AN EFFICIENT FULLY CONNECTED NEURAL NETWORK FOR MICROANEURYSM DETECTION FROM RETINAL FUNDUS IMAGES

Mihir Deshmukh[1,2]*, Shibasish Chowdhury*[1] *and Chandra Sekhar Seelamantula*[3]

[1]Department of Biological Sciences, BITS Pilani, Pilani, India
[2]Department of Electronics and Instrumentation, BITS Pilani, Pilani, India
[3]Department of Electrical Engineering, Indian Institute of Science, Bangalore, India
mihirsdeshmukh10@gmail.com, shiba@pilani.bits-pilani.ac.in, css@iisc.ac.in

ABSTRACT

We propose a deep learning technique for the prognosis of diabetic retinopathy by the automatic detection of small lesion-like blood clots called microaneurysms in high-resolution fundus images. Microaneurysms mark the onset of diabetic retinopathy and can be detected as a warning, enabling treatment to arrest the progress of the disease. We modified the fully connected U-Net architecture to a lightweight and more specialized model with Squeeze-and-Excitation blocks. The proposed model has 5.1 million parameters and the predicted False Positives per image do not exceed 1.22. We achieved a free-response receiver operating characteristic score of 0.675 on the RC-RGB-MA dataset, outperforming the state of the art by 4%.

***Index Terms*—** Microaneurysm detection, deep learning, fundus image, squeeze-excitation (SE), diabetic retinopathy (DR).

1. INTRODUCTION

Diabetic retinopathy (DR), if not treated in the early stages, can lead to partial or complete blindness. The ICDR severity scale [1] proposes microaneurysms, which are small blood clots in tertiary vessels or smaller capillaries, as the single characteristic marking the distinction between a healthy eye (DR grade 0) and a diseased eye (DR grade 1 and above) eye. Thus, the detection of MAs is the most important feature for the diagnosis of DR. The task is tedious if done manually by ophthalmologists as is being done currently. The present work is an attempt to automate it, thus reducing the error introduced by human intervention as well as saving numerous hours of manual labor. We propose a deep learning based fully automated network for the detection of microaneurysms (MAs). We trained it on the RC-IOSTAR database, RC-RGB-MA, which contains RGB fundus images and is annotated by two experts. Our method uses a variation of the U-Net [2] with squeeze-excite blocks [3], while having 5.1 million parameters to train. Our model achieved a free-response receiver operating characteristic (FROC) measure of 0.675, outperforming the existing state of the art by 4%. The proposed work is highly specific, with no more than 1.22 false positives per image.

2. PRIOR WORK

Microaneurysm detection, with a sparse distribution of true positives, has been the challenging task at the focus of various retinal studies. Significant advancements have been made over the past few years both on the conventional image processing front and the deep learning front. Recently, Chudzik et al. [4] used a fully connected network which is a variation of the U-Net [2] first proposed by Ronneberger et al., with alternate layer freezing while transfer learning. Lam et al. [5] used a CNN based classifier for the multi-class classification of image patches for detecting microaneurysms and exudates. The non-deep-learning approaches such as those proposed by Dashtbozorg et al. [6] or the works described in the Retinopathy Online Challenge [7] typically work in three steps: (i) extracting circular candidates; (ii) extracting features using pre-engineered features; (iii) classifying the candidates as MAs or not MAs.

The only steps common to both the approaches is the preprocessing of fundus images. We have listed a few of the common preprocessing techniques in Table 1.

3. PROPOSED METHOD

The proposed U-Net based architecture uses squeeze-excitation blocks [3] for giving proportionate importance to the individual channels. Histogram equalization on the green channel of the fundus image is applied to highlight microaneurysms. The images are divided into overlapping patches of size 128×128 pixels with a stride of 32×32 pixels. The network was trained only on valid positive and valid negative patches where the validity of a patch was established by hierarchical elimination and subsequently, random sampling for obtaining a balanced dataset.

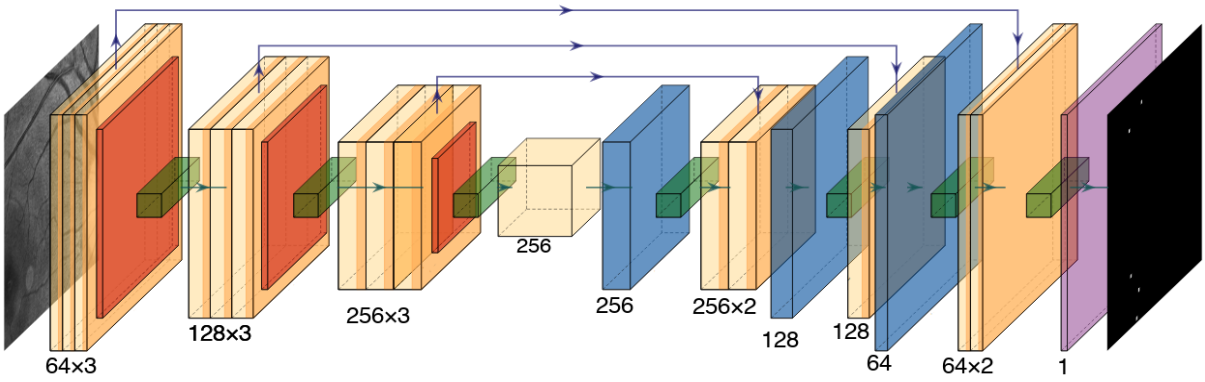

Fig. 1. The proposed modified U-Net Model structure. The yellow layers indicate Conv2D and BatchNormalization layers; The red layers indicate MaxPool layers; The blue layers indicate Conv2DTranspose. The green layers indicate SE Blocks.

Table 1. Review of common preprocessing techniques for retinal fundus image preprocessing and detection of microaneurysms and lesions

Preprocessing Steps	**Done by**
Illumination Normalization (given by Forachhia et al. [8] or Radke et al. [9])	[6], [10], [11]
Subtraction of low-frequency components, (obtained by subtraction of the Gaussian or median blur of the original image)	[4], [6], [7], [12], [13], [14]
Removal of Vasculature	[11], [15]
Green Channel Extraction	[4], [6], [10], [11], [15], [16],
Histogram Equalization	[15],[16]

3.1. Preprocessing

It is difficult to identify the presence of microaneurysms as localized features with respect to their surroundings. They can be highlighted by enhancing the contrast between dark MAs and a lighter background. We used Contrast Limited Adaptive Histogram Equalization (CLAHE) [17]. The model was trained on a concatenated input with CLAHE on the green channel of the original image and the original green channel as shown in Fig. 2. The green channel was used as it provides the highest contrast without being over-saturated [18].

3.2. Model

We adapted the U-Net architecture to the task of detecting lower-level features by removing two downsampling and corresponding upsampling layers. This is because microaneurysms are spot-like features that can be identified by the lower levels of the network. The addition of squeeze-excitation blocks aids in giving prominence to the relevant channels. This addition is done before the branch aggregation in the decoder part according to the advised placement of the SE-PRE block as described by Hu et al. [3]. Trying other variations of the Channel SE block, as performed by Roy et al. [19], like spatial squeeze-excitation or spatial channel squeeze-excitation did not give optimal performance. We used a pixel-wise mean-squared error function. The model did not converge after a Dice-based loss function, even after having weighted penalties. This can be attributed to the imbalance in the number of MA pixels to the non-MA pixels. The model was not trained on the FROC score as it is not amenable to computing at the pixel level.
Our model minimizes the following loss:

$$L(y, \hat{y}) = \frac{1}{n}\|y - \hat{y}\|^2$$

where n is the number of pixels in a image patch and y and $\hat{y}$ is the ground truth and prediction respectively.

3.3. Training

Even though the U-Net proposed by Ronenberger et al. has a large receptive field and proposes processing of the image as a whole, owing to the sparsity of the distribution of MAs and their small size, we trained our network using overlapping patches. The dataset contains 250 images which are 1944×2592 pixels. The ground truth annotations are done by two experts. The annotations are 11×11 pixels squares around each MA. We used the union of the annotation maps to create the input annotation map on which our network was trained.
The images obtained after preprocessing are divided into overlapping patches of size 128×128 pixels with a stride window of 32×32 pixels. To achieve a balanced dataset, undersampling of the patches without the presence of an MA signal was carried out. The patches that have more than 80% of their area as a part of the Field of View (FoV) are considered valid and were retained. It was empirically determined that the network attains faster convergence and maintains high specificity if a patch with at least one entire MA strictly

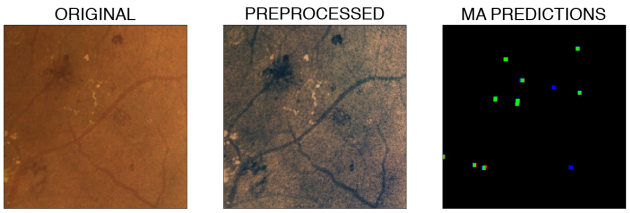

Fig. 2. Comparison of the input image, the preprocessed image, and visualization of model predictions with blue marks denoting FN, red marks denoting FP, and green marks denoting TP, while the black background being TN.

Table 2. Comparison of number of parameters for other fully connected networks used in medical imaging.

Method	Parameters (M: Million)
U-Net [2]	7.76M
Wide U-Net [20]	9.13M
UNet++[20]	9.04M
Proposed	**5.14M**

in the central one-quarter of the area as per the corresponding ground truth patch is considered a positive sample, and all other patches, even if they contain a fraction of the MA are considered negative. This reduced the number of false positives and fringe effects near the correct predictions greatly, while still being able to detect lesions around the edges of the FoV. The stride length was adjusted to get at least one valid and positive patch per annotated MA. A negative sample is a sample that is valid but does not contain a complete MA annotation in the central quarter area. They were sampled to be uniformly distributed throughout the fundus image and be the same in number as the positive samples. The valid positive and negative samples are divided into training and validation sets in the ratio 7:3.

4. RESULTS

The metric used for comparison of results is the Free Receiver Operating Characteristic. The thresholding of the probability values of the predictions gives the number of False Positives (FP) per image and True Positives (TP) per image. For this calculation, we constrain the maximum permissible distance between the center of a predicted mark and the center of the corresponding annotated mark for the prediction to be considered a TP signal. This distance is kept equal to the dataset's ground truth annotation (a square of 11×11 pixels) dimensions, thereby enforcing a positive overlap between the prediction and the ground truth. Thus, a signal is considered to be a true positive marking only if its center appears within an 11-pixel distance of the center of the original signal. Otherwise, it is classified as a false positive. A ground truth marking that does not have any prediction marking center in the 11-pixel vicinity of its center is considered a true negative (TN). Fig. 2 shows the visualization of the model predictions with the original image and the preprocessed image inputs for comparison.

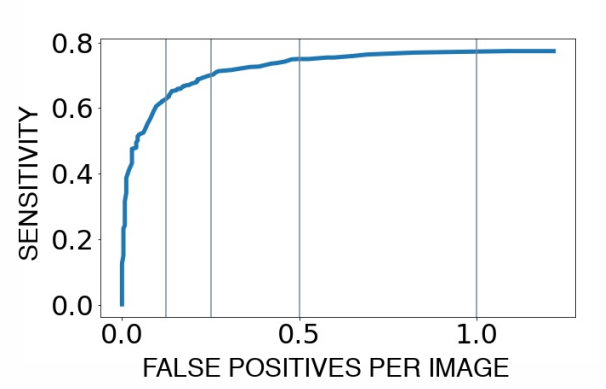

Fig. 3. FROC Curve drawn at varying values of likelihood threshold, giving a sensitivity vs number of FPs/image.

Table 3. Comparative performance versus state-of-the-art methods on RC-RGB-MA database.

Method	FROC
Dashtbozorg et al. [6]	0.647
Zhao et al. [21]	0.653
Yan et al. [22]	0.653
Proposed	**0.675**

The TPs and TNs are used to calculate sensitivities at different values of thresholding probabilities, which are plotted against corresponding FPs/image to give the FROC curve as shown in Fig. 3. The method gives highly specific predictions and does not exceed 1.22 False Positives per image, even at the lowest probability threshold. An overview of the performance accuracies of other works on the RC-RGB-MA dataset is given in Table 3. Table 2 compares the number of parameters of other fully connected networks used in medical imaging.

5. CONCLUSION

This work is aimed at finding an efficient method for the prognosis of diabetic retinopathy, the onset of which is marked by the appearance of microaneurysms. We propose a modification to the U-Net architecture, which outperforms the existing state-of-the-art models, while also having significantly fewer parameters and also being highly specific. Although the quantitative analysis of the difference in performance accuracy due to the implementation of different preprocessing strategies is beyond the scope of this study, we have described the preprocessing methods adopted by us, which are giving the best results.

6. ACKNOWLEDGEMENTS

We would like to acknowledge P. Kevin Raj from Spectrum Lab, Department of Electrical Engineering, IISc Bangalore for his contributions to this paper. His inputs and suggestions proved vital to the success of this endeavor.

7. REFERENCES

[1] C. P. Wilkinson et al., "Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales," *Ophthalmology*, vol. 110, no. 9, pp. 1677–1682, 2003.

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. on MICCAI*, 2015, pp. 234–241.

[3] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Int. Conf. on CVPR*, 2018, pp. 7132–7141.

[4] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 185–192, 2018.

[5] C. Lam, C. Yu, L. Huang, and D. Rubin, "Retinal lesion detection with deep learning using image patches," *Investigative ophthalmology & visual science*, vol. 59, no. 1, pp. 590–596, 2018.

[6] B. Dashtbozorg, J. Zhang, F. Huang, and B. M. ter Haar Romeny, "Retinal microaneurysms detection using local convergence index features," *IEEE Trans. on Image Processing*, vol. 27, no. 7, pp. 3300–3315, 2018.

[7] M. Niemeijer et al., "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE Trans. on Med. Imag.*, vol. 29, no. 1, pp. 185–195, 2009.

[8] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical Image Analysis*, vol. 9, no. 3, pp. 179–190, 2005.

[9] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE transactions on image processing*, vol. 14, no. 3, pp. 294–307, 2005.

[10] K.edir M. Adal, Peter G othersVan Etten, Jose P Martinez, Kenneth W Rouwen, Koenraad A Vermeer, and Lucas J van Vliet, "An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images," *IEEE Trans. on Biomedical Engineering*, vol. 65, no. 6, pp. 1382–1390, 2017.

[11] R. Wang, B. Chen, D. Meng, and L. Wang, "Weakly-supervised lesion detection from fundus images," *IEEE Trans. Med. Imag.*, vol. 38, no. 6, pp. 1501–1512, 2018.

[12] Yehui Yang, Tao Li, Wensi Li, Haishan Wu, Wei Fan, and Wensheng Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2017, pp. 533–540.

[13] M. JJP Van Grinsven et al., "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, 2016.

[14] C. I. Sánchez et al., "Mixture model-based clustering and logistic regression for automatic detection of microaneurysms in retinal images," in *SPIE: Computer-Aided Diagnosis*, 2009, vol. 7260, p. 72601M.

[15] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1720–1726, 2012.

[16] A. Sopharak, B. Uyyanonvara, and S. Barman, "Simple hybrid method for fine microaneurysm detection from non-dilated diabetic retinopathy retinal images," *Computerized Medical Imaging and Graphics*, vol. 37, no. 5-6, pp. 394–402, 2013.

[17] K.arel Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*. Academic Press Professional, Inc., 1994, pp. 474–485.

[18] S. Ravishankar, A. Jain, and A. Mittal, "Automated feature extraction for early detection of diabetic retinopathy in fundus images," in *Proc. IEEE Int. Conf. on CVPR*, 2009, pp. 210–217.

[19] A. G. Roy, N. Navab, and C. Wachinger, "Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks," in *Proc. Int. Conf. on MICCAI*, 2018, pp. 421–429.

[20] Z. Zhou et al., "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 3–11. Springer, 2018.

[21] Y. Zhao et al., "Uniqueness-driven saliency analysis for automated lesion detection with applications to retinal diseases," in *Proc. Int. Conf. on MICCAI*, 2018, pp. 109–118.

[22] Q. Yan et al., "Automated retinal lesion detection via image saliency analysis," *Medical Physics*, 2019.