Facial Micro-expression Recognition Algorithm Based on Big Data

Qun Xia¹*, Xiaofeng Ding¹
¹Nanchang Institute of Science and Technology, Nanchang, 330108

*Corresponding author e-mail: 384165625@ncist.edu.cn

Abstract. The 21st century is the era of big data. All aspects of society, from facial expressions to national defense and military, will generate massive amounts of data. Facial expression recognition technology, as a new technology spawned in the era of big data, has broad applications. The prospects are widely used in intelligent transportation, assisted medical care, distance education, interactive games and public safety. In recent years, it has attracted more scholars' attention and has become another research hotspot in the field of computer vision and machine learning. The purpose of this article is to study the facial micro-expression recognition algorithm based on big data. This time, big data technology is used to analyze the algorithm. Big data can better solve the small changes in face recognition and complex data processing. This paper firstly summarizes the basic theory of big data, derives the core technology of big data, and analyzes its shortcomings and shortcomings based on the current research status of facial micro-expression in my country, and finally discusses the big data based on big data. Research on facial micro-expression recognition algorithm under the following. This article takes the research situation of the face micro-expression recognition by related companies as the survey object, and analyzes it through the literature data method, questionnaire survey method, mathematical statistics method and other research methods. Experimental results show that the lower the dimensionality reduction, the less classification time is used. When the dimensionality reduction is 45 dimensions, the recognition rate of facial expressions is the highest.

Keywords: Big Data, Face Recognition, Algorithm Research, Human-Computer Interaction

1. Introduction
With the rapid progress of computer technology in the 21st century, artificial intelligence and pattern recognition technology have developed rapidly, and the research on biometric technology has become a research hotspot [1-2]. Biometric recognition technology mainly includes the recognition of human faces, expressions, voices, fingerprints, palm prints, iris, etc. These recognition technologies have begun to be widely used in all aspects of social life and play an increasingly significant role [3-4]. As a part of biometric technology, facial expression recognition plays a very important role in human-computer interaction as a kind of non-verbal information. People can understand and judge the mental state of the
other party directly through facial expressions, and develop better interactive communication. Therefore, the research on facial expressions was launched in this context [5-6].

Facial expression recognition technology first began in the research work of American psychologists Ekman and Friesen in the 1970s [7-8]. They defined the six basic facial expressions and established the Facial Action Coding System (FACS), which was widely recognized by later researchers and became the cornerstone of research in the field of facial expression recognition [9-10]. Although great progress has been made in the study of facial expressions at this stage, it has not yet entered the stage of recognizing facial expressions. And limited by the development level of science and technology at that time, artificial methods were mainly used to analyze and study facial expressions [11-12].

In this paper, through the research and analysis of big data, combined with the current situation of face recognition in my country, the face micro-expression recognition algorithm under big data is researched and analyzed, and the face recognition algorithm based on big data and our country’s current face recognition algorithm are studied and analyzed. Compared with face recognition algorithms, we will conduct a more in-depth study on the subject of this article.

2. Research on the application of facial recognition micro-expression calculation based on big data

2.1. Preprocessing of facial expression images

A complete facial expression recognition system consists of image preprocessing (face detection and geometric gray-scale normalization), feature extraction, dimensionality reduction, and classification recognition. The quality of processing at any stage will directly affect the accuracy of facial expression recognition. For facial expressions, due to various external factors, such as light, noise, angle, posture, etc., the original image of facial expressions and the target image are deviated; in addition, the difference in the distance and angle between the shooting machine and the subject will make the position, angle and size of the face in the photo have great uncertainty. Therefore, whether the first step effectively solves the problem of expression image preprocessing will directly affect the continuity and accuracy of further processing in the subsequent stages.

(1) Cascade accelerated face detection and eye location algorithm based on dual thresholds.

Human faces are a kind of non-rigid shapes. Although they have strong commonalities, due to the influence of individual expressions, illumination, concealment, and imaging angles, human faces have more subtle and complex morphological changes than other shapes. If you can well extract key features and design a well-designed classifier to detect and locate faces from complex backgrounds and various interference factors, it will provide great opportunities for solving subsequent face recognition, expression recognition, and similar complex pattern detection and recognition problems. Helpful guidance, so face detection is a step that must be pre-processed in the current facial expression recognition.

(2) Accelerated face detection of dual-threshold cascade classifier.

In order to solve the contradiction between missed detection and false detection caused by a single threshold in traditional detection methods, whether it is a single classifier or a cascade classifier, this paper takes into account the speed while setting the detection buffer by using dual thresholds to achieve improvement. The detection rate and the purpose of reducing the false detection rate and the missed detection rate.

1) Sample selection and image preprocessing.

The face training samples used by the system come from MIT and ORL training libraries. In order to facilitate training, uniformly crop it into 21×18 sample images for corresponding classifier training. During training, some alarm non-face samples can be added in time according to the later detection results, which reduces a certain degree of blindness.

2) System acceleration algorithm design.

The traditional detection method based on sliding window slides sequentially in the image to be tested, and each window is sent to the trained neural network classifier for discrimination, and a 21×18
picture is tested in the experiment. After experimental testing, a 21 × 18 picture runs in the neural network classifier trained in this article for about 0.07s. If all windows are tested, it can be seen that the amount of calculation is quite large, which is also one of the reasons for the relatively high time complexity. Based on this, this paper improves the template matching algorithm based on Euclidean distance, removes the eye template, and performs the overall template matching on the test picture using the image pyramid method of multi-scale compression, by retaining the window position with large similarity to the face. Neural network classifier test, and then remove a large amount of background, through this accelerated processing method to improve the efficiency of the previous algorithm.

3) Double threshold cascade classifier.

The first-level classifier uses a neural network method. It is a classifier obtained by BP neural network training after dimensionality reduction through PCA. Because there are some problems in the application of traditional BP neural network, such as when the convergence speed is slow and the global minimum result cannot be guaranteed, this article uses the momentum term with adaptive ability to optimize the design of the BP neural network when training the BP neural network.

(3) Face detection and positioning.

When a face image is obtained, some of the images may have certain deflection and angle problems. In order to better lay the foundation for the subsequent accurate recognition of facial expressions, precise positioning of the human eyes is required before preprocessing. This paper also uses the AdaBoost human eye detection method to obtain the position coordinates of the human eye. By applying a trained human eye detector to locate in the face range obtained above, interference factors such as hair and clothing are eliminated to the greatest extent, and the human eye coordinates positioned by the AdaBoost method are further adopted by certain geometric rules, so you can directly remove the mouth that is easily mistaken for eyes.

4) Preprocessing of facial expression images.

The geometric preprocessing of facial expression images refers to the corresponding adjustment of the size, displacement, angle, etc. of the facial image to make it meet the normalized images used for facial expression recognition in the later stage, and facilitate the extraction of facial expression features. Due to the phenomenon of partial tilt and a certain degree of deflection after the above-mentioned face detection, as well as the inconsistent uncertainty of the size, the processing of this part will directly affect the recognition accuracy in the later stage. Therefore, on the basis of face detection, the AdaBoost eye detection method is used to obtain the position coordinates of the human eye, which lays the foundation for geometric normalization.

2.2. Perceptron network

Artificial neural network is an abstraction and simulation of the human brain. Its basic unit is the perceptron. It is hoped that this structure can simulate the cognitive process of human learning.

(1) Single-layer perceptron.

A single-layer perceptron consists of only one input layer and one output layer, and its activation function is the threshold function. For a set of inputs \( X = \{ x_1, x_2, \ldots, x_n \} \), the output value is after the \( j \)-th neuron:

\[
y_j = f(\sum_{i=1}^{n} w_{ji} x_i + b)
\]

Where \( j \) is the number of neurons in the output layer, \( w_{ji} \) is the connection weight between the \( j \)-th output neuron and the \( i \)-th input neuron, and \( b \) is the bias. Then the actual output \( Y = \{ y_1, y_2, \ldots, y_m \} \).

The learning process of a single-layer perceptron is a supervised learning method. For a set of input \( X \), the target output is \( O = \{ o_1, o_2, \ldots, o_p \} \). When the input passes through the perceptron network, a set of actual outputs \( Y \), parameters The adjustment method of is obtained by comparing the actual output with the target output. The weight adjustment formula is given below:

\[
W_{ji}(k+1) = W_{ji}(k) + \Delta W_{ji}
\]

Where \( \Delta W_{ji} \) is the weight adjustment amount, which can be obtained by the following formula:
$$\Delta w_{ji} = \begin{cases} 2o_j|x_i, y_j \neq o_j \\ 0, y_j = o_j \end{cases}$$ (3)

The above formula represents the learning rate, and generally takes the smaller number between [0,1].

2.3. Convolutional neural network training process
First select the sample (X, Yp) from the sample space, input X into the network, and then calculate the corresponding output variable Op. Through this network model, the sample information is passed from the input layer to the output layer step by step. During this process, it has been processed and changed many times. The calculation rules are as follows:

$$Op=Fn(\cdots(F2(F1(XpW(1))W(2)))\cdots)W(n)$$ (4)

Among them, Fn is the activation function, and W(n) is the weight.

Define $V_{ij}$ to represent the weight from the input layer information $i$ to the hidden layer output information $j$, $W_{jk}$ to represent the weight from the hidden layer output information $j$ to the output layer information $k$, and define the thresholds of the output unit and the intermediate layer unit as $\theta_k$ and $\phi_j$, respectively. Then the output of each unit in the middle layer can be calculated by the following formula:

$$y_k = f\left(\sum_{j=0}^{L-1} V_{ki}h_j + \theta_k\right)$$ (5)

3. Experimental research on the application of face recognition and micro-expression calculation based on big data

3.1. Subjects
This article conducted two experiments. One is to directly put the dimensionality reduction results of the improved principal component analysis method directly into the SVM support vector machine for classification, namely PCA+SVM; the first is to use convolutional neural networks for feature extraction and classification of the dimensionality reduction results.

3.2. Research methods
(1) Mathematical Statistics.
This article sorts out and analyzes the collected data by using related software.
(2) Logical hierarchy method.
Through a detailed overview of the methods, approaches, and algorithms of facial micro-expression recognition, this article makes the structure of this article more rigorous and clearer.

4. Experimental analysis of facial recognition micro-expression calculation application based on big data

4.1. Analysis of the relationship between dimensionality reduction and recognition rate and recognition time
In order to make the experimental data more scientific and effective, some of the improved principal component analysis method feature extraction data dimensions and classification accuracy corresponding relationship, the results obtained are shown in Table 1.

| Table1. Analysis of the relationship between dimensionality reduction and recognition rate and recognition time |
|-------------------------------------------------|
| Recognition accuracy | Recognition time |
|----------------------|------------------|
| 43                   | 87.6%            | 8.7960s         |
| 44                   | 88.4%            | 9.9166s         |
It can be seen from Figures 1 and 2 that the accuracy rate is highest when the dimension is 45, and the time required to change the dimension after 45 is different, so the 45-dimensional recognition ability is the best.

### 4.2. Analysis of facial expression recognition results

Analyze the result analysis of the facial expression classification of the face whose dimensionality is reduced to 45 dimensions. Through experimental statistics, the classification results of facial expressions of happy and neutral facial expressions are obtained. The data is shown in Table 2.

![Figure 1. Analysis of the relationship between dimensionality reduction and recognition time](image1)

**Figure 1.** Analysis of the relationship between dimensionality reduction and recognition time

![Figure 2. Analysis of the relationship between dimensionality reduction and recognition rate](image2)

**Figure 2.** Analysis of the relationship between dimensionality reduction and recognition rate

| Dimension | Percentage | Time (s) |
|-----------|------------|----------|
| 43        | 87.60%     | 8.796    |
| 44        | 88.40%     | 8.9326   |
| 45        | 88.00%     | 8.9214   |
| 46        | 88.00%     | 8.9504   |
| 47        | 89.20%     | 8.9504   |

**Table 2.** Analysis of facial expression recognition results

| Facial Expression | Number of tests | Exact number |
|-------------------|-----------------|--------------|
| Excited           | 130             | 117          |
| neutral           | 123             | 108          |
| Sad               | 102             | 91           |
It can be seen from Figure 3 that the classification accuracy of using the improved principal component analysis method and support vector machine to directly classify facial expressions is related to the dimensionality reduction of the improved principal component analysis method, and the relationship between the classification accuracy and the dimensionality reduction dimensionality roughly shows a normal distribution trend; in addition, the lower the dimensionality reduction dimensionality, the less classification time is used. The experimental results show that the recognition rate of facial expressions is the highest when the dimensionality reduction is 45 dimensions.

5. Conclusion
As the most direct and basic way of expressing human emotions, facial expressions are a very effective way of expression in non-verbal communication. Expressions can not only directly and accurately present personal thoughts and emotions, but also judge the inner emotional state of others based on the other's expressions. Therefore, facial expressions have become one of the most abundant body language to participate in various human activities. If computers and intelligent robots have the ability to understand and express emotions like humans, this will fundamentally raise the level of human-computer interaction to a higher level, so that computers can better serve humans.

Acknowledgments
Project funding: Science and Technology General Project of Jiangxi Provincial Department of Education (No. GJJ202504).

References
[1] Malaspina M, Albonico A, Daini R. Right perceptual bias and self-face recognition in individuals with congenital prosopagnosia [J]. Laterality, 2016, 21(2):118-142.
[2] Ding C, Choi J, D Tao, et al. Multi-Directional Multi-Level Dual-Cross Patterns for Robust Face Recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(3):518-531.
[3] Solanki K, Pittalia P. Review of Face Recognition Techniques [J]. International Journal of Computer Applications, 2016, 133(12):20-24.
[4] Ding C, D Tao. A Comprehensive Survey on Pose-Invariant Face Recognition [J]. Acm Transactions on Intelligent Systems & Technology, 2016, 7(3):37.
[5] Xu Y, Li X, Yang J, et al. Integrating Conventional and Inverse Representation for Face Recognition [J]. IEEE Transactions on Cybernetics, 2017, 44(10):1738-1746.
[6] Lu J, Liong V E, Wang G, et al. Joint Feature Learning for Face Recognition [J]. IEEE Transactions on Information Forensics & Security, 2017, 10(7):1371-1383.

[7] Zhao L, Zhang Y, Yin B, et al. Fisher discrimination-based l2,1-norm sparse representation for face recognition [J]. Visual Computer, 2016, 32(9):1165-1178.

[8] Jian Y, Lei L, Qian J, et al. Nuclear Norm based Matrix Regression with Applications to Face Recognition with Occlusion and Illumination Changes [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(1):156-171.

[9] Valentine T, Lewis M B, Hills P J. Face-space: A unifying concept in face recognition research [J]. Quarterly Journal of Experimental Psychology, 2016, 69(10):1996-2019.

[10] Lu J, Gang W, Moulin P. Localized Multifeature Metric Learning for Image-Set-Based Face Recognition [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2016, 26(3):529-540.

[11] Gao Y, Ma J, Yuille A L. Semi-Supervised Sparse Representation Based Classification for Face Recognition With Insufficient Labeled Samples [J]. IEEE Transactions on Image Processing, 2017, 26(5):2545-2560.

[12] Best-Rowden L, Jain A K. Longitudinal Study of Automatic Face Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(1):148-162.