2017 IEEE 2nd International Conference on Big Data Analysis

Simple Convolutional Neural Network on Image Classification

Tianmei Guo, Jiwen Dong ,Henjian Li，Yunxing Gao
Department of Computer Science and Technology Shandong Provincial Key Laboratory of Network based Intelligent Computing University of Jinan, Jinan, China mengkeer_ad@163.com, ise_dongjw@ujn.edu.cn

Abstract—**In recent years, deep learning has been used in image classification, object tracking, pose estimation, text detection and recognition, visual saliency detection, action recognition and scene labeling. Auto Encoder, sparse coding, Restricted Boltzmann Machine, Deep Belief Networks and Convolutional neural networks is commonly used models in deep learning. Among different type of models, Convolutional neural networks has been demonstrated high performance on image classification. In this paper we bulided a simple Convolutional neural network on image classification. This simple Convolutional neural network completed the image classification. Our experiments are based on benchmarking datasets minist [1] and cifar-10. On the basis of the Convolutional neural network, we also analyzed different methods of learning rate set and different optimization algorithm of solving the optimal parameters of the influence on image classification.**

Keywords-Convolutional neural network; Deep learning; Image classification; learning rate; parametric solution

I. INTRODUCTION

Image classification plays an important role in computer vision, it has a very important significance in our study, work and life. Image classification is process including image preprocessing, image segmentation, key feature extraction and matching identification. With the latest figures image classification techniques, we not only get the picture information faster than before, we apply it to scientific experiments, traffic identification, security, medical equipment, face recognition and other fields.

During the rise of deep learning, feature extraction and classifier has been integrated to a learning framework which overcomes the traditional method of feature selection difficulties. The idea of deep learning is to discover multiple levels of representation, with the hope that high-level features represent more abstract semantics of the data. One key ingredient of deep learning in image classification is the use of Convolutional architectures.

Convolutional neural network design inspiration comes from the mammalian visual system structure[1]. Visual structure model based on the cat visual cortex was proposed by Hubel and Wiesel in 1962. The concept of receptive field has been proposed for the first time. The first hierarchical structure Neocognition used to process images was proposed by Fukushima in 1980. The Neocognition adopted the local connection between neurons, can make the network translation invariance. Convolutional neural network is first introduced by LeCun in [1] and improved in [2]. They developed a multi-layer artificial neural network called LeNet-5 which can classify handwriting number. Like other neural network, LeNet-5 has multiple layers and can be trained with the backpropagation algorithm [3]. However, due to the lack of large training data and computing power at that time. LeNet-5 cannot perform well on more complex problems, such as large-scale image and video classification.

Since 2006, many methods have been developed to overcome the difficulties encountered in training deep neural networks. Krizhevsky propose a classic CNN architecture Alexnet [4] and show significant improvement upon previous methods on the image classification task. With the success of Alexnet [4], several works are proposed to improve its performance. ZFNet [5],VGGNet [6] and GoogleNet [7] are proposed.

In recent years, the optimization of Convolutional neural network are mainly concentrated in the following aspects: the design of Convolutional layer and pooling layer, the activation function, loss function, regularization and Convolutional neural network can be applied to practical problems.

In this paper we proposed a simple Convolutional neural network on image classification. On the basis of the Convolutional neural network, we also analyzed different methods of learning rate set and different optimization algorithm of solving the optimal parameters of the influence on image classification. We also study the composition of the different Convolutional neural network on the result of image classification.

II. BASIC CNN COMPONENTS

Convolutional neural network layer types mainly include three types, namely Convolutional layer, pooling layer and fully-connected layer. Fig. 1 shows the architecture of LeNet-5[1] which is introduced by Yann LeCun.

Figure 1. The architecture of LeNet-5[1] network(adapted from [1])

A. Convolutional Layer

Convolutional layer is the core part of the Convolutional neural network, which has local connections and weights of shared characteristics. The aim of Convolutional layer is to learn feature representations of the inputs. As shown in above, Convolutional layer is consists of several feature maps. Each neuron of the same feature map is used to extract local characteristics of different positions in the former layer, but for single neurons, its extraction is local characteristics of same positions in former different feature map. In order to obtain a new feature, the input feature maps are first convolved with a learned kernel and then the results are passed into a nonlinear activation function. We will get different feature maps by applying different kernels. The typical activation function are sigmoid, tanh and Relu[15].

B. Pooling Layer

The sampling process is equivalent to fuzzy filtering. The pooling layer has the effect of the secondary feature extraction, it can reduce the dimensions of the feature maps and increase the robustness of feature extraction. It is usually placed between two Convolutional layers. The size of feature maps in pooling layer is determined according to the moving step of kernels. The typical pooling operations are average pooling[16] and max pooling[17].We can extract the high level characteristics of inputs by stacking several Convolutional layer and pooling layer.

C. Fully-connected Layer

In general, the classifier of Convolutional neural network is one or more fully-connected layers. They take all neurons in the previous layer and connect them to every single neuron of current layer. There is no spatial information preserved in fully-connected layers. The last fully-connected layer is followed by an output layer. For classification tasks, softmax regression is commonly used because of it generating a well-performed probability distribution of the outputs. Another commonly used method is SVM, which can be combined with CNNs to solve different classification tasks[18].

III. Our Simple Convolutional Neural Network

Observe the development process of the depth Convolutional neural network, it is obvious that the expression ability of the network is enhanced with the depth of the network. Time complexity of the same two network structure, the deeper the performance of the network will have a relatively improved. However, the network is not as deep as possible. As the depth of the network increases, the memory consumption will be more and more, and the network performance may not improve.

Based on this idea, we builed a simple Convolutional neural network on image classification. On the basis of the Convolutional neural network, we mainly analyzed different methods of learning rate set and different optimization algorithm of solving the optimal parameters of the influence on image classification.

The following shows the data flow of the network structure(Our data flow graph is based on benchmarking datasets minist).

Our simple Convolutional neural network introduces relu[15] and dropout[19]. The standard sigmoid output does not have sparsity. In order to produce sparse data, the standard sigmoid needs to use some penalty factors to train a lot of close to 0 redundant data. In addition unsupervised pre-training is needed. Relu[15] is a linear correction and purelin's polyline, the formula is: g (x) = max (0, x). The formula let the value equal to 0 if the output value is less than 0, otherwise keep the original values. This is a simple and brutal way to force some data to 0, but the practice has proved that the network is fully trained with a moderate sparse. Moreover, the visualization effect of training is similar to that of the traditional method, which indicates that relu[15] has the ability to guide the sparse. Training Convolutional neural network, when the iteration times increase, there will be a good fit training set, but the degree of fitting to the verification set is very poor. We follow a certain probability on the weight of the parameters of random sampling, selection of updates in the training process. Dropout[19] is the model training in the random network layer hidden layer node weights do not work, but to retain its weight, but not to be updated.

A. Data Flow Diagram of Convolutional Layer 1

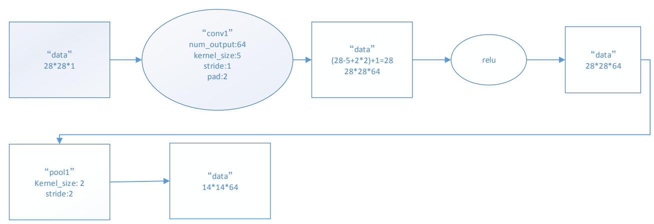

Figure 2. The data flow diagram of convolutional layer 1

B. Data Flow Diagram of Convolutional Layer 2

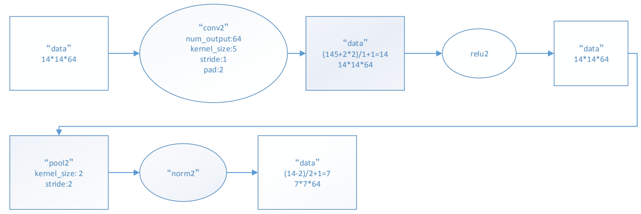

Figure 3. The data flow diagram of convolutional layer 2

C. Data Flow Diagram of Convolutional Layer 3

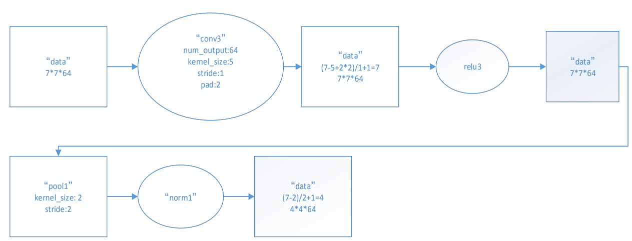

Figure 4. The data flow diagram of convolutional layer3

D. Data Flow Diagram of ip Layer 1

Figure 5. The data flow diagram of ip layer1

E. Data Flow Diagram of ip Layer 2

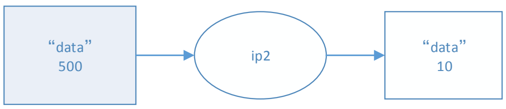

Figure 6. the data flow diagram of ip layer2

IV. EXPERIMENTAL RESULT AND ANALYSIS

A. Learning Rate Set and Algorithm of Solving the Optimal Parameters

The parameters of the convolutional neural network need to be solved by the gradient descent algorithm. In the iterative process, the basic learning rate needs to be adjusted, and the adjustment strategy has different choices. In the following, we give the influence of different solving strategies on the experimental results and give the corresponding analysis.

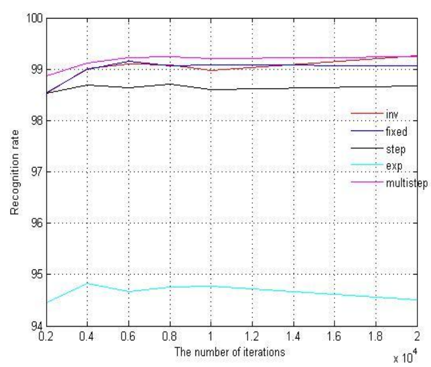

Figure.7: different learning rate set in training data

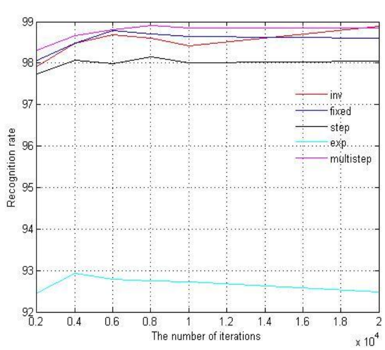

Figure.8: different learning rate set in testing data

As can be seen from the figure above, with the increase of the number of iterations, the recognition rate of each algorithm has improved and multistep is the best in those algorithms.

The basic learning rate of multistep will be adjusted with the value of stepvalue, so that we can learn more excellent parameters. The fixed method keeps the basic learning rate constant and has universal applicability. This setting has some empirical value, the general set to a smaller value because it will help the network learning and convergence.

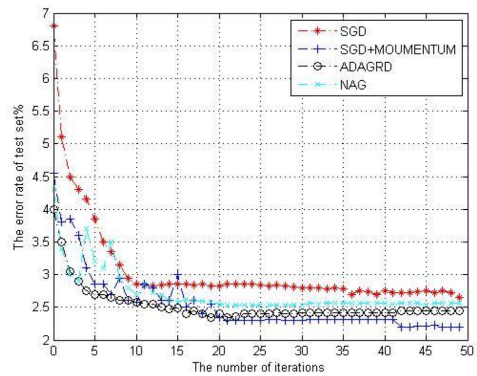

Figure.9: different solving strategies in testing data

From the above experimental results, it can be seen that the SGD can reduce the error with increasing the number of iterations, but the descending speed and effect are general. SGD with momentum and NAG are better than SGD, but the gradient descends slowly in the early stage, and then gradually becomes stable with the increase of the iteration number, and SGD with momentum is better than NAG. The error rate of test set is decline relatively stable in ADAGRAD, the curve is smooth and the effect is better.

B. The Result of MNIST

The MNIST dataset[2] consisits of 60000 28×28 grayscale images of handwritten digits 0 to 9. The classification results of our simple convolutional neural network are as follows:

TABLE I. THE RESULT OF MNIST

Method	Error rate
Regularization of Neural Networks using DropConnect[20]	0.21%
Multi-column Deep Neural Networks for Image Classification[21]	0.23%
APAC: Augmented Pattern Classification with Neural Networks[22]	0.23%
Batch-normalized Max-out Network in Network[23]	0.24%
Generalizing Pooling Functions in Convolutionalal Neural Networks: Mixed, Gated, and Tree[24]	0.29%
Fractional Max-Pooling[25]	0.32%
Max-out network (k=2) [26]	0.45%
Network In Network [27]	0.45%
Deeply Supervised Network [28]	0.39%
RCNN-96 [29]	0.31%
Our simple Convolutional neural network	0.66%

Compared with the existing methods, although our recognition rate is not the highest, but our network structure is simple, and parameters take up memory is small. We also verify that the shallow network also has a relatively good recognition effect.

V. SUMMARY

In this paper we proposed a simple Convolutional neural network on image classification. This simple convolutional neural network imposes less computational cost. On the basis of the convolutional neural network , we also analyzed different methods of learning rate set and different optimization algorithm of solving the optimal parameters of the influence on image classification. We also verify that the shallow network also has a relatively good recognition effect.

ACKNOWLEDGMENT

This work is supported by grants by National Natural Science Foundation of China (Grant No. 61303199)

REFERENCES

[1] LecunY, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.

[2] Cun Y L, Boser B, Denker J S, et al. Handwritten digit recognition with a back-propagation network[C]// Advances in Neural Information Processing Systems. Morgan Kaufmann Publishers Inc. 1990:465.

[3] Hecht-Nielsen R. Theory of the backpropagation neural network[M]// Neural networks for perception (Vol. 2). Harcourt Brace & Co. 1992:593-605 vol.1.

[4] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2):2012.

[5] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in ECCV, 2014.

[6] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in ICLR, 2015.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with Convolutionals,” Co RR, vol. abs/1409.4842, 2014.

[8] Hinton G E. Deep belief networks[J]. Scholarpedia, 2009, 4(5): 5947.

[9] Scott rozelle, Xue Lei Xu Yangming, etc. The depth of the study review [J]. Computer application research, 2012, 29 (8) : 2806-2810.

[10] lili guo, shifei ding. Deep learning research progress [J]. 2015.

[11] FanYaQin Wang Binghao, wang wei, etc. Depth study domestic research review [J]. China distance education, (6) : 2015-27 to 33.

[12] Liu Jinfeng. A concise and effective to accelerate the Convolutional of the neural network method [J]. Science, technology and engineering, 2014 (33) : 240-244.

[13] Markoff J. How many computers to identify a cat? 16,000[J]. New York Times, 2012.

[14] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(8): 1798-1828.

[15] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in ICML, 2010, pp. 807–814.

[16] T. Wang, D. Wu, A. Coates, and A. Ng, “End-to-end text recognition with Convolutionalal neural networks,” in International Conference on Pattern Recognition (ICPR), 2012, pp. 3304–3308.

[17] Y. Boureau, J. Ponce, and Y. Le Cun, “A theoretical analysis of feature pooling in visual recognition,” in ICML, 2010, pp. 111–118.

[18] Y. Tang, “Deep learning using linear support vector machines,” arXiv prprint arXiv:1306.0239, 2013.

[19] S. Wang and C. Manning, “Fast dropout training,” in ICML, 2013.

[20] Wan, L., et al. Regularization of neural networks using dropconnect. in Proceedings of the 30th International Conference on Machine Learning (ICML-13). 2013.

[21] Ciregan,D.,U.Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. 2012. IEEE.

[22] Sato, I., H. Nishimura, and K. Yokoi, APAC: Augmented PAttern Classification with Neural Networks. ar Xiv preprint ar Xiv:1505.03229, 2015.

[23] Jia-Ren Chang,Y.-S.C.,Batch-normalized Maxout Network in Network.2015.ar Xiv:1511.02583v1.

[24] Lee, C.-Y., P.W. Gallagher, and Z. Tu. Generalizing pooling functions in Convolutionalal neural networks: Mixed, gated, and tree. in International Conference on Artificial Intelligence and Statistics. 2016.

[25] Graham, B., Fractional Max-Pooling. 2014. Ar Xiv e-prints, December 2014a.

[26] Ian J. Goodfellow, D.W.-F., Mehdi Mirza,Aaron Courville,Yoshua Bengio, Maxout networks. 2013. 28: p. 1319-1327.

[27] Lin, M., Q. Chen, and S. Yan, Network In Network. Co RR, 2013. abs/1312.4400.

[28] Lee, C.-Y., Xie, Saining, Gallagher, Patrick W., Zhang, Zhengyou, and Tu, Zhuowen.,

[29] Deeply supervised nets. In Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2015, San Diego, California, USA, May9-12, 2015.

[30] Hu, M.L.a.X., Recurrent Convolutionalal neural network for object recognition. 2015.