Journal of
Clinical Medicine

MDPI

Article

Fully Automated Colorimetric Analysis of the Optic Nerve Aided by Deep Learning and Its Association with Perimetry and OCT for the Study of Glaucoma

Marta Gonzalez-Hernandez [1,2], Daniel Gonzalez-Hernandez [1], Daniel Perez-Barbudo [1], Paloma Rodriguez-Esteve [3], Nisamar Betancor-Caro [3] and Manuel Gonzalez de la Rosa [1,3,*]

1 INSOFT S.L., 25 de Julio, 34, 38004 Santa Cruz de Tenerife, Spain; martaglezhdez@gmail.com (M.G.-H.); management@insoft.es (D.G.-H.); danieltf@gmail.com (D.P.-B.)
2 Ophthalmology Department, Hospital Universitario de Canarias, Carretera Ofra s/n, 38320 San Cristobal de La Laguna, Spain
3 Facultad de Ciencias de la Salud, Universidad de La Laguna, C/Sta, María Soledad s/n, 38200 San Cristobal de La Laguna, Spain; paloma.rguezesteve@gmail.com (P.R.-E.); nisamar89@hotmail.com (N.B.-C.)
* Correspondence: mgdelarosa1950@gmail.com

Citation: Gonzalez-Hernandez, M.; Gonzalez-Hernandez, D.; Perez-Barbudo, D.; Rodriguez-Esteve, P.; Betancor-Caro, N.; Gonzalez de la Rosa, M. Fully Automated Colorimetric Analysis of the Optic Nerve Aided by Deep Learning and Its Association with Perimetry and OCT for the Study of Glaucoma. *J. Clin. Med.* **2021**, *10*, 3231. https://doi.org/10.3390/jcm10153231

Academic Editors: Jose Javier Garcia-Medina and Maria Dolores Pinazo-Duran

Received: 6 July 2021
Accepted: 20 July 2021
Published: 22 July 2021

Abstract: Background: Laguna-ONhE is an application for the colorimetric analysis of optic nerve images, which topographically assesses the cup and the presence of haemoglobin. Its latest version has been fully automated with five deep learning models. In this paper, perimetry in combination with Laguna-ONhE or Cirrus-OCT was evaluated. Methods: The morphology and perfusion estimated by Laguna ONhE were compiled into a "Globin Distribution Function" (GDF). Visual field irregularity was measured with the usual pattern standard deviation (PSD) and the threshold coefficient of variation (TCV), which analyses its harmony without taking into account age-corrected values. In total, 477 normal eyes, 235 confirmed, and 98 suspected glaucoma cases were examined with Cirrus-OCT and different fundus cameras and perimeters. Results: The best Receiver Operating Characteristic (ROC) analysis results for confirmed and suspected glaucoma were obtained with the combination of GDF and TCV (AUC: 0.995 and 0.935, respectively. Sensitivities: 94.5% and 45.9%, respectively, for 99% specificity). The best combination of OCT and perimetry was obtained with the vertical cup/disc ratio and PSD (AUC: 0.988 and 0.847, respectively. Sensitivities: 84.7% and 18.4%, respectively, for 99% specificity). Conclusion: Using Laguna ONhE, morphology, perfusion, and function can be mutually enhanced with the methods described for the purpose of glaucoma assessment, providing early sensitivity.

Keywords: glaucoma; deep learning; perimetry; optic nerve

1. Introduction

Glaucoma is a disease with a relatively high incidence, estimated between 2.09% and 5.82% of the adult population [1]. Its diagnosis in early stages is difficult due to its almost asymptomatic onset and controversy of criteria regarding its initial signs.

Intraocular pressure is an important pathogenesis factor; while not the only one, it is currently the only one that can be targeted for treatment. Many other factors seem to have an influence on the disease: genetic, morphological, vasospasm, intracranial pressure, tissue thickness and morphology, sleep apnoea, etc.

For many years, visual field assessment and optic nerve head examination were the main diagnostic procedures. More recently, morphological tests have been proposed, and many authors have argued their greater precocity [2], although, as shown in this paper, this opinion is debatable. Examples are the scanning laser polarimeter (GDx), the Heidelberg retina tomograph (HRT) and the optical coherence tomograph (OCT), which is

currently widely used. More recently, OCT angiographies have been introduced to analyse vascularization [3], especially in the retinal areas surrounding the optic nerve.

We developed a simple colorimetric procedure to estimate the presence of haemoglobin and its distribution on the optic disc, which reflects its perfusion and morphology, using conventional colour retinographies (Laguna ONhE) [4]. The method has been shown to have good sensitivity and specificity in previous studies [5–12]. However, deep learning methods for the segmentation of optic nerve edges were later incorporated to facilitate its use and improve its reproducibility [13].

Artificial intelligence, especially convolutional networks, allow the experience of experts to be transferred with great precision to other possible users in different medical specialties, most notably in ophthalmology [14]. Transfer Learning's new procedures simplify the development work, taking advantage of the previous training of networks on generic problems [15].

In a previous experiment [13], we used a deep learning U-Net architecture for semantic segmentation [16]. In this case, 40,000 images were used to segment the optic disc, identifying the inner edge of Elschnig's scleral ring.

The importance of perfusion in glaucoma is well recognised [3]. A recent paper using this same version of the algorithm showed that its results are in the range of an OCT angiography [17]. The aim of this paper was to take advantage of these new automatic procedures based on experience to facilitate the use of the application and to improve its reproducibility, sensitivity, and specificity. In particular, an attempt was made to combine the morphological and perfusion information provided by the method with functional perimetric data representative of visual field homogeneity [18] to assess whether this provided robustness in the diagnostic decision.

2. Materials and Methods

The study protocol of this cross-sectional retrospective study adhered to the principles of the 1975 Declaration of Helsinki revised in 2013, and was approved by the Research Ethics Committee of the Hospital Universitario de Canarias (CHUC_2018_09 (V3)). Consent was obtained from all participants.

2.1. Automation of the Laguna ONhE Method for Estimation of Haemoglobin in the Optic Nerve Head

The presence of haemoglobin in the tissue was estimated by Laguna ONhE with reference to the colour of the vessels. Haemoglobin mainly absorbs green radiation and reflects most of the red. Therefore, the reference colour of the vessels was calculated using the values of the red (R) and green (G) channels of their pixels, to which the formula (R−G)/R is applied. The same equation was used for the pixels of the tissue, and finally, the result was expressed as a percentage. An estimate of cup size and position was also obtained, and the results of the cup, rim sectors, vertical cup/disc ratio (Hb-C/D) and cup/disc area ratio were compared with the percentiles achieved in the normal population [10]. Each fundus camera model was calibrated to achieve an equivalent response.

To achieve full automation of the method, five neural networks were used: one to segment the edges of the optic nerve already described [13], one for vessel segmentation using 4195 optic disc images, one to identify the eye as left or right using 4201 images, one to recognize image quality using 7048 images, and one to identify between normality (using 1518 images) and confirmed or suspected glaucoma (using 1596 images). The technical method is described in detail in the Appendix A "Computing development setup".

The classification results obtained by deep learning were associated with the distribution of haemoglobin and the estimated Cup/Disc ratios to define a new value for the "Globin Distribution Function" (GDF) index, as previously described [4,5]. Once the value of the deep learning classifier was normalized to the mean values and standard deviation in the normal population, it influenced the result of the current GDF index by 45%. In the remaining 55%, the rest of the usual variables that we used in previous studies also intervened with normalized values. An example of the graphic results is shown in Figure 1.

It shows how the method automatically performs the analysis of a retinography (a). Once its quality has been checked, the position of the optic nerve is identified, the inner edge of Elschnig's scleral ring is defined, which is more internal than the apparent clear edge, and the size and shape of the cup is estimated (b). Its veins and arteries are then segmented, the colour of which is used to estimate the relative haemoglobin of the rest of the nerve tissue, shown in a colour code (c). In each sector of the optic disc, cup or rim, their areas are estimated as a percentage of the total disk area and expressed in colour if it corresponds to what is expected in a normal optic nerve (d).

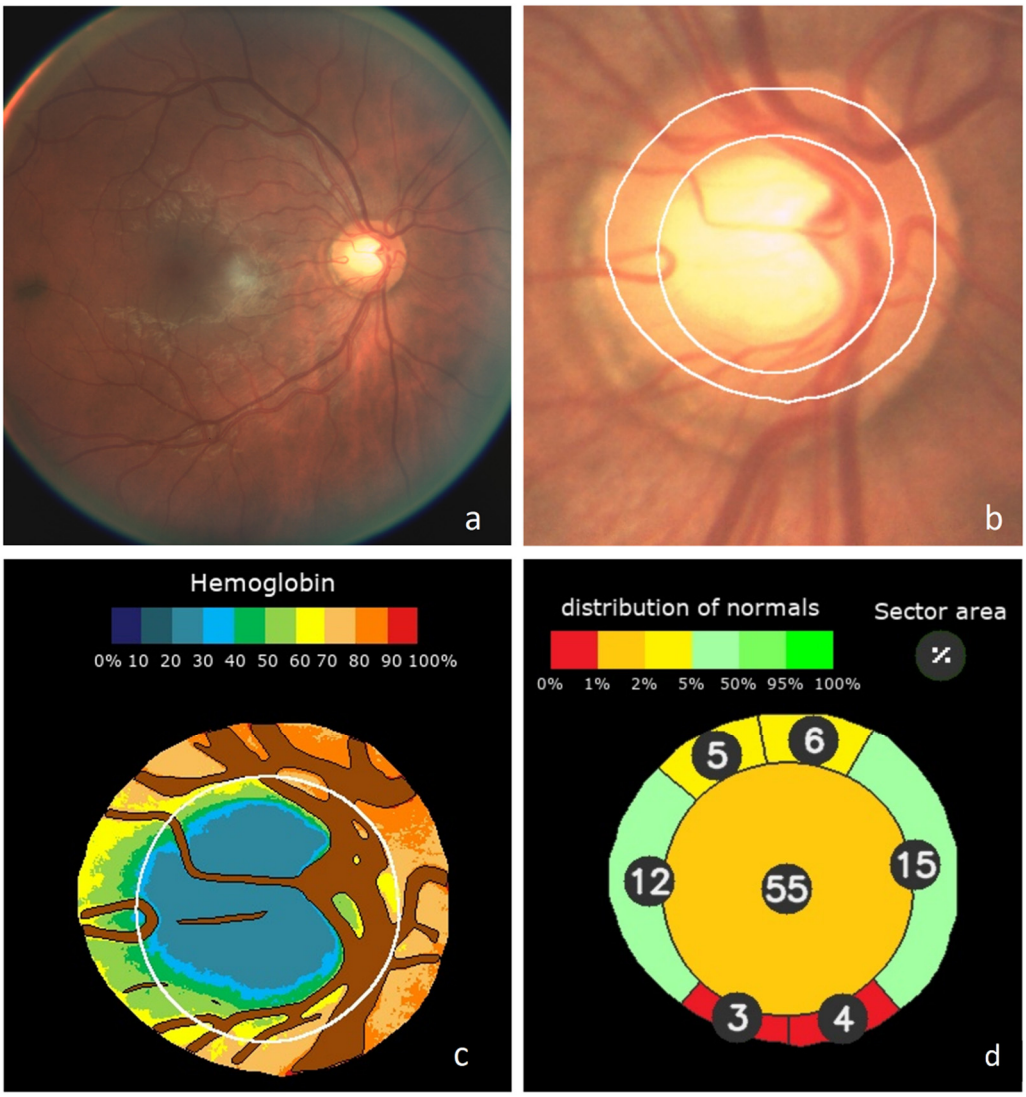


Figure 1. Example of Laguna ONhE analysis: (**a**) Original wide-field eye fundus image. (**b**) Identification of the optic disc and boundaries segmentation. Central cup estimated from the haemoglobin distribution. (**c**) Segmentation of reference vessels and pseudo-colour image of haemoglobin distribution. (**d**) Estimated sector areas as a percentage of the total area, and compared to a normal reference population.

2.2. Combination of Laguna ONhE and Perimetric Indices

The pattern standard deviation (PSD) perimetric index and its equivalent in the Octopus perimeters, the square root of loss variance (sLV), are well known. In essence, the pointwise differences between the patient's sensitivity thresholds and the expected value in a subject of the same age were calculated, and the standard deviation of all values was obtained. In general, these values increase as the disease progresses, but in advanced

defects they reduce as the number of points with no detectable sensitivity (0 dB) increases. To achieve a linear response, in advanced glaucoma, a modification was made, as previously described [19], taking into account the mean defect/deviation (MD), using Equation (1).

Equation (1): sLV modification to provide a linear response.

$$\text{If } \text{abs}(\text{MD}) > 16.33, \text{ then } \text{sLV} = \text{sLV} + \frac{\text{abs}(\text{MD}) - 16.33}{0.84} \quad (1)$$

The threshold coefficient of variation (TCV) is an index of regularity described by our research group, which represents the harmony of one's own visual field, without comparing it with patterns of normality [18]. A characteristic for normality is stability or harmony, and for pathology, is irregularity or fluctuation. We aimed to analyse them in the subject themself, regardless of whether their thresholds are close or far from the normal average, in the same way that a subject can be harmonious and of short stature or harmonious and tall. This way, TCV is a less complex index than PSD, as it is independent of the age of the subject. In normal subjects and initial and moderate glaucoma, it is calculated by dividing the standard deviation of the thresholds in 16 symmetrical positions of the visual field by their mean value, and multiplying by 100. In advanced glaucoma, where absolute scotomas are present, both values are adjusted to their number.

2.3. Datasets for the Laguna ONhE, OCT and Perimetric Indices

Two groups were included, consisting of 477 healthy eyes from 409 subjects and 333 eyes with confirmed or suspected chronic open-angle glaucoma from 246 subjects. Healthy subjects consisted of hospital staff, patient relatives, or people who needed refraction but did not have eye abnormalities.

All cases had corrected visual acuity of 20/40 or higher, refractive error with spherical equivalent of less than ±5.00 dioptres, astigmatism less than ±2 dioptres, and open anterior chamber angle. Subjects with cataracts reaching the specified visual acuity were not excluded. Previous cataract or glaucoma surgery were also not exclusion criteria. Patients with associated ocular diseases which could interfere with the interpretation of the results, such as optic neuritis, coloboma, and papillary oedema, were not included. All the participants underwent a complete examination, including visual acuity, slit lamp examination, Goldmann tonometry, and fundus examination, within a maximum time interval of one week.

All cases were examined with the Cirrus spectral-domain optical coherence tomograph (OCT; Carl Zeiss Meditec, Jena, Germany), using the optic disc cube 200 × 200 acquisition protocol (software version 5.2). All images were acquired with a quality greater than 6/10.

Three fundus cameras, two perimeters, and two different perimetric strategies were used to verify that the results of the evaluated method did not depend on the instruments used. In total, 213 normal cases and 110 glaucomas were examined with a Nidek AFC-210 non-mydriatic fundus retinograph (Nidek Co., LTD, Aichi, Japan) and a white-on-white Spark strategy in an Easyfield perimeter (Oculus Optikgeräte GmbH, Wetzlar, Germany) [20]. Then, 87 normal cases and 70 glaucomas were examined with a Kowa Wx non-mydriatic fundus retinograph (Kowa Co. Ltd., Tokyo, Japan) and with the Easyfield perimeter using Spark strategy, and 177 normal and 153 glaucomas were examined with a Horus Scope DEC-200 handheld fundus camera (MiiS, Hsinchu, Taiwan) and an Octopus 300 perimeter (Hagg-Streit AG, Bern, Switzerland) using the Tendency Oriented Perimetry (TOP) strategy [21].

All control and glaucoma patients underwent perimetric assessment, having undergone at least two previous examinations. Healthy eyes had intraocular pressure less than 21 mmHg, and no abnormal results in the visual field or on the optic disc. The "glaucoma" group comprised glaucoma and glaucoma-suspected eyes. Not all patients in the "glaucoma" group had defects characteristic of the optic nerve or visual field. In some cases, there were only signs of suspicion, such as intraocular pressure greater than 21 mmHg, associated with a family history of glaucoma, an optic disc with a dubious appearance,

or borderline visual fields, such as a mean deviation exceeding −2 dB or points outside normal limits on the defect curve. Subjects with ocular pressures greater than 25 mmHg, or pressures between 21 and 25 mmHg accompanied by a thin cornea (less than 500 µm), were also included.

Two types of analysis were carried out. In the first, no strict boundary was established between confirmed and suspected glaucoma, so as not to introduce an a priori criterion that could alter an objective interpretation of the results [22]. This methodology was widely discussed in a previous paper [23]. Additionally, two groups were analysed separately (confirmed glaucoma and glaucoma suspects): confirmed glaucoma was defined based on the presence of glaucomatous visual field loss in standard automated perimetry (pattern standard deviation or mean deviation of <5%) and signs of glaucomatous neuropathy (Laguna ONhE GDF or OCT-rim area of <5%). Those who did not meet these criteria, but met some of the criteria described in the previous paragraph, were considered as glaucoma suspects.

2.4. Statistical Analysis

Clinical statistical analyses were performed using the Excel 2016 program (Excel, Microsoft Corp., Redmond, WA, USA) and MedCalc (Version 18.9–64 bits; MedCalc software bvba, Mariakerke, Belgium). The statistical comparison between the results of the different AUCs were calculated in MedCalc using the criteria described by DeLonng et al. [24].

For statistical purposes, the sign of MD values of the Octopus Perimeter was inverted, to use the same criterion in both perimeters.

In order to associate different indices, they were previously normalized in relation to the mean value and standard deviation of all normal cases.

3. Results

The reference group was composed of 167 male eyes and 310 female eyes, aged 44.56 ± 13.23 years, with a perimetric MD of 0.18 ± 1.56 dB and PSD-sLV of 1.48 ± 0.52 dB. The rim area of the Cirrus-OCT was 1.42 ± 0.29 mm^2, the retinal nerve fibre layer thickness (RNFLT) was 91.67 ± 9.77 µm, and the vertical cup/disc ratio (OCT-C/D) was 0.45 ± 0.15.

The glaucoma group consisted of 176 male eyes and 157 female eyes, aged 63.63 ± 11.74 years, with a perimetric MD of −8.85 ± 8.45 dB and PSD-sLV of 4.43 ± 2.82 dB. The rim area was 0.77 ± 0.34 mm^2, the RNFLT was 68.75 ± 16.44 µm, and the OCT-C/D was 0.74 ± 0.14. Differences were statistically significant in all cases ($p < 0.0001$). Among them, 235 met the criteria for confirmed glaucoma, and 98 were considered to be suspected glaucoma. Glaucoma suspects had an MD of −0.92 ± 2.71 dB, and confirmed glaucoma had an MD of −12.16 ± 7.81 dB.

3.1. Results of the Indices of the Three Testing Methods on the Total Sample (without Separating Confirmed and Suspected Glaucoma)

The GDF index result in the groups of validation datasets is shown in Table 1. The Laguna ONhE GDF index obtained significantly higher results than all perimetric and Cirrus-OCT indices. The best OCT index was rim area and the best perimetric index TCV, with no significant differences between them.

Figure 2 shows the receiver operating characteristic (ROC) curves obtained for the Laguna ONhE GDF index, the OCT indices, and the perimetric indices. To improve the identification of differences, the upper left part of the curves is enlarged. GDF presented a sensitivity higher than the rim area for 95.0% ($p = 0.0121$) and 99.0% specificities ($p = 0.0131$), which was also higher than the rest of the perimetric indices and OCT ($p < 0.0001$).

Table 1. Performance of Laguna OnhE, OCT, and perimetry indices.

	GDF	RNFLT	Rim Area	OCT-C/D	MD	PSDr	TCV
AUC	0.970	0.875	0.926	0.921	0.897	0.883	0.904
SE	0.0054	0.0135	0.0104	0.0099	0.012	0.013	0.0115
CI	0.956–0.981	0.850–0.897	0.906–0.943	0.901–0.939	0.874–0.917	0.859–0.904	0.881–0.923
	GDF	**RNFLT**	**Rim Area**	**OCT-C/D**	**MD**	**PSD-sLVr**	
RNFLT	$p < 0.0001$						
Rim Area	$p < 0.0001$	$p < 0.0001$					
OCT-C/D	$p < 0.0001$	$p = 0.0006$	$p = 0.5812$				
MD	$p < 0.0001$	$p = 0.1389$	$p = 0.0422$	$p = 0.0869$			
PSDr	$p < 0.0001$	$p = 0.6013$	$p = 0.0039$	$p = 0.0083$	$p = 0.0692$		
TCV	$p < 0.0001$	$p < 0.0001$	$p = 0.0824$	$p = 0.1731$	$p = 0.5053$	$p = 0.0066$	

GDF = Laguna OnhE globin distribution factor; RNFLT = OCT retina nerve fiver layer thickness; OCT-C/D = OCT vertical cup/disc ratio; MD = perimetric mean defect or deviation; PSDr = perimetric pattern standard deviation or square root of loss variance rectified; TCV = perimetric threshold coefficient of variation; AUC = area under the receiver operating characteristic curve; SE = standard error; and CI = 5–95% confidence intervals.

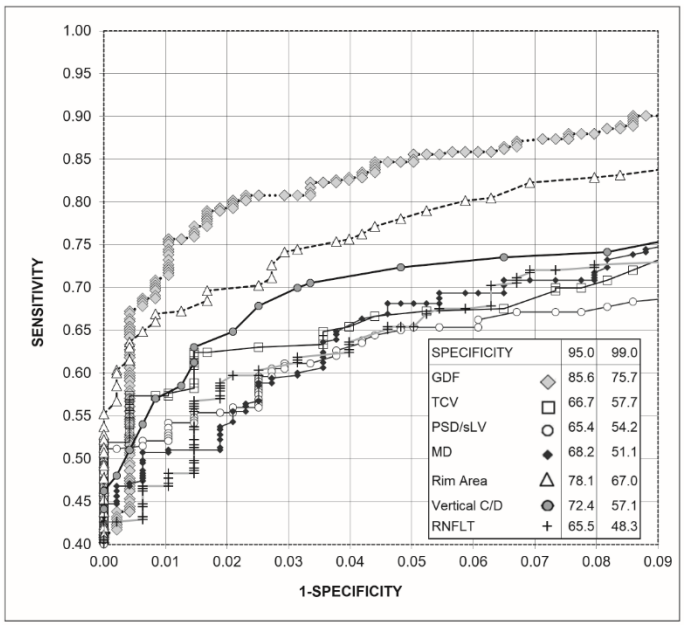


Figure 2. Receiver operating characteristic curve (ROC) obtained with Laguna ONhE (GDF), Cirrus-OCT (rim area, vertical C/D ratio, and RNFLT) and perimetry (MD, TCV, and PSD-sLV) indices, as well as percentage sensitivities for 95% and 99% specificities.

3.2. Combination of Laguna ONhE and Perimetric Indices on the Total Sample (without Separating Confirmed and Suspected Glaucoma)

Table 2 shows the results of the combinations of the two examination methods (Laguna ONhE and OCT) and the perimetric indices of uniformity and harmony. The AUCs did not show significant differences when combining GDF with each of the two perimetric indices, but these two combinations produced significantly larger curves than the combinations of perimetric and OCT indices. The RNFLT and PSDr combination performed better than the other OCT indexes combinations with the perimetric PSDr. The combinations with the perimetric MD performed worse and were thus ignored.

Table 2. Performance of the combination of Laguna ONhE and Cirrus-OCT variables and perimetric variables.

	GDF and TCV	GDF and PSDr	Rim Area and PSDr	OCT-C/D and PSDr	RNFLT and PSDr
AUC	0.978	0.976	0.945	0.947	0.915
SE	0.00459	0.00506	0.00872	0.0079	0.0109
CI	0.965–0.987	0.963–0.986	0.927–0.960	0.929–0.961	0.894–0.933
	GDF and TCV	**GDF and PSDr**	**Rim Area and PSDr**	**OCT-C/D and PSDr**	
GDF and PSDr	$p = 0.2974$				
Rim Area and PSDr	$p = 0.0001$	$p = 0.0002$			
OCT-C/D and PSDr	$p < 0.0001$	$p < 0.0001$	$p = 0.8728$		
RNFLT and PSDr	$p < 0.0001$	$p < 0.0001$	$p = 0.0015$	$p = 0.0004$	

GDF = Laguna ONhE globin distribution factor; RNFLT = OCT retina nerve fiver layer thickness; OCT-C/D = OCT vertical cup/disc ratio; MD = perimetric mean defect or deviation; PSDr = perimetric pattern standard deviation or square root of loss variance rectified; TCV = perimetric threshold coefficient of variation; AUC = area under the receiver operating characteristic curve; SE = standard error; and CI = 5–95% confidence intervals.

Figure 3 shows the ROC curves obtained for the combinations of the Laguna ONhE and OCT indices with those of perimetry. The combinations of GDF and TCV and GDF and rectified PSD-sLV did not present differences in sensitivity for 95.0% ($p = 0.790$) and 99.0% ($p = 0.674$) specificities. For 95.0% specificity, both combinations had higher sensitivities than the combinations of perimetry and OCT ($p < 0.0005$). For 99.0% specificity, the combination GDF and TCV had higher sensitivity than the rim area and rectified PSD-sLV ($p = 0.011$). GDF and rectified PSD-sLV surpassed this with $p = 0.032$. Both combinations surpassed the combinations used by the other OCT and perimetrics indices ($p < 0.0005$).

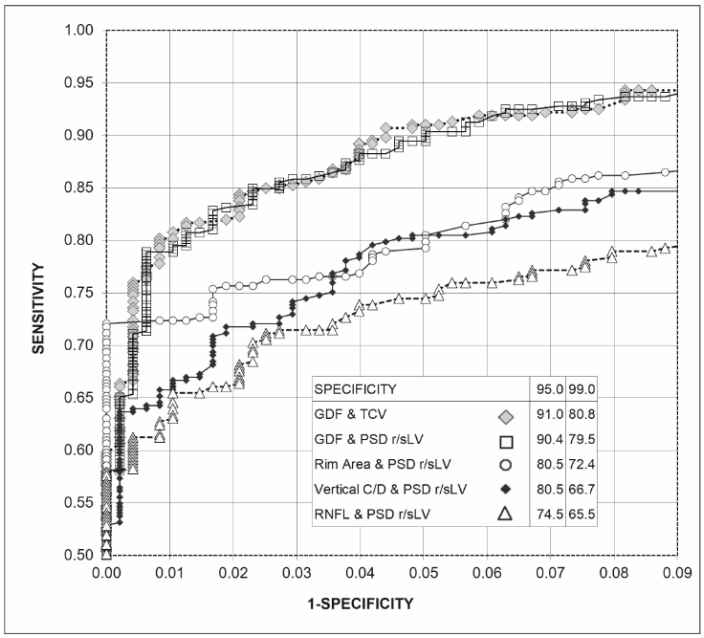
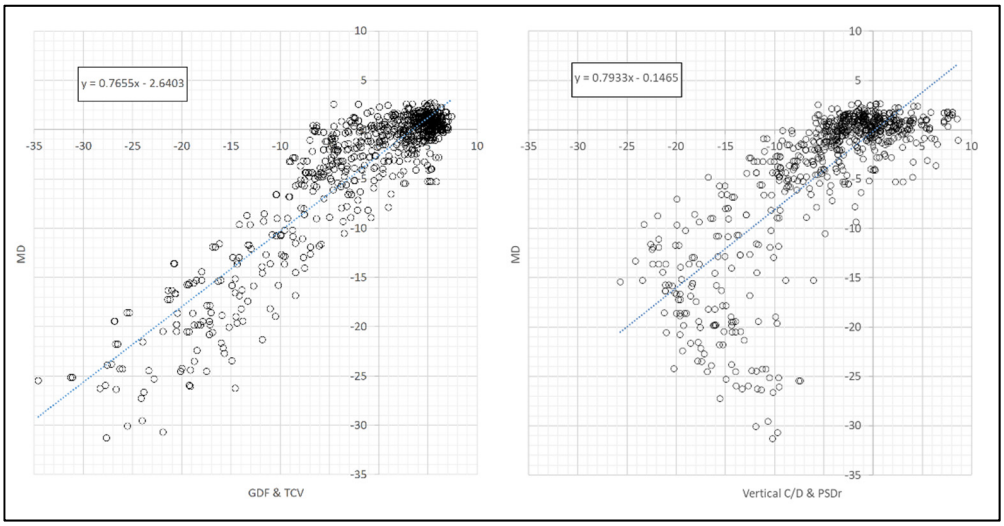


Figure 3. Graph of the receiver operating characteristic curve (ROC) of the combinations of Laguna ONhE and Cirrus-OCT indices with the perimetrics, as well as percentage sensitivities for 95% and 99% specificities.

A linear relationship was observed between perimetric MD and the GDF and TCV combination (Figure 4a), while the relationship was largely curvilinear when comparing MD with the combination of OCT morphological indices and TCV or PSDr (Figure 4b). The correlation between MD and GDF and TCV (r = 0.9132) was significantly higher than that between MD and GDF and PSDr (r = 0.8927) ($p = 0.0252$).



(**a**) (**b**)

Figure 4. Relationship between the perimetric MD index and (**a**) the combined GDF and TCV and (**b**) the vertical C/D and PSDr (right) indices.

3.3. Results of the Indices of the Three Testing Methods on the Total Sample Compared to Confirmed and to Suspected Glaucoma

Table 3 shows that Laguna ONhE GDF index obtained significantly higher results than all perimetric and Cirrus-OCT indices in suspect glaucomas. It also achieves the highest ROC area in confirmed glaucoma, but without significant differences with several of the other indices.

Table 3. Performance of Laguna ONhE, OCT, and perimetry indices.

	GDF	RNFLT	Rim Area	OCT-C/D	MD	PSDr	TCV
AUC susp.	0.932	0.709	0.827	0.834	0.687	0.651	0.718
AUC conf.	0.986	0.944	0.968	0.958	0.985	0.98	0.981
SE susp.	0.013	0.0303	0.025	0.0223	0.0297	0.0311	0.029
SE conf.	0.00437	0.0107	0.00877	0.00867	0.00316	0.00432	0.00391
CI susp.	0.908–0.951	0.670–0.746	0.794–0.857	0.801–0.864	0.647–0.725	0.611–0.690	0.679–0.754
CI conf.	0.974–0.993	0.925–0.960	0.952–0.979	0.940–0.971	0.973–0.993	0.966–0.989	0.968–0.990
	GDF	**RNFLT**	**Rim Area**	**OCT-C/D**	**MD**	**PSDr**	
RNFLT susp.	$p < 0.0001$						
RNFLT conf.	$p = 0.0002$						
RimArea susp.	$p < 0.0001$	$p = 0.0002$					
Rim Area conf.	$p = 0.0635$	$p = 0.0490$					
OCT-C/D susp.	$p < 0.0001$	$p = 0.0001$	$p = 0.7755$				
OCT-C/D conf.	$p = 0.0019$	$p = 0.2691$	$p = 0.1642$				
MD susp.	$p < 0.0001$	$p = 0.5917$	$p = 0.0004$	$p = 0.0001$			
MD conf.	$p = 0.8630$	$p = 0.0002$	$p = 0.0636$	$p = 0.0025$			
PSDr susp.	$p < 0.0001$	$p = 0.1429$	$p < 0.0001$	$p < 0.0001$	$p = 0.2399$		
PSDr conf.	$p = 0.2847$	$p = 0.0015$	$p = 0.2264$	$p = 0.0199$	$p = 0.2146$		
TCV susp.	$p < 0.0001$	$p = 0.8057$	$p = 0.0022$	$p = 0.0008$	$p = 0.2796$	$p = 0.0050$	
TCV conf.	$p = 0.3720$	$p = 0.0008$	$p = 0.1663$	$p = 0.0106$	$p = 0.2885$	$p = 0.6232$	

GDF = Laguna ONhE globin distribution factor; RNFLT = OCT retina nerve fiver layer thickness; OCT-C/D = OCT vertical cup/disc ratio; MD = perimetric mean defect or deviation; PSDr = perimetric pattern standard deviation or square root of loss variance rectified; TCV = perimetric threshold coefficient of variation; AUC = area under the receiver operating characteristic curve; SE = standard error; and CI = 5–95% confidence intervals.

3.4. Performance of Combined Indexes in Normal Subjects as Compared to Confirmed and to Suspected glaucoma

Table 4 shows the results of the best combinations of two non-psychophysical examination methods (Laguna ONhE and OCT), with the perimetric indices of uniformity and harmony in confirmed and suspected glaucoma. Significant differences between the two could be observed in glaucoma suspects. In suspected glaucomas, the combinations of the three OCT indices and the perimetric PSDr did not show significant differences. The combinations with the perimetric MD were worse and thus ignored.

Table 4. Performance of the combination of Laguna ONhE and Cirrus-OCT variables and perimetric variables between reference and confirmed (conf.) and suspect (susp.) glaucoma cases.

	GDF and TCV	GDF and PSDr	Rim Area and PSDr	OCT-C/D and PSDr	RNFLT and PSDr
AUC conf.	0.995	0.995	0.988	0.988	0.987
AUC susp	0.935	0.932	0.843	0.847	0.742
SE conf.	0.0021	0.0021	0.0049	0.0035	0.0033
SE susp.	0.013	0.015	0.024	0.022	0.029
CI conf.	0.987–0.999	0.987–0.999	0.977–0.995	0.977–0.995	0.976–0.994
CI susp	0.912–0.954	0.908–0.951	0.810–0.871	0.815–0.876	0.704–0.77
	GDF and TCV	**GDF and PSDr**	**Rim Area and PSDr**	**OCT-C/D and PSDr**	
GDF and PSDr conf.	$p = 0.695$				
GDF and PSDr susp.	$p = 0.323$				
Rim Area and PSDr conf.	$p = 0.187$	$p = 0.195$			
Rim Area and PSDr susp.	$p < 0.0001$	$p = 0.0002$			
OCT-C/D and PSDr conf.	$p = 0.060$	$p = 0.064$	$p = 0.8728$		
OCT-C/D and PSDr susp.	$p < 0.0001$	$p < 0.0001$	$p = 0.839$		
RNFLT and PSDr conf.	$p = 0.036$	$p = 0.037$	$p = 0.852$	$p = 0.859$	
RNFLT and PSDr susp.	$p < 0.0001$	$p < 0.0001$	$p = 0.0005$	$p = 0.0001$	

GDF = Laguna ONhE globin distribution factor; RNFLT = OCT retina nerve fiver layer thickness; OCT-C/D = OCT vertical cup/disc ratio; MD = perimetric mean defect or deviation; PSDr = perimetric pattern standard deviation or square root of loss variance rectified; TCV = perimetric threshold coefficient of variation; AUC = area under the receiver operating characteristic curve; SE = standard error; and CI = 5–95% confidence intervals.

4. Discussion

Several current papers that apply deep learning to the study of glaucoma exhibit interesting theoretical approaches; however, they are not always based on real-life scenarios, but on samples collected by outside groups. Several review papers [14,25] highlight the results of others that exclude suspected glaucomas from the sample, not fully reflecting the reality of the clinical problem. They often do not clarify the type of sample analysed [26], or they base the reference classification (normal or glaucoma) exclusively on subjective expert opinions on the images [27]. Others include only cases with large cups as glaucomas [28], or do not indicate the specificity achieved [29]. Achieving a good sensitivity is not sufficient in this type of task [15] if the specificity is not high. Our analysis was performed on a real sample, without excluding doubtful or intermediate cases, and trying to avoid biases that may occur when differentiating between confirmed and suspected glaucoma [23], although both groups were analysed separately, according to the most common criteria in the literature.

Selectively analysing the optic nerve and controlling image quality has undoubted advantages over the use of wider eye fundus images [30], although it has been suggested that the sectoral atrophy of fibres could enrich the diagnosis [31].

Attempting to emulate the behaviour of an OCT to obtain its most efficient indices is an interesting approach [32]. However, the classifiers obtained through deep learning produce results that tend to involve the extremes of a dichotomous series (for example: Yes/No glaucoma or Yes/No image of adequate quality) but do not provide a gradual value of the degree of defect. For this reason, in our case, we attempted to combine these results with the estimation of other indices, such as the cup/disc ratio, which is obtained by multiple regression from sectoral haemoglobin estimates [10], and/or with the results of

visual field examinations. This type of combination provides new indices with progressive gradation to better assess the level of defect and, likely, its degree of progression.

In general, published procedures require the intervention of experts for pre-processing or assessing if the images are of sufficient quality for analysis. Our procedure reached a degree of full automation in this respect, requiring no human intervention.

It is well known that the relationship between morphology and function estimated by OCT and perimetry is not linear. However, the precocity of detectable morphological damage with respect to functional damage has been questioned by several authors [33,34]. Some aspects can be complementary. Indeed, our study suggests that the association between the Laguna ONhE algorithm and perimetry seems to provide highly promising results. In particular, the linearity between the perimetric mean deviation (MD), the combination of the morphological and perfusion information presented by the Laguna ONhE GDF index, and the functional disharmony assessed by TCV provides a new perspective in the interpretation of these relationships.

It is noteworthy that this combination achieves high sensitivity while retaining high specificity. In a disease of uncertain onset, such as glaucoma, it is advisable to evaluate it with highly specific procedures, in order not to establish unnecessary treatments and to avoid overloading the health systems.

The combination of the indices provided by OCT angiography with perimetric indices should be evaluated in the future, in order to compare their results with those presented in this paper. However, an important factor to take into account in this type of comparison is the cost–benefit ratio for healthcare systems, given the low cost of some current fundus cameras, and the possibility of evaluating their images via telemedicine, without additional equipment. Such an evaluation should also be carried out in the near future.

5. Conclusions

The Laguna ONhE method arose from the observation of differences in RGB frequency histograms of optic disc structures (vessels, rim, and cup) when analysing colour photographs of the optic nerve, as can be seen in Figure 6 of our first publication [4]. The favourable results of a simple, non-invasive test, such as the one we propose for the assessment of glaucoma, were cited in the introduction, and are further evident in this new study, as they are enhanced by functional perimetric results. The application of deep learning facilitates the use of the Laguna ONhE software, taking advantage of expert experience, and improving its reproducibility, sensitivity, and specificity. Morphology, function, and perfusion can be combined for the optimal evaluation of glaucoma. However, these results should be confirmed by other independent studies.

Author Contributions: Conceptualization, M.G.-H., D.G.-H., and M.G.d.l.R.; methodology, M.G.d.l.R., D.G.-H., and M.G.-H.; software, D.G.-H. and D.P.-B.; validation, M.G.d.l.R., D.G.-H., and M.G.-H.; formal analysis, M.G.d.l.R. and M.G.-H.; investigation, M.G.d.l.R., D.G.-H., M.G.-H., P.R.-E., and N.B.-C.; resources, M.G.d.l.R., D.G.-H., and M.G.-H.; data curation, M.G.d.l.R. and D.P.-B.; writing—original draft preparation, M.G.-H. and M.G.d.l.R.; writing—review and editing, M.G.d.l.R., D.G.-H., and M.G.-H.; visualization, M.G.d.l.R.; supervision, M.G.d.l.R.; project administration, M.G.-H.; and funding acquisition, M.G.-H. and D.G.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported in part by a grant (PI12/02307) from the Instituto de Salud Carlos III with FEDER funds and by Instrumentacion y Oftalmologia (INSOFT SL) company.

Institutional Review Board Statement: The study protocol of this cross-sectional retrospective study adhered to the principles of the 1975 Declaration of Helsinki revised in 2013, and was approved by the Research Ethics Committee of the Hospital Universitario de Canarias (CHUC_2018_09 (V3)).

Informed Consent Statement: Consent was obtained from all participants.

Data Availability Statement: The datasets analysed and generated during the study can be found at the following link: https://cloud.insoft.es/s/sRaFgmjQp7S4sgR.

Acknowledgments: The authors would like to thank RetinaLyze® for their support.

Conflicts of Interest: Daniel Gonzalez-Hernandez, Marta Gonzalez-Hernandez, and Manuel Gonzalez de la Rosa participate in the patent rights of the Laguna ONhE method and partners of INSOFT S.L.; Daniel Perez-Barbudo is employed by INSOFT S.L.; Paloma Rodriguez-Esteve and Nismar Betancor-Caro have no commercial interest.

Appendix A. Computing Development Setup

For the development of the Laguna ONhE program and its neural networks, the Python programming language (ver. 3.7.1, Python Software Foundation, Beaverton, OR, USA) was selected.

Machine learning and deep learning development was carried out using images provided by RetinaLyze®, partner of INSOFT SL. For each neural network, the corresponding image data sets were divided into three subsets. The training process of the neural networks requires two different images and ground-truths subsets (training and validation) to fully pass through them in several cycles in order to complete the training process. Once the network is trained on an independent test subset, which has never been seen by the algorithm, it is used to estimate the method's efficacy.

The libraries used were Keras, with Tensorflow background for the implementation of the architectures; OpenCV, for RGB component analysis, image pre-processing, and post-processing; Numpy, for calculations; Pandas, for data manipulation; and Scikit-learn for splitting the training set between training and validation sets.

A. Optic Disc segmentation: In a previous experiment [13], we used a deep learning U-Net architecture [16] for semantic segmentation. In this case, 40,000 images were used to segment the optic disc, identifying the inner edge of Elschnig's scleral ring. This was achieved by manually identifying by an expert. Elschnig's scleral ring is observed as a thin white ring, immediately adjacent to the margin of the optic disc. An example is shown in Figure 1 of reference [13].

B. Vessel segmentation: 4195 optic disc images were used to train the neural network on vessel segmentation. The vessels were drawn by hand by a single expert on each image, using the GIMP program (Version 2.10.10. http://www.gimp.org, Accessed on 21 July 2019). The data set was then divided into training (3684), validation (921), and test (410) subsets. The training set was used to train the neural network, obtaining information on each pixel belonging to a vessel or to adjacent tissue. The same basic structure of U-Net was used with contracting path filter sizes of 16, 32, 64, 128, 256, and 512 pixels towards a centre filter of 1024 pixels with batch normalization at each convolution block and ReLU activation.

Results: The Sorensen–Dice index of coincidence between the ground truth and the network was 0.932, and its corresponding Jaccard index was 0.877, in the 410 images of the test subset.

C. Eye labelling as right or left: 4201 images of the optic disc and surrounding region were divided into subsets for training (3025), validation (756), and testing (420). The results were compared with two other criteria: the greater presence of vessels in the nasal half of the optic disc and the estimation of a greater presence of haemoglobin in such area. The optic disc position on wide-angle retinographies was not considered, in order to explicitly assess the method's ability to identify it as isolated or centred.

A pre-trained ResNet50-v2 was used, with the top layers cut out. The network was trained in the ImageNet database. The network was fine-tuned with a global average pooling layer and three fully connected layers using stochastic gradient descent (SGD).

Results: 98.5% of the eyes of the test dataset were correctly automatically classified using only the optic nerve image, observing a greater presence of haemoglobin in the nasal sector; 99.3% by the greater presence of vessels in the nasal half; and 95.9% by the neural network. The coincidence of two of these three criteria identified 99.5% of the cases.

D. Image quality: Previous experience in screening systems with multiple users showed us that it is necessary to detect and exclude images of very poor quality, incomplete

optic disc, etc. In total, 7048 images were subjectively classified into two groups according to their quality by a single expert (4780 acceptable and 2268 unacceptable quality), and divided into subsets of training (5075), validation (1268), and testing (705).

For this analysis, a ResNet50 network architecture was used with two fully connected layers. A LeakyRelu activation function [35] and an Adam optimizer were used for stochastic optimization [36]. The network was first trained for 50 epochs and a learning rate of 0.0001 and was fine tuned for an extra 20 epochs at a learning rate of 0.00001.

Results: In the test subset, an AUC of 0.958 was obtained, with a 5–95% confidence interval 0.940 to 0.971 ($p < 0.0001$). For a sensitivity of 95.0%, indicating that the image quality was correct, the specificity was 82.8%.

E. Normal vs. glaucoma discrimination: To train the network to distinguish normal cases from glaucoma, the haemoglobin values estimated in the tissue were represented as a pseudo-colour image, preserving the details of its surroundings. Normalization of its colour channels was applied to reduce the differences due to the fundus camera or flash used. In this way, the training was intended to include the atrophies that glaucoma usually produces outside the optic disc.

Then, 3114 such images belonging to subjects verified as normal (1518) and of glaucoma or glaucoma suspects (1596) were divided into subsets of training (2274), validation (529), and testing (311). Transfer learning was used, passing the images to a ResNet50 network [37], with the upper layers cut out, including the pre-trained weights in the ImageNet database [38]. Therefore, independently of the original image size, which was 1956 × 1934, all images were trimmed around the ONH segmentation and resized to 224 × 224 in order to fine tune the pretrained ResNet50 network. Random horizontal flip, rotation of up to 25°, and brightness darkening up to a 20% were implemented. A global average pooling layer and three specific layers were then included to classify the case as normal or glaucoma [39]. After each of the first two fully connected layers, a dropout layer with a rate value of 0.5 was included. Additionally, both use a hyperbolic tangent (tanh) activation function. The final one uses a softmax activation function. Categorical cross-entropy was used as a loss function, and accuracy was used as a metric to evaluate the performance of the model. Stochastical gradient decent (SGD) was used as an optimizer with a learning rate of 0.0001 and momentum of 0.9. The classification results obtained by deep learning were associated with the distribution of haemoglobin and the estimated cup/disc ratios to define a new value for the globin distribution function (GDF) index, as previously described [4,5].

Results: In the test subset, an AUC of 0.976 was obtained, with a 5–95% confidence interval of 0.953 to 0.990 ($p < 0.0001$).

In order to assess the model's efficiency [40], the learning curve of the training and validation loss as well as the confusion matrix of the test dataset are shown in Figure A1a,b. These graphical representations are the most commonly accepted methods to determine the model's accuracy and generalization capability [41,42].

These show the capacity of the dataset, and therefore seem to us to be more representative of the effectiveness of the network than other possibilities such as "heat maps", which show examples of how the network focuses on representative details of specific cases, but do not reflect the overall sensitivity of the method [43]. Figure A2 shows four examples of these heat maps.

As can be seen from the learning curve, we achieved an optimal fit, which is the goal of any learning algorithm achieving the minimum generalization gap [44]. Both the training and validation losses decreased to a point of stability. After 1000 epochs, the training was stopped in order to avoid overfitting. Additionally, dropout was activated when training, but deactivated when evaluating the validation data, which makes it reasonable for the validation error to be smaller than training error. The confusion matrix showed that the model had very good accuracy in the test data set. The learning curves show that both the training and validation losses drop rapidly until 100 epochs and then both continue to decrease at a slower pace until both curves meet. Even though the validation error in the

graph looks as though it stops lowering from epoch 100 onwards, this is a scale problem. The validation curve does not stop; it continues to descend.

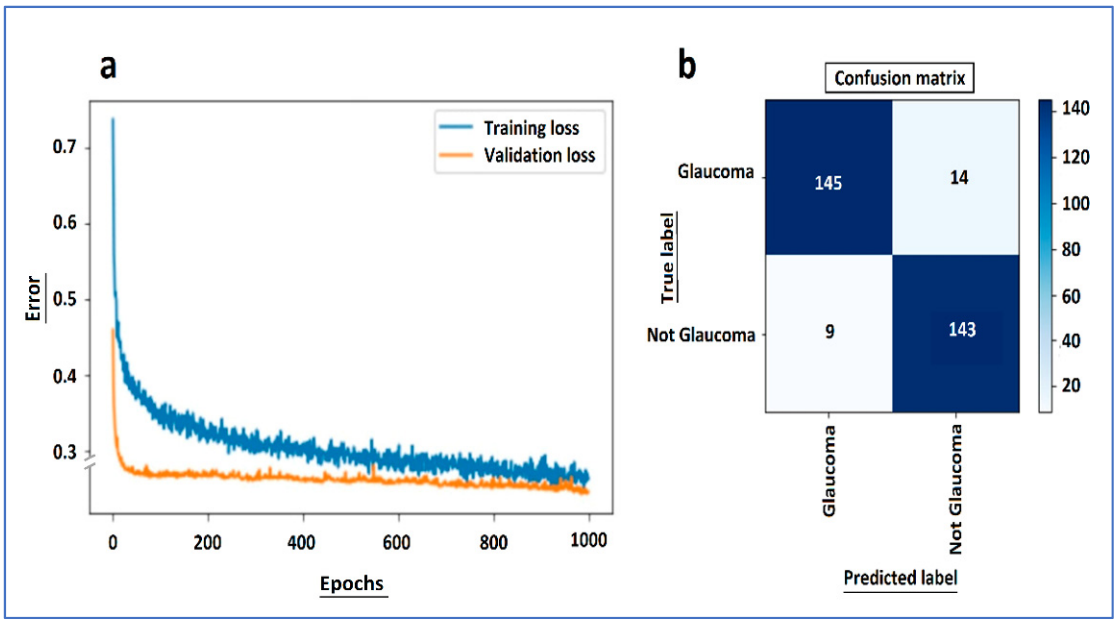


Figure A1. From the normal vs. glaucoma neural network: (**a**) learning curve of the training and validation loss and (**b**) confusion matrix of the test dataset.

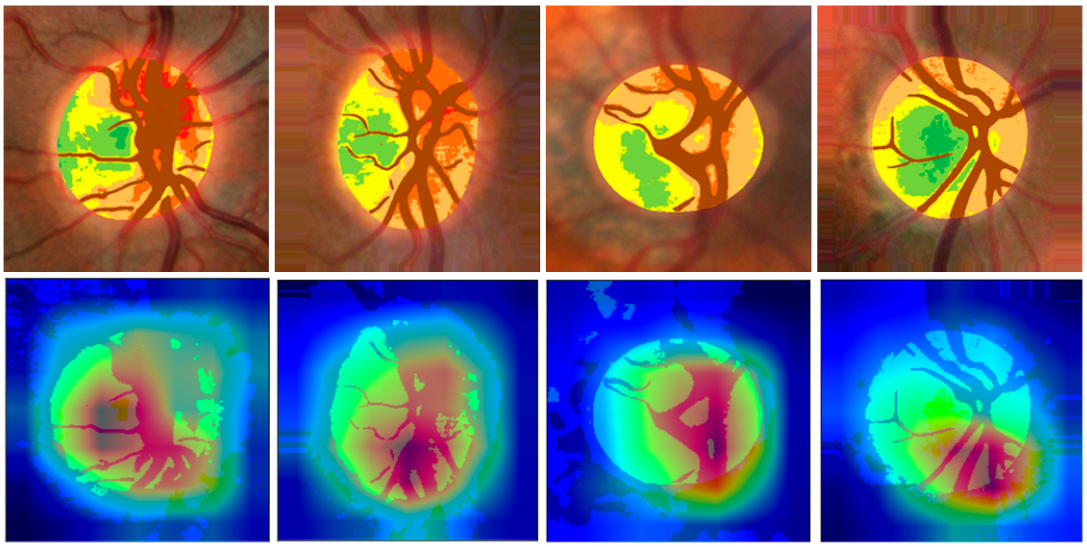


Figure A2. (**Top**): Pseudo-colour images of haemoglobin distribution in two normal optic discs (left) and two glaucomas (right) and the normalised environment, which are used for network training. The eyes are always analysed with the temporal region to the left. (**Below**): The Gradient-weighted Class Activation Maps, known as Grad-CAM saliency map, or heat maps, indicate that the networks seem to be particularly focused on the cup and the upper and lower poles.

References

1. 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]
2. Hoyt, W.F.; Frisén, L.; Newman, N.M. Fundoscopy of nerve fiber layer defects in glaucoma. *Investig. Ophthalmol.* **1973**, *12*, 814–829.
3. Jia, Y.; Morrison, J.C.; Tokayer, J.; Tan, O.; Lombardi, L.; Baumann, B.; Lu, C.D.; Choi, W.; Fujimoto, J.G.; Huang, D. Quantitative OCT angiography of optic nerve head blood flow. *Biomed. Opt. Express* **2012**, *3*, 3127–3137. [CrossRef]
4. de la Rosa, M.G.; Gonzalez-Hernandez, M.; Sigut, J.; Alayon, S.; Radcliffe, N.; Mendez-Hernandez, C.; García-Feijoo, J.; Fuertes-Lazaro, I.; Perez-Olivan, S.; Ferreras, A. Measuring haemoglobin levels in the optic nerve head: Comparisons with other structural and functional parameters of glaucoma. *Investig. Ophthalmol. Vis. Sci.* **2013**, *54*, 482–489. [CrossRef] [PubMed]
5. Pena-Betancor, C.; Gonzalez-Hernandez, M.; Fumero-Batista, F.; Sigut, J.; Medina-Mesa, E.; Alayon, S.; de la Rosa, M.G. Estimation of the relative amount of haemoglobin in the cup and neuroretinal rim using stereoscopic color fundus images. *Investig. Ophthalmol. Vis. Sci.* **2015**, *56*, 1562–1568. [CrossRef] [PubMed]
6. Uña, I.R.; Hernandez, C.M.; Saenz-Frances, F.; Feijóo, J.G. Correlación de la relación excavación/papila óptica medida mediante HRT-III, SD-OCT y el dispositivo de colorimetría fotográfica Laguna On_hE. *Arch. Soc. Esp. Oftalmol.* **2015**, *90*, 212–219.
7. Mendez-Hernandez, C.; Rodriguez-Uña, I.; Rosa, M.G.-d.-l.; Arribas-Pardo, P.; Garcia-Feijoo, J. Glaucoma diagnostic capacity of optic nerve head haemoglobin measures compared with spectral domain OCT and HRT III confocal tomography. *Acta Ophthalmol.* **2016**, *94*, 697–704. [CrossRef]
8. Gonzalez-de-la-Rosa, M.; Gonzalez-Hernandez, M.; Mendez-Hernandez, C.; Garcia-Martin, E.; Fumero-Bautista, F.; Alayon, S.; Sigut, J. Glaucoma imaging: Measuring haemoglobin levels in the optic nerve head for glaucoma management. In *Glaucoma Imaging*; Ferreras, A., Ed.; Springer: Berlin/Heidelberg, Germany, 2016; pp. 265–280.
9. Mendez-Hernandez, C.; Garcia-Feijoo, J.; Arribas-Pardo, P.; Saenz-Frances, F.; Rodriguez-Uña, I.; Fernandez-Perez, C.; de la Rosa, M.G. Reproducibility of optic nerve head haemoglobin measures. *J. Glaucoma* **2016**, *25*, 348–354. [CrossRef]
10. Medina-Mesa, E.; Gonzalez-Hernandez, M.; Sigut, J.; Fumero-Batista, F.; Pena-Betancor, C.; Alayon, S.; de la Rosa, M.G. Estimating the amount of haemoglobin in the neuroretinal rim using color images and OCT. *Curr. Eye Res.* **2016**, *41*, 798–805. [CrossRef]
11. Perucho-González, L.; Méndez-Hernández, C.D.; González-de-la-Rosa, M.; Fernández-Pérez, C.; Sáez-Francés, F.; Andrés-Guerrero, V.; García-Feijoó, J. Preliminary study of the differences in optic nerve head haemoglobin measures between patients with and without childhood glaucoma. *J. Pediatric Ophthalmol. Strabismus* **2017**, *54*, 387–394. [CrossRef]
12. Gonzalez-Hernandez, M.; Sigut Saavedra, J.; de la Rosa, M.G. Relationship between retinal nerve fiber layer thickness and haemoglobin present in the optic nerve head in glaucoma. *J. Ophthalmol.* **2017**, *2017*, 2340236. [CrossRef]
13. Gonzalez-Hernandez, D.; Diaz-Aleman, T.; Perez-Barbudo, D.; Mendez-Hernandez, C.; de la Rosa, M.G. Segmentation of the optic nerve head based on Deep Learning to determine its haemoglobin content in normal and glaucomatous subjects. *J. Clin. Exp. Opthamol.* **2018**, *9*, 1000760. [CrossRef]
14. Kapoor, R.; Walters, S.P.; Al-Aswad, L.A. The current state of artificial intelligence in ophthalmology. *Surv. Ophthalmol.* **2019**, *64*, 233–240. [CrossRef]
15. Shibata, N.; Tanito, M.; Mitsuhashi, K.; Fujino, Y.; Matsuura, M.; Murata, H.; Asaoka, R. Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. *Sci. Rep.* **2018**, *8*, 14665. [CrossRef]
16. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015*; Springer: Cham, Switzerland, 2015; pp. 234–241.
17. Mendez-Hernandez, C.; Wang, S.; Arribas-Pardo, P.; Salazar-Quiñones, L.; Güemes-Villahoz, N.; Fernandez-Perez, C.; Garcia-Feijoo, J. Diagnostic validity of optic nerve head colorimetric assessment and optical coherence tomography angiography in patients with glaucoma. *Br. J. Ophthalmol.* **2020**. [CrossRef]
18. Abreu-Gonzalez, R.; Gonzalez-Hernandez, M.; Pena-Betancor, C.; Rodriguez-Esteve, P.; De La Rosa, M.G. New visual field indices of disharmony for early diagnosis of glaucoma, alone or associated with conventional parameters. *Eur. J. Ophthalmol.* **2018**, *28*, 590–597. [CrossRef] [PubMed]
19. De La Rosa, M.G.; Gonzalez-Hernandez, M.; Diaz-Aleman, T. Linear regression analysis of the cumulative defect curve by sectors and other criteria of glaucomatous visual field progression. *Eur. J. Ophthalmol.* **2009**, *19*, 416–424. [CrossRef] [PubMed]
20. de la Rosa, M.G.; Gonzalez-Hernandez, M. A Strategy for averaged estimates of visual field threshold: Spark. *J. Glaucoma* **2013**, *22*, 284–289. [CrossRef] [PubMed]
21. de la Rosa, M.G.; Martinez, A.; Sanchez, M.; Mesa, C.; Cordovés, L.; Losada, M.J. Accuracy of the Tendency Oriented Perimetry (TOP) in the Octopus 1-2-3 Perimeter. In *Perimetry Update 1996/1997*; Wall, M., Wild, J., Eds.; Kugler Publications: Amsterdam, The Netherlands, 1997; pp. 119–123.
22. Ransohoff, D.F.; Feinstein, A.R. Problems of Spectrum and Bias in Evaluating the Efficacy of Diagnostic Tests. *N. Engl. J. Med.* **1978**, *299*, 926–930. [CrossRef] [PubMed]
23. de la Rosa, M.G.; Gonzalez-Hernandez, M. Gold standards may bias clinical test validation. *ARS Clin. Acad.* **2020**, *6*, 5–10.
24. Delong, E.R.; Delong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845. [CrossRef]
25. Hogarty, D.T.; Mackey, D.A.; Hewitt, A.W. Current state and future prospects of artificial intelligence in ophthalmology: A re-view. *Clin. Exp. Ophthalmol.* **2019**, *47*, 128–139. [CrossRef]

26. Raghavendra, U.; Fujita, H.; Bhandary, S.V.; Gudigar, A.; Tan, J.H.; Acharya, U.R. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf. Sci.* **2018**, *441*, 41–49. [CrossRef]
27. Liu, H.; Li, L.; Wormstone, I.M.; Qiao, C.; Zhang, C.; Liu, P.; Li, S.; Wang, H.; Mou, D.; Pang, R.; et al. Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs. *JAMA Ophthalmol.* **2019**, *137*, 1353–1360. [CrossRef]
28. Li, Z.; He, Y.; Keel, S.; Meng, W.; Chang, R.T.; He, M. Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs. *Ophthalmology* **2018**, *125*, 1199–1206. [CrossRef]
29. Du, X.-L.; Li, W.-B.; Hu, B.-J. Application of artificial intelligence in ophthalmology. *Int. J. Ophthalmol.* **2018**, *11*, 1555–1561. [CrossRef] [PubMed]
30. Phan, S.; Satoh, S.; Yoda, Y.; Kashiwagi, K.; Oshika, T.; The Japan Ocular Imaging Registry Research Group. Evaluation of deep convolutional neural networks for glaucoma detection. *Jpn. J. Ophthalmol.* **2019**, *63*, 276–283. [CrossRef] [PubMed]
31. Christopher, M.; Belghith, A.; Bowd, C.; Proudfoot, J.A.; Goldbaum, M.H.; Weinreb, R.N.; Girkin, C.A.; Liebmann, J.M.; Zangwill, L.M. Performance of Deep Learning Architectures and Transfer Learning for Detecting Glaucomatous Optic Neuropathy in Fundus Photographs. *Sci. Rep.* **2018**, *8*, 16685. [CrossRef] [PubMed]
32. Thompson, A.C.; Jammal, A.A.; Medeiros, F.A. A Deep Learning Algorithm to Quantify Neuroretinal Rim Loss From Optic Disc Photographs. *Am. J. Ophthalmol.* **2019**, *201*, 9–18. [CrossRef] [PubMed]
33. Gonzalez-Hernandez, M.; Pablo, L.E.; Armas-Dominguez, K.; de La Vega, R.R.; Ferreras, A.; De La Rosa, M.G. Structure-function relationship depends on glaucoma severity. *Br. J. Ophthalmol.* **2009**, *93*, 1195–1199. [CrossRef]
34. Malik, R.; Swanson, W.H.; Garway-Heath, D. Structure–function relationship in glaucoma: Past thinking and current concepts. *Clin. Exp. Ophthalmol.* **2012**, *40*, 369–380. [CrossRef] [PubMed]
35. Zhang, X.; Trmal, J.; Povey, D.; Khudanpur, S. Improving deep neural network acoustic models using generalized maxout networks. In Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Florence, Italy, 4–9 May 2014; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2014; pp. 215–219.
36. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations (ICLR), San Diego, CA, USA, 7–9 May 2015; pp. 1–15.
37. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 1 June 2016; pp. 770–778.
38. ImageNet Project. Stanford Vision Lab, Stanford University, Princeton University. 2016. Available online: http://www.image-net.org/ (accessed on 9 November 2020).
39. Shin, H.-C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef] [PubMed]
40. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; The MIT Press: Cambridge, MA, USA, 2016; pp. 96–161.
41. Anzanello, M.J.; Fogliatto, F.S. Learning curve models and applications: Literature review and research directions. *Int. J. Ind. Ergon.* **2011**, *41*, 573–583. [CrossRef]
42. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer-Verlag: New York, NY, USA, 2013; pp. 29–42.
43. Adebayo, J.; Muelly, M.; Liccardi, I.; Kim, B. Debugging Tests for Model Explanations. *arXiv* **2020**, arXiv:2011.05429.
44. Hoffer, E.; Hubara, I.; Soudry, D. Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4 December 2017; pp. 1729–1739.