Hindawi
Advances in Materials Science and Engineering
Volume 2024, Article ID 9849245, 1 page
https://doi.org/10.1155/2024/9849245

Retraction

Retracted: Construction of Facial Expression Recognition Technology of Intelligent Service Robot Based on Nanofiber Flexible Wearable Electronic Skins

Advances in Materials Science and Engineering

Received 8 January 2024; Accepted 8 January 2024; Published 9 January 2024

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] F. Li, C. Li, and Y. Liu, "Construction of Facial Expression Recognition Technology of Intelligent Service Robot Based on Nanofiber Flexible Wearable Electronic Skins," *Advances in Materials Science and Engineering*, vol. 2022, Article ID 8669298, 10 pages, 2022.

Hindawi
Advances in Materials Science and Engineering
Volume 2022, Article ID 8669298, 10 pages
https://doi.org/10.1155/2022/8669298

Research Article

Construction of Facial Expression Recognition Technology of Intelligent Service Robot Based on Nanofiber Flexible Wearable Electronic Skins

Feigao Li, Chenyang Li, and Yue Liu

Henan Polytechnic, Zhengzhou 450046, Henan, China

Correspondence should be addressed to Feigao Li; lifeigao123@sina.com

Received 8 July 2022; Revised 10 August 2022; Accepted 8 September 2022; Published 30 September 2022

Academic Editor: K. Raja

Recognizing human features is a highly significant and popular research topic in academia today, especially with regard to human intervention and computer intelligence. If you want your computer to have real intelligence, then the computer must first learn to understand human feelings, and facial expressions often contain rich emotional information. Clearly, recognizing human features is indispensable for the future intelligent development of computers. There are always a few components and many researchers are involved. Compared with the traditional facial expression recognition method, which needs to manually define features and artificially perform appropriate feature extraction, a great advantage of deep learning is to remove the disadvantages that require a lot of human intervention, thereby reducing the uncertainty of expression recognition. In recent years, deep processing has led to outstanding gains in numerous domains, including robotics, computer graphics, and natural language processing. Therefore, this article applies a popular recognition algorithm based on deep learning, combined with a nanofiber flexible wearable electronic skin, and applies it to the field of facial expression recognition. Functional expression identification is an essential goal of the external analysis of human features. Mature facial expression recognition technology has great application values.

1. Introduction

Smart service robots based on nanofiber flexible wearable electronic skins are various high-tech integrated smart devices that can provide humans with necessary services in an unstructured environment mainly focusing on the application requirements of service robots and dangerous work robots. Intelligent service robot technology based on nanofiber flexible wearable electronic skins involves many fields, such as machinery, electronics, materials, computers, sensors, control, and medicine. It requires brainstorming and multifaceted collaborative research [1]. The robot must be able to understand and respond to people's needs by listening, speaking, observing people's movements, expressions, etc. In an unstructured environment, robots must quickly collect and quickly process the necessary information for all environments, including a wide variety of static and dynamic objects, including humans and animals. Advanced hybrid drive technology, flexible robot technology, tactile technology, visual technology, and artificial intelligence technology make service robots, especially humanoid service robots, smarter, and more friendly. Not only can they walk, run, communicate and express emotions but they can also learn and grow their intelligence like children. The key technologies that China is currently mastering are, namely, environmentally aware sensors and signal processing methods. The service robot must observe and identify the surrounding environment while working. The more sensors, the more information the surrounding environment acquires and the more complementary and redundant the various sensor information is in space and time, the more they are combined by some kind of operation and optimization. To describe and accurately identify the environment, intelligent control is the key technology of the robot that mainly includes the following: neural network, paste control, and optimization calculation. Fuzzy control is

a control method that uses the basic theory of fuzzy mathematics. It generally has the following 5 principal sections: defining parameters, ambiguities, a repository of information, illogical determination, and defuzzification. Artificial neural network control is inferred from the activities of humans and creatures [2]. The characteristics of neural networks are information processing and reasoning, association, and thinking. Navigation and positioning: The multisensor fusion technology and artificial neural network control in the service robot system can realize the autonomous positioning and navigation of the service robot. Route plan: Path planning refers to finding a route between the initial situation and the desired goal situation within the service motor's construction space in order to escape from barriers. The artificial intelligence approach is a developed and productive programming approach to conventional arithmetic. It allows route programming processes using AI tools but without thinking about whether the route is best or not. Smart route programming approaches can provide greater barrier avoidance performance, accelerate programming faster, and enable real-world solutions. Human-robot interaction technology. Since the aim of smart computing research is to enable robots to work smartness and serve humans as if they had a "brain," there must be human instruction and domination, detection, and maintenance of the robot, and feedback from the robot on the outcomes of its operations. It mainly includes technologies that include texture awareness, synthesizing and recording spoken language, identifying and managing pictures, and performing computer interpretation [3]. At present, the research on the more popular intelligent service robots in foreign countries is mainly concentrated in Germany, Japan, and the United States and has already been adopted by a variety of companies. Due to the economic prosperity, the increasing demand for life in China and the increase in the proportion of elderly people caused by family planning policies, and the problem of elderly care, many robotic R&D companies have begun to shift their research focus to the development of intelligent service robots and have made relatively large achievements in this area.

Following the rapid growth of synthetic AI and computer vision, there is a growing demand for intelligent emotional feedback and anthropomorphic person-machine engagement. As one of the key components of AI human-computer communication technology, the recognition of functional facial expressions has attracted a lot of focus in recent years. Human expressions can be divided into macro expressions and microexpressions. The study of macroscopic expressions has been more than 20 years old and has achieved certain results. Microexpressions are spontaneous and unconscious expressions in a very short time. Changes, which occur for a short time, are difficult to capture by humans. Therefore, high-speed cameras need to be introduced to capture and analyze them one frame at a time. However, this work is still not something that ordinary people can do and requires some experience. The psychology expert can complete it. What we have to do is to help humans recognize expressions through deep learning. Facial expression recognition refers to the analysis of human facial expressions and changes through computer pattern recognition and machine learning algorithms and then judges people's inner emotions and thought activities, thereby obtaining a more intelligent human-computer interaction environment. The machine learning method based on traditional manual custom feature classification has gradually become the main method for the automatic recognition of facial expressions. A series of research results have been obtained, but functional features are influenced by numerous parameters, such as the influence of the natural environment, such as illumination and angle. as Along with the influence of individual differences in ethnicity, gender, age, etc., traditional methods have many difficulties in manually customizing features [4]. The concepts and advancements of profound training have provided innovative ideas and approaches for researchers working in the visual identity field. It replaces artificial signature collection in the form of effective techniques for unmonitored or quasi-supervised signature acquisition and stratified signature capture. Fruitful results. Thanks to the use of the graphic processing unit (GPU) and the improvement of computing performance, the convolutional neural network (CNN) has achieved a series of eye-catching performances in the field of image recognition, including AlexNet and VGGNet. The network structure such as ResNet has won many times in the ILSVRC competition, and its results have also been applied to various image recognition tasks. Currently, profound training, specifically convergent cognitive approaches to facial expression recognition, has become the mainstream approach to automatic facial expression recognition. However, there are also difficult problems in the research of facial expression recognition, which are summarized as follows: The human face is a flexible body in a three-dimensional space. It is easy and accurate to build a rigid body model, but the human face is a flexible body in a three-dimensional space, and it is difficult to establish a model. The existence of individual differences affects expression recognition. The experimental results are different for the same expression, due to the influence of individual differences. If the individual difference is greater, the influence on the expression recognition will be greater. When feature extraction, the interference of individual differences cannot be completely avoided, which undoubtedly increases the difficulty of extracting expression features of the same category. The expression itself is different. The expressions of Europeans and Americans are more exaggerated, the emotions are more obvious, while the Asians are more restrained and the expressions are smaller; women and children express more expressive expressions than men; the expressions of old and young people are different. The folds also affect the extraction of expression features and expression recognition in the natural environment. At present, most expression recognition is performed on the published database, and the images in these databases are expressions of the unnatural environment made by volunteers under the guidance of professionals. Most of them are exaggerated and not true to volunteers. The expression of emotions has certain performance components, which are not conducive to the application of actual scenes. Interference from external factors.

The existence of disturbance factors such as illumination, attitude change, and noise in the actual environment not only increases the difficulty of obtaining effective expression information but also is not conducive to popularizing the expression recognition system in practical applications.

About the expression recognition since the beginning of the 21th century, based on a simple knowledge distillation scheme, Gera D uses a single network to complete label inference on large-scale facial expression datasets, and the proposed framework demonstrates its effectiveness and comprehensiveness in the noisy FER dataset [5]. Li et al. fused energy feature vectors and facial feature vectors in their research study, used support vector machine (SVM) to classify the fused feature vector, and finally proved that the proposed method has higher accuracy and stronger generalization transformation ability [6]. Hossain and Yousuf proposed a facial recognition system that uses the currently developed OpenCV Haar Feature-based cascade classifier for face detection, capable of recognizing natural head nods and shakes [7]. Jan and Sajjad proposed a new descriptor called the Weber discrete wavelet transform, which can effectively identify facial expressions, and finally, they observed through experiments that the feature dimension was significantly reduced, while the recognition accuracy was significantly improved [8]. Wang et al. proposed a multimanifold discriminant analysis method for image feature extraction and finally verified the effectiveness of this analysis method for image feature extraction through experiments [9]. Rusia and Singh proposed two customized CNN models, namely, Proposed Model 1 and Proposed Model 2, to classify generic facial expressions to prevent overfitting [10]. These research methods have carried out a profound research on face recognition to a certain extent.

The emergence of deep learning in recent years has brought new vitality to the field of computer vision [11]. In fact, it has made the research of behavior recognition, face recognition, and semantic segmentation into a new stage and achieved very significant results. The artificial neural network is a very complex nonlinear network [12]. It is composed of a large number of processing units and is an abstraction and modeling of the human brain or biological neural network. Since the artificial neural network was proposed, after decades of development, researchers still cannot interpret or completely imitate the working mechanism of the brain. For a while, the research study only stays in the shallow neural network. In recent years, the development of deep neural networks has brought about another phase of neural network development. Deep learning is also a kind of neural network. Deep learning refers to the way of using artificial neural networks to learn features by relying on powerful graphics computing power and the development of big data at present. The basic idea is that through convolutional neural networks (CNN) and other methods to carry out complex calculations of a large number of training datasets and update network parameters through backpropagation, finally enabling the trained model to better reflect the data characteristics [13]. In order to solve the problem of expression recognition in traditional methods, deep learning seems to be a good tool. However, deep learning needs to be trained with a higher number of databases in order to obtain representative features, otherwise, it will cause over-fitting, etc. Problems. Following the triadic categorization scheme in profound understanding techniques and operations, the profound understanding structures are broadly classified into the following three types: (1) Asocial or spontaneous deep systems: The profound Internet produces symptoms validly through Internet coding without employing object tags, so it is also known as a generative method. For example, the defined branch neural network (DBN), revolving network nature (RNN), and the like. (2) Deep forests with monitored teaching: Significantly, superintended training means that profound systems use labeling of object types, either explicitly or obliquely, to create distinctions for the purpose of modality selection, so they are also known as distinguishing depth networks, such as curvilinear neural nets (CNNs) and reciprocal neural nets (RNNs). (3) Hybrid deep forests: "Hybrid" refers to a conjunction of unsupervised and monitored training, such as the DBN-CNN hybrid model for pretraining deep convolutional neural networks using a generative DBN.

According to the aforesaid diagnosis, this article suggests that the facial expression recognition operation based on profound acquisition improves the accuracy of expression recognition, the average accuracy rate is improved by 2%, and it is faster than the popular expression recognition algorithm. Moreover, through the combination of deep face-based face verification algorithms, different expressions of the same person can be obtained, which effectively helps to further analyze facial expressions.

2. Proposed Method

Expression recognition can be deeply studied through deep learning [14]. The profound acquisition is a division of methodological training. The purpose is to simulate the structure of the brain's neural network by building a model. When processing data (images, voices, texts, etc.), it is a multilevel structure to learn features. Deep learning can be divided into deep concessional bridging systems, profound conviction systems, profound Boltzmann systems, and compounded autoencoders according to the model structure. The deep involutional nerve structure is a type of recurrent nerve structure designed to process multidimensional data [15].

2.1. The Overall Framework of Expression Recognition Based on Deep Convolutional Neural Networks. In this section, a dynamic expression identification system using profound involutional integer systems is designed [16]. The system consists of a computer camera and software function modules. The video image information is collected in real time by the camera in a real environment, and each frame of the video is separately calculated. In the software part, first, the Haar classifier is used for face detection. Second, the obtained face image is preprocessed to obtain the face gray image; the obtained gray image is input to the trained deep convolutional neural network again. Finally, the resulting

expression features are entered into the SoftMax classifier for expression classification. Real-time dynamic facial expression recognition [17].

The block diagram of the expression recognition system based on deep convolutional neural network is shown in Figure 1.

2.2. Face Detection Using the Haar Classifier. This section uses the face detection method based on the Haar-like feature and the AdaBoost algorithm, namely, the Haar classifier algorithm [18]. This algorithm uses the Viola–Jones target detection framework, which is a real-time object detection method, which can realize the first frame from video. It positions the face to achieve real-time expression recognition. The integral graph algorithm can be used to quickly calculate Haar-like features with different sizes and positions. The integral image value at a pixel (x, y) in the image is defined in references [19, 20]:

$$\gamma\gamma(x, y) = \sum_{x' \le x, y' \le y} \gamma(x', y'), \tag{1}$$

where $\gamma\gamma(x, y)$ is the integral image and $\gamma(x, y)$ is the original image. We pass the following iteration formula [21, 22]:

$$\begin{aligned} s(x, y) &= s(x, y-1) + \gamma(x, y), \\ \gamma\gamma(x, y) &= \gamma\gamma(x-1, y) + s(x, y). \end{aligned} \tag{2}$$

Once the image is scanned, when the pixel in the lower right corner of the image is reached, the constructed image is shown in Figure 2:

Wherein, in the integral image, point a represents the sum of the pixel values of area A, point b represents the sum of the pixel values of the two areas A and B, point c represents the sum of the pixel values of the two areas A and C, and point d represents the sum of the pixel values of the two areas A and C. The sum of the pixel values of the four regions A, B, C, and D. Therefore, any rectangle can be derived from the integral of the 8 rectangular corner points [23, 24].

The AdaBoost algorithm trains the identical narrow classifier on separate practice episodes. The partial categories acquired from separate practice episodes are then combined to form a definitive and powerful categorizer. The AdaBoost algorithm is used because the Haar-like features have high dimensionality and it is necessary to choose them with a particular method and incorporate them into a powerful categorizer to examine a human face [25].

$$h(x) = \begin{cases} 0, & other, \\ 1, & \sum_{t=1}^{T} a_t h_t(x) \ge \frac{1}{2} \sum_{t=1}^{T} a_t. \end{cases} \tag{3}$$

2.3. Emoticon Image Preprocessing. Geometric normalization. The MTCNN algorithm is used to detect the position of the face and extract the key points of the face. The center point of each eye is calculated according to the above-mentioned face key points and connection, and the angle between the horizontal line is calculated. According to this angle, we affine the picture. Thereby, the horizontally aligned picture can be obtained. At this time, the face is cut according to a certain scale factor, and the degree change and the face rotation can be processed by the abovementioned method. Grayscale normalization. The picture pixels are normalized to [0, 1], mainly to reduce image quality problems caused by illumination changes.

2.4. Training Deep Convolutional Neural Network. The normal decomposition paradigm employed in building deep decomposed recurrent nets commonly consists of three primary levels, including a decomposition level, a pooling level, and a wholly concatenated level. Different layers have different functions. A neural network is capable of image recognition, its input has a lot of pixels. The first layer of computing units can only recognize a few shapes, such as the difference between light and dark, and the second layer of neurons is transmitted to the first layer of neurons As a result, the analysis obtains information such as edges and angles, and continues abstracting layer by layer and eventually obtaining high-level features that are closer to the nature of the research object [26].

Given the input, the calculation process for the entire network is as follows:

For convolutional layer input X, the convolution process is as follows [27, 28]:

$$\text{con} = f\left(\sum_{i,j \in M} x_{i,j} w_{m-i,n-j} + b \right), \tag{4}$$

where x represents the element in the convolutional region of input X, w represents the element in the convolution kernel, m, n represents the size of the convolution kernel, b represents the offset, and f represents the ReLu activation function.

For the pooling layer input y, the pooling process is [9]

$$\text{pool} = \text{down}\left(\max\left(y_{i,j}\right)\right), i, j \in p, \tag{5}$$

where y represents the pooling layer input element, and $down$ is the downsampling process, which retains the maximum value in the pooled area.

For the fully connected layer input G, the formula is as follows [29]:

$$\text{full} = f(w * z + b), \tag{6}$$

where z represents the input element, w represents the weight, b represents the offset, and f represents the ReLu activation function.

For the input layer input, the formula is as follows [30]:

$$p(y = j \mid x; w) = \frac{e^{w_l x}}{\sum_{l=1}^{K} e^{w_l x}}. \tag{7}$$

2.5. Expression Classification. This chapter uses SoftMax as the expression recognition classifier. We set the training set to $\{(x^{(1)}, y^{(1)}), \cdots, (x^{(m)}, y^{(m)})\}$ with input features $x^i \in R^{n+1}$, and the class is labeled $y^{(i)} \in \{1, 2, k, \cdots 7\}$:

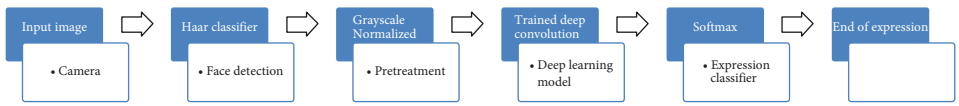

Figure 1: Block diagram of the expression recognition system based on the deep convolutional neural network.

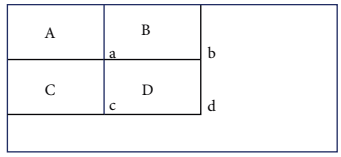

Figure 2: Integral image calculation demonstration.

$$h_\theta(x) = \begin{bmatrix} p\left(y^{(i)} = 1 \mid x^{(i)}; \theta\right) \\ p\left(y^{(i)} = 2 \mid x^{(i)}; \theta\right) \\ \vdots \\ p\left(y^{(i)} = k \mid x^{(i)}; \theta\right) \end{bmatrix} = \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix}, \tag{8}$$

where $p(y^{(i)} = j \mid x^{(i)})$ represents the probability $\theta_1, \theta_2, \cdots \theta_k \in R^{n+1}$ of the input $x^{(i)}$ of the ith sample belonging to the category j. The system cost function is as follows [31, 32]:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\left\{y^{(i)} = j\right\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^{k} e^{\theta_j^T x^{(i)}}}\right]. \tag{9}$$

Cost function gradient $\nabla_{\theta_j} J(\theta)$:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[x^{(i)}\left(1\left\{y^{(i)} = j\right\} - p\left(y^{(i)} = j \mid x^{(i)}; \theta\right)\right)\right]. \tag{10}$$

3. Experiments

In this paper, two expression image libraries are used in the experiment. One is the standard image library JAFFE, and the other is to collect the image library dataset, which is processed on the Internet. JAFFE is a database of expressions established in Japan. It is used by a large number of facial expression recognition researchers and has a certain persuasive power. The JAFFE facial expression image library is compiled by ten Japanese women as models. Each expression selects three to four images according to the size of the amplitude. Each person makes seven expressions to collect, and together, there are 213 images, depending on the resolution of the image. It is 256×256. Due to the needs of the experiment, the original pixels in the expression library will be sampled, and image libraries of four different pixels of 16×16, 32×32, 64×64, and 128×128 will be created. The comparison of the samples in the JAFFE image library at different resolutions shows that lower the resolution, the more blurred the information carried by the image.

The second database, which collected about 13 people, each had 75 emoticons, for a total of 965 images. The image resolution is 64×64. Similarly, this paper has established 16×16 and 32×32 expression libraries, respectively. Due to the mixed expressions in this expression database, a relatively obvious and representative expression image was selected, and a total of six expressions were selected, which were happy, surprised, sad, neutral, angry, and disgusted.

The experimental code implementation is in the Python language, the IDE is PyCharm graphical interface tool, the running environment is window7, memory 4G, i5 processor. Since the experiment involved image processing, a large number of matrix operations were required, so the program was selected to run on the GPU. This not only improves the efficiency of the experiment but also shortens the experiment time. Since there are many unpredictable situations during the experiment, in order to ensure the validity of the experimental results, five cross-validation methods were carried out in the experiment of the paper. Here, the image samples of the established facial expression image library are equally divided into 5 for each experiment, 4 sets of images were used for network training, and the rest were used for experimental tests. The process of training to identify such an experiment is repeated five times; then, the average of the five experimental results is the experimental result of the paper.

4. Discussion

In this paper, the face image is grayscaled and grayscale normalized before image processing. Figure 3 shows the image processing results in this paper, where Figure 3(a) is the color image. Figure 3(b) is the image after the image is grayscaled, and Figure 3(c) is the image after the image grayscale is normalized. It can be seen from the figure that after the image gray processing, there is no change in the overall structure of the image, and it is advantageous to extract the object features for analysis after preprocessing the image.

After preprocessing the image, this section of the experiment constructs the CNN structure shown in Table 1 for the original expression set of 32×32 and 64×64 resolutions, respectively. The learning rate setting includes setting the fixed learning rate, the unified local optimal learning rate, and layering. The dynamic local optimal learning rate is trained in three scenarios, and the network convergence comparison is shown in Figures 4 and 5.

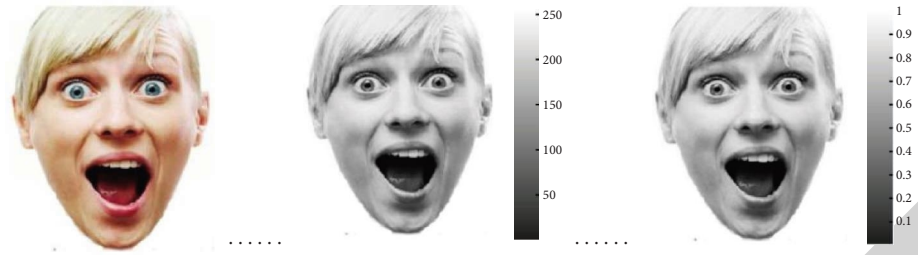

Figure 3: Image preprocessing. (a) Color image. (b) Grayscale image. (c) Normalized image.

From the experimental results of Figures 4 and 5, it can be known that this method achieves similar improvement effects in the training of image samples of 32×32 and 64×64 images with different resolutions. In general, the convergence speed of uniformly setting the local optimal learning rate is higher than the convergence speed of setting the fixed learning rate, and the convergence speed of the dynamic smooth local optimal learning rate adjustment is significantly faster than the convergence speed of uniformly setting the local optimal learning rate. The influence of various setting methods on the convergence speed is mainly reflected in the early stage. This is because the pretraining period is far from the extreme point, which makes the adjustment of the optimal learning rate show obvious advantages. The learning rate optimization accelerates the speed of network convergence and verifies that the proposed method is feasible for the shallow layer enhancement scheme, and the expression classification accuracy based on the method is higher. This dynamic learning rate adjustment method is used to jump out of the local extremum. Points have a beneficial effect.

The learning rate optimization method in this paper brings a certain performance improvement, but it also has defects: the convolutional neural network contains a large number of parameters and calculating the optimal learning rate will incur a certain time cost. If it is optimized for each iteration, it will be optimized once. However, in the case of large data volumes or complex networks, the time loss caused by this will offset the advantage of its convergence speed. In the actual training, the learning rate is updated with the inertia of the learning rate. The shorter the interval between the learning rate adjustment and the longer the convergence step of the network is, the longer the learning rate adjustment interval is, the lower the iterative average time cost is. Choosing the appropriate learning rate adjustment interval requires finding a balance of time performance.

For the purpose of ensuring the credibility of the experimental outcomes and preventing the influence of chance agents, the data in the dataset were separated into components and the cross-validation method was adopted. In this experiment, the images in the JAFFE photoset are separated into 3 sections, 2 of which are used as exercise figures as well as the other section as test figures; the images in the dataset are separated into 5 sections, 4 of which will be used as exercise figures each time. One served as the experimental figure. Table 2 demonstrates the classification

Table 1: CNN structure description.

Resolution	Network structure	Convolution kernel	Activation function
32×32	$C^6S^6C^{13}S^{13}$	5, 5	RReLU + sigmoid
64×64	$C^6S^6C^9S^9C^{12}S^{12}$	9, 5, 5	RReLU + sigmoid

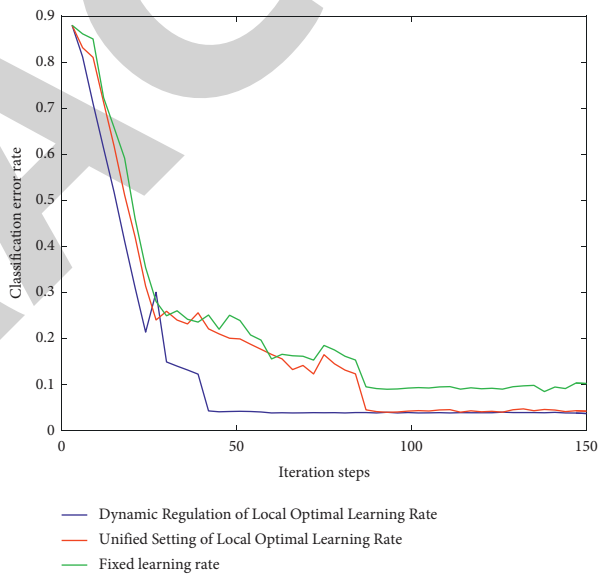

Figure 4: Convergence at 32×32 using different learning rate schemes.

performance of the algorithm for different expressions in the JAFFE representation database. Table 3 demonstrates the categorization outcomes for various representations in the dataset drainage system database. It is obvious from Tables 2 and 3. In some test sets, the whole network performs better and the correct rate is high, but it is relatively poor in others. The explanation for this is probably the expressive traits that can be derived from the exercise model material. Inadequate access to sufficient training features results in a poor classification of the classifier, a situation that mirrors the principle of the significance of example material for involutional systems.

First, the original expression set is balanced then, different degrees of amplification are obtained to obtain a plurality of amplified expression sets, and then the

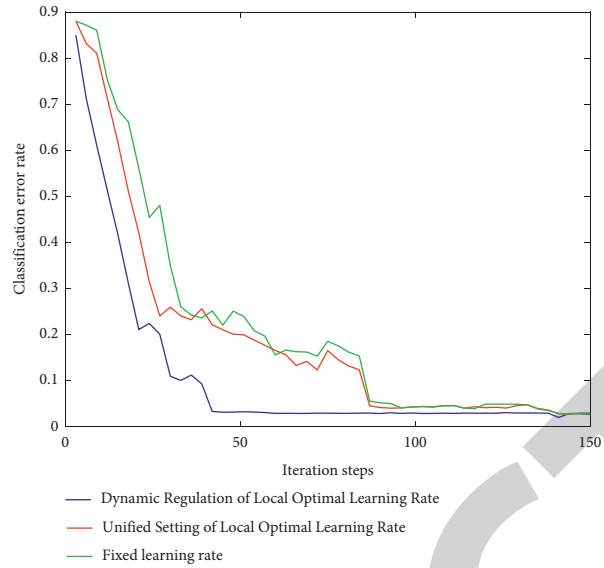

Figure 5: Convergence on different 64×64 using different learning rate schemes.

Table 2: Classification accuracy of different expressions of JAFFE expression library (%).

	Angry	Disgust	Scared	Happy	Neutral	Sad	Surprised	Overall (%)
Test set 1	100	80	100	100	100	90.91	88.89	94.37
Test set 2	100	90	90	81.82	100	100	100	92.96
Test set 3	100	100	81.82	90.91	100	100	100	95.77
Overall (%)	100	89.66	90.63	90.63	100	96.77	96.55	94.37

Table 3: Classification accuracy of different expressions of dataset database (%).

	Angry	Disgust	Scared	Happy	Neutral	Sad	Surprised	Overall (%)
Test set 1	88.89	94.44	80	92.86	70.83	96	93.94	88.89
Test set 2	70.37	77.78	80	96.30	68	84	96.97	82.32
Test set 3	77.78	85.71	84.62	100	64	72	93.94	83.33
Test set 4	62.96	94.29	88	89.29	60	82	87.88	80.81
Test set 5	84.48	85.71	72	92.86	64	79.17	100	83.33
Overall (%)	76.30	87.59	80.95	94.26	65.37	82.23	94.55	83.75

amplified expression set is trained using the optimal model determined in the previous section experiment, and some samples are subjected to (−10, the rotation of the −5, 5, and 10) degree, and the mirror flip process amplifies the sample to 151,974 sheets. The network trained a total of 1,000 rounds, with the training batch size set to 512. In training, 75% of the total sample was randomly selected as the training sample, and the remaining 25% was used as the test sample. The results of statistical experiments were compared to compare the classification accuracy and generalization ability of models based on different amplification sets. In this section, the data balance uses the SMOTE algorithm to linearly interpolate a small number of dense regions of the original expression set. The processed expression set contains a total of 3,500 expression images, and the seven types of expression samples are balanced and each has 500. It is further hoped to improve the generalization ability of the model. In this paper, the existing samples of the balanced expression set are randomly rotated and distorted to generate new samples, and the new samples are added to the existing training set to obtain the amplification set. According to the need for experimental comparison, the CNN models established in the previous section were used to train a plurality of different degrees of amplification sets, and the ability to obtain the expression classification is shown in Table 4. Generalization ability of the amplification set. Figure 6 shows the simulation results of 35,000 iterations (about 470 rounds). On the whole, the loss on the test set gradually decreases with the number of iterations. After 35,000 times, the loss tends to be stable at around 0.2 and no longer decreases. The accuracy rate on the test set gradually increases with the number of iterations. Through gradient learning and continuous learning and updating of the model weight parameters, the classification ability of the network is getting stronger and stronger, and the accuracy of the final test set is basically stable.

Table 4: Expression recognition using different amplification ratio data.

Amplification factor	Input map resolution	Classification accuracy (amplification) (%)
1	32 × 32	98.74
	64 × 64	98.71
2	32 × 32	98.52
	64 × 64	98.76
3	32 × 32	98.76
	64 × 64	98.87
4	32 × 32	98.84
	64 × 64	98.90
5	32 × 32	98.85
	64 × 64	98.91
6	32 × 32	98.85
	64 × 64	98.90

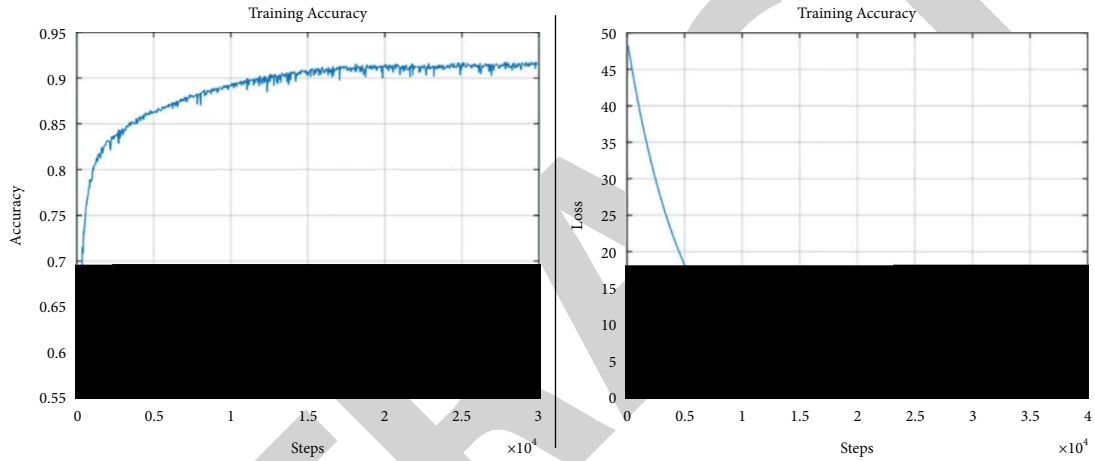

Figure 6: Accuracy and loss trend curve of training on the database.

Data amplification is beneficial to improve the generalization ability of the model, and the richer the amplified data, the stronger the generalization ability, because: The CNN structure itself is a complex network with a large number of parameters, and the richness of the dataset can alleviate its overlearning of existing samples. The random data amplification method in this paper differentiates the information contained in the same kind of expression, which is beneficial to the model because enables it to pay more attention to the common features in the process of fitting the data. In addition, as can be seen from Table 2, the data preprocessing in this paper has a positive impact on the expression classification ability, but the improvement of the classification ability does not increase linearly with the amplification multiple, because the data amplification is only limited to the existing data. A manual processing method is not a panacea, and the quality of the original data is often the bottleneck of the final recognition performance.

5. Conclusions

In today's society, facial expressions are a popular form of expressing affective messages in relation to public events. Facial expressions contain a diverse collection of signals, including an individual's affective mood state, individual traits, and perceptions of events. In addition, numerous elements relating to an individual's state of mood, mental conditions, and physiological status are also relevant to the messages contained in embedded visual representations. Psychologist Mehrabian proposed that in the process of human communication, only 7% of the information is expressed by language. Compared with this, facial expression is 55% of the information, which is the most of the heavy part of the information. It can be seen that the research on the facial expression recognition technology of the intelligent service robot of the nanofiber flexible wearable electronic skin can help us obtain a lot of valuable information and then analyze and study the human psychological and physiological state, thereby improving people's living standards.

In order to achieve the purpose of quantitative analysis of human emotions, research studies on human expression recognition technology is the key to success. At the same time, expression recognition plays an indispensable role in intelligent robots and smart cities. With the continuous development of artificial intelligence and machine learning technology, the demand for intelligent human-computer interaction in society is also expanding. With the deepening of research, many outstanding image open-source

frameworks and high-performance hardware devices have provided good technical support for the development of facial expression recognition. At present, the expression recognition technology of intelligent service robots with nanofiber flexible wearable electronic skin is not a unique research topic in the field of computer vision but a cross-domain topic in many fields such as biology and psychology. The research results have important implications for many fields, such as human-computer interaction, healthcare, and safe driving.

Although this paper uses the intelligent service robot facial expression recognition technology to conduct a profound study on the wearable electronic skin of nano-flexible fibers, there are still many deficiencies. The depth and breadth of the research in this paper is not enough. Moreover, the acquisition is carried out under absolutely ideal conditions, and the completeness and effectiveness are not enough. In future works, we will study suitable methods and means from more angles based on the existing technology and level and continuously improve the recognition quality.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

References

[1] Z. Lv and L. Qiao, "Deep belief network and linear perceptron based cognitive computing for collaborative robots," *Applied Soft Computing*, vol. 92, Article ID 106300, 2020.

[2] Z. Lv, S. Zhang, and W. Xiu, "Solving the security problem of intelligent transportation system with deep learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4281–4290, 2020.

[3] Z. Lv, D. Chen, and Q. Wang, "Diversified technologies in internet of vehicles under intelligent edge computing," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 4, pp. 2048–2059, 2021.

[4] A. admin and M. Gupta, "Ensemble learning for facial expression recognition," *Fusion: Practice and Applications*, vol. 2, no. 1, pp. 31–41, 2020.

[5] D. Gera and S. Balasubramanian, "Consensual collaborative training and knowledge distillation based facial expression recognition under noisy annotations," *International Journal of Engineering Trends and Technology*, vol. 69, no. 7, pp. 244–254, 2021.

[6] D. Li, Z. Wang, Q. Gao, Y. Song, X. Yu, and C. Wang, "Facial expression recognition based on Electroencephalogram and facial landmark localization," *Technology and Health Care*, vol. 27, no. 4, pp. 373–387, 2019.

[7] M. S. Hossain and M. A. Yousuf, "Real time facial expression recognition for nonverbal communication," *The International Arab Journal of Information Technology*, vol. 15, no. 2, pp. 278–288, 2018.

[8] Z. Jan and M. Sajjad, "Facial expression recognition using weber discrete wavelet transform. Journal of intelligent & fuzzy systems," *Applications in Engineering and Technology*, vol. 33, no. 1, pp. 479–489, 2017.

[9] Y. Wang, X. Ma, and P. Qian, "Wind turbine fault detection and identification through PCA-based optimal variable selection," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 4, pp. 1627–1635, 2018.

[10] M. K. Rusia and D. K. Singh, "An efficient CNN approach for facial expression recognition with some measures of overfitting," *International Journal of Information Technology*, vol. 13, no. 6, pp. 2419–2430, 2021.

[11] P. Y. Hong, L. M. Huang, C. Y. Chang, and C. A. Lin, "Lattice Boltzmann simulations of cavity flows on graphic processing unit with memory management," *Journal of Mechanics*, vol. 33, no. 6, pp. 863–871, 2017.

[12] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[13] C. Weng, R. B. Ghazali, S. A. Mustafa, A. N. Kareem, and B. A. Khalaf, "Weather forecasting for batu pahat using neural network," *Fusion: Practice and Applications*, vol. 6, no. 2, pp. 64–70, 2021.

[14] S. S. S. Kruthiventi, K. Ayush, and R. V. Babu, "Deepfix: a fully convolutional neural network for predicting human eye fixations," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4446–4456, 2017.

[15] H. Chen, Y. Zhang, W. Zhang et al., "A Low-dose CT via convolutional neural network," *Biomedical Optics Express*, vol. 8, no. 2, pp. 679–694, 2017.

[16] M. K. Kim, "Contactless palmprint identification using the pretrained VGGNet model," *Journal of Korea Multimedia Society*, vol. 21, no. 12, pp. 1439–1447, 2018.

[17] Z. Lu, X. Jiang, and A. Kot, "Deep coupled resnet for low-resolution face recognition," *IEEE Signal Processing Letters*, vol. 25, no. 4, pp. 526–530, 2018.

[18] R. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters, "One-bit compressive sensing of dictionary-sparse signals," *Information and Inference: A Journal of the IMA*, vol. 7, no. 1, pp. 83–104, 2017.

[19] N. Anselmi, G. Oliveri, M. A. Hannan, M. Salucci, and A. Massa, "Color compressive sensing imaging of arbitrary-shaped scatterers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 6, pp. 1986–1999, 2017.

[20] H. Rauhut and C. Schwab, "Compressive sensing Petrov-Galerkin approximation of high-dimensional parametric operator equations," *Mathematics of Computation*, vol. 86, no. 304, pp. 661–700, 2016.

[21] B. Manavalan and J. Lee, "SVMQA: support–vector-machine-based protein single-model quality assessment," *Bioinformatics*, vol. 33, no. 16, pp. 2496–2503, 2017.

[22] D. Tien Bui, Q. P. Nguyen, N. D. Hoang, and H. Klempe, "A novel fuzzy K-nearest neighbor inference model with differential evolution for spatial prediction of rainfall-induced shallow landslides in a tropical hilly area using GIS," *Landslides*, vol. 14, no. 1, pp. 1–17, 2017.

[23] T. Wiatowski and H. Bölcskei, "A mathematical theory of deep convolutional neural networks for feature extraction," *IEEE Transactions on Information Theory*, vol. 64, no. 3, pp. 1845–1866, 2018.

[24] S. Choi, J. H. Shin, J. Lee, P. Sheridan, and W. D. Lu, "Experimental demonstration of feature extraction and dimensionality reduction using memristor networks," *Nano Letters*, vol. 17, no. 5, pp. 3113–3118, 2017.

[25] J. Li, D. Chen, N. Yu, Z. Zhao, and Z. Lv, "Emotion recognition of Chinese paintings at the thirteenth national exhibition of fines arts in China based on advanced affective computing," *Frontiers in Psychology*, vol. 12, Article ID 741665, 2021.

[26] M. L. Berardi, J. J. Staples, S. H. Ferguson, and E. J. Hunter, "Speech formant changes due to repeated measurements, instruction, and simulated environments using both automated and manual feature extraction," *Journal of the Acoustical Society of America*, vol. 143, no. 3, p. 1965, 2018.

[27] X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-based edge-preserving features for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7140–7151, 2017.

[28] S. G. Meshram and S. K. Sharma, "Prioritization of watershed through morphometric parameters: a PCA-based approach," *Applied Water Science*, vol. 7, no. 3, pp. 1505–1519, 2017.

[29] A. Brazhe, J. Fordsmann, and M. Lauritzen, "Correction for non-rigid movement artefacts in calcium imaging using local-global optical flow and PCA-based templates," *Journal of Cerebral Blood Flow and Metabolism*, vol. 37, no. S1, pp. 360-361, 2017.

[30] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 3, pp. 2727–2736, 2018.

[31] X. Y. Zhang, F. Yin, Y. M. Zhang, C. L. Liu, and Y. Bengio, "Drawing and recognizing chinese characters with recurrent neural network," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 849–862, 2018.

[32] J. Zhou, W. Li, X. Yu, X. Xu, X. Yuan, and J. Wang, "Elman-based forecaster integrated by adaboost algorithm in 15 min and 24 h ahead power output prediction using pm 2.5 values, pv module temperature, hours of sunshine, and meteorological data," *Polish Journal of Environmental Studies*, vol. 28, no. 3, pp. 1999–2008, 2019.