

A Novel Approach for Detecting Abnormality in Ejection Fraction Using Transthoracic Echocardiography with Deep Learning

<https://doi.org/10.3991/ijoe.v16i13.18483>

Prattay Guha Sarkar
RIMS, Jharkhand, India

Vishal Chandra (✉)
UMU, Jharkhand, India
vcvishalchandra@gmail.com

Abstract—Cardiovascular diseases (CVD) are the prime cause of mortality in people worldwide. Mortality in CVD has been strongly linked to Ejection Fraction (EF) in various studies [1]. Left ventricular ejection fraction (LVEF) is the central measure of left ventricular systolic function. LVEF is the fraction of chamber volume ejected in systole (stroke volume) in relation to the volume of the blood in the ventricle at the end of diastole (end-diastolic volume) [2]. Evaluation of left ventricular systolic function by left ventricle ejection fraction (EF) using Transthoracic echocardiography is usually a first line investigation. Determination of Ejection fraction (EF) is done most commonly by a semi-automatic process in which echocardiographer segments the left ventricle in both systolic and diastolic frames to generate systolic and diastolic chamber dimensions. The whole process is time consuming and highly dependent on operator experience causing a lot of inter-observer and intra-observer variations. Our goal is to develop algorithms so as to reduce the time consumed during whole process and make it more reliable and reproducible. We have used M-Mode of Left ventricle in PLAX view to measure chamber dimensions and calculate EF by Teich method. EF >50% has been categorized as normal ejection fraction. EF < 50% has been categorized as reduced ejection fraction and LV systolic dysfunction. In this research we have used fine-tuned ResNet 50 and trained it with 200 cases. We observed an accuracy of 98% and a F1 score of 77% for reduced EF (<50%) and 77% for normal EF (>50%). Although this is a small dataset, it shows that deep learning algorithms can be applied to medical imaging. ResNet50 is a preferred choice in terms of accuracy. This research will serve as a stepping stone for future research and will determine other cardiac matrices.

Keywords—Echocardiography, CNNs, Ejection fraction, Deep learning

1 Introduction

Approximately 17.7 million deaths in occurred in 2015 due to Cardiovascular diseases (CVDs)¹. Mortality in cardiovascular diseases strongly correlates with left

ventricle (LV) function. There are various techniques for the assessment of cardiac function like Trans thoracic echocardiography, Trans esophageal echocardiography, MRI, Nuclear scan etc. In day to day clinical practice, LV function assessment is done routinely by trans thoracic echocardiography by calculation of ejection fraction (EF). TTE is a non-invasive, low cost technique which makes it the first choice for the evaluation of ejection fraction by physicians and cardiologists. Various methods and techniques for the automatic segmentation of left ventricle are available. Some examples are active contours [2] method, supervised learning method [3] etc. Commercial solutions like Auto LV [4] are also available. These techniques have several limitations. They require good ultrasound images with low signal to noise ratio. Also, prior knowledge about the shape of Left ventricle is needed [5]. Deep learning approach has been shown to overcome this problem⁶. However, till date deep learning has been applied mostly to cardiac MRI and 3D echo images [6]. As both these modalities of cardiac imaging are costly and not freely available, it limits their use as first line investigation. On the other hand, 2D echocardiography is routinely used, cheap and easily available. In this research, we have applied deep learning algorithms to m-mode images of left ventricle obtained in PLAX view and classified the patients into two classes i.e. normal ejection fraction (Fig 1.) and reduced ejection fraction (Fig 2.).

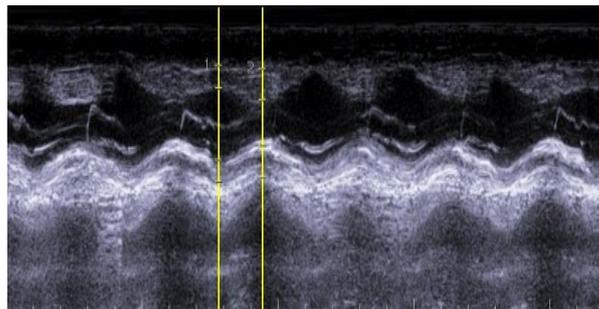


Fig. 1. Normal EF > 50%

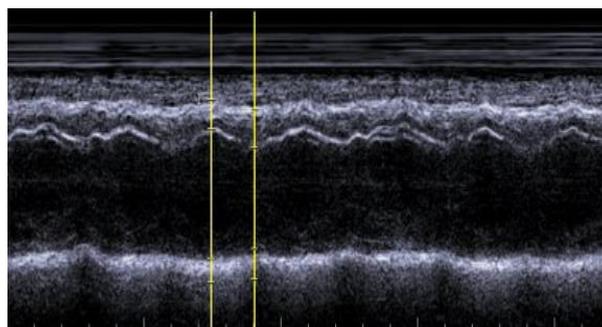


Fig. 2. Reduced EF < 50%

2 Literature Review

Transthoracic echocardiography is the cornerstone of non-invasive cardiovascular imaging using ultrasound-based technique to evaluate cardiac structure and function. It provides real time images of the heart which is evaluated by the operator to provide various cardiac diagnosis. Various deep learning models have been used to evaluate various aspects of this imaging technique. Automatic identification of endocardium, optical flow model for tracking of anterior mitral leaflet and assessment of Aortic regurgitation have been previously done by various authors. survey on applications of deep learning are easily available [7].

3 Deep Learning in Medical Imaging

Manual feature extraction nature of Conventional machine learning (ML) techniques has been used to process cardiac images in the past two decades. The main difficulty with ML is that it requires a manual feature extraction procedure before the Machine Learning pipeline. In the recent years, deep learning or deep neural networks have revolutionized this field by accelerating image processing, image classification, segmentation and object detection. Deep learning works on self and automatic feature extraction by the specific algorithms for particular task. Both ANNs and CNNs have had a huge contribution in medical imaging field. Various aspects like classification to segmentation and detection have been explored. CNNs have multiple layers of neurons that perform convolution operations which process the input image using a filter representation of the initial input images. After repeating this process several times, prediction layer gives class probabilities of input image. Compared to traditional image processing techniques such as image classification, CNN results are superior. Various studies on application of deep learning in medical imaging have been done and published. Authors [8] applied a deep learning model for classification of images of various 2D echocardiography views. The author achieved an accuracy of 92.1% in classification of eight different echo views. Authors [9] et al. described a web-service that performs image pre-processing of cardiac images. abnormalities in mitral leaflets are done by author [10][11], mitral regurgitation classification also one by authors [12]. In apical four chamber view tracking of mitral leaflets are also done by authors [13].

4 Methods of Estimation of Ejection Fraction

LVEF is calculated as the fraction of chamber volume ejected into the aorta in systole as the stroke volume in relation to the volume of the blood in the ventricle at the end of diastole (end-diastolic volume).

$$\text{LVEF} = [\text{SV}/\text{EDV}] \times 100$$

LVEF is evaluated either quantitatively using various standard methods or is estimated by eye balling. Estimating LVEF by eyeballing has several limitations with large inter observer and intra observer variations, hence it is not recommended. The

preference is to employ quantitative measures to assess LVEF to minimize variability and favour more precision and accuracy in the measurement. However, since quantitative methods are also dependent on the operator for segmentation of LV in systole and diastole, operator bias and variability are often inevitable.

Quantitative estimation of LV ejection fraction by two-dimensional echocardiography.

4.1 Modified simpson method (Biplane method of disks)

The American Society of Echocardiography recommends this method for measuring LVEF [14]. Operator traces the endocardial border in apical 4 chamber and 2 chamber views in both systole and diastole. The LV cavity into a predetermined number of disks (usually 20). Disk volumes are based on the tracings obtained from the study. Comparison of the disk volumes during systole and diastole generated LV end diastolic and end systolic volumes, which helps in calculating the ejection fraction. Since the entire tracing of LV cavity border is not done, some geometric assumptions need to be made which can be a source of some error in this method. Also, since the LV endocardial tracing is done by the operator, success of this method will depend on the accuracy of such delineation by the operator.

4.2 Teich method

This method is based on the assumption of a fixed geometric LV shape like a prolate ellipsoid. This method utilizes linear measurements. Single measurements of the LV cavity in the mid-ventricle in both end-diastole and end-systole is done in mid systole and mid diastole. Since the assumption of fixed LV geometric shape does not apply in a variety of cardiac pathologies, the calculated LVEF may not correlate with actual ejection fraction. Teich method for calculating LV volumes from LV linear dimensions is no longer recommended for clinical use.

4.3 Three-dimensional echocardiography

Because three-dimensional echo does not require geometric assumptions, it is felt to be the optimal way of measuring LVEF using echocardiography. LVEF derived from this modality would mostly require the data to be obtained over several heartbeats using special three-dimensional imaging probes. Unlike other M-mode and two-dimensional echocardiographic techniques, three-dimensional methods give a minimal explanation about the shape of LV cavity. When compared to other echocardiographic methods, three-dimensional modality is known to more accurate and far less variable because the entire LV cavity is detected.

5 Deep Learning in LV Segmentation

The application of deep learning in LV segmentation have been elaborated in several papers: Dong et al. [15] proposed a deep learning model and snake approach for LV segmentation in 3D electrocardiography. Luo et al. [16] proposed a deep learning model for the estimation of volumes using MRI. Carneiro et al. used deep learning model in 2D echocardiography for LV segmentation.

There were two major limitations in their approach: the first was to detect systolic and diastolic frame and feed it to model as input for estimation of ejection fraction and the second was LV segmentation. The obtained results were error prone. These approaches were not only used for estimation of the Ejection Fraction but also used for classification.

6 Methods

The proposed methodology is pictorially presented in the figure 3, it has two main parts: first dataset creation and second customized ReNet50 CNN model – each of them is explained in detail in the following subsections.

6.1 Data collection and pre-processing

Data was collected form Cosy Care hospital under informed consent. 400 m-mode samples were collected from distinct patients, 200 for $EF > 50$ and 200 $EF < 50$. All 400 samples annotated after measuring by using Biplane Simpson's approach. 60 samples were used for testing, results were validated by again using Simpson's formula. images were cropped to the region of interest.

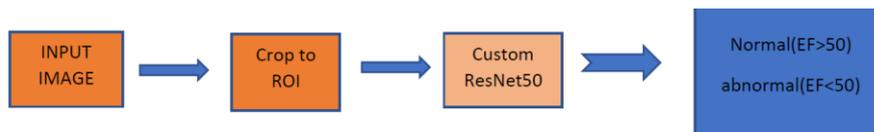


Fig. 3. Proposed pipeline for training and testing

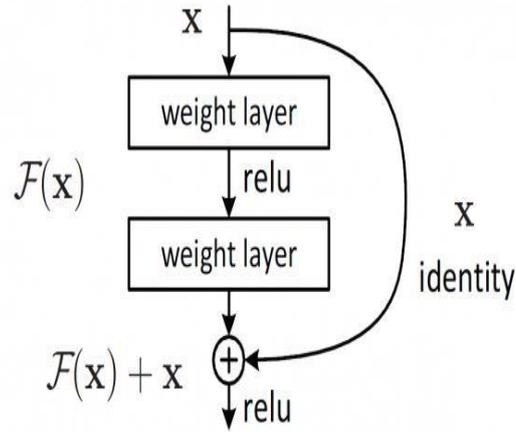


Fig. 4. Residual block

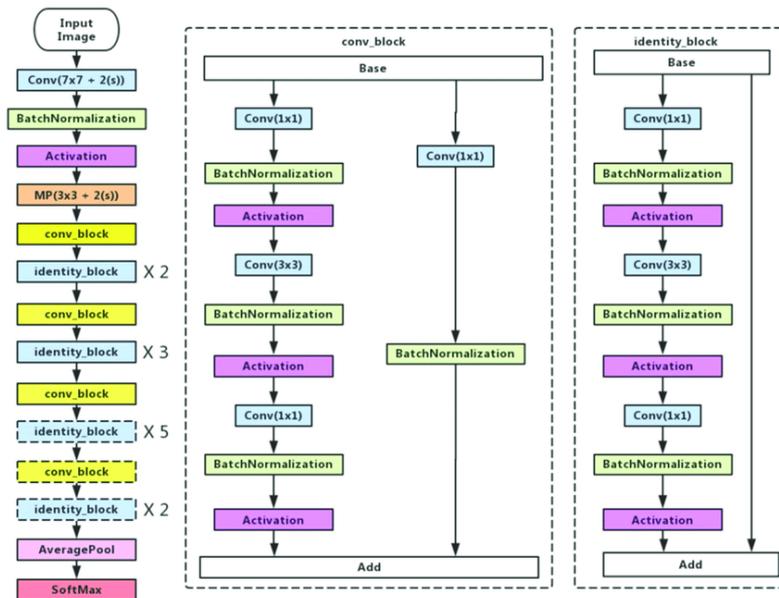


Fig. 5. ResNet50 architecture

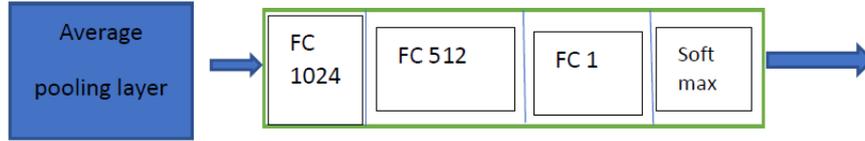


Fig. 6. Customized last layer

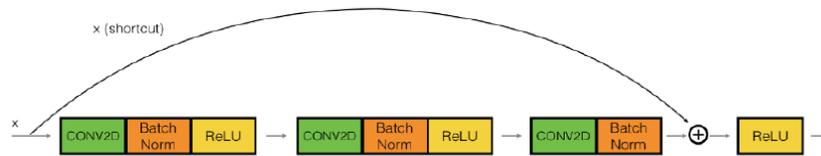


Fig. 7. Skip connection

Figure 4 is residual block of ResNet50. Figure 5 is the full architecture of classical ResNet50. We added two fully connected layer of unit 1024 and 512 before SoftMax classifier, ImageNet weight was not used.

6.2 Resnet50 and proposed resnet50

Resnet is the short form of Residual Networks is a classic deep learning model, backbone of for many computer vision applications [17], in 2015 ResNet50 won the ImageNet challenge. ResNet was the breakthrough in computer vision, allowed to train very deep model with 50 layers successfully. Before ResNet, there was a problem of vanishing gradients which made training of very deep neural networks was difficult. Skip connection was first introduced by ResNet. in figure 7 skip connection is illustrated. In skip connection, we add original input to output of convolution block.

Smaller representation of input mainly comes from bottom layers, which is fed to the middle layers. Majority of computation is done by middle layers and last feed to the softmax classifier. Since small asymmetric filters have fewer parameters [18], throughout the model Batch Normalization was used and for activation ReLU function is used. Batch normalization slows down the training of network by phenomenon of internal covariate shift. Activation function was achieved using ReLU. It is one of the variations of the rectified linear unit [19]. To avoid the phenomenon of overfitting, dropout and regularization [20] were used in our custom model. these modification makes different from classical ResNet50. Our proposed model trained over 40 epochs and 1000 steps per epoch, has 10^{-5} learning rate, dropout = 0.5 and l2 regularization 2.5×10^{-5} .

Inter observer and Intra observer variability for calculation of EF. Inter observer variability is a phenomenon whereby different operators generate different data on EF from the same image. Intra observer variability is a phenomenon whereby same operator

generates different data on EF from the same image at different points of time. This arises due to operator differences in segmentation of LV in systole and diastole. To test interobserver variability, we selected 30 medical practitioners who have sound knowledge of echocardiography and having experience more than 5 years. 40 images were taken for testing purpose in which each of the operators were asked to comment whether the ejection fraction was normal (>50%) or reduced (<50%).

7 Results and Discussion

F1 score of both classes was obtained, class1=97% and class2=97%. Results also support the use of balanced dataset. Classification report is given in table 1, and comparison with other is given in table 2. PR curve is shown in figure 8, figure 7 shows AUC, output of testing image is shown in figure 10, confusion metrics in figure 11.

Table 1 shows the inter-observer variability in predicting EF as normal or reduced correctly using 40 images for each of the 30 operators. The number of correct responses varied from 9 to 22 out of 40. Thus, low precision results were generated by human operators. In contrast, deep learning algorithm generated 38 correct responses during all the test runs. There was no variability and the result generated was highly precise. The mean correct response rate by human operators was 15.7 vs 38 for the deep learning algorithm. So, the results generated by the human operators was of low accuracy, compared to high accuracy results from the deep learning algorithms. On the basis of achievement of extremely high precision and accuracy by deep learning in classifying EF as normal or reduced, we can say that deep learning method can help in improving the utility and reliability of echocardiography for cardiac diagnosis. Our model successfully identifies 38 images out of 40, 2 images were misclassified. On repeat test runs, similar results were generated by our model without any variation.

Table 1. Inter Observer Variability

observer	O1	O2	O3	O4	O5	O6	O7	O8	O9	O10
Machine	38	38	38	38	38	38	38	38	38	38
Human	12	17	9	22	19	22	10	11	18	17
Difference	26	21	29	16	19	16	28	27	20	21

Intra observer variability was evaluated in 5 echocardiography operators, who were given 40 images set and were asked to report whether ejection fraction was normal or reduced. Reporting from each operator was done 5 times with the same 40 images set on different days.

The minimum observed intra observer variability for correct responses was 3 and maximum observed variability was 11 out of 40 images. Average intra observer variability for correct response was 8.4 out of 40 images. No intra observer variability was observed while applying the deep learning model which generated 38 out of 40 correct responses on each run.

Table 2. Interobserver Variability

Observer	O1	O2	O3	O4	O5	Variability
First	18	12	23	22	17	11
Second	19	15	20	23	20	8
Third	17	20	11	22	18	11
Fourth	22	21	19	19	22	3
Fifth	23	20	16	29	25	9
Average	19.8	17.6	17.8	23	20.4	
Actual	38	38	38	38	38	
difference	18.2	20.4	20.2	15	17.6	

Table 3. Classification report

	precision	recall	f1-score	support
0 EF>50	0.97	0.97	0.97	30
1 EF<50	0.97	0.97	0.97	30
accuracy			0.98	60
macro avg	0.97	0.97	0.97	60
weighted avg	0.97	0.97	0.97	60

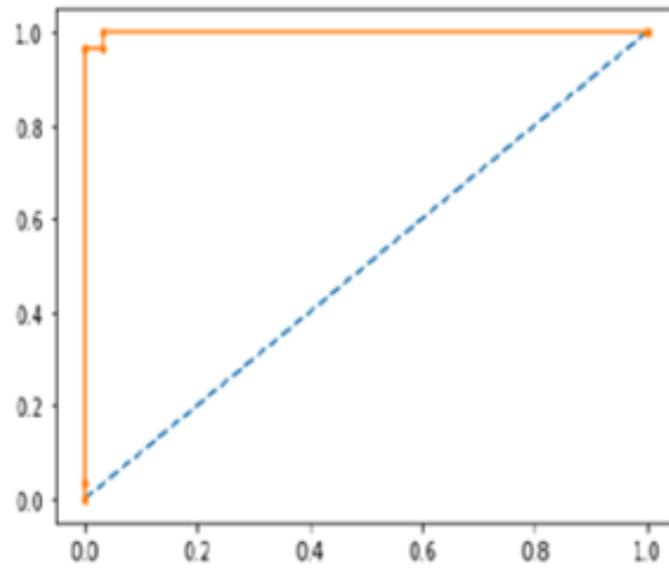


Fig. 8. AUC

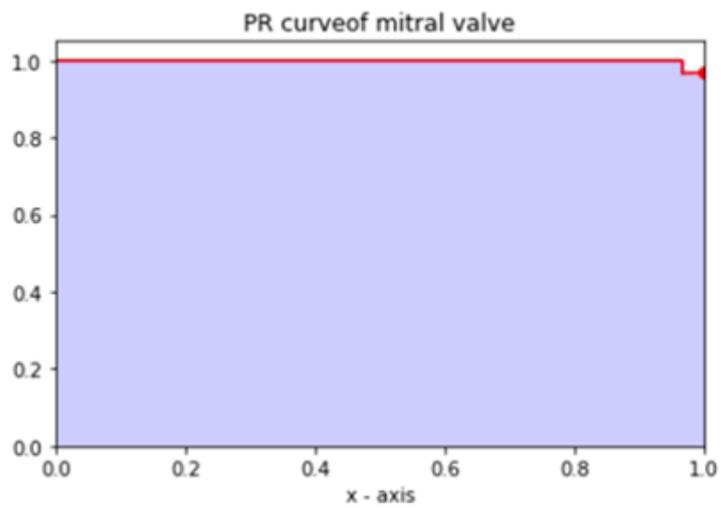


Fig. 9. PR curve

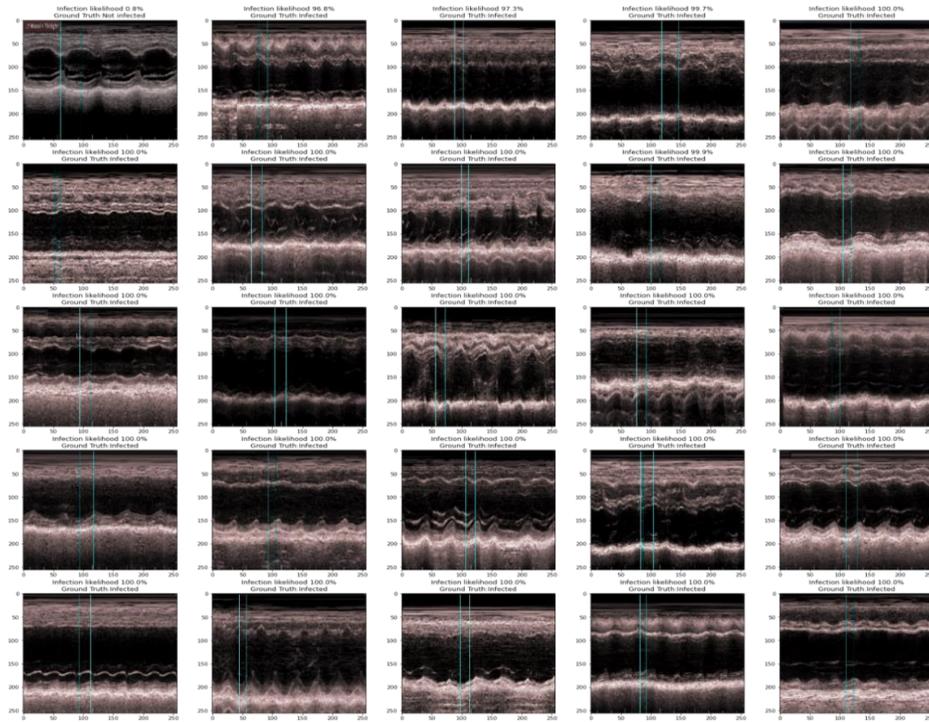


Fig. 10. Output of testing “infected” is depicted as $EF < 50$ and “Uninfected” is depicted as $EF > 50$

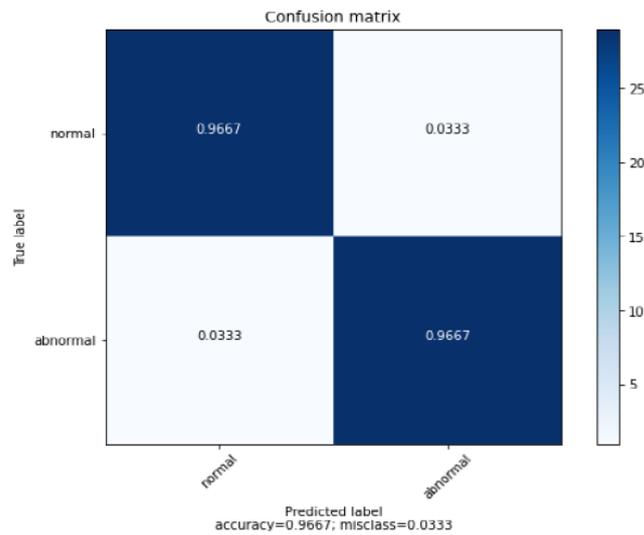


Fig. 11. Normalised confusion matrix

8 Conclusion

Left ventricular ejection fraction (LVEF) is the central measure of left ventricular systolic function. Determination of Ejection fraction (EF) is done most commonly by a semi-automatic process in which echocardiographer segments the left ventricle in both systolic and diastolic frames to generate systolic and diastolic chamber dimensions. The whole process is time consuming and highly dependent on operator experience causing a lot of inter-observer and intra-observer variations. This study shows that deep learning can help sonographers in their assessment of LV function. Deep learning algorithms are fast and provide extremely precise and accurate results. This may help to significantly reduce the inter observer and intra observer variation involved in these calculations.

9 References

- [1] Hajouli S, Ludhwani D. Heart Failure and Ejection Fraction. [Updated 2020 Jan 18]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK553115/>
- [2] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc IEEE Comput Soc Conf Computing Vis Pattern Recognit 2016; 2016-December:770–8. <https://doi.org/10.1109/cvpr.2016.90>
- [3] Kosaraju A, Goyal A, Grigороva Y, et al. Left Ventricular Ejection Fraction. [Updated 2020 May 5]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK459131/>
- [4] Lee, J. Y., Sunwoo, J. S., Kwon, K. Y., Roh, H., Ahn, M. Y., Lee, M. H., Park, B. W., Hyon, M. S., & Lee, K. B. (2018). Left Ventricular Ejection Fraction Predicts Poststroke Cardiovascular Events and Mortality in Patients without Atrial Fibrillation and Coronary Heart Disease. Korean circulation journal, 48(12), 1148–1156. <https://doi.org/10.4070/kcj.2018.0115>
- [5] O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Hernandez GV, Krpalkova L, et al. Deep Learning vs. Traditional Computer Vision. Adv Intell Syst Comput 2020; 943:128–44. https://doi.org/10.1007/978-3-030-17795-9_10
- [6] World Health Organization, “Cardiovascular diseases (CVDs),” 2017. [Online]. Available: <http://www.who.int/mediacentre/factsheets/fs317/en/>
- [7] Chandra V, Singh V, Sarkar PG. A survey on the role of deep learning in 2D transthoracic echocardiography. Int J Sci Technol Res 2020; 9:7060–5
- [8] E. d. S. Rebouças, A. M. Braga, R. M. Sarmiento, R. C. Marques, and P. P. Rebouças Filho, “Level set based on brain radiological densities for stroke segmentation in ct images,” in *2017 IEEE 30th International Symposium on Computer- Based Medical Systems (CBMS)*. IEEE, 2017, pp. 391–396. <https://doi.org/10.1109/cbms.2017.172>
- [9] R. J. Pagán, P. P. Parikh, P. J. Mergo, T. C. Gerber, R. Mankad, W. D. Freeman, and B. P. Shapiro, “Emerging role of cardiovascular ct and mri in the evaluation of stroke,” *American Journal of Roentgenology*, vol. 204, no. 2, pp. 269–280, 2015. <https://doi.org/10.2214/ajr.14.13051>
- [10] Chandra V, Singh V, Sarkar PG. Mitral Valve Abnormality Detection by fully end to end with Deep Neural Network. Int J Eng Adv Technol 2020; 9:1997–2004. <https://doi.org/10.35940/ijeat.c5372.029320> (UGC Approved).

- [11] Chandra V, Journal I, Chandra V, Sarkar PG, Singh V. Mitral Valve Abnormality Detection Using Deep Learning Approach in PLAX view in TTE 2020;8:1349–56 (scopus indexed). <https://doi.org/10.30534/ijeter/2020/67842020>
- [12] Chandra V, Sarkar PG, Singh V. Mitral Regurgitation Classification using MrNetDeep Learning Model. Int J Adv Trends Comput Sci Eng 2020; 9:1574 – 1582 (Scopus indexed). <https://doi.org/10.30534/ijatcse/2020/102922020>
- [13] Chandra V, Sarkar PG, Singh V. ScienceDirect ScienceDirect Mitral Valve Leaflet Tracking in Echocardiography using Yolo3 Mitral Valve Echocardiography using Vishal Leaflet Prattay Guha in Sarkar Custom Yolo3 Vishal. Procedia Comput Sci 2020; 171:820–8. <https://doi.org/10.1016/j.procs.2020.04.089>. (Scopus)
- [14] M. Cannesson, M. Tanabe, M. S. Suffoletto, D. M. McNamara, S. Madan, J. M. Lacomis, and J. Gorcsan, “A novel twodimensional echocardiographic image analysis system using artificial intelligence-learned pattern recognition for rapid automated ejection fraction,” *Journal of the American College of Cardiology*, vol. 49, no. 2, pp. 217–226, 2007. <https://doi.org/10.1016/j.jacc.2006.08.045>
- [15] A. Sarti, C. Corsi, E. Mazzini, and C. Lamberti, “Maximumlikelihood segmentation of ultrasound images with Rayleigh distribution,” *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 52, no. 6, pp. 947–960, 2005. <https://doi.org/10.1109/tuffc.2005.1504017>
- [16] J. G. Bosch, S. C. Mitchell, B. P. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. Reiber, “Automatic segmentation of echocardiographic sequences by active appearance motion models,” *IEEE transactions on medical imaging*, vol. 21, no. 11, pp. 1374–1383, 2002. <https://doi.org/10.1109/tmi.2002.806427>
- [17] A. Antony, A. Ramesh, A. Sojan, B. Mathews, and T. A. Varghese, “Skin Cancer Detection Using Artificial Neural Networking,” *Ijireeice*, vol. 4, no. 4, pp. 305–308, 2016.
- [18] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. S´anchez, “A Survey on Deep Learning in Medical Image Analysis,” no. 1995, 2017. [Online]. Available: <https://doi.org/10.1016/j.media.2017.07.005>
- [19] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P. M. Jodoin, and H. Larochelle, “Brain tumor segmentation with Deep Neural Networks,” *Medical Image Analysis*, vol. 35, pp. 18–31, 2017. [Online]. Available: <https://doi.org/10.1016/j.media.2016.05.004>
- [20] X. Gao, W. Li, M. Loomes, and L. Wang, “A fused deep learning architecture for viewpoint classification of echocardiography,” *Information Fusion*, vol. 36, pp. 103–113, 2017. <https://doi.org/10.1016/j.inffus.2016.11.007>

10 Authors

Prattay Guha Sarkar is from RIMS, Ranchi, Jharkhand, India

Vishal Chandra is from UMU, Ranchi, Jharkhand, India. vcvishalchandra@gmail.com

Article submitted 2020-09-12. Resubmitted 2020-10-21. Final acceptance 2020-10-28. Final version published as submitted by the authors.