

# A NOVEL AUTOMATIC CARDIAC AUSCULTATION WITH HYBRID ANT COLONY OPTIMIZATION-SVM CLASSIFICATION

Prasertsak Charoen\*, Khamin Subpinyo, and Waree Kongprawechnon

*Received: November 23, 2015; Revised date: April 05, 2016; Accepted date: April 07, 2016*

## Abstract

Cardiac auscultation is a method to examine the condition of a heart using a stethoscope. Since the method requires expertise, which is rare in suburban rural areas, to analyze the condition of a heart, an automatic cardiac auscultation system is introduced with the use of a machine learning technique to eliminate the need of for expertise. In this paper, we develop a classification model based on a support vector machine (SVM) and ant colony optimization (ACO) for an automatic cardiac auscultation system. The proposed method uses ACO to further select heart sound features, resulting from the initial feature selection by principle principal component analysis (PCA), and uses them to train the SVM for classification. Experimental results show that the proposed ACO/SVM technique is able to give higher classification accuracy on heart sound samples, compared to with a traditional SVM and a genetic algorithm (GA) -based SVM.

**Keywords:** Support vector machine, Ccardiac auscultation, Aant colony optimization, Mmachine learning

## Introduction

Cardiac disease is one of the leading causes of death in the modern world. The rate is relatively high, considering that cardiac diseases can be prevented by regular heart check-ups to find the symptoms before it becomes fatal. People do not have Rregular health check-ups are not widely done by the people, especially in suburban rural areas where the availability of check-ups is low and considered to be expensive. The availability of cardiac auscultation is even lower since traditional cardiac auscultation requires a skilled

doctor to use a stethoscope to listen to heartthe sounds of patients' hearts directly and analyze their conditions. Though the diagnostics tools for cardiac auscultation are available, skilled doctors who can conduct cardiac auscultation are rare. Moreover, they are only available in major hospitals, making cardiac auscultation inaccessible in suburban rural areas. Normally when patients need to have a cardiac auscultation procedure, they need to go to big hospitals located in downtown areas because ECG systems for

---

Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, Thailand 12121,  
E-mail: prasertsak@siit.tu.ac.th

\* Corresponding author E-mail: prasertsak@siit.tu.ac.th

cardiac auscultation treatment are expensive. Since the tools for cardiac auscultation are expensive, the availability of tools is limited (Gupta *et al.*, 2007).

The main objective of this system is to create a low-cost cardiac auscultation system by using an automatic system in order to promote public health care. If cardiac auscultations are easy to access with low cost to access, then fatalities from heart diseases will decrease because the ease of access to cardiac auscultations provides early measurements. The easy accessibility of cardiac auscultations will provide more possibilities of treatment in the early stages of heart disease before it becomes fatal. The main problem that we discovered from the previous works in of Chatunapalak *et al.*, 2012 and Banpavivhit *et al.*, 2013 is the classification accuracy of the system. Moreover, the area that we consider a flaw is the accuracy of false negatives (FNs) where the system is designed to be used in dedicated situations which will affect lives. This is a vital part of the results since FNs classify diseased heart samples as healthy ones.

The proposed system consists of 4 main parts: Ppreprocessing, Ffeature Eextraction, Ffeature Sselection, and a Ssupport Vvector Mmachine (SVM) classifier. The preprocessing part resamples, reduces noise, and normalizes heart sounds. The feature extraction part is used to extract features from each heart sound by using a Wwavelet Ppacket-based feature extraction. The feature selection part is considered very important as it is used to select the salient features by using principal component analysis (PCA) and Ant colony optimization (ACO). In this system we use Ant colony optimization (ACO) to increase the accuracy of the classifier. Since each feature of each heart sound has a different significance in identifying a health conditions, ACO is implemented to select a significant feature set and to use this set of features to train a SVM classifier in the last part.

## Materials and Methods

### Classification by Support Vector Machine

Classifying data is a common task in

machine learning. A SVM has been known for its competent algorithm for analyzing data and recognizing patterns in the areas of statistics and computer science (Burgess, 1998; Vapnik, 1998). It is used for classification and regression analysis. Given certain data points, the largest separation on a hyperplane can make the SVM a non-probabilistic binary linear classifier, or give it, the perceptron of optimal stability. The classifier can also be optimized with the input data set for generalization.

Let us consider a supervised SVM classifier with a training dataset  $\{x_i, y_i\}_{i=1}^N$ , where  $x$  is the input sample and  $y \in \{+1, -1\}$  is the label of classes. An SVM classifier is constructed by using an optimal hyperplane with far enough distance to isolate two classes of the training dataset. This hyperplane is defined as  $w \cdot x - b = 0$ , where  $x$  is a testing data point on the hyperplane,  $w$  determines the orientation of the hyperplane, and  $b$  is the bias value with respect to the origin. A margin is used to form the decision surface and is represented by the distance  $d$  between two supporting hyperplanes as  $d = \frac{2}{\|w\|}$ . The larger the margin  $d$ , the better the classifier can separate the dataset. To find the optimum hyperplane, we must minimize  $\|w\|^2$  with the constraint  $y_i(w \cdot x_i - b) \geq 1, i = 1, 2, \dots, n$ . Thus, the optimization problem for finding the optimum hyperplane is given by

$$\begin{aligned} & \underset{w}{\text{minimize}} && \frac{1}{2} \|w\|^2 \\ & \text{subject to} && y_i(w \cdot x_i - b) \geq 1, i = 1, 2, \dots, n \end{aligned} \quad (1)$$

Furthermore, for a practical classification problem, the data sample may not always be linearly separable. The positive slack variable  $\xi_i$  is introduced for a nonlinear decision surface by substitution into the optimization problem. This allows an error term in the SVM classifier as soft-margin classification. The new optimization problem is shown, as in Equation (2):

$$\begin{aligned} & \underset{w}{\text{minimize}} && \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \\ & \text{subject to} && y_i(w \cdot x_i - b) \geq 1 - \xi_i, \\ & && i = 1, 2, \dots, n \\ & && \xi_i \geq 0 \end{aligned} \quad (2)$$

where  $C$  is a non-negative cost parameter which controls the tradeoff between maximizing the margin and minimizing the error. Finally, the classification decision function becomes:

$$f(x) = \text{sing} \left( \sum_{i=1}^N L_i y_i K(x_i, x_j) - b \right) \quad (3)$$

where  $L_i$  are Lagrange multipliers, and  $(x_i, x_j) = \varphi(x)$ . For a nonlinear classifier, a kernel function  $\varphi(x)$  is used to map the nonlinear data into a higher dimensional space. In this work, we select the radial basis function (RBF) kernel function for the SVM classifier, as suggested by Phatiwuttipat *et al.*, 2011. The RBF kernel is calculated by using  $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ , where  $x_i = [x_{i1}, x_{i2}, \dots, x_{id}]^T$  and  $x_j = [x_{j1}, x_{j2}, \dots, x_{jd}]^T$  are two 2 sample data sets and  $\gamma >$  is the width of the Gaussian.

### Principle Principal Component Analysis (PCA)

PCA is applied to attain the compactness of a feature set through a data-driven basis, so as to signify the most relevant HS heart sound features, according to Equation. (4):

$$A = QB. \quad (4)$$

where  $A$  is a dimensional-reduced version of the extracted feature vector  $B$ , and  $Q$  is a matrix whose eigenvectors, arranged in decreasing

order of magnitude of their eigenvalues, are made up from the sum of eigenvalues until being they are more than 90% of all summing eigenvalues. These eigenvectors are formed via the covariance matrix of the training set  $COV$ , defined by Equation. (5):

$$COV = \frac{1}{n-1} \sum_{i=1}^n (a_i - \bar{a})(a_i - \bar{a})^T \quad (5)$$

where  $n$  is the number of feature vectors,  $a_i$  is the original feature vectors, and  $\bar{a}$  is the mean vector. Best-tree coefficients of the wavelet packet (WP) feature set are then compactly supported as the algorithm iteratively traces subsets of eigenvectors from the covariance matrix of the WP-generated features to ensure the best related features are in the record. For the selection of the WP coefficients, the chosen set of 9 features is acquired for each HS heart sound from where original features are densely aligned to the signified level using MATLAB.

### Ant Colony Optimization (ACO)

ACO, which firstly aims to search for an optimal path in a graph, is based on the behavior of ants seeking a path between their colony and a source of food (Al-Ani, 2007; Aghdam, 2008). The original idea has since diversified to solve a wider class of numerical problems, and, as a result, several procedures have emerged, drawing on various aspects of the behaviors of ants. In nature, ants initially wander randomly, and, upon finding food, return to their colony while

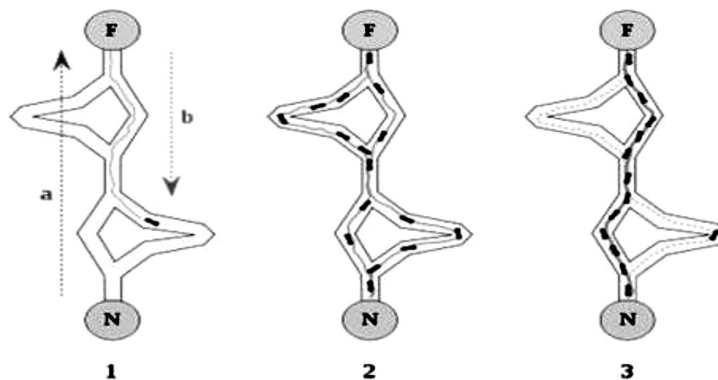


Figure 1. An example of ants laying pheromone

laying down pheromone trails. If other ants find such a path, they are likely that they will not to keep travelling at random, but to instead will follow the trail, returning and reinforcing it if they eventually find food. As time passes, however, the pheromone trail starts to evaporate, thus reducing its attractiveness. The more time it takes for an ant to travel down the path and back again, the higher the possibility of evaporation. A short route, in comparison, gets marched over more frequently, and thus the pheromone density becomes higher than on longer ones. Pheromone evaporation also has the advantage of avoiding the convergence to a locally optimal solution. If there were no evaporation at all, the paths chosen by the first ants would tend to be excessively attractive to the following ones. In that case, the exploration of the solution space would be constrained.

An example of the ACO process is shown in Figure 1. The first ant wanders randomly until it finds a food source (F). Then it returns to its nest (N), laying a pheromone trail. Afterwards, other ants follow one of the paths randomly, also laying pheromone trails. Since the ants traveling on the shortest path lays pheromone trails faster, this path gets reinforced with more pheromone, making it more appealing to the following ants. Lastly, the ants are likely to follow the shortest path more and more, since it is constantly reinforced with a larger amounts of pheromones.

The pheromone trails of longer paths evaporate.

### Proposed Methodology

The overall block diagram of the system is shown in Figure 2. It is divided into 4 main steps: Ppreprocessing, Ffeature Eextraction, Ffeature Sselection, and Cclassification. All of the processes and algorithms that are used in the system are done in MATLAB. Wavelet packet WP-based feature extraction and PCA are studied in Chatunapalak *et al.*, 2012, and this method is implemented in the proposed system. First, heart sounds go into the preprocessing stage, which is composed of resampling, noise cancellation, and normalization. After that, the heart sound features are extracted by WP, using WP decomposition, WP entropy, and best tree coefficients. Then, feature selection is done by PCA and ACO to select only the salient features. The final trained SVM can classify healthy heart sounds from abnormal heart sounds.

### Data Collection and Preprocessing

During preprocessing, the length of each heart cycle is equalized and resampled at a rate of 4 kHz. Then, the resampled heart sound enters the noise cancellation process; here a 5-level DWT discrete wavelet transform via soft-threshold is used along with the Daubechies-6 wavelet family for detailed coefficients (Chebil and Al-Nabulsi, 2007). The physiological variations

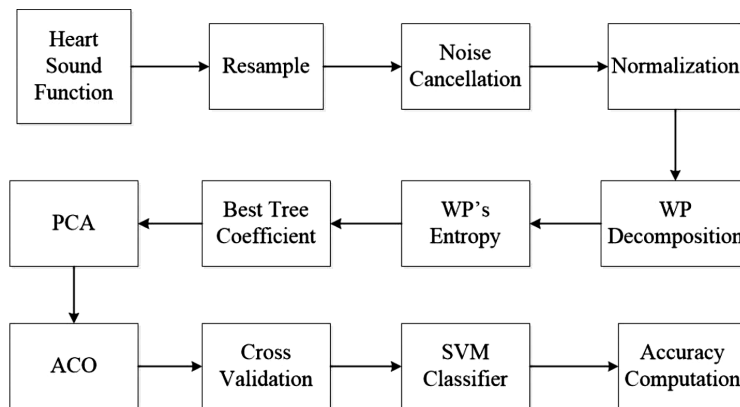


Figure 2. Overall block diagram of the proposed system

in intensity of each heart sound are discarded by normalization, using Equation. (6):

$$\hat{x} = \frac{x - \mu}{\sigma} \quad (6)$$

This has zero mean and unity variance where  $\hat{x}$  is a final heart sound signal, and  $x$  is a signal before normalization.  $\mu$  is the mean of the samples and  $\sigma$  is the variance of the samples (Bunluechokchai and Ussawongaraya, 2009).

### Feature Extraction

The feature extraction part of the proposed system is based on Chatunapalak *et al.*, 2012. Since heart sound is a non-stationary signal, in order to retain most of the time-frequency information in the signal, a WP-based filter is used because it has a high performance yield at characterizing these heart sound signals. The Hhigher-order Daubechies family wavelets are used to improve the frequency resolution and reduce aliasing. After WP decomposition, each all sub-band energy is assessed by employing the non-normalized Shannon's entropy criterion, as Equation (7):

$$E(t) = -|\hat{x}(t)| \cdot \log|\hat{x}(t)|. \quad (7)$$

where  $E(t)$  is Shannons's entropy and  $\hat{x}(t)$  is a heart sound signal. The logarithmic function

allows a greater entropy weight for signal intensity while attenuating noise and disturbances. The WP that is suitable for the heart sound signal is the 6-level decomposition on the Daubechies-3 mother wavelet function. The selected percentage of retained energy is 99.9%. This is to characterize the best-basis feature with a 96% compression rate, approaching hierarchically on the decomposition indices. The total outputs of the 2-channel filter banks are obtained, as Equation (8):

$$Total\ Features = \sum_{i=1}^n 2^i, \quad n = 6 \quad (8)$$

After the performance of WP, the final number of different extracted features on the frequency plane is 126.

### Feature Selection

PCA is the method commonly used to reduce the dimensions of the given features. The algorithm tries to locate a subset of the original extracted feature. The dimensions of these feature sets are important since they significantly shape the performance and accuracy of the system. PCA uses subsets of eigenvectors from the covariance matrix of the extracted feature. The feature that has the sum of 90% of eigenvalues (to the previous sum) is retained as related features, meaning that these features are significant in identifying each sample. The resulting number of features

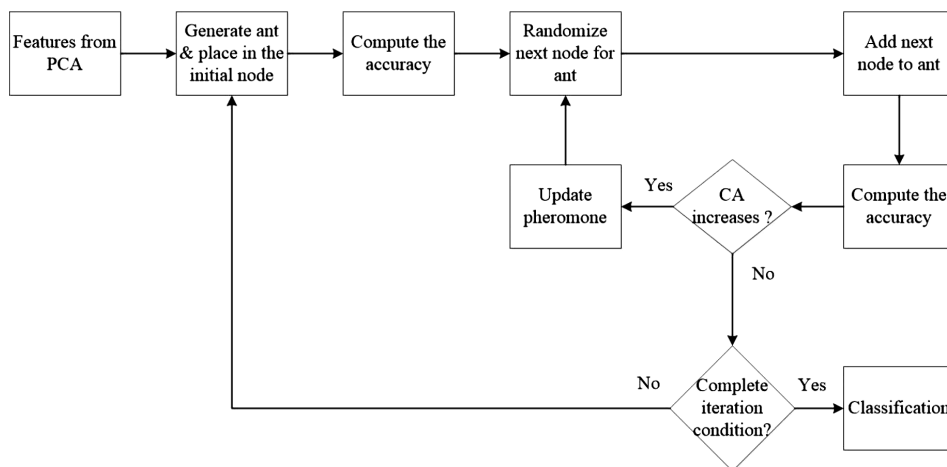


Figure 3. The process of the proposed ACO system

remaining in this particular system is 9, which is considerably smaller than the previous 126 features.

### Ant Colony Optimization Algorithm

ACO is an optimization technique based on the theory of ants finding food. The proposed system uses ACO for feature selection to choose the best set of salient features to train SVM classification (Kabir *et al.*, 2013; Majdi *et al.*, 2013). Since The 9 features selected from the heart sound samples are significant., since These factors determine the accuracy of the proposed system. Thus the selection of these features is very important. From the experiment, training the SVM using one 1 feature leads to significantly different classification accuracy of the SVM classifier. Therefore if the system selects only salient features, the accuracy of the classifier can be increased.

The ACO process diagram is shown in Figure 3. The method of ACO in the proposed system is based on a selection of the salient factors. First, the system generates ants and places them into initial nodes for all of the features. Classification accuracy (CA) is determined for all of the ants and pheromone levels and a level of 1 is assigned to each feature. Then, a random next feature for each ant is used and concatenated with the previous feature to compute the CA again. The feature associated with the ant that has the higher highest CA is given a pheromone level. Otherwise, the system terminates the ant and starts to generate the next ant. The process continues until it reaches the termination criteria. Note that the pheromone in this system is based on discrete values, with an increment of 1.

### SVM Classification

In order to classify all the heart sound samples into two 2 categories, healthy and unhealthy, the SVM method is chosen. Training of the SVM classification using the RBF kernel can be found in (Chatunapalak *et al.*, (2012). The classification accuracy is evaluated using a 10-fold cross validation method. The accuracy is then used to determine the fitness value of each feature set, provided by the ACO feature selection. The SVM classifier with the RBF-Kernel function finds the solution of the model by

mapping the nonlinear input data into a higher dimensional feature space. This is done to make the feature data separable.

After the ACO loop, the system arranges the features in descending order, corresponding to the pheromone level. A total of 9 cases of SVM are considered. The first case has 1 feature to train in the SVM and the second case has 2 features (in descending order), and so on until it reaches 9<sup>th</sup> the ninth case. As This is opposite to the work in of Banpavivhit *et al.*, (2013), where a genetic algorithm (GA) was used to find the optimal weight for each feature, and the proposed method used ACO to determine which feature is the most relevant to the system and then selected the number of features needed to maximize classification accuracy.

### Results and Discussion

The proposed method was tested with 304 individual labeled heart sounds in a digitized WAV file format with 16 bits resolution, and mono sound. The pathological heart sounds were comprised of Aortic Sstenosis, Aortic Rregurgitation, Mmitral Sstenosis, and Mmitral Rregurgitation. To evaluate the performance, ten 10-fold cross validation is used, where the data is divided into ten 10 portions as follows:

$$D = Q_1 \cup Q_2 \cup \dots \cup Q_9 \cup Q_{10} \text{ and } Q_i \cap Q_j = \emptyset; i \neq j \quad (9)$$

where  $Q_i$  is the  $i^{\text{th}}$  partition. Cross validation is carried out by taking different folds that are independent of one another. Each fold  $Q_i$  is utilized for testing exactly once, where the remaining folds of that partition  $L_i$  are used to train the models, as given in Equation 10:

$$L_i = D - Q_i \text{ where } i = 1, 2, \dots, 10 \quad (10)$$

The process continues via generating a new set of both portions until ten 10 sets of data are reached. Cross validation results can be evaluated by four 4 numbers; true positive (TP), false positive (FP), true negative (TN), and false negative: (FN):

- True Positive (TP): the proportion of actual positives that are correctly identified;



- False Positive (FP): the proportion of actual positives that are incorrectly identified;
- True Negative (TN): the proportion of negatives that are correctly identified;
- False Negative (FN): the proportion of negatives that are incorrectly identified.

In this experiment, positive means a diseased heart with disease and negative means a healthy heart. Performance rate can be calculated by using sensitivity, specificity, and accuracy in a confusion matrix. For sensitivity, it measures the percentage of the abnormal heart sound which is classified correctly. Specificity measures the percentage of the normal heart sounds which are classified correctly. Accuracy is the proximity of the measurement results to the true value.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (11)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (12)$$

$$\text{Accuracy} = \frac{\sum \text{Correctly classified HSS}}{\sum \text{HSS}} \quad (13)$$

From the simulation of the system with 50 iterations, Table 1 shows the total pheromone levels of the system. The highest pheromone

values are from feature number 8, which is the most important feature. It is arranged to be the first feature to be used in SVM training. The lowest pheromone values are from feature number 1, which is assigned to the lowest priority.

Table 2 shows the result of the SVM classification according to the number of features trained in the SVM. We simulate 9 SVM classifications to see how many feature(s) acquired the highest accuracy for SVM classification. The results indicated that the number of features trained in the SVM have an impact on the system's accuracy. An increase in features that are less relevant can decrease the accuracy. The highest accuracy is obtained when use only the 4 highest pheromone features are used.

Table 3 shows the results compared to with the previous works of Chatunapalak *et al.*, 2012 and Banpavivhit *et al.*, 2013. The proposed system with ACO has a better cultural algorithm (CA) at 1.32%, compared to with the system without the GA and ACO. The system with ACO improves the CA by 0.99%, when compared to with the system with GA. Improvement of the classification accuracy results from a reduction in the FN of the proposed method as the system

**Table 1. Pheromone levels of each feature**

Feature Number	1	2	3	4	5	6	7	8	9
Pheromone levels	8	16	18	19	13	13	17	25	19

**Table 2. Experimental results from the system**

Number of features for training SVM	FN	FP	TN	TP	Sensitivity	Specificity	Accuracy
1	57	0	144	103	64.38%	100%	81.25%
2	23	0	144	137	85.63%	100%	92.43%
3	20	0	144	140	87.50%	100%	93.42%
4	18	0	144	142	88.75%	100%	94.08%
5	22	0	144	138	86.25%	100%	92.76%
6	22	0	144	138	86.25%	100%	92.76%
7	22	0	144	138	86.25%	100%	92.76%
8	22	0	144	138	86.25%	100%	92.76%
9	22	0	144	138	86.25%	100%	92.76%

**Table 3. An experimental comparison of the proposed system and previous works**

Setting	FN	FP	TN	TP	Sensitivity	Specificity	Accuracy
Without ACO & GA	22	0	144	138	86.25%	100%	92.76%
With GA	21	0	144	139	86.88%	100%	93.09%
With ACO	18	0	144	142	88.75%	100%	94.08%

correctly classified patients with abnormal hearts. ACO emphasized on the important characteristic of the features in the heart sounds while it suppressed features that are irrelevant to the system. This led to the system making a better decision when classifying diseased heart sounds as abnormal hearts. It could be concluded that the method proposed by this study obtains promising results for an automatic cardiac auscultation classification system.

## Conclusions

This paper proposes a new alternative screening diagnostics tool that has low cost. The main objective of this system is to act as a cheap and efficient screening system so that patients with potential heart disease can be identified. The system uses a wavelet packet -based feature extraction to extract significant features from the heart sound samples to form 126 feature sets. The dimensions of the features are then further reduced using PCA. To obtain sets of salient features to be used to train an SVM classifier, ACO is used. ACO checks the importance of features by pheromone levels. These pheromone values arrange the features in descending order by the pheromones levels, to train in SVM. The highest CA result is obtained when choosing only 4 out of 9 features. The resulting classifier accuracy is improved, compared to with previous methods. Moreover, the proposed system can reduce the FN area which is an important factor that can affect the life of a patient.

In future research, feature extraction using Linear Discriminant Analysis (LDA) and Latent Semantic Analysis (LSA) will be investigated. We plan to expand combination of such techniques with SVM classification to achieve maximum classification accuracy. Also, the proposed system will be tested on raw heart sounds collected on site from patients using

a digital stethoscope. The project will be evaluated by a doctor from a hospital to verify and compare the records with the proposed method. A Graphic User Interface (GUI) can also be developed for real-time usage of the algorithm.

## References

- Aghdam, H.M., Ghasem-Adhaee, N., and Basiri, E.M. (2008). Text feature selection using ant colony optimization. *Expert Systems with Applications*, Vol. 36, Issue (3), Part 2, p. 6,843–6,853.
- Ahmed Al-Ani, A. (2007). Ant Colony Optimization for feature subset selection. *International Journal of Computer, Electrical, Automation, Control, Quantum and Information Engineering*, Vol.1, No.(4):999-1,002.
- Banpavivhit, S., Kongprawechnon, W., and Tungpimolrut, K. (2013). Cardiac Auscultation with Hybrid GA/SVM. *Proceedings of the 9<sup>th</sup> International Conference on Computing and Information Technology (IC2IT2013)*; May 9-10, 2013; Bangkok, Thailand, Vol. 209, p. 185-194.
- Bunluechokchai, C. and, Ussawongaraya, W. (2009). A wavelet-based factor for classification of heart sounds with mitral regurgitation. *International Journal of Applied Biomedical Engineering*, Vol.2, No.(1):44-48.
- Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, vol. 2, no. (2):121-167.
- Chatunapalak, I., Phatiwuttipat, P., Kongprawechnon, W., and Tungpimolrut, K. (2012). Childhood musical murmur classification with support vector machine technique. *Proceedings of the 9<sup>th</sup> International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology*; May 16-18, 2012; Petchaburi, Thailand, p. 1-4. Sirindhorn International Institute of Technology, Thammasat University, Pathum Thani, Thailand.
- Chebil, J., and Al-Nabulsi, J. (2007). Classification of heart sound signal using discrete wavelet analysis. *International Journal of Soft Computing*, 2(1), p. 37-41.
- Gupta, C.N., Palaniappan, R., Swaminathan, S., and Krishnan, S.M. (2007). Neural network classification of homomorphic segmented heart



- sounds. *Appl. Soft Comput.*, 7(1):286-297. Biomedical Engineering Research Center, Nanyang Technological University, Singapore.
- Kabir, M., Shahjahan, M., and Murase, K. (2013). Ant colony optimization toward feature selection. In: *Ant Colony Optimization - Techniques and Applications*. Barbosa, H., (Eed.). InTech, Rijeka, Croatia, DOI: 10.5772/51707.
- Majdi, M., and Eleyan, D. (2013). Ant colony optimization based feature selection in rough set theory. *International Journal of Computer Science and Electronics Engineering, (IJCSEE)*; Vol. 1, Issue (2):244-247.
- Murase, K., Monirul, K., Shahjahan, M. (2013). Ant Colony Optimization toward Feature Selection. *Ant Colony Optimization - Techniques and Applications*, Dr. Helio Barbosa (Ed.), InTech, DOI: 10.5772/51707.
- Phatiwuttipat, P., Kongprawechnon, W., and Tungpimolrut, K. (2011). Cardiac auscultation analysis system with neural network and SVM technique. *Proceedings of the 8<sup>th</sup> International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology Conference (ECTI-CON)*; May 17-19, 2011; Khonkaen, Thailand, p. 1,027-1,030.
- Vapnik, V.N. N. (1998). *Statistical Learning Theory*, New York: Wiley, Hoboken, NJ, USA, 768p.

