A Cardioid Based Technique to Identify Premature Ventricular Contractions

Vu Mai[1], Ibrahim Khalil[2]

[1,2]RMIT University, Melbourne, Victoria, Australia

Abstract

Premature Ventricular Contraction (PVC) can occur in healthy people of any age and is linked to mortality when associated with myocardial infarction. There has been voluminous research that focuses on PVC detection. However, pre-processing techniques such as wavelet filtering may cause delay by extracting features on frequency domain rather than time domain. Furthermore, many classifiers are only suitable for powerful processing systems. In this paper, we propose a new patient-specific classification technique to detect PVC by using two dimensional cardioid loops obtained from QRS complexes. Pre-processing time is reduced significantly since cardioid loops can be drawn directly from raw QRS complexes. The feature set comprises x-y coordinates of centroid, upper, lower, left and right extreme points of each cardioid loop. We conducted experiments over 20 subjects of the MIT/BIH arrhythmia database and obtained an average detection accuracy of 99.60%, average sensitivity of 97.43%, and average positive predictive value of 98.62%.

1. Introduction

Premature Ventricular Contraction (PVC) is a relatively common cardiac event where the heartbeat is initiated by the heart ventricles rather than by the sinoatrial node. Many studies have shown that PVCs, when associated with myocardial infarction, can be linked to mortality [1]. PVC is difficult to detect because patients usually seek medical attention only when symptoms persist. Doctors have to look over long records of ECG to identify and classify abnormal beats.

Automatic PVC detection techniques using neural networks as classifiers have achieved high accuracy rates. Shyu, et al. [2] achieved 97.04% accuracy for PVC beat classification using wavelet feature extraction and fuzzy neural network as classifier. They use ECG data of seven patients in Physionet's MIT/BIH arrhythmia database, two of which were in the training set of their neural network. Ham et al. [3] used the mean square value of each QRS segment and two transformed linear predictive coding coefficients as input features for PVC classification with PVC detection rates up to 97%. Their experimental results, however, were based on data of six patients in the MIT/BIH arrhythmia database.

Other researchers trained and tested their classifiers over large datasets and obtained fairly lower classification accuracy. Using self-organising maps and learning vector quantisation, Hu, et al. [4] obtained 62% accuracy over an ECG dataset of 20 patients in the MIT/BIH arrhythmia database. Inan, et al. [5] proposed an innovative approach to differentiate between normal and PVC beats. ECG timing information is combined with other features extracted on frequency domain using Wavelet transformation. They also trained and tested a Multilayer Perceptron neural network classifier with a large ECG data set. The accuracy was 95.16% over 93,281 heartbeats of 40 patients in the MIT/BIH arrhythmia database.

Pre-processing techniques such as wavelet filtering used in [2] and [5] may cause delay by extracting features on frequency domain rather than time domain. Using raw ECG data of the whole QRS complex or other Fiducial Point Detection mechanisms (i.e. measuring the on-set and off-set of P wave, QRS complex and T wave) to obtain morphological features often require considerable amount of processing power. Furthermore, many classifiers are only suitable for powerful processing systems. Significant differences in PVC and normal beat morphologies in large ECG datasets might also reduce the accuracy rates of those proposed methods.

The idea of cardiac arrhythmias detection by using features of cardioid loops drawn from QRS complexes has been proposed recently [6]. The authors suggest that their technique can be used to identify ventricular abnormality with only five points (i.e. centroid and upper, lower, left and right extreme). This research aims to take their work a step further by using a supervised data mining technique to automatically generate patterns that can differentiate normal and PVC heartbeats from the features suggested in their research. Using Multilayer Perceptron (MLP) neural network as classifier, experiments were performed over 20 subjects of the MIT/BIH arrhythmia database. We evaluated the classification accuracy of cardioid technique by common diagnostic measures often used by other researchers such as accuracy, sensitivity and positive predictive value [5].

2. Background

This section briefly introduces necessary background topics such as Premature Ventricular Contraction and its waveforms. A brief description of Cardioid feature extraction technique is provided together with background information about Multilayer Perceptron neural network which is used as our classifier.

2.1. Premature ventricular contraction

If a heartbeat starts in the ventricles, it is called premature ventricular contraction (PVC). Two typical PVC morphologies found in ECG recordings are shown in Figure 1 in which normal heartbeats are interrupted by early beats. These early beats are caused by the premature discharge of an ectopic pacemaker. When PVC occurs, circulation is inefficient because the ventricles have not received enough blood from the atria to be able to pump a maximum amount of blood both to the lungs and to the rest of the body. Although PVC is a relatively common event and can occur to healthy individuals, it may be a precursor to ventricular tachycardia, especially in groups of patients who have history of cardiac disorders.

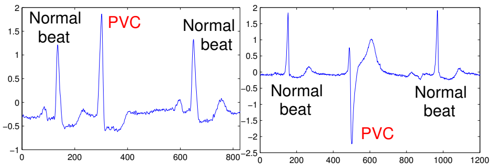

Figure 1. Two PVC morphologies found in many ECG records

2.2. Cardioid technique

Cardioid eliminates the time dimension and isolates morphological features of ECG. QRS complexes of ECG signal are extracted to create 2-dimensional closed loops. It is assumed that positions of QRS complexes of heartbeats in an ECG graph have already been detected by another beat detection algorithm or have been annotated by a cardiologist. Cardioid only predicts from QRS complex whether a heartbeat is a normal heartbeat or a PVC beat. This technique is described in the following steps:
1. QRS complexes in an ECG record are extracted. A QRS complex is represented in (1).

$$X = \{x(1), x(2), x(3), ..., x(N)\} \quad (1)$$

where X represents a set of electric current values obtained from the QRS complex with $x(1)$, $x(2)$ are particular values and N is the length of the series.

2. Subtract the value of one point from the next point in the set X and store results in set Y as in (2)

$$Y = \{x(2)-x(1), x(3)-x(2), ..., x(N)-x(N-1)\} \quad (2)$$

3. A two dimensional loop is generated as a scatted x-y graph. A point in this graph has its y-coordinate from an element in set Y and its x-coordinate from the corresponding element in set X.

These steps are repeated for each QRS complex. Cardioid loops of normal QRS complexes are very similar to each other, while cardioid loops of a PVC QRS complex will be significantly different from normal ones as in Figure 2. A neural network classifier can be trained to recognise these differences, and hence, can classify normal and PVC heartbeats.

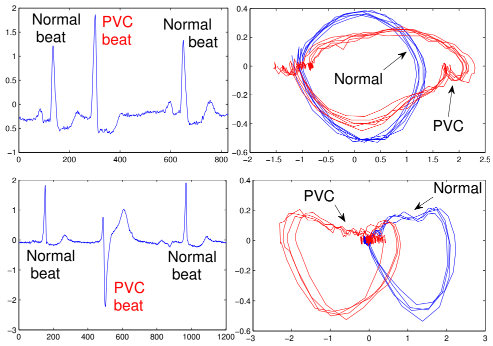

Figure 2. PVC beats display different cardioid loops compared to normal beats. Each loop represents a QRS complex of a heartbeat

3. Methods

In this section, details of our method are provided and implemented. Feature extraction is the first step in which we attempt to isolate crucial morphological features. Next, the architecture of MLP neural network classifier is described together with its input and output. Finally, the effectiveness and robustness of our method are evaluated using common performance metrics such as accuracy, sensitivity and positive predictive value.

3.1. Feature extraction

At this stage, we create cardiod loops of all ECG heartbeats to be classified. The centroid of a loop is calculated using Equation (3)

$$\text{centroid} = \left[\frac{\sum_{i=1}^{N} x_i}{N}, \frac{\sum_{i=1}^{N} y_i}{N} \right] \quad (3)$$

Then x, y coordinates of left, right, upper and lower extreme points are calculated as in Figure 3. Each cardioid loop is uniquely represented by these coordinates. Therefore, the feature set is reduced to only 10 dimension, instead of using 60 to 80 dimension if raw QRS complexes are used. This 10-dimension feature set helps Cardioid to be a fast and efficient technique and is sufficient to represent the similarity between normal QRS complexes and the differences between normal and PVC QRS complexes.

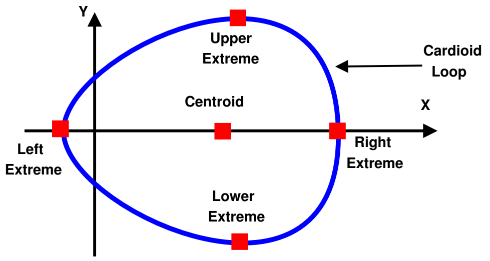

Figure 3. x, y coordinates of the extreme points and the centroid are used to form a 10-dimension feature set to uniquely represent each Cardioid loop

3.2. MLP neural network architecture

A two-layer neural network was used in our experiments. In each layer, there were many artificial neurons using sigmoid transfer functions as step functions. The number of neurons in the hidden layer was chosen manually by experiments so that it was enough for the network to converge but also not too many that overfitting might occur. There was one neuron in the output layer to classify whether a heartbeat is PVC or not. Figure 4 shows the structure of one of the neurons we used in the hidden layer. In this figure, vector p is the input connected to each

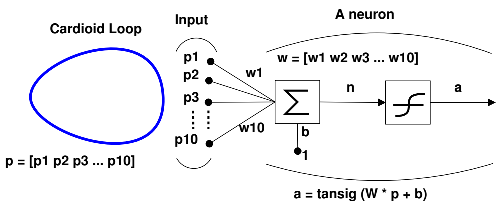

Figure 4. Structure of a neuron used in MLP neural network

neuron in the hidden layer by a weight matrix W, b is the bias. The net input n is passed through a sigmoid transfer function $tansig$ to produce the output a.

3.3. Performance metrics

Some common metrics used in many researches such as accuracy, sensitivity, and positive predictive value are used to measure the performance of our classifier. A confusion matrix such as the one shown in Figure 5 is very useful to understand these concepts. Sensitivity (Sen) and positive

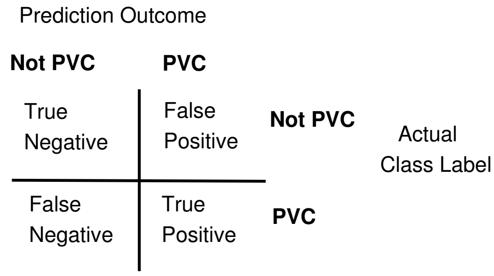

Figure 5. Confusion Matrix

predictive value (PPV) are specific to each class. They are calculated as followed:

$$\text{Sen} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{4}$$

$$\text{PPV} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{5}$$

Accuracy (Acc) shows the general performance of the classifier on all types of beats and is calculated in equation (6).

$$\text{Acc} = \frac{\text{True Negative} + \text{True Positive}}{\text{Total Number of Instances}} \tag{6}$$

4. Results

ECG data used for cardioid loop construction should be recorded via lead II so that cardioid loops are closed loops and extremes points are identified clearly. The more PVC heartbeats in each ECG record, the better the classification accuracy will be. These data selection criteria helped us to select 20 records from the MIT-BIH Arrhythmia database. Having experienced early symptoms of PVC, their 30-minute ECG records contain a significant number of PVC beats along with normal beats as shown in [7]. For each record, we used all PVC and normal beats to train and test an MLP neural network in Weka with 10-fold cross validation to obtain the best estimation of the classifier performance. Finally, results including Patient ID, Accuracy, Sensitivity and Positive Predictive Value were recorded in Table 1.

5. Discussion

Cardioid technique has proven to be effective in detecting PVC by having high average Accuracy (99.60%), Sensitivity (97.43%), and Positive Predictive Value (98.62%). It is also very efficient due to its 10-dimension feature set, which is significantly smaller compared to a 60-dimension feature set needed if raw ECG data is used directly. Cardioid technique is faster than many other classification

PatientID	Acc (%)	Sen (%)	PPV (%)
mit105	99.81	92.68	95
mit106	99.75	99.42	99.61
mit114	99.89	97.67	97.67
mit116	99.75	96.33	98.13
mit119	100	100	100
mit200	99.14	98.18	99.14
mit201	99.84	99.49	98.99
mit203	97.17	88.51	92.25
mit205	99.96	98.59	100
mit208	99.81	99.7	99.8
mit210	99.01	90.21	96.15
mit213	99.79	98.64	98.64
mit215	99.88	98.78	98.78
mit219	99.86	95.31	100
mit221	100	100	100
mit223	99.6	98.52	99.36
mit228	99.61	98.62	99.17
mit233	99.87	99.76	99.76
mit217	99.26	98.15	100
mit202	100	100	100
Average	**99.60**	**97.43**	**98.62**

Table 1. PVC detection results obtained by experiments on 20 patients

techniques using wavelet transformation, because it extracts features directly on time domain rather than frequency domain like wavelet does.

However, using neural network as classifier does have some limitations. Training data is required. Neural network cannot recognise PVC heartbeats that it hasn't been trained on. Therefore, the technique described in this paper can only be applied in some scenarios in which the patient has experienced early symptoms of PVC. For example, a patient can go to a doctor to report chest pain or palpitations. Then the doctor recommends Holter monitoring to monitor ECG of that patient in 24 hours. After reading ECG obtained from Holter monitoring, doctor might identify some PVC heartbeats and use them as training data to train the neural network so that it can be used to continue monitoring and raise alarm if PVC happens again.

Sometimes, the training data set is imbalanced because the number of PVC heartbeats could be very small compared to the number of normal heartbeats, limiting the amount of training data. On the other hand, Multilayer Perceptron neural networks can only be trained to generalise well within the range of training inputs but cannot extrapolate beyond this range. As a consequence, accuracy rate and other diagnostic measures of the technique might be reduced. To overcome this limitation, timing information (i.e. the time distances between PVC heartbeat and its previous and following beat), which is very distinctive, could be added as another important feature in the feature set.

6. Conclusion

We have achieved high accuracy rates by using cardioid technique to classify normal and PVC heartbeats. We also show that cardioid is a fast and efficient technique to detect PVC. However, many other factors such as limited training data, imbalanced dataset and vast amount of differences in ECGs of many patients still limit us from obtaining higher accuracy. This technique, however, can be very effective in certain scenarios, especially when patients have just experienced early symptoms of PVC.

In future studies, we will examine other classifier implementations which may provide higher classification results. A possible approach is using a classifier such as a delayed-input neural network, which looks at a series of beats. The timing information would then be inherent, as the waveform morphology for a series of beats will depend greatly on timing. More fast and efficient feature extraction techniques could also be explored such as compressed ECG data. We can also combine cardioid technique with ECG biometrics and many ECG data compression and encryption schemes to provide a real-time, mobile healthcare and tele-monitoring system.

References

[1] I. Atsushi M. Hwa AHTL, Narayan SM. Abnormal heart rate turbulence predicts the initiation of ventricular arrhythmias. Pacing Clin Electrophysiol ;11:1189–1197.

[2] Shyu L, Wu Y, Hu W. Using wavelet transform and fuzzy neural network for VPC detection from the Holter ECG. Biomedical Engineering IEEE Transactions on 2004; 51(7):1269–1273. ISSN 0018-9294.

[3] Ham F, Han S. Classification of cardiac arrhythmias using fuzzy ARTMAP. Biomedical Engineering IEEE Transactions on 1996;43(4):425–429. ISSN 0018-9294.

[4] Hu Y, Palreddy S, Tompkins W. A patient-adaptable ECG beat classifier using a mixture of experts approach. Biomedical Engineering IEEE Transactions on 1997;44(9):891–900. ISSN 0018-9294.

[5] Inan O, Giovangrandi L, Kovacs G. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. Biomedical Engineering IEEE Transactions on 2006;53(12):2507–2515. ISSN 0018-9294.

[6] Sufi F, Khalil I, Tari Z. A cardiod based technique to identify Cardiovascular Diseases using mobile phones and body sensors. In Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE. ISSN 1557-170X, 2010; 5500–5503.

[7] Physionet. MIT-BIH Arrhythmia Database Directory. `physionet.org/physiobank/database/html/mitdbdir/mitdbdir.htm`, 1997.