Computer Engineering and Applications Vol. 9, No. 2, June 2020

Automated ECG Waveform Annotation Based on Stacked Long Short-Term Memory

Annisa Darmawahyuni, Siti Nurmaini*, Muhammad Naufal Rachmatullah, Vicko Bhayyu

Intelligent System Research Group, Universitas Sriwijaya, Palembang 30137, Indonesia
**siti_nurmaini@unsri.ac.id*

ABSTRACT

The classification of electrocardiogram (ECG) waveform segmentation techniques can be difficult due to physiological variation of heart rate and different characteristics of the different ECG waves in terms of shape, frequency, amplitude, and duration. The P-wave, PR-segment, QRS-complex, ST-segment, and T-wave are extracted as the feature for classification algorithm to diagnose specified cardiac disorders. This requires the implementation of algorithms that identify specific points within the ECG wave. Some previous computational algorithms for automatic classification of ECG segmentation are proposed to overcome limitations of manual inspection of the ECG. This study presents new insight into the ECG semantic segmentation problem is surmounted by a deep learning approach for automatic ECG wave-form. Long short-term memory (LSTM) is proposed for this task. This experimental study has been performed for six different waveforms of ECG signal that represents cardiac disorders obtained from the Physionet: QT database. Overall, LSTM performance achieved accuracy, sensitivity, specificity, precision, F1-score, is 93.36%, 86.85%, 95.78%, 81.79%, and 83.09%, respectively.

Keywords: ECG, Segmentation, seq2seq, LSTM, Classification

1. INTRODUCTION

A cardiologist analyzes the electrical function of the cardiac via electrocardiogram (ECG). Analysis of critical segments of ECG is a crucial thing for diagnosing cardiac disorders. However, in some cases, ECG-based diagnosis can be difficult [1]. For examples, diagnosis of myocardial infarction (MI), one of coronary heart disease due to a lack of oxygen demand in the cardiac muscle tissue [2][3]. The ECG form changes in ST-elevation, T-waveform, and the ST interval length. In other cases, the detection of QT and RR interval for calculating QT corrected (QTc) to inform Long QT Syndrome (LQTS), and P-wave abnormality to diagnose Atrial Fibrillation, are difficult due to the lack of symptoms [4][5][6]. A cardiologist needs to analyze ECG recordings that acquired over several hours or even days, making the task very troublesome and time-consuming [1]. Such limitations can be reduced by advanced computing systems permit the automatic interpretation of ECG. Precise ECG is essential in getting maximum benefits to properly interpret ECG recordings. The additional features of ECG need to be analyzed the morphology of the different waves within the signal. The computing algorithms require specific points within the ECG wave segmentation, include the start (on) to end (off) of the P, QRS, and T waves.

Automatic segmentation of the ECG is challenging due to the reduced amplitude of the P-wave, the high variability in the shape of the QRS complex, and the smooth transitions of the beginning and end of T-wave [7][8][9]. The previous conventional algorithm is applied for automatic ECG segmentation in many works of literature. Laguna et al., [10] firstly proposed a very successful approach in ECG segmentation based on second-order bandpass filtering the ECG and then differentiating it. In the end, different waves would be detected based on their zero-crossings and finding the nearest points exceeding empirical thresholds. Otherwise, some method is proposed for ECG segmentation, such as hidden Markov model [11], multiple higher-order moments (MHOM) Metric [12], continuous wavelet transform [13], a band-pass Butterworth filter and the first derivative [14], etc.

This study aims at providing new insights into the ECG semantic segmentation problem (i.e seq2seq) using a deep learning approach. Deep learning is still outstanding due to an automatic process of feature calculation without predetermining some appropriate features. This study proposes Long Short-Term Memory (LSTM), a family of Recurrent Neural Network (RNN) with its time correlation for classifying three different waveforms for ECG. In our previous work [2], LSTM works well for Myocardial Infarction classification compared to other recurrent networks. For this reason, the segmentation of beat-by-beat is applied by classifying P-wave (P_{on}-P_{off}), PR-segment (P_{off}-QRS_{on}), QRS-complex (QRS_{on}-R_{peak}, R_{peak}-QRS_{off}), ST-segment (QRS_{off}-T_{on}), and T-wave (T_{on}-T_{off}) based on LSTM algorithm. The classification of ECG segmentation techniques is used for detecting specific points to diagnose wave pattern abnormality.

In the following section, Section 2, with a detailed description of ECG database and pre-processing. Section 3 presents a classification method based on LSTM and describe the proposed architecture. For model evaluation, Section 4 discusses performance metrics in terms of accuracy, sensitivity, specificity, precision, and F1-score to analyze how well the model works. Finally, Section 5 state the conclusion of this study.

2. MATERIAL AND METHOD

In this study, a new input layer with wavelet basis was designed for the multiclass classification of start to the end P-wave, QRS-complex, and T-wave (P_{on}-P_{off}, P_{off}-QRS_{on}, QRS_{on}-R_{peak}, R_{peak}-QRS_{off}, QRS_{off}-T_{on}, and T_{on}-T_{off}) from standard deep learning layers. The input signals are decomposed into wavelet levels and transferred to other layers as sequences. This study conducted the experiment on the well-known Physionet: QT Database (QTDB), which the training and testing set was generated. LSTM networks and architectures were designed for the classification process.

2.1 THE PHYSIONET: QT DATABASE

ECG raw data is obtained from the open-access of the Physionet public dataset: The QT Database (QTDB) [15]. The database consists of 105 ECG records taken in 15 min with a sampling frequency of 250 Hz from two-channel Holter ECG recordings. This database consists of 15 records of MIT-BIH Arrhythmia, 6 records of MIT-BIH ST Change, 13 records of MIT-BIH Supraventricular Arrhythmia, 33 records of European ST-T, 24 records of sudden death patients from BIH, 4 records of MIT-BIH Long-Term ECG, and 10 records of MIT-BIH Normal Sinus Rhythm.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

For this study, the records of sudden death are excluded. This database provides the input to the WFDB function *ecgpuwave*(), which gives us the exact position of all the P, R, and T peaks found in the signal. The output of the *ecgpuwave* is written as a standard WFDB-format annotation file associated with the specified annotator. It is utilized as “ground truth” or label for the proposed ECG segmentation algorithm. Otherwise, only a complete waveform pattern of P-wave, QRS-complex, and T-wave is utilized in this study. Points of interest within the ECG include the P_{on}-P_{off}, P_{off}-QRS_{on}, QRS_{on}-R_{peak}, R_{peak}-QRS_{off}, QRS_{off}-T_{on}, and T_{on}-T_{off}. The location of these points is plotted in Figure 1.

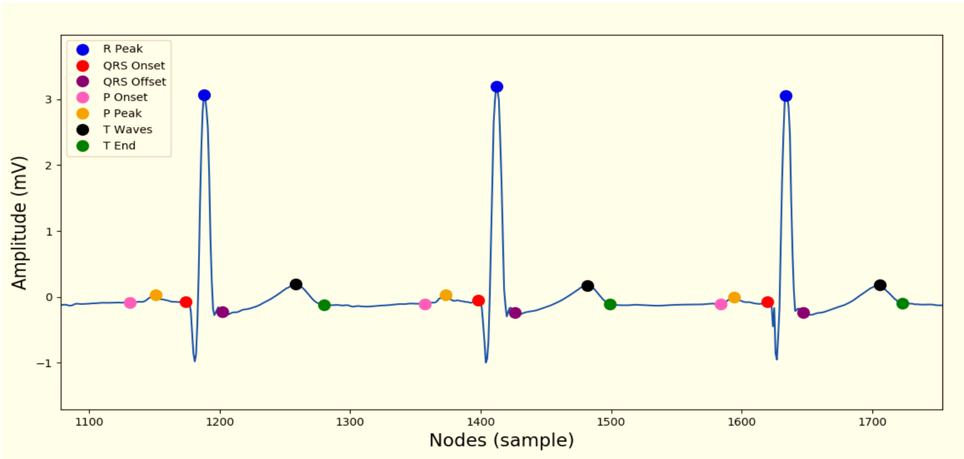

FIGURE 1. The $Point_{on}$-$Point_{off}$ ECG Segmentation

2.2 WAVELET TRANSFORM

The vital sign from cardiac patients monitored by ECG. The changes of ECG waveforms indicate the cardiac abnormality that may occur for any reason. While acquisition, ECG can get corrupted due to different types of artifacts and power line interference [16]. Before transferred to the classification process, ECG signals are enhanced by eliminating various kinds of noise and artifacts. Wavelet transform (WT) is applied to the process with reconstruction a signal from a noisy one. WT performs a correlation analysis, therefore the output is expected to be maximal when the input signal most resembles the mother wavelet. Unlike the old denoising method (i.e Fourier transform), WT provides an analysis of the signal which is localized in both time and frequency. In contrast, the Fourier transform is localized only in frequency. Given a mother wavelet $\varphi\ (t)$ (which can be considered simply as a basic function of L^2), the continuous wavelet transform (CWT) of a function $x\ (t)$ (assuming that $x \in L^2$) is defined as [17]:

$$X\ (a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \varphi \left(\frac{t-b}{a}\right) x(t)\ dt \tag{1}$$

where $\varphi\ (t)$ is basic waveforms or functions, dilation a corresponds to frequency information, and translation b relates to the locations of the wavelet function.

The algorithm of wavelet denoising mainly contains three steps: wavelet decomposition, coefficient processing, and wavelet reconstruction [18]. The common wavelet families such as *daubechies, biorthogonal, coiflet,* and *symlet* can be used for ECG signal denoising [19]. Among them, this study applies the biorthogonal ("bior") wavelet due to this wavelet function can remove noise successfully [20][21][22]. Qin et. al. proposes the decomposition procedure of 8-level WMRA using bior6.8 wavelet, soft thresholding [7]. The results of bior6.8 wavelet in QTDB can be shown in Figure 2.

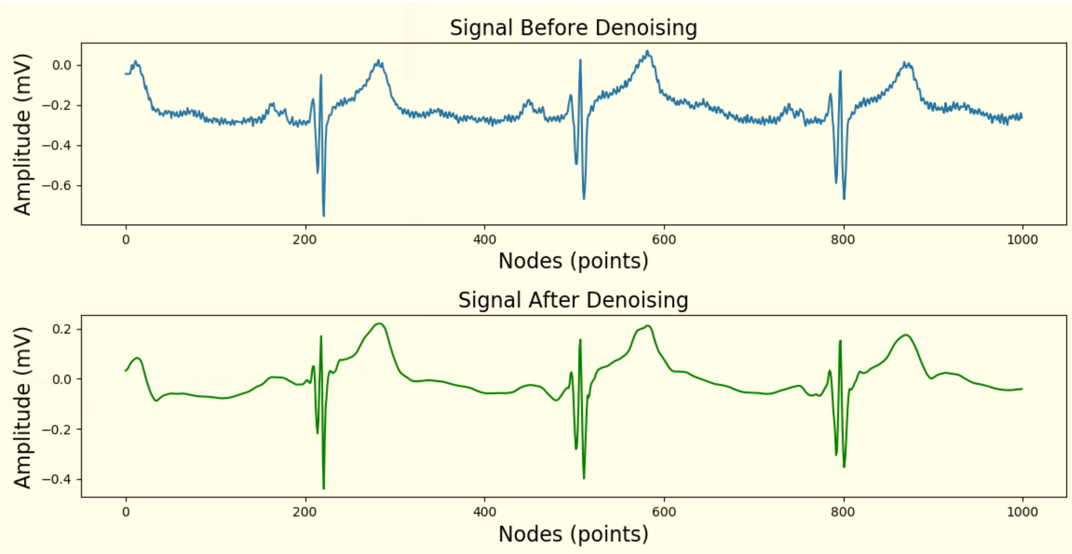

FIGURE 2. Wavelet Transform Processing of Samples in QT Database

3. LONG SHORT-TERM MEMORY

The most widely used a family of the recurrent neural network (RNN) is Long Short-Term Memory (LSTM). Each node in LSTM is a cell that comprises input, forget, and output gates. The common of standard RNN problems caused by the iterative nature, whose gradient is essentially equal to the recurrent weight matrix raised to high power. The gradient to grow or to shrink at a rate that is exponential in the number of timesteps [23]. With the gating mechanism that control the manner in which internal states are retained or discarded, LSTM overcomes the gradient problems [2]. Mathematically, the LSTM can be written as [24]:

$$c_t = \sigma\left(W_f I_t\right) c_{t-1} + \sigma(W_i I_t) \tanh(W_{in} I_t), \tag{2}$$

$$h_t = \sigma(W_o I_t) \tanh(c_t), \tag{3}$$

where σ is a sigmoid function, $c_t \in R^N$, column vector $I_t \in R^{(M+N)}$ is a concatenation of the current input, $X_t \in R^M$ and the previous output, $h_{t-1} \in R^N$.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

Under the assumption $c_0 = 0$, the hidden state vector of LSTM can be derived by:

$$c_t = \sum_{k=1}^{t}\left[\prod_{j=k+1}^{t}\sigma\left(W_f I_j\right)\right]\sigma(W_i I_k)\tanh\left(W_{in} I_k\right) \quad (4)$$

The comparation of output RNN (in Equation 5) and the gating mechanism in LSTM (in Equation 6) are given by:

$$h_t = \tanh\left(\sum_{k=1}^{t} W_c^{t-k} W_{in} X_k\right) \quad (5)$$

$$h_t = \sigma(W_o I_t)\tanh\left(\sum_{k=1}^{t}\left[\prod_{j=k+1}^{t}\sigma\left(W_f I_j\right)\right]\sigma(W_i I_k)\tanh\left(W_{in} I_k\right)\right) \quad (6)$$

where $W_i, W_f,\ W_o,\ W_{in}$ are weight matrices for the input gate, forget gate, output gate, and the input, respectively.

In this study, we propose unidirectional LSTM architecture for the multiclass classification process. The LSTM layer is stacked by three layers. Each layer had 512 nodes and loss function was selected as the categorical cross-entropy. Time steps of LSTM are 2 seconds. The Adam optimization method was used in this study. Each classifier was trained for 100 epochs to ensure consistent comparisons, with a batch size of 32. The proposed LSTM architecture can be presented in Figure 3.

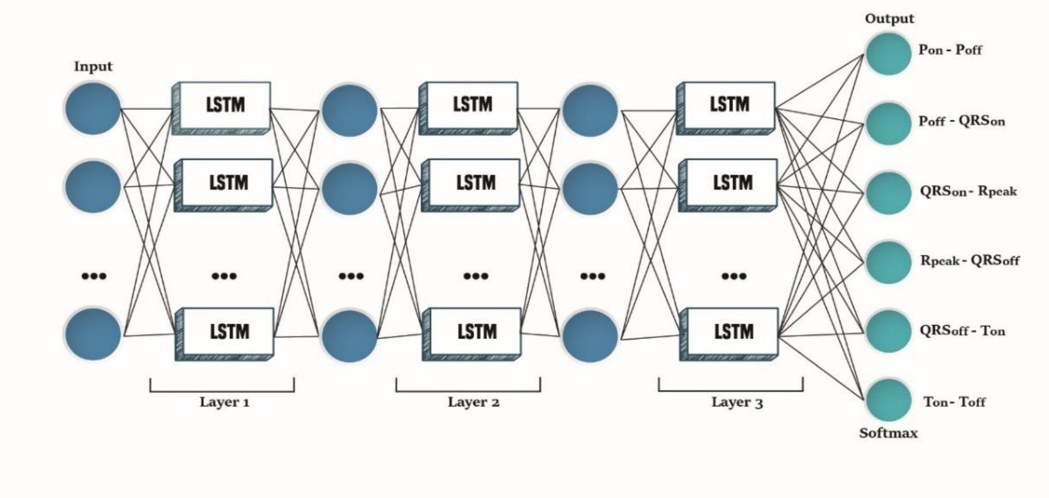

FIGURE 3. The Proposed Stacked-LSTM Architecture

4. RESULTS AND DISCUSSION

This study proposes a stacked-LSTM architecture with three layers for the automatic classification of the ECG segments. All 76 records of QTDB were divided as 90% in the training phase and the remaining 10% in the test phase with shuffle sampling. The testing as a validation set was used to tune the parameters and determine the optimal unit numbered of designed models. We utilized a workstation with NVIDIA GeForce RTX 2080 and CuDNN which is a GPU-accelerated deep neural network library that supports the training of LSTM for sequence learning. The results of LSTM performance for QTDB can be described in Table 1. Each start to the end of P-wave, QRS-complex, and T-wave is classified and measured by common performance metrics. Table 1 shows the performance of the ECG segment from six databases of QTDB. Overall, with the proposed LSTM architecture, segmentation of QRS-complex (QRS_{on}-R_{peak}, R_{peak}-QRS_{off}) perform well in all QTDB. The poor result of segmentation of P-wave and T-wave can be seen in Arrhythmia, Supraventricular Arrhythmia, European ST-T, and Long-Term ECG Database.

TABLE 1.
The LSTM Performance of Testing Set in QT Database

QTDB	Metrics	LSTM Performance (%)						
		P_{on}-P_{off}	P_{off}-QRS_{on}	QRS_{on}-R_{peak}	R_{peak}-QRS_{off}	QRS_{off}-T_{on}	T_{on}-T_{off}	Average
Arrhythmia	Accuracy	82.79	98.21	98.52	98.89	96.90	85.40	93.45
	Sensitivity	65.02	88.93	91.08	95.62	92.01	95.89	88.09
	Specificity	98.08	98.85	98.97	99.06	97.66	82.23	95.81
	Precision	96.68	84.34	84.18	84.26	85.89	61.96	82.88
	F1-Score	77.75	86.57	87.49	89.59	88.84	75.28	84.25
ST Change	Accuracy	89.92	98.02	98.16	98.87	97.44	92.12	95.75
	Sensitivity	83.15	77.08	90.80	89.41	92.47	93.72	87.77
	Specificity	95.72	99.25	98.70	99.55	98.23	91.69	97.19
	Precision	94.32	85.70	83.80	93.42	89.32	75.01	86.93
	F1-Score	88.38	81.16	87.16	91.37	90.87	83.33	87.04
Supraventricular Arrhythmia	Accuracy	80.76	96.60	98.16	98.65	94.82	85.92	92.49
	Sensitivity	63.00	83.54	84.25	91.22	91.80	95.20	84.84
	Specificity	97.50	97.56	98.96	99.06	95.36	83.78	95.37
	Precision	95.97	71.44	82.45	84.45	78.05	57.45	78.30
	F1-Score	76.07	77.02	83.34	87.71	84.37	71.65	80.02
Normal Sinus Rhythm	Accuracy	83.56	97.97	98.25	99.52	97.56	86.77	93.94
	Sensitivity	71.42	92.20	87.20	97.38	91.69	94.56	89.08
	Specificity	96.72	98.40	98.76	99.63	98.41	84.90	96.13
	Precision	95.93	81.01	76.39	92.93	89.20	59.98	82.57
	F1-Score	81.88	86.24	81.44	95.10	90.42	73.40	84.75
European ST-T	Accuracy	80.92	97.80	98.33	98.89	96.96	83.34	92.71
	Sensitivity	66.21	84.28	87.31	91.10	88.70	92.88	85.08
	Specificity	95.59	98.67	99.00	99.34	98.10	80.85	95.26
	Precision	93.73	80.32	84.19	88.89	86.49	55.92	81.59
	F1-Score	77.60	82.25	85.72	89.98	87.58	69.81	82.16
Long-Term ECG	Accuracy	78.31	96.95	98.11	98.72	96.77	82.05	91.82
	Sensitivity	57.19	92.50	93.79	97.53	82.40	94.23	86.27
	Specificity	98.30	97.36	98.41	98.83	97.88	78.74	94.92
	Precision	96.96	76.44	80.18	87.73	74.98	54.66	78.49
	F1-Score	71.95	83.71	86.45	92.37	78.51	69.18	80.36

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

For segmentation of P-wave (P_{on}-P_{off}), MIT-BIH Arrhythmia achieved accuracy, specificity, and precision, is 82.79%, 98.08%, and 96.68%, respectively. A poor result of sensitivity and F1-score, is 65.02% and 77.75%, respectively. The segmentation of P-wave performs well in ST Changes and Normal Sinus Rhythm Database. Overall, the performance of the segmentation of PR-segment (P_{off}-QRS_{on}) shows a good result in all databases. Like PR-segment, the segmentation of QRS-complex (QRS_{on}-R_{peak}, R_{peak}-QRS_{off}) is also obtained a good performance. In contrast, the segmentation of T-wave is not really good in all databases, except in ST Change. As we can see, the ST Change Database was correctly classified at accuracy, sensitivity, specificity, precision, F1-score, is 95.75%, 87.77%, 97.19%, 86.93%, 87.04%, respectively. The comparison of the average performance of LSTM can be seen in Table 2. For all QTDB, the LSTM classifiers obtained the accuracy, sensitivity, specificity, precision, F1-score, is 93.36%, 86.85%, 95.78%, 81.79%, and 83.09%, respectively.

TABLE 2.
The Average of LSTM Performance in The Testing Set

Metrics	LSTM Performance (%)						
	Arrhythmia	ST Change	Supraventricular Arrhythmia	Normal Sinus Rhythm	European ST-T	Long-Term ECG	QTDB Average
Accuracy	93.45	95.75	92.49	93.94	92.71	91.82	93.36
Sensitivity	88.09	87.77	84.84	89.08	85.08	86.27	86.85
Specificity	95.81	97.19	95.37	96.13	95.26	94.92	95.78
Precision	82.88	86.93	78.30	82.57	81.59	78.49	81.79
F1-Score	84.25	87.04	80.02	84.75	82.16	80.36	83.09

5. CONCLUSION

This study focuses on the segmentation of ECG-based annotation. The segmentation of the desired points of ECG is a challenging task due to the variation of characteristic signals. The new insight of segmentation is proposed by using stacked-LSTM networks, which recently have provided a significant achievement in the preliminary investigation. LSTM as classifier can overcome the semantic segmentation problem. The challenge to construct the right model of LSTM is feature selection. For future work, it can be more explored to obtain a better result of LSTM performance in this study.

ACKNOWLEDGEMENTS

This research is supported by the Kemenristekdikti Indonesia under the Basic Research Fund Number. 096/SP2H/LT/DRPM/2019 and Universitas Sriwijaya, Indonesia under Hibah Unggulan Profesi Fund 2019.

REFERENCES

[1] P. Sodmann, M. Vollmer, N. Nath, and L. Kaderali, "A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms," *Physiol. Meas.*, vol. 39, no. 10, p. 104005, 2018.

[2] A. Darmawahyuni, S. Nurmaini, W. Caesarendra, V. Bhayyu, M. N. Rachmatullah, and others, "Deep Learning with a Recurrent Network Structure in the Sequence Modeling of Imbalanced Data for ECG-Rhythm Classifier," *Algorithms*, vol. 12, no. 6, p. 118, 2019.

[3] A. Darmawahyuni, "Coronary Heart Disease Interpretation Based on Deep Neural Network," *Comput. Eng. Appl. J.*, vol. 8, no. 1, 2019.

[4] M. A. C. Gil, "A new, simpler and better correction formula for the QT interval," *J. Am. Coll. Cardiol.*, vol. 61, no. 10 Supplement, p. E294, 2013.

[5] B. Vandenberk *et al.*, "Which QT correction formulae to use for QT monitoring?," *J. Am. Heart Assoc.*, vol. 5, no. 6, p. e003264, 2016.

[6] N. Nuryani, M. Solikhah, A. S. Nugoho, A. Afdala, and E. Anzihory, "RR-Interval variance of electrocardiogram for atrial fibrillation detection," in *Journal of Physics: Conference Series*, 2016, vol. 776, no. 1, p. 12105.

[7] I. Beraza and I. Romero, "Comparative study of algorithms for ECG segmentation," *Biomed. Signal Process. Control*, vol. 34, pp. 166–173, 2017.

[8] S. Nurmaini, A. Gani, and others, "Cardiac Arrhythmias Classification Using Deep Neural Networks and Principle Component Analysis Algorithm.," *Int. J. Adv. Soft Comput. Its Appl.*, vol. 10, no. 2, 2018.

[9] Siti Nurmaini, Radiyati Umi Partan, Muhammad Naufal Rachmatullah, "Deep Neural Networks Classifiers on The Electrocardiogram Signal for Intelligent Interpretation System," *Sriwij. Int. Conf. Med. Sci.*, 2018.

[10] P. Laguna, R. Jané, and P. Caminal, "Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database," *Comput. Biomed. Res.*, vol. 27, no. 1, pp. 45–60, 1994.

[11] M. Akhbari, M. B. Shamsollahi, O. Sayadi, A. A. Armoundas, and C. Jutten, "ECG segmentation and fiducial point extraction using multi hidden Markov model," *Comput. Biol. Med.*, vol. 79, pp. 21–29, 2016.

[12] A. Ghaffari, M. R. Homaeinezhad, M. Khazraee, and M. M. Daevaeiha, "Segmentation of holter ECG waves via analysis of a discrete wavelet-derived multiple skewness--kurtosis based metric," *Ann. Biomed. Eng.*, vol. 38, no. 4, pp. 1497–1510, 2010.

[13] M. V\'\itek, J. Hrubeš, and J. Kozumpl\'\ik, "A wavelet-based ECG delineation with improved P wave offset detection accuracy," *Anal. Biomed. Signals Images*, vol. 20, pp. 160–165, 2010.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

[14] C. R. Vázquez-Seisdedos, J. E. Neto, E. J. M. Reyes, A. Klautau, and R. C. L. de Oliveira, “New approach for T-wave end detection on electrocardiogram: Performance in noisy conditions,” *Biomed. Eng. Online*, vol. 10, no. 1, p. 77, 2011.

[15] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, “A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG,” in *Computers in cardiology 1997*, 1997, pp. 673–676.

[16] J. Wang, Y. Ye, X. Pan, and X. Gao, “Parallel-type fractional zero-phase filtering for ECG signal denoising,” *Biomed. Signal Process. Control*, vol. 18, pp. 36–41, 2015.

[17] R. Cohen, “Signal denoising using wavelets,” 2012.

[18] Z. Wang, J. Zhu, T. Yan, and L. Yang, “A new modified wavelet-based ECG denoising,” *Comput. Assist. Surg.*, pp. 1–10, 2019.

[19] Ö. Yildirim, “A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification,” *Comput. Biol. Med.*, vol. 96, pp. 189–202, 2018.

[20] Q. Qin, J. Li, L. Zhang, Y. Yue, and C. Liu, “Combining low-dimensional wavelet features and support vector machine for arrhythmia beat classification,” *Sci. Rep.*, vol. 7, no. 1, p. 6067, 2017.

[21] M. Sharma and U. R. Acharya, “A new method to identify coronary artery disease with ECG signals and time-Frequency concentrated antisymmetric biorthogonal wavelet filter bank,” *Pattern Recognit. Lett.*, vol. 125, pp. 235–240, 2019.

[22] K. Mourad and B. R. Fethi, “Efficient automatic detection of QRS complexes in ECG signal based on reverse biorthogonal wavelet decomposition and nonlinear filtering,” *Measurement*, vol. 94, pp. 663–670, 2016.

[23] R. Jozefowicz, W. Zaremba, and I. Sutskever, “An empirical exploration of recurrent network architectures,” in *International Conference on Machine Learning*, 2015, pp. 2342–2350.

[24] Y. Su, Y. Huang, and C.-C. J. Kuo, “Dependent Bidirectional RNN with Extended-long Short-term Memory,” 2018.