Multi-class Classification Approach for Retinal Diseases

Mario G. Gualsaqui, Stefany M. Cuenca, Ibeth L. Rosero, Diego A. Almeida, Carolina Cadena, Fernando Villalba, and Jonathan D. Cruz *

Universidad de Investigación de Tecnología Experimental Yachay, Urcuquí, Ecuador;
Email: mario.gualsaqui@yachaytech.edu.ec (M.G.G.), stefany.cuenca@yachaytech.edu.ec (S.M.C.), ibeth.rosero@yachaytech.edu.ec (I.L.R.), dalmeida@yachaytech.edu.ec (D.A.A.), ccadena@yachaytech.edu.ec (C.C.), gvillalba@yachaytech.edu.ec (F.V.)
*Correspondence: jcruz@yachaytech.edu.ec (J.D.C.)

***Abstract*—Early detection of the diagnosis of some diseases in the retina of the eye can improve the chances of cure and also prevent blindness. In this study, a Convolutional Neural Network (CNN) with different architectures (Scratch Model, GoogleNet, VGG, ResNet, MobileNet and DenseNet) was created to make a comparison between them and find the one with the best percentage of accuracy and less loss to generate the model for a better automatic classification of images using a MURED database containing retinal images already labeled previously with their respective disease. The results show that the model with the ResNet architecture variant InceptionResNetV2 has an accuracy of 49.85%.**

***Keywords*—retinal diagnosis, Convolutional Neural Network (CNN), deep learning, machine learning, automated diagnosis**

I. Introduction

The retina, like many other parts of the central nervous system, is filled with a variety of different kinds of neurons and it is located at the back of the eye. Also, there are roughly 55 different cell types in mammalian retinas each with a unique purpose [1]. Moreover, the processing of color happens at the retinal cells. Our brain's billions of neurons cooperate to provide the illusion of color perception. There is no color in the real world; rather, color is a projection of neurological processes onto the world we see [2, 3]. Features of light, energy and frequency of vibration or wavelength, are used to generate color. Scientists have long been interested in the puzzle of how our brain separates these two aspects of light, energy and wavelength, before recombining them to create color vision [4].

Several retinal diseases might occur. We are going to review some typical pathologies regarding retinal dysfunction. Also, we will present statistics about the most important diseases.

Age-related Macular Degeneration (AMD), that is defined by a loss of central vision, is the most prevalent cause of severe visual loss in the world. AMD causes advanced aging-related blindness, with over 80% of people diagnosed going blind after the age of 70 [5]. Also, in people over 65, AMD is the most common cause of significant visual impairment. Globally, 25 to 30 million people suffer from severe AMD-related vision loss [6, 7]. Finally, the prevalence of AMD was observed to range from 40% in France to 39% in Germany to 36.3% in the Netherlands to 16.30% in the European North of Russia to 14% in Bulgaria [8].

Diabetic retinopathy is one of the most common disease complications caused by diabetes mellitus. It is one of the most important emerging causes of blindness. Globally, it accounts for about 2.4 million cases of blindness [9]. Moreover, Diabetic retinopathy affects 4.8% of the world's population, whereas only 3 to 4.1% of Europeans have the condition [10]. Recent epidemiological data show that Germany (10.6%) and France (16.6%) have the highest prevalence rates of diabetic retinopathy among people over 60 [11].

Open-angle and angle-closure glaucoma are the two basic kinds of glaucoma, and each can be classed as primary or secondary [12]. Glaucoma affects sixty-seven million people worldwide, 25 million of them reside in Europe. According to estimates, glaucoma affects 12.3% of the world's population and 21.8% of adults in Europe, including 18% of those over the age of 50. Overall, glaucoma is responsible for 5.2 million cases of blindness that represents 15% of global blindness [13]. Now, we will review methods to aid in the detection of retinal pathologies

Computer aided diagnostic and analysis algorithms speed up the processing of big data sets, reduce the time and cost related to visual interpretation, and aid in the early diagnosis and planning of illness treatments [14]. For our project we will use transfer Convolutional Neural Networks with architectures as GoogLeNet, VGG and ResNet. Also, the MuReD dataset contains images of retinal diseases and csv archives to train and classify the dataset.

Manuscript received October 10, 2022; revised November 23, 2022, accepted December 9, 2022; published May 5, 2023.

Artificial Neural Networks (ANNs) are computational processing systems that draw a lot of their inspiration from biological nerve systems, like the human brain. ANNs are primarily made up of a large number of connected computational nodes, also known as neurons, which collaborate in a distributed manner to learn from the input and optimize the final output [15]. Also, supervised and unsupervised learning means that the dataset needs to be preprocessed before training for the supervised learning and for the unsupervised learning, the dataset does not contain any labels. This means that unsupervised learning can learn and differentiate patterns saving time by skipping the preparation and labeling of the images [16].

Similar to conventional ANNs, Convolutional Neural Networks (CNNs) are made up of neurons that learn to optimize themselves. Each neuron will continue to process data and carry out an action. But. the main difference between them is that ANNs struggle with the processing of images [17]. Typical neural networks algorithms can be made using Python with libraries of artificial intelligence. But it is troublesome to work with the process of installing and setting the version of the modules. The solution is to use Google Colab. The site is free and runs Jupiter notes online. Also, it provides free GPU and ram to train and classify large datasets. An important part of the code is the neural network architecture. Some pre-trained neural networks are being used in the project. Some architectures by mention will be Googlenet [18], VGG [19] and Resnet [20]. They can be trained to recognize items from a thousand different categories using an image database

For the diagnosis of retinopathy, fundus imaging is used [21]. Since a few decades the use of artificial intelligence has been extended in the classification of retinopathies. The article by Chaum *et al.* [22] sought to automate diagnosis using clinical metadata and some semantic factors, however this involved a medical evaluation. On the other hand, Imran *et al.* [23] used CAD tools for retinopathy diagnosis.

In the last decades the increase in the use of new machine learning techniques has been evident [24]. For example, in the article by Gabor *et al.* [25] they look at classification by means of neural networks, the paradigm they use is binary classification. Using Bayesian neural networks, they achieve an accuracy of about 99%. Another technique that has been widely used is convolutional networks. So, it has been one of the most widely used because it allows to directly use image and extract features automatically. An example is the article by Pratt *et al.* [26] who by means of a 12-layer convolutional network sought to classify five classes of retinopathy. In this paper a precision of 75% was obtained. Some papers use neural network model transfer techniques, the most common being GoogLeNet, ResNet, Xception, among others [27]. One is the article by Wang *et al.* [28] where he used the following architectures: Alexnet, VGG16 and Inception. In their research they looked for the classification of diabetic retinopathy, so the following results were obtained in the classification of 5 classes: 37.43%, 50.03% and 63.23%, respectively. The paper by Pan *et al.* [29] uses convolutional networks with the following architectures: DenseNet, Resnet and VGG16. In this paper they use AUC as a performance metric, obtaining scores of 0.84, 0.94 and 0.96 in the 5-class classification.

Finally, we mentioned some related works were used different architectures of convolutional network. Most articles seek binary classification. However, more articles on multi-label classification have also appeared in recent years, so that the number of classes commonly used to train neural networks is 5. We will use MuReD dataset that has 2208 images with 20 different labels, with different image quality to train several neural networks.

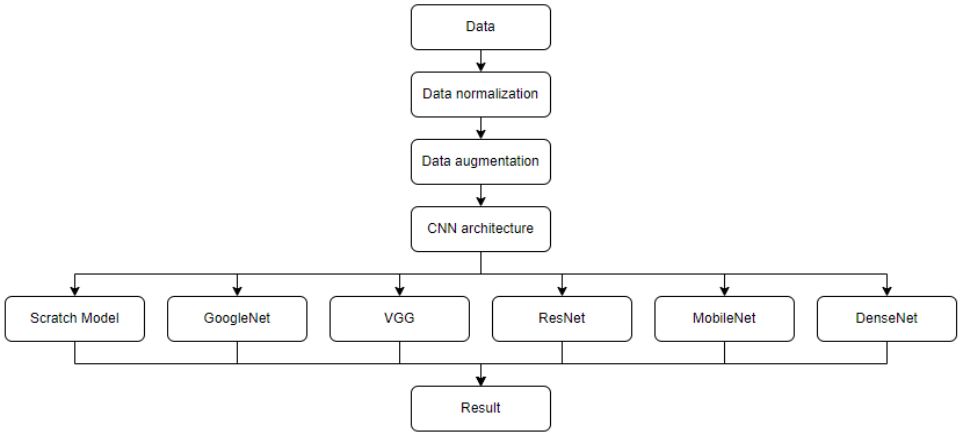

Figure 1. Process flow to transfer models with different architectures.

II. Methodology

A. Dataset Preparation

Fig. 1 shows the process flow, listing the main steps in the execution of the system. Each step was described in detail:

Dataset: Retinal fundus images were obtained from one source. The images were acquired from the IEEE Data Port standard dataset repository. Multi-label Retinal Disease Dataset (MURED) contains 2208 with 20 different labels, their distribution is shown in Fig. 2. Class descriptions and number of samples per label are given in Table I.

The dataset has several retinopathy labels, so the number of images is kept around 2208 images, in order to maintain a balance in the number of images per class. The number of images per class is around 50−150 images per class. Of the total number of images, 75% were divided into 75% for training and 25% for validation. Thus, the following samples were obtained: 1869 images for training and 1869 for validation.

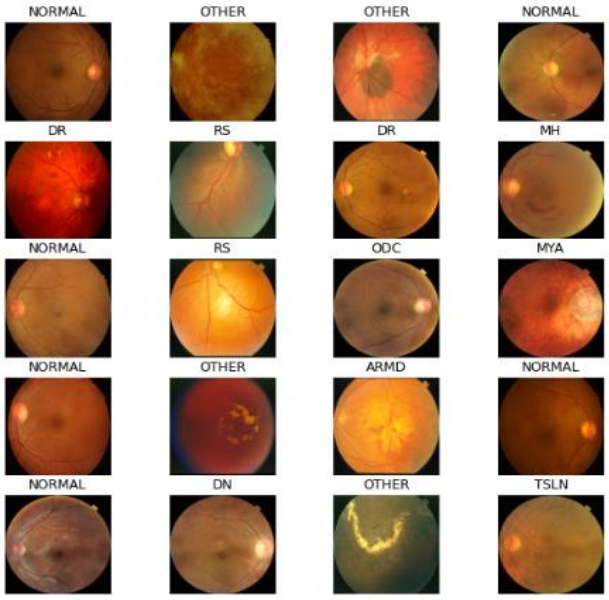

Figure 2. Visualization of the first twenty images of the MURED dataset.

TABLE I. DIVISION OF THE MURED DATASET IMAGES

Acronym	*Full Name*	*Training*	*Validation*	*Total*
DR	Diabetic Retinopathy	396	99	495
NORMAL	Normal Retina	395	98	493
MH	Media Haze	135	34	169
OCD	Optic Disc Cupping	211	52	263
TSNL	Tessellation	125	31	156
ARMD	Age-Related Macular Degeneration	126	32	158
DN	Drusen	130	32	162
MYA	Myopia	71	18	89
BRVO	Branch Retinal Vein Occlusion	63	16	79
ODP	Optic Disc Pallor	50	12	62
CRVO	Central Retinal Vein Occlusion	44	11	55
CNV	Choroidal Neovascularization	48	12	60
RS	Retinitis	47	11	58
ODE	Optic Disc Edema	46	11	57
LS	Laser Scars	37	9	46
CSR	Central Serous Retinopathy	29	7	36
HTR	Hypertensive Retinopathy	28	7	35
ASR	Asteriosclerotic Retinopathy	26	7	33
CRS	Chorioretinitis	24	6	30
OTHER	Other Diseases	209	52	261

Image normalization: the set of images with RGB channel values in the range [0, 255] is not ideal for a neural network, so the values in the range [0, 1] will be standardized.

Image augmentation: The dataset was augmented by operations such as rotating and flipping existing images. The images were rotated to generate new augmented images.

B. CNN Architecture

1) Scratch model

A CNN architecture has been built from scratch using the following procedures. In this regard, a total number of 4 convolutional layers have been built with a varied use of the number and dimension of filters. In the first and second layer, 64 filters with a dimension of 3×3 have been added. The third and fourth layers, the number of filters has been increased to 128 while maintaining the same filter dimension as before. A significant aspect of the convolutional layers is that the number of filters has increased as the layers have gone up. In each layer, a stride dimension of 1×1 is maintained along with a maximum grouping layer of dimension 2×2. Next, the flatten layer has been added to convert the image pixels into a one-dimensional vector. In addition, in each layer, the ReLU (Rectified Linear Unit) function has been imposed as an activation function, following each convolutional layer. Then, a dense layer of 512 neurons has been integrated with a dropout rate of 0.5. The dropout function reduces overfitting and improves the generalization error. Finally, the output layer together with the softmax activation function results in 27 classifications. Therefore, a total number of 13,116,501 weights/parameters have undergone the construction of the entire scratch model. The architecture of the scratch model is illustrated in Fig. 3.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 254, 254, 64)      1792

max_pooling2d (MaxPooling2D  (None, 127, 127, 64)     0
)

conv2d_1 (Conv2D)            (None, 125, 125, 64)      36928

max_pooling2d_1 (MaxPooling  (None, 62, 62, 64)       0
2D)

conv2d_2 (Conv2D)            (None, 60, 60, 128)       73856

max_pooling2d_2 (MaxPooling  (None, 30, 30, 128)      0
2D)

conv2d_3 (Conv2D)            (None, 28, 28, 128)       147584

max_pooling2d_3 (MaxPooling  (None, 14, 14, 128)      0
2D)

flatten (Flatten)            (None, 25088)             0

dropout (Dropout)            (None, 25088)             0

dense (Dense)                (None, 1024)              25691136

dense_1 (Dense)              (None, 21)                21525

=================================================================
Total params: 25,972,821
Trainable params: 25,972,821
Non-trainable params: 0
_________________________________________________________________
```

Figure 3. Visualization and complete information of scratch model.

2) GoogleNet

The Inception V3 architecture comes from the GoogleNet module. Inception version 3 is composed of 42 layers. In this model, only mixed 6, mixed 7, mixed 8, mixed 9, and mixed 10 were trained in the Inception V3 architecture. The components that designed the model were the following. The saliency layer was created, the global spatial averages grouping layer was created and then a layer connecting the number of neurons per layer was added through the ReLU activation function. Then, a dense layer of 1024 neurons with a dropout rate of 0.2 and a logistic layer where the number of labels that the images have, which are 20 classes, are specified. Finally, the output layer together with the softmax activation function with a result of 21 classifications.

3) VGG

There are a total of 13 convolutional layers and 3 fully connected layers in the VGG16 architecture which contains smaller filters (3×3) with more depth instead of having large filters. Another variation is VGG19 which has 19 weight layers consisting of 16 convolutional layers with 3 fully connected layers and the same 5 pooling layers.

A total number of 16 or 19 convolutional layers has been constructed according to the VGG variant to be used. The first two layers were added 64 filters with a dimension of 3×3. Then a clustering layer was used. This is followed by 2 convolution layers where the filters of the two previous layers are added to obtain 128 filters. This pattern is repeated until the convolution layers are completed. At the end of the pattern there is a layer connecting the number of neurons per layer through the ReLU activation function. Then, a dense layer of 1024 neurons with a dropout rate of 0.2 and a logistic layer of 20 classes are integrated. Finally, the output layer together with the softmax activation function with a result of 21 classifications. This procedure is given in the two variations VGG16 and VGG19.

4) ResNet

The ResNet152V2 version and the InceptionResNetV2 module were used.

The input layer has a 256×256×3 image. The feature extraction consists of four CNN blocks. Each of these blocks mainly has a convolution layer, a batch normalization layer, and a ReLU layer. In addition, it has a maximum clustering layer and a dropout layer with a rate of 0.2. These data are used for the two ResNet variables.

5) MobileNet

Mobile models have been built on increasingly efficient building blocks. In MobileNetV3, a combination of layers was used as building blocks to build a more efficient model. The input layer has a 256×256×3 image. Each block consists of an input and an output that have no nonlinearity. These blocks mainly have a convolution layer, a batch normalization layer, and a ReLU layer. A dense layer of 1024 neurons with a rate of 0.2.

6) *DenseNet*

It uses the above data in other architectures such as the number of neurons per layer, the size of the input layer image and the same dropout rate.

DenseNet-201 is a convolutional neural network with 201 layers deep. Where each convolutional layer takes the output of the previous layer forming a characteristic output map. This architecture, allows to connect all the layers taking as name a densely connected convolutional network. Within the model, it is divided into multiple blocks called DenseBlocks where it consists of four blocks with a variable number of layers.

C. Feature Extraction

Retinopathies are diagnosed through images of the fundus of the eye so the architectures to be trained will extract features from the image that deviate from normal. This can be obtained because the image can show changes in color, in the thickness of the blood vessels visible in the eye. Convolution filters allow a better ability to recognize these changes.

D. Metrics

To Evaluate the performance of the models, we take account the time of training, the val_loss and val_accuracy of each model. Then we performed an evaluation of the model using the trained models and test data. The metrics used in the test dataset was precision, recall, F1-score and accuracy. These metrics take into account the TP (True Positive), TN (True Negative), FP (False positive) and FN (False Negative). These metrics are used because they give us a clearer idea of how the retinopathies behave at the time of classification. In this way we can see how far the results are from reality. And thus, see if our model is feasible for medical applications. The metrics are obtained by:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\_score = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

III. Results

A. Training Performance

The training results observed were very important for the selection of the best model. We analyzed accuracy, val accuracy, val loss, number of parameters and training time. The number of parameters viewed was greater than 4 million, being the maximum 161 million parameters and the least 4 million parameters, in models Inception V3 ‘Mixed 6’ and MobileNetV3Large, respectively. On the other hand, the accuracy obtained in the models are higher than 40% in most of them. Being the highest 100% and the lowest 26.08% for models InceptionV3 ‘Mixed 9’ and MobileNetV3Large, respectively. The value of the validation accuracy was the highest in model InceptionResnetV2 and the lowest in model

MobileNetV3Large, 49.85% and 27.43% respectively. The validation loss is an important parameter to analyze the amount of failed data at the time of classification. This metric obtained a minimum in 1.75 and the maximum value was 3.20 in model DenseNet201 and Inception V3 "Mixed 6" respectively. The training times are important to analyze the efficiency of the architecture. A model that converges the solution in the shortest time with the least amount of parameters is desired. Thus, in model MobileNetV3Large a short time of 134s was obtained, in model InceptionResnetV2 a large time of 496s was obtained. Being the maximum in model InceptionResnetV2 and the minimum in model MobileNetV3Large. The values are shown in Table II.

TABLE II. COMPARISON OF PERFORMANCE BETWEEN DIFFERENT CNN ARCHITECTURES

Architecture	Sub-Network	# Parameters	Accuracy	Val_accuracy	Val_loss	Time[s]
Scratch model	-	25,972,821	0.3536	0.3392	2.1462	276.59
Inception V3	"Mixed 6"	160,994,613	0.9118	0.4307	3.2002	259.52
	"Mixed 7"	163,138,485	0.9992	0.4454	2.8888	300.74
	"Mixed 8"	57,883,317	0.8821	0.4661	2.2712	172.38
	"Mixed 9"	91,242,741	**1.0000**	0.4808	2.6250	235.34
	Base	23,922,485	0.5384	0.4454	1.9167	173.13
VGG16	-	15,261,525	0.4510	0.4484	1.8671	457.80
VGG19	-	20,571,221	0.4441	0.4425	1.8780	469.38
ResNet152V2	-	25,684,501	0.8251	0.4897	1.8939	296.24
InceptionResNetV2	-	55,932,149	0.7551	**0.4985**	2.0428	496.15
MobileNetV3Large	-	4,001,941	0.2608	0.2743	2.4110	**134.26**
DenseNet201	-	20,310,613	0.7422	0.4838	**1.7594**	368.79

The number of parameters is an important value to know the weight of the model and thus be able to export it to applications where there is not much memory. In general, you need a model that meets the required specifications regardless of the number of parameters. However, if we need to transfer our model to a mobile application the model MobileNetV3Large would be the best choice. The accuracy value in training is not so important in determining the quality of the model. However, it can help us to see if the model is improving. On the other hand, a high value does not mean that the model is better as it may be in a state of overfitting which makes the model unable to predict images that are not within the training set. One way to solve this is to introduce a validation set in the training set. In this way we obtain a validation accuracy that is usually lower than the training accuracy. This value is important in order to determine the performance of the model. In this situation it is required to have a high value similar to the training set. Here we can visualize that model InceptionResnetv2 obtained the best results in the validation set. The value of the validation loss is very important as it helps us to understand if the model performance really correlates with the validation accuracy of the model. In this case a value of 0 is desired. The minimum value obtained was in model DenseNet201.

Finally, the training time should be measured to see how optimized it is. We want the model to be optimized and to be able to convergence of weights of the neurons faster. In this case the best model was MobileNetV3Large.To obtain a balanced model, the model must have a good validation accuracy, a very low loss and a good time. However, the time can be neglected because it depends on the hardware used. In this case the model with the best results taking into account the two values are the VGG16 architecture.

B. Comparison with Other Works

To compare the performance of this work, it is necessary to identify previous work that includes multiclass classification. However, several challenges were encountered. One of them is the lack of work using the MuReD dataset. The other problem is the comparison of performance with different datasets.

This approach to multiclass classification of retinopathies is new and little explored. Thus, the dataset used in this work has only been worked by Rodriguez *et al.* [30] In their research they used the C-TRAN architecture which consists of a Transformer encoder that feeds from both visual features extracted by a CNN and a set of masked labels.

The MuReD database is formed by the union of several databases such as the ODIR dataset. This dataset consists of a set of images of the eye that have been used to create diagnostic algorithms. Thus, we can find works that generate novel architectures or use common architectures for their own models [31−33]. Some of these models are around 90% accurate. However, the number of classes does not exceed 6. For example, in the Table III. We can observe that there are different architectures that we used in this article, in addition to other architectures obtained from other authors. However, only the article by c-tran [30] made a classifier of 20 labels. The other authors have only made classifiers with a maximum of 5 labels, which is why the results have a higher performance. In our models we had a maximum accuracy of 0.4926 in the inception V3 architecture compared to the performance of the other articles is consistent because we have more classes. The accuracy we have is 0.5573 in the Inception Resnet v2 architecture, this value is very similar to those found in other articles, in fact this value is higher in some occasions, considering that we have a higher number of classes. Also, in the AUC section we also get a high value of 0.8777 which compared to other articles is higher. However, our results show promising values.

TABLE III. COMPARISON OF PERFORMANCE OF OUR MODEL WITH OTHER AUTHORS

Architecture	Sub-Network	Accuracy	Bin accuracy	Precision	F1 - Score	AUC
Scratch model	-	0.2891	0.9548	0.5457	0.1329	0.8676
Inception V3	"Mixed 6"	0.4690	0.9551	0.5504	0.1432	0.8728
	"Mixed 7"	0.4572	0.9553	0.5529	0.1402	0.8710
	"Mixed 8"	0.4690	0.9551	0.5474	0.1402	0.8706
	"Mixed 9"	**0.4926**	0.9552	0.5475	0.1434	0.8712
	Base	0.4513	0.9551	0.5446	**0.1457**	0.8734
VGG16	-	0.3658	0.9551	0.5484	0.1398	0.8725
VGG19	-	0.4307	0.9552	0.5536	0.1349	0.8724
ResNet152V2	-	0.4867	0.9553	0.5560	0.1355	0.8750
InceptionResNetV2	-	0.4454	**0.9556**	**0.5573**	0.1358	**0.8777**
MobileNetV3Large	-	0.2655	0.9555	0.5569	0.1306	0.8743
DenseNet201	-	0.4543	0.9552	0.5536	0.1332	0.8751
C-tran [30]	-	-	-	0.685	0.573	0.962
MCG-Net [31]	-	-	-	0.4564	0.8646	0.6824
MCGs-Net [31]	-	-	-	0.4627	0.8656	0.6983
Vessel Net [32]	-	-	0.7911	-	-	0.8464
EV-GCN+MCED [33]	-	0.8106	-	-	0.8286	0.8472
InceptionV3 [28]	-	0.6323	-	-	-	-
CNN Modified [26]	-	0.75	-	-	-	-
MLC model [34]	-	0.6875	-	0.8156	0.7711	
Transfer model [29]	-	-	-	-	-	0.96

IV. CONCLUSION

In this work, we seek to create a classifier using known algorithms, a new dataset was used for multi-label classification of retinal diseases in fundus images. The dataset used is a dataset containing 2208 samples for 20 different classes. We obtained the performances of several methods proposed in the literature. From the analysis of each architecture, it was found that the VGG16 architecture is the one with the best results considering its accuracy and loss. For future works, more images will be integrated to the dataset and also each architecture will be modified to have better results.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Mario G. Gualsaqui: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition. Stefany M. Cuenca: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Ibeth L. Rosero: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Diego A. Almeida: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Carolina Cadena: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Fernando. Villalba: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. Jonathan D. Cruz: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition. All authors had approved the final version.

ACKNOWLEDGMENT

The authors wish to express their gratitude to the entire team of the Biomedical Engineering program at Universidad Yachay Tech.

REFERENCES

[1] R. H. Masland, "The fundamental plan of the retina," *Nature Neuroscience*, vol. 4, no. 9, pp. 877−886, 2001.

[2] M. A. MacNeil, J. K. Heussy, R. Dacheux, *et al.*, "The shapes and numbers of amacrine cells: Matching of photo filled with Golgi-stained cells in the rabbit retina and comparison with other mammalian species," *J. Comp. Neurol.*, vol. 413, pp. 305–326, 1999.

[3] M. A. MacNeil and R. H. Masland, "Extreme diversity among amacrine cells: Implications for function," *Neuron.*, vol. 20, pp. 971–982, 1998.

[4] P. Gouras, *Color Vision*, University of Utah Health Sciences Center, Salt Lake City (UT), 2011.

[5] D. M. Gohdes, A. Balamurugan, B. A. Larsen, *et al.*, "Age-related eye diseases: An emerging challenge for public health professionals," *Prev. Chronic. Dis.*, vol. 2, p. A17, 2005.

[6] M. Villain, J. P. Nordmann, J. P. Renard, *et al.*, "The French glaucoma and ocular hypertension 1-day study (in French)," *J. Fr. Ophtalmo.*, vol. 29, pp. 520–525, 2006.

[7] L. Verma, T. Das, S. Binder, *et al.*, "New approaches in the management of choroidal neovascular membrane in age-related

macular degeneration," *Indian Journal of Ophthalmology*, vol. 48, no. 4, p. 263, 2000,
[8] E. Prokofyeva and E. Zrenner, "Epidemiology of major eye diseases leading to blindness in Europe: A literature review," *Ophthalmic Research*, vol. 47, no. 4, pp. 171–188, 2012.
[9] J. K. Kristinsson, "Diabetic retinopathy. Screening and prevention of blindness. A doctoral thesis," *Acta Ophthalmol. Scand. Suppl.*, pp. 1–76, 1997.
[10] S. Resnikoff, D. Pascolini, D. Etya'Ale, *et al.*, "Global data on visual impairment in the year 2002," *Bulletin of the World Health Organization*, vol. 82, no. 11, pp. 844–851, 2004.
[11] D. Simmons, G. Clover, and C. Hope, "Ethnic differences in diabetic retinopathy," *Diabetic Medicine*, vol. 24, no. 10, pp. 1093–1098, 2007.
[12] P. Harvey, "Common eye diseases of elderly people: Identifying and treating causes of vision loss," *Gerontology*, vol. 49, no. 1, p. 1, 2003.
[13] C. Nduaguba and R. K. Lee, "Glaucoma screening: Current trends, economic issues, technology, and challenges," *Current Opinion in Ophthalmology*, vol. 17, no. 2, pp. 142–152, 2006.
[14] N. Petrick, B. Sahiner, S. G. Armato III, *et al.*, "Evaluation of computer-aided detection and diagnosis systems[a]," *Medical Physics*, vol. 40, no. 8, 087001, 2013.
[15] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," arXiv preprint, arXiv:1511.08458, 2015.
[16] M. W. Berry, A. Mohamed, and B. W. Yap, *Supervised and Unsupervised Learning for Data Science*, Springer, 2019.
[17] H. H. Aghdam and E. J. Heravi, *Guide to Convolutional Neural Networks*, New York, NY: Springer, 2017, pp. 978–973.
[18] P. Aswathy and D. Mishra, "Deep GoogLeNet features for visual object tracking," in P*roc. IEEE 13th International Conference on Industrial and Information Systems (ICIIS)*, 2018, pp. 60–66.
[19] I. Ha, H. Kim, S. Park, *et al.*, "Image retrieval using BIM and features from pretrained VGG network for indoor localization," *Building and Environment*, vol. 140, pp. 23–31, 2018.
[20] S. Peng, H. Huang, W. Chen, *et al.*, "More trainable inception-ResNet for face recognition," *Neurocomputing*, vol. 411, pp. 9–19, 2020.
[21] E. D. Cole, E. A. Novais, R. N. Louzada, *et al.*, "Contemporary retinal imaging techniques in diabetic retinopathy: A review," *Clinical & Experimental Ophthalmology*, vol. 44, no. 4, pp. 289–299, 2016.
[22] E. Chaum, T. P. Karnowski, V. P. Govindasamy, *et al.*, "Automated diagnosis of retinopathy by content-based image retrieval," *Retina*, vol. 28, no. 10, pp. 1463–1477, 2008.
[23] I. Qureshi, J. Ma, and Q. Abbas, "Recent development on detection methods for the diagnosis of diabetic retinopathy," *Symmetry*, vol. 11, no. 6, p. 749, 2019.
[24] S. K. Padhy, B. Takkar, R. Chawla, *et al.*, "Artificial intelligence in diabetic retinopathy: A natural step to the future," *Indian Journal of Ophthalmology*, vol. 67, no. 7, p. 1004, 2019.
[25] G. M. Somfai, E. Tátrai, L. Laurik, *et al.*, "Automated classifiers for early detection and diagnosis of retinopathy in diabetic eyes," *BMC Bioinformatics*, vol. 15, no. 1, pp. 1–10, 2014.
[26] H. Pratt, F. Coenen, D. M. Broadbent, *et al.*, "Convolutional neural networks for diabetic retinopathy," *Procedia Computer Science*, vol. 90, pp. 200–205, 2016.
[27] I. Kandel and M. Castelli, "Transfer learning with convolutional neural networks for diabetic retinopathy image classification: A review," *Applied Sciences*, vol. 10, no. 6, p. 2021, 2020.
[28] X. Wang, Y. Lu, Y. Wang, *et al.*, "Diabetic retinopathy stage classification using convolutional neural networks," in *Proc. IEEE International Conference on Information Reuse and Integration (IRI)*, 2018, pp. 465–471.
[29] X. Pan, K. Jin, J. Cao, *et al.*, "Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 258, no. 4, pp. 779–785, 2020.
[30] M. A. Rodriguez, H. AlMarzouqi, and P. Liatsis, "Multi-label retinal disease classification using transformers," arXiv preprint, arXiv:2207.02335, 2022.
[31] J. Lin, Q. Cai, and M. Lin, "Multi-label classification of fundus images with graph convolutional network and self-supervised learning," *IEEE Signal Processing Letters*, vol. 28, pp. 454–458, 2021.
[32] D. Luo and L. Shen, "Vessel-net: A vessel-aware ensemble network for retinopathy screening from fundus image," in *Proc. IEEE International Conference on Image Processing (ICIP)*, 2020, pp. 320–324.
[33] Y. Huang and A. Chung, "Edge-variational graph convolutional networks for uncertainty-aware disease prediction," in *Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2020, pp. 562–572.
[34] L. Zhou, X. Zheng, D. Yang, *et al.*, "Application of multi-label classification models for the diagnosis of diabetic complications," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1–10, 2021.

Copyright © 2023 by the authors. This is an open access article distributed under the Creative Commons Attribution License (CC BY-NC-ND 4.0), which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.