2020 11th International Conference on Electrical and Computer Engineering (ICECE)

An Automatic Ocular Disease Detection Scheme from Enhanced Fundus Images Based on Ensembling Deep CNN Networks

Ishtiaque Ahmed Khan[1], Asaduzzaman Sajeeb[2], Shaikh Anowarul Fattah[1]
[1]Department of EEE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
[2]Department of CSE, BRAC University, Dhaka, Bangladesh
ishtiaqueahmedkhan@ug.eee.buet.ac.bd, sajeebasaduzzaman@gmail.com, fattah@eee.buet.ac.bd

2020 11th International Conference on Electrical and Computer Engineering (ICECE) | 978-1-6654-2254-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICECE51571.2020.9393050

***Abstract*—Millions of people around the world suffer from various ocular diseases often leading to blindness due to delayed detection and treatment. This has led to a demand for quick automated detection process from medical images including retinal fundus images. In this paper, an automatic approach for classifying normal and diseased cases from given retinal fundus images is developed based on ensembling of some suitable deep learning architectures in a transfer learning platform. Instead of directly using the raw images, it is shown that use of an enhancement technique based on adaptive histogram equalization followed by morphological operations can offer better class separation between the normal and diseased images. Some efficient deep convolutional neural network (CNN) based architectures are implemented utilizing the pre-trained weights obtained via transfer learning. In order to achieve significant improvement in the classification performance, the predictions obtained from some selected deep CNN architectures, namely ResNet50, InceptionResNetV2, EfficientNetB0 and EfficientNetB2 are combined. Comprehensive experimentation carried out on an extensive ophthalmic database show promising performance. The wide range of disease and diverse collection conditions of the fundus images affirm the suitability of the method for practical implementation.**

***Index Terms*—Retinal Fundus, Ocular, CNN, Transfer Learning, Ensemble Network, Classifier.**

I. INTRODUCTION

It has been found that over 2.2 billion people suffer from ocular disorders(OD) such as cataract, age-related macular degeneration (AMD), glaucoma and others [1]. Persons with cataract, longer-standing diabetes, Glaucoma and Macular Degeneration(MD) have a greater risk of developing complications. Cataracts are the largest cause of blindness in the world and together with Glaucoma affects over 70 million people [1, 2]. Hence, timely detection of ocular diseases is important for ensuring proper care and successful recovery.

A commonly adopted technique is fundus screening and a number of works relating to cataract detection from the fundus image have been catalogued and reported on in [3]. Their research found that in sections such as preprocessing or qualified functions, previous works could be enhanced. There are research that also assess the magnitude of cataracts. Convolutional Neural Networks (CNNs) have recently been used significantly in the processing of medical images [4]. In [5], transfer of image feature learning obtained from pre-trained neural networks is explored for the issue of AMD detection. In [6], ROI and Texture analysis and SVM technique have been used for detecting myopia through for optic nerve head detection. In [7], a transfer-learning technique has been used in publicly accessible datasets such as DRIVE, ORIGA and RIM ONE to diagnose intra ocular pressure in the optic nerve. Although the coexistence of multiple eye disorders has been found to affect the efficiency of the single-disease model classification[8], few studies have simultaneously considered various diseases[9]. In addition, several studies independently processed bilateral eyes' CFPs. In [10], an attention-based unilateral and bilateral weighting and fusion function network (AUB-Net) has been used to automatically classify patients into the required categories of disease. The authors determine severity scoring for age-related macular degeneration using an Inception-v3 based multi-part deep network in [11]. Ensemble techniques have proven to be an attractive method for improving deep learning performance. In [12], a multi-stage transfer-leaning approach has been utilized, where a pre-trained CNN on non-medical images is fine-tuned to classify malignant and benign masses in digital breast imaging data. The authors combined multiple ensemble algorithms to improve performance on various UCI databases in [13].

In this paper, we propose an ensemble approach to identify ocular disease utilizing data from fundus images to differentiate OD affected people. The raw fundus images from the database are first processed through enhancement algorithm to counter some quality loss inherent to the acquisition procedure. Secondly, some well known deep neural networks trained on large image databases for day to day object classification are modified and fine tuned to adapt to the current ocular disease recognition task. These networks individually show decent detection capability. However, for further increase in performance, the outputs of the networks are combined and final decision is made from this ensemble. The validity of the proposed scheme is tested using a publicly accessible fundus database.

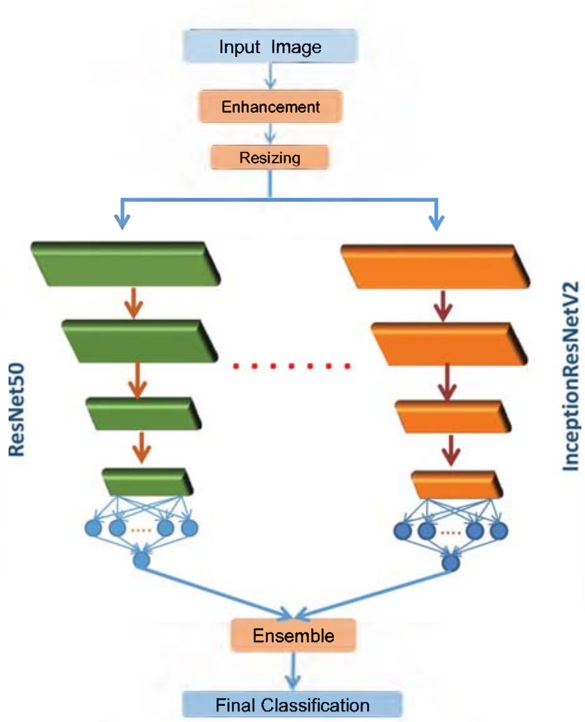


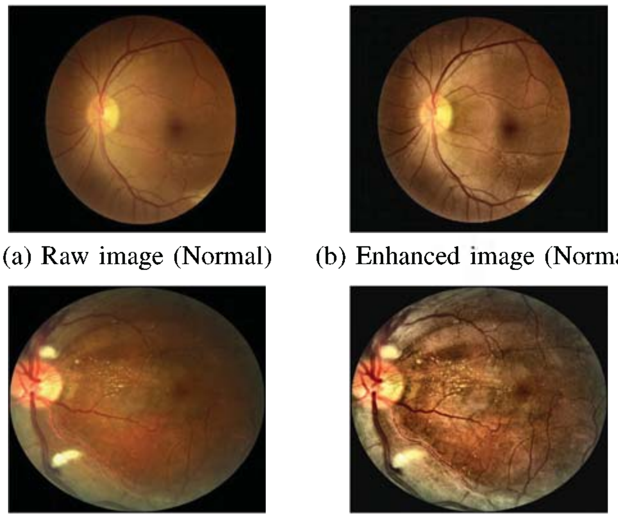
Fig. 1. **Flow chart of the proposed method.**

II. Methodology

The structural representation of the proposed method shown in Fig.1 consists of several stages. Firstly, the quality of the input image is augmented using image enhancement. To accommodate images of varying resolution, incoming images are resized. Inside the ensemble network, pre-trained models fine tuned using the database are utilized to produce preliminary scores. These are averaged and rounded to obtain the final classification label.

A. Pre-processing

Various factors during fundus image acquisition can introduce noise, low contrast and other defects into the image, further complicating the detection task. Hence, to complement the detection procedure of the network, appropriate image processing techniques can be applied increasing the learning capacity. In this paper, an image enhancement technique based on adaptive histogram equalization is applied on every image. In this case, intensity distribution of the image is modified adaptively with an objective to improve the local contrast. A given image is divided into small overlapping tiles and a clipping threshold is adapted so that the noise amplification can be controlled in case of noisy images. Moreover, in order to further enhance the characteristic features in an image differentiating the normal and disease classes, an enhancement block using some image morphological operations is also employed. The samples in Fig. 2 show that this can highlight intricate features integrated in the image. The database contains images of various resolutions.



(a) Raw image (Normal) (b) Enhanced image (Normal)

(c) Raw image (Disease) (d) Enhanced image (Disease)

Fig. 2. **Illustration of the enhancement effect.**

Hence, along with enhancement, the images are all resized to a same size.

B. Transfer Learning

Transfer Learning is a process consisting of leveraging features learned by networks trained on large databases for solving a problem similar to the task at hand. Here, the main idea is to utilize the generic features learned by the layers close to the input while fine-tuning the deeper layers to generate optimized features suited to the target domain. This is specially convenient for tasks where the target domain has very limited data. If the concerned problems are similar, fine tuning only the final few layers generally produce satisfactory results. However, the majority of the available pre-trained models are trained on regular day to day objects which are quite different from fundus images. Hence, for the current study, training of all the layers was necessary. Nevertheless, the initial weights help to kick start the process and supplement the learning, reducing the train time and helping to overcome the data limitation. We considered various such networks including ResNet50, InceptionResNetV2, EfficientNetB0 and EfficientNetB2.

C. Ensemble network

Ensemble network is an algorithm for combining multiple machine learning techniques to generate one robust predictive model. Variants of this include bagging, boosting, stacking and more modified versions. Here, we adopt stacking which exhibits a soothing influence by simultaneously considering information produced by multiple models. Initially, we train each base pre-trained network separately. During testing, we average the outputs of multiple networks and make a final decision. For a test sample x, the corresponding predicted label y' can be expressed as:

$$y' = \left\lfloor \frac{\sum_{i=1}^{N} p'_i}{N} \right\rceil \quad (1)$$

Here, p'_i is the output of the i-th model and N is the cardinality of the set of models. The ensemble concept highlights the strength of each model leading to improved performance.

For training each individual model, as we have 2 possible output classes, we use a binary cross entropy loss function defined as,

$$L(w) = -\frac{1}{n}\left[\sum_{i=1}^{n} y_i \ln(\tilde{y}_i) + (1 - y_i)\ln(1 - \tilde{y}_i)\right] \quad (2)$$

and Adam optimizer with learning rate of 10^{-4} is used.

III. Database

For evaluation of the proposed method, a large ophthalmic disease dataset of paired fundus images of 5000 patients collected from different hospitals and medical centers in China is utilized [14]. The testing data annotations have not been made public. Hence, we measure performance on the training data and partitioned it into train, test and validation sets. The database includes multiple images not properly captured due to faulty equipment and other environmental conditions which can introduce irregularity in the training. Thus, some images were ignored. We consider a binary classification between normal and disease afflicted fundus image. There are several diseases in the database such as normal, diabetes, glaucoma, cataract, AMD, hypertension, myopia and others. All disease annotations are collectively labeled as disease for our purpose. In this paper, diabetic retinopathy was excluded. The images were captured by the staff of many different institutions using equipment of varying configurations, resolutions and calibrations. This aspect along with the wide range of diseases makes this quite a challenging database.

IV. Results and Discussions

We trained separate networks for the left and right eye images. Several different networks were applied. Among them ResNet50, InceptionResNetV2, EfficientNetB0 and EfficientNetB2 showed promising results. The performance of the fine tuned models on the left and right eye test sets were averaged and are reported in Table I. Here, the training and performance evaluation were implemented on non-enhanced raw images.

Due to the diverse acquisition environments faced by the images in the database, there may be some CFPs with low quality. Hence, image enhancement can help increase the performance. To test this, we trained and tested on the same images as before but now they had been processed through enhancement technique. Improvement in the performance was observed as evidenced by the results in Table II.

TABLE I
Performance of the Networks on Raw Images.

Model Name	Accuracy	Sensitivity	Specificity
ResNet50	80.33%	78.74%	81.91%
EfficientNetB0	79.51%	76.61%	82.40%
EfficientNetB2	79.50%	77.21%	81.80%
InceptionResNetV2	80.77%	76.49%	85.04%
Ensemble	82.05%	79.34%	84.76%

TABLE II
Performance of the Networks on Enhanced Images.

Model Name	Accuracy	Sensitivity	Specificity
ResNet50	82.57%	79.34%	85.81%
EfficientNetB0	80.63%	77.37%	83.89%
EfficientNetB2	81.67%	75.45%	87.89%
InceptionResNetV2	84.22%	78.74%	89.70%
Ensemble	86.08%	81.26%	90.90%

We show the increase in accuracy graphically in Fig. 3. From both tables it can be seen that ensemble of multiple networks perform better than any individual one. One important point to note is that the ensemble network can be made up of various combinations of networks. The performance obtained will vary according to the combination. We explored the choice of ensemble combinations to find the optimal combinations. The average performance of five combination of networks for enhanced left and right eye fundus images are represented in Table III. Here, in each case there are at least three models included in the network. Using all models together achieves the best detection results.

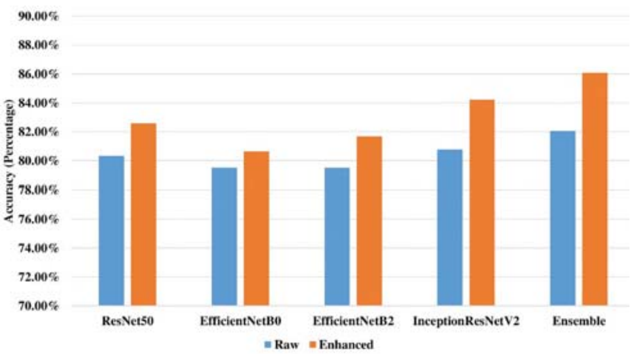
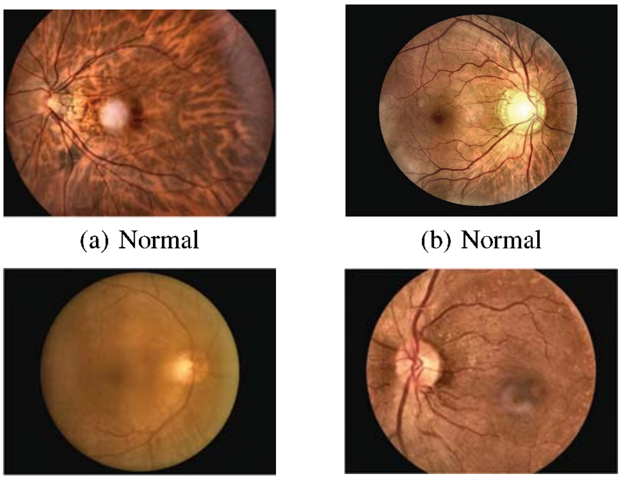


Fig. 3. **Accuracy comparison of the models for raw and enhanced images.**

TABLE III
Performance Comparison of Different Combinations of Ensemble Network.

Model Excluded	Accuracy	Sensitivity	Specificity
None	86.08	81.26	90.90
EfficientNetB2	83.40	78.58	88.22
EfficientNetB0	84.52	79.78	89.26
InceptionResNetV2	83.62	79.02	88.22
ResNet50	83.84	78.41	89.26



(a) Normal (b) Normal

(c) Cataract (d) Hypertensive Retinopathy

Fig. 4. **Sample images detected by the proposed method.**

A few sample images correctly classified by our process may be observed in Fig. 4. As can be seen, the similarity between normal and disease afflicted images can become substantial. This again highlights the efficacy of the proposed enhanced ensemble approach. A basic CNN model based on the network in [15] was applied for comparison. For fair comparison, the image size was kept same as in this paper and converted to grey scale. A few extra convolutional blocks had to be added to keep the parameters within limit. The performance obtained was lower than that of the proposed ensemble approach as shown in table IV.

TABLE IV
PERFORMANCE COMPARISON WITH A MODEL FROM [15].

Model Name	Accuracy	Sensitivity	Specificity
Proposed	86.08	81.26	90.90
CNN in [15]	77.18	79.13	75.23

V. CONCLUSION

A novel automated algorithm is proposed for detecting ocular disease from retinal fundus images. Instead of training a new network from scratch, transfer learning is applied to gain important insight from training done using other large image databases. An adaptive equalization enhancement technique has been employed to improve image quality and highlight the difference between the two classes boosting the classification performance. Ensemble algorithm helps to combine the individual deep neural networks and reinforce the strength of each one and has shown better performance than any individual network. Evaluation using a large ophthalmic disease database proves the validity of the method. Comparison using an existing network has shown this method to be superior. Such an automated process can help quickly classify between disease afflicted and healthy subjects allowing health care professionals to focus on those who need immediate attention. Inclusion of more networks into the ensemble and more complex ensmebling techniques along with an increase in data may lead to further improvement.

VI. ACKNOWLEDGMENT

The authors would like to thank the Department of Electrical and Electronic Engineering of Bangladesh University of Engineering and Technology (BUET) for constant support.

REFERENCES

[1] A. Ram and d. C. C. Reyes-Aldasoro, "The relationship between fully connected layers and number of classes for the analysis of retinal images," *arXiv preprint arXiv:2004.03624*, 2020.

[2] R. N. Weinreb, T. Aung, and F. A. Medeiros, "The pathophysiology and treatment of glaucoma: a review," *Jama*, vol. 311, no. 18, pp. 1901–1911, 2014.

[3] B. R. Kumar and M. Shimna, "Recent approaches for automatic cataract detection analysis using image processing," *Journal of Network Communications and Emerging Technologies (JNCET)*, vol. 7, no. 10, 2017.

[4] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[5] P. Burlina, D. E. Freund, N. Joshi, Y. Wolfson, and N. M. Bressler, "Detection of age-related macular degeneration via deep learning," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2016, pp. 184–188.

[6] J. Liu, D. W. Wong, J. H. Lim, N. M. Tan, Z. Zhang, H. Li, F. Yin, B. Lee, S. M. Saw, L. Tong *et al.*, "Detection of pathological myopia by pamela with texture-based features through an svm approach," *Journal of Healthcare Engineering*, vol. 1, 2010.

[7] B. Elakkiya and O. Saraniya, "A comparative analysis of pretrained and transfer-learning model for automatic diagnosis of glaucoma," in *2019 11th International Conference on Advanced Computing (ICoAC)*. IEEE, 2019, pp. 167–172.

[8] R. R. Bourne, G. A. Stevens, R. A. White, J. L. Smith, S. R. Flaxman, H. Price, J. B. Jonas, J. Keeffe, J. Leasher, K. Naidoo *et al.*, "Causes of vision loss worldwide, 1990–2010: a systematic analysis," *The lancet global health*, vol. 1, no. 6, pp. e339–e349, 2013.

[9] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, p. 158, 2018.

[10] J. He, C. Li, J. Ye, S. Wang, Y. Qiao, and L. Gu, "Classification of ocular diseases employing attention-based unilateral and bilateral feature weighting and fusion," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2020, pp. 1258–1261.

[11] Y. Peng, S. Dharssi, Q. Chen, T. D. Keenan, E. Agrón, W. T. Wong, E. Y. Chew, and Z. Lu, "Deepseenet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs," *Ophthalmology*, vol. 126, no. 4, pp. 565–575, 2019.

[12] R. K. Samala, H.-P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter, and K. H. Cha, "Breast cancer diagnosis in digital breast tomosynthesis: effects of training sample size on multi-stage transfer learning using deep neural nets," *IEEE transactions on medical imaging*, vol. 38, no. 3, pp. 686–696, 2018.

[13] G. I. Webb and Z. Zheng, "Multistrategy ensemble learning: Reducing error by combining ensemble learning techniques," *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 8, pp. 980–991, 2004.

[14] "Peking university international competition on ocular disease intelligent recognition (odir-2019)." [Online]. Available: https://odir2019.grand-challenge.org/dataset/

[15] M. T. Islam, S. A. Imran, A. Arefeen, M. Hasan, and C. Shahnaz, "Source and camera independent ophthalmic disease recognition from fundus image using neural network," in *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*. IEEE, 2019, pp. 59–63.