

Escuela Superior Politécnica del Litoral

Facultad de Ingeniería en Electricidad y Computación

Clasificación de deformidades en frutas usando modelos de aprendizaje profundo con
imágenes reales y sintéticas

Proyecto Integrador

Previo la obtención del Título de:

Ingeniero en Ciencias de la Computación

Presentado por:

Tommy David Beltrán Borbor

Raúl Josué Villao Rodríguez

Guayaquil - Ecuador

Año: 2023

Dedicatoria

Dedico el presente proyecto a mis padres Humberto y Nelly, quienes con su esfuerzo y dedicación lograron traerme hasta aquí. Siempre estuvieron para brindarme apoyo, mucho más en aquellos momentos en los que me daba por vencido. Este logro no es solo mío, también es de ellos.

A mi hermana Kelly que, mediante sus consejos y apoyo incondicional, me permitieron dar lo mejor de mí, además de sacarme miles de sonrisas a lo largo de nuestra vida universitaria. Mas que hermana, es mi mejor amiga.

A mis mascotas Mishi, Pelussa y Kira, quienes con su cariño me inspiraban a seguir adelante para cumplir mis metas y darles la vida que se merecen.

Concluyo dedicando este trabajo a mis amigos quienes, con su compañía y, palabras de aliento, hicieron que etapa que hoy concluye, sea más llevadera y agradable.

Tommy David Beltrán Borbor

Dedicatoria

El presente proyecto lo dedico mis padres Graciela y Arturo que me han mostrado un apoyo incondicional a lo largo de mi vida, a mi hermana Yaret que ha estado a mi lado de igual manera siempre. Sobre todo, también lo dedico a Jehová Dios quien me ha dado el entendimiento necesario para poder realizar este trabajo.

Raúl Josué Villao Rodríguez

Agradecimientos

Agradecemos profundamente al PhD. Boris Vintimilla y al MSc. Luis Chuquimarca quienes, con infinita paciencia, trazaron y guiaron el desarrollo de este proyecto. Sus valiosos consejos y conocimientos nos permitieron avanzar en cada etapa del desarrollo siendo nuestra luz a lo largo del camino.

Declaración expresa

“Los derechos de titularidad y explotación, nos corresponde conforme al reglamento de propiedad intelectual de la institución; Tommy David Beltrán Borbor y Raúl Josué Villao Rodríguez damos nuestro consentimiento para que la ESPOI realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual”.

A handwritten signature in black ink, enclosed in a hand-drawn oval. The signature appears to read "T. Beltrán" with a small mark below it.

Tommy David Beltrán Borbor

Raúl Josué Villao Rodríguez

Evaluadores

PhD. Miguel Realpe

Profesor de la materia

PhD. Boris Vintimilla

Profesor tutor

Resumen

En la actualidad, una fruta pasa por múltiples procesos hasta llegar a nuestro hogar. Uno de ellos es la evaluación de su calidad mediante la inspección manual que evalúa parámetros de madurez, defectos y forma. Este proceso de evaluación puede dar cabida a inconsistencias y alargar los tiempos en el proceso de calificación de una fruta. Por ese motivo, el presente proyecto propone la evaluación de arquitecturas de CNN que permitan reconocer y clasificar deformidades en las frutas manzana, mango y fresa, con el fin de automatizar esta importante fase de la industria. Para ello, se recolectó imágenes disponibles en repositorios públicos de las frutas mencionadas y se utilizó generadores de imágenes sintéticas para obtener un conjunto de datos significativo en cuanto a la forma de las frutas. Luego, se procedió con la implementación y entrenamiento de varias arquitecturas de CNN; entre ellas, VGG16, MobileNetV2 y CIDIS. Se emplearon técnicas de Transferencia de Aprendizaje y se empleó el concepto de redes siamesas para habilitar el uso de imágenes de siluetas de las frutas. Como resultados de los entrenamientos, se obtuvo que la mejor arquitectura para obtener clasificación por forma de las frutas fue la MobileNetV2, alcanzando exactitudes del 90%, 94% y 92% en manzanas, mangos y fresas respectivamente. Se concluye que la arquitectura MobileNetV2 es la más precisa y eficiente para la clasificación de deformidades en frutas y se recomienda su inclusión en aplicaciones de evaluación de calidad de las frutas.

Palabras Clave: Forma, Silueta, Redes Neuronales Convolucionales, Transferencia de Aprendizaje, Red Siamesa.

Abstract

Currently, a fruit goes through multiple processes until it reaches our home. One of them is the evaluation of its quality through manual inspection that evaluates parameters of maturity, defects and shape. This evaluation process can give rise to inconsistencies and lengthen the times in the process of qualifying a fruit. For this reason, this project proposes the evaluation of CNN architectures that allow recognizing and classifying deformities in apple, mango and strawberry fruits, in order to automate this important phase of the industry. For this, images available in public repositories of the mentioned fruits were collected and synthetic image generators were used to obtain a significant data set regarding the shape of the fruits. Then, we proceeded with the implementation and training of various CNN architectures; among them, VGG16, MobileNetV2 and CIDIS. Techniques like Transfer Learning were used and the concept of Siamese networks was used to enable the use of images of silhouettes of the fruits. As results of the training, it was obtained that the best architecture to obtain classification by shape of the fruits was the MobileNetV2, reaching accuracies of 90%, 94% and 92% in apples, mangoes and strawberries respectively. It is concluded that the MobileNetV2 architecture is the most accurate and efficient for the classification of deformities in fruits and its inclusion in fruit quality evaluation applications is recommended.

Keywords: *Shape, Silhouette, Convolutional Neural Networks, Transfer Learning, Siamese Network.*

Índice general

Resumen	I
Abstract.....	II
Índice general	III
Abreviaturas.....	V
Índice de figuras.....	VI
Índice de tablas.....	VIII
Capítulo 1	1
1.1 Introducción	2
1.2 Descripción del problema.....	3
1.3 Justificación del problema.....	4
1.4 Objetivos.....	5
1.4.1 Objetivo general.....	5
1.4.2 Objetivos específicos	5
1.5 Marco teórico	6
1.5.1 Inspección de calidad de las frutas	6
1.5.2 Deformidades en frutas.....	6
1.5.3 Técnicas para la detección de deformidades en frutas	14
Capítulo 2	20
2.1 Metodología	21
2.2 Diseño de la solución.....	21
2.2.1 Parámetros de evaluación de deformidades en frutas	22
2.2.2 Elección de arquitecturas de CNN	22
2.3 Módulo 1: Creación del dataset.....	23
2.3.1 Recopilación de imágenes reales	24
2.3.2 Generación de imágenes sintéticas	26
2.3.3 Preprocesamiento de imágenes	31

2.4	Módulo 2: Implementación y entrenamiento de modelos de CNN	35
2.4.1	Enfoque de Transfer Learning	35
2.4.2	Enfoque de Red Siamesa	36
2.5	Módulo 3: Evaluación de los modelos de CNN	37
2.6	Módulo 4: Desarrollo de Prototipo	38
	Capítulo 3	40
3.1	Resultados y análisis	41
3.2	Creación del conjunto de imágenes de entrenamiento.....	41
3.2.1	Dataset Real Versión 1 (DRV1)	41
3.2.2	Dataset Real Versión 2 (DRV2)	43
3.2.3	Dataset Sintético (DS)	44
3.2.4	Dataset de siluetas de frutas.....	45
3.3	Entrenamiento de modelos de CNN.....	46
3.3.1	Resultados con enfoque Transfer Learning	46
3.3.2	Resultados con enfoque de Red Siamesa	50
3.4	Prototipo	53
	Capítulo 4	55
4.1	Conclusiones y recomendaciones.....	56
4.2	Conclusiones	56
4.3	Recomendaciones	57
4.4	Trabajos futuros.....	58
4.5	Contribuciones.....	59
	Referencias.....	60

Abreviaturas

ESPOL	Escuela Superior Politécnica del Litoral
CNN	Convolutional Neural Network
CIDIS	Centro de Investigación, Desarrollo e Innovación de Sistemas Computacionales
OECD	Organisation for Economic Co-operation and Development
USDA	United States Department of Agriculture
IA	Inteligencia Artificial
AP	Aprendizaje Profundo
TL	Transfer Learning
SAM	Segment Anything Model
DRV1	Dataset Real Versión 1
DRV2	Dataset Real Versión 2
DS	Dataset Sintético

Índice de figuras

Figura 1.1 <i>Canales de comercialización</i>	2
Figura 1.2 <i>Diagrama de ambiente controlado para adquisición de imágenes</i>	8
Figura 1.3 <i>Manzana Clase Extra</i>	9
Figura 1.4 <i>Manzana Clase I</i>	9
Figura 1.5 <i>Manzana Clase II</i>	10
Figura 1.6 <i>Simetría de la manzana a lo largo de su formación</i>	10
Figura 1.7 <i>Mango Clase Extra</i>	11
Figura 1.8 <i>Mango Clase I</i>	12
Figura 1.9 <i>Mango Clase II</i>	12
Figura 1.10 <i>Fresa Clase Extra</i>	13
Figura 1.11 <i>Fresa Clase I</i>	13
Figura 1.12 <i>Fresa Clase II</i>	14
Figura 1.13 <i>VGG16 modificada para detección de forma y tamaño de manzanas</i>	17
Figura 1.14 <i>Capas de modelo de CNN propuesto para clasificación de calidad de manzanas</i> ..	17
Figura 1.15 <i>Arquitectura siamesa empleada en la clasificación de madurez de Papaya</i>	18
Figura 2.1 <i>Módulos de desarrollo de la solución</i>	21
Figura 2.2 <i>Ejemplos de imágenes no admitidas en el conjunto de datos de entrenamiento</i>	25
Figura 2.3 <i>Imágenes de fresas generadas por DALL-E 2</i>	26
Figura 2.4 <i>Comparación de resultados entre DALL-E mini y DALL-E 2</i>	27
Figura 2.5 <i>Interfaz gráfica de Easy Diffusion</i>	29
Figura 2.6 <i>Interfaz gráfica de Kohya</i>	30
Figura 2.7 <i>Fases para la generación de imágenes sintéticas</i>	31
Figura 2.8 <i>Preprocesamiento de imágenes producto de técnica de Rotación</i>	32
Figura 2.9 <i>Segmentación de objetos con SAM</i>	32
Figura 2.10 <i>Identificación de todas máscaras de una imagen</i>	33

Figura 2.11 <i>Preprocesamiento de imágenes producto de técnica Mosaico</i>	33
Figura 2.12 <i>Uso de método de ROI para segmentación de la fruta</i>	34
Figura 2.13 <i>Resultado de eliminación de background no uniforme con herramienta SAM</i>	34
Figura 2.14 <i>Esquema de aprendizaje por Transfer Learning</i>	35
Figura 2.15 <i>Esquema de Red Siamesa</i>	36
Figura 2.16 <i>Matriz de confusión multiclase</i>	37
Figura 2.17 <i>Boceto Inicial del prototipo web</i>	39
Figura 2.18 <i>Boceto Final del prototipo web</i>	39
Figura 3.1 <i>Reajuste de tamaño de imagen sin relleno de bordes</i>	42
Figura 3.2 <i>Reajuste de tamaño de imagen con relleno de bordes</i>	42
Figura 3.3 <i>Ejemplos de diferencias encontradas entre las imágenes de los datasets</i>	44
Figura 3.4 <i>Obtención de silueta de cada fruta</i>	45
Figura 3.5 <i>Matriz de confusión obtenida con el mejor modelo para Manzana – Enfoque Lineal</i>	49
Figura 3.6 <i>Matriz de confusión obtenida con el mejor modelo para Mango – Enfoque Lineal</i>	49
Figura 3.7 <i>Matriz de confusión obtenida con el mejor modelo para Fresa – Enfoque Lineal</i>	50
Figura 3.8 <i>Matriz de confusión obtenida con el mejor modelo para Mango – Enfoque de Red Siamesa</i>	51
Figura 3.9 <i>Matriz de confusión obtenida con el mejor modelo para Manzana – Enfoque de Red Siamesa</i>	52
Figura 3.10 <i>Prototipo aplicado en manzanas</i>	53
Figura 3.11 <i>Prototipo aplicado en mangos</i>	54
Figura 3.12 <i>Prototipo aplicado en fresas</i>	54

Índice de tablas

Tabla 1.1 <i>Parámetros evaluados para clasificar deformidades según distintos organismos internacionales</i>	7
Tabla 1.2 <i>Comparación de desempeño de modelos de CNN en la clasificación de frutas</i>	16
Tabla 2.1 <i>Fuentes de bases de datos de imágenes reales consultadas</i>	24
Tabla 2.2 <i>Características del equipo de cómputo provisto por CIDIS</i>	28
Tabla 2.3 <i>Principales entradas de texto usadas para la generación de imágenes con Stable Diffusion</i>	28
Tabla 3.1 <i>Cantidad de imágenes reales en DRV1 obtenidas para el entrenamiento</i>	43
Tabla 3.2 <i>Combinación de hiperparámetros utilizados en los entrenamientos</i>	46
Tabla 3.3 <i>Resultados con DRV1 – Enfoque Transfer Learning</i>	47
Tabla 3.4 <i>Resultados con DRV2 – Enfoque Transfer Learning</i>	48
Tabla 3.5 <i>Mejores resultados – Enfoque Lineal</i>	50
Tabla 3.6 <i>Resultados con DRV2+DS+Silueta – Enfoque de Red Siamesa</i>	51
Tabla 3.7 <i>Mejores Resultados – Enfoque de Red Siamesa</i>	52
Tabla 3.8 <i>Comparación de resultados Enfoque Transfer Learning vs Enfoque de Red Siamesa</i>	52

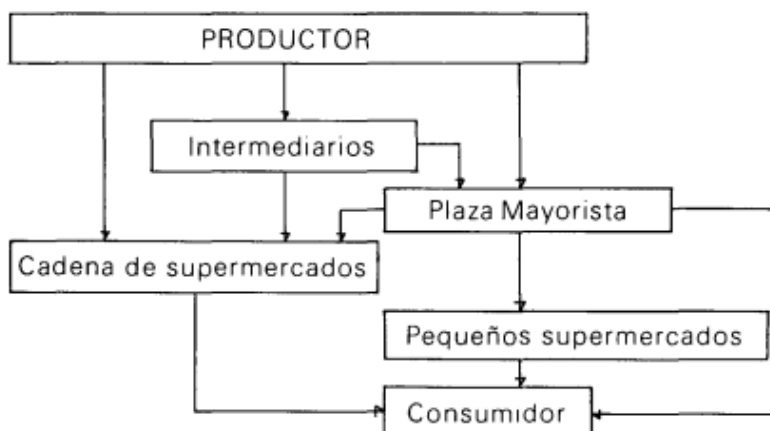
Capítulo 1

1.1 Introducción

Una de las etapas más importantes de la postcosecha de frutas y hortalizas, es la selección y clasificación. Durante este proceso, se descarta todo alimento que presente algún grado de descomposición o defecto [1], por lo que suele ser un proceso mecánico, minucioso y convencionalmente realizado por personas involucradas en esta área. La importancia de este proceso reside en el valor comercial que se le da al producto una vez haya sido evaluado, dado que, un producto con mayor calidad tenderá a ser vendido por un mayor precio. A su vez, pequeños comercios y supermercados, regidos por la exigencia comercial que les caracteriza, también evalúan parámetros para ejecutar la adquisición de estos alimentos [2].

Figura 1.1

Canales de comercialización



Nota: Esquema de los canales de comercialización más comunes alrededor del mundo. Obtenido de [2].

En el caso de las frutas, los supermercados se surten mediante canales de comercialización del tipo productor-mayorista [3], similares a los observados en la Figura 1.1, donde el precio de adquisición varía según la cantidad y calidad del fruto. El criterio de clasificación comúnmente utilizado mide parámetros externos visibles como el grado de madurez, firmeza, apariencia y deformidades.

El problema radica en que este proceso de evaluación de calidad es llevado a cabo por terceros, tornándose un proceso lento y poco consistente ya que el criterio de evaluación suele ser subjetivo. Este factor humano, reduce la eficiencia con la que se realiza la evaluación y altera el valor comercial de los frutos, afectando así a consumidores en general y agricultores/productores que podrían percibir una menor ganancia de sus cosechas.

El objetivo del presente proyecto es proponer una mejora en el proceso de clasificación de frutas mediante la elaboración de modelos de aprendizaje profundo (AP) que permitan clasificar de forma automática uno de los principales parámetros de evaluación de calidad: las deformidades. En este proyecto, las frutas de estudio consideradas serán manzanas, mangos y fresas.

1.2 Descripción del problema

La evaluación de la calidad constituye un aspecto esencial en la cadena de distribución y suministro de alimentos. Su misión es certificar que un producto determinado cumpla con los estándares y regulaciones estipulados en el lugar de producción, con el fin de asegurar que sea apto para el consumo humano y pueda obtener un rendimiento económico óptimo durante la etapa de comercialización.

En el caso específico de las frutas, la inspección puede llevarse a cabo a través de diversas técnicas. Entre ellas, las invasivas y no invasivas. El inconveniente reside en que el uso de técnicas invasivas, a menudo implican manipular, cortar o exprimir el fruto, lo que lo inhabilita para su comercialización [4]. Como alternativa, se realiza una inspección externa llevada a cabo por personas que, mediante criterios subjetivos, evalúan características como el tamaño, la forma, el nivel de madurez y presencia de defectos. Este método puede introducir errores en la clasificación, lo que puede resultar en una disminución del valor económico de muchos de estos frutos, sin olvidar mencionar que se trata de un proceso mecánico y naturalmente lento.

Para solucionar esto, el Centro de Investigación, Desarrollo e Innovación de Sistemas Computacionales (CIDIS) de ESPOL, se encuentra desarrollando un Sistema de Inspección de Calidad con el que se busca automatizar el proceso de clasificación de calidad en frutas. Dicho proyecto consta de tres componentes:

- Clasificación de grado de madurez.
- Identificación de defectos.
- Clasificación de deformidades.

Actualmente, los dos primeros componentes ya se han desarrollado, por lo que, se plantea desarrollar el tercer y último componente. Para lograr este cometido, se entrenarán varios modelos de Redes Neuronales Convolucionales (CNN por sus siglas en inglés) utilizando un conjunto de imágenes etiquetadas según el nivel o grado de deformidad establecido por un estándar internacional seleccionado.

1.3 Justificación del problema

De acuerdo con [5, 6], una buena proporción de los frutos cosechados se retiran de la cadena de producción a causa de una errónea evaluación de calidad. La inspección manual, falta de conocimiento y valoración rápida, suelen ser las principales causas de que esto ocurra. Dada la problemática, se necesita automatizar el proceso de clasificación de las frutas según su calidad. El desarrollo del componente destinado a la clasificación de deformidades del Sistema de Inspección de Calidad permitirá evaluar y clasificar con mayor precisión cada fruto, considerando su forma y verificando si presenta algún tipo de deformación.

Adicionalmente, dicha implementación dará paso al desarrollo de sistemas especializados en la inspección y control de calidad en el sector agrícola y afines. Así, se busca simplificar y automatizar el proceso de verificación de calidad de las frutas una vez cosechadas, eliminando el factor humano en la inspección, de modo que los frutos sean clasificados con mayor rapidez, precisión y que los resultados de la clasificación sea consistentes. Como consecuencia,

es posible disminuir el tiempo empleado en la clasificación de calidad de las frutas. Así, este proceso será mucho más eficiente y ocupará menos recursos de tiempo y manufactura.

El sistema va dirigido como una solución para el CIDIS. Posteriormente los modelos desarrollados formarán parte de un sistema de visión por computador, para que así el sistema intervenga en los procesos existentes de clasificación que usan las industrias alimentarias.

1.4 Objetivos

1.4.1 *Objetivo general*

Evaluar arquitecturas de CNN que permitan la clasificación de deformidades en frutas como manzanas, mangos y fresas con el fin de garantizar el cumplimiento de normas y estándares de calidad establecidos para estos alimentos.

1.4.2 *Objetivos específicos*

1. Levantar información sobre las principales deformaciones en manzanas, mangos y fresas según varios estándares internacionales para elaborar una categorización general por grados de deformidad en cada fruta mencionada.
2. Generar un conjunto de imágenes reales y sintéticas de manzanas, mangos y fresas, acordes a su grado de deformidad para la fase de entrenamiento y validación de los modelos de CNN.
3. Usar diferentes modelos de CNN existentes en el estado del arte para evaluar su rendimiento y seleccionar el modelo con mayor precisión.
4. Ejecutar pruebas para medir el correcto desempeño de los modelos evaluados y comprobar que clasifican acorde al estándar seleccionado.

1.5 Marco teórico

En esta sección se abordará el estado del arte respecto al análisis de deformidades en las frutas: manzana, mango y fresa. También se repasará las técnicas y modelos que se han implementado en otros estudios para el análisis de las deformidades en estas frutas.

1.5.1 Inspección de calidad de las frutas

La inspección de calidad es una parte esencial en la industria alimentaria. Muchos son los alimentos y productos preprocesados que requieren de frutas para su elaboración, por lo que comercios minoristas y mayoristas apuntan a adquirir las de mejor calidad y presentación. En base a ello, aseguran la satisfacción del consumidor y pueden obtener una remuneración mayor en la venta de dicho producto.

Para la evaluación de la calidad de las frutas, se pueden aplicar 2 metodologías: invasivas y no invasivas. Los métodos invasivos, también denominados destructivos, permiten analizar características como el color interno, textura y propiedades como contenido de azúcar, acidez, entre otros, mediante la manipulación del fruto. Por su parte, los métodos no invasivos permiten valorar características de forma, tamaño, grado de madurez y presencia de defectos sin alterar o dañar la fruta. La espectroscopía de reflectancia difusa, la espectroscopía de fluorescencia y la tomografía de resonancia magnética son las técnicas más comunes al emplear este enfoque [7, 8].

Estudios como [9] han mostrado que el uso de métodos no invasivos puede ser tan preciso como los métodos invasivos en la evaluación de la calidad de las frutas, además de ser más rápidos y menos costosos.

1.5.2 Deformidades en frutas

En este estudio, se investiga la detección de deformidades en frutas mediante el uso de CNN. Es esencial considerar la naturaleza de las deformidades para lograr una clasificación precisa y permitir que las CNN aprendan a reconocer frutas con deformidades.

Durante el proceso de crecimiento y desarrollo de las frutas, se pueden producir alteraciones en su forma y tamaño como resultado de diversos factores, incluyendo causas genéticas. Una fruta deformada es aquella que presenta una forma atípica que no se corresponde con las características representativas de su especie.

Cuando se trata de deformidades, generalmente se suelen valorar dos aspectos: la forma y el tamaño. Existen organismos internacionales encargados de establecer parámetros de calidad que evalúan dichos aspectos. Algunos ejemplos son la Organización para la Cooperación y el Desarrollo Económico (OECD por sus siglas en inglés), el Departamento de Agricultura de los Estados Unidos (USDA por sus siglas en inglés) y entes gubernamentales que establecen sus propios criterios como la República de China.

Suelen existir discrepancias entre los criterios establecidos por cada organismo. A continuación, se presenta una tabla que resume los principales parámetros evaluados en la mayoría de las frutas:

Tabla 1.1

Parámetros evaluados para clasificar deformidades según distintos organismos internacionales

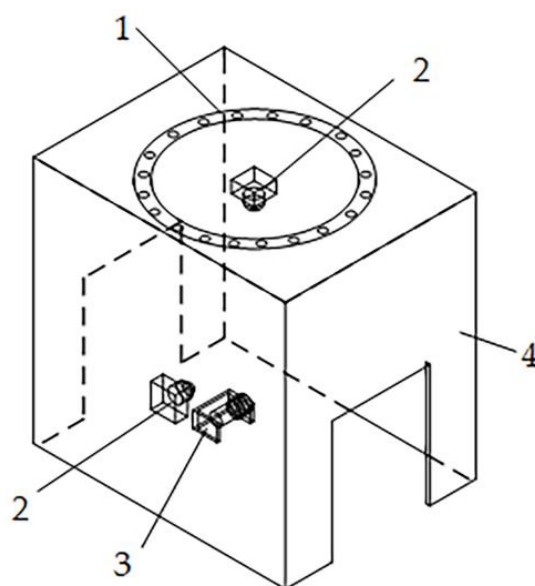
	Categorías	Forma	Tamaño
OECD	Clase-Extra	✓	✗ No evalúa
	Clase 1	Forma típica del fruto	
	Clase 2		
USDA	Extra-Fancy	✓	✓
	Fancy	Forma típica del fruto	Medición de diámetro mayor del fruto
	Utility		
República de China	Excelente	✓ Forma típica del fruto	✓
	1ra Clase		Medición de diámetro mayor del fruto. Medición de circunferencia del fruto desde vista superior
	2da Clase		
	Subestándar		

Nota: Conjunto de datos recopilados: OECD [10, 11, 12]; USDA [13, 14, 15]; R. China: [16, 17].

Como se puede observar, estándares como el estadounidense y chino tienen una forma mucho más estricta para llevar la evaluación de deformidades. Como ejemplo, en [16] y [18] se considera la forma y el diámetro de la manzana para su categorización. No obstante, la limitante que conlleva considerar el tamaño como parámetro de evaluación es que se requiere medir el diámetro de los frutos, por lo que será necesario realizar la obtención de imágenes mediante un ambiente controlado, donde se consideren parámetros como distancia de trabajo entre la cámara y el objetivo para facilitar el cálculo del diámetro.

Figura 1.2

Diagrama de ambiente controlado para adquisición de imágenes



Nota: La figura muestra un bosquejo del ambiente empleado en [16] para la adquisición de imágenes de manzanas para su posterior uso en el estudio, el punto 1 muestra las luces para iluminación de fotos, en número 2 muestra la cámara ubicada en la parte top, en la parte 3 interruptor fotoeléctrico y 4 caja negra.

La OECD expone la clasificación de la siguiente forma:

1. Extra-Clase (más alta)
2. Clase I (intermedio)

3. Clase II (más baja)

Según la OECD, para que una fruta sea clasificada como Extra-Clase, Clase I o Clase II, la fruta debe cumplir con requisitos mínimos de madurez, defectos, y forma. Si una fruta no alcanza las características para ser categorizada como Clase II, entonces se entiende que el fruto es de baja calidad y es categorizada como Sin Clase. A continuación, se muestran ejemplos gráficos de cada categoría por las frutas objeto de estudio:

1.5.2.1 Deformidades en manzanas. Según [10], la clasificación para deformidades en manzanas es la siguiente:

Figura 1.3

Manzana Clase Extra



Nota: La figura muestra el límite permitido para que una manzana sea categorizada como Clase Extra según su forma. Obtenido de [10].

Figura 1.4

Manzana Clase I



Nota: La figura muestra el límite permitido para que una manzana sea categorizada como Clase I según su forma. Obtenido de [10].

Figura 1.5

Manzana Clase II

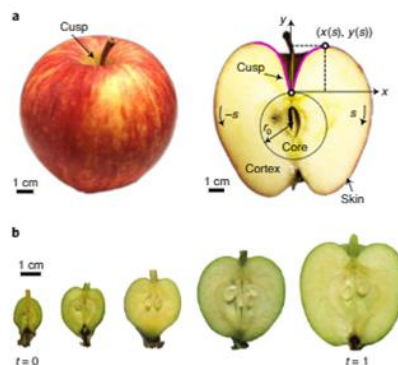


Nota: La figura muestra el límite permitido para que una manzana sea categorizada como Clase II según su forma. Obtenido de [10].

En este caso, la OECD establece que la manzana debe presentar una forma simétrica en su superficie, sin desviaciones o bultos evidentes. En las imágenes previas, se puede apreciar que las manzanas del lado izquierdo exhiben una inclinación que varía según su categoría, mientras que en el lado derecho se observan bultos que indican una deformación. Esto evidentemente describe la forma de una manzana tradicional, la que durante su desarrollo debe presentar una forma simétrica de ambos lados, es decir, como si fuese un espejo. Cuando la manzana no presenta una forma simétrica, entonces se la considera deforme [19].

Figura 1.6

Simetría de la manzana a lo largo de su formación



Nota: Se muestra las etapas de formación de la manzana donde es común que mantenga una forma simétrica. Adaptado de [19].

Como se puede observar en la Figura 1.6, ambos lados de la parte “top” de la manzana crecen de forma similar.

En Ecuador, la MALUS DOMESTICA es la variedad de manzana más consumida [20]. Aunque esta manzana no se distingue por características especiales en su forma en comparación con la manzana tradicional, en el contexto de esta investigación, todas las manzanas, independiente de su familia, deben presentar una forma simétrica. Por lo tanto, se pueden incluir en el conjunto de imágenes manzanas de diferentes familias, ya que todas son útiles para el reconocimiento de deformidades en su simetría.

1.5.2.2 Deformidades en mangos. Según [11], la clasificación para deformidades en mangos es la siguiente:

Figura 1.7

Mango Clase Extra



Nota: La figura muestra el límite permitido para que un mango sea categorizado como Clase Extra según su forma. Obtenido de [11].

Figura 1.8*Mango Clase I*

Nota: La figura muestra el límite permitido para que un mango sea categorizado como Clase I según su forma. Obtenido de [11].

Figura 1.9*Mango Clase II*

Nota: La figura muestra el límite permitido para que un mango sea categorizado como Clase II según su forma. Obtenido de [11].

De acuerdo con los parámetros establecidos por la OECD, el mango debe tener una forma ovalada. No obstante, existen distintas familias, por lo que no podemos asegurar que cada mango tendrá una forma similar. Este estudio se centrará en las propiedades del mango proveniente de *Magnifera Indica*, el mango de la variedad Tommy Atkins. Este mango es el de

mayor producción en el Ecuador y su forma tradicional es una forma ovoide [21]. Esto es de gran ventaja porque entonces se podrá afirmar que todo mango en el conjunto de imágenes que presente esta forma será catalogado con la categoría Extra-Clase.

1.5.2.3 Deformidades en fresas. Según [12], la clasificación para deformidades en fresas es la siguiente:

Figura 1.10

Fresa Clase Extra



Nota: La figura muestra el límite permitido para que una fresa sea categorizada como Clase Extra según su forma. Obtenido de [12].

Figura 1.11

Fresa Clase I



Nota: La figura muestra el límite permitido para que una fresa sea categorizada como Clase I según su forma. Obtenido de [12].

Figura 1.12*Fresa Clase II*

Nota: La figura muestra el límite permitido para que una fresa sea categorizada como Clase II según su forma. Obtenido de [12].

Al igual que con las manzanas y los mangos, también existen varias especies y familias de fresas. Ventajosamente para este estudio, las variedades de fresa se distinguen por su tamaño y características en su sabor, pero la forma cónica tradicional se mantiene [22], por lo que cualquier familia será admitida en la recopilación de imágenes.

1.5.3 Técnicas para la detección de deformidades en frutas

1.5.3.1 Generación de datasets. Al emplear modelos de CNN, es esencial contar con una base de datos amplia de imágenes con las que el modelo pueda extraer características significativas de cada clasificación. Sin embargo, no es concluyente que tener una buena cantidad de imágenes de entrenamiento es equivalente a obtener un modelo con buena precisión. En [23], se concluye que la falta de preprocesamiento adecuado a las imágenes de entrada repercutió en un modelo con bajos niveles de aprendizaje. Desde ese sentido, surge la necesidad de adoptar uno de los principios del Smart Data: dar valor a la información.

Estudios como [16], [17] y [24] brindan valor a su conjunto de imágenes mediante el uso de técnicas de preprocesamiento, donde segmentan el objeto de interés, aplican un proceso de realzado o, utilizan imágenes en diferentes espectros visibles.

Respecto a la adquisición de imágenes, se toma en cuenta la naturaleza del estudio. Si dicho estudio no es tan restrictivo, es común acceder a plataformas como Kaggle (<https://www.kaggle.com>), Mendeley (<https://data.mendeley.com>) o GitHub (<https://github.com>) para descargar imágenes desde repositorios públicos. Si, por el contrario, es necesario llevar un control de los datos de entrada como en [16], se opta por un proceso de adquisición propia, donde se prepara el entorno, y se pueden manejar parámetros de luminosidad, distancia de enfoque, entre otros.

Por otro lado, si se desea aumentar el número de imágenes de entrenamiento, es posible utilizar técnicas de Data Augmentation. El uso de imágenes sintéticas también es una opción. [25] por ejemplo, empleó el motor de juegos Unreal Engine para la creación de imágenes realistas de banano en diferentes etapas de maduración. En [26], en cambio, se utilizó DALL-E mini, una inteligencia artificial Text-to-Image para generar imágenes de manzanas y mangos con diferentes tipos de defectos como podredumbre, costras y manchas. En ambos estudios, el uso de este tipo de imágenes permitió fidelizar el resultado del modelo, obteniendo un buen nivel de precisión luego del entrenamiento.

1.5.3.2 Detección y clasificación de deformidades en frutas por medio de visión por computador usando arquitecturas CNN. Para este estudio se usarán redes CNN para la clasificación de deformidades en las 3 frutas que se han planteado: manzanas, mangos, fresas. En este sentido, es vital que para el entrenamiento se usen arquitecturas que den los mejores resultados posibles. Son varios los estudios donde se realizan entrenamientos de CNN para detectar defectos generales en las frutas, entre ellos deformidades. Aunque estos modelos no se centraron en buscar deformidades, aun así, es posible que sean un buen punto de partida para el entrenamiento de las CNN propuestas en este proyecto.

Tabla 1.2

Comparación de desempeño de modelos de CNN en la clasificación de frutas

Fruta	Clasificación	Modelos CNN	Métrica	Desempeño
Manzana [18]	Forma, tamaño	VGG16	Accuracy	99.04%
Manzana [27]	Calidad	Propio	Accuracy	99.00%
Manzana [17]	Apariencia (Color, piel)	Imp-ResNet50	Accuracy	96.50%
Manzana [26]	Defectos (Podredumbre, Costras, Puntos negros)	InceptionV3	Accuracy	93.40%
		MobileNetV2		97.50%
		DenseNet121		94.50%
		VGG16		76.20%
Mango [26]	Defectos (Podredumbre, Costras, Puntos negros)	InceptionV3	Accuracy	91.90%
		MobileNetV2		92.90%
		DenseNet121		92.50%
		VGG16		63.60%
Mango [28]	Especies	Faster R-CNN	Accuracy	90.00%
Banana [25]	Madurez	VGG19	Accuracy	56.20%
		ResNet-50		81.60%
		InceptionResNetV2		86.90%
		CIDIS		87.20%
		InceptionV3		84.90%
Papaya [24]	Madurez	Propio	F1-Score	0.97

El estudio [18] se centró en identificar la forma y tamaño en manzanas. Se usó un modelo VGG16 modificado que consistía en 2 capas fully connected (tradicional de las VGG) y se añadieron 4 capas convolucionales más. La arquitectura es la mostrada en la Figura 1.13.

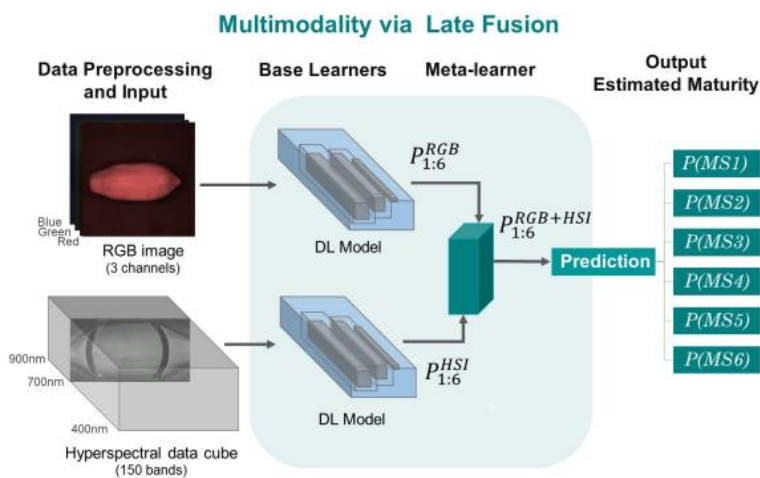
Lo interesante de esta investigación es que comparan su arquitectura propuesta con otras arquitecturas conocidas como el Google InceptionV3 y VSM Classifier. De hecho, la arquitectura propuesta contiene características de ambas arquitecturas para obtener mejores resultados.

Otra investigación usó Imp-ResNet50 para el entrenamiento de su modelo obteniendo buenos resultados para clasificar manzanas por su grado de madurez basándose en el color y características visibles en la piel como daños o manchas [17].

Para terminar, cabe destacar lo realizado en [24], dónde se implementó una arquitectura siamesa, para la clasificación de madurez de la papaya. Dicha arquitectura estaba compuesta de dos modelos base de AP, donde cada uno de ellos percibe una imagen de entrada y ejecutaba las respectivas operaciones de convolución. La regresión logística multiclase se utilizaba luego para fusionar los resultados de los modelos base. Finalmente, se ejecuta la predicción en base a 6 etapas de madurez. Lo expuesto se representa en la Figura 1.15.

Figura 1.15

Arquitectura siamesa empleada en la clasificación de madurez de Papaya



Nota: Obtenido de [24].

Esta arquitectura fue evaluada mediante la métrica F1-score donde obtuvo un resultado de 0.97.

Como se puede notar, para la evaluación de los modelos de estos estudios se usaron a lo mucho las métricas Accuracy y F1-score. No obstante, otras métricas pueden ser empleadas según el fin que se tenga con el modelo, entre ellas la matriz de confusión, de la cual se derivan otras métricas como Precision, Recall y AUC-ROC [29].

Capítulo 2

2.1 Metodología

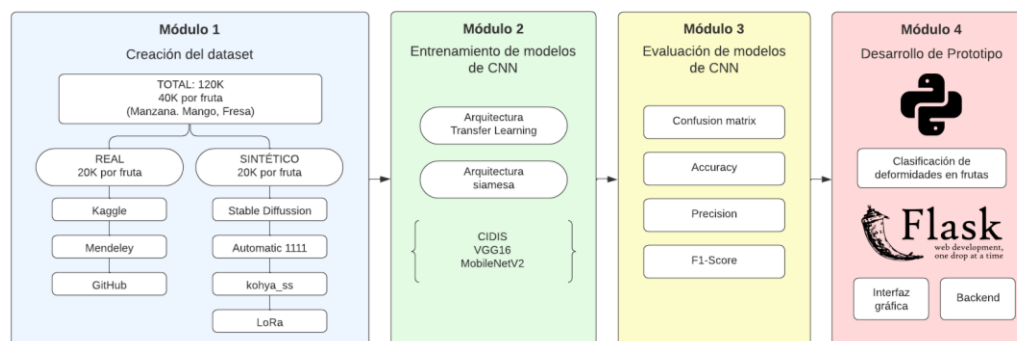
A lo largo de este capítulo, se presenta la metodología empleada para el desarrollo de la solución de este proyecto. Se empezó con un análisis exhaustivo de los recursos disponibles según la revisión literaria teniendo en cuenta las limitantes existentes, con lo cual se tomó decisiones de diseño fundamentadas en dicha evaluación. Se describen las diferentes fuentes consultadas para la creación del conjunto de datos real de imágenes, así como las tecnologías utilizadas para la generación del conjunto de datos sintéticos. Adicionalmente, se brinda una breve descripción sobre las arquitecturas de CNN seleccionadas para realizar la detección y clasificación de deformidades en las frutas: manzana, fresa y mango.

2.2 Diseño de la solución

Concerniente al entrenamiento de los modelos de AP, se decidió dividir la metodología de este proyecto en 4 módulos. Los Módulos 1 y 2 están destinados a la obtención del conjunto de datos de entrenamiento y a la implementación de los modelos de CNN respectivamente. El Módulo 3 evalúa los modelos de CNN obtenidos del módulo previo mediante distintas métricas. Y finalmente, el Módulo 4 desarrolla un prototipo de aplicación web en el que se utiliza el modelo de CNN mejor evaluado para cada fruta para la clasificación de deformidades. En la Figura 2.1 se muestran los 4 módulos usados para la solución de este proyecto

Figura 2.1

Módulos de desarrollo de la solución



Ahora bien, para el desarrollo de los módulos fue necesario definir normas a seguir, entre ellas, la elección del estándar de evaluación y las arquitecturas de CNN para la clasificación de deformidades. Esta información se detalla a continuación:

2.2.1 Parámetros de evaluación de deformidades en frutas

Los mercados internacionales exigen que las frutas sean evaluadas mediante un control de calidad [25], para aquello surgieron los estándares internacionales. En el marco teórico se mencionó algunos de los más conocidos; entre ellos, los estándares establecidos por la organización europea OECD y la organización americana USDA, además del estándar de la República Popular China.

Al tratarse de deformidades, se evalúa la forma y el tamaño de la fruta. El estándar chino y americano sí presentan normativas para ambas características, contrario al estándar europeo que sólo se centra en la forma de la fruta. La evaluación de ambas características aseguraría una predicción mucho más precisa de la deformidad, sin embargo, la consideración del tamaño conlleva un mayor nivel de complejidad en la recolección de los datos. En [16, 17], por ejemplo, optaron por elaborar su propio conjunto de datos mediante la construcción de un ambiente controlado, de manera que, el cálculo del diámetro de los frutos era mucho más sencillo y fiable.

Dada la imposibilidad de ensamblar un ambiente controlado para la adquisición de imágenes y la escasez de datos públicos respecto a la evaluación del tamaño, para este proyecto se tomó la decisión de adoptar los estándares establecidos por la OECD para la evaluación de deformidades. Por tal motivo, las categorías que se emplearon por cada fruta fueron: Extra Clase, Clase I, Clase II, y Sin Clase.

2.2.2 Elección de arquitecturas de CNN

Una vez seleccionado el estándar y, por ende, las características a considerar de la fruta, el siguiente paso fue seleccionar las arquitecturas de CNN que se iban a utilizar para la clasificación de deformidades. La revisión literaria ofreció una amplia gama de modelos CNN empleados para diferentes objetivos, entre ellos el análisis de la colorimetría (madurez), defectos

externos, calidad, entre otros. Hasta la fecha de publicación de este estudio, es escasa la información divulgada sobre el análisis de deformidades en frutas. Por tal motivo, este estudio propone dos enfoques de entrenamiento, uno de ellos mediante el uso de la técnica Transfer Learning (TL), y el otro mediante el uso de una red siamesa, los detalles de estos enfoques se dan más adelante en la sección 2.4.

El siguiente paso consistió en definir las arquitecturas CNN a usar en los enfoques de entrenamiento previamente propuestos. El estado del arte brinda buenas opciones de modelos empleados en estudios con objetivos similares, pero ninguno en el ámbito de deformidades de frutas. Por lo tanto, se propone el uso de los siguientes modelos para su posterior análisis:

- CIDIS. Utilizado en el estudio [25] con el fin de clasificar el grado de madurez del banano. Al ser un modelo exclusivo del cliente del proyecto, se decidió reutilizar para evaluar su rendimiento, ahora en la clasificación de deformidades.
- MobileNetV2. Implementado en el estudio [26] para detectar defectos como podredumbre, costas y manchas en las frutas manzana y mango. Se planteó su uso dada la eficiencia respecto a su tamaño y velocidad de predicción, resultando una gran alternativa en dispositivos con bajo poder computacional [30].
- VGG16. Red utilizada en el estudio [18]. Corresponde a uno de los pocos modelos hallados que fue utilizado para un objetivo similar al de este proyecto. Para detalles sobre la estructuras y capas, véase la Figura 1.13

2.3 Módulo 1: Creación del dataset

Este módulo abarcó la recopilación de imágenes de repositorios públicos y la generación de imágenes sintéticas para su posterior clasificación en base a los estándares de la OECD. También, se hizo uso de técnicas de preprocesamiento con el fin de eliminar ruido y garantizar la calidad de los datos, además de proveerle información más precisa a los modelos entrenados.

2.3.1 Recopilación de imágenes reales

Como todo modelo de AP, fue necesario contar con un gran número de imágenes para que los modelos sean capaces de aprender y generalizar los resultados. Partiendo desde este punto, se consideró que el número adecuado de imágenes por cada categoría de deformación era 5000; es decir, se necesitaron en total 20 000 imágenes por fruta para los 4 tipos de deformidades a clasificar. Para conseguir aquello, se consultó información en diferentes plataformas disponibles en la web como: Kaggle, Github y Mendeley. En un principio, se buscaban datasets que se encuentren clasificados de forma similar a la establecida por el estándar seleccionado (Extra, Clase I, Clase II, y Sin Clase). No obstante, esto no fue posible y no se encontró ni un solo grupo de imágenes adecuado para este caso de estudio. Por el contrario, los datasets de imágenes de frutas disponibles estaban destinados para otros campos de aplicación. Esto llevó a la decisión de acoger cualquier conjunto de imágenes y adaptarlo según lo requerido.

Tabla 2.1

Fuentes de bases de datos de imágenes reales consultadas

Base de Datos	# Manzana	# Fresa	# Mango
Fruits 360 [31]	6404	1230	916
Fruits and Vegetables dataset [32]	1200	1199	1198
Fresh and Rotten Fruits Dataset for Machine-Based Evaluation of Fruit Quality [33]	1872	1874	
Spoiled and fresh fruit inspection dataset [34]		2000	1000
Healthy-Defective-Fruits [26]	5000		5000
Fruits fresh and rotten for classification [35]	5031		
Mango Variety and Grading Dataset [36]			2200
MangoMassNet-552 Dataset [37]			552
Kesar Mango [38]			828
Sala Mango Images Dataset [39]			710

Images of individual strawberries (entire fruit, flesh and calyx) with contrasting background [40]			260	
Strawberry dataset (Good and damaged) [41]			732	
Total	19507	7295		12404

Luego, se procedió con la clasificación manual de cada una de las imágenes de las fuentes consultadas. Cada imagen fue categorizada de acuerdo con los límites admitidos para cada clase (véase Figura 1.3 hasta Figura 1.12). En paralelo, el conjunto de datos se fue refinando mediante la eliminación de imágenes de frutas que no aportaban valor significativo al entrenamiento; entre ellas, imágenes con fondo no uniforme el cual resultaba difícil de remover, e imágenes de frutas desde una vista superior o inferior.

Figura 2.2

Ejemplos de imágenes no admitidas en el conjunto de datos de entrenamiento



Al final de la clasificación, se evidenció un notorio desbalance entre las categorías de deformidad, siendo la clase Extra la que mayor número de imágenes disponía, contrario a la Clase II y Sin Clase. Para subsanar este detalle y balancear el conjunto de imágenes, se aplicó técnicas de *data augmentation* hasta un factor de 8x en las clases que lo necesitaron.

2.3.2 Generación de imágenes sintéticas

Simultáneo a la obtención del conjunto de datos reales, se procedió también con la elaboración del dataset sintético. Para esto se evaluaron diferentes herramientas y tecnologías que permitieron conseguir de forma sencilla un gran número de imágenes representativas a las categorías adoptadas. Entre ellas destacaron las IA *Text-to-image*, capaces de recibir una entrada natural de texto (o también llamado *prompt*) y generar una o varias imágenes acordes a las descripciones otorgadas.

DALL-E, lanzada en 2021 por OpenAI, es un programa de IA que genera una o varias imágenes a partir de una entrada de texto. Actualmente posee dos versiones disponibles: DALL-E mini de uso gratuito en la plataforma Hugging Face (<https://huggingface.co>), y DALL-E 2 (<https://openai.com/dall-e-2>), su versión más reciente y, por ende, más sofisticada.

Figura 2.3

Imágenes de fresas generadas por DALL-E 2

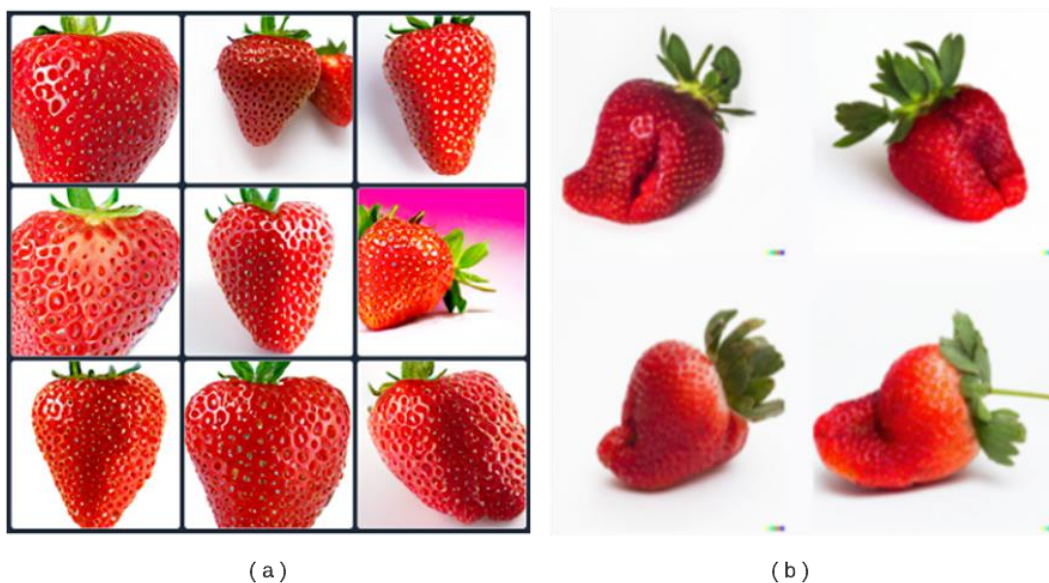


Nota: La entrada de texto empleada para estas imágenes fue “Fresa deformada”.

A pesar de ser buenas opciones, se encontró limitantes con cada una de ellas. Con DALL-E mini, por ejemplo, a pesar de la variedad de *prompts* empleados, no fue posible obtener resultados representativos para la Clase II de las frutas dado que, la mayoría de las imágenes resultantes no correspondían a lo deseado y no eran de utilidad para la clase.

Figura 2.4

Comparación de resultados entre DALL-E mini y DALL-E 2



Nota: (a) Imágenes generadas por DALL-E mini. (b) Imágenes generadas por DALL-E 2. Entrada de texto: “Misshapen strawberry”.

Nótese que en la Figura 2.4 la mayoría de las imágenes de DALL-E mini no concuerdan con lo sugerido por la entrada de texto, siendo imágenes correspondientes a Clase I e incluso Clase Extra. Con DALL-E 2 se obtuvieron resultados más precisos. Sin embargo, su principal limitante es que, en su versión gratuita, se tiene hasta un máximo de 15 pruebas por mes (Cada prueba genera 4 imágenes). Para el número de imágenes que requiere este proyecto, el uso gratuito de este software no fue posible. Dadas estas restricciones, se consideró una herramienta adicional para la generación de imágenes sintéticas, Easy Diffusion.

Easy Diffusion es una interfaz gráfica que permite el uso del software Stable Diffusion, IA Text-to-image de uso gratuito. Se procedió con la instalación de la versión 2.5 [42] en un computador con las siguientes características:

Tabla 2.2

Características del equipo de cómputo provisto por CIDIS

Característica	Valor
Procesador	Intel® Core i9-9820X
Frecuencia CPU	3.30 GHz
RAM	128 GB
Sistema Operativo	Windows 10 Pro x64 bits
GPU 0	NVIDIA GeForce Gt 1030
GPU 1	NVIDIA TITAN XP
VRAM	1982 MB
Memoria de Video	67353 MB

A diferencia de las tecnologías mencionadas, Easy Diffusion ofrece ventajas como:

- Adicionar una imagen como entrada además del texto, lo que permitió obtener resultados más afines a lo requerido.
- Ajustar el número de ejemplares que se desea generar.
- Hacer uso de la potencia del procesador y la tarjeta gráfica.
- Guardar automáticamente las imágenes generadas en un directorio específico.

Tabla 2.3

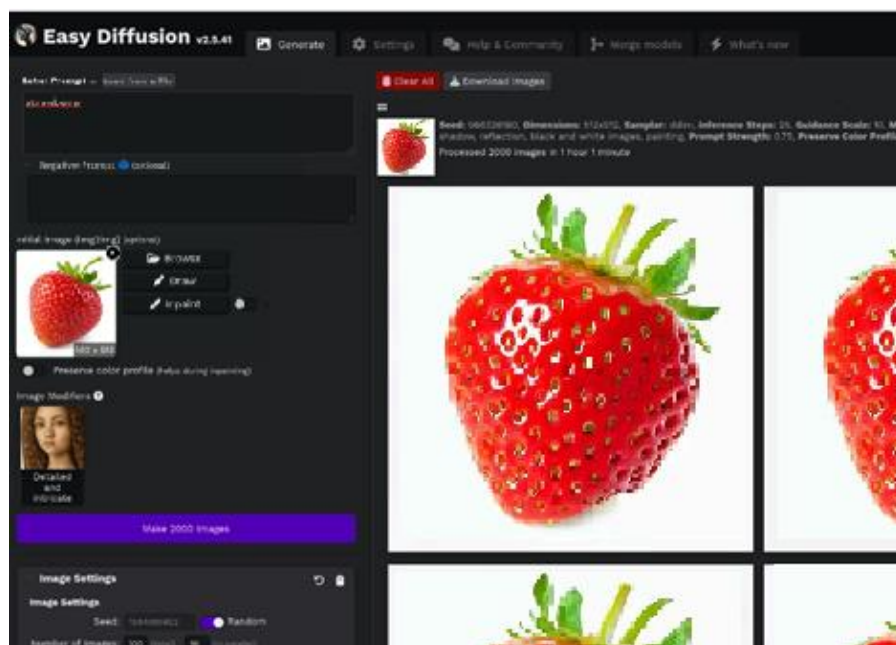
Principales entradas de texto usadas para la generación de imágenes con Stable Diffusion

Fruta	Categoría	Entradas de texto
Manzana	Extra	“apple”, “perfect apple”, “red apple”, “green apple”, “nice apple”
	Clase I	“apple with slight tilt”
	Clase II	“apple with inclination on his cusp”, “asymmetrical apple”
Mango	Extra	“mango”, “red mango”, “green mango”
	Clase I	“slightly misshapen mango”
	Clase II	“deformed mango”, “misshapen mango”

Fresa	Extra	“strawberry”, “perfect strawberry”
	Clase I	“slightly misshapen strawberry”
	Clase II	“misshapen strawberry”

Figura 2.5

Interfaz gráfica de Easy Diffusion



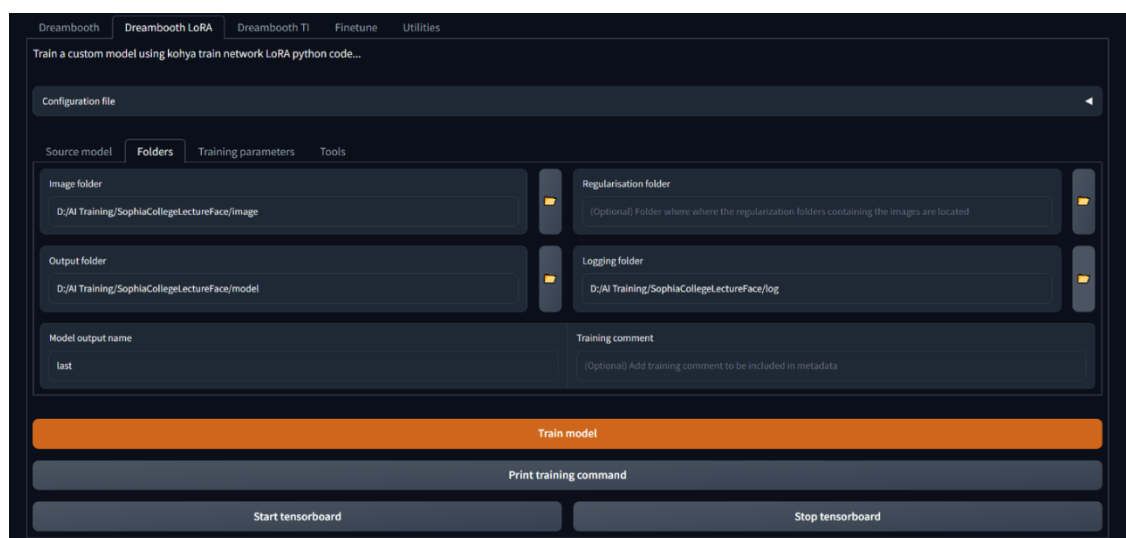
Nota: En esta imagen se ingresa la imagen de una fresa y una descripción para orientar la generación de resultados. En la parte inferior se configura el número de ejemplares deseados.

Easy Diffusion permitió obtener una gran cantidad de imágenes similares con ligeras diferencias entre ellas, simulando que fuese un dataset real. Sin embargo, se observó que no todos los resultados generados eran de utilidad. Por ejemplo, de 100 imágenes orientadas a dar resultados de manzanas Clase I, solo fue de utilidad el 50% de ellas. El resto eran imágenes caricaturescas o alejadas de la realidad. De extrapolar esto a cantidades mucho más grandes, se hubiera tenido que llevar a cabo un proceso adicional de limpieza de los datos. Por lo tanto, se decidió usar otras 2 herramientas más para generar data más precisa y evitar en lo posible el proceso de limpieza: Kohya’s Stable Diffusion trainers [43] y Automatic 1111 [44].

Basado en lo redactado por Rubén Romero en su post de Medium [45], se empezó por crear conjuntos pequeños de imágenes (entre 200 y 800) dependiendo de las diferentes escalas de formas por categoría (Extra, Clase I, Clase II, Sin Clase) y por fruta. Estas imágenes fueron obtenidas previamente con DALL-E y Easy Diffusion. Luego, se entrenó cada dataset con Kohya usando el método de entrenamiento LoRa. Esta herramienta generó un archivo “.safetensors” con los pesos para ajustar el Stable Diffusion y generar imágenes de acuerdo con lo entrenado en el modelo. Luego, con el modelo cargado se utilizó la herramienta Automatic 1111 que, similar a Easy Diffusion, permite generar imágenes mediante las modalidades Image-to-image o Text-to-image.

Figura 2.6

Interfaz gráfica de Kohya

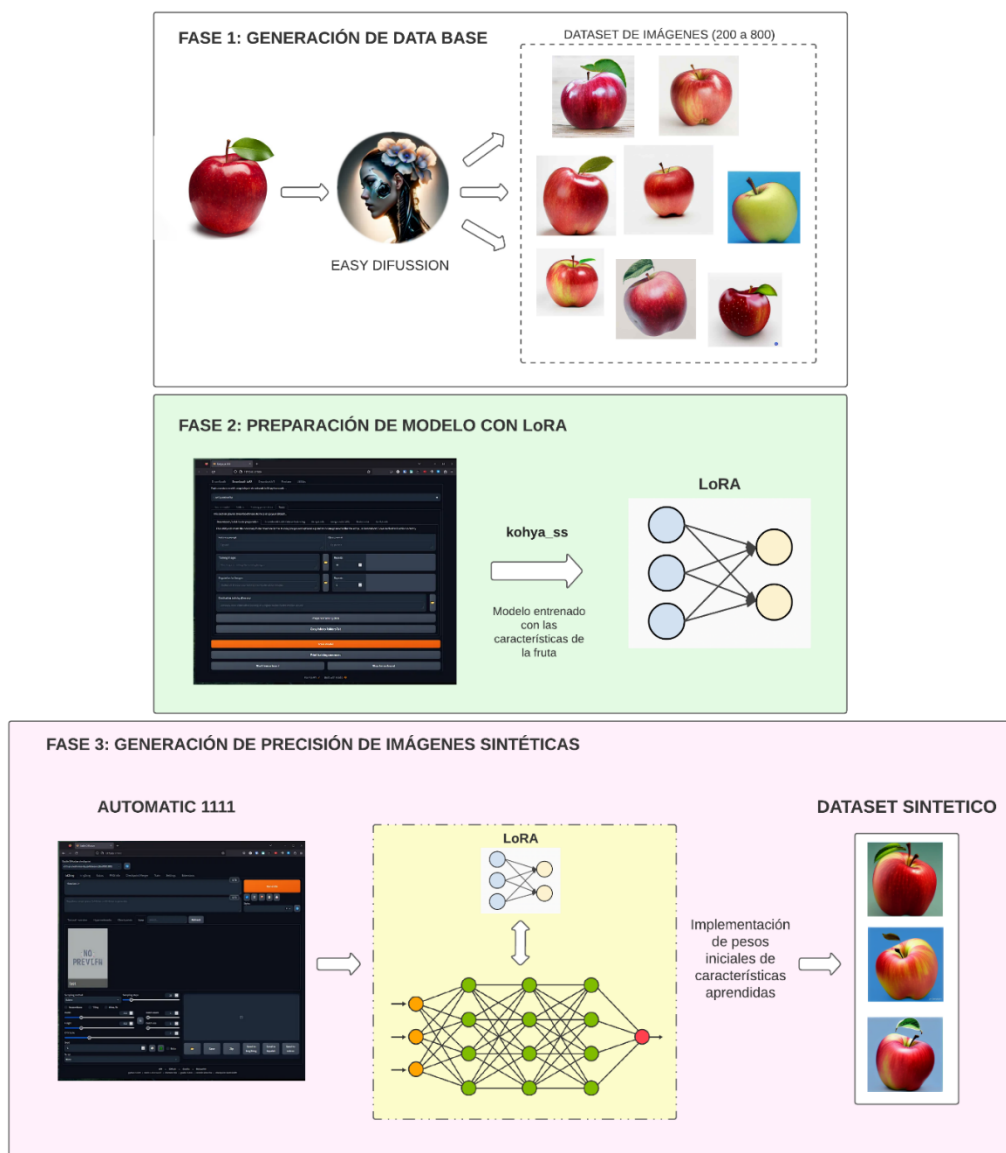


Nota: Interfaz de Kohya donde se especifica la ruta de carpeta con las imágenes para entrenar y la ruta de salida del archivo “.safetensors”.

Con el uso en conjunto de estas herramientas se logró mejorar considerablemente el resultado de las imágenes sintéticas por cada categoría, en comparación con los resultados de Easy Diffusion. Esto permitió ahorrar tiempo y llevar un mayor control de los resultados generados por Stable Diffusion. El siguiente esquema resume todo el procedimiento seguido:

Figura 2.7

Fases para la generación de imágenes sintéticas



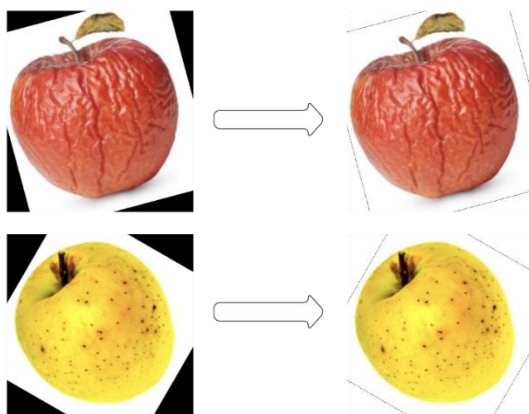
2.3.3 Preprocesamiento de imágenes

Como parte del aseguramiento de la calidad de los datos de entrenamiento, se llevó a cabo un proceso de preprocesamiento en aquellas imágenes que podían introducir ruido a los modelos. Entre las imágenes tratadas se encuentran aquellas que pasaron por un proceso de *data augmentation* mediante la técnica de rotación, de modo que, las esquinas o bordes quedaron sin intensidad en sus píxeles. Aquel cambio de contraste tan súbito entre el borde

oscuro y el background claro de la imagen, podía afectar negativamente el aprendizaje de los modelos. Por lo tanto, se optó por rellenar aquellos espacios vacíos con un color neutro, tal que, el aprendizaje sea afectado en menor medida.

Figura 2.8

Preprocesamiento de imágenes producto de técnica de Rotación



Similar al caso anterior, se encontraron imágenes en las que se utilizó la técnica de mosaico, donde porciones de otras imágenes se reflejaban en los bordes. Dado que solo interesaba la región central de la imagen, se procedió a remover estos bordes mediante la herramienta Segment Anything Model (SAM) (<https://segment-anything.com>). SAM, desarrollado por Meta, es capaz de segmentar precisamente cada objeto de una imagen, mediante la generación de máscaras por cada uno de ellos [46].

Figura 2.9

Segmentación de objetos con SAM



Figura 2.10

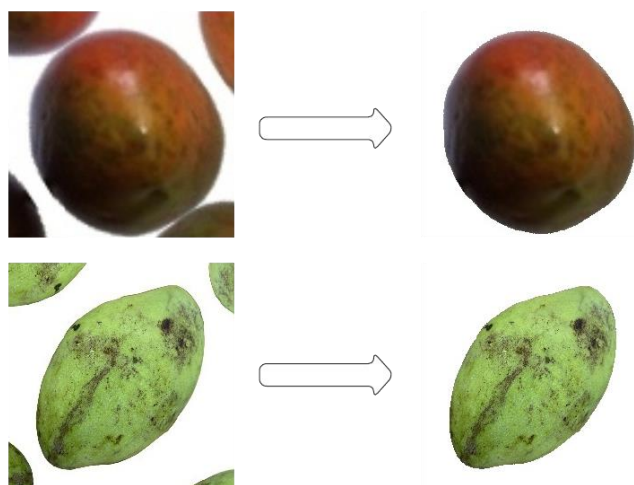
Identificación de todas máscaras de una imagen



Con el uso de las múltiples máscaras fue posible emplear operaciones binarias con la librería OpenCV de Python para remover los objetos que no eran de interés.

Figura 2.11

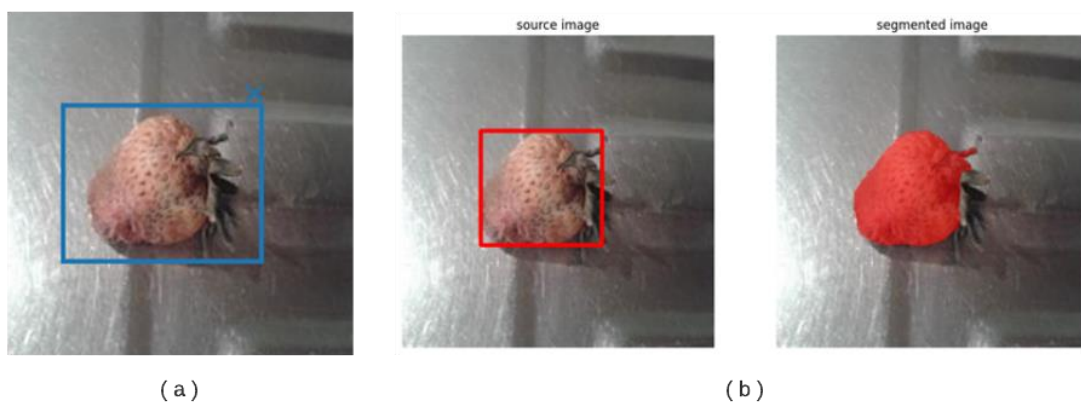
Preprocesamiento de imágenes producto de técnica Mosaico



Finalmente, se operó con aquellas imágenes cuya textura del background no era uniforme. Para esto también se utilizó SAM pero con el método de Caja, donde se delimitó manualmente la región de interés (ROI) por cada imagen, permitiéndole a la herramienta definir con mayor precisión el objeto a segmentar.

Figura 2.12

Uso de método de ROI para segmentación de la fruta

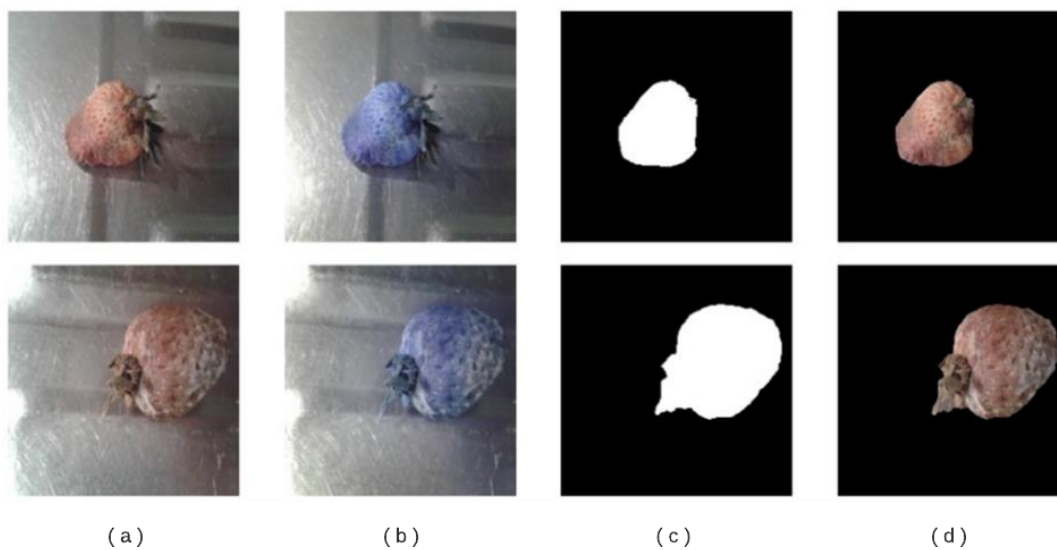


Nota: (a) ROI dibujado manualmente, (b) ROI precisado por SAM y segmentación de la fruta.

Una vez obtenida la máscara del objeto, se procedió nuevamente a utilizar OpenCV para la segmentación de la fruta de la imagen original.

Figura 2.13

Resultado de eliminación de background no uniforme con herramienta SAM



Nota: (a) Imagen RGB original, (b) Imagen BGR, (c) Máscara obtenida de la imagen, (d) Segmentación de la fruta.

Por otro parte, si fue posible encontrar arquitecturas VGG16 y MobileNetV2 previamente entrenadas para clasificación, de modo que se tomaron los pesos resultantes de aquel entrenamiento y se usaron como base para el nuevo entrenamiento con únicamente las imágenes reales de las frutas.

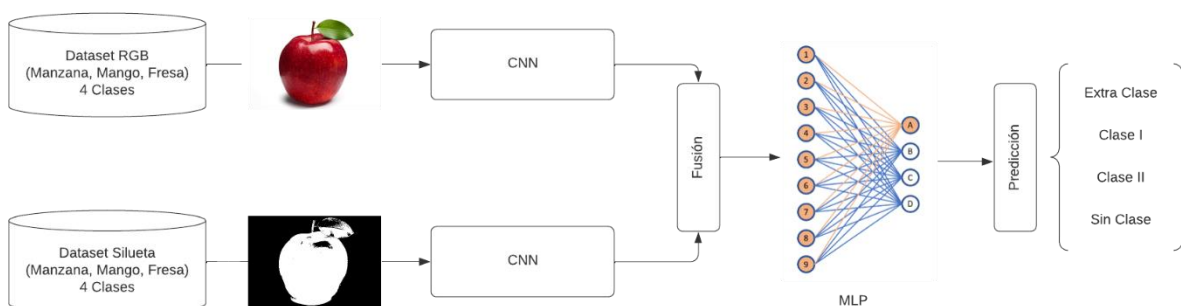
2.4.2 Enfoque de Red Siamesa

Entendiendo la naturaleza del estudio de deformidades en frutas, el segundo enfoque, ilustrado en la Figura 2.15, se centra mucho más en el análisis de la forma mediante la implementación de una arquitectura siamesa. Este enfoque es recurrentemente usado en estudios orientados al reconocimiento y detección facial [47], dónde se evalúan principalmente diferencias entre las imágenes de entrada de cada rama. No obstante, se plantea implementar una arquitectura similar a la utilizada en el estudio [24], donde se unifica el resultado obtenido por cada rama mediante un proceso de fusión de características y se logra dar una clasificación luego de una capa Multilayer Perceptron (MLP).

Dado que este enfoque prioriza el estudio de la forma de las frutas, fue necesario ignorar características usualmente reconocidas por las CNN como el color y textura de las imágenes. Por ende, se decidió que las imágenes de entrada de una de las ramas sean las imágenes RGB de las frutas (provenientes del conjunto real y sintético), mientras que la otra rama recibiera las siluetas de las frutas de esas mismas imágenes, tal y como se aprecia en la Figura 2.15.

Figura 2.15

Esquema de Red Siamesa



Para este enfoque se utilizaron nuevamente los modelos CIDIS, MobileNetV2 y VGG16 con la diferencia de que el modelo CIDIS fue entrenado sin pesos iniciales, mientras que los modelos MobileNetV2 y VGG16 fueron inicializados con los mismos pesos que se hallaron previo al entrenamiento con el enfoque de TL.

2.5 Módulo 3: Evaluación de los modelos de CNN

Todos los modelos entrenados fueron evaluados con las siguientes métricas:

Matriz de Confusión. Permite la visualización de la eficiencia de la predicción por cada categoría, define los conceptos:

- *True Positive (TP)*, *True Negative (TN)*, Frutas clasificadas correctamente según su categoría.
- *False Positive (FP)*, *False Negative (FN)*. Predicción incorrecta de la categoría de las frutas.

Figura 2.16

Matriz de confusión multiclase

MATRIZ DE CONFUSIÓN - $MC(i, j)$ de $N \times N$

		CLASE I					CLASE N-I				
Real = 1	Real = 1	TP	FN	FN	FN	FN	TN	TN	TN	FP	TN
	Real = 2	FP	TN	TN	TN	TN	TN	TN	TN	FP	TN
	Real = ...	FP	TN	...	TN	TN	TN	TN	...	FP	TN
	Real = N-1	FP	TN	TN	TN	TN	FN	FN	FN	TP	FN
	Real = N	FP	TN	TN	TN	TN	TN	TN	TN	FP	TN
		Predicted = 1	Predicted = 2	Predicted = ...	Predicted = N-1	Predicted = N	Predicted = 1	Predicted = 2	Predicted = ...	Predicted = N-1	Predicted = N

Nota: Obtenido de [48].

Accuracy. También denominada Exactitud, esta métrica muestra el porcentaje general de frutas clasificadas correctamente según su grado de deformidad.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.1)$$

Precision. Esta métrica muestra el porcentaje de frutas clasificadas correctamente según una categoría específica. Usada para valorar la confiabilidad del modelo al evaluar una de las clases de deformidad.

$$Precision = \frac{TP}{TP + FN} \quad (2.2)$$

Recall. Empleada para verificar cuan bien el modelo puede detectar una de las clases de deformidad.

$$Recall = \frac{TP}{TP + FN} \quad (2.3)$$

F1-score. Esta métrica relaciona la información del recall y la precision.

$$F1 - score = \frac{2x Precision x Recall}{Precision + Recall} \quad (2.4)$$

2.6 Módulo 4: Desarrollo de Prototipo

Para el desarrollo del prototipo se propuso un despliegue web que permitiera cargar una imagen y seleccionar el tipo de fruta a la que se desea detectar la deformidad. Dada la finalidad del proyecto, se desarrolló una pequeña aplicación web que sirva únicamente para pruebas y visualización de resultados. Se utilizó el modelo con mejor rendimiento para cada una de las frutas.

Dado que se empleó Python para el desarrollo de los modelos, fue sencilla su integración con el framework Flask para el desarrollo de la interfaz gráfica. Se desarrollaron bocetos base con los que se orientó el diseño del prototipo.

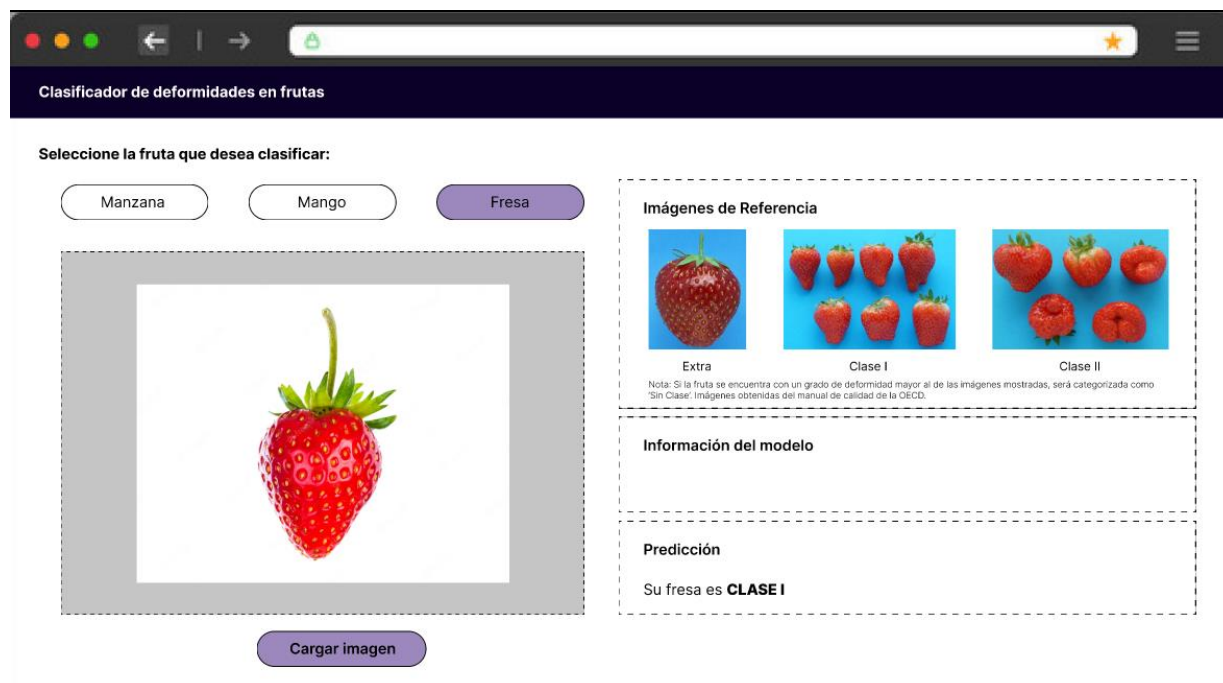
Figura 2.17

Boceto Inicial del prototipo web



Figura 2.18

Boceto Final del prototipo web



Capítulo 3

3.1 Resultados y análisis

En este capítulo se detallan los resultados obtenidos luego de la ejecución de la metodología planteada en el capítulo anterior. Se empieza por describir la generación del conjunto de imágenes reales y sintéticas, además de ciertas consideraciones adicionales que se fueron implementando a lo largo de la fase de entrenamiento.

Luego, se detallan los resultados obtenidos por cada enfoque de entrenamiento mencionando cada uno de los modelos de CNN empleados. Para cada entrenamiento se utilizó diferentes combinaciones de hiperparámetros. La selección del mejor modelo por fruta se realizó mediante la evaluación y análisis de cada una de las métricas seleccionadas.

Finalmente, se muestra el prototipo de aplicación web que utiliza los mejores modelos de CNN para realizar la clasificación de la deformidad en las frutas objeto de estudio y visualizar los resultados.

3.2 Creación del conjunto de imágenes de entrenamiento

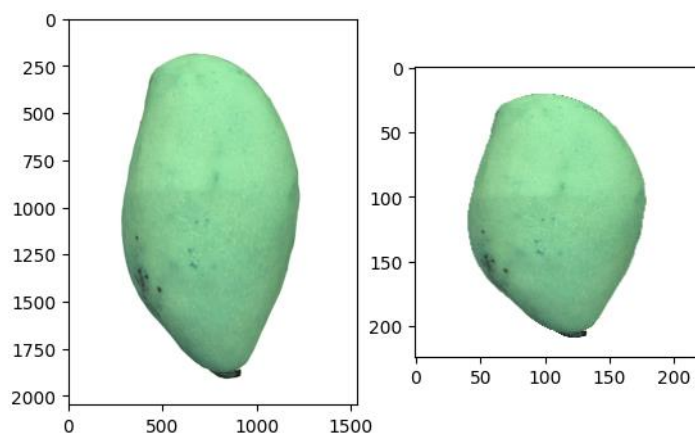
3.2.1 *Dataset Real Versión 1 (DRV1)*

Para la creación del dataset real se procedió con la clasificación manual de cada una de las imágenes encontradas en las fuentes de bases de datos listadas en la Tabla 2.1. Dada la variabilidad entre cada fuente consultada, se obtuvieron imágenes de frutas con diferentes características de color, tamaño, forma, entre otros.

Adicionalmente, muchas de las imágenes no eran cuadradas, es decir, no tenían el mismo ancho y alto de píxeles. Dada la naturaleza de los modelos de CNN los cuales requieren de imágenes cuadradas para el entrenamiento, fue necesario aplicar un proceso adicional de rellenado de bordes en ciertas imágenes, con el fin de que la forma de la fruta no se vea afectada al ingresar a los modelos en la fase de entrenamiento.

Figura 3.1

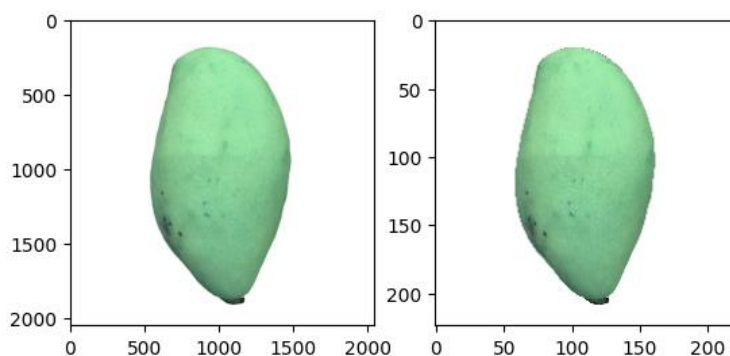
Reajuste de tamaño de imagen sin relleno de bordes



Nota: A la izquierda se presenta la imagen original con un tamaño de 1536x2048 pixeles. A la derecha se presenta la imagen reajustada a un tamaño de 224x224 pixeles (Tamaño de entrada de los modelos de CNN). Se evidencia una desproporción considerable de la fruta, afectando así su categoría de deformación.

Figura 3.2

Reajuste de tamaño de imagen con relleno de bordes



Nota: A la izquierda se presenta la imagen original luego de pasar por el proceso de relleno, presentando una forma simétrica de 2048x2048 pixeles. A la derecha se muestra la imagen reajustada a 224x224 pixeles. Se evidencia que la fruta mantiene su forma.

Una vez recopilado el conjunto de imágenes reales de frutas con los mínimos requeridos (650), se procedió al uso de técnicas de Aumento de Datos para alcanzar el total de imágenes acordado para el entrenamiento. El tamaño de las imágenes resultantes fue configurado con un tamaño de 224x224 píxeles y se aplicaron técnicas de rotación, giros horizontales y giros verticales.

En resumen, se obtuvo un total de 20 000 imágenes por fruta, distribuidas equitativamente entre las 4 categorías de deformidad.

Tabla 3.1

Cantidad de imágenes reales en DRV1 obtenidas para el entrenamiento

	Extra	Clase 1	Clase 2	Sin Clase
Manzanas	4050 (950)	1741 (3259)	651 (4349)	1961 (3039)
Fresas	2782 (2218)	875 (4125)	759 (4241)	679 (4321)
Mangos	4317 (683)	1889 (3111)	655 (4345)	734 (4266)

Nota: Las cifras en paréntesis corresponden al total de imágenes resultantes de las técnicas de Aumento de Datos.

3.2.2 Dataset Real Versión 2 (DRV2)

Una vez terminada la creación del DRV1, se procedió con el entrenamiento de los modelos del enfoque TL. No obstante, de acuerdo con los resultados mostrados en la Tabla 3.3 de la sección 3.3.1 de este documento, se evidenció la necesidad de aplicar un proceso de refinamiento a cada conjunto de imágenes generado. Por lo tanto, se procedió con el análisis de las imágenes contenidas en los datasets en busca de patrones que expliquen el estancamiento de los resultados. Luego del análisis se evidenció lo siguiente:

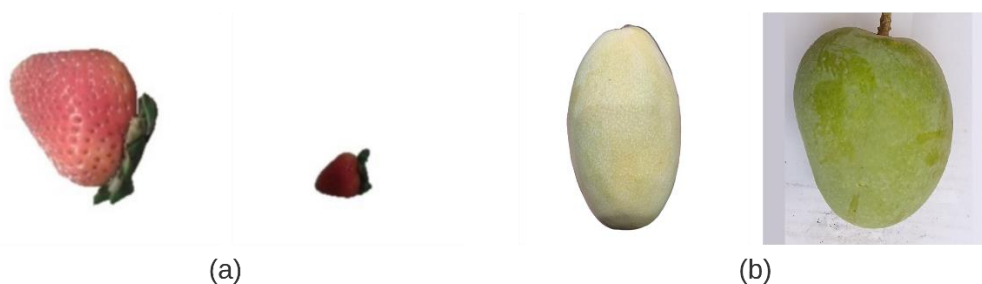
1. Distancia de trabajo variada entre las imágenes: Muchas de las imágenes discrepaban entre ellas por la distancia de trabajo aplicada al fotografiar la fruta, resultando en imágenes con frutas muy pequeñas, e imágenes con frutas ocupando la imagen casi en

su totalidad. Se decidió eliminar aquellas imágenes cuya distancia de trabajo era muy grande.

2. Distintas familias de frutas: Diferencias encontradas especialmente en mangos, donde muchos de ellos eran alargados y otros eran redondos. Si bien cada mango se encontraba correctamente categorizado, se decidió eliminar aquellas familias de mangos que tenían menor número de imágenes.
3. Clasificación errónea: Se encontraron ciertas imágenes mal categorizadas, por lo que se procedió a reasignarla a su categoría adecuada.

Figura 3.3

Ejemplos de diferencias encontradas entre las imágenes de los datasets



Nota: A la izquierda, diferencia de distancia de trabajo encontrada en fresas. A la derecha, mangos categorizados en Clase Extra pero perteneciendo a distintas familias.

Luego de esto, cada fruta obtuvo un total de 20 000 imágenes, exceptuando el de manzanas que obtuvo 16 000. Esta nueva versión del dataset (DRV2) trajo consigo resultados más certeros, por lo que, fue directamente utilizado en los entrenamientos con enfoque de red siamesa.

3.2.3 Dataset Sintético (DS)

Dada las facilidades otorgadas por Stable Diffusion y las diferentes tecnologías utilizadas, se generaron imágenes sintéticas de muy buena calidad con un tamaño de 512x512 píxeles. En muchas de ellas, se emplearon imágenes base obtenidas del conjunto de datos reales con el fin de generar imágenes similares.

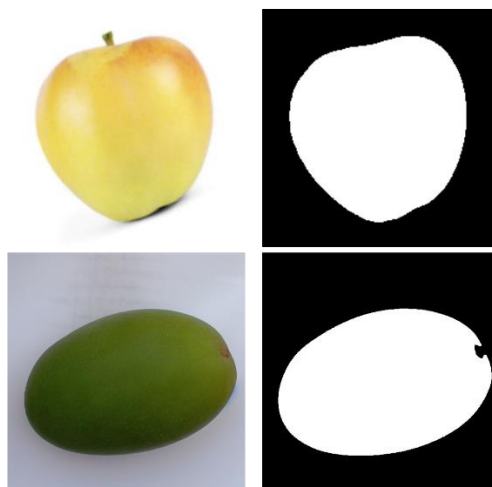
Luego, se procedió con uso de técnicas de Aumento de Datos para rotar y girar las imágenes con el fin de darle variabilidad al dataset. Similar al conjunto de imágenes reales, se obtuvo un total de 20 000 imágenes por frutas, distribuidas equitativamente entre las categorías de deformidad.

3.2.4 Dataset de siluetas de frutas

Luego de la creación del DRV2 y la generación de imágenes sintéticas, se realizó la unificación de estos conjuntos de imágenes dada la metodología planteada para el entrenamiento con enfoque de red siamesa. Después, mediante la elaboración de un script en Python y la herramienta SAM, se definieron conjuntos de reglas que permitieron automatizar el proceso de segmentación y generación de silueta de cada una de las imágenes.

Figura 3.4

Obtención de silueta de cada fruta



La generación de la silueta de la mayoría de las imágenes fue correcta. Por otro lado, aquellas imágenes de fruta, cuyas siluetas no fueron correctas, se les dio un tratamiento individual o de plano fueron descartadas.

Con los conjuntos de imágenes en RGB y silueta ya listos, se procedió al entrenamiento de los modelos de CNN mostrados en la sección 3.3.2 de este documento.

3.3 Entrenamiento de modelos de CNN

En la Tabla 3.2 se muestra el conjunto de hiperparámetros empleados para el entrenamiento de cada uno de los modelos. Cada entrenamiento fue llevado a cabo con una combinación distinta según los resultados que se iban obteniendo. Se evidencia el uso de diferentes optimizadores; entre ellos Adam, Nadam y RMSProp, y diferentes Ratios de Aprendizaje. Generalmente, los entrenamientos fueron configurados para ejecutarse hasta la época 30. No obstante, dada la inclusión de otras funcionalidades como *Early Stopping* y *Model Checkpoint* provistos por la librería Keras, los entrenamientos se detenían automáticamente en caso de no presentar signos de mejora a lo largo de las épocas, además de restaurar los pesos aprendidos de la época que mejor evaluación presentó.

Tabla 3.2

Combinación de hiperparámetros utilizados en los entrenamientos

Optimizador	Learning Rate	Épocas	Dropout
Adam, Nadam, RMSProp	0.001, 0.0001, 0.00001	1-30	NA, 0.2, 0.3, 0.5

Es necesario detallar también que se realizó una modificación a la estructura del modelo CIDIS. Este modelo fue originalmente implementado para la clasificación de nivel de madurez en bananos. No obstante, dados los resultados que iba presentando para este estudio, se decidió agregar una capa de regularización con el fin de optimizar los resultados según la pérdida y precisión de los conjuntos de entrenamiento y validación. De este modo, el modelo logró mejorar los resultados en comparación a la arquitectura antecesora. Para efectos del estudio, este nuevo modelo se denominó CIDIS V2.

3.3.1 Resultados con enfoque Transfer Learning

La Tabla 3.3 y Tabla 3.4 resumen los resultados obtenidos en los entrenamientos de este enfoque. En la primera de ellas se resalta el uso de la primera versión del conjunto de imágenes (DRV1), de modo que, los modelos no obtuvieron resultados precisos, y esto es avalado por cada una de las métricas de evaluación. No obstante, se destaca que la arquitectura MobileNetV2 fue la que mejores resultados logró obtener en cada fruta.

También, se logra observar el resultado de las métricas *Loss* y *Accuracy* del conjunto de validación (Val Loss y Val Accuracy respectivamente). El resultado de estas métricas discrepaban totalmente de las del conjunto de entrenamiento, de manera que se percibió la necesidad de revisar el conjunto de datos generado.

Tabla 3.3

Resultados con DRV1 – Enfoque Transfer Learning

Fruta	Modelo	Val Accuracy	Val Loss	Precision	Recall	F1-score	Test Accuracy
Manzana	CIDIS V2	0.6815	0.8491	0.6100	0.5900	0.5500	0.5915
	MobileNetV2	0.7848	0.5851	0.6100	0.6100	0.5700	0.6095
	VGG16	0.7747	0.6073	0.5600	0.5700	0.5200	0.5685
Mango	CIDIS V2	0.5155	1.2260	0.4400	0.4100	0.4100	0.4400
	MobileNetV2	0.5943	0.9612	0.6700	0.6700	0.6700	0.6710
	VGG16	0.5746	1.0673	0.4900	0.4800	0.4800	0.4770
Fresa	CIDIS V2	0.5601	1.3034	0.4900	0.4100	0.4000	0.4110
	MobileNetV2	0.6971	1.2005	0.7000	0.7000	0.7000	0.6980
	VGG16	0.6694	1.0141	0.6300	0.6100	0.6000	0.6100

Luego de lo expuesto en la sección 3.2.2, se procedió al reentrenamiento de los modelos con una versión mejorada del dataset de cada fruta (DRV2). Los resultados expuestos detallan una mejora considerable en la evaluación de cada una de las métricas. Tal es el caso, que todos los modelos obtuvieron una exactitud en el conjunto de pruebas por encima del 80%.

Tabla 3.4*Resultados con DRV2 – Enfoque Transfer Learning*

Fruta	Modelo	Val Accuracy	Val Loss	Precision	Recall	F1-score	Test Accuracy
Manzana	CIDIS V2	0.8956	0.3645	0.8900	0.8800	0.8800	0.8806
	MobileNetV2	0.9325	0.2087	0.9100	0.9000	0.9000	0.9044
	VGG16	0.9050	0.2511	0.9000	0.8900	0.8900	0.8931
Mango	CIDIS V2	0.8679	0.3819	0.8700	0.8700	0.8700	0.8694
	MobileNetV2	0.9420	0.1594	0.9400	0.9400	0.9400	0.9425
	VGG16	0.9022	0.2290	0.9600	0.9600	0.9600	0.9604
Fresa	CIDIS V2	0.8010	0.5258	0.8200	0.8100	0.8200	0.8135
	MobileNetV2	0.9189	0.1945	0.9200	0.9200	0.9200	0.9220
	VGG16	0.8962	0.2646	0.9100	0.9100	0.9100	0.9050

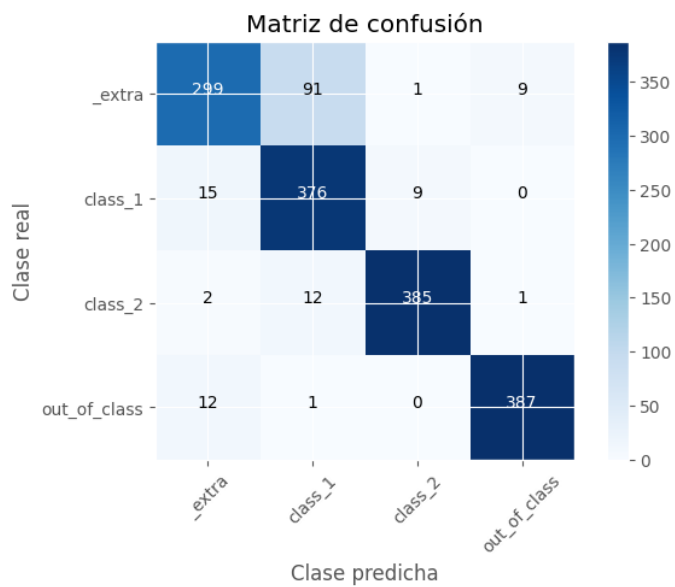
Para finalizar con este enfoque de entrenamiento, se procedió con la selección del mejor modelo obtenido para cada fruta. En el caso de las manzanas y fresas, no fue difícil seleccionar a la arquitectura MobileNetV2, dado que todas sus métricas evaluadas superan a la del resto de arquitecturas.

En el caso del mango, se dio una situación especial dónde el modelo VGG16 supera en exactitud en el conjunto de pruebas al modelo MobileNetV2. No obstante, este último modelo supera en el resto de las métricas al VGG16. Considerando una menor pérdida en el conjunto de validación y un mayor valor de recall, lo que asegura una predicción más confiable de la deformidad de la fruta, se ha seleccionado también el modelo MobileNetV2.

A continuación, se colocan las matrices de confusión obtenidas con el mejor modelo para cada fruta:

Figura 3.5

Matriz de confusión obtenida con el mejor modelo para Manzana – Enfoque Lineal

**Figura 3.6**

Matriz de confusión obtenida con el mejor modelo para Mango – Enfoque Lineal

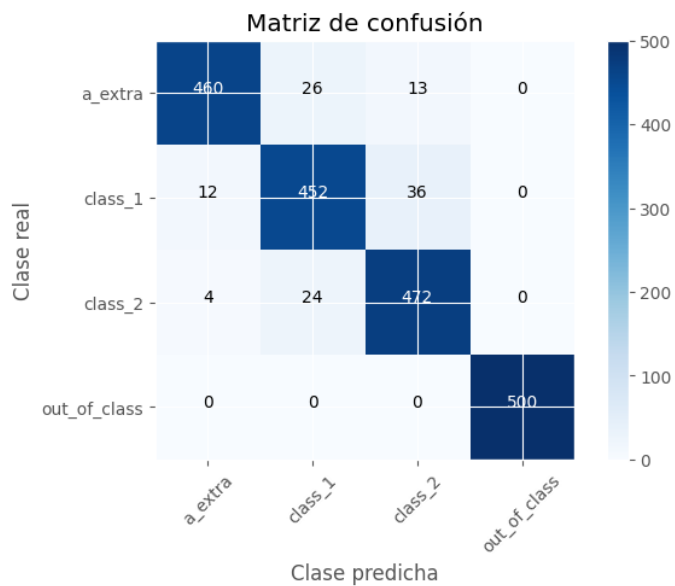
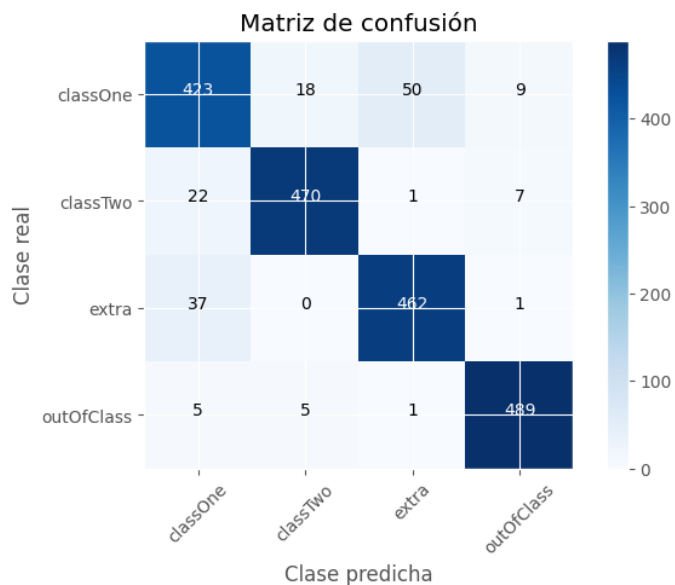


Figura 3.7

Matriz de confusión obtenida con el mejor modelo para Fresa – Enfoque Lineal

**Tabla 3.5**

Mejores resultados – Enfoque Lineal

Fruta	Modelo CNN	Accuracy
Manzana	MobileNetV2	0.9044
Mango	MobileNetV2	0.9425
Fresa	MobileNetV2	0.9220

3.3.2 Resultados con enfoque de Red Siamesa

Como se mencionó anteriormente, este enfoque empleó directamente el DRV2, sumado al DS, y de ellos se obtuvo la silueta. La Tabla 3.6 muestra los resultados obtenidos en los entrenamientos con las diferentes arquitecturas de CNN en mangos y manzanas. Se evidenció buenos resultados en mangos. Esto pudo darse dado que fue el conjunto de datos con la mayor cantidad de imágenes luego de haber realizado la segmentación y refinamiento de los datos.

Todos los modelos obtuvieron exactitudes por encima del 90%, siendo la arquitectura VGG16 la que mejor rendimiento presentó.

Tabla 3.6

Resultados con DRV2+DS+Silueta – Enfoque de Red Siamesa

Fruta	Modelo	Val Accuracy	Val Loss	Precision	Recall	F1-score	Test Accuracy
Manzana	CIDIS V2	0.7626	1.6818	0.8639	0.8366	0.8107	0.8249
	MobileNetV2	0.8068	2.2276	0.9257	0.9158	0.9161	0.9121
	VGG16	0.8060	1.9643	0.8966	0.8920	0.8832	0.8844
Mango	CIDIS V2	0.9234	0.2577	0.9237	0.9232	0.9233	0.9232
	MobileNetV2	0.9440	0.1505	0.9394	0.9393	0.9390	0.9393
	VGG16	0.9501	0.1696	0.9551	0.9539	0.9541	0.9539

En el caso de manzanas, destacó nuevamente la arquitectura MobileNetV2 obteniendo una exactitud del 91%. Con estos resultados podemos inferir que en fresas se obtendrían resultados similares, dando posibilidad a mejoras en la precisión de la predicción de una fruta.

Figura 3.8

Matriz de confusión obtenida con el mejor modelo para Mango – Enfoque de Red Siamesa

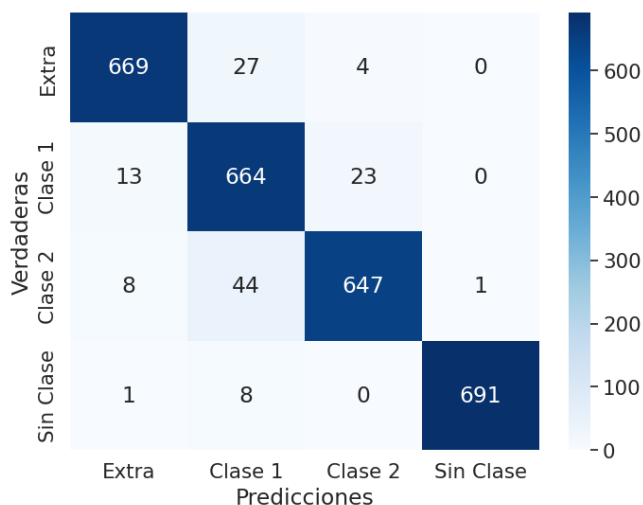


Figura 3.9

Matriz de confusión obtenida con el mejor modelo para Manzana – Enfoque de Red Siamesa



La Tabla 3.7 resume los resultados con el enfoque de Red Siamesa:

Tabla 3.7

Mejores Resultados – Enfoque de Red Siamesa

Fruta	Modelo CNN	Accuracy
Manzana	MobileNetV2	0.9257
Mango	VGG16	0.9550

Finalmente, la Tabla 3.8 realiza una comparación de los mejores resultados obtenidos en ambos enfoques de entrenamiento en cada fruta del estudio. Se evidencia mejoras en las métricas de evaluación al usar una Red Siamesa, principalmente la métrica de exactitud. Se asume que esto se debe al uso adicional de imágenes de silueta, dado que, se remueve por completo factores ajenos al análisis de este estudio, como el color y presencia de defectos en las frutas.

Tabla 3.8

Comparación de resultados Enfoque Transfer Learning vs Enfoque de Red Siamesa

	Enfoque TL	Enfoque de Red Siamesa
Manzana	MobileNetV2 (90.44%)	MobileNetV2 (92.57%)
Mango	MobileNetV2 (94.25%)	VGG16 (95.50%)

Fresa	MobileNetV2 (92.20%)	NA
-------	----------------------	----

Nota: Los porcentajes mostrados corresponden a la exactitud obtenida en el conjunto de pruebas luego del entrenamiento.

3.4 Prototipo

Luego del análisis de cada uno de los modelos de CNN obtenidos en ambos enfoques de entrenamiento, se procedió con la implementación de un prototipo de aplicación web que permitiera clasificar imágenes de manzanas, mangos y fresas según su categoría de calidad de deformación. Para el desarrollo de la interfaz gráfica y el backend, se utilizó el framework Flask con el lenguaje de programación Python. Los modelos empleados fueron los obtenidos en el enfoque TL.

Figura 3.10

Prototipo aplicado en manzanas

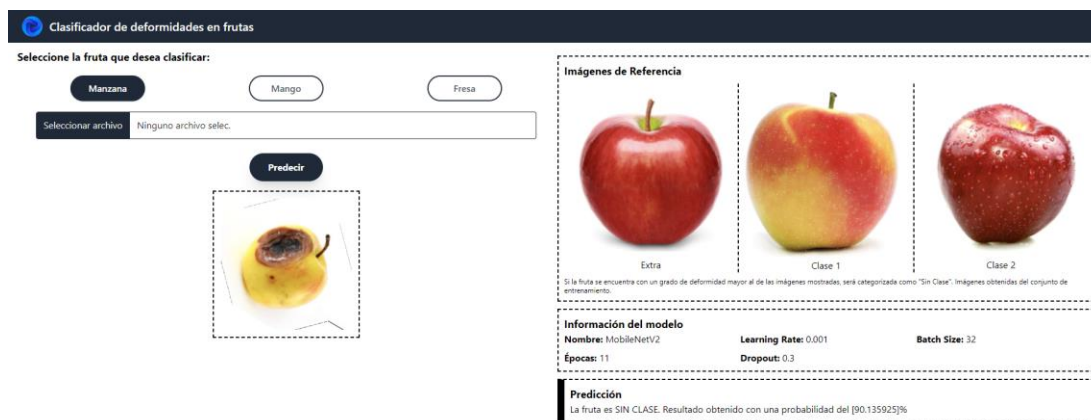


Figura 3.11

Prototipo aplicado en mangos

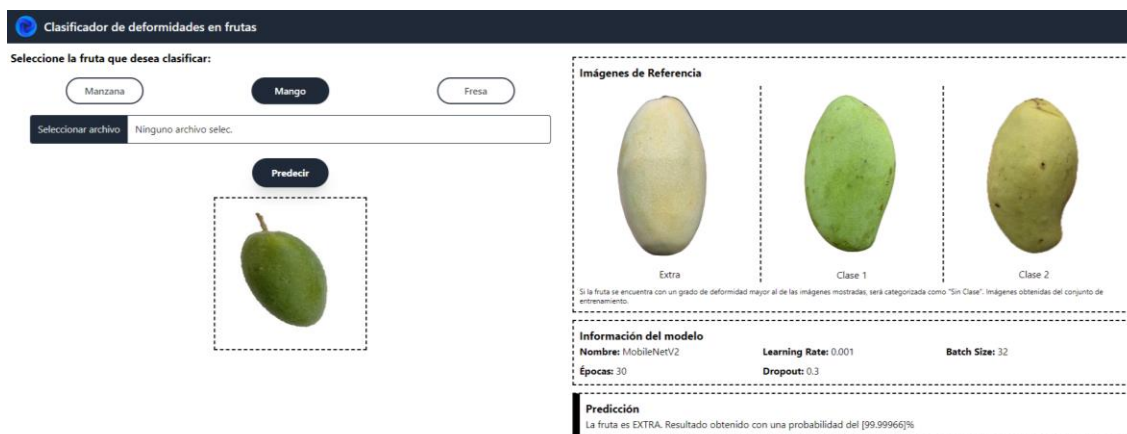
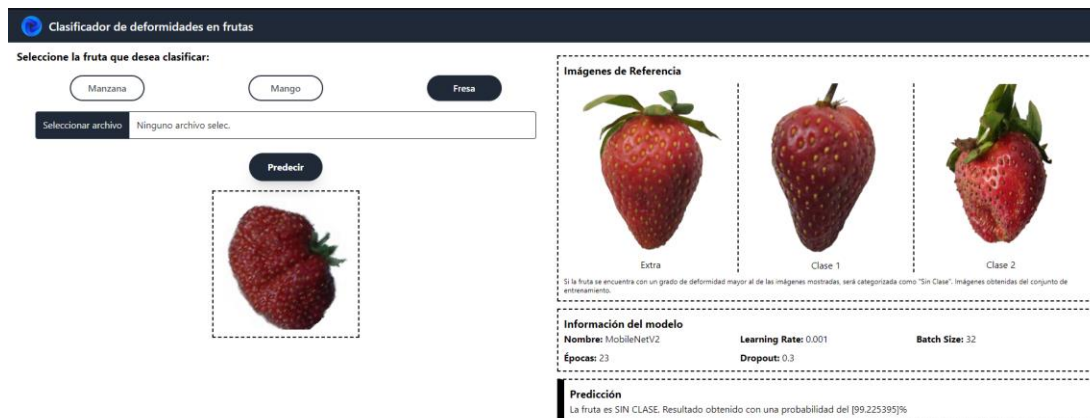


Figura 3.12

Prototipo aplicado en fresas



Capítulo 4

4.1 Conclusiones y recomendaciones

En este capítulo se redacta, en orden de relevancia, las conclusiones obtenidas a lo largo del desarrollo del proyecto. Además, se brindan recomendaciones que pueden ser aplicadas con el fin de reproducir y mejorar los resultados obtenidos. Finalmente, se brinda una visión sobre futuros proyectos enlazados y relacionados con esta área de estudio.

4.2 Conclusiones

- El análisis de la forma de las frutas conlleva una amplia cantidad de consideraciones; entre ellas, la necesidad de asegurar que la imagen posea sólo los elementos de interés para los modelos de CNN. Por lo tanto, es primordial aplicar un adecuado preprocesamiento que logre eliminar detalles que puedan introducir ruido a los modelos de CNN, como fondo u objetos ajenos a la fruta. Adicionalmente, cabe precisar que las imágenes deben ser completamente cuadradas con el fin de evitar que la fruta se vea afectada al reajustar el tamaño de la imagen.
- Se generaron conjuntos de imágenes adecuados y de calidad para las frutas: manzana, mango y fresa. Cada una de las imágenes recolectadas y/o generadas pasó por una estricta revisión con el fin de asegurar su idoneidad para este estudio. Al final se obtuvieron datasets de las frutas previamente mencionadas, rigurosamente categorizadas según la calidad de deformación establecida por la OECD.
- El uso de herramientas de IA *Text to image* como Stable Diffusion, permite diversificar el conjunto de imágenes con el fin de generalizar los resultados. Adicionalmente, permite minimizar los costos y el esfuerzo de obtención de imágenes.
- Luego del entrenamiento de cada uno de los modelos de CNN con sus distintos enfoques de entrenamiento, y la revisión y análisis de las métricas de evaluación, se puede concluir que la arquitectura MobileNetV2 es la que mejores resultados otorga al realizar la

clasificación de la deformidad en las frutas: manzana, mango y fresa a través del enfoque TL. Además, se destacan otras propiedades como su eficiente velocidad de predicción y su tamaño ligero, optimizando los recursos donde sea alojado. Su implementación en la industria alimentaria permitirá automatizar el proceso de clasificación de deformidades de las frutas, disminuyendo cuantiosamente la incidencia de errores de clasificación.

- Enfoque TL y Enfoque de Red Siamesa ofrecen buenos resultados en la clasificación de deformidades. No obstante, se concluye que el enfoque de Red Siamesa ofrece mejores resultados al realizar el análisis de deformidades en las frutas.
- Se desarrolló un prototipo de aplicación web que permite a sus usuarios subir imágenes de frutas con el fin de obtener una clasificación de su deformidad. La interfaz es amigable e intuitiva, proveyendo sólo las opciones necesarias con el fin de brindar una agradable experiencia a sus usuarios. Además, cuenta con los modelos de CNN más precisos y eficientes para realizar las distintas predicciones, de modo que, cumple con varios principios de usabilidad.

4.3 Recomendaciones

- Con el fin de obtener resultados más precisos y disminuir considerablemente las etapas de preprocesamiento, se recomienda ensamblar un ambiente controlado para la obtención de imágenes de las frutas. Esto permitiría tener un mayor control por sobre las imágenes del dataset, evitando así imágenes que puedan introducir ruido o tener distintas distancias de trabajo.
- Dada la escasa disponibilidad de las imágenes de frutas en los repositorios públicos, este proyecto consideró todo tipo de imágenes, desde frutas completamente sanas hasta aquellas con presencia de defectos. En estas últimas se evidencia una variación en la pigmentación a lo largo de su superficie, introduciendo posibles patrones que los modelos

pueden aprender, pero que no son de interés para esta investigación. Por ende, en caso de contar con un conjunto de imágenes más amplio, se recomienda descartar todas aquellas imágenes de frutas que posean defectos, con el fin de, asegurar que los modelos de CNN se enfoquen en el aprendizaje de la forma, mas no, en otras características como el color.

- Dado el objetivo del estudio que es analizar únicamente la forma de las frutas, se recomienda realizar los entrenamientos del enfoque TL solo con imágenes de silueta, de modo que, los modelos no aprendan de características que no son de interés, como el color de las frutas.

4.4 Trabajos futuros

Al haber implementado y analizado varias modelos de CNN que se encuentran en el estado del arte, es posible replicar este proyecto no sólo en otras frutas, sino también en aquellos alimentos que puedan presentar algún tipo de deformidad y pasen por procesos de inspección de calidad, tal es el caso de vegetales como el tomate, rábano, berenjena o frutas como papaya, sandía, banano, entre otros.

Adicionalmente, este proyecto puede ser empleado como componente de un sistema mucho más grande de inspección de calidad, en el que se consideren también otros parámetros como la madurez y/o los defectos externos. Este sistema podría estar regido por un conjunto de reglas que, permitan obtener una clasificación general de calidad de la fruta. De acuerdo con lo mencionado en las recomendaciones, se aconseja que este módulo de clasificación de deformidades sea el último en realizar la inspección, de modo que lleguen a él, aquellas frutas completamente libres de defectos, otorgando así, una predicción mucho más adecuada, ignorando cualquier otro tipo de característica presente en la imagen.

4.5 Contribuciones

- Se generaron datasets reales y sintéticos de manzanas, mangos y fresas, clasificados según su calidad de deformidad. Dado los varios procesos aplicados a los datos, se asegura que, cada imagen es de calidad y está correctamente categorizada (robustez). En la web, no existen repositorios públicos categorizados de esta forma. Al final se obtuvo un total de 20 000 imágenes reales (Sólo 16 000 imágenes en manzanas) y 20 000 imágenes sintéticas por cada fruta.
- Se entrega una versión mejorada del modelo CIDIS para su aplicación en la clasificación de deformidades en frutas. Este modelo es capaz de adaptar sus resultados durante la fase de entrenamiento con el uso de una capa de regularización.
- Se pone a disposición scripts de manipulación y segmentación de imágenes mediante la herramienta SAM. Estos scripts están escritos en lenguaje Python y cuentan con instrucciones detalladas para su entendimiento y uso.

Referencias

- [1] FAO, «FAO,» [En línea]. Available: <https://www.fao.org/3/a1374s/a1374s07.pdf>.
- [2] A. Parra Coronado, «Comercialización de frutas y hortalizas,» de *Ingeniería e Investigación*, 1989, pp. 14-19.
- [3] E. Martínez Castro y A. Rebollo Arévalo, «El sistema de comercialización en origen de las frutas y hortalizas en fresco,» de *Distribución y Consumo*, 2008, pp. 8-25.
- [4] O. Phey Zhen, N. Hashim y B. Maringgal, *Quality evaluation of mango using non-destructive approaches: A review*, 2020.
- [5] C. G. González G., «Frutas y verduras perdidas y desperdiciadas, una oportunidad para mejorar el consumo,» de *Revista chilena de nutrición*, Santiago, 2018, pp. 198-198.
- [6] FAO, «Global food losses and food waste - Extent, causes and prevention,» Roma, 2011.
- [7] A.-M. Bratu, C. Popa, M. Bojan, P. C. Logofatu y M. Petrus, *Non-destructive methods for fruit quality evaluation*, 2021.
- [8] D. S. S. a. D. S. Jasmeen Gill, «A review of automatic fruit classification using soft computing techniques.,» *International Conference on Computer, Systems and Electronics Engineering*, p. 91, 2014.
- [9] Y. Abbaspour-Gilandeh y A. Soltani Nazarloo, *Non-Destructive Detection of Fruit Quality Parameters Using Hyperspectral Imaging, Multiple Regression Analysis and Artificial Intelligence*, 2022.
- [10] OECD, *Apples, International Standards for Fruit and Vegetables*, Paris: OECD Publishing, 2021.
- [11] OECD, *Mangoes, International Standards for Fruit and Vegetables*, Paris: OECD Publishing, 2012.
- [12] OECD, *Strawberries, International Standards for Fruit and Vegetables*, Paris: OECD Publishing, 2021.
- [13] USDA, *United States Standards for Grades of Apples*, 2019.
- [14] USDA, *United States Standards for Grades of Mangos*, 2007.
- [15] USDA, *United States Standards for Grades of Strawberries*, 2006.

- [16] J. Wang, Y. Huo, Y. Wan y H. Zhao, «Grading detection of “Red Fuji” apple in Luochuan based on machine vision and near-infrared spectroscopy,» *PLOS ONE*, vol. 17, nº 8, 2022.
- [17] L. Sun, K. Liang, Y. Song y Y. Wang, «An Improved CNN-Based Apple Appearance Quality Classification Method With Small Samples,» *IEEE Access*, vol. 9, pp. 68054-68065, Enero 2021.
- [18] G. Hu, E. Zhang, J. Zhou y J. Zhao, «Infield Apple Detection and Grading Based on Multi-Feature Fusion,» *Horticulturae*, vol. 7, nº 9, p. 276, Septiembre 2021.
- [19] A. Chakrabarti, T. C. Michaels T., S. Yin y E. Sun, «The cusp of an apple,» *Nature Physics*, vol. 17, pp. 1125-1129, 4 Octubre 2021.
- [20] N. Soria y J. León F., «El cultivo del manzano en la zona alta del Ecuador,» INIAP, Estación Experimental Santa Catalina, Programa de Fruticultura, Quito, 1992.
- [21] R. E. Apolinario Quintana, M. G. Rodríguez Donoso y L. M. Zambrano Mejía, «La Cadena de Valor del mango ecuatoriano y su competitividad internacional,» *Compendium*, vol. 24, nº 47, 2021.
- [22] A. García Villar, «Arándanos El Cierrón,» 2017. [En línea]. Available: <https://www.arandanoselcierron.com/tipos-variedades-fresas#:~:text=Atendiendo%20al%20tamaño%20de%20las,y%20con%20un%20tamaño%20de.> [Último acceso: 2023].
- [23] Ž. Vujović, «Classification Model Evaluation Metrics,» *International Journal of Advanced Computer Science and Applications ((IJACSA)*, vol. 12, nº 6, 2021.
- [24] C. Garillos-Manliguez y J. Y. Chiang, «Multimodal Deep Learning via Late Fusion for Non-Destructive Papaya Fruit Maturity Classification,» *18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pp. 1-6, 2021.
- [25] B. X. Vintimilla, L. E. Chuquimarca y S. A. Velastin, «Banana Ripeness Level Classification Using a Simple CNN Model Trained with Real and Synthetic Datasets,» de *18th International Conference on Computer Vision Theory and Applications VISAPP*, 2023.
- [26] R. Pacheco, P. González, L. E. Chuquimarca, B. X. Vintimilla y S. A. Velastin, «Fruit defect detection using CNN models with real and virtual data.,» de *18th International Conference on Computer Vision Theory and Applications VISAPP*, 2023.
- [27] Y. Li, X. Feng, Y. Liu y X. Han, «Apple quality identification and classification by image processing based on convolutional neural networks,» *Scientific Reports*, vol. 11, nº 16618, 17 Agosto 2021.

- [28] P. Borianne, F. Borne, J. Sarron y E. Faye, «Deep Mangoes: from fruit detection to cultivar identification in colour images of mango trees,» *DISP'19 International Conference on Digital Image and Signal Processing*, Abril 2019.
- [29] S. Kumar Agrawal, «Analytics Vidhya,» 20 Julio 2021. [En línea]. Available: <https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/>.
- [30] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto y H. Adam, «MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,» 2017.
- [31] M. Oltean, «Fruit 360,» Kaggle, 2018. [En línea]. Available: <https://www.kaggle.com/datasets/moltean/fruits>. [Último acceso: 7 Junio 2023].
- [32] M. Mukhiddinov, A. Muminov y J. Cho, «Improved Classification Approach for Fruits and Vegetables Freshness Based on Deep Learning,» *Sensors* 22, nº 21, p. 8192, 26 Octubre 2022.
- [33] N. Sultana, M. Jahan y M. S. Uddin, «Fresh and Rotten Fruits Dataset for Machine-Based Evaluation of Fruit Quality,» Mendeley Data, 8 Abril 2022. [En línea]. Available: <https://data.mendeley.com/datasets/bdd69gyhv8/1>.
- [34] C. G. Pachon Suescun, J. O. Pinzón Arenas y R. Jiménez-Moreno, «Spoiled and fresh fruit inspection dataset,» Mendeley Data, 1 Noviembre 2020. [En línea]. Available: <https://data.mendeley.com/datasets/6ps7gtp2wg/1>.
- [35] S. R. Kalluri, «Fruits fresh and rotten for classification,» Kaggle, 2018. [En línea]. Available: <https://www.kaggle.com/datasets/sriramr/fruits-fresh-and-rotten-for-classification>.
- [36] H. M. Rizwan Iqbal y A. Hakim, «Mango Variety and Grading Dataset,» Mendeley Data, 6 Julio 2021. [En línea]. Available: <https://data.mendeley.com/datasets/5mc3s86982/1>.
- [37] M. H. bin Ismail y M. N. Wagimin, «MangoMassNet-552 Dataset,» Kaggle, 2022. [En línea]. Available: <https://www.kaggle.com/datasets/mypapit/mangomassnet552-dataset?select=images>.
- [38] S. Naik, «Kesar Mango,» Mendeley Data, 3 Febrero 2019. [En línea]. Available: <https://data.mendeley.com/datasets/nsjggt7tyz/1>.
- [39] M. H. bin Ismail, «Sala Mango Images Dataset,» Kaggle, 2022. [En línea]. Available: <https://www.kaggle.com/datasets/mypapit/sala-mango-image>.

- [40] A. Durand-Petiteville, S. Vougioukas, D. Slaughter y D. Sadowski, «Images of individual strawberries (entire fruit, flesh and calyx) with contrasting background,» Dryad, 2017. [En línea]. Available: <https://datadryad.org/stash/dataset/doi:10.25338/B8WC70>.
- [41] O. Mahendra, «Strawberry Dataset (Good and Damaged),» GitHub, 2018. [En línea]. Available: https://github.com/okamahendra123/dataset_strawberry.
- [42] S. AI, «stable-diffusion-ui,» Stability AI, [En línea]. Available: <https://stable-diffusion-ui.github.io/>.
- [43] bmailtais, «Github,» [En línea]. Available: https://github.com/bmailtais/kohya_ss.
- [44] AUTOMATIC1111, «Gituhb,» [En línea]. Available: <https://github.com/AUTOMATIC1111/stable-diffusion-webui>.
- [45] R. Romero, «betterprogramming.pub,» 5 mayo 2023. [En línea]. Available: <https://betterprogramming.pub/train-your-own-stable-diffusion-model-locally-no-code-needed-36f943825d23>.
- [46] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. Berg, W.-Y. Lo, P. Dollar y R. Girshick, «Segment Anything,» *arXiv:2304.02643*, 2023.
- [47] M. X. Bastidas Rodríguez, «reconoSER ID,» 7 Enero 2021. [En línea]. Available: <https://reconoserid.com/redes-siamesas-aplicacion-de-deep-learning-para-la-validacion-de-identidad/>.
- [48] W. Barrios Bustamante, «Medium,» 2 Abril 2021. [En línea]. Available: <https://wbarriosb.medium.com/calculando-la-precisión-en-un-modelo-de-clasificación-multiclase-224d96f52043>.