Detection and Classification of Cardiac Murmurs using Segmentation Techniques and Artificial Neural Networks

S. L. Strunic, F. Rios-Gutierrez, R. Alba-Flores, G. Nordehn, S. Burns

University of Minnesota Duluth
1023 University Drive
Duluth, MN 55812
stru0121@d.umn.edu

Abstract

A diagnostic system based on Artificial Neural Networks (ANN) is implemented as a detector and classifier of heart murmurs. Segmentation and alignment algorithms serve as important pre-processing steps before heart sounds are applied to the ANN structure. The system enables users to create a classifier that can be trained to detect virtually any desired target set of heart sounds. The output of the system is the classification of the sound as either normal or a type of heart murmur. The ultimate goal of this research is to develop a tool that can be used to help physicians in the auscultation of patients and thereby reduce the number of unnecessary echocardiograms-- those that are ordered for healthy patients. Testing has been conducted using both simulated and recorded patient heart sounds. Results are described for a system designed to classify heart sounds as normal, aortic stenosis, or aortic regurgitation. The system is able to classify with up to $85 \pm 7.4\%$ accuracy and $95 \pm 6.8\%$ sensitivity for a group of 72 simulated heart sounds. The accuracy rate of the ANN system for simulated sounds is compared to the accuracy rate of a group of medical students who were asked to classify heart sounds from the same group of sounds classified by the ANN system.

Background

Cardiology is a popular field for the different types of ANN applications [1-6]. Many studies have worked toward designing practical murmur detector and classifier systems to improve the diagnostic accuracy of physicians in small practice settings. As an example, one study used an ANN structure to classify pediatric heart sounds as either innocent or pathological [6]. Sounds were recorded at a hospital using a microphone, and the frequency spectrum was used as input to an ANN structure. The classifier achieved high accuracy, but the system required its users to select input vectors by hand based on the best observed cycles in terms of background noise and heart sound clarity. The need for a person to assess which signals are best to use for the ANN system makes the system non-ideal for a real-world scenario—a problem that is common to several reported classifier systems [1-2].

Various segmentation algorithms are used in an attempt to eliminate the need for manual selection of heart cycles from heart sounds [1, 4-6]. Segmentation algorithms extract individual heart cycles from heart sound recordings based on properties common to all cardiac signals. Still though, problems arise due to issues such as background noise while recording heart sounds. Also, some algorithms require the inclusion of the electrocardiogram signal to complete segmentation [1, 5-6].

The intent of this research is to develop a working classifier system using ANNs that accepts heart sound recordings directly, processes the sounds, and classifies the inputs to identify the type of cardiac event that is present in the heart sounds. An important goal is to eliminate the need for human interaction in selecting the best part of the heart sound to use as input. The research begins by developing a system which works using simulated heart sounds, and later the testing is extended to include heart sounds recorded from patients using a commercially available electronic stethoscope and a common laptop running a Windows operating system.

Introduction

Cost effective, accurate and early detection of cardiac illnesses is important to curb deaths caused by cardiovascular diseases. When a patient visits the physician for auscultation, a heart murmur is the most common abnormal finding. When a murmur is detected, the physician must decide whether it represents either a pathological or an innocent murmur—one which does not represent current or future illness. Oftentimes, a physician who suspects that a patient is healthy will still order an echocardiogram for reassurance. The result of this practice is a misallocation of healthcare funds, since echocardiograms cost between $750-1500 [7-8]. While it is clearly important to avoid type-I errors where healthy patients are sent for echocardiogram, it is also important to avoid type-II errors—where a patient who has a pathological heart murmur is sent home without proper treatment.

In this work, a diagnostic system is implemented that can help to reduce the number of echocardiograms that are ordered for healthy patients. The diagnostic system is based on an easy-to-use graphical user interface that has been designed using MATLAB software and the ANN Toolbox. The system allows the user to interactively design and create a heart sound classification system that implements a set of ANNs as a means for classification. After a classifier has been designed and created, the user is able to select any audio files on his or her computer to use as input to the ANN system.

The ultimate goal of the diagnostic system is to provide physicians with a reliable and inexpensive classification tool to use along with auscultation. In the event that a patient has a heart sound that is difficult for physicians to diagnose, the classifier could reduce the tendency of the physician to refer the patient to echocardiogram when the procedure is not necessary.

To show the system's ability to classify heart sounds, testing has been performed using the ANN classifier to distinguish between three specific types of heart sounds: normal, aortic stenosis (AS), and aortic regurgitation (AR).

Given an input heart sound of virtually any length or heart rate, a segmentation algorithm is applied to compute the representative single heart cycle for that sound. The segmentation algorithm is used to identify the heart sound components S1 and S2. Once the positions of these components are located, individual heart cycles can be identified and the average of all the cycles within the sound is computed. The heart cycle obtained from the averaging process will be the one that represents the particular sound under study. Next, the spectrogram of that heart cycle is computed. Using the spectrogram, the 195Hz frequency component (time vs. amplitude vector) of each sound in the design (or test) set are extracted and aligned. Through experimentation it was found that the 195Hz band of the heart sound contains the necessary information for the ANN to identify the type of murmurs that are considered in this study. The alignment algorithm then prepares the vectors that will be inputted into the ANN system for the design (or test) phase.

To design and test the ANN system, a heart sound library consisting of 72 simulated heart sounds is used. The simulated heart sounds were recorded from a device that has been used by medical schools to teach auscultation to students. Also, recorded patient heart sounds are available for testing from the Murmur Study Library at the University of Minnesota Duluth. The sounds from the Murmur Study Library were recorded (with IRB approval) using a commercially available electronic stethoscope provided by Welch-Allyn. Echocardiogram reports from St. Luke's Hospital in Duluth, MN, have been used to define the correct condition of each recorded patient heart sound. The patient sounds used for testing in this research include 7 examples of normal, 4 examples of AS, and 2 examples of AR.

As a means for comparison of ANN system results to human accuracy, a group of medical students at the University of Minnesota Medical School Duluth were asked to classify sounds from the same set of simulated heart sounds used to design the ANN system.

Methods

Heart Sounds Database Used for Designing and Testing

The heart sounds used to design the ANN system were recorded from a simulator provided by the University of Minnesota School of Medicine Duluth. The simulator is commonly used to aurally teach heart murmurs to medical students. The heart sound library consists of 24 normal sounds, 24 AR sounds, and 24 AS sounds (12 innocent and 12 pathological). The sounds within each group of 24 files varied in heart rate, amplitude and duration of murmur.

The Murmur Study Library from the University of Minnesota Duluth provides recorded patient sounds that can be used to test the system. These patient heart sounds were recorded from four locations on the chest using an electronic stethoscope provided by Welch Allyn and a laptop running Windows XP.

Currently, the heart sound databank includes a total of more than 110 sounds from over 28 patients, including more than four types of murmurs. However, the recorded patient sounds used for testing in this work include only 7 examples of normal, 4 examples of AS, and 2 examples of AR. These 13 patient heart sounds were selected as good representations of the heart murmurs which this research focuses on. There were several reasons why only 13 of the 110 total patient sounds were used for testing. One reason is excess background noise in many of the sounds which leads to failure of the segmentation and alignment algorithms. Another reason is that some of the echocardiogram reports for the patient sounds are vague, listing multiple murmurs of various degrees. Considerations have not yet been included for account for these situations.

Since the set of 13 patient sounds is a very limited sample size, the heart cycle averaging step was omitted for testing patient sounds. Instead, individual heart cycles were extracted from these 13 patient sounds, resulting in 22 normal cycles, 13 AS cycles, and 6 AR cycles. Each of these 41 cycles then represents the type of heart sound from which it was extracted. The heart cycle averaging technique will be applied to patient sounds once a sufficient set of useable sounds is available for testing.

Data Type and Pre-Processing Steps Used for ANN Vectors

The heart sound data that is provided to the complete system must be in .wav audio format. There are numerous commercially available electronic stethoscopes capable of recording heart sounds and transferring those sounds to a computer in .wav format. Heart sounds of virtually any heart rate or duration can be inputted to the system.

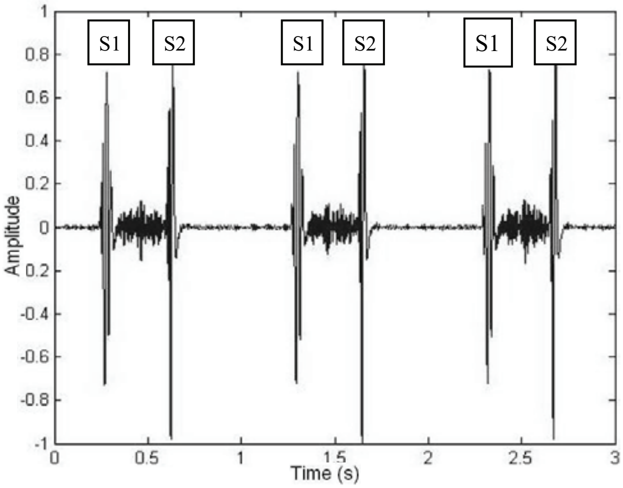

Fig 1. The time vs. amplitude plot of a heart sound with severe AS.

The first pre-processing step uses a segmentation algorithm to identify the heart sound components S1 and S2 within the original heart sound (Fig. 1). The segmentation method uses the components of S1 and S2 that occur at low frequency levels relative to murmurs. The spectrogram of the full heart sound is computed and the 45Hz frequency band is extracted (Fig. 2). The high amplitude components of the 45Hz band of the heart sound are observed, and S1 and S2 are identified based on the timing between those high amplitude components. The fact that the time from S1 to S2 (systole) is always less than the time from S2 to S1 (diastole) is the basis for this algorithm.

Once the S1s and S2s are located, the individual heart cycles within the heart sound can be identified and re-sampled to normalize their lengths. Only a single heart cycle is used for the analysis and classification of the sound. Instead of taking any of the individual cycles from a heart sound for the analysis, the average of the single cycles is computed and the resulting averaged heart cycle for that heart sound will be the one that is processed and inputted into the ANN for design (or testing) (Fig. 3). Compared to the use of a single extracted cycle from a heart sound, the heart cycle averaging technique provides a better representation of the complete heart sound. Also, the averaging technique helps to reduce the effect of anomalous heart cycles that may be present in a heart sound.

The next pre-processing step is to compute the spectrogram of the averaged heart cycle. Using the spectrogram, the 195Hz frequency band is extracted to serve as the input vector that will be fed into the ANN classifier. After studying and analyzing several frequency bands in the spectrogram for different heart sounds, the 195Hz component was found to contain the necessary information for the ANN to identify the types of murmurs are considered in this study. Fig. 4 illustrates the extraction of the 195Hz band from the spectrogram of the heart sound. The murmur is apparent in the 195Hz band—in the case of Fig. 4, the murmur occurs between S1 and

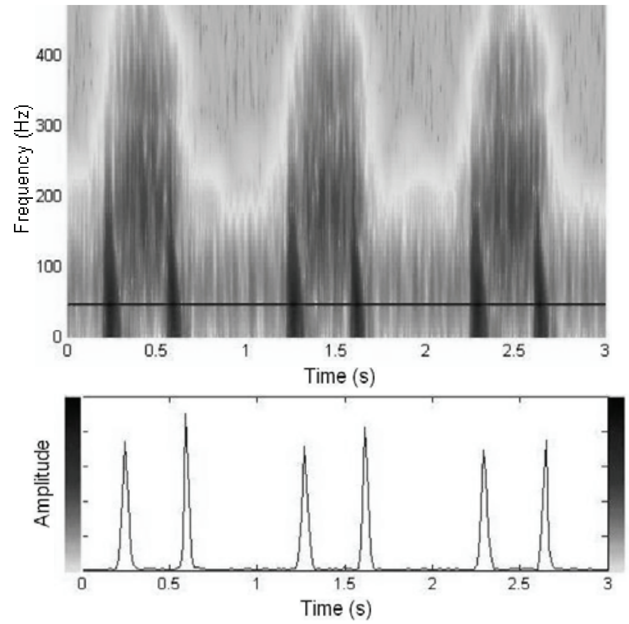

Fig 2. Illustration of the extraction of the 45Hz frequency component (bottom) of the heart sound using the spectrogram (top).

S2. The group of 195Hz frequency bands from all averaged heart cycles in the design (or test) set makes up the set of vectors that will be inputted into the ANN system for design (or testing). First, though, an alignment algorithm is applied.

In order to account for the variability in S1-S2 intervals between heart sounds with different heart rates, an alignment algorithm is applied to the design (or test) vector set. The total number of data points to use for the input vectors to the ANN structure is first set—for example, 25 data points are used for every input vector to match the 25 input neurons of the ANN. The alignment is performed by sampling 60% of the desired number of total data points equally spaced from the end of S2 to the end of S1 and sampling the remaining 40% of total data points from the end of S1 to the end of S2. The result is that for different

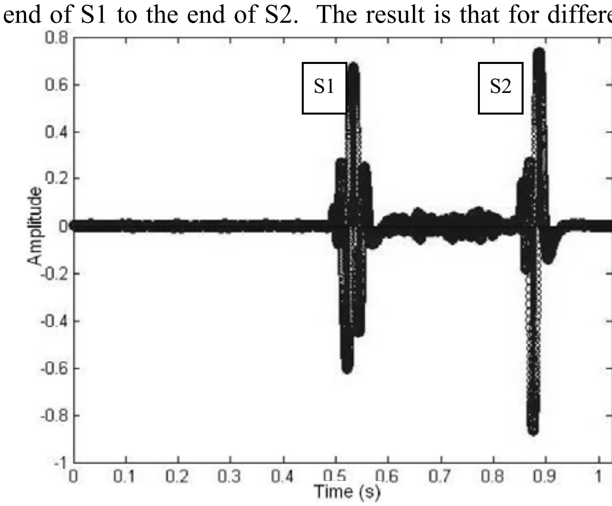

Fig 3. The average heart cycle—the result of averaging together the single cycles found by applying the segmentation algorithm to the heart sound in Fig. 1.

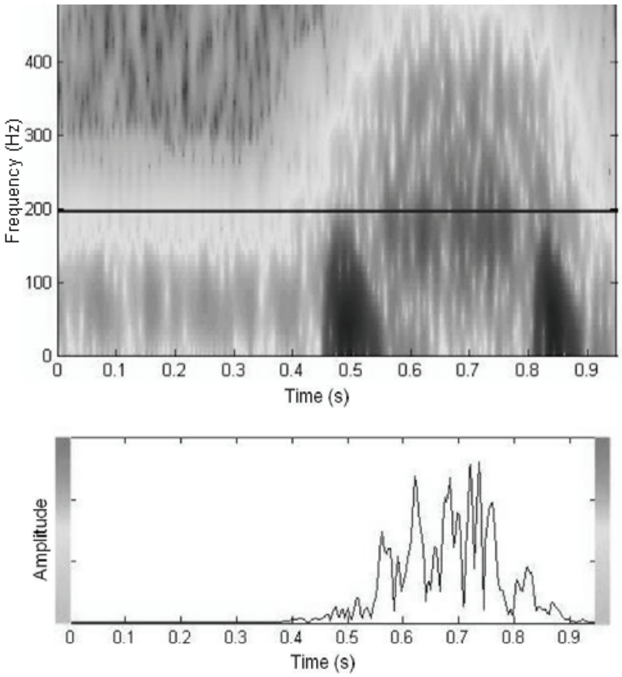

Fig 4. Illustration of the extraction of the 195Hz frequency component (bottom) of the heart sound using the spectrogram (top).

sound, regardless of the time between S1 and S2, the same number of data points are taken from before (60% of total data points) and after (40% of total data points) the end of S1 (Fig. 5).

Sampling at a fixed number of data points across the single heart cycle maintains sufficient detail that was found in the original frequency band while eliminating some redundancy of information. The alignment algorithm ensures that the input data is always presented to the input neurons of the ANN in the same order: from the end of S2 to the end of the subsequent S2.

ANN Architecture Used to Create the Classifier

Resilient back-propagation was used as the training algorithm for the feed-forward ANN classifier since it is recognized as a good algorithm for the purpose of pattern recognition [12]. The results shown in this report are from an ANN system using 3 hidden layers with 25 neurons each, 25 input neurons, one output neuron, training mean squared error goal of 0.0005, and a learning rate of 0.0005. These parameters were optimized by repetition and comparison, with consideration from previous work [13].

Design of the ANN System

The process of designing the ANN system encompasses both the training and validation steps that are commonly referred to in ANN literature [2-3, 18]. In order to design the ANN classifier, a set of training vectors based on the desired training targets must first be provided. The user specifies separate directories on their computer for each desired target. Each of these directories contains only one type of heart sound—for example the group of 24 unique

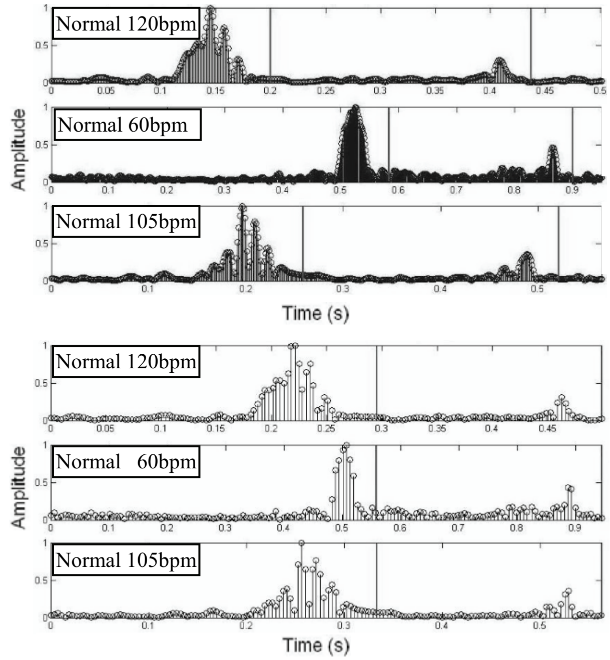

Fig 5. Before (top) and after (bottom) the alignment algorithm is applied to a set of 195Hz bands for 3 different normal heart sounds. The vertical lines mark the end of S1 (left) and S2 (right) in the plots.

normal sounds or the group of 12 innocent AS sounds. When a directory is selected to represent a training target, the system randomly selects 50% of the files within that directory to use as the design set for that target, while the other 50% of the files are left to use as the test set. Once the user specifies that they are ready to train the neural network (i.e. all of the desired targets have been defined), the system is trained and all ANN characteristics are fixed.

As an example of a classifier design set, consider the target set which trains the ANN system to detect normal, AR and AS. Three training target directories are selected. Each of the three directories contains 24 unique heart sound files, thus each of the three training target sets consists of 12 unique heart sound files. The remaining files in the three directories that were not used for the design are used as the test set.

Procedure Used to Test the ANN System

Once the ANN classifier has been designed and created, the user can observe output results using either the test set(s) or any other heart sounds on their computer. An output is then provided which indicates which training target is most similar to the current test input.

A comparison of the simulated heart sounds used in this research to heart sounds recorded from an electronic stethoscope shows that the sounds compare closely after some filtering. Thus, designing the system with simulated sounds is expected to yield useful results for testing both simulated and patient-recorded heart sounds.

Procedure Used to Test 2nd Year Medical Students

In order to compare ANN system accuracy to human accuracy, a group of 2nd year medical students at the University of Minnesota School of Medicine Duluth was asked to classify a set of simulated heart sounds taken from the same set used for ANN testing. The accuracy of the 2nd year medical students is shown for set of target sounds which corresponds to the set used for the ANN system.

Results and Analysis

Now that the ANN system has been designed and tested, sensitivity and specificity can be explored. Sensitivity is a very important measure for this particular research since it is a measure of the percentage of patients with unhealthy hearts that are recognized as such. With high sensitivity, the system has fewer Type II errors—the case when an unhealthy heart is classified as healthy.

Specificity is the percentage of healthy cases that are classified as healthy. With high specificity, the system has fewer Type I errors—the case when a healthy heart is classified as unhealthy.

For this system, high sensitivity is more important than high specificity. Higher sensitivity increases the number of patients with healthy hearts who are told they are unhealthy and sent to echocardiogram for further testing. More importantly, higher sensitivity reduces the number of patients with an unhealthy heart that are told they are healthy. On the other hand, higher specificity reduces the number of healthy patients who are referred to echocardiogram for further testing, while it increases the number of patients who have a heart murmur but are still told that they have a healthy heart—they are released from care with a potentially deadly heart condition. Therefore, high sensitivity is a much more desirable goal for this system.

The accuracy of the ANN system is compared to the accuracy of the 2nd year medical students in Fig. 6. The illustration shows the 95% confidence intervals for the true mean values based on multiple trials. The 95% confidence interval for the sensitivity of the tested ANN system was $95 \pm 6.8\%$. The range for the sensitivity of the medical students has not been reported because a detailed sensitivity analysis remains to be completed. However, the general trend is that, for the medical students, there are few false-negative diagnoses, i.e. the students rarely classified a heart sound as normal when a murmur was in fact present. The most common error is that the student recognizes that a murmur is present, but they are unable to identify the type of murmur. In short, the sensitivity of the medical students was high.

Observing these results, there are several important features to note. First, the reported auscultative accuracy rate of primary care physicians is between 20-40% [7, 9, 17]. Although this range is the accuracy of physicians who are detecting heart murmurs using patient sounds, the accuracy rate for the 2nd year medical students observed in this study is similarly low when trying to classify

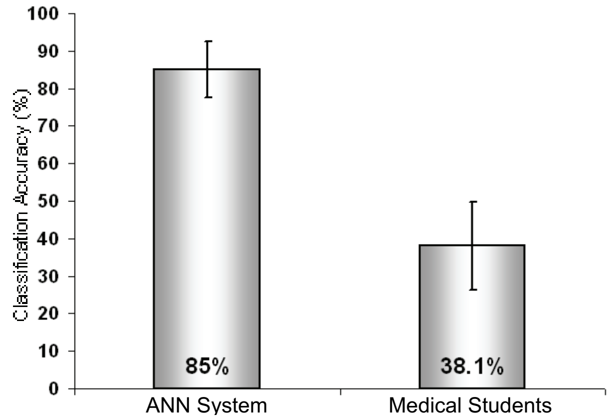

Fig. 6. Comparison of classification accuracy between ANN system and medical students.

simulated heart sounds. For each of the three target sets used in this research, the ANN system clearly has far better accuracy than the medical students.

Another important piece of information is that, in general, expert cardiologists have an auscultative accuracy rate of roughly 80% [7, 13]. The 95% confidence interval given for the accuracy rate of the ANN system includes the 80% accuracy level, which indicates that if these results could be reproduced using patient sounds, this particular classifier system could prove very useful as an aide for primary care physicians before deciding whether or not to refer a patient for an echocardiogram.

System performance using patient sounds was less impressive than performance using simulated sounds. For the system trained to detect and classify simulated sounds representing normal, AS, and AR, the system achieved an average of 85% accuracy. However, when testing this same system with recorded patient sounds, the average accuracy was only 48.5%.

There are several factors that led to the low accuracy when testing with patient sounds. After filtering the patient sounds, there is still some background noise present. The system seems to have difficulty classifying sounds with excess background noise when it has been trained using simulated sounds with a very low level of background noise. Also, the alignment algorithm was not able to properly align some of the 41 patient heart cycles. In these cases the system would be less likely to correctly classify the sounds, since testing without properly applying the alignment algorithm means that S1 and S2 are not applied to the same input neurons that were used during system training.

One potential solution to address the background noise in patient sounds is to train the system with noisy simulated sounds. White Gaussian noise could be added to the original simulated sounds, and then those noisy simulated sounds could be used to train the ANN system. Additionally, more advanced filtering methods could be applied to the patient sounds. Literature suggests that wavelet analysis is a powerful tool for extracting only the meaningful portions of heart sounds, and this may lead to successful noise suppression [2-5]. Reducing noise within the patient sounds would also help to reduce the failure rate of the alignment algorithm, since the main reason for failure of alignment is the presence of high amplitude anomalies at very low frequencies. It may also be possible to improve stethoscope recording methods so that there is less unintentional chest piece movement, which results in unwanted noise.

The current accuracy rate of the system for patient sounds is clearly not high enough to significantly improve a real-world situation. However, the suggested changes to the overall system design are expected to improve performance.

Conclusion

The auscultative accuracy rate of the average physician is clearly low, and this fact leads to the referral of healthy patients for echocardiogram. Unnecessary referral to this costly procedure could be reduced if an inexpensive yet reliable diagnostic tool were available as an aide for physicians. The software system proposed in this work attempts to provide such a tool.

The software system provides the user with an output classification for an unknown heart sound, and this information could prove useful for a physician to consider when deciding whether or not to refer a patient for echocardiogram. It is expected that future research in noise reduction methods will lead to even better rates of classification.

Future work will emphasize pre-processing steps that will reduce the background noise levels present in the recorded patient sounds. As mentioned previously, wavelet analysis may be explored, and an attempt will be made to train the system using noisy simulated sounds. When the set of available patient sounds is large enough, an attempt will be made to train the system using patient sounds and observe performance. Additionally, research will include a slight change in the type of input vectors used. Results are expected to improve if S1 and S2 are excluded from the input vector, since murmurs will occur over systole and diastole but they will be virtually undetectable during S1 and S2. In any case, it is important to show that the accuracy of the ANN system is high enough using patient sounds, not only simulated sounds, before it can be considered for a real-world application.

Acknowledgments. Thanks to the University of Minnesota Duluth Department of ECE, the University of Minnesota School of Medicine Duluth and St. Luke's Hospital for motivation and providing the means to conduct this research. Also, thanks to the Whiteside Institute for funding these research efforts so generously.

References

[1] T. S. Leung, et al., "Classification of heart sounds using time-frequency methods and artificial neural networks," Proceedings of the 22nd Annual EMBS International Conference, Chicago IL, 988-991, July 23-28,2000.

[2] I. Hartimo and H. Liang, "A Heart Sound Feature Extraction Algorithm Based on Wavelet Decomposition and Reconstruction," Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 20(3), 1998.

[3] I. Hartimo and H. Liang, "A Feature Extraction Algorithm Based on Wavelet Packet Decomposition for Heart Sound Signals," IEEE, 93-96, 1998.

[4] C. N. Gupta, et al., "Segmentation and Classification of Heart Sounds," CCECE/CCGEI, Saskatoon, IEEE, 1674-77, May 2005.

[5] D. Barschdorff, U. Femmer, and E. Trowitzsch, "Automatic Phonocardiogram Signal Analysis in Infants based on Wavelet Transforms and Artificial Neural Networks," Computers in Cardiology, 753-6, IEEE, 1995.

[6] S. R. Bhatikar, C. DeGroff, and R. L. Mahajan, "A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics," Artificial Intelligence in Medicine, 2005; 33: 251-260.

[7] E. Etchells, C. Bell, and K. Robb, "Does this Patient Have an Abnormal Systolic Murmur," JAMA, 277: 564-571, 1997.

[8] American Society of Echocardiography, "Frequently Asked Questions about Cardiography," www.seemyheart.org/faq.php.

[9] S. Mangione and L.Z. Nieman, "Cardiac Auscultatory Skills of Internal Medicine and Family Practice Trainees: A Comparison of Diagnostic Proficiency," JAMA, 278(9): 717-722, Sept 1997.

[10] R.O. Bonow, et al., "ACC/AHA Task Force Report," JACC, 32(5): 1486-1588, November 1998.

[11] A.R. Freeman and S. A. Levine, "The Clinical Significance of the Systolic Murmur", Annals of Internal Medicine VI (1933): 1371-1385.

[12] M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*, Martin Hagan, January 2002.

[13] K. Ejaz, G. Nordehn, R. Alba-Flores, F. Rios-Gutierrez, S. Burns, and N. Andrisevic, "Heart Murmur Detection System Using Spectrograms and Artificial Neural Networks," International Association of Science and Technology for Development (IASTED) Conference: Circuits, Signals and Systems (CSS 2004), Nov. 28, 2004 – December 1, 2004.

[14] AHA Scientific Statement: AHA Guidelines for Primary Prevention of Cardiovascular Disease and Stroke: 2002 Update, #71-0226 Circulation. 2002; 106: 388-391.

[15] AHA, Heart Disease and Stroke Statistics: 2006 Update. Dallas, TX, American Heart Association, 2003 [Online].

[16] MATLAB Help Documentation: Neural Networks Toolbox Help. MATLAB v6.5.0.1, June 18, 2002.

[17] Lam M. Z. C., et al., "Factors influencing cardiac auscultation proficiency in physician trainees," Singapore Medical Journal, 2005; 25: 11-14.

[18] A. Disher, C. Ornes, and J. Sklansky, "A Visual Neural Network that Learns Perceptual Relationships," IEEE, 873-875, 1998.