International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), 11-12 July, 2019

Diabetic Retinopathy Classification with a Light Convolutional Neural Network

Md. Sanaullah Chowdhury
Electronics & Communication Engineering Discipline
Khulna University
Khulna, Bangladesh
sanaullahashfat@gmail.com

Faozia Rashid Taimy
Electronics & Communication Engineering Discipline
Khulna University
Khulna, Bangladesh
faozia.rashid.taimy@gmail.com

Niloy Sikder
Computer Science & Engineering Discipline
Khulna University
Khulna, Bangladesh
niloysikder333@gmail.com

Abdullah-Al Nahid
Electronics & Communication Engineering Discipline
Khulna University
Khulna, Bangladesh
nahid.ece.ku@gmail.com

***Abstract*—The number of diabetic patients is increasing rapidly every year all around the world, and the worst fact is that these patients suffer from a wide range of physical conditions directly associated with long-term diabetes. Diabetic Retinopathy (DR) is a perfect example which affects the eyes of more than 50% of all diabetes patients to some degree. Starting from blurred vision, the effects of DR can extend to permanent blindness; and in most of the cases, victims fail to report any early symptoms. The traditional detection process of DR involves a trained clinician who takes enhanced pictures of the retina and looks for the presence of lesions and vascular abnormalities within them, which by description is a time-consuming and error-prone procedure. Alternatively, we can employ machine learning techniques that will automate the detection process as well as provide fast and more importantly, reliable results. Using a deep learning technique this paper determines the presence and severity of DR in diabetic individuals by analyzing the pictures of their retina. The CNN-based models are potent enough to carry out their tasks with accuracy up to 89.07%, even when the images are captured or provided in very low resolutions.**

Keywords*—*diabetic retinopathy detection, deep CNN, coye filter, image classification, machine learning

I. Introduction

Diabetic Retinopathy (DR) is the leading complication of blindness in adults. Statistics show that at present about 4.2 million adults had DR and 655,000 had vision-loss due to DR. The number of people who are affected by vision loss due to DR increased in every year [1]. It is known to all that DR becomes a common complication of Diabetes Mellitus (DM). The treatment of DR is not easy as there is no symptom shown at the earlier-stages of DR and patients rarely notice a vision loss [2]. Most of them could not realize that they have DR until the disease started to affect their vision, which usually occurs in the last stage. As a result, they might not go through treatment promptly. Therefore, a coordinated management strategy is crucial to optimally address the clinical challenges of DR and limiting its progression.

Early identification and classification of disorder of the retinal images always being a serious concern to the research community. There are a few works that have been performed to classify retinal image malignancy utilizing conventional shallow learning techniques. However, Deep Learning (DL) techniques have gained tremendous success in solving visual related problems [3], [4]. One of the variants of this method known as Convolutional Neural Network (CNN) has gained a tremendous momentum after the famous model "AlexNet" which is proposed in 2012 [5]. Initially, this method used to classify natural images. However, in later, several variants of this model are used to solve many real lives vision-related problems, such as identification of malignancy of retinal images [6], [7], [8]. Image classifications rate might be degraded due to different issues such as contrast, illumination etc. [9], which can be improved by various image pre-processing techniques. Chen et al. utilized linear un-sharp masking filter to enhance the edge and detailed information, then, the enhanced image is feed to a CNN method named as SI2DRNet-v1 for retinal image classification [7]. V. Raman et al. classified a set of retinal images whereas a pre-processing technique, Contrast-Limited Adaptive Histogram Equalization (CLAHE) method is utilized to improve the visibility of images [10]. D. K. Prasad et al. classify a set of retinal images (DIARETDB1 dataset) for detecting the early status of the DR and those images are preprocessed by adaptive histogram equalization techniques to improve the image contrast [1]. M. U. Akram et al. classify retinal lesions utilizing a hybrid classifier where they utilized Gabor filter as a pre-processing tool [8]. From the above discussion it turns out that, an image preprocessing stage before the classification stage allows a classifier to perform better, Coye filter has been used as an image preprocessor which enhances the contrast information of the corresponding image [11]. Then a noble light CNN model is hyper-tuned by these enhanced images to classify a set of retinal images into their respective classes.

To do so, this paper starts with Section I stating the background of DR and its severity throughout the world. Various aspects of the proposed scheme especially the entire methodology has been step-wise illustrated in section II. A brief analysis in CNN to classify DR has been incorporated in Section III along with our proposed a light noble CNN model. In Section IV, the overall analysis and comparison of the results have been presented. This paper concludes with Section V providing a discussion on overall research.

II. Methodology

This experiment has performed in a supervised manner as Fig. 1. In a general principle, supervised learning method trained based on a training dataset which is properly labeled. This labeled data is used to select probable best parameters for the classifier. Lastly, the selected or targeted data is fit with the model which provide relatively the best decisions. This experiment is performed on a publicly available dataset named as EyePACS [12]. This dataset contains five classes provided on a scale of 0 to 4. Table I shows the statistic information about the different classes of the dataset such as The table shows that almost 78 to 80 percent of the total images belong to class 0. However, another four classes of the data cover almost 20% of the total data. This indicates that this dataset is imbalanced.

978-1-7281-3060-6/19/$31.00 ©2019 IEEE

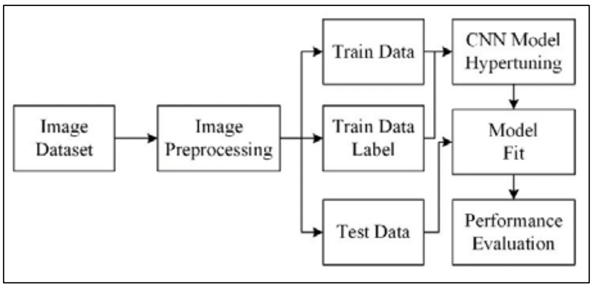

Fig. 1. Block diagram of the proposed DR classification method.

TABLE I. PROPERTIES OF THE EYEPACS DATASET

Class	No DR	Mild	Moderate	Severe	Proliferative
Samples	25770	2432	5278	867	707
Class	0	1	2	3	4

A. Image Pre-processing

To improve the contrast information of a captured RAW image as a pre-processing tool this paper utilized Coye filter which is described Fig. 2. The Coye filter follows these steps:

- In Step-1, raw image $I(x, y)$ is taken as input which is RGB in nature. The raw Red Green Blue(RGB)three channel image is converted to grayscale (GRAY) image through Lab domain and CLAHE operation consecutively [10]. This CLAHE method consists of 128 bins. Between the Lab and CLAHE operation, Principle Component Analysis (PCA) techniques are utilized [11].
- In Step-2, image color variation is corrected through background subtraction. The amount of subtracted information from the background is calculated using an average image filtering technique.
- In Step-3, the output images are calculated from the subtracted image with intermediate steps such as binary image correction, complementing and masking operation.

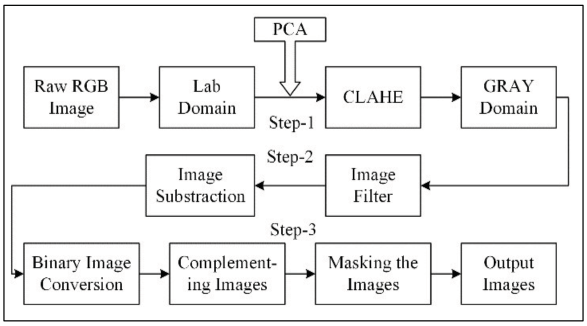

Fig. 2. Image pre-processing steps.

Before working with them, we needed to create a functional dataset where each sample contains a fraction of the original EMG signal of its corresponding class. Fig. 3 shows the formation of the experimental dataset along with the labels for each class. The figure also indicates that the dataset we are working with is balanced, meaning it contains the same amount of samples of each class which in turn, reduces the chances of getting biased decisions.

B. CNN

DL has brought revolutionary changes in the current era which continuously allow the intelligent machine to behave as like as human. The mathematical structures of DL have been utilized to solve many problems of different domains. Among the few branches of DL, Convolution Neural Network (CNN) has also made a strong footprint in many areas [5], [13]. Basically, CNN is an extension to the Neural Network (NN), where instead of utilizing a set of local features, global features are extracted and utilized. These global features are extracted by a special mathematical operation known as convolution. A convolution operation is performed by a two-dimensional $k_1 \times k_2$ matrix known as kernel k. k is scanned throughout the available data matrix. To overcome the boundary issues, zero padding operation is performed. If an input image I is convoluted by k, then the convolution operation can be represented as

$$(I * k)_{ij} = \sum_{m=-k_1/2}^{k_1/2} \sum_{n=-k_2/2}^{k_2/2} I(i-m, j-n)k(m,n) \quad (1)$$

Sometimes a pooling operation is performed to downsample the data which particularly solve the overfitting problems as like conventional NN, back-propagation technique is utilized to reduce the overall error values. To make the structure of CNN as like a NN, at least one dense layer is utilized before the decisions layer. In decision layers, conventionally softmax layer is utilized.

C. Our Proposed Model

Our proposed model is illustrated in Fig. 3. In this particular work, the original input image is scanned by a 3×3 filter, which produces 32 feature maps, where each features map is scaled down to 30×30 matrix. Then a max-pooling layer of 2×2 matrix is utilized, which downsample the feature matrix to 15×15 scale. After this down-sampling stage one convolution layer, one pooling layer and then again a convolution layer is utilized [13] [14]. The last convolution layer produces 64 feature maps. To further down-sample the data, again a 2×2 max-pooling operation is performed which produces 64 feature maps. Then this layer is flattened to make it compatible with a conventional NN. As a result of this, it produces 256 neurons. Each of these 256 neurons is connected with another 380 neurons. Lastly, 10% of these neurons are dropped out to overcome the oversampling issue. At the last layer, we have utilized five neurons (as there are five classes) with softmax operation to perform the

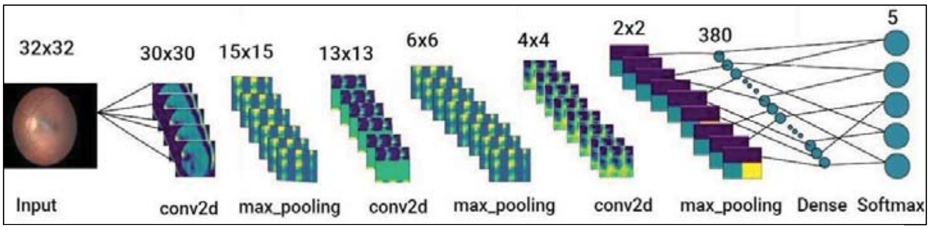

Fig. 3. Proposed CNN model for classification

classification operation. Fig 3 represents the model which has been utilized for this experiment.

III. Results and Discussion

t-Distributed Stochastic Neighbor Embedding (t-SNE) is mainly utilized for data visualization through reduction of the data dimensions [15]. It is easy to visualize data in two or three dimensions, however, if the data dimension increased it is difficult to understand the data pattern. t-SNE is a mathematical tool for data visualization. In Fig. 4, the red, blue, purple, yellow, pink dots represent the 0, 1, 2, 3, and 4 labeled classes respectively on a two-dimensional plane.

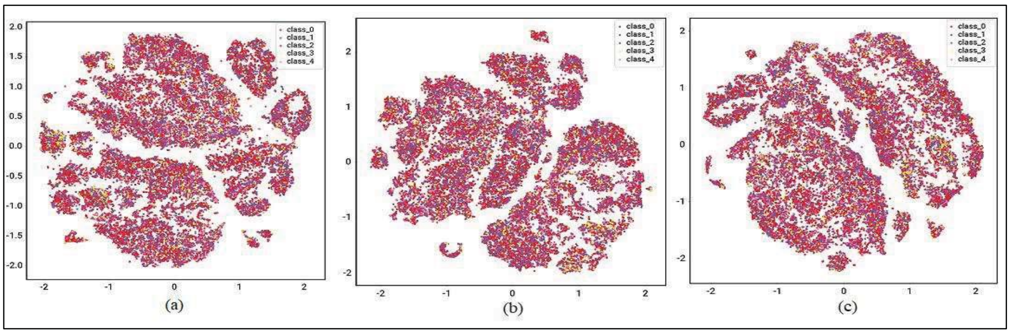

Fig. 4. t-SNE of (a) Case-I, (b) Case-II, and (c) Case-III.

This experiment is performed in a supervisory manner. To hyper-tune the weight and bias values of the model through training, a certain amount of data from the total data set have been utilized as training data. In this particular case 70% data of the total data-set has been considered as training data-set. Rest of the data are utilized for the testing purpose. This experiment considered three cases, in the first case (Case-I) the original images are directly fed to the model. In the second case (Case-II), only green channel image information are feed to the model. In the third (Case-III) Coye filtered images are feed to the model. This experiment utilized 32 × 32 pixels images to overcome computational complexity and constraints. Fig. 5 a, b, c represent the accuracy performance for Case-I, Case-II and Case-III respectively. In all the experiments, a back-propagation algorithm with optimization techniques named as "Adam" is allowed to perform up to 50 epochs. Fig. 5 shows that, train accuracy is always higher than test accuracy. As the epoch goes, in all the cases train accuracy continuously increased which is almost 100%. For Case-I, the test accuracy is almost constant which is near about 88%, with a few exceptions. Interestingly, for the Case-II and Case-III after a few initial epochs, the test accuracy continuously decreased, that means for this particular two cases the gap between the train and test accuracy increased with the epoch number. Fig. 6 a, b, c represent the loss performance for Case-I, Case-II and Case-III respectively. These figures show that, with the epoch increases the training loss continuously decreased, however test loss increased, which resembles the accuracies.

When the data-set is imbalanced, Mathews Correlation Coefficient (MCC) is an excellent performance measuring parameters both for the binary and multi-class classification problems. The value of MCC is bounded by the region in between the range –1 to +1, where –1 represented a classifier as a worst one whereas +1 marked as a classifier as a perfect one. Fig. 7 shows the MCC performances of Case-I, Case- II and Case-III respectively. All the figures show that as with the epoch the value of MCC for the training scenario.

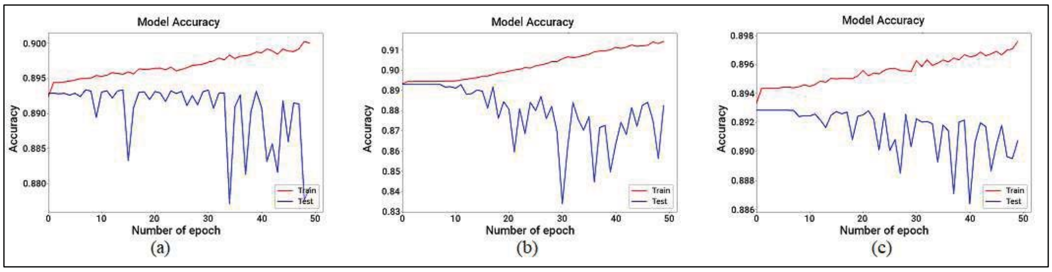

Fig. 5. Classification accuracy of (a) Case-I, (b) Case-II, and (c) Case-III.

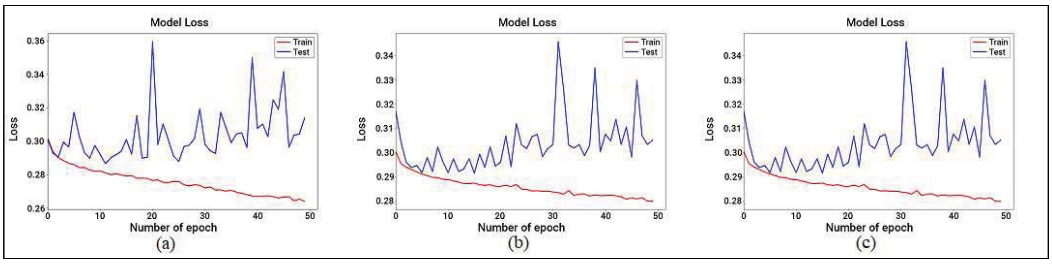

Fig. 6. Loss curve of (a) Case-I, (b) Case-II, and (c) Case-III.

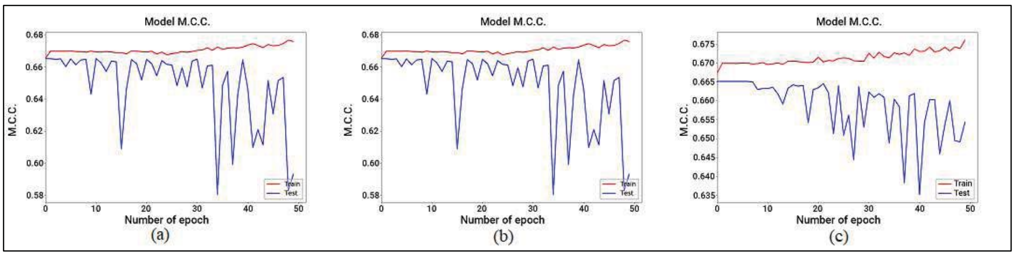

Fig. 7. MCC of (a) Case-I, (b) Case-II, and (c) Case-III.

Receiver Operating Characteristic (ROC) curve represents the ratio of the Sensitivity against Specificity [16]. The diagonal line represents the performance of a model which provides results in a random nature. Fig. 8 shows the ROC curves for the three cases discussed before.

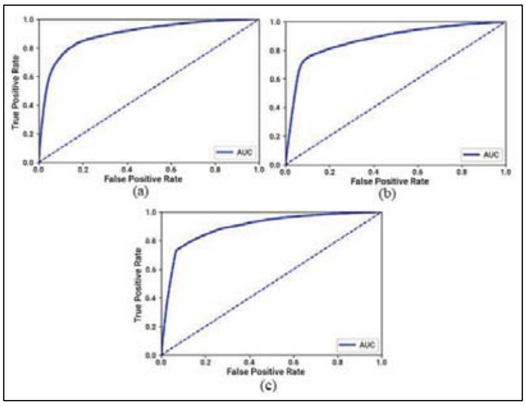

Fig. 8. ROC curve of (a) Case-I, (b) Case-II, and (c) Case-III

Table II represents the overall findings of our experiment. This table shows that Case-III provides the best accuracy (89.07%) among all the three cases. It also summarizes other findings such as Recall, Precision, F-measure, MCC and Kullback Liberal Divergence (KLL) values [17].

TABLE II. SUMMARY OF THE ACQUIRED RESULTS

Parameters	Case-I	Case-II	Case-III
Accuracy (%)	87.93	88.23	89.07
Recall (%)	69.73	68.51	71.03
Precision (%)	76.31	71.48	73.44
F-measure (%)	65.54	69.95	72.21
MCC (%)	59.32	62.68	65.43
KLD (%)	98.82	87.30	83.85

Table III compares the finding of this paper with state-of-the-art findings. The table shows that the described method outperforms the methods described in [2] and [13] by 40.87% and 4.07% respectively.

TABLE III. COMPARISON WITH OTHER METHODS

Parameters	Proposed	[2]	[13]
Accuracy (%)	89.07	48.2	85

IV. CONCLUSION

Updated knowledge of DR is essential to the doctors as well as the patient to control the risk of eye losses. This paper attempts to classify DR images in an automated way using a state-of-the-art CNN model. The utilized CNN model is light. To improve the contrast variability, as a pre-processor this paper utilized a contrast-enhanced filter. The overall accuracy of this experiment is 89.07%, which is comparable to the recent findings on this particular dataset. In the future, with fine-tuning, these results can be further improved with the addition of complexity and latency.

REFERENCES

[1] D. K. Prasad, L. Vibha, and K. Venugopal, " Early detection of diabetic retinopathy from digital retinal fundus images," 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), 2015.

[2] S. Masood, T. Luthra, H. Sundriyal, and M. Ahmed, "Identification of diabetic retinopathy in eye images using transfer learning," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017.

[3] A. Nahid, M. A. Mehrabi, and Y. Kong, “Histopathological Breast Cancer Image Classification by Deep Neural Network Techniques Guided by Local Clustering,” BioMed Research International, vol.2018, pp. 1–20, 2018.

[4] A. Nahid, A. Mikaelian, and Y. Kong, "Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation," Biomedical Research, vol. 29, no. 10, 2018.

[5] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2012.

[6] S. Dutta, B. C. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, “Classification of Diabetic Retinopathy Images by Using Deep Learning Models,” International Journal of Grid and Distributed Computing, vol. 11, no. 1, pp. 99–106, 2018.

[7] Y. W. Chen, T. Y. Wu, W. H. Wong, and C. Y. Lee, "Diabetic Retinopathy Detection Based on Deep Convolutional Neural Networks," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018.

[8] M. U. Akram, S. Khalid, A. Tariq, S. A. Khan, and F. Azam, “Detection and classification of retinal lesions for grading of diabetic retinopathy,” Computers in Biology and Medicine, vol. 45, pp. 161–171, 2014.

[9] A. Nahid, F. B. Ali, and Y. Kong, "Histopathological breast-image classification with image enhancement by convolutional neural network," 2017 20th International Conference of Computer and Information Technology (ICCIT), 2017.

[10] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color retinal image enhancement using CLAHE," International Conference on ICT for Smart Society, 2013

[11] C. Tyler, “A Novel Retinal Blood Vessel Segmentation Algorithm for Fundus Images,” MATLAB Cent. File Exch., 2016.

[12] “Diabetic Retinopathy Detection,” Kaggle. [Online]. Available: https://www.kaggle.com/c/diabetic-retinopathy-detection. [Accessed: 20-Feb-2019].

[13] A. Nahid, M. A. Mehrabi, and Y. Kong, "Frequency-domain information along with LSTM and GRU methods for histopathological breast-image classification," 2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2017.

[14] A. Nahid and Y. Kong, "Histopathological Breast-Image Classification Using Local and Frequency Domains by Convolutional Neural Network," Information, vol. 9, no. 1, p. 19, 2018.

[15] L.V.D. Maaten, and G. Hinton, “Visualizing Data using t-SNE Laurens,” J. Mach. Learn. Res., vol. 9, pp. 2579–2605, 2008.

[16] N. Sikder, K. Bhakta, A. A. Nahid, and M. M. M. Islam, “Fault Diagnosis of Motor Bearing Using Ensemble Learning Algorithm with FFT-based Preprocessing,” 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 564–569, 2019.

[17] K. Bhakta, N. Sikder, A. A. Nahid, and M. M. M. Islam, “Fault Diagnosis of Induction Motor Bearing Using Cepstrum-based Preprocessing and Ensemble Learning Algorithm,” 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019.