

Modelo clasificador de granos de café cereza en sus estados de maduración basado en Redes Neuronales Convolucionales



Carlos Fabian Paz Duque

Víctor Hugo Pinto Rodríguez

Corporación Universitaria Comfacauca - Unicomfacauca

Facultad de Ingenierías

Popayán, Colombia

2024

Modelo clasificador de granos de café cereza en sus estados de maduración basado en Redes Neuronales Convolucionales



Carlos Fabian Paz Duque

Víctor Hugo Pinto Rodríguez

Director

Diego Andrés Campo Ceballos, PhD

Profesor Asociado Ingeniería Mecatrónica

Trabajo de grado para obtener el título de Ingeniero Mecatrónico

Corporación Universitaria ComfacaUCA - UnicomfacaUCA

Facultad de Ingeniería

Ingeniería Mecatrónica

Popayán, Colombia

2024

Nota de aceptación

Aprobado por el comité de grado en cumplimiento de los requisitos exigidos por la Corporación Universitaria Comfacaucá - Unicomfacaucá para optar al título de Ingeniero Mecatrónico.



Jairo Felipe Ortiz Mosquera

Jurado

Jurado

Agradecimientos

Estamos muy agradecidos con la Corporación Universitaria Comfacauca por brindarnos la oportunidad de avanzar en esta maravillosa carrera. agradecemos especialmente al programa de Ingeniera por su constante apoyo y su fe en nuestras habilidades y su disposición para ayudarnos y guiarnos han sido fundamentales para la finalización de esta tesis.

Quisiéramos expresar nuestro más profundo agradecimiento a nuestro director de tesis, el Dr. Diego Andrés Campo Ceballos por su experiencia, comprensión y apoyo pues contribuyeron al enriquecimiento, apropiación de conocimiento y a ampliar la experiencia en este campo de la Inteligencia Artificial. Su guía constante y su fe en nuestras habilidades nos motivaron a alcanzar este gran logro, en realidad no tenemos palabras para expresar nuestra gratitud por su inmenso apoyo.

Y por último gracias infinitas a nuestros padres, por su amor incondicional y su apoyo moral, su fe en nuestra convicción y determinación, incluso en los momentos más difíciles, han sido el pilar para no desfallecer y obtener este gran logro, es por ustedes y para ustedes.

Tabla de contenidos

Capítulo 1. Introducción	15
1.1. Formulación del problema	15
1.2. Justificación	17
1.3. Objetivos	18
1.3.1. General.....	18
1.3.2. Específicos.....	18
1.4. Organización del documento	19
Capítulo 2. Marco de referencia.....	21
2.1. Marco teórico	21
2.1.1. Café especial.....	21
2.1.2. Inteligencia artificial	21
2.1.3. Neurona Artificial	22
2.1.4. Red Neuronal Artificial.....	23
2.1.5. Aprendizaje profundo	23
2.1.6. Redes neuronales convolucionales	24
2.1.7. Transfer learning.....	24
2.1.8. Arquitectura de red VGG16.....	25
2.1.9. Arquitectura de red VGG19.....	25

2.1.10. Arquitectura de red InceptionResnetV2.....	26
2.1.11. Arquitectura de red Inception V3	26
2.1.12. Arquitectura de red Densenet201.....	26
2.1.13. MobilenetV2	27
2.1.14. Gradiente descendente	27
2.1.15. Backpropagation	28
2.1.16. Desvanecimiento del gradiente	28
2.1.17. Función de activación	28
2.1.18. Validación cruzada.....	29
2.1.19. Sobre entrenamiento	29
2.1.20. Sub-entrenamiento	29
2.1.21. Regularización por parada temprana	29
2.2. Marco contextual	31
2.3. Normas técnicas del café.	31
Capítulo 3. Materiales y Métodos.....	34
3.1. Materiales.....	34
3.1.1. Hardware para la adquisición de imágenes.....	34
3.1.2. Hardware para procesamiento de imágenes y desarrollo de la red neuronal ..	34
3.2. Métodos.....	35
3.2.1. Construcción de los conjuntos de datos de entrenamiento	35

3.2.2. Experimentación con los modelos y conjuntos de datos de Tamayo.....	38
3.2.3. Entrenamiento de una Red Neuronal Convolutiva.....	39
3.2.4. Evaluación de los modelos.....	42
Capítulo 4. Resultados	43
4.1. Construcción del conjunto de datos	43
4.2. Experimentación con modelos y conjuntos de datos de Tamayo	46
4.3. Entrenamiento de una Red Neuronal Convolutiva.....	51
4.3.1. Entrenamiento utilizando el conjunto de datos 1	51
4.3.2. Entrenamiento utilizando conjunto de datos 2.....	55
4.4. Evaluación de los modelos.....	56
Capítulo 5. Conclusiones	59
Trabajos futuros	62
Referencias.....	63
Anexos	67
A1. Estructura del banco de pruebas.	67
A2. Fórmulas para el cálculo de las métricas de evaluación	71
A3. Modificaciones realizadas al código de Tamayo y microservicios del sistema	74
A4. Repositorio de trabajo	75
A5. Servicios creados para las fases de entrenamiento y evaluación de los modelos	

Lista de Figuras

Fig. 1. Elementos que constituyen una neurona artificial [13].	23
Fig. 2. Diferencia entre bloque residual y residual invertido [20].	27
Fig. 3. Metodología secuencial.	35
Fig. 4. Diagrama de flujo de la construcción del dataset 1.	36
Fig. 5. Diagrama de flujo de la construcción del dataset 2.	37
Fig. 6. Imágenes de granos de café cortadas y etiquetadas.	37
Fig. 7. Diagrama de Procesamiento de imágenes.	38
Fig. 8. Diagrama de flujo del proceso realizado para replicar los experimentos de Tamayo.	39
Fig. 9. Diagrama de flujo del entrenamiento de las redes neuronales.	41
Fig. 10. Imágenes de granos de café redimensionadas con iluminación controlada.	43
Fig. 11. Conversión de imagen a blanco y negro.	44
Fig. 12. Recorte de imagen a escala 224x224.	45
Fig. 13. Captura el grano de café en movimiento.	45
Fig. 14. Métricas de evaluación de la arquitectura VGG16 de Tamayo.	47
Fig. 15. Métricas de evaluación de la arquitectura VGG16 de Tamayo con validación cruzada.	48
Fig. 16. Métricas de evaluación de la arquitectura VGG19.	48
Fig. 17. Resultados de la validación cruzada del modelo VGG19.	49
Fig. 18. Resultados de replicar los experimentos de Tamayo con transferencia de aprendizaje, validación cruzada y el conjunto de datos desbalanceado.	50

Fig. 19. Exactitud de las arquitecturas entrenadas.	52
Fig. 20. Sensibilidad de las Arquitecturas entrenadas.	53
Fig. 21. Precisión de las arquitecturas entrenadas.	53
Fig. 22. F1-score medida que genera la diferencia entre precisión y sensibilidad.	54
Fig. 23. Evaluación de los modelos con el dataset 3_1.	58
Fig. 24. Evaluación de los modelos con el dataset 4_1.	59
Fig. 25. Diseño del banco de pruebas para toma de imagenes en movimiento.	68
Fig. 26. Diagrama de plano inclinado para la construcción del banco de pruebas.	69
Fig. 27. foto del proceso de construcción del sistema de adquisición de imágenes.	71
Fig. 28. Diagrama de flujo del servicio de entrenamiento.	76
Fig. 29. Diagrama de flujo del servicio de entrenamiento para todos los modelos.	77
Fig. 30. Servicio de evaluación de un modelo específico.	78
Fig. 31. Servicio de evaluación de un modelo específico.	79

Lista de tablas

Tabla 1. Defectos de los granos de café; tomada de [22].	32
Tabla 2. Número de registros de los conjuntos de datos construidos.	45
Tabla 3. Descripción del dataset desbalanceado de Tamayo.	47
Tabla 4. Métricas de evaluación obtenidas de replicar los experimentos de Tamayo con validación cruzada.	49
Tabla 5. Resultados de entrenar los modelos utilizando el conjunto de datos 1 sin validación cruzada.	52
Tabla 6. Resultados de entrenar los modelos utilizando el conjunto de datos 1 y validación cruzada.	55
Tabla 7. Resultados de entrenar los modelos utilizando el conjunto de datos 2, sin validación cruzada.	56
Tabla 8. Entrenamiento de las arquitecturas, utilizando el conjunto de datos 2 y validación cruzada.	56
Tabla 9. Composición de los dataset 3_1 y dataset 4_1.	58

Resumen

El café es uno de los productos que más impacto tiene en el país, debido a su alta demanda de consumo tanto a nivel nacional como internacional. Además, el departamento del Cauca produce alrededor de 1.500.000 sacos al año, lo que se traduce en un 10,46% de la producción nacional de café. A ello se suma que los estándares son cada vez más exigentes al momento de generar café excelso de tipo exportación, y la adecuada selección de los granos de café en el proceso de cosecha juega un papel fundamental en la calidad final.

Por tal motivo, la Inteligencia Artificial se presenta como una alternativa para reducir la falla humana mediante la implementación de modelos de red neuronal. El presente proyecto muestra la aplicación de redes neuronales convolucionales (CNN) en la clasificación del fruto de café cereza en diferentes estados de maduración.

Para ello, se utilizaron los mejores resultados obtenidos por Tamayo como punto de partida. En ese orden de ideas, se empleó Transfer Learning y se crearon cuatro conjuntos de datos en diferentes entornos de iluminación para el entrenamiento y las pruebas de las arquitecturas de red neuronal convolucional MobileNetV2, VGG16, VGG19, InceptionResNetV2, InceptionV3 y DenseNet201. Al realizar las pruebas con el conjunto de datos número uno este se divide en tres partes: 70% para entrenamiento, 20% para validación y 10% para prueba; como resultado, se obtuvo que las redes VGG16 y Densenet201 obtuvieron unas métricas de exactitud, precisión y sensibilidad del 0.98. Al entrenar los modelos con el Segundo conjunto de datos, el cual fue construido a partir de imágenes estáticas en superficie plana, y al usar la validación cruzada estas métricas dieron como resultado una eficiencia del 1.00, y por último se crean dos dataset externos para prueba de los modelos previamente entrenados, y como consecuencia las arquitecturas VGG 16 y VGG19 tuvieron los mejores resultados con un F1-score de 0.97 para el conjunto de datos 3 que validó el conjunto de datos 1, y 0.96 y 0.97 utilizando el conjunto de datos 4 para validar el 2.

Palabras Clave: Redes neuronales convolucionales, Clasificación del grano cereza de café mediante CNN, Redes neuronales convolucionales para la clasificación de café, aprendizaje profundo en la clasificación de café, clasificación del grano cerezo de café

Abstract

Coffee is one of the products that has the greatest impact on the country, due to its high consumption demand both nationally and internationally. Furthermore, the department of Cauca produces around 1,500,000 bags per year, which translates into 10.46% of the national production of coffee. In addition to this, the standards are increasingly demanding at the moment of generating excellent coffee for export, and the adequate selection of the coffee beans in the harvesting process plays a fundamental role in the final quality.

Por tal motivo, la Inteligencia Artificial se presenta como una alternativa para reducir la falla humana mediante la implementación de modelos de red neuronal. El presente proyecto muestra la aplicación de redes neuronales convolucionales (CNN) en la clasificación del fruto de café cereza en diferentes estados de maduración.

For this purpose, the best results obtained by Tamayo were used as a starting point. In that order of ideas, Transfer Learning was used, and four datasets were created in different lighting environments for training and testing the convolutional neural network architectures MobileNetV2, VGG16, VGG19, InceptionResNetV2, InceptionV3 and DenseNet201. When testing with data set number one this is divided into three parts: 70% for training, 20% for validation and 10% for testing; as a result, it was obtained that the VGG16 and Densenet201 networks obtained accuracy, precision and sensitivity metrics of 0.98. When training the models with the second dataset, which was built from static images on flat surface, and using cross-validation these metrics resulted in an efficiency of 1.00, and finally two external datasets are created for testing the previously trained models, and as a consequence the VGG 16 and VGG19 architectures had the best results with an F1-score of 0.97 for dataset 3 which validated dataset 1, and 0.96 and 0.97 using dataset 4 to validate dataset 2.

Translated with DeepL.com (free version)

Keywords: Convolutional neural networks, Coffee cherry bean classification using CNN, Convolutional neural networks for coffee classification, deep learning in coffee classification, coffee cherry bean classification

Capítulo 1. Introducción

1.1. Formulación del problema

El café de Colombia tiene un impacto significativo en la economía nacional ya que genera más de 9000 empleos directos e indirectos y está presente en 22 departamentos del país [1]. Por otra parte, Colombia es el tercer productor de café tipo arábico en el mundo [2]; destacando además por su calidad ya que Colombia es el país con el mejor café suave del mundo [3]. Respecto a la región del Cauca, es uno de los departamentos líderes en producción de café de alta calidad y como expresa el artículo de la FNC [4] “su clima, variada topografía, reflejo de las cordilleras Occidental y Oriental que lo atraviesan, unidos a la tradición y pasión de su gente, lo convierte en el 4º productor de café a nivel nacional y el primero en número de familias cafeteras, siendo más de 90.000 mil las familias encargadas de cultivar 93.000 hectáreas de café arábico de las variedades Castillo, Colombia, Caturra, Típica, Borbón y Tabí, en 34 municipios para lograr un café de alta calidad”; cómo se puede apreciar el cultivo del café es una de los principales actividades productivas y económicas de la región y de la cual dependen muchas familias pero para producir un café de alta calidad el primer paso se da en el proceso de recolección, En primer lugar, es necesario que se asegure la calidad de la recolección del café, conteniendo en su mayoría frutos maduros y menos del 2,5% de granos de café verdes [5], sin embargo, en la práctica sucede lo contrario, durante el proceso de beneficiado se suele encontrar mezclados granos de café en excelentes condiciones con granos defectuosos, esto se debe a diferentes factores como la iluminación en los entornos de recolección o cafetales, error humano al cosechar de forma manual y falta de capacitación en la producción de un café excelso.

Por otra parte, la tarea repetitiva de recolección y clasificación de café pueden ser optimizadas implementando la agricultura 4.0 la cual abarca la inteligencia artificial en los procesos de control y de cultivos, cosecha y pos-cosecha [6], siendo específicos para poder clasificar el café por sus estados de maduración con mayor agilidad y precisión se necesita de la visión artificial, como rama alterna de la IA, recientemente en Colombia se han realizado estudios en este campo para la clasificación de granos de café cereza en la región del departamento de Boyacá municipio de Caldas usando cámaras hiper-espectrales y técnicas de aprendizaje profundo o Deep Learning [7], este artículo realiza una comparación de 5 tipos de arquitecturas CNN que primero deben ser entrenadas y posterior a ello validadas para saber cuál tiene mejor porcentaje de precisión, para ello se seleccionan un conjunto de imágenes y se separan para entrenamiento y prueba y por tal razón surge la siguiente pregunta de investigación: ¿Cómo desarrollar un modelo clasificador de granos de café cereza en sus estados de maduración aplicando redes neuronales convolucionales?

1.2. Justificación

Según la NCA (National Coffee Association) de Estados Unidos, el café colombiano fue uno de los principales productos de exportación para Estados Unidos y Canadá, con una participación del 47.7% cuyo consumo en Estados Unidos se mantuvo durante el confinamiento en la pandemia. Sin embargo, sigue estando en el cuarto puesto del ranking de mejores productores de café especial por debajo de Indonesia, condición que se puede mejorar por medio de la implementación de tecnología en las diferentes etapas de la producción; especialmente en el control de calidad [8]. Una de los factores que reduce la calidad es que a menudo se realiza sin un adecuado control de la madurez del grano de café, haciendo que el producto pierda valor; crear un dispositivo que facilite el monitoreo de la calidad postcosecha permitirá a las familias caficultoras mejorar su proceso para agregar valor y asegurar un café de alta calidad, gracias a esto, Colombia podrá ocupar un mejor puesto en el ranking propuesto por la NCA; beneficiando a las “aproximadamente 555.692 familias cafeteras” [9] que existen en Colombia.

Además, se espera que la investigación de la inteligencia artificial aplicada a la agricultura 4.0 específicamente al clasificar los granos de café por etapa de madurez, facilite nuevas aplicaciones de la inteligencia artificial en la cadena de valor como clasificadores de café, robots cosechadores y monitores de calidad en la cosecha desde un dispositivo móvil; para ello, los conjuntos de datos generados durante el proyecto se publicaron en un repositorio digital (Anexo A4), de tal forma que la comunidad científica podrá continuar esta línea de investigación; los algoritmos de entrenamiento y evaluación también estarán públicos al finalizar el proyecto.

1.3. Objetivos

1.3.1. General. Definir un modelo de aprendizaje automático que permita la clasificación de los estados de maduración del café cereza del departamento del Cauca basado en redes neuronales convolucionales – CNN.

1.3.2. Específicos.

1. Seleccionar imágenes de granos de café cereza en entornos controlados de iluminación para usarlos como datos de entrada del modelo de clasificación por estado de maduración del mismo en el departamento del Cauca.
2. Implementar una red neuronal convolucional CNN como clasificador involucrando diferentes arquitecturas.
3. Validar la efectividad del modelo con muestras de café recolectadas en la temporada de cosecha en una finca disponible del Cauca.

1.4. Organización del documento

A continuación, se presenta la organización general del documento y un resumen de la información presentada en cada capítulo:

Capítulo 1: Introducción: En este capítulo se presenta el tema de investigación, la justificación, planteamiento del problema y pregunta de investigación, el objetivo general y objetivos específicos.

Capítulo 2: Marco de referencia: En este capítulo se da el contexto del marco teórico, todo lo relacionado a Deep Learning y las diferentes arquitecturas de red que se usaron para la clasificación de granos cerezos de café en diferentes estados de maduración, se expone el marco contextual y las normas técnicas del café para regular la calidad de este.

Capítulo 3: Materiales y Métodos, en este capítulo se expone los métodos para la construcción de cuatro dataset de imágenes usados para entrenamiento y validación, se explicará a profundidad el funcionamiento de las arquitecturas de red neuronal CNN VGG16, VGG19, InceptionResnetV2, InceptionV3, Densenet201 y MobilenetV2 para clasificación de los granos cerezos de café, usando TransferLearning con dataset de imágenes desbalanceados.

Capítulo 4: Resultados: En este apartado se realizan las pruebas de clasificación de los diferentes algoritmos previamente estudiados en el estado del arte usando lenguaje de programación Python y librerías de entrenamiento de redes neuronales keras y tensorflow y procesamiento de datos como opencv y numpy, se graficarán los resultados de las métricas de evolución de la matriz de confusión y validación cruzada del entrenamiento y validación de las arquitecturas de red mencionadas.

Capítulo 5: Conclusión y Trabajos futuros: En este capítulo se presentan las conclusiones obtenidas al evaluar el desempeño de los modelos de red para la clasificación de granos cerezos de café usando redes CNN, y se identificarán posibles trabajos o proyectos futuros que puedan desprenderse del actual trabajo de grado.

Capítulo 2. Marco de referencia

En esta sección, se describen detalladamente los fundamentos teóricos, metodológicos y legales en los que se basa el presente documento.

2.1. Marco teórico

2.1.1. Café especial.

Según la Federación nacional de cafeteros, el café especial es un producto diferenciado por características de origen, preparación o sostenibilidad en su producción. Los orígenes del concepto ‘café especial’ se dan en una conferencia internacional sobre el café en Montreal Francia, en 1978, Erna Knutsen propuso que “microclimas geográficos especiales producen granos con perfiles de sabor únicos, los cuales ella se refirió como ‘cafés especiales’” [10]. En ese orden de ideas, para obtener un café especial, se deben dar las condiciones geográficas adecuadas de clima y química del suelo, también debe cuidarse la conservación, es decir, una cereza de café madura sobre una planta saludable del linaje apropiado, en la tierra adecuada, con las condiciones climáticas apropiadas debería ser escogida en el pico de maduración con el fin de preservar el potencial de excelencia que esta mantiene. Pese a que existen otras condiciones para garantizar un café especial, para fines del presente proyecto, solo se abordará hasta la importancia de la maduración en la calidad del café.

2.1.2. Inteligencia artificial

(*Artificial Intelligence*), Adriana M. Porcelli menciona que este término fue acuñado en 1956 por primera vez por el matemático John McCarthy y su equipo de científicos de la Universidad Darmouth College en los Estados Unidos en donde afirmaron que cada aspecto del aprendizaje y de la inteligencia humana podía ser tan precisamente

descrito que una máquina podría simularlo [11]“; por otra parte, la ISO define la IA como “un campo técnico y científico dedicado al sistema de ingeniería que genera resultados como contenido, previsiones, recomendaciones o decisiones para un conjunto determinado de objetivos definidos por el ser humano”[12].

Uniendo estas dos definiciones, se puede decir que la Inteligencia Artificial es la ciencia que busca dotar de inteligencia a las máquinas para desarrollar tareas que realizan los seres vivos y que de otra forma resultaría imposible. Para el caso del presente proyecto, se evaluará el enfoque de deep learning en la clasificación de café cereza.

2.1.3. Neurona Artificial

una neurona artificial, esta se asemeja al funcionamiento biológico de una neurona cerebral recibiendo una señal de estímulo exterior la cual es procesada y analizada generando una señal estimada de salida que luego es procesada por las siguiente y de esa forma determinar una acción. La neurona cuenta con unos pesos x_i la que determina las características de entrada a la neurona, w_{ij} son los pesos que controlan el nivel de inhibición o excitación de la neurona, el parámetro b_i (sesgo) puede afectar el rendimiento de la red neuronal al predecir una salida, z_i suma de estímulos recibidos por la neurona, $f_{(z_i)}$ función de activación o transferencia y y_i determina la salida generada por la neurona [13].

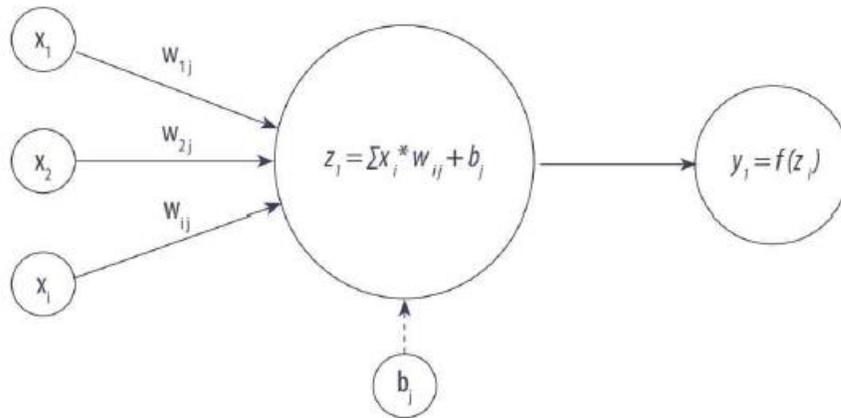


Fig. 1. Elementos que constituyen una neurona artificial [13].

2.1.4. Red Neuronal Artificial

Una red neuronal artificial es un modelo computacional basado en el comportamiento biológico de las neuronas cerebrales [13], esta conexión está conformada por un conjunto de nodos que simulan las conexiones de las neuronas en el cerebro y que normalmente están agrupadas en capas, capa de entrada y capa de salida, en una red neuronal profunda estaría conformada por la capa de entrada, capa oculta o convolucional y capa de salida.

2.1.5. Aprendizaje profundo

(*Deep Learning*). El aprendizaje profundo permite a modelos computacionales compuestos por múltiples capas de procesamiento aprender representaciones de datos con múltiples niveles de abstracción [14]. Es decir, el aprendizaje profundo es un modelo matemático basado en redes neuronales que se caracteriza por tener múltiples capas ocultas y en cada capa logra un mejor nivel de abstracción, a través de la minimización de una función de coste utilizando el gradiente descendente.

2.1.6. Redes neuronales convolucionales

(*Convolutional neural networks - CNN*). Los fundamentos de las redes neuronales convolucionales se basan en el Neocognitron, introducido por Kunihiko Fukushima en 1980. Este modelo fue más tarde mejorado por Yann LeCun et al. en 1998, Este tipo de red es una variación de un perceptron multicapa, sin embargo, debido a que su aplicación es realizada en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones.

Según [15] las CNN son en una serie de capas diseñadas con filtros en las cuales se extraen las características relevantes de la imagen de entrada, en la capa convolucional se generan 2 procesos: la misma convolución (operación para el procesamiento de señales) y la evolución en la función de activación

Carlos M. Pineda define las CNN como una potente herramienta de clasificación basado en el funcionamiento de la corteza visual del cerebro humano la cual está conformada por múltiples niveles y al atravesarlos se va reconociendo más y más la información, la red CNN empieza por la identificación de múltiples pixeles pasando por el reconocimiento de formas geométricas hasta el reconocimiento de formas abstractas como formas, objetos, animales, personas, reconocimiento de voz y procesamiento de lenguaje Natural, pues las capas convolucionales aprenden patrones locales de las características de entrada a la neurona pudiendo así reconocer patrones el cualquier punto[13].

2.1.7. Transfer learning

[16] describe el Transfer Learning como un enfoque del Deep Learning que utiliza un modelo entrenado para una tarea como punto de partida para otro modelo que realiza una tarea similar. Actualizar y volver a entrenar una red con Transfer Learning es más fácil y rápido que entrenarla desde cero, esta técnica se utiliza para clasificación de imágenes, detección de objetos, reconocimiento de voz, y otras aplicaciones.

consiste en usar modelos de red preentrenados para la clasificación de objetos, con este método de entrenamiento se usan los pesos que han sido previamente entrenados gracias a las competencias realizadas por Imagenet y su repositorio de 1000 clases y un aproximado de 1.2 millones de imágenes[17], facilitando la parametrización de los pesos que han sido previamente entrenados con repositorios de imágenes similares lo que agiliza el proceso de entrenamiento del modelo de red con el nuevo dataset de imágenes, asumiendo que los filtros de las primeras capas aprenden características generales de la entrada y a medida que la red profundiza en sus capas aprenden características más abstractas y complejas hasta llegar a la capa final donde predice y clasifica el objeto .

2.1.8. Arquitectura de red VGG16

la red VGG16 es una red que involucra 16 capas, 13 capas convolucionales 2 capas densas de 4096 neuronas cada una y una de salida con 1000 neuronas lo que impacta en la velocidad de procesamiento volviéndola más precisa con una capacidad de extracción de 138.400 parámetros [18].

2.1.9. Arquitectura de red VGG19

VGG-19 es una red neuronal convolucional con 19 capas de profundidad, El tamaño de la entrada de imagen de la red es de 224 x 224, contiene 16 capas convolucionales, 3

capas de fully-conect, 5 capas de MaxPool y una softmax, con 143.7 millones de parámetros [11].

2.1.10. Arquitectura de red InceptionResnetV2

Según (MathWorks), Inception-ResNet-v2 es una red neuronal convolucional que está entrenada con más de un millón de imágenes de la base de datos de ImageNet. La red tiene 164 capas de profundidad y puede clasificar imágenes en 1000 categorías de objetos (por ejemplo, teclado, ratón, lápiz y muchos animales). Como resultado, la red ha aprendido representaciones ricas en características para una amplia gama de imágenes. El tamaño de la entrada de imagen de la red es de 299 por 299

2.1.11. Arquitectura de red Inception V3

(MathWorks) define a Inception-v3 como una red neuronal convolucional con 48 capas de profundidad. Puede cargar una versión preentrenada de la red entrenada en más de un millón de imágenes desde la base de datos [\[1\]](#) de ImageNet. La red preentrenada puede clasificar imágenes en 1000 categorías de objetos (por ejemplo, teclado, ratón, lápiz y muchos animales). Como resultado, la red ha aprendido representaciones ricas en características para una amplia gama de imágenes. El tamaño de la entrada de imagen de la red es de 299 por 299

2.1.12. Arquitectura de red Densenet201

(MathWorks) define la arquitectura DenseNet-201 como una red neuronal convolucional con 201 capas de profundidad. Puede cargar una versión preentrenada de la red entrenada en más de un millón de imágenes desde la base de datos [\[1\]](#) de ImageNet. esta red puede clasificar imágenes en 1000 categorías de objetos (por ejemplo, teclado,

ratón, lápiz y muchos animales). Como resultado, la red ha aprendido representaciones ricas en características para una amplia gama de imágenes. El tamaño de la entrada de imagen de la red es de 224 por 224.

2.1.13. MobilenetV2

Esta arquitectura se basa en el uso de convoluciones separables en profundidad y utiliza una estructura residual invertida, donde la entrada y la salida del bloque residual son capas delgadas de cuello de botella opuestas a los modelos residuales tradicionales que utilizan representaciones expandidas en la entrada. Los bloques de cuello de botella parecen similares a los bloques residuales, donde cada bloque contiene una entrada seguida de varios cuellos de botella y luego una expansión como se muestra en la siguiente figura:

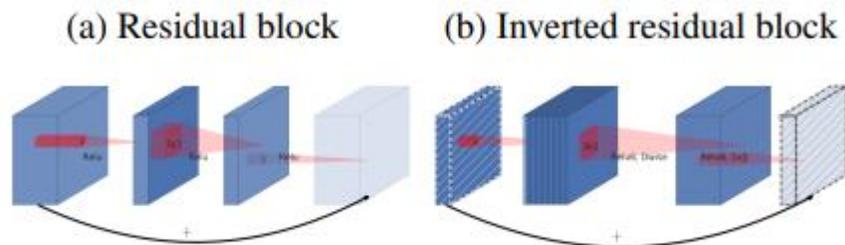


Fig. 2. Diferencia entre bloque residual y residual invertido [20].

2.1.14. Gradiente descendente

El gradiente descendente se usa en redes neuronales para la actualización de pesos al usar un subconjunto del porcentaje definido del dataset de entrenamiento, al usarlo la convergencia del error mínimo es más exacto, pero tarda más en darse, sin embargo, requiere de todo el bloque del *minibatch* (pequeño lote de entrenamiento) [15].

2.1.15. Backpropagation

La propagación inversa permite estimar el error en la actualización de los pesos en las capas ocultas, pues al no saber el valor de la capa oculta no es posible calcular el error en las neuronas específicas por lo que se vuelve imposible hallar una conexión para modificación de los pesos, al aplicar la propagación hacia atrás de los errores (*backpropagation*), se usa la estimación del error en las expresiones de las capas siguientes para encontrar el error en la propia, para ello se usa la estimación del error de la capa de salida hacia el interior de la capa oculta la cual se puede calcular para hallar la estimación en las capas ocultas [15].

2.1.16. Desvanecimiento del gradiente

Cuando se calcula el gradiente al aplicar *backpropagation* a medida en que se profundiza en la red este puede hacerse más pequeño y eventualmente se tendrán capas donde será cercano a cero lo que impide una conexión en las capas debido a que el ajuste de los pesos depende en gran medida al descenso del gradiente [15]. Lo que produce que el modelo entrenado se estanque en mínimos locales y no pueda seguir corrigiendo el error al no continuar modificando los pesos.

2.1.17. Función de activación

La función de activación determina el procesamiento no lineal de la capa de una neurona esto garantiza que la red neuronal no se sature, las funciones más utilizadas son lineal, escalón, sigmoide, RELU y tangente hiperbólica; La más común en redes neuronales con capas profundas cuando se usan algoritmos de optimización como el descenso del gradiente es la función RELU, para el procesamiento no lineal de capas ocultas y hallar los

pesos óptimos de la red, también al momento de emplear el algoritmo del backpropagation [21].

2.1.18. Validación cruzada

La validación cruzada consiste en dividir aleatoriamente el conjunto de datos en k elementos o pliegues (folds), en donde $k-1$ son los pliegues para realizar las pruebas del modelo y el restante para el entrenamiento; el número de pliegues generado será igual al número de iteraciones definidas [21], la principal ventaja de esta técnica es que permite todo el conjunto de datos en la etapa de entrenamiento.

2.1.19. Sobre entrenamiento

El sobre entrenamiento se refiere al estado en donde el modelo de red no puede responder adecuadamente a la validación cuando se usan datos diferentes a los de entrenamiento, esto sucede cuando el error es mínimo o cercano a cero y llega a un punto de memorización, lo que trae como consecuencia que el error de prueba sea mayor que el de entrenamiento [15].

2.1.20. Sub-entrenamiento

En este tipo de error de entrenamiento en donde al aplicar la técnica de regularización para mitigar el error de sobre entrenamiento puede suceder lo contrario en donde la salida trata de ajustarse a los data de entrada a través de un modelo lineal [15].

2.1.21. Regularización por parada temprana

Es una técnica en donde surge un nuevo conjunto de datos para validación de la red el cual, se tiene en cuenta para el entrenamiento del modelo, pero no para la modificación de los pesos, al aplicar esta técnica surgen dos tipos de errores de aprendizaje y de

validación, lo que sugiere la técnica es parar el entrenamiento del modelo en el momento en que el error de validación empiece una curva ascendente asegurando un punto mínimo (López, 2021).

2.2. Marco contextual

El café es un producto que ha presentado muchas variaciones en la historia de Colombia, por ejemplo, Cano et al., (2012) describe la caída del mercado del café a nivel nacional tanto en estancamiento de producción como de las exportaciones que decayeron en el periodo 1990 - 2011, el autor explica que se debió al crecimiento económico de otros sectores y que la producción del café pergamino pasó de representar el 25% al 6% del PIB en Colombia. Otro factor que influye en estas variaciones son las cooperativas caficultoras, las cuales según Vargas prieto & Castañeda Guzmán (2018), juegan un rol importante en la producción y exportación de café ya que aseguran la compra del café a los productores para su exportación, generan trabajo, inyectan capital, impulsan el desarrollo tecnológico, ofrecen posibilidades de créditos a los caficultores, e invierten en educación.

2.3. Normas técnicas del café.

En la resolución 02 de 2016 de la federación nacional de cafeteros tratan las normas técnicas para obtener un café verde almendra de calidad para exportación a partir de los siguientes criterios:

Debe ser un grano grande, retenido en malla 14, con tolerancia hasta del 5 % inferior a la misma malla, pero de este porcentaje, retenido mínimo en malla 12; con al menos 50 % de granos retenidos sobre la malla 15. Adicionalmente se deben tener en cuenta los siguientes criterios:

Humedad: no mayor al 12.5% en una muestra superior a 400g

Defectos: Al analizar una muestra de 500 gramos de café verde almendra, se debe tener menos de 24 de las faltas presentadas en la Tabla 1.

Tabla 1. Defectos de los granos de café; tomada de [22].

Grano Defectuoso	Falta
Granos Grupo I	
Un(1) Grano Negro	1.0
Un(1) Grano Vinagre	1.0
Un(1) Grano Reposado	1.0
Dos (2) Granos Parcialmente Negro	1.0
Dos (2) Granos Parcialmente Vinagre	1.0
Un(1) Grano Cardenillo	1.0
Dos (2) Granos Parcialmente Cardenillo	1.0
Cinco (5) Granos Ámbar o Mantequilla	1.0
Granos Grupo II	
Cinco (5) Granos Mordido y Cortado	1.0
Cinco (5) Granos Picado por Insecto	1.0
Diez (10) Ligeramente Picado por Broca	1.0
Cinco (5) Granos Malformado Averdaneado	1.0
Cinco (5) Granos Inmaduro	1.0
Cinco (5) Granos Aplastado	1.0
Cinco (5) Granos Decolorado Veteado	1.0
Cinco (5) Granos Puntiagudo	1.0
Cinco (5) Granos Flojos	1.0
Cinco (5) Granos Sobresecado	1.0

Infestación: el café debe estar libre de todo insecto vivo y en la eventualidad de tener perforaciones por insectos se aplicarán los procedimientos de tolerancias descritos por la FDA o la ISO en caso de ir a mercado norteamericano.

Olor: olor característico del tipo de café verde arábigo, sin contaminación.

Color: apariencia uniforme, si presenta mezclas de colores no se autoriza como tipo exportación.

Prueba en Taza: sabor y aroma característicos, debe estar libre de sabores defectuosos como el fermento, químicos, moho o reposo. (Federación Nacional de Cafeteros de Colombia)

Como se puede apreciar en (Federación Nacional de Cafeteros Colombia) La adecuada selección de los granos de café maduros es un proceso importante en la disminución de los defectos mencionados en la Tabla 1, de los cuales puede tener únicamente 24 penalizaciones ya que reduce la probabilidad de obtener granos inmaduros o paleteados, pintones, negros y mordido que es producto del beneficiado de granos verdes.

Capítulo 3. Materiales y Métodos

3.1. Materiales

3.1.1. Hardware para la adquisición de imágenes

Para la adquisición de imágenes se usaron cuatro dispositivos smartpone de diferentes características:

- Motorola G6 de 13 + 5 Mpx.
- Iphone 12 con modalidad de cámara lenta de 12 Mpx.
- Iphone 13 con modalidad de cámara lenta de 12 Mpx.
- Samsung Galaxy S23 modalidad cámara a 12 Mpx.

3.1.2. Hardware para procesamiento de imágenes y desarrollo de la red neuronal

Para realizar el entrenamiento de los modelos de clasificación, se utilizó un equipo con las siguientes características de hardware:

- Procesador Ryzen 7 3700X
- Memoria RAM de 32GB marca
- Disco duro de 500GB NVMe
- Placa madre GYGABYTE AOURUS B450 WIFI
- Tarjeta gráfica marca ZOTAC RTX 3090 24GB

Sistema operativo Ubuntu 22.0.4, este sistema facilita la compatibilidad de las librerías para entrenamiento del modelo de red con la GPU usada, en el momento que se empezó a desarrollar el código.

3.2. Métodos

Tal como se puede apreciar en la Fig. 3, para el entrenamiento de los modelos, se utilizó una metodología secuencial, lineal, determinando las fases de: construcción del conjunto de datos, preparación el conjunto de datos, entrenamiento de diferentes modelos de redes neuronales convolucionales y validación de la efectividad del modelo.

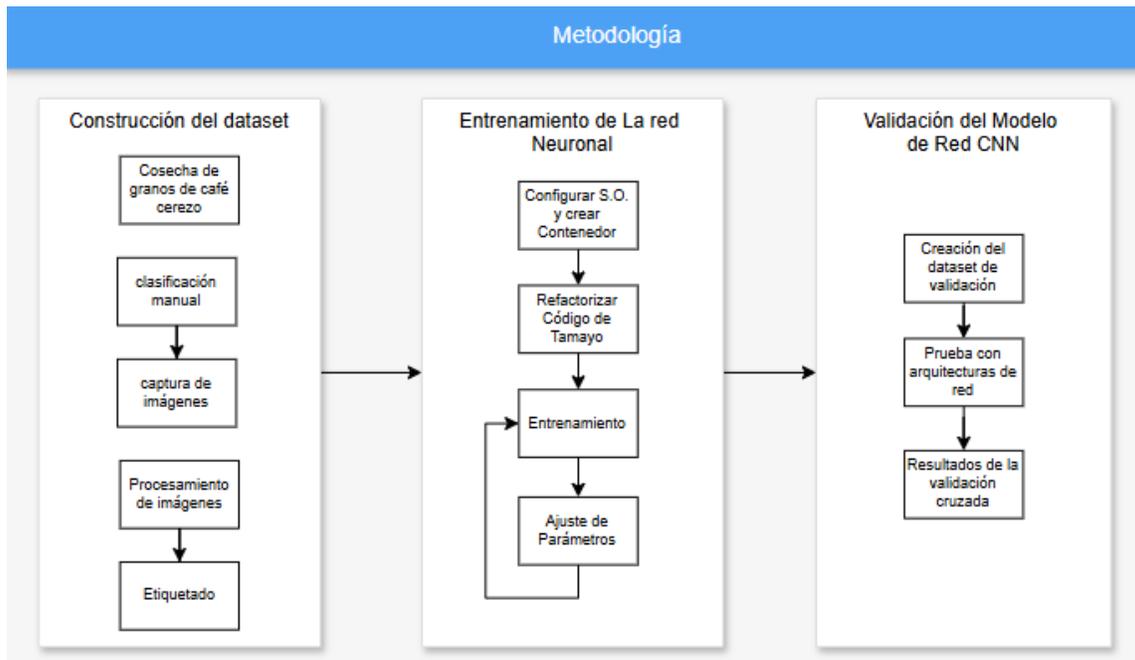


Fig. 3. Metodología secuencial.

3.2.1. Construcción de los conjuntos de datos de entrenamiento

Se crearon dos conjuntos de datos con imágenes de granos de café cerezo en sus diferentes estados de maduración (verde, pintón, maduro, sobre maduro y seco) obtenidos de una finca del municipio de Cajibío en el departamento del Cauca, el primer conjunto de datos se construyó con el dispositivo smartphone Motorola G6, el segundo conjunto es

construido a partir de imágenes que fueron tomadas con un smartphone Samsung Galaxy S23 en modalidad cámara a 12Mpx ambos dispositivos mencionados en el apartado 3.1.1.

Las imágenes son capturadas en entornos de iluminación controlados y no controlados, una vez se obtuvo el banco de imágenes se realizó el tratamiento y redimensionamiento. El segundo conjunto de datos fue obtenido a través de vídeos capturados utilizando un Smartphone Iphone 12 y un banco de pruebas (ver Anexo A1), para facilitar la toma de imágenes en movimiento capturando el video como se puede apreciar en la *Fig. 13*, posteriormente el video es procesado mediante un algoritmo de reconocimiento de objetos extrayendo frames cada segundo (*Fig. 5*), las imágenes extraídas son recortadas y redimensionadas con una resolución de 224x224 a 3 canales RGB.

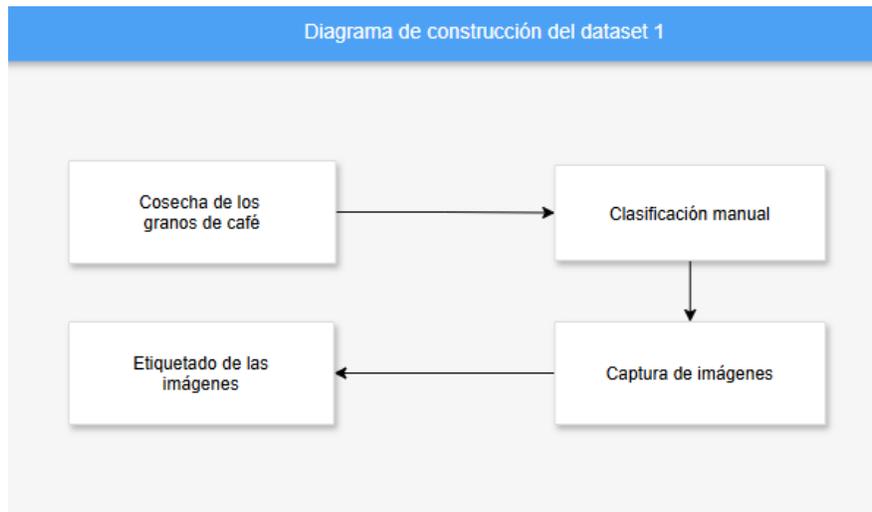


Fig. 4. Diagrama de flujo de la construcción del dataset 1.

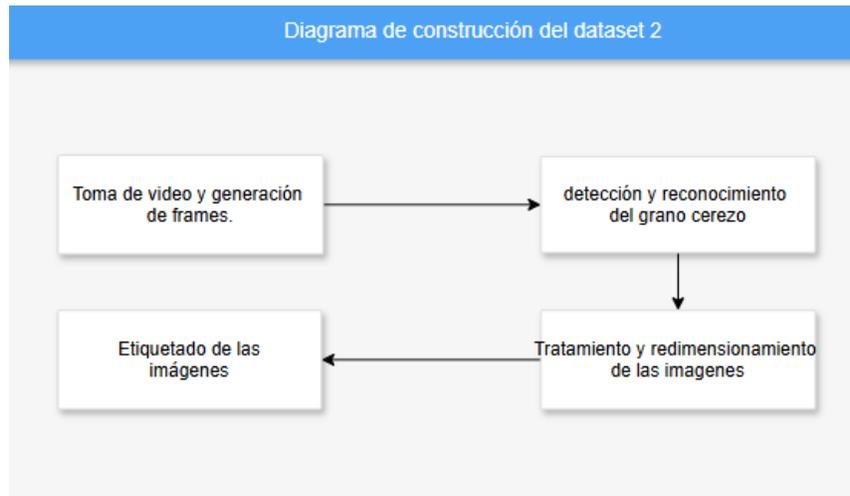


Fig. 5. Diagrama de flujo de la construcción del dataset 2

Muestra de los frames que se extraen de los videos en el banco de pruebas, con iluminación controlada (Fig. 6).

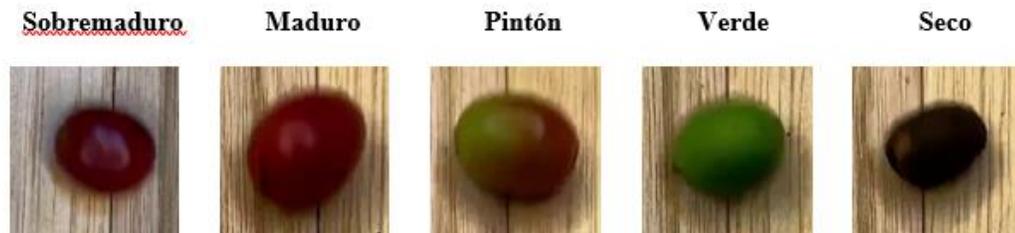


Fig. 6. Imágenes de granos de café cortadas y etiquetadas.

Después de crear el banco de imágenes con los frames generados del video, estas son llamadas y tratadas mediante el código de procesamiento de imágenes desarrollado en el software “Visual Studio Code” (Fig. 7).

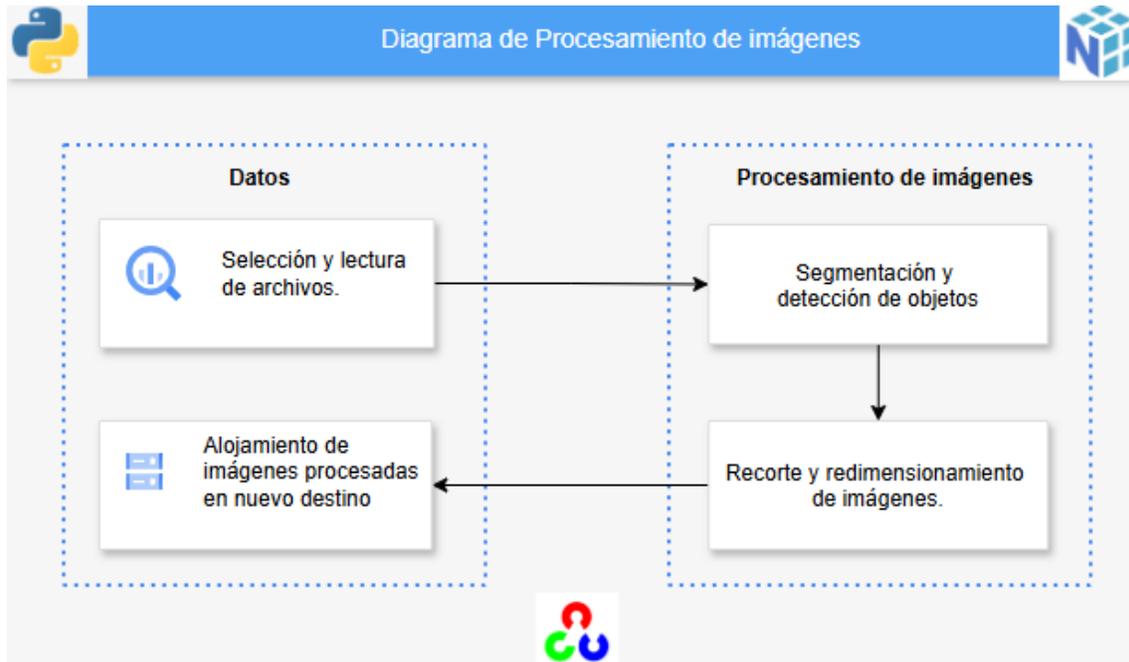


Fig. 7. Diagrama de Procesamiento de imágenes.

3.2.2. Experimentación con los modelos y conjuntos de datos de Tamayo

Inicialmente se realizó una investigación sobre las arquitecturas de red neuronal utilizadas por otros autores para resolver problemas de clasificación similares, entre los cuales destaca el de Tamayo [25] donde utilizaron diferentes modelos de red neuronal para comparar el uso de transfer Learning de los modelos, frente al no uso del mismo; también se modificó el conjunto de datos inicial para comparar la eficiencia de los modelos resultantes al entrenar con el conjunto de datos desbalanceado, comparándolo con el conjunto de datos balanceado al utilizar aumento de datos.

La experimentación se llevó a cabo sobre el conjunto de datos desbalanceado debido a que tal como expreso Tamayo es el dataset que reportó mejores resultados utilizando Transfer Learning y validación cruzada. En la Fig. 8 se resume el proceso realizado, donde se inicia con la configuración del entorno de trabajo y la creación del

contenedor, posteriormente, se incluyó el código de Tamayo en el contenedor, junto con las librerías necesarias para el entrenamiento, también se cargó el conjunto de datos de Tamayo y finalmente se realizó el entrenamiento de todos los modelos propuestos almacenando sus resultados para crear tablas y gráficas que se presentan en la sección 4.2.

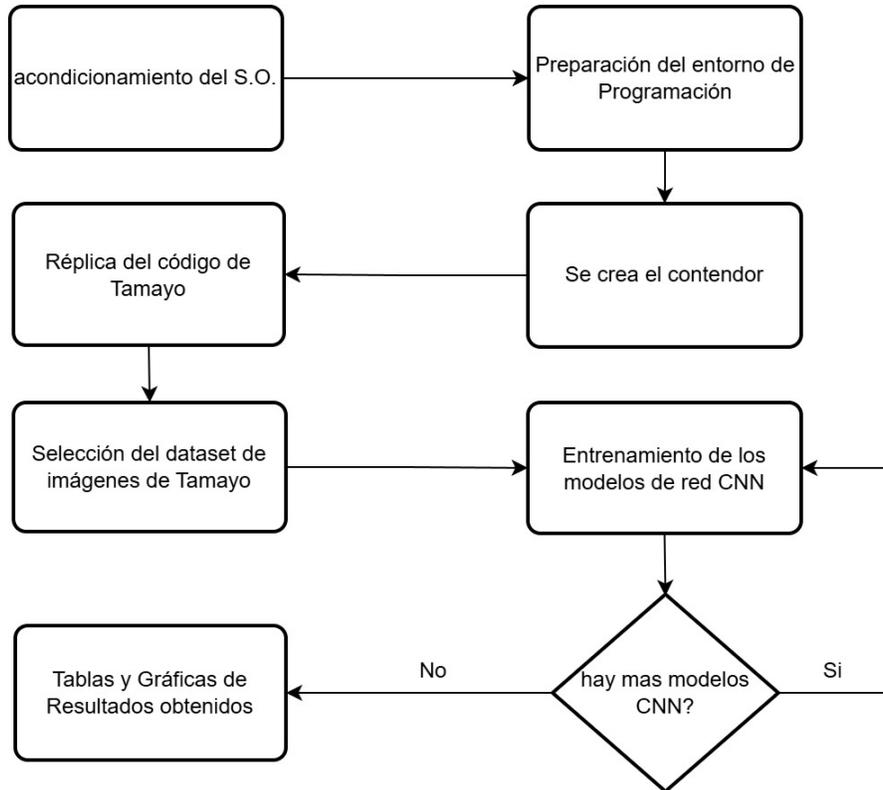


Fig. 8. Diagrama de flujo del proceso realizado para replicar los experimentos de Tamayo.

3.2.3. Entrenamiento de una Red Neuronal Convolutacional

Para realizar el entrenamiento, se siguió los pasos descritos en la Fig. 9. Primero, se configuró el sistema en contenedores de Docker que facilitaron la experimentación, el despliegue y configuración de la aplicación al utilizar la imagen de TensorFlow-GPU configurada y disponible en la página de Docker Hub; dentro de esta configuración, se

añadieron algunas automatizaciones para la instalación de paquetes de Ubuntu en el contenedor, instalación de paquetes de Python, copia del proyecto dentro del contenedor y creación de volúmenes de Docker.

Después, se utilizó como base el código publicado por Tamayo con varias modificaciones y mejoras como la modificación de su arquitectura monolítica a una arquitectura orientada a microservicios con el fin de mejorar su escalabilidad y facilitar su implementación en sistemas orientados a la web; también se incluyó la librería de Flask en Python con el fin de exponer estos microservicios en una red local y el cliente Postman para ejecutar los diferentes módulos del sistema; las demás modificaciones, junto con el funcionamiento de cada microservicio se exponen en el Anexo A3.

Con respecto al entrenamiento, este se dividió dos fases, en la primera fase de entrenamiento se dividió el conjunto de datos en 70% para entrenamiento, 20% para pruebas y 10% para validar el modelo posterior al entrenamiento; con lo anterior, se realizó el ajuste de los parámetros y en la segunda etapa se realizó la validación cruzada la cual, como se habló en la sección 2.1.18. permite utilizar todo el conjunto de datos para entrenar, lo cual aporta calidad al modelo con la desventaja de que requiere una mayor capacidad computacional.

Una vez configurado el sistema y ajustado el código, se realizó el entrenamiento de diferentes CNN, utilizando como base los 5 modelos propuestos por Tamayo (VGG16, VGG19, Inception-ResNetV2, InceptionV3 y DenseNet201); adicional a los modelos propuestos por Tamayo, se incluyó el modelo MobileNetV2 por ser el más liviano de la librería keras applications y el que menor tiempo requiere para hacer inferencias según la

documentación de Keras, disponible en: <https://keras.io/api/applications/>. Dentro de la configuración de los modelos, posterior al uso de las CNN, se utilizó una capa de aplanamiento y dos capas densas acompañadas de capas de regularización (Relu y Batch Normalization), en la salida del modelo, se utilizó una capa de 5 neuronas que corresponden a las clases que a su vez son las etapas de maduración de café y la función de activación Sotmax que permite obtener por cada neurona un valor, siendo el valor más alto el que refleja la clase que el modelo que predice.

Posterior al entrenamiento, se ajustaron los parámetros de entrenamiento teniendo en cuenta los resultados obtenidos, y luego, se repitió la etapa de entrenamiento hasta que se obtuvieron los resultados que se presentan en el Capítulo 4.

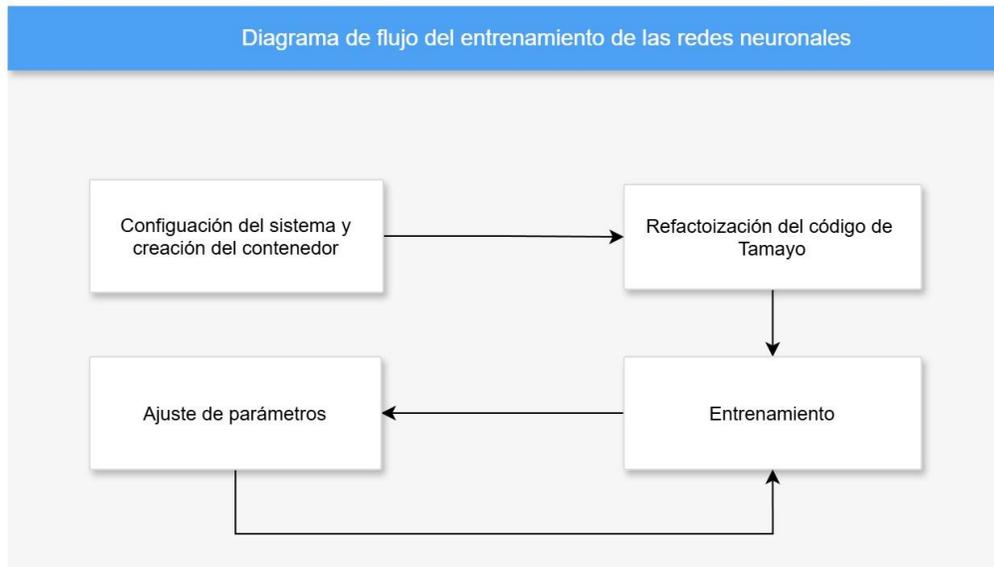


Fig. 9. Diagrama de flujo del entrenamiento de las redes neuronales.

3.2.4. Evaluación de los modelos

Inicialmente se construyó otros conjuntos de datos denominados 3_1 y 4_1 utilizando la misma metodología que la planteada para el conjunto de datos construido en la sección 4.1. construcción de los conjuntos de datos; lo anterior con el fin de evaluar los modelos entrenados utilizando nuevos datos. Es importante resaltar que se entrenaron 6 modelos por cada tipo de conjunto de datos (dataset 1 y dataset 2), de esta forma se utilizaron los nuevos conjuntos de datos para evaluar los modelos de la siguiente forma: dataset 3_1 para evaluar los modelos producto del dataset 1 y dataset 4 para evaluar los del dataset 2.

La evaluación de los modelos se realizó teniendo en cuenta el microservicio que evalúa todos los modelos (ver Anexo A5). A partir del microservicio mencionado, se evaluó todos los modelos con los nuevos conjuntos de datos creados en esta etapa.

Capítulo 4. Resultados

Teniendo en cuenta la metodología presentada en el capítulo 3, se obtuvieron los siguientes resultados:

4.1. Construcción del conjunto de datos

Tal como lo manifiesta en el apartado (3.2.1. el primer dataset es adquirido por el proceso de cosecha manual recolectando el fruto del árbol y vertiéndolo en un recipiente, este proceso se llevó a cabo en una finca en el municipio de Cajibío del departamento del Cauca, una vez es cosechado el fruto se procede a realizar la capturas de las imágenes de los granos en diferentes estados de maduración con un dispositivo smartphone Motorola G6 de 13Mpx, la captura de las imágenes se realizó en torno de iluminación controlada y luz natural, para la iluminación controlada se usó una mesa cuadrada de superficie plana de color blanco y un aro de luz con soporte para dispositivo móvil a una altura de 20 cm sobre el grano de esta forma se capturaron las imágenes con claridad Fig. 10.



Fig. 10. Imágenes de granos de café redimensionadas con iluminación controlada.

El segundo dataset es construido a partir de los frames de los videos capturados mostrando el grano cerezo de café en movimiento, pues mediante el código se genera la

captura de las imágenes cada vez que pasa el grano por la cámara, posterior a ello se realizar el procesamiento de las imágenes utilizando las librerías de OpenCV y Numpy junto con el lenguaje de programación Python en el software Visual Estudio Code. El código se encarga de realizar la difuminación de la imagen a partir de un filtros gaussianos para encontrar los contornos, luego se utiliza la transformación morfológica opening que consiste en reducir el ruido de la imagen y la operación findContours para ubicar el grano de café (Fig. 11); una vez ubicado el grano de café, se recorta la imagen a un tamaño de 300x300 alrededor del grano (Fig. 12), y finalmente se etiquetan las imágenes por clases y se guardan en carpetas que llevan el nombre de su tipo Fig. 6.

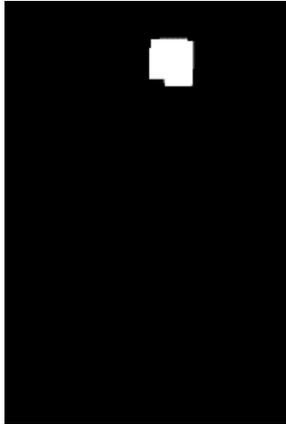


Fig. 11. Conversión de imagen a blanco y negro.

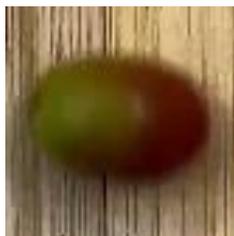


Fig. 12. Recorte de imagen a escala 224x224.



Fig. 13. Captura el grano de café en movimiento.

Según lo descrito anteriormente, se obtuvieron dos conjuntos de datos que se componen de 5 clases y se caracterizan a continuación:

Tabla 2. Número de registros de los conjuntos de datos construidos.

Cantidad de Imágenes por Dataset		
Clases	Dataset 1	Dataset 2
Verdes	265	76
Pintones	259	131
Maduros	271	277
Sobremaduros	160	99
Secos	372	51

4.2. Experimentación con modelos y conjuntos de datos de Tamayo

A partir del repositorio publicado por Tamayo [24], se reconstruyeron los experimentos con las diferentes arquitecturas de red que implementaron, entre las cuales se encuentran VGG16, VGG19, Inception-ResNet-V2, Inception V3, DenseNet201, implementando un contenedor de docker para simular los resultados de las redes antes mencionadas, sin embargo se encontró que los resultados entre ejecuciones no son los mismos debido a la ausencia de una semilla la cual permite una reproducción exacta del entrenamiento del modelo, a continuación se presenta el análisis de las arquitecturas VGG16 y VGG19 utilizando transfer Learning y el dataset desbalanceado que cuenta con 640 registros y cada registro 15 canales (ver Tabla 3); un resumen de todos los resultados se presenta en la Tabla 4.

Una vez terminado el entrenamiento del modelo se evaluaron las métricas, en la parte inferior de la Fig. 14 se observa la matriz de confusión, Tamayo utilizó aplicar esta métrica, junto con la precisión, sensibilidad, exactitud y puntaje F1 para verificar su ajuste con el dataset, en la matriz de confusión se puede notar la distribución de los valores verdaderos y las predicciones, es decir los errores y aciertos del algoritmo que se ha entrenado, la matriz es un arreglo de 5x5 que refleja la cantidad de granos de café que el modelo encontró como cierta clase, vs la clase a la que corresponde realmente; en el anexo A2 se presentan las fórmulas que se debe utilizar para hallar las métricas de evaluación.

A pesar de los resultados presentados en la Fig. 14, si se analiza el resultado después de aplicar validación cruzada Fig. 15, se evidencia que las métricas obtenidas fueron

inferiores, después de analizar el código en profundidad, se encontró que el modelo utilizado durante la validación cruzada era diferente y este puede ser uno de los factores.

Tabla 3. Descripción del dataset desbalanceado de Tamayo.

EXPERIMENTO 1	DATASET
	DESBALANCEADO
	(448,224,224,15)
	(192,224,224,15)
	(48,224,224,15)
	(192,224,224,15)

```
[0.25303220748901367, 0.9479166865348816]
(192,)
(192,)
Accuracy: 0.9479166666666666
F1 score: 0.9492183113384709
Recall: 0.9474749373433584
Precision: 0.9531346431346431

classification report:
      precision    recall  f1-score   support

0         1.00      0.92      0.96         24
1         0.91      0.98      0.94         42
2         0.98      0.90      0.93         48
3         0.95      0.97      0.96         38
4         0.93      0.97      0.95         40

 accuracy          0.95
macro avg          0.95
weighted avg       0.95

confusion matrix:
[[22  0  0  1  1]
 [ 0 41  0  0  1]
 [ 0  4 43  1  0]
 [ 0  0  0 37  1]
 [ 0  0  1  0 39]]
```

Fig. 14. Métricas de evaluación de la arquitectura VGG16 de Tamayo.

```

Training for fold 10 ...
Score for fold 19: loss of 0.6284974813461304; accuracy of 89.0625%
(64,)
2/2 [=====] - 0s 41ms/step
acc 0.6703125
acc_SD 0.20341925801469732
f1-score 0.6285547289844509
f1_SD 0.2719285438846651
recall 0.6741063199420733
recall_SD 0.2914179929377961

```

Fig. 15. Métricas de evaluación de la arquitectura VGG16 de Tamayo con validación cruzada.

Por otra parte, el entrenamiento del modelo de la red VGG19 dio como resultado una precisión de 0.95, recall de 0.94, F1-Score de 0.95 y accuracy de 0.95 como se puede apreciar en la Fig. 16; como se puede apreciar en la Fig. 17, un fenómeno similar al evidenciado para VGG16, sucede con VGG19 al utilizar validación cruzada.

```

=====
Total params: 26,489,157
Trainable params: 26,488,389
Non-trainable params: 768
-----
[0.2580578029155731, 0.9479166865348816]
(192,)
(192,)
Accuracy: 0.9479166666666666
F1 score: 0.9473931042933639
Recall: 0.944141604010025
Precision: 0.9534552845528456

classification report:
      precision    recall  f1-score   support

0         1.00      0.88      0.93         24
1         0.91      0.98      0.94         42
2         0.96      0.90      0.92         48
3         0.93      0.97      0.95         38
4         0.98      1.00      0.99         40

 accuracy
macro avg      0.95      0.94      0.95         192
weighted avg   0.95      0.95      0.95         192

confusion matrix:
[[21  0  0  3  0]
 [ 0 41  1  0  0]
 [ 0  4 43  0  1]
 [ 0  0  1 37  0]
 [ 0  0  0  0 40]]

```

Fig. 16. Métricas de evaluación de la arquitectura VGG19.

```

Training for fold 10 ...
Score for fold 15: loss of 2.397688388824463; accuracy of 59.375%
(64,)
2/2 [=====] - 0s 47ms/step
-----
Training for fold 10 ...
Score for fold 17: loss of 0.05221624672412872; accuracy of 98.4375%
(64,)
2/2 [=====] - 0s 42ms/step
-----
Training for fold 10 ...
Score for fold 19: loss of 0.8602582216262817; accuracy of 87.5%
(64,)
2/2 [=====] - 0s 43ms/step
acc 0.6984375
acc_SD 0.25175652455547204
f1-score 0.6557479489711409
f1_SD 0.2916838473406756
recall 0.7117214342076292
recall_SD 0.27255509098339037

```

Fig. 17. Resultados de la validación cruzada del modelo VGG19.

Si se compara las métricas de evaluación de la matriz de confusión obtenidas al reproducir el entrenamiento de las arquitecturas de red VGG16 y VGG19 propuestas por Tamayo, se obtuvo que la exactitud de ambas llega a un 95% en comparación con los resultados de Tamayo que dan 84.79% y 92.39% respectivamente, esta diferencia se debe en gran parte a la falta de semillas para garantizar la reproducibilidad de los experimentos.

Después de realizar los entrenamientos con validación cruzada de todos los modelos, se encontraron los resultados presentados en la Tabla 4, donde se evidencia que el mejor modelo fue Inception V2 con un 99% en precisión, sensibilidad y puntaje F1; estos resultados también se presentan de manera gráfica en la Fig. 18.

Tabla 4. Métricas de evaluación obtenidas de replicar los experimentos de Tamayo con validación cruzada.

Model	Accuracy	F1 Score	Recall
VGG16	67%	63%	67%
VGG19	70%	66%	71%
Inception V2	99%	99%	99%

Inception V3	97%	97%	98%
Densenet 201	98%	98%	98%

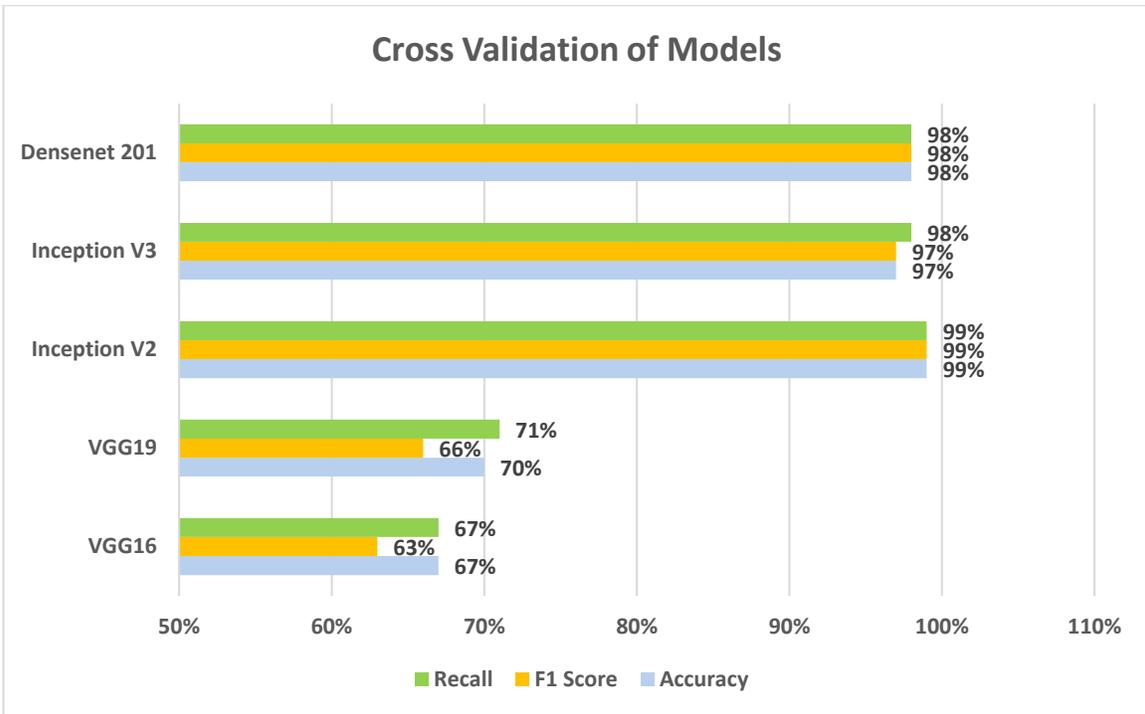


Fig. 18. Resultados de replicar los experimentos de Tamayo con transferencia de aprendizaje, validación cruzada y el conjunto de datos desbalanceado.

4.3. Entrenamiento de una Red Neuronal Convolutiva

Teniendo en cuenta lo mencionado en la sección 3.2.3. se creó un software que permite el entrenamiento, del que se destaca la aplicación de técnicas de optimización del entrenamiento y regularización como la parada temprana para evitar el sobreajuste y el decaimiento en la tasa de aprendizaje con el fin de evitar el error de decaimiento del gradiente descendiente a medida que avanzan las iteraciones del entrenamiento del modelo; el valor de decaimiento de la tasa de aprendizaje k utilizado fue de 0.03 ya que al graficar la función de decaimiento, se encontró que en la epoch 100, aún presentaba algo de ajuste. Todos los modelos se entrenaron utilizando 100 epochs en la primera etapa de entrenamiento y 20 epochs en la segunda etapa de entrenamiento que corresponde a la validación cruzada, donde se implementó 10 folds; los dataset de imágenes utilizados se componen de imágenes con una dimensión de 224 x 224, se utilizó una tasa de aprendizaje de 0.0001 y un valor k de 0.03. Se entrenó utilizando los conjuntos de datos 1 y 2 descritos en la sección 3.2.1. tal como se plantea a continuación:

4.3.1. Entrenamiento utilizando el conjunto de datos 1

Se realizó el entrenamiento de las arquitecturas antes mencionadas utilizando el conjunto de datos 1 dividido en 70% para entrenamiento, 20% para validación y el 10% para pruebas; en la Tabla 5 se presentan las métricas de evaluación en donde se puede apreciar resultados similares para todos los modelos, siendo Vgg16 y Densenet 201 los modelos que mejores resultados obtuvieron con un accuracy, precisión, recall y F1-score de 0,98. La información de estos modelos se amplía en las Figuras Fig. 19, Fig. 20, Fig. 21

y Fig. 22 ya que se graficaron utilizando todos los decimales para percibir la diferencia que existe entre los resultados.

Tabla 5. Resultados de entrenar los modelos utilizando el conjunto de datos 1 sin validación cruzada.

Arquitecturas	exactitud	precisión	sensibilidad	f1-score
Vgg16	0,98	0,98	0,98	0,98
Vgg19	0,97	0,98	0,97	0,97
InceptionResnet V2	0,94	0,95	0,94	0,95
Mobilnet V2	0,95	0,96	0,95	0,95
Densenet201	0,98	0,98	0,98	0,98
Inception V3	0,96	0,97	0,96	0,96

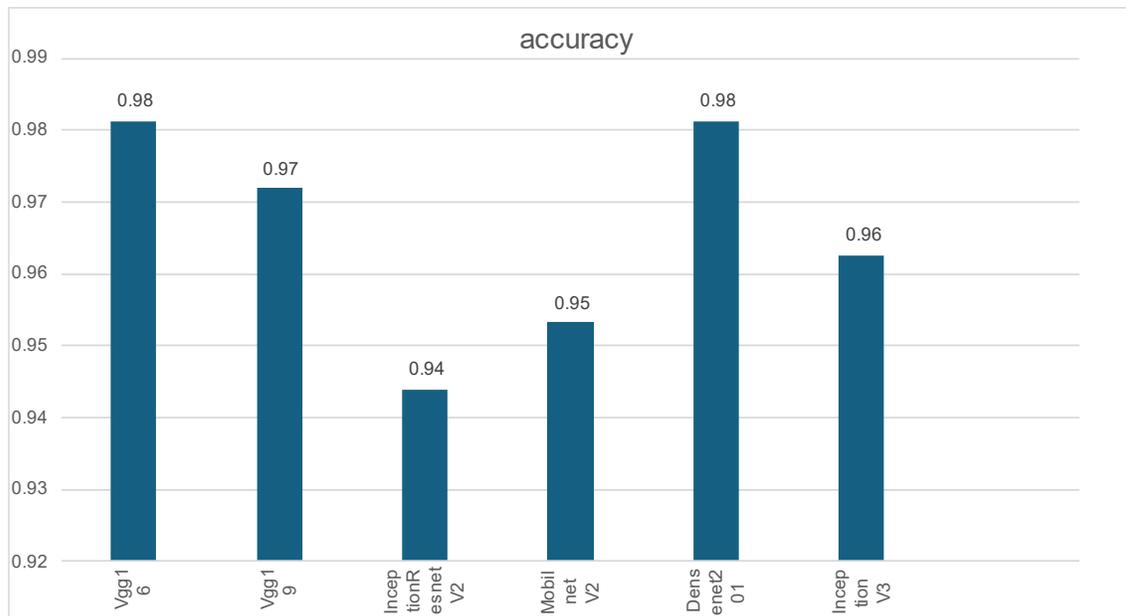


Fig. 19. Exactitud de las arquitecturas entrenadas.

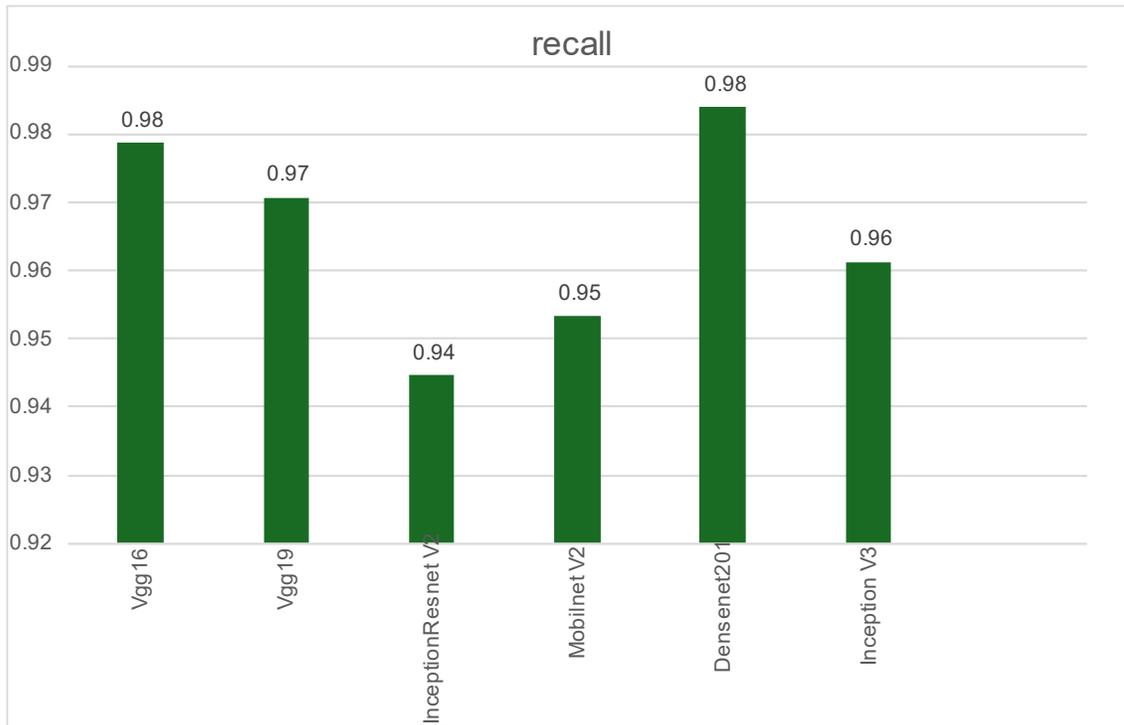


Fig. 20. Sensibilidad de las Arquitecturas entrenadas.

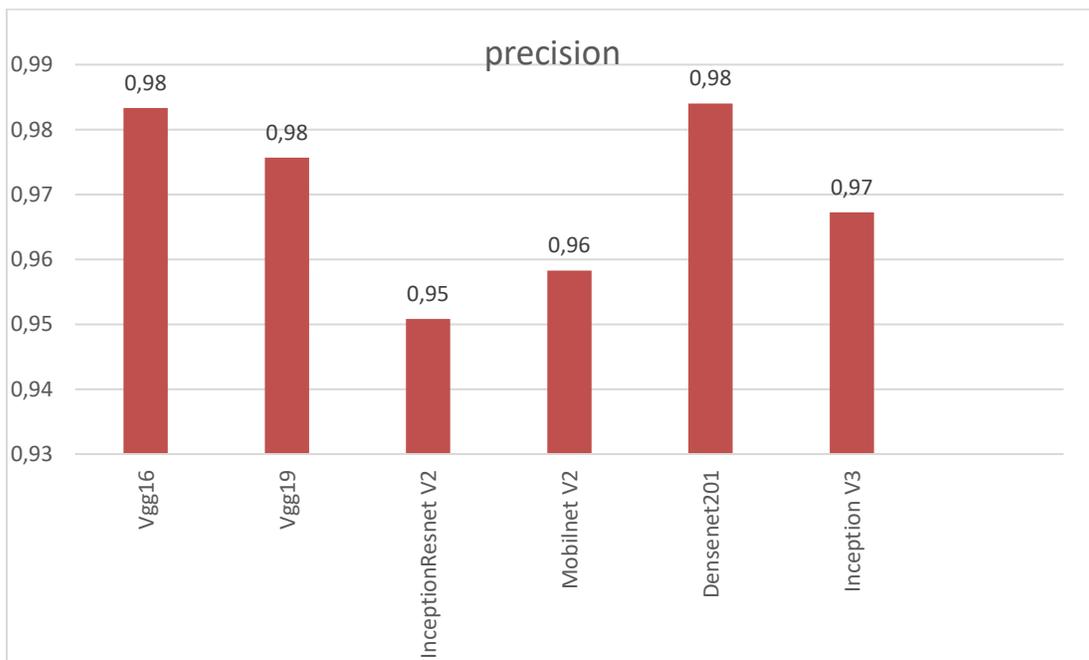


Fig. 21. Precisión de las arquitecturas entrenadas.

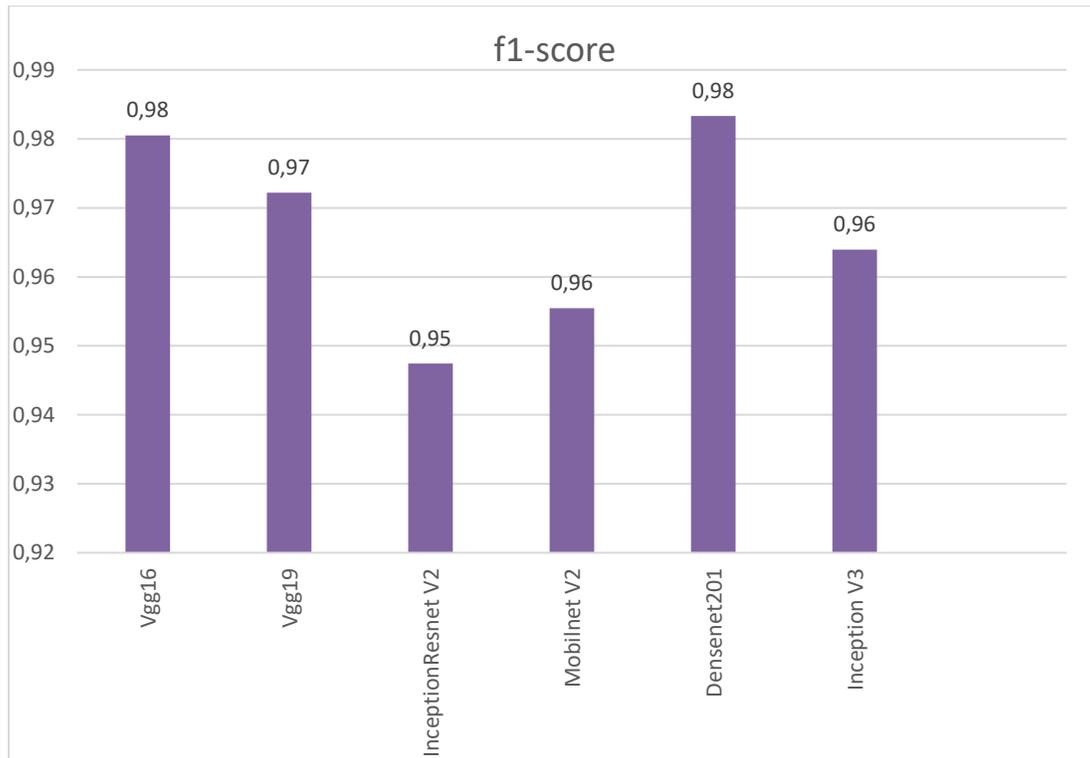


Fig. 22. F1-score medida que genera la diferencia entre precisión y sensibilidad.

Posterior al entrenamiento del conjunto de datos dividido, se utilizó el modelo resultante para entrenar de nuevo utilizando validación cruzada, en este caso, se dividió el conjunto de datos en 90% para validación cruzada y 10% para validación del modelo; como se puede apreciar en la Tabla 6, al tener en cuenta en el entrenamiento las características de una mayor cantidad de registros del conjunto de datos, los resultados obtenidos fueron más eficientes.

Tabla 6. Resultados de entrenar los modelos utilizando el conjunto de datos 1 y validación cruzada.

Arquitecturas	exactitud	precisión	sensibilidad	f1-score
Vgg16	1,00	1,00	1,00	1,00
Vgg19	1,00	1,00	1,00	1,00
InceptionResnet V2	0,99	0,99	0,99	0,99
Mobilnet V2	1,00	1,00	1,00	1,00
Densenet201	1,00	1,00	1,00	1,00
Inception V3	1,00	1,00	1,00	1,00

Como se puede observar, todas excepto InceptionResnetV2 clasificaron el 100% las clases de manera satisfactoria.

4.3.2. Entrenamiento utilizando conjunto de datos 2

Después de obtener los resultados de entrenamiento para el conjunto de datos 1, se utilizó el conjunto de datos 2 el cual se compone de imágenes ubicadas en una superficie plana de color blanco y tomadas con diferentes fuentes de iluminación como iluminación natural e iluminación artificial; se repitió un proceso similar al descrito en la sección 4.3.1. El cual entrenó utilizando el conjunto de datos dividido en 70% para entrenamiento, 20% para validación y 10% para prueba; posteriormente se entrenó la red utilizando validación cruzada, dejando solo un 10% del conjunto de datos para validar los resultados y obtener las métricas que se describen en la Tabla 7, donde se evidencian resultados más eficientes que los presentados en la Tabla 5; esto puede ser porque el experimento corresponde a un conjunto de datos menos complejo debido a que las imágenes tomadas para el conjunto de datos 2 no tenían movimiento.

Tabla 7. Resultados de entrenar los modelos utilizando el conjunto de datos 2, sin validación cruzada.

Arquitecturas	exactitud	precisión	sensibilidad	f1-score
Vgg16	1.00	1.00	1.00	1.00
Vgg19	1.00	1.00	1.00	1.00
InceptionResnet V2	0.94	0.93	0.95	0.94
Mobilnet V2	0.88	0.93	0.89	0.90
Densenet201	0.96	0.98	0.97	0.97
Inception V3	0.90	0.91	0.92	0.91

Al entrenar utilizando validación cruzada y el conjunto de datos 2, se encontró que todas las arquitecturas obtuvieron una precisión, sensibilidad, exactitud y F1-score del 100%.

Tabla 8. Entrenamiento de las arquitecturas, utilizando el conjunto de datos 2 y validación cruzada.

Arquitecturas	Accuracy	Precision	Recall	F1-score
Vgg16	1.00	1.00	1.00	1.00
Vgg19	1.00	1.00	1.00	1.00
InceptionResnet V2	1.00	1.00	1.00	1.00
Mobilnet V2	1.00	1.00	1.00	1.00
Densenet201	1.00	1.00	1.00	1.00
Inception V3	1.00	1.00	1.00	1.00

4.4. Evaluación de los modelos

Para realizar la evaluación de los modelos entrenados anteriormente se construyeron dos nuevos conjuntos de datos denominados dataset 3_1 y dataset 4_1; para su construcción, se utilizaron granos de café cereza procedentes del municipio de Cajibío en el departamento del Cauca; tal como los conjuntos de datos presentados en la sección 4.1. los dataset 3_1 y 4_1 se componen por 5 clases que corresponden a los grados de maduración del café. El 1% de estos conjuntos de datos fue incluido en el entrenamiento

planteado en la sección 4.3. Lo anterior, teniendo en cuenta que se utilizaron dispositivos diferentes a los utilizados en la construcción de los conjuntos de datos de entrenamiento; en el dataset 3_1 que se tomó con videos de granos de café cayendo, se utilizó un Iphone 13 y en el 4_1 tomado con fondo blanco y granos de café sin movimiento, un Samsung Galaxy S23 Ultra a una resolución de 12 MPx. Con respecto a la cantidad de registros del conjunto de datos, el dataset 3_1 se compone de 1525 registros y el dataset 4_1 se compone de 309 registros, como se puede apreciar en la Tabla 9.

Al evaluar los modelos entrenados a partir del dataset 1 con el dataset 3_1, donde se encuentran las imágenes en movimiento sobre un banco de pruebas las arquitecturas Vgg16 y Vgg19, se obtuvieron los resultados más altos en las métricas, tal como se puede apreciar en la Fig. 23.

Los modelos entrenados con el dataset 2, fueron evaluados a partir del dataset 4_1, se puede denotar que las arquitecturas vgg16 y vgg19 nuevamente sobresalen respecto a los otros modelos, con la arquitectura densenet201 también se obtuvieron resultados estables, pero no supera el 97% de exactitud logrado por los modelos vgg16 y vgg19 como se muestra en la Fig. 24.

De lo anterior, se encontró que a pesar de ser una de las arquitecturas más simples, la Vgg16 obtuvo los mejores resultados en la clasificación de los granos de café cereza en sus diferentes clases, esto puede deberse a que la capacidad de extracción de características de las demás arquitecturas es mayor y por tal razón, se sobre ajustan al conjunto de datos entrenado, perdiendo capacidad de generalización.

Tabla 9. Composición de los dataset 3_1 y dataset 4_1.

Dataset de validación en superficie inclinada y superficie plana		
Clases	Dataset 3_1	Dataset 4_1 cropped
Verdes	143	60
Pintones	269	45
Maduros	498	86
Sobre maduros	265	54
Secos	350	64

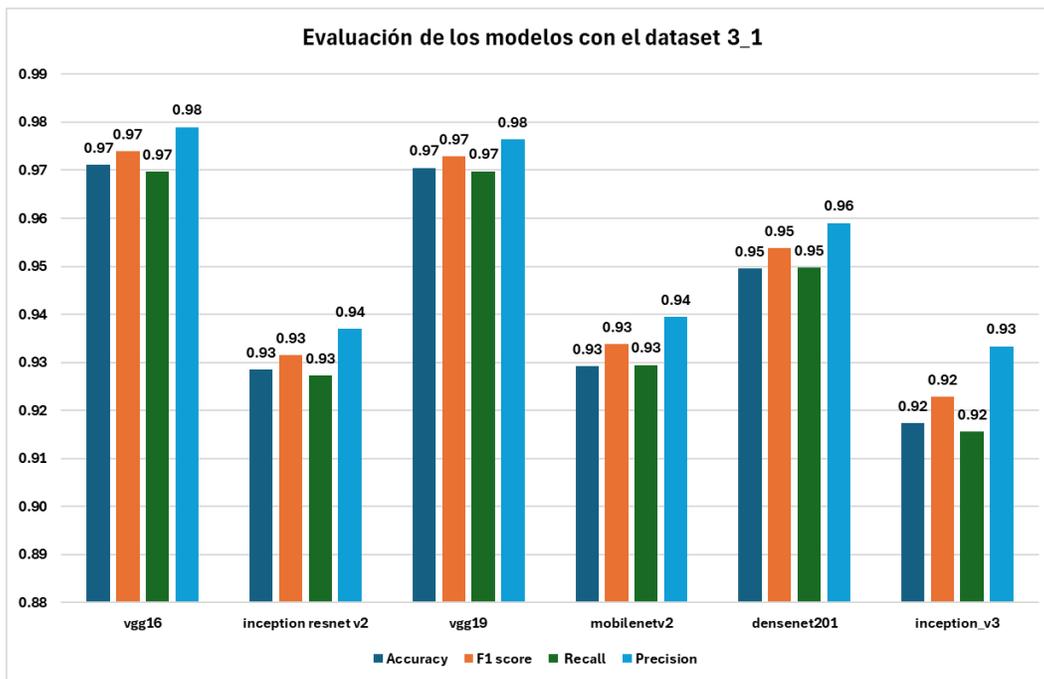


Fig. 23. Evaluación de los modelos con el dataset 3_1.

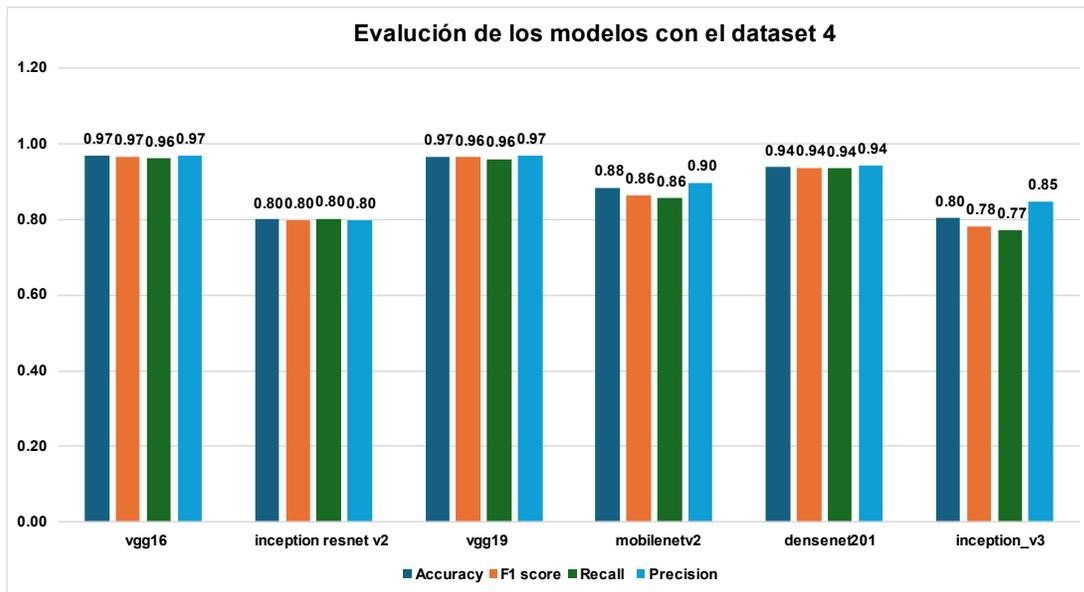


Fig. 24. Evaluación de los modelos con el dataset 4_1.

Capítulo 5. Conclusiones

La calidad y diversidad del conjunto de datos son aspectos fundamentales para el éxito de los modelos de redes neuronales convolucionales (CNN) aplicados a la clasificación del grado de maduración del café. En este proyecto, se tuvieron en cuenta dos enfoques para la captura de imágenes: la creación de un banco de pruebas que permitió tomar imágenes en movimiento y la captura de imágenes estáticas. Se encontró que el enfoque de imágenes en movimiento, aunque más eficiente en términos de tiempo y esfuerzo, puede presentar desafíos adicionales, como la necesidad de cámaras con mayor velocidad de captura. Por otro lado, la recolección de imágenes estáticas, si bien demanda más tiempo y esfuerzo, permitió un mayor control sobre las condiciones de iluminación y el encuadre de la imagen, lo que resultó en un conjunto de datos más uniforme y de alta calidad. Por tal motivo, en proyectos futuros, se debe tener en cuenta este criterio de acuerdo con la necesidad.

Después de entrenar y evaluar diferentes arquitecturas de redes neuronales convolucionales (CNN) utilizando validación cruzada, transfer learning y cuatro conjuntos de datos desbalanceados de café cereza en diferentes estados de maduración; se determinó que las arquitecturas VGG16 y VGG19 obtuvieron los mejores resultados. Estas arquitecturas destacaron gracias a su capacidad para extraer características relevantes para ajuste de pesos y entrenamiento de la red. Los modelos VGG16 y VGG19 lograron un F1-score de 0.98 y 0.99 sobre el conjunto de datos uno sin validación cruzada, el resultado que se confirmó al validar estos modelos con el conjunto de datos tres, en donde ambos modelos obtuvieron un F1-score de 0.98 y nuevamente destacaron frente a los demás. Por otra parte, al entrenar utilizando el conjunto de datos dos, los modelos VGG16 y VGG19 obtuvieron un F1-score de 1.00; resultado que se validó con el conjunto de datos número cuatro en el que se obtuvo un F1-score de 0.97 y 0.96 respectivamente. El desempeño de las arquitecturas utilizadas para entrenar los diferentes conjuntos de datos sugiere que su diseño basado en capas convolucionales profundas es adecuado para resolver el problema de clasificación de granos de café cereza por grado de maduración. En conclusión, las arquitecturas VGG16 y VGG19 fueron los modelos que mejor se ajustaron a los conjuntos de datos para la clasificación de café.

Teniendo en cuenta la experimentación realizada con la validación cruzada, se concluye que, a pesar de ser una técnica costosa computacionalmente, reduce notablemente uno de los principales problemas que pueden surgir al utilizar modelos de redes neuronales convolucionales que es el sobreajuste.

Al analizar los resultados presentados en la tabla 6 donde todos los modelos arrojaron resultados entre 0.99 y 1.00 con el conjunto de datos utilizado para entrenar y la fig. 23 con el conjunto de datos creado en diferentes épocas del año y utilizado para validar; se resalta la importancia de validar los resultados de los modelos con conjuntos de datos diferentes a los utilizados para entrenar en proyectos relacionados con clasificación; especialmente en el contexto del fruto del café, donde las imágenes pueden presentar diferentes tonos de acuerdo con la iluminación. Lo anterior garantiza un mayor grado de generalización por parte del modelo ya que los modelos se pueden ajustar correctamente a un conjunto de datos específico, pero en la práctica, presenta errores cuando varían ligeramente las condiciones del entorno.

En diferentes partes del mundo se han realizado investigaciones relacionadas con clasificación de granos de café, sin embargo, estas investigaciones utilizan cámaras multiespectrales [7] y cámaras profesionales [25], razón que puede dificultar el proceso de implementación, teniendo en cuenta el costo de estos dispositivos; en el presente proyecto, además de utilizarse cámaras de celulares configurados a 12MPx y los canales RGB, se incluyó una arquitectura orientada a microservicios, lo cual facilita la implementación de este proyecto en aplicaciones web, móviles e incluso en el uso de máquinas que consuman estos servicios para realizar los procesos de clasificación. Por otra parte, se corrobora que el uso de redes neuronales convolucionales para la solución de este problema de clasificación permite altos niveles de precisión y exactitud, incluso utilizando cámaras menos sofisticadas que las utilizadas en otras investigaciones.

Trabajos futuros

Gracias al uso de la arquitectura orientada a microservicios, el uso de contenedores, las automatizaciones realizadas para el despliegue y también a que el presente proyecto se encuentra público en un repositorio de Github; es posible redefinir su aplicación, creando microservicios para el entrenamiento de nuevos conjuntos de datos diferentes al campo de aplicación del café.

El presente proyecto puede ser utilizado como base para la creación de prototipos alrededor de la cosecha, clasificación o evaluación postcosecha del grano de café cereza.

Teniendo en cuenta la arquitectura de microservicios del presente proyecto y el uso de Flask que es una librería que permite exponer aplicaciones de Python como servicios, este también puede ser desplegado en la web para la creación de una aplicación móvil que, al consumir los servicios creados, pueda clasificar correctamente los granos de café cereza y a su vez esta aplicación podría ser utilizada para aumentar el número de registros del conjunto de datos de entrenamiento.

Se puede incluir nuevas clases como defectos en el grano de café y el grano de café cereza de color amarillo que corresponde a otra variedad, pero que también se considera como un grano de café maduro.

Referencias

- [1] Minagricultura, “Café: producto insignia de Colombia para el mundo.” Accessed: Nov. 10, 2024. [Online]. Available: <https://www.minagricultura.gov.co/noticias/Paginas/CafeProductoInsigniaColombiaMundo.aspx>
- [2] T. García, “Colombia es el tercer exportador de café en el mundo con 8% de participación global.” Accessed: Nov. 10, 2024. [Online]. Available: <https://www.agronegocios.co/mercados/colombia-exporta-8-del-cafe-consumido-en-el-mundo-3973258>
- [3] Essence of Coffee, “¿Por qué Colombia tiene el mejor café suave del mundo?” Accessed: Nov. 10, 2024. [Online]. Available: <https://www.essenceofcoffee.net/colombia-mejor-cafe-suave-del-mundo/>
- [4] Federación Nacional de Cafeteros Cauca, “Café de Cauca.” Accessed: Nov. 23, 2024. [Online]. Available: <https://cauca.federaciondefcafeteros.org/cafe-de-cauca/>
- [5] G. I. Puerta, “Calidad del café.” *Federación Nacional de Cafeteros de Colombia, Manual del cafetero colombiano: Investigación y tecnología para la sostenibilidad de la caficultura*, vol. 3, 2013.
- [6] S. Santos Valle, *Agricultura 4.0 - Robótica agrícola y equipos automatizados para la producción agrícola sostenible*, no. 24. in *Gestión integrada de cultivos*. Rome, Italy: FAO, 2021. [Online]. Available: <https://www.fao.org/publications/card/es/c/CB2186ES/>

- [7] M. A. Tamayo-Monsalve *et al.*, “Coffee Maturity Classification Using Convolutional Neural Networks and Transfer Learning,” *IEEE Access*, vol. 10, pp. 42971–42982, 2022, doi: 10.1109/ACCESS.2022.3166515.
- [8] J. Lelbovich, C. Córdoba, N. Gonzales, M. Villamil, J. Méndez, and J. M. Izquierdo, “Informe de Gestión 2020,” *Federación Nacional de Cafeteros de Colombia*, Bogotá, 2021. doi: 10.38141/10793.
- [9] Federación Nacional de Cafeteros de Colombia, “FNC en Cifras,” 2017. Accessed: Nov. 23, 2024. [Online]. Available: <https://federaciondefcafeteros.org/static/files/FNCCIFRAS2017.pdf>
- [10] F. F. Valencia, J. Arcila, A. M. Moreno Berrocal, L. F. Salazar Gutiérrez, and E. Hincapié Gómez, *Cafés Especiales*. Bogotá: Federación Nacional de Cafeteros de Colombia - Cenicafé, 2007. Accessed: Nov. 23, 2024. [Online]. Available: <https://www.cenicafe.org/es/documents/LibroSistemasProduccionCapitulo10.pdf>
- [11] A. M. Porcelli, “Artificial intelligence and robotics: its social, ethical and legal dilemmas,” Oct. 01, 2020. doi: <https://doi.org/10.32870/dgedj.v6i16.286>.
- [12] ISO/IEC, “¿Qué es la inteligencia artificial (IA)?” Accessed: Nov. 23, 2024. [Online]. Available: <https://www.iso.org/es/inteligencia-artificial/que-es-ia>
- [13] C. M. Pineda Pertuz, *Aprendizaje Automático y Profundo en Python*. Bogotá: Ediciones de la U, 2021.
- [14] Y. LeCun, Y. Bengio, and G. Hinton, “Aprendizaje profundo,” May 2015, doi: 10.1038/natureza14539.

- [15] J. A. López Sotelo, *Deep Learning Teoría y Aplicaciones*. Bogotá: Alpha editorial, 2021.
- [16] Mathworks, “Transfer Learning.” Accessed: Oct. 09, 2024. [Online]. Available: <https://la.mathworks.com/discovery/transfer-learning.html>
- [17] D. López Betancour, D. Bosco Rmebrant, C. Guerrero Méndez, Z. R. R, and T. Saucedo Anaya, “Comparación de arquitecturas de redes neuronales convolucionales para el diagnóstico de COVID-19,” *Universidad Autónoma de Zacatecas*, Dec. 2021, Accessed: Nov. 23, 2024. [Online]. Available: <https://www.scielo.org.mx/pdf/cys/v25n3/2007-9737-cys-25-03-601.pdf>
- [18] K. Simonyan and A. Zisserman, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION,” *Universidad de Oxford*, Apr. 10, 2015. Accessed: Nov. 23, 2024. [Online]. Available: <https://arxiv.org/pdf/1409.1556>
- [19] MathWorks, “MathWorks-vgg19.” Accessed: Oct. 19, 2024. [Online]. Available: <https://la.mathworks.com/help/deeplearning/ref/vgg19.html>
- [20] D. A. Pérez Aguilar, R. Risco, and L. C. Pacherez, “Transfer learning en la clasificación binaria de imágenes térmicas,” *INGENIUS*, Jun. 2021. doi: 10.17163/ingenius.n26.2021.07.
- [21] C. M. Pineda Pertuz, *Aprendizaje Automático y Profundo en Python*, 1st ed. EDU, 2021.
- [22] Federación Nacional de Cafeteros Colombia, “Calidad del café y sus Defectos”, Accessed: Oct. 09, 2024. [Online]. Available: chrome-

extension://efaidnbmnnnibpcajpcgclefindmkaj/https://biblioteca.cenicafe.org/bitstream/10778/1033/1/Calidad%20caf%C3%A9.pdf

- [23] Federación Nacional de Cafeteros de Colombia, “Normas del Calidad del Café”, Accessed: Oct. 09, 2024. [Online]. Available: chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://federaciondecafeteros.org/static/files/RESOL04.pdf
- [24] M. R. E. P. J. P. , B. O. M. Tamayo M. Manuel, “Clasificación del café maduro usando CNN,” *Manizales*, Apr. 11, 2023. Accessed: Oct. 19, 2024. [Online]. Available: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9755125>
- [25] H. L. Gope and H. Fukai, “Peaberry and normal coffee bean classification using CNN, SVM, and KNN: Their implementation in and the limitations of Raspberry Pi 3,” *AIMS Agriculture and Food*, vol. 7, no. 1, pp. 149–167, 2022, doi: 10.3934/AGRFOOD.2022010.
- [26] J. Montilla Páez *et al.*, “Propiedades Físicas Y Factores De Conversión Del Café En El Proceso De Beneficio,” *Cenicafé*, Apr. 2008. Accessed: Nov. 24, 2024. [Online]. Available: <https://biblioteca.cenicafe.org/bitstream/10778/358/1/avt0370.pdf>

Anexos

A1. Estructura del banco de pruebas.

El diseño del banco de pruebas de clasificación se realizó mediante el software CAD Solid Works versión 2018, el diseño está compuesto de 10 piezas, 3 bisagras para unir la base dividida en dos secciones de 1m de largo * 0.134m de ancho, de y darle el ángulo de inclinación para que el grano se deslice en una sola dirección, ángulos que soportan piezas perpendiculares a la base para guiar el grano en dirección de a la cámara para la toma de la imagen, 2 piezas laterales de 1m * 0.05m de ancho los cuales aseguran que el grano no salga del área de trabajo del banco de pruebas, a continuación se muestra una figura del banco de pruebas ensamblado

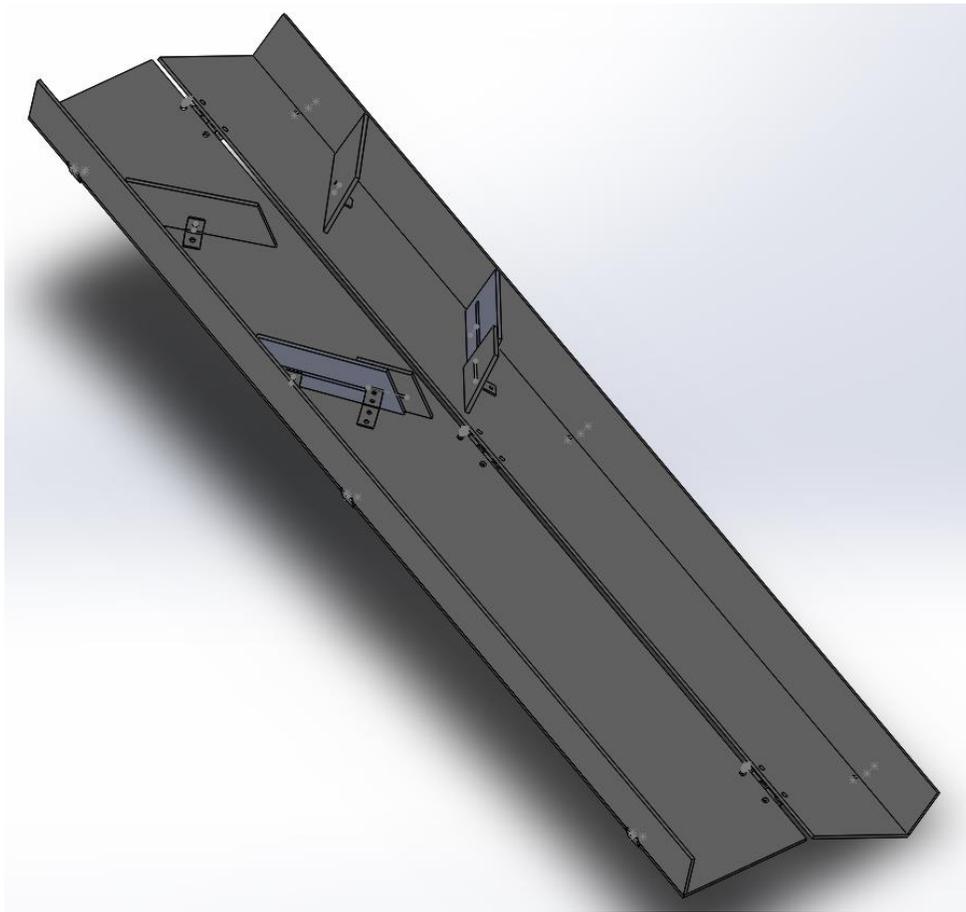


Fig. 25. Diseño del banco de pruebas para toma de imagenes en movimiento.

La altura del banco de pruebas es definida por la velocidad en la que pasa el cerezo de café por la zona de detección pues tiene relación directa, puesto que a mayor altura mayor velocidad, por esta razón se toma una altura de 10° con relación a la base del banco de pruebas, para definir la velocidad a la que pasa el grano por la zona de detección, se toma en cuenta la fórmula de conservación de energía, la medida del banco de pruebas y el peso promedio del cerezo de café, el peso de este es tomado como referencia de los estudios que realizo Cenicafé[26] donde promedia el peso del cerezo de café en 2 gramos y este será nuestra referencia en los cálculos.

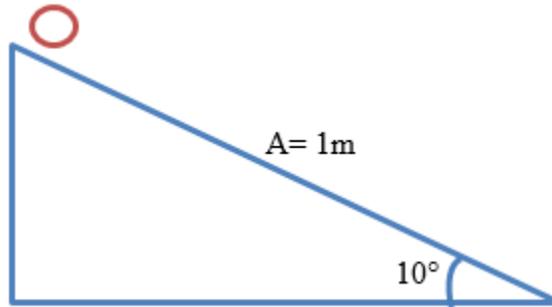


Fig. 26. Diagrama de plano inclinado para la construcción del banco de pruebas.

De acuerdo con los fundamentos de físicos de ley de la conservación de la energía se define la siguiente ecuación asumiendo un rozamiento nulo debido a la geometría del cerezo y una velocidad inicial debido a la caída del grano desde la tolva:

$$M = 2g = 0.002\text{kg.}$$

$$A = 1\text{m.}$$

$$V_o = ?$$

$$V_f = ?$$

$$h_o = ?$$

$$t_o = 1\text{s}$$

Se deben hallar los elementos anteriormente designados con signo de interrogación.

Primero se halla la altura inicial en donde cae el grano de la tolva para ello usamos la fórmula de aceleración asumiendo que la aceleración equivale a la gravedad con la que cae el cerezo:

$$a = \frac{V_f - V_o}{t}$$

$$9.8\text{m/s}^2 = \frac{V_f}{1\text{s}}$$

$$V_f = 9.8\text{m/s}$$

Se halla el valor de la altura inicial:

$$\text{sen } \sigma = \frac{\text{cat.op}}{h}$$

$$\text{sen } \sigma * h = \text{cat.op}$$

$$\text{sen}(10^\circ) * 1\text{m} = \text{cat.op}$$

$$\text{cat.op} = 0.173\text{m}$$

Reemplazando valores en formula de conservaci3n de energa asumiendo altura final nula (hf=0):

$$\Delta E_m = 0 \quad \Rightarrow \quad \underline{E_{mf}} - \underline{E_{mo}} = 0 \quad \Rightarrow \quad \underline{E_{mo}} = \underline{E_{mf}}$$

$$E_{co} = \frac{m*v^2}{2} + (m * g * h_o) = E_{cf} = \frac{m*v^2}{2} + m*g*hf$$

$$\frac{m*v^2}{2} + (m * g * h_o) = \frac{m*v^2}{2}$$

$$\frac{0.002\text{kg}*(9.8\text{m/s})^2}{2} + 0.002\text{kg} * \frac{9.8\text{m}}{\text{s}} * 0.173\text{m} = \frac{0.002\text{kg}*V_f^2}{2}$$

$$0.09604 \frac{\text{kgm}^2}{\text{s}^2} + 0.00339 \frac{\text{kgm}^2}{\text{s}^2} = \frac{0.002\text{kg}*V_f^2}{2}$$

$$\frac{2(0.0994 \text{kgm}^2/\text{s}^2)}{0.002 \text{kg}} = V_f^2$$

$$V_f = \sqrt{99.43\text{m}^2/\text{s}^2}$$

$$V_f = 9.971 \text{m/s}$$

cale

Esta es la velocidad a la cual pasa el grano cerezo por la zona de detecci3n.

Una vez realizados los c3lculos y al diseo se procede a desarrollar el banco de pruebas funcional usando como material primario l3minas de triplex que es un tipo de madera aglomerada, de una densidad de 4mm que da flexibilidad y portabilidad al banco de pruebas, en la figura N°() se muestra el desarrollo del mencionado.

Extracción del conjunto de datos de granos de café cereza



Fig. 27. foto del proceso de construcción del sistema de adquisición de imágenes.

A2. Fórmulas para el cálculo de las métricas de evaluación

Las fórmulas detrás de cada métrica evaluada se presentan a continuación:

Glosario de variables:

- VP: verdaderos positivos
- FP: falsos positivos

- VN: verdaderos negativos
- FN: falsos negativos
- Recall: sensibilidad
- Precisión: Precisión
- Accuracy; Exactitud

Clase 0: verdes

Precisión:

$$\frac{VP}{(VP+Fp)} = \frac{22}{(22+0)} = 1$$

Recall:

$$\frac{VP}{(VP+FN)} = \frac{22}{(22+2)} = 0.92$$

F1-Score:

$$2 * \frac{Precisión \times Recall}{Precisión + Recall} = \frac{1 * 0.92}{1 + 0.92} = 0.95$$

Clase 1: Pintones

Precisión:

$$\frac{VP}{(VP+Fp)} = \frac{44}{(44+4)} = 0.91$$

Recall:

$$\frac{VP}{(VP+FN)} = \frac{44}{(44+1)} = 0.98$$

F1-Score:

$$2 * \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} = \frac{0.91 \times 0.98}{0.91 + 0.98} = 0.94$$

Clase 2: Maduros

Precisión:

$$\frac{VP}{(VP+Fp)} = \frac{43}{(43+1)} = 0.98$$

Recall:

$$\frac{VP}{(VP+FN)} = \frac{43}{(43+4-1)} = 0.90$$

F1-Score:

$$2 * \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} = \frac{0.98 \times 0.90}{0.98 + 0.90} = 0.93$$

Clase 3: Sobre Maduros

Precisión:

$$\frac{VP}{(VP+Fp)} = \frac{37}{(37+2)} = 0.95$$

Recall:

$$\frac{VP}{(VP+FN)} = \frac{37}{(37+1)} = 0.97$$

F1-Score:

$$2 * \frac{\text{Precisión} \times \text{Recall}}{\text{Precisión} + \text{Recall}} = \frac{0.95 \times 0.97}{0.95 + 0.97} = 0.96$$

La precisión del modelo se determina mediante la métrica *Accuracy* una vez evaluado con el conjunto de datos de pruebas.

A3. Modificaciones realizadas al código de Tamayo y microservicios del sistema

Tal como se mencionó en la sección 3.2.3. se realizaron modificaciones al código de Tamayo, inicialmente se realizó una refactorización, eliminando librerías importadas pero no utilizadas, también se implementó un archivo de constantes que permite configurar todos los parámetros de entrenamiento y las diferentes rutas de archivos; otra modificación fue la eliminación de algunas redundancias en la configuración de las redes neuronales ya que al aplicar la validación cruzada se estaba añadiendo nuevas capas iguales a las que ya se tenía; por otra parte, se modificó el diseño del software ya que Tamayo utilizó el paradigma de programación procedimental lo que aumentaba la complejidad a la hora de entender el código, se implementó el paradigma orientado a objetos creando clases para funciones específicas de procesamiento de imágenes, entrenamiento del modelo y de validación; gracias al uso del paradigma orientado a objetos, se implementó una función que permite configurar cualquier modelo de red neuronal de la librería keras applications, utilizando pocas líneas de código.

Otra optimización realizada fue el uso de compresión de listas para reemplazar los ciclos for, lo cual reduce el tiempo de ejecución requerido, especialmente en el procesamiento de imágenes.

Con el fin de facilitar la implementación de los diferentes algoritmos, se creó una función que, de manera recursiva, carga un dataset, identifica sus clases internas de forma automática y crea las etiquetas; es importante mencionar que esta función depende de una

ruta del dataset, y de que el dataset esté debidamente separado en carpetas, donde cada carpeta corresponde a una clase. Otra modificación importante que facilita la trazabilidad y el entendimiento de los resultados fue el uso de la librería Logging de Python que permite entre otras cosas, almacenar todo lo que se imprime en la consola en modo ‘*Debug*’.

A4. Repositorio de trabajo

Para este trabajo se construyó un repositorio de Github que está disponible y publicado en la siguiente dirección:

https://github.com/Victor-Pinto/coffee_classification/tree/main

El dataset de imágenes creado se encuentra publicado en Zenodo con el siguiente dirección: <https://doi.org/10.5281/zenodo.11114229>

A5. Servicios creados para las fases de entrenamiento y evaluación de los modelos

Los principales servicios desarrollados fueron: servicio de entrenamiento de un modelo específico, servicio de entrenamiento de todos los modelos preconfigurados, servicio de evaluación de un modelo específico y servicio de evaluación de todos los modelos; estos servicios son explicados a continuación.

Servicio de entrenamiento de un modelo específico

Con el fin de realizar el entrenamiento, se creó un servicio que inicia con la entrada de la ruta del conjunto de datos y el nombre del modelo que se entrenará; posteriormente, el sistema carga el conjunto de datos y el modelo, divide el conjunto de datos en entrenamiento, pruebas y validación; después ajusta el modelo sin validación cruzada y guarda los resultados, tal como se puede observar en la Fig. 28.

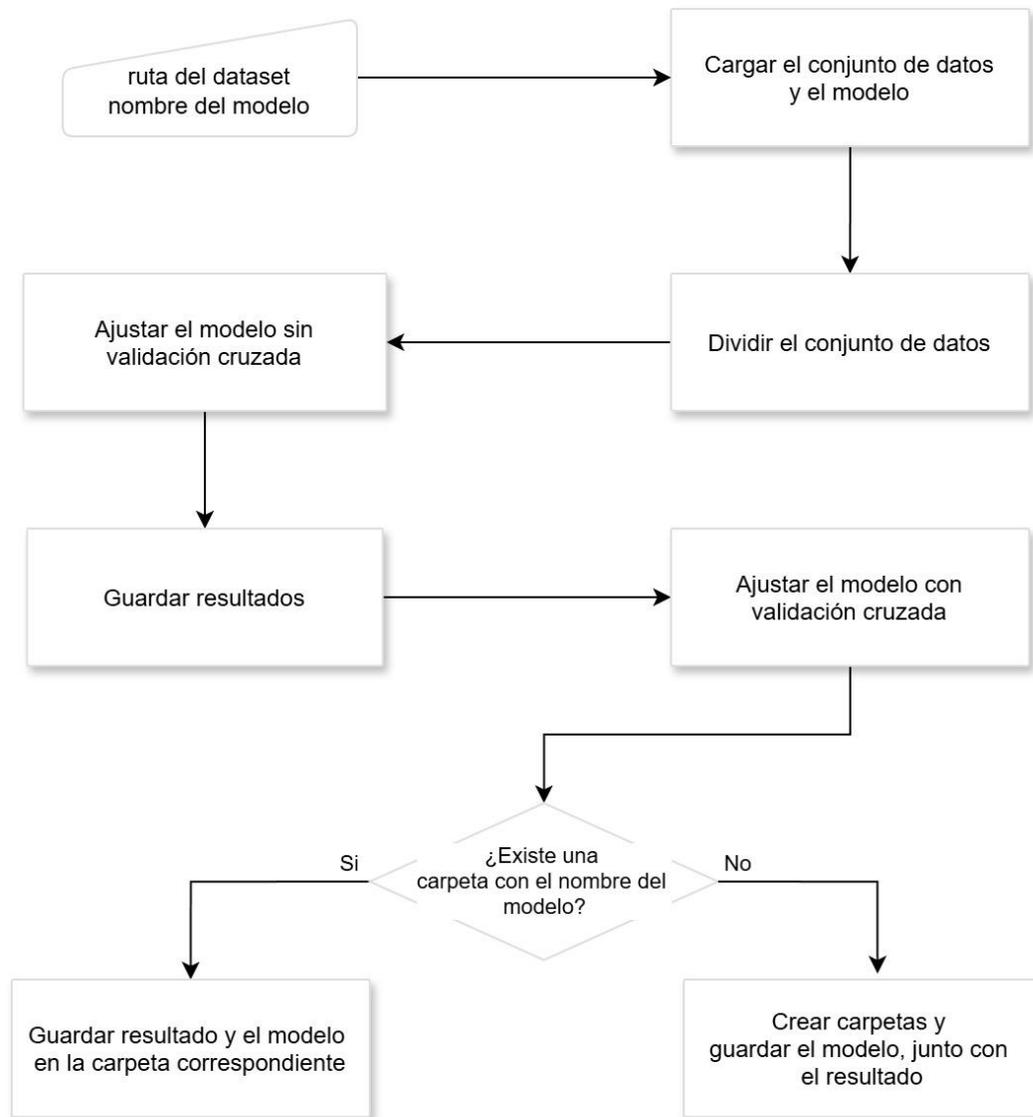


Fig. 28. Diagrama de flujo del servicio de entrenamiento.

Servicio de entrenamiento de todos los modelos

El servicio de entrenamiento para todos los modelos es similar al servicio para entrenar un único modelo, sin embargo, en lugar de recibir un modelo específico, tiene en

cuenta una lista con la definición de todos los modelos que es iterada con el fin de entrenar cada modelo; este cambio se puede apreciar al comparar la Fig. 28 con la Fig. 29.

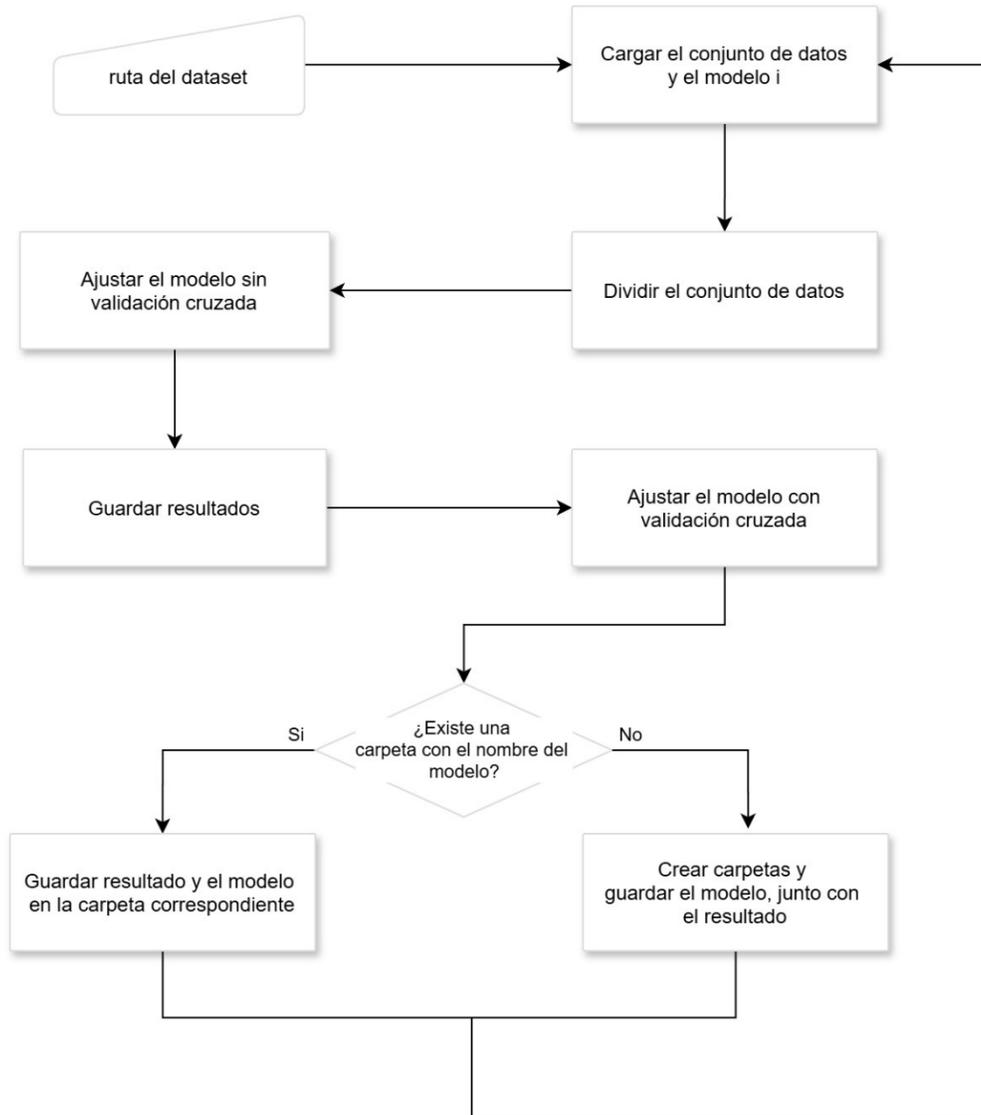


Fig. 29. Diagrama de flujo del servicio de entrenamiento para todos los modelos.

Servicio de evaluación de un modelo específico

El servicio de evaluación tal como se presenta en la Fig. 30, recibe la ruta del conjunto de datos para realizar la evaluación, la ruta del modelo; luego carga el conjunto

de datos realizando un etiquetado de acuerdo con la carpeta a la que corresponda y también carga el modelo especificado; utilizando el modelo, realiza predicciones sobre el conjunto de datos y finalmente guarda los resultados en la ruta del modelo indicado inicialmente.

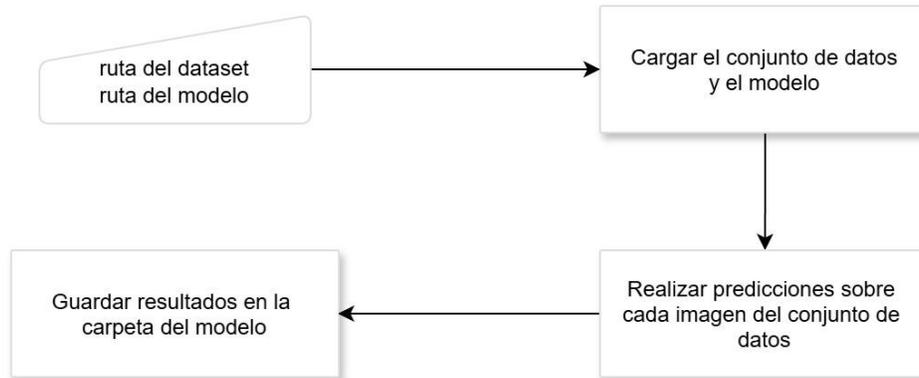


Fig. 30. Servicio de evaluación de un modelo específico.

Servicio de evaluación de todos los modelos

El servicio de evaluación de todos los modelos tal como se presenta en la Fig. 31, recibe la ruta del conjunto de datos para realizar la evaluación; luego carga el conjunto de datos realizando un etiquetado de acuerdo con la carpeta a la que corresponda y también carga el modelo que se encuentra en el principio de la lista; utilizando el modelo, realiza predicciones sobre el conjunto de datos y posteriormente guarda los resultados en la ruta del modelo específico y realiza una nueva iteración si existen más modelos en la lista.

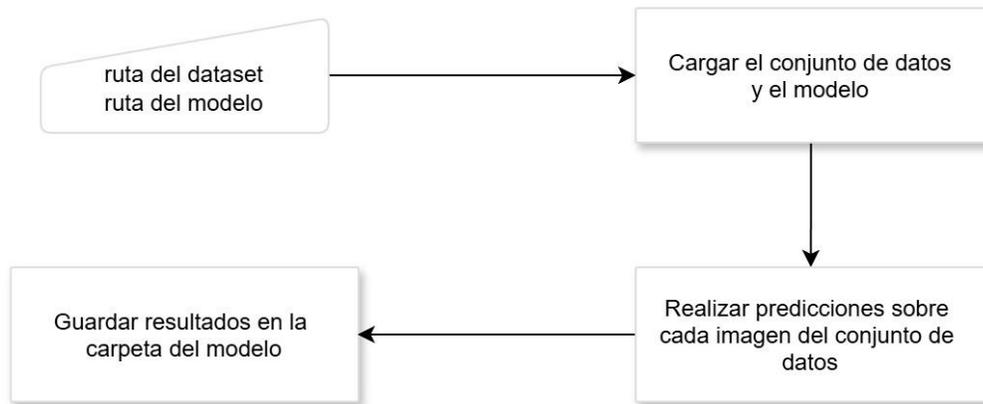


Fig. 31. Servicio de evaluación de un modelo específico.