

Optimización de imágenes

Image optimization

A. Benali, R. Martin-Pinardel, E. Romero, A. Vellido

Resumen

La fase de modificación de los datos previa a su uso en modelos de inteligencia artificial (IA) se conoce también como preprocesamiento. En esta área, se han ideado y explorado empíricamente procedimientos nuevos, usándose en retinografía, tomografía de coherencia óptica (OCT) y angiografía de tomografía de coherencia óptica (OCTA). Estos procedimientos consisten en transformaciones de datos donde los objetivos principales son: eliminar artefactos, detectar valores atípicos y mejorar la relación señal/ruido (SNR).

El objetivo es aprovechar al máximo la limitada cantidad de imágenes disponibles, lo que se puede conseguir aplicando un preprocesamiento *ad hoc*. Esto permite allanar el camino a los modelos de IA para que puedan obtener mejores resultados.

Palabras clave: Preprocesamiento de imagen. Imagen médica de retina. Aprendizaje profundo. Aumento de datos. Superresolución.

Resum

La fase de modificació de les dades prèvia al seu ús en models d'intel·ligència artificial (IA) es coneix també com a preprocessament. En aquesta àrea, s'han ideat i explorat empíricament procediments nous, usant-se en retinografia, tomografia de coherència òptica (OCT) i angiografia de tomografia de coherència òptica (OCTA). Aquests procediments consisteixen en transformacions de dades on els objectius principals són eliminar artefactes, detectar valors atípics i millorar la relació senyal/soroll (SNR).

L'objectiu és aprofitar al màxim la limitada quantitat d'imatges disponibles, cosa que es pot aconseguir aplicant un preprocessament *ad-hoc*. Això permet obrir camí als models IA per tal que aquests puguin obtenir millors resultats.

Paraules clau: Preprocessament d'imatge. Imatge mèdica de retina. Aprenentatge profund. Augmentació de dades. Superresolució.

Abstract

The phase of data modification prior to its use in artificial intelligence (AI) models is also known as preprocessing. In this area, new procedures have been devised and empirically explored and used in retinography, optical coherence tomography (OCT) and optical coherence tomography angiography (OCTA). These procedures consist of data transformations where the main objectives are to eliminate artefacts, detect outliers and improve the signal-to-noise ratio (SNR). The goal is to make the best use of the limited amount of available images, which can be achieved by applying *ad-hoc* preprocessing. This paves the way for AI models to be able to achieve better results.

Key words: Image preprocessing. Retinal imaging. Deep learning. Data augmentation. Super-resolution.

3.1. Optimización de imágenes

Image optimization

A. Benali¹, R. Martin-Pinardel², E. Romero¹, A. Vellido¹

¹*Intelligent Data Science and Artificial Intelligence Research Center-Universitat Politècnica de Catalunya (IDEAI-UPC).*

²*Fundació de Recerca Biomèdica Clínic-Institut d'Investigacions Biomèdiques August Pi i Sunyer (FCRB-IDIBAPS).*

Correspondencia:

Anass Benali

E-mail: anass.benali@estudiantat.upc.edu

Introducción

Al aplicar modelos de IA, y más específicamente, de redes neuronales profundas (DNN, *deep neural networks*) a imágenes médicas, surgen varios retos. Uno de los más destacados es que el tamaño de los conjuntos de datos disponibles suele ser bastante pequeño en comparación con lo que se requiere para entrenar de manera óptima una DNN. Además, en la mayoría de los casos, existe un desequilibrio de distribución entre las clases a predecir. Estas preocupaciones se agravan con la presencia de artefactos, patrones superfluos y valores atípicos. Tales fenómenos son problemáticos al entrenar las DNN, ya que pueden llevar fácilmente a la obtención de resultados engañosos.

Las DNN son muy flexibles en su capacidad de aprendizaje, por lo que pueden encontrar y aprender patrones complejos. Sin embargo, dicha flexibilidad es una arma de doble filo, pues las hace propensas al sobreajuste (*overfitting*). El sobreajuste es el comportamiento indeseado de un modelo de aprendizaje automático o *machine learning* (ML), que obtiene buenos resultados para los datos observados, pero no logra resultados comparables en nuevos datos (aquellos con los que la IA no ha sido entrenada). En tal caso, decimos que el modelo ha obtenido resultados que no generalizan. Para mitigar este problema y abrir paso a resultados generalizables, es importante contar con datos bien preprocesados.

Métodos propuestos, áreas de aplicación y ejemplos

Transformaciones básicas

Existen algunos pasos básicos que se aplican en la mayoría de imágenes:

- Escalar los valores de la imagen al rango de -1 a +1. Este ajuste puede ser útil para el entrenamiento de una DNN y para adecuarse al formato de coma flotante.
- Estandarización de los valores de la imagen. Esta transformación está diseñada para ayudar a que los modelos converjan más rápido y con mayor estabilidad.
- Reducción del tamaño de la imagen. Esto ayuda a limitar el consumo de memoria y la complejidad del modelo.

Estas transformaciones no cambian la esencia de las imágenes, ya que simplemente son un mapeo a un formato más adecuado para las DNN.

Aumento de datos

Es bien conocido que la adquisición de grandes conjuntos de datos en el ámbito de las imágenes médicas no es viable debido a que requiere mucho tiempo y es costoso. Una de las formas de mitigar el hecho de tener un conjunto de datos relativamente

pequeño es por vía del aumento de datos (*data augmentation*)¹. Este procedimiento consiste en tomar las imágenes y aplicarles aleatoriamente algunas transformaciones. Las transformaciones que se suelen aplicar son: rotaciones, giros, inclinaciones, recortes, cambio de tamaño, ajuste de brillo, ajuste de color, etc. Suelen aplicarse en cantidades muy pequeñas para simular, lo más fielmente posible, distorsiones reales¹.

En las imágenes médicas de retina, las transformaciones más básicas para el aumento de datos son las rotaciones de múltiplos de 90°, reflejos sobre el eje horizontal o vertical, pequeñas rotaciones y ligeros incrementos del *zoom*, entre otros (Figura 1).

Los datos generados mediante esta técnica, no serán tan efectivos como obtener una nueva muestra de datos, pero debido a la baja dificultad de su generación, resulta una forma sencilla de obtener una mejora en los resultados, como por ejemplo ayudando a mitigar el sobreajuste.

Optimización de retinografías

El método de preprocesamiento ideado por Ben Graham, ganador del concurso Kaggle de detección de retinopatía diabética de

2015², logró buenos resultados para modelos DNN en imágenes de retinografía (Figura 2). Este método tiene como objetivo normalizar la iluminación, la exposición y otros factores de las imágenes. En trabajos recientes, se han explorado variaciones del método para otros tipos de modelos, como por ejemplo, para modelos de factorización matricial (MF)³.

En la adaptación para modelos MF, en lugar de mapear el fondo de la imagen a gris, se mapea a blanco y luego se invierte, lo cual proporciona mejores resultados. Además, analizando los canales de color RGB de las imágenes, se ha observado que los canales verde y rojo son los más informativos, mientras que el azul es menos relevante, debido a su susceptibilidad a ser afectado negativamente en imágenes tomadas con mala iluminación, por lo que este es descartable (Figura 3).

Estos cambios ayudaron a que el modelo de factorización se centre en encontrar características de los vasos sanguíneos y los daños en la retina de forma más efectiva, al disminuir la prevalencia de artefactos y la variabilidad en el fondo de las imágenes³.

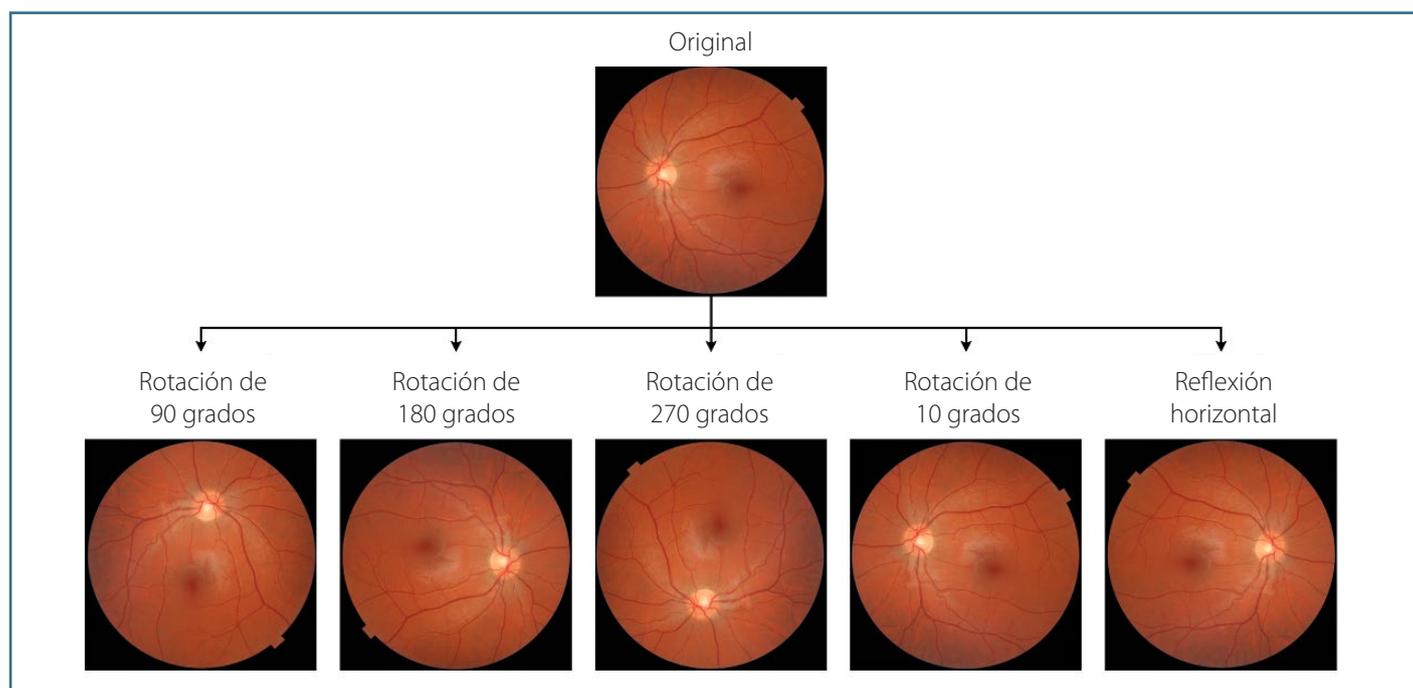


Figura 1. Ejemplo de transformaciones básicas dentro del proceso de aumento de datos en retinografías.

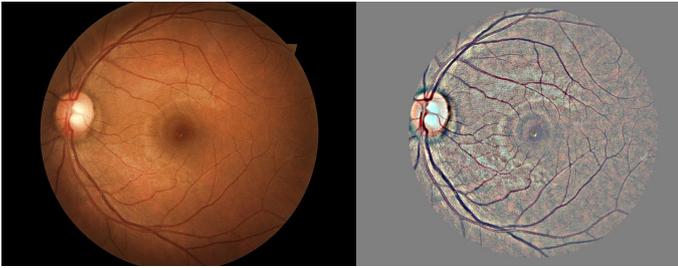


Figura 2. Ejemplo de aplicación del método de Ben Graham en una retinografía².

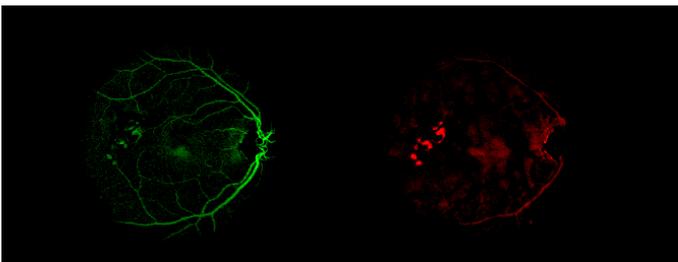


Figura 3. Resultado de aplicar el método adaptado en una retinografía, separado por canales RGB, descartando el azul⁴.

Optimización de imágenes OCT

Existen varias características en la optimización de imágenes OCT que pueden dificultar el aprendizaje de los modelos, tales como la presencia de artefactos o la curvatura (longitud axial), entre otros. Peor aún, estos artefactos pueden confundir o engañar al modelo. Por ello y para que los modelos aprendan a no centrarse en esos patrones, se han utilizado técnicas de visión artificial (*computer vision*) para normalizar las características que podrían dificultar el aprendizaje de otras más útiles.

Para imágenes OCT, se ha trabajado en un proceso personalizado para eliminar artefactos, extraer la región de interés (ROI), horizontalizar y enderezar dichas imágenes⁴.

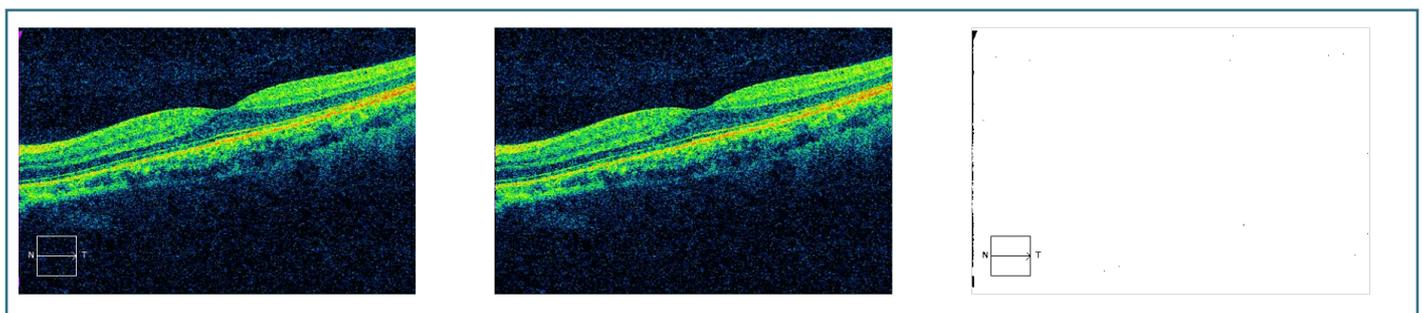


Figura 4. Imagen OCT original, imagen filtrada y la diferencia entre ambas (máscara), respectivamente⁴.

Para la eliminación de artefactos, un proceso basado en la separación de colores ha sido explorado y ha obtenido resultados satisfactorios. Las imágenes OCT exportadas pueden contener una barra magenta y un cuadro de orientación blanco. Al filtrar por color en el espacio de color HSV (en lugar de RGB), estos indicadores pueden eliminarse sin que ello afecte negativamente a la calidad de la imagen (Figura 4).

Para la extracción de la ROI, se ha ideado un proceso más elaborado. En primer lugar, para seleccionarla, se utilizan filtros de desenfoque, métodos de valor umbral y operadores morfológicos para obtener la silueta de la imagen OCT. En segundo lugar, a partir de esa silueta, se estima el rectángulo (o elipse) que mejor se ajuste a ella con tal de rotar la imagen a orientación horizontal (Figura 5). En tercer lugar, se aborda la curvatura (longitud axial). Dado que la curvatura es o convexa o cóncava, se puede ajustar una parábola para estimar y deshacer dicha curvatura (Figura 6). Finalmente, dado que la rotación de la imagen añade cierta oblicuidad en los lados, ésta se deshace con una transformación de perspectiva, dando por finalizado el proceso (Figura 7).

El resultado de aplicar estas transformaciones son imágenes mejoradas, listas para ser usadas en cualquier modelo de ML. Experimentalmente, éstas han demostrado ser más adecuadas en métodos de MF³, permitiendo a estos modelos aprender características más útiles.

Optimización de imágenes OCTA

Varios filtros y transformaciones tradicionales han sido aplicados a las imágenes OCTA con el objetivo de reducir el ruido, mejorar la SNR, extraer o acentuar algún tipo específico de vasos o resaltar la zona avascular foveal (FAZ). Los filtros utilizados para el proce-

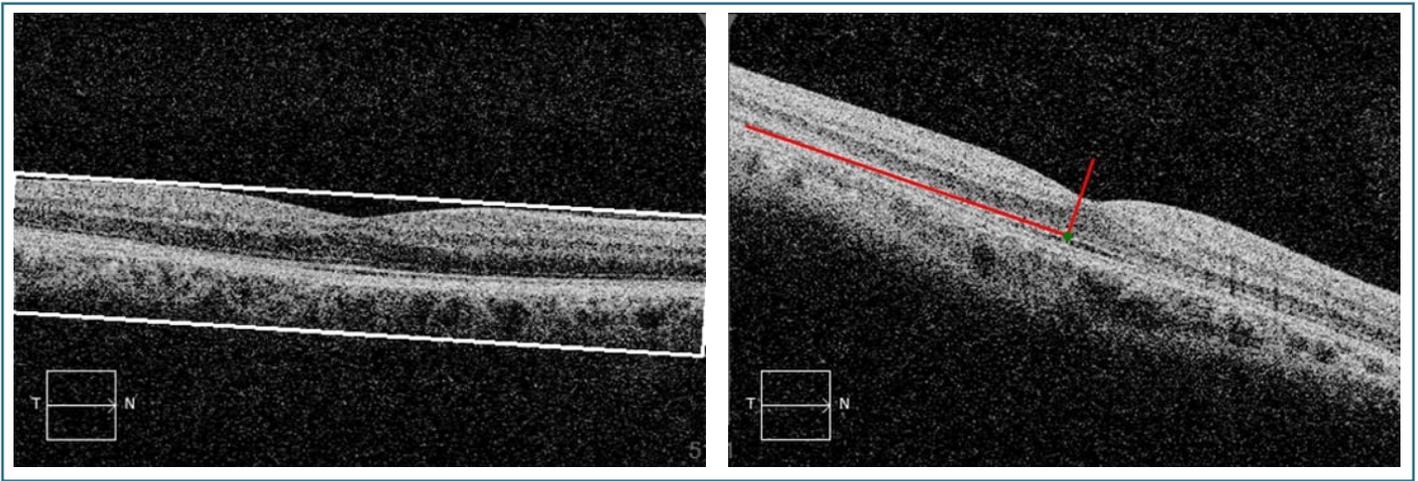


Figura 5. Resultado de la estimación del rectángulo (izquierda) y de la elipse (derecha) en imágenes OCT⁴.

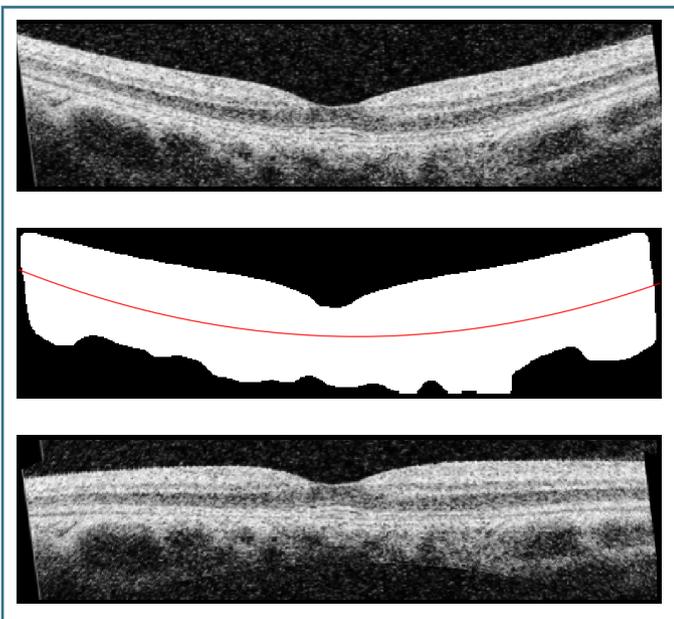


Figura 6. Proceso de enderezado de la imagen OCT⁴.

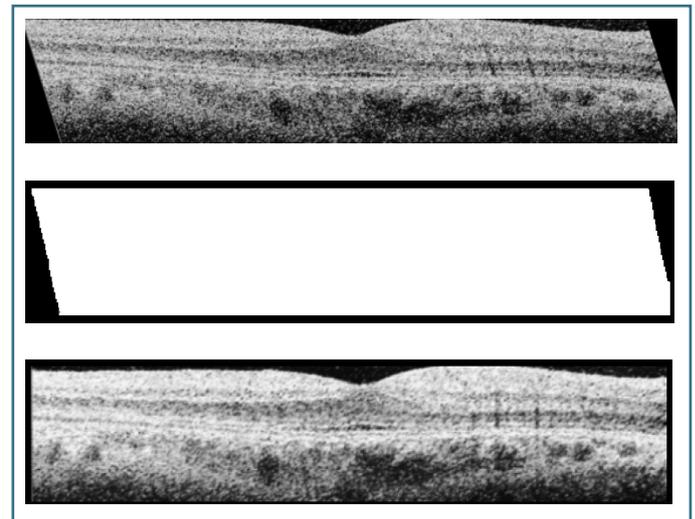


Figura 7. Proceso de corrección de la oblicuidad de la imagen OCT⁴.

sado de vasos y capilares se basan en la detección de bordes y de crestas (*ridges*), como por ejemplo, el filtro *hessiano*⁵, filtro *frangi*⁶, filtro *sato*⁷, filtro *meijering*⁸, entre otros. Algunas de las transformaciones utilizadas son: los métodos de umbralización, los operadores de morfología, los filtros de enfoque/desenfoque, la transformada rápida de Fourier y la transformada polar (Figura 8, Figura 9 y Figura 10).

Superresolución en imágenes OCTA

Si bien la idea de tener una mayor resolución de imagen puede parecer beneficioso, en realidad, para un modelo de IA, resulta más importante tener imágenes de mayor calidad, lo cual significa imágenes con máximo contraste y mínimo ruido.

Para ello, resulta interesante el procedimiento de mejora conocido como superresolución (SR). El objetivo de la SR es reconstruir una imagen para conseguir una de mayor tamaño, con mejor contraste, nitidez y SNR.⁹.

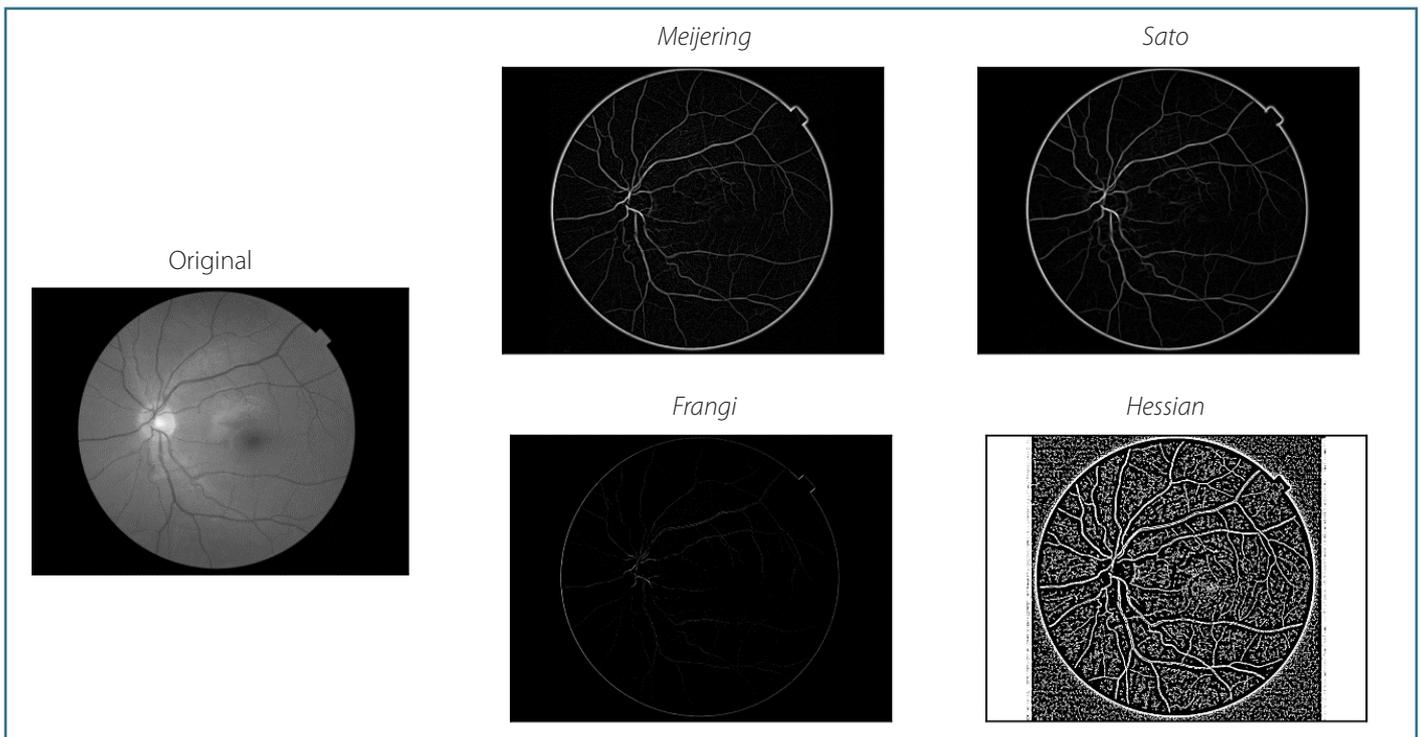


Figura 8. Ejemplo de aplicación de filtros de detección de crestas a imágenes de retinografía⁴.

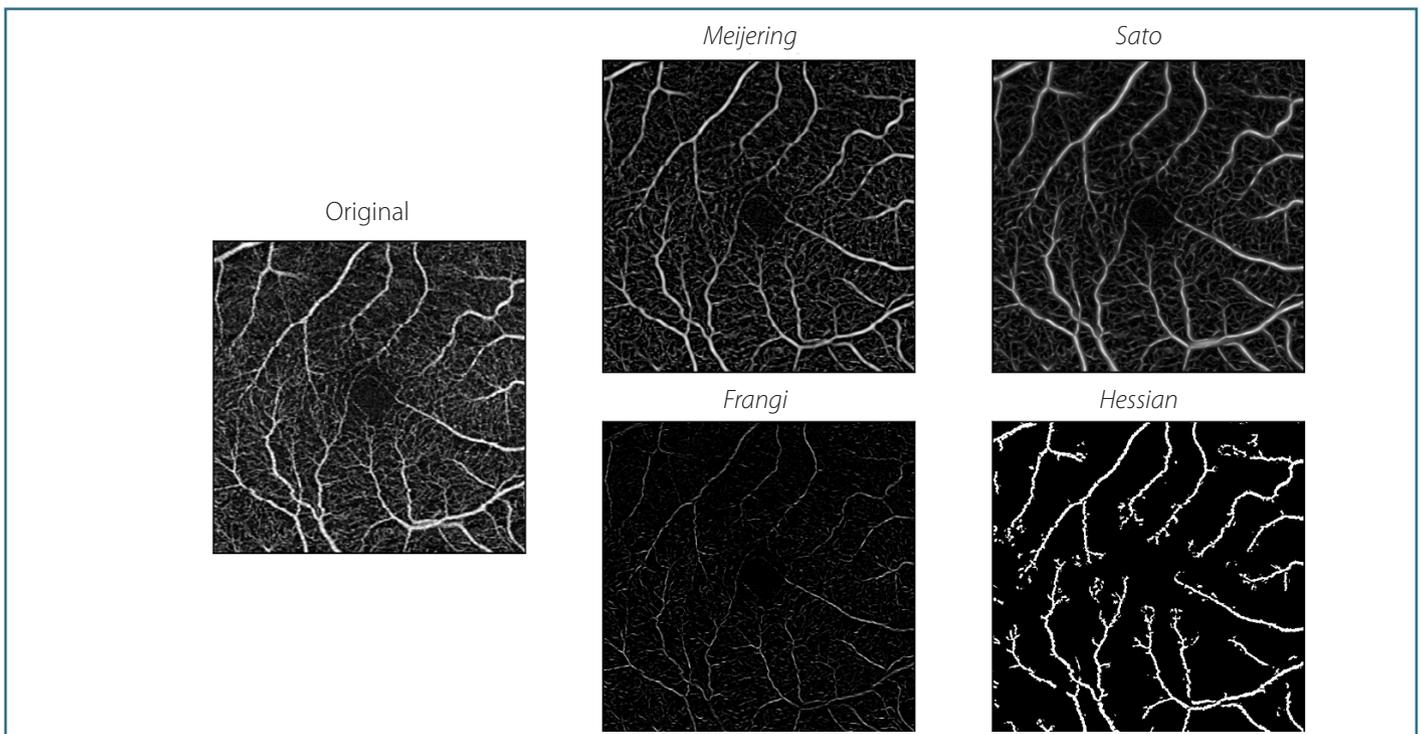


Figura 9. Ejemplo de aplicación de filtros de detección de crestas a imágenes OCTA⁴.

Para obtener las imágenes mejoradas con SR, se suele utilizar una DNN. Las redes neuronales de SR reciben una imagen de baja resolución como entrada y buscan dar como salida la correspondiente imagen mejorada (Figura 11). Usualmente, el conjunto de datos

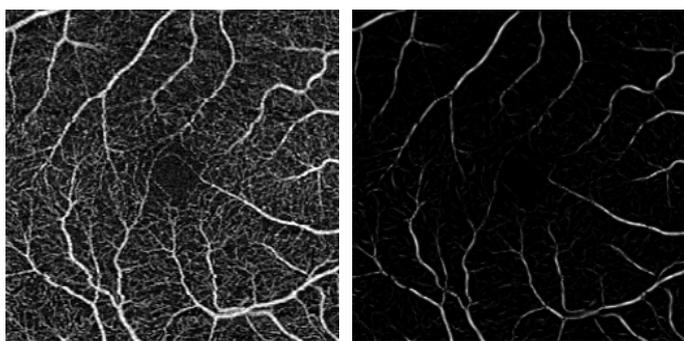


Figura 10. Ejemplo de aplicación del filtro *frangi* y filtros de desenfoque en una imagen OCTA⁴.

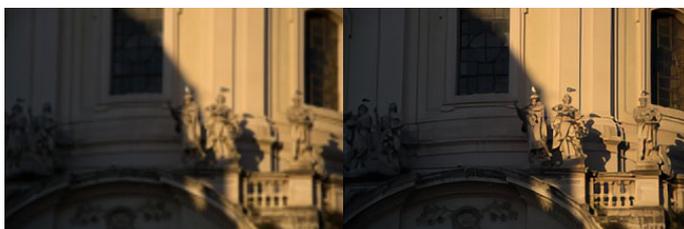


Figura 11. Comparativa entre una imagen de baja calidad y la mejorada por el *turing image super-resolution*¹⁰.

para la tarea se construye recopilando imágenes, degradándolas para obtener las que se utilizan como entrada. La degradación suele realizarse aplicando aleatoriamente distorsiones de distintos grados, tales como desenfoque, adición de ruido y adición de artefactos.

Para imágenes OCTA de alta resolución, el proceso de la SR implica mejorar el detalle y el contraste de los vasos sanguíneos respecto del fondo. El objetivo es que los vasos sanguíneos y las patologías presentes en las imágenes sean más reconocibles para el modelo de IA¹¹. Para las imágenes OCTA, lo más habitual en la literatura es utilizar una red neuronal convolucional (CNN, *convolutional neural network*), como DCARnet¹².

Las imágenes OCTA se pueden obtener a diferentes profundidades, permitiendo identificar, por ejemplo, el plexo vascular superficial (SCP, *superficial capillary plexus*) y el plexo vascular profundo (DCP, *deep capillary plexus*). Además, se pueden adquirir a diferentes resoluciones de escaneo, típicamente 3 x 3 mm, 6 x 6 mm, 8 x 8 mm, que determinan cuanta área se captura.

En el proceso de la SR, aprovechando las distintas capturas, en lugar de usar imágenes 3 x 3 mm degradadas para entrenar a la red neuronal, se ha explorado a usar las obtenidas de duplicar el tamaño de las imágenes 6 x 6 mm, alinearlas con su correspondiente de 3 x 3 mm y quedarse con el área que se superpone. El

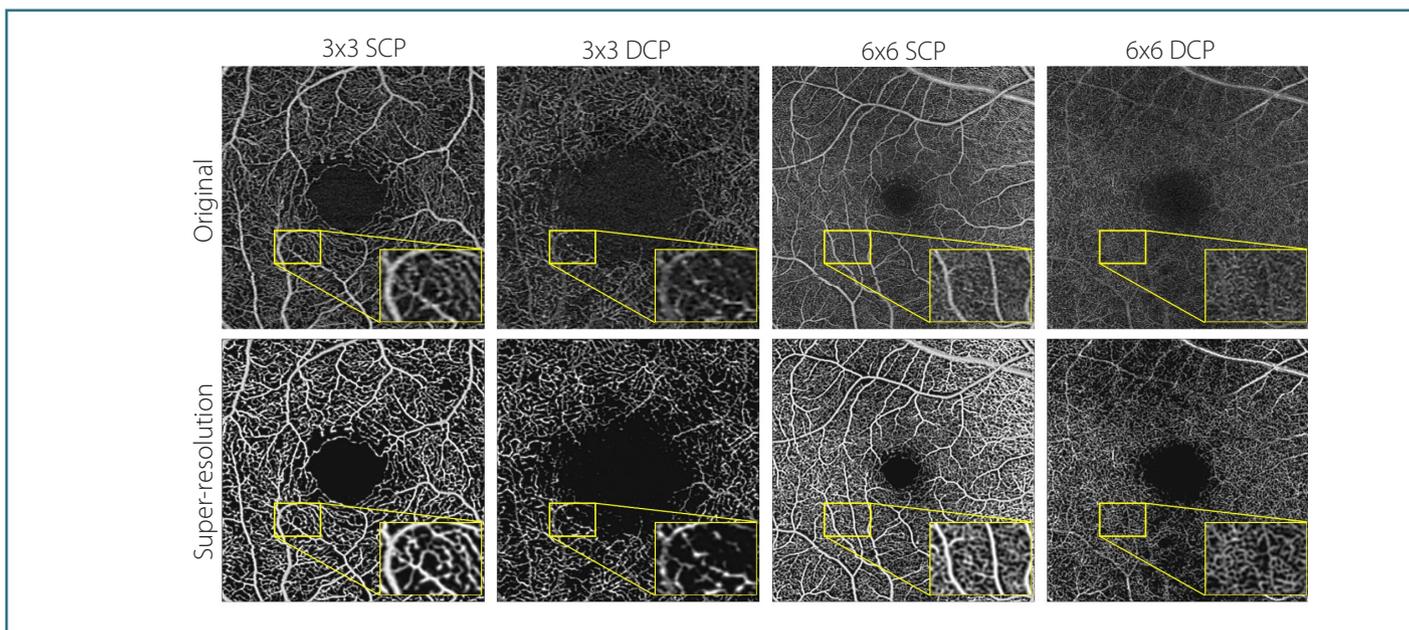


Figura 12. Comparativa entre las imágenes OCTA originales y las superresolución con DCARnet en un ojo con retinopatía diabética referible a dos resoluciones de escaneo (3x3 mm y 6x6 mm) y a dos profundidades (SCP y DCP)¹¹.

modelo de SR recibe las imágenes divididas en trozos para que el proceso sea escalable y agnóstico al tamaño de la imagen¹¹.

La imagen resultante del proceso de SR estará lista para ser usada en la DNN de clasificación (Figura 12). Experimentalmente, se ha observado que este procedimiento permite obtener mejores resultados por un pequeño pero significativo margen¹¹.

Conclusión

Debido a la dificultad de conseguir gran cantidad de datos y a la gran heterogeneidad en las muestras, la preparación y limpieza de las imágenes es, y parece que seguirá siendo, un paso decisivo previo a la aplicación de modelos de IA.

En Cataluña, la investigación y desarrollo en el ámbito de la imagen médica enfocada a la retina es sumamente activa, habiéndose investigado soluciones para abordar algunos de los retos más importantes.

Bibliografía

1. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data*. 2019;6(1):60.
2. Graham B. Kaggle Diabetic Retinopathy Detection competition report. [Internet]. En: Kaggle.com. Kaggle; 2015. Disponible en: <https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15801>
3. Benali A, Carrera L, Christin A, Martín R, Alé A, Barraso M, et al. NMF for Quality Control of Multi-modal Retinal Images for Diagnosis of Diabetes Mellitus and Diabetic Retinopathy. En: Rojas I, Valenzuela O, Rojas F, Herrera LJ, Ortuño F, eds. *Bioinformatics and Biomedical Engineering Parte II*. Cham: Springer International Publishing; 2022. p. 343-56.
4. Benali Bendahmane A. Using Machine Learning on the sources of retinal images for diagnosis by proxy of Diabetes Mellitus and Diabetic Retinopathy. [Tesis de maestría]. [Internet]. Barcelona: Facultat d'Informàtica de Barcelona. Universitat Politècnica de Catalunya; 2021. Disponible en: <http://hdl.handle.net/2117/356881>
5. Ng CC, Yap MH, Costen N, Li B. Automatic Wrinkle Detection Using Hybrid Hessian Filter. En: Cremers D, Reid I, Saito H, Yang MH, eds. *Computer Vision -- ACCV 2014*. Cham: Springer International Publishing; 2015. p. 609-22.
6. Frangi AF, Niessen WJ, Vincken KL, Viergever MA. Multiscale vessel enhancement filtering. En: Wells WM, Colchester A, Delp S, eds. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*. Heidelberg: Springer Berlin Heidelberg; 1998. p. 130-7.
7. Sato Y, Nakajima S, Shiraga N, Atsumi H, Yoshida S, Koller T, et al. Three-dimensional multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. *Med Image Anal*. 1998;2(2):143-68.
8. Meijering E, Jacob M, Sarria JCF, Steiner P, Hirling H, Unser M. Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytometry A*. 2004;58A(2):167-76.
9. Yang W, Zhang X, Tian Y, Wang W, Xue JH, Liao Q. Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Trans Multimed*. 2019;21(12):3106-21.
10. Microsoft Turing. Introducing Turing Image Super Resolution: AI powered image enhancements for Microsoft Edge and Bing Maps. [Internet]. En: Microsoft Bing Blogs. Disponible en: <https://blogs.bing.com/search-quality-insights/may-2022/Turing-Image-Super-Resolution>
11. Zurita Martel Y. Predicting Diabetic Retinopathy in OCTA Images using a Deep Learning Framework. [Tesis de maestría]. Barcelona: Universitat Politècnica de Catalunya; 2023.
12. Gao M, Hormel TT, Wang J, Guo Y, Bailey ST, Hwang TS, et al. An Open-Source Deep Learning Network for Reconstruction of High-Resolution OCT Angiograms of Retinal Intermediate and Deep Capillary Plexuses. *Transl Vis Sci Technol*. 2021;10(13):13.