IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 23, 2021, accepted June 24, 2021, date of publication July 5, 2021, date of current version July 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3094649

Classification of Eye Diseases in Fundus Images

OMAR BERNABÉ[1], ELENA ACEVEDO[1], ANTONIO ACEVEDO[1], (Member, IEEE), RICARDO CARREÑO[2], AND SANDRA GÓMEZ[3]

[1]Instituto Politécnico Nacional, ESIME Zacatenco, Mexico City 07738, Mexico
[2]Universidad del Istmo, Campus Tehuantepec, Oaxaca 70760, Mexico
[3]Instituto Politécnico Nacional, UPIITA, Mexico City 07340, Mexico

Corresponding author: Antonio Acevedo (eacevedo@ipn.mx)

ABSTRACT Eye diseases have been a severe problem worldwide, especially in developing countries where technology and finance are limited. Today, the problem is being resolved thanks to the task of classification that is part of pattern recognition. Its primary goal is to group standard features from any entity, object, phenomenon, or event belonging to the real or abstract world. Convolutional Neural Networks are a type of Artificial Neural Network used in intelligent pattern classification, Machine Learning, and Data Mining. Also, medicine and ophthalmology used these algorithms for detecting diseases in the human body. This work presents a novel intelligent pattern classification algorithm based on a Convolutional Neural network, which is validated through the K-Fold Cross Validation test. Two different groups of retinography images are given: Glaucoma and Diabetic Retinopathy. The result of accuracy percentage was 99.89%. Numerical metrics: Accuracy, Recall, Specificity Precision, and F_1 score with values close to 1, and ROC curves support the suitable performance of the proposed classifier.

INDEX TERMS Artificial intelligence, convolutional neural networks, machine learning, supervised learning.

I. INTRODUCTION

According to the first report of the World Health Organization in 2019, it is estimated that there are one thousand two hundred million persons with an eye condition. Some of the conditions are refraction error (one hundred and twenty-three million), eye cataracts (sixty-five million), Glaucoma (seven million), corneal opacities (four million), macular degeneration (ten million), diabetic retinopathy (three million), and presbyopia (eight hundred and twenty-six million).

There is inadequate attention to eye diseases due to inequalities in health coverage, low quality in prevention services, treatment and rehabilitation, and little integration of health services.

On the other hand, image processing, particularly Deep Learning techniques, is significant support for detecting and diagnosing the latter eye conditions and Glaucoma and Diabetic Retinopathy.

Convolutional Neural Networks (CNN) are algorithms from Deep Learning and are very useful in recognizing, detecting, and classifying images.

The associate editor coordinating the review of this manuscript and approving it for publication was Gina Tourassi.

A. RETINA

The retina is a thin, delicate, transparent sheet of tissue derived from the neuroectoderm [1]. It is the light-sensing tissue that resides on the back of the eye.

Retinography is a non-invasive diagnostic test that identifies eye problems such as Diabetic Retinopathy, Diabetic Macular Edema, Hypertensive Retinopathy, Age-related Macular Degeneration, and Glaucoma.

Retinographies are images that allow observing the retina. The structures of the fundus that are examined are [2]:

∗ Retinal parenchyma: the retina is an almost transparent membrane. The reddish color it presents is due to the hue of the Retinal Pigment Epithelium (RPE) and can vary depending on race, age, and skin color. It loses its shine over the years.

∗ Papilla or Optic Disc (OD): the visible part of the optic nerve that is rounded or oval in a vertical direction, 1.5 mm in diameter; White-pinkish color, with a whitish central area that corresponds to the physiological excavation known as the Optical Cup (OC) and is of variable size, not exceeding 30% of the OD diameter.

∗ Retinal vessels: in the center of the optic disc is the vascular bundle formed by the central retinal artery and vein, which are divided into the superior and inferior temporal

arteries and veins, and the superior and inferior nasal arteries; branches of the retinal vessels approach from all sides, but do not reach the fovea; the vein is of a darker wine-red color and wavier in trajectory with an artery-vein caliber ratio of 2/3.

∗ Macula: located approximately two diameters (OD) from the papilla.

The structures can be observed in Figure 1.

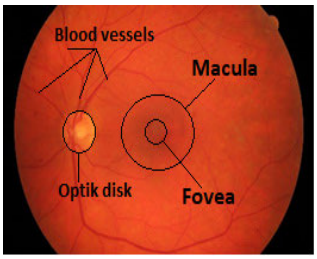

FIGURE 1. The retina of the eye is the entire orange circle. The OD is the yellowish part, blood vessels, and macula.

1) RETINAL DISEASES

There are three dominant methods practiced for the detection of ***Glaucoma***: Intraocular Pressure Measurement (IPM), Function-Based Visual Field Test (FBVFT), and Optic Nerve Head (ONH). The ONH examination is a convenient paradigm for identifying Glaucoma in its early stage and is performed contemporaneously by professional and qualified glaucoma specialists. It uses various morphological parameters such as vertical Cup-to-Disc Ratio (CDR), Edge-to-Disc Ratio (EDR), Neuro-retinal Border Ratio (NBR), and disc diameter to detect Glaucoma on retinal fundus images [3]. In other words, the diagnosis of Glaucoma using the ONH technique is achieved by examining the OD. The eye has millions of nerve fibers that run from the retina to the optic nerve, located in the center of the OD. If Glaucoma is present, the OD becomes hollow, and the optic nerve assumes the shape of a cup (OC), which exceeds the value of 0.5 concerning OD [4].

In Figure 2, the differences between an eye with Glaucoma and another healthy eye are shown.

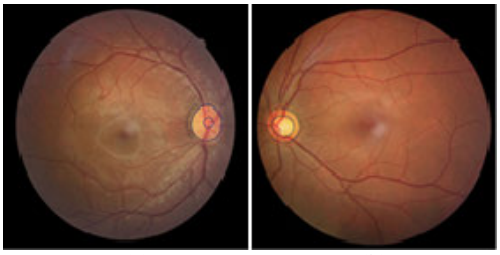

FIGURE 2. Retinography a) shows an entirely healthy eye and b) has Glaucoma. The relationship between the cup and the optic disc exceeds the value of 0.5.

Diabetes mellitus occurs when the pancreas cannot secrete enough insulin, which affects the circulatory system and generates an increase in glucose, which causes the deterioration of the blood vessel walls that leads to bubbles, obstructions, and leaks in the veins. ***Diabetic retinopathy*** is a condition caused by diabetes mellitus. Observing the retinal vascular circulation is very important to control systemic diseases before they spread [5]. The lesions that occur in the fundus of the eye due to this disease are:

- Micro-Aneurysms (MA): they are small dark red dots.
- Exudates are masses of cells, and fluid seeped out of blood vessels or organs, especially in inflammation.

Figure 3 shows the conditions caused by the disease.

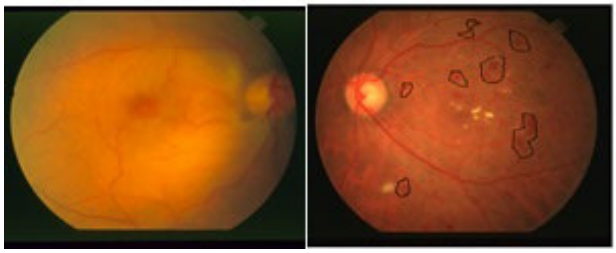

FIGURE 3. a) Healthy retinography. b) A diabetic retinography shows the dark red microaneurysms and slightly smaller red hard exudates.

Diabetic Macular Edema (DME) [6] is a potential complication of diabetic retinopathy (DR). In people with type 2 diabetes, DME causes the most vision loss. It involves the deterioration of the blood-retinal barrier in the eye and a resulting pooling of fluid within the central area of the retina.

The types of DME are:

- Clinically significant macular edema
- Non-central-involved macular edema
- Central-involved macular edema

Figure 4 shows a healthy retinography and clinically significant macular edema.

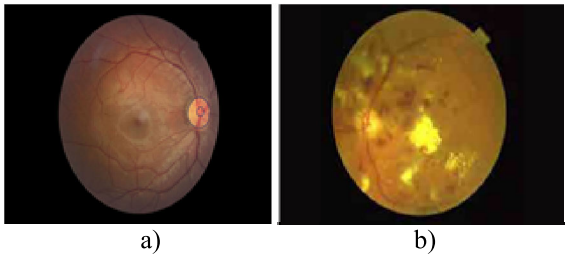

FIGURE 4. a) Healthy retinography. b) A clinically significant macular edema.

Vascular changes associated with systemic elevated blood pressure are visible in the retina as ***Hypertensive Retinopathy (HR)*** [7]. In Figure 5, the grades of HR can be observed.

Age-related macular degeneration (AMD) [8] is a disorder of the macula characterized by one or more of the following:

- Presence of at least intermediate-size drusen (63 μm or larger in diameter)

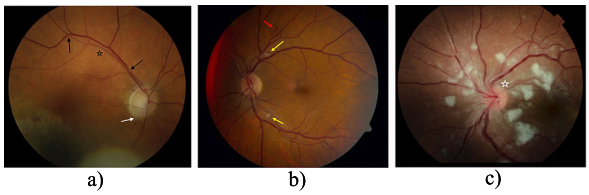

FIGURE 5. Grades of hypertensive retinopathy. a) Mild hypertensive retinopathy (in an eye with an unrelated chorioretinal lesion) with arteriolar narrowing (white arrow), copper wiring (black star), and AV nicking (black arrow). b) Moderate hypertensive retinopathy with features of mild hypertensive retinopathy as well as cotton wool spots (yellow arrow) and intraretinal hemorrhages (red arrow). c) Severe hypertensive retinopathy with features of moderate hypertensive retinopathy and optic disc swelling (white star).

- Retinal pigment epithelium (RPE) abnormalities such as hypopigmentation or hyperpigmentation
- Reticular pseudodrusen4
- Presence of any of the following features: geographic atrophy of the RPE, choroidal neovascularization (exudative, wet), polypoidal choroidal vasculopathy, or retinal angiomatous proliferation

An example of the first feature can be observed in Figure 6.

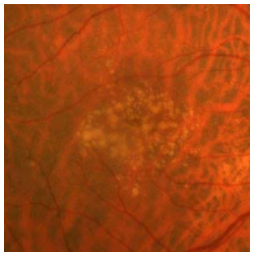

FIGURE 6. Drusen in non-neovascular age-related macular degeneration appear as round, yellow deposits deep in the retina during fundus examination.

B. RELATED WORK

Some studies have made the detection of these diseases in retinographies. However, there is also a great variety of classification structures, such as the K-Nearest Neighbors algorithm, Decision Trees, Support Vector Machines, Multilayer Perceptrons, and Convolutional Neural Networks. The summary of some related works is shown in Table 1. It can be observed that all the proposals classify one disease. This work presents a Convolutional Neural Network (CNN) that classifies two conditions: Glaucoma and Diabetic Retinopathy.

This paper is organized as follows. Section 2 presents the definitions of Image Processing and Convolutional Neural Networks. The proposed model is introduced in Section 3; Section 4 describes the dataset and the CNN architecture. Section 5 gives the Results. Finally, we highlight the Conclusions.

II. METHODS AND MATERIALS

A. IMAGE PRE-PROCESSING

Retinography images were pre-processing and came from two different databases to form the one used in this paper. The dimensionality of glaucoma retinographies is 3072 × 2048 pixels, and diabetic retinopathies have a resolution of 700 × 605 pixels. They do not have the same dimensionality. Therefore, all images must have the same size. Then, the standard size was 250 × 250 because it is optimal for analyzing this type of image according to related works [4] and [13].

As is known, an image in RGB color format contains three two-dimensional matrices, which correspond to planes of primary colors: Red (R), Green (G), and Blue (B) [4]. It is necessary to work with such planes in these types of images because all three contain essential information. Figure 7 shows a diagram of the matrix representation of these planes. Following the recommendations of the related works, the best plane to detect the different structures of each condition is the green plane (G). Also, it is the plane that contains more information and is the cleanest. However, the red plane (R) is the loudest. Consequently, a new plane was created based on the green (G) and the blue (B) planes and replaced by the red one (R). It is also convenient to eliminate noisy pixels; therefore, it is necessary to use a filter.

$$I_R(m,n,1)=\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \quad I_G(m,n,2)=\begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & g_{mn} \end{bmatrix}$$

$$I_B(m,n,3)=\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

FIGURE 7. RGB planes, each value within the matrix, represents its pixel intensity at that position in the image.

When an image is processed in the spatial domain, it can be expressed as in Equation (1):

$$g(x, y) = T(f(x, y)) \tag{1}$$

where:

$f(x, y)$ is the original image

$g(x, y)$ is the processed image

T is an operator that acts on f and is defined in some environment of (x, y).

In the case of filtering, it is an operator that acts on the original image. There are different smoothing filters. Below are the categories of these filters.

- Linear filters find average and Gaussian.
- Non-linear filters are made up of ordered statistical filters.

In the mean filter used in this proposal, each pixel value is replaced by the mean of the values of the neighboring pixels, performing a convolution with a given mask. For example, the most used mask has a 3 × 3 dimension, as is shown in Figure 8.

TABLE 1. Different approaches to classify eye diseases: Glaucoma and diabetic retinopathy.

Year	Eye disease	Algorithm	Observations	Effectiveness
2015	Glaucoma [9]	1. Support Vector Machine. 2. Artificial Neural Network	Processing techniques were used to extract characteristics. LOOCV	94.11% 89.6%
	Glaucoma [10]	1. Support Vector Machine.	Processing techniques were used to extract characteristics.	97% Max. 82% Min.
2016	Diabetic Retinopathy [11]	1. K-Nearest Neighbors. 2. Support Vector Machine (x2). 3. Naïve Bayes.	Fuzzy processing and circular Hugh transform techniques were used.	93% 70%, 93% 91.49%
	Glaucoma [12]	Support Vector Machines.	The Local Binary Pattern technique is used for the extraction of texture features.	91.20%
2017	Diabetic Retinopathy [13]	**Deep Convolutional Neuronal Network**	The classification was performed using the characteristics extracted from an increasing gradient tree and a convolutional neural network.	**94.50%**
2018	Glaucoma [14]	Neuronal Network with multiple inputs.	First, the VGG16 neural network was used, and learning was transferred into the multi-input network 5-f c-v.	91.51%
2018	Diabetic Retinopathy [4]	**Deep Convolutional Neuronal Network.**	The circular Hugh transform was used to ignore the Optical Disc, and the performance was calculated in 3 different databases.	99.17% 98.53% **99.18%**
	Glaucoma [8]	Support Vector Machine.	AlexNet based features and critical features were extracted using "Stacked Auto Encoder". 10-f c-v.	98.80%
2019	Diabetic Retinopathy [15]	1. Decision Tree 2. Multilayer Neural Network 3. Support Vector Machine. 4. -Nearest Neighbors. 5. Naïve Bayes	Image features were extracted with the Gabor filter and classified with different superstructures.	65.95% 65.95% 74.46% 73.04% 77.85%
2020	Diabetic Retinopathy [16]	1. Support Vector Machines (x2). 2. Random Forests.	Different features were extracted with processing methods.	85.33% 87.96% 95.08%

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

FIGURE 8. A 3 × 3 dimension mask for filter media.

Then, the convolution operation is applied between a filtering matrix and an image (for each of its planes), the resulting function is Equation (2):

$$g(x, y) = \frac{1}{9} \sum_{s-1}^{3} \sum_{t-1}^{3} w(s, t)x(x + s, y + t) \qquad (2)$$

Figure 9 shows the comparison of two retinographies (original and image pre-processed) when the image processing mentioned above is applied. Such retinographies show quite marked signs of Glaucoma.

Subsequently, Figure 10 presents a comparison between the original and the pre-processed images of diabetic retinopathy.

B. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks appeared in 1990. Yann LeCun *et al.* wrote the first pioneering work of CNN. This work is entitled "Gradient-Based Learning Applied

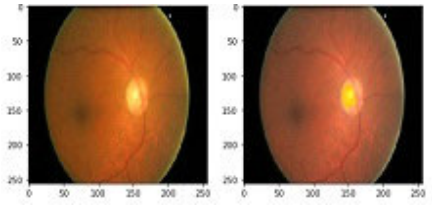

FIGURE 9. Comparison 1. Image a) has a slightly redder color compared to image b). It can be observed that the optical cup is more visible in the second image than in the first.

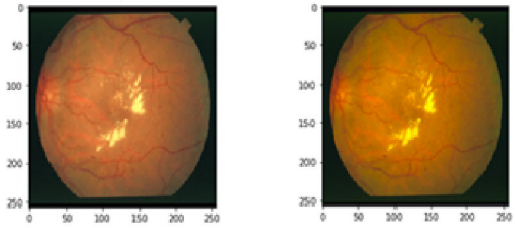

FIGURE 10. a) Original Image and b) New image. Comparison 2. Image b) has better contrast over image a).

to Document Recognition"; this work demonstrated that a CNN model that selects elementary features from a set of progressively more complicated features could be used successfully for handwritten character recognition [17]. Yann LeCun trained a CNN using the MNIST set, which comprises images of handwritten numbers that go from 1 to 9 and include their respective labels.

Now, the function of a CNN will be explained. CNN consists of series of layers of convolution, grouping, and fully connected multilayer perceptron; it also consists of specific parameters used, such as activation functions. First, from an image, in the convolution layers, feature maps are created with convolution filters, whose weights depend on their size. Figure 11 shows the convolution between an image plane with a filter.

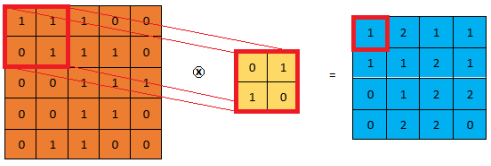

FIGURE 11. Convolution between an image plane with a 2 × 2 filter. The filter traverses the entire image performing convolution operations. The Stride is the size of the space of the filter that convolves the image and, Padding is the size of the characteristic map that depends on the padding with zeros on the edges.

In this sense, what convolution does is similar to the operation of a mask in image processing, where certain types of relationships can be highlighted.

After obtaining feature maps, the ReLU function is applied to break the linearity in the image and extract the most outstanding features from the maps. There are several functions that can be used for the extraction of features. Still, the most used in this type of structure is the ReLU function. This function detects negative values, and such values will be represented as 0, and positive values will be represented with their corresponding value. In the grouping layers, in the maps obtained, the average of the neighboring pixels is calculated, or the largest is selected, to speed up the processes and reduce the maps sizes. The most widely used function is Max-pooling, which consists of taking the highest value of a neighborhood of pixels in an image. Also, it is necessary to define the size of the neighborhood, which in turn must be square in size. Figure 12 shows how the Max-pooling function works on a characteristic map.

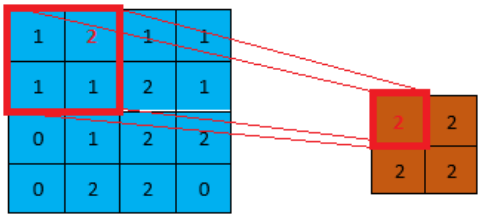

FIGURE 12. The Max-pooling function with 2 × 2 filter consists of selecting the element with the highest value from a neighborhood of numbers.

The convolution and grouping layers can be repeated several times, depending on the data bank complexity; this implies that the feature maps reduce the information that each extra layer of the image is passed through subsequent layers.

Finally, the parameters obtained are converted into a vector as an input of the Multilayer Perceptron. Then, the Convolutional Neural Network is trained with these parameters and weights using any function such as softmax or sigmoid through feed-forward and backpropagation processes [4] and [13]. An example of the architecture of a convolutional neural network is shown in Figure 13.

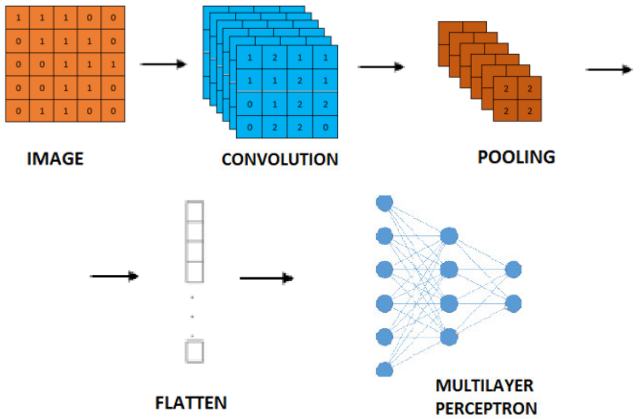

FIGURE 13. The architecture of a convolutional neural network of a layer for an image plane. The convolution and grouping processes can be repeated several times, depending on the complexity of the data bank.

C. CLASSIFICATION METRICS

Classification tasks give us a predicted value when an input pattern is presented [18]. This value can have a limited set of

values (the number of classes). This proposal works with a binary classification task; therefore, the predicted value can be Glaucoma or Diabetic Retinopathy.

The performance of a classification model can be achieved in two ways: using numerical metrics or by plotting a curve and analyze its shape.

1) NUMERICAL METRICS

Accuracy is the proportion of the correct predictions to the total number of predictions. Equation (3) shows the way to calculate this metric.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{3}$$

However, accuracy may not be the best metric to analyze the performance of a classification task in many cases.

Confusion Matrix is a table with n different combinations of predicted and actual values, where n is the number of classes; it essentially summarizes the prediction results of a classification problem. Table 2 shows an example of a confusion matrix for a binary classification problem.

TABLE 2. Confusion matrix for a binary classification problem.

		Predicted	
		0	1
Actual	0	Correctly classified	Incorrectly classified
	1	Incorrectly classified	Correctly classified

This matrix is called confusion because the correctly and incorrectly classified patterns can be viewed as true positives, true negatives, false positives, and false negatives, as shown in Table 3.

TABLE 3. Table of confusion.

		Predicted	
		0	1
Actual	0	True Negatives	False Positives
	1	False Negatives	True Positives

The size of this table is always 2×2, whether the classification is binary or multiclass. There will be many tables as the number of classes in the latter case because it is applied one-versus-all classification.

The definitions of the terms from Table 3 are as follows:

True positives and ***true negatives:*** These are the counts of the correctly predicted data points in the positive and negative classes, respectively.

False positives: They are *Type 1* errors and refer to the count of the data points that belong to the negative class but were predicted to be positive.

False negatives: They are *Type 2* errors and refer to the count of the data points that belong to the positive class but were predicted to be negative.

Three critical metrics can be calculated from the parameters in the confusion matrix. ***Precision***, ***Recall*** (or Sensitivity), and ***Specificity*** are shown in Equations (4), (5), and (6). Precision indicates the general performance of the classifier, Recall indicates the classifier ability to identify positive instances, and Specificity is the classifier ability to identify negative instances

$$Precision = \frac{True\ positives}{True\ positives + False\ positives} \tag{4}$$

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives} \tag{5}$$

$$Specificity = \frac{True\ negatives}{True\ negatives + False\ positives} \tag{6}$$

However, the importance of Precision and Recall depends on the problem statement. Therefore, a metric combines Precision and Recall into a single number that can be used for a fair judgment of the model and is equal to the harmonic mean of these two metrics. Equation (7) shows F_1 score.

$$F_1 score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{7}$$

The value of F_1 score is in the range of 0 and 1. When the value is 0, precision or recall can be 0; conversely, if F_1 score is 1, precision and recall can be 1. Therefore, if the value is close to 1, the classifier will show better performance.

2) CURVE PLOTS

Receiver Operating Characteristic (***ROC***) is a curve used for visualizing the performance of a classification model. The building of this curve includes two parameters: ***True Positive Rate*** (***TPR***) and ***False Positive Rate*** (***FPR***). The first in the Y-axis and the second is on the X-axis. TPR is equal to recall, and FPR is the complement of Specificity. Equation (8) shows the parameter FPR.

$$FPR = \frac{True\ positives}{False\ positives + True\ negatives} \tag{8}$$

Figure 14 shows an illustrative example of a ROC curve.

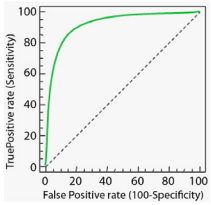

FIGURE 14. Example of ROC curve.

The best performance would be at a point in the upper left corner, i.e., no false negative or false positive. The diagonal is called a non-discrimination line, and it is a random classification. The point of a ROC random classifier will shift to the position (0.5, 0.5). Points above the diagonal represent good ranking results, points below the line for poor results.

III. THE PROPOSED ALGORITHM

A. DATASET

The intelligent pattern classifier was built, trained, and tested with two different datasets. These were obtained on the WEB and were freely accessible. The Origa dataset contains

650 renography images in JPG format, where 482 are healthy fundus images, and 168 have Glaucoma. Therefore, only glaucoma retinographies were taken. The other dataset comes from DIARETDB0 with images in PPM format, containing retinographies with diabetic retinopathy. From this dataset, 397 images were taken. Thus, the dataset contains 168 images with Glaucoma and 397 with Diabetic Retinopathy, a total of 565 retinographies.

B. CONVOLUTIONAL NEURAL NETWORK: TRAINING AND TESTING

The images have already gone through the pre-processing mentioned in Section 3.1 of this paper for CNN training and testing. For this, the Keras library was used, which is an open-source neural network library written in Python. The advantage of this library is that a pre-trained CNN can be used. However, for the present work, it was decided to build a new one. Figure 15 shows the architecture of the proposed convolutional neural network.

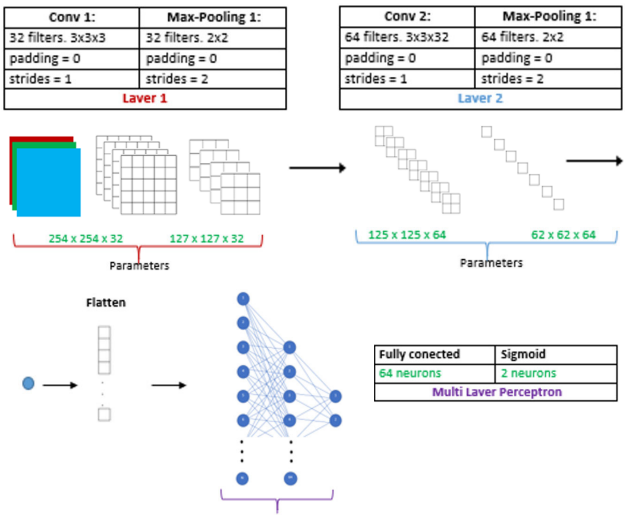

FIGURE 15. The architecture of the proposed convolutional neural network.

In this way, the Convolutional Neural Network consists of two layers of convolution, where the first layer has 32 filters for each RGB plane, and the result is 32 feature maps with dimensions 254 × 254. After this, the dimensions are reduced to 127 × 127 through the Max-pooling function. The same procedure is carried out with the resulting maps in the second layer, with 64 filters, and the dimensions of the feature maps are reduced to 64, with sizes of 62 × 62. The maps are converted into a vector using the flatten function. The dimension of this vector is 246,016 (62 × 62 × 64); these are the number of input neurons in the MLP. The perceptron consists of a single deep layer with 64 neurons. The sigmoid function was applied. The other essential parameters of the CNN are:

∗ Lot size: 32
∗ Number of epochs: 5
∗ Loss function: binary cross-entropy
∗ Optimizer: Adam
∗ Metric: Precision

With these parameters, CNN can detect the most critical features of each disease, such as micro-aneurysms or hard exudates in the case of Diabetic Retinopathy. In the case of Glaucoma, the hard disc and the optic cup are detected.

IV. RESULTS

The complete code was programmed in Python language, making use of OpenCV libraries for image pre-processing. Keras was used to built CNN, an open-source library written in the same programming language. The characteristics of the computer were the following:

∗ 8th generation Intel CORE i7 + processor at 2.2 GHz and Turbo Boost at 4GHz.
∗ NVIDIA GEFORCE GTX 1050 graphics card
∗ 8 GB RAM
∗ Windows 10 Home operating system

The validation method decided to use was the K-Fold Cross-Validation. Furthermore, it is possible to train and test all the patterns (fundus images) at different times. The dataset is close to being unbalanced, so it was decided to follow the recommendations of F. Dietrich [19]. Where

$$IR = \frac{Majority\ class}{Monority\ class} \cong 2.5$$

The dataset was divided into training $\boldsymbol{E}$ and test $\boldsymbol{P}$ sets.

The 2-Fold Cross-Validation method was applied, and the algorithm can be observed in Figure 16.

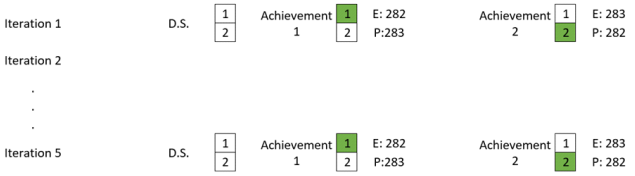

FIGURE 16. 2-fold cross-validation as a validation method.

Figures 17, 18, 19, 20, and 21 show the CNNs training performance over five epochs.

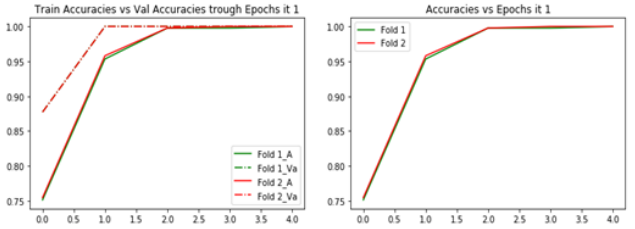

FIGURE 17. Accuracy behavior in the set P throughout the five epochs in iteration 1. The precision in the first epoch is close to 0.75. After that, the accuracy increases considerably, tending to one for two training folds at different times. On the other hand, the validation precision is also quite acceptable because it tends to one.

The concepts of validation and training precision are essential to understand CNN behavior. First, CNN was trained with

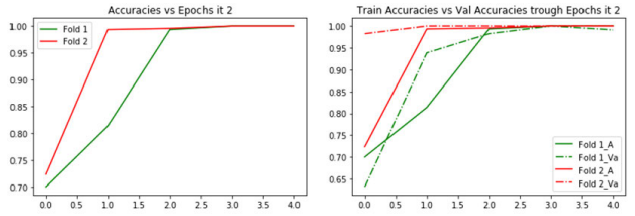

FIGURE 18. Accuracy behavior in the set P over five epochs in iteration 2. In this iteration, the precision in the first epoch is smaller. However, times two and three peaks for training on the two folds. The validation precision concerning accuracy is also quite acceptable.

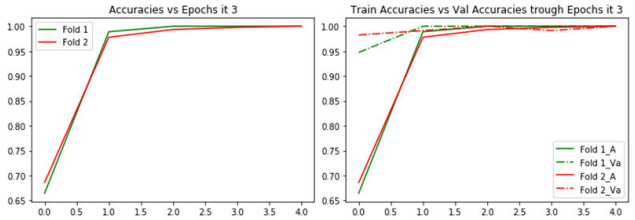

FIGURE 19. Accuracy behavior in set P over five epochs in iteration 3. Outstanding percentage of accuracy from the second epoch and excellent validation precision.

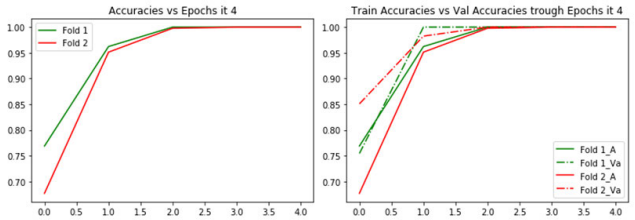

FIGURE 20. Accuracy behavior in set P over five epochs in iteration 4. In this iteration, the precision and validation precision percentages are very close to iteration one, which means that the neural network continues training correctly.

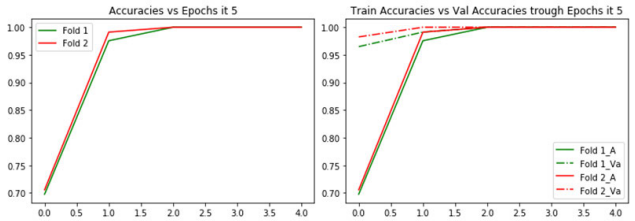

FIGURE 21. Accuracy behavior in the set P over five epochs in iteration 5. The precision and validation percentages are excellent. What remains is to test the neural network with the training set.

the set $\boldsymbol{E}$ and divided into two subsets: learning (2 executions) and validation. The first subset is used to train the Convolutional Neural Network. Remember, the weights are adjusted in each epoch, and the network could over-adjust these weights at a specific time in the process. Therefore, at the end of each epoch, the validation set is shown to CNN. Convergence to one of these parameters means that the network is not over-tuning, but it is learning correctly.

Figures show the behavior of the training and validation accuracy; validation has a pretty acceptable behavior. In other words, if it did not converge to one, it could be that CNN is over-adjusting its weights, and therefore some neurons would need to be turned off.

Figure 22 shows the confusion matrices that can be used to present numerical metrics shown in Table 4.

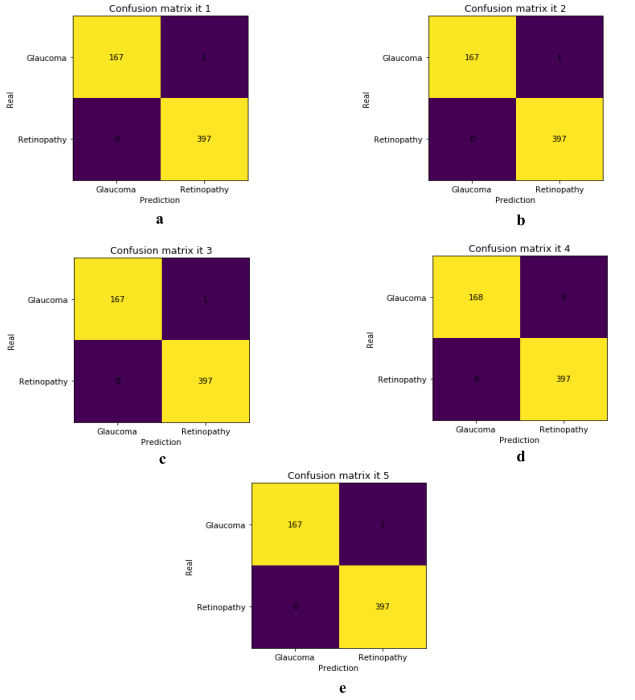

FIGURE 22. Confusion matrices for the five iterations. It can be observed that the results only vary in a single iteration (See image d)), where all the retinographies were correctly classified. However, the number of correctly classified retinographies is very high in the other iterations.

TABLE 4. Accuracy, recall, specificity precision and F_1 score results.

Iteration	Accuracy	Recall	Specificity	Precision	F_1 Score
1	0.998233229	0.9999	1	1	0.9970
2	0.9991166	0.9999	1	1	0.9970
3	0.9991166	0.9999	1	1	0.9970
4	1	1	1	1	1
5	0.998233229	0.9999	1	1	0.9970
Mean	**0.998939931**	**0.9999**	**1**	**1**	**0.9976**

In Table 4, it can be observed the behavior of the CNN with positive (Glaucoma) and negative (Diabetic Retinopathy) instances, Recall and Specificity, respectively. However, the capacity of the classifier to detect positive instances is not perfect, but it can be assumed that this happens because the positive cases belong to the minority class. The results are suitable for both cases. In the same way, the level of precision is very high. Regarding F_1 score values, it is observed that the results almost reach the value of 1.

The results obtained in this work were compared with those of the related work. These results are included in Table 5.

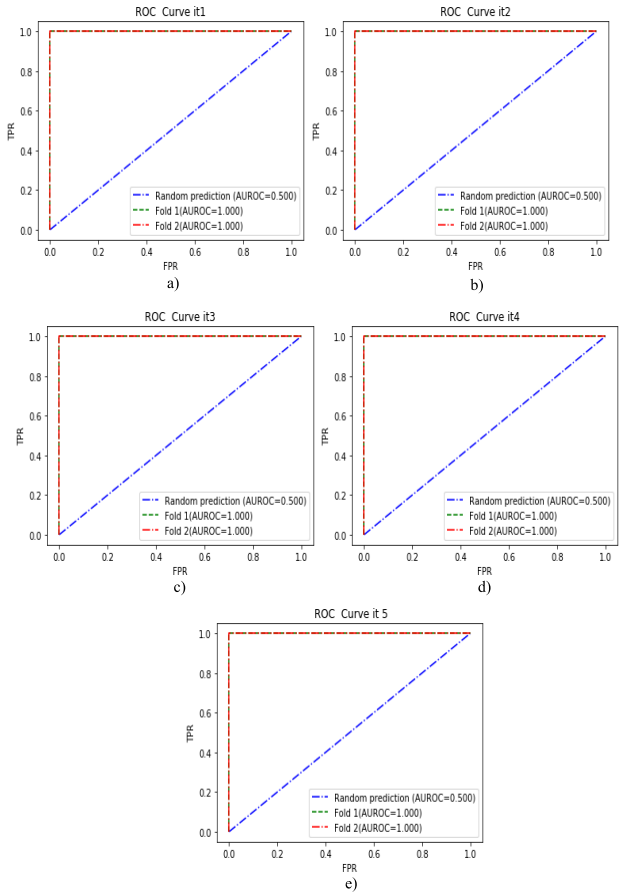

FIGURE 23. ROC curves of the five iterations. Sections a) to e) show the performance is excellent in each of the five iterations. Therefore, CNN ranks the two classes very well above the random decision threshold (0.5).

TABLE 5. Results comparisons. The papers analyze diabetic retinopathy (DR) or glaucoma (G).

Algorithm	Year	Diseases	Accuracy (%)
The proposal	2020	DR and G	**99.89%**
Convolutional Neural Network	2017	DR	94.50%
	2018		99.18%
Decision tree	2019	DR	65.95%
K-Nearest Neighbors	2016	DR	93%
	2019		73.04%
Support Vector Machine	2015	G	94.11%
	2016	G	97%
	2016	DR	93%
	2017	G	91.20%
	2019	G	98.8%
	2019	DR	74.46%
	2020	DR	87.96%
Naïve Bayes	2016	DR	91.4%
	2019		77.85%
Artificial Neural Network	2015	G	89.6%
	2016	DR	65.95%
Multiple inputs – Convolutional Neural Network	2018	G	91.51%
Random Forests	2020	DR	95.08%

The results of the ROC curve can be observed in Figure 22.

The five curves from Figure 22 show the excellent performance of the classifier at each iteration.

V. CONCLUSION

Convolutional Neural Networks are tools for classifying and can help the science of ophthalmology, especially relating to the global problem of Glaucoma and Diabetic Retinopathy.

The proposal shows numerical metrics that are almost perfect. Accuracy, Recall, Specificity Precision, and F_1 score are close to 1. The ROC curves highlight the suitable performance of the proposed classifier.

The result of accuracy was 99.89% and it is the best when compared with the results of the papers in the related work.

The training and validation results show that the training process could stop in epoch three since the learning process is stable related to the training and validation accuracy.

The contributions of this proposal were the improvement of the image by implementing a new channel for the RGB matrix, the classification of two diseases, and a high percentage of model accuracy. Also, the two most crucial eye diseases can be classified, while other papers only work with one condition. In this paper, healthy images were not analyzed because we considered it more important to distinguish between the two incident diseases.

REFERENCES

[1] M. P. Gupta, A. A. Herzlich, T. Sauer, and C.-C. Chan, "Retinal anatomy and pathology," in *Retinal Pharmacotherapeutics*, vol. 55. Basel, Switzerland: Karger, pp. 7–17, 2016.

[2] [Online]. Available: https://www.who.int/publications-detail/world-report-on-vision

[3] A. F. Revuelta, "Técnica de exploración del fondo de ojo," *Actualización en Medicina de Familia*, vol. 8, no. 7, pp. 383–387, 2012.

[4] N. Patil, V. P. Rao, and N. P. Patil, "Hybrid CNN assisted computer aided diagnosis system for glaucoma detection and classification: GlaucoNet+," *Int. J. Innov. Technol. Exploring Eng. (IJITEE)*, vol. 9, pp. 3060–3072, Mar. 2019.

[5] R. O. Bernabé, M. E. Acevedo, and M. A. Acevedo, "Técnicas de segmentación en retinografías," in *Proc. XVIII Congreso Nacional de Ingeniería Electromecánica y de Sistemas*, Mexico City, Mexico, 2019.

[6] *Clinical Practice Recommendations for Managing Diabetic Macular Edema*, International Diabetes Federation, Brussels, Belgium, 2019.

[7] M. Tsukikawa and A. W. Stacey, "A review of hypertensive retinopathy and chorioretinopathy," *Clin. Optometry*, vol. 12, pp. 67–73, May 2020.

[8] K. Adem, "Exudate detection for retinopathy with circular Hugh transformation and convolutional neural networks," *Expert Syst. Appl.*, vol. 114, pp. 285–289, Aug. 2018.

[9] A. Issaz, S. M. Partha, and D. M. Kishore, "An adaptive threshold based image processing technique for improved glaucoma detection and classification," *Comput. Methods Programs Biomed.*, vol. 122, pp. 229–244, Nov. 2015.

[10] A. A. Salam, T. Khalil, M. U. Akram, A. Jameel, and I. Basit, "Automated detection of glaucoma using structural and non structural features," *SpringerPlus*, vol. 5, no. 1, p. 1519, Dec. 2016.

[11] S. S. Rahim, V. Palade, J. Shuttleworth, and C. Jayne, "Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing," *Brain Informat.*, vol. 3, no. 4, pp. 249–267, Dec. 2016.

[12] J. Alves de Sousa, "Texture based on geostatic for glaucoma diagnosis from eye image," *Multimed Tools*, vol. 76, pp. 19173–19190, Apr. 2017.

[13] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, no. 12, p. 2054, Nov. 2017.

[14] Y. Chai, H. Liu, and J. Xu, "Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models," *Knowl.-Based Syst.*, vol. 161, pp. 147–156, Dec. 2018.

[15] K. D. Kirange, "Diabetic retinopathy detection and grading using machine learning," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 6, pp. 3570–3576, 2019.

[16] A. Colomer, J. Igual, and V. Naranjo, "Detection of early signs of diabetic retinopathy based on textural and morphological information in fundus images," *Sensors*, vol. 20, no. 4, p. 1005, 2020.
[17] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[18] B. Bateman, A. R. Jha, B. Johnston, and I. Mathur, *A New, Interactive Approach to Understanding Supervised Learning Algorithms*, 2nd ed. Birmingham, U.K.: Packt Publishing, 2020, ch. 7, pp. 342–346.
[19] T. G. Dietterich, "Approximate statistical tests for comparing supervised classification learning algorithms," *Neural Comput.*, vol. 10, no. 7, pp. 1895–1923, Oct. 1998.

OMAR BERNABÉ was born in Mexico City, Mexico, in 1994. He received the Telecommunications Technician degree from CECyT 11-Wilfrido Massieu, IPN, in 2014, and the Engineering degree in communications from ICE, ESIME-Zacatenco, IPN, in 2017. He is currently pursuing the master's degree in telecommunications with SEPI, ESIME-Zacatenco. He coursed the subject intelligent pattern classification with CIC, ESIME-Zacatenco, in 2019. In 2019, he participated at the CNIES Congress XIX as a Speaker, with his paper "Técnicas de Segmentación en Retinografías," together with the same members' article.

ELENA ACEVEDO received the B.S. degree in engineering with specialization in computing from the Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME), National Polytechnic Institute (IPN), in 1996, the M.Sc. degree in computing from the Centro de Investigación y de Estudios Avanzados (CINVESTAV), in 2001, and the Ph.D. degree from the Center for Computing Research, IPN, in 2006. She has been teaching at ESIME, since 1994. Her research interests include artificial intelligence, digital signal processing, bidirectional associative memories, and machine learning.

ANTONIO ACEVEDO (Member, IEEE) was born in Mexico City, in 1968. He received the degree in communications and electronics engineering, in 1992, the M.Sc. degree in electronic engineering from the Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional, in 1996, and the Ph.D. degree in computer science from the Centro de Investigación en Computación. He is currently a Research Professor with the Master of Science in Telecommunications Engineering, SEPI-ESIME-IPN.

RICARDO CARREÑO received the degree in communications and electronics engineering from the ESIME Zacatenco, IPN, the master's degree in administration from the Autonomous University of Querétaro, and the Ph.D. degree in systems engineering from the Graduate Section SEPI, ESIME Zacatenco. He carried out postdoctoral research stay on unconventional neural networks at the CINVESTAV Center for Research and Advanced Studies, Campus Ticomán, México City. He is currently a Research Professor in the artificial intelligence field with the University of the Istmo in the computer engineering career. His scientific production includes his JCR articles, patents, technical reports, PRODEP projects, and technological prototypes. Among his recognitions are the desirable profile granted by the SEP. He is a member of the National System of Researchers, CONACyT.

SANDRA GÓMEZ received the B.S. degree in communications and electronics engineering from the Escuela Superior de Ingeniería Mecánica y Eléctrica, Instituto Politécnico Nacional (IPN), in 2003, the M.Sc. degree in telecommunications engineering from SEPI ESIME Zacatenco, IPN, in 2008, and the Ph.D. degree in electrical engineering from the Universidad Nacional Autónoma de México, in 2014. She has worked several years as a Bachelor Professor at IPN, UPIITA. Her research interests include digital signal processing (image, video, and audio).