

Métodos de ayuda al diagnóstico de glaucoma basado en diferencias de características extraídas de ambos ojos en tomografías de coherencia óptica

David Palazón Palau, Francisco Rodríguez Robles, Oleksandr Kovalyk,
Rafael Verdú Monedero, Juan Morales Sánchez

Universidad Politécnica de Cartagena, 30202 Cartagena, España
david.palazon@edu.upct.es, franroroba@gmail.com, oleksandr.kovalyk@edu.upct.es,
rafael.verdu@upct.es, juan.morales@upct.es

Resumen

Una de las enfermedades más relevantes que impacta en el nervio óptico es el glaucoma, provocando un deterioro progresivo e irreversible que disminuye el campo de visión del paciente. El espesor de la capa de fibras nerviosas de la retina se convierte en un indicador crucial para evaluar la condición y la evolución de esta afección. En el presente artículo se muestran resultados favorables para los métodos de ayuda al diagnóstico a partir de tomografías de coherencia óptica de las cuales se han extraído los grosores de cada sector de la capa de fibras, se le han aplicado métricas de asimetría entre ambos ojos y se han realizado varios estimadores basados en modelos de Machine Learning. Las asimetrías presentes en ambos ojos han resultado ser informativas respecto a la detección de pacientes con glaucoma. Los resultados muestran la efectividad de los métodos de ayuda al diagnóstico desarrollados y una mejora evidente en la clasificación de los pacientes con el uso de Random Forest respecto de Decision Tree, logrando clasificar correctamente en media a un 87,5 % de los pacientes con neuropatía óptica glaucomatosa.

1. Introducción

El glaucoma, una enfermedad que daña el nervio óptico, es una de las principales causas de ceguera en los países desarrollados, después de las cataratas. Se estima que en 2020, alrededor de 80 millones de personas padecían glaucoma [1], y se espera que esta cifra aumente a 111,8 millones para 2040 [2]. Esta enfermedad es crónica y generalmente no presenta síntomas en sus primeras etapas, lo que resalta la importancia del diagnóstico temprano para prevenir la pérdida permanente de la visión. Por ello, se busca desarrollar métodos automatizados y simplificados para el diagnóstico, así como mejorar los ya existentes [3], con el objetivo de hacerlos accesibles en los centros de atención primaria y mejorar las posibilidades de prevención y tratamiento temprano.

La neuropatía óptica por glaucoma es una condición que afecta al nervio óptico y puede tener diversas causas. Principalmente, esta afección se desarrolla en la parte frontal del ojo debido a desequilibrios en la producción y drenaje del humor acuoso, lo que aumenta la presión intraocular y, en última instancia, daña el nervio óptico produciendo una pérdida progresiva e irreversible del campo visual que deriva en ceguera [4]. El diagnóstico clínico del glaucoma se basa en la medición de la presión intraocular mediante tonometría, la evaluación del campo visual mediante campimetría y el análisis de

la anatomía ocular, utilizando técnicas como la retinografía y la tomografía de coherencia óptica (OCT).

Estudios previos han centrado su hipótesis acerca de las asimetrías interoculares de la capa de fibras nerviosas de la retina (retinal nerve fiber layer, RNFL) con umbrales estadísticos sobre pacientes sanos y pacientes con glaucoma [5], proponiendo a los algoritmos como potenciales métodos de ayuda al diagnóstico. Estudios como [3, 6, 7] han desarrollado y confirmado dicha propuesta.

Este artículo se centra en el avance de los métodos de ayuda al diagnóstico ya desarrollados desde un paradigma de la sencillez e interpretabilidad de los modelos basados en características de asimetría entre ambos ojos de un mismo paciente, así como del avance y el uso de modelos más complejos pero potentes.

2. Materiales y Métodos

Esta sección ofrece una breve descripción de los recursos utilizados en el estudio, específicamente, las imágenes de tomografía de coherencia óptica 2D (modo B-scan), que abarcan la región alrededor del nervio óptico y que conforman la base de datos, así como el proceso de obtención de estas imágenes y su posterior procesado. Este tipo de imágenes médicas se emplea de manera extensa en el diagnóstico de la enfermedad del glaucoma, ya que permite determinar los espesores de la RNFL.

2.1. Base de datos

La base de datos, utilizada en [3, 6, 8], está conformada por las imágenes de OCT de ambos ojos de 207 pacientes de los cuales 160 han sido etiquetados como sanos y 47 como patológicos. El etiquetado manual por parte de expertos del Servicio de Oftalmología del Hospital General Universitario Reina Sofía en Murcia, España, donde se utilizó un OCT Spectralis S2610-CB de Heidelberg Engineering GmbH, ha permitido realizar estudios sobre las características y asimetrías de la RNFL mediante algoritmos de aprendizaje máquina supervisado.

2.2. Adquisición de imágenes

Las imágenes OCT-SD en el dominio espectral, también conocidas como Fourier-domain OCT (FD-OCT), se han capturado utilizando un escaneo con un diodo láser superluminiscente (SLD) de longitud de onda de 870 nm. Estas imágenes registran una sección cilíndri-

ca alrededor del disco óptico de las capas de la retina (Figura 1).

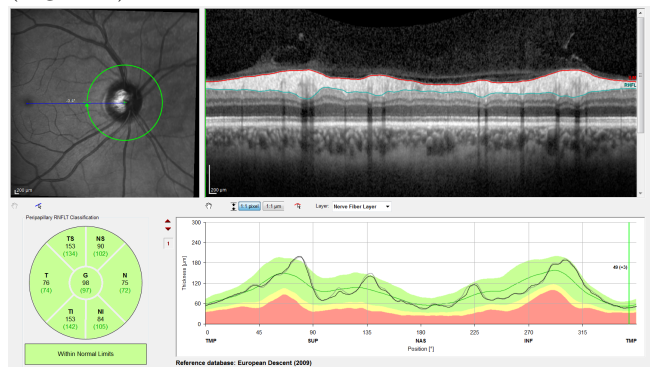


Figura 1. Captura de pantalla del espectrómetro Spectralis. En la parte superior, se muestra una imagen de la retina centrada en el disco óptico, y se indica la ubicación del barrido B peripapilar mediante un círculo verde. A la derecha, se muestra el barrido B peripapilar OCT en 2D con reconstrucción cartesiana. En las imágenes se incluyen coordenadas y valores medios estimados del grosor de la capa de fibras nerviosas de la retina (RNFL) para cada sector y el promedio general. También se muestra el contorno rectificado de la RNFL con el grosor estimado y los valores de referencia de la base de datos de ascendencia europea (2009).

Esta sección circular se divide en varios sectores, denominados temporal (T), temporal superior (TS), nasal superior (NS), nasal (N), nasal inferior (NI) y temporal inferior (TI). En cada sector se mide el grosor de la capa de fibras nerviosas de la retina.

Para facilitar el estudio y el procesamiento de estas imágenes, el B-scan circular se proyecta de coordenadas polares a coordenadas cartesianas en el eje x . En la parte inferior de la Figura 1 se detalla la correspondencia entre los valores medidos en grados en coordenadas polares y los valores medidos en píxeles en coordenadas cartesianas.

Las imágenes de OCT utilizadas en este estudio tienen una resolución de 768×496 píxeles, una profundidad de 8 bits por píxel en escala de grises y un escalado en z de $3,87 \mu\text{m}$ por píxel. El escalado en x depende del diámetro de la circunferencia peripapilar de la OCT y no es significativo para este estudio, ya que el objetivo principal es determinar el grosor vertical del RNFL en la imagen de la OCT. En la parte superior izquierda de la Figura , se muestra el trazado circular del haz del láser SLD posicionado sobre el centro del nervio óptico, también conocido como papila, en la imagen de fondo de ojo de la retina. La correspondiente proyección de coordenadas cartesianas del B-scan de OCT se presenta en la parte superior derecha de la Figura 1.

2.3. Segmentación de la RNFL y cálculo del grosor de los sectores en OCT

El siguiente paso se enfoca en la segmentación de las imágenes con el fin de estimar el grosor de la capa de fi-

bras nerviosas de la retina. La segmentación de la RNFL representa un gran desafío debido a las características de las imágenes, especialmente las propias de la capa retinal en análisis. Como se menciona en los estudios previos [9, 10], es común encontrar ruido especular, un bajo nivel de contraste y características morfológicas irregulares en los conjuntos de datos de OCT. Además, dado que la RNFL es la capa más interna de la retina, las imágenes pueden presentar artefactos debido al efecto de sombreado causado por las venas retinianas, lo cual dificulta definir con precisión las capas en sectores específicos del ojo. En [11] se realiza una revisión exhaustiva de los enfoques para la segmentación, algunos de los cuales son relevantes para el conjunto de datos utilizado en este estudio.

La segmentación utilizada en este trabajo es la proporcionada por el software Spectralis versión 6.9.4.0. Utilizando este algoritmo, se calcula el grosor de la RNFL a partir de las posiciones de los límites superior e inferior de la capa proporcionados por el algoritmo de segmentación. Estas posiciones se escalan de píxeles a micrómetros según la resolución de OCT y se proporciona el grosor promedio de la RNFL para cada sector del ojo (T, TS, NS, N, NI y TI), así como el valor promedio de grosor del contorno circumpapilar global (G).

2.4. Métricas de asimetría aplicadas a los grosores de la RNFL

Se recogen las métricas de asimetría que mejores resultados ofrecen, anteriormente propuestas en [3, 6] y se plantea una nueva en la que se tiene en cuenta una mayor implicación del sector G, el cual ha demostrado ser el sector con mayor importancia en la modalidad de imagen médica de OCT en los modelos de clasificación entrenados en otros trabajos como [6, 7]. Las métricas consideradas en este trabajo son:

$$|\Delta_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S,i}^r + w_{S,i}^l}, \quad (1)$$

$$|\bar{\Delta}_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{G,i}^r + w_{G,i}^l}, \quad (2)$$

$$|\Delta_{S,i}|^{0,5} = \sqrt{\frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S,i}^r + w_{S,i}^l}}, \quad (3)$$

$$|\tau_{S,i}| = \left| \frac{w_{S,i}^r + w_{G,i}^r}{w_{G,i}^l} - \frac{w_{S,i}^l + w_{G,i}^l}{w_{G,i}^r} \right|, \quad (4)$$

donde S indica el sector concreto (TS, T, TI, NS, N, NI, G) sobre el que se calcula la asimetría, i indica el paciente y r y l hacen referencia al ojo derecho e izquierdo, respectivamente [3].

2.5. Preprocesado y estimador

Una vez extraídas las medidas de los grosores de la RNFL y aplicada la métrica de asimetría, los datos de cada paciente están preparados para el entrenamiento

y evaluación del estimador. Dicho estimador está compuesto por un preprocesado y un algoritmo de aprendizaje máquina. La Figura 2 muestra el flujo que se ha utilizado en los datos para el entrenamiento y validación.

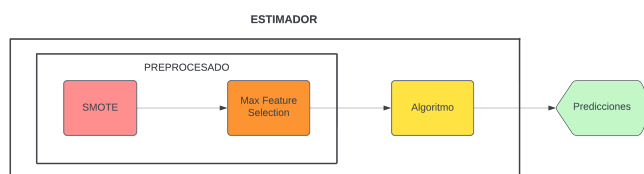


Figura 2. Flujo interno del estimador utilizado..

El flujo de ejecución es el siguiente. Inicialmente se parte de la base de datos original y se obtienen los valores de asimetría interocular de cada paciente para cada una de las métricas de asimetría propuestas. A partir de dichos valores se realiza una optimización en dos pasos del estimador completo, primero haciendo uso de una búsqueda aleatoria (*Random Search*) y posteriormente, realizando una búsqueda por cuadrícula (*Grid Search*) en un conjunto de valores definidos entorno a los mejores hiperparámetros del paso anterior. En todos los pasos del proceso se garantiza la independencia de la información del conjunto de datos de entrenamiento y de validación.

Debido al desbalanceo de la clase objetivo, donde solo hay 47 pacientes con glaucoma frente a los 160 pacientes sanos, es necesario aplicar técnicas de sobremuestreo de la clase minoritaria para realizar un balanceo de clases como SMOTE (*Synthetic Minority Over-sampling TEchnique*) [12] sobre el conjunto de entrenamiento. Además, se ha realizado una validación cruzada repetida estratificada y aleatoria (5 pliegues y 10 repeticiones), garantizando la distribución equitativa de clases y aumentando la calidad de los resultados. En adición a esta problemática, para reducir la dimensionalidad que evite el sobreajuste de los modelos utilizando un *Feature Selector*, se han seleccionado las características más importantes para la clasificación según el MDI (*Mean Decrease Impurity*) [13] obtenido de un Random Forest genérico.

3. Resultados

Esta sección presenta los mejores resultados obtenidos en ambos modelos junto con la métrica de asimetría asociada a dichos resultados, es decir, el mejor estimador basado en *Decision Tree* (DT) [14] y el mejor estimador basado en *Random Forest* (RF) [15]. El criterio de elección del mejor estimador para cada uno de los modelos utilizados se fundamenta principalmente en la métrica de clasificación conocida como *Recall*, que en este estudio denominamos *Sensitivity*.

A continuación se presenta en la Tabla 1 los mejores resultados obtenidos, mostrando el valor medio y dos veces la desviación típica, con el objetivo de tener en cuenta no solo el valor medio de dicha métrica de clasificación, sino también, la variabilidad de los resultados.

Por ello, el mejor resultado será el que tenga un mayor valor promedio de *Sensitivity* acompañado de la menor desviación típica.

ML Model	Decision Tree	Random Forest
Assymetry metric	$ \Delta $	$ \Delta ^{0,5}$
Sensitivity	0.810 ± 0.214	0.875 ± 0.154
Precision	0.619 ± 0.278	0.661 ± 0.187
Accuracy	0.822 ± 0.140	0.862 ± 0.098
Specificity	0.826 ± 0.208	0.858 ± 0.114

Tabla 1. Resultados de la validación cruzada de las métricas de asimetría destacadas para cada modelo y modalidad de imagen médica.

En primer lugar, podemos ver como las métricas de asimetría dominantes son distintas en cada modelo. Esto puede indicar que el rendimiento del modelo está ligado a la métrica de asimetría utilizada y que distintos modelos rinden mejor con distintas métricas de asimetría. Sin embargo, a la vista de la variabilidad de los resultados, este razonamiento aunque sea lógico carece de valor ya que no se puede demostrar. Por este motivo resulta interesante aumentar el número de muestras y reducir la varianza de los modelos pudiendo así ofrecer resultados más realistas y profundizar en dicha hipótesis.

En segundo lugar, se observa una mejora en el rendimiento de todas las métricas de clasificación al pasar de DT a RF tanto en promedio como en desviación típica. RF aumenta un 8% el valor promedio de *Sensitivity* y reduce su desviación típica en un 38%. La métrica de asimetría propuesta en este trabajo ($|\tau|$) no ha conseguido destacar en ninguno de los dos modelos.

En la Figura 3 se muestran las matrices de confusión de ambos estimadores. Estas se han obtenido con una validación cruzada aleatoria y estratificada de 5 pliegues donde en cada iteración se realizaba el entrenamiento y la posterior clasificación del conjunto de validación. Una vez completadas las iteraciones, mediante los identificadores de cada muestra se realizaba la comparativa entre la clasificación del modelo y la clasificación real para generar dicha matriz de confusión.

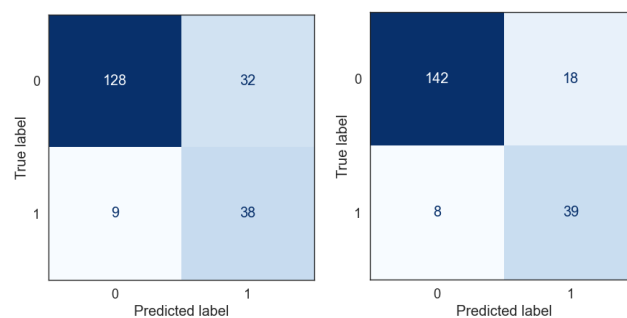


Figura 3. Matrices de confusión del mejor estimador de DT y RF.

En trabajos anteriores donde se ha utilizado la misma base de datos para entrenar y evaluar modelos basados en *Decision Trees* [3, 6], se han hallado los siguientes resultados *Sensitivity* 80.9%, *Specificity* 88.2%, *Precision* 66.7%, *Accuracy* 86.5%. Por otro lado, en [7] se obtienen los siguientes resultados en promedio para una base de datos distinta pero haciendo uso de *Random Forest*, 88.2% *Accuracy*, 91.7% *Sensitivity* y 85.1% *Specificity*.

Aunque no óptimos, los resultados obtenidos en este trabajo son favorables y evidencian que la asimetría en la morfología de la RNFL en ambos ojos junto con la ayuda del aprendizaje máquina son un potencial método de ayuda al diagnóstico precoz de glaucoma.

4. Conclusiones

Este estudio se enfocó en mejorar los métodos de ayuda al diagnóstico del glaucoma mediante el uso de técnicas de *Machine Learning*, utilizando una perspectiva diferente a la de investigaciones previas como [6] y [3]. Se emplearon diversas técnicas de análisis de datos y se llevaron a cabo en un entorno de desarrollo diferente.

El uso de la misma base de datos que los estudios anteriores destacó la importancia de las métricas de asimetría en la morfología de ambos ojos para diagnosticar el glaucoma. Sin embargo, este estudio logró resultados significativamente superiores a los previos, estableciéndose como un punto de referencia para investigaciones futuras en esta área.

La elección de árboles de decisión se basó en su sencillez y facilidad de interpretación, con la intención de consolidar y mejorar resultados anteriores. Además, se consideró el uso de *Random Forest* como una opción más compleja para mejorar aún más la precisión del diagnóstico.

Los resultados destacan la efectividad de los modelos basados en *Random Forest* como los principales métodos de apoyo en el diagnóstico de glaucoma.

La limitación de datos presentó un desafío, ya que intentar abordar el sesgo y la varianza simultáneamente resultaba complicado debido al compromiso sesgo-varianza. La implementación de *Random Forest* abordó este problema reduciendo la varianza, aunque sacrificando la interpretabilidad del modelo.

Para futuras mejoras, es esencial aumentar el tamaño de la muestra de datos disponible. También se pueden explorar estrategias de optimización de hiperparámetros y la inclusión de características adicionales para mejorar la representación de los datos. La investigación de modelos completamente nuevos, aunque menos interpretables, como *XGBoost* o *causalML*, podría conducir a avances significativos en el diagnóstico del glaucoma.

Referencias

- [1] Jesus Claramunt and Noelia Fernandez. Glaucoma y poblacion en 2020. *Gac Optom optica Oftalmica*, 2016.
- [2] Yih-Chung Tham, Xiang Li, Tien Y Wong, Harry A Quigley, Tin Aung, and Ching-Yu Cheng. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11):2081–2090, 2014.
- [3] R. Berenguer, R. Verdu, J. Morales, I. Selles, O. Kovalyk, and J.L. Sancho. Decision trees for glaucoma screening based on the asymmetry of the RNFL in OCT. *Sensors*, 22(13), 2022.
- [4] Kwon YH, Figert JH, Kuehn MH, and Alward WL. Primary open-angle glaucoma. *N Engl J Med.*, 360(11):1113–1124, 2009.
- [5] Matthew G Field, Tarek Alasil, Neda Baniyadi, Christian Que, Huseyin Simavli, Doaa Sobeih, David Soladell Valle, Matthew J Best, and Teresa C Chen. Facilitating glaucoma diagnosis with intereye retinal nerve fiber layer asymmetry using Spectral-Domain OCT. *J Glaucoma*, 25(2):167–176, February 2016.
- [6] F. Rodriguez, R. Verdu, R. Berenguer, J. Morales, I. Selles. Analysis of the asymmetry between both eyes in early diagnosis of glaucoma combining features extracted from retinal images and octs into classification models. *Sensors*, 23(10), 2023.
- [7] Chao-Wei Wu, Hsiang-Li Shen, Chi-Jie Lu, Ssu-Han Chen, and Hsin-Yi Chen. Comparison of different machine learning classifiers for glaucoma diagnosis based on spectralis oct. *Diagnostics*, 11(9), 2021.
- [8] O. Kovalyk, J. Morales, R. Verdu, I. Selles, A. Palazon, and J.L. Sancho. Papila: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. *Scientific Data*, 9(1):1–12, 2022.
- [9] Eric A. Swanson and James G. Fujimoto. The ecosystem that powered the translation of oct from fundamental research to clinical and commercial impact. *Biomed. Opt. Express*, 8(3):1638–1664, Mar 2017.
- [10] David Huang, Eric A. Swanson, Charles P. Lin, Joel S. Schuman, William G. Stinson, Warren Chang, Michael R. Hee, Thomas Flotte, Kenton Gregory, Carmen A. Puliafito, and James G. Fujimoto. Optical coherence tomography. *Science*, 254(5035):1178–1181, 1991.
- [11] R. Berenguer, R. Verdú, J. Morales, I. Sellés, R. Del Amor, G. García, and V. Naranjo. Automatic segmentation of the retinal nerve fiber layer by means of mathematical morphology and deformable models in 2D OCT imaging. *Sensors*, 21(23):8027, 2021.
- [12] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [13] Eugene Tuv, Alexander Borisov, George Runger, and Kari Torkkola. Feature selection with ensembles, artificial variables, and redundancy elimination. *The Journal of Machine Learning Research*, 10:1341–1366, 2009.
- [14] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. Classification and regression trees. *ISBN-13*, pages 978–0412048418, 1984.
- [15] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, Oct 2001.