A global extended extreme learning machine combined with electronic nose for identifying tea gas information

Chao Wang¹,², Jizheng Yang¹,² and Qing Wu¹,²

Abstract
To enhance the detection performance of electronic nose (e-nose), a recognition method of gas feature based on a global extended extreme learning machine (GEELM) is proposed, which combines the expansion factor and global balance coefficient to expand and balance the difference between categories, and improve the classification performance. Then this method is applied to identify the quality of tea. Firstly, the dragging factor and following matrix are introduced to increase the distance between classes. Secondly, the global identification coefficient is introduced further to increase the feature differences among different types of tea, and improve the classification stability. Finally, under different feature sets, the classification performance of multi-pattern recognition methods is compared to prove the effectiveness of GEELM in e-nose gas feature recognition. The results show that GEELM has the best classification accuracy of 98.20%, F1-score of 0.9871, and Kappa coefficient of 0.9775. In conclusion, GEELM can be an effective technique to identify gas features, and it also provides a new method for tea quality measurement.

Keywords
Electronic nose, feature recognition, global extended extreme learning machine, quality recognition

Introduction
Tea is rich in various chemical components with nutritional value and health functions and is one of the most popular beverages worldwide. It is very common for low-grade tea to pretend to be high-quality tea.¹ At present, the inspection method of tea quality mainly relies on artificial sensory evaluation, but the method has the disadvantages of strong subjectivity and poor reproducibility.² Meanwhile, gas chromatography,³ liquid chromatography,⁴ near-infrared spectroscopy,⁵ and other technologies can qualitatively analyze the chemical composition of tea. However, the detection of a single chemical component cannot show the overall quality of the tea. In the tea market, the smell of tea is an essential factor that affects consumers’ preferences, and the smell of tea of different qualities is different. Therefore, it is important to realize the effective identification of tea quality combined with a fast measurement method.

Electronic nose (e-nose) uses the gas array of cross sensitive to realize the classification decision of samples. It can reflect the overall gas information of the tested sample, and then reflect the overall quality.⁶ It has the advantages of simple operation, good reproducibility, and high detection accuracy. Therefore, this paper is based on the e-nose technology to detect the gas information reflecting the overall quality of tea. In using the e-nose technology to realize the gas identification of the sample, an effective gas classification method is an important factor in improving the detection performance of the e-nose. Therefore, an effective pattern recognition algorithm directly determines the detection performance of e-nose. In the study of pattern recognition methods, Sun et al. applied the integrated learning algorithm of support vector machines (SVM), random forest (RF), and linear discriminant analysis to identify gas information of Chinese herbal, it showed that the classification effect of integrated learning was better than that of single classifier.⁷ Qiao et al.⁸ used the tree parzen estimator method to optimize the light gradient boosting machine algorithm to identify the gas information of wine. Jiang et al.⁹ proposed an algorithm

¹Faculty of Electronic Information Engineering, Henan Polytechnic Institute, Nanyang, China
²Henan Material Forming Equipment Intelligent Technology Engineering Research Center, Nanyang, China

Corresponding author:
Chao Wang, Faculty of Electronic Information Engineering, Henan Polytechnic Institute, Nanyang 473000, China.
Email: wangchao751215@163.com

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
Based on the voting committee to realize intelligent classification of indoor pollutant gas. Luo et al.\textsuperscript{10} proposed a recognition algorithm of quantum particle swarm optimization to improve the performance of e-nose detection. Jia et al.\textsuperscript{11} proposed a multi-core support vector machine algorithm to identify the pollutant gas based on e-nose technology. Shi et al.\textsuperscript{12} tracked the origin of rice based on the multi-recognition technology, and the difference of gas features was visualized. Wang et al.\textsuperscript{13} proposed a transfer learning method combined with e-nose to identify the xihu longjing tea. Zhang et al.\textsuperscript{14} proposed a cross-domain discriminative subspace learning method to identify the gas information for multiple e-nose systems. In previous studies, researchers have proposed various pattern recognition methods to identify the gas features. However, due to the existence of certain nonlinearity and correlation between gas features, this requires high classification performance of the pattern recognition algorithm. The poor classification stability, slow training speed, and low classification stability are still shortcomings.

Extreme learning machine (ELM) is a supervised pattern recognition algorithm with simple structure, fast running speed, and strong classification performance,\textsuperscript{15} which was proposed by Guangbin Huang, professor of Nanyang University of technology in Singapore. It is a single hidden layer feedforward neural network, and classification can only be achieved by determining the connection weight between the hidden and output layers.\textsuperscript{15} ELM has been successfully applied in many fields. Janakiraman and Deva Priya\textsuperscript{16} used ELM to identify human DNA sequences and obtained better classification accuracy. Men et al. intelligently identified the fusion features of e-nose and electronic tongue based on the ELM algorithm to realize the classification of beer flavor information. Compared to SVM and RF, ELM obtained the best classification performance.\textsuperscript{15} Shi et al.\textsuperscript{12} obtained the gas information of rice from different origins based on the e-nose technology and traced the origin of the gas information of the rice based on the ELM algorithm. Lahoura et al. obtained online data of breast cancer based on cloud technology and combined it with ELM for online identification. Compared to other diagnostic technologies, ELM achieved the best classification performance.\textsuperscript{17} Yaseen et al.\textsuperscript{18} used the ELM algorithm to predict the drought index nonlinearly and provided an effective technical method for preventing drought. However, although ELM has achieved a good classification result, it still has the disadvantage of low classification stability because it is a kind of neural network. In the previous research, the research on the improved algorithm of ELM. Li et al.\textsuperscript{19} proposed radial basis function-extreme learning machine (RBF-ELM), combined the density peak clustering algorithm to select the center of RBF adaptively, and then realized the effective identification of mixed data. Zeng et al.\textsuperscript{20} combined ELM and dictionary learning to enhance the consistency of the data within the class, added the maximum margin standard regularization method to avoid over-fitting, and finally realized image classification. Diker et al. combined the differential evolution algorithm (DEA) and ELM to classify the EEG features extracted by the convolutional neural network. Compared with other pattern recognition algorithms, DEA-ELM achieved the best classification performance.\textsuperscript{21} Turkoglu\textsuperscript{22} proposed a multi-core ELM machine combined with a deep learning method to realize the detection of new coronaviruses in CT scan images. Hasani and Shabanlou\textsuperscript{23} proposed a weighted regularized (ELM) to estimate the side channel flow coefficient and compared it with the traditional ELM to prove its effectiveness. Zhao et al.\textsuperscript{24} proposed a weighted discriminative ELM to solve the classification problem of sample imbalance to realize the diagnosis of lung cancer gas. Wang et al.\textsuperscript{25} proposed an unsupervised discriminative domain reconstruction based on ELM to improve the detection performance of e-nose. Shi et al.\textsuperscript{26} proposed a novel L1 regularized ELM for soft-sensing of an industrial process. However, the neural network algorithms generally have the disadvantage of low classification stability. Still, most of them have not improved the structure of ELM itself, which will not essentially improve the classification stability of ELM. A small amount of research has improved the ELM itself, but the stability of the ELM has not been significantly improved. Meanwhile, the data differences between classes have not been effectively balanced. Therefore, this paper proposes a global extended extreme learning machine (GEELM) to enhance classification stability and performance. Meanwhile, compared with traditional ELM and other classical pattern recognition algorithms, the effectiveness of GEELM is verified in gas feature recognition of tea.

To improve the classification performance of the e-nose gas features, provide an effective detection method for the quality supervision of the tea market, this paper proposes a GEELM combined with e-nose for identifying tea gas information. In the GEELM, the dragging factor and following matrix are introduced to increase the distance between classes. Meanwhile, the global identification coefficient is introduced to increase the feature differences among different types of tea and improve the classification stability of ELM. The identification process of gas features based on GEELM is discussed in detail.

**Methods**

**ELM**

ELM randomly initializes the connection weight between the hidden and the output layers. When regularization learning rules are introduced to enhance the
classification performance, the ELM optimization problem can be transformed into the form of Formula (1)\textsuperscript{27}:
\[
\min \frac{1}{2} \|\beta\|^2 + C \cdot \frac{1}{2} \sum_{i=1}^{N} \|\xi_i\|^2
\]
\[
s.t. \ h(x_i)\beta = t_i - \xi_i \iff H\beta = T - \xi, \quad i = 1, 2, \cdots, N
\]
(1)

where \(\beta\) denotes the weight between the hidden layer and the output layer, \(\xi = [\xi_1, \xi_2, \cdots, \xi_i]\) represents the output error, \(C\) is the penalty factor in the training process, \(H = [h(x_1), h(x_2), \cdots, h(x_N)]\) denotes the output matrix of the hidden layer, \(T = [t_1, t_2, \cdots, t_i]\) is the category vector, and \(x_i\) is the network’s input. In the ELM classification process, output functions are divided into two categories.

Category 1: when there are few training samples, the output function can be expressed as\textsuperscript{28}:
\[
f(x) = h(x)\beta = h(x)HT\left(\frac{I}{C} + HH^T\right)^{-1}T
\]
(2)
where \(I\) is the identity matrix.

Category 2: when the number of training samples is too large, which is greater than the number of features, the output function can be expressed as\textsuperscript{28}:
\[
f(x) = h(x)\beta = h(x)\left(\frac{I}{C} + HH^T\right)^{-1}HT
\]
(3)

GEELM

Figure 1 shows the algorithm flow chart of GEELM. The improvement process of GEELM to ELM is as follows:

Set the output labels of different types of samples as \(T_i = [-1, \cdots, +1, \cdots, -1] \in \mathbb{R}^c, c\) is the number of categories, and the position corresponding to +1 is the sample category. An extended coefficient matrix \(A = a_{ij}\) is introduced to increase the distance between classes. In matrix \(A\), the elements of the matrix are greater than 0. Meanwhile, the following matrix \(Z\) is introduced, when \(T_{ij} = 1, z_{ij} = 1\), when \(T_{ij} = -1, z_{ij} = -1\). Take three categories as an example, and Table 1 shows the expanded form for the distance between classes.

For the first two categories output calculation, when the expansion coefficient matrix \(A\) is not introduced, the distance between the output categories can be expressed by Formula (4):
\[
\sqrt{(1 - (-1))^2 + (-1 - 1)^2 + (-1 - (-1))^2} = 2\sqrt{2}
\]
(4)

After adding expansion coefficient matrix \(A\), the distance between output classes of the hidden layer can be expressed by Formula (5):
\[
\sqrt{((1 + a_{11}) - (-1 - a_{21}))^2 + ((-1 - a_{12}) - (1 + a_{22}))^2 + ((-1 - a_{13}) - (1 + a_{23}))^2} \geq 2\sqrt{2}
\]
(5)

Figure 2 shows the extended form.

In this way, solving the problem can be expressed by the following Formula:
\[
\min \frac{1}{2} \|\beta\|^2 + C \cdot \frac{1}{2} \sum_{i=1}^{N} \|H\beta - T \odot A\|^2 \quad A > 0
\]
(6)

Set \(W = T + Z \odot A\), the optimization function can be expressed as:
\[
g(\beta) = \min \frac{1}{2} \|\beta\|^2 + C \cdot \frac{1}{2} \sum_{i=1}^{N} \|H\beta - W\|^2
\]
(7)

Because \(\frac{\partial g(\beta)}{\partial \beta} = 0\), the result is as follows:
\[
\beta^* = HT\left(\frac{1}{C} + H^T H\right)^{-1}W
\]
(8)
Since $A_{ij}$, the optimization problem is further transformed into:

$$
Z = \begin{bmatrix}
1, -1, -1 \\
-1, 1, -1 \\
-1, 1, -1
\end{bmatrix}
$$

Because of the calculation form of Hadamard product, Equation (10) can transform the multiplication form of elements, which can be expressed as:

$$
\min_i C \cdot \frac{1}{2} \sum_{i=1}^{N} \| U_j - Z \otimes A \|^2
$$

Set $U = H\beta - T$, Formula (9) can be transformed into:

$$
\min_i C \cdot \frac{1}{2} \sum_{i=1}^{N} \| U - Z \otimes A \|^2
$$

Because of the calculation form of Hadamard product, Formula (10) can transform the multiplication form of elements, which can be expressed as:

$$
\min_i (U_j - Z_{ij}A_{ij})^2
$$

Since $Z_{ij}^2 = 1$, therefore:

$$
(U_j - Z_{ij}A_{ij})^2 = (Z_{ij}U_j - A_{ij})^2
$$

The optimization problem is further transformed into:

$$
A_{ij} = \max (Z_{ij}U_j, 0)
$$

Since $A_{ij} > 0$, the following results are obtained:

$$
A = \max (Z \otimes U, 0)
$$

In the training process, the $\beta$ and $A$ can be obtained, and the output function can be transformed into:

$$
y = h(z)\beta = h(z)H^T(H^TH + \frac{I}{C})W
$$

To obtain the optimal $\beta$ and $A$, we set the target error is $10^{-4}$:

$$
\| A_{k+1} - A_k \|^2 - \| \beta_{k+1} - \beta_k \|^2 < 10^{-4}
$$

The optimal parameters are obtained. Where $k$ is the number of iterations, and the maximum is 50. Meanwhile, the global balance coefficient matrix $B = b_i$ is proposed to increase the difference between categories, thereby improving the classification stability of ELM.

$$
B = (x_i - \bar{x})(x_i - \bar{x})^T
$$

where $\bar{x}$ represents the mean value of all samples, and then the identification coefficient $b_i$ (i represents the i-th sample) can be calculated according to Formula (17). Then the identification coefficient and expansion factor are combined as follows:

$$
\begin{bmatrix}
a_{11} & a_{12} & a_{13} \\
a_{21} & a_{22} & a_{23} \\
a_{31} & a_{32} & a_{33}
\end{bmatrix}
\begin{bmatrix}
b_1 \\
b_2 \\
b_3
\end{bmatrix}
= \begin{bmatrix}
a_{11}x_1b_1 + a_{12}x_2b_1 + a_{13}x_3b_1 \\
a_{21}x_1b_2 + a_{22}x_2b_2 + a_{23}x_3b_2 \\
a_{31}x_1b_3 + a_{32}x_2b_3 + a_{33}x_3b_3
\end{bmatrix}
$$

Further, $W$ is transformed to $W' = T + Z \otimes A \otimes B$. So, the Formula (19) expressed the output function as follow:

$$
y = h(z)\beta = h(z)H^T(H^TH + \frac{I}{C})W'
$$

In this study, the Sigmoid is used as the transfer function. To pre-adjust the parameters for the best classification performance, the penalty factor $C$ is 0.02 after training and optimization, and the number of hidden layer neurons is 500.
Feature processing

Feature processing includes feature dimensionality reduction and feature selection. Feature dimensionality reduction is data linearly reconstruction on the original features to obtain some linearly independent feature vectors to represent the original detection data, which makes the features independent of each other. Feature dimensionality reduction is not a specific feature but new features after linear transformation of the original features. Feature selection is to delete some features that affect the classification performance from the original features to improve the classification ability of classifier.

There is an obvious correlation between the original detection features of e-nose, which will reduce the decision-making performance of the classifier. We used the feature dimensionality reduction method to process the original detection features to remove the correlation between the features. Principal component analysis (PCA), kernel principal component analysis (KPCA), and kernel entropy component analysis (KECA) were introduced to transform the original gas features. PCA transforms the original features into several linearly independent principal components by linear transformation. However, the extracted features had strong nonlinearity because the sensor array of cross sensitive. Therefore, KPCA first maps features to high-dimensional space based on the kernel function to reduce the nonlinear relationship between features and reduce the dimension of features in high-dimensional space. The radial basis kernel function was applied in this work, and the kernel function parameter was 10. KECA does not use eigenvalues to retain effective information. It uses Renyi entropy to represent the feature structure of input space, ensure the minimum information loss in dimension reduction, and retain the maximum information entropy of features. KECA is more conducive to finding the feature direction after dimensionality reduction.

In this work, to obtain the overall information of the e-nose detection feature, after pre-adjusting the cumulative variance contribution rate, we took the principal components with the cumulative method contribution rate of more than 90% to form the dimensionality reduction feature set.

Meanwhile, to compare with the feature dimensionality reduction method, based on the feature selection methods, the recursive feature elimination (RFE), maximum information coefficient (MIC), and maximum relevance minimum redundancy (MRMR) were applied to obtain the optimized feature set. The feature selection method sequentially accumulated the features according to their importance to obtain multiple feature sets. The optimized feature set was obtained when the classifier obtained the best classification accuracy.

Process of gas identification

In this work, GEELM was proposed and combined with e-nose technology to realize intelligent identification of gas information of different grades of tea. The main steps were as follows:

1. Obtain the gas information of different grades of tea based on PEN3 e-nose;
2. Feature extraction was carried out according to the characteristics of the detection signal of the e-nose;
3. The feature dimensionality reduction and feature selection methods were used to obtain the feature set after feature processing;
4. Different feature sets were identified based on GEELM and other multi-pattern recognition methods to verify the effectiveness of GEELM.

Experiment and discussion

Samples

Five different grades of Shucheng little orchid tea were used as experimental samples. The grade, type, and producing area of tea are shown in Table 2. Five grams of Shucheng little orchid tea were placed into a 200 ml beaker, 150 ml boiling distilled water was added, and covered with a surface plate. After soaking for 5 min, the tea leaves were filtered, and the tea soup was cooled to room temperature for analysis. Meanwhile, sufficient tea soup was prepared during the experiment.

E-nose

The PEN3 e-nose developed by German Airsense company was used to detect the volatile gas of tea. The e-nose system comprises a gas acquisition device, gas detection unit, and air purification device. The system structure is shown in Figure 3. The detection unit includes a sensor array and pattern recognition analysis system. The gas sensor array includes 10 metal oxide sensors, which can realize the gas information detection of cross sensitive. Its working principle is that the gas and the active material on the sensor’s surface undergo an oxidation-reduction reaction to produce a conductivity value. The response value of the sensor is defined

| No. | Grade       | Type     | Producing area |
|-----|-------------|----------|----------------|
| 1   | Superfine   | Green tea| Shucheng, Anhui|
| 2   | First-degree| Green tea| Shucheng, Anhui|
| 3   | Second-degree| Green tea| Shucheng, Anhui|
| 4   | Third-degree| Green tea| Shucheng, Anhui|
| 5   | Fourth-degree| Green tea| Shucheng, Anhui|
as \( G/G_0 \), where \( G \) is the sensor’s conductivity when the measured gas enters the sensor chamber, and \( G_0 \) is the conductivity of the sensor when the pure air enters the gas chamber. The parameters of each sensor are shown in Table 3.

### Experimental process

The ambient temperature for e-nose gas test was \( 25 \pm 5^\circ C \). The experimental parameters and steps determined by pre-experiment were as follows:

1. Take 50 ml of tea sample into 200 ml volumetric flask, seal it with preservative film and stand still for 20 min to ensure sufficient headspace air.
2. Clean air with a 300 ml/min flow rate treated by activated carbon was introduced into the sensor chamber for 60 s to clean the sensor chamber.
3. After cleaning and calibrating the sensors array, the volatile gas of tea was detected. Based on the detection method of dynamic headspace, a single sample was detected for 80 s, and the frequency of obtaining gas information was 1 Hz.
4. To avoid losing generality, 30 groups of parallel samples were prepared for tea of each grade, and 150 groups of data were obtained from five different grades of tea.

Figure 4 shows the sensor response curve of the e-nose in the process of gas detection. The \( x \)-axis is the detection time, and the \( y \)-axis is the response value of the sensors.

| Sensor | Substances detected                                      | Threshold value (ml/m³) |
|--------|----------------------------------------------------------|-------------------------|
| W1C    | Aromatics                                               | 10                      |
| W5S    | Nitrogen oxides                                         | 1                       |
| W3C    | Ammonia and aromatic molecules                          | 10                      |
| W6S    | Hydrogen                                                | 100                     |
| W5C    | Methane, propane, and aliphatic non-polar molecules      | 1                       |
| W1S    | Broad methane                                           | 100                     |
| W1W    | Sulfur-containing organics                              | 1                       |
| W2S    | Broad alcohols                                          | 100                     |
| W2W    | Aromatics and sulfur- and chlorine-containing organics   | 1                       |
| W3S    | Methane and aliphatics                                  | 10                      |

Figure 3. The structure diagram of e-nose system.

Figure 4. The response curve of e-nose sensor.
Feature extraction

Figure 5 shows the response radar chart of gas information of different grades of tea in 60 s, and the overall trend of the gas response was similar. The original detection signal contained a lot of redundant information. If the data was directly input to the pattern recognition algorithm, it would increase the classification difficulty of the classifier. Therefore, it was necessary to extract the features that could represent the overall characteristics of gas information. Figure 5 shows that WSS had the largest feature response, but it cannot be used for pattern recognition with the response features of a single sensor, which would cause overfitting of the model. Meanwhile, the high or low of sensor response information could not be used to judge the effectiveness of its features. Therefore, we extracted the different features to represent the overall properties of the original signal, including the steady-state mean value of 51–60 s, the maximum value, and the integral value of 0–60 s. Because the detection signal was stable, the time-frequency feature was not selected. The steady-state mean value represented the dynamic balance of gas volatilization and air intake in the stable detection state. The maximum value represented the peak value of gas volatilization concentration. The integral value was the dynamic feature of gas volatilization degree in the whole detection process. Formulas (20)–(22) shows
the calculation method of features. To accelerate the convergence speed of model, the 30 features contained in 10 sensors were normalized to the interval (0, 1).

\[
T_1 = \frac{1}{10} \sum_{i=51}^{60} g(t)
\]

\[
T_2 = \max |g(t)|
\]

\[
T_3 = \int_{0}^{60} g(t)dt
\]

where \(g(t)\) represents the response value at time \(t\), and \(T_1 - T_3\) represents the steady-state mean value of 51–60 s, the maximum value, and the integral value of 0–60 s.

**Feature processing**

Ten sensors contained a total of 30-dimensional gas features. If there were strong correlation between the features, it would increase the classification difficulty of the pattern recognition method. Based on the calculation method of Pearson correlation, Figure 6 shows the correlation heat map of 30-dimensional gas features. The correlation value range was \((-1, 1)\), and the value close to 1 indicated that the two features have a strong positive correlation. A value close to \(-1\) meant a strong negative correlation between the two features. It can be seen that there were positive and negative correlations between different features. Therefore, feature preprocessing methods were applied to obtain the optimal feature set of e-nose gas features.

For the feature dimensionality reduction, PCA, KPCA, and KECA were introduced to transform the original gas features to reduce the correlation between input features. Figure 7 shows the effect of feature dimension reduction. The cumulative contribution rate of PCA’s first two principal components was 68.91\%, KPCA was 74.81\%, KECA was 84.75\%. Obviously, after PCA dimensionality reduction, there were more overlapping samples between classes, KPCA had obvious dimensionality reduction effect, and KECA had the best dimensionality reduction effect. In this work, for the three feature dimensionality reduction methods, the feature vectors corresponding to the feature values whose cumulative principal component reached more than 90\% were used for pattern recognition. Finally, when the cumulative principal component reached more than 90\%, the number of principal components obtained by PCA, KPCA, and KECA were 8, 8, 5, respectively.

For the feature selection methods, the RFE, MIC, and MRMR were applied to obtain the optimized feature set. The feature selection method sequentially accumulated the features according to their importance to obtain multiple feature sets. The optimized feature set was obtained when the classifier obtained the best classification accuracy.

**Classification results**

Each grade of tea contained 30 samples. Based on the Kennard stone (KS) method, 20 samples were selected as the training set, and 10 samples were selected as the testing set. Then the training set contained 100 samples, and the testing set contained 50 samples. KS could make the sample distribution more extensive, which is conducive to strengthening the generalization ability of the model.31 The original fusion feature contained 30 dimensions. To highlight the classification effectiveness of GEELM, the backpropagation neural network (BP), radial basis function neural network (RBF), probabilistic neural network (PNN), learning vector quantization neural network (LVQ), SVM, RF, ELM, and kernel extreme learning machine (KELM) under different feature sets were compared and analyzed. After training and parameter tuning, and Table 4 shows the average accuracy of each algorithm for 10 runs. Overall, the accuracy of the feature set formed by KECA was better than other feature sets, which showed the effectiveness of the KECA feature processing method. BP had the lowest classification performance in the performance comparison of each classification method. BP is a classic neural network algorithm with many trainable
parameters, which requires many samples for training. The detection data of the e-nose had the characteristic of a few samples. Therefore, BP had over-fitted, and the classification performance was poor. Compared with BP, other algorithms had better classification performance because of their fewer trainable parameters. Importantly, GEELM achieved better classification performance in each feature set, and 98.20% accuracy was obtained in the KECA feature set, which was better than the traditional ELM and KELM method.

When a high classification accuracy is obtained, the classifier should have a good stability. To visualize classification stability, we selected the box diagram to show the classification stability of multi-classifiers. Figure 8 shows the visualization diagram of classification stability of ELM, KELM, and GEELM under different feature sets. In the 10 calculation results, the maximum classification accuracy value was indicated by the upper boundary of the box plot. The lower boundary indicated the minimum value, the black line indicated the median value, and the singular point was outside the box diagram. The traditional ELM classification stability was poor, and the singular point appeared in the original features. Compared with ELM and KELM, GEELM had obtained better classification stability under each feature set, which showed that the

Table 4. The classification accuracy of multi pattern recognition algorithm under different feature sets.

| Data set    | BP (%) | RBF (%) | PNN (%) | LVQ (%) | SVM (%) | RF (%) | ELM (%) | K-ELM (%) | GEELM (%) |
|-------------|--------|---------|---------|---------|---------|--------|---------|-----------|-----------|
| Original features | 70.20  | 79.00   | 81.20   | 83.40   | 89.60   | 87.80  | 88.60   | 90.40     | 90.60     |
| PCA         | 82.80  | 89.40   | 84.00   | 88.80   | 94.00   | 94.40  | 93.20   | 93.20     | 95.40     |
| KPCA        | 79.80  | 88.70   | 86.60   | 90.60   | 93.80   | 94.00  | 93.60   | 94.40     | 97.00     |
| KECA        | 85.80  | 91.80   | 91.20   | 90.80   | 95.00   | 94.40  | 94.00   | 95.80     | 98.20     |
| RFE         | 86.00  | 91.20   | 90.60   | 90.20   | 94.80   | 93.60  | 93.60   | 95.20     | 97.60     |
| MIC         | 85.40  | 92.00   | 90.80   | 90.60   | 95.00   | 94.40  | 93.20   | 94.00     | 97.80     |
| MRMR        | 84.60  | 90.80   | 92.40   | 89.60   | 94.00   | 94.40  | 92.80   | 95.20     | 97.80     |

Figure 8. The box diagram of classification stability: (a) ELM, (b) KELM, and (c) GEELM.
algorithm improved classification accuracy and enhanced the classification stability.

To evaluate the classification performance of the classifier under multi-indexes, this paper introduced the evaluation methods of F1-score and Kappa coefficient. The F1-score value is close to 1, which shows that the classifier has strong comprehensive classification ability. The Kappa coefficient is close to 1, indicating that the actual and predicted output is consistent. Figure 9 shows each classifier’s best F1-score and Kappa coefficient under different feature sets. After calculation, all algorithms got the best parameter values in the KECA feature set. Compared with other multi-classifiers, GEELM had obtained the best F1-score of 0.9871 and Kappa coefficient of 0.9775.

Conclusion
To identify the gas information of different grades of tea and improve the detection performance of the e-nose. The main contributions and conclusions are as follows:

(1) An improved ELM algorithm (GEELM) is proposed, which improves the ELM optimization problem. It combines the extended coefficient matrix and global identification factor to improve the accuracy and stability of e-nose feature recognition.

(2) In multiple feature sets, the classification performance of multi-pattern recognition algorithm is compared and analyzed. GEELM achieves the highest classification accuracy of 98.20%, the best F1-score of 0.9871, and the best Kappa coefficient of 0.9775.

(3) Based on the visualization results of the box plot, GEELM has the best classification stability.

In conclusion, GEELM provides an effective pattern recognition method to identify the detection information of the e-nose. Moreover, it also provides a method for monitoring the quality of tea grades. In future research, the generalization ability of ELM can be further studied. In the case of noisy input, ELM can also effectively identify key information.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Pre-Research Project of General Armaments Department (41419070201).

ORCID iD
Chao Wang https://orcid.org/0000-0002-1122-2720

References
1. Koláčková T, Sumczynski D, Bednářik V, et al. Mineral and trace element composition after digestion and leaching into matcha ice tea infusions (Camellia sinensis L). J Food Compos Anal 2021; 97: 103792.
2. Shi Y, Liu X, Yin C, et al. A novel kohonen one-class method for quality control of tea coupled with artificial lipid membrane taste sensors. Anal Methods 2020; 12: 1460–1468.
3. Ly TK, Ho TD, Behra P, et al. Determination of 400 pesticide residues in green tea leaves by Upk-Ms/Ms and Gc-Ms/Ms combined with quechers extraction and mixed-mode SPE clean-up method. Food Chem 2020; 326: 126928.
4. Cardoso VGK and Poppi RJ. Non-invasive identification of commercial green tea blends using NIR
spectroscopy and support vector machine. *Microchem J* 2021; 164: 106052.
5. Wang YJ, Jin SS, Li MH, et al. Onsite nutritional diagnosis of tea plants using micro near-infrared spectrometer coupled with chemometrics. *Comput Electron Agric* 2020; 175: 105538.
6. Shi Y, Yuan H, Xiong C, et al. Improving performance: a collaborative strategy for the multi-data fusion of electronic nose and hyperspectral to track the quality difference of rice. *Sens Actuators B Chem* 2021; 333: 129546.
7. Sun X, Liu L, Wang Z, et al. An optimized multi-classifiers ensemble learning for identification of ginsengs based on electronic nose. *Sens Actuators A Phys* 2017; 266: 135–144.
8. Qiao M, Zhang L and Mu F. Research on discriminating wine varieties based on electronic nose and lightGBM algorithm. *Food Mach* 2020; 36: 76–79.
9. Jiang X, Jia P, Luo R, et al. A novel electronic nose learning technique based on active learning: EQBC-RBFNN. *Sens Actuators B Chem* 2017; 249: 533–541.
10. Luo H, Jia P, Qiao S, et al. Enhancing electronic nose performance based on a novel QPSO-RBM technique. *Sens Actuators B Chem* 2018; 259: 241–249.
11. Jia P, Meng F, Cao H, et al. Training technique of electronic nose using labeled and unlabeled samples based on multi-kernel LAPSVM. *Sens Actuators B Chem* 2019; 294: 98–105.
12. Shi Y, Jia X, Yuan H, et al. Origin traceability of rice based on an electronic nose coupled with a feature reduction strategy. *Meas Sci Technol* 2021; 32: 025107.
13. Wang X, Gu Y and Liu H. A transfer learning method for the protection of geographical indication in China using an electronic nose for the identification of xihu longjing tea. *IEEE Sens J* 2021; 21: 8065–8077.
14. Zhang L, Liu Y and Deng P. Odor recognition in multiple E-nose systems with cross-domain discriminative subspace learning. *IEEE Trans Instrum Meas* 2017; 66: 1679–1692.
15. Men H, Shi Y, Fu S, et al. Mining feature of data fusion in the classification of beer flavor information using E-tongue and E-nose. *Sensors* 2017; 17: 1656–1672.
16. Janakiraman S and Deva Priya M. Advanced extreme learning machine-based ensemble classification scheme with enhanced data perturbation for human DNA sequences. *Comput Intell* 2021; 37: 1890–1915.
17. Lahoura V, Singh H, Aggarwal A, et al. Cloud computing-based framework for breast cancer diagnosis using extreme learning machine. *Diagnostics* 2021; 11: 241.
18. Yaseen ZM, Ali M, Sharafati A, et al. Forecasting standardized precipitation index using data intelligence models: regional investigation of Bangladesh. *Sci Rep* 2021; 11: 3435.
19. Li Q, Xiong Q, Ji S, et al. A method for mixed data classification base on RBF-ELM network. *Neurocomputing* 2021; 431: 7–22.
20. Zeng Y, Li Y, Chen J, et al. ELM embedded discriminative dictionary learning for image classification. *Neural Netw* 2020; 123: 331–342.
21. Diker A, Sönmez Y, Özyurt F, et al. Examination of the ECG signal classification technique DEA-ELM using deep convolutional neural network features. *Multimed Tools Appl* 2021; 80: 24777–24800.
22. Turkoglu M. Covid-19 detection system using chest CT images and multiple kernels-extreme learning machine based on deep neural network. *IRBM* 2021; 42: 207–214.
23. Hasani F and Shabanlou S. Weighted regularized extreme learning machine to model the discharge coefficient of side slots. *Flow Meas Instrum* 2021; 79: 101955.
24. Zhao L, Qian J, Tian F, et al. A weighted discriminative extreme learning machine design for lung cancer detection by an electronic nose system. *IEEE Trans Instrum Meas* 2021; 70: 1–9.
25. Wang Z, Yan J, Chen F, et al. Sensor drift compensation of E-nose systems with discriminative domain reconstruction based on an extreme learning machine. *IEEE Sens J* 2021; 21: 17144–17153.
26. Shi X, Kang Q, An J, et al. Novel L1 Regularized extreme learning machine for soft-sensing of an industrial process. *IEEE Trans Ind Inform* 2022; 18: 1009–1017.
27. Zhang L and Zhang D. Evolutionary cost-sensitive extreme learning machine. *IEEE Trans Neural Netw Learn Syst* 2017; 28: 3045–3060.
28. Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification. *IEEE Trans Syst Man Cybern Part B* 2012; 42: 513–529.
29. Shi Y, Liu M, Sun A, et al. A fast Pearson graph convolutional network combined with electronic nose to identify the origin of rice. *IEEE Sens J* 2021; 21: 21175–21183.
30. Jia X-M, Meng Q-H, Jing Y-Q, et al. A new method combining KECA-LDA with elm for classification of Chinese liquors using electronic nose. *IEEE Sens J* 2016; 16: 8010–8017.
31. Li Y, Zhang J, Li T, et al. Geographical traceability of wild boletus edulis based on data fusion of FT-MIR and ICP-AES coupled with data mining methods (SVM). *Spectrochim Acta Part A-Mol Biomol Spectrosc* 2017; 177: 20–27.