

Clasificación de patologías en la retina a partir de imágenes OCT con Deep Learning

Isabella Bermón Rojas
Departamento de Ingeniería Electrónica
Universidad de Antioquia
Medellín, Colombia
isabella.bermonr@udea.edu.co

Abstract—The application of machine learning techniques to OCT aims to improve the accuracy and speed of image classification, providing valuable assistance to ophthalmologists in the diagnosis and monitoring of these diseases. This project aims to train different types of CNNs and vision transformer architectures to find the best possible solution to classify diseases into OCT scans, as well as to study the effects of preprocessing on this type of medical images.

Keywords: CNNs, vision transformers, machine learning, image processing, medical imaging.

I. CONTEXTO

La tomografía de coherencia óptica o tomografía óptica coherente (OCT, por sus siglas en inglés) es una técnica de imagen tomográfica óptica, no invasiva [1]. Estas imágenes permiten detectar y diagnosticar diversas enfermedades oculares, como la degeneración macular, el edema macular, y la retinopatía diabética. Así como en otras áreas de la ciencia, la capacidad de automatizar el procesamiento de estas imágenes mediante técnicas de *deep learning* puede mejorar la precisión diagnóstica, reducir el tiempo de análisis y facilitar el acceso a servicios de salud ocular.

II. OBJETIVO DE MACHINE LEARNING

El objetivo de este proyecto de *machine learning* es identificar el desempeño de diferentes arquitecturas de *deep learning*, entre ellas redes neuronales convolucionales y transformadores de visión, para clasificar imágenes de OCT en cuatro categorías: NORMAL, CNV (neovascularización coroidea), DME (edema macular diabético) y DRUSEN. La finalidad del modelo es predecir la condición médica (una de las cuatro clases mencionadas) de la retina basándose en imágenes de escaneo OCT. Dada una imagen de retina, el modelo aprenderá a identificar características relevantes que puedan indicar la presencia de estas condiciones patológicas para asistir en el diagnóstico médico. Adicionalmente, se aplicarán diferentes filtros a la imagen de entrada con el objetivo de encontrar la mejor representación de las 4 clases.

III. DATASET

El dataset utilizado es el **Labeled Optical Coherence Tomography (OCT)** disponible en Kaggle. Este conjunto de datos contiene imágenes de OCT que han sido etiquetadas según las 4 categorías de enfermedades o estados de salud ocular. El conjunto de datos está compuesto por imágenes de OCT en formato PNG, cada una con etiquetas correspondientes a las cuatro categorías mencionadas.

A. Tamaño del dataset

El dataset consta de 109,309 imágenes de OCT, ocupando alrededor de 6.70 GB en disco. El dataset se encuentra desbalanceado entre las cuatro clases, como se muestra en la siguiente figura.

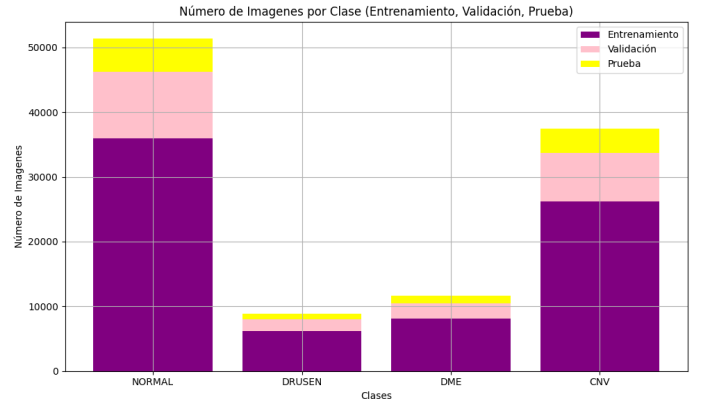


Fig. 1. Representación de base de datos dividida por patologías y su respectiva división en subconjuntos.

B. Distribución del dataset

- Entrenamiento: 76,515 imágenes (70% del total)
- Validación: 21,861 imágenes (20% del total)
- Prueba: 10,933 imágenes (10% del total)

A continuación se muestra un ejemplo de imágenes aleatorias de cada clases y cada subconjunto. Se puede evidenciar que hay en algunos casos una rotación en el ángulo de las imágenes y el tamaño varía en muchos casos.

Adicionalmente, aparte del ángulo que presenta la forma de la retina hay imágenes que tienen partes sin información de la imagen original (se ven como franjas amarillas, que en la visualización de escala de grises se verían negras por completo) y pueden generar ruido en el proceso de entrenamiento e inferencia.

Muestra de Imágenes de los conjuntos de Entrenamiento, Validación y Prueba

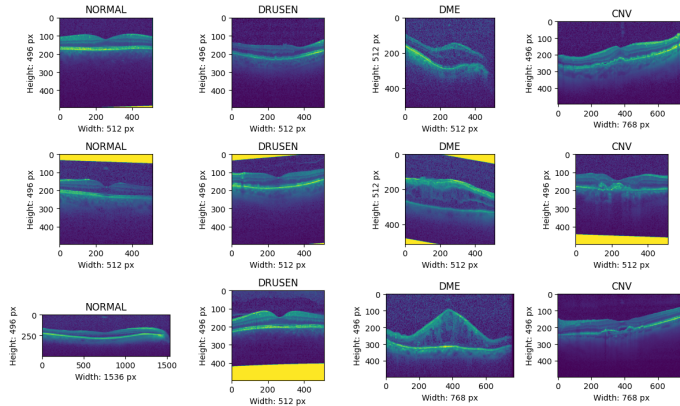


Fig. 2. Diagrama de flujo de la solución implementada.

IV. MÉTRICAS DE DESEMPEÑO

A. Machine Learning

Para evaluar el desempeño del modelo de clasificación se utilizarán las siguientes métricas en base a las métricas comúnmente reportadas en la literatura:

- **Exactitud (Accuracy):** Proporción de predicciones correctas sobre el total de predicciones.
- **Precisión (Precision):** Medida de la exactitud de las predicciones positivas.
- **Recall (Sensibilidad):** Capacidad del modelo para identificar correctamente las instancias positivas.
- **F1-Score:** Media armónica entre precisión y recall, útil para balancear ambas métricas.
- **Matriz de Confusión:** Para visualizar el desempeño del modelo en cada clase específica.

B. Negocio

Para evaluar el impacto en el negocio y en la práctica clínica se consideran las siguiente métricas:

- **Tiempo de Diagnóstico:** Reducción en el tiempo requerido para obtener un diagnóstico preciso en comparación con la evaluación de un experto.

- **Costo de Implementación:** Inversión necesaria para desplegar el modelo en un entorno clínico.
- **Tasa de Adopción:** Nivel de aceptación y uso del modelo por parte de los profesionales de la salud.
- **Mejora en la Calidad de Atención:** Incremento en la precisión diagnóstica y en la satisfacción del paciente.
- **Retorno de Inversión:** Beneficios económicos derivados de la implementación del modelo versus los costos asociados.

V. RESULTADOS PREVIOS

Estudios previos han demostrado que los modelos de *deep learning*, especialmente las redes neuronales convolucionales, superan a los métodos tradicionales en la clasificación de imágenes médicas OCT. Por ejemplo, Malliga et al. (2022) desarrollaron sistema de aprendizaje profundo que utiliza redes neuronales convolucionales para clasificar las imágenes de OCT de retina en categorías como DMAE, NVC y RD con una precisión de aproximadamente el 97% [2]. De Fauw et al. (2018) implementaron un modelo capaz de diagnosticar múltiples enfermedades oculares con una precisión que comparan significativamente con la de los profesionales humanos [3]. Estos resultados respaldan la viabilidad de utilizar técnicas de *deep learning* para mejorar el diagnóstico y el tratamiento de enfermedades oculares mediante el análisis de imágenes OCT.

REFERENCES

- [1] Elsharif, A. A. E. F., Abu-Naser, S. S. (2022). Retina diseases diagnosis using deep learning.
- [2] (2022). Classification of Retinal OCT Images Using Deep Learning. doi: 10.1109/icccci54379.2022.9740985
- [3] De Fauw, J., Ledsam, J. R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., ... Ronneberger, O. (2018). Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine*, 24(9), 1342-1350.