Sistema de detección en tiempo real de maracuyá defectuoso mediante reconocimiento visual con YOLOv5

|  |  |  |
| --- | --- | --- |
| Renato Sánchez | José Martínez | Luis Salazar |
| Facultad de Ingeniería de Sistemas | Facultad de Ingeniería de Sistemas | Facultad de Ingeniería de Sistemas |
| Universidad Peruana de Ciencias Aplicadas | Universidad Peruana de Ciencias Aplicadas | Universidad Peruana de Ciencias Aplicadas |
| Lima, Perú | Lima, Perú | Lima, Perú |
| u202020105@upc.edu.pe | u201923468@upc.edu.pe | pcislsal@upc.edu.pe |

**Resumen**

La inteligencia artificial es muy importante en los procesos de producción, ya que permite analizar datos y optimizar las labores de planificación y control con el fin de aumentar la eficiencia y reducir los costos. En este sentido, se utilizó el modelo YOLOv5 en una fábrica de producción de jugos de maracuyá, destacándose la efectividad del modelo en la detección y clasificación de la calidad de la fruta considerando la variabilidad inherente en las características de la maracuyá. El modelo fue entrenado con aproximadamente 500 imágenes y se logró una precisión del 92.869% y una sensibilidad del 91.224%. Los desafíos clave incluyeron la necesidad de equilibrar datos de las diferentes categorías de defectos que podrían estar presentes en la fruta y la alta variabilidad de sus características. Para abordar estos desafíos, el estudio aumentó el número de imágenes para categorías específicas con el fin de mejorar la precisión. El trabajo futuro se centrará en optimizar aún más el sistema mediante la ampliación del tamaño del conjunto de datos y el perfeccionamiento de los algoritmos de aprendizaje para mantener el tiempo de procesamiento dentro de los 30 minutos por sesión. Esta investigación proporciona un marco sólido para la integración de soluciones tecnológicas avanzadas en la producción de jugo, mejorando significativamente la sostenibilidad y eficiencia del procesamiento de maracuyá. Los resultados indican una dirección prometedora para la aplicación del aprendizaje profundo en el control de calidad industrial, ofreciendo métodos escalables y efectivos para mejorar la consistencia del producto y reducir el desperdicio.

**Palabras clave**

Deep learning, reconocimiento visual, maracuyá, clasificación de frutas en tiempo real, reconocimiento visual y maracuyá.

**Abstract**

Artificial intelligence is very important in production processes, as it allows data analysis and optimisation of planning and control tasks in order to increase efficiency and reduce costs. In this sense, the YOLOv5 model was used in a passion fruit juice production factory, highlighting the effectiveness of the model in detecting and classifying the quality of the fruit considering the inherent variability in the characteristics of passion fruit. The model was trained on approximately 500 images and achieved an accuracy of 92.869% and a sensitivity of 91.224%. Key challenges included the need to balance data from the different categories of defects that could be present in the fruit and the high variability of fruit characteristics. To address these challenges, the study increased the number of images for specific categories to improve accuracy. Future work will focus on further optimising the system by expanding the size of the dataset and refining the learning algorithms to keep processing time within 30 minutes per session. This research provides a solid framework for the integration of advanced techno-logical solutions in juice production, significantly improving the sustainability and efficiency of passion fruit processing. The results indicate a promising direction for the application of deep learning in industrial quality control, offering scalable and effective methods to improve product consistency and reduce waste.

**Keywords**

Deep learning, visual recognition, passion fruit, real-time fruit classification, visual recognition and passion fruit.

1. Introducción

La maracuyá es una fruta de relevancia económica y nutricional que enfrenta retos significativos en su clasificación debido a la variabilidad de sus características. En Perú, las exportaciones de zumo de maracuyá han mostrado un crecimiento sustancial, alcanzando 25,693 Miles US$ FOB en 2023 [2]. No obstante, la clasificación manual de la fruta puede conllevar a errores que terminan afectando la calidad y el valor económico del producto [6], [1]. Esto también termina afectando la cantidad final de productos para la exportación, ocasionando pérdidas financieras que podrían haberse evitado.

La clasificación manual ineficiente resulta en desperdicio y en una pérdida económica, subrayando la necesidad de métodos tecnológicamente más avanzados para la clasificación de la maracuyá. Investigaciones como las de Rashidi et al. [3] y Omid et al. [4] han resaltado el potencial de la aplicación de tecnologías de procesamiento de imágenes y modelos como el Deep Learning para mejorar la precisión en la clasificación de frutas, lo que es crucial dada la creciente importancia del maracuyá en el mercado global en los últimos años [1], [2].

Aunque los estudios existentes en el uso del Deep Learning para la clasificación de frutas, como el de Xiaoting y Xueying [5], han aportado avances significativos, se identifican limitaciones en términos de precisión y adaptabilidad en diferentes condiciones de producción. El estudio de Liu [6] sobre la clasificación de la calidad del maracuyá utilizando ATC-YOLOv5 destaca por su precisión del 95.36% y tiempo de detección medio de 3.2ms, mejorando la precisión y la velocidad de detección en comparación con la de otros modelos realizados.

Este estudio tiene como objetivo desarrollar un sistema de clasificación basado en tecnologías de Deep Learning (como YOLOv5) para la evaluación en tiempo real de la calidad de la maracuyá y lograr resultados óptimos considerando las diferentes características que pueden tener los tipos de maracuyás. Todo esto realizado con el enfoque de mejorar la eficiencia y reducir el desperdicio en el proceso de clasificación, además de que se busca aplicar este sistema en las plantas de producción de maracuyá en Perú, abordando así los desafíos identificados [2].

El documento se estructura en secciones que incluyen una revisión de otros trabajos relevantes, la descripción del modelo propuesto, la validación de la información, y finalmente las conclusiones y recomendaciones del proyecto. Se busca proporcionar una solución tecnológica que mejore significativamente la clasificación del maracuyá y elimine el factor del error humano, contribuyendo a la eficiencia y sostenibilidad de la industria global.

1. Trabajos relacionados

Esta sección presenta una revisión sistemática y literaria, dividida en formulación y búsqueda de preguntas, selección de palabras claves, y definición de los criterios de inclusión y exclusión. Nuestras preguntas de investigación definidas son Q1: ¿Qué sistemas de detección en tiempo real, que utilicen técnicas de reconocimiento visual para la identificación de frutas defectuosas, existen?; Q2: ¿Cuáles son los elementos que componen un sistema de reconocimiento visual de frutas?; Q3: ¿Qué algoritmo resulta más exacto y apropiado para el reconocimiento de características particulares en un grupo de imágenes?; y Q4: ¿Cómo se verifica la eficacia del sistema de reconocimiento visual cuando se aplica a la evaluación de frutas?.

Las palabras claves que hemos definido son: “reconocimiento visual”, “deep learning”, “redes neuronales”, “espectroscopia de infrarrojo”, “pardeamiento”, y “recall”. Las bases de datos multidisciplinarias seleccionadas fueron “Scopus”, “ScienceDirect” e “IEEE”. Al buscar artículos solo se consideraron los publicados después del 2020, sin considerar artículos del tipo “reseña” y que respondieran a por lo menos una de nuestras preguntas de investigación. Finalmente, 25 estudios fueron seleccionados y analizados para su uso como apoyo en nuestra propuesta de proyecto, elegidos por ser los más afianzados a las respuestas de nuestras preguntas de investigación.

Tabla I: *Referencias por pregunta*

|  |  |  |
| --- | --- | --- |
| **Categoría** | **Referencias** | **Cantidad** |
| Técnicas (Q1) | [23], [25], [27] | 3 |
| Componentes (Q2) | [12], [15], [16], [19], [22], [23], [27], [30] | 8 |
| Algoritmos (Q3) | [8], [9], [10], [13], [14], [17], [18], [20], [25], [26], [28], [29] | 12 |
| Eficacia (Q4) | [7], [11], [13], [14], [18], [21], [22], [24], [28], [33] | 10 |

Es crucial tener conocimiento de las técnicas de reconocimiento visual actualmente disponibles y la diferencia entre cada una de estas para definir bien sus ventajas y desventajas sobre las otras. En [23], los autores utilizaron un sistema de visión por computador combinado con Deep Learning y redes neuronales para la detección de frutas defectuosas en una línea de clasificación. Por su parte, [25], los autores. propusieron métodos como ResNet y DCGAN para la detección de defectos en peras. En [27], los autores, describieron un sistema de inspección visual en tiempo real para clasificar frutas utilizando técnicas de aprendizaje profundo y un módulo Raspberry Pi.

Los sistemas de reconocimiento visual de frutas se componen de múltiples componentes tanto de hardware como de software, esenciales para su funcionamiento eficiente y preciso. En términos de software, [15], los autores mencionaron que una Red Neuronal Convolucional (CNN) es fundamental para procesar imágenes a través de capas de entrada, ocultas y de clasificación, extrayendo características para determinar el tipo o estado de la fruta, como se demuestra en la detección de la madurez de las papayas, por su parte, [30], los autores utilizaron una CNN en la identificación de enfermedades y defectos en frambuesas.

En [23], [27], los autores, indicaron que estos sistemas también requieren un entorno controlado para la captura de imágenes, que puede incluir cámaras RGB de alta resolución y sistemas de iluminación uniforme con luces LED y halógenas. Además, en [27], los autores añadieron hardware adicional, como los módulos Raspberry Pi equipados con cámaras y pantallas táctiles, se utiliza junto con técnicas de aprendizaje profundo y visión por computadora para clasificar frutas en tiempo real.

En [22], los autores propusieron, la integración de técnicas avanzadas como la espectroscopía NIR, crucial para la evaluación de calidad, permitiendo analizar parámetros como el contenido de humedad y sólidos solubles en peras. Además, en [12], los autores utilizaron de sistemas de iluminación LED optimizados, como en el caso de la evaluación de la madurez de las uvas, y en, [19], los autores sugirieron, la implementación de tecnologías como el empaquetado inteligente para monitorear la frescura, además, en [16], los autores, destacaron la importancia de una infraestructura adecuada y tecnologías de apoyo en el reconocimiento y clasificación precisos de las frutas. Estos componentes trabajan en conjunto para adaptarse a las necesidades específicas de cada aplicación, asegurando la eficacia y precisión del sistema en una variedad de contextos agrícolas.

Diversos estudios han evaluado algoritmos para el reconocimiento en tiempo real de características en imágenes. En [8], los autores pusieron a prueba el modelo S-MGDM, basado en YOLOv3, que ha demostrado ser efectivo en la detección de uvas Shine-Muscat con una precisión del 96.73%. Además, en [13], los autores, para la identificación de hematomas en manzanas, usaron el algoritmo YOLOv5s, que sobresalió con un mAP de 99.6%, superando a otros algoritmos como YOLOv3-Tiny y Faster R-CNN. Por su parte, en [17], los autores indicaron que el algoritmo RetinaNet con la arquitectura ResNet-50 ha mostrado eficacia en la detección de objetos en imágenes de piñas, con un mAP del 94.19%.

En [28], los autores encontraron que el modelo MASK-RCNN es altamente preciso en el reconocimiento de contornos y focos de enfermedades en manzanas, bajo condiciones climáticas variables, alcanzando altos niveles de precisión y recall. Por otro lado, en [25], el modelo ResNet50 que había sido mejorado con datos sintéticos generados por DCGAN obtuvo una precisión del 97.35% en la detección de defectos en peras, superando a otros modelos como AlexNet y VGG19. En el caso de la detección de hematomas en fresas, [10] demostró que un modelo CNN optimizado supera significativamente al mejor modelo preentrenado, VGG-19, con una precisión del 98%.

En [26], los autores combinaron YOLO V4 y BiSeNet V2, encontrando que esta combinación era la más efectiva para la detección y segmentación de defectos en manzanas en tiempo real. Además, [18] reportó que el modelo ResNet50 alcanzó la mayor precisión del 95.7% en la clasificación de calidad de frutas, aunque EfficientNetB0 fue preferido por su equilibrio entre precisión y eficiencia computacional. En [29], los autores resaltaron el uso de métodos basados en Kernel, como LS-SVM, junto con técnicas de selección de variables como CARS y GA, que son altamente efectivos en la evaluación de la calidad de frutas mediante espectroscopía Vis/NIR.

Asimismo, en [14], los autores probaron que el algoritmo EfficientNetB4 y ResNet50 que han mostrado precisión en la detección de defectos en tomates. En [20], los autores utilizaron CNN para la clasificación de frutas y verduras, logrando un reconocimiento efectivo de enfermedades. Finalmente, [9] mostró que los algoritmos de detección de objetos como YOLO y RetinaNet parecen ser más adecuados debido a su capacidad de identificar características específicas rápidamente y con alta precisión en sistemas en tiempo real.

Verificar la eficacia del sistema de reconocimiento visual es importante para comprobar que en efecto es la mejor opción posible para elegir. Como en [11] los autores utilizaron un enfoque de espectroscopia de infrarrojo para clasificar el pardeamiento en las frutas y su eficacia se verificó mediante la comparación de los resultados del modelo con las observaciones visuales reales del pardeamiento en las frutas. Por otro lado, en [13] los autores utilizaron indicadores como Precisión (P), Recall (R), Average Precision (AP) y mean Average Precision (mAP) para verificar la eficacia del modelo al colocarlos como los índices de evaluación de la precisión del modelo.

En [7], los autores verificaron la efectividad del sistema de reconocimiento visual mediante la preparación y aumento de datos, extracción y selección de características, y evaluación mediante múltiples algoritmos de clasificación supervisados. El uso de matrices de confusión y análisis de eficiencia computacional demostró una precisión del 99.7%. En [14], los autores emplearon redes neuronales residuales (ResNet) para detectar defectos externos en tomates, logrando una precisión promedio del 94.6% mediante la preparación exhaustiva del conjunto de datos y métricas de rendimiento. En [18], los autores verificaron la efectividad del sistema mediante métricas de rendimiento, curvas de aprendizaje y análisis de componentes principales (PCA), logrando una precisión del 95.7% con ResNet50. En [21], los autores evaluaron diferentes algoritmos de aprendizaje automático en frutas dragón, utilizando métricas como precisión, recall y F1-score, con CNN mostrando mejoras significativas en velocidad y efectividad. En [22], los autores validaron el sistema de reconocimiento visual utilizando espectroscopia NIR con métodos de selección de variables y actualización del modelo, mejorando la precisión y reduciendo el sesgo de predicción.

En [24], los autores verificaron la efectividad del sistema en la detección de enfermedades en maracuyá mediante la preparación y etiquetado de conjuntos de datos, modelos de color y análisis estadístico, logrando una precisión promedio del 79%. Mientras que en [28], los autores validaron el sistema mediante métricas de rendimiento, experimentos de campo y pruebas de aplicación en tiempo real, logrando una precisión promedio de 0.926 y una tasa de recall de 0.961. Finalmente, en [33], los autores evaluaron un modelo de conjunto ligero basado en CNN para detectar defectos en zanahorias, utilizando estrategias de combinación y métricas como precisión, sensibilidad y F1-score, logrando una precisión del 97.04%.

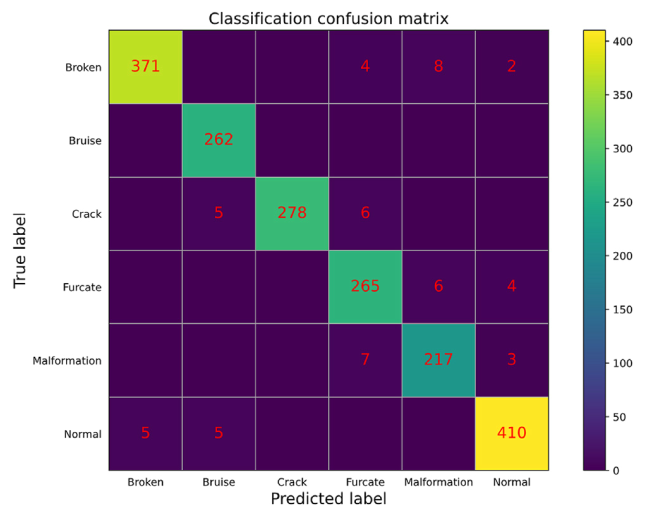


Figura I: *Matriz de confusión obtenida en el estudio realizado por W. Xie, S. Wei, Z. Zheng y D. Yang [33]*

1. Modelo propuesto

Este artículo propone una arquitectura innovadora basada en el algoritmo YOLOv5 para la detección de maracuyá defectuoso. Nuestro enfoque se centró en una variante de Spatial Pyramid Pooling (SPP) y una modificación de la Path Aggregation Network (PANet) para integrar BottleneckCSP.

Esta estructura combinó avanzadas técnicas de visión por computadora, como se detallada en la Figura II, para detectar objetos en tiempo real con alta precisión y eficiencia [6], [31]. Esto es posible pre entrenando al sistema con un dataset y una única red neuronal para detectar objetos en imágenes.

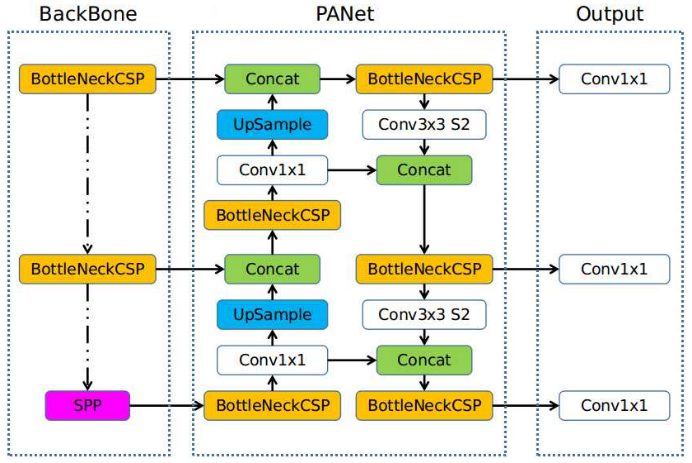


Figura II: *Estructura del algoritmo YOLOv5*

La Figura III ilustra el diagrama de la arquitectura base del sistema que se planeó desarrollar, mostrando de manera detallada la interacción entre todas las partes de dicho sistema, así como su funcionamiento y la interacción con el usuario. Cabe destacar que, para este estudio, se decidió crear una arquitectura sencilla con el objetivo de minimizar los costos de desarrollo y distribución. Esta simplificación no solo está orientada a facilitar las pruebas, sino también a garantizar una implementación práctica y eficiente en entornos reales.

Diagrama

Descripción generada automáticamente

Figura III: *Diagrama de arquitectura base*

**A. Fase 1 – Preparación y configuración inicial**

En esta fase inicial, el enfoque está en la creación de un conjunto de datos robusto y representativo. Se llevó a cabo una exhaustiva recopilación y anotación de imágenes de maracuyá bajo diversas condiciones. Este meticuloso proceso es fundamental para entrenar de manera eficaz el algoritmo YO-LOv5, garantizando que pueda reconocer una amplia gama de características y variantes en los maracuyás.

La Figura IV ilustra este proceso, mostrando la captura de imágenes de maracuyás que se utilizan para formar el dataset en un Raspberry Pi, reflejando así la diversidad y la calidad de las imágenes recopiladas. Paralelamente, se configura la infraestructura de hardware, incluyendo la cámara y el CPU, para garantizar su compatibilidad y rendimiento óptimo con el modelo de reconocimiento visual.

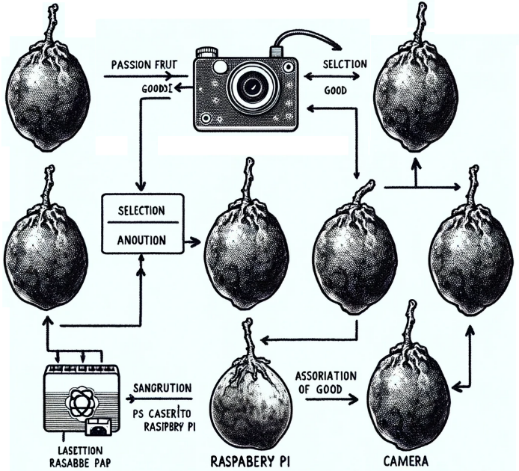


Figura IV: Muestra de cómo se tomaron fotos a diferentes maracuyás para el entrenamiento del YOLOv5

**B. Fase 2 – Desarrollo y personalización para el modelo YOLOv5**

Esta etapa se centró en la adaptación y optimización del modelo con el algoritmo YOLOv5. Se realizaron modificaciones específicas, como la integración de una variante de Spatial Pyramid Pooling y la modificación de la Path Aggregation Network con BottleneckCSP, para mejorar la detección de defectos en maracuyás [6], [32]. El proceso de entrenamiento del modelo, que se ve reflejado en la Figura V, es crucial para entender cómo las imágenes del dataset se utilizaron para afinar la capacidad del modelo de identificar maracuyás defectuosos con alta precisión.



Figura V: Muestra interna del entrenamiento con el algoritmo YOLOv5 usando un dataset con imágenes de maracuyás

**C. Fase 3 – Integración y pruebas del sistema**

En esta fase crucial, se integró el algoritmo YOLOv5 dentro del sistema completo, estableciendo una comunicación efectiva entre la cámara, el CPU, el dataset y la interfaz de usuario, tal como se ilustró en la Figura III. La cámara desempeñó un papel clave al capturar las imágenes de los maracuyás, que fueron posteriormente procesadas por el CPU utilizando el algoritmo YOLOv5 para la identificación de sus características. Este proceso fue vital para la selección precisa de maracuyás defectuosos en base a sus características únicas por cada categoría, y es que cada una de estas puede presentar características que puedan ser interpretadas correctamente como defectos en dicha categoría, pero incorrectamente no en otra categoría.

La Figura VI refleja esta integración, mostrando un ejemplo de cómo el sistema procesaba las imágenes capturadas y, basándose en las características visuales, identificaba las características únicas de cada maracuyá. Esta etapa del proyecto no solo demostró la funcionalidad técnica del modelo en un entorno de prueba, sino que también subrayó su aplicabilidad práctica en un escenario de trabajo real. Además, destacó la importancia de la interacción fluida entre los distintos componentes del sistema, facilitando su uso para el usuario.

Esto último también fue importante, ya que el sistema debía ser lo más sencillo e intuitivo posible para el usuario. Ya que en algunos casos, esta podría ser la primera vez que un sistema de este tipo fuera utilizado por los trabajadores de una fábrica (pensando en el caso de una posible distribución del sistema), y se esperaría que el uso del sistema fuese de fácil aprendizaje.

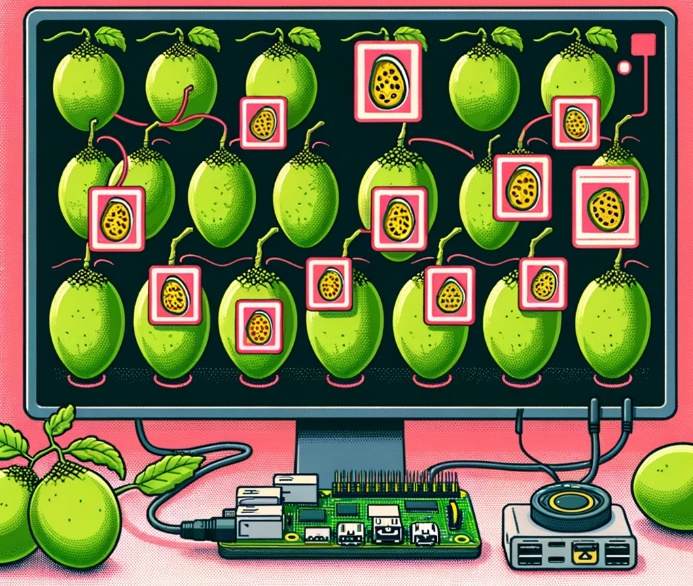


Figura VI: Muestra de cómo el sistema fue capaz de usar las fotos de maracuyás e identificar su condición, en base a sus características visuales

**D. Fase 4 – Despliegue y retroalimentación**

Esta fase representó el despliegue final del sistema en entornos de uso real, un paso crucial para validar su eficacia en condiciones prácticas. La instalación de la cámara y la configuración del sistema se llevaron a cabo en la zona de detección de maracuyás defectuosos.

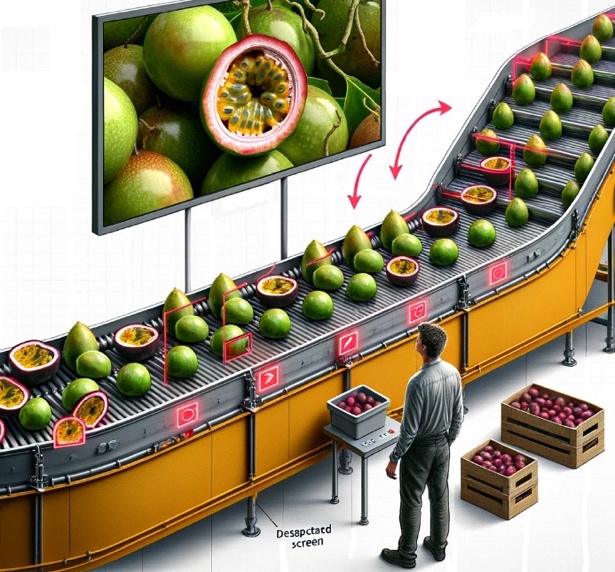


Figura VII: El sistema en funcionamiento, identificando las maracuyás como verdes, maduros, desechables, etc.

Un aspecto clave fue el desarrollo de una interfaz de usuario intuitiva y eficiente, que permitiera a los usuarios identificar rápidamente los maracuyás no aptos. Esta identificación se realizó mediante un cuadro rojo superpuesto en las imágenes de los maracuyás defectuosos, indicando claramente que debían ser retirados. Se designaron colores específicos para los cuadros según la categoría del maracuyá analizado.

La Figura VII ilustró vívidamente este proceso, demostrando el sistema en acción y su capacidad para clasificar los maracuyás en diferentes categorías basándose en su condición. La efectividad de esta interfaz y la precisión de la detección fueron esenciales para el éxito del sistema en un entorno real, proporcionando una herramienta valiosa para la clasificación y selección de maracuyás.

1. Validación

La validación del sistema de detección de maracuyá defectuoso utilizando YOLOv5 es un paso crucial para asegurar su eficacia y aplicabilidad en entornos de producción reales. Esta fase busca evaluar la precisión, la velocidad de respuesta y la integrabilidad del sistema en las operaciones existentes de clasificación de frutas.

A continuación, presentamos una propuesta detallada para validar el sistema bajo condiciones prácticas y operativas. Explicando las muestras y caso de estudio considerados, nuestras métricas de evaluación que sirvieron como instrumentos para el cálculo y validación de los resultados, y además los que fueron nuestros sujetos de estudio.

**A. Muestra y caso de estudio**

* **Selección de muestra**

Se seleccionaron 500 maracuyás durante un período de un mes en una planta de producción de jugo de maracuyá en Perú. Las muestras incluyeron maracuyás en diferentes estados como: arrugado, maduro o manchado.

* **Caso de estudio**

El sistema fue implementado en una línea operativa de clasificación de maracuyás de una instalación en funcionamiento, para monitorear y analizar su rendimiento en condiciones reales de operación.

**B. Métricas de evaluación**

* **Consideración inicial de las fórmulas utilizadas**

Normalmente se utilizarían fórmulas como estas para hallar la precisión y el recall:

*VP (Verdaderos Positivos):* Número de maracuyás defectuosos correctamente identificados por el sistema.

*FP (Falsos Positivos):* Número de maracuyás no defectuosos que el sistema identifica erróneamente como defectuosos.

*VP (Verdaderos Positivos):* Número de maracuyás defectuosos correctamente identificados por el sistema.

*FP (Falsos Positivos):* Número de maracuyás no defectuosos que el sistema identifica erróneamente como defectuosos.

Sin embargo, estas fórmulas solo deben ser utilizadas para una matriz de confusión cotidiana de solo dos clases, y debido a que la nuestra fue de varias clases, utilizamos fórmulas diferentes.

Si bien, también es posible convertir los resultados de una matriz de confusión con múltiples clases a resultados equivalentes de una de solo dos clases para seguir aplicando estas misma fórmulas, nos decidimos por optar por otro método ya que así tenemos el beneficio adicional de obtener los valores de precisión y recall por cada clase y no solo los valores generales de todo el modelo.

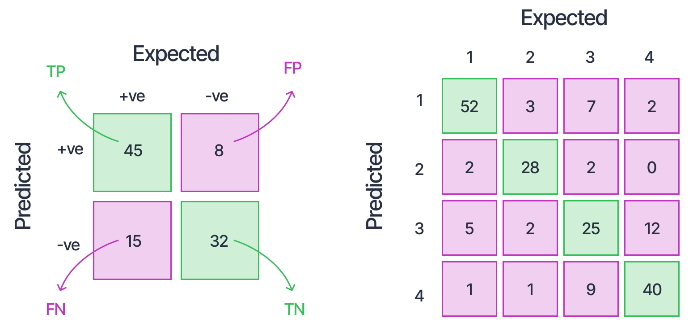


Figura VIII: *Comparación entre, selección de valores de una matriz de confusión con solo dos clases (izquierda) VS una matriz de confusión de múltiples clases (derecha)*

Por lo que se debe tener en cuenta que las fórmulas de precisión y recall que mostraremos a continuación solo son usadas para matrices de confusión de múltiples clases.

* **Precisión**

Porcentaje de maracuyás defectuosos correctamente identificados entre todas las identificaciones realizadas por el sistema.

WP: % ponderado de precisión de todo el modelo

: % de precisión de una clase específica del modelo, donde “n” representa la clase

: Valor de una clase específica del modelo que forma parte de la diagonal del modelo, donde “n” representa la clase

: Suma de valores esperados de una clase específica del modelo, donde “n” representa la clase

: Suma de valores predichos de una clase específica del modelo, donde “n” representa la clase

* **Recall (sensibilidad)**

Proporción de maracuyás defectuosos correctamente identificados entre todos los maracuyás defectuosos presentes en la muestra.

WR: % ponderado de recall (también conocido como sensibilidad) de todo el modelo

: % de recall de una clase específica del modelo, donde “n” representa la clase

: Valor de una clase específica del modelo que forma parte de la diagonal del modelo, donde “n” representa la clase

: Suma de valores esperados de una clase específica del modelo, donde “n” representa la clase

: Suma de valores predichos de una clase específica del modelo, donde “n” representa la clase

Hay que considerar que debido a los valores de la fórmula de la fórmula de WR normalmente es resumida de la siguiente forma:

Sin embargo, no lo hicimos porque, como dijimos, queríamos dejar en legado los valores obtenidos para por cada clase, para su apreciación y por si fueran necesarios.

* **F1-Score**

Métrica que combina precisión y recall para proporcionar una visión general del rendimiento del sistema. Cabe mencionar que la fórmula de F1-Score se mantiene igual, aunque sea una matriz de confusión de múltiples clases.

*P (Precisión):* La precisión mide la proporción de identificaciones correctas sobre el total de identificaciones realizadas como defectuosas por el sistema.

*R (Recall):* El recall evalúa la capacidad del sistema para identificar correctamente todos los maracuyás defectuosos presentes en la muestra.

* **Tiempo de respuesta (T Respuesta)**

Duración promedio desde que una fruta es detectada hasta su clasificación por el sistema.

*T Inicial​ (Tiempo Inicial):* Momento en que la fruta entra en el campo de visión del sistema.

*T Final​ (Tiempo Final):* Momento en que se completa la clasificación de la fruta.

**C. Sujetos de estudio**

* **Operadores de la planta**

Trabajadores del lugar, su feedback es vital para evaluar la usabilidad y la integración del sistema en el flujo de trabajo existente, así como la facilidad de uso que debe tener el sistema.

* **Expertos en la calidad del maracuyá**

Personal especializado que realizo una evaluación manual paralela de las frutas para comparar los resultados del sistema con la opinión de un profesional como lo son estos expertos.

* **Equipo de desarrollo de tecnologías de información**

Fueron quienes evaluaron el rendimiento técnico, al momento de la integración del sistema con la infraestructura de la planta en el momento.

**D. Proceso de validación**

* **Fase 1 – Calibración y pruebas iniciales**

Antes de empezar es necesario realizar algunos ajustes y calibraciones iniciales para el sistema antes de su uso completo. Ya que en nuestro caso utilizamos el sistema solo para analizar maracuyás, y el sistema puede ser programado incluso para analizar cosas que no sean frutas.

* **Fase 2 – Implementación y monitoreo**

Durante el primer mes de pruebas, el sistema se monitoreo continuamente, registrando las métricas definidas y los resultados obtenidos. De ser necesario este tiempo puede ampliarse hasta que el sistema arroje datos concisos en los resultados. Una vez que llegado a un punto en el que obtuvimos resultados óptimos, el monitoreo comenzó a hacerse con menos frecuencia.

* **Fase 3 – Evaluación comparativa**

Se realizo una comparación de los resultados obtenidos por el sistema con la opinión dada por un grupo de expertos de la planta (expertos en la calidad del maracuyá).

* **Fase 4 – Análisis y ajustes**

Basándose en los resultados y en la retroalimentación obtenida, se realizaron los últimos ajustes necesarios para optimizar el sistema. De ser necesario se pueden repetir los pasos de la Fase 2 y Fase 3, y continuar con este proceso de repetición hasta obtener un análisis adecuado en la Fase 4.

* **Fase 5 – Reporte final**

Se elaboro un informe detallado sobre el rendimiento del sistema, con recomendaciones para mejoras futuras. Si fuese necesario, se puede solicitar el apoyo y/u opinión del personal de expertos en la calidad del maracuyá para la elaboración de este informe.

En nuestra primera prueba numerada obtuvimos valores aproximados entre 70% y 75% en los resultados del sistema, sin embargo, con forme pudimos mejorar los métodos utilizados, así como aumentar el número de imágenes que eran analizadas por el sistema pudimos ir mejorando los resultados con cada prueba que seguíamos ejecutando. Finalmente, en nuestra última prueba obtuvimos unos resultados favorables de 92.869% en la consistencia de la precisión y 91.224% en la fiabilidad del recall, como validación del resultado del proyecto.

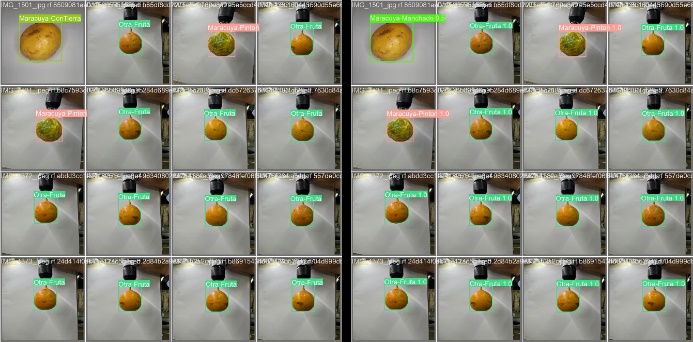


Figura IX: *Comparación entre, un remarcado hecho manualmente (izquierda) VS el remarcado hecho por el sistema en la última prueba (derecha)*

Es así pues que esta propuesta de validación busca proporcionar una evaluación rigurosa del sistema de detección de maracuyá defectuoso, asegurando su precisión, eficiencia y usabilidad en un entorno real de producción.

1. Resultados y discusión

**A. Resultados**

Con el algoritmo YOLOv5 pudimos obtener los resultados en forma de una matriz de confusión que, en una explicación simple, es una visualización que muestra cómo se está desempeñando el modelo con respecto a las clases en las que fue entrenado. En nuestro caso las clases (o categorías) elegidas para este estudio fueron elegidas en base a las posibilidades de variaciones de maracuyás, siendo concretamente: “Maracuyá Descarte”, “Maracuyá Pintón”, Maracuyá Verde”, Maracuyá Arrugado”, Maracuyá Maduro”, Maracuyá Manchado”, Maracuyá con Tierra” y “Otra Fruta”.

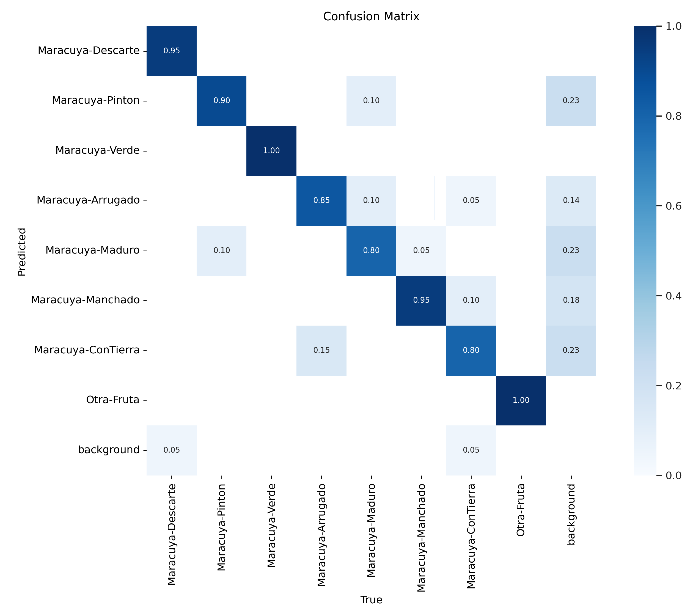


Figura X: *Matriz de confusión de la última prueba*

Luego de procesados los datos obtenidos en el grafico anterior, se obtuvieron resultados que fueron muy prometedores como se muestran en la tabla a continuación. Aunque el primer valor de la precisión no cumplió con las expectativas, eso se debió a la forma de progresión de aprendizaje que tiene el sistema.

Tabla II: *Resultados obtenidos*

|  |  |  |  |
| --- | --- | --- | --- |
| **# de la prueba** | **# de imágenes del dataset** | **Consistencia de la precisión** | **Fiabilidad del recall** |
| 2 | 200 aprox. | 75.406% | 72.510% |
| 3 | 350 aprox. | 89.832% | 87.504% |
| 4 | 500 aprox. | 91.393% | 90.136% |
| 5 | 500 aprox. | 92.869% | 91.224% |

A considerar que, se omitió la prueba 1 dentro de los resultados finales ya que esta prueba solo se ejecutó para la validación del correcto funcionamiento de los procesos del algoritmo YOLOv5.

Se logró obtener una precisión mayor al 90% con los resultados obtenidos de la última prueba, la cual uso unas 500 imágenes aproximadamente. Si bien se podría haber aumentado el porcentaje de precisión con una mayor cantidad de imágenes, no necesariamente era la única opción. Ya que, es posible también mejorar la precisión con el riesgo de aumentar el tiempo de procesamiento, sin embargo, el objetivo siempre fue lograr buenos resultados, pero sin llegar a aumentar mucho el tiempo de procesamiento. En pocas palabras, llegar a balancear bien el sistema en cuanto a los resultados obtenidos y su tiempo de procesamiento.

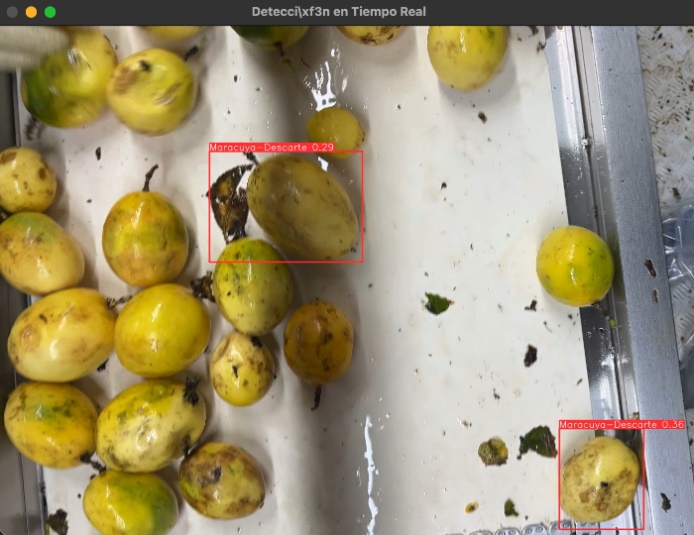


Figura XI: *Demostración de la ejecución del sistema en un entorno realista (una fábrica de jugo de maracuyá)*

En la Figura XI se puede apreciar el sistema funcionando correctamente en un entorno real, en este caso una fábrica de jugo de maracuyá, demostrando que el sistema ha funcionado eficazmente y no solo en un entorno controlado y al realizar pruebas, sino también al momento de realizar una ejecución real.

**B. Discusión**

Las pruebas realizadas demostraron que algunas categorías específicas de maracuyás, como por ejemplo el “maracuyá con tierra”, pueden generar más discordancia que otras categorías, cosa que podría pasar con otras frutas, aunque talvez esto sea solo por la alta variabilidad de las características que tienen los maracuyás a diferencia de la mayoría de las frutas. Esto fue solucionado aumentando la cantidad de imágenes por encima del promedio para que el sistema mejore su entendimiento de esta categoría especifica y así no presentar problemas de confusión con las demás, por lo tanto, sugeriríamos probar este método en futuros estudios. Sin embargo, hay que considerar que, si bien este método funcionó con nuestro sistema, esto pudo deberse a la dificultad de aprendizaje que tiene el sistema con la maracuyá, específicamente por la alta variedad en sus características.

Finalmente, comparando nuestros resultados con los de otros estudios como el [9] que uso 500 imágenes por fruta, y [10] que uso casi 3,000 imágenes de fresas, podemos apreciar que nuestros resultados se acercaron a otros en precisión, con una diferencia aproximada de solo 5% por encima de nuestros resultados.

Tabla III: *Resultados de otros trabajos*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Paper** | **Algoritmo** | **Fruta** | **Consistencia de la precisión** | **Fiabilidad del recall** |
| [9] | XGBoost | Kiwi | 98.44% | 93.75% |
| [9] | XGBoost | Naranja | 98.00% | 93.92% |
| [9] | XGBoost | Ciruela | 98.44% | 95.36% |
| [9] | XGBoost | Sandía | 97.94% | 94.44% |
| [10] | InceptionV3 | Fresa | 65% | 70% |
| [10] | VGG19 | Fresa | 69% | 76% |
| [10] | Resnet50 | Fresa | 62% | 44% |
| [10] | Optimized | Fresa | 98% | 98% |

Nuestros resultados sugieren que el algoritmo YOLOv5 puede ser usado de forma eficiente para un proceso de clasificación y descarté de maracuyás. Y como se ha demostrado en las pruebas esto se pudo lograr aun considerando la gran variedad de características diferenciales que pueden tener algunas frutas como en nuestro caso fue con los maracuyás.

Por estos motivos nos parece totalmente aceptable, utilizar el algoritmo YOLOv5 para la detección, clasificación y descarte de maracuyás según sus características; y aplicarlo en un ámbito laboral.

1. Conclusiones y trabajos futuros

**A. Conclusión general**

El estudio demostró con éxito la aplicación de un sistema de reconocimiento visual basado en el algoritmo YOLOv5 para la clasificación de maracuyá defectuoso antes de su procesamiento en la producción de jugo. La integración de técnicas avanzadas como Spatial Pyramid Pooling y Path Aggregation Network con BottleneckCSP resultó en un modelo altamente eficiente, superando significativamente las limitaciones de los métodos manuales tradicionales. Demostrando así la posibilidad de la optimización de la cadena de producción del jugo de maracuyá, y toda la mejora adicional que conllevaría al resto del proceso, y subsecuentemente a ahorros en costos de producción.

En resumen, este estudio ofreció un ejecución tecnológica en la clasificación de maracuyá con una mejora significativa con respecto a los procedimientos manuales estándares, con implicaciones directas en la mejora de la calidad del jugo de maracuyá y en la eficiencia de su producción. Las innovaciones en reconocimiento visual y procesamiento de imágenes aplicadas en este contexto abrieron nuevas perspectivas para la automatización y optimización de procesos en la industria alimentaria, lo cual como ha quedado demostrado, puede ser aprovechado con la finalidad de mejorar todo el proceso de selección, con cosas como: reducción de margen de error al momento de seleccionar los maracuyás, mayor facilidad de creación de graficas de resultados por categoría, reducción de costos de personal, y en algunos casos mejorando la calidad de insumos producidos (en caso que la empresa utilice los maracuyás para generar productos adicionales, como por ejemplo: zumo de maracuyá).

**B. Conclusiones específicas**

* **Mejora en la precisión y velocidad de clasificación**

El algoritmo YOLOv5 mejorado identificó de manera rápida y precisa los maracuyás defectuosos. Esta eficiencia en la detección contribuyó a una selección más acertada de la fruta para la producción de jugo, reduciendo el riesgo de incluir frutas de baja calidad y mejorando el producto final.

* **Reducción del desperdicio y aumento de la productividad**

La implementación del sistema automatizado mostró un impacto positivo con la reducción de maracuyá desperdiciado, debido a una mala selección, y en el incremento de la productividad final. Al disminuir la cantidad de frutas defectuosas procesadas inadvertidamente, el sistema permitió una gestión más efectiva de los recursos y un ahorro en los costos operativos.

* **Interfaz visual para asistencia en la clasificación**

La interfaz visual del sistema, que resaltaba las frutas defectuosas con un cuadro rojo, proporcionó una herramienta asistencial crucial para los trabajadores en la línea de clasificación, así también los demás cuadros de colores para cada categoría no supusieron un problema de confusión para los trabajadores de acuerdo con nuestras pruebas realizadas en el área de trabajo. Aunque la interfaz no permitía una interacción directa, su diseño intuitivo, de fácil uso y su función de señalización mejoro notablemente la eficacia del proceso de selección manual, asegurando que las frutas no aptas fueran retiradas con precisión.

**C. Trabajos futuros**

Estudios futuros deberían poder entrenar el modelo propuesto con un mayor número de imágenes y evaluar los resultados. Asimismo, se debería realizar el estudio con otras frutas con alta variabilidad de características tales como las fresas, mandarina, lúcuma, arándano o plátano.

Agradecimientos

Agradecemos el apoyo de la Universidad Peruana de Ciencias Aplicadas (UPC) por el apoyo en la realización de la presente investigación..

Referencias

1. Sierra y Selva Exportadora. (2021). Análisis de Mercado 2015 - 2020 Maracuyá. Recuperado el 13 de agosto de 2023, de https://repositorio.sierraexportadora.gob.pe/handle/SSE/385 (Ministerio de Desarrollo Agrario y Riego del Perú).
2. Estadísticas de producto - JUGO DE MARACUYA. (2023, junio). Adex Data Trade. Recuperado el 20 de agosto de 2023, de https://adexdatatrade.upc.elogim.com/Members/EstadisticaProducto.aspx?partida=2009892000#
3. M. Rashidi, M. Giholami, y S. Abbasi. “Cantaloupe volume determination through image processing”, J. Agric. Sci. Technol. Vol. 11, pp. 623-631. 2009.
4. M. Omid, M. Khojastehnazhand, A. Tabatabaeefar. “Estimating volume and mass of citrus fruits by image processing technique”, J. Food Eng. Vol. 100, pp. 315-321. 2010.
5. L. Xiaoting, J. Xueying. “Real-Time Grading of Defect Apples Using Semantic Segmentation Combination with a Pruned YOLO V4 Network”. Vol 11, pp 3150.”, Sensors, 21(4), 1375. 2021.
6. C. Liu, W. Lin, Y. Feng, Z. Guo, y Z. Xie, "ATC-YOLOV5: Fruit Appearance Quality Classification algorithm based on the improved YOLOV5 model for Passion Fruits," Mathematics, vol. 11, no. 16, p. 3615, 2023. https://doi.org/10.3390/math11163615
7. Shah, F., Khan, M. A., Sharif, M., Tariq, U., Khan, A., Kadry, S., & Thinnukool, O. (2022). A cascaded design of best features selection for fruit diseases recognition. Computers, materials & continua, 70(1), 1491-1507. <https://doi.org/10.32604/cmc.2022.019490>
8. Wei, X., Xie, F., Wang, K., Jian, S., & Bai, Y. (2023). A study on Shine-Muscat grape detection at Maturity based on Deep learning. Scientific Reports, 13(1). <https://doi.org/10.1038/s41598-023-31608-6>
9. Alharbi, A. H., Alkhalaf, S., Asiri, Y., Abdel‐Khalek, S., & Mansour, R. F. (2023). Automated fruit classification using enhanced Tunicate swarm algorithm with fusion based deep learning. Computers & Electrical Engineering, 108, 108657. <https://doi.org/10.1016/j.compeleceng.2023.108657>
10. Guo, B., Li, B., Huang, Y., Fa-Yi, H., Xu, B., & Dong, Y. (2022). Bruise detection and classification of strawberries based on thermal images. Food and Bioprocess Technology, 15(5), 1133-1141. <https://doi.org/10.1007/s11947-022-02804-5>
11. Daniels, A., Poblete-Echeverría, C., Nieuwoudt, H., Botha, N., & Opara, U. L. (2021). Classification of browning on intact table grape bunches using Near-Infrared spectroscopy coupled with partial least Squares-Discriminant analysis and artificial neural networks. Frontiers in Plant Science, 12. <https://doi.org/10.3389/fpls.2021.768046>
12. Pampuri, A., Tugnolo, A., Giovenzana, V., Casson, A., Guidetti, R., & Beghi, R. (2021). Design of cost-effective LED based prototypes for the evaluation of grape (Vitis vinifera L.) ripeness. Computers and Electronics in Agriculture, 189, 106381. <https://doi.org/10.1016/j.compag.2021.106381>
13. Yuan, Y., Yang, Z., Liu, H., Wang, H., Li, J., & Zhao, L. (2022b). Detection of early bruise in Apple using near-infrared camera imaging technology combined with deep learning. Infrared Physics & Technology, 127, 104442.
14. Da Costa, A. Z., Hernandez-Figueroa, H. E., & Fracarolli, J. A. (2020). Computer vision-based detection of external defects on tomatoes using deep learning. Biosys-tems Engineering, 190, 131-144. <https://doi.org/10.1016/j.biosystemseng.2019.12.003>
15. Gayathri, S., Ujwala, T., Vinusha, C., Pauline, N. R., & Tharunika, D. (2021). Detection of papaya ripeness using deep learning approach. 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA). <https://doi.org/10.1109/icirca51532.2021.9544902>
16. Yu, B., Zhan, P., Lei, M., Zhou, F., & Wang, P. (2020). Food quality monitoring system based on smart contracts and evaluation models. IEEE Access, 8, 12479-12490. <https://doi.org/10.1109/access.2020.2966020>
17. Siricharoen, P., Yomsatieankul, W., & Bunsri, T. (2023). Fruit Maturity Grading Framework for small dataset using single image multi-object sampling and mask R-CNN. Smart Agricultural Technology, 3, 100130. <https://doi.org/10.1016/j.atech.2022.100130>
18. Suhendar, H., Efelina, V., & Ziveria, M. (2022). Fruit quality classification using convolutional neural network. Journal of physics, 2377(1), 012015. <https://doi.org/10.1088/1742-6596/2377/1/012015>
19. Alam, A. U., Rathi, P., Beshai, H., Sarabha, G. K., & Deen, M. J. (2021). Fruit quality monitoring with smart packaging. Sensors, 21(4), 1509. <https://doi.org/10.3390/s21041509>
20. Amin, J., Anjum, M. A., Sharif, M., Kadry, S., & Nam, Y. (2022). Fruits and vegetable diseases recognition using convolutional neural networks. Computers, materials & continua, 70(1), 619-635. <https://doi.org/10.32604/cmc.2022.018562>
21. Patil, P. U., Lande, S. B., Nagalkar, V. J., Nikam, S. B., & Wakchaure, G. (2021). Grading and sorting technique of dragon fruits using machine learning algorithms. Journal of Agriculture and Food Research, 4, 100118. <https://doi.org/10.1016/j.jafr.2021.100118>
22. Mishra, P., Woltering, E., Brouwer, B., & Echtelt, E. H. (2021). Improving moisture and soluble solids content prediction in pear fruit using near-infrared spectroscopy with variable selection and model updating approach. Postharvest Biology and Technology, 171, 111348. <https://doi.org/10.1016/j.postharvbio.2020.111348>
23. Fan, S., Li, J., Zhang, Y., Tian, X., Wang, Q., He, X., Zhang, C., & Huang, W. (2020). Online detection of defective Apples using computer vision system combined with deep learning methods. Journal of Food Engineering, 286, 110102. <https://doi.org/10.1016/j.jfoodeng.2020.110102>
24. S. B. D. H. Dharmasiri and S. Jayalal, "Passion Fruit Disease Detection using Image Processing," 2019 International Research Conference on Smart Computing and Systems Engineering (SCSE), Colombo, Sri Lanka, 2019, pp. 126-133. https://doi.org/10.23919/SCSE.2019.8842799
25. Zhang, Y., Wa, S., Sun, P., & Wang, Y. (2021). Pear defect Detection Method based on ReSNEt and DCGAN. Information, 12(10), 397. <https://doi.org/10.3390/info12100397>
26. Liang, X., Jia, X., Huang, W., He, X., Li, L., Fan, S., Li, J., Zhao, C., & Zhang, C. (2022). Real-Time grading of defect Apples using semantic segmentation combination with a pruned YOLO V4 network. Foods, 11(19), 3150. <https://doi.org/10.3390/foods11193150>
27. Ismail, N., & Malik, O. A. (2022). Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. Information Processing in Agriculture, 9(1), 24-37. <https://doi.org/10.1016/j.inpa.2021.01.005>
28. Kutyrev, A., Kiktev, N., Kalivoshko, O., & Rakhmedov, R. (2022). Recognition and Classification Apple Fruits Based on a Convolutional Neural Network Model. In Information Technology and Implementation (IT&I-2022), (pp. 90-101). Kyiv, Ukraine.
29. Wang, M., Xu, Y., Yang, Y., Mo, B., Nikitina, M. A., & Xiao, X. (2022). VIS/NIR Optical Biosensors applications for fruit monitoring. Biosensors And Bioelectronics: X, 11, 100197. <https://doi.org/10.1016/j.biosx.2022.100197>
30. Naranjo-Torres, J., Mora, M., Fredes, C., & Valenzuela, A. M. (2021). Disease and defect detection system for raspberries based on convolutional neural networks. Applied sciences, 11(24), 11868. <https://doi.org/10.3390/app112411868>
31. K. Li et al., "A fast and lightweight detection algorithm for passion fruit pests based on improved YOLOv5," Comput. Electron. Agric., vol. 204, 2023.
32. G. Jocher et al., "YOLO by Ultralytics," GitHub Repository, 2023.
33. W. Xie, S. Wei, Z. Zheng y D. Yang, “A CNN-based lightweight ensemble model for detecting defective carrots”, Biosyst. Eng., vol. 208, pp. 287–299, agosto de 2021. https://doi.org/10.1016/j.biosystemseng.2021.06.008