

Study on Image Detection and Recognition based on Deep Neural Network under Cloud Computing Services

Feng Liu^{1, a}, Peiwei Wang^{2, b} and Zhixian Wang^{1, c}

¹NetWork & Information Center, Yulin University, Yulin719000, China ²College of Information Engineering, Yulin University, Yulin719000, China ^aliufeng@yulinu.edu.cn, ^bwpwkimkibun@163.com, ^cwzx@yulinu.edu.cn

Keywords: Image detection; Image recognition; Deep neural network; Cloud computing

Abstract. Deep learning model has the ability to extract image features independently, and also can extract abstract features. This paper studied on image detection and recognition based on deep neural network under cloud computing service. Firstly, the cloud computing service model was introduced. Secondly, the learning network was designed by the optimization method. Lastly, the test results were analyzed by recall rate and confusion Matrix, convergence performance and comparison of accuracy. The results show the deep neural network under cloud computing services is a valid method for image detection and recognition.

Introduction

Due to the rapid development of information technology, when network data gradually complicated, the traditional parallel technology has not required for full foot growing human development and scientific research office [1]. As some network model arose at the historic moment, cloud computing as a new parallel computing technology also appeared on the network[2]. As a new Internet, cloud computing is the further development of distributed computing, parallel computing and grid computing based computing model, is also the concept of computer science commercial implementation[3].

Distributed computing tasks will constitute a large number of computer resource pool, so that all applications can root access to computing power, storage space and software services [4]. Cloud computing resources are dynamically scalable and virtualization, through the Internet providing cloud computing is a super computing model based on Internet, and the architecture of a distributed, resource structure of global access, the data center operation in a similar calculation under the environment of Internet are stored in the personal computer, mobile phone and other equipment on the large amount of information and processor resources and work together. Cloud computing services include three levels of service, IaaS, PaaS and SaaS[5].

In recent years, virtualization is the cornerstone of cloud computing, which has also been a hot research field. Virtual tides across all areas of server, storage, network, PC machine etc.. The most prominent advantage of virtualization is to save money, the integrated service server, maximize the utilization of resources.

Cloud Computing Service Model

In the cloud computing service model, the designed user interface can capture user requirements with Web Service for various user access interface. The configuration tool can prepare task environment in the distribution of the nodes. The system management module is responsible for the management and allocation of all available resources, whose core is to ensure the load balance. The directory service is the list of services that a user can access. The monitoring module is responsible for monitoring the running state of the node and finishing the statistics of the user's node. The user interface allows users with interactive from the directory select and call a service and passes the request to the system management module. It will assign the appropriate information source for the user, and then call the



configuration tool to run environment for users. In the cloud server, computing and storage resources of all in different nodes, the system administrator is to use some configuration tools and system management software, on the one hand, it can provide convenient computing and storage resources for the user, on the other hand it can be an efficient management of the computer storage resources and improve resource utilization the rate.

For cloud computing service providers, its core technology is how to allocate the computational resources for users and the work performance and its efficiency directly affects the entire cloud computing environment. Because cloud computing resources has the characteristics of autonomy, virtualization unique, which makes the original in simple distributed grid computing, resource allocation and scheduling algorithms cannot work effectively in the environment. In cloud computing, the efficiency of resource allocation is very important, and it has a great impact on the system performance.

Optimization Design of Learning network

Error Transfer. Learning network is a cascade structure, which can diffuse to all levels of elements through output error, so as to complete the operation through descent gradient.

If the current layer of convolution layer is m, the connection mode would be sampling connection. Then, with the development of connection mode, the resolution ratio of feature map will be reduced to 1/u of the original size, and the phenomenon of enlarging by S times will appear when connection is back. Through expand function, each corresponding element in the error matrix of feature map of m+1 layer can be extended to $u \times u$ micro-grid matrix, thus obtaining that resolution ratio is S times of characteristic error. If excitation function is represented as f(x), in the feature map error of $o_{f_{map-j}}^{l}$ times of derivative at $o_{f_{map-j}}^{l}$ of j feature map of m layer, the corresponding feature map error $o_{f_{map-j}}^{l}$ can be obtained, which can be expressed as Eq.1.

$$o_{f_{map-j}}^{l} = f(\mathbf{x})_{|\mathbf{x}=o_{f_{map-j}}^{l}} \operatorname{gexp} and(\Delta o_{f_{map-j}}^{l}[\mathbf{u},\mathbf{u}])$$
(1)

Since sampling layer has no parameters, and error of feature map must be transmitted in the backward transmission process, it is necessary to assume that the current layer m is sampling layer, and the connection after m is convolution connection. For n_{outmap} convolution needs to be connected, the corresponding error is the summation of n_{outmap} component. Dimension vector function of returning to matrix is represented as size, downscaling size of dimension is $size(o^l) - (size(k) - 1)$, the corresponding matrix is o^{l+1} , operating convolution kernel reversal is represented by inv, and error matrix $o_{f_{map-j}}^{l}$ of feature map of layer m is represented convolution operation and Eq.2 is obtained.

$$O_{f_{map-i}}^{l} = \sum_{j=1}^{n_{outnap}} \operatorname{cov} n(\Delta O_{f_{map-j}}^{l}, inv(\mathbf{k}_{ij}^{l+1}), "full")$$
(2)

Elastic Momentum Learning. In the backward transmission process of error, the feature error of nodes among various layers is $\Delta O_{f_{map-j}}^l$, and in the process from m = 2 to u = n, the error of convolution operator can be obtained, kernel matrix error of reversal convolution can get output error, otherwise, reversal matrix inv $(O_{f_{map-j}}^l)$ and output error convolution are "valid". And $\Delta O_{f_{map-j}}^l$ can get nuclear error Δk_{ij}^l as Eq.3.

 $\Delta k_{ij}^{l} = \text{convn} (\text{ivn} (\Delta o_{f_{map-i}}^{l}) , \Delta o_{f_{map-j}}^{l}, \text{"valid"})$ (3)

Through a backward and onward process, layer weight error of full connection and layer nuclear error of each convolution can be obtained, and convolution operator is finally obtained. Ignoring the label of layer, and making the convolution kernel of No. n-1 adjustment be $\Delta k_{ij}^{(n-1)}$ and that after adjustment be $k_{ij}^{(n)}$, then Eq.4 is as follows.

$$k_{ij}^{(n)} = k_{ij}^{(n-1)} - \Delta k_{ij}$$
(4)

In order to accelerate the convergence speed of algorithm, the momentum method is applied, and β initializes a function β_0 . In the learning process, the size of parameter is set according to the algorithm, and the method plays a more flexible role, compared with momentum method, while in the aspect of adjusting direction, that the learning parameters should be improved still has not been realized by Eq.5 and Eq.6.

$$\Delta k_{ij}^{(n)} = convn(g) + \beta \Delta k_{ij}^{(n-1)}$$
(5)

$$\boldsymbol{\beta} = f(\boldsymbol{\beta}_0, \mathbf{n}_{hatch}, \mathbf{s}) \tag{6}$$

For the above-described method, the consistency between the error symbol and momentum aspect has not been taken into account, but in the actual process, the current error term and previous momentum and direction have a certain connection. In order to meet the shortfall, the following method is proposed.

Adjusted items are all represent as η_{ij} , the reverse current error of sign function is $\partial \eta_{ij}$, and product symbol of previous adjustment amount is $\Delta \eta_{ij}$ (n-1), so that the error term and momentum term maintain a dynamic consistency in direction. The selection of this value is the proportion in the middle of the sum of previous adjustment and error amount, such mechanism is defined as gradient descent of elastic momentum, and is represented as Eq.7 and Eq.8.

$$\beta = \operatorname{sign} \left(\partial \eta_{ij} g \Delta \eta_{ij} (n-1) \right| \frac{\Delta \eta_{ij} (n-1)}{\Delta \eta_{ij} (n-1) + \partial \eta_{ij}}$$
(7)

$$\Delta \eta_{ij}(\mathbf{n}) = \partial \eta_{ij} + \beta \Delta \eta_{ij}(\mathbf{n}-1)$$
(8)

Analysis on Test Results

Recall Rate and Confusion Matrix. The indicator formulated for determining the performance of evaluation classifier is called recall ratio. In the subset constituted by n A samples in set S, a classifier is used to test the testing set, and samples classified as Class A constitutes subset S_0 , that is, the ratio of element number m and n of $S_0 = S_T \cap S_0$ is classifier's recall ratio of A.

Convergence Performance. Convolution depth study of deep learning neural network selects three momentum algorithms, to observe the iteration round change of loss function of training objectives. So the convergence iteration times of exponential momentum is the minimum, and its consumption of the training time is also the minimum. To make decisions based on the volatility and smoothness of loss function curve has a strong advantage.

Comparison of Accuracy. In order to compare the accuracy of the method, three different methods are applied for comparison, respectively the comparative tests of Level 3,4 and 5 standard neural network, and the accuracy results obtained are shown in Table 1 below.



Image Recognition Method	Level 3 Neural Network	Level 4 Neural Network	Level 5 Neural Network
Elastic degree	73.2	98.2	98.1
Adaptive degree	66.8	97.3	97.2
Standard degree	95.6	95.1	96.5

		_				
Tahle 1	Accuracy	1 comparison	of algorithms	' image reco	omition (%)	
	Accuracy	(comparison	of argorithmis	innage reet	ginuon (70)	

Summary

Deep learning model has the ability to extract image features independently, and also can extract abstract features. And if the data scale is larger, the extracted feature has better recognition ability. Based on deep convolution updating network, this paper builds target detection and recognition framework. On the basis of the analysis on in-depth architecture error transmission, it proposes parameter learning method of self-adaptive elastic momentum of index. Simulation results show that compared with artificial feature detection algorithms, this method has higher accuracy and low loss rate. And compared with homologous error optimization algorithm, it has faster convergence speed and smoother convergence curve.

Acknowledgements

This work is supported by Key Technological Project on Agricultural Innovation of Shaanxi Science and Technology Department(2016NY-134) and National Natural Science Foundation of China(11641002).

References

- Shyam Sunder Agrawal and Vinod Yadava, Materials and Manufacturing Processes, Vol.19 (2013) No.4, p.122.
- [2] P.Asvestas, F. K. Matsopoulos and K.S. Nikita, Journal of Visualization, Vol. 24 (2007) No.7, p.84.
- [3] S Luo and B Ren, Comput Methods Programs Biomed, Vol. 34 (2016) No.130, p.154.
- [4] Bengio Y, Guyon G and Dror V, Workshop on Unsupervised & Transfer Learning, Vol.33 (2011) No.7, p.531.
- [5] Hinton G E., Momentum, Vol. 9 (2010) No.1, p.599.