Transfer Learning based Diabetic Retinopathy Detection with a Novel Preprocessed Layer

Md. Robiul Islam
dept. of Electrical & Computer Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
robiulruet00@gmail.com

Md. Al Mehedi Hasan
dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
mehedi_ru@yahoo.com

Abu Sayeed
dept. of Computer Science & Engineering
Rajshahi University of Engineering & Technology
Rajshahi, Bangladesh
Abusayeed.cse@gmail.com

***Abstract*— one of the major reasons for impaired vision in the world nowadays is diabetic retinopathy (DR). Many people could be saved from permanent blindness with early detection. The manual diagnosis is erroneous and tedious. Hence, numerous computerized vision methods for the automatic detection of diabetic retinopathy and its distinctive stages from retinal images were proposed. Various image processing techniques have been developed besides deep learning methods. In image processing techniques, complex features are manually identified. Most of the earlier works used very small dataset which has a great chance to be over-fitting and worked with grayscale image after transforming color fundus images. In our paper, we developed a deep learning model with transfer learning from VGG16 model followed by a novel color version preprocessing technique. It reduced the training time and provided an average accuracy of 0.9132683 implemented to new Kaggle dataset "APTOS 2019 Blindness Detection". Moreover, to avoid the over-fitting problem for long run we used Stratified K-fold cross validation.**

Keywords— diabetic retinopathy, transfer learning, deep learning, Convolution neural Network

I. INTRODUCTION

Individuals having diabetes could have a disease of eye referred to as diabetic retinopathy. It is the underlying source of visual impairment throughout the world today. The US Center for Disease Control and Prevention reports that there are 29.1 million people living with diabetes in the US and the World Health Organization estimates that there are 347 million human beings living with the disease around the world. Diabetic Retinopathy (DR) is a disease of eye that has long been associated with diabetes. This often occurs when elevated levels of blood sugar inside the retina cause damage to the blood vessels. Such blood vessels may swell and leak, or close, blocking the blood flow. Anomalous new blood vessels often develop on the retina. All these changes can make our vision take hold. In general, (i) non-proliferative diabetic retinopathy (NPDR) and (ii) proliferative diabetic retinopathy (PDR) are two major stages of diabetic retinopathy. NPDR is the primary stage of eye disease for diabetics. Many people suffering from diabetes have this. In case of NPDR, there are leakages from tiny blood vessels, thereby causing a swell in the retina. It is known as macular edema, when the macula swells. It's the most typical reason people suffering from diabetes lose their sight.

NPDR also is split into mild, moderate and severe phases [8]. In the mild stage, there is a small circular red dot called micro-aneurysm (MA) at the terminal point of blood vessels. The MAs rapture into the deeper layers in the moderate stage and create a flame-like shape called hemorrhage in the retina [10]. In each of the four quadrants, the severe stage holds more than 20 intra-retinal hemorrhages, possessing venous beading in one or more quadrants with leading intra-retinal micro-vascular abnormality (IRMA) [4]. In NPDR, Blood vessels from the inside of the retina can further block off. That is known as macular ischemia. Whenever this occurs, the macula cannot be penetrated by blood. Exudates are some small particles that may sometimes develop in the retina. Our vision can be affected by it as well.

PDR is a far more extreme situation of eye disease in diabetes. It originates with the beginning of new blood vessels developing in the retina. This is well-known as neovascularization. Such delicate new-fangled vessels can bleed inside the vitreous. Some dark floaters appear when they bleed just a small amount. If they bleed very much, this could obstruct all visibility. Such new-fangled blood vessels may be producing scar tissue may cause the macula problem or may lead to a detached retina.

Patients may not have any signs and symptoms at the early stages of development of DR, but as the patient's condition headways, spots or shady strings moving in the patient's vision (floaters), blurred vision, fluctuating vision, impaired vision, dark or purging areas appear gradually, finally loss of vision. Therefore, detecting the DR in the early stages is the absolute necessity to avoid the more horrifying impact of the latter stages.

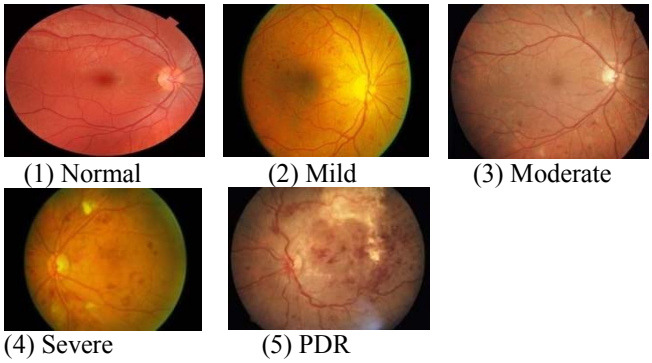

Fig. 1. Different stages of Diabetic Retinopathy [9]

Diabetic retinopathy is detected using colored fundus images. Highly skilled domain experts are needed for the manual analysis and hence is time-consuming and expensive.

978-1-7281-7366-5/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Carleton University. Downloaded on May 27,2021 at 22:11:08 UTC from IEEE Xplore. Restrictions apply.

It is then worth noting the use of computerized and precise systems for the automatic analysis of the fundus photographs and assist physicians/clinicians [10].

Various machine learning approaches such as Decision Tree, Logistical Regression, K-nearest Neighbor (K-NN), Random Forest, Support Vector Machine (SVM) etc. were used to identify diabetic retinopathy automatically [8]. The key problem of such conventional machine learning methods is that they need to manually extract the complex features using image processing strategies from the fundus photographs. But retrieving all of the complex features is troublesome from the fundus images. And more often the accuracy of the diagnosis is poor.

In [1], E. V. Carrera et al. utilized image processing technique to separate blood vessels, micro-aneurysms and hard exudates for extracting the features that was given to support vector machine (SVM) to grade DR. Here, the massive challenge for certain retinal images was the failure to detect micro-aneurysms. That was time consuming too.

Convolution Neural Network (CNN) has recently showed significant performance in image recognition and computer vision.

In [2], D. U. N. Qomariah et al. applied transfer learning technique and made use of the high-level features from the final fully connected layer of the Convolution Neural Network (CNN) as input features for the classification using support vector machine. The issue was that they only classified either typical or severe NPDR classes. They ignored all the remaining early stages before the severe NPDR. Another problem was that they utilized only 147 fundus images from Messidor database which has a great chance of over-fitting. They got the highest accuracy of 95.83% from VGGNet type 19 and 90.48% from VGGNet type 16.

In [3], P. Khojasteh et al. introduced a novel CNN architectural framework (four convolution, three max pooling and three standardization layers with SDG optimizer) by incorporating a preprocessing layer embedded before the first convolution layer. Two enhancement techniques- (a) Contrast Enhancement (CE) and Contrast-Limited Adaptive Histogram Equalization (CLAHE)-were used for pre- processing. Their model got an accuracy of 87.6% and 83.9% for CE and CLAHE respectively. The problem was that they used only 89 color fundus images to be over-fitting.

A large and new Kaggle dataset was classified in this paper based on severity, and our proposed system classified all DR stages. Our proposed system used a customized deep learning architecture with a better color version preprocessing technique. The remaining sections are organized as follows: Section II describes our working procedure. Section III presents the results and analysis. Section IV gives the summary.

II. PROPOSED METHODOLOGY

A. Dataset Description

We used the Kaggle latest dataset "APTOS 2019 Blindness Detection" provided by Aravind Eye Hospital to diagnose and prevent the diabetic retinopathy disease amongst rural people where it is tough to perform medical screening. The dataset contains 5590 high resolution fundus images. We used 3662 images for training purposes and 1928 images for testing purposes. Each image was scored on a scale of 0 to 4 for the severity of the diabetic retinopathy: 0 meant for No DR, 1 meant for Mild, 2 meant for Moderate, 3 meant for Severe, 4 meant for Proliferative DR.

TABLE I. TRAINING SET DISTRIBUTION

Training Set	
Classes (Severity)	***No. of samples***
0 (No DR)	1805
1 (Mild)	999
2 (Moderate)	370
3 (Severe)	295
4 (PDR)	193

B. Image Pre-processing

As images were collected from distinctive sources, they came with a variety of lighting conditions. Some of the images are exceptionally dark and difficult to visualize. Sometimes miscellaneous color illumination is confusing. Another problem is that, for some images, we can get the uninformative dark areas as shown in Fig. 1(3). So, in the second case, it's instinctive to crop out the uninformative areas. This is important because when we reduce the size of the image, informative areas become too small.

To begin with, we cut off the uninformative dark regions from the images using the crop circle. Here the cropped image is a color image, e.g. we did a color version of the crop circle because the color image gives better representation details than the gray scale image in this problem. Most of the previous works used changed over the grayscale image to avoid computational costs, but the color version gives better representation details with better accuracy. After that we resized our images into 512 x 512 to unify the size of the images.

Images can also contain different types of noises due to the different sources (camera sensor). First, we blurred the images for image sharpening. Image smoothing (blurring) strategies help to reduce noise. Image smoothing can be done in a number of ways. We used the Gaussian filter to smooth the image. After which we subtracted the smoothed image from the original image (the consequent distinction is referred as a mask). As a result, the resultant image (mask) had maximum number of the high-frequency components that were cancelled with the aid of the smoothing filter. The addition of this mask back to the original input images enhanced the high-frequency components. We got the sharpened images in this way.

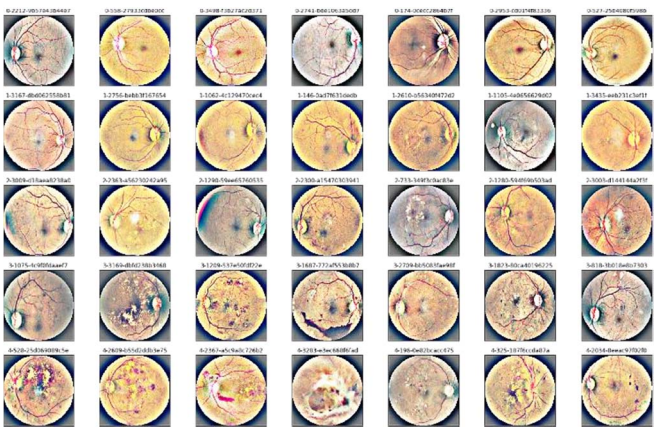

Fig. 2. Pre-processed color images

Authorized licensed use limited to: Carleton University. Downloaded on May 27,2021 at 22:11:08 UTC from IEEE Xplore. Restrictions apply.

C. *Architecture Design of Proposed CNN Model*

We utilized transfer learning technique for getting better performance and saving time. Conventional learning is separated and occurs entirely on the basis of datasets, particular tasks and the training of separate, isolated models on them. There is no knowledge left which can be moved from one model to the other. We can exploit knowledge (weights, features etc.) from earlier trained models to train newer models in transfer learning and even deal with issues such as having less data for the newer task.

Model: "model"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 512, 512, 3)]	0
block1_conv1 (Conv2D)	(None, 512, 512, 64)	1792
block1_conv2 (Conv2D)	(None, 512, 512, 64)	36928
block1_pool (MaxPooling2D)	(None, 256, 256, 64)	0
block2_conv1 (Conv2D)	(None, 256, 256, 128)	73856
block2_conv2 (Conv2D)	(None, 256, 256, 128)	147584
block2_pool (MaxPooling2D)	(None, 128, 128, 128)	0
block3_conv1 (Conv2D)	(None, 128, 128, 256)	295168
block3_conv2 (Conv2D)	(None, 128, 128, 256)	590080
block3_conv3 (Conv2D)	(None, 128, 128, 256)	590080
block3_pool (MaxPooling2D)	(None, 64, 64, 256)	0
block4_conv1 (Conv2D)	(None, 64, 64, 512)	1180160
block4_conv2 (Conv2D)	(None, 64, 64, 512)	2359808
block4_conv3 (Conv2D)	(None, 64, 64, 512)	2359808
block4_pool (MaxPooling2D)	(None, 32, 32, 512)	0
block5_conv1 (Conv2D)	(None, 32, 32, 512)	2359808
block5_conv2 (Conv2D)	(None, 32, 32, 512)	2359808
block5_conv3 (Conv2D)	(None, 32, 32, 512)	2359808
block5_pool (MaxPooling2D)	(None, 16, 16, 512)	0
flatten (Flatten)	(None, 131072)	0
dense (Dense)	(None, 1024)	134218752
dense_1 (Dense)	(None, 5)	5125

Total params: 148,938,565
Trainable params: 134,223,877
Non-trainable params: 14,714,688

Fig. 3. Proposed Model's architecture

In our proposed model, we used the VGG16 model as previously trained (pre-trained) model for the extraction of complex features. VGG16 has a total of 13 convolution layers that use 3x3 convolution kernels with a stride of 1, use same padding. Max pooling layers are used for down sampling. Then a softmax classifier will be followed by two fully connected layers each with 4,096 nodes. Because we added two fully connected dense layers used to predict image classification, we cut the last three layers from the original VGG16 model. The first five blocks were more important to us so that we could use these blocks from the VGG16 model as an effective extractor and save time of training. The five convolution blocks of VGG16 were frozen so that weights will not be updated.

With stratified K-fold cross validation, we have folded our training data set into five folds to confirm that each fold contains almost the same percentage of samples from each target class as the entire set [12]. Each time a fold was used for validation, the rest of the folds were used for model training. Cross Validation of stratified K-fold helped to avoid the problem of over-fitting.

In addition, it is important to choose an optimizer when training a deep learning model [11]. A good optimizer can speed up the training process significantly, avoid the bad local optimum and provide a better learning result. We used Adam as the optimizing algorithm for our model.

III. RESULTS AND DISCUSSION

We run our proposed CNN model in Google Colab with GPU support using Tensorflow python library. After the original fundus images were processed and saved, the proposed model was trained and validated using stratified 5-fold cross-validation of which one fold was used for validation and the remaining folds for training. The batch size was set to 40 because of the shortage of computing resources.

There were 0.9904404 and 0.9870493 respectively, the highest training and validation accuracy observed. Tuning the rate of learning and no. of epochs the model was trained and validated. With the learning rate of 0.001 and epoch size of 10, the folds were trained and validated four times; each time the split was performed with random shuffling to remove the problem of over-fitting and biasness.

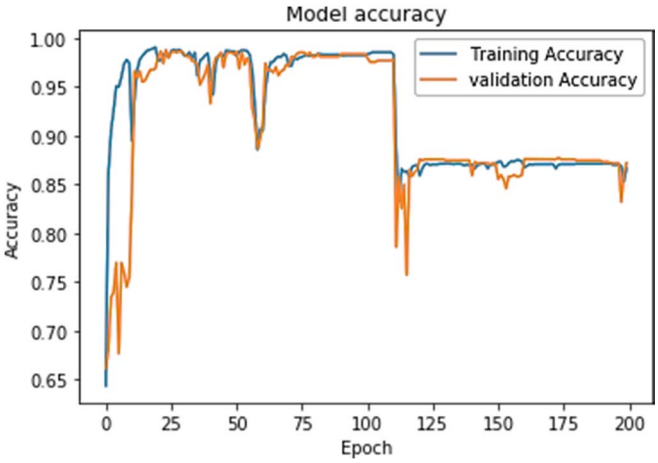

Fig. 4. Proposed Model's accuracy

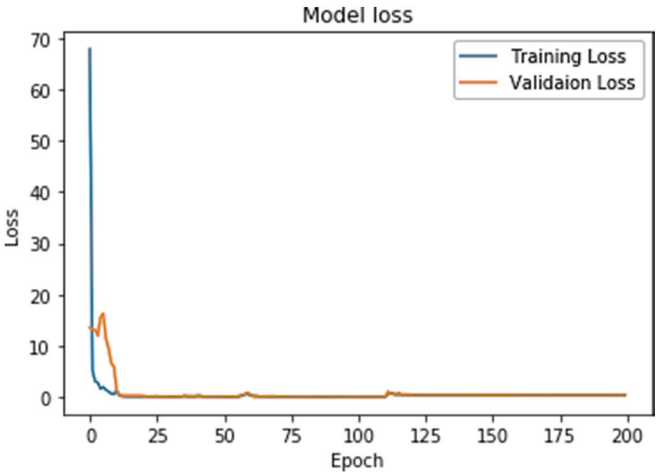

Fig. 5. Proposed Model's training loss and validation loss

Authorized licensed use limited to: Carleton University. Downloaded on May 27,2021 at 22:11:08 UTC from IEEE Xplore. Restrictions apply.

We got the average accuracy of training: 0.92577666, average accuracy of validation: 0.9132683, average loss of accuracy: 0.0640956 and average loss of validation: 0.0816171. As we can see from the Fig. 4, the validation accuracy is at a saturation level (above 0.95) up to 115 epochs. After 115 epochs, the model's accuracy has fallen but is greater than 0.80 on average. Therefore, the average accuracy is at a satisfactory level.

IV. CONCLUSION AND FUTURE WORK

We used the method of transfer learning from the pre-trained VGG16 model in this paper. But before feature extracting and classification, the fundus images were processed. We have seen that with the increase of number of epochs the accuracy rate increases before falling down into a local minimum. But our proposed model got an average validation accuracy of 0.9132683. Another limitation of our work is that the dataset was an imbalanced dataset. In our future work, we will solve the problem of local minimum. Furthermore, we will improve the accuracy far better and import more diabetic retinopathy dataset along with our dataset to get robust and more generalized results.

REFERENCES

[1] E. V. Carrera, A. González and R. Carrera, "Automated detection of diabetic retinopathy using SVM," 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON), Cusco, 2017, pp. 1-4.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] D. U. N. Qomariah, H. Tjandrasa and C. Fatichah, "Classification of Diabetic Retinopathy and Normal Retinal Images using CNN and SVM," 2019 12th International Conference on Information & Communication Technology and System (ICTS), Surabaya, Indonesia, 2019, pp. 152-157.

[3] P. Khojasteh, B. Aliahmad, S. P. Arjunan and D. K. Kumar, "Introducing a Novel Layer in Convolutional Neural Network for Automatic Identification of Diabetic Retinopathy," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, 2018, pp. 5938-5941.

[4] Y. Jiang, N. Tan, T. Peng and H. Zhang, "Retinal Vessels Segmentation Based on Dilated Multi-Scale Convolutional Neural Network," in IEEE Access, vol. 7, pp. 76342-76352, 2019.

[5] T. A. Soomro et al., "Deep Learning Models for Retinal Blood Vessels Segmentation: A Review," in IEEE Access, vol. 7, pp. 71696-71717, 2019.

[6] H. Thanati, R. J. Chalakkal and W. H. Abdulla, "On Deep Learning based algorithms for Detection of Diabetic Retinopathy," 2019 International Conference on Electronics, Information, and Communication (ICEIC), Auckland, New Zealand, 2019, pp. 1-7.

[7] K. Bhatia, S. Arora and R. Tomar, "Diagnosis of diabetic retinopathy using machine learning classification algorithm," 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, 2016, pp. 347-351.

[8] V. Raman, P. Then and P. Sumari, "Proposed retinal abnormality detection and classification approach: Computer aided detection for diabetic retinopathy by machine learning approaches," 2016 8th IEEE International Conference on Communication Software and Networks (ICCSN), Beijing, 2016, pp. 636-641.

[9] Kaggle, "APTOS 2019 Blindness Detection", https://www.kaggle.com/c/aptos2019-blindness-detection/data, August 29, 2019, accessed January 01, 2020.

[10] S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," in IEEE Access, vol. 7, pp. 150530-150539, 2019.

[11] X. Zeng, H. Chen, Y. Luo and W. Ye, "Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network," in IEEE Access, vol. 7, pp. 30744-30753, 2019.

[12] L. Li, W. Wang, S. Yu, L. Wan, Z. Xu and X. Kong, "A Modified Node2vec Method for Disappearing Link Prediction," 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech), Orlando, FL, 2017, pp. 1232-1235.

[13] X. Wang, Y. Lu, Y. Wang and W. Chen, "Diabetic Retinopathy Stage Classification Using Convolutional Neural Networks," 2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, 2018, pp. 465-471.

Authorized licensed use limited to: Carleton University. Downloaded on May 27,2021 at 22:11:08 UTC from IEEE Xplore. Restrictions apply.