Implementation of Kernel Sparse Representation Classifier for ECG Biometric System

Haryati Jaafar, Nurhidayah Ramli and Aimi Salihah Abdul Nasir
Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus,
Sg. Chuchuh, 02100, Padang Besar, Perlis, Malaysia.
haryati@unimap.edu.my

***Abstract*—In this paper, a biometric human recognition system based on Electrocardiography (ECG) signal is proposed. Three processes i.e., pre-processing, feature extraction and classification is discussed. A combination of enhanced start and end point detection namely short time energy (STE) and short time average zero crossing rate (STAZCR) is employed in the pre-processing. Subsequently, an autocorrelation method is applied in feature extraction. For the classification process, the kernel sparse representation classifier (KSRC) is proposed as a classifier to increase the system performance in high dimensional feature space. 79 recorded signals from 79 subjects are used are employed in this study. To validate the performance of the KSRC, several classifiers, i.e. sparse representation classifier (SRC), k nearest neighbor (kNN) and support vector machine (SVM) are compared. An experiment based on different sizes of feature dimensions is conducted. The classification performance for four classifiers are found to be 90.93%, 92.8%, 94.24%, 62.9%, 97.23% and 95.87% for the kNN, SVM (Polynomial and RBF), SRC and KSRC (Polynomial and RBF), respectively. The results reveal that the KSRC is a promising classifier for the ECG biometric system compared to the existing reference classifiers.**

***Index Terms*—Autocorrelation Method; ECG Signal; KSRC; SRC; STE and STAZCR.**

I. INTRODUCTION

Electrocardiography is a transthoracic interpretation of electrical activity of the heart over a period of time as sensed by electrodes attached to the surface of the skin and recorded by an external device which attached on the body. The recording formed by this noninvasive procedure is termed an electrocardiogram (ECG), which used to measure the heart's electrical conduction system. In recent times, some studies show that the use internal feature i.e. heartbeat signal which is known as electrocardiogram (ECG) signal has been documented to be suitable for biometric human recognition [1]. The validity of using ECG for biometric recognition is supported by the fact that the physiological and pathological of the heart in different individuals display certain uniqueness in their ECG signals [2].

In previous study, several methods in pre-processing and feature extraction have been suggested by researchers to prove the reliability and the robustness of ECG biometric for person recognition. However, less attention has been paid in the literature to their use for classification based on ECG signal. Therefore, this paper explores an approach that is different from the majority of the existing methods where a sparse representation classifier (SRC) is used as classifier for person identification system based on heartbeat signal. This classifier is a non-parametric learning method and it can directly assign a class label to a test sample without the training process [3]. Generally, sparse representations take account of most or all information of a signal with a linear combination of a small number of elementary signals called atoms. Often, the atoms are chosen from a so-called over-complete dictionary. The advantage of a scale-embedded dictionary is that it reduces the need to run the detector across various scales and the computation time for object detection [4]. Recently, the SRC is reported to outperform music genre recognition, phone recognition and speaker identification [5, 6]. The reason for this classifier increasingly becoming recognized in audio signal classification is because many signals are either naturally sparse, or they can be made sparse in some specific domain by using some predefined transforms such as the discrete Fourier transform (DFT) or the discrete cosine transform (DCT). This inherent or manufactured sparsity of audio signals will lead potentially to a lower computational complexity and less demand on resources [7]. Hence, this paper proposes the feasibility of using the sparse representation classifier (SRC) to identify person based on ECG signal. Nevertheless, the SRC is unable to classify a test sample successfully if the training samples belong to many different classes, as reported in Yin et al. [3]. Therefore, the kernel sparse representation classifier (KSRC) is proposed in this paper. Originally, the idea of the kernel function was used to construct the nonlinear SVM where the samples were mapped into a high dimensional feature space by nonlinear mapping [8]. Hence, the inner product does not need to be evaluated in the feature space and this provides a way of addressing the curse of dimensionality [3].

In order to identify the person based on ECG signal, a template matching is compared by two data which are called the enrolment or training and recognition or testing data. Successful template matching recognizes an individual's identity and we represent it as score as shown in Figure 1.

The overall architecture design of heartbeat biometric system is shown in Figure 2 [9]. The system consists of three important procedures, i.e. pre-processing or syllable segmentation, feature extraction and classification. . In the signal segmentation, the Short Time Energy (STE) and Short Time Average Zero Crossing Rate (STAZCR) are employed in this process. Consequently, in the feature extraction process, an autocorrelation method is used to find out similarity or relationship features. Finally, the KSRC is used as a classifier in the pattern matching process.

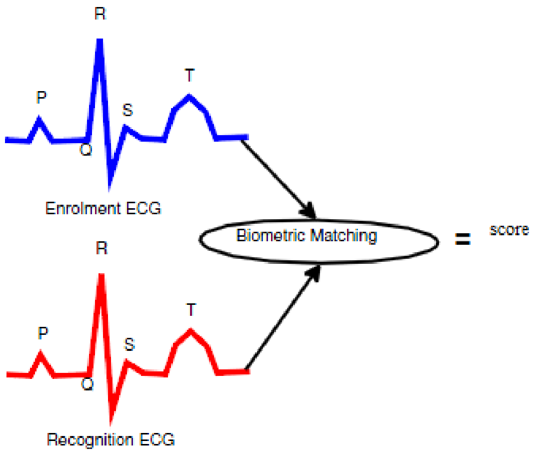


Figure 1: A typical ECG signal

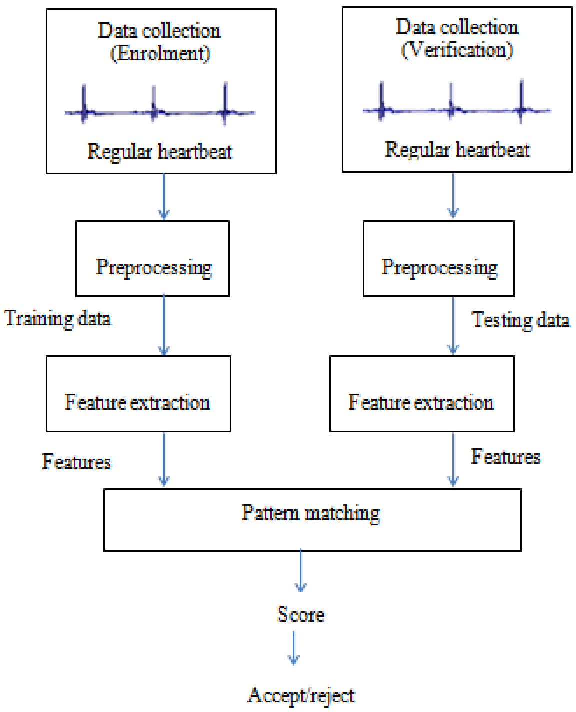


Figure 2: An architecture design of heartbeat biometric system

II. METHODOLOGY

A. Data Acquisition

The ECG database is obtained freely from public heart sound database, assembled for an international competition, the PhysioNet/Computing in Cardiology (CinC) Challenge 2016. The archive comprises nine different heart sound databases sourced from multiple research groups around the world. In this study, the Shiraz University adult heart sounds database (SUAHSDB) was used where this database was constructed using recordings made from 79 healthy subjects and 33 patients (total 69 female and 43 male, aged from 16 to 88 years). During the recording, the subjects were asked to relax and breathe normally during the recording session. The database consists of 114 recordings (81 normal recordings and 33 pathological recordings). The recording length varied from approximately 30s–60s. The sampling rate was 8000 Hz with 16-bit quantization except for three recordings at 44 100 Hz and one at 384000 Hz. The data were recorded in the wideband mode of the digital stethoscope, with a frequency response of 20 Hz–1 kHz [10].

B. Signal Segmentation

Two other techniques based on time-domain which are the Short Time Energy (STE) and Short Time Average Zero Crossing Rate (STAZCR) have been applied in this study [11]. The STE is the energy of a short desired signal segment. It is used to estimate the initial signal in the detection of desired and undesired signal segments. In the meantime, the STAZCR indicates the presence or absence of sound in the input signal [11]. If the value of STAZCR is high, the frame is considered to be undesired signal and if it is low, the frame is considered to be desired signal frame.

The STE function is defined by the following expression:

$$E_m = \frac{1}{N}\sum_{k=1}^{N}\left[x(k)w(m-k)\right]^2 \qquad (1)$$

where: E_m = Function which measures the change of voice signal amplitude
$x(m)$ = Input signal in one frame
m = Temporal length of each frame
$w(m-k)$ = Operator that represents a frequency shifted window sequence

On the other hand, the STAZCR is defined as:

$$Z_m = \frac{1}{2N}\sum_{k=1}^{N}|\operatorname{sgn} x(k) - \operatorname{sgn}[x(k-1)]|\, w(m-k) \qquad (2)$$

where: Z_m = Function which defines the zero crossing count.

C. Feature Extraction

Once the ECG signals have been segmented, the signals which consist of regular and irregular heartbeat segments will be divided into training and testing sets. A training set is used for parameter estimation and it is implemented to build up a model, while a test (or validation) set is to validate the model built [12]. Since heartbeat signals contain the useful feature, redundant features and leftover noises. It is important to pick only features that are unique, significant and least corrupted noise. For this propose, an autocorrelation method is used to find out similarity or relationship features among records of the same subject.

$$R_f(x) = \int_{-\infty}^{\infty} f(t)\, f^*(t-x) \qquad (3)$$

With regard to the series which successive observations are correlated, the first-order autocorrelation, the lag is one time unit. It is merely the correlation coefficient of the first N-1 observations, X_t, t=1,2...N-1 and the next N-1observations X_{t+1}, t=1,2...N-1. The correlation between X_t and X_{t+1} is given by:

$$r_1 = \frac{\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(1)}}\right)\left(x_{t+1} - \overline{x_{(2)}}\right)}{\left[\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(1)}}\right)^2\right]^{1/2}\left[\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(2)}}\right)^2\right]^{1/2}} \quad (4)$$

where: $\overline{x_{(1)}}$ = Mean of the first N-1 observations
$\overline{x_{(2)}}$ = Mean of the last N-1 observations

Since the autocorrelation space is a high dimensional space, an algorithm such as principle component analysis (PCA) is applied to the autocorrelation coefficients for dimensionality reduction [12].

III. ARCHITECTURE OF KSRC

SRC is a non-parametric learning process and this method directly assigns a class label to the test sample without the training model. Given a set of training and test samples, the basic idea of the SRC is to compute the sparse representation of the test sample on the training data. Then, the test sample is assigned to the class that minimizes the residual between itself and is reconstructed by the sparse representation that is associated with the training samples of each class [13]. In brief, the SRC can be formulated to solve the following optimization problem [14]:

$$\min \| x \|_1 \ \textit{subject to} \| y - Ax \|_2 \le \varepsilon \quad (5)$$

where: $\| x \|_1$ = l_1-norm
A = Matrix of training sample
y = Matrix of test sample

The test sample can be classified by minimizing the residual and yield:

$$\min r_k(y) = \| y - A\partial_k(x) \|_2 \quad (6)$$

where: ∂_k = Characteristic function to select the coefficient of the sample belonging to class k

Theoretically, the step finding sparse representation in the SRC is fast and behaves well in pattern recognition. However, it requires more effort to apply the SRC, particularly in multiclass data, since the data in the same direction would overlap each other after the normalization process. To overcome the problem, the KSRC is proposed where the kernel method is used to map samples into a high dimensional feature space; hence it gives better accuracy in classification [3].

Let the test sample $x \in R^m$ and ϕ be the nonlinear mapping functions corresponding to a kernel function. Suppose the samples are mapped from original feature space R^m into high dimensional feature space, F, by a non-linear mapping ϕ as below:

$$x \to \phi(x) \quad (7)$$

Let $D = [\phi_1(x), \phi_2(x), ..., \phi_k(x)]^T$ represent the matrix for the training sample after the mapping ϕ. Since in the SRC, the test sample can linearly be represented by the training sample, the test sample in the KSRC can similarly be represented by the training sample and is given as:

$$\min \| \vartheta \|_1 \ \textit{subject to} \ \phi(y) = \vartheta D \quad (8)$$

where: ϑ = Coefficient corresponding to the training samples
$\phi(y)$ = Test sample in the high dimensional feature space, which corresponds to y in the original feature space

However, it is not practical to directly solve the optimization problem in Equation (8). This is because the dimensionality of the feature space, F is far greater than the original feature space R^m and required a huge computation. This cause the system may slow down terribly or run out memory. Moreover, it has been observed that a large number of features may actually degrade the performance of classifiers if the number of training samples is small relative to the number of features [15]. For this reason, it is necessary to reduce the dimensionality in feature space F into low-dimensional subspace and this is given as:

$$P^T \phi(y) = P^T D\vartheta \quad (9)$$

Consider the representation of the transformation matrix. P in the kernel-based dimensionality reduction method, the transformation matrix can be expressed as:

$$P = DB \quad (10)$$

where: B = Pseudo-transformation matrix

Substituting Equation (9) into (10) yields:

$$(DB)^T \phi(y) = (DB)^T D\vartheta \quad (11)$$

where: $(DB)^T\phi(y) = k(x, y)$ = Kernel function defined as the inner product
$k(x, y)$ = $\phi(x)^T\phi(y)$
Z = D^TD
D = $\phi(x)$

Hence, Equation (11) can be obtained as:

$$B^T k(x, y) = B^T Z\vartheta \quad (12)$$

Substituting Equation (12) into Equation (8), the optimization problem is given as:

$$\min \| \vartheta \|_1 \ \textit{subject to} \ B^T k(x, y) = B^T Z\vartheta \quad (13)$$

By considering the noise data, Equation (11) can be modified as:

$$\min \| \vartheta \|_1 \ \textit{subject to} \| B^T k(x, y) - B^T Z\vartheta \|_2 \le \varepsilon \quad (14)$$

The test sample can be classified by assigning Equation (6) to the kth object class that minimizes the residual between itself and this yield:

$$\min | r_k(y) = \| B^T k(x,y) - B^T Z \vartheta_k \|_2 \leq \varepsilon \qquad (15)$$

The following algorithm summarizes the proposed recognition framework.

Algorithm: Kernel sparse representation classifier (KSC)

1. **Input:** a matrix of training sample $A_k \in R^{m \times n}$ and test sample $y \in R^m$
2. Normalize the column of A to have l_2-norm
3. Determine the kernel function $k(x,y)$
4. Solve the l_1 problem in $\min \| \vartheta \|_1 \text{ subject to } B^T k(x,y) = B^T Z \vartheta$ or $\min \| \vartheta \|_1 \text{ subject to} \| B^T k(x,y) - B^T Z \vartheta \|_2 \leq \varepsilon$
5. Compute the residual $\min | r_k(y) = \| B^T k(x,y) - B^T Z \vartheta_k \|_2 \leq \varepsilon$
6. **Output :** Identification of (y)

IV. EXPERIMENTAL RESULTS

In this paper, the principle component analysis with autocorrelation method is selected as a feature extraction due to the fact that the feature is more robust to noise compared to other feature extractions [12]. The proposed methods have been implemented in Matlab R2010(b) and have been tested in Intel Core i5, 2.1GHz CPU, 6G RAM and Windows 7 operating system. The database consists of the selected signal from 79 healthy subjects with 79 recording. In all experiments, the performance was evaluated based on the classification accuracy (C_A) which is calculated as;

$$C_A = \frac{N_C}{N_T} \times 100\% \qquad (16)$$

where: N_c = Number of syllables which is recognized correctly
N_T = Total number of test syllables

The comprehensive experiment was conducted to evaluate the effectiveness of the proposed classifier and compare it with other state-of-the-art classifiers such as the kNN, SVM, and KSRC after feature extraction. Two major experiments were conducted to evaluate the proposed methods.

In the first experiment, the optimal values of k for kNN and optimal kernel parameters for KSRC and SVM were firstly obtained. For KSRC and SVM, two popular kernels are obtained. One is the polynomial kernel, and the other is radial basis function (RBF) kernel $k(x,y) = \exp\|x-y\|^2 / \gamma$. Two aspects were compared in this experiment: (1) the performance of classifiers in different sizes of feature dimensions and (2) the performance of classifiers in different numbers of training samples.

These experiments were performed by ten-fold cross-validation by means of the CA rate. 15 training samples were randomly chosen from each ECG recording of ECG dataset, while the remaining samples formed the testing set. On each dataset, the experiments were repeated 10 times, and then nine different training and testing sets were attained for performance evaluation. The samples were extracted by MFCC and the dimension was fixed at 4,096 and were normalized to the unit norm. Here, in the kNN classifier, the values of k are presented in the odd numbers ranging from 1 to 15. The polynomial and RBF kernel were applied in the KSRC and SVM. To find the optimal kernel parameters for KSRC, the intervals were tested from 1 to 10 for both parameters d of the polynomial kernel and C of the RBF kernel. Meanwhile, various pairs of (C, γ) were tried for the SVM with the RBF kernel and the one with the best cross-validation accuracy was selected. Subsequently, the values of C and γ were used to determine parameter d in the polynomial kernel. Here, for the parameter d of the polynomial kernel, the candidate interval is from 1 to 10.

Figure 3(a) shows the CA rates of kNN in variation values of k. It was found the optimal CA rates were achieved when k=3. Meanwhile, the CA rates for SVM with the RBF kernel and polynomial kernel are shown in Figures 3(b) and 3(c), respectively. It was observed that the optimal parameter (C, γ) for RBF kernel was (2^5, 2^{-9}) and d for polynomial kernel was 3. Figures 3(d) and 3(e) show the CA rates for KSRC with the polynomial kernel and RBF kernel, respectively. From Figures 3(d) and 3(e), the optimal parameter γ was 2 and the optimal parameter d was 4.

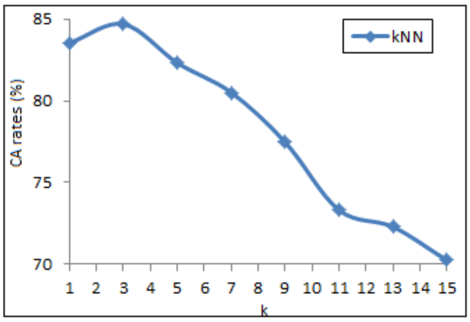
In the second experiment, a comparison between KSRC with kNN, SVM and SRC was made after determining the optimal value of k for kNN and parameters for KSRC and SVM. Here, the sizes of feature dimensions were computed at 100, 256, 1024, 4096, 6400, 7225 and 8100. Ten-fold cross-validation was applied and the experiment was repeated ten times.

Figure 4 shows the classification performance at different sizes of feature dimensions for each classifier, i.e. kNN, SVM, SRC and KSRC. Table 1 lists the maximal CA rates and the standard deviation of each classifier.

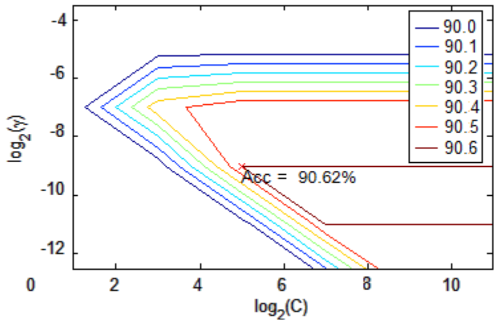
Table 1
Maximal CA Rates Based on Different Sizes of Feature Dimensions

Classifier	Parameter	CA rate (%)
kNN	k=3	90.93 ± 3.13
SVM (Polynomial)	d = 3	92.8 ± 3.78
SVM (RBF)	(C, γ) =(2^5, 2^{-9})	94.24 ± 3.07
SRC	-	62.9 ± 2.24
KSRC (Polynomial)	d = 4	97.23 ± 3.02
KSRC (RBF)	γ = 2	95.87 ± 3.26

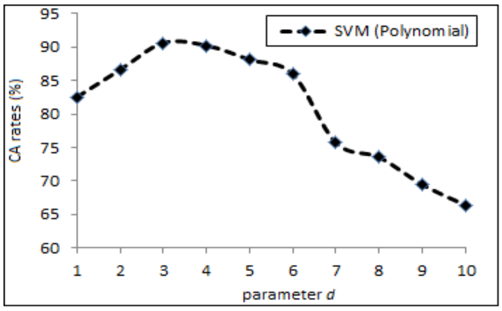
By examining Figure 4 and Table 1, some interesting points were found. First, they show the supremacy of KSRC over other classifiers. For added feature dimensions, the KSRC performs better as compared to other classifiers. The result also shows that the KSRC significantly performs better than the SVM and kNN for the 100 feature dimensions. However, most signals of practical interest have some noise in the feature that causes the trade-off in the SVM and kNN classifiers' increased computational complexity. The results also show that the KSRC outperforms the SRC for every feature dimension. It is clear to see that the KSRC can also obtain competitive classification results on high-dimensional data. Second, the kernel was able to improve the CA rates for both SVM and KSRC over the kNN and SRC. When compared with the polynomial kernel and RBF kernel, the KSRC (polynomial) performs better than KSRC (RBF). However, in the SVM, it it is inapplicable to the RBF kernel. Lastly, the SRC has the poorest performance which only achieves CA rates less 50% at 100 and 256 feature dimensions. This weakness indicates that the SRC is unable to handle many different classes compared than the KSRC as discussed in Section III.



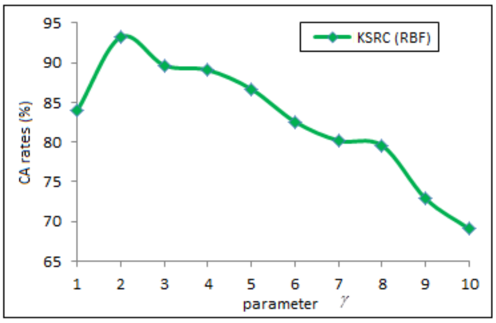
(a) kNN



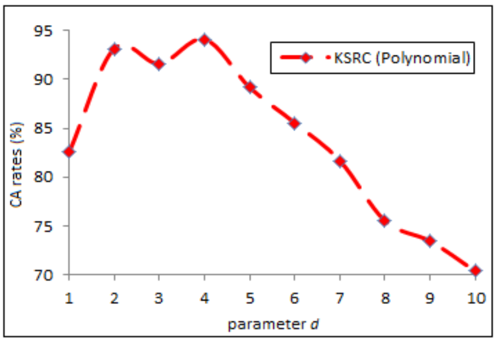
(b) SVM RBF kernel



(c)



(d) KSRC RBF kernel



(e) KSRC polynomial kernel

Figure 3: CA rates of classifiers under the variation of parameters

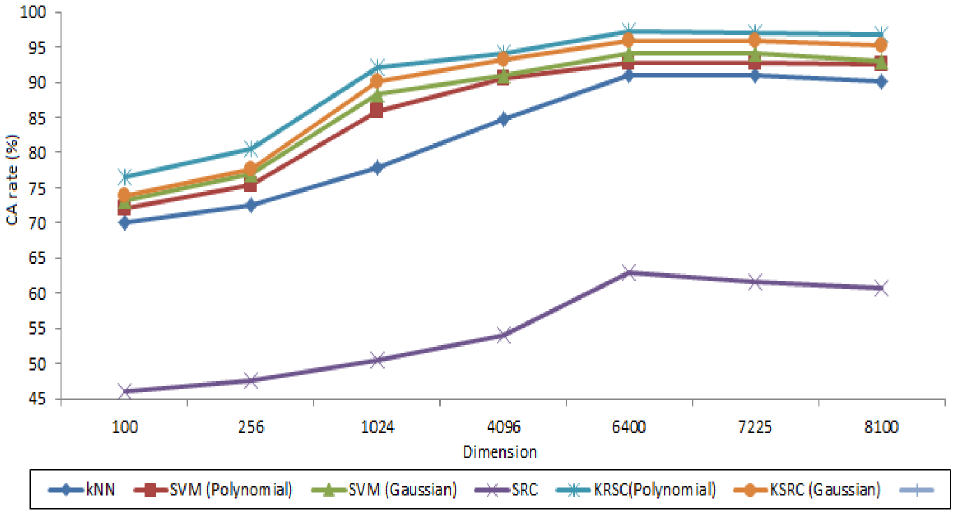


Figure 4: Performances of kNN, SVM, SRC and KSRC based on different sizes of feature dimensions

V. Conclusion

This paper presents a system which is able to perform biometrics recognition by using ECG signals. A combination of STE and STAZCR in signal segmentation has been employed. In the classification process, the KSRC has been proposed to map the samples into a high dimensional feature space. In this process, the segmented signals are first extracted with the autocorrelation method. Subsequently, three well-known classifiers, i.e. kNN, SVM and SRC, are used to validate the effectiveness of the KSRC. A series of experiments, based on different sizes of feature dimensions has been performed to determine the competence of the proposed classifier. For these experiments, the classification accuracies are up to 90.93%, 92.8%, 94.24%, 62.9%, 97.23% and 95.87% for kNN, SVM (polynomial), SVM(RBF), SRC and KSRC(polynomial) and KSRC(RBF), respectively. The results indicate that the KSRC with the polynomial kernel provides the best classification among the four classifiers.

Acknowledgment

The authors would like to express their gratitude for the financial support provided by Universiti Malaysia Perlis Short Term Grant (9001-00500).

References

[1] K. A. Sidek, V. Mai, and I. Khalil, "Data mining in mobile ECG based biometric identification," *Journal of Network and Computer Applications*, vol. 44, pp. 83-91, 2014.

[2] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing*, vol. 19, 2008.

[3] J. Yin, Z. Liu, Z. Jin, and W. Yang, "Kernel sparse representation based classification," *Neurocomputing*, vol. 77, pp. 120-128, 2011.

[4] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal of Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[5] B. L. Strum, and P. Noorzad, "On automatic genre recognition by sparse representation classification using auditory temporal modulations," 9*th* *International Symposium on Computer Music Modeling and Retrieval (CMMR 2012): Lecture Notes in Computer Sciences Series, pp.* 379-394, 2012.

[6] O. Vinyals, and L. Deng, "Are sparse representations rich enough for acoustic modeling?" *Proceedings of Interspeech 2012*, 2012.

[7] S. Zubair, F. Yan, and W. Wang, "Dictionary learning based sparse coefficients for audio classification with max and average pooling," *Digital Signal Processing*, vol. 23, pp. 960–970, 2013.

[8] S.W. Lin, Z.L. Lee, S.C. Chen, and T.Y. Tseng, "Parameter determination of support vector machine and feature selection using simulated annealing approach," *Applied Soft Computing*, vol. 8, pp. 1505–1512, 2008.

[9] Fratini, A., Sansone, M., Bifulco, P., & Cesarelli, M. (2015). Individual identification via electrocardiogram analysis. Biomedical engineering online, 14- 78.

[10] C. Liu, D. Springer, Q. Li, B. Moody, R.A. Juan, F.J. Chorro, F. Castells, J.M. Roig, I. Silva, A.E. Johnson, and Z. Syed, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, pp. 2181, 2016.

[11] D.A. Ramli, and H. Jaafar, "Peak Finding Algorithm to Improve Syllable Segmentation for Noisy Bioacoustic Sound Signal," *Procedia Computer Science*, vol. 96, pp.100-109, 2016.

[12] I., Guyon, S., Gunn, M., Nikravesh, and L. A. Zadeh, *Feature extraction: foundations and applications*, vol. 207, Springer Eds, 2008.

[13] W.L. Zuo, Z.Y. Wang, T. Liu, and H.L. Chen, "Effective detection of Parkinson's disease using an adaptive fuzzy k-nearest neighbor approach," *Biomedical Signal Processing and Control*, vol. 8, no. 4, pp. 364-373, 2013.

[14] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Analysis Machine Intelligent*, vol. 31, no. 2, pp. 210–227, 2009.

[15] L. Zhang, W.D. Zhou, P.C. Chang, and J. Liu, "Kernel sparse representation-based classifier," *IEEE Trans. On Signal Processing*, vol. 60, no. 4, pp. 1684-1695, 2012.