Decision Trees for Glaucoma Screening Based on the Asymmetry of the Retinal Nerve Fiber Layer in Optical Coherence Tomography

Rafael Berenguer-Vidal 1, Rafael Verdú-Monedero 2,*, Juan Morales-Sánchez 2, Inmaculada Sellés-Navarro 3, Oleksandr Kovalyk 2 and José-Luis Sancho-Gómez 2

1 Departamento de Ciencias Politécnicas, Universidad Católica de Murcia UCAM, 30107 Guadalupe, Spain; rberenguer@ucam.edu
2 Departamento de Tecnologías de la Información y Comunicaciones, Universidad Politécnica de Cartagena, 30202 Cartagena, Spain; juan.morales@upct.es (J.M.-S.); oleksandr.kovalyk@edu.upct.es (O.K.); josel.sancho@upct.es (J.-L.S.-G.)
3 Hospital General Universitario Reina Sofía, 30003 Murcia, Spain; inmasell@um.es
* Correspondence: rafael.verdu@upct.es

Abstract: Purpose: The aim of this study was to analyze the relevance of asymmetry features between both eyes of the same patient for glaucoma screening using optical coherence tomography. Methods: Spectral-domain optical coherence tomography was used to estimate the thickness of the peripapillary retinal nerve fiber layer in both eyes of the patients in the study. These measurements were collected in a dataset from healthy and glaucoma patients. Several metrics for asymmetry in the retinal nerve fiber layer thickness between the two eyes were then proposed. These metrics were evaluated using the dataset by performing a statistical analysis to assess their significance as relevant features in the diagnosis of glaucoma. Finally, the usefulness of these asymmetry features was demonstrated by designing supervised machine learning models that can be used for the early diagnosis of glaucoma. Results: Machine learning models were designed and optimized, specifically decision trees, based on the values of proposed asymmetry metrics. The use of these models on the dataset provided good classification of the patients (accuracy 88%, sensitivity 70%, specificity 93% and precision 75%). Conclusions: The obtained machine learning models based on retinal nerve fiber layer asymmetry are simple but effective methods which offer a good trade-off in classification of patients and simplicity. The fast binary classification relies on a few asymmetry values of the retinal nerve fiber layer thickness, allowing their use in the daily clinical practice for glaucoma screening.

Keywords: optical coherence tomography (OCT); peripapillary OCT; retinal nerve fiber layer (RNFL); RNFL thickness asymmetry; retinal imaging analysis; glaucoma; decision trees

1. Introduction

Glaucoma is an eye disease which damages the optic nerve head (ONH) and can produce complete loss of vision. After cataracts, it is the leading cause of irreversible blindness worldwide [1]. In 2013 there were 64.3 million people aged 40–80 with glaucoma worldwide, and this number is estimated to increase to 111.8 million in 2040 [2]. The term glaucoma describes ocular disorders with multi-factorial etiology which usually are united with clinical intraocular-pressure-associated optic neuropathy. There are different types of glaucoma, classically divided into the categories of primary or secondary open-angle or angle-closure glaucoma [3]. Primary open-angle glaucoma (POAG) is the most common form of glaucoma, where the drainage angle formed by the cornea and iris remains open, but the trabecular meshwork is partially blocked, causing an increase in the intraocular pressure (IOP) [4]. Secondary forms of glaucoma are caused by various ocular or systemic diseases; most of them are associated with high IOP, but not all, such as normal-tension...
glaucoma [5]. Primary open-angle glaucoma is usually associated with an increase in the intraocular pressure which modifies the appearance of the optic nerve, causing the neuroretinal rim of the optic nerve to become progressively thinner, thereby enlarging the optic-nerve cup. The different types of glaucoma have in common a slow progressive degeneration of retinal ganglion cells and their axons, resulting in a distinct appearance of the optic disc and a concomitant pattern of visual loss [6]. Patients with glaucoma typically lose peripheral vision and may lose all vision if not treated properly [7].

Glaucoma is a silent and progressive illness which is initially asymptomatic. The diagnosis is based on tonometry to assess the intraocular pressure, campimetry to evaluate the visual field, and imaging modalities to analyze the morphology of the ONH and its retinal layers. The internal structure of the eye can be imaged by retinal fundus images and optical coherence tomography [8].

Optical coherence tomography (OCT) is an imaging modality used in the assessment of the glaucomatous damage [9]. The OCT provides images with backscatter produced by the differences in the refractive index between adjacent retinal tissues [10,11]. The initial time-domain (TD)-OCT was surpassed by spectral domain (SD)-OCT, which offers increased axial resolution and faster scanning speeds [12]. SD-OCT allows one to analyze and quantify eye parameters, such as the thickness of the peripapillary XXXXX retinal nerve fiber layer (RNFL) for monitoring glaucoma [13,14], and evaluate neurostructural and vascular changes (see, e.g., [15]).

The damage produced by glaucoma alters the morphology of ocular structures, such as the RNFL and ONH, which allows for discrimination between glaucomic and healthy eyes. Those structural changes in the ONH or the thinning of the RNFL result in differences between the eyes of the same individual. For this reason, asymmetry has been analyzed and considered as an early indicator of glaucoma [16] using, e.g., the difference in values of intraocular pressure [17], central corneal thickness [18], corneal hysteresis [19], neuroretinal rim width [20] or as in this work, RNFL thickness [21]. While some studies determine the limits of the normal interocular asymmetry in retinal layers considering thickness measurements with OCT in normal subjects (see, e.g., [22–24]), other studies use RNFL thickness measurements to evaluate the diagnostic capabilities of intereye and intraeye differences to identify early primary open-angle glaucoma [25] or propose an early glaucoma discriminator index [26].

In this work, we focus on primary open-angle glaucoma and propose a set of asymmetry metrics based on the thickness of the RNFL of both eyes. Statistical analysis was performed to quantify the importance of each metric as a relevant characteristic for glaucoma diagnosis. A simple classifier was then applied to the most discriminant metric, obtaining a decision tree for the classification of healthy and glaucoma patients. This model offers a simple but effective tool for the screening and early detection of glaucoma that can be used in the daily clinical practice, similarly to the inferior-superior-nasal-temporal (ISNT) rule [27], the cup-to-disc ratio (CDR) [28] and the rim-to-disc ratio (RDR) [29] measurements. Note that these previous indicators are obtained from retinal fundus images, instead of OCT, which is one of the novelties of this work.

After this introduction, the rest of the article is organized as follows: Section 2 provides all the details of the dataset used and the proposed metric to evaluate the asymmetry. Section 3 shows the statistical analysis of the asymmetry metrics and describes the classification models, along with the metrics providing the greatest discrimination between healthy and glaucoma patients. Finally, Section 4 closes the article with a discussion and the conclusions.

2. Materials and Methods

This section describes the image acquisition procedure and the methods used in the assessment of the diagnosis of the disease.
2.1. Image Acquisition Procedure

We took a set of 2D peripapillary B-scan optical coherence tomography (OCT) images centered at the optic nerve head as raw material, from both healthy individuals and unhealthy patients with glaucoma. The OCT image dataset used in this work was acquired by the Ophthalmology Service of the Hospital General Universitario Reina Sofia (HGURS, Murcia, Spain) using a Spectralis OCT S2610-CB (Heidelberg Engineering GmbH, Heidelberg, Germany) from October 2018 to November 2020 from 287 individuals. These images were anonymized according to the criteria of the Human Ethics Committee.

The patients included in the study were divided into two groups. In the first group, G1, patients diagnosed with simple chronic glaucoma recruited in the Glaucoma Section of the HGURS; in the second group, G2, were primary care patients without any type of ocular pathology (verified by ophthalmologists) that could influence the morphology of the optic nerve. An exclusion criterion was the presence of opacities in the transparent media (corneal alterations and advanced cataracts) that prevented obtaining an assessable OCT. Another reason for excluding a patient was that the alteration in the confidence indices of the visual field persisted in two consecutive tests. The patients of G2 had IOP equal to or below 22 mmHg, and no lesions suggestive of glaucomatous neuropathy were detected. The sample size of patients was calculated by taking into account the population of Area VII of the city of Murcia and the prevalence of simple chronic glaucoma (3.5% in the population older than 40 years), while applying a confidence level of 95% and a 3% margin of error.

The spectral-domain optical coherence tomography (SD-OCT), also called Fourier-domain OCT (FD-OCT), was acquired by means of an 870 nm wavelength super-luminescent diode (SLD), scanning a cylindrical section of the retinal layers centered on the optic disc. This cylindrical section, also called a B-scan, was projected from polar to Cartesian coordinates resulting in a image with a resolution of 768 × 496 pixels, a bit depth of 8 bits/pixel in grayscale and a z-scaling of 3.87 µm/pixel, as illustrated in Figure 1. In order to facilitate the analysis, the cylindrical B-scan is normally divided into a group of 6 sectors whose names refer to their positions, namely, temporal (T), temporal superior (TS), nasal superior (NS), nasal (N), nasal inferior (NI) and temporal inferior (TI). Note that sectors (T) and (N) use an angle of 90°, whereas the rest of sectors use an angle of 45°. As shown in Figure 2, the naming of the sectors is symmetrical with respect to the right and left eyes. Thus, the analysis of the asymmetry can be performed by simple comparison of the thickness measurements between each paired eye sector. The relations between polar coordinates measured in degrees and Cartesian coordinates measured in pixels is provided in Table 1 and outlined at the bottom of Figure 1.

Figure 1. Segmented RNFL in a peripapillary B-scan OCT image. Relations of the sectors (T, TS, NS, N, NI and TI) measured in degrees and pixels.
Table 1. Relations between the polar and Cartesian coordinates in the OCT projection.

|                | Temporal (T) | Temporal Sup. (TS) | Nasal Sup. (NS) | Nasal (N) | Nasal Inf. (NI) | Temporal Inf. (TI) | Temporal (T) |
|----------------|--------------|--------------------|-----------------|-----------|----------------|-------------------|--------------|
| Polar (degrees) | Min 0°       | 45°                | 90°             | 135°      | 225°           | 270°              | 315°         |
|                | Max 45°      | 90°                | 135°            | 225°      | 270°           | 315°              | 360°         |
| Cartesian (pixels) | Min 1       | 97                 | 193             | 289       | 481            | 577               | 673          |
|                | Max 96       | 192                | 288             | 480       | 576            | 672               | 768          |

As described in [30], expert ophthalmologists diagnosed both eyes of all individuals in the database using three possible levels of glaucoma disease: healthy, not healthy and uncertain. Given the aim of this work to compare the asymmetry between the eyes of healthy and glaucoma patients, only individuals with two healthy eyes (160 individuals) and with both eyes with glaucoma disease (47 individuals) were selected (see Table 2). Hence, patients with only one eye diagnosed with glaucoma or those for whom the expert diagnosis was inconclusive were not considered.

Note that the sizes of the two groups are not comparable; i.e., the number of healthy individuals was larger than the number of unhealthy ones. This may suggest biased results in a subsequent classification process. Nevertheless, it is important to highlight that the individuals in the database were collected directly from patients attending the HGURS ophthalmology service; thus, the distribution of the two sample groups is similar to what can be found in a real clinical setting for glaucoma diagnosis.

Table 2. Summary of the OCT dataset used to evaluate the asymmetry metrics.

|                | Healthy (h) | Glaucoma (g) | Total |
|----------------|-------------|--------------|-------|
| Age (years)    | 59.10 ± 12.70 | 70.15 ± 9.21 | 61.61 ± 12.85 |
| Gender (male/female) | 53/107        | 19/28        | 72/135          |
| Total          | 160 patients (320 OCTs) | 47 patients (94 OCTs) | 207 patients (414 OCTs) |

2.2. RNFL Segmentation and Thickness Calculation

The next step is devoted to the segmentation of the images in order to estimate the thickness of the retinal nerve fiber layer (RNFL). The segmentation of the RNFL under analysis poses a great challenge due to the characteristics of the images, and more specifically, the features of the retinal layer under analysis. As described in [9,10], speckle noise, a low level of contrast and irregularly shaped morphological features are frequently present in OCT datasets. In addition, since the RNFL is the innermost layer of the retina, artifacts due to the shadowing effect of the retinal veins appear in the images, increasing the difficulty of precisely defining the layers in specific sectors of the eye. The most commonly used
approaches for this segmentation are comprehensively reviewed in [30], and some of them are relevant to the dataset used here.

The segmentation method chosen for this work is the algorithm provided by Spectralis software version 6.9.4.0. The thickness of the RNFL is calculated from the positions of the upper and lower boundary of the RNFL provided by the segmentation algorithm, scaled from pixels to µm according to the OCT resolution, as shown in Figure 3. Unlike other methods where the thickness value at each angle of the B-scan can be obtained, the Spectralis device provides only the average thickness of the RNFL for each sector of the eye (T, TS, NS, N, NI and TI) and the thickness average value of the global circumpapillary contour (G). Therefore, these mean section thickness values were used for the inter-eye asymmetry study (Section 3).

Figure 3. Screenshot of Spectralis. From left to right and top to bottom: Retinal fundus image centered on the optic disc (the green circle indicates the location of the peripapillary B-scan, which is shown on the right with the segmentation of the RNFL); 2D peripapillary B-scan OCT with Cartesian coordinates; estimated mean values for RNFL layer thickness for each sector and the overall mean; rectified outline of the RNFL with estimated thickness and reference values according to the database European Descent (2009).

2.3. Proposals for Asymmetry Metrics

In this section, a set of inter-eye asymmetry measures of RNFL thickness are proposed. The difference between the mean value of the RNFL thickness in the right eye and its corresponding thickness in the left eye is the base operation for all measures. Some metrics calculate absolute differences, in order to analyze only the difference between eyes by isolating the effect of which particular eye shows greater thickness. Nevertheless, the influence of the sign of the difference in the RNFL thickness on the discrimination between healthy patients and glaucoma patients has also been explored with other metrics. Additionally, given that the thickness of the retinal layers varies between individuals, a set of normalizers have been proposed to provide relative asymmetry values, with the objective of decoupling to some extent the specific dimensions of the eyes of the patients.
The first attempt to measure the asymmetry was with the following difference

\[ \delta_{S,i} = w_{S,i}^r - w_{S,i}^l, \quad (1) \]

where \( S \) denotes the specific sector (T, TS, NS, N, NI, TI or G) over which the asymmetry is calculated; \( i \) indicates the patient number; and \( w_{S,i}^r \) and \( w_{S,i}^l \) refer to the mean thickness of the RNFL layer in the sector \( S \) calculated for the right and left eyes of patient \( i \), respectively.

By applying the modulus to the previous equation, the direction in which the subtraction is calculated becomes irrelevant. Thus, the following measure is obtained:

\[ |\delta_{S,i}| = |w_{S,i}^r - w_{S,i}^l|. \quad (2) \]

The following proposals for the measurement of the asymmetry normalize previous measures with the sum of the RNFL thickness of both eyes in the corresponding sector \( S \) of the patient \( i \):

\[ \Delta_{S,i} = \frac{\delta_{S,i}}{w_{S}^r + w_{S}^l} = \frac{w_{S,i}^r - w_{S,i}^l}{w_{S}^r + w_{S}^l}, \quad (3) \]

and

\[ |\Delta_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S}^r + w_{S}^l}. \quad (4) \]

The normalization can also be performed with the average thickness in each sector of all the patients in the dataset:

\[ \bar{\Delta}_{S,i} = \frac{\delta_{S,i}}{w_{S}^r + w_{S}^l} = \frac{w_{S,i}^r - w_{S,i}^l}{w_{S}^r + w_{S}^l}, \quad (5) \]

and

\[ |\bar{\Delta}_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{S}^r + w_{S}^l}. \quad (6) \]

where the values \( w_{S}^r \) and \( w_{S}^l \) can be obtained as

\[ w_{S}^r = \frac{1}{N} \sum_{i=1}^{N} w_{S,i}^r, \quad (7) \]

and

\[ w_{S}^l = \frac{1}{N} \sum_{i=1}^{N} w_{S,i}^l, \quad (8) \]

i.e., the mean value of the RNFL thickness for each sector of both eyes. The values obtained with the dataset of OCTs used in this work \((N = 160 + 47 = 207)\) are gathered in Table 3.

Finally, the last proposed normalizations divide the thickness difference by the sum of the average thickness of the complete RNFL in both eyes of patient \( i \):

\[ \tilde{\Delta}_{S,i} = \frac{\delta_{S,i}}{w_{G,i}^r + w_{G,i}^l} = \frac{w_{S,i}^r - w_{S,i}^l}{w_{G,i}^r + w_{G,i}^l}, \quad (9) \]

and

\[ |\tilde{\Delta}_{S,i}| = \frac{|w_{S,i}^r - w_{S,i}^l|}{w_{G,i}^r + w_{G,i}^l}. \quad (10) \]

where \( w_{G,i}^r \) and \( w_{G,i}^l \) are, respectively, the average thickness of the complete RNFL in the right eye and in the left eye of patient \( i \).
Table 3. Mean values in µm of the RNFL thickness calculated for all sectors in both eyes. Values used in asymmetry metrics $\Delta_{S,i}$ and $|\Delta_{S,i}|$.

| Sector | Right | Left |
|--------|-------|------|
| TS     | 128.1691 | 126.3527 |
| T      | 68.7826 | 68.0531 |
| TI     | 130.0773 | 132.8454 |
| NS     | 101.2029 | 109.5266 |
| N      | 74.1932 | 70.9469 |
| NI     | 106.1353 | 106.6957 |
| G      | 93.9469 | 93.9710 |

3. Results

This section firstly addresses the statistical characterization of the RNFL thickness values using all previous asymmetry metrics. The mean, standard deviation and $p$-value were calculated for the subsets of healthy and glaucoma patients. Next, the chosen machine learning models, decision trees, are briefly described, and the design process for glaucoma screening is detailed.

3.1. Statistical Characterization of the Asymmetry Metrics

The thickness asymmetry has been calculated with the proposed metrics for each sector of the OCT and for the overall RNFL, considering two subsets, healthy and glaucoma patients (see Table 2). In order to analyze the relevance of each asymmetry metric, statistical characterization has been performed. Table 4 gathers the mean value, standard deviation and $p$-value of each metric for all sectors (TS, T, TI, NS, N and NI) and for entire OCT scan (G) for the two subsets of healthy and glaucoma patients. The values of the non-normalized metrics $\delta_{S}$ and $|\delta_{S}|$ are given in µm, whereas the values of the normalized metrics are dimensionless and range from $-1$ to $1$ in the case of $\Delta_{S}$, $\bar{\Delta}_{S}$ and $\bar{|\Delta|}$, and from $0$ up to $1$ in the case of the metrics based on the absolute value, i.e., $|\Delta_{S}|$, $|\bar{\Delta}_{S}|$ and $|\bar{|\Delta|}|$.

As can be seen in Table 4, the metrics not using absolute value ($\delta_{S}$, $\Delta_{S}$, $\bar{\Delta}_{S}$ and $\bar{|\Delta|}$) have mean values centered at zero and standard deviations lower in the group of healthy individuals than in the group of glaucoma patients. According to the definitions of these asymmetry metrics $\delta_{S}$, $\Delta_{S}$, $\bar{\Delta}_{S}$ and $\bar{|\Delta|}$, the statistical distributions of such asymmetry values will have be normal when the number of patients is high enough, due to the central limit theorem. The means of these normal distributions will always be approximately zero, considering that the direction in which the subtraction is calculated has been arbitrarily chosen, and the asymmetry must occur equiprobably in both directions (from the left eye to the right, or vice versa). Then, since the values follow a normal distribution with a zero mean, the variance of the distributions will be the distinguishing factor in the screening of healthy and glaucoma patients.

On the other hand, metrics based on the absolute value ($|\delta_{S}|$, $|\Delta_{S}|$, $|\bar{\Delta}_{S}|$ and $|\bar{|\Delta|}|$) have mean values greater than zero and standard deviations for the glaucoma subset slightly higher than for the subset of healthy patients. The absolute values of the asymmetry metrics cause normal distributions to become half-normal distributions [31]. The mean of the half-normal distribution is

$$\mu = \sigma \sqrt{\frac{2}{\pi}}$$

where $\sigma$ is the standard deviation of the initial normal distribution. That is, the mean of the new half-normal distribution is determined, proportionally, by the variance of the initial normal distribution, allowing the use of thresholds for the discrimination of healthy and glaucoma patients. Note that a greater variance will imply a higher mean value after taking the absolute value of the asymmetry for a specific sector.

The third characteristic gathered in Table 4 is the $p$-value, which determines the statistical significance of the hypothesis that the asymmetry of the RNFL thickness is statistically related to the occurrence of glaucoma disease. Then, a low $p$-value indicates that the asymmetry metric in a given sector has sufficient statistical significance in the occurrence of glaucoma between the healthy and glaucoma subset of patients. Taking
this into account, since the $p$-values of the metrics $|\Delta_S|$ and $|\bar{\Delta}_S|$ are the smallest of all the proposed metrics, $|\Delta_S|$ and $|\bar{\Delta}_S|$ are the most useful and relevant as discriminant variables in patient classification.

Table 4. Statistical characterization of the asymmetry metrics proposed for RNFL thickness. Mean value, standard deviation and $p$-value of each metric for the two subsets of healthy and glaucoma patients.

| Sector | Healthy | | | | Glaucoma | | | | | | $p$-Value | Healthy | | | | Glaucoma | | | | $p$-Value |
|--------|--------|--------|-----|--------|--------|--------|--------|-----|--------|--------|--------|--------|--------|-----|--------|--------|--------|-----|--------|--------|--------|
| TS     | 2.7563 ± 16.8122 | | | | −1.3830 ± 45.3539 | | | | 5.0247 × 10^{-1} | | | | 12.7438 ± 11.2637 | | | | 34.1489 ± 29.4522 | | | | 6.1489 × 10^{-6} |
| T      | 2.1500 ± 9.6793 | | | | −4.1064 ± 20.1022 | | | | 3.3495 × 10^{-2} | | | | 7.0125 ± 6.9896 | | | | 15.2979 ± 13.4987 | | | | 2.6754 × 10^{-4} |
| TI     | −2.3375 ± 14.9220 | | | | −4.2340 ± 51.8594 | | | | 7.0975 × 10^{-1} | | | | 11.7500 ± 9.4463 | | | | 40.0213 ± 32.7291 | | | | 1.0413 × 10^{-6} |
| NS     | −7.5938 ± 16.878 | | | | −10.8085 ± 34.9370 | | | | 8.4612 × 10^{-1} | | | | 14.3938 ± 11.3149 | | | | 27.8298 ± 23.4227 | | | | 4.0740 × 10^{-3} |
| N      | 4.2875 ± 12.5123 | | | | −0.2979 ± 26.0576 | | | | 2.9617 × 10^{-1} | | | | 10.0000 ± 8.6271 | | | | 17.3617 ± 19.2641 | | | | 2.2671 × 10^{-2} |
| NI     | −1.0625 ± 16.5372 | | | | −1.1489 ± 34.5989 | | | | 3.6061 × 10^{-1} | | | | 12.8125 ± 10.4606 | | | | 26.7660 ± 21.5973 | | | | 5.7953 × 10^{-5} |
| G      | 0.6563 ± 6.5895 | | | | −2.3404 ± 25.8728 | | | | 5.5860 × 10^{-1} | | | | 4.2188 ± 5.0937 | | | | 20.6383 ± 15.4855 | | | | 4.5656 × 10^{-9} |

| Sector | Healthy | | | | −0.0040 ± 0.2357 | | | | 6.4838 × 10^{-1} | | | | 0.0470 ± 0.0413 | | | | 0.1824 ± 0.1469 | | | | 6.7965 × 10^{-8} |
| T      | 0.0142 ± 0.0647 | | | | −0.0361 ± 0.2235 | | | | 1.0267 × 10^{-1} | | | | 0.0481 ± 0.0453 | | | | 0.1465 ± 0.1714 | | | | 3.9634 × 10^{-4} |
| TI     | −0.0075 ± 0.0531 | | | | −0.0170 ± 0.3159 | | | | 7.7715 × 10^{-1} | | | | 0.0416 ± 0.0338 | | | | 0.2396 ± 0.2036 | | | | 4.9500 × 10^{-8} |
| NS     | −0.0359 ± 0.0798 | | | | −0.0494 ± 0.2155 | | | | 8.8592 × 10^{-1} | | | | 0.0666 ± 0.0566 | | | | 0.1750 ± 0.1329 | | | | 2.9386 × 10^{-5} |
| N      | 0.0300 ± 0.0876 | | | | 0.0069 ± 0.1854 | | | | 4.8085 × 10^{-1} | | | | 0.0674 ± 0.0634 | | | | 0.1386 ± 0.1217 | | | | 4.8106 × 10^{-4} |
| NI     | −0.0052 ± 0.0772 | | | | 0.0058 ± 0.2342 | | | | 4.9736 × 10^{-1} | | | | 0.0582 ± 0.0508 | | | | 0.1779 ± 0.1502 | | | | 1.8119 × 10^{-6} |
| G      | 0.0033 ± 0.0346 | | | | −0.0162 ± 0.1804 | | | | 5.4845 × 10^{-1} | | | | 0.0214 ± 0.0272 | | | | 0.1428 ± 0.1094 | | | | 9.2247 × 10^{-10} |

With the aim of easing the understanding of the distribution of the values provides by each asymmetry metric, Figure 4 shows notched box plots of the proposed RNFL thickness asymmetry metrics for the healthy (h) and glaucoma (g) subsets of patients, depicted for each eye sector (TS, T, TI, NS, N and NI) and for the overall global thickness value (G). For each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. These percentiles delimit the so-called interquartile range. The whiskers extend to the most extreme values not considered outliers.
Outliers are values located more than 1.5 times the interquartile range outside the upper or lower boundary of the box. The tapered and shaded regions, called notches, display the variability of the median between values. The width of a notch is computed so that boxes whose notches do not overlap have different medians at the 5% significance level—i.e., the true medians do differ with 95% confidence.

Figure 4. Cont.
Figure 4. Notched box plots of the proposed RNFL thickness asymmetry metrics for the healthy (h) and glaucoma (g) subsets of patients, depicted for each eye sector (TS, T, TI, NS, N and NI) and for the overall global thickness value (G).

As can be seen in Figure 4, these graphs corroborate the insights concluded above. The values of metrics $\delta_S$, $\Delta_S$, $\bar{\Delta}_S$ and $\bar{\bar{\Delta}}_S$ are distributed around zero, and there is greater dispersion for the subset of glaucoma patients. The box plots of $|\delta_S|$, $|\Delta_S|$, $|\bar{\Delta}_S|$ and $|\bar{\bar{\Delta}}_S|$ depict how the variances of the distributions not using absolute values have been transformed into the corresponding means of the related metrics with absolute values, according to Equation (11). In addition, it can also be seen qualitatively how the notches with the least overlap for healthy and glaucoma patients are those corresponding to metrics $|\Delta_S|$ and $|\bar{\Delta}_S|$, especially in sectors TS and TI, although the distribution of asymmetry values considering the entire layer, G, has the greater separation. As will be proven numerically later, this fact graphically justifies the use of these metrics as the characteristics for the design of the classifier.

3.2. Decision Trees

Artificial intelligence (AI) is increasingly being incorporated into the diagnostic process in healthcare due to its ability to analyze data with complex artificial networks and to learn automatically, especially through machine learning (ML) and deep learning (DL) [32,33]. In this work, instead of using complex classification methods based on machine learning to diagnose glaucoma (see, e.g., [34–36]), we chose classification trees [37] for their simplicity, with the aim of transferring them to daily clinical practice in glaucoma screening.

A decision tree is a supervised machine learning algorithm with a tree-like structure which repeatedly splits the input dataset into classes, taking into account one exploratory variable at a time. These trees are used when the target variable is categorical and can assume only one of two mutually exclusive values (healthy and glaucoma, in our case) [38,39].

Considering the RNFL thickness in all sectors (TS, T, TI, NS, N and NI) and the global mean (G) of the peripapillary OCT, the proposed asymmetry metrics have been applied to generate a vector with the features of the subsets of healthy and glaucoma patients. With these features as input variables, a fitted binary classification decision tree has been designed for each asymmetry metric taking into account the labeling of each class (healthy or glaucoma patient).

The designed binary trees split branching nodes based on the values of the input features. One of the parameters of the model is the depth of the tree, which is related to the model complexity, and therefore, to the computational cost. The classification trees have been developed using five folds in the cross-validation. The results of the classification trees have been analyzed controlling the maximum depth of the trees (or maximum number of splits), $\zeta$. Initially, the weights of the inputs corresponding to the class healthy and to the class glaucoma, $w_h$ and $w_g$, respectively, were set to unity, i.e., $w_h = 1.0$ and $w_g = 1.0$. 
Table 5 gathers the classification loss for observations not used for training in deep decision trees ($\xi = 15$) and shallower trees ($\xi = 3$). The largest value of $\xi$ for deep decision trees depends on the number of input features used, which in our case, were the seven asymmetry measurements performed in the six sectors and in the global average of the RNFL thickness. As can be seen, the classification loss decreases as the maximum number of splits is reduced, providing the best values for the metrics $|\Delta|$, $\bar{\Delta}$ and $|\bar{\Delta}|$. The best values have been highlighted in bold in Table 5, and their corresponding decision trees and confusion matrices are shown in Figure 5. The representation of each model is a binary tree where each root node represents a single input variable (asymmetry feature) and a split point on that variable. The leaf nodes of the tree contain an output variable which is used to predict if the input corresponds to a healthy or a glaucoma patient. Regarding the confusion matrices, they provide four outcomes:

- True positive (TP): glaucoma patient predicted as glaucoma (top left element),
- False positive (FP): healthy patient predicted as glaucoma (bottom left element),
- True negative (TN): healthy patient predicted as healthy (bottom right element),
- False negative (FN): glaucoma patient predicted as healthy (top right element).

In order to assess the performances of these three trees, the following parameters have been computed from the confusion matrix:

- **Accuracy**: 
  \[ ACC = \frac{TP + TN}{TP + TN + FP + FN}. \]  
  \[ (12) \]
- **Sensitivity (recall or true positive rate, TPR)**: 
  \[ TPR = \frac{TP}{TP + FN}. \]  
  \[ (13) \]
- **Specificity (true negative rate, TNR)**: 
  \[ TNR = \frac{TN}{TN + FP}. \]  
  \[ (14) \]
- **Precision (positive predictive value, PPV)**:  
  \[ PPV = \frac{TP}{TP + FP}. \]  
  \[ (15) \]

Table 5. Classification loss for observations not used for training ($w_h = 1$, $w_g = 1$).

| $\xi$ | $\gamma$ | $|\gamma|$ | $\Delta$ | $|\Delta|$ | $\bar{\Delta}$ | $|\bar{\Delta}|$ | $\bar{\Delta}$ | $|\bar{\Delta}|$ |
|-------|--------|---------|--------|--------|--------|--------|--------|--------|
| 15    | 0.2174 | 0.2222  | 0.1643 | 0.1498 | 0.1643 | 0.1932 | 0.1256 | 0.1546 |
| 10    | 0.1932 | 0.1932  | 0.1594 | 0.1498 | 0.1691 | 0.1739 | 0.1256 | 0.1498 |
| 5     | 0.1787 | 0.1643  | 0.1498 | 0.1546 | 0.1546 | 0.1449 | 0.1159 | 0.1449 |
| 3     | 0.1643 | 0.1401  | 0.1643 | 0.1401 | 0.1401 | 0.1304 | 0.1304 | 0.1208 |

As can be seen in Table 6, these decision trees provide high accuracy (all of them higher than 87%), high specificity (higher than 95%) and high precision (higher than 79%). In order to improve the results, i.e., increase the sensitivity while keeping high values for the remaining parameters, the weight of the inputs corresponding to the class glaucoma has been increased from 1.0 to 1.5. Table 7 contains the classification loss values of the new models considering this weight in the glaucoma input observations.
Figure 5. Decision trees and confusion matrices of the best results of Table 5. (a) Decision tree (left) and confusion matrix (right) for the metric $\bar{\Delta}$, using $w_h = 1$, $w_g = 1$ and maximum number of splits $\xi = 5$. (b) Decision tree (left) and confusion matrix (right) for the metric $|\Delta|$, using $w_h = 1$, $w_g = 1$ and maximum number of splits $\xi = 3$. (c) Decision tree (left) and confusion matrix (right) for the metric $|\bar{\Delta}|$, using $w_h = 1$, $w_g = 1$ and maximum number of splits $\xi = 3$.

Table 6. Performances of the best classification trees ($w_h = 1$, $w_g = 1$).

| Tree Parameters | Accuracy | Sensitivity | Specificity | Precision |
|-----------------|----------|-------------|-------------|-----------|
| $\bar{\Delta}, \xi = 5$ | 0.8840 | 0.6595 | 0.9500 | 0.7948 |
| $|\Delta|, \xi = 3$ | 0.8792 | 0.6170 | 0.9562 | 0.8055 |
| $|\bar{\Delta}|, \xi = 3$ | 0.8792 | 0.5957 | 0.9625 | 0.8235 |
Table 7. Classification loss for observations not used for training ($w_h = 1, w_g = 1.5$).

| $\zeta$ | $\delta$ | $|\delta|$ | $\Delta$ | $|\Delta|$ | $\bar{\Delta}$ | $|\bar{\Delta}|$ |
|---------|----------|-------------|----------|-------------|--------------|-------------|
| 15      | 0.2386   | 0.2516      | 0.1692   | 0.1844      | 0.2082       | 0.2234      |
| 10      | 0.2343   | 0.2408      | 0.1692   | 0.1844      | 0.2082       | 0.2148      |
| 5       | 0.2169   | 0.2082      | 0.1996   | 0.1931      | 0.2104       | 0.1584      |
| 3       | 0.1822   | 0.1497      | 0.1670   | 0.1432      | 0.1996       | 0.1627      |

Based on the results of classification loss gathered in Table 7, the model which offers the best performance with the minimum number of splits is $|\Delta|$ with $\zeta = 3$. Figure 6 shows the resulting model and its corresponding confusion matrix. This classification tree provides the confusion matrix whose parameters are collected in Table 8.

Figure 6. Decision tree and confusion matrix for metric $|\Delta|$, using $w_h = 1, w_g = 1.5$ and $\zeta = 3$.

Table 8. Performance of the best classification tree ($w_h = 1, w_g = 1.5$).

| Tree Parameters | Accuracy | Sensitivity | Specificity | Precision |
|-----------------|----------|-------------|-------------|-----------|
| $|\Delta|, \zeta = 3$ | 0.8792   | 0.6595      | 0.9437      | 0.7750    |

Finally, in order to further increase the number of true glaucoma patients detected, i.e., the sensitivity, and considering that the asymmetry metric $|\Delta|$ offers the lowest classification loss and taking into account that the asymmetry values range from 0 to 1 (since this measure is normalized by the sum of the RNFL thickness of each eye in the corresponding sector), the asymmetry metric $|\Delta|^{0.5}$ is proposed. The nonlinear operation of applying the square root to the non-negative real values provided by the metric $|\Delta|$ decompresses the distribution of the values corresponding to healthy and glaucoma patients. This process of encoding with this decompressive power-law nonlinearity is called gamma compression [40]. As can be seen in Table 9, which gathers the statistical parameters of the metric $|\Delta|^{0.5}$, the mean values for healthy and glaucoma patients have increased in all sectors. In addition, the distances between the mean values for the subsets of healthy and glaucoma patients have further increased, compared to the values provided by $|\Delta|$. This effect can be also observed in Figure 7, where the notched box plots of each subset are further apart. The resulting decision tree using $|\Delta|^{0.5}$ and the corresponding confusion matrixes are shown in Figure 8. The parameters of accuracy, sensitivity, specificity and precision are gathered in Table 10.
Table 9. Statistical characterization of the asymmetry metric $|\Delta_S|^{0.5}$ proposed for RNFL thickness. Mean values, standard deviations and $p$-values of this metric for the two subsets of healthy and glaucoma patients.

| Sector | Healthy | Glaucoma | $p$-Value |
|--------|---------|----------|-----------|
| TS     | 0.1960 ± 0.0931 | 0.3908 ± 0.1741 | 4.2798 × 10^{-9} |
| T      | 0.1989 ± 0.0929 | 0.3358 ± 0.1856 | 5.9084 × 10^{-5} |
| TI     | 0.1870 ± 0.0813 | 0.4393 ± 0.2183 | 5.6436 × 10^{-9} |
| NS     | 0.2373 ± 0.1019 | 0.3813 ± 0.1739 | 5.4446 × 10^{-5} |
| N      | 0.2332 ± 0.1145 | 0.3359 ± 0.1623 | 3.5263 × 10^{-4} |
| NI     | 0.2190 ± 0.1016 | 0.3842 ± 0.1759 | 3.5045 × 10^{-7} |
| G      | 0.1254 ± 0.0758 | 0.3469 ± 0.1515 | 2.4453 × 10^{-12} |

Figure 7. Notched box plots of the proposed RNFL thickness asymmetry metric $|\Delta_S|^{0.5}$ for the healthy (h) and glaucoma (g) subsets of patients, depicted for each eye sector (TS, T, TI, NS, N and NI) and for the overall global thickness value (G).

Figure 8. Decision tree and confusion matrix for metric $|\Delta|^{0.5}$, using $w_h = 1$, $w_g = 1.5$ and $\xi = 3$.

Table 10. Performance of the best classification tree ($w_h = 1$, $w_g = 1.5$).

| Tree Parameters | Accuracy | Sensitivity | Specificity | Precision |
|-----------------|----------|-------------|-------------|-----------|
| $|\Delta|^{0.5}$, $\xi = 3$ | 0.8792    | 0.7021      | 0.9312     | 0.7500    |

4. Discussion

This article has delved into the hypothesis that the asymmetry between the two eyes can be used as an indicator of glaucoma. For this purpose, new asymmetry metrics have been defined for the thickness of the RNFL obtained from OCTs of both eyes of the same patient. In this work, the mean thickness of each sector of the RNFL was provided by
the segmentation of Spectralis; however, segmentation methods such as those described in [30,41] will be used in the future to address the relationship of the RNFL thickness at each angle of the B-scan in the diagnosis of glaucoma.

Through a statistical analysis of the RNFL thickness values, the asymmetry metrics that provide the greatest distinction between the values of the subset of healthy and glaucoma patients have been chosen. Using these metrics, supervised machine learning models have been designed—decision trees. The resulting models are classifiers that partitioning the feature space of the asymmetry values in the sectors of the RNFL and building binary decision trees recursively can create meaningful outputs that predict being healthy or suffering from glaucoma.

The proposed decision trees are simple but effective models which provide a binary classification decision (healthy or glaucoma) by splitting branching nodes based on the input asymmetry features. Their simplicity or reduced complexity is one of their main advantages. They allow for rapid classification of new observations, since it is much simpler to evaluate just one or two logical conditions than to compute scores using complex nonlinear equations for each group. For example, in [42] complex machine learning classifiers (conditional inference trees, logistic model tree, C5.0 decision tree, random forest and extreme gradient boosting XGBoost) use OCT parameters to diagnose glaucoma, achieving slightly better performances (average accuracy 0.8818, average sensitivity 0.9166, average specificity 0.8507 and average area under the curve 0.9459) than decision trees. However, these models can be seen as black boxes which produce results based solely on the input data using an algorithm, which prevents clinicians from understanding how variables are being combined to make such predictions. On the contrary, decision trees proposed in this paper allow one to study and analyze the importance of the input asymmetry features, and ophthalmologists may obtain clinical insight from the explanations. In addition, the visual diagram of a decision tree provides a simple explanation for the reason for classifying a patient as healthy or glaucomic. In daily clinical practice it is much easier to explain if–then statements than complex nonlinear equations.

Another advantage of decision trees is their implicit feature selection, since the top few nodes on which the tree is split are the most important variables within the set. In all the trees designed, the characteristic considered in the root is the asymmetry in the global (G) thickness of the RNFL, followed by the asymmetry in the TI sector.

The decision trees have some limitations, such as their tendency to overfit when very complex, in which case they generally have low bias. In this work, this drawback has been overcome by restricting the maximum number of splits. The best trees proposed for the diagnosis of glaucoma have, at most, up to three divisions without degrading the screening performance. Experimentally, it has been found that three partitions suppose an adequate trade-off between simplicity and performance, allowing, if necessary, for the possibility of being applied by hand, without the support of a computer.

Another limitation of decision trees is that they can be unstable because small variations in the input data can generate a completely different tree. This situation is also solved by limiting the maximum number of splits, which improves the ability of the model to generalize to new input data. In addition, as future work, the size of the dataset will be increased (both healthy and patients with glaucoma, keeping the balance between classes) allowing us to check the performance of the designed trees.

One more future line of work is the analysis of the performances of complex machine learning models made to estimate the upper bound of the classification results that can be obtained using the proposed database and to study the differences between the results of the proposed decision trees and these complex models, even though they cannot be easily implemented by ophthalmologists in daily clinical practice.

In conclusion, decision tree models are easy to understand and implement (even by hand if necessary), which gives them a strong advantage when compared to other analytical models. The results obtained in this paper confirm that the proposed machine learning classifier, the decision tree, offers satisfactory performance and a high capacity
References

1. Quigley, H. Number of people with glaucoma worldwide. Br. J. Ophthalmol. 1996, 80, 389–393. [CrossRef] [PubMed]
2. Tham, Y.C.; Li, X.; Wong, T.Y.; Quigley, H.A.; Aung, T.; Cheng, C.Y. Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040: A Systematic Review and Meta-Analysis. Ophthalmology 2014, 121, 2081–2090. [CrossRef] [PubMed]
3. Casson, R.J.; Chidlow, G.; Wood, J.P.; Crowston, J.G.; Goldberg, I. Definition of glaucoma: Clinical and experimental concepts. Clin. Exp. Ophthalmol. 2012, 40, 341–349. [CrossRef]
4. Weinreb, R.N.; Khaw, P.T. Primary open-angle glaucoma. Lancet 2004, 363, 1711–1720. [CrossRef]
5. Collaborative Normal-Tension Glaucoma Study Group. Natural history of normal-tension glaucoma. Ophthalmology 2001, 108, 247–253. [CrossRef]
6. Fechtner, R.D.; Weinreb, R.N. Mechanisms of optic nerve damage in primary open angle glaucoma. Surv. Ophthalmol. 1994, 39, 23–42. [CrossRef]
7. Kwon, Y.H.; Fingert, J.H.; Kuehn, M.H.; Alward, W.L. Primary Open-Angle Glaucoma. N. Engl. J. Med. 2009, 360, 1113–1124. [CrossRef] [PubMed]
8. Abramoff, M.D.; Garvin, M.K.; Sonka, M. Retinal Imaging and Image Analysis. IEEE Rev. Biomed. Eng. 2010, 3, 169–208. [CrossRef]
9. Swanson, E.A.; Fujimoto, J.G. The ecosystem that powered the translation of OCT from fundamental research to clinical and commercial impact. Biomed. Opt. Express 2017, 8, 1638–1664. [CrossRef]
10. Huang, D.; Swanson, E.; Lin, C.; Schuman, J.; Stinson, W.; Chang, W.; Hee, M.; Flotte, T.; Gregory, K.; Puliafito, C.; et al. Optical coherence tomography. Science 1991, 254, 1178–1181. [CrossRef]
11. Kriengsri, J.; Roth, B. Wavefront Shaping Concepts for Application in Optical Coherence Tomography—A Review. Sensors 2020, 20, 7044. [CrossRef] [PubMed]
12. Bussel, I.I.; Wollstein, G.; Schuman, J.S. OCT for glaucoma diagnosis, screening and detection of glaucoma progression. Br. J. Ophthalmol. 2014, 98, ii15–ii19. [CrossRef] [PubMed]
13. Badalà, F.; Nouri-Mahdavi, K.; Raoof, D.A.; Leeprechanon, N.; Law, S.K.; Caprioli, J. Optic disk and nerve fiber layer imaging to detect glaucoma. Am. J. Ophthalmol. 2007, 144, 724–732. [CrossRef]
14. Alasli, T.; Wang, K.; Keane, P.A.; Lee, H.; Baniasadi, N.; de Boer, J.F.; Chen, T.C. Analysis of normal retinal nerve fiber layer thickness by age, sex, and race using spectral domain optical coherence tomography. J. Glaucoma 2013, 22, 532–541. [CrossRef]
15. Capece, M.; Montorio, D.; Comune, C.; Aveta, A.; Melchionna, A.; Celentano, G.; Imbimbo, C.; Crocetto, F.; Calìfano, G.; Cennamo, G. Retinal and Optic Disc Vascular Changes in Patients Using Long-Term Tadalafil: A Prospective Non-Randomized Matched-Pair Study. Diagnostics 2021, 11, 802. [CrossRef] [PubMed]
16. Sommer, A.; Katz, J.; Quigley, H.A.; Miller, N.R.; Robin, A.L.; Richter, R.C.; Witt, K.A. Clinically Detectable Nerve Fiber Atrophy Precedes the Onset of Glaucomatous Field Loss. Arch. Ophthalmol. 1991, 109, 77–83. [CrossRef]
17. Williams, A.L.; Gatia, S.; Leiby, B.E.; Fahmy, I.; Biswas, A.; De Barros, D.M.; Ramakrishnan, R.; Bhardwaj, S.; Wright, C.; Dubey, S.; et al. The value of intraocular pressure asymmetry in diagnosing glaucoma. J. Glaucoma 2013, 22, 215–218. [CrossRef]
18. Lester, M.; Telani, S.; Frezzotti, P.; Manni, G.; Uva, M.; Figus, M.; Perdicchi, A.; Brusini, P.; Paoli, D.; Rolando, M.; et al. Differences in central corneal thickness between the paired eyes and the severity of the glaucomatous damage. Eye 2012, 26, 1424–1430. [CrossRef]
