International Journal of Computer Engineering and Data Science (ISSN:**2737-8543**)
Volume 1– Issue 1, 20/07/2021

Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks

AATILA Mustapha[1], LACHGAR Mohamed[1], HRIMECH Hamid[2] and KARTIT Ali[1]

[1] LTI laboratory, ENSA, Chouaib Doukkali University, El Jadida 1166, Morocco.

[2]University Hassan 1st, ENSA of Berrechid, LAMSAD, B.P 218, Morocco.

E-mail: mu.aatila@gmail.com , m.lachgar@gmail.com, hamid.hrimech@uhp.ac.ma, alikartit@gmail.com

Article history

Received April 9, 2021
Revised June 30, 2021
Accepted July 03, 2021
Published July 20, 2021

ABSTRACT

Diabetic retinopathy (DR) is considered one of the worldwide diseases of blindness, especially in the elderly. The main reason for this disease is the complication of diabetes in the retinal blood vessels. Usually, the warning signs are not observed. Screening is an important key to diagnosing the early stages of diabetic retinopathy. This work represents an intelligent system of DR classification based on deep learning (DL) tools, especially convolutional neural networks (CNN). Proposed system can assist ophthalmologists to make a preliminary decision, it allows a DR classification considering normal eyes, mild DR, Moderate DR, Severe DR and Proliferative DR. Obtained results, in terms of classification accuracy, for DR classification using the color retinal background images based on VGG-16 and ResNet50 models are in order 70% and 25% respectively.

Keywords: ***Diabetic retinopathy (DR), Deep learning, CNN, VGG-16, ResNet50.***

I. INTRODUCTION

Diabetes, a chronic disease affects various organs in the human body, including the retina. Diabetic retinopathy (DR) results from diabetes mellitus (DM) [1]. DM is the leading cause of blindness among a significant age group in Western countries. It is increasing in underdeveloped countries too. Patients with DM are much more susceptible to blindness than without DM. Progressive diabetic retinopathy and macular edema (clinically significant) can lead to severe vision loss [2]. DR affects a large diabetic population in developed countries, it is a silent disease that only appears in its later stages where treatment is very difficult and, in some cases, impossible. In the case of DR, the blood vessels that help nourish the retina begin to leak fluid and blood onto the retina [3], resulting in visual features called lesions such as microaneurysms, hemorrhages, hard exudates, cotton wool spots, vessel area blood. It can only be treated effectively in its early stages and therefore its early detection is very important through regular screening [4]. Automatic screening is highly necessary so that manual effort is reduced since the cost of this procedure is quite high. Studies have shown that the blinding complications of diabetes can be largely prevented medically, through glycemic and blood pressure control, as well as the early detection and prompt treatment of diabetic retinopathy with photocoagulation/ surgical techniques. Therefore, screening guidelines have been developed by national professional organizations such as the American Diabetes Association (ADA) [5] and American Academy of Ophthalmology Retina Panel (AAORP) [6]. Adults and children over 10 years of age with type 1 diabetes should have an initial dilated examination by an ophthalmologist within 5 years of the onset of diabetes [7]. Since people may already have type 2 diabetes before they are aware of symptoms, and up to 20% of patients with type 2 diabetes have retinopathy at the time of diagnosis, they should have an initial dilated exam with an ophthalmologist when diagnosed with diabetes [8]. Subsequent examinations should be annual or more frequent if retinopathy progresses. Pregnant women with pre-existing diabetes should have a dilated eye exam early in the first trimester of pregnancy, as pregnancy may potentiate the rapid progression of retinopathy. Close monitoring should continue throughout pregnancy and 1 year after childbirth. DR is detected by the appearance of different

types of lesions on an image retina. These lesions are microaneurysms (MA), hemorrhages (HM), soft and hard exudates (EX) [9][10][11].

• Microaneurysms (MA) are the first sign of DR that appears as small red round dots on the retina due to weak vessel walls. The size is less than 125 µm and the margins are sharp. In [12], the MA was classified six types as shown in figure 1. Types of MA were observed with AOSLO reflectance and conventional fluorescein imaging.

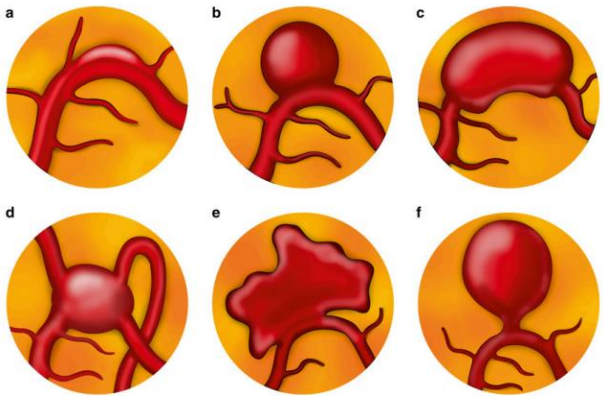

Figure 1. Focal bulge (a), Saccular (b), Fusiform (c), Mixed Saccular/Fusiform (d), Irregular (e) and Pedunculated (f) types of MA [13].

• Hemorrhages (HM) appear as larger spots on the retina, where its size is greater than 125 µm with an irregular margin. There are two types of HM, which are flame (shallow HM) and blot (deeper HM), as shown in figure 2.

• Hard exudates appear as bright yellow spots on the retina caused by plasma leakage. They have sharp margins and can be found in the outer layers of the retina as illustrated in figure 2.

• Soft exudates (also called cotton wool) appear as white spots on the retina caused by swelling of the nerve fiber. The shape is oval or round as shown in the figure 2 bellow.

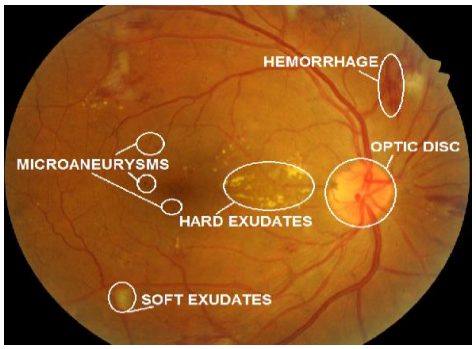

Figure 2. Retinal fundus image, illustrating examples of lesions types [14].

According to the International Diabetic Retinopathy Clinical Study (IDRC) and the Diabetic Retinopathy Early Treatment Research Group (DRETR), the various severity levels of DR are defined in the table I as follows:

TABLE I. DR SEVERITY SCALES.

Stage	Dilated Ophthalmoscopy Observable Signs	Severity
I	No Signs	No DR
II	Micro-aneurysms	Mild DR
III	Micro-aneurysms, retinal dot and blot hemorrhages, hard exudates or cotton wool spots **without** signs of severe non-proliferative diabetic retinopathy	Moderate DR
IV	More than 20 intra-retinal hemorrhages in each of 4 quadrants, definite venous beading in 2 or more quadrants, prominent intra-retinal microvascular abnormality (IRMA) in 1 or more quadrants **without** signs of proliferative retinopathy	Severe DR
V	Neovascularization or Vitreous/pre-retinal hemorrhage	Proliferative DR

II. RELATED WORKS

In the literature, various techniques have been used to build intelligent systems for diabetic retinopathy detection and classification. Several works which deal with the main subject of this study are analyzed and evaluated. Authors of [15] have developed an intelligent diagnosis system to assist ophthalmologists in screening, evaluation and treatment of DR. Trained on a dataset on of 75137 fundus images of eyes, this CNN-based proposed system provided a sensitivity and specificity in order of 94% and 98% respectively. In [16], authors proposed a CNN approach based on the analysis of color fundus retinal images to automate the screening of DR. The proposed system achieved an accuracy of 95% and 85% for the two class and five class classification respectively on the Kaggle dataset (over than 30000 images). In [17], the authors developed an intelligent system, based on Fuzzy C Means (FCM) clustering technique to segment the candidate region areas in order to detect and classify DR. Proposed system provided an accuracy varying between 82.53% and 97.05%, a sensitivity varying between 74.28% and 97.14% and a specificity between 78.57% and 100% on a dataset of 140 retinal images. Trained on a dataset of 9939 images, the system proposed in [18] provided an accuracy between 81% and 96%, this proposed system is a modified GoogLeNet deep CNN. Authors of [19] developed a DR classification system based on SVM, KNN, Naïve bayes, Bayesian logistic regression and Ensemble Bagged Tree. This proposed system achieved an accuracy of 98.50% on a dataset of 1200 retinal images. Based on neural network with moth search optimization (DNN-MSO) method, the proposed approach in [20] allowed the detection and the classification of DR with an accuracy of 99.12%, a sensitivity of 97.91% and a specificity of 99.47% on a dataset of 1200 images and the authors of [21] developed a CNN-based system of DR classification. Based on AlexNet, VGG16, and InceptionNet V3, proposed system achieved an accuracy of 37.43%, 50.03% and 63.23% for AlexNet, VGG16, and InceptionNet V3 respectively. The table II below shows the comparison between 5 selected items:

TABLE II. RELATED WORKS SUMMARY

Authors	year	Method	Dataset	Acc	Sens	Spec
Rishab G et al.[15]	2017	CNN	75137 images	N.A	94%	98%
Ghosh R et al. [16]	2017	CNN	3000 images	85% to 95%	N.A	N.A
Saha R et al. [17]	2016	FCM	140 images	82.53 % to 97.05 %	74.28 % to 97.14 %	78.57% to 100 %
Takahashi H et al. [18]	2017	Modified GoogLe Net deep CNN	9939 images	81% to 96%	N.A	N.A
Amine J et al. [19]	2017	SVM, KNN, NB, Bayesian LR and Ensemble Bagged Tree	1200 images	98.5%	N.A	N.A
Shankar K et al. [20]	2020	DNN-MSO	1200 images	99.12 %	97.91 %	99.47%
Wang X et al. [21]	2018	CNN	Kaggle	37.43 % to 63.23 %		

III. METHODOLOGY AND MATERIALS

A. *Data description*

Retinal fondus images are captured by an optical system known as a background camera, a combination of camera attached with a low-power microscope. It can be used simultaneously to illuminate the retina and its imaging. It was designed to image the inside of the eye primarily the retina, optic disc, macula and posterior pole.

The present works consists of the DR classification using the Diabetic Retinopathy Detection 2015 challenge dataset available in (https://www.kaggle.com/sovitrath/diabetic-retinopathy-2015-data-colored-resized). Retinal images were provided by EyePACS, a free retinopathy screening platform. There is a large set of high-resolution retinal images taken under various imaging conditions. A left and right field is provided for each subject. Images are tagged with a subject id as well as either left or right (i.e 1_left.png is the left eye of patient id 1). The dataset consists of 35126 retinal images to detect diabetic retinopathy and all images are resized into 224x224 pixels. A clinician rated the presence of diabetic retinopathy in each image on a scale of 0 to 4, according to the scale mentioned previously. The number of images for each stage of infection is presented in the table III bellow:

TABLE III. DESCRIPTION OF THE USED DATASET.

Class	Stage	Number of Images	Size in percentage
0	Normal eyes	25810	73%
1	Mild DR	2443	7%
2	Moderate DR	5292	15%
3	Severe DR	873	2%
4	Proliferative DR	708	2%

The figure3 bellow represents the repartition of images with respect to different DR scale.

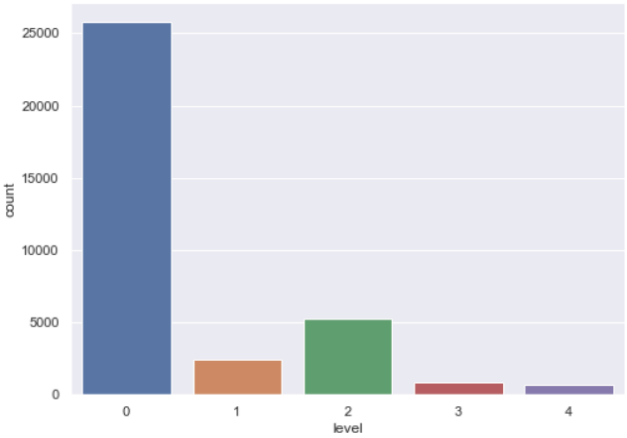

Figure 3. Number of images by DR stages.

The images in this dataset were from different models and types of cameras and were of very mixed quality. There was noise in the pictures and the labels. Some images contained artifacts - were out of focus, underexposed, or overexposed. A major objective of this project was to develop an algorithm that can work in the presence of noise and variations.

Currently, detection of DR is a manual and time-consuming process that requires a trained clinician to review and assess digital color photographs of the fundus of the retina. By the time human readers submit their reviews, often a day or two later, delayed results lead to loss of follow-up, poor communication, and delayed processing. Globally, DR causes 2.6% of blindness. With color fundus photography as an input, the goal of this study is to develop an automated detection and classification system of DR.

B. *Methodology*

Artificial neural networks have already been widely applied in medical imaging, but a special type of neural networks known as Deep Networks, especially CNN (Convolution Neural Networks), produce impressive results in extraction and classification automatic characteristics. They have already proven to be surprisingly great at recognizing handwritten characters. Deep networks are very robust using techniques like dropout which helps the network to produce correct results even if some functionality is missing in the test data. Additionally, ReLU (Rectified Linear Units) is used as transfer

functions in deep CNNs which help in effective training as they do not disappear in extremes. Finally, with dozens of layers in deep networks, they can be easily parallelized with GPUs. The model of the proposed system is presented in the figure 4 bellow.

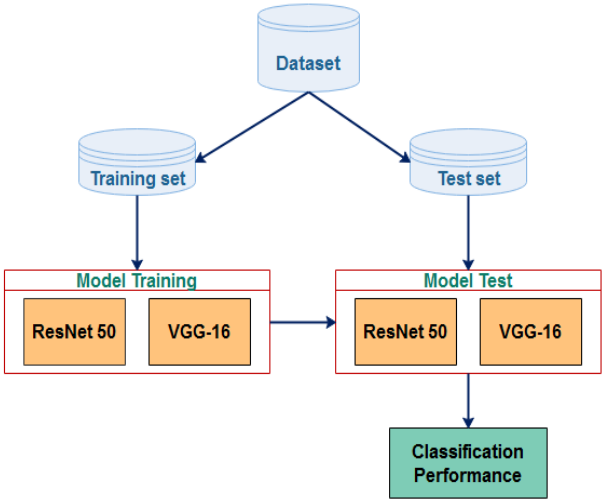

Figure 4. Classification Model.

Convolutional Neural Networks (CNN), a subclass of deep networks is a direct-feedback multilayer neural network trained with a back propagation method that automatically extracts features from images and therefore a rigorous step of extracting and selecting from characteristics is avoided using CNN. It acts as a feature extraction layer and generates both low level and high-level features in its different layers. For feature extraction and selection two layers are used with a final fully connected layer to refine the entire network.

Subsampling or pooling layer acts as a selection of features that helps reduce the spatial resolution of each separate feature map element and makes the network robust to changes in image size, orientation, background, etc.

The VGG-16 is a version of the popular convolutional neural network called VGG-Net. VGG-16 consists of several layers, including 13 convolutional layers and 3 fully connected layers. He must therefore learn the weights of 16 diapers. It takes as input a color image of size 224××224 px and classifies it in one of the 1000 classes. It therefore returns a vector of size 1000, which contains the probabilities of belonging to each of the classes. The architecture of VGG-16 is illustrated in the figure 5 below.

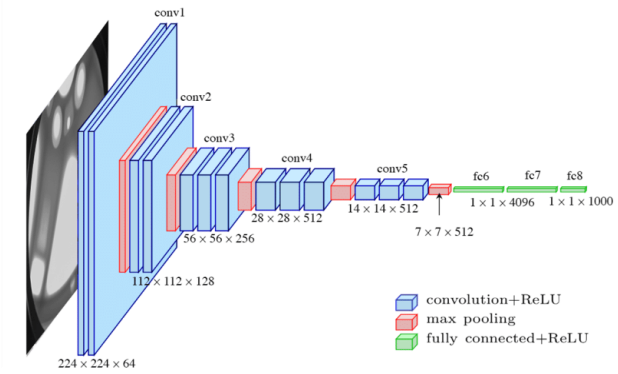

Figure 5. VGG-16 architecture [22].

Each convolutional layer uses color filters of size 3××3 px, moved with a step of 1 pixel. The zero-padding is 1 pixel so that the input volumes have the same output dimensions. The number of filters varies depending on the "block" in which the layer is located. In addition, a bias parameter is introduced into the convolution product for each filter.

Each convolutional layer has the function of activating a ReLU. In other words, there is always a ReLU correction layer after a convolution layer.

The pooling operation is performed with cells of size 2××2 px and a step of 2 px, so the cells do not overlap.

The first two fully connected layers each compute a vector of size 4096 and are each followed by a ReLU layer. The last returns the probability vector of size 1000 (the number of classes) by applying the Softmax function. In addition, these three layers use a bias parameter for each element of the output vector. In our case, we kept the same architecture with the addition of Softmax as an activation function.

The ResNet50 or Residual Network which was developed by Microsoft and winner in 2015 of the annual ILSVRC (ImageNet Large Scale Visual Recognition Competition) competition. ILSVRC is an annual competition where several teams compete to come up with the best algorithm for computer vision tasks.

ResNet50 is composed of 50 layers deep and over 25.6 million parameters. The set is a combination of convolution, identity block (input = output), and a fully connected layer. The identity x corresponds to the input value of the original block or signal. The output value of the residual block is therefore the sum of the input value of the block with the output values of the internal layers of the block. The figure 6 bellow illustrates the ResNet50 architecture.

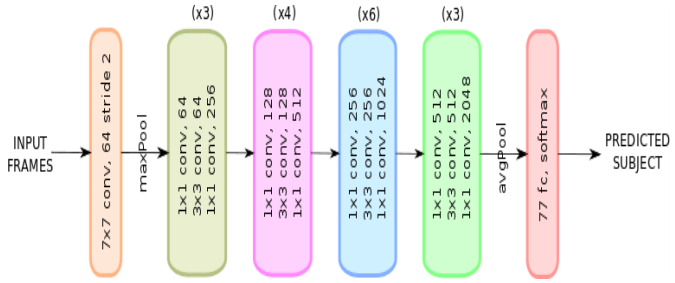

Figure 6. ResNet50 architecture [23].

C. *Performance Evaluation Measure*

In the field of health, data used in medical treatment is generally divided into two categories; one belongs to data with disease and others without disease. To evaluate performance of proposed systems, the metric used is the classification accuracy.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Where, True Positive (PT) is the number of images of diseases classified as disease, True negative (TN) is the number of normal images classified as normal , False Positive (FP) is the number of normal images classified as disease and False negative (FN) is the number of diseases of the images classified as normal.

IV. Simulation Results And Discussion

All the experiments were developed in Keras and trained using Intel Core i5-6300U CPU on 64-bit Window 8.1 OS. This section explores the results obtained using the two pre-trained networks VGG-16 and ResNet50 on the Diabetic Retinopathy dataset of Kaggle. A total of 35126 retinal fundus images are divided into training and test datasets representing 70% and 30% respectively of the original dataset.

A. *VGG-16 Results*

Training and validation accuracy obtained using VGG-16 over 15 epochs are illustrated in the figure 7 bellow.

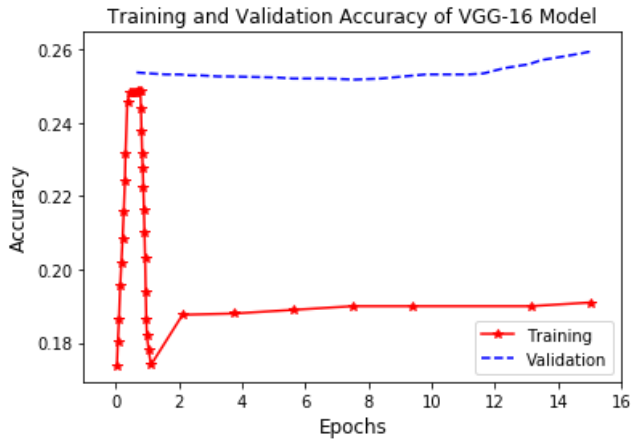

Figure 7. Tarining and validation accuracy of VGG-16.

The validation accuracy is almost stable during all epochs of classification, while the training accuracy stabilizes after a certain number of epochs.

The figure 8 bellow presents the training and the validation loss function of the VGG-16 model classification over the 15 epochs.

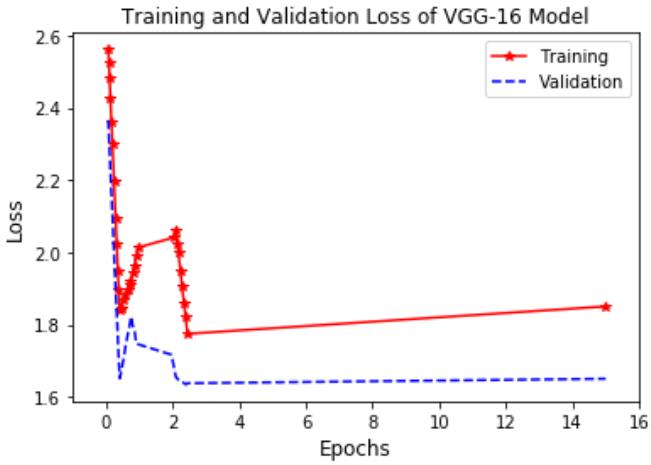

Figure 8. VGG-16 loss in training and validation steps.

The loss function decreases significantly after the first epoch during both validation and training phases.

The confusion matrix of the DR classification using the VGG-16 model is illustrated in the figure 9 bellow in order to describe the full performance of the VGG-16 model.

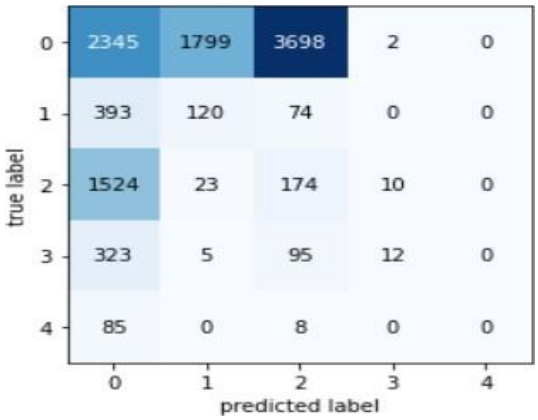

Figure 9. Validation confusion matrix of the VGG-16 model.

According to the VGG-16 model matrix, the model made the majority of predictions on class 0 (Eyes with No Diabetic Retinopathy) and no prediction on class 4 (Proliferative DR).

B. *ResNet-50 Results*

The figure 10 bellow represents the training and validation accuracy using ResNet50 model over 15 epochs.

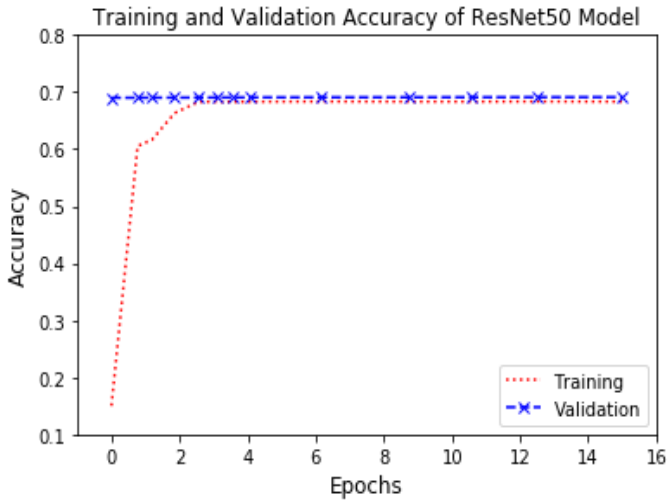

Figure 10. Tarining and validation accuracy of ResNet.

Accuracy in the validation phase remains at 70% during all epochs on the contrary in training it is not stable until epoch 2 as shown in the figure 10.

The figure 11 bellow illustrates the training and validation loss of the ResNet50 model over 15 epochs.

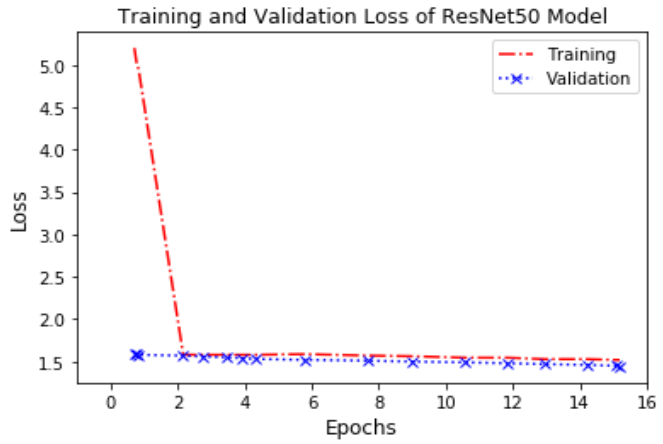

Figure 11. RestNet loss in tarining and validation steps.

The loss function is stable in the validation phase and in training it decreases from 5.5 in epoch 1 to 1.52 during the second epoch.

To best describe the full performance of the ResNet model, the confusion matrix of the DR classification using this model is presented in the figure 12 bellow.

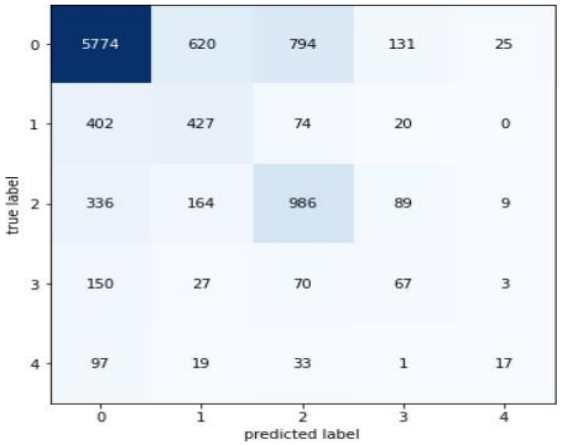

Figure 12. Validation confusion matrix of ResNet model.

ResNet50 model performed a number of predictions over all DR classes. In this case, the model made 7271 correct predictions this number represents the sum of the diagonal this matrix.

The table IV bellow represents the performance of different models used in this study.

TABLE IV. COMPARISON OF VGG-16 AND RESNET50 METRICS.

Algorithms	**Train loss**	**Train accuracy**	**Test loss**	**Test accuracy**
VGG-16	2.04	0.22	1.81	0.25
ResNet-50	2.16	0.7	1.52	0.70

Generally, ResNet50 model provided a good performance compared to the VGG-16 model in the DR Classification.

As shown in the table IV, ResNet50 model allows Diabetic Retinopathy detection and classification in different stages with an accuracy of 70% and a loss in order of 1.52, while VGG-16 model achieves an accuracy of DR classification in order of 25% and a loss around 1.81. However, ResNet50 is capable of detecting and classifying different stages of DR with a better accuracy than the VGG-16 model providing a bad classification performance. The score obtained by ResNet50 is almost three times that of VGG-16 which is normal according to the performance achieved in the training and validation phases and according to the nature of the ResNet50 model which belongs to the class of network-in-network architectures (very deep learning architecture) contrary to the classical and sequential architectures of VGG-16 model. Very deep learning architecture was introduced because the traditional models like the VGG-16 reached certain limits like in the evanescence of the gradient.

V. CONCLUSION

Automated screening systems dramatically reduce the required diagnosis time, save effort and costs for ophthalmologists, and result in rapid treatment of patients.

However, automated systems have an important role in detecting DR at an early stage.

This work made a comparison between VGG-16 and ResNet-50 networks in the classification of DR, studied models provided an accuracy of 25% and 70% using VGG-16 and ResNet50 respectively. ResNet50 provided the best performance hence the usefulness of this architecture in images classification, especially for systems involving a large number of images.

REFERENCES

[1] Chan. J.C.Y, Chee. M.L, Tan. N.Y.Q. et al., "Differential effect of body mass index on the incidence of diabetes and diabetic retinopathy in two Asian populations," Nutr & Diabetes, vol. 8, 2018.

[2] Rishi P. Singh, Michael J. Elman, Simran K. Singh, Anne E. Fung and Ivaylo Stoilov, "Advances in the treatment of diabetic retinopathy," Journal of Diabetes and its Complications, vol. 33, 2019.

[3] Suvajit Dutta, Bonthala CS Manideep, Syed Muzamil Basha, Ronnie D. Caytiles and N. Ch. S. N. Iyenga, "Classification of Diabetic Retinopathy Images by Using Deep Learning Models," International journal of Grid and Distributed Computing, vol 11, pp. 89-106, 2018.

[4] Rahim S.S, Palade V, Almakky. I and Holzinger. A, "Detection of Diabetic Retinopathy and Maculopathy in Eye Fundus Images Using Deep Learning and Image Augmentation," In: Holzinger A., Kieseberg P, Tjoa A., Weippl E. (eds) Machine Learning and Knowledge Extraction. CD-MAKE 2019. Lecture Notes in Computer Science, vol 11713. Springer, Cham.

[5] American Diabetes Association: Standards of medical care in diabetes—2009. Diabetes Care 32 (Suppl. 1): S13-S61, 2009

[6] American Academy of Ophthalmology Retina Panel: Preferred practice pattern guidelines: diabetic retinopathy. San Francisco, American Academy of Ophthalmology, 2008. Available online from http://one.aao.org/CE/PracticeGuidelines/PPP_Content.aspx?cid=d0c853d3-219f-487b-a524-326ab3cecd9a

[7] Bruggeman. B, Zimmerman. C, LaPorte. A et al., "Barriers to retinopathy screening in youth and young adults with type 1 diabetes," Pediatr Diabetes, p. 1– 5, 2020.

[8] Klein. R, Klein. BE, Moss. SE, Davis. MD and DeMets. DL, "The Wisconsin Epidemiologic Study of Diabetic Retinopathy. III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years," Arch Ophthalmol, vol. 102, p. 527-532, 1984.

[9] Taylor. R and Batey. D, "Handbook of retinal screening in diabetes: diagnosis and management," second ed. John Wiley & Sons, Ltd Wiley-Blackwell, 2012.

[10] E. T. D. R. S. R. GROUP, "Grading diabetic retinopa thy from stereoscopic color fundus photographs- an extension of the modified Airlie House classification," Ophthalmology, vol. 98, p. 786–806, 1991.

[11] Scanlon. PH, Wilkinson. CP, Aldington. SJ and Matthews. DR, "A Practical manual of diabetic retinopathy management," first ed. Wiley-Blackwell, 2009.

[12] Dubow. M, et al., "Classification of human retinal microaneurysms using adaptive optics scanning light ophthalmoscope fluorescein angiography," Investig Ophthalmol Vis Sci, vol. 55, p. 1299– 1309, 2014.

[13] Ziemssen. F and Agostini. H. T, "Diabetic Retinopathy," Essentials in Ophthalmology, p. 89–130, 2016.

[14] Borsos. B et al., "Automatic detection of hard and soft exudates from retinal fundus images," Acta Universitatis Sapientiae, Informatica, vol. 11, p. 65 – 79, 2019.

[15] Gargeya. R. and Leng. T, "Automated Identification of Diabetic Retinopathy Using Deep Learning," Ophthalmology, vol. 124, p. 962–969, 2017.

[16] Ghosh. R, Ghosh. K and Maitra. S, "Automatic detection and classification of diabetic retinopathy stages using CNN," 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2017, p. 550-554.

[17] Saha. R, Chowdhury. A. R and Banerjee. S, "Diabetic Retinopathy Related Lesions Detection and Classification Using Machine Learning Technology," Lecture Notes in Computer Science, 2016, p. 734–745.

[18] Takahashi. H, Tampo. H, Arai. Y, Inoue. Y and Kawashima. H, "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy," PLoS ONE, vol. 12, 2017.

[19] Amin. J, Sharif. M, Yasmin. M, Ali. H and Fernandes. S. L, "A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions," Journal of Computational Science, vol. 19, p. 153–164, 2017.

[20] Shankar. K, Perumal. E and Vidhyavathi. R.M, "Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images," SN Appl. Sci., 2020.

[21] Wang. X, Lu. Y, Wang. Y and Chen. WB, "Diabetic retinopathy stage classification using convolutional neural networks," In: International Conference on information Reuse and Integration for data science, 2018. p. 465–71.

[22] Nash. W, Drummond. T and Birbilis. N, "A review of deep learning in the study of materials degradation," Npj Materials Degradation, vol. 2, 2018.

[23] S. Jahromi, M. N., Buch-Cardona, P., Avots, E., Nasrollahi, K., Escalera, S., Moeslund, T. B., and Anbarjafari, G, "Privacy-Constrained Biometric System for Non-Cooperative Users," Entropy, vol. 21, 2019.