S.F., H.P., C.R.-O. and J.M.R.-C.; funding acquisition, H.P. and C.R.-O. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded by the research grant from Miguel Hernández University 79SH0010RP “Análisis de imágenes en agricultura de precisión para cultivos de interés local”.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bhakta, I.; Phadikar, S.; Majumder, K. State-of-the-art technologies in precision agriculture: A systematic review. *J. Sci. Food Agric.* **2019**, *99*, 4878–4888. [\[CrossRef\]](#)
2. Kendall, H.; Clark, B.; Li, W.; Jin, S.; Jones, G.D.; Chen, J.; Taylor, J.; Li, Z.; Frewer, L.J. Precision agriculture technology adoption: A qualitative study of small-scale commercial “family farms” located in the North China Plain. *Precis. Agric.* **2022**, *23*, 319–351. [\[CrossRef\]](#)
3. Soumitra, N. A vision of precision agriculture: Balance between agricultural sustainability and environmental stewardship. *Agron. J.* **2023**, *116*, 1126–1143.
4. Liu, J.; Xiang, J.; Jin, Y.; Liu, R.; Yan, J.; Wang, L. Boost precision agriculture with unmanned aerial vehicle remote sensing and edge intelligence: A survey. *Remote Sens.* **2021**, *13*, 4387. [\[CrossRef\]](#)
5. Leroux, C.; Jones, H.; Pichon, L.; Guillaume, S.; Lamour, J.; Taylor, J.; Naud, O.; Crestey, T.; Lablee, J.-L.; Tisseyre, B. GeoFIS: An open-source decision-support tool for precision agriculture data. *Agriculture* **2018**, *8*, 73. [\[CrossRef\]](#)
6. Guo, J.; Li, X.; Li, Z.; Hu, L.; Yang, G.; Zhao, C.; Fairbairn, D.; Watson, D.; Ge, M. Multi-GNSS precise point positioning for precision agriculture. *Precis. Agric.* **2018**, *19*, 895–911. [\[CrossRef\]](#)
7. Wolfert, S.; Ge, L.; Verdouw, C.; Bogaardt, M.-J. Big data in smart farming—A review. *Agric. Syst.* **2017**, *153*, 69–80. [\[CrossRef\]](#)
8. Toscano, F.; Fiorentino, C.; Capece, N. Unmanned Aerial Vehicle for Precision Agriculture: A Review. *IEEE Access* **2024**, *12*, 69188–69205. [\[CrossRef\]](#)
9. Adewusi, A.O. AI in precision agriculture: A review of technologies for sustainable farming practices. *World J. Adv. Res. Rev.* **2024**, *21*, 2276–2285. [\[CrossRef\]](#)
10. Krishna, K.R. *Agricultural Drones: A Peaceful Search*; CRC Press: Boca Raton, FL, USA, 2018.
11. Marinello, F.; Pezzuolo, A.; Chiumenti, A.; Sartori, L. Technical analysis of unmanned aerial vehicles (drones) for agricultural applications. *Eng. Rural Dev.* **2016**, *15*, 870–875.
12. Veroustraete, F. The rise of drones in agriculture. *EC Agric.* **2015**, *2*, 325–327.

13. Daponte, P.; Vito, L.; Glielmo, L.; Iannelli, L.; Liuzza, D.; Picariello, F.; Silano, G. A review on the use of drones for precision agriculture. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *275*, 012022. [\[CrossRef\]](#)
14. Furnitto, N.; Ramírez-Cuesta, J.M.; Sottosanti, G.; Longo, D.; Schillaci, G.; Failla, S. Potentiality of Multispectral Vegetation Indexes for Evaluating the Influence of the Sowing Technique on Durum Wheat Cultivation Density. In Proceedings of the Safety, Health and Welfare in Agriculture and Agro-Food Systems, Ragusa, Italy, 6–9 September 2023; pp. 313–321.
15. Manfreda, S. On the use of unmanned aerial systems for environmental monitoring. *Remote Sens.* **2018**, *10*, 641. [\[CrossRef\]](#)
16. Payares, A.; Tarquis, A.H.; Peraló, R.; Canó, J.; Cámara, J.; Gómez del Campo, M. Multispectral and thermal sensors on board UAVs for heterogeneity in Merlot vineyard detection: Contribution to zoning maps. *Remote Sens.* **2023**, *15*, 4024. [\[CrossRef\]](#)
17. Pérez-Zavala, R.; Torres-Torriti, M.; Cheein, F.A.; Troni, G. A pattern recognition strategy for the visual detection of grape clusters in vineyards. *Comput. Electron. Agric.* **2018**, *151*, 136–149. [\[CrossRef\]](#)
18. Sozzi, M.; Cantalamessa, S.; Cogato, A.; Kayad, A.; Marinello, F. Automatic cluster detection in white grape varieties using YOLOv3, YOLOv4, and YOLOv5 deep learning algorithms. *Agronomy* **2022**, *12*, 319. [\[CrossRef\]](#)
19. Squeri, C.; Poni, S.; Matese, A.; Gatti, M. Comparison and verification of the terrain of different remote and proximal sensing platforms to characterize variability in a hedgerow-grown vineyard. *Remote Sens.* **2021**, *13*, 2056. [\[CrossRef\]](#)
20. Cogato, A.; Meggio, F.; Collins, C.; Marinello, F. Sentinel-2 medium-resolution multispectral data to assess damage and recovery time from late frosts in vineyards. *Remote Sens.* **2020**, *12*, 1896. [\[CrossRef\]](#)
21. Payne, A.B.; Walsh, K.B.; Subedi, P.P.; Jarvis, D. Estimation of mango crop yield using image analysis—Segmentation method. *Comput. Electron. Agric.* **2013**, *91*, 57–64. [\[CrossRef\]](#)
22. Robson, A.; Rahman, M.M.; Muir, J. Using worldview satellite imagery to map yield in avocado (*Persea americana*): A case study in Bundaberg, Australia. *Remote Sens.* **2017**, *9*, 1223. [\[CrossRef\]](#)
23. Shi, R.; Li, T.; Yamaguchi, Y. An attribution-based pruning method for real-time mango detection with YOLO network. *Comput. Electron. Agric.* **2020**, *169*. [\[CrossRef\]](#)
24. Ballesteros, R.; Intrigliolo, D.S.; Ortega, J.F.; Ramírez-Cuesta, J.M.; Buesa, I.; Moreno, M.A. Vineyard yield estimation by combining remote sensing, computer vision and artificial neural network techniques. *Precis. Agric.* **2020**, *21*, 1242–1262. [\[CrossRef\]](#)
25. Wang, Y.; Yi, Y.; Wang, X.-F.; Chen, J.; Li, X. Fig fruit recognition method based on YOLOv4 deep learning. In Proceedings of the 2021 18th International Conference on Electrical Engineering, Electronics, Computer, Telecommunications and Information Technology: Smart Electrical System and Technology, ECTI-CON, Bangkok, Thailand, 24–26 March 2021; pp. 303–306. [\[CrossRef\]](#)
26. Íñiguez, R.; Gutiérrez, S.; Poblete-Echeverría, C.; Hernández, I.; Barrio, I.; Tardáguila, J. Deep learning modelling for non-invasive grape bunch detection under diverse occlusion conditions. *Comput. Electron. Agric.* **2024**, *226*, 109421. [\[CrossRef\]](#)
27. Jiménez, A.; Ceres, R.; Pons, J.L. A study of computer vision methods for locating fruit on trees. *Trans. ASAE* **2000**, *43*, 1911–1920. [\[CrossRef\]](#)
28. Xiong, J.; Lin, R.; Liu, Z.; He, Z.; Tang, L.; Yang, Z.; Zou, X. The recognition of groups of lychee and the calculation of the collection point in a natural night-time environment. *Biosyst. Eng.* **2018**, *166*, 44–57. [\[CrossRef\]](#)
29. Wang, W.; Shi, Y.; Liu, W.; Che, Z. An unstructured orchard grape detection method utilizing YOLOv5s. *Agriculture* **2024**, *14*, 262. [\[CrossRef\]](#)
30. Gavrilović, M.; Jovanović, D.; Božović, P.; Benka, P.; Govedarica, M. Vineyard Zoning and Vine Detection Using Machine Learning in Unmanned Aerial Vehicle Imagery. *Remote Sens.* **2024**, *16*, 584. [\[CrossRef\]](#)
31. Sun, L.; Gao, F.; Anderson, M.C.; Kustas, W.P.; Alsina, M.M.; Sanchez, L.; Sams, B.; McKee, L.; Dulaney, W.; White, W.A.; et al. Daily Mapping of 30 m LAI and NDVI for Grape Yield Prediction in California Vineyards. *Remote Sens.* **2017**, *9*, 317. [\[CrossRef\]](#)
32. Guan, S.; Fukami, K.; Matsunaka, H.; Okami, M.; Tanaka, R.; Nakano, H.; Sakai, T.; Nakano, K.; Ohdan, H.; Takahashi, K. Assessing Correlation of High-Resolution NDVI with Fertilizer Application Level and Yield of Rice and Wheat Crops Using Small UAVs. *Remote Sens.* **2019**, *11*, 112. [\[CrossRef\]](#)
33. Labus, M.P.; Nielsen, G.A.; Lawrence, R.L.; Engel, R.; Long, D.S. Wheat Yield Estimates Using Multi-Temporal NDVI Satellite Imagery. *Int. J. Remote Sens.* **2002**, *23*, 4169–4180. [\[CrossRef\]](#)
34. Vannoppen, A.; Gobin, A. Estimating Farm Wheat Yields from NDVI and Meteorological Data. *Agronomy* **2021**, *11*, 946. [\[CrossRef\]](#)
35. Karkauskaite, P.; Tagesson, T.; Fensholt, R. Evaluation of the Plant Phenology Index (PPI), NDVI and EVI for Start-of-Season Trend Analysis of the Northern Hemisphere Boreal Zone. *Remote Sens.* **2017**, *9*, 485. [\[CrossRef\]](#)
36. Jaafar, H.H.; Zurayk, R.; King, C.; Ahmad, F.; Al-Outa, R. Impact of the Syrian Conflict on Irrigated Agriculture in the Orontes Basin. *Int. J. Water Resour. Dev.* **2015**, *31*, 436–449. [\[CrossRef\]](#)
37. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016.
38. Santos, T.T.; de Souza, L.L.; dos Santos, A.A.; Ávila, S. Detection, segmentation and tracking of grapes using deep neural networks and three-dimensional association. *Comput. Electron. Agric.* **2020**, *170*, 105247. [\[CrossRef\]](#)
39. Liu, G.; Nouaze, J.T.; Kim, J.P. A robust algorithm for tomato detection based on YOLOv3. *Sensors* **2020**, *20*, 2145. [\[CrossRef\]](#)

40. Dong, X.; Lan, J.; Zhu, Y. Method for identifying the collection position of lychee based on YOLOv5 and PSPNet. *Remote Sens.* **2022**, *14*, 2004.
41. Regulatory Council of the Protected Designation of Origin Vinalopó Bagged Table Grapes. Available online: <https://uva-vinalopo.org> (accessed on 3 October 2024).
42. General Directorate of Rural Development and Common Agricultural Policy. *Specifications of the Protected Designation of Origin Vinalopó Bagged Table Grapes*; Ciutat Administrativa 9 d'Octubre: Valencia, Italy, 2015.
43. Koirala, A.; Jha, M.; Bodapati, S.; Mishra, A.; Chetty, G.; Sahu, P.K.; Mohanty, S.; Padhan, T.K.; Mattoo, J.; Hukkoo, A. Deep learning for real-time malaria parasite detection and counting using YOLO-mp. *IEEE Access* **2022**, *10*, 102157–102172. [\[CrossRef\]](#)
44. Muhlashin, M.N.I.; Stefanie, A. Klasifikasi penyakit mata berdasarkan citra fundus menggunakan YOLO V8. *J. Mahasiswa Tek. Inform.* **2023**, *7*, 1363–1368. [\[CrossRef\]](#)
45. Badeka, E.; Karapatzak, E.; Karampatea, A.; Bouloumpasi, E.; Kalathas, I.; Lytridis, C.; Tziolas, E.; Tsakalidou, V.N.; Kaburlasos, V.G. A deep learning approach for precision viticulture: Assessing grape maturity via YOLOv7. *Sensors* **2023**, *23*, 8126. [\[CrossRef\]](#)
46. Codes-Alcaraz, A.; Puerto-Molina, H.; Rocamora-Osorio, C.; Furnitto, N.; Failla, S.; Ramírez-Cuesta, J.M. Deciphering the role of UAV flight height and sensor inclination in the prediction of table grape production through computer vision techniques. In Proceedings of the IEEE International Workshop on Metrology for Agriculture and Forestry, MetroAgriFor 2024, Padua, Italy, 29–31 October 2024. Proceedings, 2024, in press.
47. Vijayasekaran, D. SEN2-AGRI—Crop type mapping pilot study using Sentinel-2 satellite imagery in India. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2019**, *XLII-3/W6*, 175–180. [\[CrossRef\]](#)
48. Sentinel User Handbook and Exploitation Tools (SUHET). Sentinel-2 user handbook (Issue: 1, Revision: 2). European Space Agency Standard Document 2015. Available online: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook (accessed on 5 December 2024).
49. Planet Labs. A Bigger Crayon Box with SuperDove's 8 Spectral Bands. In *Manual Planet*; Planet Labs: San Francisco, CA, USA, 2024.
50. Planet Team. Planet Team. Planet Application Program Interface. In *Space for Life on Earth*; Planet Team: San Francisco, CA, USA, 2017.
51. Waltuch, M.; Cameron, E.; Lafframboise, A.; Zeiler, M. *Exploring ArcObjects Vol. 1—Applications and Cartography*; ESRI Press: Redlands, CA, USA, 2001.
52. Farmonov, N.; Amankulova, K.; Szatmári, J.; Urinov, J.; Narmanov, Z.; Nosirov, J.; Mucsi, L. Combining PlanetScope and Sentinel-2 Images with Environmental Data for Improved Wheat Yield Estimation. *Int. J. Digit. Earth* **2023**, *16*, 847–867. [\[CrossRef\]](#)
53. Amankulova, K.; Farmonov, N.; Abdelsamei, E.; Szatmári, J.; Khan, W.; Zhuran, M.; Rustamov, J.; Akhmedov, S.; Sarimsakov, M.; Mucsi, L. A Novel Fusion Method for Soybean Yield Prediction Using Sentinel-2 and PlanetScope Imagery. *Remote Sens.* **2024**, *17*, 13694–13707. [\[CrossRef\]](#)
54. Angelini, R.; Angelats, E.; Luzi, G.; Ribas, F.; Masiero, A. Shoreline extraction methods from Sentinel-2 and PlanetScope images. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2024**, *XLVIII-1-2024*, 1–6. [\[CrossRef\]](#)
55. Furnitto, N.; Ramírez-Cuesta, J.M.; Sottosanti, G.; Schillaci, G.; Failla, S. Appraising the use of remote and proximal platforms for wheat biomass estimation from multispectral imagery. In Proceedings of the International Mid-Term Conference 2024 of the Italian Association of Agricultural Engineering (AIIA), Padova, Italy, 17–19 June 2024; Lecture Notes in Civil Engineering (LNCE). Springer Nature: Cham, Switzerland, 2025. in press.
56. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring vegetation systems in the Great Plains with ERTS. In Proceedings of the Third Earth Resources Technology Satellite—1 Symposium, Washington, DC, USA, 10–14 December 1973. NASA SP-351.
57. Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, *58*, 289–298. [\[CrossRef\]](#)
58. Huete, A.R. A Soil-Adjusted Vegetation Index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [\[CrossRef\]](#)
59. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A Modified Soil Adjusted Vegetation Index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [\[CrossRef\]](#)
60. Haboudane, D. Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [\[CrossRef\]](#)
61. Huete, A.R.; Liu, H.Q.; Batchily, K.; van Leeuwen, W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* **1997**, *59*, 440–451. [\[CrossRef\]](#)
62. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [\[CrossRef\]](#)
63. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [\[CrossRef\]](#)
64. McFeeters, S.K. The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [\[CrossRef\]](#)

65. Wu, D.; Jiang, S.; Zhao, E.; Liu, Y.; Zhu, H.; Wang, W.; Wang, R. Detection of *Camellia oleifera* fruit in complex scenes by using YOLOv7 and data augmentation. *Appl. Sci.* **2022**, *12*, 11318. [\[CrossRef\]](#)
66. Ortuani, B.; Mayer, A.; Bianchi, D.; Sona, G.; Crema, A.; Modina, D.; Bolognini, M.; Brancadoro, L.; Boschetti, M.; Facchi, A. Effectiveness of management zones delineated from UAV and Sentinel-2 data for precision viticulture applications. *Remote Sens.* **2024**, *16*, 635. [\[CrossRef\]](#)
67. Khaliq, A.; Comba, L.; Biglia, A.; Ricauda Aimonino, D.; Chiaberge, M.; Gay, P. Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sens.* **2019**, *11*, 436. [\[CrossRef\]](#)
68. Matese, A.; Toscano, P.; Di Gennaro, S.; Genesio, L.; Vaccari, F.; Primicerio, J.; Belli, C.; Zaldei, A.; Bianconi, R.; Gioli, B. Intercomparison of UAV, aircraft, and satellite remote sensing platforms for precision viticulture. *Remote Sens.* **2015**, *7*, 2971–2990. [\[CrossRef\]](#)
69. Di Gennaro, S.F.; Toscano, P.; Gatti, M.; Poni, S.; Berton, A.; Matese, A. Spectral comparison of UAV-based hyper and multispectral cameras for precision viticulture. *Remote Sens.* **2022**, *14*, 449. [\[CrossRef\]](#)
70. Lucero, M.F.; Hernández, C.M.; Carcedo, A.J.P.; Zajdband, A.; Guillevic, P.C.; Houborg, R.; Hamilton, K.; Ciampitti, I.A. Enhancing alfalfa biomass prediction: An innovative framework using remote sensing data. *Remote Sens.* **2024**, *16*, 3379. [\[CrossRef\]](#)
71. Wang, Z.; Wu, W.; Liu, H. Spatial estimation of soil organic carbon content utilizing PlanetScope, Sentinel-2, and Sentinel-1 data. *Remote Sens.* **2024**, *16*, 3268. [\[CrossRef\]](#)
72. Manivasagam, V.S.; Sadeh, Y.; Kaplan, G.; Bonfil, D.J.; Rozenstein, O. Studying the feasibility of assimilating Sentinel-2 and PlanetScope imagery into the SAFY crop model to predict within-field wheat yield. *Remote Sens.* **2021**, *13*, 2395. [\[CrossRef\]](#)
73. Huang, J.; Ma, H.; Sedano, F.; Lewis, P.; Liang, S.; Wu, Q.; Su, W.; Zhang, X.; Zhu, D. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST–PROSAIL model. *Eur. J. Agron.* **2019**, *102*, 1–13. [\[CrossRef\]](#)
74. Zhou, T.; Geng, Y.; Ji, C.; Xu, X.; Wang, H.; Pan, J.; Bumberger, J.; Haase, D.; Lausch, A. Prediction of soil organic carbon and the C:N ratio on a national scale using machine learning and satellite data: A comparison between Sentinel-2, Sentinel-3 and Landsat-8 images. *Sci. Total Environ.* **2021**, *755*, 142661. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Ziliani, M.G.; Altaf, M.U.; Aragon, B.; Houborg, R.; Franz, T.E.; Lu, Y.; Sheffield, J.; Hoteit, I.; McCabe, M.F. Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model. *Agric. For. Meteorol.* **2021**, *313*, 108736. [\[CrossRef\]](#)
76. Li, F.; Miao, Y.; Chen, X.; Sun, Z.; Stueve, K.; Yuan, F. In-season prediction of corn grain yield through PlanetScope and Sentinel-2 images. *Agronomy* **2022**, *12*, 3176. [\[CrossRef\]](#)
77. Rouault, P.; Courault, D.; Pouget, G.; Flamain, F.; Diop, P.-K.; Desfonds, V.; Doussan, C.; Chanzy, A.; Debolini, M.; McCabe, M.; et al. Phenological and biophysical Mediterranean orchard assessment using ground-based methods and Sentinel 2 data. *Remote Sens.* **2024**, *16*, 3393. [\[CrossRef\]](#)
78. Rouault, P.; Courault, D.; Flamain, F.; Pouget, G.; Doussan, C.; Lopez-Lozano, R.; McCabe, M.; Debolini, M. High-resolution satellite imagery to assess orchard characteristics impacting water use. *Agric. Water Manag.* **2024**, *295*, 108763. [\[CrossRef\]](#)
79. Elfarkh, J.; Johansen, K.; El Hajj, M.M.; Almashharawi, S.K.; McCabe, M.F. Evapotranspiration, Gross Primary Productivity, and Water Use Efficiency over a High-Density Olive Orchard Using Ground and Satellite-Based Data. *Agric. Water Manag.* **2023**, *287*, 108423. [\[CrossRef\]](#)
80. Padilla-Díaz, C.M.; Rodríguez-Domínguez, C.M.; Hernández-Santana, V.; Pérez-Martin, A.; Fernandes, R.D.M.; Montero, A.; García, J.M.; Fernández, J.E. Water status, gas exchange and crop performance in a super high density olive orchard under deficit irrigation scheduled from leaf turgor measurements. *Agric. Water Manag.* **2018**, *202*, 241–252. [\[CrossRef\]](#)
81. Rodenhizer, H.; Yang, Y.; Fiske, G.; Potter, S.; Windholz, T.; Mullen, A.; Watts, J.D.; Rogers, B.M. A Comparison of Satellite Imagery Sources for Automated Detection of Retrogressive Thaw Slumps. *Remote Sens.* **2024**, *16*, 2361. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



2024 IEEE INTERNATIONAL WORKSHOP ON

Metrology for Agriculture and Forestry

PADOVA, ITALY / OCTOBER 29-31, 2024

**[PROCEEDINGS]**

© 2024. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to use any copyrighted component of this work in other works must be obtained from the IEEE.

IEEE Catalogue Number
CFP24U22-USBISBN
979-8-3503-5543-7

© 2024 IEEE

Proceedings of 2024 IEEE International Workshop on Metrology for Agriculture and Forestry

IEEE MetroAgriFor 2024

© 2024. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to use any copyrighted component of this work in other works must be obtained from the IEEE.

Copyright and Reprint Permission: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For reprint or republication permission, email to IEEE Copyrights Manager at pubs-permissions@ieee.org. All rights reserved. Copyright ©2024 by IEEE.

IEEE Catalog Number: CFP24U22-USB

ISBN: 979-8-3503-5543-7

WORKSHOP PROGRAM

Tuesday, October 29

Plenary Session - Keynote Speaker

Room: Sala Luna

1 The Role of Geometrical Metrology in Manufacturing-An Update

Leonardo De Chiffre and Danilo Quagliotti (Technical University of Denmark, Denmark)

Session 1.1 - Contactless Measurements and Computer Vision for Agriculture, Forestry and Zootechnics - PART I

Room: Sala Luna

7 Rapid Monitoring of Maize Plant Status Using UAV-Based Multispectral and Hyperspectral Data

Massimiliano Gargiulo, Claudia Savarese, Marco De Mizio, Francesco Tufano and Sara Parrilli (CIRA - Italian Aerospace Research Centre, Italy)

13 An Energy-Based Approach to Cushioning Material Selection for Table Olive Mechanical Harvesting

Alessandro Annessi, Francesco Belluccini, Vittoria Medici, Veronica Giorgi and Enrico Maria Lodolini (Università Politecnica Delle Marche, Italy); Milena Martarelli and Paolo Castellini (Università Politecnica delle Marche, Italy); Davide Neri (Università Politecnica Delle Marche, Italy)

19 Deciphering the Role of UAV Flight Height and Sensor Inclination in the Prediction of Table Grape Production Through Computer Vision Techniques

Ana Maria Codes Alcaraz and Herminia Puerto Molina (University Miguel Hernandez & Centre for Agri-Food and Agro-Environmental Research and Innovation CIAGRO, Spain); Carmen Rocamora Osorio (University of Miguel Hernandez & Centre for Agri-Food and Agro-Environmental Research and Innovation CIAGRO, Spain); Nicola Furnitto (Università degli studi di Catania, Italy); Sabina Failla (University of Catania, Italy); Juan Miguel Ramírez Cuesta (Università degli Studi di Catania, Italy)

24 Olive Tree Metrics Assessment by UAV Photogrammetric Approach

Francesco Belluccini and Alessandro Annessi (Università Politecnica Delle Marche, Italy); Roberta Cacciatore (Università Politecnica delle Marche, Italy); Veronica Giorgi and Enrico Maria Lodolini (Università Politecnica Delle Marche, Italy); Milena Martarelli and Paolo Castellini (Università Politecnica delle Marche, Italy); Davide Neri (Università Politecnica Delle Marche, Italy)

Session 1.2 - Soil monitoring

Room: Sala Ganimede

29 Experimental Investigations on the Conditions Where Radial Expansion Benefits Soil Penetration for Plant Root-Inspired Diggers

Venkata Rithwick Puranam (Scuola Superiore Sant'Anna, Italy); Jules Sebastiaan de Ruiter (Delft University of Technology, The Netherlands); Serena Rosa Maria Pirrone (Italian Institute of Technology & Scuola Superiore Sant'Anna, Italy); Emanuela Del Dottore (IIT, Italy); Barbara Mazzolai (Istituto Italiano di Tecnologia, Italy)

35 Predictive Modeling of Soil Chromium Content Using Hyperspectral Spectroscopy

Mariam Alcibahy and Abdel Rahman Satei Hussein Alsaleh (Khalifa University, United Arab Emirates); Hamed Al Hashmi (UAE Space Agency, United Arab Emirates); Bayan Athamneh (Environment Agency Abu Dhabi, United Arab Emirates); Ali Al Hammadi and Lakmal Seneviratne (Khalifa University, United Arab Emirates); Maryam Rashed AlShehhi (Khalifa University, Civil, Infrastructure and Environmental Engineering & MIT, United Arab Emirates)

Deciphering the role of UAV flight height and sensor inclination in the prediction of table grape production through computer vision techniques

Ana María Codes Alcaraz
Engineering Department
University Miguel Hernandez
Orihuela, Spain
acodes@umh

Herminia Puerto Molina
Engineering Department
University Miguel Hernandez
Orihuela, Spain
hpuerto@umh.es

Carmen Rocamora Osorio
Engineering Department
University Miguel Hernandez
Orihuela, Spain
rocamora@umh.es

Nicola Furnitto
Department of Agriculture, Food and
Environment (Di3A)
University of Catania
Catania, Italy
nicola.furnitto@phd.unict.it

Sabina Failla
Department of Agriculture, Food and
Environment (Di3A)
University of Catania
Catania, Italy
sfailla@unict.it

Juan Miguel Ramírez-Cuesta
Department of Agriculture, Food and
Environment (Di3A)
University of Catania
Catania, Italy
ramirezcueta.jm@gmail.com

Abstract— Due to the new and promising technological advances in Artificial Intelligence (AI), there is a great opportunity to apply these advancements in agriculture. One of the various issues that arise in agriculture is production estimation, as this data allows for a more efficient crop management, optimizing harvest planning, and reducing the need for intensive labor. This study evaluates a computer vision-based methodology for predicting the production of bagged table grapes in the Vinalopó region through oblique RGB images acquired from an unmanned aerial vehicle. Specifically, two different configurations were used for the acquisition of the images: (i) flight altitude (h) of 10 m and sensor tilt (θ) of 35° ; and (ii) h of 17 m and θ of 45° . The research also presents a comparison of the latest versions of YOLO, where YOLOv8 offers the best performance in terms of precision and efficiency, particularly at a flight altitude of 10 meters and a camera inclination of 35° with a Root Mean Square Error (RMSE) of 0.62, and at 17 m and 45° with a RMSE of 1.39. This (RMSE) is correlated with the number of grapes in the plot. However, YOLO v.7x showed greater consistency and precision at both altitudes with a RMSE of 0.75 at 10 m and 0.78 at 17 m, making it the most robust and versatile model for both altitudes. The implementation of these computer vision algorithms, supported by drone technology, not only allows for more effective monitoring of large cultivation areas but also provides real-time information that is crucial for quick and effective decision-making in the agricultural field.

Keywords—YOLO, prediction, production, grape, drone.

I. INTRODUCTION

The use of artificial intelligence (AI) in agriculture is gaining importance, presenting a significant potential for various practical applications. One of the uses of AI in agriculture is crop yield prediction, either through an algorithm trained with soil properties or by counting flowers [1, 2].

Estimating agricultural production is critical for a proper management of a farm. However, this task is often laborious and requires time, experienced personnel or specific training.

Numerous studies have focused on estimating production by detecting fruits on trees. There are specific examples for cherries, grapes, avocados, mangoes, and other types of fruits, where agricultural production prediction is carried out through the detection of trees using aerial images taken with unmanned aerial vehicles (UAV) and the subsequent counting of these trees or the detection of fruits using images captured from the ground [2-6].

In the specific case of grapes, several studies have successfully identified and detected grape clusters by training algorithms with images taken from the ground. However, this approach has the disadvantage of requiring a high amount of time for the image acquisition process [7, 8].

In the province of Alicante, there are emblematic crops of high economic value, such as the bagged table grapes of Vinalopó, which have a Protected Designation of Origin (PDO). The bagged table grapes are exclusive to this province and are characterized by their unique ripening process, in which the bagging technique protects the grapes and provides qualities such as sweetness, thin skin, and natural preservation, allowing for later harvest dates. Accurate harvest prediction of these grapes is crucial, as it can significantly improve yield management, market planning, and economic returns, further reinforcing the importance of this crop for the entire region.

The main objective of this work was to evaluate an effective and rapid computer vision-based methodology to estimate the production of grapes in plots belonging to the PDO of bagged table grapes of Vinalopó. This main aim included two specific objectives: (i) evaluate crop yield estimation by using two different UAV flight height-tilt configurations; and (ii) assess the influence that the version of the object detection algorithm (YOLO) had on the final estimated yield.

II. MATERIALS AND METHODS

A. Description of the study area

The study plot (1.03 ha) is located in the municipality of Agost (Alicante, Spain) (Figure 1). The planting frame is 1.8 x 2 m with a total of 2.742 plants and a registered production of 12000 kg and average cluster weight of 598 g. All grape clusters are bagged in white paper bags, following standard practice and complying with the PDO requirements for Vinalopó Table Grapes.

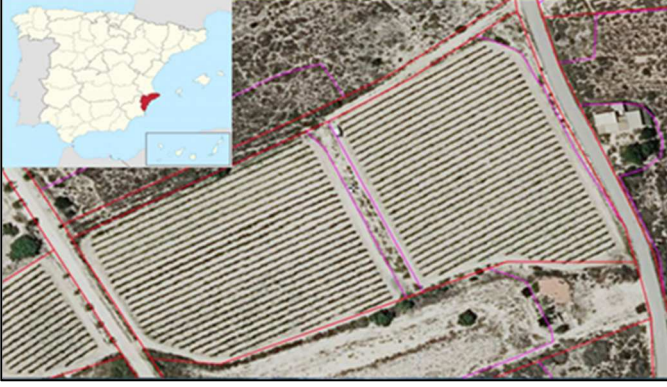


Fig. 1. Study site location at Alicante, Spain ($38^{\circ}25'5.46''N$; $0^{\circ}38'22.56''W$).

B. UAV flight campaign

The aerial images were taken on October 27, 2022, using a drone model Autel EVO 2 Pro (UUC Technology Co., Ltd., Hong Kong), equipped with a 1-inch RGB CMOS camera (Sony Group Corporation, Tokyo), adjustable aperture F2.8-F11 capable of 20 MP resolution and 4K/30 FPS video recording.

The drone flights were conducted following a path parallel to the crop lines using two sensor height (h) - tilt (θ) configurations: (i) $h=10$ m and $\theta=35^{\circ}$; and (ii) $h=17$ m and $\theta=45^{\circ}$ (Figure 2).

C. YOLO model

The UAV images were subsequently processed using the object detection algorithm known as YOLO. This algorithm is a state-of-the-art single-stage object detection model that approaches detection as a regression problem [9]. This model can detect multiple objects in a single image by dividing the image into a grid. Initially, each grid cell predicts the probability of a class. Then, the model determines bounding box coordinates and calculates confidence scores for each. As a single-stage detector, YOLO models achieve high computational speed [9].

In this research, three of the most recent versions of the YOLO model were considered, including YOLO v.7x, YOLO v.7-E6, and YOLO v.8. YOLO v.7x is recognized as the official model by the computer vision and machine learning communities, as it was developed by the same group of researchers who initially proposed this idea [10]. On the other hand, YOLO v.8 was developed by [11] and combines the best features from various real-time object detection approaches [12]. YOLO v.7-E6 is obtained through a new composite scaling method from YOLO v7-W6 and YOLO v.7x is also obtained by the composite scaling method via YOLO v.7 [13]. The YOLO v.7 model outperforms all previous object detection techniques, including earlier versions of the YOLO

algorithms, in terms of speed and accuracy [13]. As for the YOLO v.8 model, it introduces several improvements such as mosaic augmentation, C3 convolutions, and anchor-free detection to optimize performance and inference speed.

In this study, the performance of the YOLO v.7x, YOLO v.7-E6, and YOLO v.8 models was evaluated. The models were trained on the MS COCO (Microsoft Common Objects in Context) dataset [14] using pre-trained weights.

The models were retrained with our specific dataset and we compared the performance of YOLO v.7x, YOLO v.7-E6 and YOLO v.8 in detecting and locating grape clusters in the plot.

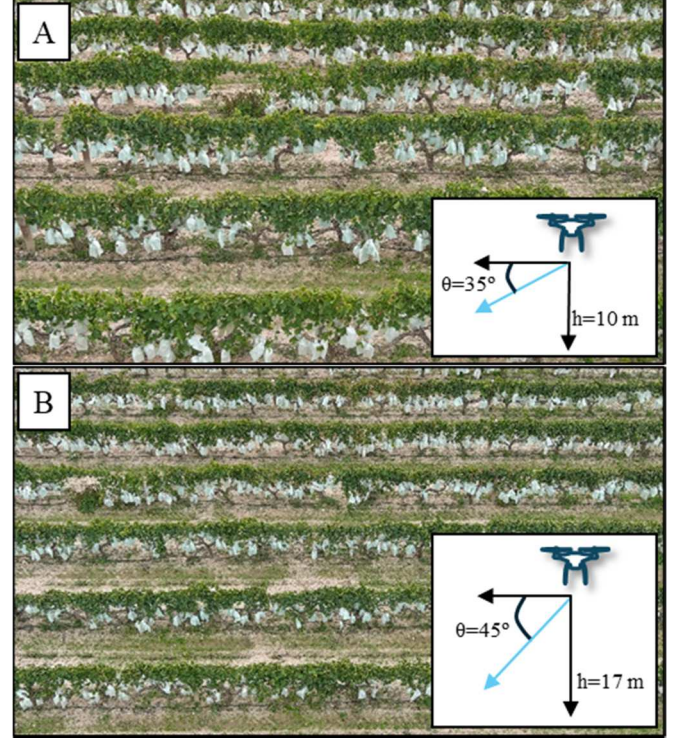


Fig. 2. UAV images acquired with (A) the $h=10$ m and $\theta=35^{\circ}$ configuration; and (B) the $h=17$ m and $\theta=45^{\circ}$ configuration.

D. Model training

The UAV images acquired from each flight configuration were labelled using the Makesense application, where each cluster was annotated with rectangles. A total of 58 images were labelled, out of which 90% were used for training the model and 10% for validation [15, 16]. These images were annotated in YOLO format.

The training procedure was performed using Google Colaboratory, a cloud product from Google Research that allows any user to write and execute arbitrary Python code in the browser.

The YOLO v.7x and YOLO v.8 versions were trained with an image size of 640 pixels, as that is the size these models are originally trained on [11-13]. For YOLO v.7-E6, an image size of 1280 pixels was used, as this version is trained for that size [13]. The number of epochs, referring to the iterations of the training algorithm, was set to a minimum of 250 in this study, as this was the point at which the algorithm produced the best results.

Lastly, the grape clusters number and the total yield estimated for each different flight configuration and YOLO version were compared to the cluster number visually identified in each selected image and with the yield provided by the cooperative. To make an exact comparison between the different versions, 50 images were selected: 25 images with the $h=10\text{ m}$ and $\theta=35^\circ$ configuration; and 25 images with the $h=17\text{ m}$ and $\theta=45^\circ$ configuration. These images obtained from each flight configuration were cropped to match the same location at both heights in order to determine which height was the most effective for counting grape clusters.

E. Statistical analysis

Linear relationships between the estimated grapes cluster number estimated at the different UAV flight configurations, different YOLO versions and the cluster number visually identified were performed. The linear trend, the coefficient of determination (R^2) and the Root Mean Square Error (RMSE) were calculated for evaluating the agreement between the assessed methodologies.

Moreover, the differences between the total yield derived from the estimated number of cluster and the yield provided by the cooperative, were analysed by means of an Analysis of Variance (ANOVA). The mean separation was performed following the Two-sided Dunnett's Multiple Comparisons with a Control (i.e. yield provided by the cooperative) ($p \leq 0.05$).

III. RESULTS

Once the different versions of YOLO were trained, indicators such as precision, recall, mean average precision (mAP), and training time were used to evaluate the model's efficiency. The mAP value is the area under the Precision-Recall curve, and a higher value means better model performance [5].

The results obtained (Figure 3) showed that the best-performing model in terms of training was the latest version, YOLO v.8, with a mAP of 0.849, followed by YOLO v.7-E6 (mAP = 0.821), and YOLO v.7x (mAP = 0.806).

As for the training time, YOLO v.8 was the fastest version taking 50 min; whereas the YOLO v.7x took 1 hour and YOLO v.7-E6 took 2 hours.

For all evaluated YOLO versions, the estimated number of clusters per vine was more accurate using the $h=10\text{m}$ and $\theta=35^\circ$ configuration than when using the $h=17\text{m}$ and $\theta=45^\circ$ one. More specifically, the $h=10\text{m}$ and $\theta=35^\circ$ configuration provided slope terms in the range 0.90-1.03, intercepts lower than 1.00, R^2 values between 0.79 and 0.84 and RMSE values lower than 1.17 clusters vine⁻¹ (Table 1). Contrarily, the $h=17\text{m}$ and $\theta=45^\circ$ configuration resulted in lower slope terms (0.56-0.77) and higher intercept terms (except for Yolo v.7-E6; Table 1), lower R^2 values (0.53-0.71) and higher RMSE values (0.78-2.36 clusters vine⁻¹) than the $h=10\text{m}$ and $\theta=35^\circ$ configuration.

These results evidence the better capability of the model for identifying the clusters at low altitudes and lower inclination angles probably due to the greater proximity to objects, which facilitates their precise detection. However, the cluster identification in vines located in the background of the image is more affected by the obstruction of the canopy of closer vines when using the $h=10\text{m}$ and $\theta=35^\circ$ configuration than when using the $h=17\text{m}$ and $\theta=45^\circ$ configuration.

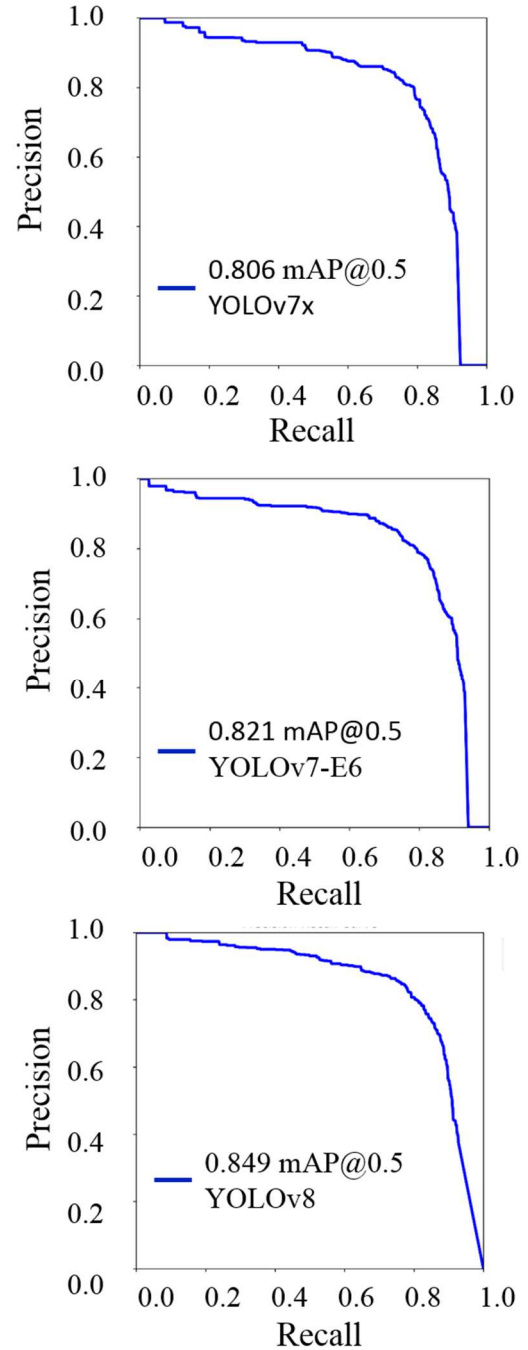


Fig. 3. Precision-Recall graphs obtained for the training of each different YOLO versions.

The accuracy of the estimation of the grapes number per vine was also dependent on the version of the model utilized. In this sense, the best accuracy was obtained with YOLOv.8 and the $h=10\text{m}$ and $\theta=35^\circ$ configuration (RMSE= 0.62 clusters vine⁻¹), although the precision decreased when utilizing the $h=17\text{m}$ and $\theta=45^\circ$ configuration (RMSE= 1.39 clusters vine⁻¹). A more constant behavior was observed with YOLO v.7x, which provided clusters per vine estimations with a RMSE of 0.75-0.78 clusters vine⁻¹, independently on the flight configuration (Table 1). Yolo v.7-E6 provided the worst accuracies among the tested versions, with RMSE values ranging from 1.17 to 2.36 clusters vine⁻¹ for the $h=10\text{m}$ and $\theta=35^\circ$; and the $h=17\text{m}$ and $\theta=45^\circ$ configurations, respectively.

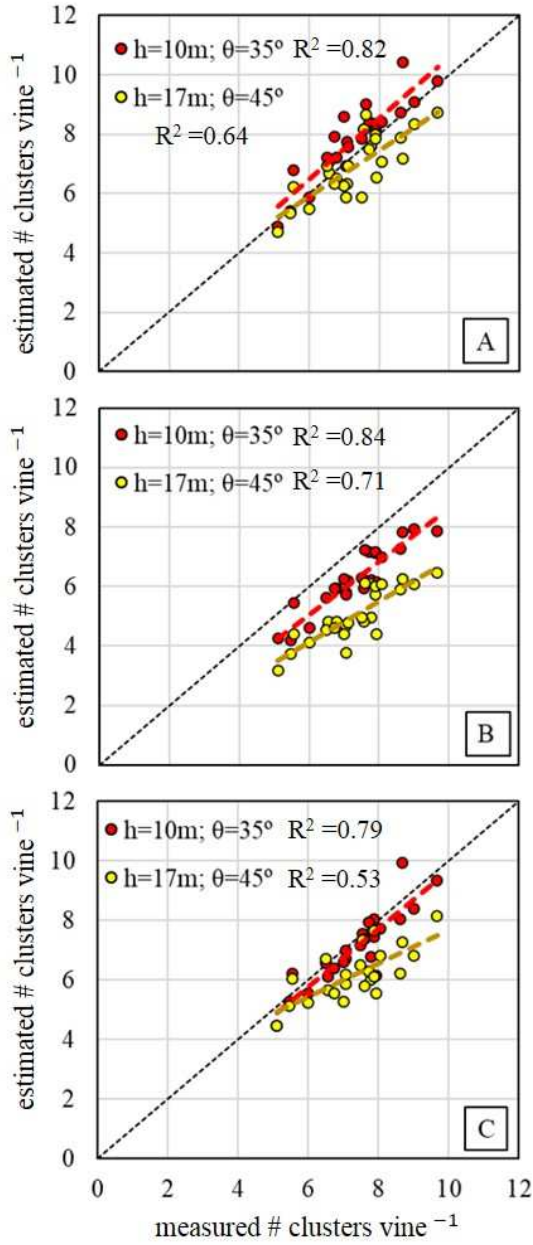


Fig. 4. Correlation between the estimated number of clusters per vine (A: YOLO v.7x; B: YOLO v.7-E6; and C: YOLO v.8) and the number of clusters measured in the field; for each flight configuration.

TABLE I. CORRELATION EQUATIONS, COEFFICIENT OF DETERMINATION (R^2) AND ROOT MEAN SQUARE ERROR (RMSE) FOR THE DIFFERENT YOLO VERSIONS AND FLIGHT CONFIGURATIONS.

Model version	Flight config.	Equation	R^2	RMSE (clusters vine ⁻¹)
v.7x	h=10m θ=35°	$y = 1.03x + 0.29$	0.82	0.75
	h=17m θ=45°	$y = 0.77x + 1.26$	0.64	0.78
v.7-E6	h=10m θ=35°	$y = 0.90x - 0.32$	0.84	1.17
	h=17m θ=45°	$y = 0.69x - 0.04$	0.71	2.36
v.8	h=10m θ=35°	$y = 0.98x - 0.09$	0.79	0.62
	h=17m θ=45°	$y = 0.56x + 2.10$	0.53	1.39

It is important to note that, except for the h=10m and θ=35° configuration using YOLO v.7x, the model underestimated the number of clusters per vine (Figure 4 and Table 1), probably due to the existent overlapping between the clusters.

When evaluating the total yield estimated from the different flight configurations and YOLO versions, it was observed that the yield obtained from YOLO v.7x (8736-9902 kg ha⁻¹; Figure 5) did not differ from the measured yield (9231 kg ha⁻¹; Figure 5) in any of the flight configurations. Similar results were obtained from YOLO v.8 for the h=10 and θ=35° configuration (8933 kg ha⁻¹; Figure 5), whereas a significant yield underestimation was obtained for this version with the h=17 and θ=45° configuration (7819 kg ha⁻¹; Figure 5), and for both flight configurations using YOLO v.7-E6 (6345-7866 kg ha⁻¹; Figure 5).

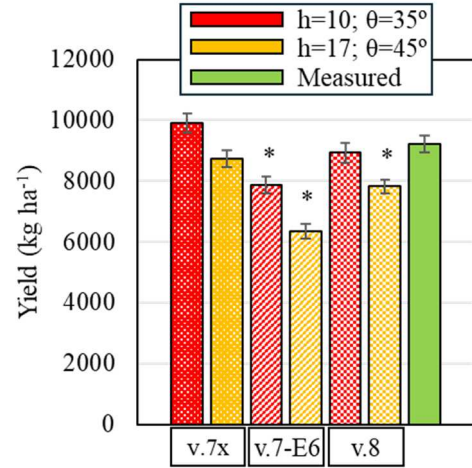


Fig. 5. Yield predicted from each different combination of model version-flight configuration. Asterisk identify significant differences with the yield measured on the field ($p \leq 0.05$).

IV. CONCLUSION

The study presents an evaluation of different flights configuration to estimate the grape clusters number based on computer vision-based techniques. The main conclusions that can be drawn from the present study are:

- The UAV flight configuration of 10m height and 35° sensor tilt provided more accurate grape cluster number estimates than using the flight configuration of 17m height and 45° sensor tilt.
- The use of YOLO v.8 with the 10m height and 35° sensor tilt configuration offered the most accurate yield estimations.
- YOLO v.7x, showed a greater consistency and precision at both evaluated UAV flight configurations making it the most robust and versatile model for different flight conditions.

These findings highlight the importance of considering flight height and sensor tilt when selecting a detection model for predicting grape production in the case of Vinalopó grapes with Protected Designation of Origin (PDO). Additionally, it underscores the advantages of combining UAV technology with detection algorithms to improve agricultural productivity, optimize harvest planning, and reduce the need for intensive labour and product waste. Together, these technological advancements promote smarter and more sustainable agriculture.

REFERENCES

- [1] A. Robson, M. Rahman, J. Muir, "Using Worldview Satellite Imagery to Map Yield in Avocado (*Persea americana*): A Case Study in Bundaberg, Australia", *Remote Sens.*, vol. 9, n.º 12, p. 1223, 2017.
- [2] D. Wu, S. Lv, M. Jiang, H. Song, "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments", *Comput. Electron. Agriculture*, vol. 178, p. 105742, 2020.
- [3] R. Shi, T. Li, Y. Yamaguchi, "An attribution-based pruning method for real-time mango detection with YOLO network", *Comput. Electron. Agriculture*, vol. 169, p. 105214, 2020.
- [4] L.P. Osco, M.D.S. de Arruda, D.N. Gonçalves, A. Dias, J. Batistoti, M. de Souza, F.D.G. Gomes, A.P.M. Ramos, L.A. de Castro Jorge, V. Liesenberg, J. Li, "A CNN approach to simultaneously count plants and detect plantation-rows from UAV imagery", *ISPRS J. Photogrammetry Remote Sens.*, vol. 174, pp. 1–17, 2021..
- [5] W. Yijing, Y. Yi, W. Xue-Fen, Jian, C., Xinyun, L. Fig fruit recognition method based on YOLO v4 deep learning. ECTI-CON 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology: Smart Electrical System and Technology, Proceedings, 303–306. 2021.
- [6] K. L. Wiggers, C. D. Pohlod, R. Orlovski, R. Ferreira, T. A. Santos, "Detection and counting of plants via deep learning using images collected by RPA", *Revista Brasileira de Ciências Agrárias - Brazilian J. Agricultural Sci.*, vol. 17, n.º 2, pp. 1–9, 2022.
- [7] J. Wang, Z. Zhang, L. Luo, W. Zhu, J. Chen, W. Wang, "SwinGD: A Robust Grape Bunch Detection Model Based on Swin Transformer in Complex Vineyard Environment", *Horticulturae*, vol. 7, n.º 11, p. 492, 2021.
- [8] M. Sozzi, S. Cantalamessa, A. Cogato, A. Kayad, F. Marinello, "Automatic Bunch Detection in White Grape Varieties Using YOLOv3, YOLOv4, and YOLOv5 Deep Learning Algorithms", *Agronomy*, vol. 12, n.º 2, p. 319, 2022.
- [9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, "Ssd: single shot multibox detector". European Conference on Computer Vision, Springer, pp. 21-37. 2016. 10.1007/978-3-319-46448-0_2.
- [10] I. Sinapan, C. Lin-Kwong-Chon, C. Damour, J.J.A. Kadjo, M. Benne, "Oxygen bubble dynamics in pem water electrolyzers with a deep-learning-based approach", *Hydro*, 4, pp. 556-572. 2023. 10.3390/hydrogen4030036.
- [11] G. Jocher, A. Chaurasia, J. Qiu YOLO by Ultralytics. 2023.
- [12] H. Lou, X. Duan, J. Guo, H. Liu, J. Gu, L. Bi, H. Chen, "Dc-yolov8: small-size object detection algorithm based on camera sensor", *Electronics*, 12, p. 2323. 2023. 10.3390/electronics12102323.
- [13] C.Y. Wang, A. Bochkovskiy, H.Y.M. Liao, "Yolov7: Trainable Bag-Of-Freebies Sets New State-Of-The-Art for Real-Time Object Detectors". In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 7464-7475. 2023.
- [14] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, "Microsoft coco: common objects in context". In European Conference on Computer Vision, Springer, pp. 740-755. 2014. 10.1007/978-3-319-10602-1_48.
- [15] A. Koirala *et al.*, "Deep Learning for Real-time Malaria Parasite Detection and Counting Using YOLO-mp", *IEEE Access*, p. 1, 2022..
- [16] M.N.I. Muhlashin, A. Stefanie, "Klasifikasi Penyakit Mata Berdasarkan Citra Fundus Menggunakan YOLO V8", *JATI (J. Mhs. Tek. Inform.)*, vol. 7, n.º 2, pp. 1363–1368, 2023.