

Fusion based Feature Extraction Analysis of ECG Signal Interpretation – A Systematic Approach

Dr. T. Vijayakumar,

Professor,

Department of ECE,

GNIT, Hyderabad, India.

Email id: vishal_16278@yahoo.co.in

Mr. R. Vinothkanna,

Department of ECE, Vivekanandha College of Technology for Women,

Namakal, India.

Email: rvinothkannaphd@gmail.com

Dr. M. Duraipandian,

Head of Department, Department of Computer Science and Engineering,

Vivekanandha College of Technology for Women,

Namakal, India.

Email: svsduraipandian@gmail.com

Abstract- Our human heart is classified into four sections called the left side and right side of the atrium and ventricle accordingly. Monitoring and taking care of the heart of every human is the very essential part. Therefore, the early prediction is essential to save and give awareness to humans about diet plan, lifestyle schedule. Also, this is used to improve the clinical diagnosis and treatment of any patients. To predict or identifying any cardiovascular problems, Electro Cardio Gram (ECG) is used to record the electrical signal of the heart from the body surface of humans. The algorithm learns the dataset from before cluster is called supervised; The algorithm learns to train the data from the set of a dataset is called unsupervised. Then the classification of more amount of heartbeat for different category of normal, abnormal, irregular heartbeats to detect cardiovascular diseases. In this research article, a comparison of various methods to classify the dataset with a fusion-based feature extraction method. Besides, our research work consists of a denoising filter to reconstruct the raw data from the original input. Our proposed framework performing preprocessing that consists of a filtering approach to remove noises from the raw data

set. The signal is affected by thermal noise and instrumentation noise, calibration noise due to power line fluctuation. This interference is high in many handheld devices which can be eliminated by de-noising filters. The output of the de-noising filter is input for fusion-based feature extraction and prediction model construction. This workflow progress has given good results of classifier effectiveness and imbalance arrangement conditions. We achieved good accuracy 96.5% and minimum computation time for classification of ECG signal.

Keywords: *ECG classification, Random Forest, Feature Extraction*

1. INTRODUCTION

In the world, 30% of death is due to cardiovascular problems disease according to World Health Organization (WHO). The disease mechanism is very crucial during the death of any human in dangerous circumstances. Therefore, any strange activity or abnormal condition of the heart routine process can be determined by the radiologist to predict the heart disease [1]. There are lots of ECG formats to understand the normal and abnormal state of the heartbeat in the body. The probe leads fix temporarily in the body and record reading from the heartbeat in the body surface [2]. Based on the number of leads the information's are vast that both of them is depending on each other. Recently this ECG recording includes standard 12 lead probe's that processing by deep learning technique to improve the prediction rate with high accuracy. This large amount of probe leads provides more information about heartbeat for getting higher accuracy [3]. Deal this amount of data, we need to go for a powerful tool to compute that information. The ECG formats are needed powerful computational methods to extract information from the complete dataset. This research article is focusing to describe the machine learning method to analyze the ECG dataset for the diagnosis of all cases of the patient. We are trying to extend our proposed framework for this analysis. This analysis method is used to learn and train networks to make an accurate prediction of diagnosis [4].

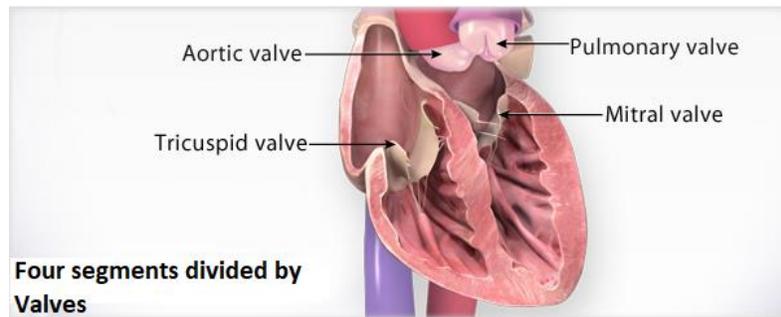


Figure 1 Cross sectional view of human heart picture

There are two type of learning methods as follows;

1. Supervised
2. Unsupervised

The machine learning methods are a capable approach to deal with medical ECG signals with various directions of simulation results. Besides, the lack of a database is one of the great challenges in the medical field to solve the disease interpretation issues [5]. For training and validation process of the dataset is interpreting the ECG signal and it depends on the generation of synthetic data. This simulation results have several points and task which includes the deal with noise in the signal for QRS detection, P waves, and T waves. The noise reduction can be done in this article for getting clear signals for training and learning further [6]. Many transform and filter techniques are carried out to deal with this noise reduction in a signal. The wavelet transform method increases a good prediction and detection rate over 99%. To the analysis of ECG signals, the position and shape of the signal are very important components [7]. The advanced method of Kalman filtering for detection is reaching good accuracy results in ECG signal processing. The classification of ECG signal ensembles many class testing and training process in pattern recognition [8]. The simplified block diagram of the proposed architecture framework is showing in figure 2.

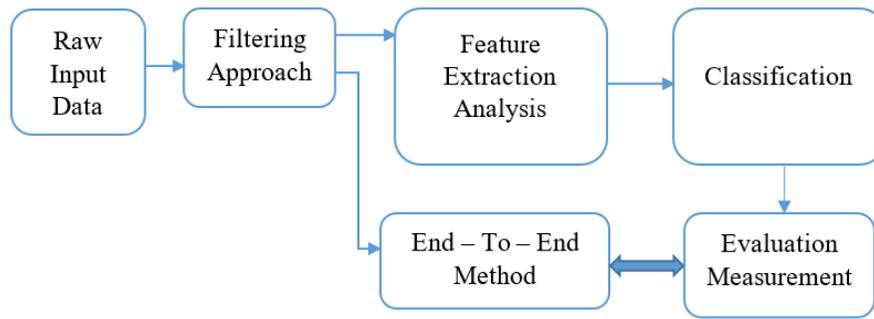


Figure 2 Block diagram of overview of proposed architecture construction

The changes in shape and position of ECG signal are very tough to analyze with a simple and single algorithm. The threshold-based analysis is suffering from slight changes of ECG recording that the new person's recorded signal delay or early start of the heartbeat signal [9]. Also if changes of new devices provide different threshold values based on the system specification. Therefore, the data-driven or adaptive method is essential to handle this analysis with a hybrid type of machine learning approach. The right side of the heart is containing a tricuspid valve used to control the motion of the blood from the atrium to the ventricles which are showing in figure 1. ECG is the process of recording the electrical signal from the heart based on the body surface to monitor the activities [10]. The electrodes are called probe lead fixing in the body near to the heart to get the information of heartbeat. This fixing of probe electrodes is near to atrioventricular and ventricle nodes. The healthy heart has a specific order of division and intervals between each heartbeat. If there are any deviations in the heartbeat will be noted as an abnormal condition [11]. The normal condition of the heartbeat is showing in figure 3.

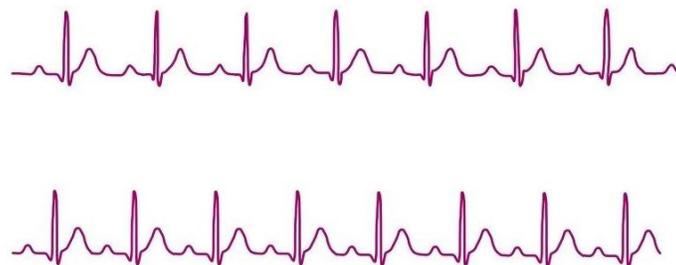


Figure 3 Normal condition of heartbeat

The figure shows the voltage and time sequence of a signal produced in ECG. The sequence of ECG indicates our heartbeat health and its conditions. The periodic PQRST signal is indicated in the heartbeat activities graph. An among these PQRST signals consist of; the P indicates atrial contraction, QRS indicates a contraction of ventricles, T indicates an expansion of the ventricle action respectively [12]. This wave signal is depending on the heartbeat rate. The typical heartbeat points are showing in figure 4.

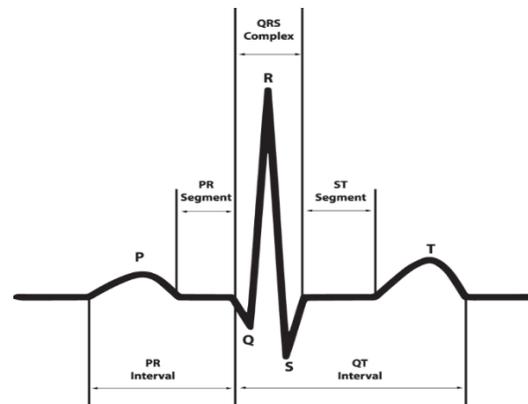


Figure 4 Typical ECG complexes

The interval measures method is measuring the distance between each complex. Also, the irregular and regular rhythm is classifying based on heartbeat rate. The heartbeat rate is in between 60 to 90 bpm is called the normal heartbeat rate [13]. The many ECG complexes classification problems are solving by the hybrid computational algorithm with neural network solution. Because of its strong robustness, the prediction rate is high. Also, this type of solution can solve problems of the complex nonlinear structure easily [14].

2. ORGANIZATION OF THE RESEARCH

The structure of the research article is organized as follows; Section 3 gives related work of ECG classification based on a machine learning algorithm. Section 4 discusses proposed methodologies for ECG classification. Section 5 consists of a result discussion from the obtained output. The conclusion and the future task delivers in section 6.

3. PRELIMINARIES

Pinho Andre et al propose artificial intelligence analysis for ECG signal classification. They tested with SVM of a different set of extracted features. They compared with artificial Intelligence Neural Network (ANN) and proven the performance is better than SVM [15]. Moavenian et al introduces Multi-Layered Perceptron (MLP) and compared it with SVM and proving higher performance in the speed process. But the accuracy of the process is slower than the existing method [16]. Zhang et al propose SVM with the genetic method is performing for the classification and achieved 93% true positive with moderate data set around 750 ECG signal record [17]. Chauhan et al tested a deep LSTM network with the unknown size of the dataset for signal classification. The measuring metric is calculated as 97.5% precision and 96.45% F-score. Leutheuser et al classifies the abnormal and normal condition of heartbeat with the MIT-BIH label method and obtained 89.6% accuracy which is validated on the dependent dataset [18]. Oresko et al compute the normal and abnormal condition detection of the heartbeat by many morphological filter patterns and get 99% accuracy for the normal method [19]. J.Wang et al introduce the RR interval feature normalization method to finding the distance between the PQRST complexes. The measuring metric is computed for various classification methods. They achieved good accuracy, specificity with a moderate amount of dataset [20]. A.Dallali et al review the RR interval feature normalization with many metric measures. They introduce the removing noise in the raw electrical signal from the ECG device. They obtained high accuracy and precision [21]. M.Vijayavanan et al calculate intervals between PQRST duration segments from signal feature extraction. They trained MIT-BIH arrhythmias for forwarding classifier technique and tested with high accuracy [22]. V.Srivastava et al implement with same arrhythmias feed-forward transformation method incorporated with fuzzy logic method provide good specificity rate [23]. M.Korurek et al propose 5R peak with noise removing algorithm of linear phase filtering techniques. They constructed the framework with baseline correction for signal extraction and its classification; Therefore, they achieved good higher sensitive and specificity rate [24].

4. METHODOLOGIES

4.1 Filtering Approach

Usually, ECG signal is affected by noise due to various sources. The variation of the signal is depending on the changes in the position of electrodes. These small changes in the signal create shifting the signal from the baseline which interprets the wrong diagnosis. These small signal changes will not be in the specified range of frequencies. Therefore, elimination is a big challenge here. The construction of the bandpass filtering approach is performed here to eliminate the noise from the signal. The proposed model is based on many features that are computing for the prediction level. To reduce the noise in the data, the time domain co-efficient has been taken to get a higher accuracy level of the data set [25]. The model has developed with many features based on fusion technique is showing a good prediction rate. The analysis of noisy datasets is performing with feature coefficients for many decomposition iterations of bandpass filtering. To achieve higher accuracy, rate the bandpass filtering techniques is one of the great choices for clean data classification. The heartbeat variability measure also considering in the construction of the framework. The filter decomposition level is variable based on the raw data input ECG signal variation. It is an adaptive function approach transformation can provide better visibility after denoising the signal. The interval is computed includes time-domain components and frequency domain components for good visibility structure. This all processing flow is showing in figure 5.

4.2 Fusion based Feature Selection

In this research article, feature selection has been included for detection. Mainly QRS detection is a challenging task in this extraction process. Figure 6 shows the feature extraction methods in our proposed architecture. To find R peaks in the electrical heartbeat signal of ECG, our proposed algorithm is appropriate for a higher accuracy detection rate. The main standard algorithm is the Pan-Tomkins algorithm which is used to detect QRS complexes [26]. This heartbeat detection should be near to the ideal perfection of our proposed algorithm. Our research work concentrates on filtering processing for raw data of input ECG signal in our algorithm. The data is filtered with a range of frequencies is mentioned as equation below,

$$W_{bp}(t) = \frac{1}{\sqrt{2\pi a}} \exp\left(-\frac{t^2}{2a^2}\right) \{\exp[j2\pi f_L t] + \dots + \exp[j2\pi(f_L + k\Delta f)t] + \dots + \exp[j2\pi f_H t]\}$$

Where $k=0,1,2,\dots,N$ and $f_H = f_L + N\Delta f$

This can help to obtain information about the slope of the QRS. The classification method here is used to classify the data based on the training set and class labels in the proposed structure and it helps to predicts the class label also. Finally, our proposed system construct with a prediction segment to predict the structured models with continuous-valued function [27]. Therefore, the missing or unlabeled values can be re-calculated for medical diagnosis.

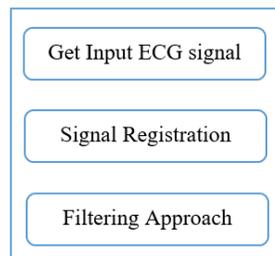


Figure 5 Preprocessing Segment of Proposed Work

Besides, to put a boundary condition for minima and maxima for ECG signal to feature extraction vector. The feature extraction computes the R peaks appearing in the length of the signal sequences. The standard deviation of the R peak value is directly proportional to obtained and predicted R-Peak values with a minus sign [28]. Based on our approach the heartbeat finds the best representation of heartbeat from the input signal. Our filtering approach is also more resistant to the noise signal.

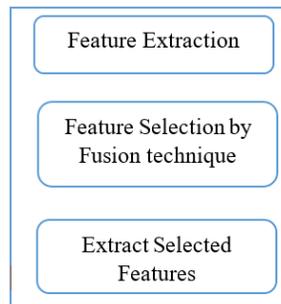


Figure 6 Feature fusion technique of Proposed Work

This feature selection is one of the essential parts of our proposed architecture to select the subset from the raw feature set. This selection procedure can be efficient and given data makes weak cross-correlation function among all the features. Therefore, the categorizing of feature selection made it easier to find appropriate standards algorithm. The evaluation measure metric gives better results compared to the existing method. Besides, our proposed architecture is reducing overfitting problems in a model generation to improve our algorithm. This filtering approach, feature, and fusion selection technique is used to provide better metric values compared to another existing method [29]. Figure 7 is showing the final stage of our proposed architecture.

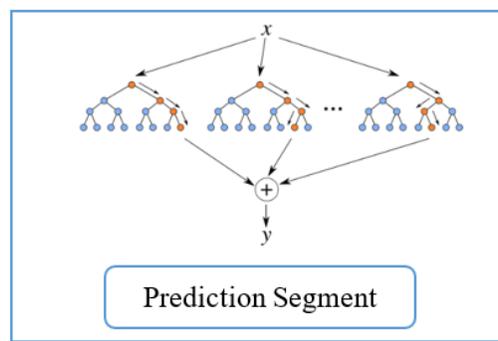


Figure 7 Final stage of Our Proposed Framework

The Random Forest algorithm is incorporated with our proposed structure based on decision trees. This approach will be appropriated for our structure model to train the data very fast. This is a very fast computation algorithm consist of more than 60 decision trees which is shown in figure 2. Also, this approach classification process is used to construct the framework with an optimized structure. Our proposed architecture is constructed based on the decision tree method. This supervised learning algorithm is used to make an ensemble process for regression [30]. This technique is used to combine the classification and predictions from the multiple input machine learning algorithm for making accurate predictions by any single model. Feature vector transformation is showing in figure 8.

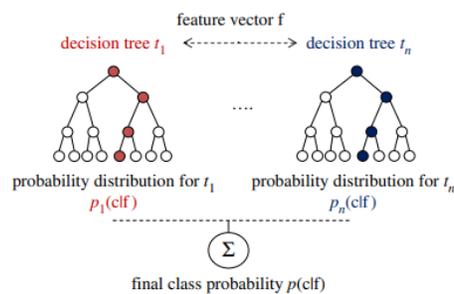


Figure 8 Decision Tree at Proposed Algorithm

5. RESULTS DISCUSSION

Our proposed model compares with 350 samples to analyze for training and testing the dataset. The measurement of class balance, classifier completeness, decision optimization, and model estimation. Our proposed research work is splitting the data into train and test for minimizing the overlap problem in classes. This segmentation approach provides highly interpretable and classes are balanced with high accuracy [31]. The formula for measurement metrics is given below as an equation description for our proposed model. Figure 9 shows the distance calculation method at each interval of ECG signal for classification.

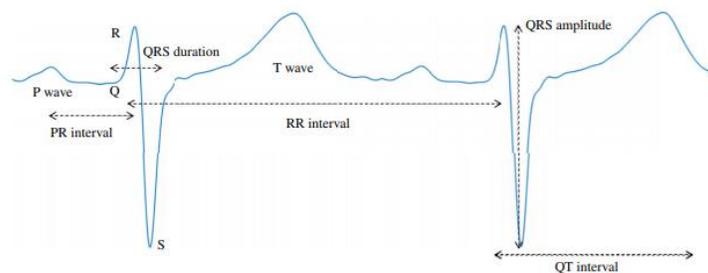


Figure 9 Distance calculation at interval of ECG Signal

Our framework procedure is an independent and data classification from 5 lead probe sensor output as raw data. This structure performs on two decision trees with the help of the Random Forest method to fuse the final classification decision. We achieved higher accuracy, sensitivity, and minimum execution time of our algorithm. Except for computation time other metric parameters are a bit lower than the hybrid version of the learning algorithm. But the

application of the proposed method is directly the medical field diagnosis sector [32][33]. Therefore, this sector should be provided the result with time constraints. The overall performance of the model is improved and compared to the existing method for optimization purposes. These measuring metrics are performing to find the effectiveness of class with the help of labels. The result is calculated and tabulated in Table 1 below.

Table 1 Comparison of Our techniques

S.NO	TASK	MODEL PROPOSED	SIGNAL SAMPLES	PERFORMANCE MEASURE
1	Myocardial infarction detection	Neural Network	375 Samples	Accuracy: 93.5% Sensitivity:91.3% Specificity:90.83%
2	Classification of ECG Signal	Our Fusion technique Architecture	350 Samples	Accuracy:96.5% Sensitivity:94.35% Specificity:97.2%
3	AF detection	CNN & RNN	325 Samples	Accuracy:98% Sensitivity:97.3% Specificity:98.7%

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FP}$$

$$Specificity = \frac{TN}{TN + FP}$$

The frequency power calculation of our proposed architecture is based on an estimation of power spectrum performance. The interpolation of intervals between in the ECG signal is calculated by integration of composite rule-based method. Our proposed architecture is very fast as well as computation time is very less compared to the hybrid learning method. Our proposed method is good in the overall performance chart with execution time. Because the medical field needs to record results immediately. Therefore, the time constraint is very important. As per the perception of point of view, our proposed architecture is fulfilling those tasks very successfully compared to the hybrid version of the learning method. The comparison of performance metric measures is showing in figure 10.

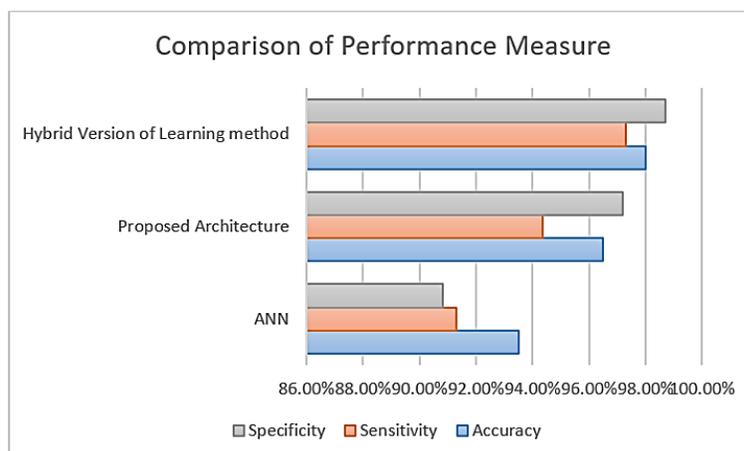


Figure 10 Comparison chart of Performance measure

6. CONCLUSION

Our proposed framework is concentrating on the dataset class imbalance problem which arises in many algorithms during classification. The random forest is a more efficient method to classify class-appropriate balancing segments. Our research article is showing ECG signal analysis with an RF machine learning algorithm. This approach has processed including imbalanced class and overfitting problem with suitable decision tree approach. Our research article is detected QRS complexes from the input ECG signal to verify the specified features for classification. The authors believe strongly this research article can help for a better understanding of medical field challenges especially cardiovascular research. Besides the machine learning approach has given better clinical advancement to the analysis of ECG signals.

Future work

The proposed architecture is used to help upgrade the automatic disease diagnosis monitoring and predicting based on ECG recording further. The manual handling of the ECG recording segment is a very large computation and more time expensive. One of the problems of our proposed work occupies a huge amount of storage to record ECG signal and its processing also. The higher number of data gives more accuracy leads to high dimension data storage. The classifying of the huge amount of irrelevant data or redundant data set is a challenging task in our

proposed algorithm. We can be achieved better accuracy and computation time result by removing those features and correlated vector function features.

REFERENCES

- [1] Emelia, J.; Benjamin, S.S.V.; Clifton, W.; Callaway, A.M.C.; Alexander, R.; Chang, S.C.; Stephanie, E.; Chiuve, M.C.; Francesca, N.; Delling, R.D.; et al. Heart Disease and Stroke Statistics—2018 Update: A Report from the American Heart Association. *Circulation* 2018, 137, e67–e492.
- [2] Hannun, A.Y.; Rajpurkar, P.; Haghpanahi, M.; Tison, G.H.; Bourn, C.; Turakhia, M.P.; Ng, A.Y. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat. Med.* 2019, 25, 65–69. [CrossRef]
- [3]. Yang, C.; Veiga, C.; Rodriguez-Andina, J.J.; Farina, J.; Iniguez, A.; Yin, S. Using PPG Signals and Wearable Devices for Atrial Fibrillation Screening. *IEEE Trans. Ind. Electron.* 2019, 66, 8832–8842. [CrossRef]
- [4]. Lyon, A.; Mincholé, A.; Martínez, J.P.; Laguna, P.; Rodriguez, B. Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances. *J. R. Soc. Interface* 2018, 15, 20170821. [CrossRef]
- [5]. Singh, B.N.; Tiwari, A.K. Optimal selection of wavelet basis function applied to ECG signal denoising. *Digit. Signal Process.* 2006, 16, 275–287. [CrossRef]
- [6]. Friesen, G.M.; Jannett, T.C.; Jadallah, M.A.; Yates, S.L.; Quint, S.R.; Nagle, H.T. A comparison of the noise sensitivity of nine QRS detection algorithms. *IEEE Trans. Biomed. Eng.* 1990, 37, 85–98. [CrossRef]
- [7]. Van Alste and, T.S.; Schilder, J.A. Removal of base-line wander and power-line interference from the ECG by an efficient FIR filter with a reduced number of taps. *IEEE Trans. Biomed. Eng.* 1985, BME-32, 1052–1060. [CrossRef]
- [8]. Oster, J.; Behar, J.; Sayadi, O.; Nemati, S.; Johnson, A.E.W.; Clifford, G.D. Semisupervised ECG Ventricular Beat Classification with Novelty Detection Based on Switching Kalman Filters. *IEEE Trans. Biomed. Eng.* 2015, 62, 2125–2134. [CrossRef]

- [9]. Dalin Tang, Z.T.; Canton, G.; Hatsukami, T.S.; Dong, L.; Yuan, X.H.A.C. Local critical stress correlates better than global maximum stress with plaque morphological features linked to atherosclerotic plaque vulnerability: An in vivo multi-patient study. *Biomed. Eng. Online* 2009, 8, 15. [CrossRef]
- [10]. Frolich, L.; Dowding, I. Removal of muscular artifacts in EEG signals: A comparison of linear decomposition methods. *Brain Inf.* 2018, 5, 13–22. [CrossRef] [PubMed]
- [11]. Blanco-Velasco, M.; Weng, B.; Barner, K.E. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Comput. Biol. Med.* 2008, 38, 1–13. [CrossRef]
- [12]. Alfaouri, M.; Daqrouq, K. ECG Signal Denoising By Wavelet Transform Thresholding. *Am. J. Appl. Sci.* 2008, 5, 276–281. [CrossRef]
- [13]. Xu, Y.; Luo, M.; Li, T.; Song, G. ECG Signal De-noising and Baseline Wander Correction Based on CEEMDAN and Wavelet Threshold. *Sensors* 2017, 17, 2754. [CrossRef] [PubMed]
- [14]. Han, G.; Xu, Z. Electrocardiogram signal denoising based on a new improved wavelet thresholding. *Rev. Sci. Instrum.* 2016, 87, 084303. [CrossRef]
- [15]. Pinho, A.; Pombo, N.; Silva, B.M.C.; Bousson, K.; Garcia, N. Towards an accurate sleep apnea detection based on ECG signal: The quintessential of a wise feature selection. *Appl. Soft Comput.* 2019, 83, 105568. [CrossRef]
- [16] Moavenian, M.; Khorrami, H. A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification. *Expert Syst. Appl.* 2010, 37, 3088–3093. [CrossRef]
- [17]. Appathurai, A.; Jerusalin Carol, J.; Raja, C.; Kumar, S.N.; Daniel, A.V.; Jasmine Gnana Malar, A.; Fred, A.L.; Krishnamoorthy, S. A study on ECG signal characterization and practical implementation of some ECG characterization techniques. *Measurement* 2019, 147, 106384. [CrossRef]
- [18]. Chauhan S, Vig L. 2015 Anomaly detection in ECG time signals via deep long short-term memory networks. In 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 1 – 7. 19– 21 October, Paris, France. Piscataway, NJ: IEEE. (doi:10.1109/DSAA.2015.7344872).

- [19]. Oresko JJ, Jin Z, Cheng J, Huang S, Sun Y, Duschl H, Cheng AC. 2010 A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. *IEEE Trans. Inf. Technol. Biomed.* 14, 734– 740. (doi:10.1109/TITB.2010.2047865)
- [20] J. S. Wang, W. C. Chiang, Y. T. Yang, and Y. L. Hsu, "An effective ECG arrhythmia classification algorithm," *Bio-Inspired Computing and Applicat*, Springer Berlin Heidelberg, pp. 545-550, 2012.
- [21] A. Dallali, A. Kachouri, and M. Samet, "Classification of Cardiac Arrhythmia Using WT, HRV, and Fuzzy C-Means Clustering," *Signal Processing: An Int. J. (SPJI)*, vol. 5, no. 3, pp. 101-109, 2011.
- [22] M. Vijayavanan, V. Rathikarani, and P. Dhanalakshmi, "Automatic Classification of ECG Signal for Heart Disease Diagnosis using morphological features," *Int. J. of Comput. Sci. and Eng. Technology (IJCSET)*, vol. 5, no. 4, pp. 449-455, 2014.
- [23] V. K. Srivastava and D. Prasad, "Dwt-Based Feature Extraction from ecg Signal," *American J. of Eng. Research (AJER)*, vol. 2, no. 3, pp. 44- 50, 2013.
- [24] M. Korurek and B. Dogan, "ECG beat classification using particle swarm optimization and radial basis function neural network," *Expert syst. with Applicat.*, vol. 37, no. 12, pp. 7563-7569, 2010.
- [25]. Üstündağ, M.; Gökbulut, M.; Sengür, A.; Ata, F. Denoising of weak ECG signals by using wavelet analysis and fuzzy thresholding. *Netw. Modeling Anal. Health Inform. Bioinform.* 2012, 1, 135–140. [CrossRef]
- [26] A. T. Sadiq and N. H. Shukr, "Classification of Cardiac Arrhythmia using ID3 Classifier Based on Wavelet Transform," *Iraqi J. of Sci.*, vol. 54, no. 4, pp. 1167-1175, 2013.
- [27] N. P. Joshi and P. S. Topannavar, "Support vector machine based heartbeat classification," *Proc. of 4th IRF Int. Conf.*, pp. 140-144, 2014.
- [28] E. Zeraatkar et al., "Arrhythmia detection based on Morphological and time-frequency Features of t-wave in Electrocardiogram," *J. of medical signals and sensors*, vol. 1, no. 2, pp. 99-106, 2011.
- [29] S. N. Yu and K. T. Chou, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Syst. with Applicat.*, vol. 34, no. 4, pp. 2841-2846, 2008.

[30] D. Joshi and R. Ghongade, "Performance analysis of feature extraction schemes for ECG signal classification," *Int. J. of Elect., Electron. and Data Commun.*, vol. 1, pp. 45-51, 2013.

[31] M. K. Das and S. Ari, "ECG Beats Classification Using Mixture of Features" *Int. Scholarly Research Notices*, 2014.

[32] X. Tang and L. Shu, "Classification of Electrocardiogram Signals with RS and Quantum Neural Networks," *Int. J. of Multimedia and Ubiquitous Eng.*, vol. 9, no. 2, pp. 363-372, 2014.

[33] Wang, Zh., Yan, W., Oates, T. Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline // arXiv preprint. arXiv:1611.06455, 2016.