

Deep learning methods for screening patients' S-ICD implantation eligibility

Anthony J. Dunn^a, Mohamed H. ElRefai^b, Paul R. Roberts^b, Stefano Coniglio^a,
Benedict M. Wiles^c, Alain B. Zemkoho^{a,*}

^a University of Southampton, School of Mathematical Sciences, United Kingdom

^b University Hospital of Southampton, United Kingdom

^c St George's University Hospitals NHS Foundation Trust, United Kingdom

ARTICLE INFO

Keywords:

Subcutaneous implantable cardioverter-defibrillators
Sudden cardiac death
Ventricular arrhythmia
Electrocardiogram
Deep learning
Convolutional neural networks
Phase space reconstruction
Patient screening

ABSTRACT

Subcutaneous Implantable Cardioverter-Defibrillators (S-ICDs) are used for prevention of sudden cardiac death triggered by ventricular arrhythmias. T Wave Over Sensing (TWOS) is an inherent risk with S-ICDs which can lead to inappropriate shocks. A major predictor of TWOS is a high T:R ratio (the ratio between the amplitudes of the T and R waves). Currently, patients' Electrocardiograms (ECGs) are screened over 10 s to measure the T:R ratio to determine the patients' eligibility for S-ICD implantation. Due to temporal variations in the T:R ratio, 10 s is not a long enough window to reliably determine the normal values of a patient's T:R ratio. In this paper, we develop a convolutional neural network (CNN) based model utilising phase space reconstruction matrices to predict T:R ratios from 10-second ECG segments without explicitly locating the R or T waves, thus avoiding the issue of TWOS. This tool can be used to automatically screen patients over a much longer period and provide an in-depth description of the behavior of the T:R ratio over that period. The tool can also enable much more reliable and descriptive screenings to better assess patients' eligibility for S-ICD implantation.

1. Introduction

Sudden Cardiac Death (SCD) is one of the leading causes of death in the modern world. Most of these deaths can be attributed to Ventricular Arrhythmias (VA) [1]. The key to survival in patients affected by VA is adequate Cardiopulmonary Resuscitation (CPR) and early defibrillation [11]. Medical guidelines recommend the use of Implantable Cardioverter-Defibrillators (ICDs) for prevention of SCD triggered by VA in high risk populations [17,23]. Conventional transvenous ICDs (TV-ICDs) consist of a *can* and transvenous leads implanted into the right ventricle to treat the arrhythmia by delivering a voltage shock. TV-ICDs are associated with the risk of complications with potentially fatal consequences.

The Subcutaneous ICD (S-ICD), which comprises an electrically active can and a single subcutaneous lead (see Fig. 1a) was designed to avoid complications of the TV-ICD by utilising a totally avascular approach. The sensing mechanism of the S-ICD has been shown to be equally effective to that of the TV-ICD [4] and demonstrated less incidence of device-related complications when compared with conventional ICDs [15]. However, a consequence of the algorithm used by the S-ICD to detect VA is an inherent risk of T Wave Over Sensing (TWOS),

whereby the T wave, one of the 5 main waveforms of the PQRST complex (the electrocardiogram (ECG) of a single heartbeat) which follows a QRS complex (comprised of the Q, R and S waves of a PQRST complex) is misinterpreted as a second R wave, which can lead to inappropriate shocks. Inappropriate shocks are associated with increased morbidity and mortality [28].

Not all patients are eligible for S-ICD therapy and eligibility is identified during a pre-implant screening process that is undertaken in potential S-ICD recipients. Surface ECGs are used as a surrogate marker of future S-ICD vectors (as shown in Fig. 1a) to non-invasively assess S-ICD eligibility. These ECG recordings are assessed. Vectors with lower T:R ratios (the ratio between the amplitude of the T wave and that of the R wave) are more likely to pass the screening, while patients with an ECG morphology that does not meet this screening criteria are deemed at high risk of TWOS and are ineligible for an S-ICD. Despite the current screening practice, the most common cause of inappropriate shocks in patients implanted with S-ICDs remains TWOS [15]. The T:R ratio—a major predictor of S-ICD eligibility—is not fixed in the same individual because of frequently observed temporal variations in the amplitudes of the R and T waves which are influenced by multiple factors [2,9,21,22]. In patients with S-ICDs, variations in the T:R ratio often go undetected as

* Corresponding author.

E-mail address: a.b.zemkoho@soton.ac.uk (A.B. Zemkoho).

<https://doi.org/10.1016/j.artmed.2021.102139>

Received 9 March 2021; Received in revised form 20 June 2021; Accepted 3 August 2021

Available online 9 August 2021

0933-3657/© 2021 University of Southampton. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

“silent” episodes of TWOS in spite of carrying a considerable risk of leading to the development of clinically relevant oversensing, which manifests in inappropriate shocks.

Under the current screening process, ECG electrodes are placed on the chest wall using the anatomical landmarks which correspond to the location of the sensing electrodes of the potential future S-ICD implantation shown in Fig. 1b. A short three lead ECG recording (several PQRST complexes), corresponding to the three vectors utilised by the S-ICD, is evaluated using the manual S-ICD screening tool shown in Fig. 2. Henceforth, when discussing leads, we refer only to the three leads corresponding to the three vectors used by the S-ICD. The patient is deemed eligible for S-ICD implantation if, for at least one of the leads (representing one of the S-ICD vectors), the patient's QRS complex (the PQRST complex excluding the P wave) sits entirely within the boundary of the template. These templates correspond roughly to a maximum acceptable T:R ratio of 1:3. As mentioned previously, the T:R ratio could fluctuate, as the amplitudes of both R waves and T waves may vary according to other factors (e.g. electrolyte levels). Due to the short duration (several PQRST complexes) of the current screening, it is possible that a patient with a typically high T:R ratio could pass this screening and likewise a patient with a typically low T:R ratio could fail it.

1.1. Related works

Machine learning methods have been used for ECG analysis in a variety of applications. There has been a wealth of work in the classification of various Cardiovascular Diseases (CVDs) from ECG data [8,14,25–27,29–31,33]. Other applications of machine learning in ECG analysis include detecting seizures and heart attacks [18,19], predicting patients' blood pressure [24], detecting a patients facial expressions [6] and analysis of ECG of the brain has been used for creating brain computer interfaces (BCI) capable of detecting which body part the subject was completing a task with [5,7].

A popular technique for preprocessing ECG data is to create its Phase Space Reconstruction (PSR) matrix. Typically, features are extracted from the PSR matrix of ECG data which can then be used as inputs for a classification model. Box counting as well as column and row statistics are features often extracted from the PSR matrix of ECG data. These methods have been used in the prediction of CVD [26,27,29,30], creating BCIs [5,7] and detecting facial expressions [6]. These approaches are all centred around manually selecting features to extract from the PSR matrix. Our proposed method diverges from this by using

the whole PSR matrix as the input to a model which is itself capable of extracting features. Convolutional neural networks (CNN) are an example of this. During training, convolutional layers learn to extract features of the input image which are most impactful in accurately determining the model's output. As such, these models can replace the need for time consuming feature extraction and can arrive at much more descriptive features. CNNs have been used in ECG analysis for classifying heart attacks [19,42], CVDs [38], atrial fibrillation [8,25,41] and other arrhythmias and rhythm abnormalities [14,31–33,40] as well as for predicting blood pressure [24], locating cardiac accessory pathways [35], recognising early signs of heart failure with reduced ejection fraction [37] and detecting digoxin toxicity [34], electrolyte imbalance [36] and anaemia [39]. CNNs are typically used for the classification of 2D images and, as such, use 2D filters for feature extraction. All of these methods use the filtered ECG as the model input, where each lead corresponds to a single 1D signal. By creating the PSR matrix for each lead, a 2D input is created from each lead.

1.2. Contributions and outline of the paper

The work in this paper is a proof of concept designed to demonstrate how the adoption of AI in developing an automated screening tool for performing prolonged ECG screenings can allow clinicians to better scrutinise the S-ICD implantation eligibility of patients from a range of patient groups. Our proposed method uses each lead of a Holter recording to assess if the corresponding vector has a low enough T:R ratio to be a viable candidate for S-ICD implementation. Holter monitors are frequently used in the risk stratification assessment for ICDs. In such cases, no additional data collection is required for our proposed screening tool. In cases where Holter monitors are not already being used, the test is considered to be inexpensive, non-invasive and routine.

In order to ensure that our screening tool is capable of accurately predicting T:R ratios in patients suffering from heart conditions which make them likely candidates for S-ICD implantation, we require ECG data from those patient groups for the training and evaluation of our models. To this end, we have collected ECG data from a range of patient groups which potential S-ICD implantation candidates are likely to belong to. Further details of this data set are given in Section 3.1. This new data set, collected for the purpose of this project and related analysis, enables us to train models with a high degree of robustness across a range of different patient groups. This automated screening allows for a drastically extended screening window which, in turn, can enable clinicians to better scrutinise patients' eligibility for S-ICD implantation.

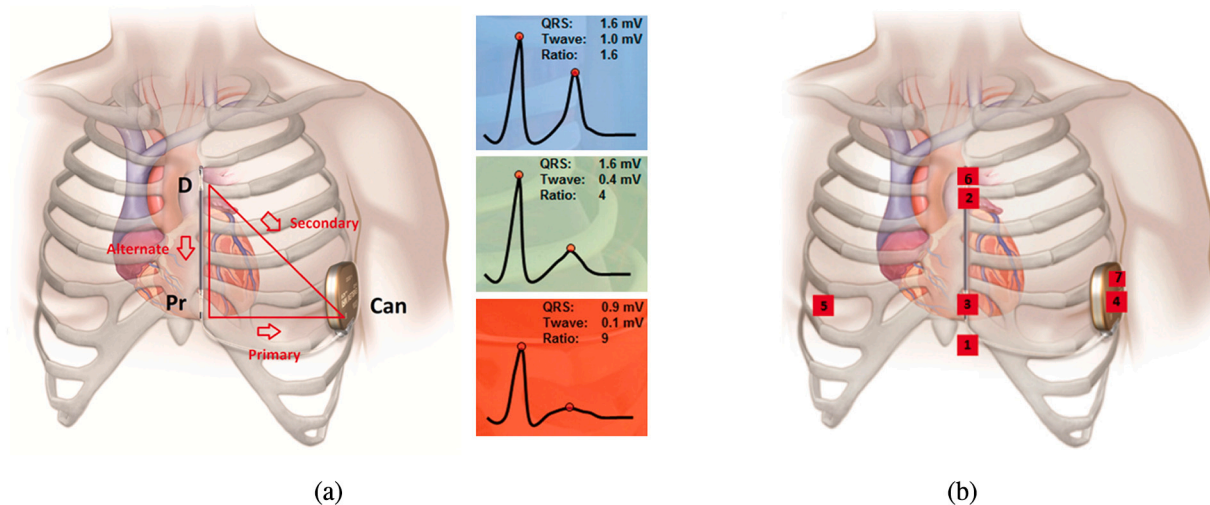


Fig. 1. (a) S-ICD sensing electrodes and vectors between them. An implanted S-ICD with underlying anatomical features showing the location of the can (pulse generator), the proximal (Pr) and distal (D) sensing electrodes and the shocking coil (located between the electrodes). (b) Holter recorder surface ECG positions. Images (prior to annotation) © Boston Scientific Corporation or its affiliates. Reproduced with permission.

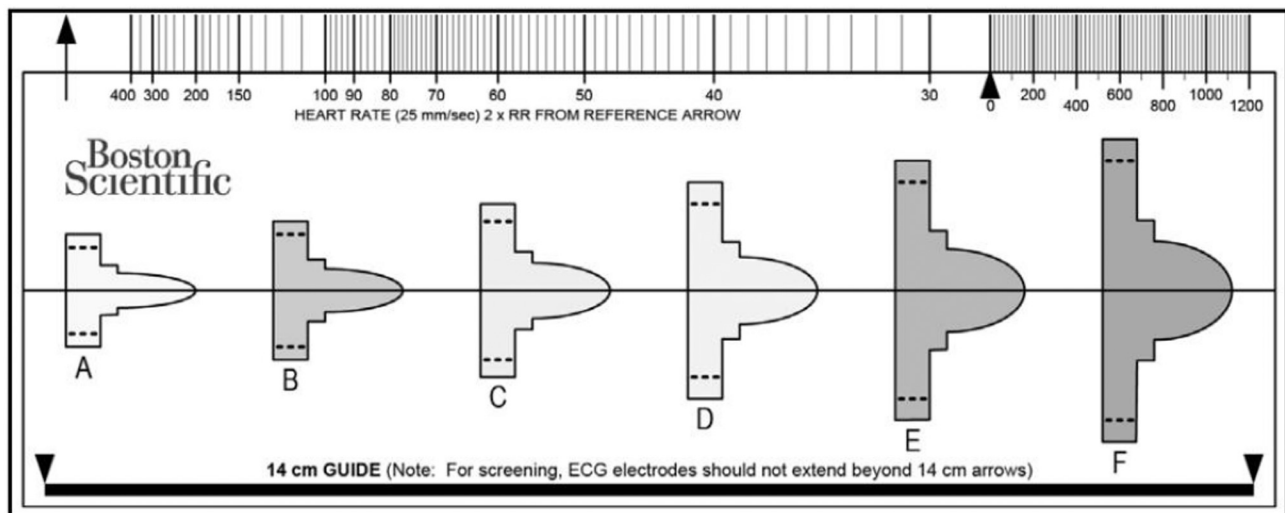


Fig. 2. S-ICD screening tool. The recorded QRST morphology in every vector is then compared to the acceptable templates. The template is aligned to the isoelectric line of the ECG, and the QRST complexes are viewed through the appropriately sized template. The R wave peak of the ECG must be placed within either hashed box (positive or negative) of any template. A vector passes screening if the remainder of the QRST complex sits entirely within the boundary of the template. (this manual screening method is now largely replaced by the manufacturer with an automatic screening tool following the same principals).

This tool also has the potential to give insight into how the behavior of patients' T/R ratios vary between patient groups.

To be precise, in this paper, we propose an accurate, reliable and robust method that utilises the concept of prolonged screening for S-ICD eligibility to better scrutinise the selection criteria in an attempt to find a cohort of patients with low probability of TWOS and inappropriate shocks. Central to this prolonged screening is a deep learning based method for predicting the T/R ratio—the main determinant of S-ICD eligibility—of a 10-second segment of a single lead ECG. Solving this regression problem is at the core of this paper. The aim of our proposed screening method is to use the model developed in this paper to analyse the T/R ratio of each single lead 10-second ECG segment within a three lead 24-hour ECG recording to determine if any of the three leads have a suitably low risk of TWOS. For a given lead, should a patient have a T/R ratio above 1:3 for a continuous period of at least 20 s (the duration of TWOS at which the S-ICD would deliver an inappropriate shock), that lead would fail the screening. If all three leads fail the screening, the patient would be deemed not eligible for an S-ICD. Temporal variations in the T/R ratio make the current 10-second screening process unreliable. Our proposed method allows for a much more robust screening, as it would allow for the analysis of variations in the T/R ratio over a 24-hour period. [Section 3.4](#) gives an example of how the model for predicting T/R ratio from ECG segments developed in this paper could be used within a screening tool to access patients' eligibility for S-ICD implantation.

In contrast with most of the literature, in which, with only a few exceptions (see, e.g., [\[3\]](#)), CNNs are used for image classification rather than for regression, we propose using 2D PSR images of the filtered ECG signals as input to CNNs with 2D filters. In particular, rather than extracting features from PSR images which are then used as model inputs, as typically done in previous works, we use the entire PSR image as input to our models. While CNNs have been used to analyse ECGs, this has been only done using data in which each lead corresponds to a single 1D-ECG signal. To the best of our knowledge, this is the first work in ECG analysis where a PSR matrix is generated for each lead that serves as input to a CNN based deep-learning model. Further separating this work from the literature is our focus on predicting T/R ratios. This is a very unique goal in ECG analysis, where various other ECG factors are identified and examined. The intuitive approach to calculating T/R ratios by detecting and measuring the amplitude of the R and T waves individually is vulnerable to the same TWOS problem that the algorithm

used within the S-ICDs suffers from. Our novel methodology for predicting the T/R ratio does not share this vulnerability as it considers multiple PQRST complexes simultaneously.

In [Section 2](#), we outline the preprocessing techniques used to create PSR images of the filtered 10-second ECG segments, detail our selection of deep learning model architectures and describe the process we use to train and evaluate these models. [Section 3](#), illustrates the results of an extensive set of experiments, demonstrating the capabilities of our models to predict T/R ratios from PSR images as well as providing examples of how such models could be integrated into a clinical tool. In [Section 4](#), we conclude the paper and outline directions for future work.

2. Methodology

We propose a new screening process where a Holter®—a portable ECG recording device—is used to record 24 h of ECG data from three leads corresponding to the three vectors utilised by the S-ICD. This data is then split into 10-second segments and an artificial neural-network based model is used to predict the T/R ratio for each 10-second segment. The cardiologist would then be able to review the behavior of the patient's T/R ratio over the 24-hour period and evaluate their eligibility. We rely on a number of filtering techniques for removing noise from the ECG signals. We then use PSR, a popular technique in waveform analysis, to convert the ECG signal into an image of its PSR matrix with which we then train a Convolutional Neural Network (CNN) to predict the T/R ratio from these images.

The most straightforward approach to measuring the T/R ratio is to locate the R and T waves and measure their amplitudes. However, by explicitly detecting and measuring the peaks of the R and T waves, we run the risk of TWOS when the characteristics of the T wave becomes similar to those of the R wave. To avoid this, our model aims to predict the T/R ratio without ever explicitly locating or measuring the R or T waves.

2.1. Preprocessing

Preprocessing involves filtering the ECG data, performing transformations to emulate the methods used within S-ICDs and creating images by plotting the PSR of the data. [Fig. 2](#) gives an overview of the preprocessing techniques used to prepare our data for the training of the regression models.

2.1.1. Data structure

For this paper, we consider data in the form of 10-second segments of single lead ECG recordings from Holter leads corresponding to the vectors used by the S-ICD. These 10-second ECG segments are annotated with the positions of the peaks of the T and R waves occurring in this period.

From these annotations, we are able to calculate the dependant variable for our regression problem: the average T:R ratio. As mentioned previously, the T:R ratio of a single PQRS complex is simply the ratio between the amplitudes of the T and R waves. For the purposes of this paper, we will consider this ratio in the form of a fraction. The average T:R ratio for a 10-second ECG segment $\{x(1), \dots, x(10 \cdot f)\}$, where f is the sampling frequency of the signal, with T-peak annotations at indexes $\{T_1, \dots, T_n\}$ and R-peak annotations at indexes $\{R_1, \dots, R_n\}$, is given by

$$\frac{\sum_{i=1}^n x(T_i)}{\sum_{i=1}^n x(R_i)}.$$

When the T wave has negative amplitude, this fraction is negative. From a clinical perspective, we are only interested in the magnitude of the T:R ratio. However, from a signal processing perspective, there is a great difference between a PQRS complex with a negative T wave and one with a positive T wave. For this reason, we will preserve the sign of the T:R ratio, as the loss of information resulting from considering only the magnitude of the ratio would lead to a reduction in the accuracy of our models. Having built a model capable of predicting T:R ratios from 10-second ECG segments, we take the magnitude of this model's outputs for use in a clinical tool. Our choice to consider the T:R ratio rather than the R:T ratio, which is more common in the literature, is well motivated. As the T wave amplitude approaches 0, very small changes in the T wave amplitude can result in extreme changes in the R:T ratio. This massive variation in R:T ratio for very similar ECG signals makes it inappropriate for use as a label in our regression problem. Typically, the R wave is of greater amplitude than the T wave. Because of this, the T:R ratios of a set of ECG segments are well distributed between 0 and 1. For this reason, we use the T:R ratio as our dependent variable in our regression problem. If the situation requires it, the R:T ratio can of course be derived from our model by simply taking the inverse of the model's output.

2.1.2. Filtering

Fig. 3 gives an overview of the filtering methods we will use to remove noise from our ECG signal [20].

Firstly, baseline drift correction is implemented using one-dimensional Discrete Wavelet Transformation (DWT). The ECG signal is decomposed at 9 levels, using the Daubechies 8 (db8) wavelet, then reconstructed using only level 9 coefficients. This reconstructed signal is the low frequency component for the ECG signal which is assumed to be the drifting baseline. Subtracting this from the original signal leaves us with an ECG signal with a stable baseline of value 0.

Adaptive bandstop filtering is used to suppress power-line noise with a frequency of 50 Hz, while a lowpass filter is used to remove the remaining high-frequency noise. Having applied these filters, the locations of the R and T peak markers may no longer be correct. To account for this, a small region around the R peak is searched for a maximum and this maximum is taken as the new R peak. Similarly, a small region around the T peak is searched and the maximum or minimum value in this region is taken as the new T peak if the T peak is positive or negative, respectively.

Fig. 4 gives an example of an unfiltered 10-second ECG segment as well as the same segment after filtering has been applied.

2.1.3. Negative QRS peak flipping

While R waves are strictly positive, a PQRS complex with a small R wave could be prone to the T wave being labeled as an R wave, leading to double counting. For this reason, when implementing R peak detection, it may be reasonable to look for both positive and negative peaks and simply flip the ECG when the peak detected is negative. This approach is followed by the algorithm used in S-ICDs. Fig. 5 shows some possible QRS complexes where the amplitude of the Q or the S wave is greater than that of the R wave.

In cases where the peak of greatest magnitude of a QRS complex is negative and it is labeled as the R peak, we are, in actuality, labeling either the Q or the S peak. In this case, the ECG signal of the entire PQRS complex is flipped (multiplied by -1) and the R peak is assigned as the previously negative peak. To implement this on our data, we search within a region of each of the R peak markers for a negative peak of greater magnitude. If one is found, it is assigned as the new R peak and the signal is flipped. Fig. 6a gives an example of a filtered 10-second ECG segment while Fig. 6b shows the same segment after the negative QRS peaks have been detected, the signal has been flipped and the peaks have been reassigned.

2.1.4. Phase space reconstruction

Phase space reconstruction is a technique for representing non-linear characteristics of a time series set of data using delay maps. For a given time series $x(1), x(2), \dots, x(n)$, the time lagged phase space vectors are given by

$$X(i) = \{x(i), x(i + \tau), \dots, x(i + (d - 1)\tau)\} \\ \text{for } i \in \{1, \dots, n - (d - 1)\tau\},$$

where τ is the time delay between points in the series and d is the number of dimensions of the phase space which we are mapping this data to. While using high dimensional PSR has given good results in the field of BCI [5,7], the majority of the work in ECG analysis uses two-dimensional PSR, corresponding to $d = 2$ [16,26,27,30]. For this reason, we transform our time series data into a normalized matrix of two-dimensional phase space vectors

$$B = \begin{bmatrix} x(1)/q & x(1 + \tau)/q \\ x(2)/q & x(2 + \tau)/q \\ \vdots & \vdots \\ x(n - \tau)/q & x(n)/q \end{bmatrix},$$

where $q = \max \{|x(i)| : i=1, \dots, n\}$.

In two dimensions, the phase space plot corresponding to this matrix ranges from -1 to $+1$ on each axis. This area can now be divided into N^2 small square areas, $g(i, j)$, of size $R \times R$, for $i, j=1, \dots, N$, where R is given by $R = 2/N$ and N is an integer number. The phase space matrix C (of dimension $N \times N$) is constructed with each of its elements $c(i, j)$ equal to the number of phase space vectors in B which fall within the square area $g(i, j)$. C is then normalized to give P , wherein each element $p(i, j)$ gives the probability of a phase point falling within $g(i, j)$. Formally,

$$P = \frac{1}{M} C, \quad M = \sum_{i=1}^N \sum_{j=1}^N c(i, j)$$

A typical approach is to extract features from these PSR by either box-counting [30], calculating the spatial filling index [16] or calculating statistics of the distributions of values within each column of C [27]. These features are then used as model inputs for classifying various different categories of ECG. In this paper, we use tools typically used for



Fig. 3. Flowchart of the methods in our data preprocessing step.

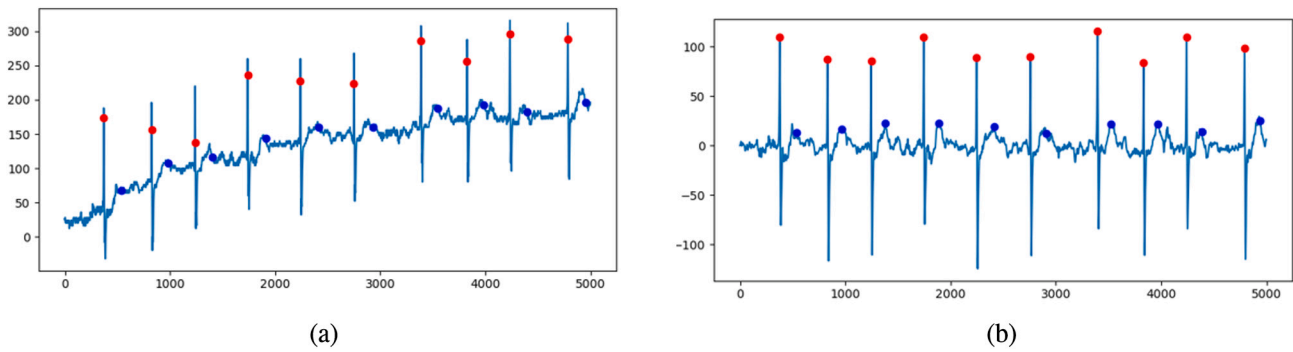


Fig. 4. (a) Example of a 10-second ECG segment pre-filtering. R and T peaks shown in red and blue, respectively. (b) The same 10-second ECG segment post-filtering. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

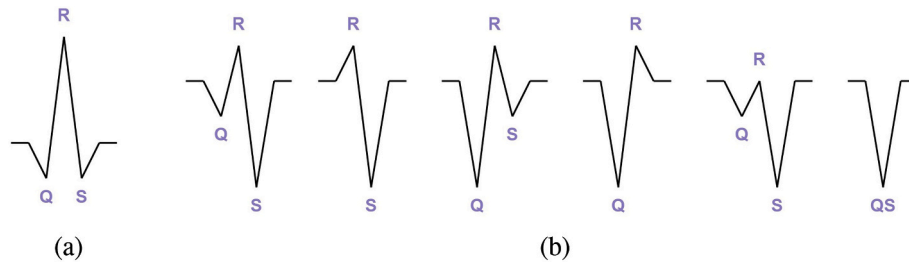


Fig. 5. (a) An example of a normal QRS complex. (b) Examples of QRS complexes in which the R wave is not the wave of greatest amplitude.

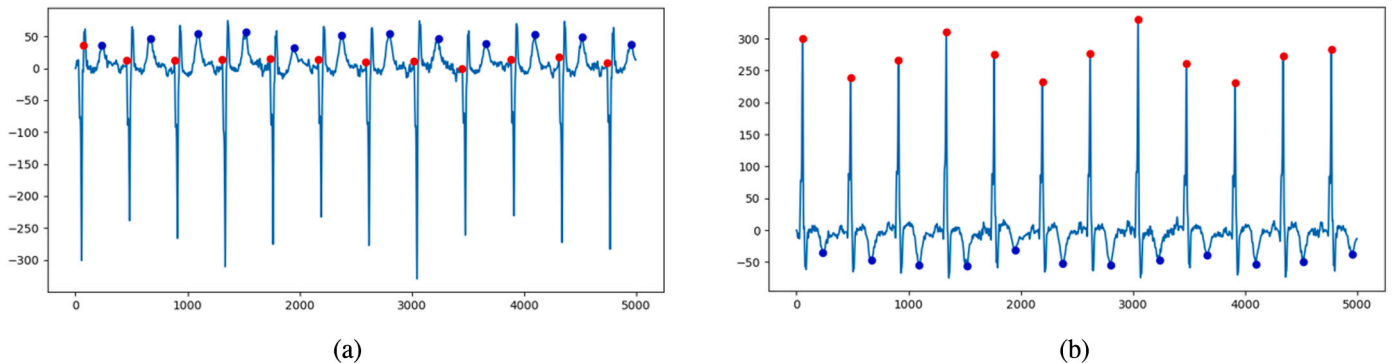


Fig. 6. (a) Example of a filtered 10-second ECG segment. R and T peaks shown in red and blue, respectively. (b) The same 10-second ECG segment post-flipping. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

computer vision to allow us to use the $N \times N$ pixel images derived from P as the input for our model.

Fig. 7 gives an example of a PSR image with $N = 32$ as well as a darkened version of the same image for readability. As one can see in the darkened image, there are some connected bands of higher probability. With values of N greater than 32, these bands become disconnected as no phase space vectors land within that portion of the grid.

2.2. MLP and CNN for regression

In this section, we discuss the architecture of the models used to predict T:R ratios from 32×32 pixel PSR images. The general structure of each model is laid out in Fig. 8. Each model is made up of N feature-extraction blocks, followed by a regression block.

The regression block is used by all models to derive the T:R ratio from the extracted features. The outputs from the preceding feature-extraction blocks are flattened to a 1D vector and fed into a series of fully connected (dense) layers of neurons to arrive at the final regression output: the T:R ratio. Table 1 gives an overview of the layers comprising

this block. We use batch normalisation for regularisation as it has been shown to be superior to dropout for use in CNNs [10]. We perform batch normalisation before applying the activation function as proposed in the original paper on batch normalisation [13].

The first and most basic of our feature-extraction blocks is the MLP feature-extraction block shown in Table 2. It is comprised of a single layer of fully connected neurons followed by a batch normalisation and activation layer. The input and output of these blocks are 1D. As such, when using these blocks, we flatten the images before the first feature-extraction block rather than before the regression block.

The basic CNN feature-extraction blocks utilise convolutional layers, which exploit the 2D structure of the PSR images, as opposed to fully connected layers. These layers are followed by the batch normalisation and activation layers mentioned previously and finally a maximum pooling layer to reduce the size of the output images. As shown in Table 3, the output of each layer and the kernel size of the convolutional layer depends on the number of feature-extraction blocks.

The complex CNN feature-extraction block is based on the basic CNN feature-extraction block. The convolutional layer is replaced by a pair of

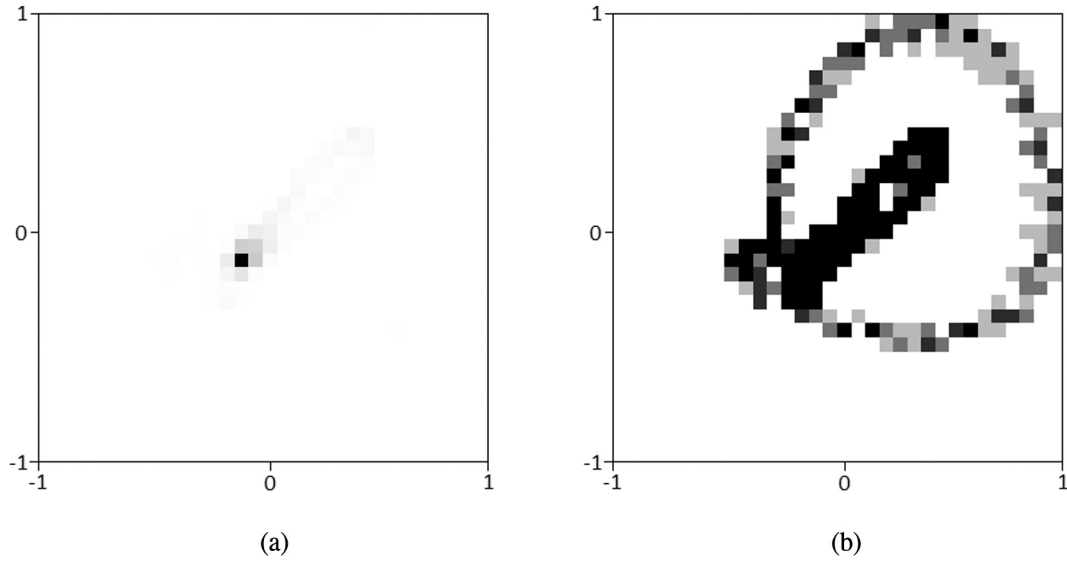


Fig. 7. (a) Image of the PSR matrix formed from a 10-second segment ECG signal with $N = 32$ (b) Darkened image of the same PSR matrix.

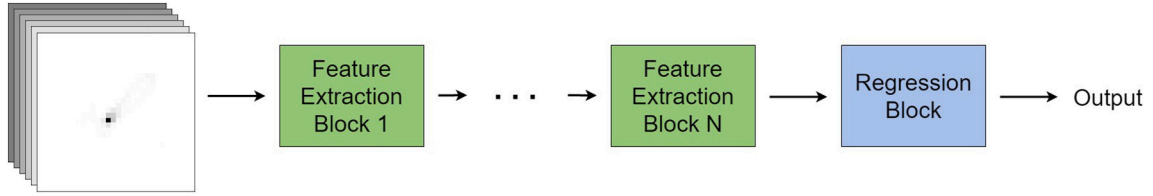


Fig. 8. Diagram of a model comprised of N feature-extraction blocks.

Table 1

Regression block.

Layer	Type	Output size
1	Dense	256
2	BatchNorm	256
3	Activation(ReLu)	256
4	Dense	64
5	BatchNorm	64
6	Activation(ReLu)	64
7	Dense	1

Table 2

MLP feature extraction block n .

Layer	Type	Output size
1	Dense	1024
2	BatchNorm	1024
3	Activation(ReLu)	1024

Table 3

Basic CNN feature extraction block n .

Layer	Type	Output size	Kernel size	Stride
1	Convolutional	$2^{n+4} \times 2^{6-n} \times 2^{6-n}$	$(7-n) \times (7-n)$	1
2	BatchNorm	$2^{n+4} \times 2^{6-n} \times 2^{6-n}$		
3	Activation(ReLu)	$2^{n+4} \times 2^{6-n} \times 2^{6-n}$		
4	MaxPooling	$2^{n+4} \times 2^{5-n} \times 2^{5-n}$	2×2	2

convolutional layers with smaller kernels. The maximum pooling layer is replaced by an additional convolutional layer with stride equal to 2, which reduces the size of the output whilst continuing to extract

features. Finally, *addition skip connections*, as popularised in ResNet [12] and utilised in other works using CNNs for ECG analysis [36,40,41], are added over the first two convolutional layers in order to speed up training. Table 4 gives an overview of the layers of this block.

The Deep CNN feature-extraction block is very similar to the complex CNN feature-extraction block. Before the convolutional layer with stride equal to 2 which is used for pooling, we include an additional pair of convolutional layers with smaller kernels as well as a second skip connection. Table 5 gives an overview of the layers of this block. When referring to the model, we give the type of feature-extraction block used followed by the number of feature-extraction blocks. For example, a model comprised of 5 MLP feature-extraction blocks followed by the regression block would be referred to as MLP5, while a model comprised of 3 Complex CNN feature-extraction blocks followed by the regression block would be referred to as Complex CNN3.

2.3. Model enhancement

To improve the performance of our models described in the previous

Table 4

Complex CNN feature extraction block n .

Layer	Type	Output size	Kernel size	Stride
1	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
2	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
3	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
4	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
5	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
6	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
7	Skip(Input)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
8	Convolutional	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$	$(7-n) \times (7-n)$	2
9	BatchNorm	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$		
10	Activation(ReLu)	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$		

Table 5
Deep CNN Feature Extraction Block n.

Layer	Type	Output size	Kernel size	Stride
1	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
2	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
3	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
4	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
5	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
6	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
7	Skip 1(Input)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
8	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
9	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
10	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
11	Convolutional	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$	$(6-n) \times (6-n)$	1
12	BatchNorm	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
13	Activation(ReLu)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
14	Skip 2(Skip 1)	$2^{n+2} \times 2^{6-n} \times 2^{6-n}$		
15	Convolutional	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$	$(7-n) \times (7-n)$	2
16	BatchNorm	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$		
17	Activation(ReLu)	$2^{n+3} \times 2^{5-n} \times 2^{5-n}$		

subsection, we consider the following two techniques.

Firstly, we consider image augmentation, which can strengthen the training of models by randomly introducing small distortions to the training data, thus resulting in models which are more robust to distortions which may occur naturally in our unaltered data. While there are many image augmentation strategies, only a small number of them are appropriate for the sort of images we are handling. We test 4 image augmentation strategies: shifting, zooming, rotating and shearing. Fig. 9 gives an example of how each of these augmentations could affect a PSR image. When implementing image augmentation, a model may be trained on data which, at each mini-batch, is either shifted randomly within a range of $[1 - \phi/100, 1 + \phi/100]$, zoomed randomly within a range of $[1 - \phi/100, 1 + \phi/100]$, rotated randomly by an angle within a range of $[-\phi^\circ, +\phi^\circ]$ or sheared randomly with a shear angle within a range of $[-\phi^\circ, +\phi^\circ]$. We denote the augmentation as either tiny ($\phi=2.5$), small ($\phi=5$) or moderate ($\phi=10$). For instance, MLP5 Small Rotation would refer to a model consisting of 5 MLP feature-extraction blocks followed by a regression block which has been trained on images which have been, at each mini-batch, rotated randomly by an angle within a range of $[-5^\circ, 5^\circ]$.

Secondly, we consider ensemble models. In order to avoid overfitting, 20% of our training data is randomly reserved for validation. Models are trained until their accuracy on the validation set is no longer increasing. This means that 20% of our original training data is not being used for training. To avoid this, we can iteratively reserve a different 20% of our training data for validation, training 5 models with the same architecture but which have all used a different set of data for validation. In doing so, while some data may be used for validation by one model, it will be used for training by the remaining 4. By taking the average of these models' predictions, we aim to obtain a higher prediction accuracy than would be achieved by any single model. When referring to an ensemble model comprised on 5 sub-models, we simply append the word ensemble onto the name of the sub-model.

2.4. Model evaluation

To evaluate these models, we use 10-fold cross validation. At each iteration, 10% of the data is used as the testing set for ultimately evaluating the models we train, 20% of the remaining data forms the validation set to help avoid overfitting during training and the remaining data is used to form the training set. We assess the accuracy of each model using mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE) and STD of errors, respectively defined as

$$\begin{aligned}
 \text{MSE} &= \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \\
 \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}, \\
 \text{MAE} &= \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}, \\
 \text{STD of errors} &= \sqrt{\frac{\sum_{i=1}^n \left(y_i - \hat{y}_i - \frac{\sum_{j=1}^n y_j - \sum_{j=1}^n \hat{y}_j}{n} \right)^2}{n-1}},
 \end{aligned}$$

where y are the true T:R ratios, \hat{y} are the predicted T:R ratios and n is the number of PSR images.

3. Experiments

3.1. SGH-ECG and ECG-ID data sets

Our aim in this study is to show that sophisticated ML tools can be used to better scrutinise patients' eligibility for S-ICD implantation. We achieve this by constructing deep learning models capable of accurately predicting T:R ratios of patients belonging to a range of patient groups; namely, patients suffering from congestive heart failure, underlying congenital heart disease and "hypertrophic cardiomyopathy", as well as patients with structurally normal hearts. In order to develop such models and ensure their accuracy across all of these patient groups, we have collected our own data set. We collected this data set at Southampton General Hospital as part of a study on S-ICD eligibility and as such we refer to it as the Southampton General Hospital ECG (SGH-ECG) data set. The SGH-ECG data set consists of 390 10-second ECG segments, sampled at 500 Hz, annotated with R and T peaks. These signals were obtained at random intervals from the 24-hour ECG recordings of 18 different participants and were collected as a part of a clinical research study – HEART TWO¹ conducted by the Cardiac rhythm management research department at the University Hospital of Southampton. A detailed break down of the participants demographics and heart conditions can be found in Table 6. The participants ages range from 20 to 80 years old with a mean of 53.16 years. Using the preprocessing methods laid out in Section 2.1, 32×32 -pixel PSR images and their corresponding T:R ratios are derived from these 10 ECG segments. Our aim is to build a model capable of accurately predicting T:R ratios from these PSR images.

In order to increase the amount of training data, at each round of cross validation, after testing and validation sets have been reserved, we bolster our training set by combining it with the 32×32 pixel PSR images and T:R ratios derived from the ECG Identification (ECG-ID) Database. The ECG-ID data set, collected by Lugovaya [20], consists of 310 20-second ECG segments sampled at 500 Hz, with R and T peak annotations for the first 10 heartbeats. These signals are obtained from 90 participants, 44 men and 46 women, with ages ranging from 13 to 75.

¹ This study was performed with favourable opinion from the REC (17/SC/0623) and with R&D (RHMCA0528) approval. This study was conducted in accordance with the Research Governance Framework for Health and Social Care (2005), Good Clinical Practice and their relevant updates.

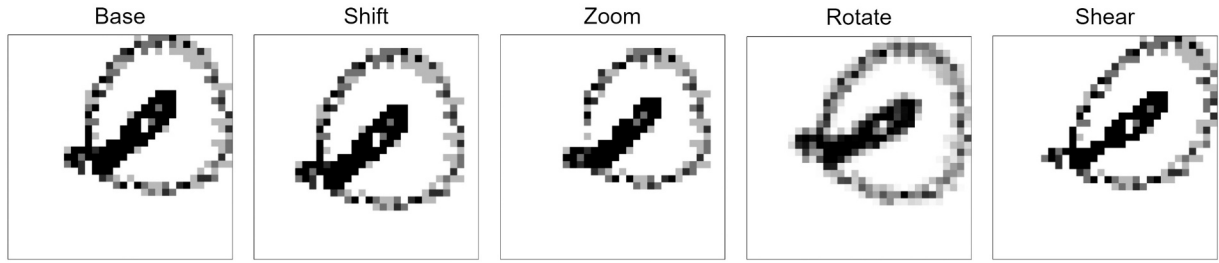


Fig. 9. Examples of how shifting down and to the left by 12.5%, zooming by -12.5% , shearing with an angle of 15° and rotating by 15° effect a PSR image.

Table 6

Detailed breakdown of the demographics and underlying aetiology of patients in the SGH-ECG data set.

Total number of participants		$n = 18$
Demographics	Mean age [years]	53.16
	Male	9 (50.00%)
Heart condition	Structurally normal heart	4 (22.22%)
	Adult congenital heart disease	3 (16.67%)
	Hypertrophic cardiomyopathy	3 (16.67%)
	Congestive heart failure	8 (44.44%)

3.2. Experiment 1: architecture selection

Our first experiment is to assess which model architectures, as detailed in Section 2.2, most accurately predict T:R ratios from 32×32 pixel PSR images. Fig. 10 shows the average accuracy across each of the 10 rounds of cross validation of some of the best performing models using the accuracy metrics laid out in Section 2.4 and the model naming convention laid out in Section 2.2.

As can be seen from the results in Fig. 10, the MLP5 model is capable of accurately predicting T:R ratios with a mean absolute prediction error of only 0.0558. We recall that T:R ratios range between 0 and 1 and the threshold for failing a screening is 0.33. In switching to a Basic CNN structure, we see a considerable drop in accuracy. Adding more convolutional layers by using Complex CNN feature-extraction blocks allows us to recoup this loss. The Complex CNN5 model outperforms the MLP5 model for all tested metrics. However, continuing to add convolutional layers by using Deep CNN feature-extraction blocks results in a drop in accuracy for all tested metrics. This is likely due to the fact that our data set is not sufficiently large to enable the training of such deep neural-networks.

Fig. 10 shows the average cross validation accuracies of only the best performing model for each type of feature-extraction block. There are several models with architectures utilising the MLP and Complex CNN feature-extraction blocks with a MAE under 0.06 indicating that, on average, these models are able to predict T:R ratios within 0.06 of their true value.

3.3. Experiment 2: model enhancement

In this experiment, we test the effect of image augmentation and creating ensemble models on our two best performing models from the previous experiment: MLP5 and Complex CNN5. We test 4 different image augmentation schemes (shifting, zooming, rotating and shearing) at three different magnitudes (tiny, small and moderate). We also evaluate ensemble models, created by averaging the prediction of 5 sub-models which each use a different portion of the training set for validation. The details of both of these methods are given in Section 2.3.

Neither method for model enhancement had a significant positive impact on the accuracies of the Complex CNN5 model. For this reason, Fig. 11 compares the average accuracy across the 10 rounds of cross validation of the enhanced MLP5 models which were able to outperform the base MLP5 model with that of the MLP5 and the Complex CNN5 models.

The only two image augmentation schemes which improve the performance of the MLP5 model are small rotations and small shears, which both only result in small increases in some accuracy metrics, along with small decreases in others. The MLP5 Ensemble model gives modest improvements in performance, outperforming all other variants of the MLP5 model on each metric. However, even the MLP5 Ensemble model is outperformed by the Complex CNN5 model on every metric aside from MAE, where the MLP5 Ensemble model is able to achieve an error of

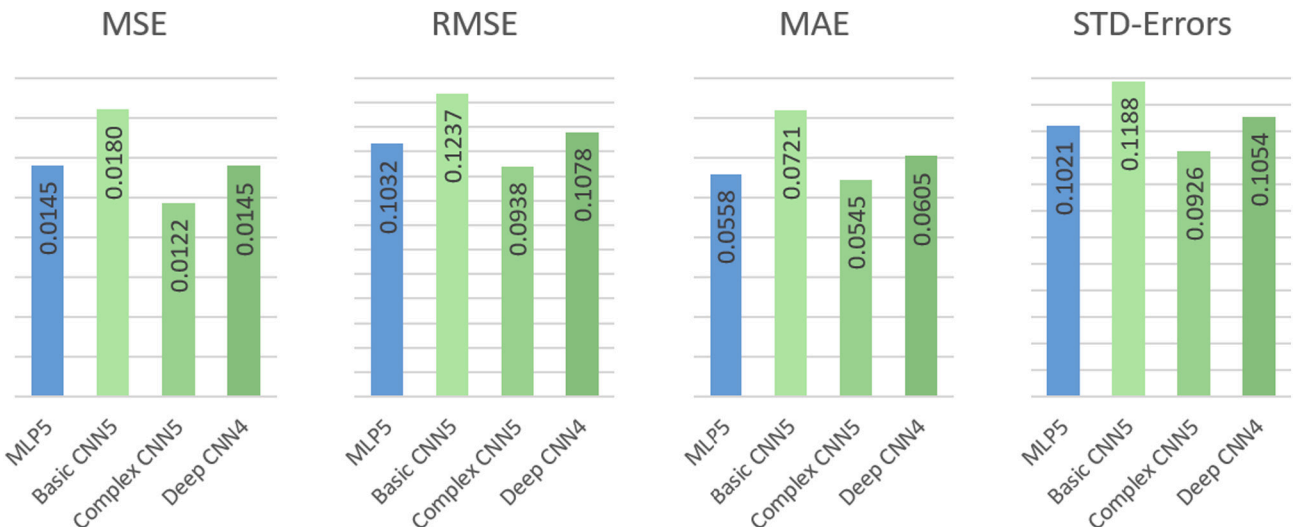


Fig. 10. Performances of the best models using each type of feature extraction block.

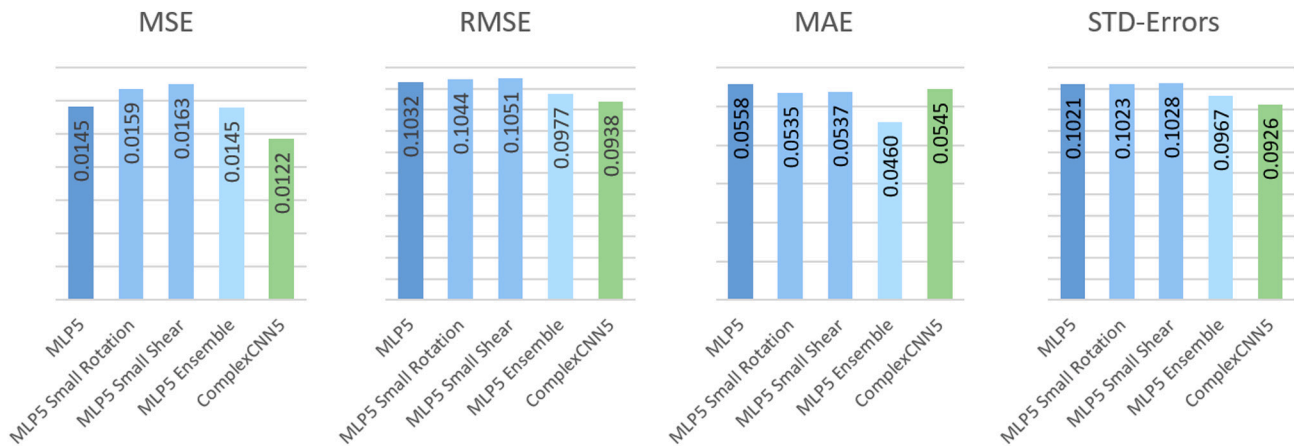


Fig. 11. Comparison between MLP5 model variants and the Complex CNN5 model.

only 0.460.

As MSE and RMSE are generally considered better measures of the regression capabilities of a model than MAE, we determine that the Complex CNN5 model is the best performing regression model as it significantly outperforms all tested models on these metrics. An additional drawback to the MLP5 Ensemble model is that, while the MLP5 model is less computationally expensive to train than the Complex CNN5 model, due to its more simplistic architecture, training the 5 MLP5 models required for creating the MLP5 Ensemble model is more expensive than training a single Complex CNN5 model.

Although, individually, rotating and shearing results in improved performance for the MLP5 model, we find that using both of these image augmentation schemes in concert does not result in improved performance. Likewise, we do not observe any improvement in performance when using image augmentation whilst creating ensemble models.

3.4. Screening tool

The previous two experiments have shown that the Complex CNN5 model outperforms all other tested models and is capable of predicting T:R ratios across a range of relevant patient groups within 0.0545 of their true value with an average MSE of only 0.122. In this section, we illustrate how this model could be used for the purpose of screening patients for S-ICD implantation eligibility.

After recording 24-hour ECG signals on multiple leads, each continuous 24-hour signal would be broken into 8640 non-overlapping 10-second segments which would then be processed using the methodology detailed in Section 2.1, resulting in 8640 PSR images for each lead, which are ordered chronologically. We would then input these images to our model and produce predictions for the T:R ratio of each 10-second segment represented by a single PSR image. As mentioned previously, from a clinical perspective only the magnitude of the T:R ratio is considered. Our model would output positive and negative T:R ratios (depending on the sign of the T wave), where the magnitude of these outputs would be used for the screening analysis. This would enable us to examine the T:R ratio for each 10-second segment of each lead from the 24-hour screening. To demonstrate this, we use this method to predict the T:R ratio for each 10-second segment of each lead of a 3-lead, 24-hour ECG recording. The primary aim of the screening is to determine if any of the leads are at a low enough risk of TWOS to be used by an S-ICD. As stated in Section 1, for a given lead, should a patient have a T:R ratio above 1:3 for a continuous period of at least 20 s, that lead would fail the screening. This means that, for each lead, if our model predicts a T:R ratio of over 0.33 for two or more consecutive 10-second segments, then that lead has failed the screening.

In the event that multiple leads pass the screening, the secondary aim of the screening is to determine which of the leads that pass are at the

lowest risk of TWOS. To examine how the behavior of the T:R ratio differs between each lead, we may wish to plot a histogram of what proportion of the 24-hour screening period the T:R ratio of a particular lead spent in each range of T:R ratios. To give an example of how our tool is capable of performing this task, Fig. 12 shows the histogram of the T:R ratios predicted by the Complex CNN5 model as well as the histogram of true T:R ratios for PSR images reserved for testing in one of the rounds of cross validation. As one can see, our model is able to predict the proportion of PSR images in the testing set whose T:R ratio belongs to each range to within 2.5% of the true value. Fig. 13 contains three histograms showing the proportion of the 24-hour screening that each lead spent with a predicted T:R ratio in each range. This would enable the cardiologist to assess which of the leads that passed the screening spends the smallest proportion of the 24-hour screening with high values of T:R ratio and, as such, is at the lowest risk of TWOS.

Additionally, while less directly applicable to the screening, our model allows for further analysis of the variation of the T:R ratio over the 24-hour screening. Fig. 14 shows one tool that our model facilitates. The variation of the T:R ratio is plotted for each lead, over the 24-h period. For readability, the lines in this graph are smoothed in such a way that each point gives the average T:R ratio for the preceding half hour. This could enable a cardiologist to detect any period during the 24 h where the T:R ratio was consistently high and, as such, the patient was at greater risk of TWOS. Our model could also allow cardiologists to further examine any single 10-second segment from within the 24-hour screening period across all 3 leads and view the ECG signal alongside its predicted T:R ratio. This is shown in Fig. 15, where, for a single 10-second segment, for each lead, the predicted T:R ratio is given alongside a plot of the original ECG signal, a plot of the filtered ECG signal and the

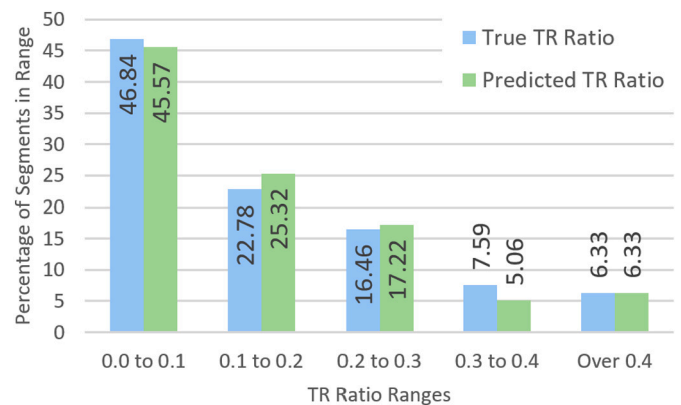


Fig. 12. Histogram of true and predicted TR ratios of 10-second ECG segments in the testing set.

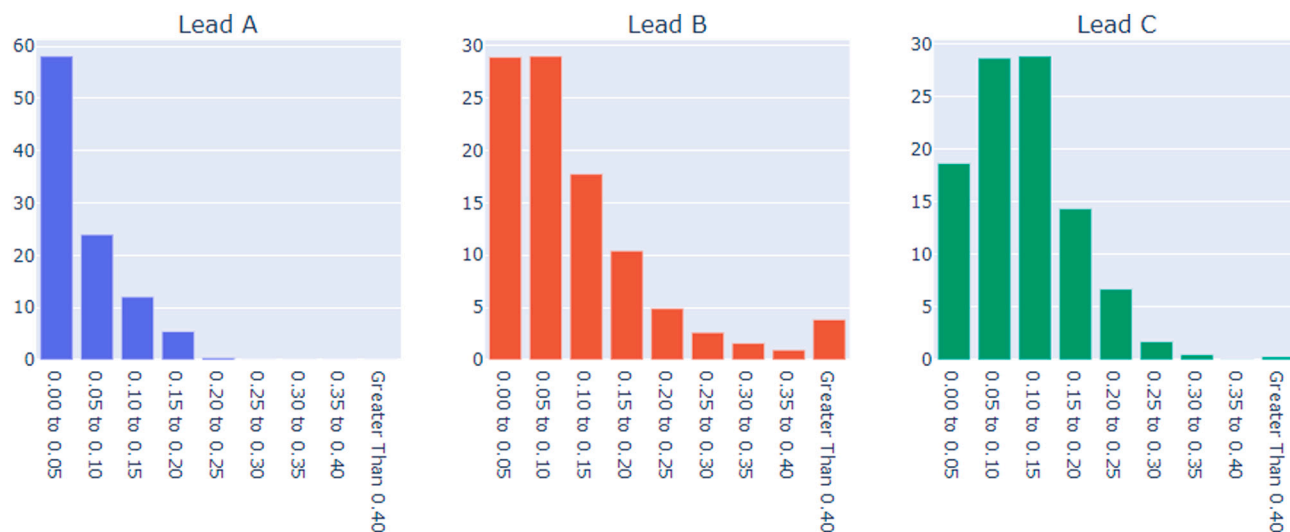


Fig. 13. Histogram of the predicted TR ratio over the 24-hour screening period for each lead

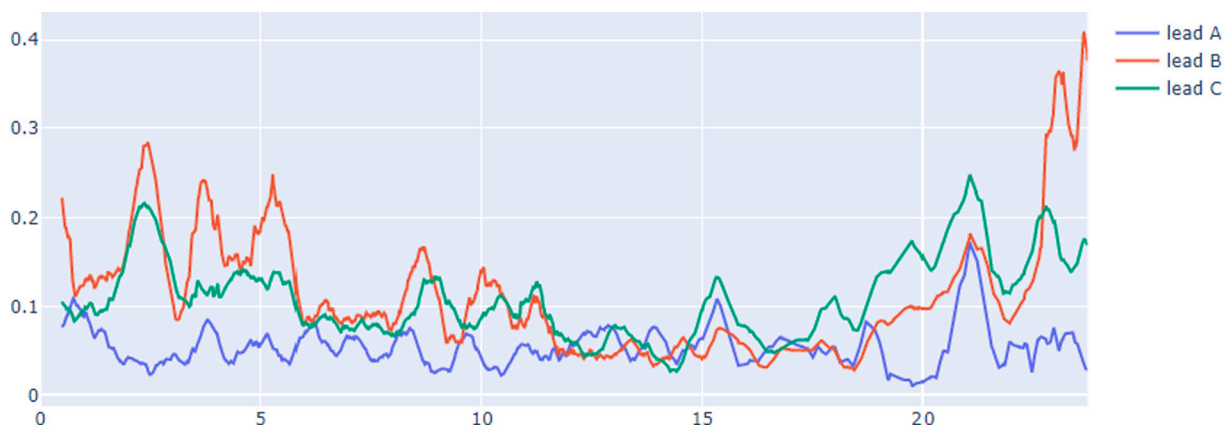


Fig. 14. Graph of the variation of the predicted TR ratio over the 24-hour screening period for each lead.

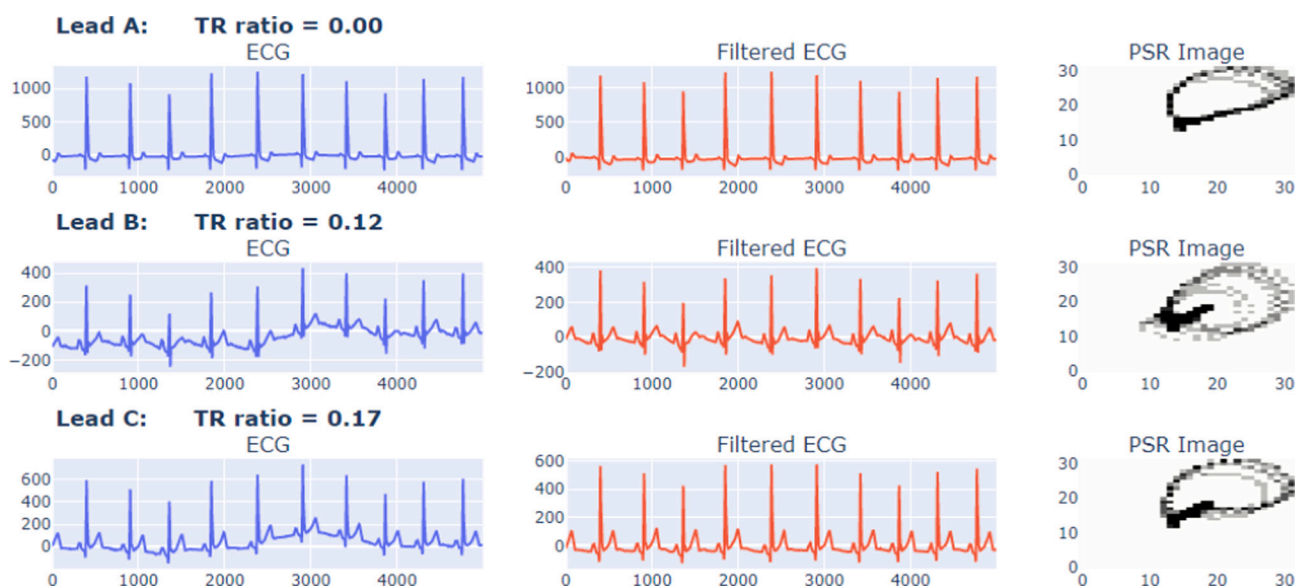


Fig. 15. Visualizations of data from a single 10-second segment of the 24 h ECG recording.

PSR image of that filtered ECG signal.

4. Conclusion

TWOS is an inherent risk with S-ICDs and can lead to inappropriate shocks. The current method for determining if a patient is likely to suffer from TWOS is to perform a 10-second screening of the patient's ECG, where the T:R ratio, a major predictor of TWOS, is examined. Temporal variations in the T:R ratio make this single 10-second screening unreliable.

In this paper, we have collected the SGH-ECG data set (detailed in Section 3.1) consisting of 10-second ECG segments from patients belonging to a range of patient groups which S-ICD implantation candidates are likely to belong to. As is shown in Section 3, we have used this data set to train and evaluate a range of deep learning models capable of predicting T:R ratios from 10-second ECG segments with a high degree of accuracy. The best performing of these models is a convolutional neural network (CNN) based model for predicting T:R ratios from 32x32 pixel PSR images derived from 10-second ECG segments. This is a novel approach to ECG analysis. While CNNs and PSR transformations have both individually been used in ECG analysis, to the best of our knowledge, this is the first work using them both in conjunction. We have also shown that this model can be integrated into a clinical tool for performing automated screening over much longer periods than the 10 s window currently used. For example, this tool would enable clinicians to perform 24-hours screenings. This prolonged screening period would enable one to much more reliably determine the normal range of a potential implantation candidate's T:R ratio than the current 10-second screening, as well as providing insight into the variation of the T:R ratio over the screening period. The increased reliability and descriptiveness of this tool can allow cardiologists to better assess patients' risk of TWOS and hence their eligibility for S-ICD implantation.

As future work, we will be using this tool to analyse the ECG of patients belonging to a range of patient groups in order to determine how the behavior of the T:R ratio differs between them. We also hope to determine if the T:R ratio of a patient's ECG can be an indicator of impending cardiac episodes such as VT or VF. Finally, we are looking to formally conduct cynical trials to determine if this method can be used in place of the current screening procedure.

Declaration of competing interest

To the best of our knowledge, the authors declare that there is no conflict of interest.

Acknowledgments

The work of Anthony J. Dunn is jointly funded by Decision Analysis Services Ltd. and EPSRC through the Studentship with Reference EP/R513325/1. The work of Alain B. Zemkoho is supported by the EPSRC grant EP/V049038/1 and the Alan Turing Institute under the EPSRC grant EP/N510129/1.

The feedback provided by Sion Cave (DAS Ltd) on the initial draft of the paper is gratefully acknowledged.

References

- [1] Adabag A, Selcuk, Luepker Russell V, Roger Véronique L, Gersh Bernard J. Sudden cardiac death: epidemiology and risk factors. *Nat Rev Cardiol* 2010;7(4):216.
- [2] Assanelli Deodato, Castelnovo Augusto Di, Rago Livia, Badilini Fabio, Vinetti Giovanni, Gianfagna Francesco, et al. T-wave axis deviation and left ventricular hypertrophy interaction in diabetes and hypertension. *J Electrocardiol* 2013;46(6):487–91.
- [3] Babu Giduthuri Sateesh, Zhao Peilin, Li Xiao-Li. Deep convolutional neural network based regression approach for estimation of remaining useful life. In: *International conference on database Systems for Advanced Applications*. Springer; 2016. p. 214–28.
- [4] Boersma Lucas, Barr Craig, Knops Reinoud, Theuns Dominic, Eckardt Lars, Neuzil Petr, et al. Implant and midterm outcomes of the subcutaneous implantable cardioverter-defibrillator registry: the effortless study. *J Am Coll Cardiol* 2017;70(7):830–41.
- [5] Chen Minyou, Fang Yonghui, Zheng Xufei. Phase space reconstruction for improving the classification of single trial EEG. *Biomedical Signal Processing and Control* 2014;11:10–6.
- [6] Dawid Aleksander. Psr-based research of feature extraction from one-second eeg signals: a neural network study. *SN Applied Sciences* 2019;1(12):1536.
- [7] Djemal Ridha, Bazyed Ayad G, Belwafi Kais, Gannouni Sofien, Kaaniche Walid. Three-class eeg-based motor imagery classification using phase-space reconstruction technique. *Brain Sci* 2016;6(3):36.
- [8] Fan Xiaomao, Yao Qihang, Cai Yunpeng, Miao Fen, Sun Fangmin, Li Ye. Multiscale fusion of deep convolutional neural networks for screening atrial fibrillation from single lead short eeg recordings. *IEEE J Biomed Health Inform* 2018;22(6):1744–53.
- [9] Fosbøl Emil Loldrup, Seibæk Marie, Brendorp Bente, Torp-Pedersen Christian, Køber Lars. Danish Investigations and Arrhythmia on Dofetilide (Diamond) Study Group. Prognostic importance of change in qrs duration over time associated with left ventricular dysfunction in patients with congestive heart failure: the diamond study. *J Card Fail* 2008;14(10):850–5.
- [10] Christian Garbin, Zhu Xingquan, Marques Oge. Dropout vs. batch normalization: an empirical study of their impact to deep learning. *Multimed Tools Appl* 2020;79: 12777–815.
- [11] Hazinski Mary F, Idris Ahamed H, Kerber Richard E, Epstein Andrew, Atkins Dianne, Tang Wanchun, et al. Lay rescuer automated external defibrillator ("public access defibrillation") programs: lessons learned from an international multicenter trial: advisory statement from the american heart association emergency cardiovascular committee; the council on cardiopulmonary, perioperative, and critical care; and the council on clinical cardiology. *Circulation* 2005;111(24):3336–40.
- [12] He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. Deep residual learning for image recognition. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition* 2016:770–8.
- [13] Ioffe Sergey, Szegedy Christian. Batch normalization: accelerating deep network training by reducing internal covariate shift. *PLMLR* 2015;37:448–56. <http://proceedings.mlr.press/v37/loffe15.html>.
- [14] Kiranyaz Serkan, Ince Turker, Gabbouj Moncef. Real-time patient-specific eeg classification by 1-d convolutional neural networks. *IEEE Trans Biomed Eng* 2015; 63(3):664–75.
- [15] Knops Reinoud E, Nordkamp Louise Ra Olde, Delnoy Peter-Paul HM, Boersma Lucas VA, Kuschyk Jürgen, El-Chami Mikhail F, et al. Subcutaneous or transvenous defibrillator therapy. *N Engl J Med* 2020;383(6):526–36.
- [16] Krishnan Shankar M, Dutt D Narayana, Chan YW, Anantharaman V. Phase space analysis for cardiovascular signals. In: *Advances in cardiac signal processing*. Springer; 2007. p. 339–54.
- [17] Kusumoto Fred M, Bailey Kent R, Chaouki Ahmad Sami, Deshmukh Abhishek J, Gautam Sandeep, Kim Robert J, et al. Systematic review for the 2017 AHA/ACC/HRS guideline for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: a report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society. *J Am Coll Cardiol* 2018;72(14):1653–76.
- [18] Lee Sang-Hong, Lim Joon S, Kim Jae-Kwon, Yang Junggi, Lee Youngho. Classification of normal and epileptic seizure eeg signals using wavelet transform, phase-space reconstruction, and euclidean distance. *Comput Methods Prog Biomed* 2014;116(1):10–25.
- [19] Liu Wenhan, Zhang Mengxin, Zhang Yidan, Liao Yuan, Huang Qijun, Chang Sheng, et al. Real-time multilead convolutional neural network for myocardial infarction detection. *IEEE J Biomed Health Inform* 2017;22(5):1434–44.
- [20] Lugovaya TS. Biometric Human Identification Based on Electrocardiogram. Master's thesis. Faculty of Computing Technologies and Informatics, Electrotechnical University "LETI", Saint-Petersburg, Russian Federation; 2005.
- [21] Madias John E. QTC interval in patients with changing edematous states: implications on interpreting repeat QTC interval measurements in patients with anasarca of varying etiology and those undergoing hemodialysis. *Pacing Clin Electrophysiol* 2005;28(1):54–61.
- [22] Madias John E, Bazaz Raveen, Agarwal Himanshu, Win Moethu, Medepalli Lalitha. Anasarca-mediated attenuation of the amplitude of electrocardiogram complexes: a description of a heretofore unrecognized phenomenon. *J Am Coll Cardiol* 2001; 38(3):756–64.
- [23] Priori Silvia G, Blomström-Lundqvist Carina, Mazzanti Andrea, Blom Nico, Borggrefe Martin, Camm John, et al. ESC guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death: the task force for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death of the european society of cardiology (ESC) endorsed by: association for european paediatric and congenital cardiology (AEPC). *Ep Europace* 2015;17(11):1601–87 (2015).
- [24] Miao Fen, Wen Bo, Hu Zhejing, Fortino Giancarlo, Wang Xi-Ping, Liu Zeng-Ding, et al. Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques. *Artif Intell Med* 2020; 108:101919.
- [25] Pourbabae Bahareh, Roshtkhari Mehrsan Javan, Khorasani Khashayar. Deep convolutional neural networks and learning eeg features for screening paroxysmal atrial fibrillation patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 2018;48(12):2095–104.

- [26] Roberts Felice M, Povinelli Richard J, Ropella Kristina M. Identification of eeg arrhythmias using phase space reconstruction. In: European Conference on Principles of Data Mining and Knowledge Discovery. Springer; 2001. p. 411–23.
- [27] Rocha T, Paredes S, de Carvalho P, Henriques J, Antunes M. Phase space reconstruction approach for ventricular arrhythmias characterization. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE; 2008. p. 5470–3.
- [28] van Rees Johannes B, Borleffs C Jan Willem, de Bie Mihály K, Stijnen Theo, Erven Lieselot van, Bax Jeroen J, et al. Inappropriate implantable cardioverter-defibrillator shocks: incidence, predictors, and impact on mortality. *J Am Coll Cardiol* 2011;57(5):556–62.
- [29] Vemishetty Naresh, Acharyya Amit, Das Saptarshi, Ayyagari Shivteja, Jana Soumya, Maharatna Koushik, et al. Classification methodology of CVD with localized feature analysis using phase space reconstruction targeting personalized remote health monitoring. In: 2016 Computing in Cardiology Conference (CinC). IEEE; 2016. p. 437–40.
- [30] Vemishetty Naresh, Gunukula Ramya Lakshmi, Acharyya Amit, Puddu Paolo Emilio, Das Saptarshi, Maharatna Koushik. Phase space reconstruction based CVD classifier using localized features. *Sci Rep* 2019;9(1):1–18.
- [31] Zhang Jing, Liu Aiping, Gao Min, Chen Xiang, Zhang Xu, Chen Xun. ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network. *Artif Intell Med* 2020;106:101856.
- [32] Sangaiah Arun Kumar, Arumugam Maheswari, Bian Gui-Bin. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. *Artif Intell Med* 2020;103:101788.
- [33] Oh Shu Lih V, Jahmunah, San Tan Ru, Ciacchio Edward J, Yamakawa Toshitaka, Tanabe Masayuki, Kobayashi Makiko, et al. Comprehensive electrocardiographic diagnosis based on deep learning. *Artif Intell Med* 2020;103:101789.
- [34] Chang Da-Wei, Lin Chin-Sheng, Tsao Tien-Ping, Lee Chia-Cheng, Chen Jiann-Torng, Tsai Chien-Sung, et al. Detecting digoxin toxicity by artificial intelligence-assisted electrocardiography. *Int J Environ Res Public Health* 2021;18(7):3839.
- [35] Nishimori Makoto, Kiuchi Kunihiro, Nishimura Kunihiro, Kusano Kengo, Yoshida Akihiro, Adachi Kazumasa, et al. Accessory pathway analysis using a multimodal deep learning model. *Sci Rep* 2021;11(1):1–8.
- [36] Kwon Joon-myung, Jung Min-Seung, Kim Kyung-Hee, Jo Yong-Yeon, Shin Jae-Hyun, Cho Yong-Hyeon, et al. Artificial Intelligence for Detecting Electrolyte Imbalance Using Electrocardiography. *Annals of Noninvasive Electrocardiology*; 2021, e12839.
- [37] Cho Jinwoo, Lee ByeongTak, Kwon Joon-Myoung, Lee Yeha, Park Hyunho, Oh Byung-Hee, et al. Artificial intelligence algorithm for screening heart failure with reduced ejection fraction using electrocardiography. *ASAIO J* 2021;67(3): 314–21.
- [38] Siontis Konstantinos C, Noseworthy Peter A, Attia Zach I, Friedman Paul A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol* 2021:1–14.
- [39] Kwon Joon-myung, Cho Younghoon, Jeon Ki-Hyun, Cho Soohyun, Kim Kyung-Hee, Baek Seung Don, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. *The Lancet Digital Health* 2020;2(7): e358–67.
- [40] Zhu Hongling, Cheng Cheng, Yin Hang, Li Xingyi, Zuo Ping, Ding Jia, et al. Automatic multilabel electrocardiogram diagnosis of heart rhythm or conduction abnormalities with deep learning: a cohort study. *The Lancet Digital Health* 2020;2(7):e348–57.
- [41] Jo Yong-Yeon, Cho Younghoon, Lee Soo Youn, Kwon Joon-myung, Kim Kyung-Hee, Jeon Ki-Hyun, et al. Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram. *Int J Cardiol* 2021;328:104–10.
- [42] Cho Younghoon, Kwon Joon-myung, Kim Kyung-Hee, Medina-Inojosa Jose R, Jeon Ki-Hyun, Cho Soohyun, et al. Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography. *Sci Rep* 2020;10(1): 1–10.
- [43] Department of Health. Research governance framework for health and social care. Crown, Produced by COI for the Department of Health 2005. gov.uk; 2005. [Accessed 17 August 2021].