

NOMBRE: Oswaldo Rivero Romero
CC.: 1.065.009.838

1. TÍTULO: Clasificación de hojas de tomate a partir de imágenes de área foliar con afectaciones de diversas enfermedades.

1. INTRODUCCIÓN

La producción agrícola es un pilar fundamental de la economía global, ya que representa la principal fuente de alimentos para la humanidad. En numerosas regiones, especialmente en los países en desarrollo, la agricultura juega un papel crucial en el desarrollo económico, generando empleo, promoviendo la producción de materias primas y contribuyendo significativamente al Producto Interno Bruto (PIB) [1]. En este contexto, la implementación de estrategias que optimicen los resultados agrícolas mediante el monitoreo continuo de los cultivos puede resultar de gran valor, proporcionando recomendaciones y consejos acertados para mejorar el rendimiento de estos. Por lo tanto, la implementación de tecnologías informáticas y la inteligencia artificial (IA) pueden ser una técnica viable para el rendimiento agrícola [2].

El tomate (*Solanum lycopersicum* L.) es uno de los cultivos más populares a nivel mundial, siendo superado únicamente por la papa [3]. Este fruto se utiliza en una amplia variedad de alimentos, desde ensaladas hasta salsas, e incluso como producto natural en el cuidado de la piel [4]. Sin embargo, uno de los principales desafíos que enfrentan los agricultores es la aparición de enfermedades foliares, las cuales pueden reducir el tamaño de los frutos e incluso provocar la muerte de las plantas. Además, el tiempo necesario para diagnosticar dichas enfermedades y aplicar medidas correctivas afecta negativamente el rendimiento de los cultivos. Por lo tanto, en este trabajo, a partir de un *dataset* de imágenes disponibles de dominio público se pretendió emplear modelo basado en redes neuronales convolucionales (CNN) para la clasificación de hojas de tomate a partir de imágenes de área foliar con afectaciones de diversas enfermedades.

2. OBJETIVO DE MACHINE LEARNING

Considerando el contexto previamente expuesto y la base de datos proporcionada por el conjunto de datos de *Tomato leaf disease detection* disponible en Kaggle.com, el objetivo propuesto es el siguiente:

- "Crear un modelo de machine learning utilizando redes neuronales convolucionales (CNN) para clasificar de hojas de tomate a partir de imágenes de área foliar con afectaciones de diversas enfermedades.

Disponible en <https://www.kaggle.com/datasets/kaustubhb999/tomatoleaf/data>

3. DATASET

Tipo de datos: Imágenes etiquetadas y anotadas que incluyen información sobre la enfermedad o el problema específico que afecta a la planta.

Tamaño: El conjunto de datos de *Tomatoleaf* contiene 10,000 imágenes de hojas enfermas, organizadas en 10 clases según la enfermedad. Su tamaño total en disco es de 186.22 MB. Además, se cuenta con un total de 1000 imágenes para la validación del modelo, con 100 imágenes por cada clase. La distribución de las clases (10) por enfermedad se detalla en la siguiente tabla:

Clases	Imágenes para entrenamiento	Imágenes para validación
Tomato_mosaic_virus	1000	100
Target_Spot	1000	100
Bacterial_spot	1000	100
Tomato_Yellow_Leaf_Curl_Virus	1000	100
Late_blight	1000	100
Leaf_Mold	1000	100
Early_blight	1000	100
Spider_mitesTwo_spotted_spider_mite	1000	100
Tomato___healthy	1000	100
Septoria_leaf_spot	1000	100

4. MÉTRICAS DE DESEMPEÑO

Las métricas utilizadas para evaluar el desempeño del modelo son las siguientes:

- **Accuracy:** Esta métrica tiene como objetivo medir qué tan bien el modelo clasifica las imágenes de entrada en las categorías correctas. Se determina con la siguiente fórmula:

$$Accuracy = \frac{No. de predicciones correctas}{No. total de predicciones}$$

5. REFERENCIAS Y RESULTADOS PREVIOS

El *Deep Learning* ha adquirido un papel fundamental en diversas disciplinas, y su aplicación en agronomía ha dado lugar a desarrollos importantes, especialmente en la detección temprana de enfermedades en plantas. Para abordar este problema, diversos investigadores han propuesto el uso de técnicas avanzadas de inteligencia artificial, como el *Machine Learning* (ML), para ayudar a los agricultores a tomar decisiones oportunas que optimicen la producción de tomate. En un estudio de Lu et al. (2021), se analizaron los problemas más comunes y las soluciones asociadas al uso de redes neuronales convolucionales (CNN) en la clasificación de enfermedades vegetales. Los autores describieron los métodos implementados con CNN para abordar esta clasificación, proporcionando una visión general de los principios de *Deep Learning* aplicados específicamente a este tipo de problemas en el ámbito agronómico [5].

En el ámbito de los cultivos de tomate, se ha demostrado que las técnicas avanzadas de procesamiento de imágenes pueden ser altamente efectivas para la detección temprana de enfermedades. Por ejemplo, Ashok et al. (2020) desarrollaron un sistema que emplea redes neuronales convolucionales (CNN) para identificar enfermedades en las hojas de tomate, logrando una precisión del 98% [6]. Además, un estudio reciente utilizó filtros de preprocesamiento como *Gaussian Blur* y *Gaussian Noise*, junto con modelos de color como HSI y CMYK, para mejorar la precisión en la clasificación de imágenes de hojas. Este trabajo evaluó modelos de redes neuronales convolucionales profundas (DCNN) como *Vgg-19*, *MobileNet-V2* y *ResNet-50*, alcanzando una precisión máxima de 99.53% con *ResNet-50*, al combinar filtros de ruido gaussiano con la conversión de color de RGB a CMYK. Estos resultados subrayan el gran potencial de estas metodologías para su aplicación en el ámbito agrícola [7].

6. REFERENCIAS

- [1] Falaschetti, L., Manoni, L., Di Leo, D., Pau, D., Tomaselli, V., & Turchetti, C. (2022). A CNN-based image detector for plant leaf diseases classification. *HardwareX*, 12. <https://doi.org/10.17605/OSF.IO/UCM8D>
- [2] Geetharamani, G., & Arun Pandian, J. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional neural network. *Computers and Electrical Engineering*, 76, 323–338. <https://doi.org/10.1016/j.compeleceng.2019.04.011>
- [3] National Plant Data Center, NRCS, USDA. (2023). *The scientific name of Tomato*. The plants database (version 5.1.1). Baton Rouge, LA 70874-4490 USA. <http://plants.usda.gov/Reference> (Accessed 13 November 2023). For: *Solanum lycopersicum* var. *lycopersicum*.
- [4] National Plant Data Center, NRCS, USDA. (2023). *The scientific name of Potato*. The plants database (version 5.1.1). Baton Rouge, LA 70874-4490 USA. <http://plants.usda.gov/Reference> (Accessed 13 November 2023). For: *Solanum tuberosum*.
- [5] Lu, J., Tan, L., & Jiang, H. (2021). Review on convolutional neural network (CNN) applied to plant leaf disease classification. *Agriculture (Switzerland)*, 11(8). <https://doi.org/10.3390/agriculture11080707>
- [6] Ashok, S., Kishore, G., Rajesh, V., Suchitra, S., Sophia, S. G. G., & Pavithra, B. (2020). Tomato leaf disease detection using deep learning techniques. *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, 979-983. <https://doi.org/10.1109/ICCES48766.2020.9137986>.
- [7] Hossain, M. I., Jahan, S., Al Asif, M. R., Samsuddoha, M., & Ahmed, K. (2023). Detecting tomato leaf diseases by image processing through deep convolutional neural networks. **Smart Agricultural Technology*, 5*, 100301. <https://doi.org/10.1016/j.atech.2023.100301>.