International Journal of Innovative
Computing, Information and Control
Volume **7**, Number **12**, December **2011**

ICIC International ©2011 ISSN 1349-4198
pp. 7131–7144

CLASSIFICATION OF HEART SOUNDS BASED ON THE LEAST SQUARES SUPPORT VECTOR MACHINE

Gür Emre Güraksin[1] and Harun Uğuz[2]

[1]Department of Computer Engineering
University of Afyon Kocatepe
03200 Afyonkarahisar, Turkey
emreguraksin@aku.edu.tr

[2]Department of Computer Engineering
University of Selçuk
42020 Çimenlik, Konya, Turkey
harun_uguz@selcuk.edu.tr

Received August 2010; revised December 2010

Abstract. *The heart is of crucial significance to human beings. Auscultation with a stethoscope is regarded as one of the pioneer methods used in the diagnosis of heart diseases. However, the fact that auscultation via a stethoscope depends on the skills of the physician's auscultation or his/her experience may lead to some problems in diagnosis. Therefore, the use of an artificial intelligence method in the diagnosis of heart sounds may help the physicians in a clinical environment. In this study, primarily, heart sound signals in numerical format were separated into sub-bands through discrete wavelet transform. Next, the entropy of each sub-band was calculated by using the Shannon entropy algorithm to reduce the dimensionality of the feature vectors with the help of the discrete wavelet transform. The reduced features of three types of heart sound signals were used as input patterns of the least square support vector machines and they were classified by least square support vector machines. In the method used, 96.6% of the classification performance was obtained. The classification performance of the method used was compared with the classification performance of previous studies which were applied to the same data set, and the superiority of the system used was demonstrated.*
Keywords: Least squares support vector machine, Discrete wavelet transform, Shannon entropy, Heart sounds

1. Introduction.

The heart is one of two organs which are crucial for human life. Therefore, a disorder of the heart is of great importance to human health. All over the world, in the period between 1985 and 2006, the mortality rate stemming from heart diseases ranked second after brain embolisms [1].

The heart is a hollow muscle which pumps the blood through all the body [2]. The primary and most important duty of the heart is that it delivers blood into the circulatory system [3]. This shows us how crucial it is for human life. The cycles of the heart are known as the systole and diastole. The stage occurring when the heart contracts is called the systole, and the stage occurring when the heart relaxes is called the diastole. The heart sounds heard as "lub, dub" occur because of the closing of the heart valves. "Lub" is known as the first heart sound (S1) and "dub" is known as the second heart sound (S2). The third heart sound (S3) occurs immediately after the S2, and it is of lower energy than the second one. The fourth heart sound (S4) occurs before the S1 and it has a lower scale of amplitude than the other sounds [2]. In addition, the sounds due to the flow of blood in the vessels and in the heart are components of the heart sounds. However, in fact, how they occur is still a subject of discussion [3].

Abnormal heart sounds such as murmurs are extra heart sounds heard among real heart sounds and they are caused by damaged valves [2]. The murmurs are the first signs of pathologic changes occurring in the heart valves, and they can be detected by the auscultation method in primary health organizations [6]. The common heart diseases result from heart valve disorders. Heart valve disorders are important heart diseases. Thus, in heart valve disorders, early diagnosis is one of the chief areas of study in the field of medicine [5].

In distinguishing normal and abnormal heart sounds, the auscultation method is one of the primary ones used by the physicians [4]. These sounds are listened to and interpreted by the doctors using stethoscopes, and they help to detect whether the patient has any heart disorder [10]. Dopplers, echocardiography and magnetic resonance imaging techniques are effective methods used nowadays in determining the anomalies of heart valves. With the use of these methods, the importance of auscultation and phonocardiography, both conventional methods, has started to decrease according to expert doctors. The auscultation method is a major and substantial diagnosis method for physicians but it is relatively more expensive, inaccessible and time-consuming than the other methods. However, in primary health organizations, auscultation still has an important place in determining whether patients need expert interference [6].

The heart sounds determined via the auscultation method by expert physicians are interpreted, and it is determined whether the patient has any heart disorder. However, there are also some disadvantages to the auscultation method. By using this method, the diagnosis of the disease depends on the skills and experiences of the physician in hearing and interpreting the different heart sounds [8]. These necessary experiences and skills have been acquired as a result of long examinations made by the physician. In addition, the fact that the environmental conditions are not suitable for and are incompatible with the patient can lead to a lack of diagnosis. Due to these difficulties, it is observed that the auscultation method has not been very successful in determining heart diseases [7]. The auscultation method may contribute to the removal of the problems stemming from the disadvantages we mentioned above with the support of artificial intelligence and may also make a significant contribution to helping physicians in primary health organizations.

There are a number of studies in the literature about the classification of heart diseases by means of artificial intelligence methods. In one of these studies, Leung et al. defined digitally recorded pathologic and non-pathologic phonocardiograms using the time-frequency method, and made a classification using the probability neural networks method. In conclusion, in the removal of the pathologic systolic blasts, they obtained 97.3% sensitivity and 94.4% specificity [18]. In another study carried out by Folland et al., in the course of auscultation, they applied Lewinson-Durbin algorithms and fast Fourier transform to the heart sounds in order to analyse abnormalities in the heart sounds, and they applied the data to Artificial Neural Networks (ANN) of Radial Basis Function (RBF) networks and Multi Layer Perceptron (MLP) in order to classify abnormal sounds. In conclusion, the sensitivity values obtained during the classification of the heart sounds of MLP and RBF neural networks were 84% and 88% respectively [20]. In another study, Chauhan et al. described an automatic heart sound classification system which uses a probabilistic approach based on Mel-Frequency Cepstral Coefficiency (MFCC) and Hidden Markov Models (HMM) [19]. In another study, Reed et al. developed a prototype system in order to analyse the sounds and their classification. They concentrated on the wavelet transform of the sounds and used a neural network based classifier and classified different heart sounds [21]. In a study by Güraksın et al., a heart sound classification system was developed by using the Discrete Fourier Transform (DFT) and ANN. In total, data comprising 120 normal, mitral stenosis and pulmonary stenosis heart sounds were used in this study, and 91.6% classification performance was obtained [9]. In another study carried out by using the same data set as Güraksın et al., Uğuz developed a biomedical system for the classification of heart sounds using DFT and Burg and Principal Component Analysis (PCA) with the ANN classifier, and obtained a classification performance of 95% [24].

In addition, when the studies carried out on the classification of heart sounds through the use of the Support Vector Machine (SVM) method in the literature are examined, we understand that Maglogiannis et al. developed an SVM classifier and classified a global data set of 198 heart sound signals, which come from both healthy medical cases and cases suffering from the four most usual heart valve diseases: aortic stenosis, aortic regurgitation, mitral stenosis and mitral regurgitations. Besides, the alternative classifiers they examined for comparison purposes backed propagation neural networks and the k-nearest neighbourhood and the naive Bayes classifiers exhibited much lower performance for the same diagnostic problems when using the same dataset as their SVM based classifier [22]. In another study carried out Çomak et al., by using the SVM and the Least-Squares Support Vector Machine (LS-SVM), Doppler signals belonging to heart valves were classified as normal and abnormal. The study produced a set of data comprising 215 heart sounds and it was emphasized that the LS-SVM method became more advantageous than the Back Propagation Artificial Neural Network method (BP-ANN) with regards to the training duration [5].

In this study, an artificial intelligence based system was formed for the classification of heart sounds. The study was carried on a total 120 heart sounds, and the sounds were examined in three groups: normal heart sounds, pulmonary stenosis heart sounds and mitral stenosis heart sounds. The study consisted of three stages. In the first stage, by using wavelet transform, feature extraction was realized. In this stage, the heart sound signals obtained were separated into sub-bands by using DWT. The feature extraction stage is one of the most significant stages affecting the performance of the classification. The results of the classification will be very successful, if the features are chosen well [23]. At the same time, with the classifier, a large number of input parameters will increase computationally and be intensive and time consuming. Therefore, the aim was to choose fewer features to represent the data set rather than a large number of features obtained with DWT. In this way, in the second stage, which is the dimension reduction stage, the entropy of each sub-band was calculated by using the Shannon entropy algorithm to reduce the dimensionality of the feature vectors via DWT [24]. In the third stage, the stage of classification, by using the LS-SVM method, the sounds were classified into three groups: mitral stenosis, pulmonary stenosis and normal sounds. SVM is a method developed for classification and regression, and it is a powerful alternative to ANN. Compared with ANN, the generalization ability of SVM is better. The LS-SVM method was used in this study and an optimal separating hyperplane was found through using a number of linear equations. In the standard SVM, however, an optimal separating hyperplane was found by solving the quadratic optimization problem [5]. In the results for the classification of heart sounds via LS-SVM, the classification success was estimated and the classification performance of the method used was compared with the classification performance of the previous studies [9,24] which were applied to the same data set.

2. **Materials and Methods.**

The flow diagram of the method used is seen in Figure 1. In this study, at first, the sounds obtained were digitized via software developed in a visual studio environment. Later, DWT, one of the signal processing methods, was applied to the digitized heart sounds, and the feature extraction process was realized. After the feature extraction process, dimension reduction was materialized using the Shannon entropy algorithm. Finally, the feature vectors were classified with the use of the LS-SVM method.

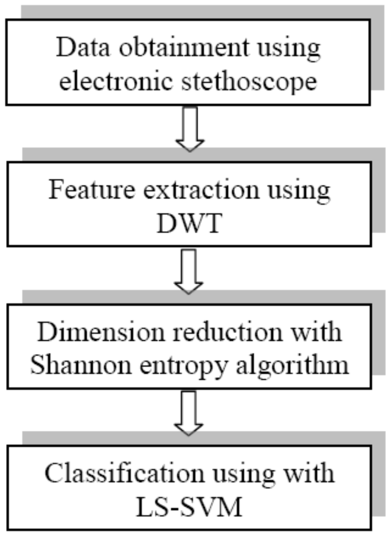

FIGURE 1. Flow chart of the method used

2.1. **Raw data obtainment.** The sounds used in this study were recorded using the Littmann 4100 model electronic stethoscope from Afyon Kocatepe University with the 07.AFMYO.01 numbered scientific research project [10]. It is possible to record six different sounds by means of the model 4100 Littmann electronic stethoscope. Thanks to this, heart sounds recorded in six consecutive patients can be stored in the stethoscope's own body. Heart sounds taken by the Littmann 4100 stethoscope have been recorded under the format of "*e*4*k*". Thus, these sounds were studied by making them into a "wav" formation by means of a program provided by Littmann.

2.2. **Wavelet transform.** Wavelet transform is a tool used for analysing unstable signals in the time frequency domain [11]. This is based on the analysis of base functions formed by shifts and the scaling of the sample functions of a signal $\Psi(t)$. The base functions are that they change fast with high frequency, and they change slowly with low frequency functions. Wavelets were used, for the first time, in Haar's thesis in 1909 [12]. Wavelet transform is a powerful alternative method to the Fourier method and is used in a number of areas such as biology, telecommunications and medicine.

Standard wavelet decomposition is expressed by the formula:

$$d_m(t_m) = x(t)\psi_m\left(\frac{t - t_m}{2^m}\right) \tag{1}$$

where m shows the level, t_m shows a section of the time period and Ψ_m represents the decomposition filter for the m level. The time-frequency coefficients are given as follows:

$$d[n] = x[n]h[n], \quad c[n] = x[n]g[n] \tag{2}$$

where $h(n)$ represents the impulse for the high pass filter and $g[n]$ represents the impulse for the low pass filter [5].

Wavelet transform can be examined in two groups: Continuous Wavelet Transform (CWT) and DWT. The CWT of a signal is the integral of the signal multiplied by scaled and shifted versions of a wavelet function Ψ, and it can be expressed as follows:

$$CWT(a.b) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{|a|}} \psi \left(\frac{t-b}{a} \right) dt \tag{3}$$

where Ψ represents the mother wavelet, a and b represent scaling and shifting parameters in the wavelet transform, respectively. However, calculating wavelet coefficients for every possible scale also fetches unnecessary information from the signal. Thus, it causes a longer transform process, which leads to a longer analysis process [13]. If scales and shifts are selected based on powers of two, which are called dynamic scales and positions, then it will be much more efficient during the wavelet analysis [15]. This means that fewer coefficients are being obtained in terms of quantity; however, these coefficients give the signal's frequency-scale variation over time. This analysis is called DWT and is defined as follows:

$$DWT(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{|a|}} \psi \left(\frac{t-2^j k}{2^j} \right) dt \tag{4}$$

where a and b are replaced by 2^j and $2^j k$, respectively. An effective way of implementing this schema by using filters was developed by Mallat [16]. DWT separates a signal into two types — Approximation (A) and Detail (D) — by using an Orthogonal Quadratic Mirror Filter (QMF) [14]. DWT enables the division of the whole time-frequency table into distinct time frequency particulars instead of separating low frequency components [17].

Two-levelled wavelet decomposition belonging to a signal is shown in Figure 2, where D1 represents the level-1 detail coefficient and D2 represents the level-2 detail coefficients.

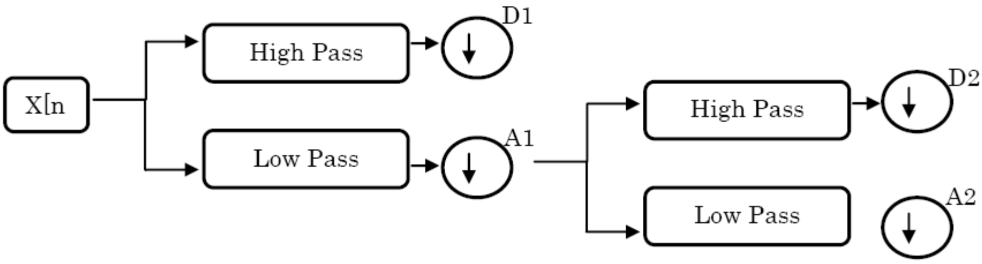

Figure 2. A signal's two-level wavelet decomposition

2.3. **Shannon entropy.** Entropy is a quantity determining the amount of irregularities in a thermodynamic system. Entropy measurement is an ideal method for measuring the level of disorder of a non-stationary signal [25]. An entropy-based criterion describes information-related properties for an accurate representation of a given signal [23]. From the information theory viewpoint, the concept of entropy is generalized as the amount of information stored in a more general probability distribution. Shannon is the first entropy concept to be applied to the science of information theory [26]. The Shannon entropy is defined using the following equation:

$$E(s) = -\sum_i s_i^2 \log(s_i^2) \tag{5}$$

where s represents the signal and s_i represents the coefficients of s on an orthonormal basis. In this study, the features obtained from DWT were exposed to the Shannon entropy and the dimension reduction process was realized.

2.4. **Support vector machines.** SVM was first discovered by Vapnik in 1979. Later, in 1995, it was recommended by Vapnik for classification and regression [5]. It is based on theoretical learning theory [27]. The main concept of the SVM is based on the formation of a Lagrange multiplier equation, combining both objective terms and constraints [31]. The aim in SVM is to find the optimum separating hyperplane which is able to classify data points as well as possible and again to separate them into two classification points as much as possible; that is, it aims to find the state in which the distance between two groups is the maximum. The hallmarks of this classification reasoning are the support vectors chosen from the training set and they are located on the end points of both groups [27]. An example of support vectors and the maximum margin hyperplane is shown in Figure 3.

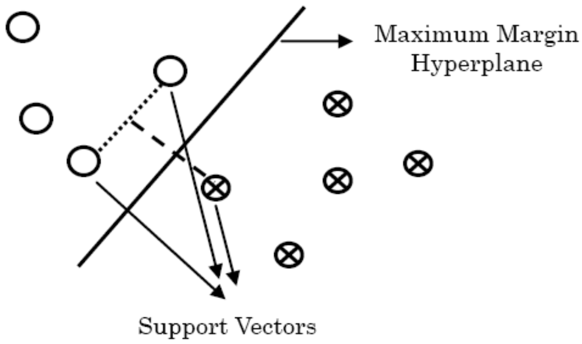

FIGURE 3. Maximum margin hyperplane

SVM maps the input patterns into a higher dimensional feature space through some non-linear mapping chosen a priori. A linear decision surface is then formed in this high dimensional feature space. Thus, SVM is a linear classifier in the parameter space, but it becomes a nonlinear classifier as a result of nonlinear mapping of the space for the input patterns into the high dimensional feature space [28]. SVM has been receiving an increasing interest in areas ranging from its original application in pattern recognition to regression estimation due to its good generalization performance, absence of local minima and sparse representation of solutions [33].

LS-SVM was developed a very short time ago [29]. LS-SVM involves equality instead of inequality constraints and works with a least squares cost function [32]. LS-SVM involves the equality constraints only. Thus, it is possible to obtain the result by solving a number of linear equations. Efficient and scalable algorithms, such as those based on the conjugate gradient, can be applied to solve LS-SVM. As regards generalization performance, LS-SVM is similar to SVM [29]. However, LS-SVM is distinct from standard SVM and is used for linear equations instead of second-degree programming methods. In addition, due to the formulaic distinctions between them and the fact that more time is required for training, LS-SVM classifiers are different from standard SVM classifiers [27].

3. **Application of the Method Used.** This study can be examined in three stages: feature extraction, dimension reduction and classification. The features representing the heart sounds signals in the feature extraction stage were obtained using DWT. Then, the obtained features were reduced to lower dimensions by using the Shannon entropy. Finally, reduced features were classified into three groups — normal, mitral stenosis and pulmonary stenosis — by using the LS-SVM method. The methods DWT, Shannon entropy and LS-SVM were realized using the Matlab Software Package. All experiments were run on a machine with 2.8 GHz CPU, 4 GB of RAM, 500 GB HDD space and the

Windows 7 operating system. In advanced sections, these three stages will be explained in the following subsections.

3.1. **Feature extraction using DWT.** In this study, in order to obtain feature vectors from heart sound signals, DWT was used. DWT is often used in areas such as signal processing, coding, denoising and image compression [14]. DWT, as it is a developing practice technique, is used naturally in a number of practices. DWT provides detailed knowledge on the content of time frequency during a cardiac cycle of heart sound signal. The dilatation factor of DWT shows the low and high pass filters in the branch shape. In each step, it turns a low pass filter into lower and higher frequency components. The high frequency component of a signal is called the details (Ds), and that of the low frequency is called approximations (As) [30]. Generally, the high frequency knowledge is represented between d_1 and d_4, the first levels of decomposition, whereas low frequency knowledge between d_5 and d_8 is at later levels. In this study, decomposition is taken as level 5. Thus situated, heart sound signals are separated into (d_1-d_2-d_3-d_4-d_5) detailed sub-bands and finally into approximation sub-band a_5. Since the accuracy of the classification is closely related to the selected wavelet type for the application, it is important to work out which wavelet type to use. Usually, wavelet type selection is done after various tests are performed with different wavelet types when the one which gives the maximum efficiency for the application in question is worked out. In both the temporal and frequency domains, Daubechies wavelet has good localizing properties. The Order 4 (db4) Daubechies wavelet leads to more suitable detection of changes in the signals being studied. As a result, in this study, db4 was used to compute the wavelet coefficients.

3.2. **Application of the entropy to wavelet-transformed heart sound signals.** The feature extraction stage is the one which influences the performance of classification directly. According to the signal obtained nowadays, removing better and lower features from the patterns in pattern recognition systems is preferred rather than forming complicated classification structures. This is because the success of the classifier depends on the chosen feature domain rather than how to design it. At the same time, for the classifier, a large number of input parameters will increase so that they will be computationally intensive and time consuming. Besides this, the presence of excessive numbers of redundant, irrelevant and noisy input variables may hide the meaningful variables in the data set. Therefore, the aim was to choose fewer features to represent the data set rather than a large number of features.

In this study, as we have mentioned previously, the wavelet coefficients corresponding to the d_1-d_5 and a_5 as frequency bands of the heart sound signals were computed. Thanks to this, wavelet coefficients belonging to heart sound signals were used as feature vectors representing the signal. Here, as we are faced with high dimensional feature vectors, it is necessary to reduce these feature vectors. At this point, to reduce the dimensionality of the feature vectors, the Shannon entropy was applied. Shannon entropy values of heart sound signals in each sub-band were computed by the formula in Equation (5). In this way, the feature vector was extracted by calculating the six Wavelet entropy (five detailed — d_1-d_2-d_3-d_4-d_5 — and one approximation — a_5) values per heart sound signal. Thus, the feature vector had a lower dimension which included most of the useful information from the original vector.

3.3. **Classification of the heart sounds using LS-SVM.** Having completed the dimension reduction process by the Shannon entropy, we moved to the LS-SVM method used in this study for the classification process. LS-SVM classifiers were used in the classification of normal heart sounds, mitral stenosis heart sounds and pulmonary stenosis heart sounds.

In the classification process, generally, an input space shown as $x_1, x_2, x_3, \ldots, x_n$ is classified as an output space of $C_1, C_2, C_3, \ldots, C_n$. In the classification, we deal with the interconnection between the label y and feature vector x. The aim of the classification is to briefly estimate the mapping of $x \to y$.

In order to classify the data in the input space, SVM attempts to find the optimal separating hyperplane from all the probable separating hyperplanes. Thus, SVM maximizes the margin and in this way it gains a good generalization ability. Separating the hyperplane is a linear function which is able to divide the training sets into two groups. The following equation defines a separating hyperplane function:

$$D(x) = (w * x) + w_0 \tag{6}$$

In addition, all separating hyperplanes have to provide the following equation:

$$y_i[(w * x_i) + w_o] \geq 1, \quad i = 1, \ldots, n \tag{7}$$

There are some differences between LS-SVM and SVM. The biggest difference between these two methods is that, while LS-SVM uses a number of linear equations, SVM uses a quadratic programming problem.

In order to classify the feature chosen in this study, the Radial Basis Function (RBF), along with the LS-SVM method, is used as a kernel type. The RBF kernel function's equation is given as follows:

$$K(x, x') = \exp\left(-Ix - x'I^2/\sigma^2\right) \tag{8}$$

Kernels have been used in the transformation of the data from the input space into a high dimensional feature space [5].

4. **Results and Discussion.** In this study, a biomedical system based on DWT, the Shannon entropy and LS-SVM was developed in order to diagnose three different heart sounds. A total of 120 heart sounds (normal, mitral stenosis and pulmonary stenosis) were studied. The sounds studied were chosen from individuals aged between 4 and 65. 65 of them were males and 55 of them were females. 40 of the heart sounds were normal, 40 of them were mitral stenosis and the rest of them were pulmonary stenosis. As in previous studies, the training and test sets were adjusted as 50% training and 50% test. The distribution of the training and test data used in LS-SVM is shown in Table 1.

TABLE 1. Training and test sets

Class	Training set	Test set	Total
Normal heart sound	20	20	40
Pulmonary stenosis heart sound	20	20	40
Mitral stenosis heart sound	20	20	40
Total	60	60	120

Consisting of only one period with detail coefficients of the first five decomposition levels and wavelet approximations belonging to normal heart sounds, mitral stenosis heart sounds and pulmonary stenosis heart sounds are shown, respectively, in Figures 4-6.

As seen in Figures 4-6, there are remarkable differences between the graphs of normal heart sounds and mitral stenosis or pulmonary stenosis heart sounds. These differences are

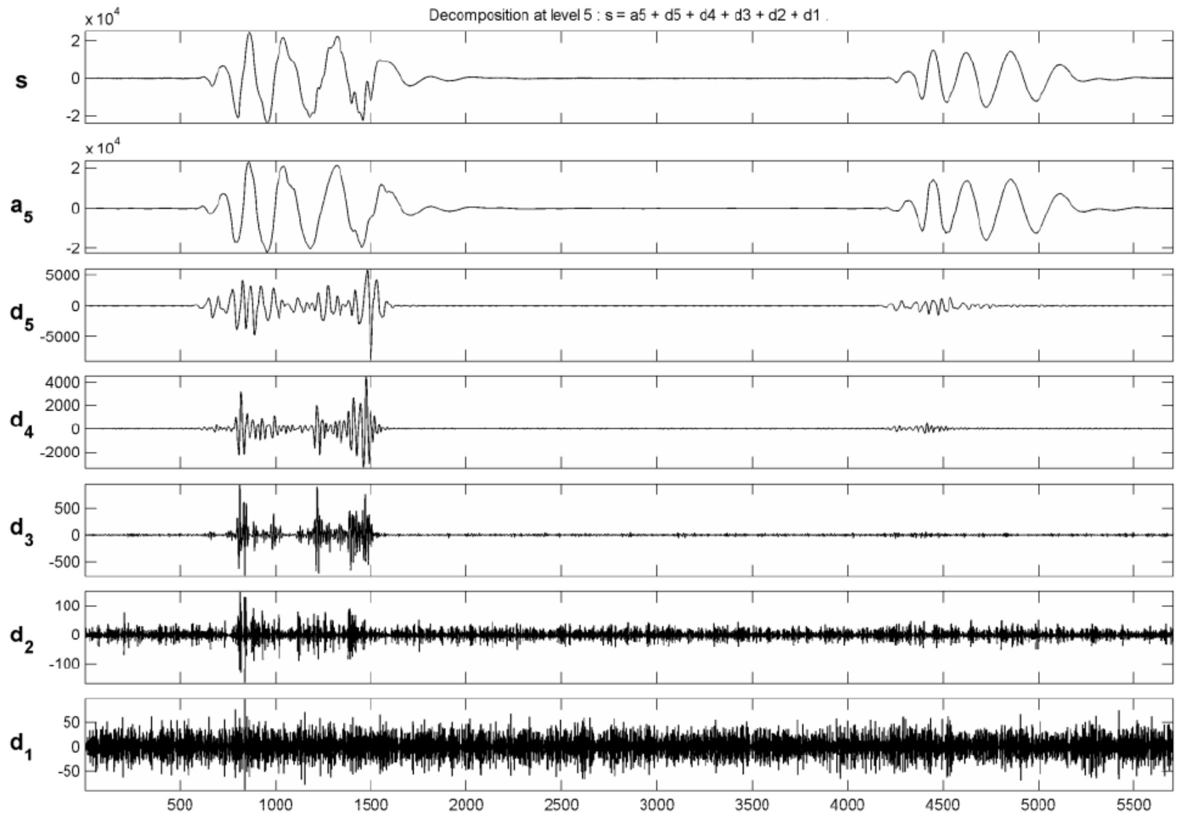

FIGURE 4. Wavelet detail coefficients and wavelet approximation coefficients at the first five decomposition levels of a normal heart sound

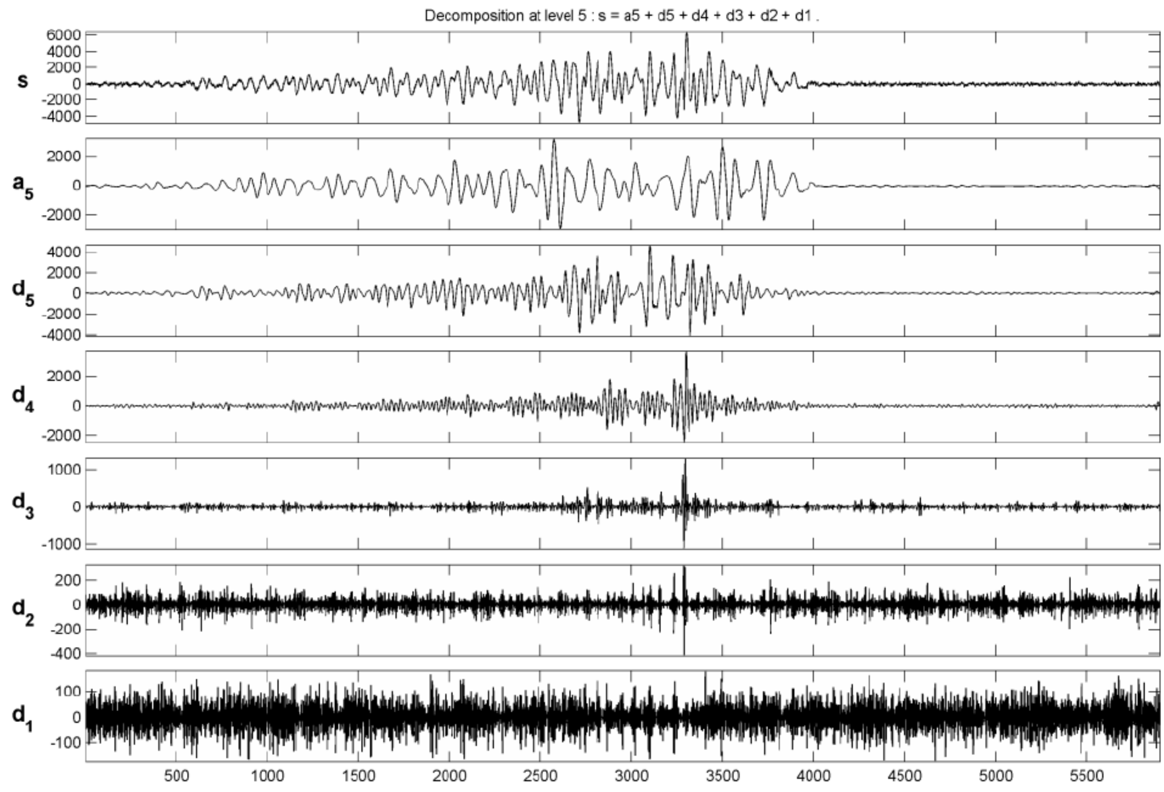

FIGURE 5. Wavelet detail coefficients and wavelet approximation coefficients at the first five decomposition levels of a mitral stenosis heart sound

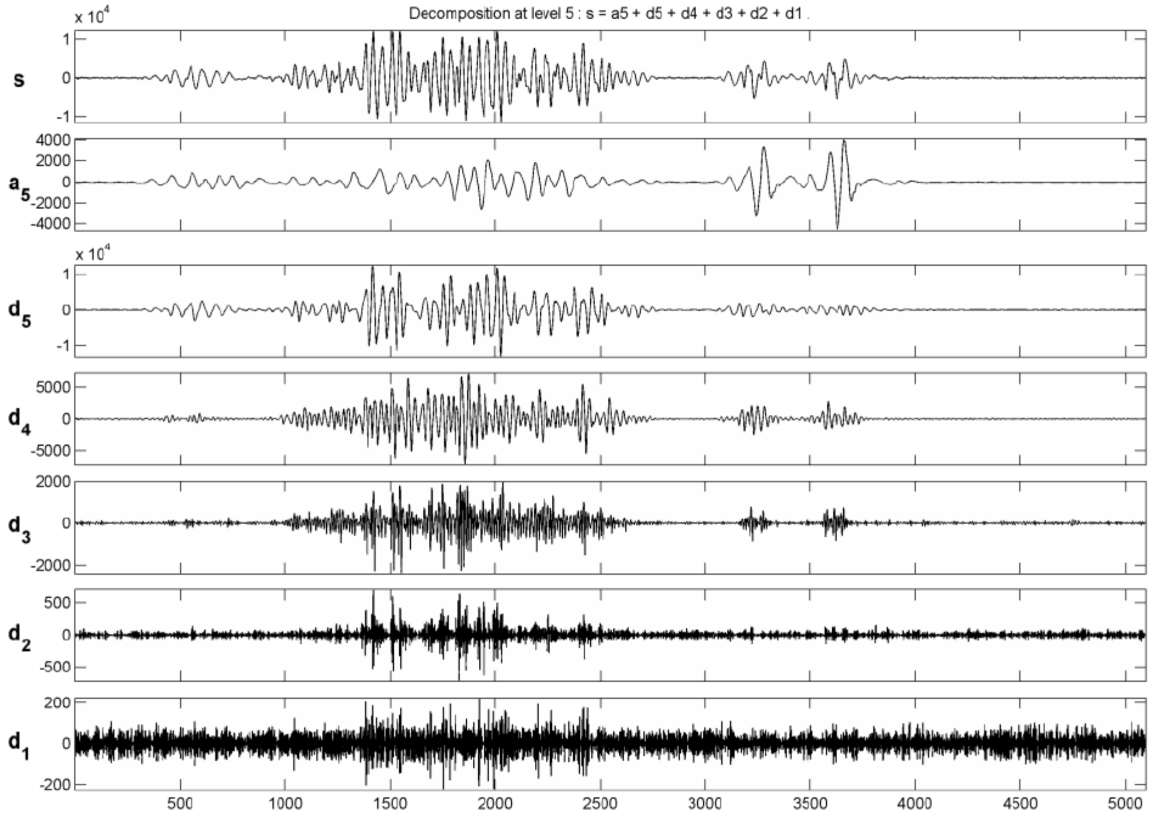

Figure 6. Wavelet detail coefficients and wavelet approximation coefficients at the first five decomposition levels of a pulmonary stenosis heart sound

reflected in the DWT graphics. By considering these variations in DWT, a classification system was formed and it was used in diagnosing the different heart sounds.

In this study, 20 normal heart sounds, 20 mitral stenosis heart sounds and 20 pulmonary stenosis heart sounds, totalling 60 heart sounds, were chosen for the training stage of LS-SVM. The sounds chosen for the training set were recorded using an equal number for each sound. The rest of the 60 sounds were used with the aim of testing them.

The classifier performance of SVM largely depends on the choice of kernels. There are many types of kernels, such as the Radial Basis Function (RBF) kernel, the polynomial kernel and B-splines kernel. In this study, the RBF kernel, which is one of the most popular kernel functions, is used. For the implementation of the LS-SVM using RBF kernel functions, one has to assume a value for σ. The optimal σ can only be found by systematically varying its value in the different training sessions [28]. The sigma (σ) value used in LS-SVM according to the performance is given in Figure 7. As seen in Figure 7, when the sigma (σ) value is 1.5, the best classification performance is obtained.

ROC curves are given in Figure 8. The ROC curve is a reliable technique for evaluating classification performance depending on the negative false values and true positive values. In this study, in the results of the classification, the area remaining under the ROC curve was used and so the assessment of the classification performance was actualized. As the area remaining under the ROC curve is nearly 1, it shows the success of the classifier's performance.

To evaluate the performance of the classifier, sensitivity, specificity and total classification accuracy analyses were carried out. Sensitivity and specificity analyses are important measures for the performance of diagnostic tests. The sensitivity, specificity and total

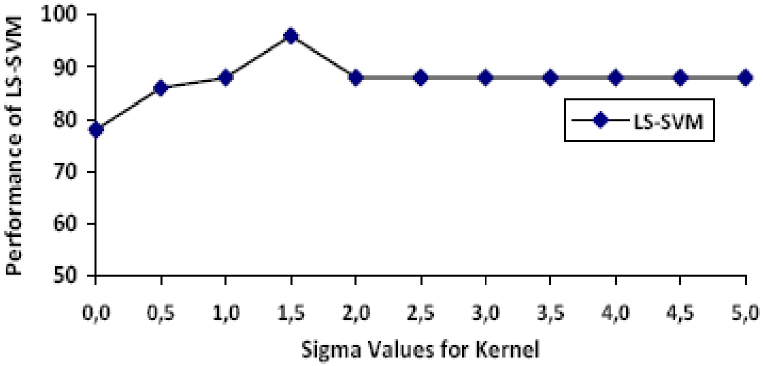

FIGURE 7. Performance of LS-SVM against σ values

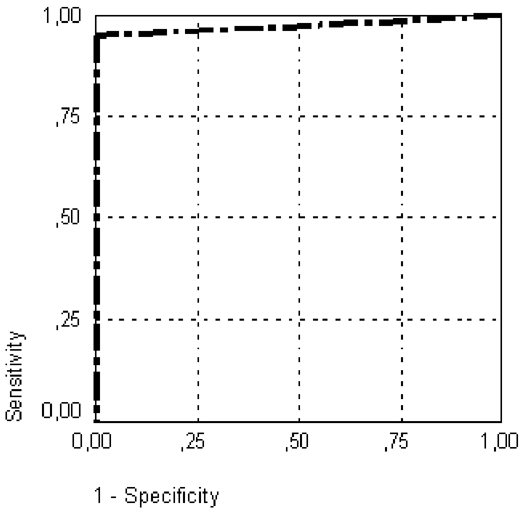

FIGURE 8. ROC curves for LS-SVM

classification accuracies are defined by the following equations:

$$Sensitivity = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \tag{9}$$

$$Specificity = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \tag{10}$$

where TP, TN, FP and FN denote true positives, true negatives, false positives and false negatives, respectively.

$$\text{Total Classification Accuracy} = \text{number of correct decisions/total number of cases} \tag{11}$$

The developed system provided 96.66% classification accuracy for the sensitivity and specificity values obtained in our study; the number of features and the areas under the ROC curves with total accuracy values are given in Table 2. In addition, there are total accuracy values with areas under the ROC curves, and sensitivity and specificity values obtained from the previous studies [9,24] can be seen in Table 2.

As seen in Table 2, the study carried out before by Güraksın et al. [9] on the same data set using ANN and DFT brought about 91.6% success. In another study carried out by Uğuz on the same data set [24], along with DFT, Burg and PCA, ANN was also used in the classification process, and 95% success was obtained. In this study, the classification accuracy obtained was found to be higher than that of previous studies.

Table 2. Statistical parameters of the method used and the methods of previous studies

Classifier	Number of features	Specificity %	Sensitivity %	The areas under the ROC curves	Total accuracy %
Method used	6	90.9	100	0.975	96.66
DFT/Burg AR-PCA-ANN [24]	33	90.48	97.44	0.950	95
DFT-ANN [9]	300	82.60	97.29	0.925	91.67

According to Table 2, higher classification accuracy was obtained with less features when the wavelet and entropy based methods were used. Consequently, feature extraction via the DWT-Shannon entropy based method removed redundant or irrelevant features from the feature space, thereby reducing the high dimensionality of the features and decreasing the computational complexity of the SVM algorithm used in the classification. Thus, the feature vector has a lower dimension which includes most of the useful information from the original vector, and the classification performance is increased.

Table 3 shows the LS-SVM classification's confusion matrix. As seen in Table 3, one subject with mitral stenosis was incorrectly classified as a pulmonary stenosis subject, and one subject with pulmonary stenosis was incorrectly classified as a patient suffering from mitral stenosis.

In this study, a biomedical system in search of the classification of normal, mitral stenosis and pulmonary stenosis heart sounds was designed. In the system designed to classify the heart sounds, heart sound signals were divided into sub-bands through DWT for the very first stage of the feature extraction process. Next, each of the sub-bands' entropies was computed using the Shannon entropy algorithm to reduce the feature vectors obtained from DWT. The reduced features obtained from three different heart sounds were used as input patterns in the LS-SVM classifier.

Table 3. Confusion matrix for the LS-SVM classifier

Output/Desired	Normal heart heart sound	Mitral stenosis heart sound	Pulmonary stenosis heart sound
Normal heart sound	20	0	0
Mitral stenosis heart sound	0	19	1
Pulmonary stenosis heart sound	0	1	19

In comparison with the previous studies in terms of the classification performance, the method developed is found to be more successful than the other classifiers. For further improvement, the DWT-Shannon entropy based method was utilized as a feature extraction tool and it is shown that the results were considerably enhanced using the reduced dimension feature set. The classification findings show that the developed system could provide successful results in heart-valve diseases. In addition, this type of decision-support system will help physicians in the clinical environment diagnose heart diseases.

Further improvement can be obtained when the different entropy functions such as the norm and threshold entropy are used. Alternatively, in DWT based feature extraction stage, feature selection and dimension reduction techniques such as information gain, genetic algorithm, and principal component analysis may be used for feature reduction. Thus, classifier accuracy may be increased even more by removing noisy, redundant or irrelevant features from feature space.

The number and variety of sounds used in this study may not be sufficient to demonstrate the performance of the classification. Therefore, the plan is to carry out some studies on the data from more and much more different heart diseases in the future. Working with a bigger dataset, it may be possible to design a biomedical system which may be able to classify all of the types of heart disease. Additionally, the combined DWT/Shannon entropy/LS-SVM approaches which were used can be applied to other biomedical signals such as Doppler signals and lung sounds.

This diagnostic system could help other researchers who want to study the classification of heart sounds. Besides being used in the clinical environment, such systems can be used for the training of medical students.

Acknowledgment. The authors wish to acknowledge the Afyon Kocatepe University Scientific Research Council with the Project number "07.AFMYO.01" for the collection of the heart sounds used in this study. This study has been supported by Scientific Research Project of Selcuk University.

REFERENCES

[1] Z. Jiang and S. Choi, A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope, *Expert Systems with Applications*, vol.31, pp.286-298, 2006.

[2] Z. Sharif, M. S. Zainal, A. Z. Sha'ameri and S. H. S. Salleh, Analysis and classification of heart sounds and murmurs based on the instantaneous energy and frequency estimations, *Proc. of IEEE TENCON*, vol.2, pp.130-134, 2000.

[3] C. Ahlström, *Processing of the Phonocardiographic Signal-Methods for the Intelligent Stethoscope*, Ph.D. Thesis, Linköping University, Institute of Technology, Linköping, Sweden, 2006.

[4] R. K. Sinha, Y. Aggarwal and B. N. Das, Backpropagation artificial neural network classifier to detect changes in heart sound due to mitral valve regurgitation, *Journal of Medical Systems*, vol.31, pp.205-209, 2007.

[5] E. Çomak, A. Arslan and İ. Türkoğu, A decision support system based on support machines for diagnosis of the heart valve diseases, *Computers in Biology and Medicine*, vol.37, pp.21-27, 2007.

[6] C. Ahlstrom, P. Hult, P. Rask, J. E. Karlsson, E. Nylander, U. Dahlström and P. Ask, Feature extraction for systolic heart murmur classification, *Annals of Biomedical Engineering*, vol.34, no.11, 2006.

[7] G. E. Güraksın, U. Ergün and Ö. Deperlioğu, Performing discrete fourier transform of the heart sounds on the pocket computer, *The 14th National Biomedical Engineering Meeting*, İmir, Turkey, pp.124-127, 2009.

[8] A. Kandaswamy, C. Kumar, R. Ramanathan, S. Jayaraman and G. Malmurugan, Neural classification of lung sounds using wavelet coefficients, *Computers in Biology and Medicine*, vol.34, pp.523-537, 2004.

[9] G. E. Güraksın, U. Ergün and Ö. Deperlioğlu, Classification of the heart sounds via artificial neural network, *International Symposium on Innovations in Intelligent Systems and Applications*, Trabzon, Turkey, pp.507-511, 2009.

[10] G. E. Güraksın, *Classification of the Heart Sounds via Artificial Neural Network*, Master Thesis, Institute of Natural and Applied Sciences, Afyon Kocatepe University, Afyonkarahisar, Turkey, 2009.

[11] A. Subasi, EEG signal classification using wavelet feature extraction and mixture of expert model, *Expert System with Application*, vol.32, no.4, pp.1084-1093, 2007.

[12] O. N. Uçan, O. Osman and A. M. Albora, *Image Processing Techniques and Applications in Engineering*, 1st Edition, Nobel Publications, 2006.

[13] O. Rioul and M. Vetterli, Wavelets and signal processing, *Signal Processing Magazine IEEE*, vol.8, no.4, pp.14-38, 1991.

[14] S. Babaei and A. Geranmayeh, Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transform of PCG signals, *Computer in Biology and Medicine*, vol.39, pp.8-15, 2009.

[15] H. Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, *Expert System with Application*, vol.36, pp.2027-2036, 2009.

[16] S. G. Mallat, A theory for multi resolution signal decomposition: The wavelet representation, *IEEE Transactions on Pattern Anal Machine Intelligence*, vol.11, no.7, pp.674-693, 1989.
[17] I. Turkoglu, A. Arslan and E. Ilkay, An intelligent system for diagnosis of the heart valve diseases with wavelet packet neural networks, *Computers in Biology and Medicine*, vol.33, pp.319-331, 2003.
[18] T. S. Leung, P. R. White, W. B. Collis, E. Brown and A. P. Salmon, Classification of heart sounds using time-frequency method and artificial neural network, *The 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol.2, pp.988-991, 2000.
[19] S. Chauhan, P. Wang, C. S. Lim and V. Anantharaman, A computer aided MFCC based HMM system for automatic auscultation, *Computers in Biology and Medicine*, vol.38, no.2, pp.221-233, 2008.
[20] R. Folland, E. L. Hines, P. Boilot and D. Morgan, Classifying coronary dysfunction using neural networks through cardiovascular auscultation, *Med. Biol. Eng. Comput.*, vol.40, pp.339-343, 2002.
[21] T. R. Reed, N. E. Reed and P. Fritzson, Heart sound analysis for symptom detection and computer-aided diagnosis, *Simulation Modelling Practive and Theory*, vol.12, pp.129-146, 2004.
[22] I. Maglogiannis, E. Loukis, E. Zafiropoulos and A. Stasis, Support vectors machine-based identification of heart valve diseases using heart sounds, *Computer Methods and Programs in Biomedicine*, vol.95, pp.47-61, 2009.
[23] I. Turkoglu, A. Arslan and E. Ilkay, An expert system for diagnosis of the heart valve diseases, *Expert Systems with Applications*, vol.23, pp.229-236, 2002.
[24] H. Uğuz, A biomedical system based on artificial neural network and principal component analysis for diagnosis of the heart valve diseases, *Journal of Medical Systems*, 2010.
[25] S. Tong, A. Bezerianos, J. Paul, Y. Zhu and N. Thakor, Nonextensive entropy measure of EEG following brain injury from cardiac arrest, *Physica A: Statistical Mechanics and Its Applications*, vol.305, no.3-4, pp.619-628, 2002.
[26] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, vol.27, pp.623-656, 1948.
[27] E. Çomak, *Proposals to Solve Support Vector Machines Multi Class Problems*, Master Thesis, Graduate School of Natural and Applied Sciences, Selçuk University, Konya, Turkey, 2004.
[28] E. D. Übeyli, Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals, *Expert Systems with Applications*, vol.37, pp.233-239, 2010.
[29] J. Ye and T. Xiong, SVM versus least squares SVM, *Proc. of International Conference on Artificial Intelligence and Statistics*, pp.640-647, 2007.
[30] Z. Yan, Z. Jiang, A. Miyamoto and Y. Wei, The moment segmentation analysis of heart sound pattern, *Computer Methods and Programs in Biomedicine*, vol.98, no.2, pp.140-150, 2010.
[31] H. Tamura and K. Tanno, Midpoint validation method for support vector machines with margin adjustment technique, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(A), pp.4025-4032, 2009.
[32] Y. Xing, X. Wu and Z. Xu, Multiclass least squares auto-correlation wavelet support vector machines, *ICIC Express Letters*, vol.2, no.4, pp.345-350, 2008.
[33] Q. She, H. Su, L. Dong and J. Chu, Support vector machine with adaptive parameters in image coding, *International Journal of Innovative Computing, Information and Control*, vol.4, no.2, pp.359-367, 2008.