Reconstruction of Body Surface Potential From 12-Lead ECG: A Conditional GAN Based Approach

Rohan Banerjee
TCS Research
Tata Consultancy Services
Kolkata, India
rohan.banerjee@tcs.com

Oishee Mazumder
TCS Research
Tata Consultancy Services
Kolkata, India
oishee.mazumder@tcs.com

Ayan Mukherjee
TCS Research
Tata Consultancy Services
Kolkata, India
ayan.m4@tcs.com

Soumitra Sinhahajari
Department of Computer Science & Engineering
Jadavpur University
Kolkata, India
soumitrasinhahajari09@gmail.com

Aniruddha Sinha
TCS Research
Tata Consultancy Services
Kolkata, India
aniruddha.s@tcs.com

***Abstract*—Body Surface Potential Map (BSPM) is an augmented version of 12-lead Electrocardiogram (ECG) with an increased number of electrodes that provides high density spatial information of the cardiac potential on the torso surface for source localization of cardiac abnormalities. A total reconstruction of BSPM is a challenging task. In this paper, we propose a novel Generative Adversarial Network (GAN) architecture to reconstruct 65-lead BSP from standard 12-lead ECG. We present Time-Series GAN (TSGAN), a specially designed modified pix2pix GAN for an accurate reconstruction of time-series BSP data. Further, we propose certain regularization terms in the generator loss function to preserve the key morphological properties of the generated waveform which is a major contribution of this work. The proposed architecture outperforms a Variational Autoencoder (VAE) and a baseline GAN on publicly available dataset in reconstructing 65-lead BSP with morphological preservation.**

***Index Terms*—Body Surface Potential, Electrocardiogram, ECG Morphology, Generative Adversarial Network**

I. Introduction

Cardiovascular Diseases (CVD) have huge prevalence and is regarded as the leading cause of death globally [1]. Among the vast plethora of CVD diagnosis tools, Electrocardiogram (ECG) remains the most commonly conducted procedure. The 12-lead ECG is the standard assessment for cardiac disorders and is used for screening as well as monitoring, aiding healthcare professionals provide both prevention and treatment [2]. Despite being the first level diagnostic tool, 12-lead ECG lacks spatial resolution and is insufficient in arrhythmia localization, cardiac activation pattern mapping through myocardium, atrial and ventricular activation abnormalities, to name a few. A recent approach towards non invasively generating cardiac activation maps for arrhythmic source localization, termed as Electrocardiographic Imaging (ECGI) has gained immense research and medical interests [3].

Electrical activity at myocardium level and successive activation mapping can be reconstructed non-invasively from dense Body Surface Potential Maps (BSPM). BSP, like conventional 12-lead ECG, measures electrical potential of heart at body surface but employs relatively large number of electrodes (often 50 to 300) distributed throughout the thorax surface. Dense distribution of electrodes results in higher degree of accuracy in detecting cardiac conditions and source level abnormalities [4]. In spite of the rich information derived in terms of diagnostic yield, BSP is still not used widely in clinical practices. Main reason for its non inclusion in standard medical practice is management of huge number of leads spread across the torso surface. Standardization of these leads in terms of placement, signal to noise ratio, type of electrodes, etc. are generic hindrance that has prevented the use of BSP over standard 12-lead ECG [5]. BSP electrodes are often made available in form of electrode vests, but custom electrode manufacturing across multiple institutions leads to incompatible electrode interfaces apart from the high cost associated [6]. Selection of lead numbers and location also varies among manufactures and several studies have argued on the optimal number of electrodes required on the torso surface to regenerate cardiac activation maps [7].

A way to overcome the limitations associated with managing large number of physical electrodes and yet generate dense cardiac activation information can be by use of a partial and optimal set of body surface electrodes and reconstruct the complete BSPM data matrix synthetically [8], often using Machine Learning (ML) techniques [9]. ML coupled with information from 12-lead ECG have been mostly utilized to localize ventricular activation origin, improve efficacy of localizing Atrial Fibrillation (AF) ablation sites [10]. There have been limited studies on reconstruction of BSPM from a reduced set of electrodes by using deep generative models like Variational Autoencoders (VAE) [11]. In VAE, the encoding distribution is regularized to generate new data from the latent space [12]. However, the newly generated data are often noisy. Generative Adversarial Networks (GAN) often outperform

the traditional VAE. In spite of being extensively used in realistic image synthesis, GANs are yet to be fully explored in biomedical applications dealing with time-series data. In this paper, we propose a novel conditional GAN architecture to generate 65-lead BSP data from standard 12-lead ECG. We claim the following novelties:

- We propose Time-Series GAN (TSGAN), a modified pix2pix GAN [13] architecture for reconstruction of BSP time-series from reduced set of ECG leads. We have compared the results with the only available prior work for generation of BSP from 12-lead ECG using VAE [11].
- Morphology-preserving regularization terms in the loss function for an accurate reconstruction.

A brief description of our dataset is provided in Section II. A detailed description of the proposed TSGAN architecture is provided in Section III, followed by experimental results and a conclusion.

II. Dataset

We have used BSP and ECG data from EDGAR database, contributed by Radboud University, Nijmegen, Netherland [14]. The dataset, popularly known as 'Nijmegen'data contains 65-lead BSP signals acquired using 53 electrodes on front and 12 on back of the torso. Fig.1a shows the 9 electrode locations for a standard 12-lead ECG configuration and the Nijmegen BSP configuration. In 12-lead ECG configuration, there are three limb leads, six precordial leads and three augmented leads. The limb leads namely, lead I ($= LA - RA$), lead II ($= LL - RA$), lead III ($= LL - LA$) are computed from the three limb electrodes namely, Right Arm (RA), Left Arm (LA), Left Leg (LL). The precordial leads are measured directly from the electrodes $V1$ to $V6$ placed at specific anatomically standardized locations. The rest three augmented leads $aVR(= RA - (LA + LL)/2), aVL(= LA - (LL + RA)/2)$ and $aVF(= LL - (LA + RA)/2)$ are calculated from the limb electrodes [15].

In the Nijmegen BSP dataset, along with the potential profiles of 65 leads, the co-ordinate information to extract the 9 leads (limb and precordial) of 12-lead ECG are also provided. We extract the 12-lead ECG information and use this as input to reconstruct the complete 65-lead BSP data using our proposed TSGAN model discussed next.

The original dataset has 16 independent recordings, each having 9999 data-points. We rearrange them in two equal parts in a random manner to form the training and the test set. Each recording is broken into three non-overlapping segments of 3000 data-points which are considered as individual instances to be applied to our model. The waveforms in the dataset are sampled at 1000 Hz. It is strictly ensured that multiple segments from a single recording are not mixed up in training and test set. We apply standard data augmentation techniques like band-pass filtering, addition of white Gaussian noise, baseline shift, random cropping etc. [16] to enhance the number of instances in the training set by 20 times.

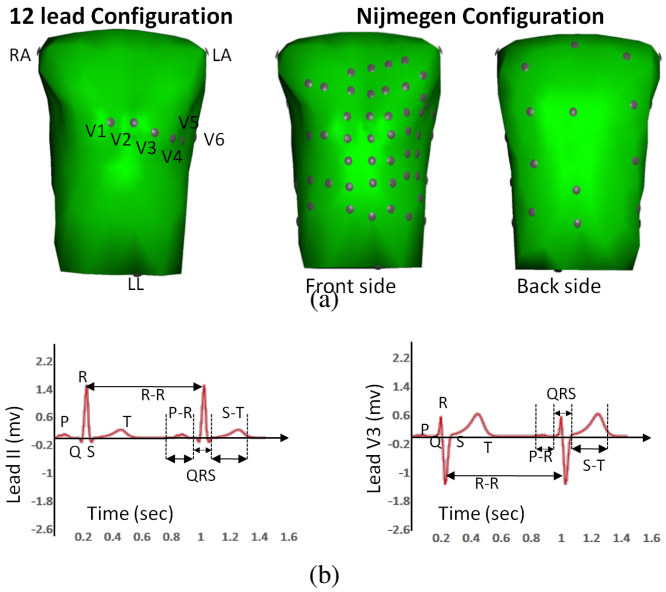

Fig. 1. a) Electrode placement in conventional 12-lead ECG and the Nijmegen BSP configuration; b) Cardiac electrical activity recorded at torso from lead II and lead V3

III. Proposed TSGAN Architecture

A GAN has two components, a generator (G) and a discriminator (D) [17]. The generator takes random noise as input and creates new plausible synthetic examples, the discriminator classifies whether its input is fake (generated) or real. Traditional GANs have no control over the types of generated examples which can be improved by conditional GANs [19].

The pix2pix model is a conditional GAN popularly used in image to image translation [13]. Unlike traditional GANs, the generator of pix2pix GAN takes a source image as input and transforms into a translated image. The discriminator determines whether the translated image is a plausible transformation of the source image. In this paper, we propose TSGAN, a modified pix2pix model to generate 65-lead BSP time-series from 12-lead ECG. If x is the input (12-lead ECG), z is the random noise, y is the target output (original 65-lead BSP), $G(x, z)$ is the generated 65-lead BSP, then the objective function of a conditional GAN is represented by:

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[logD(x, y)] + \mathbb{E}_{x,z}[log(1 - D(x, G(x, z)))] \quad (1)$$

In the original pix2pix model, the L1 loss of the pixel difference between the target and the generated image is added with the generator loss function for minimizing during training. In TSGAN, we minimize the L2 loss of the amplitude difference between the target and generated 65-lead BSP which causes a more accurate and noise-free reconstruction of time-series. A comparison between the L1 and L2 loss functions are given in the results sections to justify our approach. In addition, few more regularization terms are also added to preserve the key morphological properties in the generated waveform.

Fig.1b shows the normal ECG template of lead II and lead V3 along with major morphological points of interest like P,Q,R,S,T points and their intervals. These specific points in

ECG correspond to major electrophysiological events leading in to generation of the ECG patterns, like P wave signifies atrial depolarization, QRS complex represents ventricular depolarization and T wave corresponds to ventricular repolarization. The morphological patterns of these specific regions along with their intervals vary for normal and pathological conditions [15]. The L2 loss of the position differences of P,Q,R,S,T peaks in the target and generated waveform on time axis are also minimized along with the generator loss function. The final objective of the proposed GAN is:

$$G* = arg \min_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda_1.\mathcal{L}_{L2_{sig}}(G) + \lambda_2.\mathcal{L}_{L2_P}(G) + \lambda_3.\mathcal{L}_{L2_Q}(G) + \lambda_4.\mathcal{L}_{L2_R}(G) + \lambda_5.\mathcal{L}_{L2_S}(G) + \lambda_6.\mathcal{L}_{L2_T}(G) \quad (2)$$

where,

$$\mathcal{L}_{L2_{sig}}(G) = \mathbb{E}_{x,y,z}[||y - G(x,z)||_2] \quad (3)$$

The other terms, $\mathcal{L}_{L2_P}, \mathcal{L}_{L2_Q}, \mathcal{L}_{L2_R}, \mathcal{L}_{L2_S}, \mathcal{L}_{L2_T}$ denote the L2 loss of time difference between the P,Q,R,S,T locations in the target and the generated waveform. The constants, denoted by $\lambda_i, 1 <= i <= 6$, in eqn. 2 can be adjusted during training to assign higher weightage to selected portions of the BSP waveform requiring more accurate reconstruction.

A. The Generator Model

The generator takes 12-lead ECG time-series having 3000 samples in each lead (channel) as input (dimension = (3000, 1, 12)) and maps it into 65-lead BSP (dimension = (3000, 1, 65)) using a 1D convolutional encoder-decoder structure shown in Fig.2. The output tensor dimension of each block is provided at the bottom of the block. The encoder has three

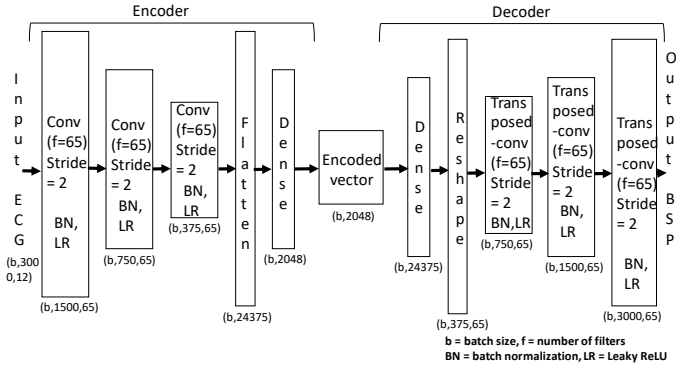

Fig. 2. Generator architecture of proposed TSGAN

1D convolutional layers, with associated batch normalization and leaky ReLU activation layer. In each layer, convolution is performed with a kernel dimension of 5 and stride length of 2 by applying zero-padding to the inputs. We apply 65 filters to each convolutional layer. The output is flattened after the third layer and is applied to a dense layer with leaky ReLU activation to get the encoded vector. The decoder comprises a series of transposed-convolution layers having similar kernel dimension with stride length of 2 to increase the output dimension to eventually map into the desired output. In order to maintain stochasticity in the output, we add 40% dropout to different layers of the encoder. Additionally, the input data (x) is added with low amplitude random Gaussian noise (z).

B. The Discriminator Model

The discriminator is a 1D convolutional patchGAN discriminator which takes two sets of paired inputs. The input 12-lead ECG (x) paired with the real (y -the actual) and fake ($G(x,z)$ -generated) 65-lead BSP data conditioned on the input are applied to the discriminator in separate batches. It aims to classify whether the pair of data is real or fake by minimizing the likelihood of a negative log identifying real and fake data. We closely follow the patchGAN discriminator architecture proposed in the original pix2pix model in [13]. In patchGAN discriminator, the model outputs a tensor where each element is corresponding to a patch of the input and the value indicates whether the patch is real or fake. The discriminator is applied convolutionally across the input data, averaging all responses to provide the final output prediction. The patch size is selected as 70x1. The discriminator in our model comprises 8 1D convolution layers with associated batch normalization and activation layers. Leaky ReLU is used in the first seven layers. The final layer uses sigmoid function for classification. The output at the end of the final block is a (30,1) tensor, where each point represents a 70x1 patch in the input.

C. Training of Proposed TSGAN

The proposed TSGAN is implemented in Python 3.8.10 using TensorFlow 2.6.0 library. The model is trained on a computer system having Intel® Xeon(R) 16-core processor, 64 GB of RAM and an NVidia GeForce GTX 1080 Ti graphics processing unit. NeuroKit2 [18], a Python package for neurophysiological data analysis is used to extract the P,Q,R,S,T regions. The standard procedure of training a GAN [17], [13] is followed. The generator takes x as input and generates $G(x,z)$. The discriminator takes x and $G(x,z)$ (fake) as well as x and the actual target data y (real) as inputs. The real and fake data are labeled by arrays of ones and zeros. Binary cross entropy loss is defined to model the objectives of the generator and discriminator. For the generator, we measure the cross entropy loss of the generated images and an array of ones. The discriminator loss is calculated by averaging the sum of the cross entropy loss for the real and the fake data. In each iteration, we first measure the discriminator loss followed by the generator loss. Next, the gradient of the loss is measured with respect to model weights using backpropagation to update the weights. The regularization constants are empirically chosen for an optimum reconstruction. The value of λ_1 is selected as 50. The other constants are set as $\lambda_2 = 1, \lambda_3 = 10, \lambda_4 = 20, \lambda_5 = 15, \lambda_6 = 20$ for an optimum reconstruction. It can be seen that that to preserve the overall morphology of ECG, the λ_1 needs to be given the maximum weight. However, among the 5 key points (P, Q, R, S, T) in ECG, the R and T gets the maximum weights in regularization as they corresponds to the peaks during ventricular depolarization and repolarization respectively.We use Adam optimizer with a learning rate of 0.0002, and

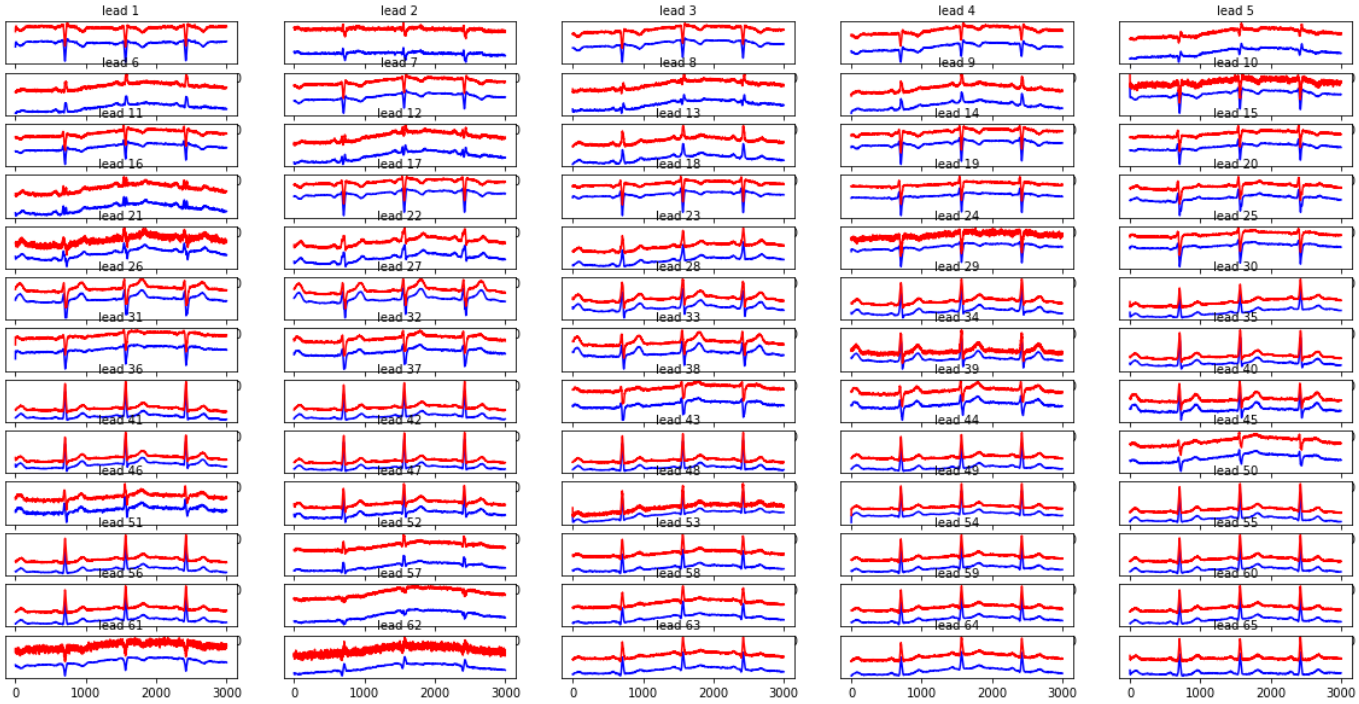

Fig. 3. Sample plot of reconstructed 65-lead BSP waveform from a 12-lead input ECG data in the test set as generated by our proposed TSGAN. The actual waveform are shown in blue and the reconstructed waveform are in red

momentum parameters $\beta_1 = 0.5, \beta_2 = 0.999$ for training. The batch size is taken as 4. The model weights are initialized from a normal distribution with zero mean and standard deviation of 0.02. The end-to-end GAN is trained for 200 epochs. Once the training is done, the discriminator is discarded and the generator model is used for BSP reconstruction on test data.

IV. Experimental Results

In Table I, we show a comparative analysis of overall reconstruction loss obtained by using L1 and L2 loss in the generator loss function as a regularization term. Here, we measure the overall reconstruction loss on the test set in terms of Normalized Root Mean Square Error (NRMSE) between the original and the reconstructed BSP data across various leads. Since we are dealing with reconstruction of time-series data, it can be observed that the reconstruction loss is much lower in using L2 loss, as proposed in our approach, compared to L1 loss based regularization proposed in the original pix2pix model which was proposed for image to image translation.

Sample waveform of reconstructed 65-lead BSP data from a 12-lead ECG in the test set along with the original waveform are shown in Fig.3. The reconstructed waveform closely matches the morphology of the original waveform in almost all leads except lead: 10, 21, 24, 34, 48, 61, 62, where the reconstructed data are noisy. In general, the output leads, located far from any of the base 9 leads of the input ECG yield higher reconstruction loss.

A detailed quantitative performance analysis is shown in Table II. The proposed TSGAN is compared with two other approaches, the VAE-based approach in prior art [11] and a baseline conditional GAN (cGAN). The baseline cGAN has an identical structure to our proposed TSGAN, but it does not consider the morphology-specific regularization terms in eqn.2. We also present the lead position and the NRMSE value corresponding to the minimum and the maximum reconstruction loss along with the average reconstruction loss for all the leads. It can be clearly observed that the TSGAN yields the least reconstruction loss. The GAN-based models inevitably outperform the VAE model. However, utilizing

TABLE I
Comparison between using L1 loss and L2 loss as a regularization term in the generator loss in terms of reconstruction loss in NRMSE

Reconstruction loss NRMSE	L1 loss	L2 loss (Proposed)
Average reconstruction loss across all leads	0.13	0.07

TABLE II
Comparison of the proposed TSGAN with VAE and baseline cGAN in terms of reconstruction loss in NRMSE

Model architecture	Minimum reconstruction loss	Maximum reconstruction loss	Average reconstruction loss
VAE [11]	0.08 (lead 13)	0.61 (lead 48)	0.31
Baseline cGAN	0.06 (lead 24)	0.32 (lead 62)	0.18
Proposed TSGAN	0.02 (lead 30)	0.12 (lead 62)	0.07

the morphology-preserving regularization, which is a key contribution of the paper, we can achieve a much improved reconstruction compared to the baseline cGAN.

In Table III, we report the maximum and average NRMSE for RR, QRS, QT and PR interval time distances between the actual and target waveform in the test set. An accurate detection of the P and S points are often unreliable on noisy ECG data. Hence, our model reports higher error values in reconstruction of the PR and QRS segments compared to the RR and QT segments. The TSGAN still reports the least error for all the features, justifying its usability in morphology-preserving reconstruction.

TABLE III
MORPHOLOGICAL ANALYSIS OF EXTRACTED FEATURES FROM THE RECONSTRUCTED DATA W.R.T THE ORIGINAL DATA IN TERMS OF NRMSE (ME: MAXIMUM ERROR, AE: AVERAGE ERROR, L=LEAD)

	RR		QRS		QT		PR	
Model	ME	AE	ME	AE	ME	AE	ME	AE
VAE [11]	0.26 (L24)	0.08	0.55 (L10)	0.36	0.25 (L62)	0.19	0.91 (L62)	0.77
Baseline cGAN	0.08 (L24)	0.03	0.82 (L62)	0.35	0.26 (L24)	0.12	0.68 (L61)	0.47
Proposed TS-GAN	0.08 (L61)	0.02	0.38 (L62)	0.13	0.12 (L30)	0.06	0.42 (L62)	0.29

The weights for the regularizers (λ_1 to λ_6) are empirically chosen based on the healthy ECG data used in the current experiment. These weights need to be tuned for disease conditions. For example in the case of ischemic disease, the weight of the λ_6 (corresponding to T) needs to be higher compared to others. Similarly, for the scenario of Atrial Fibrillation, where the P waves are missing or having a different morphology in certain ECG cycles and the R-R intervals are irregular in nature, the λ_2 and λ_4 would need special focus. Such enhancements are planned as future extensions of the present work.

V. CONCLUSION

The TSGAN model proposed in this paper can faithfully reconstruct 65-lead BSP from standard 12-lead ECG and also surpassed the reconstruction accuracy of other related approaches. The reconstruction reported in this paper is based on a healthy population. The L2 loss function and incorporation of the regularizers related to the keys points in ECG leads to a better performance in reconstruction of BSP as compared to previous approaches.

In future, we plan to extend the generation of BSP for disease conditions by means of the disease specific morphological preservation through adjustment in the loss function. The proposed method of BSP generation can potentially remove the hindrance associated with multiple electrode management and yet provide high resolution cardiac activity information that could aid in source localization of arrhythmia using noninvasive estimations and ECGI techniques.

REFERENCES

[1] Roth.GA, Mensah.GA, Fuster.V et al; GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol. 2020 Dec 22;vol:76(25); pp:2982-3021.

[2] Fye WB. A history of the origin, evolution, and impact of electrocardiography; Am J Cardiol. 1994; vol:73; pp:937–949

[3] Shah AJ, Hocini M, Pascale P, Roten L, Komatsu Y, Daly M et al. Body surface electrocardiographic mapping for non-invasive identification of arrhythmic sources. Arrhythm Electrophysiol Rev 2013; vol:2, pp:16–22

[4] Cluitmans M.J.M., Peeters R.L.M, Westra R.L, Volders P.G.A; Noninvasive reconstruction of cardiac electrical activity: update on current methods, applications and challenges Neth Heart J (2015); vol: 23; pp:301–311

[5] Ambroggi, L.D.; Musso, E.; Taccardi, B. Body Surface Potential Mapping. In Comprehensive Electrocardiology; Macfarlane, P., Veitch Lawrie, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; Vol: 32, pp: 1375–1413.

[6] Bergquist, J.; Rupp, L.; Zenger, B.; Brundagem, J.; Busatto, A.; MacLeod, R.S. Body Surface Potential Mapping: Contemporary Applications and Future Perspectives. Hearts 2021,vol: 2; pp: 514–542.

[7] Salud Guillem M, Bollmann A, Climent A, Husser D, Millet-Roig J, Castells F. How many leads are necessary for a reliable reconstruction of surface potentials during atrial fibrillation? IEEE Transactions on Information Technology in Biomedicine 2009;vol:13(3); pp:330–340.

[8] Peter M. van Dam, Roderick Tung, Kalyanam Shivkumar, Michael Laks, Quantitative localization of premature ventricular contractions using myocardial activation ECGI from the standard 12-lead electrocardiogram, Journal of Electrocardiology, Vol:46(6), pp:574-579

[9] Karoui A, Bendahmane M and Zemzemi N (2021) Cardiac Activation Maps Reconstruction: A Comparative Study Between Data-Driven and Physics-Based Methods. Front. Physiol. 12:686136. doi: 10.3389/fphys.2021.686136

[10] Boyle, P. M., Zghaib, T., Zahid, S., Ali, R. L., Deng, D., Franceschi, W. H., et al. (2019). Computationally guided personalized targeted ablation of persistent atrial fibrillation. Nat. Biomed. Eng. vol:3, pp:870–879. doi: 10.1038/s41551-019-0437-9

[11] D. Bizcaino et al., "Neural Network-Based Matrix Completion for Minimal Configuration of Body Surface Potential Mapping," 2019 Computing in Cardiology (CinC), 2019, pp:1-4, doi: 10.23919/CinC49843.2019.9005744.

[12] Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. Foundations and Trends in Machine Learning, vol:12(4), pp:307-392.

[13] Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1125-1134).

[14] Aras K, Good W, Tate J, Burton B, Brooks D, Doessel O, Schulze W, Patyogaylo D, Wang L, Van Dam P, MacLeod R, Coll-Font J. Experimental Data and Geometric Analysis Repository: EDGAR. Journal of Electrocardiology 2015; vol:48(6), pp:975–981

[15] Kligfield P, Gettes LS, et.al; Recommendations for the standardization and interpretation of the electrocardiogram: part I: the electrocardiogram and its technology a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; J Am Coll Cardiol. 2007 Mar 13; vol:49(10), pp:1109-27. doi: 10.1016/j.jacc.2007.01.024. PMID: 17349896.

[16] Yang, S., Xiang, H., Kong, Q., & Wang, C. (2020, September). Multi-label classification of electrocardiogram with modified residual networks. In 2020 Computing in Cardiology (pp. 1-4). IEEE.

[17] Goodfellow, I. J., et al. "D., WardeFarley, S. Ozair, A. Courville, and Y. Bengio,"Generative Adversarial." Networks,", ArXiv e-prints (2014).

[18] Makowski, D., Pham, T., Lau, Z.J., Brammer, J.C., Lespinasse, F., Pham, H., Schölzel, C. and Chen, S.H., 2021. NeuroKit2: A Python toolbox for neurophysiological signal processing. Behavior research methods, 53(4), pp.1689-1696.

[19] Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.