Research on the Willingness of Wuhan Residents to Use Face Recognition in the Era of Artificial Intelligence
—— Based on Principal Component Analysis

Yihong Xu*[1,a], Yifan Liu[1,b] and Rongxiao Yu[2,c]

[1]*Business school, Jianghan University, Wuhan, Hubei, China*
[2] *Business school, Jianghan University, Wuhan, Hubei, China*
*[a]*alanyhxu@jhun.edu.cn*, [b]*1329326804@qq.com*, [c]*714544173@qq.com*
Corresponding Author: Yihong Xu

ABSTRACT
Based on the innovative development of artificial intelligence, face recognition has become one of the most popular topics today, but the technology also has many security risks.The research of residents' willingness to use face recognition products has practical and theoretical significance for the development of face recognition field.Taking Wuhan residents as the research object, 1006 samples were sampled by combining hierarchical sampling and three-stage sampling, and the principal component analysis method was used to dig deeply to analyze the relationship between the basic information of residents and the willingness to use face recognition products.The results show that the user characteristics of face recognition products are significantly different and the willingness to use is affected by many factors.

Keywords: *Artificial intelligence, Face recognition, Willingness to use, Principal component analysis*

1. INTRODUCTION

As the Internet promotes digital life and the further improvement of computer computing power, artificial intelligence (AI) technology has become an important driving force for a new round of economic development and the core competitiveness of social innovation. Face recognition is a kind of biometric identification technology based on human facial feature information. As the fastest growing branch of artificial intelligence, it is widely used in various industries and scenarios, such as banks, airports, criminal investigations, access control, attendance, etc. The use of face for intelligent identity recognition has become an important way of personal identity authentication, which to a certain extent improves the level of public security and intelligent management in the city. However, with the rapid development of face recognition technology, face recognition technology also faces challenges. In particular, residents are worried about personal information being leaked and privacy rights being violated.

According to the "Public Research Report on Face Recognition Applications (2020)", 60% of the respondents believe that face recognition technology has a tendency to be abused, and 30% of the respondents said that they have suffered privacy or property losses due to the leakage and abuse of face information; The most unacceptable scenario for the respondents was "the mall uses face recognition technology to collect customers' behavior and purchasing methods" (42.68%), followed by "colleges and universities use face recognition technology to collect posture" (28.36%), and "applications such as face changing, beauty makeup, personality judgment, and health status prediction based on face image analysis" (19.01%). Relatively speaking, respondents are more receptive to face recognition applications based on security scenarios, such as public safety cameras and red-light-running recording systems. From AI video surveillance into classrooms, face-changing apps, to consumers suing "face-scanning" businesses, the security risks brought by face recognition commercial applications are attracting attention.

A. El-Hashash et al. (Eds.): IEIT 2022, ACSR 100, pp. 607–614, 2023.
https://doi.org/10.2991/978-94-6463-058-9_97

As an emerging technology, face recognition should be developed. However, while face recognition products bring convenience to residents, the security risks brought by them cannot be underestimated. Therefore, it is particularly important to explore the willingness to use face recognition and its development countermeasures.

2. LITERATURE REVIEW

2.1. Foreign Research

Experts and scholars in some developed countries have carried out research and practice on improving the biometric recognition performance of face recognition technology. Ross (2003) used geometric features such as faces and fingerprints to improve the performance of biometric systems; Jain et al. (2005) believed that multi-pattern recognition systems using user-specific weights performed better.

With the widespread application of artificial intelligence, face recognition technology has also continued to improve. Taigman (2014) first proposed a face recognition model using a convolutional neural network structure, which achieved 97.35% accuracy on the face dataset Labeled Faces in the Wild (LFW). Schroff (2015) proposed the FaceNet system and adopted a novel online triplet mining method, which achieved 95.12% and 99.63% recognition accuracy on YouTube Faces DB and LFW face datasets, respectively.

2.2. Research State in China

The domestic face recognition research is relatively late. Some universities and research institutes have carried out extensive research on face recognition, and have made important breakthroughs in theoretical research and industrialization. They have independently developed a number of face recognition technologies, and some technical indicators have reached the international advanced level.

With the popularity of face recognition applications, the security of face recognition systems has become a social focus. Wu Jianfeng (2018) found that users' willingness to use the technology will be affected by public opinion; Shi Jiayou (2021) found that the risks of face recognition technology will lead to personal, property and psychological hidden dangers of users.

In response to the current face recognition verification database security problems and prosthetic attack problems, Wang Yanqiu (2021) and others proposed a big data face recognition algorithm, and designed the system architecture and implementation process of the big data face recognition algorithm; Shen Xi (2021) et al. proposed a new secure face recognition solution. Through a special secure face acquisition module, the face encryption private key is used to ensure the security of face biometrics.

To sum up, the current face recognition technology has been able to achieve better recognition results in specific scenarios. However, the high demand for privacy protection and secure identification in real life will objectively affect users' willingness to use face recognition products, and face recognition technology still faces many difficulties and challenges.

3. RESEARCH DESIGN

According to the needs of the research purpose, a questionnaire survey method was adopted to investigate the willingness of face recognition use among Wuhan citizens from different online and offline methods.

A three-stage sampling method was adopted to divide Wuhan into the main urban area and the new urban area. The streets of each district were selected from each urban area, and the resident residents were selected from each street. Under the estimation of the overall parameter P, the best results were obtained.

Sample size:

$$n_0 = \frac{t^2 PQ / d^2}{1 + \frac{1}{N}\left(\frac{t^2 PQ}{d^2} - 1\right)} \approx 591.5752 \tag{1}$$

Among them, n is the number of overall units, that is, the total number of permanent residents in Wuhan, which is 11,212,000. When the confidence level is 95%, t=1.96, d=0.04, and the pre-investigation result is p=0.56.

Since deff is difficult to calculate in the actual process, after the results of the pre-investigation combined with many references, this paper determines , so the adjusted optimal sample size n is obtained:

$$n = n_0 \times deff \approx 1006 \tag{2}$$

4. DATA ANALYSIS

4.1. Reliability and Validity Test

4.1.1. Reliability Test

According to the survey characteristics and specific needs, the α reliability coefficient method is used to estimate the consistency of the survey test. The formula is:

$$\alpha = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum S_i^2}{S_x^2}\right) \tag{3}$$

Where k is the total number of items in the survey, S_i^2 is the intra-item variance of the score for the ith item, and S_x^2 is the variance of the total score of all items.

Quantify the degree problem in the questionnaire, and divide it into two aspects: the analysis of Wuhan residents of face recognition products and the analysis of improvement of face recognition products. The Cronbach reliability analysis is carried out. The test results are as follows:

Table 1: Reliability Test.

project	Cronbach coefficient	number of items	reliability evaluation
Analysis of Wuhan Residents of Facial Recognition Products	0.877	14	The scale has good reliability
Improvement analysis of face recognition products	0.764	3	The scale has good reliability

Table 1 shows that the Cronbach coefficient of each item is higher than 0.8, and the overall reliability coefficient of the scale is 0.898, which shows the rationality of the questionnaire structure and question design.

4.1.2. Validity Test

Table 2: Validity Test.

KMO measures sampling suitability		0.785
Bartlett's sphericity test	approximate chi-square	9765.740
	degrees of freedom	136
	salience	0.000

It can be seen from the test results that the KMO value of 0.785 is between 0.7-0.8, indicating that the validity is acceptable, and the p-value of the Bartlett sphericity test is less than 0.05, which is statistically significant.

In conclusion, the questionnaire has passed the validity test and can be analyzed by factor.

4.1.3. Rotation Circular Matrix

Table 3: Rotation Circular Matrix.

Rotation element matrix					
	element				
	1	2	3	4	5
experience feelings	0.666	0.381	-0.016	0.198	0.244
Difficulty to perceive	0.784	-0.027	0.007	0.107	0.103
degree of liking	0.642	0.315	-0.009	0.356	0.235
Willingness to recommend	0.701	0.383	0.121	0.346	0.049
degree of recognition	0.699	0.048	0.312	-0.161	-0.103
Terms of Use Concerns	-0.055	0.419	0.724	0.250	-0.133
level of security concerns	-0.092	0.192	0.624	0.233	-0.160
degree of dependence	0.452	-0.002	0.661	-0.177	0.279
trust level	0.456	0.209	0.577	0.114	0.306
personal information leakage	0.293	0.824	0.263	0.091	0.040
Face unlock failed	0.077	0.335	0.075	0.295	0.744
Identify errors	0.129	0.051	-0.010	0.147	0.873
safety	0.142	0.862	0.173	0.286	0.106
accuracy	0.105	0.715	0.220	0.052	0.382
Convenience	0.264	0.128	0.552	0.260	0.333
privacy protection	0.148	0.173	0.226	0.780	0.226
propaganda power	0.163	0.167	0.174	0.849	0.202

It can be seen intuitively from the rotating element matrix that the polarization effect of each column element is obvious.

To sum up, through reliability analysis and validity analysis, this questionnaire has certain credibility and reference significance in content and testing. It can be filled out differently for different objects, and the questionnaire design is relatively successful.

4.2. Descriptive Statistical Analysis

The descriptive statistical analysis of the current situation of face recognition cognition of Wuhan residents was carried out by SPSSAU software, and the following results were obtained:

(1) Basic information of the subjects. Among the surveyed residents, 9.86% are under 18 years old, 57.75% are 19-29 years old, 7.04% are 30-39 years old, 16.90% are 40-49 years old, and 8.45% are over 50 years old. , young people and middle-aged people make up the majority, which is in line with social reality.

(2) Subjects' willingness to use face recognition products. According to the participation of the respondents and the realistic background of residents using face recognition, the group is divided by "whether face recognition has been used". The data shows that among the valid samples, the proportions of those who have used face recognition and those who have not used face recognition are 84.51% and 15.49%, respectively.

(3) The degree of residents' perception. Most of the residents' perceptions of use are concentrated in the good and general levels, reaching 40.00% and 41.67%; the degree of difficulty and ease of use by residents is mostly concentrated in the good and good levels, reaching 33.33% and 45%; the use of human faces by residents The favorite degree of recognition is concentrated in the general degree, reaching 53.33%.

(4) Willingness to promote face recognition. 15% of the respondents are very willing to promote, 21.67% of the respondents are willing, 38.33% of the respondents expressed a general state, and 25% of the respondents expressed their unwilling intention. This reflects that face recognition technology has given residents a certain convenience in the current market. Maybe there are still some imperfect aspects, but there is still hope for future development.

(5) Concerns brought by face recognition. The rise of face recognition technology has also brought a series of hazards to residents, which may result in the leakage of personal information due to the illegal acquisition of biometrics. Under this circumstance, 33.33% of the residents surveyed were very worried, and 45.00% were worried. Regarding the level of concern about encountering unreasonable (fraudulent) terms in the process of using face recognition products, the surveyed residents are very worried about 30.00%, and the worry level is also as high as 38.33%. It can be seen that face recognition technology is still in an immature stage, and some hidden problems still need to be solved.

(6) Residents' trust in face recognition. Residents' trust in face recognition payment providers and the efforts of payment providers is mostly concentrated in the general level of trust, accounting for 40.00% and 43.33% respectively, and the rest of the distribution of trust and less trust is almost evenly divided. It shows that Wuhan residents have a certain degree of trust in face recognition technology, and they have not rejected face recognition technology because of some existing problems.

(7) Residents use face recognition frequency. 21.67% of the residents never use face recognition every week, 43.33% of the residents use it 1-4 times, 13.33% of the residents have used it 5-9 times, 6.67% of the residents have used it 10-15 times, 15 15% of the residents are the second or above. Taken together, the majority use face recognition technology 1-4 every week. It can be seen from this that many people still contact and use the face recognition system in the daily life of residents, and they are not completely excluded because of some existing hidden dangers.

(8) How to use face recognition. Alipay payment accounted for 53.33%, mobile phone lock screen accounted for 43.33%, and face recognition accounted for 38.33%. In addition to some aspects of the investigation, some residents said that they have also used face recognition technology in bank transactions, account registration, etc., and face recognition technology has completely penetrated into every corner of our lives.

(9) Satisfaction with face recognition products. Residents' recognition of the accuracy of face recognition reached 50%. In general, the satisfaction with the security and publicity of face recognition reached 60% and 62%, reflecting the residents' attitude towards face recognition. Technology has a certain recognition.

4.3. Empirical Analysis

After substituting the data from the questionnaire into SPSSAU, the following rubble diagram and variance interpretation table were obtained through principal component analysis.

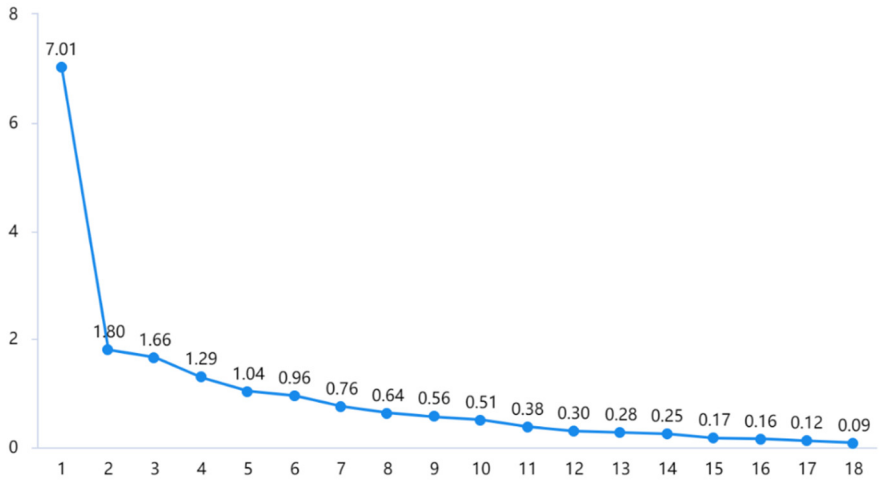

Figure 1: gravel diagram.

Table 4: Rotation Circular Matrix.

Element	Characteristic Root			Principal Component Extraction		
	Characteristic root	Variance explained rate %	Accumulation %	Characteristic root	Variance explained rate %	Accumulation %
1	7.015	38.970	38.970	7.015	38.970	38.970
2	1.802	10.013	48.983	1.802	10.013	48.983
3	1.664	9.242	58.225	1.664	9.242	58.225
4	1.292	7.176	65.401	1.292	7.176	65.401
5	1.040	5.779	71.180	1.040	5.779	71.180
6	0.957	5.314	76.494	-	-	-
7	0.760	4.220	80.714	-	-	-
8	0.639	3.549	84.263	-	-	-
9	0.565	3.138	87.401	-	-	-
10	0.512	2.847	90.247	-	-	-
11	0.383	2.128	92.375	-	-	-
12	0.300	1.668	94.043	-	-	-
13	0.277	1.539	95.582	-	-	-
14	0.252	1.401	96.983	-	-	-
15	0.173	0.963	97.947	-	-	-
16	0.160	0.891	98.838	-	-	-
17	0.122	0.676	99.514	-	-	-
18	0.088	0.486	100.000	-	-	-

Combining the above gravel diagram and variance interpretation table, we can see that the first five components have basically covered most of the variance, and these five principal components cover more than 71.180% of the information in the 18 questions.

After that, the data was processed with SPSSAU to obtain the following load coefficient table, which is shown in Table 5 below.

Table 5: Load Factor Table.

Question	Load Factor					Common Factor Variance
	1	2	3	4	5	
Q7	0.690	0.368	-0.079	-0.312	-0.019	0.716
Q8	0.478	0.614	0.078	-0.071	0.162	0.642
Q9	0.718	0.319	-0.141	-0.203	0.098	0.688
Q10	0.765	0.276	0.077	-0.287	0.141	0.770
Q11	0.423	0.458	0.486	0.047	-0.030	0.628
Q12	0.545	-0.548	0.424	0.064	0.097	0.790

Q13	0.339	-0.446	0.372	0.047	0.244	0.515
Q14	0.551	0.210	0.401	0.472	-0.109	0.743
Q15	0.739	0.086	0.223	0.241	-0.021	0.662
Q17	0.580	-0.062	-0.007	0.505	-0.123	0.611
Q18.1	0.651	-0.062	0.083	0.250	0.124	0.513
Q18.2	0.651	-0.223	-0.283	0.029	0.437	0.745
Q18.3	0.647	-0.210	-0.323	-0.047	0.509	0.828
Q19.1	0.734	-0.197	0.188	-0.378	-0.306	0.849
Q19.2	0.639	-0.092	-0.524	0.222	-0.203	0.781
Q19.3	0.463	0.152	-0.592	0.319	-0.174	0.720
Q19.4	0.739	-0.349	-0.022	-0.393	-0.238	0.879
Q19.5	0.678	-0.240	-0.088	-0.089	-0.447	0.733

Remarks: If the numbers in the table are colored: blue indicates that the absolute value of the load factor is greater than 0.4, and red indicates that the common degree (common factor variance) is less than 0.4.

Table 5 shows that the common degree values corresponding to all research items are higher than 0.4, which means that there is a strong correlation between the question and the principal component, and the principal component can effectively extract information.

Let the questions of the questionnaire be x1~x18 respectively, and then according to the component score matrix obtained by processing (Table 6).

Table 6: Component Score Coefficient Matrix.

Question	Element				
	1	2	3	4	5
Q7	0.098	0.204	-0.047	-0.242	-0.018
Q8	0.068	0.340	0.047	-0.055	0.156
Q9	0.102	0.177	-0.084	-0.157	0.094
Q10	0.109	0.153	0.047	-0.222	0.136
Q11	0.060	0.254	0.292	0.036	-0.029
Q12	0.078	-0.304	0.255	0.050	0.094
Q13	0.048	-0.248	0.224	0.036	0.235
Q14	0.079	0.116	0.241	0.365	-0.105
Q15	0.105	0.048	0.134	0.186	-0.020
Q17	0.083	-0.035	-0.004	0.391	-0.118
Q18.1	0.093	-0.034	0.050	0.194	0.119
Q18.2	0.093	-0.124	-0.170	0.022	0.420
Q18.3	0.092	-0.117	-0.194	-0.037	0.489
Q19.1	0.105	-0.109	0.113	-0.292	-0.294
Q19.2	0.091	-0.051	-0.315	0.172	-0.195
Q19.3	0.066	0.084	-0.356	0.247	-0.167
Q19.4	0.105	-0.193	-0.013	-0.304	-0.229
Q19.5	0.097	-0.133	-0.053	-0.069	-0.429

According to the above table, the following expression is obtained:

$$y_1 = 0.098 * x_1 + 0.068 * x_2 + 0.102 * x_3 + 0.109 * x_4 + 0.060 * x_5 + 0.078 * x_6 + 0.048 * x_7 + 0.079 * x_8 + 0.105 * x_9 + 0.083 * x_{10} + 0.093 * x_{11} + 0.093 * x_{12} + 0.092 * x_{13} + 0.105 * x_{14} + 0.091 * x_{15} + 0.066 * x_{16} + 0.105 * x_{17} + 0.097 * x_{18} \quad (4)$$

$$y_2 = 0.204 * x_1 + 0.340 * x_2 + 0.177 * x_3 + 0.153 * x_4 + 0.254 * x_5 - 0.304 * x_6 - 0.248 * x_7 + 0.116 * x_8 + 0.048 * x_9 - 0.035 * x_{10} - 0.034 * x_{11} - 0.124 * x_{12} - 0.117 * x_{13} - 0.109 * x_{14} - 0.051 * x_{15} + 0.084 * x_{16} - 0.193 * x_{17} - 0.133 * x_{18} \quad (5)$$

$$y_3 = -0.047 * x_1 + 0.047 * x_2 - 0.084 * x_3 + 0.047 * x_4 + 0.292 * x_5 + 0.255 * x_6 + 0.224 * x_7 + 0.241 * x_8 + 0.134 * \quad (6)$$

$$x_9 - 0.004 * x_{10} + 0.050 * x_{11} - 0.170 * x_{12} - 0.194 * x_{13} + 0.113 * x_{14} - 0.315 * x_{15} - 0.356 * x_{16} - 0.013 * x_{17} - 0.053 * x_{18}$$

$$y_4 = -0.242 * x_1 - 0.055 * x_2 - 0.157 * x_3 - 0.222 * x_4 + 0.036 * x_5 + 0.050 * x_6 + 0.036 * x_7 + 0.365 * x_8 + 0.186 * x_9 + 0.391 * x_{10} + 0.194 * x_{11} + 0.022 * x_{12} - 0.037 * x_{13} - 0.292 * x_{14} + 0.172 * x_{15} + 0.247 * x_{16} - 0.304 * x_{17} - 0.069 * x_{18} \quad (7)$$

$$y_5 = -0.018 * x_1 + 0.156 * x_2 + 0.094 * x_3 + 0.136 * x_4 - 0.029 * x_5 + 0.094 * x_6 + 0.235 * x_7 - 0.105 * x_8 - 0.020 * x_9 - 0.118 * x_{10} + 0.119 * x_{11} + 0.420 * x_{12} + 0.489 * x_{13} - 0.294 * x_{14} - 0.195 * x_{15} - 0.167 * x_{16} - 0.229 * x_{17} - 0.429 * x_{18} \quad (8)$$

The proportion of the variance explained rate in the cumulative contribution rate is regarded as the weight of each principal component, and the comprehensive score calculation formula is obtained：

$$S = 0.547 * y_1 + 0.141 * y_2 + 0.130 * y_3 + 0.101 * y_4 + 0.081 * y_5 \quad (9)$$

Based on the above formula, the collected sample data is evaluated for willingness, and the comprehensive evaluation score of 1065 respondents is calculated. Using SPSSAU for data statistics and MATLAB for drawing, it can be known that the willingness distribution map basically conforms to the normal distribution, and the willingness to the vast majority of the surveyed population use face recognition.

5. CONCLUSIONS

Most residents have used face recognition products, and their characteristics are significantly different. Older women, with a high level of education, good economic conditions, and average working environment, are more willing to use face recognition products (18.3%); younger women, with a low level of education, better economic conditions, and good working environment, very reluctant to use face recognition products (4.2%); young men, with low educational level, poor economic conditions and good working environment, are reluctant to use face recognition products (18.31%); older men, with high educational level, average economic conditions and poor working environment, are very willing to use face recognition products (59.15%).

Wuhan residents' willingness to use face recognition products is affected by many factors. Usefulness has a positive effect on the willingness to use face recognition products; risk awareness has a negative effect; the surrounding environment is affected by other factors, which can affect the willingness to use; publicity efforts have a negative impact on usefulness to a certain extent.

REFERENCES

[1] Jain A, Nandakumar K, Ross A. (2005). Score normalization in multimodal biometric systems. J. Pattern Recognition, 38(12): 2270-2285.

[2] Krizhevsky A, Sutskever I, Hinton G E. (2017).ImageNet Classification with Deep Convolutional Neural Networks. J. Communications of the Acm, 60(6): 84-90.

[3] Li S Z, Chu R F, Liao S C, et al.(2007). Illumination invariant face recognition using near-infrared images. J. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(4): 627-639.

[4] Qin Y B, Tian C W. (2018). Weighted Feature Space Representation with Kernel for Image Classification. J. Arabian Journal for Science and Engineering, 43(12): 7113-7125.

[5] Ross A, Jain A. (2003).Information fusion in biometrics. J. Pattern Recognition Letters, 24(13): 2115-2125.

[6] Simonyan K, Zisserman A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. C. Proceedings of the International Conference on Learning Representations.

[7] Schroff, F., Kalenichenko, D. and Philbin, J. (2015) FaceNet: A Unified Embedding for Face Recognition and Clustering. IEEE Conference on Computer Vision & Pattern Recognition, Boston, 7-12 June 2015, 815-823.

[8] Taigman, Y., Yang, M., Ranzato, M.A., et al. (2014) Deepface: Closing the Gap to Human-Level Performance in Face Verification. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1701-1708.

[9] Xu Y, Fei L K, Zhang D. (2015). Combining Left and Right Palmprint Images for More Accurate Personal Identification. J. Ieee Transactions on Image Processing, 24(2): 11.

[10] Zhang Weijia.(2021). Research on the Application and Dilemma of Face Recognition in Investigation .J. Light Industry Science and Technology,37(04):114-115.