Gabor Representation Combination of PCNN Deep Learning Method for Expression Recognition

▯▯ ▯ (jsjyaolisha@163.com)
Anhui Xinhua University https://orcid.org/0000-0002-6006-3092
Haifeng Zhao
Anhui University

Research Article

Keywords: Gabor representation, Pulse Coupled Neural Network (PCNN), facial expression recognition

Posted Date: July 6th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1257571/v1

Gabor Representation Combination of PCNN Deep Learning Method for Expression Recognition

Lisha Yao*
School of Big Data and Artificial Intelligence
Anhui Xinhua University
Hefei, Anhui, China
jsjyaolisha@163.com

Haifeng Zhao
School of Computer Science and Technology
Anhui University
Hefei, Anhui, China
senith@163.com

Abstract—**Traditional recognition methods are simple to extract features and need to be manually extracted with high complexity and unstable accuracy. The expression recognition method of deep learning still has the problems of poor network representation ability and low recognition rate. In order to fully represent the complex texture and edge features of expression images, a deep learning method of expression recognition based on Gabor representation combined with PCNN was proposed. Firstly, different Gabor representations are obtained through a set of Gabor filter banks with different proportions and directions, and the corresponding convolutional neural network model is trained to generate G-CNNs. Then, the Pulse Coupled Neural Network (PCNN) was introduced to fuse the different outputs of G-CNNs. Experiments in CK+ and JAFFE databases show that the average recognition rates of this method obtained 94.87% and 96.91%. Compared with other methods, the proposed method achieves a better recognition effect.**

Keywords- Gabor representation; Pulse Coupled Neural Network (PCNN);facial expression recognition

I. INTRODUCTION

With the continuous improvement of artificial intelligence technology, the demand for human-computer interaction technology also increases. Emotional expression is a basic way for human beings to express emotions, and it is the primary medium for emotional communication between individuals and people at the beginning. Facial expression recognition has injected new blood into artificial intelligence, and the study of artificial intelligence has begun to revive and set off another research climax [1]- [2].Expression recognition is widely used in the fields of human-computer interaction, auxiliary medicine, intelligent transportation, distance education and safe driving.

At present, facial expression recognition is mainly divided into two categories: the first is the traditional expression recognition method, which mainly focuses on the extraction and classification algorithm of facial expression features. Research on feature extraction mainly includes Local Binary Pattern (LBP) [3]- [4], Active Appearance Model (AAM) [5], Gabor wavelet transform [6]- [7], Lacal Discriminative Component Analysis (LDCA) [8], etc. The classification algorithms mainly include K-Nearest Neighbor (KNN) [9], Sup-port Vector Machine (SVM) [10], Hidden Markov Model (HMM) [11], etc. However, the traditional feature extraction methods generally require manual operation, and the uncertainty of these artificial feature extraction methods loses the original expression feature information to some extent, which makes the recognition robustness poor in practical ap-plication. The second type is the expression recognition method based on deep learning [12-14]. Deep learning abandons the method of manual extraction of features, and realizes the feature representation method of machine autonomous learning. The features learned have better ability of representation and generalization. Lopes et al. [12] proposed a new facial expression recognition method. The method based on convolutional neural network achieves good classification effect, but it can only process a small amount of data and training sequence samples. Ding Mingdu et al. [13] proposed a dual-path feature fusion model, which combined the convolutional neural network (CNN) and directional gradient histogram (HOG) method to extract the HOG features of facial expressions, and then input them to the full connection layer of the convolutional neural network for expression classification, achieving a good effect. He Zhichao [14] et al. proposed a convolutional net-work expression recognition method based on multi-resolution feature fusion, which used the convolutional network to learn features of different resolutions, and then sent them into the full connection layer for feature fusion and expression classification. Experiments showed that the network had strong generalization ability and good robustness.

Deep Convolutional Neural Network (DNC) is an end-to-end overall recognition process that combines data feature extraction, model training, classification and other operations. It has achieved good results in the recognition of large-scale data sets (such as face recognition), but for the small data sets such as facial expression recognition, direct end-to-end training recognition is easy to cause overfitting and poor recognition effect. At the same time, the features extracted from the traditional CNN model are used as the input of the full connection layer. When the network structure is deep, abstract features are mainly extracted, while facial features pay more attention to small features. Therefore, the traditional CNN model is not effective in facial expression classification directly. In summary, the traditional recognition methods are simple to extract features, re-quiring manual extraction, high complexity of features and unstable accuracy. The ex-pression recognition method of deep learning still has the problems of poor network representation ability and low recognition rate.

In order to analyze users' emotions, realize facial expression recognition in real time, efficiently and accurately, and fully extract the rich detail features of facial expression images, a deep learning method of facial expression recognition based on Gabor representation combined with PCNN was proposed. This method first constructs G-CNNs to obtain a set of Gabor representations of facial expressions at different scales and directions, and then uses this set of Gabor representations to train the corresponding CNN. Then use PCNN to carry out feature fusion of multiple outputs of G-CNNs. In this method, Gabor filter was used to enhance and retain detail features of the image. G-CNNs were constructed to improve the feature representation ability of the model, and PCNN was used to fully integrate and retain the features of the image.

II. Related Technologies

A. Gabor Representation

Gabor filters are mainly used for texture analysis and edge detection, with good directional sensitivity, and can obtain features of different scales and directions of images [15]. Studies have shown that the brain responds to visual signals very much like a Gabor filter. Gabor filter can obtain the local structure features of the corresponding orientation and scale selectivity of the image through Gabor filter, so Gabor filter is often ap-plied in computer vision. Gabor filter has good characteristics of rotation, translation and scale invariance. Kernel function of Gabor filter adopted in this paper is expressed as:

$$g(x, y : \lambda, \theta, \varphi, \sigma, \gamma) = exp(-\frac{x^{'2} + \gamma^2 + y^{'2}}{2\sigma^2}) exp(i(2\pi \frac{x^{'}}{\lambda} + \varphi)) \quad (1)$$

Equation (1) is formed by real and imaginary components representing orthogonal directions, as shown in Equations (2) and (3) respectively.

$$g(x, y : \lambda, \theta, \varphi, \sigma, \gamma) = exp(-\frac{x^{'2} + \gamma^2 + y^{'2}}{2\sigma^2}) cos(i(2\pi \frac{x^{'}}{\lambda} + \varphi)) \quad (2)$$

$$g(x, y : \lambda, \theta, \varphi, \sigma, \gamma) = exp(-\frac{x^{'2} + \gamma^2 + y^{'2}}{2\sigma^2}) sin(i(2\pi \frac{x^{'}}{\lambda} + \varphi)) \quad (3)$$

Where, λ represents the wavelength of sine wave, θ represents the direction of image rotation, φ represents the phase shift of sine wave, σ represents the standard deviation of Gaussian function, and γ represents the aspect ratio of image.

$$x^{'} = xcos\theta + ysin\theta \quad (4)$$

$$y^{'} = -xsin\theta + ycos\theta \quad (5)$$

Gabor filter has good directional and scale characteristics [23-26]. In order to make full use of the multi-scale and multi-direction texture and edge information of Gabor filter under multi-parameters, Gabor filter banks of different scales and directions were established by setting several parameters for Gabor filter, so as to obtain detail features of images at different scales and in different directions.

B. Convolutional Neural Network

Convolutional Neural Network (CNN) has its own advantages over networks with interconnected neurons. First of all, CNN can avoid explicit feature extraction and carry out implicit self-learning and feature extraction. Secondly, CNN's shared weights reduce the number of weights and provide the possibility to build a deeper network. The input layer is the two-dimensional data of the image, and the one-dimensional feature vectors of feature mapping extracted by layer transformation are generated alternatively by layer transformation between the luminal layer and the sub-sampling layer. CNN's training is divided into forward communication and backward communication. Forward propagation mainly convolves and subsamples the trained two-dimensional image data and parameters to calculate the error loss between the predicted output and the real output. Gradient descent iterative optimization is used to adjust the weight and bias parameters for backward propagation.

The convolution operation of CNN's convolutional layer is relatively complex, but its purpose is relatively simple, which is to simplify complex data expression, filter out complex data noise, and extract key features. Convolution is the process of generating the function h from the functions f and g. The mathematical definition of convolution is shown in Equation (6) :

$$h(x) = f(x) * g(x) = \int_{-\infty}^{+\infty} f(t) g(x-t) dt \quad (6)$$

The function $h(x)$ is looking for the relationship between the curve of $f(t)$ and $g(x-t)$ and the area enclosed by $y=0$ (the t-axis) and x. With the change and movement of x, the corresponding change process of $h(x)$ generated by the change and change of $g(x-t)$ is the process of feature extraction. And the convolution kernel is extracting features during this movement.Its calculation process is expressed as follows:

$$y(x) = f(x^T W + b) \quad (7)$$

In Equation (7), $y(x)$ is the output of the current layer neuron, x is the output of neurons at the previous layer, W represents the convolution kernel, b represents the bias, and $f(\bullet)$ represents the activation function.

Identify applicable sponsor/s here. *(sponsors)*

Depth of convolution neural network to classify the data feature extraction, model training, such as operating together an end-to-end overall recognition process, its identity in the massive data sets (such as face recognition) has obtained the good effect, but for facial expression recognition data such as small data sets, end-to-end training recognition directly, it is easy to cause a fitting, to identify the effect not beautiful. At the same time, the features extracted by the traditional CNN model are used as the input of the full connection layer. When the network structure is deep, the extracted features are mainly abstract features, while the expression features pay more attention to small features. Therefore, the direct expression classification effect of the traditional CNN model is not good.

C. PCNN Model

The Pulse Coupled Neural Network (PCNN) model can simulate the ability of human brain autonomous learning, and adjust the weight coef-ficient of each eigenvalue flexibly and adaptively by training appropriate parameters suitable for different sample sets through autonomous learning.

The PCNN model is shown in Figure 1.The learning process of a single neuron in PCNN is as follows:

- Receiving unit:

It mainly receives feedback input F_{ij} and connection input L_{ij} . F_{ij} is feedback input of position (i, j), whose value comes from the image information I_{ij} of external input, and n is the number of iterations, as shown in Equation (8). The connection input L_{ij} receives information from the neighborhood neuron through the connection matrix W . α_L is the attenuation coefficient, V_L is the amplitude, (k,l) is the location of the neighborhood neurons, and $Y_{kl}(n)$ is the neuron output after the last iteration, as shown in Equation (9).

$$F_{ij}(n) = I_{ij} \tag{8}$$

$$L_{ij}(n) = e^{-\alpha_L} L_{ij}(n-1) + V_L \sum_{k,l} W_{ij} Y_{kl}(n-1) \tag{9}$$

- Modulator unit：

The internal active item U_{ij} is formed by coupling the modulator to modulate feedback input F_{ij} and connect input L_{ij}, as shown in Equation (10).Where, β is the connection strength.

$$U_{ij}(n) = F_{ij}(n)(1+\beta)L_{ij}(n) \tag{10}$$

- Pulse generator：

It mainly produces dynamic threshold θ_{ij} and neuron output Y_{ij}, and α_θ is the time attenuation coefficient of θ_{ij}, as shown in Equations (11)-(12).

$$\theta_{ij}(n) = e^{-\alpha_\theta} \theta_{ij}(n-1) \tag{11}$$

$$Y_{ij}(n) = \begin{cases} 1 & U_{ij}(n) > \theta_{ij}(n) \\ 0 & U_{ij}(n) \le \theta_{ij}(n) \end{cases} \tag{12}$$

θ_{ij} starts to decay from the initial value. When θ_{ij} decays to less than U_{ij}, Y_{ij} is 1, and the function T_{ij} recording the firing times is also increased by 1, otherwise, the neuron output Y_{ij} is 0.

After iterative calculation, the ignition times corresponding to position (i, j) in the feature graph are T_{ij}.The function T_{ij} is expressed as follows:

$$T_{ij}(n) = T_{ij}(n-1) + Y_{ij}(n) \tag{13}$$

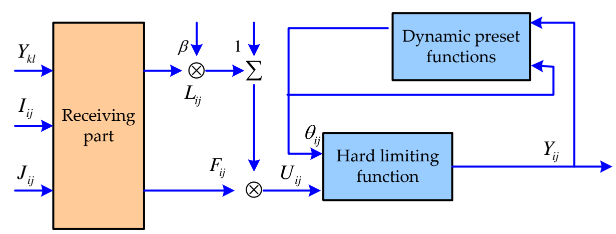

Figure 1. Single neuron structure model of PCNN.

III. G-CNNs Combined with PCNN Expression Recognition

In order to realize facial expression recognition in real time, efficiently and accurately, combining with the multi-feature characteristics of Gabor filter, the image detail features are fully represented. At the same time, PCNN is introduced to use PCNN to fuse the different out-puts of G-CNNs, and facial expression recognition is completed through the full connection layer.

The method is divided into two steps.

In the first step, the expression training data set is filtered by 16 Gabor filter banks of different scales and directions to obtain the different Gabor representations of 16 images, which are respectively the image detail information under different parameters, so as to strengthen the image detail information. Gabor_i ($i = 1, 2, \ldots, 16$)was used to train the corresponding 16 CNNs respectively to construct G-CNNs. Because different Gabor representations extract and enhance the specific texture and edge information, the corresponding CNN is trained to obtain the detailed features of enhanced image expressions in a specific direction and scale.

In the second step, in order to make full use of the features of different scales and directions under the multi-parameter Gabor filter, PCNN is introduced to fuse the features of different Gabor representations through the corresponding G-CNNs, and finally it is sent into the full connection layer to complete expression recognition. The overall identification process is shown in Figure 2.

A. *Gabor Representation of the Image*

The expression images represented by Gabor can enhance the enhanced detail features of images, which improves the feature representation ability of the G-CNNS model. In order to obtain Gabor representations of facial expressions in different directions and different scales, a group of multiparameter Gabor filter banks should be set.

Two-dimensional Gabor filter can well describe the receptive field characteristics of human visual cortex cells, so a two-dimensional Gabor filter is defined in this paper, as follows:

$$\psi_{u,v}(z)=\frac{\left\|k_{u,v}\right\|^2}{\sigma^2}e^{(\left\|k_{u,v}\right\|^2\left\|z\right\|^2/2\sigma^2)}\times[e^{ik_{u,v}z}e^{\sigma^2/2}] \qquad (14)$$

In Equation (14), $z=(x,y)$ represents pixel position, u represents direction of the filter bank, and v represents the scale of the filter bank. $k_{u,v}=k_v e^{i\phi_u}$, $k_v=k_{\max}/f^v$, $\phi_u=\pi u/8$, $k_{\max}$ represents the maximum frequency and f represents the spacing factor between filters in the frequency domain.

Therefore, filter can obtain different Gabor representations of the image by setting parameters U and V. The specific form of Gabor is as follows:

$$Gabor_{u,v}(z)=I(z)*\psi_{u,v}(z),$$
$$for \quad 0\le u\le U-1, 0\le v\le V-1 \qquad (15)$$

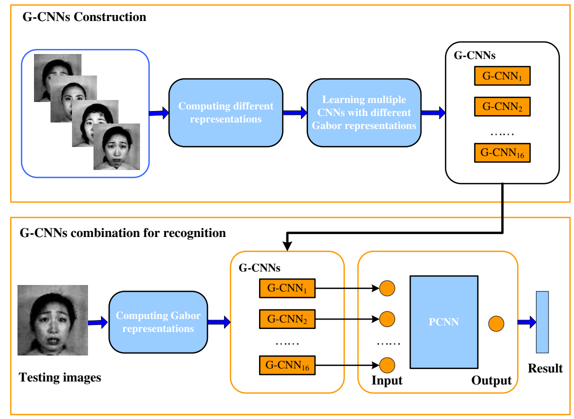

Figure 2. G-CNNs combined with PCNN expression recognition process.

$Gabor_{u,v}(z)$ is the Gabor representation of a given image position z in a specific direction U and scale V, and its value is the convolution result of a given image and Gabor kernel function. The given facial expression image is filtered by Gabor filter. The Gabor filter bank is a 4×4 filter bank (that is, 16 Gabor filters in different directions and different scales). The direction of Gabor filter bank is set as 0°, 90°, 180°, 270°, and the scale is set as 4, 8, 16, 32.

The facial expression images are convolved by Gabor filter banks to obtain 16 Gabor representations. Figure 3 shows 16 different Gabor representations of facial image. Different Gabor representations enhanced the key detail features of expression images from different directions and scales, and trained the corresponding G-CNNs to improve the representation ability of the model.

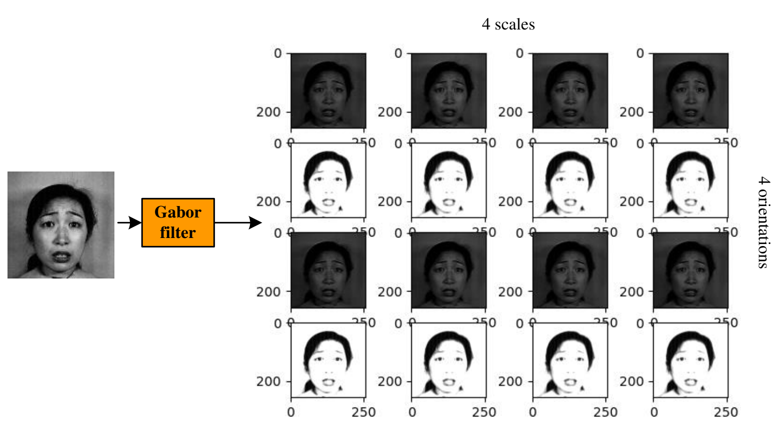

Figure 3. Different Gabor representations of facial images.

B. G-CNNS Structure Design

Different Gabor representations enhance the key details of expression images. In order to fully extract the information of expression images, 16 groups of different Gabor representations were used to train the corresponding CNN. Each group of CNNs has the same structure, and 16 $G-CNN_i (i=1,2,\ldots,16)$ are obtained to construct G-CNNs. Each G-CNN extracts enhanced features in a specific direction and scale, and the process of constructing G-CNNs is shown in Figure 4.

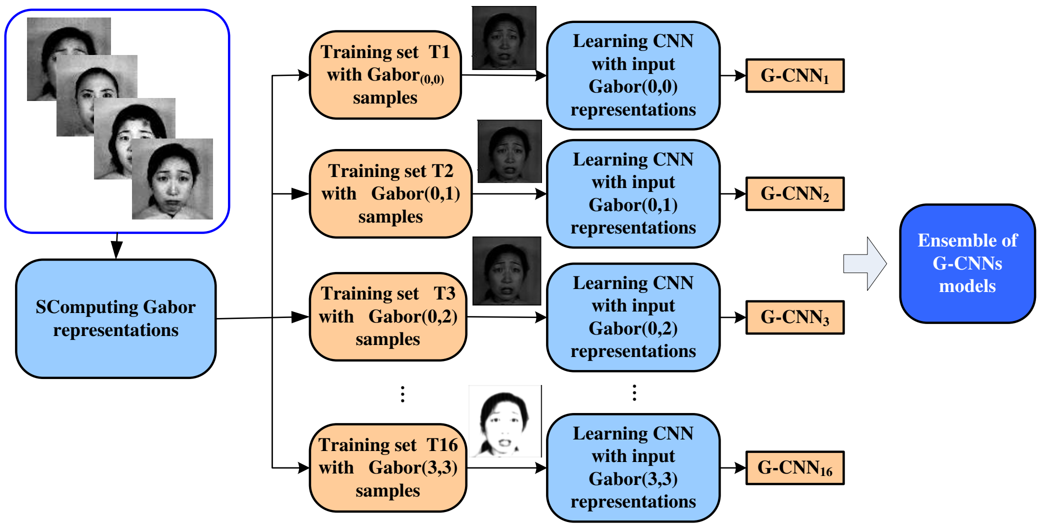

Figure 4. G-CNNs construction process.

Different Gabor representations are used to extract rich texture and edge detail features of facial expressions.

The CNN designed in this paper is shown in Figure 5. The CNN model structure has 8 layers. As shown in TABLE I, the size of the input image is 96 × 96. Convolution kernels are 5 × 5 size. They each have 32,64,128 convolution kernels. Pooling layer S1, S2, S3 uses random pooling, the pooling is 2 × 2 window. The full connection layer has 300 neurons. In order to prevent overfitting, Dropout operation is adopted and its value is set to 0.4 to improve the generalization ability of the model. ReLU is used as activation function and Softmaxloss function is used. After the connection layer the characteristics of the output vector to 1 × 1 × 300.

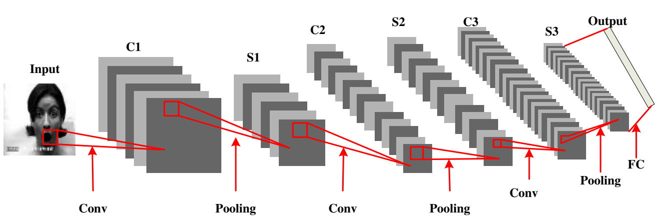

Figure 5. CNN model structure.

TABLE I. PARAMETER SETTINGS OF MODEL

Model Design Structure
input 96×96
C1：5×5 conv,32
S1：2×2，random-pooling
C2：5×5 conv,64
S2：2×2，random-pooling
C3：5×5 conv,128
S3：2×2，random-pooling
Dropout：0.4
FC:300

C. PCNN Feature Fusion

The G-CNNs structure was obtained by training different Gabor representations of facial expression training data set and corresponding G-CNNs. When using test facial expression images for facial expression recognition, enhanced feature information in different directions and sizes can be obtained through trained G-CNNs. PCNN is introduced to perform feature fusion of different Gabor representations through corresponding G-CNNs, which makes full use of the multi-feature information of Gabor filters.

In this paper, multiple features represented by different Gabor were fused by PCNN, and feature vectors extracted by G-CNNs were fed into PCNN. According to the importance of different features, weight coefficients are allocated through PCNN autonomous learning, and the final eigenvalues are obtained by weight fusion. Reasonable allocation of adaptive weight coefficients will improve the overall performance of the model.

The specific eigenvalues of 16 feature vectors extracted by G-CNNs at the same position correspond to one neuron of PCNN. The function T_{ij} represents the information content contained in the eigenvalue. It is used to calculate the fusion coefficient, as shown below:

- After iterative calculation, PCNN obtains 16 specific eigenvalues $f_i (i=1,2,\ldots,16)$ at the same position of eigenvectors, and the ignition times are denoted as t_i respectively.

- The ignition times t_i is normalized by sigmoid threshold function, which is used as the fusion weight coefficient of each eigenvalue, as shown in Equation (16).

$$w_i = \frac{1}{1+e^{-t_i}} \tag{16}$$

- The fusion eigenvalue S of the final specific position is determined, as shown in Equation (17).

$$S = \sum_{i=1}^{16} w_i \times f_i \tag{17}$$

IV. RESULTS

A. Experimental Environment and Data Preparation

In this paper, JAFFE database and CK+ database in Japan are used for experimental evaluation. The JAFFE database in Japan contains seven emojis from 10 women in Japan. CK+ database is an extension of CK database. It is an expression sequence database of angry, bored, fearful, happy, ordinary, sad and surprised. The sample image is shown in Figure 6 and 7.

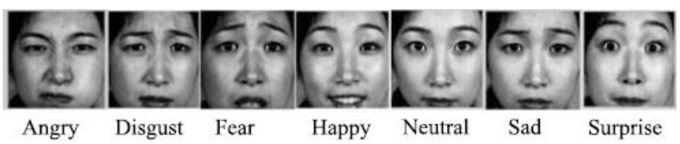

Figure 6. 7 kinds of facial expression image in JAFFE expression dataset.

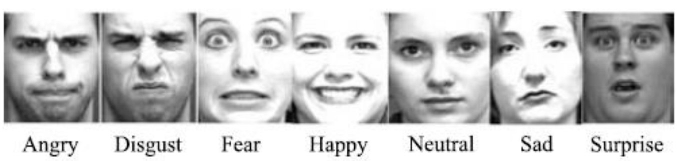

Figure 7. 7 kinds of facial expression image in CK+ expression dataset.

The specific parameter Settings are shown in TABLE II.

TABLE II. NETWORK TRAINING PARAMETER SETTING

Data Set	CK+	JAFFE
Learning rate	0.00001	0.000001
Number of test iterations	200	2000
Attenuation of weights	0.0005	0.0005
Batch number	20	10
Maximum number of iterations	100000	100000

B. Performance Analysis

A deep learning method of expression recognition based on Gabor representation combined with PCNN was used to complete expression classification. In order to reduce the randomness of the experiment, the average recognition rate of the experiment was calculated by 10 times 10 fold cross validation. The confusion matrix of experimental 7 types of facial expressions in CK+ database and JAFFE database is shown in TABLE III and TABLE IV.

TABLE III. CK+ CONFUSION MATRIX %

Expression	anger	disgust	fear	happy	neutral	sad	surprise
anger	95.9	1.49	0.62	0.33	0.27	1.01	0.46
disgust	2.18	93.9	0.89	0.22	0.83	0.89	0.56
fear	1.11	0.68	95.1	0.21	0.58	1.45	0.32
happy	0.15	0.25	0.28	97.6	0.91	0.19	0.76
neutral	1.11	0.89	0.63	0.78	93.1	2.67	1.11
sad	1.19	1.19	2.08	0.26	2.22	92.6	0.39
surprise	0.47	0.79	1.08	0.31	0.78	0.39	95.9

TABLE IV. JAFFE CONFUSION MATRIX %

Expression	anger	disgust	fear	happy	neutral	sad	surprise
anger	98.9	0	0	0	0	0	0
disgust	0.97	97.2	0.73	0.14	0.49	0.73	0.17
fear	0.77	0.51	96.9	0.28	0.74	0.59	0.37
happy	1.58	1.62	1.48	91.1	2.13	1.05	1.29
neutral	0	0	0	0	100	0	0
sad	0.34	0.63	0.99	0.29	0.51	97.2	0.29
surprise	0.52	0.49	0.69	0.36	0.69	0.27	97.1

As can be seen from Table III and Table IV, the average recognition rate of the proposed method in CK+ database and JAFFE database is over 90%. The method presented in this paper has a high misjudgment rate for some categories that do not change obviously and are not easy to distinguish. As shown in Table III, the error rate between classes was higher in disgust and anger (2.18%), sad and neutral (2.22%), and sad and fear (2.08%). The recognition effect of the same algorithm in JAFFE database is better than CK+ database. The reason for this result is that the CK+ database comes from 123 face samples of different genders and countries, which is complicated, while the JAFFE database is only 10 Face samples of Japanese women.

C. Comparison of Different Methods

In order to compare the performance of the proposed facial expression recognition method with other facial expression recognition methods, the experiment was conducted based on the same CK+ data set and the operating platform.

In the experiment, the traditional Gabor feature based machine learning method (reference [6]) and the convolutional neural network based expression recognition method (reference [12] and reference [13]) were respectively used to compare with the proposed method. In literature [6], Gabor transform embedded in principal component analysis was used to extract features from hyperspectral images, local Fisher discriminant analysis or local protected non-negative matrix separation was used to reduce the dimension of Gabor features, and gaussian mixture model classifier was used to classify the features after dimensionality reduction. Literature [12] uses the combination of convolutional neural network and specific image pretreatment steps to complete real-time facial expression recognition. The HOG feature of facial expression of human face was extracted in literature [13] and then input to the full connection layer of convolutional neural network. Finally, the fusion feature was transferred to the output layer, and Softmax classifier was used for recognition and output results.

The depth model training and testing of the model and comparison in this paper carried out iterative training for 200 cycles each time. The sample number of each training batch was set to 24 sample pictures, and the initial value of the neural network learning rate was adjusted to 0.01. In the process of network model training and learning, dynamic attenuation learning rate is used to dynamically update the parameters and weights of the network, which is convenient to find the optimal model. The final classification function of the experimental network model is Sofmax function. Fig. 8 shows the average recognition rate of seven types of expressions in different methods.

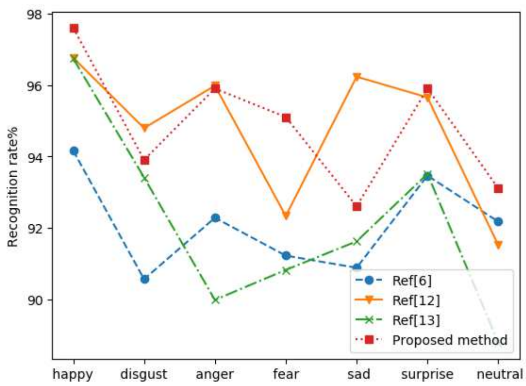

Figure 8. Average recognition rate of 7 kinds of facial expressions in different algorithms%.

Figure 9 shows the variation of accuracy of the algorithm in this paper, literature [12], literature [13] and literature [6] in a certain training process on CK+ expression data set. According to the changes of recognition accuracy in the training process shown in Figure 9, the algorithm in this paper has a faster convergence speed and a better recognition rate. The comprehensive performance data of different algorithms are shown in Table 5. As can be seen from Table V, the recognition effect of the proposed method is better than that of traditional machine learning recognition methods and recent expression recognition results based on deep learning. The method using Gabor filter to enhance and keep the image texture and edge features, and the depth of the face image feature extracting by CNN, by building the G CNNs could improve the ability of the characteristics of the model representation, on the premise of guarantee the recognition efficiency rate has improved, but the recognition efficiency is still not as good as the method of literature [6] single feature recognition.

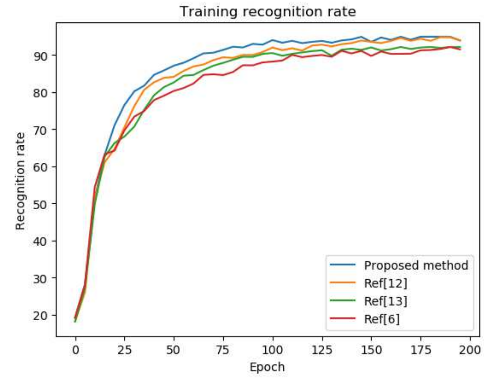

Figure 9. The accuracy trend diagram of different algorithms CK+ data sets.

TABLE V. COMPREHENSIVE PERFORMANCE COMPARISON OF DIFFERENT ALGORITHMS

Algorithm	Recognition rate	Recognition time (ms)
Ref.[6]	92.10%	1690
Ref. [12]	94.75%	1879
Ref. [13]	92.12%	2688
Proposed method	94.87%	2097

V. CONCLUSIONS

This paper proposes a deep learning method of expression recognition based on Gabor representation combined with PCNN. The fuzzy weighted entropy attention model was used to fuse the low-level features of multi-channel expression images into high-dimensional features to complete expression classification. The average recognition rate in CK+ database and JAFFE database is 94.87% and 96.91%, respectively. Compared with other recognition methods, it can be seen that the Gabor representation method combined with PCNN method in this paper fully extracts rich features of ex-pression images, improves the robustness and characterization ability of the model, and improves the recognition rate on the basis of ensuring recognition efficiency. Although the method in this paper has achieved some results, further exploration and research on different databases and refining facial expression features are needed in the future work.

ACKNOWLEDGMENT

Special thanks to the following funds for their support: Key Research Project of Natural Science in Universities of Anhui Province(No.KJ2020A0782); Provincial quality engineering in Anhui Province Grass-roots teaching and research office demonstration project(No.2018jyssf111); Data Science and Big Data Technology University First-Class Undergraduate Program Construction Center (No. 2020ylzyx02);University-level Quality Engineering Demonstration Experiment and Training Center "Big Data Comprehensive Experiment and Training Center" (No. 2020 sysxx01); 2020 Anhui Provincial College Student Innovation Plan Project (No. 202012216083).

DECLARATION

-Ethical Approval
Not applicable

-Consent to Participate

Not applicable

-Consent to Publish

On behalf of all authors, I consent to publish our manuscript.

-Authors Contributions

Conceptualization, L.Y. and H.Z.; methodology, L.Y. and H.Z.; software, L.Y.; validation, L.Y.; formal analysis, L.Y.; investigation, L.Y.; resources, H.Z.; data curation, L.Y.; writing—original draft preparation, L.Y.; writing—review and editing, L.Y.; visualization, L.Y.; supervision, H.Z.; project administration, L.Y.; funding acquisition, L.Y. All authors have read and agreed to the published version of the manuscript.

-Funding

Special thanks to the following funds for their support: Key Research Project of Natural Science in Universities of Anhui Province(No.KJ2020A0782); Provincial quality engineering in Anhui Province Grass-roots teaching and research office demonstration project(No.2018jyssf111); Data Science and Big Data Technology University First-Class Undergraduate Program Construction Center (No. 2020ylzyx02);University-level Quality Engineering Demonstration Experiment and Training Center "Big Data Comprehensive Experiment and Training Center" (No. 2020 sysxx01); 2020 Anhui Provincial College Student Innovation Plan Project (No. 202012216083).

-Competing Interests

There exist no competing interests.

-Availability of data and materials

Some or all data, models, or code generated or used during the study are available from the corresponding author by request.

REFERENCES

[1] Liu Q M, Xin Y Y. Face Expression Recognition Based on End-to-End Low-quality Face Images[J]. Journal of Chinese Computer Systems, 2020, 041(003):668-672.

[2] Li T T,Hu Y L,Wei F L. Improved Facial Expression Recognition Algorithm Based on GAN and Application[J]. Journal of Jilin University(Science Edition), 2020,058(003):605-610.

[3] Qiang W,Zhou X Y,Xu H N,Li D P,An H R. Micro-expression Recognition Based on LBP and Two-stream Spatial-temporal Neural Network[J]. Information and Contro, 2020, 49(06):37-43.

[4] Wu H,Hu M,Gao Y, et al. Facial Expression Recognition Based on DCLBP and HOAG[J]. Journal of Electronic Measurement and Instrumentation,2020,(2):73-79.

[5] Zhou H L,Kinman L,He X J. Shape-appearance-correlated Active Appearance Model[J]. Pattern Recognition,2016,56(C):88-99.

[6] Ye ZH,BAI L,Zhan Y J. Hyperspectral Image Classification Algorithm Based on Gabor Feature and Locality-preserving Dimensionality Reduction[J]. Acta Optica Sinica,2016,36(10):1028003.

[7] Paczos T, Bonham A, Canavesi C, et al. Near-Histologic Resolution Images of Cervical Dysplasia Obtained With Gabor Domain Optical Coherence Microscopy[J]. Journal of Lower Genital Tract Disease, 2021,25(2):137-141.

[8] Jiang B, Jia K B. A Local Discriminative Component Analysis Algorithm for Facial Expression Recognition[J]. Acta Electronica Sinica,2014,42(1):155- 159.

[9] Huang H, Zheng X L. Hyperspectral Image Classification with Combination of Weighted Spatial-spectral and KNN[J]. Optics and Precision Engineering, 2016, 24(4):873-881.

[10] Xu Y, Gu Y, Peng D L,et al. SAR ATR Based on Disentangled Representation Learning Generative Adversarial Networks and Support Vector Machine[J]. Optics and Precision Engineering,2020,28(3):727-735.

[11] Su B C,Chen G,Che R S. Kinetic People Model Learning of Modified HMM[J]. Optics and Precision Engineering, 2009, 17(6):1485-1495.

[12] Lopes A T, Aguiar E D, Souza A, et al. Facial Expression Recognition with Convolutional Neural Networks: Coping with Few Data and the Training Sample Order[J]. Pattern Recognition, 2017, 61:610-628.

[13] Ding M D,Li L. CNN and HOG Dual-path Feature Fusion for Face Expression Recognition[J]. Information and Control, 2020, 49(1):47-54.

[14] He Z C,Zhao L Z,Chen C. Convolution Neural Network with Multi-Resolution Feature Fusion for Facial Expression Recognition[J]. Laser & Optoelectronics Progress,2018,55(07):370-375.

[15] Sreeshan K, Dinesh R, Renji K. Enhancement of Thermographic Images of Composite Laminates for Debond Detection: An Approach Based on Gabor Filter and Watershed[J]. Ndt & E International, 2019, 103, 68-76.