

Received August 2, 2020, accepted August 9, 2020, date of publication August 13, 2020, date of current version August 31, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3016301

# Research on Screening of Empathy Information Based on Image Recognition and Data Mining

GONG YING<sup>1</sup>, YAN JIN<sup>1</sup>, CUIYU BAO<sup>2</sup>, ZHAORONG CHEN<sup>1</sup>, AND YUNLONG DENG<sup>1</sup>

<sup>1</sup>The Third Xiangya Hospital of Central South University, Changsha 410000, China

<sup>2</sup>Scientific Research Office, Hubei University of Science and Technology, Xianning 437100, China

Corresponding author: Yan Jin (yanjin0163@163.com)

This work was supported by the Scientific Research Project for young and middle-aged Talents of Hubei Provincial Department of Education under Grant Q20192803.

**ABSTRACT** The Internet not only provides help for people to understand the world and facilitate life, but also provides a convenient way for widespread dissemination of bad information. It follows that young people are often harassed by pornographic, violent and other bad images, which affects the development of young people's empathy. In this study, from the perspective of the combination of bad image screening and youth empathy ability, the effect of bad image on youth empathy ability is studied. In this article, a new empathy analysis model is constructed based on traditional empathy theory, combined with image recognition and data mining technology. First, the theoretical principles of bad image recognition technology and their application in the evaluation of empathy ability are expounded. Then, based on image recognition and fusion particle swarm optimization algorithm, the classification of bad images was studied. Finally, on the basis of image classification, a data envelopment model is used to grade the young people's empathy ability. The actual case analysis and performance test results illustrate the superiority of the implemented image classification and empathy evaluation method based on image recognition and data mining. This research has certain theoretical significance for the research of enriching empathy ability and interpersonal relationship. At the same time, it has certain practical significance for improving the youth's interpersonal trust, realizing the harmonious interpersonal relationship and the healthy development of body and mind.

**INDEX TERMS** Youth empathy, image recognition, bad information, particle swarm optimization, data envelope.

## I. INTRODUCTION

With the rapid development of information technology and Internet technology, network information has become a well-known convenient source of information and a leisure method [1]. At the same time, a large amount of unhealthy information such as pornography and obscenity on the Internet has seriously interfered with the normal online life and seriously poisoned the physical and mental health of young people and the development of empathy [2]. Empathy refers to the individual's ability to identify, understand, and cope with the emotional state of others in order to produce an emotional experience consistent with others [3]. In modern society, empathy has an important impact on individual moral development, interpersonal communication, and interpersonal relationships. How to purify the network environment, increase monitoring methods for network

activities, and improve the ability of information recognition have become a strong demand. As its technical support, content-based bad information recognition technology has increasingly attracted people's attention. The content-based bad image recognition and detection technology has recently aroused great interest. At the same time, it is also an important and urgent research topic faced by content-based network filtering systems [4]. The identification of pornographic images is actually an image classification problem. We use content-based methods to study images and use statistical classification methods to identify pornographic images.

Bad image recognition is to determine whether an image is an image with bad information. The bad image information recognition technology mainly uses image analysis technology to segment the skin color area in the image [5]. Forsyth used computer vision and image understanding technology to study the recognition of bad images, and judged whether the images contained pornographic content by detecting the geometric features of skin color segmentation and human

The associate editor coordinating the review of this manuscript and approving it for publication was Zhihan Lv.

posture [6]. Wang implemented a pornographic image analysis system that uses Daubechies wavelet and color histogram features to detect bad images [7]. At the same time, in terms of empathy assessment H. Smith believes that when an individual observes that another person is in a certain strong emotional state, they are naturally born to experience the same feeling as the observed emotional state [8]. Spencer believes that empathy can make all members of the group experience the same kind of emotional state quickly, so that it is possible to make consistent behaviors [9]. Titchener mentioned the concept of “empathy” in English for the first time. He believed that empathy is a process in which an individual actively and diligently enters another person’s inner world [10]. Stoutland believes that empathy is an emotional reaction in which the observer perceives that others are or will experience a certain emotion. Hoffman regards empathy as an emotional response that is more suitable for the situation of others than his own situation [11]. Fables believes that empathy is the perception of the emotional state or situation of others. Greenfield pointed out that empathy is the perception of the emotions of others and the experience and behavior caused by emotions [12].

Generally, image processing itself is a resource- and time-consuming task, and the complexity of the recognition algorithm makes the performance of the image recognition system low. How to improve the performance of the system so that the real-time requirements for identifying and filtering information will also be the main problem in future research. The research of empathy has different forms under different historical conditions. At present, most of the relevant researches are conducted from the perspective of a single subject or a single visual form, and there is no research from the perspective of the combination of graphic images and mathematical statistics [13]. Based on this, based on the theory of empathy ability, combined with image recognition and data mining technology, this article builds a new empathy ability classification model. This method can provide scientific reference and basis for the research of modern empathy ability.

## II. OVERVIEW OF RELATED TECHNOLOGIES

### A. SCREENING OF BAD INFORMATION BASED ON IMAGE CATEGORY

Through research, it is found that the identification of bad information images is best to determine whether the images contain specified objects or objects, which is also the most direct requirement of the content-based image recognition system. In this way, the image matching technology can be used for reference based on the identification of bad information content. For example, we can judge whether an image is a bad image by judging whether an image contains a designated sex organ. This problem is actually an optimization problem. You can use particle swarm optimization or other matching algorithms to find the fitness value to determine whether it contains the specified object. However, due to the uncertainty of the scale of pornographic images and the complexity of the

image background, the study found that directly using matching technology to identify any bad images is not ideal [14].

Target detection refers to the process of detecting and identifying specific target objects in an image, that is, not only to obtain the category information of the objects in the image, but also to output the specific position information of the target object on this basis. At present, there are some traditional detection methods for target detection tasks, as well as some emerging detection methods. The former is to use traditional machine learning methods to generate some candidate frames and perform feature extraction, and then use some common classifiers to complete the detection task. The latter is based on deep learning target detection method, which has strong feature extraction ability and learning ability, the disadvantage is that it contains a large number of parameters and complex levels, the training and learning process is time-consuming, etc., so often need to use high-performance computing Platform and massive data sets to assist neural network learning and training [15]. The calculation process is shown in Figure 1.

Different target objects often have different shapes and are susceptible to lighting conditions, noise and background effects. This makes it difficult to manually design a feature that is suitable for all applications, so different features need to be adopted according to the actual testing scenario and testing requirements. Common features include SIFT features, HOG features, Haar-like features and other classic features.

Scale-invariant feature transform (SIFT) is a classic feature extraction algorithm. It uses differential Gaussian Difference of Gaussian function to construct different scale spaces to detect the feature points existing under these scales, and describes the local feature points. It has a very high tolerance to other interfering factors such as image rotation changes, light brightness changes, etc., which can reduce the impact caused by target deformation, so that the feature extraction operation can be successfully completed, so its robustness is very good. The Haar-like feature is proposed by Papageorge *et al.* It is a feature extraction method based on the gray value of the image, and describes the characteristics of the object according to the change of the sum of the gray values between different blocks in the image. It is a feature method commonly used to describe human faces. For example, compared to other areas, the color of the eyes is darker, and the color of the cheeks is lighter than the area around the eyes. The commonly used features are shown in Figure 2.

Human judgment of bad images is based on the high-level semantics of image expression, while computer vision technology can extract the low-level visual features of the image. There is a large semantic gap between high-level semantics and low-level visual features. If machine learning methods can be used to automatically learn the bad images and find the regularity, then this regularity can be used to solve the problem of bad image recognition. Moreover, bad image recognition is a small sample problem, and how to solve the small sample problem is also a problem to be considered. Particle swarm optimization has the ability of learning

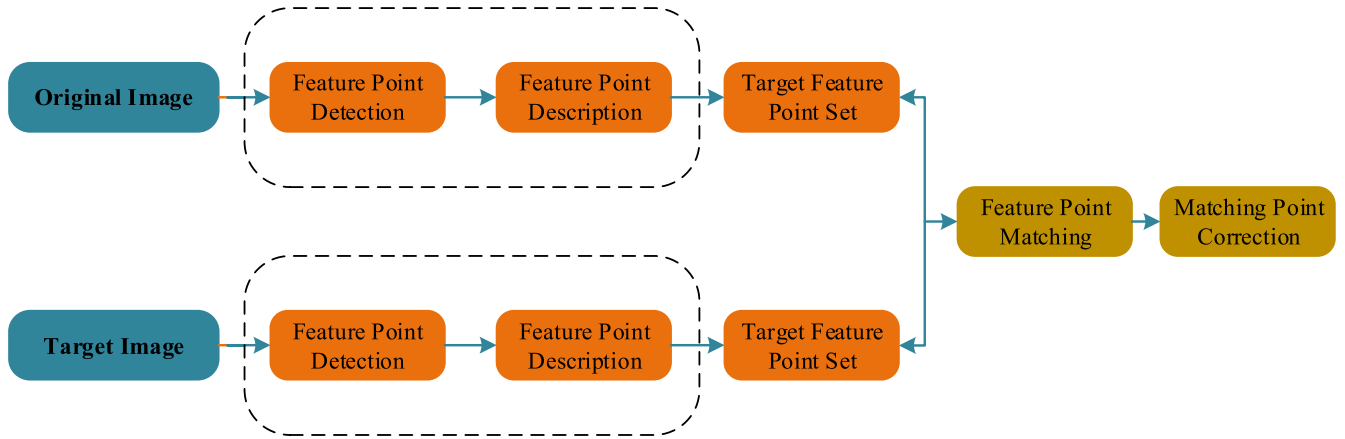


FIGURE 1. SIFT feature calculation process in image recognition.

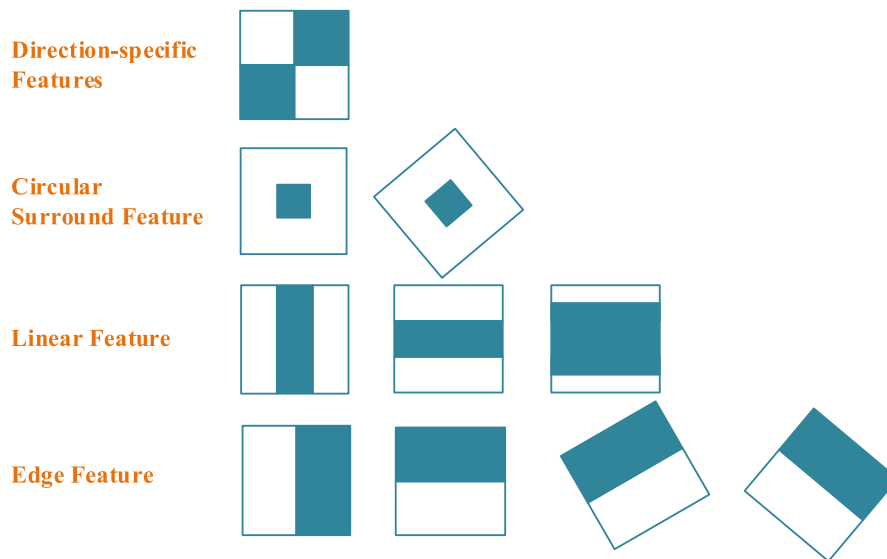


FIGURE 2. Schematic diagram of Haar-like features in image recognition.

and memory, and has great advantages for solving small sample, nonlinear and high-dimensional pattern recognition problems. This article introduces the principle of particle swarm algorithm, and then describes the basic framework of particle swarm algorithm training and poor image recognition algorithm based on particle swarm algorithm. Finally, the corresponding experimental results are given.

**B. EVALUATION OF EMPATHY**

Empathy, also known as empathy, empathy, empathy, empathy, empathy, etc. There have been nearly 100 years of research history in the West, and many branches of philosophy, sociology, and psychology have conducted extensive research on it. Empathy is considered by humanist psychologists as a key characteristic that affects the effectiveness of the psychological counseling process [16]. In the Encyclopedia of Psychology, the definition of empathy refers to the emotional interaction between people in interpersonal

communication. When an individual sees that others endure pain or good luck, they can also experience uneasy or happy emotions. This ability to share the emotions of others is the ability of empathy. Only when the individual can look at the situation from the perspective of the person concerned can the meaning expressed by the other party's emotions be clarified, thereby producing the corresponding emotional experience and behavioral response. Scholars at home and abroad have conducted some psychological theory and empirical research in the field of empathy, but so far there is no unified understanding of this concept [17]. Throughout previous research, there are at least three interpretation orientations for the interpretation of empathy: emotional orientation, cognitive orientation and multidimensional orientation. The conceptual structure of empathy is shown in Figure 3.

The structure of empathy content is complex and diverse, and the measurement methods are also diverse, generally

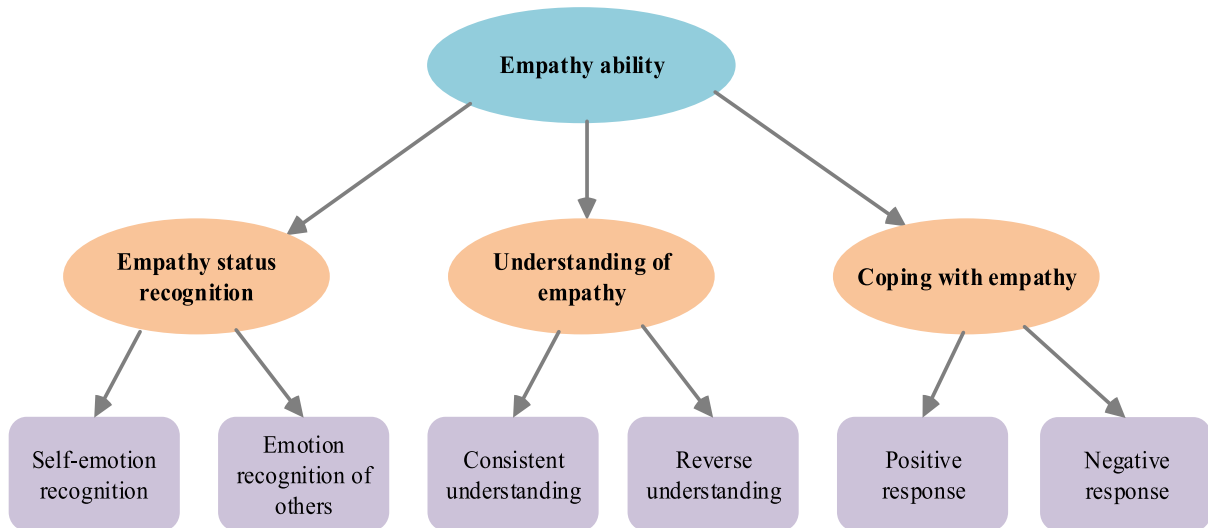


FIGURE 3. The conceptual structure of empathy.

including: self-reporting method, observer evaluation method and physiological measurement method. The indicators used are mainly language indicators, facial indicators and physiological indicators. The tool for measuring adult emotional empathy was once popular with the emotional empathy questionnaire compiled by Mehrabian in 1972. The questionnaire consists of susceptibility to other people's emotional infections, understanding of strangers' feelings, extremely high emotional responsiveness, tendency to be influenced by others' positive emotions, tendency to be influenced by others' negative emotions, sympathy tendency, and spontaneous assistance to others in difficulty 7 The three dimensions constitute [18].

A lot of research has been conducted on the factors that affect empathy. Through the existing empirical research, it is found that the individual differences that affect empathy and related responses can be attributed to factors such as genetics, family, social environment, personality, peers, age, and gender. Previous scholars believe that genetics will affect many personality dimensions, and empathy as a personality dimension is also affected by genetic factors. Empathy is a socialized emotion, and the family is the first place for children to socialize. Parents are the first teachers of children to the society, so the influence of family factors is particularly obvious and important [19]. At the same time, the specific scenes in the process of interpersonal communication will also have a timely impact on empathy.

The measurement of interpersonal relationship generally includes the measurement of a group's interpersonal relationship structure and the measurement of individual interpersonal relationship status. The social measurement method was first founded by the American scholar Moreno in the 1930s. It asks group members to ask questions and ask them to choose other members to understand the interpersonal relationship between the entire group and each member from

the perspective of the group [20]. Behavior measurement method is to observe the factual state of communicative behavior direction, communicative frequency and communicative level without any awareness of the observed person, in order to study people's mutual selection and the characteristics of their actual relationship.

Motivation measurement is a method about the internal reasons for people to interact with each other and establish a good relationship. This method adopts two methods: direct motivation selection and indirect motivation selection. The hierarchical measurement method is to quantitatively determine the proportion of people who hold positive or negative opinions to the total number of people who provide opinions, and then qualitatively analyze the level of people who hold positive or negative opinions, and infer and judge interpersonal relationships from the level analysis Whether it is healthy or negative, it can be more realistic to judge whether the interpersonal relationship is healthy [21]. AHP framework for empathy is shown in Figure 4.

### III. EVALUATION OF BAD IMAGES AND EMPATHY BASED ON DATA MINING

#### A. BAD IMAGE CLASSIFICATION BASED ON PARTICLE SWARM OPTIMIZATION

Suppose there is a D-dimensional target search space, and particles form a particle swarm. Each particle is described by a D-dimensional vector, and its spatial position can be expressed as  $m_i = (m_{i1}, m_{i2}, \dots, m_{iD})$ . This can be regarded as the first solution of the objective optimization problem. Substituting the fitness function to calculate the fitness value can measure the quality of the particles. The flying speed of the particles is also a D-dimensional vector, denoted as  $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ . The position experienced by the particle with the best fitness value is called the historical best position of the individual. The best position experienced by the entire particle group is called the best position in the global history,

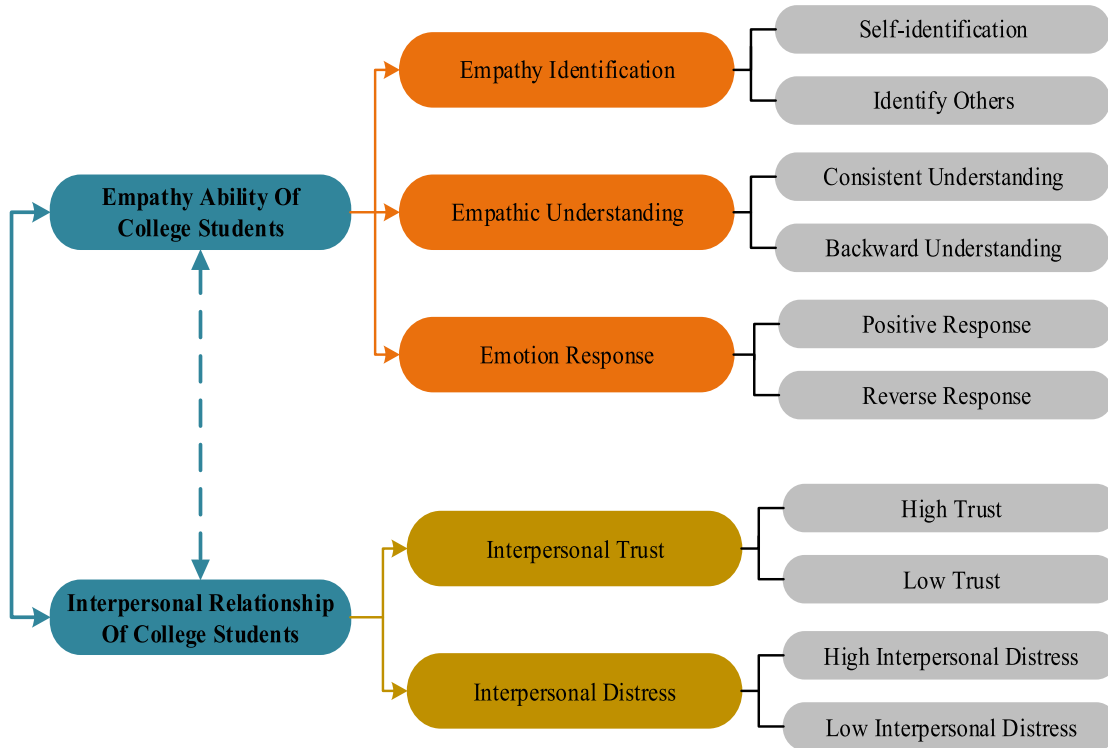


FIGURE 4. AHP framework for empathy.

denoted as  $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ .

$$\begin{cases} v_{ij}(t+1) = v_{ij}(t) + c_1 r_1(t)(p_{ij}(t) - x_{ij}(t)) + c_2 r_2(t)(p_{gj}(t) - x_{ij}(t)) \\ m_{ij}(t+1) = m_{ij}(t) + v_{ij}(t+1) \end{cases} \quad (1)$$

where  $i$  represents the  $i$ -th particle,  $j$  represents the  $j$ -th dimension of particle  $j$ ,  $t$  represents the  $t$ -th generation,  $c_1, c_2$  is two acceleration constants, and  $r_1 \sim U(0, 1), r_2 \sim U(0, 1)$  is two independent random functions.

It can be seen from the above formula that the particles can be adjusted to the best position around itself, and the particles can be adjusted to the best position that the entire particle group can find. The “position” in the algorithm is the optimal result obtained by the price and its four indicators in this article. That is, the “particle” is the dynamic variable that stores and finds the solution. In order to find a good enough optimal solution, and to avoid the increase in the amount of calculation. Particles are set in this model to perform the optimal solution search. Set the particles to store five parameters. First of all, we must initialize all the particles, that is, let the different parameters in the particles take random values within the feasible range. This process is equivalent to spreading the “seed” to find the optimal solution at some points in the entire “position” area. Where to spread the “seed” is closely related to whether it can find the “optimal solution” later. When the distribution position is unreasonable, the “optimal solution” may be a local optimal solution rather than a global optimal solution.

Second, calculate the fitness value of each particle. The fitness value of this question can be calculated using the objective function. The fitness value is defined as:

$$Q = \begin{cases} Y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 \\ a_1 + a_2 + a_3 + a_4 + a_5 = 1 \\ 0 < a < 1 \end{cases} \quad (2)$$

Then record these values. For each particle, compare its fitness with the individual’s best global history. If it is better than the best record, then set this value to the best record. Next, referring to Equation (3), the destination of each particle is adjusted [22]. Every time it is adjusted, it must be re-randomly selected to ensure the randomness of particle behavior. Not only can there be a tendency to approach the optimal solution, but there is also the possibility of finding other more optimal solutions. After finding the first group of optimal solutions, the group of data is removed from the particle swarm, and the remaining data is optimized again to obtain the second group of optimal solutions [23], [24]. That is, the second-best solution after the first group, this operation is performed 20 times, and finally 20 sets of optimized solutions are obtained. In order to ensure that the solution found is optimal, the particles need to be given enough action time to find the solution. Here let the number of iterations be 1000.

A step size of 0.1 is selected, and 20 sets of optimized solutions are obtained by programming. Each set of data is normalized separately, and the new data obtained is used to obtain the relationship between pricing and four indicators

through data regression analysis.

$$Z_1 = -0.002Z_2 - 0.0869Z_3 + 1.6Z_4 + 0.385Z_5 \quad (3)$$

Randomly select 100 unfinished tasks in Annex I, and use these data to get the pricing under the optimal solution. When the following inequality is satisfied, it is certain that the task can be completed under the new pricing scheme, otherwise it cannot be completed.

**B. EVALUATION OF EMPATHY ABILITY BASED ON DATA ENVELOPE**

Data envelopment analysis is a non-parametric estimation method used to evaluate the relative effectiveness of work performance of decision-making units of the same type. With the help of mathematical methods and statistical data to determine the production frontier of relative effectiveness, each decision unit is projected onto the production frontier of DEA, and their relative effectiveness is evaluated by comparing the degree to which the decision unit deviates from the frontier of DEA [25], [26]. Flow chart of particle swarm optimization is shown in Figure 5.

There are decision-making units, each decision-making unit has the same input and the same output:

$$\begin{cases} 1 \rightarrow \\ 2 \rightarrow \\ 3 \rightarrow \\ \vdots \\ m \rightarrow \end{cases} \begin{pmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3n} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \dots & x_{mn} \end{pmatrix} \quad (4)$$

If  $v_i$  is used to represent the weight of input  $i$  and  $u_r$  is the weight of output  $r$ , then the expression of the input-output ratio of decision unit  $h_j$  is:

$$h_j = \frac{\sum_{r=1}^s u_r y_{rj0}}{\sum_{i=1}^m v_i x_{ij0}} \quad (5)$$

Then the performance evaluation of the first decision unit can be classified as the following optimization model:

$$\max h_{j0} = \frac{\sum_{r=1}^n u_r y_{rj0}}{\sum_{i=1}^m v_i x_{ij0}} \quad (6)$$

$$st. \begin{cases} \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1 & (j = 1, \dots, n) \\ v_i \geq 0 (i = 1, \dots, m), u_r \geq 0 (r = 1, \dots, s) \end{cases} \quad (7)$$

By transforming into an equivalent linear programming problem:

$$t = \frac{1}{\sum_{i=1}^m v_i x_{ij}}, \quad w_i = tv_i, \quad \mu_r = tu_r \quad (8)$$

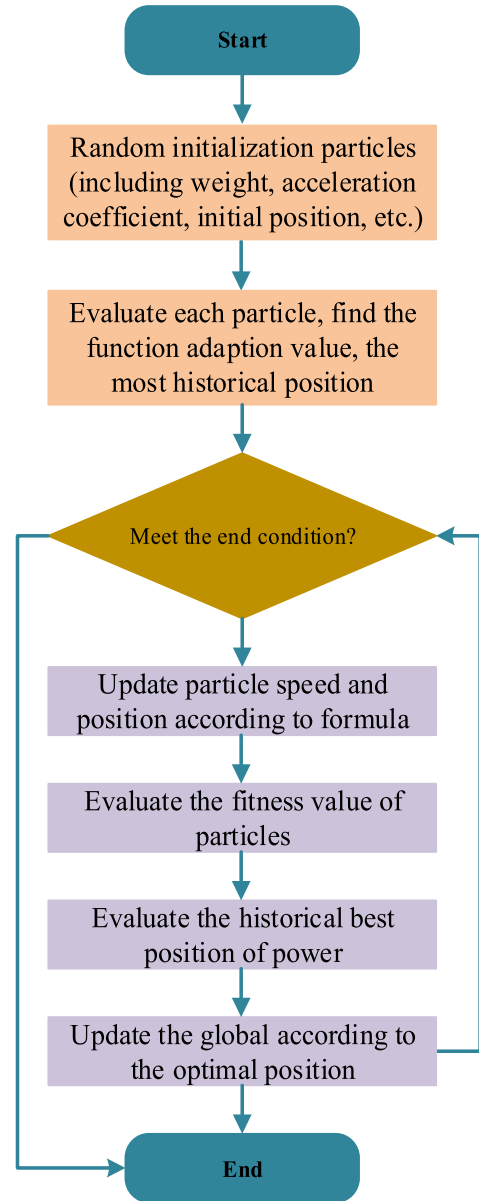


FIGURE 5. Flow chart of particle swarm optimization.

Introducing dual variables, the dual problem of the above model can be written as:

$$\min \theta \quad (9)$$

$$st. \begin{cases} \sum_{j=1}^n \lambda_j x_{ij} \leq \theta x_{ij} & (i = 1, \dots, m) \\ \sum_{j=1}^n \lambda_j y_{rj} \geq y_{rj0} & (r = 1, \dots, m) \\ \lambda_j \geq 0 & (j = 1, \dots, n) \end{cases}$$

In which,  $\theta(0 < \theta \leq 1)$  is the completion rate of the decision-making unit, and the larger the  $\theta$ , the more output the unit inputs. That is, the closer  $\theta$  is to 1, the higher the completion rate of the decision unit.

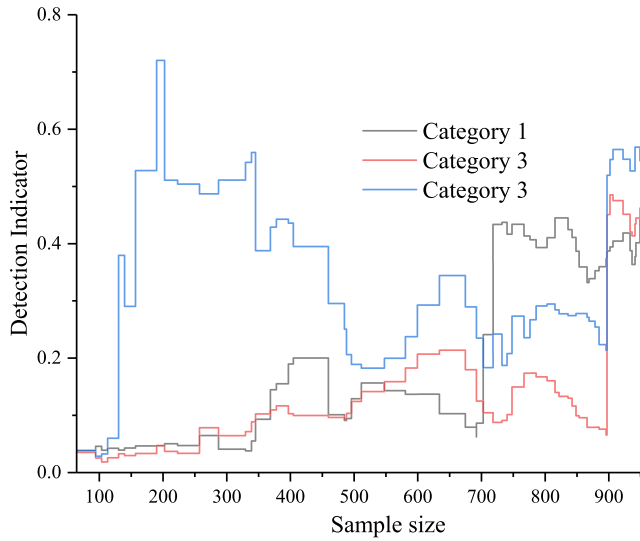


FIGURE 6. Empathy ability classification accuracy in different categories.

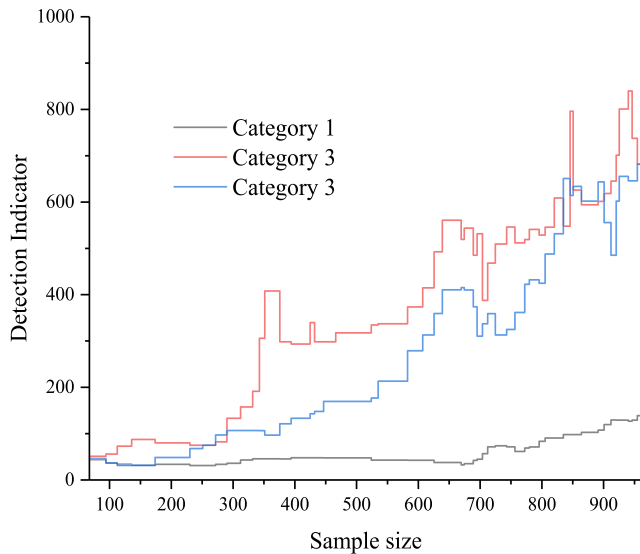


FIGURE 7. Standard value of different types of empathy test.

**IV. CASE ANALYSIS OF BAD INFORMATION CLASSIFICATION AND EMPATHY ABILITY**

**A. CLASSIFICATION OF BAD IMAGES BASED ON DATA MINING**

Empathy ability classification accuracy in different categories is shown in Figure 6. Standard value of different types of empathy test is shown in Figure 7.

The comparison of the classification effect between particle swarm optimization and other algorithms is shown in Figure 8 and Figure 9. The three dimensions of empathy ability, interpersonal trust, and interpersonal distress have no significant differences in professional categories. However, further analysis of the six factors of empathy ability revealed that there are significant differences in the factors of reverse comprehension, forward coping, and reverse coping. There is a significant difference in the reverse understanding factor,

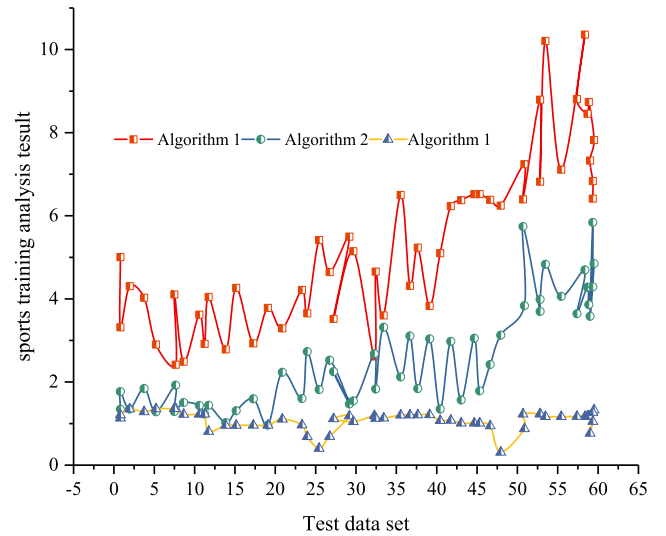


FIGURE 8. Comparison between particle swarm optimization and other algorithms-category accuracy.

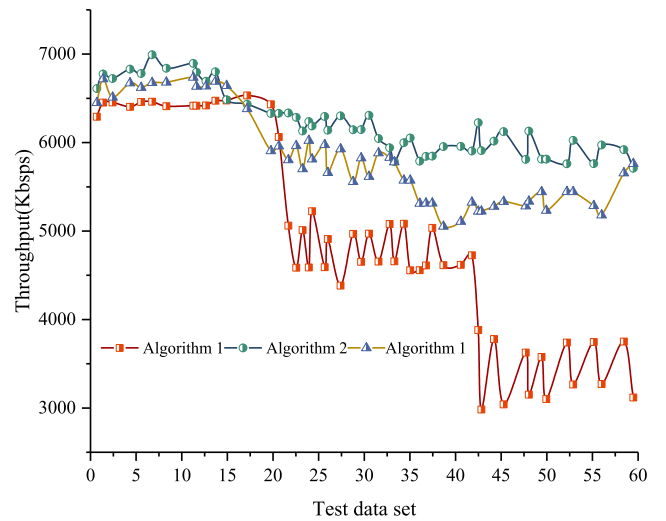


FIGURE 9. Comparison between particle swarm optimization and other algorithms-response efficiency.

and the science and engineering students score higher than the literature and history. The positive coping factors are significantly different, and the scores of literature and history students are higher than those of science and engineering students. The reverse coping factors are significantly different, and science and engineering students score higher than literature and history students. In this article, through the preparation, prediction and testing of the empathy ability questionnaire of youth, a formal survey questionnaire is formed to conduct relevant research and analysis on the empathy ability, interpersonal trust and interpersonal distress of youth, and prove the empathy ability and various dimensions of youth, Factors have different correlations and influences on interpersonal trust and interpersonal distress, which can effectively predict and affect the interpersonal relationship of youth.

**TABLE 1.** Descriptive statistics of empathy, interpersonal trust, and interpersonal distress.

	<i>N</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Standard Deviation</i>
<i>Empathy Ability</i>	500	2.47	4.78	2.98	0.7594
<i>Interpersonal Trust</i>	500	1.54	3.67	4.76	0.4590
<i>Interpersonal Distress</i>	500	1.67	2.98	4.23	0.5353
<i>Empathy Recognition</i>	500	1.33	4.76	5.00	0.2546
<i>Self-Identification</i>	500	1.55	3.27	5.02	0.7945
<i>Identify Others</i>	500	1.05	3.63	4.67	0.7923
<i>Empathic Understanding</i>	500	1.06	3.18	4.89	0.4317
<i>Consistency</i>	500	1.24	2.98	2.98	0.8734
<i>Backward Understanding</i>	500	1.26	4.76	4.76	0.6324
<i>Empathy Response</i>	500	1.33	4.23	4.23	0.5390
<i>Positive Response</i>	500	1.02	5.00	4.23	0.8033
<i>Reverse Response</i>	500	1.00	5.02	5.00	0.4673
<i>Effective N</i>	500				

## B. EVALUATION OF EMPATHY ABILITY BASED ON DATA ENVELOP

In the specific analysis, factor analysis is used to test the structural validity of the forecast questionnaire. The purpose is to extract factors through factor analysis to see whether the structure between these factors is consistent with the conceived structure. If they are completely consistent, the structural validity is high [27], [28]. First, a factor suitability test was performed, and the Bartlett spherical test was performed on the data of the empathy ability questionnaire of the youth. The test value was 3127,  $p < 0.001$ , indicating that there is a possibility of sharing factors in each project. At the same time, the sample suitability measure KMO is 0.803, indicating that the data sample is suitable for factor analysis. Secondly, the principal component orthogonal rotation method was used for exploratory factor analysis, and 11 factors were extracted [29], [30]. The factor load is discrete and distributed among more than two factor items, and further factor analysis is performed. As a result, 8 factors are extracted. Then remove the discrete items with low factor load, and then conduct factor analysis, and then 6 factors are extracted. Then, these 6 factors were analyzed by second-order factors, and it was found that there were 3 higher-order factors. Combining the main content of the content of the item, we named the three factors as empathy recognition, empathy understanding, and empathy coping [31], [32]. Finally, the empathy recognition, empathy understanding, and empathy should be analyzed by the factors contained in each of the three factors. The description statistics of empathy ability, interpersonal trust and interpersonal distress are shown in Table 1.

There is a very significant correlation between the six second-order factors of empathy and interpersonal trust and

interpersonal distress [33], [34]. Specifically, interpersonal distress and the recognition of others in empathic ability, the consistency understanding factor showed a significant negative correlation at the 0.01 level, and the P values were 0.002 and 0.008, respectively, and the correlation coefficients were 0.177 and 0.151. Interpersonal distress and reverse comprehension are positively correlated at the 0.05 level, with a P value of 0.039 and a correlation coefficient of 0.117. Interpersonal distress and reverse coping factors are positively correlated at 0.01 level, and the correlation is significant. Interpersonal trust and reverse coping factors are negatively correlated at 0.01 level, and the correlation is significant, with a correlation coefficient of 0.198. Interpersonal trust and interpersonal distress are negatively correlated at the 0.01 level, and the correlation is significant, with a correlation coefficient of  $-0.219$ .

The results in the analysis table show that there are six types of empathy capabilities after division, namely 001 (cluster 1, 2), 111 (cluster 3, 5), 000 (cluster 4), 010 (cluster 6), 011 (cluster 7) and 100 (cluster 8). Empathy, interpersonal trust, interpersonal distress and demographic variables also have a certain degree of correlation. There is a significant positive correlation between gender and interpersonal trust, and the interpersonal trust of girls is higher than that of boys. It is negatively related to the recognition of others and consistent understanding. The empathy recognition and understanding ability of boys is generally better than girls. Grades are negatively correlated with empathy ability, empathy recognition, and coping [35]. In the sample surveyed, with the gradual improvement of grades, the overall empathy ability of youth and empathy recognition, understanding, and coping dimensions all showed a downward trend. Major is



**TABLE 2. Multi-indicator RFM segmentation results.**

Clustering	Factor 1	Factor 2	Factor 3	Quantity	Proportion%	Type
1	-1.037	-4.214	25.308	1	0.02	001(1)
2	-1.085	-5.657	39.602	1	0.02	001(1)
3	0.6702	2.852	0.0846	227	4.24	111(7)
4	-0.5124	-0.497	-0.128	2570	47.99	000(0)
5	0.6924	0.25	9.9656	13	0.24	111(7)
6	-0.3458	0.625	-0.024	1372	25.62	010(2)
7	-0.1142	0.548	2.9686	106	1.98	011(3)

negatively correlated with positive coping, and positively correlated with reverse comprehension and reverse coping, indicating that youth of literature and history have higher empathy and positive empathy coping ability than science and engineering, but reverse comprehension and reverse coping ability are higher than science and engineering to be low. The only child has a negative correlation with empathy ability, empathy recognition, and identification of others, that is, the only child has a higher empathy ability and recognition dimension. In addition, the source area is only positively related to interpersonal distress, that is, the interpersonal distress of rural youth is the most serious, followed by cities and towns, and cities. Student cadres are not related to empathy, interpersonal trust, and interpersonal distress. In summary, the empathy ability of youth has significant differences in demographic variables, and has a significant correlation with the interpersonal trust and interpersonal distress of youth, and can use six second-order factors to interpersonal trust and interpersonal distress of youth Have a significant predictive effect.

## V. CONCLUSION

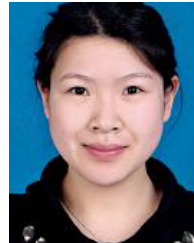
Many researches have been conducted on the factors that affect empathy. Through the existing empirical research, it is found that the individual differences affecting empathy and related reactions can be attributed to genetic, family, social environment and other factors. Based on the conclusions of the research on the empathy ability and interpersonal relationship of youth, this article combines graphic image and data mining technology to construct a new empathy ability analysis model. First, the theoretical principles of bad image recognition technology and their application in the evaluation of empathy ability are expounded. Then, based on image recognition and fusion particle swarm optimization method, the classification of bad images was studied. Finally, on the basis of image classification, data envelopment method is used to grade the young people's empathy ability. The actual case analysis and performance test results illustrate the superiority of the implemented image classification and empathy evaluation method based on image recognition and data mining. Among the surveyed youths, girls' empathy recognition ability and interpersonal trust are higher than that of boys. This conclusion is similar to the relevant research results at

home and abroad, confirming Eisenberg's "gender is one of the factors that cause empathy individual differences. And women have a higher emotional component of empathy than men." This conclusion reveals that college educators should pay attention to the difference between boys and girls in fostering empathy and interpersonal trust. It is hoped that it will have certain practical educational significance for the empathy ability and interpersonal attention of youth.

## REFERENCES

- [1] Ö. N. Yalçın, "Empathy framework for embodied conversational agents," *Cognit. Syst. Res.*, vol. 59, no. 12, pp. 123–132, Jan. 2020.
- [2] W. Xing, N. Deng, B. Xin, Y. Chen, and Z. Zhang, "Investigation of a novel automatic micro image-based method for the recognition of animal fibers based on wavelet and Markov random field," *Micron*, vol. 33, no. 15, pp. 142–153, Jun. 2019.
- [3] D. Vertsberger, L. Abramson, and A. Knafo-Noam, "The longitudinal Israeli study of twins (LIST) reaches adolescence: Genetic and environmental pathways to social, personality and moral development," *Twin Res. Hum. Genet.*, vol. 22, no. 6, pp. 567–571, Dec. 2019.
- [4] N. M. Thompson, A. Uusberg, and B. Chakrabarti, "Empathy and emotion regulation: An integrative account," *Prog. Brain Res.*, vol. 247, no. 32, pp. 134–143, Aug. 2019.
- [5] S. W. Teng, R. M. N. Sadat, and G. Lu, "Effective and efficient contour-based corner detectors," *Pattern Recognit.*, vol. 48, no. 7, pp. 2185–2197, Jul. 2015.
- [6] S. L. Jones et al., "The sepsis early recognition and response initiative (SERRI). Critical elements of a school report to parents on body mass index. Expanded HIV testing coverage is associated with decreases in late HIV diagnoses. Using PGD to reduce surgical infect," *Life Sci.*, vol. 30, no. 16, pp. 1379–1384, Jan. 2016.
- [7] J. P. Santos, C. Cremona, and P. Silveira, "Early damage detection based on pattern recognition and data fusion," *J. Struct. Eng.*, vol. 143, no. 2, pp. 162–171, Jun. 2017.
- [8] K. Nilsuwan, S. Benjakul, and T. Prodpran, "Physical/thermal properties and heat seal ability of bilayer films based on fish gelatin and poly(lactic acid)," *Food Hydrocolloids*, vol. 77, pp. 248–256, Apr. 2018.
- [9] K. Martijn, P. Merel, and H. N. Van, "F90. Emotion processing and white matter structural connectivity in schizophrenia," *Schizophrenia Bull.*, vol. 12, no. 24, pp. 167–173 Jun. 2019.
- [10] M. Golparvar-Fard, V. Balali, and J. M. de la Garza, "Segmentation and recognition of highway assets using image-based 3D point clouds and semantic texton forests," *J. Comput. Civil Eng.*, vol. 15, no. 31, pp. 256–263, Jan. 2015.
- [11] M. E. Kret, J. J. Stekelenburg, and K. Roelofs, "From face to hand: Attentional bias towards expressive hands in social anxiety," *Biol. Psychol.*, vol. 56, no. 13, pp. 123–134, Nov. 2015.
- [12] K. H. Kim, S. Lee, J. B. Shim, K. H. Chang, D. S. Yang, W. S. Yoon, Y. J. Park, C. Y. Kim, and Y. J. Cao, "A text-based data mining and toxicity prediction modeling system for a clinical decision support in radiation oncology: A preliminary study," *J. Korean Phys. Soc.*, vol. 71, no. 4, pp. 231–237, Aug. 2017.

- [13] M. Kathleen, V. Sophia, and K. Erich, "T221. Causal relationships between empathy, motivation and functional outcomes in early schizophrenia identified with causal discovery modeling," *Schizophrenia Bull.*, vol. 32, no. 7, pp. 134–143, Jan. 2020.
- [14] A. Karakilic, S. Kizildag, S. Kandis, G. Guvendi, B. Koc, G. B. Camsari, U. M. Camsari, M. Ates, S. G. Arda, and N. Uysal, "The effects of acute foot shock stress on empathy levels in rats," *Behavioural Brain Res.*, vol. 349, pp. 31–36, Sep. 2018.
- [15] M. B. Horzum, T. Ayas, C. Randler, and B. Düşünceli, "The effects of empathy and circadian preference on cyberbullying of adolescents in turkey," *Biol. Rhythm Res.*, pp. 1–14, Apr. 2019, doi: 10.1080/09291016.2019.1603839.
- [16] S. Hargreaves, P. Bath, and J. Ellis, "Sharing and empathy in digital spaces: Qualitative study of online health forums for breast cancer and motor neuron disease (amyotrophic lateral sclerosis)," *J. Med. Internet Res.*, vol. 20, no. 6, pp. 112–123, Aug. 2018.
- [17] J. K. Elrod and J. L. Fortenberry, "Am i seeing things through the eyes of patients? An exercise in bolstering patient attentiveness and empathy," *BMC Health Services Res.*, vol. 18, no. S3, pp. 402–412, Dec. 2018.
- [18] Y. Cao, X. Kou, and J. Du, "Research of a sub-assembly recognition algorithm based on a weighted undirected connection graph and an adjacency incidence matrix," *Basic Clin. Pharmacol. Toxicol.*, vol. 12, no. 9, pp. 233–241, Aug. 2016.
- [19] Z. Dong, Z. Zhou, Z. Li, C. Liu, P. Huang, L. Liu, X. Liu, and J. Kang, "Convolutional neural networks based on RRAM devices for image recognition and online learning tasks," *IEEE Trans. Electron Devices*, vol. 66, no. 1, pp. 793–801, Jan. 2019.
- [20] C. Detotto, D. Paolini, and J. D. Tena, "Do managerial skills matter? An analysis of the impact of managerial features on performance for Italian football," *J. Oper. Res. Soc.*, vol. 69, no. 2, pp. 270–282, Feb. 2018.
- [21] L. Deng, Y. Wang, Z. Han, and R. Yu, "Research on insect pest image detection and recognition based on bio-inspired methods," *Biosyst. Eng.*, vol. 169, pp. 139–148, May 2018.
- [22] J. Chen, L. Feng, and G. Sheng, "Research on pattern recognition of partial discharge based on KNN and MSR," *Electr. Eng.*, vol. 54, no. 32, pp. 134–143, Jun. 2018.
- [23] M. O. Bernardo, A. R. D. A. Lima, and M. J. Costa, "Investigating the relation between self-assessment and patients' assessments of physicians-in-training empathy: A multicentric, observational, cross-sectional study in three teaching hospitals in Brazil," *BMJ Open*, vol. 9, no. 6, pp. 412–423, Apr. 2019.
- [24] C. Banzhaf, F. Hoffmann, P. Kanske, Y. Fan, H. Walter, S. Spengler, S. Schreiter, and F. Bormpohl, "Interacting and dissociable effects of alexithymia and depression on empathy," *Psychiatry Res.*, vol. 43, no. 3, pp. 234–243, Jan. 2018.
- [25] B. Amir, A. Insaf, and B. Abdelhani, "Experiments and improvements of ear recognition based on local texture descriptors," *Opt. Eng.*, vol. 20, no. 6, pp. 212–223, Apr. 2017.
- [26] W. Yu, L. Jiao, and D. Liu, "Sports motion recognition using MCMR features based on interclass symbolic distance," *Int. J. Distrib. Sensor Netw.*, vol. 2016, no. 3, pp. 1–15, Nov. 2016.
- [27] G. Li and C. Zhang, "Automatic detection technology of sports athletes based on image recognition technology," *EURASIP J. Image Video Process.*, vol. 2019, no. 1, pp. 1–9, Dec. 2019.
- [28] K. H. Hu, F. H. Chen, and T. Gwo-Hsiung, "Evaluating the improvement of sustainability of sports industry policy based on MADM," *Sustainability*, vol. 8, no. 7, pp. 606–615, Jan. 2016.
- [29] F. M. Verhoeven and K. M. Newell, "Coordination and control of posture and ball release in basketball free-throw shooting," *Hum. Movement Sci.*, vol. 49, pp. 216–224, Oct. 2016.
- [30] A. Delextrat and A. Martinez, "Small-sided game training improves aerobic capacity and technical skills in basketball players," *Int. J. Sports Med.*, vol. 35, no. 05, pp. 385–391, Oct. 2013.
- [31] H. Liang, J. Zou, K. Zuo, and M. J. Khan, "An improved genetic algorithm optimization fuzzy controller applied to the wellhead back pressure control system," *Mech. Syst. Signal Process.*, vol. 142, no. 7, Aug. 2020, Art. no. 106708.
- [32] J. Hu, Y. Sun, G. Li, G. Jiang, and B. Tao, "Probability analysis for grasp planning facing the field of medical robotics," *Meas.*, vol. 69, no. 2, pp. 270–282, Nov. 2019.
- [33] F. Hu and G. Wu, "Distributed error correction of EKF algorithm in multi-sensor fusion localization model," *IEEE Access*, vol. 8, pp. 93211–93218, Jan. 2020.
- [34] D. Jiang, G. Li, Y. Sun, J. Kong, and B. Tao, "Gesture recognition based on skeletonization algorithm and CNN with ASL database," *Multimedia Tools Appl.*, vol. 12, no. 24, pp. 167–173 Jun. 2018.
- [35] H. Liang, J. Zou, Z. Li, M. J. Khan, and Y. Lu, "Dynamic evaluation of drilling leakage risk based on fuzzy theory and PSO-SVR algorithm," *Future Gener. Comput. Syst.*, vol. 95, pp. 454–466, Jun. 2019.



**GONG YING** was born in Linqing, Shandong, in November 1985. She is currently pursuing the Ph.D. degree major in clinical psychology with The Third Xiangya Hospital of Central South University.



**YAN JIN** was born in September 1972. He is currently a Professor of social psychology and nursing and the Director of the Nursing Department, The Third Xiangya Hospital of Central South University.



**CUIYU BAO** was born in Wuhan, Hubei, in August 1969. She is currently a Professor of cardio-cerebrovascular disease and the Director of the Scientific Research Office, Hubei University of Science and Technology.



**ZHAORONG CHEN** was born in Changsha, Hunan, in February 1989. She is currently pursuing the Ph.D. degree major in clinical psychology with The Third Xiangya Hospital of Central South University.



**YUNLONG DENG** was born in Liuyang, Hunan. He is currently a Professor of Clinical Psychology and the Vice-President of The Third Xiangya Hospital of Central South University. He is also a Professor with the Clinical Psychology Department.