CHAPTER II
EXISTING SYSTEM

The investigation of the ECG has been extensively used for diagnosing many cardiac diseases. The ECG is a realistic record of the direction and magnitude of the electrical commotion that is generated by depolarization and repolarization of the atria and ventricles. Classification of arrhythmia is generally composed of the pre-processing part, the feature extraction and selection part and the classification part. Several researches have been carried out for preprocessing, feature extraction and selection and classification.

2.1. ECG SIGNAL PREPROCESSING TECHNIQUES

The pre-processing phase removes noise components and does other forms of processing for better accurate feature extraction or classification. The major noise components of ECG include baseline drift, power line interference and moving artifacts. Several authors have investigated on the filtering techniques for removing noise components at the same time preserving both ECG morphology and fast processing.

A novel approach for ECG feature extraction was put forth by Castro et al., in [79]. This paper presented an algorithm, based on the wavelet transform, for feature extraction from an ECG signal and recognition of abnormal heartbeats. Since wavelet transforms can be localized both in the frequency and time domains, developed a method for choosing an optimal mother wavelet from a set of orthogonal and bi-orthogonal wavelet filter bank by means of the best correlation with the ECG signal. The foremost step of this approach is to denoise (remove noise) the ECG signal by a soft or hard threshold with limitation reconstructs ability and then each PQRST cycle is decomposed into a coefficients vector by the optimal wavelet function. The coefficients, approximations of the last scale level and the details of the all levels, are used for the ECG analysis. This approach
divided the coefficients of each cycle into three segments that are related to P-wave, QRS complex and T-wave. The summation of the values from these segments provides the feature vectors of single cycles.

The shape of ECG is used to classify ECG beat in four types such as normal beat (N), left bundle branch block beat (L), right bundle branch block beat (R) and ventricular premature beat (V). To extract the shape of ECG, the discrete wavelet transform with level 3 of Daubechies 1 is used after digital filter was applied to remove noise from ECG signal by Thanapatay et al., [32]. After that Principle Components Analysis (PCA) and SVM are adapted to create model of classifier for using with paper based ECG printout. The ECG image from ECG printout is processed by some image processing techniques such as red grid removing, noise rejection, image thinning and time-series ECG extraction to obtain the time-series ECG signal before classification.

Increasing needs and possibilities for serial comparison of ECG computer analysis results bring up the question of the reproducibility and stability. Joseph et al., [51] report here results from systematic measurement manipulation experiments and corresponding changes in diagnostic statements of the HES algorithm. Adding noise or filtering the original data resulted in totally less than 4% changes in any of the diagnostic categories NOR, HYP and INF. Shifting the relevant wave onsets or offsets creates changes in diagnostic statements as well.

A rule-based rough-set decision system for the development of a disease inference engine is described by Mitra et al., [61]. For this purpose, an offline-data-acquisition system of paper ECG records is developed using image-processing techniques. The ECG signals may be corrupted with six types of noise. Therefore, at first, the extracted signals are fed for noise removal. A QRS detector is also developed for the detection of R-R interval of ECG waves. After the detection of this R-R interval, the P and T waves are detected based on a syntactic
approach. The isoelectric-level detection and base-line correction are also implemented for accurate computation of different attributes of P, QRS and T waves. A knowledge base is developed from different medical books and feedbacks of reputed cardiologists regarding ECG interpretation and essential time-domain features of the ECG signal. Finally, a rule-based rough-set decision system is generated for the development of an inference engine for disease identification from these time-domain features.

Sufi et al. in [81] formulated a new ECG obfuscation method for feature extraction and corruption detection. This paper presents a new ECG obfuscation method, which uses cross correlation based template matching approach to distinguish all ECG features followed by corruption of those features with added noises. It is extremely difficult to reconstruct the obfuscated features without the knowledge of the templates used for feature matching and the noise. Therefore, considered three templates and three noises for P wave, QRS Complex and T wave comprise the key, which is only 0.4%-0.9% of the original ECG file size. The key distribution among the authorized doctors is efficient and fast because of its small size.

A robust ECG feature extraction scheme was put forth by Olvera in [91]. This method utilizes a matched filter to detect different signal features on a human heart electrocardiogram signal. The detection of the ST segment, which is a precursor of possible cardiac problems, was more difficult to extract using the matched filter due to noise and amplitude variability. By improving on the methods used; using a different form of the matched filter and better threshold detection, the matched filter ECG feature extraction could be made more successful. The detection of different features in the ECG waveform was much harder than anticipated but it was not due to the implementation of the matched filter. The more complex part was creating the revealing method to remove the feature of interest in each ECG signal.
2.2. FEATURE EXTRACTION AND SELECTION TECHNIQUES

The feature extraction and selection part makes feature vectors that are used later in the classification part. Several signal compression algorithms are used to represent the signal's characteristics effectively with a small computational burden.

ECG feature selection and extraction has been studied from early time and lots of advanced techniques as well as transformations have been proposed for accurate and fast ECG feature extraction. This section discusses various techniques and transformations proposed earlier in literature for extracting feature from ECG.

The ECG feature extraction system provides fundamental features (amplitudes and intervals) to be used in subsequent automatic analysis. In recent times, a number of techniques have been developed to detect these features [73-75]. The previously proposed method of ECG signal analysis was based on time domain method. But this is not always adequate to study all the features of ECG signals. Therefore the frequency representation of a signal is required. The deviations in the normal electrical patterns indicate various cardiac disorders. Cardiac cells, in the normal state are electrically polarized [77].

One cardiac cycle in an ECG signal consists of the P-QRS-T waves. The majority of the clinically useful information in the ECG is originated in the intervals and amplitudes defined by its features (characteristic wave peaks and time durations). The improvement of precise and rapid methods for automatic ECG feature extraction is of chief importance, particularly for the examination of long recordings [75].

Zhao et al. [78] developed a feature extraction method using wavelet transform and support vector machines. This paper presented a new approach to the feature extraction for reliable heart rhythm recognition. This system of classification is comprised of three components including data preprocessing,
feature extraction and classification of ECG signals. Two diverse feature extraction methods are applied together to achieve the feature vector of ECG data. The wavelet transform is used to extract the coefficients of the transform as the features of each ECG segment. Concurrently, Autoregressive Modeling (AR) is also applied to get hold of the temporal structures of ECG waveforms. Then at last the SVM with Gaussian kernel is used to classify different ECG heart rhythm. The results of computer simulations provided to determine the performance of this approach.

Mahmoodabadi et al. in [73] described an approach for ECG feature extraction which utilizes Daubechies Wavelets transform. The approach had been developed and evaluated an ECG feature extraction system based on the multi-resolution wavelet transform. The ECG signals from Modified Lead II (MLII) were chosen for processing. The wavelet filter with scaling function further intimately similar to the shape of the ECG signal achieved better detection. The foremost step of this approach was to de-noise the ECG signal by removing the equivalent wavelet coefficients at higher scales. Then, QRS complexes are detected and each one complex is used to trace the peaks of the individual waves, including onsets and offsets of the P and T waves which are present in one cardiac cycle. Experimental results revealed that this approach for ECG feature extraction in terms of and predictivity.

Saxena et al in [82] described an approach for effective feature extraction from ECG signals. Competent composite method which has been developed for data compression, signal retrieval and feature extraction of ECG signals. After signal retrieval from the compressed data, it has been found that the network not only compresses the data, but also improves the quality of retrieved ECG signal with respect to elimination of high-frequency interference present in the original signal. With the implementation of Artificial Neural Network (ANN) the compression ratio increases as the number of ECG cycle increases. Moreover the
features extracted by amplitude, slope and duration criteria from the retrieved signal match with the features of the original signal. Experimental results at every stage are steady and consistent and prove beyond doubt that the composite method can be used for efficient data management and feature extraction of ECG signals in many real-time applications.

A feature extraction method using Discrete Wavelet Transform (DWT) was developed by Emran et al. in [83]. This method used a DWT to extract the relevant information from the ECG input data in order to perform the classification task. This approach includes the following modules data acquisition, pre-processing beat detection, feature extraction and classification. In the feature extraction module the Wavelet Transform (DWT) is designed to address the problem of non-stationary ECG signals. It was derived from a single generating function called the mother wavelet by translation and dilation operations. Using DWT in feature extraction may lead to an optimal frequency resolution in all frequency ranges as it has a varying window size, broad at lower frequencies and narrow at higher frequencies. The DWT characterization will deliver the stable features to the morphology variations of the ECG waveforms.

Jovic and Bogunovic in [85] developed chaos theory that can be successfully applied to ECG feature extraction. This paper also discussed numerous chaos methods, including phase space and attractors, correlation dimension, spatial filling index, central tendency measure and approximate entropy, and created a new feature extraction environment called ECG chaos extractor to apply the above mentioned chaos methods. A new semi-automatic program for ECG feature extraction has been implemented and is presented in this article. Graphical interface is used to specify ECG files employed in the extraction procedure as well as for method selection and results saving. The program extracts features from ECG files.
A method for automatic extraction of both time interval and morphological features, from the ECG to classify ECGs into normal and arrhythmic was described by Alexakis et al. in [89]. The method utilized the combination of ANN and Linear Discriminant Analysis (LDA) techniques for feature extraction. Five ECG features namely RR, RTc, T wave amplitude, T wave skewness and T wave kurtosis were used in this method. These features are obtained with the assistance of automatic algorithms. The onset and end of the T wave were detected using the tangent method. The three feature combinations used had very analogous performance when considering the average performance metrics.

Correlation analysis for abnormal ECG signal feature extraction was explained by Ramli and Ahmad in [93]. This paper investigated the technique to extract the important features from the 12 lead system ECG signals and chose II for the entire analysis due to its representative characteristics for identifying the common heart diseases. The analysis technique chosen is the cross-correlation analysis. Cross-correlation analysis measures the similarity between the two signals and extracts the information present in the signals. Test results suggested that this technique could effectively extract features, which differentiate between the types of heart diseases analyzed and also for normal heart signal.

Ubeysi et al. in [94] described an approach for feature extraction from ECG signal. This approach developed an automated diagnostic systems employing dissimilar and amalgamated features for ECG signals were analyzed and the accuracies were determined. The classification accuracies of Mixture of Experts (ME) trained on composite features and Modified Mixture of Experts (MME) trained on diverse features were also compared. The inputs of these automated diagnostic systems were composed of diverse or composite features and these were chosen based on the network structures. The achieved accuracy rates of this approach were higher than that of the ME trained on composite features.
Fatemian et al. [95] developed an approach for ECG feature extraction. This paper described a new wavelet based framework for automatic analysis of single lead ECG for application in human recognition. This system utilized a robust preprocessing stage, which enables it to handle noise and outliers. This facilitates it to be directly applied on the raw ECG signal. In addition this system is capable of managing ECGs regardless of the Heart Rate (HR) which renders making presumptions on the individual's stress level unnecessary. The substantial reduction of the template gallery size decreases the storage requirements of the system appreciably. Additionally, the categorization process is speeded up by eliminating the need for dimensionality reduction techniques such as PCA or LDA. Experimental results revealed the fact that this technique outperformed other conventional methods of ECG feature extraction.

Many of the cardiac problems are visible as distortions in the ECG. Since the abnormal heart beats can occur randomly it becomes very tedious and time-consuming to analyze say a 24 hour ECG signal, as it may contain hundreds of thousands of heart beats. Hence it is desired to automate the entire process of heart beat classification and preferably diagnose it accurately. Ghongade and Ghatol [40] have focused on the various schemes for extracting the useful features of the ECG signals for use with artificial neural networks. Arrhythmia is one such type of abnormality detectable by an ECG signal. The three classes of ECG signals are Normal, Fusion and Premature Ventricular Contraction (PVC). The task of an ANN based system is to correctly identify the three classes, most importantly the PVC type, this being a fatal cardiac condition. Discrete Fourier Transform, Principal Component Analysis and Discrete Wavelet Transform and Discrete Cosine Transform are the four schemes discussed and compared. For comparison the statistical techniques like linear discriminant analysis and tree clustering are also evaluated.
2.3. CLASSIFICATION TECHNIQUES

The classification phase makes an arrhythmia diagnosis using obtained feature vectors from the previous phase. Statistical approaches, fuzzy inference approaches and neural network approaches are the typically utilized techniques for ECG classification.

The classifying method which have been developed during the last decade and under evaluation includes digital signal analysis, Fuzzy Logic methods, Artificial Neural Network, Hidden Markov Model, Genetic Algorithm, Support Vector Machines, Self-Organizing Map, Bayesian and other method with each approach exhibiting its own advantages and disadvantages. This section provided an over view on various techniques and transformations used for extracting the feature from ECG signal.

A Mathematical morphology for ECG feature extraction was developed by Tadejko and Rakowski in [80]. The primary focus of this approach is to evaluate the classification performance of an automatic classifier of the ECG for the detection abnormal beats with new concept of feature extraction stage. The obtained feature sets were based on ECG morphology and RR-intervals. Configuration adopted a well known Kohonen Self-Organizing Maps (SOM) for examination of signal features and clustering. A classifier was developed with SOM and Learning Vector Quantization (LVQ) algorithms using the data from the records recommended by ANSI/AAMI EC57 standard. In addition this approach compares two strategies for classification of annotated QRS complexes: based on original ECG morphology features and developed new approach - based on preprocessed ECG morphology features. The mathematical morphology filtering is used for the preprocessing of ECG signal.

Tayel and Bouridy together in [84] put forth a technique for ECG image classification by extracting the feature using wavelet transformation and neural
networks. Features are extracted from wavelet decomposition of the ECG images intensity. The obtained ECG features are then further processed using artificial neural networks. The features are: mean median, maximum, minimum, range, standard deviation, variance and mean absolute deviation. The introduced ANN was trained by the main features of the 63 ECG images of different diseases. The test results showed that the classification accuracy of the introduced classifier and the extracted features of the ECG signal using wavelet decomposition were effectively utilized by ANN in producing the classification accuracy.

An algorithm was presented by Chouhan and Mehta in [86] for detection of QRS complexities. The recognition of QRS complexes forms the origin for more or less all automated ECG analysis algorithms. The presented algorithm utilizes a modified definition of slope, of ECG signal, as the feature for detection of QRS. A succession of transformations of the filtered and baseline drift corrected ECG signal is used for mining of a new modified slope-feature. In the presented algorithm, filtering procedure based on moving averages [87] provides smooth spike-free ECG signal, which is appropriate for slope feature extraction. The foremost step is to extract slope feature from the filtered and drift corrected ECG signal, by processing and transforming it, in such a way that the extracted feature signal is significantly enhanced in QRS region and suppressed in non-QRS region. This method has been analyzed in terms of detection rate and positive predictivity. Xu et al. in [88] described an algorithm using Slope Vector Waveform (SVW) for ECG QRS complex detection and RR interval evaluation. In this method variable stage differentiation is used to achieve the desired slope vectors for feature extraction and the non-linear amplification is used to get better of the signal-to-noise ratio. The method allows for a fast and accurate search of the R location, QRS complex duration and RR interval and yields excellent ECG feature extraction results. In order to get QRS durations, the feature extraction rules are needed.
A modified combined wavelet transforms technique was developed by Saxena et al. in [90]. The technique has been developed to analyze multi lead electrocardiogram signals for cardiac disease diagnostics. Two wavelets have been used, i.e. a Quadratic Spline Wavelet (QSWT) for QRS detection and the Daubechies Six Coefficient (DB6) wavelet for P and T detection. A procedure has been evolved using electrocardiogram parameters with a point scoring system for diagnosis of various cardiac diseases. The consistency and reliability of the identified and measured parameters were confirmed when both the diagnostic criteria gave the same results.

Jen et al. in [92] formulated an approach using neural networks for determining the features of ECG signal. This paper presented an integrated system for ECG diagnosis. The integrated system comprised of cepstrum coefficient method for feature extraction from long-term ECG signals and ANN models for the classification. Utilizing this method, one can identify the characteristics hiding inside an ECG signal and then classify the signal as well as diagnose the abnormalities. To explore the performance of this method various types of ECG data from the MIT/BIH database were used for verification. The experimental results showed the accuracy of diagnosing cardiac disease. In addition, this method successfully extracted the corresponding feature vectors, distinguished the difference and classified ECG signals.

Nazmy et al., [31] presented an intelligent diagnosis system using hybrid approach of Adaptive Neuro-Fuzzy Inference System (ANFIS) model for classification of ECG signals. Feature extraction using Independent Component Analysis (ICA) and Power spectrum, together with the RR interval then serve as input feature vector, this feature was used as input of ANFIS classifiers. Six types of ECG signals are Normal Sinus Rhythm (NSR), PVC, Atrial Premature Contraction (APC), VT, Ventricular Fibrillation (VF) and Supraventricular Tachycardia (SVT). ANFIS model combined the neural network adaptive
capabilities and the fuzzy inference system. The results indicate a high level of efficient of tools used with more accuracy.

Nait-Hamoud and Moussaoui [33] developed two novel methods of ECG classification to discriminate five heart beat types. The first approach combines PCA and Modified Fuzzy One-Against-One (MFOAO) method for multiclass categorization. The Fuzzy One-Against-One (FOAO) method converts the n-class problem of classification to n(n-1)/2 two-class problems and performs the binary classification with SVM. It was introduced to solve the problem of the unclassified regions induced by the classical pairwise classification one-against-one. This modified algorithm of FOAO uses Fuzzy Support Vector Machine (FSVM) for the binary classification in order to discard outliers. The second approach integrates PCA, Unbalanced Clustering (UC) and FOAO algorithms. PCA is used to extract the principal characteristics of the signal and reduce its dimension. UC algorithm is used to discard outliers and reduce the training set by replacing samples with prototypes. The first goal of this work is to compare the ability of the two novel methods to discard outliers and enhance the performance of the classification with PCA and FOAO; the second one is to highlight the efficiency of the combined method PCA-UC-FOAO in the classification of long term ECG records.

Wei Jiang and Seong Kong [34] presented evolvable Block-based Neural Networks (BbNNs) for personalized ECG heartbeat pattern classification. A BbNN consists of a 2-D array of modular component NNs with flexible structures and internal configurations that can be implemented using reconfigurable digital hardware such as Field-Programmable Gate Arrays (FPGAs). Signal flow between the blocks determines the internal configuration of a block as well as the overall structure of the BbNN. Network structure and the weights are optimized using local gradient-based search and evolutionary operators with the rates changing adaptively according to the effectiveness in the previous evolution period. Such adaptive operator rate update scheme ensures higher fitness on average compared
to predetermined fixed operator rates. The Hermite transform coefficients and the
time interval between two neighboring R-peaks of ECG signals are used as inputs
to the BbNN. A BbNN optimized with the Evolutionary Algorithm (EA) makes a
personalized heartbeat pattern classifier that copes with changing operating
environments caused by individual difference and time-varying characteristics of
ECG signals. Simulation results using the MIT-BIH arrhythmia database
demonstrate high average detection accuracies of ventricular ectopic beats and
supraventricular ectopic beats patterns for heartbeat monitoring, being a
significant improvement over previously reported ECG classification results.

A novel multi-lead ECG classification method is developed by Shen et al.,
[35]. At the feature extracting stage, an improved ICA method is introduced. In
this method, a heartbeat is intercepted into 3 segments (P wave, QRS interval, ST
segment). ICA is used to extract the features of each segment separately. These
three feature vectors construct the feature of single lead firstly. Then, twelve single
lead feature vectors are combined to generate a multi-lead feature vector one by
one. At last, the SVM is used for multi-classification and 2-classification
experiments. For MIT-BIH data, multi-classification result is discussed. The final
average accuracy of the testing data and the average sensitivity is discussed. For
practical data, 2-classification experiment result is discussed.

Vornicu and Goras [36] discussed the possibility of using the dynamics of a
class of Cellular Neural Networks (CNN's) for ECG signals classification. The
main idea is that of segmentation and transformation of the temporal signal into a
1D spatial one which is further processed by means of a bank of linear spatial
filters using a parallel architecture of CNN type. A major advantage of this
solution is the independence of the filters spatial frequency characteristics on the
number of samples of the ECG pattern, which allows dealing very easily with the
heart rate variability. The principle of this architecture is briefly discussed and the
design of a bank of spatial filters for ECG classification is presented. Transistor
level simulation and considerations regarding the architecture reconfiguration are given as well.

Privacy protection is a crucial problem in many biomedical signal processing applications. For this reason, particular attention has been given to the use of secure multiparty computation techniques for processing biomedical signals, whereby nontrusted parties are able to manipulate the signals although they are encrypted. Barni et al., [37] focused on the development of a privacy preserving automatic diagnosis system whereby a remote server classifies a biomedical signal provided by the client without getting any information about the signal itself and the final result of the classification. This paper presented and compared two methods for the secure classification of ECG signals: the former based on linear branching programs (a particular kind of decision tree) and the latter relying on neural networks. This method dealt with all the requirements and difficulties related to working with data that must stay encrypted during all the computation steps, including the necessity of working with fixed point arithmetic with no truncation while guaranteeing the same performance of a floating point implementation in the plain domain. A highly efficient version of the underlying cryptographic primitives is used, ensuring a good efficiency of the two methods, from both a communication and computational complexity perspectives. This system proved that carrying out complex tasks like ECG classification in the encrypted domain efficiently is indeed possible in the semihonest model, paving the way to interesting future applications wherein privacy of signal owners is protected by applying high security standards.

Barni et al., [38] described a privacy-preserving system where a server can classify an ECG signal without learning any information about the ECG signal and the client is prevented from gaining knowledge about the classification algorithm used by the server. The system relies on the concept of Linear Branching Programs (LBP) and a recently developed cryptographic protocol for secure
evaluation of private LBPs. This approach investigated the trade-off between signal representation accuracy and system complexity both from practical and theoretical perspective. As a result, the inputs to the system are represented with the minimum number of bits ensuring the same classification accuracy of a plain implementation. This paper revealed how the overall system complexity can be strongly reduced by modifying the original ECG classification algorithm. Two alternatives of the underlying cryptographic protocol are implemented and the corresponding complexities are analyzed to show suitability of this system in real-life applications for current and future security levels.

Bozzola et al., [39] presented an approach to the automated ECG classification based on a hybrid neuro-fuzzy model. The classification power of the connectionist paradigm has been coupled with the ability of the fuzzy set formalism to treat in a quantitative way natural language. This allows to build up a system capable of both good classification accuracy and to give meaningful explanations of the diagnoses, in the form of symbolic IF-THEN rules.

In cardiology, determining whether an ECG is normal or not is sometimes referred to as ECG classification. ECG is the most frequently-used means of cardiac diagnosis. It is the cheapest and the most widely-available; it is also crucial for detecting rhythmic problems. Bensaid et al., [41] derived fuzzy rules for ECG classification from ID3-induced decision trees. The system of fuzzy rules is designed based on 106 ECGs and it is evaluated using a validation set of 48 ECGs carefully selected by cardiologists. Using the same 106 ECGs for design and the same 48 ECGs for validation, an ID3-generated decision tree yields and a neural network trained with the feed forward cascade-correlation algorithm are used. On the other hand, the derived fuzzy rules, combined with an optimized defuzzification using the cascade correlation neural network.
This study by Mar et al., [42] tackles the ECG classification problem by means of a methodology, which is able to enhance classification performance while simultaneously reducing the computational resources, making it specially adequate for its application in the improvement of ambulatory settings. For this purpose, the Sequential Forward Floating Search (SFFS) algorithm is applied with a new criterion function index based on linear discriminants. This criterion has been devised specifically to be a quality indicator in ECG arrhythmia classification. Based on this measure, a comprehensive feature set is analyzed with the SFFS algorithm and the most suitable subset returned is additionally evaluated with a Multilayer Perceptron (MLP) to assess the robustness of the model. Aiming at obtaining meaningful estimates of the real-world performance and facilitating comparison with similar studies, the present contribution follows the Association for the Advancement of Medical Instrumentation standard EC57:1998 and the same interpatient division scheme used in several previous studies.

Silipo et al., [43] presented a supervised and unsupervised learning for diagnostic ECG classification. A hybrid system with RBF pre-processing, a system with supervised learning, is compared with some Kohonen SOM in a subtle ECG classification task. Based on ECG measures, they are supposed to detect normal condition, presence of infarction and of hypertrophy and at the same time to sub-classify those pathologies. During the evaluation process the hybrid system produces better results in terms of average sensitivity and specificity, but Kohonen maps allow a detailed description of the similarities among input data. An integration of the two techniques should improve the final results.

Bortolan et al., [44] described a technique for the automatic acquisition of expert knowledge in order to set up a knowledge base for the diagnostic classification of ECG signals. The method is indirect, because the knowledge of the expert, in contrast with the general approach which learns through the direct communication of rules and facts, is derived from a learning set of classified
ECGs. It is, on the other hand, different from conventional statistical techniques, because (1) the reference classification is given by experts and not by independent exams like autopsy, coronaryography, echocardiography, cardiac surgery and so on and (2) this classification can be uncertain, i.e. the various classes are associated with each ECG with certainty factors which can differ from 0 or 1. The data are derived from the CSE pilot diagnostic library. In this preliminary study, the results of the method, which is based on fuzzy pattern matching, show a global type-4 error (complete disagreement).

Ceylan et al., [45] in this study, a new structure formed by Complex Wavelet Transform (CWT) with different levels and Complex-Valued Artificial Neural Network (CVANN) is developed for classification of ECG arrhythmias. In this structure, features of ECG data are extracted using CWT and data size is reduced. After then, four statistical features (maximum value, minimum value, mean value and standard deviation) are obtained from extracted features. These new statistical features are presented to CVANN as inputs. Data set used in this study, including five different arrhythmias (normal sinus rhythm, right bundle branch block, left bundle branch block, atrial fibrillation and atrial flutter), are selected from MITBIH ECG database. Number of samples in training and test sets for each pattern is reduced from 200 real-valued samples to 100, 50 and 25 complex-valued samples using first level CWT, second level CWT and third level CWT, respectively.

Automatic analysis of cardiac arrhythmias is very important for diagnosis of cardiac abnormalities. Leigang Zhang et al., [46] presented a novel approach that classifies ECG signals with the combination of Wavelet transform and Decision tree classification. This approach has two aspects. In the first aspect, this approach utilized the wavelet transform to extract the ECG signals wavelet coefficients as the first features and utilize the combination of PCA and ICA to remove the first features relativity and search this features independence as the
new features, then added the RR interval as the final features. In the second aspect, this approach utilized the ID3 algorithm which is one of analysis decision tree methods as the classifier to recognize the different heartbeat arrhythmias. This approach utilized the MIT-BIH Arrhythmia Database to create the classification and test the classification. The results confirm its high reliability and high accuracy.

The objective of this work by Llamedo Soria et al., [47] is to develop a model for ECG classification based on multilead features. The MIT-BIH Arrhythmia database was used following AAMI recommendations and class labeling. This method used for classification of classical features as well as features extracted from different scales of the wavelet decomposition of both leads integrated in an RMS manner. Step-wise and a randomized method were considered for feature subset selection and LDA was also used for additional dimensional reduction. Three classifiers: linear, quadratic and Mahalanobis distance were evaluated, using a k-fold like cross validation scheme. Results in the training set showed that the best performance was obtained with a 28-feature subset, using LDA and a Mahalanobis distance classifier. This model was evaluated in the test dataset with the performance measurements for supraventricular beats Ventricular beats Sensitivity. This results show the feasibility of classification based on the multilead wavelet features, although further development is needed in subset selection and classification algorithms.

In this study, Sheikh and Taj [48] presented a new approach to analyze and classify ECG signals for diagnosis of five cardiac conditions. Instead of following the conventional approach of beat-to-beat classification, this method classified cardiac rhythms/segments of ECG based on statistical and morphological features extracted from them. This method selected and extracted seven suitable features and reduced the feature space by using PCA and LDA to optimize this problem. In classification phase, two types of classifiers are used, i.e. Euclidean and
Manhalanobis and compared the results. A new algorithm for segmentation of ECG lengths is also presented. This study provides a fundamental step for the development of preliminary automated diagnostic system for cardiac disorders.

A procedure for beat detection and classification was developed using ECG recordings. This procedure can be used for beat detection with one or two leads and then a portion of each detected beat is used for classifying. This task is performed by a neural network. In Risk et al., [49] work the morphology of the QRS portion of the ECG feeds a SOM. The SOM was previously trained with different QRS complexes such as normal and ectopic beat morphologies. The beat classification is very important in Heart Rate Variability (HRV) analysis, because one must use only the normal beats and reject the ectopic ones for the construction of the RR intervals beat series.

The 12-lead ECG, as well as the patient history, plays an important role in the early diagnosis of Acute Myocardial Infarction (AMI). A hybrid neuro-fuzzy approach to the diagnostic classification of 12-lead ECGs is presented by Lu [50]. The architecture used is a combination of fuzzy logic and neural network theory. For ECG diagnosis, the system benefits from the reasoning capabilities of fuzzy logic as well as the learning ability of neural networks. This hybrid system consists of two phases: (1) Use fuzzy logic to establish the diagnosis system in the form of symbolic IF-THEN rules based on expert cardiac knowledge; (2) Through a training process, use a back propagation network to automatically adjust the parameters of the system. A total of 124 ECGs from patients with or without acute myocardial infarction have been studied and eight diagnostic classes have been taken into account regarding the different locations of AMI. Sensitivity, specificity, partial and total accuracy are used for evaluation of the system. The results confirmed that AMI can be diagnosed with reasonable accuracy. While the method recognized that the diagnosis of AMI varies according to clinical
circumstances, the hybrid system has the potential for automatic classification of AMI.

This study by de Chazel et al., [52] investigates the automatic classification of the Frank lead ECG into different disease categories. A comparison of the performance of a number of different feature sets is presented. The feature sets considered include wavelet-based features, standard cardiology features and features taken directly from time-domain samples of the EGG. The classification performance of each feature set was optimised using automatic feature selection and choosing the best classifier model from linear, quadratic and logistic discriminants. The ECG database used contains 500 cases classed into seven categories. Using multiple runs of ten-fold cross-validation, the overall seven-way accuracy of different feature sets and classifier model combinations are determined. The best performing classifier used linear discriminants processing selected time-domain features. This is also found to be the simplest and fastest classifier to implement.

Habboush et al., [53] compared neural networks designed for ECG compression and classification with optimum linear methods. It is found that simple neural networks with one hidden layer approach the performance of linear methods, but offer no advantage over them. Suitably constructed networks with more than one hidden layer, however, can perform more efficient ECG compression than is possible using linear methods under the same constraints.

Neural and traditional techniques have been compared for the particular task of automatic ECG analysis. A large validated ECG database has been used. Statistical methods, neural architectures with supervised and unsupervised learning, and a neuro fuzzy architecture have been considered by Silipo and Bortolan [54]. The results from the connectionist approach are always at least comparable with those coming from more traditional classification methods. But
the best performances have been obtained by the combination of the connectionist with the fuzzy approach.

Llamedo et al., [55] studied the improvement achieved by including information from the 12 ECG leads, in a previously developed classification model. This model includes features from the RR interval series and morphology descriptors calculated from the wavelet transform. The experiments were carried out in the INCART database, available in Physionet and the generalization was corroborated in a private database. In both databases the AAMI recommendations for class labeling and results presentation were followed. Different approaches to integrate the additional information available in the 12-leads were studied. The best performing approach obtained for normal beats, Supraventricular beats, Ventricular beats. The generalization capability was confirmed in a private database with comparable results. The performance of the reference two-lead classifier was improved by taking into account additional information from the 12-leads.

The performance of the neural network approach in the diagnostic classification of 12-lead ECG is investigated by Bortolan et al., [56]. For this study, a validated ECG database established at the University of Leuven is used. Previous results obtained from the same database to derive two classifiers based on statistical models (linear discriminant analysis and logistic discriminant analysis) are taken as reference points in the evaluation. Simple neural network architecture is chosen: the feed-forward structure with the use of the back-propagation algorithm. Sensitivity, specificity, total and partial accuracy are the indices used for the assessment of the performance. The results show a comparable behavior with the two statistical methods.

The single-layer Feedforward Neural Network (FFNN) in conjunction with the Backpropagation Training Algorithm (BPTA) is developed by Sakk et al., [57]
for ECG classification. It has been observed that, for such a problem, the values of the input weights are closely related to the input training set. An implication of this observation is that, rather than choosing initially random weights for the BPTA, one may choose initial weights that are actually quite close to an optimal solution. An advantage of such a choice is faster convergence time based on knowledge of the incoming training data. Decreasing convergence time makes more promising the use of the FFNN to classify ECGs for arrhythmia detection, ambulatory monitoring and analysis and front-line physician support instrumentation.

Soria et al., [58] studied the classification performance of models based on intervals, angles and amplitudes. These features were extracted from both ECG leads and different scales of the wavelet decomposition. The MIT-BIH Arrhythmia database was used, following AAMI recommendations for class labeling and results presentation. The training and testing set and any cross-validation division of the database was made patient-oriented. A floating feature selection algorithm was used to obtain best performing models in the training set. This model was evaluated in the test set obtaining a global accuracy for normal beats, Supraventricular beats, Ventricular beats. This classifier model based on multidomain features performs better than other state of the art methods, with a fraction of the features.

Messadeg et al., [59] developed a method for the classification of the ECG signal. This method is based on the use of the ergodic Markov model. The data used was obtained from the MIT-BIH arrhythmia database for two categories. Each beat is isolated and its discrete wavelet transform is calculated and the vectorial quantization is applied. The parameters of the Markov model are computed for different number of states. The results presented are discussed.
Using an array of three Probabilistic Neural Networks (PNNs), this approach successfully identified NSR and Atrial Fibrillation (AF), as well as normal and PVC waveforms. Training and test waveforms were obtained from the MIT-BIH Arrhythmia Database. Kramer et al., [60] applied various preprocessing techniques to reduce the dimension of the training sets. Combining independent PNNs, each classifying based on either shape or rhythm, this method enhanced the integrated system performance by diminishing PNN element misclassifications. Most notably, the percentage of correctly classified PVCs from testing record 116, the worst performance based on shape, increased the classification when adding rhythm information. Similarly, the amount of correctly classified NSR from testing record 201, the worst performance based on rhythm, rose when shape information was added.

Medical engineering support systems that are controlled by neural networks are being applied with increasing frequency in medical practice. However, a solution still needs to be found for the problem of constructing medical support systems that can be set up by the physicians themselves without the need of knowledge of the mathematical theories of neural networks and signal processing. Reuter et al., [62] describe a medical support system which can be set up by novices in the field of neural networks and which can be controlled and the results correctly interpreted by nurses and other medical staff. On the basis of typical pathological types of ECG signals, simulated by a common simulator which is used in Europe for heart beat monitoring, this paper shown the basic structure of normal and pathological heart beat signatures and how they can be presented in a new and readily interpretable display. Furthermore, this paper explained how this support system can be used to create networks which are typical for most patients, are small and can be quickly set up to monitor the patient's condition during therapy, or to provide a measure of the depth of anaesthesia of the patient.
Teck Wee Chua and Woei Wan Tan [63] studied the ability of a Non-Singleton Fuzzy Logic System (NSFLS) that is evolved using Genetic Algorithm (GA) to handle the uncertainties in pattern classification problems. The performance of non-singleton and singleton systems for cardiac arrhythmias classification is compared. Results show that NSFLS can deal with uncertainty within its framework more efficiently, thereby enabling classification to be performed using features that are easier to extract.

Raghav and Mishra [64] developed a method for the classification of ECG arrhythmia using local fractal dimensions of ECG signal as the features to classify the arrhythmic beats. The heart beat waveforms were extracted within a fixed length window around the R-peak of the signal and local fractal dimension is calculated at each sample point of the ECG waveform. The method is based on matching these fractal dimension series of the test ECG waveform to that of the representative ECG waveforms of different types of arrhythmia, by calculating Euclidean distances or by calculating the correlation coefficients. The performance of the classifier was tested on independent MIT-BIH arrhythmia database. The achieved performance is represented in terms of the percentage of correct classification. The performance was found to be competitive to other published results. The current classification algorithm proved to be a computationally efficient and hence a potential technique for automatic recognition of arrhythmic beats in ECG monitors or Holter ECG recorders.

Accurate ECG beat classification is essential for automated detection of arrhythmias. A novel classification algorithm of the ECG beats, applying Mirrored Gauss Model (MGM) had been developed by Quanyi Zhou et al., [65]. The MGM has strong morphological representation ability for QRS complex waves using curve fitting. With the MGM, the width of QRS complex wave could be extracted and applied to ECG beat classification easily, effectively and automatically. It was proved by experiment carrying out using all of ECG records in MIT-BIH
Arrhythmia Database that the MGM is a promising algorithm for ECG beat classification. The whole classification accuracy for normal beats and PVC beats are analyzed.

Table 2.1
Overview of Existing Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Researcher</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DWT</td>
<td>Castro et al., [79]</td>
<td>Used DWT for feature extraction from an ECG signal and recognition of abnormal heartbeats.</td>
</tr>
<tr>
<td></td>
<td>Mahmoodabadi et al., [73]</td>
<td>Utilized Daubechies Wavelets transform for ECG feature extraction.</td>
</tr>
<tr>
<td></td>
<td>Emran et al., [83]</td>
<td>Used DWT to extract the relevant information from the ECG input data in order to perform the classification task.</td>
</tr>
<tr>
<td>SVM</td>
<td>Thanapatay et al., [32]</td>
<td>PCA and SVM are adapted to create model of classifier.</td>
</tr>
<tr>
<td></td>
<td>Zhao et al. [78]</td>
<td>SVM with Gaussian kernel is used to classify different ECG heart rhythm.</td>
</tr>
<tr>
<td></td>
<td>Shen et al., [35]</td>
<td>ICA is used to extract the features and SVM is used for multi-classification.</td>
</tr>
<tr>
<td>NN</td>
<td>Saxena et al., [82]</td>
<td>Competent composite method is used for data compression, signal retrieval and feature extraction of ECG signals and ANN for classification.</td>
</tr>
<tr>
<td></td>
<td>Alexakis et al., [89]</td>
<td>Combination of ANN and Linear Discriminant Analysis (LDA) techniques for feature extraction</td>
</tr>
<tr>
<td></td>
<td>Kramer et al., [60]</td>
<td>PNN is used for ECG classification.</td>
</tr>
<tr>
<td>GA</td>
<td>Teck Wee Chua and Woei Wan Tan [63]</td>
<td>GA to handle the uncertainties in pattern classification problems</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>SOM</td>
<td>Tadejko and Rakowski [80]</td>
<td>Classifier was developed with SOM and Learning Vector Quantization (LVQ).</td>
</tr>
<tr>
<td>Risk et al., [49]</td>
<td>SOM was previously trained with different QRS complexes such as normal and ectopic beat morphologies for heart beat detection and classification.</td>
<td></td>
</tr>
<tr>
<td>Fuzzy</td>
<td>Nazmy et al., [31]</td>
<td>ANFIS model for classification of ECG signals.</td>
</tr>
<tr>
<td></td>
<td>Nait-Hamoud and Moussaoui [33]</td>
<td>FSVM for the binary classification.</td>
</tr>
<tr>
<td>Other Methods</td>
<td>Jovic and Bogunovic [85]</td>
<td>Chaos theory is applied to ECG feature extraction.</td>
</tr>
<tr>
<td></td>
<td>Ramli and Ahmad [93]</td>
<td>Cross-correlation analysis measures the similarity between the two signals and extracts the information.</td>
</tr>
<tr>
<td></td>
<td>Chouhan and Mehta [86]</td>
<td>Slope Vector Waveform (SVW) for ECG QRS complex detection.</td>
</tr>
<tr>
<td></td>
<td>Barni et al., [38]</td>
<td>Linear Branching Programs (LBP) to classify an ECG signal without learning.</td>
</tr>
</tbody>
</table>

**2.4. SUMMARY**

Existing studies on the classification of arrhythmia is generally composed of pre-processing part, feature extraction part and classification part. From the
approaches mentioned above, each method has its own advantages and disadvantages. It is clear that research in the field of ECG classification has reached a good level of maturation but it needs further evaluation in feature selection and classification for accurate and reliable arrhythmia identification. In the design of an ECG classification part, there are still some open issues that are to be addressed for the development of more robust and efficient classifiers. One of these issues is related to the choice of the classification approach to be adopted.

The major drawbacks that need to be addressed in the existing systems are,

1) The methods of removing noise used are not so efficient.
2) Accuracy of the classification results is not adequate for reliable arrhythmia identification. Limited number of MIT BIH records and arrhythmia types are considered.
3) Feature selection is not performed in a completely automatic way.
4) The selection of the best free parameters of the adopted classifier is generally done empirically.