Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 18s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Ecg Image Based Predictions for Heart Care Using Machine Learning Based Framework

Hariom Patidar[1], Anshul Shrotriya[2], Yogesh Payasi[3], Divya Chouhan[4]

[1]Assistant Professor, Department of Electronics and communication, Medi-Caps University, Indore, Madhya Pradesh, India, hariom.patidar@medicaps.ac.in

[2]Assistant Professor, Department of Electronics and communication, Medi-Caps University, Indore, Madhya Pradesh, India, anshul.shrotriya@medicaps.ac.in

[3]Assistant Professor, Department of Electronics and communication, Medi-Caps University, Indore, Madhya Pradesh, India, yogesh.payasi@medicaps.ac.in

[4]Assistant Professor, Department of Electronics and communication, Medi-Caps University, Indore, Madhya Pradesh, India, divya.chouhan@medicaps.ac.in

Cite this paper as: Hariom Patidar, Anshul Shrotriya, Yogesh Payasi, Divya Chouhan, (2025) Ecg Image Based Predictions for Heart Care Using Machine Learning Based Framework. *Journal of Neonatal Surgery*, 14 (18s), 334-339.

ABSTRACT

Early detection of left ventricular dysfunction (LVD), especially in asymptomatic individuals, is critical for timely intervention and improved cardiovascular outcomes. However, widespread access to echocardiography remains limited, particularly in low-resource settings. In this study, we developed and externally validated a deep learning (DL) model that predicts left ventricular ejection fraction (LVEF) from 12-lead ECG trace images, eliminating the need for raw signal data. A total of 1,19,281 ECG-echocardiogram pairs from 1,04,697 patients formed the training and test datasets, while 24,319 pairs from a distinct cohort were used for external validation. The ECG trace plots were processed via multi-otsu thresholding to extract the region of interest and standardized using Z-score normalization. The model architecture was based on DenseNet121, trained with class-weighted focal binary cross-entropy loss to address data imbalance. On internal test data, the model achieved a receiver operating characteristic area under curve (ROCAUC) of 0.92 and precision-recall AUC (PRAUC) of 0.78 in identifying LVEF < 50%. External validation yielded comparable performance with ROCAUC and PRAUC of 0.88 and 0.74, respectively. Notably, the algorithm demonstrated 97% sensitivity in detecting severe LVD (EF ≤ 35%) and maintained robust performance across age, sex, and paced ECG subgroups. With a diagnostic odds ratio of 31.7 on test data and a high negative predictive value (NPV ~0.94), the model ensures low false negative rates—critical for triaging in high-volume clinical settings. This study highlights the feasibility of using ECG image-based DL models for LVD screening, especially in resource-constrained environments. The ability to extract LVEF-related features from trace images offers practical scalability in primary and tertiary healthcare centers and introduces a new paradigm in accessible, AI-powered cardiac diagnostics.

Keywords: ECG, ANN, Machine learning, Prediction, Test data

1. INTRODUCTION

Usually going unreported, left ventricular dysfunction (LVD) might be silent or asymptomatic [1]. An improved result depends on the diagnosis of asymptomatic heart failure with a lower ejection fraction or those whose LV function is compromised due to undetected ischemia episodes. It is also known that asymptomatic coronary disease and LVD are related.

Fig. 1 and showed the Electro Cardio Graphs and region of ineterest to detect the hear6t related abnormal symptoms. A considerable number of patients with LVD would be identified with the use of echocardiography for routine screening for decreased LV function. However, there are limitations in terms of cost, availability, and skill when it comes to echocardiography [2, 3]. In most rural areas, the only way to screen for heart disease is through ECG monitoring. ECG is utilized for screening, and echocardiographic evaluation is not common, even in urban areas. However, ECGs appear normal to the human eye when there are no ischemia symptoms present.

Due to the limited sensitivity and negative predictive values, prior attempts to measure EF using ECG and electrocardiographic data such as voltage ratios or QRS duration have mostly failed. Additionally, attempts have been made to measure EF utilizing wearable technology, digital stethoscopes, hemodynamic parameters, and pulse wave morphology. 6. However, because of their higher cost and evaluation complexity, these models have serious drawbacks [4-7]. The potential of AI algorithms using raw ECG data—numerical values of several ECG parameters—for LVD screening has also been highlighted in recent papers. Acquiring stored ECG signal data is the main drawback, particularly in healthcare systems with limited resources where EMR use is relatively minimal [8-10]. Photographing an ECG is an easy task, and it would be economical and take little technical expertise to predict the left ventricular ejection fraction (LVEF) from an image. To address these issues and raise the precision and usability of LVEF estimate, we thus suggest creating and testing an AI program that uses ECG pictures to identify LVD [11-14].

2. METHODOLOGY

The study aims to examine the potential of ECG images using neural networks in predicting left ventricular dysfunction. To determine if a patient has EF ≥ 50% or <50%, a deep learning model was trained using the ECG trace images as input.

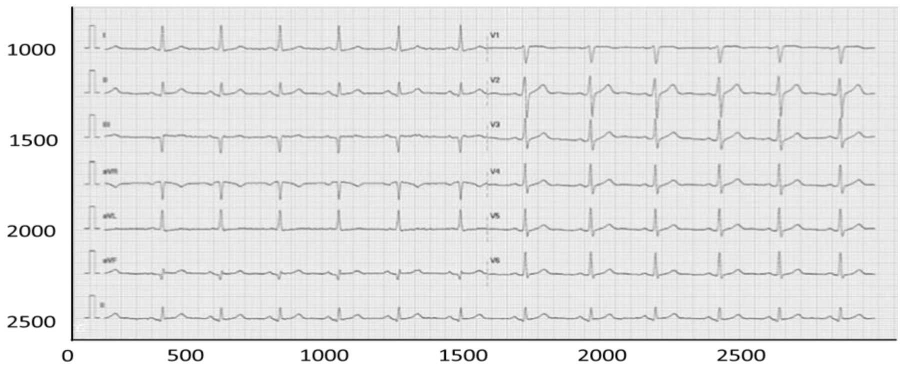

Fig. 1. Sample Greyscale image of ECG

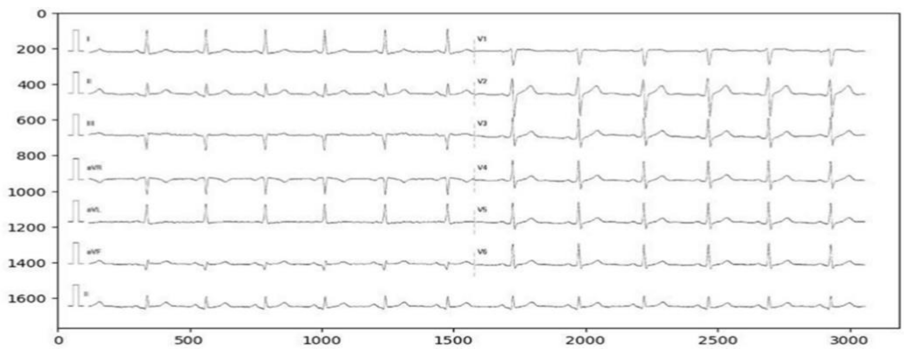

Fig. 2. Cropped region of interest (ROI) selecting the ECG trace plot

Certified echocardiography reports completed that same day provided information on the patient's EF. An external validation dataset and an internal test dataset were used to evaluate the model's performance after training.

Table 1 Data acquisition and compute specifications

Baseline Feature	Mean (+1 Standard Deviation)	Frequency N (%)	Missing Values N (%)
Hypertension	-	31,601 (30)	0 (0)
Diabetes mellitus	-	25,313 (25)	0 (0)
Total Cholesterol	172.21 (48.1)	-	43,725 (41.8)
Serum Creatinine	0.94 (0.56)	-	17,939 (17.1)
BMI	26 (4.76)	-	66,477 (64%)
Gender	-	Male - 75,099 (72) Female - 29,596 (28)	2 (~0)
Age	51.9 (13.5)	-	0 (0)
Age Groups, n (%) <40	-	19,314 (18.4)	-
40-49	-	23,360 (22.3)	-
50-59	-	28,985 (27.7)	0 (0)
60-69	-	23,939 (22.9)	-
70-79	-	8179 (7.8)	-
≥80	-	920 (0.9)	-

All patients who visited the hospital between December 2022 and March 2023 and had both their echocardiogram and ECG done on the same day were included in the data. The patients came from two centers' outpatient departments for cardiac sciences. Our cohort's ECG acquisition specifications included a 1000 Hz sample frequency, a paper speed of 25 mm/s, and an amplitude of 10 mm/mV. Patients with bundle branch block, atrial fibrillation, premature ventricular contractions, and early atrial contractions were among the ECG abnormalities we included. The BENEHEART R12 devices from Shenzhen Mindray Bio-medical Electronics were used to record ECG data. The information was then saved on the internal Picture Archiving and Communication Systems (PACS) servers as Digital Imaging and Communications in Medicine (DICOM) images. Our study's objective was to create a model that can forecast a patient's left ventricular performance regardless of any underlying cardiac conditions, including the frequently observed anomalies in the ECG. The results of patients who had echocardiograms were entered into the Electronic Medical Record system. Based on the modified Simpson's criteria, the biplane approach of disk summation was employed to compute the LV function. For the purposes of this investigation, the ejection fraction was used to quantify the LV function.

We identified and linked the ejection fractions of 1,30,464 archived 12-channel ECG DICOM pictures with estimates from 2D/color Doppler echocardiograms. The final cohort size was 1,19,281 ECG-echo pairs from 1,04,697 distinct patients after 11,183 ECG-echo pairs were eliminated due to faults and corruptions in the ECG picture data. Python 3.8, scikit-learn 1.2.0, Tensorflow 2.11.0, and scikit-image 0.20.0 were used for all data pre-processing and model construction. To train the model, one NVIDIA Tesla M60 GPU was used.

3. DATA PRE-PROCESSING

ECGs had a "YBR_FULL_422" photometric interpretation and were saved as "Secondary Capture Image" DICOMs. The images have three channels (YBR) and measured 2550 × 3299 pixels. The first channel was selected for additional examination. A sample ECG that is used for analysis is shown in Fig. 1. To isolate the portion of the image that only contains the ECG graph plot, multi-otsu thresholding was used to separate the several ECG regions. The region of interest was obtained by cropping the image (Fig. 2). The ECG trace plots were either 3 × 4 or 6 × 2 (with or without a rhythm strip). In order to preserve the aspect ratio and improve the input shape for effective model training, the cropped image was scaled to a tensor of shape (1,300,540,1). After testing a range of resolutions, these dimensions provided the optimal balance between model performance, training time, and processing cost. To normalize the input, Z-score scaling was used for each image and batch. Batch normalization was then used for training. After that, the LVEF was binarized, with class 0 being given to > 50% and class 1 to < 50%.

4. MODEL IMPLEMENTATION AND PERFORMANCE

The DenseNet121 architecture was used to construct the model. 300 × 540 x 1 (1 channel) input dimensions were used, and the weights were initialized at random (He normal distribution). A global average pooling was carried out following the last convolutional layer, producing a single output neuron with a sigmoid activation function. As a result, the model generates a probability between 0 and 1 after sampling image inputs. A focal sigmoid class-weighted binary cross-entropy loss ($\alpha = 0.795$; $\gamma = 0.2$)12 with an Adam optimizer was used to train the neural network. Given the lower frequency of patients with EF < 50%, the custom loss function was employed to reduce the impacts of class imbalance in the data. For the duration of the training, the minibatch size was 16. When the validation area under the precision-recall curve (PRAUC) did not improve for more than three consecutive epochs, the training was terminated.

4.1. External validation

An attempt was made to validate the model in a different healthcare facility situated 1800 kilometers away from the original model development site in order to determine the model's generalizability. The pre-processing processes for the ECG images and all other inclusion criteria were identical to those in the internal dataset.

4.2. Model performance and statistical analysis

The test data and external validation data were used to assess the model's performance. At the threshold maximizing Youden's Index, performance measures included the receiver-operating characteristic (ROCAUC), precision-recall curve, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and confusion matrices. 1000 iterations of bootstrapping were used to establish 95% confidence intervals. The performance of the model was also examined in subgroups according to pacemaker presence, gender, and age.

5. RESULTS AND DISCUSSION

The entire dataset has an average age of 52.27 ± 13.46 years. Seven.3% of the cohort's total population had an LVEF of less than 35%, and 20.3% had an LVEF of less than 50%. The ratio of males to females was 2.6. Table 1 has more information on the baseline characteristics. Using stratified sampling, the ECGs and associated ground truths were divided into three datasets: training (n = 1,07,147), validation (n = 2020), and test (n = 10,114). The train and validation sets were used to build the model, and the test data was used to evaluate its performance.

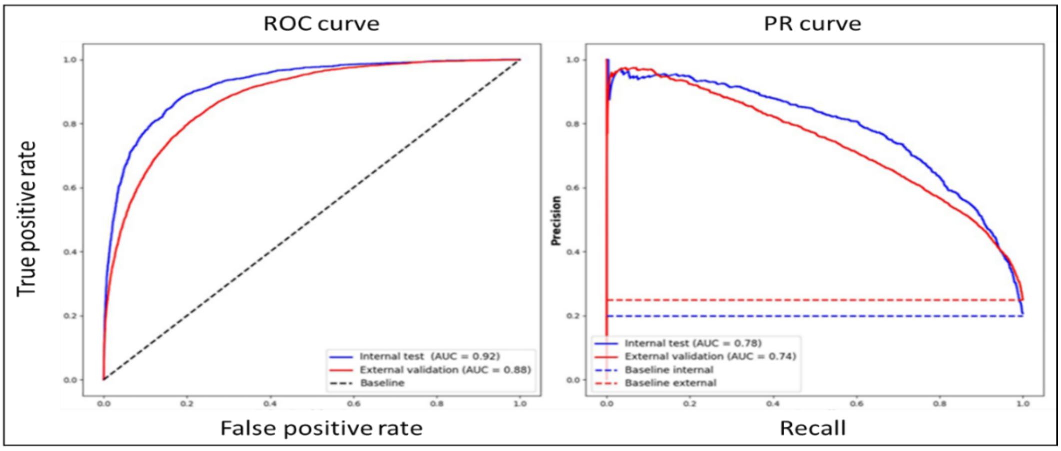

Fig. 3. Using the model's probability outputs, the receiver operating characteristic curve and precision-recall curve on internal and external validation data

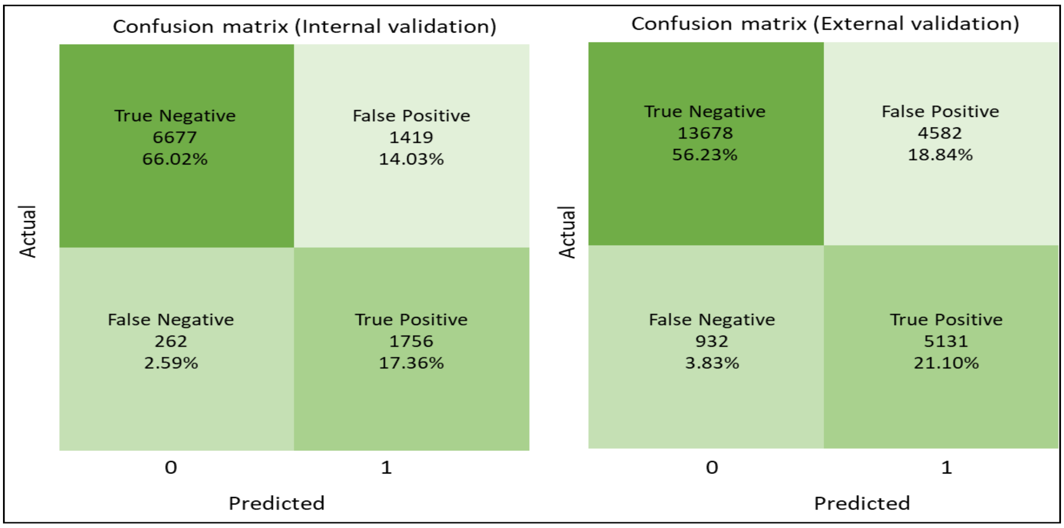

Fig. 4. Confusion matrices of internal and external data

For the test data, the model's ROCAUC and PRAUC to identify LVEF <50% were 0.92 and 0.78, respectively (fig. 3). You-den's score was determined to be at its highest using the ROCAUC at 0.18; if it was higher, the probability output was categorized as having an LVEF <50%. This threshold was used to create the confusion matrices for the internal and external validation data (fig. 4). For the test data, we found that an ECG showing LV systolic dysfunction was linked to a greater than 31-fold increase in the likelihood of having LV systolic dysfunction (diagnostic odds ratio: 31.7, 95% CI: 27.46–36.38). Furthermore, when the EF was less than 35%, the model accurately identified 97% of the samples. The model's performance was similar for all sexes and ages. It was found that the mean LVEF decreased and the fraction of LVD increased for each decile when the estimated probabilities on the test data. Table 2 presents the model's performance on the external validation data.

Table 2 External data validation and the performance of model

SUBSET	No. of Samples (n)	SENSITIVITY	SPECIFICITY	PPV	NPV	ROCAUC	PRAUC	DOR#
All	24319	0.85	0.75	0.53	0.94	0.88	0.74	16.44
Male	17185	0.84	0.75	0.59	0.92	0.88	0.78	16.12
Female	7131	0.88	0.74	0.33	0.98	0.89	0.56	20.59
≥65 Y	6861	0.88	0.6	0.5	0.92	0.84	0.73	11.49
<65 Y	17457	0.83	0.8	0.55	0.94	0.9	0.76	19.04
Paced ECGs	2202	0.88	0.52	0.45	0.91	0.78	0.61	7.81
No Paced ECGS	22117	0.84	0.77	0.54	0.94	0.89	0.76	17.96

External validation revealed that the model performed robustly and comparably. A total of 24,323 ECG-echo pairs were chosen from a different population between January 2022 and March 2024. About 25% of people had LVD (below than 50% EF), while 7.5% of them had severe LVD (EF ≤ 35%). This sample also demonstrated a better PRAUC (0.76) when compared to the held-out test data, where the prevalence of patients with pacemakers was 0.6%. The algorithm properly recognized 96% of samples with ≤35% EF, regardless of whether paced ECGs were included or not.

We have created and externally verified a deep learning algorithm that may be applied to ECG image screening for LVD. Both the external validation dataset and the internal test data show reliable and consistent performance from our model. Furthermore, our model showed similar accuracy across several demographic groups and performed well regardless of age and gender.

Although it has been reported, the correlation between EF and ECG using various ML (Machine Learning) techniques has not been investigated in the Indian context. Our model has the added benefit of having been evaluated against external data, and it produces findings that are equivalent to those of other models. The model exhibits resilient and constant high sensitivity and NPV despite the variable prevalence of dysfunction (between the external validation data and the internal test data). Both age and sex showed similar performance in the test data.

To collect and input into the model, complex data pipelines are needed for the numerical computation of raw numerical measures derived from ECG tracings. Obtaining ECG images is far easier than obtaining ECG signal data, particularly in the Indian population. Since ECG trace images are readily available at the primary, secondary, and tertiary care levels, they were utilized as the model's input.

Although it has been reported, the correlation between EF and ECG using various ML (Machine Learning) techniques has not been investigated in the Indian context. Our model has the further benefit of having been evaluated against external data, and it produces findings that are comparable to those of previous models. The model exhibits strong and constant high sensitivity and NPV despite the variable prevalence of dysfunction (between the internal test data and external validation data). Both age and sex showed similar performance in the test data. To collect and input into the model, complex data pipelines are needed for the numerical computation of raw numerical measures derived from ECG tracings. Obtaining ECG images is far easier than obtaining ECG signal data, particularly in the Indian population. Since ECG trace images are readily available at the primary, secondary, and tertiary care levels, they were utilized as the model's input. It is essential to diagnose LVD with ECG, especially in situations with limited resources when ECG is very accessible and reasonably priced. ECG is frequently the only assessment method available to patients in urban areas undergoing non-cardiac procedures. Nevertheless, ECG has a limited sensitivity for LVD detection, which frequently results in postoperative cardiac problems. By identifying patients who are more likely to experience perioperative cardiac problems, the use of ML-ECG models to diagnose LVD improves patient safety. Preventive management is made easier, anesthesia planning is optimized, and patient risk classification and prognostication are demonstrated.

In this study, we tried to maintain a respectable PPV while optimizing the model's sensitivity-specificity trade-off. With a PPV of about 0.55, approximately 1.8 echocardiograms would be required to confirm 1 case of low EF if the model suggests that the ECG is dysfunctional. It is demonstrated that almost all cases of severe LVD (≤35% EF) are captured by the model. In situations when there is an excessive demand for echocardiography services, this can serve as an additional triaging mechanism and allow for the proper diagnosis of severe LVD cases, which may require an urgent referral. Our model's high NPV guarantees few false negatives. This study acts as a proof of concept for further research in which the ejection fraction can be determined by taking ECG pictures using a mobile device. Unlike some earlier studies, this method does not require raw ECG data as model input. This approach increases the model's potential application and makes it more practical for a range of clinical settings by reducing the complexity of data needs.

6. CONCLUSION

This study demonstrates the successful implementation of an AI-driven model that accurately predicts LVEF from ECG images, providing a non-invasive and accessible solution to screen for LVD in both urban and resource-limited clinical settings. The model, developed using over 1.19 lakh ECG-echo pairs and externally validated on a separate 24k sample, exhibited high sensitivity, specificity, and diagnostic power, including a 97% sensitivity for severe dysfunction (EF ≤ 35%). Its comparable performance across age groups, sexes, and ECG abnormalities reinforces its generalizability. Unlike traditional approaches that rely on raw ECG signals, our model simplifies deployment by utilizing image-based data readily available at all healthcare levels.

By reducing reliance on echocardiography for initial LVD screening, this approach enhances early diagnosis and risk stratification, particularly in settings where echocardiographic resources are scarce. With a positive predictive value of ~0.55, the model ensures clinical utility while maintaining high NPV to avoid missed diagnoses. The results establish a proof-of-concept for broader integration of AI-based ECG interpretation into mobile and outpatient care workflows, facilitating preventive cardiology. Future work can explore real-time LVEF estimation using smartphone-based ECG capture to enhance outreach and telemedicine-based cardiac care.

REFERENCES

[1] Savarese G, Stolfo D, Sinagra G, Lund LH. Heart failure with mid-range or mildly reduced ejection fraction. Nat Rev Cardiol. 2022; 19:100–116.

[2] Haas P, Santos ABS, Can˜on-Montan˜ez W, et al. Associations between coronary Artery calcification and left ventricular global longitudinal strain and diastolic parameters: the ELSA-Brasil study. Am J Cardiol. 2023; 204:215–222.

[3] Panicker GK, Narula DD, Albert CM, et al. Validation of electrocardiographic criteria for identifying left ventricular dysfunction in patients with previous myocardial infarction. Ann Noninvasive Electrocardiol. 2021;26, e12812.

[4] Liu S-H, Yang Z-K, Pan K-L, Zhu X, Chen W. Estimation of left ventricular ejection fraction using cardiovascular hemodynamic parameters and pulse morphological characteristics with machine learning algorithms. Nutrients. 2022; 14:4051.

[5] Attia ZI, Dugan J, Rideout A, et al. Automated detection of low ejection fraction from a one-lead electrocardiogram: application of an AI algorithm to an electrocardiogram-enabled Digital Stethoscope. Eur Heart J Digit Health. 2022;3: 373–379.

[6] Khunte A, Sangha V, Oikonomou EK, et al. Detection of left ventricular systolic dysfunction from single-lead electrocardiography adapted for portable and wearable devices. NPJ Digit Med. 2023; 6:124.

[7] Kagiyama N, Piccirilli M, Yanamala N, et al. Machine learning assessment of left ventricular diastolic function based on electrocardiographic features. J Am Coll Cardiol. 2020; 76:930–941.

[8] Potter EL, Rodrigues CHM, Ascher DB, Abhayaratna WP, Sengupta PP, Marwick TH. Machine learning of ECG waveforms to improve selection for testing for asymptomatic left ventricular dysfunction. JACC Cardiovasc Imaging. 2021;14: 1904–1915.

[9] Sangha V, Nargesi AA, Dhingra LS, et al. Detection of left ventricular systolic dysfunction from electrocardiographic images. Circulation. 2023.

[10] Liao P-S, Chen T-S, Chung P. A fast algorithm for multilevel thresholding. J Inf Sci Eng. 2001;17.

[11] Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks [Internet]. arXiv; 2018 [cited 2024 Jun 17]. Available from: http://arxiv.org/abs/1608.06993.

[12] Lin T-Y, Goyal P, Girshick R, He K, Doll´ar P. Focal Loss for Dense Object Detection [Internet]. arXiv; 2018 [cited 2023 Jul 18]. Available from: http://arxiv.org/abs/1 708.02002.

[13] Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med. 2019; 25:70–74.

[14] Cho J, Lee B, Kwon J-M, et al. Artificial intelligence algorithm for screening heart failure with reduced ejection fraction using electrocardiography. ASAIO J. 2021;67: 314.