International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Cardiovascular Abnormalities Classification Model Using Machine Learning and Signal Processing Techniques

B. Venkataramanaiah[1*], M. Anuradha[2], K. Balasubramanian[3], C. Gnanaprakasam[4], D. Praveen Kumar[5], R. Palanikumar[6]

Submitted: 26/09/2023 **Revised**: 16/11/2023 **Accepted**: 27/11/2023

Abstract: Unexpected impulsive cardiac death (ISD) can occur when a person has cardiovascular illness. Electrocardiogram (ECG) signal can be used to identify impulsive cardiac mortality risks.This paper describes an intelligent human cardiac monitoring approach based on machine learning. Existing method suffers from misclassification of heart diseases. To reduce misclassification, we have proposed two innovative models for cardiovascular disease (CVD) identification. In First model, Principal component Analysis (PCA) features and Wavelet transform (WT) features are applied for machine learning classifiers such as Multinomial logistic regression (MLR) and Random Forest (RF) to find CVD. In model 2: Heart Rate Variability (HRV) and WT features are applied to Nave Bayes (NB), Decision Tree (DT) and k nearest neighbour (KNN) machine learning classifiers for classification in order to create an intelligent machine learning based cardiovascular diseases risk monitoring system. Effective features are important when Data is classified into normal or abnormal subjects. Proposed novel approach identified risks with the highest degrees of accuracy: 99.6(model 1 by MLR), and 99.3% (model 2 by DT).The outcomes demonstrate that the proposed strategy is reliable and effective for identifying impulsive cardiac risk. Effectively identifying risk factors for impulsive cardiac death is the goal of the proposed research

Keywords: *ECG, biomedical signal processing, heart rate variability, wavelet transform, PCA, Impulsive cardiac death (ISD)*

1. Introduction

This In order to find accurate learning, artificial intelligence, machines, and neural networks have been introduced to the biomedical area. Impulsive cardiac death (ISD) refers to unanticipated fatalities from cardiovascular illness that have a date and time but are not foreseen in advance. The majority of fatalities are caused by severe abrupt cardiac arrest; these cases should be found quickly.

An electrocardiogram (ECG) is a recording of the electrical activity of the human heart that includes data on P, Q, R, S, T, and U waves. Using this data, biomedical signal processing may filter out noise and get precise information. Transforms are used to process the data that was obtained after filtering in order to obtain the ECG signal's characteristics. By employing features extracted from the ECG signal, machine learning classifiers are utilized to categories anomalies in the ECG signal. Several criteria, including time domain features, frequency domain features, and wavelet analysis, are used to determine abnormalities in ECG signals. Amplitude, intervals, peaks, mean, variance, standard deviation, median, energy, and noise power are some examples of these properties. Features selection performance are measured in terms of True positive, True Negative, Falsepositive, False Negative, Precision, Recall, F- measures, sensitivity, precision and specificity.

The current state of ECG research is difficult since there are so many restrictions on the signals that may be analyzed, including sampling frequency, signal duration, sample count, filter selection for noise removal, and classifier selection for classifying the signal into normal or abnormal patients. These elements will determine the features used for classification's accuracy, sensitivity, and specificity. A dimension-reduction method, PCA also includes Eigen values, Eigen vectors, and other characteristics. Deep neural networks and machine learning classifiers are also applied to image analysis to categories target detection results.

In order to identify and categories diseases accurately, artificial intelligence, machine learning, and neural networks have been introduced to the biomedical profession. The rate of sudden cardiac mortality has been rising rapidly over the past 20 years. Death occurs shortly

*[1]Department of Electronics and Communication Engineering, Vel Tech ngarajan Dr Sagunthala R&D institute of Science and Technology, Avadi-)0062, Chennai, Tamil Nadu, India. Email: enkataramanaiah@veltech.edu.in**
[2]Department of Computer Science and Engineering, S.A.Engineering ollege, Chennai-600077.Tamil Nadu India. Email: anuparini@gmail.com
[3]Department of Computer Applications, Kalasalingam Academy of esearch and Education (Deemed to be University), Krishnankoil-626126. mil Nadu, India Email:ksbala75@gmail.com
[4]Department Of Artificial Intelligence and Data Science, Panimalar ngineering College Poonthamalli, Chennai – 600123.Tamil Nadu ndia,Email:cgn.ds2021@gmail.com
[5]Assistant Professor, Department of Electrical and Electronics ngineering, Mohan Babu University (Erstwhile Sree Vidyanikethan ngineering College),Tirupati, 517102.Email: aveenkumar.d@vidyanikethan.edu
[6]Associate Professor, Department of Computer Science and ngineering,P.S.R. Engineering College (Autonomous), Sevalpatti, Sivakasi - 26140, Tamil Nadu,India, Email:palanikumar@psr.edu.in

after the patient's symptoms are identified. Currently, an embedded hardware module is required to recognize sudden cardiac death and diagnose cardiovascular illnesses. Even though there are many hardware modules described in the literature, effective hardware architectures were still needed to detect anomalies in the ECG signal. In this study, we suggest a brand-new hardware component for diagnosing cardiovascular conditions. Architecture's approaches for classifying diseases play a significant role in addition to hardware module methods.

The motivation of the focusing work is Disease classification. Feature extraction and disease identification are measured by accuracy in terms of true positive rate, true negative, false positive and false negative

2. Literature Survey

Automatic atrial fibrillation identification has been established based on heart rate variability and ECG spectral characteristic. There are some difficulties in the work proposed by Abbasi [1] that will be the subject of future study on the EEG signal. ECG-based biometrics are a new trend for identifying attendance of a certain person, and biometric is attendance for many organization personnel. Adeli et al. [2] present a brief review of artificial intelligence, machine learning, fuzzy systems, and neural networks. Impulsive cardiac death (ISD) refers to unanticipated fatalities from cardiovascular illness that have a date and time but are not foreseen in advance.Convolutional neural networks are used to anticipate patient abnormalities through simultaneous ECG and EEG monitoring of human health. The majority of fatalities are caused by severe abrupt cardiac arrest, which should be detected quickly. Ahmadlou et al. presented neural network-based cardiac disease identification [3]. Our future work will focus on the aberrant identification of heart disease using neural networks. A. Rahman et al. [4] perform the classification of heartbeats and the detection of cardiomyopathy.

A person with ventricular arrhythmia is the cause of rapid cardiac death. Various machine learning classifiers have automated strategies in place to identify this danger. Ahmadlou et al. designed a deep neural network model for local decision circles. Today's classification and regression-based industry is led by machine learning and deep learning algorithms. A smart vest system with an IOT module was created to assess signal quality and detect light-weight QRS to diagnose cardiovascular disorders [5]. [QRS fragmentation is one way used by different machine learning classifiers to recognize and quantify ECG data. C.H. Hung introduced a health management strategy that uses blood pressure measurements on older adults to identify cardiovascular problems. Arterial Stiffness is monitored thorough wearable sensor by machine learning algorithms' module and PPG module is integrated with Micro controller unit to monitor Arterial Stiffness of the patient [6]. Wireless monitoring systems have been employed in optical remote sensing to track patients' cardiovascular illnesses using ECGs received from sensors placed on their bodies. A person with ventricular arrhythmia is the cause of sudden cardiac mortality; different machine learning classifiers have automated strategies in place to identify this risk [7]. Applications for signal processing and picture processing use the effective wavelet transform technology. Machine learning and neural networks are two distinct artificial intelligence techniques for applications in classification and regression. Intelligent machine monitoring based on wavelet transforms and neural networks is shown with sound signal processing.A sudden cardiac death analysis model was put up by Ebrahimzadeh et al. [8] to detect sudden cardiac arrest utilizing linear techniques and time frequency methodologies. A smart vest system was created using an IOT module for light-weight QRS detection and signal quality assessment in order to detect cardiovascular illnesses. Ebrahimzadeh et al. implemented the MIT BIH database-based feature selection-based HRV for forecasting sudden cardiac death [9]–[10].Machine learning classifiers are evaluated and trained to identify diseases with the best degree of accuracy. Here, KNN and SVM machine learning classifiers produce relatively low values for accuracy, sensitivity, and specificity. The neural network classifier Multilayer Perception (MLP) is one of the most efficient ones. Data was classified into sudden cardiac arrest or normal using time domain ECG characteristics applied to MLP [11].ECG signal database for normal and abnormal subjects is obtained from MIT BIH arrhythmia, morphological and dynamic features are extracted with heart beat classification by using wavelet-based support vector machine classifier which has low accuracy to find cardiovascular abnormalities [12].

Numerous publications have utilized the MIT BIH database-based feature selection-based HRV for forecasting sudden cardiac death. Machine learning algorithms use wearable sensors to assess arterial stiffness. Microcontroller unit and ECG and PPG modules are integrated to track the patient's arterial stiffness [13].Here, KNN and SVM machine learning classifiers produce relatively low values for accuracy, sensitivity, and specificity. The neural network classifier Multilayer Perception (MLP) is one of the most efficient ones. using categories data into sudden cardiac arrest or normal, time domain derived features from the ECG were applied using MLP.Implemented random survival forest classifier for patients' cardiac arrhythmia detection [14].Pan Tompkins algorithm is used to find HRV from preprocessed ECG, then machine learning classifiers like DT, SVM, KNN, and NB classifies data into normal or abnormal. Machine learning classifier model implemented for left ventricular hypertrophy detection in young adults [15].

Machine learning and deep learning algorithms leads present industry which depends on classification and regression issues. Bio medical industry mainly depends on artificial intelligence which identifies diseases based on image or signal received. Griet. et. al. [16] introduced QRS fragmentation as one of the methods to detect and quantify ECG by various machine learning classifiers, new methodology which rely on wavelet packet transform for recognizing impulsive cardiovascular illnesses. supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning are the four categories of machine learning classifiers. The top three most precise machine learning classifiers are KNN, SVM, and DT. Only 97.3% accuracy was obtained using the non-linear HRV feature-based sudden cardiac death prediction technique proposed by DT, KNN, and SVM [17]. Houshyarifar. et. al. [18] published a review on ECG analysis to predict coronary artery disorders.

The rate of sudden cardiac mortality has been rising rapidly over the past 20 years. Death occurs shortly after the patient's symptoms are identified. Currently, an embedded hardware module is required to recognise sudden cardiac death and diagnose cardiovascular illnesses. Even though there are many hardware modules described in the literature, effective hardware architectures were still needed to detect anomalies in the ECG signal. In this study, we suggest a brand-new hardware component for diagnosing cardiovascular conditions. Architecture's approaches for classifying diseases play a significant role in addition to hardware module methods. Early risk identification of cardiac patients using discrete wavelet transforms and nonlinear features using DT, SVM, and KNN classifiers. Supervised learning, semi-supervised learning, unsupervised learning, and reinforcement learning are the four categories of machine learning classifiers. The top three most precise machine learning classifiers are KNN, SVM, and DT. Only 97.3% accuracy was reached by the non-linear HRV feature-based sudden cardiac death prediction technique proposed by DT, KNN, and SVM[19].There are techniques based on genetic algorithms, fuzzy systems, sparse coding, and ensemble learning for automatically identifying cardiovascular disorders. The use of independent component analysis led to the proposal of a deep neural network for ECG biometrics [20]. Disease difficulties brought on by the human body are also covered. To characterize signal properties, multiple frequency transformations are available. A convolution neural network based on the short time Fourier transform (STFT) was introduced to detect arrhythmia in the human heart's ECG signal [21].

Bio medical industry mainly depends on artificial intelligence which identifies diseases based on image or signal received. J. Amezquita.et.al [22] presented new methodology which depends on wavelet packet transform for identifying impulsive cardiovascular diseases. Muhammed amin et.al. proposed drivers stress detection model using fuzzy systems and deep neural network [23]. Presented risk prediction model architecture to identify cardiac arrhythmias in patients by random survival forest [24]. Pulse transit time is the method to measure blood pressure level which shows severity of heart diseases in human.

In Optical remote sensing, wireless monitoring system has been implemented to monitor cardiovascular diseases of the patient with ECG obtained by sensors attached in body [25]. Machine learning methods, point care plots and recurrence plots are new estimation methods in biomedical signal processing for feature extraction, feature construction and feature classification. Methods bases on genetic algorithms, fuzzy systems, sparse coding and ensemble learning for automatic cardiovascular diseases identification were exist [26]. Pan Tompkins algorithm is used to find HRV from preprocessed ECG, then machine learning classifiers like DT, SVM, KNN, and NB classifies data into normal or abnormal [27].M Rizwan et.al. presented risk prediction model architecture to identify cardiac arrhythmias in patients by random survival forest [28].

ECG database taken from MIT BIH physio net website. Architecture for proposed method which detects since the work focuses on the cardiovascular diseases, reviews on the cardiovascular diseases advances in ECG are discussed in literature. The disease challenges resulting from human body are also discussed Prasad et al. [29] proposed wavelet and RR based neural network architecture to automatic classification of ECG arrhythmia with the help of multi resolution analysis. Pulse transit time is the method to measure blood pressure level which shows severity of heart diseases in human [30]. Review on ECG analysis to predict coronary artery diseases presented by R liew [31]. R. Sparapani et al. proposed a model for detection of Left ventricular hypertrophy by using regression tree algorithm [32].

Now a days may people are getting affected with brain tumor worldwide, it is not just limited to old age people but also found in the early age peoples also. A tumor developed inside the brain due to irregular growth of cells within the brain cranium. Sambath Kumar and Rajendran proposed a model for automatic brain tumor segmentation using neural network [33]. ECG based biometric is new trend for identifying attendance of particular person [34]-[36]. Cost of the module, area occupied by the overall module and time taken to execute function were problems in existing works. Robustness and sensitivity are also problems found in literature. But our proposed model only rectifies cost, area and executed time to find diseases. We try to overcome these drawbacks in future. Some problems found in existing works [37]. An electrocardiographic system presented for identifying left ventricular hypertrophy in young adults by using machine learning algorithms [38]**.** Biometric is attendance for several organization employees,Human health monitoring through simultaneous ECG and EEG signals are monitored to predict abnormalities of patient based on convolution neural network [39].

Discrete wavelet transforms and nonlinear features based early risk identification of cardiac patient by DT, SVM and KNN classifiers [40].Wavelet transform is efficient transform technique used in signal processing and image processing applications. Neural network and machine learning two different methods in Artificial intelligence for classification and regression applications. Neural network and wavelet transform based intelligent machine monitoring presented with sound signal processing [41**].** KNN classifier-based abnormality detection has been done and novel LMS algorithm based adaptive filter to eliminate noise and interference in ECG signal [42]. SVM classifier-based abnormality detection has been done and DENLMS algorithm based adaptive filter to eliminate noise and interference in ECG signal [43] Rapid technological advancements in health care is due to smart health care devices. Remote real time cardiac health monitoring by IOT based platform using tuned twin support vector machine [44].Machine learning methods, point care plots and recurrence plots are new estimation methods in biomedical signal processing for feature extraction, feature construction and feature classification. Statistical and machine learning methods presented for fault identification and also comparative analysis has taken [45]-[46]**.** Based on heart rate variability and spectral feature obtained from ECG, automatic atrial fibrillation detection has been identified [47].Virtual doctor module presented for identifying heart diseases. This module is useful for people lives in remote areas [48]-[49].

After detailed literature survey, we have identified various research gaps such as higher false positive rate, under fitting, disease misclassification and lack of real time implementation. We have considered disease misclassification which shows major issue for addressing cardiovascular disease abnormalities. We have proposed a innovative models for identifying heart diseases. MLR and DT machine learning classifiers are effective in identification of CVD

3. Proposed Methodology

3.1 Method 1

Proposed Methodology is an innovative idea and it is represented in terms of components is shown in Figure 1. The ECG signal is processed by various stages are Preprocessing, Feature Extraction, Feature selection and Machine learning classifier. In preprocessing stage ECG signal is filtered to eliminate noise. After preprocessing stage feature extraction stage is used to extract feature samples. Wavelet transform and Principal component analysis are used to apply to the preprocessed signal for feature extraction. From feature extraction features are selected for testing and training for classifying signal. Machine learning classifiers are random forest and Multinomial logistic regression used for detection and classification of cardiovascular abnormalities in ECG signal

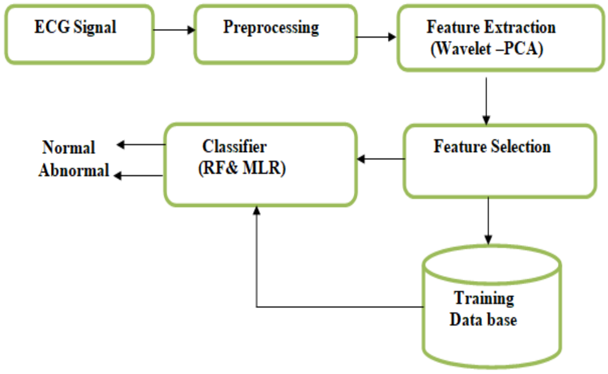

Fig .1. Block diagram for Detection and classification of cardiovascular Abnormalities

3.1.1Preprocessing

It is necessary to process ECG signal through Preprocessing stage to eliminate various types of noise that are presented in ECG signal. There are various types of noise are base line wander, power line interference and artifacts in muscle contraction. Baseline wander correction and band pass filtering operations are performed to eliminate all types of noise that are present in Signal

3.1.2Wavelet Transform

In time domain signal analysis is difficult but in frequency domain signal statistics can analysis is easy. The first basic frequency transform is Fourier transform (FT) is cannot used for analyzing heart beat signal because it is nonstationary in nature. Wavelet transform is best transform for feature extraction over other transforms like DFT, DCT and FFT etc. In our research Daubechies wavelet of order 8 and 3 level decomposition are chosen for feature extraction. Sampling frequency of the signal is 360 Hz and approximation level 3 wavelet transform bas.

3.1.3 Principal component analysis

Principal component analysis (PCA) is used for dimensionality reduction, visualization of high dimension data, feature extraction, noise filtering and feature selection. In our research wavelet features and temporal features are applied to PCA for feature selection and further feature extraction. In PCA with the help of ranker method correlation matrix can obtained. From correlation matrix various features are obtained, they are Eigen value, Eigen vector, proportion, cumulative and ranked attributes with min and max values. All extracted features are mentioned in table 1 are applied to machine learning classifier to assessment classification results for detection of cardiovascular Abnormalities in ECG signal.

3.1.4 Random Forest

Random forest is a machine learning classifier, which is formed by collection of decision trees. Datasets are divided into features and it is divided into number of trees. Each tree is categorized as separate class. Among different classes majority voting from classes are finalized as final class for classification. Random forest learning classifier working flow is shown in figure 2.

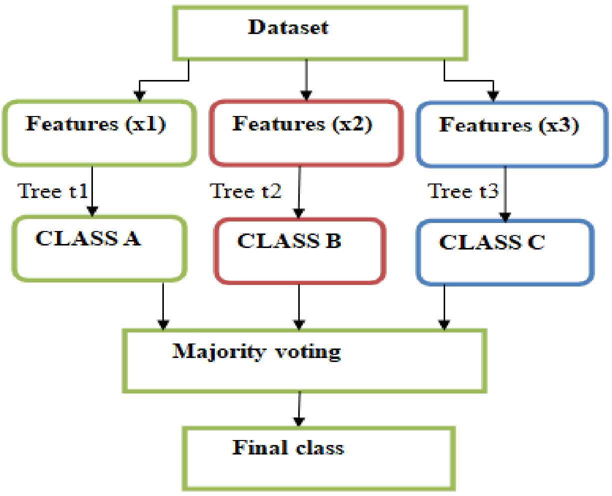

Fig.2. Random Forest classifier flow diagram

3.1.5 Multinomial Logistic Regression

Multinomial Logistic Regression (MLR) is a supervised machine learning algorithm. It is extension to the logistic regression which is used for predicting nominal dependent variable. Logistic Regression is two classes function whereas Multinomial Logistic Regression is a multi-class function. Multinomial Logistic Regression is also used for unordered categories. Steps in MLR classifier is shown below

ECG signal downloaded from physio net website [50]-[51]. Table 1 shows different class for dataset collected from public available database. There are 154 single lead ECG signal recording available in database. There are four different types of cardiovascular diseases labeled in table. Data portion for Normal sinus Rhythm is 12%, data portion for Supraventricular arrhythmia is 52%, data portion for ventricular tachyarrhythmia is 28% and data portion for Atrial Fibrillation is 13%.One sample reading for each class is shown in figure 3.

Table 1. Dataset class distribution

Class	Number of recordings	Portion
Normal sinusRhythm (N)	19	12.3%
Supraventricular arrhythmia (S)	80	51.9%
ventricular tachyarrhythmias (V)	35	22.7%
Atrial Fibrillation (A)	20	12.98%

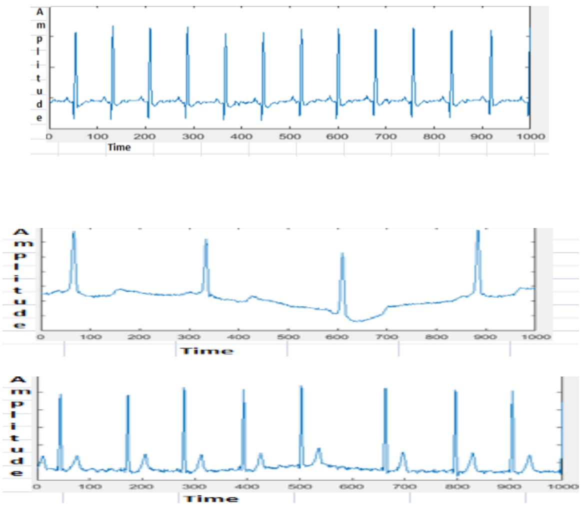

Fig.3. various class sample recording

3.2 Method 2

Our proposed work aim is to develop innovative and efficient method for identifying cardiovascular diseases is shown in figure 4. ECG signal consists of PQRST sample points. In this ECG, R peak gives significant information. Because most of diseases are identified with the help of R-peak information. Before finding R-peaks, ECG signal processed through pre processing stage to eliminate noise and interference from unwanted devices. R peak detection obtained by differentiation, squaring and moving window integration. HRV is another feature to differentiate normal and abnormal patient. HRV extracted from R peaks. Wavelet transform is also applied to the ECG to obtain various features like mean, standard deviation, variance, median L1-norm and L2-norm etc. These features are fused separated into training and testing phases. Training and testing data applied to various machine learning classifiers. Applied machine learning classifiers classifies received data into normal or abnormal subjects. We have tested four types of data classes, they are normal sinus rhythm, supra ventricular arrhythmia, ventricular tachy arrhythmias and atrial fibrillation. In this data class 70% data set used for training and remaining 30% dataset used for testing.The key factor for choosing NB, KNN and DT algorithms is only for accurate prediction and also it is shown in results.

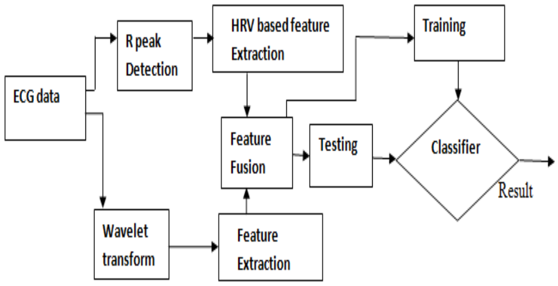

Fig.4. Peak detection- WT model

Figure 5 shows R-peaks identified from ECG and Heart rate variability determined from peaks. Figure 6 shows wavelet transform method applied to ECG data for feature extraction. The first five major diseases are Ventricular Flutter, Sinus Tachycardia, Sinus bradycardia, Atrial Flutter and tachycardia tested with our module. Threshold level for identifying these devices is different from one to another. They are 1) Ventricular Flutter is due to no visible P-wave presence, 2) Sinus Bradycardia is due to heart beating slower than normal beat, 3) Sinus Tachycardia is due to heart rate crossing more than 100 Beats per Minute (BPM),4)Atrial Flutter is due to multiple P-waves in the ECG and 5)tachycardia disease presents due to irregular heart beat

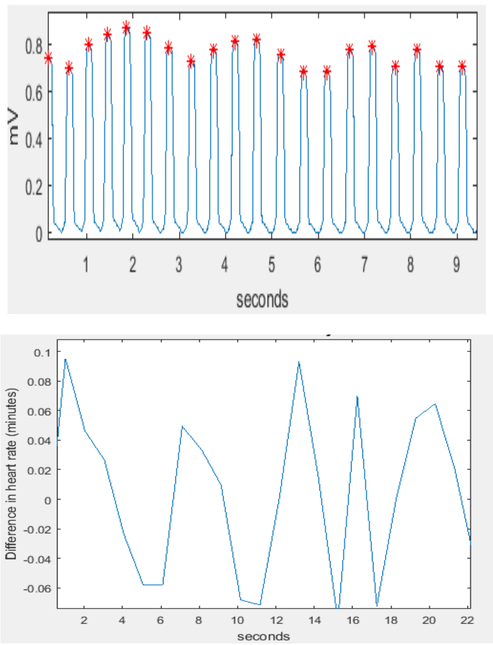

Fig.5. a) R –Peak detection b) HRV

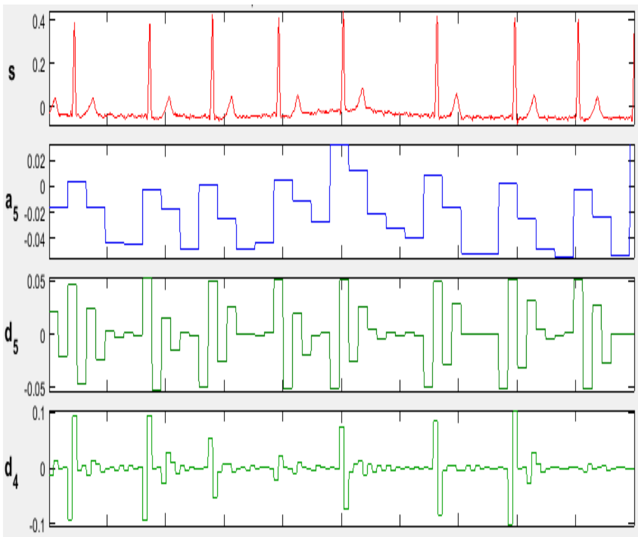

Fig. 6. Wavelets transform method for feature extraction

3.2.1 Machine learning classification

ECG signal is analyzed and abnormalities classified by machine learning classifiers. Classifier performance measured in terms of accuracy, sensitivity, specificity, precision and F-measure .In our proposed work, we have applied three machine learning classifiers such as KNN, DT and NB. Accuracy is key parameter to judge performance of machine learning classifier.

i. K-Nearest Neighbor (KNN)

KNN algorithm is one of the effective and simple machine learning classifier. Accurate classification of KNN algorithm is depends on k values because if k value varies accuracy changes. Based on training and testing data, minimum distance calculated for k-nearest neighbors.

ii. Decision Tree (DT)

Decision tree is machine learning algorithm that contains conditional control statements. It represents decisions of outcome visually and explicitly. Performance of DT depends only designed tree for accurate classification. It produces classification results based on disorder training samples.

iii. Naïve Bayes (NB)

Naïve bayes classifier is a machine learning algorithm works based on bayes theorem. This algorithm is a statically independent feature classification machine learning classifier. Bayesian additive regression trees are also machine learning classifier to detect left ventricular hypertrophy in cardiac diseases suffering Patients.

4. Results

Various machine learning open source tools are exist for artificial intelligence. Some of most popular tools are Tensor Flow, Keras, scikit-learn, PyTorch and Weka. In our research, we have chosen Weka tool for cardiovascular diseases abnormalities classification. Performance metrics are precision, sensitivity, specificity, F-measure and Accuracy. These key metrics are defined in terms of True positive (TP), True negative (TN), False positive (FP) and False negative (FN). Accuracy is important parameter to judge performance of machine learning classifiers.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

$$\text{F} - \text{measure} = 2\,\frac{\text{Precision.Recall}}{\text{Precision+Recall}} \quad (4)$$

Accuracy = (TP+TN)/ (TP+TN+FP+FN) (5)

Heart disease presents due to various factors.Ventricular Flutter is due to no visible P-wave presence,Sinus Bradycardia is due to heart beating slower than normal beat, Sinus Tachycardia is due to heart rate crossing more than 100 Beats per Minute (BPM),Atrial Flutter is due to multiple P-waves in the ECG andtachycardia disease presents due to irregular heart beat

Performance metrics varies for categorized features from two classification problems to three classification problems. If all features are applied for disease classification, the maximum accuracy shows 98.99% for two classification problem and 87.8% for three classification problem by DT classifier. If top 10 features are applied for disease classification, the maximum accuracy shows 99.1% for two classification problem and 78.4% for three classification problem by DT classifier. If top 5 features are applied for disease classification, the maximum accuracy shows 99.3% for two classification problem and 88.5% for three classification problem by DT classifier

Table 2. Performance comparison between proposed method with other classifiers

Literature	Year	Classifier	Feature/method	Accuracy
Can Ye	2012	SVM	Wavelet+ICA +RR	99.3
Felipe Alonso	2014	SVM	HOS	98.6
Sandeep Raj	2020	SVM	Hybrid+PSO	98.8
Venkatesan	2018	KNN	DWT with HRV	97.5
Venkatesan	2018	SVM	HRV	97.5
Jingshan Huang	2019	CNN	STFT	99
Prasad	2003	NN	Wavelet+RR	94
Salem	2022	DN	Naset model	97.49
Proposed	2023	MLR	Wavelet+PCA	99.6

Table 2 shows Accuracy comparison between proposed method with other literature techniques. In literature SVM classifier with various feature extraction methods exist and KNN classifier based discrete wavelet transform exist.

Table 3. Comparison of our method and other recent existing method

Author	Year	Methodology			
		Data type	Feature Extraction	Classifier	Acc (%)
Present research	2023	ECG	WT, HRV	DT	99.3
Ebrahimzadeh et.al.	2018	HRV	HRV, TD	MLP	90.18
M. Khazaei et.al.	2018	HRV	WPT,	KNN	95
Ebrahimzadeh et.al.	2019	HRV	HRV,TD	MLP	88.29
H.Fujita et.al.	2016	HRV	NL features	SVM	94.7
U. R. Acharya et.al.	2015	ECG	SCDI	DT	92.1

Neural network and convolution neural network classifiers with different feature extraction methods exist. Proposed wavelet-PCA based Multinomial logistic regression machine learning classifier produce high accuracy than other approaches. Existing authors uses MLP, SVM, FPNN algorithms for disease identification and classification. We have used wavelet transform- HRV based methodology for feature extraction. Existing author's uses wavelet packet transform, nonlinear HRV, entropy and time domain methods for feature extraction. Table 3 is performance comparison between proposed model and traditional model; it is clearly showing proposed model produces highest accuracy.This table shows comparison between proposed method and existing works. Based on ECG, HRV different diseases has different threshold level. For example tachycardia disease presents due to irregular heart beat (RR-segment). Irregular heart beat also causes chest pain .Our proposed work shows better results than existing work. In this table material described in terms of data type and length of the signal, methodology described in terms of feature extraction and classifier. Best performance is listed between our work and existing work from 2015 to 2019. Accuracy of proposed work is 99.3 % which is higher than other existing works. Our proposed work tested with KNN, NB and DT machine learning classifiers but DT algorithm produces highest accuracy. We have compared our research work with existing work and we have achieved highest accuracy [52]-[55]. The proposed technique is concluded by obtained results software simulations. It is observed that maximum accuracy has been achieved using DT and MLR machine learning classifier which is higher than other machine learning algorithms. Features combined from wavelet transform and HRV are applied to DT which produces highest accuracy 99.3. Features combined from wavelet transform and PCA are applied to MLR which produces highest accuracy 99.6.

5. Conclusion

In this research Paper, An Impulsive cardiac death (ISD) risk identification by machine learning approach based on heart rate variability and wavelet transform is proposed for cardiovascular diseases monitoring. Extracted features from HRV and wavelet transform are combined and applied to machine learning classifiers. The proposed method uses machine learning algorithms naïve bayes, Decision tree and k nearest neighbor (KNN) for identifying abnormalities in ECG of the patient. Accuracy varies from 97.46% to 99.3% for two classification problem (N versus A) and 86.76% to 88.5% for three classification problem (N versus S versus A). Our proposed method shows improved accuracy than other existing methods. The proposed work with machine learning classifiers is concluded by obtained results with high accuracy and good specificity for identifying cardiovascular diseases. In this research paper various challenges are faced for selection of appropriate morphological features, spectral features and wavelet features. This paper also presents a new features extraction implemented method to contribute more to accuracy. Most effective 32 features are chosen with Multinomial Logistic regression classifier and Random Forest in our proposed method, 15 features are newly extracted features to classify better performance for real time biomedical application. The proposed method is concluded by obtained results of Multinomial logistic regression classifier is better than random forest because MLR achieved 99.6% accuracy higher than other literature methods like SVM, KNN, CNN and also including random forest classifier. Proposed method is planning to implement in hardware can be integrated with IOT module in future

References

[1] Abbasi, H., Bennet, L., Gunn, A. J., and Unsworth, C. P., "Robust wavelet stabilized footprints of uncertainty for fuzzy system classifiers to automatically detect sharp waves in the EEG after hypoxia ischemia" International Journal of Neural Systems 27(3):1650051, 2017 .

[2] Adeli, H., and Hung, S. L., "Machine learning - neural networks, genetic algorithms, and fuzzy systems, " New York: Wiley, 1995.

[3] Ahmadlou, M., and Adeli, H., "Enhanced probabilistic neural network with local decision circles: A robust classifier, " Integrated Computer-Aided Engineering 17(3):197–210, 2010.

[4] A. Rahman, L. G. Tereshchenko, M. Kongkatong, T. Abraham, M. R. Abraham, and H. Shatkay, ''Utilizing ECG-based heartbeat classification for hypertrophic cardiomyopathy identification,'' IEEE Trans. Nanobiosci., vol. 14, no. 5, pp. 505–512, Jul. 2015.

[5] Chengyu Li ,Xiangyu Zhang, Lina Zhao, Feifei Liu, Xingwen Chen, Yingjia Yao and Jianqing Li, "Signal Quality Assessment and Lightweight QRS Detection for Wearable ECG SmartVest System" IEEE Internet of Things Journal, DOI 10.1109/JIOT.2018.2844090.

[6] C. H. Hung, Y.-W. Bai, and R.-Y. Tsai, "Design of blood pressure measurement with a health management system for the aged," IEEE Trans. Consum. Electron., vol. 58, no. 2, pp. 619–625, May 2012

[7] DakunLai , Yifei Zhang , Xinshu Zhang , Ye Su , And Md Belal Bin Heyat , "An Automated Strategy for Early Risk Identification of Sudden Cardiac Death by Using Machine Learning Approach on Measurable Arrhythmic Risk Markers" IEEE Access, Volume 7, 2019

[8] Ebrahimzadeh, E., and Pooyan, M., "Early detection of sudden cardiac death by using classical linear techniques and time-frequency methods on electrocardiogram signals". Journal of Biomedical Science and Engineering 4(11):699, 2011.

[9] E. Ebrahimzadeh., Pooyan, M., and Bijar, A., "A novel approach to predict sudden cardiac death (SCD) using nonlinear and time frequency analyses from HRV signals. PloS one 9(2):e81896, 2014

[10] E. Ebrahimzadeh, A. Foroutan, M. Shams, R. Baradaran, L. Rajabion, M. Joulani, and F. Fayaz, ‘‘An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from HRV signal,’’ Comput. Methods Programs Biomed., vol. 169, pp. 19–36, Feb. 2019.

[11] E. Ebrahimzadeh, M. S. Manuchehri, S. Amoozegar, B. N. Araabi, and H. Soltanian-Zadeh, ‘‘A time local subset feature selection for prediction of sudden cardiac death from ECG signal,’’ Med. Biol. Eng. Comput., vol. 56, no. 7, pp. 1253–1270, 2018

[12] Felipe Alonso-Atienza∗, Eduardo Morgado, Lorena Fern´andez-Mart´ınez, Arcadi Garc´ıa-Alberola, and Jos´e Luis Rojo-´Alvarez, ” Detection of Life-Threatening Arrhythmias Using Feature Selection and Support Vector Machines” IEEE Transactions On Biomedical Engineering, Vol. 61, No. 3, March 2014

[13] Fen Miao ,Xurong Wang, Liyan Yin and Ye Li ,” A Wearable Sensor for Arterial Stiffness Monitoring Based on Machine Learning Algorithms”, IEEE sensors, DOI 10.1109/JSEN.2018.2880434

[14] F. Miao, Y.P. Cai, Y.X. Zhang, Y. Li, and Y.T. Zhang. "Risk prediction of one-year mortality in patients with cardiac arrhythmias using random survival forest." Computational and mathematical methods in medicine 2015 (2015).

[15] Gen-Min Lin , and Kiang Liu, “An Electrocardiographic System With Anthropometrics via Machine Learning to Screen Left Ventricular Hypertrophy among Young Adults” Medical Imaging And Diagnostic Radiology, volume 8,2020

[16] Griet Goovaerts ,SibasankarPadhy , Bert Vandenberk, Carolina Varon , Rik Willems, and Sabine Van Huffel,” A Machine-Learning Approach for Detection and Quantification of QRS Fragmentation” IEEE Journal Of Biomedical And Health Informatics, Vol. 23, No. 5, September 2019

[17] G. H. Tison, J. Zhang, F. N. Delling, and R. C. Deo, ‘‘Automated and interpretable patient ECG profiles for disease detection, tracking, and discovery,’’ 2018, arXiv:1807.02569. [Online]. Available: http://arxiv.org/abs/1807.02569

[18] Houshyarifar V, Amirani MC. Early detection of sudden cardiac death using Poincaré plots and recurrence plotbased features from HRV signals. Turkish Journal of Electrical Engineering & Computer Sciences 2017;25 (2):1541–53

[19] H. Fujita, U. R. Acharya, V. K. Sudarshan, D. N. Ghista, S. V. Sree, L. W. J. Eugene, and J. E. W. Koh, ‘‘Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index,’’ Appl. Soft Comput., vol. 43, pp. 510–519, Jun. 2016.

[20] Jae-Neung Lee1 And Keun-Chang Kwak ”ECG-based biometrics using a deep network Based on Independent Component Analysis”, IEEE Acsess, volume 10, 2022, pp. 12913-12926.

[21] Jingshan Huang, BinqiangChen , Bin Yao and Wangpeng He,” ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network” IEEE Access, Volume 7, 2019

[22] J. Amezquita-Sanchez, M. Valtierra-Rodriguez, H. Adeli, and C. A. Perez-Ramirez, ‘‘A novel wavelet transform-homogeneity model for sudden cardiac death prediction using ECG signals,’’ J. Med. Syst., vol. 42, no. 10, p. 176, 2018

[23] J. J. Oresko et al., “A wearable smartphone-based platform for realtime cardiovascular disease detection via electrocardiogram processing,” IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 3, pp. 734–740, May 2010

[24] J. R. Quinlan, ‘‘Simplifying decision trees,’’ Int. J. Man-Mach. Stud., vol. 27, no. 3, pp. 221–234, Sep. 1987.

[25] Muhammad Amin, Khalil Ullah, Muhammad Asif1 , Abdul Waheed, Sana Ul Haq , Mahdi Zareei, R.R. Biswal, “ECG-Based Driver’s Stress Detection Using Deep Transfer Learning and Fuzzy Logic Approaches” ”, IEEE Acsess, volume 1, 2022, pp. 1-24.

[26] Murugappan M, Murukesan L, Iqbal O, Sabira K. Time domain features based sudden cardiac arrest prediction using machine learning algorithms. Journal of Medical Imaging and Health Informatics 2015;5(6):1267–71.

[27] M. Khazaei, K. Raeisi, A. Goshvarpour, and M. Ahmadzadeh, ‘‘Early detection of sudden cardiac death using nonlinear analysis of heart rate variability,’’ Biocybernetics Biomed. Eng., vol. 38, no. 4, pp. 931–940, 2018.

[28] M. Rizwan, B. M. Whitaker, and D. V. Anderson, ‘‘AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning,’’ Physiol. Meas., vol. 39, no. 12, Dec. 2018, Art. no. 124007

[29] Prasad G K and J.S.Sahambi, “Classification of ECG arrhythmias using multi-resolution analysis and neural networks,” in Proc.Conf.Convergent

Technol. Asia-Pacific Region, Oct. 2003, pp. 227–231
[30] R. Jegan And K. V. Anusuya , “Heart signal monitoring for remote patient assistant based on optical remote sensing” Optoelectronics And Advanced Materials – Rapid Communications Vol. 9, No. 5-6, May – June 2015, p. 570 – 574
[31] R. Liew, ‘‘Electrocardiogram-based predictors of sudden cardiac death in patients with coronary artery disease,’’ Clin. Cardiol., vol. 34, no. 8, pp. 466–473, 2011
[32] R. Sparapani et al., ‘‘Detection of left ventricular hypertrophy using Bayesian additive regression trees: The MESA (Multi-Ethnic Study of Atherosclerosis),’’ J. Am. Heart Assoc., vol. 8, no. 9, May
[33] Sambath Kumar, K. and Rajendran, A. ‘An Automatic Brain Tumor Segmentation Using Modified Inception Module Based U-Net Model’ Journal of Intelligent & Fuzzy Systems , 1 Jan. 2022 : 2743 – 2754.2019.
[34] Sandeep Raj , ” An Efficient IoT-Based Platform for Remote Real-Time Cardiac Activity Monitoring” IEEE Transactions On Consumer Electronics, Vol. 66, No. 2, May 2020
[35] Song-KyooKim , Chan Yeob Yeun , Ernesto Damiani1 and Nai-Wei Lo , “A Machine Learning Framework for Biometric Authentication Using Electrocardiogram, ” IEEE Access, Volume 7, 2019
[36] S.-B. Kim, K.-S. Han, H.-C. Rim, and S. H. Myaeng, ‘‘Some effective techniques for Naive Bayes text classification,’’ IEEE Trans. Knowl. Data Eng., vol. 18, no. 11, pp. 1457–1466, Nov. 2006
[37] S. Raj, K. Maurya, and K. C. Ray, “A knowledge-based real-time embedded platform for arrhythmia beat classification,” Biomed. Eng. Lett., vol. 5, no. 4, pp. 271–280
[38] S. Raj, G. Chand, and K. C. Ray, “Arm-based arrhythmia beat monitoring system,” Microprocess. Microsyst., vol. 39, no. 7, pp. 504–511, 2015
[39] T. Denoeux, ‘‘A k-nearest neighbor classification rule based on Dempster– Shafer theory,’’ IEEE Trans. Syst., Man., Cybern., vol. 25, no. 5, pp. 804–813, May 1995.
[40] U. R. Acharya, H. Fujita, V. K. Sudarshan, V. S. Sree, L. W. J. Eugene, D. N. Ghista, and R. S. Tan, ‘‘An integrated index for detection of sudden cardiac death using discrete wavelet transform and nonlinear features,’’ Knowl.-Based Syst., vol. 83, pp. 149–158, Jul. 2015
[41] Vahid Nasir , Julie Cool, and Farrokh Sassani, “Intelligent Machining Monitoring Using Sound Signal Processed With the Wavelet Method and a Self-Organizing Neural Network” IEEE Robotics And Automation Letters, VOL. 4, NO. 4, OCTOBER 2019
[42] Venkatesan C, P. Karthigaikumarand R. Varatharajan, “A novel LMS algorithm for ECG signal preprocessing and KNN classifier based abnormality detection” Multimed Tools Appl, February 2018
[43] Venkatesan C, P. Karthigaikumarand R. Varatharajan, “ECG Signal Preprocessing and SVM Classifier-Based Abnormality Detection in Remote Healthcare Applications” IEEE Access, Volume 6,2018
[44] WangpengHe , Geng Wang , Jie Hu , Cheng Li, Baolong Guo , And Fengping Li, “Simultaneous Human Health Monitoring and Time-Frequency Sparse Representation Using EEG and ECG Signals” IEEE Access, Volume 7, 2019
[45] X. Ding, Y. Zhang, and H. K. Tsang, "Impact of heart disease and calibration interval on accuracy of pulse transit time-based blood pressure estimation, " Physiological Measurement, vol. 37, no. 2, pp. 227-237, 2016.
[46] Y. Suresh, L. Kumar, and S. K. Rath, ‘‘Statistical and machine learning methods for software fault prediction using CK metric suite: A comparative analysis,’’ ISRN Softw. Eng., vol. 2014, pp. 1–15, 2014
[47] ZhenningMei , Xiao Gu1 , Hongyu Chen2 , And Wei Chen , ” Automatic Atrial Fibrillation Detection Based on Heart Rate Variability and Spectral Features” IEEE Access, Volume 6, 2018
[48] Venkataramanaiah, B. et.al, “Efficient prediction of heart diseases using mahhine learning classifiers”, IEEE Xplore,2022 **DOI:** 10.1109/ICTACS56270.2022.9988582
[49] Venkataramanaiah, B., Kamala, J. ECG signal processing and KNN classifier-based abnormality detection by VH-doctor for remote cardiac healthcare monitoring. *Soft Comput* **24,** 17457–17466 (2020).
[50] MIT/BIH-NSR (2018), https://www.physionet.org/physiobank/ database/nsrdb/, database accessed January 1, 2018.
[51] MIT/BIH-SCDH (2018), https://physionet.org/physiobank/ database/sddb/#clinical information/databased accessed January 1, 2018.

[52] Salem Nermin et.al ,"Improved Bat Algorithm with Deep Learning-Based Biomedical ECG Signal Classification Model"Computer and materials,2022

[53] [Cheng,YC, Hartman, T, Tsai, PY.et al., "Population based ant colony optimization for reconstructing ECG signals", Evol.intel.9,55-66,2016

[54] Firmino,J.V, Melo, M., Salemi, V.et al, " Heart failure recognition using human voice analysis and artificial intelligence" Evol.intel,2023

[55] Aggarwal, R., Kumar, S, "Classification model for meticulous presaging of heart disease through NCA using machine learning", Evol.Intel,2023