

## CHARACTERIZATION OF POST-MI ELECTROCARDIOGRAM USING POWER RATIO FEATURES AND K-NEAREST NEIGHBOR CLASSIFIER

M. S. A. M. Ali<sup>1,\*</sup>, A. H. Jahidin<sup>2</sup>, I. M. Yassin<sup>1</sup>, I. Pasya<sup>3</sup>, M. F. A. Khalid<sup>3</sup> and Z. Awang<sup>3</sup>

<sup>1</sup>Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

<sup>2</sup>Centre for Foundation Studies in Science, University of Malaya, Kuala Lumpur, Malaysia

<sup>3</sup>Microwave Research Institute, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia

Published online: 05 October 2017

---

### ABSTRACT

Myocardial infarction (MI) is the irreversible necrosis of heart muscles caused by prolonged ischemic condition. Subsequently, the presence of damaged tissues in post-MI patients is expected to have an effect on their electrocardiogram (ECG). Hence, this paper proposes characterization of post-MI ECG from bipolar and augmented limb leads. Data from patients of inferior and anterior MI, as well as healthy controls are acquired from PTB Diagnostic ECG Database. Noise removal is performed using band-pass filter. Power ratio is then extracted from energy spectral density of bipolar and augmented limb leads. Classification via k-nearest neighbor method revealed that features from the augmented limb leads outperform bipolar limb leads in terms of training and testing accuracies. Hence, power ratio features from the augmented limb leads exhibit higher degree of separation between the ECG classes.

**Keywords:** myocardial infarction; electrocardiogram; power ratio; k-nearest neighbor.

---

Author Correspondence, e-mail: [megatsyahirul@salam.uitm.edu.my](mailto:megatsyahirul@salam.uitm.edu.my)

doi: <http://dx.doi.org/10.4314/jfas.v9i4s.56>

---

## 1. INTRODUCTION

Electrocardiogram (ECG) is a non-invasive electrical recording of cardiac activities. The bio-potential signal originates from ionic fluctuations that occur during propagation of impulse throughout the heart's conduction system. The fluctuations can be detected by specialized electrodes that are placed on the left and right arms as well as left leg, which effectively forms the limb lead system. While bipolar configuration measures the potential difference between pairs of electrodes, the augmented limb leads are unipolar configurations that utilize Goldberger's central terminal as a common reference. The six leads from bipolar and augmented limb leads view the electrical activities of the heart from the frontal plane. The multiple lead systems provide opportunity for cardiologists to localize sources of abnormality within the heart's conduction system [1]. Arrhythmias are a manifestation of anomalies that occurs within the anatomical structure of the heart. Among the widely researched heart conditions include bundle branch blocks[2-3], cardiomyopathy[4-7]and premature ventricular contraction [8]. Explicitly, necrosis of the myocardium is triggered by deprivation of oxygen for prolonged periods of time. The ischemic conditions are usually caused by coronary artery disease. Without proper treatment, acute myocardial infarction often results in cardiac arrest and death [9].

In an effort to develop computerized ECG diagnostic systems, extensive studies have been conducted to identify suitable features for distinguishing arrhythmias from normal sinus rhythms. Various pre-processing and feature extraction techniques have since been proposed which include time-domain morphological features [10-11], wavelet transforms [8] and spectral analysis. Frequency range of 1 Hz to 20 Hz which is generally associated with QRS complex has shown to be very reliable for discriminating between arrhythmias and normal sinus rhythm. Frequency range lower than 1 Hz is associated with baseline drift and breathing. Meanwhile, the power starts to deteriorate beyond 12 Hz. Hence, the frequency content of less than 20 Hz is not exposed to power line interference and noise from electromyogram [12]. Signal power of individual ECG varies due to the dissimilar body composition [1]. As the information in one lead is correlated with each other, one suitable approach is to normalize the energy from individual lead against the total energy within the same lead system. The obtained results would thus, quantify the inter-relationship of energy between the various

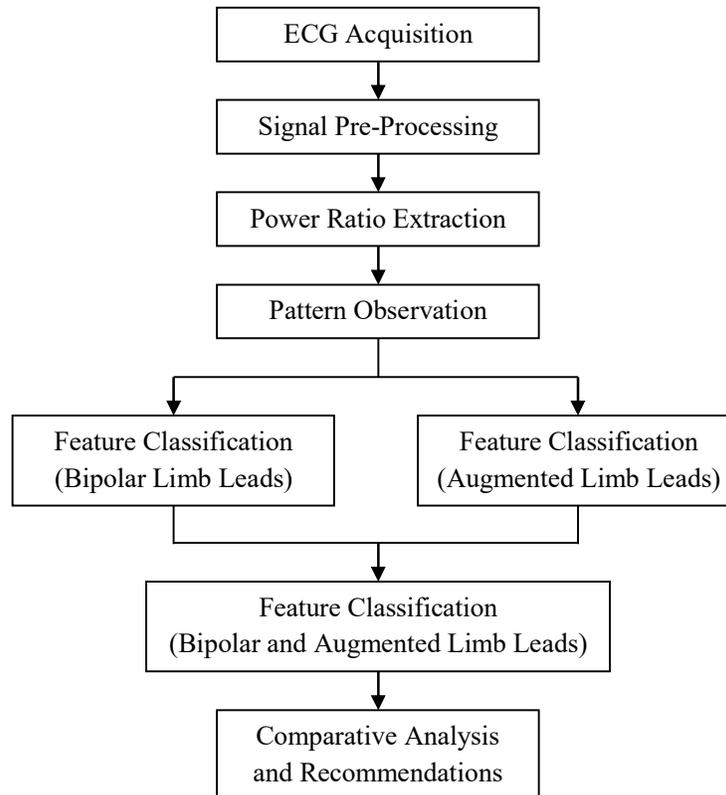
leads. Such approach has yet to be implemented in ECG studies, but has been successfully applied to electroencephalogram (EEG) to quantify the inter-relationship of different EEG bands in various cognitive processes [13-14].

Establishment of new pre-processing and feature extraction algorithms are often complemented with statistical or intelligent modelling methods. These include instance-based learning algorithms such as Bayesian network and k-nearest neighbor (kNN) [15], to advanced techniques such as artificial neural network [3-4] and support vector machines [16]. Each of these methods has unique advantages and is implemented to address specific research needs. In a study requiring investigation of new ECG features, kNN would be suitable as it requires minimal processing and proven robust with sufficiently large dataset [17]. kNN is a non-parametric classification algorithm that has been implemented in a wide range of pattern recognition studies. The method is based on the rule that instances within a set of data should exist in close proximity to other instances of similar attributes. Nevertheless, such approach is highly reliant on selection of neighborhood distance,  $k$ . While a simple approach is to conduct repetitive trials for varying  $k$  and assess the performance, a more constructive approach is proposed via cross-validation techniques [18].

Despite the recent progress and developments, none have focused on characterizing the pattern of ECG with history of acute MI. Therefore, the study aims to investigate the ECG of post-MI patients and healthy controls. Samples from post-MI will also be segregated into inferior and anterior MI. The pattern of features will be observed using standard error of mean and classified using kNN with  $k$ -fold cross-validation.

## 2. METHODOLOGY

The overall framework of the study is comprised of ECG acquisition, signal pre-processing and extraction of power ratio from bipolar and augmented limb leads, pattern observation via standard error of mean, and feature classification using kNN with  $k$ -fold cross-validation. Classification is performed for various combinations of features based on the bipolar and augmented limb leads systems. Fig. 1 summarizes the flow of research methods.



**Fig.1.** Flowchart of research methods

## 2.1. ECG Acquisition

ECG from thirty-five subjects is acquired from the PTB Diagnostic ECG Database [19]. Seven are patients of inferior MI, while another eight are survivors of anterior MI. Based on the historical records, all fifteen patients experienced acute MI once. Samples from the other twenty subjects are used as healthy control.

## 2.2. Signal Pre-Processing and Extraction of Power Ratio Features

All signal pre-processing and feature extraction is executed in MATLAB. Baseline correction and removal of power line interference is performed by applying band-pass finite impulse response filter (FIR) between 1 Hz to 20 Hz. The frequency range is vital as it is associated with the behavior of QRS complex. Hence, any deviation from normal sinus rhythm can be quantified within the specified spectrum. FIR filter is being adopted due to its linear phase response property which avoids distortion of phase information among the lead systems [20]. Subsequently, the entire length of signal is segregated into smaller 5 seconds segments.

Fast Fourier transform (FFT) algorithm is then implemented to estimate the power spectral

density (PSD) via Welch method. Each signal segment is divided into 50% overlapping epochs, with FFT length of 1024. The processing approach adopts the Hamming window function. Energy spectral density (ESD) is then obtained by determining the area under the PSD curve. Subsequently, the information between the bipolar limb leads is normalized through the power ratio relationships expressed by Equation (1), (2) and (3). For  $ESD_I$ ,  $ESD_{II}$  and  $ESD_{III}$ , each represent the ESD of Lead I, Lead II and Lead III.

$$\text{Power Ratio I} = \frac{ESD_I}{ESD_I + ESD_{II} + ESD_{III}} \quad (1)$$

$$\text{Power Ratio II} = \frac{ESD_{II}}{ESD_I + ESD_{II} + ESD_{III}} \quad (2)$$

$$\text{Power Ratio III} = \frac{ESD_I}{ESD_I + ESD_{II} + ESD_{III}} \quad (3)$$

Likewise, the power ratio relationship between the augmented limb leads is expressed by Equation (4), (5) and (6). For  $ESD_{aVR}$ ,  $ESD_{aVL}$  and  $ESD_{aVF}$ , each represent the ESD of Lead aVR, Lead aVL and Lead aVF.

$$\text{Power Ratio aVR} = \frac{ESD_{aVR}}{ESD_{aVR} + ESD_{aVF} + ESD_{aVL}} \quad (4)$$

$$\text{Power Ratio aVL} = \frac{ESD_{aVL}}{ESD_{aVR} + ESD_{aVF} + ESD_{aVL}} \quad (5)$$

$$\text{Power Ratio aVF} = \frac{ESD_{aVR}}{ESD_{aVR} + ESD_{aVF} + ESD_{aVL}} \quad (6)$$

Based on historical data, power ratio features from bipolar and augmented limb leads are then clustered into the respective ECG classes. The patterns are observed in SPSS using standard error of mean. These reveal the variability of features within the control groups.

### 2.3.k-Nearest Neighbor and k-Fold Cross-Validation

The classification algorithm for kNN is relatively unsophisticated. The unlabeled power ratio features will be identified based on the rule of majority. The classifier initially memorizes the features from training dataset based on the corresponding ECG class labels. Subsequently, the unlabeled power ratio features from the testing dataset will be classified by assigning the most frequent ECG class label with k nearest training samples. In this study, the classification

adopts Euclidean distance and is performed for  $k=1$  to  $k=5$ . 80% of the samples are used for training and the other 20% for testing.

The overall classification performance is assessed using accuracy (Acc), positive predictivity (Pp) and sensitivity (Se). The parameters are each expressed by Equation (7), (8) and (9) where TP, is the true positives, TN is true negatives, FP is false positives and FN is false negatives.

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100\% \quad (7)$$

$$\text{Pp} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100\% \quad (8)$$

$$\text{Se} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (9)$$

Subsequently, the true performance is assessed by incorporating k-fold cross-validation into the kNN algorithm. Prior to classification, a disjointed training and datasets are formed using random sampling method. The cross-validation estimate is the total of correct classification that is averaged over the number of folds within the dataset. Hence, features are assumed stable for specific set of data and perturbations if similar prediction is being induced with different set of perturbations.

In this study, the dataset is randomly segregated into five disjointed folds. At any one time, four folds are used for training while the remaining one fold is used for testing. These match the 80:20 ratio of training and testing dataset that is implemented in kNN. At varying k iteration, different combination of folds will form the datasets for training and testing. Therefore, the classifier is trained and tested for five instances. The results are then averaged to obtain the true classification performance.

In an attempt to compare the effectiveness of different lead systems for discriminating ECG of post-MI patients and the healthy controls, the classifier will be trained and tested separately for bipolar and augmented limb leads. Extended investigation is also performed using combination of both lead systems.

### 3. RESULTS AND DISCUSSION

The findings initially discusses on clustering of data. This is followed by observation of mean power ratio patterns for bipolar and augmented limb leads. Subsequently, results pertaining to

classification of power ratio features are elaborated based on the different lead systems.

### 3.1.Data Clustering

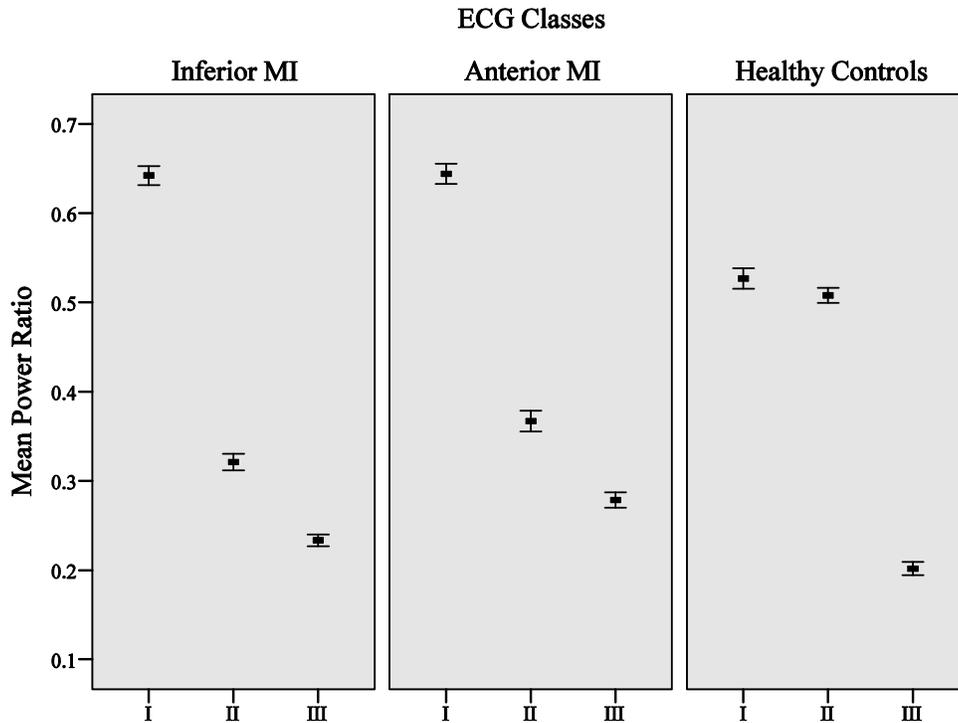
Samples are clustered into inferior MI, anterior MI and healthy controls based on the historical record from PTB Diagnostic ECG Database. For classification purposes, each ECG class has been assigned unique index label. Extraction of power ratio features has resulted in 1080 samples for each group. Table 1 summarizes the distribution of samples according to the ECG classes.

**Table 1.**Distribution of sample based on the ECG classes

ECG Class	No. of Subjects	Index	No. of Samples
Inferior MI	7	1	1080
Anterior MI	8	2	1080
Healthy Controls	20	3	1080

### 3.2.Power Ratio for Bipolar Limb Leads

Mean power ratio and standard error for bipolar limb leads are shown in Fig. 2. Through visual inspection, the highest mean for Lead I is attained by ECG with anterior MI, followed by inferior MI. Healthy controls attained the lowest mean power ratio for Lead I. It is notable that the feature for Lead I have high degree of overlapping between standard error mean for inferior and anterior MI. Meanwhile, the highest mean for Lead II is attained by the ECG of healthy controls. This is followed by those with anterior MI and then, the inferior MI. Conversely, the highest mean power ratio for Lead III is achieved by ECG with anterior MI, followed by inferior MI. The healthy controls have attained the lowest mean power ratio for Lead III. No overlapping of standard error has been observed for the features in Lead II and Lead III.



**Fig.2.**Standard error of mean for power ratio features from bipolar limb leads

### 3.3.Power Ratio for Augmented Limb Leads

Separately, Fig. 3 demonstrates the mean power ratio and standard error for augmented limb leads. Through similar visual inspection, the highest mean for Lead aVR is attained by healthy controls. This is followed by those with inferior MI and subsequently, anterior MI. Meanwhile, the highest mean power ratio for Lead aVL is achieved by ECG with inferior MI, followed by anterior MI. The lowest mean for Lead aVL is attained by healthy controls. Conversely, the healthy controls obtained the highest mean power ratio for Lead aVF. This is subsequently followed by ECG with anterior MI. Those with inferior MI attained the lowest mean power ratio for Lead aVF. Comparatively, no overlapping in standard error has been observed for the features in Lead aVR, Lead aVL and Lead aVF, therefore indicating better separation than the bipolar limb leads.

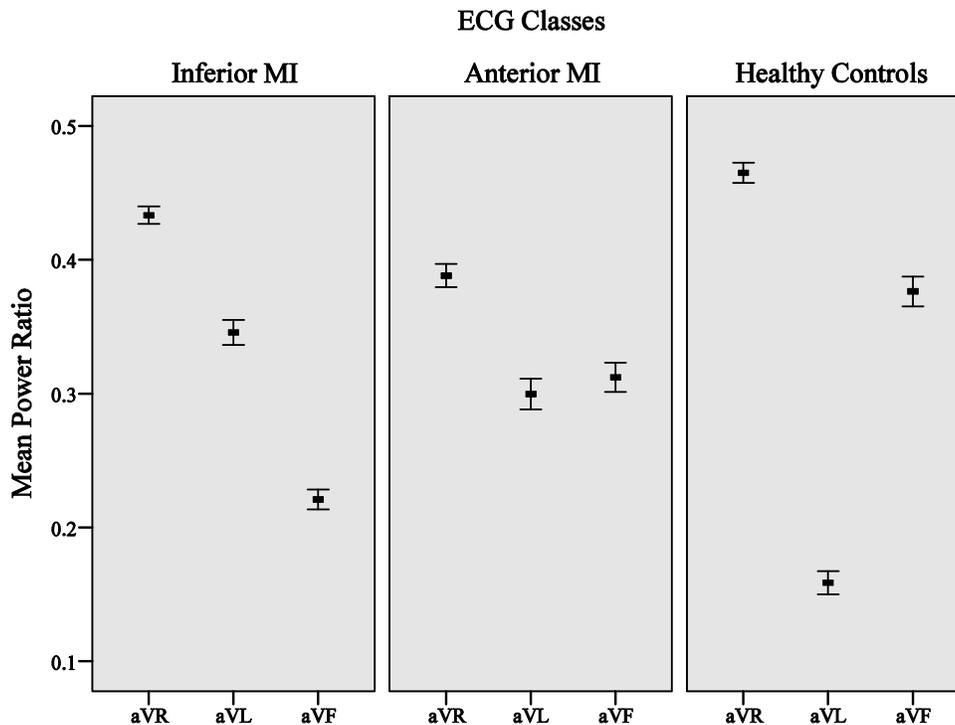


Fig.3. Standard error of mean for power ratio features from augmented limb leads

### 3.4. Classification of Power Ratio: Bipolar Limb Leads

Initially, classification is performed separately between power ratio features of bipolar and augmented limb leads. The mean classification accuracies for power ratio from bipolar limb leads are shown in Fig. 4. In principle,  $k=1$  indicates classification of similarly labeled features. This produces optimum results with minimum intrusion from features of other ECG classes. The neighboring distance however is positively correlated with probability of disturbance. Should the classification attain similar performance even with increasing distance, this indicates good separation of features among the ECG classes. Hence, the optimum performance is assessed based on the highest accuracy for the furthest neighboring distance.

The best classification performance for power ratio from bipolar limb leads are achieved at  $k=2$  with 100% accuracy for training and 89.4% accuracy for testing. The performance however deteriorates with increasing distance. At  $k=5$ , the training accuracy recorded a reduction of 7.5% and testing accuracy declines by 2.3%. This reflects the effect of interference from features of other ECG classes. For each  $k$ , testing accuracy yielded lower results than training due to smaller sample size and overlapping of features between the groups. Positive predictivity and

sensitivity for each control groups at k=2 is shown in Table 2. The classifier attained perfect performance during the training stage. For testing however, the ECG classes attained positive predictivity and sensitivity of approximately 89.4%.

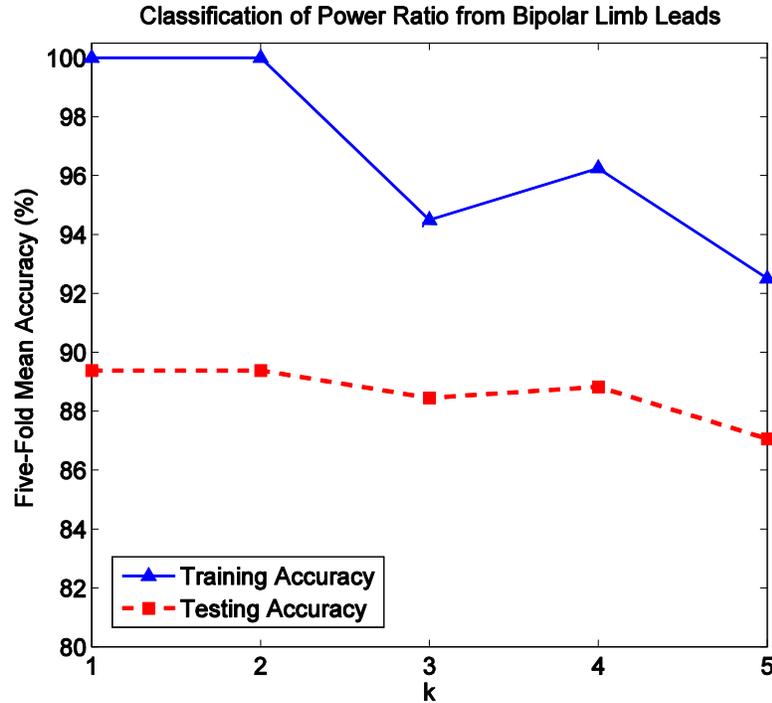


Fig.4. Mean accuracies for classification of power ratio features from bipolar limb leads

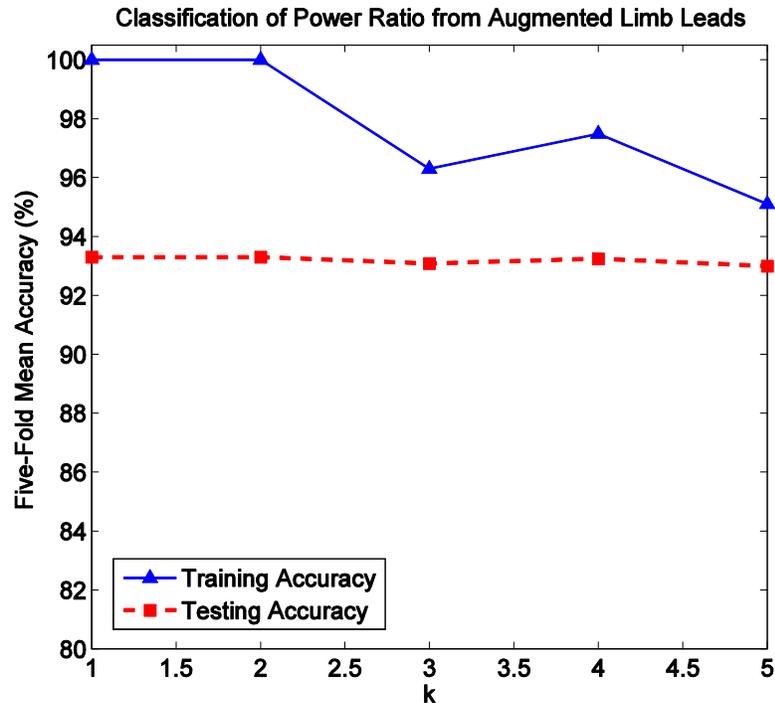
Table 2. Five-fold mean positive predictivity and sensitivity for classification of power ratio features from bipolar limb leads at k=2

ECG Classes	Training		Testing	
	Pp (%)	Se (%)	Pp (%)	Se (%)
Inferior MI	100	100	90.3	89.9
Anterior MI	100	100	88.4	89.0
Healthy Controls	100	100	89.5	89.3

### 3.5. Classification of Power Ratio: Augmented Limb Leads

The best classification performance for power ratio from augmented limb leads are attained at k= 2 with 100% training accuracy and 93.3% testing accuracy. Similarly, the performance deteriorates with increasing neighboring distance. At k=5, the training accuracy declines by4.9% and testing accuracy by0.3%. Mean positive predictivity and sensitivity for each ECG

classes at  $k=2$  is shown in Table 3. The training stage yielded perfect results. However, during testing, the control groups attained positive predictivity of approximately 93.4% and sensitivity of approximately 93.3%.



**Fig.5.**Mean accuracies for classification of power ratio features from augmented limb leads

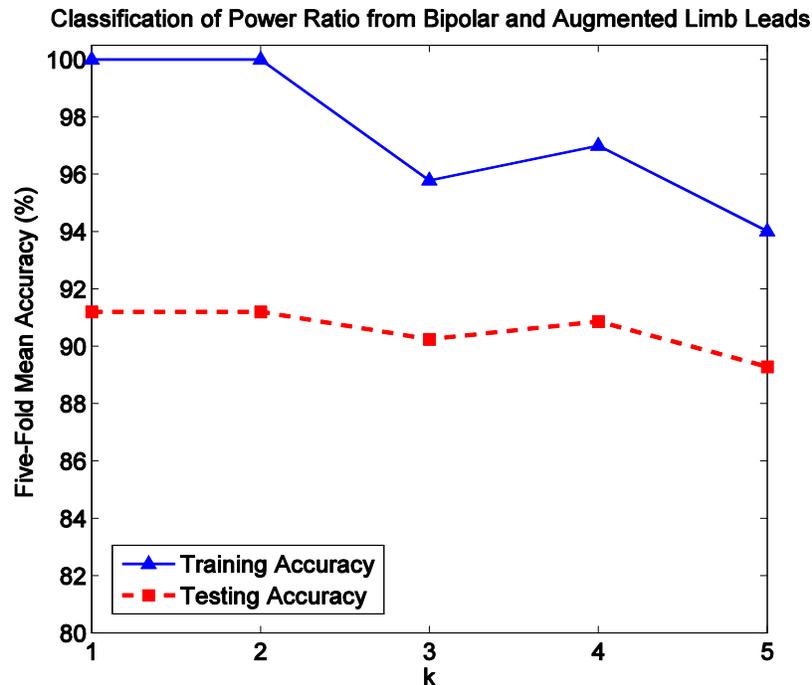
**Table 3.**Five-fold mean positive predictivity and sensitivity for classification of power ratio features from augmented limb leads at  $k=2$

ECG Classes	Training		Testing	
	Pp (%)	Se (%)	Pp (%)	Se (%)
Inferior MI	100	100	92.7	93.1
Anterior MI	100	100	94.3	93.3
Healthy Controls	100	100	93.1	93.4

### 3.6. Classification of Power Ratio: Combination of Bipolar and Augmented Limb Leads

An extended investigation is conducted to assess the effect of combining both lead systems for discriminating ECG from healthy controls and post-MI patients. The best classification performance are attained at  $k=2$  with 100% accuracy for training and 91.2% accuracy for testing. At  $k=5$ , training accuracy recorded a decline of 6.0% and testing accuracy at

1.9%. The overall performance is superior to the bipolar limb lead system, but is inferior to the augmented limb leads. Similar pattern in performance can also be observed in the deteriorating accuracies at  $k=5$ .



**Fig.6.** Mean accuracies for classification of power ratio features from bipolar and augmented limb leads

Meanwhile, positive predictivity and sensitivity for each ECG classes at  $k=2$  is shown in Table 4. The classifier attained 100% performance for the training stage. During testing however, the control groups attained positive predictivity and sensitivity of approximately 91.2%. The performance is also better than bipolar limb lead system, but is inferior to the augmented limb leads.

**Table 4.** Five-fold mean positive predictivity and sensitivity for classification of power ratio features from bipolar and augmented limb leads at  $k=2$

ECG Classes	Training		Testing	
	Pp (%)	Se (%)	Pp (%)	Se (%)
Inferior MI	100	100	90.7	90.9
Anterior MI	100	100	91.1	90.5
Healthy Controls	100	100	91.9	92.2

#### 4. CONCLUSION

Based on limited number of subjects, the study has successfully characterized the ECG of post-MI patients and healthy controls using the power ratio features. Distinct pattern between inferior and anterior MI is also observed for both bipolar and augmented limb leads. Feature classification via kNN has revealed superior discriminative capabilities of augmented limb lead system compared to the bipolar limb leads. The findings provide useful information that could be utilized in the development of automated systems that classify post-MI survivors from the healthy subjects.

#### 5. ACKNOWLEDGEMENTS

Authors extend their appreciation to the Microwave Research Institute, Faculty of Electrical Engineering and Institute of Research Management and Innovation, Universiti Teknologi MARA, Malaysia for supporting this work through the LESTARI research grant scheme (600-IRMI/MyRA 5/3/LESTARI (015/2017)).

#### 6. REFERENCES

- [1] Webster J. G. Medical instrumentation: Application and design. New Jersey: Wiley, 2009
- [2] Megat AMS A, Jahidin AH, Norali AN. Hybrid multilayered perceptron network for classification of bundle branch blocks. In International Conference on Biomedical Engineering, 2012, pp. 149-154
- [3] Megat A M S A, Jahidin A H, Norali A N, Mat S M H. Classification of bundle branch blocks using multilayered perceptron network. In IEEE International Conference on Control System, Computing and Engineering, 2011, pp. 531-535
- [4] Megat A M S A, Che Z C Z A, Husman A, Saaid M F, Noor M Z H, Jahidin A H. Detection of cardiomyopathy using multilayered perceptron network. In IEEE 8th International Colloquium on Signal Processing and its Applications, 2012, pp. 436-440
- [5] Megat A M S A, Rani M F, Jahidin A H, Saaid M F, Noor M Z H. Identification of cardiomyopathy disease using hybrid multilayered perceptron network. In IEEE International Conference on Control System, Computing and Engineering, 2012, pp. 23-27
- [6] Ahmad S M H, Megat A M S A, Noor M Z H, Jahidin A H, Saaid M F, Zolkapli M.

Investigation on Elman neural network for detection of cardiomyopathy. In IEEE Control and System Graduate Research Colloquium, 2012, pp. 328-332

[7] Megat A M S A, Shaari N F, Julai N, Jahidin A H, Amiruddin A I, Noor M Z H, Saaid M F. Robust arrhythmia classifier using hybrid multilayered perceptron network. In IEEE 9th International Colloquium on Signal Processing and its Applications, 2013, pp. 304-309

[8] Inan O T, Giovangrandi L, Kovacs G T A. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Transactions on Biomedical Engineering, 2016, 53(12):2507-2515

[9] Tunstall-Pedoe H, Kuulasmaa K, Amouyel P, Arveiler D, Rajakangas A M, Pajak A. Myocardial infarction and coronary deaths in the World Health Organization MONICA Project. Circulation, 1994, 90(1):583-612

[10] Mohamad F N, Megat A M S A, Jahidin A H, Saaid M F, Noor M Z H. Principal component analysis and arrhythmia recognition using Elman neural network. In IEEE 4th Control and System Graduate Research Colloquium, 2013, pp. 141-146

[11] Amiruddin A I, Megat A M S A, Saaid M F, Jahidin A H, Saaid M F, Noor M Z H. Feature reduction and arrhythmia classification via hybrid multilayered perceptron network. In IEEE 3rd International Conference on System Engineering and Technology, 2013, pp. 290-294

[12] Lin C H. Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier. Computer and Mathematics with Applications, 2008, 55(4):680-690

[13] Jahidin A H, Taib M N, Md Tahir N, Megat Ali M S A, Lias S, Fuad N, Omar WRW. Brainwave sub-band power ratio characteristics in intelligence assessment. In IEEE Control and System Graduate Research Colloquium, 2012, pp. 318-321

[14] Megat A M S A, Jahidin A H, Taib M N, Md Tahir N. EEG sub-band spectral centroid frequency and amplitude ratio features: A comparative study in learning style classification. Jurnal Teknologi, 2016, 78(2):15-23

[15] Agarwal C. C. Instance-based learning: A survey. In C. C. Agarwal (Ed.), Data classification: Algorithm and applications. New York: CRC Press, 2014, pp. 157-185

[16] Melgani F, Bazi Y. Classification of electrocardiogram signals with support vector machines and particle swarm optimization. IEEE Transactions on Information Technology in

Biomedicine, 2008, 12(5):667-677

[17] Wettschereck D, Aha D W, Mohri T. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms. *Artificial Intelligence Review*, 1997, 11(1-5):273-314

[18] Wang J, Duntsch I, Gediga G, Guo G. Nearest neighbours without k. In *Monitoring, security, and rescue techniques in multiagent systems*. In B. Dunin-Keplicz, A. Jankowski, A. Skowron, & M. Szczuka (Eds.), Heidelberg: Springer Berlin Heidelberg, 2005, pp. 179-189

[19] Goldberger A L, Amaral L A N, Glass L, Hausdorff J M, Ovanov P C H, Mark R G, Mietus J E, Moody G B, Peng C K, Stanley H E. Physiobank, PhysioToolkit, and PhysioNet: Components of a new resource for complex physiological signals. 2000, <https://www.physionet.org/physiobank/database/ptbdb/>

[20] Proakis J G, Manolakis D G. *Digital signal processing: Principles, algorithms and applications*. New Jersey: Prentice Hall, 2006

**How to cite this article:**

Ali M S A M, Jahidin A H, Yassin I M, Pasya I, Khalid M F A, Awang Z. Characterization of post-mi electrocardiogram using power ratio features and k-nearest neighbor classifier. *J. Fundam. Appl. Sci.*, 2017, 9(4S), 937-951.