








## Employing Generative Networks for Synthetic Phonocardiogram and Electrocardiogram Signal Creation: A Privacy-Ensured Approach to Data Augmentation in Heart Diagnostics

Swarajya Madhuri Rayavarapu<sup>1\*</sup> , Tammineni Shanmukha Prasanthi<sup>1</sup> , Gottapu Santosh Kumar<sup>2</sup> , Gottapu Sasibhushana Rao<sup>1</sup> , Aruna Singham<sup>1</sup> 

<sup>1</sup> Department of Electronics and Communication Engineering, Andhra University, Visakhapatnam 530003, India

<sup>2</sup> Department of Civil Engineering, Gayatri Vidya Parishad College of Engineering, Visakhapatnam 530048, India

Corresponding Author Email: [madhurirayavarapu.rs@andhrauniversity.edu.in](mailto:madhurirayavarapu.rs@andhrauniversity.edu.in)

<https://doi.org/10.18280/isi.280408>

### ABSTRACT

**Received:** 30 April 2023

**Accepted:** 30 July 2023

**Accepted:** 16 August 2023

**Available online:** 31 August 2023

#### Keywords:

*auscultation, data augmentation, electrocardiogram, generative networks, phonocardiogram, synthetic data generation*

The diagnosis of various cardiac conditions necessitates meticulous analysis of Phonocardiogram (PCG) and Electrocardiogram (ECG) signals. In light of this, artificial intelligence and machine learning, coupled with computer-assisted diagnostic techniques, have been progressively integrated into modern healthcare systems, facilitating clinicians in making crucial diagnostic decisions. However, the effectiveness of these deep learning applications hinges on the availability of extensive training data, which exacerbates the risk of privacy violations. In response to this dilemma, research into methodologies for synthetic patient data generation has witnessed a surge. It has been observed that most attempts to generate synthetic ECG and PCG signals focus on modeling the statistical distributions of the available real training data, a process known as Data Augmentation. Among the various data augmentation techniques, Generative Adversarial Networks (GANs) have gained significant traction in recent years. This paper conducts an in-depth exploration and evaluation of GANs, specifically Deep Convolutional GANs and Conditional GANs, for the generation of synthetic ECG and PCG signals.

## 1. INTRODUCTION

Cardiac complications contribute significantly to global mortality rates [1]. Amidst the rise of wearable technologies and artificial intelligence, the precision and automation of cardiovascular detection have become increasingly pertinent in medical practice. Arrhythmias, abnormalities in heart rhythms, can range from minor inconveniences to life-threatening emergencies. Arrhythmias develop when the normal electrical impulses of the heart are disrupted, leading to heartbeats that are exceptionally slow, rapid, or irregular. Utilization of Electrocardiogram (ECG) and Phonocardiogram (PCG) signals can facilitate effective prediction and diagnosis of heart disease [2, 3]. Artificial auscultation, while providing a cost-effective and time-efficient tool for cardiac diagnostics, demands considerable training for clinicians.

The electrocardiogram serves as a critical diagnostic tool for arrhythmias, tracking the heart's electrical activity. Conversely, the PCG records heart sounds. Figure 1 presents a visual representation of the fundamental components of an ECG waveform that recur with each rhythmic contraction and relaxation of the heart [2]. The P, T waves, and QRS complex, including the R peak, signify the three primary constituents of a standard electrocardiogram waveform that encapsulate an entire cardiac cycle.

By employing a phonocardiograph, clinicians can trace the sounds and murmurs generated by the heart, including those induced by the closure of the atrioventricular, pulmonary, and aortic valves, as illustrated in Figure 2. Sound S1 is produced during atrioventricular valve closure; sound S2 occurs during semilunar valve closure; the interval between S1 and S2 is

stystole; and the interval between S2 and S1 of the following cycle is diastole [3].

Due to the complexity of the PCG signal waveform in comparison to the ECG, and the propensity for significant disturbances during collection, PCG has been less extensively explored [4]. Visual evaluation of an ECG for cardiac irregularities may not yield reliable results. As deep learning methodologies gained traction, automatic processing of ECG and PCG data, along with other medical and healthcare-related sectors, began to incorporate deep network topologies.

The advent of artificial intelligence, particularly for medical applications, has seen significant progress in recent years [5]. Incorporation of attention mechanisms into deep learning models has greatly benefitted both image and sequence data [6]. Despite the burgeoning popularity of deep learning algorithms for e-health applications, the identification of key factors that enhance prediction or diagnostic accuracy still necessitates large datasets. Deep learning algorithms strive to emulate new synthetic data by first identifying patterns that emerge from suitable training data. This process, known as Data Augmentation, allows for the expansion of a training set through synthetic generation of additional samples. When crucial modalities are unavailable for multimodal medical image segmentation, data generation is employed as a substitute [7]. Consequently, the model is taught to generalize higher quality images while concurrently circumventing overfitting. The Generative Adversarial Network (GAN), a popular form of generative modeling, is used to create both time series data and visuals [8]. GANs find application in healthcare [5, 9], financial technologies [10, 11], image inpainting [12, 13], fashion technology [14, 15], among others.

This study explores the application of GANs in synthesizing ECG and PCG signals de novo. Section III delves into the architecture of GAN and Deep Convolutional GAN (DCGAN);

Section IV details evaluation measures; Section V presents the results and the datasets used for simulation. A brief review of relevant literature is provided in Section II.

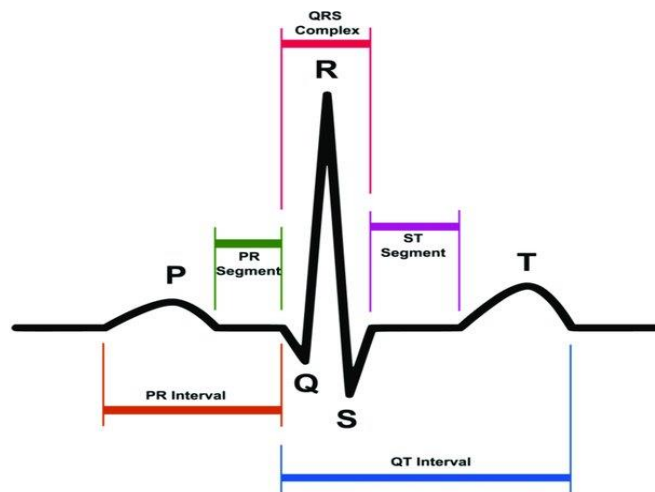


Figure 1. ECG signal and its characteristic waveforms [2]

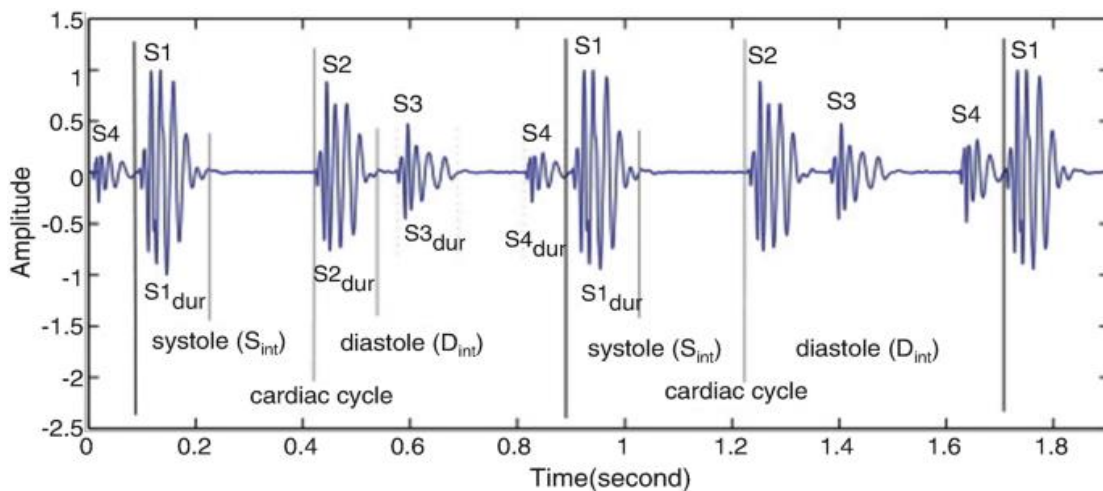


Figure 2. PCG signal and its recordings [3]

## 2. LITERATURE REVIEW

Contemporary methodologies for the classification of PCG and ECG signals primarily deploy traditional supervised machine learning approaches. A fusion of Mel-frequency cepstral coefficient (MFCC) features and dynamic time wrapping (DTW) was paired with a linear Support Vector Machine (SVM) to accomplish the classification of PCG signals [16]. The SVM method has also been extensively employed by researchers for the classification of ECG signals. Rizal et al. in used Short-time Fourier Transform for classification of PCG signals [17].

However, these methods have faced challenges in adequately scaling to incorporate ECGs from an expansive variety of individuals. Notwithstanding the time and effort required, each patient's ECG signal exhibits unique dynamics and morphology, significantly influenced by the patient's current health condition. Fully connected networks have demonstrated superior performance on recognized clinical benchmarks when compared to the state-of-the-art [17]. The simultaneous acquisition of ECG and PCG data, as

implemented by Martins et al., facilitated enhanced diagnosis of heart disease [17]. Zhang et al. [18] proposed the detection of coronary artery disease using multi-modal features. Multimodal feature models were suggested based on ECG and PCG signals, followed by a comparison of the relative accuracy of different modes in data classification.

Despite the promise of numerous deep learning methods for ECG or PCG classification and detection, their efficacy is predicated on the availability of ample training instances. The infrequency of potentially lethal arrhythmias results in a dearth of training data for deep learning algorithms, necessitating the use of medical image augmentation. Simple data augmentation techniques such as snipping, flipping, and adding noise have been attempted.

However, the management of complex data like medical imaging challenges these elementary methods. The variational autoencoder (VAE), an alternative deep model, has been overlooked in favour of more popular approaches [19]. Nevertheless, a significant drawback is that their output images are often blurry and indistinct.

A recently introduced model in the field of data

augmentation is the Diffusion Model, based on the diffusion mechanism [20]. While these models are still nascent in the field of medicine, they are hampered by significant computational costs and prolonged sample times. Figure 3 illustrates various data augmentation methods for biomedical images.

The extensive heartbeat categorization challenge has resulted in a scarcity of deep ECG and PCG models suitable

for clinical application. In the case of phonocardiograms, the classification of individual types of heart murmurs has seen no significant breakthroughs due to the shortage of labelled PCG data [21].

To mitigate this challenge, this study proposes the training of deep learning models with simulated ECG and PCG signals, representing a range of arrhythmias, using generative modelling.

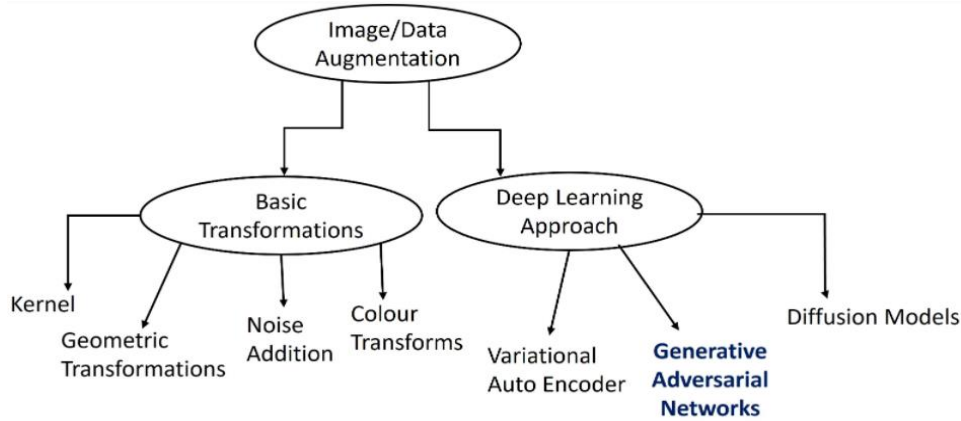


Figure 3. Various approaches for synthetic data generation/data augmentation

### 3. ARCHITECTURE

For the purpose of generating artificial ECG or PCG signals, we implemented and compared different architectures of GAN, DC GAN and conditional GAN.

#### 3.1 Generative adversarial networks

Generative adversarial networks were first proposed in 2014 by a team of academics led by Goodfellow et al. [8]. The field of “generative models” includes GANs. The GANs come under semi-supervised deep learning model. In contrast to other methods, GANs use a game-theoretic concept to train a synthetic image model.

As such, GANs are founded on the idea of zero-sumation games, in which each participant actively seeks to maximize his or her personal benefit at the detriment of all others. The GAN’s output is the collaborative effort of the Discriminator and Generator networks. The Discriminator’s goal is to master the ability to recognize and reject fake distributions generated by the Generator, whereas the Generator’s goal is to master the art of fooling the Discriminator with genuine-looking fakes. GAN is a deep learning technique that has risen in popularity in recent years due to its widespread use in fields such as image identification, video creation, anomaly detection, and security applications such as steganography.

The GAN is a well-liked deep generative model because of its two core components, the generator and the discriminator. The generator (G) takes as input a latent vector of N dimensions (z) with a Gaussian distribution and outputs a vector of the same shape. After been trained on instances of both sorts of data x, the discriminator’s output (D) is the likelihood of correctly determining whether the created data output is fake or real. The discriminator and the generator engage in a zero-sum game that ultimately results in a convergence condition. The objective function of a GAN is expressed in terms of min-max optimization, as indicated in Eq. (1).

$$\min_G \max_D V(D, G) = E_z \log(1 - D(G(z))) + E_x \log(D(x)) \quad (1)$$

D(x) represents the probability that D is applied to the real data x, while D(G(z)) is the probability that D is applied to the generated data G(z). E stands for the expectation, and G(z) is the data generated by the generator G. To improve the D, we want to make D(G(z)) to zero, and to improve the G, we want to make it to 1. The architecture of the GAN is shown in the Figure 4.

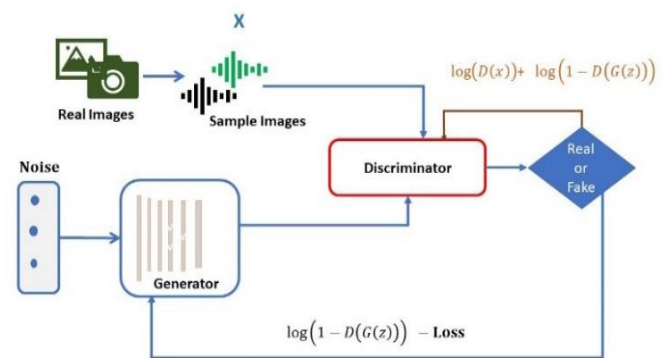


Figure 4. Architecture of GAN

The generator strives to lessen the possibility that the discriminator would incorrectly predict its output as fake, and both it and the discriminator want to boost the probability of accurate classification of actual and fake data. The discriminator gives each set of false and authentic data a probability of 0.5 under ideal conditions because their distributions are statistically identical. In this research, we introduced a GAN that comprises of a generator that can generate accurate ECG or PCG signals and a discriminator that can tell the difference between real and artificial signals.

### 3.2 DCGAN

Deep Convolutional Generative Adversarial Networks (DCGANs) are a new form of CNNs developed by Narváez and Percybrooks [22] with certain architectural requirements. These requirements compelled CNN to implement three architectural shifts. In order to work under these constraints, CNN modified its architecture in three ways. One way to improve the network's accuracy is to replace Discriminator strided convolutions and generator fractional strided convolutions are used in place of fully linked hidden layers and pooling layers. The second adjustment is the implementation of LeakyReLU activations across the entire discriminator network and ReLU activations across all except the final layer of the generative model. Thirdly, both the

generator and discriminator will now make advantage of batch normalisation.

### 3.3. Conditional GAN

The extension of GANs to a conditional GANs is possible by conditioning both the generator and discriminator on additional information  $y$  [23]. This additional details, which can take the form of labelled classes or data from other modalities, allows for a more versatile and context-aware generative process. The conditioning process can be achieved by introducing  $y$  as an additional input layer for both the discriminator and generator. The architecture of CGAN is shown in the Figure 5.

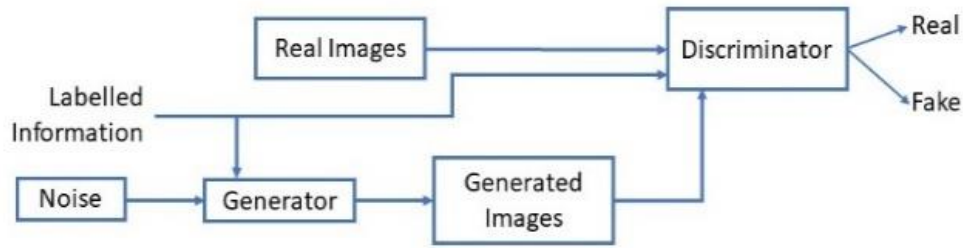


Figure 5. Architecture of CGAN

## 4. EVALUATION METRICS

The metrics used to evaluate GAN performance are generally consistent with those used to evaluate more conventional image creation methods. Simply, this involves calculating an image-to-image distance.

Because of its superior accuracy over the more commonplace Euclidean metric and other distance metrics, we choose to use the structural similarity measure (SSIM) in this research.

### 4.1 Structural similarity index

The SSIM index [24] is a statistic for judging the quality of digital still images and movies. The three terms of brightness, contrast, and structure are the basis for the SSIM Index's quality evaluation. Multiplying these three quantities together yields the overall index.

If we have two photos,  $x$  and  $y$ , we can get their SSIM using the formula below.

$$SSIM(x, y) = \frac{(2 \cdot \mu_x \mu_y + c_1) \cdot (2 \cdot \sigma_{xy} + c_2)}{(c_1^2 + \mu_x^2 + \mu_y^2) \cdot (c_2^2 + \sigma_x^2 + \sigma_y^2)} \quad (2)$$

$\sigma_y$  2-variance of  $y$ .

$\sigma_x$  2-variance of  $x$ .

$\sigma_{xy}$  covariance of  $y$  and  $x$ .

$\mu_y$  mean of  $y$ .

$\mu_x$  mean of  $x$ .

The constant constants  $c_1$  and  $c_2$  are determined based on the dynamic range of pixel values. The SSIM value is equal to one, if and only if both of  $x$  and  $y$  are equal.

### 4.2 Cross correlation coefficient

By measuring the degree of similarity between two signals, cross-correlation analysis can be used to infer information about the signals [25]. The cross correlation of matrices is  $X$  and  $H$  is given in the Eq. (3).  $X$  is of size  $M \times N$  and  $H$  is of size  $P \times Q$ .

$$C(k, l) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(M, n) H^*(m - k, n - l) \quad (3)$$

where,  $-(P-1) \leq k \leq M-1$ ,  $-(Q-1) \leq l \leq N-1$ .

## 5. RESULTS

### 5.1 Datasets

Both the PTB and MIT-BIH databases [26, 27] are used to assess the efficacy of the technique. There are a total of 549 entries in the PTB diagnostic database, representing 290 people (209 men, with an average age of 55.5; age range of 17 to 87, with 57 as the mean, and 81 females, with a mean of 61.6). Fifteen simultaneous measurements of a signal are contained in each record. Every signal is digitally sampled at a rate of one thousand per second. Myocardial infarction, Fetal ECG, PCG recordings, atrial fibrillation, myocardial hypertrophy, myocarditis, healthy individuals, and other data are among the many diagnostic categories available. The MIT-BIH database contains 23 healthy recordings with sequential numbers between 100 and 124, and 25 recordings with junctional, heart blocks and supra-ventricular arrhythmias.



## 5.2 Simulation results

The genuine and synthetic ECG and PCG signals generated by GAN are displayed in Figure 6 and Figure 7, respectively whereas Figures 8 and 9 shows real and artificial ECG and PCG signals generated using DCGAN. Figure 10 shows the ECG signal generated by CGAN. Table 1 gives a comparison

of the evaluation metrics used for three different architectures GAN, DCGAN and CGAN. Since DCGAN uses convolutional networks instead of a fully connected layer as in GAN, DCGAN gives better performance compared to GAN. In CGAN, since labelled or auxiliary information is given to the Generator, it comparatively gives a good performance compared to GAN.

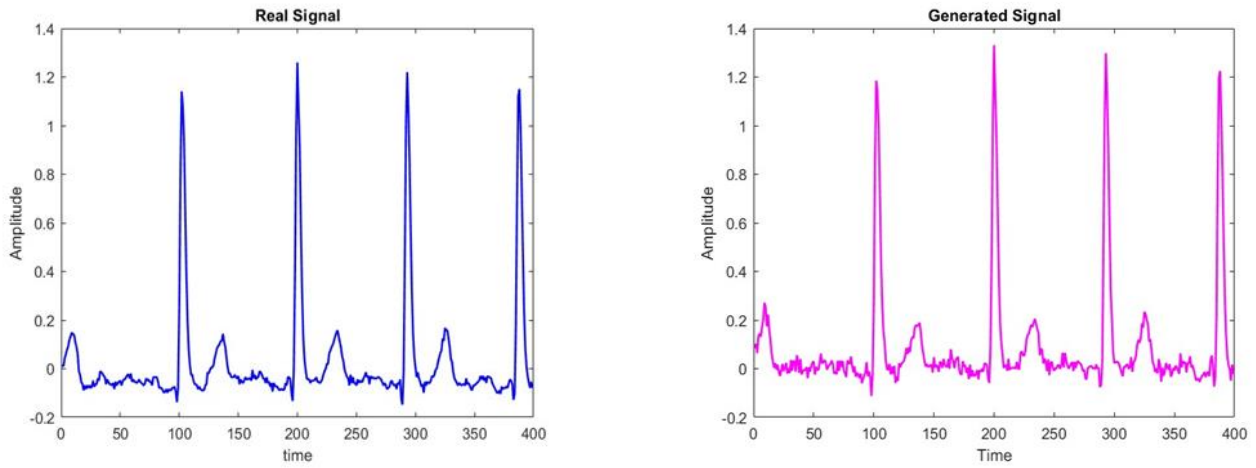


Figure 6. Real and synthesised ECG signals using GAN

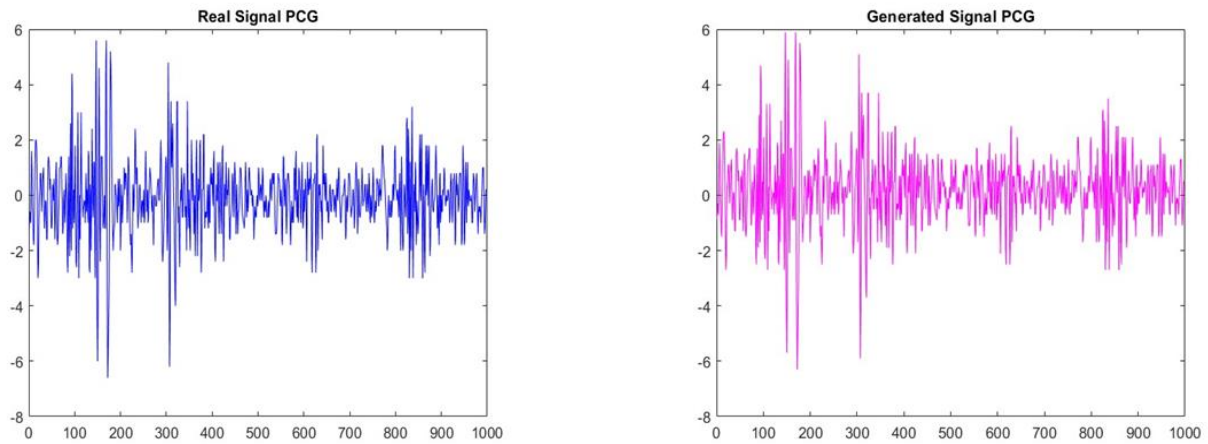


Figure 7. Real and synthesised PCG signals using GAN

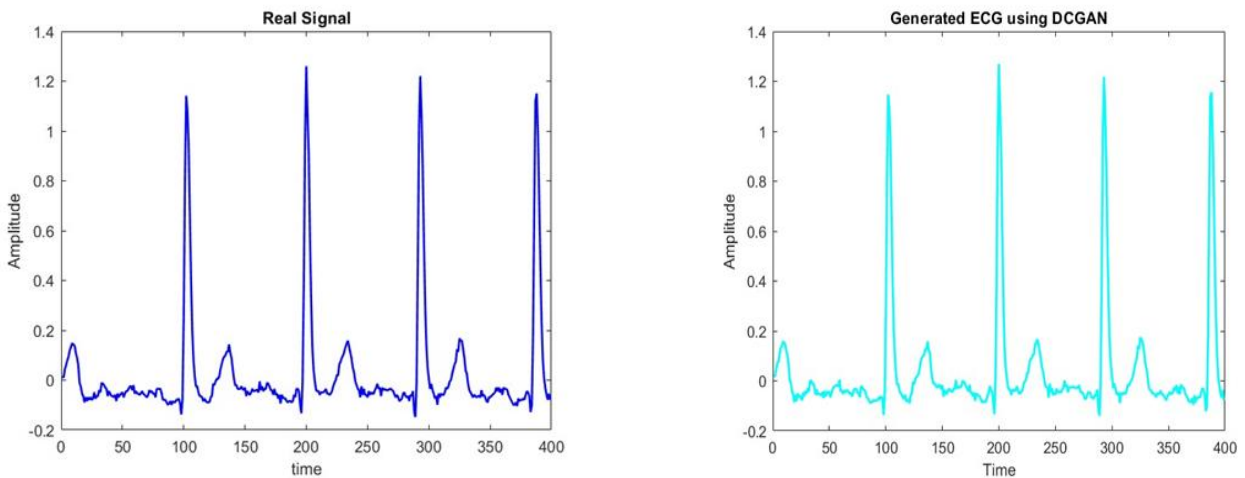


Figure 8. Real and synthesized ECG signals using DCGAN

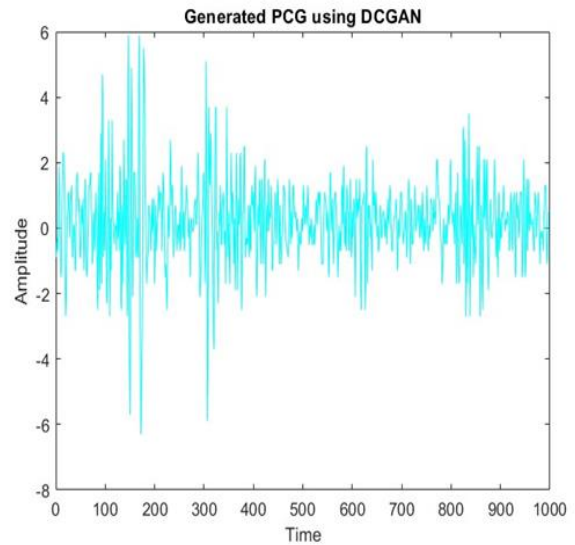
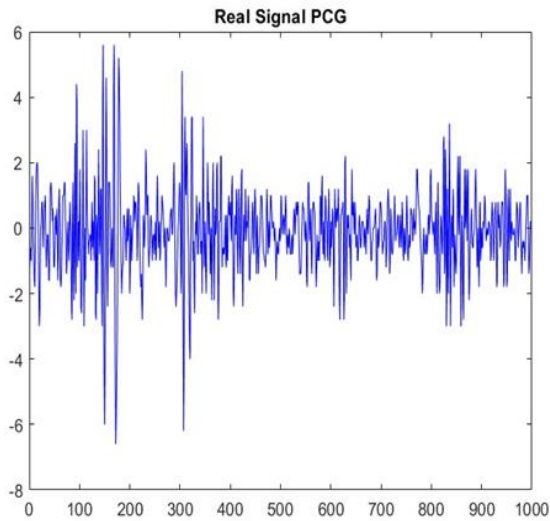


Figure 9. Real and synthesized PCG signals using DCGAN

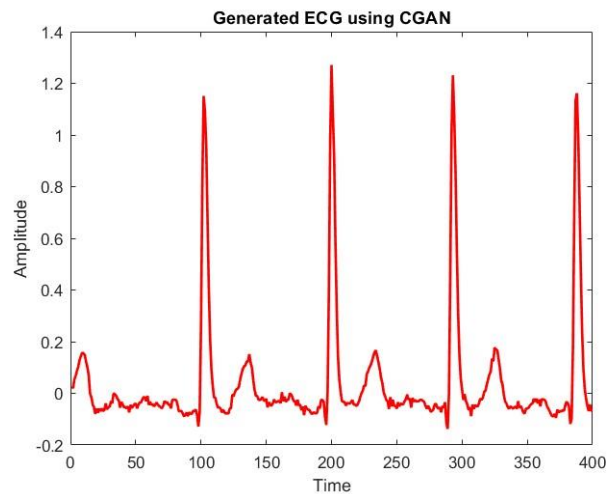
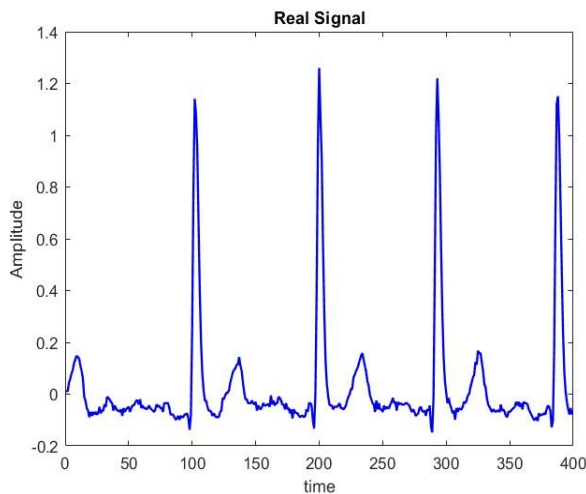


Figure 10. Real and synthesized ECG signals using CGAN

Table 1. Similarity results between real and generated signals

Method	Signal	Cross Correlation Coefficient	SSIM
GAN	ECG	0.9532	0.9621
GAN	PCG	0.9605	0.9699
DCGAN	ECG	0.9652	0.9756
DCGAN	PCG	0.9638	0.9732
CGAN	ECG	0.9785	0.9814

## 6. CONCLUSION

For several reasons, collecting large amounts of patient data may prove challenging. To better train supervised machine learning classifiers on datasets, the synthesis of accurate information or data has arisen as an interesting area of research in healthcare, particularly medicine. To create synthetic heart sounds and ECG signals for use in training/testing classification models, a GAN-based architecture was implemented. Moreover, DCGAN and CGAN architecture are also implemented and the generated ECG and PCG signals shown better performance with CGAN with respect to GAN and DCGAN architectures. These synthetically generated

ECG and PCG samples are given to a classifier as a substitution of original data. The synthetically generated samples showed a 98% similarity to the original data. In the future direction, 99.9 percent comparable images can be generated with the help of improved GAN architectures or by employing diffusion models.

## REFERENCES

- [1] Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning. *Computational Intelligence and Neuroscience*, 2021: 8387680. <https://doi.org/10.1155/2021/8387680>
- [2] Avanzato, R., Beritelli, F. (2020). Automatic ECG diagnosis using convolutional neural network. *Electronics*, 9(6): 951. <https://doi.org/10.3390/electronics9060951>
- [3] Mubarak, Q.U.A., Akram, M.U., Shaukat, A., Ramazan, A. (2019). Quality assessment and classification of heart sounds using PCG signals. *Applications of Intelligent*

- Technologies in Healthcare, 1-11. [https://doi.org/10.1007/978-3-319-96139-2\\_1](https://doi.org/10.1007/978-3-319-96139-2_1)
- [4] Huang, Q.J., Yang, H.R., Zeng, E., Chen, Y.R. (2022). A deep-learning-based multi-modal ECG and PCG processing framework for cardiac analysis. <https://doi.org/10.36227/techrxiv.19668621.v1>
- [5] Mouhni, N., Elkalay, A., Chakraoui, M., Abdali, A., Ammoumou, A., Amalou, I. (2022). Federated learning for medical imaging: An updated state of the art. *Ingénierie des Systèmes d'Information*, 27(1): 143-150. <https://doi.org/10.18280/isi.270117>
- [6] Weerakody, P.B., Wong, K.W., Wang, G.J., Ela, W. (2021). A review of irregular time series data handling with gated recurrent neural networks. *Neurocomputing*, 441: 161-178. <https://doi.org/10.1016/j.neucom.2021.02.046>
- [7] Iqbal, A., Sharif, M., Yasmin, M., Raza, M., Aftab, S. (2022). Generative adversarial networks and its applications in the biomedical image segmentation: a comprehensive survey. *International Journal of Multimedia Information Retrieval*, 11(3): 333-368. <https://doi.org/10.1007/s13735-022-00240-x>
- [8] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11): 139-144. <https://doi.org/10.1145/3422622>
- [9] Skandarani, Y., Jodoin, P.M., Lalande, A. (2023). Gans for medical image synthesis: an empirical study. *Journal of Imaging*, 9(3): 69. <https://doi.org/10.3390/jimaging9030069>
- [10] Naritomi, Y., Adachi, T. (2020). Data augmentation of high frequency financial data using generative adversarial network. In 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), IEEE, 641-648. <https://doi.org/10.1109/WIAT50758.2020.00097>
- [11] Eckerli, F., Osterrieder, J. (2021). Generative adversarial networks in finance: an overview. *arXiv Preprint arXiv: 2106.06364*. <https://doi.org/10.48550/arXiv.2106.06364>
- [12] Weng, Y., Ding, S.Y., Zhou, T. (2022). A survey on improved GAN based image inpainting. In 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), IEEE, Guangzhou, China, pp. 319-322. <https://doi.org/10.1109/ICCECE54139.2022.9712740>
- [13] Liu, H.M., Lu, G.M., Bi, X.H., Yan, J.J., Wang, W.L. (2018). Image inpainting based on generative adversarial networks. In 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), IEEE, Huangshan, China, pp. 373-378. <https://doi.org/10.1109/FSKD.2018.8686914>
- [14] Yuan, C.X., Moghaddam, M. (2020). Attribute-aware generative design with generative adversarial networks. *IEEE Access*, 8: 190710-190721. <https://doi.org/10.1109/ACCESS.2020.3032280>
- [15] Lomov, I., Makarov, I. (2019). Generative models for fashion industry using deep neural networks. In 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), IEEE, Riyadh, Saudi Arabia, pp. 1-6. <https://doi.org/10.1109/CAIS.2019.8769486>
- [16] Ortiz, J.J.G., Phoo, C.P., Wiens, J. (2016). Heart sound classification based on temporal alignment techniques. In 2016 Computing in Cardiology Conference (CinC), IEEE, Vancouver, BC, Canada, pp. 589-592.
- [17] Rizal, A., Handzah, V.A.P., Kusuma, P.D. (2022). Heart sounds classification using short-time Fourier transform and gray level difference method. *Ingénierie des Systèmes d'Information*, 27(3): 369-376. <https://doi.org/10.18280/isi.270302>
- [18] Martins, M., Gomes, P., Oliveira, C., Coimbra, M., da Silva, H.P. (2019). Design and evaluation of a diaphragm for electrocardiography in electronic stethoscopes. *IEEE Transactions on Biomedical Engineering*, 67(2): 391-398. <https://doi.org/10.1109/TBME.2019.2913913>
- [19] Zhang, H., Wang, X.P., Liu, C.C., Liu, Y.Y., Li, P., Yao, L.K., Li, H., Wang, J.K., Jiao, Y. (2020). Detection of coronary artery disease using multi-modal feature fusion and hybrid feature selection. *Physiological Measurement*, 41(11): 115007. <https://doi.org/10.1088/1361-6579/abc323>
- [20] Elbattah, M., Loughnane, C., Guérin, J.L., Carette, R., Cilia, F., Dequen, G. (2021). Variational autoencoder for image-based augmentation of eye-tracking data. *Journal of Imaging*, 7(5): 83. <https://doi.org/10.3390/jimaging7050083>
- [21] Kebaili, A., Lapuyade-Lahorgue, J., Ruan, S. (2023). Deep learning approaches for data augmentation in medical imaging: a review. *Journal of Imaging*, 9(4): 81. <https://doi.org/10.3390/jimaging9040081>
- [22] Narváez, P., Percybrooks, W.S. (2020). Synthesis of normal heart sounds using generative adversarial networks and empirical wavelet transform. *Applied Sciences*, 10(19): 7003. <https://doi.org/10.3390/app10197003>
- [23] Radford, A., Metz, L., Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv Preprint arXiv: 1511.06434*. <https://doi.org/10.48550/arXiv.1511.06434>
- [24] Mirza, M., Osindero, S. (2014). Conditional generative adversarial nets. *arXiv Preprint arXiv: 1411.1784*. <https://doi.org/10.48550/arXiv.1411.1784>
- [25] Shahriari, Y., Fidler, R., Pelter, M.M., Bai, Y., Villaroman, A., Hu, X. (2017). Electrocardiogram signal quality assessment based on structural image similarity metric. *IEEE Transactions on Biomedical Engineering*, 65(4): 745-753. <https://doi.org/10.1109/TBME.2017.2717876>
- [26] Verulkar, N.M., Ambalkar, R.R. (2016). Discriminant analysis of electrocardiogram (ECG) signal using cross correlation. In 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), IEEE, Pune, India, pp. 523-526. <https://doi.org/10.1109/ICACDOT.2016.7877640>
- [27] Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23): e215-e220. <https://doi.org/10.1161/01.CIR.101.23.e215>