Article

A Parallel Cross Convolutional Recurrent Neural Network for Automatic Imbalanced ECG Arrhythmia Detection with Continuous Wavelet Transform

Tabassum Islam Toma and Sunwoong Choi *

School of Electrical Engineering, Kookmin University, Seoul 02707, Korea
* Correspondence: schoi@kookmin.ac.kr

Abstract: Automatic detection of arrhythmia using electrocardiogram (ECG) and deep learning (DL) is very important to reduce the global death rate from cardiovascular diseases (CVD). Previous studies on automatic arrhythmia detection relied largely on various ECG features and have achieved considerable classification accuracy using DL-based models. However, most previous research has ignored multi-class imbalanced problems in ECG arrhythmia detection. Therefore, it remains a challenge to improve the classification performance of the DL-based models. This paper proposes a novel parallel cross convolutional recurrent neural network in order to improve the arrhythmia detection performance of imbalanced ECG signals. The proposed model incorporates a recurrent neural network and a two-dimensional (2D) convolutional neural network (CNN) and can effectively learn temporal characteristics and rich spatial information of raw ECG signals. Continuous wavelet transform (CWT) is used to transform the ECG signals into a 2D scalogram composed of time–frequency components, and subsequently, the 2D-CNN can learn spatial information from the 2D scalogram. The proposed model is not only efficient in learning features with imbalanced samples but can also significantly improve model convergence with higher accuracy. The overall performance of our proposed model is evaluated based on the MIT-BIH arrhythmia dataset. Detailed analysis of evaluation metrics reveals that the proposed model is very effective in arrhythmia detection and significantly better than the existing hierarchical network models.

Keywords: cardiac arrhythmia; convolutional neural network; continuous wavelet transform; electrocardiogram; recurrent neural network

1. Introduction

Cardiovascular diseases (CVD) have become a global concern because of their deadly consequences, either death or disability which severely affects people’s life. A report published by the World Health Organization (WHO) manifests that approximately 17.9 million people died from CVD in 2019, accounting for 30% of global deaths [1]. Recently, WHO has warned mass people announcing CVD as the top ten health threats in the world and it is estimated that the death toll caused by CVD will be increased to 23 million by 2030. Cardiac arrhythmia, which causes heart rhythmic problems, is the fundamental disease group of CVD. Due to coronary artery disease, high blood pressure, changes in heart muscle, and valve disorders, electrical pulses that coordinate heartbeats can cause the heart to beat irregularly, resulting in cardiac arrhythmia [2]. According to reports, some forms of arrhythmia, such as ventricular fibrillation and tachycardia, are outrageous and life-threatening and can trigger cardiac arrest and sudden death. Therefore, beat-by-beat examination utilizing electrocardiogram (ECG) has become a vital modern medical tool for early diagnosis of cardiovascular conditions, such as the process of cardiac excitability, transmission, and recovery.

The ECG is used to record the electrical activities and states of the heart over time through electrodes attached to the skin surface [3]. Thus, the presence of abnormal heart
electrical activity can be detected by analyzing the ECG, which can assist to derive the
type of arrhythmia in turn. Previously, the ECG analysis as well as arrhythmia recognition
and the diagnosis was performed with the help of human intervention. However, this
procedure is very laborious and time-consuming [4]. Recently, the emergence of intelli-
gent healthcare requires the incorporation of artificial intelligence to automatically detect
irregular heart rhythms from the ECG. Over the past decade, many researchers have investi-
gated and developed several schemes for automatic arrhythmia detection [5–10]. Although
these developed schemes could achieve higher accuracy on the ECG signal classification,
researchers also investigated their performances by introducing minor variations in the
machine learning (ML) classifier and fusion of different hand-crafted features in [11–13].
A disease specific feature-based ECG heartbeat classification procedure was developed
in [11]. A one-versus-one rule based binary support vector machine (SVM) classifier was
used to perform classification. In [12], a novel ECG heartbeat classification method was
proposed that combines morphological features and RR interval information. An SVM
classifier was used to perform the classification in this paper. A combination of projected
and dynamic features was used to classify the ECG heartbeat in [13]. A random projection
matrix was employed to extract projected features, while three weighted RR intervals
were considered as the dynamic features. Finally, using those two kinds of features, an
SVM classifier was used to cluster heartbeats into one of fifteen or five classes. However,
these schemes are based on ML algorithms (e.g., neural network, SVM, etc.), with various
features, including hand-crafted time and frequency features extracted empirically, which
require extensive professional knowledge. In addition, it becomes more difficult to select
the model parameter when the feature dimension increases.

With the development of the deep learning (DL) algorithm, it has become predom-
nant in the ECG classification because of its exemplary performance, especially that of
convolutional neural network (CNN) using one-dimensional (1D) and two-dimensional
(2D) convolution. The ability of DL-based models to automatically learn invariant and
hierarchical features directly from data has increased their applicability in designing end-
to-end learning mechanisms that take data as input and class prediction as output. Recent
DL-based models have been developed depending on the 1D ECG signal or 2D representa-
tion of ECG by transforming the ECG signal into images or some matrix form. Numerous
studies on 1D ECG classification have recently been conducted [14–20]. The authors demon-
strated a performance comparison among five models: 1D-CNN, CNN combined with
linear discriminant analysis, SVM classifier, CNN combined with SVM, and CNN combined
with SVM and RR interval input in [14]. A 1D-CNN classifier was designed to identify
atrial fibrillation in [15]. Before using the ECG recordings as input, Nurmaini et al. [15] used
the discrete wavelet transform to remove noise from the ECG recordings. In [16], Acharya
et al. developed a typical nine-layer 1D-CNN model to classify five categories of heartbeats,
i.e., non-ectopic, supraventricular ectopic, ventricular ectopic, fusion, and unknown beats.
The authors performed filtering, R-peak extraction, and heartbeat segmentation before
feeding the data into the 1D-CNN classifier to identify heartbeat categories. In another
study, an 11-layer 1D-CNN model was proposed to detect five types of arrhythmias, and
its overall performance was investigated by Acharya et al. [17]. At this time, their exper-
iment included ECG recording durations of 2 and 5 s instead of heartbeats. However,
adopting a hybrid model and variations of CNN can further improve the performance
of the classic CNN. The authors of [18] proposed a neural network-based method that
involves two steps for precise classification of heartbeats by following the Association
for the Advancement of Medical Instrumentation (AAMI) inter-patient standards. Data
preprocessing and features extraction were the task of the first step, while the classification
was performed by a two-layer classifier in which each layer consists of two independent
fully connected neural networks in the second step. The experimental results depicted that
the proposed model precisely detects arrhythmia conditions. In [19], Chen et al. proposed a
DL-based hybrid model that combines CNN and long short-term memory (LSTM) network
to classify heartbeats. Their model requires the input of two forms of 1D data from the ECG
signals, namely the ECG heartbeats that provide the morphological information and the adjacent segment of the ECG beats that enrich temporal information. In another study [20], a complex DL-based model consisting of a CNN and an LSTM network was developed to detect six types of ECG signals by processing 10 s ECG slices of the MIT-BIH arrhythmia dataset. Experimental results validated that the proposed model was very effective in automatically detecting arrhythmia. However, the outstanding performance of 2D-CNN on image classification or pattern recognition attracted researchers to convert raw ECG data to images and utilize them for precise arrhythmia detection. Fourier transform is a well-known and widely used method for extracting time–frequency components from non-stationary ECG signals. [21] studied the performance of time–frequency spectrogram-based 2D-CNN model on ECG classification. In this study, the authors utilized the short-time fourier transform method to convert ECG signals into time–frequency spectrograms. An elaborate exhibition of the experiment proves that 2D-CNN can achieve higher classification accuracy than 1D-CNN. In [22], spectro-temporal images were used to train and optimize a dense CNN. The developed model in [22] can achieve significantly improved performance by capturing both beat-to-beat and single-beat information. Another widely accepted technique to generate time–frequency representation from a given segment of the ECG is wavelet transform (WT). The authors in [23] adopted WT to transform heartbeat time intervals of ECG signals into images and then trained a six-layer CNN employing these images for heartbeat classification. Very recently, a generative adversarial neural network has become popular in the field of image super-resolution and image reconstruction. The capability of image reconstruction to convert the raw 1D ECG signal data into a 2D image has been exploited in [24]. Afterward, the authors fed these images as input to DenseNet which produces highly accurate classification with high sensitivity and specificity using four classes. In [25], treating multi-lead ECG as 2D matrices for input, a novel model called multi-lead-CNN (ML-CNN) was designed, which employs sub-2D convolutional layers and lead asymmetric pooling layers. Incorporating both hand-crafted RR interval and spatial features extracted by CNN from the 2D scalogram, a novel method was developed in [26] for AAMI standard-based ECG heartbeat classification. In this study, their method shows poor performance for the highly skewed class.

Despite the large number of studies in the area of arrhythmia detection exploiting the ECG signal and DL-based models, there are still many challenges because of the differences in the recording environment, variations of disease patterns among the subjects during testing, complex, nonstationary, and noisy nature of the ECG signal. Moreover, the aforementioned methods ignored multi-class imbalanced problems in arrhythmia detection. Imbalanced data learning [27] has always been a concerning issue in the field of ML, and the adverse effect of data imbalance on the performance of the learning algorithms is not negligible. Data imbalance issues also exist in the field of arrhythmia detection from the ECG data analysis [28]. Therefore, solving the imbalance problem can improve the overall performance. Considering the data imbalance problem, [28] proposed a novel multi-module neural network system AAMI standard-based ECG heartbeat classification. To tackle the imbalance problem, the authors introduced the borderline-SMOTE algorithm, a novel context feature module, and two-phase training (referred to as 2PT). In [29], a transfer learning-based heartbeat classification method utilizing 2D time–frequency diagrams was proposed. The authors employed a hybrid time–frequency analysis of the Hilbert transform and the Wigner–Ville distribution to convert 1D ECG recordings into 2D time–frequency diagrams.

However, it can be noticed from the aforementioned studies ([19,20,22,26,28]) that multi-type feature fusion, while training the DL-based model, can lead to improved performance. In a recent work [30], the authors developed a parallel cross CNN to improve the detection performance of imbalanced traffic flows by fusing the flow features learned from two-branch CNNs. Motivated by this fact, this paper proposes a novel parallel cross-convolutional recurrent neural network to improve the arrhythmia detection performance of imbalanced ECG signals. The proposed model consists of two branches: a recurrent neural network (RNN) and a 2D CNN for temporal characteristics and spatial features.
Three LSTM layers and three convolutional blocks are used to form the top and bottom branches of the model, respectively. After each block and layer operation, the features are fused and passed through the subsequent layers and blocks. Hence, the model becomes more distinguishable and robust for imbalanced ECG signal classification, especially for the AAMI standard-based classes, which are largely skewed. The main contributions of this paper are summarized as follows:

- Considering ECG arrhythmia detection as a multi-class imbalance problem, this paper proposes a novel parallel cross convolutional RNN classifier to improve detection performance.
- The proposed model includes two branches consisting of the LSTM network and CNN to capture both temporal characteristics from the ECG segment and spatial characteristics from the 2D scalogram. The network becomes more distinguishable and robust through cross-fusion of the features.
- Continuous wavelet transform (CWT) has been adopted to convert the preprocessed ECG segments into 2D scalograms. Subsequently, both ECG segment and corresponding 2D scalogram have been used as input for training and testing the proposed model.
- Finally, we have evaluated the performances of the proposed model on the MIT-BIH arrhythmia dataset. We also compared the performance with other existing works and investigated the performance by varying the wavelet function, such as Gaussian wavelets of order 8 (gaus8) and order 4 (gaus4), Mexican hat wavelet (mexh), Morlet wavelet (morl).

The remainder of this paper is structured as follows: Section 2 describes the overall methodology of ECG arrhythmia detection along with the details of data preprocessing and the proposed model used in this paper. In Section 3, the experimental results followed by implementation details have been elaborately illustrated. Finally, we draw the conclusions of this paper in Section 4.

2. Methodology

This section describes ECG arrhythmia detection exploiting a deep neural network if a set of ECG samples and their corresponding arrhythmia classes are given. Figure 1 depicts the architecture of the proposed model for ECG arrhythmia classification method. This method includes several steps before enabling the proposed model to use ECG samples for training. Each essential component of arrhythmia detection is elaborately demonstrated in the following sections.

![Figure 1. The architecture of arrhythmia recognition and classification.](image)

2.1. ECG Signal Preprocessing and Segmentation

When raw ECG data are collected in the clinical environment, the data contains various types of noise, including baseline wandering, electromyogram signals, and power-line disturbance [26]. Without removing the noise contamination, the extracted features may lead to performance degradation. Therefore, filtering is required to denoise the ECG signal before
segmentation. However, attention should be provided when filtering is applied, otherwise the ECG signal may lose useful information. In this study, the baseline wandering technique has been adopted for noise removal. To obtain the baseline, two consecutive filters, including a 200 ms width median filter followed by a 600 ms median filter, are utilized. Next, the computed baseline is subtracted from the raw ECG signal. Various studies show that the median filter is effective in outlier elimination without distorting the phase of the signal compared with other filter techniques, such as regular infinite impulse response and finite impulse response. Figure 2 depicts the ECG signal filtering step.

![ECG signal denoising](image)

Figure 2. ECG signal denoising (a) before filtering and (b) after filtering.

The ECG signal segmentation involves the extraction of a complete heartbeat from the ECG signal. This step usually requires the detection of position of R-peak because R-peak is regarded as an anchor point for locating a complete heartbeat [31]. In this study, no algorithm is adopted for R-peak detection rather we have exploited the annotated R-peak location to segment the ECG signal into a series of heartbeats. Figure 3 depicts an example of a heartbeat segmentation from the ECG signal. The preceding and succeeding time lengths, $\Delta t_1$ and $\Delta t_2$, respectively, of any R-peak $R_i$ help determine the length of an ECG segment. Consequently, the number of data samples involved in each segment can be calculated as follows:

$$N_{\text{sample}} = (\Delta t_1 + \Delta t_2) \times f + 1$$  \hspace{1cm} (1)

![ECG heartbeat segmentation](image)

Figure 3. ECG heartbeat segmentation.
For each heartbeat segment, we have obtained $\Delta t_1 = 0.3$ s and $\Delta t_2 = 0.5$ s in this study.

2.2. Scalogram Obtained Using CWT

The ECG signal is nonstationary data composed of time-varying frequency components. Therefore, time–frequency representation of an ECG segment can be a useful tool for arrhythmia detection. STFT is one of the most popular and widely used techniques for transforming the ECG signal into a 2D time–frequency image. The STFT uses a fixed-size window function when it converts the ECG signal into time–frequency image. However, choosing the size of the window function often limits its effectiveness in many areas [32]. To overcome this limitation, CWT has been explored in different application areas in recent years. Unlike the STFT method, CWT can choose a smooth analytical mother wavelet that can identify the dynamic frequency properties of ECG signals at different scales.

In addition, CWT can provide high-time resolution and low-frequency resolution in high frequencies, as well as high-frequency resolution and low-time resolution in low frequencies, by adjusting the scale and translation parameters [33]. Given an ECG signal $x(t)$, the CWT is defined as the inner product between the signal and the mother wavelet $\psi(t)$ that yields the 2D matrix of wavelet coefficients $C_{a,b}$ as follows:

$$C_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi \left( \frac{t - b}{a} \right) dt$$  \hspace{1cm} (2)

where $a$ represents the scale parameter, $b$ is a time-shifting or translation parameter, and $\psi(t)$ is the mother wavelet function. The scale can be replaced by frequency as follows:

$$F = \frac{f_c \times f_s}{a}$$  \hspace{1cm} (3)

where $f_c$ and $f_s$ represent the center frequency of the mother wavelet and the sampling frequency of the signal $x(t)$, respectively. However, choosing the mother wavelet function is crucial because, depending on the application areas, the classifier’s performance largely relies on the mother wavelet function. In this paper, four wavelet functions, namely $gaus8$, $gaus4$, $mexh$, and $morl$, have been employed to generate a 2D scalogram and to distinctly investigate the performance of the proposed model. The mathematical expression for $gaus8$, $gaus4$, $mexh$, and $morl$ wavelet functions can be represented as follows [34–36]:

$$\psi_{gaus}(mt) = C_m \frac{d^m}{dt^m} g_c(t)$$  \hspace{1cm} (4)

$$\psi_{mexh}(t) = e^{2\pi it} e^{\frac{-t^2}{2\sigma^2}}$$  \hspace{1cm} (5)

$$\psi_{morl}(t) = \left(1 - \frac{t^2}{\sigma^2}\right) e^{\frac{-t^2}{2\sigma^2}}$$  \hspace{1cm} (6)

where $\psi_{gaus}(mt)$, $\psi_{mexh}(t)$, and $\psi_{morl}(t)$ denote Gaussian wavelet, $mexh$, and $morl$, respectively. The parameter $m$ denotes the order of the Gaussian wavelet and $g_c(t)$ denotes Gabor function, which is the product of the Gaussian function and the complex exponential function. Figure 4 shows the four types of wavelet functions.
Sensors 2022, 22, x FOR PEER REVIEW 7 of 17

where \( f \) which includes memory cells to overcome the vanishing gradient of the RNN and is we have used the LSTM layer to construct the top branch. LSTM is a variant of RNN, particularly suitable for processing time series signals [37]. There are three LSTM layers in the addition of information to LSTM’s current state, and \( i(t) \) controls the update of the \( h(t) \)

\[
f(t) = \sigma\left(\omega_f \cdot [h(t-1), x(t)] + \beta_f\right)
\]

\[
i(t) = \sigma\left(\omega_i \cdot [h(t-1), x(t)] + \beta_i\right)
\]

\[
o(t) = \sigma\left(\omega_o \cdot [h(t-1), x(t)] + \beta_o\right)
\]

\[
c(t) = f(t) \odot c(t-1) + i(t) \odot \tanh(\omega_c \cdot [h(t-1), x(t)]) + \beta_c
\]

\[
h(t) = o(t) \odot \tanh(c(t))
\]

Figure 4. Continuous wavelet functions. (a) Gaussian of order 8 (gaus8), (b) Gaussian of order 4 (gaus4), (c) Mexican hat (mexh), and (d) Morlet (morl).

2.3. Model Description

The proposed model is composed of two branches, as shown in Figure 5. The top branch comprises several RNN layers to detect temporal changes within a given heartbeat. The RNN layer exploits cell mechanisms to remember the previous pattern of a heartbeat signal and utilize it to capture the next-step temporal characteristics. More specifically, we have used the LSTM layer to construct the top branch. LSTM is a variant of RNN, which includes memory cells to overcome the vanishing gradient of the RNN and is particularly suitable for processing time series signals [37]. There are three LSTM layers in the top branch; each contains 16, 32, and 64 output units, respectively. Each LSTM layer is composed of several LSTM cells. Figure 6 depicts the internal structure of the LSTM cell. For an ECG signal \( x(t) \) at time \( t \), the corresponding output \( h(t) \) and internal state \( c(t) \) can be calculated as follows:

\[
f(t) = \sigma\left(\omega_f \cdot [h(t-1), x(t)] + \beta_f\right)
\]

\[
i(t) = \sigma\left(\omega_i \cdot [h(t-1), x(t)] + \beta_i\right)
\]

\[
o(t) = \sigma\left(\omega_o \cdot [h(t-1), x(t)] + \beta_o\right)
\]

\[
c(t) = f(t) \odot c(t-1) + i(t) \odot \tanh(\omega_c \cdot [h(t-1), x(t)]) + \beta_c
\]

\[
h(t) = o(t) \odot \tanh(c(t))
\]

where \( f(t) \), \( i(t) \), and \( o(t) \) denote the forget, input, and output gates of the LSTM cell, respectively, which are jointly utilized to control the update of the \( h(t) \). More specifically, \( f(t) \) assists in removing useless information contained in LSTM’s past state, \( i(t) \) controls the addition of information to LSTM’s current state, and \( o(t) \) determines the final output. Due to these three gates, the temporal characteristics in the ECG signal can be effectively extracted.
The last layer of each convolutional block is the batch normalization layer. Batch normal-
pooling layer, a dropout layer, and a batch normalization layer. All of the convolutional

Table 1. Configuration of the proposed network architecture.

| Layer          | Output Volume | Layer          | Output Volume |
|----------------|---------------|----------------|---------------|
| Input_1        | (200, 1)      | Input_2        | (100, 100, 1) |
| LSTM_unit_1    | (200, 16)     | Conv_block_1   | (50, 50, 16)  |
| Reshape_1      | (4, 50, 16)   |                |               |
|                | (135, 25, 32) | Concenate_2    |               |
| Reshape_4      | (3375, 32)    |                |               |
| Reshape_3      | (108, 25, 32) |                |               |
| LSTM_unit_2    | (2700, 16)    |                |               |
| Reshape_3      | (200, 16)     |                |               |
| LSTM_unit_2    | (2700, 32)    | Conv_block_2   | (27, 25, 32)  |
| Reshape_3      | (108, 25, 32) |                |               |

Figure 5. Architecture of the proposed parallel cross convolutional recurrent neural network.

Figure 6. Architecture of the LSTM cell internal structure.

The bottom branch is composed of several convolutional blocks. Each block comprises
a 2D convolutional layer with a rectified linear unit (ReLU) activation function, a 2D max-
pooling layer, a dropout layer, and a batch normalization layer. All of the convolutional
layers use $3 \times 3$ 2D convolutional filter to extract spatial features. The number of convo-
lutional filters at each layer is 16, 32, and 64, respectively. A $2 \times 2$ max-pooling layer is
deployed after each convolutional layer to extract more representative features, reducing
the size of feature maps. In addition, to alleviate the bottom branch from overfitting during
training, a dropout layer with a dropout rate of 0.1 is inserted in every convolutional block.
The last layer of each convolutional block is the batch normalization layer. Batch normal-
ization is considered another type of regularization technique that enables a network to
converge faster, avoiding overfitting. The overall configuration of our proposed model is
shown in Table 1.
The proposed model combines high-level spatial features extracted from the CWT scalogram with the temporal characteristics of the upper layers to improve arrhythmia detection results. The model performs three feature fusion operations on the feature maps. The number of output units in the LSTM layers and the number of filters in the convolutional blocks (e.g., 16, 32, 64) are taken in such a way that the pairs of the feature map after each LSTM layer and convolutional block can be concatenated by applying the feature fusion technique. For the first two pairs of the feature map, a cross-fusion operation is adopted so that the fused features can be passed through both convolutional blocks and LSTM layers. The first two fusion operations do not change the size of the feature map, only doubling the number of channels. The last pair of the feature map is flattened before performing a linear fusion operation. The final fused feature is then passed through a dropout layer (with a dropout rate of 0.1) and a dense layer to reduce the size of the fused feature. The network’s output uses the dense and softmax layers to perform multi-class imbalanced ECG arrhythmia classification.

### 3. Experimental Results

This section first describes the ECG dataset and evaluation metrics for evaluating the proposed model performances. Finally, the overall performances, along with the comparison with other existing network performances, have been presented to investigate the effectiveness of the proposed model and unveil its superiority over other existing networks.

#### 3.1. ECG Dataset Description

Although many open-source datasets are available for research in the field of arrhythmia detection, the ECG signal used in this paper is taken from the well-known MIT-BIH arrhythmia database [38]. The database contains 48 fully annotated half-hour two-lead ECG recordings sampled at 360 Hz obtained from 47 subjects. Twenty-three recordings are randomly selected from the subjects’ daily routine. The remaining 25 recordings, containing less common but clinically significant arrhythmias, are also obtained from the above stated ambulatory ECG recordings. Each recording involves two channels of ECG data with 11-bit resolution over a 10-mV range signal. The first channel includes modified lead II (MLII), while the other channel includes lead depending on the recording (e.g., VI, VII, V2, V4, or V5). In the experiment, ECG signals obtained from the MLII channel are used because MLII is ubiquitous in the above records. Two or more cardiologists carefully investigated these recordings and labeled them into 15 arrhythmia types (e.g., normal beat, left bundle branch block beat, right bundle branch block beat, atrial escape beat, nodal (junctional) escape beat, atrial premature beat, aberrant atrial premature, nodal (junctional) premature, premature or supraventricular premature, premature ventricular contraction, ventricular escape, fusion of ventricular and normal, fusion of paced and normal, paced beat, and unclassifiable beat). Recommended by the AAMI standards, all 15 types of ECG heartbeat labels are mapped to 5 AAMI categories, i.e., non-ectopic cardiac beat (NB), supraventricular ectopic beat (SVEB), ventricular ectopic beat (VEB), fusion beat (FB), and unknown beat (QB), as shown in Table 2. However, four ECG recordings (i.e., 102, 104, 107, and 217), which include paced

| Layer           | Output Volume | Layer       | Output Volume |
|-----------------|---------------|-------------|---------------|
| Lstm_unit_3     | (3375, 64)    | Conv_block_3| (68, 13, 64)  |
| Flatten_1       | (216000)      | Flatten_2   | (56576)       |
| Concatenate_3   | (272576)      |             |               |
| Dropout         | (272576)      |             |               |
| Dense_1         | (128)         |             |               |
| Dense_2         | (4)           |             |               |
beats, as well as the QB class, are discarded in this paper, similar to other studies [26,39,40].

As we described above, CWT is applied to transform the segmented ECG heartbeat into a 2D scalogram, Figure 7 illustrates the scalogram, decomposed by gaus8 wavelet function, of four types of ECG heartbeat. A total of 100 continuous values ranging from 2.16 to 216 are used as the scaling factor of the mother wavelet. Each of the scalograms in this study is resized into a (100 × 100) matrix in order to feed into our proposed classifier. The scalograms are represented as a function of time and frequency in the figure.

Table 2. ECG class mapping using the AAMI standard.

| AAMI Classes          | Class Symbol | Heartbeat Types                                      | Training Samples | Test Samples |
|-----------------------|--------------|------------------------------------------------------|------------------|--------------|
| Normal Beat           | NB           | Normal beat (NOR)                                    |                  |              |
|                       |              | Right bundle branch block beat (RBBB)                |                  |              |
|                       |              | Left bundle branch block beat (LBBB)                 |                  |              |
|                       |              | Atrial escape beat (AE)                               |                  |              |
|                       |              | Nodal (junctional) escape beat (NE)                  |                  |              |
| Supraventricular      | SVEB         | Atrial premature beat (AP)                            | 2160             | 540          |
| Ectopic Beat          |              | Premature or ectopic supraventricular beat (SP)      |                  |              |
|                       |              | Nodal premature beat (NP)                             |                  |              |
|                       |              | Aberrated atrial premature beat (AAP)                |                  |              |
| Ventricular Ectopic   | VEB          | Ventricular escape beat (VE)                          | 5600             | 1400         |
| Beat                  |              | Premature ventricular contraction (PVC)              |                  |              |
| Fusion Beat           | FB           | Fusion of ventricular and normal beat (FVN)          | 641              | 160          |

Figure 7. Scalogram of four types of ECG classes based on the AAMI standard (a) NB, (b) SVEB, (c) VEB, and (d) FB.
The spike in the scalograms indicates that there is a change in the frequency of the ECG signal. It is also evident from the depth of the spike that the change in frequency is not similar for all four classes. Therefore, the distinct pattern of these scalograms can be very impactful, and carrying out the classification with these scalograms can lead to satisfactory performances.

3.2. Implementation Details and Performance Metrics

After having the preprocessed ECG heartbeats along with their corresponding 2D scalograms, 80% of each category are selected for the training set, and the remaining 20% are used for the test set. After splitting, the number of total training samples is 80,401, and that of total testing samples is 20,100, including all four AAMI categories. Next, the classifier is trained and optimized using the training data samples and an optimizer. Finally, the performance has been evaluated using the test data samples. All training and testing programs have been performed in an Anaconda Python 3.7 environment on a system equipped with a 3.80 GHz CPU, 256 GB RAM, and a single Nvidia Quadro RTX 6000 GPU. All hyper-parameters used to train the proposed model are summarized in Table 3. However, varying the learning rate \( (lr) \), the performance of the proposed model has been evaluated at the beginning. Keeping the batch size = 512, the performances for three \( lr \) (e.g., 0.0005, 0.0001, and 0.00005) are investigated. \textit{Adam} optimizer, along with categorical-cross entropy loss function, is used during the training of the proposed model. Furthermore, 60 iterations or epochs are used to optimize the network.

### Table 3. Hyper-parameter of the model training.

| Hyper-Parameter           | Values                                      |
|---------------------------|---------------------------------------------|
| Training data shape       | (80401, 100, 100, 1) and (80401, 200, 1)    |
| Test data shape           | (20100, 100, 100, 1) and (20100, 200, 1)    |
| Batch size                | 512                                         |
| Learning rate             | 0.0005, 0.0001, 0.00005                     |
| Epoch                     | 60                                          |
| Optimizer                 | \textit{Adam}                               |
| Loss function             | Categorical-cross entropy                   |

Afterward, the optimum model performance has been compared with those of six existing studies. Finally, the classifier performance is also demonstrated by changing the wavelet functions (e.g., \textit{gaus8}, \textit{gaus4}, \textit{mexh}, and \textit{morl}).

To evaluate the performances, four standard statistical indices, also known as evaluation metrics (e.g., accuracy (ACC), positive predictive value (PPV), sensitivity (SE), and F1-score (F1)), are used in this study. These metrics are defined as follows:

\[
ACC_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (12)
\]

\[
PPV_i = \frac{TP_i}{TP_i + FP_i} \quad (13)
\]

\[
SE_i = \frac{TP_i}{TP_i + FN_i} \quad (14)
\]

\[
F1_i = \frac{2PPV_i\cdot SE_i}{PPV_i + SE_i} \quad (15)
\]

where \( TP_i \) and \( FN_i \), also known as true positive and false negative, respectively, refer to the number of the \( i \)-th class correctly predicted and the number of the \( i \)-th class classified into other classes, respectively. In addition, \( TN_i \) and \( FP_i \), also known as true negative and false
positive, respectively, are the number of other classes that are not classified as the $i$-th class and the number of other classes predicted as the $i$-th class, respectively.

3.3. Performance Analysis

To investigate the convergence performance of the proposed model, we depicted the accuracy value curves for different $lr$, setting the batch size as 512 and wavelet as gaus8 in Figure 8. It can be seen that $lr = 0.0001$ helps converge with higher accuracy compared with the other two $lr$ (e.g., 0.0005 and 0.00005) after 60 iterations. For the other two values of $lr$ (e.g., 0.0005 and 0.00005), multiple numbers of fluctuations are observed while $lr = 0.0001$ can attain a smaller number of fluctuations, indicating that the learning process for the proposed model is smooth and the convergence of the proposed model is better for the $lr = 0.0001$ value.

Figure 8. Accuracy value curves for different learning rates.

Figure 9 shows the confusion matrices for $lr$ values. In Table 4, we compute the performance of the proposed model numerically in terms of four evaluation metrics taking the values of $TP_i$, $FP_i$, $TN_i$, and $FN_i$ from the confusion matrices. The last four rows of the table compute the average value for all metrics. The average value of the evaluation metric reveals that the proposed model achieves the highest average value for PPV, SE, and F1 scores for batch size $= 512$ and $lr = 0.0001$. Table 5 shows 10-fold cross validation score for the proposed classifier. In this process, the whole data are split into ten number of folds and one fold is considered for testing and the rest will be for training and moving on with iterations in the entire process. In this study, 10-fold cross validation is applied to justify the proposed model generalization ability. From the table, it can be seen that the difference between the average value of the performance metrics (e.g., PPV, SE, F1) for all the four classes after applying 10-fold cross validation and the actual value of the performance metrics is very small.

Figure 9. Confusion matrices for (a) $lr = 0.0005$, (b) $lr = 0.0001$, and (c) $lr = 0.00005$. 

| Class | NB | SVEB | VEB | FB |
|-------|----|------|-----|----|
| Positive | 17955 | 20 | 6 | 19 |
| SVEB | 71 | 466 | 3 | 0 |
| VEB | 16 | 2 | 1369 | 13 |
| FB | 18 | 0 | 3 | 139 |

| Class | NB | SVEB | VEB | FB |
|-------|----|------|-----|----|
| Positive | 17963 | 33 | 2 | 2 |
| SVEB | 31 | 508 | 1 | 0 |
| VEB | 16 | 1 | 1776 | 7 |
| FB | 21 | 1 | 4 | 134 |

| Class | NB | SVEB | VEB | FB |
|-------|----|------|-----|----|
| Positive | 17990 | 7 | 2 | 1 |
| SVEB | 64 | 474 | 2 | 0 |
| VEB | 20 | 1 | 1368 | 11 |
| FB | 33 | 0 | 3 | 124 |
Table 4. Classification performance of the proposed model by varying learning rate ($lr$).

| Classes | Metrics | $lr = 0.0005$ | $lr = 0.0001$ | $lr = 0.00005$ |
|---------|---------|--------------|--------------|----------------|
| NB      | ACC     | 99.25        | 99.48        | 99.37          |
|         | PPV     | 99.41        | 99.62        | 99.35          |
|         | SE      | 99.75        | 99.79        | 99.94          |
|         | F1      | 99.58        | 99.71        | 99.64          |
| SVEB    | ACC     | 99.52        | 99.67        | 99.63          |
|         | PPV     | 95.49        | 93.56        | 98.34          |
|         | SE      | 86.29        | 94.07        | 87.77          |
|         | F1      | 90.67        | 93.81        | 92.75          |
| VEB     | ACC     | 99.72        | 99.94        | 99.80          |
|         | PPV     | 99.13        | 99.49        | 99.49          |
|         | SE      | 97.79        | 98.28        | 97.71          |
|         | F1      | 98.45        | 98.89        | 98.59          |
| FB      | ACC     | 99.74        | 99.83        | 99.76          |
|         | PPV     | 81.29        | 93.71        | 91.17          |
|         | SE      | 86.88        | 83.75        | 77.50          |
|         | F1      | 83.99        | 88.45        | 83.78          |
| Average | ACC     | 99.58        | 99.71        | 99.64          |
|         | PPV     | 93.83        | 96.60        | 97.09          |
|         | SE      | 92.68        | 93.97        | 90.73          |
|         | F1      | 93.17        | 95.22        | 93.69          |

Table 5. Ten-fold cross-validation score of our proposed model.

| K-Folds | NB | SVEB | VEB | FB |
|---------|----|------|-----|----|
|         | PPV | SE   | F1  | PPV | SE   | F1  | PPV | SE   | F1  | PPV | SE   | F1  |
| Fold = 1| 99.41 | 99.94 | 99.68 | 96.80 | 89.63 | 93.08 | 99.27 | 97.00 | 98.12 | 97.10 | 82.72 | 89.33 |
| Fold = 2| 99.51 | 99.80 | 99.66 | 97.54 | 88.15 | 92.61 | 98.14 | 98.00 | 98.07 | 85.19 | 86.25 | 85.71 |
| Fold = 3| 99.67 | 99.88 | 99.77 | 96.98 | 95.19 | 96.07 | 99.42 | 97.57 | 98.49 | 88.61 | 87.50 | 88.05 |
| Fold = 4| 99.33 | 99.92 | 99.62 | 98.73 | 86.67 | 92.31 | 98.83 | 96.43 | 97.61 | 78.95 | 75.00 | 76.92 |
| Fold = 5| 99.22 | 99.88 | 99.55 | 96.76 | 88.52 | 92.46 | 98.67 | 95.57 | 97.10 | 89.23 | 72.50 | 80.00 |
| Fold = 6| 99.40 | 99.81 | 99.61 | 99.40 | 99.81 | 99.61 | 98.70 | 97.57 | 98.13 | 88.46 | 85.19 | 86.79 |
| Fold = 7| 99.57 | 99.82 | 99.69 | 94.25 | 91.11 | 92.66 | 98.86 | 99.00 | 98.93 | 98.46 | 80.00 | 88.28 |
| Fold = 8| 99.60 | 99.78 | 99.69 | 94.74 | 93.68 | 94.21 | 97.58 | 98.00 | 97.79 | 95.31 | 76.25 | 84.72 |
| Fold = 9| 99.31 | 99.80 | 99.56 | 92.02 | 89.30 | 90.64 | 99.12 | 96.71 | 97.90 | 91.80 | 70.00 | 79.43 |
| Fold = 10| 99.57 | 99.78 | 99.67 | 93.04 | 94.07 | 93.55 | 98.56 | 98.21 | 93.65 | 73.75 | 82.52 | 82.52 |
| Average | 99.46 | 99.84 | 99.65 | 96.03 | 91.61 | 93.72 | 98.72 | 97.37 | 98.04 | 90.68 | 78.92 | 84.18 |

This outcome proves that the proposed classifier is generalizing and actively learning. A comparison of the performances with other existing works [11,12,14,26,28,29] has been provided in Table 6. Based on the average value, it can be stated that the proposed model achieves outstanding performance compared to those of studies mentioned in the table in terms of all evaluation metrics. In terms of performance metrics, the proposed model can attain an average improvement of 1.45–3.57%. However, it is noticeable that achieving higher
performance for the FB class is difficult. Since the FB class is composed of the fusion of ventricular and normal beat, which is very close to the normal heartbeat, existing studies predict a large number of ECG samples from the FB class as NB class. Therefore, the overall performance in terms of ACC, PPV, SE, and F1 scores for the FB class becomes deteriorated. In our study, the proposed model allows fused temporal characteristics and rich spatial features to be passed through the LSTM layers and convolutional blocks; hence, the model becomes more robust and distinguishable for the FB class. From the confusion matrix, it can be seen that a few ECG samples from the FB class are misclassified as NB class, resulting in higher ACC, PPV, and SE. Thus, our model shows more promising results for the FB class than that of other works, as shown in Table 6. For this class, our model has achieved a 5.46% improvement in the F1 score than [29].

Table 6. Comparison of the classification performance of existing studies and our studies \((lr = 0.0001)\) in terms of all classes.

| Classes | Metrics | Liu et al. [14] | Zhang et al. [11] | Ye et al. [12] | Wang et al. [26] | Jiang et al. [28] | Y. Zhang et al. [29] | Ours |
|---------|---------|-----------------|-------------------|----------------|------------------|-------------------|-----------------|------|
| NB      | PPV     | 96.66%          | 98.98%            | 97.55%         | 98.17%           | 98.39%           | 99.36%          | 99.62%         |
|         | SE      | 94.06%          | 88.94%            | 88.61%         | 99.42%           | 97.64%           | 99.62%          | 99.79%         |
|         | F1      | 95.34%          | 93.69%            | 92.87%         | 98.79%           | 98.01%           | 99.48%          | 99.71%         |
| SVEB    | PPV     | 39.87%          | 35.98%            | 52.34%         | 89.54%           | 63.70%           | 97.48%          | 93.56%         |
|         | SE      | 33.12%          | 79.06%            | 61.02%         | 74.56%           | 64.40%           | 98.44%          | 94.07%         |
|         | F1      | 36.18%          | 49.46%            | 56.34%         | 81.37%           | 64.04%           | 97.96%          | 93.81%         |
| VEB     | PPV     | 76.51%          | 92.75%            | 61.45%         | 93.25%           | 90.00%           | 92.66%          | 99.49%         |
|         | SE      | 90.20%          | 85.48%            | 81.82%         | 95.65%           | 91.00%           | 87.27%          | 98.28%         |
|         | F1      | 82.79%          | 88.96%            | 70.19%         | 94.43%           | 90.50%           | 89.88%          | 98.89%         |
| FB      | PPV     | 12.99%          | 13.73%            | 2.50%          | 2.04%            | 43.99%           | 91.04%          | 93.71%         |
|         | SE      | 40.72%          | 93.81%            | 19.69%         | 2.06%            | 76.70%           | 76.25%          | 83.75%         |
|         | F1      | 19.70%          | 23.96%            | 4.43%          | 0.46%            | 55.91%           | 82.99%          | 88.45%         |
| Average | PPV     | 56.51%          | 60.36%            | 53.46%         | 70.75%           | 74.02%           | 95.14%          | 96.60%         |
|         | SE      | 64.53%          | 86.82%            | 62.79%         | 67.47%           | 82.44%           | 90.40%          | 93.97%         |
|         | F1      | 58.50%          | 64.02%            | 55.96%         | 68.76%           | 77.12%           | 92.58%          | 95.22%         |

In this paper, we have also observed and analyzed three other wavelet functions' performances in the ECG classification. Taking the values of \(TP_i\), \(FP_i\), \(TN_i\), and \(FN_i\) from the confusion matrices depicted in Figure 10, a comparison of four wavelet functions' performances is listed in Table 7. Significant variation among the performance of the four types of wavelet function for all metrics except that of ACC appeared. Specifically, \(gaus8\) performs better than other wavelets for the VEB and FB classes, while \(mexh\) performs better for the SVEB class. Since the number of samples is large in the NB class, the performance is almost the same for all the wavelets. Because of the higher similarity between the waveform of the ECG signal and the \(gaus8\) wavelet function, the \(gaus8\) based scalogram is the most suitable for arrhythmia classification.
Since the number of samples is large in the NB class, the performance is almost the same for all the wavelets. Because of the higher similarity between the waveform of the ECG signal and the `gaus8` wavelet function, the `gaus8` based scalogram is the most suitable for arrhythmia classification.

**Figure 10.** Confusion matrices for four wavelet functions (a) `gaus8`, (b) `gaus4`, (c) `mexh`, and (d) `morl` keeping $lr = 0.0001$.

**Table 7.** Comparison of the classification performance of the four wavelet functions in terms of all classes (a) NB class, (b) SVEB class, (c) VEB class, and (d) FB class.

| Classes | Wavelet Functions | ACC | PPV | SE | F1  |
|---------|-------------------|-----|-----|----|-----|
| NB      | `gaus8`           | 99.48 | 99.62 | 99.79 | 99.70 |
|         | `gaus4`           | 99.38 | 99.38 | 99.93 | 99.66 |
|         | `mexh`            | 99.50 | 99.53 | 99.91 | 99.72 |
|         | `morl`            | 99.52 | 99.63 | 99.84 | 99.73 |
| SVEB    | `gaus8`           | 99.67 | 93.55 | 94.07 | 93.81 |
|         | `gaus4`           | 99.63 | 98.74 | 87.40 | 92.73 |
|         | `mexh`            | 99.72 | 98.01 | 91.67 | 94.74 |
|         | `morl`            | 99.71 | 96.35 | 92.96 | 94.62 |
| VEB     | `gaus8`           | 99.84 | 99.49 | 98.28 | 98.88 |
|         | `gaus4`           | 99.82 | 98.92 | 98.50 | 98.71 |
|         | `mexh`            | 99.81 | 99.78 | 98.50 | 98.64 |
|         | `morl`            | 99.82 | 98.50 | 98.92 | 98.71 |
| FB      | `gaus8`           | 99.82 | 93.70 | 83.75 | 88.44 |
|         | `gaus4`           | 99.79 | 95.41 | 78.13 | 85.91 |
|         | `mexh`            | 99.81 | 95.55 | 80.62 | 87.48 |
|         | `morl`            | 99.77 | 94.53 | 75.62 | 84.02 |

Therefore, this investigation recommends that using a wavelet function close to the ECG signal (e.g., `gaus8`) will lead to the best ECG arrhythmia detection performance.
4. Conclusions

In this article, we proposed a novel parallel cross convolutional RNN to improve the arrhythmia detection performance by modeling ECG signal classification as an imbalanced problem. The proposed model incorporates extracted temporal characteristics and spatial features, utilizes several LSTM layers in the top branch and several convolutional blocks in the bottom branch, and allows for flowing throughout the whole model, making the model’s learning more distinguishable and robust. We have evaluated the detailed performance of the proposed model in terms of four evaluation metrics on the well-known MIT-BIH dataset. Additionally, a comparison of the performances of other existing works with the optimum performance of the proposed model has been presented. The proposed model achieves higher performances compared with others, especially since the classifier shows an outstanding performance for the FB class. Finally, the model’s performance depending on the four distinct wavelet functions has been elaborately described. Our proposed classifier can distinguish NB, SVEB, VEB, and FB classes with an average ACC, average PPV, average SE, and average F1 of 99.58%, 96.60%, 93.97%, and 95.22%. For the most skewed class, i.e., FB class, we have obtained promising results (PPV = 93.71%, SE = 83.75%, F1 = 88.45%) as compared to other state-of-the-art approaches. This experimental result validates the effectiveness of the proposed model. Moreover, we observed that gaus8 leads to higher classification performance in this study compared to other wavelet functions. However, the model is not lightweight because the training parameter is as high as about 34 million, which leads to high training time. In the future, we will research reducing the training parameter to make the proposed parallel cross convolutional RNN classifier lightweight and will investigate its performance on other physiological signal analysis.

Author Contributions: T.I.T.: conceptualization, methodology, software, writing—original draft. S.C.: validation, supervision, project administration, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Research Foundation of Korea (NRF) Grant, funded by the Korean Government (MSIT) (Nos. 2016R1A5A1012966 and 2021R1F1A1062285).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used in this study are obtained from public datasets. Readers should be able to obtain these data by requesting the dataset sources described in this study.

Conflicts of Interest: The authors hereby declare that they have no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the publication of the results.

References

1. Cardiovascular Diseases (CVDs). Available online: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed on 17 May 2017).
2. Ahmad, Z.; Tabassum, A.; Guan, L.; Khan, N.M. ECG Heartbeat Classification Using Multimodal Fusion. IEEE Access 2021, 9, 100616–100626. [CrossRef]
3. Zhang, J.; Liang, D.; Liu, A.; Gao, M.; Chen, X.; Zhang, X.; Chen, X. MLBF-Net: A Multi-Lead-Branch Fusion Network for Multi-Class Arrhythmia Classification Using 12-Lead ECG. IEEE J. Transl. Eng. Health Med. 2021, 9, 1–11. [CrossRef] [PubMed]
4. Zhu, W.; Chen, X.; Wang, Y.; Wang, L. Arrhythmia Recognition and Classification Using ECG Morphology and Segment Feature Analysis. IEEE/ACM Trans. Comput. Biol. Bioinform. 2019, 16, 131–138. [CrossRef] [PubMed]
5. Osowski, S.; Hoa, L.T.; Markiewic, T. Support vector machine based expert system for reliable heartbeat recognition. IEEE Trans. Biomed. Eng. 2004, 51, 582–589. [CrossRef]
6. Chazal, P.D.; O’Dwyer, M.; Reilly, R.B. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng. 2004, 51, 1196–1206.
7. Jun, T.J.; Park, H.J.; Kim, Y.H. Premature ventricular contraction beat detection with deep neural networks. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 859–864.
8. Monasterio, V.; Laguna, P.; Martinez, J.P. Multilead analysis of t-wave alternans in the ecg using principal component analysis. IEEE Trans. Biomed. Eng. 2009, 56, 1880–1890. [CrossRef]

9. Özbay, Y.; Tezel, G. A new method for classification of ECG arrhythmias using neural network with adaptive activation function. Dig. Sig. Process. 2010, 20, 1040–1049. [CrossRef]

10. Tuncer, T.; Doğan, S.; Plawiak, P.; Acharya, U.R. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. Knowl. Based Syst. 2019, 186, 104923. [CrossRef]

11. Zhang, Z.; Dong, J.; Luo, X.; Choi, K.S.; Wu, X. Heartbeat classification using disease-specific feature selection. Comput. Biol. Med. 2014, 46, 79–89. [CrossRef]

12. Ye, C.; Kumar, B.V.; Coimbra, M.T. Heartbeat classification using morphological and dynamic features of ECG signals. IEEE Trans. Biomed. Eng. 2012, 59, 2930–2941.

13. Chen, S.; Hua, W.; Li, Z.; Li, J.; Gao, X. Heartbeat classification using projected and dynamic features of ECG signal. Biomed. Signal Process. Control 2017, 31, 165–173. [CrossRef]

14. Jian, L.; Shuang, S.; Guozhong, S.; Yu, F. Classification of ECG Arrhythmia Using CNN, SVM and LDA. In International Conference on Artificial Intelligence and Security; Lecture Notes in Computer Science (LNCS 11633); Springer: Berlin/Heidelberg, Germany, 2019; pp. 191–201.

15. Nurmaini, S.; Tondas, A.E.; Darmawahyuni, A.; Rachmatullah, M.N.; Khoirani, R. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. Future Gener. Comput. Syst. 2020, 113, 304–317. [CrossRef]

16. Acharya, U.R.; Oh, S.L.; Hagiwara, Y.; Tan, J.H.; Tan, R.S. A deep convolutional neural network model to classify heartbeats. Comput. Biol. Med. 2017, 89, 389–396. [CrossRef] [PubMed]

17. Acharya, U.R.; Fujita, H.; Lih, O.S.; Hagiwara, Y.; Tan, J.H.; Adam, M. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Inf. Sci. 2017, 405, 81–90. [CrossRef]

18. Wang, H.; Shi, H.; Lin, K.; Qin, C.; Zhao, L.; Huang, Y.; Liu, C. A high precision arrhythmia classification method based on dual fully connected neural network. Biomed. Signal Process. Control 2020, 58, 101874. [CrossRef]

19. Chen, A.; Wang, F.; Liu, W.; Chang, S.; Wang, H.; He, J.; Huang, Q. Multi-information fusion neural networks for arrhythmia automatic detection. Comput. Meth. Programs Biomed. 2020, 193, 105479. [CrossRef]

20. Chen, C.; Hua, Z.; Zhang, R.; Liu, G.; Wen, W. Automated arrhythmia classification based on a combination network of CNN and LSTM. Biomed. Signal Process. Control 2019, 57, 101819. [CrossRef]

21. Huang, J.; Chen, B.; Yao, B.; He, W. ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network. IEEE Access 2019, 7, 92871–92880. [CrossRef]

22. Hao, C.; Wibowo, S.; Majmudar, M.; Rajput, K.S. Spectro-temporal feature based multi-channel convolutional neural network for ECG beat classification. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; pp. 5642–5645.

23. Oliveira, A.T.; Nobrega, E.G.O. A novel arrhythmia classification method based on convolutional neural networks interpretation of electrocardiogram images. In Proceedings of the 2019 IEEE International Conference on Industrial Technology (ICIT), Melbourne, VIC, Australia, 13–15 February 2019; pp. 841–846.

24. Al Rahhal, M.M.; Bazi, Y.; Almubarak, H.; Alajlan, N.; Al Zuair, M. Dense convolutional networks with focal loss and image generation for electrocardiogram classification. IEEE Access 2019, 7, 182225–182237. [CrossRef]

25. Liu, W.; Zhang, M.; Zhang, Y.; Liao, Y.; Huang, Q.; Chang, S.; Wang, H.; He, J. Real-time multilead convolutional neural network for myocardial infarction detection. IEEE J. Biomed. Health Inform. 2018, 22, 1434–1444. [CrossRef]

26. Wang, T.; Lu, C.; Sun, Y.; Yang, M.; Liu, C.; Ou, C. Automatic ECG Classification Using Continuous Wavelet Transform and Convolutional Neural Network. Entropy 2021, 23, 119. [CrossRef] [PubMed]

27. He, H.; Garcia, E.A. Learning from imbalanced data. IEEE Trans. Knowl. Data Eng. 2009, 21, 1263–1284.

28. Jiang, J.; Zhang, H.; Pi, D.; Dai, C. A novel multi-module neural network system for imbalanced heartbeats classification. Expert Syst. Appl. X 2019, 1, 1–15. [CrossRef]

29. Zhang, Y.; Li, J.; Wei, S.; Zhou, F.; Li, D. Heartbeats Classification Using Hybrid Time-Frequency Analysis and Transfer Learning Based on ResNet. IEEE J. Biomed. Health Inform. 2021, 25, 4175–4184. [CrossRef]

30. Zhang, Y.; Chen, X.; Guo, D.; Song, M.; Teng, Y.; Wang, X. PCCN: Parallel Cross Convolutional Neural Network for Abnormal Network Traffic Flows Detection in Multi-Class Imbalanced Network Traffic Flows. IEEE Access 2019, 7, 119904–119916. [CrossRef]

31. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng. 1985, 32, 230–236. [CrossRef]

32. Lee, H.K.; Choi, Y-S. Application of Continuous Wavelet Transform and Convolutional Neural Network in Detecting Motor Imagery Brain-Computer Interface. Entropy 2019, 21, 1199. [CrossRef]

33. Guo, M.F.; Zeng, X.D.; Chen, D.Y.; Yang, N.C. Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems. IEEE Sens. J. 2017, 18, 1291–1300. [CrossRef]

34. Wang, L.; Wang, C.; Fu, F.; Yu, X.; Guo, H.; Xu, C.; Jing, X.; Zhang, H.; Dong, X. Temporal Lobe seizure prediction based on a complex Gaussian wavelet. Clin. Neurophysiol. 2011, 4, 656–663. [CrossRef]

35. Auger, F.; Patrick, F.; Paulo, G.; Olivier, L. Time-Frequency Toolbox; CNRS France-Rice University: Paris, France, 1996.

36. Meignen, S.; Oberlin, T.; McLaughlin, S. A New Algorithm for Multicomponent Signals Analysis Based on Synchro Squeezing: With an Application to Signal Sampling and Denoising. IEEE Trans. Signal Process. 2012, 60, 5787–5798. [CrossRef]
37. Kong, W.; Dong, Z.Y.; Hill, D.J.; Luo, F.; Xu, Y. Short-term residential load forecasting based on resident behavior learning. *IEEE Trans. Power Syst.* 2018, 33, 1087–1088. [CrossRef]

38. Moody, G.B.; Mark, R.G. The impact of the MIT-BIH arrhythmia database. *IEEE Eng. Med. Biol. Mag.* 2001, 20, 45–50. [CrossRef] [PubMed]

39. Mar, T.; Zaunseder, S.; Martínez, J.P.; Llamedo, M.; Poll, R. Optimization of ECG classification by means of feature selection. *IEEE Trans. Biomed. Eng.* 2011, 58, 2168–2177. [CrossRef] [PubMed]

40. Mondéjar-Guerra, V.; Novo, J.; Rouco, J.; Penedo, M.G.; Ortega, M. Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. *Biomed. Signal Process. Control* 2019, 47, 41–48. [CrossRef]