

Mederos Barrera, A.; Marcello Ruiz, F.; Eugenio González, F. Cartografía precisa de viñedos mediante U-Net y datos multimodales PNOA

# Cartografía precisa de viñedos mediante U-Net y datos multimodales PNOA

**Mederos Barrera, Antonio Ramón<sup>1</sup> Marcello Ruiz, Francisco Javier<sup>1</sup> Eugenio González, Francisco<sup>1</sup>**

<sup>1</sup> Universidad de Las Palmas de Gran Canaria (ULPGC), España

ORCID: Mederos Barrera 0000-0003-1680-0726 Marcello Ruiz 0000-0002-9646-1017 Eugenio González 0000-0002-0010-4024

Correspondencia: [mederosbarrera.antonio@gmail.com](mailto:mederosbarrera.antonio@gmail.com) [javier.marcello@ulpgc.es](mailto:javier.marcello@ulpgc.es) [francisco.eugenio@ulpgc.es](mailto:francisco.eugenio@ulpgc.es)

## RESUMEN

La identificación y mapeo de los viñedos en islas volcánicas presenta una elevada complejidad debido a la fuerte variabilidad topográfica, la fragmentación de las parcelas y la coexistencia de múltiples sistemas de cultivo tradicionales. En este contexto, la teledetección y las técnicas de Deep Learning constituyen herramientas prometedoras para mejorar la monitorización agrícola. El presente estudio evalúa el rendimiento del modelo U-Net para la segmentación de viñedos utilizando datos del Plan Nacional de Ortofotografía Aérea (PNOA) de 2021, combinando información espectral (RGB y NIR) y datos derivados de LiDAR (DEM y CHM). Se entrenaron cuatro configuraciones del modelo U-Net con distintas combinaciones de datos de entrada con el fin de analizar el impacto de cada fuente de información. Los resultados muestran que todas las configuraciones alcanzan métricas bastante robustas (F1-Score >71%, IoU >55%). Sin embargo, la integración de datos multimodales mejora sistemáticamente la segmentación. La mejor combinación (RGB + NIR + DEM + CHM) obtuvo un F1-Score del 75.32% y una IoU del 60.41%, evidenciando el impacto de incorporar información topográfica y de estructura de la vegetación. Este trabajo demuestra el potencial de la fusión de datos espectrales y LiDAR junto con modelos U-Net para la detección y monitorización precisa de viñedos en entornos agrícolas heterogéneos.

**Palabras clave:** Cartografía de viñedos, Deep Learning, U-Net, Datos multimodales, PNOA


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


# Cartografía precisa de viñedos mediante U-Net y datos multimodales PNOA

Mederos Barrera, Antonio Ramón <sup>(1)</sup>, Marcello Ruiz, Francisco Javier <sup>(1)</sup>, Eugenio González, Francisco <sup>(1)</sup>

<sup>(1)</sup> Universidad de Las Palmas de Gran Canaria (ULPGC), España.

 0000-0003-1680-0726, mederosbarrera.antonio@gmail.com

 0000-0002-9646-1017, javier.marcello@ulpgc.es

 0000-0002-0010-4024, francisco.eugenio@ulpgc.es

**Resumen:** La identificación y mapeo de los viñedos en islas volcánicas presenta una elevada complejidad debido a la fuerte variabilidad topográfica, la fragmentación de las parcelas y la coexistencia de múltiples sistemas de cultivo tradicionales. En este contexto, la teledetección y las técnicas de *Deep Learning* constituyen herramientas prometedoras para mejorar la monitorización agrícola. El presente estudio evalúa el rendimiento del modelo U-Net para la segmentación de viñedos utilizando datos del Plan Nacional de Ortofotografía Aérea (PNOA) de 2021, combinando información espectral (RGB y NIR) y datos derivados de LiDAR (DEM y CHM). Se entrenaron cuatro configuraciones del modelo U-Net con distintas combinaciones de datos de entrada con el fin de analizar el impacto de cada fuente de información. Los resultados muestran que todas las configuraciones alcanzan métricas bastante robustas (F1-Score >71%, IoU >55%). Sin embargo, la integración de datos multimodales mejora sistemáticamente la segmentación. La mejor combinación (RGB + NIR + DEM + CHM) obtuvo un F1-Score del 75.32% y una IoU del 60.41%, evidenciando el impacto de incorporar información topográfica y de estructura de la vegetación. Este trabajo demuestra el potencial de la fusión de datos espectrales y LiDAR junto con modelos U-Net para la detección y monitorización precisa de viñedos en entornos agrícolas heterogéneos.

**Palabras clave:** Cartografía de viñedos, Deep Learning, U-Net, Datos multimodales, PNOA.

## ***Precise vineyard mapping using U-Net and PNOA multimodal data***

**Abstract:** *The identification and mapping of vineyards on volcanic islands is highly complex due to strong topographic variability, parcel fragmentation, and the coexistence of multiple traditional cultivation systems. In this context, remote sensing and Deep Learning techniques represent promising tools to improve agricultural monitoring. This study evaluates the performance of the U-Net model for vineyard segmentation using data from Spain's 2021 National Plan for Aerial Orthophotography (PNOA), combining spectral information (RGB and NIR) and LiDAR-derived data (DEM and CHM). Four U-Net configurations were trained with different combinations of input data to analyze the contribution of each information source. The results show that all configurations achieve robust metrics (F1-Score > 71%, IoU > 55%). However, integrating multimodal data consistently improves segmentation. The best combination (RGB + NIR + DEM + CHM) achieved an F1-Score of 75.32% and an IoU of 60.41%, highlighting the impact of incorporating topographic and vegetation-structure information. This work demonstrates the potential of fusing spectral and LiDAR data together with U-Net models for accurate vineyard detection and monitoring in heterogeneous agricultural environments.*

**Keywords:** *Vineyard Mapping, Deep Learning, U-Net, Multimodal data, PNOA.*

## 1. INTRODUCCIÓN

La viticultura constituye una actividad agrícola de gran relevancia económica, cultural y paisajística en las Islas Canarias, especialmente en la isla de Tenerife, donde la elevada pendiente, la orografía compleja y la fragmentación parcelaria dan lugar a sistemas vitícolas heterogéneos y de difícil monitorización (García-

Verdugo, 2020). Tradicionalmente, la identificación y cartografía de cultivos se ha basado en trabajos de campo o fotointerpretación manual, procesos costosos en tiempo y recursos. En este contexto, la teledetección se presenta como una alternativa eficiente para la monitorización agrícola (Khosravi, 2025).

Las técnicas convencionales de *Machine Learning* basadas en píxel, que explotan principalmente la información espectral, presentan limitaciones en entornos complejos como el de Tenerife, debido a la variabilidad espacial, la mezcla de cubiertas y la influencia del relieve. Sin embargo, el avance de las técnicas de *Deep Learning* ha permitido mejorar significativamente los resultados en tareas de segmentación semántica mediante la integración conjunta de información espectral y espacial (Barros *et al.*, 2022). Estudios recientes han mostrado el potencial del *Deep Learning* para la segmentación agrícola y, específicamente, para la detección de viñedos mediante imágenes RGB y multispectrales. Asimismo, revisiones recientes subrayan la creciente relevancia de la segmentación profunda en agricultura (Barros *et al.*, 2022).

En este contexto, y dado que la información disponible por las Administraciones no siempre está actualizada ni presenta el detalle requerido, este trabajo aborda la identificación y segmentación de viñedos en Tenerife mediante el modelo U-Net y datos PNOA 2021. A diferencia de estudios previos centrados principalmente en imágenes UAV RGB o multispectrales para segmentación de viñedo, esta contribución analiza, sobre datos PNOA de cobertura regional y en un entorno volcánico altamente fragmentado, el efecto incremental de combinar información espectral (RGB y NIR) y variables derivadas de LiDAR (DEM y CHM) dentro de una misma arquitectura de segmentación. En particular, se evalúa qué combinación de datos de entrada ofrece el mejor compromiso para la cartografía precisa de viñedos en entornos agrícolas complejos.

El principal objetivo de este estudio es evaluar la capacidad del modelo U-Net para segmentar viñedos en un entorno agrícola complejo, analizando el potencial de la fusión de datos multimodales en la mejora de los resultados. Especialmente, se realiza un estudio de la mejor combinación de datos de entrada para la identificación de viñedos.

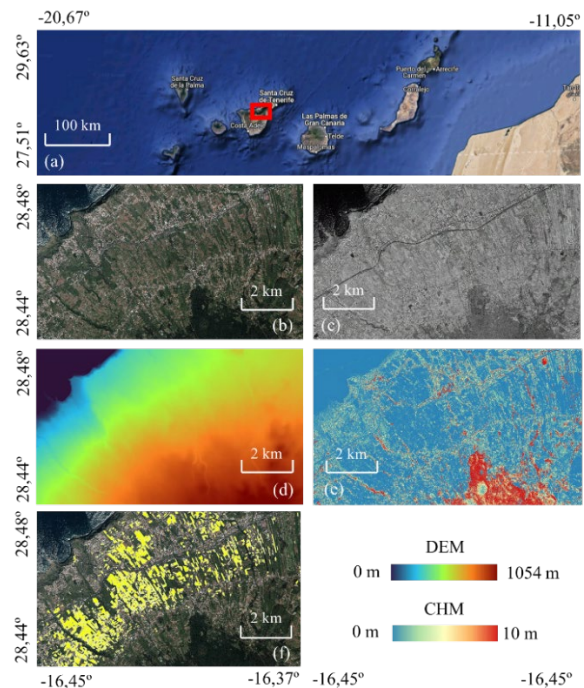
## 2. MATERIAL Y MÉTODOS

### 2.1. Zona de estudio

La zona de estudio corresponde a la isla de Tenerife (Islas Canarias, España), caracterizada por una orografía compleja, con fuertes pendientes, elevada variabilidad altitudinal y una marcada fragmentación del territorio, contemplando un rango de alturas de 0 a 3715 metros en un área de 2034 km<sup>2</sup>. En Tenerife, la viticultura se distribuye principalmente en zonas de medianías (de 300 a más de 1000 metros aproximadamente), y presenta una gran heterogeneidad espacial, con parcelas de reducido tamaño y sistemas de cultivo tradicionales (García-Verdugo, 2020), como banales o terrazas, lo que dificulta su identificación mediante métodos convencionales. En concreto, el área elegida se encuentra al norte de la isla de Tenerife (Figura 1), con una elevada presencia de diferentes parcelas de viñedos, y famosa por la Denominación de Origen (D.O.) Tacoronte-Acentejo.

### 2.2. Datos

Para el desarrollo del presente estudio se han empleado datos procedentes del Plan Nacional de Ortofotografía Aérea (PNOA) correspondientes al año 2021. En concreto, se han utilizado ortofotografías aéreas con información RGB y NIR a 20 cm de resolución, que proporcionan información espectral para la discriminación de cubiertas vegetadas. Adicionalmente, se han utilizado los datos de las nubes de puntos LiDAR PNOA (3ª cobertura), para la generación del Modelo Digital de Elevaciones (DEM) y el Modelo de Altura de Dosel (CHM), útiles como información topográfica y de altura de la vegetación.



**Figura 1.** Zona de estudio y datos. (a) Zona de estudio en la isla de Tenerife remarcada en rojo. (b) Imagen RGB PNOA 2021. (c) Imagen NIR PNOA 2021. (d) DEM generado con datos LiDAR PNOA. (e) CHM generado con datos LiDAR PNOA. (f) Parcelas de viñedos en amarillo del Gobierno de Canarias 2021.

### 2.3. Metodología

#### 2.3.1. Preprocesado LiDAR

Inicialmente, se implementó una etapa de preprocesado LiDAR. En este caso, los datos RGB y NIR se obtienen preprocesados. Respecto a los datos LiDAR se generaron los modelos DEM y CHM a 20 cm de resolución, a partir de la nube de datos PNOA. La nube LiDAR PNOA (3ª cobertura) se procesó en LiDAR360 y en el sistema de referencia común a todas las imágenes REGCAN95 / UTM zona 28N (EPSG:4083), con una densidad media aproximada de 5 puntos/m<sup>2</sup>. El empleo de *Kriging* para interpolar DEM y DSM se adoptó para obtener superficies continuas y suavizadas, coherentes con la resolución espacial de la ortofoto PNOA (20 cm), minimizando artefactos locales de interpolación.

Para la generación del DEM se utilizó el método *Kriging* con los parámetros: radio variable, distancia de 40 píxeles, utilizando 12 puntos y con un buffer de 15 píxeles. Posteriormente, se generó el Modelo Digital de Superficie (*Digital Surface Model*; DSM), con el modelo *Kriging* y los mismos parámetros anteriores. Finalmente, se obtuvo el CHM mediante la resta del DSM y el DEM.

### 2.3.2. Datos de entrenamiento y validación

Seguidamente, se obtuvieron los datos para el entrenamiento y validación. Para ello, se emplearon los datos de cultivos del Gobierno de Canarias de 2021. En concreto, se obtuvo una máscara binaria de la zona con la presencia de viñedos. Se unieron las zonas de monocultivo, así como de asociación de vid y otros suelos (barbecho, cítricos, hortaliza, huerta limpia, papa, templados y tunera).

Posteriormente, se unieron las imágenes (RGB, NIR, DEM, CHM) y máscara de referencia, y se realizó una teselación en parches de 1024x1024 píxeles (512x512 metros) sin solape. Estos parches son los utilizados para el entrenamiento y validación del modelo *Deep Learning* U-Net. En total, se generaron 836 teselas, de las cuales 444 tienen la presencia de vid, y 392 no presentan vid. Se puede observar que, a nivel de tesela, existe un balanceo. A su vez, se dividió cada grupo (con/sin vid) en 80/20% para el entrenamiento y validación, uniéndose cada agrupación. Esto resultó en un dataset de 668 parches de entrenamiento y 168 de validación. Además, cabe mencionar que se utilizó *Data Augmentation* para evitar sobreajustes. En concreto, se emplearon rotaciones de 90° y volteos horizontales y verticales aleatorios en el 50% de las imágenes, solamente en el conjunto de entrenamiento.

### 2.3.3. Modelo de segmentación U-Net

Para la segmentación de los viñedos, se empleó un modelo U-Net (Ronneberger *et al.*, 2015), con 4 etapas de reducción de escala. Cabe mencionar que la base de kernels es de 64, multiplicando por 2 en cada reducción de escala (*downscaling*). Respecto al bloque convolutivo, se empleó una doble etapa. Cada etapa consiste en una convolución, normalización por grupo y la aplicación de una ReLU. También se utilizó una regularización por abandono (*dropout*) del 20% al final del bloque convolutivo durante el entrenamiento.

Respecto al entrenamiento del modelo U-Net, se optó por un optimizador AdamW (con una tasa de aprendizaje de  $3 \cdot 10^{-3}$  y un decaimiento de peso de  $10^{-4}$ ), así como un programador de la tasa de aprendizaje *Reduce on Plateau* (utilizando el valor máximo, un factor de 0.5 y una paciencia de 7 épocas). Adicionalmente, es de especial interés la función de pérdida implementada, consistente en una combinación de la Entropía Cruzada Binaria (*Binary Cross Entropy*; BCE) y la función Tversky (Salehi *et al.*, 2017):

$$Loss = \lambda \cdot BCE + (1 - \lambda) \cdot Tversky, \quad (1)$$

donde  $\lambda$  es el parámetro que controla el impacto de cada función de pérdida. En este caso se eligió un valor de 0.3. En concreto, BCE penaliza a nivel de píxel, y ayuda en la estabilidad y convergencia del entrenamiento, mientras

que Tversky se emplea para mitigar el sesgo derivado del desbalanceo y controlar la penalización de FP y FN. Las imágenes con presencia de viñedos poseen un 20-40% de píxeles con la clase. Tversky ayuda a controlar los Falsos Positivos (FP) frente a los Falsos Negativos (FN), así como los Positivos Verdaderos (*True Positives*; TP):

$$Tversky = \frac{TP}{TP + \alpha FN + \beta FP}, \quad (2)$$

donde  $\alpha$  y  $\beta$  controlan los FN y FP. En este caso, se eligen valores de 0.55 y 0.44 respectivamente, por lo que se penalizan en mayor medida los FN para no perder viñedos, a la vez que se penalizan los viñedos detectados que no existen (FP).

Cabe mencionar que, para observar el impacto de diferentes fuentes de datos, se entrena el modelo U-Net con diferentes combinaciones de entradas: RGB (3 canales), RGB+NIR (4 canales), RGB+DEM+CHM (5 canales) y RGB+NIR+DEM+CHM (6 canales). Además, los experimentos se implementaron en PyTorch, con un tamaño de lote de 8 imágenes, durante 150 épocas, y se entrenaron en una GPU NVIDIA GeForce RTX 4080 SUPER.

### 2.3.4. Métricas de evaluación

Finalmente, respecto a las métricas de evaluación de los modelos, se optó por F1-Score e Intersección sobre Unión (*Intersection over Union*; IoU):

$$F1 = \frac{2TP}{2TP + FN + FP}, \quad (3)$$

$$IoU = \frac{TP}{TP + FN + FP}.$$

F1-Score es útil para analizar la capacidad de detección y en casos de desbalanceo de clases, e IoU es de interés en la evaluación de la precisión y coincidencia espacial de las parcelas segmentadas.

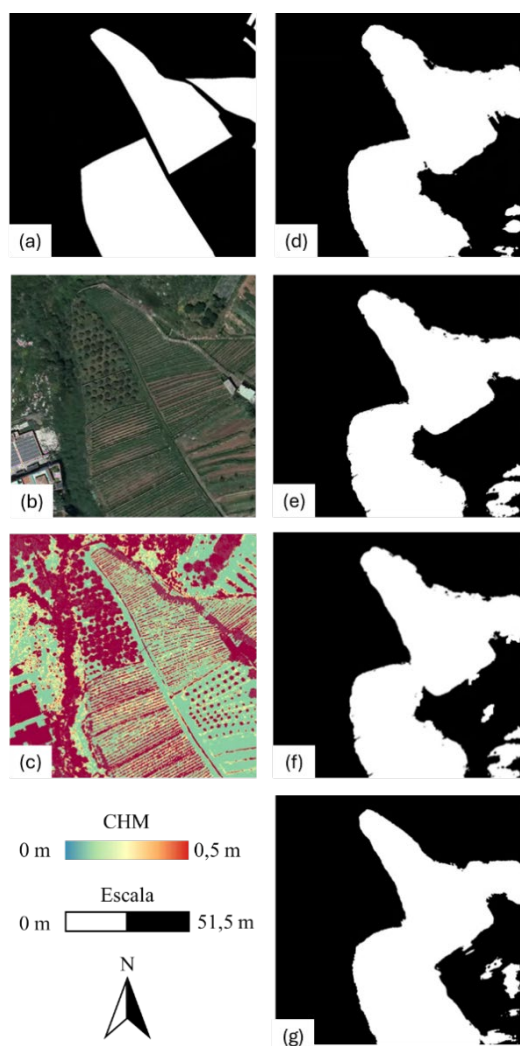
## 3. RESULTADOS

Respecto a los resultados, se presentan las métricas obtenidas con el modelo U-Net para cada combinación de entrada (Tabla 1). Se puede observar que las métricas obtenidas son bastante robustas utilizando cualquier combinación de datos de entrada (F1-Scores > 71% e IoU > 55%). Además, los resultados muestran una mejora progresiva del rendimiento a medida que se incorporan nuevas fuentes de información. Mientras que el uso exclusivo de datos RGB proporciona un rendimiento base razonable, la incorporación de datos NIR o LiDAR (DEM y CHM) mejoran las métricas, tanto en F1-Score como en IoU. Se destaca que las métricas son superiores añadiendo los datos DEM y CHM (datos LiDAR), que con el uso de NIR. La diferencia absoluta entre ambos es de 1.4% y 1.7% en F1-Score e IoU respectivamente. No obstante, la mejor combinación se obtiene al integrar todos los datos de entrada (RGB + NIR + DEM + CHM), alcanzando un F1-Score del 75.32% y una IoU del 60.41%, lo que evidencia la importancia de la fusión de datos multimodales para la segmentación de viñedos en entornos agrícolas complejos.

**Tabla 1.** Análisis del uso de diferentes combinaciones de entradas en la segmentación de viñedos con el modelo U-Net.

Datos de entrada	F1-Score (%) ↑	IoU (%) ↑
RGB	71.06	55.10
RGB + NIR	72.12	56.40
RGB + DEM + CHM	73.51	58.12
<b>RGB + NIR + DEM + CHM</b>	<b>75.32</b>	<b>60.41</b>

Por otro lado, en la Figura 2 se presenta un ejemplo de segmentación de un parche. Cualitativamente, también se observa que el uso de todos los datos multimodales mejora el resultado. No obstante, el uso exclusivo de la imagen RGB también ofrece un resultado adecuado.



**Figura 2.** Ejemplo de segmentación U-Net: (a) Máscara binaria de referencia (en blanco los viñedos). (b) Imagen RGB. (c) Modelo CHM. (d) Resultado con RGB. (e) Resultado con RGB y NIR. (f) Resultado con RGB, DEM y CHM. (g) Resultado con RGB, NIR, DEM y CHM.

#### 4. CONCLUSIONES

En este trabajo se ha evaluado el uso del modelo de *Deep Learning* U-Net para la segmentación de viñedos en la isla de Tenerife a partir de datos PNOA,

demostrando su robustez en un entorno agrícola caracterizado por una elevada complejidad. Los resultados obtenidos muestran que, aunque el uso exclusivo de imágenes RGB proporciona un rendimiento satisfactorio, la incorporación progresiva de información adicional mejora de forma consistente la segmentación. En particular, la integración de datos LiDAR (DEM y CHM) aporta información clave sobre el relieve y la estructura de la vegetación, superando incluso la mejora obtenida mediante el uso del canal NIR de forma aislada. La mejor configuración, basada en la fusión de datos espectrales y topográficos (RGB + NIR + DEM + CHM), alcanza un F1-Score del 75.32% y una IoU del 60.41%, evidenciando el valor de la fusión multimodal para la identificación precisa de viñedos.

#### 5. AGRADECIMIENTOS

Este trabajo ha sido financiado por el proyecto "Agricultura de precisión para el control del estrés hídrico y la prevención de enfermedades en cultivos de vid y olivo" (Convenio CC 25/63), de la Consejería de Agricultura, Ganadería, Pesca y Soberanía Alimentaria del Gobierno de Canarias. La participación en el congreso ha sido financiada por el Organismo Autónomo Parques Nacionales (Proyecto SPIP2022-02897). Adicionalmente, también agradecemos la disponibilidad de los datos LiDAR y ortofotos del PNOA, y el Mapa de Cultivos de Canarias del Gobierno de Canarias.

#### REFERENCIAS

- Barros, T., Conde, P., Gonçalves, G., Premebida, C., Monteiro, M., Ferreira, C. S. S., & Nunes, U. J. (2022). Multispectral vineyard segmentation: A deep learning comparison study. *Computers and Electronics in Agriculture*, *195*, 106782. <https://doi.org/10.1016/j.compag.2022.106782>
- García-Verdugo, D. L. (2020). Diagnóstico de la Viticultura en Canarias. Gobierno de Canarias: Santa Cruz de Tenerife, Spain. <https://www.gobiernodecanarias.org/agricultura/detalles/Diagnostico-de-la-viticultura-en-Canarias/>
- Khosravi, I. (2025). Advancements in crop mapping through remote sensing: A comprehensive review of concept, data sources, and procedures over four decades. *Remote Sensing Applications: Society and Environment*, *38*, 101527. <https://doi.org/10.1016/j.rsase.2025.101527>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer International publishing. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- Salehi, S. S. M., Erdogmus, D., & Gholipour, A. (2017). Tversky loss function for image segmentation using 3D fully convolutional deep networks. In *International workshop on machine learning in medical imaging* (pp. 379-387). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-67389-9\\_44](https://doi.org/10.1007/978-3-319-67389-9_44)