Retinal Disease Classification from Optical Coherence Tomographical Scans using Multilayered Convolution Neural Network

Rajarshi Bhadra
Dept. of Electrical Engineering
Future Inst. of Engg. and Mgmt.
Kolkata, India
rajarshibhadra11@gmail.com

Subhajit Kar, *Member, IEEE*
Dept. of Electrical Engineering
Future Inst. of Engg. and Mgmt
Kolkata, India
subhajit.kar@teamfuture.in

***Abstract*—Classification of retinal diseases using Optical Coherence Tomographical (OCT) scans is a crucial task. Accurate detection and classification of these diseases is necessary for patient's survival. Presently, the analysis of retinal diseases are carried out by doctors by examining the OCT images regularly. However the manual diagnosis procedure is tedious. Therefore, in this paper, an automatic detection and classification technique of retinal diseases has been proposed to assist doctors in their diagnosis. A deep multilayered convolutional neural network (CNN) has been used to detect and classify the retinal abnormalities using OCT scans. The proposed technique has been applied on an open source retinal OCT dataset containing 59,142 images and 96.5% blind test accuracy has been achieved.**

***Index Terms*—Retinal Abnormalities, OCT scans, Deep Learning, multilayered CNN**

I. Introduction

OCT is a non-invasive imaging test [1]. It uses light waves to capture cross section images of retina in human eyes. It is also often used to evaluate disorders of optic nerve. OCT has a vast application in medical imaging. OCT has been used to treat eye abnormalities such as- Choroidal NeoVascularization (CNV) [2], Diabetic Macular Edema (DME) [3], Drusen [4]. Diagnosis of retinal diseases require trained medical professionals to analyze the OCT images manually for detecting retinal abnormalities. This manual procedure is complex, time consuming and also prone to error. Therefore, an automatic detection and classification procedure is required to assist the medical professionals with their diagnostic prediction.

In this context, deep learning plays a vital role in medical image analysis. A novel method combining deep learning and graph propagation has been proposed by Girard et al. to classify artery/vein diseases from the fundus images [5]. The DRIVE database has been used by them to test the model performance and 93.3% test accuracy is achieved. A region based CNN approach has been used by Sambaturu et al. to mark the exudates and hemorrhages automatically for analyzing diabetic retinopathy [6], [7]. A recall of 90% has been achieved by them by marking the lesions successfully. A pretrained CNN model has been used by Calimeri et al. to detect the position in retinal fundus images [8]. A specialized CNN model has been proposed by Athar et al. to localize and classify Fluid Filled Regions (FFRs) in retinal OCT [9]. They have used Dense Blocks and Scaled Exponential Unit (SeLU) activations to achieve a mean average precision of 0.75 on true positive images and test accuracy of 94.8%. A CNN model has also been proposed by Gopinath et al. to design a mapping function by combining multiple motions to produce a probability map for cyst location in retinal OCT scans [10].

In contrast, a deep multilayered CNN has been proposed in this paper for retinal disease detection and classification. The state of the art image processing techniques have been applied to extract the ROI's from the OCT scans. The one dimensional layers of the trained model has been retrained to elevate the blind classification accuracy. Further, the Adam optimizer [11] has also been used in the proposed methodology to optimize the overall training time. In this work, 59,142 images have been used for training, 16,896 images have been used for validation and 8446 images have been used to test the efficacy of the proposed methodology.

II. Proposed Approach

The block diagram of the proposed methodology has been shown in Fig. 1. First, the region of interests (ROIs) have been segmented from the retinal OCT scans. Then the segmented ROIs have been used as input to the deep multilayered CNN. The deep multilayered CNN then classifies the retinal OCT scans in 4 classes such as CNV, DME, DRUSEN, and NORMAL.

Fig. 1. Block Diagram of Proposed Methodology.

A. Data Preprocessing

The retinal OCT scans contain artifacts that decrease the classification performance. Therefore, binary transformation [12] has been applied using an experimental threshold value on the retinal OCT images to remove the artifacts. Then, the complete white pixels have been transformed to complete dark pixels. Thereafter, a bounding box has been applied around the dark pixels and the region inside the bounding box has been segmented from the retinal OCT scan. Finally, the processed image has been resized for using in the deep multilayered CNN architecture. The binary transformation and data preprocessing algorithms have been shown in Algorithm 1 and Algorithm 2 respectively. Fig. 2 shows different steps of data preprocessing techniques

Algorithm 1: Binary Transformation

Choose thresh = T_i, maxval = M_i;
2D image f(x,y) of size M $\times$ N
if $f(x, y) > thresh$ **then**
  $f(x, y) = 0$;
else
  $f(x, y) = maxval$;
end

Algorithm 2: Data Preprocessing Algorithm

- Apply Binary Transformation.
- Transform complete white pixels i.e $f(x, y) = 255$ to complete dark pixels i.e. $f(x, y) = 0$.
- Apply bounding box across the dark pixels.
- Crop the region inside the bounding box.
- Resize the cropped image to $128 * 128$

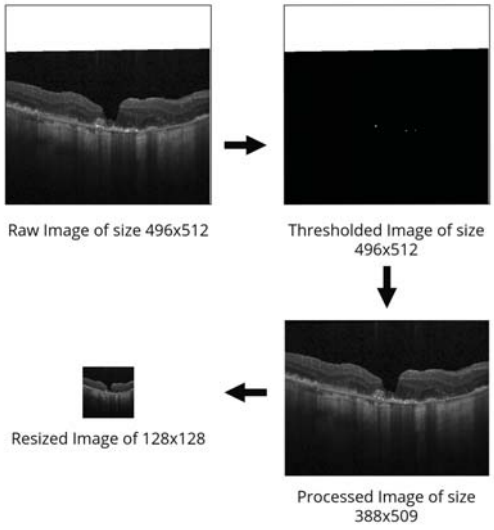

Fig. 2. Data Preprocessing.

B. Proposed Deep Multilayered CNN Architecture

The proposed deep multilayered CNN has been shown in Fig. 3. The processed images have been used as input to the deep multilayered CNN architecture. Two convolution operations with 16 filters of 3 x 3 kernels have been performed in each layer using unit stride sliding window and single padding. Rectified linear unit (ReLU) [13] has been used as the activation function. A 2 x 2 max pooling operation has also been performed to pool the results of the convolutions from each layer into a more compact tensor. After the multilayered structure, the three dimensional tensors have been converted to one dimensional feature vectors consisting 4096 neurons. Thereafter, a dense layer of 256 neurons, a dropout layer with a dropout rate of 30% and an output layer consists of 4 neurons have been used. ReLU has been used as an activation function in the dense layer and SoftMax has been used as an activation function in the output layer. ReLU or Rectified Linear Unit is defined as,

$$y = max(0, x) \tag{1}$$

where, x is the input to ReLU activation function and y is it's output. The SoftMax activation function is defined as,

$$\hat{y}_i = \frac{e^{y_i}}{\sum_{j=1}^{k} e^{y_j}} \tag{2}$$

where, y_i is i^{th} logit value, k is total number of logits, $\hat{y_i}$ denotes pedicted probability of a particular sample. The output layer of the proposed deep multilayered CNN calculates the probabilities of each class. A particular output node (Logit) having the highest probability signifies the classification of the corresponding class.

III. Experimental Result

A. Database Preparation

The retinal OCT database described in [14] has been used in this work. The database consist of four classes: CNV, DME, DRUSEN and NORMAL. These classes have been shown in Fig. 4 and Fig. 5. The dataset has been split into training set, validation set and blind test set to prove the efficacy of the proposed method.
The retinal OCT dataset contains 84,484 images.

- Training set: The training set contains 59,142 images (CNV: 26,219; DME: 8120; DRUSEN: 6207 and NORMAL: 18,596).
- Validation set: The validation set contains 16,896 images (CNV: 7491; DME: 2319; DRUSEN: 1773 and NORMAL: 5313).
- Blind test set: The blind test set contains 8446 images (CNV: 3745; DME: 1159; DRUSEN: 886 and NORMAL: 2656).

B. Training

In this paper, a deep multilayered CNN has been used for training purpose using Adam optimizer [11] which is a

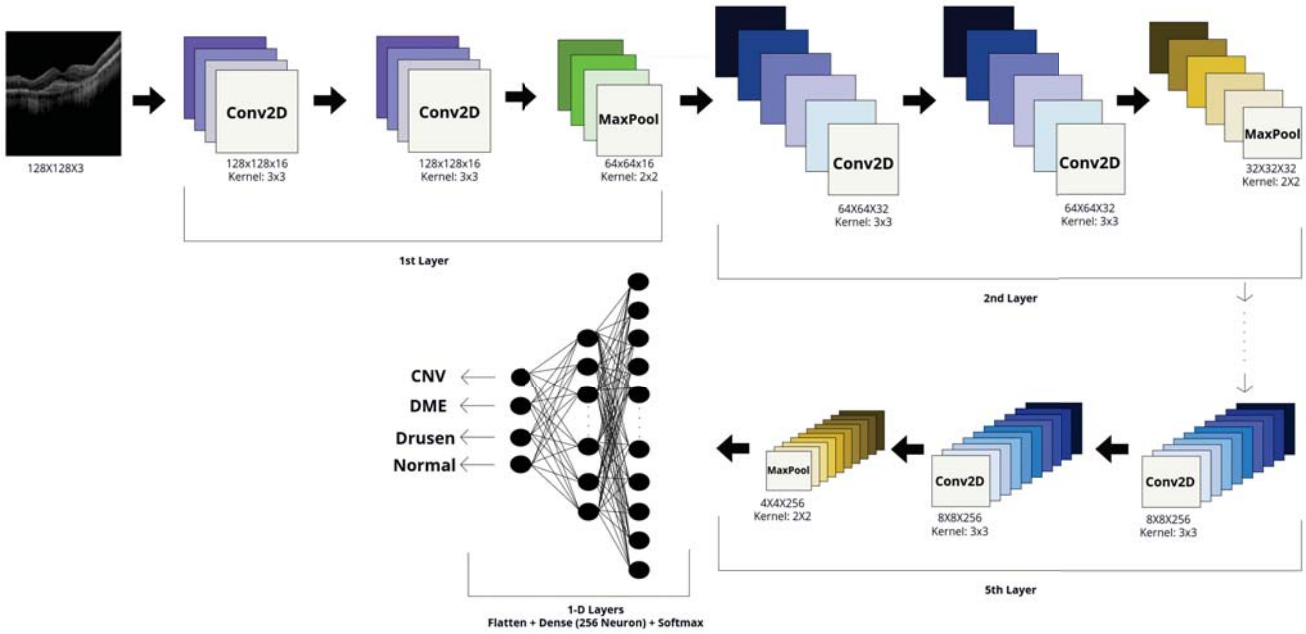

Fig. 3. Proposed Deep Multilayered CNN Architecture.

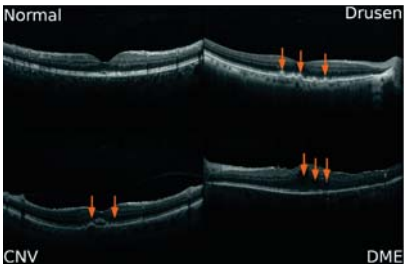

Fig. 4. Different classes of retinal abnormalities.

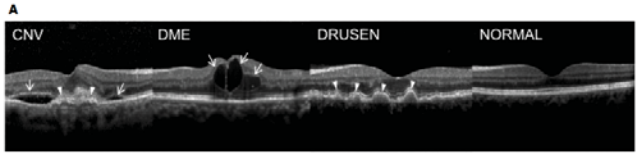

Fig. 5. ROIs of different retinal abnormalities.

first order gradient based stochastic optimization process. The Adam optimizer is given by:

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{3}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{4}$$

where, m_t and v_t are estimates of the first moment (the mean) and the second moment (the uncentered variance) of the gradients respectively, g_t is gradient at time t, β_1 and β_2 are exponentially decay rates for moment estimates. The β_1 and β_2 values have been set to 0.9 and 0.999 respectively in the proposed algorithm. The categorical cross entropy loss has been used as cost function in the proposed algorithm. The cost function $(J(\theta))$ can be defined as:

$$J(\theta) = -\sum_{c=1}^{M} y_{i,c} log(p_{i,c}) \tag{5}$$

where, M denotes number of classes (in this case which is 4), $y_{i,c}$ is a binary indicator (0 or 1) that indicates whether c is the correct class, $p_{i,c}$ denotes predicted probability between 0 and 1. In the proposed approach, the training is converged after 12 epochs while all the layers of the model have been trained and no improvement over categorical crossentropy loss has been monitored for 5 more epochs. However, the training is converged after 11 epochs by applying re-training on the one dimensional layers of the proposed model and losses with no improvement has also been monitored for 5 more epochs. The initial learning rate has been selected as 0.0001 and the learning rate is decreased by a factor of 10 after every 10 epochs. The training accuracy of 99.9%, validation accuracy of 95.9% and blind test accuracy of 96.5% have been evaluated by using the proposed method.

The training accuracy has been evaluated on the training dataset of the model. In contrast, the validation accuracy has been calculated on the validation dataset that has never been

used for the training purpose. Hence, the training accuracy is higher than the validation accuracy. The training characteristics of the base model and the retrained model have been shown in Fig. 6 and Fig. 7 respectively.

The proposed algorithm using deep multilayered CNN architecture has been shown in Algorithm 3. The classification results have further been reconfirmed by precision, recall and f1 score. The classification report on blind test set for different retinal OCT classes has been shown in Table 1. Furthermore, the classification result has also been analyzed by calculating the confusing matrix shown in Fig. 8. The confusion matrix shows the number of blind test samples classified correctly and the recall values for each class.

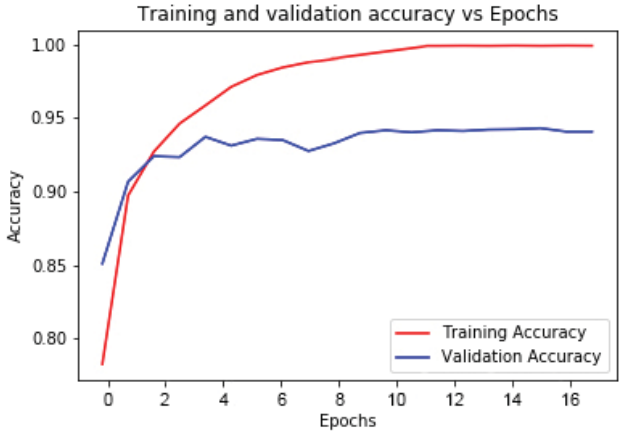

Fig. 6. Training Characteristics for Base Model.

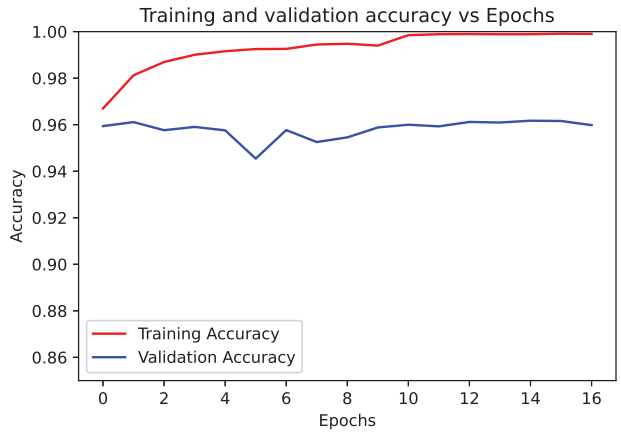

Fig. 7. Training Characteristics after applying retraining on 1-D Layers.

IV. Conclusion

In this paper, a deep multilayered CNN with Adam optimizer has been proposed to detect and classify retinal abnormalities. First, the classical image processing techniques have been applied to segment the ROIs in the retinal OCT images. Thereafter, the network has been trained with retraining of one dimensional layers to classify the images accurately. A blind test accuracy of 96.5% has been obtained by the proposed technique. Therefore, the proposed methodology can be used as a supportive measure to help medical professionals with their diagnostic prediction and classification.

Algorithm 3: Proposed Training Algorithm using Deep Multilayered CNN Architecture

- Create a 5-layer CNN with 2 Convolution layers and 1 MaxPool layer in each layer.
- Train using Adam Optimizer.
- Reduce the learning rate by a factor of 10 after 10 epochs.
- If training loss doesn't reduce for 5 consecutive epochs then stop training.
- Retrain the 1-D layers of the model.

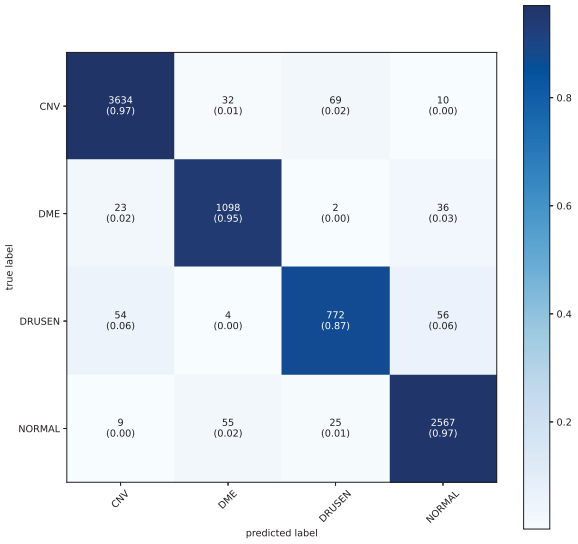

Fig. 8. Confusion Matrix.

TABLE I
Classification Performance

Training Accuracy	Validation Accuracy	Testing Accuracy
0.999	0.959	0.965

Class	Precision	Recall	F1-Score
CNV	0.98	0.97	0.97
DME	0.92	0.95	0.94
Drusen	0.89	0.87	0.88
Normal	0.96	0.97	0.96

Sensitivity	Specificity
0.991	0.979

ISBN: 978-1-7281-6881-4

Part No: CFP20P52-USB

REFERENCES

[1] W. Geitzenauer, C. K. Hitzenberger, and U. M. Schmidt-Erfurth, "Retinal optical coherence tomography: past, present and future perspectives," *British Journal of Ophthalmology*, vol. 95, no. 2, pp. 171–177, 2011. [Online]. Available: https://bjo.bmj.com/content/95/2/171

[2] L. Jian, Y. Panpan, and X. Wen, "Current choroidal neovascularization treatment," *Ophthalmologica*, vol. 230, no. 2, pp. 55–61, 2013.

[3] C. Wilkinson, F. L. Ferris III, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, D. Dills, A. Kampik, R. Pararajasegaram, J. T. Verdaguer *et al.*, "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.

[4] R. L. Novack and R. Y. Foos, "Drusen of the optic disk in retinitis pigmentosa," *American journal of ophthalmology*, vol. 103, no. 1, pp. 44–47, 1987.

[5] F. Girard and F. Cheriet, "Artery/vein classification in fundus images using cnn and likelihood score propagation," in *2017 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2017, pp. 720–724.

[6] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep learning*. MIT press Cambridge, 2016, vol. 1.

[7] B. Sambaturu, B. Srinivasan, S. M. Prabhu, K. T. Rajamani, T. Palanisamy, G. Haritz, and D. Singh, "A novel deep learning based method for retinal lesion detection," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2017, pp. 33–37.

[8] F. Calimeri, A. Marzullo, C. Stamile, and G. Terracina, "Optic disc detection using fine tuned convolutional neural networks," in *2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2016, pp. 69–75.

[9] S. Athar, A. Vahadane, A. Joshi, and T. R. Dastidar, "Weakly supervised fluid filled region localization in retinal oct scans," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 1467–1470.

[10] K. Gopinath and J. Sivaswamy, "Segmentation of retinal cysts from optical coherence tomography volumes via selective enhancement," *IEEE journal of biomedical and health informatics*, vol. 23, no. 1, pp. 273–282, 2018.

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[12] P. K. Sahoo, S. Soltani, and A. K. Wong, "A survey of thresholding techniques," *Computer vision, graphics, and image processing*, vol. 41, no. 2, pp. 233–260, 1988.

[13] R. Arora, A. Basu, P. Mianjy, and A. Mukherjee, "Understanding deep neural networks with rectified linear units," *arXiv preprint arXiv:1611.01491*, 2016.

[14] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.