

DETECCIÓN DE MALARIA

Marian Vela, Aarón Menchú, Otto Chamo

Problema

El diagnóstico de malaria presenta un desafío, en especial en regiones de escasos recursos, en donde está limitada la disponibilidad de equipo especializado y profesionales con conocimientos y experiencia suficiente para diagnosticar adecuadamente la enfermedad. Por esto, se busca emplear técnicas de Machine Learning para facilitar el diagnóstico, identificando a partir de las imágenes de las muestras cuáles están infectadas con malaria de manera que se pueda optimizar el tiempo de respuesta oportuna para iniciar el tratamiento.

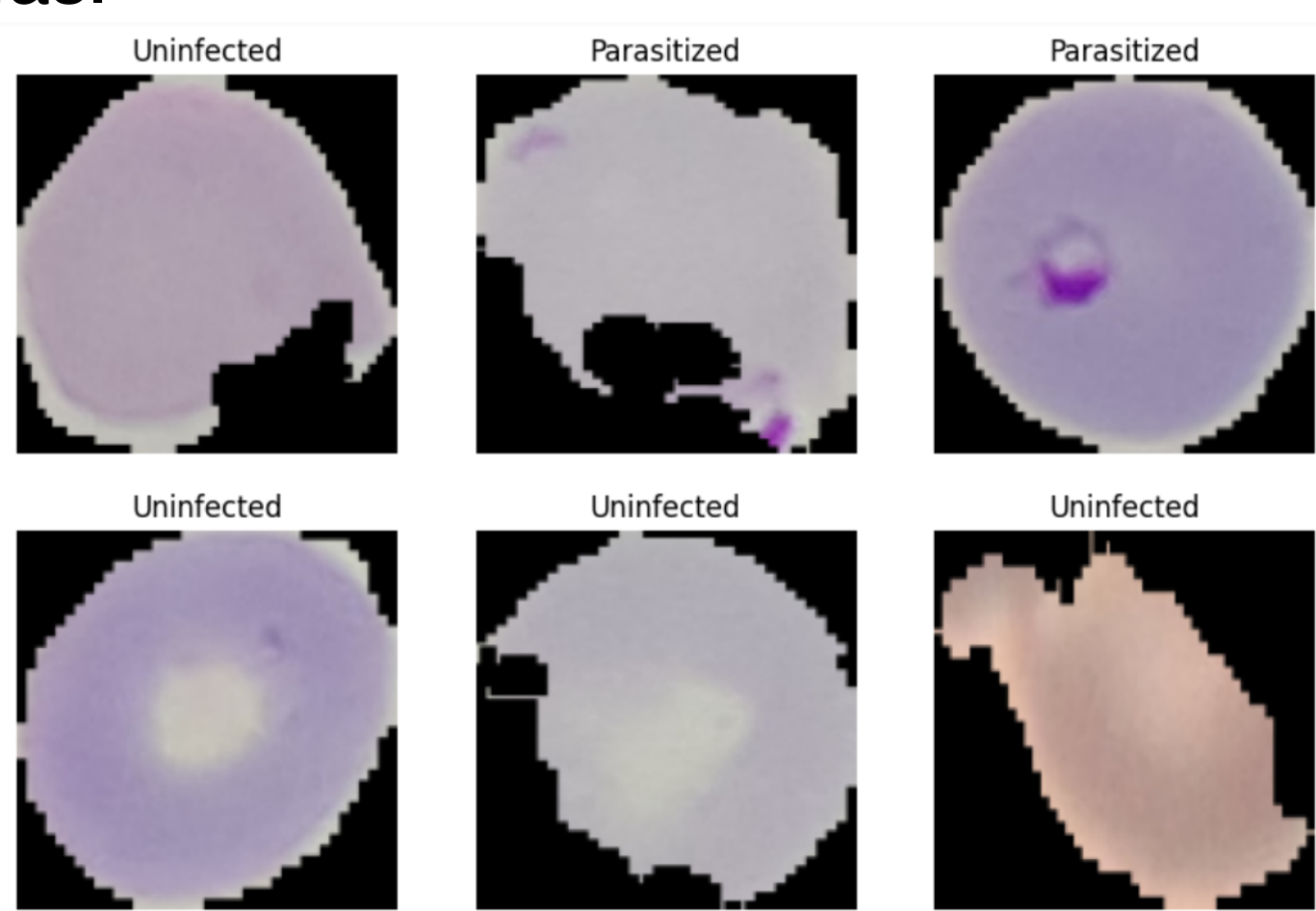
Malaria

La Malaria es una enfermedad infecciosa causada por el parásito *Plasmodium*, cuyo vector de transmisión son los mosquitos hembras, los cuales depositan los parásitos en la sangre de las personas infectadas al momento de la picadura. Los típicos síntomas son fiebres altas, escalofríos, dolor de cuerpo y fatiga.

La enfermedad principalmente afecta en regiones tropicales, cálidas y en climas con depósitos de agua estancada. El diagnóstico y tratamiento oportuno son fundamentales para eliminar los parásitos y evitar la manifestación de síntomas severos, ya que anualmente, alrededor de 380 mil personas mueren debido a la enfermedad a pesar de que existen tratamientos efectivos anti malaria.

Dataset

Para este problema se trabajó con un total de 27,560 imágenes provistas por Kaggle, las cuales muestran casos en que el paciente cuenta con la enfermedad de Malaria o no. Para poder trabajar con esta información, se clasificaron las imágenes en 2 categorías: Uninfected y Parasitized. Se utilizó una partición de 80% a 20%: se utilizó 11,024 imágenes para cada una de las clases de Uninfected y Parasitized para el entrenamiento, y para test se utilizó 2,756 imágenes en ambas categorías.



Visualización de la Data

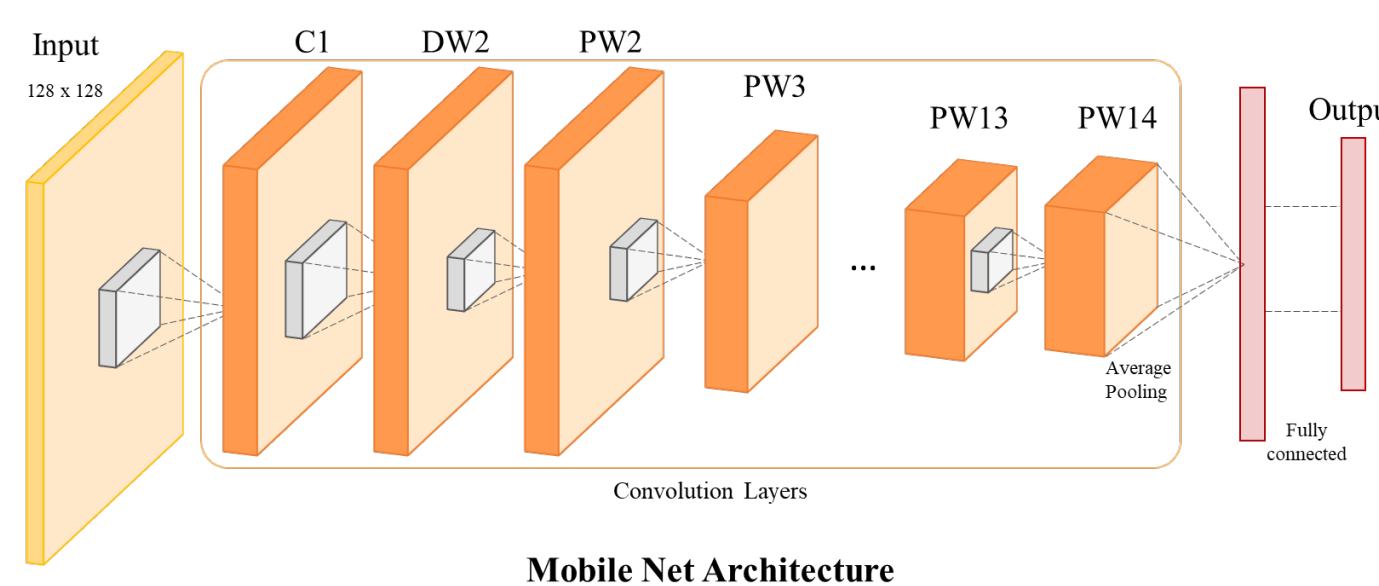
Metodología

El objetivo que se quiere lograr es identificar rápidamente y con la mayor precisión posible cuando la imagen muestra una célula infectada con malaria, por lo tanto se busca un modelo con el mejor rendimiento posible en cuanto a precisión. Se realizaron en total 3 modelos en el lenguaje de Python, utilizando librerías necesarias para trabajar con modelos de Machine Learning. Dos modelos fueron hechos desde cero utilizando redes convolucionales, y en el tercer modelo se utilizó Transfer Learning.

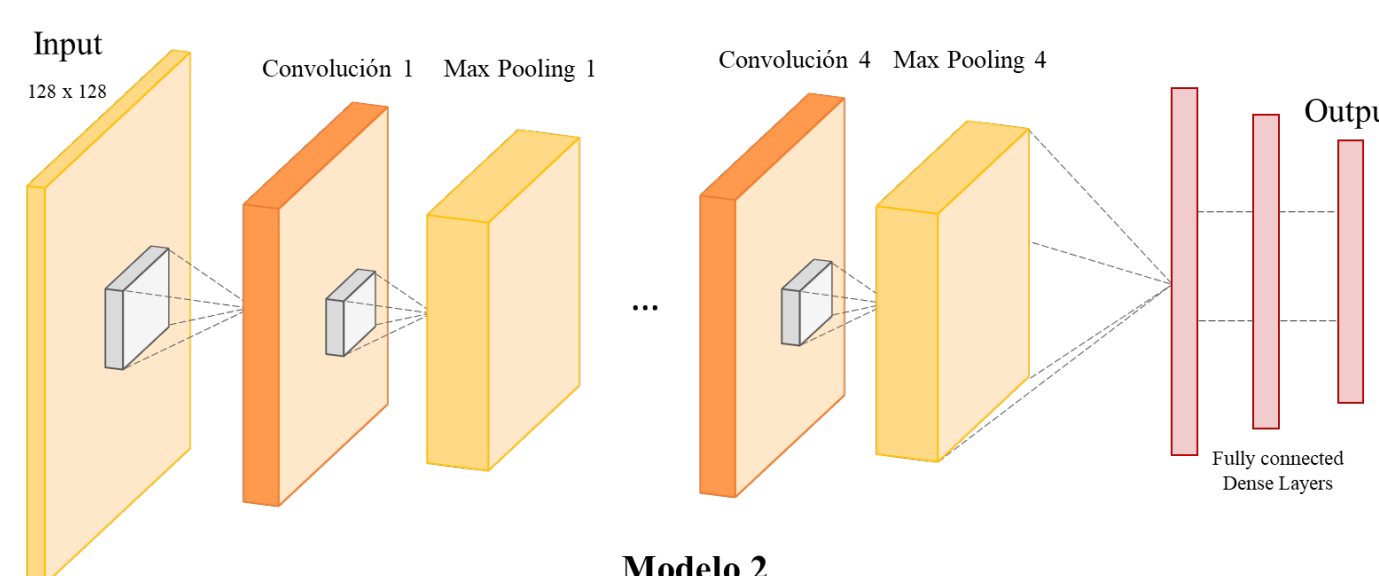
Arquitecturas

En el modelo que se utilizó Transfer Learning, se optó por utilizar el modelo de **MobileNet V2**, el cual fue entrenado con las imágenes de ImageNet, ya que este brinda varias ventajas:

- Gran conjunto de datos que se encuentran pre-entrenados
- Bajo consumo de memoria
- Rendimiento y precisión
- Eficiencia en recursos



Para el Modelo 2 se diseñó una arquitectura con 4 capas de convolución, 4 de Max Pooling para feature extraction y 3 capas densas en la red neuronal.



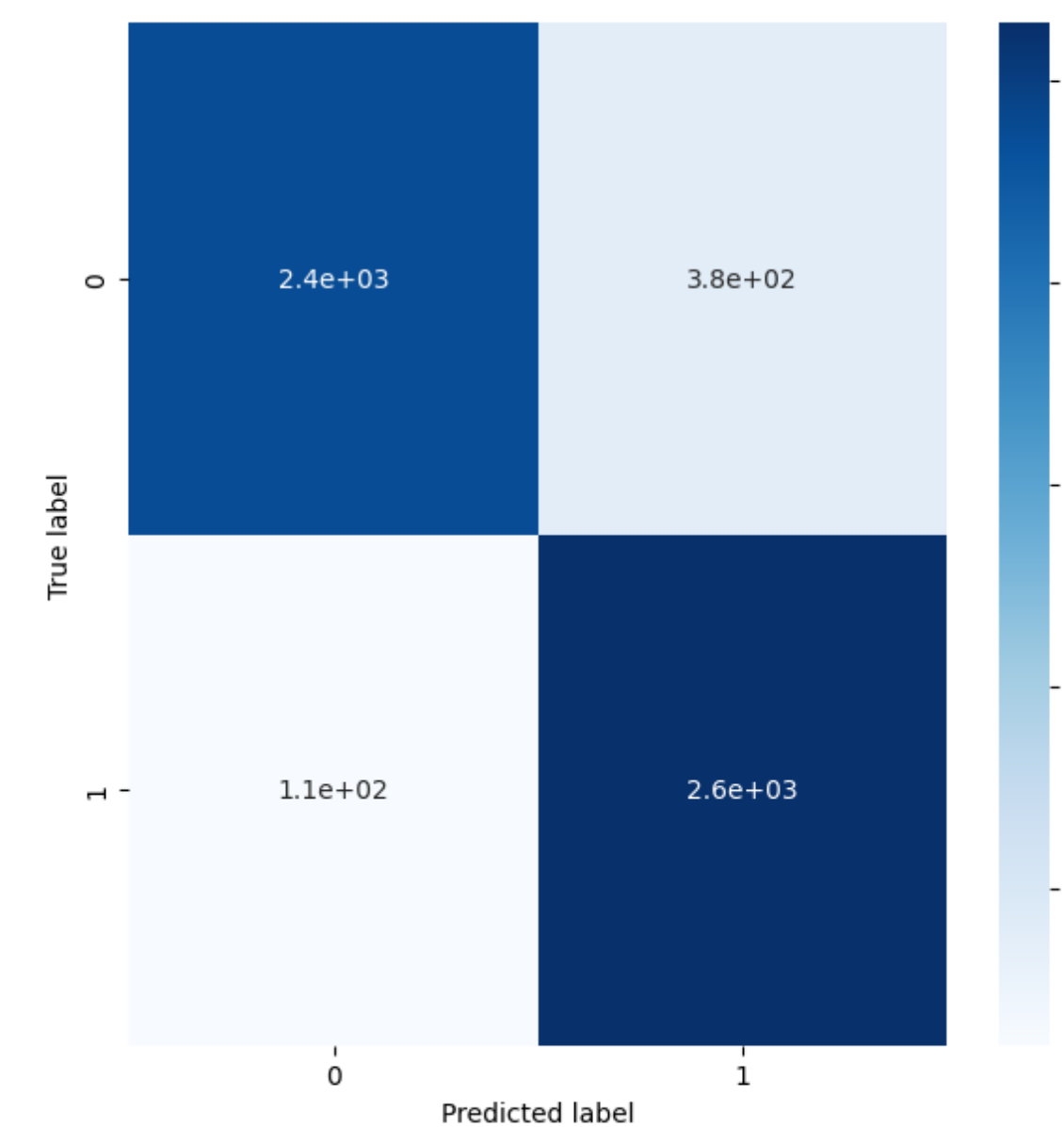
Resultados

Se observa que el mejor modelo es el Modelo 2, debido a que devolvió las mejores métricas. No obstante el modelo de Transfer Learning también tiene un buen desempeño.

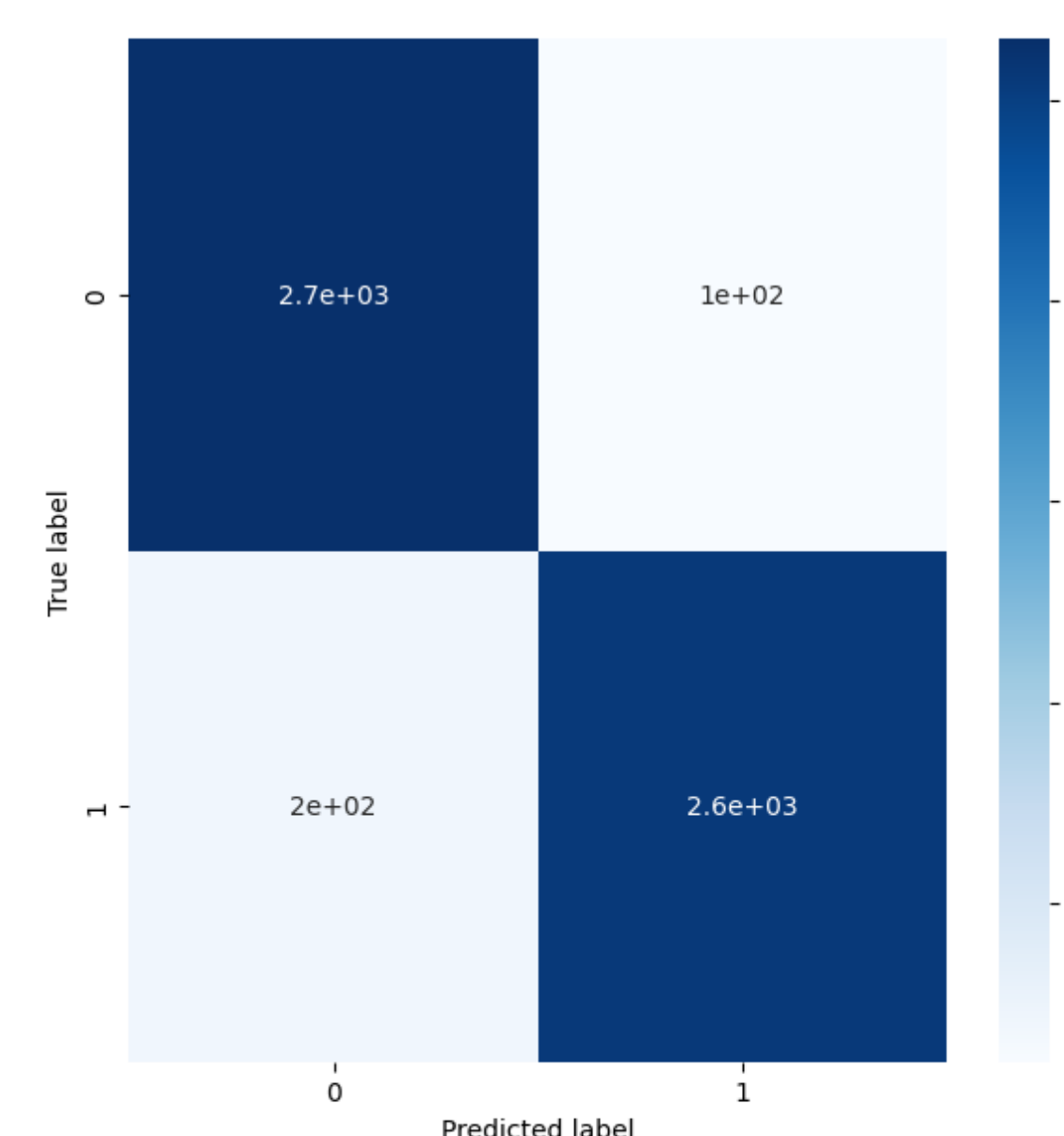
Comparación de métricas				
Modelo	Loss	Accuracy	Specificity	Sensibility
MobNet	0.2278	0.9102	0.86	0.96
Modelo 1	0.6416	0.6204	0.54	0.71
Modelo 2	0.1689	0.9452	0.96	0.93

Resultados

Se obtuvo las siguientes matrices de correlación y se observa que la matriz de correlación para el Modelo 2 es mejor para predicciones.



Matriz correlación TL



Matriz correlación modelo 2

Mejoras a Futuro

Considerando que los modelos obtuvieron un resultado satisfactorio, pueden llegar a ser de gran utilidad para el área médica, algunas mejoras a futuro serían:

- Entrenar el modelo con un dataset de gran volumen, que contenga solo imágenes de malaria.
- Agregar más capas para extracción de features para aumentar la sensibilidad del modelo a un 99%.
- Realizar un apartado visual en donde la persona pueda observar las imágenes donde se indique el área afectada y si esta posee la enfermedad o no.

Conclusiones

Machine Learning puede ser útil para diagnóstico médico de manera ágil y con gran precisión. En el caso del experimento se observa que, a pesar que existan modelos entrenados con una gran cantidad de imágenes y con una gran cantidad de capas de convolución, no necesariamente resultan siendo los mejores modelos ya que al utilizar un modelo de AI diseñado específicamente para el problema se obtuvo un mejor resultado que empleando Transfer Learning.

References

- [1] WHO. Malaria, 2003. URL <https://www.who.int/news-room/fact-sheets/detail/malaria>.
- [2] NIH Library. Detecting Malaria. URL <https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria>.
- [3] Google. MobileNet V2. URL https://tfhub.dev/google/imagenet/mobilenet_v2_00128/classification/5.
- [4] Jijon Sofia. A very simple poster. URL <http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm>.

