

Research on image retrieval using deep convolutional neural network combining L1 regularization and PRelu activation function

Wei QingJie, Wang WenBin

Chongqing University of Posts and Telecommunications, China
Tel:18716320057

Abstract. In this paper, the image retrieval using deep convolutional neural network combined with regularization and PRelu activation function is studied, and improves image retrieval accuracy. Deep convolutional neural network can not only simulate the process of human brain to receive and transmit information, but also contains a convolution operation, which is very suitable for processing images. Using deep convolutional neural network is better than direct extraction of image visual features for image retrieval. However, the structure of deep convolutional neural network is complex, and it is easy to over-fitting and reduces the accuracy of image retrieval. In this paper, we combine L1 regularization and PRelu activation function to construct a deep convolutional neural network to prevent over-fitting of the network and improve the accuracy of image retrieval

1 Introduction

Most of the search engines are based on text retrieval, but the text expression is limited, one-sided and can not fully express users' intention. The user can not receive desired results and affect the search accuracy. Compared with text, image containing very rich content can better express users intention. Using image retrieval can meet users' requirements and improves search efficiency. Now the research is mainly extracting image visual features^[1,2] then using feature similarity measure to get the result of image retrieval. Using this method obtains low image retrieval rate and can not meet the needs of users. With the development of neural network, research using BP neural network for image retrieval appeared. First of all, extract the visual features of the image, and then enter the features to BP neural network^[3] for image retrieval. The effect of using this method is better than the first method. But the network structure is simple, so the retrieval efficiency is not ideal. In recent years, the deep learning network^[4] has become a research hotspot, it has more complex network structure than traditional neural network, closer to the process of human brain transmission and understanding of information, so there are a lot of researches through deep convolutional neural network for image retrieval. The practice has proved that the image retrieval efficiency is significantly improved.

Deep convolutional neural network^[5,6] is a kind of deep learning network, not only has complex structure, but also has convolutional operation for image processing, which is suitable for processing images. Deep convolutional neural network includes input layer, hidden layer, all connection layer and output layer. The hidden layer is an important structure of the deep convolutional neural network to process the image and transfer the image information, including the feature convolution layer and the feature sampling layer. The input of the deep convolutional neural network is a complete image, from the processing of the hidden layer to the full connection layer, the full connection layer is similar to



the BP neural network connection layer, the final output image retrieval results. Because deep convolutional neural network has large number of network layers, it is easily over-fitting, which reduces the efficiency of image retrieval and the generalization ability of the network. In this paper, we use the method of combining regularization and PRelu activation function to prevent the network from over-fitting. After experiment, the efficiency of image retrieval is better than original deep convolutional neural network.

In this paper, Part 2 introduces the method of deep convolutional neural network, regularization and activation function. Part 3 introduces the method of image retrieval using deep convolutional neural network combining with L1 regularization and PRelu activation function. Part 4 introduces the experimental results and last part is a summary.

2 Theoretical principle

2.1 Deep convolutional neural network

Deep convolutional neural network^[7,8,9] is based on neural network, so it includes the basic structure of neural network, input layer, hidden layer and output layer. Each layer consists a number of neurons connecting to each other. Each neuron contains an input signal, an internal activation function, and an output signal. Deep convolutional neural network than the neural network contains more hidden layer, hidden layer structure is more complex, including two types of processing layer. One is the feature convolutional layer, mainly carries on the convolutional operation to the image; the other is the feature sampling layer, mainly carries on the sampling to the image feature after the convolutional processing, reducing the network training parameters and enhancing the network training speed.

Deep convolutional neural network has two most important features, local feelings domain and parameter sharing. The local feelings domain, that is, the neurons of each hidden layer do not connect to each pixel of the image, but only through the convolution template to the local image area, and then at higher layer will combine neurons feeling different local image to get the overall information. Through the local feelings of the domain can reduce the network connection parameters. Parameter sharing refers to the same neurons connecting the image have the same weight, for example, convolutional template is 10*10, then the 100 connection weight with the image are the same. This can effectively reduce the training parameters and improve training time.

2.2 Regularization

Regularization^[6] refers to adding the additional penalty item on the basis of the cost function. The two most commonly used regularization methods are L1 regularization and L2 regularization. L1 regularization penalty is the absolute value of all network weights. The purpose of training is to reduce the cost function, which means that the penalty items should actually be small, and the absolute minimum is 0. For the network, L1 regularization forces some of the connection weight is 0, which can reduce the complexity of the network and avoid over-fitting. The regularization penalty for L2 is the sum of squares of all weights. L2 regularization will make each connection weight smaller, but will not let the weight to 0. Reduce the impact of high-order items by reducing the weight, reducing the possibility of over-fitting. In this paper, we use L1 regularization, because the network structure is complex, L2 regularization is only to reduce the connection weight, and can not effectively solve the network over-fitting.

2.3 Activation function

The activation function^[10] is an important part of a neuron. The neurons of the neural network are similar to neurons in the brain and receive the input signal. If the response threshold is reached by the activation function, the output signal is generated and the signal is passed to the neurons connected to it. The activation function adds non-linear factors to the network and improves the model fitting and expression ability. The activation function is nonlinear, informative and monotonic. The main activation function of the current deep learning network are the Sigmoid function and the Tanh function. The functions' expression is shown in Equation 1, 2.

$$f(x) = \frac{1}{1 + e^{-ax}} \quad (1)$$

$$f(x) = \frac{2}{1 + e^{-ax}} - 1 \quad (2)$$

The output range of the Sigmoid function is between [0-1]. The output range of the Tanh function is between [-1-1]. As long as there is an input signal, Sigmoid and Tanh functions will respond to the output, that is, in the network training, all neurons are in the active state, easily to cause over-fitting. At the same time Sigmoid and Tanh functions will appear gradient disappearance, when the neuron's output value close to the maximum, the function curve becomes very smooth. When the error feedback the derivative is close to zero and the error of the next layer can not be transmitted, training failed. For the above phenomenon, there are some studies that use Relu activation function instead of Sigmoid and Tanh functions. The expression for the Relu function is shown in Equation 3.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (3)$$

Relu function is unilateral activation, when the neuron value is less than zero, the neuron without any output, thus reducing the complexity of the network and reducing the probability of over-fitting. When the value of the neuron is greater than zero, the output of the neuron is linearly related to the input. During the error transfer, the gradient decreases rapidly, which reduces the training time and avoids the disappearance of the gradient. But Relu has a drawback that is likely to cause neuronal death. As long as the input of the neuron is negative, there will be no output, and if there is a large number of neurons dead, the deep learning network will lose its meaning.

In order to solve the above problem, this paper uses the PRelu activation function, which is an improvement of the Relu function, a Relu activation function with parameters. The function expression is shown in Equation 4.

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{if } x < 0 \end{cases} \quad (4)$$

PRelu adds a negative response on the basis of Relu, which avoids the direct death of neurons. It not only reduces training time function but also avoids over-fitting like Relu. In summary, the PRelu activation function can overcome the shortcomings of the commonly used activation function. It can solve the fitting problem and improve the efficiency of image retrieval in deep convolutional neural network.

3 Method of image retrieval using deep convolutional neural network combined with regularization and PRelu activation function

3.1 Construct deep convolutional neural network

In this paper, four kinds of deep convolutional neural networks are constructed by combining L1 regularization and PRelu activation function, including input layer, hidden layer, all connection layer and output layer. The main difference is that the number of hidden layers is different, including the layers 3, 4, 5 and 6 respectively. The specific hidden layers' structure are shown in Figure1.

| Network Ordinal | Number of Hidden Layer | Structure of Hidden Layer | Activation Function | Method of Sampling |
|-----------------|------------------------|---|---------------------|--------------------|
| 1 | 3 | Feature Convolutional Layer-Feature Sampling Layer-Feature Convolutional Layer | PReLU | Max Sampling |
| 2 | 4 | Feature Convolutional Layer-Feature Sampling Layer-Feature Convolutional Layer-Feature Convolutional Layer | | |
| 3 | 5 | Feature Convolutional Layer-Feature Sampling Layer-Feature Convolutional Layer-Feature Convolutional Layer-Feature Sampling Layer | | |
| 4 | 6 | Feature Convolutional Layer-Feature Sampling Layer-Feature Convolutional Layer-Feature Convolutional Layer-Feature Sampling Layer-Feature Convolutional Layer | | |

Figure1: The structure of deep convolutional neural network for image retrieval experiment

The cost function of constructing a deep convolutional neural network is shown in Equation 5.

$$f(x) = \frac{1}{2} \sum (y-t)^2 + \alpha |w| \quad (5)$$

Where y denotes the image category predicted by the network, t denotes the true category of the image, and w denotes the absolute value of the connection parameter. The function represents the objective of network training to minimize the sum of the difference between the predicted and true values, while minimizing the absolute values of the connection parameters.

3.2 A method of image retrieval using deep convolutional neural network

Deep convolutional neural network for image retrieval steps: first, image database is divided to training set and test set, then normalize all the images and the pixel value reduced to 0-1, to avoid the value range of pixel having influence for network weights training. Second, forward train deep convolutional neural network. Input the training set images into the deep convolutional neural network which has been constructed. After processed by the convolutional layer and the sampling layer of the deep convolutional neural network, the image information of the image is output as the network predictive value, while calculating the accuracy of image retrieval. Third, the backward feedback. Calculates the sum of the squared error of the predicted values of the images and the true values and the absolute values of all the connecting parameters as the total error value, and then adjusts the parameters using the backward error feedback method. Fourth, repeat the second step with the adjusted parameters, and stop the training when the image retrieval accuracy reaches a threshold. Fifth, the test set images are input to the trained network and calculate the retrieval accuracy of the test set images by the output of the network.

4 Experiment

The experiment's image database is Oxford landmark library, containing 17 landmark buildings, a total of 5038 images. The images of the Oxford image library are not designed for image experimentation, including a large number of tourist images, which means the target object has a large number of occlusion, tilt, fuzzy and other issues, thus having a great impact for the efficiency of image retrieval. If the retrieval efficiency of this image library is ideal, it can prove that the method of constructing the model is reasonable.

In the experiment, the images in the image library are randomly divided into training set and test set according to the ratio of 7: 3. The training set contains a total of 3530 images and the test set contains 1508 images. All pixel values are normalized. The training set images are input to the four deep convolutional neural networks for training. According to the training error to adjust the network

parameters, when image retrieval rate of the training set is greater than 95% stop training, thus completing the deep convolutional neural network training. For testing, the images of the test set are input to the trained network, and the retrieval accuracy of the test set image is calculated. The experimental results are shown in Table 1.

| Network Ordinal | Number of Hidden Layer | Activation Function | Accuracy of Image Retrieval |
|-----------------|------------------------|---------------------|-----------------------------|
| 1 | 3 | PReLU | 65.77% |
| 2 | 4 | PReLU | 67.97% |
| 3 | 5 | PReLU | 70.49% |
| 4 | 6 | PReLU | 71.95% |

Table1: The accuracy of image retrieval using a deep convolutional neural network that combines L1 regularization and PReLU activation functions.

At the same time, we use the same number of network layers, but do not use L1 regularization and use Sigmoid activation function for deep convolutional neural network to experiment on the same image set. The experiment's results are shown in Table 2.

| Network Ordinal | Number of Hidden Layer | Activation Function | Accuracy of Image Retrieval |
|-----------------|------------------------|---------------------|-----------------------------|
| 1 | 3 | Sigmoid | 65.25% |
| 2 | 4 | Sigmoid | 67.04% |
| 3 | 5 | Sigmoid | 66.78% |
| 4 | 6 | Sigmoid | 65.72% |

Table2: The accuracy of image retrieval using the deep convolutional neural network with Sigmoid activation function

Figure 2 compares the accuracy of image retrieval using deep convolutional networks combined with L1 regularization and PReLU activation functions and the accuracy of image retrieval using deep convolutional networks with Sigmoid activation functions. From Figure 2 we can see that the network which number of hidden layer is only 3,4, the two groups of network image retrieval accuracy are very similar. However, with the increase of network layer, the accuracy of image retrieval using Sigmoid activation function begins to decline, and the over-fitting phenomenon appeared. While the image retrieval accuracy of the deep convolutional neural network using L1 regularization and PReLU activation function shows a more stable improvement. Therefore, it can be proved that the method of this paper can effectively solve the problem of over-fitting of deep convolutional neural network and improve the accuracy of image retrieval.

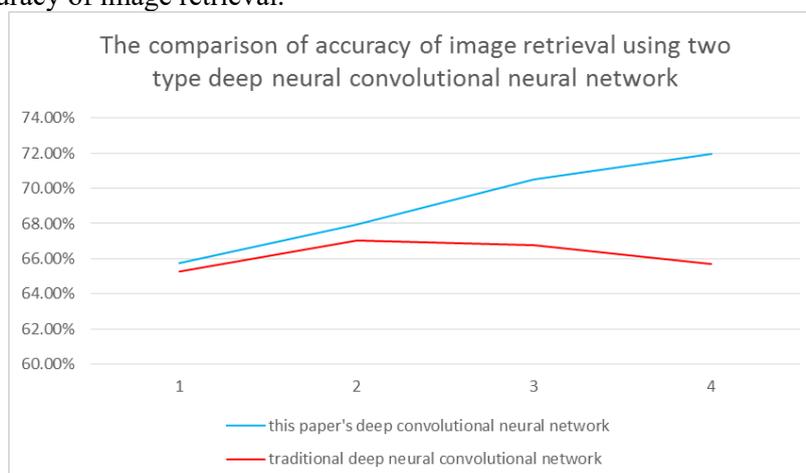


Figure 2 The comparison of the accuracy of image retrieval between this paper's deep convolutional neural network and traditional deep convolutional neural network

5 Conclusion

In this paper, we use the deep convolutional neural network combined with L1 regularization and PRelu activation function and the traditional deep convolutional neural network to carry on the image retrieval experiment. The experimental results show that the traditional deep convolutional neural network is prone to over-fitting, which leads to the decrease of the accuracy of image retrieval. The method of this paper can prevent the network from over-fitting, so that with the increase of network structure complexity, the corresponding image retrieval accuracy is also improving. Over-fitting easily occur in the deep learning network, but also a very important factor. In order to use deep learning network, we must solve over-fitting problem. Experiments shows that the method of this paper can prevent deep convolutional neural network from over-fitting and improve the accuracy of image retrieval.

References

- [1] Ms,Shilpa P.Pant.Content Based Image Retrieval Using Color Feature[J],*Engineering Research and Technology*,2013,10(2).
- [2] C Kavitha, MB Rao, BP Rao. Content Based Image Retrieval by Using an Integrated Matching Technique Based on Most Similar Highest Priority Principle on the Color and Texture Features of the Image Sub-blocks[B],2011.
- [3] H Karamti, M Tmar, F Gargouri. Content-based image retrieval system using neural network[J], *IEEE/ACS International Conference on Computer and Science*,2014:723-728.
- [4] Dong Yu,Li Dong. Deep Learning and Its Applications to Signal and Information Processing[J], *IEEE Signal Processing Magazine*,2013,1(145).
- [5] H Karamti, M Tmar, F Gargouri. Content-based image retrieval system using neural network[J], *IEEE/ACS International Conference on Computer and Science*,2014:723-728.
- [6] Matthew D.Zeiler,Rob Fergus.Stochastic Pooling for Regularization of Deep Convolutional Neural Networks[J],*Computer Science*,2013,2(16).
- [7] D.C.Cire,san, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification[J],*Arxiv preprint arXiv:1202.2745*, 2012.
- [8] Y. LeCun, K. Kavukcuoglu, and C. Farabet. Convolutional networks and applications in vision[J]. *InCircuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pages 253–256.IEEE, 2010.
- [9] Duong Quoc Thang LE, Shriman Narayan Tiwari. Deep Learning Image Recognition[D], *Stuwww.eurecom.fr*,2015.
- [10] SL Shan, M Khalil-Hani, R Bakhteri. Bounded activation functions for enhanced training stability of deep neural networks on visual pattern recognition problems[J], *Neurocomputing*,2016,216:718-734.