Received 7 September 2022, accepted 22 September 2022, date of publication 10 October 2022, date of current version 20 October 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3213675

RESEARCH ARTICLE

Image Recognition and Analysis: A Complex Network-Based Approach

ZHUANG MA AND GUANGDONG HUANG
College of Mathematics and Physics, China University of Geosciences, Beijing 100083, China
Corresponding author: Guangdong Huang (gdhuang@cugb.edu.cn)

ABSTRACT The placement and order of image pixels play a significant role in the accuracy of image recognition in current algorithms. Complex networks will significantly reduce the impact of images on classification recognition accuracy when rotation, translation, and scaling occur. Complex networks' topological invariance has made it clear that using them to analyze image recognition will considerably increase image classification accuracy. However, most studies of complex networks for image classification have focused on individual networks, neglecting the combination of multiple networks. This paper proposes a new complex network classification method that combines complex networks and convolutional neural networks(CNN) to train classification using deep learning. We show that the method has high classification accuracy and distinct network features and compares well with a single complex network approach. In addition, to make the distribution of the degree histogram of the image more uniform and concentrated, the original formula for calculating the power value was optimized to reduce the influence of the radius parameter on the power value.

INDEX TERMS Complex networks, convolutional neural networks, degree histogram, image classification, image recognition.

I. INTRODUCTION

Image recognition is the process of processing and analyzing an image first, then describing and classifying it. There are many methods of image recognition, such as methods based on image key points, based on image texture recognition [1], based on Karhunen-Loeve(K-L) Transform [2], and based on geometric image features, based on image model recognition [3], [4], and based on image edge contour recognition [5]. Among them, keypoint-based refers to the keypoint extraction of images, such as extracting Harris corner points, Scale-invariant feature transform(SIFT), and structure construction based on the extracted key points to obtain the spatial location information of these critical points. In addition, the recognition classification of images can be identified by measuring the similarity between images, such as by extracting feature points of images using degree histograms [6].

With the continuous development of technology, complex networks have gradually attracted the attention of several researchers [7], [8]. Unlike neural networks, complex networks are based on a graph-theoretic structure and consist of vertices and connected edges, where the interconnections between vertices constitute directed or undirected edges. The corresponding connection rules were set in the connection process to simplify the complex network. Therefore, based on the complex network used to describe the image, it has the characteristics of better stability and robust noise immunity to efficiently complete image recognition.

For the first time, complex networks have been applied in the field of image recognition, taking grayscale images as an example, using each location point of the image to construct a complex network model and simplify the network based on its pixel values and locations to obtain topological features for recognition. The complex network maintains topological invariance even when the image is physically changed, which is why a large training dataset is not required compared to deep learning [5]. In contrast, pixel points are used as vertices for the network construction. [8]. This study [9] used multiple sets of threshold parameters to extract multiple topological features for network construction, and the articles [10] built

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

on this still used pixel points as vertices by combining multiple image recognition methods to improve the recognition rate. However, the recognition speed was slower because of the large number of vertices. Moreover, the network structure in the identification process is more complex than that in the previous network structure. The literature no longer uses pixel points as vertices, the gray matrix of the image as the linking rule of a complex network, or the topology of structural balance to accomplish image recognition.

In addition, various methods can be used to extract critical points for complex network construction, such as the corner point extraction method. The literature [11] uses weight thresholds to perform dynamic evolution to achieve feature extraction of the entire image based on the obtained multi-group network. Of course, this method has significant limitations for disconnected subgraphs, which also adds difficulties to subsequent feature extraction. The literature proposes a new evolutionary approach, the USA, in a minimum spanning tree for evolutionary analysis. The article even adds pixel and distance thresholds to dynamically evolve and combine the matrix for feature extraction. However, because of the manual setting of thresholds, the optimal parameters cannot be selected, which affects the recognition results. Cout et al. [12] optimized the algorithm to establish the degree matrix to extract the features of the texture image and establish the average degree matrix of the complex network of the image. Gao et al. fused multiple texture features to achieve recognition analysis of an image. Zhou et al. [13] used the properties of nodes and connected edges to perform dynamic evolution, combined with mathematical methods such as multiscale wavelets to extract features.

This paper proposes a multi-network-based image recognition method that combines complex networks and neural networks to recognize images. The pixel points of an image are used to construct a complex network, analyze its static statistics and dynamic evolution process, and combine it with convolutional neural network training to improve image recognition accuracy. The remainder of this paper is organized as follows. Section 2 introduces the related research. Section 3 presents the model framework and experimental data. Section 4 presents the experiments and results. Section 5 discusses the results and concludes the study.

II. RELATED WORKS

A. COMPLEX NETWORKS

Complex networks are self-organizing, self-similar, attractive, small-world, scale-free networks with some or all properties. They are mainly used to explain the complexity of network phenomena, which is not only limited to the field of mathematics but can also be applied to, for example, the fields of economics, engineering, and biology. It can describe a wide range of systems and networks in society. We observe that the vertices of the units are different for different networks. For example, in a social network, each individual is the vertex of the complex network. Complex networks

TABLE 1. Bacon count of film actors (as of May 2020).

Number of Bacon	Number of actors
0	1
1	861
2	63509
3	240356
4	120927
5	9751
6	761
7	152
8	37
9	16

have some characteristics, such as small-worldness, and the complexity of the structure evolves into different structural features. A small-world phenomenon (also called six degrees of separation)refers to social relationships in which we go through others to meet strangers, on average, no more than six people. A psychology professor at Harvard University in the United States conducted a similar experiment. The chain letter experiment illustrates that it takes only up to five intermediate parties to contact two people who do not know each other. For example, Jackie Chan starred in Around the World in 80 Days, directed by American director Michele, and Luke Wilson was also present. In addition, Luke Wilson and Kevin Bacon co-starred in "My Dog Skip" in 2000, giving Chan a Bacon number of two.

As shown in Table 1, as of May 2020, more than 5 million actors and actresses were counted to obtain the results. As shown in the figure, the left column shows the number of Bacons in order, and the right column shows the number of actors corresponding to the number of Bacons. The average number of Bacons was calculated as only 3.18.

Among all the actors counted, some are more central than Kevin Bacon, meaning their values are lower if we consider them the center of the movie industry. For example, if we take Samuel L. Jackson as the center, Jackson's average = 2.997.

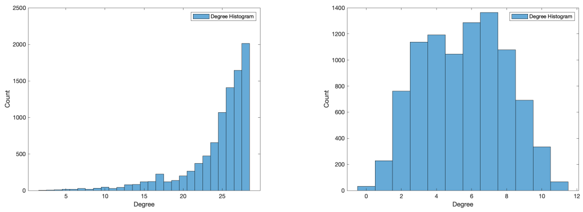
First, the image is defined as a two-dimensional matrix of pixels. For the gray image P, each pixel has an integer value p = 0,1,...L, representing the intensity of light in that pixel, where L is the largest pixel value. Set P(x,y) = p, where(x,y) is the pixel P(x,y) position coordinate. By considering each pixel of an image P as a vertex of a network N, the distance between two pixels and the corresponding vertices is defined as follows:

$$dij = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{1}$$

The following relation determines whether the two vertices have previously connected edges.

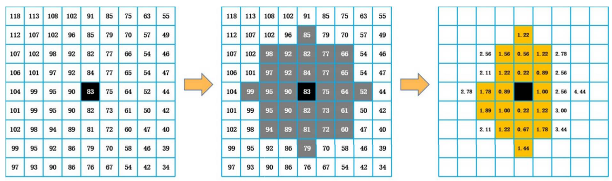
$$e(p_i, p_j) = \begin{cases} 1 & \text{if } d_{ij} \leq r \\ 0 & \text{if } d_{ij} \geq r \end{cases} \tag{2}$$

where $e(p_i, p_j) = 1$ denotes an edge connection between two vertex sums, and $e(p_i, p_j) = 0$ denotes no edge connection between the two vertex sums. Next, the weights of the connected edges are calculated. The literature [2] provides the



(a) Degree histogram under the original formula (b) Degree histogram under the modified formula

FIGURE 1. Degree histogram.



a.Image pixel values (partial) b.Connecting edges of vertices c.Weights of connected edges

FIGURE 2. CNN training progress.

following calculation method:

$$W(e(p_i, p_j)) = (x_i - x_j)^2 + (y_i - y_j)^2 + R^2 * |\frac{P_i(x_i, y_i) - P_j(x_j, y_j)}{255}| \quad (3)$$

To make the degree distribution of the vertices more uniform and concentrated, we defined the weights as follows: The modified effects are illustrated in Fig. 1.

$$W(e(p_i, p_j)) = (x_i - x_j)^2 + (y_i - y_j)^2 + |P_i(x_i, y_i) - P_j(x_j, y_j)| \quad (4)$$

However, not all the connected edges that satisfy the connection rules are retained. We want to provide a weight parameter to eliminate some connected edges, thus simplifying the network and reducing complexity. The connection relation is as follows: The connected edges are retained if the weight is less than the given parameter W. The connected edges were eliminated if the weight value exceeded the parameter W.

As shown in Figure 2, a 9 × 9 regions of an image is considered as an example. Here, the radius parameter R is given as 3, and the center point is connected to its vertices whose Euclidean distance is not greater than 3. The weights of the connected edges were calculated using a weighted formula. Here, the weight parameter W is given as 2, the connected edges with weights greater than two are eliminated, and only the connected edges with weights less than two are retained, thereby simplifying the complex network. The specific weights were calculated as follows:

$$W(e(p_i, p_j)) = \frac{(x_i - x_j)^2 + (y_i - y_j)^2 + |P_i(x_i, y_i) - P_j(x_j, y_j)|}{R^2} \quad (5)$$

TABLE 2. The effect of different activation functions and training algorithms on the correctness of the validation set.

diagboxFunctionMethod	sgdm(%)	adam(%)	rmsprop(%)
relu	93.40(1'09)	91.28(1'28)	91.49(1'25)
tanh	92.18(1'19)	90.65(1'28)	90.49(1'29)
sigmoid	92.27(1'25)	91.92(1'30)	91.35(1'31)

For example, the calculation of the weight between $P_{54} = 90$ and $P_{55} = 83$

$$W(e(p_i, p_j)) = \frac{(5-5)^2 + (5-4)^2 + |90 - 83|}{9} = 0.89 \quad (6)$$

B. CONVOLUTIONAL NEURAL NETWORKS

Among the more mature approaches to neural network image recognition, deep learning is currently used for image training recognition classification. In deep learning, the most commonly used neural networks are recurrent neural networks, convolutional neural networks, and other essential algorithms. The convolutional neural network (CNN) structure generally consists of a convolutional layer, sampling layer, and fully connected layer, and it also includes more than five hidden layers. Initially, when applied to image recognition, convolutional neural networks were not as productive as expected. It was not until 2012 when deep neural networks were built, improved algorithms and added the concept of weight decay to improve the computational power effectively. Convolutional neural networks generally do not require manual design or feature selection when recognizing images. They are all performed by neural network training and learning the required feature information. In addition, the number of parameters is taken in various ways, such as weight sharing or local connectivity, to reduce the computational complexity. Convolutional neural networks have been widely researched worldwide and have made significant progress in various fields, such as face recognition, digital recognition, and image recognition. This paper shows the superiority of complex networks in image recognition based on complex networks compared with convolutional neural networks.

The dataset used was NOTMNIST, similar to MNIST, with an image size of 28 × 28. However, compared to MNIST, NOTMNIST contains 10 categories of artistic typographic characters from A-J, with different shapes of characters, more noise, and more challenging to handle, and consists of two subsets, large and small. The small dataset was manually cleaned and contained 19,000 images with a misclassification rate of 0.5%. In contrast, the large dataset was not manually cleaned and contained 500,000 images with a misclassification rate of approximately 6.5%. The images of the dataset are shown in Figure 3.

In addition, the size of the dataset here was 28*28*1*18724, and the label data type was categorical. The Randperm function generates a random sequence with the same number of samples. We chose 50% of the dataset as the training

TABLE 3. Initializing the training process.

Number of iterations	Time spent (hh:mm:ss)	Small batch accuracy(%)	Verification accuracy(%)	Small lot losses	Verification loss
1	0:00:13	12.50	9.47	3.0805	2.7120
50	0:00:15	80.47	83.39	0.6170	0.6073
100	0:00:16	91.41	87.54	0.4073	0.4259
150	0:00:17	95.31	88.53	0.2283	0.3919
200	0:00:18	94.53	89.96	0.1892	0.3534
250	0:00:19	92.97	89.92	0.2626	0.3475
300	0:00:20	90.62	90.24	0.3403	0.3401
350	0:00:22	88.28	90.88	0.3427	0.3247
400	0:00:23	95.31	90.80	0.1757	0.3218
450	0:00:24	89.84	90.81	0.2931	0.3215
500	0:00:25	90.62	90.80	0.2504	0.3317
...	...	...	...	...	...
3500	0:01:25	100.00	91.12	0.0101	0.4791
3550	0:01:26	100.00	91.21	0.0058	0.4732
3600	0:01:27	100.00	91.01	0.0183	0.4861
3650	0:01:28	100.00	91.19	0.0127	0.4795
Test set classification accuracy of 90.2804%					

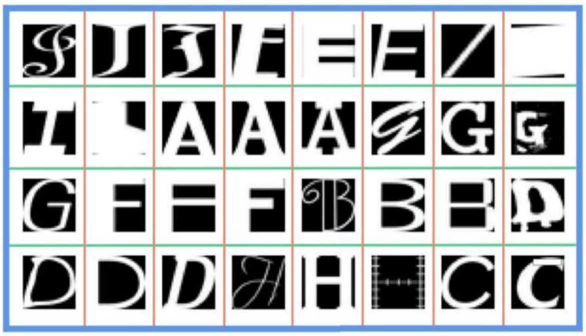


FIGURE 3. Some images of the dataset are shown.

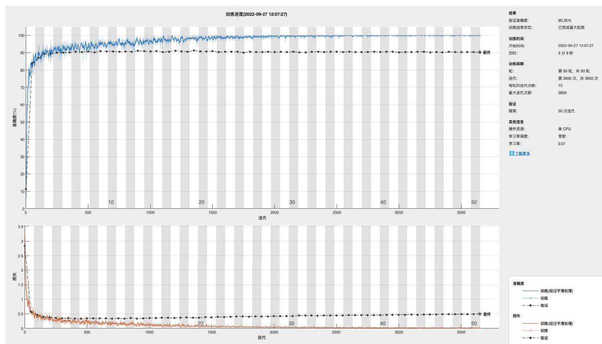


FIGURE 4. CNN training progress.

set, 30% as the validation set, and the remaining 20% as the test set. After dividing the dataset, a CNN network was constructed. The CNN defined here has one convolution layer, one pooling layer, one complete connection layer, and a batch normalization operation. The ReLU was selected as the activation function, and dropout was added to prevent overfitting. Next, we iterated and selected the sgdm method. We can also use Adam, rmsprop, etc. The maximum number of iterations was set to 50, and the training progress and results are shown in Figure 4.

It can be noted that without the validation set, the training process will iterate to the maximum number of iteration steps. However, when we encounter more complex problems, the final result will most likely be overfitted if the model lacks operations to prevent overfitting. Therefore, we must introduce a validation set to determine whether the model has been trained correctly.

Here, we discuss the effect of the different activation functions and training algorithms on the correctness of the validation set (the number of iterations was 50, and the learning rate was 0.01 to see the accuracy and time used for the three activation functions with the three training algorithms). We can see that the training process stopped early after adding the validation set. Epoch ended the training process after only eight iterations because when Patience was five, the training was terminated if the loss of the validation set did not decrease in five iterations. Table 3 shows the three training algorithms' accuracy and time used for the three activation functions.

Table 2 lists the specific initialization training process, including the number of iterations, training time, sample accuracy, and loss.

III. THE PROPOSED METHOD

Next, a complex network is built. First, each image pixel is used as the vertex of the complex network to build the average pixel function. The image's pixel value and radius information is obtained from the input image. Subsequently, the weights between the vertices of the complex network were calculated, and a weight matrix was constructed. The weights were calculated using the weight formula given the radius and threshold. Consider the D1 image of the Brodatz dataset as an example, given our previously configured parameters

TABLE 4. The image pixel value of image P(Intercepted section).

Location	1	2	3	4	5	6	7
1	0.3077	0.4231	0.5000	0.5385	0.5385	0.5000	0.4231
2	0.4231	0.5000	0.6923	0.6154	0.5769	0.5769	0.5385
3	0.5385	0.6538	0.9231	0.8462	0.7692	0.9231	0.9615
4	0.5769	0.7692	0.5000	0.3846	0.6923	0.5385	0.5385
5	0.5769	0.6154	0.7692	0.7308	0.6538	0.7308	0.8462
6	0.5769	0.6154	0.8462	0.8077	1.0000	0.9231	1.0000
7	0.5385	0.6154	0.8462	0.4231	0.9231	1.0000	0.9615

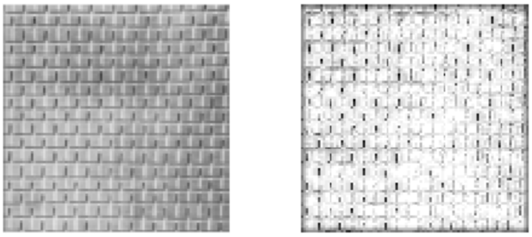


FIGURE 5. CNN training progress.

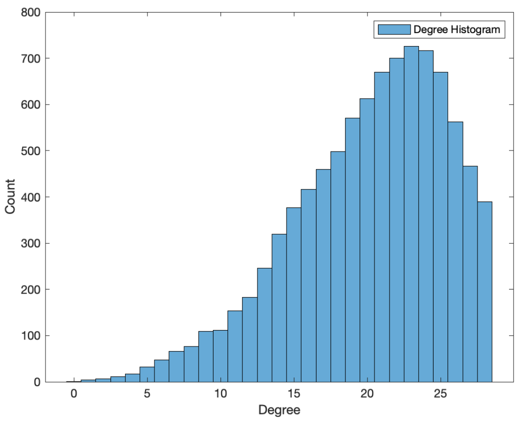


FIGURE 6. CNN training progress.

R = 3 and T = 3, to calculate the degree of histogram and image processing, as shown in Figures 5 and 6.

In addition, we can see the image pixel value variable I data, as shown in Table 4, which were all calculated after pixel averaging.

Looking at the weights between vertices again, as shown in Table 5, the distances between vertices and their connected vertices are within the given thresholds when a radius threshold is given.

As shown in Figure 7, given the radius parameter, the vertices that match the connection rule with a particular vertex have connected edges. The weight matrix was examined, as shown in Table 8.

The processed images were then packed into a dataset for training and testing. Some of the processing results are shown in Figure 8. The processed data were verified again through convolutional neural network training, and the training process is shown in Figure 9. Table 6 shows the modified initialization training process.

TABLE 5. Weights between partial vertices.

vertices	vertices	weights
1	2	0.3333
1	3	0.2778
1	4	0.3889
1	97	0.1667
1	98	0.3563
1	99	0.6798
1	193	0.3889
1	194	0.6242
1	195	0.7127
1	289	0.6111
2	3	0.1667
2	4	0.1667
2	5	0.2222
2	97	0.4675

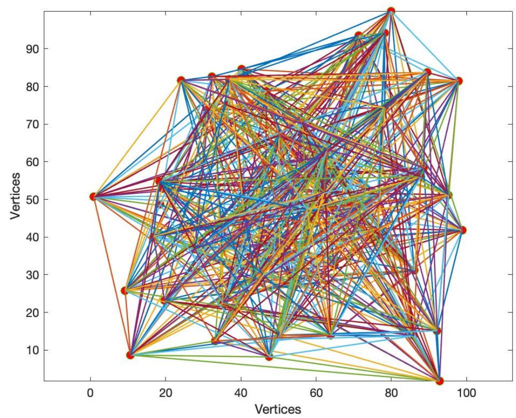


FIGURE 7. Weight matrix.

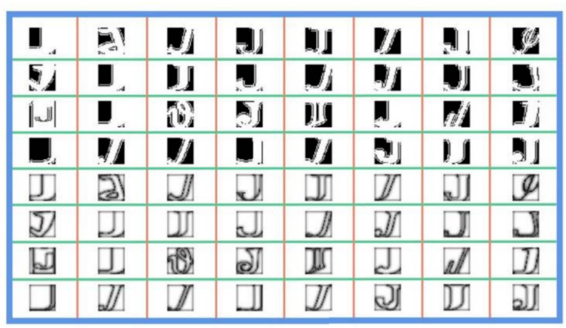


FIGURE 8. CNN training progress.

TABLE 6. Corrected initialization training process.

Number of iterations	Time spent (hh:mm:ss)	Small batch accuracy(%)	Verification accuracy(%)	Small lot losses	Verification loss
1	0:00:02	8.59	11.73	2.9618	2.5016
50	0:00:03	89.84	87.88	0.3833	0.4778
100	0:00:04	89.84	89.85	0.2989	0.3795
150	0:00:05	92.19	90.37	0.2138	0.3491
200	0:00:06	89.84	90.97	0.2142	0.3325
250	0:00:07	90.62	91.22	0.3248	0.3352
300	0:00:08	96.88	91.37	0.0965	0.3138
350	0:00:09	93.75	90.55	0.1648	0.3397
400	0:00:10	94.53	91.65	0.1674	0.3124
450	0:00:12	93.75	91.92	0.1634	0.3026
500	0:00:13	95.31	91.58	0.2362	0.3116
...	...	...	...	...	...
3500	0:01:03	100.00	92.72	0.0054	0.2545
3550	0:01:04	100.00	92.74	0.0384	0.2534
3600	0:01:05	100.00	93.02	0.0065	0.2552
3650	0:01:06	100.00	93.86	0.0051	0.2727
Test set classification accuracy of 92.2563%					

TABLE 7. Comparison of the correctness of five methods to recognize five image datasets.

Method	Mnist(%)	Notmnist(%)	Fashion-mnist(%)	ORL(%)
Logistic regression	87.28±0.39	86.34±0.20	85.28±0.39	86.09±0.24
Image analysis method	89.42±0.21	88.92±0.34	88.24±0.28	88.28±0.15
CNN	92.95±0.14	91.38±0.26	90.28±0.45	**91.09±0.21**
proposed method	**93.21±0.29**	**92.25±0.31**	**91.88±0.18**	88.98±0.32

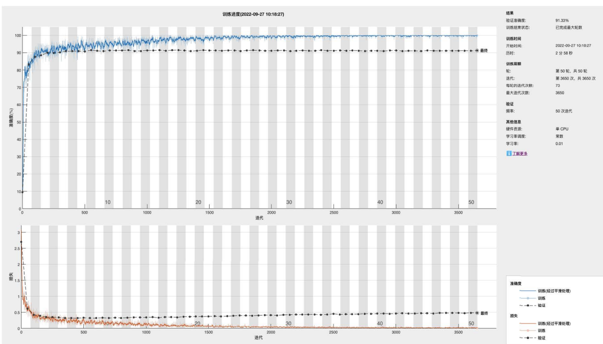


FIGURE 9. CNN training progress.

TABLE 8. The weight matrix of image P(Intercepted section).

Location	1	2	3	4	5	6	7	8	9
1	10	13	15	16	16	15	13	13	8
2	13	15	20	18	17	17	16	16	12
3	16	19	26	24	22	26	27	27	12
4	17	22	15	12	20	16	16	17	24
5	17	18	22	21	19	21	24	27	24
6	17	18	24	23	28	26	28	24	24
7	16	18	24	13	26	28	27	27	25
8	16	19	24	12	28	28	28	28	26
9	15	22	26	15	26	25	21	21	21

Two percentage points improved the recognition accuracy compared to the previous model, and all the result parameters were improved, which shows the accuracy of the recognition classification accuracy of the model after adding the complex network.

We then applied the model to the other three datasets and compared it with the other three image recognition methods. The results are shown in Table 7.

The table shows that our method performs well on the Mnist, Notmnist, and Fashion-Mnist datasets.

IV. CONCLUSION

This paper proposes a complex network-based CNN image recognition method that builds a complex network from an image's pixel points and then trains a neural network to recognize images more accurately. For a small sample classification picture dataset, the Euclidean distance formula is used to determine the distance between each pixel point in the image, which is then treated as a vertex of the complex network. The connection rules determine the connection of edges between vertices given the distance parameters. The weights of the linked edges are then computed. The original weight calculation algorithm is changed to lessen its sensitivity to the radius parameter to enhance the uniformity and concentration of the degree histogram distribution. The linked edges that did not follow the connection rules were then deleted the connection rules, thereby simplifying the complicated network given the weight values. The weights were normalized to obtain equal weights for the radius and pixel information of the picture. The complicated network for the picture data was created in MATLAB and then integrated with a convolutional neural network to train image recognition.

The final results demonstrated that, compared to other image recognition techniques, the training time was significantly decreased, and the recognition accuracy was significantly increased.

However, the correct rate in this aspect is unstable in terms of different types of images. It is worth considering adjusting the complex network description parameters for different characteristic images again. Second, using multiple evolutionary methods to obtain as comprehensive and practical features as possible is a problem worth exploring.

REFERENCES

[1] A. R. Backes, A. S. Martinez, and O. M. Bruno, "Texture analysis using graphs generated by deterministic partially self-avoiding walks," *Pattern Recognit.*, vol. 44, no. 8, pp. 1684–1689, Aug. 2011.

[2] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, May 2011, pp. 302–306.

[3] J. Espinal, M. Ornelas, H. J. Puga, J. M. Carpio, and J. A. Munoz, "3D object reconstruction using structured light and neural networks," in *Proc. IEEE Electron., Robot. Automot. Mech. Conf.*, Sep. 2010, pp. 74–79.

[4] R. A. Brooks, "Model-based three-dimensional interpretations of two-dimensional images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-5, no. 2, pp. 140–150, Mar. 1983.

[5] A. R. Backes, D. Casanova, and O. M. Bruno, "A complex network-based approach for boundary shape analysis," *Pattern Recognit.*, vol. 42, no. 1, pp. 54–67, 2009.

[6] W. Pan, K. Qin, and Y. Chen, "An adaptable-multilayer fractional Fourier transform approach for image registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 400–414, Mar. 2009.

[7] L. D. F. Costa, "Complex networks: New concepts and tools for real-time imaging and vision," 2006, *arXiv:cs/0606060*.

[8] W. N. Gonçalves, J. D. A. Silva, and O. M. Bruno, "A rotation invariant face recognition method based on complex network," in *Iberoamerican Congress on Pattern Recognition*. Berlin, Germany: Springer, 2010, pp. 426–433.

[9] T. Xiao, W. Yin-He, and W. Qin-Ruo, "A rotation and scale invariance face recognition method based on complex network and image contour," in *Proc. 12th Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Dec. 2012, pp. 371–376.

[10] L. Zhou, C. Zhang, Z. Wang, Y. Wang, and Z. Lu, "Hierarchical palmprint feature extraction and recognition based on multi-wavelets and complex network," *IET Image Process.*, vol. 12, no. 6, pp. 985–992, Jun. 2018.

[11] D. Cartwright and F. Harary, "Structural balance: A generalization of Heider's theory," *Psychol. Rev.*, vol. 63, no. 5, pp. 277–293, 1956.

[12] L. N. Couto, A. R. Backes, and C. A. Z. Barcelos, "Texture characterization via deterministic walks' direction histogram applied to a complex network-based image transformation," *Pattern Recognit. Lett.*, vol. 97, pp. 77–83, Oct. 2017.

[13] L. Zhou, "Palmprint feature extraction based on multi-wavelet and complex network," *J. Inf. Hiding Multim. Signal Process.*, vol. 8, no. 3, pp. 589–598, 2017.

ZHUANG MA was born in Cangzhou, Hebei, China. He is currently pursuing the Ph.D. degree with the China University of Geosciences, Beijing, China. His research interests include image recognition and processing, machine learning and deep learning, isotope double diluents, and first-principles computing.

GUANGDONG HUANG is currently the Ph.D. Associate Professor of mathematics with the School of Science, China University of Geosciences, Beijing, China. He has published more than 20 research papers independently or cooperatively, and presided over and participated in nine projects. His research interests include mathematical modeling, modern intelligent algorithm, financial engineering,and probability theory and statistics. He is mainly engaged in the research and teaching of applied statistics and machine learning.