

UNIVERSIDAD TÉCNICA DEL NORTE



Facultad de Ingeniería en Ciencias Aplicadas

Carrera de Software

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

Trabajo de grado previo a la obtención del título de Ingeniero en Software.

Autor:

Cevallos Sánchez Marcel Alejandro

Director:

PhD. Iván Danilo García Santillán

Ibarra – Ecuador

2023



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:	100332646-7		
APELLIDOS Y NOMBRES:	MARCEL ALEJANDRO CEVALLOS SÁNCHEZ		
DIRECCIÓN:	IBARRA, SAN FRANCISCO		
EMAIL:	macevalloss@utn.edu.ec		
TELÉFONO FIJO:	5000286	TELÉFONO MÓVIL:	0959090926

DATOS DE LA OBRA	
TÍTULO:	Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia artificial y Visión por computador.
AUTOR(ES):	MARCEL ALEJANDRO CEVALLOS SÁNCHEZ
FECHA:	16/02/2022
PROGRAMA:	PREGRADO
TÍTULO POR EL QUE OPTA:	INGENIERO DE SOFTWARE
DIRECTOR:	PhD. IVÁN GARCÍA
ASESOR 1:	Ing. MARCO R. PUSDÁ
ASESOR 2:	Ing. PEDRO GRANDA

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 24 días del mes de abril de 2023

EL AUTOR:



ESTUDIANTE

Marcel Alejandro Cevallos
Sánchez

C.I: 100332646-7

CERTIFICACIÓN DIRECTOR

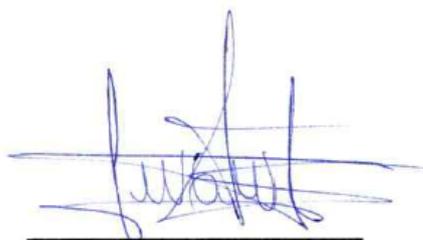
Ibarra 24 de abril del 2023

CERTIFICACIÓN DIRECTOR DEL TRABAJO DE TITULACIÓN

Por medio del presente yo PhD. Iván García, certifico que el Sr. Marcel Alejandro Cevallos Sánchez portador de la cedula de ciudadanía número 1003326467, ha trabajado en el desarrollo del proyecto de grado "Detección de defectos morfológicos en granos de café utilizando técnicas de Inteligencia artificial y Visión por computador", 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

Este trabajo está dedicado a mis padres Marcelo e Ima, mi hermana Mica y mi familia en general, pero en especial para aquellos que no pudieron ver la culminación del mismo, mi bisabuela Elina Vallejos y mi abuelo Juan Sánchez.

AGRADECIMIENTO

A mi familia por apoyarme y darme consejos para poder completar, aumentar y mejorar mi trabajo conforme avancé en el proceso de desarrollo, escritura y pruebas.

A todos quienes me ayudaron a realizar las pruebas necesarias para realizar mi trabajo les agradezco por prestarme su atención y tiempo.

A mis amigos y compañeros de carrera que me apoyaron realizando tareas tanto de fotografías, etiquetación, y hasta ayudándome a leer y corregir formatos en este documento como una que otra figura aquí presentada.

Tabla de contenidos

AUTORIZACIÓN DE USO Y PUBLICACIÓN A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTEII	
CERTIFICACIÓN DIRECTOR..... IV	IV
DEDICATORIA V	V
AGRADECIMIENTO VI	VI
Tabla de contenidos VII	VII
Índice de figuras VIII	VIII
Índice de tablas X	X
ResumenXI	XI
Abstract.....XI	XI
Introducción.....1	1
Antecedentes1	1
Situación actual1	1
Planteamiento del problema.....2	2
Objetivos3	3
Objetivo general3	3
Objetivos específicos3	3
Alcance3	3
Justificación.....3	3
Justificación tecnológica.....5	5
Justificación agrícola5	5
CAPÍTULO 1 Marco Teórico6	6
1.1 El Café y su Calidad.....6	6
1.1.1 Características Generales del Cultivo.....6	6
1.1.2 Características Morfológicas del Grano7	7
1.1.3 Estándares de Selección y Calidad del Grano10	10
1.2 Generalidades la de Visión por Computador12	12
1.2.1 Definición12	12
1.2.2 Proceso13	13
1.2.3 Algoritmos:14	14
1.3 Redes Neuronales Convulocionales.....19	19
1.3.1 Redes Neuronales.....19	19
1.3.2 CNN21	21
1.3.3 Aplicaciones de las CNN.....22	22
1.4 Trabajos relacionados.....23	23

1.4.1	A sugar Beet Leaf Disease Classification Method Based on Image Processing and Deep Learning	23
1.4.2	Dry Beans Classification Using Machine Learning.....	23
1.4.3	Review of Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification	23
1.4.4	Smart Agriculture: Real-time Classification of Green Coffee Beans by Using a Convolutional Neural Network.....	24
1.4.5	An automatic Bean Classification System Based on Visual Features to Assist the Seed Breeding Process	25
Capítulo 2 Desarrollo.....		26
2.1	Adquisición de imágenes.....	26
2.2	Preprocesamiento de imágenes	28
2.3	Arquitectura del modelo	29
2.3.1.	SSD Mobilenet	30
2.3.2.	SSD Resnet 50.....	30
2.4	Entrenamiento del modelo.....	30
Capítulo 3 Resultados y validación		34
3.1	Métricas cuantitativas	34
3.1.1	Matriz de confusión.....	34
3.1.2	Precisión	35
3.1.3	Recall	36
3.1.4	F1-score	36
3.2	Gráficas de rendimiento	37
3.3	Discusión	42
Conclusiones y recomendaciones		45
Conclusiones		45
Recomendaciones		46
Referencias.....		47

Índice de figuras

Figura 1	<i>Defectos en el café</i>	2
Figura 2	<i>Árbol de problemas</i>	2
Figura 3	<i>Proceso de recolección, actividades realizadas</i>	4
Figura 4	<i>Arquitectura de una CNN</i>	4
Figura 5	<i>Comportamiento histórico de las exportaciones de café en el Ecuador</i>	5
Figura 6	<i>Cinturón del café</i>	7
Figura 7	<i>Partes del grano de café</i>	7
Figura 8	<i>Proceso tradicional de tostado y producto obtenido</i>	8
Figura 9	<i>Herramienta “Canguro” utilizada para la recolección manual de café</i>	8

Figura 10	<i>Granos con cortes y rotos</i>	10
Figura 11	<i>Grano con pulpa o cascara pegada</i>	11
Figura 12	<i>Grano con daños por humedad</i>	11
Figura 13	<i>Grano afectado por insectos</i>	11
Figura 14	<i>Granos con malformaciones</i>	12
Figura 15	<i>Granos afectados por hongos</i>	12
Figura 16	<i>Grano afectado por variación de temperatura</i>	12
Figura 17	<i>Representación del espacio de color RGB</i>	14
Figura 18	<i>Descomposición del espacio RGB</i>	15
Figura 19	<i>Representación del espacio de color HSI</i>	15
Figura 20	<i>Descomposición del espacio HSV</i>	16
Figura 21	<i>Descomposición del espacio XYZ</i>	16
Figura 22	<i>Proceso de dilatación</i>	17
Figura 23	<i>Proceso de erosión</i>	17
Figura 24	<i>Proceso de apertura</i>	18
Figura 25	<i>Operación de relleno</i>	18
Figura 26	<i>Modelo de un MLP</i>	20
Figura 27	<i>Modelo básico de una CNN</i>	21
Figura 28	<i>Proceso de la capa convolucional</i>	22
Figura 29	<i>Proceso de la capa de Pooling al aplicar max pooling</i>	22
Figura 30	<i>Proceso de segmentación y separación de canales</i>	25
Figura 31	<i>Proceso de captura de imágenes de los granos de café</i>	26
Figura 32	<i>Esquema del proceso</i>	26
Figura 33	<i>Procesos de data augmentation realizados en roboflow</i>	29
Figura 34	<i>Arquitectura mobilenet</i>	30
Figura 35	<i>Arquitectura resnet 34 capas</i>	30
Figura 36	<i>Documento label_map</i>	31
Figura 37	<i>Campos a modificar en archivo config</i>	31
Figura 38	<i>Gráfico segundos por época mobilenet 5000</i>	32
Figura 39	<i>Gráfico segundos por época mobilenet 10000</i>	33
Figura 40	<i>Gráfico segundos por época resnet 50000</i>	33
Figura 41	<i>Matriz de confusión mobilenet</i>	34
Figura 42	<i>Matriz de confusión resnet</i>	35
Figura 43	<i>Pérdida total mobilenet</i>	37
Figura 44	<i>Pérdida de clasificación mobilenet</i>	38
Figura 45	<i>Pérdida de localización mobilenet</i>	38
Figura 46	<i>Pérdida total resnet</i>	38
Figura 47	<i>Pérdida de clasificación resnet</i>	39
Figura 48	<i>Pérdida de localización resnet</i>	39
Figura 49	<i>Map mobilenet</i>	40
Figura 50	<i>Recall mobilenet</i>	40
Figura 51	<i>Map resnet</i>	40
Figura 52	<i>Recall resnet</i>	41
Figura 53	<i>Pruebas de selección de café realizadas manualmente</i>	43

Índice de tablas

Tabla 1 <i>Estudios que aplican métodos de CNN en identificación de enfermedades de plantas</i> ..	24
Tabla 2 <i>Resumen comparativo de las características disponibles en cada herramienta</i>	27
Tabla 3 <i>Características de QA y gestión en las distintas herramientas</i>	28
Tabla 4 <i>Resumen métricas cuantitativas</i>	37
Tabla 5 <i>Comparativa de exactitud alcanzada</i>	42
Tabla 6 <i>Resultados de pruebas de selección de café realizadas manualmente</i>	43

Resumen

En el presente trabajo se plantea como objetivo el detectar defectos morfológicos en granos de café con este fin se propuso la utilización de una CNN o red neuronal convolucional trabajando en la misma desde el punto de la adquisición de imágenes hasta su entrenamiento. Tras recolectar las imágenes, mediante investigación y apoyo de un experto se etiquetó y clasificó 3276 las mismas en granos de café buenos y malos. Sobre estos se realizó un cambio de tamaño como preprocesamiento para cambiar su tamaño a 640x640 px y se aplicó data augmentation aumentando las imágenes disponibles a 20284. Al analizar los trabajos relacionados y las opciones disponibles se realizó entrenamiento sobre dos modelos, SSD mobilenet v2 y SSD resnet50, el ajuste final realizado sobre ambas fue de 5000 y 50000 épocas respectivamente, de igual manera cada uno alcanzó una exactitud de 76.92% y 52.01%, al comparar estos resultados con trabajos similares y con pruebas a mano son cotas bajas al confrontarlas con otros estudios pero se encuentra en la media de eficiencia de los sujetos con los que se realizó selección manual aunque no los supera en su punto álgido, mostrando problemas especialmente para realizar clasificación de granos buenos, clase en la que tiene f1-score 0.27 en su punto más bajo, el modelo resnet. Los resultados muestran amplia capacidad de mejora, pero también quiere decir que de momento no es posible el uso práctico de los mismos.

Palabras clave: Etiquetar, Data augmentation, Entrenamiento, Exactitud, F1-score, Clasificación.

Abstract

In this work the detection of morphological defects of coffee beans was set as objective in order to achieve this the development of a CNN or convolutional neural network was proposed working on it from image acquisition until training. After acquiring the images, through investigation and the help of an expert 3276, were labeled and classified between good and bad. Using this their resolution was changed to 640x640 px and data augmentation techniques were applied, increasing the number of available images to 20284. After analyzing related works and the available options, training was done on two models, SSD mobilenet v2 and SSD resnet50, the final fit was done on 5000 and 50000 epochs each, reaching an accuracy of 76.92% and 52.01% respectively, contrasting our results with the ones of similar works and tests done manually we found our results are low but fall in the middle of the tests done with people although, our models aren't able to get over their best results, showing issues especially with the coffee beans classified as good, class in with it has an f1-score of 0.27 at its lowest, the resnet model. This results show that there's still ample opportunity to improve the models, this also means that it has no practical use has of right now

Keywords: Label, Data augmentation, Training, Accuracy, F1-score, Classification.

Introducción

Antecedentes

El café, es una bebida con características organolépticas que cambian en dependencia de la variedad que se emplea, las mezclas de granos que se realizan, así como el manejo durante el proceso de postcosecha. Dentro de este último, el fruto se clasifica por el tamaño y densidad, lo que es aplicado en la mayoría de asociaciones de productores y empresas distribuidoras o cafetaleras, y se puede llevar a cabo de manera manual o mecánica; aunque la selección del grano seco se realiza fundamentalmente de forma tradicional. En los procesos industriales se utilizan distintas técnicas, dentro de las que encontramos las artificiales, sin embargo, estas pueden llegar a ser bastante costosas y difíciles de implementar para los pequeños y medianos productores tal como las describen Rosas-Echevarría et al. (2019). La imposibilidad de usar estas últimas y recurrir a los métodos más arcaicos, puede llevar a pérdidas económicas, bien por el tiempo que se demora en la postcosecha o por la dificultad de comercializar en un mercado que exige altos estándares en la calidad y selección del grano.

Situación actual

En la Asociación Agroartesanal de Caficultores Rio Intag (AACRI), se recolecta café de cerca de 130 fincas, acopiando de ellas entre 300 a 400 quintales en cada cosecha (Arcos et al., 2010). Pero no todo el café que se compra a los productores está en las mismas condiciones, aunque la asociación incentiva que el mismo se entregue ya seco y seleccionado. No todos los caficultores están en la posición de hacerlo, existen varias razones para esto, la falta de tiempo y recursos necesarios para realizar el proceso de postcosecha, ya sean estos económicos, para contratar mano de obra, o tecnológicos para verificar correctamente la humedad del grano. Además de lo anterior, hay productores que no realizan el proceso de secado del producto, puesto que comercializan el café directamente después de ser recolectado de la planta.

Teniendo en cuenta todo lo anterior, el café en sí mismo se separa en tres clases A, B y C, las que son de acuerdo al grado de selección que se haya realizado sobre el mismo, incluida la limpieza, sin embargo, antes de comercializarlo la asociación realiza otra selección que puede llevarse a cabo desde un quintal al día en el caso del café tipo C hasta 3 o 4 al día en el grado más alto, A.

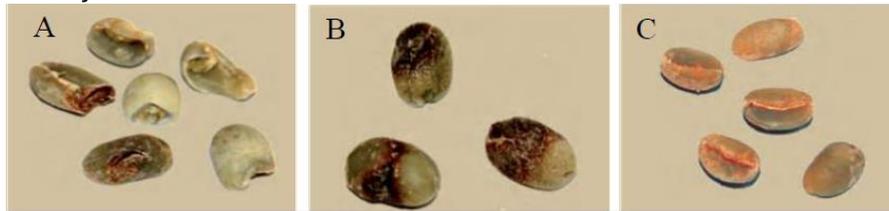
La razón por la que el producto pasa por varios filtros antes de ser comercializado se debe a las ventas del café sin procesar (sin tostar) hacia el mercado internacional, debido a las certificaciones que se requieren para lograr el cumplimiento de distintos estándares de calidad, en el caso de la AACRI posee actualmente las siguientes: "JAS" por la certificadora alemana BCS *Öko-Garantie*, la "Certificación Orgánica Ecuador" y "USDA organic".

Las pruebas de esta investigación se realizaron en la finca "Conrralá", ubicada en la Parroquia Peñaherrera, cantón Cotacachi, la misma cuenta con dos hectáreas dedicadas a la producción de café, sin embargo, a pesar de esta extensión la cantidad que se produce y vende es bastante menor de la esperada, debido a varios motivos como la falta de organización en la recolección, o la ausencia de replantación tras las cosechas. Además de lo anterior, el proceso de postcosecha no está estandarizado, ya que quienes lo realizan lo han hecho solo con conocimiento empírico, esto lleva a que exista un tiempo considerable antes de que se venda el producto a la asociación, debido a que pueden existir discrepancias con respecto al estado de un grano o errores de apreciación por el efecto de la edad, teniendo en cuenta que el proceso es llevado a cabo por los dueños, ya adultos mayores.

Planteamiento del problema

Los inconvenientes generados durante la postcosecha, conllevan a que los pequeños productores de café deban invertir tiempo y recursos para el proceso de selección manual de los granos sin malformaciones, cortados o de colores incorrectos (Figura 1), para alcanzar niveles de calidad que responda a las exigencias del mercado nacional e internacional.

Figura 1
Defectos en el café

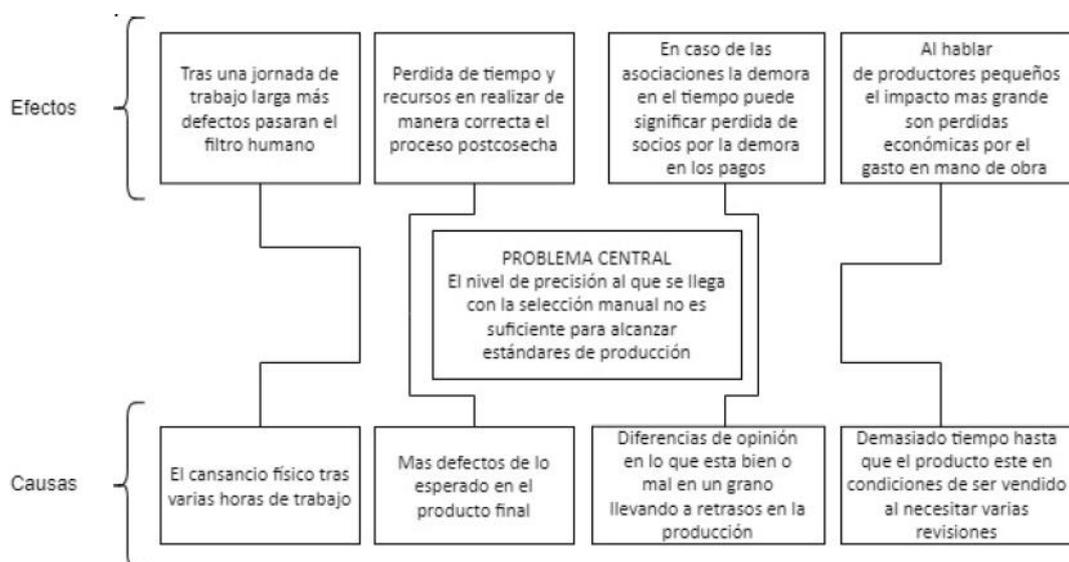


Nota. "A": granos cortados por causa de la despulpadora. "B": granos negros por efecto de enfermedades. "C": granos del color incorrecto por errores en el proceso de secado. Tomado de *Control de calidad del café Manual técnico* por Marín, 2013.

Lo anterior, también influye en que se incrementen los costos generales del proceso de postcosecha, lo que implica una pérdida de competitividad, tanto a nivel individual como asociativo. Si el proceso no se realiza de forma adecuada, los parámetros de calidad bajan, lo cual incide en el precio que reciben por su producto y en algunos casos la imposibilidad de comercializar en el mercado internacional.

Si se analiza el contexto anterior podemos plantear el árbol de problemas sobre el que se sustentó esta investigación (Figura 2).

Figura 2
Árbol de problemas



Objetivos

Objetivo general

Detectar defectos morfológicos en granos de café mediante la utilización de técnicas de visión artificial e inteligencia artificial.

Objetivos específicos

- Elaborar un marco teórico sobre las características de calidad del café y los procesos de cosecha y postcosecha.
- Desarrollar un modelo de red neuronal convolucional (CNN) para la detección de defectos en granos de café.
- Validar los resultados del modelo obtenido utilizando gráficas y métricas cuantitativas.

Alcance

El presente trabajo se llevó a cabo en un ambiente controlado, en el que las pruebas se realizaron con iluminación óptima, y la cámara RGB empleada se ubicó en una posición fija para tomar las imágenes en condiciones similares a las pruebas, las imágenes recolectadas sirvieron para poder desarrollar el entrenamiento de los modelos. Tanto el lenguaje como los frameworks a usar en el proyecto se decidieron en base a investigaciones previas en el área.

Es necesario señalar que se eligió de entre los lenguajes enfocados hacia la Inteligencia Artificial tales como Python, el más común en el área, o matlab, por su facilidad para matemática y estadística (Manrique, 2020); mientras que en el caso del framework se tomó en cuenta aquellos enfocados hacia el procesamiento de imágenes, tales como Keras/TensorFlow, orientado a redes neuronales (Arora & Bhatia, 2018), pytorch, similar a keras pero más simple (Jiang & Zhang, 2021) u openCV, por su arquitectura abierta y compatibilidad con varios lenguajes (Rosas-Echevarría et al., 2019). En cuanto a los criterios de aceptación del grano de café, en este caso específico sobre las variedades del café arábigo tales como caturra rojo, caturra amarillo, colombia, bourbon, entre otros, se definieron en el marco teórico en el que se aborda información sobre el mismo y su clasificación.

En la figura 3 se describe el proceso de cosecha y postcosecha así como el punto en el que se intervino para realizar el trabajo.

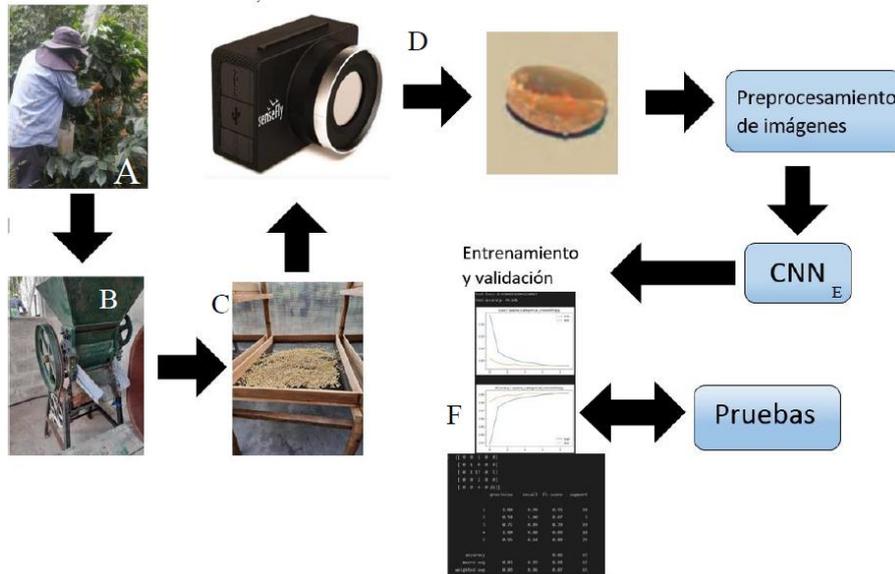
La arquitectura base de una CNN consta de la capa de entrada, la extracción de características (feature maps) y su clasificación (Figura 4), sin embargo, aunque sirva como guía tanto el número de capas como su complejidad deberán ser definidas en el trabajo, al igual que las características del grano a tomar en cuenta y las clases en las que se clasificó.

Justificación

El trabajo realizado al incorporarse 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” (Organización de las Naciones Unidas [ONU], 2018), al promover una tecnificación en el área agrícola al igual que con el plan de creación de oportunidades del Ecuador al fortalecer un proceso establecido como se habla en el objetivo 3 (Secretaría Nacional de Planificación y Desarrollo [SENPLADES], 2021).

Figura 3

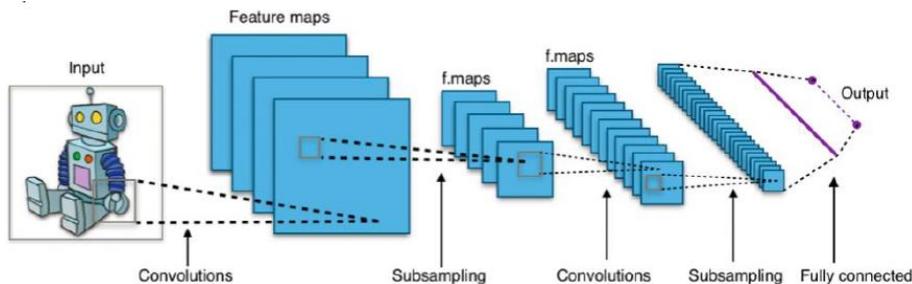
Proceso de recolección, actividades realizadas



Nota. “A”: recolección de grano “cereza” de la planta. “B”: Despulpadora usada en el proceso. “C”: secado de los granos de café tras el despulpado. “D”: Recolección de imágenes. “E”: Ver figura 4. “F”: Ejemplo de los resultados a obtener como graficas de precisión, matriz de confusión y datos como F1-score (precisión y sensibilidad combinadas) necesario cuando los datos están desbalanceados.

Figura 4

Arquitectura de una CNN



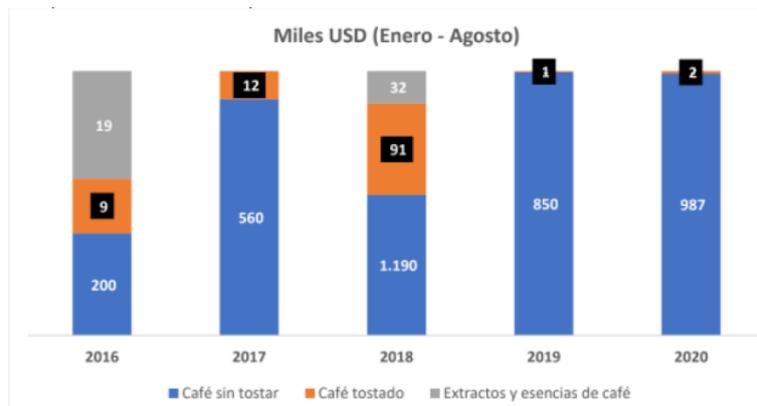
Nota. Tomado de *Handwriting recognition using Deep Learning in Keras* por Arora & Bhatia, 2018.

Se justifica el proyecto por el área a la que está enfocado, la comunidad productora de café, considerando que éste y sus derivados forman parte de los tres principales productos consumidos a nivel mundial con una certificación en especial cuando se trata de café sin tostar, que en Ecuador se ha vuelto la mayor parte de las exportaciones dentro el ámbito cafetalero (Figura 5), y que tanto productores como asociaciones deben apegarse a estas y sus reglas, son las razones principales por las que se propone este proyecto. Además del fácil acceso que se tiene para la toma de datos por la cercanía que existe con el medio.

Los modelos desarrollados en este trabajo, aunque no funcionan como solución directa a la necesidad de fortalecer el proceso de postcosecha si puede ser parte de una solución futura que apoye esta actividad en asociaciones cafetaleras como como la ACCRI o incluso para productores independientes.

Figura 5

Comportamiento histórico de las exportaciones de café en el Ecuador



Nota. Tomado de *Impacto en las exportaciones de organizaciones de pequeños productores, afectaciones, desafíos y oportunidades* por Farías et al., 2020.

Justificación tecnológica

Aunque como se mencionó antes los modelos no solucionan directamente el problema de la tecnificación la futura solución en la que se apliquen puede llegar a tener tanto un impacto económico al igual que social por el posible ahorro en el tiempo, haciendo que el productor pueda aprovechar mejor sus recursos o que las asociaciones puedan aceptar más socios al realizar más rápidamente el proceso.

Justificación agrícola

Al mejorar el proceso de postcosecha se espera que ayude al sector cafetalero, en especial en Ecuador, tan dependiente del sector agrícola, a mantener el ritmo de una creciente demanda mundial de café.

CAPÍTULO 1 Marco Teórico

1.1 El Café y su Calidad

1.1.1 Características Generales del Cultivo

Existen más de 100 especies de café, dentro de estas existen dos que destacan por ser cultivadas de forma significativa con fines comerciales, estas son *Coffea arabica* L. y *Coffea canephora* P., más conocidas por sus nombres comunes arábigo y robusta respectivamente, ambas, al igual que la mayoría, originarias de África, desde donde eventualmente se diversificaron en las variedades que conocemos hoy en día. Sin embargo, en América Latina y por ende Ecuador, no existe una gran variedad en lo que a especies respecta, pues en su gran mayoría todas son descendientes de una única planta que llegó en 1690, causando que exista una gran uniformidad genética (Cevallos, 2000).

Es esta misma uniformidad genética la que nos da ciertas generalidades sobre las características de la planta, como las que nos mencionan Wallis-García et al. (2004). El cafeto es un arbusto perennifolio que puede llegar hasta 9 metros de alto en algunos casos, con hojas que pueden variar entre una coloración verde o amarilla, y los frutos que produce la planta pueden ser amarillos o rojos en su madurez, la que es alcanzada alrededor de las 32 semanas desde el inicio de la floración.

Cevallos et al. (2018) nos dicen que existen ciertos aspectos, que proporcionan una mayor calidad del cultivo, tales como:

- **Altitud:** la altitud ideal para esta planta se encuentra entre los 900 y 1800 msnm. Se tiene como regla general que el café producido a mayor altura tiende a ser más grande y a su vez realza sus cualidades organolépticas.
- **Temperatura:** la temperatura idónea el cultivo oscila entre 22 y 28 °C. Este aspecto afecta directamente el crecimiento de la planta, una temperatura demasiado baja da como resultado una maduración de frutos tardía, mientras que el caso contrario acelera senescencia del producto, llevando a menos producción y un mayor riesgo de enfermedades.
- **Precipitación:** Para tener un adecuado crecimiento se espera entre 1700 a 2500 mm de precipitación anuales. Aunque no se puede evitar la sobre precipitación en ocasiones de sequía o en zonas más secas existen varias prácticas que ayudan en este problema.
- **Luminosidad:** El cafeto necesita de 8 a 13 horas de iluminación para la floración, considerándose de esta manera una planta de día corto, cuando existe una sobreexposición se cierran las estomas en las hojas de la planta, para protegerse de la transpiración excesiva, esto lleva a una disminución en la producción de alimentos de la planta y obviamente una disminución de la productividad.

Todas estas características las podemos encontrar con algunas variaciones en los países que se encuentran dentro de lo que se denomina el cinturón del café como se ve en la figura 6.

Figura 6
Cinturón del café

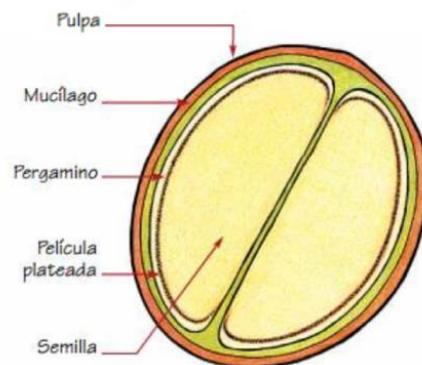


Nota. Tomado de *La geología del café* por Valbuena, 2019.

1.1.2 Características Morfológicas del Grano

El Instituto Interamericano de Cooperación para la Agricultura (IICA, 2019) define el fruto del café como una baya drupácea, conformada por la pulpa, epidermis o exocarpio y el mucilago o mesocarpio que es una sustancia azucarada que recubre los dos granos enfrentados con cara plana. La almendra o semilla que es de donde se extrae la bebida, está cubierta por dos capas: una película plateada (espermoderma) y el pergamino o endocarpio. En la figura 7 podemos observar estas partes.

Figura 7
Partes del grano de café



Nota. Tomado de *Manual de producción sostenible de café en la República Dominicana* por Instituto Interamericano de Cooperación para la agricultura (IICA), 2019.

La semilla es una nuez, de forma oblonga, plano convexo y de tamaño variable, es esta la que pasa por el proceso de tostado, ilustrado en la figura 8, y trituration para eventualmente convertirse en la bebida del café.

Podemos dividir los métodos de cosecha en dos grandes grupos, selectivos y no, los primeros, realizados de manera manual y se refiere a los métodos en los que se solo se recolectan los granos de coloración roja o amarilla, es decir aquellos maduros. Para realizar esta tarea se pueden utilizar distintas herramientas como la que se presenta en la figura 9 tanto para la recolección como almacenamiento entre fases. Aunque este proceso puede tardar, en

especial en plantaciones grandes o cuando no se dispone de suficiente mano de obra, se compensa al no necesitar una separación más adelante.

Figura 8

Proceso tradicional de tostado y producto obtenido



Figura 9

Herramienta “Canguro” utilizada para la recolección manual de café



Los métodos no selectivos suelen ser utilizados casi exclusivamente de manera industrial a gran escala, utilizando máquinas especializadas, sin embargo, también existen aquellos métodos manuales que entran dentro de esta categoría, como por ejemplo el “ordeño” total o parcial, procedimientos en los que se ignora la madurez del grano y se recolectan todos los frutos que se encuentran en las ramas.

En el estudio realizado por Isaza-Gil et al. (2006) se comparó el tradicional método selectivo con el “ordeño” total y parcial, obteniendo los siguientes resultados.

Se midió el promedio de frutos maduros recolectados en 3 fincas utilizando los 3 métodos antes mencionados obteniendo como resultado los porcentajes de 97.89 para el método selectivo tradicional, 88.05 al utilizar ordeño parcial y 84.52 en el total. Además, como

medidor de productividad, se obtuvieron los kilos recolectados por hora, obteniendo 10, 26.7 y 24.7 respectivamente.

Resultados similares a los que obtuvieron Wallis-García et al. (2004) quienes indican que, aunque la eficiencia aumenta al emplear los métodos no selectivos la calidad de la recolección disminuye. Es por esto que al menos en el caso manual, en el que la deficiencia no se puede ignorar como cuando la mayor parte del proceso esta automatizado, se debe tener en cuenta las prioridades con respecto al producto a alcanzar al momento de elegir la forma de cosechar. También debemos considerar que los métodos no selectivos pueden llegar a ser contraproducentes cuando no se realizan en el tiempo adecuado, ya que es necesario tanto cierto nivel de madurez en la rama, así como que exista cierta concentración de frutos en la misma.

Tras la recolección existen ciertas tareas que se realizan para garantizar la calidad del grano. El conjunto de estas actividades es conocido como el proceso de postcosecha, Bastidas (2017) nos describe los siguientes:

- Selección de cerezas que floten por menor densidad (vanos): Se colocan los frutos en piscinas o tinajas con agua. Aquellos frutos con menor densidad o vanos, es decir vacíos, flotan hasta la superficie haciendo más fácil el reconocerlos y separarlos del producto principal.
- Selección entre verdes y pintones: Es necesario separar ambos ya que los verdes son aquellos frutos inmaduros y disminuyen la calidad del producto final, el tiempo de este proceso depende del tipo de recolección que se realice, pues, aunque se utilicen métodos selectivos aún existe cierto porcentaje de frutos que no han llegado a la madurez dentro del producto.
- Clasificación por tamaño: Esta tarea se realiza cuando se desea obtener una mayor uniformidad en los procesos posteriores, la misma se puede realizar bien junto con el despulpado mediante el uso de zarandas o antes de realizar el mismo.
- Despulpado: Proceso necesario para obtener las semillas del fruto que eventualmente se convertirán en la bebida, es posible llevarlo a cabo de varias maneras, a mano o utilizando una despulpadora, se puede utilizar agua para minimizar la cantidad de granos sin despulpar o mordidos por la máquina a cambio de aumentar la dificultad de manejo de la pulpa, es decir sus residuos.
- Fermentación: Para la obtención de cafés especiales este paso es fundamental, pues dependiendo del tiempo o lavados que se le realicen, ciertas características se verán afectadas, pudiendo bien atenerse a un estándar o intentar obtener un sabor diferenciado de los demás productores y distribuidores.
- Secado: En esta etapa se debe mantener una humedad entre 10 y 12 % dependiendo de la variedad del fruto. Una humedad muy alta puede causar problemas de moho u otras enfermedades, mientras que en el caso contrario el pergamino queda dañado afectando la semilla en el interior. Cuando se realiza de manera manual no es posible controlar ciertos aspectos ambientales, en estos casos es necesario un monitoreo constante.
- Selección: Durante los distintos procesos es casi imposible que no existan granos que no pasen un corte de calidad o residuos que se mezclen con el producto final. Ciertos errores, como la pulpa seca o pergamino desprendido en el conjunto final se resuelven rápidamente mediante el uso de zarandas en algún punto anterior o actual, o al “aventar”, donde se toma ventaja del viento para eliminar estos desperfectos. Existen

varios problemas con semillas en sí que deben ser separadas a mano, estos desperfectos serán abordados más tarde en el trabajo.

También debemos tomar en cuenta que ninguno de estos procesos es necesario para la comercialización, ya que también se vende el café en cereza, sin embargo, la cantidad de procesos y la atención que se le dedique a cada uno de ellos, dependerá del mercado al que se apunte, que puede ser nacional, internacional, la venta a un intermediario o en varios casos para consumo personal.

1.1.3 Estándares de Selección y Calidad del Grano

Los cafés especiales o de calidad ocupan un lugar fundamental en el mercado por el que se paga un precio superior, la comercialización de estos requiere una certificación dependiendo de cual es el destino final del producto (Otero, 2019), algunas de estas certificaciones son: “JAS” por la certificadora alemana BCS *Öko-Garantie*, la “Certificación Orgánica Ecuador” o “USDA organic”. Estas certificaciones evalúan algunas características relacionadas con el cultivo (suelo, clima y abonos que se usan), así como con el proceso de postcosecha y la calidad del grano que se obtiene. En el caso del grano, lo que se realiza para obtener estas certificaciones es una selección del mismo que, aunque existe cierta variación entre certificaciones y algunos tipos especiales de café sobre lo que se considera una falla existen varios defectos generales a tomar en cuenta, a continuación, se describen los mismos y algunas de las posibles razones que los causan.

- Granos con cortes o rotos: Al usar la máquina de despulpado algunos granos son cortados por la misma, afectando su forma o al pergamino, se diferencian de los sanos por sus partes faltantes (Figura 10).

Figura 10

Granos con cortes y rotos



- Pulpa o cascara pegada: En algunos casos quedan restos de la cáscara o pulpa al realizarse el proceso de despulpado, estos restos se adhieren a los granos o no se desprenden del todo, es importante intentar separar la cascara mientras aún no se ha secado pues caso contrario al estar seco se dificulta el realizarlo y se debe descartar ese grano. Además, se pueden generar posibles efectos de la humedad por estos restos al separarlos (Figura 11).
- Daños por humedad: Un mal secado u objetos extraños en el grano puede generar una acumulación de humedad, lo que puede llevar a daños por la misma, los que son identificables por manchas en el pergamino (Figura 12).

Figura 11

Grano con pulpa o cascara pegada



Figura 12

Grano con daños por humedad



- Insectos: Algunos insectos se consideran plagas para el cultivo, aunque los más importantes son los barrenadores, que afectan el grano en estado de maduración, siendo el más común el conocido como broca de café (*Hypothenemus hampei* Ferrari) (Figura 13).

Figura 13

Grano afectado por insectos



- Malformaciones: Durante la maduración del grano se pueden producir afectaciones en la forma del mismo, tales como caracolillos, granos vanos, entre otras malformaciones sin clasificar (Figura 14).

Figura 14
Granos con malformaciones



Nota: A Caracolillo, B vano

- Hongos: La afectación más común en el grano es aquella producida por el hongo conocido como la roya (*Hemileia vastatrix* Berk. & Broome) (Figura 15).

Figura 15
Granos afectados por hongos



- Variación brusca de temperatura: Los cambios bruscos durante el proceso de secado llevan como consecuencia la rotura del pergamino (Figura 16).

Cabe destacar que un grano puede estar afectado por varios de estos de defectos a la vez.

Figura 16
Grano afectado por variación de temperatura



1.2 Generalidades la de Visión por Computador

1.2.1 Definición

Para poder aclarar el término de visión por computador y poder desarrollar una definición propia sobre el tema se cita a varios autores que han hablado sobre el tema

Learned-Miller (2011) define a la visión por computador como la ciencia de otorgar vista a computadoras u otras máquinas, sin embargo, el ver no significa únicamente el grabar o tomar fotografías con una cámara, para este caso se puede decir que la visión es la habilidad de interpretar imágenes. Atencio (2009) coincide con esta definición al incluir en la suya que estos sistemas constan de los procesos de captura de imagen, procesamiento e interpretación de la misma, también nos explica que, aunque es importante la existencia de una inteligencia artificial no es necesaria para su funcionamiento.

Villa y Cruz (2013) señalan que este campo consiste en la deducción de la estructura y propiedades de un mundo tridimensional, a partir de una o varias imágenes bidimensionales del mismo.

De igual manera tenemos autores como Mery (2004) quien nos explica que el proceso de la visión por computador engloba varias técnicas de reproducción de imágenes bajo distintas condiciones, cambiando así sus características.

Así pues, tomando en cuenta estos autores y sus definiciones podemos decir que la visión por computador es el conjunto de técnicas por el cual se intenta replicar el sentido de la vista humano y por lo tanto una de las bases para la inteligencia artificial, al tomar sus interpretaciones para realizar acciones o tomar decisiones, intentando así parecerse a nuestro razonamiento, en el que tomamos estímulos externos y hacemos lo propio basados en nuestra experiencia y conocimientos.

1.2.2 Proceso

Al igual que varios autores dan definiciones sobre lo que es la visión por computador existen de la misma forma otros que dan su aporte en lo que respecta al proceso general que se lleva a cabo para alcanzar lo que describen.

Moreno y Ramírez (2017) descomponen la idea propuesta inicialmente por Villa y Cruz (2013), para poder explicar el proceso de visión por computadora explicándolo en 5 pasos: adquisición de imágenes, pre-procesamiento de imágenes, extracción de características, identificación y/o generación de conocimiento y toma de acciones.

- A) Adquisición de imágenes: puede ser tanto una recolección en vivo mediante un video o al encontrar un dataset o un grupo de imágenes que se adecuen al objetivo a alcanzar o en caso de que no existan se deben adquirir estas por cualquier medio, tales como fotografías, ultrasonidos, resonancias magnéticas, entre otras técnicas.
- B) Pre-procesamiento de imágenes: el ruido o imperfecciones en la imagen pueden alterar las decisiones o interpretaciones eventualmente tomadas por el sistema. Para esto existen varias operaciones para eliminar el mismo, esto se realiza mediante algoritmos propios de distintas tecnologías o técnicas más visuales como cambios en el espectro de color, incremento de contraste, etc.
- C) Extracción de características: consiste en extraer las características que faciliten el identificar zonas de interés, lo cual puede ser tan simple como el buscar bordes, líneas, segmentación en base al color o con técnicas más complejas como el uso de histogramas, identificación de patrones, o texturas. El objetivo es identificar ciertos descriptores que se utilizan en etapas siguientes. El autor nos menciona que en este punto es posible terminar el proceso ya que algunas de las técnicas de extracción ya se podrían considerar interpretación en algunos casos.

- D) Identificación y/o generación de conocimientos: aquí se toma lo adquirido en la fase previa y se utilizan mecanismos de información para intentar identificar elementos complejos como animales, estructuras, y otros. Se puede decir que es aquí donde se convierte en un proceso multidisciplinario donde se aplica la inteligencia artificial.
- E) Toma de acciones: como se da a entender esto requiere la toma de una acción en base a las respuestas. Como reproducir música, sonar una alerta o acciones más complejas aplicadas en sistemas como conducción asistida, cirugía o reconstrucción de objetos.

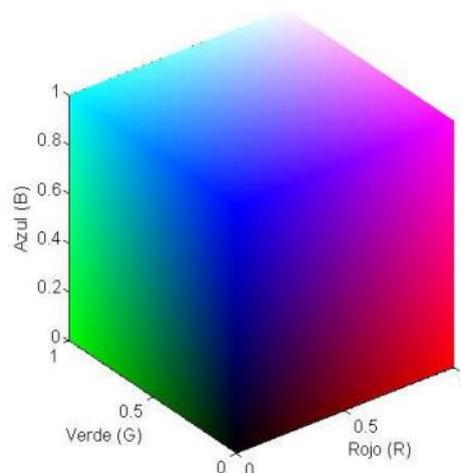
1.2.3 Algoritmos:

Aquí se describen los distintos algoritmos propios de la visión por computador es decir aquellos usados en las fases de pre-procesamiento y extracción de características.

- Cambios en el espacio de color: un espacio de color es un método de representación, por el cual, como lo dice su nombre, se visualizan colores. Dependiendo del tipo de captación de imágenes y la aplicación que se desarrolle, puede ser necesario cambiarlo. La forma en que esto se aplica es tanto al cambiar entre los varios espacios mediante ecuaciones, así como al separar los distintos canales de cada uno de los espacios de color con el objetivo de mitigar o eliminar el ruido (Martín et al., 2016), como se ejemplifica en la figura 17. Algunos de los espacios son:
 - ❖ RGB (rojo, verde y azul, siglas en inglés): se refiere a los tres colores primarios (rojo, verde y azul), los cuales al variar sus combinaciones forman distintos colores representado en la figura 18, este se utiliza en la mayoría de pantallas y cámaras.

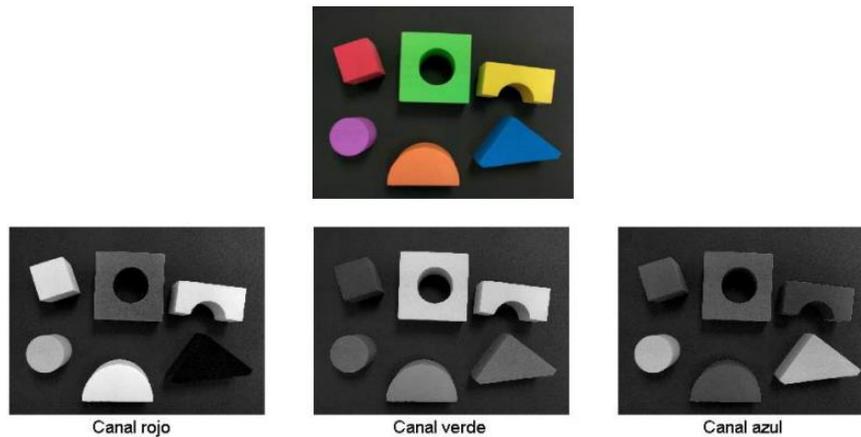
Figura 17

Representación del espacio de color RGB



Nota. Tomado de *El color: Modelos y transformaciones en el espacio de color en Conceptos y métodos en visión por computador* por Martín et al., 2016.

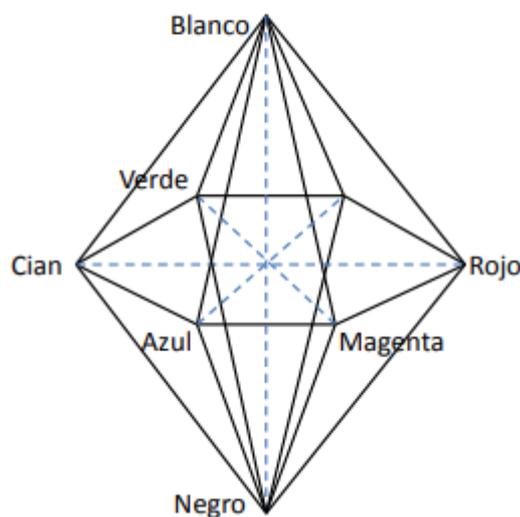
Figura 18
Descomposición del espacio RGB



Nota. Tomado de *El color: Modelos y transformaciones en el espacio de color en Conceptos y métodos en visión por computador* por Martín et al., 2016.

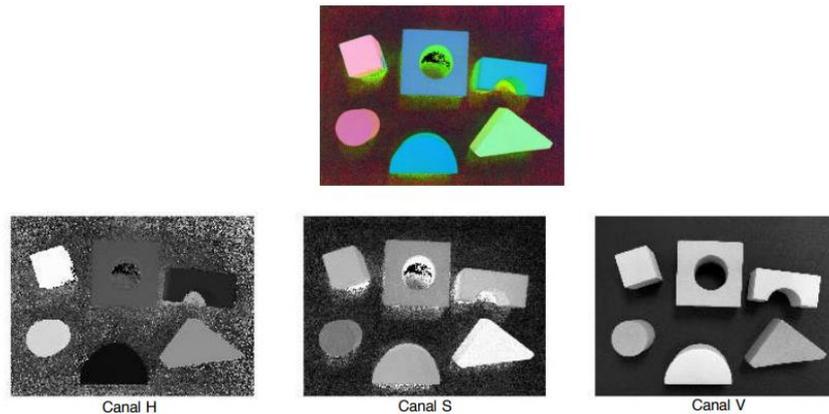
- ❖ HSI (tono (H), saturación (S) y brillo (I), siglas en inglés): espacio de color representado por la figura 19 posee distintas variaciones sobre el mismo tales como: HSL (H, S y luminosidad), HSV (H, S y valor (V)), HCl (H, croma (C), I), o HVC (H, V, C). No es usado comúnmente, pero se suele cambiar RGB hacia estos modelos por la descomposición de sus canales por ejemplo la de HSV en la figura 20 y el distanciamiento que tienen de la intensidad lumínica para los casos en los que esta puede alterar el resultado

Figura 19
Representación del espacio de color HSI



Nota. Tomado de *El color: Modelos y transformaciones en el espacio de color en Conceptos y métodos en visión por computador* por Martín et al., 2016.

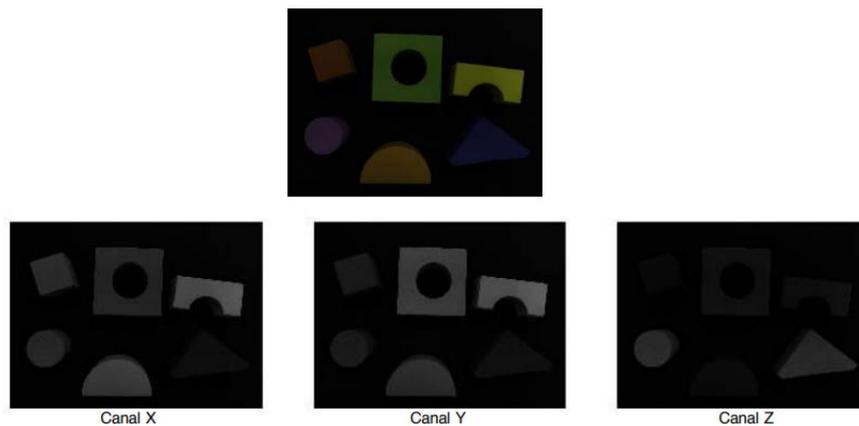
Figura 20
Descomposición del espacio HSV



Nota. Tomado de *El color: Modelos y transformaciones en el espacio de color en Conceptos y métodos en visión por computador* por Martín et al., 2016

- ❖ XYZ: espacio avanzado para distintos colores en determinadas longitudes de onda, se utiliza con la finalidad de evitar coeficientes negativos en la colorimetría, se usa sobre sistemas de 16 bits y sus variaciones sobre 8 bits (figura 21).

Figura 21
Descomposición del espacio XYZ



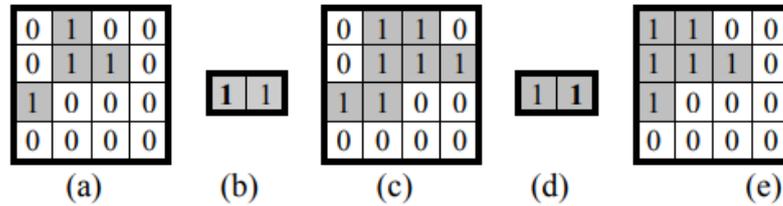
Nota. Tomado de *El color: Modelos y transformaciones en el espacio de color en Conceptos y métodos en visión por computador* por Martín et al., 2016.

- Procesamiento morfológico: según Cancelas et al. (2016) estas son un grupo de técnicas que analizan las imágenes basándose en la forma, su uso principal es el pre-procesamiento para la eliminación de ruido tanto en imágenes binarias, a color o en escala de grises, las transformaciones morfológicas matemáticas poseen varias operaciones básicas como:
 - ❖ Dilatación: conocida también como adición de Minkowski, es una operación en la que se aplica un crecimiento sobre cada uno de los puntos de la imagen,

es posible orientar este crecimiento, sin embargo, es preferible no privilegiar ninguna dirección, en la figura 22 podemos observar la dilatación en dos sentidos distintos.

Figura 22

Proceso de dilatación

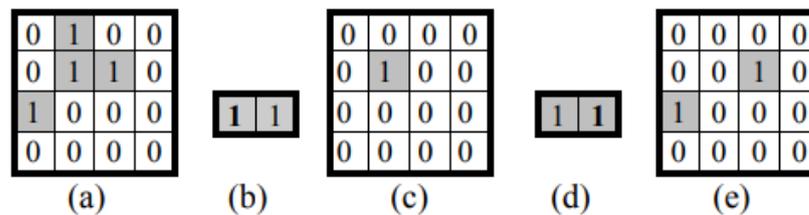


Nota. (a) Imagen original (b) EE centrado con origen en el píxel izquierdo; (c) Imagen dilatada con (b); (d) EE centrado con origen en el píxel derecho (d); (e) Imagen dilatada con (d). Tomado de *Procesamiento morfológico en Conceptos y métodos en visión por computador* por Cancelas et al., 2016

- ❖ Erosión: operación contraria a la dilatación, en la que se reduce el tamaño del objeto, útil para eliminar ruido de pequeño tamaño, al igual que en su cálculo opuesto se puede orientar, ejemplo en la figura 23.

Figura 23

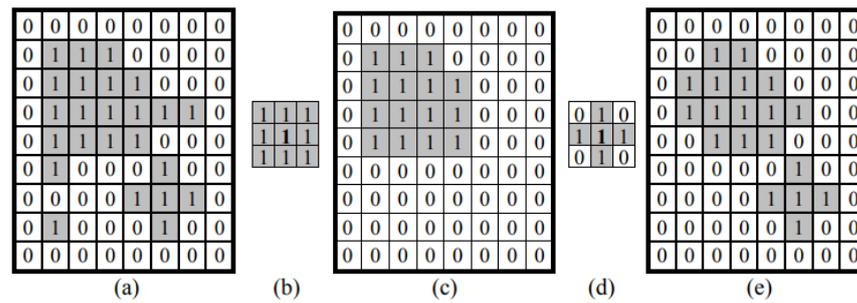
Proceso de erosión



Nota. (a) Imagen original; (b) EE centrado con origen en el píxel izquierdo; (c) Imagen erosionada con (b); (d) EE centrado con origen en el píxel derecho (d); (e) Imagen erosionada con (d). Tomado de *Procesamiento morfológico en Conceptos y métodos en visión por computador* por Cancelas et al., 2016.

- ❖ Apertura: aplicación conjunta las dos operaciones anteriores en la que se realiza una erosión seguida de una dilatación, esta se utiliza para realzar detalles, ya que tras la erosión aquellos que queden serán más notables, ilustración en la figura 24.
- ❖ Cierre: Similar a la apertura, pero se realizan las operaciones en distinto orden, se utiliza al binarizar una imagen para rellenar espacios faltantes y regresar los objetos a un tamaño similar al original.
- ❖ Obtención de contornos: Aquí se utiliza tanto dilatación como erosión, bien sea para encontrar el borde interior o exterior de la figura al realizar la operación y eliminar la imagen original de la misma.

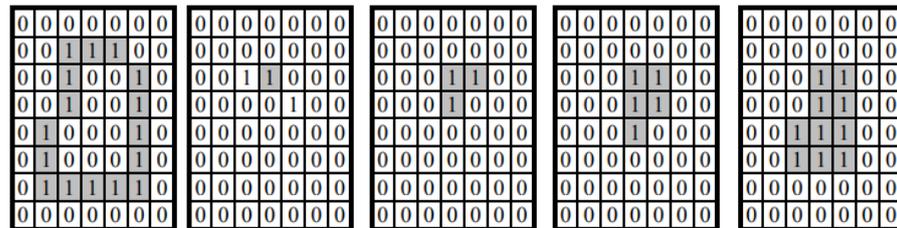
Figura 24
Proceso de apertura



Nota. (a) Imagen original; (b) EE vecinos a 8; (c) Apertura de la imagen (a) con b; (d) EE vecinos a 4; (e) Apertura de la imagen (a) con (d). Tomado de *Procesamiento morfológico en Conceptos y métodos en visión por computador* por Cancelas et al., 2016.

- ❖ Rellenado de regiones: aunque existe la operación de cierre, la misma solo es aplicable en regiones faltantes de poco tamaño. Este método se basa en algoritmos iterativos para ocupar espacios mayores. Al buscar un punto vacío y aumentar su dilatación hasta que se encuentre con la figura original. Representación gráfica en la figura 25.

Figura 25
Operación de relleno



Nota. Proceso de relleno aumentando desde un punto cada vez. Tomado de *Procesamiento morfológico en Conceptos y métodos en visión por computador* por Cancelas et al., 2016

Las operaciones más complejas se basan en encontrar patrones sobre la imagen para encontrar puntos de referencia, esqueletos o resaltar elementos irregulares.

- Descripción de texturas: se utiliza en varios problemas, aunque por su complejidad la descripción o catalogado deben ser más simples. Siendo posible clasificar los descriptores que realizan esta tarea por la forma en que se etiquetan las mismas (García-Olalla y Alegre, 2016).
 - Microtexturas o macrottexturas.
 - Irregulares o regulares.
 - Repetitivas o no repetitivas.

El principal exponente de esto es el método de Local Binary Pattern (LBP), que analiza cada pixel y el vecindario asignado estudiando si el mismo supera un determinado umbral, se destaca por sobre los demás métodos por la robustez que este posee frente a variaciones en la escala de grises por diferencias en la iluminación, entre otros factores.

- Segmentación de regiones: según Guijarro et al. (2016) la segmentación de regiones se utiliza para dividir una imagen en sus distintos componentes y separarlas en varias, en

a las que se resalte el elemento principal. No existe un solo criterio para realizarla o elegir sus juicios de valor, algunos de estos métodos son:

- ❖ Binarización: no se refiere a un único método si no a aquellos que utilizan umbrales para realizar histogramas y realizar su división en base a la intensidad de los mismos.
- ❖ Etiquetado de componentes conexas: aquí se consideran componentes conexas aquellos que posean un camino por el que se pueda recorrer sin salirse de la región. Cada uno de estos forman un conjunto que se etiqueta y segmenta.
- ❖ Crecimiento y división: utiliza una semilla por la cual se va expandiendo un espacio entre las regiones para eventualmente eliminar las conexiones para poder dividir las regiones en base a los elementos eliminados.
- ❖ Extracción de regiones por color: aquí no se habla de una sola técnica ya que podemos basarnos en distintos tonos o colores predominantes en la imagen.

Los distintos lenguajes nos ofrecen diferentes operaciones y librerías además de las básicas, la mejor forma de aprender el uso de las mismas es mediante la experimentación para encontrar la más conveniente para nuestros propósitos.

En la actualidad existen varias librerías de código especializadas para implementar funciones de visión por computador, algunas de estas son: OpenCv, Pillow, Scikit-image, Matlab, sus descripciones según su documentación oficial son:

- OpenCV: librería de código abierto que incluye varios algoritmos basados en C++, esta funciona de manera modular con los siguientes módulos: Procesamiento de imagen, Análisis de video, Calibración de cámara y reconstrucción 3D, Framework de características 2D, Detección de objetos, GUI, Video I/O.
- Pillow: añade características de procesamiento de imágenes a Python, además de esto también posee usos principales como archivar y mostrar imágenes, al incluir herramientas como PhotoImage o BitmapImage para esto.
- Scikit-image: trabaja sobre scipy.ndimage librería Python, y crece al usar proposiciones de la comunidad que pasan por un proceso antes de ser aceptas e implementadas sobre la librería principal.
- Matlab: al ser su propio lenguaje aquí es posible utilizar tanto sus librerías como por ejemplo simulink e incluso apoyarse en opciones externas como las antes mencionadas al usar herramientas de integración y APIs interoperables.

1.3 Redes Neuronales Convulocionales

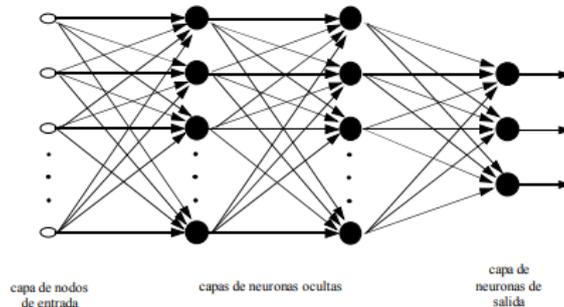
1.3.1 Redes Neuronales

Se puede realizar una interpretación propia sobre las redes neuronales o también conocidas como perceptrones multicapa basándonos en las definiciones que le dan distintos autores. Matich (2001) nos ofrece dos explicaciones, comentando que las podemos definir tanto, como un modelo matemático cuyos elementos están organizados en niveles, o de forma más compleja que son elementos de procesos interconectados, que procesan información mediante su estado en base a estímulos externos. También existen interpretaciones menos técnicas como la que indica Vintimilla (2014) que dice que estas son simplemente algoritmos intentando replicar el pensamiento humano a través de la repetición.

Quizás, una definición completa basándonos en lo que se propone en la bibliografía, podría ser una combinación entre ambas clases tanto la técnica como la simple, al decir que se

trata de varios nodos o neuronas interconectados para intentar replicar el cerebro humano y su razonamiento. Pero como dicen Acevedo et al. (2007) aunque cada definición es importante para entenderlas realmente deberíamos aprender cómo funcionan y realizan sus funciones. En este caso para entenderlas usaremos un modelo presentado por Vintimilla (2014) representado en la figura 26.

Figura 26
Modelo de un MLP



Nota. Tomado de *Desarrollo e implementación de una aplicación que traduzca el abecedario y los números del uno al diez del lenguaje de señas a texto para ayuda de discapacitados auditivos mediante dispositivos móviles android* por Vintimilla, 2014.

Como se puede observar en el caso del modelo mencionado anteriormente, el mismo consta de cuatro capas (pueden ser más o menos capas), y las mismas están compuestas por neuronas, en orden las clases de estas capas son:

- Capa de entrada: aquellas que reciben la información externa.
- Capas ocultas o intermedias: se encarga de realizar el trabajo es decir analizar, se les da este nombre debido a que desconocemos las entradas y salidas de las mismas.
- Capa de salida: proporciona el resultado del trabajo de las neuronas anteriores.

1.3.1.1 Modelos.

Existen varias maneras de clasificar los tipos de redes neuronales, la forma principal de clasificarlas es por los dos métodos de aprendizaje que existen, que puede ser supervisados o no supervisados.

Algunos de los modelos o topologías principales a tomar en cuenta son los convolucionales (CNN), recurrentes (RNN) o de base radial (RBF).

- CNN: surge como variación del perceptrón multicapa, usada principalmente para algoritmos de percepción visual, sobre la base que existe se añaden las capas de convolución y pooling las cuales repiten un proceso para obtener conjuntos de matrices (Ghosh et al., 2019).
- RNN: se identifican por intentar establecer una correspondencia entre sus secuencias de entrada y salida. Dentro de su grupo estas se pueden diferenciar entre totalmente recurrentes y parcialmente recurrentes, en el primer caso las conexiones entre las neuronas son variables mientras que en el segundo estas son estáticas (Bonet et al., 2007).

- RBF: llamadas así por las funciones que utilizan que tal como su indica son de base radial, usada mayoritariamente en aplicaciones de tiempo real por la particularidad de que poseen una única capa oculta (Cevallos, 2008).

1.3.1.2 Funciones de Activación

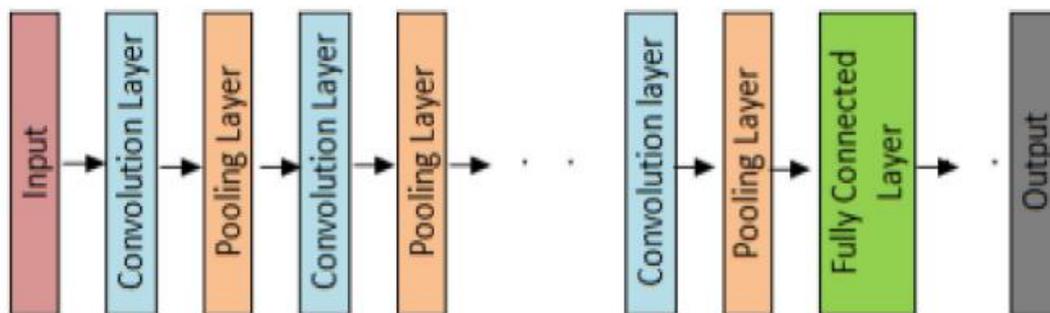
La función de activación es lo que define la salida de cada neurona y transmite la información de entre sus conexiones, existen varias de estas funciones, pero las más comunes son, sigmoide, tanh y Relu, siendo esta última la más usada en las CNN.

1.3.2 CNN

Los modelos de CNN consisten en un proceso finito de capas de procesamiento que mediante la abstracción aprenden y extraen características, más o menos específicas dependiendo del número de capas, en la figura 27 podemos observar el modelo conceptual básico de una CNN con sus distintas capas.

Figura 27

Modelo básico de una CNN



Nota. Tomado de Fundamental concepts of convolutional neural network por Ghosh et al., 2019.

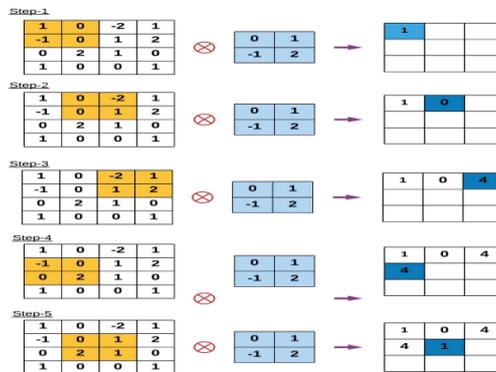
Tal como se menciona, estas aprenden y mejoran, esto se realiza a través del proceso conocido como entrenamiento o ajuste, que según Matich (2001) se define como la transformación de los pesos de la red en base a la información que se le proporciona. Para realizar esta actividad se pueden utilizar distintas herramientas, bien sean los lenguajes orientados a la inteligencia artificial como Matlab utilizado por Adem et al. (2022), las librerías existentes, como las que se aplican en Python empleado por Słowiński (2021) o Garcia et al. (2022), o el usar APIs las cuales se valen de las librerías y el lenguaje para facilitar el proceso como lo realizó Chicaiza (2022).

1.3.2.1 Componentes

- Capa convolucional: considerada la capa más importante de este modelo contiene varios núcleos o filtros, donde se asignan pesos en un principio aleatorios, que junto con la imagen de entrada genera mapas de características. En una operación de convolución utilizamos un filtro que pasa por la imagen o parte de la misma para obtener mapas de características podemos observar un ejemplo de este funcionamiento en la figura 28.
- Capa de pooling: en esta capa se toman los mapas de características producidas anteriormente y permite reducir los mapas de características y al simplificarlos

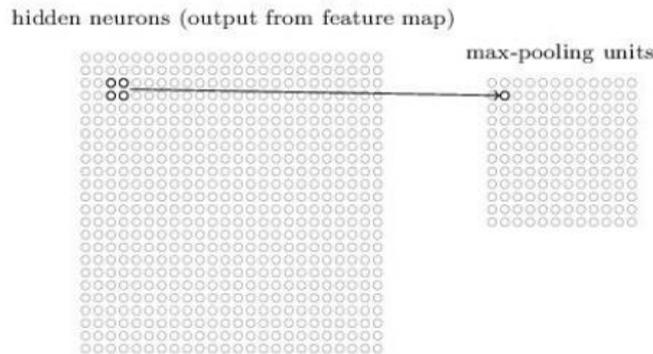
evitar el sobreajuste, no existe un único método para llevar a cabo esta capa los más comunes son el max pooling que toma el mayor punto de cada región mapa y el promedio en el que se toman como indica el nombre los promedios de la región para hacer un nuevo punto, un ejemplo del mapa reducido se puede observar en la figura 29.

Figura 28
Proceso de la capa convolucional



Nota. Tomado de Fundamental concepts of convolutional neural network por Ghosh et al., 2019.

Figura 29
Proceso de la capa de Pooling al aplicar max pooling



Nota. Tomado de Redes convolucionales por Bonilla, 2020.

- Capa de clasificación: esta sería la parte más parecida a las redes neuronales tradicionales, funciona como clasificador de la última capa convolucional o de pooling, toma las decisiones en base a los pesos y traduce esto hacia la capa de salida, es decir la clasificación.

1.3.3 Aplicaciones de las CNN

Al haberse desarrollado durante varios años y tener importantes exponentes tales como AlexNet, YOLO, o LeNet, podemos encontrar el uso de este modelo en varias áreas, como pueden ser:

- Clasificación de imágenes
- Reconocimiento de texto
- Análisis médico

- Seguridad y vigilancia
- Entre otras

1.4 Trabajos relacionados

Para poder determinar ciertas herramientas y técnicas a utilizar en este trabajo tomaremos en cuenta lo que han realizado distintos autores en el ámbito de clasificación mediante el uso de CNN o en agricultura de precisión, que, aunque se refiere a una amplia gama de técnicas y aplicaciones, en este caso nos centraremos en aquellos que entren dentro del tema de la detección de defectos mediante el uso de inteligencia artificial. Aunque aquí solo se analizaran algunos de los más relevantes, mencionaremos más y sus métodos al momento de discutir los procesos realizados y las decisiones tomadas sobre los mismos, así como las herramientas a usar.

1.4.1 A sugar Beet Leaf Disease Classification Method Based on Image Processing and Deep Learning.

En este trabajo realizado por Adem et al. (2022) se llevó a cabo una clasificación de hojas de remolacha, para detectar puntos en las mismas, producidos por *Cercospora beticola* Sacc. Para desarrollar la investigación, los autores utilizaron, RCNN, SSD, VGG16 y Yolov4, que junto con técnicas de deep-learning ayudaron para entrenar y probar los modelos, utilizando un dataset de 1040 imágenes.

El trabajo se realizó en una computadora con un procesador Intel i7 6700, 16 GB de ram, y una tarjeta gráfica Nvidia GTX 960 M, sobre el lenguaje de Matlab.

En lo que respecta al pre-procesamiento de las imágenes las mismas se cambiaron a xyz y después a luv color utilizando los mismos valores, para luego descomponerlo y usar un histograma sobre el canal "v" de este.

1.4.2 Dry Beans Classification Using Machine Learning

Trabajo de análisis que se llevó a cabo utilizando Python sobre la versión gratuita de Google collab. Se empleó un dataset de 13611 muestras de distintas, para la clasificación solo se tomó en cuenta la forma de los granos más no el color, esto, aunque impráctico es suficiente para el objetivo planteado de realizar una comparación, se emplearon distintas técnicas para esto, como Naive Bayers, Clasificador de vector (SVM), arboles de decisión, Random Forest, y Redes Neuronales. Al comparar estos métodos el autor encontró tanto el clasificador con menor precisión, así como los granos más fáciles de identificar dentro de las 7 especies utilizadas (Słowiński, 2021).

1.4.3 Review of Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification

Lu et al. (2021) nos presentan en este artículo una revisión sobre algunos métodos de CNN aplicados a la detección de enfermedades en plantas, para esto señalan el proceso que estas suelen seguir además de las dificultades que estos pueden afrontar, como datasets insuficientes, aquí como conclusión se nos presentan varios trabajos, el tamaño de sus datasets, el método que emplearon, la precisión que se alcanzó y, los pros y contras de estos, como podemos ver en la tabla 1.

Tabla 1
Estudios que aplican métodos de CNN en identificación de enfermedades de plantas

Tarea	DatASET	Método	Precisión	Pros y contras
Identificación de 14 cultivos y 26 enfermedades	54306 imágenes del dataset PlantVillage	AlexNet, GoogleNet	99.35%	No es útil para la aplicación práctica
Identificación de pestes y enfermedades en plantas de tomate	5000 imágenes tomadas en distintas circunstancias	VGGNet y ResNet	83%	La precisión es menor en aplicación práctica
Identificación de enfermedades re arroz y maíz	500 imágenes de arroz y 466 de maíz	VGGNet, Inception	92%	Trabajos futuros se centraran en utilizarlo mediante dispositivos móviles
Identificación de 4 enfermedades de hoja de manzana	13689 imágenes de hojas enfermas	Arquitectura basada en AlexNet	97.62%	Se generó una red neuronal robusta
Clasificación de enfermedades de hoja de tomate	14828 imágenes	AlexNet, GoogleNet	99.18%	No hubo suficientes muestras
Diagnóstico de enfermedades de maíz	15408 imágenes de kaggle	VGG16 y 19	98.2%	Se necesitó más diversidad en el dataset
Reconocimiento de dos enfermedades de hoja de maíz	Imágenes de PlantVillage e imágenes a tiempo real	DCNN	88.46%	Solo dos enfermedades reconocidas y clasificadas
Detección y diagnóstico de enfermedades de plantas	87848 imágenes de PlantVillage	MCT, AlexNet, GoogleNet	99.53%	Alto porcentaje de éxito

Nota. Lu et al. (2021)

1.4.4 Smart Agriculture: Real-time Classification of Green Coffee Beans by Using a Convolutional Neural Network

Huang et al. (2020) mediante el uso de una cámara y una cinta transportadora recolectaron las imágenes y se entrenó una CNN. Ya que las imágenes estaban en movimiento la precisión disminuyó. El entrenamiento y pruebas se realizaron utilizando un procesador i5, una GPU NVIDIA Geforce GTX 1080 Ti y una memoria de 48 Gigas.

Sobre el dataset de aproximadamente 1000 imágenes se realizó una segmentación, al cambiar las imágenes a escala de grises y remover el fondo, fue esto lo que pasó por la red neuronal. Las conclusiones a las que llegaron los autores fueron los posibles errores al utilizar escala de grises, así como la imposibilidad de usar de forma práctica el sistema completo, ya que únicamente uno de los dos lados del grano pasa por la cinta.

1.4.5 An automatic Bean Classification System Based on Visual Features to Assist the Seed Breeding Process

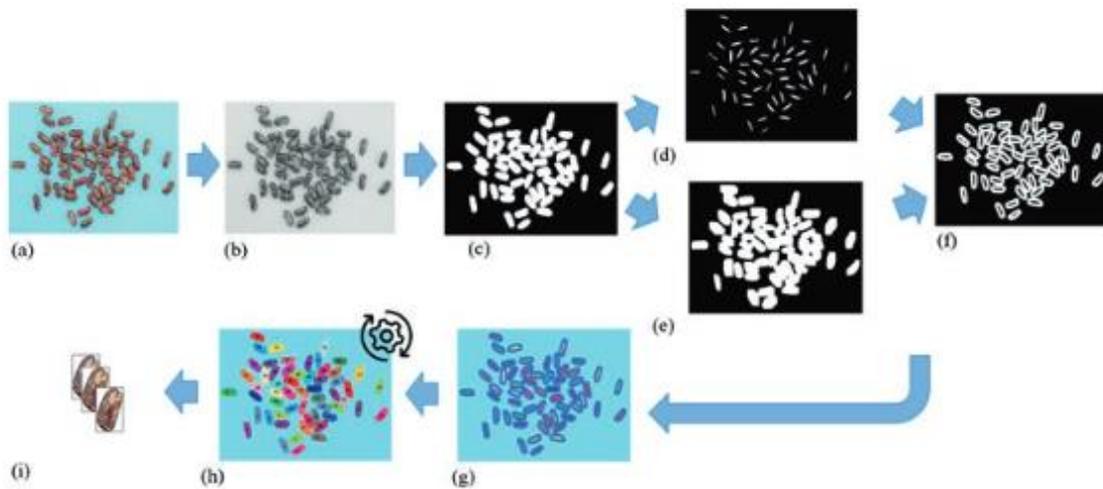
Trabajo realizado por Garcia et al. (2022) que toma 6 variedades de frijoles para su clasificación. Basándose en trabajos relacionados realizaron un clasificador random forest en el que les asignaron pesos a distintas características como área, textura o forma.

Este trabajo recolecto su dataset de 46440 fotografías, tras data augmentation, mediante el uso de una cámara Nikon a 30 cm del objetivo con distintos puntos de iluminación, después de la adquisición de las imágenes estas se segmentaron con el proceso de la figura 30.

El sistema codificado en Python llego a alcanzar una precisión de 98.5 con el mejor clasificador desarrollado.

Figura 30

Proceso de segmentación y separación de canales



Nota. Tomado de Redes neuronales recurrentes para el análisis de secuencias por Garcia et al., 2022.

Capítulo 2 Desarrollo

2.1 Adquisición de imágenes

Se adquirió un total de 3276 imágenes entre granos sanos y defectuosos con un porcentaje de 46.06% y 53.94% respectivamente. Se tomaron las fotografías empleando una cámara digital “Nikon D3400” sobre un trípode a una distancia 14 cm del objeto sobre un fondo blanco como se ve en la figura 31 y en el esquema (Figura 32), con una lente AF-P DX NIKKOR 18-55mm f/3.5-5.6G VR aplicando una distancia focal de 55mm y una apertura de 5.6, valores alcanzados mediante prueba y error hasta obtener una imagen clara en la que se aprecien los defectos del grano.

Figura 31

Proceso de captura de imágenes de los granos de café



Figura 32

Esquema del proceso



Las imágenes obtenidas de este proceso se recortaron para eliminar objetos extraños y centrarlas utilizando el programa GIMP, seleccionado al ser libre y gratuito además de poseer la función necesaria para facilitar el proceso de recorte, la repetición de acciones sobre grupos de

imágenes, tomando en cuenta que pocas características fueron fundamentales, esta opción sobresalió entre las más complejas. Tras esto, se etiquetaron las mismas utilizando la aplicación online “CVAT”. Se decidió utilizar esta herramienta, tomando como base la comparación realizada en las tablas 2 y 3 presentadas por Lima de Sousa (2021).

Se optó por “CVAT” considerando los criterios de gestión y visión por computador presentados por las tablas 2 y 3 y descartando las opciones que no fue posible utilizar. Primero se contempló el uso de “diffgram” por su capacidad colaborativa y resúmenes estadísticos sobre las actividades de etiquetado, esta opción se descartó por el límite de imágenes en su versión gratuita haciendo imposible su uso. Otra alternativa que se tomó en cuenta fue “labelMe” que se eliminó pues su versión online ya no estaba disponible y únicamente funciona como aplicación de escritorio lo que dificultaría o imposibilitaría la colaboración.

Tabla 2

Resumen comparativo de las características disponibles en cada herramienta

Nombre de la herramienta	Visión por computador	Procesamiento de lenguaje	Automatización y desarrollo	Administración y QA
Colabeller	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%

Nota. Lima de Sousa (2021)

Con apoyo del experto se clasificaron las imágenes en dos etiquetas, “bueno” y “malo”, esto pensando en su uso general a largo plazo pues, aunque es útil conocer el tipo de defecto del grano, solo se necesitan estas dos clases para la separación ya que no todos los agricultores le dan el mismo fin al producto de menor calidad.

Tabla 3*Características de QA y gestión en las distintas herramientas*

Nombre de la herramienta	Gestión de calidad	Gestión de proyecto	Gestión de datos	Consenso	Benchmark	Métricas de rendimiento
Colabeller	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

Nota. ✓=disponible X=no disponible ¿=inconcluso. Lima de Sousa (2021)

El dataset en formato tfrecord y como imágenes etiquetadas y recortadas se pueden encontrar en [este link](#).

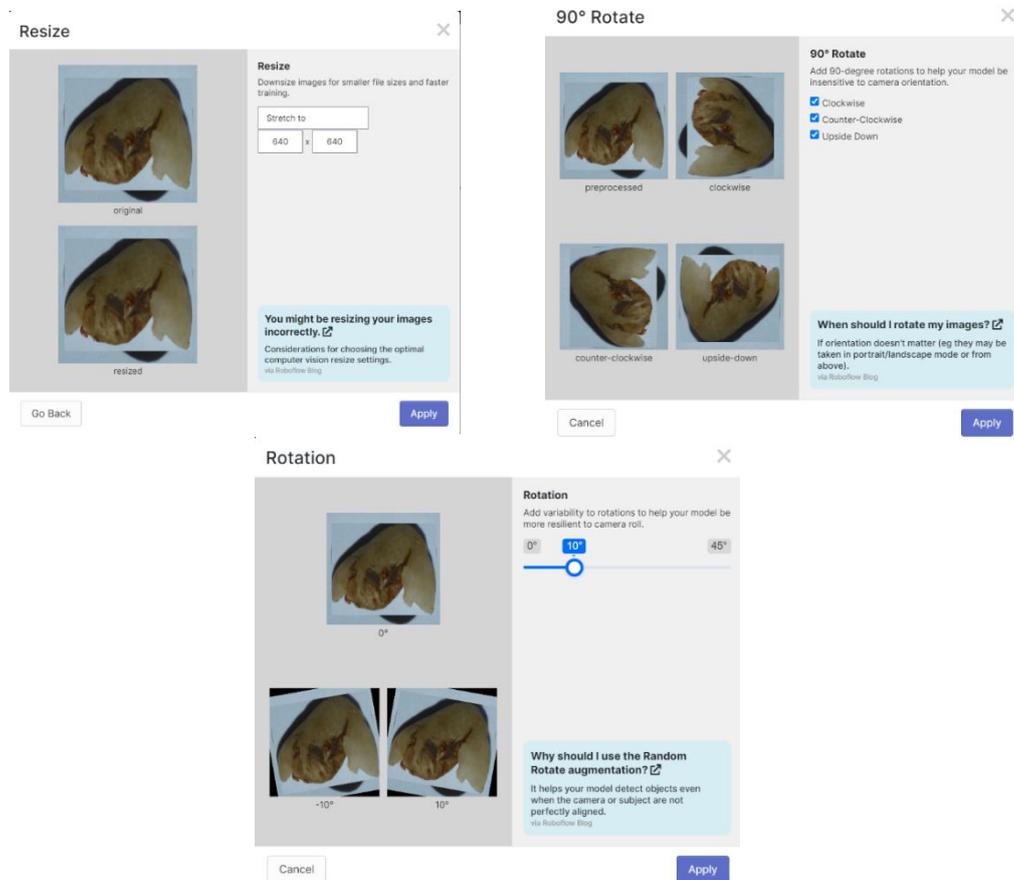
2.2 Preprocesamiento de imágenes

Tras terminar el etiquetado de las imágenes es posible descargar las mismas en distintos formatos basados en varias herramientas, como “LabelMe” o “COCO”, u modelos como “YOLO”, en este caso se descargó el dataset en el formato “Pascal VOC” el cual devuelve tanto la imagen como un archivo xml en donde se detalla el nombre y dimensiones de la imagen, y la clasificación y coordenadas del bounding box. Se eligió este formato por la herramienta utilizada para el preprocesamiento.

Se eligió la herramienta roboflow para realizar las transformaciones necesarias como se realizó en los trabajos de Chicaiza (2022), Haque et al. (2022) o Jubayer et al. (2021). Mediante su uso se realizaron las acciones de, cambio de tamaño y rotación. La primera como opción para agilizar el proceso de entrenamiento al reducir el peso y utilizar el tamaño óptimo de los modelos seleccionados para entrenamiento, la segunda opción, aplicada sobre la imagen entera

o sobre el bounding box, se aplicó con el fin de realizar data augmentation es decir, ampliar el dataset y darle variabilidad sin necesidad de recolectar más imágenes. Tras estos procesos (figura 33) nuestro dataset en un inicio de 3276 muestras paso a 20284, más de seis veces su tamaño original, estas se dividieron en 82% para entrenamiento, 12% de validación y 7% para pruebas.

Figura 33
Procesos de data augmentation realizados en roboflow



2.3 Arquitectura del modelo

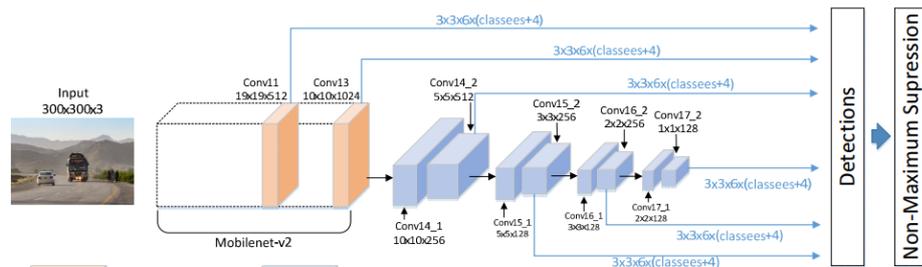
Para este trabajo se decidió utilizar la API de detección de objetos de tensorflow, esta resolución se tomó en base a que dentro de esta existen varios modelos que la misma soporta, además de que se considera también que la compatibilidad con la herramienta tensorboard lo que facilitó la obtención de varios gráficos. Los modelos compatibles con la misma se encuentran el repositorio de github “Tensorflow 2 model Detection model zoo”, de entre los disponibles aquí se escogieron dos modelos sobre los que trabajar, uno ligero, “SSD Mobilenet V2”, y uno más profundo, “SSD resnet50 v1”.

Se determinaron estos en base a los trabajos relacionados ya que dentro de los mismos podemos encontrar ambas clases sin una clara preferencia

2.3.1. SSD Mobilenet

Según Chiu et al. (2020) esta es una red neuronal ligera basada en la más pesada, VGG 16 al reducir su complejidad con el objetivo de aumentar la velocidad de procesamiento y entrenamiento de la misma. Se eligió para este trabajo por estas mismas razones con el fin de que sea posible probar rápidamente distintos cambios sobre su proceso de ajuste o al reentrenarse en caso de modificar nuestro dataset. En la figura 34 podemos ver la arquitectura del modelo, 6 capas de las que el mismo se compone.

Figura 34
Arquitectura mobilenet

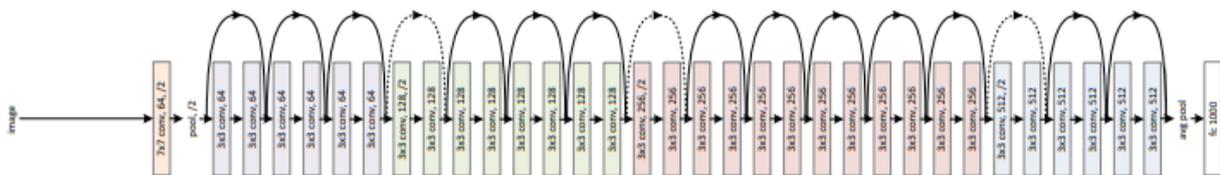


Nota. Tomado de *Mobilenet-SSDv2: An Improved Object Detection Model for Embedded Systems* por Chiu et al., 2020.

2.3.2. SSD Resnet 50

Resnet o “Residual Network” se refiere como su nombre indica a una red que se ensambló usando partes de distintos modelos además la misma aplica saltos entre sus capas, volviéndola no lineal (Kaiming et al., 2006). Se eligió Resnet50 para nuestro modelo pesado de entre las opciones disponibles en model zoo por el número de clases que tenemos en este caso solo dos por lo que nos favorece el uso de una las alternativas más simples. La figura 35 nos muestra una versión de 34 capas de Resnet.

Figura 35
Arquitectura resnet 34 capas



Nota. Tomado de *Deep residual learning for image recognition* por Kaiming et al., 2006

2.4 Entrenamiento del modelo

Ya que se empleó la API de detección de objetos de tensorflow en primer lugar se descargó nuestro dataset aumentado y dividido desde Roboflow en el formato compatible que en este caso

consiste de un archivo .pbtxt con las etiquetas que existen como se ve en la figura 36 y un documento .tfrecord que consiste tanto de las imágenes, así como sus bounding boxes.

Figura 36

Documento label_map

```
1 item {
2     name: "badrotation",
3     id: 1,
4     display_name: "badrotation"
5 }
6 item {
7     name: "goodrotation",
8     id: 2,
9     display_name: "goodrotation"
10 }
11 }
```

Para entrenar, los modelos descargados nos proporcionan un archivo de configuración en el que encontramos parámetros editables de estos cambiamos el número de clases a dos, el tipo de trabajo a detección y el tamaño de las imágenes se mantuvo pues las nuestras ya aplicamos este cambio con anterioridad. Otro cambio realizado de manera general son las ubicaciones tanto de las partes de nuestro dataset así como del modelo a utilizar, campos en un principio vacíos como se muestra en la figura 37. Además de estos Chicaiza (2022) nos describe otros a campos a considerar como:

Figura 37

Campos a modificar en archivo config

```
fine_tune_checkpoint: "PATH_TO_BE_CONFIGURED"
num_steps: 50000
startup_delay_steps: 0.0
replicas_to_aggregate: 8
max_number_of_boxes: 100
unpad_groundtruth_tensors: false
fine_tune_checkpoint_type: "classification"
fine_tune_checkpoint_version: V2
}
train_input_reader {
  label_map_path: "PATH_TO_BE_CONFIGURED"
  tf_record_input_reader {
    input_path: "PATH_TO_BE_CONFIGURED"
  }
}
```

- **Batch Size:** también conocido como tamaño de lote en español este determina la cantidad de muestras que se procesan a la vez en cada una de las iteraciones durante el entrenamiento, esto afecta directamente a el tiempo y eficiencia del mismo. Para nuestro trabajo esto se modificó su valor a 1 en el caso de Resnet y para mobilenet se mantuvo el valor por defecto, 128.
- **Optimizer:** o la función de optimización es la que cambia los pesos de nuestro modelo al final de cada época en base a los resultados obtenido. Existen varias tales como ADAM, momentum o "RMSProp. En nuestro caso se aplicó la función de optimización momentum opción por defecto en ambos casos.

- Activation: como se explica en el capítulo anterior existen varias de estas cada una enfocada a distintas funciones como la sigmoide que es utilizada para clasificación binaria con una sola clase. Aquí aplicamos la función ReLu tanto por ser esta la más común además de que en los trabajos relacionados esta también fue bastante utilizada.

Habiendo realizado los cambios necesarios sobre los parámetros de nuestra configuración, actividad que se puede realizar a mano sobre el archivo o mediante el uso de comandos en Python, se procedió al entrenamiento el cual se realizó sobre una maquina con un procesador Intel i7 10750H, procesador de decima generación, 16 GB de ram, y una tarjeta gráfica Nvidia GTX 1650 ti, sobre el lenguaje ya antes mencionado.

Se inició con el ajuste de mobilenet con imágenes 320x320 px, en el que se utilizó en un principio 5000 épocas mediante el uso de línea de comando con el código de la ecuación 1. A continuación se corrió el script de la API, *model_main_tf2.py*, al que se le proporciona los campos de *model_dir*, la dirección donde se guardarán varios ciclos, cada mil es la configuración por defecto, esto sirve para continuar desde uno de estos puntos en caso de que se pare el proceso o para exportar el modelo, *pipeline_config_path*, la dirección de nuestro archivo de configuración, *num_train_steps*, el número de iteraciones para entrenamiento, aunque este campo también se puede modificar en el archivo config la línea de comando sobrescribirá esto y es más simple cambiar el campo en este momento. El primer intento demoró 45 minutos (figura 38). Tras este se probaron con 10000 para ver si mejoraba su rendimiento, este tardó casi dos horas (figura 39).

```

py Tensorflow/models/research/object_detection/model_main_tf2.py
--model_dir
= Tensorflow/workspace/models/my_ssd_mobnet
--pipeline_config_path
= Tensorflow/workspace/models/my_ssd_mobnet/pipeline.config
--num_train_steps = 5000

```

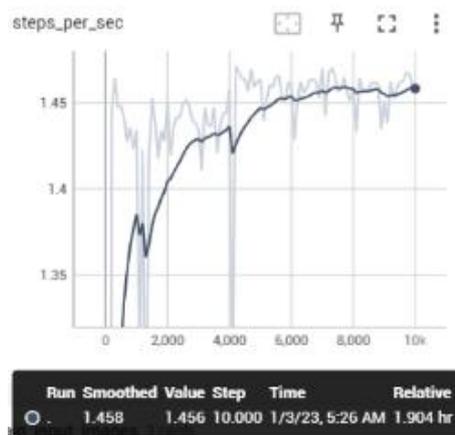
(1)

Figura 38

Gráfico segundos por época mobilenet 5000



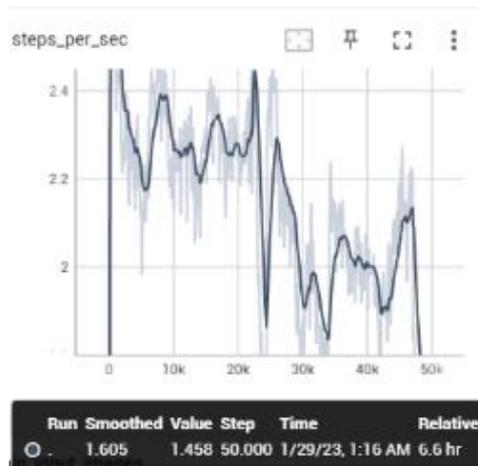
Figura 39
Gráfico segundos por época mobilenet 10000



Después de estos intentos es cuando se cambiaron las imágenes a 640x640 px y se probó de nuevo con 5000 y 10000 iteraciones tardando tiempos similares, pero con mejores resultados, entre estos dos se tomó el primero para exportar y realizar nuestras pruebas finales, esto debido a que en un principio se sospechaba de overfitting.

Tras el procedimiento anterior se iniciaron las pruebas con resnet de igual manera con 5000 y 10000 épocas, sin embargo, debido a los bajos resultados se aplicaron 50000, lo cual mejoró considerablemente el rendimiento, este último tardó más de seis horas (figura 40).

Figura 40
Gráfico segundos por época resnet 50000



Los resultados de las pruebas de ambos modelos se expondrán y discutirán en el siguiente capítulo.

Capítulo 3

Resultados y validación

Los datos y gráficos discutidos y expuestos en este capítulo se obtuvieron mediante tensorboard, la API de tensorflow, su script de evaluación, y otro llamado `confusion_matrix_tf2` obtenido en github https://github.com/svpino/tf_object_detection_cm.

3.1 Métricas cuantitativas

Al ser este un problema de detección tomaremos en cuenta los datos de precisión, f1-score, recall y la matriz de confusión.

3.1.1 Matriz de confusión

El contenido de la misma consta de $m \times m$ campos siendo m la cantidad de clases en las que es posible clasificar, dentro de la matriz encontramos una diagonal con valores correctos y aquellos que se encuentran fuera de la misma son denominados celdas de confusión es decir aquellos que fueron asignados erróneamente en la clase (Ariza-López et al., 2018). En base a esta es posible determinar los distintos campos antes mencionados.

Como se señaló anteriormente se ejecutaron las matrices de confusión con el script de `confusion_matrix_tf2.py` mediante el comando ejecutado de la ecuación 2. Que consta de `input_tfrecord_path`, la dirección de nuestro archivo `.tfrecord` del dataset, `output_path`, el nombre y dirección del archivo que se genera al final de proceso, `inference_graph`, la ubicación del modelo exportado, y `class_labels`, dirección del documento `.pbtxt`. Al ejecutar el comando se obtiene la matriz de confusión, así como un archivo en el que se almacenan los resultados.

```
py RealTimeObjectDetection --main\confusion_matrix_tf2.py
  --input_tfrecord_path
  = Tesis3.v1i.tfrecord\valid\coffee.tfrecord
  --output_path
  = confusion_matrix.csv
  --inference_graph
  = RealTimeObjectDetection --main\exported
  --models\resnet5050\saved_model
  --class_labels = Tesis3.v1i.tfrecord\valid\coffee_label_map.pbtxt
```

(2)

Se aplicó sobre la versión de 5000 épocas de mobilenet y en resnet a sus 50000, los resultados de estas pruebas se observan en la figura 41 y la figura 42 respectivamente.

Figura 41

Matriz de confusión mobilenet

```
Confusion Matrix:
[[1242.  12.  13.]
 [ 288. 555. 195.]
 [  12.  19.   0.]
```

Figura 42

Matriz de confusión resnet

```
Confusion Matrix:  
[[1034.  55. 178.]  
 [ 425. 177. 436.]  
 [   8.  15.  0.]]
```

Estas matrices se interpretan de la siguiente manera, en la primera fila se encuentran las muestras identificadas como malas y en la segunda aquellas que se reconocieron como buenas, y de igual manera con las columnas. Como se puede observar se encuentra un patrón en los modelos, que es que tiene dificultades con aquellos granos marcados como buenos, ya que existe una gran cantidad que se clasifican de manera errónea.

3.1.2 Precisión

Al necesitar la clasificación de ambas clases se realizó un cálculo para ambas en, en la siguiente parte de este capítulo se analizará de manera más general el modelo con precisión de clasificación y localización.

La forma en que se encontró la precisión es mediante la fórmula de la ecuación 3 tomada de Lukic et al. (2020) citado por Chicaiza (2022), en la que TP es el número de imágenes clasificadas correctamente en la clase y FP aquellas que no, aquí se consideró también nuestra tercera fila y columna que son ocasiones en que se intuyó un objeto y se clasificó, pero no existía nada en esa posición para calcular esto el script de la matriz de confusión toma en cuenta un IOU del 50%.

$$\text{Precisión} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

En el primer modelo se obtuvieron los siguientes resultados.

$$\text{Precisión granos malos} = \frac{1242}{1242 + 288 + 12} = 0.805447$$

$$\text{Precisión granos buenos} = \frac{555}{555 + 12 + 19} = 0.947099$$

En el segundo, las cotas fueron las siguientes.

$$\text{Precisión granos malos} = \frac{1034}{1034 + 425 + 8} = 0.704840$$

$$\text{Precisión granos buenos} = \frac{177}{177 + 55 + 15} = 0.716599$$

Como se observa hay una precisión relativamente alta en el primer caso, lo cual no lo es todo, en especial para los modelos presentados.

3.1.3 Recall

De igual forma se calculó recall en ambas clases este mismo da el porcentaje de una clase que el modelo es capaz de identificar correctamente.

Para el cálculo fue utilizada la ecuación 4 tomada de Lukic et al. (2020) citado por Chicaiza (2022). Los componentes de esta son TP, el número de imágenes clasificadas correctamente en la clase, y FN aquellas que pertenecen a la clase, pero fueron clasificadas de manera errónea.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

Los resultados del primer caso son

$$\text{Recall granos malos} = \frac{1242}{1242 + 12 + 13} = 0.980268$$

$$\text{Recall granos buenos} = \frac{555}{555 + 288 + 195} = 0.534682$$

Para resnet se obtuvo lo siguiente

$$\text{Recall granos malos} = \frac{1034}{1034 + 55 + 178} = 0.816101$$

$$\text{Recall granos buenos} = \frac{177}{177 + 425 + 436} = 0.170520$$

Al interpretar los datos se corrobora lo que se observó en las matrices de confusión, que para los modelos se dificulta el diferenciar los granos buenos ya que en ambos casos el recall de esta clase es bajo, en mobilenet por ejemplo de toda la clase de café apto solo se identificó correctamente un poco más de la mitad.

3.1.4 F1-score

Según Ordóñez y Roggen (2016) esta es la relación entre precisión y recall lo que facilita el análisis tomando en cuenta que se le dé la misma importancia a los dos campos. Este cálculo se realiza mediante la ecuación 5.

$$\text{F1 - score} = \frac{2 \cdot \text{recall} \cdot \text{precisión}}{\text{recall} + \text{precisión}} \quad (5)$$

Sabiendo esto, los siguientes resultados se obtuvieron para mobilenet.

$$\text{Granos malos} = \frac{2 \cdot 0.80 \cdot 0.98}{0.80 + 0.98} = 0.88$$

$$\text{Granos buenos} = \frac{2 \cdot 0.94 \cdot 0.53}{0.94 + 0.53} = 0.67$$

Y se alcanzaron los siguientes para el segundo modelo.

$$\text{Granos malos} = \frac{2 \cdot 0.70 \cdot 0.81}{0.70 + 0.81} = 0.75$$
$$\text{Granos buenos} = \frac{2 \cdot 0.71 \cdot 0.17}{0.71 + 0.17} = 0.27$$

Como se ve, en base a estos resultados la clasificación de los sujetos aptos aún necesita ser mejorada tanto en mobilenet como en resnet.

En la tabla 4 se encuentra un resumen sobre las métricas cuantitativas obtenidas en este capítulo

Tabla 4
Resumen métricas cuantitativas

Modelo	Precisión	Recall	F1-score
Mobilenet v2 Malo	0.80	0.98	0.88
Mobilenet v2 Bueno	0.94	0.53	0.67
Resnet 50 Malo	0.70	0.81	0.75
Resnet 50 Bueno	0.71	0.17	0.27

3.2 Gráficas de rendimiento

En primer lugar, se exponen los gráficos de entrenamiento obtenidos durante el mismo, al generarse un archivo que es observable gracias a tensoboard. Las figuras 43, 44 y 45 nos indican la pérdida total, de clasificación y localización de mobilenet respectivamente.

Figura 43
Pérdida total mobilenet

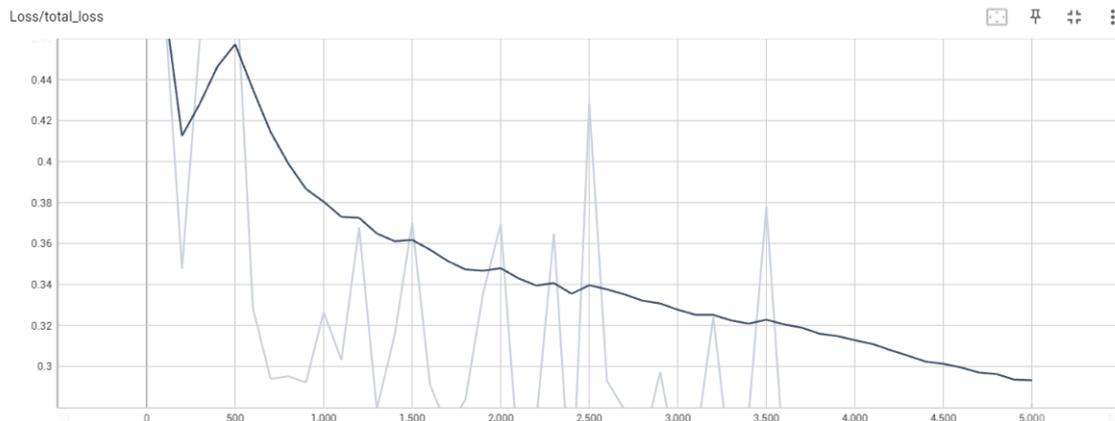


Figura 44
Pérdida de clasificación mobilenet

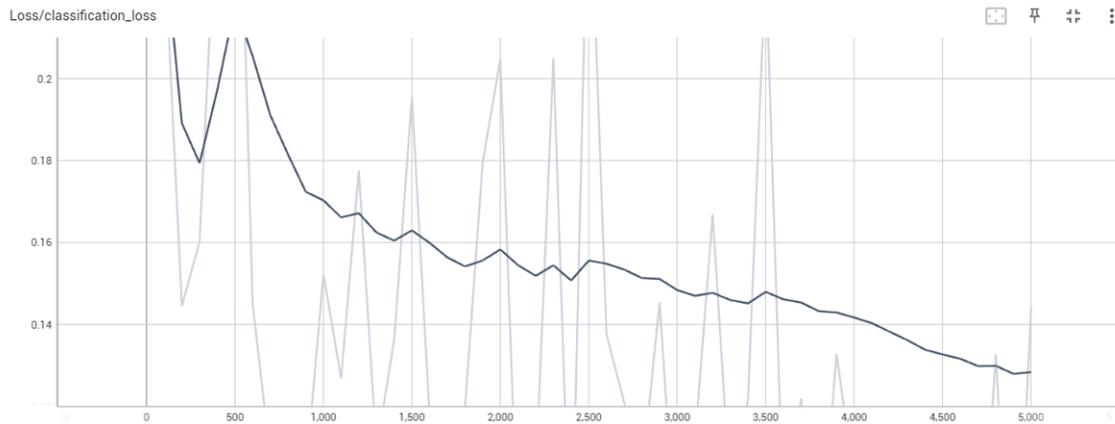
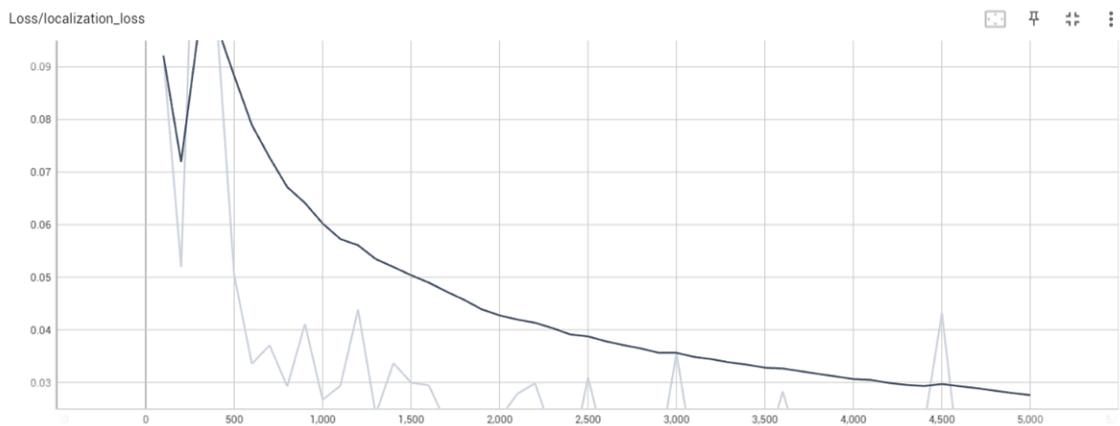


Figura 45
Pérdida de localización mobilenet



Con el mismo orden, las figuras 46, 47 y 48 que corresponden a resnet

Figura 46
Pérdida total resnet

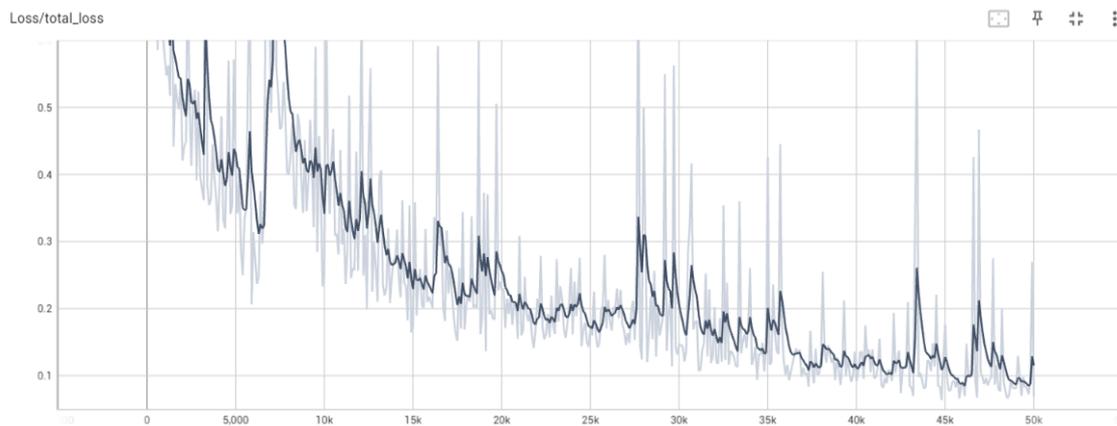


Figura 47
Pérdida de clasificación resnet

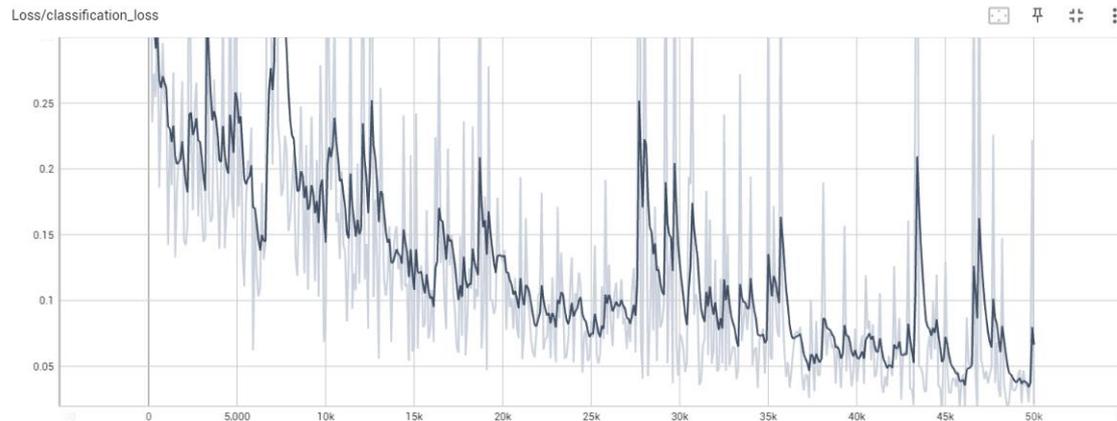
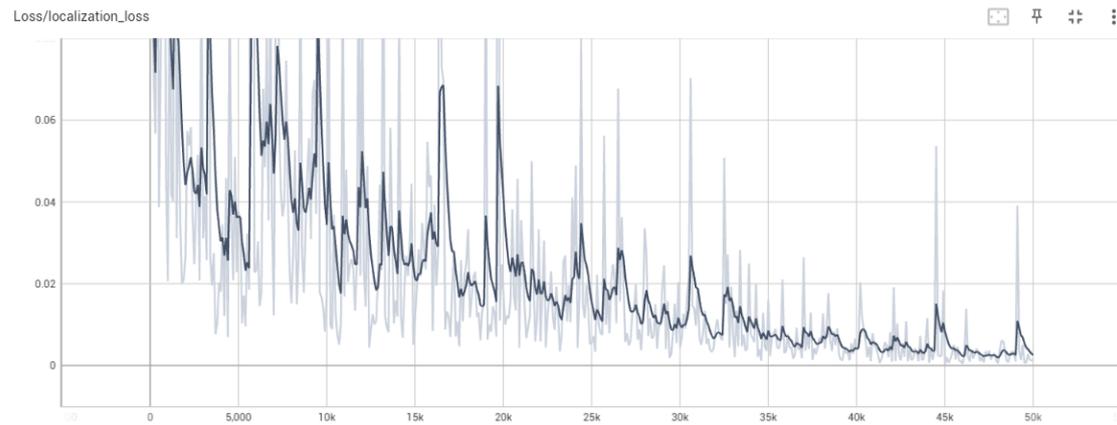


Figura 48
Pérdida de localización resnet



Mientras más cercano a cero mejor es la pérdida que representa el error, o la diferencia entre las predicciones y los datos reales, sin embargo, una vez más, para este trabajo no es lo único a tomar en cuenta.

Las siguientes figuras fueron obtenidas mediante el comando de evaluación de la ecuación 6, donde se utilizó el dataset de evaluación, siendo `model_dir` y `checkpoint_dir`, la dirección de los puntos de entrenamiento, y `pipeline_config_path`, la dirección del archivo de configuración. Esto genera de igual manera un archivo para visualizar mediante tensorboard.

```
py Tensorflow\models\research\object_detection\model_main_tf2.py  
  --model_dir  
  = Tensorflow\workspace\models\my_ssd_mobnet  
  --pipeline_config_path  
  = Tensorflow\workspace\models\my_ssd_mobnet\pipeline.config  
  --checkpoint_dir = Tensorflow\workspace\models\my_ssd_mobnet
```

(6)

En las figuras 49 y 50 se puede ver el map o “mean average precision”, y el gráfico de recall obtenido del primer modelo.

Figura 49
Map mobilenet

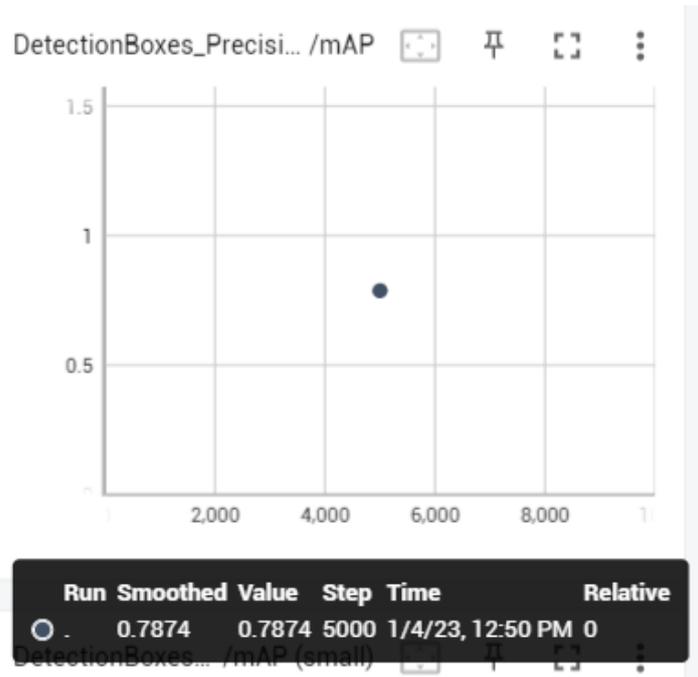
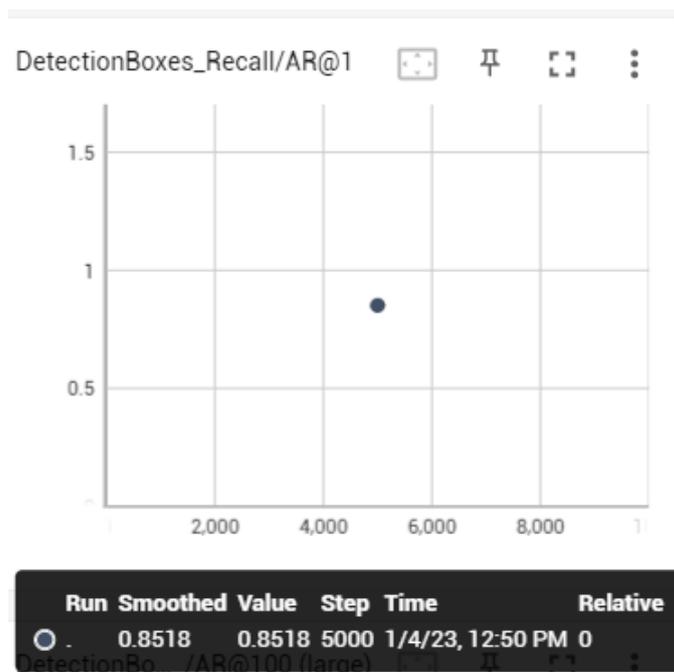


Figura 50
Recall mobilenet



Y con la misma disposición las figuras 51 y 52 corresponden a resnet.

Figura 51
Map resnet

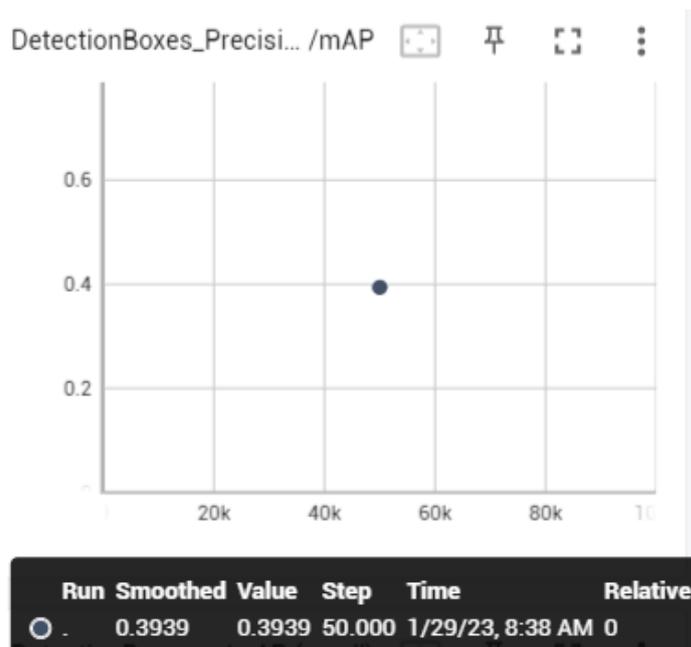
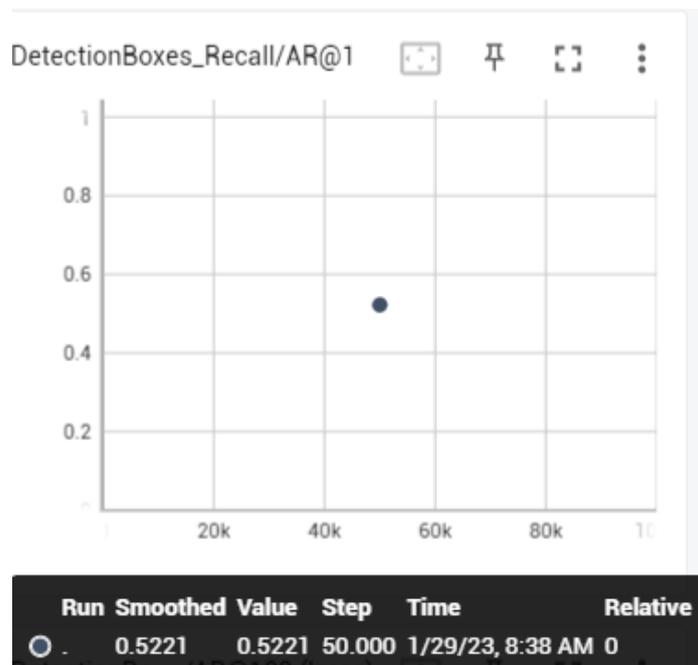


Figura 52
Recall resnet



Gracias a estos gráficos se ve que tanto por el cálculo de manera general como el cambio de dataset obtenemos resultados distintos a los encontrados anteriormente. Se analizarán los resultados obtenidos, así como las posibles causas de los mismos en la siguiente parte del capítulo.

3.3 Discusión

Para poder hablar sobre los resultados obtenidos y las posibles razones de los mismos se tomarán en cuenta las limitaciones del modelo, los resultados de trabajos relacionados, pruebas con individuos para ampliar la discusión al realizar una comparativa con la precisión obtenida, y algunos de los inconvenientes que se encontraron durante el proceso.

Hablando sobre las limitaciones que presentan los modelos. Durante pruebas de imágenes fuera del dataset al igual que con las que se presentaron anteriormente se encuentran problemas para la identificación de granos sanos, por este motivo se paró el entrenamiento de resnet al inferir una tendencia a pesar de la posibilidad de obtener resultados más altos. Además de esto se encuentran inconvenientes al clasificar imágenes con varios sujetos en las que los modelos no son capaces de trabajar correctamente en fotografías con más de tres. También se debe tomar en cuenta el peso y velocidad de inferencia que se alcanzó, en el caso de mobilenet es de 26.5 mb y 2.3 fps, mientras que en resnet es 248 mb y 2.8 fps. Aunque esta es una velocidad normal es posible que sea necesaria mejorarla en base al uso que se les de en un futuro

También existieron problemas durante el entrenamiento de resnet puesto que la visualización de imágenes en tensorboard llevo erróneamente a la conclusión de que el problema era el dataset en el ajuste de 5000 y 10000 épocas, ya que los colores de las fotografías se observaban de manera saturada. Antes de la utilización del equipo especificado se probó el uso del servicio de google collab, pero en su versión gratis no fue posible realizar estos procesos por su límite de tiempo en el que verifica con un captcha para evitar bots, esto quita la GPU asignada e interrumpe el proceso.

Hay varios puntos desde los que comparar este trabajo con los realizados por Słowiński (2021) y Huang et al. (2020) que son los más parecidos a lo desarrollado. Empezando por la exactitud podemos discutir distintos elementos, la que ambos alcanzaron es de 92.58 y 94.63 respectivamente, estos siendo sus resultados más altos, a continuación, se comenta sobre lo que hacen distinto. Un resumen de las cotas alcanzadas se ve en la tabla 5.

Tabla 5
Comparativa de exactitud alcanzada

Trabajo	Exactitud	Modelo
Actual (1)	76.92	Mobilenet v2
Actual (2)	52.01	Resnet 50
Słowiński (2021)	92.58	Red propia de 6 capas
Huang et al. (2020)	94.63	Yolo v3

En el caso de Słowiński (2021) la primera consideración, es el hecho de que no toma en cuenta el color de las imágenes teniendo como paso de preprocesamiento el cambio de estas a blanco y negro, además de que el tamaño de dataset es menor con 13611 muestras, se utilizó en el caso de la CNN una red ligera de 6 capas.

Huang et al. (2020) se parece más a lo aquí presentado, sin embargo, existen varias diferencias, tales como, el fondo de las fotografías del dataset que es negro al igual que su tamaño que es mayor, 72000 imágenes, pero quizá las diferencias más significativas son su

hardware, superior al utilizado en esta investigación, y el hecho de que para la clasificación de los granos únicamente se tomó en cuenta uno de los lados, este trabajo se desarrolló con Yolov3.

Al analizar ambos autores y las limitaciones es posible inferir que para mejorar los modelos presentados se necesite más variedad en el dataset, aumentándolo al añadir otras variables como distintos fondos, imágenes con varios sujetos o agregar más alteración en el entrenamiento cosa que no fue posible por lo ya discutido.

En la tabla 6 se observan los resultados de las pruebas que realizaron seis individuos sobre 3 muestras de café al inicio y al final de una jornada de trabajo, en la figura 53 se ve el proceso de selección que llevaron a cabo, en el que se utilizó 100 gramos de granos con defectos en 3 kg para realizar una separación. En este proceso se utilizó una balanza para medir la precisión en base al peso del café recolectado y luego cuántos de estos fueron descartados de manera errónea.

Figura 53
Pruebas de selección de café realizadas manualmente



Tabla 6
Resultados de pruebas de selección de café realizadas manualmente

Encargado	Muestra total	Defectos	Escogido/Escogido o correctamente muestra 1 (mañana)	Escogido/Escogido correctamente muestra 1 (tarde)	Escogido/Escogido correctamente muestra 2 (mañana)	Escogido/Escogido correctamente muestra 2 (tarde)	Escogido /Escogido correctamente muestra 3 (mañana)	Escogido/Escogido correctamente muestra 3 (tarde)
Sujeto 1	3kg	100g	42g / 40g	41g / 32g	26g / 24g	25g / 16g	38g / 36g	51g / 22g
Sujeto 2	3kg	100g	68g / 60g	53g / 45g	33g / 30g	56g / 30g	67g / 64g	39g / 28g
Sujeto 3	3kg	100g	54g / 53g	59g / 37g	40g / 34g	47g / 33g	63g / 58g	44g / 25g
Sujeto 4	3kg	100g	81g / 77g	65g / 52g	61g / 55g	84g / 40g	72g / 66g	76g / 59g
Sujeto 5	3kg	100g	70g / 62g	60g / 49g	67g / 65g	59g / 38g	102g / 88g	63g / 34g
Sujeto 6	3kg	100g	40g / 39g	28g / 20g	46g / 42g	56g / 53g	50g / 44g	36g / 17g
Promedio			59g / 54g	51g / 39g	45g / 41g	54g / 35g	65g / 59g	51g / 30g

Existen varias posibles explicaciones para la caída en la confianza y recolección de café que se observa durante las pruebas de la tarde, pueden ser varios factores, tanto externos al sujeto como no, estos pueden ser el cansancio, apresurarse por terminar la jornada o que no hay tanta luz como en las pruebas que se hicieron al comienzo del día.

Gracias a estos resultados se observa que, aunque la mayoría de los sujetos son eficientes en su trabajo no pueden mantener una consistencia con respecto a las muestras de la tarde y la mañana, algunos de los datos a destacar son la selección más baja por parte del sujeto 6 de 28g o la única vez en la que la cantidad escogida superó la muestra de defectos. Si se clasificaran los modelos junto con los sujetos estos estarían en la media al clasificar varios granos buenos de manera incorrecta, pero aquellos sujetos con resultados altos los superan al menos en las pruebas durante la mañana.

Conclusiones y recomendaciones

Conclusiones

El apoyo del experto para el análisis y clasificación del dataset probó ser tanto útil como necesario, ya que, aunque se citan manuales en este trabajo su contenido puede variar entre ellos, y con ayuda del especialista se pudo generalizar y aplicar una separación en base a las variedades de café recolectadas.

La investigación y análisis de los trabajos relacionados fue clave para conocer las tendencias dentro de la ciencia de la inteligencia artificial, así como las nuevas tecnologías, para saber elegir las bases con las que se experimentó al igual que las opciones de herramientas para realizar el trabajo.

Aunque los trabajos relacionados superan los modelos presentados debemos tomar en cuenta las diferencias que se marcan en la discusión y aprender de estas para intentar mejorar en un futuro.

Al analizar los resultados obtenidos a través de las evaluaciones de precisión y la matriz de confusión de las distintas pruebas de entrenamiento de ambos modelos podemos ver que todavía existe amplia oportunidad de mejora. Sin embargo, esto también significa que en su estado actual no sería posible su uso práctico.

Es posible para los modelos superar la eficiencia de los sujetos probados tras la jornada de trabajo, gracias a su consistencia su precisión no decae, pero no se compara con los mejores resultados de nuestros individuos.

Aunque la máquina utilizada para entrenamiento fue suficiente, el ajuste de la red más pesada podría haber mejorado su tiempo con un equipo de mejor calidad, lo que hubiera permitido más pruebas ya que se aplicaron los requerimientos mínimos con el fin de evitar errores de falta de memoria.

Recomendaciones

En caso de realizar una continuación sobre este trabajo existen varios puntos que se pueden tomar en cuenta para mejorar los resultados de futuros modelos.

- **Dataset:** Lo ideal sería ampliar el dataset presentado al aplicar más técnicas de data augmentation como el añadir ruido, con más fotografías, con varios sujetos, fondos distintos, más granos y más diferenciados de la clase problemática, o en el caso de que se desee utilizar una variedad de café con cualidades distintas, el reemplazo total.
- **Hardware:** el uso de un hardware superior permitirá más pruebas, lo que puede llevar a cotas mayores. También existe la opción de utilizar servicios de entornos de ejecución virtuales como Google Collab, para considerar el uso de estas debemos tomar en cuenta que la opción de suscripción sería la más idónea para este tipo de tarea.
- **Modelos:** Aunque es posible seguir el entrenamiento con los modelos aquí desarrollados también existe la opción tanto de cambiar de modelo y configuración de entrenamiento hasta de reemplazar el framework utilizado.

Si se llega a alcanzar una mejora en el futuro se plantea como posibilidad los usos prácticos de los modelos como herramienta, bien como reemplazo total o parcial en el proceso de clasificación mediante la ayuda de una maquina o como herramienta de apoyo para quienes realizan este trabajo.

Referencias

- Acevedo, G., Caicedo, E. y Loaiza, H. (2007). Selección de personal mediante redes neuronales artificiales. *Revista de Matemática: Teoría y Aplicaciones*, 14(1), 7–20.
<https://doi.org/10.15517/RMTA.V14I1.278>
- Adem, K., Metin Ozguven, M. y Altas, Z. (2022). A sugar beet leaf disease classification method based on image processing and deep learning. *Multimedia Tools and Applications*, 1–18.
<https://doi.org/10.1007/S11042-022-13925-6>
- Arcos, M., Cajas, G., Carrión, C., Chapi, J., Ellen Fieweger, M. y Karen, D. (2010). Íntag. *Caficultores Inteños Producen Café de Calidad*, pp. 19–24. Recuperado de www.intagnewspaper.org
- Ariza-López, F., Rodríguez-Avi, J. y Alba-Fernández, V. (2018). Control estricto de matrices de confusión por medio de distribuciones multinomiales. *Revista internacional de ciencia y tecnología de la información geográfica*, 21, 215–226.
- Arora, S. y Bhatia, M. (2018). Handwriting recognition using Deep Learning in Keras. En *Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018* (pp. 142–145). Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ICACCCN.2018.8748540>
- Atencio, P. (2009). *Desarrollo de un sistema de visión artificial para el control de calidad del mango de azúcar de exportación* [Tesis de pregrado, Universidad del Magdalena]. <http://repositorio.unimagdalena.edu.co/handle/123456789/362>
- Bastidas, M. (2017). *Caracterización de las técnicas de manejo en postcosecha para la obtención de cafés especiales* [Tesis de pregrado, Universidad de Nariño]. <https://sired.udenar.edu.co/5474/1/ARTICULO%20POSTCOSECHA%20FINAL%20%20MABEL%20BASTIDAS.pdf>
- Bonet, I., Salazar, S., Rodriguez, A., Grau, R. y Garcia, M. (2007). Redes neuronales recurrentes para el análisis de secuencias. *Revista Cubana de Ciencias Informáticas*, 1(4), 48–57.
<https://www.redalyc.org/pdf/3783/378343634004.pdf>
- Bonilla, C. (2020). *Redes convolucionales* [Tesis de pregrado, Universidad de Sevilla]. <https://idus.us.es/bitstream/handle/11441/115221/TFG%20DGMMyE%20Bonilla%20Carri%C3%B3n%20Carmelo.pdf?sequence=1&isAllowed=y>
- Cancelas, J., Gonzáles, R., Álvarez, I. y Enguita, J. (2016). Procesamiento morfológico. En E. Alegre, G. Pajares y A. De la Escalera (Eds.), *Conceptos y métodos en visión por computador* (pp. 77-97). CETA
- Cevallos, J. (2008). Redes Neuronales de Base Radial aplicadas a la mejora de la calidad. *Industrial Data*, 11(2), 63–72. <http://www.redalyc.org/articulo.oa?id=81619829009>
- Cevallos, Á., Sánchez, I., Lalama, J., Echeverría, J. y Montes, S. (2018). *Embriogénesis somática de Coffea arabica L. Var. Caturra Rojo, Bourbon Cidra y SL-28, de plantaciones cafetaleras de la Provincia del Carchi, Zona 1*. Mawill. Recuperado de <https://isbn.cloud/9789942787217/embriogenesis-somatica-de-coffee-arabica-l-var-caturra-rojo-bourbon-cidra-y-sl-28-de-plantacione/>
- Cevallos, A. (2000). *Establecimiento de una metodología eficiente en el proceso de embriogénesis somática del café (Coffea esp.), mediante el uso de marcadores morfohistológicos y moleculares* [Tesis de doctorado, Universidad Agraria de la Habana].
- Chicaiza, D. (2022). *Desarrollo de un modelo de red neuronal profunda para detección de meteoritos del sistema integrado AllSkyCams en Estados Unidos* [Tesis de pregrado, Universidad Técnica del norte]. <http://repositorio.utn.edu.ec/handle/123456789/12917>
- Chiu, Y., Tsai, C., Ruan, M., Shen, G. y Lee, T. (2020). Mobilenet-SSDv2: an improved object detection model for embedded systems. *2020 International Conference on System Science and Engineering, ICSSE 2020*, 0–4.
<https://doi.org/10.1109/ICSSE50014.2020.9219319>
- Farías, R., Muñóz, L., Marcillo, C., Viteri, M., Vinueza, J., Galarza, C. y Cevallos, J. (2020). Covid-

- 19: *Impacto en las exportaciones de organizaciones de pequeños productores, afectaciones, desafíos y oportunidades*. Ministerio de Producción, Comercio Exterior, Inversiones y Pesca. <https://www.produccion.gob.ec/wp-content/uploads/2020/12/Doc-completo-Impacto-Exportaciones-EPS.pdf>
- García-Olalla, O. y Alegre, E. (2016). Descripción de textura en imágenes utilizando local binary pattern (LBP). En E. Alegre, G. Pajares y A. De la Escalera (Eds.), *Conceptos y métodos en visión por computador* (pp. 115-130). CETA
- García, M., Chaves, D. y Trujillo, M. (2022). An Automatic Bean Classification System Based on Visual Features to Assist the Seed Breeding Process, 165–176.
- Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A. y De, D. (2019). Fundamental concepts of convolutional neural network. *Intelligent Systems Reference Library*, 172, 519–567. https://doi.org/10.1007/978-3-030-32644-9_36
- Guijarro, M., Herrera, P. y Montalvo, M. (2016) En E. Alegre, G. Pajares y A. De la Escalera (Eds.), *Conceptos y métodos en visión por computador* (pp. 99-114). CETA
- Haque, E., Rahman, A., Junaid, I., Hoque, S. y Paul, M. (2022). Rice leaf disease classification and detection using yolov5. *arXiv preprint arXiv:2209.01579*.
- Huang, N., Chou, D., Lee, C., Wu, F., Chuang, A., Chen, Y. y Tsai, Y. (2020). Smart agriculture: real-time classification of green coffee beans by using a convolutional neural network. *IET Smart Cities*, 2(4), 167–172. <https://doi.org/10.1049/IET-SMC.2020.0068>
- Instituto Interamericano de Cooperación para la Agricultura (2019). *Manual de producción sostenible de café en la República Dominicana*. Instituto Interamericano de Cooperación para la Agricultura (IICA)/RD. Recuperado de <https://bvearmb.do/handle/123456789/1379>
- Isaza-Gil, L., Montoya-Restrepo, E., Vélez-Zape, J. y Oliveros-Tascón, E. (2006). Evaluación de la concentración de los frutos maduros de café empleando técnicas no selectivas de recolección manual. *Cenicafé*, 57(4), 274–287.
- Jiang, L. y Zhang, Z. (2021). Research on image classification algorithm based on pytorch. In *Journal of Physics: Conference Series* (Vol. 2010, p. 12009). IOP Publishing. <https://doi.org/10.1088/1742-6596/2010/1/012009>
- Jubayer, F., Soeb, J., Mojumder, A., Paul, M., Barua, P., Kayshar, S. y Islam, A. (2021). Detection of mold on the food surface using YOLOv5. *Current Research in Food Science*, 4, 724–728. <https://doi.org/10.1016/j.crfs.2021.10.003>
- Kaiming, H., Xiangyu, Z., Shaoqing, R. y Jian, S. (2006). Deep residual learning for image recognition. *Indian Journal of Chemistry - Section B Organic and Medicinal Chemistry*, 45(8), 1951–1954. <https://doi.org/10.1002/chin.200650130>
- Learned-Miller, E. (2011). Introduction to Computer Vision. *University of Massachusetts*
- Lima de Sousa, P. (2021). Data Labeling tools for Computer Vision: a review [Tesis de maestría, Universidade de Lisboa]. <https://run.unl.pt/bitstream/10362/135873/1/TCDMAA0144.pdf>
- Lu, J., Tan, L. y Jiang, H. (2021). Review on Convolutional Neural Network (CNN) Applied to Plant Leaf Disease Classification. *Agriculture*, 11(8), 707. <https://doi.org/10.3390/AGRICULTURE11080707>
- Manrique, E. (2020). Machine Learning: análisis de lenguajes de programación y herramientas para desarrollo. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E28), 586–599. Recuperado de <https://search.proquest.com/openview/c7e24c997199215aa26a39107dd2fe98/1?pq->
- Marín, G. (2013). *Control de calidad del café Manual técnico*. Editorial Desco.
- Martín, D., García, F. y Armingol, J. (2016). El color: modelos y transformaciones de los espacios de color. En E. Alegre, G. Pajares y A. De la Escalera (Eds.), *Conceptos y métodos en visión por computador* (pp. 47-60). CETA
- Matich, D. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. En C. Ruiz y M. Basualdo (Comp.), *Informática Aplicada a la Ingeniería de Procesos-Orientación I*.

- Universidad Tecnológica Nacional.
- Mery, D. (2004). *Visión por Computador*. Recuperado de <http://www.ing.puc.cl/>
- Moreno, F. y Ramírez, E. (2017). Algoritmos de Visión por Computador para un SBC. *Lecturas en ciencias de la computación*
- Organización de las Naciones Unidas (2018). *Agenda 2030 y los Objetivos de Desarrollo Sostenible. Una oportunidad para América Latina y El Caribe*. (LC/G.2681-P/Rev.3), Santiago
- Ordóñez, F. y Roggen, D. (2016). Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors (Switzerland)*, 16(1).
<https://doi.org/10.3390/s16010115>
- Otero, I. (2019). *Análisis de la capacitación y asistencia técnica a los productores de café, caso de estudio el noroccidente de Quito 2014-2018* [Tesis de pregrado, Pontificia Universidad Católica del Ecuador] <http://repositorio.puce.edu.ec:80/handle/22000/16587>
- Rosas-Echevarría, C., Solís-Bonifacio, H. y Cerna-Cueva, A. (2019). Efficient and low-cost system for the selection of coffee beans: An application of artificial vision. *Scientia Agropecuaria*, 10(3), 347–351. <https://doi.org/10.17268/sci.agropecu.2019.03.04>
- Secretaría Nacional de Planificación y desarrollo (2021). Plan de Creación de oportunidades 2021-2025. *Secretaría Nacional de Planificación*. Quito: Secretaría Nacional de planificación. Recuperado de <https://www.planificacion.gob.ec/wp-content/uploads/2021/09/Plan-de-Creación-de-Oportunidades-2021-2025-Aprobado.pdf>
- Słowiński, G. (2021). Dry beans classification using machine learning. *CEUR Workshop Proceedings*, 2951, 166–173.
- Valbuena, L. (2019). De la geología al café. *Anekumene*, (18), 80–81.
<https://doi.org/10.17227/anezumene.2019.num18.13195>
- Villa, E. y Cruz, C. (2013). *Visión por computador aplicado al módulo de inspección de la célula minitek de la Facultad de Ingeniería Industrial de la Universidad Tecnológica de Pereira* [Tesis de pregrado, Universidad tecnológica de Pereira].
<https://repositorio.utp.edu.co/server/api/core/bitstreams/134b208b-e32b-401e-85c1-f529ee4a313a/content>
- Vintimilla, M. (2014). *Desarrollo e implementación de una aplicación que traduzca el abecedario y los números del uno al diez del lenguaje de señas a texto para ayuda de discapacitados auditivos mediante dispositivos móviles android*. Universidad de las Fuerzas Armadas, Sangolquí.
- Wallis-García, J., Montoya-Restrepo, C., Vélez-Zape, J. y Oliveros-Tascón, E. (2004). Calidad y eficacia de dos métodos no selectivos de recolección manual de café (*Coffea arabica*). *Cenicafé*, 55(1), 45–51.