Forward-Solution Aided Deep-Learning Framework for Patient-Specific Noninvasive Cardiac Ectopic Pacing Localization

Yashi Li, Huihui Ye, and Huafeng Liu(✉)

State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China
liuhf@zju.edu.cn

Abstract. Accurate localization of the ectopic pacing is the key to effective catheter ablation for curing cardiac diseases such as premature ventricular contraction (PVC) and tachycardia. Invasive localization method can achieve high precision but has disadvantages of high risk, high cost, and time-consuming process, therefore, a non-invasive and convenient localization method is in demand. Noninvasive methods have been developed to utilize electrophysiological information provided by 12-lead electrocardiogram (ECG), and most of them are purely based on end-to-end data-driven architecture. This architecture generally needs a substantial and comprehensive labeled dataset, which is very difficult to obtain for whole ventricular ectopic beats in clinical setting. To address this issue, we propose a framework that combines cardiac forward-solution simulation and deep learning network for patient-specific noninvasive ectopic pacing localization. For each patient, it only requires his/her own CT images to establish a specific heart-torso model and to simulate various ECG data from different ectopic pacing locations and uses this simulated ECG data as the training dataset for our designed network. The network mainly contains time-frequency fusion module and local-global feature extraction module. Five PVC patient ECG data are tested with high precision and accuracy for ectopic pacing localization, which shows its high-potential in clinical setting.

Keywords: Ectopic pacing localization · Forward-solution · Deep Learning

1 Introduction

Catheter ablation surgery is a common operation that treats premature ventricular contraction arrhythmia effectively, and it mainly includes two steps: electrical physiological measurement and radiofrequency ablation. For electrical physiological measurement, it requires doctors to insert an electrode catheter in patient's heart, analyze the causes and parts of arrhythmia through the electrical signal obtained by the catheter, and finally determine the specific location of ectopic pacing and catheter ablation [1]. It highly depends on the doctors' experience and is very time-consuming [2]. Thus a noninvasive localization method that determines the ectopic pacing location in advance will significantly shorten the process of catheter ablation surgery and brings benefits for both doctors and patients.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14226, pp. 195–205, 2023.
https://doi.org/10.1007/978-3-031-43990-2_19

There is a trend to use computational tools based on 12-lead electrocardiogram (ECG) to analyze ablation location since ECG has been proven that can provide information about pacing areas in patients [3–5]. Earlier studies focused on extracting features of QRS axis through mathematical statistics and machine learning methods [6]. In recent years, deep learning has become a research hotspot due to its excellent feature extraction ability, which can automatically learn the mapping relationship between signal characteristics and pacing position [7–10]. Most methods are purely end-to-end data-driven architecture based on a large number of clinical databases. However, that will be greatly influenced by the comprehensiveness and size of the labeled data [11], especially when obtaining clinical data of whole ventricular ectopic beats is difficult.

Rapid developments in computer performance and theoretical knowledge have enabled detailed, physiologically realistic whole-heart simulations of arrhythmias and pacing [12, 13]. Based on this, a forward-solution computational mapping system for accurate localization of atrial and ventricular arrhythmias has been proposed, where a comprehensive arrhythmia simulation library is generated [14]. It can eliminate the impact of insufficient clinical trial data on the algorithm's accuracy. However, since the computational models are not specific, it will introduce errors in the ectopic pacing location process.

In this paper, inspired by work in [14], we propose a forward-solution aided deep-learning framework to realize noninvasive prediction of ectopic pacing from 12-lead ECG. Patient-specific heart-torso forward model is built, and a time-frequency fusion network based on the local-global feature extraction module is designed. The advantages of this paper can be summarized from three aspects:

1. Propose a framework that is trained based on ECG simulation data from the specific patient's CT for noninvasive cardiac ectopic pacing localization. It can eliminate the effect of insufficient clinical data and patient variance error on location accuracy.
2. Propose a network that combines time-frequency information and local-global information to achieve precise ectopic pacing location based on a small training data set.
3. Proposed method achieves great performance on PVC patient data, which demonstrates its potential for clinical cardiac treatment.

2 Methodology

2.1 Overview of Location Framework

Figure 1 outlines our framework for cardiac ectopic pacing localization. In brief, we construct a computational model of patient-specific anatomical structures, using a suitable cardiac source model and a transfer matrix H to obtain the solution of the electrocardiography forward problem for simulating whole-ventricle focal pacing as well as the corresponding ECGs. We then utilized this modeled dataset to train a deep-learning structure to locate the ectopic pacing. The structure is first evaluated on the simulated data and subsequently tested on clinical PVC patients' ECG data. We will describe the specific principles of our pipeline in detail in the following sections.

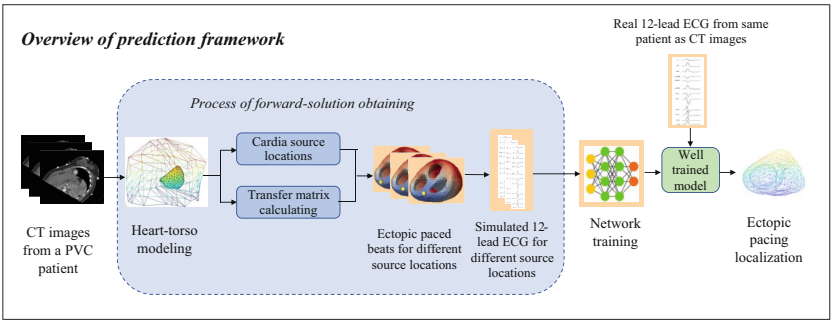

Fig. 1. Proposed forward-solution aided deep-learning framework for noninvasive cardiac ectopic pacing localization

2.2 Process of Forward-Solution Obtaining

Establishing a specific heart-torso model requires the geometric structure of both patient's heart and torso, which can be constructed from CT imaging data from the common preoperative scanning. The solution to the forward problem is obtained from the process of calculating torso surface potentials based on known cardiac source parameters [15], which can be mathematically represented as:

$$y = A(x) \tag{1}$$

where y denotes the torso surface and, x denotes the cardiac source, A is a transfer function dependent on the source model, which has been demonstrated in prior research that can be expressed as a transfer matrix $H \in R_{N*M}$. . This transfer matrix can be calculated through the resolution of the relationship between the cardiac and torso domains based on the finite element method (FEM) or boundary element method (BEM). In terms of the cardiac source, the two most common models are the activation-based model and the potential-based model [16]. One regards the arrival time of the depolarized wavefront as the main characteristic of the electrical activity of the heart; the latter uses the time-varying potential of the cardiac surface.

Compared to activation-based model, the potential-based model considers spatial distribution of cardiac potentials, offering insights into cardiac electric field and potential formation and propagation. Its comprehensive representation of cardiac electrical activity is advantageous for analyzing complex arrhythmias and cardiac disorders. In our experiment, we considered a simple two-variable potential-based model [17]; the process of cardiac excitation can be described as the follows:

$$\frac{\partial u}{\partial t} = \nabla(\boldsymbol{D}\nabla u) + f_1(u, v) \tag{2}$$

$$\frac{\partial v}{\partial t} = f_2(u, v) \tag{3}$$

where u is the transmembrane potentials, v is the conduction current, D is the diffusion tensor dependent on 3-D myocardial structure, and tissue conductive anisotropy, $\nabla(\boldsymbol{D}\nabla u)$

is the diffusion term. Functions f_1 and f_2 produce TMP shapes, which can be represented as:

$$f_1(u, v) = ku(u - a)(1 - u) - uv \tag{4}$$

$$f_2(u, v) = -e(v + ku(u - a - 1)) \tag{5}$$

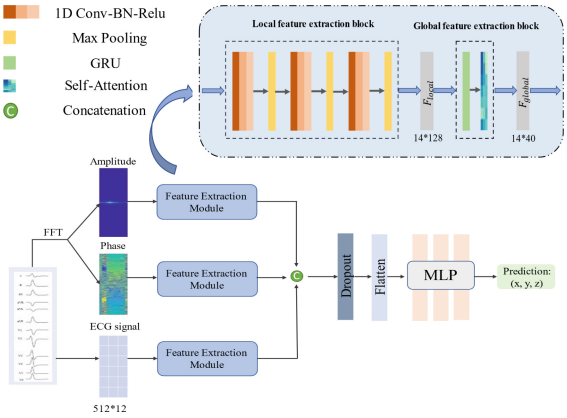

Fig. 2. Diagram of our proposed network

According to[17], parameter k, a, e is set as: $a = 0.15, k = 8, e = 0.01$. The longitudinal and transverse tensors of the diffusion tensor D are set to 4 and 1.

2.3 Network Structure

Figure 2 shows the structure of the proposed network. It mainly consists of the feature extraction part and the feature fusion part. Briefly, the 12-lead signal data (512 sampling points) is represented as a 512 * 12 matrix. Each lead signal undergoes Fourier transform to obtain frequency domain information, resulting in matrices of signal amplitude and signal phase with the same dimensions as the time domain matrix. These three matrices are input into the local-global feature extraction module to obtain encoded time and frequency information. An early fusion method is selected to concatenate these multi-level features, which is transformed into 3-dimensional coordinates of the predicted ectopic pacing location $\tilde{P}(x, y, z)$ by multilayer perceptron (MLP). The loss of the proposed network is defined as:

$$Loss = \frac{1}{N} \sum_{i=1}^{N} \|\tilde{P}(x, y, z) - P(x, y, z)\|_2^2 \tag{6}$$

where N is the number of samples and $P\ (x, y, z)$ is the labeled coordinates. To minimize the loss function, we can finally get a well-trained model. The training uses the Adam optimizer with a learning rate of 1e−4 and a batch size of 32.

Feature Fusion in Time-Frequency Domain. The electrocardiogram signal is a temporal signal, and therefore its analysis can be performed in both time and frequency domains.

The algorithm proposed in reference [18] was the first to convert the ECG signal to a Fourier spectrum to achieve discrimination of ventricular tachyarrhythmia. Currently, most algorithms process ECG signals in the time domain but ignore information in the frequency domain. Therefore, we choose the feature fusion method to combine time and frequency information for better accuracy localization. The Fourier transform breaks down a function into its constituent frequencies [19]. Given an ECG signal vector $\{\phi_t\}$ with $t \in [0, T-1]$ where T is the total number of samples, its Discrete Fourier transform (DFT) can be expressed as:

$$\Phi_k = \sum_{t=0}^{T-1} \phi_t e^{-\frac{2\pi i}{T} tk}, 0 \leq k \leq T-1 \tag{7}$$

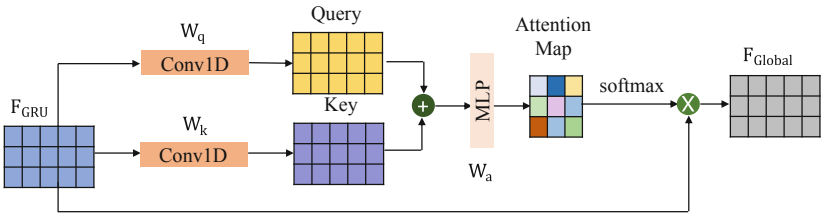

Fig. 3. Detailed illustration of our self-attention module

For each sample moment t, DFT algorithm generates a new representation Φ_k as the sum of all the original input ϕ_t. $\{\Phi_k\}$ is a complex matrix that can be represented as:

$$\{\Phi_k\} = \{r_k\} e^{i\{\theta_k\}}, 0 \leq k \leq T-1 \tag{8}$$

where r_k and θ_k, respectively, denote amplitude and phase. We then separate amplitude matrix $\{r_0, r_1, \ldots, r_{T-1}\}$ and phase matrix $\{\theta_0, \theta_1, \ldots, \theta_{T-1}\}$ as new features that are sent to the network along with signal matrix $\{\phi_0, \phi_1, \ldots, \phi_{T-1}\}$ before normalization to $[-1, 1]$.

Local-Global Feature Extraction Module. Clinical diagnosis of heart disease based on 12-lead ECG mainly depends on its local features, such as P wave, T wave, or QRS complex. Convolutional neural network (CNN) has proven effective for extracting local information of ECG waveforms [20]. However, CNN's limited receptive field size makes it challenging to capture global feature representations, including dependence on long-distance signals and different leads. On the other hand, models like GRU [21] and attention mechanism [22], typical models in natural language processing, excel at capturing long-range dependencies but may compromise local feature details[23]. Inspired by these, our feature extraction module is divided into local feature extraction module and global feature extraction module, ensuring the acquisition of detailed and comprehensive information and improving the network's ability to encode ECG signals.

The local feature extraction block consists of three Conv1D blocks, each containing a 1DConv-BN-Relu layer and a max pooling layer. The Conv1D-BN-Relu layer comprises a one-dimensional convolutional layer, a batch normalization layer, and an activation layer with a Relu activation function. The kernel size of the convolutional layer is sequentially set to 7, 5, and 3; the stride is sequentially set to 2, 2, and 1.

Figure 2 shows the global feature extraction block. Gate Recurrent Unit (GRU)is a kind of recurrent neural network that occupies less memory and is more suitable for small data training [21], which is in line with our needs. The GRU layer is defined as:

$$H_{GRU}^{t} = GRU(\begin{matrix} H_{GRU}^{t-1}, F_{local}^{t}; \\ W_{GRU}^{t-1} \end{matrix}) \tag{9}$$

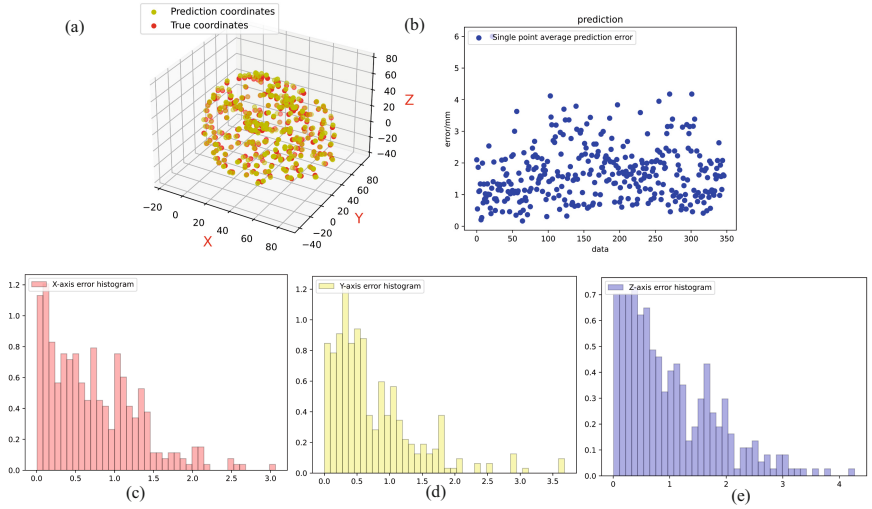

Fig. 4. Overall performance on testing data. (a) 3D display of predicted vs. labeled coordinates. (b) Location error point plot. (c) X-axis error histogram. (d) Y-axis error histogram. (e) Z-axis error histogram.

where $GRU(.)$ is the GRU network, H_{GRU} is the hidden state, F_{local} is the input and W_{GRU} represents the corresponding weight and biases.

In order to extract a broader dependence relationship, we introduced the additive self-attention mechanism after the GRU layer. Its network structure is shown in Fig. 3, which can be represented by the formula:

$$F_{Global} = softmax(MLP(W_q F_{GRU} + W_k F_{GRU}))F_{GRU} \tag{10}$$

where, F_{GRU} is the input of this part; more precisely, it denotes the output matrix from the previous GRU layer, with dimensions set to (14,40). W_* represents the weights to be learned and set to be (40,32); in this case, they are configured as a matrix with dimensions (40, 32). $MLP(.)$ means single layer MLP used for calculating the similarity of the

Query matrix and Key matrix. This operation yields an attention map that captures the correlation among the elements of the matrix. Finally, by applying the softmax function, we obtain the output F_{Global}, which encompasses the global feature.

3 Experimental Result

3.1 Overall Performance

The experimental data of this part are generated by ECGSIM software based on the UDL source model [24]. The solution of the electrocardiography forward problem is calculated to simulate 3485 sets of 12-lead ECG with 697 nodes of ectopic pacing on a 3-D heart model, and we set the active and resting TMP to 15 and 20, respectively. The training, validation, and test sets were divided into an 8:1:1 ratio.

Figure 4 shows the visual predictive performance of the test data. Figure 4(a) displays the coordinates of labeled data and the predicted location, showing that the predicted coordinate points nearly covered the actual value coordinate points. Figure 4(b) shows each test point's prediction error, indicating that the localization error was mainly below 2 mm. Figure 4(c)~(e) respectively displays the error in the *x*, *y*, and *z*-axis. The localization errors for each axis are mostly concentrated in the 1.5 mm range. Quantitative error data are presented in Table 1, which shows that the mean localization error for all test data is 1.76 mm, and the mean localization error in the x-axis, y-axis, and z-axis, respectively, is 0.84 mm, 0.95 mm, and 0.91 mm.

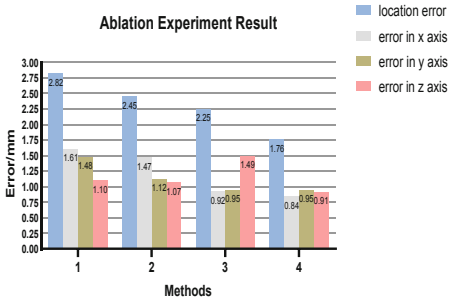

Fig. 5. Ablation study result.

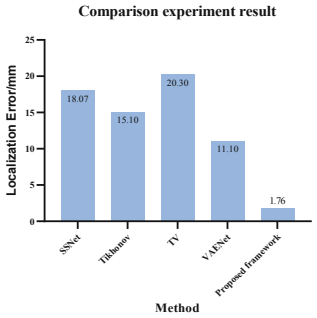

Fig. 6. Comparison experiment result.

Noise Robustness Experiment. In practical clinical measurement processes, noise is inevitably introduced; therefore, the prediction model needs robustness to noise.

In this section, we added Gaussian white noise with 5 dB, 15 dB, and 25 dB to the testing 12-lead data, respectively. From the results, we can see that network has certain robustness to different noise levels. At 25 dB noise, the prediction error is still only 2.97 mm. Though the error reached 10 mm at 5 dB, it is still within the clinically acceptable range considering the small signal-to-noise ratio (SNR).

Table 1. Localization error under different simulated levels of noise

SNR	Location Error/mm	Error in x axis/mm	Error in y axis/mm	Error in z axis/mm
Without noise	1.76	0.84	0.95	0.91
25DB	2.97	1.98	1.43	0.97
10DB	4.45	3.16	2.17	1.18
5DB	10.56	6.78	5.69	3.14

Ablation and Comparison Experiment. We compared four experiments' results to verify the effectiveness of the feature-fusion module and the local-global feature-extraction module: (1) CNN with feature fusion module: Remove the global feature extraction module; (2) GRU-Attention with feature fusion module: Remove the local feature extraction module; (3) Network without feature fusion module: only original ECGs sent to the local-global feature extraction module (4) Proposed Network. As the results shown in Fig. 5, single feature extraction networks, method (1) and method (2), are inferior compared to the proposed network. Method (3) only processes temporal signals and has lower error metrics than the above two methods, but the localization error is still 27.8% higher than that of the proposed network. These results demonstrate that combining temporal and frequency domain information of lead signals can improve the localization accuracy of ectopic pacing sites.

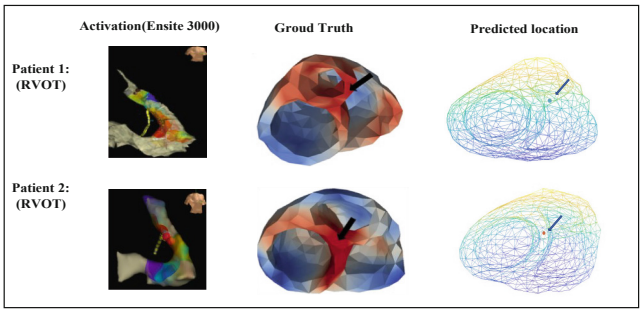

Fig. 7. Intuitive experiment result on clinical data.

To further demonstrate the capacity of the forward-solution aided method for ectopic pacing localization, we compared our method with the typical inverse solution methods shown in [25], which is also based on ECGSIM to simulate ectopic pacing data [26]. As is shown in Fig. 6, SSNet, Tikhonov, TV, and VAENet are inverse solution methods, and their localization errors are all >10 mm, while that of our method can reach within 2 mm.

Table 2. Quantitative results for the simulated patient pacing data

Patient ID	Location Error/mm	Error in x axis/mm	Error in y axis/mm	Error in z axis/mm
1	7.50	3.52	3.30	4.42
2	4.44	1.70	2.23	2.85
3	4.76	2.27	2.14	2.75
4	5.02	2.35	2.58	2.64
5	4.70	1.71	3.50	1.69

3.2 Clinical Data Experiment

Previous experiments confirmed the efficacy of our prediction framework on simulated cardiac data and the network's superior ability to learn the relationship between 12-lead data and ectopic pacing location. We now transfer the method to clinical PVC patients' data using the double variation cardiac source for calculating the forward solution of the specific heart-torso model and use them as the training data.

Table 2 shows the quantitative results of simulated patient pacing data. The location error ranges from 7.50 mm to 4.44 mm, with an average error of 5.28 mm. The network demonstrates precision and accuracy in localizing ectopic pacing in different cardiac models. Figure 7 shows the ectopic pacing location results for patient 1 and 2. The first column displays the activation of cardiomyocytes on the heart surface measured by the gold standard Ensite3000 system, with the red area indicating the earliest activation at the pacing location. The second column shows the activation mapping of the whole heart by Ensite3000. Our framework's localization results are consistent with the gold standard, with both pacing positions located in the posterior right ventricular outflow tract (RVOT), highlighting the practical significance of our study.

4 Discussion

In this paper, we developed a forward solution-aided deep learning framework for analyzing ectopic pacing from 12-lead ECG data. Only CT data is needed to establish the specific heart-torso model for simulating ECG data as the training set for the designed network. Time-frequency fusion module and local-global feature extraction module are the core component of the network. Experiments have shown that the framework performs well on both simulated and clinical data. In the future, to enhance the robustness of our proposed method, additional datasets and comprehensive simulations involving a broader spectrum of cardiac conditions should be incorporated. This paper primarily emphasizes the clinical potential of the proposed approach rather than extensively comparing actual and simulated data. The discrepancies observed could be attributed to noise, slight variations in physiological parameters, and consistent electrode placement in clinical practice. Further investigation and discussion on these aspects will be addressed in future research endeavors.

Acknowledgements. This work is supported in part by the National Natural Science Foundation of China (No: U1809204, 61525106) and by the Talent Program of Zhejiang Province (No: 2021R51004).

References

1. Soejima, K., Suzuki, M., Maisel, W.: Catheter ablation in patients with multiple and unstable ventricular tachycardias after myocardial infarction. Short ablation lines guided by reentry circuit isthmuses and sinus rhythm mapping. ACC Curr. J. Rev. **1**(11), 66–67 (2002)
2. Yamada, T.: Twelve-lead electrocardiographic localization of idiopathic premature ventricular contraction origins. J. Cardiovasc. Electrophysiol. **30**(11), 2603–2617 (2019)
3. Waxman, H.L., Josephson, M.E.: Ventricular activation during ventricular endocardial pacing: I. electrocardiographic patterns related to the site of pacing. Am. J. Cardiol. **50**(1), 1–10 (1982)
4. Miller, J.M., Marchlinski, F.E., Buxton, A.E., Josephson, M.E.: Relationship between the 12-lead electrocardiogram during ventricular tachycardia and endocardial site of origin in patients with coronary artery disease. Circulation **77**(4), 759–766 (1988)
5. Ito, S., et al.: Development and validation of an ECG algorithm for identifying the optimal ablation site for idiopathic ventricular outflow tract tachycardia. J. Cardiovasc. Electrophysiol. **14**(12), 1280–1286 (2003)
6. Feng, Q., Hu, H., Liu, H.: Multi-level information for non-invasive identification of exit site of ventricular tachycardia. In: 2020 Computing in Cardiology, pp. 1–4. IEEE (2020)
7. Gyawali, P.K., Horacek, B.M., Sapp, J.L., Wang, L.: Learning disentangled representation from 12-lead electrograms: application in localizing the origin of ventricular tachycardia. arXiv preprint arXiv:1808.01524 (2018)
8. Yang, T., Yu, L., Jin, Q., Wu, L., He, B.: Localization of origins of premature ventricular contraction by means of convolutional neural network from 12-lead ECG. IEEE Trans. Biomed. Eng. **65**(7), 1662–1671 (2017)
9. Monaci, S., et al.: Non-invasive localization of post-infarct ventricular tachycardia exit sites to guide ablation planning: a computational deep learning platform utilizing the 12-lead electrocardiogram and intracardiac electrograms from implanted devices. Europace **25**(2), 469–477 (2023)
10. Strodthoff, N., Strodthoff, C.: Detecting and interpreting myocardial infarction using fully convolutional neural networks. Physiol. Meas. **40**(1), 015001 (2019)
11. Sun, C., Shrivastava, A., Singh, S., Gupta, A.: Revisiting unreasonable effectiveness of data in deep learning era. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 843–852 (2017)
12. Gonzales, M.J., Vincent, K.P., Rappel, W.J., Narayan, S.M., McCulloch, A.D.: Structural contributions to fibrillatory rotors in a patient-derived computational model of the atria. Europace **16**(suppl_4), iv3–iv10 (2014)
13. Villongco, C.T., Krummen, D.E., Stark, P., Omens, J.H., McCulloch, A.D.: Patient-specific modeling of ventricular activation pattern using surface ECG-derived vectorcardiogram in bundle branch block. Prog. Biophys. Mol. Biol. **115**(2–3), 305–313 (2014)
14. Krummen, D.E., et al.: Forward-solution noninvasive computational arrhythmia mapping: the VMAP study. Circulat. Arrhythmia Electrophysiol. **15**(9), e010857 (2022)
15. Bear, L.R., et al.: Forward problem of electrocardiography: is it solved? Circulat. Arrhythmia Electrophysiol. **8**(3), 677–684 (2015)
16. Langrill, D.M., Roth, B.J.: The effect of plunge electrodes during electrical stimulation of cardiac tissue. IEEE Trans. Biomed. Eng. **48**(10), 1207–1211 (2001)

17. Aliev, R.R., Panfilov, A.V.: A simple two-variable model of cardiac excitation. Chaos, Solitons Fract. **7**(3), 293–301 (1996)
18. Minami, K.I., Nakajima, H., Toyoshima, T.: Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Trans. Biomed. Eng. **46**(2), 179–185 (1999)
19. Lee-Thorp, J., Ainslie, J., Eckstein, I., Ontanon, S.: Fnet: mixing tokens with Fourier transforms. arXiv preprint arXiv:2105.03824 (2021)
20. Hannun, A.Y., et al.: Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat. Med. **25**(1), 65–69 (2019)
21. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 (2014)
22. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30 (2017)
23. Peng, Z., et al.: Local features coupling global representations for visual recognition. In: CVF International Conference on Computer Vision, ICCV, pp. 357–366. IEEE (2021)
24. Van Oosterom, A., Oostendorp, T.: ECGsim: an interactive tool for studying the genesis of QRST waveforms. Heart **90**(2), 165–168 (2004)
25. Huang, X., Yu, C., Liu, H.: Physiological model based deep learning framework for cardiac TMP recovery. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) MICCAI 2022. LNCS, vol. 13432, pp. 433–443. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16434-7_42
26. Ramanathan, C., Ghanem, R.N., Jia, P., Ryu, K., Rudy, Y.: Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia. Nat. Med. **10**(4), 422–428 (2004)