Mongolian Handwriting Character Recognition Based On Convolutional Neural Network(CNN)

Narankhuu Natsagdorj

College Of Computer Engineering, Shanghai Maritime University,

Pudong District, Shanghai, China

email: narankhuu0308@gmail.com

Key words: Convolutional neural network, LeNet-5, Mongolian Cyrillic alphabet.

Abstract. The ability of multilayer networks with increasing complexity, multidimensional and nonlinear mappings from large collections of examples can be good for image recognition tasks. In this case, Conventional neural networks(CNN) is compatible for image recognition task as it contains multiple layers of small neuron collection. In this paper, we accomplish Mongolian character recognition by using algorithm of Convolutional neural network. The input of network is normalized with the images of Mongolian Cyrillic character. Using back-propagation method efficient for programs with their own databases. LeNet-5 neural network regulate weight value and threshold value of the convolutional neural network and enable the program to achieve the minimum error. Finally, this article analyzes the training and testing error rate with research data.

Introduction

The main purpose of this paper is to show that it is possible to create CNN algorithm for Mongolian handwriting. Convolutional neural networks have three main purposes to ensure of some degree of shift, scale and distortion invariance like local receptive fields, shared weights and temporal sub-sampling. Convolutional neural network always has more layer as like LeNet-5 is a specific convolutional neural network which has 7 layers not counting input layers. Each layer contains trainable parameter such as weight. Input is 32x32 pixel image. LeNet-5 is suitable for as it recognizes characters one by one.

In this paper, The Mongolian Cyrillic alphabet is the writing system used for the standard dialect of the Mongolian language in the modern state of Mongolia. It has a largely phonemic orthography, meaning that there is a fair degree of consistency in the representation of individual sounds. But in the Inner Mongolia, they still use the Traditional Mongolian script. Only the Outer Mongolia use the Mongolian Cyrillic alphabet. The database described in here was constructed from the own created database. The own created database containing binary images of handwritten characters. This database contains 24 character which each one is duplicated more than 300. The original black and white images size is set to fit in 32x32 pixel. So that background corresponds to a value of -0.1 and the foreground corresponds to 1.175. The mean input is roughly 0, with a variance of roughly 1, which accelerates learning. Database used by three versions as following:

Firstly, the version the images were centered in 28x28 image by computing the center of mass of the pixels. And it translates the image so as to position this point at the center of the 28x28 field. In some situation, this field was extended 32x32 with background pixels. That database version will be referred to the regular database.

Secondly, the character images were cropped to 20x20 pixels, so that it computes the second moment of inertia of the pixel which counting a foreground pixel as 1 and background pixel 0. Also, the shear images shift horizontally, so that the principles axis is vertical. That version database will be referred to as the slanted database.

Thirdly, the Database's version is used some early experiments, the images were reduced to 16x16 pixel. In the following, picture 1 shows an example of randomly picked test set.

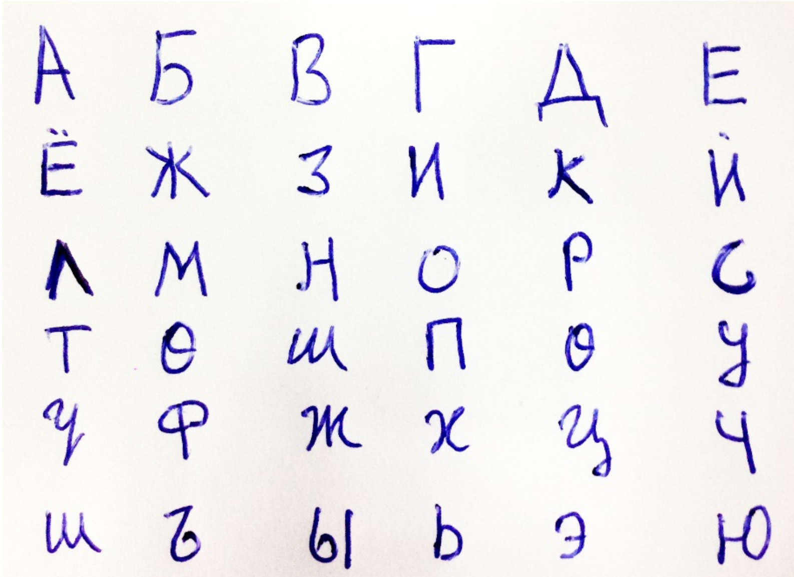

Figure 1. Randomly picked image examples from Mongolian Cyrillic character database.

Network model

In LeNet-5, all the layers are divided into three groups as Convolutional and are labeled as C_x, Sub-Sampling layer are labeled S_x , Fully Connected layer are labeled F_x Where x is layer's index.

The Function of Convolutional layer. The layer is characterized by a convolution kernel maps which can convolute, and through an activation function, we can get the output feature map. Each output value map may be a combination of a plurality of input maps of convolution

$$x_i^l = f(\sum_{i \in M_j} x_i^{l-1} * K_{ij}^l + B_j^l)$$

M_j maps indicating the selection of the set of input. We have a choice of a pair or three. But here we will discuss how to automatically select the desired combination of feature maps. Each output map will give an additional bias B, but for a specific convolution map, inputs are not the same. If the output feature map j and output feature map k from the input map i are the summation of convolution, the convolution kernel corresponding is not the same.

The Function of Sub-Sampling layer. For sub-sampling layer, there are n input and output maps, but each output map is smaller.

$$x_i^l = f(\frac{1}{n} \sum_{i \in M_j} X_i^{l-1} + B^i)$$

So that the output image in two dimensions are reduced n times. Each output corresponds to a map of their own bias β and a multiplicative additive bias b.

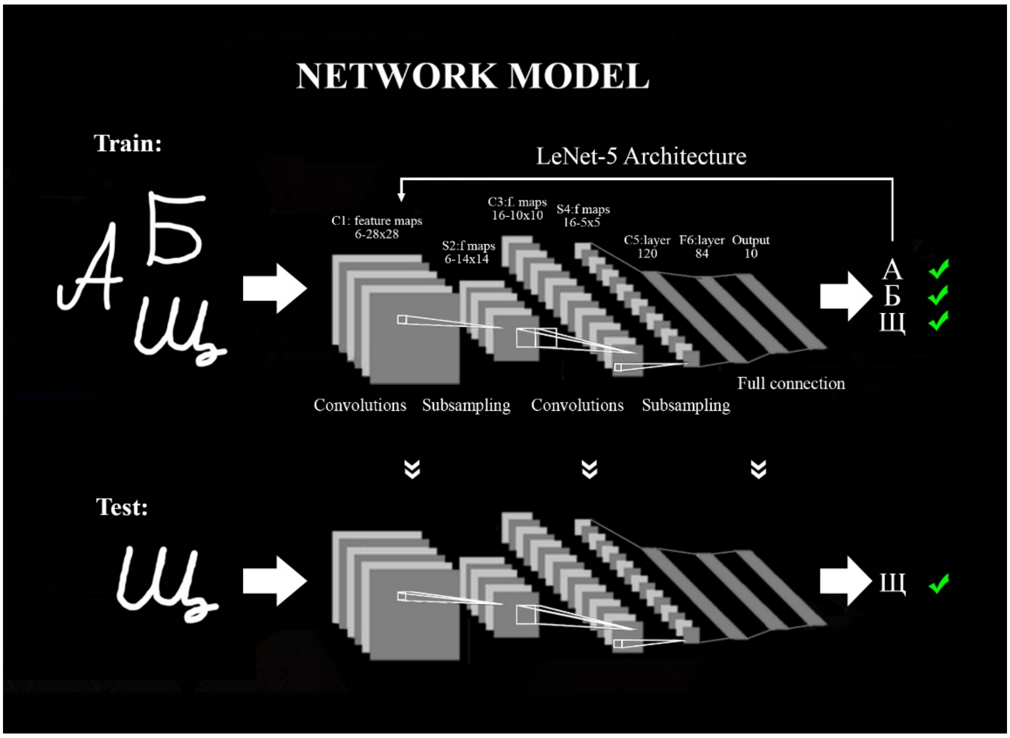

Figure 2. Showing network main model, first doing train by every character then doing test one by one.

Experiments and Result Analysis

Trial and error result. As Following table 1 showing trial error rate by every single Cyrillic character. The error rate on the training set reaches around 8.1% after 300 passes. Result is clear that a larger training epoch set could improve the performance of LeNet-5.

Table 1. Trial error rate by every single Cyrillic character

№	Character	Epoch		
		300	500	1000
1	"А"	8.1% Error rate	5.375% Error rate	2.82% Error rate
2	"Б"	8.2% Error rate	5.335% Error rate	2.91% Error rate
3	"Д"	8.4% Error rate	5.375% Error rate	2.91% Error rate
4	"З"	8.2% Error rate	6.375% Error rate	2.83% Error rate
5	"П"	8.1% Error rate	5.345% Error rate	2.71% Error rate
6	"М"	8.2% Error rate	5.375% Error rate	2.81% Error rate
7	"Н"	8.3% Error rate	5.375% Error rate	2.72% Error rate
8	"О"	8.8% Error rate	5.375% Error rate	2.91% Error rate
9	"У"	8.8% Error rate	5.375% Error rate	2.98% Error rate
11	"Ф"	8.7% Error rate	5.345% Error rate	2.91% Error rate
12	"Х"	8.7% Error rate	5.311% Error rate	2.92% Error rate
13	"Ү"	8.6% Error rate	5.113% Error rate	2.98% Error rate
14	"Ъ"	8.3% Error rate	5.375% Error rate	2.93% Error rate
15	"Э"	8.2% Error rate	6.120% Error rate	2.98% Error rate
16	"Щ"	8.8% Error rate	5.175% Error rate	2.91% Error rate
17	"Ш"	8.7% Error rate	5.298% Error rate	2.92% Error rate
18	"П"	8.3% Error rate	5.275% Error rate	2.91% Error rate
19	"Г"	8.4% Error rate	5.180% Error rate	2.88% Error rate
20	"Ж"	8.5% Error rate	5.208% Error rate	2.88% Error rate
21	"Ө"	8.6% Error rate	5.301% Error rate	2.89% Error rate
22	"В"	8.5% Error rate	5.375% Error rate	2.94% Error rate
23	"Р"	8.3% Error rate	5.350% Error rate	2.82% Error rate

Test error rate result. The following table 2 shows test error rate by every single Cyrillic character. The error rate on each test set reaches around 19.4% after 300 passes. Result is clear that a larger trial epoch set could improve the performance of LeNet-5

Table 2. Trial error rate by every single Cyrillic character

№	Character	Epoch		
		300	500	1000
1	“А”	19.1% Error rate	14.375% Error rate	11.3% Error rate
2	“Б”	21.2% Error rate	13.375% Error rate	11.3% Error rate
3	“Д”	8.4% Error rate	6.375% Error rate	1.9% Error rate
4	“З”	15.2% Error rate	11.375% Error rate	7.9% Error rate
5	“П”	14.1% Error rate	7.345% Error rate	3.7% Error rate
6	“М”	7.2% Error rate	4.375% Error rate	1.1% Error rate
7	“Н”	24.3% Error rate	13.375% Error rate	10.7% Error rate
8	“О”	3.8% Error rate	2.375% Error rate	1% Error rate
9	“У”	15.8% Error rate	9.375% Error rate	6.9% Error rate
10	“Ф”	8.8% Error rate	5.375% Error rate	1.1% Error rate
11	“Х”	6.7% Error rate	3.345% Error rate	1.1% Error rate
12	“Ү”	21.7% Error rate	18.311% Error rate	15.7% Error rate
13	“Ъ”	19.6% Error rate	13.113% Error rate	11.3% Error rate
14	“Э”	20.3% Error rate	15.375% Error rate	6.1% Error rate
15	“Щ”	28.2% Error rate	26.120% Error rate	22.2% Error rate
16	“Ш”	27.8% Error rate	25.175% Error rate	21.1% Error rate
17	“П”	29.7% Error rate	27.298% Error rate	24.2% Error rate
18	“Г”	7.3% Error rate	5.275% Error rate	2.4% Error rate
19	“Ж”	17.4% Error rate	14.180% Error rate	11.3% Error rate
20	“Ө”	6.5% Error rate	5.208% Error rate	2.2% Error rate
21	“В”	13.6% Error rate	11.301% Error rate	8.1% Error rate
22	“С”	13.5% Error rate	12.375% Error rate	8.4% Error rate
23	“Р”	16.3% Error rate	12.350% Error rate	10.8% Error rate

The following table 3 shows 100 tested “Ө” Mongolian Cyrillic character. 98 answer is correct, but 2 answer is incorrect. Reason is that “Ө” Cyrillic character and “B” Cyrillic character’s stroke so similar. And these two character’s stroke style different from another characters, that is why correct answer has a rating high.

Table 3: **“Ө” Correct 98 ----** Incorrect 2

1-10	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
11-20	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
21-30	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
31-40	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
41-50	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
51-60	Ө	Ө	Ө	B	Ө	Ө	Ө	Ө	Ө	Ө
61-70	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
71-80	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
81-90	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө
91-100	Ө	Ө	Ө	Ө	Ө	Ө	Ө	Ө	B	Ө

Conclusion

In Conclusion, we accomplish Mongolian character recognition by using convolutional neural network. We mainly apply the algorithm of Convolutional neural network to recognize Mongolian character one by one. We can conclude that LeNet-5 architecture has suitable structure for image recognition task. Experiment results show that we have achieved good results by evaluating trial error rate and testing error compered by every single character, as well as analyzing every character recognizing result.

References

[1] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 81(11).
[2] Y. Lecun, B. Boser, J.S.Denker, D.Henderson, R.E. Howard, W. Hubbard, and L.D Jackel, "Backpropagation applied to handwritten zip code recognition," Neural Computation, vol.1 no.4, pp. 541-551, Winter 1989.
[3] Y. Lecun and Y. Bengio. Convolutional networks for images, speech, and time-series. In M. A. Arbib editor, The handbook of Brain Theory and Neural Networks. MIT Press, 1995
[4] L. Bottou, Y. LeCun, and Y.Bengio. Global training of document processing systems using graph transformer networks. In Proc. of Computer Vision and Pattern Recognition, Puerto-Rico, 1997. IEEE.
[5] Y. LeCun, L. Bottou, and Y.Bengio. Reading checks with graph transformer networks. In International Conference on Austics, Speech, and Signal Processing, volume 1, pages 151-154, Munich, 1997. IEEE.
[6] Y.LeCun, L.D. Jackel, L. Bottou, A. Brunot, C.Cortes, J.S. Denker, H. Drucker, I. Guyon, U.A. Muller, E. Sackinger, P.Simard and V. Vapnik. Comparison of learning algorithms for handwritten digit recognition. In F. Fogelman and P. Gallinari, editors, Internation Conferences on Artificial Neural Networks, page 53-60, Paris, 1995. EC2 & Cie.
[7] R.Valliant, C.Monrocq, and Y. LeCun, Original approach for the localisation of objects in images. IEEE Proc on Vision, Image and Signal Processing, 141(4):245-250, August 1994.
[8] Ofer Matan, Christoper J.C. Burges, Yann LeCun, and John S. Denker. Multi-digit recognition using a space displacement neural network. In J. M. Moody, S. J. Hanson, and R. P. Lippman, editors, Neural Information Processing Systems, volume 4. Morgan Kaufman Publishers, San Mateo CA, 1992.
[9] Y. Bengio, Y.LeCun, C. Nohl, C.Burges, AT&T Bell Laboratories. A NN/HMM Hybrid for On-Line Handwriting Recognition.
[10] Y. LeCun, P.Sermanet, S. Chintala, Convolutional Neural Network Applied to House Number Digit Classification.
[11] J. Fan, W. Xu, Y. Wu and Y.Gong. Human Tracking Using Convolutional Neural Networks. Neural Networks, IEEE Transaction on 2010.
[12] Y. Lecun and C. Cortes. The MNIST Database of handwritten digits.
[13] P. Sermanet, Y. Lecun. Traffic Sign Recognition with Multi-Scale Convolutional Networks. In Proceedings of International Joint Conferences on Neural Networks, 2011.
[14] Y. Bengio, R.De Mori and G. Flammia, R.Kompe 1992. Global Optimization of a Neural Network- Hidden Markov Model Hybrid. IEEE Transaction on Neural Networks v.3.
[15] L. Bottou, and P. Gallinari, 1991, A Framework for the Cooperation Algorithms. In M. Lippman and Touretzky editors, Advances in Neural Information Processing System volume 3. Denver, Co.
[16] C.Tappert, C. Suen, and T. Wakahara, 1990. The State Of The Art In On-Line Handwriting recognition. IEEE Trans.