



UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE GUAYAQUIL
CARRERA DE MECATRÓNICA

IMPLEMENTACIÓN DE UN SISTEMA DE CLASIFICACIÓN DE FRESAS EMPLEANDO VISIÓN ARTIFICIAL PARA EL CONTROL DE CALIDAD EN FUNCIÓN DE SUS ETAPAS DE MADURACIÓN

Trabajo de titulación previo a la obtención del
Título de Ingeniero en Mecatrónica

AUTORES: Diego Sebastián Cortez Martínez
Giuliano Antonio Tutiven Patiño
TUTOR: Alberto Santiago Ramírez Farfán

Guayaquil - Ecuador
2026

19/02/26

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA
UNIVERSIDAD POLITÉCNICA SALESIANA**

Nosotros, **Diego Sebastián Cortez Martínez** con documento de identificación N° **0952810248** y **Giuliano Antonio Tutiven Patiño** con documento de identificación N° **0951024264**, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del **Dispositivo Tecnológico: IMPLEMENTACIÓN DE UN SISTEMA DE CLASIFICACIÓN DE FRESAS EMPLEANDO VISIÓN ARTIFICIAL PARA EL CONTROL DE CALIDAD EN FUNCIÓN DE SUS ETAPAS DE MADURACIÓN**, el cual ha sido desarrollado para optar por el título de: Ingeniero en Mecatrónica, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribimos este documento en el momento que hacemos la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 18 de Febrero del año 2026

Atentamente,



Diego Sebastián Cortez Martínez
0952812048



Giuliano Antonio Tutiven Patiño
0951024264

CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Nosotros, **Diego Sebastián Cortez Martínez** con documento de identificación N° **0952810248** y **Giuliano Antonio Tutiven Patiño** con documento de identificación N° **0951024264**; manifestamos que:

Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo.

Guayaquil, 18 de Febrero del año 2026

Atentamente,



Diego Sebastián Cortez Martínez
0952812048



Giuliano Antonio Tutiven Patiño
0951024264

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, **Alberto Santiago Ramírez Farfán**, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: **IMPLEMENTACIÓN DE UN SISTEMA DE CLASIFICACIÓN DE FRESAS EMPLEANDO VISIÓN ARTIFICIAL PARA EL CONTROL DE CALIDAD EN FUNCIÓN DE SUS ETAPAS DE MADURACIÓN**, realizado por **Diego Sebastián Cortez Martínez** con documento de identificación N° **0952812048** y por **Giuliano Antonio Tutiven Patiño** con documento de identificación N° **0951024264**, obteniendo como resultado final el trabajo de titulación bajo la opción **Dispositivo Tecnológico** que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 18 de Febrero del año 2026

Atentamente,



Ing. Alberto Santiago Ramírez Farfán.
0923348890

DEDICATORIA

Este trabajo de titulación está dedicado a mi padre Franklin Stalin Cortez Martínez y a mi madre Priscila Jazmín Martínez Parrales, quienes me han guiado con amor y un apoyo incondicional en esta etapa de mi formación académica.

Finalmente agradezco a mi familia por siempre haber creído en mi y apoyarme siempre a soñar en grande.

Diego Sebastián Cortez Martínez

Dedico este trabajo de titulación primordialmente a mis padres, quienes han sido el pilar fundamental de mi vida, brindándome su apoyo incondicional y sacrificios para permitirme alcanzar esta meta profesional. A mi abuelita, por su amor infinito, sus oraciones y por ser mi fuente de inspiración diaria. De igual manera extendiendo este sentimiento a toda mi familia por su apoyo constante en este proceso de aprendizaje.

Giuliano Antonio Tutiven Patiño

AGRADECIMIENTO

Agradezco a Dios, a mis padres por sus consejos y haberme sembrado el valor de los estudios y estar siempre presente en mi carrera.

Agradezco a los docentes y compañeros de la carrera de mecatrónica que han formado parte de mi desarrollo como profesional y han sido parte de esta increíble etapa de mi vida.

Agradezco al Ing. Alberto Ramírez por su presencia y guía durante proyecto de titulación.

Diego Sebastián Cortez Martínez

Expreso mi gratitud al Ing. Alberto Ramírez Farfán, por su guía y disposición como tutor de este trabajo de titulación. Sus conocimientos técnicos, su rigor profesional y sus oportunas orientaciones fueron fundamentales para esta investigación. Gracias por la confianza depositada en mi capacidad y por compartir su experiencia académica con generosidad. De manera muy especial, quiero agradecer a mi novia, quien ha sido mi compañera y mi apoyo emocional durante este camino. Gracias por la paciencia, y por el apoyo en este proceso.

Giuliano Antonio Tutiven Patiño

I. RESUMEN

Este estudio presenta un sistema para clasificar fresas utilizando tecnología de visión artificial enfocado en el control de calidad según sus etapas de madurez. La propuesta incluye un módulo para la adquisición de imágenes que consta de una cámara OV5647-62 y un dispositivo Grove Vision AI V2, los cuales están montados sobre una estructura mecánica destinada a analizar las fresas mientras se desplazan en una banda transportadora.

Para generar el conjunto de datos, se tomaron fotografías en un ambiente con iluminación controlada, las cuales fueron etiquetadas y procesadas en la plataforma Edge Impulse a fin de entrenar una red neuronal convolucional. El modelo resultante fue luego implementado en el dispositivo embebido a través de SenseCraft, lo que permite la evaluación en tiempo real del sistema.

Se llevó a cabo una evaluación del rendimiento del modelo usando métricas cuantitativas como precisión, recall y F1-score, alcanzando resultados superiores al 97% en entornos controlados. Además, se realizaron pruebas en diversas condiciones de iluminación, incluidas iluminación controlada, ambiental y la ausencia de luz artificial, para examinar la resistencia del sistema frente a cambios en el entorno.

Los hallazgos indican que el sistema es altamente eficiente al clasificar fresas tanto maduras como inmaduras en condiciones de iluminación apropiadas, logrando una efectividad del 100% durante las pruebas en tiempo real. Sin embargo, se encontraron limitaciones en situaciones sin iluminación controlada, lo que resalta la relevancia del entorno para el funcionamiento efectivo del sistema.

Este proyecto demuestra que es posible implementar un sistema de visión artificial asequible para controlar la calidad de los productos agrícolas, lo cual permite automatizar los procesos de inspección visual y sentar las bases para optimizaciones futuras en su aplicación en contextos productivos reales.

Palabras claves: Visión artificial, clasificación de fresas, control de calidad, redes neuronales convolucionales, sistemas embebidos, procesamiento de imágenes, Edge Impulse.

II. ABSTRACT

This study presents a strawberry classification system based on computer vision technology, focused on quality control according to ripeness stages. The proposed solution includes an image acquisition module consisting of an OV5647-62 camera and a Grove Vision AI V2 device, mounted on a mechanical structure designed to analyze strawberries as they move along a conveyor belt.

To build the dataset, photographs were captured under controlled lighting conditions. The images were labeled and processed using the Edge Impulse platform to train a convolutional neural network. The resulting model was deployed on the embedded device through SenseCraft, enabling real-time system evaluation.

The model's performance was assessed using quantitative metrics such as precision, recall, and F1-score, achieving results above 97 % in controlled environments. Additionally, tests were conducted under various lighting conditions, including controlled lighting, ambient light, and the absence of artificial light, to evaluate the system's robustness to environmental changes.

The findings indicate that the system is highly effective in classifying both ripe and unripe strawberries under appropriate lighting conditions, achieving 100 % effectiveness during real-time testing. However, limitations were observed in scenarios without controlled lighting, highlighting the importance of environmental conditions for optimal system performance.

This project demonstrates the feasibility of implementing an affordable computer vision system for agricultural quality control, enabling the automation of visual inspection processes and laying the groundwork for future optimization in real-world production environments.

Keywords: computer vision, strawberry classification, quality control, convolutional neural networks, embedded systems, image processing, Edge Impulse..

ÍNDICE

I.	Resumen	3
II.	Abstract	4
III.	Introducción	1
IV.	Problema	2
V.	Justificación	3
VI.	Objetivos	5
VI-A.	Objetivo general	5
VI-B.	Objetivos específicos	5
VII.	Marco Teórico	6
VII-A.	Sistemas de clasificación	6
VII-B.	Etapas de maduración de la fresa	6
VII-C.	Calidad de fruta poscosecha	6
VII-D.	Perdidas poscosechas por madurez	7
VII-E.	Aprendizaje automático	7
VII-F.	Redes Neuronales convolucionales.	8
VII-G.	Estadística cuantitativa	9
VII-H.	Comparación entre Hardwares de procesamiento	10
VII-H1.	Grove Vision AI V2	10
VII-H2.	Raspberry Pi	10
VII-H3.	ESP32-CAM	10
VII-H4.	Jetson Nano	11
VII-H5.	Cámara OV5647-62 FOV	11
VII-H6.	Selección de Hardware	11
VII-I.	Selección de software	12
VII-I1.	SenseCraft AI	12
VII-I2.	Edge Impulse	12
VII-I3.	Google Colab	12
VII-J.	Comparación entre modelos de detección.	12
VII-J1.	FOMO (Fast Oriented Motion Object)	12
VII-J2.	YOLO(You Only Look Once)	13
VII-J3.	Comparación entre FOMO y YOLO	13
VII-K.	Visión artificial en procesos de clasificación por detección.	13
VII-L.	Matriz de confusión	13
VII-L1.	Verdaderos Positivos (VP)	13
VII-L2.	Verdaderos Negativos (VN)	13
VII-L3.	Falsos Positivos (FP)	13
VII-L4.	Falsos Negativos (FN)	14
VII-M.	Interpretación de métricas de aprendizaje automático	14

VIII. Marco Metodológico	15
VIII-A. Fase de Análisis	15
VIII-A1. Diagrama de flujo	15
VIII-B. Fase de Diseño	16
VIII-B1. Diseño Mecánico	16
VIII-B2. Diseño y análisis estructural del soporte del sistema de visión artificial	16
VIII-B3. Metodología del análisis por elementos finitos de estructura para el Grove Vision	16
VIII-B4. Diseño y análisis estructural de la base para el transporte de fresas	19
VIII-B5. Metodología del análisis por elementos finitos de base para fresas	20
VIII-B6. Resultados del análisis de tensiones (Von Mises) de base para fresas	20
VIII-B7. Resultados del análisis de desplazamiento de base para fresas	21
VIII-B8. Resultados del análisis de deformación unitaria de base para fresas	22
VIII-B9. Factor de seguridad	23
VIII-C. Software y Entrenamiento del Modelo	25
VIII-C1. Sense Craft como recolector de data.	25
VIII-C2. Edge Impulse entorno de clasificación y entrenamiento.	26
VIII-D. Fase de Proceso Post-Entrenamiento	29
VIII-D1. Conversión del formato archivo TensorFlow Lite a formato Vela	29
VIII-E. Fase de Construcción del Prototipo	30
IX. Resultados	31
IX-A. Resultados del proceso del entrenamiento	31
IX-A1. Matriz de confusión	32
IX-B. Modelo entrenado	33
IX-C. Pruebas en entorno real.	35
IX-C1. Resultados del sistema en tiempo real bajo iluminación controlada.	35
IX-C2. Resultados con variación en la iluminación.	36
IX-C3. Resultados con iluminación ambiental (sin caja y sin luces LED)	36
IX-C4. Resultados sin caja de ambientación y con iluminación LED	37
X. Cronograma	39
XI. Presupuesto	40
XII. Conclusiones	41
XIII. Recomendaciones	42
Anexo A: Diseños en SolidWorks	43
Anexo B: Estructura	44
Anexo C: Programación Google Colab	45
C-A. Código para obtener la matriz de confusión	45
C-B. Código de conversión a formato Vela	47
Anexo D: Adquisición de datos	48
Anexo E: Guía de uso Grove Vision AI V2 mediante Sensecraft y Edge Impulse	48
Referencias	62

ÍNDICE DE FIGURAS

1.	Etapas de maduración de fresas.[22].	6
2.	Perdidas post produccion	7
3.	Redes neuronales convucionales.[23].	8
4.	Grove Vision AI V2.[27].	10
5.	OV5647-62 FOV.[31].	11
6.	Diagrama de Flujo	15
7.	Tensión de Von Mises	17
8.	Desplazamientos resultantes	18
9.	Deformacion unitaria	19
10.	Propiedades del material	20
11.	Von Mises	21
12.	Desplazamientos resultantes	22
13.	Deformación Unitaria	23
14.	Factor de seguridad	24
15.	Dataset del manejo de datos	25
16.	Clasificación de imágenes según la maduración de la fresa.	26
17.	Dataset del manejo de datos	26
18.	Distribución de datos según las clases	27
19.	Configuración para el entrenamiento de la red	28
20.	Estructura armada por dentro.	30
21.	Set de validacion, matriz de confusión y métricas.	31
22.	Matriz de confusión	32
23.	Resultados del modelo entrenado	34
24.	Clasificación en tiempo real de una fresa madura mediante el sistema propuesto.	35
25.	Clasificación en tiempo real de una fresa inmadura mediante el sistema propuesto.	35
26.	Clasificación errónea por falta de iluminación.	36
27.	No detección por falta de iluminación.	37
28.	No detección por falta de iluminación.	37
29.	No detección por falta de iluminación.	38
30.	No detección por falta de iluminación.	38
31.	No detección por falta de iluminación.	39
32.	Cronograma	40
33.	Diseño de bases de fresas	43
34.	Diseño de base de la cámara y el Grove Vision	43
35.	Bases de fresas	44
36.	Perfiles de aluminio con cámara instalada.	44
37.	45
38.	46
39.	46
40.	47
41.	47
42.	47
43.	Prueba de luces LED	48

ÍNDICE DE TABLAS

I. Presupuesto 41

III. INTRODUCCIÓN

La agricultura representa una de las principales actividades económicas para el progreso de las naciones, ya que asegura la provisión de alimentos y crea empleo en diversos segmentos de la sociedad. En este contexto, la fresa se destaca como uno de los frutos más solicitados por sus beneficios nutricionales y su amplia aceptación en el mercado. Sin embargo, los procedimientos para seleccionar y clasificar este fruto se llevan a cabo, en muchas ocasiones, de manera manual, lo que limita la productividad y tiene un impacto en la homogeneidad de la calidad del producto terminado.

La selección de forma manual de fresas depende en gran medida de la evaluación visual del operador, lo que añade un nivel significativo de subjetividad en el proceso decisional. Elementos como el cansancio, la luz y la experiencia del trabajador afectan la exactitud de este procedimiento, generando inconsistencias que pueden llevar a pérdidas financieras y a una reducción de la competitividad de los productores. Este panorama representa un reto importante, sobre todo para pequeños y medianos agricultores en busca de mejorar sus operaciones sin incurrir en altos costos tecnológicos.

La aplicación de tecnologías que se fundamentan en la visión artificial y la inteligencia artificial surge como una solución eficaz para modernizar los sistemas de control de calidad en el sector agrícola, ante esta problemática. Estas tecnologías permiten que la inspección de productos se realice automáticamente, lo cual mejora la precisión en la clasificación y reduce el grado de dependencia de la evaluación humana. De este modo, se favorecen procesos más objetivos, repetitivos y eficazmente ejecutados.

El presente estudio sugiere la creación de un sistema para clasificar fresas utilizando visión artificial con el fin de controlar la calidad en relación con sus etapas de madurez. El sistema propuesto incluye un módulo para la captura de imágenes mediante una cámara OV5647-62 y un dispositivo Grove Vision AI V2, ambos montados en una estructura mecánica diseñada específicamente para analizar las fresas mientras se mueven sobre una cinta transportadora. Las imágenes se capturaron en un ambiente de iluminación controlada y fueron procesadas mediante plataformas de aprendizaje automático, permitiendo la formación de un modelo capaz de clasificar en tiempo real.

Así mismo, en la elaboración del sistema se consideraron factores higiénicos y mecánicos, empleando materiales idóneos para el contacto con alimentos frescos, como el acero inoxidable AISI 304. Este material es frecuente en la industria alimentaria por su facilidad de limpieza y resistencia a la corrosión. Esta decisión obedece no solo a exigencias técnicas, sino también a la necesidad de garantizar condiciones seguras para manejar productos agrícolas.

Los resultados obtenidos indican que el sistema puede clasificar correctamente fresas maduras e inmaduras en diversas condiciones de iluminación, lo que resalta su potencial para ser utilizado como una herramienta auxiliar en los procesos de control de calidad. La automatización de esta labor ayuda a minimizar errores humanos, acelerar la inspección y mejorar la calidad del producto final. Este trabajo constituye una contribución a la mejora del ámbito agrícola a través de la introducción de tecnologías asequibles y de bajo costo, enfocadas en aumentar la producción y la competitividad de los agricultores. Además, favorece la transferencia de saberes de la ingeniería a usos prácticos que afectan de manera directa la calidad de los alimentos y la efectividad de los procesos de producción.

Por último, este proyecto permite la posibilidad de crear sistemas más sofisticados que incorporen mecanismos automáticos para la separación y análisis en tiempo real, apoyando así la modernización del sector agroindustrial y la creación de soluciones tecnológicas con un efecto social y económico positivo.

IV. PROBLEMA

A nivel mundial, la desnutrición se relaciona con la limitada disponibilidad de alimentos frescos y saludables, como las frutas. Garantizar la calidad de productos como las fresas es vital para prevenir enfermedades y promover su exportación. Sin embargo, la producción enfrenta retos debido a la falta de tecnología y automatización en muchos países, lo que provoca pérdidas y dificulta el cumplimiento de las regulaciones internacionales.

A nivel nacional, la exportación de frutos y vegetales enfrenta desafíos en el control de calidad debido a la limitada innovación tecnológica y a la deficiente automatización. Esto genera inconsistencias en la calidad, retrasos, pérdidas y problemas de trazabilidad, dificultando el cumplimiento de estándares internacionales fitosanitarias como Global GAP y el análisis de peligros y puntos críticos (H.A.C.C.P. por sus siglas en inglés) exigidas por la Unión Europea y Estados Unidos [1].

En la evaluación de la calidad de las fresas, uno de los aspectos más importantes es el grado de madurez. Generalmente, las fresas se clasifican en una escala de cinco categorías de maduración. En la categoría 1, la fruta aparece verde, sin haber madurado, sin aroma ni dulzor, y no es apta para el consumo. La categoría 2 indica el inicio del cambio hacia un tono rosado, aunque la fresa mantiene un sabor ácido y no tiene mucha jugosidad. En la categoría 3, se puede ver un color rosado más intenso y un aroma mejorado, aunque el dulzor todavía no alcanza el ideal. La categoría 4 se asocia a la madurez óptima, con un color rojo vivo, un sabor dulce, adecuada firmeza y es ideal para la venta. Finalmente, en la categoría 5, la fresa está en estado de sobremadurez, exhibiendo un color rojo oscuro, una textura muy blanda y una mayor probabilidad de sufrir daños por moho o compresión. Reconocer adecuadamente estas categorías es esencial para decidir el momento ideal de cosecha, reducir pérdidas y garantizar un producto de alta calidad para el consumidor [2]. Las fresas, por su alta sensibilidad a la temperatura, humedad y manipulación, ven comprometida su calidad durante la cadena logística. A 20 °C pueden perder hasta un 5 % de peso diario y reducir su firmeza, afectando su valor comercial [3].

Las exportaciones de frutas, como las fresas en Ecuador, enfrentan desafíos importantes, especialmente en términos de control de calidad, falta de innovación tecnológica y automatización de procesos ineficientes. Las regiones como Chilla ubicada en la provincia del Oro, tienen un bajo rendimiento en comparación con otros países [4]. Anualmente, en el Ecuador se cultiva una área de 108 ha, alcanzando un rendimiento de 16.27 T/ha (cantidad de toneladas de un cultivo que se producen por hectárea), siendo muy baja en relación a otros países productores de fresa, como por ejemplo Estados Unidos con 66,90 T/ha de producción; España (47,6 T/ha); Egipto 46,6 T/ha; Israel con 43,5 T/ha, y a nivel de América del Sur está Colombia con 36,5 T/ha de producción de fresa al año [5].

Las fresas o frutillas también siguen siendo vulnerable a ataques microbianos y lesiones mecánicas, lo que conlleva pérdidas económicas y desperdicios de hasta el 50 % de la cosecha. El reto que enfrentan los productores en todo el mundo, incluido Ecuador, es satisfacer esta demanda de forma sostenible y rentable, siendo imperativo el incorporar tecnologías postcosecha [6].

La identificación de irregularidades a través de imágenes tomadas por cámaras es una de las aplicaciones más frecuentes en el sector de exportación industrial. Este método juega un papel importante en la preservación de la calidad y en la mejora de la eficiencia en varios procesos de producción [7].

No obstante, a pesar de su prometedor uso, las técnicas fundamentadas en DL (Deep Learning) todavía son raramente implementadas en los sistemas automáticos de visión. Esto se debe principalmente a que los datos disponibles tienden a ser insuficientes para formar redes neuronales profundas, y la generación de estos conjuntos de datos resulta costosa debido al gran esfuerzo manual requerido para etiquetar la información. [8].

La habilidad de las personas para descifrar imágenes es bastante alta, lo que les permite adaptarse con facilidad a nuevos contextos. No obstante, las tareas monótonas y repetitivas producen cansancio y hacen que la eficacia y confiabilidad de la inspección efectuada por los operarios se deterioren con rapidez. [9]. Para la evaluación manual de la producción frutal es necesario invertir una buena cantidad de recursos humanos, lo que supone costos laborales y puede ser una labor que consume mucha mano de obra, por lo que eleva los gastos operativos y puede resultar ineficiente en términos de tiempo. Los productores deben gastar considerablemente para contratar y capacitar personal extra. [10].

V. JUSTIFICACIÓN

Los sistemas basados en visión artificial tienen bastantes beneficios en el área de producción a grandes escalas, en los cuales se requiere un control más minucioso y más detallado cuando se trata de alimentos. Existe evidencia de que la visión artificial puede ayudar a la producción de alimentos. Por ejemplo, en un estudio aplicado al café se propone el uso de un mecanismo que emplea una cinta transportadora para mover los granos mientras que un sistema de captura de imágenes obtiene fotografías en tiempo real del café. Estas fotografías son evaluadas mediante algoritmos de análisis de imágenes que obtienen atributos de color y características específicas del grano. Este procedimiento facilita la evaluación de la calidad y condición del café de manera veloz y exacta, mejorando la clasificación y disminuyendo las pérdidas, lo que contribuye a optimizar el control de calidad en productos agrícolas a través de visión por computadora [11].

La industria frutícola es un pilar esencial de la economía global y un componente fundamental del suministro mundial de alimentos. La creciente demanda de productos frescos de alta calidad ha impulsado la necesidad de implementar sistemas eficientes de inspección y control que garanticen la inocuidad y la ausencia de contaminantes en las frutas destinadas al consumo. La clasificación precisa y la detección temprana de defectos y contaminantes no solo garantizan la satisfacción del consumidor, sino que también se ajustan a las normativas internacionales de seguridad alimentaria [12]. El énfasis de la tecnología de clasificación ha estado en la percepción no destructiva, la extracción y la evaluación de la calidad de las frutas y hortalizas. En los últimos años, se han puesto en producción muchas líneas comerciales de clasificación de frutas y verduras que utilizan estas tecnologías de detección. Estas líneas de producción son eficientes, pero también requieren costos iniciales y de mantenimiento más altos. Tomando como ejemplo el actual proceso de producción de algunos huertos, después de la cosecha a gran escala en el huerto, las frutas se transportan a la instalación para su ulterior procesamiento comercial, como el almacenamiento, la clasificación y el envasado [13]. Los sistemas de visión artificial y el procesamiento de imágenes son un área de investigación en constante crecimiento, una técnica de análisis importante para la pre y postcosecha de cultivos [14]. La visión artificial ofrece una alternativa como técnica automatizada, no destructiva y rentable para cumplir con estos requisitos [15]. La precisión del rendimiento se garantizó mediante el diseño óptimo de los componentes mecánicos, realizado mediante síntesis cinemática y análisis dimensional [16].

La madurez es un indicador importante de la calidad de las frutas, relacionado con su calidad de consumo, cosecha, almacenamiento y comercialización. Las fresas completamente maduras tienen la vida útil más corta, las fresas inmaduras pueden madurar y sobre madurar rápidamente. Son vulnerables a daños físicos, especialmente las maduras. Explorar los grados de madurez adecuados para la cosecha es importante para la gestión del crecimiento, el almacenamiento y la comercialización [17].

La disminución de la fuerza laboral, causada por el envejecimiento de la población, las restricciones a la migración o las cambiantes aspiraciones de los trabajadores agrícolas, es motivo de gran preocupación. La automatización y la robotización del sector se consideran una solución viable a este problema, pero requieren la resolución de varios desafíos técnicos para tener éxito [18].

El rápido crecimiento de las fresas y la laboriosa labor de la cosecha manual subrayan la importancia de la automatización para mantener altos rendimientos y prevenir el rápido deterioro de las plantas de fresa [19].

Sin embargo, el área está en un continuo desarrollo y genera nuevos modelos que ofrecen funciones y un rendimiento superior. Esta evolución constante en la tecnología ha dado lugar a herramientas complejas que lideran el sector agrícola. Entre ellas, el algoritmo de YOLO (you only look once) se destaca como un progreso significativo en la automatización del campo. Yolo, que originalmente fue creado para reconocer objetos en tiempo real, ha sido extensamente modificado para varios fines agrícolas, incluyendo aspectos específicos relacionados con la identificación de fresas mediante la segmentación de eventos [20].

Por lo tanto, el presente proyecto justifica el proponer la implementación de un sistema de control de calidad haciendo uso de visión artificial.

VI. OBJETIVOS

VI-A. *Objetivo general*

Implementar un sistema de clasificación de fresas empleando visión artificial para el control de calidad mediante las etapas de maduración.

VI-B. *Objetivos específicos*

- Crear un conjunto de datos mediante la captura de imágenes en un prototipo con banda transportadora para el entrenamiento del clasificador de las etapas de maduración de fresas.
- Entrenar una red neuronal convolucional a partir de datos tomados en tiempo real para el sistema de control de calidad.
- Validar el funcionamiento del proyecto evaluando la efectividad en la clasificación mediante estadística cuantitativa.

VII. MARCO TEÓRICO

VII-A. *Sistemas de clasificación*

En el centro de empaque, casi todos los artículos son organizados y agrupados según su tamaño para cumplir con los requisitos de calidad y dimensiones del mercado al que se dirigen. La selección para descartar aquellos productos que no alcanzan la calidad estándar y aquellos de distintas calidades en numerosos casos, se lleva a cabo de manera manual. La división en función del tamaño, peso, longitud o diámetro se realiza, en muchas ocasiones, a través de un proceso automatizado, para el cual existe una amplia gama de equipos, mayormente diseñados para cada tipo de cultivo. En el sector del tomate, se aplica un sistema óptico automático para clasificar según colores y la división por tamaño se efectúa antes o después de esta clasificación usando maquinaria. Varias instalaciones de empaque también implementan un proceso de clasificación inicial en el área de suministro a la línea de empaque, para eliminar los productos realmente defectuosos y los desperdicios. Esta primera clasificación normalmente se realiza manualmente y a menudo solo requieren de uno o dos operarios. [21]

VII-B. *Etapas de maduración de la fresa*

Según estudios de horticultura, la maduración de la fresa se clasifica generalmente en tres etapas: Inmadura: color verde o blanco; baja concentración de azúcares.[21] Intermedia o semi-madura: presencia de zonas rojas y verdes; textura firme. Madura: color rojo uniforme; textura blanda y sabor dulce. La determinación automática de estas etapas permite mejorar la gestión poscosecha, optimizar la selección para exportación y reducir pérdidas por manipulación inadecuada.



Figura 1: Etapas de maduración de fresas.[22].

VII-C. *Calidad de fruta poscosecha*

La calidad global de los alimentos debe estar presente en todas las etapas de la comercialización y en los participantes del sector lo que hace imprescindible hacer uso de un sistema que asegure la calidad del producto. En el entorno de la comercialización de frutas frescas estas se pueden ver afectadas en cualquier etapa del proceso. El punto más crítico es en el momento que pasa del productor al mayorista ya que el producto puede llegar en condiciones diferentes a las observadas en los campos de cultivos, estado dado por diferentes factores.

En las frutas incluso si no se daña, cambia el estado de madurez al llegar al distribuidor el cual necesita saber de su nivel de maduración y la evolución que esta desarrollara de esta manera decidir en que lote debería ser enviada o incluso descartarla. Es el por esto la importancia del control de calidad en el mercado ya que de esta manera se asegura que la fruta contenga la suficiente calidad para el consumo.

VII-D. Perdidas poscosechas por madurez

Es complicado calcular la pérdida de frutas en general debido a los diversos factores que afectan el proceso. De acuerdo con investigaciones llevadas a cabo por el departamento de agricultura estadounidense, las principales causas de las pérdidas poscosechas son la maduración de los productos y las enfermedades. El proceso de maduración de los productos genera pudriciones debido a la invasión de microorganismos patógenos que dañan su calidad; asimismo, durante este período se acelera la tasa respiratoria, lo cual provoca que el producto pierda agua y se marchite.



Figura 2: Perdidas post produccion

VII-E. Aprendizaje automático

El aprendizaje automático es un segmento de la inteligencia artificial que capacita a los sistemas informáticos para reconocer patrones y hacer elecciones sin requerir programación directa. Su funcionamiento se fundamenta en algoritmos que examinan grandes cantidades de información para extraer datos importantes y optimizar su eficacia con el tiempo. Hay diversos métodos en el aprendizaje automático, entre los cuales se encuentra el aprendizaje supervisado, donde el modelo se entrena utilizando datos etiquetados, y el aprendizaje no supervisado, en el que el sistema descubre patrones de forma autónoma, sin ayuda externa. En el sector industrial, el aprendizaje automático ha transformado áreas como la manufactura, la logística y la agricultura, posibilitando la optimización de procesos y la automatización de tareas complejas. En la labor de clasificar bananos, la adopción de esta tecnología incrementa la eficiencia al sustituir los métodos tradicionales de inspección visual, que son subjetivos y demandan mucha intervención humana. A través de modelos de redes neuronales, se pueden analizar atributos visuales como color, tamaño y textura, logrando una clasificación más exacta y fiable.

Además, el aprendizaje automático facilita la identificación temprana de enfermedades y defectos, lo que ayuda a disminuir pérdidas financieras y a elevar la calidad del producto. La adopción de estas tecnologías en la industria del banano representa un progreso relevante en la modernización del sector, permitiendo no solo mejorar la calidad del producto, sino también reducir los costos operativos y aumentar la competitividad en el mercado internacional.

VII-F. *Redes Neuronales convolucionales.*

Las redes neuronales convolucionales (CNN, por sus siglas en inglés) constituyen una arquitectura de aprendizaje profundo especialmente diseñada para el análisis y procesamiento de imágenes. Estas redes se inspiran en el funcionamiento del córtex visual humano, lo que les permite reconocer patrones y características visuales con un alto grado de precisión. Su estructura está compuesta por diferentes tipos de capas —convolucionales, de agrupamiento (*pooling*) y totalmente conectadas— que actúan de manera conjunta para extraer y procesar atributos relevantes como bordes, texturas, formas y colores.

En el ámbito industrial, las CNN han probado su eficacia en tareas como la detección de objetos, la identificación de defectos y la clasificación de productos agrícolas. Su aplicación en la clasificación de frutillas permite determinar con exactitud su grado de madurez y detectar irregularidades superficiales, incrementando la eficiencia del proceso y reduciendo la subjetividad asociada a la inspección visual manual. Gracias a su capacidad para adaptarse a diferentes condiciones de iluminación, posición y tamaño, resultan especialmente adecuadas para entornos productivos dinámicos.

Así mismo, la incorporación de las CNN en sistemas de visión artificial facilita el procesamiento de imágenes en tiempo real, posibilitando el desarrollo de soluciones escalables y sostenibles. La combinación de estas redes con tecnologías como el Internet de las Cosas (IoT) y la computación en la nube ofrece nuevas oportunidades para mejorar la trazabilidad y el control de calidad en la cadena productiva agrícola. En el caso de la industria de las frutillas, su implementación no solo optimiza la clasificación del producto, sino que también contribuye a disminuir costos operativos y fortalecer la competitividad en el mercado global.

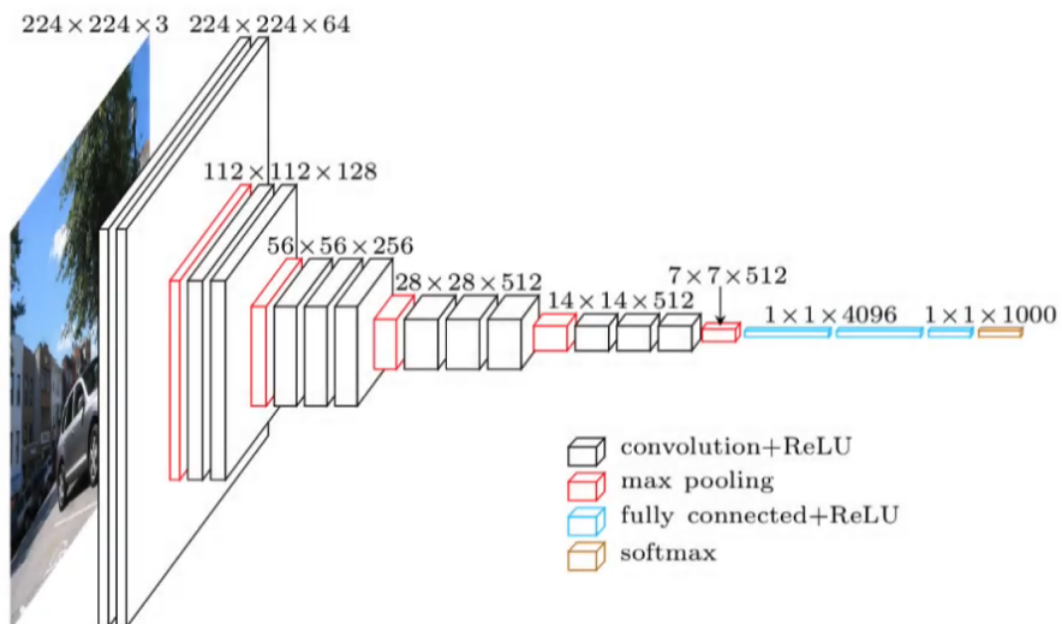


Figura 3: Redes neuronales convolucionales.[23].

VII-G. *Estadística cuantitativa*

Es un campo de la estadística que se centra en el examen de datos numéricos con el propósito de describir, inferir y modelar fenómenos que pueden ser observados. En el ámbito del aprendizaje automático y la visión por computadora, la estadística cuantitativa facilita la evaluación objetiva del rendimiento de los modelos a través de métricas, distribuciones de probabilidad y pruebas de hipótesis. [24]

Este método se fundamenta en la recolección ordenada de datos cuantificables, su estructuración y análisis con herramientas matemáticas, lo que permite detectar patrones, tendencias y conexiones entre diferentes variables. En tareas de clasificación, la estadística cuantitativa es esencial para entender los resultados de experimentos, comparar diferentes modelos y verificar su habilidad para aplicar su conocimiento a información que no se ha analizado antes. [25]

VII-H. Comparación entre Hardwares de procesamiento

VII-H1. Grove Vision AI V2: Grove - Módulo de IA de Visión V2: El módulo de Visión AI V2 de Grove es una alternativa sofisticada en el ámbito de la visión artificial, diseñado para llevar a cabo el procesamiento de inteligencia artificial de manera local, sin necesidad de contar con servidores externos. Este dispositivo incorpora el microcontrolador Himax WiseEye2 HX6538, que tiene un núcleo dual Arm Cortex-M55 junto con una red neuronal Arm Ethos-U55 ya integrada. Gracias a esta configuración, se logra la ejecución de modelos de aprendizaje automático de forma local, lo que asegura rapidez y autonomía en el manejo de datos visuales.

Lo que hace único al Grove Vision AI V2 en comparación con otros sistemas de visión artificial es su habilidad para combinar un potente procesamiento con sensores de imagen en un solo equipo, logrando alta eficiencia que posibilita la implementación en tiempo real de algoritmos de IA. Además, la plataforma SenseCraft AI hace que el despliegue de modelos sea más accesible, disminuyendo la necesidad de conocimientos técnicos profundos en programación, lo que permite que desarrolladores de diversos niveles puedan utilizarla con facilidad.

En conclusión, el Grove Vision AI V2 representa una opción versátil y sólida para desarrollar proyectos relacionados con la visión por computadora, brindando una excelente combinación de usabilidad, eficiencia y capacidad de procesamiento. Esta integración de características lo transforma en una alternativa atractiva tanto para principiantes en el campo de la inteligencia artificial como para usuarios experimentados en busca de soluciones prácticas y efectivas para sus iniciativas. [26]

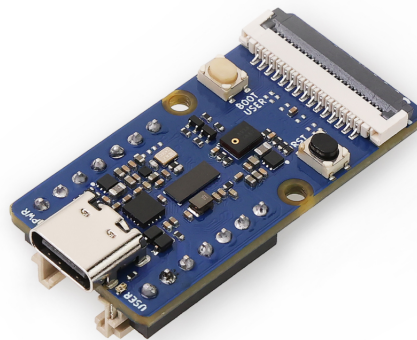


Figura 4: Grove Vision AI V2.[27].

VII-H2. Raspberry Pi: La Raspberry Pi es un ordenador compacto (PCB) creado para operar con sistemas operativos que se basan en Linux. Cuenta con un procesador basado en la arquitectura ARM y la habilidad para ejecutar software de análisis de imágenes y visión artificial mediante bibliotecas como OpenCV y plataformas de aprendizaje automático como TensorFlow. Su diseño permite la conexión de dispositivos externos, incluyendo cámaras digitales, lo que la hace una opción popular en iniciativas educativas, estudios de investigación y desarrollos de visión artificial.[28]

VII-H3. ESP32-CAM: El ESP32-CAM es un controlador que combina conectividad inalámbrica WiFi, un procesador eficiente en el uso de energía y una cámara, diseñado para aplicaciones embebidas económicas. Este

aparato posibilita tomar fotos y realizar funciones simples de procesamiento, siendo utilizado en sistemas de monitoreo, identificación de objetos y en iniciativas relacionadas con el Internet de las Cosas(Iot).[29]

VII-H4. Jetson Nano: El Jetson Nano es una solución de computación integrada desarrollada por NVIDIA, enfocada en aplicaciones de visión por computadora e inteligencia artificial. Comprende un chip ARM y una unidad gráfica exclusiva, lo que posibilita el diseño de modelos de redes neuronales y procedimientos sofisticados para examinar imágenes en tiempo real. Se emplea frecuentemente en proyectos relacionados con la robótica, los vehículos autónomos y los sistemas de evaluación visual.[30]

VII-H5. Cámara OV5647-62 FOV: La OV5647-62 FOV es un módulo de cámara que combina el sensor de imagen OV5647 con una lente tipo ojo de pez, permitiendo la captura de imágenes con un rango visual de 62 grados. Este ángulo es más extenso que el que ofrecen cámaras tradicionales, lo que posibilita que el módulo cubra una área más amplia de la escena. Esto es beneficioso para usos que necesitan un campo de visión más amplio.



Figura 5: OV5647-62 FOV.[31].

Entre las características destacadas de este módulo de cámara se encuentra el sensor OV5647, conocido por su excelente calidad de imagen y su capacidad para tomar fotos detalladas con una resolución de 2592 x 1944 píxeles. Este sensor se asocia con una lente ojo de pez, que proporciona al módulo su amplio rango visual, siendo adecuado para diversas aplicaciones, como la vigilancia, la robótica y la visión artificial. Además, la OV5647-62 FOV es compatible con plataformas como Raspberry Pi 3B+ y 4B, facilitando su uso en diferentes proyectos. Debido a su amplio rango de visión y su habilidad para capturar imágenes de alta resolución, esta cámara resulta adecuada para una variedad de aplicaciones.[26]

VII-H6. Selección de Hardware: Para llevar a cabo este proyecto, se eligió el dispositivo Grove Vision AI V2 junto a la cámara OV5647-62, ya que se centra en aplicaciones de visión artificial embebida y es compatible directamente con plataformas como Edge Impulse y SenseCraft. En contraste con Raspberry Pi y Jetson Nano, el Grove Vision AI V2 consume menos energía, tiene un costo más bajo y presenta una arquitectura más sencilla al no necesitar un sistema operativo completo. Además, en comparación con la ESP32-CAM, proporciona una mayor capacidad para ejecutar modelos de redes neuronales convolucionales con una precisión superior.

La combinación del Grove Vision AI V2 y la cámara OV5647-62 crea un sistema especializado y compacto para

la clasificación de imágenes en tiempo real, lo cual facilita la creación del prototipo y disminuye la complejidad del desarrollo. Por estas razones, esta plataforma fue elegida como la opción más adecuada para lograr los objetivos del proyecto.

VII-I. Selección de software

VII-I1. SenseCraft AI: SenseCraft AI es una solución integral confeccionada para ayudar a programadores y creadores a desarrollar e implementar proyectos de inteligencia artificial sin complicaciones. El portal proporciona diversas herramientas y opciones que facilitan el desarrollo de IA, haciéndolo fácil para personas con distintos grados de conocimiento. En este documento, analizaremos las secciones principales del portal de SenseCraft AI, ofreciendo un resumen de sus características y funciones esenciales.[32]

VII-I2. Edge Impulse: Edge Impulse es una plataforma especializada en el desarrollo, entrenamiento y despliegue de modelos de aprendizaje automático para sistemas embebidos y aplicaciones de computación en el borde (edge computing). Proporciona un entorno integrado que permite la gestión de datos, el etiquetado de imágenes, el entrenamiento de modelos de visión artificial y la optimización de estos para su ejecución en dispositivos con recursos limitados. Asimismo, Edge Impulse facilita la validación del desempeño de los modelos y su exportación a distintos formatos compatibles con hardware embebido, lo que la convierte en una herramienta adecuada para el desarrollo de sistemas de visión artificial en tiempo real.[33]

VII-I3. Google Colab: Google Colab, que es la versión corta de Google Colaboratory, es un servicio en la nube que posibilita la creación y la ejecución de cuadernos Jupyter desde el navegador. Está diseñado para programar y correr código en Python, con un enfoque particular en campos como la ciencia de datos, el machine learning y la investigación educativa. Google Colab permite la colaboración en tiempo real, proporciona acceso a recursos de computación como GPU y TPU, y se integra con Google Drive, haciendo de él una herramienta útil y disponible para el desarrollo experimental y la reproducibilidad en contextos académicos. [34]

VII-J. Comparación entre modelos de detección.

VII-J1. FOMO (Fast Oriented Motion Object): En el ámbito del análisis visual, FOMO (Fast Oriented Motion Object) es un método de identificación de objetos creado específicamente para reconocer elementos pequeños y que se mueven rápidamente en secuencias de imágenes o vídeos. A diferencia de los enfoques convencionales que utilizan cajas delimitadoras, FOMO se basa en una representación mediante mapas de calor, donde cada píxel indica la probabilidad de que pertenezca a un objeto particular. [35]

Este método se adapta especialmente a situaciones en las que los objetos ocupan una pequeña porción de la imagen y tienen movimientos considerablemente grandes entre frames sucesivos, momentos en que los detectores tradicionales suelen perder eficacia debido a la reducción de información espacial que ocasionan las técnicas de muestreo descendente.[36] FOMO mantiene una resolución espacial más alta en las últimas capas de la red neuronal, permitiendo de esta manera identificar patrones locales eficazmente. [37]

Así mismo, los modelos de FOMO están diseñados para funcionar en entornos de Edge AI, donde la capacidad de computación es restringida. Su estructura liviana facilita su implementación en microcontroladores y aparatos embebidos, alcanzando un balance entre el tiempo de respuesta, la exactitud y el consumo energético. Por estas razones, FOMO se ha establecido como una opción eficiente para aplicaciones como la detección de drones, el monitoreo de fauna, el análisis deportivo y los sistemas de vigilancia en tiempo real.[38]

VII-J2. YOLO(You Only Look Once): YOLO, que significa You Only Look Once, es un sistema de detección de objetos que funciona en tiempo real. Este modelo aborda la tarea como un problema de regresión única, previendo simultáneamente las probabilidades de clase y las coordenadas de las cajas directamente a partir de la imagen introducida. Gracias a este método, se pueden analizar imágenes completas en una sola pasada a través de la red neuronal, logrando un buen balance entre rapidez y precisión. [39] Las versiones más recientes de YOLO han introducido mejoras en su arquitectura que aumentan tanto la precisión como la eficiencia en computación, lo que permite su uso en sistemas integrados y en aplicaciones de visión artificial en tiempo real.[40]

VII-J3. Comparación entre FOMO y YOLO: YOLO (You Only Look Once) y FOMO (Fast Oriented Motion Object) son metodologías empleadas en la detección de objetos dentro del ámbito de la visión por computadora, diferenciándose principalmente en su estructura, su enfoque de detección y su uso.[39] YOLO se caracteriza por ser un detector de objetos versátil que utiliza cajas delimitadoras, segmentando la imagen en una rejilla y prediciendo en paralelo tanto las posiciones de las cajas como probabilidades en diferentes clases. Detectando así en tiempo real con un alto grado de precisión, siendo ampliamente utilizado en aplicaciones de visión por computadora que requieren identificar objetos de tamaño medio o grande en entornos con suficiente potencia computacional.[39] [40]

Por otra parte, FOMO utiliza un método de detección densa fundamentado en mapas térmicos, donde cada pixel o área local indica la probabilidad de la presencia de un objeto. Esta metodología es especialmente efectiva para identificar objetos pequeños y de rápido movimiento, que a menudo son eliminados en modelos que utilizan cajas delimitadoras debido a la disminución de la resolución espacial. [35] [36]

Con respecto a su implementación, YOLO está diseñado para funcionar en GPUs o dispositivos embebidos con un mayor poder de procesamiento, mientras que FOMO ha sido ajustado para Edge AI, facilitando su instalación en microcontroladores con capacidades limitadas. Por lo tanto, la elección entre YOLO y FOMO depende del tipo de objeto que se desea detectar, sus dimensiones y las limitaciones de hardware del sistema. [40] [35]

VII-K. Visión artificial en procesos de clasificación por detección.

Diversos estudios han demostrado la eficacia de la visión artificial en la clasificación de frutas. Según Cubero et al. (2011), las características de color son las más utilizadas para determinar el grado de madurez, mientras que la forma y la textura permiten detectar deformaciones o daños. En el caso de las fresas, el color es un indicador clave de maduración. Las fresas inmaduras presentan tonalidades verdes o blanquecinas, mientras que las maduras tienden a tonalidades rojas intensas.

VII-L. Matriz de confusión

Esta es un recurso estadístico para analizar el rendimiento de un modelo de clasificación al comparar las etiquetas generadas con las etiquetas verdaderas. Este recurso organiza los resultados en cuatro elementos clave: verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos. [41]

VII-L1. Verdaderos Positivos (VP): situaciones en las que el modelo tiene éxito al prever la existencia de un objeto o clase de interés.

VII-L2. Verdaderos Negativos (VN): situaciones en las que el modelo acierta al predecir que no está presente la clase o el objeto de interés.

VII-L3. Falsos Positivos (FP): situaciones en las que el modelo anticipa la existencia de un objeto o clase que no está presente realmente.

VII-LA. Falsos Negativos (FN): situaciones donde el modelo no identifica un objeto o clase que sí existe.

Estas definiciones constituyen la base para el cálculo de las métricas de evaluación y el análisis de la matriz de confusión

El estudio de la matriz de confusión ayuda a descubrir errores sistemáticos en el modelo, tales como inclinaciones hacia una categoría particular o dificultades para identificar categorías menos representadas. De esta matriz se obtienen métricas esenciales como la precisión, la sensibilidad y la especificidad, las cuales ofrecen una comprensión más profunda del desempeño del clasificador que la simple tasa de aciertos general. [24] [42]

En el campo de la visión por computadora, especialmente en situaciones donde hay clases desiguales o elementos pequeños, la matriz de confusión se vuelve crucial para evaluar el verdadero efecto de los errores de clasificación en el sistema final. [41]

VII-M. Interpretación de métricas de aprendizaje automático

Son indicadores cuantitativos que ayudan a verificar la efectividad de un modelo predictivo. Entre las métricas más comunes en tareas de clasificación se encuentran la exactitud, la precisión, la sensibilidad, la puntuación F1 y el área bajo la curva ROC. [43]

La correcta interpretación de estas medidas depende del contexto en el que se apliquen. Por ejemplo, en situaciones donde los falsos negativos tienen consecuencias significativas, la sensibilidad se convierte en un aspecto más crítico que la exactitud. La métrica F1 es útil para equilibrar la precisión y la sensibilidad, lo que la hace especialmente valiosa en conjuntos de datos desequilibrados. [42]

Usar múltiples métricas de manera simultánea en el ámbito del aprendizaje profundo y la visión por computadora proporciona una evaluación más robusta del modelo, previniendo conclusiones erróneas que se apoyen únicamente en un único cálculo de su desempeño.[41]

VIII. MARCO METODOLÓGICO

El presente proyecto se desarrolló siguiendo un enfoque metodológico basado en la investigación aplicada y experimental. Su propósito principal fue diseñar, desarrollar e implementar un sistema de clasificación de frutillas mediante visión artificial, optimizando el proceso de identificación del estado de madurez a través de la automatización y el aprendizaje profundo.

El método se dividió en varias fases secuenciales e interdependientes: análisis, diseño, entrenamiento del modelo, proceso post-entrenamiento, y construcción del prototipo. Cada una de estas etapas permitió estructurar el desarrollo técnico y experimental del sistema.

VIII-A. Fase de Análisis

En esta primera etapa se llevó a cabo el estudio y comprensión del problema. Se analizaron las condiciones físicas de las frutillas, considerando sus variaciones de color según el nivel de maduración. La información recolectada permitió establecer los parámetros de clasificación y los requerimientos del sistema de visión artificial.

El análisis contempló la selección de herramientas y dispositivos tecnológicos que respondieran a las necesidades del proyecto. Se identificaron como componentes principales el módulo *Grove Vision AI Module v2* para el procesamiento de imágenes.

VIII-A1. Diagrama de flujo: En la "Figura 6" se presenta el diagrama de flujo correspondiente al funcionamiento general del sistema de clasificación de fresas mediante visión artificial. El proceso inicia con la activación de la banda transportadora y la cámara, permitiendo la adquisición continua de imágenes de los frutos que circulan por el sistema. Posteriormente, se ejecuta la etapa de preprocesamiento de imagen y clasificación mediante el modelo de visión artificial, donde cada imagen capturada es analizada para determinar la presencia de una fresa y su correspondiente etapa de maduración. En caso de no detectarse ningún fruto, el sistema retorna al ciclo de adquisición. Cuando se identifica una fresa, el sistema procede a su clasificación de acuerdo con los criterios establecidos. Si el fruto cumple con los requisitos de calidad, es considerado apto para el consumo; caso contrario, se clasifica como falla de maduración o falla por pudrición. Finalmente, el sistema retorna al ciclo de adquisición de imágenes, permitiendo una operación continua y automatizada del procedimiento de control de calidad.

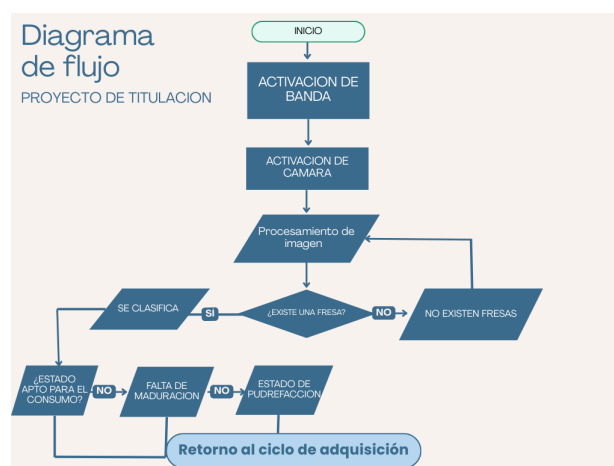


Figura 6: Diagrama de Flujo

VIII-B. Fase de Diseño

VIII-B1. Diseño Mecánico: Se utilizó una estructura metálica y un caja ya preexistente proporcionada por el club de robótica, la cual sirvió como soporte para la cámara en la banda transportadora y para mantener el ambiente al momento de la adquisición de datos, esto para no tener problemas de variaciones en las imágenes capturadas, así que se mantuvo un solo ambiente para el entrenamiento.

Durante el diseño mecánico se realizaron mejoras a la estructura física que sostiene el sistema de visión. Se emplearon perfiles de aluminio y uniones de soporte. También se mejoró la caja ya que esta presentaba orificios por donde fácilmente podría entrar la luz.

VIII-B2. Diseño y análisis estructural del soporte del sistema de visión artificial: Para el diseño del soporte del módulo Grove Vision AI V2 y la cámara OV5647-62 se seleccionó como material de fabricación el polímero ABS (Acrilonitrilo Butadieno Estireno), debido a su amplia utilización en procesos de impresión 3D, su adecuada resistencia mecánica y su bajo costo. Este material resulta apropiado para aplicaciones de prototipado funcional, ya que ofrece una buena relación entre rigidez, resistencia y facilidad de manufactura.

El soporte fue fabricado mediante impresión 3D con un porcentaje de relleno (infill) del 100 %, con el objetivo de garantizar una estructura sólida y homogénea, reduciendo la presencia de vacíos internos y aumentando su capacidad para soportar cargas mecánicas. Esta configuración permite asegurar una mayor rigidez estructural y minimizar deformaciones durante la operación del sistema.

Las propiedades mecánicas del material ABS consideradas en la simulación incluyen un módulo de elasticidad aproximado de 2 GPa, una densidad de 1020 kg/m³ y un límite de tracción cercano a 30 MPa, valores que representan adecuadamente el comportamiento del material utilizado en la fabricación del soporte.

VIII-B3. Metodología del análisis por elementos finitos de estructura para el Grove Vision: El análisis estructural fue realizado mediante un estudio estático utilizando el software SolidWorks Simulation. La carga aplicada corresponde al peso del conjunto conformado por el módulo Grove Vision AI V2 y la cámara OV5647-62, los cuales presentan una masa reducida debido a su naturaleza como dispositivos electrónicos compactos. Esta carga fue aplicada sobre las superficies de apoyo del soporte, mientras que las zonas de fijación fueron definidas como restricciones mecánicas (condiciones de frontera).

El objetivo del análisis fue determinar la distribución de tensiones, desplazamientos y deformaciones unitarias generadas por la carga, con el fin de verificar la integridad estructural del diseño y asegurar que el soporte pueda operar sin presentar fallas mecánicas durante el funcionamiento del sistema de clasificación de fresas.

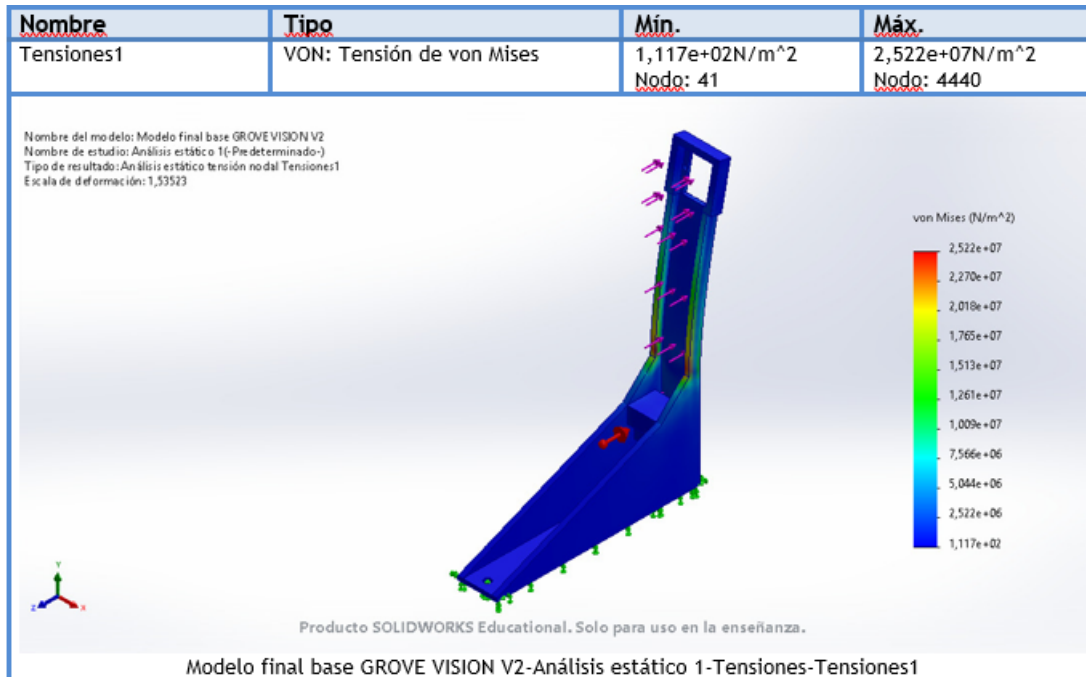


Figura 7: Tensión de Von Mises

Los resultados de la tensión máxima de Von Mises se concentran, sobre todo, en áreas adyacentes a los puntos de fijación y en regiones donde se producen alteraciones geométricas en el diseño. Sin embargo, los valores obtenidos se mantienen significativamente por debajo del límite elástico del material ABS, lo que señala que no existe peligro de deformación permanente o falla estructural en el soporte.

La distribución de tensiones es coherente con el comportamiento mecánico esperado de una estructura sometida a carga estática, concentrándose los esfuerzos en las áreas de mayor restricción mecánica.

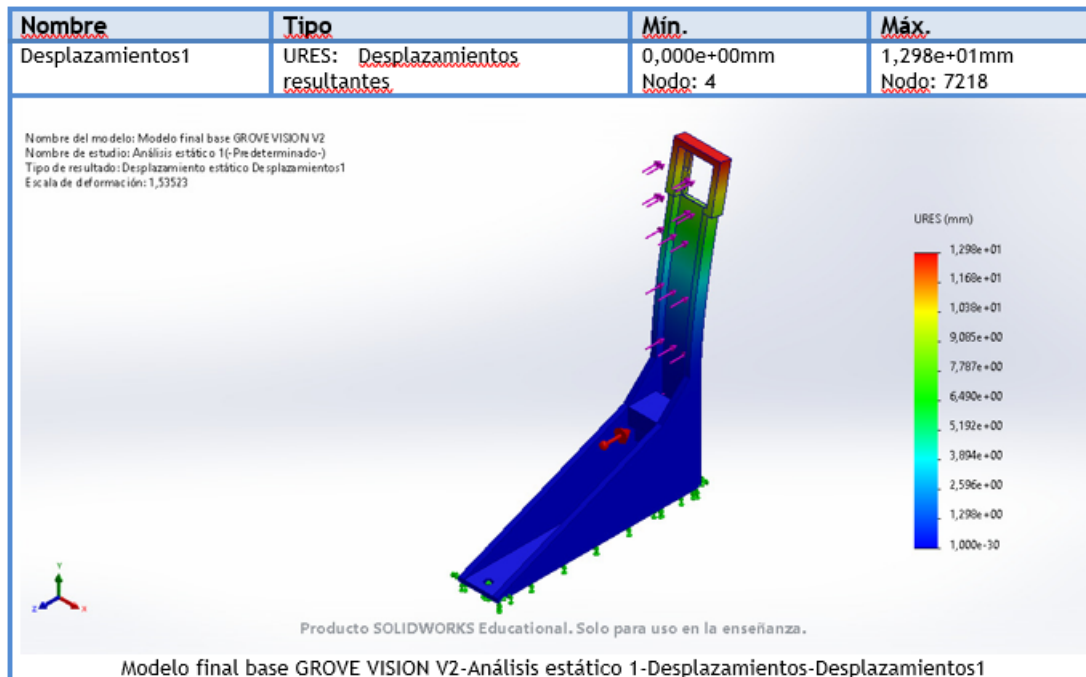


Figura 8: Desplazamientos resultantes

El análisis de desplazamiento evidencia que las mayores deformaciones se presentan en las zonas más alejadas de los puntos de sujeción, mientras que en las regiones restringidas el desplazamiento es prácticamente nulo. La magnitud del desplazamiento máximo es reducida y no compromete la estabilidad ni la alineación de la cámara respecto a la banda transportadora.

Este comportamiento garantiza que la posición del sistema de visión artificial se mantenga estable durante la operación, permitiendo una correcta captura de imágenes sin afectar la precisión del proceso de clasificación.

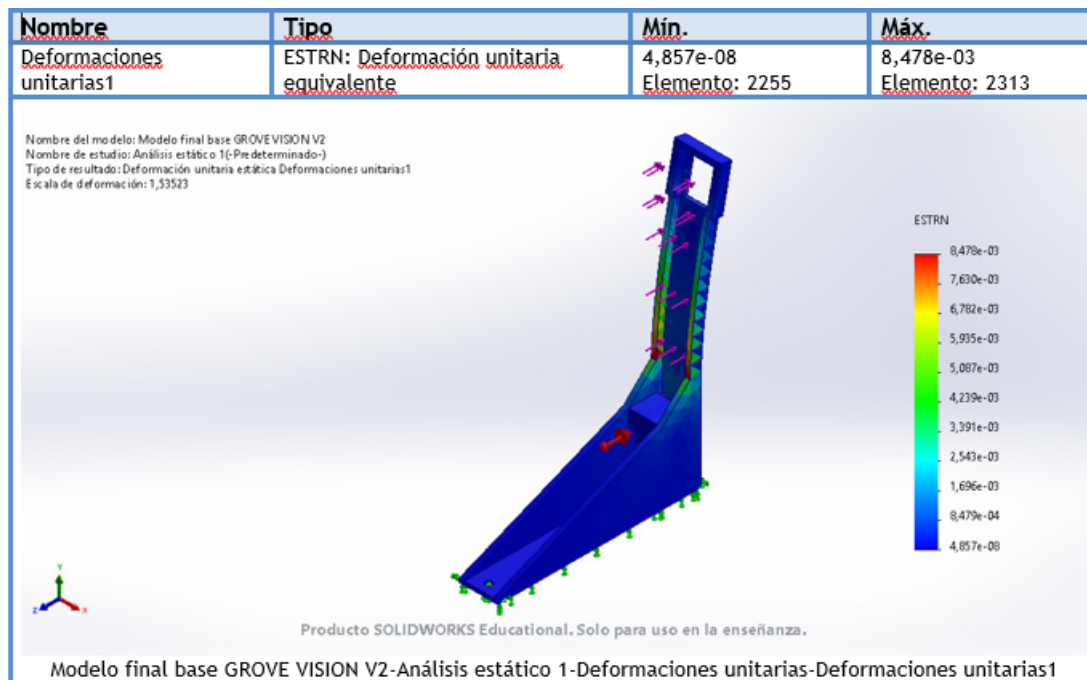


Figura 9: Deformacion unitaria

En cuanto a la deformación unitaria, los valores obtenidos son bajos y se concentran en zonas específicas del soporte, principalmente donde se presentan concentraciones de esfuerzo. Dichos valores se encuentran dentro del rango elástico del material, confirmando que el soporte trabaja sin presentar deformaciones plásticas ni daño estructural.

Esto demuestra que el diseño posee una adecuada rigidez y estabilidad mecánica para soportar la carga aplicada.

Considerando que el peso del módulo Grove Vision AI V2 y de la cámara OV5647-62 es reducido, los esfuerzos generados sobre la estructura impresa resultan mínimos en comparación con la capacidad resistente del material. La utilización de un relleno del 100 % contribuye significativamente a incrementar la rigidez del soporte y a disminuir las deformaciones, proporcionando una estructura sólida y confiable.

Los resultados del análisis confirman que el diseño es adecuado para su implementación en el prototipo, ya que garantiza la estabilidad mecánica necesaria para el correcto funcionamiento del sistema de visión artificial en condiciones reales de operación. Las tensiones generadas se mantienen dentro de límites seguros, los desplazamientos son reducidos y las deformaciones unitarias no comprometen la integridad estructural del diseño.

Por lo tanto, el soporte diseñado cumple con los requerimientos mecánicos necesarios para su aplicación en el sistema de clasificación, asegurando una correcta sujeción del sensor y una operación estable durante el procedimiento de control de calidad.

VIII-B4. Diseño y análisis estructural de la base para el transporte de fresas: Se eligió el acero inoxidable AISI 304 como material para la base de las fresas que se colocan en la banda transportadora, por su adecuación para utilizarse en contacto con alimentos frescos y por sus propiedades mecánicas. Dado que es fácil de limpiar, resiste a la corrosión y cumple con condiciones higiénicas, este material se usa mucho en el sector alimentario, lo cual lo vuelve apropiado para los sistemas de control de calidad de productos agrícolas.

La selección del material se fundamenta además en experiencia previa en entornos industriales dedicados al procesamiento de productos marinos, específicamente en el diseño y fabricación de piezas mecánicas para camarónicas,

donde el acero inoxidable AISI 304 es empleado debido a su durabilidad y compatibilidad con ambientes húmedos y productos perecederos. Esta experiencia permitió validar su aplicación en un sistema de clasificación de fresas, garantizando condiciones apropiadas para el manejo de alimentos frescos.

Propiedades de material

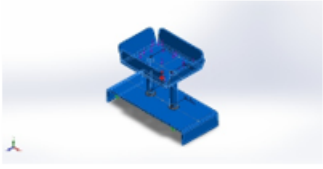
Referencia de modelo	Propiedades	Componentes
	Nombre: AISI 304	Sólido 1(Redondeo1)(Pieza1-6),
	Tipo de modelo: Isotrópico elástico lineal	Sólido 1(Brida de arista1)(Pieza2-3),
	Criterio de error predeterminado: Tensión de von Mises máx.	Sólido 1(Saliente-Extruir1)(Pieza3-6),
	Límite elástico: 2,06807e+08 N/m ²	Sólido 1(Saliente-Extruir1)(Pieza3-7)
	Límite de tracción: 5,17017e+08 N/m ²	
	Módulo elástico: 1,9e+11 N/m ²	
	Coefficiente de Poisson: 0,29	
	Densidad: 8.000 kg/m ³	
	Módulo cortante: 7,5e+10 N/m ²	
	Coefficiente de dilatación térmica: 1,8e-05 /Kelvin	
Datos de curva:N/A		

Figura 10: Propiedades del material

Las propiedades mecánicas consideradas en la simulación incluyen un módulo de elasticidad de 190000000000 N/m², un límite elástico de 206000000 N/m² y una densidad de 8000 kg/m³, valores representativos del comportamiento real del material.

VIII-B5. Metodología del análisis por elementos finitos de base para fresas: El análisis estructural fue realizado mediante un estudio estático utilizando el software SolidWorks Simulation. Se definieron condiciones de frontera de tipo geometría fija en las zonas de sujeción de la base, representando el anclaje real de la estructura al sistema de la banda transportadora.

Como carga principal se consideró el peso promedio de una fresa, así como la fuerza normal ejercida durante su desplazamiento sobre la banda. Para efectos de simulación se aplicó una carga equivalente de 15 N, la cual representa un escenario conservador frente al peso real de las fresas, con el fin de evaluar el comportamiento estructural bajo condiciones de operación desfavorables.

El propósito del análisis fue determinar la distribución de tensiones, desplazamientos, deformaciones unitarias y el factor de seguridad para comprobar que la estructura sea capaz de soportar las cargas relacionadas con el transporte del fruto sin sufrir fallos mecánicos o deformaciones que perjudiquen el proceso de clasificación.

VIII-B6. Resultados del análisis de tensiones (Von Mises) de base para fresas: Los resultados obtenidos indican que la tensión máxima de Von Mises alcanza un valor aproximado de 4190000 N/m², localizada principalmente en las zonas cercanas a los puntos de aplicación de la carga y en regiones con cambios geométricos de la estructura.

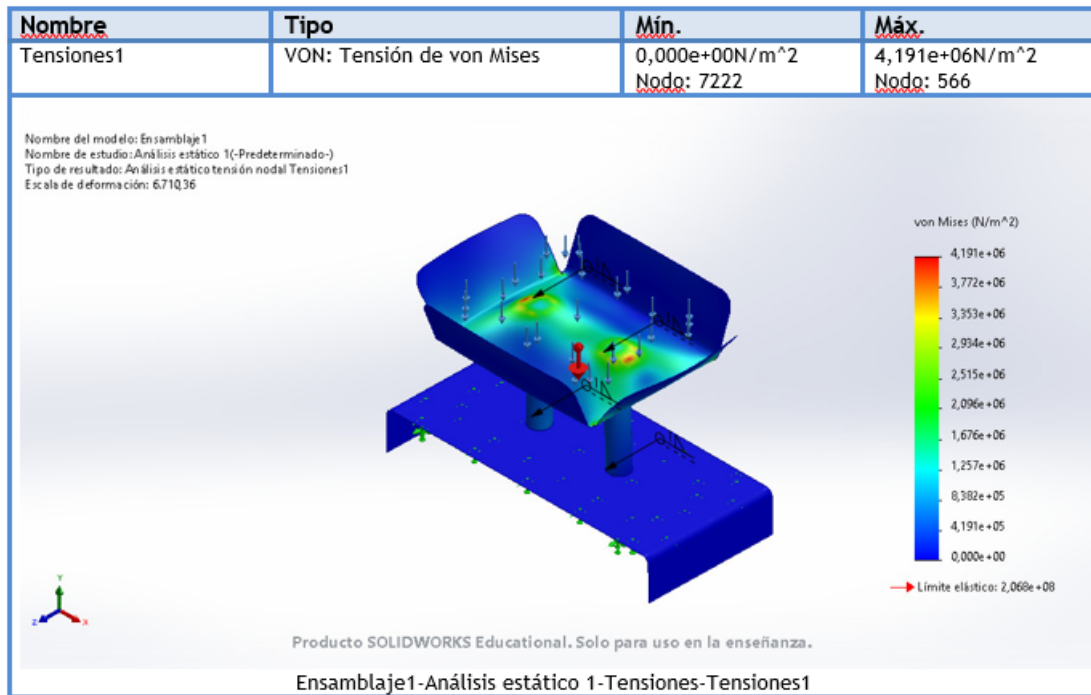


Figura 11: Von Mises

Este valor es menor al límite elástico del acero inoxidable AISI 304, lo que evidencia que la base trabaja completamente dentro del régimen elástico y no presenta riesgo de falla estructural. La distribución de tensiones concuerda con el comportamiento mecánico esperado de una estructura sometida a cargas estáticas de baja magnitud.

VIII-B7. Resultados del análisis de desplazamiento de base para fresas: El análisis de desplazamiento mostró un valor máximo de aproximadamente 0.0012 mm, localizado en las zonas más alejadas de las restricciones mecánicas. Este desplazamiento es prácticamente despreciable y no compromete la estabilidad ni la geometría de la base donde se apoyan las fresas.

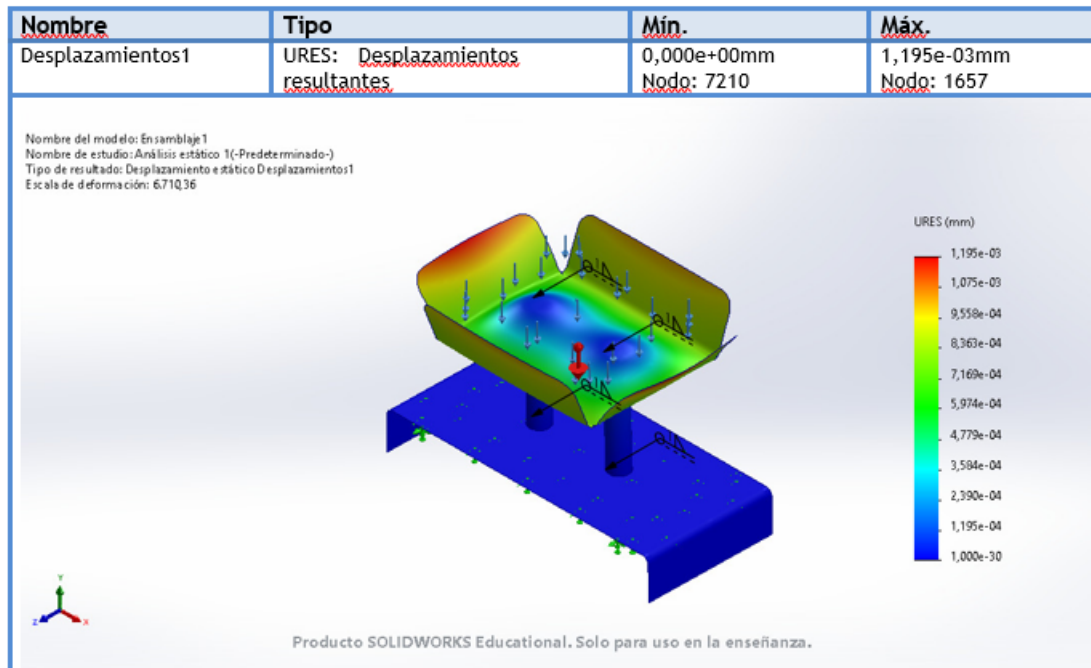


Figura 12: Desplazamientos resultantes

La reducida magnitud del desplazamiento garantiza que el fruto mantenga una posición estable durante su transporte sobre la banda, favoreciendo una correcta captura de imágenes por parte del sistema de visión artificial.

VIII-B8. Resultados del análisis de deformación unitaria de base para fresas: La deformación unitaria máxima obtenida fue del orden de 0.0000153, concentrada en regiones específicas del diseño. Este valor está ubicado dentro del intervalo elástico del material, lo que sugiere que no hay deformaciones permanentes ni daño estructural bajo las condiciones de carga analizadas.

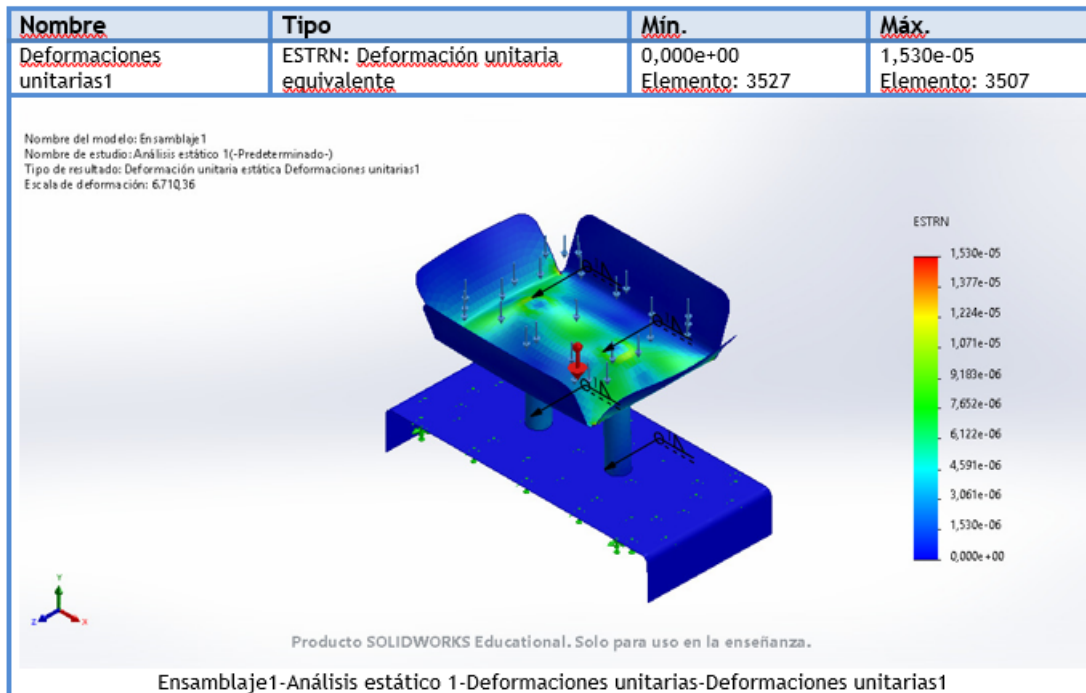


Figura 13: Deformación Unitaria

VIII-B9. *Factor de seguridad:* El análisis del factor de seguridad mostró valores mínimos cercanos a 44, lo que representa un margen amplio entre los esfuerzos generados por el peso de las fresas y la capacidad resistente del material. Este resultado confirma que la estructura posee una elevada confiabilidad mecánica y puede operar con holgura frente a las condiciones reales de trabajo.

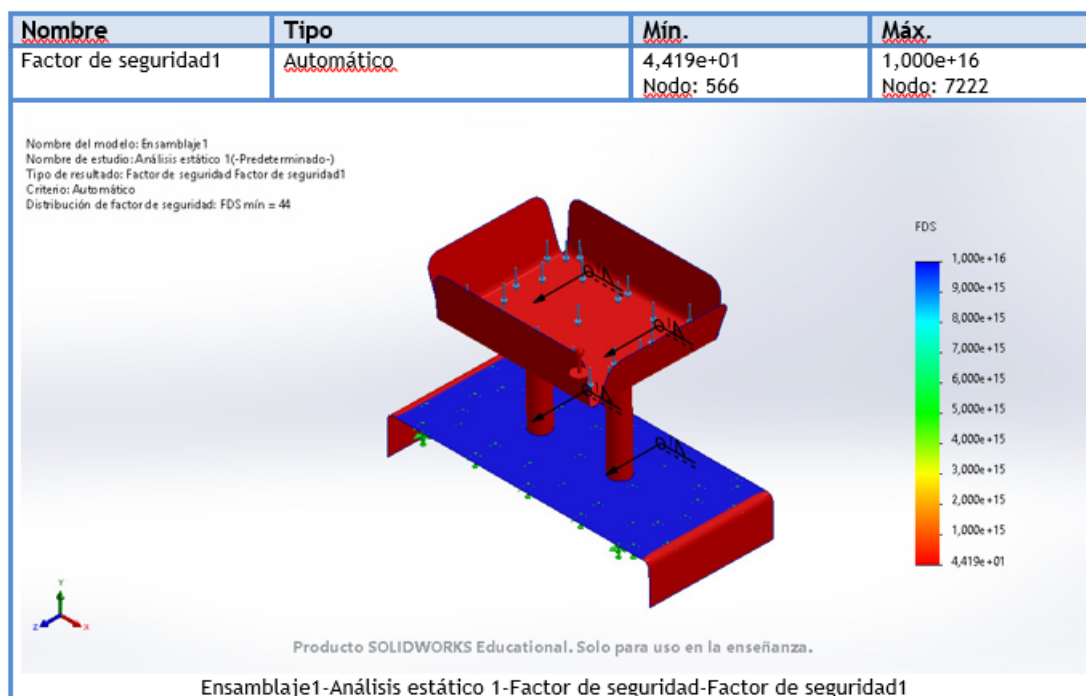


Figura 14: Factor de seguridad

Considerando que el peso real de una fresa es reducido en comparación con la capacidad resistente del acero inoxidable AISI 304, los esfuerzos generados sobre la base son mínimos. La simulación realizada bajo una carga superior al peso real del fruto permite concluir que el diseño presenta un comportamiento estructural seguro incluso en escenarios conservadores.

Además, la selección del acero inoxidable AISI 304 aporta ventajas adicionales relacionadas con la higiene y la durabilidad del sistema, aspectos fundamentales para aplicaciones de control de calidad en productos alimenticios. Su resistencia a la corrosión y facilidad de limpieza permiten su uso continuo sin comprometer la integridad del producto ni el funcionamiento del prototipo. Se concluye que la base diseñada en acero inoxidable AISI 304 es capaz de soportar sin inconvenientes el peso de las fresas durante su transporte sobre la banda, manteniendo su estabilidad estructural y funcional. Las tensiones generadas se mantienen dentro de límites seguros, los desplazamientos son despreciables y el factor de seguridad elevado confirma la confiabilidad del diseño.

En consecuencia, el diseño sugerido satisface los requisitos mecánicos e higiénicos requeridos para ser puesto en funcionamiento en un sistema de clasificación de fresas a través de visión artificial, lo que favorece la solidez y viabilidad real del prototipo creado.

VIII-C. Software y Entrenamiento del Modelo

VIII-C1. *Sense Craft como recolector de data.*: El proceso de adquisición de imágenes se realizó utilizando la plataforma SenseCraft, a través del módulo *Grove Vision AI Module v2* para tomar fotos de fresas en un entorno estático. Las muestras fueron colocadas manualmente en la cinta transportadora, manteniéndose inmóviles durante la captura. Para asegurar que las condiciones de adquisición fueran controladas y consistentes, se diseñó e implementó una estructura cerrada con iluminación LED, que permitió mantener un nivel de luz constante y minimizar el impacto de la luz externa. Además, se llevaron a cabo capturas variando la posición y la orientación de las muestras, con el objetivo de aumentar la variabilidad del conjunto de datos y mejorar la capacidad de generalización del sistema de visión artificial.

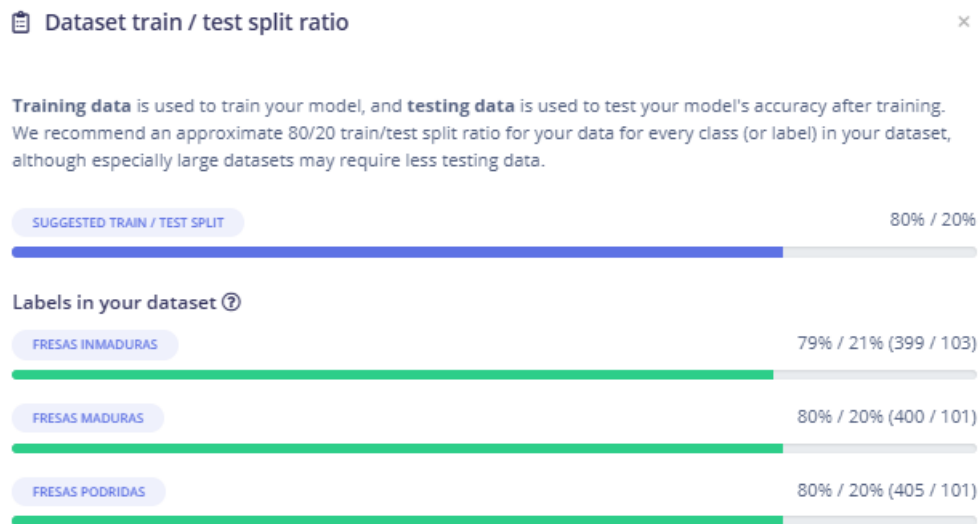


Figura 15: Dataset del manejo de datos

Al término de la adquisición de datos, SenseCraft nos permite exportar una carpeta de todas con todas las fotos capturadas.

VIII-C2. *Edge Impulse entorno de clasificación y entrenamiento.*: Para la creación del modelo de entrenamiento, las imágenes obtenidas en un entorno controlado fueron incorporadas a la plataforma Edge Impulse, donde se realizó la clasificación de cada muestra según su nivel de madurez. Este proceso facilitó la organización del conjunto de datos en categorías bien definidas, lo que ayudó a preparar un conjunto de entrenamiento sólido y representativo. Una vez que la clasificación se completó, comenzó el adiestramiento del modelo de visión artificial.

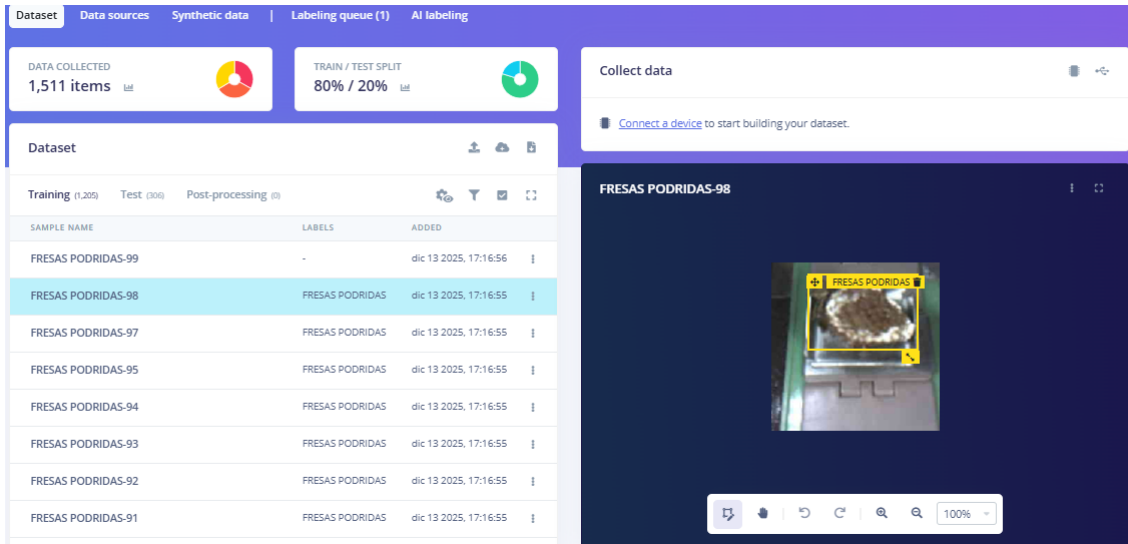


Figura 16: Clasificación de imágenes según la maduración de la fresa.

Se consideraron tres categorías principales: fresas inmaduras, fresas maduras y fresas podridas, además de la clase background para representar regiones sin presencia del objeto de interés. Para el manejo de los datos, se aplicó una división aproximada del 80 % para entrenamiento y 20 % para pruebas (train/test split), recomendada para este tipo de aplicaciones.

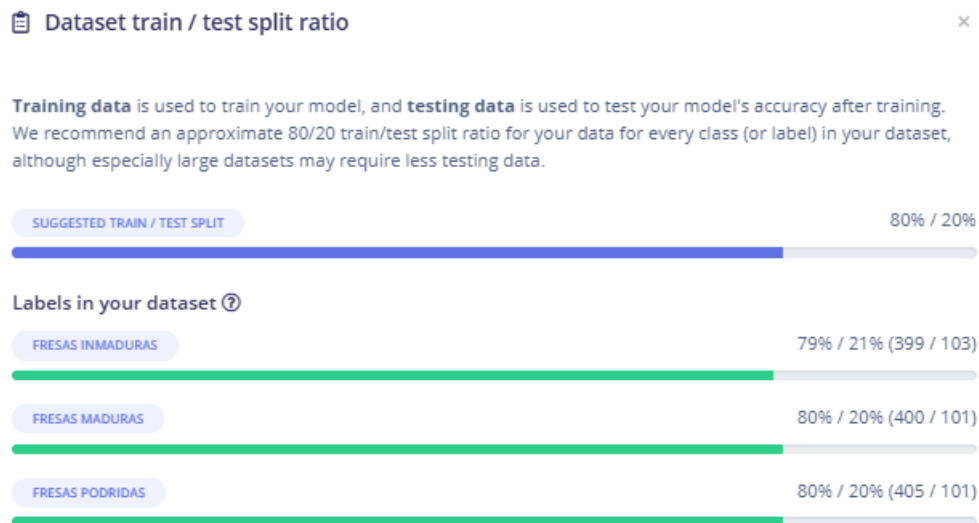


Figura 17: Dataset del manejo de datos

De manera específica, el conjunto de datos quedó distribuido de la siguiente forma: para la clase de fresas inmaduras se utilizaron 399 imágenes para entrenamiento y 103 para pruebas (79 % / 21 %); para las fresas maduras, 400 imágenes para entrenamiento y 101 para pruebas (80 % / 20 %); y para las fresas podridas, 405 imágenes para entrenamiento y 101 para pruebas (80 % / 20 %). Esta distribución permitió mantener un balance adecuado entre las clases, garantizando una evaluación confiable del desempeño del modelo.



Figura 18: Distribución de datos según las clases

El entrenamiento del modelo se realizó utilizando la plataforma Edge Impulse. Se configuró un total de 150 ciclos de entrenamiento (epochs), valor seleccionado tras realizar pruebas preliminares con diferentes cantidades de ciclos (50, 100 y 300), observándose que con 150 ciclos el modelo alcanzaba una adecuada convergencia sin evidencias significativas de sobreajuste. Valores inferiores no permitían una correcta estabilización de la función de pérdida, mientras que valores superiores incrementaban el tiempo de entrenamiento sin mejoras sustanciales en la precisión. La tasa de aprendizaje (learning rate) se estableció en 0.001 debido a que proporciona un equilibrio adecuado entre la velocidad de aprendizaje y la estabilidad del proceso de optimización. Durante las pruebas iniciales, valores mayores generaron oscilaciones en la función de pérdida, mientras que valores menores ralentizaron el entrenamiento sin aportar mejoras relevantes en el desempeño final del modelo. El entrenamiento se ejecutó empleando CPU, considerando que la arquitectura seleccionada presenta un bajo costo computacional y está orientada a su implementación en dispositivos embebidos con recursos limitados, como el Grove Vision AI V2. Esta decisión permite asegurar que un sistema sea viable en aplicaciones reales donde no se dispone de unidades de procesamiento gráfico (GPU).

Para incrementar la diversidad del conjunto de datos de entrenamiento, se implementó el método de "Data augmentation". Esto permite simular diferentes condiciones visuales y ayuda a optimizar la habilidad de generalización del modelo, disminuyendo así el peligro de sobreajuste.

Se apartó el 20 % del total de los datos para validar el modelo, siguiendo una proporción que se usa frecuentemente en sistemas de aprendizaje automático para analizar cómo funciona el modelo con datos que no se emplearon durante la capacitación. Además, se definió el tamaño del lote (batch size) en 32 muestras porque esta cifra proporciona un equilibrio apropiado entre la eficiencia computacional y la estabilidad del gradiente.

The image shows a configuration interface for a neural network. It is divided into several sections:

- Neural Network settings**: A header with a menu icon.
- Training settings**:
 - Number of training cycles: 150
 - Use learned optimizer:
 - Learning rate: 0.001
 - Training processor: CPU
 - Data augmentation:
- Advanced training settings**:
 - Validation set size: 20 %
 - Split train/validation set on metadata key: (empty field)
 - Batch size: 32
 - Profile int8 model:
- Neural network architecture**:
 - Input layer (27,548 features)
 - Model: FOMO (Faster Objects, More Objects) MobileNetV2 0.35
 - Choose a different model (dashed box)
 - Output layer (3 classes)
- Save & train**: A button at the bottom right.

Figura 19: Configuración para el entrenamiento de la red

La capa de entrada del modelo está compuesta por un total de 27 648 características (features), las cuales corresponden a la representación numérica de las imágenes capturadas por la cámara tras el proceso de preprocesamiento realizado por la plataforma Edge Impulse. Durante este proceso, las imágenes originales son redimensionadas, normalizadas y transformadas en un vector de características que contiene información relevante sobre color, textura y patrones visuales del fruto. Esta representación permite reducir la dimensionalidad de los datos y facilitar el aprendizaje del modelo, optimizando el uso de recursos computacionales. La utilización de características extraídas previamente, en lugar de píxeles crudos, resulta especialmente adecuada para sistemas embebidos, ya que disminuye la carga de procesamiento y mejora la eficiencia del modelo durante la inferencia en tiempo real.

Para clasificar las fresas de acuerdo a su nivel de maduración, se utilizó el modelo FOMO, que tiene como base la arquitectura MobileNetV2 y un factor de anchura de 0.35. Se eligió esta arquitectura por su buen rendimiento en tareas de visión artificial en dispositivos con recursos limitados y por el bajo costo computacional que tiene. La red neuronal está compuesta por una capa de entrada correspondiente al vector de características extraídas de las imágenes, seguida por el modelo convolucional MobileNetV2, y finaliza con una capa de salida conformada por tres neuronas, cada una asociada a una clase de maduración de la fresa. Este ajuste facilita la clasificación de manera efectiva y al instante, asegurando un balance apropiado entre la exactitud y el uso de recursos, lo cual es esencial para su utilización en sistemas de control de calidad automatizado. La capa de salida del modelo está conformada por tres neuronas, cada una correspondiente a una categoría de clasificación asociada a las etapas de maduración de la fresa: inmadura, en proceso de maduración y madura. Cada neurona genera una probabilidad que indica el grado de pertenencia de la muestra de entrada a una de las clases definidas. La clase final asignada a cada fresa se determina seleccionando aquella con el mayor valor de probabilidad (argmax). Esta estructura permite realizar una clasificación multiclase de forma directa y eficiente, facilitando la identificación automática del estado de maduración del fruto en tiempo real. La definición de tres clases responde a criterios prácticos de control de calidad, ya que dichas categorías representan los estados más relevantes para la toma de decisiones en procesos de selección y comercialización de frutas, permitiendo separar aquellas que no cumplen con los requisitos mínimos de madurez de aquellas aptas para su distribución. Adicionalmente, se aplicó la técnica de cuantización del modelo mediante el perfilado en formato INT8 (Profile int8 model), con el fin de reducir el tamaño del modelo y optimizar su ejecución en el dispositivo embebido. Esta estrategia permite disminuir el consumo de memoria y acelerar el tiempo de inferencia, manteniendo una precisión aceptable, lo cual resulta fundamental para la operación del sistema en tiempo real.

VIII-D. Fase de Proceso Post-Entrenamiento

VIII-D1. Conversión del formato archivo TensorFlow Lite a formato Vela: Una vez concluido el entrenamiento del modelo de visión artificial creado en la plataforma Edge Impulse, se convirtió al formato Vela para satisfacer los requerimientos de optimización y compatibilidad del hardware utilizado. Esta adaptación permite que el modelo funcione adecuadamente en dispositivos embebidos, dado que el compilador Vela mejora las operaciones del modelo para ser utilizado en aceleradores neuronales, disminuyendo el tiempo de inferencia y el uso de recursos computacionales. Así, la conversión a Vela garantiza un rendimiento eficiente del sistema de visión artificial, preservando la exactitud del modelo y asegurando su operatividad en tiempo real dentro del entorno designado.

Posteriormente, el modelo convertido fue importado nuevamente a la plataforma SenseCraft, donde se integró al sistema de clasificación propuesto. Esta etapa permitió la implementación del modelo en el entorno físico, posibilitando la clasificación de las fresas según su etapa de maduración y verificando el correcto funcionamiento del sistema.

VIII-E. Fase de Construcción del Prototipo

La etapa final del enfoque metodológico incluyó el montaje y la unión del prototipo operativo, en el que se fusionaron los elementos mecánicos, eléctricos y de software del sistema. Se utilizaron perfiles de aluminio para la estructura mecánica, lo que aportó rigidez y estabilidad al prototipo, además de actuar como soporte para la cerradura de iluminación empleada durante la toma de imágenes. En la parte electrónica, se incorporaron los dispositivos necesarios para la captura y análisis de imágenes, mientras que la parte de software permitió la implementación del modelo de visión artificial que había sido previamente capacitado. Esta fase integró el funcionamiento conjunto de todos los subsistemas, facilitando la validación del sistema que clasifica fresas según su grado de maduración en un entorno controlado.

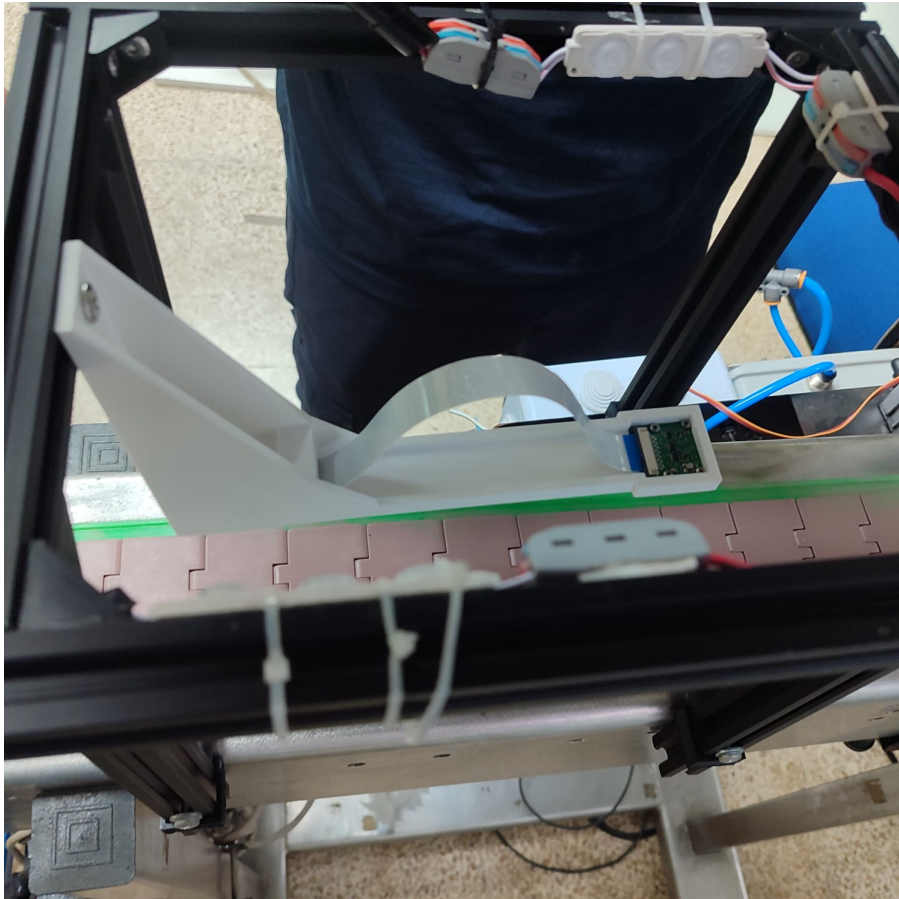


Figura 20: Estructura armada por dentro.

IX. RESULTADOS

IX-A. Resultados del proceso del entrenamiento

El desempeño del modelo fue evaluado utilizando el conjunto de validación, obteniéndose un valor de F1-score del 98.4%, lo que evidencia un equilibrio adecuado entre la precisión y la capacidad de recuperación (recall) del sistema. Este resultado indica que el modelo no solo clasifica correctamente la mayoría de las muestras, sino que también presenta una baja tasa de falsos positivos y falsos negativos.

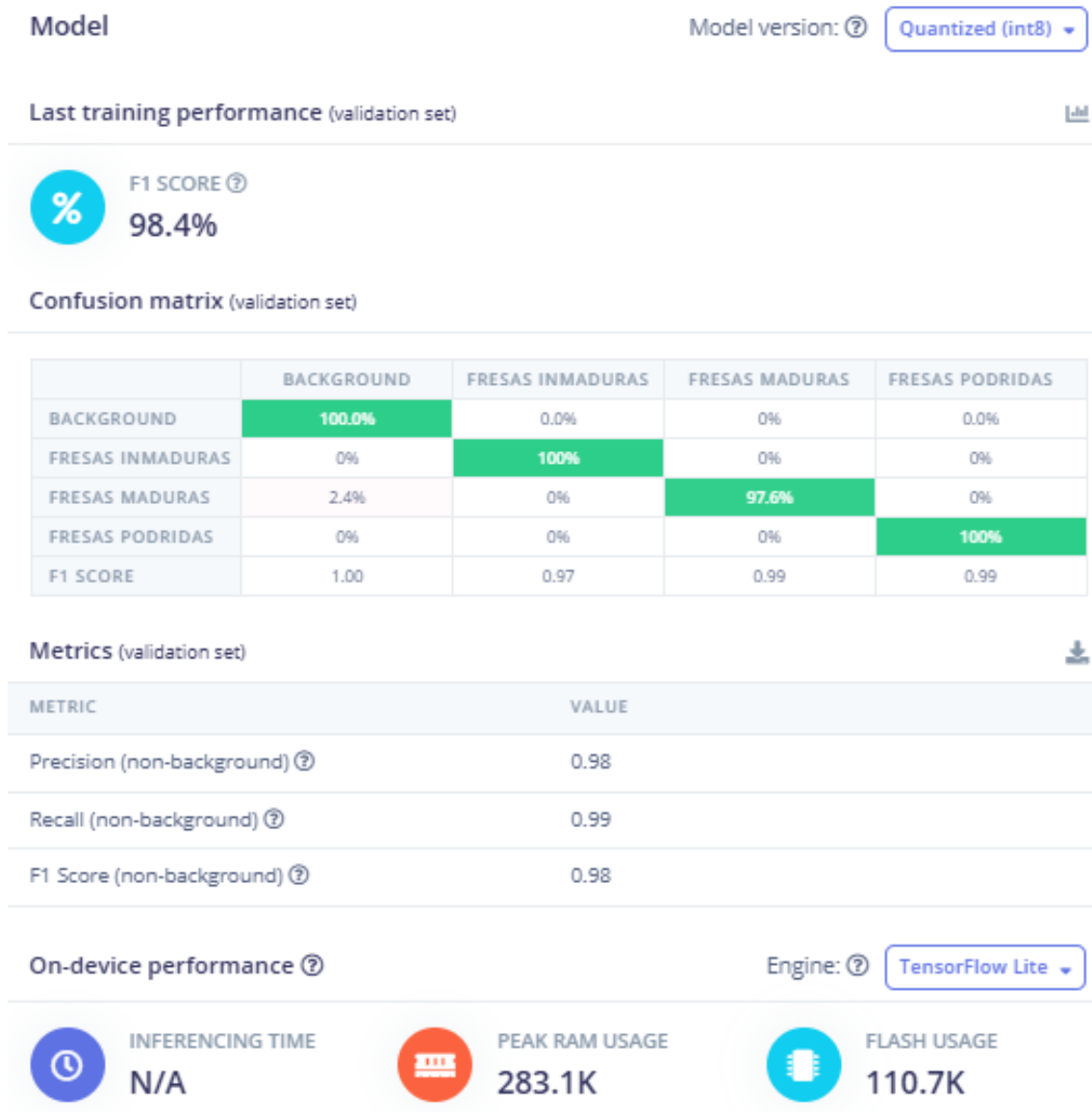


Figura 21: Set de validacion, matriz de confusión y métricas.

Al considerar las métricas excluyendo la clase background, se alcanzaron valores de precisión cercanos al 98% y valores de recall de aproximadamente 99%, lo que demuestra una alta confiabilidad en la identificación de las etapas de maduración de las fresas. La matriz de confusión muestra una clasificación perfecta para las clases correspondientes a fresas inmaduras y fresas podridas, alcanzando un 100% de aciertos.

En el caso de la clase de fresas maduras, se observa una tasa mínima de error del 2.4 %, asociada principalmente a confusión con el fondo de la imagen. Este comportamiento puede atribuirse a la similitud cromática entre ciertas áreas del fruto maduro y el entorno de iluminación, así como a la presencia de reflejos luminosos en la superficie del fruto. Estos resultados evidencian una clara separación entre las clases definidas y una adecuada capacidad del modelo para discriminar los distintos estados de maduración bajo condiciones controladas de iluminación. No obstante, que haya errores en la clase de fresas maduras indica que el sistema es más sensible a los cambios visuales durante las etapas intermedias de maduración, cuando las diferencias de color y textura son menos notables. En cuanto al rendimiento del modelo en el dispositivo embebido, se registró un consumo aproximado de 283 KB de memoria RAM y 110 KB de memoria Flash, lo que confirma que el modelo es compatible con plataformas de bajo costo y recursos limitados, como el Grove Vision AI V2. La aplicación de cuantización en formato INT8 permitió reducir significativamente el tamaño del modelo sin comprometer de manera significativa la precisión obtenida. En conjunto, los resultados del entrenamiento indican que el modelo presenta un desempeño satisfactorio en un entorno controlado; no obstante, su comportamiento podría verse afectado en escenarios reales donde existan variaciones de iluminación, velocidad de desplazamiento de la banda transportadora o presencia de imperfecciones físicas en el fruto, aspectos que deberán ser considerados en futuras implementaciones del sistema.

IX-A1. Matriz de confusión: Para una mejor presentación de los resultados adquiridos durante el entrenamiento, se utilizó un código en Google Colab que permitió visualizar la matriz de confusión de esta. La matriz de confusión del modelo optimizado en INT8, revela una predominancia notable de datos en la diagonal principal, lo que sugiere un elevado grado de aciertos en la clasificación de todas las categorías analizadas. Las clases de fresas inmaduras, maduras y en estado de descomposición muestran tasas de verdaderos positivos casi completas, presentando pocos errores de clasificación. Específicamente, la categoría de fresas maduras obtiene una precisión de 0.9808, un recall de 0.9968 y un F1-score de 0.9887, lo que demuestra un balance satisfactorio entre sensibilidad y precisión del modelo. La precisión general de 0.9998 respalda un desempeño excepcional del sistema. Los errores identificados se agrupan esencialmente en la clase de fondo, lo cual es comprensible debido a la variedad visual de esta categoría y no afecta el rendimiento del modelo en las clases de interés. Estos hallazgos refuerzan la solidez y la fiabilidad del modelo de visión artificial para clasificar fresas según su grado de madurez.

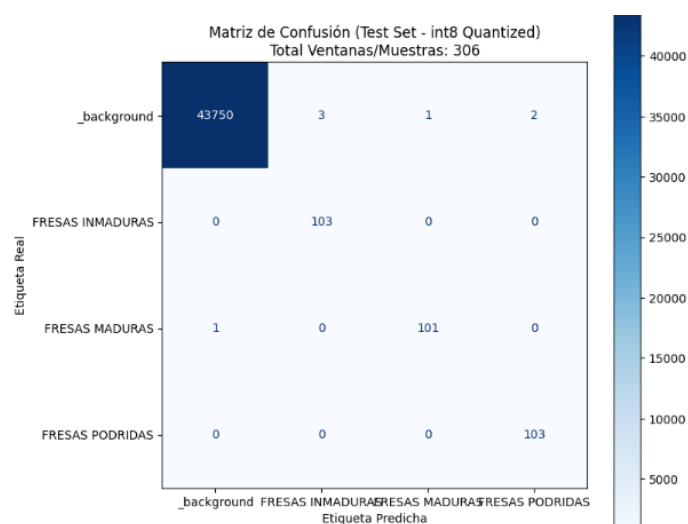


Figura 22: Matriz de confusión

IX-B. Modelo entrenado

El modelo ya entrenado fue evaluado utilizando el conjunto de prueba proporcionado por la plataforma Edge Impulse, alcanzando una accuracy global del 97.71 %, lo que indica un alto nivel de acierto en la clasificación de las muestras analizadas.

Si solamente tomamos en cuenta las predicciones de las clases de fresas (sin incluir la clase background), los valores obtenidos son: precisión 0.98, recall 1.00 y F1-score 0.99; esto indica un desempeño sólido y equilibrado del sistema de clasificación. Estos hallazgos muestran que el modelo tiene una capacidad casi total para detectar la presencia de fresas en las imágenes analizadas y una tasa de falsos positivos muy baja.

El análisis visual mediante la herramienta Feature Explorer permitió identificar las predicciones correctas e incorrectas del modelo, representadas mediante puntos verdes y rojos, respectivamente. Esta visualización facilita la inspección de los casos en los que el sistema presentó errores de clasificación, permitiendo analizar patrones asociados a dichas fallas.

Las predicciones incorrectas se concentraron principalmente en muestras con condiciones visuales atípicas, tales como variaciones en la iluminación, presencia de sombras o superposición parcial del fruto con el fondo. Estos factores afectan la extracción de características visuales, generando confusión entre clases con atributos cromáticos similares.

El valor de precisión obtenido (0.98) indica que, cuando el sistema predice la presencia de una fresa en una determinada categoría, dicha predicción es correcta en el 98 % de los casos, lo cual se traduce en una baja cantidad de falsos positivos. Por su parte, el recall del 100 % demuestra que el modelo fue capaz de detectar todas las fresas presentes en las imágenes evaluadas, sin omitir muestras reales. El F1-score de 0.99 confirma un equilibrio adecuado entre ambas métricas.

% ACCURACY ?
97.71%

Metrics for Object detection



METRIC	VALUE
Precision (non-background) ?	0.98
Recall (non-background) ?	1.00
F1 Score (non-background) ?	0.99

Feature explorer ?

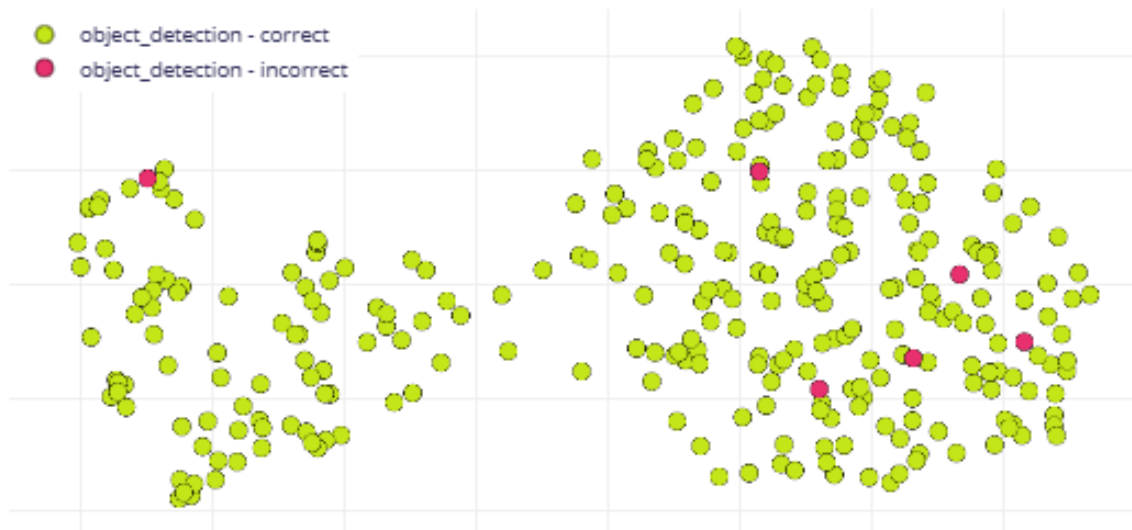


Figura 23: Resultados del modelo entrenado

Estos resultados validan la capacidad del sistema para operar en tiempo real bajo condiciones controladas; sin embargo, su desempeño podría verse afectado en un entorno de producción real, donde existirían variaciones más pronunciadas en la iluminación, cambios en la velocidad de la banda transportadora y posibles imperfecciones físicas en los frutos, tales como manchas, deformaciones o suciedad superficial.

En conjunto, los resultados obtenidos demuestran que el modelo presenta un desempeño satisfactorio para aplicaciones de control de calidad automatizado en un entorno experimental, aunque se recomienda realizar pruebas adicionales en condiciones reales de operación para evaluar su estabilidad y confiabilidad a largo plazo.

IX-C. Pruebas en entorno real.

IX-C1. Resultados del sistema en tiempo real bajo iluminación controlada.: El sistema fue evaluado en tiempo real mediante la ejecución del prototipo físico bajo las mismas condiciones de iluminación y entorno controlado utilizadas durante la adquisición de las imágenes de entrenamiento. Como se observa en la "Figura 24z "Figura 25", el modelo fue capaz de detectar y clasificar correctamente fresas maduras e inmaduras mientras se desplazaban sobre la banda transportadora, asignando la categoría correspondiente con niveles de confianza superiores al 90 %.

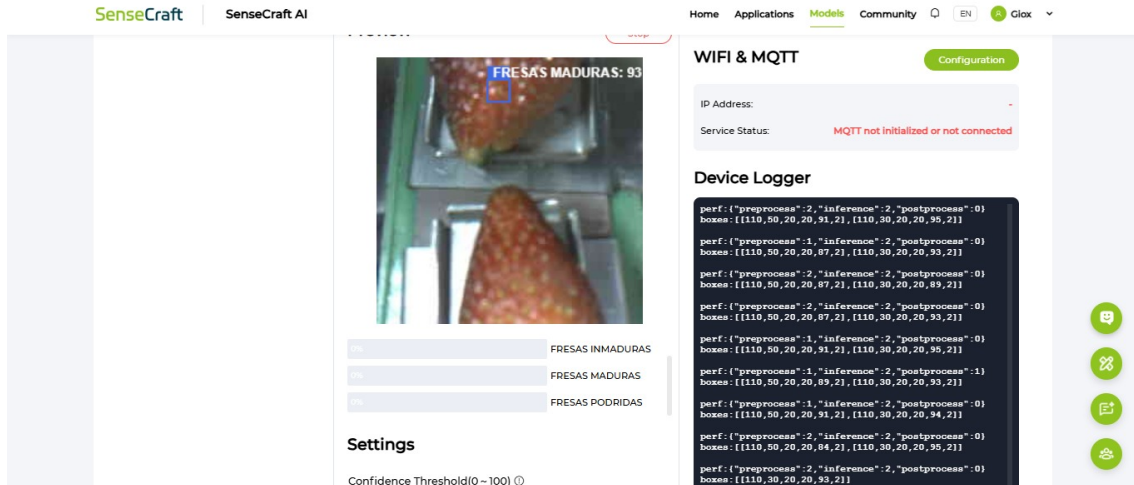


Figura 24: Clasificación en tiempo real de una fresa madura mediante el sistema propuesto.

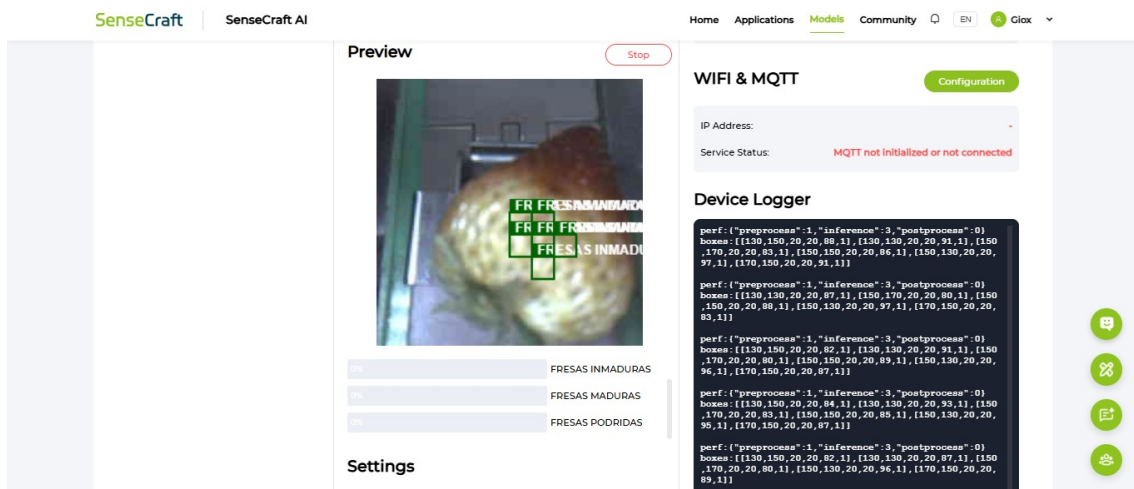


Figura 25: Clasificación en tiempo real de una fresa inmadura mediante el sistema propuesto.

Durante estas pruebas experimentales no se registraron errores de clasificación, alcanzándose una efectividad del 100% en las muestras evaluadas. Este resultado confirma que el modelo presenta un desempeño óptimo cuando opera en condiciones consistentes con su fase de entrenamiento, validando la correcta integración entre el hardware (cámara OV5647-62 y Grove Vision AI V2) y el modelo de visión artificial implementado en SenseCraft.

Así mismo, la visualización de los resultados en tiempo real permitió verificar la estabilidad del sistema durante la operación continua, demostrando su capacidad para procesar imágenes, ejecutar inferencias y emitir resultados de clasificación de manera inmediata, lo que evidencia su viabilidad para aplicaciones de control de calidad automatizado en entornos experimentales.

IX-C2. Resultados con variación en la iluminación.: Con el objetivo de evaluar la robustez del modelo ante cambios en las condiciones de iluminación, se realizaron pruebas sin el uso de las luces LED dentro de la caja de ambientación, utilizando únicamente la iluminación ambiental del entorno. Bajo estas condiciones, se evidenció una disminución significativa en el desempeño del sistema. En la "Figura 18" se observa un caso de clasificación incorrecta, donde una fresa inmadura es identificada como fresa podrida. Este comportamiento indica que el modelo presenta confusión entre clases cuando las características visuales (color, textura y brillo) se ven alteradas por una iluminación insuficiente o no controlada. La reducción de contraste y la presencia de sombras influyen directamente en las características extraídas por la red neuronal, afectando el proceso de inferencia.

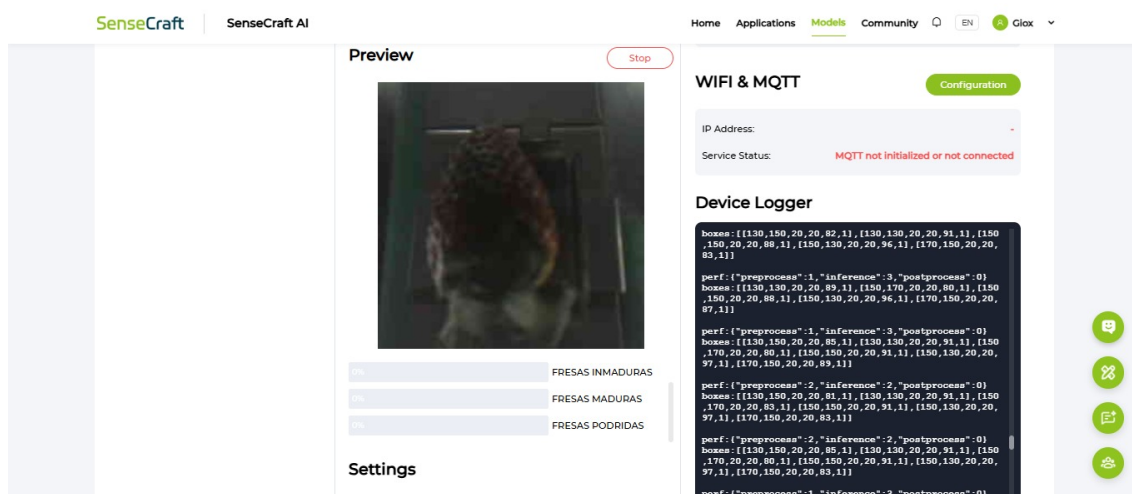


Figura 26: Clasificación errónea por falta de iluminación.

Por otro lado, en la "Figura 19" se presenta un escenario en el cual el modelo no logra detectar la presencia de la fresa en la imagen, lo que evidencia una pérdida total de capacidad de reconocimiento bajo condiciones de iluminación no similares a las utilizadas durante el entrenamiento. Este resultado sugiere que el modelo fue altamente dependiente de la iluminación controlada empleada en la etapa de adquisición de datos. Estos resultados demuestran que, si bien el sistema alcanza una efectividad del 100% bajo las mismas condiciones de iluminación utilizadas durante el entrenamiento, su desempeño se ve comprometido cuando existen variaciones en la iluminación, sobre todo si esta se ve reducida. Lo anterior resalta la importancia de incorporar mayor diversidad de condiciones lumínicas en el conjunto de entrenamiento, así como la necesidad de un sistema de iluminación controlado para garantizar una clasificación confiable en aplicaciones reales.

IX-C3. Resultados con iluminación ambiental (sin caja y sin luces LED): Se realizaron pruebas adicionales eliminando la caja de ambientación y las luces LED, empleando únicamente la iluminación proporcionada por una lámpara convencional de una habitación. En estas pruebas se evaluaron únicamente las clases de fresas maduras e inmaduras. Bajo estas condiciones, el modelo alcanzó una efectividad del 100% en la clasificación de las muestras analizadas.

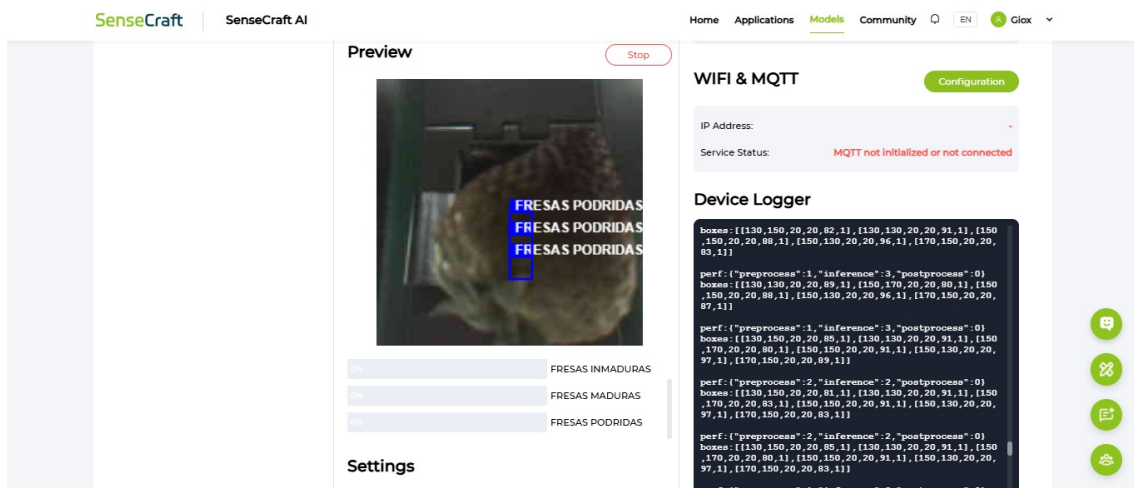


Figura 27: No detección por falta de iluminación.

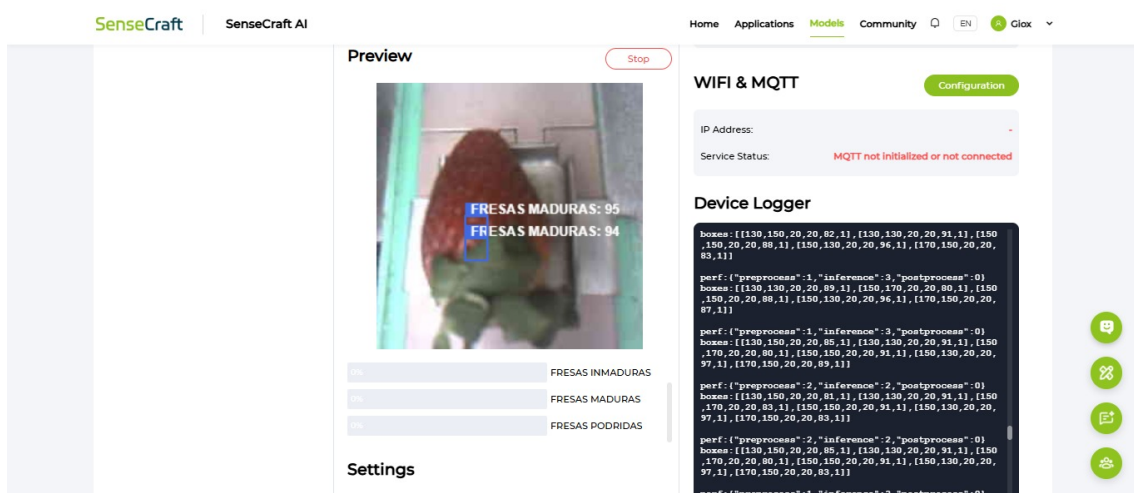


Figura 28: No detección por falta de iluminación.

Los resultados muestran que el sistema logró identificar correctamente ambas clases sin presentar errores de detección ni confusión entre categorías, lo que evidencia una adecuada capacidad de generalización frente a variaciones moderadas en las condiciones de iluminación. Esto sugiere que el modelo no depende exclusivamente de un entorno de iluminación controlado para realizar la clasificación de estos dos estados de maduración.

Los resultados obtenidos validan el funcionamiento del sistema en un entorno más cercano a condiciones reales para las clases evaluadas.

IX-C4. Resultados sin caja de ambientación y con iluminación LED: Se realizaron pruebas experimentales retirando la caja de ambientación, pero manteniendo la iluminación mediante luces LED blancas, con el fin de evaluar el comportamiento del sistema ante cambios en la configuración física del entorno. Los resultados obtenidos mostraron una efectividad del 100% en la clasificación de las muestras evaluadas, sin presentarse errores de detección ni confusión entre clases. Este rendimiento señala que el modelo es sólido a pesar de no tener un ambiente cerrado, siempre que la iluminación sea uniforme y apropiada, suministrada por las luces LED.

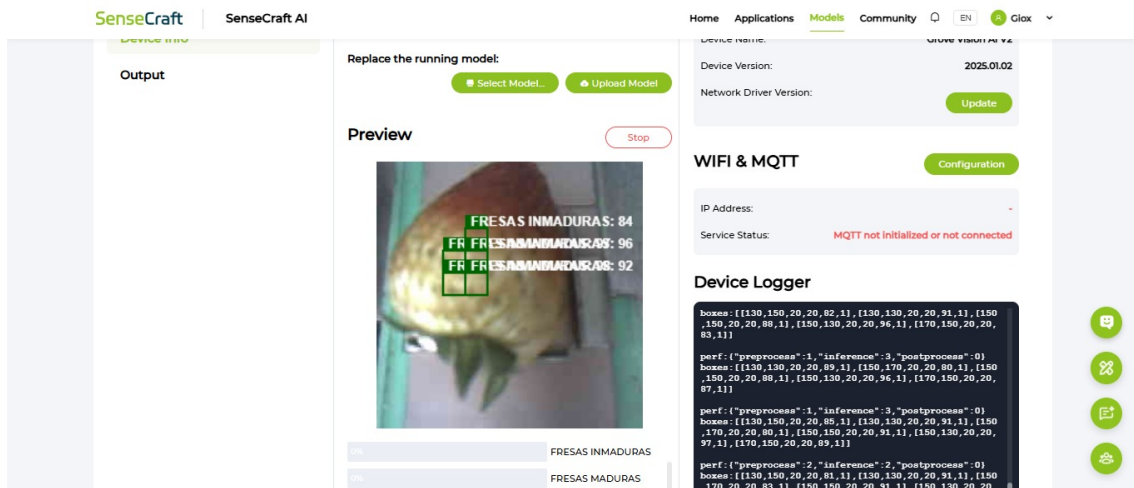


Figura 29: No detección por falta de iluminación.

Estos resultados confirman que el sistema no depende estrictamente de la caja de ambientación para operar correctamente, sino que su desempeño está principalmente asociado a la estabilidad de las condiciones de iluminación. No obstante, al igual que en las pruebas con iluminación ambiental, la clase de fresas podridas no fue evaluada bajo estas condiciones, lo cual se establece como una limitación del presente estudio.

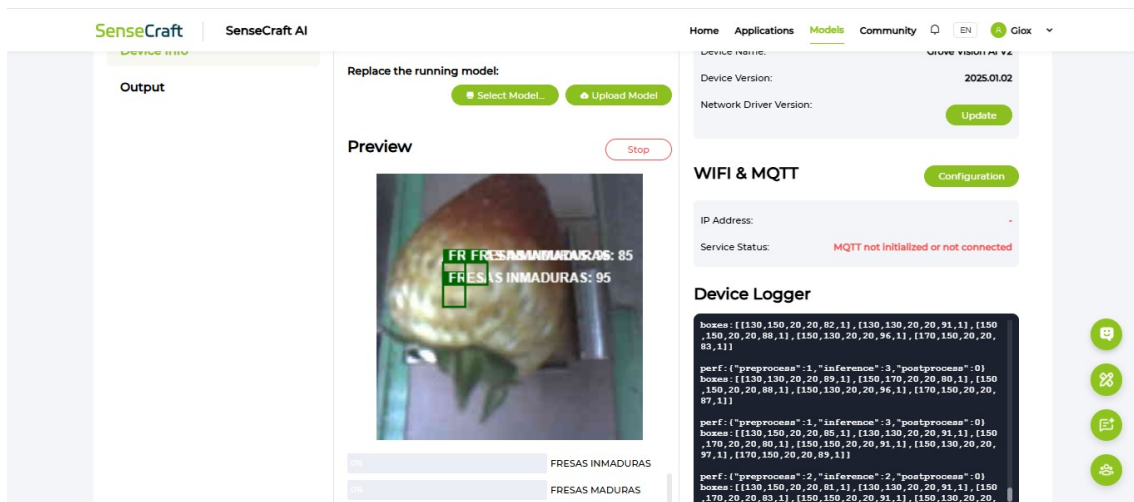


Figura 30: No detección por falta de iluminación.

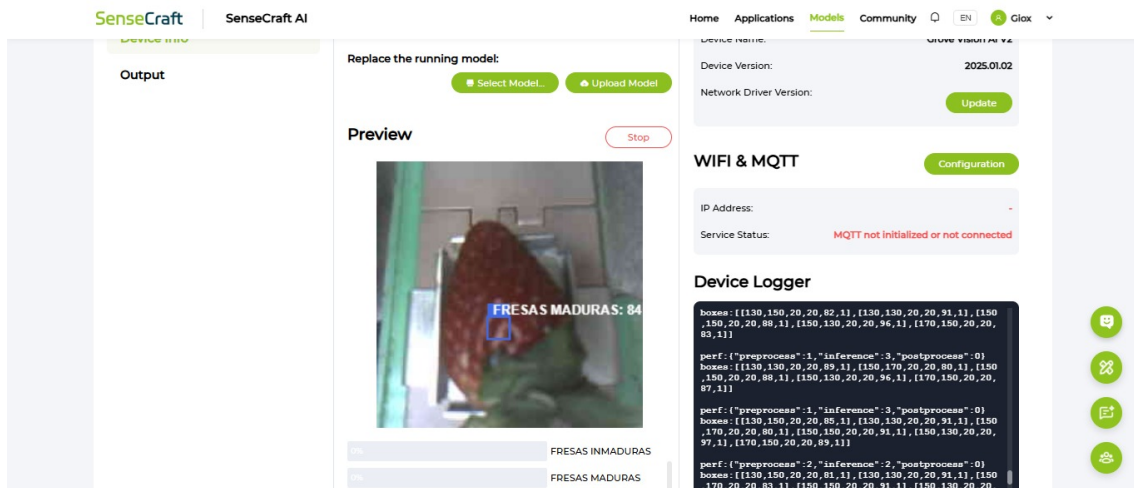


Figura 31: No detección por falta de iluminación.

X. CRONOGRAMA

A continuación se muestra el cronograma de trabajo en la figura.

CRONOGRAMA		
SEMANA	FASE	ACTIVIDAD
27-31 OCTUBRE	DISEÑO	Analisis de propuesta a presentar y recopilacion de datos.
3-7 NOVIEMBRE	DISEÑO	Mejora de diseño mecanico y planificacion de montaje de estructura en la banda
10-28 NOVIEMBRE	ENTRENAMIENTO	Inicio de toma de fotos de fresas maduras.
1-5 DICIEMBRE	ENTRENAMIENTO	Inicio de toma de fotos de fresas podridas.
8-12 DICIEMBRE	ENTRENAMIENTO	Inicio de toma de fotos de fresas verdes.
22-26 DICIEMBRE	PROTOTIPO	Montaje de componentes mecanicos y electricos
5-9 ENERO	PROTOTIPO	Programacion e inicio de pruebas del sistema de vision artificial.
19-23 ENERO	VALIDACION	Evaluacion de la precision del sistema de clasificacion de fresas.
2-5 FEBRERO	VALIDACION	Analisis de resultados
2-5 FEBRERO	DOCUMENTACION	Finalizacion de documentacion de los resultados obtenidos y elaboracion del documento final.

Figura 32: Cronograma

XI. PRESUPUESTO

A continuación se muestra la tabla de presupuesto del proyecto. ??.

TABLA DE PRESUPUESTO				
NOMBRE DEL ELEMENTO	DESCRIPCION	CANTIDAD	PRECIO UNITARIO	VALOR TOTAL
CAMARA GROVE VISION V2	UNIDAD	1	1	50
MODULO XIAO ESP32	UNIDAD	1	1	50
FUENTE DE 12V	UNIDAD	1	1	8
BASE DE FRUTILLAS	UNIDAD	1	1	45
BASE DE MODULOS	UNIDAD	1	1	5
LUCES DE ILUMINACION	UNIDAD	1	2	9
SUBTOTAL				167
MANO DE OBRA	CORTE LASER	3	30	30
MANO DE OBRA	PLEGADO	1	30	30
MANO DE OBRA	SOLDADURA LASER	1	30	30
SUBTOTAL COSTOS LOGISTICOS				90
TOTAL				257

Tabla I: Presupuesto

XII. CONCLUSIONES

El presente trabajo cumplió satisfactoriamente con el objetivo general de implementar un sistema de clasificación de fresas empleando visión artificial para el control de calidad mediante sus etapas de maduración. A través del desarrollo de un prototipo físico con banda transportadora, una cámara OV5647-62 y el uso de plataformas de Edge AI como SenseCraft y Edge Impulse, se logró construir un sistema funcional capaz de operar en tiempo real y clasificar fresas en función de su estado de madurez.

En relación con el primer objetivo específico, se logró crear un conjunto de datos propio mediante la captura de imágenes directamente en el prototipo desarrollado. Las imágenes fueron adquiridas bajo un entorno controlado de iluminación con luces LED blancas dentro de una caja de ambientación, garantizando condiciones uniformes para el entrenamiento del modelo. Este conjunto de datos permitió representar adecuadamente las clases de fresas inmaduras, maduras y podridas, constituyendo una base sólida para el proceso de aprendizaje del clasificador.

Respecto al segundo objetivo específico, se entrenó una red neuronal convolucional utilizando los datos obtenidos en tiempo real, alcanzando métricas de desempeño elevadas. El modelo presentó un F1-score de hasta 98.4 %, con valores de precisión cercanos al 98 % y recall aproximado del 99 %, lo que evidencia una adecuada capacidad del sistema para aprender las características visuales asociadas a cada etapa de maduración de la fresa. La cuantización del modelo en formato INT8 permitió además reducir el consumo de recursos computacionales, posibilitando su implementación en dispositivos embebidos de bajo costo y bajo consumo energético.

En cumplimiento del tercer objetivo específico, se validó el funcionamiento del sistema mediante la evaluación cuantitativa de su efectividad en la clasificación. Las pruebas realizadas en tiempo real bajo las mismas condiciones de iluminación empleadas durante la captura de datos mostraron una efectividad del 100 % en la clasificación de las muestras evaluadas. Asimismo, se llevaron a cabo pruebas adicionales variando las condiciones de iluminación, tanto con luces LED sin caja de ambientación como con iluminación ambiental, donde se mantuvo una efectividad del 100 % para las clases de fresas maduras e inmaduras.

Sin embargo, al eliminar completamente la iluminación controlada y trabajar bajo condiciones de baja iluminación, el sistema presentó errores de clasificación, principalmente confundiendo fresas inmaduras con fresas podridas y

omitiendo la detección de fresas maduras. Este comportamiento evidencia que la iluminación constituye un factor crítico para el desempeño del modelo, ya que las variaciones de brillo y contraste afectan directamente las características visuales utilizadas por la red neuronal para la clasificación. Desde el punto de vista práctico, los resultados obtenidos demuestran que el sistema propuesto es viable para aplicaciones de control de calidad automatizado en procesos agroindustriales, siempre que se mantengan condiciones mínimas de iluminación adecuadas. El bajo consumo de memoria RAM (aproximadamente 283 KB) y memoria Flash (110 KB) confirma la factibilidad de su implementación en plataformas embebidas, lo que representa una ventaja frente a soluciones basadas en sistemas de alto costo computacional. Como principales limitaciones del proyecto se identifican la dependencia de un entorno de iluminación similar al utilizado durante el entrenamiento y la evaluación parcial de la clase de fresas podridas bajo condiciones de iluminación variable. Asimismo, el conjunto de datos fue capturado en un entorno específico, lo que podría limitar la generalización del sistema ante escenarios con fondos, cámaras o condiciones lumínicas distintas. Finalmente, se concluye que el sistema desarrollado cumple con los objetivos planteados y valida el uso de técnicas de visión artificial y aprendizaje automático para el control de calidad de productos agrícolas. Este trabajo aporta una solución funcional, de bajo costo y con alta precisión para la clasificación de fresas por su estado de maduración, sentando los fundamentos para futuras indagaciones enfocadas en mejorar la robustez del sistema mediante la ampliación del conjunto de datos, la incorporación de nuevas clases y la evaluación en entornos productivos reales.

XIII. RECOMENDACIONES

Para trabajos futuros, se aconseja enriquecer el conjunto de datos al incluir imágenes tomadas en una mayor variedad de situaciones reales, como cambios drásticos en la iluminación, diversos fondos, velocidades variadas de las bandas transportadoras y tamaños distintos de fresas. Esto ayudaría a mejorar la capacidad del modelo para generalizar y a disminuir los errores en la clasificación en situaciones no controladas.

De igual manera, es recomendable analizar el rendimiento del sistema al incluir todas las categorías consideradas durante el entrenamiento (fresas maduras, inmaduras y en mal estado) en condiciones reales en el campo, con el objetivo de obtener métricas de rendimiento más completas y confirmar su aplicabilidad en procesos industriales de aseguramiento de la calidad.

Desde un enfoque mecánico, se sugiere perfeccionar el diseño del sistema para facilitar la limpieza y el mantenimiento, teniendo en cuenta las normativas de higiene relacionadas con el manejo de alimentos frescos. Aunque el uso de materiales como el acero inoxidable AISI 304 ha demostrado ser efectivo, podrían buscarse opciones alternativas que disminuyan gastos sin afectar ni la resistencia mecánica ni la seguridad del producto.

Por último, se propone incorporar el sistema con mecanismos automáticos de separación (como actuadores o compuertas) para clasificar físicamente las fresas según su nivel de madurez, convirtiendo así el prototipo en un sistema que funcione completamente automatizado para la supervisión de calidad en tiempo real.

ANEXO A
DISEÑOS EN SOLIDWORKS

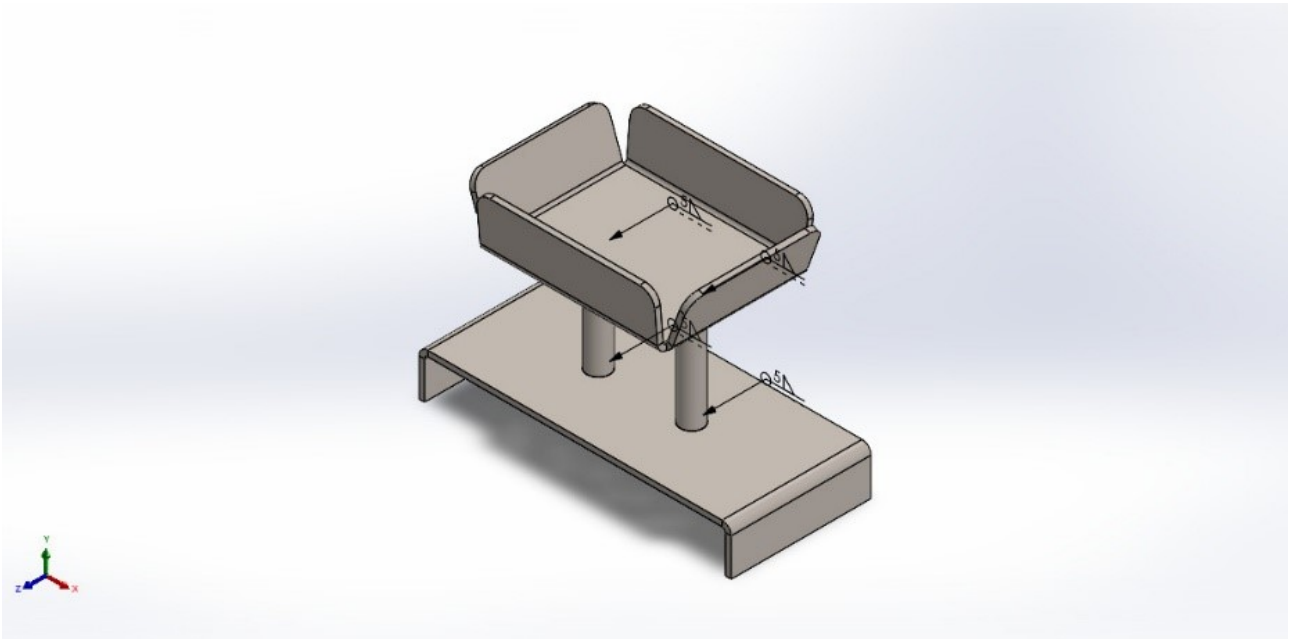


Figura 33: Diseño de bases de fresas

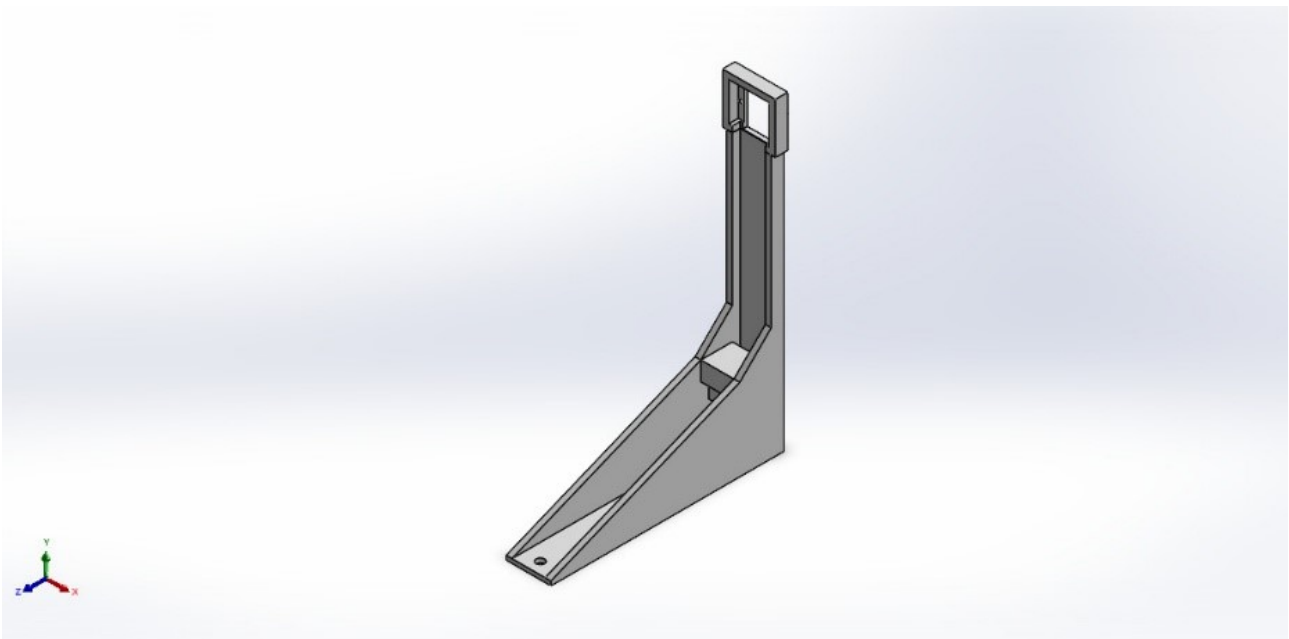


Figura 34: Diseño de base de la cámara y el Grove Vision

ANEXO B
ESTRUCTURA



Figura 35: Bases de fresas



Figura 36: Perfiles de aluminio con cámara instalada.

ANEXO C
PROGRAMACIÓN GOOGLE COLAB

C-A. Código para obtener la matriz de confusión

```
import json
import numpy as np
import matplotlib.pyplot as plt
from sklearn.metrics import ConfusionMatrixDisplay

# --- CONFIGURACIÓN ---
# 1. Asegúrate de que este nombre coincida con el archivo JSON que subiste.
file_name = 'FRESAS.3 (4).json'
# -----

# 1. Cargar el archivo JSON
with open(file_name, 'r') as f:
    data = json.load(f)

# 2. Navegar a los resultados del Test Set para el modelo cuantizado (int8)
# El modelo int8 es el que usarás en el Grove Vision AI V2.
test_data = data['test']['int8']

# 3. Extraer los datos necesarios
cm_values = test_data['confusion_matrix']
class_names = test_data['class_names']
support = test_data['support'] # Número total de ventanas de prueba
```

Figura 37

```

# Convertir la matriz a un array de numpy
cm_array = np.array(cm_values)

# 4. Visualizar la Matriz de Confusión
fig, ax = plt.subplots(figsize=(8, 8))

# Usar ConfusionMatrixDisplay de scikit-learn para un formato estándar
disp = ConfusionMatrixDisplay(confusion_matrix=cm_array, display_labels=class_names)

# Configurar y plotear el gráfico
disp.plot(cmap=plt.cm.Blues, ax=ax, values_format='d')
ax.set_title(f'Matriz de Confusión (Test Set - int8 Quantized)\nTotal Ventanas/Muestras: {support}')
plt.xlabel('Etiqueta Predicha')
plt.ylabel('Etiqueta Real')

# 5. Guardar la imagen (para que se muestre en Colab)
output_image_file = 'matriz_confusion_test_final.png'
plt.savefig(output_image_file)

# 6. Imprimir las métricas clave extraídas
precision = test_data['non_background']['precision']
recall = test_data['non_background']['recall']
f1_score = test_data['non_background']['f1']
accuracy = test_data['classification_report']['accuracy']

```

Figura 38

```

print("\n--- Métricas Clave (int8) ---")
print(f"Precisión (FRESAS MADURAS): {precision:.4f}")
print(f"Recall (FRESAS MADURAS): {recall:.4f}")
print(f"F1 Score (FRESAS MADURAS): {f1_score:.4f}")
print(f"Accuracy General: {accuracy:.4f}")

```

Figura 39

C-B. Código de conversión a formato Vela

```
!pip install ethos-u-vela
!!vela --version

Collecting ethos-u-vela
  Downloading ethos_u_vela-4.5.0-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (10 kB)
Collecting flatbuffers==24.3.25 (from ethos-u-vela)
  Downloading flatbuffers-24.3.25-py2.py3-none-any.whl.metadata (850 bytes)
Requirement already satisfied: numpy>=1.23.0 in /usr/local/lib/python3.12/dist-packages (from ethos-u-vela) (2.0.2)
Downloading ethos_u_vela-4.5.0-cp312-cp312-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (1.9 MB)
 1.9/1.9 MB 75.9 MB/s eta 0:00:00
Downloading flatbuffers-24.3.25-py2.py3-none-any.whl (26 kB)
Installing collected packages: flatbuffers, ethos-u-vela
Attempting uninstall: flatbuffers
  Found existing installation: flatbuffers 25.9.23
  Uninstalling flatbuffers-25.9.23:
    Successfully uninstalled flatbuffers-25.9.23
Successfully installed ethos-u-vela-4.5.0 flatbuffers-24.3.25
['4.5.0']
```

Figura 40

```
!vela FRESAS_PROYECTO1.tflite --accelerator-config ethos-u55-64

... Warning: No configuration file specified. Using a default of ['/usr/local/lib/python3.12/dist-packages/ethosu/config_files/Arm/vela.ini']. Compilation may be invalid or non-optimal.
Warning: No system configuration specified. Using a default of Ethos_U55_High_End_Embedded. Compilation may be invalid or non-optimal.
Warning: No memory mode specified. Using a default of Shared_Sram. Compilation may be invalid or non-optimal.

Network summary for FRESAS_PROYECTO1
Accelerator configuration      Ethos_U55_64
System configuration          Ethos_U55_High_End_Embedded
Memory mode                   Shared_Sram
Accelerator clock              500 MHz
Design peak SRAM bandwidth    3.73 GB/s
Design peak Off-chip Flash bandwidth 0.47 GB/s

Total SRAM used                146.02 KiB
Total Off-chip Flash used      44.33 KiB

CPU operators = 0 (0.0%)
NPU operators = 55 (100.0%)
```

Figura 41

```
... CPU operators = 0 (0.0%)
NPU operators = 55 (100.0%)

Average SRAM bandwidth          1.39 GB/s
Input SRAM bandwidth           0.96 MB/batch
Weight SRAM bandwidth          0.19 MB/batch
Output SRAM bandwidth          0.68 MB/batch
Total SRAM bandwidth           1.83 MB/batch
Total SRAM bandwidth           per input 1.83 MB/inference (batch size 1)

Average Off-chip Flash bandwidth 0.02 GB/s
Input Off-chip Flash bandwidth  0.00 MB/batch
Weight Off-chip Flash bandwidth 0.03 MB/batch
Output Off-chip Flash bandwidth 0.00 MB/batch
Total Off-chip Flash bandwidth  0.03 MB/batch
Total Off-chip Flash bandwidth  per input 0.03 MB/inference (batch size 1)

Neural network macs            6210432 MACs/batch
```

Figura 42

ANEXO D
ADQUISICIÓN DE DATOS



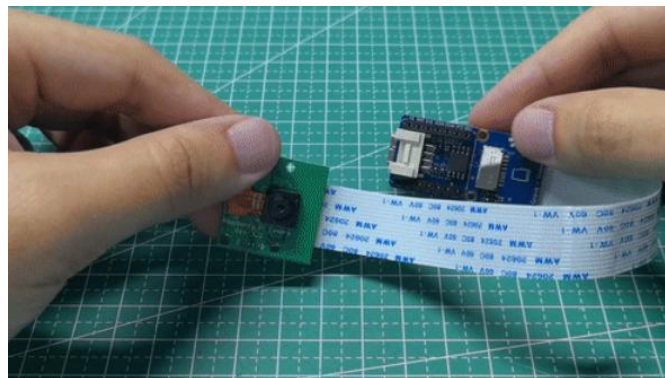
Figura 43: Prueba de luces LED

ANEXO E
GUÍA DE USO GROVE VISION AI V2 MEDIANTE SENSECRAFT Y EDGE IMPULSE

GUIA DE USO GROVE VISION AI V2 MEDIANTE SENSECRAFT Y EDGE IMPULSE

Es un módulo de IA de visión basado en MCU (Microcontrolador de un solo chip), con tecnología Arm Cortex-M55 y Ethos-U55. Es compatible con los frameworks TensorFlow y PyTorch, y con el IDE de Arduino. Con la plataforma de algoritmos de IA SenseCraft, los modelos de aprendizaje automático entrenados se pueden implementar en el sensor sin necesidad de programación. Cuenta con una interfaz CSI estándar para que la conexión desde la cámara al microcontrolador sea inmediata lo cual es crucial para que las imágenes sean analizadas en tiempo real, un micrófono digital integrado y una ranura para tarjetas SD, lo que lo hace ideal para diversos proyectos de visión con IA integrada.

1. Para crear el dataset se utilizo el software o plataforma de desarrollo en la nube SENSECRAFT el cual es el entorno nativo desarrollado por la empresa de SeedStudio. Se hace la toma de datos mediante la conexión de la cámara al microcontrolador Grove Vision AI V2 como se muestra a continuación.



2. Luego ingresamos en el navegador a la pagina de SENSECRAFT mediante el link <https://sensecraft.seed.cc/ai/home> una vez dentro de la plataforma buscamos la opción de Training.

SenseCraft | SenseCraft AI

Home Applications Models Community EN Clox

Start your journey: Deploy a pretrained Model

Step 1: Discover Pretrained Models

Grove Vision AI V2
XIAO ESP32S3 Sense

Trafficcamnet Detection
Deploy SenseCraft AI
This model was trained using the DetectNet-2 architecture in TAO. The training algorithm optimizes the network to...

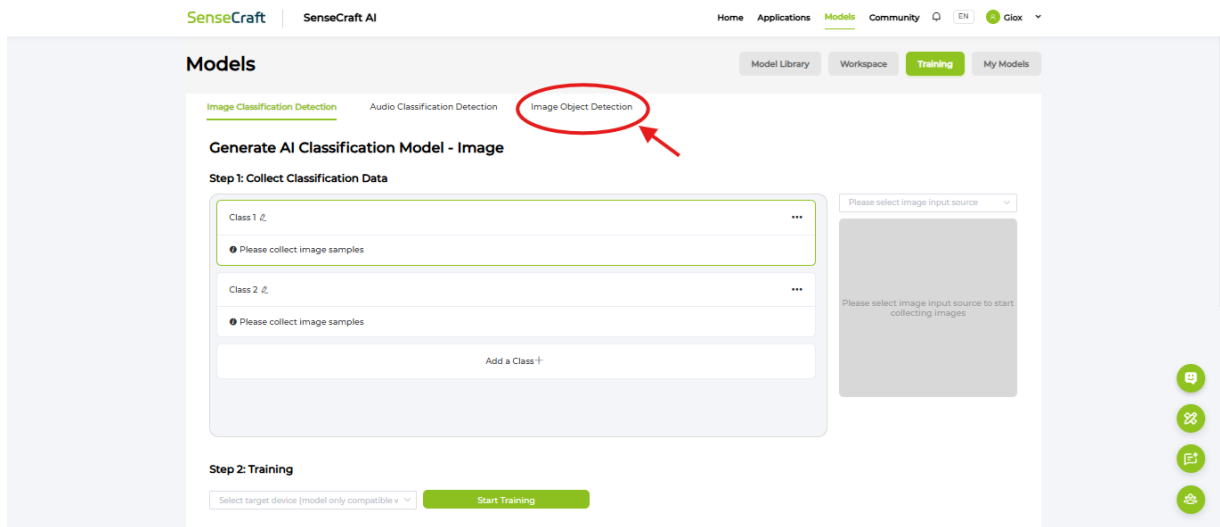
Explore more Pretrained Models → Training →

Step 2: Deploy and Preview Vision

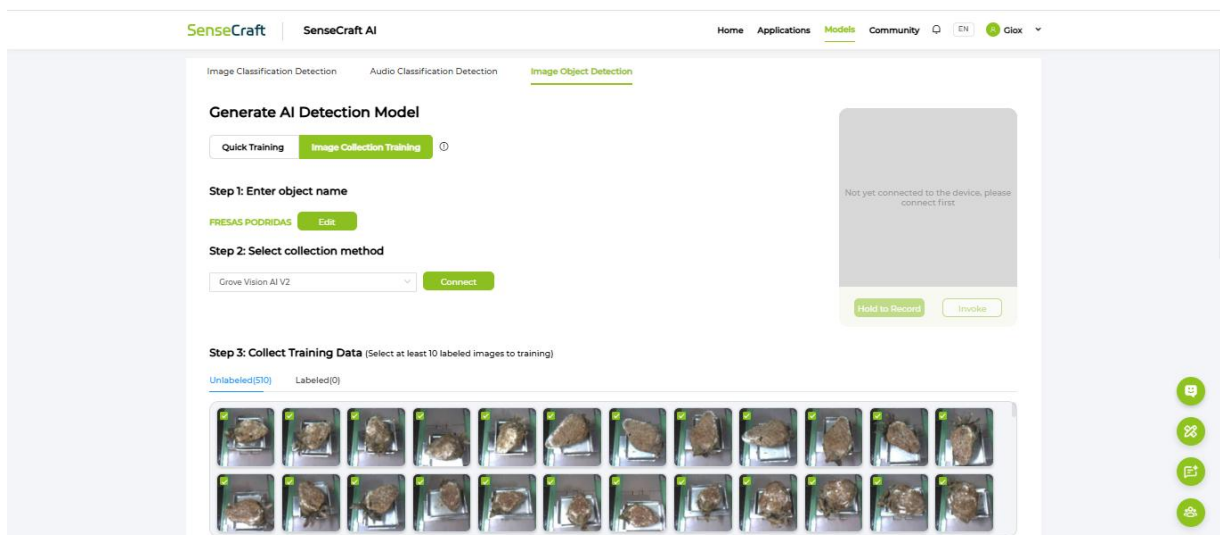
Application Inference Effects

scissors: Prepare and Deploy

- Una vez dentro de Training buscamos la opción de Image Object Detection.

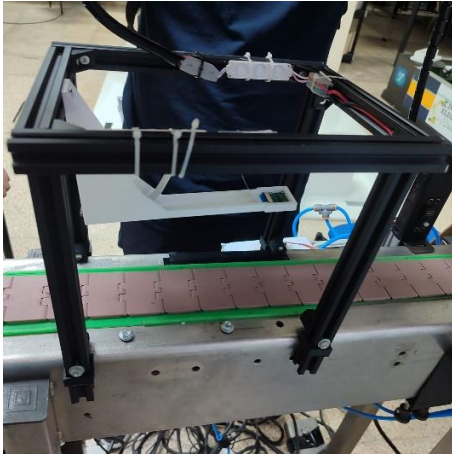


- Luego se abre la interfaz para la toma de datos en nuestro caso se conecta el modulo de Grove Vision AI V2 y se muestra lo siguiente.



- En este apartado tenemos varias opciones como podemos editar el nombre del Dataset que vamos crear tenemos el Step 2 el cual es la opción para conectar el Grove Vision.

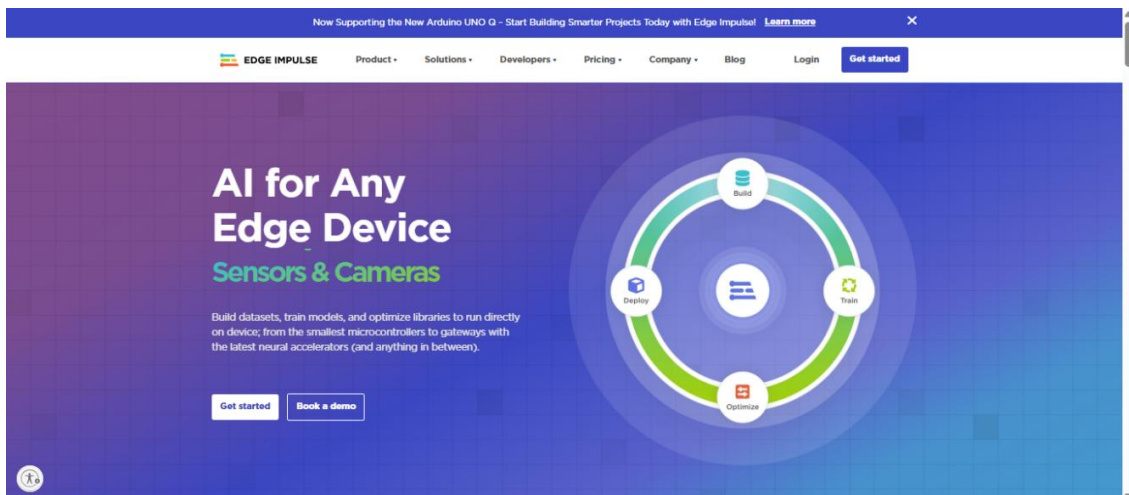
- Las fotos se toman mediante la opción de “Hold to Record” lo cual significa que debemos mantener el clic en esa opción y la toma de imágenes será automática para esto ya tenemos que tener todo listo la base donde ira el módulo de Grove Vision, el entorno el cual será fijo para obtener mejores resultados al momento de entrenar y las luces en caso de ser necesarias como se ve en la imagen.



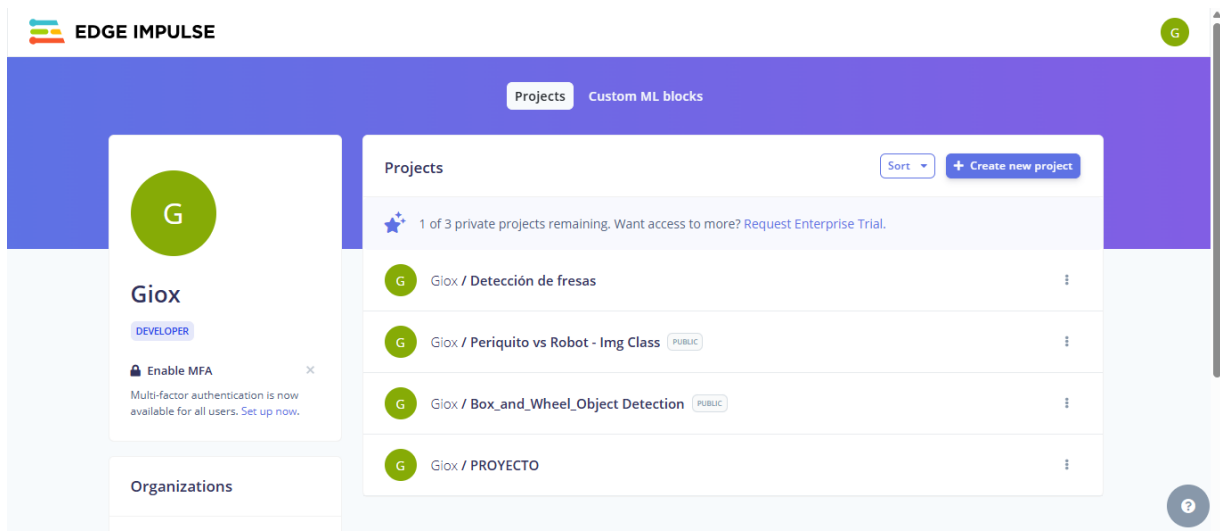
- La toma de imágenes tiene un retraso de aproximadamente 1 segundo es decir cuando se mantiene presionado debemos esperar 1 segundo por foto para que la imagen sea captada por el Grove Vision, una vez que ya están las 500 imágenes por clases, así como se ve en la imagen nos vamos a la opción de exportar y automáticamente se nos descargara una carpeta con todas las imágenes que han sido tomadas.

SenseCraft AI interface showing the training process. The 'Export' button is highlighted with a red circle and arrow.

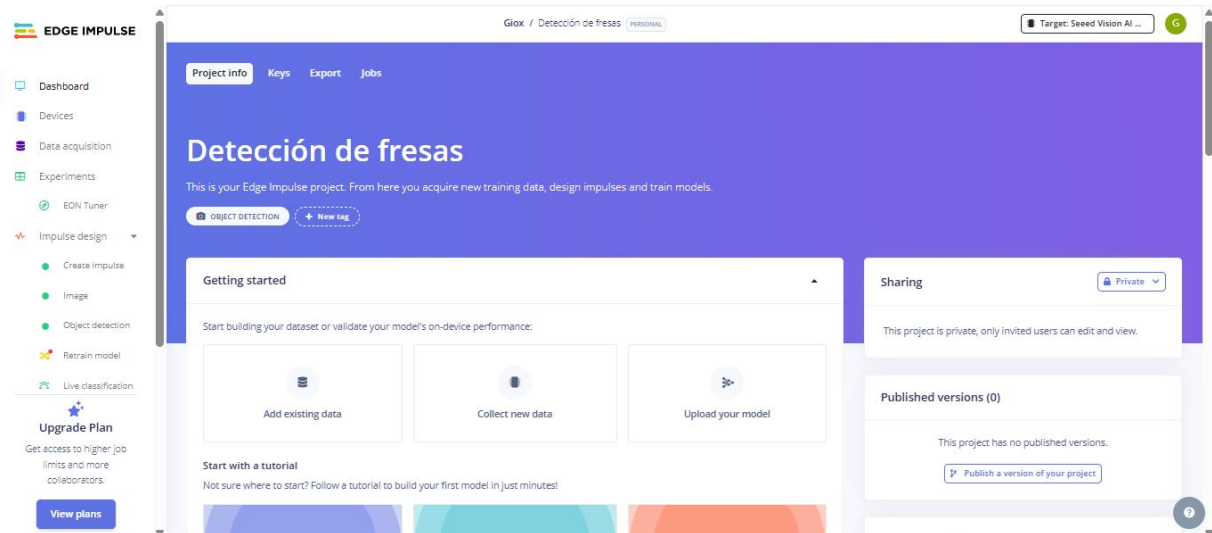
8. Ingresamos a la página de Edge Impulse y es aquí donde haremos el entrenamiento, se decidió utilizar esta plataforma puesto que Sensecraft solo nos deja entrenar bajo una clase cuando queremos proyectos de visión artificial en tiempo real, si no necesitamos que la clasificación sea en tiempo real podemos usar la opción de Image Classification Detection la cual si permite crear varias clases como vimos al principio. Ingresar al siguiente link: <https://www.edgeimpulse.com/>



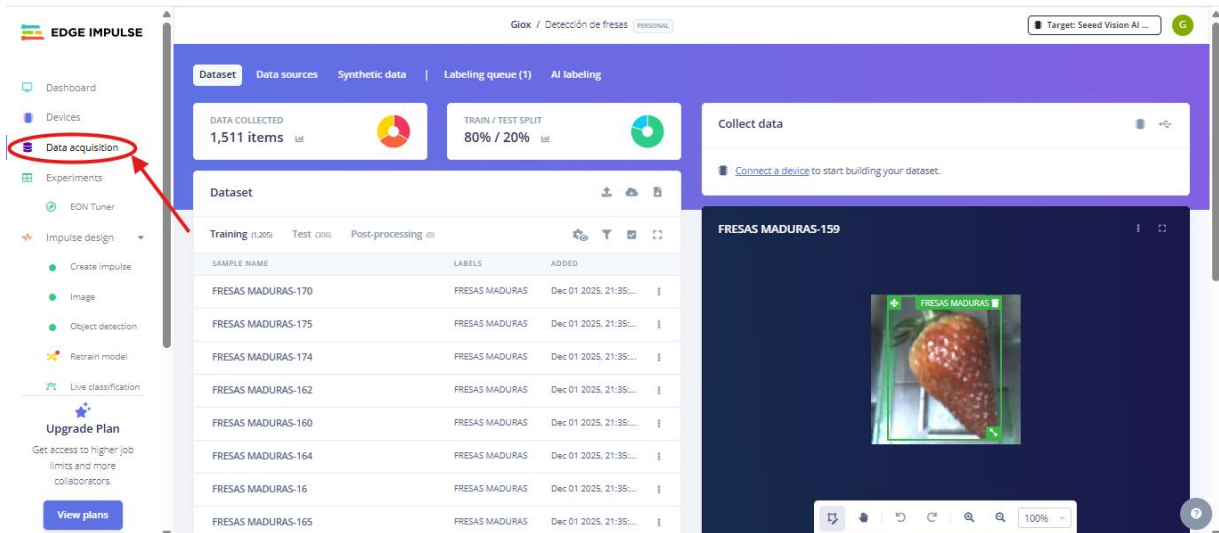
9. Una vez aquí dentro se crea una cuenta para poder crear los proyectos y se verá así como esta en la siguiente imagen.



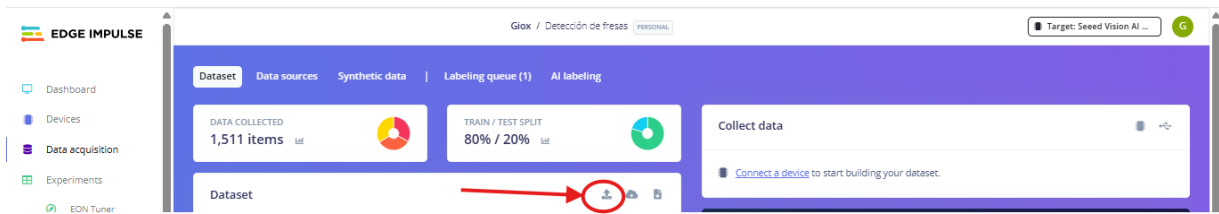
10. Escogemos el proyecto Detección de fresas. En esta interfaz tenemos varias opciones sobre cómo vamos a crear nuestro proyecto y bajo qué condiciones de entrenamiento, tamaño de imagen, entre otros.



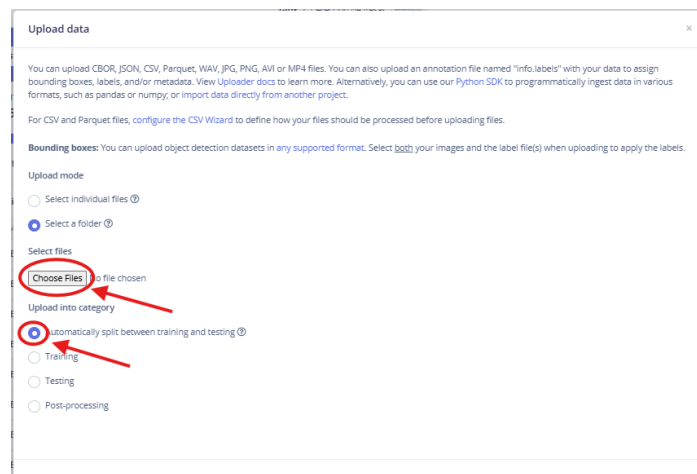
11. Ingresamos a la opción de Data Acquisition and se tienen varias opciones, se muestra a continuación:



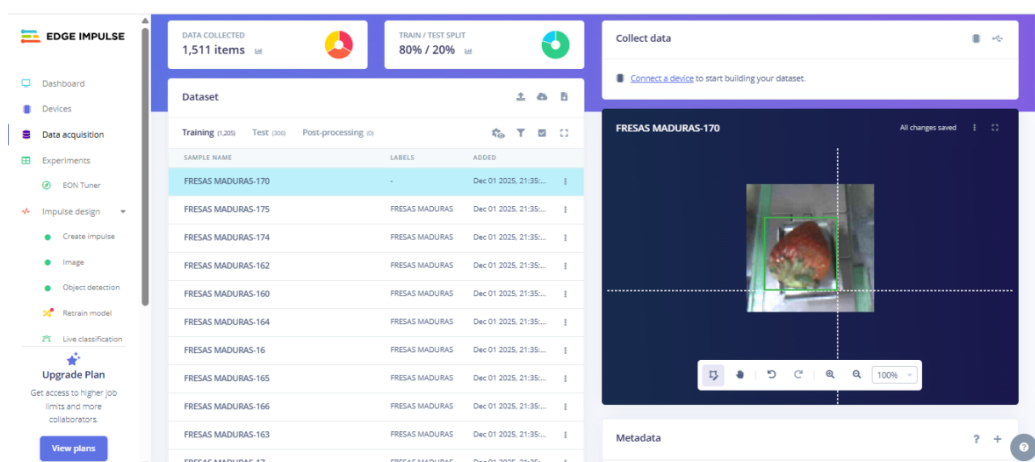
12. Debemos subir las imágenes que previamente ya fueron tomadas en la opción de Upload Data.



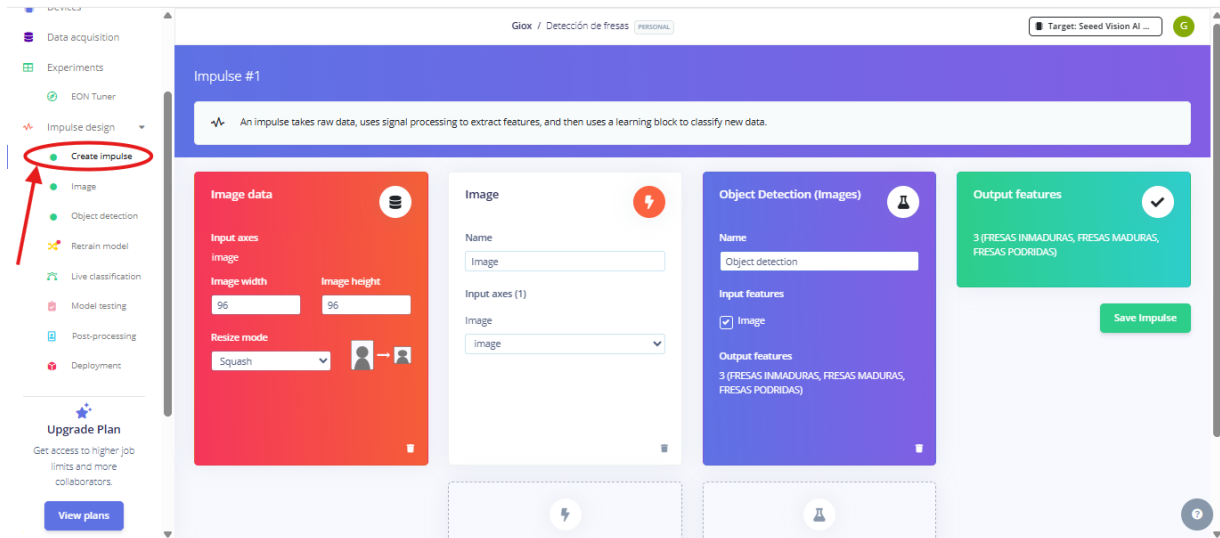
13. Cargamos las imágenes desde la carpeta donde se encuentran las imágenes y en las opciones que se ven sobre la distribución de imágenes por tipo le damos a **Automatically split between training and testing**, esto para mantener el equilibrio entre entrenamiento y prueba el apartado de validación el cual se verá más adelante.



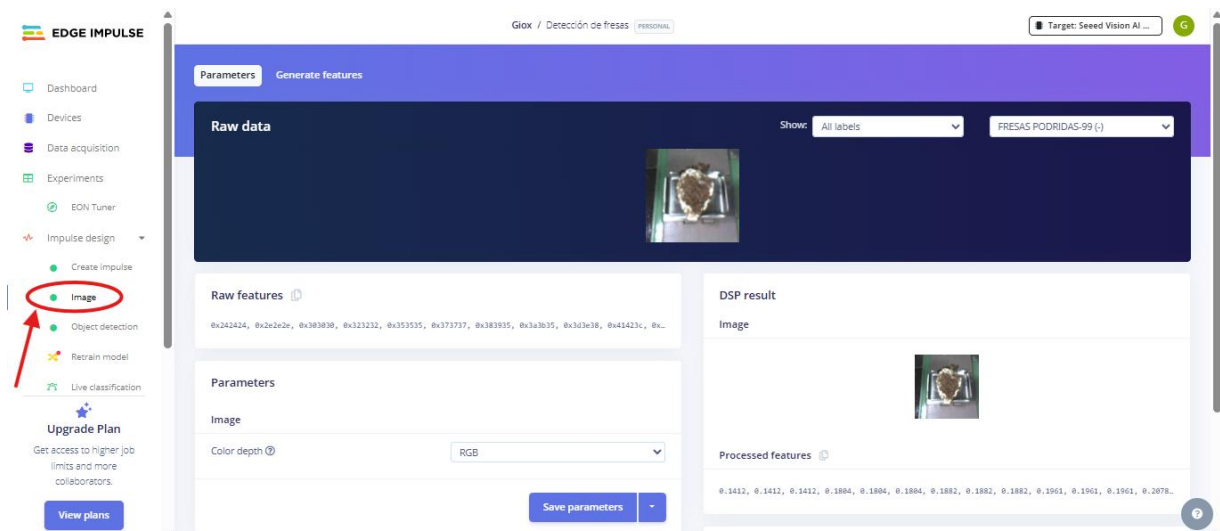
14. Una vez cargadas las imágenes procedemos al etiquetamiento por cada clase que hayamos creado, para este proyecto serán tres, fresas maduras, fresas inmaduras y fresas podridas. El etiquetamiento se hace por cajas delimitadoras y es de forma manual.



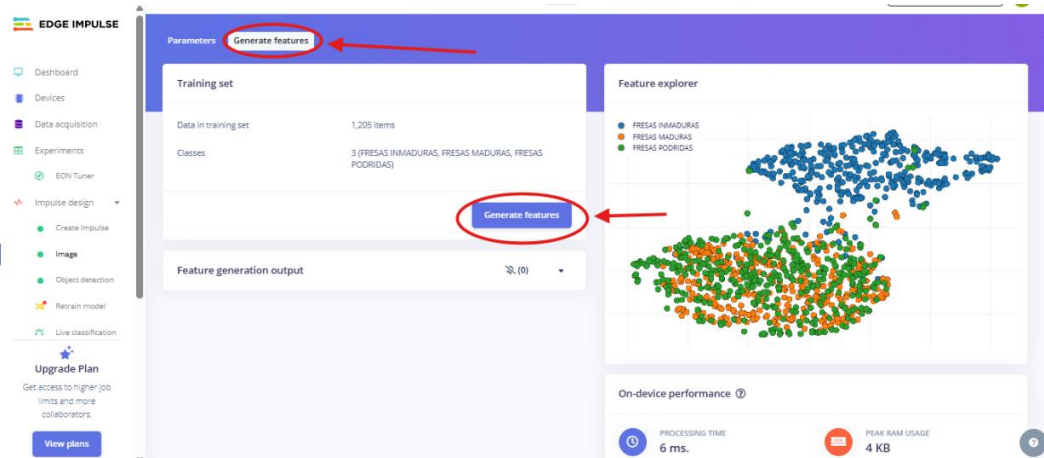
15. Debemos hacer el etiquetamiento por cada parte tanto como entrenamiento como las que están en pruebas (TEST). Luego en la opción de Create Impulse es donde vamos a definir las características por imagen del entrenamiento. Se escogió una altura y anchura por imagen de 96x96 por las limitaciones de la cámara, otra técnica usada es la de reducción por anchura la cual cambiara el tamaño de la imagen e ignorara la relación de aspecto y como características de salida se tiene las tres clases a detectar.



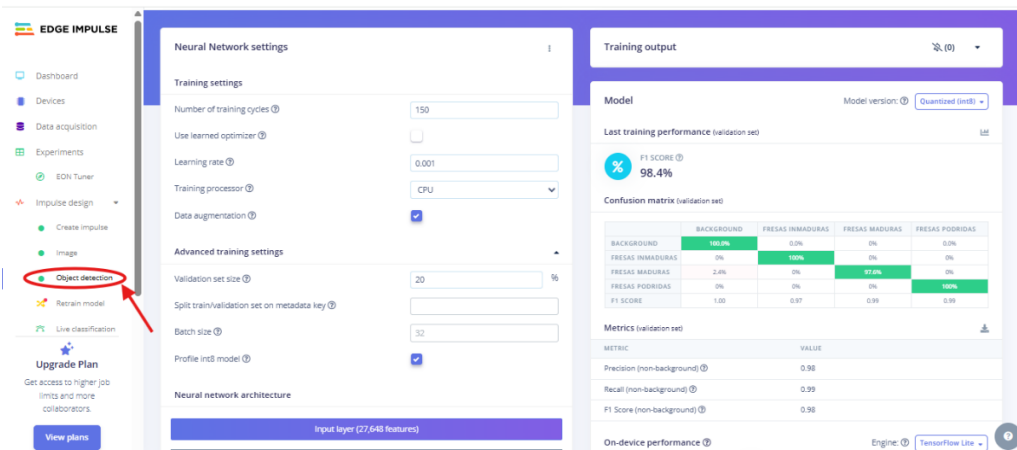
16. Luego vamos a la opción de Image en la cual podremos definir la profundidad de color en este caso se necesita RGB ya que las características son por color no por su forma en su mayoría, esto incrementara el tamaño de RAM en el proceso, pero es soportable para el Grove Vision.



17. Luego se ingresa en la parte de Generate features y de nuevo en la parte de abajo Generate features su función principal es el análisis de separabilidad que tanto el modelo distingue entre una clase y otra por ejemplo en este caso los puntos azules están bien separados de los demás por lo cual no va a confundirse, pero los puntos verdes y naranjas probablemente si, pero esto solo nos sirve para saber que puede pasar al momento de hacer el entrenamiento.



18. En la siguiente sección se tienen las diferentes opciones de cómo va a estudiar la red neuronal, se utilizo 150 épocas de entrenamiento porque la red se estaba confundiendo en dos clases a pesar de eso se tiene una precisión del 98.4% por clase del apartado de validación(Validation) lo cual nos indica que durante el entrenamiento la clasificación esta siendo correcta, la tasa de aprendizaje (Learning rate) es del 0.001 controla qué tan grandes son los "pasos" que da el modelo para corregir sus errores se definió este numero para que el aprendizaje sea gradual y estable, se utilizo la CPU como procesador y nos indica que el entrenamiento se está realizando mediante procesadores centrales para procesar los datos, el data augmentation es muy útil lo que hace es que crea variaciones de las fotos (las gira, les cambia el brillo o les hace zoom) de forma artificial. Esto ayuda a que el modelo sea más robusto y no se confunda si la luz cambia o la fresa está en otra posición y finalmente se tiene la opción "Profile int8 model" calcula cómo funcionará el modelo una vez optimizado (cuantizado) para hardware limitado.



19. En modelo de entrenamiento se escogió el de FOMO dado su optimización y versatilidad al momento de realizar el entrenamiento, este modelo está basado en círculos centroides, pero en Sensecraft se vera reflejado como cajas centroides el cual esta basado en el mismo principio, en las siguientes imágenes se vera el modelo y como ingresar específicamente el de FOMO.

The screenshot shows the EDGE IMPULSE training interface. On the left is a navigation menu with options like Dashboard, Devices, Data acquisition, Experiments, and Impulse design. The main area is titled 'Advanced training settings' and includes fields for 'Validation set size' (20%), 'Batch size' (32), and 'Profile int8 model' (checked). Below this is the 'Neural network architecture' section, which shows an 'Input layer (27,648 features)' and an 'Output layer (3 classes)'. The selected model is 'FOMO (Faster Objects, More Objects) MobileNetV2 0.35'. A red oval highlights the 'Choose a different model' link, with a red arrow pointing to it. To the right, a 'Confusion matrix (validation set)' is displayed, showing high accuracy for the classes: BACKGROUND (100.0%), FRESAS INMADURAS (100%), FRESAS MADURAS (97.6%), and FRESAS PODRIDAS (100%). Below the matrix, 'Metrics (validation set)' are listed: Precision (0.98), Recall (0.99), and F1 Score (0.98). 'On-device performance' is also shown, with values for Inference time (N/A), Peak RAM usage (283.1K), and Flash usage (110.7K).

This screenshot shows the 'Choose a different model' dialog box. It lists several models with their descriptions and 'Add' buttons. The 'FOMO (Faster Objects, More Objects) MobileNetV2 0.35' model is circled in red, with a red arrow pointing to it. The dialog also includes a 'Did you know?' section and a 'Want access to all learning blocks? Upgrade now.' prompt at the bottom. In the background, the training interface is visible, showing the same confusion matrix and metrics as in the previous screenshot.

20. Finalmente, para ver los resultados del apartado de prueba (TEST) se tomaron 100 imágenes por clases las cuales no han sido vistas ni usadas para el entrenamiento, se da click en la siguiente opción.

The screenshot shows the Edge Impulse interface. On the left sidebar, the 'Model testing' option is highlighted with a red circle and an arrow. The main area is divided into two panels. The left panel, titled 'Test data', contains a table with the following data:

SAMPLE NAME	EXPECTED OUTCOME	F1 SCORE	RESULT
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	
FRESAS PODRIDA...	FRESAS PODRIDAS	100%	

The right panel, titled 'Model testing output', shows the following results:

- Results:** ACCURACY 97.71%
- Metrics for Object detection:**

METRIC	VALUE
Precision (non-background)	0.98
Recall (non-background)	1.00
F1 Score (non-background)	0.99
- Feature explorer:** A scatter plot showing data points for 'object_detection - correct' (yellow) and 'object_detection - incorrect' (red).

21. Esta opción es muy importante para ver las métricas de estudio de la red neuronal pues indica que errores tuvo al momento de clasificar y que porcentajes de exactitud tiene, en este caso los puntos amarillos son las imágenes correctamente clasificadas y los puntos rojos son las imágenes mal clasificadas o no las clasifico, se tiene una exactitud del 97.71% y 5 errores por imágenes mal clasificadas.

22. Luego en el apartado de Dashboard se descarga el archivo con denominación "TensorFlow Lite (int8 quantized)" el cual servirá para hacer la conversión a un modelo específico de Tensor Flow Lite para cargarlo en el Grove Vision mediante Sensecraft.

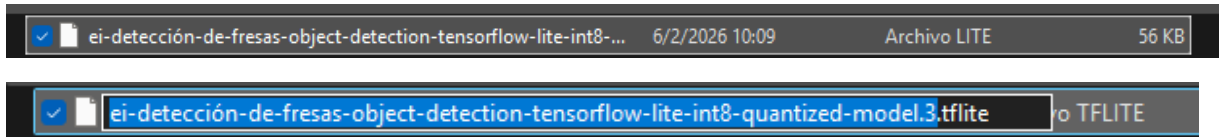
The screenshot shows the Edge Impulse Dashboard. On the left sidebar, the 'Dashboard' option is highlighted with a red circle and an arrow. The main area is divided into two panels. The left panel, titled 'Download block output', contains a table with the following data:

TITLE	TYPE	SIZE
Image training data	NPY file	1205 windows
Image training labels	JSON file	1205 windows
Image testing data	NPY file	306 windows
Image testing labels	JSON file	306 windows
Object detection model	TensorFlow Lite (float32)	83 KB
Object detection model	TensorFlow Lite (int8 quantized)	55 KB
Object detection model	TensorBoard logs	37 MB
Object detection model	TensorFlow SavedModel	108 KB
Object detection model	Keras h5 model	90 KB
Object detection model	Model evaluation metrics (JSON file)	-

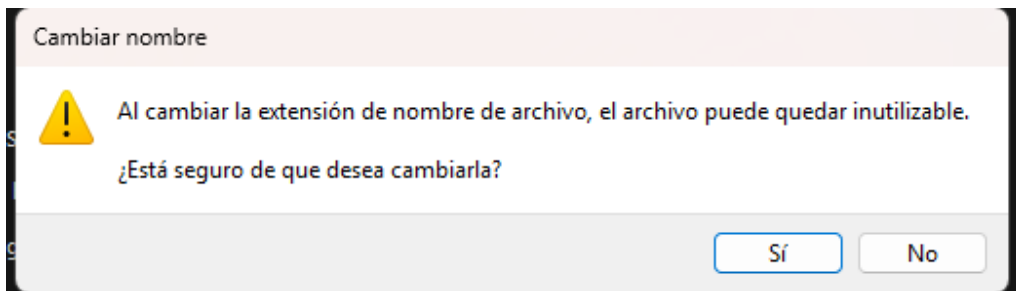
The 'Object detection model' row with 'TensorFlow Lite (int8 quantized)' is highlighted with a red circle and an arrow. The right panel shows project information:

- Collaborators (1/3):** Giox (owner)
- Summary:** DEVICES CONNECTED: 0; DATA COLLECTED: 1,511 items
- Project info:** Project ID: 837844; Labeling method: Bounding boxes (object c)

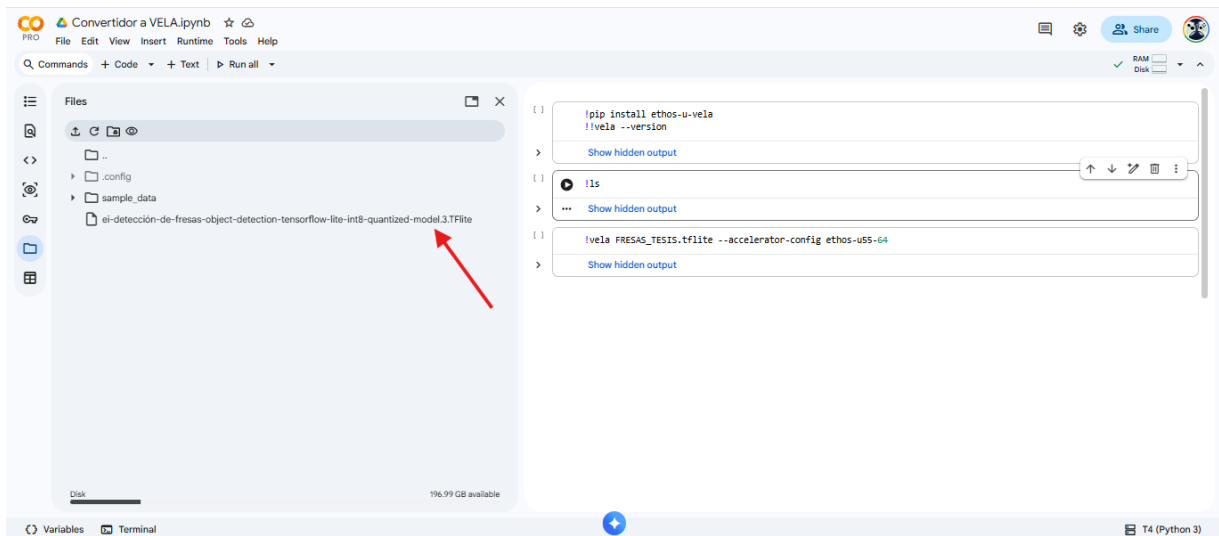
23. Una vez descargado este archivo se le cambia la denominación de .lite a .tflite como se ve a continuación.



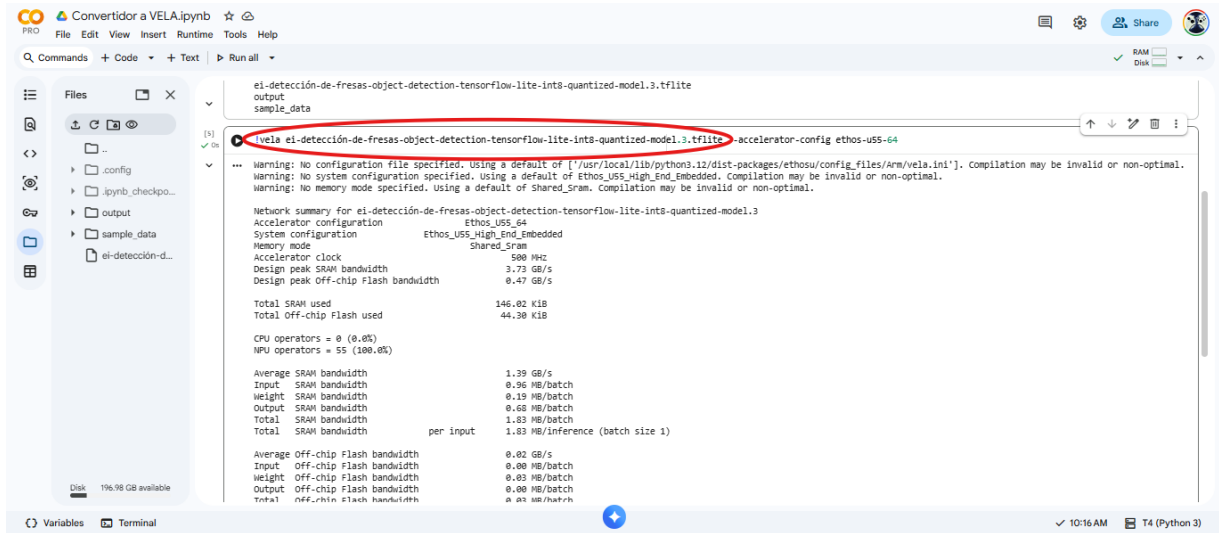
24. Le damos sí.



25. Se abre en Google Colab el código de convertidor a Vela.ipynb y se agrega el archivo al espacio compartido de Colab.



26. Se ejecuta el código y se coloca el nombre del archivo a convertir en la tercera línea de código en Colab.



```
ei-detección-de-fresas-object-detection-tensorflow-lite-int8-quantized-model.3.tflite
output
sample_data

[5]
[5]
❌ vela ei-detección-de-fresas-object-detection-tensorflow-lite-int8-quantized-model.3.tflite -accelerator-config ethos-u55-64
Warning: No configuration file specified. Using a default of ['/usr/local/lib/python3.12/dist-packages/ethosu/config_files/Arm/vela.ini']. Compilation may be invalid or non-optimal.
Warning: No system configuration specified. Using a default of Ethos_U55_High_End_Embedded. Compilation may be invalid or non-optimal.
Warning: No memory mode specified. Using a default of shared_sram. Compilation may be invalid or non-optimal.

Network summary for ei-detección-de-fresas-object-detection-tensorflow-lite-int8-quantized-model.3
Accelerator configuration      Ethos_U55_64
System configuration           Ethos_U55_High_End_Embedded
Memory mode                   Shared_SRAM
Accelerator clock              500 MHz
Design peak SRAM bandwidth    3.73 GB/s
Design peak Off-chip Flash bandwidth 0.47 GB/s

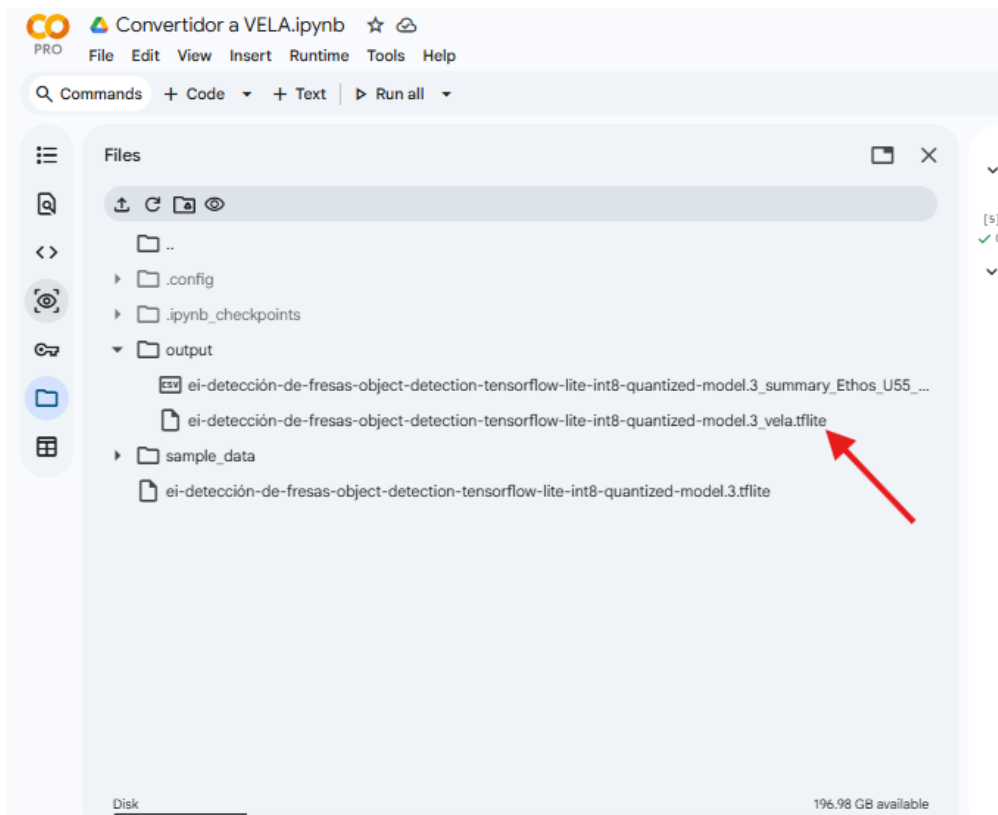
Total SRAM used                146.02 KiB
Total Off-chip Flash used     44.30 KiB

CPU operators = 0 (0.0%)
NPU operators = 55 (100.0%)

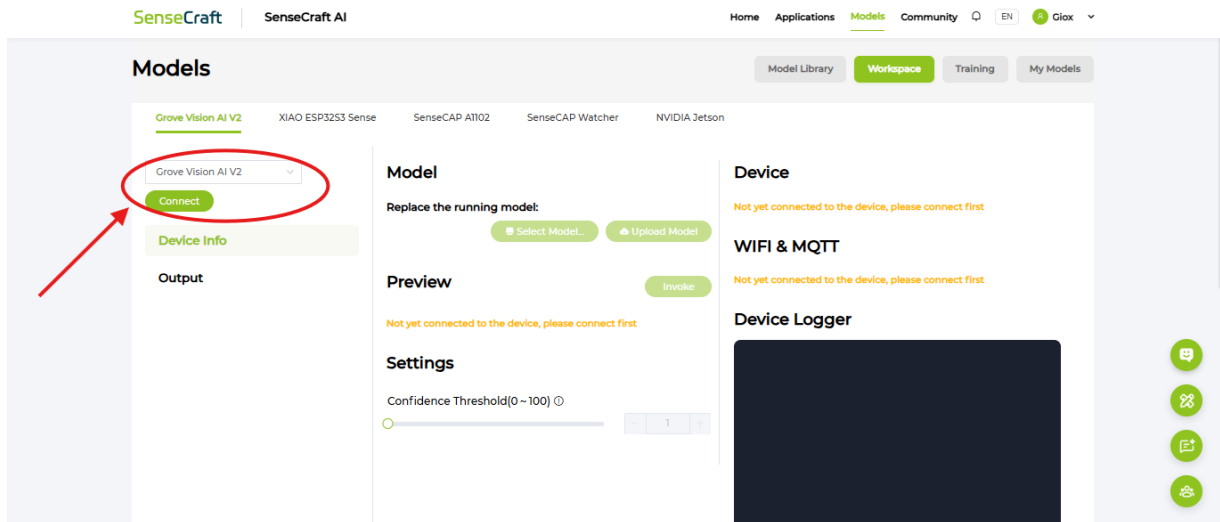
Average SRAM bandwidth        1.39 GB/s
Input SRAM bandwidth          0.96 MB/batch
Height SRAM bandwidth         0.19 MB/batch
Output SRAM bandwidth         0.68 MB/batch
Total SRAM bandwidth          1.83 MB/batch
Total SRAM bandwidth          per input 1.83 MB/inference (batch size 1)

Average Off-chip Flash bandwidth 0.02 GB/s
Input Off-chip Flash bandwidth 0.00 MB/batch
Height Off-chip Flash bandwidth 0.03 MB/batch
Output Off-chip Flash bandwidth 0.00 MB/batch
Total Off-chip Flash bandwidth 0.01 MB/batch
```

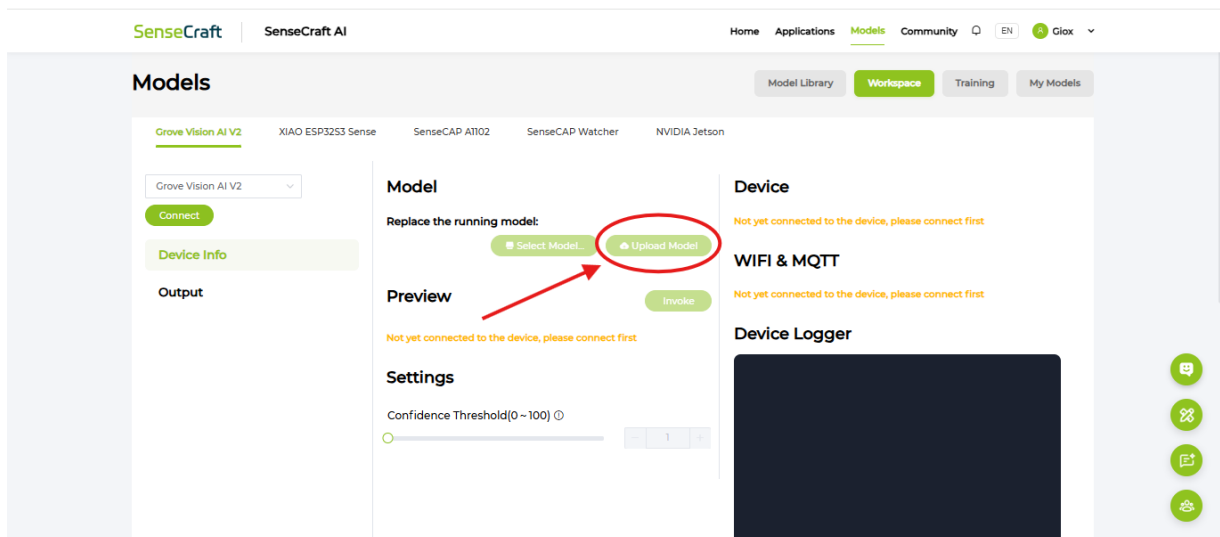
27. Se descarga el archivo con la denominación “vela.tflite” dentro de la carpeta compartida en Colab.



28. Luego en Sensecraft se abre la opción de “Workspace” en la cual se puede elegir el dispositivo a flashear en este caso el Grove Vision AI V2.



29. Luego click en la opción “Upload Model” allí se cargará el archivo previamente convertido en Colab



30. Listo ya se podrá realizar proyectos de visión artificial en tiempo real.

REFERENCIAS

- [1] M. A. Pérez, J. R. López y L. F. Sánchez, «Desafíos y oportunidades en la exportación de frutos y vegetales: Control de calidad y automatización en Ecuador,» 2025.
- [2] M. I. Ashqui Balseca, B. P. Aucatoma Matias y E. P. Córdova Córdova, «Sistema de control de calidad de cultivo de fruta de temporada para etapa de pre cosecha empleando robótica aérea con planificación de trayectorias y visión artificial,» 2024.
- [3] T. S. Roy, N. Imtiaz, R. Chakraborty y B. C. Kundu, «Applying biochar with different form of nitrogen: Be a good agricultural practice for better yield and processing quality of potato,» 2022.
- [4] E. G. Riofrío Vega, F. D. Galarza Calle, C. M. Morales Macas y J. B. Calva Jara, «Evaluación de la adaptación y producción de cultivares de fresa con efecto microtúnel en el Cantón Chilla, El Oro, Ecuador,» 2023.
- [5] C. F. Abad-Abad, M. C. Álvarez Cedeño y J. González, «Efecto de la cubierta (microtúnel) en la productividad de dos variedades de fresa (*Fragaria vesca*) en el sector Cajanuma, cantón Loja,» 2020.
- [6] J. E. Criollo Otavalo, «Estudio del aprovechamiento industrial del cultivo de fresas,» 2025.
- [7] A. Alzarooni, E. Iqbal, S. U. Khan, S. Javed, B. Moyo e Y. Abdulrahman, «Anomaly Detection for Industrial Applications, Its Challenges, Solutions, and Future Directions: A Review,» 2025.
- [8] D. Martin, S. Heinzl, J. Kunze von Bischhoffshausen y N. Kühn, «Deep Learning Strategies for Industrial Surface Defect Detection Systems,» 2021.
- [9] D. J. Mendoza Chipantasi y J. E. Salazar Castillo, «Sistema de clasificación y control de calidad en un proceso de producción industrial usando visión artificial,» 2019.
- [10] V. A. Estudio Técnico-Económico, «La mano de obra representa cerca del 50 % de los costos totales en la producción frutal,» 2024.
- [11] L. A. Gutiérrez Sánchez, «Sistema inteligente de selección de cosecha en cultivo de pitahaya mediante visión artificial,» 2024.
- [12] I. Rojas Santelices, S. Cano, F. Moreira y Peña Fritz, «Artificial Vision Systems for Fruit Inspection and Classification: Systematic Literature Review,» 2025.
- [13] J. Feng, Q. Yang, H. Tian, Z. Wang, S. Tian y H. Xu, «Promising real-time fruit and vegetable quality detection technologies applicable to manipulator picking process,» 2024.
- [14] A. Bhargava y A. Bansal, «Fruits and vegetables quality evaluation using computer vision: A review,» 2021.
- [15] P. Pathmanaban, B. K. Gnanavel y S. S. Anandan, «Quality Assessment of Fruits and Vegetables Based on Spatially Resolved Spectroscopy: A Review,» 2024.
- [16] Authors Various, «Development of an optimally designed real-time automatic citrus fruit grading–sorting machine leveraging computer vision-based adaptive deep learning model,» 2023.
- [17] Z. Gao, Y. Shao, G. Xuan, Y. Wang, Y. Liu y X. Han, «Real-time hyperspectral imaging for the in-field estimation of strawberry ripeness with deep learning,» 2020.
- [18] J. Le Louëdec y G. Cielniak, «3D shape sensing and deep learning-based segmentation of strawberries,» 2021.
- [19] J. López, S. Fernández y A. Ruiz, «Implementación de sistemas automatizados para la cosecha y manejo postcosecha de fresas,» 2024.
- [20] A. K. A. Abdul-Razak Alhassan Gamani I. Arhin, «Performance Evaluation of YOLOv8 Model Configurations, for Instance Segmentation of Strawberry Fruit Development Stages in an Open Field Environment,» 2024.

- [21] Food y A. O. of the United Nations. Oficina Regional para América Latina y el Caribe, «Manual para el mejoramiento del manejo poscosecha de frutas y hortalizas,» FAO, Oficina Regional para América Latina y el Caribe, inf. téc., 1987, Parte I: Cosecha y Empaque; traducido del original en inglés. dirección: <https://www.fao.org/4/x5055s/x5055S05.htm>.
- [22] Departamento Técnico de Cultifort, *¿Cómo conseguir una óptima maduración del fresón?* <https://www.cultifort.com/optima-maduracion-freson/>, Cultinews de Cultifort sobre manejo agronómico en el cultivo de fresón, 2024.
- [23] Codificando Bits, *¿Qué son las Redes Convolucionales? — Introducción a las Redes Convolucionales*, <https://codificandobits.com/blog/redes-convolucionales-introduccion/>, Publicado el 23/03/2019, 2019.
- [24] D. C. Montgomery y G. C. Runger, *Applied Statistics and Probability for Engineers*. Wiley, 2014.
- [25] G. James, D. Witten, T. Hastie y R. Tibshirani, *An Introduction to Statistical Learning*. Springer, 2013.
- [26] U. P. Salesiana, *Módulo de Cámara OV5647-62 FOV*, <https://dspace.ups.edu.ec/bitstream/123456789/30353/1/UPS-GT006353.pdf>, Documento institucional, s.f.
- [27] Seeed Studio, *Using Grove Vision AI V2 with AT Commands*, https://wiki.seeedstudio.com/grove_vision_ai_v2_at/, Consultado el 19 de noviembre de 2025, 2025.
- [28] *Raspberry Pi Documentation*, Accessed: 2026-02-10, Raspberry Pi Foundation, 2023. dirección: <https://www.raspberrypi.com/documentation/>.
- [29] *ESP32-CAM Technical Reference Manual*, Accessed: 2026-02-10, Espressif Systems, 2023. dirección: <https://www.espressif.com/en/products/socs/esp32>.
- [30] *NVIDIA Jetson Nano Developer Kit User Guide*, Accessed: 2026-02-10, NVIDIA Corporation, 2023. dirección: <https://developer.nvidia.com/embedded/jetson-nano-developer-kit>.
- [31] Robot Italy, *OV5647 62° FOV Camera Module for Raspberry Pi 3B/4B (Fisheye Lens, 2592×1944)*, https://robot-italy.com/products/114110127_ov5647-62-fov-camera-module-for-raspberry-pi-3b-4b-fisheye-lens-2592-x-1944-suitable-for-large-or-night-landscape-surveillance, Módulo de cámara con sensor OV5647, adecuado para paisajes amplios o vigilancia nocturna, 2025.
- [32] Seeed Studio, *SenseCraft AI Platform*, <https://wiki.seeedstudio.com/sensecraft-ai/overview/>, Plataforma basada en web para entrenar y desplegar modelos IA embebidos, 2025.
- [33] EdgeImpulse, *Edge Impulse Documentation: Machine Learning for Edge Devices*, Consultado en 2025, 2024. dirección: <https://docs.edgeimpulse.com/>.
- [34] Google. «Google Colaboratory.» Accedido: 2026-01-13. dirección: <https://colab.research.google.com/>.
- [35] Edge Impulse, *FOMO: Fast Objects, More Objects*, Technical documentation, 2022. dirección: <https://docs.edgeimpulse.com>.
- [36] W. Liu, D. Anguelov, D. Erhan, C. Szegedy y S. Reed, «Small Object Detection in Computer Vision: A Survey,» *Pattern Recognition*, vol. 112, pág. 107799, 2021.
- [37] T.-Y. Lin, P. Goyal, R. Girshick, K. He y P. Dollár, «Focal Loss for Dense Object Detection,» *Proceedings of the IEEE International Conference on Computer Vision*, 2017.
- [38] Y. Chen, T.-J. Yang, Y. Zhang, J. Chen y J. Emer, «Efficient Neural Networks for Embedded Vision Applications,» *ACM Computing Surveys*, vol. 53, n.º 6, 2020.
- [39] J. Redmon, S. Divvala, R. Girshick y A. Farhadi, «You Only Look Once: Unified, Real-Time Object Detection,» *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

- [40] A. Bochkovskiy, C.-Y. Wang y H.-Y. M. Liao, «YOLOv4: Optimal Speed and Accuracy of Object Detection,» *arXiv preprint arXiv:2004.10934*, 2020.
- [41] M. Sokolova y G. Lapalme, «A Systematic Analysis of Performance Measures for Classification Tasks,» *Information Processing & Management*, vol. 45, n.º 4, págs. 427-437, 2009.
- [42] D. M. W. Powers, «Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness and Correlation,» *Journal of Machine Learning Technologies*, vol. 2, n.º 1, págs. 37-63, 2011.
- [43] T. Fawcett, «An Introduction to ROC Analysis,» *Pattern Recognition Letters*, vol. 27, n.º 8, págs. 861-874, 2006.