Biosystems Engineering 238 (2024) 78–88

Contents lists available at ScienceDirect

Biosystems Engineering

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Automatic detection of continuous pig cough in a complex piggery environment

Xipeng Wang, Yanling Yin [**], Xinpeng Dai, Weizheng Shen [*], Shengli Kou, Baisheng Dai

School of Electrical Engineering and Information, Northeast Agricultural University, Harbin, 150030, China

ARTICLE INFO

Keywords:
Pig respiratory diseases
Voice activity detection
Multi-classifier selection
Multi-classifier fusion

ABSTRACT

Pig cough sound monitoring is an effective means of early warning for respiratory diseases. Until present, most studies focused on the investigation of high precision pig cough recognition algorithms based on manually segmented individual sound datasets. However, the recognition of continuous sound was mostly ignored, which cannot apply in practical engineering. Meanwhile, less consideration has been given to complex scenarios, such as the overlap of multiple sounds, which occur frequently in large-scale piggeries. To this end, we explored an automatic detection of continuous sound algorithm and proposed a continuous pig cough recognition method which has a significant role in the diagnosis of diseases. Initially, we proposed a voice activity detection (VAD) method to automatically segment continuous sound. Subsequently, we investigated a multi-classifier fusion strategy to promote recognition accuracy. Finally, we proposed a low-complexity continuous pig cough recognition method. The experimental results show that the recall and precision of pig cough in continuous sound is 93.1% and 91.6% respectively, which is much higher than the 67.3% and 90.6% of the baseline detection method. The recognition accuracy of continuous pig cough reached 91.4%. From the perspective of practical application, our algorithm development considered the complex environment of a real pig barn.

1. Introduction

Pig respiratory diseases seriously restrict healthy pig breeding due to their highly infectious and lethal characteristics (Assavacheep & Thanawongnuwech, 2022; Ji, Yin, et al., 2022; Sassu et al., 2018). The traditional early warning methods mainly rely on manual inspection. In recent years, an increasing number of studies have focused on the automatic detection of pig coughs to provide warnings (Racewicz et al., 2021). Previous studies have given many solutions in terms of pig cough recognition and achieved promising results. In the early stage, researchers tried different acoustic features and classifiers in speech recognition for pig cough recognition, such as mel-frequency cepstrum coefficient (MFCC) (Chung, Oh, Lee, Park, & Kim, 2013), power spectral density (PSD) (Exadaktylos, Silva, Aerts, Taylor, & Berckmans, 2008) features and fuzzy c-means (Exadaktylos et al., 2008; Hirtum & Berckmans, 2003), support vector machine (SVM) (Chung et al., 2013), and dynamic time warping (Guarino, Jans, Costa, Aert, & Berckmans, 2008) classifiers. The highest recognition accuracy was approximately 94% (Chung et al., 2013). With the development of deep learning models, the cough recognition accuracy was further improved. Different mature neural networks and self-built networks were used for cough recognition,such as convolutional neural networks (Shen, Wang, et al., 2022; Song, Zhao, Hu, Sun, & Zhou, 2022) and deep belief networks (Li et al., 2018), etc. The accuracy reached 96.7% (Shen, Wang, et al., 2022). Our team has also carried out many studies in recent years. We first studied the methods based on convolutional neural network (Shen, Tu, Yin, & Bao, 2021; Yin, Tu, Shen, & Bao, 2021). Then we focused on feature selection, feature fusion (Ji, Shen, et al., 2022; Shen ,Ji, et al., 2022), and classifier fusion strategies (Yin et al., 2023) to promote accuracy in complex piggery environments. The best accuracy was 99.2% based on the self-built dataset including 2500 individual sounds.

Although previous studies have achieved high recognition accuracy, there are still some obvious shortcomings. First, most of the methods are based on isolated single sound recognition. However, due to the lack of continuous sound automatic segmentation, it cannot achieve cough detection in continuous sound, which makes it cannot be applied to practical engineering. Second, the existing research mainly focused on the recognition of a single cough and lacked continuous cough (multiple single coughs with very short intervals) recognition, which is important for the diagnosis and treatment of the diseases. Third, there is a lack of a datasets of labelled continuous sounds. The majority of the datasets used

* Corresponding author.
** Corresponding author.
E-mail addresses: yinyanling@neau.edu.cn (Y. Yin), wzshen@neau.edu.cn (W. Shen).

https://doi.org/10.1016/j.biosystemseng.2024.01.002
Received 17 September 2023; Received in revised form 6 December 2023; Accepted 5 January 2024
Available online 15 January 2024

Nomenclature

Acc	Accuracy of the base classifier
$Acc_{\max}$	Maximum values of Acc
$Acc_{\min}$	Minimum values of Acc
AD	Accuracy–diversity
AdaBoost	Adaptive boosting
a_i	Threshold parameters of low threshold in VAD
b_i	Threshold parameters of high threshold in VAD
BiLSTM	Bidirectional long short-term memory
C_i	The i^{th} base classifier
CQT	Constant-Q transforms
D	The number of deletions in continuous coughs
D_{ij}	Time interval between two neighbouring coughs (s)
$D_{\max}$	Maximum time interval between two neighbouring coughs (s)
Dv	Diversity metric
$Dv_{\max}$	Maximum values of the Dv
$Dv_{\min}$	Minimum values of the Dv
DTED	Double threshold energy detection
DTFBVD	Double threshold frequency band variance detection
DTMISD	Double threshold MFCC inverse spectral distance detection
E_{noise}	The normalised energy of N frame background noise
FP	False positive
FN	False negative
K	Count the number of individual sounds for continuous coughs detection
I	The number of insertions in continuous coughs
LPCC	Linear prediction cepstral coefficient
MFCC	Mel-frequency cepstrum coefficient
M	Number of individual sounds in the buffered segment
N	Count the number of continuous coughs for continuous coughs detection
N_f	The number of total frames of sound segment
N_{miss}	The number of missing frames
N_{false}	The number of error frames
N_{wd}	The number of detected coughs by VAD
N_w	The total number of coughs in data set
N_{cs}	The total number of coughs in data set
N_{dcs}	The number of correctly detected cough
N_{fcs}	The number of non-coughs detected as coughs
N_c	The total number of coughs in a continuous cough
N_{mc}	The number of missed detections of continuous cough
N_{fc}	The number of false detections of continuous cough
N_{cc}	The total number of continuous coughs
PSD	Power spectral density
P_{miss}	Frame miss detection probability of VAD
P_{false}	Frame false detection probability of VAD
P_0	Indicator for VAD threshold search of VAD
P_d	Frame detection probability of VAD
P_{wd}	Cough detection rate in units of word vector of VAD
P_{mcs}	Cough miss detection probability of continuous sound recognition
P_{fcs}	Cough false detection probability of continuous sound recognition
P_{dcs}	Cough detection accuracy of continuous sound recognition
P_{mc}	Miss detection probability of continuous cough recognition
P_{fc}	False detection probability of continuous cough recognition
P_{dc}	Detection accuracy of continuous cough recognition
S	The number of substitutions in continuous coughs
SVM	Support vector machine
SNR	Signal-to-noise (dB)
TN	True negative
TP	True positive
TH	High threshold in DTED method
TL	Low threshold in DTED method
T	Duration of the individual sound segment (s)
$T_{\max}$	Maximum duration of individual cough (s)
VAD	Voice activity detection
WER	Word error rate of continuous cough recognition
α	Update rate for E_{noise}
β	Weight for balancing the accuracy and diversity

in the existing studies was manually segmented into distinct individual sounds based on the experts' annotations, which lost some import information in continuous sounds. Last, the datasets used in the existing study is relatively simple compared to the sounds in a large-scale piggery. Coughs are generally interference-free with a relatively high signal-to-noise ratio (SNR), and non-coughs contain only a limited variety. However, the acoustic environment in the real pig barn is much more complex, mainly reflected in the following two aspects: first, the sound sources in the pig barn are very complex, often with multiple sounds overlapped together. For example, coughs may be overlapped with fan noise, human talking, or pig screams. According to the statistical results of the dataset in this research, nearly one-fifth of coughs are overlapped with other sounds. Second, there are many kinds of non-coughs, up to a dozen. There is a serious imbalance in the number of some non-coughs. Meanwhile, it is time-consuming and challenging to annotate non-coughs.

To solve the above problems, we proposed a continuous cough automatic detection method to detect single cough and continuous cough in a complex piggery environment. The contributions of this paper can be summarised as follows.

(1) We built a continuous sound dataset in a complex piggery environment.
(2) We proposed an automatic cough detection method in continuous sounds.
(3) We significantly enhanced the cough detection probability by the improved VAD and multi-classifier fusion algorithm.
(4) We proposed a low-complexity continuous pig cough recognition method.

The remainder of this paper is organised as follows. Section 2 describes the dataset used in this work. Section 3 presents the details of our proposed method for pig cough detection. The experimental results are shown in Section 4. A discussion of the results is provided in Section 5. Conclusions are drawn in Section 6.

2. Datasets

2.1. Data Collection

The data were collected in a large-scale fattening pig barn, as shown in Fig. 1. The pig barn measured 27.5 m in length, 12.8 m in width and 3.2 m in height and consisted of 12 large pens and 9 small pens. The large pen was 4.15 m long and 3.6 m wide, while the small pen was 3.6 m long and 2.75 m wide. The barn was mechanically ventilated, and the pigs were automatically fed. The floor was a half-slatted concrete floor. The workers shovelled and flushed the floor in the morning and afternoon. There were 128 pigs distributed in 12 large pens, as shown in Fig. 1. Two pigs with severe coughs were isolated in pen 13 near the door. The cough had been present for several days when we collected the

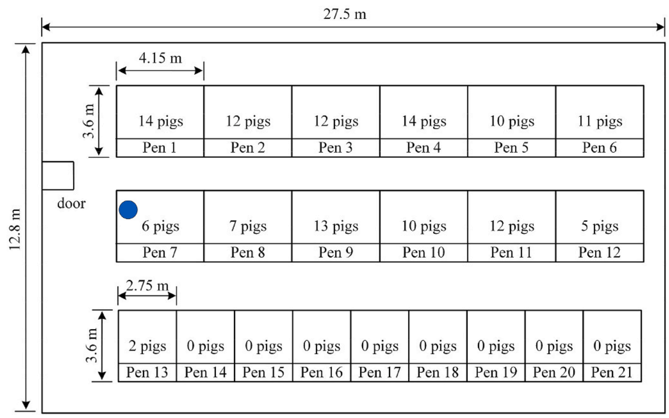

Fig. 1. Sketch of the pig barn in this study. The microphone was placed in pen 7.

data, and there were several coughing pigs in each pen.

We connected a microphone (LIQI LM320E, Cardioid electret microphone, frequency range 100Hz-16kHz) to a laptop with a sound card (Conexant Smart Audio HD) for sound recording. Limited by experimental conditions, we fixed the microphone to a large pen near the door. The microphone was approximately 1.4 m from the ground and 0.8 m from the pigs' back. We set the audio sample rate to 44100 Hz and the resolution to 16 bits. The sounds were saved in ".wav" format. We collected data for 7 consecutive days.

2.2. The dataset

With the help of veterinarians, we manually segmented and labelled the sound segments from the entire continuous audio files by using "Cool Edit Pro 2.1" software. The onset and offset of each individual cough were further labelled in the continuous sound segments. The composition of the dataset is shown in Table 1. The dataset consisted of two parts. The first part was 316 continuous cough segments, ranging from 7 – 40 s each, included a large number of coughs and a small number of non-coughs. The second part consisted of continuous non-cough segments randomly and uniformly selected in the morning, afternoon, evening, and night. Each of them lasted 10 min, for a total of 25 segments, contained a large number of non-coughs and a small number of coughs. The two parts of the data contained a total of 3408 coughs and 3133 non-coughs, of which 2898 single coughs constituted a total of 332 continuous coughs. The proportion of complex coughs (the cough overlapped with other sounds) was approximately one-fifth. Non-coughs included a variety of sounds, such as pig squeals, snoring, sneezing, human talking, clearing noises, and other sounds. We divided the dataset into a training set and a test set.

2.3. Analysis of the dataset

In a large-scale pig barn, the general rearing density is high, and the number of pigs is large. Pigs are vocally active animals and often scream as a result of biting and fighting. Meanwhile, there is much mechanical noise in the piggery. The complex piggery environment poses a great challenge to the recognition of pig cough. The complex pig cough can be summarised as the following three situations.

Table 1
Composition of the dataset.

Dataset	Total number of segments	Duration of each segment	Total number of single coughs	Total number of non-coughs	Total number of continuous coughs
Part 1	316	7–40 s	3155	768	313
Part 2	25	10 min	253	2365	19
All	341	–	3408	3133	332

(a) The continuous coughs from two different pigs are overlapping, as shown in Fig. 2 (a). Two coughs in close proximity may be detected as one sound, which is difficult to correctly recognise as a cough.
(b) A continuous cough is overlapped with a weak non-cough, as shown in Fig. 2 (b). The overlapped area is reflected as partial overlap or full overlap, and there may be multiple overlaps of different sounds. In this case, the traditional VAD algorithm may lead to missed detection of some coughs.
(c) A continuous cough is overlapped with a strong non-cough, as shown in Fig. 2 (c). It is more challenging for the VAD algorithm in this case. When two sounds are fully overlapped and the overlapped non-cough intensity is much greater than the cough, it is difficult to detect.

3. Proposed methods

In this section, we present the details of the proposed continuous pig cough recognition method. We first introduce the overall framework of the proposed method in this study. We then describe the proposed continuous sound automatic detection method, including the improved VAD algorithm and the improved classification algorithm, and then describe the continuous cough detection technique. Finally, we elaborate the evaluation metrics in this paper.

3.1. Framework of the proposed method

The framework of the proposed method is shown in Fig. 3. The recorded continuous sound is first buffered into segments with an overlap. The overlap is to prevent a single cough from being split into two segments. The typical buffer size is 5 – 15 s, and the overlap can be set to 1 – 3 s. The segment is then pre-processed, which includes pre-emphasis, filtering, framing and windowing. Subsequently, the active sound is detected by the proposed improved VAD algorithm. We recorded and stored the start and end times of the active sound used for the judgement of a continuous cough. Then, the features, including acoustic features and deep features, are extracted from the detected active sound segment. By inputting different features into different classifiers, we can obtain several base classifiers. These base classifiers are selected using the selection algorithm based on accuracy–diversity (AD) metrics, and the selected base classifiers are fused using soft voting to make a decision. Finally, if a cough is detected, we further distinguish between a single cough and a continuous cough based on the information stored in the VAD step.

3.2. Improved VAD algorithm

The traditional double threshold energy detection-based (DTED) VAD method can achieve good detection results under the condition that the background noise is stationary and the SNR is high (Exadaktylos et al., 2008; Li et al., 2018; Song et al., 2022). However, we found that there are two obvious limitations of this method in complex piggery environments.

(1) The noise is nonstationary, and the SNR changes rapidly with time.
(2) The silence duration between two consecutive coughs is short, and the energy is much higher than the background noise due to the overlap of other sounds.

This may lead to two outcomes.

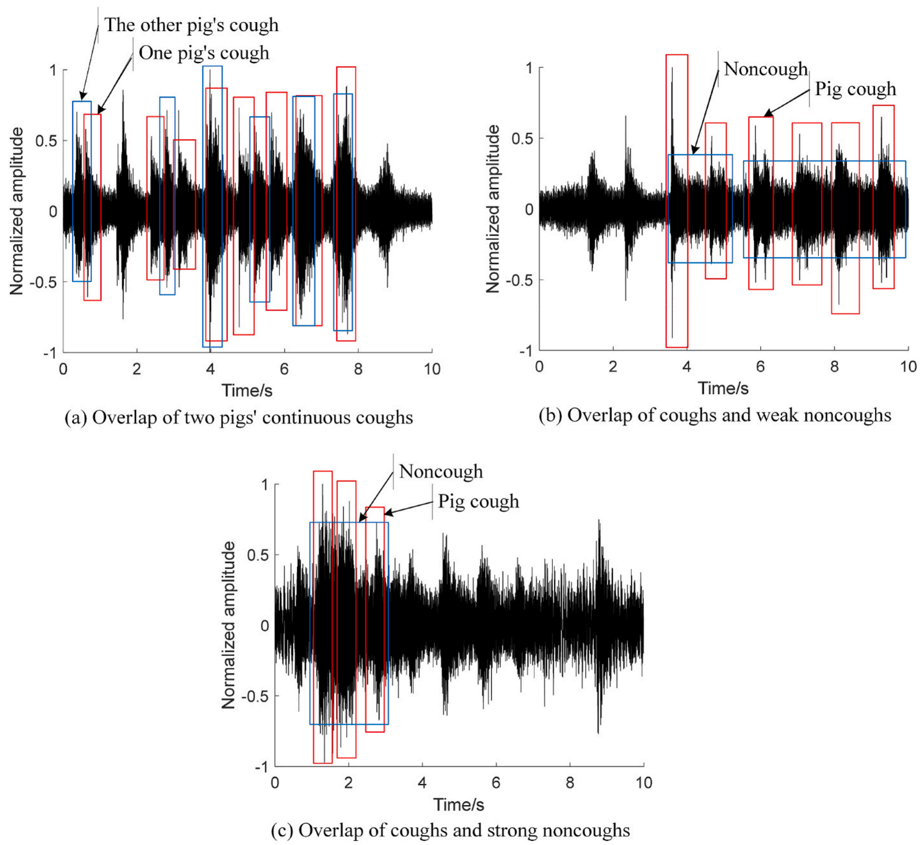

Fig. 2. Examples of some complex situations. (a) Overlap of two pigs' continuous coughs. The red square represents one pig's cough, and the blue square represents the other pig's cough. (b) Overlap of strong coughs and weak non-coughs. The red square represents pig cough, and the blue square represents non-cough. (c) Overlap of weak coughs and strong non-coughs. The red square represents pig cough, and the blue square represents non-cough. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

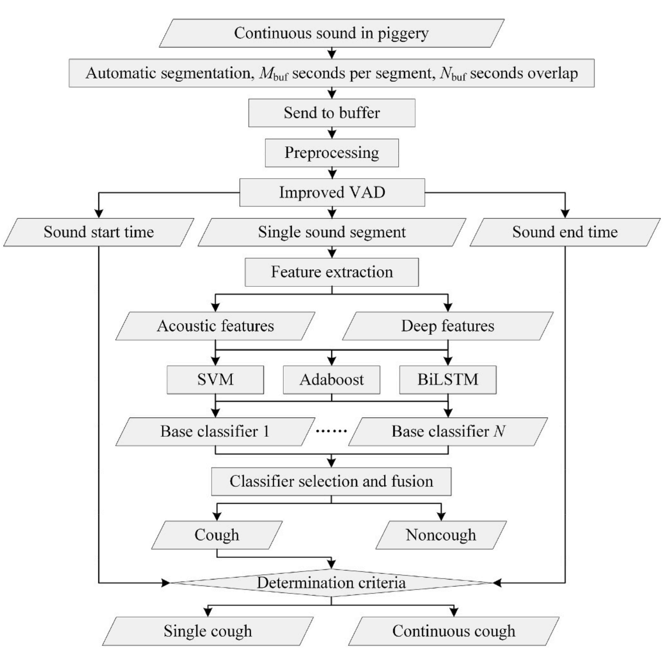

Fig. 3. Overall framework of the proposed method.

(1) A high threshold may result in the misdetection of some low- SNR sounds.
(2) A low threshold may result in multiple sounds being detected as one individual sound.

Therefore, we proposed an improved VAD algorithm to solve the problems by setting adaptive multiple thresholds, the flowchart of the algorithm is shown in Fig. 4.

There are two thresholds in the DTED method. The high threshold *TH* is used for the first rough judgment to determine whether there is an active sound. The low threshold *TL* is used to determine the onset and offset of the active sound. Define the energy of *N* frame background noise as E_{noise}:

$$TH = b \times E_{noise} \tag{1}$$

$$TL = a \times E_{noise} \tag{2}$$

where a and b are constants obtained by experience setting or searching.

We first set the low threshold parameters a_1 and b_1 and perform DTED on continuous sound for the first time to obtain single sound segments. Judging the duration of each sound segment, it will end if the durations of all sound segments are less than $T_{\max}$. Otherwise, we reset a_1 and b_1, set higher threshold parameters a_2 and b_2, and perform the second DTED for the sound segments with durations greater than $T_{\max}$. If the obtained detection probability is greater than the previous probability, continue the above steps until the duration of all sound segments is less than $T_{\max}$ or the number of DTED is greater than 5. $T_{\max}$ is the maximum duration of an individual cough, which is a statistic obtained from the experimental data. We use the linear search to optimise all the threshold parameters a and b.

In the research, we found that the missing frame has a greater impact on the classification performance than the error frame. We hope to minimise the number of missing frames. In this case, frame detection probability as the threshold search indicator cannot accurately reflect

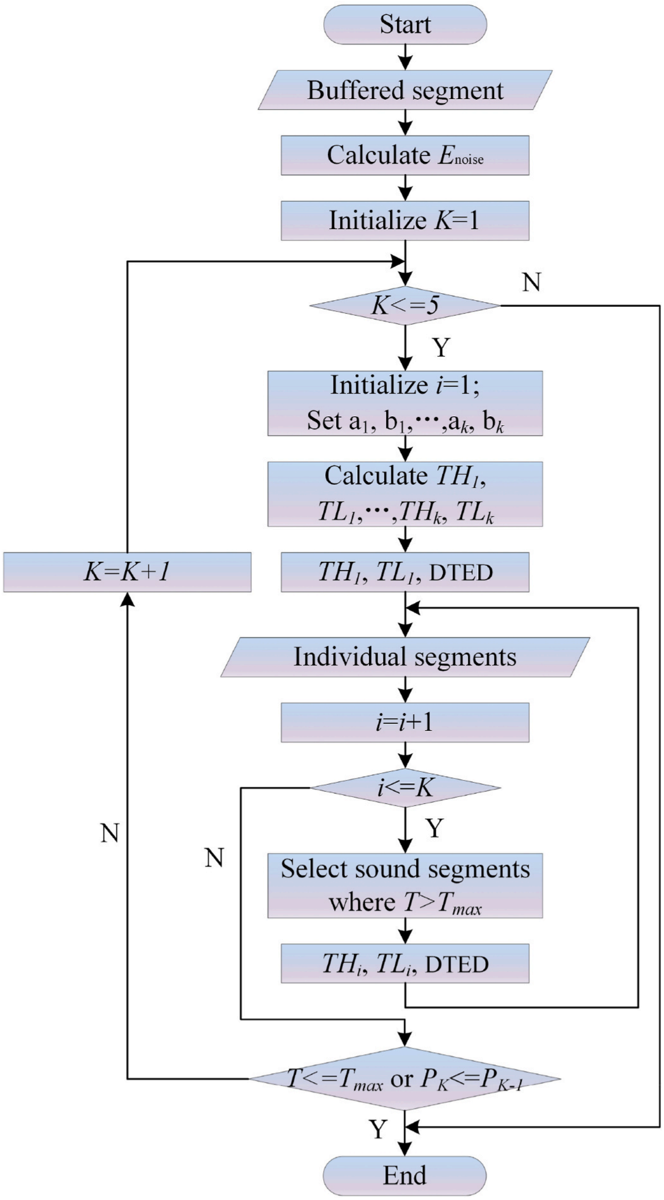

Fig. 4. Flowchart of the proposed improved VAD algorithm. K is used to count the total times of DTED. i is used to count the number of DTED. P_K denotes the threshold search indicator. TH_i and TL_i denote the high and low thresholds of the i^{th} DTED, and b_i and a_i denote the threshold parameters of TH_i and TL_i. T denotes the duration of the individual sound segment, T_{max} denotes the maximum duration of an individual cough.

the missing frame and error frame. Therefore, we define an indicator for the threshold search, as shown in Equ. (3),

$$P_0 = 1 - P_{miss} - 0.2 * P_{false} \tag{3}$$

where P_{miss} and P_{false} denote the frame miss detection probability and frame false detection probability, respectively. They are defined as follows:

$$P_{miss} = N_{miss} / N_f \tag{4}$$

$$P_{false} = N_{false} / N_f \tag{5}$$

where N_f represents the number of total frames of a sound segment and N_{miss} and N_{false} represent the number of missing frames and error frames, respectively.

Since the background noise changes rapidly, we update the noise energy in real time. The current noise energy is updated as

$$E_{noise} = \alpha \times E_{noise} + (1 - \alpha) \times E \tag{6}$$

where α denotes the update rate, which is generally taken as [0.9, 0.98], with a smaller α indicating a faster update rate. E is the energy of the current noise frames.

3.3. Improved classification algorithm

The complex piggery environment presents a great challenge to accurately recognise pig cough. It is difficult to achieve high accuracy by relying only on a single classifier. The fusion of multiple classifiers provides an effective way to enhance the accuracy (Cui et al., 2023; Ding, Wu, & Nugent, 2023; Jiang et al., 2021; Velpula & Sharma, 2023). We have performed research on multi-classifier fusion and achieved an accuracy of 99.2% based on an individual sound dataset (Yin et al., 2023). In that work, we constructed multiple base classifiers from multiple features and one classifier. However, the dataset was relatively simple compared to the dataset in this work. The number of non-coughs was limited, and complex coughs were not considered. Meanwhile, the features and classifiers considered in that method were limited. To enhance the robustness of the model against complex datasets, we provided more distinctive features and more robust classifiers for ensemble learning, and optimise the selection of all base classifier combinations.

We treat the previously proposed classifier fusion method as the baseline method. In the next section, we first introduce the framework of the baseline method and then present the proposed improved classifier fusion algorithm.

3.3.1. Baseline multiple classifier fusion method

The framework of the baseline method (Yin et al., 2023) is shown in Fig. 5. We preprocess the individual sound segments and then extract acoustic features and deep features. The acoustic feature vectors are formed by MFCC and linear prediction cepstral coefficient (LPCC). The deep features are extracted from the mel-spectrogram using fine-tuned Alexnet (Krizhevsky, Sutskever, & Hinton, 2017). We fused the MFCC and LPCC, then input the fused acoustic features and deep features into two SVMs, and finally performed soft voting to make a decision.

3.3.2. Improved multiple classifier fusion method

The framework of the improved multiple classifier fusion method is shown in Fig. 6. We carefully chose the more distinctive features and more robust classifiers, and design a classifier selection algorithm to optimise the selection of all base classifier combinations. We first preprocess the sound segments and extract the acoustic features and deep features. We found that LPCC performed poorly in complex piggery environments, so we chose MFCC and PSD as acoustic features. The deep features are extracted from the image features of the spectrogram, melspectrogram, constant-Q transforms (CQT) and MFCC colour matrix map. We used the lightweight network SqueezeNet (Iandola et al., 2016) instead of AlexNet to extract deep features to reduce the complexity. By feeding acoustic features and deep features into different classifiers, we can obtain multiple classification models, where the classifiers include SVM, adaptive boosting (AdaBoost) and bidirectional long short-term memory (BiLSTM). The classifier selection method is used to select the base classifier for fusion. Soft voting is used for the final decision.

3.3.3. Evaluation metrics for classifier selection

In the multi-classifier fusion algorithm, the selection of the base classifier is the key step that affects the classification performance. When selecting a base classifier, we should not only choose the one with high accuracy but also choose the one with great difference between each base classifier (Sagi & Rokach, 2018; Shiue, You, Su, & Chen,

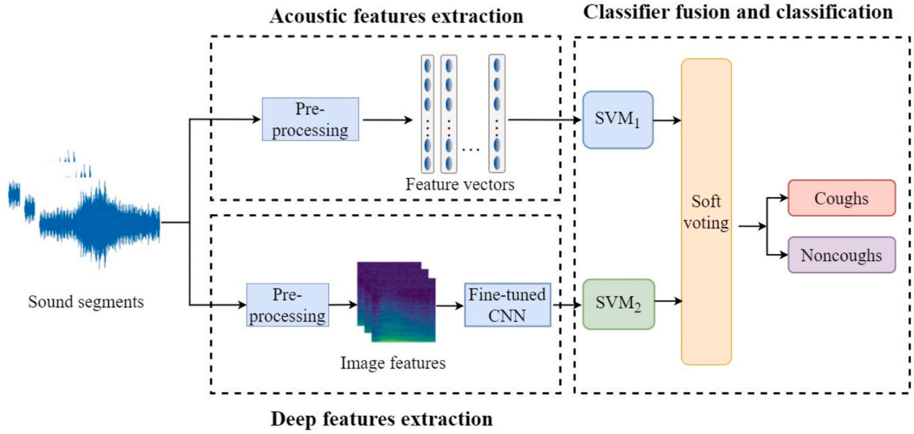

Fig. 5. Framework of the baseline classification method.

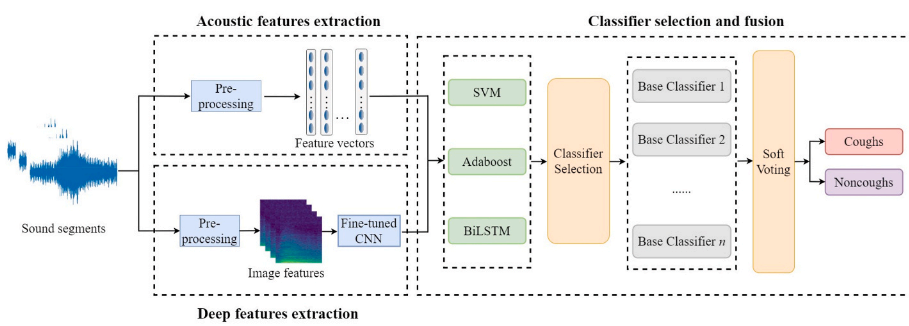

Fig. 6. Framework of the proposed classification method.

2021). In this study, we defined an accuracy–diversity (AD) metric to select the base classifiers.

$$AD_{ij} = \left(Acc_i + Acc_j\right) / 2 \pm \beta * Dv \tag{7}$$

where *Acc* is the accuracy of the base classifier and *Dv* is the diversity metric. The commonly used diversity evaluation metrics include Q statistics, correlation coefficients, disagreement measures, and double fault measures (Kuncheva & Whitaker, 2003). If a large *Dv* indicates poor diversity, the formula uses a negative sign; otherwise, it uses a positive sign. β is a weight for balancing the accuracy and diversity. Let the maximum and minimum values of *Acc* in all base classifiers be $Acc_{\max}$ and $Acc_{\min}$, respectively, and let the maximum and minimum values of the diversity measure be $Dv_{\max}$ and $Dv_{\min}$, respectively; then, β can be calculated as follows.

$$\beta = \frac{Acc_{max} - Acc_{min}}{Dv_{max} - Dv_{min}} \tag{8}$$

3.3.4. Classifier selection

The complexity of the piggery environment leads to a great diversity of data samples, which in turn leads to a certain fluctuation in the diversity between base classifiers. Therefore, the classifier selection algorithm should not only be efficient but also stable and reliable. The flowchart of the proposed base classifier selection algorithm is shown in Fig. 7. Suppose there are n base classifiers $C_1, C_2, \ldots, C_n$. We start with each base classifier in turn and select the second base classifier according to the classifier selection metric AD. Then, we calculate the average AD between the remaining base classifiers and these two base classifiers and select the third base classifier. Loop this process until n base classifiers are selected. We evaluate all combinations selected by the algorithm and choose the optimal combination as the final fusion scheme. The method is much more efficient than selecting all combinations. We also used data perturbation (Jiang et al., 2021) to enhance the diversity of the selected base classifier combinations.

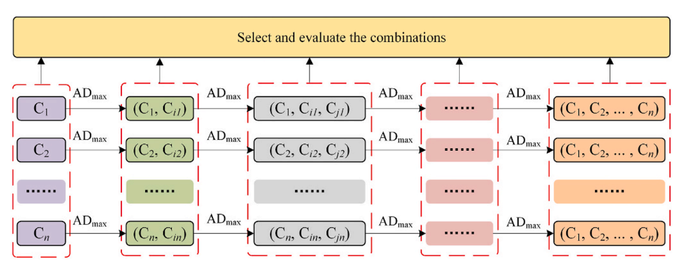

Fig. 7. Flowchart of the base classifier selection algorithm based on AD metrics. C_{ij} represents the i^{th} base classifier selected when C_j is the first base classifier.

3.4. Detection of continuous coughs

The continuous cough detection process is shown in Fig. 8. The individual segments are first detected by the proposed VAD from the buffered segment. Classify each individual segment and determine whether it is a cough. If it is a non-cough, then proceed to make the classification. If the number of coughs detected is greater than or equal to three and the time intervals between two neighbouring coughs are less than the maximum cough interval, it is judged to be a continuous cough; otherwise, it is a single cough. The maximum cough interval is a statistic obtained from experimental data.

3.5. Evaluation metrics

There are four groups of evaluation metrics in different processes.

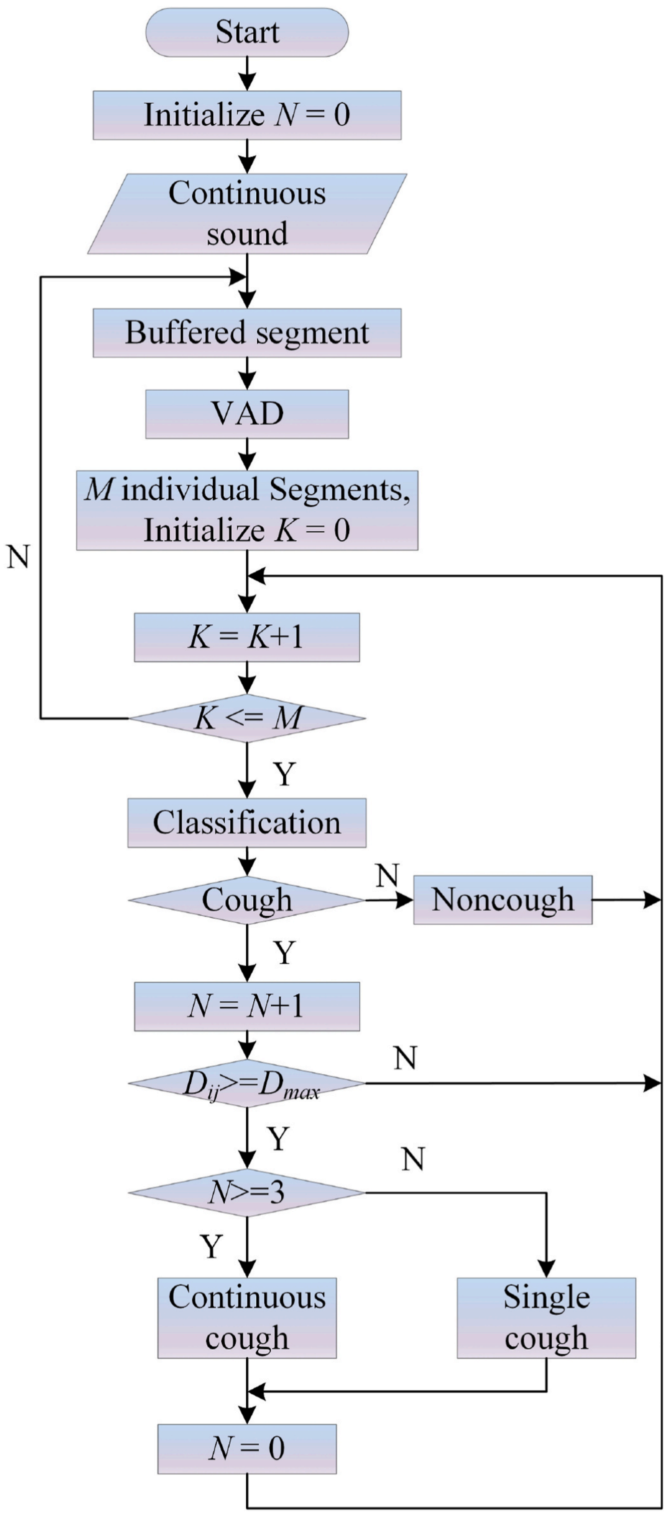

Fig. 8. Continuous cough detection process. Assume there are M individual sounds in the buffered segment. K is used to count the number of individual sounds, and N is used to count the number of continuous coughs. D_{ij} and $D_{\max}$ denote the time interval and the maximum time interval between two neighbouring coughs, respectively.

They are shown in Fig. 9. We evaluate VAD and individual sound classification algorithm used metrics 1 and 2, respectively. The continuous sound automatic detection process consists of VAD and individual sound classification, it is evaluated by metrics 3. The metric 4 is used for the evaluation of continuous cough recognition.

3.5.1. Evaluation metrics for VAD (metrics 1)

According to the definition of frame miss detection probability P_{miss} and frame false detection probability P_{false} in Equ. (4) – (5), we can obtain the frame detection probability P_d, which is defined as

$$P_d = 1 - P_{miss} - P_{false} \tag{9}$$

Furthermore, we also define a cough detection probability in units of word vector, which is defined as

$$P_{wd} = N_{wd} / N_w \tag{10}$$

where N_{wd} and N_w represent the number of detected coughs and the total number of coughs, respectively. Here, we do not consider the frame error; as long as a sound segment is detected as a cough, it is counted as a correctly identified cough.

3.5.2. Evaluation metrics for classification (metrics 2)

We used four widely used evaluation metrics: accuracy, recall, precision, and F1_score to evaluate the individual sound classification. Let the number of coughs correctly identified as coughs be *TP*, the number of non-coughs correctly identified as non-coughs be *TN*, the number of coughs wrongly identified as non-coughs be *FN*, and the number of non-coughs wrongly identified as coughs be *FP*; then, we have

$$Accuracy = (TP + TN) \,/\, (TP + TN + FP + FN) \tag{11}$$

$$Recall = TP/(TP + FN) \tag{12}$$

$$Precision = TP \,/\, (TP + FP) \tag{13}$$

$$F1\_score = 2 \times (Precision \times Recall) \,/\, (Precision + Recall) \tag{14}$$

3.5.3. Evaluation metrics for continuous sound automatic detection (metrics 3)

We used recall (Recall_C) and precision (Precision_C) to evaluate the continuous sound automatic detection system. Let the number of individual coughs contained in all continuous sounds is N_{cs}. Considering the missed detection of VAD and the misidentification in single sound classification, the number of correctly detected cough is N_{dcs}. The number of non-coughs detected as coughs is N_{fcs}, then we have:

$$Recall\_C = N_{dcs} / N_{cs} \tag{15}$$

$$Precision\_C = N_{dcs} \,/\, \left(N_{dcs} + N_{fcs}\right) \tag{16}$$

3.5.4. Evaluation metrics for continuous cough recognition (metrics 4)

We used the word error rate (WER) in continuous speech recognition as the evaluation metric of pig continuous cough recognition. The WER is defined as:

$$WER = (S + I + D) \,/\, N_c \tag{17}$$

where S is the number of substitutions, I is the number of insertions, D is the number of deletions and N_c is the total number of coughs in a continuous cough. Suppose a continues cough is marked as (cough, cough, cough, cough). If the detection result is (cough, cough, cough, cough, cough), an extra cough is an insertion error. If the detection result is (cough, cough, non-cough, cough), the non-cough is a substitution error. If the detection result is (cough, cough, cough), the missing cough is a deletion error.

We also define a continuous cough detection accuracy to evaluate the performance of the proposed continuous cough detection algorithm. It is defined as

$$P_{dc} = 1 - P_{mc} - P_{fc} \tag{18}$$

where P_{mc} is the missed detection probability of continuous cough, and P_{fc} is the false detection probability of continuous cough. They are defined as

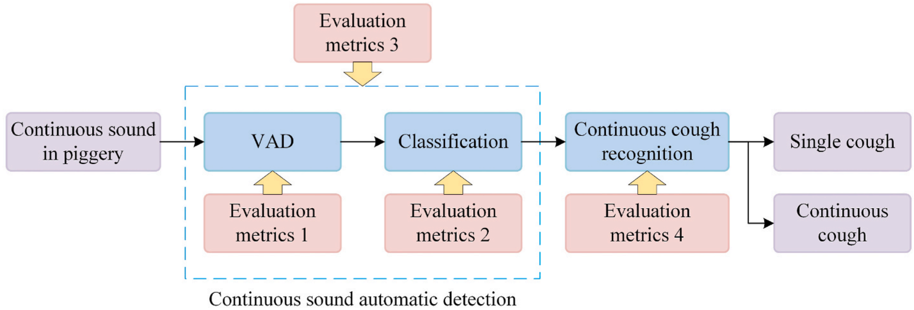

Fig. 9. Evaluation metrics in this study.

$$P_{mc} = N_{mc}/N_{cc} \tag{19}$$

$$P_{fc} = N_{fc}/N_{cc} \tag{20}$$

where N_{cc} represents the total number of continuous coughs, N_{mc} represents the number of missed detections of continuous cough, and N_{fc} represents the number of false detections of continuous cough. In a continuous cough, even if one or more single coughs are wrongly identified, as long as one continuous cough is identified, it is judged to be correctly identified. For example, there is a continuous cough expressed as (cough, cough, cough, cough, cough, cough, cough). If we wrongly identify it as (cough, cough, cough, cough, non-cough, non-cough) and detect the first four coughs as a continuous cough, we still assume that the continuous cough is correctly identified.

4. Results

In this section, we describe the experimental results of this study. First, we describe the experimental parameter settings. Second, we analyse the performance of improved VAD and classification method respectively. Third, we describe the performance of continuous sound automatic detection. Finally, we present the results of continuous cough recognition.

4.1. Experimental parameters

The experiments were conducted using a configuration of an Intel (R) Core (TM) i5-11400H CPU running at 2.70 GHz with 16 GB of memory and an NVIDIA GeForce RTX 3050 GPU with 4 GB of memory. The software used in this work was MATLAB 2021b.

The tool box used in this study included the *Deep Learning Toolbox* and the *Statistics and Machine Learning Toolbox*. We carefully chose all parameters by the optimization process and selected the parameters from the perspective of computation and performance. The SVM 'kernelFunction' was set to 'RBF', and the 'kernelScale' was set to 'auto'. The AdaBoost 'NumLearningCycles' was set to 500, the 'Learners' was 'templateTree' and the 'MaxNumSplits' was set to 5. In the preprocessing, the pre-emphasis coefficient was 0.9375, the FIR bandpass filter range was 100–16000 Hz, the time duration of each frame was 20 ms, the overlap was 10 ms, and the Hamming window was used. The dimension of MFCC was 13, and the fast Fourier transform window length was 1024 when extracting PSD. The parameters used for fine-tuning the pretrained convolutional neural network were as follows: batch size: 128, learning rate: 0.0001, epoch: 30, optimiser: Adam, and loss function: cross entropy. The parameters used for BiLSTM were as follows: batch size: 64, learning rate: 0.001, epoch: 300, number of hidden layers: 200, optimiser: Adam, loss function: cross entropy, and output mode: 'last'.

4.2. Results of VAD

In the detection of pig cough, we do not consider the sound clips below and above the length of the cough to filter out part of the interference signal. According to the data in the database, we calculated the statistical range of cough duration to be 0.22 – 0.97 s. Therefore, we set the effective sound segment length to 0.2 – 1.1 s in the VAD process. For the proposed VAD method, we obtained three groups of thresholds: (2, 2.4), (5, 5.5), and (15, 16). The update rate α was set to 0.92.

We compared the proposed method with the DTED method, the double threshold frequency band variance-based detection (DTFBVD) method, and the double threshold MFCC inverse spectral distance-based detection (DTMISD) method (Zhang, Shao, Wu, Geng, & Fan, 2020). The results are shown in Table 2. It can be seen that the traditional methods show very obvious limitations in complex environments, and the cough detection probability is very low, with the highest being the DTED method with a word detection probability of only 75.8%. Our method achieves a cough detection probability of 97.6%, which is significantly higher than that of traditional methods. Although the frame false detection probability of the proposed method is slightly higher, the frame miss detection probability is greatly reduced, resulting in a higher frame detection probability of 66.7%. Analysing the results, we found that most of the unsuccessfully detected coughs were disturbed by strong interference.

To more clearly describe the advantages of the proposed VAD algorithm, we gave an energy curve of a continuous sound, as shown in Fig. 10. The solid red line represents the high threshold, and the dashed red line represents the low threshold. For the traditional double threshold method, one group threshold is generally selected. A high threshold will cause some sound segments to be missed, as shown in the blue box, and a low threshold will cause multiple sounds to be detected as one sound, as shown in the green box. The proposed VAD method can solve this contradiction and enhance the detection probability.

4.3. Results of classification

We selected two acoustic features, four image features and three classifiers for constructing the base classifier. There will be eighteen combinations for selection. We first chose eight base classifiers for further selection and fusion according to the rank of accuracy and the diversity of features. Deep features were extracted from image features by SqueezeNet. The performance of the chosen base classifiers is shown in Table 3.

Next, we further select the optimal number of fusion classifiers and the fusion scheme. According to the accuracy–diversity evaluation metrics and classifier selection method described in Section 3.3, the

Table 2
Performance of the proposed VAD method and traditional method.

Method	P_{wd}	P_{miss}	P_{false}	P_d
DTED	75.8%	33.7%	13.0%	53.3%
DTFBVD	75.3%	34.3%	12.5%	53.2%
DTMISD	69.2%	43.3%	12.0%	44.7%
Proposed method	97.6%	13.3%	20.0%	66.7%

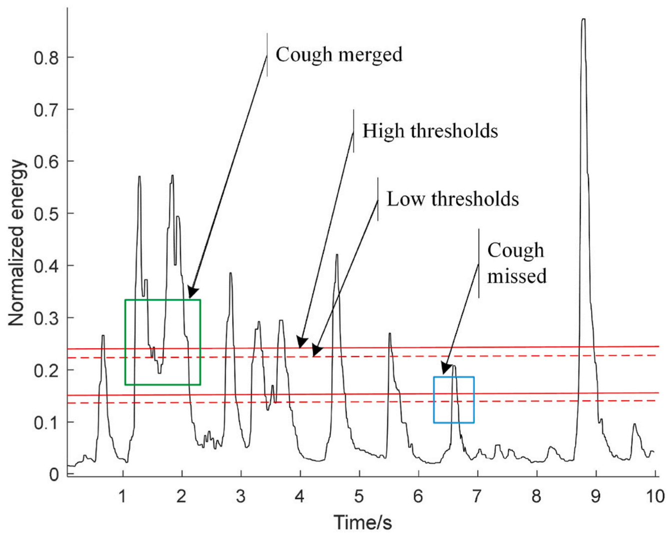

Fig. 10. Energy curve of a continuous sound. The solid red line represents the high threshold, and the dashed red line represents the low threshold. The blue square indicates the case of misdetection, and the green square indicates the case of multiple signals detected as one. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

accuracy of the optimal combination for the fusion of different numbers of base classifiers is given in Fig. 11, where the horizontal coordinate represents the number of fusion base classifiers, and the vertical coordinate represents the accuracy of the selected optimal base classifier combination. It can be seen that Q statistics perform best in all combinations. Correlation and disagreement have similar performance. Double faults present different performances with different numbers of fusion classifiers. When the number of fusion classifiers is seven, the four indicators achieve the same accuracy. To further validate the reliability of our proposed selection method and its adaptability to complex environments, we fused all the base classifier combinations and found that the base classifier combinations obtained by our selection method are basically the optimal combinations when using Q statistics.

We compared the proposed multi-classifier fusion method (Q statistics) with the baseline method (Yin et al., 2023), and the results are shown in Table 4. The baseline achieved an accuracy of 99.2% in the previous study. However, due to the complexity of the database in this study, the accuracy dropped to 92.1%. The proposed method achieved 93.2% with two classifiers fused. As the number of fused base classifiers increases, the accuracy gradually increases and drops slightly when fusing eight classifiers. The highest accuracy of 94.3% is achieved when seven base classifiers are fused. It can be seen from the results that by increasing the diversity of features and classifiers and optimizing the selection and fusion of base classifiers, the anti-interference ability and the accuracy of the model can be effectively improved.

4.4. Results of continuous sound automatic detection

The performance of continuous sound automatic detection is shown in Table 5. Among them, the baseline detection method is composed of the DTED algorithm and the baseline classification algorithm. The proposed detection method is composed of the improved VAD algorithm and the improved multi-classifier fusion algorithm. It can be seen that the proposed continuous sound automatic detection algorithm achieved high detection accuracy, the recall and precision are 93.1% and 91.6%, respectively, which are significantly higher than the baseline method. Due to the poor VAD performance of the baseline method, the recall of baseline method is only 67.3%, which is much lower than that of proposed method. The proposed method ensures the accuracy of VAD and classification, hence, realises the high-precision automatic detection of cough sound.

4.5. Results of continuous cough recognition

The continuous pig cough recognition results are shown in Table 6. According to the statistics, the maximum time interval between two neighbouring coughs was set to 1.5 s. It can be seen that with the increase in the number of fusion classifiers, the WER significantly

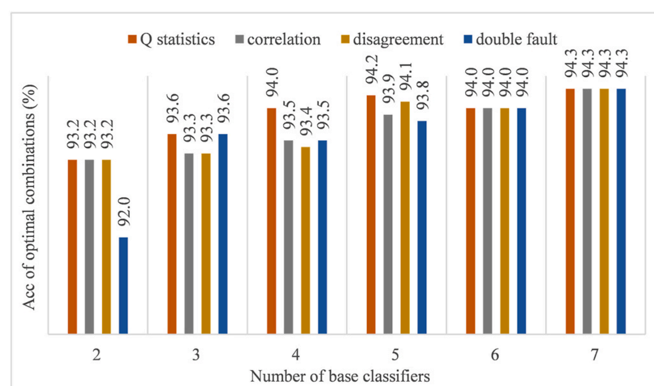

Fig. 11. Performance of different diversity indicators.

Table 4
Performance of multiclassifier fusion.

Method	Number of base classifiers	Accuracy	Recall	Precision	F1-score
Baseline	2	92.1%	89.8%	91.9%	90.8%
Proposed method	1	91.7%	94.7%	87.2%	90.8%
	2	93.2%	94.8%	89.9%	92.3%
	3	93.6%	95.9%	90.0%	92.9%
	4	94.0%	95.6%	91.0%	93.2%
	5	94.2%	95.4%	91.6%	93.5%
	6	94.0%	96.0%	90.8%	93.3%
	7	94.3%	95.6%	91.7%	93.6%
	8	93.9%	95.8%	90.7%	93.2%

Table 3
Performance of eight selected base classifiers.

Number	Feature	Classifier	Accuracy	Recall	Precision	F1-score
1	MFCC	SVM	90.1%	89.3%	88.0%	88.6%
2	MFCC	BiLSTM	88.9%	90.1%	85.3%	87.6%
3	PSD	AdaBoost	87.9%	86.3%	85.9%	86.1%
4	PSD	BiLSTM	89.6%	90.3%	86.4%	88.3%
5	Spectrogram	SVM	91.7%	94.7%	87.2%	90.8%
6	MelSpectrogram	SVM	90.7%	90.5%	88.4%	89.4%
7	CQT	SVM	90.9%	92.2%	87.5%	89.8%
8	MFCC	SVM	89.9%	91.2%	86.4%	88.7%

Table 5
Performance of continuous sound automatic detection.

Method	P_{wd}	Accuracy	Recall_C	Precision_C
Baseline method	75.8%	92.1%	67.3%	90.6%
Proposed method	97.6%	94.2%	93.1%	91.6%

Table 6
Performance of proposed continuous cough recognition.

Number of base classifiers	P_{dc}	P_{mc}	P_{fc}	WER
1	87.6%	3.3%	9.1%	12.0%
2	91.0%	2.9%	6.2%	10.6%
3	91.0%	2.4%	6.7%	8.8%
4	90.0%	3.8%	6.2%	8.2%
5	91.4%	2.9%	5.7%	7.9%
6	88.6%	3.3%	8.1%	8.6%
7	90.5%	2.9%	6.7%	8.2%
8	89.5%	3.3%	7.1%	9.0%

declined, and the minimum WER was 7.9%. The best detection accuracy of continuous cough recognition reached 91.4%. The results demonstrated that the proposed method is effective and reliable in complex piggery environments.

5. Discussion

The major contribution of this work is to realise the automatic detection of a continuous sound and the recognition of a continuous cough sound in complex piggery environment, which is of great significance to the engineering application of the algorithm and the diagnosis of diseases. Most of the previous research work has focused on the classification of individual sounds, ignoring the recognition process of continuous sounds. Although high precision has been achieved in the individual sound classification process, the precision loss caused by automatic sound segmentation has not been taken into account. Through the study of this paper, we found that the detection probabilities of the traditional VAD methods were very low in complex piggery environment, which led to a high miss detection probability of pig cough. Li et al. (2019) and Zhao et al. (2020) investigated the BiLSTM-connectionist temporal classification model and deep neural networks-hidden Markov model for continuous sound recognition. In these works, they also used traditional DTED method for VAD and they did not consider the miss detection in the VAD process.

In the case of a high SNR and no other sound overlap, the VAD algorithm proposed in this paper is equivalent to the DTED method, and the detection accuracy can reach more than 99%. However, in a large-scale piggery, the SNR changes rapidly with time, and the cough sound is easily overlapped with other sounds, resulting in a low detection probability of traditional VAD methods. We also tried other state-of-the-art VAD methods, and they all performed poorly in complex piggery environments.

In previous studies, only several limited features and classifiers were considered in the multi-classifier fusion. In this work, we fully considered more features and classifiers, and selected more distinctive features and more robust classifiers for ensemble learning. Although the LPCC feature performed well in our previous studies, we found that it is sensitive to noise in the complex dataset. Therefore, we did not use it in this study. We also tried other classifiers, such as k-nearest neighbour and random forest. Their performance was slightly inferior to or close to that of SVM, so we chose only the SVM classifier. Meanwhile, for the selection of the base classifier, empirical selection may result in suboptimal fusion results, and the calculation of all combinations will lead to a large computational complexity. The classifier selection algorithm proposed in this paper greatly reduces the computational complexity while ensuring the classifier fusion performance.

At present, there is no public pig cough datasets, and all researches were based on their private datasets. The datasets in this paper fully considers the complexity of the actual piggery environment, making the proposed method more suitable for practical engineering applications. Datasets have a great influence on classification results. The same algorithm shows different classification accuracies on different datasets. We test the proposed classification method used the simple dataset in the previous study (Yin et al., 2023), the classification accuracy was also higher than 99%. However, the cough recognition accuracy dropped to 94.2% in our new designed complex dataset which is closer to the sound in a real pig barn environment. It is still 2.1% higher than the baseline (92.1%) method. The classification accuracy is affected by the number of samples in the dataset. If the number of samples is too low or unbalanced, the classification accuracy may be reduced. In this research, we did not further label the non-cough. Since there are many kinds of non-coughs, we cannot guarantee that the number of each non-cough is the same. We tried to increase the number of non-coughs in the training dataset and found that with the increase in the number of non-coughs, the accuracy gradually increased, and when it reached approximately 2000, the accuracy basically remained stable.

In our previous studies (Shen ,Ji, et al., 2022; Yin et al., 2023), we have deeply discussed the performance of different neural networks with different inputs of image features, and we concluded that shallow convolutional neural networks (AlexNet) performed better than deep convolutional neural networks (VGG16, VGG19, ResNet152), and mel-spectrograms performed slightly better than CQTs and spectrograms. To reduce the complexity and facilitate engineering implementation, we replaced AlexNet with the lightweight SqueezeNet. The experimental results reveal that SqueezeNet and AlexNet have similar performance. In the ensemble learning model, we used the SVM and the lightweight deep learning model to reduce the occupied memory. The model has good practicability in engineering application. Although the performance gradually improves with the increase in the number of fusion classifiers, the complexity will also increase. In practical applications, we can consider using two or three base classifiers to compromise the complexity and accuracy.

In the future, we will try to explore the continuous cough automatic detection method based on end-to-end model to further improve the accuracy and reduce the complexity of the model. Meanwhile, we would like to further study the recognition of different types of coughs, such as dry cough and wet cough. Additionally, we will consider the study of pig cough recognition and location based on microphone array.

6. Conclusions

In this study, we investigated an automatic detection method of continuous pig cough in a complex piggery environment. The proposed improved VAD algorithm can effectively solve the limitations of traditional methods by using multiple adaptive dynamic thresholds and greatly improves the detection probability, with a result of 97.6%. The improved multi-classifier fusion method effectively enhanced the recognition accuracy in complex environments, with an accuracy of 94.3%. Under the condition of improving the accuracy of VAD and classification, the proposed continuous sound automatic detection method obtained a recall and precision of 93.1% and 91.6% respectively, and the proposed continuous cough recognition method obtained a high accuracy of 91.4% and a low WER of 7.9%. The research of this paper provides strong technical support for practical engineering applications. In future work, we will further improve the accuracy and reduce the complexity of the model.

CRediT authorship contribution statement

Xipeng Wang: Conceptualization, Software, Methodology, Writing – original draft, Investigation, Formal analysis. **Yanling Yin**: Writing – review &editing, Methodology, Visualization, Funding acquisition.

Xinpeng Dai: Software, Methodology, Data curation. **Weizheng Shen**: Conceptualization, Methodology, Funding acquisition, Investigation. **Shengli Kou**: Supervision, Visualization, Data curation. **Baisheng Dai**: Software, Funding acquisition.

Funding

This work was supported by the project of the National Natural Science Foundation of China [grant number 32172784, 31902210]; the Outstanding Youth Program of the Natural Science Foundation of Heilongjiang Province of China (grant number YQ2023C012); the University Nursing Program for Young Scholars with Creative Talents in Heilongjiang Province (grant number UNPYSCT-2020092); the Academic Backbone Project of Northeast Agricultural University; and the earmarked fund (grant number CARS-36, CARS-35).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Assavacheep, P., & Thanawongnuwech, R. (2022). Porcine respiratory disease complex: Dynamics of polymicrobial infections and management strategies after the introduction of the African swine fever. *Frontiers in Veterinary Science, 9*, Article 1048861. https://doi.org/10.3389/fvets.2022.1048861

Chung, Y., Oh, S., Lee, J., Park, D., & Kim, S. (2013). Automatic detection and recognition of pig wasting diseases using sound data in audio surveillance systems. *Sensors, 13*, 12929–12942. https://doi.org/10.3390/s131012929

Cui, L., Chen, S., Mu, Y., Xu, X., Zhang, B., & Zhao, X. (2023). Tree sSpecies classification over Cloudy Mountainous Regions by Spatiotemporal fusion and ensemble classifier. *Forests, 14*(1), 107. https://doi.org/10.3390/f14010107

Ding, W., Wu, S., & Nugent, C. (2023). A multimodal fusion enabled ensemble approach for human activity recognition in smart homes. *Health Informatics Journal, 29*(2), Article 14604582231171927. https://doi.org/10.1177/14604582231171927

Exadaktylos, V., Silva, M., Aerts, J. M., Taylor, C. J., & Berckmans, D. (2008). Real-time recognition of sick pig cough sounds. *Computers and Electronics in Agriculture, 63*, 207–214. https://doi.org/10.1016/j.compag.2008.02.010

Guarino, M., Jans, P., Costa, A., Aert, J. M., & Berckmans, D. (2008). Field test of algorithm for automatic cough detection in pig house. *Computers and Electronics in Agriculture, 62*, 22–28. https://doi.org/10.1016/j.compag.2007.08.016

Hirtum, A. V., & Berckmans, D. (2003). Fuzzy approach for improved recognition of citric acid induced piglet coughing from continuous registration. *Journal of Sound and Vibration, 266*, 677–686. https://doi.org/10.1016/s0022-460x(03)00593-5

Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv preprint arXiv:160207360. doi:10.48550/arXiv.1602.07360.

Ji, N., Shen, W., Yin, Y., Bao, J., Dai, B., Hou, H., et al. (2022). Investigation of acoustic and visual features for pig cough classification. *Biosystems Engineering, 219*, 281–293. https://doi.org/10.1016/j.biosystemseng.2022.05.010

Ji, N., Yin, Y., Shen, W., Kou, S., Dai, B., & Wang, G. (2022). Pig sound Analysis: A measure of Welfare. *Smart Agriculture, 4*(2), 19. https://doi.org/10.12133/j.smartag.SA202204004

Jiang, F., Yu, X., Du, J., Gong, D., Zhang, Y., & Peng, Y. (2021). Ensemble learning based on approximate reducts and bootstrap sampling. *Information Sciences, 547*, 797–813. https://doi.org/10.1016/j.ins.2020.08.069

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90. https://doi.org/10.1145/3065386

Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine Learning, 51*, 181–207. https://doi.org/10.1023/A:1022859003006

Li, X., Zhao, J., Gao, Y., Lei, M., Liu, W., & Gong, Y. (2018). Recognition of pig cough sound based on deep belief nets. *Transactions of the Chinese Society for Agricultural Machinery, 49*(3), 179–186. https://doi.org/10.6041/j.issn.1000-1298.2018.03.022

Li, X., Zhao, J., Gao, Y., Liu, W., Lei, M., & Tan, H. (2019). Pig continuous cough sound recognition based on continuous speech recognition technology. *Transactions of the Chinese Society of Agricultural Engineering, 35*, 174–180. https://doi.org/10.11975/j.issn.1002-6819.2019.06.021

Racewicz, P., Ludwiczak, A., Skrzypczak, E., Składanowska-Baryza, J., Biesiada, H., Nowak, T., et al. (2021). Welfare health and productivity in commercial pig herds. *Animals, 11*, 1176. https://doi.org/10.3390/ani11041176

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8*(4), e1249. https://doi.org/10.1002/widm.1249

Sassu, E. L., Bossé, J. T., Tobias, T. J., Gottschalk, M., Langford, P. R., & Hennig-Pauka, I. (2018). Update on Actinobacillus pleuropneumoniae-knowledge, gaps and challenges. *Transboundary and Emerging Diseases, 65*, 72–90. https://doi.org/10.1111/tbed.12739

Shen, W., Ji, N., Yin, Y., Dai, B., Tu, D., Sun, B., et al. (2022a). Fusion of acoustic and deep features for pig cough sound recognition. *Computers and Electronics in Agriculture, 197*, Article 106994. https://doi.org/10.1016/j.compag.2022.106994

Shen, W., Tu, D., Yin, Y., & Bao, J. (2021). A new fusion feature based on convolutional neural network for pig cough recognition in field situations. *Information Processing in Agriculture, 8*, 573–580. https://doi.org/10.1016/j.inpa.2020.11.003

Shen, M., Wang, M., Liu, L., Chen, J., Tai, M., & Zhang, W. (2022b). Recognition method of pig cough based on deep neural network. *Transactions of the Chinese Society for Agricultural Machinery, 53*(5), 257–266. https://doi.org/10.6041/j.issn.1000-1298.2022.05.026

Shiue, Y. R., You, G. R., Su, C. T., & Chen, H. (2021). Balancing accuracy and diversity in ensemble learning using a two-phase artificial bee colony approach. *Applied Soft Computing, 105*, Article 107212. https://doi.org/10.1016/j.asoc.2021.107212

Song, H., Zhao, B., Hu, J., Sun, H., & Zhou, Z. (2022). Research on improved DenseNets pig cough sound recognition model based on SENets. *Electronics, 11*(21), 3562. https://doi.org/10.3390/electronics11213562

Velpula, V. K., & Sharma, L. D. (2023). Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion. *Frontiers in Physiology, 14*, Article 1175881. https://doi.org/10.3389/fphys.2023.1175881

Yin, Y., Ji, N., Wang, X., Shen, W., Dai, B., Kou, S., et al. (2023). An investigation of fusion strategies for boosting pig cough sound recognition. *Computers and Electronics in Agriculture, 205*, Article 107645. https://doi.org/10.1016/j.compag.2023.107645

Yin, Y., Tu, D., Shen, W., & Bao, J. (2021). Recognition of sick pig cough sounds based on convolutional neural network in field situations. *Information Processing in Agriculture, 8*, 369–379. https://doi.org/10.1016/j.inpa.2020.11.001

Zhang, T., Shao, Y., Wu, Y., Geng, Y., & Fan, L. (2020). An overview of speech endpoint detection algorithms. *Applied Acoustics, 160*, Article 107133. https://doi.org/10.1016/j.apacoust.2019.107133

Zhao, J., Li, X., Liu, W., Gao, Y., Lei, M., Tan, H., et al. (2020). DNN-HMM based acoustic model for continuous pig cough sound recognition. *International Journal of Agricultural and Biological Engineering, 13*(3), 186–193. https://doi.org/10.25165/j.ijabe.20201303.4530