

dx.doi.org/10.17488/RMIB.42.3.2

E-LOCATION ID: 1188

Image-based Glaucoma Classification Using Fundus Images and Deep Learning

Clasificación de Glaucoma Basada en Imágenes de Fondo de Ojo y Aprendizaje Profundo

H. J. Sandoval-Cuellar¹, G. Alfonso-Francia¹, M. A. Vázquez-Membrillo², J. M. Ramos-Arreguín¹, S. Tovar-Arriaga¹

¹Universidad Autónoma de Querétaro

²Instituto Mexicano de Oftalmología, IAP

ABSTRACT

Glaucoma is an eye disease that gradually affects the optic nerve. Intravascular high pressure can be controlled to prevent total vision loss, but early glaucoma detection is crucial. The optic disc has been a notable landmark for finding abnormalities in the retina. The rapid development of computer vision techniques has made it possible to analyze eye conditions from images enabling to help a specialist to make a diagnosis using a technique that is non-invasive in its initial stage through fundus images. We propose a methodology glaucoma detection using deep learning. A convolutional neural network (CNN) is trained to extract multiple features, to classify fundus images. The accuracy, sensitivity, and the area under the curve obtained using the ORIGA database are 93.22%, 94.14%, and 93.98%. The use of the algorithm for the automatic region of interest detection in conjunction with our CNN structure considerably increases the glaucoma detecting accuracy in the ORIGA database.

KEYWORDS: Deep Learning, Glaucoma diagnosis, Image-based classification, Convolutional Neural Network

RESUMEN

El glaucoma es una enfermedad que afecta gradualmente al nervio óptico. La presión intravascular se puede controlar para prevenir la pérdida de visión, por lo que la detección temprana del glaucoma es crucial. El disco óptico ha sido un punto de referencia importante para encontrar anomalías en la retina. El rápido desarrollo de técnicas de visión por computadora ha hecho posible el analizar las condiciones del ojo ayudando al especialista a realizar un diagnóstico utilizando una técnica no invasiva en su estadio inicial en imágenes de fondo de ojo. En este artículo, se propone una arquitectura para la detección de glaucoma utilizando aprendizaje profundo. Una red neuronal convolucional (RNC) es entrenada para extraer múltiples características, para clasificar imágenes de fondo de ojo. La exactitud, sensibilidad, y el área bajo la curva obtenidos en la base de datos ORIGA son 93.22%, 94.14% y 93.98%. El uso del algoritmo para la detección automática de la región de interés, incrementa considerablemente la exactitud de detección de glaucoma en la base de datos ORIGA.

PALABRAS CLAVE: Aprendizaje profundo, Diagnóstico del glaucoma, Clasificación de imágenes, Red neuronal convolucional

Corresponding author

TO: Saúl Tovar Arriaga

INSTITUTION: Universidad Autónoma de Querétaro

ADDRESS: Cerro de las Campanas S/N,

Cto. Universitario, C. P. 76110, Las Campanas,

Santiago de Querétaro, México

E-MAIL: saul.tovar@uaq.mx

Received:

8 June 2021

Accepted:

11 September 2021

INTRODUCTION

The World Health Organization (WHO) states that glaucoma is the second leading cause of blindness in the world, only after cataracts [1]. Since there is no cure for glaucoma [2], it is necessary to diagnose this disease early to delay its development [3].

Glaucoma is characterized by optic nerve damage due to increased degeneration of nerve fibers [4]. Since symptoms appear until the disease is severe, glaucoma is called a silent thief of sight [5]. Typically, aqueous humor drains out of the eye through the trabecular meshwork, but when the passage is obstructed, aqueous humor accumulates. The increase in this fluid will cause pressure to grow and cause ganglion cell damage [6]. However, it was found that the pressure measurement was not specific or sensitive enough to be the only useful indicator for detecting glaucoma because visual impairment would occur without increasing pressure. Therefore, a comprehensive examination should also include the use of images and visual field tests to analyze the retina [7].

Fundus imaging is the process of obtaining a two-dimensional projection of the retinal tissue employing reflected light. Finally, the resulting 2D image contains the image intensity that indicates the reflected light amount. Color fundus imaging is used to detect disease, where the image has R, G, and B bands, and their intensity changes highlight different parts of the retina [8].

Due to their non-invasiveness, funduscopy and optical coherence tomography (OCT) has become the imaging methods of choice for the detection and evaluation of glaucoma [9]. Nevertheless, due to its cost, OCT is not readily available.

There are different types of glaucoma, such as open-angle, closed-angle, secondary, normal-tension, pigmentary, and congenital glaucoma [10]. Some types require surgical treatment [11]. The optic disc and cup,

peripapillary atrophy, and retinal nerve fiber layer [12] are four structures that are considered essential for detecting glaucoma. Family history has also been shown to be genetically related to the onset of the disease [13].

In most cases, treatment in the early stage of the disease can prevent total vision loss in glaucoma patients [14], so a system that can help ophthalmologists make a diagnosis could increase the chances of saving people's vision. However, designing a system that provides reliable tests for glaucoma diagnosis is a complicated and stimulating task in clinical practice [15].

Figure 1 shows an image of the fundus in a healthy condition where the most important parts can be seen, such as the optic nerve, which is where the study of this work will be focused. The outer part of the optic nerve is called the optic disk (OD), and the smaller blurry inner circle is called the optic cup (OC). In the OD area, we can see the main arteries and veins, whereas the veins have a darker color than the arteries. Veins usually are larger in caliber than arteries, having an average AVR ratio of 2:3 [17].

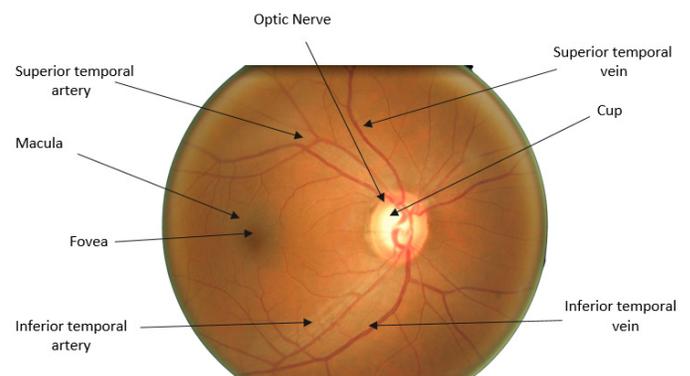


FIGURE 1. Image of the optic nerve [16].

The region of interest (ROI) usually is less than 11% of the retina's fundus image's total size. By decreasing the size of the image with the ROI's detection, a reduction in the computational resources used is possible

^[18]. Both optic disc and cup segmentation are essential components of optic nerve segmentation and, together, form the basis of a glaucoma evaluation. In ^[19], a reference data set for evaluating the cup segmentation method was published. Nevertheless, it does not provide free access to the general public.

There are relatively few public data sets for glaucoma evaluation compared to available data sets for diabetic retinopathy ^[20] and vascular segmentation ^[21]. The ORIGA database contains 650 retinal images labeled by retina specialists from the Singapore Eye Research Institute. An extensive collection of image signs usually taken into account for the diagnosis of glaucoma are annotated. In the Drishti-GS dataset ^[22], all images were collected from the Madurai Aravind Eye Hospital of the hospital visitors with their consent. The selection of glaucoma patients is made based on the clinical findings during the study. The selected subjects are between 40 and 80 years old, and the number of males and females is roughly equal. Patients who choose to undergo routine optometry and who are not glaucoma represent the normal category.

Usually, the fundus image analysis with artificial intelligence is carried out from two approaches, 1) the classification at the image level and 2) the classification at the pixel level. In image-level classification, the learning model is trained with images previously classified by an expert, in this case, a retina specialist. This classification generally contains the disease progression ^[23]. For example, the severity of DR is classified into five grades, and each of these grades is associated with a number. The learning model begins to associate the patterns in the image to their labels. The second approach is the anatomical and lesion segmentation, such as separating the blood vessels from the rest of the retina in order to measure their caliber. For example, in the case of Hypertensive Retinopathy, the vein/artery relationship plays an essential role in the diagnosis of the disease ^[24].

There are several works carried out for the detection of glaucoma. Both images based and pixel-based classification have been used. We will follow an image-based classification approach in this research.

A six-layer CNN architecture was proposed by Chen *et al.* ^[5], where four layers are convolutional, and the final two are fully connected. ORIGA and SCES are used in this study. From the ORIGA database, 99 images were randomly selected for training and the remaining 551 images for the testing, obtaining results of 0.831 for AUC. In a second experiment, 650 images from the ORIGA database are used for training, and 1676 images from the SCES database for the testing, the area under the curve obtained is 0.887.

Acharya *et al.* ^[6] proposed to use a Support Vector Machine for classification and the Gabor transform that will notice the subtle changes in the image's background. The database used was a private database of Kasturba Medical College, Manipal, India, with 510 images. 90% of the images were used for training, while the remaining 10% for testing. The results obtained were 93.10% accuracy, 89.75% sensitivity, and a specificity of 96.20%.

Raghavendra *et al.* ^[1] proposes to perform the automatic recognition of glaucoma utilizing a convolutional neural network of 18 layers. This work consists of a standard CNN, with convolution layers and max-pooling, and a fully connected layer where classification is performed. Initially, 70% of the randomly selected samples are used for training and 30% for testing. 589 healthy images and 837 with glaucoma of a private database were deployed. The process was repeated fifty times with random training and test partitions. The results obtained were 98.13% accuracy, 98% sensitivity, and 98.3% specificity.

Gour *et al.* ^[8] proposes an automatic glaucoma detection system using Support Vector Machine (SVM) for classification. Combine GIST and PHOG to extract features in images. This technique eliminates the need for

image segmentation. Instead, it works with a diagnostic system that makes use of characteristics such as texture and shape to detect the disease. This technique yielded an accuracy of 83.4% using the Drishti-GS1 and High-Resolution Fundus (HRF) databases.

Gheisari *et al.* [25] implement two architectures, the VGG16 and ResNet, concatenating LSTM blocks. To determine the best one, they carry out several experiments varying the number of epochs and learning rate. The best results are achieved with the VGG16 network, achieving 95% sensitivity and 96% specificity.

Gómez-Valverde *et al.* [26] use architectures such as VGG19, ResNet, GoogleNet and Denet Disc. The best results obtained were obtained using VGG19 with transfer learning, obtaining 94.20 of AUC, 87.01 of sensitivity, and 89.01 of specificity.

Pinto *et al.* [27] use 5 databases adding a total of 1707 images. They carry out the experimentation with each of the databases separately, but the best results were obtained by putting together all the available images. They achieve an AUC of 96.05, a specificity of 85.8 and a sensitivity of 93.46 using Xception architecture.

We propose an image-based classification approach using a deep neural network for glaucoma diagnosis in fundus images in the present work. A preprocessing step will be carried out where the region of interest will be extracted, specifically the region where the optic disc is located. Once that region is obtained, a neural network is fed with the cropped images in order to classify if the image has or not glaucoma.

MATERIALS AND METHODS

Preprocessing

This section introduces a method to locate the retina optic disc, which contains the necessary features to diagnose glaucoma. These image areas will be the in-

put to the neural network proposed in this work. Since the images' size is 3072 x 2048 pixels, it is reduced to four times its original size, leaving the images' dimensions in 768 x 512. The input images are converted to grayscale. Therefore, it is reasonable to obtain a higher contrast of the optic disc than the original image. For this, the red and green channels are used, which are the ones that have the most significant impact on the optic disc. To do this, Equation 1 is applied.

$$Img_{gray} = R * 0.9 + G * 0.5 \quad (1)$$

Where Img_{gray} is the grayscale image, and R and G are the corresponding red and green channels. To determine which channels would be used, the histograms per channel of the image were obtained, in which it is determined which of them have the greatest impact on the contrast of the image. Conventional grayscale conversion is obtained using the luminance coefficients in ITU-R BT. 601-7, which is a recommendation that specifies digital video signal encoding methods [28].

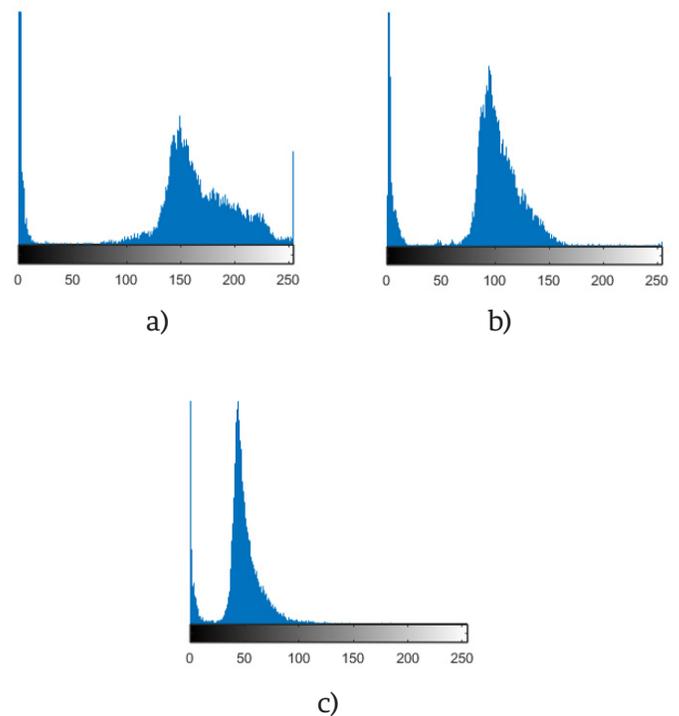


FIGURE 2. The histogram of the red channel is shown in a), the green channel in b), and the blue channel in c).

The coefficients and channels used in our method were obtained by analyzing the histograms in each channel of the images. These histograms are shown in Figure 2.

The results obtained after the transformation are shown in Figure 3.

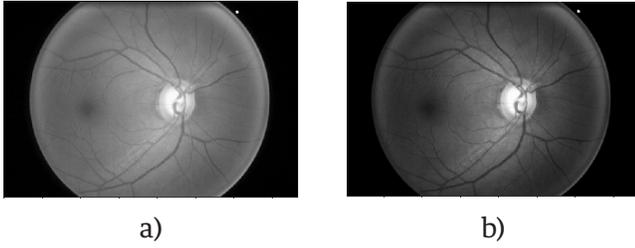


FIGURE 3. The conventional transformation to grayscale is represented by a), and the one proposed by b).

The next step is to scroll a kernel through the image to divide it into different sub-images to determine where the optic disc is located as shown in Figure 4. The optic disc is the most brilliant part of the retina. Excess brightness can be eliminated in parts where the optic disc is not located.

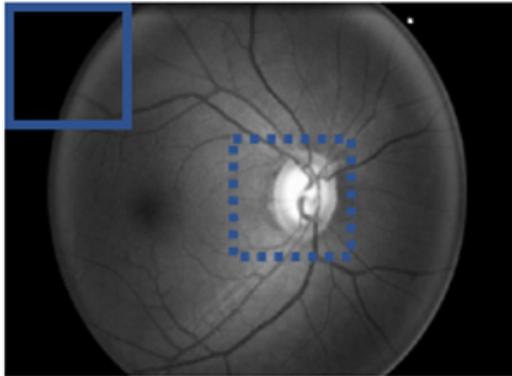


FIGURE 4. Kernel scrolling through original image.

Because the optical disc represents approximately 10% of the original image, the scrolling box's size, or kernel, was chosen to ensure that any sub-images would be selected. The selected size was 192 x 192 pixels, four times smaller than the image width after size reduction.

Equation 2 is applied to calculate the sub-images, depending on the size of the scrolling box and the step used

$$ubImg(x, y) = Img(x + x_0, y + y_0), \quad \forall (x, y) \in R \quad (2)$$

Where *SubImg* is the image inside the kernel, *Img* is the image after resizing, *x* is the kernel's horizontal starting point, *y* the vertical starting point, x_0 the horizontal endpoint, that is, $x + 192$ and y_0 the vertical endpoint, that is, $y + 192$.

To reduce the processing time, the kernel step will be 165 pixels in the horizontal direction and 150 in the vertical direction. With this step, we guarantee that the optical disc is in one of the sub-images at least once. Each time a new sub-image is selected, the average of all the pixels is calculated by applying Equation 3. This average is stored, and once all the sub-images have been analyzed, it is located where the highest average is. In this way, the sub-image where the optic disc is predicted is chosen.

$$promPix = \frac{\sum_{i=1}^n p_i}{n} \quad (3)$$

Where *promPix* is the average of the pixel values, p_i is the current pixel value, and *n* is the total number of pixels. With this new sub-image selected, the pixels' average value is taken, but now of each column and row.

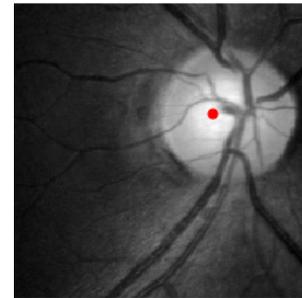


FIGURE 5. A red dot represents the center of the image to be cropped.

This average will give us x and y coordinates. These coordinates represent the new image center that will be cropped to obtain the ROI. An example of this can be found in Figure 5.

The result after cropping an image can be found in Figure 6.



FIGURE 6. Result after cropping the image and obtaining the region of interest.

The entire process explained above for the ROI detection is shown in the pseudo-code in Algorithm 1.

Algorithm: Region of interest

```

Input: number of images Images
Output: Region of interest ROI
1: while Images is not empty
2:   Images ← Images/4
3:   Images ← (RedChannel * 0.9) + (GreenChannel * 0.5)
4:   kernel = 192x192
5:   if kernel ∈ Images then
6:     average = averageKernel
7:     if average > previousAverage
8:       previousAverage → average
9:       Images → kernel
10:    end
11:  end
12: valMax = where (valMaxPix = Images)
13: x = mean(posValMaxX)
14: y = mean(posValMaxY)
15: ROI = center(x, y)
16: end

```

The preprocessing described in Algorithm 1 is applied to the 650 images of the ORIGA database [16] to obtain the ROI. With the region of interest located, we pro-

ceed to perform normalization to the images. What is sought is that the values are within a smaller range, since the CNN's do not perform well when the input numerical attributes have a very large range. The normalization used (Equation 4) was the min-max [23].

$$valNorm = \frac{val - min_{img}}{max_{img} - min_{img}} \quad (4)$$

Where *valNorm* is the normalized value, *val* is the current value, min_{img} and max_{img} are the minimum and maximum image values. The values before normalization range from 0 to 255. After normalization, they will be in a range from 0 to 1.

Architecture

The images fed to the network have a size of 128 x 128 pixels. Convolution layers, ReLU, max-pooling, and fully connected layers comprise the proposed neural network. The output of each layer is the input of another, thus allowing the extraction of features. Since it is sought to differentiate between small and local image characteristics, which may differentiate a person's distinctive features with glaucoma, small 3x3 and 5x5 filters are used in the network. The final two layers will unite these characteristics to make a classification.

The order of the layers; convolution, ReLU and max-pooling is used in several well-known architectures, such as VGG16 [29]. The idea of using two fully connected layers came from the AlexNet [30] network, which uses 3 fully connected layers. Based on these two articles, an architecture is proposed.

As activation function at the output of the network, the sigmoidal function was used. This function was chosen because we are interested in a binary classification, so it will give us a probability of which class it can belong to. This activation function should not be confused with the one used between each convolutional layer. ReLU is used to add non-linearity to the

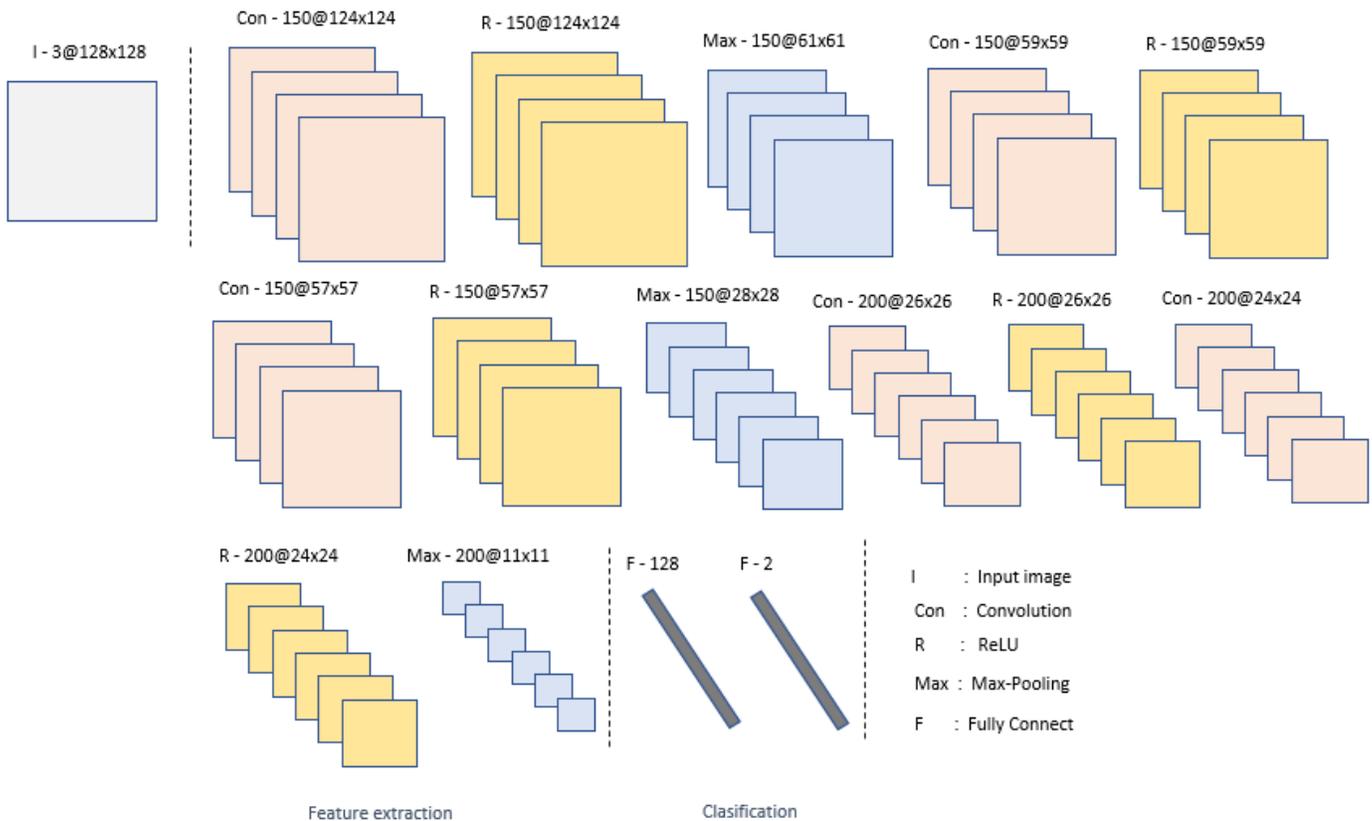


FIGURE 7. Overview of the proposed convolutional neural network architecture.

network and sigmoidal is used to regularize the output. The first fully connected layer has 128 neurons and the second has 2 neurons for classification.

As a loss function, we use Binary Cross Entropy. This was chosen due to the fact that binary classification is sought. This function will give us the prediction error, which will indicate to the network how correct its classification, which will help the network to improve as they pass the epochs. To avoid overfitting, dropout was used, which randomly turns off neurons in each training period, in order to distribute the power of selection of characteristics to all neurons equally, in this way, the model to learn independent characteristics. Adam was used as an optimizer, which combines the advantages of some others, such as the Momentum, maintaining the exponential moving average of the gradients, and the RMSprop, which maintains the exponential moving average, but of squared gradients.

The proposed neural network consists of fifteen layers. This number of layers was determined by experimentation. We have to take into account that with a shallower network, the extraction of features could be poor and therefore increase the difficulty to identify the desired features. However, with a deeper architecture, there is a risk of overfitting.

Convolutional layers represent high-level data features, and combining with a fully connected layer is one way to learn non-linear combinations of those features. In the network, two completely connected layers are used, seeking more significant learning of these combinations. To choose the best learning rate, we carried out tests with different values, and the one that had a better relationship between learning rate and performance is the selected one. The two images where the optic disc was not correctly detected are shown in Figure 8.

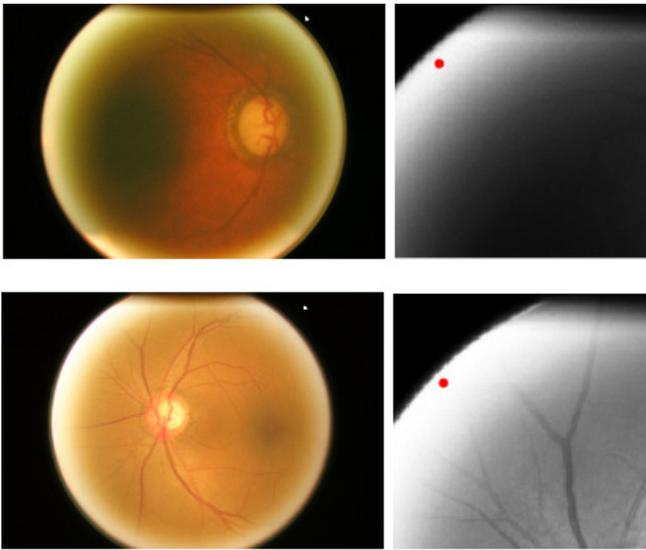


FIGURE 8. Unidentified optic disc.

In images that were not appropriately cropped, excess brightness may be noticed on the circumference of the fundus image. For this reason, the algorithm places the optic disc in that position and does not cut the image correctly. Of the 650 used images, 482 are from healthy patients and 168 from glaucoma patients.

Training and testing

455 images were used for training and 195 for validation, that is, 70% for training and 30% for validation. The images were chosen at random. The algorithm was developed in Python and ran on a computer with an Intel Core i7-3160QM 2.3 GHz CPU and 8 GB of RAM. The training execution time was 18.5 hours. Python 3 and TensorFlow were used. To implement the algorithm, it was necessary to use the Keras, Pandas, Numpy and Matplotlib libraries.

To ensure the obtained results were correct, we implemented cross-validation. For this validation, the algorithm was executed ten times for a k-fold with a value of $k = 10$. Each of the executions had 50 epochs. The advantage that it gives our model the opportunity to perform multiple training-test partitions. This can better show how our model performs on unseen data.

RESULTS AND DISCUSSION

For evaluation of the obtained results, some metrics such as accuracy, sensitivity, and area under the curve were used:

- True-Positives (TP): The predicted value was positive and agreed with the true value.
- True-Negatives (TN): The predicted value was negative and agrees with the true value.
- False-Positives (FP): The model is classified as a positive class, and the true value is negative.
- False-Negatives (FN): The model is classified as a negative class, and the actual value is positive.

Accuracy tells us the percentage of predictions that were made correctly.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

The sensitivity tells us how many of the positive cases the model was able to predict correctly.

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

The ROC-AUC is the metric that shows us a probability curve, which is shown by sketching the range of true positives against the range of false positives at various threshold values. The results obtained with different learning rates without ROI are shown in Table 1.

TABLE 1. Performance without ROI.

Learning rate	Accuracy	Sensitivity	ROC-AUC
0.01	88.72	89.42	89.71
0.001	90.01	90.01	90.35
0.0001	91.02	93.01	89.14

The results obtained with different learning rates and ROI are shown in Table 2.

TABLE 2. Performance with ROI.

Learning rate	Accuracy	Sensitivity	ROC-AUC
0.01	91.36	92.62	92.87
0.001	92.13	93.18	92.57
0.0001	93.03	94.02	93.56

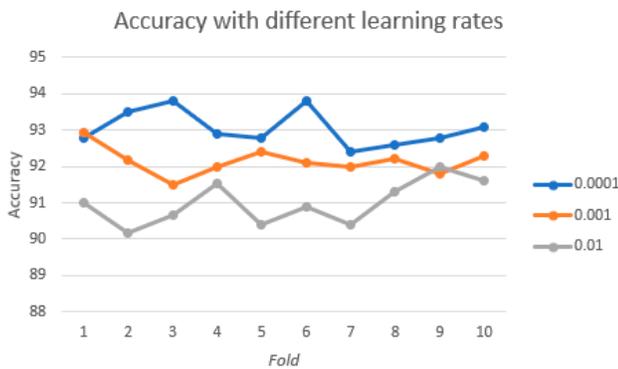


FIGURE 9. Performance with different learning rate.

Figure 10 shows the graph obtained for ROC-AUC, threshold values are in intervals of 0.01.

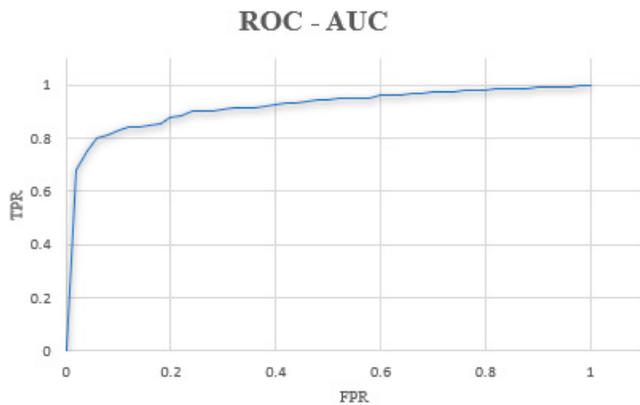


FIGURE 10. ROC-AUC.

In order to check on which part of the image the network was more focused to perform the classification, it was decided to obtain a method of visualization of neuron activations. There are several options to perform this task. Erhan *et al.* [31] propose a method that seeks to

produce an image that shows the activation of the specific neuron we are trying to visualize. Zeiler *et al.* [32] propose a method to find the patterns of the input image that activate a specific neuron in a layer of a CNN. Dosovitskiy *et al.* [33] propose a method that consists of training a neural network that performs the steps in the opposite direction to the original one so that the output of this new network is the reconstructed image.

The method used in this work is the one proposed by Selvaraju *et al.* [34] that adopts a CAM architecture, which generates a class activation map that indicates the most used image regions. Retraining is required for this purpose. In Figure 11 we can visualize the regions in the image that have the most weight while making the classification in some of the images.

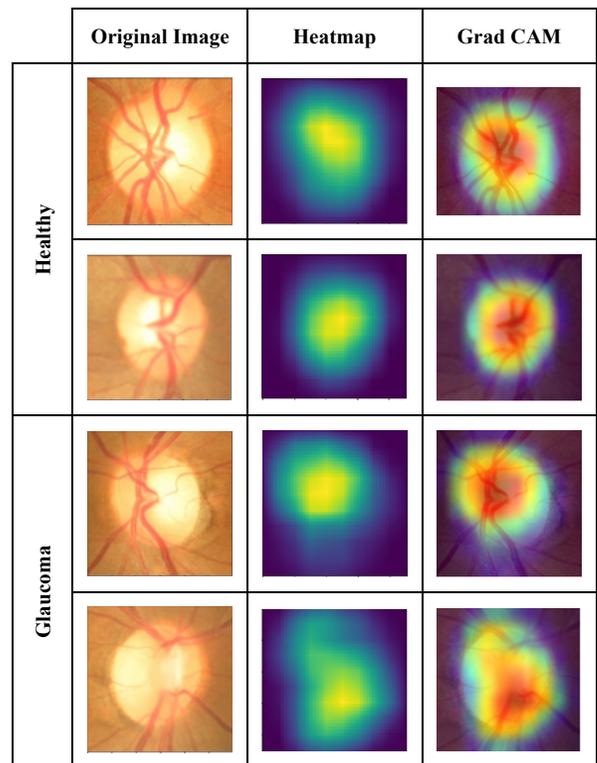


FIGURE 11. Class activation map display.

For the color map, the Jet configuration was used, which returns a color map in a three-column matrix, with the same number of rows as the color map of the original image.

Each row will have different intensities between the color red, green, and blue.

Table 3 compares the performance of our method against other methods found in the literature. It outperforms all other results in any of the displayed metrics.

TABLE 3. Comparison of our method against other methods of state-of-the-art.

Method	Performance
Six layers CNN ^[5]	ROC-AUC: 88.7
Different features from Gabor transform (SVM) ^[6]	Accuracy: 93.1 Sensitivity: 89.75 Specificity: 96.2
Eighteen layers CNN ^[11]	Accuracy: 98.13 Sensitivity: 98 Specificity: 98.3
GIST y PHOG (SVM) ^[8]	Accuracy: 83.4 ROC-AUC: 88
VGG16 + LSTM ^[25]	Sensitivity: 95 Specificity: 96
VGG19 + Transfer learning ^[26]	ROC-AUC: 94.2 Sensitivity: 87 Specificity: 89.01
Xception ^[27]	ROC-AUC: 96.05 Specificity: 85.8 Sensitivity: 93.46
Proposed	Accuracy: 93.03 Sensitivity: 94.02 ROC-AUC: 93.56

CONCLUSIONS

As we stated, glaucoma is one of the principal causes of blindness globally. It could be vitally important to have a tool capable of supporting ophthalmologists to diagnose this condition more quickly.

The proposed method achieves excellent metrics with a not-so-deep neural network, achieving an accuracy, sensitivity, and area under the curve of 93.22, 94.14, and 93.98, respectively. To corroborate the performance of our approach, we did our analysis on the ORIGA database, which is public, and it is one of the most used databases for glaucoma analysis.

The preprocessing of the images to obtain the ROI also helps the algorithm be more effective. It obtains the region where the optical disc is located in almost all the images of the database used, being a method to use in future work.

We will further examine different alternatives to increase the classification performance, either by preprocessing or modifying the network structure.

The purpose of this investigation is to get a high classification method that could be implemented for automatic glaucoma detection; this would save specialists time and speed up the diagnostic process.

By obtaining the characteristics map, we were able to visualize in which part of the image the decision of the network has more weight. Therefore, we conclude that the decision is made depending on the disc and optical cup characteristics, as shown in the results.

For any reader that would like to see the implementation code, it can be provided by requesting it to the corresponding author.

AUTHOR CONTRIBUTIONS

Authors H.S.-C., S.T.-A. and M.A.V.-M. conceptualized the project from the beginning. H.S.-C. acquired data and performed its curation; and participated in the following activities: data analysis, proposed the use of a CNN as classifier and the algorithm for finding the ROI, carried out the classification and research, compared the method with other glaucoma classifications, included methodology in the software, also with a visualization interface, and wrote the original draft. G.A.-F. analyzed the data and carried out the classification, validate the methods and wrote the original draft. S.T.-A. also wrote the original draft. M.A.V.-M. and J.M.R.-A. reviewed and edit the final version of the manuscript.

ACKNOWLEDGMENT

We would like to thank CONACyT for partially funding the project.

ETHICAL STATEMENT

The database used in this work is public; therefore, all ethical considerations are met.

REFERENCES

- [1] Raghavendra U, Fujita H, Bhandary SV, Gudigar A, et al. Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. *Inf Sci [Internet]*. 2018;441:41-49. Available from: <https://doi.org/10.1016/j.ins.2018.01.051>
- [2] Belghith A, Balasubramanian M, Bowd C, Weinreb RN, et al. Glaucoma progression detection using variational expectation maximization algorithm. In: 2013 IEEE 10th International Symposium on Biomedical Imaging [Internet]. San Francisco: IEEE. 2013:876-879. Available from: <https://doi.org/10.1109/ISBI.2013.6556615>
- [3] Yin F, Liu J, Wong DWJ, Tan NM, et al. Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis. In: 2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS). Rome: IEEE. 2012:1-6. Available from: <https://doi.org/10.1109/CBMS.2012.6266344>
- [4] Zilly J, Buhmann JM, Mahapatra D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Comput Med Imaging Graph [Internet]*. 2017;55:28-41. Available from: <https://doi.org/10.1016/j.compmedimag.2016.07.012>
- [5] Chen X, Xu Y, Kee Wong DWK, Wong TY, et al. Glaucoma detection based on deep convolutional neural network. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [Internet]. Milan: IEEE. 2015:715-718. Available from: <https://doi.org/10.1109/EMBC.2015.7318462>
- [6] Acharya UR, Ng EYN, Eugene LW, Noronha KP, et al. Decision support system for the glaucoma using Gabor transformation. *Biomed Signal Process Control [Internet]*. 2015;15:18-26. Available from: <https://doi.org/10.1016/j.bspc.2014.09.004>
- [7] Díaz Pinto AY. Machine Learning for Glaucoma Assessment using Fundus Images [Ph.D.'s thesis]. [Valencia]: Universitat Politècnica de València, 2019. 117p.
- [8] Gour N, Khanna P. Automated glaucoma detection using GIST and pyramid histogram of oriented gradients (PHOG) descriptors. *Pattern Recognit Lett [Internet]*. 2020;137:3-11. Available from: <https://doi.org/10.1016/j.patrec.2019.04.004>
- [9] Nawaldgi S. Review of automated glaucoma detection techniques. In: 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) [Internet]. Chennai: IEEE. 2016:1435-1438. Available from: <https://doi.org/10.1109/WiSPNET.2016.7566373>
- [10] Septiarini A, Harjoko A, Pulungan R, Ekantini R. Automated Detection of Retinal Nerve Fiber Layer by Texture-Based Analysis for Glaucoma Evaluation. *Healthc Inform Res [Internet]*. 2018;24(4):335. Available from: <https://doi.org/10.4258/hir.2018.24.4.335>
- [11] Del Portillo C. Cómo Entender y Vivir Con Glaucoma. Glaucoma Research Foundation. [Internet]. 2013; Available from: https://www.glaucoma.org/GRF_Understanding_Glaucoma_ES
- [12] Li A, Wang Y, Cheng J, Liu J. Combining Multiple Deep Features for Glaucoma Classification. In: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [Internet]. Calgary: IEEE. 2018:985-989. Available from: <https://doi.org/10.1109/ICASSP.2018.8462089>
- [13] Carrillo J, Bautista L, Villamizar J, Rueda J, et al. Glaucoma Detection Using Fundus Images of the Eye. In: 2019 XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA) [Internet]. Bucaramanga: IEEE. 2019:1-4. Available from: <https://doi.org/10.1109/STSIVA.2019.8730250>
- [14] Sarhan A, Rokne J, Alhadj R. Glaucoma detection using image processing techniques: A literature review. *Comput Med Imaging Graph [Internet]*. 2019;78:101657. Available from: <https://doi.org/10.1016/j.compmedimag.2019.101657>
- [15] Li L, Xu M, Liu H, Li Y, et al. A Large-Scale Database and a CNN Model for Attention-Based Glaucoma Detection. *IEEE Trans Med Imaging [Internet]*. 2020;39(2):413-424. Available from: <https://doi.org/10.1109/TMI.2019.2927226>
- [16] Zhang Z, Lee BH, Liu J, Wong DWK, et al. Optic disc region of interest localization in fundus image for Glaucoma detection in ARGALI. In: 2010 5th IEEE Conference on Industrial Electronics and Applications [Internet]. Taichung: IEEE. 2010:1686-1689. Available from: <https://doi.org/10.1109/ICIEA.2010.5515221>
- [17] Rani A, Mittal D. Measurement of Arterio-Venous Ratio for Detection of Hypertensive Retinopathy through Digital Color Fundus Images. *J Biomed Eng Med Imaging [Internet]*. 2015;2(5):35. Available from: <https://doi.org/10.14738/jbemi.25.1577>
- [18] Zhang Z, Yin FS, Liu J, Wong WK, 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 [Internet]. Buenos Aires: IEEE. 2010:3065-3068. Available from: <https://doi.org/10.1109/IEMBS.2010.5626137>
- [19] Sng CC, Foo L-L, Cheng C-Y, Allen JC, et al. Determinants of Anterior Chamber Depth: The Singapore Chinese Eye Study. *Ophthalmology [Internet]*. 2012;119(6):1143-1150. Available from: <https://doi.org/10.1016/j.ophtha.2012.01.011>
- [20] Decencièrre E, Zhang X, Cazuguel G, Lay B, et al. Feedback on a publicly distributed image database: the Messidor database. *Image Anal Stereol [Internet]*. 2014;33:231-234. Available from: <https://doi.org/10.5566/ias.1155>
- [21] Brandon L, Hoover A. Drusen Detection in a Retinal Image Using Multi-level Analysis. In: Ellis R.E., Peters T.M. (eds). *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003* [Internet]. Springer, Berlin, Heidelberg: Lecture Notes in Computer Science. 2003:618-625p. Available from: https://doi.org/10.1007/978-3-540-39899-8_76
- [22] Sivaswamy J, Krishnadas SR, Joshi GD, Jain M, et al. Drishti-GS: Retinal image dataset for optic nerve head segmentation. In: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI) [Internet]. Beijing: IEEE. 2014:53-56. Available from: <https://doi.org/10.1109/ISBI.2014.6867807>
- [23] Han J, Kamber M, Pei J. *Data Mining: Concepts and Techniques* [Internet]. Waltham: Elsevier. 2011:113-114p. Available from: <https://doi.org/10.1016/B978-0-12-381479-1.00016-2>

- [24] Bhatkalkar BJ, Reddy DR, Prabhu S, Bhandary S. Improving the Performance of Convolutional Neural Network for the Segmentation of Optic Disc in Fundus Images Using Attention Gates and Conditional Random Fields. *IEEE Access* [Internet]. 2020;8:29299-29310. Available from: <https://doi.org/10.1109/ACCESS.2020.2972318>
- [25] Gheisari S, Shariflou S, Phu J, Kennedy PJ, et al. A combined convolutional and recurrent neural network for enhanced glaucoma detection. *Sci Rep* [Internet]. 2021;11(1):1995. Available from: <https://doi.org/10.1038/s41598-021-81554-4>
- [26] Gómez-Valverde J, Antón A, Fatti G, Liefers B, et al. Automatic glaucoma classification using color fundus images based on convolutional neural networks and transfer learning. *Biomed Opt Express* [Internet]. 2019;10(2):892-913. Available from: <https://doi.org/10.1364/BOE.10.000892>
- [27] Diaz-Pinto A, Morales S, Naranjo V, Köhler T, et al. CNNs for automatic glaucoma assessment using fundus images: an extensive validation. *BioMed Eng Online* [Internet]. 2019;18(1):1-19. Available from: <https://doi.org/10.1186/s12938-019-0649-y>
- [28] Unión Internacional de Telecomunicaciones. Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9. ITU [Internet]. 2011; Available from: RECOMENDACIÓN UIT-R BT.601-7 - Parámetros de codificación de televisión digital para estudios con formatos de imagen normal 4:3 y de pantalla ancha 16:9 (itu.int)
- [29] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* [Internet]. 2014; Available from: <https://arxiv.org/abs/1409.1556>
- [30] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun* [Internet]. 2017; 60(6):84-90. Available from: <https://doi.org/10.1145/3065386>
- [31] Erhan D, Bengio Y, Courville AC, Vincent P. Visualizing Higher-Layer Features of a Deep Network [Internet]; 2009. Report 1341. Available from: (PDF) Visualizing Higher-Layer Features of a Deep Network (researchgate.net)
- [32] Zeiler MD, Taylor GW, Fergus R. Adaptive deconvolutional networks for mid and high level feature learning. In: 2011 International Conference on Computer Vision [Internet]. Barcelona: IEEE. 2011:2018-2025. Available from: <https://doi.org/10.1109/ICCV.2011.6126474>
- [33] Dosovitskiy A, Brox T. Inverting Visual Representations with Convolutional Networks. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. Las Vegas: IEEE. 2016:4829-4837. Available from: <https://doi.org/10.1109/CVPR.2016.522>
- [34] Selvaraju RR, Cogswell M, Das A, Vedantam R, et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In: 2017 IEEE International Conference on Computer Vision (ICCV) [Internet]. Venice: IEEE. 2017:618-626. Available from: <https://doi.org/10.1109/ICCV.2017.74>