Detecting beats in the ECG: A comparison of time domain and morphological features using Support Vector Machines and MultiLayer Perceptron

Aunsa Shah, Seral Özşen, Abbas Shah

Abstract— **The ElectroCardioGraph (ECG) is the most widely used diagnostic test for determining heart related disease prognosis. This paper presents a comparison of two types of feature extraction methods and two types of classifiers for the detection of four types of heart beats in the ECG. The four types of heart beats considered in this work are Normal, Right Bundle Branch Block Beat, Left Bundle Branch Block Beat and the Premature Ventricular Contraction beat. The first set of features computed for each beat type are statistical in nature in the time domain and the second set of features are morphological in nature. The values of the features in these two sets are then sent to two different classification algorithms, the Support Vector Machine (SVM) and the MultiLayer Perceptron (MLP) Neural Network. The classification results demonstrate that when comparing the chosen set of statistical and morphological features, the statistical values of each beat provide a higher detection accuracy for all beat types. Furthermore, it was also observed that when comparing the performance of the SVM and MLP algorithms for heart beat classification, the MLP was found to outperform SVM when using statistical features and when both feature sets were combined, however, the opposite was observed when only morphological features were used in which case, the SVM outperformed the MLP network.**

Index Terms— **ECG, Heart beat detection, MultiLayer Perceptron, Support Vector Machine**

I. INTRODUCTION

According to World Health Organization, the number of deaths in the world due to Cardiac problems in 2015 was 17.7 million [1]. Moreover, the Centre for Disease Control (CDC) in the United States suggests that 1 in 4 deaths in 2016 were heart related [2] and it was the leading cause of death in the world. A major cause of heart related deaths is Cardiac Arrhythmia, which is life threatening [3], [4]. However, if detected and treated timely, life could be saved. The major mechanism of heart monitoring is the use of ElectroCardioGrams (ECG), any abnormality in the ECG pattern may indicate to an aberration in heart activity. Different types of heart beats occur in a specific sequence to make up the heart rhythm, thus a change in the beats (for e.g. order) would result in a change in the heart rhythm and can be used to determine the heart health condition. It would therefore be very useful if algorithms could be developed that would facilitate processing of biological signals such as the ECG to enable automated diagnosis.

The paper is organized as follows, Section II provides a brief discussion of the previously used approaches, Section III discusses the methodology of the proposed work, Section IV provides the results of the current work and Section V concludes the discussion.

Aunsa Shah, Institute of Information and Communication Technologies, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan

Dr. Seral Özşen (PhD), Selcuk University, Electrical and Electronics Engineering Department, Konya, Turkey. She was born in 1980 in Aydın, Turkey. She graduated from Electrical and Electronics Department of Ege University in 2002. Then she fullfilled master and PhD programs in 2004 and 2008 respectively in Selcuk University Eelctrical & Electronics Engineering Department. She is now working in the same department as Associate Prof.

Abbas Shah, Department of Electronic Engineering, Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan

II. LITERATURE REVIEW

The task of determining heart beat types in an ECG is pivotal for heart disease determination Unsurprisingly, the processing of ECG signals has therefore been of much interest to researchers in the data science community.

The authors in [5] perform QRS complex detection and compare the performance of Fuzzy Rough Nearest Neighbor and Multi-Layer Perceptron for use as classifiers. In [6], the authors use Auto-Regressive (AR) [7] modeling of ECG beats along with neural networks to detect ventricular arrhythmia in ECGs. This work considers two types of arrhythmias (Ventricular TachyCardia and Ventricular Flutter). Another technique utilizing AR Modeling is presented in [8] where they use Principal Component Analysis (PCA) [9] in addition to AR modeling for feature extraction and then use various classifiers to differentiate between different types of heartbeats.

The authors in [10] use a multi-resolution wavelet transform along with amplitude thresholding and two classifiers, a Multi-layer Perceptron Neural Network (MLP NN) [11] and the Support Vector Machines (SVM) [12]. They consider the extracted features for the conditions of normal (N), left bundle branch block (LBBB), right bundle branch block (RBBB) and Paced beats (P) for detection through MLP NN and SVM. They found that the SVM outperforms the MLP NN. Another method based on wavelets is presented in [13], where the authors use Discrete Wavelet Transforms to extract different frequency bands of the heartbeat in an ECG. Computing statistical features of those frequency components they use a Random Forest (RF) classifier [14] to classify different types of heartbeats. They conclude that the RF classifier performs better than the C4.5 [15] and the Classification and Regression Tree (CART) [16] classifiers for this task.

A comparison of five different feature detection algorithms was performed by the authors in [17]. The authors use the Principal Component Analysis, Auto-Regressive modelling and three types of wavelets, Debauchies [18], Haar [19] and

Bioorthogonal [20] along with SVM to detect three types of ECG beats with a short and a long window. The authors observed from their results that AR modelling and the three wavelets were good approaches for use with short windows as compared to longer windows. On the other hand, PCA provided suitable results for both window sizes. In [21], the authors use spectral correlation and PCA to extract features from ECG heartbeats before passing it on to a SVM for classification. In [22], the authors use Discrete Cosine Transform (DCT) [23] and PCA along with K-Nearest Neighbours (K-NN) [24] to perform classification of heartbeats in the ECG.

The authors in [25] use the Independent Component Analysis [26] to extract features from the ECG heartbeat. Combined with the power of the QRS complex and the R-R interval power, the authors utilize a Neural Network to classify multiple types of heartbeats. The authors in [27] provide a comparison of different combinations of PCA, ICA and DWT feature extraction methods and three types of classifiers, namely Optimal Path Forest (proposed by them), SVM-RBF and a Bayesian Classifier (BC) for the detection of heartbeats. They find that the SVM-RBF combination provides better performance compared to the other classifiers.

As is clear from the above discussion, various techniques have been used and proposed for the feature extraction from heartbeats in the ECG and their subsequent classification. It would be useful if a consolidated comparison be provided for the most commonly used feature extraction and classification algorithms used for heartbeats. In this work, we present such a comparison by using statistical time domain and morphological features and providing a comparison of their performance using two different classifiers, the Support Vector Machines and the MultiLayer Perceptron Neural Network.

III. MIT-BIH ARRHYTHMIA DATABASE

The ECG waves used in this work are derived from the Massachusetts Institute of Technology - Beth Israel Hospital (MIT-BIH) Arrhythmia Database [28], [29]. The database contains 48 recordings of two channel ECG waves for different arrhythmia types. Every ECG recording present in this database is 30 minutes long and recorded with a sampling frequency of 360 Hz. The ECG waves are annotated for arrhythmias and the beats that constitute them. The MIT-BIH database has been widely used for testing of algorithms for ECG signal processing and classification and therefore, this database has been used in this work.

IV. METHODOLOGY

The methodology followed in this work is the conventional three step process true for typical detection/classification scenarios as shown in Fig. 1. In the first stage, the ECG signals from the database are suitably conditioned so that they can be used for further processing. In the second stage, two different types of feature sets are computed, the first feature set is a group of statistical time domain features and the second set is a group of morphological features. Lastly, each set is sent to a SVM classifier and a MultiLayer Perceptron Neural Network classifier for the detection of heart beats.

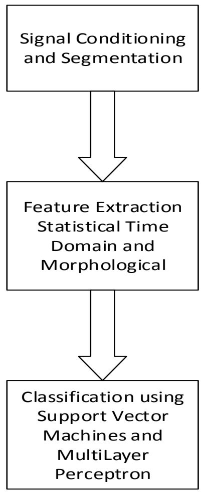

Figure 1: Detection of Heart beats in the ECG wave

A. *Signal Conditioning and Segmentation*

The process of recording ECGs is susceptible to noise and thus signal aberrations of unwanted nature are acquired while recording the heart rhythm. This includes power line interference and DC offsets also called baseline interference [30]. In order to remove these unwanted effects, each ECG wave used from the MIT-BIH ECG database has been zero phase filtered with a notch filter having a cutoff frequency of 60Hz to remove power line interference. Furthermore, baseline interference is removed by passing each ECG wave through a high pass filter having a cutoff frequency of 0.5 Hz. Once these two noise signals are removed from the ECG wave, individual beats are extracted from each wave.

B. *Feature Extraction*

After signal conditioning and segmentation, the next task is the extraction of features from the individual beats of the ECG. The feature selection process is of pivotal importance as the extracted features should be characteristic of the beat types considered. In this work, a total of eight features have been computed for each ECG beat. These features encompass both statistical time domain and as well as the shape of beat, i.e. of morphological nature. Table I lists the features computed for each beat.

Table I: List of features computed

Time domain	**Morphological**
Time Mean	QRS Interval
Time Kurtosis	RR Interval
Time Skewness	
Time Energy	
Time Std	
Zero crossings	

C. *Classification*

Two different algorithms have been used for the purpose of classification in this work. These are the Support Vector Machine and the Multilayer Perceptron Neural Network. These were chosen based on the literature review and consists of choices which have been deemed suitable for speed as well as computational complexity

V. RESULTS AND DISCUSSION

To compare the performance of each of the features in characterizing heart beat type, five tests have been performed in total using each of the two classification algorithms using 10-fold cross-validation to provide a comprehensive observation of the performance of each type of feature and each classification algorithm.

The total number of beats for each type are listed in Table II.

Table II: No of beats for each beat type

Beat Type	No of beats considered
Normal Beat Right	7849
Left Bundle Branch Block Beat	1734
Bundle Branch Block Beat	4364
Premature Ventricular Contraction	1464
Total beat samples	15411

A. *Time Domain Features*

In this test, the two classifiers were tested using features extracted in the time domain only. The results for each of the classifiers are presented in Tables II and IV. The multilayer perceptron had 9 sigmoid layers.

Table III: SVM Results for Time domain features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	7623	47	99	80	97.12065
1734	**LBBB**	35	1685	11	3	97.17416
4364	**RBBB**	169	7	4162	26	95.37122
1464	**PVC**	552	33	63	816	55.7377

Table IV: Multilayer Perceptron Results for Time domain features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	7615	40	125	69	97.01873
1734	**LBBB**	31	1683	16	4	97.05882
4364	**RBBB**	47	15	4278	24	98.02933
1464	**PVC**	250	28	228	958	65.43716

It can be observed from Tables III and IV that the MLP network outperforms the SVM for the detection of heart beats when using time domain features only. The classification percentages for the N and the LBBB is very similar, however the detection accuracy for the RBBB and PVC beat types is higher than that of SVM.

B. *Morphological Features*

In this test, the two classifiers were tested using the morphological features only. The results for each of the classifiers are presented in Tables V and VI. The multilayer perceptron had 7 sigmoid layers.

Table V: SVM Results for Morphological features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	7214	249	385	1	91.9098
1734	**LBBB**	958	774	2	0	44.63668
4364	**RBBB**	2258	56	2016	34	46.19615
1464	**PVC**	954	27	342	141	9.631148

Table VI: Multilayer Perceptron Results for Morphological features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	6506	51	1292	0	82.88954
1734	**LBBB**	1653	60	21	0	3.460208
4364	**RBBB**	2269	2	2093	0	47.96059
1464	**PVC**	1072	2	390	0	0

From Tables V and VI, it can be observed that for the case of morphological features, the SVM outperforms the MLP network. However, the performance for detection of heart beats for both these algorithms is below par, with the MLP completely failing to correctly classify PVC beats.

C. *Time and Morphological Features*

In this test, the two classifiers were tested using features from both the time and morphological domains. The results for each of the classifiers are presented in Tables VII and VIII. The multilayer perceptron had 10 sigmoid layers.

Table VII: SVM Results for Morphological features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	7628	51	97	73	97.18435
1734	**LBBB**	32	1690	10	2	97.46251
4364	**RBBB**	153	7	4179	25	95.76077
1464	**PVC**	511	34	60	859	58.67486

Table VIII: Multilayer Perceptron Results for Morphological features

Total beats	Beat type	N	LBBB	RBBB	PVC	Accuracy (%)
7849	**N**	7603	40	85	121	96.86584
1734	**LBBB**	19	1697	16	2	97.86621
4364	**RBBB**	36	10	4305	12	98.67064
1464	**PVC**	263	27	142	1032	70.4918

As can be observed from Table VII and Table VIII, when time domain and morphological features are combined the MLP network provides much better performance than the SVM with the biggest difference observed for the case of the PVC beat which is detected with an accuracy of 58% and 70% for the SVM and MLP respectively.

D. *Comparison of SVM and MLP performance*

Table IX lists the average detection percentages for each of the classifiers with respect to the two types of feature sets computed for this work.

Table IX: Average detection accuracies for each classifier and each feature set

Feature Set	SVM (%)	Multilayer Perceptron (%)
Time	86.35094	89.38601
Morphological	48.09344	33.57758
Both	87.27063	90.97362

As can be observed from Table IX, combining statistical time domain and morphological features provides much better performance for determining heart beat type than using each set of features individually. This is reflected in the result obtained from both the algorithms, the MLP and the SVM. Surprisingly, the MLP provides subpar performance when only morphological features are considered whereas the MLP provides relatively better performance when either time domain only features or both the combined set of features is considered.

VI. CONCLUSION

This work addresses the problem of heart beat detection in ECG monitoring. Considering four different types of beats, Normal, Left Bundle Branch Block, Right Bundle Branch Block and It provides a comparison of two popular feature sets computed for heart beat classification. The first set consists of six statistical features computed in the time domain, namely Time Mean, Time Kurtosis, Time Skewness, Time Energy, Time Std, Zero crossings and the second feature set consists of two morphological features, the QRS

Interval and the RR Interval. Tests have been performed in providing the two feature sets individually and togather to two different classification algorithms, the SVM and the MLP neural network. The results indicate that the MLP provides better results when either statistical time domain features are used or when the sets of features are combined together. On the other hand, SVM provides relatively better performance when considering then case of using morphological features only.

Although the current work provides a comprehensive discussion on the use of time domain and morphological features and the use of the SVM and MLP networks for classification of different beat types of the heart, future work in this area may involve the computation of other types of features and the inclusion of more classification algorithms.

REFERENCES

[1] (Website)World Health Organization, Cardiovascular diseases- Key Facts. Available http://www.who.int/mediacentre/factsheets/fs317/en/

[2] (Website)Centre for Disease Control and Prevention, Deaths and Mortality. Available https://www.cdc.gov/nchs/fastats/deaths.htm

[3] (Website) American Heart Association, About Arrythmia. Available http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/About-Arrhythmia_UCM_002010_Article.jsp#.WHsd5i9_e00

[4] (Website) American Heart Association, Ventricular Fibrillation. Available http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/Ventricular-Fibrillation_UCM_324063_Article.jsp#.WHseBC9_e00

[5] T. Barman, R. Ghongade, and A. Ratnaparkhi, "Rough set based segmentation and classification model for ECG," in *Advances in Signal Processing (CASP), Conference on*, , pp. 18–23, 2016

[6] M. S. Rattar, S. M. Shehram Shah, B.S. Chowdhry, S. M. Z. Abbas Shah,"Detection of Ventricular Arrhythmia from ECG" Pakistan Journal of Computer and Information Systems (PJCIS), vol. 1, no. 1, pp. 17-28, October 2016

[7] R. Acharya, S. M. Krishnan, J. A. E. Spaan, and J. S. Suri, *Advances in cardiac signal processing*. Springer, 2007.

[8] E. Alickovic and A. Subasi, "Effect of multiscale PCA de-noising in ECG beat classification for diagnosis of cardiovascular diseases," *Circuits, Syst. Signal Process.*, vol. 34, no. 2, pp. 513–533, 2015.

[9] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, no. 4, pp. 433–459, 2010.

[10] S. Sahoo, B. Kanungo, S. Behera, and S. Sabut, "Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities," *Meas. J. Int. Meas. Confed.*, vol. 108, pp. 55–66, 2017.

[11] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification," *IEEE Trans. neural networks*, vol. 3, no. 5, pp. 683–697, 1992.

[12] I. Steinwart and A. Christmann, *Support vector machines*. Springer Science & Business Media, 2008.

[13] E. Alickovic and A. Subasi, "Medical Decision Support System for Diagnosis of Heart Arrhythmia using DWT and Random Forests Classifier," 2016.

[14] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[15] J. R. Quinlan, *C4. 5: programs for machine learning*. Elsevier, 2014.

[16] B. Surendiran and A. Vadivel, "Classification and regression Tree classifier for classifying benign and malignant masses in digital mammogram using shape properties," in *proc. Of International conference on Computing Technologies (ICONCT'09), India*, , pp. 47–52, 2009

[17] K. A. Alfarhan, M. Y. Mashor, A. Rahman, M. Saad, H. A. Azeez, and M. M. Sabry, "Effects of the Window Size and Feature Extraction Approach for Arrhythmia Classification," vol. 30, pp. 1–11, 2017.

[18] I. Daubechies, *Ten lectures on wavelets*. SIAM, 1992.

[19] C. F. Chen and C. H. Hsiao, "Haar wavelet method for solving lumped and distributed-parameter systems," *IEE Proceedings-Control Theory Appl.*, vol. 144, no. 1, pp. 87–94, 1997.

[20] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 2, pp. 186–200, 1996.

[21] A. F. Khalaf, M. I. Owis, and I. A. Yassine, "Expert Systems with Applications A novel technique for cardiac arrhythmia classification using spectral correlation and support vector machines," *Expert Syst. Appl.*, no. July, 2015.

[22] U. Desai, R. J. Martis, C. G. Nayak, K. Sarika, S. G. Nayak, A. Shirva, V. Nayak, and S. Mudassir, "Discrete cosine transform features in automated classification of cardiac arrhythmia beats," in *Emerging research in computing, information, communication and applications*, Springer, pp. 153–162, 2015

[23] G. Strang, "The discrete cosine transform," *SIAM Rev.*, vol. 41, no. 1, pp. 135–147, 1999.

[24] J. M. D. Bullas, "K-nearest neighbours with weighted linear regression.," 1999.

[25] M. Sarfraz, A. A. Khan, and F. F. Li, "Using independent component analysis to obtain feature space for reliable ECG Arrhythmia classification," *Proc. - 2014 IEEE Int. Conf. Bioinforma. Biomed. IEEE BIBM 2014*, pp. 62–67, 2014.

[26] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*, vol. 46. John Wiley & Sons, 2004.

[27] V. H. C. de Albuquerque *et al.*, "Robust automated cardiac arrhythmia detection in ECG beat signals," *Neural Comput. Appl.*, pp. 1–15, 2016.

[28] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," in IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45-50, May-June 2001.

[29] A.L. Goldberger, L.A.N. Amaral, L. Glass, et al., "Physiobank, physiotoolkit, and physionet.," *Circulation*. vol. 101, no. 23, pp. e215–e220, 2000.

[30] A. Gacek and W. Pedrycz, *ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence*. Springer Science & Business Media, 2011.

Aunsa Shah is a Masters student of Electronic Systems Engineering at Mehran University of Engineering and Technology, Jamshoro, Pakistan and completing research requirements of her Masters degree as an exchange student in Selcuk University, Konya, Turkey. She recieved a Bachelor's degree in Electronics Engineering from the Mehran University of Engineering and Technology, Jamshoro in 2016. She has worked on several projects during and after her undergraduate students. Moreover, she has been volunteering in various seminars and conferences held her department. After completing her bachelors, she served as an intern at the University of Sindh, Jamshoro, Pakistan. She has a publication in her name. Her research interests include signal processing for health and data science.

Dr. Dr. Seral Özşen (PhD), Selcuk University, Electrical and Electronics Engineering Department, Konya, Turkey. She was born in 1980 in Aydın, Turkey. She graduated from Electrical and Electronics Department of Ege University in 2002. Then she fullfilled master and PhD programs in 2004 and 2008 respectively in Selcuk University Eelctrical & Electronics Engineering Department. She is now working in the same department as Associate Prof.

Syed Muhammad Zaigham Abbas Shah is working as an Assistant Professor in the Department of Electronics Engineering, Mehran University of Engineering and Technology, Pakistan. He received his BEng. in Electronic Engineering from Mehran University of Engineering and Technology, Pakistan as a silver medalist and an MSc in Electrical and Electronic Engineering from the University of Strathclyde, UK. He has been teaching subjects related to Digital Instrumentation, Embedded System Design in Undergraduate as well as Postgraduate programs at Mehran UET. His research interests include Image processing for Monitoring applications, Embedded Systems for assisted living, Machine learning for disease diagnosis. He has publications in the mentioned areas in international conferences and journals. He is also the co-author of the book "The First practical book on Electronics Workshop" which has been published by River Publishers, Germany