

UNIVERSIDAD TÉCNICA DEL NORTE



Facultad de Ingeniería en Ciencias Aplicadas

Carrera de Software

Detección de defectos morfológicos en múltiples granos de café en una misma imagen utilizando redes neuronales convolucionales.

Trabajo de grado previo a la obtención del título de Ingeniero de Software
presentado ante la ilustre Universidad Técnica del Norte.

Autor:

Anderson Javier Chamorro Pinchao

Director:

Ing. Iván Danilo García Santillán, PhD

Ibarra – Ecuador

2025



UNIVERSIDAD TÉCNICA DEL NORTE

BIBLIOTECA UNIVERSITARIA

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE

1. IDENTIFICACIÓN DE LA OBRA

En cumplimiento del Art. 144 de la Ley de Educación Superior, hago la entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
CÉDULA DE IDENTIDAD:	0401865860		
APELLIDOS Y NOMBRES:	Chamorro Pinchao Anderson Javier		
DIRECCIÓN:	San Gabriel		
EMAIL:	ajchamorro1@utn.edu.ec andersonchamorro2000@live.com		
TELÉFONO FIJO:	0000-000	TELÉFONO MÓVIL:	0993536108

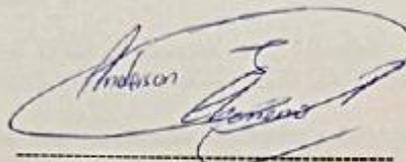
DATOS DE LA OBRA	
TÍTULO:	Detección de defectos morfológicos en múltiples granos de café en una misma imagen utilizando redes neuronales convolucionales.
AUTOR(ES):	Anderson Javier Chamorro Pinchao
FECHA:	01/07/2025
PROGRAMA:	PREGRADO
TÍTULO POR EL QUE OPTA:	INGENIERO DE SOFTWARE
DIRECTOR:	PhD. Iván Danilo García Santillán
ASESOR 1:	PhD. Marco Remigio Pusedá Chulde

2. CONSTANCIAS

El autor (es) manifiesta (n) que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es (son) el (los) titular (es) de los derechos patrimoniales, por lo que asume (n) la responsabilidad sobre el contenido de esta y saldrá (n) en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 1 días del mes de julio de 2025

EL AUTOR:

A handwritten signature in black ink, appearing to read "Anderson J. Chamorro Pinchao", written over a horizontal dashed line.

ESTUDIANTE

Anderson Javier
Chamorro Pinchao

C.I 0401865860

CERTIFICACIÓN DIRECTOR

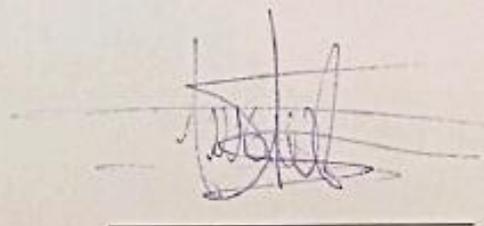
Ibarra, 1 de Julio del 2025

CERTIFICACIÓN DIRECTOR DEL TRABAJO DE TITULACIÓN

Por medio del presente yo PhD. Iván García, certifico que el Sr. Anderson Javier Chamorro Pinchao portador de la cedula de ciudadanía número 0401865860, ha trabajado en el desarrollo del proyecto de grado **"Detección de defectos morfológicos en múltiples granos de café en una misma imagen utilizando redes neuronales convolucionales"**, previo a la obtención del Título de Ingeniero en Software realizado con interés profesional y responsabilidad que certifico con honor de verdad.

Es todo en cuanto puedo certificar a la verdad

Atentamente



PhD. Iván García

DIRECTOR DE TRABAJO DE GRADO

DEDICATORIA

Al concluir esta importante etapa, dedico este logro con todo mi corazón a mi familia, cuyo apoyo y esfuerzo conjunto hicieron posible este camino.

A mis padres, mi ejemplo constante de perseverancia, honestidad y amor incondicional, gracias por creer en mí y sostenerme en cada paso y decisión.

A mis hermanas, compañeras inseparables, por sus valiosos consejos, su presencia constante y su apoyo inquebrantable.

Y a todas las personas que me brindaron su ayuda y ánimo durante este proceso, su respaldo fue fundamental para alcanzar esta meta.

Anderson Chamorro

AGRADECIMIENTO

Estoy profundamente agradecido con Dios por su guía y bendiciones en cada etapa de mi vida, y por darme la fortaleza necesaria para no rendirme en los momentos difíciles, lo que me permitió alcanzar esta importante meta.

A mis padres, Ramiro y Mirian, cuyo sacrificio aseguró que nunca me faltara nada, y cuyo ejemplo de amor y valores me formó como una persona honesta y responsable, apoyándome siempre en cada paso hacia este logro tan esperado.

A mis hermanas, Jacqueline y Hady, por su cariño constante, su compañía, orientación y apoyo en todas mis decisiones.

A mis amigos más cercanos, quienes me brindaron su amistad y compañía incondicional a lo largo de todo este camino.

A mi novia, por su respaldo inquebrantable durante todo el proceso de esta aventura.

A la Universidad Técnica del Norte, por la formación profesional recibida; y a mi tutor, Ing. Iván García PhD., por su paciencia y dedicación, que fueron fundamentales para la culminación exitosa de este trabajo.

Al Ingeniero Alberto Enríquez por darme la apertura a la finca “La Esperanza” para la realización de este trabajo, en calidad de propietario.

Anderson Chamorro

TABLA DE CONTENIDOS

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE	2
CERTIFICACIÓN DIRECTOR.....	4
DEDICATORIA.....	5
AGRADECIMIENTO.....	6
TABLA DE CONTENIDOS	7
ÍNDICE DE FIGURAS	11
ÍNDICE DE TABLAS	13
RESUMEN	14
ABSTRACT	15
INTRODUCCIÓN	16
Tema	16
Problema	16
Antecedentes	16
Situación Actual	16
Planteamiento del problema.....	17
Objetivos	18
Objetivo General	18
Objetivos Específicos	18
Alcance.....	19
Metodología.....	21
Justificación	23
Justificación Tecnológica.....	23
Justificación Agrícola	25

CAPÍTULO 1	26
1. Marco Teórico.....	26
1.1. Calidad de los granos de café.....	26
1.1.1. Características generales de los granos de café.....	26
1.1.2. Características morfológicas en los granos de café	28
1.1.3. Estándares de selección y calidad del grano de café.....	37
1.2. Visión por computador	41
1.2.1. Definición	41
1.2.2. Proceso.....	42
1.2.3. Algoritmos.....	44
1.3. Redes neuronales convolucionales.....	50
1.3.1. Redes neuronales.....	50
1.3.2. CNN.....	55
1.3.3. Aplicaciones de las CNN.....	57
1.4. Trabajos relacionados.....	58
1.4.1. Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia artificial y Visión por computador.....	58
1.4.2. Procesamiento de imágenes para la clasificación de granos de café por forma y color.....	59
1.4.3. Comparative analysis of YOLO models for green coffee bean detection and defect classification.....	60
1.4.4. Implementing a deep learning model for defect classification in Thai Arabica green coffee beans	61
1.4.5. Classification of Coffea Liberica Quality Using Convolution Neural Networks (Slim-CNN, YOLOv5, and VGG-16).....	62
1.4.6. Development of YOLO-Based Mobile Application for Detection of Defect Types in Robusta Coffee Beans.....	63
1.4.7. Enhanced Coffee Bean Defect Detection Using Super-Resolution GANs and Deep Learning	64

1.4.8.....	65
1.4.9. Coffee Bean Defects Automatic Classification Realtime Application Adopting Deep Learning	66
CAPÍTULO 2	68
2. Desarrollo	68
2.1. Adquisición de imágenes	68
2.2. Preprocesamiento de las imágenes	71
2.2.1. Data augmentation.....	72
2.2.2. Dataset	73
2.3. Arquitectura del modelo	74
2.3.1. YOLO versión 11 nano.....	75
2.3.2. YOLO versión 11 small	75
2.4. Entrenamiento del Modelo	76
2.4.1. Herramientas utilizadas.....	76
2.4.2. Proceso de Entrenamiento.....	78
2.4.3. Modelos entrenados.....	81
CAPÍTULO 3	85
3. Resultados.....	85
3.1. Métricas cuantitativas	85
3.1.1. Matriz de confusión.....	85
3.1.2. Precisión	89
3.1.3. RECALL.....	91
3.1.4. mAP.....	94
3.2. Discusión	97
3.3. Conclusiones	104
3.4. Recomendaciones	105

BIBLIOGRAFÍA 106

ÍNDICE DE FIGURAS

Figura 1 Árbol de problemas	18
Figura 2 Estructura de una CNN	20
Figura 3 Proceso de desarrollo del proyecto	20
Figura 4 Proceso de desarrollo.....	22
Figura 5 Exportaciones de café en Ecuador.....	23
Figura 6 Cinturón del café	28
Figura 7 Partes del grano de café	28
Figura 8 Proceso manual de tostado del producto	29
Figura 9 Herramienta utilizada para la recolección manual de café	30
Figura 10 Proceso de recolección y separación	32
Figura 11 Diferenciación entre frutos verdes y pintones.....	32
Figura 12 Herramienta para separación de tamaño	33
Figura 13 Maquina para despulpar el grano de café	34
Figura 14 Proceso de fermentación del café	34
Figura 15 Invernadero para el secado del café	35
Figura 16 Almacenamiento del grano de café	36
Figura 17 Grano cortado o dañado.....	38
Figura 18 Grano con restos de pulpa o cáscara adheridos	38
Figura 19 Grano dañado por humedad	39
Figura 20 Grano afectado por insectos	39
Figura 21 Grano con malformaciones	40
Figura 22 Grano con presencia de hongos.....	40
Figura 23 Grano con rotura de pergamino.....	41

Figura 24 Representación del modelo de color RGB en el espacio espacial.	45
Figura 25 Imagen de “Lenna” en color	46
Figura 26 Representación en canales R (a), G (b) y B (c) de ‘Lenna’.	46
Figura 27 Histogramas de la imagen cromática ‘Lenna’ de los mapas rojo (a), verde (b) y azul (c).	47
Figura 28 Funcionamiento de una CNN	52
Figura 29 Captura de imágenes de los granos de café	68
Figura 30 Esquema de captura de imágenes de los granos de café	69
Figura 31 Fotografías de café con fondo blanco	73
Figura 32 Fotografías de café con fondo variable	74
Figura 33 Notebook para permitir el acceso a los datos.....	76
Figura 34 Archivo data.yaml	77
Figura 35 Instalación de dependencias para el entrenamiento	78
Figura 36 API key para la obtención de los datos	79
Figura 37 Parámetros configurables para el entrenamiento	80
Figura 38 Gráfica de precisión de YOLO v11 Nano y diversos fondos.....	81
Figura 39 Gráfica de precisión de YOLO v11 Nano y fondo blanco	82
Figura 40 Gráfica de precisión de YOLO v11 Small y diversos fondos	83
Figura 41 Gráfica de precisión de YOLO v11 Small y fondo blanco.....	83
Figura 42 Matriz de confusión de yolo v11 nano fondo blanco.....	87
Figura 43 Matriz de confusión de yolo v11 nano fondo variable	87
Figura 44 Matriz de confusión de yolo v11 small fondo blanco	88
Figura 45 Matriz de confusión de yolo v11 SMALL fondo variable	88

ÍNDICE DE TABLAS

Tabla 1 Una descripción general de las funciones accesibles en cada herramienta, comparadas entre sí.....	69
Tabla 2 Capacidades de garantía de calidad y gestión dentro de las diversas herramientas.	70
Tabla 3 Diferencia entre Yolo nano y small.....	76
Tabla 4 Resultados de los modelos entrenados	84
Tabla 5 Matriz de confusión explicada	86
Tabla 6 Precisión modelo Yolo v11 nano fondo blanco	90
Tabla 7 Precisión modelo Yolo v11 nano fondo variable	90
Tabla 8 Precisión modelo Yolo v11 small fondo blanco	90
Tabla 9 Precisión modelo Yolo v11 small fondo variable	91
Tabla 10 Recall modelo Yolo v11 nano fondo blanco	92
Tabla 11 Recall modelo Yolo v11 nano fondo variable	92
Tabla 12 Recall modelo Yolo v11 small fondo blanco	93
Tabla 13 Recall modelo Yolo v11 small fondo variable.....	93
Tabla 14 mAP modelo Yolo versión 11 nano fondo blanco	95
Tabla 15 mAP modelo Yolo versión 11 nano fondo variable.....	96
Tabla 16 mAP modelo Yolo versión 11 small fondo blanco	96
Tabla 17 mAP modelo Yolo versión 11 small fondo variable	96
Tabla 18 Comparativa de trabajos propios y trabajos relacionados en datasets en granos de café	99
Tabla 19 Acceso y resumen de los datasets de los trabajos relacionados	100

RESUMEN

El café y los productos derivados del café forman parte de la alimentación mundial y son también uno de los principales productos de exportación de Imbabura (Ecuador). Los productores y vendedores de café deben asegurarse de que los granos de café que venden son de buena calidad, ya que un solo grano puede presentar varios defectos morfológicos. Sin embargo, seleccionar a mano los granos de buena calidad es un proceso largo, tedioso y laborioso, en el que es fácil cometer errores, ya que las personas se cansan y tienen diferentes formas de clasificar los productos. Utilizando redes neuronales convolucionales modernas, YOLO v11 (variantes Nano y Small), este trabajo tiene como objetivo clasificar automáticamente los granos de café de una imagen según su calidad. Recopilamos 7099 imágenes (512 × 640) y las dividimos en dos grupos según la calidad de los granos: buenos y malos. Los fondos de las imágenes se cambiaron a blanco, gris y negro para que se diferenciaron mejor. Utilizamos una GPU NVIDIA Tesla P100 para entrenar los modelos en Kaggle. YOLO v11 Nano tuvo una precisión de 0,880 y YOLO v11 Small tuvo una precisión de 0,871. Los tiempos de inferencia fueron de 7,89 ms (126,7 fps) y 10,08 ms (99,20 fps), respectivamente. Estos resultados son similares a los descritos en la bibliografía, teniendo en cuenta el problema de la superposición de los granos de café en cada imagen.

Palabras claves: clasificación de granos de café, defectos del café, CNN, YOLO, Pytorch.

ABSTRACT

Coffee and coffee products are part of the global diet and are also one of the main export products of Imbabura (Ecuador). Coffee producers and sellers must ensure that the coffee beans they sell are of good quality, as a single bean can have several morphological defects. However, selecting good quality beans by hand is a long, tedious, and laborious process in which it is easy to make mistakes, as people get tired and have different ways of classifying products. Using modern convolutional neural networks, YOLO v11 (Nano and Small variants), this work aims to automatically classify coffee beans in an image according to their quality. We collected 7099 images (512 × 640) and divided them into two groups according to the quality of the beans: good and bad. The backgrounds of the images were changed to white, gray, and black for better differentiation. We used an NVIDIA Tesla P100 GPU to train the models on Kaggle. YOLO v11 Nano had an accuracy of 0.880 and YOLO v11 Small had an accuracy of 0.871. The inference times were 7.89 ms (126.7 fps) and 10.08 ms (99.20 fps), respectively. These results are similar to those described in the literature, taking into account the problem of coffee bean overlap in each image.

Keywords: coffee bean classification, coffee defects, CNN, YOLO, Pytorch.

INTRODUCCIÓN

Tema

Detección de defectos morfológicos en múltiples granos de café en una misma imagen utilizando redes neuronales convolucionales

Problema

Antecedentes

El café es una de las bebidas más populares a nivel mundial en la actualidad, y debido a su amplio consumo, se busca constantemente mejorar su calidad. Según la FAO (Organización de las Naciones Unidas para la Alimentación y la Agricultura), se prevé un crecimiento significativo y sostenido en la producción mundial de café. En este sentido, es muy probable que América Latina y el Caribe continúen siendo la principal región productora de café a nivel global [1].

Con la esperanza de que la producción siga creciendo. El Ministerio de Agricultura y Ganadería (MAGAP) de Ecuador está trabajando para mejorar tanto la calidad como la cantidad del café. Se enfrentan a muchos problemas, como granos con defectos morfológicos que dificultan su selección y clasificación [2].

Situación Actual

En este caso, la finca La Esperanza, situada en Santa Catalina de Salinas, en la provincia de Imbabura, cultiva café en las mejores condiciones climáticas posibles. Sin

embargo, tiene problemas económicos y de personal que dificultan la selección de los granos, lo que reduce la calidad y el valor comercial del producto. Esta ineficiencia en la clasificación no solo reduce los beneficios económicos esperados, sino que también encarece el proceso debido al tiempo y recursos requeridos para llevarlo a cabo. Además, se desarrolló la tesis "Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia Artificial y Visión por Computador" [3]. Donde se identificó que la falta de diversidad en los conjuntos de datos de imágenes ya que solo presenta un grano por imagen representa un obstáculo importante para la detección eficaz y más cercano a la realidad debido a esta limitación.

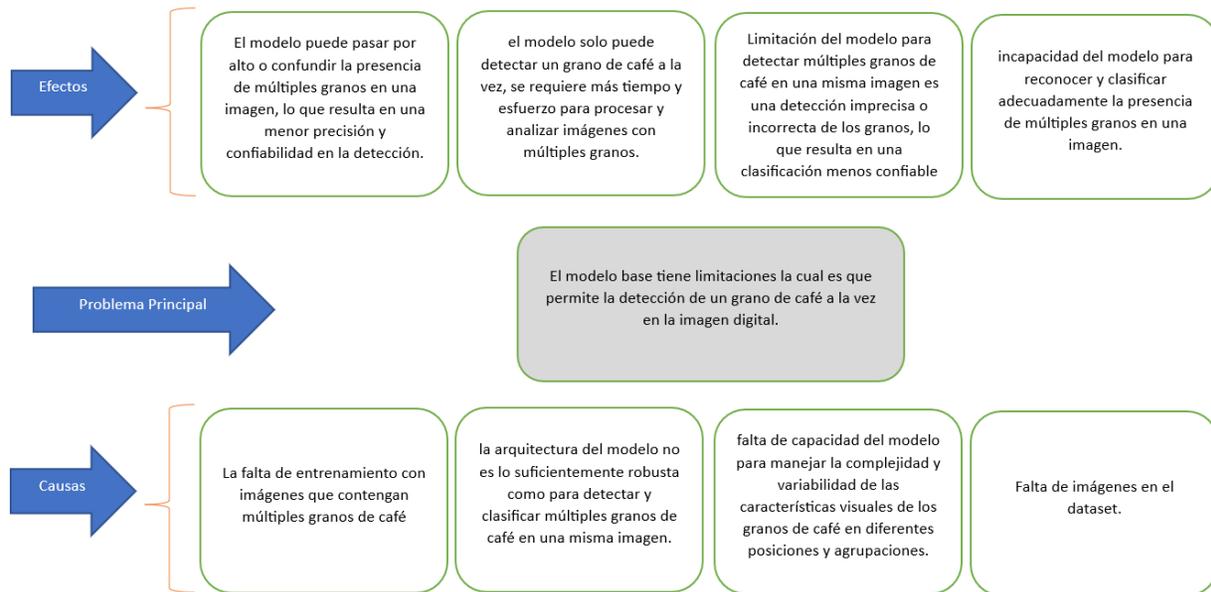
Planteamiento del problema

Las dificultades que experimentan los pequeños productores de café en la etapa de postcosecha para los obligan a dedicar una cantidad considerable de tiempo y recursos a la selección manual de los granos: deben comprobar que no haya deformaciones, heridas o colores inapropiados para cumplir con los estándares y requisitos de calidad de los mercados nacional e internacional. Esto no solo aumenta el costo total de la postcosecha, sino que también reduce la competitividad de los productores. Si el proceso de selección no se realiza adecuadamente, la calidad del café disminuye, lo que impacta negativamente en el precio recibido y puede limitar su comercialización en mercados internacionales. Además, factores como la variabilidad climática y la falta de infraestructura adecuada agravan estos problemas, dificultando aún más la capacidad de los productores para mantener un café de alta calidad. En este contexto, se ha desarrollado un árbol de problemas (Figura 1), que respalda esta

investigación, destacando la necesidad urgente de abordar estos desafíos para mejorar la situación del sector cafetalero.

Figura 1

Árbol de problemas



Objetivos

Objetivo General

Identificar defectos morfológicos en múltiples granos de café en una misma imagen digital utilizando redes neuronales convolucionales.

Objetivos Específicos

- Actualizar el marco teórico con los últimos trabajos desarrollados relacionado a la detección de defectos morfológicos en granos de café.

- Ampliar el dataset de granos de café para el entrenamiento del modelo de detección de defectos morfológicos, incluyendo varios granos de café y una mayor variedad de características.
- Desarrollar un modelo de CNN para la detección de defectos morfológicos considerando la existencia de múltiples granos de café en la misma imagen.
- Validar los resultados del modelo entrenado mediante métricas y gráficas cuantitativas de rendimiento.

Alcance

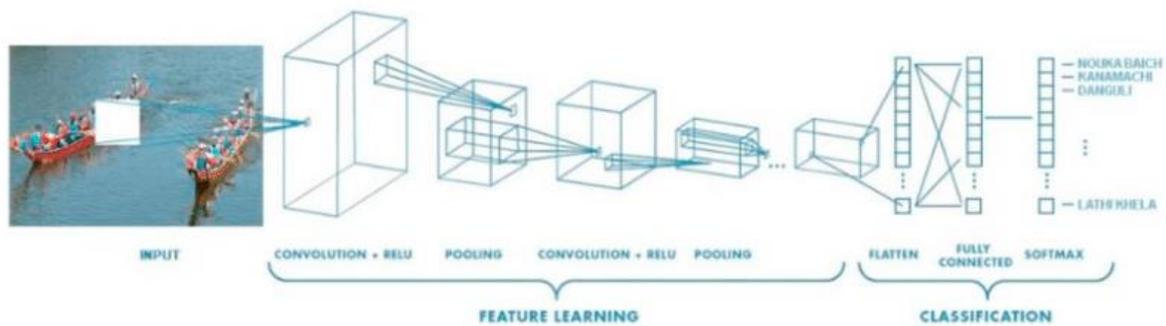
El proyecto se llevará a cabo en un entorno controlado con la mejor iluminación y una cámara RGB fija que siempre tomará las mismas imágenes. Utilizaremos estas imágenes para crear y enseñar un modelo de IA. El objetivo principal es facilitar la búsqueda y clasificación de diferentes tipos de granos de café en una sola imagen, sin los problemas que surgieron en estudios anteriores, como el realizado por [3]. Para alcanzar este objetivo, tenemos previsto recopilar unas 3500 imágenes y utilizar técnicas avanzadas de preprocesamiento, como el ajuste del brillo y el contraste, la eliminación del ruido y la corrección de la distorsión, para aumentar el conjunto de datos a 25 000. Python es una buena opción porque es adecuado para las matemáticas y la estadística y cuenta con marcos especializados como TensorFlow/Keras, PyTorch, OpenCV y Label Studio, que son buenos para las redes neuronales y se pueden utilizar de muchas maneras diferentes. Puedes ver los pasos que debes seguir en la (Figura 3).

La arquitectura del modelo será una red neuronal convolucional (CNN) en la (Figura 2) se puede observar la estructura de una CNN, la cual será adaptada

específicamente para este proyecto, con un diseño que incluye un número adecuado de capas y características para el análisis y clasificación de los granos. El proceso abarcará desde la recolección de datos hasta la mejora continua del modelo, buscando optimizar tanto la precisión como la adaptabilidad del sistema para lograr resultados más eficientes en la clasificación de granos de café.

Figura 2

Estructura de una CNN



Nota: esta imagen se obtuvo de [4].

Figura 3

Proceso de desarrollo del proyecto



Metodología

El primer objetivo del proyecto es establecer un marco teórico robusto que integre los conceptos fundamentales relacionados con la calidad de los granos de café, la visión por computador, las redes neuronales convolucionales y el análisis del modelo propuesto. Para lograrlo, se realizará una investigación exhaustiva que incluirá fuentes primarias y secundarias, como manuales técnicos, artículos científicos y datos recopilados directamente de pequeños productores. Este marco teórico proporcionará una comprensión integral de los factores que influyen en la calidad del café, así como un examen de las técnicas de visión por computador y redes neuronales aplicadas a su clasificación.

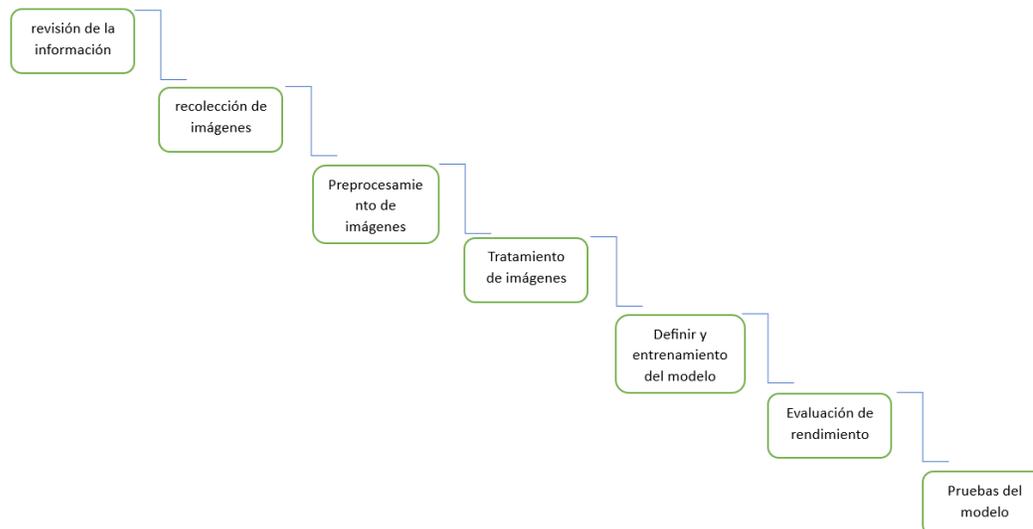
El segundo objetivo se centra en la recopilación de imágenes representativas de los granos de café, seguido de su anotación y preprocesamiento. Además, se planificará la división del conjunto de datos en diferentes grupos para entrenamiento, validación y prueba, asegurando así una adecuada preparación para el desarrollo del modelo.

Para el cumplimiento del tercer objetivo el cual tiene como meta elegir las mejores herramientas para desarrollar el proyecto, donde se utilizará un proceso sistemático para garantizar el cumplimiento de todas sus etapas, basándose en el alcance establecido anteriormente. Como se muestra en la (figura 4).

El cuarto objetivo es comprobar los resultados cuantitativos (Precisión, Acurracy, Recall, f1-score, matriz de confusión), para asegurarse de que son correctos, las cuales se utilizarán tanto durante el entrenamiento como en las pruebas finales, lo que permitirá evaluar exhaustivamente el funcionamiento del modelo y su análisis para su respectiva discusión de resultados.

Figura 4

Proceso de desarrollo



Justificación

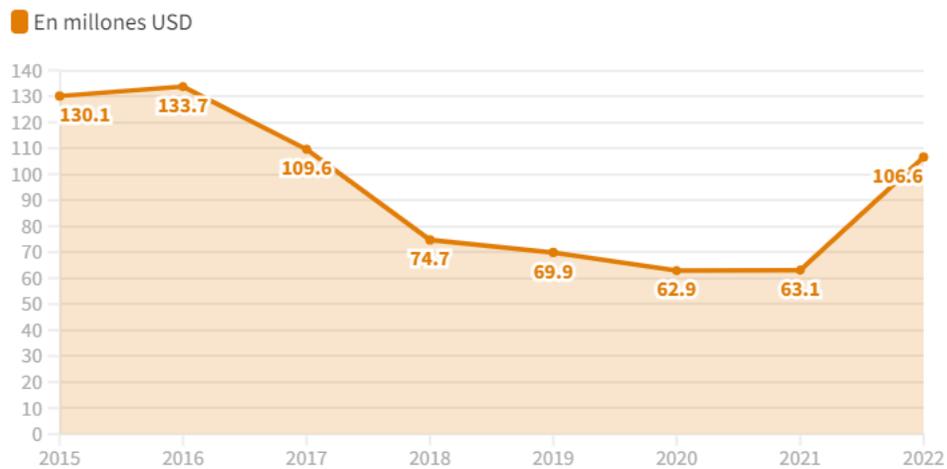
El trabajo planteado se incorporará a una solución en un futuro estaría en concordancia con el objetivo de desarrollo sostenible (ODS) número 9: “Construir infraestructuras resilientes, promover la industrialización inclusiva y sostenible y fomentar la información”[5]. El objetivo 3 del proyecto busca fortalecer un proceso establecido, promoviendo la tecnificación en el área agrícola y alineándose con el plan de creación de oportunidades del Ecuador el cual se colinda con dicho objetivo [6].

Uno de los productos más consumidos en todo el mundo es el café y sus derivados según [6], y contar con certificaciones. En Ecuador, el café representa una parte significativa de las exportaciones como lo dice en el informe del banco central del Ecuador (Figura 5) Es por eso por lo que el proyecto se justifica debido al enfoque en el área de la comunidad productora de café muestra que el café ocupa un lugar importante dentro del ámbito cafetalero.

Figura 5

Exportaciones de café en Ecuador

De enero a noviembre de cada año



Nota: esta imagen se obtuvo de [7].

Los productores como las asociaciones siempre buscan cumplir con las regulaciones y normas establecidas en el ámbito cafetalero, lo que refuerza la necesidad de este proyecto. Se debe recalcar que, ya que el café es uno de los cultivos más producidos en el país, este hecho respalda la recolección de los datos de forma más fácil.

El modelo creado en este trabajo puede ser visto como un componente de una futura solución para apoyar esta actividad en las asociaciones de café como ACCRI e incluso para los productores individuales, a pesar de que no aborda directamente el fortalecimiento del proceso de postcosecha.

Justificación Tecnológica

Aunque el modelo no ofrece una solución directa para abordar el problema de la tecnificación, su implementación futura podría tener un impacto significativo tanto en términos económicos como sociales. El uso del modelo podría resultar en ahorro de tiempo, lo que permitiría a los productores optimizar sus recursos y a las asociaciones

acelerar el proceso. Esto podría generar beneficios económicos al hacer que la producción sea más eficiente y facilitar la incorporación de más socios a las asociaciones. Además, el ahorro de tiempo puede tener un efecto positivo en la sociedad, ya que libera recursos para otras actividades importantes que pueden realizar los miembros de la comunidad cafetera.

Justificación Agrícola

El objetivo de mejorar el proceso es ayudar a la industria cafetera, especialmente en Ecuador, donde la agricultura es muy importante, permitiéndole mantenerse al día con la demanda en constante aumento en todo el mundo.

CAPÍTULO 1

1. Marco Teórico

1.1. Calidad de los granos de café

1.1.1. Características generales de los granos de café

La agricultura es una de las ocupaciones más importantes para el ser humano, ya que no sólo proporciona alimentos a la población, sino que también suministra materias primas a los fabricantes de productos agrícolas [8]. el café tiene un valor nutricional donde estimula el sistema nervioso y tiene la capacidad de mejorar la concentración mental; El árbol del café es procedente del norte de Etiopia, en oriente de África; En América Latina no existen gran variedad por lo cual lleva a que dos tipos de granos sean populares entre los productores los cuales son café robusto (*Coffea canephora P.*) y arábica (*Coffea arabica L.*). La producción de café es una de las actividades agrícolas más importantes de Ecuador. Este producto básico ha sido clasificado entre los diez cultivos con mayor superficie cosechada, y se ha producido en más de 19 provincias de Ecuador en el transcurso de los últimos 15 años [9].

El café se clasifica como planta arbustiva, y puede alcanzar una altura de hasta nueve metros en algunos casos. Las hojas del café varían de color, del verde al amarillo, según el tipo de café de que se trate. El café se estudia en diversos campos, como la agronomía, la genética, la cultura, la economía y el turismo, de ahí que su calidad sea tan importante [10].

Según [11], menciona que además de una forma correcta de cultivar existen varios aspectos a tomar en cuenta:

- **Clima:** Un ambiente cálido y húmedo, con temperaturas que oscilan entre los 17 y los 23 grados celsius, es bueno para el cultivo de la vegetación cafetera. Este tipo de clima es típico de regiones subtropicales o tropicales, es crucial para los granos Arábica y Robusta, que requieren diferentes altitudes. Estas condiciones climáticas estables son esenciales para prevenir enfermedades y asegurar la calidad del grano de café, características comunes en las principales regiones cafeteras [12].
- **Altitud:** el cultivo de café tiene cierta dificultad para mantenerse en una producción óptima, las áreas por encima de los 1800 m. s. n. m. es atractiva para la producción de café robusta, al contrario, el café arábico en su mayoría es óptima a una altura de 900 m. s. n. m. la cual es una gran diferencia entre los tipos de café más producidos en Latinoamérica [13].
- **Precipitación:** La fuente más importante de agua para cultivos como el café es la precipitación; sin embargo, las variaciones en la distribución de las precipitaciones como resultado de la variabilidad climática podrían tener un efecto sobre su crecimiento. Esto se debe a que se ha demostrado que la producción de café es sensible a los cambios en los patrones climáticos [13].

Un punto para recalcar es la ubicación donde se siembra dicho cultivo, existe el cinturón del café (Figura 6), el cual se refiere a una zona geográfica alrededor del mundo

donde las condiciones climáticas son óptimas para el cultivo del café, es decir zonas cálidas húmedas

Figura 6

Cinturón del café



Nota: tomado de [14].

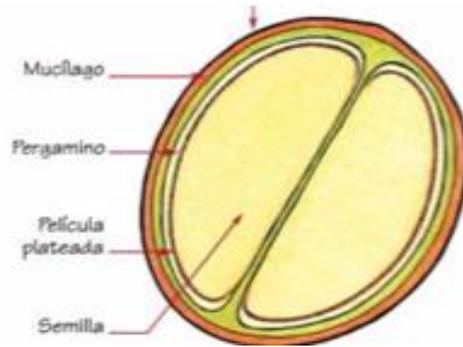
1.1.2. Características morfológicas en los granos de café

El fruto del café se describe como una baya parecida a la drupa que está formada por numerosas capas, según el Instituto Interamericano de Cooperación para la Agricultura [15]. Estas incluyen la pulpa, la epidermis o exocarpio, y el mucílago o mesocarpio, que es un material dulce que cubre los dos granos que se colocan uno frente al otro y revela sus lados planos. Otras dos capas de protección rodean la semilla, que es la fuente de la bebida de café. Estas capas consisten en una película plateada llamada espermoderma y una segunda capa que se encuentra en el exterior de la semilla y se denomina pergamino o endocarpio.

Estas partes están explícitas en detalle (Figura 7).

Figura 7

Partes del grano de café



Nota: Manual de producción sostenible de café en la República Dominicana [15].

La semilla, que es un fruto seco de forma oblonga, plano convexo y tamaño variado, es la que se somete al proceso de tostado (Figura 8) y posterior molienda, para crear finalmente la bebida que consume el ser humano.

Figura 8

Proceso manual de tostado del producto



Nota: tomado de [3].

En términos generales, existen dos categorías principales de técnicas de recolección: procesos selectivos y no selectivos. Las técnicas que pertenecen a la primera categoría son las que se realizan manualmente. Estos métodos se utilizan para cosechar únicamente los granos de color rojo o amarillo, que son los granos que están

listos para ser cosechados. Para la realización de este proceso en la cosecha del café se puede utilizar diferentes herramientas (Figura 9), la cual es tanto para la recolección y almacenamiento entre fases. El Principal inconveniente de este proceso en plantaciones grandes es la falta de tiempo o de mano de obra suficiente para la tarea a realizar, se equilibra con el hecho de que no debe exigirse una separación posterior.

Figura 9

Herramienta utilizada para la recolección manual de café



Los métodos manuales que entran en esta categoría incluyen los procedimientos de “ordeño” total o parcial, en los que se ignora la madurez del grano y se recoge todo el fruto de las ramas. Los métodos no selectivos se utilizan principalmente en la industria a gran escala, que es donde se utilizan máquinas especializadas. Sin embargo, también hay métodos manuales que entran en esta categoría.

En la investigación de [16] donde fue comparado el método tradicional selectivo con el “ordeño” parcial y total, el cual arrojó los siguientes resultados:

En tres explotaciones se midieron los frutos maduros recolectados utilizando las tres formas corrientes que se han comentado anteriormente. Los resultados obtenidos se midieron en porcentajes, y fueron los siguientes: El 97,89% se adquirieron utilizando

el método clásico seleccionado, el 88,05% se lograron utilizando el ordeño parcial, y el 84,52% se obtuvieron en conjunto. Además, para determinar el nivel de producción, se determinó que el promedio de kilogramos recogidos en una hora era de 10, 26,7 y 24,7, respectivamente.

Los resultados fueron similares a los que obtuvieron en

Los resultados coinciden con los comunicados por [17] en su investigación titulada “Calidad y eficiencia de dos métodos no selectivos de recolección manual del café *Coffea arábica*”, en la que indicaron que, si bien la eficiencia mejora con los métodos no selectivos, la calidad de la cosecha disminuye. Además, es crucial reconocer que las técnicas de recolección no selectiva pueden resultar improductivas si no se ejecutan en el momento adecuado, ya que necesitan un nivel de madurez específico en las ramas de las que se realiza la recolección, junto con una concentración adecuada en las mismas.

Luego de la recolección, se llevan a cabo distintas actividades destinadas a asegurar la calidad del grano de café. Este conjunto de procedimientos se lo denomina proceso de postcosecha, y según [18] donde incluyen las siguientes etapas:

- Separación de cerezas de baja densidad (vanos): Los frutos recolectados se colocan en recipientes los cuales contienen agua, ya puede ser piscinas o tinas. Aquellos con tengan menor densidad, se conocen como canos o vacíos, estos flotan en la superficie, este proceso facilita su identificación y ayuda a la separación del producto. Este proceso se puede visualizar (Figura 10).

Figura 10

Proceso de recolección y separación



- Diferenciación entre frutos verdes y pintones: Es necesario llevar a cabo la separación de frutos inmaduros (verdes), ya que estos afectan a la calidad del producto final. Según el método de recolección que se utilizó, será el tiempo que implicará esta actividad, ya que incluso en cosechas selectivas pueden incluirse algunos frutos que no han alcanzado su estado de madurez (Figura 11).

Figura 11

Diferenciación entre frutos verdes y pintones



Nota: esta ilustración se obtuvo de [19]

- Clasificación según su tamaño: El resultado de esta acción es proporcionar una mayor coherencia en los procedimientos futuros. Esta operación puede realizarse al mismo tiempo que el despulpado mediante tamices, o puede preceder a esta fase, según las necesidades del productor. En esta etapa se emplean técnicas de separación (Figura 12).

Figura 12

Herramienta para separación de tamaño



- **Despulpado:** en esta fase se extraen las semillas del fruto, las cuales posteriormente se convertirán en el producto consumible el cual es el café. Este paso puede llevarse de forma manual o con una maquina despulpadora. Ocasionalmente, se emplea agua para reducir la cantidad de granos los cuales no fueron despulpados o que hayan sido dañados por la máquina, aunque esto puede llevar a dificultades en la gestión de los residuos de la pulpa. En la (Figura 13) podemos ver la maquinaria necesaria para este procedimiento.

Figura 13

Maquina para despulpar el grano de café



- **Fermentación:** Este procedimiento es esencial para la producción de cafés especiales, debido a que la duración de los lavados realizados durante el proceso influye en las características finales del grano. Este paso permite el ajustar el perfil del grano de café para cumplir estándares establecidos o diferenciarlo de otro tipo de grano (Figura 14).

Figura 14

Proceso de fermentación del café



Nota: Esta imagen se obtuvo de [19].

- Secado: En función del tipo de grano, es esencial mantener el nivel de humedad entre el 10 y el 12%. Una humedad excesiva conlleva propiciar el crecimiento de moho o enfermedades las cuales lo pueden afectar, mientras que los niveles muy bajos son los responsables de que el pergamino se pueda dañar las semilla. En procesos manuales en donde no se controlen ambientales, es requerido un monitoreo constante. Es por eso por lo que en ocasiones se usan invernaderos de secado (Figura 15).

Figura 15

Invernadero para el secado del café



- Selección final: Durante el proceso posterior a la cosecha, es inevitable que algunos granos de café no cumplan los requisitos de calidad establecidos,

o que los residuos de las operaciones anteriores se mezclen con los granos de café. Estos defectos, como la pulpa seca, pergamino desprendido, suelen ser eliminados utilizando técnicas como el “aventado” el cual aprovecha el viento para la separación de los residuos. Sin embargo, algunos defectos deben ser corregidos manualmente, tarea donde se aborda con mayor detalle en la investigación establecida. Después son almacenados de forma que puedan mantener sus características para su correcto uso (Figura 16).

Figura 16

Almacenamiento del grano de café



Sin embargo, la atención que se preste a cada uno de los procesos, así como la dedicación que se dé a cada uno de ellos dependerá del mercado al que se dirija, ya sea un mercado nacional, un mercado internacional, o la venta a un intermediario o a una microempresa. Esto es algo que debemos tener en cuenta porque ninguno de los

procesos que se han descrito son necesarios para la comercialización. Esto se debe a que este producto también se comercializa en diferentes etapas de cultivo, al igual que la café cereza.

1.1.3. Estándares de selección y calidad del grano de café

Las especialidades de café, también conocidas como cafés de alta calidad, desempeñan un papel importante en el mercado, donde a menudo pueden encontrarse a precios muy elevados. Para su comercialización, es necesario contar con certificaciones específicas las cuales varían según el destino final de los granos de café como lo menciona en [20]. Por ejemplo, la certificación “JAS”, que concede la certificadora alemana BCS Öko-Garantie, la “Certificación Orgánica Ecuador” o la certificación “USDA organic” son algunas de las más significativas. Estas certificaciones examinan diversas variables asociadas al cultivo. Algunas de las características más importantes son las condiciones del suelo, el clima y los fertilizantes que se utilizan, así como el proceso posterior a la cosecha y la calidad del grano que se adquiere.

Cuando se trata del grano, se realiza un proceso de selección riguroso para cumplir con los estándares de certificación. Aunque los criterios específicos pueden variar entre las diferentes certificaciones y los diferentes tipos de café, existen defectos generales que deben ser identificados y gestionados. A continuación, se detallan los principales defectos y sus posibles causas:

- Granos cortados o dañados: Durante el proceso del despulpado, algunos granos pueden ser cortados o pueden ser dañados por la máquina, lo que

afecta su forma y el pergamino. Estos defectos se diferencian por las partes faltantes del grano (Figura 17).

Figura 17

Grano cortado o dañado



- Restos de pulpa o cáscara adheridos: Ocasionalmente, la pulpa o cáscara no se eliminan por completo durante el proceso de despulpado y se adhieren al grano. Esto se debe solucionar antes que el grano se seque, ya que una vez seco, la remoción se hace más difícil, pudiendo generar problemas de humedad y obligando a descartar los granos afectados (Figura 18).

Figura 18

Grano con restos de pulpa o cáscara adheridos



- Daños por humedad: La acumulación por humedad, ya sea por un secado inadecuado o por la presencia de objetos extraños, puede provocar manchas visibles en el pergamino, lo que indica un grano dañado (Figura 19).

Figura 19

Grano dañado por humedad



- Afectado por insectos: Algunos insectos, como la broca del café o (*Hypothenemus hampei*) atacan los granos durante su maduración, causando así daños significativos. Uno de los tipos de insectos que más daños causan en los cafetales es la mariposa. (Figura 20).

Figura 20

Grano afectado por insectos



- Malformaciones: En el curso del crecimiento del grano, es posible que éste adquiera anomalías en su forma, como granos huecos o caracoles, además de malformaciones adicionales que aún no han sido reconocidas (Figura 21).

Figura 21

Grano con malformaciones



- Presencia de hongos: Los granos de café son sensibles a varias enfermedades, una de las cuales es la roya, causada por el hongo *Hemileia vastatrix*. Esta enfermedad tiene un impacto considerable en la calidad de los granos (Figura 22).

Figura 22

Grano con presencia de hongos



- Rotura del pergamino: Cambios drásticos en la temperatura durante el secado son los causantes de la rotura de pergamino, afectando la estructura del grano. (Figura 23)

Figura 23

Grano con rotura de pergamino



Es importante saber que un mismo grano puede presentar varios de estos defectos al mismo tiempo, Por consiguiente, cada vez es más necesario un estricto control de calidad en todo el proceso.

1.2. Visión por computador

1.2.1. Definición

La visión por computador, comúnmente denominada visión artificial, es una rama multidisciplinar de la informática que tiene como objetivo capacitar a los ordenadores para interpretar y comprender el mundo visual de forma análoga a como el ojo humano percibe la realidad de su entorno. Se maneja en su obra titulada “Conceptos y métodos en visión por computador” [21]. Por esta noción de la definición de lo que es la visión por computador. Este tema abarca la creación de algoritmos y sistemas capaces de

capturar, procesar y analizar las fotos y vídeos asociados, extrayendo así la información más relevante para la toma de decisiones o la realización de acciones. El objetivo es garantizar que los dispositivos no solo vean las imágenes, sino que también las comprendan y puedan interactuar con ellas, tal como sucede en las tareas de reconocimiento de objetos, segmentación, clasificación o seguimiento de movimientos, como lo menciona [22].

Esto se consigue mediante el uso de sofisticados métodos de procesamiento de señales, geometría, estadística y aprendizaje automático, como las redes neuronales convolucionales (CNN), que utilizan los sistemas de visión por ordenador para detectar patrones de imágenes intrincados, estas tecnologías son aplicadas en números campos, des la medicina, donde ayudan a detectar enfermedades a partir de imágenes médicas, hasta la robótica. Además, se utilizan en la seguridad, la automoción y la manufactura, mejorando así la automatización, precisión y eficiencia de las tareas en diversas industrias. Así, la visión por computador ha revolucionado la manera en que interactuaba con el mundo digital y físico [23].

1.2.2. Proceso

El libro “Conceptos y métodos en visión por computador” [21] organizó este proceso en cinco pasos clave, que son los siguientes: la adquisición de imágenes, el preprocesamiento de imágenes, la extracción de características, la identificación o creación de conocimiento y la toma de decisiones son ejemplos de técnicas de procesamiento de imágenes.

- Adquisición de las imágenes: Durante esta fase, se recopilan datos visuales en tiempo real, como vídeos, o se utilizan conjuntos de imágenes relevantes. Si no había buenas imágenes, había que crearlas utilizando fotos, ecografías o resonancias magnéticas.
- Preprocesamiento de imágenes: este punto trata de usar ciertos algoritmos o cambios visuales para el cumplimiento del objetivo principal, el cual es mejorar la calidad de la imagen eliminando cualquier defecto o ruido que pueda hacer que el sistema interprete los datos de forma incorrecta.
- Captura de características distintivas: En este paso se identifican atributos relevantes de las imágenes (bordes, líneas, patrones, texturas). Estas características ayudaron a resaltar áreas de interés y, en algunos casos, esta etapa fue suficiente para cumplir con los objetivos del sistema.
- Identificación y generación de conocimiento: En este punto para el reconocimiento de elementos complejos (Estructuras, objetos, patrones específicos), se utilizan técnicas avanzadas como la inteligencia artificial para procesar las características extraídas, de modo que puedan
- Toma de acciones: Finalmente, los resultados obtenidos guiaron acciones concretas que podían incluir desde emitir alertas o reproducir sonidos hasta realizar tareas más complejas como asistencia en conducción, intervenciones quirúrgicas o reconstrucción de objetos.

Este enfoque estructurado destaca el carácter interdisciplinario del campo y su capacidad para integrar múltiples tecnologías con el fin de resolver problemas específicos.

1.2.3. Algoritmos

A continuación, se describen los métodos y técnicas que se utilizan en el campo de la visión por ordenador. Se utilizan sobre todo en las fases de preprocesamiento y extracción de características del proceso.

1.2.3.1. Cambios en el espacio de color

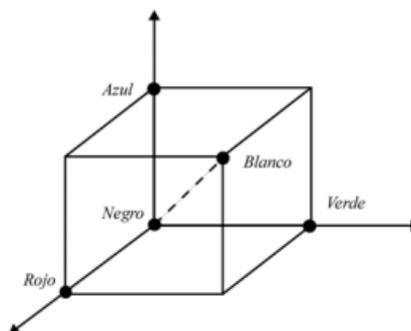
Además, el espacio de color es el enfoque de representación que permite la visualización de los colores en una variedad de dimensiones. La aplicación que se creará, que en este caso será en imágenes de granos de café, dependerá del tipo de captura de imagen que se esté realizando; Pueden ser necesarios realizar conversiones entre espacios de color o trabajar con algunos canales específicos para reducir el ruido de las imágenes [24].

- Modelos RGB: Además de la síntesis aditiva, el modelo triestímulo sirve de base para el modelo de color RGB, muy utilizado. Los colores se representan como cantidades de rojo, verde y azul en un espacio tridimensional que tiene forma de cubo. Este método depende del dispositivo que se utilice. En la mayoría de los casos, las coordenadas del modelo están en el rango $[0,1]$, pero en aplicaciones que tratan con multimedia y procesamiento de imágenes, se utilizan valores discretos en

el rango [0,255]. En RGB, las coordenadas que unen el punto negro con el punto blanco corresponden al píxel de escala de grises. Este modelo concreto se compone de tres planos separados, uno para cada uno de los colores básicos. Estos planos, cuando se fusionan en una pantalla RGB, dan lugar a la producción de una imagen en color. El procesamiento de imágenes es un uso especialmente beneficioso de esta técnica, sobre todo cuando las imágenes se representan en términos de los tres planos. Además, muchas cámaras digitales adquieren imágenes en formato RGB (Figura 24), lo que refuerza su relevancia en aplicaciones prácticas [25].

Figura 24

Representación del modelo de color RGB en el espacio espacial.



Nota: Procesamiento morfológico de imágenes en color: aplicación a la reconstrucción geodésica [25].

Es posible ver la imagen en color de "Lenna" con ayuda de la (Figura 25). Se puede observar cómo el canal R (rojo) muestra valores grandes, lo que indica que la imagen original tiene un color razonablemente rojizo. Esto es posible gracias a la división de la imagen en canales RGB, como se

muestra en la Figura ##. Los histogramas de los canales R, G y B se muestran en la (Figura 26), que puede consultar aquí. Estos histogramas muestran el porcentaje de uso de cada valor numérico en los canales R, G y B, proporcionando información cuantitativa sobre la distribución del color en la imagen.

Figura 25

Imagen de "Lenna" en color



Nota: Ilustración de la imagen "Lenna" obtenida de [25].

Figura 26

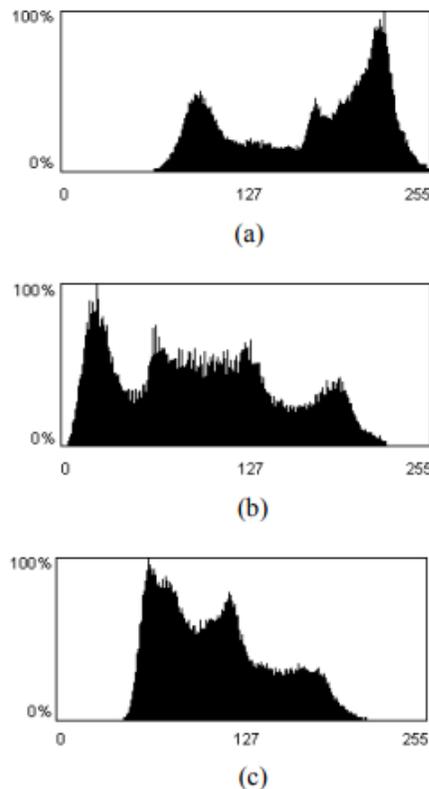
Representación en canales R (a), G (b) y B (c) de 'Lenna'.



Nota: Ilustración de la imagen “Lenna” obtenida de [25].

Figura 27

Histogramas de la imagen cromática ‘Lenna’ de los mapas rojo (a), verde (b) y azul (c).



Nota: Ilustración de la imagen “Lenna” obtenida de: Procesamiento morfológico de imágenes en color: aplicación a la reconstrucción geodésica en [25].

Varios factores contribuyen al atractivo del sistema RGB, entre ellos su capacidad para transmitir la tricromaticidad y su importancia en el proceso de captura y presentación de imágenes, sobre todo en el contexto de aplicaciones de visión artificial como la fotografía por satélite. El examen de los distintos planos de estas fotos permite identificar características y detalles importantes. Estas imágenes están formadas por numerosas

fotografías espectrales del mismo sujeto. Sin embargo, el modelo RGB presenta limitaciones en ciertos tipos de procesamiento, como en técnicas de realce de imágenes que requieren ecualizar la intensidad luminosa, ya que modificar el histograma en cada plano puede alterar la composición espectral. En tales casos, un modelo cromático que incluya una componente específica para la intensidad resulta más adecuado.

Las desventajas del modelo RGB son:

- ❖ Falta de uniformidad: Este modelo no permite medir la distancia entre colores usando los componentes R, G y B.
- ❖ Poca intuitividad: Puede resultar difícil describir un color haciendo referencia a los números y porcentajes de los colores principales que componen esa tonalidad.
- Modelo CMY: El cian, el amarillo y el magenta son los colores básicos en la síntesis sustractiva del color, mientras que el cian, el amarillo y el magenta son los colores secundarios desde la perspectiva de la síntesis aditiva de la luz. Los modelos CMY y CMYK, este último que incluye una componente de tinta negra, son ampliamente utilizados en sistemas de impresión, aunque su aplicación en visión por computador es limitada. La transformación del espacio RGB al modelo CMY se puede realizar mediante la ecuación:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

En ambos modelos, las componentes se encuentran en el rango [0,1]. Esta transformación ilustra que una superficie de color amarillo puro (Y=1, C=0, M=0) no refleja luz azul, lo que implica que B=0, G=1 y R=1. Un pigmento magenta puro no refleja la luz verde, mientras que un pigmento cian no refleja la luz roja. Ambos pigmentos son similares en este sentido [25].

- **Modelo XYZ** La CIE estableció el espacio de color XYZ para resolver las limitaciones de los triestímulos espectrales R, G y B. Esto se hizo así porque algunos colores necesitaban niveles de estímulo negativos para poder ser representados. El espacio de color XYZ está formado por una transformación lineal del sistema RGB. Los colores primarios XYZ de este modelo son ficticios por naturaleza y no se correlacionan con luces físicas reales. Estos colores se crean transformando los primarios RGB y utilizando funciones de coincidencia de color $x(\lambda)$, $y(\lambda)$ y $z(\lambda)$. Ambos métodos se utilizan para obtener estos colores. Este espacio permite una representación más completa y precisa del color en comparación con los modelos anteriores, como muestra el desarrollo de los nuevos valores triestímulos para varios tonos del espectro, que se ve en la (Figura 24) [25].

La matriz que permite realizar la transformación lineal del sistema RGB al espacio de color XYZ, considerando valores dentro del rango [0,1].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4125 & 0.3576 & 0.1804 \\ 0.2127 & 0.7156 & 0.0722 \\ 0.0193 & 0.1192 & 0.9502 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

La transformación inversa de (2.23) se puede definir como:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.2405 & -1.5372 & -0.4985 \\ -0.9693 & 1.8760 & 0.0416 \\ 0.0556 & 0.2040 & 1.0573 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

El modelo cromático XYZ es un sistema independiente del dispositivo que se desarrolló para aplicaciones en las que la representación del color no depende del equipo o hardware que se esté utilizando. Un ejemplo de este tipo de aplicación procede de los sistemas industriales de medición del color. A efectos de los cálculos colorimétricos, este modelo sirve como componente vital de la definición normalizada de las características del color, permitiendo una comparación precisa y consistente de colores entre diferentes dispositivos. Para asegurando que los colores se representen de manera uniforme independientemente del medio de visualización o impresión, la Comisión Internacional de la Iluminación (CIE) desarrollado, el modelo XYZ la cual proporciona una base común que facilita la conversión entre distintos espacios de color.

1.3. Redes neuronales convolucionales

1.3.1. Redes neuronales

El funcionamiento del cerebro humano sirve de inspiración para el modelo matemático conocido como red neuronal artificial. Este modelo se desarrolló para analizar información y encontrar soluciones a problemas mediante el uso de un sistema de aprendizaje. Estas estructuras de red se construyen a partir de una gran cantidad de componentes sencillos, conocidos como neuronas artificiales, que se organizan en

niveles o capas interconectadas. Su operación imita la manera en que las neuronas biológicas procesan señales y generan respuestas [26].

las características más destacadas en la investigación “Redes neuronales: Conceptos básicos y aplicaciones” [27] son:

- Estructura jerárquica: Las neuronas se organizan en capas (de entrada, ocultas y de salida), permitiendo el procesamiento de información en paralelo.
- Adaptabilidad: Trata sobre las conexiones entre neuronas, llamadas pesos, se ajustan a través de procesos de aprendizaje (aprendizaje supervisado).
- Resolución de problemas complejos: Las redes neuronales se centran la resolución de problemas como encontrar patrones, clasificar datos y predecir tendencias en situaciones no lineales.

Debido a su funcionamiento, las redes neuronales pueden utilizarse en muchos ámbitos diferentes, como la visión artificial, el procesamiento del lenguaje natural y la predicción del futuro del dinero.

Hay tres tipos principales de capas en las redes neuronales: de entrada, de salida y ocultas (las más esenciales).

- Capa de entrada: También se llama capa sensorial y está hecha de neuronas que reciben información o señales del entorno que las rodea. Estas neuronas son el primer punto de procesamiento.

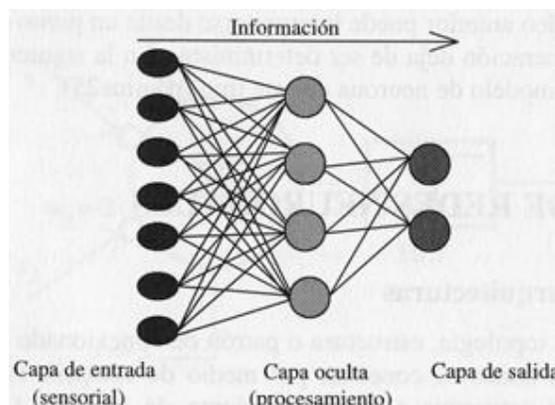
- Capa de salida: Las neuronas que componen esta capa son las responsables de producir las respuestas de la red neuronal, que a su vez proporciona los resultados del procesamiento.
- Capas ocultas: Entre las funciones más importantes es dar a la red más grados de libertad. Estas capas no están directamente relacionadas con el entorno, no interactúan directamente con los sensores ni con los efectores, ya que no están conectadas directamente con el entorno.

Las capas ocultas son importantes porque la red neuronal en esta etapa son las encargadas de encontrar patrones difíciles y sacar características importantes durante el entrenamiento y la operación.

Para entender de mejor manera como funciona una red Neuronal se presenta el siguiente esquema (Figura 28).

Figura 28

Funcionamiento de una CNN



Nota: esta ilustración se obtuvo de [28] en su trabajo sobre las redes neuronales

En términos de su estructura, las redes pueden dividirse en dos categorías: redes de una sola capa, que están formadas por una sola capa de neuronas, y redes de varias capas, que están formadas por muchas capas a través de las cuales se organizan las neuronas. Se puede establecer una diferencia entre las redes unidireccionales, también conocidas como redes (feedforward), y las redes recurrentes o de retroalimentación, dependiendo del flujo de datos. En las redes unidireccionales, la información solo viaja en una dirección. Por otro lado, en las redes recurrentes o de retroalimentación (feedback), la información puede circular entre las distintas capas de neuronas en cualquier dirección, incluso de la salida a la entrada [28].

1.3.1.1. Modelos existentes

Las redes neuronales se pueden clasificar de muchas maneras diferentes, y una de las diferencias más significativas se basa en las técnicas de aprendizaje que se utilizan: supervisadas y no supervisadas. Las tres de las topologías o modelos más destacados son: La red neuronal convolucional (CNN), la red neuronal recurrente (RNN) y la red de función de base radial (RBF) son [29]:

- CNN: Estas redes se basan en el perceptrón multicapa y se utilizan principalmente en algoritmos para la percepción visual. Su estructura tiene capas de agrupación y convolución. Estas capas crean conjuntos de matrices realizando la misma operación una y otra vez a intervalos regulares [30].
- RNN: Estas redes se caracterizan por establecer correspondencias entre secuencias de entrada y salida. Pueden ser totalmente recurrentes, con

conexiones neuronales dinámicas, o parcialmente recurrentes, donde las conexiones son estáticas [31].

- RBF: referidos por sus funciones de base radial, que utilizan, son especialmente efectivas en aplicaciones en tiempo real debido a su estructura simple, que consta de una única capa oculta [32].

Cada una de estas topologías está diseñada para resolver un problema concreto utilizando sus características únicas en diferentes áreas.

1.3.1.2. Funciones de Activación

La función de activación es un componente esencial de las redes neuronales, ya que su objetivo principal es restringir la variedad de salidas que puede producir cada neurona. Dependiendo de la naturaleza del problema que se deba abordar, puede ser lineal o no lineal. La selección de la función viene determinada por los criterios establecidos por el investigador y está influenciada por diversos elementos, entre los que se incluyen la naturaleza del problema, la precisión deseada, la velocidad del modelo y el algoritmo de aprendizaje que se está utilizando. Por ejemplo, funciones no lineales como AdamW son preferidas en muchas aplicaciones por su eficiencia computacional y capacidad para modelar relaciones complejas. Además, con el fin de mejorar el rendimiento del modelo y protegerlo contra problemas como la desaparición del gradiente, se recomienda ajustar los datos para que se encuentren dentro del rango de la función de activación que se haya elegido [30].

1.3.2. CNN

El análisis de imágenes hace un uso extensivo de métodos de aprendizaje profundo, en particular de redes neuronales convolucionales, debido a su capacidad para llevar a cabo la extracción automática de características (también conocida como “extracción de rasgos”). A diferencia de los métodos de aprendizaje automático, que requieren que el analista de datos predefina las características, este enfoque no lo requiere [33]. Debido a esta ventaja, las redes neuronales convolucionales han experimentado un gran aumento de su importancia relativa en los últimos años. Además, el desarrollo de una tecnología más potente ha permitido la aceptación y el uso de estos algoritmos, lo que ha facilitado aún más su empleo [34].

1.3.2.1. Componentes principales de una CNN

En el estudio de [35] nos menciona que las capas importantes que componen las redes neuronales convolucionales (CNN). Son las siguientes:

- **Capas Convolucionales (Convolutional Layers):** En las redes neuronales convolucionales, las capas convolucionales cumplen un papel esencial, ya que son las que se encargan de aplicar filtros (kernels), a las entradas con el fin de extraer determinadas características las cuales son bordes texturas y patrones que pueden llegar a ser más complejos.
- **Capas de Pooling (Pooling Layers):** Mediante el uso de capas de agrupación, se reduce la dimensionalidad de las características recuperadas, pero se conservan las más significativas. Este proceso no solo disminuye la cantidad de datos que deben procesarse, lo que ayuda a

reducir el costo computacional, sino que también contribuye a prevenir el sobreajuste al simplificar la representación de los datos.

- Capas de Activación (Activation Layers): En esta capa la red es capaz de adquirir la capacidad de comprender conexiones complejas dentro de los datos mediante la inclusión no linealidades en el modelo, las capas de activación utilizan funciones no lineales, como la función Unidad Lineal Rectificada (ReLU). Gracias a esto permite reconocer patrones que no podrían representarse mediante funciones lineales simples.
- Capas Completamente Conectadas (Fully Connected Layers): Estas capas además de que están conectadas a todas las neuronas de la capa inferior, también son esenciales para interpretar la información que se ha recuperado y producir el resultado que se debe lograr. Esto sucede durante las fases finales de la red, las capas totalmente conectadas toman las características aprendidas y las combinan para llegar a juicios finales, como clasificaciones.
- Capas de Normalización (Normalization Layers): Estas capas cumplen la función de reducir la influencia de las variaciones en los datos al modificar los valores de las características a un rango estándar, su principal contribución se resume en la estabilización y aceleración del proceso de entrenamiento.
- Capas Softmax o Sigmoid: Las capas de salida utilizan las funciones Softmax y Sigmoid. La función Softmax se utiliza más en problemas

multiclase, mientras que la función Sigmoid es comúnmente utilizada en problemas de clasificación binaria (0 y 1). En estas capas se centran más en las tareas de clasificación.

1.3.3. Aplicaciones de las CNN

Numerosas aplicaciones han demostrado que estas estructuras son muy eficaces en sus respectivos campos gracias a su capacidad para extraer de forma automática y eficaz características de imágenes y otros tipos de datos espaciales, las redes neuronales convolucionales (CNN) han revolucionado el campo del aprendizaje profundo y la visión artificial. Sus usos más relevantes son:

- Agricultura de precisión.
- Control de calidad industrial.
- Vehículos autónomos.
- Clasificación de imágenes.
- Detección y segmentación de objetos.
- Reconocimiento facial.

1.4. Trabajos relacionados

1.4.1. Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia artificial y Visión por computador

Mediante el uso de varias técnicas de visión artificial e inteligencia, el objetivo principal de la investigación era identificar defectos morfológicos en los granos de café. Pudimos lograr este objetivo utilizando las arquitecturas SSD Mobilenet V2 y SSD ResNet50. Se utilizaron redes neuronales convolucionales (CNN) en la construcción del modelo. Las fotografías se tomaron en un entorno controlado con un fondo blanco, una iluminación constante y una cámara situada a una distancia precisa para capturar un solo grano en cada una de las fotografías. Mediante la aplicación de la ampliación de datos, se transformaron 3276 fotografías originales en 20 284 imágenes. A continuación, estas imágenes se separaron manualmente en dos categorías: granos excelentes y granos defectuosos. Los resultados mostraron que Mobilenet tenía una tasa de precisión del 76,92 %, mientras que ResNet solo alcanzaba una tasa de precisión del 52,01 %. Una puntuación f1 mínima de 0,27 indica que ResNet obtuvo unos resultados muy deficientes a la hora de localizar frijoles en buen estado. Según los resultados de la investigación, las CNN tienen el potencial de ser beneficiosas para la identificación de granos; sin embargo, existe una limitación importante: el sistema solo puede examinar un grano a la vez. Debido a esta circunstancia, no es factible automatizar los procedimientos a gran escala, lo que los hace menos eficientes para la industria del café, que requiere el procesamiento de muchos granos al mismo tiempo. Dicho esto, a pesar de que el estudio representa un avance significativo en la industria, aún no es una solución viable ni

escalable para los grupos agrícolas o los agricultores. No obstante, proporciona una base tecnológica sólida para futuras soluciones que serán más duraderas y adecuadas al entorno en el que se desarrolla la realidad [36].

1.4.2. Procesamiento de imágenes para la clasificación de granos de café por forma y color

Esta investigación centra su objetivo principal en mejorar los métodos manuales convencionales que utiliza la industria cafetalera mediante el uso de una red neuronal Convolutiva (CNN) basada en el modelo MobileNetV1 con el fin de clasificar los granos de café según su forma y color. MobileNetV1 se resume como una arquitectura diseñada para optimizar dispositivos con recursos limitados. Su principal distinción respecto a otras es por el uso de convoluciones separables en profundidad, la cual reduce de manera drástica el número de parámetros y el coste de cálculo sin sacrificar la calidad del modelo. El proceso de entrenamiento consistió en utilizar un dataset de 3000 fotos de granos de café individuales, con un tamaño de 350 x 350 píxeles. El dataset se dividió de la siguiente manera: 80 % para el entrenamiento (train), mientras que el 20 % para la validación (valid). Con el fin de mejorar la generalización del modelo, se aplicaron técnicas de Data argumentation y se utilizó el optimizador Adam para realizar ajustes en los pesos respectivos de la red. Los resultados obtenidos reflejan que MobileNetV1 fue capaz de alcanzar una precisión del 88,72 % y una validación del 81,25 %, lo que demuestra que es capaz de clasificar correctamente los granos de café. Sin embargo, La limitación central respecto a nuestro trabajo es que esta investigación se enfoque en imágenes individuales, lo que impide la detección simultánea de múltiples granos en una sola imagen [37].

1.4.3. Comparative analysis of YOLO models for green coffee bean detection and defect classification

Este estudio analizó y comparó varios modelos YOLO diferentes, como YOLOv3, YOLOv4, YOLOv5, YOLOv7, YOLOv8 y YOLOv8n personalizado. El objetivo de este estudio era automatizar y mejorar el control de calidad en la industria del café, así como clasificar los defectos que se encuentran en los granos de café verde. Se utilizaron 5044 fotos, que se clasificaron en grupos según el tipo de defecto que presentaban. Estos grupos incluían granos negros, granos rotos, granos descoloridos y granos agrios. El conjunto de entrenamiento, el conjunto de validación y el conjunto de prueba eran conjuntos de datos diferentes. Como parte del plan, se mejoró la arquitectura del modelo personalizado (por ejemplo, se cambió la columna vertebral), se optimizaron los hiperparámetros (tasa de aprendizaje, optimizador y refuerzo) y se utilizaron métodos de aumento de datos para que el modelo fuera más generalizable. El modelo YOLOv8n personalizado fue el que presentó mejores resultados a comparación de los otros presentando los siguientes resultados: mAP de 0,995, precisión de 0,977 y recuperación de 0,990. Estas cifras muestran gran competitividad al momento de la comparación de métodos similares. Por otro lado, se observaron limitaciones importantes como los largos tiempos de entrenamiento y un conjunto de datos limitado, estas podrían afectar a su capacidad para aplicarse a una gama más amplia de situaciones. Sugerimos que, como parte de nuestro trabajo futuro, diversifiquemos el conjunto de datos, busquemos formas más avanzadas de normalizar y etiquetar los datos, y combinemos datos de diferentes fuentes para crear un sistema de alerta temprana basado en el Internet de las cosas. Este sistema no solo ayudaría a detectar mejor los defectos, sino que también podría

utilizarse en otras áreas de la agricultura para hacer los procesos más eficientes y reducir los costes [38].

1.4.4. Implementing a deep learning model for defect classification in Thai Arabica green coffee beans

Queríamos crear un modelo de aprendizaje profundo que utilizara redes neuronales convolucionales para clasificar diecisiete tipos diferentes de defectos que podían encontrarse en los granos de café verde arábica cultivados en Tailandia. El primer conjunto de datos contenía 979 imágenes. métodos de data argumentation el número de imágenes aumentó a 6853. Después de analizar arquitecturas CNN avanzadas, MobileNetV3 destacó por su capacidad de generalización y su gran precisión tras el ajuste de hiperparámetros. El modelo alcanzó una precisión global del 88,63 %, ya que funcionó bien con datos que nunca había visto antes y obtuvo una precisión de validación del 90,19 %. Además, se añadió a una aplicación web que clasificaba los defectos en tiempo real, lo que hizo muy felices a los agricultores locales. El estudio encontró algunos problemas con el uso de un conjunto de datos que, aunque mejorado, no tenía suficiente variedad para mostrar con precisión cómo son las cosas en el mundo real. Para que el estudio sea más útil en la vida real, se sugiere que el conjunto de datos sea más diverso, que se utilicen métodos avanzados de aumento de datos y que se utilicen métodos de aprendizaje conjunto [39].

1.4.5. Classification of Coffea Liberica Quality Using Convolution Neural Networks (Slim-CNN, YOLOv5, and VGG-16)

El objetivo de la investigación posterior fue crear un sistema autónomo de identificación de defectos en los granos de café Liberica (Kapeng Barako) mediante el uso de métodos de procesamiento de imágenes con redes neuronales convolucionales (CNN), más concretamente los modelos Slim-CNN, YOLOv5 y VGG-16. El propósito fue clasificar los granos en tres categorías físicas: bueno, medio y malo. Dentro de la categoría de granos malos, se contemplaron subcategorías como normales, negros, rotos, infectados por insectos y descoloridos; sin embargo, el sistema no reconoce cada una de estas subcategorías de forma individual. Este enfoque busca mejorar la calidad del café exportado y reducir el margen de error humano en el proceso de clasificación. Se diseñó un prototipo con una Raspberry Pi 4, cámara Logitech C922, iluminación LED y una pantalla LCD, el cual fue utilizado para capturar imágenes en un entorno controlado. El conjunto de datos constó de 3,040 imágenes con 25,047 anotaciones, las cuales fueron aumentadas con técnicas de transformación para mejorar la robustez del modelo. Se dividieron en proporción 8:1:1 para entrenamiento, validación y prueba. Las imágenes fueron procesadas y clasificadas con los tres algoritmos mediante una interfaz que permitía elegir el modelo a utilizar. Los resultados mostraron una precisión del 93.63% para Slim-CNN, 98.52% para YOLOv5 y 96.89% para VGG-16, destacando a YOLOv5 como el modelo más eficaz. Las métricas de F1-score, precisión y Recall corroboraron los resultados propuestos. No obstante, una limitación significativa fue el hardware empleado, ya que equipos más potentes como los kits de desarrollo NVIDIA Jetson podrían mejorar el rendimiento. Además, algunas clasificaciones, como los

granos rotos y descoloridos, eran más propensas a errores, lo que indica que es necesario mejorar la calidad del etiquetado y ampliar la base de datos para lograr una mayor generalización [40].

1.4.6. Development of YOLO-Based Mobile Application for Detection of Defect Types in Robusta Coffee Beans

El objetivo de este estudio era crear una aplicación móvil que pudiera detectar automáticamente problemas en los granos de café Robusta. Para ello, los investigadores utilizaron el algoritmo YOLO. El objetivo era crear una herramienta rápida, precisa y fácil de usar que pudiera sustituir a los métodos humanos, que a menudo son erróneos y lentos. Para alcanzar este objetivo, utilizamos Extreme Programming (XP), más conocido como método de desarrollo ágil. Para ello, se creó una aplicación que funciona con Android. A continuación, esta aplicación se vinculó a un modelo de aprendizaje profundo que se había entrenado con 1860 fotos divididas en cuatro grupos: granos normales, mohosos, negros y perforados. El modelo se entrenó con el objetivo de reducir el número de veces que se sobreajustaba y mejorar su capacidad de generalización. Durante la fase de entrenamiento del proceso, se utilizaron dos estrategias: la detención temprana y la regularización L2. El programa permite a los usuarios tomar fotos, enviarlas al servidor para su análisis y encontrar automáticamente y en tiempo real cualquier defecto que pueda haber. La clase de granos negros obtuvo muy buenos resultados en el experimento, con un 95,3 % de respuestas correctas. Por otro lado, la clase de granos mohosos tuvo la precisión más baja, con un 62,2 %. Esto podría deberse a que se parece mucho al grano normal. Aunque el rendimiento general fue bastante bueno, el mayor problema encontrado fue que al modelo le costaba distinguir los defectos cuando las

diferencias entre las clases se basaban en diferencias de color muy pequeñas. Los datos no pasaron por ninguna fase de argumentación y la imagen no muestra ningún grano que se haya añadido [41].

1.4.7. Enhanced Coffee Bean Defect Detection Using Super-Resolution

GANs and Deep Learning

Este estudio muestra cómo un sistema inteligente compuesto por redes generativas antagónicas de superresolución (SR-GAN) y redes neuronales convolucionales (CNN) podría ayudar a detectar defectos en los granos de café con mayor facilidad. El objetivo de este sistema en el sector cafetero es automatizar y mejorar el control de calidad para obtener los mejores resultados. Tradicionalmente, este proceso se realizaba de manera manual o mediante fotografías de mala calidad, para solucionar este problema, los autores idearon un método que comienza tomando fotografías de los granos, luego preprocesa las imágenes para separarlos, utiliza SR-GAN para mejorar la nitidez y el detalle de los granos y, finalmente, utiliza una CNN entrenada para clasificar los granos en grupos defectuosos y no defectuosos. Con este método, fue posible ver varios defectos (manchas, deformaciones y daños en la superficie) los cuales antes eran difíciles de ver en fotografías borrosas. Estos problemas incluían defectos en la superficie. Gracias a ello, el sistema fue capaz de detectar cosas con mucha más precisión que los modelos que utilizan fotos sin editar. Se consiguió una tasa de precisión global del 72,37 %. Por otro lado, un gran problema es que SR-GAN no funciona bien a menos que el sistema parta de fotos de buena calidad. Además, no se indica si el modelo puede adaptarse bien a otros tipos de defectos o variaciones del café, lo que lo hace

menos universal. También sería importante probar su funcionamiento en entornos de producción reales, especialmente cuando factores como la luz, la suciedad o el movimiento de la cámara pueden afectar al rendimiento [42].

1.4.8. Coffee bean graded based on deep net models

Este trabajo presenta un enfoque basado en redes neuronales profundas para clasificar los granos de café Robusta en nueve grados de calidad diferentes. Se desarrolló un sistema automatizado utilizando imágenes capturadas en condiciones controladas, con un conjunto de datos compuesto por 450 imágenes distribuidas uniformemente entre las categorías. Se utilizaron cuatro arquitecturas de aprendizaje profundo para desarrollar el modelo: ResNet-34, Inception v3, VGG-16 y EfficientNet-B0. Estas redes se entrenaron utilizando diferentes ratios de entrenamiento/prueba y hasta 50 épocas por configuración. Entre los resultados más destacados, EfficientNet-B0 alcanzó una precisión del 100 % en la clasificación de los nueve grados, incluso en condiciones de iluminación y fondo difíciles, mientras que ResNet-34 también mostró un rendimiento excepcional, con una precisión de hasta el 99,07 %. Por el contrario, VGG-16 e Inception v3 obtuvieron peores resultados, con precisiones que oscilaron entre el 57,82 % y el 75,43 %. A pesar de los buenos resultados, una limitación importante del estudio es que no detecta un grano a la vez, al contrario, evalúa todo el conjunto, es decir no es práctico al momento de usarlo, otro punto es el pequeño tamaño del conjunto de datos, lo que puede afectar a la generalización del modelo en escenarios reales. Además, el uso de imágenes de teléfonos móviles, aunque práctico, podría comprometer la coherencia de los datos. También se observó que algunas arquitecturas, como

Inception v3, mostraron una alta pérdida de validación, lo que sugiere posibles problemas de sobreajuste. Por lo tanto, las investigaciones futuras deberían centrarse en ampliar el conjunto de datos, mejorar la variabilidad de las condiciones de captura y aplicar técnicas de regularización o aprendizaje por transferencia para reforzar el modelo [43].

1.4.9. Coffee Bean Defects Automatic Classification Realtime Application

Adopting Deep Learning

A partir de los resultados de este estudio, se sugiere crear una aplicación en tiempo real que clasifique automáticamente los defectos de los granos de café utilizando YOLOv8 y otros métodos de procesamiento de imágenes. El objetivo de este estudio en la industria del café es reducir la necesidad de análisis humano y mejorar la eficacia de los métodos tradicionales de inspección visual. Estos procesos de selección se los han hecho de manera manual el cual es costoso en términos de mano de obra, sino que también está sujeto a juicios subjetivos y presenta problemas de precisión en el diagnóstico. Debido a este problema, ahora podemos utilizar tecnologías que emplean visión artificial e inteligencia para crear soluciones que detectan mejor los defectos en los granos de café y pueden utilizarse a mayor escala. Para este modelo, recopilamos imágenes de granos de café con nueve tipos diferentes de defectos comunes en el sector cafetero. Utilizamos técnicas de preprocesamiento y aumento de datos para hacer el modelo más robusto y garantizar su generalización. Se utilizó la arquitectura YOLOv8 para entrenar el sistema de clasificación. Este diseño es famoso por su capacidad para clasificar objetos de forma rápida y precisa. También utilizamos OpenCV para procesar las imágenes, lo que facilitó la segmentación y la localización de características importantes en los granos de café. Los estudios demostraron que el sistema creado

puede detectar y clasificar los problemas de los granos de café en uno o tres segundos, con una precisión del 38 %. Este sistema es una buena solución que puede utilizarse en muchos lugares diferentes, como granjas y fábricas, pero presenta limitaciones las cuales pueden influir en la selección. El modelo es más preciso para defectos visibles, como granos rotos o negros, pero menos preciso para defectos más sutiles, como hongos o daños causados por insectos. Los granos rotos y los granos rotos son otros dos ejemplos de defectos. También puede resultar difícil para los pequeños fabricantes utilizar este método debido a los altos costes de la instalación inicial y a la necesidad de contar con la infraestructura técnica adecuada [44].

CAPÍTULO 2

2. Desarrollo

2.1. Adquisición de imágenes

Cada una de las 4090 fotos que se recalentaron presentaban granos sanos y granos defectuosos, con un porcentaje de granos sanos del 56,48 % y de granos defectuosos del 43,52 %. Tal y como se indica en la (Figura 29) y en su esquema (Figura 30), las imágenes se tomaron con una cámara digital Canon SX50 HS colocada sobre un trípode. La distancia entre el objetivo y el fondo era de 14 centímetros, y el fondo era blanco. Para esta investigación se utilizó el objetivo integrado en la cámara. Se ajustó para tener una distancia focal de 55 milímetros y una apertura de 5. Este conjunto de parámetros se obtuvo mediante un proceso de prueba y error hasta que se consiguió una imagen nítida que permitía detectar los defectos de los granos de café.

Figura 29

Captura de imágenes de los granos de café



Figura 30

Esquema de captura de imágenes de los granos de café



El tamaño de las imágenes obtenidas se procesó para la correspondiente eliminación de objetos extraños. Posteriormente, las imágenes se etiquetaron utilizando la aplicación online “CVAT”, Se etiquetó un total de 16235 granos. El motivo para usar esta herramienta se expresa en base a la comparación realizada en las tablas 1 y 2 [45].

Los criterios de gestión y visión artificial que se muestran en las tablas 2 y 3 fueron los que determinaron la elección de la herramienta “CVAT”. Esta decisión nos permitió descartar las alternativas que no eran adecuadas para su implementación.

Debido a su capacidad de cooperación y a los resúmenes estadísticos del proceso de etiquetado de imágenes, “Labellmg” fue la primera alternativa que se tuvo en cuenta. Por otro lado, esta alternativa no se consideró porque no contaba con ningún tipo de soporte, lo que la hacía obsoleta. “LabelStudio” fue la otra opción que se tuvo en cuenta; sin embargo, no se eligió debido al gasto que habría supuesto para el desarrollo de este proyecto.

Tabla 1

Una descripción general de las funciones accesibles en cada herramienta, comparadas entre sí.

Nombre de la herramienta	Visión por computador	Procesamiento de lenguaje	Automatización y desarrollo	Administración y QA
Colabeler	57%	100%	50%	0%
CVAT	86%	0%	50%	17%
diffgram	100%	100%	100%	83%
ImageTagger	57%	0%	0%	0%
Label Studio	71%	100%	100%	17%
Labelbox	86%	100%	100%	67%
LabelD	29%	0%	0%	0%
Labelimg	29%	0%	0%	0%
LabelMe	86%	0%	0%	17%
makesense.ai	86%	0%	50%	0%
Playment	86%	0%	0%	17%
Ratsnake	86%	0%	50%	17%
RectLabel	71%	0%	50%	0%
Remo.ai	57%	0%	100%	33%
V7 Darwin	100%	0%	100%	67%
VGG Image Annotation	100%	0%	50%	0%
VoTT	57%	0%	50%	33%
COCO Annotator	43%	0%	50%	0%
EVA	29%	0%	0%	0%
SuperAnnotate	100%	0%	50%	0%

Con la ayuda del especialista, pudimos dividir las fotografías en dos categorías distintas: «buenas» y «malas». Procedimos de esta manera teniendo en cuenta su uso general a largo plazo, tomando en consideración el tipo de defecto del grano y rechazándolo. Esto se debe a que los agricultores no tienen en cuenta estos defectos al separar los granos, sino que simplemente los clasifican para su procesamiento específico.

Tabla 2

Capacidades de garantía de calidad y gestión dentro de las diversas herramientas.

Nombre de la herramienta	Gestión de calidad	Gestión de proyecto	Gestión de datos	Consenso	Benchmark	Métricas de rendimiento
Colabeler	X	X	X	X	X	X
CVAT	X	✓	X	X	X	X
diffgram	✓	✓	X	✓	✓	✓
ImageTagger	X	X	¿	X	X	X
Label Studio	X	X	X	✓	X	X
Labelbox	✓	X	X	✓	✓	✓
LabelD	¿	¿	X	¿	¿	¿
Labelimg	X	X	X	X	X	X
LabelMe	X	X	X	✓	X	X
makesense.ai	X	X	X	X	X	X
Playment	X	X	✓	X	X	X
Ratsnake	X	X	✓	X	X	X
RectLabel	X	X	X	X	X	X
Remo.ai	X	✓	✓	X	X	X
V7 Darwin	✓	X	X	✓	✓	✓
VGG Image Annotation	X	X	X	X	X	X
VoTT	X	✓	X	X	✓	X
COCO Annotator	X	X	X	X	X	X
EVA	X	X	X	X	X	X
SuperAnnotate	X	X	X	X	X	X

2.2. Preprocesamiento de las imágenes

Una vez completado el trabajo de etiquetado de imágenes, fue posible descargar las imágenes en diversos formatos basados en diversas herramientas, como “LabelMe” o “COCO”, o modelos como “YOLO”, que resultaron útiles para el desarrollo de este proyecto. Sin embargo, en ese caso concreto, el conjunto de datos se descargó en formato “Pascal VOC”, que devuelve tanto la imagen como un archivo XML que detalla las diversas características de las imágenes, como el nombre, el tamaño y la clasificación de las coordenadas del cuadro de Bounding. Se seleccionó este formato concreto debido a que la utilidad se utilizó para el preprocesamiento.

2.2.1. Data augmentation

La herramienta de roboflow fue la elegida para realizar las transformaciones necesarias para el proceso como lo hicieron en el trabajo de [3]. Esta herramienta se utilizó para aplicar varios métodos de aumento de datos con el fin de hacer que el conjunto de datos fuera más robusto. Estos métodos contribuyeron a ampliar el conjunto de datos, mejorar la capacidad de generalización del modelo y aumentar la variedad de las fotos.

- Generación de Imágenes Aumentadas: Para la creación de tres nuevas imágenes aumentadas para cada imagen original se aplican cambios aleatorios a cada imagen original de la colección.
- Recorte (Crop): El nivel de zoom puede ser de entre 0 % (sin recorte) y 20 %. Para ello, se puede seleccionar una parte aleatoria de la imagen original.
- Corte Inclinado (Shear): Las imágenes se modifican mediante una transformación de cizallamiento, que mueve los píxeles horizontal o verticalmente. El ángulo de cizallamiento se puede cambiar dentro de un rango de ± 15 grados en ambos ejes.
- Escala de Grises (Grayscale): El 15% de las imágenes del conjunto de datos se convierte a escala de grises. Esta transformación ayuda al modelo a aprender características relevantes sin depender únicamente de la información cromática.
- Ajuste de Tono (Hue): Para la variación de colores se modificó el tono de las imágenes en un rango de -50° a $+50^\circ$.

2.2.2. Dataset

Con el fin de alcanzar el objetivo de mejorar la solidez y representatividad de la recopilación de datos que se está utilizando en este proyecto, se consideró la inclusión de fotografías del trabajo realizado por [3]. Estas imágenes aportaron una mayor diversidad en características visuales y contextos, complementando significativamente el conjunto de datos original. Como resultado de esta incorporación, el conjunto final se compuso de un total de 7,044 imágenes.

2.2.2.1. Versión 1

En la versión 1 del conjunto de datos, se logró un total de 16,916 imágenes tras aplicar las técnicas de aumentación de datos antes mencionadas. Lo que mejora la capacidad para entrenar modelos de detección que sean robustos y generalizables. Una característica destacada de esta versión del dataset es que todas las imágenes cuentan con un fondo blanco uniforme, lo que permite centrarse exclusivamente en las características morfológicas de los granos de café, eliminando distracciones o interferencias visuales del entorno.

Figura 31

Fotografías de café con fondo blanco



2.2.2.2. Versión 2

En la segunda iteración del conjunto de datos, se obtuvieron un total de 18 548 fotos mediante el uso de las metodologías de aumento de datos que se describieron en la versión 1. Esta versión se distingue por introducir una mayor variabilidad en los fondos, ya que el 15% de las imágenes presenta fondos negros y grises, mientras que el resto conserva el fondo blanco uniforme. Este cambio se implementó con el objetivo de aumentar la representatividad del dataset y evaluar la capacidad del modelo para adaptarse a diferentes condiciones de iluminación y contraste.

Figura 32

Fotografías de café con fondo variable



2.3. Arquitectura del modelo

Se eligió utilizar la API de detección de objetos de PyTorch debido a las múltiples ventajas que ofrece. Destaca su soporte para diversos modelos, incluidos aquellos basados en la arquitectura YOLO, lo que permite una amplia flexibilidad en la implementación. Además, la compatibilidad con la librería Ultralytics facilita una integración más sencilla y eficiente, mejorando la generación de gráficos y análisis visuales relacionados con los resultados obtenidos. Esta elección fue estratégica para

aprovechar la potencia y versatilidad de estas herramientas en el desarrollo del proyecto, asegurando un enfoque efectivo en la detección de objetos [46].

2.3.1. YOLO versión 11 nano

YOLO v11 Nano se presenta como una opción excelente debido a su eficiencia computacional y rapidez. Esta versión está optimizada para ser más ligera que otras variantes de YOLO, lo que permite un procesamiento más rápido sin comprometer significativamente la precisión del modelo. Gracias a su bajo requerimiento de recursos, YOLO v11 Nano resulta especialmente adecuada para entornos con limitaciones de hardware. La elección de esta arquitectura se basa en los resultados obtenidos en el estudio realizado por [47], donde se demostró que YOLO v11 Nano superó a versiones anteriores en términos de rendimiento, evidenciando su potencial para obtener resultados mejorables en tareas de detección de defectos en granos de café.

2.3.2. YOLO versión 11 small

YOLO v11 Small es una excelente opción para detectar defectos morfológicos en los granos de café, ya que tiene el tamaño, la velocidad y la precisión adecuados. Aunque es más grande que la versión Nano, sigue siendo ligera y funciona bien. Esto le permite realizar tareas más difíciles y obtener mejores resultados en términos de precisión sin perder mucha velocidad. Esta versión es ideal para situaciones en las que se necesita un modelo que pueda detectar y clasificar muy bien, pero que no consuma muchos recursos. Por ello, funciona bien con dispositivos que tienen una tecnología más potente que la que necesita la versión Nano, pero no es tan pesada como las versiones más

grandes de YOLO. También acelera el proceso de inferencia, que es una parte importante del procesamiento de imágenes que funciona bien [47].

Tabla 3

Diferencia entre Yolo nano y small

Model	size (pixels)	mAP ^{val} 50-95	Speed CPU ONNX (ms)	Speed T4 TensorRT10 (ms)	params (M)	FLOPs (B)
YOLO11n	640	39.5	56.12 ± 0.82 ms	1.55 ± 0.01 ms	2.6	6.5
YOLO11s	640	47.0	90.01 ± 1.17 ms	2.46 ± 0.00 ms	9.4	21.5

Nota: esta comparativa de las diferentes versiones de Yolo11 se obtuvo de [48].

2.4. Entrenamiento del Modelo

2.4.1. Herramientas utilizadas

Para la implementación del modelo de detección de objetos, se utilizó la API de PyTorch. En primer lugar, se exportó el conjunto de datos a través de la API de Roboflow, lo que facilitó la organización y estructuración eficiente de las imágenes y etiquetas. El conjunto de datos se dividió en tres subconjuntos: el 80% se destinó al entrenamiento (train), el 10% a pruebas (test) y el 10% restante a validación (valid).

Figura 33

Notebook para permitir el acceso a los datos

```
!pip install roboflow

from roboflow import Roboflow
r_f = Roboflow(api_key="████████████████████████████████████████")
project = r_f.workspace("anderson-chamorro-ycqsk").project("deteccion_fondo-ooafh")
version = project.version(1)
dataset = version.download("yolov11")
```

Una vez descargado, el archivo contiene tanto las imágenes como sus respectivas etiquetas, organizadas para facilitar su uso en el proceso de entrenamiento. Además, incluye un archivo de configuración en formato. YAML que proporciona información detallada sobre las etiquetas y su estructura, permitiendo una integración adecuada con los modelos de detección de objetos.

Figura 34

Archivo data.yaml

```
1 train: /kaggle/working/DETECCION_FONDO-1/train/images
2 val: /kaggle/working/DETECCION_FONDO-1/valid/images
3 test: /kaggle/working/DETECCION_FONDO-1/test/images
4
5 nc: 2
6 names: ['badrotation', 'goodrotation']
7
8 ▾ roboflow:
9   workspace: anderson-chamorro-ycqsk
10  project: deteccion_fondo-ooafh
11  version: 1
12  license: CC BY 4.0
13  url: https://universe.roboflow.com/anderson-chamorro-ycqsk/deteccion\_fondo-ooafh/dataset/1
```

Kaggle fue la plataforma seleccionada para entrenar el modelo propuesto en este trabajo, siendo ideal para proyectos de machine learning y visión por computadora. Esta plataforma ofrece acceso gratuito de hasta 30 horas semanales, con una capacidad máxima de 16 GB en tarjeta gráfica, 20 GB de almacenamiento en disco y 30 GB de memoria RAM. Estas características permiten un rendimiento adecuado para entrenar modelos de redes neuronales profundas sin necesidad de contar con hardware propio de alto rendimiento. Además, Kaggle proporciona un entorno interactivo basado en notebooks, lo que facilita tanto el desarrollo como la ejecución del código de manera eficiente. La plataforma también cuenta con una comunidad activa que comparte

conocimientos, soluciones y tutoriales, lo que resulta especialmente valioso para resolver problemas y mejorar el rendimiento de los proyectos.

2.4.2. Proceso de Entrenamiento

En primer lugar, se instalaron las bibliotecas necesarias para llevar a cabo el entrenamiento del modelo de detección de defectos morfológicos en granos de café. Para implementar la arquitectura YOLO v11 Nano, se utilizó la biblioteca Ultralytics, que ofrece la implementación oficial de esta versión de YOLO, optimizada para tareas de detección de objetos con alta eficiencia. Además, se incorporó la biblioteca Roboflow, que facilita la gestión, organización y preparación de conjuntos de datos, permitiendo una integración fluida entre el dataset, la arquitectura del modelo y el proceso de entrenamiento. Ambas bibliotecas fueron fundamentales para el desarrollo del modelo y aseguraron un flujo de trabajo eficiente y bien estructurado.

Figura 35

Instalación de dependencias para el entrenamiento

```
!pip install ultralytics  
  
from ultralytics import YOLO  
  
model = YOLO("yolo11n.pt")  
  
!pip install roboflow
```

Utilizando la API Key proporcionada por Roboflow, se facilitó el acceso directo al proyecto creado en la plataforma, lo que permitió descargar versiones específicas del dataset diseñado para la tarea de detección de defectos en los granos de café. Gracias a esta integración, se pudo obtener el conjunto de datos adecuado, con las imágenes previamente etiquetadas, y se incorporaron de manera eficiente al flujo de trabajo. Esto simplificó el proceso de entrenamiento y validación del modelo, garantizando que las imágenes y sus correspondientes etiquetas estuvieran correctamente alineadas para optimizar el rendimiento del modelo de detección.

Figura 36

API key para la obtención de los datos

```
from roboflow import Roboflow
rf = Roboflow(api_key=" ")
project = rf.workspace("deteccion-ivjfi").project("deteccion")
version = project.version(3)
dataset = version.download("yolov11")
```

El modelo YOLO v11 se entrenó utilizando el dataset previamente descargado y configurado. Para optimizar el proceso de entrenamiento, se ajustaron los siguientes parámetros clave:

- data: Se especificó la ruta del archivo.YAML, que contiene información detallada sobre las clases y las imágenes utilizadas para el entrenamiento, facilitando la integración y configuración del dataset.
- batch: Para un equilibrio entre el uso de memoria y la velocidad de entrenamiento el tamaño del lote se configuró a 16.

- epochs: Con la ayuda de la función patience para su entrenamiento más optimo se definió un total de 50 épocas.
- imgsz: para que el modelo procesara imágenes de alta resolución sin sobrecargar los recursos computacionales se estableció El tamaño de las imágenes con una dimensión máxima de 640 píxeles.
- plots: Se habilitó la opción para guardar gráficos de entrenamiento, lo que permite monitorear el progreso y visualizar métricas clave como la precisión y la pérdida a lo largo de las épocas.
- optimizer: Se utilizó el optimizador AdamW, conocido por su capacidad para mejorar el rendimiento en tareas de visión computacional.
- workers: Se configuraron 4 hilos de trabajo para cargar los datos de manera eficiente durante el entrenamiento.

Figura 37

Parámetros configurables para el entrenamiento

```
results = model.train(data="/kaggle/input/script/data.yaml",
                    batch= 64,
                    epochs=150,
                    imgsz=640,
                    plots=True,
                    optimizer="AdamW",
                    workers=4,)
```

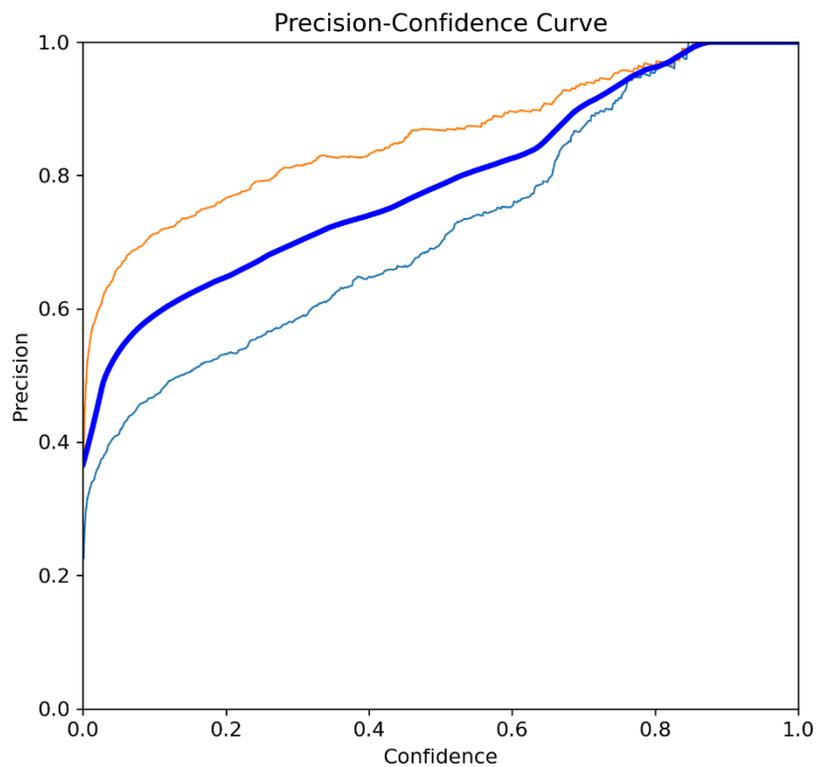
2.4.3. Modelos entrenados

Se procedió al entrenamiento de cuatro modelos específicos, cada uno con diferentes configuraciones, con el objetivo de comparar su rendimiento bajo diversas condiciones. Los modelos entrenados son los siguientes:

- Modelo con arquitectura YOLO v11 Nano y diversos fondos: Este modelo se entrenó utilizando imágenes con diferentes fondos (blanco, negro y gris). La gráfica correspondiente muestra el comportamiento del modelo a lo largo de las épocas de entrenamiento, destacando la precisión y la pérdida del modelo.

Figura 38

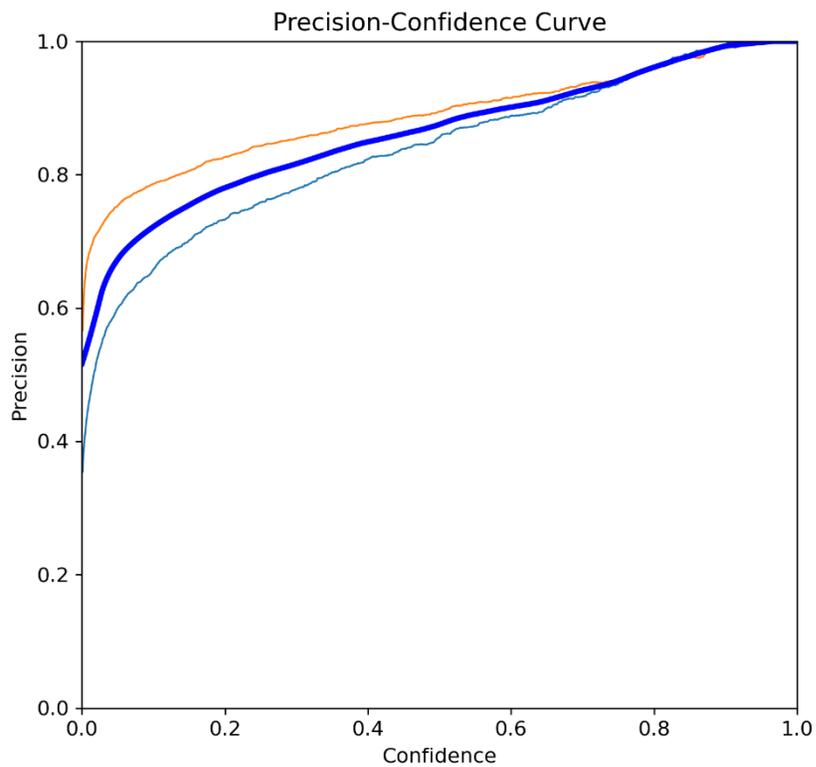
Gráfica de precisión de YOLO v11 Nano y diversos fondos



- Modelo con arquitectura YOLO v11 Nano y fondo blanco: Este modelo se entrenó exclusivamente con imágenes que tienen un fondo blanco. La gráfica asociada ilustra el rendimiento del modelo durante el entrenamiento, proporcionando una comparación directa con el modelo anterior que usó diversos fondos.

Figura 39

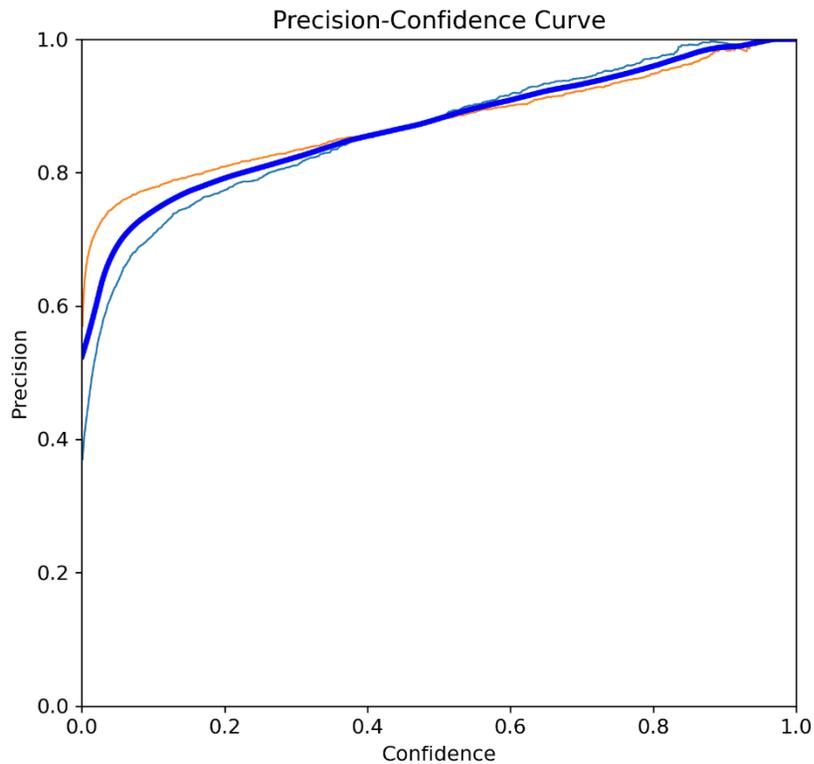
Gráfica de precisión de YOLO v11 Nano y fondo blanco



- Modelo con arquitectura YOLO v11 Small y diversos fondos: Similar al primer modelo, este modelo utilizó la versión más grande de YOLO (v11 Small) con imágenes que contienen una variedad de fondos. La gráfica resultante permite observar cómo se comporta esta arquitectura frente a diferentes fondos durante el proceso de entrenamiento.

Figura 40

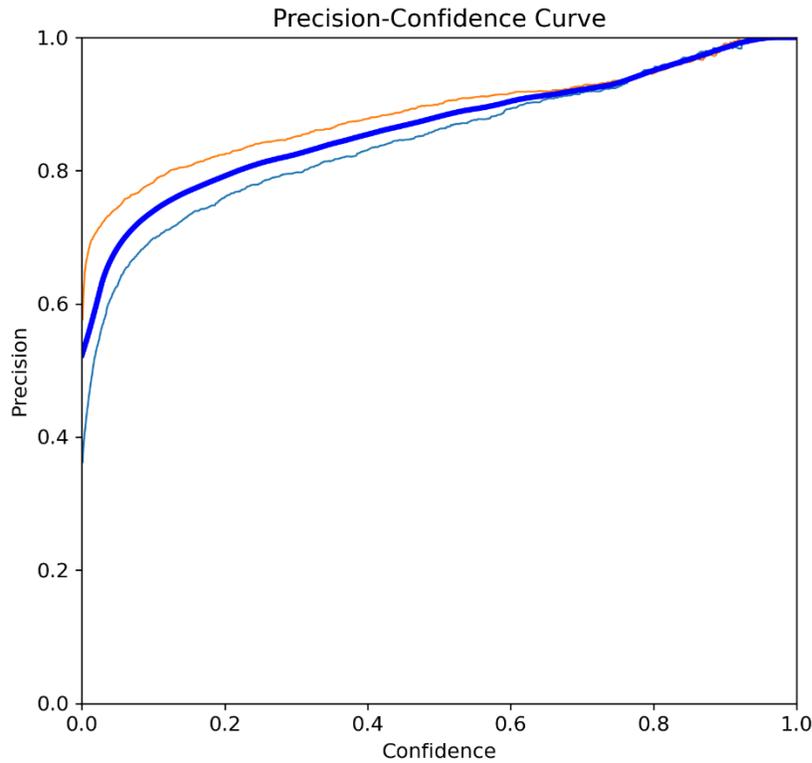
Gráfica de precisión de YOLO v11 Small y diversos fondos



- Modelo con arquitectura YOLO v11 Small y fondo blanco: Finalmente, se entrenó un modelo con YOLO v11 Small, pero esta vez con imágenes que solo tenían fondo blanco. La gráfica correspondiente permite comparar su desempeño con el modelo anterior y evaluar su capacidad para generalizar bajo condiciones específicas de fondo.

Figura 41

Gráfica de precisión de YOLO v11 Small y fondo blanco



En el proceso de llevar a cabo las respectivas sesiones de formación, se obtuvieron los resultados de cada modelo. Estos resultados se muestran en la tabla 4, que incluye las métricas clave, como la precisión, la pérdida y otras estadísticas pertinentes que son necesarias para evaluar el rendimiento de los modelos que se entrenaron con diversas modificaciones. Estas representaciones tabulares permiten realizar una comparación directa entre los resultados obtenidos y los distintos fondos y diseños arquitectónicos que se utilizaron.

Tabla 4
Resultados de los modelos entrenados

	Versión del modelo	Precisión	Recall	mAP	Tamaño de la imagen	Batchsize	Epoch	Optimizer	Dataset	Tiempo	Peso
YOLO V11 SMALL	V1	0,876	0,896	0,948	640x640	16	50	AdamW	16916	7:47:02	38,01
	V2	0,854	0,914	0,95	640x640	16	50	AdamW	18548	8:09:45	38,01
YOLO V11 NANO	V1	0,854	0,904	0,951	640x640	16	50	AdamW	16916	5:55:28	10,6
	V2	0,86	0,905	0,947	640x640	16	50	AdamW	18548	6:11:41	10,6

CAPÍTULO 3

3. Resultados

3.1. Métricas cuantitativas

Dada la naturaleza del problema relacionado con la identificación de defectos en granos de café, se considerarán métricas esenciales para la evaluación, tales como la matriz de confusión, precisión, mAP, y el Recall. Estas métricas facilitarán un análisis exhaustivo del rendimiento del modelo, permitiendo evaluar su habilidad para detectar correctamente los defectos y reducir al mínimo los errores de omisión y de comisión.

3.1.1. Matriz de confusión

La matriz de confusión es una herramienta muy importante para probar la eficacia de los modelos de clasificación cuando se trata de problemas de aprendizaje supervisado. Esto es especialmente cierto cuando se analizan modelos de clasificación que funcionan bien. En este caso, cuando se utilizan redes neuronales convolucionales para detectar problemas morfológicos en los granos de café, esta matriz permite comparar las etiquetas reales de las muestras con las etiquetas que ha predicho la red neuronal. Esto permite obtener una visión completa del funcionamiento del modelo. Esta comparación facilita la detección de problemas de forma en los granos de café.

Para comprender la matriz de confusión del campo $m \times m$, donde m representa el número de categorías en las que se puede clasificar, es necesario localizar una diagonal dentro de esta matriz que tenga los valores adecuados, que se pueden definir de la manera que se muestra en la Tabla 5.

Tabla 5

Matriz de confusión explicada

	Predicho Positivo (Malo)	Predicho Negativo (Bueno)
Real Positivo (Malo)	Verdaderos Positivos (TP)	Falsos Negativos (FN)
Real Negativo (Bueno)	Falsos Positivos (FP)	Verdaderos Negativos (TN)

- TP (Verdaderos Positivos): Casos realmente malos que el modelo clasificó correctamente como malos.
- FN (Falsos Negativos): Casos realmente malos que el modelo clasificó incorrectamente como buenos.
- FP (Falsos Positivos): Casos realmente buenos que el modelo clasificó incorrectamente como malos.
- TN (Verdaderos Negativos): Casos realmente buenos que el modelo clasificó correctamente como buenos.

Una cosa que hay que destacar es que las matrices que se muestran a continuación son matrices de tres por tres. Es posible interpretar la existencia de la tercera fila (Background) como el modelo que identifica un grano en ausencia de cualquier otro grano, independientemente de si el grano es bueno o malo.

La matriz de confusión obtenida de los modelos entrenados son los siguientes:

Figura 42

Matriz de confusión de yolo v11 nano fondo blanco

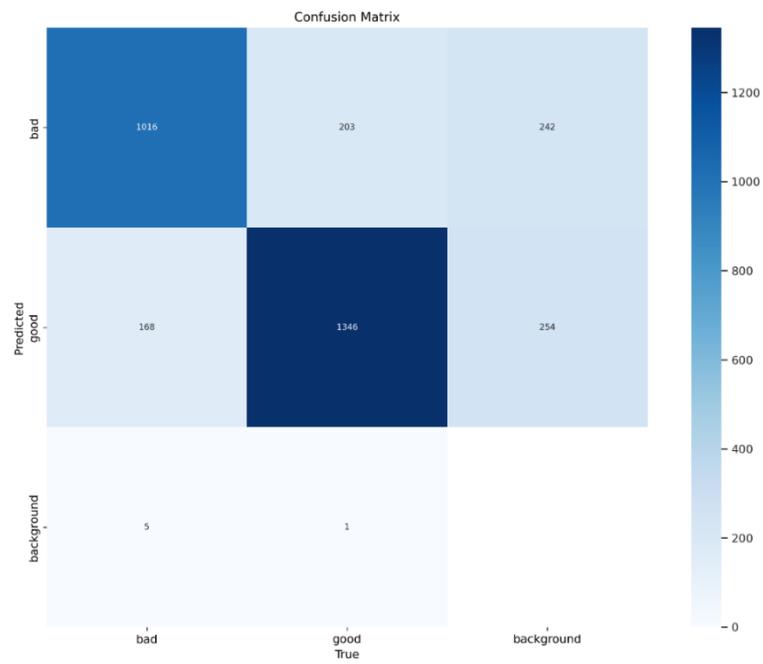


Figura 43

Matriz de confusión de yolo v11 nano fondo variable

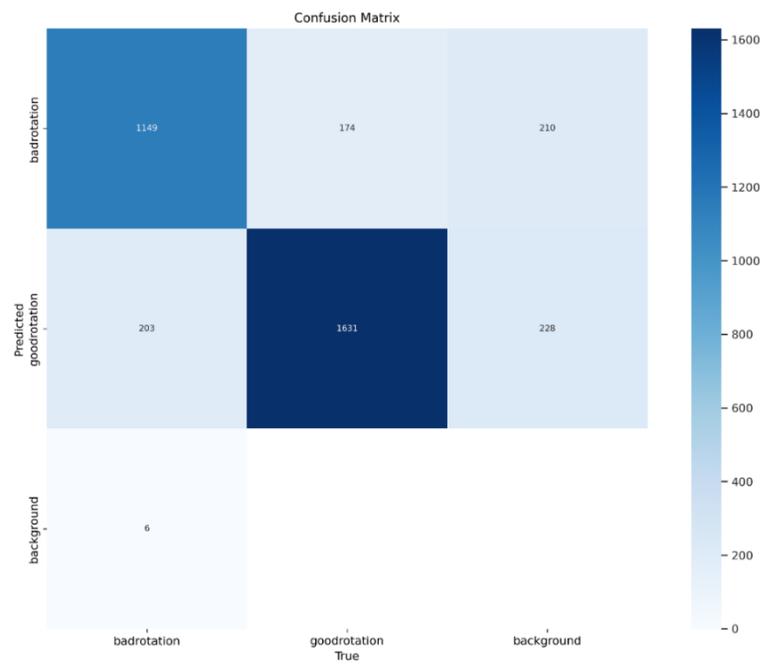


Figura 44

Matriz de confusión de yolo v11 small fondo blanco

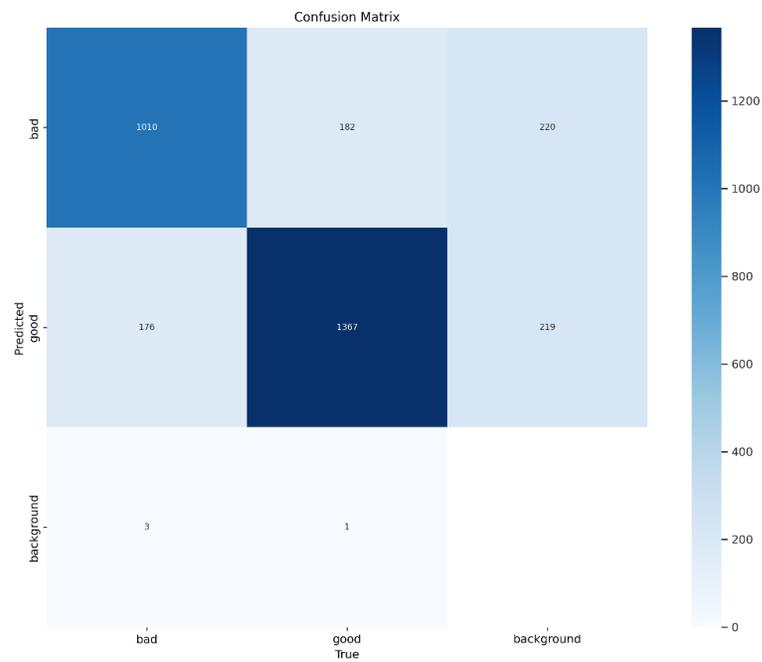
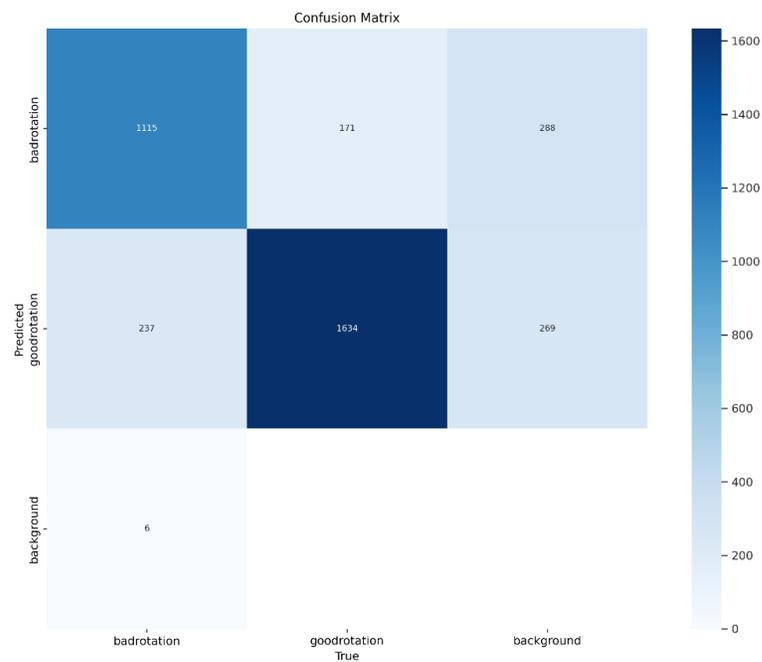


Figura 45

Matriz de confusión de yolo v11 SMALL fondo variable



Como se puede observar en las distintas matrices de confusión, el modelo muestra un desempeño preciso en la clasificación de los casos analizados. En la mayoría de los escenarios evaluados, logra identificar correctamente tanto los granos de buena calidad como aquellos defectuosos. Esto sugiere un alto nivel de confiabilidad en su capacidad predictiva, lo que refuerza su utilidad para la clasificación automática.

3.1.2. Precisión

Debido a que la categorización abarca ambos grupos, se realizaron dos cálculos separados para cada uno de ellos. A continuación, se llevará a cabo un análisis más exhaustivo del modelo, teniendo en cuenta tanto la precisión de la clasificación como la precisión de la localización.

La precisión se determinó utilizando la fórmula como lo dice en el trabajo de [49]. En esta ecuación, el término «TP» se refiere al número de fotos que se han clasificado correctamente en su clase respectiva, mientras que el término «FP» engloba el número de imágenes que se han clasificado erróneamente. Además, se ha tenido en cuenta una tercera categoría, que incluye los casos en los que el modelo ha identificado y clasificado un elemento en una ubicación en la que no había ninguno.

$$\text{Precisión} = \frac{TP}{TP+FP}$$

3.1.2.1. Modelo Yolo versión 11 nano fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 6

Precisión modelo Yolo v11 nano fondo blanco

CLASE	PRESICIÓN
Todos	0,854
Malos	0,828
Buenos	0,879

3.1.2.2. Modelo YOLO versión 11 nano fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 7

Precisión modelo Yolo v11 nano fondo variable

CLASE	PRESICIÓN
Todos	0,86
Malos	0,85
Buenos	0,87

3.1.2.3. Modelo YOLO versión 11 Small fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 8

Precisión modelo Yolo v11 small fondo blanco

CLASE	PRESICIÓN
Todos	0,876
Malos	0,854
Buenos	0,897

3.1.2.4. Modelo YOLO versión 11 Small fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 9

Precisión modelo Yolo v11 small fondo variable

CLASE	PRESICIÓN
Todos	0,854
Malos	0,852
Buenos	0,856

Los resultados obtenidos indican que el modelo YOLO v11 Small con fondo blanco logró la mayor precisión global de 0.876, superando a otras configuraciones, con 0.854 en la detección de granos defectuosos y 0.897 en granos sanos. En general, los modelos YOLO v11 Small mostraron un mejor desempeño en comparación con los modelos YOLO v11 Nano, especialmente en condiciones de fondo uniforme. Sin embargo, el modelo YOLO v11 Nano con fondo variable alcanzó una precisión competitiva de 0.86, sugiriendo una mejor adaptación a escenarios con mayor variabilidad de fondo. Además, en todas las configuraciones, la precisión en la detección de granos sanos fue superior a la de los granos defectuosos, lo que podría reflejar una mayor complejidad en la identificación de defectos o un posible sesgo en el conjunto de datos

3.1.3. RECALL

La recuperación, que a veces se denomina sensibilidad o tasa de verdaderos positivos, es una de las métricas fundamentales que se utilizan en la evaluación de los modelos de clasificación. Esto es especialmente importante en circunstancias en las que existe un desequilibrio entre las clases. La capacidad del modelo para identificar con precisión todas las ocurrencias positivas de una clase determinada se evalúa mediante esta medida.

En términos matemáticos, la recuperación se define como el porcentaje de casos que se identifican correctamente como positivos (verdaderos positivos) en relación con el número total de ejemplos que son verdaderamente positivos.

$$\text{Recall} = \frac{TP}{TP+FN}$$

3.1.3.1. Modelo Yolo versión 11 nano fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 10

Recall modelo Yolo v11 nano fondo blanco

CLASE	RECALL
Todos	0,904
Malos	0,886
Buenos	0,922

3.1.3.2. Modelo YOLO versión 11 nano fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 11

Recall modelo Yolo v11 nano fondo variable

CLASE	RECALL
Todos	0,905
Malos	0,87
Buenos	0,94

3.1.3.3. Modelo YOLO versión 11 Small fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 12

Recall modelo Yolo v11 small fondo blanco

CLASE	RECALL
Todos	0,896
Malos	0,881
Buenos	0,912

3.1.3.4. Modelo YOLO versión 11 Small fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 13

Recall modelo Yolo v11 small fondo variable

CLASE	RECALL
Todos	0,914
Malos	0,875
Buenos	0,953

Los resultados de los modelos en términos de recall indican que el modelo YOLO v11 Small con fondo variante obtuvo el mejor desempeño global, alcanzando un recall total de 0.914 y destacándose especialmente en la detección de granos sanos, con un recall de 0.953. Por su parte, el modelo YOLO v11 Nano con fondo variante logró un recall de 0.905 en la detección de granos defectuosos, colocándose a la par del modelo Small en términos globales, pero mostrando un desempeño superior en la detección de granos sanos con un recall de 0.94. En contraste, el modelo YOLO v11 Small con fondo blanco obtuvo un recall global de 0.896, inferior al de su contraparte con fondo variante,

mientras que el YOLO v11 Nano con fondo blanco logró un recall global de 0.904, ligeramente superior al del modelo Small con fondo blanco. En general, los modelos con fondo variante demostraron un mejor rendimiento en recall, especialmente en la detección de granos sanos, lo que sugiere una mayor capacidad de adaptación a condiciones de fondo más variables.

3.1.4. mAP

La precisión media, a menudo conocida como mAP, es una estadística que se utiliza ampliamente para evaluar la precisión y la capacidad de un modelo para identificar con exactitud instancias de diversas clases existentes dentro de una imagen. Esta medida es especialmente útil en situaciones en las que el modelo no solo debe categorizar correctamente los elementos, sino también localizarlos con precisión, como cuando se identifican defectos morfológicos en los granos de café. En tales situaciones, el modelo no solo debe clasificar correctamente. Se realiza un cálculo para determinar la precisión media absoluta (mAP) tomando la media de las precisiones obtenidas en varios niveles de confianza para cada tipo de objeto. En su forma más básica, se define como la media de la Average Precision (AP) de todas las clases involucradas en el modelo de detección.

La Average Precision (AP), primero se utilizan las curvas de precisión-recuerdo generadas por las predicciones del modelo. Mientras que el recuerdo es una medida de la capacidad del modelo para identificar correctamente todos los casos positivos, la precisión se refiere al porcentaje de predicciones correctas que son proporcionales al número total de predicciones generadas. Para cada clase, se obtiene una precisión

promedio a través de diferentes niveles de recall, resultando en el valor AP para esa clase. La fórmula general es:

$$AP = \frac{1}{N} \sum_{i=1}^N P(i) \cdot \Delta R(i)$$

Donde:

- $P(i)$: es la precisión en el i-ésimo punto de la curva de precisión-recall.
- $\Delta R(i)$: es el cambio en el Recall en el i-ésimo punto

El Mean Average Precisión (mAP) se obtiene promediando las AP de todas las clases. Si un modelo detecta múltiples clases, como los defectos de distintos tipos en los granos de café, el mAP proporciona una métrica global que resume la precisión del modelo para todas las clases al mismo tiempo. Su fórmula es:

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c$$

3.1.4.1. Modelo Yolo versión 11 nano fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 14

mAP modelo Yolo versión 11 nano fondo blanco

CLASE	mAP50
Todos	0,951
Malos	0,943
Buenos	0,959

3.1.4.2. Modelo YOLO versión 11 nano fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 15

mAP modelo Yolo versión 11 nano fondo variable

CLASE	mAP50
Todos	0,947
Malos	0,94
Buenos	0,955

3.1.4.3. Modelo YOLO versión 11 Small fondo blanco

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 16

mAP modelo Yolo versión 11 small fondo blanco

CLASE	mAP50
Todos	0,948
Malos	0,939
Buenos	0,957

3.1.4.4. Modelo YOLO versión 11 Small fondo variante

Los resultados obtenidos de este modelo fueron los siguientes:

Tabla 17

mAP modelo Yolo versión 11 small fondo variable

CLASE	mAP50
Todos	0,95
Malos	0,943
Buenos	0,956

Los resultados de los modelos en términos de recall indican que el modelo YOLO v11 Small con fondo variante obtuvo el mejor desempeño global, alcanzando un recall total de 0.914 y destacándose especialmente en la detección de granos sanos, con un recall de 0.953.

Por otro lado, el modelo YOLO v11 Nano con un fondo variado obtuvo una recuperación del 0,905 en la detección de granos defectuosos. Esto significa que, en general, funciona tan bien como el modelo Small. Sin embargo, obtuvo mejores resultados en la detección de granos sanos, con una recuperación del 0,94. El modelo YOLO v11 Small con fondo blanco, por su parte, obtuvo una recuperación global inferior, de 0,896, que el modelo con un fondo diferente. El modelo YOLO v11 Nano con fondo blanco, por su parte, obtuvo una tasa de recuperación de 0,904, ligeramente superior a la del modelo Small con fondo blanco. En general, los modelos con fondo variable obtuvieron mejores resultados en las pruebas de memoria, especialmente en lo que respecta al reconocimiento de granos sanos. Esto sugiere que se adaptan mejor a diferentes situaciones de fondo. Esto es especialmente cierto cuando se trata de encontrar granos sanos.

3.2. Discusión

Los resultados obtenidos al utilizar YOLO v11 Small y YOLO v11 Nano para detectar defectos morfológicos en granos de café nos permiten comprender cómo las diferentes condiciones del fondo influyen en el rendimiento del modelo. Analizamos dos configuraciones alternativas: fotografías con fondo blanco (V1) e imágenes con fondo variable (V2). Los datos sugieren que las versiones con fondos distintos al blanco son ligeramente menos precisas que las que tienen fondo blanco. Esto es cierto en lo que

respecta a la precisión. Para YOLO v11 Small, la precisión disminuyó de 0.876 a 0.854, mientras que para YOLO v11 Nano, la reducción fue de 0.854 a 0.86. No obstante, se observó una mejora en la métrica de recall en las versiones con fondo variable, lo que sugiere una mayor capacidad para identificar una cantidad superior de granos, aunque con un pequeño sacrificio en la precisión. Esta variación puede atribuirse a la complejidad adicional en la segmentación y clasificación que introduce un fondo no uniforme, dificultando la discriminación entre granos sanos y defectuosos.

Al comparar ambas versiones, queda claro que YOLO v11 Nano, aunque es una versión más ligera, funciona tan bien como YOLO v11 Small. Por ejemplo, YOLO v11 Nano obtiene un mAP de 0,951 en la versión V1, lo que es un poco mejor que el 0,948 de YOLO v11 Small. Este resultado demuestra que YOLO v11 Nano sigue siendo bueno para clasificar granos de café, aunque su diseño sea más sencillo. Podría ser una buena opción cuando no se dispone de suficientes recursos informáticos. Es interesante observar que YOLO v11 Nano, que pesa 10,6 MB, es mucho más ligero que YOLO v11 Small, que pesa 38,01 MB. Además, YOLO v11 Nano tarda menos tiempo en entrenarse, lo que puede ser una gran ventaja cuando el tiempo y la potencia informática son limitados.

Durante los experimentos, quedó claro que a ambos modelos les costaba más encontrar granos defectuosos que granos buenos. Esto se debe a que los defectos morfológicos pueden ser difíciles de ver y menos consistentes que los rasgos de un grano sano. Este resultado implica que las investigaciones futuras podrían centrarse en la creación de métodos de preprocesamiento o enfoques híbridos que ayuden a resaltar

estos rasgos defectuosos. Esto mejoraría la capacidad del modelo para reconocer los granos defectuosos con mayor precisión.

Tabla 18

Comparativa de trabajos propios y trabajos relacionados en datasets en granos de café

	Model	Detección Múltiple	Numero de clases	Acurracy
Literatura [36]	SSD-MobileNet V2	1	Binario (2): bueno y malo	0,917
	SSD-Resnet50	1		0,831
Literatura [39]	MobileNetV3	1	Multiclase (17): Roto, Cortado, Cereza seca, Descolorido, Flotador, Totalmente negro, Totalmente agrio, Daño por hongos, Cascarilla, Inmaduro, Pergamino, Parcialmente negro, Parcialmente agrio, Daño severo por insectos, Cáscara, Daño leve por insectos, Marchito	0,902
Literatura [38]	custom-YOLOv8n	1	Multiclase (5): normal, negro, roto, desvanecido, agrio	0,952
Literatura [37]	MobilNet v1	>1	Multiclase (6): normal, negro, agrio, roto, caracolillo, desvanecido	0.887
Literatura [44]	Yolo v8	>1	Multiclase (10): normal, negro, roto, grande roto, inmaduro, insectos, mezclado, moho, marrón, parte negro	0.38
Literatura [40]	Slim-CNN	>1	Multiclase (5): bueno, medio, malo	0.936
	Yolo v5	>1		0.985
	VGG-16	>1		0.968
Literatura [41]	Yolo	>1	Multiclase (5): normal, negro, roto, descolorido, infectado por insectos	0.783
Literatura [42]	SR-GAN y CNN	>1	Multiclase (4): normal, negro, marrón, roto	0.824
Trabajo propio	Yolo v11 nano (fondo blanco)	>1	Binario (2): bueno y malo	0,864
	Yolo v11 nano (fondo variable)	>1		0,880

	Yolo v11 small (fondo blanco)	>1		0,869
	Yolo v11 small (fondo variable)	>1		0,871

Tabla 19

Acceso y resumen de los datasets de los trabajos relacionados

	Model	Acceso	descripción
Trabajo propio	Yolo v11 (fondo blanco)	Link	Se utilizo un dataset propio en este trabajo, contiene 4099 imágenes, se etiquetaron en dos clases “bueno” y “malo”, el dataset contiene imágenes que muestran un grano por imagen, hasta 11 granos por imagen.
Trabajo propio	Yolo v11 (fondo variable)	Link	el dataset utilizado para este modelo cuenta con la variación de que cambia el fondo blanco, negro y gris.
Literatura	SSD-MobileNet, SSD-Resnet50, V2 [36]	Link	El dataset contiene 3276 imágenes de granos de café, capturadas en condiciones controladas y etiquetadas en dos clases: “bueno” y “malo”. El dataset tiene la limitación que solo muestra un grano por imagen.
Literatura	MobileNetV3 [39]	Link	El dataset contiene 979 imágenes de granos de café verde Arábica tailandés con 17 tipos de defectos. Las imágenes están en 500x500 px, tomadas en condiciones controladas, y fueron aumentadas por rotación a un total de 6,853 imágenes para entrenamiento de modelos de clasificación
Literatura	custom-YOLOv8n [38]	Contacto con hlgope@sau.ac.bd	El dataset contiene 5044 imágenes JPG de granos de café verde con anotaciones en formato YOLO. Las imágenes

			están clasificadas en cuatro tipos de defectos: black, broken, fade y sour, y fueron capturadas en condiciones controladas con fondo blanco. Se encuentra dividido en 80% entrenamiento, 10% validación y 10% prueba
Literatura	MobilNet v1 [37]	Contacto con ING. ESTEBAN FABRICIO GONZABAY JIMENEZ	Se utilizó un dataset de 3000 imágenes de granos de café (buenos, defectuosos y trizados), seleccionadas aleatoriamente de un total de 4000, provenientes de bases de datos públicas y etiquetadas para clasificación con redes neuronales.
Literatura	Yolo v8 [44]	Link	El conjunto de datos utilizado era público y contaba con 4640 imágenes de granos de café Robusta verde sobre fondo blanco en condiciones controladas. Enmarcamos las imágenes y las clasificamos en nueve grupos según los problemas más comunes, como moho, insectos, granos rotos y granos negros. Kaggle tiene el conjunto de datos disponible.
Literatura	Slim-CNN, Yolo v5, VGG-16 [40]	acceso restringido	El dataset contiene 3,040 imágenes de granos de café Coffea Liberica, clasificadas en cinco tipos: normal, negro, roto, infectado por insectos y descolorido. Fue recolectado por agricultores locales, dividido en sets de entrenamiento, validación y prueba (80:10:10), y se aplicaron técnicas de aumentación para mejorar la precisión del modelo
Literatura	Yolo [41]	Acceso restringido	Se utilizó un conjunto de datos compuesto por 1860 imágenes de granos de café robusta recolectadas en Lampung,

			Indonesia. Las imágenes fueron clasificadas manualmente en cuatro categorías: normal, mohoso, negro y agujereado. El dataset fue dividido en proporciones de 90% para entrenamiento, 5% para validación y 5% para prueba
Literatura	SR-GAN y CNN [42]	Acceso restringido	El dataset incluye imágenes de granos de café con y sin defectos, pero el artículo no indica cuántas son, de dónde provienen, ni si están disponibles. Esta falta de información es una limitación.
Literatura	[50]	Link	El artículo muestra un conjunto de datos con 450 imágenes de granos de café que se han clasificado en 9 grupos según su calidad. Su objetivo es ayudar a los modelos de visión artificial a aprender a detectar problemas morfológicos en granos negros, rotos, agrios o dañados por insectos u hongos. EfficientNet-B0 fue capaz de alcanzar una precisión del 100 % en diversas situaciones.

Algunos estudios que intentan encontrar defectos en los granos de café solo analizan un grano en cada imagen. Por ejemplo, utilizan modelos como SSD-MobileNet V2 y SSD-Resnet50 para separar los granos en dos grupos: buenos y malos. Pueden hacerlo con una tasa de éxito de hasta el 91,7 %. De forma similar, [39] emplean MobileNetV3 para clasificar hasta 17 tipos de defectos en imágenes con un solo grano, logrando un accuracy cercano al 90%. En comparación al estudio realizado podemos recalcar que los trabajos que solo detectan un grano a la vez no están cercanos a aplicaciones prácticas reales, el cual es un limitante importante.

En contraste, otros trabajos abordan la detección múltiple de granos dentro de una misma imagen, lo que representa un reto mayor, pero se acerca más a condiciones reales. [37] usa MobileNet v1 para detectar y clasificar cinco tipos de defectos en múltiples granos con un accuracy de 88.7%. [40] presentan resultados destacados con YOLO v5 y VGG-16, alcanzando precisiones superiores al 96% en detección multiclase y múltiple. Sin embargo, [44] reportan un desempeño bajo con YOLO v8 en detección multiclase múltiple, reflejando la dificultad de manejar imágenes con alta complejidad, es decir, que los trabajos que manejan más de una detección de granos por imagen representan más complejidad al momento de realizar la predicción.

El estudio [36] analiza una tarea binaria que separa los granos buenos de los malos en el proceso de clasificación. Utilizamos este método en nuestro propio trabajo con YOLO v11 nano y small, y funciona bastante bien (aproximadamente el 86 % de precisión). Esto facilita la creación de modelos y los hace más sólidos, pero también dificulta la diferenciación entre algunos tipos de fallos. La clasificación binaria suele ser suficiente para operaciones industriales que solo necesitan conocer la diferencia entre la calidad general.

Por otro lado, la clasificación multiclase se enfoca en identificar distintos tipos de defectos morfológicos. [39], [38] y otros autores abordan hasta 17 clases distintas, proporcionando un diagnóstico detallado que puede apoyar un control de calidad más exhaustivo. Esta mayor granularidad implica un desafío técnico mayor y a menudo resulta en una reducción de la precisión debido a la complejidad añadida, como muestra el bajo desempeño de algunos modelos en detección múltiple y multiclase.

Este estudio [43] no es exactamente igual al otro, ya que no identifica granos individuales, sino que clasifica imágenes de todo el mundo. Esta tesis utiliza un enfoque diferente, ya que emplea un modelo para detectar y clasificar objetos morfológicos a nivel de objeto. Esto permite ver muchos granos en una sola imagen y observar sus defectos uno por uno, lo que supone un problema técnico más complejo.

El estudio realizado en Imbabura también reveló que la recolección manual de granos de café tiene una precisión media de solo el 78,9 %. Esto sugiere que las redes neuronales convolucionales (CNN) podrían acelerar el proceso de recolección de granos. Para los pequeños y medianos productores de café que desean clasificar los granos de café con mayor precisión y en menos tiempo, los modelos YOLO v11 podrían ser una buena opción.

3.3. Conclusiones

- Los conocimientos aportados por la persona cualificada fueron esenciales para lograr una clasificación e identificación precisas de los granos sanos y defectuosos. La intervención realizada permitió llevar a cabo un etiquetado más preciso, teniendo en cuenta los aspectos morfológicos que eran difíciles de identificar automáticamente. Además, el análisis exhaustivo de trabajos previos permitió posicionar el modelo propuesto dentro del contexto actual de la investigación. Gracias a esta revisión, fue posible identificar métricas de comparación y validar que los resultados obtenidos con YOLOv11 son competitivos frente a otras propuestas.

- Aunque se utilizaron dos copias del conjunto de datos, una con fondo blanco y otra con fondos variados, no se observó una mejora sustancial en la precisión del modelo. Esto significa que el modelo es bastante resistente a los cambios de color del fondo, aunque haya algunas pequeñas diferencias en la recuperación y la duración del entrenamiento.
- Uno de los mayores problemas del modelo era que las fotos tenían granos superpuestos. Esto dificultaba la localización exacta de los bordes de cada grano, lo que provocaba errores en la predicción de los rectángulos delimitadores. Sin embargo, el modelo seguía siendo capaz de distinguir entre diferentes ocurrencias en situaciones complicadas.
- El modelo YOLOv11 Nano demostró un rendimiento competitivo frente a la versión Small, a pesar de ser más liviano y requerir menos tiempo de entrenamiento.

3.4. Recomendaciones

- Se recomienda continuar experimentando con datasets alternativos, como los mencionados en trabajos relacionados. Ampliar el conjunto de datos con imágenes provenientes de diferentes fuentes y condiciones puede mejorar la generalización y adaptabilidad del modelo a escenarios reales.
- Una implementación que se podría hacer en el futuro es hacer que el modelo funcione con más de una clase de defectos. Esto permitiría no solo encontrar granos defectuosos, sino también determinar qué tipo de defecto

es y con qué frecuencia se produce. Esto ayudaría a evaluar la calidad de la cosecha.

- Para comprobar si el modelo funciona, se recomienda probar en la vida real utilizándolo con dispositivos embebidos o cámaras inteligentes. Esta validación le permitirá ver cómo funciona en tiempo real y en diferentes situaciones, como cuando cambia la iluminación, el movimiento o el fondo.
- Sería útil examinar diferentes arquitecturas de detección o cambiar los hiperparámetros del modelo existente. Se podría mejorar el rendimiento sin realizar grandes cambios en la estructura del modelo cambiando la tasa de aprendizaje, el tamaño del lote o el tipo de aumento.
- Se recomienda estudiar el impacto de la resolución y calidad de las imágenes en el desempeño del modelo. Esto permitiría determinar si es posible mantener una alta precisión con imágenes de menor tamaño, optimizando así el uso de recursos computacionales. O si mejorar la calidad de la imagen puede representar una mejora.

BIBLIOGRAFÍA

- [1] FAO, “Perspectivas a Plazo Medio de los Productos Básicos Agrícolas: Proyecciones al año 2010”, *Fao*, vol. 1, p. 114, 2004, Consultado: el 14 de enero de 2025. [En línea]. Disponible en: <http://www.fao.org/3/y5143s/y5143s0z.htm>
- [2] MAGAP, “MAGAP ejecuta ‘Proyecto de Reactivación de la Caficultura Ecuatoriana’ – Ministerio de Agricultura y Ganadería”. Consultado: el 14 de enero de 2025. [En línea]. Disponible en: <https://www.agricultura.gob.ec/magap-ejecuta-proyecto-de-reactivacion-de-la-caficultura-ecuatoriana/>
- [3] M. A. Cevallos Sánchez, “Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia artificial y Visión por computador”, jun. 2023, Consultado: el 2 de diciembre de 2024. [En línea]. Disponible en: <https://repositorio.utn.edu.ec/handle/123456789/14099>

- [4] M. S. Islam, F. A. Foysal, N. Neehal, E. Karim, y S. A. Hossain, "InceptB: A CNN Based Classification Approach for Recognizing Traditional Bengali Games", *Procedia Comput Sci*, vol. 143, pp. 595–602, ene. 2018, doi: 10.1016/J.PROCS.2018.10.436.
- [5] N. Unidas, "La Agenda 2030 y los Objetivos de Desarrollo Sostenible: una oportunidad para América Latina y el Caribe. Objetivos, metas e indicadores mundiales", el 7 de enero de 2019, *CEPAL*. Consultado: el 25 de junio de 2025. [En línea]. Disponible en: <https://hdl.handle.net/11362/40155>
- [6] B. D. EFEAGRO, "El café modera su precio en un año con más producción prevista". Consultado: el 25 de junio de 2025. [En línea]. Disponible en: <https://efeagro.com/cafe-precio-2023/>
- [7] Banco Central del Ecuador, "La exportación de café ecuatoriano de especialidad crece - El Comercio", La exportación de café ecuatoriano de especialidad crece Este contenido ha sido publicado originalmente por EL COMERCIO. Si vas a hacer uso del mismo, por favor, cita la fuente y haz un enlace hacia la nota original en la dirección: <https://www.elcomerci...> Consultado: el 25 de junio de 2025. [En línea]. Disponible en: <https://www.elcomercio.com/actualidad/negocios/exportacion-cafe-ecuatoriano-especialidad-crecio/>
- [8] P. Vázquez-López, J. de J. Espinoza-Arellano, A. González-Mancilla, y L. A. Guerrero-Ramos, "Characteristics of coffee producers and plantations in the northern region of the state of Chiapas[Características de productores y plantaciones de café en la zona norte de Chiapas]", *Rev Mex De Cienc Agric*, vol. 2022, núm. 28, pp. 101–111, sep. 2022, doi: 10.29312/remexca.v13i28.3266.
- [9] L. A. León-Serrano *et al.*, "Ecuador: producción de banano, café y cacao por zonas y su impacto económico 2013-2016", *Revista Científica UISRAEL*, vol. 7, núm. 3, pp. 103–121, sep. 2020, doi: 10.35290/RCUI.V7N3.2020.324.
- [10] J. P. Enríquez, R. F. Retes-Cálix, y E. F. Vásquez-Reyes, "Importancia, genética y evolución del café en Honduras y el mundo", *Innovare: Revista de ciencia y tecnología*, vol. 9, núm. 3, pp. 149–155, dic. 2020, doi: 10.5377/INNOVARE.V9I3.10649.
- [11] J. Garcia Lopez, "Aptitud para cultivo de café arábigo en el departamento de Santander y variaciones de los escenarios por futuros incrementos de temperatura.", Diplomado de profundización para grado. Consultado: el 21 de abril de 2024. [En línea]. Disponible en: <https://repository.unad.edu.co/handle/10596/56111>
- [12] C. Bunn, M. M. Lundy, P. R. D. Läderach, F. A. Castro-Llanos, y E. H. Girvetz, "Café sostenible adaptado al clima en Honduras", 2018, *International Center for Tropical Agriculture*. Consultado: el 21 de abril de 2024. [En línea]. Disponible en: <https://hdl.handle.net/10568/105436>

- [13] P. C. Parada Molina, C. R. Cerdán Cabrera, y G. C. Ortiz Ceballos, “Condiciones microclimáticas en dos sistemas de producción de café”, *UVserva*, núm. 12, 2021, doi: 10.25009/uvs.vi12.2808.
- [14] “Presentación del cinturón de países productores del Café – Coffee Bean and Birds”. Consultado: el 19 de noviembre de 2024. [En línea]. Disponible en: <https://coffeebeanandbirds.com/es/blogs/notas-de-interes/presentacion-del-cinturon-de-paises-productores-del-cafe>
- [15] I. I. de C. para la A. (IICA), “Manual de producción sostenible de café en la República Dominicana”, 2019, Consultado: el 19 de noviembre de 2024. [En línea]. Disponible en: <https://bvearmb.do/handle/123456789/1379>
- [16] L. E. ISAZA G., E. C. MONTOYA R., J. C. VELEZ Z., y C. E. OLIVEROS T., “Evaluación de la concentración de los frutos maduros de café empleando técnicas no selectivas de recolección manual”, jul. 2007, Consultado: el 25 de noviembre de 2024. [En línea]. Disponible en: <https://biblioteca.cenicafe.org/handle/10778/106>
- [17] J. A. WALLIS G., E. C. MONTOYA R., J. C. VELEZ Z., y C. E. OLIVEROS T., “Calidad y eficacia de dos métodos no selectivos de recolección manual de café *Coffea arabica*”, sep. 2004, Consultado: el 25 de noviembre de 2024. [En línea]. Disponible en: <https://biblioteca.cenicafe.org/handle/10778/113>
- [18] P. A. Leguizamo Cuchimba, A. K. Perdomo Valderrama, L. Y. Medina Trujillo, y A. Ducuara Manrique, “PRÁCTICAS EN LA POST COSECHA DEL CAFÉ Y SU INFLUENCIA EN LAS FINANZAS DEL CAFICULTOR”, *Creceer Empresarial Journal of Management and Development*, vol. 0, núm. 02, pp. 16–28, dic. 2018, [En línea]. Disponible en: <https://journalusco.edu.co/index.php/cempresarial/article/view/1877>
- [19] C. Oliveros-Tascón *et al.*, “SEPARACIÓN DE FRUTOS DE CAFÉ VERDES POR MEDIOS MECÁNICOS”, *Cenicafé*, vol. 61, pp. 260–269, ene. 2010.
- [20] G. I. Otero Lastra, “Análisis de la capacitación y asistencia técnica a los productores de café, caso de estudio el noroccidente de Quito 2014-2018”, el 27 de mayo de 2019, *PUCE-Quito*. Consultado: el 2 de diciembre de 2024. [En línea]. Disponible en: <https://repositorio.puce.edu.ec/handle/123456789/34090>
- [21] E. Alegre, G. Pajares, y A. de la Escalera, “Conceptos y métodos en visión por computador”, *Conceptos y métodos en visión por computador*, 2016.
- [22] G. Enrique Urrego, F. C. Calderón, A. Forero, y J. A. Quiroga, “Adquisición de variables de tráfico vehicular usando visión por computador”, *Revista de Ingeniería*, núm. 30, pp. 7–15, 2009, [En línea]. Disponible en: <https://www.redalyc.org/articulo.oa?id=121015710006>
- [23] L. A. Ashame, S. M. Youssef, M. N. Elagamy, A. Othman, y S. M. El-Sheikh, “A Computer-Aided Model for Dental Image Diagnosis Utilizing Convolutional Neural Networks”, *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 52, núm. 2, pp. 15–25, 2025.

- [24] E. A. (coord.), G. P. (coord.), y A. de la E. (coord.), *Conceptos y métodos en visión por computador*. Grupo de Visión del Comité Español de Automática (CEA), 2016. [En línea]. Disponible en: <https://produccioncientifica.ucm.es/documentos/6353371f978f296ba7a95a83>
- [25] F. G. Ortiz Zamora, “Procesamiento morfológico de imágenes en color: aplicación a la reconstrucción geodésica”, 2002.
- [26] F. O. Pérez Ramírez y H. Fernández Castaño, “Las redes neuronales y la evaluación del riesgo de crédito”, *Revista Ingenierías Universidad de Medellín*, vol. 6, núm. 10, pp. 77–91, 2007.
- [27] D. J. Matich, “Redes Neuronales: Conceptos básicos y aplicaciones”, *Universidad Tecnológica Nacional, México*, vol. 41, pp. 12–16, 2001.
- [28] P. Larranaga, I. Inza, y A. Moujahid, “Tema 8. redes neuronales”, *Redes Neuronales, U. del P. Vasco*, vol. 12, p. 17, 1997.
- [29] S. I. Fernández Velasco y K. A. Ramos Casanova, “Implementación de un sistema con inteligencia computacional para identificar dificultad respiratoria a partir del procesamiento digital de señales de voz”, 2022.
- [30] P. Loncomilla, “Deep learning: Redes convolucionales”, *Recuperado de <https://ccc.inaoep.mx/~pgomez/deep/presentations>*, 2016.
- [31] A. Sherstinsky, “Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network”, *Physica D*, vol. 404, p. 132306, 2020.
- [32] M. J. Er, S. Wu, J. Lu, y H. L. Toh, “Face recognition with radial basis function (RBF) neural networks”, *IEEE Trans Neural Netw*, vol. 13, núm. 3, pp. 697–710, 2002.
- [33] M. Valizadeh y S. J. Wolff, “Convolutional Neural Network applications in additive manufacturing: A review”, *Advances in Industrial and Manufacturing Engineering*, vol. 4, p. 100072, 2022, doi: <https://doi.org/10.1016/j.aime.2022.100072>.
- [34] S. Coulibaly, B. Kamsu-Foguem, D. Kamissoko, y D. Traore, “Deep Convolution Neural Network sharing for the multi-label images classification”, *Machine Learning with Applications*, vol. 10, p. 100422, 2022, doi: <https://doi.org/10.1016/j.mlwa.2022.100422>.
- [35] S. Qiu *et al.*, “Multi-sensor information fusion based on machine learning for real applications in human activity recognition: State-of-the-art and research challenges”, *Information Fusion*, vol. 80, pp. 241–265, 2022, doi: <https://doi.org/10.1016/j.inffus.2021.11.006>.
- [36] M. Cevallos, L. Sandoval-Pillajo, V. Caranqui-Sánchez, C. Ortega-Bustamante, M. Pusedá-Chulde, y I. García-Santillán, “Morphological Defects Classification in Coffee Beans Based on Convolutional Neural Networks”, en *Technologies and Innovation*, R. Valencia-García, T. Borodulina, J. Del Cioppo-Morstadt, C. E. Moran-Castro, y N. Vera-Lucio, Eds., Cham: Springer Nature Switzerland, 2024, pp. 3–15.

- [37] E. F. Gonzabay Jiménez, “Procesamiento de imágenes para la clasificación de granos de café por forma y color”, 2023, Consultado: el 20 de marzo de 2025. [En línea]. Disponible en: <http://repositorio.utmachala.edu.ec/handle/48000/21622>
- [38] H. L. Gope, H. Fukai, F. M. Ruhad, y S. Barman, “Comparative analysis of YOLO models for green coffee bean detection and defect classification”, *Sci Rep*, vol. 14, núm. 1, 2024, doi: 10.1038/s41598-024-78598-7.
- [39] S. Arwatchananukul, D. Xu, P. Charoenkwan, S. Aung Moon, y R. Saengrayap, “Implementing a deep learning model for defect classification in Thai Arabica green coffee beans”, *Smart Agricultural Technology*, vol. 9, 2024, doi: 10.1016/j.atech.2024.100680.
- [40] J. Manansala y C. Paglinawan, *Classification of Coffea Liberica Quality Using Convolution Neural Networks (Slim-CNN, YOLOv5, and VGG-16)*. 2024. doi: 10.1109/ICCCNT61001.2024.10723931.
- [41] E. D. Nugroho *et al.*, “Development of YOLO-Based Mobile Application for Detection of Defect Types in Robusta Coffee Beans”, *Journal of Applied Informatics and Computing*, vol. 9, núm. 1, pp. 153–160, ene. 2025, doi: 10.30871/jaic.v9i1.8886.
- [42] S. Prabu, P. S. Thanigaivelu, S. Poonkodi, y M. Anand, “Enhanced Coffee Bean Defect Detection Using Super-Resolution GANs and Deep Learning”, *2025 2nd International Conference on Computational Intelligence, Communication Technology and Networking, CICTN 2025*, pp. 827–832, ene. 2025, doi: 10.1109/CICTN64563.2025.10932332.
- [43] B. Jayakumari, A. Mambilamthoda, S. A. S, P. Venkitesan, y V. Raghavendra, “Coffee bean graded based on deep net models”, *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, p. 3084, jun. 2024, doi: 10.11591/ijece.v14i3.pp3084-3093.
- [44] H.-D. Thai, H.-J. Ko, y J.-H. Huh, “Coffee Bean Defects Automatic Classification Realtime Application Adopting Deep Learning”, *IEEE Access*, vol. PP, p. 1, ene. 2024, doi: 10.1109/ACCESS.2024.3452552.
- [45] P. M. Lima y S. Reis, “Data Labeling tools for Computer Vision: a Review”, abr. 2022, Consultado: el 26 de mayo de 2025. [En línea]. Disponible en: <https://run.unl.pt/handle/10362/135873>
- [46] R. Khanam y M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements”, 2024. [En línea]. Disponible en: <https://arxiv.org/abs/2410.17725>
- [47] A. Sharma, V. Kumar, y L. Longchamps, “Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species”, *Smart Agricultural Technology*, vol. 9, p. 100648, dic. 2024, doi: 10.1016/J.ATECH.2024.100648.

- [48] Z. Shariff, "YOLOv11 Is Officially Out! What You Need To Know! | by Zain Shariff | Medium". Consultado: el 23 de enero de 2025. [En línea]. Disponible en: <https://medium.com/@zainshariff6506/yolov11-is-officially-out-what-you-need-to-know-6738c5d25be1>
- [49] D. Patiño-Pérez, L. Armijos-Valarezo, L. Chóez-Acosta, y F. Burgos-Robalino, "Artículo Científico / Scientific Paper Redes neuronales convolucionales para detección de retinopatía diabetica Convolutional neural networks for diabetic retinopathy detection", vol. 33, pp. 91–101, 2025, doi: 10.17163/ings.n33.2025.08.
- [50] B. N. B. J, A. N. K. M, S. A. S, y V. Raghavendra, "CBD: Coffee Beans Dataset", *Data Brief*, vol. 59, p. 111434, abr. 2025, doi: 10.1016/j.dib.2025.111434.