



“DIAGNÓSTICO AUTOMATIZADO DE ENFERMEDADES EN TOMATES MEDIANTE DEEP LEARNING”

30 de octubre de 2023

Hector Navarro Barboza





“La **urgencia global de intensificar** la producción alimentaria responde al **incremento poblacional**, que para **2050** exigirá un **aumento del 70 %** para nutrir a más de **9 mil millones de seres humanos.**”(FAO 2009) ¹

¹FAO: Food and Agriculture Organization



1 Introducción

2 objetivos

3 Metodología


- Dataset
- Modelo simplificado
- Modelo ResNet50

4 Resultados

- Modelo simplificado
- Modelo ResNet50

5 Conclusiones

6 Referencias

An aerial photograph of a rural landscape. The image is divided into several rectangular agricultural plots. A central plot is highlighted in white. To the right of this white plot is a large, vibrant green field. The surrounding plots are a light tan or beige color, suggesting bare soil or dry vegetation. A dark, winding path or ditch runs along the left side of the white plot. The word 'Introducción' is centered in black text over the white plot.

Introducción



¿Qué ves en la imagen?



¿Qué ves en la imagen?



- Los humanos pueden identificar imágenes complejas fácilmente.
- ¿Cómo enseñamos a las máquinas a interpretar imágenes?

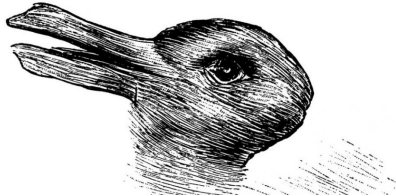


Imagen adaptada de Jastrow 1900. The mind's eye.
Popular Science Monthly, 54, 299-312.

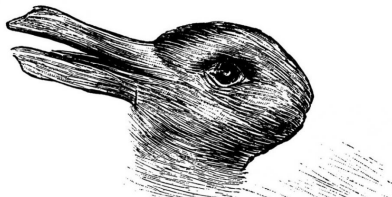
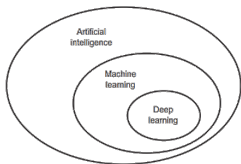


Imagen adaptada de Jastrow 1900. The mind's eye.
Popular Science Monthly, 54, 299-312.

“En esta imagen, un modelo de Deep Learning, al igual que el cerebro, debe identificar el animal presentado basándose en patrones conocidos, destacando el desafío en IA de procesar e interpretar ambigüedades.”

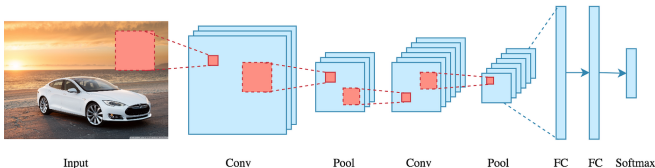
Relación entre IA, Machine Learning y Deep Learning



- **Inteligencia Artificial (IA):** Busca crear o simular inteligencia en máquinas, permitiéndoles realizar tareas que normalmente requieren inteligencia humana.
- **Machine Learning (ML):** Subconjunto de IA que proporciona a las máquinas la habilidad de aprender automáticamente y mejorar con la experiencia sin ser explícitamente programadas.
- **Deep Learning (DL):** Rama de Machine Learning basada en redes neuronales artificiales con múltiples niveles de abstracción.

Estructura General de una CNN

- **Capas Convolucionales:** Aplican una serie de filtros a la imagen para crear mapas de características.
- **Capas de Pooling (Submuestreo):** Reducen la cantidad de parámetros, agrupando información.
- **Capas Fully Connected (Totalmente Conectadas):** Funcionan como una red neuronal clásica, clasificando la información basada en las características aprendidas.




Fuente: Imagen tomada de AI 2023.

Ejemplo de Aplicación de una CNN

Reconocimiento de Imágenes

Las CNN son fundamentales en sistemas modernos de reconocimiento de imágenes, permitiendo tareas como:

- Identificación de objetos en imágenes.
- Reconocimiento facial.
- Diagnósticos médicos a partir de imágenes clínicas.


An aerial photograph of agricultural fields. A central square field is shaded in a light grey color. To its right and bottom-right is a vibrant green field. The surrounding fields are a light tan or beige color, showing signs of being recently plowed or harvested. The word "objetivos" is written in a black, sans-serif font across the center of the grey square.

objetivos



Objetivos

- Comparar estrategias de clasificación mediante un modelo simple y ResNet50 para identificar enfermedades en tomates.
- Evaluar la precisión y eficiencia de ambos modelos, utilizando métricas estandarizadas y tiempos de entrenamiento.

An aerial photograph of agricultural fields. A central square field is shaded in light grey. To its right and bottom-right is a vibrant green field. The surrounding fields are a light tan or beige color, showing signs of being recently plowed or harvested. A dark, straight line, possibly a ditch or road, runs diagonally from the bottom left towards the center. The word 'Metodología' is centered over the grey square.

Metodología



Datos

Se empleó un subset del PlantVillage Dataset de Kaggle, específico para tomates, incluyendo 18,345 imágenes para entrenamiento y 4,585 para validación, abarcando diez categorías de salud y enfermedades.

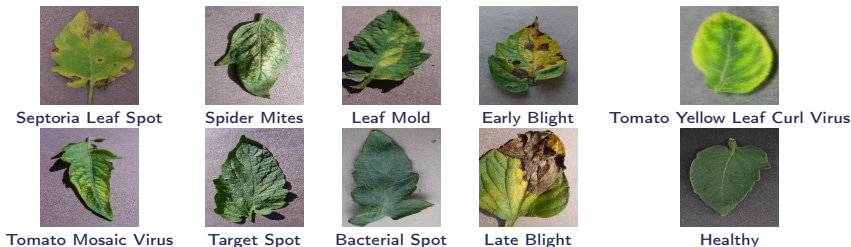


Figura 1: Representación visual de las clases de enfermedades y condiciones de salud en hojas de tomate.

Análisis del Modelo de Red Neuronal

Descripción del Modelo:

- Modelo de tipo secuencial.
- Posee tres capas convolucionales para la extracción de características; cada una seguida por una capa de max-pooling para la reducción de dimensionalidad.
- Una capa Flatten para convertir las matrices 2D en un vector.
- Una capa densa con 128 neuronas para aprendizaje profundo.
- Capa de salida con función de activación 'softmax' para la clasificación entre las diferentes clases.

Código del Modelo:

```
model = Sequential([
    layers.Conv2D(16, 3, padding='same',
                  activation='relu',
                  input_shape=(256, 256, 3)),
    layers.MaxPooling2D(),
    layers.Conv2D(32, 3, padding='same',
                  activation='relu'),
    layers.MaxPooling2D(),
    layers.Conv2D(64, 3, padding='same',
                  activation='relu'),
    layers.MaxPooling2D(),
    layers.Flatten(),
    layers.Dense(128, activation='relu'),
    layers.Dense(num_clases,
                  activation='softmax')
])
```



Uso de Modelo Preentrenado en la Red Neuronal

Descripción del Proceso:

- Carga de ResNet50 preentrenado.
- *Fine-tuning*: se descongelan y entrenan las últimas 20 capas.
- Se añaden capas personalizadas al final para adaptarlo a nuestro conjunto de datos específico.
- Compilación del modelo con métricas específicas.

Código del Modelo:

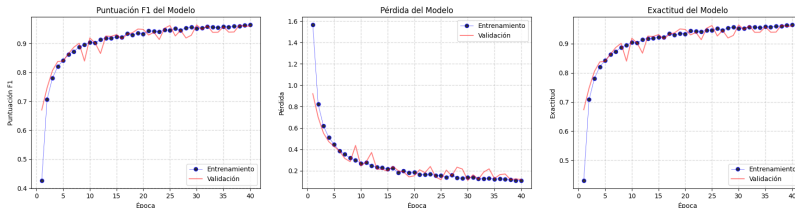
```
base_model = tf.keras.applications.ResNet50(  
    weights='imagenet',  
    input_shape=(256, 256, 3),  
    include_top=False  
)  
  
for layer in base_model.layers[-20:]:  
    layer.trainable = True  
  
model = Sequential([  
    base_model,  
    layers.Flatten(),  
    layers.Dense(128, activation='relu'),  
    layers.Dense(self.num_clases,  
                  activation='softmax')  
)  
  
model.compile(...)
```

An aerial photograph of agricultural fields. A central square field is highlighted in a light grey color. It is surrounded by other fields: a tan-colored field to the top left, a brownish-grey field to the top right, and a vibrant green field to the bottom right. A dark, winding path or ditch runs along the left and bottom edges of the central grey field.

Resultados



Resultados del Entrenamiento/Validación



| loss | accuracy | f1_score | recall | val_loss | val_accuracy | val_f1_score | val_recall |
|-------|----------|----------|--------|----------|--------------|--------------|------------|
| 0.107 | 0.964 | 0.964 | 0.962 | 0.117 | 0.960 | 0.960 | 0.958 |

Los resultados muestran una mejora consistente en todas las métricas clave, indicando un modelo robusto y preciso. El modelo exhibe un alto rendimiento, especialmente en las métricas de validación, lo que sugiere una buena generalización.

Tiempo de entrenamiento/validación: 215minutos

Matriz de Confusión - Modelo Simple

- Alto rendimiento en la clasificación, como indican los valores elevados en la diagonal principal.
- Confusiones notables entre *Target_Spot* y *Spider_mites* y *Two-spotted_spider_mite*.
- Leve confusión entre *Early_blight* y *Septoria_leaf_spot*.
- Excelente clasificación de *Tomato_Yellow_Leaf_Curl_Virus* y *healthy* con mínima confusión.

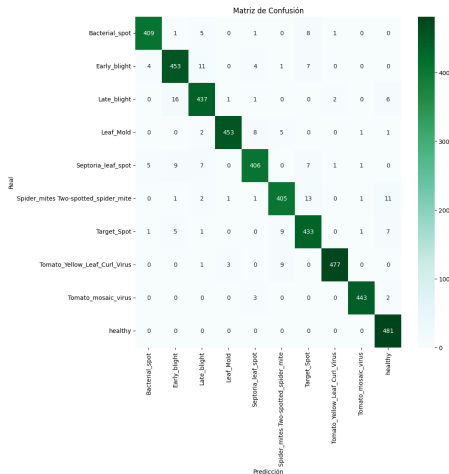
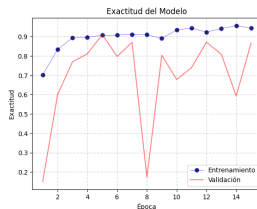
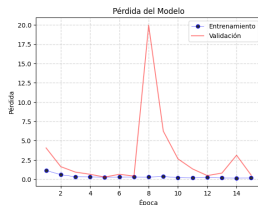
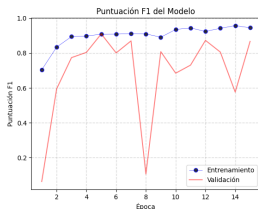


Figura 2: Precisión y los errores de clasificación del modelo.

Resultados del Entrenamiento/Validación - Resnet50



| loss | accuracy | f1_score | recall | val_loss | val_accuracy | val_f1_score | val_recall |
|-------|----------|----------|--------|----------|--------------|--------------|------------|
| 1.139 | 0.703 | 0.703 | 0.621 | 4.056 | 0.150 | 0.063 | 0.082 |
| 0.591 | 0.834 | 0.833 | 0.799 | 1.631 | 0.600 | 0.594 | 0.564 |
| 0.178 | 0.945 | 0.945 | 0.941 | 0.523 | 0.866 | 0.866 | 0.860 |

Se observa una tendencia de mejora en todas las métricas clave. Sin embargo, los valores de pérdida de validación (**val_loss**) sugieren posibles episodios de sobreajuste, dada la inconsistencia entre entrenamiento y validación. A pesar de esto, el modelo muestra un rendimiento prometedor, evidenciado en las métricas de exactitud (accuracy) y precisión (recall).

Tiempo de entrenamiento/validación: 595minutos

Matriz de Confusión - Resnet50

- Precisión destacada en clases como *Leaf Mold*, *Spider mites* *Two-spotted spider mite*, y *Tomato Yellow Leaf Curl Virus*, con clasificaciones correctas en la mayoría de los casos.
- Confusión observada en la clase *Bacterial Spot* con *Target Spot*, indicando dificultades en la diferenciación entre estas condiciones.
- Desafíos en la clasificación precisa de *Early Blight*, *Septoria Leaf Spot*, y *Target Spot*, posiblemente debido a similitudes en características visuales entre estas enfermedades.

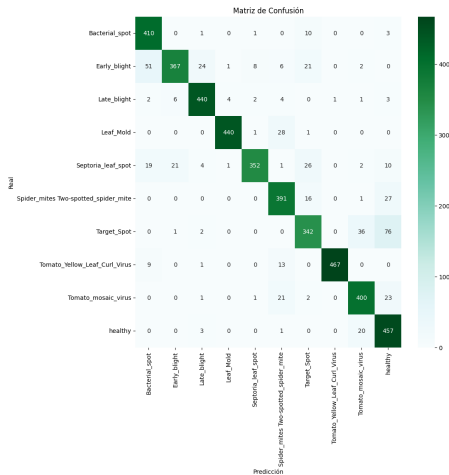


Figura 3: Análisis detallado de los errores de clasificación del modelo.

An aerial photograph of a rural landscape. A large, bright white square is positioned in the center-left of the frame. To its right is a vibrant green field, and above it are several brown, tilled agricultural plots. A dark, winding path or ditch separates the white square from the green field. The word "Conclusiones" is written in a black, sans-serif font across the white square.

Conclusiones



Conclusiones

- Se desarrollaron y evaluaron dos modelos para identificar 10 condiciones en hojas de tomate: uno simplificado y ResNet50.
- A pesar de su rapidez, el modelo simplificado superó a ResNet50 con una precisión del 95.9 % frente al 88.7 %.
- El modelo ResNet50 mostró dificultades en clases específicas, sugiriendo la necesidad de más datos o ajustes finos.
- Importancia de seleccionar y optimizar la arquitectura del modelo según el conjunto de datos, evidenciada por el alto rendimiento del modelo más simple.
- Potencial de los modelos de clasificación de imágenes en agricultura, con posibilidades de mejora mediante técnicas avanzadas y optimización.

Agradecimientos

¡Gracias por su atención!

Para más información o preguntas:
h.navarrobarboza@gmail.com



 github.com/hectornav

 linkedin.com/in/hectornavarrobarboza

An aerial photograph of agricultural land. A central square area is covered in grey material, possibly sand or a specific type of soil, contrasting with the surrounding brownish-tan tilled fields. A vibrant green field is located in the bottom right corner. A dark, straight line, likely a ditch or road, runs diagonally from the left towards the center. The word "Referencias" is centered over the grey square.

Referencias



Referencias I



AI, Bootcamp (2023). *Redes Neuronales Convolucionales*. Bootcamp AI Medium Blog. url: <https://bootcampai.medium.com/redes-neuronales-convolucionales-5e0ce960caf8>.



FAO (2009). *2050: un tercio más de bocas que alimentar*. Organización de las Naciones Unidas para la Alimentación y la Agricultura. url: <https://www.fao.org/news/story/es/item/35675/icode/> (visitado 25-10-2023).



Jastrow, Joseph (1900). *Fact and fable in psychology*. Houghton, Mifflin y Company.