A novel intelligent system based on adjustable classifier models for diagnosing heart sounds

Shuping Sun1✉, Tingting Huang2, Biqiang Zhang2, Peiguang He2, Long Yan2, Dongdong Fan2, Jiale Zhang2 & Jinbo Chen2

A novel intelligent diagnostic system is proposed to diagnose heart sounds (HSs). The innovations of this system are primarily reflected in the automatic segmentation and extraction of the first complex sound ($CS_1$) and second complex sound ($CS_2$); the automatic extraction of the secondary envelope-based diagnostic features $\chi$, $\gamma$, and $\beta$ from $CS_1$ and $CS_2$; and the adjustable classifier models that correspond to the confidence bounds of the Chi-square ($\chi^2$) distribution and are adjusted by the given confidence levels (denoted as $\beta$). The three stages of the proposed system are summarized as follows. In stage 1, the short time modified Hilbert transform (STMHT)-based curve is used to segment and extract $CS_1$ and $CS_2$. In stage 2, the envelopes $CS_{1FE}$ and $CS_{2FE}$ for periods $CS_1$ and $CS_2$ are obtained via a novel method, and the frequency features are automatically extracted from $CS_{1FE}$ and $CS_{2FE}$ by setting different threshold value ($Thv$) lines. Finally, the first three principal components determined based on principal component analysis (PCA) are used as the diagnostic features. In stage 3, a Gaussian mixture model (GMM)-based component objective function $f_{CS}(x)$ is generated. Then, the $\chi^2$ distribution for component $k$ is determined by calculating the Mahalanobis distance from $x$ to the class mean $\mu_k$ for component $k$, and the confidence region of component $k$ is determined by adjusting the optimal confidence level $\beta_k$ and used as the criterion to diagnose HSs. The performance evaluation was validated by sounds from online HS databases and clinical heart databases. The accuracy of the proposed method was compared to the accuracies of other state-of-the-art methods, and the highest classification accuracies of 99.43%, 98.93%, 99.13%, 99.85%, 98.62%, 99.67% and 99.91% in the detection of MR, MS, ASD, NM, AS, AR and VSD sounds were achieved by setting $\beta_k (k = 1, 2, \ldots, 7)$ to 0.87, 0.65, 0.67, 0.65, 0.67, 0.79 and 0.87, respectively.

**Background.** As an efficient method, using heart sound (HS) analysis is often used to evaluate heart function; this approach has been widely used to diagnose heart disease and evaluate heart functions, such as congenital heart disease classification1, ventricular septal defect detection2, blood pressure estimation3 and congenital heart disease screening4, for children and adults. A normal HS is primarily composed of two basic sounds: the first sound ($S_1$) which is generated by the closing of aortic valves and the vibrations associated with tensing of the chordate trendiness and the ventricular walls, the second sound ($S_2$) is produced by the closure of the aortic and pulmonic valves at the beginning of is volumetric ventricular relaxation. However, HSs with unitary murmurs generally occur between $S_1$ and $S_2$ with different noise patterns5. Therefore, analyses of $S_1$, $S_2$, and the period between $S_1$ and $S_2$ play important roles in characterizing HS features with different types of information. Detailed information for $S_1$, $S_2$, and the sounds between $S_1$ and $S_2$ can be used to accurately classify HS. Additionally, to avoid analyzing the sounds between $S_1$ and $S_2$, which are generally segmented from HSs with low accuracy, $S_1$ and part of the period between $S_1$ and $S_2$ are integrated to obtain $CS_1$, and $S_2$ and the part of the period between $S_1$ and $S_2$ are integrated to form $CS_2$. Then, the features are efficiently extracted from $CS_1$ and $CS_2$. Finally, a classification method is established to diagnose heart diseases.

1School of Information Science and Engineering, Hunan Institute of Science and Technology, 414006 Yueyang, China. 2Department of Information Engineering, Nanyang Institute of Technology, 473004 Nanyang, China. *Correspondence: Shuping.Sun@IEEE.org
Need for research.

- **CS1 and CS2 extraction** The studies regarding HS segmentation can be summarized into two branches: one branch includes studies that segment each cardiac cycle into a sequence of four heart stages: \( S_1 \rightarrow \text{Systole period} \rightarrow S_2 \rightarrow \text{Diastole period} \). As a result, the four fundamental stages to be segmented are different due to the nonstationary nature of an abnormal HS signal and the effect of background noise. The other branch includes studies that segment a periodic HS into a sequence of two heart stages, which are expressed as \( \gamma_1 \rightarrow \gamma_2 \). The remainder of this paper is organized as follows. Section “Methodology” presents the approach for determining the diagnostic features and the features are determined based on PCA; and (5) the confidence region of the \( \chi^2 \) distribution, which are adjusted based on the desired \( \beta \), is determined and used as the classification criterion for diagnosing a given HS. The other branch includes studies that segment an abnormal HS into a sequence of two heart stages, and \( S_1 \rightarrow \text{CS1} \rightarrow S_2 \rightarrow \text{CS2} \). Moreover, study noted that the use of frequency features was more efficient in distinguishing normal from abnormal sounds than was the use of time features. Therefore, an efficient frequency feature extraction method should be developed.

- **Feature extraction** As an important component of efficient feature extraction, the frequency width of the envelope over a given threshold value (Thv) has been verified to be useful for detecting heart diseases. However, for many types of HSs, it is difficult to extract frequency widths with unsuitable Thv values due to the existence of a non smooth envelope. To extract the frequency widths for a smooth envelope without setting different Thv values, the smooth envelope can be treated as a secondary envelope, as proposed in, and used to automatically extract the frequency feature matrix based on the STMHT technique; this method was successfully applied to detect different types of heart diseases. However, for mitral stenosis and mitral regurgitation noises, the feature matrix was not easily extracted because the second frequency component component component was missing. Therefore, to improve the classification accuracy for diagnosing different types of heart disease and simplify the complexity of the diagnostic method, the smooth envelopes for CS1 and CS2 extraction in the frequency domain must be considered; additionally, more frequency widths corresponding to different Thv values should be used, and dimensionality reduction should be employed to reduce the number of features considered. Such a classification method could be applied in the efficient extraction of features for diagnosing heart diseases.

- **Classifier model** Gaussian mixture models (GMMs) have been used in a wide variety of clustering applications due to their powerful mathematical characteristics. Confidence regions are used to diagnose the detection data in GMMs, and the optimal confidence regions is determined based on Mahalanobis distance following the Chi-square (\( \chi^2 \)) distribution. Thus, classifier models with adjustable sizes corresponding to the confidence bounds of the Chi-square (\( \chi^2 \)) distribution, which can be adjusted by changing the desired confidence level (denoted as \( \beta \)), are proposed. The \( \chi^2 \) confidence bounds used as the classification criteria are employed to diagnose heart diseases.

Major contributions and organization. In summary, this study proposes an innovative and intelligent system. The major contributions in this study are (1) the STMHT-based CS1 and CS2 are automatically located and extracted; (2) a novel method for obtaining the secondary curves of CS1 and CS2 are extracted in the frequency domain; (3) frequency features are automatically extracted over the given threshold value; (4) the diagnostic features \( \gamma_1, \gamma_2 \) and \( \gamma_3 \) are determined based on PCA; and (5) the confidence region of the \( \chi^2 \) distribution, which are adjusted based on the desired \( \beta \), is determined and used as the classification criterion for diagnosing a given HS. The remaining paper is organized as follows. Section “Methodology” presents the approach for determining the diagnostic features \( \gamma_1, \gamma_2, \gamma_3 \), and a definition of the confidence region-based diagnostic method for diagnosing heart diseases. In “Performance evaluation” section, the performance of the proposed method is compared that of other efficient methods for diagnosing heart diseases. In “Conclusion” section, the conclusions are provided. Finally, the future study is pointed out in “Future study”.

Methodology

This study was approved by the ethics committee of Nanyang Institute of Technology (Approval Number:2016-06) and the informed consent was waived by the ethics committee of Nanyang Institute of Technology. The present study was also conducted in accordance with the tenets of the 1975 Declaration of Helsinki, as revised in 2008.

The flow chart of the proposed intelligent system, shown in Fig. 1, consists of three stages: the automatic location and extraction of CS1 and CS2; the automatic determination of frequency features \( \gamma_1, \gamma_2 \) and \( \gamma_3 \); and the establishment of the Mahalanobis distance criterion-based diagnostic method. In stage 1, the STMHT-based curve (denoted as \( \text{HS}_{\text{STMHT}} \)), which is extracted for the HS envelope generated by the HS, is used to segment and extract CS1 and CS2 from the HS (Fig. 1A). In stage 2, the envelopes CS1E,F and CS2E,F for every period CS1 and CS2 are obtained via a novel method, and the frequency features are automatically extracted from CS1E,F and CS2E,F by setting different Thv lines. Finally, the first three principal components, \( \gamma_1, \gamma_2 \) and \( \gamma_3 \), which express 86.7% of the 1F information, are determined and used as diagnostic features (Fig. 1B, C). In stage 3, the GMM-based mixed classification objective function \( f_k(x) \) which combines component \( k \) with respect to the parameters \( \pi_k, \mu_k, \) and \( \Sigma_k \) and the features \( x = [\gamma_1, \gamma_2, \gamma_3] \), is generated. Then, the \( \chi^2 \) distribution for component \( k \) is determined by calculating the Mahalanobis distance from \( x \) to the class mean \( \mu_k \) of component \( k \), and the adjustable confidence bound (denoted as MDC_k shown in Fig. 1E) is determined to diagnose heart diseases.
Figure 1. Flow chart of the proposed methodology.
Stage 1: Automatic extraction of $CS_1$ and $CS_2$. As shown in Fig. 2, five steps consisting of heart sound auscultation, heart sound preprocessing, heart sound envelope extraction, STMHT extraction, and $CS_1$ and $CS_2$ extraction, which is used to construct the procedure of the $CS_1$ and $CS_2$ extraction and is detailed in the following steps.

Step A: Heart sound auscultation. Auscultation is performed for the purposes of examination cardiovascular. As described in previous study, the original heart sound, denoted as $ST$ (colored in blue line as shown in Fig. 2), are collected by 3M-3200 electronic stethoscope with a $FS = 44.1\, kHz$ sample rate which is widely used by many doctors and produced by American 3M company, and the tricuspid area is selected as the auscultation area due to the tricuspid area reported to supply more important information. Meanwhile, you can hear the sounds when auscultating heart sounds, ensuring that we avoid as much environmental noise as possible during the auscultation procedure. Even so, the collected heart sounds still need to be preprocessing for canceling the invalid components.

Step B: WD-based heart sound preprocessing. HSs are reported to be primarily dispersed in the frequency range of $20\sim700\, Hz$. Therefore, according to the sampling frequency ($FS = 44.1\, kHz$), WD-based HSs are filtered to obtain the efficient frequency components ($21.5\sim689\, Hz$). The Daubechies wavelet $d_B^{10}$ has been used to give the maximum signal-to-noise ratio and minimum root-mean-square error for HSs. Therefore, $d_B^{10}$ is selected for use as the mother wavelet for preprocessing HSs. A filtered and normalized sound, colored by gray and denoted as $HTS_T$, is shown in Fig. 2.

Step C: heart sound envelope ($HS_E$) extraction. The Viola integral-based envelope, denoted as $HS_E$, is extracted from the heart sound $HTS_T$, as reported in studies; this envelope can effective overcome amplitude variations and complex backgrounds and noise. This concept is described as follows. Consider a filtered sound $HTS_T[m]$ for $m = 0, 1, \ldots, M - 1$, where $M$ denotes the number of HSs. In a $W_m$ neighborhood of time $m$, called the width $W_m$ time scale, the $M$-point envelope $HS_E[m]$ is obtained by Eq. (1):

$$HS_E[m] = \frac{1}{2W_m + 1} \sum_{k=m-W_m}^{m+W_m} (HS_T[k] - \overline{HS_T}[m])^2, \quad m = W_m, W_m + 1, \ldots M - 1 - W_m, \quad (1)$$

where

$$\overline{HS_T}[m] = \frac{1}{2W_m + 1} \sum_{k=m-W_m}^{m+W_m} HS_T[k], \quad (2)$$

$W_m = 2205$ if the duration of $CS_1$ or $CS_2$ greater than 0.13 s. Finally, normalization is performed by setting the maximum amplitude of $HS_E$ to 1 (Fig. 2).

Step D: STMHT extraction for HS. Given an $M$-point HS, the STMHT for the HSs, $HS_{STMHT}$, is computed from Eq. (3)

$$HS_{STMHT}[n] = \sum_{m=n-W_m}^{m+n-W_m} HS_E[m] W_N[m-n] W_E \left[ m - \left( n - \frac{N-1}{2} \right) \right], \quad (3)$$

Figure 2. Flow chart of the $CS_1$ and $CS_2$ extraction.
Figure 3. The automatic extraction procedures for CS1 and CS2. A-B show the procedure for an example of a typical AR from the database in study25. C-D show the procedure for an example of a typical normal sound database27.

\[ W_E[i] = \begin{cases} \cos\left(\frac{n \pi}{N}\right) - \cos\left(\frac{n+1 \pi}{N}\right) & \text{for } i = 0, 1, \ldots, N - 1, \\ 0 & \text{for } i = \frac{N-1}{2}, \frac{N-3}{2}, \ldots, \frac{N-1}{2}. \end{cases} \]  

where \( n = (N - 1)/2, \ldots, M = 1 - (N - 1)/2, \) and \( W_E[i] = -(N - 1)/2, \ldots, (N - 1)/2 \) is a moving window of odd length \( N \). According to studies23, the length \( N \) is set to 44101.

**Step E: Automatic extraction of CS1 and CS2.** The characteristics of HSSTMHT considered in studies24,8, as shown in Fig. 3A, C, are summarized as follows: ① The negative-to-positive (N2P) points of HSSTMHT, denoted by ▲, correspond to the geometry center peaks of S1 and S2; ② The geometry center between S1 and S2, denoted by ▼ is determined by the positive-to-negative P2N points of HSSTMHT. Moreover, the interval from S2 to S1 is generally greater than that from S1 to S2 in one period of an HS23,24. Therefore, the N2P and P2N-based CS1 and CS2 features can be automatically segmented from one period of an HS and extracted by two procedures, as described as follows.

1. **N2P and P2N location**

The algorithm for detecting N2P and P2N is detailed as follows.

① First, the signum function of HSSTMHT, denoted as \( S_{HSSTMHT} \), is calculated by

\[ S_{HSSTMHT} = \begin{cases} -1 & \text{if } HSSTMHT < 0, \\ 0 & \text{if } HSSTMHT = 0, \\ +1 & \text{if } HSSTMHT > 0. \end{cases} \]  

(5)

② Then, the variation in \( S_{HSSTMHT} \) (DSHSSTMHT) is determined from Eq. (6)

\[ D_{SHSSTMHT}[i] = S_{HSSTMHT}[i + 1] - S_{HSSTMHT}[i], i = 1, 2, \ldots, n. \]  

(6)

③ Finally, N2P and P2N are determined by

\[ \begin{cases} \text{N2P} = i/F_s & \text{if } D_{SHSSTMHT}[i] = +2 \\ \text{P2N} = i/F_s & \text{if } D_{SHSSTMHT}[i] = -2 \end{cases} \]  

(7)

2. **Automatic extraction of CS1 and CS2**

① Calculate the difference between two adjacent N2Ps, denoted as \( D_{N2P} \), with Eq. (8)

\[ D_{N2P}[i] = \text{N2P}[i + 1] - \text{N2P}[i], i = 1, 2, \ldots, n. \]  

(8)

② Determine the points CS12 and CS21 that are used for segmentation from CS1 to CS2 and from CS2 to CS1, respectively, by using Eq. (9).

\[ \begin{cases} \text{CS12}[i] = \text{P2N}[i], \text{CS12}[i] = \text{P2N}[i + 1], & \text{if } D_{N2P}[i] < D_{N2P}[i + 1] \\ \text{CS21}[i] = \text{P2N}[i + 1], \text{CS21}[i] = \text{P2N}[i + 2], & \text{Otherwise} \end{cases} \]  

(9)

③ Extract CS1 (denoted as CS1_i) and CS2 (denoted as CS2_i) for the ith period of an HS as follows
Figure 4. Example of feature definition and automatic extraction.

The automatic extraction procedures for CS1 and CS2 are illustrated in Fig. 3. Figure 3(A, B) show a typical AR sound, and the typical NM sound is shown in Fig. 3(C, D).

Stage 2: Automatic feature generation. Feature definition. To extract the efficient frequency widths, as shown in Fig. 4, the smooth envelopes for CS1 and CS2 in the frequency domain are firstly generated, and then the frequency widths corresponding to different Thv values are extracted.

1) Secondary envelopes CS1F_H and CS2F_H generation: Given an M-point HS, the secondary envelope in the frequency-domain, denoted as HSF_j[k], can be calculated from Eq. (11):

\[
CSF_H[k] = \frac{\sum_i (L_1 + L_2 + 1 - |l_i|)HSF_i[k + l_i] - \sum_i HSF_i[k + l_i]}{(2L_1 + 1)(2L_2 + 1)},
\]

where \(l_1, l_2\) and HSF[k] are defined by Eq. (12):

\[
\begin{align*}
& l_1 = -(L_1 + L_2), -(L_1 + L_2) + 1, \ldots , (L_1 + L_2) \\
& l_2 = -(L_1 - L_2 - 1), -(L_1 - L_2 - 1) + 1, \ldots , (L_1 - L_2 - 1) \\
& HSF[k] = \sum_{m=0}^{M-1} HSF[m + 1]e^{-j \frac{2\pi m}{2M}}, k = 0, 1, 2, \ldots , M - 1
\end{align*}
\]

| \(i\) | absolute value sign, \(2L_1 + 1\) is the first window width, and \(2L_2 + 1\) is the second window width. According to studies\(^2,^8\), \(L_1\) and \(L_2\) are set to 9 and 17, respectively. Moreover, \(HSF[k]\) is also normalized by setting the maximum amplitude of \(HSE\) to 1. The secondary envelopes for CS1 and CS2, denoted as CS1F_H and CS2F_H respectively, are illustrated by using the examples described in Fig. 3 which are first automatically generated based on Eq. (11), are shown in Fig. 4, where the plots in Fig. 4A.2 describe the results of CS1F_H corresponding to CS1 in Fig. 4A.1, the plots in Fig. 4B.2 describe the results of CS1F_H which corresponds to CS1 in Fig. 4B.1, and the plots in Fig. 4A.3 describe the results of CS2F_H corresponding to CS2 in Fig. 4A.1, and the plots in Fig. 4B.3 describe the results of CS2F_H which corresponds to CS2 in Fig. 4B.1.

2) Definition and automatic extraction of frequency features: The frequency features are illustrated in Fig. 4(B, C), and their gravities are calculated by

\[
\begin{align*}
CS1G &= \frac{\sum_{k=0}^{M-1} k \times CS1F_H[k]}{\sum_{k=0}^{M-1} CS1F_H[k]} \\
CS2G &= \frac{\sum_{k=0}^{M-1} k \times CS2F_H[k]}{\sum_{k=0}^{M-1} CS2F_H[k]}
\end{align*}
\]

The frequency widths over a given threshold value are defined and calculated by

\[
\begin{align*}
CS1FWi &= R_{pi} - L_{pi}, i = 1, 2, 3 \\
CS2FWi &= R_{pi} - L_{pi}
\end{align*}
\]

where \(L_{pi}\) and \(R_{pi}\) are the \(i\)th left and right intersections, respectively, of CS1F_H and CS2F_H over the Thv lines (Thv = 0.3, 0.5 and 0.8). Moreover, the frequency features are expressed based on Eq. (15) and described in Table 1.

\[
FF = \{CS1FW1, CS1FW2, CS1FW3, CS1G, CS2FW1, CS2FW2, CS2FW3, CS2G\}
\]
Experimental results for several typical types of heart disease. The features FF of six typical and normal sounds are illustrated in Fig. 5. From Fig. 5, CS₁ and CS₂ are first automatically located and extracted, then, the envelopes for every CS₁ and CS₂ are extracted by Eq. (11). Finally, the features defined by Eq. (15) for CS₁ and CS₂ in the frequency domain are automatically extracted with Eqs. (13-14). The experimental sounds are 665-period AR sounds (3M database, medical sound library, heart auscultation sounds, continuing medical implementation, sounds Database of the University of Dundee, and patients only with AR disease from the Nanyang First People's Hospital), 381-period AS sounds (continuing medical implementation, sounds database of the University of Dundee, 3M database, medical sound library, auscultation sound, and patients only with AS disease from the Nanyang first People's Hospital, and heart auscultation sounds), 315-period ASD sounds (Medical sound library, heart auscultation sounds, 3M database, patients only with ASD disease from the Nanyang First People's Hospital, and medical sound library), 769-period MR sounds (3M database, sounds database of the University of Dundee, heart auscultation sounds, medical sound library, and auscultation sound), 439-period MS sounds (3M database, auscultation sound, medical sound library, and continuing medical implementation), and 1056-period NM sounds (3M database, Michigan database, medical sound library, ThinkLabs database, and healthy undergraduates from Nanyang Institute of Technology, China). Moreover, the boxplots for the features are plotted in Fig. 6, where Fig. 6A shows the features extracted from CS₁ and Fig. 6B shows the features from CS₂ for each type of heart disease. The scatter plots of features in Fig. 6 illustrate the discrimination ability of the model in distinguishing among different heart diseases and highlighting the following findings: 1) The MS and VSD sounds are easy to distinguish from the other sounds by using CS₁FW₁ (Fig. 6A), and by using the CS₂FW₁ (Fig. 6B), the VSD sound is easy to distinguish from the other sounds; 2) The MS sound is easy to distinguish from the other sounds based on CS₁FW₂ (Fig. 6A), and by using the CS₂FW₂ (Fig. 6B), the AR and VSD sounds

| Feature index | Feature's symbol | Feature description | Unit |
|---------------|------------------|---------------------|------|
| 1             | CS₁FW₁           | The frequency width of CS₁ corresponding to Thv = 0.3 Hz |      |
| 2             | CS₁FW₂           | The frequency width of CS₁ corresponding to Thv = 0.5 Hz |      |
| 3             | CS₁FW₃           | The frequency width of CS₁ corresponding to Thv = 0.8 Hz |      |
| 4             | CS₁G             | The center of gravity of CS₁ in frequency-domain | Hz   |
| 5             | CS₂FW₁           | The frequency width of CS₂ corresponding to Thv = 0.3 Hz |      |
| 6             | CS₂FW₂           | The frequency width of CS₂ corresponding to Thv = 0.5 Hz |      |
| 7             | CS₂FW₃           | The frequency width of CS₂ corresponding to Thv = 0.8 Hz |      |
| 8             | CS₂G             | The center of gravity of CS₂ in frequency-domain | Hz   |

Table 1. Description of the frequency domain feature matrix FF.
are distinguished from other sounds; (3) The NM sound is easy to distinguish from other sounds using CS1FW3 (Fig. 6A). (4) The AR and VSD sounds are easy to distinguish from the other sounds using CS2FW3, as shown in Fig. 6B; (5) Fig. 6A indicates that CS1G can be used to easily distinguish MR from other sounds and the AS and ASD sounds from other sounds; (6) Fig. 6B shows that the distribution of CS2G from AS sounds is different from that for other sounds, except NM sounds. The analysis results discussed above indicate that different combinations of several features defined by Eq. (15) can be used to distinguish among various types of heart disease. Therefore, to simplify features and develop a diagnostic method that is simple and effective, dimension reduction is used to determine new features; this process is described in detail as follows.

**Algorithm 1: principal components analysis on the data set FF**

1. Calculate the standardized features $Z$

   $$Z = \frac{\text{FF} - \mu_{\text{FF}}}{\sigma_{\text{FF}}}$$

   where $\mu_{\text{FF}}$ and $\sigma_{\text{FF}}$ are the mean and standard deviation along each column of FF, respectively.

2. Calculate the covariance matrix $\Sigma = Z^T Z$

3. Compute the eigenvalues $\lambda_i$ and eigenvectors $\xi_i$

   $$\Sigma \xi_i = \lambda_i \xi_i (i = 1, 2, \ldots, p),$$

   where $p$ is the number of features.

4. Estimating high-valued Eigenvectors

   (1). Arrange all the Eigenvalues ($\lambda_i$) in descending order

   (2). Choose a threshold value, $\theta$

   (3). Number of high-valued $\lambda_i$ can be chosen so as to satisfy the relationship $(\sum_{i=1}^m \lambda_i) \geq \theta$, where $m$ is the number of high valued $\lambda_i$ chosen.

   (4). Select eigenvectors corresponding to selected high valued $\lambda_i (i = 1, 2, \ldots, m)$.

5. Extract low dimensional feature vectors (principal components) from raw feature matrix.

   $$\gamma_i = \frac{\text{FF} - \mu_{\text{FF}}}{\sigma_{\text{FF}}} \xi_i$$
Diagnostic feature determination. To simplify the computation when using features to diagnose heart diseases, PCA, a linear dimensionality reduction technique for finding principal components and replacing high-dimensional data in many studies, such as studies on heart arrhythmias classification\textsuperscript{35}, heart disease classification\textsuperscript{2,26}, emotion recognition\textsuperscript{33}, respiratory rate extraction\textsuperscript{28} and electrocardiogram heart disease diagnosis\textsuperscript{39}, is employed to generate a few efficient principal components to characterize HS features and diagnose heart diseases. The algorithm corresponding to the generation of new features via PCA for a given data set FF is described as Algorithm 1. The eigenvector $\hat{e}_i$ in Algorithm 1, which corresponds to the eigenvalue $\lambda_i$ and is calculated for the matrix $Z$ in step 2, as shown in Table 3, is the actual weighted coefficient for the $i$th principal component $\gamma_i$. Table 3 shows that the largest absolute coefficients in the first principal component $\gamma_1$ are $CS_{1FW1}$, $CS_{2FW2}$ and $CS_{1FW2}$; the second principal component $\gamma_2$ is mainly weighted based on $CS_{1G}$, $CS_{2FW2}$, $CS_{1FW3}$ and $CS_{2FW3}$ and the third component $\gamma_3$ is mainly weighted based on $CS_{1FW2}$, $CS_{1FW3}$ and $CS_{2FW3}$ (Table 3). To determine the smallest number of principal components $m$ should be considered, the Pareto chart is used; this chart provides a tool for visualizing the Pareto principle, which states that observing a small set of variables that influence a common outcome is more common than detecting many variables that influence the same outcome. This approach has been used to determine the percent variability explained by each principal component (Fig. 7A). Therefore, according to the smallest $m$ value such that $\eta_m > 80\%\textsuperscript{80}$, combined with the scatter plot for the first $m$ principal components, the smallest $m$ is determined. The Pareto chart of the PCA results in Fig. 7A shows the explained variance and accumulated variance for each principal component $\gamma_i$, where $i = 1, 2, 3, \ldots, 8$. According to Fig. 7A, 67.58% of the total variance is captured by the first two components, $\gamma_1$ and $\gamma_2$, and 86.73% of the total variance is captured by the first three components $\gamma_1$, $\gamma_2$ and $\gamma_3$. Therefore, the following conclusions can be obtained.

- $\gamma_1$ and $\gamma_2$ lead to a dimensionality reduction of 75% (from 8 to 2 variables) and only 32.42% information loss. The scatter diagram of $\gamma_1$ and $\gamma_2$ given in Fig. 7B indicates that although the distribution region corresponding to each type of heart disease is obviously different and the overlaps between MR and other diseases, AR and other diseases, and VSD and other diseases are small, the overlaps among MS, ASD, NM, and AS are relatively large; therefore, it is difficult to accurately distinguish among these four types of heart diseases.
- However, the scatter diagram of $\gamma_1$, $\gamma_2$ and $\gamma_3$, plotted in Fig. 7C, shows that there are different distribution regions for these types of heart diseases. In addition, $\eta(3) = 86.73\%$, as shown in Fig. 7A, based on feature number determination\textsuperscript{80}. Thus, $\gamma_1$, $\gamma_2$ and $\gamma_3$ lead to a dimensionality reduction of 62.5% (from 8 to 3 variables) with only 13.27% information loss. The scatter diagram of $\gamma_1$, $\gamma_2$ and $\gamma_3$ in Fig. 7C is used to verify the different distribution regions corresponding to these types of heart diseases.

Therefore, $m$ is set to 3, and the new 3-dimensional feature matrices consisting of $\gamma_1$, $\gamma_2$ and $\gamma_3$ (see Fig. 7C) are used to diagnose heart diseases.

Stage 3: classification based on the squared Mahalanobis distance criterion. Classifier determination. The squared Mahalanobis distance classification criterion-based methodology, consisting of the five sequential steps as shown in the flow chart (Fig. 8A), is proposed to diagnose HSs and is described in the following 5 steps.

Step 1: GMM-based $\mu_k$ and $\Sigma_k$ generation. In the design step of GMM, the estimated target function, $f_d(x)$, is a mixture of $d$-dimensional normal Gaussian distributions $p(x|\mu_k, \Sigma_k)$ that reflect the training pattern of each component; it is assumed that components can be modeled by mixtures of normal Gaussian distributions by

$$f_d(x) = \sum_{k=1}^{K} \pi_k p(x|\mu_k, \Sigma_k),$$

(19)

where

$$p(x) = \frac{1}{\sqrt{(2\pi)^d | \Sigma_k |}} \exp(-\frac{1}{2} (x-\mu_k)^T \Sigma_k^{-1} (x-\mu_k))$$

(20)

expresses the posterior probabilities corresponding to each component; $K$ is the number of components; $\pi_k$ corresponds to the mixed weights, such that $\sum_{k=1}^{K} \pi_k = 1$; and $\mu_k$ and $\Sigma_k$ are the mean value and covariance matrix of the $k$th component, respectively. Because the goal is to maximize the function $f_d(x)$, the parameters ($\pi_k$, $\mu_k$ and $\Sigma_k$) are determined based on the EM algorithm\textsuperscript{80} for a set of sample records. Based on the types of heart disease described in Sect. 2.2 and the scatter diagram plotted in Fig. 7C, the number of Gaussian mixture components is set to $K = 7$, and the \textit{fitgmdist} function in MATLAB 2018b is used to return a GMM with $K = 7$ components fitted to the features $[\gamma_1, \gamma_2, \gamma_3]$ established in Sect. 2.2 using the EM algorithm by assigning a posterior probability to each component density with respect to each observation. Furthermore, the regularization value is set as 0.01 to avoid ill-conditioned covariance estimates, and the number of optimization iterations is set to 1000 based on experience. The Gaussian mixture parameter estimates for $\pi_k$, $\mu_k$, and $\Sigma_k$ are obtained and shown in Table 4. To characterize the 3-dimensional interspace corresponding to each 3-dimensional Gaussian component for diagnosing heart diseases, the 3-dimensional interspaces can be used as 3-dimensional classifiers to diagnose heart diseases with high classification accuracy; the overlapping interspace between two random
components is made as small as possible, and the independent 3-dimensional interspace corresponding to each component is considered.

Step 2: $\chi^2$ determination for the $k$th component in 3-dimensional interspace. Since the squared Mahalanobis distances for each Gaussian component follow the Chi-square distribution ($\chi^2$) in 3-dimensional interspace, to determine the decision region for classifying the test data $x$ via the components estimated in the above step, the squared Mahalanobis distance in 3-dimensional interspace for the $k$th component with mean $\mu_k$ and full covariance matrix $\Sigma_k$, $d^2(x|\mu_k, \Sigma_k)$, is computed as follows:

### Table 2. Mean ($\mu_{\text{FF}}$) and standard deviation ($\sigma_{\text{FF}}$) of the features.

| Features from CS1FW1 | Features from CS1FW2 | Features from CS1FW3 | Features from CS1G | Features from CS2FW1 | Features from CS2FW2 | Features from CS2FW3 | Features from CS2G |
|----------------------|----------------------|----------------------|-------------------|----------------------|----------------------|----------------------|-------------------|
| $\mu_{\text{FF}} \pm \sigma_{\text{FF}}$ | 45.3 $\pm$ 11.8 | 33.1 $\pm$ 5.8 | 18.8 $\pm$ 3.6 | 80.6 $\pm$ 21.7 | 44.1 $\pm$ 23.1 | 32.2 $\pm$ 9.1 | 18.4 $\pm$ 6.5 | 79.8 $\pm$ 18.9 |

### Table 3. Eigenvector $\xi_i$ and eigenvalue $\lambda_i$ ($i = 1, \ldots, 8$) for $\Sigma$ in descending order of eigenvalues.

| Features | Eigenvector ($\xi_i$ in descending order of eigenvalues) | Eigenvalue ($\lambda_i$) |
|----------|----------------------------------------------------------------|--------------------------|
| S1_FW1   | 0.4309 $\xi_1 = 3.6423$ | 0.2169 $\xi_2 = 1.7643$ | 0.0818 $\xi_3 = 1.5316$ | 0.2179 $\xi_4 = 0.4671$ | 0.8062 $\xi_5 = 0.3226$ | 0.2394 $\xi_6 = 0.0606$ | 0.0616 $\xi_7 = 0.0368$ | 0.0616 |
| S1_FW2   | 0.3716 0.0757 0.5303 0.0431 0.0306 0.6778 0.2873 0.1735 |
| S1_FW3   | 0.3411 -0.0983 0.5431 0.0444 -0.4364 -0.6126 -0.1001 -0.0375 |
| S1_G     | 0.2385 0.6501 -0.0157 -0.2122 0.0212 0.0986 -0.5500 -0.4031 |
| S2_FW1   | 0.3313 0.0487 -0.2471 0.8958 0.0977 -0.0299 -0.0699 -0.0944 |
| S2_FW2   | 0.4475 -0.2390 -0.2924 -0.1628 -0.3044 0.2633 -0.3928 0.5607 |
| S2_FW3   | 0.3517 -0.2191 -0.5142 0.2650 -0.2357 -0.1047 0.3750 -0.5353 |
| S2_G     | 0.2632 0.1026 -0.0768 -0.0662 -0.0207 -0.1308 0.5506 0.4368 |

### Figure 7. PCA results. A shows the Pareto chart of the variance by contribution of each principal component, B plots the scatter diagram of the first two components $\gamma_1$ and $\gamma_2$, and C shows the first three components $\gamma_1$, $\gamma_2$, and $\gamma_3$.  

- **Step 2**: $\chi^2$ determination for the $k$th component in 3-dimensional interspace. Since the squared Mahalanobis distances for each Gaussian component follow the Chi-square distribution ($\chi^2$) in 3-dimensional interspace, to determine the decision region for classifying the test data $x$ via the components estimated in the above step, the squared Mahalanobis distance in 3-dimensional interspace for the $k$th component with mean $\mu_k$ and full covariance matrix $\Sigma_k$, $d^2(x|\mu_k, \Sigma_k)$, is computed as follows:
Therefore, $\chi^2_k$, which is constructed based on component $k$ and denoted as $\chi^2_k(\mu_k, \Sigma_k)$, is determined by

$$d^2_k(x|\mu_k, \Sigma_k) = (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k). \quad (21)$$

Therefore, the squared Mahalanobis distance $d^2_k(x|\mu_k, \Sigma_k)$ specified based on the desired confidence level, denoted as $\beta_k$, can be used as the $k$th classifier criterion for determining whether feature $x$ belongs to the $k$th class.

Table 4. The Gaussian mixture parameter estimates are achieved for the new features $[y_1, y_2, y_3]$ by setting the number of Gaussian mixture components as 7.

| Components | Component number | Gaussian mixture parameter estimates $\pi_k, \mu_k, \Sigma_k$ |
|------------|------------------|-------------------------------------------------------------|
| MR Classifier | $k = 1$ | $\pi_1 = 0.1947, \mu_1 = 0.7056, \Sigma_1 = 2.7126, 1.4950, 0.0425, -0.0007, 0.0013$ |
| MS Classifier | $k = 2$ | $\pi_2 = 0.0827, \mu_2 = 3.2981, \Sigma_2 = 2.6064, -3.7382, 0.3310, -0.094, -0.0122$ |
| ASD Classifier | $k = 3$ | $\pi_3 = 0.1130, \mu_3 = 2.3453, \Sigma_3 = 0.3484, 0.5773, 0.5373, -0.0172, -0.0039$ |
| NM Classifier | $k = 4$ | $\pi_4 = 0.1683, \mu_4 = 2.7874, \Sigma_4 = 1.8620, -0.9829, 0.1403, 0.0107, 0.0063$ |
| AS Classifier | $k = 5$ | $\pi_5 = 0.0783, \mu_5 = 0.7511, \Sigma_5 = 0.3199, -0.5341, 0.0972, 0.0077, -0.0161$ |
| AR Classifier | $k = 6$ | $\pi_6 = 0.2676, \mu_6 = -1.2294, \Sigma_6 = 0.1198, 0.3222, 0.3301, -0.0011, 0.0025$ |
| VSD Classifier | $k = 7$ | $\pi_7 = 0.0954, \mu_7 = -0.1631, \Sigma_7 = -1.1167, 0.9454, 0.1338, 0.0048, 0.0095$ |

Figure 8. Flow chart of the diagnostic determination and 3-dimensional surface classifier results.
Step 3: The $k$th confidence level ($\beta_k$) determination. Actually, the $k$th confidence region, as specified by the $k$th desired confidence level $\beta_k$, is surrounded by the $k$th ellipsoid, and this relation is expressed as

$$d_2^2(x | \mu_k, \Sigma_k) \leq MDC_k = \chi_{\frac{3}{2}, \beta_k}^2,$$

where $\chi_{\frac{3}{2}, \beta_k}$ is the inverse of $\chi_{\frac{3}{2}}^2$ for a given confidence level $\beta_k$, and $MDC_k$ represents the classification criterion for component $k$ and satisfies the following equation

$$\chi_{\frac{3}{2}}^2(d_2^2(x | \mu_k, \Sigma_k) = MDC_k | \mu_k, \Sigma_k) = \beta_k$$

For the $\chi_{\frac{3}{2}}^2$ distribution, although the confidence regions corresponding to the confidence levels of 68.3%, 95%, and 97.5% are widely used classification criteria in many studies$^{2,42-45}$, the optional $\beta_k$ is identified by setting $\beta_k \in [0.63, 0.97]$ combined with the following rules: 1) each ellipsoid should be as large as possible; 2) each common region should be as small as possible; and 3) the classification accuracy defined in Eq. (26) should be as high as possible. The classification accuracies for classifying sound data summarized in Sect. 2.2 are plotted in Figs. 9, and 9 shows the following results:

1. For VSD sounds, high accuracy can be achieved by setting the desired confidence level $\beta$ to each value within the interval of $(0.71 < \beta < 0.89)$, as shown in Fig. 9(VSD);
2. For AR and MR sounds, by setting the desired confidence level $\beta$ based on $\beta \in [0.69, 0.81]$, high classification accuracy could be achieved (Fig. 9(MR and AR));
3. For MS, AS and NM sounds, to achieve the accurate classification of HSs, the interval of the desired confidence level $\beta$ should be set as $[0.63, 0.65]$ (Fig. 9); and
4. For ASD sounds, Fig. 9 shows that the highest classification accuracy is achieved by setting the desired confidence level $\beta$ to each value within the interval of $(0.69 < \beta < 0.81)$, as shown in Fig. 9(A). Furthermore, the desired confidence level $\beta$ can be adjusted to improve the classification accuracy and fit new datasets without reperforming the computations for the objective function, especially for VSD sounds and MR sounds (Fig. 9(VSD and MR)). In this study, according to the rules described above combined with the accuracy analysis results plotted in Fig. 9, the $\beta_k (k = 1, \ldots, 7)$ values are set as $0.87, 0.65, 0.67, 0.65, 0.67, 0.79$ and $0.87$, respectively.

Step 4: MDC$_k$ determination corresponding to $\beta_k$. Based on the $k$th confidence level achieved for $\beta_k$ in the above step, by using the function 'chi2inv' in MATLAB 2018b, the inverse of $\chi_{\frac{3}{2}, \beta_k}^2$, denoted as MDC$_k = \chi_{\frac{3}{2}, \beta_k}^2$, is determined. The analysis results for the $k$th confidence region in the 3-dimensional interspace, which is surrounded by the $k$th ellipsoid corresponding to the $k$th desired confidence level $\beta_k$, are determined and shown in Fig. 8B. Furthermore, Fig. 8B shows that the common regions between two random ellipsoids are almost zero; thus, a faulty decision process is avoided because the input will not fall into two or more categories.

Step 5: MDC$_k$-based diagnostic result determination. Based on the ellipsoid surfaces region shown in Fig. 8B, the diagnosis method is described as follows.

Figure 9. The achieved accuracies corresponding to classifying the heart sounds described in Sect. 2.2 by setting $\beta$ form 0.63 to 0.97 with a step of 0.02.
Performance evaluation criteria. 

Table 5. Experimental sounds used to evaluate the performance.

| Data source                  | Period numbers of every type of heart disease/Patients |
|------------------------------|--------------------------------------------------------|
| Sounds in Sect. 2.2          | 769/10 439/5 315/7 1056/45 381/10 665/15 327/10  |
| New sounds                   | 156/3 132/2 82/2 183/8 126/3 153/4 70/3             |
| Total sounds                 | 925/13 571/7 397/9 1239/53 507/13 818/19 397/13  |

Table 6. Efficient methods successfully used in diagnosing normal sounds from other common heart diseases.

1. The 3-dimensional diagnostic features \( \gamma_1, \gamma_2, \gamma_3 \) are first transformed from the features \( FF \) (denoted as \( FF_k \)) of the testing sample and calculated with the following equation

\[
\gamma_i = \frac{FF_k - \mu_{FF}}{\sigma_{FF}} \times \xi_i, \quad i = 1, 2, 3
\]

where \( \mu_{FF} \) and \( \sigma_{FF} \) are shown in Table 2.

2. Then, according to the confidence region shown in Fig. 8B, the MDC\(_k\)-based diagnostic result for a test feature \( x = [\gamma_1, \gamma_2, \gamma_3] \) is determined.

3. MDC\(_k\)-based diagnostic results for a test feature \( x = [\gamma_1, \gamma_2, \gamma_3] \) are determined by

\[
\begin{align*}
\text{Class } k, & \quad d^2_k(x|\mu_k, \Sigma_k) \leq \text{MDC}_k, \\
\text{Unknown class, otherwise } & .
\end{align*}
\]

where class \( k \) corresponding to the type of heart disease is detailed in Table 4, and MDC\(_k\) (1, 2, ..., 7) is 5.6489, 3.2831, 3.4297, 3.2831, 3.4297, 4.5258 and 5.6489.

Performance evaluation criteria. To evaluate the performance of these ellipsoids in 3-dimensional space, the classification accuracy (CA), sensitivity (Se) and specificity (Sp) values are calculated by

\[
\begin{align*}
\text{CA} &= \frac{TP+TN}{TP+FP+FN+TN}, \\
\text{Se} &= \frac{TP+TN}{TP} \times 100, \\
\text{Sp} &= \frac{TP+TN}{FP+TN} \times 100.
\end{align*}
\]

where TP, FP, TN and FN are the numbers of true positives, false positives, true negatives and false negatives, respectively.

Performance evaluation

To evaluate the performance of the proposed methodology, the comparison between the proposed methodology and the state-of-the-art methods on the clinical sounds and online sounds data was conducted as follows.

- Total sounds: The total sounds, consisting of sounds described in Sect. 2.2 and new sounds, were summarized in Table 5 to evaluate the performance of this proposed methodology.
- State-of-the-art methods: To highlight the efficiency of the proposed methodology for diagnosing the seven typical heart diseases, the state-of-the-art methods, published in recent five years and described in Table 6, were comparatively analyzed.
A novel intelligent system was proposed for diagnosing heart diseases with high CA. The innovation of this approach is primarily reflected in: 1) the automatic extraction of secondary envelope-based frequency features; 2) the automatic determination of PCA-based diagnostic features $\gamma_1$, $\gamma_2$ and $\gamma_3$; and 3) the determination of adjustable confidence regions corresponding to the $\chi^2$ distribution. The confidence regions are obtained by calculating the Mahalanobis distance, which is adjusted by the desired confidence level $\beta$, and the results were used as the classification criteria for diagnosing heart diseases. The procedure for the implementation of the intelligent

| Method | MR | Se% | CA(%) | Sp% |
|--------|----|-----|-------|-----|
|        |    | 92.1| 86.34 | 87.6 |
|        |    | 90.6| 89.93 | 88.3 |
|        |    | 90.1| 88.6 | 87.5 |
|        |    | 90.1| 87.34| 86.7 |
|        |    | 88.6| 85.93| 85.3 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.3| 90.4 | 90.3 |
|        |    | 100 | 99.43| 99.3 |
|        |    | 92.1| 86.53| 87.4 |
|        |    | 83.9| 86.21| 87.2 |
|        |    | 83.1| 81.67| 81.8 |
|        |    | 85.2| 83.52| 83.3 |
|        |    | 92.0| 83.13| 82.5 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 92.0| 83.13| 82.5 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |
|        |    | 90.0| 83.96| 90.2 |
|        |    | 89.7| 91.64| 92.1 |

Table 7. Comparative analysis of eight different methods for the diagnosis of heart diseases summarized in Table 5.

| Method | Performance evaluation |
|--------|------------------------|
| 1⁶ | The highest classification accuracies were obtained by using the features described in Table 3 on page 28. |
| 2⁷ | The highest classification accuracies were obtained by using the 13-features extracted using MFCC algorithm. |
| 3⁸ | The highest classification accuracies were obtained by using the proposed WaveNet model consists of 6 residual blocks. |
| 4⁹ | The highest classification accuracies were obtained based on the rules described in a previous study⁸. |
| 5⁹ | The highest CA results were obtained based on the following rules. |
|     | Rule 1: If the $8^{th}$ value of Lyapunov exponent ($LPE_8$) ≥ 0.79 and ($LPE_9$) ≤ 0.38 then the heart is normal. |
|     | Rule 2: If $LPE_5 ≤ 0.17$ and ($LPE_9$) ≥ 0.1, then the heart disease is VSD. |
|     | Rule 3: If $LPE_5 ≥ 0.17$, $LPE_6 ≤ 0.39$, and $LPE_7 ≥ 0.56$, then the heart disease is MR. |
|     | Rule 4: If $LPE_5 ≥ 0.17$, $LPE_6 ≥ 0.67$, and $LPE_7 ≥ 0.37$, then the heart disease is MS. |
|     | Rule 5: If $LPE_5 ≥ 0.54$, $LPE_6 ≥ 0.29$, and $LPE_7 ≥ 0.49$, then the heart disease is AR. |
|     | Rule 6: If $LPE_5 ≥ 0.39$, $LPE_6 ≥ 0.72$, and $LPE_7 ≤ 0.68$, then the heart disease is ASD. |
|     | Rule 7: If $LPE_5 ≥ 0.64$, $LPE_6 ≥ 0.39$, and $LPE_7 ≥ 0.21$, then the heart disease is AS. |
|     | Rule 8: If none of these conditions are met, the HS is undefined. |
| 6⁹ | The most accurate results were obtained by the structure consisting of one input layer with 60 neurons, one hidden layer with 11 neurons and one output layer with five neurons. |
| 7⁹ | The most accurate results were obtained by setting the number of features at each node, the number of trees and the maximum depth of trees to 1, 108, and 36, respectively. |

Table 8. The highest accuracies corresponding to the parameters set in every state-of-the-art method.

- Comparison results: The comparison results were summarized in Table 7, where the parameters corresponding to the state-of-the-art methods were described in Table 8. The results in Table 7 support the following conclusions.
- Although using the method #1 to diagnose AS yielded a higher Sp than that of the proposed method, the CA was lower than that of the proposed method, partially due to the high Se achieved by the proposed method.
- Although using the method #3 to diagnose MS yielded a higher Sp than that of the proposed method, the CA was lower than that of the proposed method, partially due to the high Se achieved by the proposed method.
- Although using the method #5 to diagnose NM yielded a higher Sp than that of the proposed method, the CA was lower than that of the proposed method, partially due to the high Se achieved by the proposed method.
- For other sounds, the classification accuracies achieved in the proposed method were all greater than those of the other methods listed in Table 7.

Overall, the efficiency of the proposed method in diagnosing MR, MS, ASD, NM, AS, AR and VSD diseases was evaluated by comparison with the other efficient methods listed in Table 7.

Conclusion

A novel intelligent system was proposed for diagnosing heart diseases with high CA. The innovation of this approach is primarily reflected in: 1) the automatic extraction of secondary envelope-based frequency features; 2) the automatic determination of PCA-based diagnostic features $\gamma_1$, $\gamma_2$ and $\gamma_3$; and 3) the determination of adjustable confidence regions corresponding to the $\chi^2$ distribution. The confidence regions are obtained by calculating the Mahalanobis distance, which is adjusted by the desired confidence level $\beta$, and the results were used as the classification criteria for diagnosing heart diseases. The procedure for the implementation of the intelligent
system involved three stages. Stage 1 described the location and extraction of STMHT-based CS1 and CS2. In stage 2, in the frequency domain, a novel method was first proposed to generate the envelopes CS1_kFE and CS2_kFE then, based on the Thr lines, ThF was automatically extracted. Finally, based on PCA, the first three principal components, γ₁, γ₂ and γ₃, which expressed 86.7% of the FF information, were determined and used as diagnostic features. In stage 3, the GMM-based objective function JS(x) with respect to the features $x = [γ₁, γ₂, γ₃]$ and the parameters $[π_k, μ_k, Σ_k]$, where $k = 1, 2, \ldots, K$, was generated. Then, the $χ^2$ distribution for component $k$ was determined by calculating the Mahalanobis distance from $x$ to the class mean $μ_k$ of component $k$, and the confidence region for component $k$ was determined by adjusting the optimal confidence level $β_k$ and used as the criterion (denoted as MDC_k) to diagnose a given HS. The performance evaluation was validated by sounds from online HS databases and clinical heart databases. The accuracy of the proposed method was compared to the accuracies of other well-known classifiers, and the highest classification accuracies of 99.43%, 98.93%, 99.13%, 99.85%, 98.62%, 99.67% and 99.91% in the detection of MR, MS, ASD, NM, AS, AR and VSD sounds were achieved by setting $β_k(k = 1, 2, \ldots, 7)$ to 0.87, 0.65, 0.67, 0.65, 0.67, 0.79 and 0.87, respectively. Therefore, this proposed intelligent diagnosis system provided an efficient way to diagnose seven types of heart diseases.

The advantages and limitations were summarized as follows:

- **Advantages:** (1) CS1 and CS2 were automatically extracted to reduce difficulty in segmenting each cardiac cycle into a sequence of four heart stages: $S₁ \rightarrow$ Systole period $\rightarrow S₂ \rightarrow$ Diastole period. (2) More features could be extended by setting even more threshold values for the unknown heart diseases, especially for the heart sound with the compound heart diseases; (3) Every classifier achieved in this study could be adjusted based on the desired $β$ for fitting incremental new features without being retrained via huge training features.

- **Limitations:** (1) This methodology was impossible to diagnose the sounds when CS1 and CS2 cannot be segmented and extracted via the STMHT method for a given heart sound such as that plotted in Fig. 10; (2) The proposed classifier might not be satisfied with the compound heart diseases due to the distribution of features extracted from which can not fit a single Gaussian distribution.

**Future study**

Future study focused on how to handle the sounds (such as some AR sounds) when CS1 and CS2 cannot be segmented and extracted via the STMHT method will be explored, and on how to build the classifier model for fitting the compound heart diseases will be further studied.

**Research statement**

The study was conducted at Nanyang Institute of Technology and Nanyang First People's Hospital, Henan, China from December 2017 to June 2021, and was approved by the ethics committee of Nanyang Institute of Technology and First People's Hospital (Approval Number: V6.0). Informed consent was waived due to the retrospective design of the study. The study complies with the Declaration of Helsinki.

Received: 17 February 2021; Accepted: 5 October 2021
Published online: 25 January 2022

**References**

1. Tan, Z., Wang, W., Zong, R., Pan, J. & Yang, H. [classification of heart sound signals in congenital heart disease based on convolutional neural network]. Shengwu yixue gongcheng xue za zhi = J. Biomed. Eng. Shengwu yixue gongchengxue zazhi 36, 728–736, https://doi.org/10.7507/1001-5515.201806031 (2019).

2. Sun, S., Wang, H., Chang, Z., Mao, B. & Liu, Y. On the Mahalanobis distance classification criterion for a ventricular septal defect diagnosis system. IEEE Sens. J. 19, 2665–2674. https://doi.org/10.1109/JSEN.2018.2882582 (2019).

3. Omari, T. & Bereksi-Reguig, F. A new approach for blood pressure estimation based on phonocardiogram. Biomed. Eng. Lett. 9, 395–406. https://doi.org/10.1007/s13534-019-00113-z (2019).

4. Notario, P. M. et al. Home-based telemedicine for children with medical complexity. Telemed. e-Health (2019).

5. Coviello, J. S. Auscultation Skills: Breath & Heart Sounds. -5th Ed. (Lippincott Williams & Wilkins, 2013), 5 edn.

6. Liu, Q., Wu, X. & Ma, X. An automatic segmentation method for heart sounds. Biomed. Eng. Online 17, 1–22. https://doi.org/10.1186/s12938-018-0538-9 (2018).

7. Marston, E., Zohrer, M. & Pernkopf, F. Heart sound segmentation: an event detection approach using deep recurrent neural networks. IEEE Trans. Biomed. Eng. 65, 1964–1974. https://doi.org/10.1109/TBME.2018.2843258 (2018).

8. Sun, S. & Wang, H. Principal component analysis-based features generation combined with ellipse models-based classification criterion for a ventricular septal defect diagnosis system. Aust. Phys. Eng. Sci. Med. 31, 655–662. https://doi.org/10.1007/s13246-018-0676-1 (2018).

9. Sun, S. An innovative intelligent system based on automatic diagnostic feature extraction for diagnosing heart diseases. Knowl. Based Syst. 75, 224–238. https://doi.org/10.1016/j.knosys.2014.12.001 (2015).

10. Shuping Sun, Z. J. Y., Haibin, W. & Tao, T. Segmentation-based heart sound feature extraction combined with classifier models for a VSD diagnosis system. Exp. Syst. Appl. 41, 1769–1780. https://doi.org/10.1016/j.eswa.2013.08.076 (2014).
11. Choi, S. & Ijang, Z. Cardiac sound murmurs classification with autoregressive spectral analysis and multi-support vector machine technique. *Comput. Biol. Med.* 40, 8–20. https://doi.org/10.1016/j.compbiomed.2009.10.003 (2010).
12. Zhang, J., Yin, Z. & Wang, R. Pattern classification of instantaneous cognitive task-load through GMM clustering, Laplacian eigenmap, and ensemble SVMs. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 14, 947–965. https://doi.org/10.1109/TCBB.2016.256197 (2017).
13. Li, Z., Xia, Y., Ji, Z. & Zhang, Y. Brain voxel classification in magnetic resonance images using niche differential evolution based Bayesian inference of variational mixture of Gaussians. *Neurocomputing* 269, 47–57. https://doi.org/10.1016/j.neucom.2016.08.147 (2017).
14. Ortiz-Rosario, A., Adeli, H. & Buford, J. A. MUSIC-Expected maximization gaussian mixture methodology for clustering and detection of task-related neuronal firing rates. *Behav. Brain Res.* 317, 226–236. https://doi.org/10.1016/j.bbr.2016.09.022 (2017).
15. Davari, A., Aftoula, E., Yanikoglu, B., Maier, A. & Riess, C. GMM-based synthetics samples for classification of hyperspectral images with limited training data. *IEEE Geosci. Remote Sens. Lett.* 15, 942–946. https://doi.org/10.1109/LGRS.2018.2817361 (2018). arXiv:1712.04778.
16. Simms, L. M. et al. Nuclear Inst. and Methods in Physics Research. A Pulse discrimination with a Gaussian mixture model on an FPG. *Nucl. Inst. Methods Phys. Res.* A900, 1–7. https://doi.org/10.1016/j.nima.2018.05.039 (2018).
17. Xue, W. & Jiang, T. An adaptive algorithm for target recognition using Gaussian mixture models. *Meas. Int. Meas. Conf.* 124, 233–240. https://doi.org/10.1016/j.measurement.2018.04.019 (2018).
18. Zhang, S. et al. Segmentation of small ground glass opacity pulmonary nodules based on Markov random field energy and Bayesian probability difference. *Biomed. Eng. Online* 19, 1–20. https://doi.org/10.1186/s12938-020-00793-0 (2020).
19. The World Medical Association Inc. DECLARATION OF HELSINKI Ethical Principles for Medical Research Involving Human Subjects Adopted. *WMA General Assembly, Somerset West, Republic of South Africa* 1–5 (2008).
20. 3MCompany. 3m health care company. http://www.3M.com/ (2019).
21. Bernard Karnath, W. T. Auscultation of the heart. http://www.turner-white.com/pdf/hp_sep02_heart.pdf (2002).
22. Yeo, T. C. http://www.cctoolbox.com/index.html (2019).
23. Máttar, J. A. http://www.med.umich.edu/lrc/psb/heart sounds/index.htm (2019).
24. Medical Sound Library. Auscultate - Learn Heart Sounds, Murmurs and Medical Auscultation (2019).
25. ThinkLabs. Thinklabs heart sound library. https://www.thinklabs.com/heart-sounds?hc_location=us (2019).
26. MacWalter, D. & MacWalter, G. Human heart sounds and murmurs. http://www.dundee.ac.uk/medther/Cardiology/hsmurl.html (2019).
27. Shuping Sun, H. W. A novel method-based secondary envelope extraction for heart sound analysis (2020).
28. 3MDatabase. 30 heart and lungs sounds library. http://solutions.3m.com/wps/portal/3M/en_US/3M-Littmann-EMEA/stethoscope/littmann-learning-institute/heart-lung-sounds/heart-lung-sound/ (2019).
29. Medical Sound Library. Auscultate - Learn Heart Sounds, Murmurs and Medical Auscultation (2019).
30. AMBOSSMed. AMBOSS: Medical Knowledge Distilled (2019).
31. Auscultation Sound. Heart Murmur-Mitral Regurgitation Auscultation Sound !!! Complete (2019).
32. Medical Sound Library. Auscultate - Learn Heart Sounds, Murmurs and Medical Auscultation (2019).
33. MacWalter, D. & MacWalter, G. Human heart sounds and murmurs. http://www.dundee.ac.uk/medther/Cardiology/hsmurl.html (2019).
34. ThinkLabs. Thinklabs heart sound library. https://www.thinklabs.com/heart-sounds?hc_location=us (2019).
35. Mohseni, S. S. Heart arrhythmias classification via a sequential classifier using neural network , principal component analysis and heart rate variation. *IEEE 8th International Conference on Intelligent Systems* Heart 715–722 (2016).
36. Ketabha, R. & Kannan, E. An efficient framework for heart disease classification using feature extraction and feature selection technique in data mining. In: Proceedings of the 1st International Conference on Emerging Trends in Engineering, Technology and Science. ICETETS 2016 https://doi.org/10.1109/ICETETS.2016.7603000 (2016).
37. Guo, H.-W. et al. Heart rate variability signal features for emotion recognition by using principal component analysis and support vectors machines. In: Proceedings of the 2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE) 274–277. https://doi.org/10.1109/BIBE.2016.40 (2016).
38. Motin, M. A Principal component analysis : a novel approach for extracting respiratory rate and heart rate from photoplethysmographic signal.22, 766–774 (2018).
39. H. El-Saadawy, H. A. S., M. Tantawi & Tolba, M. F. Electrocardiography (ecg) heart disease diagnosis using pnn, svm and softmax regression classifiers. In: The Eighth International Conference on Intelligent Computing and Information Systems (ICICIS) 106–110 (2017).
40. Johnson, R. A. & Wichern, D. W. *Applied Multivariate Statistical Analysis* (6th Edition) (Pearson, 2007).
41. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* 39, 1–38 (1977).
42. Pinto, R. C. & Engel, P. M. A fast incremental gaussian mixture model. *PLoS ONE* 10, 1–12. https://doi.org/10.1371/journal.pone.0139931 (2015).
43. Proia, E., Perera, A., Thouroude, T., Michel, G. & Clotault, J. On the characterization of flowering curves using Gaussian mixture models. *J. Theor. Biol.* 402, 75–88. https://doi.org/10.1016/j.jtbi.2016.04.022 (2016).
44. Mangi, P. K. Using Keystroke Dynamics in a Multi-level Architecture to Protect Online Examinations from Impersonation. In: *Proceedings of the 2017 IEEE 2nd International Conference on Big Data Analysis* 622–627 (2017).
45. Aryafar, A., Mikael, R., Doulali Ardejani, E., Shaffiee HaghiShenas, S. & Jafapour, A. Application of non-linear regression and soft computing techniques for modeling process of pollutant adsorption from industrial wastewaters. *J. Min. Environ.* 210–220. https://doi.org/10.1016/j.cmpb.2020.105604 (2020).
46. Sawyer, N. K., Patidar, S., Nesraghi, N. & Acharya, U. R. Automated detection of abnormal heart sound signals using Fano-factor constrained tunable quality wavelet transform. *Biocybern. Biomed. Eng.* 41, 111–126. https://doi.org/10.1016/j.bisei.2020.12.007 (2021).
47. Deperfioglou, O. Heart sound classification with signal instant energy and stacked autoencoder network. *Biomed. Signal Process. Control* 64, 102311. https://doi.org/10.1016/j.bspc.2020.102311 (2021).
48. Oh, S. J. et al. Classification of heart sound signals using a novel deep WaveNet model. *Comput. Methods Prog. Biomed.* https://doi.org/10.1016/j.cmpb.2020.105604 (2020).
49. Karar, M. E., El-Khatif, S. H. & El-Brawany, M. A. Automated diagnosis of heart sounds using rule-based classification tree. *J. Med. Syst.* 41, 60. https://doi.org/10.1007/s10916-017-0704-9 (2017).
50. S. Oh, S. L. et al. Heart sound anomaly and quality detection using ensemble of neural networks without segmentation. *Comput. Cardiol. Conf.* (Conf.Ci) (2016) https://doi.org/10.22489/Conf.Ci.2016.180-213.
51. Karar, M. E., El-Khatif, S. H. & El-Brawany, M. A. Automated diagnosis of heart sounds using rule-based classification tree. *J. Med. Syst.* 41, 60. https://doi.org/10.1007/s10916-017-0704-9 (2017).
52. Zabih, M. et al. Heart sound anomaly and quality detection using ensemble of neural networks without segmentation. *Comput. Cardiol. Conf.* (Conf.Ci) (2016) https://doi.org/10.22489/Conf.Ci.2016.180-213.
53. Witten, I. H., Frank, E. & Hall, M. *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, 2016), fourth edn. arXiv:1011.1669v3.
Acknowledgements
We would like to express our thanks to the medical ethics committee of Nanyang Institute of Technology and the First Peoples Hospital of Nanyang for their approval to conduct this research in Nanyang Institute of Technology's biomedical information unit. The authors would like to thank Vaibhav Deshmukh for his suggestions on submitting this manuscript to Scientific Reports. This work was supported by the Key Projects of Hunan Provincial Department of Education (Nos. 21A0403 and 21A0405), and the university-industry collaborative project (No. 202102211006).

Author contributions
Sun, Biqiang Zhang and Huang: collection, organizing, and review of the literature. Sun, Biqiang Zhang and Huang: preparing the manuscript. Sun, Huang, Biqiang Zhang, He, Yan, Fan, Jiale Zhang and Chen: manuscript revision and modification. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to S.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022