World Academy of Science, Engineering and Technology
International Journal of Medical and Health Sciences
Vol:14, No:8, 2020

Attention Based Fully Convolutional Neural Network for Simultaneous Detection and Segmentation of Optic Disc in Retinal Fundus Images

Sandip Sadhukhan, Arpita Sarkar, Debprasad Sinha, Goutam Kumar Ghorai, Gautam Sarkar, Ashis K. Dhara

Abstract—Accurate segmentation of the optic disc is very important for computer-aided diagnosis of several ocular diseases such as glaucoma, diabetic retinopathy, and hypertensive retinopathy. The paper presents an accurate and fast optic disc detection and segmentation method using an attention based fully convolutional network. The network is trained from scratch using the fundus images of extended MESSIDOR database and the trained model is used for segmentation of optic disc. The false positives are removed based on morphological operation and shape features. The result is evaluated using three-fold cross-validation on six public fundus image databases such as DIARETDB0, DIARETDB1, DRIVE, AV-INSPIRE, CHASE DB1 and MESSIDOR. The attention based fully convolutional network is robust and effective for detection and segmentation of optic disc in the images affected by diabetic retinopathy and it outperforms existing techniques.

Keywords—Ocular diseases, retinal fundus image, optic disc detection and segmentation, fully convolutional network, overlap measure.

I. Introduction

Visual impairment and blindness are a major problem in developing countries [1]. Diabetic retinopathy, hypertensive retinopathy, glaucoma are common causes of visual impairment and blindness [2]. Early diagnosis and appropriate referral for treatment of these diseases can prevent visual loss. Research is going on the development of a computer-aided diagnosis system for accurate identification of different parts and pathologies in retinal fundus image to assist ophthalmologists.

The Optic disc is the entry point of the major blood vessels in the retina [3] and considered a landmark in the retinal fundus image. Disc size and cup area are used for diagnosis of glaucoma [4], [5]. The centre of the optic disc is an important reference for detecting the macula and grading macular pathologies, such as diabetic maculopathy, macular edema, and macular ischemia [6]. Disc size is also an important parameter for determination of the region of interest, where the width of artery and vein needs to be computed for diagnosis of hypertensive retinopathy [7]. Along with the position of the optic disc, the vessel origin is another important feature for vasculature analysis [8]. Automated detection and segmentation of the optic disc is a challenging problem due to the variation in size, shape, colour, and the variation introduced by the field of view, inhomogeneous illumination and pathological abnormalities. Shape and brightness [9], [10], [11], convergence of blood vessels [12], [13] and orientation of blood vessels [13], [14] have been investigated for detection of optic disc. The assumption of the circular shape of the optic disc does not hold good, where the optic disc is partly present in the retinal image. Hoover *et al.* resolved this issue of poor contrast of optic disc by considering the convergence of blood vessel in the optic disc. Orientation of blood vessels have been used by Foracchia *et al.* [13] and Youssif *et al.* [14] to improve the result of optic disc localization. Vessel templates were also investigated by Osareh *et al.* [15] and Lowell *et al.* [3]. The Active shape model is used to extract the main blood vessels for localization of optic disc [16]. Brightness characteristics of the optic disc and vessel density in the optic disc region are utilized by Giachetti *et al.* [17]. Soares *et al.* [18] focused on the local appearance of the optic disc region and orientation of main blood vessels to determine the centre of the optic disc. Vessel directional and distribution of blood vessel are used by Zhang *et al.* [19] to improve the accuracy of optic disc localization. Roychowdhury *et al.* [20] used region-based features to classify the bright areas as optic disc and non-optic disc regions. The region with maximum vessel density and solidity is considered as the optic disc candidate.

The success of convolutional neural network in object segmentation [21], [22], [23], [24] has motivated us to investigate the performance of attention based fully convolutional network for optic disc detection and segmentation. The attention modules help to increase the performance of segmentation for natural images. To suppress false-positive regions and to highlight informative regions, we have developed attention based fully convolutional network. The spatial attention provides the location of the features, while the channel attention utilises the features that can be found in the available channels. Attention module helps to focus on relevant features to improve the segmentation results. Channel and spatial attention combine both the spatial context as well as the semantic information of optic disc. The purpose of the proposed framework is the development of attention

Sandip Sadhukhan, Arpita Sarkar, Debprasad Sinha and Gautam Sarkar are from Electrical Engineering Department, Jadavpur University, India-700032 (*e-mail: sansad@gmail.com).

Goutam Kumar Ghorai is from Electrical Engineering Department, Ghani Khan Choudhury Institute of Engineering and Technology, India-732102.

Ashis Kumar Dhara is from Electrical Engineering Department, National Institute of Technology Durgapur, India-713209.

based deep network for determining informative region of fundus image similar to optic disc and selection of informative images for fast and efficient transfer learning.

II. Methodology

A. Preprocessing

The images of different databases have different sizes. Therefore, the images are resized to 512×512 pixels for all databases. This process of resizing not only reduces the storage space of the database but also decreases the computational time without hampering the performance of the algorithm. The remaining operations are carried on these resized images. Red channel image is thresholded. Morphological opening, closing and erosion operations with square structuring element are used to create a mask of circular retinal fundus region-of-interest, which allows focusing only on the foreground of retinal images. The fundus image is cropped based on the bounding box of this mask. The segmentation algorithm is applied on the cropped image to reduce the processing time.

B. Segmentation Using Fully Convolutional Network

In case of convolution neural networks (CNN), a fully connected layer is added at the end of the network, whereas an FCN uses a convolutional layer only without adding the fully connected layers. In classification tasks, CNN works well because the output required is a label to which the class belongs. In contrast to this, semantic segmentation outputs an image where each pixel have been labeled. Thus it is required that the convolutions are performed to pixel-level and also that both spatial and semantic information is preserved. By removing the fully connected layer and replacing it with convolutional layers an upsampling procedure can be performed once the pixels in the image have been labeled.

Attention modules provide a way to improve features that are relevant for classification tasks. Different approaches have been explored in recent years. The attention modules can either work in the direction of hard attention or soft attention. Hard attention is non-differentiable and stochastic. Soft attention is probability-based [25] and deterministic, which makes it easier to use during training. In this work, soft attention is discussed as it is better suited for optimisation as it is compatible with back-propagation [26]. Early works on attention modules, where both local features and global features are extracted and the classification decision is based on the weighted local features. By using the weighted local features the network is guided towards learning only relevant features. Fan et al. [27] worked on re-designing CNN architectures to improve semantic segmentation tasks. They used channel attention and spatial attention modules [27].

The channels containing high-level features could be improved by channel attention whereas the spatial attention could enhance the spatial connectivity of features. The proposed fully convolutional neural network with attention module is shown in Fig. 1. The network consists of four blocks and three attention block are added to refine the feature map of last three blocks. In the convolution block

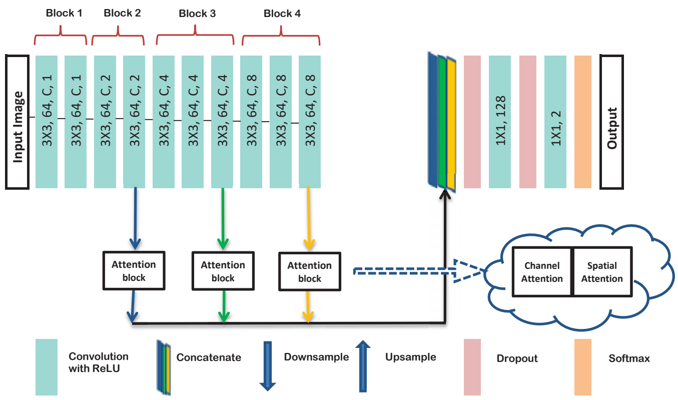

Fig. 1 Block diagram of Attention based fully convolutional neural network

attention module (CBAM) [28], a channel attention module is applied on a feature map. The result is further refined by a spatial attention module. The architecture focuses on both channel and spatial attention to learn about what and where about features. The features extracted from the last three blocks are refined by CBAM and final feature map is obtained by depthwise concatenation of refined features.

A stochastic gradient-based optimization *ADAM* [29] is applied to minimize the cross-entropy based cost function. ADAM utilizes the first and second moments of gradients for updating and correcting the moving average of the current gradients. The learning rate for ADAM optimizer is set to 0.0001, weights of background and foreground are initialized as 1:10, and training was performed up to 30 epochs. Over-fitting is reduced by using dropout [30].

III. Results and Discussions

Training has been performed in a Linux environment using an 8 GB GPU on a system with Core-i7 processor and 32 GB RAM. The network architecture is implemented in Python using the PyTorch library. The segmentation results are evaluated by considering 3-fold cross-validation.

The performance of optic disc detection is evaluated using Success Rate (SR) and it represents the percentage of retinal images in a dataset where the centroid of the optic disc is successfully localized within the boundary of the ground truth mask of the optic disc. The performance of optic disc segmentation is evaluated in terms of a region-based metric Overlap Measure (OM) and a contour-based metric Mean Absolute Distance (MAD) [20]. The OM represents the ratio of the intersecting area between the actual optic disc and segmented optic disc and it is defined as

$$\mathrm{OM} = \frac{N_{tp}}{N_{tp} + N_{fp} + N_{fn}}. \tag{1}$$

where N_{tp}, N_{fp} and N_{fn} are number of true positive, false positive and false negative pixels, respectively. MAD represents the mean of the shortest distances from the boundary of the actual optic disc to the boundary of the segmented optic disc. Let $\mathrm{P}_{i'}$, where i' = 1, . . ., m_C and $\mathrm{Q}_{j'}$

where, $j' = 1, \ldots, n_{\hat{C}}$ are the boundary points obtained on the segmented and actual optic disc boundaries respectively. The shortest distance from each boundary point of the segmented optic disc to the boundary points on the boundary points of ground truth mask of the optic disc is calculated using (2) and the shortest distance from each boundary point of the ground truth mask optic disc to the boundary points on the boundary points of the segmented optic disc is calculated using (3). Mean of the obtained shortest distances is calculated using (4).

$$m_{C,\hat{C}}(i') = \min_{j'} ||P_{i'} - Q_{j'}||_2 \tag{2}$$

$$m_{\hat{C},C}(j') = \min_{i'} ||Q_{j'} - P_{i'}||_2 \tag{3}$$

$$\text{MAD} = \frac{1}{2}\{\frac{1}{n_C}\sum_{i'}^{m_C} m_{C,\hat{C}}(i) + \frac{1}{n_{\hat{C}}}\sum_{j'}^{n_{\hat{C}}} m_{\hat{C},C}(j)\} \tag{4}$$

The proposed framework consists of an off-line training step for learning global features and transfer learning step for learning database-specific features. The network is trained from scratch using the fundus images of extended MESSIDOR database and the trained model is fine-tuned by a few images of a particular database to learn database-specific features. Multi-scale features extracted in off-line training are concatenated with database-specific features for improved segmentation. The false positives are removed based on the geometrical features of salient regions in initial segmentation. The result is evaluated on six public fundus image databases such as DIARETDB0, DIARETDB1, DRIVE, AV-INSPIRE, CHASE DB1 and MESSIDOR. The comparison of the proposed and competing techniques is provided in Table I. The attention based fully convolutional network is robust and effective for detection and segmentation of optic disc in the images affected by diabetic retinopathy and it outperforms existing techniques. The method is successful in optic disc localization and segmentation when tested on both dilated and non-dilated types of fundus images acquired from different medical centres. The performance of this algorithm does not degrade while handling images containing strong distractors like yellowish exudates which prove the effectiveness and robustness of the proposed process.

TABLE I
COMPARATIVE RESULT OF OPTIC DISC SEGMENTATION

Method	Author	OM	MAD	SR
DIARETDB1	**Proposed**	**0.92**	**1.98**	**100**
	Roychowdhury *et al.* [20]	0.80	4.82	100
	Morales *et al.* [31]	0.82	2.88	100
	Salazar *et al.* [32]	0.76	6.38	96.7
	Welfer *et al.* [33]	0.43	8.31	97.7
DIARETDB0	**Proposed**	**0.83**	**4.26**	**97.75**
	Roychowdhury *et al.* [20]	0.78	4.91	-
DRIVE	**Proposed**	**0.86**	**2.58**	**97.5**
	Roychowdhury *et al.* [20]	0.81	5.01	100
	Morales *et al.* [31]	0.72	5.85	100
	Salazar *et al.* [32]	0.71	6.68	97.5
	Welfer *et al.* [33]	0.42	5.74	100
MESSIDOR	**Proposed**	**0.92**	**1.95**	**99.92**
	Roychowdhury *et al.* [20]	0.84	3.9	100
	Marin *et al.* [34]	0.87	6.17	99.75
	Giachetti *et al.* [17]	0.88	-	99.83
	Aquino *et al.* [8]	0.86	-	98.83
	Yu *et al.* [35]	0.83	7.7	99.08
CHASE_DB1	**Proposed**	**0.81**	**7.89**	**100**
	Roychowdhury *et al.* [20]	0.81	5.19	-
AV_INSPIRE	**Proposed**	**0.83**	**4.63**	**100**

IV. CONCLUSION

Attention based fully convolutional network is developed for segmentation of optic disc. Attention modules are a reasonably new approach used in object detection and classification. Most of the reported architectures have been designed to work well for object detection. The placement of attention modules in a fully convolutional network could enhance the features of different layers. The method is successful in optic disc detection and segmentation when tested on both dilated and non-dilated types of fundus images acquired from different medical centres. The performance of this algorithm does not degrade while handling images containing strong distractors like yellowish exudates which prove the effectiveness and robustness of the proposed process.

REFERENCES

[1] R. R. Bourne, S. R. Flaxman, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg *et al.*, "Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 9, pp. e888–e897, 2017.

[2] D. Pascolini and S. P. Mariotti, "Global estimates of visual impairment: 2010," *British Journal of Ophthalmology*, vol. 96, no. 5, pp. 614–618, 2012.

[3] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy, "Optic nerve head segmentation," *IEEE Transactions on medical Imaging*, vol. 23, no. 2, pp. 256–264, 2004.

[4] G. D. Joshi, J. Sivaswamy, and S. Krishnadas, "Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment," *IEEE transactions on medical imaging*, vol. 30, no. 6, pp. 1192–1205, 2011.

[5] J. Cheng, J. Liu, Y. Xu, F. Yin, D. W. K. Wong, N.-M. Tan, D. Tao, C.-Y. Cheng, T. Aung, and T. Y. Wong, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Transactions on Medical Imaging*, vol. 32, no. 6, pp. 1019–1032, 2013.

[6] R. J. Qureshi, L. Kovacs, B. Harangi, B. Nagy, T. Peto, and A. Hajdu, "Combining algorithms for automatic detection of optic disc and macula in fundus images," *Computer Vision and Image Understanding*, vol. 116, no. 1, pp. 138–145, 2012.

[7] R. Estrada, C. Tomasi, S. C. Schmidler, and S. Farsiu, "Tree topology estimation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 8, pp. 1688–1701, 2015.

[8] A. Aquino, M. E. Gegúndez-Arias, and D. Marín, "Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques," *IEEE transactions on medical imaging*, vol. 29, no. 11, pp. 1860–1869, 2010.

[9] S. Barrett, E. Naess, and T. Molvik, "Employing the hough transform to locate the optic disk." *Biomedical sciences instrumentation*, vol. 37, pp. 81–86, 2001.

[10] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson, "Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images," *British Journal of Ophthalmology*, vol. 83, no. 8, pp. 902–910, 1999.

[11] M. Blanco, M. G. Penedo, N. Barreira, M. Penas, and M. J. Carreira, "Localization and extraction of the optic disc using the fuzzy circular hough transform," in *International Conference on Artificial Intelligence and Soft Computing*. Springer, 2006, pp. 712–721.

[12] A. Hoover and M. Goldbaum, "Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels," *IEEE transactions on medical imaging*, vol. 22, no. 8, pp. 951–958, 2003.

[13] M. Foracchia, E. Grisan, and A. Ruggeri, "Detection of optic disc in retinal images by means of a geometrical model of vessel structure," *IEEE transactions on medical imaging*, vol. 23, no. 10, pp. 1189–1195, 2004.

[14] A. A.-H. A.-R. Youssif, A. Z. Ghalwash, and A. A. S. A.-R. Ghoneim, "Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter," *IEEE Transactions on Medical imaging*, vol. 27, no. 1, pp. 11–18, 2008.

[15] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham, "Comparison of colour spaces for optic disc localisation in retinal images," in *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 1. IEEE, 2002, pp. 743–746.

[16] H. Li and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Transactions on biomedical engineering*, vol. 51, no. 2, pp. 246–254, 2004.

[17] A. Giachetti, L. Ballerini, and E. Trucco, "Accurate and reliable segmentation of the optic disc in digital fundus images," *Journal of Medical Imaging*, vol. 1, no. 2, pp. 024 001–024 001, 2014.

[18] I. Soares, M. Castelo-Branco, and A. M. Pinheiro, "Optic disc localization in retinal images based on cumulative sum fields," *IEEE journal of biomedical and health informatics*, vol. 20, no. 2, pp. 574–585, 2016.

[19] D. Zhang and Y. Zhao, "Novel accurate and fast optic disc detection in retinal images with vessel distribution and directional characteristics," *IEEE journal of biomedical and health informatics*, vol. 20, no. 1, pp. 333–342, 2016.

[20] S. Roychowdhury, D. D. Koozekanani, S. N. Kuchinka, and K. K. Parhi, "Optic disc boundary and vessel origin segmentation of fundus images," *IEEE journal of biomedical and health informatics*, vol. 20, no. 6, pp. 1562–1574, 2016.

[21] G. Papandreou, I. Kokkinos, and P.-A. Savalle, "Modeling local and global deformations in deep learning: Epitomic convolution, multiple instance learning, and sliding window detection," in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on*. IEEE, 2015, pp. 390–399.

[22] X. Chen, R. Mottaghi, X. Liu, S. Fidler, R. Urtasun, and A. Yuille, "Detect what you can: Detecting and representing objects using holistic models and body parts," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1971–1978.

[23] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3213–3223.

[24] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv preprint arXiv:1606.00915*, 2016.

[25] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[26] S. Jetley, N. A. Lord, N. Lee, and P. H. Torr, "Learn to pay attention," *arXiv preprint arXiv:1804.02391*, 2018.

[27] L. Fan, W.-C. Wang, F. Zha, and J. Yan, "Exploring new backbone and attention module for semantic segmentation in street scenes," *IEEE Access*, vol. 6, pp. 71 566–71 580, 2018.

[28] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 3–19.

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

[30] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *preprint arXiv:1207.0580*, 2012.

[31] S. Morales, V. Naranjo, J. Angulo, and M. Alcañiz, "Automatic detection of optic disc based on pca and mathematical morphology," *IEEE transactions on medical imaging*, vol. 32, no. 4, pp. 786–796, 2013.

[32] A. G. Salazar-Gonzalez, Y. Li, and X. Liu, "Optic disc segmentation by incorporating blood vessel compensation," in *Computational Intelligence In Medical Imaging (CIMI), 2011 IEEE Third International Workshop On*. IEEE, 2011, pp. 1–8.

[33] D. Welfer, J. Scharcanski, C. M. Kitamura, M. M. Dal Pizzol, L. W. Ludwig, and D. R. Marinho, "Segmentation of the optic disk in color eye fundus images using an adaptive morphological approach," *Computers in Biology and Medicine*, vol. 40, no. 2, pp. 124–137, 2010.

[34] D. Marin, M. E. Gegundez-Arias, A. Suero, and J. M. Bravo, "Obtaining optic disc center and pixel region by automatic thresholding methods on morphologically processed fundus images," *Computer methods and programs in biomedicine*, vol. 118, no. 2, pp. 173–185, 2015.

[35] H. Yu, E. S. Barriga, C. Agurto, S. Echegaray, M. S. Pattichis, W. Bauman, and P. Soliz, "Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets," *IEEE Transactions on information technology in biomedicine*, vol. 16, no. 4, pp. 644–657, 2012.

Sandip Sadhukhan has passed B.Tech and M.Tech from the University of Calcutta, India in Computer Science and Engineering and currently working as Assistant Professor. He is also currently pursuing PhD from Jadavpur University, India. Prior to this, he has worked in the IT industry for 18 years. His research interest is in the area of Machine Learning, Medical Image Analysis.

Arpita Sarkar is an Assistant Professor of School of Mobile Computing and Communication at Jadavpur University. She received her M.Tech degree from Calcutta University, India in 2010. She is presently pursuing her PhD degree in Jadavpur University. Her research interests include energy optimization techniques in Mobile Crowd Sensing, Networks, Security and privacy and Bio-medical Engineering.

Debprasad Sinha received the B.Tech. and M.Tech. degrees in optoelectronics from Calcutta University, Kolkata, India, in 2008 and 2010, respectively. He is currently pursuing PhD. degree in Engineering from Jadavpur University, Kolkata, India [in the Department of Electrical Engineering]. His current research interests include image processing and Bio-medical engineering.

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 B.Tech in Electrical Engineering course. He obtained B.Sc Physics (Hons) degree from Vidyasagar University, B.Tech and M.Tech in Electrical Engineering from the University of Calcutta. His teaching experience is about 14 years in degree level and presently pursuing PhD. He has published different research papers at the National and International level. His special interest is in Machine learning, Deep learning, and Bio-Medical Engineering. He has organized training programs as well as provided training on Solar Energy system design and installation for the village area.

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, design of real-time systems and Bio-medical engineering. He has authored or co-authored almost 57 technical papers, including 19 international journal papers. Dr. Sarkar has guided several students leading to PhD.

Ashis Kumar Dhara is presently serving as Assistant Professor in Department of Electrical Engineering of National Institute of Technology Durgapur. He received PhD. from Indian Institute of Technology Kharagpur, Kharagpur, India, on medical imaging analysis. His research interest includes image processing, machine learning and deep learning.