



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
<b>1. Introducción</b>	<b>1</b>
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
<b>2. Estado del Arte</b>	<b>6</b>
2.1. Introducción . . . . .	6
2.2. HSI en Agricultura de Precisión . . . . .	6
2.2.1. Imágenes RGB en Agricultura de Precisión . . . . .	8
2.2.2. Machine Learning en Imagenología Agrícola . . . . .	9
<b>3. Desarrollo</b>	<b>12</b>
3.1. Introducción . . . . .	12
3.1.1. Adquisición de Imágenes Hiperespectrales . . . . .	12
3.1.2. Arquitectura del Sistema . . . . .	13
3.2. Primera Fase: Localización y Segmentación de Figuras . . . . .	13
3.2.1. Objetivo de la Fase . . . . .	13
3.2.2. Herramientas y Tecnologías Empleadas . . . . .	14
3.2.3. Implementación del <i>Pipeline</i> . . . . .	16
3.2.4. Resultados y Productos Intermedios . . . . .	18
3.2.5. Desafíos y Observaciones Técnicas . . . . .	18
3.3. Fases Futuras del Desarrollo . . . . .	18



3.3.1. [Sección reservada para Fase 2: Selección de Bandas con Algoritmo Genético] . . . . .	18
<b>4. Resultados</b>	<b>19</b>
<b>5. Conclusiones y Trabajo Futuro</b>	<b>20</b>
5.1. Conclusiones . . . . .	20
5.1.1. Logros Principales . . . . .	20
<b>6. Agradecimientos</b>	<b>21</b>
<b>Referencias</b>	<b>22</b>

# Índice de figuras

3.1. Your image caption . . . . .	16
-----------------------------------	----

# 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 [1, 2]. Entre las amenazas más graves se encuentran las micotoxinas, sustancias producidas naturalmente por ciertos tipos de hongos bajo condiciones particulares de humedad y temperatura [2]. 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 [2].

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, cacahuètes, arroz, frutos secos e higos [2]. 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 [2]. 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 Ex-

tremadura, 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 [3, 2]. Estos métodos presentan limitaciones significativas: son lentos, costosos, y no permiten el análisis en tiempo real durante el proceso productivo [1]. 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 [4, 5, 2]. 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 [6, 7]. Esta capacidad proporciona información detallada sobre el objeto y revela diferencias sutiles en textura y composición química que no son detectables mediante métodos convencionales [3].

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 [2, 3].

## 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 [4, 8, 2].

## 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 [9].
- 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 [8].
- 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 [4, 2].

## 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 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 [4, 5, 3, 2].
- **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

La agricultura de precisión ha experimentado una transformación significativa en las últimas décadas, impulsada por el desarrollo de tecnologías emergentes que permiten el monitoreo y análisis automatizado de cultivos [1]. En este contexto, las tecnologías de imagenología avanzada, particularmente las imágenes hiperespectrales (HSI) y RGB, han emergido como herramientas clave para la detección temprana de contaminantes y patógenos en productos agrícolas. La integración de estas tecnologías con algoritmos de machine learning ha abierto nuevas posibilidades para el desarrollo de sistemas de detección no invasivos, precisos y eficientes [5, 6].

Este capítulo presenta una revisión sistemática del estado del arte en tecnologías de detección de contaminantes en productos agrícolas, centrándose en la aplicación de imágenes hiperespectrales y RGB en combinación con técnicas de aprendizaje automático. El análisis ofrece el marco teórico necesario para comprender las contribuciones de este proyecto y su relevancia dentro del panorama científico actual.

### 2.2. HSI en Agricultura de Precisión

La imagenología hiperespectral constituye un avance notable en el análisis remoto, ya que permite capturar información extremadamente detallada a través de cientos de bandas espectrales contiguas por cada píxel [7]. Este enfoque mide cómo un amplio rango del espectro electromagnético interactúa con un objeto, proporcionando información sobre su composición química y revelando variaciones sutiles que los métodos convencionales no pueden detectar [3]. La capacidad de HSI para capturar características a través de múltiples

bandas permite crear firmas únicas que representan cómo diferentes materiales responden a cada longitud de onda. Según la literatura científica, cuanto mayor es el número de bandas, más detalladas son las características que pueden ser identificadas, aunque no todas las bandas incluyen información relevante para mejorar la precisión de detección [8].

Los sistemas HSI típicamente operan en diferentes rangos espectrales, incluyendo el visible (VIS: 400-700 nm), infrarrojo cercano (NIR: 700-1000 nm), infrarrojo de onda corta (SWIR: 1000-2500 nm), y otros rangos especializados. Esta versatilidad espectral permite identificar características específicas de materiales y cambios químicos imperceptibles para el ojo humano o sistemas de imagen convencionales.

La aplicación de HSI en la detección de contaminantes agrícolas ha mostrado resultados prometedores en numerosos estudios. Estas investigaciones han abordado la identificación de infecciones fúngicas, micotoxinas y otros patógenos que comprometen la calidad y la seguridad de los productos alimentarios. En el contexto específico de detección de aflatoxinas, varios estudios han explorado el potencial de HSI para la identificación temprana de contaminación. La investigación ha demostrado que las aflatoxinas, particularmente la Aflatoxina B1 producida por *Aspergillus flavus*, pueden ser detectadas utilizando análisis espectral no invasivo. Estudios recientes han aplicado HSI con cámaras VNIR (400-1000 nm) y SWIR (1000-2500 nm) en diversos cultivos, logrando resultados prometedores en la identificación de muestras contaminadas frente a controles sanos.

La distribución superficial de las aflatoxinas representa una ventaja particular para el análisis mediante HSI, ya que permite detectar cambios químicos y estructurales en las capas exteriores de los productos agrícolas. Esta característica facilita la implementación de sistemas de detección que no requieren la destrucción de las muestras, manteniendo la integridad del producto para su comercialización.

A pesar de las ventajas evidentes de HSI, existen desafíos significativos asociados con su implementación práctica. El principal desafío radica en la alta dimensionalidad de los datos hiperespectrales, que puede representar un obstáculo considerable para los algoritmos de clasificación tradicionales [8]. Los cubos hiperespectrales generan volúmenes masivos de datos complejos que requieren técnicas especializadas de procesamiento y análisis. La gestión de la dimensionalidad espectral requiere estrategias de selección de características y reducción dimensional para identificar las bandas espec-

trales más informativas. La eliminación de bandas espectrales redundantes no solo facilita el análisis computacional, sino que también puede mejorar la precisión de clasificación al reducir el ruido y la información irrelevante. Adicionalmente, las condiciones de adquisición de imágenes hiperespectrales requieren un control cuidadoso de factores ambientales como la iluminación, temperatura y humedad, que pueden afectar la calidad y consistencia de los datos espectrales. La calibración y normalización de los datos espectrales son procedimientos críticos para garantizar la reproducibilidad y confiabilidad de los resultados.

### **2.2.1. Imágenes RGB en Agricultura de Precisión**

En el ámbito agrícola, las imágenes RGB siguen siendo la tecnología más extendida [9]. Aunque limitadas en comparación con HSI, destacan por su bajo costo, simplicidad de uso y rapidez en el procesamiento, cualidades que han facilitado su implementación a gran escala. Los sistemas RGB capturan únicamente tres bandas espectrales correspondientes a los colores primarios, generando representaciones visuales similares a la percepción humana. A pesar de su simplicidad, esta información resulta suficiente para detectar variaciones morfológicas y de color asociadas con infecciones fúngicas u otros contaminantes.

La literatura científica documenta numerosas aplicaciones exitosas de imágenes RGB en la detección de contaminantes agrícolas. Estos sistemas han demostrado eficacia particular en la identificación de cambios visuales asociados con infecciones fúngicas, decoloración y alteraciones morfológicas que preceden o acompañan la contaminación por micotoxinas. En el contexto de detección de aflatoxinas, algunos estudios han explorado el uso de imágenes RGB para identificar cambios visuales en productos contaminados. Aunque la información espectral limitada de RGB puede restringir la detección de cambios químicos sutiles, la tecnología ha mostrado utilidad en la identificación de síntomas visuales avanzados de contaminación fúngica.

Las aplicaciones RGB se han extendido también a sistemas de clasificación automatizada para el control de calidad en líneas de producción, donde la velocidad de procesamiento y la simplicidad del sistema son factores críticos. Estos sistemas pueden proporcionar una primera línea de defensa en la detección de productos visiblemente afectados.

Las principales limitaciones de los sistemas RGB radican en su capacidad limitada para detectar cambios químicos sutiles que no se manifiestan

visualmente. La contaminación por micotoxinas puede ocurrir sin síntomas visuales evidentes en las etapas tempranas, limitando la efectividad de los sistemas RGB para la detección precoz. Adicionalmente, los sistemas RGB son susceptibles a variaciones en las condiciones de iluminación y pueden requerir normalización cuidadosa para mantener la consistencia en diferentes entornos. La dependencia de características visuales también puede resultar en falsos positivos cuando se presentan variaciones naturales en color o textura que no están relacionadas con contaminación.

### **2.2.2. Machine Learning en Imagenología Agrícola**

La aplicación de deep learning (DL) en el procesamiento de imágenes hiperespectrales ha revolucionado las capacidades de análisis y clasificación en agricultura de precisión [4]. Como subconjunto del machine learning, el deep learning utiliza redes neuronales profundas que consisten en múltiples capas interconectadas de neuronas artificiales capaces de aprender representaciones de alto nivel a partir de datos de entrada.

La implementación de DL para el procesamiento y análisis de imágenes hiperespectrales fue inicialmente descrita en investigaciones pioneras que propusieron enfoques de clasificación utilizando información espacialmente dominante [10]. Desde entonces, un gran número de estudios han reflejado el interés creciente de la comunidad científica en esta área de investigación.

Las redes neuronales convolucionales tridimensionales han emergido como una arquitectura particularmente efectiva para el procesamiento de cubos hiperespectrales [11]. Estas redes pueden capturar simultáneamente características espaciales y espectrales, aprovechando la naturaleza tridimensional inherente de los datos hiperespectrales. Las CNN 3D operan mediante la aplicación de filtros convolucionales tridimensionales que se desplazan a través de las dimensiones espaciales ( $x$ ,  $y$ ) y espectral ( $z$ ) del cubo hiperespectral. Esta capacidad permite la extracción de características que consideran tanto la variabilidad espacial local como las relaciones espectrales entre bandas adyacentes.

Una alternativa prometedora al uso directo de redes 3D consiste en la aplicación de transformadas matemáticas al espectro hiperespectral antes del procesamiento con redes neuronales [2]. Las transformadas wavelet han mostrado particular eficacia en este contexto, permitiendo la descomposición del espectro en componentes de frecuencia que pueden ser procesados mediante arquitecturas de red más simples. El enfoque basado en transfor-

Las transformadas wavelet ofrecen ventajas computacionales significativas, ya que permiten la conversión de firmas espectrales unidimensionales en representaciones bidimensionales que pueden ser procesadas eficientemente mediante CNN 2D convencionales. Esta metodología puede mantener la información espectral crítica mientras reduce la complejidad computacional del procesamiento.

Las redes neuronales convolucionales bidimensionales (CNN 2D) representan el estándar establecido para el procesamiento de imágenes RGB en aplicaciones agrícolas [9]. Estas arquitecturas han demostrado eficacia excepcional en tareas de clasificación, detección de objetos y segmentación semántica aplicadas a productos agrícolas. Las CNN 2D operan mediante la aplicación de filtros convolucionales que capturan características espaciales locales en las imágenes RGB. La jerarquía de capas permite la extracción progresiva de características, desde detectores de bordes y texturas en capas tempranas hasta representaciones semánticas complejas en capas profundas.

El desarrollo de arquitecturas especializadas para detección de objetos y segmentación semántica ha facilitado la implementación de sistemas automatizados de análisis agrícola. Arquitecturas como YOLO (You Only Look Once), R-CNN y sus variantes han demostrado eficacia en la detección y localización automática de productos agrícolas en imágenes RGB. Para tareas de segmentación semántica, arquitecturas como U-Net, SegNet y DeepLab han mostrado resultados prometedores en la delimitación precisa de regiones de interés en imágenes agrícolas. Estas capacidades son fundamentales para el análisis posterior de características específicas de productos individuales.

La integración de información RGB e hiperespectral representa una frontera emergente en el análisis agrícola automatizado [5]. Los enfoques híbridos pueden aprovechar las ventajas complementarias de ambas modalidades: la simplicidad y velocidad del RGB para detección y localización, y la riqueza espectral de HSI para análisis químico detallado. Estos sistemas multimodales típicamente implementan arquitecturas de procesamiento en cascada, donde la información RGB se utiliza para la detección inicial y segmentación de productos, seguida por análisis hiperespectral detallado de las regiones de interés identificadas. Esta estrategia puede optimizar tanto la eficiencia computacional como la precisión de detección.

Las técnicas de fusión de características permiten la combinación sistemática de información extraída de diferentes modalidades de imagen. Estos enfoques pueden implementarse a diferentes niveles del pipeline de procesamiento: fusión temprana (combinación de datos raw), fusión intermedia

(combinación de características extraídas) o fusión tardía (combinación de decisiones de clasificadores independientes). La fusión efectiva de características RGB e hiperespectrales requiere consideración cuidadosa de las diferencias en resolución espacial, rango dinámico y características estadísticas entre las modalidades. Técnicas de normalización y alineamiento espacial son críticas para el éxito de estos enfoques.

El análisis de la literatura revela fortalezas y limitaciones distintas en los enfoques actuales para la detección de contaminantes agrícolas. Los sistemas basados en HSI ofrecen capacidades superiores de detección química pero requieren recursos computacionales significativos y equipos especializados costosos [6]. Los sistemas RGB proporcionan soluciones más accesibles y eficientes pero con capacidades limitadas de detección temprana.

Las arquitecturas de deep learning han demostrado capacidades excepcionales en ambas modalidades, pero su implementación efectiva requiere datasets grandes y representativos que pueden ser costosos y tiempo-intensivos de generar. La transferibilidad de modelos entre diferentes cultivos, condiciones ambientales y sistemas de adquisición permanece como un desafío significativo.

Las oportunidades de innovación identificadas incluyen el desarrollo de arquitecturas híbridas que combinen eficientemente información RGB e hiperespectral, la implementación de técnicas de selección inteligente de características espectrales y el desarrollo de sistemas adaptativos que puedan operar efectivamente en condiciones variables de campo [3]. La integración de técnicas de optimización evolutiva, como algoritmos genéticos, para la selección automática de características espectrales representa una dirección prometedora para mejorar tanto la eficiencia como la precisión de los sistemas de detección. Estos enfoques pueden automatizar el proceso de identificación de bandas espectrales óptimas para aplicaciones específicas.

# 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*.

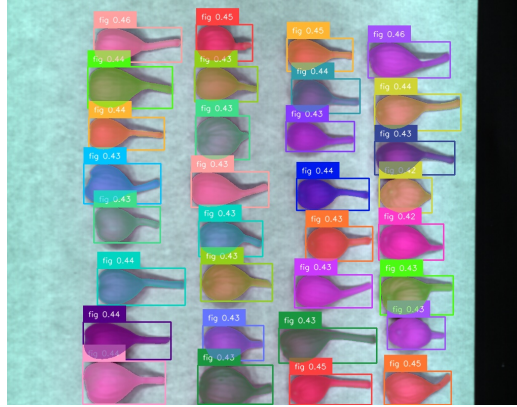


Figura 3.1: Your image caption

### 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  $250 \times 150$  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 hiperspectrales:** 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 hiperspectrales 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

- [1] I. Cisternas, I. Velásquez, A. Caro, and A. Rodríguez, “Systematic literature review of implementations of precision agriculture,” *Computers and Electronics in Agriculture*, vol. 176, p. 105626, 2020.
- [2] C. Cruz-Carrasco, J. Díaz-Álvarez, F. Chávez de la O, A. Sánchez-Venegas, and J. Villegas Cortez, “Detection of aspergillus flavus in figs by means of hyperspectral images and deep learning algorithms,” *AgriEngineering*, vol. 6, no. 4, pp. 3969–3988, 2024.
- [3] J. Wieme, K. Mollazade, I. Malounas, M. Zude-Sasse, M. Zhao, A. Gowen, D. Argyropoulos, S. Fountas, and J. Van Beek, “Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review,” *Biosystems Engineering*, vol. 222, pp. 156–176, 2022.
- [4] M. Paoletti, J. Haut, J. Plaza, and A. Plaza, “Deep learning classifiers for hyperspectral imaging: A review,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 279–317, 2019.
- [5] A. Signoroni, M. Savardi, A. Baronio, and S. Benini, “Deep learning meets hyperspectral image analysis: A multidisciplinary review,” *Journal of Imaging*, vol. 5, no. 5, 2019.
- [6] A. Khan, A. D. Vibhute, S. Mali, and C. Patil, “A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications,” *Ecological Informatics*, vol. 69, p. 101678, 2022.
- [7] R. Sahoo, S. Ray, and M. R., “Hyperspectral remote sensing of agriculture,” *Current science*, vol. 108, pp. 848–859, 03 2015.
- [8] D. Hong, N. Yokoya, J. Chanussot, J. Xu, and X. X. Zhu, “Learning to propagate labels on graphs: An iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction,” *ISPRS*

- Journal of Photogrammetry and Remote Sensing*, vol. 158, pp. 35–49, 2019.
- [9] K. P. Ferentinos, “Deep learning models for plant disease detection and diagnosis,” *Computers and Electronics in Agriculture*, vol. 145, pp. 311–318, 2018.
- [10] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, “Deep learning-based classification of hyperspectral data,” *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 7, pp. 2094–2107, 06 2014.
- [11] Z. Zhong, J. Li, Z. Luo, and M. Chapman, “Spectral-spatial residual network for hyperspectral image classification: A 3-d deep learning framework,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, pp. 847–858, 10 2017.