

Memoria fin de prácticas

Curso 2025/2026

Datos del alumno

Nombre y apellidos	Lucas Nicolás Descalzo Puchalt
Titulación	Máster Universitario en Ingeniería de Telecomunicación (MUIT)
Curso	2025/2026

Datos del convenio

Entidad	Cátedra ENIA-UPV en IA para el Desarrollo Sostenible
Proyecto marco	Agroalnext
Periodo	01/09/2025 – 04/02/2026 , con continuidad extracurricular 05/02/2026 – 30/06/2026
Tutora de empresa	María Belén Picó Sirvent
Tutor UPV	Valeriana Naranjo Ornedo

Valencia, 9 de septiembre de 2026

Índice

1. Introducción y objetivos	3
2. Línea 1 — Extracción de métricas por planta individual	3
2.1. Planteamiento	3
2.2. Desarrollo	3
2.3. Resultados y aprendizajes	4
3. Línea 2 — Segmentación semántica multispectral: dataset y benchmark	4
3.1. Conjuntos de datos	4
3.2. Preprocesado y protocolo experimental	5
3.3. Modelos evaluados	5
3.4. Preentrenamiento autosupervisado	5
4. Línea 3 — Transferencia: aplicación de escritorio	6
4.1. Motivación	6
4.2. Evolución	6
4.3. Arquitectura y funcionalidad	6
4.4. Empaquetado y distribución	7
5. Línea 4 — Fenotipado de melón	7
6. Infraestructura, herramientas y metodología de trabajo	7
7. Competencias adquiridas	8
8. Conclusiones y continuidad	8

1 Introducción y objetivos

El trabajo desarrollado se enmarca en la aplicación de inteligencia artificial y visión por computador a la agricultura de precisión, dentro del proyecto Agroalnext y bajo la Cátedra ENIA-UPV en IA para el Desarrollo Sostenible. El objetivo general ha sido estudiar la viabilidad de extraer, de forma automática y no destructiva, indicadores fenotípicos de cultivos de la familia de las Cucurbitáceas a partir de imágenes multiespectrales adquiridas con dron, en ensayos sometidos a estrés hídrico y salino.

Los objetivos concretos han sido:

1. Automatizar la extracción de métricas por planta individual (superficie foliar e índices de vegetación) a partir de ortofotos multiespectrales.
2. Desarrollar y comparar modelos de segmentación semántica sobre imagen multiespectral RGBNIR, cuantificando la aportación del canal infrarrojo cercano.
3. Construir un conjunto de datos propio, con anotaciones y *benchmarks* reproducibles, orientado a su publicación abierta.
4. Transferir el *pipeline* resultante a una herramienta de escritorio utilizable por personal investigador.

El trabajo se ha organizado en cuatro líneas que se solapan temporalmente y se alimentan entre sí.

2 Línea 1 — Extracción de métricas por planta individual

2.1 Planteamiento

El punto de partida fue la necesidad de comparar variedades de distintas cucurbitáceas a lo largo de varias jornadas de vuelo, midiendo superficie ocupada y respuesta espectral de cada planta por separado. Esto exige pasar de una máscara de vegetación global a una segmentación por individuo.

2.2 Desarrollo

- **Modelo de datos.** Se estructuró la información experimental (estudio, parcela, año, zona, lugar, jornada, identidad de planta y variedad) en una base de datos en formato Excel consultada mediante `sqlite3`, con las tablas ESTUDIO y PLANTA como núcleo relacional. La trazabilidad de cada planta a lo largo del tiempo se apoya en los identificadores de orden y de variedad.

- **Segmentación por individuo.** Se implementó y depuró un algoritmo de *Region Growing* multisemilla que crece regiones no sobre los valores crudos de píxel, sino sobre los mapas de características (*embeddings*) de una U-Net previamente entrenada para segmentación de vegetación. Las semillas se obtienen de las coordenadas de centro de planta registradas en campo. Se evaluaron tres métricas de similitud —coseno, euclídea y Mahalanobis— y se organizó la salida en una estructura de máscaras *.tif* por métrica y localización.
- **Comparativa de métodos.** El *Region Growing* sobre *embeddings* se comparó con dos alternativas clásicas: Watershed y una segmentación basada en umbral de índice. Se construyó un marco de evaluación con métricas de IoU, *scripts* de análisis estadístico y visualizaciones homogéneas que permitieran una comparación justa.
- **Extracción de indicadores.** Se desarrollaron dos herramientas complementarias: una para el análisis de una jornada aislada y otra para el análisis temporal multi-jornada, ambas produciendo estadísticos de NDVI y superficie por planta y variedad.

2.3 Resultados y aprendizajes

- Filtrar las variedades de interés *antes* del crecimiento de regiones, en lugar de descartarlas en post-proceso, reduce sustancialmente el coste computacional del *pipeline*.
- Cargar cada ortofoto y su máscara una única vez por jornada, en lugar de una vez por planta, resultó ser el patrón crítico de rendimiento en este flujo de trabajo.
- Se corrigieron errores estructurales en las consultas al modelo de datos que afectaban a la correspondencia planta-variedad.

3 Línea 2 — Segmentación semántica multispectral: dataset y benchmark

Esta es la línea de mayor componente investigador y la que sustentará el Trabajo Final de Máster.

3.1 Conjuntos de datos

- **CUCURDRONE.** Conjunto propio de imagen multispectral RGBNIR (4 canales) adquirida con dron, organizado por zonas geográficas y campañas anuales. Se ha preparado para su publicación abierta acompañado de *benchmarks* de segmentación.
- **WeedyRice.** Conjunto público de malas hierbas en arroz, también RGBNIR, empleado como contraste externo para validar la generalización de los modelos.

Se realizó una caracterización radiométrica del conjunto propio que reveló inconsistencias entre

campañas (ortomosaicos sistemáticamente más oscuros y con rango dinámico comprimido en determinadas campañas) y permitió identificar ortomosaicos defectuosos. Se comprobó que el canal NIR es el más robusto radiométricamente del conjunto, lo que refuerza el argumento sobre su papel en la detección de vegetación.

3.2 Preprocesado y protocolo experimental

- **Generación de etiquetas.** Etiquetado multiclase (suelo / cultivo / mala hierba) apoyado en una heurística basada en NDVI. Se estableció como requisito metodológico que el umbral de NDVI se calcule sobre la escala radiométrica original y *antes* de la normalización, condición necesaria para la reproducibilidad.
- **Normalización.** Por ortomosaico y no por parche (percentiles 1–99), de forma que se preserve la coherencia radiométrica interna de la escena y las diferencias relativas entre zonas de sombra y zonas iluminadas.
- **Particiones.** Estratificación por ortofoto y zona geográfica mediante una función de partición estandarizada con semilla fija, en lugar de un reparto aleatorio por parche. Esto evita la fuga de información entre entrenamiento y test derivada de la fuerte correlación entre parches vecinos de una misma ortofoto.
- **Formato de trabajo.** Parches TIFF de 512×512 píxeles, 4 canales, `float32` normalizado, con máscaras `uint8` y NDVI calculado al vuelo.

3.3 Modelos evaluados

Se implementó y entrenó un conjunto de arquitecturas bajo un protocolo idéntico de entrenamiento y evaluación (mismas funciones de cálculo de IoU, Dice, exactitud y sensibilidad; mismos formatos de visualización):

- Arquitecturas convolucionales: U-Net, U-Net++, DeepLabV3, DeepLabV3+ y FPN.
- Arquitecturas basadas en *transformers*: ViT-S con decodificador tipo DPT, DINOv3-SAT adaptado a entrada de 4 canales, y SAM3 con decodificador DPT.

3.4 Preentrenamiento autosupervisado

Dado el elevado volumen de imagen multispectral sin etiquetar disponible (aproximadamente 82.000 parches), se exploró el preentrenamiento autosupervisado como vía para mejorar los codificadores:

1. **DINO sobre ViT-S/16.** Primera aproximación, descartada tras detectarse colapso de prototipos (un único prototipo activo y rango efectivo muy bajo), un riesgo conocido en

conjuntos etiquetados pequeños.

2. **Masked Autoencoder (MAE) sobre Swin-T/8.** Implementación completa y propia del *pipeline*: codificador, enmascaramiento por ventanas, decodificador, bucle de entrenamiento y *script* de verificación. El MAE se seleccionó por aprender representaciones densas y espaciales, más adecuadas para segmentación, y por eliminar el riesgo de colapso.

Se estableció además el criterio de diagnóstico que distingue un rango efectivo bajo —normal en MAE— de una degeneración real del entrenamiento.

4 Línea 3 — Transferencia: aplicación de escritorio

4.1 Motivación

Los resultados anteriores solo tienen impacto real si el personal agronómico puede ejecutar el *pipeline* sin conocimientos de Python ni acceso a la infraestructura de cálculo. Esta línea consistió en convertir el código de investigación en una aplicación de escritorio distribuable.

4.2 Evolución

El desarrollo pasó por tres etapas: *pipeline* de evaluación por lotes, prototipo web en Streamlit y aplicación de escritorio en PySide6. El cambio a escritorio se motivó por limitaciones bloqueantes del prototipo web (recarga completa de página en cada interacción y necesidad de rodeos para la interacción con los visores gráficos).

4.3 Arquitectura y funcionalidad

- Separación estricta entre un módulo de lógica pura (entrada/salida, modelo, planificación de parches, inferencia, *region growing*, visualización, exportación y métricas) y un módulo que concentra toda la capa gráfica.
- Inferencia por parches de 512×512 obligatoria por restricciones de memoria de GPU, con un sistema de rejilla auto-optimizada que busca el desplazamiento óptimo entre múltiples combinaciones y reduce drásticamente el número de parches necesarios (por ejemplo, de 868 centros de planta a unos 156 parches).
- Carga de centros de planta mediante *shapefile* con reproyección automática de sistema de referencia, y normalización por ortofoto calculada en tiempo de ejecución.
- Concurrencia mediante trabajadores en hilos separados para mantener la interfaz responsiva durante la inferencia.

4.4 Empaquetado y distribución

Se empaquetó la aplicación como ejecutable de Windows con PyInstaller, incluyendo las dependencias CUDA, y se resolvieron los problemas de despliegue asociados: incompatibilidades entre distribuciones de OpenCV en el empaquetado, ganchos de tiempo de ejecución para las bibliotecas geoespaciales, orden de importación de bibliotecas nativas respecto a la inicialización de la interfaz gráfica, y persistencia de configuración entre sesiones. La versión se ha validado en máquinas de terceros distintas a la de desarrollo.

5 Línea 4 — Fenotipado de melón

En paralelo se inició una línea de fenotipado de melón basada en imagen de alta calidad:

- *Pipeline* de segmentación semántica basado en SAM3 (ViT-H/14) con decodificador DPT.
- Análisis colorimétrico en espacio Hunter Lab de los píxeles de pulpa como descriptor fenotípico.
- Ampliación del conjunto de datos con imágenes del banco de germoplasma del COMAV.
- Selección de genotipos mediante muestreo por punto más lejano, con el fin de maximizar la diversidad representada en el conjunto anotado.

6 Infraestructura, herramientas y metodología de trabajo

- **Cálculo:** sistema DGX con 6 GPU y 1 TB de RAM; entrenamientos ejecutados en contenedores Docker. Desarrollo local sobre GPU portátil NVIDIA RTX 2000 Ada.
- **Aprendizaje automático:** PyTorch, `segmentation-models-pytorch`, implementaciones propias de U-Net, Swin-T MAE y decodificadores DPT.
- **Geoespacial y de imagen:** rasterio, GDAL, geopandas, OpenCV, scikit-image y tiffle.
- **Datos y documentación:** pandas, sqlite3 y L^AT_EX.
- **Metodología:** configuración parametrizada al inicio de cada *script*, registro de eventos con marcas de tiempo e informe de consumo de memoria (RAM y VRAM) en los puntos clave del *pipeline*, validación incremental de cada módulo con datos sintéticos antes de su integración, y generación de documentos de traspaso al cerrar cada fase de desarrollo.

Se afrontaron además incidencias propias del trabajo sobre infraestructura compartida: agotamiento de hilos, conflictos de bibliotecas de álgebra lineal bajo carga de memoria elevada y gestión de memoria en entrenamientos de larga duración.

7 Competencias adquiridas

- **Técnicas:** aprendizaje profundo aplicado a segmentación semántica, aprendizaje autosupervisado, tratamiento digital de imagen multiespectral, teledetección y análisis geoespacial, ciencia de datos, desarrollo de aplicaciones de escritorio y empaquetado de software científico.
- **Metodológicas:** diseño de protocolos experimentales reproducibles, prevención de sesgos por fuga de datos, comparación justa entre arquitecturas, documentación técnica y redacción científica.
- **Transversales:** trabajo en equipo multidisciplinar con perfiles agronómicos, comunicación de resultados a audiencias no especializadas, y autonomía en la resolución de problemas mediante documentación y consulta.

8 Conclusiones y continuidad

Las prácticas han permitido recorrer el ciclo completo de un problema de investigación aplicada: desde la caracterización del dato bruto y la construcción de un conjunto anotado, pasando por el diseño y comparación de modelos, hasta la transferencia del resultado en forma de herramienta utilizable por el usuario final. La aportación más consolidada es el conjunto de datos CUCURDRONE con sus *benchmarks* de segmentación, en preparación para publicación abierta.

El trabajo continúa con carácter extracurricular hasta el final del curso y constituye la base del Trabajo Final de Máster, cuyas líneas abiertas son: completar el preentrenamiento autosupervisado a escala completa y su ajuste fino en segmentación, mejorar el etiquetado de la clase de mala hierba más allá de la heurística actual, y cerrar la publicación del artículo y del conjunto de datos.