GLAUCOMA GRADING VIA MEAN DEFECT BACK PROPAGATION FROM OCT IMAGES

Elisa Ramírez*, Rocío del Amor*, Gabriel García*, and Valery Naranjo*
*Instituto de Investigación e Innovación en Bioingeniería, I3B, Universitat Politècnica de València, Valencia, Spain

***Abstract*—Glaucoma is one of the most prevalent causes of irreversible blindness worldwide. In clinical practice, the most commonly used methods for diagnosing glaucoma are based on fundus and Optical Coherence Tomography (OCT) imaging, tonometry, perimetry and campimetry. OCT image is better suited to detect changes in disease progression than other existing imaging techniques. In addition, using the mean defect index (MD), the campimetry test allows evaluating the evolution of visual field loss. Convolutional Neural Networks (CNNs) can facilitate automatic diagnosis and grading of glaucoma. However, few studies have addressed this issue using OCT images. In this paper, we aim to improve the grading of glaucoma patients through a neural network that combines the information provided by both tests, OCT images and MD index. Using a two-term loss function ($\mathcal{L}_c$ and $\mathcal{L}_R$), the proposed framework obtains from an OCT image the glaucoma grade as well as its corresponding MD value. This methodological core outperforms the classification baseline framework by 3.21%, getting a test accuracy of 0.9259. This approach would not only improve the classification of glaucoma but also extracts the MD index from OCT images without the need for a visual field (VF) test.**

***Index Terms*—Glaucoma grading, Mean defect, Optical Coherence Tomography, Convolutional Neural Networks**

I. Introduction

Glaucoma has been recognised as one of the leading causes of blindness worldwide. It is defined as a chronic progressive neurodegenerative disease characterized by the progressive loss of retinal ganglion cells (GCs) and the nerve fiber layer (RNFL) [1]. Early stages are usually asymptomatic, and thus most patients with glaucoma already have late-stage disease with substantial irreversible visual loss when they present to ophthalmologists [2]. This is because the damage to the optic nerve fibers becomes noticeable and reduction in visual field is detected when about 40% of axons are already lost [3]. Therefore, an early treatment of this chronic disease is essential for timely treatment to prevent the complete and irreversible vision loss.

Currently, the primary methods of glaucoma detection include measurement of pressure in the eye (tonometry), assessment of the impact on functional vision through visual field tests (campimetry or perimetry), optic disc examination on fundus images and evaluation of retinal thickness changes via optical coherence tomography (OCT). Regarding the detection of glaucomatous structural changes by imaging, it has relied on the assessment of optic disc photographs. In fact, the majority of studies in literature are dedicated to the analysis on fundus images [4]–[8]. However, agreement among glaucoma specialists in judging change on disc photographs is only slight to fair, and photographs do not allow quantification of rates of change [9]. Although fundus imaging is cheaper and more useful for diseases such as age-related macular degeneration or diabetic retinopathy, OCT analysis is the technique par excellence for detecting and evaluating glaucoma. Unlike fundus imaging, this technique provides a 3D projection of the retina to measure the deterioration of the nerve fibre layers, which in the clinical literature has been identified as the structure that evidences the progression of glaucoma [10]–[12].

In the last few years, deep learning has been at the forefront of imaging techniques for glaucoma assessment. Ali et al presented an automatic detection of early and advanced glaucoma. They used the transfer learning technique by fine-tuning over the ResNet-50 and GoogLeNet architecture for the early as well as advanced glaucoma detection using fundus images [8]. In line with this work, in [13], the authors used a pre-trained GoogleNet Inception v3 model for transfer learning over a dataset of 1542 images.

Most glaucoma grading studies have been done with fundus imaging [3], [7], [8], [13], but only [14] and [10] use OCT images for this task. Therefore, OCT imaging has been used mainly for the detection of the disease [15]–[18].

Several criteria have been used to stating the glaucoma [19], among them stand up the Hodapp-Parrish-Anderson (HPA) [20]. The main metric for measuring vision loss is the mean defect (MD) index, mathematically represented as the average of the individual visual field defects of all test location, expressed in decibels (dBs), as follows:

$$MD = 10 \cdot log\left(\frac{1}{N}\sum_{i=1}^{N} d_i\right) \qquad d_i = n_i - x_i \tag{1}$$

where n_i and d_i are the normal value and the sensitivity loss at test location i, respectively. This index allows for the overall assessment of the glaucoma severity, being essential to determine the progression of the disease [21]. If a visual field defect worsens, independent of whether it is a local or a diffuse defect, MD will worsen too (see Fig. 1). Therefore, MD is relevant for the classification of patients according to the degree of glaucoma progression.

This work has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation, under grant agreement No. 860627 (CLARIFY), the Spanish Ministry of Economy and Competitiveness through project PID2019-105142RB-C21 (AI4SKIN) and GVA through projects PROMETEO/2019/109 and INNEST/2021/321 (SAMUEL). Rocío del Amor work have also been supported by the Spanish Government under FPU Grant [FPU20/05263].

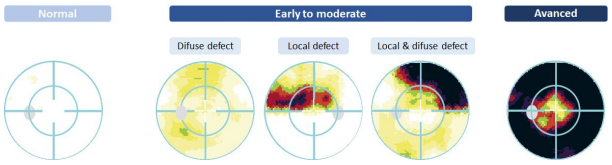

Fig. 1. Different visual fields with increasingly severe visual field loss [21].

To the best of the authors' knowledge, the proposed method is the first aimed at grading glaucoma from OCT images using the MD value to improve the predictions. By adding a regression module, the proposed framework is able to estimate the MD value from OCT images. The whole method is optimized through a custom weighted loss function composed of two terms: classification and regression. This study compare results between only classification and the proposed methodology by adding a MD prediction module. The ablation experiments conducted in this paper evidences the improvement of the predictive ability when the MD information is included in the network via backward-propagation.

II. Methods

A. Baseline for staging Glaucoma

Formally, the training set can be denoted as $D = \{X, A, Y\}$ in which X and A are the inputs and Y the associated ground-truth set. $X = \{x_1, x_2, ..., x_i, ..., x_N\}$ is the set of OCT input images, where each x_i has 248 $\times$ 384 dimensions. Second input, $A = \{a_1, a_2, ..., a_j, ..., a_P\}$ represents the age of each patient j, where P is the total number of patients. Finally, $y_i \in Y$ denotes the glaucoma grade class to which each sample belongs, which can take three values: 0 - normal, 1 - early to moderate or 2 - advanced, depending on the MD index.

As shown in Fig. 2, the backbone of the proposed network is based on the VGG16 architecture. According to [22], fine-tunning with VGG16 gives better performance in glaucoma classification than other state-of-art networks. During the training process, a fine-tuning strategy was employed, freezing VGG16 layers and transferring knowledge acquired when the network was trained with the *ImageNet* dataset [23]. This module aims to extract features from the input images (x_i) from by applying convolution and pooling operations, returning a vector $z_i \in Z$. Therefore, this module could be expressed as $f_\theta : X \rightarrow$ Z.

An attention module g_ϕ was included to focus on glaucoma-specific features aimed at improving the learning process. This module is composed of convolutions and squeeze-excitation blocks which improve the performance of the models adaptatively recalibrating the feature maps in the spatial axis [24]. This module returns a vector $t_i \in T$.

Lastly, the top model (h_γ) returns the glaucoma stage by using a softmax activation function which calculates the probability of the samples to belong to each of the n classes, according to $\hat{Y} = max[P_j(Y_j|x_i)]$, $j = 1, 2...n$.

For the optimization of the proposed architecture, we used the Categorical Cross-Entropy loss function, which calculates the difference between *n* distributions, being *n* the number of classes of the output as:

$$\mathcal{L}_c = -\sum_{i=1}^{n} y_i \cdot log(\hat{y_i}), \quad n = 3 \tag{2}$$

B. MD index constraint

In some situations, the analysis of OCT images is insufficient for glaucoma grading and, therefore, information from the VF test should be included. The present methodology aims to add a constraint in the baseline model to predict the MD value from the OCT-extracted features. Based on these ideas, the proposed MD constraint lies in the inclusion of a regression module at the end of the baseline model to predict the MD value. The combination of classification and regression losses for the model optimization enhances the glaucoma classification.

The proposed architecture can be observed in Fig. 2 A). The regression module is introduced after the classification layer to predict MD index from the probability of belonging to each glaucoma class (p_i) as $Reg(p; w^r) = \sum_{i=1}^{N}(p_i \cdot w_i^r)$, being w_i^r the regression weights per class adjusted during training.

The ground truth set in this module can be denoted as B, where $b_i \in B$ is each of the MD values associated with each image. The loss function selected in this phase was the Minimum Squared Error (MSE):

$$\mathcal{L}_r = \frac{1}{N}\sum_{i=1}^{N}(b_i - \hat{b_i})^2 \tag{3}$$

During the training, this term was combined with the classification loss function as shown in Figure 2 B). The λ_r is a parameter to optimizes the effect of regression on learning. The final loss function can be expressed as:

$$\mathcal{L} = \mathcal{L}_c + \lambda_r \mathcal{L}_r \tag{4}$$

Note that MD values are used for optimizing the parameters of the network and not for the prediction phase. In this way, with the proposed framework, it is possible to obtain the MD value without performing a VF test.

III. Ablation Experiments

A. Dataset

The experiments detailed in this paper were performed on a private database composed of 258 OCT images of $496 \times 768 \times 3$ pixels. In particular, 117 normal, 55 early to moderate and 86 severe glaucomatous circumpapillary samples were analysed. Each B-scan was diagnosed by experts ophthalmologists. Note that *Heidelberg Spectrallis* OCT system was employed to acquire the circumpapillary OCT images with an axial resolution of 4-5 μm. In he Hodapp-Parrish-Andersib system, early-moderate defects are characterized by an MD ranging from 6 to 12 dB, and severe visual filed defects have an MD worse than 12 dB. Therefore, classes can be denoted as:

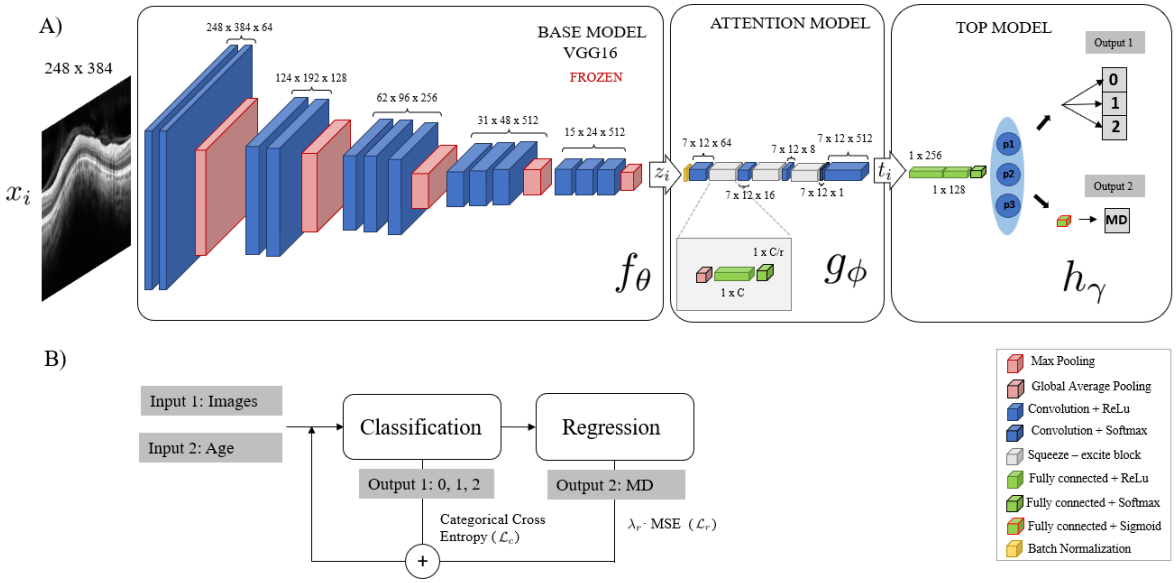

Fig. 2. A) Proposed CNN architecture to predict glaucoma grade and MD index from OCT input images. B) Combining loss functions of glaucoma grade and MD index predictions.

$$y_i = \begin{cases} Normal : 0 & \text{if } MD < 6dB \\ Early - moderate : 1 & \text{if } 6 < MD < 12dB \\ Advanced : 2 & \text{if } MD > 12dB \end{cases} \quad (5)$$

The data was partitioned into two sets: train (80%) and test (20%). K-Fold technique was used for cross-validation with $K = 5$, dividing train data into five partitions, using four for train and one for validation in each of the five iterations. The database contains OCT images from the right and left eyes of the patients as well as MD parameters obtained from the VF test. The data partitioning was carried out at the patient level to ensure that images from both eyes of a specific patient belong to the same fold. We carried out this procedure to avoid including the same patient's images in the training and validation dataset. The database contains more patients belonging to the normal class; however, the early-stage class comprises significantly fewer images. Classes were balanced, being randomly removed from each class in the different folds to guarantee equal learning across all classes during training. Before discarding samples, we performed a weighted classification loss function, enhancing the under-represented classes; however, this methodology led to worse results.

B. Selection of regression parameters

We performed an in-depth empirical exploration to select the optimal hyperparameter combination. The model was trained by optimising the result of the expression (4), an adaptive learning rate of 10e-4, with a decrease of 0.9 after 10 not-improved epochs in terms of validation loss. The learning process was addressed utilizing the Adadelta optimizer, using batches of 16 samples during 350 epochs.

In order to determine the optimal network configuration in terms of accuracy, a series of experiments have been carried out by changing different parameters. Fig. 3 shows the average accuracy and standard deviation achieved in the validation samples depending on the weighting factor λ_r included in the regression loss function, and the activation function of the regression layer. These results evidence that the use of a regression term improves the performance of the glaucoma classification. Nevertheless, using a too large slope once the constraint is satisfied ($\lambda_r = 2$) can lead to a worsening of the results. Thus, we selected $\lambda_r = 2$, which provides the best results in the validation cohort. The sigmoid activation resulted in the most accurate function for glaucoma grading, as depicted in Fig. 3. After this ablation experiment, we concluded that the best network configuration was weigh the regression loss function by 2 and use the sigmoid as the loss function.

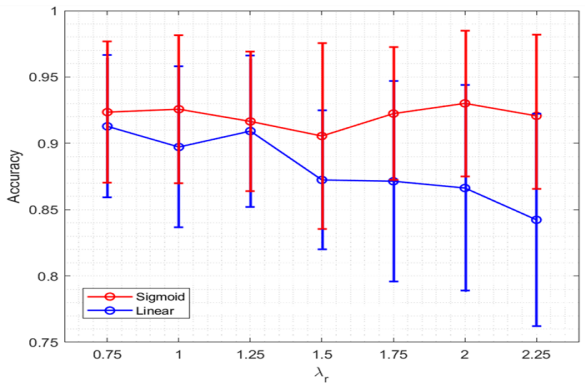

Fig. 3. Mean $\pm$ standard deviation of the accuracy value according to the weighting factor λ_r of the regression loss function and the regression layer activation function (linear or sigmoid).

TABLE I
VALIDATION RESULTS OF BASELINE (BL) AND PROPOSED MODEL (PM) USING $\lambda_r = 2$.

	Normal		Early to moderate		Advanced		MicroAvg		MacroAvg	
	BL	PM	BL	PM	BL	PM	BL	PM	BL	PM
SN	**0.99±0.02**	**0.99±0.02**	0.36±0.24	**0.83±0.18**	**0.92±0.05**	0.79±0.22	0.84±0.05	**0.89±0.08**	0.76±0.09	**0.86±0.10**
SP	0.95±0.04	**1.00±0.00**	**0.97±0.02**	0.90±0.10	0.82±0.07	**0.94±0.05**	0.92±0.03	**0.95±0.04**	0.92±0.03	**0.95±0.04**
PP	0.94±0.05	**1.00±0.00**	**0.74±0.17**	0.74±0.21	0.74±0.10	**0.91±0.08**	0.84±0.06	**0.89±0.08**	0.81±0.04	**0.86±0.09**
NP	0.99±0.02	**0.99±0.01**	0.86±0.06	**0.96±0.04**	**0.95±0.03**	0.90±0.08	0.92±0.03	**0.94±0.04**	0.93±0.02	**0.95±0.04**
FS	0.97±0.03	**0.99±0.01**	0.45±0.19	**0.76±0.17**	0.81±0.05	**0.82±0.15**	0.84±0.06	**0.89±0.08**	0.74±0.10	**0.85±0.10**
AC	0.97±0.02	**0.99±0.01**	0.84±0.05	**0.89±0.08**	0.86±0.04	**0.89±0.07**	0.89±0.04	**0.93±0.05**	0.89±0.04	**0.93±0.05**

IV. PREDICTION RESULTS

A. Validation Results

In this stage, the proposed architecture is compared with the baseline to determine the positive effect that MD prediction has on the classification. Therefore, different figures of merit as sensitivity (SN), specificity (SP), positive predictive value (PP), negative predictive value (NP), F1-score (FS), and accuracy (AC) were considered.

Table 1 shows the metrics provided by the baseline (BL) classification architecture and the proposed method (PM), of each class, and the micro-average and macro-average. Focusing on the accuracy value, the proposed method improves results in each of the classes, increasing on average approximately 4.4%. Furthermore, it is noteworthy that the sensitivity and specificity are improved by adding the regression constraint, reaching micro-average values of 0.8908 and 0.9454, respectively. The proportion of images from each class in the validation set is not equal, with more samples in the healthy class than in the other two. This is why in some figures of merit in the early to moderate and advanced categories, the standard deviation between folds is higher than expected.

The proposed model can predict the MD value with losses of 0.1155, which places the proposed model as the most suitable predictor for future occasions.

B. Test Results

The model was evaluated with the test dataset during the prediction stage. In Table 2, we report the results of each class and the micro average and macro average achieved from the baseline and proposed method. These results demonstrate that the proposed methodology outperforms the baseline for all the figures of merit in all stages of glaucoma development. By optimally adjusting the weighting value of the regression loss function, the model aims to serve as a tool for healthcare professionals with an accuracy of 92.59%, being higher in the detection of normal patients (100%) and maintaining accuracy of 88.89% in the classification of patients with early to moderate and advanced glaucoma.

In addition, the results have been compared with other state-of-the-art methods, such as those presented in [14]. In this paper, they proposed different approaches to glaucoma grading: Conventional Multi-class (CM), Static Prototypes (SP) and Dynamic Prototypes (DP). The latter is the one that provided the best performance, and that is why in Table 3 the proposed model (PM) is compared with DP. Our approach improves most of the figures of merit; however, it is neither a direct nor a completely objective comparison because the databases are different. Moreover, it would be of valuable interest to perform an external validation with other databases to evaluate the performance of the proposed method; however, there are no public databases of OCT images containing glaucoma annotations.

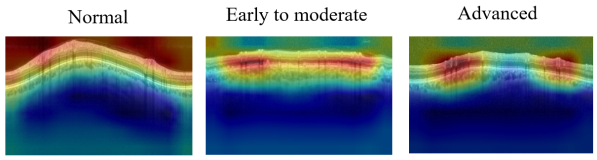

Fig. 4. Class activation maps (CAMs). Heat maps extracted from the proposed architecture at the output from the attention module.

C. Qualitative results

The qualitative class activation maps (CAMs) shown in Fig. 4 represent the activation provided by the attention module. It can be seen that for class differentiation, the activation is mainly in the RNFL structure, which is directly in line with the clinicians' opinion, who perform the classification according to the thickness of the RNLF. The tendency of activations varies by class. In the normal class, the model focuses its activation on the whole retina, while in the early to moderate class, it is on the RNLF. Finally, activation is concentrated in more specific areas in advanced glaucoma.

V. CONCLUSION

In this study, a new methodology has been proposed to improve the classification of glaucoma progression by adding a regression constraint which will affect the optimisation of the parameters in the training process. By combining the loss classification and weighted regression loss functions, the proposed model has demonstrated to improve the glaucoma grading performance. In addition, the model can predict the MD index with a value of 0.1155 in the losses, which provides a great added value to the body of knowledge, as glaucoma can be graded without incurring additional visual field tests.

TABLE II
TEST RESULTS OF BASELINE (BL) AND THE PROPOSED MODEL (PM).

	Normal		Early to moderate		Advanced		MicroAvg		MacroAvg	
	BL	PM	BL	PM	BL	PM	BL	PM	BL	PM
SN	**1.0000**	**1.0000**	0.5952	**0.7143**	0.8333	**0.8750**	0.8457	**0.8889**	0.8095	**0.8631**
SP	0.9667	**1.0000**	0.9417	**0.9500**	0.8684	**0.8947**	0.9228	**0.9444**	0.9256	**0.9482**
PP	0.9600	**1.0000**	0.7812	**0.8333**	0.7273	**0.7778**	0.8457	**0.8889**	0.8228	**0.8704**
NP	**1.0000**	**1.0000**	0.8692	**0.9048**	0.9252	**0.9444**	0.9228	**0.9444**	0.9315	**0.9497**
FS	0.9796	**1.0000**	0.6757	**0.7692**	0.7767	**0.8235**	0.8457	**0.8889**	0.8107	**0.8643**
AC	0.9815	**1.0000**	0.8519	**0.8889**	0.8580	**0.8889**	0.8971	**0.9259**	0.8971	**0.9259**

TABLE III
TEST RESULTS OF THE PROPOSED MODEL (PM) AND THE DYNAMIC PROTOTYPES (DP) MODEL PROPOSED IN [14].

	Normal		Early to moderate		Advanced		MicroAvg		MacroAvg	
	DP	PM	DP	PM	DP	PM	DP	PM	DP	PM
SN	0.9333	**1.0000**	0.6667	**0.7143**	0.8333	**0.8750**	0.8182	**0.8889**	0.8112	**0.8631**
SP	**1.0000**	**1.0000**	0.9048	**0.9500**	0.8519	**0.8947**	0.9091	**0.9444**	0.9189	**0.9482**
PP	**1.0000**	**1.0000**	0.8000	**0.8333**	0.5556	**0.7778**	0.8182	**0.8889**	0.7857	**0.8704**
NP	0.9474	**1.0000**	0.8210	**0.9048**	**0.9583**	0.9444	0.9091	**0.9444**	0.9106	**0.9497**
FS	0.9655	**1.0000**	0.7273	**0.7692**	0.6667	**0.8235**	0.8788	**0.8889**	**0.8788**	0.8643
AC	0.9697	**1.0000**	0.8182	**0.8889**	0.8485	**0.8889**	0.8788	**0.9259**	0.8788	**0.9259**

REFERENCES

[1] K. A. Thakoor, X. Li, E. Tsamis, P. Sajda, and D. C. Hood, "Enhancing the accuracy of glaucoma detection from OCT probability maps using convolutional neural networks," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 2036–2040.

[2] A. R. Ran, C. Y. Cheung, X. Wang, H. Chen, L.-y. Luo, P. P. Chan, M. O. Wong, R. T. Chang, S. S. Mannil, A. L. Young *et al.*, "Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis," *The Lancet Digital Health*, vol. 1, no. 4, pp. e172–e182, 2019.

[3] M. N. Bajwa, M. I. Malik, S. A. Siddiqui, A. Dengel, F. Shafait, W. Neumeier, and S. Ahmed, "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning," *BMC medical informatics and decision making*, vol. 19, no. 1, p. 136, 2019.

[4] S. Phan, S. Satoh, Y. Yoda, K. Kashiwagi, T. Oshika *et al.*, "Evaluation of deep convolutional neural networks for glaucoma detection," *Japanese journal of ophthalmology*, vol. 63, no. 3, pp. 276–283, 2019.

[5] N. Shibata, M. Tanito, K. Mitsuhashi, Y. Fujino, M. Matsuura, H. Murata, and R. Asaoka, "Development of a deep residual learning algorithm to screen for glaucoma from fundus photography," *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018.

[6] A. Cerentinia, D. Welfera, M. C. d'Ornellasa, C. J. P. Haygertb, and G. N. Dottob, "Automatic identification of glaucoma using deep learning methods," in *MEDINFO 2017: Precision Healthcare Through Informatics: Proceedings of the 16th World Congress on Medical and Health Informatics*, vol. 245. IOS Press, 2018, p. 318.

[7] Y. Zhen, L. Wang, H. Liu, J. Zhang, and J. Pu, "Performance assessment of the deep learning technologies in grading glaucoma severity," *arXiv preprint arXiv:1810.13376*, 2018.

[8] A. Serener and S. Serte, "Transfer learning for early and advanced glaucoma detection with convolutional neural networks," in *2019 Medical Technologies Congress (TIPTEKNO)*. IEEE, 2019, pp. 1–4.

[9] A. J. Tatham and F. A. Medeiros, "Detecting structural progression in glaucoma with optical coherence tomography," *Ophthalmology*, vol. 124, no. 12, pp. S57–S65, 2017.

[10] G. García, A. Colomer, R. Verdú-Monedero, and V. Dolz, Jose Naranjo, "A self-training framework for glaucoma grading in OCT b-scans," *29th European Signal Processing Conference (EUSIPCO)*, pp. 1281–1285, 2021.

[11] A. e. a. El-Naby, "Correlation of retinal nerve fiber layer thickness and perimetric changes in primary open-angle glaucoma," *Journal of the Egyptian Ophthalmological Society*, vol. 111, 2018.

[12] G. García, A. Colomer, and V. Naranjo, "Glaucoma detection from raw SD-OCT volumes: A novel approach focused on spatial dependencies," *Computer Methods and Programs in Biomedicine*, 2021.

[13] J. M. Ahn, S. Kim, K.-S. Ahn, S.-H. Cho, K. B. Lee, and U. S. Kim, "A deep learning model for the detection of both advanced and early glaucoma using fundus photography," *PloS one*, vol. 13, no. 11, p. e0207982, 2018.

[14] G. García, R. del Amor, A. Colomer, R. Verdú-Monedero, M.-S. Juan, and V. Naranjo, "Circumpapillary OCT - focused hybrid learning for glaucoma grading using tailored prototypical neural networks," *Artificial Intelligence in Medicine*, vol. 118, 2021.

[15] R. Asaoka, H. Murata, K. Hirasawa, Y. Fujino, M. Matsuura, and A. Miki, "Using deep learning and transfer learning to accurately diagnose early-onset glaucoma from macular optical coherence tomography images," *American Journal of Ophthalmology*, vol. 198, pp. 136–145, 2019.

[16] H. Muhammad, T. Fuchs, N. De Cuir, C. De Moraes, D. Blumberg, J. Liebmann, and et al., "Hybrid deep learning on single wide-field optical coherence tomography scans accurately classifies glaucoma suspects," *J Glaucoma*, vol. 26, p. 1086–94, 2017.

[17] J. Lee, Y. Kim, K. Park, and J. Jeoung, "Diagnosing glaucoma with spectral-domain optical coherence tomography using deep learning classifier," *J Glaucoma*, vol. 29, p. 287–94, 2020.

[18] N. Akter, J. Fletcher, S. Perry, and et al., "Glaucoma diagnosis using multi-feature analysis and a deep learning technique," *Sci Rep*, vol. 12, p. 8064, 2022.

[19] R. P. Mills, D. L. Budenz, P. P. Lee, R. J. Noecker, J. G. Walt, S. J. Siegartel, Lisa R.and Evans, and J. J. Doyle, "Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease," *American Journal of Ophthalmology*, vol. 141, pp. 24–30, 2006.

[20] T. T. Hoang, A. Van Bui, V. Nguyen, P. J. McCluskey, J. R. Grigg, and S. E. Skalicky, "Comparison of perimetric glaucoma staging systems in asians with primary glaucoma," *Eye*, pp. 1–6, 2020.

[21] J. Flammer, "The concept of visual field indices," *Graefe's archive for clinical and experimental ophthalmology*, vol. 224, no. 5, pp. 389–392, 1986.

[22] G. García, R. del Amor, A. Colomer, and V. Naranjo, "Glaucoma detection from raw circumapillary OCT images using fully convolutional neural networks," *arXiv preprint arXiv:2006.00027*, 2020.

[23] J. Deng, W. Dong, R. Socher, L.-J. Li, L. Kai, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.

[24] G. S. Jie Hu, Li Shen, "Squeeze-and-excitation networks," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7132–7141, 2018.