

# Heart Sound Classification Based on Nonlinear Time-frequency Features

Aaron Raymond See, Inah Salvador Cabili, and Yeou-Jiunn Chen\*

Department of Electrical Engineering, Southern Taiwan University of Science and Technology,  
Tainan 71005, Taiwan

(Received May 26, 2021; accepted December 2, 2021)

**Keywords:** heart sound classification, Shannon entropy, spectral entropy, support vector machine

Cardiovascular disease (CVD) has been the most common factor of death for decades, and one method to detect CVD is through heart sound auscultation. Numerous studies have investigated improvements in precision and accuracy for heart sound classification using machine learning. Nonetheless, most methods utilize many features in their machine learning to increase the accuracy of their predictive model to address challenges associated with signals acquired through sensors placed at different locations. In this paper, we propose the use of heart sounds segmented into three frequency bands and the extraction of features, namely, the Shannon entropy and spectral entropy of each frequency band, to serve as an input to our support vector machine (SVM). The focus of the study is to examine the use of only six features to achieve a satisfactory score in heart sound classification. The technique is assessed using an online heart sound database. The features that were extracted are trained and tested using the SVM to predict normal and abnormal heart sounds. Results demonstrated accuracies of 95 and 78% for normal and abnormal heart sounds, respectively. Subsequently, the testing results achieved an overall accuracy of 82.5% with a sensitivity of 85% and a specificity of 80%.

## 1. Introduction

Despite encouraging advances in the prevention and treatment of diseases, cardiovascular disease (CVD) continues to be the leading factor of death and disability worldwide. Subsequently, as the incidence in low- and middle-income countries increases, CVD will continue to become more prevalent.<sup>(1)</sup> According to the World Health Organization (WHO), more than 17.9 million deaths worldwide were due to CVD in 2016, representing 31% of all global deaths.<sup>(2)</sup> Owing to the rapid development of modern technology, machine learning algorithms can be used to predict heart sound recordings. To classify heart sounds, we have been searching for abnormal signal features such as those corresponding to mitral stenosis, mitral insufficiency, aortic stenosis, aortic insufficiency, tricuspid stenosis, and tricuspid insufficiency.<sup>(3)</sup> To achieve this, heart sound data were acquired from sensors placed in different locations. A large amount of research on extracting these features of heart sounds has been carried out. Many of these features used different signal processing methods. For example, Goda and Hajas<sup>(3)</sup> used wavelet

---

\*Corresponding author: e-mail: [chenyj@stust.edu.tw](mailto:chenyj@stust.edu.tw)  
<https://doi.org/10.18494/SAM3558>

envelope features and a support vector machine (SVM), attaining an accuracy of 81.2%. Uğuz<sup>(4)</sup> used an artificial neural network (ANN) based on the discrete wavelet transform (DWT) and a fuzzy logic approach, achieving an average accuracy of 98.33%. Springer *et al.*<sup>(5)</sup> employed a hidden semi-Markov model (HSMM) and logistic regression for the accurate segmentation of heart sound recordings. Chen *et al.*<sup>(6)</sup> proposed the usage of Mel-frequency cepstral coefficients (MFCCs) after the segmentation of S1 signals, then applied a deep neural network, achieving an accuracy of 91%. Hamidi *et al.*<sup>(7)</sup> used fractal dimension features with MFCCs to improve the classification of heart sounds, gaining an accuracy of 92%. Because of the many input features used in machine learning algorithms, predictive tasks are challenging to model<sup>(8)</sup> and the computational burden may be high. The aim of this study was to test the feasibility of classifying heart sounds using a small number of features, namely, Shannon entropy and spectral entropy.

## 2. Methodology

Figure 1 illustrates the process flow to classify normal heart sound (NHS) and abnormal heart sound (AHS) signals. First, each heart sound is preprocessed by detrending the signal, then a low-pass Butterworth filter with a frequency range of 0 to 400 Hz is used to remove unwanted signals. Murmurs are abnormal signals that are often heard in the mid- to high-frequency region (120–600 Hz),<sup>(9)</sup> while typical NHS signals such as S1 and S2 are heard in the low-frequency region (0–250 Hz).<sup>(9, 10)</sup> Because of their pitch variations, NHSs and AHSs can be distinguished from their differences in frequency. The frequency band intervals considered in this study are 0–138, 138–276, and 276–400 Hz. A total of 121000 samples of filtered data obtained from heart sound recording datasets are used in this study. Second, two features are extracted from the filtered data, namely, Shannon entropy and spectral entropy, to calculate the similarities and differences of heart sound data. Lastly, the performance of the proposed method in evaluating heart sound abnormality is evaluated using an SVM.

Since feature extraction is the key to extracting the characteristics found in signals, we assess signals in terms of their pattern and the overall behavior based on the selected features, namely, Shannon entropy and spectral entropy. Shannon entropy ( $H$ ) is a measure of randomness and complexity in the time domain, and is defined as

$$H(x) = -\sum_{i=1}^n P(x_i) \cdot \log_2 P(x_i), \quad (1)$$

where  $P(x_i)$  is the probability weighting associated with  $x_i$ . The logarithm base is 2 since it is used in communication or information. Note that  $H$  is non-negative and that the sum of the probabilities must be 1.<sup>(11)</sup>



Fig. 1. (Color online) Flow chart of the proposed method.

Spectral entropy (*SE*) is the power distribution in the frequency domain and is defined as

$$SE = -\sum_{i=1}^N p_i \cdot \log_2 p_i, \quad (2)$$

$$p_i = \frac{X(i)}{\sum_{j=1}^N X(j)}, \quad (3)$$

$$\sum_{i=1}^N p_i = 1, \quad (4)$$

where  $X_i$ ,  $i = 1, 2, \dots, N$  is the Fourier transformation of the signal, and  $p_i$  is the probability distribution.<sup>(11)</sup>

To better understand the distribution of the entropies, the next step is to visualize the features. Thus, we implement t-distributed stochastic neighbor embedding (t-SNE) to visualize how the features of NHSs and AHSs are clustered. t-SNE collects high-dimensional datasets into low-dimensional datasets while maintaining the structure of the high-dimensional data. It models each high-dimensional data into a 2D or 3D point in such a way that similar data are clustered together and dissimilar data are clustered at points distant from each other. The t-SNE algorithm consists of two main steps. First, we compute the input similarities  $x_i$  and  $x_j$  of the given set of high-dimensional data  $N$  to construct a conditional probability  $p_{j|i}$  that similar data are assigned to a higher probability and dissimilar data are assigned to a lower probability. We define the probability as

$$p_{j|i} = \frac{\exp\left(\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)}{\sum_{j' \neq i} \exp\left(\frac{\|x_i - x_{j'}\|^2}{2\sigma_i^2}\right)}, \quad (5)$$

where  $\sigma_i$  is the bandwidth of the Gaussian kernel so that the condition has a fixed perplexity for the model to adapt to different densities. To symmetrize the conditional probability distribution of the high-dimensional space, we use  $p_{ij}$  defined as

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}. \quad (6)$$

Second, t-SNE defines a similar probability distribution of the low-dimensional map points as

$$q_{ij} = \frac{\left(1 + \|y_i - y_j\|^2\right)^{-1}}{\sum_k \sum_{l \neq k} \left(1 + \|y_i - y_l\|^2\right)^{-1}}. \quad (7)$$

To measure the similarities of the high-dimensional space  $p_{ij}$  and the low-dimensional space  $q_{ij}$  so that the structure of the low-dimensional map is the same as that of the original high-dimensional map, we use the Kullback–Leibler divergence defined as<sup>(12)</sup>

$$KL(P \parallel Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}. \quad (8)$$

Several machine learning classifiers such as the SVM can be used for signal processing to classify heart sound signals.<sup>(9)</sup> The SVM has been predominantly used in complex situations because it excels at solving nonlinear learning problems as well as sample learning problems.<sup>(13)</sup> When dealing with a nonlinear algorithm, various types of SVM kernels can handle high-dimensional data such as cubic, fine Gaussian, and medium Gaussian data. The computation in this study is performed on a computer with a Windows 10 operating system and a clock speed of 2.4 GHz. The tool used for the computation is MATLAB version R2019b.

### 3. Experimental Results

The CVD database employed in this study utilized the datasets from PhysioNet – Computing in the Cardiology Challenge 2016, which comprises more than 3000 datasets. The database consisted of a total of 3241 heart sound recordings with a duration from 5 to 121 s and a sampling frequency of 2 kHz. Furthermore, the heart sound recordings were collected from typical locations for heart sound auscultation, including the aortic, pulmonic, tricuspid, and mitral areas. The database has two categories: NHSs from healthy subjects and AHSs from patients with confirmed coronary artery disease.<sup>(14,15)</sup> The data used in the training database were recording samples with durations from 10 to 121 s. Subsequently, each signal was segmented into three bands with intervals of 0 to 138 Hz, 138 to 276 Hz, and 276 to 400 Hz. After segmenting the signals, the Shannon entropy and spectral entropy were applied to extract the features, thus obtaining six features that served as the input into the SVM. Then, 10-fold cross-validation was conducted on the training dataset, where 90% of the dataset was selected at random for training and 10% was reserved for validation. In addition, the medium Gaussian SVM was chosen since it provides a higher accuracy than the cubic and fine Gaussian kernels. The classification results are summarized in Table 1. The results demonstrate that the accuracy of classifying both NHSs and AHSs using the six features was superior to that obtained by extracting only three features. Also, according to the results for the medium Gaussian SVM classifier in Table 1, using the combined features significantly increased the accuracy of AHS classification.

Table 1  
Heart sound classifier accuracy using the proposed method.

Feature type	Feature number	Classifier	Abnormal (%)	Normal (%)
Spectral entropy	3	Cubic SVM	35	73
		Fine Gaussian SVM	35	95
		Medium Gaussian SVM	32	97
Shannon entropy	3	Cubic SVM	59	56
		Fine Gaussian SVM	72	95
		Medium Gaussian SVM	73	94
Spectral entropy + Shannon entropy	6	Cubic SVM	77	94
		Fine Gaussian SVM	48	96
		Medium Gaussian SVM	78	95

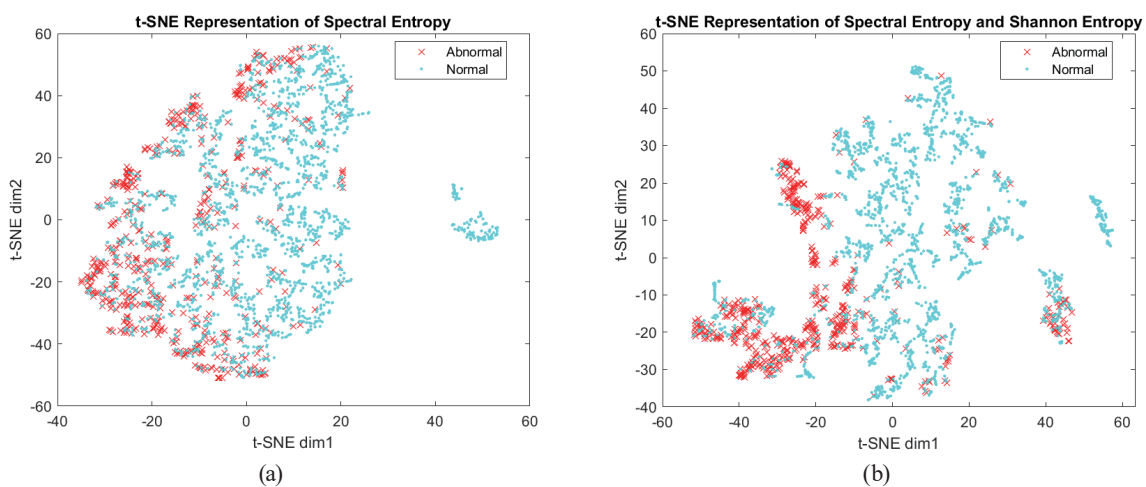


Fig. 2. (Color online) t-SNE plots visualizing cluster assignment of heart sounds with the points representing NHSs and the crosses exhibiting AHSs.

Figure 2 shows a visualization of the features of the NHS and AHS data. The feature graph visualization used t-SNE, which visualized the high-dimensional data as low-dimensional data.<sup>(16)</sup> In Fig. 2(a), only three features from the *SE* of the three frequency bands were used, while six features from both the spectral entropy and Shannon entropy in the three frequency bands were used in Fig. 2(b). t-SNE displays clustering across different heart sound categories and shows the shifting of symbols towards specific clusters. According to the visualization, Fig. 2(b) creates a clearer distinction and better separation between NHSs and AHSs.

Table 2 shows the performance of the trained model. The testing dataset was taken from the validation folder of PhysioNet – Computing in Cardiology Challenge 2016, which consists of more than 3000 test datasets, and the outcome of machine learning was tested using a medium Gaussian SVM classifier. The performance of the proposed algorithm was analyzed by determining the sensitivity, specificity, and overall score of the classification results. These three parameters are essential indicators of the performance of diagnostic tests compared with those in other studies. Recent studies of Goda and Hajas<sup>(3)</sup> and Tang *et al.*<sup>(17)</sup> focused on the detection of NHSs and AHSs using an SVM classifier. The key characteristics of these studies were the use of 25 features, which included combined DWT, time, and frequency features, in Ref. 3, and the

Table 2  
Comparison of performance characteristics with other studies.

Authors	Number of features used	Sensitivity (%)	Specificity (%)	Overall score (%)
Goda and Hajas <sup>(3)</sup>	36	77.2	85.2	81.2
Tang <i>et al.</i> <sup>(17)</sup>	515	88	87	88
Proposed	6	85	80	82.5

use of 515 features, which were extracted from multiple domains and included preprocessing by the ensemble of features, in Ref. 17. From Table 2, it can be observed that the performance of the proposed algorithm using the SVM as the machine learning classifier is comparable to that in the previous studies despite its smaller number of features.

#### 4. Conclusion

Heart sound classification is very helpful and can be used to aid the physician in the early detection and diagnosis of patients suffering from CVD or other heart diseases. Various methods have achieved good performance in classifying heart sound signals. In this paper, we mainly focused on the effect of inputting a smaller number of features, namely, Shannon entropy and spectral entropy, in machine learning for heart sound classification on the classification performance. We successfully implemented a heart sound classification system based on a limited number of nonlinear time-frequency features using an SVM. The parameter performance obtained an overall accuracy of 82.5%, comparable to the results obtained in studies that employed many more features. In addition, the extraction of a smaller number of features will be beneficial for realizing compact systems with a limited number of sensors to assist in remote health monitoring. Furthermore, the heart sound classification approach proposed in this study can be improved by employing more datasets. Additional features can also be introduced to find new ways to realize a fast and efficient heart sound classification system.

#### Acknowledgments

This work was supported by the Ministry of Education of Taiwan under grant number MOE 1300-108P097.

#### References

- 1 S. Bansilal, J. M. Castellano, and V. Fuster: *Int. J. Cardiol.* **201** (2015) S1. [https://doi.org/10.1016/S0167-5273\(15\)31026-3](https://doi.org/10.1016/S0167-5273(15)31026-3)
- 2 Cardiovascular Diseases (Cvds): [https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)) (accessed September 2020).
- 3 M. A. Goda and P. Hajas: *Proc. 2016 Computing in Cardiology Conf. (CinC)* (2016) 1133.
- 4 H. Uğuz: *Neural Comput. Appl.* **21** (2012) 1617. <https://doi.org/10.1007/s00521-011-0610-x>
- 5 D. B. Springer, L. Tarassenko, and G. D. Clifford: *IEEE T. Biomed. Eng.* **63** (2015) 822. <https://doi.org/10.1109/TBME.2015.2475278>
- 6 T.-E. Chen, S.-I. Yang, L.-T. Ho, K.-H. Tsai, Y.-H. Chen, Y.-F. Chang, Y.-H. Lai, S.-S. Wang, Y. Tsao, and C.-C. Wu: *IEEE T. Biomed. Eng.* **64** (2016) 372. <https://doi.org/10.1109/TBME.2016.2559800>
- 7 M. Hamidi, H. Ghassemanian, and M. Imani: *Biomed. Signal Process. Control* **39** (2018) 351. <https://doi.org/10.1016/j.bspc.2017.08.002>

- 8 Introduction to Dimensionality Reduction for Machine Learning: <https://machinelearningmastery.com/dimensionality-reduction-for-machine-learning/> (accessed September 2020).
- 9 A. K. Dwivedi and E. Rodriguez-Villegas: IEEE Access **7** (2018) 8319. <https://doi.org/10.1109/ACCESS.2018.2889437>
- 10 A. L. Padilla-Ortiz and D. Ibarra: Crit. Rev. Biomed. Eng. **46** (2018) 38. <https://doi.org/10.1615/CritRevBiomedEng.2018025112>
- 11 Y. N. Pan, J. Chen, and X. L. Li: (2013). Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci. **223** (2008) 1224. <https://doi.org/10.1243/09544062JMES1224>
- 12 L. V. C. Maaten and G. Hinton: J. Mach. Learn. Res. **9** (2008) 2579.
- 13 P. Lu, D.-P. Xu, and Y.-B. Liu: Proc. 2005 Int. Conf. Mach. Learn. **7** (2005) 4319. <https://doi.org/10.1109/ICMLC.2005.1527697>
- 14 MIT Laboratory for Computational Physiology: Classification of Heart Sound Recordings - the Physionet Computing in Cardiology Challenge 2016, <https://physionet.org/content/challenge-2016/1.0.0/> (accessed September 2020).
- 15 C. Liu, D. Springer, Q. Li, B. Moody, R. A. Juan, F. J. Chorro, F. Castells, J. M. Roig, I. Silva, and A. E. Johnson: Physiol. Meas. **37** (2016) 2181. <https://doi.org/10.1088/0967-3334/37/12/2181>
- 16 A. Violante: An Introduction to T-Sne with Python Example, <https://towardsdatascience.com/an-introduction-to-t-sne-with-python-example-5a3a293108d1> (accessed September 2020).
- 17 H. Tang, Z. Dai, Y. Jiang, T. Li, and C. Liu: Biomed Res. Int. **2018** (2018) 1. <https://doi.org/10.1155/2018/4205027>

## About the Authors

**Aaron Raymond See** is an assistant professor in the Department of Electrical Engineering, Southern Taiwan University of Science and Technology. He concurrently holds the position of Director of Student Affairs in the Office of International Affairs. He was born in Manila, Philippines, and received a B.S. degree in electronics and communications engineering from De La Salle University, Manila, in 2006. He obtained his master's and Ph.D. degrees in electrical engineering with a major in biomedical engineering at Southern Taiwan University of Science and Technology in 2010 and 2014, respectively. His research interests are in assistive device design and development, biomedical image and signal processing, and engineering education.

**Inah Salvador H. Cabili** received her B.S. degree in electronics engineering from Mindanao State University – Iligan Institute of Technology, Iligan City, Philippines, in 2018. She is currently pursuing a master's degree in electrical engineering at Southern Taiwan University of Science and Technology, Taiwan. Her current research interests include biomedical engineering, machine learning, and integrated circuit design.

**Yeou-Jiunn Chen** received his B.S. degree in mathematics from Tatung Institute of Technology, Taipei, Taiwan, and his Ph.D. degree from the Institute of Information Engineering, National Cheng Kung University, Tainan, Taiwan, in 1995 and 2000, respectively. He was with the Advanced Technology Center, Computer and Communications Laboratories, Industrial Technology Research Institute, from 2001 to 2005 as a researcher. He is currently a professor at the Department of Electrical Engineering, Southern Taiwan University of Science and Technology, Tainan, Taiwan. His research interests include biomedical signal processing, spoken language processing, and artificial intelligence. Dr. Chen is a member of the Biomedical Engineering Society, Taiwan Rehabilitation Engineering and Assistive Technology Society, and the Association for Computational Linguistics and Chinese Language Processing.