IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 25, 2019, accepted December 12, 2019, date of publication December 16, 2019, date of current version December 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2960116

Dense Convolutional Networks With Focal Loss and Image Generation for Electrocardiogram Classification

MOHAMAD MAHMOUD AL RAHHAL[1], (Member, IEEE), YAKOUB BAZI[2], (Senior Member, IEEE), HAIDAR ALMUBARAK[2], (Member, IEEE), NAIF ALAJLAN[2], (Senior Member, IEEE), AND MANSOUR AL ZUAIR[2], (Member, IEEE)

[1]Information System Department, College of Applied Computer Science, King Saud University, Riyadh 11543, Saudi Arabia
[2]Computer Engineering Department, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Mohamad Mahmoud Al Rahhal (mmalrahhal@ksu.edu.sa)

This work was supported by the Deanship of Scientific Research at King Saud University through the Research Group under Grant RG-1435-050.

ABSTRACT In this paper, we propose a novel end-to-end learnable architecture based on Dense Convolutional Networks (DCN) for the classification of electrocardiogram (ECG) signals. This architecture is based on two main modules: the first is a generative module and the second is a discriminative one. The task of the generative module is to convert the one dimensional ECG signal into an image by means of fully connected, up-sampling, and convolution layers. The discriminative module takes as input the generated image and carries out feature learning and classification. To handle the data imbalance problem characterizing the ECG data, we propose to use the focal loss (FL) that is based on the idea of reshaping the standard cross-entropy loss such that it reduces the loss assigned to well-classified ECG beats. In the experiments, we validate the method using the well-known MIT-BIH arrhythmia database in four different scenarios, using four classes in the first scenario, five in the second and 12 in the third. Finally, supraventricular versus the other three and ventricular versus the other three from the scenario with four classes are used as the fourth scenario. The results obtained show that the method proposed here achieves a significant accuracy improvement over all previous state-of-the-art methods.

INDEX TERMS Generative, discriminative, ECG, classification, arrhythmia.

I. INTRODUCTION

The classification of electrocardiogram signals is one of the areas that has received the most attention in the field of biosignal analysis. Cardiac arrhythmias refer to a large group of conditions in which there is abnormal activity or behavior in the heart and represent an important group of cardiovascular diseases (CVD). The algorithms in computer-aided diagnosis systems play an important role in the detection and classification of cardiac arrhythmias. These algorithms have been designed to automate the process of ECG classification. This, in turn, will help greatly the cardiologists to monitor the physiological conditions of the heart at regular intervals. Towards this end, various strategies have been put forward to handle the classification problem [1]–[15]. Some of these strategies focus on signal processing techniques, such as frequency analysis [9], wavelet transform [10], hidden Markov models [12], support vector machines [13] and mixture-of-experts methods [15]. Because of the inter-subject variability of the ECG signals, the aforementioned techniques have not performed well in classifying a new patient's ECG signal. To address this problem new strategies have recently been introduced [16]–[22]. A multi-view learning approach for heartbeat classification was proposed in [16], it consists of two models, a general classification model, and a specific classification model. The general model is trained using similar subjects out of a population dataset, where a pattern matching based algorithm is developed to select the subjects that are "similar" to the particular test subject. The two models complement each other and are combined to achieve an improved subject-specific ECG analysis. Marinho in [17] also used handcrafted features extracted

The associate editor coordinating the review of this manuscript and approving it for publication was Amjad Mehmood.

using Fourier analysis, Goertzel, Higher-Order Statistics, and structural co-occurrence matrix. These features were then used to train several machine-learning algorithm including support vector machine and multilayer perceptron networks. In [18] an energy-efficient electrocardiogram (ECG) processor with weak-strong hybrid classifier was put forwards for arrhythmia detection, the proposed method uses a weak linear classifier (WLC), which is only used to identify beats with distinct characteristics. It does this by performing a simple threshold comparison based on the beat interval features and a novel morphology feature called the QRS area ratio. Luo et al. introduced a similar method, incorporating a subject-specific constraint to improve the classification performance of the deep neural network [21]. In [22], P. Li et al. implemented a parallel general regression neural network (GRNN) to classify heartbeats and designed an online learning program to form a personalized classification model for each patient. Recently deep learning techniques have generated a lot of interest as powerful computer-based methods capable of solving various recognition problems. First introduced by Hinton in [23], they focus on obtaining a good feature representation automatically from the input data [23]–[27].

Deep convolutional neural networks (CNNs) have performed well on a variety of applications including image classification [28]–[30], object detection [31]–[34] and image segmentation [35], [36]. In the biomedical engineering field, several authors have used deep learning methods to solve various problems such as detection and classification of brain tumors in MR images [37]–[40], breast cancer diagnosis and mass classification [41], [42], abdominal adipose tissue extraction [43], and skeletal bone age assessment in X-ray images [44].

Within the field of ECG research, there have been notable studies using deep learning for ECG signal analysis and arrhythmia detection. [20], [45]–[55]. A support vector machine (SVM) classifier has been used to classify the beats that are left unclassified by the WLC. In [20], Kiranyaz proposed a patient-specific ECG classification and monitoring system, by applying the learning model 1-D CNN to each patient in an adaptive manner. This method takes advantage of the additional patient-specific information prior to tackling the inter-class data variations caused by intersubject variability. In [46] the authors propose a classification module for paroxysmal atrial fibrillation (PAF) based on deep convolutional neural networks (CNN). The features are learned directly from the raw ECG time series data by using a CNN with one fully-connected layer. The learned features can effectively replace the traditional ad-hoc and time-consuming processes of hand-crafting user features. For a long time manually handcrafted features have been used and are still used to classify arrhythmia such as in [48] were they used a downsampled signal to generate handcrafted features including RR intervals, heartbeat intervals, and segmented morphologies. These features were the used to train a deep belief network. Another work that uses handcrafted feature is done by Sannino and De Pietro [49], they used features

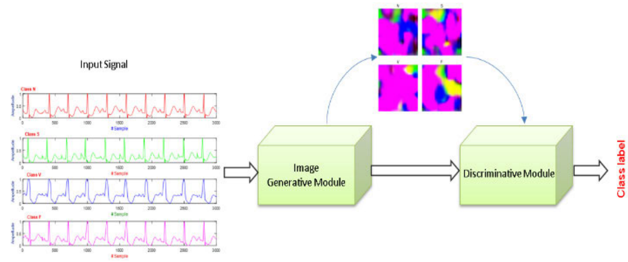


FIGURE 1. **Flowchart of the proposed method.**

such as Pre-RR interval and local average RR interval as input to a 7-layers multilayer perceptron networks and to combat the imbalanced data problem they used a subsample of the majority class in the data. In some papers no manual features engineering was used for example Zhou and Tan [50] used the raw signals as input to a convolutional neural network trained in two stages first using backpropagation then extreme learning machine fine turning of the last layer. The convolutional network is a 1-D network which received vector of length 250 as input.

In other work [51] raw ECG signal was also used as input to a deep learning network constructed as a restricted Boltzmann machine. Each heartbeat is centered around R peak and either padded or truncated to make all signal of the same length. They try to improve the result by duplicating the minority classes.

Recurrent neural network alongside convolutional network was also used to classify ECG signals [52]. Here the input is also a 1-D vector signal taken as a 10 seconds interval from the original signal and labeled with the most occurring label in the interval. Li *et al.* [53] tried to create patient-specific network by first training in a large corpus of multiple patient data then fine-tuning a network for individual patients. They also used a 1-D vector from the ECG signals as input to the network.

While most deep learning approaches including the ones using convolutional networks, use 1-D signal approach, there were some attempts to utilize 2-D convolution network by converting the 1-D ECG signal into a 2-D one. For example in [54] each signal is divided into 10 seconds intervals then a short-time Fourier transform was used in the small chunks to transform them into 2-D signal as spectrogram. These spectrograms are then used to train convolutional neural network to classify arrhythmias. Kim *et al.* [55] also used 2-D signal but not to classify arrhythmias. They used it to recognize users based on ECG feature using MIT-BHI NSRDB database. They projected the 1-D signal into 2-D space by minimizing a loss function then used the resulted 2-D signal as input to an ensemble of deep convolutional network. While the work in [54] and [55] converted the signal into a 2-D image using handcrafted features, the method used is not robust enough to discover latent features of the signal and convert them into pixel values.

In this paper, we propose an alternative approach based on deep learning for the classification of ECG signals. Typically,

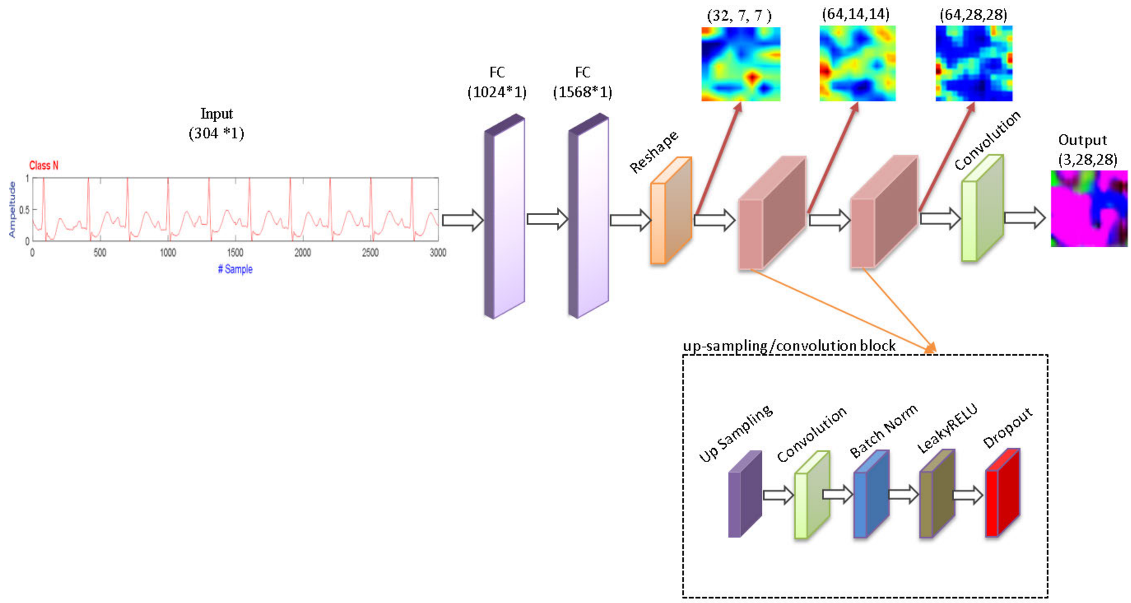


FIGURE 2. Generator network.

our approach uses a two-stage CNNs for carrying classification. The first module aims to convert the one dimensional ECG signal to an image using set an opportune generative network. The second one called discriminative network mainly based on dense convolutional networks (DCNs) takes the output of the generative module and carries out classification as in standard image classification paradigms. To handle the class imbalance problem, we propose to exploit the focal loss (FL) instead of the cross-entropy loss to down-weight the loss for the well-classified ECG beats. This paper conveys the following main contribution:

1) Propose a novel end-to-end learnable architecture for the classification of ECG signals with signal to image conversion using a generator network;
2) Handle the class unbalance problem using the focal loss;
3) The experimental results obtained on the MIT-BIH database confirm its promising capabilities compared to state-of-the-art method in terms of classification accuracy.

The remainder of the paper is organized as follows. In Section II, a detailed description of the proposed method is presented. Results and discussion are shown in Section III. Finally, our conclusions and future developments are sketched out in Section IV.

II. PROPOSED METHOD

The Let $Tr = \{\mathbf{x}_i, y_i\}_{i=1}^{n}$ be a training set where $\mathbf{x}_i \in \mathcal{R}^d$ is an ECG beat signal , $y_i \in \{1, 2, \ldots, K\}$ is its corresponding class label, K is the number of classes and n is the number of training samples. Our aim is to develop a CNN architecture that allows the classification of the test ECG record $Ts = \left\{\mathbf{x}_j\right\}_{j=n+1}^{n+m}$ based on the available training set. Fig. 1 shows a flowchart of the proposed method, which consists of two modules. The detailed descriptions for these modules are provided in the next subsections.

A. GENERATIVE MODULE

The task of the generative network is to convert the one dimensional ECG beat into an image. Fig. 2 shows the main architecture of the generator network. It is composed of two fully-connected (FC) layers, a reshape layer, two sets of up-sampling / convolution layer blocks and a final convolution layer. In the first stage, the signal is fed into two consecutive fully-connected layers to generate the ECG features, with dimensions 1024 and 1568 respectively. The second step is to reshape the 1D feature tensor of shape (1568,1) into a tensor of dimensions (32,7,7) (channels, height, width), as shown in figure 2. We can describe the process with the equation:

$$X_t^r = \text{reshape}\left(\text{ReLU}\left(X_t\right)\right)_7^7 \tag{1}$$

After that, the signal is fed through two up-sampling/ convolution blocks to obtain a tensor of dimensions (128,28,28) (channels, height, width). The final tensor is then passed through a convolution layer to obtain an image of dimension (3,28,28) (channels, height, width).

It is worth recalling that the first FC layer is followed by a batch-normalization, a ReLU activation function, and a dropout. The reshape layer is also followed by batch normalization, activation, and dropout layers. Each of the up-sampling/convolution blocks consist of a layer of up-sampling followed by convolution, batch normalization, ReLU activation then dropout regularization layers. Batch normalization allows each layer of the network to learn by itself a little more independently of other layers, while

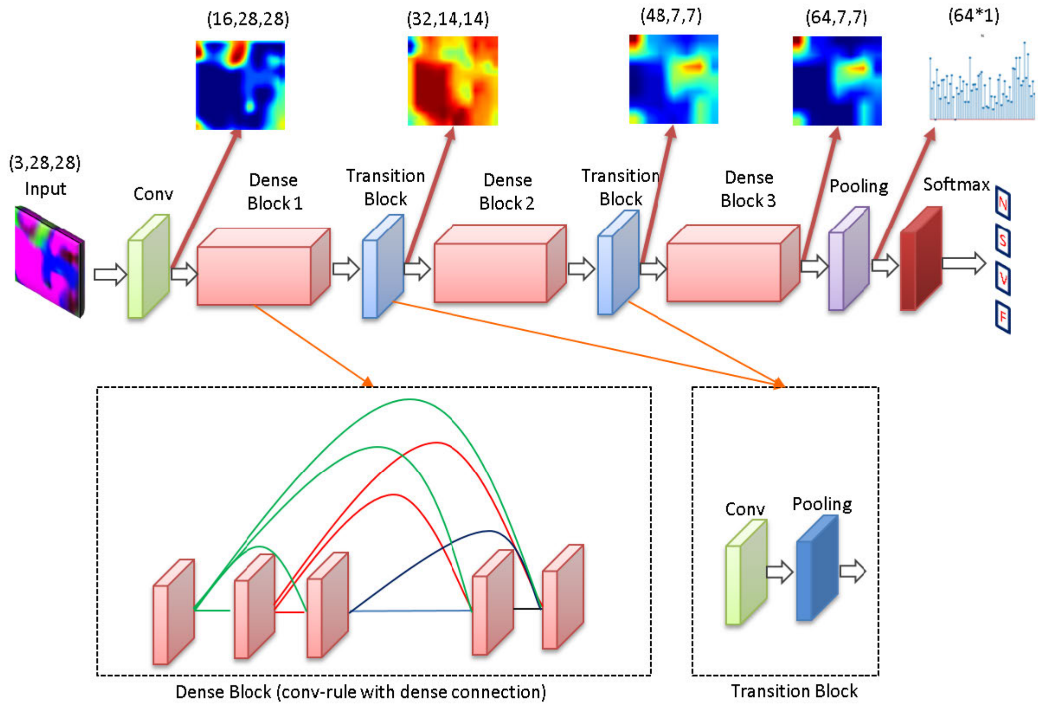


FIGURE 3. Discriminative module.

dropout is a regularization technique used to prevent overfitting during the training phase, by randomly dropping nodes from the hidden layers.

B. DISCRIMINATIVE MODULE

The discriminative module takes the images produced by the generative module as its input and classifies them into their respective classes. Fig 3. shows the discriminative module, it consists mainly of DenseNet blocks [56].

In traditional convolutional networks, a single image x_0 or a single batch of images is passed through the network, which is comprised of L layers. Each layer performs a non-linear transformation $H_\ell(.)$ on the image or batch of images, where ℓ is the layer index. $H_\ell(.)$ can be a composite function of operations, such as Batch Normalization (BN) [57], rectified linear units (ReLU) [58], Pooling [59] or Convolution (Conv). We denote the output of the ℓ^{th} layer as x_ℓ. In traditional convolutional networks, the input of layer ℓ comes exclusively from the layer $\ell - 1$.

To enhance the information sharing and information flow between layers, direct connections from any layer to its subsequent layers were introduced [56]. Fig. 4 shows the DenseNet block connectivity between layers.

With this configuration, the l^{th} layer receives the feature-maps of all preceding layers , $x_0, x_1, \ldots, x_{\ell-1}$ as input, and its output becomes:

$$x_\ell = H_\ell([x_0, x_1, \ldots, x_{\ell-1}]) \tag{2}$$

where $[x_0, x_1, \ldots, x_{\ell-1}]$ refers to the concatenation of feature-maps from previous layers.

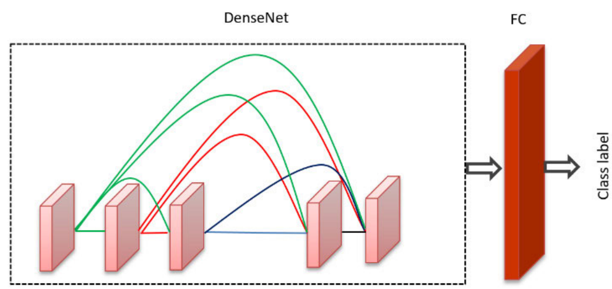


FIGURE 4. DensNet.

The concatenation operation used in Eq. (2) will not function correctly if the size of the feature-map from one of the previous $\ell - 1$ layers is not the same as the current layer ℓ. As down-sampling is an essential part of convolutional networks, changing the size of feature-maps, the connectivity after these down-sampling layers is treated differently. To include down-sampling, the network is divided into multiple densely-connected dense blocks, as shown in Fig3. The layers between the dense blocks are referred to as transition layers; they perform the convolution and pooling.

C. HANDLING CLASS IMBALANCE USING FL

Imbalanced data sets can negatively affect the overall performance of classification systems. To overcome this issue, one can apply sampling strategies such as a Synthetic Minority Over-sampling Technique (SMOTE) [60]. Another possible solution is to exploit the focal loss introduced recently in the context of object recognition. It relies mainly on

The focal loss technique [61] is designed to address the scenario in which there is an extreme imbalance between the classes during the dataset training (e.g., in the DS1 large class imbalance, N = 54777, S = 973, V = 3769, and F = 414). In focal loss, a modulating factor, $(1-p_t)^\gamma$ is added to the cross entropy loss, with tunable focusing parameter $\gamma \geq 0$. The focal loss is defined as:

$$\mathrm{FL}(p_t) = -(1-p_t)^\gamma \log(p_t) \tag{3}$$

With the values of $\gamma \in [0, 5]$.

III. EXPERIMENTAL RESULTS

A. DATASET DESCRIPTION

In experiments, we used three ECG databases to evaluate our method. The first is the MIT-BIH Arrhythmia database (MIT-BIH)) [62], [63]. This Database contains 48 excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects. The recordings were collected from a mixed population of inpatients ($\sim$ 60%) and outpatients ($\sim$ 40%). Of the 47 patients there were 25 men aged from 32 to 89 years old, and 22 women aged from 23 to 89 years old. Each record is slightly over 30 minutes long and sampled at 360 Hz. The original raw dataset consists of 4000 24-hour ambulatory ECG recordings, for the MIT-BIH dataset the first 23 records were chosen at random from this raw dataset, they were numbered from 100 to 124 inclusive with some numbers missing. The remaining 25 records were selected from the same original set to include a variety of rare but clinically significant arrhythmias.

The second database is INCART. It consists of 75 records which are annotated and derived from 32 Holter records. Each record includes 12 generic leads and was obtained from number of patients (17 men and 15 women, aged between 18 to 80) undergoing tests for coronary artery disease. Each record is 30 minutes and sampled at 257 Hz. An automatic algorithm generated the initial annotations, then the annotation was corrected manually.

The third database is called MITBIH Supraventricular Arrhythmia Database (SVDB) and consists of 78 two-lead records of approximately 30 minutes and 128 Hz sampling rate. The recordings ' beat type annotations were first performed automatically by the Marquette Electronics 8000 Holter scanner and then checked and updated by a medical student.

Table 1. shows classes distribution in this scenario, it is clear that class S and F are minor classes comparing to classes N, and V.

B. EXPERIMENTAL SETTINGS

The training was done for 250 epochs and the batch size was set to 100. Adam optimizer with learning rate of 0.001 was used for weight update. The gamma value for the focal loss (in equation 3) was set to 0.5. The DenseNet consisted of 3 blocks and a total of 16 layers. The initial filter size in the DenseNet was 32 with a growth rate of 4 and dropout of 0.5. For a detailed explanation of these parameters refer to the DenseNet paper [47]. The experiments were carried out using a computer with Intel Xeon E5620 processer, 24 GB RAM, and NVIDIA GeForce GTX 1060 GPU with 6 GB of memory.

TABLE 1. Classes distribution in the training and testing datasets.

Dataset	N	S	V	F	No. of rec
MIT(BIH) DS1	45,777	973	3769	414	22
MIT(BIH) DS2	44,011	2049	3216	388	22
INCART	153,545	1958	2000	219	75
SVDB	145,436	10733	8281	23	70

The experiments were carried out using three strategies: cross-entropy, focal loss, and cross-entropy with resampling. The cross-entropy strategy takes the unmodified training dataset and uses the cross-entropy loss to update the network weights. The focal loss strategy also takes the unmodified training dataset but uses instead the focal loss to update network weights, to combat the unbalanced classes problem. The last strategy is the cross-entropy with the resampling method; this strategy uses the cross-entropy loss, but to deal with unbalanced classes it uses the resampling technique to oversample the classes with low numbers of samples. Fig 5. shows the convergence when using focal loss and resampling strategies. These strategies were used in four different scenarios: 4 classes, 5 classes, 12 classes, and a scenario with class S (Supraventricular) versus the other classes and class V (ventricular) versus the other classes. More details about these scenarios are provided in the next section.

C. RESULTS

According to state-of-the-art ECG classification techniques, performance can be measured using standard metrics such as classification accuracy (Acc), sensitivity (Sen), specificity (Spe) and positive predictivity (Ppr). While accuracy measures the system performance across all classes of ECG beats, the other metrics are specific to each class, and they measure the ability of the classification algorithm to distinguish certain events. The respective definitions of these four common metrics using true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are written as:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Ppr = \frac{TP}{TP + FP} \tag{5}$$

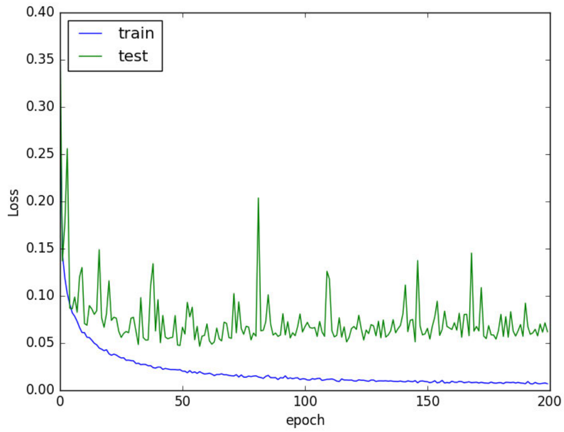
$$Sen = \frac{TP}{TP + FN} \tag{6}$$

$$Spe = \frac{TN}{TN + FP} \tag{7}$$

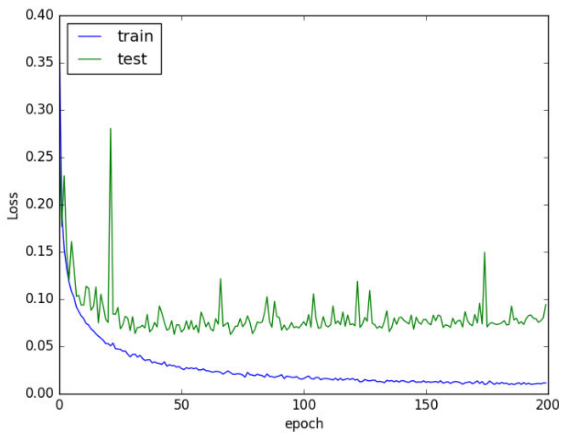
In order to evaluate the proposed approach, we used six different scenarios as shown in Table 1. These are described in the following sections:

1) SCENARIO 1

In this scenario the performance measurements are based in terms of four classes: 1 - Normal (N); 2 - Supraventricular (S);



(a)



(b)

FIGURE 5. Loss convergence plot by using: a) focal loss strategy, b) resampling strategy.

TABLE 2. Training and test datasets for proposed scenarios.

Scenarios	Training	Test
1	DS1	DS2
2	50% (17 records)	50% (17 records)
3	50% (Ds1+DS2)	50% (Ds1+DS2)
4	DS1	11 Rec(VEB)
		14 Rec (SVEB)
	DS1	24 Rec
	DS1	(Ds1+DS2)44 Rec
5	DS1	SVDB (7 0Rec)
6	DS1	INCART(75 Rec)

3 - Ventricular (V) and 4 - Fusion (F). Fig 6. shows examples of these classes.

In all experiments, similar to [64], we constructed the training set DS1 = {101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230} of the MIT-BIH dataset. The remaining records of this database are taken to form DS2 = {100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234}, which is used as the testing set.

TABLE 3. Classification performance for scenario 1 vs. methods in the literature.

Methods	N		S		V		F	
	Sen	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*
Jiang et al [65]	0.98	0.98	0.64	0.64	0.91	0.90	0.77	0.44
FL	0.99	0.99	0.77	0.94	0.97	0.95	0.80	0.68
CE	0.99	0.99	0.77	0.92	0.97	0.95	0.78	0.67
CERE	0.99	0.99	0.80	0.80	0.97	0.94	0.73	0.75

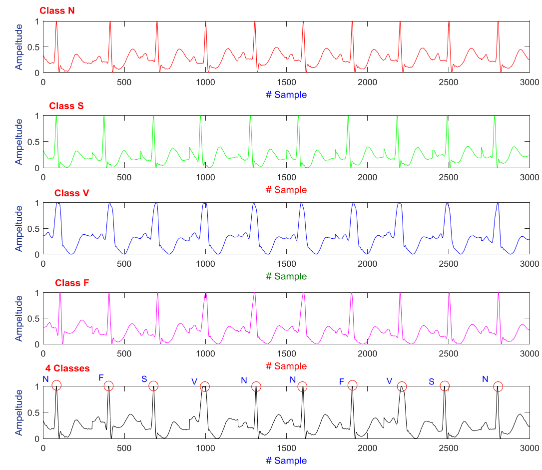


FIGURE 6. ECG example for four different classes.

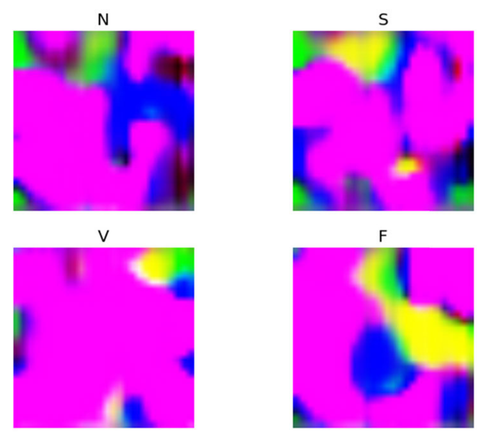


FIGURE 7. Example of the generator's output images for the classes N, S, V, and F.

Table 2. shows the distribution of classes in this scenario, it is clear that class S and F are minor classes comparing to classes N, and V.

Table 3. shows the classification performance of the proposed method in terms of four classes in three different strategies (focal loss, cross entropy, and cross entropy with resampling).

Fig. 7 shows an example of the image produced from the generator network for the different classes. Fig. 8 shows an example of the intermediate feature-maps from the

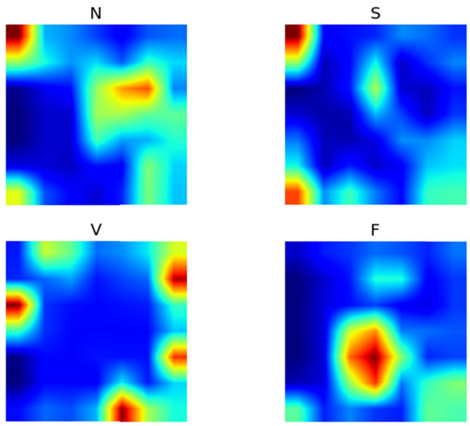


FIGURE 8. Example of feature maps form the discriminator (dense block 2 output) for the classes N,S,V, and F.

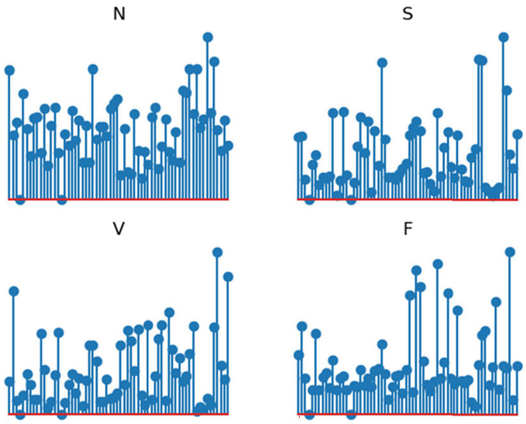


FIGURE 9. Example of the final features vector used for classification form the discriminator for the classes N,S,V, and F.

discriminator network for the different classes, taken from the output of the second dense block. Fig 9 shows an example of the final feature-maps (after the global average pooling, but before the Softmax in Fig3) which are used for classification.

As can be seen in Table 3, the values of (Sen, Ppr) for all classes using the focal loss strategy are equal to (0.99, and 0.99) for class N, (0.77, and 0.94) for class S, (0.97, and 0.95) for class V and (0.80, and 0.68) for class F. With cross entropy, the values of (Sen, Ppr) for all classes are equal to (0.99, and 0.99) for class N, (0.77, and 0.92) for class S, (0.97, and 0.95) for class V and (0.78, and 0.67) for class F. Finally, by using cross entropy with resampling the values of (Sen, Ppr) are equal to (0.99, and 0.99) for class N, (0.80, and 0.80) for class S, (0.97, and 0.94) for class V and (0.73, and 0.75) for class F. We can conclude that using focal loss achieves better results than the other two strategies.

2) SCENARIO 2

In this experiment, in accordance with [22], 17 records were selected from the MIT-BIH arrhythmia database [62], [63], the serial numbers of these records were 100, 103, 104, 106, 112, 119, 122, 200, 203, 208, 209, 217, 222, 223, 230, 232 and 233. We then randomly selected 50% of the data as the training set and the other 50% as the test data, with the results presented in terms of five classes: 1 - N (Normal); 2 – S; 3 - Ventricular (V); 4 - F (Fusion) and 5 – Unknown (Q). Table 4. presents confusion matrices for the ECG beat classification of the 17 records in terms of the 5 classes using focal loss. Table 5. presents the classification performances using the three different strategies comparing with previous proposed method in [22], one can see that our method performs better.

TABLE 4. Confusion matrix for scenario 2 (5 classes): a) without normalization b) with normalization.

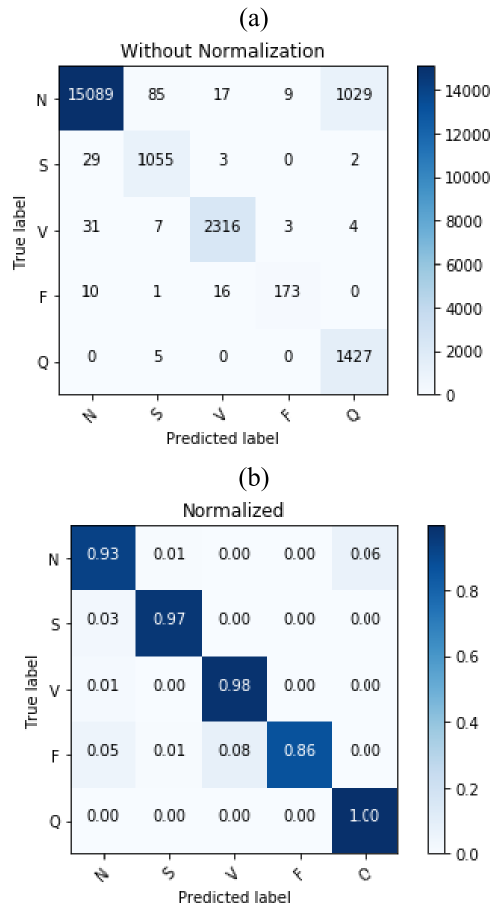


TABLE 5. Classification performance for scenario 2 vs. methods in the literature.

Method	N		S		V		F		Q	
	Sen	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*
P. Li et al [22]	0.97	0.97	0.98	0.89	0.92	0.88	0.82	0.84	0.84	0.76
FL	0.99	1.00	0.97	**0.92**	0.98	0.99	0.87	0.94	1.00	0.99
CE	0.99	0.99	0.98	**0.85**	0.98	0.98	0.78	0.95	1.00	1.00
CERE	1.00	*0.99*	0.94	**0.96**	0.97	*0.99*	0.90	*0.91*	*1.00*	*0.99*

3) SCENARIO 3

In this scenario, we have 12 ECG heartbeats classes. Each class is randomly halved, 50% are used for training and the other 50% are used for testing. The confusion matrices

TABLE 6. Classification performance for scenario 3 ((A): First 6 classes), and (B) The remaining classes).

(a)

Method	N		L		R		A		a		J	
	Sen	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*
FL	0.99	0.99	1.00	0.99	1.00	0.99	0.95	0.89	0.54	0.85	0.89	0.70
CE	0.99	0.99	0.99	1.00	0.99	1.00	0.92	0.93	0.61	0.84	0.95	0.51
CERE	1.00	0.99	0.99	0.99	1.00	0.99	0.90	0.95	0.55	0.84	0.87	0.97

(b)

Method	V		F		j		E		f		Q	
	Sen	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*	*Sen*	*Ppr*
FL	0.94	0.99	0.85	0.80	0.76	0.57	0.88	0.98	0.95	097	0.06	0.14
CE	0.97	0.96	0.78	0.84	0.71	0.65	0.92	0.80	0.92	0.99	0.00	0.00
CERE	0.98	0.95	0.67	0.98	0.58	0.79	0.84	1.00	0.93	0.99	0.13	0.50

TABLE 7. Confusion matrix for the ECG beat classification for scenario 3 using focal loss: a) Without normalization b) With normalization.

(a)

Without Normalization

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	38276	6	7	59	3	13	60	34	42	4	0	1
L	18	3968	4	5	0	0	14	0	0	3	1	1
R	16	0	4582	1	0	9	4	0	0	0	0	0
A	75	0	5	1206	4	7	7	0	4	3	0	1
a	11	0	0	8	51	0	13	0	0	0	0	0
j	0	0	0	1	0	35	1	0	0	0	0	0
V	51	3	1	10	3	0	3448	22	0	1	0	0
F	50	1	0	1	0	0	32	302	0	0	0	0
J	27	0	1	1	0	4	0	0	84	1	0	0
E	0	0	3	0	0	1	0	0	0	45	0	0
f	17	7	0	1	0	0	3	0	0	0	333	0
Q	8	1	1	0	0	0	3	0	0	1	2	0

(b)

Normalized

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	0.06	0.00	0.00	0.92	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00
a	0.13	0.00	0.00	0.10	0.61	0.00	0.16	0.00	0.00	0.00	0.00	0.00
j	0.00	0.00	0.00	0.03	0.00	0.95	0.03	0.00	0.00	0.00	0.00	0.00
V	0.01	0.00	0.00	0.00	0.00	0.00	0.97	0.01	0.00	0.00	0.00	0.00
F	0.13	0.00	0.00	0.00	0.00	0.00	0.08	0.78	0.00	0.00	0.00	0.00
J	0.23	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.71	0.01	0.00	0.00
E	0.00	0.00	0.06	0.00	0.00	0.02	0.00	0.00	0.00	0.92	0.00	0.00
f	0.05	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.92	0.00
Q	0.50	0.06	0.06	0.00	0.00	0.00	0.19	0.00	0.00	0.06	0.12	0.00

below show the classification results of these classes. Table 7-9 shows the confusion matrix for this 12 class scenario using a) focal loss, b) Entropy and c) Entropy with resampling. Table 6 shows the obtained classification results.

TABLE 8. Confusion matrix for the ECG beat classification for scenario 3 using entropy: A) Without normalization B) With normalization.

(a)

Without Normalization

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	38276	6	7	59	3	13	60	34	42	4	0	1
L	18	3968	4	5	0	0	14	0	0	3	1	1
R	16	0	4582	1	0	9	4	0	0	0	0	0
A	75	0	5	1206	4	7	7	0	4	3	0	1
a	11	0	0	8	51	0	13	0	0	0	0	0
j	0	0	0	1	0	35	1	0	0	0	0	0
V	51	3	1	10	3	0	3448	22	0	1	0	0
F	50	1	0	1	0	0	32	302	0	0	0	0
J	27	0	1	1	0	4	0	0	84	1	0	0
E	0	0	3	0	0	1	0	0	0	45	0	0
f	17	7	0	1	0	0	3	0	0	0	333	0
Q	8	1	1	0	0	0	3	0	0	1	2	0

(b)

Normalized

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	0.06	0.00	0.00	0.92	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00
a	0.13	0.00	0.00	0.10	0.61	0.00	0.16	0.00	0.00	0.00	0.00	0.00
j	0.00	0.00	0.00	0.03	0.00	0.95	0.03	0.00	0.00	0.00	0.00	0.00
V	0.01	0.00	0.00	0.00	0.00	0.00	0.97	0.01	0.00	0.00	0.00	0.00
F	0.13	0.00	0.00	0.00	0.00	0.00	0.08	0.78	0.00	0.00	0.00	0.00
J	0.23	0.00	0.01	0.01	0.00	0.03	0.00	0.00	0.71	0.01	0.00	0.00
E	0.00	0.00	0.06	0.00	0.00	0.02	0.00	0.00	0.00	0.92	0.00	0.00
f	0.05	0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.92	0.00
Q	0.50	0.06	0.06	0.00	0.00	0.00	0.19	0.00	0.00	0.06	0.12	0.00

4) SCENARIO 4

In this scenario the results have been proposed in terms of VEB (V versus the other three classes) and SVEB (S versus the other three classes) detection using three different cases for building the test set: Case 1: Using the 11 common testing records for VEB (records 200, 202, 210, 213, 214, 219, 221, 228, 231, 233 and 234), and 14 testing records for SVEB (records 200, 202, 210,212, 213, 214, 219, 221, 222, 228, 231, 232, 233 and 234). Case 2: Using the 24 common testing records from 200 up to 234. Case 3: Using the 48 records as test set (i.e. DS1 and DS2).

Of the state-of-the-art techniques, the method in [20] achieved the highest sensitivity score with 95.1% for class V (VEB), and 68.8% for class S (SVEB). The proposed methods

TABLE 9. Confusion matrix for the ECG beat classification for scenario 3 using entropy resampling: a) Without normalization b) With normalization.

(a)

Without Normalization

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	38355	0	10	75	3	7	22	14	15	0	2	2
L	100	3876	5	12	0	2	16	0	0	0	3	0
R	31	0	4579	1	0	1	0	0	0	0	0	0
A	80	0	7	1211	2	5	6	0	1	0	0	0
a	19	0	0	14	41	4	5	0	0	0	0	0
j	0	0	0	2	0	35	0	0	0	0	0	0
V	115	1	3	10	5	7	3382	15	0	0	0	1
F	65	0	0	1	0	1	321	288	0	0	0	0
J	49	0	3	4	0	0	0	0	61	0	1	0
E	0	0	3	0	0	0	0	0	0	46	0	0
f	18	1	4	0	0	4	0	0	0	0	333	1
Q	9	0	3	0	0	0	1	0	0	0	2	1

(b)

Normalized

True label \ Predicted label	N	L	R	A	a	j	V	F	J	E	f	Q
N	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
L	0.02	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
R	0.01	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
A	0.06	0.00	0.01	0.92	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
a	0.23	0.00	0.00	0.17	0.49	0.05	0.06	0.00	0.00	0.00	0.00	0.00
j	0.00	0.00	0.00	0.05	0.00	0.95	0.00	0.00	0.00	0.00	0.00	0.00
V	0.03	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.00	0.00	0.00	0.00
F	0.10	0.00	0.00	0.00	0.00	0.00	0.47	0.43	0.00	0.00	0.00	0.00
J	0.42	0.00	0.03	0.03	0.00	0.00	0.00	0.00	0.52	0.00	0.01	0.00
E	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.94	0.00	0.00
f	0.05	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.92	0.00
Q	0.56	0.00	0.19	0.00	0.00	0.00	0.06	0.00	0.00	0.00	0.12	0.06

TABLE 10. VEB and SVEB classification performance for scenario 4 vs. methods in the literature.

Methods	VEB				SVEB			
	Acc	***Sen***	***Spe***	***Ppr***	***Acc***	***Sen***	***Spe***	***Ppr***
Hu et al.	94.8	78.9	96.8	75.8	N/A	N/A	N/A	N/A
[6]	98.8	94.3	99.4	95.8	97.5	74.9	98.8	78.8
Ince et al. [7]	97.9	90.3	98.8	92.2	96.1	81.8	98.5	63.4
S.Kiranyaz [20]	98.9	95.9	99.4	96.2	96.4	68.8	99.5	79.2
S. S. Xu et al. [51]	N/A	90.5	98.1	N/A	N/A	66.2	98.6	N/A
Proposed 01 FL	**99.4**	**96.5**	**99.7**	**97.9**	**98.3**	**79.7**	**99.4**	**87.4**
Proposed 01 CE	**99.3**	**95.6**	**99.7**	**98.2**	**98.2**	**80.5**	**99.1**	**84.0**
Proposed 01 CERE	**99.4**	**98.2**	**99.6**	**97.1**	**98.7**	**78.0**	**99.8**	**95.6**
Jaing and Kong [6]	98.1	86.6	93.3	93.3	96.6	50.6	98.8	67.9
Ince et al [7]	97.6	83.4	87.4	87.4	96.1	62.1	98.5	56.7
S. Kiranyaz [20]	98.6	95.0	98.1	89.5	96.4	64.6	98.6	62.1
Proposed 02 FL	**99.2**	**95.0**	**99.7**	**96.9**	**98.7**	**69.8**	**100**	**99.1**
Proposed 02 CE	**99.1**	**93.4**	**99.8**	**97.5**	**98.9**	**79.4**	**99.8**	**94.1**
Proposed 02 CERE	**99.1**	**94.0**	**99.7**	**96.6**	**98.6**	**69.7**	**99.9**	**97.9**
Ince et al [7]	98.3	84.6	87.4	87.4	97.4	63.5	99.0	53.7
S. Kiranyaz [20]	99	93.9	90.6	90.6	97.6	60.3	99.2	63.5
Proposed 03 FL	**98.8**	**94.9**	**99.1**	**88.8**	**98.9**	**83.4**	**99.4**	**78.4**
Proposed 03 CE	**99.1**	**87.7**	**99.9**	**98.6**	**99.1**	**76.4**	**99.8**	**90.0**
Proposed 03 CERE	**99.4**	**94.1**	**99.8**	**97.7**	**99.3**	**76.9**	**99.9**	**96.0**

Proposed 1: 11 records for VEB, and 14 records for SVEB
Proposed 2: 24 records
Proposed 3: 44 records

obtain 98.2% for VEB and 78.0% for SVEB, in addition to successes in the other cases. The present results found across different methods outperform the state-of-the-art techniques in terms of Acc, Sen, Spe and Ppr in VEB and SVEB as shown in Table 10.

5) SCENARIO 5

In this scenario, INCART dataset was used as a test set while the training was DS1. Table 1 shows the distribution classes for this testing data set. Figure 10 shows example of the original input signal for the classes (N, S, V, and F) Fig. 11. shows an example of the image produced from the generator network for the input classes. Fig. 12 shows an example of the intermediate feature-maps from the discriminator network for these classes, taken from the output of the second dense block. Fig. 13 shows an example of the final feature-maps (after the global average pooling, but before the Softmax in Fig. 3) which are used for classification. Table 11 shows

TABLE 11. Classification performance of the proposed method in the term of (TN, FN, TP, and FP) on incart as test set.

Method	N	S	V	F
T_N	22168	173765	155713	175493
F_N	1316	438	1443	184
T_P	152229	1510	18557	36
F_P	1720	374	1157	130

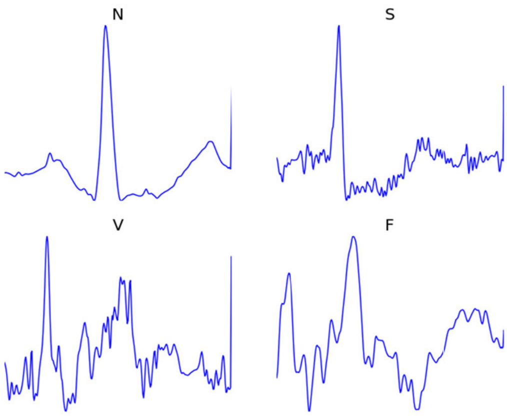


FIGURE 10. Example of original input signal for the classes N, S, V, and F (INCART).

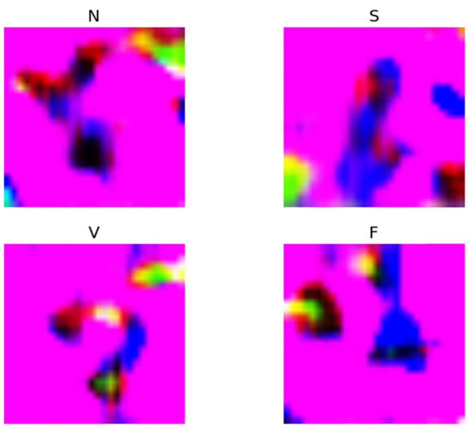


FIGURE 11. Example of the generator's output images for the classes N, S, V, and F (INCART).

TABLE 12. VEB and SVEB classification performance for scenario 5 vs. methods in the literature.

Methods	VEB				SVEB			
	Acc	*Sen*	*Spe*	*Ppr*	*Acc*	*Sen*	*Spe*	*Ppr*
Mariano et al [66]	N/A	**0.84**	N/A	**0.94**	N/A	**0.76**	N/A	**0.07**
Rahhal et al [45]	**0.82**	**0.75**	**0.83**	**0.37**	**0.92**	**0.15**	**0.93**	**0.02**
Proposed	**0.99**	**0.93**	**0.99**	**0.94**	**0.99**	**0.78**	**0.99**	**0.80**

the classification results in terms of (T_N, F_N, T_P, and F_P). The method proposed in [66] achieved the highest sensitivity score of 84% for class V (VEB), and 77% for class S (SVEB). The proposed method obtains 93% for VEB and 78% for SVEB. The present results found across different methods is shown in Table 12.

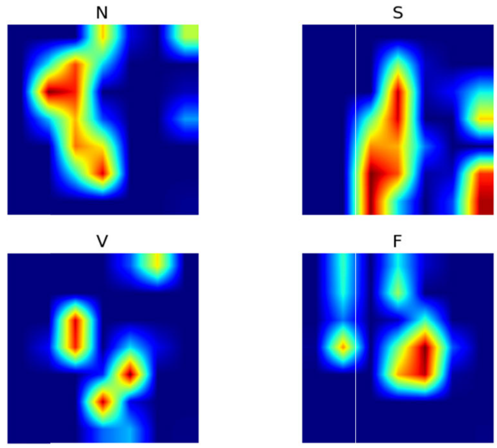


FIGURE 12. Example of feature maps form the discriminator (dense block 2 output) for the classes N,S,V, and F (INCART).

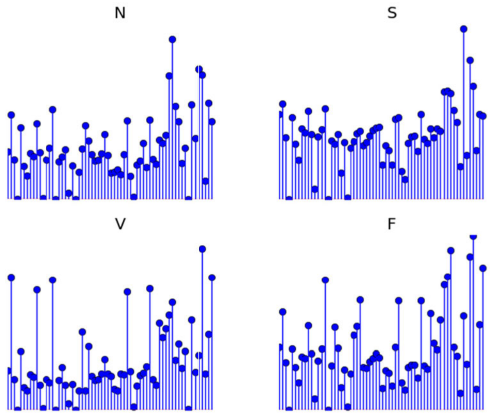


FIGURE 13. Example of the final features vector used for classification form the discriminator for the classes N,S,V, and F (INCART).

TABLE 13. Classification performance of the proposed method in the term of (TN, FN, TP, AND FP) on SVDB as test set.

Method	N	S	V	F
T_N	19037	153740	156192	164450
F_N	7045	3473	1208	22
T_P	138391	7260	7073	1
F_P	2574	6714	2460	0

TABLE 14. VEB and SVEB classification performance for scenario 6 vs. methods in the literature.

Methods	VEB				SVEB			
	Acc	*Sen*	*Spe*	*Ppr*	*Acc*	*Sen*	*Spe*	*Ppr*
Mariano et al [66]	N/A	***0.79***	N/A	***0.49***	N/A	***0.51***	N/A	***0.46***
Rahhal et al [45]	***0.66***	***0.65***	***0.66***	***0.09***	***0.90***	***0.08***	***0.96***	***0.14***
Proposed	***0.97***	***0.85***	***0.98***	***0.79***	***0.74***	***0.67***	***0.95***	***0.51***

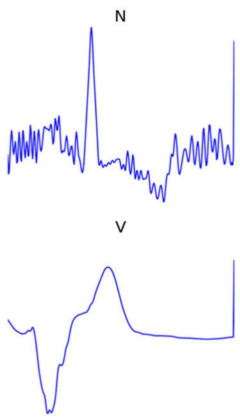


FIGURE 14. Example of orginal input signal for the classes N, S, V, and F (SVDB).

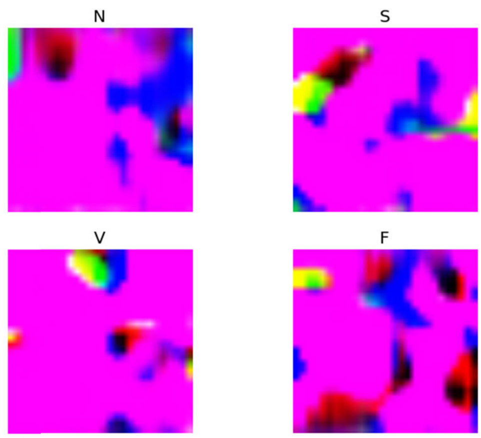


FIGURE 15. Example of the generator's output images for the classes N, S, V, and F (SVDB).

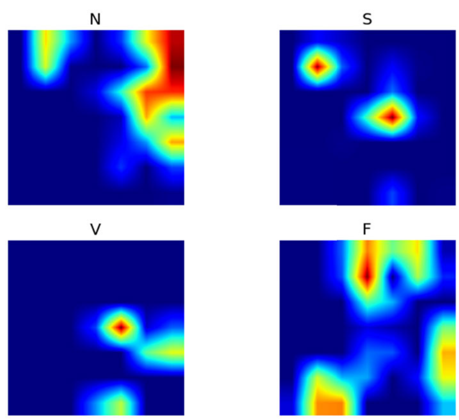


FIGURE 16. Example of feature maps form the discriminator (dense block 2 output) for the classes N,S,V, and F (SVDB).

6) SCENARIO 6

In this scenario SVDB dataset was also used to test the generalization of the method. Table 1 shows the distribution classes for this testing data set. Fig. 14 shows example of original

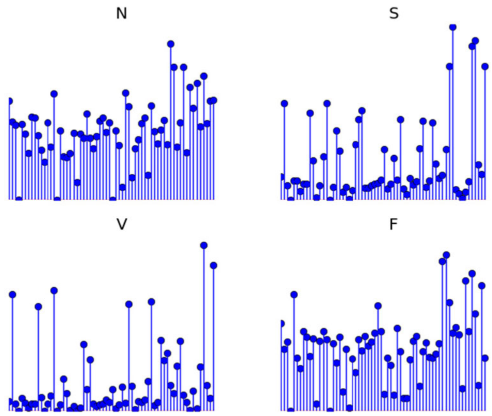


FIGURE 17. Example of the final features vector used for classification form the discriminator for the classes N,S,V, and F (SVDB).

input signal for the classes (N, S, V, and F) it is clear that the input signal is a noisy signal. Fig. 15 shows an example of the image produced from the generator network for the input classes. Fig. 16 shows an example of the intermediate feature-maps from the discriminator network for these classes, taken from the output of the second dense block. Fig. 17 shows an example of the final feature-maps.

Tables 13 shows the automatic classification results in terms of (T_N, F_N, T_P, *and* F_P) Table 14 presents the results in the terms of VEB and SVEB samples, the (Acc, Sen, Spe, and Ppr) are (97%, 85%, 98%, and 79%) and (74%, 67%, 98%, and 51%) for VEB and SVEB respectively. Table 14 confirms that the obtained results are better than stat of the methods.

IV. CONCLUSION

Our experiments show that by converting the raw 1D ECG signal data into a 2D image using a generative neural network the image can be easily fed into a state of the art convolutional neural network such as DenseNet. This produces a highly accurate classification ability, with high sensitivity and specificity. Using 4 classes N, S, V, F as well as focal loss to deal with the shortcoming of data balance performed better than oversampling the minority classes or using cross-entropy loss. When classifying V versus the other three classes and S versus the other three classes, the proposed method outperforms the state-of-the-art methods in terms of accuracy, sensitivity, and specificity.

REFERENCES

[1] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ECG complexes using Hermite functions and self-organizing maps," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 838–848, Jul. 2000.

[2] J. Rodriguez, A. Goni, and A. Illarramendi, "Real-time classification of ECGs on a PDA," *IEEE Trans. Inf. Technol. Biomed.*, vol. 9, no. 1, pp. 23–34, Mar. 2005.

[3] H. F. Huang, G. S. Hu, and L. Zhu, "Sparse representation-based heartbeat classification using independent component analysis," *J. Med. Syst.*, vol. 36, no. 3, pp. 1235–1247, Jun. 2012.

[4] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 7, pp. 1196–1206, Jul. 2004.

[5] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2535–2543, Dec. 2006.

[6] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ECG signal classification," *IEEE Trans. Neural Netw.*, vol. 18, no. 6, pp. 1750–1761, Nov. 2007.

[7] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1415–1426, May 2009.

[8] C. Ye, B. V. K. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2930–2941, Oct. 2012.

[9] K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 179–185, Feb. 1999.

[10] O. T. Inan, L. Giovangrandi, and G. T. A. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2507–2515, Dec. 2006.

[11] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 192–202, Feb. 1999.

[12] D. A. Coast, R. M. Stern, G. G. Cano, and S. A. Briller, "An approach to cardiac arrhythmia analysis using hidden Markov models," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 826–836, Sep. 1990.

[13] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 4, pp. 582–589, Apr. 2004.

[14] Y. H. Hu, W. J. Tompkins, J. L. Urrusti, and V. X. Afonso, "Applications of artificial neural networks for ECG signal detection and classification," *J. Eletrocardiol.*, vol. 26, pp. 66–73, Jan. 1990.

[15] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 9, pp. 891–900, Sep. 1997.

[16] C. Ye, B. V. K. V. Kumar, and M. T. Coimbra, "An automatic subject-adaptable heartbeat classifier based on multiview learning," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 6, pp. 1485–1492, Nov. 2016.

[17] L. B. Marinho, N. D. M. M. Nascimento, J. W. M. Souza, M. V. Gurgel, P. P. R. Filho, and V. H. C. de Albuquerque, "A novel electrocardiogram feature extraction approach for cardiac arrhythmia classification," *Future Gener. Comput. Syst.*, vol. 97, pp. 564–577, Aug. 2019.

[18] Z. Chen, "An energy-efficient ECG processor with weak-strong hybrid classifier for arrhythmia detection," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 7, pp. 948–952, Jul. 2018.

[19] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 48, no. 12, pp. 2095–2104, Dec. 2018.

[20] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.

[21] K. Luo, J. Li, Z. Wang, and A. Cuschieri, "Patient-specific deep architectural model for ECG classification," *J. Healthcare Eng.*, vol. 2017, May 2017, Art. no. 4108720.

[22] P. Li, Y. Wang, J. He, L. Wang, Y. Tian, T.-S. Zhou, T. Li, and J.-S. Li, "High-performance personalized heartbeat classification model for long-term ECG signal," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 78–86, Jan. 2017.

[23] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.

[24] X. Sun, N. M. Nasrabadi, and T. D. Tran, "Task-driven dictionary learning for hyperspectral image classification with structured sparsity constraints," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4457–4471, Aug. 2015.

[25] J.-C. Li, W. W. Y. Ng, D. S. Yeung, and P. P. K. Chan, "Bi-firing deep neural networks," *Int. J. Mach. Learn. Cybern.*, vol. 5, no. 1, pp. 73–83, Feb. 2014.

[26] J. Zhang, S. Ding, N. Zhang, and Z. Shi, "Incremental extreme learning machine based on deep feature embedded," *Int. J. Mach. Learn. Cybern.*, vol. 7, no. 1, pp. 111–120, Feb. 2016.

[27] M. M. A. Rahhal, N. A. Ajlan, Y. Bazi, H. A. Hichri, and T. Rabczuk, "Automatic premature ventricular contractions detection for multi-lead electrocardiogram signal," in *Proc. IEEE Int. Conf. Electro/Inf. Technol. (EIT)*, May 2018, pp. 0169–0173.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Proc. Comput. Vis. (ECCV)*, 2014, pp. 346–361.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[30] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2012.

[31] S. Ren, K. He, R. Girshick, X. Zhang, and J. Sun, "Object detection networks on convolutional feature maps," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1476–1481, Jul. 2017.

[32] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 142–158, Jan. 2016.

[33] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[34] W. Ouyang, X. Zeng, X. Wang, S. Qiu, P. Luo, Y. Tian, H. Li, S. Yang, Z. Wang, H. Li, K. Wang, J. Yan, C.-C. Loy, and X. Tang, "DeepID-Net: Object detection with deformable part based convolutional neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 7, pp. 1320–1334, Jul. 2017.

[35] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, Aug. 2013.

[36] W. Sun and R. Wang, "Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 474–478, Mar. 2018.

[37] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, and H. Larochelle, "Brain tumor segmentation with deep neural networks," *Med. Image Anal.*, vol. 35, pp. 18–31, Jan. 2017.

[38] X. W. Gao, R. Hui, and Z. Tian, "Classification of CT brain images based on deep learning networks," *Comput. Methods Programs Biomed.*, vol. 138, pp. 49–56, Jan. 2017.

[39] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein, and A. Biller, "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016.

[40] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. Rueckert, B. Glocker, and D. K. Menon, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Med. Image Anal.*, vol. 36, pp. 61–78, Feb. 2017.

[41] Z. Jiao, X. Gao, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification," *Neurocomputing*, vol. 197, pp. 221–231, Jul. 2016.

[42] W. Sun, T. B. Tseng, J. Zhang, and W. Qian, "Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data," *Comput. Med. Imag. Graph.*, vol. 57, pp. 4–9, Apr. 2017.

[43] F. Jiang *et al.*, "Abdominal adipose tissues extraction using multi-scale deep neural network," *Neurocomputing*, vol. 229, pp. 23–33, Mar. 2017.

[44] C. Spampinato, S. Palazzo, D. Giordano, M. Aldinucci, and R. Leonardi, "Deep learning for automated skeletal bone age assessment in X-Ray images," *Med. Image Anal.*, vol. 36, pp. 41–51, Feb. 2017.

[45] M. M. Al Rahhal, Y. Bazi, H. AlHichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Inf. Sci.*, vol. 345, pp. 340–354, Jun. 2016.

[46] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 48, no. 12, pp. 2095–2104, Dec. 2018.

[47] M. M. Al Rahhal, Y. Bazi, and M. Al Zuair, *Convolutional Neural Networks for Electrocardiogram Classification | SpringerLink*. Accessed: Nov. 26, 2018. [Online]. Available: https://link.springer.com/article/10.1007/s40846-018-0389-7

[48] S. M. Mathews, C. Kambhamettu, and K. E. Barner, "A novel application of deep learning for single-lead ECG classification," *Comput. Biol. Med.*, vol. 99, pp. 53–62, Aug. 2018.

[49] G. Sannino and G. De Pietro, "A deep learning approach for ECG-based heartbeat classification for arrhythmia detection," *Future Gener. Comput. Syst.*, vol. 86, pp. 446–455, Sep. 2018.

[50] S. Zhou and B. Tan, "Electrocardiogram soft computing using hybrid deep learning CNN-ELM," *Appl. Soft Comput.*, Sep. 2019, Art. no. 105778.

[51] S. S. Xu, M.-W. Mak, C.-C. Cheung, "Towards end-to-end ECG classification with raw signal extraction and deep neural networks," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 4, pp. 1574–1584, Jul. 2019.

[52] L. Guo, G. Sim, and B. Matuszewski, "Inter-patient ECG classification with convolutional and recurrent neural networks," *Biocybern. Biomed. Eng.*, vol. 39, no. 3, pp. 868–879, Jul. 2019.

[53] Y. Li, Y. Pang, J. Wang, and X. Li, "Patient-specific ECG classification by deeper CNN from generic to dedicated," *Neurocomputing*, vol. 314, pp. 336–346, Nov. 2018.

[54] J. Huang, B. Chen, B. Yao, and W. He, "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network," *IEEE Access*, vol. 7, pp. 92871–92880, 2019.

[55] M. Kim, H. Ko, and S. B. Pan, "A study on user recognition using 2D ECG based on ensemble of deep convolutional neural networks," *J. Ambient Intell. Hum. Comput.*, 2019, doi: 10.1007/s12652-019-01195-4.

[56] G. Huang, Z. Liu, L. van der Maaten and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261–2269.

[57] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," Feb. 2015, *arXiv:1502.03167*. [Online]. Available: https://arxiv.org/abs/1502.03167

[58] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks," in *Proc. PMLR*, 2011, pp. 315–323.

[59] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[60] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[61] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, 2017, pp. 2999–3007.

[62] A. L. Goldberger, "PhysioBank, PhysioToolkit, and PhysioNet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[63] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May 2001.

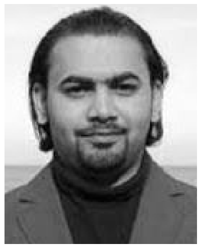
[64] M. M. A. Rahhal, "Classification of AAMI heartbeat classes with an interactive ELM ensemble learning approach," *Biomed. Signal Process. Control*, vol. 19, pp. 56–67, May 2015.

[65] J. Jiang, H. Zhang, D. Pi, and C. Dai, "A novel multi-module neural network system for imbalanced heartbeats classification," *Expert Syst. Appl. X*, vol. 1, Apr. 2019, Art. no. 100003.

[66] M. Llamedo and J. P. Martínez, "An automatic patient-adapted ECG heartbeat classifier allowing expert assistance," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2312–2320, Aug. 2012.

MOHAMAD MAHMOUD AL RAHHAL received the B.Sc. degree in computer engineering from Aleppo University, Aleppo, Syria, in 2002, the M.Sc. degree from Hamdard University, New Delhi, India, in 2005, and the Ph.D. degree in computer engineering from King Saud University, Riyadh, Saudi Arabia, in 2015. From 2006 to 2012, he was a Lecturer at Al-Jouf University, Saudi Arabia. He is currently an Associate professor with the College of Applied Computer Engineering, King Saud University, Riyadh. His research interests include signal/image medical analysis, remote sensing, and computer vision.

YAKOUB BAZI (S'05–M'07–SM'10) received the State Engineer and M.Sc. degrees in electronics from the University of Batna, Batna, Algeria, in 1994 and 2000, respectively, and the Ph.D. degree in information and communication technology from the University of Trento, Trento, Italy, in 2005. From 2000 to 2002, he was a Lecturer with the University of M'sila, M'sila, Algeria. From January to June 2006, he was a Postdoctoral Researcher with the University of Trento. From August 2006 to September 2009, he was an Assistant Professor with the College of Engineering, Al Jouf University, Al Jouf, Saudi Arabia. He is currently an Associate Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His current research interests include remote sensing, signal/image medical analysis, and computer vision. He is a referee for several international journals. He is also an Associate Editor of the IEEE Geoscience and Remote Sensing Letters, and IEEE Access.



HAIDAR ALMUBARAK (M'03) was born in Al Ahsa, Saudi Arabia, in 1983. He received the B.S. degree in computer engineering from the King Fahd University of Petroleum and Minerals, Saudi Arabia, in 2005, the M.S degree in computer, information, and network security from DePaul University, Chicago, IL, USA, in 2011, and the Ph.D. degree in computer engineering from the Missouri University of Science and Technology, Rolla, MO, USA, in 2018. In 2005, he joined SAAD Hospital, as a Computer Engineer, and in 2008, he moved to SABIC, as a Systems Engineer, before moving to USA to finish his M.S. and Ph.D. degrees. He joined King Saud University, in 2019, as a Postdoctoral Fellow. His current research interests include computer vision, applied machine/deep learning, and medical image analysis

NAIF ALAJLAN (M'11–SM'13) received the B.Sc. (Hons.) and M.Sc. degrees in electrical engineering from King Saud University, Riyadh, Saudi Arabia, in 1998 and 2003, respectively, and the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2006. He was a Systems and Control Engineer with Saudi Basic Industries Company (SABIC), Riyadh, from 1998 to 2000. Then he joined the Electrical Engineering Department, King Saud University, where he was a Lecturer, from 2000 to 2003, and an Assistant Professor, from 2007 to 2010. He is currently a Professor with the Computer Engineering Department, King Saud University. He is also the Founder and the Director of the Advanced Lab for Intelligent Systems Research, King Saud University. He has authored or coauthored more than 50 journals articles (some with high impact factors) and 25 conference papers. His current research interests include shape retrieval, machine learning, pattern recognition, and remote sensing.

MANSOUR AL ZUAIR received the B.S. degree in computer engineering from King Saud University and the M.S. and Ph.D. degrees in computer engineering from Syracuse University. He is currently an Associate Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He served as the CEN Chairman, from 2003 to 2006, the Vice Dean, from 2009 to 2015, and has been the Dean, since 2016. His research interests are in the areas of computer architecture, computer networks, and signal processing.