Less is More: A Novel Feature Extraction Method for Heart Sound Classification via Fractal Transformation

Cuiping Zhu[1,2,3], Zhonghao Zhao[1,2], Yang Tan[1,2], Mengkai Sun[1,2],
Kun Qian*[1,2], *Senior Member, IEEE*, Tao Jiang*[3], Bin Hu*[1,2], *Fellow, IEEE*,
Björn W. Schuller[4,5], *Fellow, IEEE*, and Yoshiharu Yamamoto[6], *Member, IEEE*

***Abstract*— Cardiovascular diseases (CVDs) are the leading cause of death globally. Heart sound signal analysis plays an important role in clinical detection and physical examination of CVDs. In recent years, auxiliary diagnosis technology of CVDs based on the detection of heart sound signals has become a research hotspot. The detection of abnormal heart sounds can provide important clinical information to help doctors diagnose and treat heart disease. We propose a new set of fractal features – fractal dimension (FD) – as the representation for classification and a Support Vector Machine (SVM) as the classification model. The whole process of the method includes cutting heart sounds, feature extraction, and classification of abnormal heart sounds. We compare the classification results of the heart sound waveform (time domain) and the spectrum (frequency domain) based on fractal features. Finally, according to the better classification results, we choose the fractal features that are most conducive for classification to obtain better classification performance. The features we propose outperform the widely used features significantly ($p < .05$ by one-tailed z-test) with a much lower dimension.**

***Clinical relevance*—The heart sound classification model based on fractal provides a new time-frequency analysis method for heart sound signals. A new effective mechanism is proposed to explore the relationship between the heart sound acoustic properties and the pathology of CVDs. As a non-invasive diagnostic method, this work could supply an idea for the preliminary screening of cardiac abnormalities through heart sounds.**

This work was partially supported by the Ministry of Science and Technology of the People's Republic of China with the STI2030-Major Projects (No. 2021ZD0201900), the National Natural Science Foundation of China (No. 62227807 and 62272044), the Teli Young Fellow Program from the Beijing Institute of Technology, China, and the Grants-in-Aid for Scientific Research (No. 20H00569) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. *Corresponding authors*: K. Qian, T. Jiang and B. Hu.

[1,2]Cuiping Zhu, Zhonghao Zhao, Yang Tan, Mengkai Sun, Kun Qian and Bin Hu are with Key Laboratory of Brain Health Intelligent Evaluation and Intervention, Ministry of Education (Beijing Institute of Technology), Beijing 100081, China, and also with the School of Medical Technology, Beijing Institute of Technology, Beijing 100081, China. Cuiping Zhu is also with the School of Physics and Electronic Engineering, Sichuan Normal University, No. 1819, Section 2, Chenglong Avenue, Longquanyi District, Chengdu 610101, China. zhu@stu.sicnu.edu.cn, {zhonghao.zhao, yang_tan, smk, qian, bh}@bit.edu.cn

[3]Tao Jiang is with the School of Physics and Electronic Engineering, Sichuan Normal University, No. 1819, Section 2, Chenglong Avenue, Longquanyi District, Chengdu 610101, China. jiangtao@sicnu.edu.cn

[4,5]Björn W. Schuller is with GLAM – the Group on Language, Audio, & Music, Imperial College London, 180 Queen's Gate, Huxley Bldg., London SW7 2AZ, UK, and also with the Chair of Embedded Intelligence for Health Care and Wellbeing, University of Augsburg, Eichleitnerstr. 30, Augsburg 86159, Germany. schuller@ieee.org

[6]Yoshiharu Yamamoto is with the Educational Physiology Laboratory, Graduate School of Education, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. yamamoto@p.u-tokyo.ac.jp

I. Introduction

Out of the 17 million premature deaths (under the age of 70 years) due to noncommunicable diseases in 2019, 38 % were caused by CVDs [1]. In addition, the high prevalence of CVDs and high medical costs have increased the socio-economic burden, posing huge challenges for developing countries and their families. The initial diagnosis of CVDs can be made by auscultation of heart sounds. However, due to the lack of fully automatic diagnostic tools, cardiac auscultation and interpretation of results depend on the subjectivity and training of the human ear. Meanwhile, the diagnosis of heart sounds requires a physician with years of clinical experience and specialised medical equipment. Therefore, auscultation became a suitable method of heart examination because of its simplicity and low cost widely accepted by the public [2].

There are many artificial intelligence algorithms to classify heart sound. Tang et al. used multidomain features and Support Vector Machines (SVM) for classification of heart sound [3]. Nassralla et al. considered random forests to classify time and frequency features of heart sounds [4]. Alqudah et al. proposed a bispectrum analysis approach and used a convolutional neural network (CNN) to achieve heart sound classification [5]. Almost universally, most of the proposed algorithms segment the recording into characteristic heart sounds S1, S2, related systolic and diastolic intervals. Although this segmentation provides many classification features, which may help to identify abnormal heart sounds, it also brings considerable complexity and increases computational burden to the algorithm. In the heart sound classification model based on fractals, we do not need to consider this complex segmentation process. Fractal is a mathematical structure that shows self-similarity on a series of scales and non-integer dimensions. With these features, fractal geometry can be used to effectively estimate the geometric complexity of the object, as well as the irregularity of the shape and pattern observed in the heart sound recording (changing with space or time). The phonocardiogram (PCG) signal is considered as a fractal signal where the FD is a measure of signal complexity [6]. Therefore, we are inspired to try to extract the FD characteristics of heart sounds from multiple perspectives.

The main contributions of this work can be summarised as follows: Firstly, as far as we know, this is the first study that only uses fractal information as the feature to classify heart sounds. Secondly, we explore the performance of fractal information as a classification feature in the time domain, frequency domain, and time-frequency domain to obtain better classification results. Thirdly, we observe that the fractal feature type performs better in the time domain, which could provide some useful ideas for abnormal heart sound classification from this domain. The rest of this paper will be organised as follows: Firstly, Section II describes the data and methods used in this study. Subsequently, the experimental results are shown in Section III followed by a discussion in Section IV. Finally, this work is concluded in Section V.

II. METHODS

A. Dataset

Our proposed approaches are evaluated on the database of the PhysioNet/CinC Challenge 2016 [7]. The data consists of six independent databases that are non-independent identically distributed (Non-IID) [8]. As the test set labels for this data are not publicly available, we use the training set of the database and split it into a new training/development/test set. There are totally 3 240 heart sound recordings collected from 947 pathological patients and healthy individuals [9]. All recordings were resampled to 2 000 Hz and have been provided in an uncompressed wav format, with recording times ranging from several seconds to minutes [10]. The dataset consists of six sub-databases from different research groups: the MIT, AAD, AUTH, UHA, DLUT, and SUA heart sounds database [9]. A detailed overview of the database is given in Table I.

TABLE I

AN OVERVIEW OF THE DATASET USED.

Dataset	Database	Recordings	Normal	Abnormal
Training set	AUTH	31	7	24
	DLUT	2 141	1 958	183
Total		2 172	1 965	207
Development set	UHA	55	27	28
	AAD	490	386	104
Total		545	413	132
Test set	SUA	114	80	34
	MIT	409	117	292
Total		523	197	326

TABLE II

DIFFERENT FRACTAL FEATURES. FD: FRACTAL DIMENSION.

Name	Feature Description
Fractal (wave)	FD of the original waveform of each audio frame
Fractal (amp)	FD of the waveform amplitude of each audio frame
Fractal (fre)	FD of the instantaneous frequency of each audio frame waveform
Fractal (amp+fre)	FD of the waveform amplitude and the instantaneous frequency of each audio frame

B. Preprocessing

Heart sound is vulnerable to noise interference from the acquisition equipment during the recording process and the heartbeat frequency components are below 420 Hz [11]. In this section, we design a low-pass Butterworth filter with a cut-off frequency of 420 Hz to remove high-frequency noise [12]. Referring to the work of [13], we cut all heart sounds into 5-second segments for subsequent processing and analysis. The main reason is that the heart sound segment contains the complete basic heart sound. Heart sound is also a non-stationary time-varying signal. Therefore, the heart sound signal is divided into a group of frames to analyse its characteristic parameters. We set the frame length to 256, the frame shift to 128, and choose Hamming window as the window function.

C. Fractal

A fractal is a morphological feature that fills the space in the form of a non integer dimension. It is a self-similar pattern [14], which means that it is completely or approximately similar to a part of itself. Heart murmur is an abnormal heart sound. If there is murmur, the PCG signal will be more confused. This situation is very similar to a fractal. The FD is also sensitive to the change of information contained in the signal sequence. Therefore, we try to classify normal and abnormal heart sounds by using FD as the classification feature of heart sounds. There are several measures of FD, we choose the commonly used box-counting method to calculate it [15]. The FD of an object is defined as:

$$FD = \lim_{r \to 0} \frac{\log(N(r))}{\log(1/r)}, \tag{1}$$

where $N(r)$ is the least number of boxes of size r needed to completely cover the fractal object. Fig. 1 shows the FD of two curves.

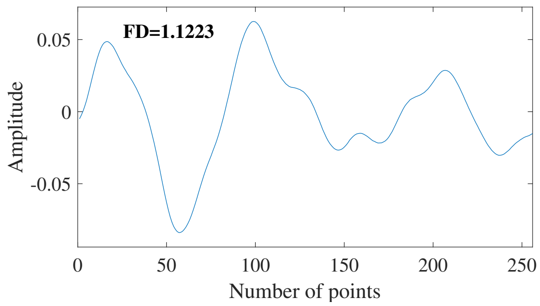
D. Feature Extraction

In this section, we try to extract the FD from the time domain and frequency domain as classification features, explore four kinds of fractal features for heart sound classification, and analyse classification results. To facilitate comparison, we use a baseline model. This model uses the COMPARE feature set (780 features, through functionals) of the widely used OPENSMILE tool, and SVM as the classifier. All FD calculations are implemented using MATLAB.

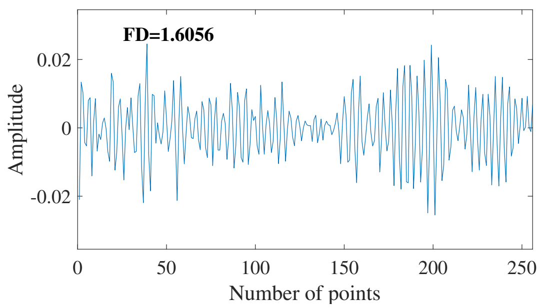
To verify the influence of different fractal features on the classification effect of the heart sound, Table II gives detailed information. Among them, Fractal (amp+fre) is obtained by combining the results of Fractal (amp) and Fractal (fre). The general framework of the proposed method is depicted in Fig. 2.

E. Functionals

When analysing general audio signals, the changes of low-level descriptors (LLDs) in a certain period of time can



(a) A curve with less folding.



(b) A curve with more folding.

Fig. 1. The fractal dimension of two curves with different complexity.

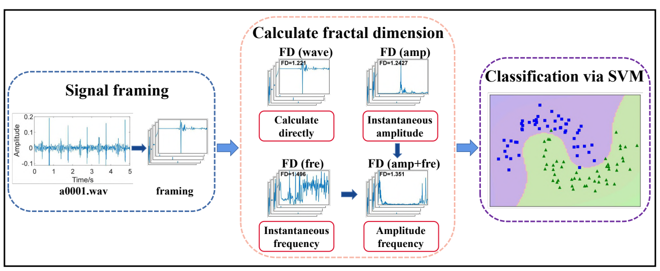


Fig. 2. A frame is used to calculate the fractal dimension of the heart sound in the time and frequency domain. Firstly, After preprocessing, the heart sound signal is divided into frames. Then, the fractal dimension features are calculated based on the frame level. Finally, heart sounds are classified.

provide important information for further model building step. If the number and dimension of LLD vectors are too large, the calculation cost and test time will increase when the analysed audio signal is long. On the contrary, functionals can be applied to the time series of LLDs (frame-level LLDs) to obtain a single fixed dimension vector independent of the input length [16]. The functionals we use include the arithmetic mean, minimum value, maximum value, range, variance, standard deviation, skewness, kurtosis, coefficient of variation, and quartile.

III. Experimental Results

A. Setup

We adopt MATLAB and Sklearn to build our experimental environment. In order to avoid over-optimistic results, we consider the independence of subjects and divide the dataset into train, development, and test set with the proportions of nearly 60 %, 20 %, and 20 %, respectively. The details are shown in Table I. In this paper, feature extraction and classification of all models (including baselines) follow the data division in Table I.

According to the official scoring mechanism of the 2016 PhysioNet/CinC Challenge [7], our model is evaluated by both Sensitivity (Se) and Specificity (Sp). For two-class classification, Se and Sp are defined as:

$$Se = \frac{TP}{TP+FN}, \tag{2}$$

$$Sp = \frac{TN}{TN+FP}, \tag{3}$$

where TP denotes the number of true positive abnormal samples, FN denotes the number of false negative abnormal samples, TN denotes the number of true negative normal samples, and FP denotes the number of false positive normal samples.

Then, precision (Prec) and weighted average recall (WAR) (or accuracy) are used as complementary metrics for evaluating the proposed model's performance, which are defined as:

$$Prec = \frac{TP}{TP+FP}. \tag{4}$$

$$WAR = \frac{TP+TN}{TP+TN+FP+FN}. \tag{5}$$

Finally, considering the impact of the imbalance of the given sample and to render the results more authentic, we add the unweighted average recall rate (UAR) [17], i. e., the average recall rate of each classes, as the main evaluation metric.

TABLE III

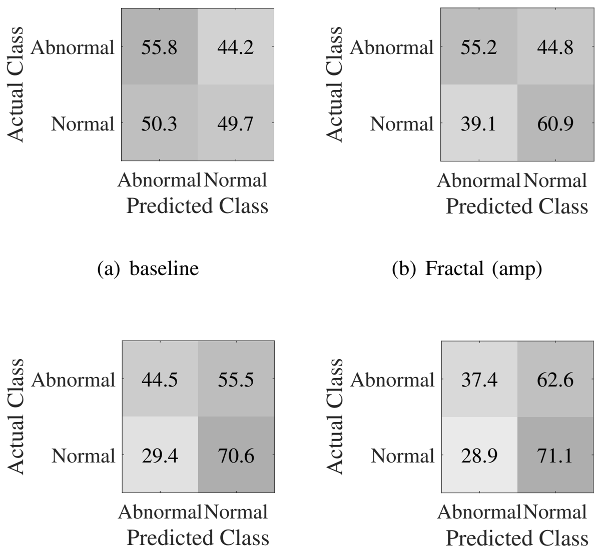
Comparison of different classification features on the test set ([%]), with SVM as classifier (Dim represents the dimension of the feature space).

Feature	Se	Sp	Prec	WAR	UAR	Dim
ComParE (baseline)	55.8	49.7	64.8	53.5	52.8	780
Fractal (wave)	44.5	70.6	71.4	54.3	57.5	12
Fractal (amp)	55.2	60.9	70.0	**57.4**	**58.1**	**12**
Fractal (fre)	24.8	68.0	56.3	41.1	46.4	12
Fractal (amp+fre)	37.4	71.1	68.2	50.1	54.2	24

B. Results

Table III shows the classification performance of different fractal features. According to the results of UAR, the choice of the Fractal (fre) model obtained the best performance, with a UAR of 58.1 %. At the same time, the UAR of the Fractal (amp) model and the Fractal (amp+fre) model both exceed 53.0 %. It is worth noting that among the four fractal based feature methods, the performance in time domain is better than that in the frequency domain.

The normalised confusion matrices of the FD characteristics of time domain and time-frequency domain on the test set are shown in Fig. 3. Among them, Fractal (fre) and Fractal (wave) have high UAR values – 5.3 % and 4.7 % higher than baseline –, respectively, which is significant



(a) baseline

(b) Fractal (amp)

(c) Fractal (wave)

(d) Fractal (amp+fre)

Fig. 3. Normalised confusion matrices of on the test set (in [%]).

($p < .05$ by one-tailed z-test). In addition, the two methods also use a lower feature space dimension than the baseline (12 features equalling, resembling only 1.5 % of the baseline space). Therefore, this feature space appears easier to deploy on hardware.

IV. Discussion

Interestingly, the performance of heart sound classification directly from the perspective of the time domain superseded that of the frequency domain. The experimental result is encouraging and promising. It shows that we can use FD as a feature to provide heart sound classification from the time domain.

The limitation and perspective of this pilot study are: It can be seen from Table III that the performance of FD in time-frequency domain features on the UAR test set is better than the baseline, the performance of the Fractal (fre) model is not as good as the other three models. The first reason for this result could be we solve the instantaneous frequency by Hilbert transform. Due to the latter's limitations [18], we cannot obtain the accurate instantaneous frequency of the heart sound signal. The second reason may be the repetitive (self-similar) patterns characterising the shape of the signal in time domain, and which FD tries to capture, are lost when moving to the frequency domain. In future work, we will study how to combine multifractal analysis with FD to obtain better results. Further, we will investigate FD feature types with others.

V. Conclusion

In this work, we studied the use of fractal feature methods to classify heart sounds. The classification feature based on the FD was introduced into the field of heart sound analysis for the first time, and it was observed that the highest UAR was 58.1 %($p < .05$ by one-tailed z-test) in a strictly subject independent setting. This promising result shows that calculating the FD of the heart sound signal as a classification feature can help for heart sound classification. As a fractal can describe the whole as well as the details of shapes, fractal features have fewer dimensions. The advantage of this method is that feature extraction does not need to divide the recording into characteristic heart sounds and systolic and diastolic intervals. This may significantly reduce the complexity and computational burden of algorithms, and promote their implementation as embedded algorithms in PCG devices.

References

[1] W.H.O.News, "Media centre-cardiovascular diseases (CVDs) fact sheet," http://www.who.int/mediacentre/factsheets/fs317/en/, accessed: 2021-06-11.

[2] K. Qian *et al.*, "Artificial intelligence internet of things for the elderly: From assisted living to health-care monitoring," *IEEE Signal Processing Magazine*, vol. 38, no. 4, pp. 78–88, 2021.

[3] H. Tang *et al.*, "Pcg classification using multidomain features and svm classifier," *BioMed research international*, vol. 2018, 2018.

[4] M. Nassralla, Z. El Zein, and H. Hajj, "Classification of normal and abnormal heart sounds," in *2017 Fourth International Conference on Advances in Biomedical Engineering (ICABME)*. IEEE, 2017, pp. 1–4.

[5] A. M. Alqudah, H. Alquran, and I. A. Qasmieh, "Classification of heart sound short records using bispectrum analysis approach images and deep learning," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 9, no. 1, pp. 1–16, 2020.

[6] M. Hamidi, H. Ghassemian, and M. Imani, "Classification of heart sound signal using curve fitting and fractal dimension," *Biomedical Signal Processing and Control*, vol. 39, pp. 351–359, 2018.

[7] C. Liu *et al.*, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, p. 2181, 2016.

[8] G. D. Clifford *et al.*, "Classification of normal/abnormal heart sound recordings: The physionet/computing in cardiology challenge 2016," in *Proceedings of 2016 Computing in Cardiology Conference (CinC)*. Vancouver, Canada: IEEE, 2016, pp. 609–612.

[9] Z. Ren *et al.*, "Learning image-based representations for heart sound classification," in *Proceedings of the 2018 International Conference on Digital Health*, Lyon, France, 2018, pp. 143–147.

[10] T. Koike *et al.*, "Audio for audio is better? an investigation on transfer learning models for heart sound classification," in *Proceedings of 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. Montreal, Quebec, Canada: IEEE, 2020, pp. 74–77.

[11] S. Debbal and F. Bereksi-Reguig, "Time-frequency analysis of the first and the second heartbeat sounds," *Applied Mathematics and Computation*, vol. 184, no. 2, pp. 1041–1052, 2007.

[12] K. M. Gaikwad and M. S. Chavan, "Removal of high frequency noise from ecg signal using digital iir butterworth filter," in *Proceedings of 2014 IEEE Global Conference on Wireless Computing & Networking (GCWCN)*. Lonavala, India: IEEE, 2014, pp. 121–124.

[13] T. Nilanon *et al.*, "Normal/abnormal heart sound recordings classification using convolutional neural network," in *Proceedings of 2016 Computing in Cardiology Conference (CinC)*. Vancouver, Canada: IEEE, 2016, pp. 585–588.

[14] J. E. Hutchinson, "Fractals and self similarity," *Indiana University Mathematics Journal*, vol. 30, no. 5, pp. 713–747, 1981.

[15] L. S. Liebovitch and T. Toth, "A fast algorithm to determine fractal dimensions by box counting," *Physics Letters A*, vol. 141, no. 8-9, pp. 386–390, 1989.

[16] F. Eyben, *Real-time speech and music classification by large audio feature space extraction*. Springer, 2015.

[17] K. Qian, "Automatic general audio signal classification," Ph.D. dissertation, Technische Universität München, 2018.

[18] J. C. Goswami and A. E. Hoefel, "Algorithms for estimating instantaneous frequency," *Signal Processing*, vol. 84, no. 8, pp. 1423–1427, 2004.