Biomedical Signal Processing and Control 68 (2021) 102692

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Detection of Ventricular Arrhythmia by using Heart rate variability signal and ECG beat image

Saurav Mandal *, Pulak Mondal, Anisha Halder Roy

Department of Radio Physics and Electronics, University of Calcutta, Kolkata, West Bengal, India

ARTICLE INFO

Keywords:
Ventricular Arrhythmia
Ventricular Tachycardia
Ventricular Fibrillation
Heat Rate Variability
ECG beat image
Cardiac scoring
Ensemble classifier
Computer-aided diagnosis system

ABSTRACT

Ventricular Arrhythmia (VA) such as Ventricular Tachycardia (VT) and Ventricular Fibrillation (VF) are the common type of arrhythmia in infants and children. Electrocardiogram (ECG) signal is used for the diagnosis of such type of cardiac abnormality. Manual ECG assessment is an error prone task because of vast difference in ECG morphology. Hence, a computer aided diagnosis (CAD) system for classification of cardiac abnormality can be useful in cardiac care units. Main goal of our study is to design a CAD system to classify ECG signals of VT and VF patients. This study proposed a technique for the process of cardiac scoring which is a significant diagnosis step in VA detection. ECG signals are used for the extraction of Heart Rate Variability (HRV) signals and ECG beat images. The features are computed from the HRV signals and ECG beat images using thirty different feature extraction methods. Skewness, Kurtosis, shape factor, fractal dimension, entropy, contrast and dissimilarity are the selected features used for classification process. The best classification accuracy has been achieved by the use of ensemble classifier. According to performance analysis, the system can be operated with 99.99% accuracy rate for the separation between healthy and VA persons.

1. Introduction

Cardiac arrhythmia is the irregular rhythm of heart which is occurred when electrical impulses of the normal heart beats are disrupted [1]. Among the many cardiac arrhythmias, most common type of arrhythmia is Ventricular Arrhythmia (VA) which is occurred in people with a heart-related problem, such as scars or damage from a heart attack [2]. VA such as Ventricular tachycardia (VT) and Ventricular fibrillation (VF) are dangerous rhythm and it is responsible for sudden cardiac death of a person with heart problems [3]. It needs early medical attention as the progression of VA may lead to long term paralysis and stroke. COVID-19 has the potential to cause of VA with at least 17% found to have an elevated troponin and 23% noted to have heart failure in a study of 191 inpatients from Wuhan, China [4]. It can be resolved with a proper treatment [63].

1.1. Motivation of the work

ECG signals are used for clinical examination of cardiac disease. Clinicians find difficulties to assess cardiac condition of the patient from ECG signal because of the vast difference in ECG morphology [5,6]. Therefore, automatic detection and diagnosis of VA is easy approach for identifying the severity of the disease. Motivation of our work is to overcome the limitations of double reading by another radiologist in different hospitals and reduce the false negative rate of screening ECG signals. Double reading cannot be widely adopted due to cost-effectiveness and practicality in many countries [7]. Efficient feature selection process as well as efficient classifier are required for the improvement of true positive accuracy of the system. Computer aided diagnosis system is an alternative of double reading and to improve true positive accuracy for the diagnosis of VA patients.

1.2. State-of-the-art

Numerous techniques like variational mode decomposition, empirical mode decomposition, adaptive variation mode decomposition have been developed for computer aided disease detection system in last two decades [53–58]. ECG signals primarily require several processing steps like pre-processing, segmentation, feature extraction and classification. Preprocessing of raw ECG signal is performed for the removing of outliers and noise from the ECG signal. Mar et al. [8] have proposed Daubechies wavelet transform for the removal of noise from the ECG signal.

* Corresponding author.
E-mail addresses: souravsourav.cu@gmail.com (S. Mandal), pmrpe@caluniv.ac.in (P. Mondal), halder.anisha@gmail.com (A.H. Roy).

https://doi.org/10.1016/j.bspc.2021.102692
Received 15 October 2020; Received in revised form 18 April 2021; Accepted 29 April 2021
Available online 19 May 2021

Author in [9] used Kalman Filter [10] model to improve signal to noise ratio. In Kandala et al. [11] high-pass and low pass Butterworth filters have been used with 1 Hz and 45 Hz cut-off frequency for the separation of noisy ECG signal. Outlier [12] is the data points that appears to deviate from the other member of the samples. In [13] Khoma et al. have proposed window outlier correction algorithm for the ECG signal processing. This approach improved the model performance after removing the outliers from the ECG signal. The morphological changes of heart beats are occurred due to heart related problems [5]. Segmentation of heart beat is a crucial approach for the classification of heart diseases [14]. ECG signal is segmented by the use of average length difference between two successive R-R waves [15]. The change of RR interval is observed due to irregular rhythm of heart that indicated abnormal cardiac activity [16]. One of the popular single lead first derivative based ECG beat detection method is the Hamilton-Tompkins algorithm [17]. Benitez et al. [18] have proposed a Hilbert transform based method for the improvement of first derived based ECG beat detection method. The method in [18] needs human intervention for the determination of the patient-specific threshold. Traditional transformation techniques like discrete wavelet transform, Fourier transform, empirical mode decomposition, stockwell transform are used to extract features from the segmented ECG data [17–20]. Features are categorized into temporal [21], morphological [22] and wavelet transform domain [23]. Clinically based tree scheme model has been proposed by Perlman et al. [24] using 12 lead ECG signal for the detection of VA patients. This model has been achieved 95.08% accuracy. Ortn et al. [25] have proposed automated real time method for VA beat detection. This method is faster than the previous method [25] but the accuracy decreased to 94.3%. Performance of the classifier depends on the quality of extracted features. ECG signals are non-stationary in nature [26]. Therefore, more complex methods are needed for the features extraction from non-stationary time series data. Convolution neural networks (CNNs) can be used for extraction of high quality features from ECG signals [27–29]. Acharya et al. [30] and Baloglu et al. [31] have proposed a deep learning model using CNN for the diagnosis of cardiac arrhythmia and these models are insensitive to noise. Extracted mixed feature sets are given to the input of classifier for the discrimination of respective classes of ECG signals. Numerous classifiers like Support Vector Machine (SVM) [33], Artificial Neural Networks (ANN) [34], Pattern Classifier [35], Random Forest (RF) classifier [37], Bayesian Approach [36] are used to discriminate VA, congenital heart disease, coronary artery disease, myocardial infarction using ECG signals analysis. Mondejar-Guerra et al. [51] have proposed VA beat classification method by the use of temporal and morphological features of ECG waveform based on multiple SVM classifiers. The model has achieved 85% accuracy for the detection of arrhythmia patient.

1.3. Contribution

Although many researchers already have proposed different VA detection technique from the ECG signals but still it is a challenging work for wearable sensor application for cardiac health monitoring system. Challenges like a CAD system need to be developed with less number of features but with satisfactory classification accuracy to detect VA disease. This study developed a machine learning based VA detection algorithm using cardiac score technique. Moreover, our proposed method does not require large number of data to train the model. On the other hand, our proposed disease detection system with seven features has given reasonable satisfactory classification accuracy. This literature focuses on the removing of powerline interface, baseline wander as well as outliers to rule out the noisy segments from the decision system and enhance the performance of the CAD system. In this work features are extraction from HRV signals as well as ECG beat images without pretained CNN architecture model. Gray-Level Co-Occurrence Matrix (GLCM) description [32] is used to extract texture feature for the ECG beat classification. HRV signal is analyzed effectively to capture the pattern changes in heart rate for adding better distinction between VA and normal heart condition.

Main contributions of this work are:

- ECG beat image and HRV signal together have been used for feature extraction.
- Cardiac scoring technique is used for disease detection.

The next sections of the literature are organized as follows.

Section 2 presents the methodology to detect VA from the ECG signal. This section emphasizes the technique for feature extraction using ECG signals. VA classification techniques using machine learning algorithms are shown in Section 3. In Section 4, results of our experiment are discussed. Finally, Section 5 presents the conclusions of the literature.

2. Methodology

The ECG signals used in this study were taken from physiobank public database [38]. Table 1 shows the details of ECG data for Normal, VT and VF patients.

The main steps of our study are shown in Fig. 1. Firstly, the ECG signals are collected from the physiobank public database [38]. Then the collected raw ECG signals are used for preprocessing to remove the unwanted noise signals. HRV signals are extracted from the preprocessed ECG signals. On the other hand ECG beat is extracted from the preprocessed ECG signal. In this stage ECG signals are converted to image data and the ECG image is used for ECG beat extraction process. After that feature extraction process is carried out on the extracted ECG beat as well as HRV signals. Next is the feature selection process. In this stage, cardiac score feature selection algorithm is used to select efficient features from the extracted features of ECG beats and HRV signals. In the last stage, classifier is used for decision process. The proposed CAD system performance is tested using the performance measure parameters.

2.1. Preprocessing

Maximum time ECG signals are corrupted by noise signal [39]. The common source of noise is power line interference and baseline wander [40]. The difference of frequency spectrum of the individual ECG signal has been found in the range of 1 Hz to 20 Hz. Power line interference of 50 or 60 Hz is the caused by the electromagnetic interference by power line during the recording of ECG. Baseline wander is a low frequency artifact in the ECG signal. Baseline wander is caused by the patient movement. We have applied same high pass filter configuration mentioned in other studies [41,12]. For removing the baseline wandering, we have applied highpass filter with 1 Hz cutoff frequency. A band-stop filter is used with 50 Hz to 60 Hz cutoff frequency to eliminate power line interface. Outliers due to high frequency noise signal also reduces the classifier performance [42]. In this work, we have used ℓ=200 ms median filter. The main advantage of the median filter is that it removes outliers while no phase distortion. Infinite impulse response filter and finite response filter are not effective to remove outliers from

Table 1
Details of ECG data used in this study.

Type of database	Male	Female	Duration
Normal (NSRDB)	12	6	30 min
VT (CUDB)	18	12	30 min
VF (VFDB)	12	10	30 min
Age	50 ± 06	51 ± 06	
Weight	100 ± 10	80 ± 20	
Height	160 ± 05	170 ± 06	
BMI	38.46 ± 04	34.1 ± 02	
AHI	8.52 ± 07	20.8 ± 10	

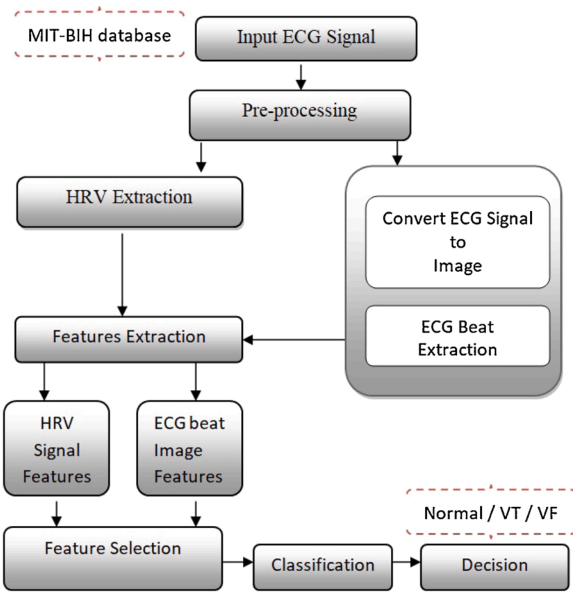


Fig. 1. Flow graph of ECG signal classification.

the signal even with the use of higher order filter structure [13]. Median filter with window length ℓ is applied to the ECG signal x and for step step in the window the median value is returned. Mathematical expressions are as follows: for odd number of window length

$$x(i-(l-1)/2):x(i+(l-1)/2) \tag{1}$$

for even number of window length

$$x(i-l/2), x(i-(l/2)+1), \ldots, x(i+(l/2)-1) \tag{2}$$

2.2. HRV signal extraction

After filtering the ECG signal, we have applied segmentation procedure for the extraction of HRV signal. In segmentation process, our aim is to detect R-wave from the ECG waveform. After segmentation process HRV signals are extracted from the waveform. The position of the R-wave is estimated by utilising Pan and Tompkins algorithm [43]. Further, the heart rate variability (HRV) signal is formed by calculating the RR waves from 4 min segmented ECG signal. Now HRV signal is ready for the next stage feature extraction to evaluate information on the cardiac condition of normal and VA patient. Next section describes the image segmentation process.

2.3. ECG beat extraction

In this work, Features like dissimilarity, shape factor, contrast are required for the VA diagnosis process. Therefore time series signal has been converted to an image of ECG signal. In this process, ECG beat is extracted from the image of ECG signal. The ECG beat extraction process is performed in two steps. These are as follows:

- Convert ECG signal to image data:

We have applied median filter for the denoising of ECG signal. ECG signal has been cropped from the position R1-125 to R2 +124 (R1 and R2 are two consecutive R peaks [43] of the ECG signal) for the conversion of that signal into an image. Fig. 2 shows ECG signal images of normal, VT and VF patients. The images of this diagram are saved as JEPG format. In this work, we have collected 18 subjects of NSR patients ECG signal images, 30 subjects of VT patients and 22 subjects of VF patients ECG signal images.

- Image data to ECG beat extraction:

In this work, a beat is constructed by considering signal segment from R-125 to R+124. Hybrid technique [44] is used to isolate one ECG beat from the ECG signal image. This technique is the combination of boundary detection [45] and region growing image segmentation methods [46,47]. Region growing approach is used to extract region of interest from the image. The region growing method is based on the Seed

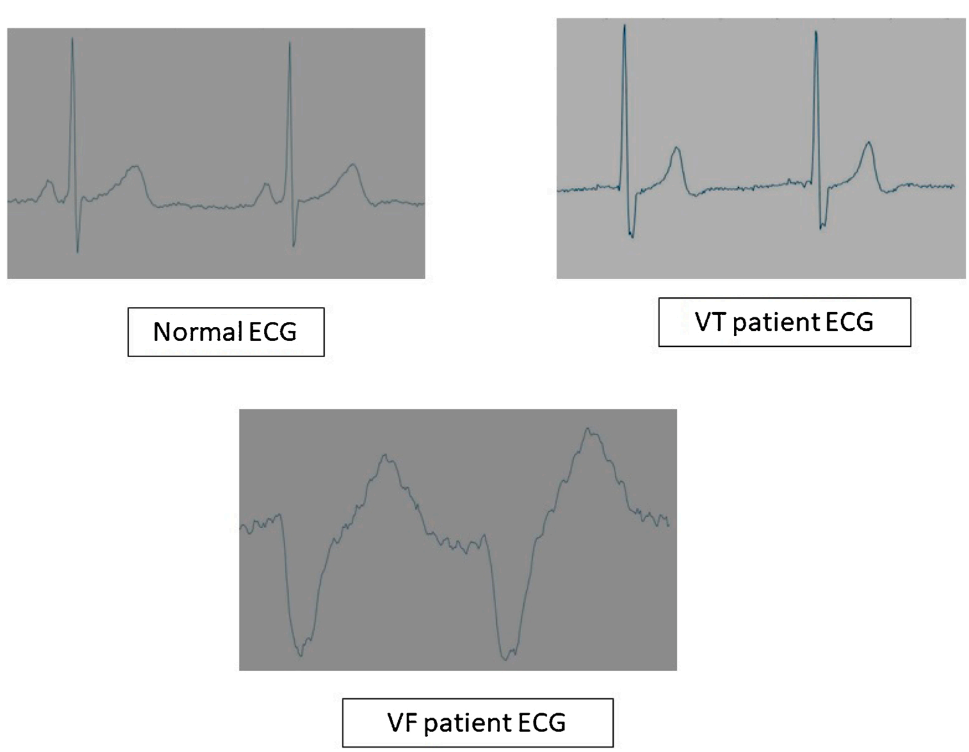


Fig. 2. ECG Signal image of Normal,VT and VF patients.

Region Growing (SRG) algorithm [65]. The centroid between the two adjacent edge regions is taken as an initial seed point. The region growing starts from the initial seed point and connected with all the pixels with same intensity. Fig. 3 shows that the technique is capable to isolate one ECG beat from the ECG signal. To achieve this aim, code is written in MATLAB software. Normalisation of each beat is done using min to max normalization procedure (range −1 to +1).

After segmentation, threshold method is used to convert gray scale image into binary image. Fig. 4 shows segmented output of normal, VT and VF subjects ECG beat images.

After that the segmented ECG beat images are used for next stage features extraction process.

2.4. Feature extraction

Feature extraction is a process of representation a high volume of data with in a limited number of samples. These samples are called features. Efficient features which have low *P*-value (<0.0001) [64] that enhance the system performance [59]. Feature captures the morphology of the signal and reveal the hidden information that lies in the ECG signal. In this study, two procedures are followed for features extraction.

- Features extraction using HRV signal
- Features extraction using ECG beat image

After extraction of HRV signals, 15 features extraction methods are applied to capture the information of morphological changes in ECG signals. 15 features extraction methods(Method 16 to Method 30) are shown in Table 3. HRV signal features are used for next stage feature selection process.

After the image segmentation of ECG beat, second order statistical texture features are extracted by the use of GLCM description. Processed binary images of ECG beats are used for GLCM feature extraction. GLCM are stored in $i*j*n$ matrix, where i is the reference pixel and neighbour pixel j in various orientation for n number of GLCMs. It is used to capture the information about structural arrangement of surfaces and their relationships with surrounding. GLCM features extraction methods (Method 1 to Method 15) are shown in Table 2. Extracted GLCM features are used for next stage feature selection process. In this study, total 30 features are extracted from the two procedures.

Each method has been applied to extract value of features from the 18 normal patients HRV signals and ECG beat images. Same procedures are applied for features extraction from the 30 VT patients and 22 VF patients data sets. Features extraction steps are shown in Fig. 5. Next stage is feature selection process. In this literature, we have proposed cardiac score feature selection technique to select efficient features.

2.5. Feature selection using cardiac score

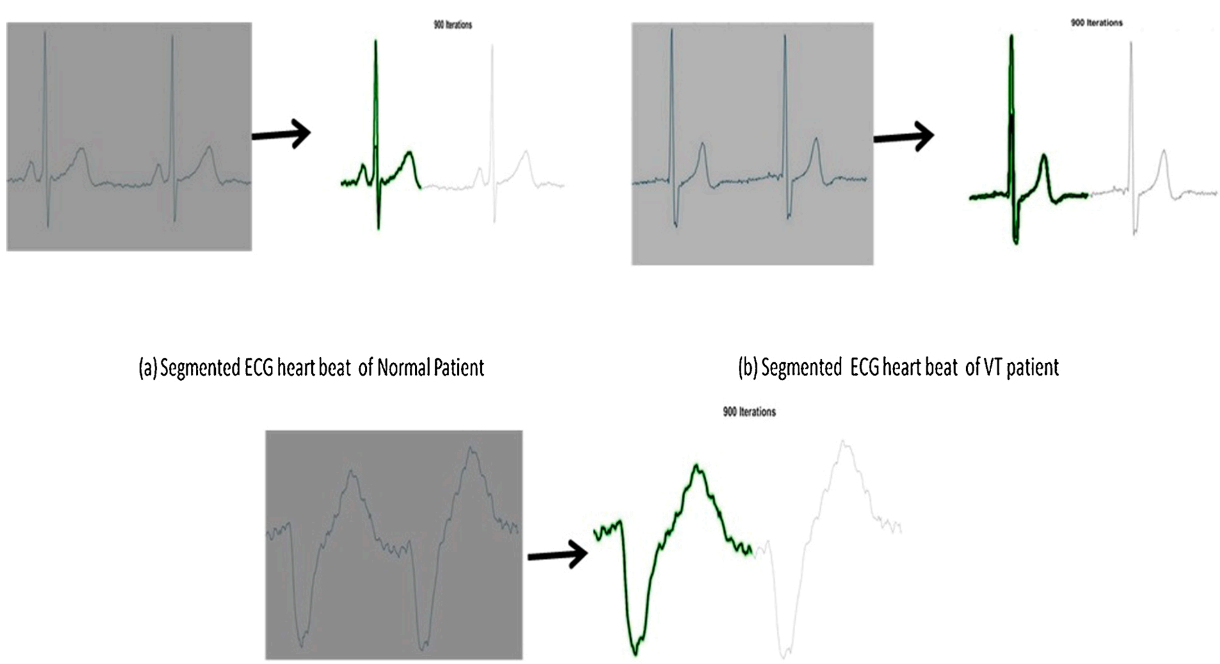
A large number of features are extracted using different feature extraction methods. All of the features are not useful for classification. To increase classification accuracy, it is necessary to reject irrelevant features from the large number of extracted features. Therefore, we have used Cardiac-score feature selection algorithm to reduce number of features.

Using Eq. (3), we have calculated average distance between center location and each class label for each possible combination.

$$P(m_k) = \sum_{1\leq a,b\leq c}\left(\frac{S_a+S_b}{N}\right)\left(\overline{\mathrm{m}_a^k}-\overline{\mathrm{m}_b^k}\right)^2 \quad (3)$$

Where a and b denote class types. a and $b = 1, 2, 3, 4, \ldots, c$ class, N denotes number of data samples, S_a represents the number of samples in a class and S_b represents the number of samples in class b. $\overline{\mathrm{m}_a^k}, \overline{\mathrm{m}_b^k}$ denote mean of class a and b for all the features m_k. Class scatter function $S(m_k)$ is calculated for all the features of m_k.

$$S(m_k) = \frac{\frac{1}{S_a}\sum_{l=1}^{m_a}\left((m_a^k)_l-\overline{\mathrm{m}_a^k}\right)^2 - min_{l\leq S_a}(m_a^k)_l - \left(\overline{\mathrm{m}_a^k}\right)^2}{max_{l\leq S_a}(m_a^k)_l - \left(\overline{\mathrm{m}_a^k}\right)^2 - min_{l\leq S_a}(m_a^k)_l - \left(\overline{\mathrm{m}_a^k}\right)^2} \quad (4)$$



(a) Segmented ECG heart beat of Normal Patient

(b) Segmented ECG heart beat of VT patient

(c) Segmented ECG heart beat of VF patient

Fig. 3. Segmentation process of (a)Normal heart beat (b) VT patients heart beat and (c) VF patients heart beat.

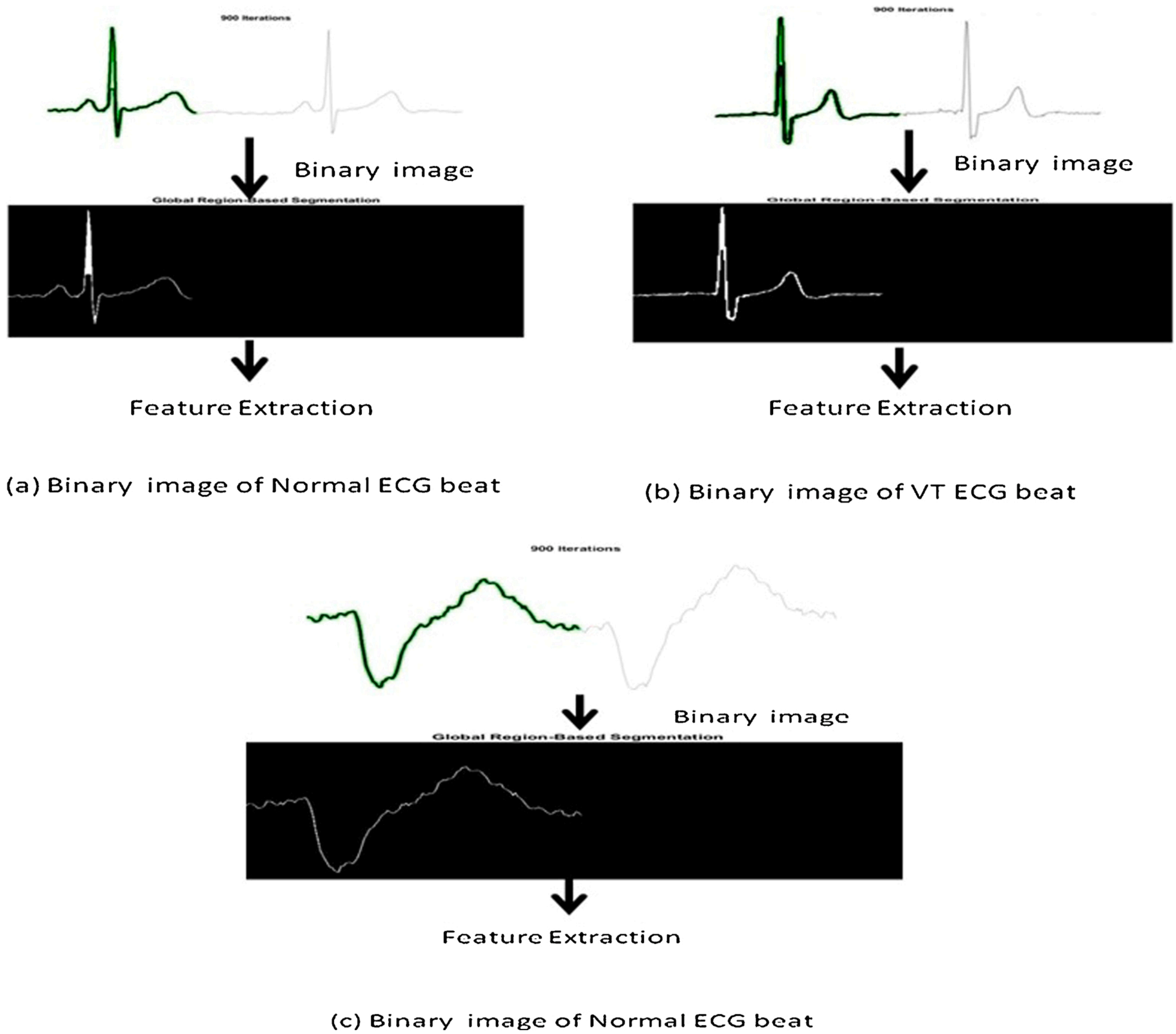


Fig. 4. Gray scale image to binary image conversion process (a)Normal heart beat (b) VT patients heart beat and (c) VF patients heart beat.

Where $(M_a^k)_l$ represents the lth observation of class a corresponding to features m_k. $S(m_k)$ evaluates the degree of tightness with in class for the features m_k. Now, $F(m_k)$ evaluates feature acceptance ratio using Eqs. (3) and (4).

$$F(m_k) = \frac{\sum_{a=1}^{c} P(m_k)_a}{\sum_{a=1}^{c} S(m_k)_a} = \text{Cardiac Score} \geq F_{\text{thresh}} \tag{5}$$

$F(m_k)$ represents how well the feature m_k is correlated with the class. Larger value of $F(m_k)$ represents high correlation of the feature with the class. We set threshold value $F_{thresh} = 0.5$ by ranking criteria $F(m_k)$.

3. Classification

Ensemble classifier consists of many classifiers. Main objective is to design such type of classifier to increase the classification accuracy. Fig. 6 presents working steps of assemble classifier. In this diagram 'n' number of classifiers are shown. Training data sets are used by all the classifiers. The training data sets are selected by the use of feature selection algorithm. Generally the classifier number may even or odd.

In this study, we have used even number of classifiers ($n = 4$). Decision of the ensemble classifier depends on majority votes [48]. In this work, class-1 represents normal class, class-2 represents VT class and class-3 represents VF class. If classifier-1 to classifier-4 respectively gives decision class output 1,1,3,2. Then, majority votes decision class is 1 (normal). Other possibilities may arise, if decision class of the four classifiers are 1,1,2,2. In such cases, ensemble classifier makes no decision. Overcome this limitation, in this work we have taken average result of majority votes. In this case, it is 1.5. Then, we take the next rounded integer 2 and decision is VT class.

In this work, we have taken $n = 4$ numbers of classifiers as follows.

- Support Vector Machine (SVM) classifier [49]

In this study, Radial Basis Function (RBF) kernel is used for SVM with their best parameters leading to higher accuracy.

- Probabilistic Neural Network(PNN) [50]

In this study, features are used as an input data of PNN network. In PNN network, number of input layers are same as number of features. This work have required three different output value. Therefore, the output layer of the network had three neurons. At the end of the work, best performance result is evaluated by the use of suitable PNN parameters.

- *K*-nearest neighbor (KNN) [37]

In this work, we have selected the value of k in such a way that it produces highest classification rate. Original data samples are randomly partitioned into k-sub samples. $(K - 1)$ sub samples are used for training and rest sub sample is used for testing. This process is repeated for k-times. Finally, we have taken average value for k number of results of the folds to produce single estimation. In this study, five different k values are taken i.e. 1,3,5,7,9 for the computation of five fold cross validation technique. Euclidean distance, city block and correlation methods are

Table 2
Methods for features extraction.

Method applied for features extraction from ECG beat image		
SL number	Feature	Equation
Method 1	F_{1B} (Mean)	$\overline{Z} = \sum_{i,j=0}^{L-1} iZ_{ij}$
Method 2	F_{2B} (Variance)	$\sigma^2 = \sum_{i,j=0}^{L-1} Z_{ij}(i-\overline{Z})$
Method 3	F_{3B} (Energy)	$E = \sum_{i,j=0}^{L-1}(Z_{ij})^2$
Method 4	F_{4B} (Entropy)	$\mathrm{En} = \sum_{i,j=0}^{L-1} - ln(Z_{ij})(Z_{ij})$
Method 5	F_{5B} (Correlation)	$C = \sum_{i,j=0}^{L-1} Z_{ij}\frac{(i-\overline{Z})(j-\overline{Z})}{\sigma^2}$
Method 6	F_{6B} (Homogeneity)	$H = \sum_{i,j=0}^{L-1}\frac{Z_{ij}}{1+(i-j)^2}$
Method 7	F_{7B} (Contrast)	$\mathrm{Cont} = \sum_{i,j=0}^{L-1} Z_{ij}(i-j)^2$
Method 8	F_{8B} (Dissimilarity)	$\mathrm{Dsim} = \sum_{i,j=0}^{L-1} Z_{i,j}\lvert i-j\rvert$
Method 9	F_{9B} (Joint Maxima)	$Jmax = max(Z_{i,j})$
Method 10	F_{10B} (Cluster Tendency)	$C_t = \sum_{i,j=0}^{L-1}((i-\overline{Z}+)(i-\overline{Z}))^2 Z_{i,j}$
Method 11	F_{11B} (Cluster Shade)	$C_s = \sum_{i,j=0}^{L-1}((i-\overline{Z}+)(i-\overline{Z}))^3 Z_{i,j}$
Method 12	F_{12B} (Cluster Prominence)	$C_p = \sum_{i,j=0}^{L-1}((i-\overline{Z}+)(i-\overline{Z}))^4 Z_{i,j}$
Method 13	F_{13B} (Joint Entropy)	$J_{\mathrm{En}} = \sum_{i,j=0}^{L-1} Z_{i,j}\log_2 Zi,j$
Method 14	F_{14B} (Inverse Difference)	$I_{\mathrm{diff}} = \sum_{i,j=0}^{L-1}\frac{Z_{i,j}}{1+\lvert i-j\rvert}$
Method 15	F_{15B} (Inverse Variance)	$I_{\mathrm{var}} = \sum_{i,j=0}^{L-1}\frac{Z_{i,j}}{\lvert i-j\rvert^2}$

Note: The sample input ($Z_{i,j}$) image consists of N gray levels. We have calculated the relationship value between reference pixel *i* and neighbour pixel *j* in various orientation.

Table 3
Continued… Methods for features extraction.

Method applied for features extraction from HRV signal		
SL Number	Feature	Equation
Method 16	F_{1H} (Standard Deviation)	$\sigma^* = \sqrt{\frac{1}{L}\sum_{i=1}^{L-1} Z_i - \overline{Z}}$
Method 17	F_{2H} (Skewness)	$S = \frac{1}{L}\sum_{i=0}^{L-1}(\frac{Z_i-\overline{Z}}{\sigma^*})^3$
		Where, $\overline{Z} = \frac{1}{L}\sum_{i=0}^{L-1} Z_i$
Method 18	F_{3H} (Kurtosis)	$S = \frac{1}{L}\sum_{i=0}^{L-1}(\frac{Z_i-\overline{Z}}{\sigma^*})^4$
Method 19	F_{4H} (Shape Factor)	$\mathrm{SH}_F = \frac{Z_{\mathrm{rms}}}{\frac{1}{L}\sum_{i=0}^{L-1} Z_i}$
Method 20	F_{5H} (Geometric Mean)	$G_m = L\sqrt{Z_1+Z_2+Z_3+\cdot+\cdot+Z_L}$
Method 21	F_{6H} (Harmonic Mean)	$H_r = \frac{L}{\frac{1}{Z_1}+\frac{1}{Z_2}+\frac{1}{Z_3}+\cdot+\cdot+\frac{1}{Z_L}}$
Method 22	F_{7H} (Impulse Factor)	$I_F = \frac{\mathrm{Max}\lvert Z_i\rvert}{\frac{1}{L}\sum_{i=0}^{L-1} Z_i}$
Method 23	F_{8H} (Crest Factor)	$C_F = \frac{\sqrt{\frac{1}{L}\sum_{i=0} L-1(Z_i)^2}}{\frac{1}{L}\sum_{i=0}^{L-1}\lvert Z_i\rvert}$
Method 24	F_{9H} (Margin Factor)	$M_F = \frac{\mathrm{Max}\lvert Z_i\rvert}{(\frac{1}{L}\sum_{i=0}^{L-1} Z_i)^2}$
Method 25	F_{10H} (Hjort's Parameter)	$H_j = \frac{\sigma_x'}{\sigma}$
	- Mobility	σ_x = variance of signal amplitude
		σ_x'= 1st order derivative
Method 26	F_{11H} (Complexity)	$H_j = \frac{\sigma_x''/\sigma'}{\sigma_x'/\sigma_x}$
Method 27	F_{12H} (Frequency Center)	$F_c = \frac{\sum_{i=2}^{L} Z_i' Z_i}{2\pi\sum_{i=1}^{L} Z_i^2}$
		Z_i'= Derivative of difference between two samples
Method 28	F_{13H} (Mean Square Frequency)	$M_S = \frac{\sum_{i=2}^{L} Z_i'^2}{4\pi\sum_{i=1}^{L} Z_i^2}$
Method 29	F_{14H} (Root Variance Frequency)	$R_v = \sqrt{(M_s)^2-(F_c)^2}$
Method 30	F_{15H}(Fractal Dimension)	Higuchi Method $D = \frac{\log L(k)}{\log k}$, Reference [22]

Note: Z_i is input data sets (HRV signal).

used to compute nearest distance [38]. We have found that $k = 3$ is efficient for highest classification accuracy.

- Random Forest (RF) [52]

In this work, RF with 75 decision trees with maximum depth of 3 is applied.

The main concept is to use ensemble classifier for the correction of errors of the previous four classifiers. Ensembles provide an extra degree of freedom in the classical bias/variance tradeoff and allow a solution that would be difficult to reach with only a single hypothesis.

4. Result and discussion

In this work, seven significant features are selected using cardiac scoring technique. Table 5 shows seven selected features that are extracted from HRV signals and image data.

From Table 4, it is clearly found that 12 features are selected from HRV signal features after 1st cardiac score analysis. Also, 11 features are selected from 15 ECG beat image features after 1st cardiac score analysis. Selected features are presented in 'selected feature' column. Total 30 features are found after combined of HRV and beat image features. Then, the feature selection algorithm is applied using new threshold value. In that case, threshold value is the average of all cardiac score that was obtained in 1st cardiac score treatment. Finally, we have got seven features using 2nd cardiac score analysis. These seven features are used for the classification of VA patients. Range of the seven sub-features (mean ± standard deviation) is extracted from the 18 normal, 30 VT and 22 VF patients ECG signals. We observed that the range of a few features is higher in normal class and range of a few features is higher in VT and VF class. Different classifiers are used for classification of above mentioned features.

Fig. 7 shows classification accuracy of the classifiers for different feature combinations. In this study, the proposed system has given better classification accuracy by the use of selected features. Another observation is that the system classification is improved by the use of ensemble classifier. Therefore, with the combination of selected features and ensemble classifier the system has achieved best classification accuracy. Performance of all the classifiers are analysed using six performance measure parameters.

There are several terms which are commonly used for the description of accuracy, specificity and sensitivity. True positive (p) is considered when the classifier correctly diagnose the presence of the disease. Similarly, if the classifier correctly detect the absent of the disease, result is true negative (q). A false positive (r) occurs when classifier incorrectly indicates the presence of a disease when the condition is not present. False negative (r) is considered when the test result of the classifier fails to detect the presence of the condition when the disease is present.

- Accuracy (A_{cc}) = (p+q)/(p+q+r+s)
- Specificity (S_{pe}) = q/(q+r)
- Sensitivity (S_{en}) = p/(p+s
- Area Under the Curve (AUC):

  It is related to training and testing performance of the classifier. Area of the curve 1 represents ideal test. Value of AUC = 1.0 represents 100% accuracy of the test data.

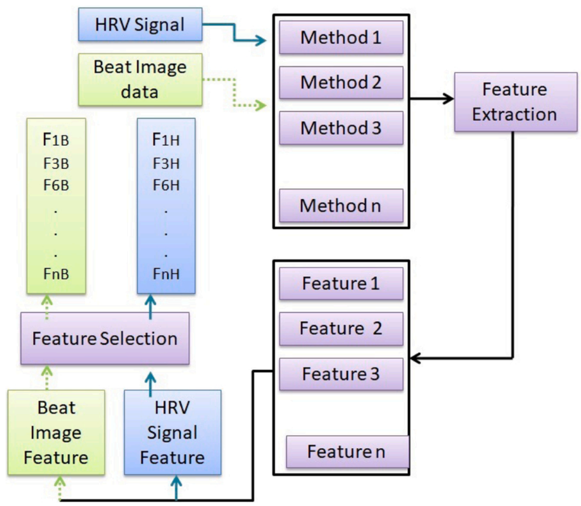


Fig. 5. Flow graph for feature extraction.

- Kappa:
  Kappa is a statistical parameter that is based on the difference between expected and observed accuracy statistics.
- *K* fold cross validation:
  System is trained *K*-times with *K*-training data sets and apply to the corresponding testing data sets for prediction of output.

Firstly, 11 HRV features and 12 beat image features are classified using four types of classifiers and the performance of the classifiers are measured using these six performance measure parameters. Then, selected seven combined features are used for performance measure using the same procedure. Later, ensemble classifier is used to check the system performance using seven selected features. We observed that the ensemble classifier using reduced number of features (seven) gives better classification results compare to the classification results using individual HRV features or image beat features. Performance analysis using $n = 4$ numbers of classifiers are shown in Tables 6 and 7.

Specificity and sensitivity are basically counter part of each other of VA detection system. Balancing both of them as final test criteria, *F*-score

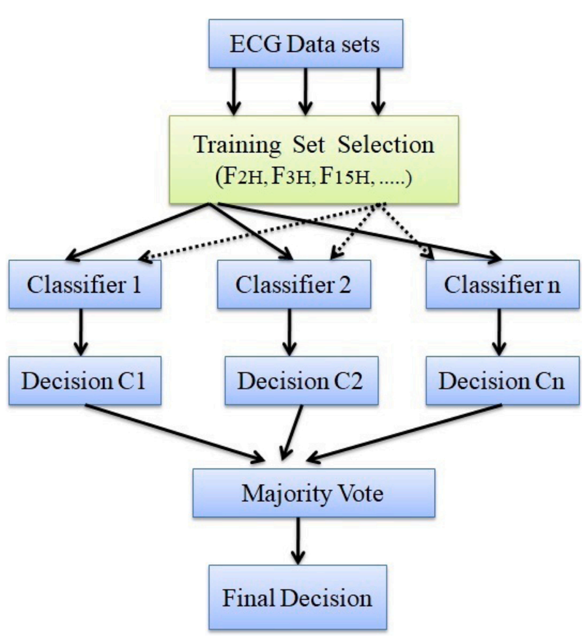


Fig. 6. Flow diagram of ensemble classifier.

Table 4
Features selection results using cardiac-score.

Signal type	Feature number	Feature selection	Number of selected features	**Selected features name**	**Selected features**
HRV Signal	15	Cardiac score ≥ 0.5	11	HRV signal features	F_{2H}, F_{3H}, F_{4H}, F_{5H}, F_{8H} F_{10H}, F_{11H}, F_{12H}, F_{13H} F_{14H}, F_{15H}
ECG beat image data	15	Cardiac score ≥ 0.5	12	Beat image features	F_{3B}, F_{4B}, F_{5B}, F_{6B}, F_{7B} F_{8B}, F_{10B}, F_{11B} F_{12}, F_{13B}, F_{14B}, F_{15B}
HRV + Image beat	30	Cardiac score ≥ average value of all cardiac score	7	Combined Features	F_{2H}, F_{3H}, F_{4H}, F_{15H} F_{4B}, F_{7B}, F_{8B}

Table 5
Selected features range for three classes, normal, VT and VF.

Selected features	Normal	VT	VF	*P*-value
F_{2H}	30 ± 02	10 ± 02	3.0023 ± 0.6	<0.0001
F_{3H}	89 ± 02	41 ± 02	59 ± 05	<0.0001
F_{4H}	0.9707 ± 0.10	0.6532 ± 0.10	0.4652 ± 0.10	<0.0001
F_{15H}	1.7021 ± 0.10	1.9156 ± 0.10	2.11 ± 0.10	<0.0001
F_{4B}	1.50 ± 0.10	1.86 ± 0.10	1.3098 ± 0.09	<0.0001
F_{7B}	0.1127 ± 0.05	0.3260 ± 0.05	0.0901 ± 0.03	<0.0001
F_{8B}	0.3156 ± 0.10	0.6266 ± 0.10	0.1038 ± 0.08	<0.0001

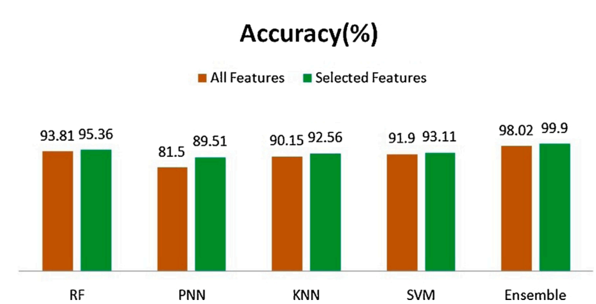


Fig. 7. Accuracy plot of classifiers for different feature combinations.

has been calculated to check the system performance. *F*-score is a statistical analysis to measure the test accuracy. It is calculated from the harmonic mean of positive predictive value and sensitivity. It is observed that our system has given better test accuracy by the use of ensemble classifier with selected features. Fig. 8 shows the system performance analysis by the use of *F*-score.

The ECG data set is divided into two different categories. 80% of the data is used for learning and 20% of the data is used for testing. 20% of learning data is used as a validation set and rest 80% is used as a training set. The model was trained by the training set as input and it was validated by the use of validation set. Test set is used to verify and validate the model through accuracy estimation. The performance of the model is evaluated by *K*-fold cross validation approach.

In this literature, following three measures have been taken to avoid over fitting.

Table 6
Classification result analysis using extracted features.

Performance parametes	HRV signal features ($N = 15$)	ECG beat image feature ($N = 15$)	Combined features ($N = 30$)
Random Forest (RF) classifier			
A_{cc}(%)	85.35	90.97	94.81
S_{pe}	0.88	0.91	0.984
S_{en}	0.83	0.90	0.890
AUC	0.85	0.90	0.94
K_{appa}	0.75	0.96	0.88
$K(5)Fold$(%)	84.42	89.26	93.00
Probabilistic Neural Network (PNN) classifier			
A_{cc}(%)	77.18	79.69	81.50
S_{pe}	0.94	0.95	0.93
S_{en}	0.58	0.65	0.66
AUC	0.66	0.78	0.81
K_{appa}	0.64	0.96	0.70
$K(5)Fold$(%)	-	-	-
K-nearest neighbor (KNN) classifier			
A_{cc}(%)	86.60	86.75	90.15
S_{pe}	0.87	0.88	0.95
S_{en}	0.77	0.78	0.85
AUC	0.86	0.87	0.90
K_{appa}	0.61	0.63	0.70
$K(5)Fold$(%)	81.35	82.34	89.05
Support Vector Machine (SVM) classifier			
A_{cc}(%)	86.90	90.50	91.90
S_{pe}	0.84	0.91	0.92
S_{en}	0.88	0.89	0.90
AUC	0.86	0.90	0.9
K_{appa}	0.64	0.71	0.79
$K(5)Fold$(%)	–	–	–
Ensemble classifier			
A_{cc}(%)	96.01	97.04	98.02
S_{pe}	0.98	0.99	1.00
S_{en}	0.91	0.94	0.93
AUC	0.96	0.97	0.98
K_{appa}	0.73	0.77	0.79
$K(5)Fold$(%)	94.90	94.05	97.90

Table 7
Classification result analysis using selected features.

Performance parametes	HRV signal features ($N = 11$)	ECG beat image feature ($N = 12$)	Combined features ($N = 07$)
Random Forest (RF) classifier			
A_{cc}	91.67	92.00	95.36
S_{pe}	0.96	0.93	0.96
S_{en}	0.85	0.90	0.94
AUC	0.91	0.92	0.95
K_{appa}	0.82	0.87	0.88
$K(5)Fold$(%)	90.00	91.05	93.05
Probabilistic Neural Network (PNN) classifier			
A_{cc}(%)	80.29	84.51	89.51
S_{pe}	0.83	0.88	0.96
S_{en}	0.79	0.81	0.82
AUC	0.80	0.84	0.89
K_{appa}	0.58	0.61	0.71
$K(5)Fold$(%)	–	–	–
K-nearest neighbor(KNN) classifier			
A_{cc}(%)	87.27	88.02	92.56
S_{pe}	0.89	0.90	0.97
S_{en}	0.86	0.86	0.88
AUC	0.87	0.88	0.92
K_{appa}	0.61	0.70	0.77
$K(5)Fold$(%)	82.51	84.88	89.59
Support Vector Machine (SVM) classifier			
A_{cc}(%)	88.98	92.6	93.11
S_{pe}	0.91	0.94	0.97
S_{en}	0.86	0.91	0.89
AUC	0.88	0.92	0.93
K_{appa}	0.70	0.77	0.81
$K(5)Fold$(%)	–	–	–
Ensemble classifier			
A_{cc}(%)	98.20	98.90	99.90
S_{pe}	0.99	0.99	1.00
S_{en}	0.93	0.89	0.98
AUC	0.98	0.98	1.00
K_{appa}	0.82	0.96	0.89
$K(5)Fold$(%)	97.50	98.10	99.80

(a) Cross-validation is a powerful method to prevent over fitting [60]. In this work K-fold cross-validation method has been used. We randomly split the dataset into $K = 5$ non overlapping folds. In each experiment 20% of learning dataset is used as a validation set and rest of 80% data is used as a training set. We iteratively train the algorithm on $K - 1$ folds while using the remaining fold as a validation set. This experiment is repeated for $k = 5$ times. Cross validation allows the tuning of the hyper parameter. For the different values of the parameters, the model is trained and it is tested using the validation set. After finding out the value of parameters, model is again trained using the training dataset. After that finally the model is tested using test dataset for $k = 5$ fold and the mean performance is reported in Tables 6 and 7.

(b) Regularization is a process to combine all the strong learner models together to smooth out the prediction model. This work combines the prediction of the four classifiers (RF, PNN, KNN, SVM). It trains the models in parallel.

(c) Dropping out the features which are not efficient (high variance and high bias) to train the model. In this work seven features are selected to train the model.

Most researchers have used NSRDB,CUDB,VFDB database for the detection of Ventricular Arrhythmia [53–58]. Therefore, in our study same database is used for the performance validation of the proposed system. Performance comparison of our model with other research works is shown in Table 8. Almost all studies have used machine learning algorithms for the detection of VA patient. Mohanti et al. [53] proposed Variational Mode Decomposition technique for feature extraction. Temporal, spectral and statistical features ware extracted

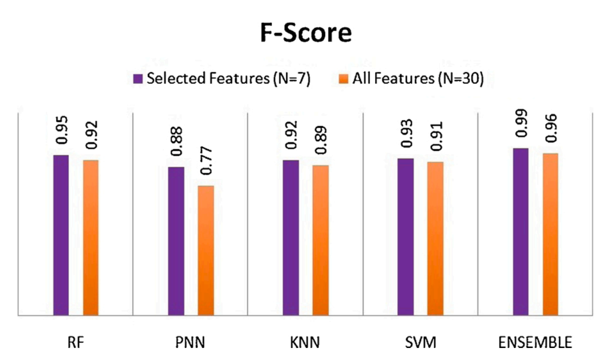


Fig. 8. F-score plot of the classifiers for the different feature combinations.

and ranker search method was used for feature ranking. They reported that the C4.5 classifier has given better classification accuracy compare to SVM classifier. But this model uses 24 informative features to achieve 99.18% accuracy for VA detection. Our proposed model uses 7 features to achieve 99.9% accuracy for VA detection. Taye et al. [54] extracted features from the QRS complex to capture the morphological changes to diagnose AV patients. In another work of Mohanti [55] for the detection of VA patient, same classifier is used with time frequency based feature extraction technique. His proposed model was tested using same training sets, still it has been achieved poor sensitivity (Sen 90.97%, Spe 97.86%, Acc 97.02%). Adaptive variational mode decomposition technique for feature extraction proposed by Yang et al. [56] is obtained maximum accuracy of 98.29%. Ibtchaz et al. [57] presented EMD and DFT technique for feature extraction and obtained 99.19% accuracy

Table 8
Comparison study for VA detection.

	Studies conducted	Features extraction technique	Database	Classifier	Class	Accuracy
01	Mohanti et al. 2019 [53]	Variational Mode Decomposition (VMD)	NSRDB CUDB VFDB	C 4.5	NSR VT VF	99.18 %
02	Taye et al. 2019 [54]	QRS complex shape Based feature	NSRDB CUDB	ANN	NSR VT	98.6 %
03	Mohanti et al. 2018 [55]	Thirteen Time-Freq Based feature	NSRDB CUDB VFDB	C 4.5	NSR VT VF	97.02 %
04	Yang et al. 2017 [56]	Adaptive Variational Mode Decomposition (AVMD)	NSRDB CUDB VFDB	Boosted - CART	NSR VT VF	98.29 %
05	Ibtchaz et al. 2019 [57]	Empirical Mode Decomposition and DFT	NSRDB CUDB	SVM	NSR VF	99.19 %
06	Hatradat et al. 2019 [58]	Empirical Mode Decomposition	NSRDB CUDB VFDB	Binary Pattern	NSR VT VF	99.70 %
07	Proposed Method	Combined Features (HRV and Image beat) 30 time-freq based features	NSRDB CUDB VFDB	Ensemble	NSR VT VF	99.99 %

with only VF and NSR signal in the database. But the limitation of the system is to detect only VF patients, while VT signal is quite similar in appearance of VA patients. Hatradat et al. [58] studied Empirical Mode Decomposition technique and obtained 99.70% accuracy by the use of binary pattern classification approach. But this scheme uses many features. It is noted that VA detection system in above studies are based on ECG signal analysis. Whereas, this work focuses on combined feature extraction technique by the use of ECG signal as well as ECG beat image.

The proposed cardiac score based feature selection technique along with ensemble classifier outperform traditional signal processing algorithm. From the above comparative research, it is found that our proposed model performed better compared to that of other models. In addition, our proposed model uses significantly less number of features compared to the presented models in Table 8. Significant features not only improves system performance, it also reduces computational burden of the system. Hence, it can be considered as an essential reference for cardiac health monitoring system.

4.1. Statistical hypothesis test

Statistical hypothesis testing is used to find out whether the results of the two methods are statistically significant or not. McNemar's test is a statistical test to analyze the significant difference of the two classifier performance [61,62]. This work shows the statistical difference of the proposed method (method A) over the existing ones [53] (method B) in Table 8.

Let, n_{AB} be the number of examples misclassified by method A but correctly classified by the method B. n_{BA} be the number of examples misclassified by method B but not by method A. Under the null hypothesis both the methods have same error rate, the test statistics ChiSqure(X^2) is estimated with 1 degree of freedom from the following equation:

$$X^2 = \frac{(|n_{\mathrm{AB}} - n_{\mathrm{BA}}| - 1)^2}{n_{\mathrm{AB}} + n_{\mathrm{BA}}} \tag{6}$$

If the estimated test value is more than 3.85 or less than 0.05, then it indicates that the null hypothesis is rejected. In this work, Table 9 shows that the X^2 value is 16.82 which is clearly the indication of null hypothesis rejection. Therefore, the performance difference of the two methods A and B is statistically relevant.

5. Conclusion

In this study, the ECG data is splitted in two different characteristics of ECG signals. The extracted features from two data sets are used for classification of VA. This study suggests that HRV signals and ECG beat images are useful to detect VA patient by using machine learning algorithms. Seven significant features are selected using cardiac score technique. The combination of these selected features and ensemble classifier provide high classification accuracy. Also, the robustness and credibility of the study have been ensured by 5-fold cross validation results. This work has achieved relatively better performance compare to other existing work of VA detection system, but the detection accuracy of the system is restricted to samples. Our future study will focus on design a VA detection system for large number of diverge database.

Table 9
Statistical analysis of two methods using McNemar's test.

Algorithm B used for comparison	Reference algorithm A = ensemble classifier		
	X^2	Comment on hypothesis test	Database
C4.5 classifier [53]	16.82	Reject	MIT-BIH

Author contributions

Conception and design of study: M. Saurav, M. Pulak, R.H. Anisha
Acquisition of data: M. Saurav, M. Pulak, R.H. Anisha
Analysis and/or interpretation of data: M. Saurav, M. Pulak, R.H. Anisha
Drafting the manuscript: M. Saurav, M. Pulak, R.H. Anisha
Revising the manuscript critically for important intellectual content: M. Saurav, M. Pulak, R.H. Anisha
Approval of the version of the manuscript to be published: M. Saurav, M. Pulak, R.H. Anisha

Declaration of Competing Interest

The authors report no declarations of interest.

References

[1] A.S. Go, E.M. Hylek, K.A. Phillips, Y. Chang, L.E. Henault, J.V. Selby, D.E. Singer, Prevalence of diagnosed atrial fibrillation in adults: national implications for rhythm management and stroke prevention: the anticoagulation and risk factors in atrial fibrillation (ATRIA) study, JAMA 285 (2001) 2370–2375.
[2] C. Richardson, E.S. Silver, Management of supraventricular tachycardia in infants, Pediatr. Drugs 19 (2017) 539–551, https://doi.org/10.1007/s40272-017-0254-0.
[3] American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (ACC/AHA/HRS Writing Committee to Develop Data Standards on Electrophysiology), A.E. Buxton, H. Calkins, D.J. Callans, J.

P. DiMarco, J.D. Fisher, H.L. Greene, D.E. Haines, D.L. Hayes, P.A. Heidenreich, J. M. Miller, A. Poppas, E.N. Prystowsky, M.H. Schoenfeld, P.J. Zimetbaum, D. C. Goff, F.L. Grover, D.J. Malenka, E.D. Peterson, M.J. Radford, R.F. Redberg, ACC/AHA/HRS 2006 key data elements and definitions for electrophysiological studies and procedures: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Data Standards (ACC/AHA/HRS Writing Committee to Develop Data Standards on Electrophysiology), Circulation 114 (December (23)) (2006) 2534–2570.
[4] F. Zhou, T. Yu, R. Du, Clinical course and risk factors for mortality of adult in patients with COVID-19 in Wuhan, China: a retrospective cohort study, Lancet 395 (2020) 1054–1062, https://doi.org/10.1016/S0140-6736(20)30566-3.
[5] V. Krasteva, I. Jekova, Assessment of ECG frequency and morphologyparameters for automatic classification of life-threatening cardiacarrhythmias, Physiol. Meas. 26 (5) (2005) 707–723.
[6] S. Schuch, S.P. Tipper, On observing another person's actions: influences of observed inhibition and errors, Percept. Psychophys. 69 (5) (2007) 828–837, https://doi.org/10.3758/bf03193782.
[7] C. Belfield, Cost-benefit analysis and cost-effectiveness analysis, Int. Encycl. Educ. (2010) 199–203, https://doi.org/10.1016/b978-0-08-044894-7.01245-8.
[8] T. Mar, S. Zaunseder, J.P. Martínez, M. Llamedo, R. Poll, Optimization of ECG classification by means of feature selection, IEEE Trans. Biomed. Eng. 58 (2011) 2168–2177.
[9] Andreotti, F. Grasser, F. Malberg, H. Zaunseder, Sebastian, Non-invasive fetal ECG signal quality assessment for multichannel heart rate estimation, IEEE Trans. Biomed. Eng. (2017) 1, https://doi.org/10.1109/TBME.2017.2675543.
[10] R.E. Kalman, A new approach to linear filtering and prediction problems, J. Basic Eng. 82 (1) (1960) 35–45.
[11] R.N.V.P.S. Kandala, R. Dhuli, P. Pławiak, G. Naik, H. Moeinzadeh, G.D. Gargiulo, S. Gunnam, Towards real-time heartbeat classification: evaluation of nonlinear morphological features and voting method, Sensors 19 (23) (2019) 5079, https://doi.org/10.3390/s19235079.
[12] F.E. Grubbs, Procedures for detecting outlying observations in samples, Technometrics 11 (1) (1969) 1–21, https://doi.org/10.1080/00401706.1969.10490657.
[13] V. Khoma, M. Pelc, Y. Khoma, D. Sabodashko, Outlier correction in ECG-based human identification, Adv. Intell. Syst. Comput. (2018) 11–22, https://doi.org/10.1007/978-3-319-75025-52.
[14] J.A.K. Suykens, J. Vandewalle, Least square support vector machine classifiers, Neural Process. Lett. 9 (1999) 293–300.
[15] R. Kumar, A. Kumar, G.K. Singh, Hybrid method based on singular value decomposition and embedded zero tree wavelet technique for ECG signal compression, Comput. Methods Programs Biomed. 129 (2016) 135–148, https://doi.org/10.1016/j.cmpb.2016.01.006.
[16] D.G. Giri, U.R. Acharya, R.J. Martis, V.S. Sree, T.C. Lim, T.V.I. Ahamed, J.S. Suri, Automated diagnosis of coronary artery disease affected patients using LDA, PCA ICA and discrete wavelet transform, Knowl. Based Syst. 37 (2013) 274–282.
[17] P.S. Hamilton, W.J. Tompkins, Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database, IEEE Trans. Eng. Biomed. Eng. 12 (1986) 1157–1165.
[18] D.S. Benitez, A new QRS detection algorithm based on the Hilbert transform, Comput. Cardiol. 27 (2000) 379–382.
[19] P. Sharma, K.C. Ray, Efficient methodology for electrocardiogram beat classification, IET Signal Process. 10 (2016) 825–832.
[20] A. Bajaj, S. Kumar, QRS complex detection using fractional Stockwell transform and fractional Stockwell Shannon energy, Biomed. Signal Process. Control 54 (2019) 101628, https://doi.org/10.1016/j.bspc.2019.101628.
[21] S. Shadmand, B. Mashoufi, A new personalized ECG signal classification algorithm using block-based neural network and particle swarm optimization, Biomed. Signal Process. Control 25 (2016) 12–23, https://doi.org/10.1016/j.bspc.2015.10.008.
[22] S. Yucelbas, C. Yucelbas, G. Tezel, S. Ozşen, Ş. Yosunkaya, Automatic sleep staging based on SVD, VMD, HHT and morphological features of single-lead ECG signal, Expert Syst. Appl. (2018), https://doi.org/10.1016/j.eswa.2018.02.034. S0957417418301234-.
[23] M. Rai Hari, A. Trivedi, S. Shukla, ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier, Measurement 46 (9) (2013) 3238–3246, https://doi.org/10.1016/j.measurement.2013.05.021.
[24] O. Perlman, A. Katz, G. Amit, Y. Zigel, Supraventricular tachycardia classification in the 12-lead ECG using atrial waves detection and a clinically based tree scheme, IEEE J. Biomed. Health Inform. 20 (6) (2016) 1513–1520, https://doi.org/10.1109/jbhi.2015.2478076.
[25] S. Ortín, M.C. Soriano, M. Alfara, C.R. Mirasso, Automated real-time method for ventricular heartbeat classification, Comput. Methods Programs Biomed. (2018), https://doi.org/10.1016/j.cmpb.2018.11.005.
[26] F.A. Elhaj, N. Salim, A.R. Harris, T.T. Swee, T. Ahmed, Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals, Comput. Methods Programs Biomed. 127 (2016) 52–63, https://doi.org/10.1016/j.cmpb.2015.12.024.
[27] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, R.S. Tan, Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals, Appl. Intell. (2018), https://doi.org/10.1007/s10489-018-1179-1.
[28] M. Hammad, K. Wang, Parallel score fusion of ECG and fingerprint for human authentication based on convolution neural network, Comput. Secur. 81 (2019) 107–122, https://doi.org/10.1016/j.cose.2018.11.003.
[29] O. Yildirim, U. Baloglu, U. Acharya, A deep learning model for automated sleep stages classification using PSG signals, Int. J. Environ. Res. Public Health 16 (4) (2019) 599, https://doi.org/10.3390/ijerph16040599.
[30] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, A. Gertych, R.S. Tan, A deep convolutional neural network model to classify heartbeats, Comput. Biol. Med. 89 (2017) 389–396, https://doi.org/10.1016/j.compbiomed.2017.08.022.
[31] U.B. Baloglu, M. Talo, O. Yildirim, R.S. Tan, U.R. Acharya, Classification of myocardial infarction with multi-lead ECG signals and deep CNN, Pattern Recognit. Lett. (2019), https://doi.org/10.1016/j.patrec.2019.02.016.
[32] F. Rebouças, P.P. Rebouças, E.S. de Marinho, L.B. Sarmento, J.M.R.S. Tavares, V.H. C. de Albuquerque, Analysis of human tissue densities: a new approach to extract features from medical images, Pattern Recognit. Lett. 94 (2017) 211–218, https://doi.org/10.1016/j.patrec.2017.02.005.
[33] S. Dutta, A. Chatterjee, S. Munshi, Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification, Med. Eng. Phys. 32 (10) (2010) 1161–1169.
[34] S. Das, S. Pal, M. Mitra, Supervised model for Cochleagram feature based fundamental heart sound identification, Biomed. Signal Process. Control 52 (2019) 32–40, https://doi.org/10.1016/j.bspc.2019.01.028.
[35] M. Rasooli, F.H. Foomany, K. Balasundaram, S. Masse, N. Zamiri, A. Ramadeen, X. Hu, P. Dorian, K. Nanthakumar, S. Krishnan, S. Beheshti, K. Umapathy, Analysis of electrocardiogram pre-shock waveforms during ventricular fibrillation, Biomed. Signal Process. Control 21 (2015) 26–33, https://doi.org/10.1016/j.bspc.2015.05.003.
[36] J. Lee, D.D. McManus, P. Bourrell, L. Sörnmo, K.H. Chon, Atrial flutter and atrial tachycardia detection using Bayesian approach with high resolution time-frequency spectrum from ECG recordings, Biomed. Signal Process. Control 8 (6) (2013) 992–999, https://doi.org/10.1016/j.bspc.2013.04.002.
[37] S.A. Dudani, The Distance-weighted k-nearest-neighbor rule, IEEE Trans. Syst. Man Cybern. SMC-6 (4) (1976) 325–327, https://doi.org/10.1109/tsmc.1976.5408784.
[38] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[39] P. Singh, G. Pradhan, Denoising of ECG signal by non-local estimation of approximation coefficients in DWT, Biocybern. Biomed. Eng. 37 (3) (2017) 599–610, https://doi.org/10.1016/j.bbe.2017.06.001.
[40] A. Singhal, P. Singh, B. Fatimah, R.B. Pachori, An efficient removal of power-line interference and baseline wander from ECG signals by employing Fourier decomposition technique, Biomed. Signal Process. Control 57 (2020) 101741, https://doi.org/10.1016/j.bspc.2019.101741.
[41] S. Cuomo, G. De Pietro, R. Farina, A. Galletti, G. Sannino, A revised scheme for real time ecg signal denoising based on recursive filtering, Biomed. Signal Process. Control 27 (2016) 134–144.
[42] G.M. Friesen, T.C. Jannett, M.A. Jadallah, S.L. Yates, S.R. Quint, H.T. Nagle, A comparison of the noise sensitivity of nine QRS detection algorithms, IEEE Trans. Biomed. Eng. 37 (1) (1990) 85–98.
[43] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. BME-32 (3) (1985) 230–236, https://doi.org/10.1109/tbme.1985.325532.
[44] H. Jia, C. Lang, D. Oliva, W. Song, X. Peng, Hybrid grasshopper optimization algorithm and differential evolution for multilevel satellite image segmentation, Remote Sens. 11 (9) (2019) 1134, https://doi.org/10.3390/rs11091134.
[45] S. Wang, F. Ge, T. Liu, Evaluating edge detection through boundary detection, EURASIP J. Adv. Signal Process. (2006), https://doi.org/10.1155/asp/2006/76278.
[46] T. Pavlidis, Y.T. Liow, Integrating region growing and edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 225–233.
[47] Q. Xu, P. Fu, Q. Sun, T. Wang, et al., A fast region growing based superpixel segmentation for hyperspectral image classification, in: Z. Lin (Ed.), Pattern Recognition and Computer Vision. PRCV 2019. Lecture Notes in Computer Science, vol. 11858, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-31723-266.
[48] A. Narasimhamurthy, Theoretical bounds of majority voting performance for a binary classification problem, IEEE Trans. Pattern Anal. Mach. Intell. 27 (12) (2005) 1988–1995, https://doi.org/10.1109/tpami.2005.249.
[49] D. Giveki, M. Karami, Scene classification using a new radial basis function classifier and integrated SIFT-LBP features, Pattern Anal. Appl. (2020), https://doi.org/10.1007/s10044-020-00868-7.
[50] B. Mohebali, A. Tahmassebi, A. Meyer Baese, A.H. Gandomi, Probabilistic neural networks. Handbook of Probabilistic Models, 2020, pp. 347–367, https://doi.org/10.1016/b978-0-12-816514-0.00014-x.
[51] G.V. Mondejar, J. Novvo, J. Rouco, M.G. Penedo, M. Ortega, Heartbeat classification fusing temporal and morphological information of ECG via ensemble of classifiers, Biomed. Signal Process. Control 47 (2019) 41–48.
[52] Z. Masetic, A. Subasi, Congestive heart failure detection using random forest classifier, Comput. Methods Programs Biomed. 130 (2016) 54–64, https://doi.org/10.1016/j.cmpb.2016.03.020.
[53] M. Mohanty, P. Biswal, S. Sabut, Machine learning approach to recognize ventricular arrhythmias using VMD based features, Multidimens. Syst. Signal Process. (2019), https://doi.org/10.1007/s11045-019-00651-w.
[54] G.T. Taye, E.B. Shim, H.-J. Hwang, K.M. Lim, Machine learning approach to predict ventricular fibrillation based on QRS complex shape, Front. Physiol. 10 (2019) 1193, https://doi.org/10.3389/fphys.2019.01193.
[55] M. Mohanty, S. Sahoo, P. Biswal, S. Sabut, Efficient classification of ventricular arrhythmias using feature selection and C4.5 classifier, Biomed. Signal Process. Control 44 (2018) 200–208, https://doi.org/10.1016/j.bspc.2018.04.005.

[56] Y. Xu, D. Wang, W. Zhang, P. Ping, L. Feng, Detection of ventricular tachycardia and fibrillation using adaptive variational mode decomposition and boosted-CART classifier, Biomed. Signal Process. Control 39 (2018) 219–229, https://doi.org/10.1016/j.bspc.2017.07.031.

[57] N. Ibtehaz, M.S. Rahman, VFPred: a fusion of signal processing and machine learning techniques in detecting ventricular fibrillation from ECG signals, Biomed. Signal Process. Control 49 (2019) 349–359, https://doi.org/10.1016/j.bspc.2018.12.016.

[58] M. Hotradat, K. Balasundaram, S. Masse, K. Nair, K. Nanthakumar, K. Umapathy, Empirical mode decomposition based ECG features in classifying and tracking ventricular arrhythmias, Comput. Biol. Med. 112 (2019) 103379, https://doi.org/10.1016/j.compbiomed.2019.103379.

[59] R. Gale, A. Hochhaus, M.J. Zhang, What is the (p-) value of the *p*-value? Leukemia 30 (2016) 1965–1967, https://doi.org/10.1038/leu.2016.193.

[60] G. Benyamin, C. Mark, The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial, 2019. arXiv:1905.12787.

[61] S. Kim, W. Lee, Does McNemar's test compare the sensitivities and specificities of two diagnostic tests? Stat. Methods Med. Res. 26 (1) (2016) 142–154, https://doi.org/10.1177/0962280214541852.

[62] R.A. Halder, R. Mandal, A. Konar, Fuzzy based hierarchical algorithm for template matching in emotional facial images, J. Intell. Fuzzy Syst. 24 (2013) 201–214, https://doi.org/10.3233/IFS-2012-0547.

[63] S.A. Wilbert, Treatment of ventricular arrhythmias. Cardiac arrhythmias – mechanisms, Pathophysiol. Treat. (2014), https://doi.org/10.5772/57545.

[64] R.L. Wasserstein, N.A. Lazar, The ASA's statement on *p*-values: context, process, and purpose, Am. Stat. 70 (2) (2016) 129–133, https://doi.org/10.1080/00031305.2016.1154108.

[65] J. Fan, D.K.Y. Yau, A.K. Elmagarmid, W.G. Aref, Automatic image segmentation by integrating color-edge extraction and seeded region growing, IEEE Trans. Image Process. 10 (10) (2001) 1454–1466, https://doi.org/10.1109/83.951532.