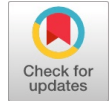# Multi-Scale Feature Pyramid for Detection of Red Lesions in Fundus Images

Goutam Kumar Ghorai, Swagata Kundu, Gautam Sarkar, Ashis Kumar Dhara

*Abstract**: Diabetic retinopathy (DR) is increasing rapidly around the world, but there is a shortage of experienced ophthalmologists. Therefore, computer-based diagnosis of the fundus images is essential to the screening of referable DR. Automated detection of red lesions is very important for the screening of DR. This paper deals with a novel method for automatic detection of red lesions. The main contribution is the development of a deep learning-based detection framework to handle severe class imbalance and size imbalance in red lesions. The multi-scale features are extracted using the feature pyramid network. A pyramid of features is generated with strong semantics. The proposed network is end-to-end trainable at the image level across multiple scales and is effective for a wide range of red lesions, achieving acceptable performance. The sensitivity of the proposed method is 0.76, with six false positives per image on the test set of the publicly available DIARECTDB1 database, and it outperforms state-of-the-art approaches. A potential benefit of a deep learning-based detection framework could be its use in screening programs for referable diabetic retinopathy (DR).*

Keywords: Diabetic Retinopathy, Fundus Images, Red Lesions Detection, Feature Pyramid Network, Focal Loss

I. INTRODUCTION

In the United States, blindness is caused mainly by DR [1]. Early diagnosis of DR is the only way to avoid blindness caused by DR [1]. About 77 million people have diabetes in India, and 18% of them are suffering from DR. DR may lead to blindness if treatment does not start in time [2]. There is a need for regular screening of diabetic patients for early detection of DR because DR is asymptomatic in the early stage. Microaneurysms and hemorrhages (Fig. 1) are called red lesions. Microaneurysms are the early sign of DR, and haemorrhages indicate the severity of DR [2]. Several challenges for detecting red lesions are poor image contrast, non-uniform illumination, and a wide variety of red lesions in terms of shape, which can lead to the missing of red lesions. In India, screening for diabetic retinopathy (DR) is conducted through eye camps, where fundus evaluation is performed for diabetic patients attending the camps.

Goutam Kumar Ghorai*, Department of Electrical Engineering, Jadavpur University, Kolkata (West Bengal), India. E-mail: goutamghorai79@gmail.com, ORCID ID: 0009-0009-7656-0323

Swagata Kundu, Department of Electrical Engineering, National Institute of Technology Durgapur, Durgapur (West Bengal), India. E-mail: swagatakundu2103@gmail.com, ORCID ID: 0000-0003-4453-5552

Gautam Sarkar, Department of Electrical Engineering, Jadavpur University, Kolkata (West Bengal), India. E-mail: sgautam63@gmail.com, ORCID ID: 0000-0002-1330-2636

Ashis Kumar Dhara, Department of Electrical Engineering, National Institute of Technology Durgapur, Durgapur (West Bengal), India. E-mail: akdhara.ee@nitdgp.ac.in, ORCID ID: 0000-0001-8776-3526

Experts and referable DR cases analyse the digital retinal images, which are then segregated for further treatment. Healthy subjects are advised to attend the following eye camps for follow-up. Automated detection of red lesions is affected by low image contrast, non-uniform illumination, and the tiny structures of microaneurysms. The junction of blood vessels has a similarity to red lesions and can cause false positives. In classical image analysis, Mathematical morphology [3], template matching [4], cross-section profiles analysis [5], and multiscale filtering [6] were explored for red lesion detection.

The micro-aneurysms were extracted from other dot-like structures using the Watershed transform [7]. The haemorrhages of several shapes were taken care of by template matching [4]. Radon Cliff operator detected the circular red lesions [8]. Wang et al. [5] performed singular spectrum analysis of cross-sections profiles of microaneurysms. Wavelet transform [9] and curvelet-based method [10] were also explored for red lesion detection.

Deep learning methods are highly popular and widely explored for object detection in natural images. The region-based convolutional neural network (R-CNN) [11] and the spatial pyramid pooling (SPP) [12] networks are examples of two stage detectors. Faster R-CNN [14] uses a region proposal network and is an improvement over Fast R-CNN [13]. The deep features extracted using a region proposal network are utilised to suggest approximate objects, and finally, bounding boxes are detected. It is a single, unified network that is end-to-end trainable and shares computations across all proposals, rather than performing calculations independently for each proposal. The last convolutional layer is used to predict the class scores and corresponding bounding boxes.

The single-shot detector (SSD) [15] and you only look once (YOLO) [16] are examples of single stage detectors, where detection is performed over a dense sampling. The performance of SSDs is not suitable for small objects. Object detection in YOLO is considered a regression problem and exhibits better generalisation for new domains. The main improvements include the use of batch normalisation, higher-resolution inputs, convolutional layers with anchors, dimensionality clustering, and fine-grained features.

Deep neural network is taught to recognise pathological lesions from retinal fundus images for the diagnosis of DR. Grinsven et al. [17] applied CNN to detect haemorrhages on patch classification, where the negative patches are extracted from images without haemorrhages, and the positive patches are extracted only from haemorrhage locations. Training was made faster by dynamically selecting misclassified negative samples. This method operates at the patch level and is not trainable in an end-to-end manner.
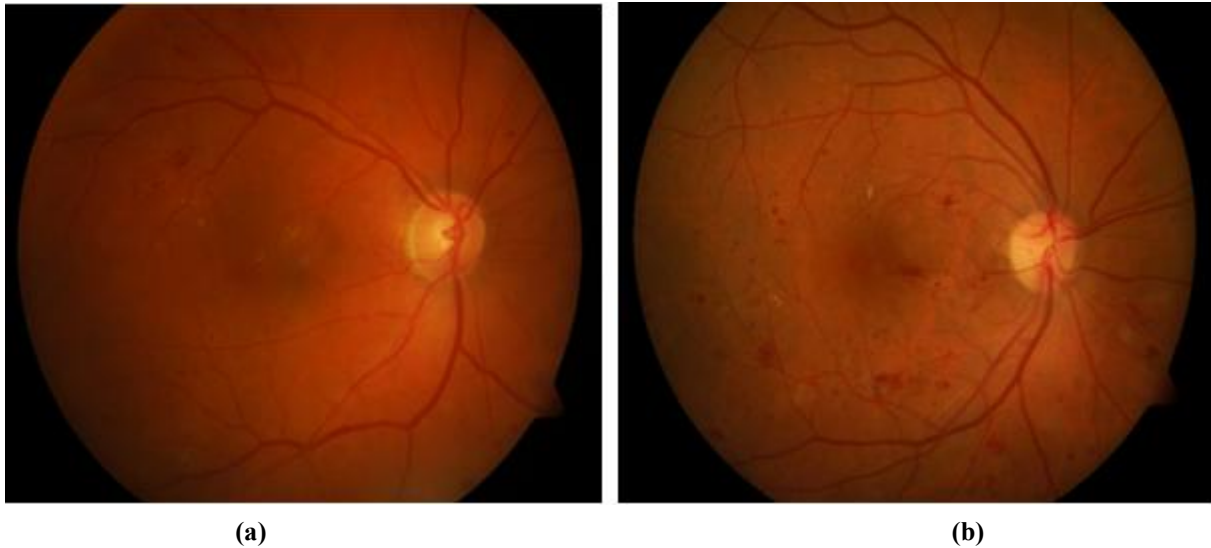
(a) **(b)**

Fig. 1: (a) and (b) Sample Relational Fundus Images from DIARECTDB1 [20] with Non-Uniform Illumination

We present a single-stage detection framework that works at the image level and is end-to-end trainable. The architecture combines a backbone for generating a feature map and a feature pyramid network (FPN) [19] with skip connections for effectively combining the features of lower resolution with features of the higher resolution images. Finally, classification and regression heads are integrated to recognize the red lesions with their locations. The FPN addresses several limitations of previous architectures and creates a feature pyramid with strong semantic features at multiple scales. The proposed network can be trained at the image level for all possible scales and is effective for a wide range of red lesions.

II. METHODS

Object detection at a wide range of scales is a fundamental challenge, and the feature pyramid is used as a neck in the architecture, along with a classification and regression head. The feature extracted by convolutional networks is used to create a feature pyramid. The pyramids are scale-invariant in the model.

Can detect objects of different shapes. The proposed network features a top-down and a bottom-up architecture with skip connections, enabling the generation of more robust semantic features. Non-uniform illumination and poor contrast are observed in fundus images. Therefore, a preprocessing method has been developed to enhance image contrast, making red lesions visible in the preprocessed images.

A. Preprocessing

The maximum illumination in the fundus image is observed near the optic disc, and the boundary region has less illumination. Therefore, illumination equalisation is crucial as a preprocessing step. The fundus portion is extracted from the large-sized fundus images using template matching, and then the images are resized to 512×512 to expedite the process. A contrast enhancement is performed as follows.

$$I'(x, y; \sigma) = \alpha I(x, y) - \beta G(x, y; \sigma) * I(x, y) + \gamma \quad (1)$$
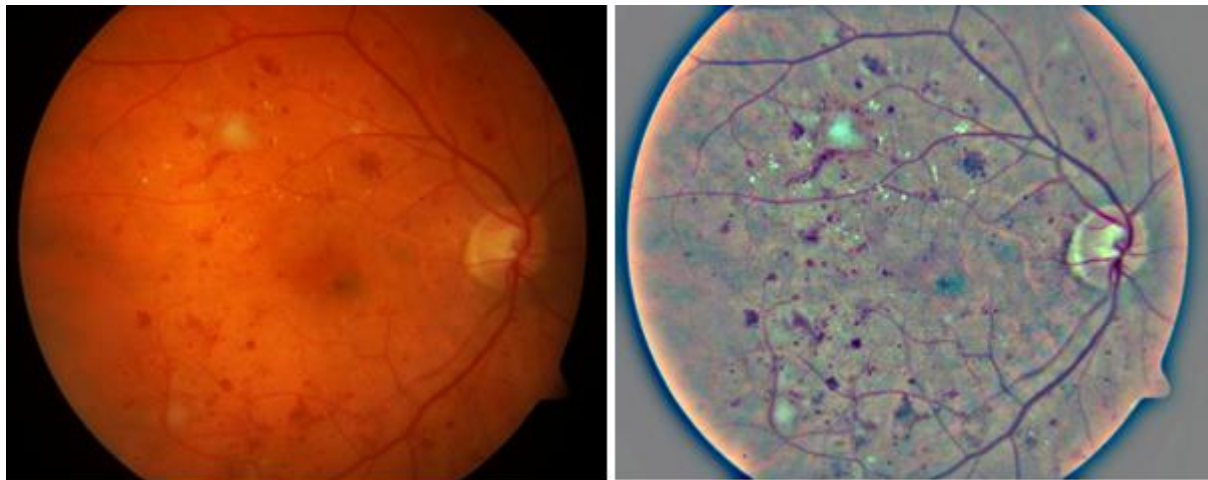

Fig. 2: Preprocessing Step: (a) Fundus Image and (b) Fundus Image After Preprocessing

Where the convolution operation is performed with a Gaussian filter of scale σ. The values of the parameters are $\alpha = \beta = 4$, and $\gamma = 128$ For DIARECTDB1 database. The parameters are chosen empirically. The preprocessed image in Fig. 2(b) exhibits high contrast, providing better visibility of the red lesions compared to the low-contrast input image in Fig. 2(a).

B. Network Architecture

One-stage detectors are very popular due to their simpler design and acceptable performance. The proposed detection framework (Fig. 3) is one stage and has similarities with RetinaNet [18]. The network comprises a backbone for feature extraction, known as the Feature Pyramid Network (FPN), and an object detector head. The CNN is used to build the feature pyramid, and throughout the pyramidal structure, high-level semantics are observed. The proposed method takes a preprocessed fundus image as input and creates a feature map with several resolutions. The FPN helps combine lower-resolution features with higher-resolution features. The FPN provides a semantically strong, multi-scale feature map that encompasses all possible scales.

The bottom-up pathway computes a feature map with multiple resolutions, generating a feature pyramid. The top-down up-samples spatially coarser features, then combines them with the feature pyramid via skip connections. The skip connection merges feature maps from the bottom-up pathway and the top-down pathway, considering the same spatial size during the merging process. The bottom-up feature map has the semantics of lower-level as it is sub-sampled fewer times. The top-down pyramid contains high semantic information and has fine resolutions. The classification and regression heads are similar to those used in standard object detection methods.

The proposed network is end-to-end trainable across multiple scales, enabling the detection of red lesions with a wide range of shapes and sizes. The focal loss [18] is considered for training of the network on complex examples as compared to easy negatives lesions or background.

The focal loss is an improvement of cross entropy by dynamically scaling the cross entropy loss so that little loss is assigned to well-classified examples and incurs a high loss value for red lesion candidates, which are difficult to detect.

C. Anchor Box Tuning

Several detectors, such as Faster R-CNN [15], SSD [14], Retina Net [16], and YOLOv3 [18] use pre-defined boxes and could hamper the ability of generalizing detectors because they need tuning onmfresh detection tasks. In this work, suitable anchor boxes were generated adaptively based on the statistical analysis of the sizes of red lesions on the test set of DIARECTDB1 [20] public dataset. In this application, anchor boxes are available in three scales and four aspect ratios to detect both large red lesions with imbalances in object sizes and shape variations, as well as tiny red lesions.

D. Network Training

The proposed method is implemented in PyTorch, and training is performed on a workstation equipped with a 16 GB GPU and 64 GB RAM. The training is performed using the pre-processed fundus images of the training set of DIARECTDB1 [20] public dataset up to 150 epochs. The learning rate is 0.00001. The area of the regions of interest is tiny compared to the size of the image and could result in severe class imbalance. Therefore, an irrelevant background region dominates training, and the easily classifiable background region leads to poor model learning. The focal loss can handle class imbalance, and the network is being trained using complex examples, i.e.. These red lesions are challenging to detect, compared to easy examples, i.e., red lesions that are easy to detect.
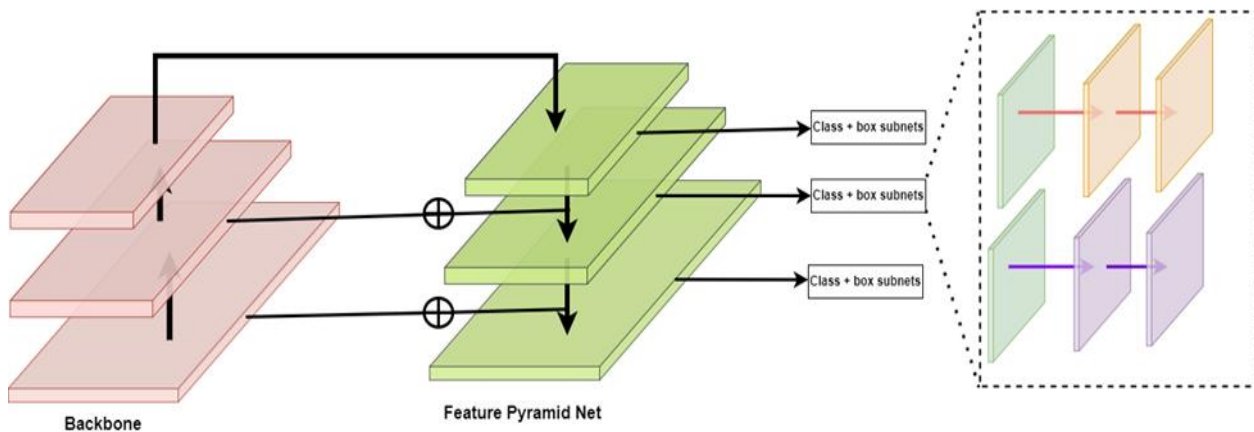

Fig. 3: Multi-Scale Feature Pyramid with Classification and Regression Head for Red Lesion Detection

III. RESULTS AND DISCUSSIONS

The proposed method is compared with the technique of Seoud et al. [21], and both methods are evaluated on the publicly available DIARETDB1 dataset [20]. A total of 89 fundus images are available in the database, which were taken using a 50-degree field of view. Out of 89 images, 84 are classified as mild or non-proliferative diabetic retinopathy (DR), and five are normal. Four expert ophthalmologists annotated several diabetic fundus lesions, including soft exudates, hard exudates, microaneurysms, and haemorrhages, and the ground truths are at a coarse level. The detection algorithms are evaluated using those annotations. In this dataset, detecting red lesions is particularly challenging due to variations in imaging conditions, including saturation, lighting conditions, and different types of blurs. In this experiment, the lesions annotated in the fundus image by at least one ophthalmologist are considered as ground truth. The methods have been evaluated on a per-lesion basis, where a delineation of all the lesions was provided. The test set of the

DIARETDB1 dataset [20] is used for evaluating the methods. The performance is analysed for every single lesion using free-response receiver operating characteristic (FROC) analysis. In the per-lesion evaluation strategy, sensitivities are plotted against the average number of false positives per image for several operating points. The FROC plot of the proposed method and the method of Seoud et. al. [21] is shown in Fig.4. The results depict that the proposed method outperforms the competing methods. The proposed network is end-to-end trainable at the image level with all possible scales and achieves a sensitivity of 0.76 with six false positives per image on the test set of the publicly available DIARECTDB1 database and outperforms the method of Seoud et. al. [21]. The visual results of the proposed method are shown in Fig. 5, and it is observed that red lesions with a wide range of shapes are detected. The performance of detection is better for haemorrhages compared to microaneurysms. A few vessel junctions have similarities with red lesions and appear as false positives. The method has been evaluated with and without preprocessing, and it is observed that preprocessing has a significant role in the performance of red lesion detection.
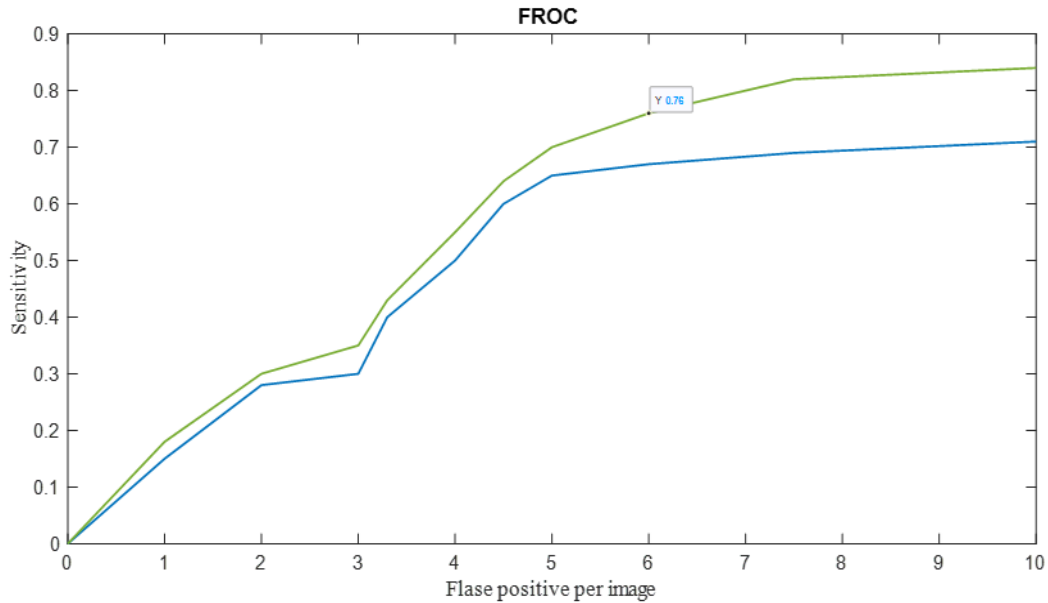

Fig. 4: FROC for the Test Set of DIARECTDB1 Database with the False Positive Per Case Along X-Axis and Sensitivity Along Y-Axis.
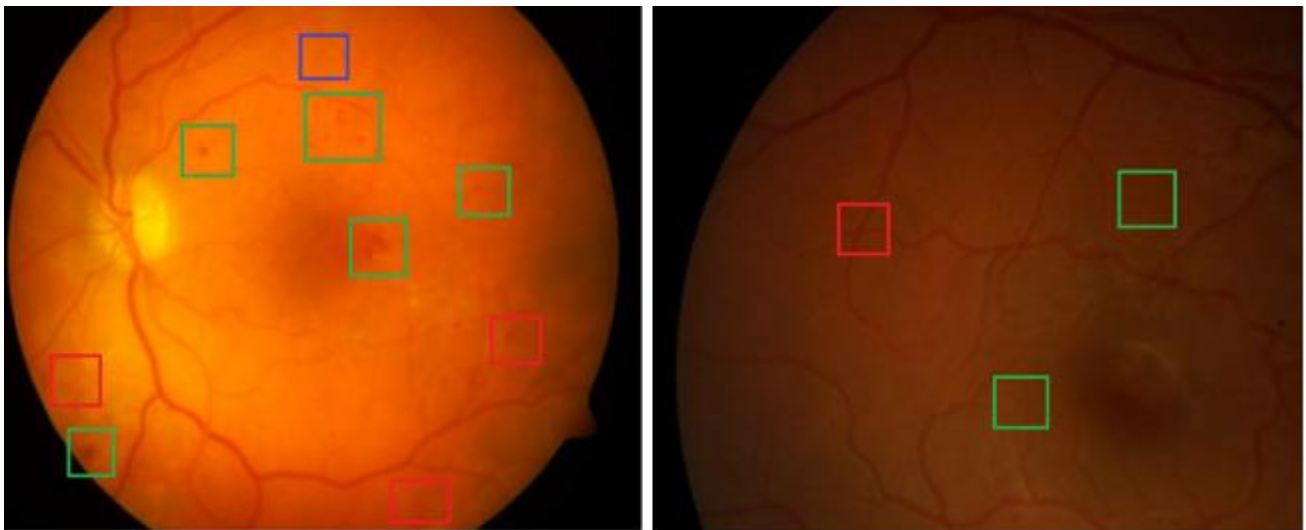

Fig. 5: Visual Results for a Few Test Images of DIARECTDB1 Database, Where True Positives, False Positives, and False Negatives are Marked with Green, Red and Blue Boxes.

IV. CONCULSION

Early detection of red lesions in fundus images has immense potential in clinical practice fead screening programs. Computer-aided screening of DR is essential in handling a large number of patients who need fundus examination. The improved performance of the methods for detecting red lesions could enhance the screening performance. The proposed network detects red lesions with a lower number of false negatives and could be part of a DR screening tool. More work is needed to focus on developing red lesion detection methods with a minimal number of false negatives and false positives, as well as developing robust methods for grading fundus images to aid in the diagnosis of diabetic retinopathy (DR). The preprocessing technique could be improved to enhance the visualisation of small pathologies in fundus images.

DECLARATION STATEMENT

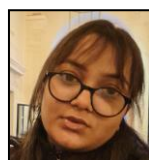|	
Funding	No, I did not receive.
Conflicts of Interest	No conflicts of interest to the best of our knowledge.
Ethical Approval and Consent to Participate	No, the article does not require ethical approval or consent to participate, as it presents evidence.
Availability of Data and Materials	Not relevant.
Authors Contributions	All authors have equal participation in this article.

REFERENCES

1. Ting, D. S. W., Pasquale, L. R., Peng, L., Campbell, J. P., Lee, A. Y., Raman, R., Tan, G.S.W., Schmetterer, L., Keane, P.A. and Wong, T.Y. & Wong, T. Y. (2019). Artificial intelligence and deep learning in ophthalmology. British Journal of Ophthalmology, 103 (2), 167-175. https://doi.org/10.1136/bjophthalmol-2018-313173
2. Raman, R., Srinivasan, S., Virmani, S., Sivaprasad, S., Rao, C., & Rajalakshmi, R. (2019). Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. Eye, 33(1), 97-109. https://doi.org/10.1038/s41433-018-0269-y
3. Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A., & Sharp, P. F. (2006). Automated microaneurysm detection using local contrast normalization and local vessel detection. IEEE transactions on medical imaging, 25(9), 1223-1232. https://doi.org/10.1109/TMI.2006.879953
4. Bae, J. P., Kim, K. G., Kang, H. C., Jeong, C. B., Park, K. H., & Hwang, J. M. (2011). A study on haemorrhage detection using a hybrid method in fundus images. Journal of Digital Imaging, 24, 394-404. https://doi.org/10.1007/s10278-010-9274-9
5. Lazar, I., & Hajdu, A. (2012). Retinal microaneurysm detection through local rotating cross-section profile analysis. IEEE transactions on medical imaging, 32(2), 400-407. https://doi.org/10.1109/TMI.2012.2228665
6. Wang, S., Tang, H. L., Hu, Y., Sanei, S., Saleh, G. M., & Peto, T. (2016). Localising microaneurysms in fundus images through singular spectrum analysis. IEEE Transactions on Biomedical Engineering, 64(5), 990-1002. https://doi.org/10.1109/TBME.2016.2585344
7. Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A., & Sharp, P. F. (2006). Automated microaneurysm detection using local contrast normalization and local vessel detection. IEEE transactions on medical imaging, 25(9), 1223-1232. https://doi.org/10.1109/TMI.2006.879953
8. Giancardo, L., Mériaudeau, F., Karnowski, T. P., Tobin, K. W., Li, Y., & Chaum, E. (2010, March). Microaneurysms detection with the Radon Cliff Operator in retinal fundus images. In Medical Imaging 2010: Image Processing (Vol. 7623, pp. 292-299). SPIE. https://doi.org/10.1117/12.844442
9. Quellec, G., Lamard, M., Josselin, P. M., Cazuguel, G., Cochener, B., & Roux, C. (2008). Optimal wavelet transforms for the detection of microaneurysms in retina photographs. IEEE transactions on medical imaging, 27(9), 1230-1241. https://doi.org/10.1109/TMI.2008.920619
10. Kar, S. S., & Maity, S. P. (2017). Automatic detection of retinal lesions for screening of diabetic retinopathy. IEEE Transactions on Biomedical Engineering, 65(3), 608-618. https://doi.org/10.1109/TBME.2017.2707578
11. Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440-1448). https://doi.org/10.1109/ICCV.2015.169
12. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence, 37(9), 1904-1916. https://doi.org/10.1109/TPAMI.2015.2389824
13. Girshick, R. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440-1448). https://doi.org/10.1109/ICCV.2015.169
14. Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.
15. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). SSD: Single-shot multibox detector. In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14 (pp. 21-37). Springer International Publishing. https://doi.org/10.11648/j.sd.20160404.17
16. Fang, W., Wang, L., & Ren, P. (2019). Tinier-YOLO: A real-time object detection method for constrained environments. IEEE Access, 8, 1935-1944. https://doi.org/10.1109/ACCESS.2019.2961959
17. Van Grinsven, M. J., van Ginneken, B., Hoyng, C. B., Theelen, T., & Sánchez, C. I. (2016). Fast convolutional neural network training using selective data sampling: Application to haemorrhage detection in colour fundus images. IEEE transactions on medical imaging, 35(5), 1273-1284. https://doi.org/10.1109/TMI.2016.2526689
18. Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2980-2988). https://doi.org/10.1109/ICCV.2017.324
19. Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2117-2125). https://doi.org/10.1109/CVPR.2017.106
20. Kauppi, T., Kalesnykiene, V., Kamarainen, J. K., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kälviäinen, H. & Pietilä, J. (2007, September). The Diabetic Retinopathy Database and Evaluation Protocol. In BMVC (Vol. 1, No. 1, p. 10). https://doi.org/10.5244/C.21.15
21. Seoud, L., Hurtut, T., Chelbi, J., Cheriet, F., & Langlois, J. P. (2015). Red lesion detection using dynamic shape features for diabetic retinopathy screening. IEEE transactions on medical imaging, 35(4), 1116-1126. https://doi.org/10.1109/TMI.2015.2509785

AUTHORS PROFILE

Goutam Kumar Ghorai is an Assistant Professor and former Head of the Department of Electrical Engineering, Ghanikhan Chowdhury Institute of Engineering and Technology, Malda, West Bengal. He has developed a Modular pattern course curriculum for the B.Tech program in Electrical Engineering. He obtained a B.Sc. (Hons.) degree in Physics from Vidyasagar University, as well as B.Tech and M.Tech degrees in Electrical Engineering from the University of Calcutta. His teaching experience spans approximately 17 years at the degree level, and he is currently pursuing a Ph.D. He has published various research papers at the National and international levels. His particular interests include machine learning, Deep learning, and Biomedical Engineering.

Swagata Kundu is a Senior Research Fellow in the Department of Electrical Engineering at the National Institute of Technology Durgapur, West Bengal, India. She obtained a B.Sc. (Hons.) degree in Physics, a B.Tech degree, and an M.Tech degree in Electrical Engineering from the University of Calcutta. She has published different research papers at the National and International levels. Her research interests include biomedical image analysis, artificial intelligence, machine learning, and deep learning.

Gautam Sarkar received the B.Tech. and M.Tech. Degrees in Electrical Engineering from the University of Calcutta, Kolkata, India, in 1999 and 2001, respectively and a PhD degree in Engineering from Jadavpur University, Kolkata, India, in 2015. He is currently a Professor in the Department of Electrical Engineering, Jadavpur University. His current research interests include electrical capacitance tomography, digital signal processing, internet-based instrumentation, the design of real-time systems, and Biomedical engineering. He has authored or co-authored almost 57 technical papers, including 19 international journal papers. Dr. Sarkar has guided several students to a PhD.

Ashis Kumar Dhara is presently serving as Assistant Professor in the Electrical Engineering Department of the National Institute of Technology Durgapur. He received his PhD from the Indian Institute of Technology, Kharagpur, India, in the field of medical imaging analysis. His research interests include biomedical image analysis, machine learning, and deep learning.