Content-Based E-Commerce Image Classification Research

XIAOLI ZHANG
School of Business, Xuchang University, Xuchang 461000, China
e-mail: wuhen6020@163.com

This work was supported by the Soft Science Research Program of Henan Province: Research on the Linkage Development Path between characteristic industries of Henan Province and E-commerce under Project 124.

ABSTRACT The 21st century is the era of big data in the Internet. Online shopping has become a trend, and e-commerce has developed rapidly. With the exponential increase of the amount of commodity image data, the management of massive commodity image database restricts the development of e-commerce to some extent. In order to effectively manage goods and improve the accuracy and efficiency of product image retrieval, this paper uses content-based methods to classify e-commerce images. Aiming at the problems of insufficient classification accuracy and long classification training time in e-commerce image classification, an adaptive momentum learning rate based LBP-DBN training algorithm–AML-LBP-DBN and commodity image classification method based on image local feature multi-level clustering and image-class nearest neighbor classifier are proposed. By simulating the commodity identification dataset RPC, the results show that the proposed method has obvious advantages in the classification training time and classification accuracy of e-commerce images.

INDEX TERMS E-commerce, commodity image classification, adaptive training, local feature multi-level clustering.

I. INTRODUCTION

Big data is one of the mainstream topics in the current information age [1]. Our living environment is full of vast amounts of information, and people live in the ocean of information. Among the kinds of information that people receive, the most intuitive and most important thing is the image information received through the vision [2]. More than 70% of the human receiving information is received by the visual [3]. In Internet e-commerce, the user is faced with the product itself, but the product image information and some simple annotations, such as the name, origin, size, price and other basic information of the product, so the image becomes the main information for the transmission of commodity information carrier. Image is the reproduction of people’s visually perceived information. Compared with sound and text, it contains a large amount of information, a more direct transmission method, and a simpler receiving method [4], [5]. Image classification is a technique that allows a computer to imitate human vision to process, analyze, and understand image information, and to understand various image information and divide it into information similar to the human brain [6]. Therefore, the recognition and processing of image information has attracted more and more research and attention. In the field of e-commerce, automatic classification of product images can provide fast transaction query for both parties, determine the placement strategy of products and intelligent recommendation of products of interest to users, thus effectively improving the overall efficiency of the e-commerce market. This is an urgent requirement for e-commerce intelligence. Therefore, the research on e-commerce image classification has very important practical significance.

There are many experts and scholars in the field of image classification [7]. In the 1970s, a model commonly used in text categorization appeared in the study of computer vision. This is the Bag of Words (BoW) [8]. In the 1980s, some methods of image classification, such as the word package model, still studied the overall shape of the image, but a fully industrialized system for identifying objects has been developed [9], [10]. In the 1990s, image classification mainly studied global features, which can use the overall information as the description information of the image. The more common features are color features and texture features [11], [12].
For example, the feature matching method based on color histogram proposed by Ballard and Seain [13]. Due to the single use of global features, the background recognition effect is very poor. In order to overcome the shortcomings of global features, scholars have begun to pay attention to the component information of objects, which are called local feature descriptors. For example, the scale-invariant feature transform (SIFT) algorithm proposed by Yanxi and Juan [14], the Histogram of Oriented Gradient (HOG) algorithm proposed by Dalal and Triggs [15]. The local feature descriptor ignores the spatial information between the images, and later scholars began to combine global features and local features. For example, Lazebnik et al. proposed Spatial Pyramid Matching (SPM) [16], using the word package model for image blocks of different scales, and finally integrating the entire image blocks to obtain the final feature representation of the image. In feature extraction, scholars began to select features, which made sparse coding appear more and more in image classification. For example, Qiang Chen et al. proposed the use of vector quantization coding [17] and Fisher vector coding [18] to replace the original single vector quantization coding. In 1990, image classification showed a new direction, and Schapire proved that a number of relatively weak models can be combined to obtain higher than all single models [19]. In 1996, Breiman proposed a Bagging (Bootstrap Aggregating) algorithm [20], which is a sampling algorithm with a return, training different sets of different classifiers after the difference of sampling results, and finally using voting to get the final result of. There are also many scholars in China who have studied image classification. Huilin [21] constructed a new convolutional neural network structure based on principal component analysis (PCA) whitening. Hui et al. [22] proposed a medical image classification algorithm based on RBF neural network. The feature information of the pixel is used as the training sample to train the RBF neural network, and the classification of the RBF neural network image after training is used to identify different The classification result gives different RGB values for display [23]. Aiming at the serious quantization error of vector quantization coding, Yongwei et al. [24] proposed an image classification method based on deep learning feature coding model.

At this stage, the number of high-definition images has increased significantly, and massive data image libraries have emerged. Traditional methods are far from meeting the needs of large databases [25]. As an image of one of the important components of multimedia information, it has rich visual and semantic information, and it is difficult to fully express the image content in a simple language. The incomplete and accurate description of the image information is a key issue encountered in the image classification process.

This paper uses content-based and deep learning methods to classify e-commerce images. Content-based image classification has been widely studied. For example, Baidu’s mapping is mainly for image classification and retrieval of faces, landscapes, stars, etc. [26], and Tutu search mainly classifies and retrieves clothes [27]. At the same time, effective classification of images can greatly improve the efficiency of e-commerce image retrieval, and expand and enrich the image database. In 2013, Baidu established the Institute of Deep Learning (IDL), the most important part of the Baidu Institute, specializing in applied research in the field of deep learning. Based on deep learning, they developed Baidu’s image intelligent recognition technology, which can find the products to be purchased through a photo, and can also realize AI searching through face recognition [28]. In December 2015, Baidu Deep Learning Institute developed the unmanned vehicle technology [29], which realized the automatic driving of unmanned vehicles in various complicated road conditions. In March 2014, Professor Tang Xiaou from the Chinese University of Hong Kong proposed a Gaussian Face-based algorithm for face recognition experiments on the LFW database with an accuracy of 98.52% [30], [31]. In June of the same year, Professor Tang Xiaou’s team conducted a series of experiments based on the Deep ID algorithm, which improved the face recognition accuracy rate to 99.55%, promoted the development of face recognition in China [32]. In summary, deep learning plays an important role in the field of image classification, especially when the era of big data is approaching; combining image classification with deep learning is of great significance.

In order to solve the problems of insufficient classification accuracy and long training time in e-commerce image classification, this paper proposes a content-based e-commerce image classification method. In the traditional LBP-DBN method, the LBP algorithm is introduced to extract the texture features of the local structure of the image, and the extracted texture features are used as the input of DBN, which effectively overcomes the influence of external factors on the features. Aiming at the problem that learning rate selection is difficult in the greedy learning stage, the training efficiency of the network is not high. An LBP-DBN training algorithm based on adaptive momentum learning rate, AML-LBP-DBN, is proposed. After the adaptive improvement, compared with the original algorithm, the classification accuracy of the AML-LBP-DBN algorithm proposed in this paper is improved, and when the iteration training reaches 3000 times, the classification accuracy of the network can reach 99.21%. It is 0.93% higher than the original algorithm. Moreover, in order to improve the image classification speed, this paper proposes a product image classification method based on image local feature multi-level clustering and image-class nearest neighbor classifier. Through experiments, it is found that the number of cluster centers will affect the classification accuracy rate and classification speed.

II. METHOD

A. CONTENT-BASED IMAGE CLASSIFICATION PROCESS

Effective management of e-commerce images is an urgent problem to be solved. Content-based image classification technology can solve this problem, providing more efficient
and accurate product management and retrieval technology to better serve customers. The content-based image classification technology mainly includes the following steps:

1) Image preprocessing: Denoising, smoothing, correcting, enhancing and removing backgrounds, etc., in order to obtain images with sharper edges, more prominent edges and stronger contrast, thereby improving the accuracy and efficiency of classification.

2) Image feature extraction and feature selection: various features are extracted from the image, such as space, texture, shape and color. Generally, there are many image features extracted. If feature selection is not performed, the efficiency and accuracy of classification will be reduced. Therefore, it is necessary to make feature selection.

3) Classification model establishment: According to the characteristics and characteristics of the classified images, appropriate algorithms are used to put together the same type of images.

B. IMAGE CONTENT DESCRIPTION (LBP TEXTURE FEATURE EXTRACTION)

1) LBP OPERATOR

DBN is a deep network structure, usually composed of several layered RBMs. The image features are extracted layer by layer through unsupervised training of each RBM layer structure. LBP is a classical texture feature description operator, which has strong rendering ability for local texture features of images. The operator has low complexity, simple calculation, high robustness, rotation invariance and illumination invariance features.

a: RECTANGULAR NEIGHBORHOOD LBP OPERATOR

In the original LBP method, a rectangular neighborhood of 3 × 3 size is usually defined as an LBP operator. Set the gray value of the pixel in the center of the neighborhood to \( g_c(x_c, y_c) \), and compare it with the surrounding pixels as the threshold. The point with the large gray value is marked as 1, otherwise the flag is 0. Then arrange the numbers in a uniform direction to get an 8-bit binary number, as shown in Figure 1. This binary number is recorded as the LBP value of the pixel center point, which is used to reflect the texture feature information of the neighborhood.

The specific calculation process is as follows:

\[
LBP(x_c, y_c) = \sum_{p=0}^{p=1} 2^p s(i_p - i_c) \tag{1}
\]

where \((x_c, y_c)\) represents the two-dimensional coordinate index of the central pixel, \(p\) is the \(p\)-th pixel in the rectangular neighborhood, \(i_c\) is the gray value of the central pixel, and \(i_p\) is the gray value of the surrounding pixels in the neighborhood. \(s(x)\) is a symbolic function:

\[s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{else} \end{cases} \tag{2}\]

b: CIRCULAR NEIGHBORHOOD LBP OPERATOR

Considering the shortcomings of the original rectangular LBP operator in practical applications, Ojala T et al. proposed a circular neighborhood LBP operator. This operator improves the original LBP immutable pixel space and changes it from a rectangle to a circular variable neighborhood, giving it a rotation-invariant characteristic, which not only expands the calculation range of the LBP operator, but also effectively avoids the center. The loss of spatial structure information around the pixel, and greatly improve the accuracy, can display the texture features of the image from multiple scales.

The improved circular operator is generally represented by the symbol \(LBP_{p-R}\), as shown in Figure 2, where \(P\) represents the number of pixels located on a circular neighborhood around the central pixel point, and \(R\) represents the radius of the circular neighborhood. In binary sorting, each circular LBP has \(2p\) arrangements and grows geometrically as \(P\) increases. To this end, Ojala T has proposed a uniform mode LBP operator \(LBP_{p-2}\). The uniform mode LBP requires that the total number of 0 ~ 1 or 1 ~ 0 transitions in the aligned binary sequence does not exceed two times. The type of arrangement is reduced from \(2^P\) to \((P - 1) + 2\), which not only reduces the feature dimension, but also reduces the influence of high frequency noise and error factors in high dimensional space.

2) ALGORITHM DESIGN

The traditional LBP-DBN method introduces the LBP algorithm to extract the texture features of the local structure of the image, and uses the extracted texture features as the input of the DBN, effectively overcoming the influence of external factors on the features. However, in the greedy learning stage, there is still a problem of difficulty in learning rate selection. Under the traditional fixed learning rate, each weight adjustment of the network does not always proceed toward the direction in which the loss function decreases, resulting in low training efficiency of the network [33]. In view of the above problems, this section proposes an LBP-DBN training algorithm based on adaptive momentum learning rate, AML-LBP-DBN. The pseudo code design of the RBM adaptive training phase is as follows:

Input: initialization weight \(W_{ij}\), offset \(b_i, c_j\), number of iterations \(epochs\), training batches \(batches\), number of training samples per batch \(batchsize\), momentum \(m\), learning rate \(\eta\) and other parameters.

for \(R = 1, 2, 3, \ldots, epochs\) do
  for \(L = 1, 2, 3, \ldots, batches\) do
1) Calculate the random unit state of the visible and hidden layers in each RBM:

\[ P(h_j = 1|v) = \sigma(c_j + \sum_i W_{ij}v_i) \]

\[ P(v_i = 1|h) = \sigma(b_i + \sum_j W_{ij}h_j) \quad (3) \]

Wherein

\[ \sigma(x) = \frac{1}{1 + \exp(-x)} \quad (4) \]

2) Derivation of the three parameters \( W_{ij}, b_i, c_j \) from the log \( p(v|\theta) \) likelihood probability log of the visible layer unit \( v \):

\[ \frac{\partial \log p(v|\theta)}{\partial W_{ij}} = (v_i h_j)_{data} - (v_i h_j)_{model} \]

\[ \frac{\partial \log p(v|\theta)}{\partial b_i} = (v_i)_{data} - (v_i)_{model} \]

\[ \frac{\partial \log p(v|\theta)}{\partial c_j} = (h_j)_{data} - (h_j)_{model} \quad (5) \]

3) Calculate the weight and offset increment of the RBM training:

\[ \Delta W_{ij} = m^* \Delta W_{ij} + \eta^*(v_i h_j)_{data} - (v_i h_j)_{recon}) / \text{batchsize} \]

\[ \Delta b_i = m^* \Delta b_i + \eta^*(v_i)_{data} - (v_i)_{recon}) / \text{batchsize} \]

\[ \Delta c_j = m^* \Delta c_j + \eta^*(h_j)_{data} - (h_j)_{recon}) / \text{batchsize} \quad (6) \]

4) Update weights and offsets:

\[
\begin{align*}
W_{ij}' &= W_{ij} + \Delta W_{ij} \\
b_i' &= b_i + \Delta b_i \\
c_j' &= c_j + \Delta c_j
\end{align*}
\]

End for

5) Calculate the reconstruction error increment and weight update direction before and after the iteration, and adaptively adjust the momentum term and the learning rate as the training parameters of the lower RBM:

\[ error = error + \sum (v - v')^2 \]

\[ \Delta = ((v_i h_j)_0 - (v_i h_j)_k) \cdot ((v_i h_j)_0' - (v_i h_j)_k') \quad (8) \]

if \( R \geq 2 \) do

\[ \Delta E = error(R) - error(R - 1) \]

\[ m' = \begin{cases} 
\alpha m, & \text{if } \Delta E > 0 \\
\beta m, & \text{if } \Delta E < 0 
\end{cases} \quad (9) \]

\[ \eta' = \eta(1 + \gamma), \quad \text{if } \Delta > 0 \]

\[ \eta(1 - \gamma), \quad \text{if } \Delta < 0 \quad (10) \]

end if

output: Weight and offset after training

\section{C. LOCAL FEATURES MULTI-LEVEL CLUSTERING AND IMAGE–CLASSICAL PRODUCT IMAGE CLASSIFICATION OF NEAREST NEIGHBORS}

In order to improve the image classification speed, this paper proposes a product image classification method based on image local feature multi-level clustering and image-class nearest neighbor classifier. The image-to-class distance is calculated by the local feature multi-level clustering method. Firstly, the image and image, the class is represented as an unordered set of local descriptors, then the class feature space is divided by multi-level clustering method, and then the image–class distance is calculated according to different levels.

1) MULTI-LEVEL CLUSTERING

The multi-level clustering method is used to recursively classify the feature space, and each level of clustering uses the K-value clustering method:

1) Randomly select K local descriptors as the cluster center, which is recorded as:

\[ \{m_k\}, \quad k = 1 \ldots K \quad (11) \]

2) Calculate the distance between all local descriptors \( x_i \) and the cluster center, find the cluster center with the smallest distance of each local descriptor \( L_2 \), and all the descriptors closest to the cluster center form a cluster, that is, \( x_j \) belongs to the cluster C (i):

\[ C(i) = \arg \min_{1 \leq k \leq K} |x_i - m_k|^2, \quad i = 1, \ldots n \quad (12) \]

where: \( n \) is the number of local feature descriptors;
3) For each cluster, recalculate the cluster center.

\[
m_k = \frac{\sum_{i \in C(i) = k} x_i}{n_k}, \quad k = 1 \ldots K
\]  

(13)

\[
W(C) = \sum_{k=1}^{K} \frac{1}{K} \sum_{C(i) = k} \sum_{C(j) = k} \| x_i - x_j \|^2
\]

\[
= \sum_{k=1}^{K} n_k \sum_{C(i) = k} \| x_i - x_j \|^2
\]

(14)

where: \( n_k \) is the number of descriptors belonging to the k-th cluster center;

4) Repeating step (2), that is, re-synthesis according to the distance between each re-calculated descriptor and each cluster center;

5) Repeat step (3) until the error of the cluster center reaches the threshold requirement (this value: 0.01), or the iteration reaches a certain number of times (the value of this paper: 1000).

In the first-level clustering, the K-means clustering method is used to divide the class feature space into several cluster regions. At the second level, the descriptors of each segment region are further divided into several sub-regions by K-means clustering method. This recursive clustering process can continue until the subregion is small enough. The number of rankings and the number of clusters per level should be set according to the size of the descriptor. For large-scale feature space, the K-means clustering process is inevitably very time-consuming. In order to reduce the computational complexity, a certain number of descriptors can be randomly selected for clustering, instead of all descriptors participating in clustering.

2) IMAGE-CLASS DISTANCE CALCULATION

It is set to perform two-level clustering on each image class feature space. The clustering center number of each level clustering is \( K_1, K_2 \), and the image-class distance can be calculated at different levels:

\( a: \) FIRST LEVEL DISTANCE

For each descriptor extracted from the test image, calculate the minimum distance from the first-level cluster center, and the sum of the minimum distances of the descriptors as the first-level distance between the images–classes:

\[
\sum_{k=1}^{K} \min_{d \in \chi_c^1} \| d_k - d \|^2
\]

(15)

where \( \chi_c^1 \) represents the first-level cluster center set. The time complexity for calculating the first-order distance is \( O(K_1^4 n) \), where \( n \) is the total number of test image descriptors.

\( b: \) SECOND LEVEL DISTANCE

Calculate the L_2 distance between the test image descriptor and the first-level cluster center, and find the first-level cluster center closest to the distance;

Calculate the L_2 distance between the test image descriptor and each secondary cluster center in the corresponding region in the nearest neighboring cluster center, and find the nearest secondary cluster center with its distance;

Calculate the sum of the distances from the descriptors of the test image to the L_2 distances of the nearest neighboring secondary cluster centers, that is, the second-level distance from the test image to the image class:

\[
\sum_{k=1}^{K} \min_{d \in \chi_c^2} \| d_k^l - d \|^2
\]

(16)

where \( \chi_c^2 \) represents the second-level cluster center set. Calculate the second-order distance time complexity as \( O((K_1 + K_2)^4 n) \).

\( c: \) THIRD LEVEL DISTANCE

Calculate the L_2 distance between the test image descriptor and the first-level cluster center, and find the first-level cluster center closest to the distance;

Calculate the L_2 distance between the test image descriptor and each secondary cluster center in the corresponding region in the nearest neighboring cluster center, and find the nearest secondary cluster center with its distance;

Calculating a minimum distance between the test image descriptor and each descriptor of the nearest neighbor secondary cluster center, and finding the descriptor whose distance is closest;

Calculating the sum of the L_2 distances of the third level of each descriptor of the test image is the third-order distance from the test image to the image class:

\[
\sum_{k=1}^{K} \min_{d \in \chi_c^3} \| d_k^l - d \|^2
\]

(17)

where \( \chi_c^3 \) represents the sub-set of descriptions to which the secondary cluster center belongs. Calculate the second-level distance time complexity as \( O((K_1 + K_2 + K_3)^4 n) \). \( K_3 \) is the average number of descriptors included in each cluster center of the second level.

III. EXPERIMENTS

A. DATA SOURCE

The experimental data in this paper is based on RPC, the largest product identification dataset in the academic world published by Defiance Technology Nanjing Research Institute. The product image range is up to 200 and the total image size is 83k. 200 categories of goods belong to 17 categories of commodities (such as instant noodles, paper towels, drinks, etc.). The number of items in the image will also be reflected, as shown in Figure 3.

B. CLASSIFICATION PERFORMANCE EVALUATION INDICATORS

In order to better classify images, it is often necessary to experiment with a variety of effective features or multiple classification methods to get the best classification
results. When different features or different features are combined and classified, the effect of classification is different. Therefore, effective evaluation criteria are needed to evaluate the quality of the classification, and the classification performance of the classifier is evaluated by the classification accuracy.

\[
\text{Classification accuracy}(\%) = \frac{\text{Correctly categorized sample size}}{\text{Total number of test samples}} \times 100\% \quad (18)
\]

C. EXPERIMENTAL PLATFORM

The server used is a Dell workstation with GPU Tesla K20c. The server trains a 20,000 training material, which takes about 3.5 hours. The steps to build the system platform are as follows:

- Install and deploy Caffe on Windows 64 and test it easily;
- Install and deploy Django and the database;
- Develop the corresponding program.

IV. RESULTS

Result 1: Comparison of AML-LBP-DBN algorithm and LBP-DBN algorithm

First, the image is \(4 \times 4\), and then the LBP operator is used to extract the feature of each sub-block. The radius \(R\) of the LBP operator is set to 1, and the number of sampled pixels \(P\) is 8. The \(LBP_{8-1}\) operator can generate 58 kinds. The equivalence model, plus a class that is not part of the equivalence model, is a total of 59 dimensions. The dimension of each sub-block feature extracted by LBP is 59 dimensions. Therefore, the 16 pieces of features can be connected to obtain a complete set of commodity texture features with a dimension of 944 dimensions, that is, the structure of the DBN network is 944-100-100-40. The DBN parameters are initialized as follows: the initial momentum \(m\) is 0.1, the initial learning rate \(\eta\) is 0.01, and the adaptive parameters are based on experience, where the momentum term increase coefficient \(\alpha\) and the decrease coefficient \(\beta\) are set to 0.8 and 0.6, and the learning rate gain factor \(\gamma\) and debuff \(\lambda\) are taken as 0.006 and 0.002, respectively, and the number of RBM iterations is 30. Train 280 sample data, one sample at a time. The loop iterative training is performed on the last logical regression layer 100-40 of the network, and the two algorithms are compared under different iteration times. The results are shown in Table 1.

![FIGURE 3. RPC dataset product image proofs.](image)

**TABLE 1.** Comparison of the two algorithms.

| Number of iterations | LBP-DBN  | AML-LBP-DBN |
|----------------------|----------|-------------|
|                      | RBM training time(s) | total time(s) | Error rate (%) | RBM training time(s) | total time(s) | Error rate (%) |
| 1000                 | 9.39     | 30.45       | 6.9           | 9.26            | 30.22         | 5.2           |
| 1500                 | 9.47     | 42.38       | 5.77          | 9.28            | 40.73         | 3.26          |
| 2000                 | 9.36     | 51.23       | 3.8           | 9.24            | 50.97         | 1.72          |
| 2500                 | 8.43     | 62.15       | 2.6           | 8.25            | 61.38         | 1.59          |
| 3000                 | 9.48     | 72.47       | 1.72          | 9.18            | 71.87         | 0.79          |

It can be seen from the data of different iterations in Table 1. After the adaptive improvement, the classification accuracy of the AML-LBP-DBN algorithm proposed in this paper is improved compared with the original algorithm, and the number of iterations is reached. At 3000 times, the classification accuracy of the network can reach 99.21%, which is 0.93% higher than the original algorithm.

Result 2: The effect of the number of cluster centers on the correct rate of classification

The average classification correct rate under the setting of the number of different cluster centers is shown in Figure 4.

**FIGURE 4.** Average classification accuracy under different cluster center number settings.

Result 3: The effect of the number of cluster centers on the classification speed

The average classification test time under the setting of the number of different cluster centers is shown in Figure 5.

V. DISCUSSION

Because the problem of improper setting of learning parameters still exists in the RBM greedy training stage, the effect
of network pre-training is not good. In this paper, based on the LBP-DBN algorithm, the adaptive learning method is used to improve the RBM training process. The parameters such as momentum and learning rate are adaptively adjusted by using the similarities and differences of reconstruction error and weight update direction before and after iteration. Further improve the training effect and overall classification performance of the network. It can be seen from the results in Table 1 that after the LBP-DBN is improved by the adaptive adjustment strategy, the training accuracy of the network is significantly improved.

It can be seen from the results 2 that for the level 1 and level 2, the average classification accuracy rate increases with the number of cluster centers, but the increase is gradually reduced. For level 3, the average classification accuracy rate decreases slightly with the increase of the number of cluster centers. This is because Level 1 and Level 2 calculate the distance between the descriptor and the cluster center. The distance of the image-class is based on the difference between the descriptors of the test image and the nearest neighbor cluster center of each image class. The larger the number of class centers is, the more the calculation error caused by quantization is naturally smaller; Level 3 calculates the distance between the descriptor and the descriptor, and the distance of the image-class is based on the distance between the descriptor of each test image and the nearest neighbor descriptor of the region to which each nearest neighbor cluster belongs. The larger the number of cluster centers is set, the smaller the search range of nearest neighbor descriptors is, which increases the deviation of image-class distance calculation.

The classification test time mainly includes two parts, the local feature extraction of the test image, the description and the distance calculation of the image-class. It can be seen from the result 3 that for the level 1 and level 2, the classification test time increases as the number of cluster centers increases; for level 3, the classification test time decreases with the increase of the number of cluster centers. This is also because the level 1 and level 2 calculate the distance between the descriptor and the cluster center, while the level 3 calculates the distance between the descriptor and the descriptor. The number of cluster centers is much larger, and the number of calculations required is decreased.

VI. CONCLUSION

With the rapid development of the Internet and e-commerce, e-commerce platforms are constantly emerging. However, when these shopping platforms provide information to customers, they are presented to shoppers in a combination of images and texts. Compared with the traditional text form, the image is more able to express the attributes of the product, and it is easier for the user to find the product information that he needs. Faced with massive images in the era of big data, effective and accurate classification of commodity image databases is an important factor in improving commodity retrieval. The content-based e-commerce image classification method proposed in this paper can effectively classify the product according to the image information of the product, and help the user to quickly and accurately find the desired product. Through simulation experiments, the proposed method achieves satisfactory image classification results. However, how to improve the accuracy of classification more greatly needs to be further studied in this aspect.

REFERENCES

[1] Z. Lv, H. Song, P. Basanta-Val, A. Steed, and M. Jo, “Next-generation big data analytics: State of the art, challenges, and future research topics,” IEEE Trans. Ind. Informat., vol. 13, no. 4, pp. 1891–1899, Aug. 2017, doi: 10.1109/TII.2017.2650204.
[2] Q. Sun, “Research on image retrieval based on deep learning,” J. Jingdezhen Univ., vol. 33, no. 141(03), pp. 15–18, Jul. 2018.
[3] Z. Wei, “The status quo and exploration of data graphics application in meteorological film and TV programs,” Meteorol. Sci. Technol. Prog., vol. 7, no. 2, pp. 70–72, 2017.
[4] Z. Shuyuan, Z. Weiyi, Y. Shanshan, and F. Yuchao, “Application status and development trend of image processing technology in information agriculture,” Jiangsu Agric. Sci., vol. 45, no. 22, pp. 9–13, 2017.
[5] Q. S. Yuan, D.-M. Zhang, G.-Q. Jin, and F. Liu, “Measurement of image information in video retrieval,” Trans. Commun., vol. 37, no. 2, pp. 80–87, 2016.
[6] L. Zuoren, W. Juay, and A. Yutong, “Application of extreme learning machine in image segmentation,” Comput. Knowl. Technol., vol. 12, no. 3, pp. 207–209, 2016.
[7] B. Wang, F. C. Zou, and X. W. Liu, “New algorithm to generate the adversarial example of image,” Optik, vol. 207, Apr. 2020, Art. no. 164477.
[8] Q. Zezhao, “An improved text topic clustering algorithm based on continuous word package model,” Comput. Knowl. Technol., vol. 14, no. 17, pp. 232–234, 2018.
[9] Y. Min, Y. Fengqin, and C. Ying, “Image labeling of superpixel word package model and SVM classification,” J. Sensors Microsyst., vol. 35, no. 12, pp. 63–65, 2016.
[10] K. Wei, Y. Fengqin, and C. Ying, “Vehicle identification algorithm based on improved word package model,” Comput. Digit. Eng., vol. 46, no. 5, pp. 16–20, 2018.
[11] T. Ojala, M. Pietikainen, and D. Harwood, “Performance evaluation of texture measures with classification based on Kullback discrimination of distributions,” in Proc. 12th Int. Conf. Pattern Recognit., Oct. 1994, pp. 582–585.
[12] Z. Huijuan, “Natural image classification and labeling algorithm based on color features and SVM,” J. Muuyang Normal Univ., vol. 231, no. 5, pp. 18–22, 2018.
[13] Y. Chengrui, “Design of moving image recognition system based on CMOS+ARM,” Comput. Digit. Eng., vol. 44, no. 4, pp. 725–729, Apr. 2016.
W. Yanxi and S. Juan, “SAR image registration method based on speckle reduction suppression SIFT,” *Syst. Eng. Electron.*, vol. 39, no. 12, pp. 2697–2703, 2017.

N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 886–893.

W. Junsheng and X. Yan, “Research progress of medical image segmentation based on middle layer features.” *Beijing Biomed. Eng.*, vol. 35, no. 3, pp. 325–328, 2016.

Z. Meiqing, T. Guangqiang, and Q. Chunhua, “Vector quantization neighbor search strategy for CBIR applications.” *Comput. Eng. Des.*, vol. 38, no. 11, pp. 169–173, 2017.

C. Zhenzhou, W. Shuangyan, and F. Bingbing, “Fisher discriminant analysis based on multilayer automatic encoder,” *J. South China Normal Univ. (Natural Sci.)*, vol. 49, no. 3, pp. 117–122, 2017.

Z. Pengfei, H. Xiaoqun, and W. Yanfei, “Application of a new combined model in credit score,” *Math. Pract. Theory*, vol. 46, no. 5, pp. 20–25, 2016.

L. Breiman, “Bagging predictors,” *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

G. Huilin, “Military image classification based on convolutional neural network,” *J. Comput. Appl.*, vol. 11, no. 21, pp. 323–325, Nov. 2017.

Z. Hui, Z. Yousa, and G. Yi, “Research on medical image classification algorithm based on RBF neural network,” *Electron. Des. Eng.*, vol. 25, no. 3, pp. 113–116, 2017.

C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, “A probability density function generator based on neural networks,” *Phys. A, Stat. Mech. Appl.*, vol. 541, Mar. 2020, Art. no. 123344.

Z. Yongwei, L. Ting, and Y. Boyu, “Image classification method based on deep learning coding model,” *Eng. Sci. Technol.*, vol. 49, no. 1, pp. 213–220, 2017.

W. Chunjing and X. Shengmei, “Similarity measurement method for content-based image retrieval,” *Data Acquisition Process.*, vol. 32, no. 1, pp. 104–110, 2017.

Y. Xingping, “Searching based on image content, ‘Baidu knowledge map is coming.’” *Comput. Knowl. Technol. (Exper. Skills)*, vol. 2, no. 17, p. 26, 2011.

Z. Zengmin et al., “An index structure for large-scale clothing image retrieval based on GMM-cluster tree,” *J. Beijing Inst. Clothing Technol., Natural Sci. Ed.*, vol. 3, no. 10, pp. 35–44, 2016.

D. Yingkui, “Application of face recognition technology based on deep learning in security field,” *China Secur.*, vol. 11, no. 22, pp. 72–74, 2017.

L. Yong, “The development of artificial intelligence promotes the paradigm shift of information security—A case study based on Baidu driverless vehicle,” *Inf. Secur. Res.*, vol. 2, no. 11, pp. 958–968, 2016.

W. Wang, R. Wang, Z. Huang, S. Shan, and X. Chen, “Discriminant analysis on Riemannian manifold of Gaussian distributions for face recognition with image sets,” *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 151–163, Jan. 2018.

Y. Zhang, X. Xiao, L.-X. Yang, Y. Xiang, and S. Zhong, “Secure and efficient outsourcing of PCA-based face recognition,” *IEEE Trans. Forensics Security*, vol. 15, pp. 1683–1695, 2020.

M. J. Zheng, “Research on face recognition technology in video surveillance system,” *Integr. Circuit Appl.*, vol. 36, no. 9, pp. 37–39, Sep. 2019.

S. K. Lakshmanaprabu, M. Elhosery, and K. Shankar, “Optimal tuning of decentralized fractional order PID controllers for TITO process using equivalent transfer function,” *Cognit. Syst. Res.*, vol. 58, pp. 292–303, Dec. 2019, doi: 10.1016/j.cogsys.2019.07.005.

XIAOLI ZHANG was born in Xuchang, Henan, China, in 1985. She received the master’s degree from the Chongqing University of Technology, China. She is currently working with the School of Business, Xuchang University. Her research interests include e-commerce and network marketing.

* * *

**IEEE Access**

X. Zhang: Content-Based E-Commerce Image Classification Research

160220

VOLUME 8, 2020