

Diseño e implementación de un sistema automatizado de clasificación de frutos cítricos en
banda transportadora mediante aprendizaje profundo e información espectral.

Johan Esneider Garzón Espejo y Meyer José Suarez Monroy

Trabajo de Grado para optar el título de ingenieros electrónicos

Director

Prof. Hans Yecid Garcia Arenas

Ingeniero Electrónico y Doctor en Ingeniería

Codirector

Pablo Andrés Gómez Toloza

Ingeniero Electrónico y PhD. (c) en ingeniería electrónica

Universidad Industrial de Santander

Facultad de ingeniería físico mecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones

Ingeniería Electrónica

Bucaramanga

2026

Dedicatoria

A mis padres, Martha y Antonio, por su amor incondicional, por el esfuerzo realizado para brindarme las oportunidades que hoy me permiten alcanzar este logro y por estar siempre presentes. Su apoyo, confianza y ejemplo fueron fundamentales durante mi formación.

A mi novia, Tatiana, por acompañarme durante toda mi carrera, por creer en mí incluso en los momentos más difíciles y por brindarme siempre su comprensión, apoyo y motivación para seguir adelante.

A mis hermanos, Dayana y Andrés, por su cariño, compañía y constante apoyo, que han hecho este camino mucho más llevadero.

A Kira, por su compañía incondicional, por hacer más amenas las largas jornadas de estudio y por permanecer siempre a mi lado, incluso durante las noches de trabajo.

Johan Esneider Garzón Espejo

A mis padres, Dora y José, porque desde niño alentaron mi curiosidad por entender cómo funcionan las cosas y me brindaron la oportunidad de convertirla en una profesión. Gracias por cada esfuerzo, consejo y muestra de confianza. Este logro es tan suyo como mío.

A mi familia, por ser el lugar al que siempre pude regresar para encontrar apoyo, tranquilidad y motivación en los momentos más exigentes de esta carrera.

A todas las experiencias, retos y aprendizajes que encontré durante estos años de formación, porque cada uno de ellos contribuyó a formar no solo al ingeniero que soy hoy, sino también a la persona que quiero seguir siendo.

Meyer José Suarez Monroy

Agradecimientos

Expresamos nuestro más sincero agradecimiento al profesor Hans Yecid García Arenas, director de este trabajo de grado, por su invaluable orientación, dedicación y compromiso durante todo el desarrollo de esta investigación. Su amplio conocimiento, experiencia y criterio académico fueron fundamentales para fortalecer cada etapa del proyecto, brindándonos las herramientas necesarias para afrontar los desafíos presentados durante el proceso. Agradecemos profundamente la confianza depositada en nuestro trabajo, su disposición para guiarnos, sus valiosas observaciones y el tiempo dedicado a compartir sus conocimientos, los cuales contribuyeron significativamente tanto a nuestro crecimiento profesional como personal.

De igual manera, expresamos nuestro más profundo agradecimiento a Pablo Andrés Gómez Toloza, codirector de este trabajo de grado, por su constante acompañamiento, disposición y apoyo durante el desarrollo de la investigación. Sus valiosos aportes, recomendaciones y observaciones enriquecieron este proyecto y permitieron fortalecer su calidad científica y académica. Agradecemos especialmente su disposición para resolver nuestras inquietudes, compartir su experiencia y motivarnos a enfrentar cada reto con compromiso y perseverancia. Su respaldo y orientación fueron esenciales para culminar satisfactoriamente este trabajo, dejando en nosotros importantes enseñanzas que trascenderán nuestra formación profesional.

Tabla de contenido

Introducción	12
1. Objetivos	14
1.1 Objetivo general	14
1.2 Objetivos específicos	14
2. Marco Teórico Y Antecedentes	15
2.1 Clasificación de calidad en frutos cítricos	16
2.2 Visión por computador en la agroindustria	17
2.3 Imágenes espectrales en la agroindustria	17
2.4 Procesamiento de imágenes	19
2.5 Aprendizaje profundo (DL)	19
2.6 Redes Neuronales Convolucionales (CNN)	20
2.7 Modelos de detección de objetos	21
2.8 Métricas de evaluación de modelos	22
2.9 Sistemas mecatrónicos y automatización	22
3. Desarrollo De La Solución	23
3.1 Adquisición de datos	24
3.2 Pre procesamiento de imágenes	27
3.3 Desarrollo de los modelos	32
3.3.1 Modelo de clasificación RGB	33
3.3.3 Adaptación del detector YOLO al dominio espectral	36

CLASIFICACIÓN AUTOMATIZADA DE FRUTOS CÍTRICOS MEDIANTE APRENDIZAJE PROFUNDO E INFORMACIÓN ESPECTRAL

5

3.4 Evaluación del desempeño de los modelos.....	38
3.4.1 Evaluación del modelo RGB	38
3.4.2 Evaluación del modelo de clasificación espectral	40
3.4.3. Evaluación del detector YOLO.....	42
3.5 Implementación del prototipo mecatrónico	44
3.6 Integración y sincronización del sistema de adquisición e inferencia	47
3.7 Pruebas experimentales.....	51
4. Conclusiones	53
5. Recomendaciones	55
Referencias Bibliográficas	56

Lista de figuras

Figura 1. Evolución del área sembrada y del área cosechada de mandarina en Colombia durante el periodo 2007–2024.	15
Figura 2. Principio de funcionamiento de un sistema de adquisición espectral de barrido lineal (<i>pushbroom</i>).	18
Figura 3. Esquema simplificado de una red neuronal artificial multicapa.	20
Figura 4. Secuencia metodológica desarrollada para la implementación y evaluación del sistema de clasificación.	24
Figura 5. Montaje experimental empleado para la adquisición de los cubos espectrales mediante la cámara SWIR.	26
Figura 6. Resultados de análisis espectral: perfil espectral	29
Figura 7. Ejemplo de detección y determinación de bounding boxes.	30
Figura 8. Ejemplos de regiones de interés extraídas de los cubos espectrales.	31
Figura 9. Arquitectura general de la red neuronal convolucional desarrollada para la clasificación de las muestras RGB.	34
Figura 10. Arquitectura híbrida CNN–Transformer desarrollada para la clasificación de las muestras espectrales.	36
Figura 11. Flujo general del proceso de adaptación del detector YOLO al dominio espectral mediante fine-tuning.	37
Figura 12. Curvas de entrenamiento y validación obtenidas durante el procedimiento de validación cruzada estratificada del modelo RGB.	39

Figura 13. Curvas de entrenamiento obtenidas durante el proceso de fine-tuning del modelo YOLOv5m.	43
Figura 14. Mecanismo de accionamiento seleccionado.....	45
Figura 15. Prototipo mecatrónico de separación implementado sobre la banda transportadora.	46
Figura 16. Arquitectura general del sistema integrado de adquisición, procesamiento, inferencia y accionamiento.	48
Figura 17. Interfaz del sistema durante la detección, seguimiento y clasificación automática de los frutos.....	49
Figura 18. Representación temporal de la planificación no bloqueante de eventos de accionamiento para frutos consecutivos.	50
Figura 19. Configuraciones evaluadas durante las pruebas experimentales del sistema integrado.	52

Lista de tablas

Tabla 1. Distribución del conjunto de imágenes RGB adquiridas.....	25
Tabla 2. Distribución de los cubos espectrales adquiridos.	27
Tabla 3. Características de los conjuntos de datos construidos.	32
Tabla 4. Desempeño del modelo RGB durante la etapa de evaluación.	40
Tabla 5. Métricas obtenidas por el modelo CNN–Transformer.	41
Tabla 6. Métricas de desempeño obtenidas por el detector YOLOv5m después del proceso de fine-tuning.....	42
Tabla 7. Escenarios considerados durante las pruebas experimentales.	52

Lista de apéndices

Apéndice A. Procedimientos complementarios de captura y conformación del dataset.

Apéndice B. Procedimientos complementarios de preprocesamiento

Apéndice C. Procedimientos complementarios para el diseño y entrenamiento del modelo CNN y YOLO.

Apéndice D. Evaluación experimental de los modelos de inteligencia artificial.

Apéndice E. Procedimientos complementarios de implementación del prototipo mecatrónico.

Apéndice F. Procedimientos complementarios de integración y sincronización del sistema.

Apéndice G. Procedimientos complementarios de pruebas experimentales.

Resumen

Título: Diseño e implementación de un sistema automatizado de clasificación de frutos cítricos en banda transportadora mediante aprendizaje profundo e información espectral.

Autores: Johan Esneider Garzon Espejo y Meyer José Suarez Monroy.

Palabras Clave: Aprendizaje profundo, imágenes espectrales, clasificación de frutos, redes neuronales convolucionales, YOLO, sistema mecatrónico, visión por computador, SWIR

Descripción:

Este trabajo presenta el diseño e implementación de un sistema automatizado para la clasificación de mandarinas sobre una banda transportadora mediante la integración de técnicas de visión por computador, aprendizaje profundo e imágenes espectrales en el rango SWIR (Infrarrojo de onda corta). Para el desarrollo del sistema se construyó un conjunto de datos propio a partir de capturas RGB y espectrales realizadas bajo condiciones controladas en el laboratorio HDSP. A partir de los cubos espectrales adquiridos se realizó un proceso de segmentación, selección de bandas espectrales y generación de representaciones pseudo-RGB, permitiendo adaptar la información espectral a modelos de inteligencia artificial previamente entrenados. La arquitectura propuesta está compuesta por dos etapas principales: detección y clasificación. La detección de los frutos se realizó mediante un modelo YOLO ajustado utilizando imágenes Falso RGB, mientras que la clasificación de calidad se efectuó mediante una red neuronal convolucional adaptada al dominio espectral mediante técnicas de transferencia de aprendizaje y una representación espectral tipo RGB+ 2. El desempeño de los modelos fue evaluado mediante métricas de clasificación y detección, incluyendo exactitud, recall, puntaje F1 y matriz de confusión. Adicionalmente, se desarrolló un sistema de adquisición e inferencia en tiempo real capaz de reconstruir imágenes a partir de información capturada por una cámara espectral tipo pushbroom, integrando procesos de reconstrucción, detección, clasificación y toma de decisiones. Finalmente, se implementó un prototipo mecatrónico compuesto por una banda transportadora, mecanismos de separación, servomotores y un sistema de control basado en Arduino, permitiendo la clasificación automática de los frutos. Los resultados obtenidos evidencian la viabilidad de integrar información espectral y técnicas de aprendizaje profundo en sistemas automatizados de clasificación para aplicaciones agroindustriales.

*Bachelor Thesis

** Facultad de ingeniería físico mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director Prof. Hans Yecid Garcia Arenas. Codirector Pablo Andrés Gómez Toloza

Abstract

Title: Design and Implementation of an Automated Citrus Fruit Classification System on a Conveyor Belt Using Deep Learning and Spectral Information.

Authors: Johan Esneider Garzon Espejo and Meyer José Suarez Monroy.

Keywords: Deep learning, spectral imaging, fruit classification, convolutional neural networks, YOLO, mechatronic system, computer vision, SWIR

Description:

This work presents the design and implementation of an automated mandarin classification system operating on a conveyor belt through the integration of computer vision, deep learning, and spectral imaging techniques in the Short-Wave Infrared (SWIR) range. A proprietary dataset was developed using RGB and spectral image acquisitions performed under controlled conditions at the HDSP laboratory. From the acquired spectral cubes, segmentation, spectral band selection, and False RGB generation processes were carried out, enabling the adaptation of spectral information to pre-trained artificial intelligence models. The proposed architecture consists of two main stages: detection and classification. Fruit detection was performed using a YOLO model fine-tuned with pseudo-RGB images, while quality classification was carried out using a convolutional neural network adapted to the spectral domain through transfer learning techniques and an RGB+2 spectral representation. Model performance was evaluated using classification and detection metrics, including accuracy, recall, F1-score, and confusion matrix analysis. Additionally, a real-time acquisition and inference system was developed to reconstruct images from data captured by a pushbroom spectral camera, integrating image reconstruction, detection, classification, and decision-making processes. Finally, a mechatronic prototype composed of a conveyor belt, sorting mechanisms, servomotors, and an Arduino-based control system was implemented to enable automatic fruit classification. The obtained results demonstrate the feasibility of integrating spectral information and deep learning techniques into automated classification systems for agro-industrial application.

*Bachelor Thesis

** Facultad de ingeniería físico mecánicas. Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Director Prof. Hans Yecid Garcia Arenas. Codirector Pablo Andrés Gómez Toloza

Introducción

El sector agroindustrial dedicado a la producción y comercialización de frutos cítricos enfrenta el desafío de garantizar altos estándares de calidad durante todo el proceso productivo. Para cumplir con los requisitos de comercialización y exportación, los frutos deben presentar características como uniformidad de color, tamaño adecuado, integridad y ausencia de defectos superficiales. Sin embargo, en numerosos procesos productivos la clasificación continúa realizándose de manera manual, dependiendo de la percepción y experiencia del operario, lo que introduce variabilidad en los criterios de selección y afecta la consistencia del producto final (Aguilar, 2017).

En este contexto, la visión por computador y el aprendizaje profundo se han consolidado como herramientas capaces de automatizar procesos de inspección y clasificación, proporcionando evaluaciones más rápidas, objetivas y reproducibles (Patricio & Rieder, 2018; Bhargava & Bansal, 2021). Además, organismos internacionales como el Codex Alimentarius y la UNECE establecen parámetros de calidad para frutas frescas, donde la apariencia externa y la uniformidad del color constituyen criterios relevantes para su aceptación comercial (CODEX Alimentarius, 2023; UNECE, 2023). Esto plantea la necesidad de desarrollar sistemas que permitan evaluar dichas características de forma automática y consistente.

Con el fin de abordar esta problemática, el presente trabajo propone el diseño e implementación de un sistema automatizado para la clasificación de mandarinas mediante la integración de técnicas de visión por computador, aprendizaje profundo e información espectral. La arquitectura desarrollada combina un modelo de detección basado en YOLO con una red neuronal convolucional para la clasificación de calidad, incorporando además información proveniente de imágenes espectrales en el rango SWIR para fortalecer el proceso de análisis.

La metodología adoptada comprende la construcción de un conjunto de datos propio, el preprocesamiento de imágenes RGB y espectrales, el entrenamiento y adaptación de modelos de inteligencia artificial, y su integración con un sistema mecatrónico de separación sobre una banda transportadora. Finalmente, este trabajo busca contribuir al desarrollo de soluciones tecnológicas aplicables al sector agroindustrial, reduciendo la dependencia de la inspección manual y promoviendo procesos de clasificación más objetivos, reproducibles y eficientes.

1. Objetivos

1.1 Objetivo general

Diseñar e implementar un sistema automatizado de clasificación de frutos cítricos en banda transportadora, basado en una red de aprendizaje profundo que utilice información espectral para estimar la uniformidad de color como indicador de la calidad comercial de los frutos.

1.2 Objetivos específicos

Diseñar e implementar el módulo mecatrónico de separación en banda transportadora, integrando electrónica de potencia y control de actuadores para el accionamiento de separadores en línea.

Implementar una red de aprendizaje profundo que estime el índice de uniformidad de color y clasifique los frutos en dos estados: apto y no apto.

Integrar el sistema completo para operación en línea, abarcando adquisición con la cámara existente, preprocesamiento, inferencia, lógica de decisión y accionamiento del módulo de separación.

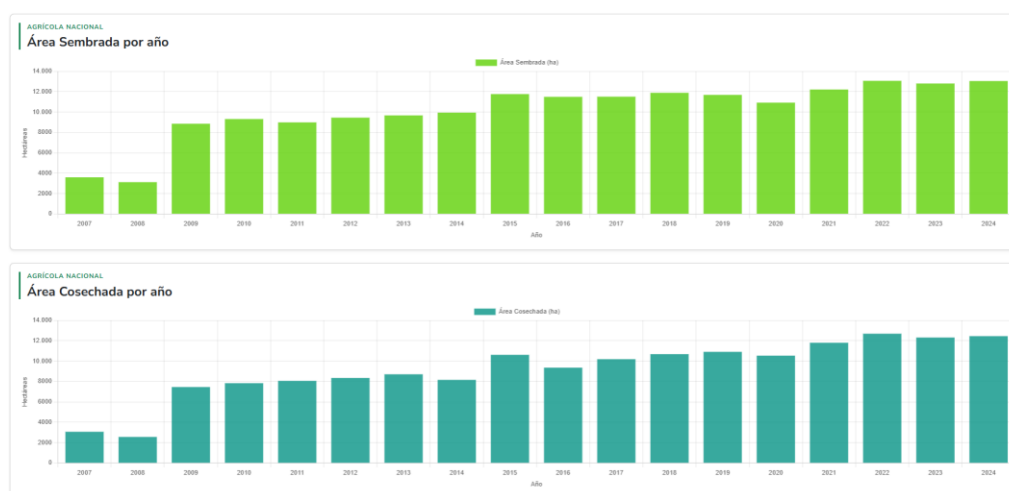
Validar el desempeño del sistema y del modelo de clasificación mediante métricas de visión por computador y pruebas experimentales en prototipo sobre la banda transportadora.

2. Marco Teórico Y Antecedentes

El sector cítrico constituye una actividad de gran importancia dentro de la agroindustria colombiana debido a su aporte a la producción agrícola y al abastecimiento de mercados nacionales e internacionales. En este contexto, la mandarina representa uno de los cultivos de mayor relevancia en el país, con un crecimiento sostenido en el área sembrada y cosechada durante los últimos años, como se muestra en la Figura 1. Entre las principales variedades cultivadas se encuentran Arrayana, Oneco y diferentes selecciones regionales, destacándose el departamento de Santander como uno de los principales productores nacionales (Agronet, 2024; Mateus-Cagua 2015).

Figura 1.

Evolución del área sembrada y del área cosechada de mandarina en Colombia durante el periodo 2007–2024.



Nota. Adaptado de de Agronet (2024).

El incremento en la producción de este fruto ha elevado la necesidad de optimizar las etapas de inspección y clasificación, tradicionalmente realizadas de forma manual y sujetas a

la experiencia del operario, lo que puede generar variabilidad en los resultados(Aguilar, 2017). En este contexto, las técnicas de visión por computador y aprendizaje profundo constituyen una alternativa para automatizar estos procesos, permitiendo realizar clasificaciones más objetivas, consistentes y reproducibles (Patricio, D & Rieder, 2018; Bhargava & Bansal, 2021).

2.1 Clasificación de calidad en frutos cítricos

La clasificación de calidad en frutos cítricos consiste en separar los productos de acuerdo con sus características visibles, considerando criterios como la uniformidad de la superficie, el color, el tamaño y la presencia de defectos externos, factores que influyen directamente en su aceptación y valor comercial (Kondo, 2010). Tradicionalmente, este proceso se realiza de forma manual, dependiendo de la experiencia del operario, lo que introduce variabilidad en la toma de decisiones y reduce la consistencia del proceso, especialmente en sistemas de producción continua donde se requiere evaluar grandes volúmenes de frutos (Blasco et al., 2003)

En este trabajo, la clasificación se plantea como un problema binario, en el que cada mandarina es catalogada como apta o no apta según la uniformidad de su superficie y la presencia de defectos visibles (Cubero et al., 2011). Debido a que algunos defectos no presentan un contraste suficiente en imágenes RGB, se emplean imágenes hiperespectrales en el rango SWIR, las cuales proporcionan información adicional relacionada con la composición y estructura del fruto, mejorando la capacidad de clasificación (Qui et al., 2013). En este contexto, las técnicas de visión por computador y aprendizaje profundo permiten automatizar este proceso, logrando clasificaciones más objetivas, consistentes y adecuadas para aplicaciones en tiempo real.

2.2 Visión por computador en la agroindustria

La visión por computador se ha consolidado como una tecnología clave en la agroindustria para la evaluación de la calidad de productos agrícolas, especialmente en etapas de pre y postcosecha. Tradicionalmente, la inspección de frutas y verduras se realiza de manera manual, lo que introduce subjetividad, variabilidad e inconsistencias en los resultados, además de ser un proceso laborioso y dependiente de las condiciones del entorno (Bhargava & Bansal, 2021). En este contexto, la calidad externa de los productos, determinada por atributos como el color, la textura, la forma y la presencia de defectos superficiales, influye directamente en su valor comercial y en su aceptación en el mercado, por lo que una clasificación inadecuada puede generar pérdidas económicas significativas (Zhang et al., 2014).

La automatización de este proceso mediante sistemas de visión por computador permite adquirir y procesar imágenes digitales para extraer características relevantes de los productos y apoyar su clasificación de forma objetiva, rápida y consistente. Aunque los sistemas basados en imágenes RGB permiten evaluar características visibles, presentan limitaciones para identificar defectos que no son evidentes en el espectro visible. Por esta razón, se han desarrollado sistemas multiespectrales y espectrales, capaces de incorporar información adicional del producto y mejorar la detección de defectos y la evaluación de calidad, ampliando así las posibilidades de automatización en aplicaciones agroindustriales (Bhargava & Bansal, 2021; Zhang et al., 2014).

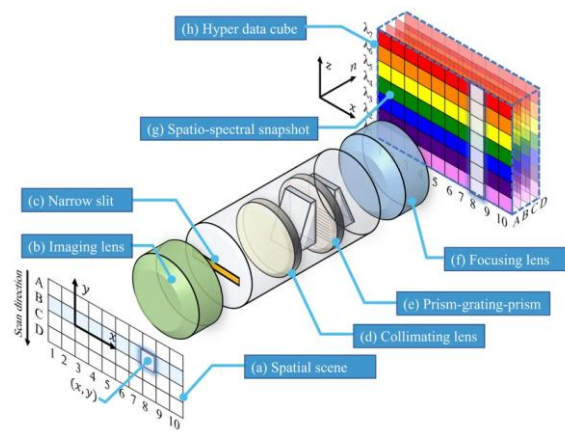
2.3 Imágenes espectrales en la agroindustria

Las imágenes espectrales se han consolidado como una herramienta para la evaluación no destructiva de productos agrícolas, al proporcionar información adicional sobre características como color, textura, estado de maduración y presencia de defectos, superando

las limitaciones de los sistemas de visión basados únicamente en imágenes RGB (Bhargava & Bansal, 2021; Zhang et al., 2014).

Figura 2.

Principio de funcionamiento de un sistema de adquisición espectral de barrido lineal (pushbroom).



Nota. Adaptado de Al-Hourani, et al(2023).

Uno de los métodos de adquisición más utilizados corresponde al sistema de barrido lineal (pushbroom), en el cual cada línea de la escena es descompuesta en múltiples longitudes de onda y, mediante el desplazamiento relativo entre el objeto y la cámara, se reconstruye un cubo espectral compuesto por dos dimensiones espaciales y una dimensión espectral (Al-Hourani, A., et al 2023). Esta representación permite asociar a cada píxel una firma espectral, facilitando la identificación de propiedades físicas y químicas del fruto que no son perceptibles en el espectro visible. Por esta razón, las imágenes espectrales se han integrado con técnicas de procesamiento de imágenes y aprendizaje automático para desarrollar sistemas automatizados de inspección y clasificación con mayor precisión en aplicaciones agroindustriales (Bhargava & Bansal, 2021; Zhang et al., 2014; Vignati et al., 2023).

2.4 Procesamiento de imágenes

El procesamiento de imágenes constituye una etapa fundamental dentro de los sistemas de visión por computador, ya que permite transformar la información visual adquirida para extraer características relevantes de los objetos de interés. En aplicaciones agroindustriales, las imágenes suelen verse afectadas por variaciones de iluminación, ruido y otras condiciones de adquisición, por lo que se emplean técnicas como filtrado, mejora de contraste y segmentación para resaltar la información útil y facilitar la evaluación de productos agrícolas (Bhargava & Bansal, 2018; Padilla Jimenez et al., 2017; Zhang et al., 2014).

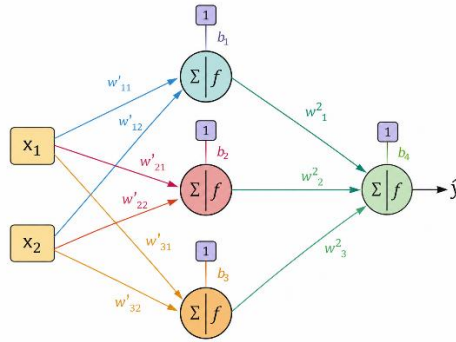
A partir de estas técnicas es posible identificar atributos como el color, la textura, la forma y la presencia de defectos superficiales, así como generar representaciones adecuadas para su procesamiento mediante modelos de aprendizaje automático y aprendizaje profundo (Bhargava & Bansal, 2018). De esta manera, el procesamiento de imágenes constituye el vínculo entre la adquisición de datos y los modelos inteligentes utilizados para la detección y clasificación automática.

2.5 Aprendizaje profundo (DL)

El aprendizaje profundo (Deep Learning, DL) es una rama del aprendizaje automático basada en redes neuronales artificiales con múltiples capas, capaces de aprender representaciones complejas mediante la extracción automática de características a partir de grandes volúmenes de datos (LeCun et al., 2015). A diferencia de los métodos tradicionales, donde las características deben definirse manualmente, los modelos de DL aprenden de manera jerárquica y automática, permitiendo identificar patrones directamente desde los datos (Goodfellow et al., 2016).

Figura 3.

Esquema simplificado de una red neuronal artificial multicapa.



Nota. La figura presenta una representación simplificada de una red neuronal artificial multicapa, en la que se muestran las entradas, los pesos, los sesgos, las funciones de activación y la salida del modelo. Elaboración propia.

Gracias a esta capacidad, el aprendizaje profundo ha alcanzado un alto desempeño en tareas de visión por computador como clasificación, detección y segmentación de objetos (Bhargava & Bansal, 2021). En aplicaciones agroindustriales, las redes neuronales convolucionales (Convolutional Neural Networks, CNN) se han consolidado como una de las arquitecturas más utilizadas para el análisis y clasificación automática de imágenes, mejorando la precisión y la consistencia de los sistemas de inspección frente a los métodos tradicionales (Zhang et al., 2014; Krizhevsky et al., 2012).

2.6 Redes Neuronales Convolucionales (CNN)

Las redes neuronales convolucionales son arquitecturas de aprendizaje profundo diseñadas para el análisis y procesamiento de imágenes, capaces de extraer automáticamente características relevantes a partir de los datos de entrada sin necesidad de definir manualmente los atributos utilizados durante la clasificación (Krizhevsky et al., 2012).

Su funcionamiento se basa principalmente en la combinación de capas convolucionales, encargadas de aprender patrones espaciales como bordes, texturas y formas, y capas de agrupamiento (pooling), que reducen la dimensionalidad de la información conservando las características más representativas (De la Cruz-Rocca et al., 2024). Gracias a estas propiedades, las CNN han demostrado un alto desempeño en tareas de clasificación y detección de objetos, siendo ampliamente utilizadas en aplicaciones agroindustriales para la inspección automática de frutas y otros productos agrícolas (Bhargava & Bansal, 2018; Abade et al., 2021).

2.7 Modelos de detección de objetos

La detección de objetos es una tarea de visión por computador que permite identificar y localizar automáticamente elementos dentro de una imagen mediante regiones delimitadas (bounding boxes), proporcionando simultáneamente la categoría y la ubicación de cada objeto detectado (Gunawan et al., 2022).

Entre los modelos de detección más utilizados se encuentra YOLO (You Only Look Once), una arquitectura de una sola etapa capaz de realizar la detección de objetos en tiempo real mediante una única inferencia, reduciendo significativamente el tiempo de procesamiento frente a otros enfoques (Vijayajumar et al., 2024). Gracias a estas características, YOLO ha sido ampliamente empleado en aplicaciones de automatización e inspección agrícola para la detección de frutos y otros productos (Qin et al., 2013). En el sistema desarrollado, YOLOv5 se utiliza para detectar automáticamente las mandarinas presentes en la escena y extraer las regiones de interés que posteriormente son enviadas al modelo de clasificación.

2.8 Métricas de evaluación de modelos

La evaluación del desempeño de modelos de aprendizaje profundo permite cuantificar su capacidad para realizar correctamente tareas de clasificación. En problemas de clasificación binaria, esta evaluación se basa en la matriz de confusión, la cual relaciona las predicciones del modelo con las clases reales del conjunto de datos, permitiendo identificar verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos (Sokolova & Lapalme, 2009).

A partir de la matriz de confusión se calculan métricas como la exactitud (Accuracy), la precisión (Precision), la sensibilidad (Recall), el puntaje F1 (F1-score), ROC (AUC) y $\text{map}@0.5:0.95$, ampliamente utilizadas para evaluar el desempeño de modelos de clasificación (Sokolova & Lapalme, 2009; Powers, 2020). En el presente trabajo, estas métricas fueron empleadas para analizar el comportamiento de los modelos de clasificación y detección desarrollados.

2.9 Sistemas mecatrónicos y automatización

Los sistemas mecatrónicos aplicados a la agroindustria integran estructuras mecánicas, sistemas de control, sensores y tecnologías de visión por computador para automatizar procesos de inspección y clasificación de productos agrícolas. Su implementación permite reducir la intervención manual, mejorar la precisión de clasificación y aumentar la eficiencia de las líneas de procesamiento.

Diversos trabajos han demostrado la viabilidad de estas tecnologías para la clasificación automática de frutas mediante bandas transportadoras y sistemas de visión artificial. En Colombia, (Romero Acero et al. (2015) desarrollaron un sistema de clasificación de mangos

basado en visión artificial, mientras que (Montoya Holguin, Cortés y Chaves, 2014) implementaron técnicas de reconocimiento automático de frutas tropicales utilizando características de color y forma. Asimismo, la incorporación de estructuras modulares, sistemas de control y tecnologías de manufactura aditiva ha favorecido el desarrollo de prototipos más flexibles y de bajo costo (Romero Acero et al., 2015).

3. Desarrollo De La Solución

El desarrollo de la solución se estructuró en siete etapas principales orientadas al diseño, implementación y validación de un sistema automatizado para la clasificación de frutos cítricos sobre banda transportadora. Estas etapas comprenden la adquisición de los datos, el preprocesamiento y construcción de los conjuntos de datos, desarrollo, evaluación de los modelos de aprendizaje profundo, la implementación del prototipo mecatrónico, la integración y sincronización de los diferentes componentes y, finalmente, la realización de pruebas experimentales sobre el sistema completo. En cada etapa se presentan los procedimientos implementados junto con los principales resultados obtenidos y su correspondiente análisis, mientras que los procedimientos, experimentos y resultados complementarios se presentan en los apéndices respectivos. De esta manera, se establece una relación directa entre las actividades desarrolladas, los resultados alcanzados y el cumplimiento progresivo de los objetivos específicos del proyecto. La Figura 4 presenta el flujo general seguido durante el desarrollo de la solución.

Figura 4.

Secuencia metodológica desarrollada para la implementación y evaluación del sistema de clasificación.



3.1 Adquisición de datos

Debido a la limitada disponibilidad de conjuntos de datos públicos orientados específicamente a la clasificación de mandarinas según criterios de calidad comercial, fue necesario realizar campañas de adquisición para conformar conjuntos de datos propios. Para ello, se trabajó con dos modalidades de información: imágenes convencionales en el espacio RGB e información espectral adquirida mediante una cámara SWIR. Las condiciones, procedimientos y configuraciones empleadas durante las campañas de captura se describen de manera complementaria en el **Apéndice A**.

Inicialmente, se realizaron salidas de campo para adquirir imágenes RGB de mandarinas bajo diferentes condiciones de madurez y apariencia superficial. Las muestras fueron clasificadas visualmente en las categorías “apta” y “no apta”, y posteriormente organizadas en directorios independientes de acuerdo con la clase correspondiente. Este procedimiento permitió obtener un conjunto compuesto por 4097 imágenes, de las cuales 2237 correspondieron a mandarinas aptas y 1860 a mandarinas no aptas.

Tabla 1.*Distribución del conjunto de imágenes RGB adquiridas*

Categoría	Número de imágenes	Porcentaje
Aptas	2237	54,6 %
No aptas	1860	45,4 %
Total	4097	100 %

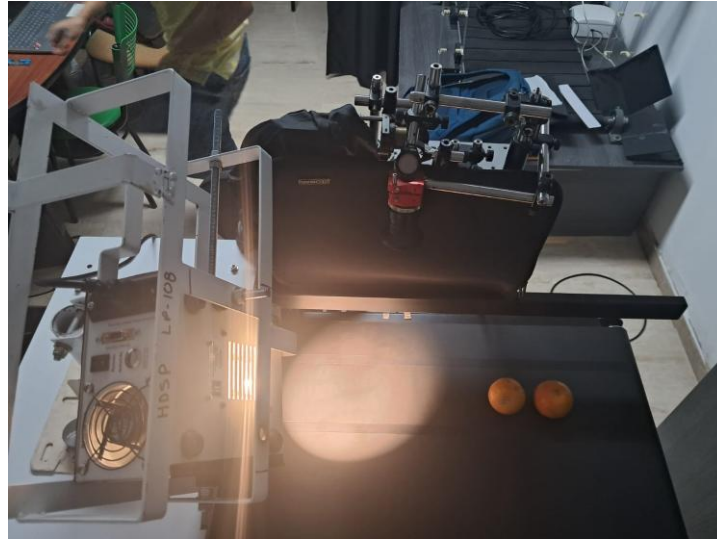
Nota. La categoría “Aptas” corresponde a imágenes de mandarinas clasificadas como aptas para el consumo, mientras que “No aptas” corresponde a imágenes de mandarinas que presentan características de deterioro o defectos. Los porcentajes se calcularon respecto al total de 4097 imágenes.

La distribución obtenida presenta una diferencia de 377 imágenes entre ambas categorías, equivalente a una proporción aproximada de 1,2 muestras aptas por cada muestra no apta. Aunque el conjunto no presenta una distribución exactamente equilibrada, ambas categorías cuentan con una representación considerable dentro de los datos adquiridos. La organización y almacenamiento del conjunto de imágenes RGB se presentan con mayor detalle en el **Apéndice A**.

Posteriormente, se realizó la adquisición de información espectral en condiciones controladas de laboratorio utilizando la cámara Daheng Imaging MER3-138-136G3M-P-SWIR con sensor InGaAs. A diferencia de las imágenes RGB, cada proceso de adquisición generó un cubo de datos compuesto por una secuencia de líneas espectrales capturadas durante el desplazamiento de los frutos sobre la banda transportadora. Los datos resultantes fueron almacenados en formato .npy, conservando la información necesaria para las posteriores etapas de análisis y procesamiento espectral.

Figura 5.

Montaje experimental empleado para la adquisición de los cubos espectrales mediante la cámara SWIR.



Como resultado de las campañas de adquisición espectral, se obtuvieron 156 cubos con aproximadamente 5 mandarinas cada uno, de los cuales 77 correspondieron a la categoría aptas y 79 a no aptas. La distribución obtenida presentó una diferencia de 2 muestras entre ambas categorías, lo que permitió disponer de un conjunto de datos con una representación similar de las dos clases para las posteriores etapas de procesamiento y construcción de los conjuntos de datos.

Tabla 2.

Distribución de los cubos espectrales adquiridos.

Categoría	Número de cubos	Porcentaje
Aptas	77	49,36 %
No aptas	79	50,64 %
Total	156	100 %

Nota. La categoría “Aptas” corresponde a los cubos espectrales asociados a mandarinas clasificadas como buenas, mientras que “No aptas” corresponde a los cubos asociados a mandarinas clasificadas como malas según los criterios establecidos para la construcción del conjunto de datos. Los porcentajes se calcularon respecto al total de 156 cubos espectrales.

La adquisición de los datos espectrales permitió disponer de información adicional a la proporcionada por las imágenes RGB convencionales y constituyó la base para la posterior selección de bandas, generación de representaciones espectrales y construcción de las entradas utilizadas por los modelos de detección y clasificación. Los procedimientos de captura, la estructura de almacenamiento de los cubos y las condiciones experimentales utilizadas se presentan detalladamente en el **Apéndice A**.

3.2 Pre procesamiento de imágenes

A partir de los datos obtenidos durante las campañas de adquisición, se implementaron procedimientos de preprocesamiento diferenciados para las imágenes RGB y los cubos espectrales, de acuerdo con las características de cada modalidad y los requerimientos de entrada de los modelos de aprendizaje profundo. Esta etapa tuvo como propósito transformar los datos adquiridos en conjuntos de muestras organizadas y estandarizadas que pudieran ser utilizadas durante las posteriores etapas de entrenamiento, validación y evaluación.

En el caso de las imágenes RGB, las muestras previamente organizadas según las categorías apta y no apta fueron redimensionadas a una resolución de 100×100 píxeles. Esta transformación permitió unificar las dimensiones espaciales de las imágenes y adaptarlas al tamaño de entrada establecido para la red de clasificación RGB. El procedimiento implementado y un ejemplo del resultado obtenido después del redimensionamiento se presentan detalladamente en el **Apéndice B**.

En el caso de los datos espectrales, el procesamiento partió de los 156 cubos adquiridos durante las campañas de captura. Debido a que cada cubo contenía información espacial y espectral de los frutos presentes en la escena, fue necesario implementar una secuencia de procesamiento orientada a identificar las regiones correspondientes a las mandarinas y extraer la información requerida para la construcción de las entradas de los modelos.

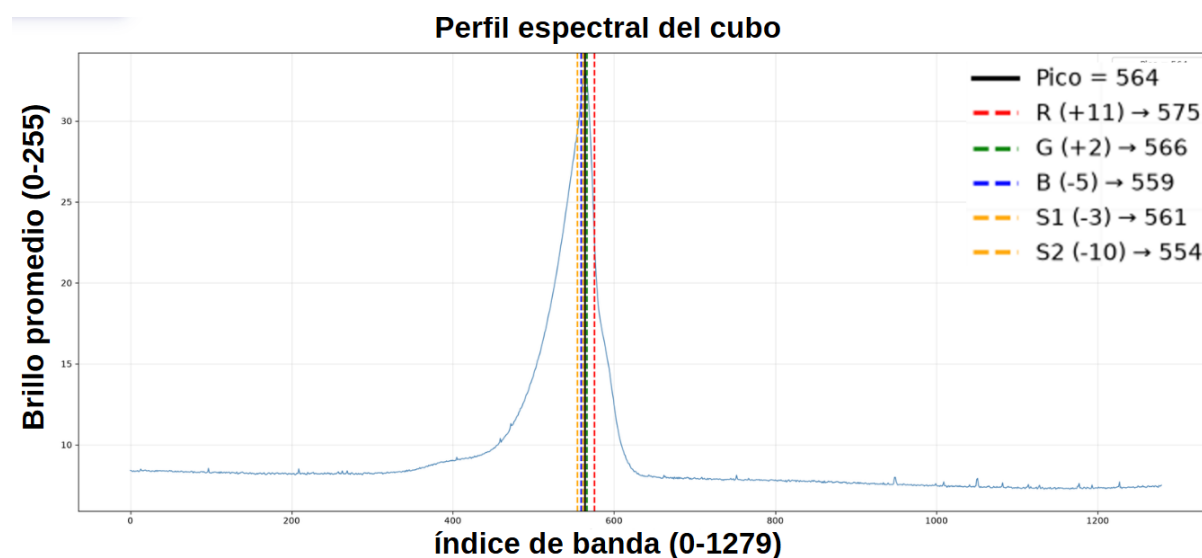
Inicialmente, se realizó el análisis de la información contenida en los cubos espectrales. Las bandas individuales fueron extraídas a partir del cubo C mediante:

$$I_b = C[:, \lambda, :] \text{ (Ec. 1)}$$

Donde λ corresponde al índice de la banda espectral seleccionada e I_b representa la imagen bidimensional resultante. A partir de las bandas extraídas se analizó la respuesta espectral de las escenas adquiridas con el propósito de identificar las regiones del espectro que proporcionaban información útil para diferenciar los frutos del fondo y establecer las bandas empleadas en las etapas posteriores del procesamiento. Los procedimientos y resultados complementarios del análisis espectral se presentan en el **Apéndice B**.

Figura 6.

Resultados de análisis espectral: perfil espectral

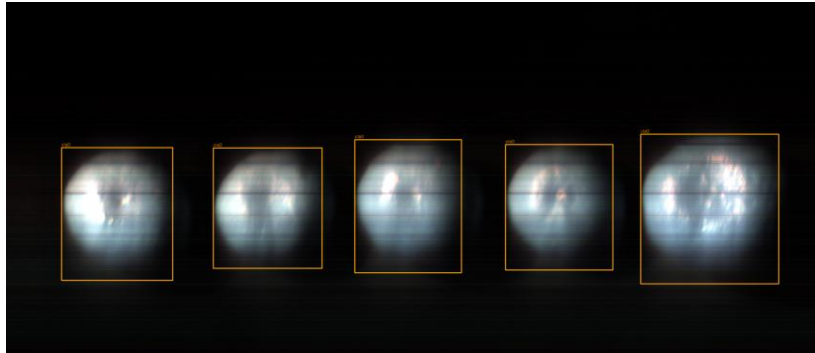


Nota. La figura muestra el perfil espectral promedio de un cubo hiperespectral, destacando el pico de respuesta y las bandas seleccionadas para la generación de la representación Pseu.

Con base en el análisis realizado, se implementó un procedimiento de segmentación automática para identificar las mandarinas presentes en cada cubo y extraer regiones de interés individuales. Para ello, se aplicaron procedimientos de umbralización y operaciones morfológicas que permitieron generar máscaras binarias, identificar los contornos correspondientes a los frutos y determinar sus bounding boxes. Como resultado, los cubos que originalmente contenían múltiples mandarinas fueron transformados en muestras individuales asociadas a cada fruto.

Figura 7.

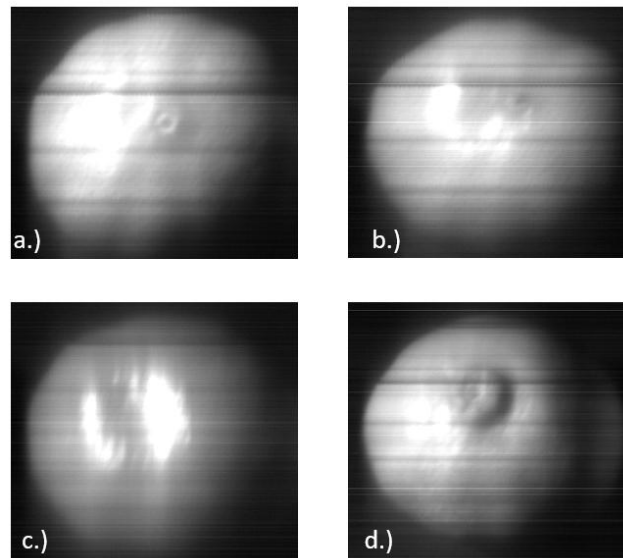
Ejemplo de detección y determinación de bounding boxes.



Como resultado del proceso de segmentación y extracción de las regiones de interés, las mandarinas presentes inicialmente en los cubos espectrales fueron aisladas de manera individual a partir de las bounding boxes obtenidas. Este procedimiento permitió generar muestras independientes para cada fruto, conservando la información espacial y espectral necesaria para las etapas posteriores de procesamiento y construcción de los conjuntos de datos. En la Figura 8 se presentan ejemplos de las regiones de interés extraídas, donde las muestras a) y b) corresponden a mandarinas clasificadas como aptas, mientras que las muestras c) y d) corresponden a mandarinas clasificadas como no aptas. Las diferencias observadas entre las muestras evidencian la variabilidad presente dentro de ambas categorías.

Figura 8.

Ejemplos de regiones de interés extraídas de los cubos espectrales: a) y b) mandarinas aptas; c) y d) mandarinas no aptas.



La aplicación de estos procedimientos permitió transformar los datos inicialmente adquiridos en conjuntos de muestras estructuradas de acuerdo con los requerimientos de los modelos implementados. De esta manera, se obtuvo un conjunto de imágenes pseudo-RGB destinado al entrenamiento y operación del detector YOLO y un conjunto de tensores espectrales de cinco bandas destinado al modelo de clasificación. En la tabla 3. se resume la estructura final de los conjuntos de datos y los formatos utilizados para su almacenamiento.

Tabla 3.

Características de los conjuntos de datos contruidos.

Modelo	Representación de entrada	Número de canales	Formato de almacenamiento
YOLO	Pseudo-RGB	3	uint8
Modelo de clasificación	Tensor espectral	5	float32

Nota. La representación Pseudo-RGB corresponde a la combinación de tres bandas espectrales seleccionadas para generar una imagen de tres canales. El tensor espectral está conformado por cinco bandas espectrales utilizadas como entrada del modelo de clasificación.

Los procedimientos detallados de redimensionamiento de las imágenes RGB, análisis espectral, selección de bandas, segmentación automática, extracción de regiones de interés, construcción de las representaciones pseudo-RGB y organización de los conjuntos de datos se presentan de manera complementaria en el **Apéndice B**.

3.3 Desarrollo de los modelos

Con el propósito de implementar el sistema automatizado de clasificación de frutos cítricos, se desarrollaron tres modelos de inteligencia artificial encargados de realizar las etapas de detección y clasificación durante el procesamiento de las imágenes adquiridas por la cámara SWIR. El desarrollo inició con una red neuronal convolucional entrenada utilizando imágenes RGB, la cual permitió establecer una línea base para el problema de clasificación binaria entre frutos aptos y no aptos. Posteriormente, se implementó un modelo de clasificación basado en información espectral, diseñado para aprovechar las cinco bandas seleccionadas durante la etapa de preprocesamiento. Finalmente, el detector YOLO fue adaptado al dominio espectral

mediante técnicas de transferencia de aprendizaje, permitiendo su integración dentro del sistema de clasificación en tiempo real.

Los procedimientos detallados relacionados con el diseño de las arquitecturas, configuración de los entrenamientos, validación cruzada, almacenamiento de los modelos y estrategias de optimización implementadas se presentan de manera complementaria en el **Apéndice C**.

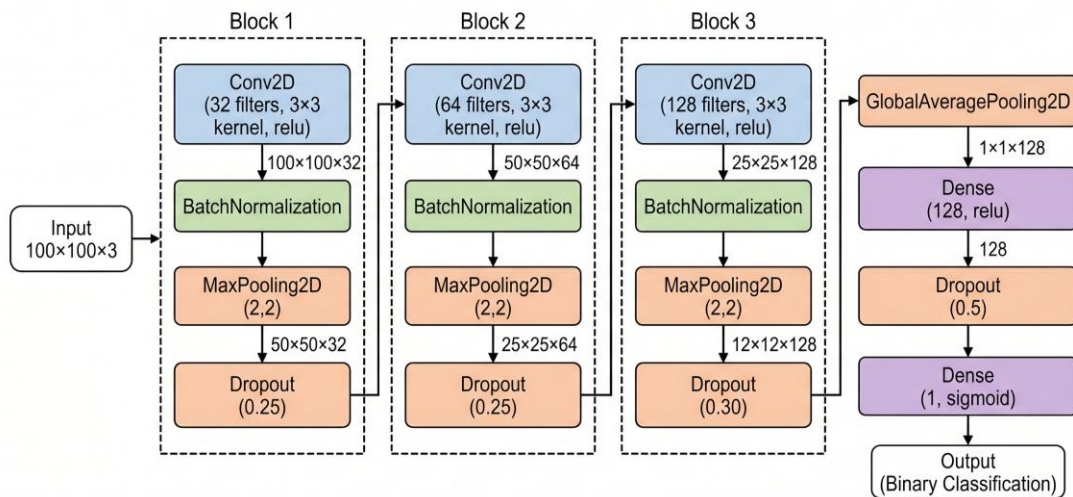
3.3.1 Modelo de clasificación RGB

Como etapa inicial del desarrollo, se diseñó e implementó una red neuronal convolucional entrenada con el conjunto de imágenes RGB construido durante la adquisición de datos. Este modelo tuvo como propósito establecer una línea base de desempeño para el problema de clasificación binaria, además de verificar la viabilidad del enfoque antes de incorporar información espectral al sistema.

La arquitectura desarrollada estuvo compuesta por bloques convolucionales encargados de extraer características espaciales de las imágenes, seguidos de operaciones de agrupamiento, normalización y regularización. Finalmente, las características obtenidas fueron procesadas mediante capas densamente conectadas para estimar la probabilidad de que cada muestra perteneciera a la categoría apta o no apta.

Figura 9.

Arquitectura general de la red neuronal convolucional desarrollada para la clasificación de las muestras RGB.



Nota. La figura muestra la arquitectura de la red neuronal convolucional utilizada para la clasificación binaria de las muestras RGB. El modelo está compuesto por tres bloques convolucionales con capas de normalización, agrupamiento y dropout, seguidos de capas de clasificación mediante promedio global y capas densas.

El entrenamiento del modelo se realizó mediante validación cruzada estratificada de cinco particiones, permitiendo evaluar la estabilidad del proceso de aprendizaje sobre diferentes distribuciones de las muestras disponibles. Posteriormente, el modelo definitivo fue entrenado utilizando la totalidad del conjunto de datos y almacenado para ser utilizado como referencia durante la comparación con los modelos desarrollados a partir de información espectral.

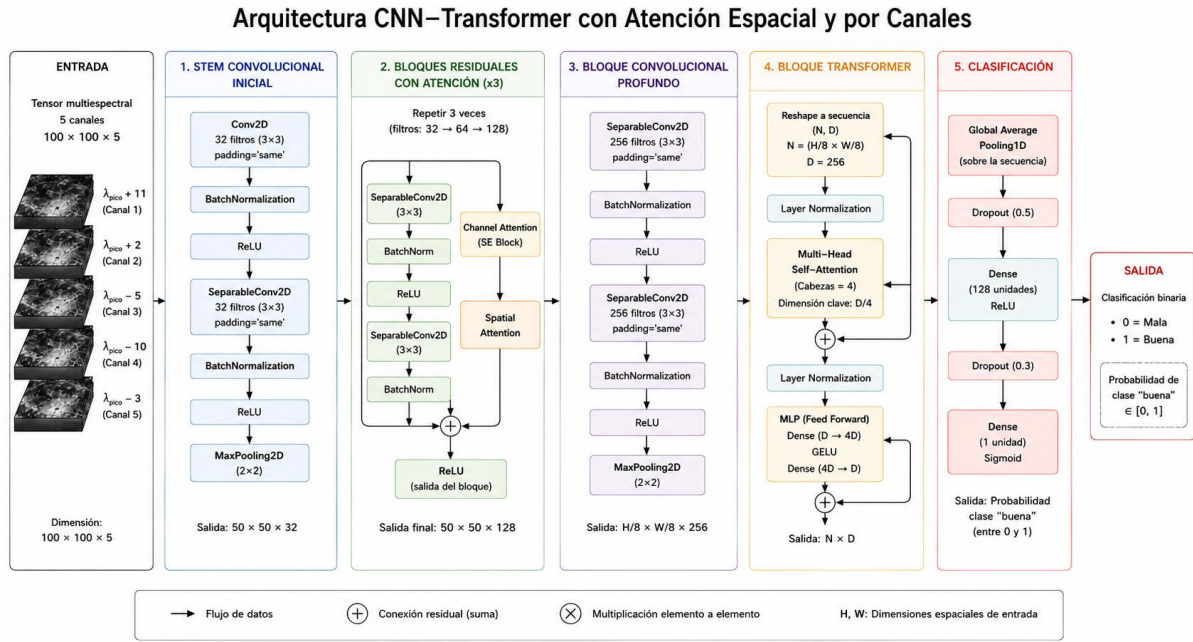
3.3.2 Modelo de clasificación espectral

Con el propósito de aprovechar la información contenida en las cinco bandas espectrales seleccionadas durante el preprocesamiento, se desarrolló un modelo de clasificación basado en información SWIR. A diferencia del modelo RGB, esta arquitectura fue diseñada para explotar tanto las características espaciales como la información espectral presente en cada muestra, buscando mejorar la capacidad de discriminación entre frutos aptos y no aptos.

Durante el desarrollo del modelo se evaluaron diferentes configuraciones de aprendizaje profundo, motivadas principalmente por la variabilidad espectral observada entre las distintas sesiones de adquisición. Esta condición incrementó la complejidad del problema y requirió analizar diferentes estrategias antes de seleccionar la arquitectura definitiva. Como resultado, se implementó una arquitectura híbrida CNN–Transformer con mecanismos de atención espacial y por canales, compuesta por convoluciones separables, bloques residuales, mecanismos de atención y un bloque Transformer encargado de modelar relaciones globales entre las características extraídas por la red.

Figura 10.

Arquitectura híbrida CNN–Transformer desarrollada para la clasificación de las muestras espectrales.



Nota. La figura presenta la arquitectura híbrida CNN–Transformer utilizada para la clasificación binaria de muestras espectrales de cinco canales. El modelo integra bloques convolucionales, conexiones residuales con mecanismos de atención espacial y por canales, un bloque Transformer y capas densas para generar la clasificación final. Elaboración propia.

Los detalles correspondientes a la evolución de las arquitecturas evaluadas, la configuración del modelo seleccionado y el procedimiento de entrenamiento se presentan en el **Apéndice C**.

3.3.3 Adaptación del detector YOLO al dominio espectral

El sistema de detección empleado durante este trabajo se basó en un modelo YOLO desarrollado previamente por el grupo de investigación HDSP para la localización de frutos cítricos en imágenes RGB. Debido a que el sistema propuesto utiliza información proveniente

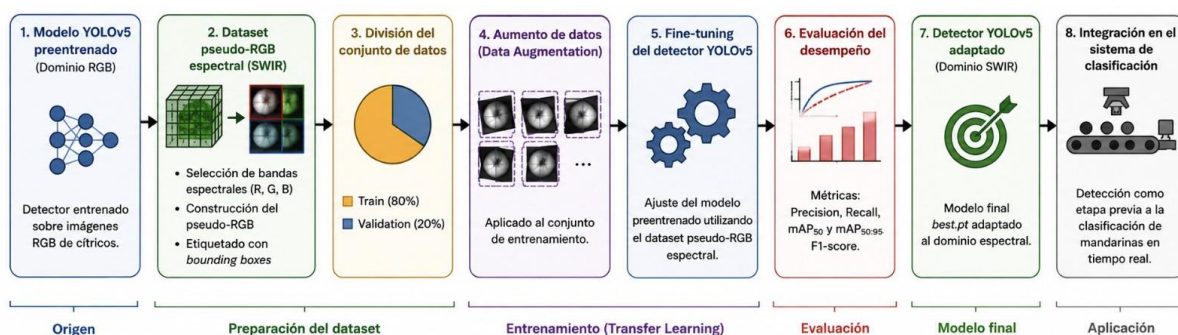
de una cámara espectral SWIR, fue necesario adaptar dicho modelo al nuevo dominio mediante un proceso de transferencia de aprendizaje (fine-tuning).

Para ello, el modelo fue reentrenado utilizando las representaciones pseudo-RGB generadas a partir de las bandas espectrales seleccionadas durante la etapa de preprocesamiento. Este procedimiento permitió conservar las capacidades de localización del detector original, al tiempo que mejoró su desempeño sobre las nuevas representaciones espectrales utilizadas por el sistema.

Una vez finalizado el proceso de ajuste, el modelo fue integrado al flujo de inferencia en tiempo real, permitiendo detectar automáticamente las mandarinas presentes en la escena y generar las regiones de interés utilizadas posteriormente por el modelo de clasificación.

Figura 11.

Flujo general del proceso de adaptación del detector YOLO al dominio espectral mediante fine-tuning.



Los procedimientos complementarios relacionados con la preparación del conjunto de datos, configuración del entrenamiento y almacenamiento del modelo ajustado se presentan en el **Apéndice C**.

3.4 Evaluación del desempeño de los modelos

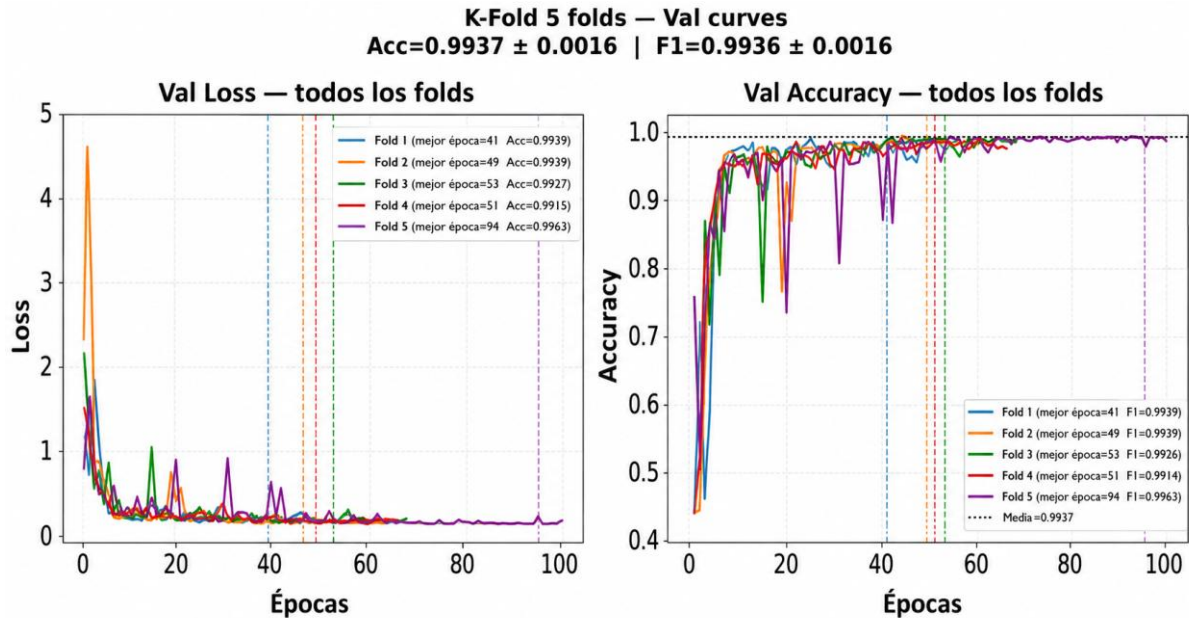
Con el propósito de evaluar el desempeño de los modelos desarrollados durante la etapa anterior, se emplearon diferentes métricas ampliamente utilizadas en problemas de clasificación binaria y detección de objetos. La evaluación permitió analizar la capacidad de generalización de cada modelo y seleccionar la configuración con mejor desempeño para su integración dentro del sistema automatizado de clasificación.

3.4.1 Evaluación del modelo RGB

La evaluación del modelo RGB se realizó mediante un procedimiento de validación cruzada estratificada de cinco particiones, permitiendo estimar la capacidad de generalización de la arquitectura sobre diferentes distribuciones del conjunto de datos. Durante cada iteración se registró la evolución de las métricas de entrenamiento y validación, así como el comportamiento de la función de pérdida, con el propósito de verificar la estabilidad del proceso de aprendizaje y detectar posibles fenómenos de sobreajuste.

Figura 12.

Curvas de entrenamiento y validación obtenidas durante el procedimiento de validación cruzada estratificada del modelo RGB.



Nota. La figura presenta las curvas de pérdida y exactitud de validación obtenidas para los cinco folds durante la validación cruzada estratificada del modelo RGB. Las líneas representan el comportamiento de cada fold a lo largo de las épocas, mientras que las líneas verticales indican la época correspondiente al mejor desempeño de cada fold.

Como se observa en la Figura 12, las curvas de entrenamiento y validación presentaron un comportamiento estable a lo largo de las diferentes particiones, evidenciando una adecuada convergencia del modelo y una diferencia reducida entre ambas etapas. Estos resultados indican que la arquitectura propuesta logró aprender representaciones discriminativas de las imágenes RGB sin presentar un sobreajuste significativo durante el entrenamiento.

Posteriormente, a partir de la información obtenida durante la validación cruzada, se entrenó un modelo definitivo utilizando la totalidad del conjunto de entrenamiento, el cual fue

empleado como referencia para las comparaciones posteriores con el modelo desarrollado a partir de información espectral.

Tabla 4.

Desempeño del modelo RGB durante la etapa de evaluación.

Métrica	Valor
Accuracy	0.9976
Precision	0.9656
Recall	1
F1-score	0.9975

Nota. Las métricas se calcularon sobre el conjunto de datos utilizado para la evaluación del modelo RGB. Accuracy corresponde a la exactitud, Precision a la precisión, Recall a la sensibilidad y F1-score a la media armónica entre precisión y sensibilidad.

Los resultados obtenidos evidencian que el modelo RGB constituye una línea base adecuada para el problema de clasificación planteado. No obstante, debido a que únicamente utiliza información correspondiente al espectro visible, su desempeño fue posteriormente comparado con un modelo desarrollado a partir de información espectral SWIR, con el propósito de analizar las ventajas asociadas al empleo de este tipo de información en la estimación de la calidad comercial de los frutos. El análisis detallado del procedimiento de entrenamiento, las configuraciones empleadas y las métricas obtenidas durante cada una de las particiones de validación cruzada se presentan en el **Apéndice D**.

3.4.2 Evaluación del modelo de clasificación espectral

El desempeño de modelo de clasificación espectral se evaluó mediante validación cruzada estratificada de cinco particiones y un conjunto Holdout independiente, conformado por cubos espectrales provenientes de sesiones de adquisición diferentes a las utilizadas durante

el entrenamiento, permitiendo estimar la capacidad de generalización del modelo frente a datos completamente nuevos.

La arquitectura CNN-Transformer obtuvo un F1 de 0.8075 durante la validación cruzada y un F1 de 0.7749 sobre el conjunto Holdout y un ROC (AUC) de 0.8125. Estos resultados evidencian un comportamiento consistente entre etapas de entrenamiento y evaluación, permitiendo seleccionar esta arquitectura para su integración dentro del sistema automatizado de clasificación.

Tabla 5.

Métricas obtenidas por el modelo CNN-Transformer.

Métrica	Validación cruzada	Holdout
Accuracy	0.8102	0.7750
Precision Macro	0.8126	0.7757
Recall Macro	0.8096	0.7750
F1 Macro	0.8075	0.7749
ROC-AUC	0.8870	0.8125

Nota. Los valores de validación cruzada corresponden al promedio obtenido durante el procedimiento de validación cruzada, mientras que los valores de Holdout corresponden al desempeño del modelo sobre el conjunto de datos reservado para evaluación. Accuracy representa la exactitud, Precision Macro la precisión promedio entre las clases, Recall Macro la sensibilidad promedio, F1 Macro la media armónica de precisión y sensibilidad, y ROC-AUC el área bajo la curva característica operativa del receptor.

Los procedimientos de entrenamiento y la evaluación detallada del modelo, incluyendo las curvas de aprendizaje y la matriz de confusión, se presentan en los **Apéndices C y D**, respectivamente.

3.4.3. Evaluación del detector YOLO

Con el propósito de evaluar el desempeño del detector de objetos implementado durante el desarrollo del sistema, se analizaron las métricas obtenidas después del proceso de fine-tuning realizado sobre el modelo YOLOv5m utilizando las representaciones pseudo-RGB generadas a partir de la información espectral. La evaluación permitió verificar la capacidad del detector para localizar automáticamente las mandarinas presentes en la escena antes de la etapa de clasificación.

Tabla 6.

Métricas de desempeño obtenidas por el detector YOLOv5m después del proceso de fine-tuning.

Métrica	Valor
Precisión	0,979
Recall	0,999
mAP@0.5	0,993
mAP@0.5:0.95	0,917

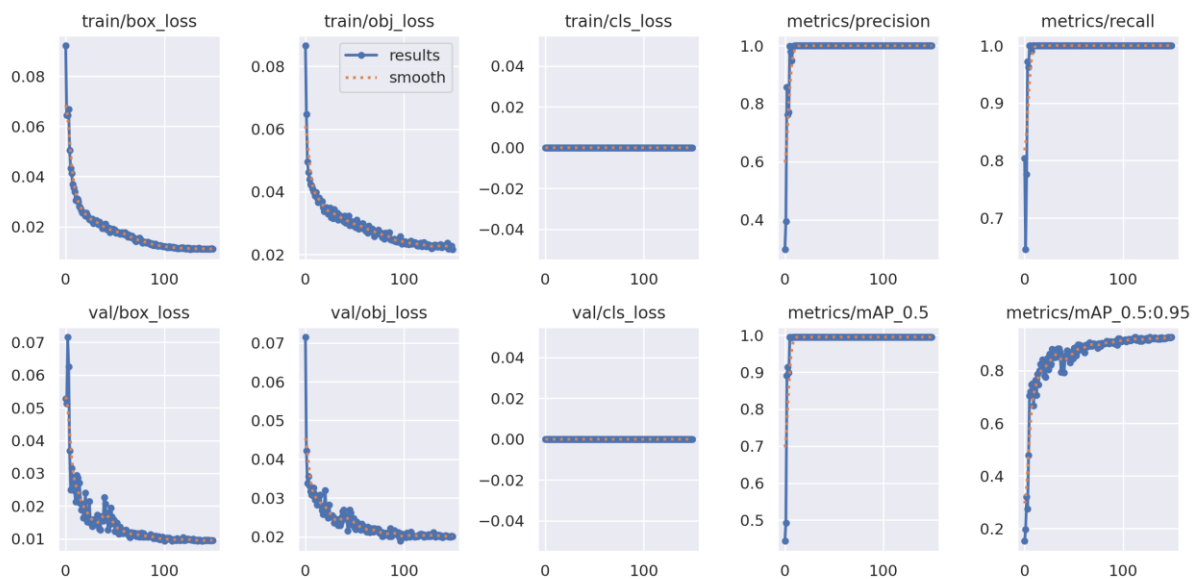
Nota. Las métricas corresponden al desempeño del detector YOLOv5m después del proceso de fine-tuning. Precisión corresponde a la proporción de detecciones positivas correctas, Recall a la proporción de objetos correctamente detectados, mAP@0.5 al promedio de precisión media con un umbral de IoU de 0,5 y mAP@0.5:0.95 al promedio de precisión media calculado para umbrales de IoU entre 0,5 y 0,95.

Los resultados obtenidos evidencian un elevado desempeño del detector sobre las representaciones pseudo-RGB, alcanzando valores superiores al 97 % en precisión y una sensibilidad cercana al 100 %. De igual manera, el valor de mAP@0.5 obtenido demuestra que el modelo fue capaz de localizar correctamente las mandarinas presentes en las imágenes

utilizadas durante la evaluación, mientras que el valor de $mAP@0.5:0.95$ confirma un comportamiento consistente incluso al considerar diferentes niveles de solapamiento entre las predicciones y las anotaciones de referencia.

Figura 13.

Curvas de entrenamiento obtenidas durante el proceso de fine-tuning del modelo YOLOv5m.



Nota. La figura presenta la evolución de las métricas de entrenamiento y validación del modelo YOLOv5m durante el proceso de fine-tuning. Las curvas muestran la pérdida asociada al entrenamiento y la validación, así como las métricas de precisión, recall, $mAP@0.5$ y $mAP@0.5:0.95$ a lo largo de las épocas.

Durante el análisis del comportamiento del detector también se evaluó la relación entre el umbral de confianza y el desempeño del modelo. A partir de la curva F1–Confidence se observó que el máximo valor de F1 se alcanzó para un umbral de confianza aproximado de 0,343. Sin embargo, el detector mantuvo un comportamiento estable en un amplio intervalo de valores de confianza, permitiendo seleccionar un umbral operativo de 0,60 para la

implementación final del sistema. De manera complementaria, el umbral de intersección sobre unión (IoU) se fijó en 0,45, garantizando un equilibrio entre la sensibilidad del detector y la reducción de detecciones redundantes durante la operación en tiempo real.

Los procedimientos detallados relacionados con la estrategia de transferencia de aprendizaje, configuración del entrenamiento, selección de hiperparámetros y almacenamiento del modelo ajustado se presentan en el **Apéndice C**.

3.5 Implementación del prototipo mecatrónico

Para la implementación del módulo de separación se partió de la banda transportadora disponible en el laboratorio HDSP, sobre la cual se diseñó una estructura adaptable a las dimensiones y condiciones de operación existentes. A partir de las medidas obtenidas directamente sobre la banda, se planteó un sistema de clasificación compuesto por dos mecanismos de accionamiento independientes, ubicados al final del recorrido de los frutos.

Durante la etapa inicial se evaluaron dos mecanismos de apertura accionados mediante servomotores y fabricados mediante impresión 3D. Ambos diseños fueron instalados sobre una plataforma provisional y sometidos a pruebas mecánicas para analizar su funcionamiento frente al desplazamiento de las mandarinas. Como resultado de esta comparación, se seleccionó el mecanismo con apertura ascendente, debido a que su configuración permitía mantener la compuerta apoyada durante la posición de reposo y reducía las perturbaciones ocasionadas por el peso de los frutos. El segundo mecanismo, aunque presentaba un movimiento adecuado, fue descartado debido a que la apertura descendente permitía que el peso de las mandarinas alterara la posición de reposo de la compuerta. Los procedimientos de selección, comparación y adaptación del mecanismo se presentan con mayor detalle en el **Apéndice E**.

Figura 14.

Mecanismo de accionamiento seleccionado.



Una vez seleccionado el mecanismo, se diseñó la estructura definitiva para su integración con la banda transportadora. El módulo fue dispuesto al final de la banda con una inclinación aproximada de 10° , permitiendo aprovechar el desplazamiento por gravedad de las mandarinas hacia las salidas de clasificación. La estructura fue inicialmente construida y evaluada con materiales de bajo costo, lo que permitió verificar sus dimensiones, puntos de fijación y comportamiento mecánico antes de realizar la fabricación definitiva. A partir de los resultados obtenidos, las piezas fueron diseñadas utilizando software CAD y fabricadas en MDF de 4 [mm] por medio de corte láser, obteniendo una estructura con mayor precisión dimensional, rigidez y estabilidad.

El accionamiento del sistema se implementó mediante dos servomotores MG90S, uno para cada mecanismo de apertura, controlados desde una tarjeta Arduino UNO. Las señales de control fueron conectadas a los pines digitales del microcontrolador, mientras que la alimentación de los actuadores se realizó a través una fuente externa de 5 [V] y 2 [A]. Posteriormente, se realizó la calibración individual de los servomotores empleando una

herramienta desarrollada en Python, obteniendo experimentalmente los ángulos de reposo y apertura requeridos por cada mecanismo. Los procedimientos de diseño, fabricación, integración electrónica y calibración se presentan detalladamente en el **Apéndice E**.

Figura 15.

Prototipo mecatrónico de separación implementado sobre la banda transportadora



Como resultado de esta fase se obtuvo un módulo mecatrónico compuesto por dos mecanismos de separación independientes, integrado estructural y electrónicamente con la banda transportadora. Las pruebas de accionamiento realizadas permitieron verificar la repetibilidad del movimiento de las compuertas y el retorno de los servomotores a sus posiciones de reposo. Adicionalmente, la incorporación de una guía central en la entrada del módulo permitió mantener la separación entre frutos próximos provenientes de los dos carriles virtuales y orientar su desplazamiento hacia el mecanismo correspondiente.

El funcionamiento final del módulo se estableció de acuerdo con la decisión generada por el sistema de clasificación. Cuando un fruto es identificado como apto, se acciona el servomotor correspondiente y la compuerta permanece abierta durante aproximadamente 400 ms, permitiendo su caída hacia la canasta de aceptación. Para los frutos clasificados como no aptos, la compuerta permanece en posición de reposo y la mandarina continúa su desplazamiento hasta la salida de rechazo. De esta manera, la implementación obtenida

permitió realizar físicamente la separación de los frutos en las dos categorías establecidas por el sistema.

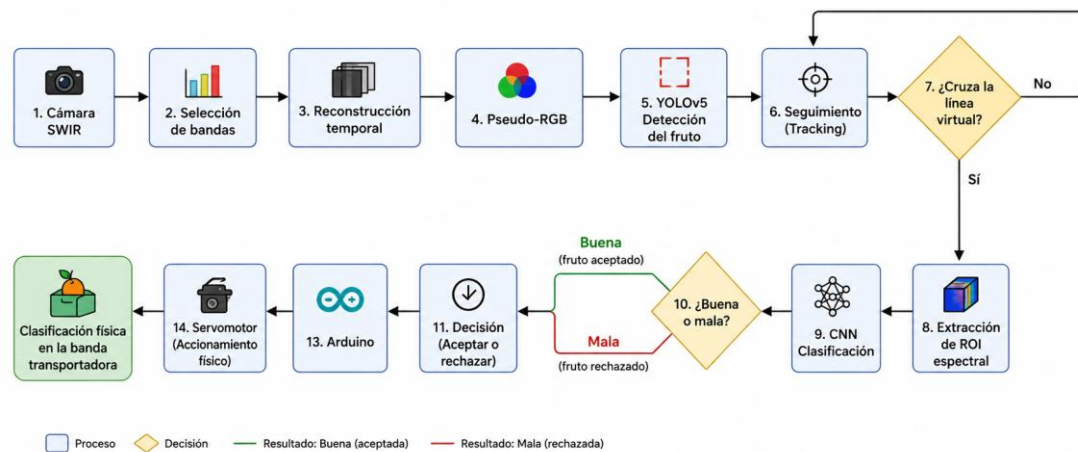
3.6 Integración y sincronización del sistema de adquisición e inferencia

Una vez desarrollados los módulos de adquisición, procesamiento, detección, clasificación y accionamiento, se realizó su integración para permitir la operación automatizada del sistema sobre la banda transportadora. El flujo implementado inicia con la adquisición continua de líneas espectrales mediante la cámara SWIR, a partir de las cuales se reconstruye progresivamente la escena correspondiente a la zona de captura utilizando las cinco bandas espectrales seleccionadas durante el análisis previo.

A partir de esta reconstrucción se genera una representación pseudo-RGB mediante la combinación de tres de las bandas seleccionadas, sobre la cual el modelo YOLOv5 realiza la detección de las mandarinas presentes. Posteriormente, el sistema de seguimiento asigna un identificador a cada fruto y mantiene su correspondencia durante el desplazamiento por la zona de captura. Cuando el centroide de un fruto cruza la línea virtual de clasificación, se determina el carril correspondiente, se extrae la región espectral de interés y el modelo CNN procesa el conjunto de cinco bandas para establecer la categoría apta o no apta.

Figura 16.

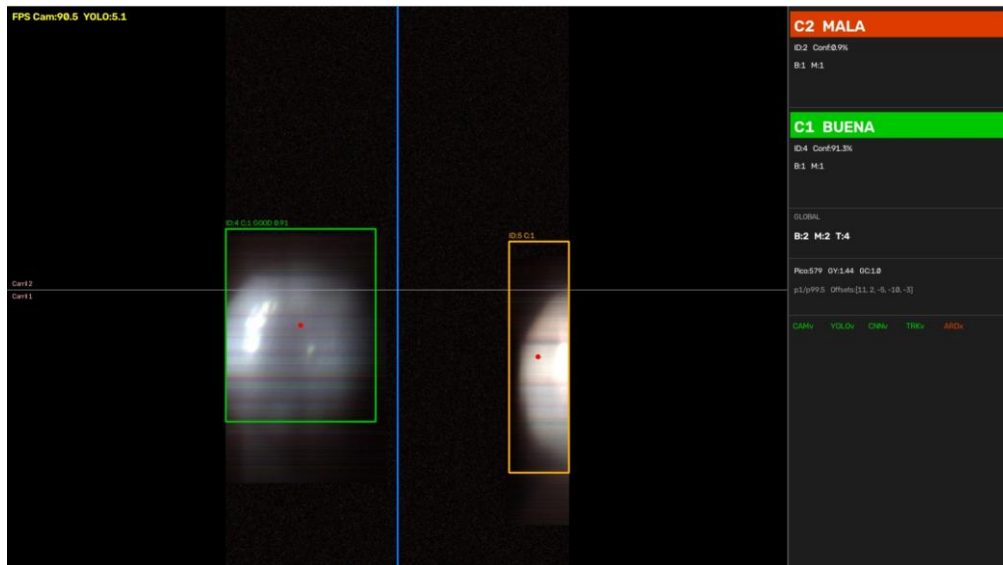
Arquitectura general del sistema integrado de adquisición, procesamiento, inferencia y accionamiento.



La integración de estos módulos permitió mantener la correspondencia entre cada fruto detectado, su ubicación sobre la banda transportadora, la decisión generada por el modelo de clasificación y el mecanismo encargado de realizar la separación física. La Figura 17 presenta la interfaz desarrollada para supervisar la operación del sistema, donde se visualizan la reconstrucción pseudo-RGB, las detecciones realizadas, el seguimiento de los frutos, la asignación de carriles y el resultado de clasificación de cada muestra.

Figura 17.

Interfaz del sistema durante la detección, seguimiento y clasificación automática de los frutos.

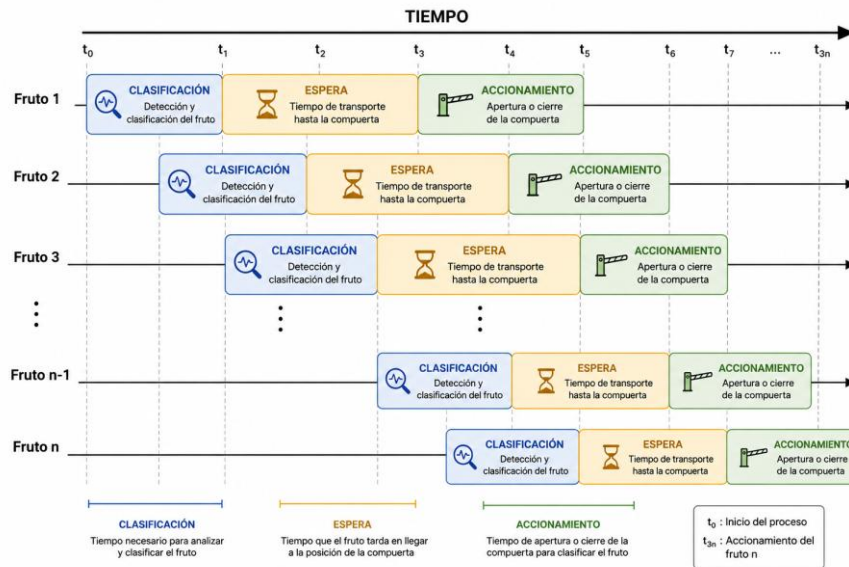


Con el fin de sincronizar la clasificación con la llegada de los frutos al mecanismo de separación, se implementó una planificación temporal no bloqueante basada en eventos independientes. El tiempo de retardo entre el cruce del centroide por la línea virtual de clasificación y el accionamiento del mecanismo fue determinado experimentalmente para una velocidad de operación constante de 120 rpm, obteniéndose un valor aproximado de 40,5 s. Este parámetro puede ajustarse cuando cambian las condiciones de operación del sistema.

La estrategia implementada permite que cada fruto clasificado genere un evento programado de manera independiente, mientras el sistema continúa realizando simultáneamente las tareas de adquisición, reconstrucción de la escena, detección, seguimiento y clasificación de nuevos frutos, evitando interrupciones en el flujo principal de procesamiento.

Figura 18.

Representación temporal de la planificación no bloqueante de eventos de accionamiento para frutos consecutivos.



Nota. La figura muestra la planificación temporal de los procesos de clasificación, espera y accionamiento de las compuertas para frutos que ingresan de manera consecutiva. La ejecución no bloqueante permite gestionar de forma independiente los tiempos asociados a cada fruto, evitando que el procesamiento de uno interfiera con el seguimiento de los siguientes.

Una vez alcanzado el instante programado, la aplicación principal envía la orden de accionamiento al Arduino UNO mediante comunicación serial, el cual ejecuta el movimiento del servomotor asociado al carril correspondiente. En ausencia de una orden de accionamiento, las compuertas permanecen en su posición de reposo y el fruto continúa hacia la salida de rechazo.

Finalmente, la integración desarrollada permitió establecer un flujo continuo desde la adquisición de la información espectral hasta el accionamiento físico del mecanismo de separación. El sistema ejecuta de manera coordinada las etapas de reconstrucción de la escena,

detección, seguimiento, extracción de la información espectral, clasificación, planificación temporal y accionamiento. Los procedimientos de implementación relacionados con la arquitectura software, reconstrucción mediante RollingBuffer, seguimiento, planificación de eventos, comunicación serial y funcionamiento de la interfaz se presentan de forma detallada en el **Apéndice F**.

3.7 Pruebas experimentales

Con el propósito de validar el funcionamiento del sistema desarrollado, se realizaron pruebas experimentales sobre el prototipo integrado, evaluando progresivamente los módulos de seguimiento, sincronización, clasificación y separación física antes de la validación completa del sistema. Esto permitió verificar el funcionamiento de cada etapa y realizar los ajustes necesarios durante la integración.

Inicialmente, se efectuaron pruebas del mecanismo de separación modificando temporalmente el umbral de decisión del modelo para generar órdenes de accionamiento controladas. Con ello se comprobó la comunicación entre el sistema de inferencia y el Arduino, así como la correcta planificación temporal y el accionamiento de las compuertas.

Posteriormente, se evaluó el desempeño del sistema bajo diferentes configuraciones de operación, incluyendo frutos simultáneos en ambos carriles, frutos consecutivos y distintas separaciones sobre la banda transportadora, verificando el seguimiento de múltiples objetos y la correcta ejecución de la planificación temporal.

Figura 19.

Configuraciones evaluadas durante las pruebas experimentales del sistema integrado.



Tabla 7.

Escenarios considerados durante las pruebas experimentales.

Escenario evaluado	Objetivo
Frutos simultáneos en ambos carriles	Verificar el funcionamiento independiente de cada mecanismo de separación.
Frutos consecutivos	Evaluar la planificación temporal de eventos de accionamiento.
Frutos con diferentes separaciones	Comprobar el seguimiento y la clasificación de múltiples frutos.
Sistema integrado	Validar la operación continua del prototipo completo.

Nota. Los escenarios fueron definidos para evaluar de manera independiente y conjunta el funcionamiento de los mecanismos de separación, la planificación temporal de los eventos de accionamiento y la operación integrada del prototipo.

Finalmente, se comprobó que el sistema requiere mantener condiciones estables de adquisición, especialmente en la posición de la cámara SWIR y la iluminación, ya que variaciones en estas afectan la respuesta espectral y pueden requerir una nueva selección de

bandas para la clasificación. Bajo condiciones controladas, el sistema ejecutó de forma continua las etapas de adquisición, detección, clasificación y separación física de los frutos. Los resultados experimentales y el análisis detallado de las pruebas se presentan en el **Apéndice G**.

4. Conclusiones

Se diseñó e implementó un sistema funcional de clasificación automática de mandarinas que integra adquisición espectral, reconstrucción de imágenes, detección mediante YOLO, clasificación basada en aprendizaje profundo y actuación física mediante un sistema mecatrónico controlado por Arduino. La integración de estos componentes permitió construir un prototipo capaz de ejecutar de forma continua el proceso de clasificación automática sobre una banda transportadora.

Se implementó un modelo de clasificación basado en imágenes RGB y un modelo de clasificación espectral utilizando una arquitectura híbrida CNN-Transformer. El modelo RGB obtuvo una exactitud (Accuracy) de 99.76% y un F1-score de 99.75% sobre el conjunto de evaluación, mientras que el modelo espectral alcanzó una exactitud (Accuracy) de 77.5%, un F1-score de 77.49% y un ROC-AUC de 81.25% sobre el conjunto de test, demostrando la viabilidad del uso de información espectral para la clasificación automática de frutos.

La adaptación del YOLOv5 al dominio espectral mediante representaciones pseudo-RGB permitió localizar automáticamente las mandarinas sobre la banda transportadora con desempeño elevado (Precision = 97.9%, Recall = 99.9%, mAP@0.5 = 99.3%, y mAP@0.5:0.95 = 91.7%), confirmando que la transferencia de aprendizaje es una estrategia efectiva para adaptar modelos entrenados sobre imágenes RGB al dominio espectral.

La arquitectura modular desarrollada permitió desacoplar las etapas de adquisición, preprocesamiento, detección, seguimiento, clasificación, planificación temporal y accionamiento, facilitando la integración del sistema completo y permitiendo el procesamiento continuo de múltiples frutos sin interrumpir el flujo principal de operación.

Las pruebas experimentales realizadas sobre el prototipo verificaron el funcionamiento integrado del sistema bajo diferentes escenarios de operación, incluyendo frutos consecutivos, múltiples frutos simultáneos y diferentes separaciones sobre la banda transportadora. Asimismo, la comparación entre la solución RGB y la solución espectral evidenció que, aunque el procesamiento RGB presenta menores requerimientos computacionales y tiempos de inferencia más reducidos, la incorporación de información espectral proporciona información complementaria que no puede obtenerse únicamente a partir del color visible, justificando su utilización en aplicaciones de clasificación automática de calidad de frutos cítricos.

5. Recomendaciones

Ampliar el conjunto de datos espectral, incorporando nuevas campañas de adquisición, condiciones de iluminación y diferentes variedades de cítricos, con el fin de fortalecer la capacidad de generalización de los modelos.

Evaluar arquitecturas de clasificación más complejas cuando se disponga de un mayor volumen de datos, ya que modelos con mayor capacidad podrían beneficiarse de una representación espectral más diversa.

Validar el sistema en condiciones operativas más cercanas a un entorno industrial, evaluando su desempeño frente a variación de velocidad de banda, iluminación y condiciones ambientales.

Implementar mecanismos de calibración automática para reducir la intervención manual requerida durante la puesta en marcha y mantener la estabilidad de la respuesta espectral entre diferentes sesiones de adquisición.

Extender el sistema a otros productos agrícolas, aprovechando la modularidad de la arquitectura desarrollada para adaptar el pipeline a nuevos problemas de clasificación mediante reentrenamiento de los modelos correspondientes.

Se recomienda realizar una verificación previa de la calibración del sistema de adquisición espectral antes de cada sesión experimental. En las pruebas desarrolladas fue necesario comprobar el pico de respuesta espectral media para garantizar una referencia consistente en la selección de las bandas espectrales, ya que pequeñas variaciones en la configuración del sistema pueden modificar la respuesta registrada respecto a sesiones anteriores.

Referencias Bibliográficas

- Abade, A., Ferreira, P. A., & de Barros Vidal, F. (2021). Plant diseases recognition on images using convolutional neural networks: A systematic review. *Computers and Electronics in Agriculture*, 185, 106125. <https://doi.org/10.1016/j.compag.2021.106125>
- Aguilar, I. (2017). Sistema automático de selección de limón (*Citrus latifolia tanaka*) basado en discriminación por color. *Agroproductividad*, 10(9). <https://revista-agroproductividad.org/index.php/agroproductividad/article/view/85>
- Al-Hourani, A., Balendhran, S., Walia, S., & Hourani, T. (2023). Line scan hyperspectral imaging framework for open source low-cost platforms. *Remote Sensing*, 15(11), 2787. <https://doi.org/10.3390/rs15112787>
- Bhargava, A., & Bansal, A. (2021). Fruits and vegetables quality evaluation using computer vision: A review. *Journal of King Saud University - Computer and Information Sciences*, 33(3), 243–257. <https://doi.org/10.1016/j.jksuci.2018.06.002>
- Blasco, J., Aleixos, N., Gómez, J., & Moltó, E. (2007). Citrus sorting by identification of the most common defects using multispectral computer vision. *Journal of Food Engineering*, 83(3), 384–393. <https://doi.org/10.1016/j.jfoodeng.2007.03.027>
- CODEX Alimentarius FAO-WHO. (s. f.). Página inicial. <https://www.fao.org/fao-who-codexalimentarius/es/>
- Cubero, S., Lee, W.S., Aleixos, N. et al. Automated Systems Based on Machine Vision for Inspecting Citrus Fruits from the Field to Postharvest—a Review. *Food Bioprocess Technol* 9, 1623–1639 (2016). <https://doi.org/10.1007/s11947-016-1767-1>
- De la Cruz-Rocca, M. A., Muñoz-Robles, W., & Trinidad-Malpartida, M. A. (2024). Análisis de la aplicación de redes neuronales convolucionales en el área de la visión artificial. *Investigación Valdizana*, 18(1), e2105. <https://doi.org/10.33554/riv.18.1.2105>

- Diaz Ramirez, R. (2009). Herramienta computacional para el apoyo en la clasificación de frutas en la agroindustria [Trabajo de grado de pregrado, Universidad Pontificia Bolivariana]. Repositorio Institucional.
<https://repository.upb.edu.co/server/api/core/bitstreams/e7dbb181-c7d5-45ed-82c3-0f3d89958f74/content>
- Duan, Z., Zhang, L., Ji, J., Wang, G., & Wang, Y. (2026). Design of fruit sorting system based on PLC. Journal of Physics: Conference Series, 3204(1), 012025.
<https://doi.org/10.1088/1742-6596/3204/1/012025>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. The MIT Press.
- Gunawan, C. R., Nurdin, & Fajriana. (2022). Design of a Real-Time Object Detection Prototype System With YOLOv3 (You Only Look Once). International Journal of Engineering, Science & Information Technology, 2(3), 96–99. <https://doi.org/10.52088/ijesty.v2i3.309>
- Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., Michael, K., TaoXie, Fang, J., imyhxy, Lorna, Yifu, Z., Wong, C., Abhiram, V., Montes, D., Wang, Z., Fati, C., Nadar, J., Laughing, ... Diaconu, D. (2022). ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation (v7.0) [Software]. Zenodo.
<https://doi.org/10.5281/zenodo.7347926>
- Kondo, Naoshi & Ahmad, Usman & Monta, Mitsuji & Murase, Haruhiko. (2000). Machine vision based quality evaluation of Iokan orange fruit using neural networks. Computers and Electronics in Agriculture. 29. 135-147. 10.1016/S0168-1699(00)00141-1.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. En Advances in neural information processing systems (pp. 1097–1105).
https://papers.nips.cc/paper_files/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
<https://doi.org/10.1038/nature14539>
- Mateus-Cagua, D., & Ordúz-Rodríguez, J. O. (2015). Mandarina Dancy: una nueva alternativa para la citricultura del piedemonte llanero de Colombia. *Corpoica Ciencia y Tecnología Agropecuaria*, 16(1), 105–118. http://www.scielo.org.co/scielo.php?pid=S0122-87062015000100010&script=sci_arttext&tlng=es
- Montoya Holguin, C., Cortés Osorio, J. A., & Chaves Osorio, J. A. (2014). Sistema automático de reconocimiento de frutas basado en visión por computador. *Ingeniare. Revista Chilena de Ingeniería*, 22(4), 504–516. <https://doi.org/10.4067/s0718-33052014000400006>
- Observatorio de Información Rural y Agropecuaria - Agronet. (s. f.). Agrícola. https://agronet.gov.co/estadisticas/agricola?producto_nacional=05070401&anio_fin_nacional=2024&ciclo_nacional_filter=
- Padilla Jimenez, A. V., Rojas, C. A. P., & Guzman, L. E. S. (2016). Procesamiento de imágenes para la identificación de plagas en los cultivos de espinaca. En 2016 IEEE Colombian Conference on Robotics and Automation (CCRA) (pp. 1–7). <https://ieeexplore.ieee.org/document/7811412>
- Patrício, D. I., & Rieder, R. (2018). Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. *Computers and Electronics in Agriculture*, 153, 69–81. <https://doi.org/10.1016/j.compag.2018.08.001>
- Pineda Cardozo, J. L., & Jiménez Buitrago, S. A. (2024). Diseño e implementación de un prototipo de banda transportadora para la identificación automática de productos cítricos mediante inteligencia artificial [Trabajo de grado de pregrado, Universidad Industrial de Santander]. Repositorio Institucional Noesis. <https://noesis.uis.edu.co/handle/20.500.14071/42452>

- Qin, J., Chao, K., Kim, M. S., Lu, R., & Burks, T. F. (2013). Hyperspectral and multispectral imaging for evaluating food safety and quality. *Journal of Food Engineering*, 118(2), 157–171. <https://doi.org/10.1016/j.jfoodeng.2013.04.001>
- Romero Acero, A., Marín Cano, A., & Jiménez Builes, O. A. (2015). Sistema de clasificación por visión artificial de mangostipo Tommy. *Revista UIS Ingenierías*, 14(1), 21–31. <https://revistas.uis.edu.co/index.php/revistausingenierias/article/view/21-31/4963>
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- UNECE. (2023). Frutas y verduras frescas - Normas. <https://unece.org/trade/wp7/FFV-Standards>
- Vignati, S., Tugnolo, A., Giovenzana, V., Pampuri, A., Casson, A., Guidetti, R., & Beghi, R. (2023). Hyperspectral imaging for fresh-cut fruit and vegetable quality assessment: Basic concepts and applications. *Applied Sciences*, 13(17), 9740. <https://doi.org/10.3390/app13179740>
- Vijayakumar, A., & Vairavasundaram, S. (2024). YOLO-based Object Detection Models: A Review and its Applications. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-18872-y>
- Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., & Liu, C. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. *Food Research International*, 62, 326–343. <https://doi.org/10.1016/j.foodres.2014.03.012>