Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 171 (2020) 679–688

Procedia Computer Science

www.elsevier.com/locate/procedia

Third International Conference on Computing and Network Communications (CoCoNet'19)

Classification of ECG Heartbeat Arrhythmia: A Review

Jagadeeswara Rao Annam[a], Srinivas Kalyanapu[a], Sureshbabu Ch.[a], Jayaprada Somala[a], S. Bapi Raju[b]

[a]*Gudlavalleru Engineering College, Gudlavalleru, AP., India, ajagarao@gmail.com*
[b]*IIIT, Gachibowli, Hyderabad - 500 032, India*

Abstract

Manual identification of ECG heart-beat classes by cardiologists is time consuming and cumbersome. These professionals rely on computer based methods for determination of these heart-disease types. In this work, existing literature is organized into a proposed taxonomy based on dichotomies involving **full time series-based** versus **feature-based**, AAMI versus Non-AAMI, and inter-patient versus intra-patient based distinctions. The basic contributions of this work are systematic review of literature on heart-beat abnormality detection, identifying research gaps and the research issues unmet sofar in the literature to propose novel approaches for addressing these gaps.

© 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the Third International Conference on Computing and Network Communications (CoCoNet'19).

Keywords: ECG; AAMI; MIT-BIH; clustering; classification;

1. Introduction

The cardiac rhythm abnormalities are different in shape from the normal rhythm of the heart-beat cycle and are known as heart-beat arrhythmia. And the symptoms of these heart diseases are sparse and infrequent. So the analysis of wearable long-term ECG recordings over hours, days and months is obligatory for detection of these infrequently occurring symptoms of heart diseases that would not be detected with short-term ECG recordings. Manual identification of these heart-beat classes by cardiologists is time consuming and cumbersome. These professionals rely on computer based classification for determination of these heart-disease types. To identify the potential gaps and un-met needs in the ECG heart-beat arrhythmia analysis problem, the past studies are organized into a taxonomical framework.

Conforming to Liao [13] and Patrick [24], the works are organized into two major categories as either raw **time series-based** approaches or **feature-based** approaches. Time series-based approaches directly use the original time series data and feature-based approaches indirectly use the original time series to compute the required features. As

* Corresponding author. Tel.: +91-9849401785; Email: ajagarao@gmail.com

1877-0509 © 2020 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Peer-review under responsibility of the scientific committee of the Third International Conference on Computing and Network Communications (CoCoNet'19).
10.1016/j.procs.2020.04.074

Patrick [24] characterized these approaches by **shape-based vs structure-based** respectively. In the next level the two main paradigms of training scheme **inter-patient vs intra-patient** [15] categorization is followed. Intra-patient supervised classification partitions the current patient's data set into training and testing subsets. Where as inter-patient supervised classification scheme uses independent training and testing sets from different patients to assess strong predictive generalization capability for unknown patient's data. And some works used two-part mixture of experts (MOE) [12] training to utilize patient adaptation: a general training set extracted from benchmark datasets and a specific patient training set taken from the patient under testing. This conforms to the AAMI guidelines that permits the utilization of first 5 minute data from the patient's record for training the classifier. And further, the previous works are segregated based on the heart beat classes they used, either classes using AAMI or non-AAMI classes of 15 beat types or less from MIT-BIH beat types.

Salem [23] summarized an extensive review on machine learning approaches to ECG classification by evaluating the features they use, their classification accuracies and the diseases to which they are applied. And Salem concluded that there is *a need for a systematic study* to examine the accuracy of variety of classifiers on a *single dataset* to directly compare each method. The current work attempts to fulfill this need.

2. Literature Review of ECG Analysis

Figure 1 presents the proposed taxonomy of solutions for the arrhythmia analysis problem using a tree diagram in which each leaf node indicates the representative publications for that path. Detailed list of publications for classification are given in Table 1.The following subsection present the review on classification solutions published in the literature.

2.1. Arrhythmia Classification approaches

This subsection discusses various solutions proposed in the literature for ECG classification and summarized in Table 1.

Hu, Palreddy and Tompkins [8] presented a hybrid of two different classifiers as mixture of experts (MOE), originally proposed by Jacobs [12]. The task of MOE as a linear combination of diverse estimates, performs better than any individual estimate. Two classifiers one is trained on general set (general classifier) and another on patient-personalized classifier are combined to form a third classifier called MOE classifier. A voting output by a gate network, dynamically weights the classification results of both classifiers to evolve a collective decision in MOE classifier. A 29 sample (i.e, 0.16 milli second of 180 samples/sec sampling rate ECG signal) QRS time series is extracted taking 14 points on each side of the R point and reduced to 9 points using Karhunen–Loeve (KL) transform for principle components. A 12 element feature vector using nine features corresponding to KL transform and the next three elements representing the interval features, was constructed for each beat. First a self organizing map (SOM) is applied on training data and its output was pipelined to learning vector quantization (LVQ) based classifier. The weight update rule used in SOM for the neighbouring ith neuron n_{ci} is $w_i(t+1)$:

$$w_i(t+1) = w_i(t) + n_{ci}(t)[x(t) - w_i(t)] \tag{1}$$

$$c = arg\min_j\{||x(t) - y_j(t)||\} \tag{2}$$

$$y_c(t+1) = [1 - f(t)\alpha_c(t)]y_c(t) + f(t)\alpha_c(t)x(t) \tag{3}$$

In LVQ y_c is updated as shown in 3 where feedback, f(t) =1 for true positive train data or -1 for false positives, α_c is a learning rate parameter. Taking 13 records each of 30 minutes are used for training the general classifier, first 2.5 minute data from 20 test records are used for training the patient-specific classifier and second 2.5 minute data of 20 test records are used for training the MOE classifier. Each of the classifiers uses both SOM and LVQ stages using the publicly available code from Helsinki univeristy[1] And MOE outperformed the other two classifiers having an overall average sensitivity, positive predictive value and classification rates as 82.6%, 77.7% and 94% respectively.

[1] ftp://cochlea.hut.fi/pub/

Chazal, Dwyer and Reilly (2004) [5] proposed a hybrid paradigm using two linear discriminant analysis (LDA) classifiers in a pure inter-patient supervised classification approach. Seven interval features were calculated taking 4 features from R-to-R intervals and 3 features from intra-beat segments. 19 morphological amplitudes were extracted for each beat time series that includes 10 samples between QRS_{onset} and QRS_{offset} and 9 samples from QRS_{offset} to $Twave_{offset}$ by uniformly sampling the segments. Using this 26 element feature vector (7 intervals and 19 amplitudes) for each beat, 12 different feature combinations were experimented with normalization, with-out normalization of features, using two channel data available in MIT-BIH database. The achieved sensitivities (Se%) were 86.8, 75.9, 77.7, 89.4 and positive predicted values (PPV%) 99.1, 38.5, 81.5, 7.8 for AAMI based classes N, S, V and F respectively. Chazal has pointed out that the performance in detecting AAMI S class is low compared to AAMI V class and he attributed that there were 944 AAMI S beats in training data which was less compared to 3788 AAMI V beats available in the training data.

Chazal and Reilly (2006) [2] also proposed a hybrid paradigm using linear discriminant analysis (LDA) classifier. A general (global) classifier that was first trained and a patient specific (local) classifier was then employed to tune the global classifier. This scheme seems to have outperformed those solutions using purely inter-patient classifiers. They addressed the issue of unbalanced class distribution by introducing weights using the training data in their proposed model. The aforesaid 26 feature set used in their previous work was employed in experimentation. For patient specific training 500 beats from testing set, in addition to the training set beats, were also used during training, to improve the results compared to the aforesaid previous work by the same authors. Using AAMI grouping the reported results on test dataset of 22 patients were Se% of 94.2, 87.7, 94.3, 73.9 and PPV% of 99.3, 46.9, 94.3, 29.1 for AAMI based N, S, V and F classes respectively. Compared to their earlier work [5] that used pure inter-patient classification, this approach which used patient-specific training, improved with large margins both in Se% and PPV% except the fusion (F) class.

Park, Cho, Lee and Song (2008) [21] presented a hierarchical SVM to reduce the effect of unbalanced classes using AAMI classes and the data division scheme used by Chazal et al. [5]. Taking 90 samples (250 milli seconds signal at 360 samples per second rate) at either side of R peak in each beat, a 181 sample time series were segmented for feature extraction for each beat. Using higher order statistics (HOS), cumulants of order two, three and four were calculated for each heart-beat using 181 sample beat time series. From 181 sample cumulant time series, 10 samples were extracted at multiples of 15, $\{C_{i,15xj} | j = 1..10, i = 2, 3, 4\}$ from each of second, third and fourth order cumulant time series. Using hermite basis functions (HBF) of order 20, 15 hermite coefficients were extracted as features. Three features were computed using RtoR intervals. A one-stage and hierarchical two-stage classifications were experimented using different combinations of HBF features, HOS features and RR-interval features. In two-stage approach, first stage discriminates the four classes into two groups each containing two classes, N and S in one group, V and F in second group. The second stage discriminates each of the two classes in each group. The reported sensitivities in hierarchical two-stage approach by SVM using HBF and interval features for N, S, V and F were 86.3, 82.6, 80.9, 54.9 respectively.

Ince, Kiranyaz and Gabbouj (2009) [11] proposed a multi dimensional swarm optimization (PSO) neural network in hybrid MOE paradigm where the training data contains both common representative beats randomly selected from the training recordings and patient-specific beats segmented from the first 5 minutes of each recording conforming the AAMI practice. Feature vector of 180 sample (500 millisecond at 360 samples per second sampling rate) centered at R-peak, is extracted from discrete wavelet transform (DWT) at scale 2^4. To tackle time-varying problem of discrete wavelet transforms, they chose time invariant discrete wavelet transform (TI-DWT) proposed by Mallat and Zhong for wavelet. Sensitivities% reported were 97.1, 62.1, 83.4, 61.3 for N, S, V and F classes respectively and PPV% are reported for S and V as 98.3 and 97.4% respectively.

Eduardo, Thiago, Victor, Joao & David (2011) [15] applied a supervised optimum path forest (OPF) based classification using a connected acyclic graph based on minimum spanning tree (MST). Each node represents a pattern in complete graph and distances between them are represented by edges. An acyclic connected graph is constructed whose nodes are all patterns of Z_1 using undirected arcs weighted by distances d between adjacent patterns. Using image forest transform, the key prototypes are exercised to cut the graph into optimum-path trees (OPT) that are descendants at each prototype using a path-cost function f_{max} by computing a MST in the complete graph Z_1. OPF was compared with MLP, SVM and Bayesian approaches on 6 different datasets extracted that are used by different researches namely Chazal et al. [5]; Guler and Ubeyli [7]; and Yu & Chou [29]. Interval features used by Chazal et al., seven level wavelet coefficients used by Song et al., 33 coefficients of ICA from 200 samples of each beat used by Yu and Chou, mean and standard deviation based features used by Guler and Ubeyli, 26 features composed of wavelet and PCA used by Ye et al., and 10 features from three level wavelet sub-bands used by Yu and Chen are extracted for comparison. In terms of accuracy, SVM is shown superior and is followed by MLP than others. And both OPF and bayesian are shown similar performances. And SVM with interval features used by Chazal are reported with Se% of 99.6, 0, 48.0, 48.7, 0 for 5 AAMI classes.

Llamedo and Martinez [14] (2011) proposed a linear discriminant classifier (LDC) using a compensated weight scheme (LDC-C) to represent minority classes. Llamedo explored features that present in two leads of MITBIH database using the vector-cardiogram magnitude (VCG_M) and angle (VCG_ϕ). The features used are grouped into three categories. First group were extracted from the ECG signal, the second group were taken from the 2-dimensional VCG formed using two lead data and the third group were taken from the discrete wavelet transform (DWT) using the fourth scale. Autocorrelation signal of DWT using both leads were computed to obtain locations of zero-crossings using 330 millisecond window around R point of each beat.

Using an alternating feature selection approach consisting of two steps namely forward selection (FS) and backward elimination (BE) from extracted 36 features, 8 best features are modeled. And using 3 AAMI categories by merging V class with F class, as F class are of very insignificant in number, results were reported on N, S and V' classes and obtained 46283 True positives from the used 49629 test beats. The sensitivities reported were 95, 77, 81, PPV were 98, 39, 87 for N, S and V' respectively. And TCA as global accuracy (A) was 93.0%. Lannoy experimented the classifier using top 3, top 5, top 6 and top 9 features. Sensitivities reported were 79.8, 92.6, 85.2, 84.5 for N, S, V, F respectively and BCR of 85.39. BCR is defined as the geometrical mean of the class accuracies.

Alvarado, Choudur and Principe (2012) [1] proposed a time representational encoding of ECG using integrate and fire (IF) sampling technique for discrimination of AAMI classes using linear discriminant approach. The ECG signal is convolved with a leaky averaging function $l(t)$ and the output is compared for positive and negative thresholds θ_p, θ_n. And if the output signal exceeds threshold, a pulse is generated for a refractory period τ. And the period τ limits the pulse rate independent of input. And then the output is transformed into bins of 35 ms . (N + 4)-dimension. Considering 20 features from each IF time-series beat, four features from RR-intervals, 24 dimensional feature data were obtained.

$$\theta_k = \int_{t_k+\tau}^{t_k+\tau} x(t)\, e^{\alpha(t-t_{k+1})}\, dt \tag{4}$$

$$l_k = e^{\alpha\ (t-t_{k+1})} \tag{5}$$

Classification by LDA is based on the posterior probability of class membership of a new pattern.

$$log(P(k|x)) \approx C - \frac{1}{2}(x-\mu_k)^T \Sigma_p^{-1} (x-\mu_k)^T + log\pi_k \tag{6}$$

The results reported were Se% 94.24, 86.19, 92.4, 66.4 on 38712 beats of test data. PPV were reported 56.68 and 94.82 for S and V respectively.

Martis, Acharya, Lim and Suri (2013) [18] employed "feed-forward neural network" (NN) and support vector machine (SVM) on three types of features independently. Principal components are extracted from ECG signals, linear prediction coded (LPC) based error and discrete wavelet transform (DWT). Principal components from ECG beats as features is shown superior with Se% of 99.9%, Sp% of 99.1%, PPV% of 99.6% and accuracy of 98.1%.

Huang, Liu, Zhu, Wang and Guangshu [10] proposed a hybrid of support vector based ensemble and a decision rule for classification of arrhythmia in AAMI framework. They used fifteen support vector machines to tackle the underlying 15 sub-classes of 5 AAMI classes using random projection based features in an ensemble fashion. They used 0.556 second data taking 0.278 seconds data from either side of R point, for each beat. (0.556 second data at 360 samples per second sampling frequency gives 200 samples). Using a random matrix $A \in R^{50\times200}$, the data matrix, $X \in R^{n\times200}$ is projected to obtain feature matrix $F \in R^{n\times59}$ using where $F = XA^{\dagger}$ where $F = XA^{\dagger}$ is a pseudo-inverse of A. By generating 15 feature matrices applying radial basis function kernel in SVM ensemble and combined with a decision rule on R to R previous interval, a two-stage hybrid of ensemble SVM and a decision rule was proposed. They achieved sensitivities of 99.2, 91.1, 93.9 and PPV% of 95.2, 42.2, 90.9 for N, S, V classes respectively.

Zhang, Dong, Luo, Choi & Wu (2014) [30] proposed a four-class AAMI based SVM classifier using one-vs-another (OVA) SVM classifiers. As selection of unique 2 class configurations from 4 classes require $C(4,2)$=6 combinations, so 6 one-vs-another (OVA) binary SVM classifiers were applied using a majority voting strategy. First each feature k is ranked in a supervised manner by dividing the patterns into positive, P_k and negative, N_k classes for a given combination of 2 classes using F_{score} as follows.

$$F\_score(k) = \frac{(P_k - X_k)^2 + (N_k - X_k)^2}{\frac{1}{n_p-1}\sum_{i=1}^{n_p}(P_{ik} - X_k)^2 + \frac{1}{n_n-1}\sum_{j=1}^{n_n}(N_{jk} - N_k)^2} \tag{7}$$

where P_k, N_k and X_k are the means of the positive, negative and whole instances, P_{ik} and N_{jk} are the i'th and j'th instances of k'th feature from the positive and negative instances for each two class combination. The classification results reported are Se%88.9, 79.1, 85.5, 93.8 and PPV% 98.9, 35.9, 92.8, 13.7.

Das and Ari (2014) [4] proposed a LMS-based SVM classifier on AAMI grouped classes using patient specific training. Using Pan-Tompkins algorithm the fiducial points Q, R and S were delineated for each beat. A 180 sample time-series (500 millisecond of 360 samples per second sampling rate) centered at R point, was considered for each beat for calculation of discrete S-transform based features. The features from intervals and S-transform were extracted for each ECG beat. Four features from RR-intervals were extracted and by concatenating to 180 sample S-transform features a 184 dimensional feature vector was constructed. A bacteria foraging optimization (BFO) was used for feature reduction. using patient-specific data containing 245 general training beats were used SVM using RBF-kernel in one-against-rest (OAR) 25 min beats Results reported were Se% of 98.1, 74.0%, 91.4% and 67.0 respectively and PPV% of 97.9%, 73.6%, 91.4% and 80.9% for N, S, V and F classes respectively.

2.2. Non-AAMI based classification of ECG

Guler and Ubeyli (2005) [7] proposed a 2-stage stacked neural network to classify four non-AAMI classes. Taking 256 sample time series (0.71 second data of 360 sample per second rate) from each beat Daubechies wavelet transform of order 2 was applied to compute four level wavelet coefficients. Using sub-bands of 4 levels means, standard deviations, average powers and ratios of means of adjacent sub-bands were extracted for each beat to form 19 element feature vector. In second stage NN, the number of outputs was four and the number of hidden neurons was chosen to be 30. Features from symmlet of order 6 were also experimented for comparison. Total classification accuracy was reported 96.94%.

Song, Lee, Cho, Kyoung and Sun (2005) [25] presented SVM and MLP using linear discriminant (LD) based feature selection on six classes of MIT-BIH. Seventeen features were extracted by wavelet transform and were reduced to 4 features, the linear combination of original features by LDA. And SVM classifier is shown superior to MLP classifier. The classification results reported for normal, arterial premature contraction (A), supra-ventricular tachycardia (SVT), ventricular premature contraction (V), ventricular tachycardia (VT) and ventricular fibrillation (VF) were 99.3%, 99.2%, 99.8%, 98.3%, 99.4% and 99.8% respectively.

Yu and Chou (2008) [29] proposed probabilistic neural network (PNN) and support vector machines (SVM) using independent component analysis (ICA) based features on eight classes of MIT-BIH. A FastICA algorithm was used to estimate the demixing matrix B for calculation of non-gaussian components represented by IC matrix, I. Here the feature vector, $X \in R^{n\times 1}$, $B \in R^{m\times n}$ and $I \in R^{n\times 1}$ and I is calculated as follows: X = A.I, I = B.X and $I_k = \sum_{i=1}^{m} b_{ik}X_i = B_k.X$. The overall accuracy reported was 99.66%.

Christov, Gomez, Krasteva, Jekova, Gotchev and Egiazarian (2006) [3] presented a k-nearest neighbour on two types of heartbeat features. First QRS recognition method was performed for computation of 26 morphological features (MF). Considering 90 samples before R and 165 samples after R, a total of 256 sample time series including R-point, (710 ms at sampling frequency of 360 samples per second) the heart beat time-series were extracted. A matching pursuits approach was applied on 256 sample time series for calculation of 10 features from time-frequency (TF) domain and one residual energy feature. Classification results were reported using MF and TF separately for N, LBBB, RBBB, PVC and PB classes. The sensitivities were reported for features of MF as 98.5, 99.1, 97.3, 96.27, and 99.84 respectively and using TF are 98.8, 98.9, 98.7, 94.7 and 99.9 respectively. PPV% on MF were reported as 99.6, 98.6, 96.3, 89.8, 97.97 and TF based features were 99.4, 99.6, 99.8, 89.2 and 99.5 respectively.

Ubeyli [27] presented a recurrent neural network (RNN) using Levenberg - Marquardt (LM) algorithm to classify four classes of MIT-BIH data. For each ECG beat of 256 point time series, 128 Lyapunov exponents were extracted using Jacobi matrices method. Distances of two points in phase space at time 0 and t being $|\delta S_j(0)|$ and $|\delta S_j(t)|$, the Lyapunov exponent in direction j, λ_j is defined as

$$\lambda_j = \lim_{t\to\infty} \frac{1}{t} \log_2 \frac{|\delta S_j(t)|}{|\delta S_j(0)|} \tag{8}$$

From 128 exponents of each beat minimum, maximum, mean and standard deviation were extracted as features. The results reported considering 360 beats for training and 360 beats for testing to classify four classes of beats namely normal, ‘congestive heart failure (CHF)’, ventricle tachy-arrhythmia (VT) and atrial fibrillation (AF) are Se% 94.44, 93.33, 95.56, 95.56 and Sp% 99.61, 98.09, 96.59 and 98.46, respectively with a total classification accuracy of 94.72%.

Babak, Seyed, and Maryam (2008) [20] presented a SVM classifier using a proposed ‘generalized discrimination analysis based feature selection’ (GDAFS). Heart rate variability (HRV) is extracted from ECG time series taking R-to-R intervals from every two consecutive R points. By depicting time verses RR intervals, divided into segments each containing 32 RR intervals. And 15 features are computed. GDAFS converts the feature data into high dimension space F so that N classes will become linearly separable. Using eigen decomposition, the N-1 eigen vectors assiciated with N-1 non-zero largest eigen values are computed from Kernel matrix. SVM was applied using the selected 5 features i.e N-1, to classify N=6 classes namely normal, PVC, AF, ‘sick sinus syndrome (SSS), ventricle fibrillation (VF) and second degree block. The accuracies % of 98.9, 98.96, 98.53, 98.51, 100 and 100 respectively were reported.

Ye, Coimbra and Kumar [28] presented a Support Vector Machine (SVM) based classifier considering all 15 classes of MIT-BIH data. From each beat, 300 sample time series (0.83 seconds) were used taking 100 points (0.27 seconds data) before R and 200 points (0.56 seconds data) after R point. Features from discrete wavelet transform (DWT) and independent component analysis (ICA) were extracted separately for each beat. Any function f(x) $\in L^2(R)$ can be represented by the wavelet function $\psi(x)$ and its scaling function $\phi(x)$ as follows:

$$f(x) = \exp\left(\sum_k A_J(k)\, \phi_{Jk}(x) + \sum_{j=J}^{\infty} \sum_k D_j(k)\, \psi_{jk}(x)\right) \tag{9}$$

J is the initial scale, A and D refers to the approximation and detail coefficients. Wavelets coefficients of level-3 and level-4 were extracted with a total of 118 wavelet features for each beat. Using the FastICA algorithm, 18 ICA coefficients were extracted from each heartbeat. A principle component analysis (PCA) approach was applied to reduce the set of 136 features (118 wavelet + 18 ICA) to 26 relevant features. Using 4 RR based interval features, a feature vector of 30 elements (26+4) was used in gaussian radial based kernel SVM for classification. The experimental results were reported with an average accuracy of 99.25%.

Huang, Liu, Zhu, Wang and Guangshu [9] proposed a mixture of 3 classifiers using majority vote approach to discriminate three classes namely Normal, LBBB and RBBB using inter-patient supervised classification approach. The classes N and RBBB were applied on a weighted LDA classifierr. LBBB and RBBB are applied on a weighted SVM. LBBB and N are applied on a minimum distance classifier (MDC). The annotation data available in the dataset for R-peak positions for each beat were considered to segment each beat. Different feature sets were used independently on these three classifiers. A 0.278 second signal of 360 sample per second discrete ECG data (100 samples) were extracted on either side of R-point so-that 200 samples timeseries were taken for each beat from each of the two-channel data. A 401 dimensional feature vector that includes a RR interval feature, was constructed as input to MDC. For LDA and SVM classifiers, 200 sample timeseries of each beat were reduced to 100 by using independent components analysis (ICA)code available at http://research.ics.aalto.fi/ica/fastica/ algorithm. For the vector $\mathbf{x} = \{x_1, x_2, .., x_m\}^T$ to be transformed to $\mathbf{y} = \{y_1, y_2, .., y_n\}^T$ by using an estimated weight matrix, $W \in R^{n\times m}$ the ICA is formulated as: $y = Wx$ The results reported were sensitivities 81.4%, 91.4% and 92.8% and positive predictive values of 98.0%, 37.3% and 88.8% for N, L and R beats respectively.

Tran, Pham and Vuong (2014) [26] used three neural approaches namely multi layer perceptron (MLP), support vector machine (SVM) and a ‘modified Takagi–Sugeno–Kang (TSK) network’ using coefficients of 15-th order Hermite basis functions (HBF) and the RR durations from each beat. A 250 millisecond time series of 91 samples, taking 45 samples around R point of each beat is segmented to extract HBF coefficients.Then ‘iterative dichotomiser 3 decision tree (ID3)’ was applied to process the outputs of three models to perform the final processing stage. Using 19 patients in non-AAMI based normal versus abnormal classes the reported results were Se% of 99.02 and Sp% of 99.25.

Mateo, Torres, Aparicio and Santos (2016) [19] proposed a radial basis function neural net (RBFNN) to discriminate ectopic originated beats from non-ectopic originated beats using 2 class patient-specific (intra-patient) classification. Two different activation functions namely gaussion RBF (GRBF) and raised-cosine RBF (RCRBF) as shown in 10 and 11 were used in experimentation. Here, $m_j(t)$ is the center or mean that was assigned by a fuzzy-c-means clustering approach.

$$\phi_j(t) = \exp\left(-\frac{\|(x(t) - m(t))\|}{2\sigma^2(t)}\right) \tag{10}$$

$$\phi_j(t) = \begin{matrix} \frac{1}{2}\left(1 + cos\left(\frac{\|(x(t) - m(t))\|}{\sigma(t)}\right)\right) |(x(t)| \leq \sigma(t) \\ 0|(x(t)| > \sigma(t) \end{matrix} \tag{11}$$

$$y_k(t) = \sum_{j=1}^{M} w_{kj}\phi_j(t) \tag{12}$$

And $\sigma(t)$ is the variance of input. And $y_k(t)$ is the output with M basis functions. Using intra-patient training scheme for 2 classes of ectopic beats and non-ectopic beates, the reported results were Se% 99, 72 and Sp% 99, 87.

Table 1 presents the publications on classification of arrhythmia. And in the next section ECG Clustering solutions in the Literature are detailed. Based on the summary of various research solutions proposed for classification and clustering of ECG time-series in the previous sections, research gaps and issues based on performance are identified and summarized in the next section 3.

Inter-patient supervised classification schemes can directly use labeled training set from the annotated database repositories like MIT-BIH arrhythmia dataset and test or predict the class label for the test data of any new patient. This avoids the tedious, impractical need for manual labeling of the beats of the patient who is under diagnosis. But in clinical practice the classification performance of these methods trained on specific data sets, especially the annotated benchmark datasets, declines on new patient data due to *inter-patient variations* because of the patient-specific characteristics. In intra-patient data, each class may have variations and different classes have similarities. By organizing the extant literature in the form of a taxonomy, gaps could be easily identified. And Table 2 summarizes the issues observed during the literature review. It is clear from the literature survey that there are three main challenging issues in heartbeat classification, namely, inter-patient beat differences (variability) for the same class, intra-patient beat similarities across different classes [1] and unbalanced data class problem because of the sporadic or sparse or intermittent disease symptoms. In addition, some leads or channels (the sensors) of of ECG may not capture the variations or discriminations across different classes [22], adding to the complications.

Among the research solutions proposed in the literature, there seems to be room for pushing the classification accuracy to the ideal 100% accuracy to further boost confidence for the usage of such solutions in clinical diagnosis [5]. Another issue is the low positive prediction value (PPV) of *atrial premature contraction (APC)*, an abnormal heart condition [14], possibly because the features of APC are very similar to those of *healthy beats*.

3. Research Issues and Gaps Identified in Literature

Intra-patient supervised classification partitions the current patient’s data set into training and testing subsets. Using this scheme, although the classifiers usually produce over-optimistic results, however manual class labeling required for training set is time consuming in emergency conditions. As discussed earlier, supervised approaches for inter-patient classification are challenged by large inter-subject variations. On the other hand intra-patient approaches are

Table 1: ECG Classification Solutions in the Literature

Method	Features	Classifier, classes	Train Data Test Data	Training Scheme	Results Se, PPV
Hu [8] (1997)	Intervals Morphology(12)	SOM+LVQ (12) KL-PCA, AAMI	15 records Train 20 records Test	First 5-min Patient-specific	Overall Se% 82.6 Overall PPV% 77.7
Chazal[5] (2004)	Amplitudes Intervals	Hybrid LD lead A +lead B	22 records train 22 records test	AAMI Inter Patient	86.8, 75.9, 77.7, 89.4 99.1, 38.5, 81.5, 8.6
Chazal[2] (2006)	Amplitudes Intervals	Hybrid LD general+specific	DS1+500 beats 22 records test	AAMI Patient-specific	94.2, 87.7, 94.3, 73.9 99.3, 46.9, 94.3, 29.1
Park [21] (2008)	RR intervals HOS and HBF	SVM 250 ms	Train 22 Rec (30 min) Test 22 Rec (30 min)	AAMI Inter Patient	86.3, 82.6, 80.9, 54.9
Ince [11] (2009)	wavelet TI-DWT intervals, PCA	MD PSO ANN 5-min train set	20 records train 24 records test	AAMI patient-specific	Se% 97, 84.6, 63.5,61 Acc%:-, 97.4, 98.3, -
Eduardo [15] (2011)	Intervals 6 methods	LDC,OPF,SVM	Chazal scheme 22 + 22	AAMI Inter Patient	99.6, 0, 48.0, 48.7, 0
Llamedo [14] (2011)	Intervals DWT scale-4	LDC projection pursuit	22+22 Inter Patient	AAMI Tca: 78	77.5, 76.4, 82.9, 95.3 99.4, 41.3, 88.0, 4.0
Mar [16] (2011)	Amplitudes ecgpuwave	MLP, RBF classifier	Chazal scheme 22+22	AAMI Inter Patient	89.6, 83.2, 86.8, 61.1 99.3, 33.5, 75.9, 16.6
Lannoy [6] (2012)	Intervals, HOS Morphology	wCRF $+L_1$ MI ranking	Chazal scheme 22+22	AAMI Inter-patient	79.8, 92.6, 85.2, 84.5 BCR% 85.4
Alvarado [1] (2012)	Amplitudes IF sampling	LDA	Chazal	AAMI TCA: 93.6	Se:94.2, 86.2, 92.4, 66.4 Ppv: S:56.68, V: 94.82
Martis [17] (2013)	Amplitudes DCT, PCA	PNN,FNN, SVM Six features	Non Chazal 10 subsets of 44 rec	AAMI intra-patient	Se 98.69% Sp 99.91%,
Huang [10] (2014)	Amplitudes, RR Random proj (51)	Ensemble SVM 15 svms+Rule	Chazal scheme 22+22	3 class AAMI Inter-patient	se 99.2, 91.1, 93.9, ppv 95.2, 42.2, 90.9,
Zhang [30] (2014)	Intervals Morphology	RBF SVM classifier	Inter-patient 22+22	AAMI Inter-patient	88.9, 79.1, 85.5, 93.8 98.9, 35.9, 92.8, 13.7
Das [4] (2014)	RR intervals 500 ms beats	RBF SVM S-transform	20 records train data 24 records testing	AAMI first 5 min intra-patient	98.1, 74.0, 91.4, 67.0 97.9, 73.6, 91.8, 80.9
Non-AAMI Classification					
Guler [7] (2005)	Intervals wavelet	CNN Intra-patient	Intra-patient Four classes	Non-AAMI 360 Testing	Sp: 97.78% TCA: 96.94%
Christov [3] (2006)	Intervals 256 points	k-NN VCG MF	First 3 to 12 min 424 beats general	Non-AAMI Intra-patient	98.5 99.6
Ubeyli [27] (2007)	wavelet-4 (70) Daub 2 (10)	RNN L-M training	small dataset intra-patient	Non-AAMI 4 class	Se% 94.44 Sp% 99.61
Babak [20] (2008)	Intervals, Morphology	SVM classifier	small dataset intra-patient	Non-AAMI 6 class	98.9, 98.9, 98.5, 98.5, 100, 100
Ye [28] (2010)	Amplitudes Wavelet	SVM 15	Non Chazal 10 subset intra	Non-AAMI 15 classes	Se: 99.91% 99.66%
Huang [9] (2014)	Intervals, 200, 278	LDA, SVM, MDC classifier	2 class Inter-patient	Non-AAMI 3 class	Se 91.4, 92.8 ppv 37.3, 88.8
Tran [26] (2014)	HBF Amplitudes (9)	SVM, MLP, TSK ID3 Decision Tree	19 patients of MITBIH	Non-AAMI 7 class	Se: 99.02% PPV: 99.25%
Mateo [19] (2016)	Amplitudes correlation	MLP raised-cosine RBF	small set Ectopic	Intra Patient non-AAMI	Se% 99, 72 Sp% 99, 87

[1] HBF: Hermite basis function, IE: Immune evolutionary, HOS: higher order statistics, KL: Karhunen–Loeve wCRF: Weighted Conditional Random Fields, EM: Expectation maximization, LVQ: Learning Vector Quantization, TSK: Takagi–Sugeno–Kang, ID3: Iterative Dichotomiser-3, PNN: Probabilistic Neural Net, CHF: Congestive Heart Failure, MI: mutual information

Table 2: Research Issues

#	Approach	Advantage	Disadvantage	Research Issues	Issues
1	Inter-patient Supervised Classification	No need of labels Test patient train labels	patient differences Low Performance	Intra-patient variations Inter-patient variations Low Performance	Issue 1
2	Intra-patient Supervised	High Performance	Need of patient train labels in emergencies also	Intra-patient variations	
3	Clustering	No Need of train labels	Parameters Low Performance	Intra-patient issue Parameters Low Performance	Issue 2

limited by the time consuming requirement for manual labeling of beats from the same patient, a task that becomes impractical in emergency scenarios. Further, as the symptoms of heart diseases are sparse, infrequent and intermittent, wearable *long term* ECG recorders are required to identify infrequently occurring symptoms of heart diseases that would not be detected with *short-term* ECG recordings. The analysis of long-term ECG recordings spanning over hours, days and months is obligatory for early diagnosis and detection of heart diseases. Manual identification of these

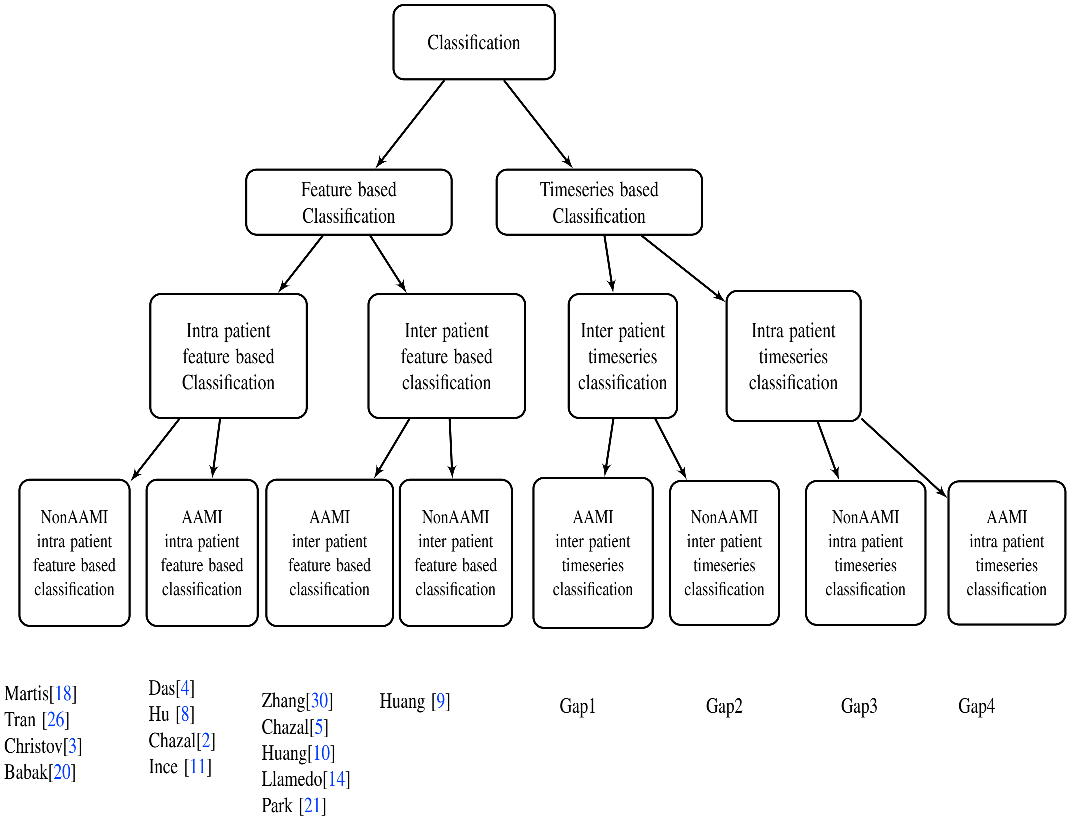

Fig. 1: Taxonomy of Research solutions in literature

heart-beat classes by cardiologists is time consuming, cumbersome and infeasible. Consequently, these professionals rely on computer-based automated classification for identifying various normal and abnormal heart-disease types. However, the sparse nature of such symptoms of heart disease poses the challenge of having to cope with unbalanced data class problem where normal class (majority) in data severely outnumbers the abnormal or heart disease (minority) class.

During the review work presented in this paper, various approaches are implemented to address the gaps identified from literature review for the problem of arrhythmia detection.

Since no publicly available feature sets are available for the ECG arrhythmia detection problem, researchers have to explicitly extract the feature sets from the raw signals available in Physio Bank repository. Both the Matlab code and resulting feature sets for all records have been deposited and made publicly available in the *Github*[2] repository. This is one of the major contributions of this work. The github repository offers a valuable resource for ECG researchers.

4. Conclusion

To identify the potential gaps and un-met needs in the ECG heart-beat arrhythmia analysis problem, the past studies are organized into a taxonomical framework. Various approaches proposed in the literature related to clustering and

[2] https://github.com/ajagarao/ECG-feature-set-of-46-features

classification of ECG heart-beat arrhythmia have been reviewed. Such organized literature survey allows researchers for the identification of gaps and the research issues unmet so-far. Gaps and unmet issues have been summarized so that the researchers could explore approaches for addressing them in this work.

References

[1] Alvarado, A.S., Lakshminarayan, C., Principe, J.C., 2012. Time-based compression and classification of heartbeats. IEEE transactions on biomedical engineering 59, 1641–1648.

[2] de Chazal, P., Reilly, R.B., 2006. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. IEEE transactions on biomedical engineering 53, 2535–2543.

[3] Christov, I., Jekova, I., Bortolan, G., 2005. Premature ventricular contraction classification by the kth nearest-neighbours rule. Physiological measurement 26, 123.

[4] Das, M.K., Ari, S., 2014. Patient-specific ecg beat classification technique. Healthcare technology letters 1, 98–103.

[5] De Chazal, P., O'Dwyer, M., Reilly, R.B., 2004. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE transactions on biomedical engineering 51, 1196–1206.

[6] De Lannoy, G., François, D., Delbeke, J., Verleysen, M., 2012. Weighted conditional random fields for supervised interpatient heartbeat classification. IEEE Transactions on Biomedical Engineering 59, 241–247.

[7] Güler, İ., Übeylı, E.D., 2005. Ecg beat classifier designed by combined neural network model. Pattern recognition 38, 199–208.

[8] Hu, Y.H., Palreddy, S., Tompkins, W.J., 1997. A patient-adaptable ecg beat classifier using a mixture of experts approach. IEEE transactions on biomedical engineering 44, 891–900.

[9] Huang, H., Liu, J., Zhu, Q., Wang, R., Hu, G., 2014a. Detection of inter-patient left and right bundle branch block heartbeats in ecg using ensemble classifiers. Biomedical engineering online 13, 72.

[10] Huang, H., Liu, J., Zhu, Q., Wang, R., Hu, G., 2014b. A new hierarchical method for inter-patient heartbeat classification using random projections and rr intervals. Biomedical engineering online 13, 90.

[11] Ince, T., Kiranyaz, S., Gabbouj, M., 2009. A generic and robust system for automated patient-specific classification of ecg signals. IEEE Transactions on Biomedical Engineering 56, 1415–1426.

[12] Jacobs, R.A., 1995. Methods for combining experts' probability assessments. Neural computation 7, 867–888.

[13] Liao, T.W., 2005. Clustering of time series data—a survey. Pattern recognition 38, 1857–1874.

[14] Llamedo, M., Martínez, J.P., 2011. Heartbeat classification using feature selection driven by database generalization criteria. IEEE Transactions on Biomedical Engineering 58, 616–625.

[15] Luz, E.J.d.S., Nunes, T.M., De Albuquerque, V.H.C., Papa, J.P., Menotti, D., 2013. Ecg arrhythmia classification based on optimum-path forest. Expert Systems with Applications 40, 3561–3573.

[16] Mar, T., Zaunseder, S., Martínez, J.P., Llamedo, M., Poll, R., 2011. Optimization of ecg classification by means of feature selection. IEEE transactions on Biomedical Engineering 58, 2168–2177.

[17] Martis, R.J., Acharya, U.R., Lim, C.M., Suri, J.S., 2013. Characterization of ecg beats from cardiac arrhythmia using discrete cosine transform in pca framework. Knowledge-Based Systems 45, 76–82.

[18] Martis, R.J., Acharya, U.R., Mandana, K., Ray, A.K., Chakraborty, C., 2012. Application of principal component analysis to ecg signals for automated diagnosis of cardiac health. Expert Systems with Applications 39, 11792–11800.

[19] Mateo, J., Torres, A., Aparicio, A., Santos, J., 2016. An efficient method for ecg beat classification and correction of ectopic beats. Computers & Electrical Engineering 53, 219–229.

[20] Mohammadzadeh, A.B., Setarehdan, S.K., Mohebbi, M., 2008. Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal. Artificial intelligence in medicine 44, 51–64.

[21] Park, K., Cho, B., Lee, D., Song, S., Lee, J., Chee, Y., Kim, I., Kim, S., 2008. Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function, in: 2008 Computers in Cardiology, IEEE. pp. 229–232.

[22] Rockstroh, C., 2013. Novel algorithms and rating methods for high-performance ecg classification .

[23] Salem, A.B.M., Revett, K., El-Dahshan, E., 2009. Machine learning in electrocardiogram diagnosis. international multiconference on computer science and information technology.

[24] Schäfer, P., 2014. Towards time series classification without human preprocessing, in: International Workshop on Machine Learning and Data Mining in Pattern Recognition, Springer. pp. 228–242.

[25] Song, M.H., Lee, J., Cho, S.P., Lee, K.J., Yoo, S.K., 2005. Support vector machine based arrhythmia classification using reduced features .

[26] Tran, H.L., Vuong, H.N., et al., 2014. Multiple neural network integration using a binary decision tree to improve the ecg signal recognition accuracy. International Journal of Applied Mathematics and Computer Science 24, 647–655.

[27] Übeyli, E.D., 2010. Recurrent neural networks employing lyapunov exponents for analysis of ecg signals. Expert systems with applications 37, 1192–1199.

[28] Ye, C., Kumar, B.V., Coimbra, M.T., 2012. Heartbeat classification using morphological and dynamic features of ecg signals. IEEE Transactions on Biomedical Engineering 59, 2930–2941.

[29] Yu, S.N., Chou, K.T., 2008. Integration of independent component analysis and neural networks for ecg beat classification. Expert Systems with Applications 34, 2841–2846.

[30] Zhang, Z., Dong, J., Luo, X., Choi, K.S., Wu, X., 2014. Heartbeat classification using disease-specific feature selection. Computers in biology and medicine 46, 79–89.