



UNIVERSIDAD DE EXTREMADURA

CENTRO UNIVERSITARIO DE MÉRIDA

GRADO EN INGENIERÍA INFORMÁTICA EN
TECNOLOGÍAS DE LA INFORMACIÓN

TRABAJO FIN DE GRADO

OPTIMIZACIÓN DE PROBLEMA DE RIEGO
EFICIENTE EN HIGUERA MEDIANTE
INTELIGENCIA ARTIFICIAL

PABLO SETRAKIAN BEARZOTTI

Mérida, Junio de 2024



UNIVERSIDAD DE EXTREMADURA

CENTRO UNIVERSITARIO DE MÉRIDA

GRADO EN INGENIERÍA INFORMÁTICA EN
TECNOLOGÍAS DE LA INFORMACIÓN

TRABAJO FIN DE GRADO

OPTIMIZACIÓN DE PROBLEMA DE RIEGO
EFICIENTE EN HIGUERA MEDIANTE
INTELIGENCIA ARTIFICIAL

Autor: Pablo Setrakian Bearzotti
Fdo:

Director: Francisco Chávez de la O
Fdo:

Agradecimientos

En primer lugar, quiero agradecer a mis padres por animarme a estudiar ingeniería informática y por su apoyo incondicional durante toda mi formación académica.

A mi profesor Francisco Chávez de la O por sus clases, las cuales me hicieron empezar a interesarme por la inteligencia artificial y por su propuesta para participar en este proyecto.

Resumen

La seguridad alimentaria representa uno de los desafíos más críticos de la actualidad, especialmente en el contexto de la detección temprana de contaminantes que pueden representar riesgos significativos para la salud humana. Las micotoxinas, sustancias tóxicas producidas por hongos como el *Aspergillus flavus*, constituyen una amenaza considerable en la cadena alimentaria, siendo clasificadas por la Agencia Internacional para la Investigación del Cáncer como sustancias carcinógenas del grupo 1.

La detección tradicional de micotoxinas en productos agrícolas requiere métodos invasivos que implican la destrucción de las muestras, lo que resulta en pérdidas económicas significativas y limitaciones en el control de calidad durante el proceso productivo. En este contexto, el desarrollo de técnicas no invasivas para la detección temprana de contaminación por aflatoxinas se presenta como una necesidad imperante para la industria agroalimentaria.

Este Trabajo Fin de Grado se centra en el desarrollo de un sistema de inteligencia artificial capaz de detectar la contaminación por micotoxinas en higos frescos mediante el análisis de imágenes hiperespectrales y técnicas de inteligencia artificial. El proyecto aborda específicamente la clasificación de estados de enfermedad en frutos de higuera, utilizando un algoritmo genético junto a una red neuronal para identificar la presencia de *Aspergillus flavus* en diferentes concentraciones de contaminación.

La metodología propuesta se estructura en múltiples fases, comenzando con la localización y segmentación de higos individuales mediante detección RGB, seguida de la selección de bandas espectrales más informativas utilizando un algoritmo genético, y culminando con el procesamiento de parches espectrales completos mediante transformadas wavelet. Esta aproximación multimodal permite aprovechar tanto la información espectral específica como las características espaciales de las imágenes hiperespectrales.

Los resultados preliminares demuestran la viabilidad del enfoque propuesto para la detección no invasiva de contaminación por micotoxinas, estableciendo un marco de trabajo que contribuirá significativamente a mejorar los estándares de seguridad alimentaria en la producción de higos, especialmente relevante para regiones productoras como Extremadura, que representa el 55.5 % de la producción nacional española.

Palabras clave - inteligencia artificial, imágenes hiperespectrales, aprendizaje automático, visión por computador, algoritmo genético

Abstract

Artificial intelligence is no longer just a futuristic promise, but a reality that is redefining the way we interact with the world and conduct business. From healthcare to the manufacturing industry, AI has demonstrated its ability to radically transform diverse fields, unleashing a series of changes that are revolutionising today. Artificial intelligence will enable the automation of a wide range of complex processes that currently still require human intervention.

Artificial intelligence which, through methods, techniques and algorithms, provides computers the ability of identifying patterns in massive data and make predictions (predictive analytics). This learning allows computers to carry out specific tasks autonomously, i.e. without the need to be programmed.

The constant innovation and evolution of the emerging technologies, paired with the use by the vast majority of the world's population, has led to organisations exploiting the power of the automatic learning methods to help them extract better quality information, increase productivity, reduce the costs and get more value from their data.

In an ever-evolving world where technology is redefining the way we live, agriculture is no exception. Artificial intelligence is emerging as a key tool to help farmers adapt to the effects of climate change and lessen its impact. By capitalising on artificial intelligence, farmers can make informed, data-driven decisions and anticipate the climate shifts.

This research deals with sustainable production systems in the cultivation of fig trees and the study of water needs in order to design irrigation strategies using a set of automatic learning techniques belonging to the world of artificial intelligence.

Lastly, once the results generated by the entire system created have been obtained, these will be analysed and the solutions to the problems presented will be identified. These will allow us to observe the degree of efficiency of the experiments developed and to know if they meet the proposed targets.

Keywords - Artificial Intelligence, Agriculture, Climate change

Índice general

Agradecimientos	I
Índice de figuras	VII
Índice de tablas	IX
1. Introducción	1
1.1. Introducción	1
1.2. Motivación	2
1.3. Objetivo general	2
1.4. Objetivos específicos	3
1.5. Planificación	4
1.6. Organización del documento	4
2. Estado del Arte	6
2.1. Introducción	6
3. Desarrollo	7
3.1. Introducción	7
3.1.1. Adquisición de Imágenes Hiperespectrales	7
3.1.2. Arquitectura del Sistema	8
3.2. Primera Fase: Localización y Segmentación de Figuras	8
3.2.1. Objetivo de la Fase	8
3.2.2. Herramientas y Tecnologías Empleadas	9
3.2.3. Implementación del <i>Pipeline</i>	11
3.2.4. Resultados y Productos Intermedios	13
3.2.5. Desafíos y Observaciones Técnicas	13
3.3. Fases Futuras del Desarrollo	13
3.3.1. [Sección reservada para Fase 2: Selección de Bandas con Algoritmo Genético]	13
4. Resultados	14

5. Conclusiones y Trabajo Futuro	15
5.1. Conclusiones	15
5.1.1. Logros Principales	15
6. Agradecimientos	16
Referencias	17

Índice de figuras

3.1. Your image caption	11
-----------------------------------	----

Índice de tablas

Capítulo 1

Introducción

1.1. Introducción

Las enfermedades y plagas en cultivos representan un desafío económico significativo para los sectores agrícola y alimentario a nivel mundial. Entre las amenazas más graves se encuentran las micotoxinas, sustancias producidas naturalmente por ciertos tipos de hongos bajo condiciones particulares de humedad y temperatura. La presencia de micotoxinas en alimentos constituye un problema serio tanto para la salud humana como animal. La Agencia Internacional para la Investigación del Cáncer (IARC) ha clasificado un grupo de aflatoxinas como sustancias carcinogénicas del grupo 1, siendo la vía común de exposición a micotoxinas la ingesta de alimentos contaminados.

El hongo *Aspergillus flavus*, que prolifera a temperaturas entre 12°C y 27°C con 85 % de humedad, se multiplica en diversos alimentos incluyendo maíz, cacahuetes, arroz, frutos secos e higos. Aunque su presencia es típica de climas tropicales, también prolifera bajo ciertas condiciones de riego. El ciclo de crecimiento de la aflatoxina es de entre 3 y 5 días. Incluso si los higos van a ser secados, la introducción de higos infectados con aflatoxinas en el proceso puede provocar la contaminación de otros frutos. Por tanto, la detección de aflatoxinas en el producto fresco se considera crucial tanto para el consumo directo como para su procesamiento posterior.

El cultivo de la higuera (*Ficus carica* L.) tiene sus orígenes en la región de Caria en Asia, habiéndose extendido a otras áreas como la región mediterránea, África y América. España es actualmente el sexto mayor productor mundial, representando el 3.5 % de la producción global. La región de Extremadura, con 12,771 hectáreas cultivadas, representa el mayor productor

en España, alcanzando el 55.5 % de la producción nacional. El aumento en la productividad está vinculado a la adopción de técnicas innovadoras como fertilización, poda, tratamiento del suelo e irrigación. Sin embargo, los cambios en la humedad facilitan la propagación de la micotoxina *Aspergillus flavus*, requiriendo investigación adicional para analizar y prevenir que higos infectados entren en la cadena alimentaria humana.

1.2. Motivación

La detección tradicional de aflatoxinas se realiza mediante métodos invasivos que requieren la destrucción de la muestra, o mediante inspección visual en etapas avanzadas de contaminación. Estos métodos presentan limitaciones significativas: son lentos, costosos, y no permiten el análisis en tiempo real durante el proceso productivo. Además, los higos frescos son perecederos, tienen una vida útil limitada y son más sensibles al crecimiento microbiano que los higos secos, alterando la calidad del producto y representando un riesgo serio para la salud humana.

El uso de imágenes hiperespectrales (HSI) combinado con técnicas de inteligencia artificial, particularmente el deep learning, ofrece una alternativa prometedora. La tecnología HSI mide la interacción de un amplio espectro de luz con un objeto determinado, adquiriendo cientos de bandas espectrales contiguas para cada píxel en una imagen. Esta capacidad proporciona información detallada sobre el objeto y revela diferencias sutiles en textura y composición química que no son detectables mediante métodos convencionales.

La necesidad de desarrollar métodos no invasivos y precisos para la detección temprana de contaminación por aflatoxinas en higos frescos es crítica para garantizar la seguridad alimentaria, reducir pérdidas económicas en la cadena de producción, y proteger la salud pública.

1.3. Objetivo general

Desarrollar un sistema de inteligencia artificial basado en el análisis de imágenes hiperespectrales para la detección temprana de contaminación por micotoxinas en higos frescos, utilizando técnicas de deep learning y algoritmos genéticos para optimizar la selección de características espectrales relevantes.

1.4. Objetivos específicos

- Implementar un sistema de detección y segmentación automática de higos individuales en imágenes RGB mediante técnicas de visión por computador, generando máscaras y anotaciones para su posterior extracción de datos hiperespectrales.
- Desarrollar e implementar un algoritmo genético para la selección óptima de las tres bandas espectrales más informativas del cubo hiperespectral, reduciendo la dimensionalidad de los datos mientras se mantiene la capacidad discriminativa.
- Diseñar y entrenar modelos de redes neuronales profundas capaces de clasificar el estado de contaminación de los higos basándose en las bandas espectrales seleccionadas, evaluando diferentes arquitecturas y configuraciones.
- Implementar un sistema de procesamiento basado en transformada wavelet para el análisis del espectro completo a nivel de patches, preparando la infraestructura para futuras investigaciones comparativas con el método de referencia.

1.5. Planificación

El desarrollo del proyecto se estructura en las siguientes fases principales, diseñadas para abordar progresivamente los desafíos técnicos y científicos:

1. **Fase de preparación y adquisición de datos:** Recolección del dataset de imágenes hiperespectrales incluyendo muestras contaminadas con diferentes niveles de micotoxinas y muestras de control no contaminadas, capturadas durante un período de dos semanas para garantizar diversidad y robustez.
2. **Fase 0 - Detección y segmentación:** Desarrollo del sistema de detección automática de higos individuales mediante modelos de object detection y segmentación aplicados a versiones RGB de las imágenes hiperespectrales, generando máscaras y anotaciones en formato COCO.
3. **Fase 1 - Selección de bandas con algoritmo genético:** Implementación del algoritmo genético para identificar las tres bandas espectrales más informativas del cubo hiperespectral, construyendo imágenes reducidas para el entrenamiento de redes neuronales.
4. **Fase 2 - Procesamiento de espectro completo con patches:** Desarrollo del sistema de extracción de patches y aplicación de transformada wavelet sobre el espectro completo, estableciendo la base para análisis comparativos futuros.
5. **Fase de validación y documentación:** Evaluación exhaustiva de los modelos desarrollados, análisis comparativo de resultados, y preparación de la documentación técnica y científica del proyecto.

1.6. Organización del documento

El presente documento se estructura en los siguientes capítulos para presentar de manera sistemática el desarrollo y resultados del proyecto:

- **Capítulo 2. Marco teórico y estado del arte:** Presenta los fundamentos teóricos de las imágenes hiperespectrales, técnicas de deep learning aplicadas a la agricultura de precisión, y una revisión exhaustiva de trabajos relacionados con la detección de aflatoxinas mediante métodos no invasivos.

- **Capítulo 3. Desarrollo:** Detalla la metodología implementada en cada fase del proyecto, incluyendo la arquitectura del sistema de detección y segmentación, el diseño del algoritmo genético, y la implementación de los modelos de redes neuronales profundas.
- **Capítulo 4. Resultados:** Presenta los resultados experimentales obtenidos en cada fase, incluyendo métricas de rendimiento, análisis comparativo entre diferentes aproximaciones, y evaluación del impacto computacional y energético de los modelos.
- **Capítulo 5. Conclusiones y trabajo futuro:** Resume las contribuciones principales del proyecto, discute las limitaciones encontradas, y propone líneas de investigación futuras para mejorar y extender el sistema desarrollado.

Capítulo 2

Estado del Arte

2.1. Introducción

Capítulo 3

Desarrollo

3.1. Introducción

El desarrollo del proyecto se ha estructurado en múltiples fases secuenciales, cada una diseñada para abordar aspectos específicos del proceso de análisis hiperespectral aplicado a la detección de aflatoxinas en higos frescos. La metodología desarrollada implementa técnicas de *computer vision* de última generación combinadas con procesamiento especializado de datos hiperespectrales para crear un sistema automatizado de análisis de muestras.

3.1.1. Adquisición de Imágenes Hiperespectrales

Las imágenes fueron capturadas utilizando una cámara hiperespectral SPECIM, específicamente el modelo FX10 VNIR, cuyas características técnicas principales incluyen: resolución espacial de 1024 píxeles ($800 \text{ width} \times 1024 \text{ height}$), rango espectral de 400 nm a 1000 nm (visible y parte del infrarrojo cercano), 448 bandas espectrales, y un salto espectral de 1.339 nm.

El conjunto de datos comprende 320 higos cosechados de la plantación de la variedad calabacita ubicada en la “Finca La Orden-Valdesequera” ($38^{\circ}51' \text{ N}$, $6^{\circ}40' \text{ W}$, altitud 184 m) en Guadajira, España, donde CICYTEX tiene su sede central. Las imágenes hiperespectrales se capturaron durante un período de 2 semanas, utilizando cada semana 160 higos cosechados en diferentes etapas de madurez.

Cada semana, los 160 higos se dividieron en cuatro subconjuntos de 40 especímenes cada uno. El primer grupo correspondió a los controles sanos (clase 0), mientras que los tres grupos siguientes fueron inoculados con concentraciones de 10^3 UFC/mL (clase 1), 10^5 UFC/mL (clase 2), y 10^7 UFC/mL (clase 3), respectivamente. El proceso de inoculación se realizó mediante inmersión del área durante aproximadamente 3 segundos, siguiendo el protocolo

establecido por CICYTEX.

Las imágenes hiperespectrales se capturaron *post*-inoculación cada 24 horas durante cinco días consecutivos. Entre cada sesión de adquisición, las muestras se almacenaron en una cámara de incubación controlada a 25°C, con humedad relativa entre 80 y 90 % para promover el crecimiento fúngico. Cada clase consistió de 380 imágenes hiperespectrales, generando un total de 1520 imágenes hiperespectrales para el *dataset* completo.

Una vez adquiridas, las imágenes originales fueron preprocesadas con corrección blanco/negro para normalizar y corregir los datos, eliminando inconsistencias introducidas por factores ambientales, condiciones de iluminación y sensibilidad del sensor, facilitando así el análisis espectral posterior.

[referencia al proyecto nacional](#)

3.1.2. Arquitectura del Sistema

La arquitectura del sistema desarrollado combina detección de objetos mediante *text prompts*, segmentación semántica avanzada [refencia a esta técnica](#), y procesamiento especializado de cubos hiperespectrales. El enfoque metodológico permite la creación de un *dataset* estructurado que facilita el análisis automatizado y escalable de muestras, estableciendo las bases para las fases posteriores que implementarán algoritmos de *machine learning* para la clasificación y detección de contaminación.

El sistema se ha diseñado para manejar eficientemente el volumen considerable de datos hiperespectrales, garantizando tanto la precisión en la localización de objetos como la integridad radiométrica de los datos espectrales extraídos. La primera fase, completada y documentada en el repositorio `create-dataset`, establece los fundamentos para el procesamiento posterior mediante la localización y segmentación automatizada de higos individuales.

3.2. Primera Fase: Localización y Segmentación de Figuras

3.2.1. Objetivo de la Fase

La primera fase del desarrollo se centra en la localización y segmentación automatizada de higos individuales dentro de imágenes hiperespectrales que contienen múltiples especímenes dispuestos sobre una superficie de trabajo. El objetivo principal consiste en generar anotaciones precisas en formato COCO <https://cocodataset.org/> que incluyan cuadros delimitadores y máscaras de segmentación para cada higo detectado, junto con la extracción

de subcubos hiperespectrales radiométricamente corregidos correspondientes a cada instancia individual.

Esta fase es fundamental para el flujo de trabajo completo, ya que permite el aislamiento automatizado de regiones de interés que posteriormente serán procesadas por algoritmos de análisis espectral y clasificación. La precisión en esta etapa determina directamente la calidad de los datos de entrada para las fases subsiguientes.

3.2.2. Herramientas y Tecnologías Empleadas

La implementación de esta fase se basa en la integración de modelos de visión por computador de última generación, complementados con librerías especializadas para el procesamiento de datos hiperespectrales y manipulación de anotaciones.

Grounding DINO

Grounding DINO representa una arquitectura innovadora que combina capacidades de *grounding* de lenguaje natural con detección de objetos basada en *transformers*. Este modelo permite la localización de objetos mediante *text prompts*, eliminando la necesidad de entrenamiento específico para nuevas clases de objetos.

La arquitectura se basa en un *backbone Swin Transformer* que procesa las características visuales de la imagen, integrado con un módulo de *grounding* que establece correspondencias entre la entrada de texto y las regiones visuales relevantes. El modelo implementa *attention mechanisms* bidireccionales que permiten la fusión eficiente entre modalidades visual y textual.

Para este proyecto, se empleó el modelo GroundingDINO SwinB con pesos preentrenados `groundingdino.swinb_cogcoor.pth`, utilizando la entrada de texto “fig.” que significa higo en inglés para la detección de higos. Esta aproximación *zero-shot* resulta particularmente ventajosa para aplicaciones especializadas donde la disponibilidad de *datasets* anotados es limitada.

La optimización de los parámetros de inferencia se realizó mediante experimentación sistemática, evaluando diferentes combinaciones de `box_threshold` y `text_threshold`. El `box_threshold` controla la confianza mínima requerida para considerar una detección válida, mientras que el `text_threshold` determina el umbral de similaridad semántica entre el *text prompt* y las regiones detectadas. Tras un proceso de optimización por prueba y error, se determinó que los valores `box_threshold=0.25` y `text_threshold=0.25` proporcionan el balance óptimo entre sensibilidad de detección y precisión para el *dataset* específico de higos.

La Tabla muestra los resultados obtenidos para diferentes combinaciones de umbrales, donde se puede observar que los valores seleccionados minimizan tanto falsos positivos como falsos negativos, maximizando la precisión general del sistema de detección en el contexto específico de localización de higos individuales.

SAM2 (Segment Anything Model 2)

SAM2 constituye la evolución del paradigma “*Segment Anything*”, implementando capacidades de segmentación universal que permiten generar máscaras precisas para cualquier objeto en una imagen. El modelo se basa en una arquitectura *Vision Transformer* modificada que procesa *prompts* visuales (como *bounding boxes* o puntos) para generar segmentaciones de alta calidad.

La arquitectura de SAM2 incluye un *image encoder* basado en *Hierarchical Vision Transformer* (HiT) que genera *embeddings* densos de la imagen, un *prompt encoder* que procesa los *inputs* de *guidance*, y un *mask decoder* que combina ambos para producir las máscaras de segmentación finales. Esta separación modular permite flexibilidad en los tipos de *prompts* soportados.

Para la implementación, se utilizó la configuración `sam2.1_hiera_l.yaml` con el *checkpoint* `sam2.1_hiera_large.pt`, que proporciona el mejor balance entre precisión de segmentación y eficiencia computacional. La inferencia se ejecutó aprovechando *mixed precision* (`bfloat16`) para optimizar el uso de memoria GPU sin comprometer la calidad de las segmentaciones.

El modelo recibe como *input* los *bounding boxes* generados por *Grounding DINO*, convertidos al formato `xyxy` requerido, y produce máscaras binarias de alta resolución que delimitan precisamente los contornos de cada higo detectado. La configuración `multimask_output=False` se empleó para obtener una única máscara por detección, simplificando el procesamiento posterior.

Librerías de Soporte

El desarrollo incorporó librerías especializadas para diferentes aspectos del *pipeline*: `spectral` para manejo eficiente de datos hiperespectrales en formato ENVI, `pycocotools` para manipulación y validación de anotaciones en formato COCO, `supervision` para visualización avanzada de detecciones y segmentaciones, `OpenCV` para operaciones de procesamiento de imágenes, y `PyTorch` como *framework* base para la ejecución de los modelos de *deep learning*.

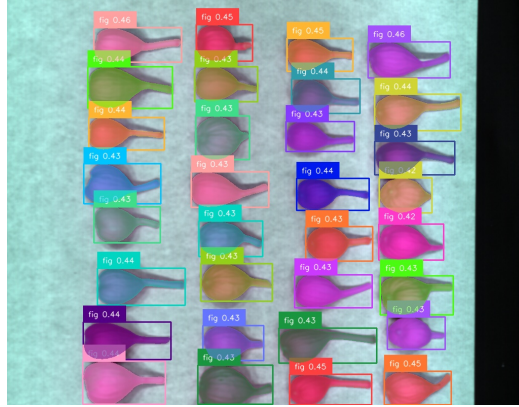


Figura 3.1: Your image caption

3.2.3. Implementación del *Pipeline*

Estructura del *Workflow*

El *pipeline* implementado consta de dos componentes principales ejecutados secuencialmente:

Componente 1 - Detección y Segmentación (`detect_and_segment_coco_annotations.py`)

La implementación comienza con la inicialización de ambos modelos en GPU, seguida del procesamiento *batch* de imágenes organizadas por clase. Para cada imagen, se ejecuta la siguiente secuencia:

1. **Carga y preprocesamiento:** Las imágenes RGB derivadas de los cubos hiperespectrales se cargan utilizando la función `load_image()` de *Grounding DINO*, que maneja automáticamente la normalización y conversión de formato requerida.
2. **Detección con *Grounding DINO*:** Se aplica el *text prompt* “fig.” para localizar instancias de higos en la imagen. El modelo retorna coordenadas de *bounding boxes* en formato normalizado (cx, cy, w, h), *scores* de confianza y *labels* correspondientes.
3. **Filtrado por tamaño:** Se implementó un filtro dimensional para eliminar detecciones erróneas, limitando las dimensiones máximas a 250×150 píxeles para asegurar la detección de higos individuales y evitar regiones que abarquen múltiples especímenes.
4. **Segmentación con SAM2:** Los *bounding boxes* filtrados se convierten al formato `xyxy` requerido por SAM2, que genera máscaras de segmentación de alta precisión utilizando *mixed precision* para optimizar el rendimiento computacional.

5. **Generación de anotaciones COCO:** Las detecciones se convierten al formato COCO estándar, incluyendo la conversión de máscaras a formato RLE (*Run-Length Encoding*) para almacenamiento eficiente.

Componente 2 - Extracción de Subcubos Hiperespectrales (`create_cropped_cubes.py`)

Este componente procesa las anotaciones generadas para extraer subcubos hiperespectrales:

1. **Carga de datos hiperespectrales:** Para cada imagen anotada, se localizan y cargan los archivos HDR correspondientes (imagen principal, referencia blanca y referencia oscura) desde la estructura de directorios organizada.
2. **Corrección radiométrica:** Se aplica la corrección línea por línea utilizando la fórmula estándar:

$$I_{corrected} = \frac{I_{raw} - R_{dark}}{R_{white} - R_{dark}} \quad (3.1)$$

donde las referencias blanca y oscura se promedian espacialmente para reducir el ruido.

3. **Extracción de subcubos:** Para cada anotación, se extraen subcubos hiperespectrales utilizando las coordenadas del *bounding box* y se aplican las máscaras de segmentación para aislar únicamente los píxeles correspondientes al higo.
4. **Almacenamiento estructurado:** Los subcubos se guardan en formato *NumPy* (`.npy`) con nomenclatura sistemática que preserva la trazabilidad hacia las imágenes originales.

Organización de Metadatos

El sistema implementa extracción automática de metadatos a partir de la nomenclatura de archivos mediante el módulo `file_processing.py`. Esta funcionalidad parsea información crítica incluyendo:

- Identificación de clase (C0: control sano, C1-C3: diferentes concentraciones de inoculación)
- *Timestamps* de captura con precisión temporal
- Información de condiciones experimentales (riego, días *post*-inoculación)

3.2.4. Resultados y Productos Intermedios

La ejecución completa de la primera fase genera los siguientes productos estructurados:

Anotaciones COCO: Archivos JSON por clase conteniendo metadatos completos de detección, con un total de anotaciones distribuidas *across* las cuatro clases experimentales. Cada anotación incluye *bounding box coordinates*, *segmentation masks* en formato RLE, y metadatos temporales extraídos automáticamente.

Visualizaciones de control: Imágenes anotadas que muestran las detecciones y segmentaciones superpuestas sobre las imágenes originales, facilitando la validación visual del *pipeline* y la identificación de casos *edge*.

Dataset de subcubos hiperespectrales: Colección estructurada de subcubos radiométricamente corregidos organizados por clase, con cada subcubo conteniendo el espectro completo (448 bandas espectrales) para la región correspondiente a un higo individual.

3.2.5. Desafíos y Observaciones Técnicas

Durante la implementación se identificaron y resolvieron varios desafíos técnicos significativos:

Optimización de memoria GPU: El procesamiento conjunto de *Grounding DINO* y SAM2 requirió implementación cuidadosa de *autocast contexts* para prevenir *overflow* de memoria, aplicando *mixed precision* selectivamente a SAM2 mientras se mantiene precisión completa para *Grounding DINO*.

Manejo de inconsistencias dimensionales: La variabilidad en dimensiones de imágenes hiperespectrales entre sesiones de captura requirió implementación de lógica robusta para *handling* de *shapes mismatch* durante la corrección radiométrica.

Precisión de segmentación: La calidad de las máscaras de segmentación mostró alta dependencia de la calidad de los *bounding boxes* de entrada, validando la importancia del filtrado dimensional implementado en la etapa de detección.

3.3. Fases Futuras del Desarrollo

3.3.1. [Sección reservada para Fase 2: Selección de Bandas con Algoritmo Genético]

Capítulo 4

Resultados

Capítulo 5

Conclusiones y Trabajo Futuro

5.1. Conclusiones

5.1.1. Logros Principales

Capítulo 6

Agradecimientos

Bibliografía