

Research Article

Automatic Algorithm for Fractal Plant Art Image Similarity Feature Generation

Zhizhe Liu ¹ and Luo Sun²

¹School of Art and Design, Wuhan University of Technology, Wuhan 430070, China

²Glasgow School of Art, Glasgow, Scotland G4 9LE, UK

Correspondence should be addressed to Zhizhe Liu; liuzhizhe@whut.edu.cn

Received 29 September 2021; Revised 29 October 2021; Accepted 30 October 2021; Published 16 November 2021

Academic Editor: Miaochao Chen

Copyright © 2021 Zhizhe Liu and Luo Sun. 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 popularity of smart devices and the Internet, the volume of multimedia data is growing rapidly, and content-based image retrieval (CBIR) can search for similar images from large-scale images to realize the utilization of the data. For data owners, outsourcing the management and maintenance of image data to cloud service providers can effectively reduce costs, but there is a privacy leakage problem. In this paper, we focus on image feature extraction, index design, and image similarity recognition methods under a dual server model with content-based image security similarity recognition as the research topic, the work done such as proposing a BOVW (Bag of Visual Word) feature-based image security similarity recognition scheme. The scheme combines SIFT (scale-invariant feature transform) feature secure extraction and locally sensitive hashing algorithm to achieve secure extraction of BOVW features of images. To protect the BOVW features of images, an inverted index based on word frequency division is designed, the index is stored in chunks, and an image secure similarity recognition scheme based on CNN (convolutional neural networks) features is proposed. The scalable hash index based on dimensional division is designed based on the image CNN features secure extraction algorithm. The security and performance of the proposed scheme are theoretically analyzed and experimentally verified. Based on different image datasets, the impact of different parameters on the performance of the scheme is tested, and optimized parameters are given. The experimental results show that the proposed scheme in this paper can effectively improve the efficiency of analyzing the similarity of plant botanical art images compared to the existing schemes.

1. Introduction

With the rapid development of the Internet, the process of information technology is becoming faster and faster, and a large amount of image data can be generated every second. As an important information carrier, images have important applications in various aspects such as fractal plant art image similarity. With the increase in the number of images [1], it is a very meaningful thing to find out the effective image resources quickly and accurately from these massive image data. The study of image similarity algorithms can solve this problem to some extent. Due to the specificity of image similarity algorithms, there is also no unified standard for defining image similarity at present, and image similarity can only

be studied in various ways. Most of the current image similarity algorithms use low-level visual features.

It is especially important to obtain high-quality image data, because nowadays, smart life cannot be lived without image data processing. Blurred image data not only affects the visual perception of human eyes but also may increase some workload on important occasions. For example, in the automatic ticket channel of the railway station, the camera needs to collect clear face image and ID card image comparison, if the camera or the environment makes the collected face image blurred, it may need to take the manual ticket channel; monitoring system if not checked for a long time, the camera may be out of focus or due to the influence of weather, environment, etc. resulting in unclear

monitoring images; if encountered in an emergency, blurred medical images play a great supporting role in the process of treating patients, and if the quality of the obtained human tissue to science images is not high or clear, it may affect the doctor's judgment of the patient's disease; in the learning research of various image algorithms [2], it is also necessary to obtain high-quality image data first in order to improve the efficiency of the research work, for example, in the industrial image defect detection, high-quality defect images need to be obtained first before subsequent algorithm research can be conducted. It can be seen that it is quite important to obtain high-quality clear images.

As providers and users of web resources, users try to use cloud computing technology to achieve storage and sharing of image data and get rid of the burden caused by local data management. Due to the openness and heterogeneity of the network, users' private data face threats, and secure similarity recognition of images in the outsourcing environment becomes a problem that needs to be studied. The currently proposed image security similarity recognition schemes suffer from high computational complexity, high user workload, high communication overhead, and many interaction rounds and do not support the update of image data; therefore, studying image feature extraction, index construction, and image similarity comparison methods under cryptographic images is important and valuable to reduce the computational complexity and communication overhead of image security similarity recognition, and making full use of cloud server resources is of great significance and value [3].

2. Related Work

Image similarity is a relatively abstract and general concept. The subjective nature of human visual senses makes the visual medium of images highly ambiguous, and different people may have different understandings and perceptions of the same image [4]. Therefore, there are various methods and forms of describing the definition of similarity between images. However, the result and purpose of all of them are to quantitatively analyze and measure the degree of similarity between images. The similarity of an image depends on the degree of similarity of specific content on the image, such as similarity obtained by pixel-level comparison or by comparison and analysis of some specific points. The semantic similarity-based image similarity calculation gets some basic information [5] about the image for comparison by linking the image spatial context and situational context; it is a high-level linkage of image target entities with a high level of abstraction and immature theory. The expense of converting one graph into another graph is calculated, i.e., predefining the set of operations for various transformations and defining the minimum change step between two graphs as the similarity of the two graphs, i.e., the image edit distance [6].

The images are structured, the maximum common subgraph of the two graphs is derived by using the formula, which maximally expresses the common information of the two graphs, and the maximum common subgraph is

defined as the similarity of the two graphs [7]. A large graph that can contain two images is defined at the same time, called the maximum association graph of the two graphs, and the maximum subcluster from the association graph was obtained to represent the similarity of the two graphs. The image is decomposed into several parts, the similarity of each part is calculated separately, and then they are combined to get the similarity of the whole image [8].

Literature image search engine, which can be based on the input semantics or images, can find out much-related semantics or similar image resources, and the similarity recognition is very efficient. Many scholars in the domestic computer field have also done a lot of work in image similarity algorithm research. The literature mainly applies various algorithms to images of different situations to analyze the similarity of images under different algorithms, and it gives what kind of algorithms should be used in what kind of scenarios; the literature uses MapReduce, a distributed computing framework for big data, to study the similarity of images; the literature introduces a maximum connected region composed of the same color and its edge color roughness by which the distribution of image colors can be obtained and thus compensate for the deficiencies of the color histogram algorithm [9]; the literature combines two low-level features of images, texture, and color, to improve the accuracy of the image similarity algorithm; the literature constructs a HIS spatial color histogram and then uses an accumulative histogram by which the accuracy of image similarity recognition can be optimized. The literature developed an image search engine [10].

Based on the image uploaded by the user, the fast similarity is identified to many image resources related to this image. Literature uses the singular value decomposition theory for solving image similarity algorithm; literature uses the Skewed-Split k -trees; this tree structure quickly builds visual vocabulary feature set and then classifies the images based on the established vocabulary feature set; based on the literature by using an iterative framework [11], the main idea of this iterative framework is to keep recording the result of the last time, and then based on the last result, the size of the similarity recognition image is continuously expanded; and thus, the similarity identifies various image resources that meet the requirements. The literature considers that if two images are similar, then they belong to the same Flickr group combination and can be trained by the algorithm SIKMA to eventually categorize the similar images [12]. Due to the specificity of image data and the different understanding of image similarity, thus the study of image similarity from different directions has some shortcomings, and there is no standard to study the image similarity. Such as that from the image color features, using the number of colors in the image to calculate the image similarity [13], this research does not take into account the specific distribution of each color in the whole image; in some special scenarios, the calculated image similarity does not match the actual situation; from the spatial local features of the image to study the image similarity, you can use the SIFT algorithm or other matching feature point algorithm to study the image similarity, although the accuracy of these algorithms is different.

The SIFT algorithm or other matching feature point algorithms can be used to study the image similarity; although these algorithms have high accuracy, the calculation is too complicated. For example, the SIFT algorithm needs to calculate the 128-dimensional vector of the image, which is very complicated and does not meet people's expectations [14]. The performance of image retrieval does not only depend on the extracted image features. After the image features such as color, texture, and shape are extracted and indexed, the key to image retrieval lies in the similarity metric (or distance metric) function used. It is directly related to the results of image retrieval and retrieval efficiency. The text-based retrieval method uses the exact matching of the text, while the content-based image retrieval system is an inaccurate matching, which is completed by calculating the similarity of the visual features between the query example image and the candidate image. After image features such as color, texture, and shape are extracted, a feature vector is formed, and this feature vector can be used to characterize the corresponding image. In image retrieval, judging whether images are similar is done by comparing whether these feature vectors are similar. That is, the comparison between image feature vectors is regarded as image similarity comparison. A good feature vector comparison algorithm has a great influence on the image retrieval results.

3. Algorithm for Automatic Generation of Fractal Plant Art Image Similarity Features

3.1. Automatic Generation Based on Plant Image Similarity Algorithm. For similar species of plants, this algorithm is used to carry out a fast recognition of them. The grayscale histogram is a function of gray levels, which represents the number of pixels of each level in the image, reflecting the probability distribution of the gray values of the image pixels; it is a global information feature of the image, usually represented by a bar graph, the process of computing the image grayscale histogram is easy, and we define the histogram of a digital image with gray levels in the range [0-L-1] as a discrete function, with the number of pixels of level gray [15].

Image search engines include (1) image retrieval based on text: the characteristics of image information make it different from text information in retrieval; (2) image retrieval based on image content features: content-based image retrieval is indexed and retrieved mainly based on the image's screen content features and the subject and object features (i.e., the actual content of the image); and (3) image information content features that can be used for network retrieval include (a) static image content features, such characteristics include color, shape, texture, and structure, and (2) dynamic image content characteristics.

Similarity measures are one of the important foundations in the field of machine learning. In many pattern recognition and computer vision studies, it is necessary to measure the similarity between pairs of samples, for example, for many clustering and classification problems built on the concept of "feature space," each sample is described by a numerical attribute vector corresponding to a point in

the feature space. The distance between two points reflects the similarity between the two samples. Similarly, in image retrieval studies, the similarity measure of an image is an important step after image feature extraction. When computing query image features and dataset image features, different similarity measures can be used to obtain different similarity rankings and thus different retrieval results. Therefore, it is especially important to define and use appropriate similarity metrics on the defined feature space to effectively complete the image retrieval task. In this paper, we focus on several distance metric functions commonly used in similarity metrics [16], and similarity learning methods based on diffusion processes. The classical histogram similarity metric is used to calculate the normalized correlation coefficient between two histograms, and the idea of this method is simple, which is to judge the similarity by the difference between vectors mathematically. So the difference in distance between two histograms can also be expressed as the degree of nonsimilarity between the two images, and the principle of the formula is

$$\text{SIM}(G) = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i) + \sum_{i=1}^n (X_i + Y_i). \quad (1)$$

The common correlation coefficients normalized between histograms are Euclidean distance, Barclay's distance, histogram intersection distance, chi-square coefficient, etc., and we will give their metric formulas separately below. Given two images and their histograms H1 and H2, the Euclidean distance between the two images is calculated as follows.

$$D(H1) = \frac{\Delta y}{\Delta x} \times \frac{\partial^2 \Omega}{\partial u \partial v} \times \nabla |\Omega|. \quad (2)$$

There is no clear definition for the feature points of an image. It can also be called key points, points of interest, etc. of an image. And in fact, feature points of an image are the location identification of points on an image. The feature of a point represented by the position of the point is one of the simplest image features. Each image has its feature points, and what they represent is some of the more important locations on the image, such as the inflection points, corner points, or intersections of the image. Point features of images are the basis of many computer vision algorithms and are used in various fields, such as motion target tracking, object recognition, 3D reconstruction, and image alignment. And feature point-based image matching means extracting feature points containing important features from the matched images and then matching image features according to the corresponding similarity metric.

The specific characteristics of the SIFT algorithm are as follows: (1) The local features of the image remain unchanged for rotation, scale scaling, and brightness changes, and a certain degree of stability for viewing angle changes, affine transformations, and noise; (2) good uniqueness, rich information, and suitable for fast and accurate matching of massive feature libraries; (3) massiveness, even

a few objects can produce a large number of SIFT features; (4) high speed, the optimized SIFT matching algorithm can even be achieved real-time; and (5) expandability, which can be easily combined with other feature vectors.

The common image feature points are the Harris corner points, SIFT feature points, etc. Among them, the Harris corner points (Harris is the point where the local window moves along each direction and produces a significant change in the local curvature of the image, i.e., the point where the local curvature of the image changes abruptly). The corner point is a typical corner point detection algorithm, which has the advantages of rotational invariance and partial invariance of affine change, but not invariance to changes in image geometric scale, whereas the SIFT algorithm is also called scale-invariant feature transformation, which has deflating rotational and affine invariant properties and has strong adaptability to illumination and image deformation; therefore, SIFT algorithm becomes the most commonly used image feature point descriptor and is used in studies such as image matching [17].

3.2. Automatic Generation Process of Similarity Algorithm.

SIFT algorithm's full name is scale-invariant feature transformation algorithm. The core idea of the algorithm is to find extreme value points in the spatial scale and extract their position [18], scale, and rotation invariants for detecting and characterizing local features in the image. Its important feature is that it is invariant to image translation, rotation, scaling, and even affine transformations. It has good stability, so this algorithm can be easily used in the field of image matching. We can determine how similar two images are by the number of image matching points found. The steps of SIFT algorithm can be shown in Figure 1.

Good stability and ease of computation are the obvious advantages of its algorithm, but the computation of each feature point requires a feature value of significant length. This also makes its algorithm more time-consuming, and the application has some limitations in dealing with real-time problems [19].

Similarity algorithms based on specific theories, this class of similarity algorithms, are mostly based on the graph structure. Assuming that the segmented images whose regions are all independent and unique, then the images can be described correspondingly as graph structures by attribute feature extraction and description of region space relations. Thus, the similarity calculation of graph structures can represent the similarity between images to some extent. To calculate the graph similarity, the literature proposes the theory that if two graph nodes have similar neighboring nodes, then these two nodes are also similar. In other words, a portion of the similarity of these two nodes is propagated to their respective neighbors. Thus, after a series of iterations, this similarity is propagated throughout the graph, and thus, we can arrive at the final overall similarity of the two graphs. For the improved algorithm, the step-by-step diagram is shown in Figure 2.

Therefore, according to this similarity view, a similarity matrix can be constructed for both graphs, treating the sim-

ilarity between both two nodes in the graph as an element of the matrix, and then a final formula for calculating the similarity can be derived by the correlation operation of the matrix. The process can be as follows.

$$F = \frac{1}{n} \sum_{i=1}^n (X_i - Y_i)(X_i + Y_i) + e^{-((x-\mu)/\sigma)}. \quad (3)$$

The graph matching view of association graphs was first proposed for solving matching between homomorphic relational structures. He proposed to build an association graph between two graphs by a preset function mapping and to find the maximal subcluster in the association graph for matching, first to build an association graph G of two graph structures; the nodes in G consist of G_1 in G_1 and V in G_2 that meet certain compatibility conditions together, and it represents a node pair. The two nodes are judged to be adjacent. The nodes and the relationships between nodes of the association graph can be derived by filtering. A maximal subcluster is a subcluster with the maximum number of node pairs in the association graph, i.e., it means that any vertex in a subgraph of the association graph is connected to the other vertices of that subgraph. The structural similarity between two images can be obtained by the higher of the maximum subcluster and thus the similarity of the images. The structure of the image is shown in Figure 3.

Concerning the principle that the overall similarity is equal to the sum of the similarity of the parts, the node iterative matching algorithm proposes the idea that the image matching error is equal to the sum of the node error and the corresponding edge error [20]. The algorithm iterates the matching process for times, which is determined by the number of graph nodes. The node matching between the two graphs with the smallest matching error can be obtained by subiterations, and its matching error is calculated. The first several matrices must be defined to represent the error difference of the nodes, the possible node matching pairs, etc., respectively. Then the formula for calculating the node matching error and the formula for calculating the edge matching error must also be defined, and finally, the similarity of the two graphs is determined based on the magnitude of the matching error. This graph matching algorithm is applied to image similarity recognition, and good experimental results are achieved. The definition of image similarity is the starting point of various algorithms in current image research, and understanding the concept of image similarity is the only way to define similarity algorithms and formulas that match the characteristics of the image objects understudy from a practical point of view. Starting from the concept of similarity, this paper introduces various algorithms and their classifications since the study of image similarity, including traditional similarity theories based on image pixel gray values and picture feature points, and various new similarity theories based on graph structure. The latter theory differs from the traditional theory in that it defines similarity from a new perspective and proposes a series of quite fruitful formulas for similarity calculation, which are applied to practical research.

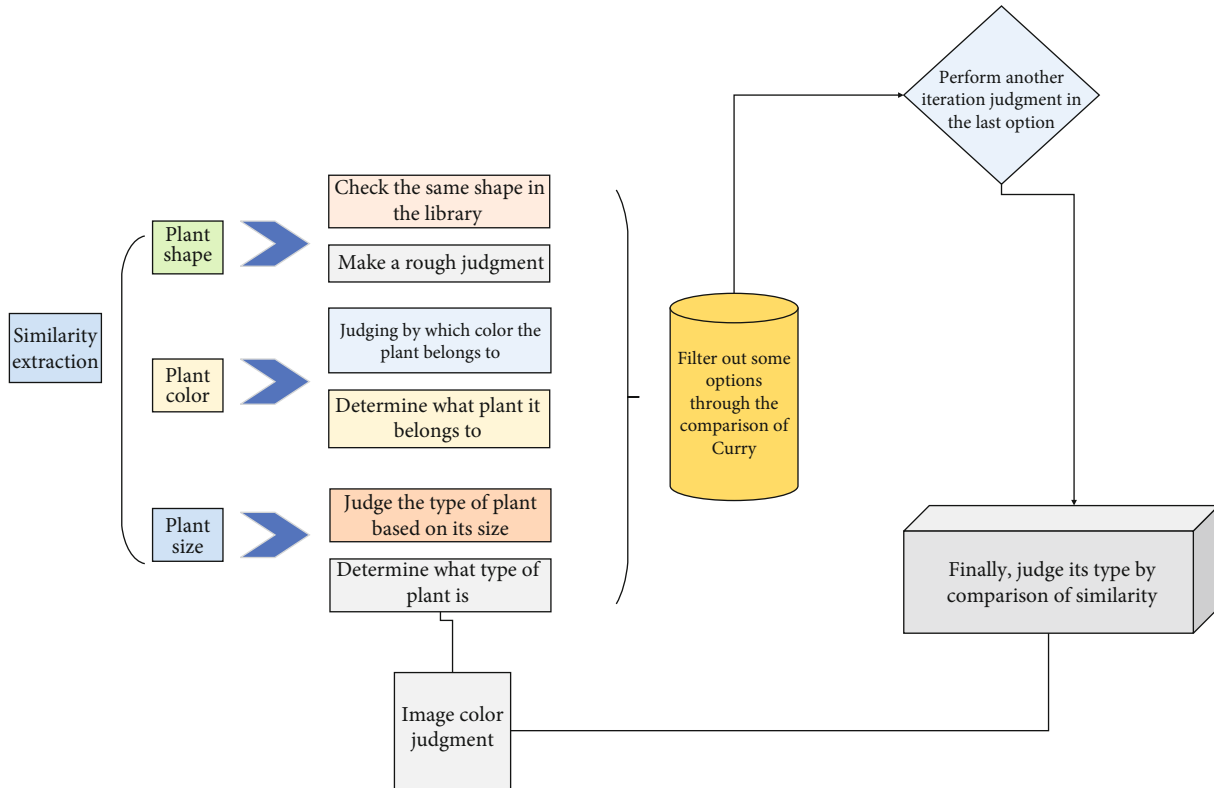


FIGURE 1: Step-by-step diagram of the SIFT algorithm.

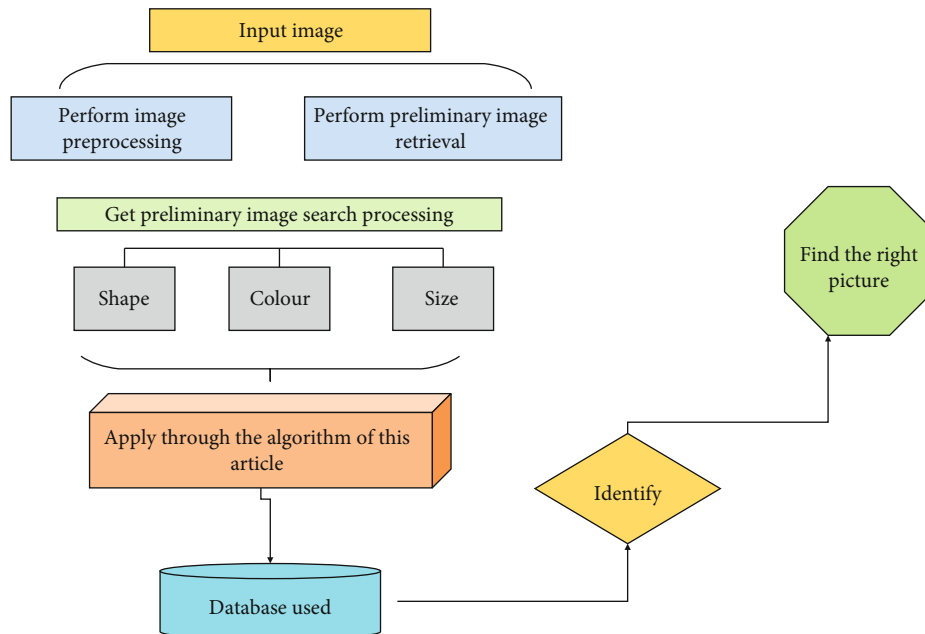


FIGURE 2: Flow chart of the improved algorithm.

Region consists of a pixel and its neighbors. By neighborhood of a pixel, it meant the set consisting of a pixel and its neighboring pixels, and there are usually several kinds of neighborhoods, neighbors, and diagonal neighborhoods. At the beginning of an image, a pixel can be consid-

ered to represent a region, and then the neighbors of the pixel are merged to form a region based on the judgment of certain similarity criteria between pixels. Thus, any two pixels within a region must have some aspect of similarity and must be connected. In other words, a region is a

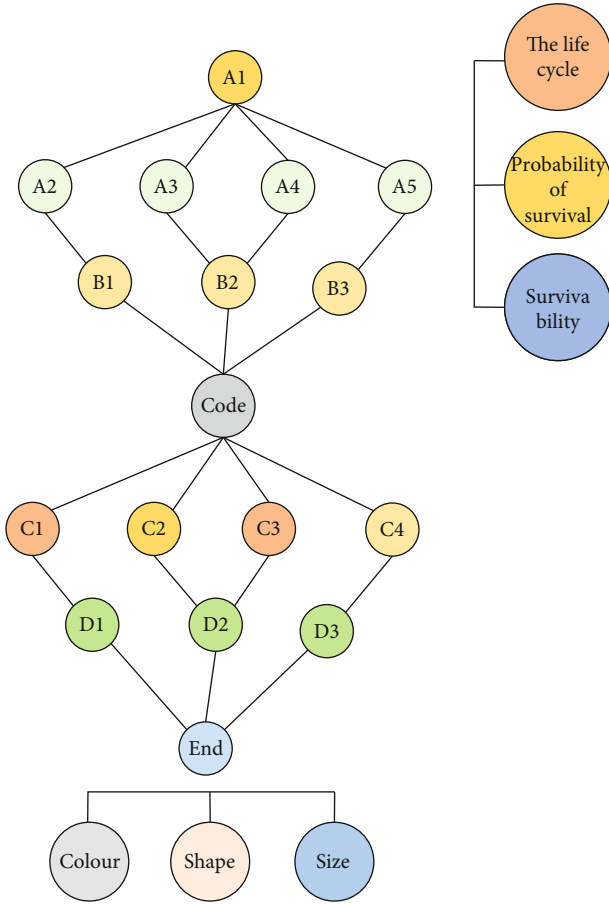


FIGURE 3: Structural similarity diagram.

collection of pixels that are similar and connected, and its image is made up of a matrix, as shown in Figure 4.

4. Experimental Results and Analysis

4.1. Experimental Results. In this experiment, the similarity recognition performance of the proposed method is experimentally validated and analyzed on the walk dataset. First, we compare the effectiveness of the fused features proposed in the paper with some single features in performing image similarity recognition to verify the effectiveness of the fused features, and N time, and we also verify the performance of the diffusion process in combining these different features for similarity recognition. Figure 5 shows the comparison of the respective ARP metrics for dataset t when similarity recognition returns image $NL = 100$ using different image features with and without the diffusion process optimization.

From the results in the figure, we can see that the fusion of multiple features improves the descriptive and discriminatory ability of features compared to single features, which significantly improves the picture similarity recognition accuracy; in addition, the distance optimization based on the diffusion process can effectively improve the similarity recognition performance, especially when combined with the multifeature fusion, which shows superior performance. This is not difficult to understand. Taking single-feature

1	2	0	8	9	2	1
3	7	3	1	2	7	9
6	5	7	6	1	7	5
0	4	0	6	1	3	9

FIGURE 4: Image matrix diagram.

LDP as an example, due to its weak feature description capability, it can only obtain an average similarity recognition accuracy of 42.7%, which leads to the inaccuracy of the similarity relationship in the initial distance matrix obtained, making the space for distance optimization by the diffusion process on this basis small. And with the gradual improvement of feature expression ability, the obtained initial similarity relationship also gradually becomes accurate, so the diffusion process based on a more accurate nearest-neighbor relationship can play its advantage and role well. Secondly, we compare the similarity recognition performance of this paper's method with more than ten other similar image similarity recognition methods based on multifeature fusion and list when 20 images are returned. The average similarity recognition accuracy of various methods is listed, and Figure 6 shows the comparison of the variation of similarity recognition performance metrics ARP and ARR when the number of images returned more from 10 to 100.

As can be seen from the results, the highest average similarity recognition accuracy can be obtained almost by using only the fused features proposed in this paper, in addition to the original method. It is worth emphasizing that the similarity recognition model proposed in the literature not only employs the fusion of region-based multiscale curvelet texture features and primary color features but also uses a fusion matching model in the picture matching process based on the Most Similar Highest Priority (MSHP) fusion matching pattern. Even so, the ARP obtained when the fused features proposed in the text are effectively combined with a diffusion-based process is nearly 4% higher than this method.

To appeal this algorithm, an improved algorithm is proposed that uses a reordering based on the matching target localization method AML (approximate max-pooling localization). Whereas the AML process generates a large number of candidate regions, whitening each region feature vector would greatly increase the computational cost. Therefore, the R-MAC method chooses max-pooling to aggregate multiple regions on the feature map. In the SCAD method, the image similarity recognition performance is improved by concatenating the aggregated vectors formed by both max-pooling and avg-pooling; although the two different aggregation methods can complement each other to a certain extent, they also bring about an increase in the dimensionality of the feature vectors. Figure 7 is the data plot of the increase in the dimensionality of the feature vectors.

Value takes about 0.95 seconds on average, and the increase does not bring much increase in time; on the Cor-el5k dataset, when the diffusion process is optimized for an

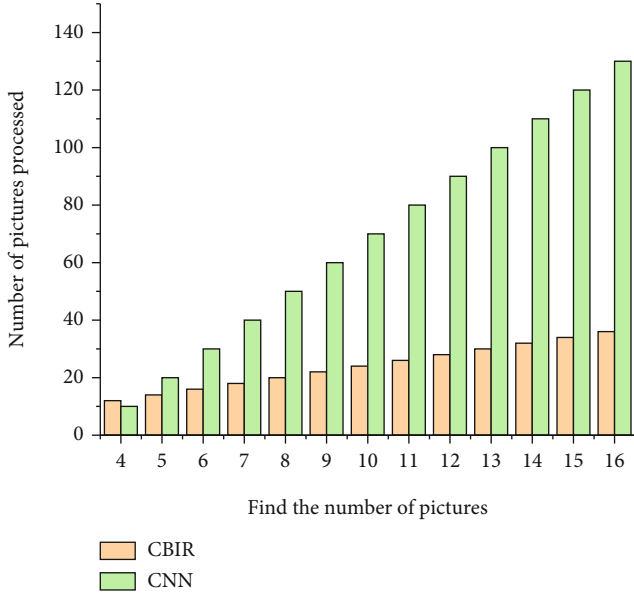


FIGURE 5: Comparison of efficiency after algorithm improvement.

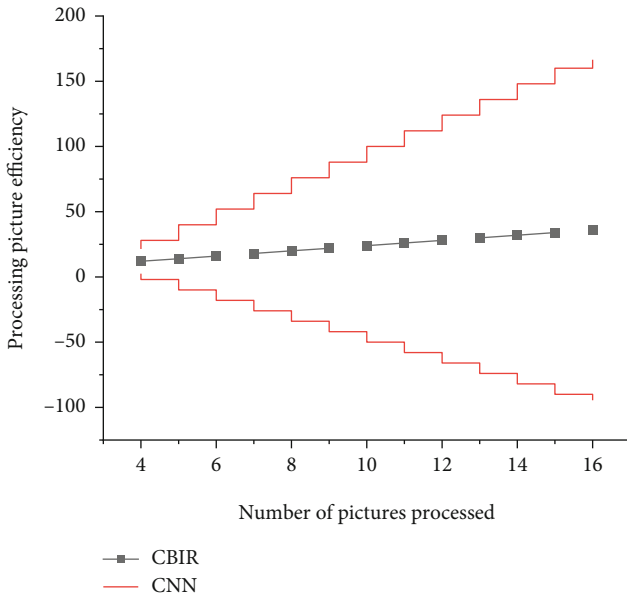


FIGURE 6: Comparison of similarity feature finding efficiency.

intimacy matrix of 5000×5000 , the time required is about 5.15 seconds, while when $K = 8$ the time required is 10.55 seconds, i.e., it increases to more than two times; for both the CorellOk and GHIMIOk datasets, the $(10000 \times K)$ among the diffusion process of the intimate matrices, the time required for $K = 5$ is 15.09 seconds, and it will rise to 92.77 seconds, which is more than 5 times higher. It can be seen that for a larger dataset, the growth of the parameter t of the diffusion process results in a larger time N consumption.

4.2. Analysis of Experimental Results. This experiment is based on multifeature fusion and diffusion process sorting

for a lofty image similarity recognition framework to enhance picture similarity recognition from two aspects: image feature representation and image matching. Disobedient first, this method proposes a hit-effective multifeature representation method that effectively fuses low-level visual features such as color, texture, and shape of an image into one. The feature description method has better image portrayal and discrimination ability and has certain stability and robustness for image description of different visual contents and features, which can effectively alleviate the existence of “feature semantic differences” in image similarity recognition; secondly, combined with the hit-effective fusion features, we adopt a diffusion-based process. Secondly, we use a distance optimization method based on the diffusion process in the image: matching to overcome the limitations of traditional base distance image matching. This experiment is aimed at solving the key problems in CNN feature-based image security similarity recognition. Feature extraction is a fundamental step in various vision tasks such as image classification and image similarity recognition, and CNN features are outstanding in image similarity recognition; however, convolutional neural network models usually consist of multiple convolutional layers, fully connected layers, and pooling layers, which have a complex structure and are computationally intensive in extracting features, which will lead to a large computational and storage overhead on the end device. Outsourcing the CNN extraction task to a cloud server can greatly reduce the computational overhead on the client-side, and there are several schemes for outsourcing feature extraction. Their different efficiency comparison graphs for different schemes are shown in Figure 8.

In the proposed CNN feature-based picture similarity recognition scheme, the image encryption is consistent with Section 3, and based on the existing analysis, it is known that the image content is secure under this encryption scheme. However, the features of the image may leak the image content, so the security of the image features is mainly discussed below. Security of image features in the index construction phase is as follows: during the index construction, based on E2LSH, the hash result of the subdimension of image CNN features is calculated, and a 2-dimensional vector is mapped to a value, and since image CNN features tend to be high-dimensional, the process is considered irreversible and does not leak image feature information in the index construction phase image features that are privacy-secure in the security of the labeled image pattern: The secure index structure designed in this paper can avoid the cloud server to infer the similarity and category of the cloud images by itself if the cloud server wants to accurately determine whether the target images I_1 and I_2 are similar or not; then, we must know the image CNN features.

However, under the semihonest model, it is considered that two cloud servers are not complicit. Also, the target image pattern is protected because the indexing based on dimensional division reduces the probability of similarity of images under the same bucket, and the indexing from the subdimensions of image features known to the cloud server does not determine whether the images are similar

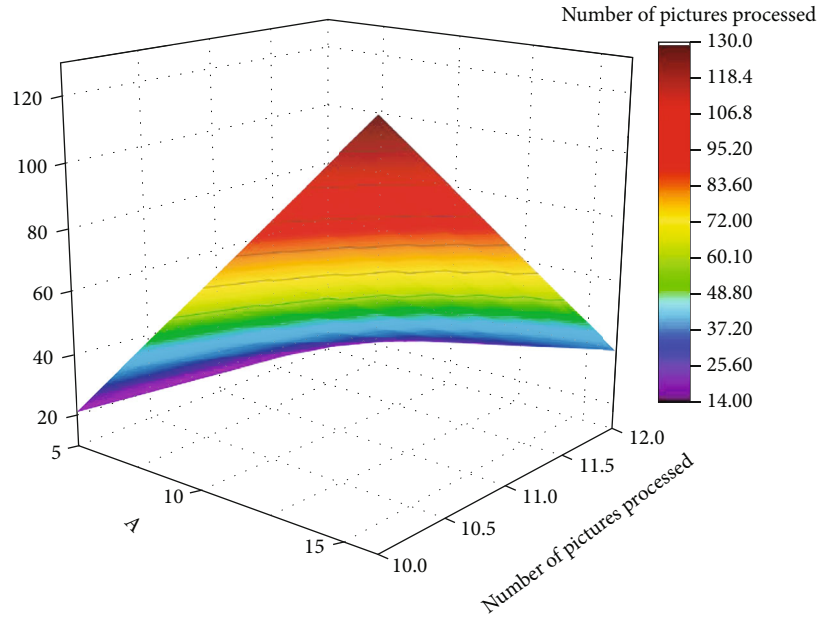


FIGURE 7: Data plot for increasing dimensionality of feature vectors.

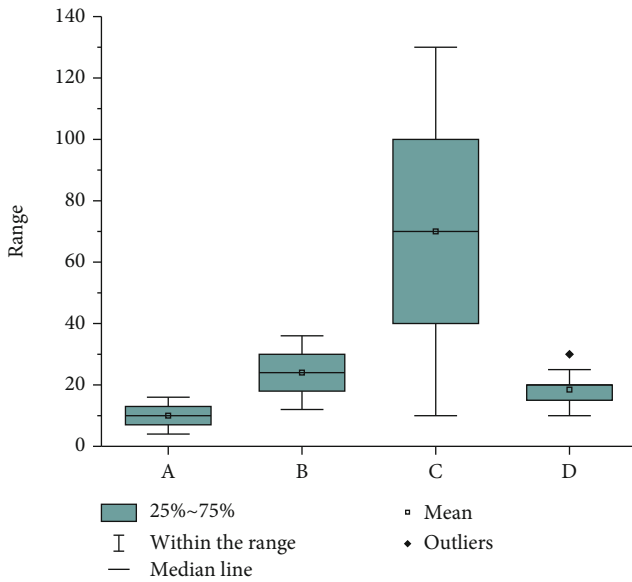


FIGURE 8: Comparison of different efficiency of different schemes.

or not. When querying users to make more queries, the server may perform statistical analysis of historical queries, causing leakage of data access patterns, and privacy protection of such access patterns is not considered in this paper. A simple solution is to send a large number of additional dummy queries per query for the server to delegate the task of sorting the candidate results to the querying user, but this increases the burden on the querying user and the number of interaction rounds. In this case, a trade-off needs to be made between the security and efficiency of the similarity recognition scheme.

5. Conclusion

A semantic-based similarity metric is proposed by introducing a theory, which is completely different from the traditional image similarity recognition model. On the one hand, the method maps the original feature space to a new semantic space through the embedding of semantic concepts and replaces the traditional full-feature space-based approach. The method selects different samples according to the structure and distribution of the original image attribute data, which is more suitable for describing their different semantic features, thus enhancing the discriminability between samples in the semantic feature space; on the other hand, to overcome the limitations of the traditional point-pair-based metric in the distance metric, the local nearest-neighbor relationship is considered in the semantic similarity-based metric, which extends the original similarity metric between image point pairs to the similarity metric between image sets and collections, by capturing the underlying data stability structure to more accurately describe the similarity between images and further improve the stability and robustness of image similarity relationship description. This method shows its good performance in completing face similarity recognition tasks that include low-level features and natural image similarity recognition tasks that employ depth features. The effectiveness of the semantic-based distance metric proposed in this chapter is verified through a large number of experiments, and at the same time, the semantic similarity metric-based similarity recognition method proposed in this paper has significant advantages over some hand-designed feature-based and deep learning-based picture similarity recognition methods, which provides a new idea for semantic-based picture similarity recognition. In this method, although the feature extraction phase

done in the offline stage is completed, both the higher feature dimension and the larger number of picture samples make the process of constructing the set of feature semantic descriptions in the AFS framework more time-consuming, both of which have been improved with the improvements of this algorithm.

The development direction of the algorithm used in the article is mainly to improve the recognition efficiency and recognition speed of the algorithm. The improvement of the algorithm is that there are 4 main steps: (1) extremum detection in scale space, search for images on all scale spaces and identify potential through Gaussian differential function of constant interest in scale and choice; (2) feature point location at each candidate location, a fine fitting model is used to determine the location scale, and the selection of key points are based on their degree of stability; (3) the feature direction assignment is based on the local gradient direction of the image and is assigned to each key point position in one or more directions. All subsequent operations are to transform the direction, scale, and position of the key point to provide the invariance of these features; and (4) feature point description in the neighborhood around each feature point, the local gradients of the image are measured on selected scales, and these gradients are transformed into a representation that allows relatively large local shape deformation and illumination transformation.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Y. Chen, L. Liu, J. Tao et al., "The improved image inpainting algorithm via encoder and similarity constraint," *The Visual Computer*, vol. 37, no. 7, pp. 1691–1705, 2021.
- [2] L. Zou, L. T. Song, T. Weise et al., "A survey on regional level set image segmentation models based on the energy functional similarity measure," *Neurocomputing*, vol. 452, pp. 606–622, 2021.
- [3] W. Sun, G. Yang, K. Ren et al., "A label similarity probability filter for hyperspectral image postclassification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6897–6905, 2021.
- [4] Y. Ge, Z. Yang, Z. Huang, and F. Ye, "A multi-level feature fusion method based on pooling and similarity for HRRS image retrieval," *Remote Sensing Letters*, vol. 12, no. 11, pp. 1090–1099, 2021.
- [5] S. Vetova, "A comparative study of image classification models using NN and similarity distance," *International Journal of Electrical Engineering and Computer Science (EEACS)*, vol. 3, pp. 109–113, 2021.
- [6] A. Higaki, T. Kurokawa, T. Kazatani et al., "Image similarity-based cardiac rhythm device identification from X-rays using feature point matching," *Pacing and Clinical Electrophysiology*, vol. 44, no. 4, pp. 633–640, 2021.
- [7] H. J. Kwon, D. H. Shin, and K. Chung, "PGGAN-based anomaly classification on chest X-ray using weighted multi-scale similarity," *IEEE Access*, vol. 9, pp. 113315–113325, 2021.
- [8] R. S. Somai and P. J. B. Hancock, "Exploring perceived face similarity and its relation to image-based spaces: an effect of familiarity," *Journal of Vision*, vol. 21, no. 9, pp. 2149–2149, 2021.
- [9] M. J. Kim, "Pothole region extraction based on similarity evaluation scale classification using image processing," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 6, pp. 871–877, 2021.
- [10] J. Kang and C. Lim, "Using similarity based image caption to aid visual question answering," *The Korean Journal of Applied Statistics*, vol. 34, no. 2, pp. 191–204, 2021.
- [11] H. Li, R. Cong, S. Kwong, C. Chen, Q. Xu, and C. Li, "Stereo superpixel: an iterative framework based on parallax consistency and collaborative optimization," *Information Sciences*, vol. 556, pp. 209–222, 2021.
- [12] J. Li, "Deep semantic cross modal hashing based on graph similarity of modal-specific," *IEEE Access*, vol. 9, pp. 96064–96075, 2021.
- [13] Y. I. Daradkeh, I. Tvoroshenko, V. Gorokhovatskyi, L. A. Latiff, and N. Ahmad, "Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic," *IEEE Access*, vol. 9, pp. 13417–13428, 2021.
- [14] J. Zhang, W. Xia, Q. Jin, and X. Gao, "A 2D/3D non-rigid registration method for lung images based on a non-linear correlation between displacement vectors and similarity measures," *Journal of Medical and Biological Engineering*, vol. 41, no. 3, pp. 331–342, 2021.
- [15] G. Li, Y. Lin, and X. Qu, "An infrared and visible image fusion method based on multi-scale transformation and norm optimization," *Information Fusion*, vol. 71, pp. 109–129, 2021.
- [16] P. Singh, P. N. Hrisheeksha, and V. K. Singh, "CBIR-CNN: content-based image retrieval on celebrity data using deep convolution neural network," *Recent Advances in Computer Science and Communications*, vol. 14, no. 1, pp. 257–272, 2021.
- [17] X. Zenggang, T. Zhiwen, C. Xiaowen, Z. Xue-min, Z. Kaibin, and Y. Conghuan, "Research on image retrieval algorithm based on combination of color and shape features," *Journal of Signal Processing Systems*, vol. 93, no. 2-3, pp. 139–146, 2021.
- [18] S. Kollem, K. R. Reddy, and D. S. Rao, "Improved partial differential equation-based total variation approach to non-subsampled contourlet transform for medical image denoising," *Multimedia Tools and Applications*, vol. 80, no. 2, pp. 2663–2689, 2021.
- [19] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Comparison of full-reference image quality models for optimization of image processing systems," *International Journal of Computer Vision*, vol. 129, no. 4, pp. 1258–1281, 2021.
- [20] E. H. Houssein, B. E. D. Helmy, A. A. Elngar, D. S. Abdelminaam, and H. Shaban, "An improved tunicate swarm algorithm for global optimization and image segmentation," *IEEE Access*, vol. 9, pp. 56066–56092, 2021.