Glaucoma Grading Using Fundus Images

Mackele Lourrane Jurema da Silva[1(✉)], Marcos Melo Ferreira[1,2], Geraldo Braz Junior[1], João Dallyson Sousa de Almeida[1], and Arthur Guilherme Santos Fernandes[1]

[1] Federal University of Maranhao, Sao Luis, MA, Brazil
{mackele.ljs,marcos.ferreira}@discente.ufma.br,
{geraldo,jdalyson,arthurgsf}@nca.ufma.br
[2] Federal Institute of Maranhao, Sao Luis, MA, Brazil

Abstract. Glaucoma is a chronic, progressive eye disease caused by gradual damage to the optic nerve and is considered the major cause of irreversible visual damage. Because it is impossible to reverse the loss of vision caused by the disease, early detection is essential that interventions can be carried out in the early stages of the disease to stop its progression. Fundus imaging is one of the main methods used to diagnose the disease, making it possible to assess the cup-to-disc ratio by a specialist. In this work, we propose a method based on deep learning, which uses fundus images to help detect the disease in its early stages. In this way, the proposed method can have clinical use and be used to develop tools for classifying more serious disease cases. As a best result, the proposed method achieved a kappa value of 0.83.

Keywords: Glaucoma · Diagnosis · Deep Learning

1 Introduction

Glaucoma is an ophthalmological disease identified as the second leading cause of blindness and the main cause of irreversible visual damage, [8,28]. It is estimated that the total number of people with glaucoma worldwide is approximately 80 million [24], with 1.5 million cases registered in Brazil, according to the Brazilian Council of Ophthalmology (CBO) [7]. However, the number of registered cases does not reflect the total number of people with the disease due to the difficulty of early diagnosis since there are no symptoms in the early stages [20]. According to a survey [21] carried out by the Brazilian Society of Glaucoma (SBG), 41% of the people interviewed do not know what glaucoma is, and 39% are unaware of the probability of blindness. These data suggest that there is little concern with

This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financing Code 001, Maranhão Research Support Foundation (FAPEMA), National Council for Scientific and Technological Development (CNPq) and Brazilian Company of Hospital Services (Ebserh) Brazil (Proc. 409593/2021-4).

A. Cunha et al. (Eds.): MobiHealth 2023, LNICST 578, pp. 163–172, 2024.
https://doi.org/10.1007/978-3-031-60665-6_12

disease prevention and that most cases will only be detected when symptoms are present. This indicates an advanced stage of the disease in which visual damage and blindness cannot be avoided.

The disease is a disorder in which excessive intraocular pressure (IOP) causes damage to the optic nerve. This damage leads to a progressive loss of peripheral vision, which can progress to total vision loss [29]. CBO estimates point out that in Brazil, there are 1.6 million blind people, with glaucoma being one of the three main causes of blindness [6]. Although irreversible, the loss can be avoided if medical procedures are performed in the early stages of the disease. According to the World Health Organization (WHO) [28], only 11% of people who received timely diagnosis and treatment reported having acquired moderate or severe damage resulting from the most severe forms of the disease.

The necessary interventions to avoid visual damage caused by increased intraocular pressure must be carried out in the early stages of the disease. Glaucoma can be detected using medical images such as fundus images and Optical Coherence Tomography (OCTs) [1]. Analyzing fundus images makes it possible to detect changes in clinical parameters that indicate the disease, such as increased excavation of the optic nerve [22]. However, the analysis of a large number of exams takes a lot of time, being exhausting for the specialist [13]. In this context, automatic methods that can help specialists detect glaucoma using medical images may have great potential for clinical use. Approaches based on deep learning have shown promising results in image classification tasks [14,27].

This study presents a method based on deep learning for automatically classifying the glaucoma stage using fundus images from different datasets. The main objective of this work is to present a method that can be used to develop tools that can facilitate the early diagnosis of the disease. In this work, the performance of an optimized Convolutional Neural Network with an architecture that combines convolutional blocks from two SOTA models [16,23], which received fundus images as input, was investigated.

The main contributions of this work are a) a CNN architecture optimized for grading the glaucoma stage based on medical images (fundus images); b) an easily configurable method and expandable convolutional neural network architecture that combines Dense and Inception blocks for glaucoma grading.

The remaining sections are organized as follows. Section 2 presents some related works. Section 3 presents the proposed method. Section 4 presents the results and evaluation of our method. Section 5 concludes this paper.

2 Related Work

Several methods have been proposed for the detection of glaucoma using fundus images. Most methods are based on deep learning, and more recently, *Vision Transformers* [9]. Another characteristic present in most of the proposed methods is the capture of the optic disc region for classification since the main biological marker for the detection of glaucoma is the excavation of the optic disc [5].

An architecture that combines features extracted from fundus images and images containing only the optic disc region was presented in [12]. The work aims to classify the images from the GAMMA dataset into early, progressive, and non-glaucoma. Feature extraction was performed using a network with two extraction levels, with the Resnet34 CNN as the backbone.

The work developed in [17] aims to segment and classify fundus images. The model was named EffUnet-SpaGen and contains two stages, a U-shaped convolutional neural network where the method segments the optical disc and the cup, EffUNet. The model also presents a spatial generative algorithm, SpaGen. The model outputs 99.7% AUROC in the Origa dataset.

In [10], a method was proposed that performs the decomposition of the optic disc region present in fundus images, the BEMD, (*Bi-dimensional Empirical Mode Decomposition*). The research presents a VGG 19 with multiple inputs, thus being a multilevel network. The original image will feed the network, and from the image, the BEMD will be made, and the decompositions will serve as input for the other network inputs; after the convolution stage, they have concatenated all the outputs. Finally, the SVM classifier classifies in glaucoma or not glaucoma. The proposed method, trained on the Refuge dataset, presents 99.0% and 96.9% accuracy when tested on the Refuge and Origa-light datasets, respectively.

Li et al. [18] evaluates the performance of four CNN Resnet models (Resnet34/50/101/152) for glaucoma stage classification on the GAMMA test set. The best result was a kappa value of 0.699, using the ResNet34 model. In [4], supervised contrastive learning was used to train a ResNet model to perform feature extraction from fundus images. As the best result, a kappa value of 0.728 and an accuracy of 0.830 were achieved.

Tian et al. [25] present a GC-Net to classify images from the GAMMA dataset into non-glaucoma, early-stage glaucoma, and progressive glaucoma, using as input optic disc regions. A pre-trained CNN forms the proposed architecture used as a feature extractor and an attention module formed by a global attention block and a class attention block, achieving as best result a kappa value of 0.894.

In [11], a comparison was made to evaluate the classification capacity between a CNN (ResNet50 model) and a DeiT network (*Distilling Vision Transformers*) [26]. The two networks were trained on a private dataset and were tested on five public datasets, with the Deit network achieving superior results compared to CNN ResNEt50.

3 Materials and Method

The proposed method uses a convolutional network based on the DenseNet [16] and Inception [23] models for predicting stages in early or progressive glaucoma (intermediate and advanced glaucoma). The proposed method uses fundus images for model training and evaluation. In this work, we optimized the model's hyperparameters, evaluating the number of dense blocks and the number of layers per block and searching for the best architecture for prediction. The method used for developing this research comprises five main steps: Image acquisition, image preprocessing, model construction, model tuning, and model evaluation.

The steps of the proposed method, along with some keywords that summarize each step, are shown in Fig. 1.

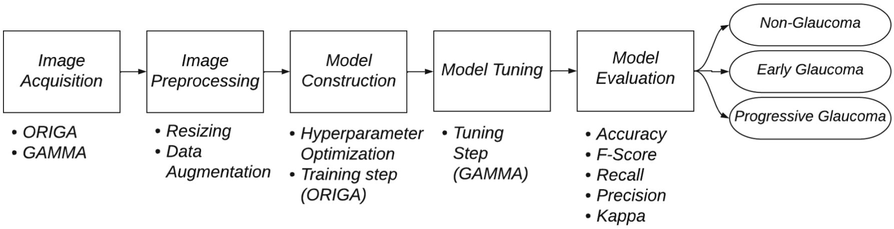

Fig. 1. Steps of the proposed method.

3.1 Image Acquisition

Origa [30] and Gamma [29] image datasets, both publicly available, were used in this work. The Origa dataset comprises 650 fundus images, divided into two classes (482 non-glaucoma samples and 168 glaucoma samples. The gamma dataset comprises 100 fundus images, divided into three classes (50 non-glaucoma samples, 26 early glaucoma samples, and 24 progressive glaucoma samples). In this work, the training set with available labels was used. The GAMMA test set only has images, which prevents the use of supervised learning techniques and restricts the evaluation of results to challenge participants only.

3.2 Image Preprocessing and Data Augmentation

The pre-processing was carried out with the purpose of preparing the data that will be used in the following steps. Each sample was pre-processed, where the image pixels' normalization was performed, color channel standardization, and resizing to the 224×224 resolution due to hardware limitations.

Data Augmentation was used to improve the model's performance in which it was applied. Therefore, synthetic images were created to reduce the imbalance between existing classes. We used the Albumentations [3] library to apply two image transformations: GaussNoise and RandomGamma. Data augmentation was only applied to the Origa dataset. At the end of the data augmentation process, 336 synthetic images were generated, totaling 504 glaucoma samples.

3.3 Model Construction

In this step, the search for the best hyperparameters of a neural network that has its architecture based on the DenseNet and Inception networks was carried out. The proposed architecture has dense blocks, each containing convolutional blocks similar to those of the Inception network. The architecture is shown in Fig. 2. During the optimization process, a search was carried out by the number of dense blocks and the number of layers (Inception blocks) per block to find the best architecture to achieve the proposed objective. The complete hyperparameter search space is presented in Table 1. Models were trained for classifying fundus images in glaucoma and non-glaucoma using samples from the Origa dataset, which was divided into 70% for training, 10% for validation, and 20% for testing.

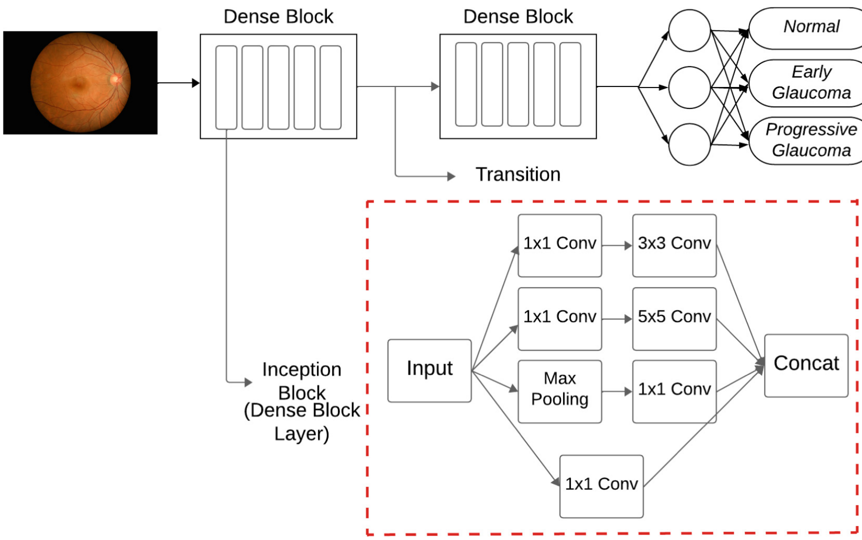

Fig. 2. Network Architecture.

Table 1. Hyperparameter Search Space.

Parameter	Search Space	Distribution
Number of Dense Blocks	[2, 3]	Categorical
Layer per Block	[2, 3]	Categorical
Growth Rate	[16, 32]	Categorical
Compress Factor	[0.5, 1]	Categorical
Dropout	[0.2, 0.3, 0.4]	Categorical

3.4 Model Tuning

In the previous step, models were trained using the samples from the ORIGA dataset [31], balanced after the data augmentation process. For the purpose of carrying out transfer learning, the model that achieved the best results in the classification of images in glaucoma and non-glaucoma was chosen to be adjusted for the classification of samples from the GAMMA dataset, which has samples from three classes, non-glaucoma, early-stage glaucoma, and progressive-stage glaucoma. At this stage, the chosen model has its classifier replaced by one adjusted for the three-class classification task. This classifier consists of two fully connected layers followed by one each of Global Average Pooling and a last layer composed of three neurons with softmax activation. A new training of the model was carried out, using 90% of the samples from the GAMMA dataset, with 10% of the samples being separated to evaluate the models.

3.5 Model Evaluate

After training, the models were evaluated using the split test, formed by 10% of fundus image samples from the GAMMA dataset. Models were evaluated in terms of precision (1), recall (2), f1-score (3), accuracy (4) and kappa score (5).

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$S = \frac{TP}{TP + FN} \tag{2}$$

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{3}$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$kappa = k = \frac{p_0 - p_e}{1 - p_e} \tag{5}$$

in which TP is the number of True Positive predictions; TN is the number of True Negative predictions; FP is the number of False Positive predictions; FN is the number of False Negative predictions; p_0 is accuracy and p_e is the sum of the products of the actual and predicted numbers corresponding to each category, divided by the square of the total number of samples.

4 Results and Discussion

This work proposes a method for grading the glaucoma stage using fundus images and transfer learning. In the model-building stage, the hyperparameters and architecture optimization of CNNs models for classifying fundus images in glaucoma and non-glaucoma was performed using samples from the ORIGA dataset. In the tuning step, models that achieved the highest metrics in the previous step were fine-tuned for classifying fundus images in early, progressive, and non-glaucoma using samples from the GAMMA dataset. The Hyperopt [2] framework was used to perform the hyperparameter search. The results achieved using CNN model DenseNet121 and the best results achieved using the optimized architecture in the split test are presented in Table 2. The confusion matrix is presented in Fig. 3.

Table 2. Best results from the methodology proposed in the GAMMA dataset.

Model	Precision	Recall	F1-Score	Accuracy	Kappa
Dense121	0.65	0.70	0.64	0.70	0.51
Proposed Method	0.90	0.90	0.89	0.90	0.83

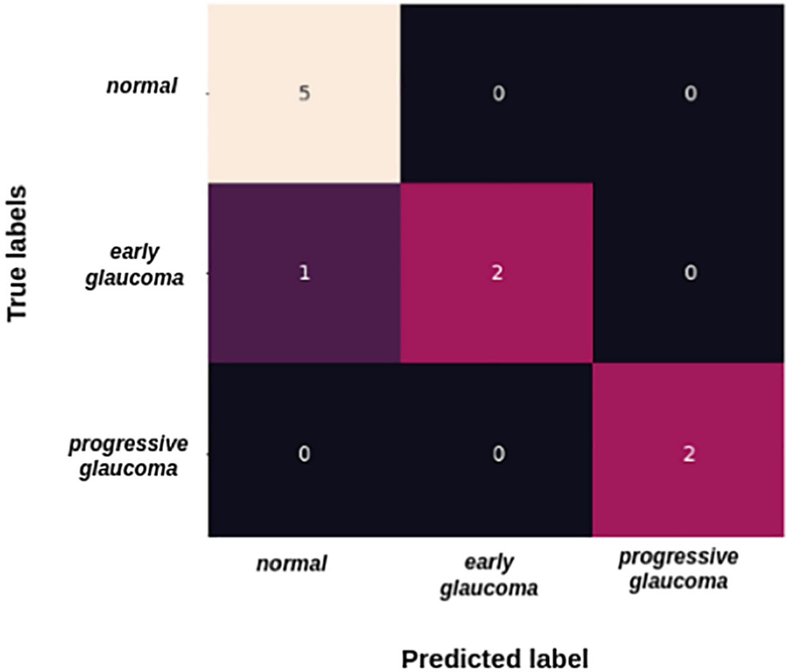

Fig. 3. Confusion matrix, glaucoma grading stage task using the optimized architecture.

The results show that the pre-trained model built with dense and inception blocks achieved promising results in classifying the stage of glaucoma using fundus images. Figure 4 presents an activation map created to visualize which regions of the images were decisive for the predictions made by the model. Through them, it is possible to see that the optic disc region was decisive for classification, which was expected due to the cup-to-disc ratio being a biological marker for detecting and classifying the glaucoma stage [25].

As the GAMMA dataset was released recently, few related works used the dataset. One of the factors that makes comparison difficult is the unavailability of the image labels that form the preliminary test dataset, which are only available to challenge participants. Furthermore, some of the studies aim to segment the disc and optical cup [15,19], without presenting results on the glaucoma grading stage.

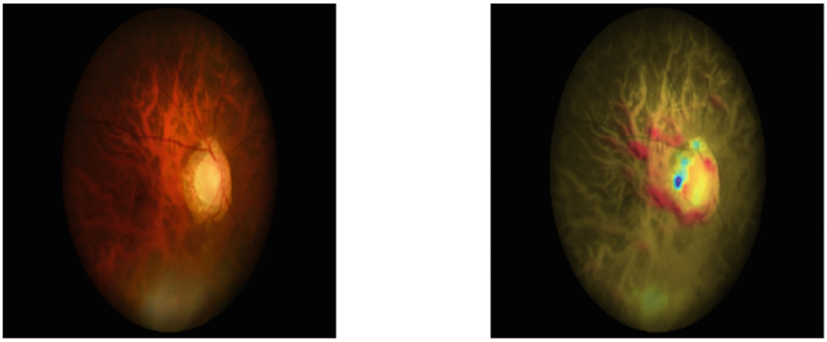

Fig. 4. Class Activation Map.

Despite the results achieved by the proposed method, it is important to highlight that it needs to be evaluated on other datasets, which have a greater number of images, making it possible to evaluate the generalization capacity of the models better. However, this is difficult because no other datasets with fundus images labeled early, progressive, and non-glaucoma are publicly available.

5 Conclusion

In this paper, we present a method in which it was proposed to create a specialized model capable of distinguishing between eyes with glaucoma and healthy eyes, optimizing the hyperparameters of the neural network during this phase. Furthermore, the developed model was subsequently used to grade the glaucoma stage using fundus image samples in non-glaucoma, early glaucoma, and advanced glaucoma.

The results show that the use of a network that combines dense and inception blocks achieved good results, enabling not only the detection of glaucoma but also the classification of stages of the disease, which allows more severe cases to be identified more quickly, and that interventions be made to stop the progress of the disease. Furthermore, it allows cases of early-stage glaucoma to be identified to begin treatment, avoiding permanent vision impairment.

This work evaluated a network that uses convolutional blocks from other CNNs widely used in image classification and segmentation tasks. To find an optimal architecture to achieve the proposed objective, a search for hyperparameters was carried out, with the main purpose of finding the number of dense blocks and layers per block to build the classification model. The results show that using dense blocks formed by Inception blocks increases the classification capacity of the models. However, since there were few images for training, only models with few blocks achieved good results. Deeper models with many blocks performed poorly, with gradient disappearance as the likely cause of overfitting. Despite results close to related works, new tests need to be carried out to analyze the robustness and generalization capacity of the models.

In future works, we intend to evaluate the use of multilevel architectures requiring optical disk capture, which would be used as input for a second level of feature extraction. It is also necessary to evaluate other convolutional blocks, such as VGG networks, which could be evaluated as layers of dense blocks.

References

1. An, G., et al.: Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images. J. Healthc. Eng. **2019** (2019)
2. Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., Cox, D.D.: Hyperopt: a python library for model selection and hyperparameter optimization. Comput. Sci. Discov. **8**(1), 014008 (2015). http://stacks.iop.org/1749-4699/8/i=1/a=014008
3. Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A.: Albumentations: fast and flexible image augmentations. Information **11**(2) (2020). https://doi.org/10.3390/info11020125, https://www.mdpi.com/2078-2489/11/2/125
4. Cai, Z., Lin, L., He, H., Tang, X.: COROLLA: an efficient multi-modality fusion framework with supervised contrastive learning for glaucoma grading. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI), pp. 1–4. IEEE (2022)
5. Camara, J., Neto, A., Pires, I.M., Villasana, M.V., Zdravevski, E., Cunha, A.: Literature review on artificial intelligence methods for glaucoma screening, segmentation, and classification. J. Imaging **8**(2), 19 (2022)
6. CBO: A importância de combater e prevenir a cegueira (2020). https://www.vejabem.org/noticia/a-importancia-de-combater-e-prevenir-a-cegueira1657730359
7. CBO: Casos de glaucoma aumentaram 26% no primeiro trimestre de 2022 (2022)
8. Coan, L., et al.: Automatic detection of glaucoma via fundus imaging and artificial intelligence: a review. Surv. Ophthalmol. (2022)
9. Dosovitskiy, A., et al.: An image is worth 16×16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
10. Elmoufidi, A., Skouta, A., Jai-Andaloussi, S., Ouchetto, O.: CNN with multiple inputs for automatic glaucoma assessment using fundus images. Int. J. Image Graph. 2350012 (2022)
11. Fan, R., et al.: Detecting glaucoma from fundus photographs using deep learning without convolutions: Transformer for improved generalization. Ophthalmol. Sci. **3**(1), 100233 (2023)
12. Fang, H., Shang, F., Fu, H., Li, F., Zhang, X., Xu, Y.: Multi-modality images analysis: a baseline for glaucoma grading via deep learning. In: Fu, H., Garvin, M.K., MacGillivray, T., Xu, Y., Zheng, Y. (eds.) OMIA 2021. LNCS, vol. 12970, pp. 139–147. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-87000-3_15
13. Ferreira, M.M., Esteve, G.P., Junior, G.B., de Almeida, J.D.S., de Paiva, A.C., Veras, R.: Multilevel CNN for angle closure glaucoma detection using AS-OCT images. In: 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), pp. 105–110. IEEE (2020)
14. Gonçalves, C., et al.: Computer vision in automatic visceral leishmaniasis diagnosis: a survey. IEEE Lat. Am. Trans. **21**(2), 310–319 (2022)
15. He, H., Lin, L., Cai, Z., Cheng, P., Tang, X.: JOINEDTrans: prior guided multi-task transformer for joint optic disc/cup segmentation and fovea detection. arXiv preprint arXiv:2305.11504 (2023)

16. Huang, G., Liu, Z., Weinberger, K.Q.: Densely connected convolutional networks. CoRR abs/1608.06993 (2016). http://arxiv.org/abs/1608.06993
17. Krishna, A., et al.: EffUnet-SpaGen: an efficient and spatial generative approach to glaucoma detection. J. Imaging **7**(6), 92 (2021)
18. Li, Y., et al.: Multimodal information fusion for glaucoma and diabetic retinopathy classification. In: Antony, B., Fu, H., Lee, C.S., MacGillivray, T., Xu, Y., Zheng, Y. (eds.) OMIA 2022. LNCS, vol. 13576, pp. 53–62. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16525-2_6
19. Li, Z., Lu, S., Li, H., Liu, H., Wang, N.: A pyramid spatial attention network for fovea localization. In: 2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC), pp. 1–5. IEEE (2022)
20. Sarhan, A., Rokne, J., Alhajj, R.: Glaucoma detection using image processing techniques: a literature review. Comput. Med. Imaging Graph. 101657 (2019)
21. SBG: Pesquisa de opinião pública sobre glaucoma (2020)
22. Shanmugam, P., Raja, J., Pitchai, R.: An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier. Appl. Soft Comput. **109**, 107512 (2021)
23. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–9 (2015)
24. 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 **121**(11), 2081–2090 (2014)
25. Tian, H., Lu, S., Sun, Y., Li, H.: GC-net: Global and class attention blocks for automated glaucoma classification. In: 2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA), pp. 498–503 (2022). https://doi.org/10.1109/ICIEA54703.2022.10005946
26. Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., Jégou, H.: Training data-efficient image transformers & distillation through attention. In: International Conference on Machine Learning, pp. 10347–10357. PMLR (2021)
27. Valencia-Moreno, J.M., González-Fraga, J.Á., Febles-Rodríguez, J.P., Gutierrez-Lopez, E.: Review of intelligent algorithms for breast cancer control: a Latin America perspective. IEEE Lat. Am. Trans. **21**(2), 226–241 (2022)
28. WHO: World report on vision (2019). www.who.int
29. Wu, J., et al.: Gamma challenge: glaucoma grading from multi-modality images. arXiv preprint arXiv:2202.06511 (2022)
30. Zhang, Z., et al.: ORIGA(-light): An online retinal fundus image database for glaucoma analysis and research, vol. 2010, pp. 3065–3068 (2010). https://doi.org/10.1109/IEMBS.2010.5626137
31. Zhang, Z., et al.: ORIGA-light: an online retinal fundus image database for glaucoma analysis and research. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, pp. 3065–3068. IEEE (2010)