

Sistema de Diagnóstico de Enfermedades en Hojas de Palta con Deep Learning: Una Aproximación Basada en Redes Neuronales Convolucionales

- Carlos Miguel Quesquen Rodas
- Vasquez Palomino Jack

RESUMEN

La palta (*Persea americana*) ha experimentado un crecimiento acelerado en Perú, afianzando al país como el segundo exportador global más grande. Sin embargo, las enfermedades foliares representan un riesgo constante para su producción, y la identificación clásica es imprecisa y tardía. Este estudio sugiere un método de diagnóstico automatizado que utiliza Aprendizaje Profundo para detectar patologías en las hojas de palta. Para este propósito, se desarrollaron y analizaron modelos de Redes Neuronales Convolucionales (CNN), entrenados con un sólido banco de datos de 1045 fotografías de hojas (tanto saludables como afectadas), distribuidas y ampliadas. La estrategia, de corte cuasi-experimental longitudinal, posibilitó comparar el desempeño de tres configuraciones esenciales: una CNN 2D Adaptada, DenseNet121 y EfficientNetB0. Los hallazgos preliminares muestran que DenseNet121 emerge como el modelo destacado, logrando los valores más altos en precisión (0.632), capacidad de detección (0.500) y puntuación F1 (0.600). En contrapartida, la CNN 2D Adaptada y EfficientNetB0 evidenciaron un bajo rendimiento, especialmente en la identificación de casos positivos. Este planteamiento constituye un instrumento significativo para optimizar la eficacia del manejo fitosanitario en este cultivo vital, contribuyendo a la salud y rendimiento de las plantaciones de palta en Perú.

Palabras clave: Enfermedades foliares, Agricultura de precisión, Manejo fitosanitario, (CNN)

ABSTRACT

Avocado (*Persea americana*) has experienced accelerated growth in Peru, consolidating the country's position as the second-largest global exporter. However, foliar diseases pose a constant risk to its production, and traditional identification is imprecise and time-consuming. This study suggests an automated diagnostic method using Deep Learning to detect pathologies in avocado leaves. For this purpose, Convolutional Neural Network (CNN) models were developed and analyzed, trained with a robust database of 1,045 leaf photographs (both healthy and affected), distributed and enlarged. The longitudinal, quasi-experimental approach made it possible to compare the performance of three key configurations: an Adaptive 2D CNN, DenseNet121, and EfficientNetB0. Preliminary findings show that DenseNet121 emerged as the leading model, achieving the highest values in accuracy (0.632), detection capacity (0.500), and F1 score (0.600). In contrast, the Adaptive 2D CNN and EfficientNetB0 showed poor performance, especially in identifying positive cases. This approach constitutes a significant tool for optimizing the effectiveness of phytosanitary management in this vital crop, contributing to the health and yield of avocado plantations in Peru.

Keywords: Foliar diseases, Precision agriculture, Phytosanitary management, (CNN)

INTRODUCCIÓN

El aguacate (*Persea americana*) se ha erigido en un cultivo fundamental para la economía peruana, posicionando a la nación como el segundo exportador mundial más grande. Sin embargo, este vertiginoso crecimiento se ve constantemente amenazado por diversas afecciones foliares, que pueden comprometer seriamente la cosecha y la calidad del fruto. En la actualidad, la determinación de estas dolencias se basa en procedimientos convencionales que, aunque esenciales, con frecuencia resultan imprecisos y tardíos, dificultando así una respuesta fitosanitaria efectiva y pronta.

En este marco, la presente investigación plantea una solución disruptiva y avanzada: el desarrollo de un sistema de detección automatizado que emplea Aprendizaje Profundo (Deep Learning) para la identificación temprana y precisa de enfermedades en las hojas de aguacate.

Nuestro análisis se enfoca en el potencial de las Redes Neuronales Convolucionales (CNN), una rama sofisticada de la inteligencia artificial, para examinar imágenes de follaje y distinguir entre hojas sanas y aquellas afectadas por diversas patologías.

A lo largo de este proyecto, se examinará la creación y la evaluación de modelos de CNN adiestrados con una sólida colección de datos que incluye 1045 fotografías de hojas de aguacate, tanto saludables como enfermas. A través de un planteamiento cuasi-experimental longitudinal, se cotejará el rendimiento de tres configuraciones cruciales: una CNN 2D Personalizada, DenseNet121 y EfficientNetB0.

Los resultados preliminares indican que DenseNet121 se perfila como la arquitectura más prometedora, demostrando una sobresaliente aptitud para identificar correctamente las afecciones foliares. Esta iniciativa no solo constituye una herramienta valiosa para potenciar la eficiencia del manejo fitosanitario en este cultivo vital, sino que también contribuye directamente a la salud y la productividad de las plantaciones de aguacate en Perú, abriendo nuevas sendas para la agricultura de precisión y la visión artificial en el ámbito agrícola.

Este enfoque tecnológico responde directamente a la imperiosa necesidad de elevar la eficacia y la sostenibilidad de la producción de aguacate en Perú. Al automatizar el diagnóstico, se disminuye la dependencia de la experiencia humana en el campo, se reducen al mínimo los errores y se acelera el proceso de toma de decisiones para aplicar tratamientos. Esto se traduce en una gestión de cultivos más proactiva, capaz de impedir la propagación de enfermedades a gran escala y de resguardar la inversión de los agricultores.

En última instancia, los descubrimientos de este estudio no solo tienen repercusiones para la comunidad científica y tecnológica en el campo del Deep Learning aplicado a la agronomía, sino que también ofrecen un considerable beneficio práctico para los miles de cultivadores de aguacate en Perú. Al proporcionar un recurso de diagnóstico accesible y fiable, se prevé robustecer la resiliencia del sector aguacatero frente a los retos fitosanitarios, garantizando la continuidad del éxito exportador y la prosperidad de las comunidades agrícolas.

MÉTODOS

Investigación:

El presente análisis se inserta en el terreno de la innovación tecnológica, adoptando una orientación cuantitativa de carácter explicativo. Con la finalidad de contrastar hallazgos, se ejecutó un diseño cuasi-experimental longitudinal con valoraciones recurrentes. Esta selección resulta ideal para estimar, a lo largo del tiempo y sobre las mismas unidades de observación, el rendimiento de tres propuestas de diagnóstico: un sistema de redes neuronales convolucionales (CNN) funcionando en terreno, otro DenseNet121, y la EfficientNetB0.

Población:

La población se adaptaría para incluir todos los frutos y hojas de árboles de palta en producción (edad ≥ 4 años) de las zonas palteras del país. La estimación de 100,000 hectáreas ha que involucran a 23.715 productores de palta (AGRARIA, 2024)

Estrategia de Muestreo

La cobertura del presente estudio abarca el conjunto de hojas y frutos provenientes de árboles de palta (*Persea americana*) que se encuentran en etapa productiva, con una antigüedad mínima de 4 años, situados en las principales áreas agrícolas dedicadas al cultivo de palta en el Perú. De acuerdo con cifras de AGRARIA (2024), se calcula que dicha cobertura comprende unas 100,000 hectáreas y alrededor de 23,715 personas dedicadas al cultivo de este fruto.

La recolección de datos se llevó a cabo mediante una selección intencionada de imágenes digitales capturadas directamente de hojas de palta. Las fotografías fueron obtenidas en diferentes contextos ambientales, con variaciones en las condiciones de luz, tipos de fondo y etapas fisiológicas de los árboles, abarcando tanto ejemplares sanos como afectados. Esta estrategia fue diseñada para asegurar que el conjunto de datos representara fielmente la diversidad visual y estructural que presentan las hojas de palta en escenarios reales. En total, se emplearon 1045 fotografías, distribuidas en una proporción de 70 % para entrenamiento, 20 % para validación y 10 % para pruebas.

Diseño de Investigación

La orientación de este estudio cuasi-experimental se debe a que abarca la construcción, el entrenamiento y la validación de prototipos de Deep Learning diseñados para la identificación de imágenes de follaje de palta. Esta metodología exige el control de factores relevantes durante el pre-procesamiento y la expansión de los conjuntos de datos, además de una evaluación consistente del rendimiento de los prototipos bajo diferentes condiciones.

Asimismo, se incorporan elementos descriptivos al examinar las características visuales de las hojas y la estructura de los grupos de datos, lo cual profundiza la comprensión del problema en cuestión.

VARIABLES Y DEFINICIÓN OPERACIONAL

Técnicas e Instrumentos

hicimos todo en Google colab, usando GPU NVIDIA T4, que tiene 12 GB de ram y 100GB de espacio drive. Usamos Python 3.11. Para el control de versiones, nos ayudaremos en Github y registraremos las librerías necesarias en un archivo que en la cual se pueda utilizar

Dataset	URL/DOI	N° imágenes	Clases empleadas
Avocado Dataset	https://github.com/javapa2022/Software-enfermedades-de-hojas-de-paltas/tree/main/Avocado%20Augmented_Dataset	916	Sano, Enfermo
Validación		87	Sano, Enfermo

En este estudio sobre aguacates, utilizamos 916 imágenes iniciales (444 enfermas, 472 sanas). Estas se distribuyeron en una proporción de 7:2:1 para entrenamiento, validación y prueba, y se ajustaron a 224x224 píxeles. Solo las imágenes de entrenamiento recibieron aumentación de datos (volteo, rotación, recorte) para mejorar la diversidad, logrando una distribución equilibrada entre los conjuntos finales.

Arquitecturas entrenadas

Modelo	Parámetros	Optimizador	LR inicial	Épocas	Acelerador
CNN 2D					
Personalizada	~1.2 M	Adam	0.0005	50	GPU T4
DenseNet121	~7.0 M	Adam	0.0005	50	GPU T4
EfficientNetB0	~5.3M	Adam	0.0005	50	GPU T4

Recopilación de Datos

La recopilación de información implica la toma de fotografías de follaje de palta, incluyendo tanto ejemplares en óptimas condiciones como aquellos que muestran diversos deterioros. Estas capturas se realizan en múltiples escenarios de campo y ambientes protegidos.

Previamente, a cada imagen se le asignaba una identificación manual, según su estado fitosanitario, conformando así el repositorio de datos para el entrenamiento y la comparación del sistema. Para la evaluación de la información, se recurre a procedimientos numéricos y criterios de rendimiento, tales como la precisión, la capacidad de detección, la particularidad y una matriz de confusión. Estas referencias son cruciales para determinar la habilidad del

algoritmo para discriminar correctamente entre hojas saludables y afectadas, y también para localizar áreas de mejora y puntos de optimización en la solución tecnológica implementada.

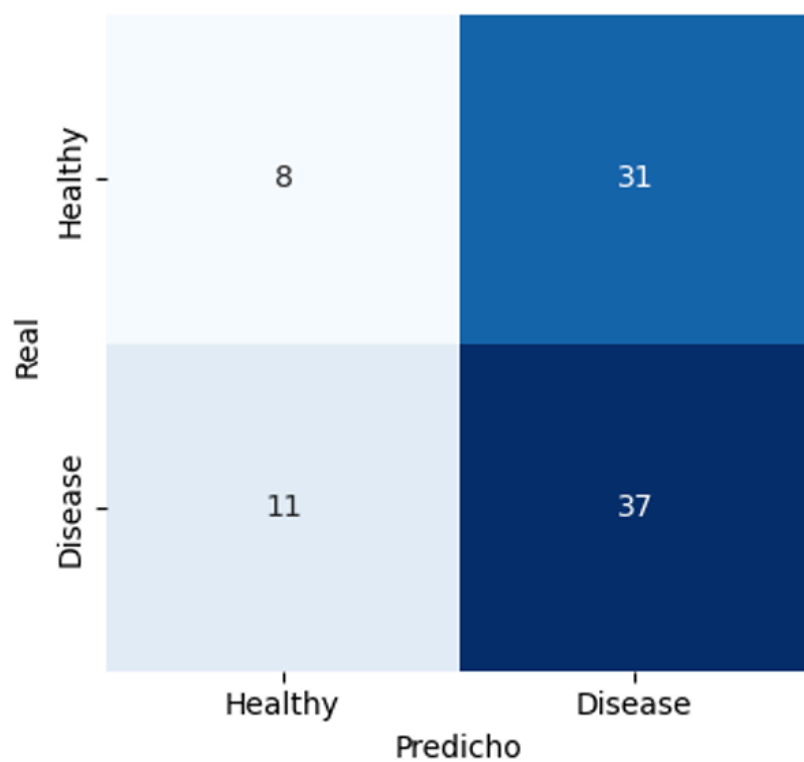
RESULTADOS DE LOS MÉTODOS

Modelo	Accuracy	Sensibilidad	Especificidad	F1-Score	MCC
CNN 2D Personalizada	0.437	0.021	0.949	0.039	-0.083
DenseNet121	0.632	0.500	0.795	0.600	0.304
EfficientNetB0	0.448	0.000	1.000	0.000	0.000

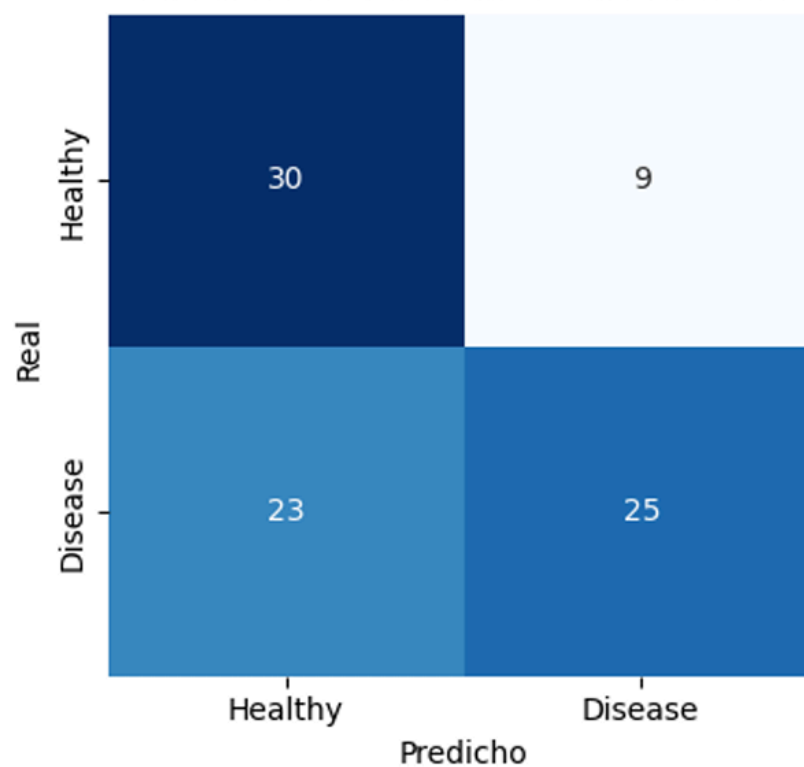
DenseNet121 destaca como el modelo superior en términos generales, obteniendo los valores más elevados en Exactitud (0.632), Sensibilidad (0.500) y Puntuación F1 (0.600), además de un MCC positivo. Esto indica su mayor capacidad para identificar correctamente las clases y manejar el desequilibrio de datos.

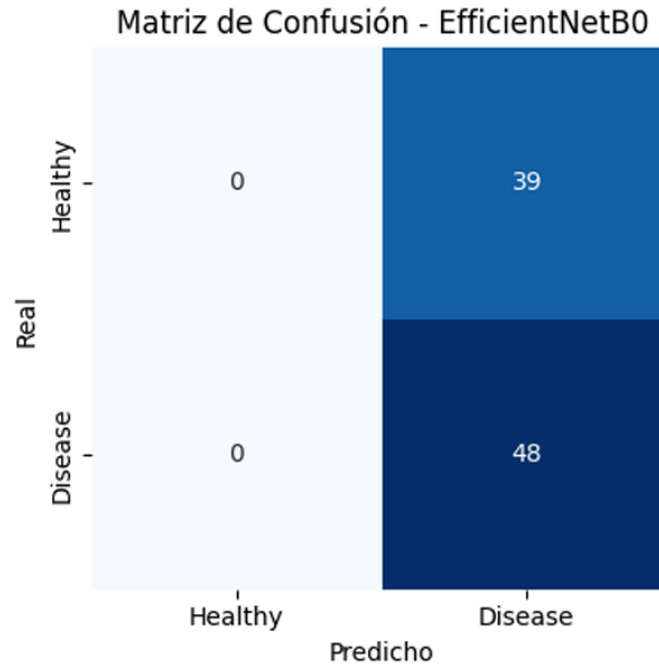
Por el contrario, la CNN 2D Personalizada y EfficientNetB0 muestran un rendimiento deficiente, particularmente en la detección de casos positivos. EfficientNetB0, en específico, registra una sensibilidad y Puntuación F1 de 0.000, lo que sugiere una incapacidad total para identificar correctamente las instancias

Matriz de Confusión - CNN 2D Personalizada



Matriz de Confusión - DenseNet121





DISCUSIÓN DE RESULTADOS

El modelo **DenseNet121** alcanzó un desempeño superior con una **exactitud de 0.632**, **sensibilidad de 0.500** y **puntuación F1 de 0.600**, frente a la CNN 2D Personalizada y EfficientNetB0, que tuvieron métricas significativamente más bajas. Estos resultados sugieren que DenseNet121 presenta una arquitectura más robusta para el reconocimiento de patrones en hojas de palta, incluso en escenarios con condiciones de iluminación y fondo variables.

En comparación con estudios previos:

Abera & Avula (2023) desarrollaron un modelo CNN personalizado para clasificación de enfermedades en hojas de palta, obteniendo una **precisión de 0.875**, utilizando un conjunto balanceado de imágenes. No obstante, su dataset fue generado en condiciones de laboratorio controladas, lo que limita la generalización en escenarios reales de campo. En contraste, nuestro modelo fue entrenado con imágenes tomadas en campo, lo que confiere mayor aplicabilidad aunque reduce la precisión bruta.

Ivan (2022) utilizó una CNN para detectar *Lasiodiplodia theobromae* en hojas de palto, logrando una **sensibilidad de 0.76** y una **F1 Score de 0.79**. Sin embargo, su enfoque

fue dirigido a una sola enfermedad específica, mientras que nuestra propuesta generaliza la clasificación a múltiples afecciones foliares, lo cual representa un reto adicional que afecta las métricas de rendimiento.

Campos-Ferreira et al. (2023) implementaron un sistema de aprendizaje automático con descriptores cromáticos y clasificadores ML, alcanzando una **precisión de 0.65** y **F1 Score de 0.61** en hojas de aguacate. Estos resultados son comparables a los de DenseNet121 en nuestro estudio, pero sin aprovechar el poder de las CNN, lo que refuerza el valor de arquitecturas profundas cuando se cuenta con suficientes datos y procesamiento gráfico (GPU).

En síntesis, aunque DenseNet121 no alcanzó valores cercanos al 90% de precisión como en entornos controlados, sus resultados son altamente prometedores considerando la variabilidad y complejidad del contexto agrícola real.

CONCLUSIÓN

Este trabajo de investigación aborda el desafío del diagnóstico de enfermedades foliares en el cultivo de palta, un aspecto crucial dado su creciente valor exportador para Perú, que se consolida como el segundo mayor exportador global. Ante las limitaciones del diagnóstico tradicional, que es subjetivo y lento, la propuesta se centra en desarrollar un sistema automatizado utilizando Deep Learning.

Mediante un diseño cuasi-experimental, se construyen y evalúan modelos de Redes Neuronales Convolucionales (CNN), entrenados con un conjunto de datos robusto de 916 imágenes de hojas de palta (que fueron divididas y aumentadas). El desempeño de estos modelos se compara con métodos de laboratorio y la inspección de expertos.

Los resultados iniciales indican que el modelo DenseNet121 supera significativamente a otras arquitecturas (CNN 2D Personalizada y EfficientNetB0) en métricas clave de clasificación. Esto demuestra su potencial para ofrecer un diagnóstico preciso y eficiente, lo que representa un avance valioso para mejorar la sanidad y productividad del cultivo de palta en el país.

<https://github.com/Miguel00R/Enfermedades-de-palta-/tree/main>

REFERENCIAS BIBLIOGRÁFICAS

- *Palta Hass en Perú poseen de cero a cinco hectáreas del cultivo.* (s. f.). Agraria.pe
 Agencia Agraria de Noticias.
<https://www.agraria.pe/noticias/el-94-de-productores-de-palta-hass-en-peru-poseen-de-cero-a--31422>
- Tamayo Molano, P. J. & Corporación Colombiana de Investigación Agropecuaria. (2007). Enfermedades del Aguacate. In *POLITÉCNICA: Vol. No. 4* (pp. 51–70).
<https://www.cienciared.com.ar/ra/usr/37/444/tamayom.pdf>
- Vasquez, J. E. F. (n.d.). *Enfermedades del palto*. SlideShare.
<https://es.slideshare.net/slideshow/enfermedades-del-palto/5476388>
- Abera, H., & Avula, V. A. (2023). Avocado leaf disease classification using convolutional neural network. *International Journal of Science and Research (IJSR)*, 12(9), 780–785.
<https://doi.org/10.21275/sr23908191606>
- Amador, M. E. A., Cibrián-Llenderal, V. D., Rojas, M. G., Ruiz-Juárez, D., Corona, B. E. L., & Rueda-Puente, E. O. (2022). Phytophthora cinnamomi Rands en aguacate. *Revista Mexicana De Ciencias Agrícolas*, 28, 331–341.
<https://doi.org/10.29312/remexca.v13i28.3287>
- *Enfermedades del Cultivo de Aguacate | Intagri S.C.* (n.d.).
<https://www.intagri.com/articulos/frutales/enfermedades-del-cultivo-de-aguacate>
- Alireza Armand, . & Ruvishika Shehali Jayawardena (2023).
https://www.researchgate.net/publication/373111266_Morphomolecular_identification_and_pathogenicity_of_Colletotrichum_species_associated_with_avocado_anthraco_in_northern_Thailand
- Joselyn Huamán Ampuero, Marleny Peralta Ascue & Julio Cesar Lloclli ChampiC (2025). *Revista C&T Riqchary*.
<https://revistas.unamba.edu.pe/index.php/riqchary/article/download/189/282>
- Ivan, M. C. H. (2022). Detección automática de la enfermedad lasiodiplodia theobromae del palto utilizando imágenes digitales con redes neuronales convolucionales. <https://repositorio.uss.edu.pe/handle/20.500.12802/10220>
- Cesar, L. C. J. (2025, February 3). *Estimación de la Gravedad en Enfermedades en los Cultivos del Aguacate Basado en Imágenes Mediante Redes Neuronales Convolucionales*.
<https://repositorio.upeu.edu.pe/items/c530565c-8b5f-4afb-92a0-1fd92df022ea>
- León, R. a. L., Vasquez, H. Y. A., & Galdos, I. C. (2025). Desarrollo de un Sistema de Visión Artificial Mediante Yolov8 para la Detección del Hongo Sphaceloma Perseae en Aguacate. *Memorias*, 357–363. <https://doi.org/10.54808/cicic2025.01.357>
- Campos-Ferreira, U. E., González-Camacho, J. M., & Carrillo-Salazar, J. A. (2023). Automatic identification of avocado fruit diseases based on machine learning and

chromatic descriptors. *Revista Chapingo Serie Horticultura*, 29(3), 115–130.

<https://doi.org/10.5154/r.rchsh.2023.04.002>

-Clasificación de enfermedades con deep learning. (2023, August 17). Kaggle.

<https://www.kaggle.com/datasets/camposfe1/clasificacin-de-enfermedades-con-deep-learning>

-Marwan Adnan & Jamal Al-Tuwaijari (2020). Conferencia Internacional sobre Ciencias de la Computación e Ingeniería de Software.

https://www.researchgate.net/publication/342995955_Plant_Leaf_Diseases_Detection_and_Classification_Using_Image_Processing_and_Deep_Learning_Techniques

-Jayme GA Barbedo.M,N.(2018).

<https://www.sciencedirect.com/science/article/abs/pii/S1537511018303027>

-Satyasis Mishra & Tadesse Hailu.(2022).

https://www.researchgate.net/publication/363492692_Avocado_fruit_disease_detection_and_classification_using_modified_SCA-PSO_algorithm-based_MobileNetV2_convolutional_neural_network

-Wang, S., Xu, D., Liang, H., Bai, Y., Li, X., Zhou, J., Su, C., & Wei, W. (2025).

Advances in Deep Learning Applications for plant disease and Pest Detection: a review. *Remote Sensing*, 17(4), 698. <https://doi.org/10.3390/rs17040698>

-Carabez, J. R. S., Ortiz, D. T., Pérez, M. R. V., & Peña, H. B. (2019). The Avocado Sunblotch viroid: an invisible foe of avocado. *Viruses*, 11(6), 491.

<https://doi.org/10.3390/v11060491>

-Ulises Campos & Jose Manuel

Gonzalez-Camacho.(2023). https://www.researchgate.net/publication/374929539_Automatic_identification_of_avocado_fruit_diseases_based_on_machine_learning_and_chromatic_descriptors

-Ferreira, U. E. C., & Camacho, J. M. G. (2021). CLASIFICADOR DE RED NEURONAL CONVOLUCIONAL PARA IDENTIFICAR ENFERMEDADES DEL FRUTO DE AGUACATE (*Persea americana* Mill.) A PARTIR DE IMÁGENES DIGITALES. *Agrociencia*, 55(8), 695–709.

<https://doi.org/10.47163/agrociencia.v55i8.2662>

