Sensors and Materials, Vol. 36, No. 11 (2024) 4741–4755
MYU Tokyo

S & M 3828

Combining Endpoint Detection with a Convolutional Neural Network Classifier for the Automatic Recognition of Cardiac Arrhythmias in Electrocardiogram Signals

Yu-En Cheng,[1] Chih-Te Tsai,[1] Chia-Hung Lin,[1*] Ching Chou Pai,[2]
Pi-Yun Chen,[1] Chien-Ming Li,[3] and Neng-Sheng Pai[1**]

[1]Department of Electrical Engineering, National Chin-Yi University of Technology, Taichung City 41170, Taiwan
[2]Show-Chwan Memorial Hospital, Division of Cardiovascular Surgery, Changhua 500, Taiwan
[3]Division of Infectious Diseases, Department of Medicine of Chi Mei Medical Center, Tainan City 710, Taiwan

(Received May 8, 2024; accepted September 12, 2024)

Keywords: cardiac arrhythmias, endpoint detection (EPD), convolutional neural network (CNN), electrocardiogram (ECG), and visualization color pattern

A cardiac arrhythmia is an abnormal heart rhythm caused by irregular heartbeats. Cardiac arrhythmias include atrial or ventricular fibrillation, right or left bundle branch block beats, and premature atrial or ventricular contractions. Different cardiac arrhythmias have distinct causes and clinical presentations. The type of cardiac arrhythmia must be identified to enable further intervention and treatment for addressing its underlying causes. In this study, we developed a convolutional neural network (CNN) model that extracts and classifies time-domain features to detect cardiac arrhythmias automatically in electrocardiogram (ECG) signals. This model employs endpoint detection to detect the activity of time-domain signals in accordance with a threshold for identifying the peak wave in ECG signals. These features are then transferred to two-dimensional (2D) color patterns that indicate abnormal heartbeats. Subsequently, a one-dimensional (1D) or 2D CNN classifier is employed to distinguish normal heartbeats from cardiac arrhythmias in raw ECG data. The proposed model was trained, tested, and validated on the Massachusetts Institute of Technology–Beth Israel Deaconess Medical Center Arrhythmia Database (commonly known as the MIT-BIH Arrhythmia Database), and it exhibited promising performance in cardiac arrhythmia recognition, as indicated by its precision, recall, *F1* score, and accuracy.

1. Introduction

Electrocardiogram (ECG) signals are recordings of the electrical activity of the heart from the atria to the right and left ventricles. These signals contain information regarding numerous time- and frequency-domain physiological parameters [e.g., heart rate (HR), HR variability, and respiration rate[1–5]] that can be used for clinical monitoring, the diagnosis of cardiac diseases, the evaluation of cardiac risk, and the quantification of the efficacy of cardiac drug therapies.

*Corresponding author: e-mail: eechl53@gmail.com
**Corresponding author: e-mail: pai@ncut.edu.tw
https://doi.org/10.18494/SAM5144

ISSN 0914-4935
https://myukk.org/

ECG signals can be measured using noninvasive limb leads, such as limb leads I, II, and III, and chest leads.[6–10] An ECG waveform comprises three components: P wave, QRS complex wave, and T wave. ECG signals can indicate an individual's emotional state, cardiovascular disorders, and sleep-related conditions (e.g., sleep apnea). Moreover, they can be used to diagnose cardiac arrhythmias, premature atrial contractions (PACs), premature ventricular contractions (PVCs), atrial fibrillation (AF), and ventricular fibrillation (VF).[6–10]

ECG measurements can be acquired using contact sensing methods, which involve placing sensors or electrodes on multiple parts of the human body, such as the upper and lower limbs and chest. Limb electrodes or adhesive patches are connected to form limb or chest leads, which typically comprise up to 12 electrodes (ECG leads). The configuration of these leads enables the continuous measurement of the electrical activity of the heart. Commercial microchip firmware (with low power and voltage) facilitates the noninvasive monitoring of the electrical potentials passing through the heart. However, contact noise (skin–contact impedance), electromagnetic interference, and interference from power source harmonics[11–13] can result in errors in potential measurement. Moreover, long-duration ECG monitoring results must be interpreted by experienced medical staff with relevant expertise. Accordingly, in this study, we developed a model that can automatically extract and classify ECG signal features and filter out noise to recognize cardiac arrhythmias with high levels of accuracy.

In signal preprocessing, baseline-drift, power line interference, and high-frequency noise filters are used to eliminate low- and high-frequency noises, and noise resulting from power source interference, respectively. Baseline drift is caused by low-frequency signal variations that occur in the baseline signal because of breathing (<1.0 Hz) and muscle movements (6.0–100.0 Hz). Power source interference is caused by the power supply frequency (50 or 60 Hz) of an electricity grid. High-frequency noise is caused by electromagnetic interference from the environment. A previous study[14] proposed the use of a band-pass filter with a cutoff frequency of 5–12 Hz to address baseline drift (within the frequency range of 0.0–0.8 Hz). A Butterworth low-pass filter with a cutoff frequency of 30 Hz can reduce or eliminate the high-frequency components of a signal, and a band-reject filter can be employed to eliminate 50 or 60 Hz harmonic components caused by the power source.[11,14–17] A filter's cutoff frequency and order (up to three orders) can be varied to improve its performance and minimize noise and distortion in time-domain signals. However, filters with analog electronic circuits (including operational amplifiers, resistors, and capacitors)[18–20] are susceptible to nonlinear properties, temperature variations, and environmental factors, all of which can lead to instability in filter performance. The filtering scheme of these filters requires calibration and compensation for practical applications. In addition, increasing the filter orders enhances the complexity of designing and implementing hardware circuits, which restricts the application of the aforementioned filters in signal preprocessing. Accordingly, in this study, we adopted digital filters incorporating mathematical algorithms to preprocess raw ECG signals to offer an altering model for baseline-drift filter and Butterworth filter designs, as adjusting parameters in amplification gains, cutoff frequencies, and filter orders. Compared with analog filters, digital filters have greater stability, flexibility, and applicability. As shown in Fig. 1(a), raw ECG signals may include high-frequency and baseline-drift components. Baseline drift and noise can be identified after detrend

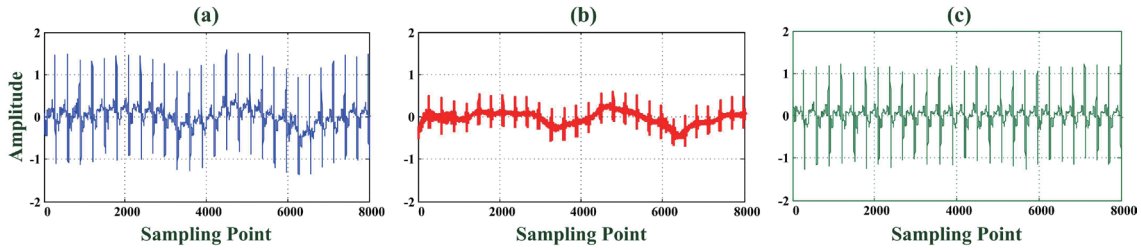

Fig. 1. (Color online) Preprocessing of raw ECG data. (a) Raw ECG data with high-frequency and baseline-drift components, (b) baseline-drift components obtained after detrend filtering, and (c) ECG data obtained after digital signal processing.

processing and the application of digital Butterworth filters [Fig. 1(b)]. Subsequently, low- and high-frequency noises can be reduced or eliminated, as depicted in Fig. 1(c). Therefore, ECG signals can be interpreted automatically and accurately by using appropriate methods.

Frequency-domain methods such as the fast Fourier transform (FFT) and wavelet transform (WT)[17,19,21,22] can be used to extract time- and frequency-domain features from signals. The FFT is used to capture global frequency information, and the WT provides a multiscale method for decomposing time-domain signals into frequency ranges, which can then be visualized as frequency feature patterns for the detection of PACs, PVCs, AF, and VF. In practical applications, appropriate wavelet basis functions, wavelet parameters (dilation and location parameters), and decomposition levels must be selected because these factors affect the effectiveness of feature extraction in different frequency ranges. The WT is sensitive to noise within incoming signals; thus, signal preprocessing must be performed to eliminate noise when the WT is employed. In addition, the WT incorporates complex computation processes; therefore, it requires considerable time to process long-term data, and thus the scope of its practical applications is limited. Accordingly, in this paper, we propose an endpoint detection (EPD) method that involves using a specific threshold[23–25] first to identify the beginning and ending points of ECG activities in amplitude variations, and then to fragment each ECG segment with these substantial activities. Subsequently, the time-domain or frequency-domain feature extraction methods, such as the short-time energy (STE), short-time zero-crossing rate (STZCR), and spectral characteristics, can be extracted from these signal segments by using energy-based methods, zero-crossing-rate analysis, the WT or FFT. Furthermore, machine learning or deep learning models—such as support vector machine, *k*th nearest neighbor, and probabilistic neural network models[21,22,26,27]—can be employed to recognize the type of cardiac arrhythmia automatically.

In this study, we combined the EPD method with 1D and 2D convolutional neural network (CNN) classifiers for the automatic recognition of cardiac arrhythmias from ECG signals (Fig. 2). In feature extraction, the EPD method[23–25] is used to detect the activities of a time-domain signal with a specific threshold to extract the QRS-complex waves from an ECG data stream. These QRS-complex waves are then transferred to colored 2D visualization feature patterns to indicate normal or abnormal heartbeats [seen as normal (•) and PVC (V) in Fig. 2]. The model then uses a 1D- or 2D-CNN-based classifier[28–30] to separate the normal condition from cardiac arrhythmias in the ECG data stream. The developed model was trained, tested, and validated on

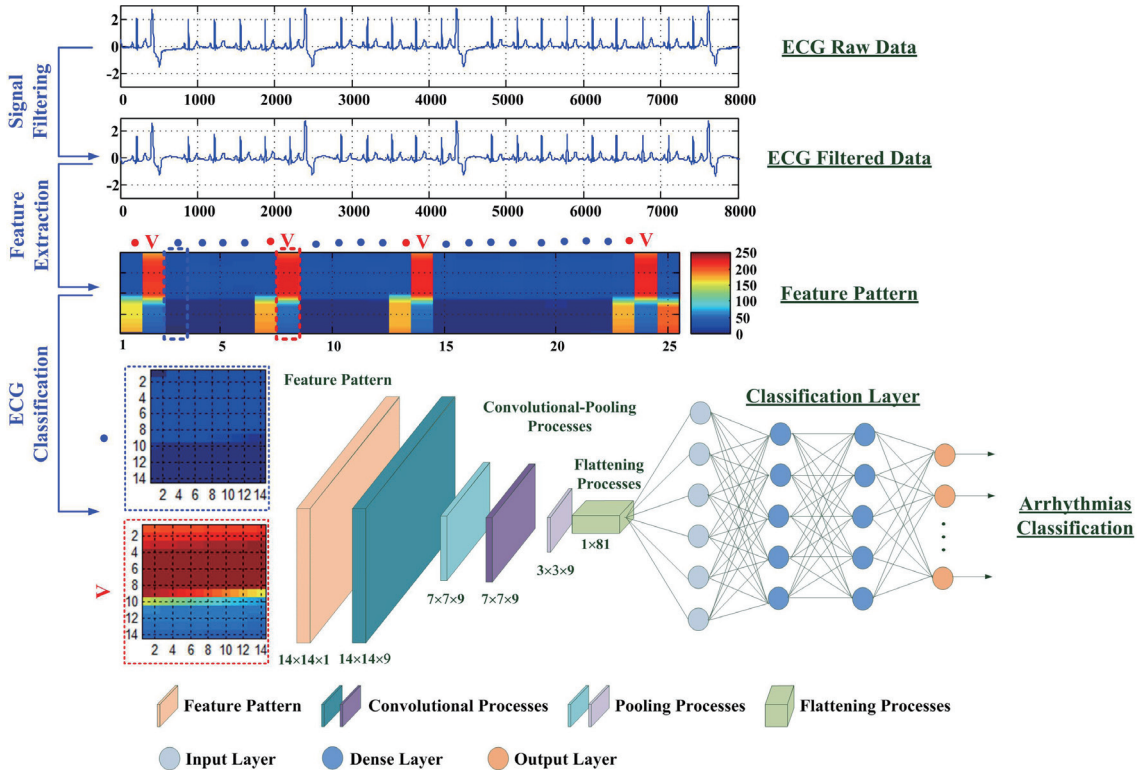

Fig. 2. (Color online) Structure of the proposed CNN model for the automatic recognition of cardiac arrhythmias in ECG signals.

the Massachusetts Institute of Technology (MIT)–Beth Israel Deaconess Medical Center (MIT-BIH) Arrhythmia Database,[26,31] and the precision (%), recall (%), *F1* score, and accuracy (%) of this model indicated that it was able to accurately recognize cardiac arrhythmias.

2. Materials and Methods

2.1 ECG raw data collection

In this study, raw ECG signals were measured from limb lead II (a bipolar lead), as displayed in Fig. 3. This lead measured the voltage difference between left limb (LL) and right arm (RA) electrodes (the LL and RA were considered the positive and negative ends of the LL–RA axis, respectively, and this axis was inclined 60° relative to the heart). However, the aforementioned signals were susceptible to various types of interference, including baseline drift, muscle artifacts, power line interference, high-frequency electromagnetic interference from the environment, and unknown white Gaussian noise[11,14–17,26] [Fig. 3(a)]. Therefore, appropriate filters had to be designed to eliminate these different types of interference, which have different frequency ranges [Fig. 3(a)]. In this study, digital Butterworth filters were employed to preprocess raw ECG signals with the syntax butter (•) of MATLAB (MathWorks, Natick, MA, USA).[32] In addition, we designed a fifth-order band-pass filter with lower and upper cutoff

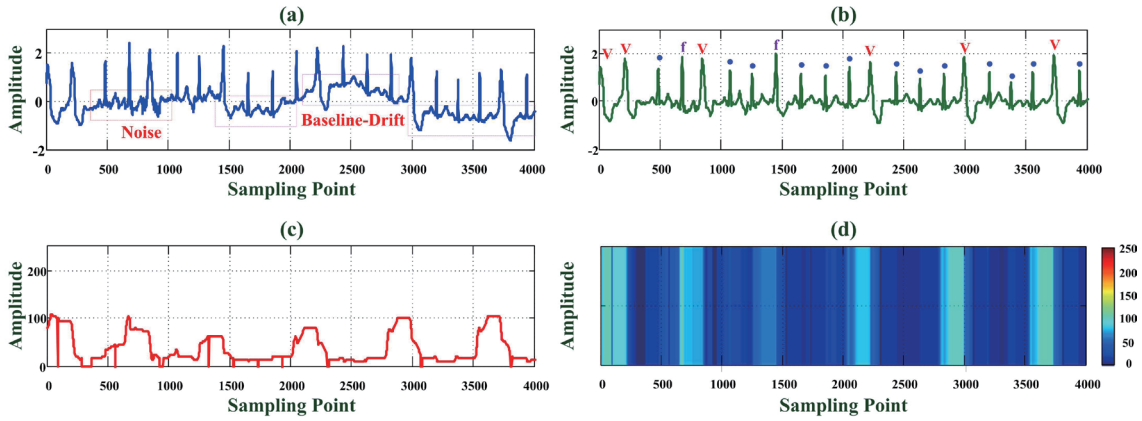

Fig. 3. (Color online) ECG signal processing and feature extraction. (a) Raw ECG signals with artifacts (baseline drift and muscle artifact) and noise, (b) preprocessed ECG signal obtained after removal of baseline drift and noises, (c) feature extraction through EPD, and (d) color patterns for normal and abnormal heartbeats.

frequencies of 1.0 and 30.0 Hz, respectively, and a sampling rate of 500.0 Hz. After signal filtering, the voltage amplification gain was set to 60 dB to obtain stable voltage gain characteristics. Moreover, the syntax detrend (•) of MATLAB was used to eliminate the effect of baseline drift caused by respiration and muscle artifacts. As shown in Fig. 3(b), after noise removal, each ECG data stream exhibited improved stability while retaining crucial features for diagnosis applications.

We conducted experimental tests by using raw ECG signals from the MIT-BIH Arrhythmia Database,[26,31] which consists of 48 ECG recordings from nine individuals, with each recording obtained from two leads, namely, limb lead II and chest lead V1, V2, V4, or V5. The duration of each recording was approximately 30 min, and each recording was digitized with a resolution of 11 bits over a 10 mV range at a sampling rate of 360 Hz per channel. Each data entry was annotated with a biomarker, which indicated whether a heartbeat's R peak belonged to a normal rhythm (normal beat) or an arrhythmia. In the annotations, •, V, A, L, R, P, F, and f represent a normal beat, a PVC, an atrial premature beat, a left bundle branch block beat, a right bundle branch block beat, a paced beat, the fusion of ventricular and normal beats, and the fusion of paced and normal beats, respectively.[31] QRS segmentations were cut using a 196-sampling-point window centered on the R peak. The EPD method was adopted to detect signal activities, after which the start and end points of signals were identified for QRS segmentation [Fig. 3(c)]. The QRS segments were then transferred to colored 2D feature patterns [Fig. 3(d)]. These feature data sets were divided into training and testing data sets. The training data sets were used to train the developed model, and the testing data sets were used to test and verify the feasibility of this model.

2.2 EPD method

The EPD method was used to extract feature patterns from time-domain signals with a specific threshold. Subsequently, the start and end points of these signals were determined to

obtain QRS segments that were then used to extract time- and frequency-domain features from the collected ECG data. The extracted features included STE, STZCR, short-time average amplitude, and cumulative spectral difference. These features can be applied in signal recognition, echo cancellation, and audio coding applications. In dynamic time-domain signals, the energy of the QRS complex wave is usually higher than that of noise. STE can be directly obtained from a time-domain ECG signal by focusing on each time point, expressed as

$$E_n = \sum_{k=1}^{N_k} x_n^2(i) \geq \theta,\ n = 1, 2, 3, \ldots, N - N_k,\ k = 1, 2, 3, \ldots, N_k, \tag{1}$$

$$\theta = \max(E_n) \times 0.1. \tag{2}$$

For a data stream, amplitude variations are more apparent during intense activity than during relatively static periods. Therefore, short-term amplitude changes over a finite period indicate energy variations for intense activities. For a short-term signal segment of length $N - N_k$, E_n represents the STE at the nth sampling point, $x(i)$ is the time-domain signal, and N_k is the length of the short-time frame processing. Equation (1) with a specific threshold θ [as Eq. (2)] is used to distinguish segments with heartbeat signals from those without such signals. Thus, different arrhythmia classes might exhibit distinct STE variations, as shown in Fig. 4. Moreover, a linear transformation is applied using the maximum energy [$E_{max} = \max(E_n)$] and minimum energy [$E_{min} = \min(E_n)$] to map energy variations to the range of [0, 1]. This transformation is expressed as

$$C_n = \frac{E_n - E_{min}}{E_{max} - E_{min}},\ C_n \in [0,\ 1],\ n = 1, 2, 3, \ldots, N - N_k, \tag{3}$$

$$\Phi_n = colormap(E_n),\ \Phi_n \in [0, 255]. \tag{4}$$

The normalized C_n value can be used alongside the colormap(•) operator of MATLAB to map QRS segments to various color feature patterns [Figs. 3(c) and 3(d)]. These patterns depict energy variations and feature patterns in color to provide clear representations of normal heartbeats and cardiac arrhythmias, respectively (Fig. 4).

2.3 CNN-based classifier

As shown in Fig. 2, we conducted digital signal processing and EPD to train a CNN model to automatically perform feature extraction, feature enhancement, and signal classification tasks. In this model, an EPD-based extractor extracts crucial QRS segments; subsequently, multiple layers of convolutional–pooling operations (CPOs) are conducted with kernel windows with different weights (convolutional operation) to perform feature extraction and enhancement. The use of multiple convolutional layers and multiple kernel windows in convolution operations can

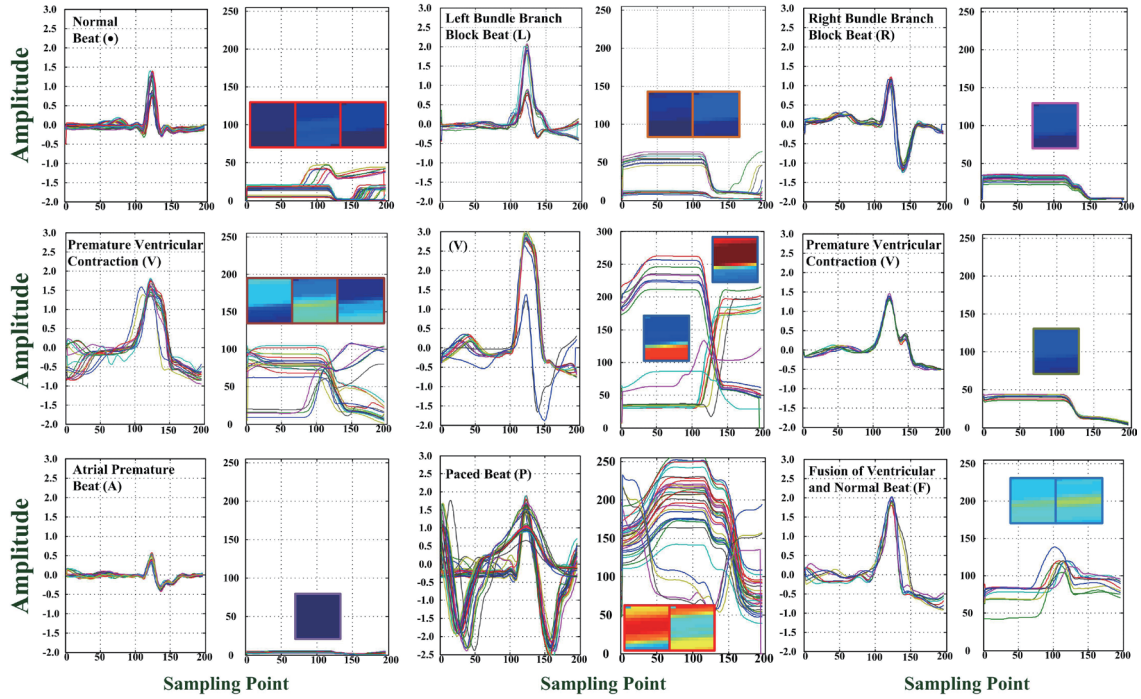

Fig. 4. (Color online) Different QRS segments and feature patterns for normal heartbeats and various types of cardiac arrhythmia [• (normal beat), V, A, L, R, P, and F].

increase the depth and breadth of features, leading to improved feature pattern dimensions, nonlinearity, and increased feature recognition accuracy. Subsequently, key feature parameters are selected using maximum-pooling (max-pooling) processes,[28–30] in which maximum values are extracted using a 2 × 2 max-pooling window to reduce the size of feature patterns (Fig. 5). After max-pooling processes, the number of features is reduced to one-fourth of the original number, with key features being retained. Thus, the developed CNN model can automatically extract crucial feature patterns. In the classification layer of this model, the different arrhythmia classes of feature patterns are identified using a backpropagation algorithm (BPA) with gradient descent optimizator (GDO) or adaptive moment estimation (ADAM) optimizator[33,34] to train, test, and validate the classifier model, namely, a 1D or 2D CNN classifier.

In this study, we designed a multilayer CNN model (Fig. 2) by using MATLAB. This model contains two CPO layers. In the first CPO layer, convolutional operations are performed with 3 × 3 × 9 kernel and 2 × 2 × 9 max-pooling windows to reduce the size of input feature patterns from 14 × 14 × 9 to 7 × 7 × 9. In the second CPO layer, feature patterns with a size of 3 × 3 × 9 are extracted. Following a flattening process, feature patterns are transformed from a matrix form (3 × 3 × 9) to a 1D pattern (vector form, 1 × 81). The classification layer of the developed CNN model consists of an input layer, multiple hidden layers, and an output layer. A Gaussian error linear unit was the activation function selected for the hidden layers, and softmax was the activation function selected for the output layer. The proposed model can classify normal heartbeats and multiple classes of cardiac arrhythmias.[26,31] The output target vector of this model is encoded using binary values; for example, the binary coding [1, 0, 0, 0, 0, 0, 0, 0] is

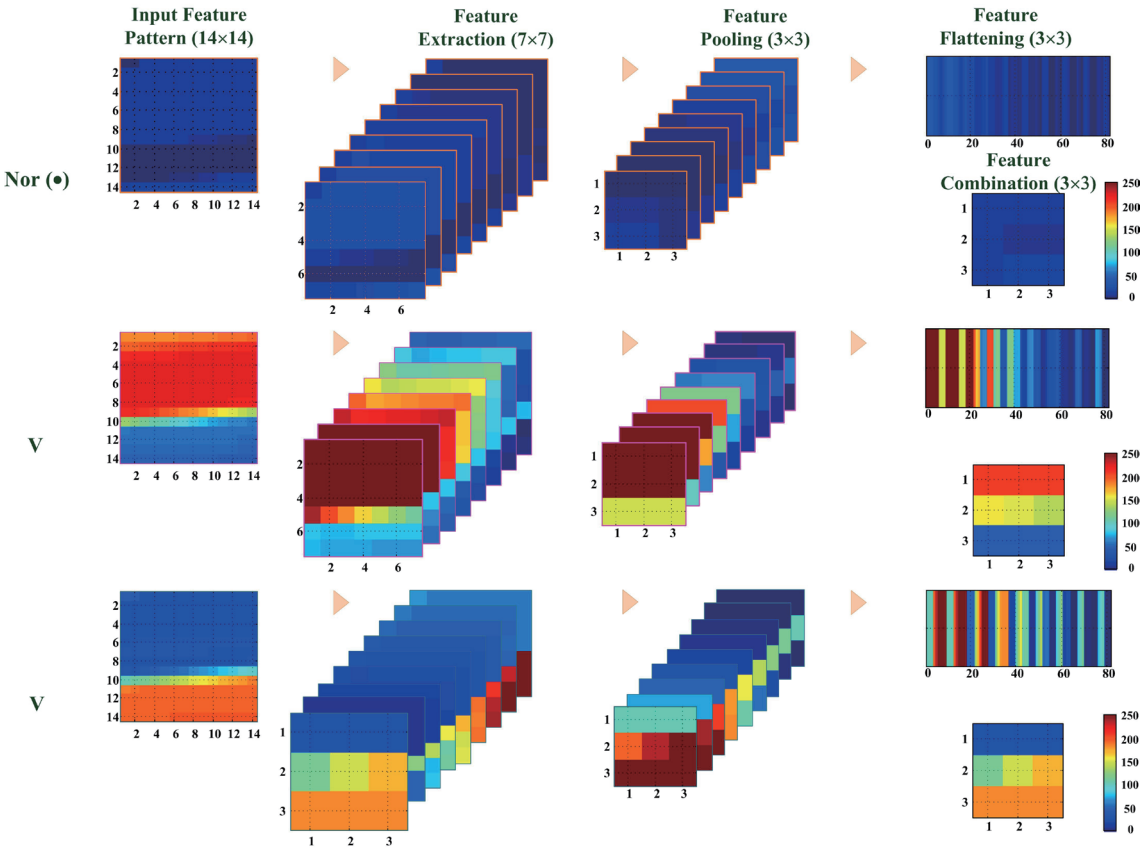

Fig. 5. (Color online) Convolutional–pooling operations for feature patterns of • (normal beat) and V classes.

used for normal beats (•). The training performance of the proposed model was evaluated in this study by using the categorical cross-entropy (CCE) loss function (LF).

The BPA or a swarm optimization algorithm (e.g., the particle or egret swarm optimization algorithm)[5,35,36] is used to update the weighted and bias parameters in the fully connected layer of the proposed model. The training process is terminated when a predetermined convergence condition is reached or when the maximum number of iterations is completed. Subsequently, the CCE LF is employed to evaluate the model's training performance. In the testing stage, the model outputs a confusion matrix, which specifies the numbers of true positives (TPs), true negatives (TNs), false positives (FPs), and false negatives (FNs). On the basis of these parameters, the precision (%), recall (%), $F1$ score, and accuracy (%) of the model can be obtained to evaluate the model's feasibility for cardiac arrhythmia recognition.

3. Experimental Results

In this study, we used ECG data sets from the MIT-BIH Arrhythmia Database to train, test, and validate the proposed CNN model. The adopted data sets comprised data regarding eight classes: normal beats (•), PVCs (V), atrial premature beats (A), left bundle branch block beats

(L), right bundle branch block beats (R), paced beats (P), the fusion of ventricular and normal beats (F), and the fusion of paced and normal beats (f). On the basis of the Advancement of Medical Instrumentation (AAMI) standard, these heartbeats were grouped into normal (N: •, L, and R), ventricular ectopic (V: V), supraventricular ectopic (S: A), and fusion/unknown (F/Q: P, F, and f) heartbeats.[37] We obtained the 1D and 2D feature patterns shown in Fig. 5 for the • and V classes; these patterns exhibited distinct color distributions for different classes. By using these patterns, we identified normal and abnormal ECG signals in different ECG data sets. Under tenfold cross-validation, the total number of input–output paired feature patterns was 1000. These patterns were divided into eight classes and into training and testing data sets. The proposed digital signal preprocessing, EPD, feature extraction, and classification algorithms were designed on a tablet PC by using MATLAB.

In the training stage, the BPA with GDO[33,34] was used to determine the optimal network parameters by using iteration computations to adjust the optimal network parameters under multiple learning rates (h = 0.05, 0.10, 0.20, 0.30, 0.40, and >0.40) for 260 randomly selected samples in each fold (40 •, 50 V, 30 A, 50 L, 30 R, 30 P, 20 F, and 10 f samples). The training process was terminated when the LF was $\leq 10^{-2}$ or when the number of iterations reached 5000. As shown in Fig. 6, under h values of 0.05, 0.10, 0.20, and 0.30, convergence was achieved after 227, 99, 50, and 196 iterations, respectively. In Fig. 6, the blue, green, red, and brown solid lines represent the convergence curves obtained under the aforementioned h values, respectively. As h was increased from 0.05 to 0.20, the model execution time decreased from 30.174 to 2.296 s and the number of iterations required to train the 2D classifier decreased from 227 to 50. A training accuracy of 100% was achieved for 130 random testing samples (20 •, 25 V, 15 A, 25 L, 15 R, 15 P, 10 F, and 5 f samples) when h was set as 0.20 (Table 1). Therefore, in the training stage, we set h as 0.20 to achieve rapid convergence.[37–39]

According to the AAMI standard, approximately 80% of heartbeats belong to the • class, and the remaining 20% are ventricular ectopic, supraventricular ectopic, or fusion heartbeats.[25] Thus, in this study, the EPD method was used to extract feature patterns from continuous ECG waveforms to segment these waveforms into classes. Figure 7 shows the color patterns obtained

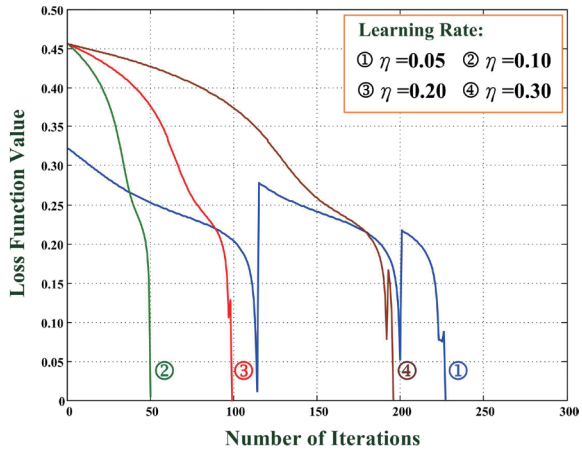

Fig. 6. (Color online) Training curves for LF value versus number of iterations under multiple learning rates.

Table 1
Performance comparisons with different learning rates.

Learning rate	Training dataset	Number of iterations	CPU time (s)
0.05	260	227	30.174
0.10	260	99	14.544
0.20	260	50	2.296
0.30	260	196	18.609
0.40	260	895	123.076
>0.40 (0.50)	260	3832	479.908

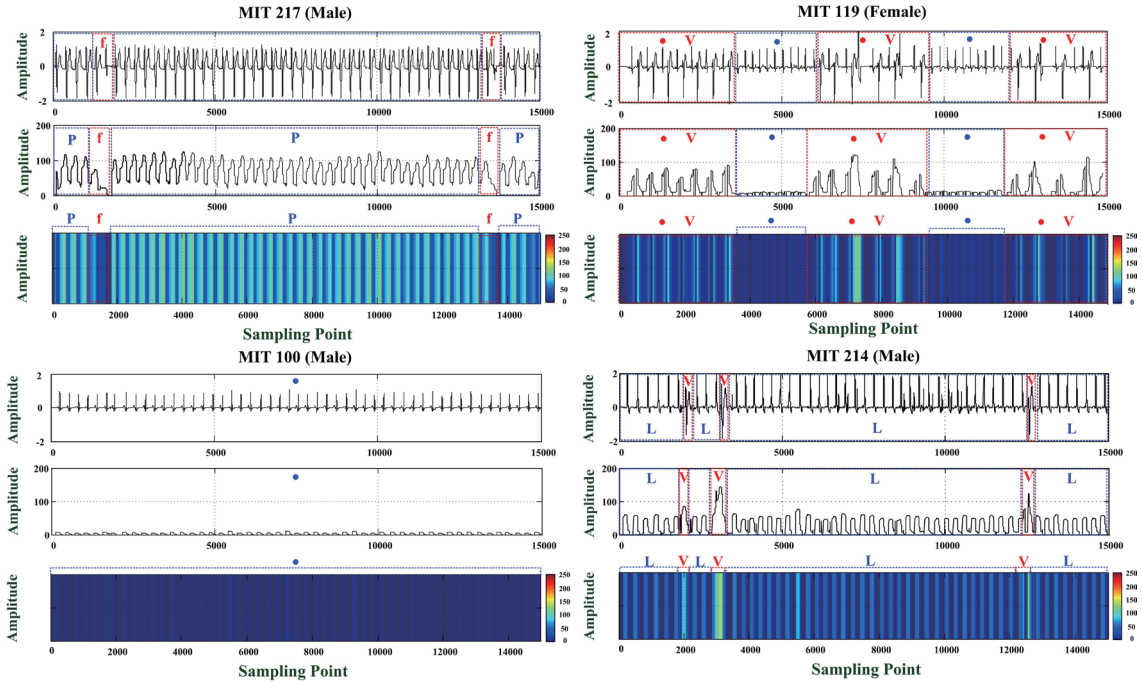

Fig. 7. (Color online) ECG feature extraction through EPD for the MIT 217 (P and f), MIT 119 (• and V), MIT 100 (•), and MIT 214 (L and V) files.

for normal, fusion, and different ectopic heartbeats from different ECG records. The proposed model used the pathological information contained in each feature pattern to recognize cardiac arrhythmias.[37–39]

Table 2 presents the experimental results obtained using the proposed model for 10 ECG files from the MIT-BIH Arrhythmia Database. From these 10 ECG files, we randomly selected raw ECG data for a duration of approximately 1 min (Fig. 7) to verify the proposed model's feasibility for automatic cardiac arrhythmia recognition. For the recognition of ventricular ectopic heartbeats (V) from the MIT 119, MIT 200, and MIT 221 records, the accuracies of the proposed model were 99.00% (24 TPs and 75 TNs), 92.00% (32 TPs and 60 TNs), and 99.00% (15 TPs and 84 TNs), respectively. In addition, the TN rates of these models (for the identification of normal heartbeats) were 98.68, 95.23, and 98.82%, respectively. For the MIT 111, MIT 118, MIT 214, and MIT 231 files, the proposed model identified 190 and 196 left and right bundle branch block

Table 2
Experimental results of cardiac arrhythmia recognition for ECG recording files, including MIT#100, #104, #111, #118, #119, #200, #214, #220, #221, and #231.

Record MIT		Beat Classes and Number of Rhythms							Accuracy (%)	Precision (%)	True negative rate (%)	Recall (%)	*F*1 score
		N	V	A	R	L	F	P					
100 Male	Actual	99	0	1	0	0	0	0	—	—	—	—	—
	Test 1	99	0	0	0	1	0	0	99.00	0.00	99.00	0.00	0.00
	Test 2	99	0	0	0	1	0	0	99.00	0.00	99.00	0.00	0.00
104 Female	Actual	0	0	0	0	0	31	69	—	—	—	—	—
	Test 1	14	1	0	0	0	16	69	85.00	85.00	0.00	100.00	0.9189
	Test 2	14	1	0	0	0	16	69	85.00	85.00	0.00	100.00	0.9189
111 Female	Actual	0	0	0	0	100	0	0	—	—	—	—	—
	Test 1	0	0	0	3	97	0	0	97.00	97.00	0.00	100.00	0.9848
	Test 2	0	0	0	3	97	0	0	97.00	97.00	0.00	100.00	0.9848
118 Male	Actual	0	1	0	99	0	0	0	—	—	—	—	—
	Test 1	0	0	0	96	4	0	0	96.00	96.00	0.00	100.00	0.9796
	Test 2	0	0	0	96	4	0	0	96.00	96.00	0.00	100.00	0.9796
119 Female	Actual	75	25	0	0	0	0	0	—	—	—	—	—
	Test 1	75	24	0	0	0	0	1	99.00	99.00	98.68	100.00	0.9796
	Test 2	75	24	0	0	0	0	1	99.00	99.00	98.68	100.00	0.9796
200 Male	Actual	65	35	0	0	0	0	0	—	—	—	—	—
	Test 1	60	32	0	0	5	3	0	92.00	91.43	95.23	86.49	0.8889
	Test 2	60	32	0	0	5	3	0	92.00	91.43	95.23	86.49	0.8889
214 Male	Actual	0	3	0	0	97	0	0	—	—	—	—	—
	Test 1	0	3	0	4	93	0	0	96.00	96.00	0.00	100.00	0.9648
	Test 2	0	3	0	4	93	0	0	96.00	96.00	0.00	100.00	0.9648
220 Female	Actual	100	0	0	0	0	0	0	—	—	—	—	—
	Test 1	100	0	0	0	0	0	0	100.00	0.00	100.00	0.00	0.00
	Test 2	100	0	0	0	0	0	0	100.00	0.00	100.00	0.00	0.00
221 Male	Actual	84	16	0	0	0	0	0	—	—	—	—	—
	Test 1	84	15	0	0	0	1	0	99.00	93.75	98.82	100.00	0.9677
	Test 2	84	15	0	0	0	1	0	99.00	93.75	98.82	100.00	0.9677
231 Female	Actual	0	0	0	100	0	0	0	—	—	—	—	—
	Test 1	0	0	0	100	0	0	0	100.00	100.00	0.00	100.00	1.0000
	Test 2	0	0	0	100	0	0	0	100.00	100.00	0.00	100.00	1.0000
Total	Actual	423	80	1	199	197	31	69	—	—	—	—	—
	Test	432	75	0	203	200	20	70	93.90	90.29	88.18	97.38	0.9317

Note: (1) Test#1: experimental results for 2D-CNN-based classifier; (2) Test#2: experimental results for 1D-CNN-based classifier.

beats (L and R, respectively), with the model accuracy exceeding 95.00% for both types of beat (Table 2). Ten bundle branch block beats were incorrectly classified by the model. For the MIT 104 file, the proposed model identified 16 F and 69 P samples, with the model's precision, recall, and *F*1 score being 85.00%, 100.00%, and 0.9189, respectively, for these samples. A total of 15 F or P samples were incorrectly classified by the model. For the MIT 100 and MIT 220 files, the proposed model exhibited a TN rate of 99.00% (for normal heartbeats). For 1000 heartbeats (423 normal and 577 abnormal heartbeats), the proposed model achieved precision, recall, and *F1* score of 90.29%, 97.38%, and 0.9317, respectively, for the identification of abnormal heartbeats (521 TPs, 56 FPs, and 14 FNs). Moreover, it exhibited a TN rate of 88.18% (418 TNs) for normal

heartbeats. Thus, the experimental results indicated that the proposed model is feasible for the automatic detection of cardiac arrhythmias from ECG signals.

Table 3 shows a comparison of the proposed model with other models from the relevant literature. Deep-learning–based classifiers, such as 1D and 2D CNN models, have been widely applied in ECG and cardiac arrhythmia classification.[37–42] In previous studies,[37,41] cardiac arrhythmia classification models have been created by combining the R-peak detection algorithm with a 1D CNN. These models have used the R-peak detection algorithm to segment each heartbeat and then extract QRS complex waves from ECG signals. The aforementioned models were tested on the MIT-BIH ECG Database[26,31] to identify different types of cardiac arrhythmia (N, V, S, F, and Q). In this process, 75% of the database was used for training, and the remaining 25% was used for validation or testing. The aforementioned models exhibited accuracy, precision, sensitivity, and specificity values greater than 90%. In other studies,[38,40] a bidirectional long short-term memory (BiLSTM) model and a 24-layer deep CNN model have

Table 3
Comparison of different classifiers for cardiac arrhythmia recognition with different ECG databases, methods, and purposes.

Reference	ECG Data	Method	Purpose
37	MIT-BIH ECG Database[26,31]	R-Peak Detection Algorithm + 1D Convolutional Deep Residual Neural Network	ECG Classification (N, V, S, F, & Q) Tenfold cross validation: Average accuracy: 98.63%, Average precision: 92.86%, Average sensitivity: 92.41%, Average specificity: 99.06%.
38	2017 PhysioNet/ CINC Challenge (8528 Single-lead ECG Records) Datasets[39]	Bidirectional Long Short-Term Memory (BiLSTM)[38,40] + 24-Layer Deep CNN (DCNN)	Normal and AF Classification Tenfold cross validation: Accuracy rate: 89.3%, *F1* score: 0.891.
40	MIT-BIH ECG Database[26,31]	LSTM + 1D-CNN	ECG Classification (N, L, R, A, & V) Tenfold cross validation: Average accuracy: 98.10%, Average sensitivity: 97.50%, Average specificity: 98.70%, Average PPV (positive predictive value): 98.69%.
41	MIT-BIH ECG Database[26,31]	R-Peak Detection Algorithm + 1D-CNN	Cardiac Arrhythmia Classification (F, N, S, & V) Average accuracy: 0.99, Average sensitivity: 0.94, Average specificity: 0.99.
42	MIT-BIH and Sudden Death Cardiac Holter (SCDH) Arrhythmia Databases	R-Peak Detection Algorithm + 1-D CNN	ECG Classification (N, L, R, V) MIT-BIH: 99.5% (97.6% → 99.5%), SCDH: 88.5% (80.2% → 88.5%), Training time decrease: 67.2 and 64.2% for MIT-BIH and SCDH.
Proposed Method	MIT-BIH ECG Database[26,31]	EPD + 2D-CNN EPD + 1D-CNN	Cardiac Arrhythmia Recognition Accuracy: 93.90%, Precision: 90.29%, Recall: 97.38%, *F1* score: 0.9317, True negative rate: 88.18%.

been combined to establish a dense heart rhythm model for the automatic classification of heartbeats, including normal heartbeats and heartbeats indicating AF. The aforementioned BiLSTM model consists of a forward LSTM network and a backward LSTM network, both of which use sequence information in the backward (future to past) and forward (past to future) directions to extract time-sensitive features of ECG signals. In addition, a multilayered CNN with 32, 64, and 128 convolution kernels was used to extract detailed feature patterns for normal heartbeats and AF. In one study,[39] tenfold cross-validation was conducted on the 2017 PhysioNet/CINC Challenge data sets (comprising 8528 single-lead ECG records) to verify the "BiLSTM + deep CNN" model, which exhibited the accuracy and *F1* score of 89.3% and 0.891, respectively, for automatic AF monitoring.[39] When the "R-peak detection algorithm + 1D CNN" model was used on the MIT-BIH Arrhythmia Database and Sudden Cardiac Death Holter Database for arrhythmia classification (N, L, R, and V),[42] the classification accuracy increased from 97.6 to 99.5% and from 80.2 to 88.5%, respectively (Table 3). In addition, the training time on these data sets decreased by 67.2 and 64.2%, respectively. However, the aforementioned model has a seven-layer structure; thus, it has a complex structure and requires an excessive number of feature data sets for its training. The proposed model exhibited promising results in automatic cardiac arrhythmia recognition for 423 normal and 577 abnormal heartbeats, with its accuracy, precision, and recall exceeding 90.00% for the recognition of abnormal heartbeats.

4. Conclusions

In this study, the EPD method was combined with a CNN classifier to develop a model for the automatic recognition of cardiac arrhythmias. This model uses digital filters to remove baseline drift and noise from ECG signals, and it employs the EPD method to extract QRS complex waves from time-domain ECG data. The proposed model detects evident electrical activities on the basis of appropriate threshold values and then transfers them to 2D or 1D color feature patterns. The proposed model was used to classify ECG signals (into the normal, V, A, L, R, P, F, and f classes) for cardiac arrhythmia recognition. It exhibited the precision, recall, and *F1* score of 90.29%, 97.38%, and 0.9317, respectively, for abnormal heartbeat recognition. In addition, it achieved an accuracy rate of 93.90% for the recognition of both normal and abnormal heartbeats and exhibited a TN rate of 88.18% for the identification of normal heartbeats. On the basis of these results, the proposed model has promising results for application in the clinical detection of cardiac arrhythmias. The proposed method can accurately extract the feature patterns from typical QRS complex waves; however, some special rhythms, such as AFs, atrial flutter, and ventricular flutter, are difficult to detect under time-varying and noisy ECG signals and heart rates. Thus, for typical classes, the proposed model ("EPD + 2D-CNN" or "WEPD + 1D-CNN") has significant results in cardiac arrhythmia recognition.

Acknowledgments

This work was supported by the National Science and Technology Council (NSTC) under contract number NSTC 112-2221-E-167-015 (duration: August 1, 2023–July 31, 2024).

References

1 J. Allen: Physiol. Meas. **28** (2007) R1-39. https://doi.org/10.1088/0967-3334/28/3/R01
2 P. Kyriacou and J. Allen: Signal Analysis and Applications (2021) 1st ed.
3 J.-X. Wu, C.-H. Lin, C.-D. Kan, and W.-L. Chen: IET Sci. Meas. Technol. **13** (2019) 1277. https://doi.org/10.1049/iet-smt.2018.5330
4 S. M.-González, J. L. N.-Mesa, G. J.-Serdá, J. F. Kraemer, N. Wessel, and A. G. R.-García: Comput. Biol. Med. **91** (2017) 47. https://doi.org/10.1016/j.compbiomed.2017.10.004
5 C.-H. Lin, J.-X. Wu, P.-Y. Chen, C.-M. Li, N.-S. Pai, and C.-L. Kuo: IEEE Access **9** (2021) 26451. https://doi.org/10.1109/ACCESS.2021.3057586
6 U. Satija, B. Ramkumar, and M. S. Manikandan: IEEE Sens. J. **19** (2019) 277. https://doi.org/10.1109/JSEN.2018.2877055
7 H. Jin, S. Yang, F. Yang, L. Zhang, H. Weng, S. Liu, F. Fan, H. Li, X. Zheng, H. Yang, Y. Zhang, J. Zhou, and J. Li: J. Transl. Int. Med. **9** (2021) 285. https://doi.org/10.2478/jtim-2021-0042
8 C. R. da S. Araújo, J. Fernandes, D. S. Caetano, A. E. V. do R. Barros, J. A. F. de Souza, M. da G. R. Machado, M. I. R. de Aguiar, S. C. S. Brandão, S. L. Campos, A. de F. D. de Andrade, and D. C. Brandão: Heart Lung **58** (2023) 210. https://doi.org/10.1016/j.hrtlng.2022.12.016
9 T. Paul, O. Hassan, K. Alaboud, H. Islam, K. Z. Rana, S. K. Islam, and A. S. M. Mosa: AMIA Jt Summits Transl. Sci. Proc. **23** (2022) 379.
10 M. Faal and F. Almasganj: Biomed. Signal Process. Control **68** (2021) 1. https://doi.org/10.1016/j.bspc.2021.102685
11 C. Levkov, G. Mihov, R. Ivanov, I. Daskalov, I. Christov, and I. Dotsinsky: Biomed. Eng. Online **4** (2005) 1. https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-4-50.
12 J. Wannenburg, R. Malekian, and G. P. Hancke: IEEE Sens. J. **18** (2018) 6023. https://doi.org/10.1109/JSEN.2018.2844122
13 A. Paul, M. S. Lee, Y. Xu, S. R. Deiss, and G. Cauwenberghs: IEEE Trans. Biomed. Circuits Syst. **17** (2023) 483. https://doi.org/10.1109/TBCAS.2023.3272649
14 C. Ye, B. V. K. V. Kumar, and M. T. Coimbra: IEEE Trans. Biomed. Eng. **59** (2012) 2930. https://doi.org/10.1109/TBME.2012.2213253
15 Q. Li, C. Rajagopalan, and G. D. Clifford: IEEE Trans. Biomed. Eng. **61** (2014) 1607. https://doi.org/10.1109/TBME.2013.2275000
16 M. K. Das and S. Ari: Healthcare Technol. Lett. **1** (2014) 98. https://doi.org/10.1049/htl.2014.0072
17 J. Kim, S. D. Min, and M. Lee: Biomed. Eng. Online **10** (2011) 1. https://biomedical-engineering-online.biomedcentral.com/articles/10.1186/1475-925X-10-56
18 L. Rozaqi, A. Nugroho, K. H. Sanjaya, and A. I. Simbolon: 2019 Int. Conf. Sustainable Energy Engineering and Application (ICSEEA) (2019) 186. https://doi.org/10.1109/ICSEEA47812.2019.8938645
19 S. Selvaraj, P. Ramya, R. Priya, and C. Ramya: 2021 3rd Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (2021) 185. https://doi.org/10.1109/ICICV50876.2021.9388515
20 M. Nazari and K.-S. Lee: 2021 IEEE Int. Midwest Symp. Circuits and Systems (MWSCAS) (2021) 916. https://doi.org/10.1109/MWSCAS47672.2021.9531689
21 U. Satija, B. Ramkumar, and M. S. Manikandan: IEEE. J. Biomed. Health. Inf. **22** (2018) 722. https://doi.org/10.1109/JBHI.2017.2686436.
22 U. Satija, B. Ramkumar, and M. S. Manikandan: Biocybern. Biomed. Eng. **38** (2018) 54. https://doi.org/10.1016/j.bbe.2017.10.002
23 J. Wu, G. Chong, W. Pang, and L. Wang: 2023 5th Int. Conf. Natural Language Processing (ICNLP) (2023) 126. https://doi.org/10.1109/ICNLP58431.2023.00029
24 Q. Wu and Y. Li: 2021 Int. Conf. Electronic Information Technology and Smart Agriculture (ICEITSA) (2021) 1. https://doi.org/10.1109/ICEITSA54226.2021.00010
25 T. Zhang, Y. Shao, Y. Wu, Y. Geng, and L. Fan: Appl. Acoust. **160** (2020) 1. https://doi.org/10.1016/j.apacoust.2019.107133
26 P. Kumar and V. K. Sharma: Healthcare Technol. Lett. **7** (2020) 18. https://doi.org/10.1049/htl.2019.0096
27 D. Hayn, B. Jammerbund, and G. Schreier: Physiol. Meas. **33** (2012) 1449. https://doi.org/10.1088/0967-3334/33/9/1449
28 M. A. Uddin, R. K. Pathan, M. Hossain, and M. Biswas: J. Inf. Telecommun. **6** (2022) 27. https://doi.org/10.1080/24751839.2021.1983318
29 A. A. Ahmed, W. Ali, T. A. A. Abdullah, and S. J. Malebary: Mathematics **11** (2023) 1. https://doi.org/10.3390/math11030562

30 D. Li, J. Zhang, Q. Zhang, and X. We: 2017 IEEE 19th Int. Conf. e-Health Networking, Applications and Services (Healthcom) (2017) 1. https://doi.org/10.1109/HealthCom.2017.8210784
31 G. B. Moody and R. G. Mark: Comput. Cardiol. **17** (1990) 185. https://doi.org/10.1109/CIC.1990.144205
32 Butterworth filter design - MATLAB butter (2024). https://www.mathworks.com/help/signal/ref/butter.html
33 P.-Y. Chen, Y.-C. Cheng, Z.-H. Zhong, F.-Z. Zhang, N.-S. Pai, C.-M. Li, and C.-H. Lin: IEEE Access **12** (2024) 9757. https://doi.org/10.1109/ACCESS.2024.3351373
34 F.-Z. Zhang, C.-H. Lin, P.-Y. Chen, N.-S. Pai, C. Su, C.-C. Pai, and H.-W. Ho: Processes **10** (2022) 1. https://doi.org/10.3390/pr10091867
35 Z. Chen, A. Francis, S. Li, B. Liao, D. Xiao, T. T. Ha, J. Li, L. Ding, and X. Cao: Biomimetics **7** (2022) 1. https://doi.org/10.3390/biomimetics7040144
36 C.-H. Lin, C.-D. Kan, J.-N. Wang, W.-L. Chen, and P.-Y. Chen: IEEE Access **6** (2018) 52652. https://doi.org/10.1109/ACCESS.2018.2870689
37 F. Khan, X. Yu, Z. Yuan, and A. Rehman: PLOS ONE. **18** (2023) 1. https://doi.org/10.1371/journal.pone.0284791
38 J. Cheng, Q. Zou, and Y. Zhao: BMC Med. Inf. Dec. Making **21** (2021) 1. https://Bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01736-y
39 G. D. Cliford, C. Liu, B. Moody, L. H. Lehman, I. Silva, Q. Li, A. E. W. Johnson, and R. G. Mark: Comput. Cardiol. **44** (2017) 1. https://doi.org/10.22489/CinC.2017.065-469
40 Shu Lih Oha, Eddie Y. K. Ngb, Ru San Tanc, and U. Rajendra Acharya: Comput. Biol. Med. **102** (2018) 278. https://doi.org/10.1016/j.compbiomed.2018.06.002
41 Adel A. Ahmed, Waleed Ali, Talal A. A. Abdullah, and Sharaf J. Malebary: Mathematics **11** (2023) 1. https://doi.org/10.3390/math11030562
42 S. W. Chen, S. L. Wang, X. Z. Qi, T. F. Ng, and H. Ibrahim: Multimedia Tools Appl. **82** (2023) 45811. https://doi.org/10.1007/s11042-023-15407-9