Global Journal of Computer and Engineering Technology (GJCET)

Vol. 1, No. 2, July, 2021 ISSN: 2767-1933

# ENSEMBLE CLASSIFICATION FOR ARRHYTHMIA DETECTION

Shivani Kashyap, Kurukshetra University, India (shivanikashyap8989@gmail.com)
Ekta, Kurukshetra University, India (ekta.tamak@gmail.com)
Reeta Devi, U.I.E.T, Kurukshetra University, India (reetadahiya@kuk.ac.in)

**ABSTRACT**

Heart diseases cause deaths in most of the countries, thus needs proper methods to detect cardiac condition of the patient. Electrocardiography (ECG) is used by the cardiologist to measure how much electrical activity the heart has. Cardiac arrhythmia is a common cardiovascular disease that may involve atrial fibrillation or ventricle fibrillation. Irregular heartbeat (too slow or too fast) constitutes another kind of arrhythmias. These irregularities of the heart can be detected with single lead (preferably lead II) ECG. The presented work, therefore, focused on to identify multiple abnormalities such as atrial fibrillation, ventricular fibrillation, bradycardia and tachycardia by the use of multiple features of single lead ECG using an ensemble classifier. The features included are R-R interval, R-R amplitude, R-R speed, gender and age. The classification accuracy of the designed algorithm has been measured as 99.41% which is a significant improvement over the previously published studies.

**Keywords:** Annotation file, Arrhythmia, Electrocardiogram, Ensemble classifier and Features.

## 1. INTRODUCTION

Electrocardiogram is the database of heart-produced electrical activity. It is obtained by non- invasive method of placing electrodes on the body at standardized positions. The electrical waves are generated by depolarizing and repolarizing of the certain cells due to the movement of sodium ($Na^+$) and potassium ($K^+$) ions in the blood (Xie, 2020). ECG signals and heart rate represents heart safety, and any changes in the rate, rhythm or morphological pattern of ECG signals indicate heart arrhythmia. It is detected and diagnosed by analyzing ECG waveform (Rangayyan, 2002) (Devi et al., 2016, 2017, 2019a, 2019b). ECG records are revealed for morphological statements and rhythm statements for diagnostic purpose. Morphological statements are based on the ECG wave form to define the working state of muscle masses. Ventricular fibrillation and atrial fibrillation are two common types of morphological statements. Ventricular fibrillation as shown in Figure 1 is a quick life-threatening rhythm beginning in the heart ventricles that can be caused by heart attack, whereas, atrial fibrillation (Figure 2) is an abnormal, often rapid rate resulting in inadequate blood flow (Chen et al., 2017).

On the other hand rhythm statement concerns the location and rate of pacemaker and the transmission of impulses by cardiac conduction system (Rjni & Kaur, 2013). It is also of two types i.e. bradycardia (lower heart rate, under sixty beats per minute) and tachycardia (quick heart rhythm, which can be normal or abnormal) as shown in Figure 3.
An arrhythmia monitor is used to detect the occurrence of these arrhythmias in the heart rhythm. Figure 4 shows a block diagram of arrhythmia monitor (Khandpur, 2014). As shown, they are designed to detect a premature or widened QRS Complex. These instruments store QRS width and R-R interval for the reference. External ECG record runs automatically during normal store mode, so that reference heart beat are examined and determined as to whether they are truly representative. An alarm is started after the detection of ectopic beats (up to the rate of 6 per minutes or 12 per minutes) like ventricular premature or widened varieties (Waugh & Grant, 2001).

Now a days, many machine learning and deep learning techniques have been practiced to classify the arrhythmias automatically. The main challenge in ECG classification is to choose a proper classifier that can classify the arrhythmia with acceptable accuracy. In this work, we present an efficient algorithm based on ensemble classification to detect the cardiac arrhythmia with improved accuracy. Ensemble learning is a method of constructing a new classifier from a set of base classifiers that performs better than any of the constituent classifiers (Tapas et al., 2017). Two methods are generally used for ensemble learning i.e. averaging methods and boosting methods. In averaging methods, the driving principle is to build several estimators independently and then to average their predictions. On average, the combined estimator is usually better than any of the single base estimator because its variance is reduced.

Bagging method and random forests methods are common examples of averaging based ensemble classifier. In boosting approaches, on the other hand, base estimators are produced progressively and the total estimator's bias is reduced. The goal is to create a powerful ensemble by combining multiple weak models. Adaboost and gradient tree boosting are popular examples of boosting ensemble classifier. In this work, we propose to classify multiple arrhythmias like atrial fibrillation, ventricular fibrillation, tachycardia and bradycardia with a designed algorithm of an ensemble classifier for efficient classification with improved accuracy.

In the following sections, we explain the literature review on related studies, the experimental work including results and discussions and conclusions of the present work.

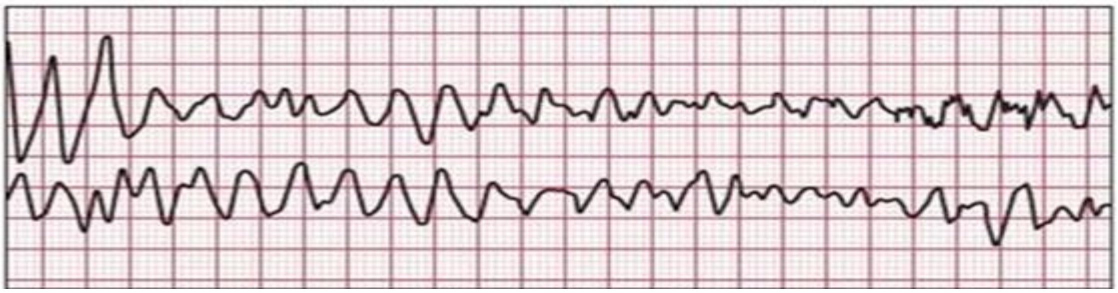

**Figure 1:** Ventricular fibrillation

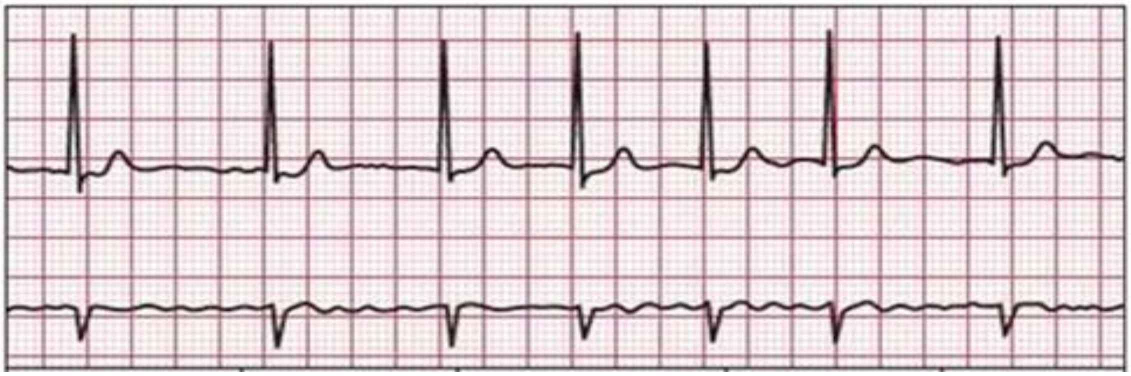

**Figure 2:** Atrial fibrillation

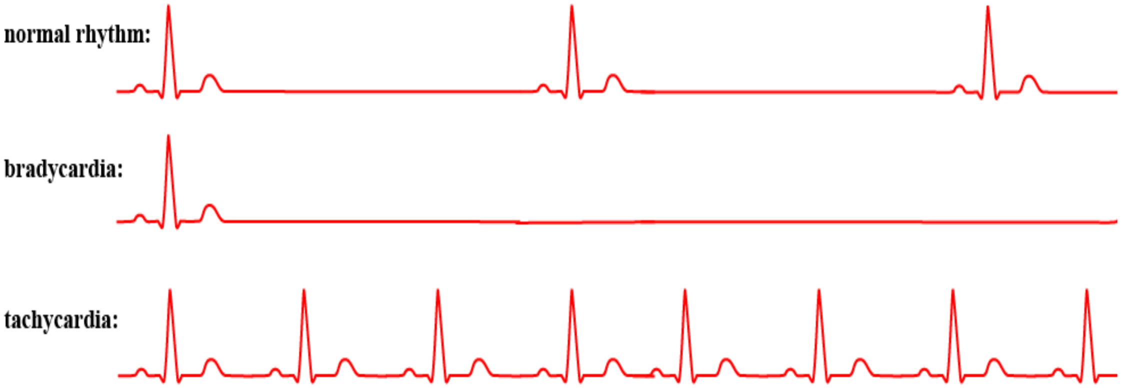


**Figure 3:** Rhythm statement

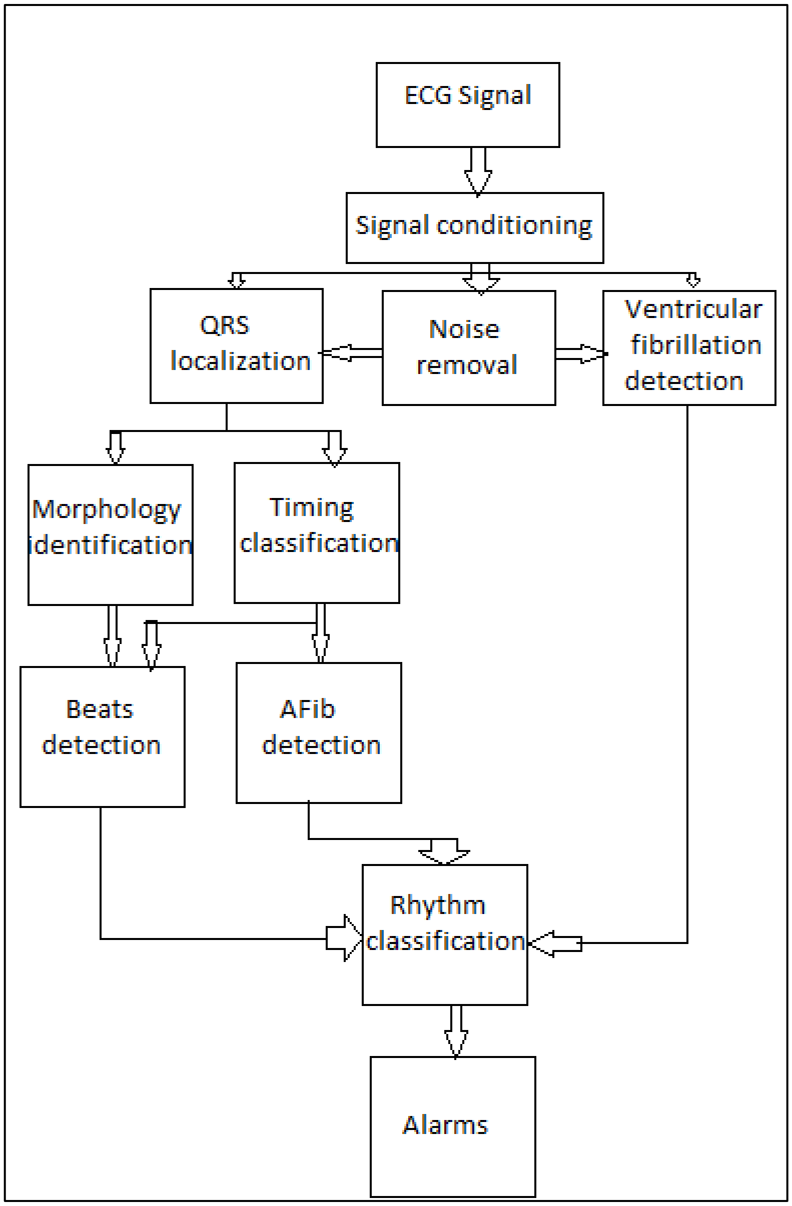


**Figure 4:** Block diagram of arrhythmia monitor (Khandpur, 2014)

## 2. LITERATURE REVIEW

As discussed above, the arrhythmia monitor detects the underlying arrhythmia by measuring the QRS complexes. Digital filter algorithms that are used to distinguish normal and abnormal QRS complexes are Pan-Tompkins algorithm (Pan & Tompkins, 2007) and ST/AR arrhythmia algorithm. In the recent times, after measuring the QRS complex, machine learning and deep learning techniques are used to classify the occurrence of arrhythmia. A summary of some popular techniques identifying cardiac arrhythmia using machine/deep learning techniques is presented in Table 1.

**Table 1:** Literature review

| S. No. | Reference | Technique used | Sensitivity | Accuracy | Specificity |
|---|---|---|---|---|---|
| 1. | (Vanitha et al., 2014) | Hybrid classifier using pNN, kNN and SVM | 97% | 90% | 96% |
| 2. | (Kalidas & Tamil, 2015) | Combining local and machine learning techniques | - | Real time TPR=94% TNR= 82% | - |
| 3. | (Dash & Rao, 2016) | Three layer feed-forward back propagation neural network | 94.90% | 99.24% | 99.57% |
| 4. | (Paradkar & Chowdhury, 2017) | Photoplethysmography | - | TPR= 93% TNR= 53.78% | - |
| 5. | (Isin & Ozdalili, 2017) | Deep learning using AlexNet | – | 92% | – |
| 6. | (Tapas et al., 2017) | Random forest and logistic regression based ensemble classifier | 87% with ECG dataset | 97% with ECG dataset | |
| 7. | (Verma & Agarwal, 2018) | Deep learning algorithm integrating CNN and LSTM with oversampling | - | F1 Score= 93.6% | - |
| 8. | (Lui & Chow, 2018) | CNN and RNN networks for portable ECG devices | 92.4% | F1 score= 94.6% | 97.7% |
| 9. | (Neha et al., 2019) | ECG and PPG based sensor and SVM | 98% | 97.674% | 98% |
| 10. | (Sai Krishna & Nithya Kalyani, 2019) | Deep learning techniques such as Recurrent neural network | – | - | – |

## 3. EXPERIMENTAL WORK

In this analysis, the approach used to detect arrhythmia differs slightly from the methods mentioned in Table 1. Here the segmentation strategy is based on annotation file, which has been downloaded along with the data from the MIT-BIH database (Goldberger et al., 2000). Therefore, the following steps are implemented in order to segment the real 30 minute ECG signal obtained from the database. The age and gender is extracted for each record from it's annotation file. The leftover three features i.e. R-R amplitude, R-R speed, R-R interval is extracted by using the below method illustrated below in flow chart shown in Figure 5.

A description of these steps is presented below:

I. At first ECG signals are obtained from the standard international MIT-BIH database and loaded into the MATLAB environment. A plot of this raw ECG signal is shown in Figure 6.
II. The downloaded signal appears to be crude. Therefore, pre-processing of the crude signal is performed by removing the base and normalizing by gain. Then, moving average filter is used to cancel the baseline wander noise and comb notch filter to remove the power-line interference from the ECG signal as shown in Figure 7.
III. After pre-processing the ECG signal, the feature extraction process is performed. The extracted features are used to retrieve diagnostic information from the underlying ECG signal. To classify an electrocardiogram signal as normal or abnormal, the very first step will be to define its attributes and store the values in particular variable and then classify the different abnormalities. QRS complex detection is performed by using the

pan-tompkins algorithm as shown in Figure 8 (a-b). In QRS complex, since R is considered to have the largest amplitude in a normal Lead-II ECG signal and is the sharpest component w.r.t. all other peaks. R peak detection is done by obtaining the local minima greater than the adaptively set threshold by estimating the amplitudes, temporal positions and duration. Then the R-R interval is calculated using R-spike detection method which basically calculates the duration from one R-point to the next R-point (successive R's). The extracted features are presented in Table 2.

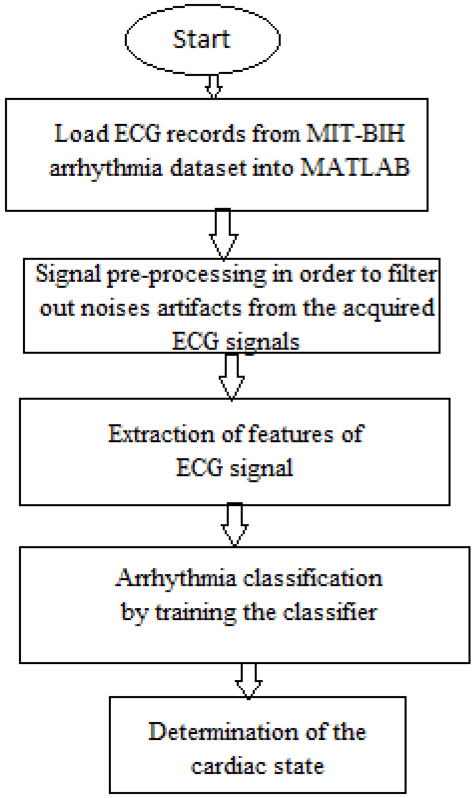


**Figure 5:** Flow chart of the proposed methodology

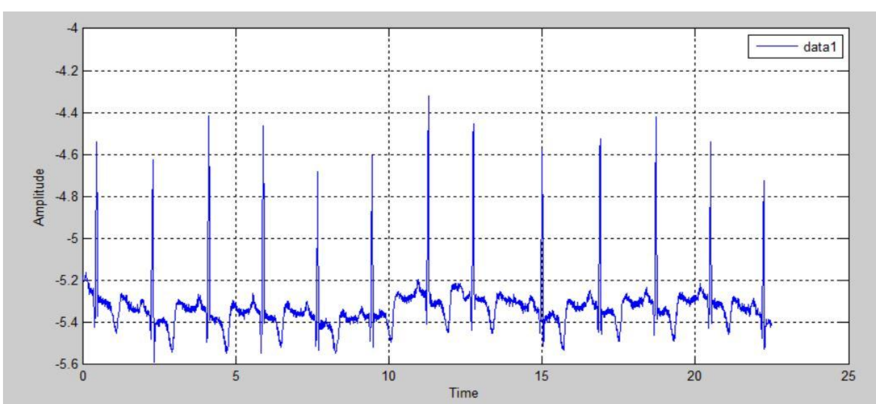


**Figure 6:** Raw ECG signal

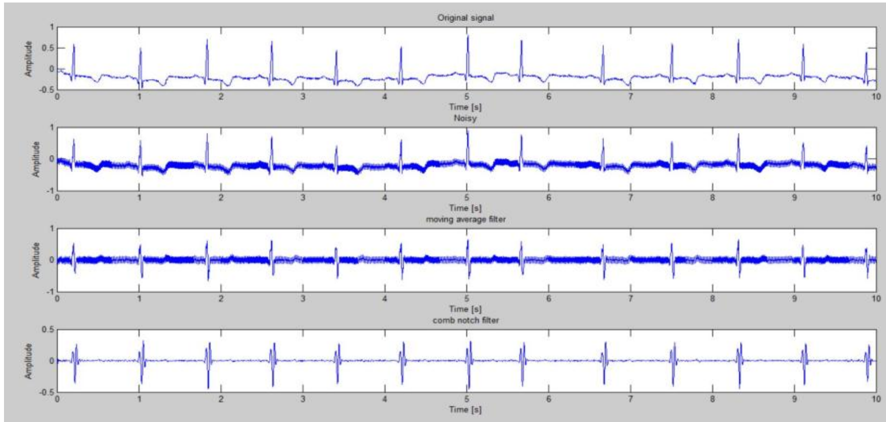


**Figure 7:** Filtered ECG output

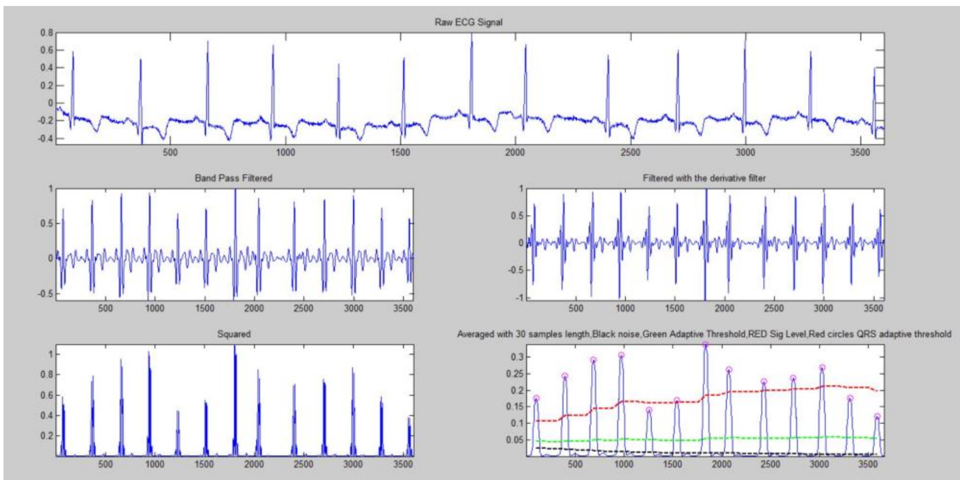


**Figure 8 (**a): QRS complex detection using Pan-Tompkins algorithm

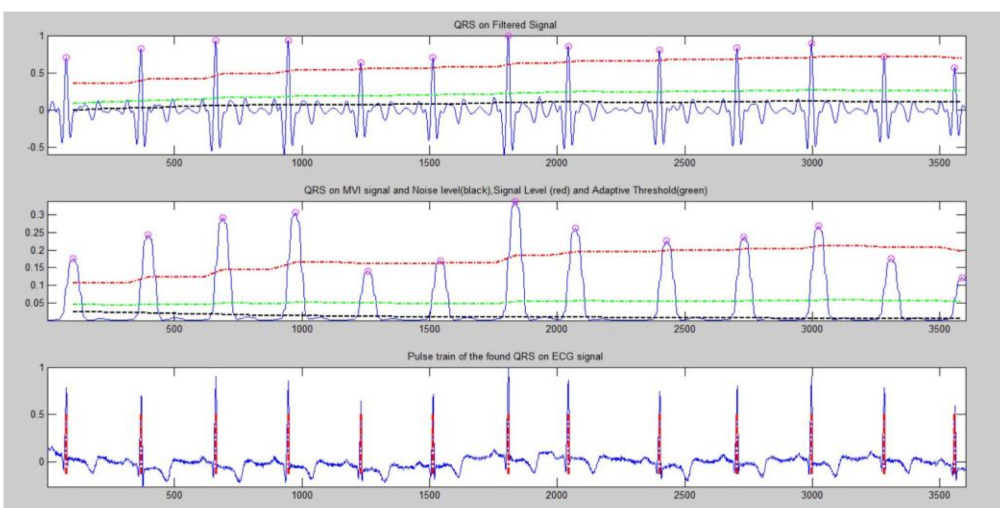


**Figure 8 (b):** Detected R-peaks by using pan-Tompkins algorithm

**Table 2:** Different features of ECG signals

| Record no. | Amplitude | R-R interval | R-R speed | Age | Gender |
|---|---|---|---|---|---|
| 106 | 2.2845 | 1.68 | 1.36 | 21 | F |
| 119 | 1.3059 | 1.77 | 0.74 | 54 | F |
| 201 | 0.5155 | 0.56 | 0.91 | 89 | M |
| 100 | 0.8049 | 0.69 | 1.15 | 69 | M |
| 101 | 1.4347 | 0.99 | 1.44 | 75 | F |
| 119 | 1.3838 | 0.94 | 1.46 | 51 | F |
| 201 | 0.9021 | 2.31 | 0.39 | 68 | M |
| 208 | 1.4629 | 1.32 | 1.10 | 23 | F |
| 200 | 2.0880 | 0.97 | 2.13 | 64 | M |
| 203 | 1.4766 | 0.62 | 2.35 | 43 | M |
| 233 | 0.7865 | 0.63 | 1.24 | 57 | M |
| 205 | 0.5488 | 0.63 | 0.86 | 59 | M |
| 207 | 0.7882 | 0.6 | 1.31 | 89 | F |
| 213 | 0.7374 | 0.51 | 1.43 | 61 | M |
| 214 | 1.0590 | 0.64 | 1.63 | 53 | M |
| 215 | 1.6109 | 0.69 | 2.31 | 81 | M |
| 223 | 1.9354 | 0.70 | 2.73 | 73 | M |

In this work ensemble classifier is used for classification. The word Ensemble is a Latin-inferred word which signifies 'association of parts'. The ordinary classifiers that are utilized frequently by other researchers are inclined to make compromised accuracy. As discussed in the introduction section above, ensemble learning is a method of creating different base classifiers from which another classifier is determined which performs superior to its constituent classifiers. These base classifiers may vary in the calculation utilized, hyper boundaries or the preparation set. The structure of the designed ensemble classifier is shown in Figure 9. The trained ensemble classifier classifies the arrhythmia whether it is normal, tachycardia, bradycardia, atrial fibrillation or ventricular fibrillation.

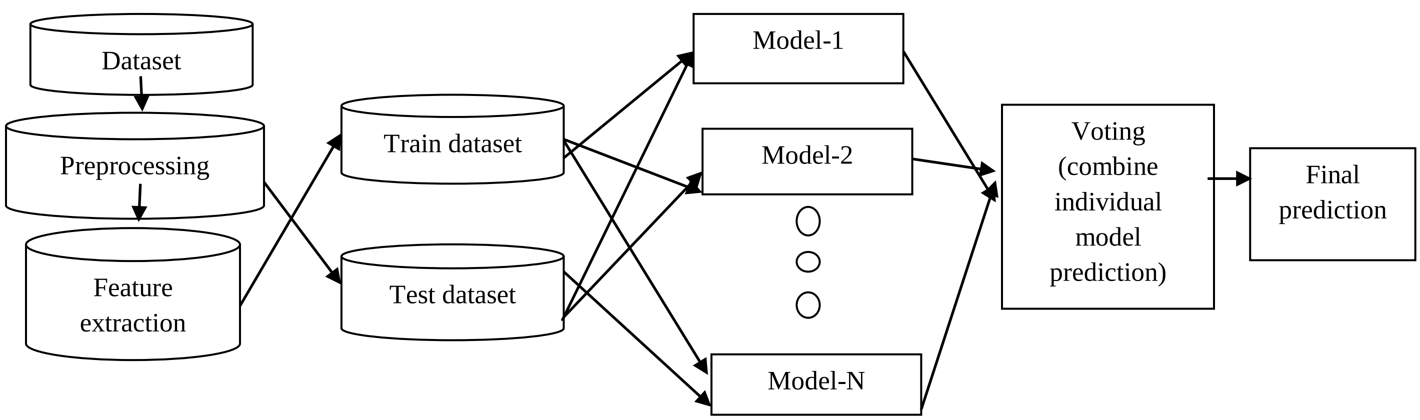


**Figure 9:** Ensemble classifier

Finally, the predictions made by ensemble classifier are measured in terms of accuracy (Acc.), sensitivity (Sen), precision (P) and specificity (Sp) using the formulae given in equations (1-4).

$$Acc= \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Sen= \frac{TP}{TP+FN} \tag{2}$$

$$Sp= \frac{TN}{TN+FP} \tag{3}$$

$$P= \frac{TP}{TP+FP} \tag{4}$$

where TP is True Positive (i.e. predict right arrhythmias), TN is True Negative (i.e. predict normal signal correctly), FP is False Positive (i.e. predict arrhythmia incorrectly) and FN shows False Negative (i.e. predict wrong normal ECG signal ). For the ensemble classifier the value of FP, FN, TP and TN are given below in Table 3 and measured values of Acc, Sen, Sp, and P are plotted in Figure 10.

**Table 3:** Parameters of the Ensemble classifier

| Parameters | TP | TN | FP | FN |
|---|---|---|---|---|
| Ensemble classifier | 273.25 | 824.25 | 2.25 | 2.25 |

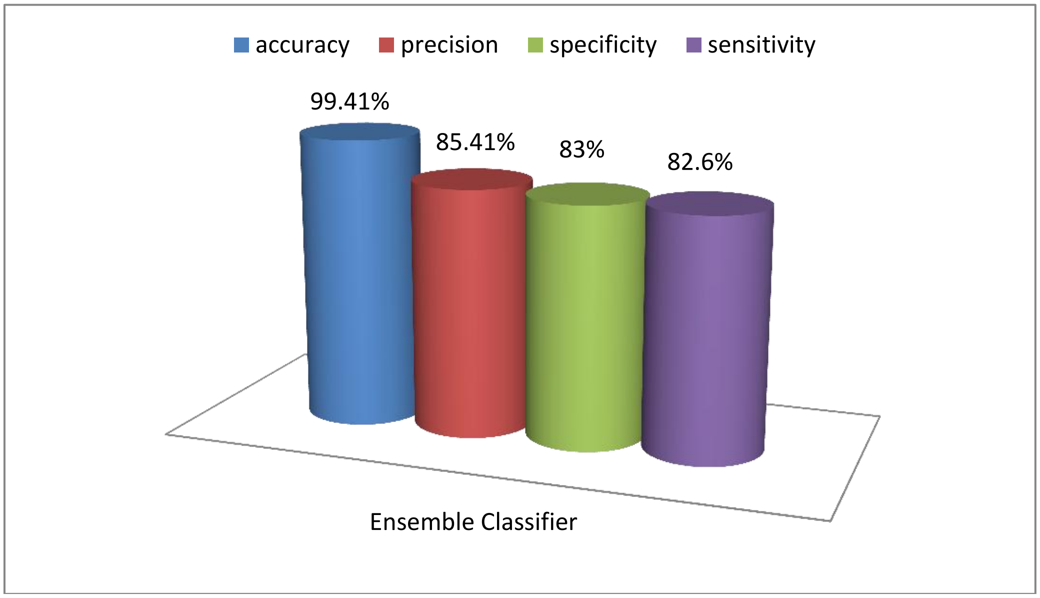


**Figure 10:** Performance parameters of the ensemble classifier

**Table 4:** Comparison with other published results

| Method | Technique used | Accuracy | Specificity | Sensitivity | Precision |
|---|---|---|---|---|---|
| Proposed | Ensemble classifier | 99.41% | 83% | 82.6% | 85.41% |
| Neha2019 at el., | SVM | 97.674% | 98% | 98% | 98% |
| Verma2018 at el., | CNN + LSTM | F1 Score: 93.6% | - | - | - |

A fair comparison of the proposed approach is presented in Table 4 with different methods of arrhythmia classification. From this table, it can be observed that the proposed methodology obtained the superior performance than the other published methods.

## 4. CONCLUSION

Worldwide cardiac disease are rising day-by-day. Occurrence of arrhythmia is a common problem in the general population.Diagnosis of arrhythmia includes patient ECG and other clinical data. Consequent detection of multiple arrhythmias is a tough job. Present work proposed an ensemble classifier for detecting four different kinds of arrhythmia i.e. ventricular fibrillation, atrial fibrillation, tachycardia and bradycardia. The classification accuracy have been measured as 99.41% which is a significant improvement over previously published results. The sensitivity, specificity and precision values are obtained as 82.6%, 83%, and 85.41% respectively. Therefore, it can be concluded from the obtained results that the proposed ensemble classifier can be efficiently used with improved accuracy in detection of multiple arrhythmias. Future work of the present study would like to focus on other features of the ECG signal that can give better performance in terms of improved values of the sensitivity, specificity and precision in arrhythmia detection.

## REFRENCES


Chen, S., Hua, W., Li, Z., Li, J., & Gao, X. (2017). Heartbeat classification using projected and dynamic features of ECG signal. *Biomedical Signal Processing and Control*, *31*. https://doi.org/10.1016/j.bspc.2016.07.010

Dash, S. K., & Rao, G. S. (2016). Robust multiclass ECG arrhythmia detection using balanced trained neural network. *International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016*, 186–191. https://doi.org/10.1109/ICEEOT.2016.7754994

Devi, R., Tyagi, H. K., & Kumar, D. (2016). Heart Rate Variability Analysis for Early Stage Prediction of Sudden Cardiac Death. *World Academy of Science, Engineering and Technology International Journal of Electrical and Computer Engineering*, *10*(3), 432–435. https://doi.org/doi.org/10.5281/zenodo.1124097

Devi, R., Tyagi, H. K., & Kumar, D. (2017). Early Stage Prediction of Sudden Cardiac Death Using Linear and Non-Linear Features of Heart Rate Variability. *International Journal of Electronics, Electrical and Computational System*, *6*(9), 742–754.

Devi, R., Tyagi, H. K., & Kumar, D. (2019a). A novel multi-class approach for early-stage prediction of sudden cardiac death. *Biocybernetics and Biomedical Engineering*, *39*(3), 586–598. https://doi.org/10.1016/j.bbe.2019.05.011

Devi, R., Tyagi, H. K., & Kumar, D. (2019b). Signal Processing Methods for Identification of Sudden Cardiac Death. *FTNCT 2018, CCIS 958, Springer Nature Singapore*, *958*(January), 57–72. https://doi.org/10.1007/978-981-13-3804-5

Goldberger, A. L., Amaral, L. A., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*, *101*, e215–e220. https://doi.org/10.1161/01.cir.101.23.e215

Isin, A., & Ozdalili, S. (2017). Cardiac arrhythmia detection using deep learning. *Procedia Computer Science*, *120*, 268–275. https://doi.org/10.1016/j.procs.2017.11.238

Kalidas, V., & Tamil, L. S. (2015). Enhancing accuracy of arrhythmia classification by combining logical and machine learning techniques. *Computing in Cardiology*, *42*(March 2016), 733–736. https://doi.org/10.1109/CIC.2015.7411015

Khandpur, R. S. (2014). *Handbook of Biomedical Instrumentation* (III). McGraw Hill Education (India) Private Limited.

Lui, H. W., & Chow, K. L. (2018). Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices. *Informatics in Medicine Unlocked*, *13*, 26–33. https://doi.org/10.1016/j.imu.2018.08.002

Neha, Kanawade, R., Tewary, S., & Sardana, H. K. (2019). Photoplethysmography Based Arrhythmia Detection and Classification. *2019 6th International Conference on Signal Processing and Integrated Networks, SPIN 2019*, 944–948. https://doi.org/10.1109/SPIN.2019.8711737

Pan, J., & Tompkins, W. J. (2007). A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, *BME-32*(3), 230–236. https://doi.org/10.1109/tbme.1985.325532

Paradkar, N., & Chowdhury, S. R. (2017). Cardiac arrhythmia detection using photoplethysmography. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 113–116. https://doi.org/10.1109/EMBC.2017.8036775

Rangayyan, R. M. (2002). *Biomedical Signal Analysis, A Case -Study Approach* (M. Akay (ed.)). IEEE Press Series on Biomedical Engineering.

Rjni, & Kaur, I. (2013). Electrocardiogram Signal Analysis - An Overview. *International Journal of Computer Applications*, *84*(7), 22–25.

Sai Krishna, V., & Nithya Kalyani, A. (2019). Prediction of cardiac arrhythmia using artificial neural network. *International Journal of Recent Technology and Engineering*, *8*(1 Special Issue 4), 462–466.

Tapas, N., Lone, T., Reddy, D., & Kuppili, V. (2017). Prediction of cardiac arrest recurrence using ensemble classifiers. *Sadhana*, *2*. https://doi.org/10.1007/s12046-017-0683-z

Vanitha, L., Suresh, G. R., & Jenefarsheela, C. (2014). Sudden Cardiac Death prediction system using Hybrid classifier. *2014 International Conference on Electronics and Communication Systems, ICECS 2014*. https://doi.org/10.1109/ECS.2014.6892677

Verma, D., & Agarwal, S. (2018). Cardiac Arrhythmia Detection from Single-lead ECG using CNN and LSTM assisted by Oversampling. *2018 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018*, 14–17. https://doi.org/10.1109/ICACCI.2018.8554541

Waugh, A., & Grant, A. (2001). *Ross and Wilson Anatomy and Physiology in Health and Illness* (Ninth). Elsevier Ltd.

Xie, C. (2020). Biomedical signal processing: an ECG application. In *Leveraging Data Science for Global Health* (pp. 285–303). Springer International Publishing. https://doi.org/10.1007/978-3-030-47994-7_17