Journal of Network Intelligence ©2023 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **8**, Number **4**, November **2023**

Emotion Recognition Based on Optimized Generalized Regression Neural Network

You-Yi Chai*

School of Automotive Technology
ZheJiang Technical Institute of Economics
Hangzhou 310018, P.R. China
13588472699@163.com

Yu-Hao Ma

Business School
The University of Sydney
Sydney 2008, Australia
yhma.leo@gmail.com

*Corresponding author: You-Yi Chai

Received April 17, 2023, revised June 26, 2023, accepted August 29, 2023.

Abstract. *The existing intelligent learning system not only provides personalized service, but also begins to pay attention to the personalized analysis of learners' emotional level. Therefore, intelligent learning system needs to make adaptive adjustment according to learners' emotional state. Therefore, it is of great practical significance to realize the emotion recognition ability of intelligent learning system. In order to solve the above problems, this work introduces generalized regression neural network in the field of deep learning to analyze the emotional characteristics of multimedia resources in intelligent learning system and its interaction with learners' emotions. Firstly, the emotional classification coding system and coding strategy of multimedia resources are designed according to certain purposes and principles. Secondly, in order to better remove noise and improve the accuracy of emotion recognition, the only adjustable parameter (smoothing factor) in generalized regression neural network is optimized by using fruit fly optimization method. Finally, the optimized generalized regression neural network is used to identify the emotion of multimedia resources and learners' facial emotion synchronously, and the influence degree of emotion on learners' emotion is obtained by calculating the correlation coefficient. The experimental results show that the proposed method can quickly and accurately judge the emotional types of multimedia resources and learners, and the emotional types of multimedia resources have a great influence on learners' emotions (positive/negative).*

Keywords: Emotional recognition; Deep learning; Generalized regression neural network; Fruit fly optimization

---

1. **Introduction.** Cloud computing, internet plus, artificial intelligence, big data and other emerging industries are increasingly applied in the field of education [1-4]. The learning system based on artificial intelligence needs to perceive the emotional state of learners and make adaptive knowledge content push. In the intelligent learning system, it is necessary to stimulate the emotional state of learners reasonably to ensure that the learning efficiency is maintained at a high level [5, 6]. Various forms of electronic screens in intelligent learning environment have affected emotional communication [7-10]. In addition to providing personalized learning resources for learners, intelligent learning

system should also make adaptive adjustments according to the emotional state to ensure that learners maintain a relatively positive emotional state [11-14]. Therefore, it is of great practical significance to realize emotion recognition in intelligent learning system, and it has become an important topic to be solved urgently in intelligent learning system.

In an ideal intelligent learning system, each learner learns by holding a smart device with a screen size similar to that of a paper textbook, or by relying on other forms of electronic screens to access knowledge content [15]. In this condition, learners acquire knowledge primarily by viewing various multimedia resources displayed by the smart device, and undergo cognitive and emotional changes in the process of viewing [16, 17]. Research has shown that human senses in the proportion of access to knowledge information, vision accounted for 83%, hearing accounted for 11%. Therefore, multimedia resources are effective in eliciting learners' visual perception and helping learners to actively participate in classroom activities in an intelligent learning system. In addition, multimedia resources have a subtle influence on learners' emotions through their vivid images, clear presentation and colourful layouts, while conveying knowledge. For example, multimedia resources can create a psychological feeling of excitement and pleasure for learners, thus effectively stimulating their interest and enthusiasm for learning. In the smart classroom, various electronic screens of different sizes and forms will become the main tool for learners to interact with smart devices [18, 19]. The multimedia resources presented on the screen are an important channel for transmitting teaching information and influencing emotional attitudes.

However, most of the existing research on image emotional features has focused on natural images, and less on computer-generated images such as multimedia graphic images. Furthermore, validating the emotional validity of multimedia resources based on learners' emotional feedback (learners' facial expressions) remains the main issue to be addressed in emotional interaction techniques for intelligent learning systems. Therefore, the purpose of this study is to analyze the emotional characteristics of multimedia resources and their effects on learners at the emotional and cognitive levels, so as to provide theoretical and technical support for improving the issue of emotional interaction in intelligent learning systems.

1.1. **Related Work.** At present, with the rapid development of technologies such as Internet+, cloud computing, Internet of Things, artificial intelligence and mobile Internet, intelligent learning systems are also beginning to be widely used. For example, Zhao [20] suggest that smart classrooms with smartphones and tablets can meet the individual needs of different learners. Learners can access task requirements and course materials anytime and anywhere. Teachers can assess learning outcomes and monitor the learning process anytime and anywhere, thus effectively facilitating effective teacher-student communication.

Multimedia resources are an important interface and medium for learners to receive knowledge and information. Learners can complete their knowledge understanding and emotional construction in the process of viewing multimedia resources. It is of great value to investigate the influence of emotional factors of multimedia resources on learners' emotions. For this reason, many experts and scholars have conducted research on the emotional design and application of multimedia resources from different aspects. Weng et al. [21] investigated whether multimedia resources based on emotional design could improve students' learning outcomes. The emotional design of multimedia resources refers to the anthropomorphic treatment of graphic elements, aiming to increase the visual impact and attractiveness, so that multimedia resources have certain emotional characteristics. Wright and Richards [22] demonstrated that emotionally charged images or videos had a positive effect on learners' memory, learning and knowledge transfer. This shows that emotionally designed multimedia resources can effectively stimulate learners' positive emotions, so that learners can quickly enter the learning state.

In order to improve the emotional interaction technology in intelligent learning systems, this paper analyses the emotional characteristics of multimedia resources and designs an emotional classification coding system and coding strategies for multimedia resources according to certain purposes and principles. In addition, the interaction between the emotions of multimedia resources and learners' emotions is also investigated. The emotion recognition of multimedia resources and the facial emotion recognition of learners are the prerequisites for realising emotional interaction in a smart learning environment.

For emotion recognition in multimedia resources and facial emotion recognition for learners, the data source is digital images. Recently, several researchers have proposed the use of radial basis function neural networks in the field of image recognition. A radial basis function neural network is a forward neural network with a three-layer structure. Alam et al. [23] used a fuzzy radial basis function neural network to achieve segmentation of ultrasound images and improve the recognition accuracy. Luo et al. [24] proposed a remote sensing image classification method based on an improved radial basis function neural network. Compared with radial basis function neural networks, generalized regression neural networks require fewer parameters to be adjusted and converge faster. Therefore, this paper uses generalized regression neural networks in the field of deep learning to identify the emotions of multimedia resources and learners' facial emotions simultaneously, and calculates the degree of influence of picture emotions on learners' emotions by correlation coefficients.

1.2. **Motivation and contribution.** (1) The emotion classification coding system and coding strategy of multimedia resources were designed according to certain purposes and principles, in order to subsequently adopt generalized neural network for emotion recognition of multimedia resources. At the same time, the generalized regression neural network is used to identify the emotions of multimedia resources and learners' facial emotions simultaneously, and the degree of influence of picture emotions on learners' emotions is calculated by correlation coefficients.

(2) In order to better remove the noise and thus improve the accuracy of emotion recognition, the unique adjustable parameter (smoothing factor) in the generalized regression neural network is optimized by using the fruit fly optimization method. The experimental results show that the proposed algorithm has higher peak signal-to-noise ratio and lower normalized mean square error than radial basis neural network and traditional generalized regression neural network, which verifies its effectiveness and advancement.

2. Emotional classification of multimedia resources.

2.1. **A model for describing the emotions of multimedia resources.** By establishing a model for describing the emotion of multimedia resources, the emotion characteristics and intensity values of any digital image can be described, thus effectively helping us to classify multimedia resources at the emotion level.

To study the emotion of multimedia resources, it is necessary to set up a suitable feature vocabulary to describe the emotion of the image. There is no uniformity in the description of image emotion, and researchers have mostly defined emotion models according to different domains. In addition, existing research on image emotion has been mixed, including both natural images acquired using digital cameras and computer-generated images.

In this paper, based on Ekman's emotion classification theory [25], 12 separate terms of cheerful, lively, funny, exaggerated, humorous, interesting, boring, dull, tedious, unreal, surprised and fearful are used to describe the emotional characteristics of multimedia resources. In addition, each emotional characteristic is classified into 6 levels of intensity, where 0 represents the lowest intensity and 5 represents the highest intensity. The constructed emotion description model is shown in Figure 1.

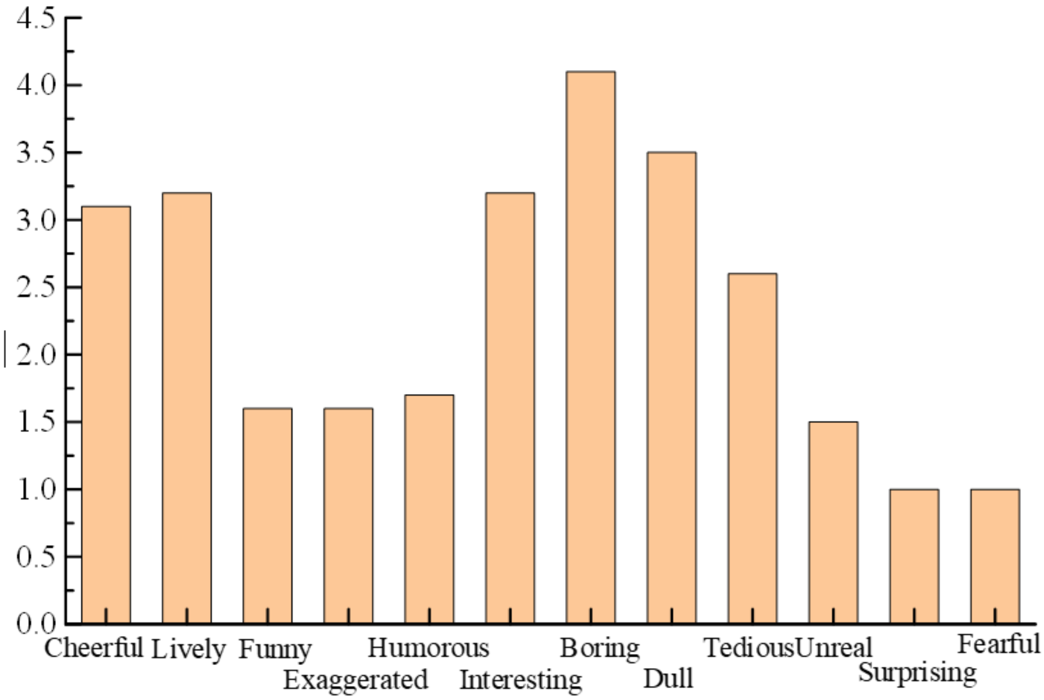

Figure 1. Emotional description model of multimedia resources

2.2. **Classification and coding of multimedia resources.** The current classification of digital learning resources is mostly based on the attributes and characteristics of the resources, and on this basis other classification features are added in conjunction with the research direction. The classification of multimedia resources requires precise positioning in order to meet the classification habits of resource users.

At the same time, the classification of multimedia resources should also have a certain degree of extensibility and universality [26, 27]. On this basis, combined with the existing classification standards and principles, this study believes that the classification of multimedia resources must contain four dimensions: applicable objects, subject classifications, teaching methods and emotional types.

This study divided the levels of the applicable target population. The finalized applicable targets include five levels of early childhood, primary, secondary, university, and social education. The subject classifications were identified as humanities, engineering, arts, skills and accomplishment. Accomplishment includes character education, social education and integrated practice. Combining the classification of existing resource forms, this study classifies the teaching methods of multimedia resources into 10 types of PPT, text, image, video, animation, webpage, online examination, APP software, practical training and extracurricular activities. As shown above, this study classifies the emotions of multimedia resources into 12 types: cheerful, lively, funny, exaggerated, humorous, interesting, boring, dull, tedious, unreal, surprising and fearful.

With the rapid development of the Internet, the total amount of multimedia resources has also shown an exponential growth trend. In the research process, it is difficult to manage such a large number of multimedia resources efficiently without a reasonable coding method. After the coding of multimedia resources, regular expressions can be used to quickly retrieve and classify multimedia resources with the same attributes. In addition, when using deep learning techniques for classification training of sentiment features, the input data can be automatically labelled in the data pre-processing stage to improve training efficiency. Guided by the principles of uniqueness, rationality and expandability, the coding structure of multimedia resources is shown in Figure 2.

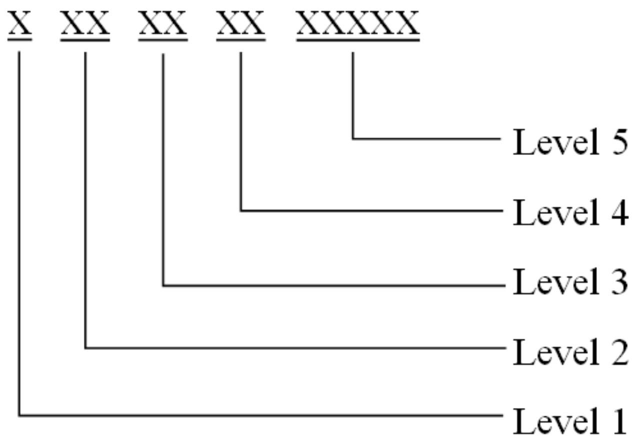

Figure 2. Coding structure of multimedia resources

The coding structure is clear and easy to understand, which will help the efficiency of managing multimedia resources and support the next step of emotion recognition research. In the coding structure designed in this paper, the first level represents the applicable object, the second level represents the subject classifications, the third level represents the teaching method, the fourth level represents the type of emotion and the fifth level is the sequential incremental code. The sequential incremental codes ensure the uniqueness of the coded object in the classification system. The coding principles for multimedia resources are shown in Table 1. For example, if the object of a multimedia resource is a university, the subject is humanities, the teaching method is PPT, the emotion type is funny and the type screen is not numbered, the code of the multimedia resource is expressed as 401010300001.

Table 1. Coding principles for multimedia resources

Level	Projects	Coding	Placeholder
1	Applicable objects	1- Early childhood; 2- Primary; 3- Secondary; 4- University; 5- Social education	1
2	Subject classifications	01-Humanities, 02-Engineering, 03-Arts, 04-Skills ; 05-Accomplishment	2
3	Teaching style	01-PPT; 02-Text; 03-Images; 04-Videos; 05-Animations; 06-Webpages; 07-Online exams; 08-APP software; 09-Practical training ; 10-Extracurricular activities	2
4	Type of emotion	01 -Cheerful; 02 - Lively; 03 - Funny; 04 - Exaggerated; 05 - Humorous; 06 - Interesting; 07 - Boring; 08 - Dull; 09 - Tedious; 10 - Unreal; 11 -Surprised; 12 -Fearful	2
5	No.	00001-99999	5

3. Emotion recognition based on optimised generalised regression neural networks.

3.1. **Generalized regression neural network model.** For both emotion recognition of multimedia resources and facial emotion recognition of learners, the data sources are digital images. As there is only one adjustable parameter (smoothing factor) in the generalised regression neural network, there are very few parameters to adjust.

In addition, generalised regression neural networks can largely reduce the influence of subjective factors, which helps to improve the accuracy of emotion recognition [28, 29]. Therefore, this paper uses generalised regression neural networks in the field of deep learning to simultaneously recognise the emotions of multimedia resources and learners' facial emotions. The emotion recognition and interaction in the intelligent learning environment is shown in Figure 3.

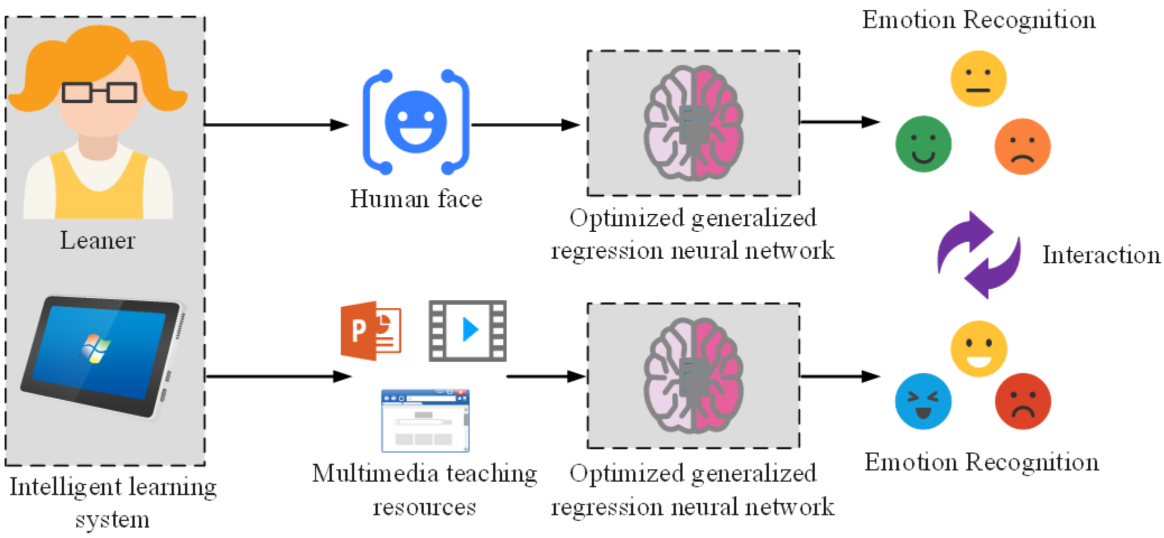

Figure 3. Emotion recognition and interaction in an intelligent learning environment

The structure of a generalised regression neural network is similar to that of a radial basis neural network, but consists of four layers: an input layer, a pattern layer, a summation layer and an output layer. A generalised regression neural network is essentially a non-linear regression radial basis neural network. When the number of training samples is small, the generalised regression neural network still has good local non-linear approximation ability and is fast to train. Figure 4 shows the structure of the generalized regression neural network network used in this paper. Where the vector $X = [x_1, x_2, ..., x_n]^T$ denotes the input to the network and the vector $Y = [y_1, y_2, ..., y_n]^T$ denotes the output of the network. The number of neurons in the input layer is the same as the dimensionality of the input vector (training sample), and the neurons in the input layer are fully connected to the neurons in the pattern layer. The number of neurons in the mode layer is kept the same as the number of neurons in the input layer. In addition, each neuron of the generalized regression neural network is represented as a radial basis function. The transfer function in the radial basis function is p_i .

$$\begin{cases} I_i = \sum_{j=1}^{n} w_{ji} x_j - \theta_i \\ y_i = f(I_i) \end{cases} \tag{1}$$

$$p_i = \exp\left[-\frac{(X - X_i)^T (X - X_i)}{2\sigma^2}\right] \quad i = 1, 2, ..., n \tag{2}$$

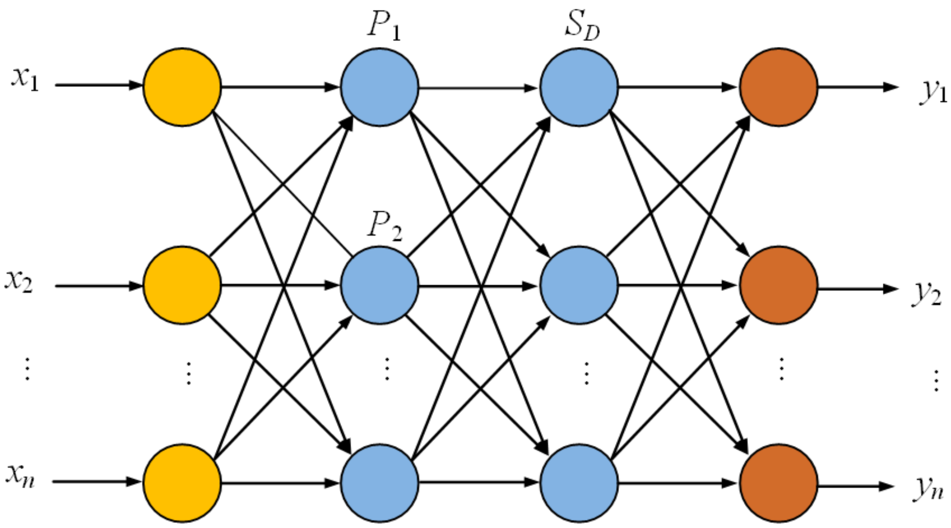

Figure 4. Structure of the generalized regression neural network

where X is the system input variable, X_i is the training sample used by neuron i, and is the diffusion parameter (smoothing factor). The output of neuron i in the pattern layer is a value expressed as an exponential square.

The activation function of the neural network is as follow:

$$\phi\left(x_p - c_i\right) = \exp\left(-\frac{\left\|x_p - c_i\right\|^2}{2\sigma^2}\right) \tag{3}$$

where $\|\cdot\|$ is the Euclidean norm and σ is the variance of Gaussian function.

$$\sigma = \frac{1}{p}\sum_{j}^{m}\left\|d_j - y_j c_i\right\|^2 \tag{4}$$

$$\sigma_i = \frac{c_{\max}}{\sqrt{2h}} i = 1, 2, \cdots, h \tag{5}$$

where c_{max} is the maximum distance from the selected center and h is the number of initial training data.

Assuming that the values of the weight parameters are all 1, the expression of the transfer function in the summation layer is shown as follow.

$$S_D = \sum_{i=1}^{n} P_i \tag{6}$$

where P_i denotes the transfer function of neuron i. In the output layer, the output value of the generalized regression neural network $\hat{Y}(x)$ is calculated as shown as follow:

$$y_j = \sum_{i=1}^{h}\omega_{ij}\exp\left(-\frac{1}{2\sigma^2}\left\|x_p - c_i\right\|^2\right) \quad 1 \le j \le m \tag{7}$$

$$\hat{Y} = E(y/X) = \frac{\int_{-\infty}^{\infty} y f(X, y)dy}{\int_{-\infty}^{\infty} f(X, y)dy} \tag{8}$$

where $\{x_i, y_i\}_{i=1}^{n}$ denotes the training sample data, n denotes the number of training samples, $\hat{Y}(x)$ denotes the predicted output value and the function $f(X, y)$ is the joint density function.

Since $\int_{-\infty}^{\infty} ze^{-t^2}dz = 0$, the output of the generalized regression neural network after integration calculation can be expressed as $\hat{Y}(x)$.

$$\hat{Y} = \hat{f}(x) = \frac{\sum_{i=1}^{n} Y_i \cdot \exp\left[\frac{(X_i - X_i)^T (X_i - X_i)}{2\sigma^2}\right]}{\sum_{i=1}^{n} \exp\left[\frac{(X_i - X_i)^T (X_i - X_i)}{2\sigma^2}\right]} \tag{9}$$

The larger the value of the smoothing factor σ, the smoother the resulting function approximation. Depending on the practical application area, several trials can be used to find the optimum value of σ.

We can represent the digital image to be recognised as a collection of light pulses (related to the information of the pixel points in the image). Thus, a mathematical function *F(x,y)* can be used to represent a digital image. In a two-dimensional coordinate system, each pixel point has a defined range of colour space values. That is, the image can be understood as a mapping *f:U→C.* where U is a subset of the plane and C is a vector. To simplify the analysis, assuming that C is a one-dimensional vector, the image function can be expressed as *z.*

$$z = f(x, y) \tag{10}$$

where z denotes the intensity of each pixel point. Since there is only one adjustable parameter (smoothing factor) in the generalised regression neural network. Therefore, the number of parameters to be adjusted is very small. Generalised regression neural networks are able to reduce the influence of subjective factors to a large extent and thus improve the accuracy of image emotion recognition. In the toolbox provided by MATLAB, the corresponding function for a generalised regression neural network is *newgrnn().*

$$[net, tr] = newgrnn(\mathbf{P}, \mathbf{T}, Spread) \tag{11}$$

where T denotes the output vector, P denotes the input vector, Spread denotes the smoothing factor, net denotes the corresponding network and tr denotes the final return. The size of the training sample image is 256 × 256, so the input vector P is set to a 256 × 256 matrix. The output vector T is a 1 × 256 matrix.

Often, the digital images that people see are images that have been contaminated with noise. Noise contamination of digital images usually occurs during the generation and transmission of the image. In general, noise directly affects the interpretation and representation of information in the original image to some extent. The adverse effects of noise on the original image are divided into two main areas: (1) interference with people's subjective visual effects. The subjective visual effect of an original image contaminated by noise is often much reduced. If the noise affects the image to a high degree, then local detailed information in the image will be lost to some extent. (2) Noise can lead to misrepresentation of an image, which in turn reduces the accuracy of various image processing applications, such as image recognition.

It should be noted that the choice of the Spread value affects the curve results of the radial basis function, which directly determines the noise immunity of the generalised regression neural network. In order to reduce the influence of noise and thus improve the accuracy of image emotion recognition, this paper uses the fruit fly optimization method to optimize the smoothing factor *Spread.*

3.2. **Parameter optimisation methods.** The fruit fly optimization algorithm is a new approach to global optimization [30] that has been widely used in science and engineering. Nowadays, the fruit fly optimisation algorithm has been successfully applied to solving extreme values of mathematical functions, support vector machine parameter optimisation and generalised regression parameter optimisation. The principle of fruit fly population searching for food is shown in Figure 5. In order to reduce the influence of noise and

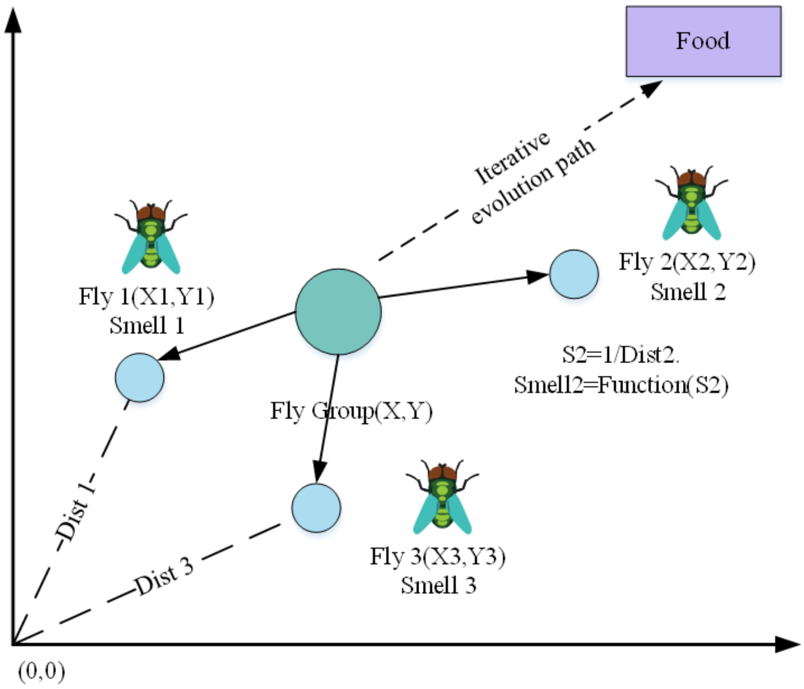

Figure 5. Principles of food search in fruit fly colonie

thus improve the accuracy of image emotion recognition, this paper uses the fruit fly optimization method to optimize the smoothing factor *Spread*, as follows.

Step 1: Set the number of fruit fly populations and the number of iterations to 100. The initialized positions of the fruit fly population are randomly distributed in the range [0,2] and the flight range is [-10,10].

Step 2: Initialize the smoothing factor Spread, the optimal position (X_{bset}, Y_{bset}) and the individual flight range of the fruit fly.

Step 3: Calculate the flavour concentration value S.

$$S_i = \frac{1}{Dist_i} \tag{12}$$

where $Dist_i$ represents the distance from the coordinate point of the individual to the origin of the coordinate system, i.e. $Dist_i = \sqrt{x_i^2 + y_i^2}$.

Step 4: Construct an adaptive degree function D_i using root mean square error in order to determine the flavour concentration value S, i.e. $D_i = F(S_i)$.

Step 5: Seek the initial extreme value based on the adaptive degree function, i.e. find the location of the individual fruit fly with the highest flavour concentration.

Step 6: Compare the current best flavour concentration with the best flavour concentration of the previous generation. If the former is greater than the latter, go to Step 7, otherwise go to Step 3. Save the best flavour concentration value and the corresponding coordinates.

$$\begin{cases} Smellbest = bestSmell \\ X\_axis = X\,(bestindex) \\ Y\_axis = Y\,(bestindex) \end{cases} \tag{13}$$

where $Smellbest$ is the concentration value of the best taste, [bestSmell, bestIndex]=max(Smell).

Step 7: Determine if the maximum number of iterations has been reached. If yes, then the current optimal flavour concentration value is the optimal *Spread* parameter.

Step 8: The optimal *Spread* parameters are brought into the designed generalized regression neural network model to obtain an optimized generalized regression neural network in order to improve the accuracy of image emotion recognition.

3.3. **Emotion recognition in multimedia resources.** In this study, annotators were recruited for a fee to annotate a sample of media teaching resources. Necessary training was given to the annotators before the annotation process began in order to minimise the influence of subjectivity. Mean absolute error (MAE) was used in this study to effectively monitor the degree of difference between the predicted and observed values.

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m} |x_i - m(x)| \tag{14}$$

Where MAE denotes the mean absolute error value derived from the final calculation, m denotes the number of samples involved in the calculation, x denotes the predicted value of the i-th multimedia resource sentiment, and $m(x)$ denotes the corresponding actual observed value.

A Fruit fly-optimised generalised regression neural network was used to train and test the multimedia resources dataset. All samples in the multimedia resources dataset were standardised in terms of image size and image type in order to facilitate the inputting and training of the generalised regression neural network. The number of samples of each type was 1000, making a total of 12,000. From these 12,000 samples, 80% were randomly selected as training samples and the rest as test samples. The sentiment recognition results of the multimedia resources are shown in Figure 6. It can be seen that the error

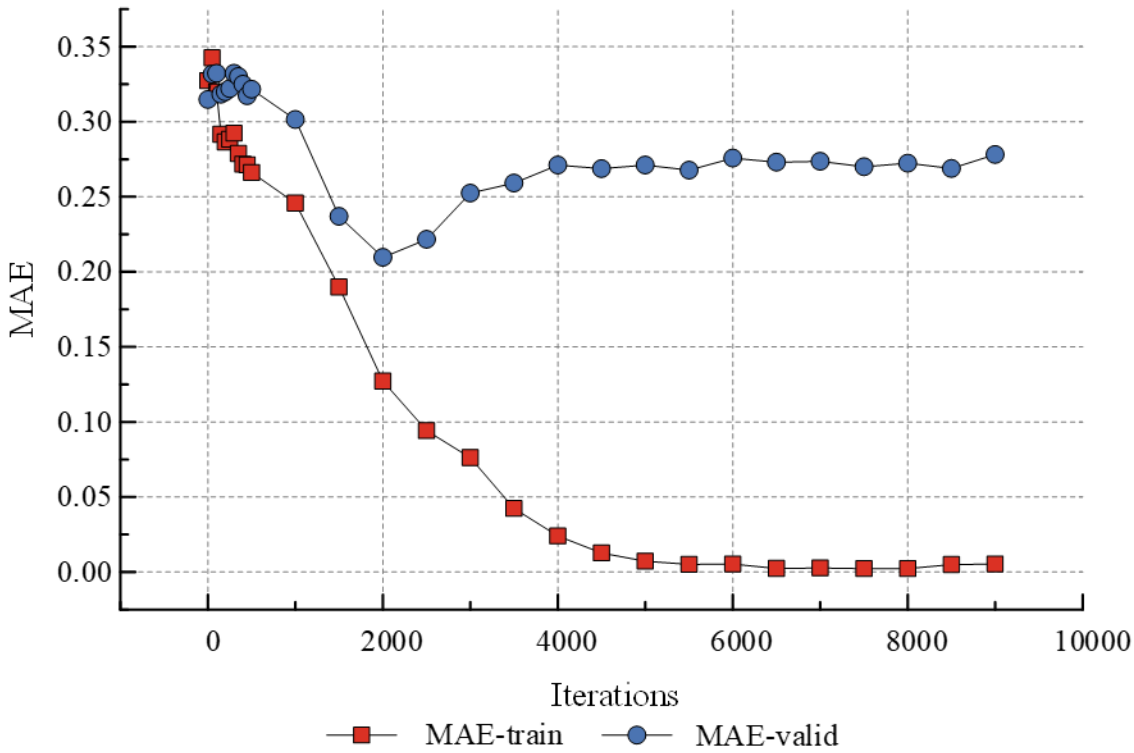

Figure 6. Emotion recognition results of multimedia resources

value of the training set gradually decreases with the increase of the number of iterations, and the error drops to the lowest and stabilizes at 2000-2500 iterations. The error of the test set at 2000 iterations was 0.20960, indicating that the Fruit fly-optimised generalised regression neural network can effectively identify the emotions of multimedia resources and can provide technical support for subsequent research.

3.4. **Learners' facial emotions.** Facial expressions are the main channel through which learners express their emotions and are currently one of the most important ways to determine learned emotions.

To enable fast face detection, this paper uses OpenCV to implement the AdaBoost algorithm. After the face image is detected by the AdaBoost algorithm, an optimised generalised regression neural network is used to perform expression recognition on the face image. The face expression recognition dataset was obtained from an autonomously constructed learner face expression database containing six basic expressions of anger, disgust, fear, happiness, sadness, surprise and neutral expressions. The subjects were all aged between 20-29 years old, with a male to female ratio of 1:3, and all subjects were university and postgraduate students. The learners' facial emotions were divided into seven types for the training and test sets, with the training set accounting for 80% of the total number of images and the test set accounting for 20% of the total number of images. The training dataset contains 40398 images, while the test images contain 10113 images. The image size is all set to 180*180, which can speed up the training while ensuring clarity. The face recognition accuracy is shown in Figure 7. As the number of

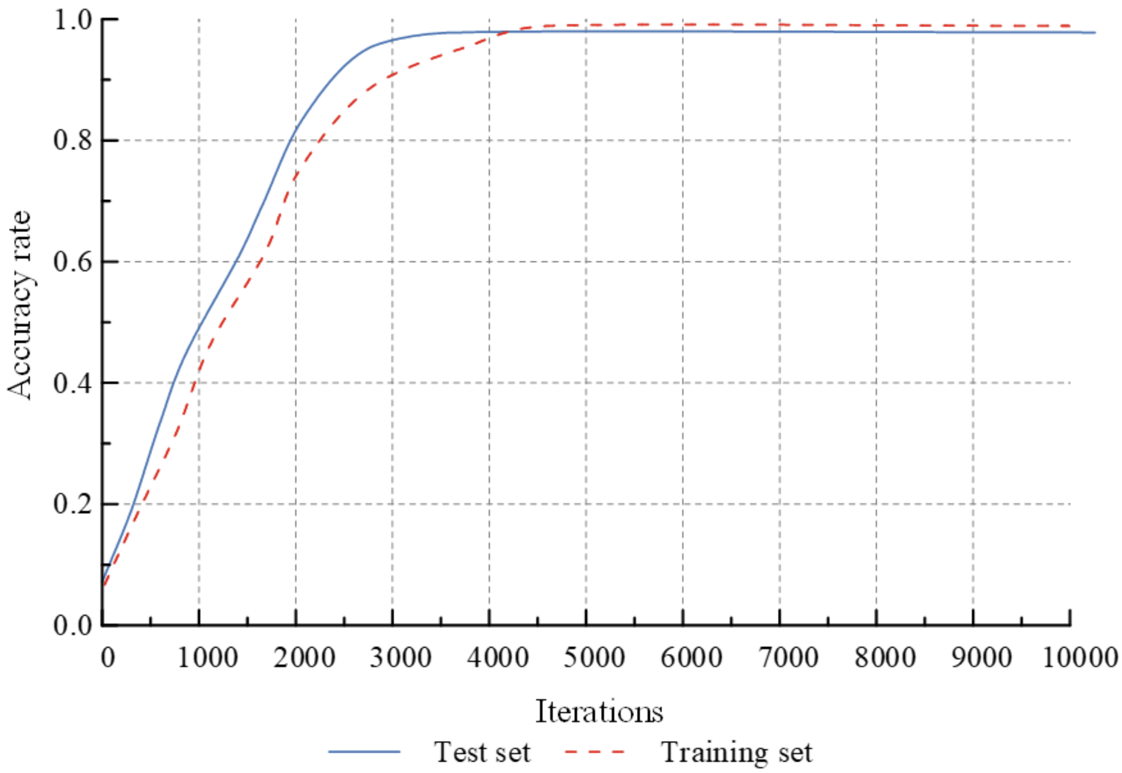

Figure 7. Face recognition accuracy rate

training sessions increases, the training accuracy of the neural network model increases gradually on both the training and test sets. Accuracy increases faster in the early stages and decreases in the later stages. After 5000 iterations of training, the test set accuracy reached 97.78%, while the training set accuracy reached 98.88%. This indicates that the Fruit fly-optimised generalised regression neural network achieves good accuracy on this dataset.

4. Experimental results and analysis.

4.1. Experimental design. The subjects selected for this study were 41 university students with an average age of 20 years. 41 students were divided into 2 classes, Class A with 23 students and Class B with 18 students. The experiment was conducted for a total of 18 weeks from 25 February 2021 to 25 June 2021. Teachers will complete their normal teaching duties during the experiment.

The computer used was equipped with a digital camera as it was necessary to capture images of the subjects' facial expressions during the experiment. The camera resolution was 640*480 and supported manual adjustment from left to right and up and down to ensure that a complete image of the face could be acquired. The computer used in this study was a Window 7 64-bit operating system with 24G of RAM and an Intel Core i7-6700 CPU (up to 3.4G HZ). The deep learning environment was set up using TensorFlow.

In this study, a large amount of data was collected during the experiment, including images of multimedia resources and images of learners' expressions, as shown in Table 2.

Table 2. Statistical results of multimedia resources and learner expression images

Courses	Number of multimedia images	Number of face images	Number of faces detected	Number of experiments	Detection rate
Class A	99309	99325	27059	23	27.24%
Class B	110265	110234	41046	18	37.24%
Total.	209574	209559	68105	41	

4.2. Performance validation of optimized generalized regression neural networks. In order to quantitatively evaluate the performance of the optimised generalised regression neural network, normalised mean square error (NMSE) and peak signal to noise ratio (PSNR) were used.

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N} 255^2}{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N}\left[A\left(i,j\right)-B\left(i,j\right)\right]^2}\right) dB \tag{15}$$

$$\mathrm{NMSE}(A,B) = \frac{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N}\left[A(i,j)-B(i,j)\right]^2}{\sum\limits_{i=1}^{M}\sum\limits_{j=1}^{N}\left[A(i,j)\right]^2} \tag{16}$$

where A is the original input image, B is the filtered image, (i,j) represents the coordinates of the pixels in the image, and M and N represent the length and width of the image, respectively. Larger values of PSNR are preferred, while smaller values of NMSE are preferred.

The experimental results show that the optimised generalised regression neural network shows the best performance in terms of both NMSE and PSNR. Figure 8 shows a comparison of the PSNR results of different models processing the same image when the noise density is increasing. It can be seen that the PSNR values of radial basis neural network,

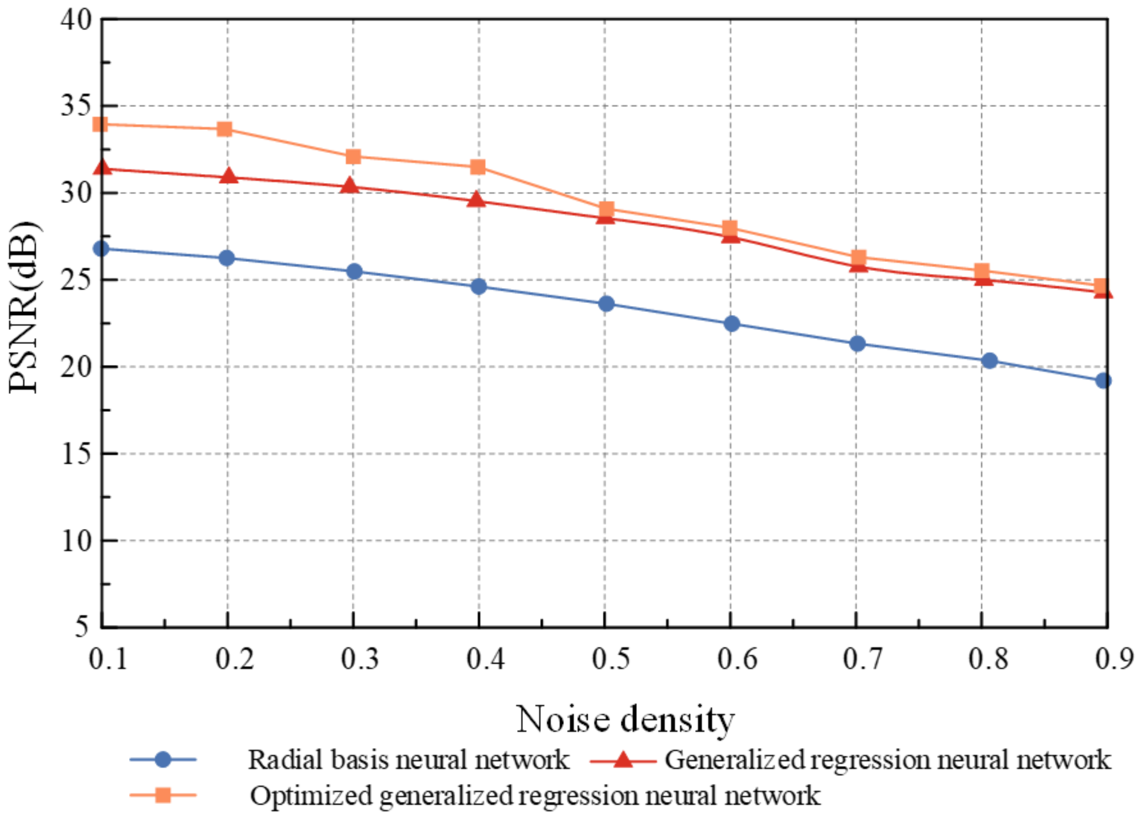

Figure 8. Comparison of PSNR results for different models

conventional generalized regression neural network and optimized generalized regression neural network all decrease with the increase of noise level. For the same noise condition, the optimized generalized regression neural network has the highest PSNR values compared to the other models. A comparison of the NMSE results for the different models is shown in Figure 9, where the larger the PSNR value, the better the NMSE value. It can be seen that the optimised generalised regression neural network has the smallest NMSE values compared to the other models under the same noise conditions. The main steps for optimising a generalised regression neural network are: (1) Optimising the parameter *Spread*, which needs $O(S + C)$; (2) Running the generalised regression neural network, which needs *O(A-N-C)*. As both N and C are smaller than S, the complexity of optimising the generalised regression neural network is approximately $S + A - N - C$, which shows no significant increase in complexity.

4.3. **The impact of emotional interactions.** Analysis. Firstly, the trained optimized generalized regression neural network model is used to identify multimedia resources and learners' expressions synchronously.

Then, the correlation coefficient $r\left(x_i, y_j\right)$ between the two emotion values was calculated.

$$r\left(x_i, y_j\right) = \frac{\sum_{t=1}^{n}\left(X_{it} - \bar{X}_i\right)\left(Y_{jt} - \bar{Y}_j\right)}{\sqrt{\sum_{t=1}^{n}\left(X_{it} - \bar{X}_i\right)^2}\sqrt{\sum_{t=1}^{n}\left(Y_{jt} - \bar{Y}_j\right)^2}} \tag{17}$$

where n is the total number of images, $\bar{X}_i$ is the emotional mean of the multimedia resources, and is the emotional mean of the learners' expressions.

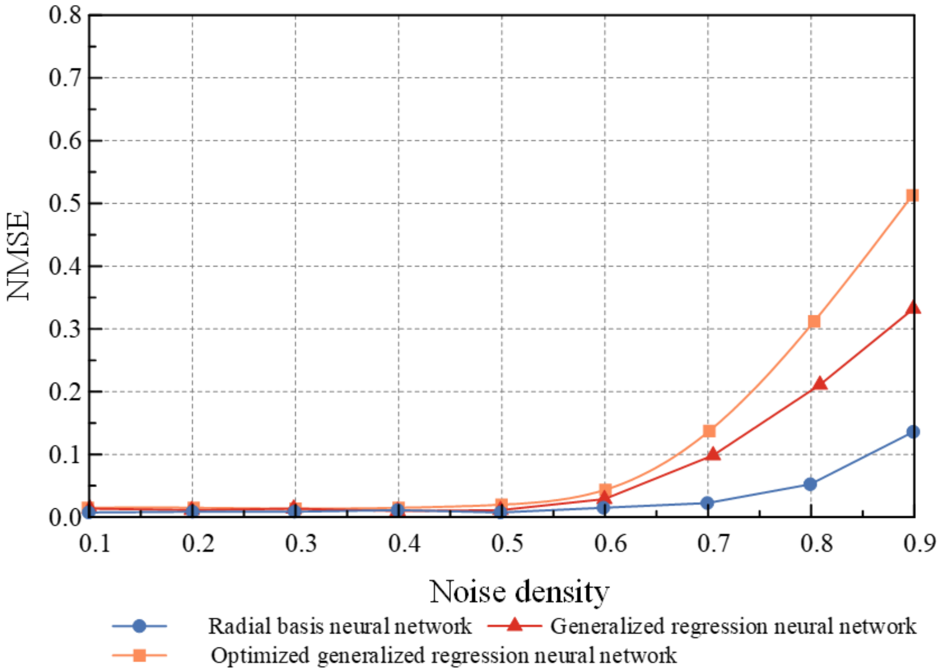

Figure 9. Comparison of NMSE results for different models

The magnitude of the absolute value of the correlation coefficient can, to a certain extent, indicate the previous correlation between the two. In this study, the correlation coefficients were used to explore the relationship between the emotions of 12 multimedia resources and learners' emotions, as shown in Table 3. In general, absolute values of correlation coefficients of 0-0.09 indicate no correlation, 0.1-0.3 weak correlation, 0.3-0.5 moderate correlation, and 0.5 and above strong correlation. The correlation coefficients

Table 3. Correlation coefficient levels

Relevance	**Unrelated**	**Weakly correlated**	**Moderately relevant**	**Strongly related**
Numerical range	$0.0 < \|r\| < 0.09$	$0.0 < \|r\| < 0.3$	$0.3 < \|r\| < 0.5$	$0.5 < \|r\| < 1.0$

derived from this experiment by pairwise calculations are shown in Table 4. The correlation coefficients are all mean values. The learners' facial emotions were divided into seven types and the emotions of the multimedia resources were divided into 12 types. It can be seen that there is a strong correlation between the emotions of some multimedia resources and learners' emotions. Analyzing from the perspective of learning emotions, it can be found that six emotions of multimedia resources: cheerful, lively, funny, humorous, interesting and unreal present a strong correlation with happiness in learning emotions. The six emotions of boring, dull, surprise, fear and thrill show strong correlation with disgust in learning emotion. Dull, depressing, chaotic showed a weak correlation with happiness in the learning emotion. Brighter, more lively and brighter colors are more likely to cause positive emotions in learners. Multimedia resources with more vivid image elements are more likely to elicit positive emotions from learners. Multimedia resources with darker tones and lower color saturation tend to arouse learners' negative emotions.

In most cases, teachers tend to focus on the selection of knowledge content when designing multimedia resources, while neglecting the important role of visual features of the images in regulating learners' emotions. By exploring the impact of multimedia resources

Table 4. Correlation coefficients between emotions of multimedia resources and learners' facial emotions

	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**	**Neutral**
cheerful	-0.062	-0.022	-0.070	0.530	-0.184	-0.391	-0.037
Lively	-0.014	-0.051	-0.057	0.440	-0.168	-0.351	-0.095
Funny	-0.019	-0.037	-0.016	0.438	-0.024	-0.277	-0.091
exaggerated	-0.026	-0.064	-0.078	0.507	-0.035	-0.228	-0.051
Humour	-0.041	-0.060	-0.063	0.469	-0.057	-0.094	-0.116
Interesting	-0.028	-0.086	-0.079	0.546	-0.028	-0.184	-0.029
Dull	0.141	0.514	0.293	-0.020	0.214	0.121	-0.183
Depressing	0.397	0.470	0.397	-0.046	0.384	0.149	0.096
chaotic	0.362	0.551	0.204	-0.015	0.125	0.513	0.105
Unreal	-0.065	-0.057	-0.046	0.373	-0.184	-0.492	0.154
Surprise	0.140	0.410	0.400	-0.015	0.157	0.447	0.063
Fear	0.279	0.432	0.325	-0.027	0.202	0.517	-0.106
Thrill	0.164	0.424	0.284	-0.014	0.501	0.418	-0.105

on learners' emotions, it was found that learners have different emotional cognitive experiences when viewing multimedia resources with different emotions. Therefore, teachers should choose multimedia resources with brighter, more lively and vibrant colours when designing their teaching, which will help stimulate learners' interest and motivation, and promote significant improvement in learners' academic performance.

5. **Conclusion.** This paper explores the emotional characteristics of multimedia resources and their impact on learners' emotions, so as to effectively improve the emotional interaction problem in existing intelligent learning systems. The paper designs an emotion classification coding system and coding strategy for the subsequent emotion recognition of multimedia resources using generalised neural networks. In order to improve the accuracy of emotion recognition by removing noise, the smoothing factor in the generalised regression neural network is optimised using a fruit fly optimisation method. The optimised generalised regression neural network is used to identify the emotion of the multimedia resources and the learner's facial emotion simultaneously, and the correlation coefficient is used to calculate the degree of influence of the emotion on the learner's emotion. The experimental results show that the optimised generalised regression neural network can accurately identify the emotions of the multimedia resources and the learners' facial emotions. The multimedia resources with cheerful, lively, funny, humorous, interesting and unreal emotion types can evoke positive emotion in learners to a certain extent. Multimedia resources with boring, dull and confusing emotions can elicit negative emotions from learners.

REFERENCES

[1] T.-Y. Wu, A. Shao, and J.-S. Pan, "CTOA: Toward a Chaotic-Based Tumbleweed Optimization Algorithm," *Mathematics*, vol. 11, no. 10, 2339, 2023.

[2] T.-Y. Wu, H. Li, and S.-C. Chu, "CPPE: An Improved Phasmatodea Population Evolution Algorithm with Chaotic Maps," *Mathematics*, vol. 11, no. 9, 1977, 2023.

[3] T.-Y. Wu, F. Kong, Q. Meng, S. Kumari, and C.-M. Chen, "Rotating behind security: an enhanced authentication protocol for IoT-enabled devices in distributed cloud computing architecture," *EURASIP Journal on Wireless Communications and Networking*, vol. 2023, no. 1, 36, 2023.

[4] T.-Y. Wu, L. Wang, X. Guo, Y.-C. Chen, and S.-C. Chu, "SAKAP: SGX-Based Authentication Key Agreement Protocol in IoT-Enabled Cloud Computing," *Sustainability*, vol. 14, no. 17, pp. 11054, 2022/09/05, 2022.

[5] H. C. Davies, R. Eynon, and C. Salveson, "The Mobilisation of AI in Education: A Bourdieusean Field Analysis," *Sociology*, vol. 55, no. 3, pp. 539–560, 2020.

[6] D. McCarthy, "Adding Social Emotional Awareness to Teacher Education Field Experiences," *The Teacher Educator*, vol. 56, no. 3, pp. 287–304, 2021.

[7] C.-M. Chen, S. Lv, J. Ning, and J. M.-T. Wu, "A Genetic Algorithm for the Waitable Time-Varying Multi-Depot Green Vehicle Routing Problem," *Symmetry*, vol. 15, no. 1, 124, 2023.

[8] A. L. H. P. Shaik, M. K. Manoharan, A. K. Pani, R. R. Avala, and C.-M. Chen, "Gaussian Mutation–Spider Monkey Optimization (GM-SMO) Model for Remote Sensing Scene Classification," *Remote Sensing*, vol. 14, no. 24, 6279, 2022.

[9] J. M.-T. Wu, Q. Teng, S. Huda, Y.-C. Chen, and C.-M. Chen, "A Privacy Frequent Itemsets Mining Framework for Collaboration in IoT Using Federated Learning," *ACM Transactions on Sensor Networks*, vol. 19, no. 2, pp. 1-15, 2023.

[10] B. D. Boulay, "Escape from the Skinner Box: The case for contemporary intelligent learning environments," *British Journal of Educational Technology*, vol. 50, no. 6, pp. 2902–2919, 2019.

[11] S. A. Abdulkareem, Y. T. Mustafa, E.-W. Augustijn, and T. Filatova, "Bayesian networks for spatial learning: a workflow on using limited survey data for intelligent learning in spatial agent-based models," *GeoInformatica*, vol. 23, no. 2, pp. 243–268, 2019.

[12] H. Li, H. Li, S. Zhang, Z. Zhong, and J. Cheng, "Intelligent learning system based on personalized recommendation technology," *Neural Computing and Applications*, vol. 31, no. 9, pp. 4455–4462, 2018.

[13] M. A. Mohammed, D. A. Ibrahim, and A. O. Salman, "Adaptive intelligent learning approach based on visual anti-spam email model for multi-natural language," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 774–792, 2021.

[14] F. Jiao, J. Song, X. Zhao, P. Zhao, and R. Wang, "A Spoken English Teaching System Based on Speech Recognition and Machine Learning," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 14, pp. 68–82, 2021.

[15] A. Saleem, M. K. Afzal, M. Ateeq, S. W. Kim, and Y. B. Zikria, "Intelligent learning automata-based objective function in RPL for IoT," *Sustainable Cities and Society*, vol. 59, 102234, 2020.

[16] Q. Zhu and M. Wang, "Team-based mobile learning supported by an intelligent system: case study of STEM students," *Interactive Learning Environments*, vol. 28, pp. 543–559, 2019.

[17] J. C.-Y. Sun, S.-J. Yu, and C.-H. Chao, "Effects of intelligent feedback on online learners' engagement and cognitive load: the case of research ethics education," *Educational Psychology*, vol. 39, no. 10, pp. 1293–1310, Nov. 2018.

[18] L. Rong, "Design of Ideological and Political Multimedia Network Teaching Resources Integration System Based on Wireless Network," *Scientific Programming*, vol. 2021, 4293771, 2021.

[19] A. Y. Rahmat, "The Effect of Using Educational Multimedia in Dictation on Students' Listening Comprehension at MA Darul Hikmah Pekanbaru," *Journal of English for Academic*, vol. 5, no. 1, pp. 33–42, 2018.

[20] Y. Zhao, "A Personalized English Teaching Design Based on Multimedia Computer Technology," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 8, pp. 210–222, 2020.

[21] F. Weng, R.-J. Yang, H.-J. Ho, and H.-M. Su, "A TAM-Based Study of the Attitude towards Use Intention of Multimedia among School Teachers," *Applied System Innovation*, vol. 1, no. 3, 36, 2018.

[22] P. M. Wright and K. A. R. Richards, "Social and Emotional Learning as an Integrated Part of Physical Education," *Journal of Physical Education, Recreation & Dance*, vol. 93, no. 2, pp. 5–7, 2022.

[23] J. Alam, M. Hassan, A. Khan, and A. Chaudhry, "Robust fuzzy RBF network based image segmentation and intelligent decision making system for carotid artery ultrasound images," *Neurocomputing*, vol. 151, pp. 745–755, 2015.

[24] X. Luo, W. Zhao, S. Wei, and Q. Fu, "Study on Urban Remote Sensing Classification Based on Improved RBF Network and Normalized Difference Indexes," *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 8, no. 10, pp. 257–270, Oct. 2015.

[25] V. Colonnello, P. Iacobucci, T. Fuchs, R. C. Newberry, and J. Panksepp, "Octodon degus. A useful animal model for social-affective neuroscience research: Basic description of separation distress, social attachments and play," *Neuroscience & Biobehavioral Reviews*, vol. 35, no. 9, pp. 1854–1863, 2011.

[26] G. Makransky and G. B. Petersen, "The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality," *Educational Psychology Review*, vol. 33, no. 3, pp. 937–958, 2021.

[27] D. R. Lara, L. W. Bisol, M. G. Brunstein, and C. T. Reppold, "The Affective and Emotional Composite Temperament (AFECT) model and scale: A system-based integrative approach," *Journal of Affective Disorders*, vol. 140, no. 1, pp. 14–37, 2012.

[28] Y. Wang and H. Peng, "Underwater acoustic source localization using generalized regression neural network," *The Journal of the Acoustical Society of America*, vol. 143, no. 4, pp. 2321–2331, 2018.

[29] Y. Fan et al., "Rationalized fruit fly optimization with sine cosine algorithm: A comprehensive analysis," *Expert Systems with Applications*, vol. 157, 113486, 2020.

[30] L. Peng, Q. Zhu, S.-X. Lv, and L. Wang, "Effective long short-term memory with fruit fly optimization algorithm for time series forecasting," *Soft Computing*, vol. 24, no. 19, pp. 15059–15079, 2020.