Real-Time Classification for Cardiac Arrhythmia ECG Beat

Conference Paper · July 2019

DOI: 10.1109/QIR.2019.8898298

Kalamullah Ramli
University of Indonesia

Dodi Sudiana
University of Indonesia

Real-Time Classification for Cardiac Arrhythmia ECG Beat

Eko Ihsanto
Department of Electrical Engineering
Universitas Indonesia.
Depok, Indonesia
eko.ihsanto@ui.ac.id

Kalamullah Ramli
Department of Electrical Engineering
Universitas Indonesia
Depok, Indonesia
kalamullah.ramli@ui.ac.id

Dodi Sudiana
Department of Electrical Engineering
Universitas Indonesia
Depok, Indonesia
dodi.sudiana@ui.ac.id

***Abstract*—Because the ECG contains much of information regarding various heart diseases, the interpretation of an ECG is essential for monitoring heart health. Sensors, data acquisition, and pre-processing, such as filtering and denoising, can be done on a single chip. In line with that, the interpretation technique is also being developed, related to the speed of calculations and accuracy. In general, there are four stages of processing ECG information, namely, pre-processing, QRS detection, feature extraction, and classification. This paper proposes another alternative for ECG beat classification, that eliminates the pre-processing stage and combines feature extraction and classification in a single calculation stage, namely ensemble MLP. This method is expected to reduce computational costs while maintaining accuracy of 97% or more and a large number of classes, with 10 or more.**

Keywords—ECG, feature extraction, classification, MLP

I. Introduction

Among efforts to reduce mortality from heart disease, cardiac monitoring through electrocardiogram (ECG) readings is considered an efficient and effective way. Related to cardiac arrhythmia, or heartbeat abnormalities, there are two ways to analyze an ECG: analysis of a single irregular beat, or morphological arrhythmia, and analysis of multiple irregular beats or rhythmic arrhythmias. This paper reviews a variety of methods for classifying single irregular ECG beats.

Since the mid-1980s, fast and effective beat detection in an ECG has been found with detection capabilities of up to 99.3% [1], where QRS complex detection is done through calculation of slope, amplitude, and time width. This detection algorithm is implemented in a small device such as the Z80 8-bit microprocessor and can be executed in real time. Since then, ECG research has been developed into ECG beat feature extraction and classification using, for example, self-organization map (SOM) and learning vector quantization (LVQ) [2], time-based judgment rules [3], Fourier and wavelets transforms [4], multilayer perceptron (MLP) [5-7], a Kalman filter [8], and convolutional neural networks (CNNs) [9-11].

A. QRS Detection and Beat Segmentation

An ECG is a signal associated with electrical activity in the heart, and ECG beat classification must be preceded by QRS detection and beat segmentation. QRS detection is arranged in several stages [1], specifically:

- Bandpass filter
- Calculating the gradient of the R wave
- Squaring function, point by point
- Moving window integrator
- Fiducial mark measurement
- Adjusting the thresholds
- Adjusting the average RR Interval and rate limits
- T wave identification

This QRS detection process can be modified with several changes. For example, bandpass filters can be replaced with denoising wavelets or the moving window integrator can be removed and replaces the recursion calculation into a single calculation [12].

After the QRS is precisely detected, including the R position, beat segmentation can be done based on this RR-interval, for example with a simple assumption taking the starting point at ¼ of the previous RR interval to the last point at ¾ of the next RR interval. Another way to arrange beat segmentation is by using the hidden Markov model [13].

B. Feature Extraction

Every single beat of results from beat segmentation can be analyzed based on its features, i.e., points P, Q, R, S, T, QRS complex, and RR-intervals. Extraction of these time-based features can be done during QRS detection and beat segmentation, including R-peak, RR-interval, and QRS Complex [14, 15]. Feature extraction can also be done at special stages apart from QRS detection and Classification, for example by calculation using the Teager Energy Operator [16], Discrete Wavelet Transform [17, 18], or other methods. In addition to being combined with the stages of QRS detection and beat segmentation, feature extraction can also be done simultaneously with the classification process, which is carried out during training in the MLP model [19] or CNN [9].

C. Multilayer Perceptron (MLP) Classifier

Differing from support vector machines (SVMs), the MLP classifier is more flexible. It can be used separately from or integrated with the features extraction process. MLP and back propagation have been successfully applied to the classification of an arrhythmia beat ECG with an accuracy of 98.19% [19].

To get better accuracy, the classification process can be realized using ensemble MLP [20]. Every MLP with the same structure, trained with the same or different data, according to its ECG beat class. After all MLP parameters have been optimized, the MLPs are combined by choosing the largest MLP output. Regarding efforts to overcome high-dimensional manifolds on ECG beats, the MLP can be preceded by a local regionalization process. Local regionalization can focus on the particularity of local variations, while deep MLP can break down complexity in the local distribution environment. This two-step combination of local regionalization and deep MLP is called a local deep field [21]

II. PROPOSED METHODS

This paper proposes ECG beat classification methods that combine features of the extraction stage with features of the classification stage. Because the data used are taken from the MIT-BIH Arrhythmia Database, which has been preprocessed by filtering or denoising, preprocessing is not needed. Therefore, only two stages are needed, the QRS detection stage and the classification stage. Beat segmentation is done during the QRS detection stage, while feature extraction is done simultaneously with classification. There are at least two options for a QRS detection method, Pan-Tompkins [1] and Faezipour et al. [12], while for the integrated feature extractor and classifier, ensemble MLP can be applied.

A. *QRS Detection and Beat segmentation*

The QRS detection method we recommend is the method proposed by Pan-Tompkins, which has been applied to the WFDB API [22] and can be used with Matlab, C ++, or Python programming. If the QRS complex has been detected, the position of the R peak is also known. Based on the RR interval, beat segmentation can be done easily by, for example, taking many sample points starting from ¼ of the previous RR interval to ¾ of the next RR interval. If there is no worry corrupting the ECG feature, simple downsampling can also be done, without extrapolation using Fourier, wavelet, or Stockwell transformation.

In this experiment, data are extracted from 48 files containing 112,647 beats that have been labeled by experts. Data were taken from the MIT-BIH database [23] using the WFDB library [22].

TABLE I. DATA FOR EXPERIMENTS EXTRACTED FROM THE MIT-BIH DATABASE

No	Symbol	Annotation Description	Total	Training	Test
1	N	Normal Beat	75,052	11,257	63,795
2	L	Left bundle branch block beat	8,075	2,826	5,249
3	R	Right bundle branch block beat	7,259	2,540	4,719
8	A	Atrial premature contraction	2,546	891	1,655
5	V	Premature ventricular contraction	7,130	2,495	4,635
12	/	Paced beat	7,028	2,459	4,569
4	a	Aberrated atrial premature beat	150	75	75
31	!	Ventricular flutter wave	472	236	236
6	F	Fusion of ventricular and normal beat	803	401	402
37	x	Non-conducted P-wave (blocked APB)	193	96	97
		Total	108,708	23,276	85,432
		Total Beats number :	112,647		

As seen in Table I, out of 112,647 beats, only 10 types of beats were taken, totaling 108,708 beats. Data are divided into two parts, the part for training and the part for testing or validation.

B. *Beat Classification Using MLP*

The beats obtained from the QRS detection and beat segmentation stage can be classified directly using an MLP classifier. MLP can be implemented by a single model or by multiple models as an ensemble with a particular function, such as maximum, average, or other arithmetic functions. If multiple models are used with the same or different structures, the final part of all of the models must have the same shape at the output layer, usually the softmax fully connected layer.

Regarding input, if the amount of training data are considered sufficient, mini batching can be carried out, with sizes 16, 32, 64, or others. If the data for training are considered too little, bagging and k-fold cross-validation can be implemented.

As seen in Fig. 1, an ensemble MLP classifier consists of one input layer, one output layer, and six fully connected layers. This figure is generated by executing a program written in Python. This program uses many libraries, including Keras [24, 25] and TensorFlow [26]. There are two MLP models combined in this classifier, both of which have a similar structure and are trained with the same data. There are 1,740 total parameters of this MLP classifier, as shown in Table II.

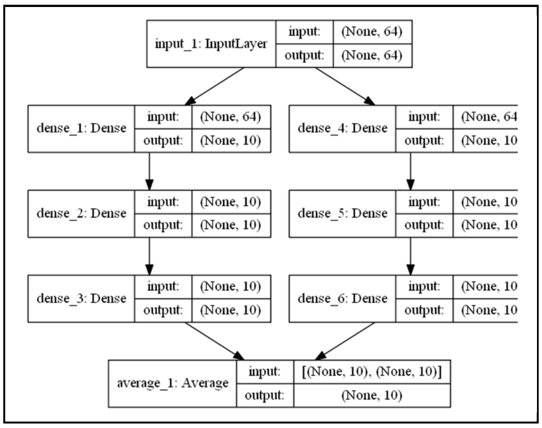

Fig. 1. The Ensemble MLP Classifier

Aside from being a classifier, the ensemble MLP shown in Fig. 1 also acts as a feature extractor. In this experiment, all beats originating from MIT-BIH were immediately processed in this ensemble MLP without needing any stages for transformation or dimension reduction.

TABLE II. LAYERS AND NUMBER OF PARAMETERS OF ENSEMBLE MLP

Layer (type)	Param #
input_1 (InputLayer)	0
dense_1 (Dense)	650
dense_4 (Dense)	650
dense_2 (Dense)	110
dense_5 (Dense)	110
dense_3 (Dense)	110
dense_6 (Dense)	110
average_1 (Average)	0
Total params:	1,740

III. PRELIMINARY RESULTS AND DISCUSSION

In this section, the preliminary results of the ECG beat classification experiment using an ensemble MLP will be presented. The training and validation history from the ensemble MLP can be seen in Fig. 2. The confusion matrix results of classification can be seen in Table III, while the performance assessment table from each class can be seen in Table IV. Images of the training and validation history show accuracy throughout the iteration. The confusion matrix table shows a comparison between ground truth and prediction results of the classification, while the performance assessment table shows sensitivity (Se) and accuracy (Acc) of the validation results

From Fig. 2, it can be seen that for both MLP models, accuracy at validation was always better than accuracy during training. This means there was no overfitting, and there are still opportunities to increase accuracy by modifying the model structure or parameters.

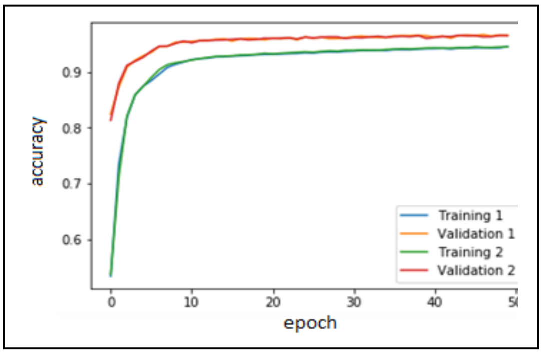

Fig. 2. Training and Validation History of the MLP Model

As seen in Tables III and IV, the ECG beats classification result was terrible for the seventh class, or aberrated atrial premature beat class. The accuracy of the seventh-class prediction result is zero even though the overall accuracy is 97.08%. This is possibly because there was insufficient data for training or because the seventh beat morphology is too difficult to be learned by this classifier, and it needs modification in the model.

TABLE III. CONFUSION MATRIX OF ENSEMBLE MLP

Ground Truth		Prediction											
		N	L	R	A	V	/	a	!	F	x	Σ	Acc
		0	1	2	3	4	5	6	7	8	9		
N	0	63,100	163	127	73	193	23	0	17	85	14	63,795	98.91
L	1	53	5,127	2	5	52	3	0	3	4	0	5,249	97.68
R	2	64	6	4,500	148	0	1	0	0	0	0	4,719	95.36
A	3	528	4	74	1,022	25	2	0	0	0	0	1,655	61.75
V	4	204	99	26	21	4,198	16	0	46	22	3	4,635	90.57
/	5	8	1	1	0	20	4,528	0	11	0	0	4,569	99.10
a	6	41	2	1	3	8	12	0	8	0	0	75	0.00
!	7	45	6	1	0	39	5	0	137	0	3	236	58.05
F	8	89	7	0	1	52	0	0	1	252	0	402	62.69
x	9	18	0	0	0	0	5	0	3	0	71	97	73.20
	Σ	64,150	5,415	4,732	1,273	4,587	4,595	0	226	363	91		97.08

Table IV shows the performance assessment of the ensemble MLP. Calculation of sensitivity (Se) and accuracy (Acc) were as follows: The three variables of FN, TP, and FP, which stand for, respectively, false negative, true positive, and false positive, were taken from the confusion matrix, while Se and Acc were calculated using the following formulas:

$$Se = TP/(TP + FN)$$
$$Acc = TP/(TP + FP)$$

TABLE IV. PERFORMANCE ASSESSMENT OF ENSEMBLE MLP

ECG class	Total Beats	Trained Beats	Test Beats	FN	TP	FP	*Se (%)*	*Acc (%)*
N	75,052	11,257	63,795	1,050	63,100	695	98.36	98.91
L	8,075	2,826	5,249	288	5,127	122	94.68	97.68
R	7,259	2,540	4,719	232	4,500	219	95.10	95.36
A	2,546	891	1,655	251	1,022	633	80.28	61.75
V	7,130	2,495	4,635	389	4,198	437	91.52	90.57
/	7,028	2,459	4,569	67	4,528	41	98.54	99.10
a	150	75	75	0	0	75	0.00	0.00
!	472	236	236	89	137	99	60.62	58.05
F	803	401	402	111	252	150	69.42	62.69
x	193	96	97	20	71	26	78.02	73.20
Σ	108,708	23,276	85,432	2,497	82,935	2,497	97.08	97.08

Even though it is only a preliminary result, the proposed method seems to outperform when compared to other methods, as shown in Table V. Among the 14 methods recapitulated in Table V, 10 of them can be compared because they use the same data for experiments, both for training and for testing. Of these 10 methods, only three have low computational costs because they combine the stages of feature extraction and classification, and therefore, these three methods are more affordable to implement for real-time ECG beats classification. These methods are written in rows 9, 12, and 14 in Table V. Among the three methods, the proposed method, which is at row 14, has the best classification ability in terms of the number of classes, with 10 classes. In the next modification, it is necessary to attempt to reach 16 classes to equal the methods on rows 6 and 11, while still achieving high levels of accuracy.

IV. CONCLUSION

As described in this paper, there are various methods for classifying ECG beats related to cardiac arrhythmia. In general, there are four stages to classifying the ECG beat, namely preprocessing, QRS detection, feature extraction, and classification. Preprocessing is usually in the form of filtering and denoising so that data are efficiently processed in the next stages. At this stage, there are various transformations and thresholding methods that can be used.

After obtaining beats from the segmentation process, the features extraction process can be done in various ways; for example, by applying a Fourier or wavelet transform. Dimension reduction can be realized by PCA. Feature extraction and scaling dimensions can be done using the autoencoding method. The autoencoding process is done to obtain the best feature descriptor representing the diversity of ECG beats to be classified.

TABLE V. VARIOUS CLASSIFICATION METHODS AND ACCURACY

No	Author	# class	Methods	Train data	Test data	Accuracy
1	Dokur and Olmez [27]	10	MLP, Wavelet, and GA	MIT-BIH	Real-time measurement	96.00%
2	Ozbay et.al. [28]	10	Fuzzy Clustering NN (MLP)	MIT-BIH	Cardiology Department Seljuk University Turkey	98.00% - 99.90%
3	Ceylan and Ozbay [29]	10	FCM, PCA, WT, and ANN (MLP)	MIT-BIH	Cardiology Department Seljuk University Turkey	99.00%
4	Melgani and Bazi [30]	10	SVM and PSO	MIT-BIH	MIT-BIH	89.72%
5	Ince et.al [31]	5	DWT, PCA, and ANN (MLP)	MIT-BIH	MIT-BIH	93.63%
6	Wen et.al. [32]	16	SOCMAC self-organizing cerebellar model articulation controller	MIT-BIH	MIT-BIH	98.21%
7	Ozbay and Tezel [19]	10	MLP with adaptive activation function	MIT-BIH	Cardiology Department Seljuk University Turkey	98.19%
8	Sarfraz et.al [33]	8	ICA and MLP	MIT-BIH	MIT-BIH	99.61%
9	Kiranyaz et.al [9]	5	1D-CNN	MIT-BIH	MIT-BIH	95.14%
10	Nanjun and Meshram [34]	2	DWT and Deep NN (MLP)	MIT-BIH	MIT-BIH	98.33%
11	Raj, Ray [35]	16	DCT, DOST, PCA, SVM, and PSO	MIT-BIH	MIT-BIH	98.82%
12	Zhai and Tin [10]	5	2D-CNN	MIT-BIH	MIT-BIH	96.05%
13	Rangappa and Agarwal [36]	2	k-NN	MIT-BIH	MIT-BIH	98.40%
14	Proposed	10	Ensemble MLP	MIT-BIH	MIT-BIH	97.08%

Features extraction is usually done before classification. In some cases, the feature dimensions are also reduced to facilitate kernel calculations. However, for MLP, the feature extraction and classification processes can be integrated into a single neural network, either single model or ensemble models. Care should be taken because although ensemble models can increase accuracy, they can also increase the number of parameters that affect the computational cost.

From the results of comparisons with various methods, it can be concluded that the proposed method has quite good performance. By paying attention to the quality of the classifier with regard to the three issues of computational cost, number of classes, and accuracy, the proposed method can be considered to outperform 10 other methods. Nevertheless, further modifications still need to be realized to increase the number of classes up to 16 or more, while still achieving high accuracy and low computational cost.

ACKNOWLEDGMENT

This research is partly funding by PDD Grant of Ristekdikti and Universitas Indonesia under contract number: No: NKB-1845/UN2.R3.1/HKP.05.00/2019.

The publication of this article is partly supported by the United States Agency for International Development (USAID) through the Sustainable Higher Education Research Alliance (SHERA) Program for the Universitas Indonesia Scientific Modeling, Application, Research, and Training for City-centered Innovation and Technology (SMART CITY) Project, Grant #AID-497-A-1600004,
Sub-grant #IIE-00000078-UI-1

REFERENCES

[1] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng,* vol. 32, pp. 230-236, 1985.

[2] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE transactions on biomedical engineering,* vol. 44, pp. 891-900, 1997.

[3] X. Fan, R. Chen, C. He, Y. Cai, P. Wang, and Y. Li, "Toward automated analysis of electrocardiogram big data by graphics processing unit for mobile health application," *IEEE Access,* vol. 5, pp. 17136-17148, 2017.

[4] Z. Dokur, T. Olmez, and E. Yazgan, "Comparison of discrete wavelet and Fourier transforms for ECG beat classification," *Electronics Letters,* vol. 35, pp. 1502-1504, 1999.

[5] H. Yan, Y. Jiang, J. Zheng, C. Peng, and Q. Li, "A multilayer perceptron-based medical decision support system for heart disease diagnosis," *Expert Systems with Applications,* vol. 30, pp. 272-281, 2006.

[6] K. Hanbay, "Deep neural network based approach for ECG classification using hybrid differential features and active learning," *IET Signal Processing,* 2018.

[7] E. D. Übeyli, "Support vector machines for detection of electrocardiographic changes in partial epileptic patients," *Engineering Applications of Artificial Intelligence,* vol. 21, pp. 1196-1203, 2008.

[8] J. Oster, J. Behar, O. Sayadi, S. Nemati, A. E. Johnson, and G. D. Clifford, "Semisupervised ECG ventricular beat classification with novelty detection based on switching Kalman filters," *IEEE Transactions on Biomedical Engineering,* vol. 62, pp. 2125-2134, 2015.

[9] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering,* vol. 63, pp. 664-675, 2015.

[10] X. Zhai and C. Tin, "Automated ECG classification using dual heartbeat coupling based on convolutional neural network," *IEEE Access,* vol. 6, pp. 27465-27472, 2018.

[11] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of EEG signals," *IEEE Access,* vol. 6, pp. 25399-25410, 2018.

[12] M. Faezipour, A. Saeed, S. C. Bulusu, M. Nourani, H. Minn, and L. Tamil, "A patient-adaptive profiling scheme for ECG beat classification," *IEEE transactions on information technology in biomedicine,* vol. 14, pp. 1153-1165, 2010.

[13] R. V. Andreao, B. Dorizzi, and J. Boudy, "ECG signal analysis through hidden Markov models," *IEEE Transactions on Biomedical engineering,* vol. 53, pp. 1541-1549, 2006.

[14] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ECG morphology and heartbeat interval features," *IEEE transactions on biomedical engineering,* vol. 51, pp. 1196-1206, 2004.

[15] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE transactions on biomedical engineering,* vol. 53, pp. 2535-2543, 2006.

[16] C. Kamath, "ECG beat classification using features extracted from Teager energy functions in time and frequency domains," *IET signal processing,* vol. 5, pp. 575-581, 2011.

[17] O. T. Inan, L. Giovangrandi, and G. T. Kovacs, "Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features," *IEEE transactions on Biomedical Engineering,* vol. 53, pp. 2507-2515, 2006.

[18] C. Venkatesan, P. Karthigaikumar, A. Paul, S. Satheeskumaran, and R. Kumar, "ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications," *IEEE Access,* vol. 6, pp. 9767-9773, 2018.

[19] Y. Özbay and G. Tezel, "A new method for classification of ECG arrhythmias using neural network with adaptive activation function," *Digital Signal Processing,* vol. 20, pp. 1040-1049, 2010.

[20] J. Zheng, Y. Jiang, and H. Yan, "Committee machines with ensembles of multilayer perceptron for the support of diagnosis of heart diseases," in *2006 International Conference on Communications, Circuits and Systems*, 2006, pp. 2046-2050.
[21] W. Li and J. Li, "Local deep field for electrocardiogram beat classification," *IEEE Sensors Journal,* vol. 18, pp. 1656-1664, 2017.
[22] I. Silva and G. B. Moody, "An open-source toolbox for analysing and processing physionet databases in matlab and octave," *Journal of open research software,* vol. 2, 2014.
[23] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine,* vol. 20, pp. 45-50, 2001.
[24] F. Chollet, "Keras: Deep learning library for theano and tensorflow," *URL: https://keras. io/k,* vol. 7, p. T1, 2015.
[25] F. Chollet, *Deep Learning with Python*. Shelter Island, NY 11964: Manning Publications, 2018.
[26] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean*, et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265-283.
[27] Z. Dokur and T. Ölmez, "ECG beat classification by a novel hybrid neural network," *Computer methods and programs in biomedicine,* vol. 66, pp. 167-181, 2001.
[28] Y. Özbay, R. Ceylan, and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine,* vol. 36, pp. 376-388, 2006.
[29] R. Ceylan and Y. Özbay, "Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network," *Expert Systems with Applications,* vol. 33, pp. 286-295, 2007.
[30] F. Melgani and Y. Bazi, "Classification of electrocardiogram signals with support vector machines and particle swarm optimization," *IEEE transactions on information technology in biomedicine,* vol. 12, pp. 667-677, 2008.
[31] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Transactions on Biomedical Engineering,* vol. 56, pp. 1415-1426, 2009.
[32] C. Wen, T.-C. Lin, K.-C. Chang, and C.-H. Huang, "Classification of ECG complexes using self-organizing CMAC," *Measurement,* vol. 42, pp. 399-407, 2009.
[33] M. Sarfraz, A. A. Khan, and F. F. Li, "Using independent component analysis to obtain feature space for reliable ECG Arrhythmia classification," in *2014 IEEE international conference on bioinformatics and biomedicine (BIBM)*, 2014, pp. 62-67.
[34] R. Nanjundegowda and V. A. Meshram, "Arrhythmia Detection Based on Hybrid Features of T-wave in Electrocardiogram," *International Journal of Intelligent Engineering & Systems,* vol. 11, pp. 153-162, 2018.
[35] S. Raj and K. C. Ray, "ECG signal analysis using DCT-based DOST and PSO optimized SVM," *IEEE Transactions on instrumentation and measurement,* vol. 66, pp. 470-478, 2017.
[36] V. G. Rangappa, S. V. A. V. Prasad, and A. Agarwal, "Classification of Cardiac Arrhythmia stages using Hybrid Features Extraction with K-Nearest Neighbour classifier of ECG Signals," *learning,* vol. 22, 2018.