

Proyecto Final

DIAGNÓSTICO AUTOMATIZADO DE ENFERMEDADES EN TOMATES MEDIANTE DEEP LEARNING: UN ENFOQUE DE CLASIFICACIÓN DE IMÁGENES

Hector Navarro Barboza

Curso: Data Science

23 de octubre de 2023

IT ACADEMY

Resumen

En este trabajo, se investigaron dos estrategias de clasificación para identificar diversas enfermedades en hojas de tomate, categorizadas en 10 clases, donde 9 representan diferentes patologías y una el estado saludable. Se diseñó un modelo desde cero con capas convolucionales y se adaptó un modelo preentrenado, ResNet50, descongelando ciertas capas para el aprendizaje afinado. La iniciativa procuró comparar la eficacia y precisión entre ambos modelos en el diagnóstico automatizado, utilizando un conjunto de datos de Kaggle compuesto por imágenes etiquetadas que facilitaron una evaluación robusta en distintas condiciones y estados de enfermedad. El entrenamiento del modelo simplificado, más rápido con 215 minutos, contrastó con los 595 minutos del ResNet50, no obstante, mostró una superioridad significativa con una precisión, F1-Score y Recall del 95.9 %, 95.9 % y 95.7 % respectivamente, frente al 88.7 %, 88.7 % y 86.8 % de ResNet50. Las matrices de confusión revelaron un desempeño superior del modelo simplificado en la mayoría de las clases, mientras que ResNet50 exhibió más errores de clasificación, especialmente en clases como Early Blight, Septoria Leaf Spot, Target Spot y Tomato Mosaic Virus, sugiriendo la necesidad de más datos o ajustes adicionales para mejorar su rendimiento. A pesar de la complejidad de ResNet50, no superó al modelo simplificado, resaltando la importancia de comprender la naturaleza del conjunto de datos y cómo diferentes arquitecturas responden ante éste. Los hallazgos enfatizan la viabilidad de aplicar modelos de clasificación de imágenes en la agricultura para la detección temprana de enfermedades en plantas y la importancia de seleccionar y optimizar la arquitectura del modelo en función de las características del dataset. La superioridad del modelo simplificado sugiere que explorar arquitecturas avanzadas, junto con técnicas de augmentación de datos y optimización, podría conducir a mejoras significativas en futuros proyectos.

ÍNDICE

1. Introducción	4
2. Metodología	5
2.1. Dataset	5
2.2. Modelos de Aprendizaje profundo	6
2.2.1. Modelo de Clasificación Simplificado	6
2.2.2. Modelo Preentrenado ResNet50	7
3. Resultados	7
3.1. Modelo de Clasificación Simplificado	8
3.2. Modelo Preentrenado ResNet50	10
4. Conclusiones	11
Referencias	13

INTRODUCCIÓN

La clasificación automatizada de imágenes para la detección de enfermedades en plantas, especialmente en cultivos de tomate, se ha convertido en un área de investigación prometedora, facilitada por los avances recientes en Redes Neuronales Profundas (Deep Neural Networks) y el aprendizaje profundo (Deep Learning). La visión por computadora, combinada con modelos de aprendizaje profundo, ha demostrado ser eficaz para detectar y clasificar diversas enfermedades en las hojas de tomate, lo que contribuye a un diagnóstico temprano y, por ende, a una intervención oportuna para mitigar pérdidas en la producción agrícola (Fuentes et al., 2017; Alhaj Ali et al., 2021).

El enfoque general de estas metodologías se centra en aplicar técnicas de aprendizaje profundo a problemas de visión por computadora específicos, como la detección y clasificación de enfermedades en tomates mediante el procesamiento de imágenes digitales (Valenzuela Cámara, 2022). La adopción de modelos de aprendizaje profundo, como las redes convolucionales, ha permitido mejorar drásticamente la precisión de los sistemas de detección y reconocimiento de objetos, incluyendo la identificación de enfermedades y plagas en plantas de tomate utilizando imágenes capturadas en tiempo real por dispositivos con cámara de diversas resoluciones (Fuentes et al., 2017).

En algunos trabajos, se ha explorado la identificación de enfermedades en etapas muy tempranas para evitar pérdidas económicas significativas en la economía agrícola. Además, se ha presentado el aprendizaje por transferencia como una técnica crucial para acelerar la tarea de entrenamiento y aumentar la precisión y estabilidad de los modelos de clasificación, lo que resulta esencial para una detección eficaz y oportuna de enfermedades (Alhaj Ali et al., 2021).

En otros estudios, se ha llevado la detección de enfermedades un paso más allá, explorando la implementación de algoritmos de aprendizaje profundo en robots que pueden recorrer campos de tomate o invernaderos en tiempo real, identificando enfermedades en las plantas mientras se desplazan por el área (Durmuş et al., 2017). Esta innovación no solo permite un monitoreo constante sino que también facilita una respuesta rápida ante cualquier signo de enfermedad, lo que podría ser crucial para preservar la salud de los cultivos y, en última instancia, la producción agrícola.

En el presente trabajo, se exploraron dos modelos de clasificación distintos con el propósito de identificar enfermedades en las hojas de los tomates, considerando 10 clases, donde 9 representan enfermedades y una el estado saludable. El primer modelo fue desarrollado desde cero, compuesto por varias capas convolucionales, mientras que el segundo empleó un modelo preentrenado, específicamente ResNet50, en el cual se descongelaron ciertas capas para facilitar el aprendizaje afinado (fine-tuning). Esta estrategia dual tuvo como objetivo investigar y comparar la eficacia y precisión entre los modelos construidos desde cero y los modelos preentrenados.

nados en la tarea de diagnóstico automatizado. Para el entrenamiento y validación de estos modelos, se utilizó un conjunto de datos disponible en Kaggle, que proporcionó una amplia gama de imágenes etiquetadas, permitiendo así una evaluación robusta de los modelos en diferentes condiciones y estados de enfermedad en las hojas de tomate.

2

METODOLOGÍA

2.1 DATASET

Para el entrenamiento y validación de los modelos se utilizó un subset específico del dataset disponible en Kaggle denominado PlantVillage Dataset [PlantVillage Dataset](#). Este dataset es conocido por su amplia colección de imágenes de plantas bajo diversas condiciones de salud y enfermedad. Sin embargo, para los fines de este proyecto, se extrajo y utilizó exclusivamente la porción del dataset correspondiente a las plantas de tomate. Esta selección incluye imágenes categorizadas en diez clases distintas, que representan diversas condiciones de salud y tipos de enfermedades que afectan a las hojas de tomate.

La Figura 1 muestra las clases específicas consideradas: Septoria Leaf Spot, Spider Mites Two-spotted Spider Mite, Leaf Mold, Early Blight, Tomato Yellow Leaf Curl Virus, Tomato Mosaic Virus, Target Spot, Bacterial Spot, Late Blight y Healthy.

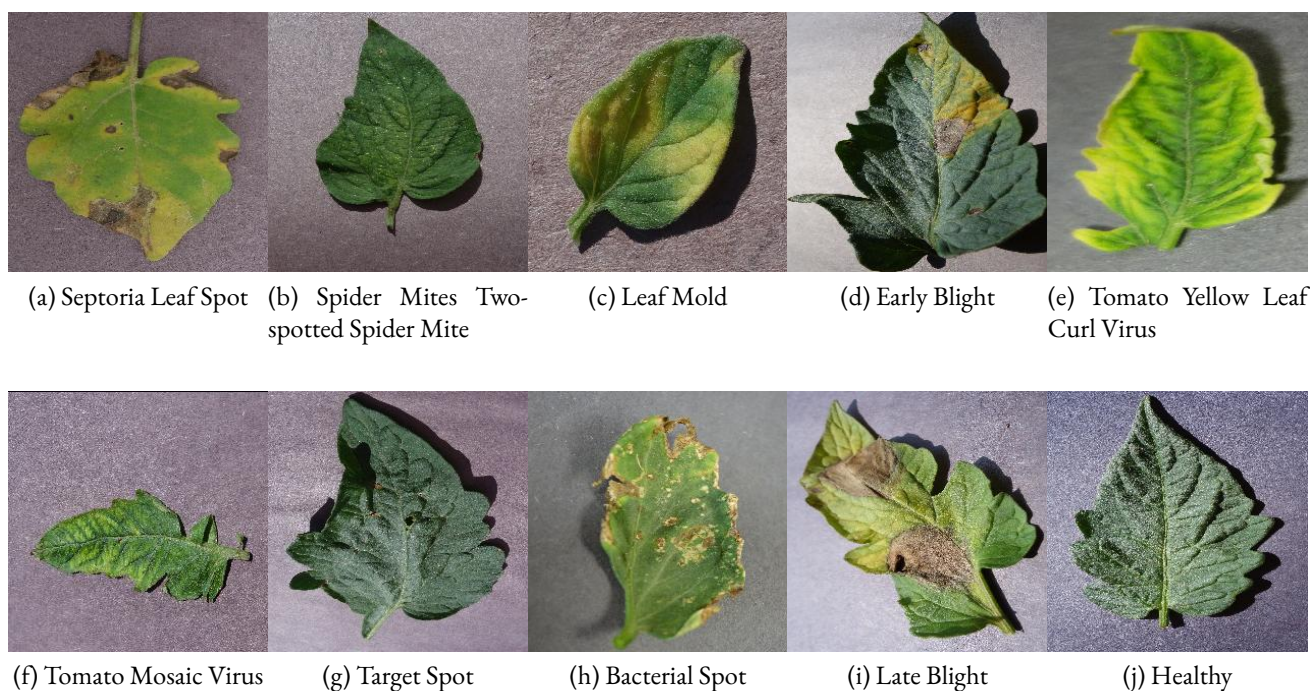


Figura 1: Representación visual de las clases de enfermedades y condiciones de salud en hojas de tomate.

La categorización en estas diez clases permite a los modelos desarrollados aprender y diferenciar entre una variedad de condiciones patológicas y el estado saludable de las hojas de tomate. El dataset proporciona una base de datos robusta y diversificada, facilitando la tarea de entrenamiento supervisado para los modelos. La distribución de las imágenes en estas clases específicas facilita la evaluación del desempeño del modelo en términos de su capacidad para identificar y clasificar correctamente las enfermedades, así como distinguir las hojas saludables de las enfermas.

Para el entrenamiento de los modelos, se ha dispuesto de un total de 18,345 imágenes, mientras que para la fase de validación se han utilizado 4,585 imágenes. Esta separación permite un entrenamiento robusto y una validación precisa para garantizar que los modelos desarrollados tengan una capacidad generalizable y confiable para diagnosticar las condiciones de las hojas de tomate a partir de imágenes.

2.2 MODELOS DE APRENDIZAJE PROFUNDO

2.2.1. Modelo de Clasificación Simplificado

Se emplea un modelo de clasificación básico construido utilizando la biblioteca TensorFlow/Keras, diseñado para identificar enfermedades en las hojas de tomate a partir de imágenes. El modelo es de tipo secuencial, lo que significa que las capas están ordenadas de manera secuencial, donde cada capa recibe la entrada de la capa anterior. El modelo se compone de tres bloques convolucionales, cada uno con una capa convolucional seguida de una capa de agrupación máxima (MaxPooling). La primera capa convolucional tiene 16 filtros, la segunda tiene 32 y la tercera tiene 64, todas con un tamaño de kernel de 3x3 y función de activación ReLU. Tras las capas convolucionales, la red se aplanar y se conecta a una capa densamente conectada con 128 neuronas y función de activación ReLU, seguida por la capa de salida con una cantidad de neuronas igual al número de clases y función de activación softmax, que proporciona las probabilidades de cada clase. El modelo se compila utilizando el optimizador Adam, una función de pérdida de entropía cruzada categórica y métricas de precisión categórica, F1-Score y Recall para evaluar su rendimiento durante el entrenamiento y la validación.

Se opta por una estructura convolucional debido a la eficacia comprobada de las redes convolucionales (ConvNets) en tareas de clasificación de imágenes. La elección de la función de activación ReLU (Rectified Linear Unit) se debe a su capacidad para reducir el problema del desvanecimiento del gradiente, permitiendo un entrenamiento más rápido y eficiente. La capa de agrupación máxima (MaxPooling) contribuye a reducir la dimensionalidad de la entrada y a conservar las características esenciales detectadas por los filtros convolucionales. La capa densa antes de la capa de salida actúa como un clasificador, mientras que la capa de salida con activación softmax proporciona una distribución de probabilidad sobre las clases de enfermedades. El uso del optimizador Adam, conocido por su eficiencia y rendimiento en la optimización de problemas de deep learning, facilita la convergencia rápida durante el entrenamiento.

2.2.2. Modelo Preentrenado ResNet50

En el marco de este proyecto, se ha optado por implementar un modelo preentrenado ResNet50, aprovechando las capacidades avanzadas que este tipo de modelos ofrece para la clasificación de imágenes. ResNet50, un modelo preentrenado proporcionado por la biblioteca TensorFlow/Keras, ha sido entrenado previamente en el dataset ImageNet, lo que le permite reconocer una amplia gama de características en imágenes. El modelo ResNet50 se integra como la base del modelo secuencial desarrollado, excluyendo su capa completamente conectada (fully-connected) final para permitir una adaptación específica a la tarea en cuestión. Se procede a descongelar las últimas 20 capas del modelo base, permitiendo que estas capas se entrenen y ajusten a los datos específicos del proyecto. Posteriormente, la salida del modelo base se aplanan y se pasa a través de una capa densa con 128 neuronas y función de activación ReLU, seguida por la capa de salida que tiene una cantidad de neuronas igual al número de clases y utiliza una función de activación softmax. Este modelo se compila utilizando el optimizador Adam, junto con una función de pérdida de entropía cruzada categórica y métricas de precisión categórica, F1-Score y Recall para monitorear su desempeño durante el entrenamiento y la validación.

El empleo de un modelo preentrenado como ResNet50 proporciona una base sólida y avanzada para la clasificación de imágenes. Gracias a su entrenamiento previo en un dataset extenso y diversificado como ImageNet, ResNet50 ya posee una capacidad considerable para reconocer características complejas en imágenes. Al descongelar las últimas 20 capas del modelo, se permite que el modelo ajuste y refine estas características para la tarea específica de clasificación de enfermedades en hojas de tomate. Esta técnica, conocida como "fine-tuning" o ajuste fino, permite transferir el conocimiento adquirido en tareas generales a la tarea específica de interés, lo cual es especialmente útil cuando se dispone de un dataset de tamaño limitado para el entrenamiento. La inclusión de una capa densa adicional antes de la capa de salida permite una transición suave entre las características extraídas por ResNet50 y la clasificación final en las clases de enfermedades. Finalmente, la compilación del modelo con el optimizador Adam y las métricas seleccionadas facilita un entrenamiento eficiente y una evaluación precisa del desempeño del modelo en la tarea de diagnóstico de enfermedades en tomates, aprovechando la robustez y la eficacia probada del modelo preentrenado ResNet50 en tareas de clasificación de imágenes.

RESULTADOS

Los modelos fueron entrenados en una laptop ASUS ZenBook con un sistema operativo Ubuntu 22.04.3 LTS (Jammy Jellyfish) y kernel versión 6.2.0-34-generic. Esta máquina está equipada con un procesador Intel Core i7-1165G7 de 11ª generación, que es un procesador de cuatro núcleos con una velocidad que varía entre

400 MHz y 4700 MHz. La laptop tiene una memoria RAM de 15.31 GiB, de la cual se utilizó el 88,9 % durante el entrenamiento. El almacenamiento principal es proporcionado por un SSD Intel de 476.94 GiB. El código fuente relacionado con este proyecto está disponible en el repositorio de [GitHub](#).

3.1 MODELO DE CLASIFICACIÓN SIMPLIFICADO

El entrenamiento del modelo se llevó a cabo durante un periodo de 215 minutos. Se implementaron callbacks para almacenar el mejor modelo obtenido en cada época, estableciendo un valor de paciencia (patience) de 10. Aunque se programó el entrenamiento para 100 épocas, el modelo alcanzó la convergencia en la época 40.

En la fase de validación, se obtuvieron las siguientes métricas: Pérdida (Loss) de 0.126, Exactitud (Accuracy) de 0.959, Puntuación F1 (F1-Score) de 0.959 y Sensibilidad (Recall) de 0.957. Estos valores indican un desempeño notable del modelo en datos no vistos, con una baja tasa de error y una alta tasa de predicciones correctas, evidenciando un balance entre precisión y sensibilidad.

La figura 2 muestra el performance del modelo a lo largo de las épocas durante el entrenamiento y la validación. Se observa una mejora en la precisión (a) y en la puntuación F1 del modelo (b) tanto en el conjunto de entrenamiento como en el de validación. Esto indica de que el modelo que el modelo esta aprendiendo y generalizando bien a los datos no vistos (validación). No obstante, es importante notar que, a partide de la época 20 aproximadamente, la precisión y la puntuación F1 en el conjunto de validación parecen estabilizarse, lo que podría ser un indicio de que el modelo ha alcanzado su capacidad de aprendizaje.

En la figura 2(c) se muestra la pérdida del modelo, se observa una disminución constante en el conjunto de entrenamiento, lo cual es esperado ya que el modelo está optimizando la función de pérdida. Sin embargo, en el conjunto de validación, la pérdida alcanza su punto más bajo alrededor de la época 20 y luego muestra una ligera tendencia al alza. Esto podría ser un signo de sobreajuste, lo cual implica que el modelo podría estar empezando a memorizar los datos de entrenamiento en lugar de aprender características generalizables. Sería prudente considerar técnicas como la regularización o el abandono (dropout) para mitigar este sobreajuste potencial y mejorar la generalización del modelo.

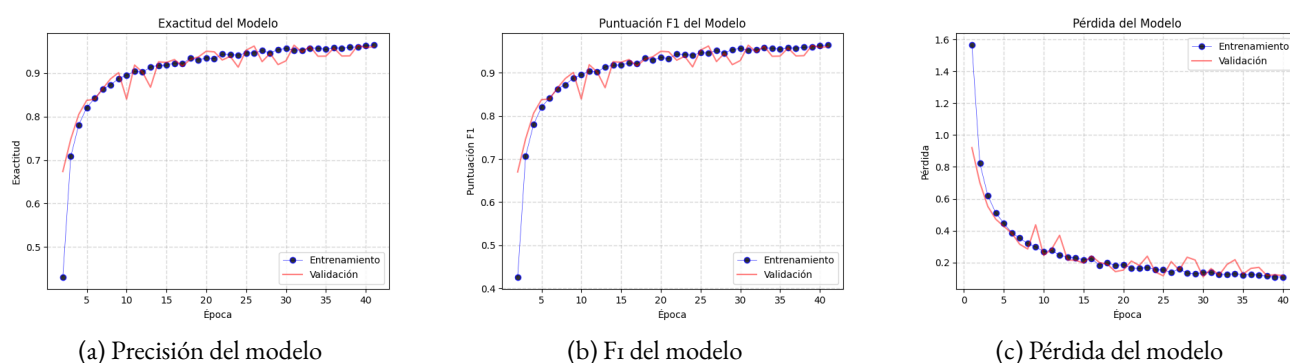


Figura 2: Análisis del rendimiento del modelo a lo largo de las épocas.

En la figura 3 se observa la matriz de confusión para el modelo, el cual muestra un desempeño robusto en la clasificación. Los altos valores en la diagonal principal indican correctas predicciones. Sin embargo, hay ciertas confusiones notables; por ejemplo, *Target_Spot* y *Spider_mites Two-spotted_spider_mite* se confunden entre sí en varias ocasiones. También, hay una confusión menor entre *Early_blight* y *Septoria_leaf_spot*. *Tomato_Yellow_Leaf_Curl_Virus* y *healthy* están muy bien clasificadas con poca confusión.

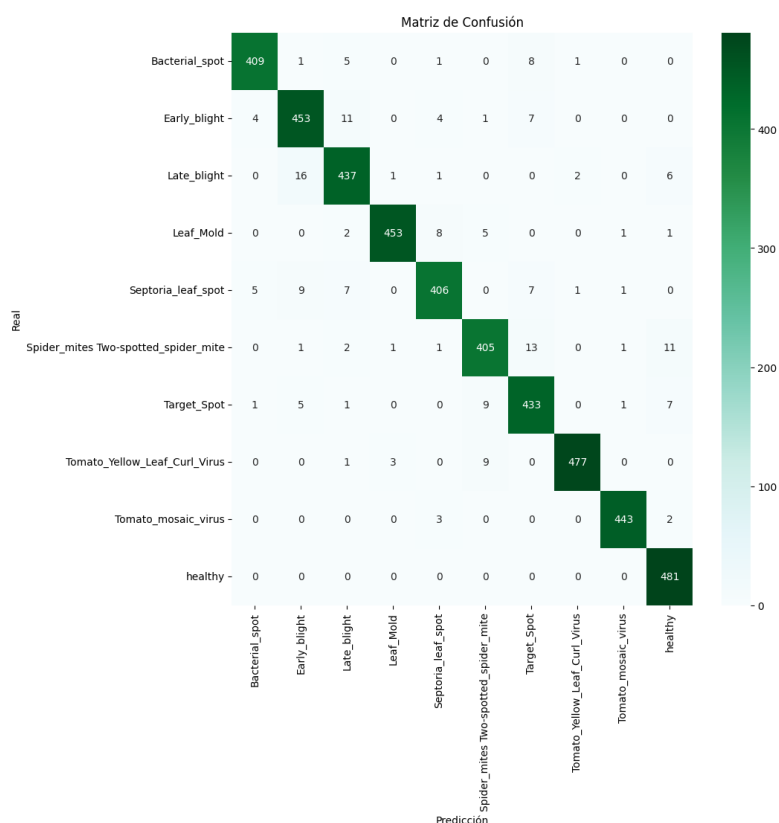


Figura 3: Matriz de Confusión: Representación visual del desempeño del modelo en clasificar correctamente las clases. Los valores en la diagonal principal representan las clasificaciones correctas, mientras que los valores fuera de la diagonal representan las clasificaciones incorrectas.

3.2 MODELO PREENTRENADO RESNET50

El proceso de entrenamiento del modelo fue extenso, abarcando un total de 595 minutos. Durante este proceso, se aplicaron técnicas de callback para guardar el mejor modelo obtenido tras cada época, estableciendo un valor de paciencia de 10 para evitar el sobreajuste y optimizar el proceso de entrenamiento. Aunque se tenía programado ejecutar el entrenamiento a lo largo de 100 épocas, el modelo logró alcanzar la convergencia mucho antes, específicamente en la época 15.

En la fase de validación, se obtuvieron indicadores que reflejan un buen desempeño del modelo. Con una pérdida (Loss) de 0.43, el modelo demostró una precisión (Accuracy) de 0.89. Además, el F1-Score, que es una métrica que combina tanto la precisión como la exhaustividad (recall), fue de 0.89, lo cual es indicativo de un balance adecuado entre estas dos métricas. El valor de Recall obtenido fue de 0.87, lo que sugiere que el modelo tiene una buena capacidad para identificar correctamente las clases positivas.

En la figura 4 se observa que el modelo aprende efectivamente a lo largo de las épocas, mostrado por el incremento en la exactitud (a) y el score F1 (b), junto con una disminución en la pérdida (c). Sin embargo, hay fluctuaciones notables en la pérdida de validación, especialmente en la época 8 y 14, lo que sugiere que el modelo podría estar experimentando alguna variabilidad en su desempeño en el conjunto de validación. Además, la disminución en el score F1 y la exactitud en la época 8 sugiere que el modelo pudo haber enfrentado desafíos con ciertos patrones en los datos en esa etapa particular.

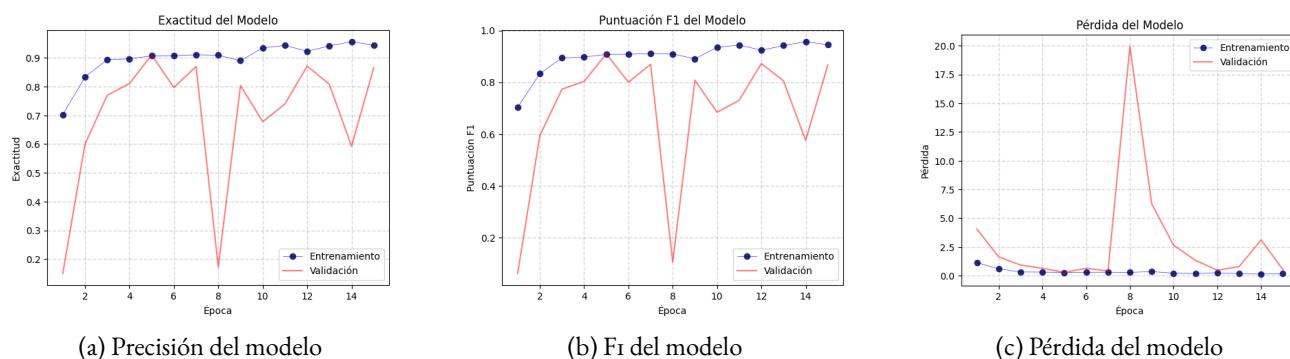


Figura 4: Análisis del rendimiento del modelo a lo largo de las épocas.

La matriz de confusión proporcionada (figura 5) revela información valiosa sobre el desempeño del modelo en la clasificación de enfermedades y condiciones en las hojas de tomate. En primer lugar, es notorio que el modelo tiene una precisión bastante alta en la clasificación de algunas clases específicas, como *Leaf Mold*, *Spider mites Two-spotted spider mite*, y *Tomato Yellow Leaf Curl Virus*. Por ejemplo, en la clase *Leaf Mold*, de los 470 casos, 440 fueron clasificados correctamente, con una desviación mínima hacia otras clases. Sin embargo, la clase *Bacterial Spot* muestra una cierta confusión con la clase *Target Spot*, donde 10 instancias fueron

mal clasificadas.

En segundo lugar, las clases *Early Blight*, *Septoria LeafSpot*, y *Target Spot* parecen presentar desafíos para el modelo. Por ejemplo, en la clase *Early Blight*, de 480 casos, sólo 367 fueron clasificados correctamente, con una desviación notable hacia *Bacterial Spot*, *Septoria LeafSpot* y *Target Spot*. Similarmente, *Septoria LeafSpot* y *Target Spot* tienen confusión con otras clases, lo que indica que el modelo podría estar teniendo dificultades para distinguir entre estas enfermedades específicas. Esto podría ser resultado de características visuales similares entre estas clases,

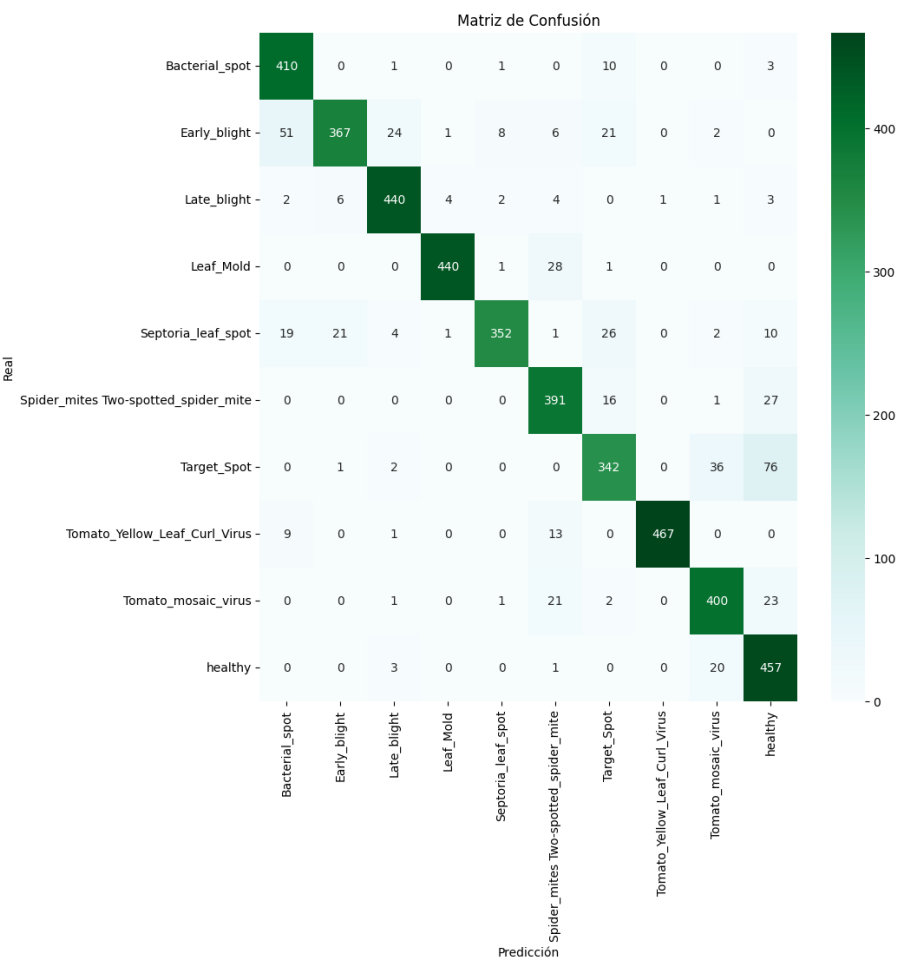


Figura 5: Evaluación del desempeño del modelo en la clasificación de enfermedades y condiciones de hojas de tomate. Las diagonales representan clasificaciones correctas, mientras que las desviaciones indican confusio- nes entre clases.

El objetivo principal de este proyecto fue desarrollar y evaluar modelos de clasificación de imágenes para identificar 10 diferentes estados de salud y enfermedades en hojas de tomate. Se experimentó con dos arquitect-

turas de modelos, un modelo simplificado y ResNet50. Aunque el entrenamiento del modelo simplificado fue mucho más rápido, tomando 215 minutos, comparado con los 595 minutos que tomó entrenar el ResNet50, el rendimiento de ambos modelos mostró diferencias significativas. El modelo simplificado logró una precisión del 95.9 %, un F1-Score del 95.9 % y un Recall del 95.7 %, mientras que ResNet50 alcanzó una precisión del 88.7 %, un F1-Score del 88.7 % y un Recall del 86.8 %.

Al analizar las matrices de confusión, se observa que el modelo simplificado tuvo un desempeño superior en la mayoría de las clases, con una matriz de confusión que muestra una fuerte diagonal principal y pocos errores de clasificación. Por otro lado, el ResNet50, aunque tuvo un desempeño notable en varias clases, mostró un número mayor de errores de clasificación, especialmente en clases como *Early Blight*, *Septoria Leaf Spot*, *Target Spot* y *Tomato Mosaic Virus*. Esto puede indicar que el modelo ResNet50 puede requerir más datos o una mayor fine-tuning para mejorar su rendimiento en estas clases específicas.

Es notable que, a pesar de la arquitectura más compleja y el mayor tiempo de entrenamiento, ResNet50 no logró superar al modelo simplificado en términos de rendimiento. Esto subraya la importancia de entender la naturaleza del conjunto de datos y cómo diferentes arquitecturas de modelos pueden responder a dicho conjunto de datos. También resalta la eficacia de modelos más simples cuando se manejan datos bien estructurados o cuando el problema de clasificación no es extremadamente complejo.

Los resultados obtenidos aquí resaltan la viabilidad de utilizar modelos de clasificación de imágenes en el ámbito agrícola para la detección temprana de enfermedades en plantas. También enfatizan la importancia de elegir y optimizar adecuadamente la arquitectura del modelo en función de las características del conjunto de datos. Aunque el modelo simplificado mostró un desempeño superior en esta instancia, la exploración de arquitecturas de modelos más avanzadas, junto con técnicas de augmentación de datos y optimización, podría llevar a mejoras significativas en futuros trabajos.

REFERENCIAS

- DURMUŞ, H., GÜNEŞ, E. O., & KIRCI, M. (2017). Disease detection on the leaves of the tomato plants by using deep learning. *2017 6th International conference on agro-geoinformatics*, 1-5.
- FUENTES, A., YOON, S., KIM, S. C., & PARK, D. S. (2017). A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17(9), 2022.
- ALHAJ ALI, A., CHRAMCOV, B., JASEK, R., KATTA, R., KRAYEM, S., & AWWAMA, E. (2021). Tomato Leaf Diseases Detection Using Deep Learning. En R. SILHAVY, P. SILHAVY & Z. PROKOPOVA (Eds.), *Data Science and Intelligent Systems* (pp. 199-208). Springer International Publishing.
- VALENZUELA CÁMARA, S. H. (2022). *Detección y clasificación de enfermedades en el tomate mediante Deep Learning y Computer Vision* [Tesis doctoral, Universidad Nacional de La Plata].