ITC 1/55
Information Technology and Control
Vol. 55 / No. 1/ 2026
pp. 166-187
DOI 10.5755/j01.itc.55.1.42647

EGOOSE-LightGBM Model-Based Identity Recognition Using Cough Acoustic Signal Analysis

Received 2025/08/27 | Accepted after revision 2025/11/12

HOW TO CITE: Zhang, M., Zhu, P., Tang, R. (2026). EGOOSE-LightGBM Model-Based Identity Recognition Using Cough Acoustic Signal Analysis. *Information Technology and Control*, 55(1), 166-187. https://doi.org/10.5755/j01.itc.55.1.42647

EGOOSE-LightGBM Model-Based Identity Recognition Using Cough Acoustic Signal Analysis

Mingjian Zhang*
Department of Information Technology, Hunan Police Academy, Changsha 410138, China; e-mail: mingjianzhang@hnpa.edu.cn

Pengliang Zhu
Department of Information Technology, Hunan Police Academy, Changsha 410138, China; e-mail: zplzpl0929@163.com

Ru Tang
Department of Information Technology, Hunan Police Academy, Changsha 410138, China; e-mail: tangru2025@163.com

Corresponding author: mingjianzhang@hnpa.edu.cn

Speech signals-based identity recognition is an active research area, which has a wide range of important applications such as human–computer interaction, forensic investigation, sound surveillance, transaction authentication, and health monitoring. As a special speech event, cough plays a crucial supplementary role in identity recognition scenarios, especially when traditional biometric features are unavailable. However, compared to traditional speaker recognition, very few research efforts have been made by researchers to explore identity recognition using cough acoustic signal analysis. In this work, we propose an EGOOSE-LightGBM model for cougher recognition, in which the hyperparameters of LightGBM are optimized by EGOOSE algorithm. Chaotic mapping, Gaussian mutation, and crisscross optimization are employed in the proposed metaheuristic-type EGOOSE algorithm, which demonstrates superior performance in terms of convergence speed and accuracy compared to GOOSE algorithm. A static-dynamic feature fusion (SDFF) technique is used to fuse cough sound characteristics including Mel frequency cepstral coefficient (MFCC), fundamental frequency, formant,

ΔMFCC, and jitter of formant to improve recognition effectiveness and noise robustness. Using the self-recorded cough sounds, we establish a dataset, the size of which is enlarged by data augmentation techniques. Performance improvement of 3.167% is obtained due to the exploitation of cough data augmentation. Experimental results show that EGOOSE-LightGBM outperforms other existing machine learning models such as SVM, XGBoost, RF-Adaboost, LSTM, and LightGBM, achieving remarkable recognition accuracy of 99.500%.

KEYWORDS: Identity recognition, EGOOSE-LightGBM model, Static-dynamic feature fusion (SDFF), Cough data augmentation, Cough acoustic signal analysis.

1. Introduction

Coughing is a biological protective reflex action, occurring when the mucous membranes of the larynx, trachea, and bronchi are irritated by inflammation, allergies, or physical chemical stimuli. Cough is a well-known symptom as sociated with various respiratory disorders. Therefore, cough sound signals are widely used in disease detection/diagnosis [7], [18], [24], [30], [36], [38], and patient monitoring/caring [3], [40].

Identity recognition based on speech signals, termed speaker recognition (SR), is an active research field [10], [43], which has drawn wide attention from researchers. As a type of bioacoustics signal determined by physiological structures such as the vocal cords, trachea, lungs, and resonant cavities, cough sound exhibits individual specificity. Due to the distinct pronunciation organ and coughing style of each individual, a cougher's voice inherently encapsulates personal traits. Leveraging these individual-specific characteristics, it becomes feasible for computer-based systems to automate the process of identity recognition by analyzing the discriminative properties embedded in cough sound signals. Therefore, as a special speech event, cough sound signal has also been investigated for speaker recognition (SR) [46], cougher identification (CI) [22], [31], cougher verification (CV) [42], and security supervision.

The approach to identity recognition from cough sound signals mainly involves two crucial stages, i.e., the extraction of suitable features and the employment of classification algorithms to accurately recognize cougher's identity information. MFCC [18], Mel-spectrogram [3], [22], i-vector [31], and formant [28] are acoustic features widely used in cough acoustic signal analysis. In [31], Pahar et al. pointed out that i-vectors demonstrates superior performance compared to both d-vectors [46] and x-vectors for the task of cougher identification. The extracted features are subsequently fed to classification algorithms [31] such as support vector machine (SVM), multi-layer perception (MLP), convolutional neural network (CNN), long short-term memory (LSTM), ResNet, etc.

Zhang et al. [46] presented a deep neural network model to learn the deep acoustic feature for speaker recognition using cough, laugh, and "Wei" sound recordings. The approach in [46] achieved EER of 10.99% for cougher recognition. Joki's et al. [22] proposed a triplet network for cougher identification, in which the CNN model was designed to learn an embedding function that maps Mel-spectrograms of cough audio recordings into a discriminative embedding space. This approach enables easier differentiation between embedded samples compared to original data. Pahar et al. [31] recorded cough sounds under quiet and noisy acoustic conditions to build the dataset, and put forward a system to detect coughs among other triggering phrases and simultaneously identify the individual who is coughing. The system in [31] obtained an accuracy up to 99.91%. Whitehill et al. [42] employed a ResNet model for the task of cougher verification, which gave 82.15% accuracy for four users. Tran et al. [39] presented a contrastive learning-based network for identity recognition using both temporal and spatial features of cough sound signals, which obtained an accuracy of 97.18%. Notably, the implementation of the cross-entropy loss method contributed to the superior recognition performance. Laska et al. [25] employed a distillation technique to extract cough prints for identifying cough acoustic signals. In recent years, multimodal signal processing methods that incorporate cough sounds and other types of signals have drawn increasing attention [4], [20].

Forensic SR serves as evidence in criminal inquiries or judicial proceedings [5], [33], [41]. Using cough sounds for individual identification holds unique potential application value in forensics science. Compared to active vocal signals like speech, cough sounds are harder to deliberately disguise and easier to record in involuntary scenarios. Moreover, the production mechanism of cough sounds relies more on physiological instincts, which makes their individual characteristics more stable. Therefore, cougher recognition is expected to serve as a low-cost, non-invasive auxiliary tool in criminal forensics, particularly playing a supplementary role in scenarios where traditional biometric features are unavailable. From a feature perspective, on one hand, the duration of a cough sound is extremely short, typically ranging from 0.1 to 0.5 seconds. On the other hand, a cough sound has a broader frequency range than speech signal, making its frequency domain spectrum more dispersed.

In this work, we investigate identity recognition by using cough acoustic signal analysis. The major contributions of this work are summarized as follows.

- Unlike traditional speaker recognition [19], investigations on leveraging cough sound signals for the purpose of identity recognition remain scarce. In this paper, the EGOOSE-LightGBM model is proposed for cougher recognition, in which the hyperparameters of LightGBM are optimized with an enhanced GOOSE (EGOOSE) algorithm. We use three techniques, i.e., chaotic mapping, Gaussian mutation, and crisscross optimization, to enhance the basic LightGBM algorithm. As a result, the proposed EGOOSE algorithm is capable of effectively evading local optimum traps, boosting the search efficiency, and accelerating the convergence speed.
- In this paper, we adopt the static-dynamic feature fusion (SDFF) technique. The static features, i.e., MFCCs, formants (F1, F2, F3), and fundamental frequency, are fused with the dynamic features, i.e., ΔMFCC and jitter, to enhance the feature representation ability. Its main advantages include enhancing the model's ability to perceive the dynamic changes of features and improving noise robustness.
- A self-recorded cough sound dataset is established. Cough data augmentation techniques such as time stretching, addition of white noise, and pitch shifting are utilized to enrich the dataset. By introducing the varieties and diversities of characteristics of cough voice data, these augmentation methods enhance the model's cougher recognition accuracy, mitigate overfitting, and improve its generalizability.
- Comprehensive studies on different model parameters, different noise levels, and comparisons with the existing models are presented to demonstrate the remarkable performance of our proposed model.

The rest of this paper is organized as follows. Section 2 introduces the cough sound dataset we established. Feature extraction methods for cough sound signals are elaborated in Section 3. The principles and design process of the proposed EGOOSE-LightGBM model are presented in Section 4. Details on experimental setup procedure and a comprehensive analysis of experimental results are provided in Section 5. Finally, concluding remarks of this study are given in Section 6.

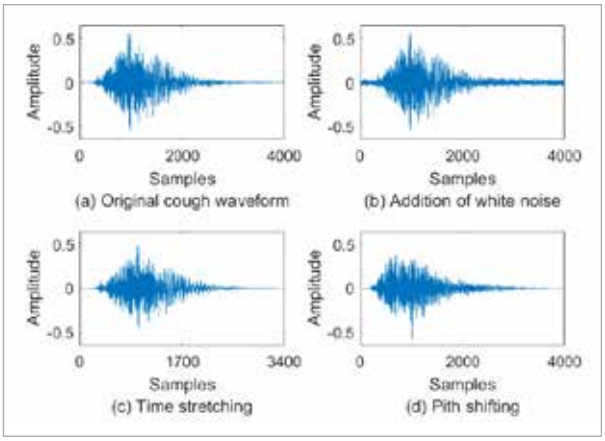
2. Cough Sound Dataset

2.1. Data Acquisition

We constructed a cough dataset using the self-recorded dry cough sounds. Dry cough is characterized by the absence of phlegm, whereas wet cough involves the expulsion of phlegm. Compared with wet cough, the sound of dry cough has higher stability and more prominent individual characteristics, which enables more accurate extraction of individual-specific voiceprint features. We recorded the dry cough sound signals using a microphone that was linked to a laptop equipped with a Conexant Smart Audio HD sound card. The recording software used for cough sound data collection was Cool Edit Pro. The cough sounds were recorded in a 5 m8 m meeting room (Figure 1), with the sampling rate of 16,000 Hz and resolution of 16 bits. The dataset consists of the cough sound recordings of 6 volunteers, including 3 male and 3 female coughers. Each cougher got a long cough sound recording, which was divided into short cough segments by using the voice activity detection (VAD) technique. The information for the six coughers and their cough sounds is given in Table 1.

Figure 1
Data acquisition environment for cougher recognition.

Figure 2
A segment of original cough sound waveform and the waveforms after data augmentation.



2.2. Cough Data Augmentation

The original dataset consists of 1200 short cough sound segments in total. Given its relatively small scale, it is susceptible to overfitting. In order to enrich the diversity of characteristics of cough sound and enlarge the size of dataset, we chose to use cough data augmentation technique.

Table 1
Gender information and number of segments of cough sound for six coughers.

Cougher	Gender	Number of Segments of Cough Sound
Cougher 1	female	225
Cougher 2	female	210
Cougher 3	female	202
Cougher 4	male	181
Cougher 5	male	182
Cougher 6	male	200

This diversifies the acoustic feature space during training, thereby curbing overfitting and enhancing the model's generalization capability. In this study, we adopted the following three types of augmentation options:

1 **Time stretching:** Its principle involves altering the temporal spacing of audio samples through techniques like resampling. This enables the adjustment of audio duration.

2 **Addition of white noise:** Characterized by a flat spectrum and Gaussian distribution, additive white noise is imposed onto cough sound by linearly combining its amplitude samples with the cough acoustic waveform.

3 **Pitch shifting:** As an audio signal processing technique, pitch shifting involves altering a sound's perceived pitch [19], [34]. It analyzes frequency components, shifts frequency components, and then resynthesizes the audio signal.

For time stretching augmentation, the cough sound is accelerated by 20% using a stretching rate of 1.2. For augmentation of addition of white noise, the value of signal-to-noise ratio (SNR) is set to be 15 dB. For pitch shifting augmentation, the cough sound is shifted by 6 semitones. As an example, a segment of original cough sound waveform and the waveforms after augmentation are shown in Figure 2.

After the implementation of cough data augmentation, the original dataset was enlarged to cover 4,800 segments of cough sound recordings in total.

3. Extraction of Static and Dynamic Features of Cough signals

3.1. Cough Signal Pre-processing

Over a relatively long period, cough acoustic signal exhibits non-stationary characteristics. However, within a very short time period, the characteristic

parameters of cough signal can be approximately regarded as unchanged. Therefore, cough signal is divided into a series of frames of 16 ms with 50% overlap. The i-th frame of the cough signal is represented as $s_i(n)$. Then, the Hamming window is applied to the framed cough signal to ensure smooth transitions at the frame edges:

$$x_i(n) = s_i(n) w_H(n). \quad (1)$$

3.2. Static Features of Cough Signals

3.2.1. MFCC

MFCC is a widely used short-term spectral feature in audio signal processing [14], [17] [32], [37]. The extraction process of MFCCs is shown in Figure 3. MFCC is based on the Mel scale analysis of cough signal. The Mel-frequency is derived from frequency as follows:

$$Mel(f) = 2595 \times \log_{10}\left(1 + \frac{f}{700}\right). \quad (2)$$

The graphical representation of Mel-scale filter bank is shown Figure 4.

$x_i(n)$ is transformed into frequency domain signal $X_i(k)$ by using Discrete Fourier Transform (DFT). The energy $E_i(k)$ of the corresponding spectral line is calculated as follows:

$$E_i(k) = [X_i(k)]^2. \quad (3)$$

Subsequently, $E_i(k)$ is filtered by the Mel-scale filter bank. The filtering operation is carried out by using $Y_i(m) = \sum_{k=0}^{N-1} E_i(k) H(m,k)$, wheren $H(m,k)$ denotes the m-th Mel scale filter in filter bank. Then, applying Discrete Cosine Transform (DCT) to $Y_i(m)$,we have the coefficients of MFCC, i.e.,

$$MFCC(i,n) = \sqrt{\frac{2}{M}} \sum_{m=0}^{M-1} \log_{10}\left(Y_i(m)\right) \cos\left[\frac{\pi n(2m-1)}{2M}\right], \quad (4)$$

where $n = 0,1,...,C-1$, and C denotes the number of MFCCs.

3.2.2. Formant

Formants contain abundant acoustic information [21], [28], [37], which is conducive to improving the recognition accuracy. Meanwhile, formants are less susceptible to noise interference, which can enhance the anti-noise performance of feature-based recognition. The extraction process of cough voice formants is shown in Figure 5. As can be seen from Figure 5, the first three steps are the same as those in the extraction process of MFCCs. Subsequently, we calculate $\hat{X}_i(k) = \log_{10}\left(\left|X_i(k)\right|\right)$. By computing inverse DFT (IDFT) of $\hat{X}_i(k)$, the cepstrum $\hat{x}_i(n)$ is obtained. Then, a window function $W_F(n)$ is applied to $\hat{x}_i(n)$, i.e., $h_i(n) = \hat{x}_i(n)\, W_F(n)$, where $W_F(n)$ is given as follows:

$$W_F(n) = \begin{cases} 1, & n \le b-1 \text{ or } n \ge N-b+1 \\ 0, & b-1 < n < N-b+1 \end{cases}, \quad (5)$$

in which b denotes the width of $W_F(n)$.

By computing the DFT of $h_i(n)$, we achieve the envelope curve $H_i(k)$. By identifying the local maxima of the amplitude on $H_i(k)$, the formants can be obtained. In this paper, we consider the lowest three formants ($F1$,$F2$,$F3$).

3.2.3. Fundamental Frequency

The period of vocal cord vibration is referred to as the fundamental period, and its reciprocal is the fundamental frequency, which is denoted as $F0$. The feature of $F0$ mainly represents the timbre information of a speaker and is of great significance in speaker recognition [37].

3.3. Dynamic Features of Cough Signals

The MFCC coefficients and formants described above only represent the static features of cough acoustic signals. However, as the human auditory system exhibits enhanced sensitivity towards the dynamic components of acoustic signals, the dynamic information embedded within vocal spectra holds valuable details that can improve recognition accuracy.

3.3.1. The First-Order Difference of MFCC

To capture the dynamic information of MFCCs, the first order difference of the MFCC (ΔMFCC) [11] is calculated as follows:

$$Dm(n) = \frac{1}{\sqrt{\sum_{i=-k}^{i=k} i^2}} \sum_{i=-k}^{i=k} i \cdot M(n+i) \quad (6)$$

Figure 3
Diagram of the computation of MFCCs.

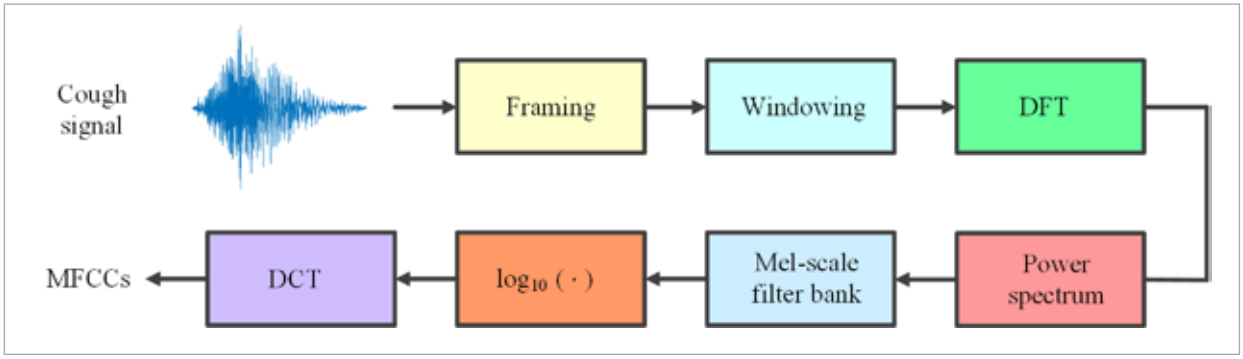


Figure 4
Filter bank on the Mel-scale.

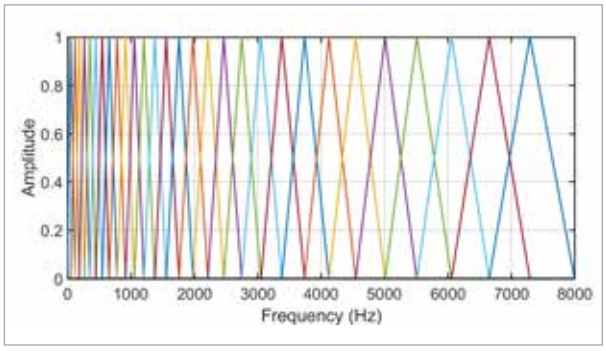
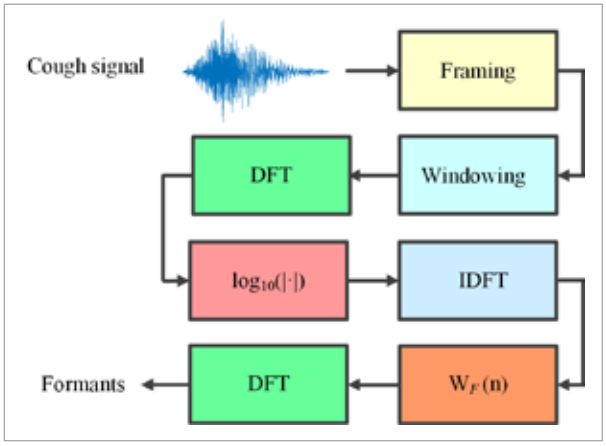


Figure 5
Diagram of the computation of formants.



where $M(n+i)$ denotes a frame of MFCC coefficients, and $k = 2$. It is worth noting that the features of the first and last frames should be removed because the first-order differences of these two frames are zeros.

3.3.2. Jitter of Formant

To capture the dynamic information of formants $F = [F1,F2,F3]$, we employ the jitter [6] defined as follows:

$$J(k) = \frac{\frac{1}{N-1}\sum_{i=1}^{N-1}\left|F_{ki} - F_{k(i+1)}\right|}{\frac{1}{N}\sum_{i=1}^{N}F_{ki}} \times 100, \quad (7)$$

where N is the number of frames of a cough voice segment, and F_{ki} is the k-th formant frequency of the i-th frame, k = 1,2,3

3.4. Static-Dynamic Feature Fusion

By conducting statistical analysis, we obtain the mean, minimum, maximum, variance of cough signal features described in the Subsections 3.2-3.3. We employ a technique called static-dynamic feature fusion to construct a hybrid feature vector *sd_vector* for cougher recognition. *sd_vector* has the form *sd_vector* = [*F*0_*vec*, *F*1_*vec*, *F*2_*vec*, *MFCC_vec*, Δ*MFCC_vec*, *jitter*],

where,

$$F0\_vec = [F0_{mean}, F0_{min}, F0_{max}, F0_{var}],$$
$$F1\_vec = [F1_{mean}, F1_{min}, F1_{max}, F1_{var}],$$
$$F2\_vec = [F2_{\text{mean}}, F2_{\text{min}}, F2_{\text{max}}, F2_{\text{var}}]$$
$$F3\_vec = [F3_{mean}, F3_{min}, F3_{max}, F3_{var}]$$
$$MFCC\_vec = [MFCC_{\text{mean}}, MFCC_{\text{min}}, MFCC_{\text{max}}, MFCC_{\text{var}}]$$

, and

$$\Delta MFCC\_vec = [\Delta MFCC_{mean}, \Delta MFCC_{\text{min}}, \Delta MFCC_{\text{max}}, \Delta MFCC_{\text{var}}]$$

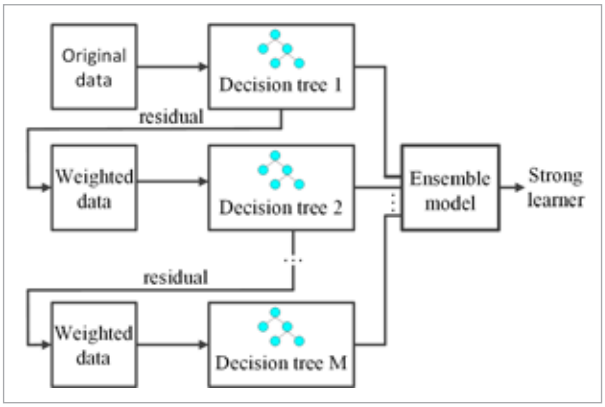
4. Proposed EGOOSE-LightGBM Method

4.1. LightGBM

LightGBM algorithm, proposed in [23], is a gradient boosting framework based on decision trees [13], [27], [47]. The diagrammatic sketch of gradient boosting decision tree architecture is shown in Figure 6. Due to its remarkable performance in terms of efficiency and accuracy, LightGBM is extensively utilized in a variety of machine learning (ML) tasks [2], [29], [44].

LightGBM employs the strategy of histogram optimization, the process of which [29] is shown in Figure 7. It constructs a histogram with a width of K. In terms of time complexity, histogram-based algorithm in LightGBM only requires iterating through K feature values (bins) when searching for the optimal split point. In terms of spatial complexity, LightGBM requires only 8 bits to store these K discrete values. In contrast, XGBoost [8] needs to use floating-point numbers to store all sorted feature values.

Figure 6
Gradient boosting decision tree architecture.



XGBoost adopts the level-wise growth strategy (Figure 8), which treats all leaves at the same tree level without discrimination. LightGBM instead employs the leaf-wise growth strategy (Figure 9) and imposes a constraint on the tree depth, which can alleviate overfitting.

Compared to XGBoost, LightGBM boasts several advantages, such as higher training efficiency, lower memory usage, and superior accuracy, particularly when dealing with high- dimensional features and large-scale datasets.

4.2. GOOSE Algorithm

The GOOSE algorithm, proposed in [16], is a novel meta heuristic optimization approach, which draws inspiration from the behavioral patterns of goose flocks in natural environments. Goose flocks exhibit distinct resting behaviors characterized by close huddling and unilateral leg stance. Particularly, some geese adopt the strategy of placing small stones on their raised legs. Once a goose falls asleep, the sound of the stone striking the ground serves as a wake-up signal. The GOOSE algorithm deeply simulates the social aggregation, adaptive monitoring, and efficient group communication behaviors of goose flocks, which constructs a computational model and algorithmic framework with robustness and extraordinary adaptability.

4.2.1. Exploitation Phase

The algorithm randomly determines the mass of the stone that the goose carries, and the weight varies from 5 to 25 g.

$$S_w = randi([5,25],1,1)\,. \tag{8}$$

In order to compute the average duration for a sound to propagate through a flock of geese and make a round trip, it is necessary to find out the overall time T_{total} of sound propagation.

$$T_{total} = \frac{\sum T_o + T_s}{dim}\,, \tag{9}$$

where T_o denotes the time needed for the stone to reach the ground, T_s denotes the time needed for a certain goose to pick up the sound of the stone striking the ground, and dim denotes the number of dimensions.

Subsequently, the average time T_a can be obtained as follows:

$$T_a = \frac{T_{total}}{2}\,. \tag{10}$$

The GOOSE algorithm assigns a variable *pro*, randomly chosen from the range of 0 to 1, for controlling the evolution process. In a flock of geese, there are behaviors of safeguarding and waking up individual members within the group. When the conditions $pro > 0.2$ and $S_w \geq 12$ are satisfied simultaneously, this type of behaviors are simulated by F_{fs}.

$$F_{fs} = T_o * \frac{\sqrt[2]{S_w}}{9.81}. \quad (11)$$

The distance D_t that sound propagates in the air is expressed as:

$$D_t = V_s * T_s. \quad (12)$$

where V_s denotes sound speed.

The distance D_g between a goose on guard duty and another goose can be represented as half of D_t, as it is sufficient to consider only the one-way propagation distance of the sound.

$$D_g = 0.5 * D_t. \quad (13)$$

For the purpose of waking individual members in a group of geese, it is necessary to find an optimal spot X_{it+1}, the formula of which is presented as follows:

$$X_{it+1} = F_{fs} + D_g * T_a^2. \quad (14)$$

On the other hand, if the conditions $S_w < 12$ or $pro \leq 0.2$ are satisfied, F_{fs} is revised into a new form:

$$F_{fs} = T_o * \frac{S_w}{9.81}. \quad (15)$$

Correspondingly, the new value of X is also adjusted to the following form:

$$X_{it+1} = F_{fs} * D_g * T_a^2 * Coe, \quad (16)$$

where Coe is a coefficient.

4.2.2. Exploration Phase

The goose starts screaming as it wakes up and then moves into the exploration phase. To enhance the search efficiency of the GOOSE algorithm, if $rnd < 0.5$, the formula for calculating the optimal position during this phase can be rewritten as follows:

$$X_{it+1} = randn(1, \dim) * M_t * alpha + Pos_{best}, \quad (17)$$

where Pos_{best} denotes the best position we found so far within the search region, and M_t represents the minimum total time. The parameter *alpha*, the formula of which is given in Equation (18), is employed to adjust the search step size, which in turn has an impact on the convergence of the algorithm. Specifically, it plays a crucial role in determining how the algorithm progresses during each iteration, and its proper adjustment can significantly contribute to achieving a more rapid and stable convergence of the algorithm.

$$alpha = 2 - \frac{2 * loop}{Max\_It}, \quad (18)$$

where Max_It denotes the maximum number of iterations performed in the GOOSE algorithm. The parameter loop stands for the current number of iterations conducted. Figure 10 exhibits a representative illustration of GOOSE behavior [16].

Figure 7
Histogram optimization.

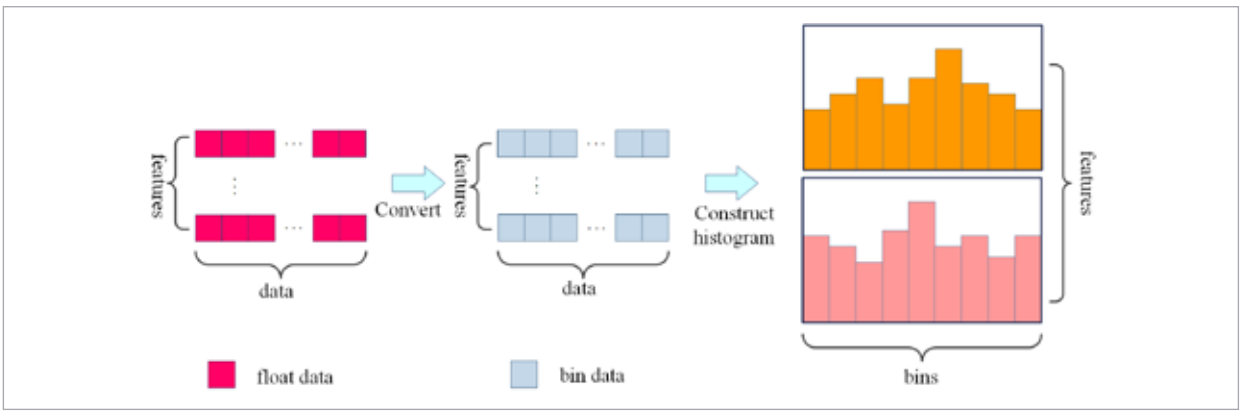


Figure 8
Level-wise growth strategy adopted in XGBoost.

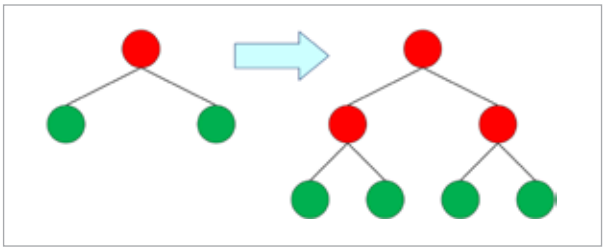


Figure 9
Leaf-wise growth strategy adopted in LightGBM.

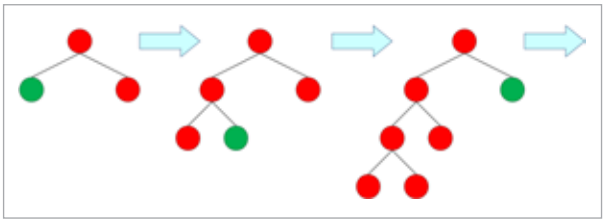
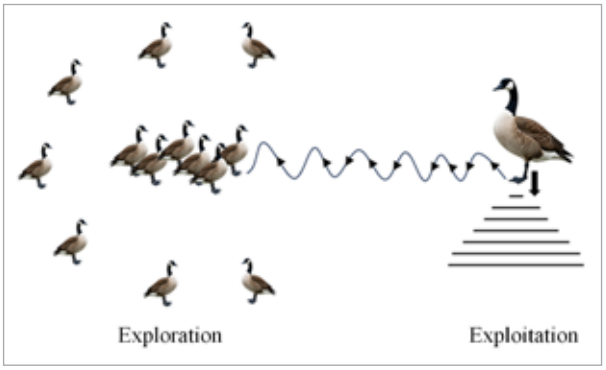


Figure 10
Graphical representation of GOOSE behavior.



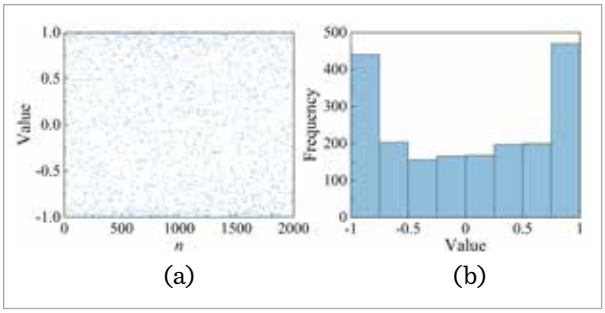
4.3. Enhanced GOOSE (EGOOSE) Algorithm

To improve population diversity and alleviate the risk of getting stuck in local optima, we propose a novel method termed Enhanced GOOSE (EGOOSE) algorithm, which adopts three strategies including chaotic mapping, Gaussian mutation, and crisscross optimization.

4.3.1. Chaotic Mapping

Due to the ergodic property of chaotic mapping, metaheuristic swarm intelligence optimization can cover different regions of the search space during the initialization stage, which is conducive to discovering a better solution space range and laying a solid foundation for global search. When the algorithm starts to search from a population initialized by chaos, it can more effectively avoid the premature convergence problem caused by the poor distribution of the initial population. In this paper, we use Chebyshev chaotic mapping, the definition of which is given as follows:

Figure 11
(a) Samplings of Chebyshev chaotic mapping, (b) Histogram of samplings of Chebyshev chaotic mapping.



$$x_{n+1} = \cos\left(k * \cos^{-1}\left(x_n\right)\right), \tag{19}$$

where k is an integer, and $|k|>1$. In the experiments of this work, the parameter k is set to 4. The range of the state value x_n of the chaotic orbit is $x_n \in [-1,1]$. Figure 11(a) shows the 2000 samplings of Chebyshev chaotic mapping, and Figure 11(b) shows the histogram of samplings of Chebyshev chaotic mapping.

4.3.2. Gaussian Mutation

Population diversity is of vital importance for preventing the algorithm from getting trapped in local optimal solutions. By increasing the diversity of populations, Gaussian mutation enables the algorithm to explore the solution space more comprehensively, thus enhancing the algorithm's robustness and adaptability. Particularly when the algorithm is approaching the optimal solution, Gaussian mutation can prompt individuals to converge rapidly. The Gaussian mutation operation applied to X_{it} is defined as follows:

$$X_{it+1} = X_{it} * Gauss\left(X_{it}, Pos_{best}, \sigma\right), \tag{20}$$

where

$$Gauss(X_{it}, Pos_{best}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(X_{it} - Pos_{best})^2}{2\sigma^2}\right) \tag{21}$$

During the iterative process of the algorithm, σ is dynamically adjusted according to the following formula:

$$\sigma = \frac{2}{1+10*it/Max\_It}. \quad (22)$$

During the initial stage of iteration procedure, a relatively large variance is employed to enhance the global search ability. As the iteration procedure moves into its later phase, the variance is progressively reduced to strengthen the local search capability.

4.3.3. Crisscross Optimization

The crisscross optimization [26] enhances inter-individual differences and population diversity, thereby facilitating the algorithm's search in complex solution spaces. This mechanism mitigates the risk of entrapment in local optima, accelerating convergence toward global optima while maintaining a balanced trade-off between exploratory search and exploitative refinement. The crisscross strategy consists of horizontal crossover and vertical crossover.

In horizontal crossover stage, individuals in the population are first randomly paired, followed by performing horizontal crossover on each pair. For parent individuals G_{i1} and G_{i2} in a pair, their offspring G_{i1}^{hc} and G_{i2}^{hc} are generated as follows:

$$G_{i1,j}^{hc} = \alpha_1 \cdot G_{i1,j} + (1-\alpha_1)\cdot G_{i2,j} + \beta_1 \cdot (G_{i1,j} - G_{i2,j}) \quad (23)$$

$$G_{i2,j}^{hc} = \alpha_2 \cdot G_{i2,j} + (1-\alpha_2)\cdot G_{i1,j} + \beta_2 \cdot (G_{i2,j} - G_{i1,j}) \quad (24)$$

where $G_{i1,j}$ and $G_{i2,j}$ represent the j-th dimensions of $G_{i1}$and G_{i2}, respectively. $G_{i1,j}^{hc}$ and $G_{i2,j}^{hc}$ are the j-th dimensions of the offspring generated by horizontal crossover. $\alpha_1 \in (0,1)$, $\alpha_2 \in (0,1)$, $\beta_1 \in (-1,1)$, $\beta_2 \in (-1,1)$ are random numbers. The offspring individual is compared with its parent counterpart, and the one exhibiting a smaller fitness function value is reserved.

In vertical crossover stage, each individual updates only one specific dimension while preserving the others unchanged. Suppose that $j1$ and $j2$ are randomly selected, then the $j1$-th dimension of the offspring G_i^{vc} is updated according to the following formula:

$$G_{i,j1}^{vc} = \alpha \cdot G_{i,j1} + (1-\alpha)\cdot G_{i,j2}, \quad (25)$$

where $\alpha \in (0,1)$ denotes a random number. Similarly, off spring individuals with a fitness function value smaller than its parent counterpart is retained.

4.3.4. Application of Three Strategies for Enhancement of GOOSE Algorithm

The EGOOSE algorithm is developed to enhance the performance of GOOSE algorithm through three strategies. First, initial population is generated with Chebyshev chaotic map ping in Equation (19). Second, in the exploration phase, if $0.45 \leq rnd < 0.5$, the formulation for calculating the individual optimal position with Gaussian mutation is given by Equation (20). On the other hand, if $rnd<0.45$, the formula for calculating the individual optimal position still follows Equation (17). Third, horizontal and vertical crossover operations, defined in Equations (23)-(25), are applied to update X_{it}. The whole flow of EGOOSE algorithm is shown in Figure 12.

4.3.5. Outline of the Proposed Work

Figure 13 demonstrates the outline of the implementation of the proposed EGOOSE-LightGBM model. There are five phases in the procedure of the EGOOSE-LightGBM model-based cougher recognition.

Phase 1: The cough sound signal recording is conducted to establish the original dataset, the size of which is subsequently enlarged by using cough data augmentation techniques.

Phase 2: Cougher-specific acoustic characteristics are extracted to construct the hybrid feature vector *sd_vector* by using SDFF technique.

Phase 3: The EGOOSE-LightGBM model for cougher recognition is built.

Phase 4: The training dataset is employed to perform iterative optimization of EGOOSE-LightGBM model parameters.

Phase 5: The model evaluation is dedicated to assess the model's performance for cougher recognition.

It is worth noting that LightGBM serves as the foundational model for cougher recognition, with its hyperparameters optimized through the EGOOSE algorithm.

5. Results and Discussion

5.1. Evaluation Metrics

To assess the performance of multi-class cougher recognition model, five evaluation metrics, i.e., accuracy, precision, F1 Score, ROC-AUC, and MCC [9], [15], [35], are used.

$$\text{Accuracy} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i + TN_i}{TP_i + FP_i + TN_i + FN_i} \tag{26}$$

$$\text{Precision} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FP_i} \tag{27}$$

$$\text{F1-Score} = \frac{1}{N}\sum_{i=1}^{N}\frac{2\times \text{Precision}_i \times \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i}, \tag{28}$$

where TP_i, FP_i, TN_i, and FN_i denote True Positive, False Positive, True Negative, and False Negative, respectively. N denotes the number of coughers.

$$\text{MCC} = \frac{a\times b - \sum_{k=1}^{N} r_k \times q_k}{\sqrt{(b^2 - \sum_{k=1}^{N} r_k^2)(b^2 - \sum_{k=1}^{N} q_k^2)}}, \tag{29}$$

where $a = \sum_{k=1}^{N} c_{kk}$, $b = \sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij}$, $r_k = \sum_{i=1}^{N} c_{ki}$, and $q_k = \sum_{i=1}^{N} c_{ik}\cdot c_{ij}$ denotes the element in the i-th row and j-th column of confusion matrix. ROC-AUC denotes the area under ROC curve.

5.2. Comparison between GOOSE and EGOOSE

The experiments were performed on a computer equipped with 16 GB of memory and an Intel(R) Core(TM) i7-9700 CPU. This work exploited Matlab 2024b as the software for experiments.

Figure 12
The flowchart of the proposed EGOOSE algorithm.

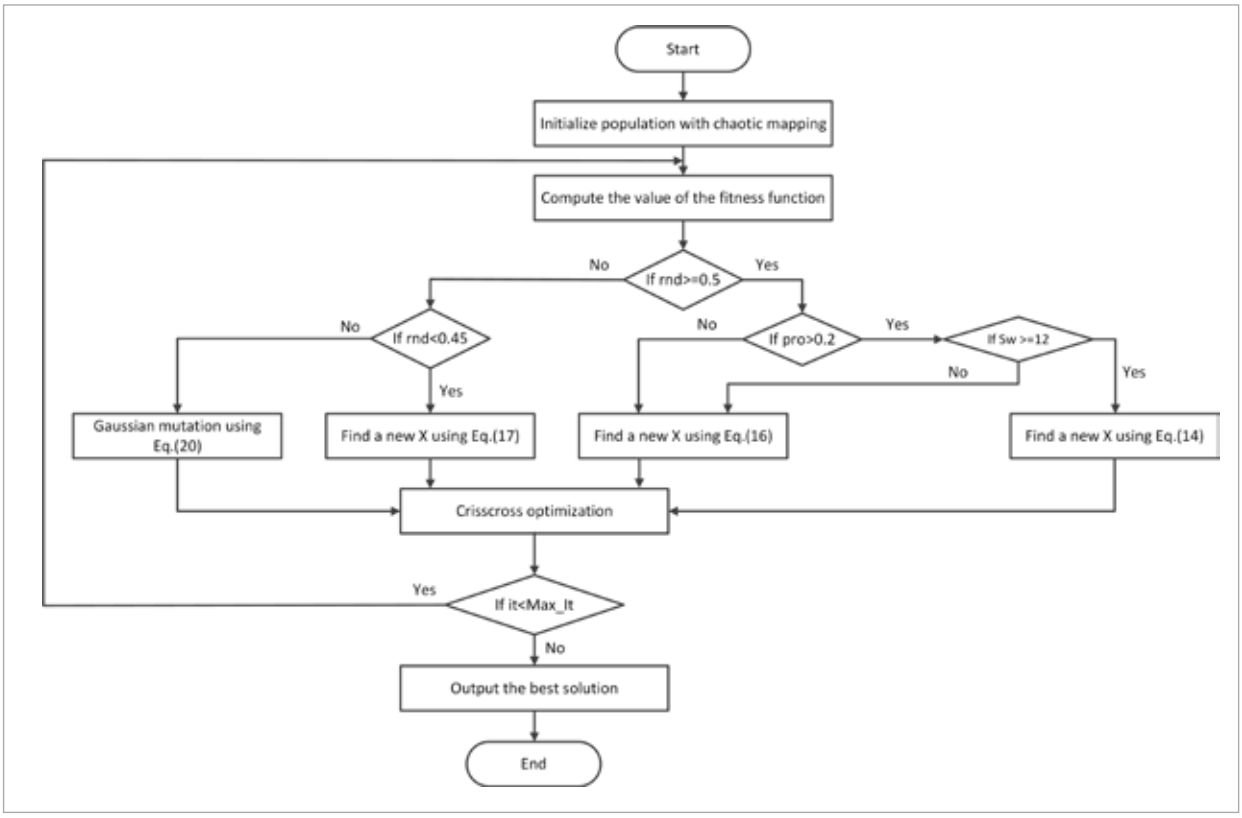


Figure 13
The overall analysis process of EGOOSE-LightGBM model-based cougher recognition method.

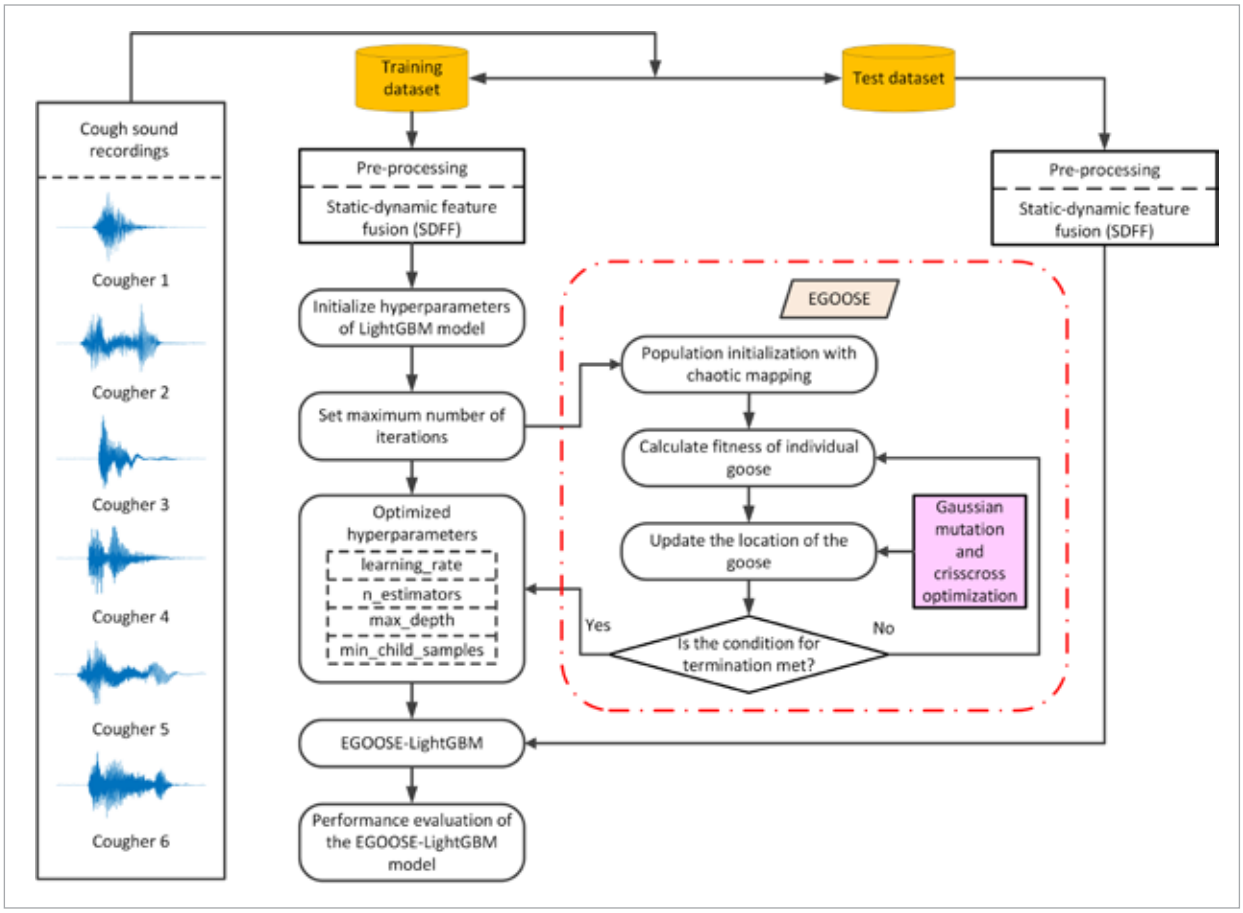


Table 2
Benchmark test functions for the comparison of GOOSE and EGOOSE.

Test Function	n	*range*	$f_{\min}$
$F_1(x)=\sum_{i=1}^{n}x_i^2$	30	$[-100,100]^n$	0
$F_2(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	30	$[-10,10]^n$	0
$F_7(x)=\sum_{i=1}^{n}ix_i^4+random[0,1)$	30	$[-1.28,1.28]^n$	0
$F_8(x)=\sum_{i=1}^{n}-x_i\sin\left(\sqrt{\lvert x_i\rvert}\right)$	30	$[-500,500]^n$	-12569.5
$F_9(x)=\sum_{i=1}^{n}[x_i^2-10\cos(2\pi x_i)+10]$	30	$[-5.12,5.12]^n$	0
$F_{10}(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\right)-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos 2\pi x_i\right)+20+e$	30	$[-32,32]^n$	0

Test Function	n	*range*	$f_{\min}$
$F_{15}(x)=\sum_{i=1}^{11}\left[a_i-\frac{x_1(b_i^2+b_ix_2)}{b_i^2+b_ix_3+x_4}\right]^2$	4	$[-5,5]^n$	0.0003075
$F_{18}(x)=\left[1+(x_1+x_2+1)^2\left(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2\right)\right]\times$ $\left[30+(2x_1-3x_2)^2\left(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2\right)\right]$	2	$[-5,5]^n$	3

To comprehensively compare the performance of GOOSE and EGOOSE, we selected 8 classical test functions from the CEC2005 [12], [45] benchmark function suite for algorithm evaluation. In this study, three unimodal functions, i.e., $F_1(x)$, $F_2(x)$, and $F_7(x)$, focused on assessing the convergence speed and local search ability of algorithms. Meanwhile, three multimodal functions, i.e., $F_8(x)$, $F_9(x)$, and, $F_{10}(x)$ focused on evaluating global search ability of algorithms. Additionally, two hybrid composition functions, i.e., $F_{15}(x)$ and $F_{18}(x)$, focused on gauging algorithm's balancing capacity of local and global search within complex search spaces. The definitions and properties of the 8 functions are described in Table 2. In $F_{15}(x)$, the coefficient $[a_1,a_2,...,a_{11}]$=[0.1957,0.1947,0.1735, 0.16,0.0844,0.0627,0.0456,0.0342,0.0323,0.0235,0.0246], $b_i=4\times2^{1-i}$, $i=1,...,4$, and $b_i=1/(2i-6)$, $i=5,...,11$.

In the experiment conducted for performance evaluation of GOOSE and EGOOSE, the swarm size was chosen as $N=100$, and the maximum number of iterations was set to $Max\_It=1000$. Figure 14 illustrates the results of performance comparison between GOOSE and EGOOSE. It should be noted that the vertical axis in Figure 14(e) employs a logarithmic scale, and the fitness value of EGOOSE drops precisely to zero when the number of iterations exceeds 150. The convergence curves intuitively present the trend of performance indicators and convergence status during algorithm iteration. It can be seen that, for unimodal, multimodal, and hybrid composition test functions, our proposed EGOOSE algorithm exhibits superior performance in terms of convergence speed and accuracy compared to GOOSE algorithm. The improvement in performance highlights the effectiveness of strategies, i.e., chaotic mapping, Gaussian mutation, and crisscross optimization, adopted in EGOOSE. The EGOOSE convergence curves in Figure 14(b), Figure 14(c), Figure 14(e), and Figure 14(f) show step-type decline, demonstrating its adaptive capability to break away from local optima and ultimately converge toward global optimal solution.

5.3. Performance Evaluation for Different Values of Swarm Size

To assess the impact of swarm size on the EGOOSE-LightGBM recognition model, we applied the model to scenarios with swarm sizes of 10, 20, 50, 100, 150, and 200, respectively, for comparison. The comprehensive ranking of the performance metrics of the EGOOSE-LightGBM recognition model with different swarm sizes is shown in Table 3. Since there are six cases of swarm size in total, the top performance level for each evaluation metric is assigned a rank of 6. The lower the performance level, the lower the corresponding rank it receives. By summarizing the rank values of all evaluation metrics for each swarm size, we can obtain its total rank value. The graphical representation of the ranking of performance metrics for the model optimized with different swarm sizes is shown in Figure 15. It can be observed that the highest comprehensive performance ranking is achieved when the swarm size is 100, accompanied by the following performance metrics: accuracy of 99.500%, precision of 99.508%, F1-Score of 99.498%, ROC-AUC of 0.99700, and MCC of 0.99402. As the swarm size increases, it not only leads to additional computational overhead but also elevates the likelihood of search agents clustering in a single region, which may potentially weaken the model's local search capability.

The hyperparameters of LightGBM model in EGOOSE-LightGBM to be optimized include learning rate, nestimators, maxdepth, and minchildsamples. When the swarm size is set to 100, the search ranges and optimal values of hyperparameters are demonstrated in Table 4.

Figure 14

3D maps and convergence curves for 8 classical benchmark test functions: (a) $F_1(x)$, (b) $F_2(x)$, (c) $F_7(x)$, (d) $F_8(x)$, (e) $F_9(x)$, (f) $F_{10}(x)$, (g) $F_{15}(x)$, and (h) $F_{18}(x)$.

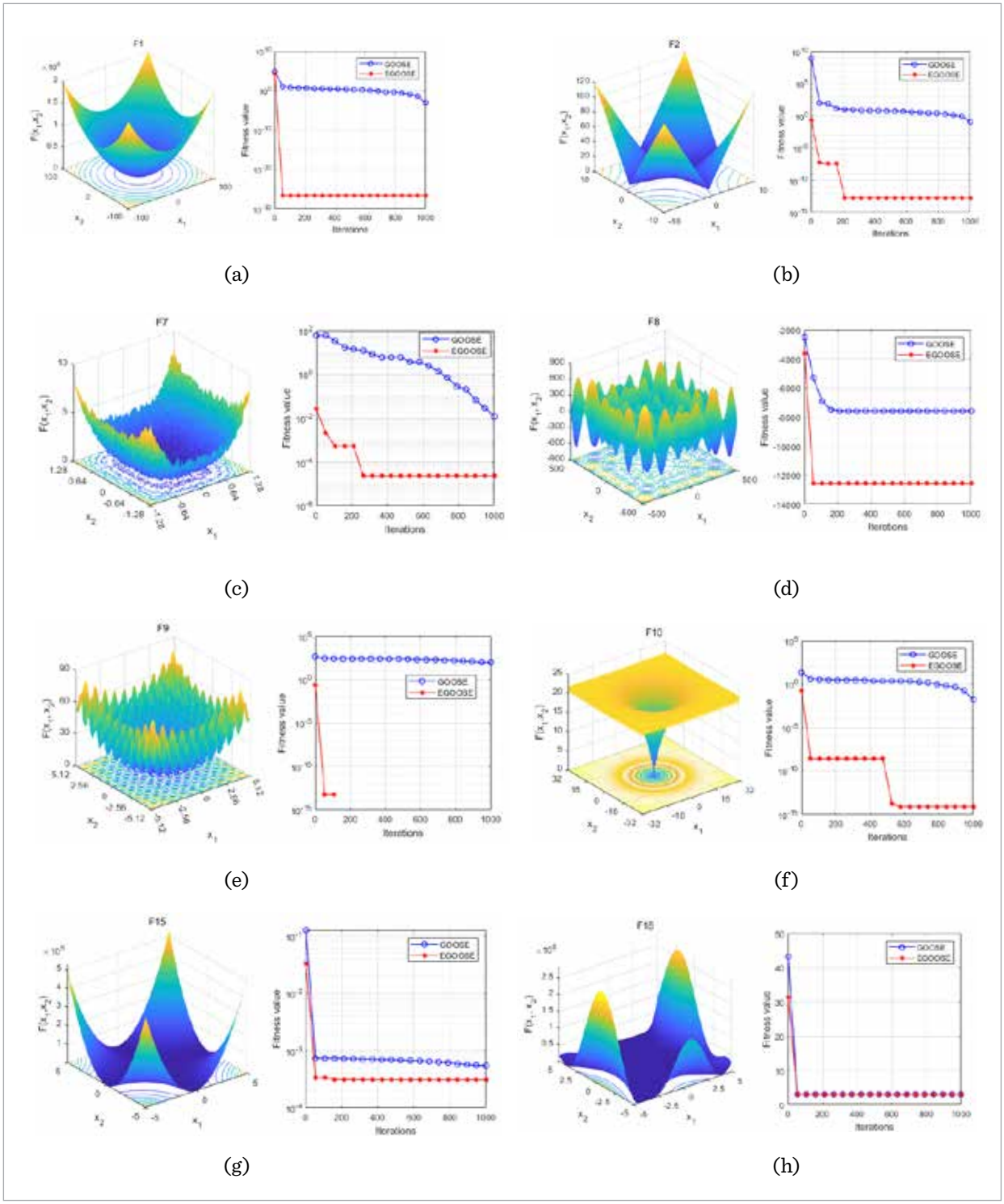


Table 3
Evaluation metrics of the EGOOSE-LightGBM cougher recognition model for different swarm sizes.

Swarm Size	Performance and Rank									
	Accuracy (%)	Rank	Precision (%)	Rank	F1Score (%)	Rank	ROC-AUC	Rank	MCC	Rank
10	98.833	2	98.838	1	98.834	2	0.99300	2	0.98601	1
20	99.333	5	99.336	5	99.332	5	0.99600	5	0.99201	5
50	99.333	5	99.333	4	99.332	5	0.99600	5	0.99200	4
100	99.500	6	99.508	6	99.498	6	0.99700	6	0.99402	6
150	99.167	4	99.168	3	99.165	4	0.99500	4	0.99001	3
200	99.000	3	98.998	2	98.998	3	0.99400	3	0.98800	2

5.4. Comparison with Existing Models

In this experiment, six models including SVM, XGBoost, RF-Adaboost, LSTM, LightGBM, and EGOOSE-LightGBM were evaluated for comparison

Figure 15
Comprehensive ranking of model performance evaluation metrics with different swarm sizes.

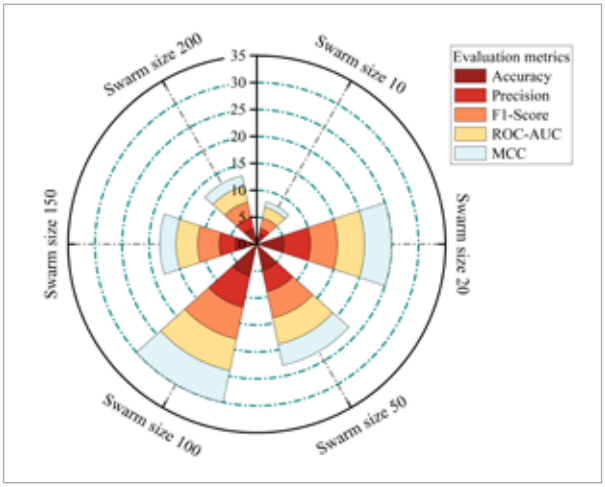


Table 4
Search ranges and optimal values of hy-perparameters of LightGBM in EGOOSE-LightGBM.

Hyperparameter	Search Range	Optimal Value
learning_rate	[0.01,1]	0.23107
n_estimators	[5,300]	188
max_depth	[3, 20]	5
min_child samples	[8, 40]	10

To train the recognition models and test the trained models, the dataset was split into training and test subsets. The test subset was composed of 600 segments of cough sound data, and the rest of cough sound segments were used for training. Five performance evaluation metrics, namely, accuracy, precision, F1-Score, ROC-AUC, and MCC, were employed to assess the performance of models. Notably, the six recognition models shared the identical training and test datasets to ensure fairness in comparison.

The results are shown in Table 5, from which it can be seen that the accuracies for SVM, XGBoost, RF-Adaboost, LSTM, LightGBM, and EGOOSE-LightGBM are 97.833%, 91.667%,98.167%, 98.000%, 93.667%, and 99.500%, respectively. This demonstrates that the EGOOSE-LightGBM is the most accurate model among them. Likewise, the values of precision, F1-Score, ROC-AUC, and MCC of the EGOOSE-LightGBM model are higher than that of SVM, XGBoost, RF-Adaboost, LSTM, and LightGBM. This indicates that our proposed EGOOSE-LightGBM model yields the best performance for cougher recognition among all models in Table 5.

5.5. Comparison of Six Models in Terms of Their Confusion Matrices

The confusion matrix serves as a visual aid that vividly presents a model's outcomes for a recognition task. Figure 16 displays the confusion matrices of the six recognition models. As depicted in Figure 16, the rows of the confusion matrix denote the true labels, while the columns represent the predicted labels. The value within each cell indicates the proportion of correct or incorrect predictions. When prediction

Figure 16
Confusion matrices of different models.

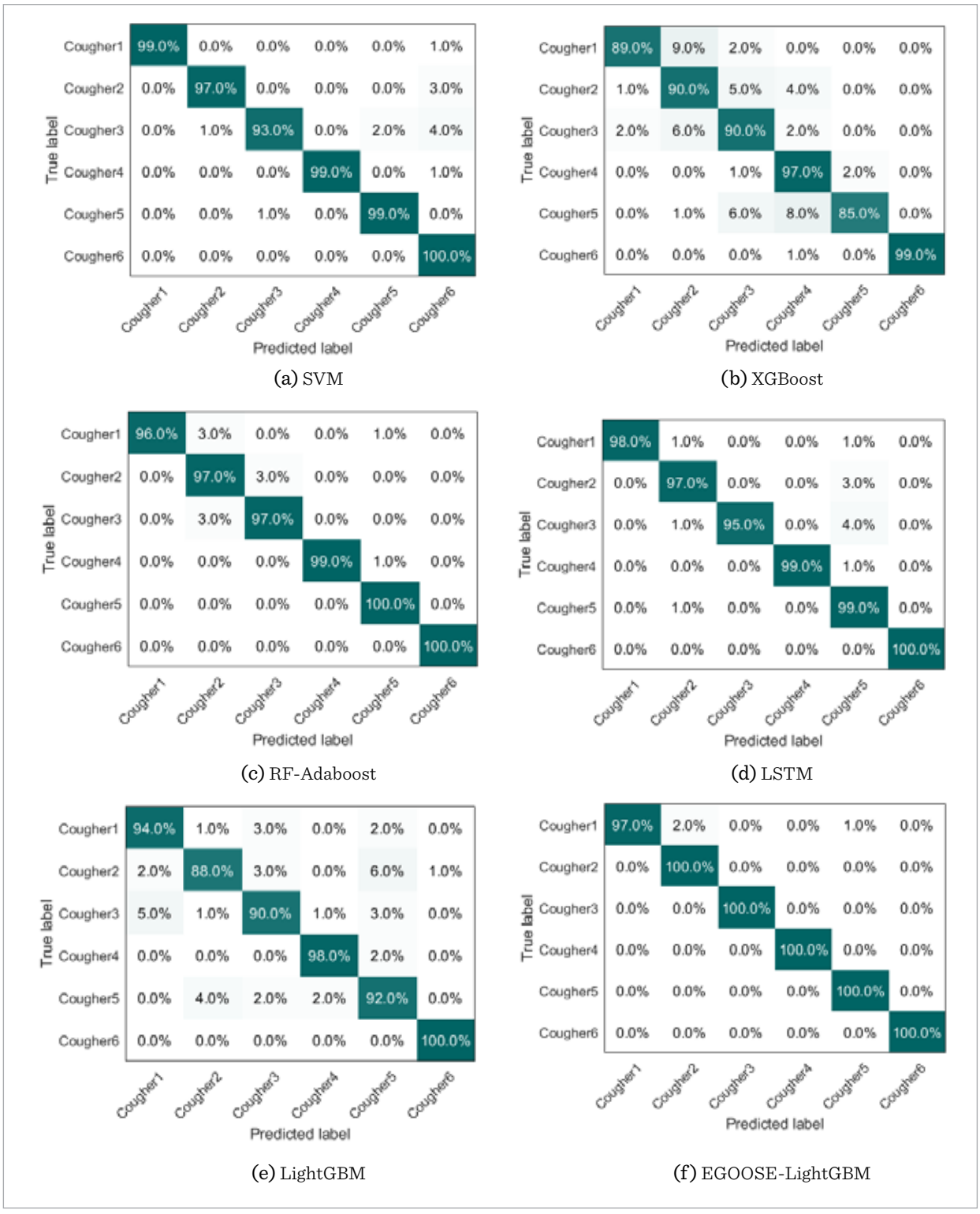


Table 5
Performance comparison with other machine learning models.

Model	Performance				
	Accuracy (%)	Precision (%)	F1-Score (%)	ROC-AUC	MCC
SVM	97.833	97.946	97.842	0.00987	0.97420
XGBoost	91.667	92.082	91.706	0.95000	0.90065
RF-Adaboost	98.167	98.202	98.172	0.98900	0.97805
LSTM	98.000	98.111	98.091	0.98800	0.97615
LightGBM	93.667	93.697	93.656	0.96200	0.92410
EGOOSE-LightGBM	99.500	99.508	99.498	0.99700	0.99402

accuracy of a category is higher, it appears as a darker shade in the visualization. Especially, all correct predictions are aligned along the diagonal, whereas incorrect predictions are placed off the diagonal.

It has been observed from Figure 16 that out of 600 segments of cough sounds, 13, 50, 11, 12, 38, and 3 segments were misclassified for SVM, XGBoost, RF-Adaboost, LSTM, LightGBM, and EGOOSE-LightGBM, respectively. In addition, our proposed model has the highest recognition accuracies for cougher 2, cougher 3, cougher 4, cougher 5, and cougher 6 among these six models. Accordingly, the experimental results validate the proposed EGOOSE-LightGBM model's capability to classify six distinct categories with the best overall recognition performance on the test dataset.

Furthermore, based on the confusion matrices of the six different recognition models, it can be found that cougher 6 has the highest average recognition rate, from which it can be inferred that the sound-related features of cougher 6 are significantly different from those of other coughers.

5.6. Performance Evaluation for Noise Robustness

To evaluate the noise robustness performance of SDFF against static features, we carried out comparison experiment for the EGOOSE-LightGBM model under different signal-to-noise ratios (SNRs). The definition of SNR is given as follows:

$$\mathrm{SNR} = 10\log_{10}\frac{E\left\{\| s(t)\|^2\right\}}{E\left\{\| n(t)\|^2\right\}}. \tag{30}$$

In Figure 17, we provide the results for the cougher recog-nition accuracies of static features and SDFF with SNR values of 0, 10, 20, 30, and 40 dB, respectively. As demonstrated in Figure 17, the EGOOSE-LightGBM model with SDFF, which fuses static and dynamic features of cough voice, achieves better noise robustness compared to the EGOOSE-LightGBM model that uses only static features.

We also conducted experiment to investigate the impact of various noise levels on the recognition performance of LightGBM and EGOOSE-LightGBM. Table 6 shows the recognition accuracy versus SNR. It can be seen from the table that, under SNRs of 0, 10, 20, 30, and 40 dB, EGOOSE-LightGBM outperforms LightGBM in terms of recognition accuracy

Figure 17
Cougher recognition accuracy comparison of static features and SDFF for the proposed EGOOSE-LightGBM model.

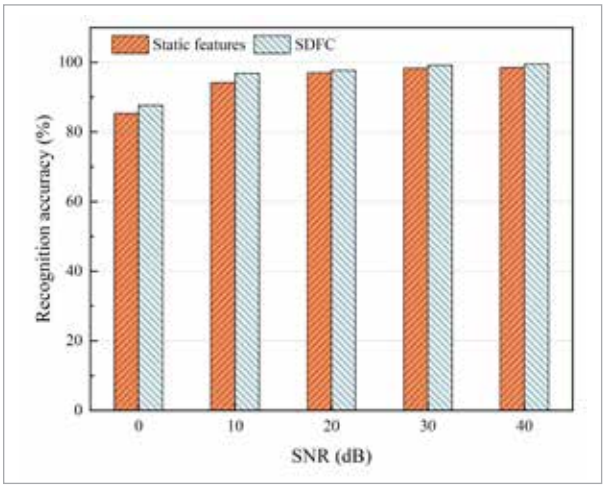


Table 6
Comparison of LightGBM and EGOOSE-LightGBM under different SNRs.

Model	Accuracy (%) under Different SNRs				
	0 dB	10 dB	20 dB	30 dB	40 dB
LightGBM	77.333	87.833	88.333	91.667	93.333
EGOOSE-LightGBM	87.667	96.833	97.667	98.500	99.167

by 10.334%, 9.000%, 9.334%, 6.833%, and 5.834%, respectively. As demonstrated by the analysis results above, EGOOSE-LightGBM exhibits significantly stronger noise robustness compared to LightGBM, particularly under low SNR conditions.

5.7. Comparison of Recognition Effects with Data Augmentation

To enhance the model's generalization ability and robustness, three types of audio augmentation methods, i.e., time stretching, addition of white noise, pitch shifting, were applied to construct dataset for training the model. In this experiment.

three audio augmentation methods were gradually added to the original dataset in sequence. For the sake of brevity, we use A, B, C, and D to represent original dataset, dataset with time stretching, dataset with addition of white noise, dataset with pitch shifting, respectively. Hence, four datasets, i.e., A, A+B, A+B+C, A+B+C+D, were built for validating the effectiveness of data augmentation in improving algorithm performance. Table 7 shows that, as datasets B, C, and D are progressively incorporated into the original dataset A to form more comprehensive datasets, all performance evaluation metrics of EGOOSE-LightGBM exhibit consistent improvement. This verifies that data augmentation enriches the diversity of data, thus boosting the model's generalization capability.

6. Conclusion

As a special speech event, cough signal conveys a wealth of valuable acoustic characteristic information, which can be dedicated to identity verification, disease diagnosis, forensic investigation, patient monitoring, security supervision, etc. Unlike traditional speaker recognition, studies focusing on cougher recognition through acoustic cough recordings are still scarce. This paper presented the EGOOSE-LightGBM model for cougher recognition using the characteristics of cough sound signals.

This study employed SDFF technique, in which the static acoustic features (MFCC, formant, and fundamental frequency) are fused with the dynamic acoustic features (ΔMFCC and jitter of formant) to combine the merits of the two types of acoustic characteristics of cough recordings. It was found that SDFF yields better recognition performance than individual static features.

Based on GOOSE algorithm, we proposed a new metaheuristic algorithm called EGOOSE. Compared to GOOSE algorithm, EGOOSE boosts population diversity, improves the ability to escape local optimum traps, and elevates global search efficiency due to the utilization of three enhancement strategies, i.e., chaotic mapping, Gaussian mutation, and crisscross optimization. In the proposed EGOOSE-LightGBM model for cougher recognition, the hyperparameters of Light-GBM are optimized with EGOOSE algorithm. Experimental results show that, EGOOSE-LightGBM obtains higher accuracy rates than LightGBM at the SNR levels of 0, 10, 20, 30, and 40 dB, which validates that our proposed model achieves stronger robustness against noise and maintains better performance in cougher recognition.

We established a cough dataset, which consists of coughing voices of 3 male and 3 female coughers. The original 1200 segments of cough sound data are extended to 4800 enhanced segments by using the audio data augmentation techniques including time stretching, addition of white noise, and pitch shifting, which contribute to the performance improvement of the EGOOSE-LightGBM model. Specifically, it achieves a recognition accuracy of 96.333% on the original dataset and 99.500% on the augmented dataset, which implies 3.167% increase in performance.

Table 7
Performance comparison for data augmentation.

Dataset	Performance				
	Accuracy (%)	Precision (%)	F1-Score (%)	ROC-AUC	MCC
A	96.333	96.379	96.343	0.97800	0.95605
A+B	97.833	97.911	97.837	0.98700	0.97145
A+B+C	98.167	98.207	98.167	0.98900	0.97808
A+B+C+D	99.500	99.508	99.498	0.997700	0.99402

We have carried out extensive experiments on the cough dataset to validate the superiority of the proposed EGOOSE-LightGBM model. The comparison results show that the devised EGOOSE-LightGBM model outperforms other models such as SVM, XGBoost, RF-Adaboost, LSTM, and LightGBM in terms of accuracy, precision, F1-Score, ROC-AUC, and MCC.

In the future, we will deepen the research on cough sound-based identity recognition from three aspects. First, we will study the utilization of richer and more effective cough sound. features. Second, we will explore the implementation of efficient deep neural networks with promising performance for identity recognition. Third, we will employ multimodal tech neologies to further enhance the accuracy and noise robustness of the algorithm.

Acknowledgement

This work was supported in part by the Key Research and Development Program of Hunan Province, China, under Grants 2025AQ2022; in part by the Social Science Fund Project of Hunan Province, China, under Grant 22YBA283; in part by the Changsha Natural Science Foundation, China, under Grant kq2208064; and in part by the Education Science Planning Project of Hunan Province, under Grant XJK24BGD009.

References

1. Bai, Z., Zhang, X.-L. Speaker Recognition Based on Deep Learning: An Overview. Neural Networks, 2021, 140, 65-99. https://doi.org/10.1016/j.neunet.2021.03.004
2. Bakir, R., Orak C., Yuksel, A. Optimizing Hydrogen Evolution Prediction: A Unified Approach Using Random Forests, LightGBM, and Bagging Regressor Ensemble Model. International Journal of Hydrogen Energy, 2024, 67, 101-110. https://doi.org/10.1016/j.ijhydene.2024.04.173
3. Bao, Z., Xu, B., Zhang, X., Yin, Y., Yang, X., Niu, Q. Potential Pneumoconiosis Patients Monitoring and Warning System with Acoustic Signal. Sensors, 2025, 25, 1874. https://doi.org/10.3390/s25061874
4. Cai, J., Wang R., Zhao, D., Yuan, Z., McKenna, V., Friedman, A., Foot, R., Storey, S., Boente, R., Vhaduri, S., Min, B.-C. Multimodal Audio-Based Disease Prediction With Transformer-Based Hierarchical Fusion Network. IEEE Transactions on Audio, Speech and Language Processing. 2025, 33, 1170-1182. https://doi.org/10.1109/TASLPRO.2025.3543975
5. Campbell, J. P., Shen, W., Campbell, W. M., Schwartz, R., Bonastre, J. F., Matrouf, D. Forensic Speaker Recognition A Need for Caution. IEEE Signal Processing Magazine, 2009, 26(2), 95-103. https://doi.org/10.1109/MSP.2008.931100
6. Chaiwongyen, A., Duangpummet, S., Karnjana, J., Kongprawechnon, W., Unoki, M. Potential of Speech-Pathological Features for Deepfake Speech Detection. IEEE Access, 2024, 12, 121958-121970. https://doi.org/10.1109/ACCESS.2024.3447582
7. Chen, Y., Attri, P., Barahona, J., Hernandez, M. L., Carpenter D., Bozkurt, A., Lobaton, E. Robust

Cough Detection with Out-of-Distribution Detection. IEEE Journal of Biomedical and Health Informatics, 2023, 27(7), 3210-3221. https://doi.org/10.1109/JBHI.2023.3264783

8. Chen, T., Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 2016 ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 13-17 August 2016, 785-794. https://doi.org/10.1145/2939672.2939785
9. Chicco, D., Jurman, G. The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation. BMC Genomics, 2020, 21(1), 1-13. https://doi.org/10.1186/s12864-019-6413-7
10. Cho, S., Wee, K. Multi-Noise Representation Learning for Robust Speaker Recognition. IEEE Signal Processing Letters, 2025, 32, 681-685. https://doi.org/10.1109/LSP.2025.3530879
11. Deng, M., Meng, T., Cao, J., Wang, S., Zhang, J., Fan, H. Heart Sound Classification Based on Improved MFCC Features and Convolutional Recurrent Neural Networks. Neural Networks, 2020, 130, 22-32. https://doi.org/10.1016/j.neunet.2020.06.015
12. Digalakis, J. G., Margaritis, K. G. On Benchmarking Functions for Genetic Algorithms. International Journal of Computer Mathematics, 2001, 77, 481-506. https://doi.org/10.1080/00207160108805080
13. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. Annals of Statistics, 2001, 29, 1189-1232. https://doi.org/10.1214/aos/1013203451
14. Gambhir, P., Dev, A., Bansal, P., Sharma, D. K., Gupta, D. Residual Networks for Text-Independent Speaker Identification: Unleashing the Power of Residual Learning. Journal of Information Security and Applications, 2024, 80, 103665. https://doi.org/10.1016/j.jisa.2023.103665
15. Grandini, M., Bagli, E., Visani, G. Metrics for Multi-Class Classification: An Overview. arXiv 2020, arXiv:2008.05756. https://doi.org/10.48550/arXiv.2008.05756
16. Hamad, R. K., Rashid, T. A. GOOSE Algorithm: A Powerful Optimization Tool for Real-World Engineering Challenges and Beyond. Evolving Systems, 2024, 15(4), 1249-1274. https://doi.org/10.22541/au.169333786.66092666/v1
17. Hanifa, R. M., Isa, K., Mohamad, S. A Review on Speaker Recognition: Technology and Challenges. Computers & Electrical Engineering, 2021, 90, 107005. https://doi.org/10.1016/j.compeleceng.2021.107005
18. Hoang, N. D., Tran-Anh D., Luong M., Luong C., Pham C. Federated Few-Shot Learning for Cough Classification with Edge Devices. Applied Intelligence, 2023, 53(23), 28241-28253. https://doi.org/10.1007/s10489-023-05006-4
19. Huang, W.-C., Wu Y.-C., Toda, T. Multi-Speaker Text-to-Speech Training with Speaker Anonymized Data. IEEE Signal Processing Letters, 2024, 31, 2995-2999. https://doi.org/10.1109/LSP.2024.3482701
20. Issahaku, F. L. Y., Liu, X., Lu, K., Fang, X., Danwana, S. B., Asimeng, E. Multimodal Deep Learning Model for Covid-19 Detection. Biomedical Signal Processing and Control, 2024, 91, 105906. https://doi.org/10.1016/j.bspc.2023.105906
21. Jahangir, R., Teh, Y. W., Nweke, H. F., Mujtaba, G., Al-Garadi, M. A., Ali, I. Speaker identification through artificial intelligence techniques: A Comprehensive Review and Research Challenges. Expert Systems with Applications, 2021, 171, 114591. https://doi.org/10.1016/j.eswa.2021.114591
22. Jokić, S., Cleres, D., Rassouli, F., Steurer-Stey, C., Puhan, M.A., Brutsche, M., Fleisch, E., Barata, F. Triplet Cough: Cougher Identification and Verification from Contact-Free Smartphone-Based Audio Recordings Using Metric Learning. IEEE Journal of Biomedical and Health Informatics, 2022, 26(2), 2746-2757. https://doi.org/10.1109/JBHI.2022.3152944
23. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the 2017 Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4-9 December 2017, 3147-3155. https://proceedings.neurips.cc/paper/2017/ hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

24. Kumar, S., Shvetsov A. V., Alsamhi, S. H. Fuzzy Guard: A Novel Multimodal Neuro-Fuzzy Framework for COPD Early Diagnosis. IEEE Internet of Things Journal, 2025, 12(8), 9627-9637. https://doi.org/10.1109/JIOT.2024.3467176

25. Laska, B., Xi, P., Valdés, J. J., Wallace, B., Green, J., Goubran, R., Knoefel, F. Coughprint: Distilled Cough Representations from Speech Foundation Model Embeddings. IEEE Transactions on Instrumentation and Measurement. 2025, 74, 1-10. https://doi.org/10.1109/TIM.2025.3568985

26. Liang, X., Zhang, Y., Long, W. Spider Monkey Optimization Algorithm with Crisscross Optimization. Mathematics in Practice and Theory, 2022, 52(12), 144-158.

27. Meng, Q., Ke, G., Wang, T., Chen, W., Ye, Q., Ma, Z.-M., Liu, T. Y. A Communication-Efficient Parallel Algorithm for Decision Tree. In Proceedings of the 2016 Conference on Neural Information Processing Systems (NIPS), Barcelona, Spain, 5-10 December 2016; pp. 1279-1287.

28. Nellore, B. T., Sreeram, G., Dhawan, K., Reddy, P. B. Evaluating Speech Production-based Acoustic Features for COVID-19 Classification using Cough Signals. In Proceedings of the 18th IEEE India Council International Conference (INDICON), Guwahati, India,19-21 December 2021, 1-5. https://doi.org/10.1109/INDICON52576.2021.9691699

29. Ni, C., Huang, H., Cui, P., Ke, Q., Tan, S., Ooi, K. T., Liu, Z. Light Gradient Boosting Machine (LightGBM) to Forecasting Data and Assisting the Defrosting Strategy Design of Refrigerators. International Journal of Refrigeration, 2024, 160, 182-196. https://doi.org/10.1016/j.ijrefrig.2024.01.025

30. Pah, N. D., Indrawati, V., Kumar, D. K. Voice Features of Sustained Phoneme as COVID-19 Biomarker. IEEE Journal of Translational Engineering in Health and Medicine, 2022, 10, 4901309. https://doi.org/10.1109/JTEHM.2022.3208057

31. Pahar, M., Klopper, M., Reeve, B., Warren, R., Theron, G., Diacon, A., Niesler, T. Wake-Cough: Cough Spotting and Cougher Identification for Personalized Long-Term Cough Monitoring. In Proceedings of the 2022 30th European Signal Processing Conference (EUSIPCO), https://doi.org/10.23919/EUSIPCO55093.2022.9909522 Belgrade, Serbia, 29 August-2 September 2022, 185-189. https://doi.org/10.23919/EUSIPCO55093.2022.9909522

32. Rejaibi, E., Komaty, A., Meriaudeau, F., Agrebi, S., Othmani, A. MFCC Based Recurrent Neural Network for Automatic Clinical Depression Recognition and Assessment from Speech. Biomedical Signal Processing and Control, 2022, 71, 103107. https://doi.org/10.1016/j.bspc.2021.103107

33. Saleem, S., Subhan, F., Naseer, N., Bais, A., Imtiaz, A. Based on Extracting Accent and Language Information from Short Utterances. Forensic Science International: Digital Investigation, 2020, 34, 300982. https://doi.org/10.1016/j.fsidi.2020.300982

34. Stefanowska, A., Zielinski S. K. Speech Emotion Recognition Using a Multi-Time-Scale Approach to Feature Aggregation and an Ensemble of SVM Classifiers. Archives of Acoustics, 2024, 49(2), 153-168. https://doi.org/10.24425/aoa.2024.148784

35. Stoica, P., Babu, P. Pearson-Matthews Correlation Coefficients for Binary and Multinary Classification. Signal Processing, 2024, 222, 109511. https://doi.org/10.1016/j.sigpro.2024.109511

36. Swain, B. K., Khan M. Z., Chowdhary C. L., Alsaeedi A. SRC: Superior Robustness of COVID-19 Detection from Noisy Cough Data Using GFCC. Computer Systems Science and Engineering, 2023, 46, 2337-2349. https://doi.org/10.32604/csse.2023.036192

37. Tirumala, S. S., Shahamiri S. R., Garhwal A. S., Wang R. Speaker Identification Features Extraction Methods: A Systematic Review. Expert Systems with Applications, 2017, 90, 250-271. https://doi.org/10.1016/j.eswa.2017.08.015

38. Topuz, E. K., Kaya Y. SUPER-COUGH: A Super Learner-Based Ensemble Machine Learning Method for Detecting Disease on Cough Acoustic Signals. Biomedical Signal Processing and Control, 2024, 93, 106165. https://doi.org/10.1016/j.bspc.2024.106165

39. Tran, V. T., You, T. H., Tsai W. H. Acoustic Cues for Person Identification Using Cough Sounds. Computer Methods and Programs in Biomedicine Update, 2025, 100195. https://doi.org/10.1016/j.cmpbup.2025.100195

40. Vhaduri, S., Dibbo, S. V., Kim, Y. Environment Knowledge-Driven Generic Models to Detect Coughs from Audio Recordings. IEEE Open Journal of Engineering in Medicine and Biology, 2023, 4, 55-66. https://doi.org/10.1109/OJEMB.2023.3271457

41. Wang, Z., Hansen, J. H. L. Multi-Source Domain Adaptation for Text Independent Forensic Speaker Recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022, 30, 60-75. https://doi.org/10.1109/ICASSP40776.2020.9053268

42. Whitehill, M., Garrison J., Patel S. Whosecough: In-The-Wild Cougher Verification Using Multitask Learning. In Proceedings of the 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4-8 May 2020, pp. 896-900.https://doi.org/10.1109/ICASSP40776.2020.9053268

43. Yang, Z., Li D., Cai, Y., Wang, Z., Yang, H. EX-Vector: Emotional X Vector Transfer Learning for Speaker Recognition with Emotion Domain Adaption. IEEE Transactions on Audio, Speech, and Language Processing, 2025, 33, 1415-1426. https://doi.org/10.1109/TASLPRO.2025.3540666

44. Yang, J., Liao C., Hu X., Zhu W., Zhang X., Liu B. Transformer Fault Diagnosis Based on DGA and TPE-LightGBM. Journal of Electric Power Science and Technology, 2024, 39, 70-77. https://link.cnki.net/doi/10.19781/j.issn.1673-9140.2024.04.008

45. Yao, X., Liu, Y., Lin, G. Evolutionary Programming Made Faster. IEEE Transactions on Evolutionary Computation, 1999, 3(2), 82-102. https://doi.org/10.1109/4235.771163

46. Zhang, M., Chen, Y., Li, L., Wang, D. Speaker Recognition with Cough, Laugh and 'Wei'. In Proceedings of the 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Aloft Kuala Lumpur Sentral, Malaysia,12-15 December 2017; pp. 497-501. https://doi.org/10.1109/APSIPA.2017.8282083

47. Zhang, T., Huang, Y., Liao, H., Liang, Y. A Hybrid Electric Vehicle Load Classification and Forecasting Approach Based on GBDT Algorithm and Temporal Convolutional Network. Applied Energy, 2023, 351, 121768. https://doi.org/10.1016/j.apenergy.2023.121768