# Application of Unsupervised Deep Learning in Color Image Recognition

**Li Bingjie, Gao Jingyang, Shi Xinyu**

Information Engineering University, Zhengzhou Henan 450000

**Keywords:** Deep Learning, Convolution Neural Network, Perceived Loss, Image Recognition

**Abstract:** Today is the Internet information age, now deep learning has made a breakthrough in the field of image, and convolution neural network is the leader of this wave of deep learning wave. In reality, images are not in the environment of simple background and single target. It is difficult to recognize such images by using CNN model. In this paper, the application of unsupervised deep learning in color image recognition is proposed. In order to verify the performance of SVM image recognition method, this paper takes two training samples and test samples, and selects different hidden layer nodes and layers to test. Through analysis, it is concluded that the accuracy of two hidden layers is higher than that of only one hidden layer. In order to further test the feasibility of the algorithm, in cooperation with a University of traditional Chinese medicine, typical tongue images were selected as training samples. Experiments are carried out on facial expression database and tongue image data respectively. The recognition rate of all images can reach more than 92%, and the recognition time is about 43M / s. The results show that the algorithm is in the international high level, and the image recognition is more intelligent. Through the analysis, the research in this paper has achieved ideal results, and made a contribution to the application of unsupervised deep learning in color image recognition method.

## 1. Introduction

At present, artificial intelligence is developing rapidly and making breakthroughs. Since 2006, machine learning, big data, cloud computing and other concepts and methods have made great progress, which is inseparable from a powerful algorithm, namely deep learning [1]. Through DL algorithm, people have found a way to deal with "abstract concept"[2]. Major science and technology companies have invested a lot of money to research deep learning and made great progress, so deep learning is a "diamond" [3]. In the scientific community, through it, we can solve many problems that could not be realized before. With the deepening of research, the field of deep learning will be gradually expanded, including the field of color image recognition [4].

Deep learning method is an effective method to solve the problems of computer vision and image processing, and has been widely used [5]. In recent years, the application of deep learning method in image recognition research has become a hot field [6]. In the implementation of traditional recognition methods, it is necessary to manually design and adjust the activity level measurement method (design filter to extract high-frequency details) and fusion rules [7], and often combined with multi-scale decomposition of images, the implementation process is more complex [8]. The recognition method based on deep learning can automatically extract the features of the input source image by establishing a deep neural network model[9], and output the ideal recognition image result after multiple training and optimization under the constraint of loss function [10].

This paper analyzes the practical application of color image recognition method in unsupervised deep learning, and finds that compared with developed countries, there are still deficiencies in technology promotion and guarantee. Therefore, this paper establishes the research on unsupervised deep learning in color image recognition. In the research, according to the actual situation of image recognition, the introduction of unsupervised deep learning can improve the work efficiency and core competitiveness. Through the analysis of the survey results, this paper considers that unsupervised deep learning can optimize the color image recognition method and achieve good results.

  DOI: 10.25236/mmmce.2020.044

## 2. Unsupervised Deep Learning and Mesh Optimization for Image Recognition

### 2.1 Unsupervised Deep Learning

In unsupervised learning, training samples only contain features without corresponding labels. Therefore, clustering, density estimation, anomaly detection and other methods or certain models are needed to determine the relationship between data and determine the optimization objective of loss function. When deep learning is applied to infrared and visible image recognition, there is no standard reference image that can be used, so supervised learning cannot be used directly for training. In order to solve the above problems, the existing infrared and visible image data sets are used to construct the complex loss function image and the constraint recognition between the source image and the image feature to complete the unsupervised learning of the model.

### 2.2 Mesh Optimization Process for Image Recognition

In deep learning, the process of minimizing the loss function (L) is also a process of network optimization and continuous improvement of model capability.

L1 loss refers to the average absolute error between the predicted value of the model and the tag value (target true value). L1 loss has good adaptability and robustness to outliers, but its derivative is discontinuous and non-differentiable, and its solution efficiency is low, the calculation method is as follows (1):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y^{(i)} - f\left(x^{(i)}\right)\right| \tag{1}$$

L2 loss is the expected value of the square of the difference between the predicted value of the model and the label value. At present, L2 loss is widely used in practical application. L2 loss is more efficient and stable than L1 loss, but it is sensitive to outliers. The calculation method of L2 loss is shown in formula (2), and the parameter definition is consistent with L1 loss.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y^{(i)} - f\left(x^{(i)}\right)\right)^2 \tag{2}$$

Square root loss is calculated on the basis of the square root. It can directly reflect the discrete degree between the predicted value and the real value of the model. The calculation method is shown in formula (3), and the parameter definition is consistent with L1 loss.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y^{(i)} - f\left(x^{(i)}\right)\right)^2} \tag{3}$$

In addition to the common loss functions mentioned above, many researchers have developed their own loss functions according to the actual network structure characteristics and task objectives in image classification and recognition tasks. These functions play an important role in network optimization, greatly improving the ability and effect of image processing.

## 3. Application and Result Analysis of Image Recognition

As for the problem of color image classification and recognition, a color image recognition method based on image feature data field and deep trust network is proposed. First, about the human visual characteristics, the image color data domain is composed. Second, image reading method, combined with and using wavelet transform to describe multi-scale texture features. Thirdly, the method of image recognition uses unsupervised deep trust network to realize image recognition. In order to verify the feasibility of the algorithm, the algorithm is tested by facial expression database and tongue data.

In this paper, in cooperation with a University of traditional Chinese medicine, typical tongue images were selected as training samples, which were preliminarily verified by Chinese medicine experts, including tooth marks, fat and thin, spot, aging and tenderness. The experimental results are

shown in Table 1. The experimental results in Table 1 are not only affected by the subjective interpretation of doctors, but also related to the standardization of training samples and learning samples. The experimental results also show that the proposed method not only has a good recognition rate, but also has a high level of acceptable recognition time. The recognition rate of all images can reach more than 92%, and the recognition time is about 43M / s, which is at a high level. Therefore, it has certain clinical diagnostic value.

Table 1 Experimental results of tongue image identification.

| Tongue image | Recognition rate (%) | Identification time (M/S) |
|---|---|---|
| Tooth marks | 94.5 | 44.1 |
| Fat and thin | 94.7 | 42.9 |
| ecchymosis | 92.8 | 43.1 |
| Old and young | 93.1 | 43.0 |

## 4. Discuss

### 4.1 Image Recognition Method Based on Deep Learning and SVM

First, simple scaling. The purpose of simple scaling is to readjust the value of each dimension of the data so that the final data vector falls into a predetermined interval, such as [0,1]. In order to avoid the difference between the input data and the output data, the order of magnitude of the difference is avoided. For example, for a natural image, the obtained pixel values are in the range of [0,255] and are scaled to the interval of [0,1], just divide these pixel values by 255. This is useful for future data processing and is necessary because many default parameters assume that the data is in a reasonable range.

Second, feature standardization. Feature standardization refers to changing the average value of each dimension of the data to zero and the variance to 1 change. The specific method of feature standardization is to use all data to calculate the average value of data in each dimension, and then subtract the average value from each dimension. Next, using the standard deviation of the data on the dimension will remove the data from each dimension. In some cases, in order to balance the effects of various components, standardization is usually used separately for each component of a feature.

Third, sample by sample mean subtraction. For images, this standardization can be used to eliminate the average brightness value in the image, because in most cases, we are more interested in the content of the image, and this is not lighting. This method is effective for natural images, but not for non-stationary data.

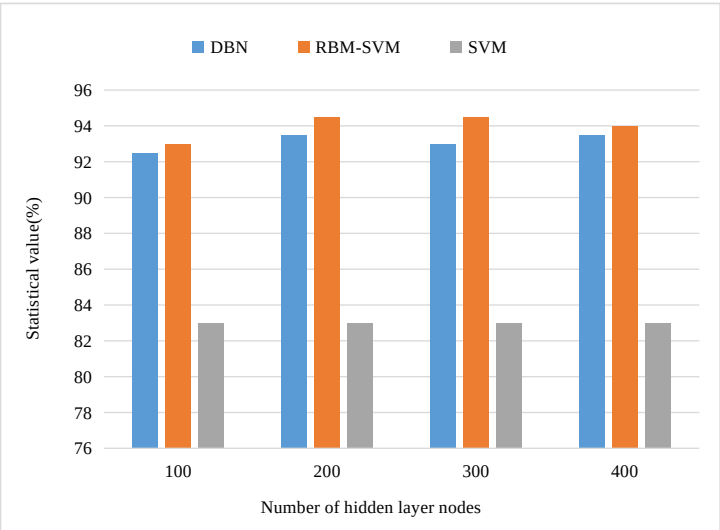


Figure 1. Prediction accuracy of different methods for single hidden layer

In order to verify the performance of SVM image recognition method, two training samples and

test samples are selected to test different hidden layer nodes and layers. The experimental results are shown in Figure. 1. The analysis in Figure 1 shows that the accuracy of the method is higher than that of the DBN method. When there is only one hidden layer and the other parameters are the same, the accuracy of this method is basically higher than that of the DBN method with only one hidden layer. Moreover, the accuracy of this method is much higher than that of SVM. With the increase of the number of samples, the accuracy of the method is obviously improved. When there are two hidden layers, the accuracy is higher than that of only one hidden layer.

In addition, this paper further tests the influence of C value on SVM method. The experimental results are shown in Figure 2. As can be seen from Figure 2, with the increase of C value, the number of support vectors in SVM method decreases sharply, and then gradually tends to be stable, with little change. As can be seen from the figure, C value has a great influence on SVM. Because the penalty factor C represents the punishment degree of the wrong samples, the higher the C value, the higher the classification accuracy of the training samples, while the lower the classification accuracy of the test samples, the lower the generalization ability of the model. Low C value will lead to the lack of learning ability. The accuracy of training samples and test samples is very low.

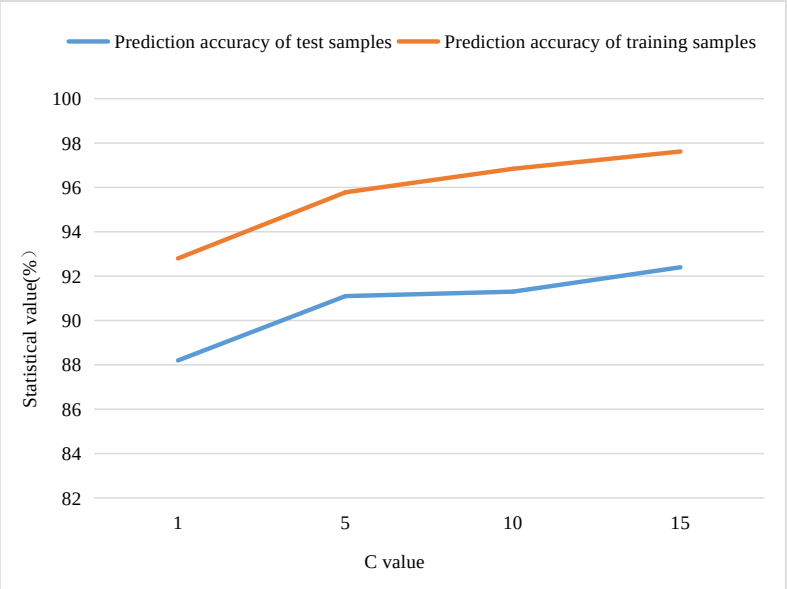


Figure 2. Influence of C value of SVM method on accuracy

### 4.2 Improvement Strategy of Image Recognition

(1) Convolution and pooling

Natural image has its inherent characteristics. In the image, the statistical characteristics of the inherent part of the image are the same as other parts. The features learned in this part can also be applied to other parts of the image. Therefore, all positions on the image can be learned using the same features. For example, in a 36x36 image, a small image (such as 9x9) is randomly selected as the sample, and some features are learned from the small sample. At this time, the features learned from 9x9 small samples can be applied to other parts of the whole image as a detector, and then the activation values of different features in the original image position can be obtained..

(2) Basic idea of alternation of unsupervised and supervised learning

General training process of DBN network is: training one RBM at a time and adopting unsupervised learning. The improved learning method has the advantages of fast convergence speed, less training steps and short training time. The basic idea is: firstly, after the first layer of RBM network is unsupervised learning, the label is added and the parameters of RBM network are adjusted. Then, the output of the adjusted RBM network is used as the visual layer of the next RBM network. This method needs only one unsupervised and supervised learning for each additional hidden layer, which can fully learn the underlying parameters of the network, thus speeding up the convergence speed of the supervised process of the RBM network.

## 5. Conclusions

In the research of color image recognition method, this paper takes unsupervised deep learning as the main line of research, through different data testing and combined with SVM image recognition scheme research, this paper optimizes and improves the influence factor C value of SVM, so that the image recognition process does not need human participation, realizes more intelligent, and is gradually expanded and applied. At the end of this study, the system performance is tested. In the test, a number of experiments including the prediction accuracy of different methods of single hidden layer and the influence of C value of SVM method on the accuracy are carried out. Through the analysis of test data, we can see that no matter the size of C value has a great impact on the accuracy of SVM. With the increase of the number of samples, the accuracy rate is significantly improved. When there are two hidden layers, the accuracy rate is higher than that of single hidden layer, reaching a higher level in the industry. Compared with the traditional method, the optimized system improves the comprehensive performance and robustness. This study has achieved ideal results and provided technical support for the application of color image recognition method.

## References

[1] Ter Hoeven C L, Van Zoonen W, Fonner K L. The practical paradox of technology: The influence of communication technology use on employee burnout and engagement [J]. Communication Monographs, 2016, 83(2):239-263.

[2] Shi Y, Yang J, Hu X. How economic globalization affects urban expansion: an empirical analysis of 30 Chinese provinces for 2000–2010[J]. Quality & Quantity, 2016, 50(3):1117-1133.

[3] ŁęgowikMałolepsza, Małgorzata, Sałek, Robert, Bubel, Dagmara. DISRUPTIONS OF THE FLOW OF INFORMATION IN BUSINESS MANAGEMENT[J]. Journal of Clinical Investigation, 2016, 113(9):1271-1276.

[4] Wooliscroft B, Shaw E H. Marketing strategy[J]. Journal of Historical Research in Marketing, 2012, 4(1):30-55.

[5] Bredow, Sabrina Monique Schenato, Lélis, Marcos Tadeu Caputi, Cunha, André Moreira. Super cycle of commodity prices and the Brazilian economy: an analysis of the external transmission mechanisms between 2002 and 2014[J]. Econ Soc, 2016, 25.

[6] Xin-Shu G, Hong-Zhen Z, Ming-Ming P. Market competition,environmental regulation and industrial pollution emissions[J]. China Population,Resources and Environment, 2017:págs. 67-78.

[7] Laird J, Venables A J. Transport investment and economic performance: A framework for project appraisal [J]. Transport Policy, 2017, 56(MAY):1-11.

[8] Hashem I A T, Chang V, Anuar N B, et al. The role of big data in smart city[J]. International Journal of Information Management, 2016, 36(5):748-758.

[9] Valencia V S, Buddlemeyer H, Coelli M, et al. Corporate Governance and Innovation[J]. Journal of Financial & Quantitative Analysis, 2012, 47(2):397-413.

[10] Jain A, Hautier G, Ong S P, et al. New opportunities for materials informatics: Resources and data mining techniques for uncovering hidden relationships[J]. Journal of Materials Research, 2016, 31(8):977-994.