

## Research Article

# Intelligent Retrieval Method of Approximate Painting in Digital Art Field

Jixin Wan  and Yu Xiaobo

*Xiamen Academy of Arts and Design, Fuzhou University, Fujian, Xiamen 361021, China*

Correspondence should be addressed to Jixin Wan; [t04303@fzu.edu.cn](mailto:t04303@fzu.edu.cn)

Received 13 September 2021; Revised 13 October 2021; Accepted 20 October 2021; Published 20 November 2021

Academic Editor: Bai Yuan Ding

Copyright © 2021 Jixin Wan and Yu Xiaobo. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of Internet technology and the wide application of image acquisition equipment, the number of digital artwork images is exploding. The retrieval of near-similar artwork images has a wide application prospect for copyright infringement, trademark registration, and other scenes. However, compared with traditional images, these artwork images have the characteristics of high similarity and complexity, which lead to the retrieval accuracy not meeting the demand. To solve the above problems, an intelligent retrieval method of artwork image based on wavelet transform and dual propagation neural network (WTCPN) is proposed. Firstly, the original artwork image is replaced by the low-frequency subimage after wavelet transform, which not only removes redundant information and reduces the dimension of data but also suppresses random noise. Secondly, in order to make the network assign different competition winning units to different types of modes, the dual propagation neural network is improved by setting the maximum number of times of winning neurons. Experimental results show that the proposed method can improve the accuracy of image retrieval, and the recognition accuracy of verification set can reach over 91%.

## 1. Introduction

With the advent of the 21st century, computer technology has been rapidly improved, and the wide popularization of image technology has led to an explosive growth in the number of digital artwork images. There are a large number of near-similar images in a large number of picture resources, and the retrieval of near-similar images has broad application prospects, such as detecting copyright infringement of works of art on the Internet and detecting whether new trademarks are similar to existing trademarks [1–6]. In recent years, near-similar image retrieval has gradually become an important branch of image retrieval and attracted more and more attention.

Nowadays, image retrieval technology is the fastest-growing technology in all information retrieval applications, and the number of image retrieval in several major retrieval engines all over the world is increasing very fast. This also makes the near-similar image retrieval technology become the focus of extensive attention at home and abroad and also

becomes a key research issue in the development of information industry and digital media technology. Near-similar image detection uses an image as a query keyword [7–10] and retrieves all database images that are all similar or partially similar to the query image. There are many methods for near-similar image detection, among which feature-based detection, index-based detection, and content-based detection are the most commonly used. At present, content-based image retrieval technology has entered a new research stage. Content retrieval based on image visual features begins to use machine learning to extract image features [11–13]. Artificial neural network (ANN), which belongs to machine learning technology [14, 15], is a widely used network model in the development of machine learning and can extract higher-level features of pictures. Artificial neural network is a neural network constructed artificially and capable of realizing a certain function. It is a theoretical mathematical model of human brain neural network based on the understanding of human brain neural network. It is an information processing system established by imitating

the structure and function of brain neural network. Artificial neural network is a brand-new system with logical thinking, fuzzy processing, and accurate calculation. It is a complex network formed by connecting a large number of simple neurons.

The traditional neural network is based on BP algorithm, but the BP algorithm has been studied with it, and some problems have been found [16–18]. (1) The signal of error correction will become smaller and smaller from top to bottom, and gradient dispersion will occur. (2) It can only converge to the local minimum. However, in practice, the use of tags is usually presented in the form of no data, while BP algorithm can only use tagged data for training. Therefore, it is difficult for BP algorithm to be widely used in practical environment. Counter propagation network (CPN) is an advanced artificial neural network proposed by Robert Hecht-Nielsen, an American scholar, in 1987.

To solve the problem of poor retrieval accuracy of approximate artwork images, an intelligent retrieval method of artwork images based on wavelet transform and dual propagation neural network (WTCPN) is proposed. The main contributions of this paper are as follows:

- (1) Replacing the original artwork image with the low-frequency subimage after wavelet transform not only removes redundant information but also reduces the dimension of data and the irrelevant factors that interfere with the recognition performance of CPN
- (2) The CPN is improved by setting the maximum times of winning neurons, which not only effectively overcomes the instability of the competition layer of CPN network but also avoids the local minimum points in the training process

## 2. Literature Review

Content-based image retrieval technology is the research direction. Dubey et al. [19] realized the local binary mode of multichannel decoding through the adder and the decoder, which can be effectively used for content-based image retrieval and is superior to other multichannel-based methods in average retrieval accuracy and average retrieval rate. Alshehri [20] proposed an image retrieval method based on BP neural network prediction technology, which classified and predicted the retrieved data by fuzzy inference of neural network. However, according to the above analysis, the research and development of traditional neural networks are limited by the problems of BP algorithm. Recently, the traditional neural network has been replaced by new network models, such as Boltzmann machine, convolutional neural network (CNN), and residual neural network (RNN), which have promoted the further development of this field.

Filip et al. [21] proposed a CNN image retrieval method without manual annotation, which expanded the maximum average pool through the trainable generalized average pool layer and showed better image retrieval performance on Oxford Buildings and Holiday Datasets. Liu et al. [22] aimed at the retrieval of large-scale network image resources and realized a simple and effective image indexing framework by

combining the pretrained large-scale convolutional neural network with the structured support vector machine. However, because this method is aimed at large-scale image retrieval, the retrieval speed has been greatly improved, but the retrieval accuracy is not high. Rajkumar and Sudhamani [23] proposed an image retrieval system based on residual neural network and used Euclidean distance measure to measure similarity. The retrieval test is carried out on a dataset with 50,000 network images in 250 categories. Experimental results show that compared with Google's random query image retrieval system, the performance of the proposed system is improved by 15%. Wang et al. [24] proposed a large-scale similar image retrieval method based on deep neural network, which mainly used deep frame learning multilevel nonlinear transformation to obtain advanced image features and achieved good retrieval results.

Usually, the dimensions of digital artwork images are very high, which make the algorithm need a long time and a large amount of computation. If the distribution of artwork image points is not very compact, it is not conducive to feature classification because artwork images usually have high-dimensional spatial points. To solve the above problems, this paper uses wavelet transform, which can not only reduce the dimension of artwork image but also filter the high-frequency interference information, highlight the main features of the image, and obtain a low-dimensional image suitable for neural network recognition. In addition, by combining the main features of wavelet transform and CPN network, a higher retrieval recognition rate is achieved.

## 3. Image Retrieval Based on Wavelet Transform and Improved CPN Network

*3.1. Operation Principle and Learning Algorithm of CPN.* CPN has three layers of standard structure, and neurons in each layer are all connected with each other. Figure 1 shows the topology of CPN network, which is composed of the input layer, the competition layer, and the output layer. Among them, the input layer and the competition layer form a feature mapping network. The competition layer and the output layer constitute a basic competitive network.

In each layer of CPN, the input vector is represented by  $X$ :

$$X = (x_1, x_2, \dots, x_n)^T. \quad (1)$$

After the competition, the output of the competition layer is expressed by  $Y$ :

$$Y = (y_1, y_2, \dots, y_m)^T, \quad y_i \in \{0, 1\}, \quad i = 1, 2, \dots, m. \quad (2)$$

The output of the network is represented by  $O$ :

$$O = (o_1, o_2, \dots, o_l)^T. \quad (3)$$

The expected output of the network is denoted by  $d$ :

$$d = (d_1, d_2, \dots, d_l)^T. \quad (4)$$

The weight matrix between the input layer and the competition layer is expressed by  $V$ :

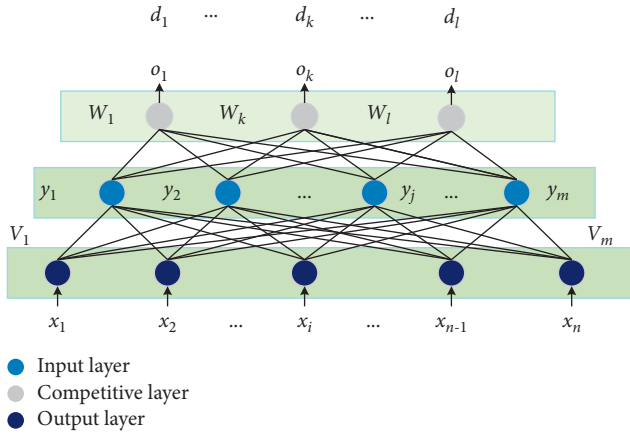


FIGURE 1: Topological structure of CPN.

$$V = (v_1, v_2, \dots, v_j, \dots, v_m), \quad (5)$$

where the column vector  $v_j$  is the inner star weight vector corresponding to the  $j$ th neuron in the competition layer.

The weight matrix between the competition layer and the output layer is represented by  $W$ :

$$W = (w_1, w_2, \dots, w_k, \dots, w_l), \quad (6)$$

where the column vector  $w_k$  is the weight vector corresponding to the  $k$ th neuron in the output layer.

It can be seen from Figure 2(a) that, after each layer of the network is trained according to the learning rules, it sends input vectors to the network in the running stage, and then, the competition layer performs competition calculation on them. When the net input value of a neuron is the maximum, it wins the competition, becomes the representative of the current input mode class, and becomes the active neuron shown in the figure at the same time, with an output value of  $L$ . While the rest neurons are inactive, and the output value is 0.

It can be seen from Figure 2(b) that, after the competing neurons win the competition, the neurons in the output layer are excited to produce the output pattern shown in the figure. The output value of failed neurons is 0, and the neurons in the output layer do not contribute to the net input and do not affect the output value, so the output is determined by the alien vector corresponding to the neurons that compete for victory.

The learning rules of CPN are composed of unsupervised learning and supervised learning, so the input vector and the expected output vector in the training sample set should be paired. Training is divided into two stages, and each stage adopts a learning rule. In the first stage, the competitive learning algorithm is used to train the inner star weight vector from the input layer to the competitive layer, and the steps are as follows:

- (1) All inner star weights are randomly assigned with initial values between 0 and 1 and normalized to unit length, and all input modes in the training set are also normalized.

- (2) Enter a pattern  $X^p$ ,  $p = 1, 2, \dots, P$ , where  $P$  is the total number of patterns in the training set.
- (3) Determine the competition winning neuron. The competition algorithm of CPN has no winning neighborhood, so only the inner star weight vector of the winning neuron is adjusted. The adjustment rule is as follows:

$$V(t+1) = \hat{V}(t) + \eta(t)[\hat{X} - \hat{V}(t)], \quad (7)$$

where  $\eta(t)$  is the learning rate which is an annealing function that decreases with time.

- (4) Repeat steps (2) to (3) until it drops to 0. It should be noted that the weight vector must be normalized again after adjustment.

In the second stage, an alien learning algorithm is adopted to train the alien weight vector from the competition layer to the output layer, and the steps are as follows:

- (1) Input a mode pair  $X^p$  and  $d^p$ , wherein the weight matrix from the input layer to the competition layer keeps the training result of the first stage.
- (2) Determine the neuron  $net_{j^*} = \max_j \{net_j\}$  that wins the competition, and satisfy

$$y_j = \begin{cases} 0, & j \neq j^* \\ 1, & j = j^* \end{cases} \quad (8)$$

- (3) Adjust the alien weight vector from the competition layer to the output layer, and the adjustment rule is

$$W_{j^*}(t+1) = \hat{W}_{j^*}(t) + \beta(t)[d - O(t)], \quad (9)$$

where  $\beta(t)$  is the learning rate of alien rules, and it is also an annealing function that decreases with time;  $O = (o_1, o_2, \dots, o_l)$  is the output value of the neurons out of the layer, which is calculated by the following formula:

$$o_k(t) = \sum_{j=1}^m w_{jk} y_j, \quad k = 1, 2, \dots, l, \quad (10)$$

$$O_k(t) = w_{j^*k} y_{j^*} = w_{j^*k}.$$

The alien weight vector adjustment rules are as follows:

$$W_{j^*}(t+1) = W_{j^*}(t) + \beta(t)[d - W_{j^*}(t)]. \quad (11)$$

- (4) Repeat steps (1) to (3) until  $\beta(t)$  drops to 0.

**3.2. Principle of Wavelet Transform.** Firstly, the wavelet transform filters the signal [25], using a group of high-pass and low-pass filters with different scales. Then, we analyze and process the signals decomposed into different frequency bands, which are decomposed from the high-frequency and low-frequency components in the original signal. Finally, in order to reach the preset threshold, the above filtering process should be repeated.

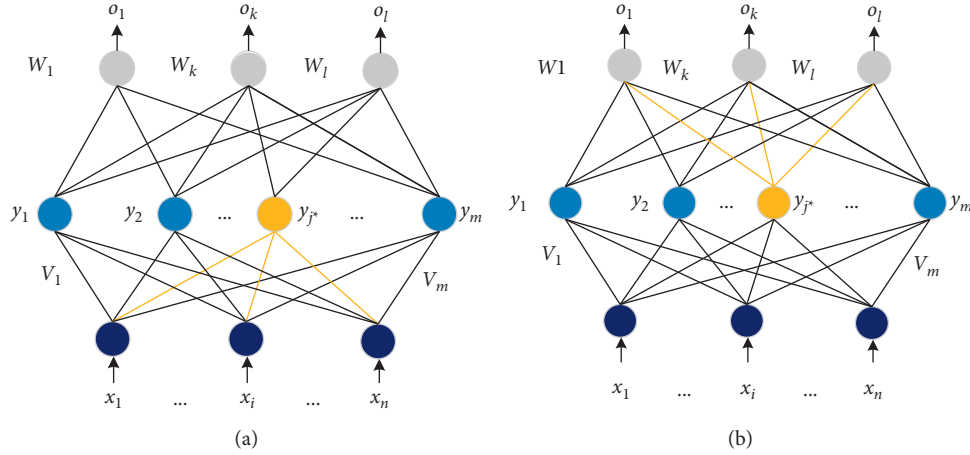


FIGURE 2: Operation process of CPN network. (a) Competition generates winning nodes; (b) win node's alien vector decision output.

At different scales  $\alpha$ , the function  $\psi(t)$  of a basic wavelet (mother wavelet) and the signal  $x(t)$  to be analyzed carry out the inner product, which constitutes the wavelet basis function:

$$WT_x(\alpha, \tau) = \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{+\infty} x(t) \psi^* \left( \frac{t - \tau}{\alpha} \right) dt, \quad \alpha > 0. \quad (12)$$

The equivalent frequency domain is expressed as follows:

$$WT_x(\alpha, \tau) = \frac{\sqrt{\alpha}}{2\pi} \int_{-\infty}^{+\infty} x(\omega) \psi^*(\alpha\omega) e^{j\omega\tau} d\omega. \quad (13)$$

Images in real life are generally two-dimensional signals. Therefore, extending wavelet from one dimension to two dimensions is the basic idea of applying two-dimensional wavelet transform to image processing. Two-dimensional wavelet transform is used to decompose the artwork image in frequency domain, and four regions can be obtained: low-frequency region is different from high-frequency region and the former LL is an approximate component, while the latter LH, HL, and HH are the horizontal component, the vertical component, and the diagonal component, respectively. Among them, the transformed low-frequency region LL can also be subjected to wavelet transform again.

Figure 3 is a schematic diagram of primary and secondary wavelet decomposition, respectively, in which LL1 is the low-frequency subimage of the original image, LH1 and HL1 are horizontal and vertical subimages, and HH1 is the high-frequency subimage of the image.

**3.3. Proposed WTCPN.** The proposed WTCPN method is divided into two processes: training process and recognition process. In the training process, firstly, the training sample of the image is read, and the original training sample set is  $X$ , and the artwork image is decomposed by wavelet transform. The wavelet basis function selected here is Daubechies [26], and the wavelet coefficients obtained are independent of each other. Finally, the transformed LL part is selected as the approximation of the original image.

Then, the improved CPN algorithm is used to classify the images. In the process of identification, firstly, the test

samples are subjected to wavelet transform, and their low-frequency component subgraphs are selected as the approximation of the test samples, and then, the test sample identification space is constructed, which is classified and identified according to the improved CPN algorithm.

Let  $\mathbf{U}_k = [u_1^k, u_2^k, \dots, u_n^k]^T$  be the input mode of CPN network and  $\mathbf{V}_k = [v_1^k, v_2^k, \dots, v_m^k]^T$  be the output of the competition layer, but the actual output of the output layer is  $\mathbf{C}_k = [c_1^k, c_2^k, \dots, c_z^k]^T$  and the desired output of the output layer is  $\mathbf{Y}_k = [y_1^k, y_2^k, \dots, y_z^k]^T$ . The number of neurons in the input layer, the competition layer, and the output layer is  $n$ ,  $m$ , and  $z$ , respectively.  $P$  is the number of input modes.  $\mathbf{W}_j = [w_{j1}, w_{j2}, \dots, w_{jn}]^T$  is the connection weight vector from the input layer to the competition layer.  $\mathbf{Q}_l = [q_{11}, q_{12}, \dots, q_{lm}]^T$  is the connection weight vector from the competition layer to the output layer.

The improved CPN learning algorithm includes the following steps:

- (1) Assign each component of  $\mathbf{W}_j$ , ( $j = 1, 2, \dots, m$ ) and  $\mathbf{Q}_l$  ( $l = 1, 2, \dots, z$ ) to a random value in the  $[0, 1]$  interval for initialization. At the same time, a variable  $t$  (with an initial value of 0) is added to each neuron in the competition layer to record the number of times the neuron won. The maximum number of neuron wins is set as  $T$ , and the error tolerance is specified as  $e$ .
- (2) The  $k$ th input pattern  $U_k$  is provided to the network input layer.
- (3) Carry out normalization processing on the connection weight vector  $\mathbf{W}_j$  ( $j = 1, 2, \dots, m$ ).
- (4) Find the input activation value of neurons in the competition layer:

$$S_j = \sum_{i=1}^n w_{ji} u_i^k, \quad (j = 1, 2, \dots, m). \quad (14)$$

- (5) Find out the maximum activation value  $S_\alpha$  from the calculated  $S_j$ . If  $t$  of  $S_\alpha$  is less than  $T$ ,  $t = t + 1$ , and the neuron corresponding to  $S_\alpha$  is taken as the winning neuron  $g$  in the competition layer. Otherwise, if  $t \geq T$ ,

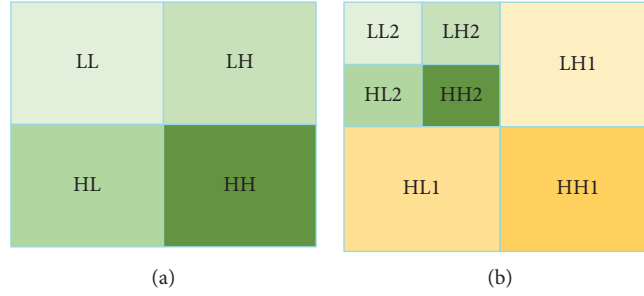


FIGURE 3: Schematic diagram of wavelet decomposition. (a) Primary wavelet decomposition; (b) secondary wavelet decomposition.

select the maximum activation value  $S_b$  except  $S_a$ . If  $t$  of  $S_b$  is less than  $T$ ,  $t = t + 1$ , and the neuron corresponding to  $S_b$  is regarded as the winning neuron  $g$  in the competition layer. Otherwise, search in  $S_j$  in order of activation value from large to small. Set the output of the winning neuron  $g$  in the competition layer as 1, and the rest as 0, and its corresponding connection weight is  $W_g$ .

- (6) Adjust  $W_g$  in the following ways:

$$w_{gi}(t+1) = w_{gi}(t) + \alpha(u_i^k - w_{gi}(t)) \quad (15)$$

$$(i = 1, 2, \dots, n),$$

where  $\alpha$  is the learning rate.

- (7) Adjust the connection weight vector  $\mathbf{Q}_g = [q_{1g}, q_{2g}, \dots, q_{lg}]$  from the winning neuron  $g$  in the competition layer to the neurons in the output layer, while other connection weights remain unchanged as follows:

$$q_{lg}(t+1) = q_{lg}(t) + \beta v_j(y_1^k - c_l^k), \quad l = 1, 2, \dots, z. \quad (16)$$

- (8) Calculate the weighted sum of the comprehensive input signals of each neuron in the output layer and take it as the actual output of the output neuron.  
 (9) Calculate the error between the actual output  $\mathbf{C}_k$  of the network and the desired output  $\mathbf{Y}_k$ :

$$\text{err} = \sqrt{\sum_{i=1}^z (c_i^k - y_i^k)^2}. \quad (17)$$

- (10) Judging whether the error calculated in step (9) is less than the error tolerance, and if so, continue step (11) to learn the next mode. If it is greater than the error tolerance, return to step (3) to continue learning.  
 (11) Return to step (2) until all the  $p$  input modes are provided to the network.

## 4. Experimental Results and Analysis

**4.1. Common Datasets and Performance Evaluation Methods.** The effectiveness of the WTCPN method is analyzed experimentally on two public datasets, specifically Oxford [27]

and Holiday [28]. Oxford architecture dataset contains 55 query images corresponding to 11 different buildings. Each query image has a rectangular area to define the building. Holiday dataset includes 1491 amplitude false pictures, which are divided into 500 groups, and each group has a different scene or object. The error tolerance  $e = 0.01$ , learning rate  $\alpha = 0.4$ , and learning rate  $\beta = 0.5$  in WTCPN algorithm. The number of neuron nodes in the input layer is 4096, the number of neuron nodes in the output layer is 21, and the number of neuron nodes in the competition layer is 30. The experimental environment is carried out on our own server, the operating system is Ubuntu14.04, the CPU is I5, and the graphics card is GTX1060. Divide the training dataset and verification dataset according to the ratio of 4:1.

The average correct rate is used to evaluate the performance of image retrieval. The average accuracy is calculated by using the area of the P-R curve, where  $P$  represents accuracy, which is the ratio of the number of retrieved positive samples to the number of all retrieved images, and  $R$  represents the recovery rate, which is the ratio of the number of retrieved positive samples to the number of all positive samples in the dataset:

$$\text{Average accuracy rate} = \frac{1}{Q} \sum_{i=1}^Q \frac{\sum_{k=1}^n P_i(k) * rel_i(k)}{R_i}, \quad (18)$$

where  $Q$  represents the number of query images,  $R_i$  represents the number of dataset images belonging to the same group as the  $i$ th query image,  $P_i(k)$  represents the accuracy of  $k$  and  $i$ th query images,  $rel_i(k)$  is an indicator, and its value is 1 when the query result of the  $i$ th query image belongs to the same group as the  $k$ th image, otherwise it is 0, and  $n$  represents the total number of all images.

**4.2. Experimental Results of Public Datasets.** Because the convolution kernel size and activation function have great influence on the performance of CPN, two different control experiments are conducted to summarize the CPN architecture with the best recognition effect. In the first group, the convolution kernel sizes were set to  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ , and  $9 \times 9$ , respectively. There were four groups, and the unified activation function was ReLU. All experimental results are shown in Table 1.

Through the analysis of the experimental results shown in Table 1, it is found that the  $3 \times 3$  small convolution kernel has a better effect, but with the increase of the convolution

TABLE 1: Network training results for different convolution kernel sizes.

Kernel layer	Average accuracy rate (%)
$3 \times 3$	94.73
$5 \times 5$	93.87
$7 \times 7$	91.44
$9 \times 9$	89.56

kernel size, the experimental results are getting worse and worse because the large-sized convolution kernel is obviously too “rough” for the small-sized input image. This leads to the unsatisfactory experimental results of large-scale convolution kernels.

In the second group of experiments, the convolution kernel size was set to  $3 \times 3$ , and the activation function was changed. The network training results of different activation functions are shown in Table 2.

Through the experimental results, it is easy to find that convolution neural network with ReLU activation function has better effect. Figure 4 shows the results retrieved from two datasets by the WTCPN method. The first and second rows represent the images in Holiday dataset, and the third and fourth rows represent the images in Oxford dataset. The word Query below each image represents a query image, while TP represents a related image and FP represents an error image.

**4.3. Retrieval Performance of Art Datasets.** Some highly similar artwork images were obtained from open source websites, with a total of more than 3,600 images, which were divided into six categories, with 600 images in each category. These six categories are national clothing (clothing), saddle (saddle), leather hip flask (pnjh), national craft ornaments (gongyi), high hat (gdm), and Ma Touqin (mtq). Divide into training dataset and verification dataset according to the ratio of 4:1. The convolution kernel size is  $3 \times 3$ , and the activation function is ReLU. The retrieval result of the WTCPN method on artwork image dataset is shown in Figure 5.

In order to show the advantages of WTCPN, WTCPN is compared with BP network and CNN network which are widely used. These three network models have basically the same topological structure and all adopt three-tier structure, thus ensuring that the three methods are under the same conditions and are comparable. When the error capacity during training is the same, the number of neurons in the input layer and the output layer of both networks is the same. Accuracy indicates the accuracy in the verification process, while loss indicates that the loss value in the verification process is obvious. Generally, the higher the accuracy value and the lower the loss value, the better the trained neural network is. Comparison of retrieval performance of three network models on artwork image dataset is shown in Figures 6 and 7, respectively.

It can be seen from Figure 6 that the accuracy value of WTCPN is higher than that of BP network and CNN network, which is 91.83%. It can be seen from Figure 7 that

TABLE 2: Network training results for different activation functions.

Activate function layer	Average accuracy rate (%)
Sigmoid	93.09
ReLU	94.73
Tanh	88.67

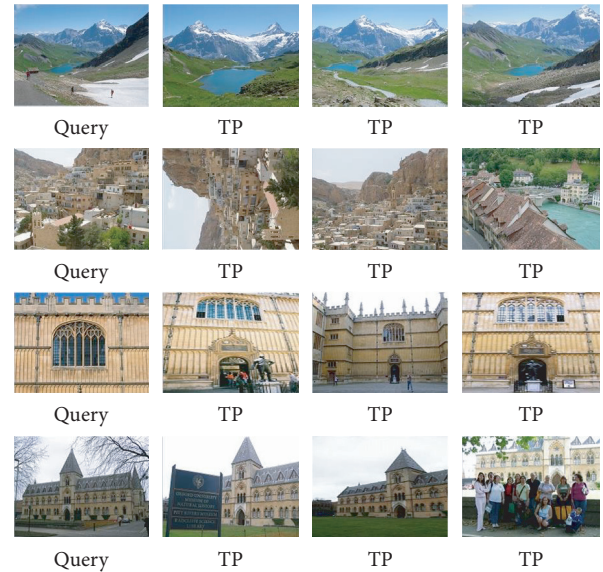


FIGURE 4: Results retrieved by the WTCPN method on two datasets.



FIGURE 5: Retrieval results of the WTCPN method on artwork image dataset.

the loss value of WTCPN is the lowest among the three network models, which is about 0.2. Therefore, it is obvious that the training results of BP network and CNN network are not as good as those of WTCPN, which means that WTCPN can get the best retrieval accuracy.

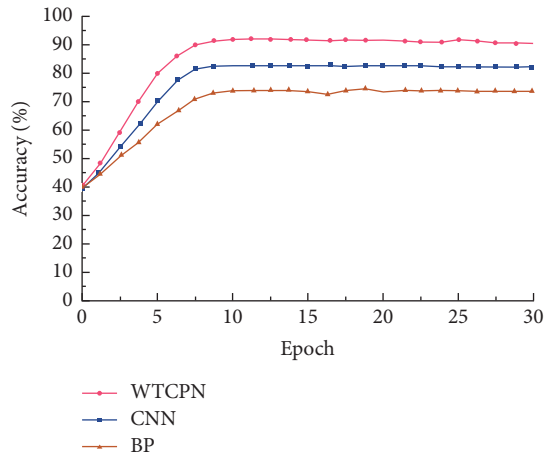


FIGURE 6: Accuracy results.

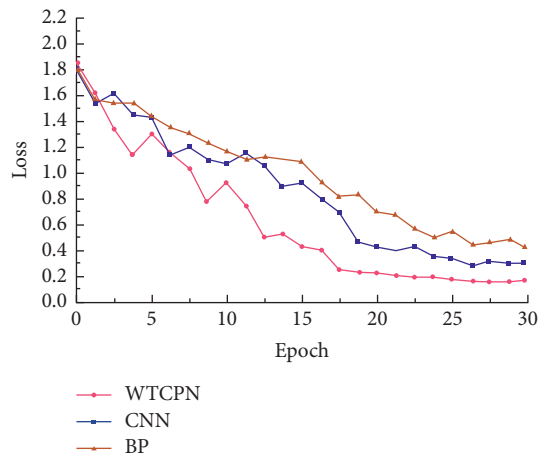


FIGURE 7: Loss results.

## 5. Conclusion

In this paper, an intelligent retrieval method of artwork image based on wavelet transform and dual propagation neural network (WTCPN) is proposed. By combining the main features of wavelet transform and CPN, the retrieval recognition rate is high. Test results on Oxford and Holiday datasets show that the optimal convolution kernel size of the WTCPN method is  $3 \times 3$ , and the optimal activation function is ReLU. Test results on art datasets show that the WTCPN method has the best retrieval accuracy compared with BP network and CNN network, with an accuracy of 91.83% and a loss value of about 0.2. In the future, aiming at the image retrieval in concurrent environment, based on the WTCPN method, we can increase message queue (MQ), MapReduce, and cache strategy, to improve the access and computing ability in concurrent environment and further improve the art image retrieval system.

## Data Availability

The datasets used in the findings of the study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] G. Amato, F. Fa Lchi, and L. Vadicamo, "Image retrieval," *Multimedia Tools and Applications*, vol. 77, no. 5, pp. 5385–5415, 2016.
- [2] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. M. Sun, and R. Kui, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2594–2608, 2017.
- [3] X.-S. Wei, J.-H. Luo, J. Wu, and Z.-H. Zhou, "Selective convolutional descriptor aggregation for fine-grained image retrieval," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2868–2881, 2017.
- [4] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "EPCBIR: an efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Information Sciences*, vol. 387, pp. 195–204, 2017.
- [5] M. H. Memon, J. P. Li, I. Memon, and Q. Ali, "GEO matching regions: multiple regions of interests using content based image retrieval based on relative locations," *Multimedia Tools and Applications*, vol. 76, no. 14, pp. 1–35, 2017.
- [6] Z. Zeng, "A novel local structure descriptor for color image retrieval," *Information*, vol. 7, no. 1, p. 9, 2016.
- [7] W. Zhou, S. Newsam, C. Li, and Z. Shao, "Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval," *Remote Sensing*, vol. 9, no. 5, p. 489, 2016.
- [8] M. Paulin, J. Mairal, M. Douze, Z. Harchaoui, F. Perronnin, and C. Schmid, "Convolutional patch representations for image retrieval: an unsupervised approach," *International Journal of Computer Vision*, vol. 121, no. 1, pp. 1–20, 2016.
- [9] Y. Wang, L. Zhang, X. Tong et al., "A three-layered graph-based learning approach for remote sensing image retrieval," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 1–15, 2016.
- [10] S. R. Dubey, S. K. Singh, and R. K. Singh, "Local bit-plane decoded pattern: a novel feature descriptor for biomedical image retrieval," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 4, pp. 1139–1147, 2016.
- [11] X. Wang, X. Duan, and X. Bai, "Deep sketch feature for cross-domain image retrieval," *Neurocomputing*, vol. 207, pp. 387–397, 2016.
- [12] A. K. Dhara, S. Mukhopadhyay, A. Dutta, M. Garg, and N. Khandelwal, "Content-based image retrieval system for pulmonary nodules: assisting radiologists in self-learning and diagnosis of lung cancer," *Journal of Digital Imaging*, vol. 30, no. 1, pp. 63–77, 2017.
- [13] L. Zhu, J. Shen, L. Xie, and Z. Cheng, "Unsupervised topic hypergraph hashing for efficient mobile image retrieval," *IEEE Transactions on Cybernetics*, vol. 47, pp. 1–14, 2016.
- [14] L. Xie, L. Zhu, P. Pan, and Y. Lu, "Cross-Modal Self-Taught Hashing for large-scale image retrieval," *Signal Processing*, vol. 124, no. 7, pp. 81–92, 2016.
- [15] A. Kumar, S. Dyer, J. Kim et al., "Adapting content-based image retrieval techniques for the semantic annotation of medical images," *Computerized Medical Imaging and Graphics*, vol. 49, pp. 37–45, 2016.

- [16] Z. Gao, J. Xue, W. Zhou, S. Pang, and Q. Tian, "Democratic diffusion aggregation for image retrieval," *IEEE Transactions on Multimedia*, vol. 18, no. 8, pp. 1661–1674, 2016.
- [17] R. Hong, L. Li, J. Cai, D. Tao, M. Wang, and Q. Tian, "Coherent semantic-visual indexing for large-scale image retrieval in the cloud," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4128–4138, 2017.
- [18] X. Li, "Tag relevance fusion for social image retrieval," *Multimedia Systems*, vol. 23, no. 1, pp. 29–40, 2017.
- [19] S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel decoded local binary patterns for content-based image retrieval," *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4018–4032, 2016.
- [20] M. Alshehri, "A content-based image retrieval method using neural network-based prediction technique," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 2957–2973, 2020.
- [21] R. Filip, T. Giorgos, and C. Ondrej, "Fine-tuning CNN image retrieval with no human annotation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, pp. 1655–1668, 2017.
- [22] R. Liu, Y. Zhao, S. Wei, and Y. Yang, "Indexing of CNN features for large scale image search," *Pattern Recognition*, vol. 48, no. 10, pp. 2983–2992, 2018.
- [23] R. Rajkumar and M. V. Sudhamani, "Image retrieval system using residual neural network in a distributed environment," *International Journal of Recent Technology and Engineering*, vol. 8, no. 6, pp. 2277–3878, 2020.
- [24] X. Wang, F. Lee, and Q. Chen, "Similarity-preserving hashing based on deep neural networks for large-scale image retrieval," *Journal of Visual Communication and Image Representation*, vol. 61, no. 5, pp. 260–271, 2019.
- [25] M. Verma and B. Raman, "Local tri-directional patterns: a new texture feature descriptor for image retrieval," *Digital Signal Processing*, vol. 51, pp. 62–72, 2016.
- [26] Y. Uchida, S. Sakazawa, and S. Satoh, "Image retrieval with Fisher vectors of binary features," *Ite Transactions on Media Technology & Applications*, vol. 4, no. 4, pp. 23–28, 2016.
- [27] X. Li, Y. Ye, and M. K. Ng, "MultiVCRank with applications to image retrieval," *IEEE Transactions on Image Processing*, vol. 25, no. 3, pp. 1396–1409, 2016.
- [28] F. Zhang, Y. Song, W. Cai et al., "Dictionary pruning with visual word significance for medical image retrieval," *Neurocomputing*, vol. 177, pp. 75–88, 2016.