Effects of Parallel Structure and Serial Structure on Convolutional Neural Networks

Xiaoya Chen1a, Baoheng Xu1, Han Lu1
Southwest Petroleum University
School of Sciences
Chengdu, China

*Corresponding author’s e-mail: 202021000610@stu.swpu.edu.cn

Abstract—Based on KERAS’ FASHION_MNIST data set, a convolutional neural network with serial structure and parallel structure was set up for training. The serial structure can be used for deep feature mining and the parallel structure can be used to describe the first impression. The results show that the parallel structure has better effect and faster speed for simple images.

1. Introduction
As we all know, at the beginning of the development of neural network, it is hoped that computers can replace human beings to distinguish and fit the basic things existing in the world. For example, in 1943, Warren McCulloch, a psychologist, and Walter Pitts [1], a mathematical logician proposed and gave the concept of artificial neural network and the mathematical model of artificial neuron, thus opening the era of artificial neural network research.

In 1949, psychologist Donald Hebb described the principles of neuron learning in a paper in The Organization of Behavior. Further, Frank Rosenblatt[3], an American neuroscientist, proposed a machine that could simulate human perception ability and called it "perceptron". In 1957, at Cornell Aeronautical Laboratory, he successfully simulated the perceptron on IBM704, and in 1960, he implemented a perceptron based neural computer, Mark1, that could recognize some English letters. Mark1’s ability to categorize simple shapes (such as triangles and quadrilaterals) is increasingly recognized as a trend toward using machines similar to human senses, learning, memory, and recognition.

In 1985, Geoffrey Hinton [4] used multiple hidden layers to replace the original single feature layer in the perceptron, and calculated network parameters using BP algorithm (Station in 1974).

In 1989, Yann Lecun et al. used deep neural networks to identify the handwritten characters of postcodes in letters. Later, Lecun further used CNN to complete handwritten character recognition of bank checks, and the recognition accuracy reached the commercial level.

Traditional neural networks are Sequential, meaning that the Nth layer of input is the N-1th layer of output, and can be represented by the technical term Sequential structure. However, researchers gradually found that the use of parallel structure would enhance the features in the process of convolution kernel scanning data, making the model output more accurate.

Zhuang Jiayi parallelized the convolutional neural network and the gated circulation unit neural network, extracted local features and timing features respectively, splitted the outputs of the two network structures and input them into the deep neural network, and carried out ultra-short-term load prediction by DNN. The results show that compared with GRU NN network structure, long and short term memory...
network structure, serial CNN-LSTM network structure and serial CNN-GRU network structure, the method proposed by the dealer has better predictive performance.

In this paper, based on the FASHION_MNIST data set over Keras, we will train the neural network with traditional serial structure and the neural network with parallel structure respectively, and compare the performance and efficiency of the two.

2. The principle of CNN

In the process of image processing, matrix convolution is often used to calculate the image feature. There are two kinds of matrix convolution: full convolution and valid Convolution.

The definition of full convolution is:

\[
z(u,v) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i,j} \cdot k_{u-i,v-j}
\]  

(1)

Assuming that \( \chi \) is \( m \times m \) order matrix, \( K \) is \( n \times n \) order matrix, \( K_{rot} \) is obtained by rotating \( K \) by 180°.

The definition of valid Convolution is:

\[
z(u,v) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i+u,j+v} \cdot k_{rot,i,j} \cdot \chi(i,j)
\]  

(2)

In type 2, \( \chi(i,j) = 1, 0 \leq i, j \leq n \\
0, \text{ others} \)

In the convolutional layer, the data exists in three dimensions. At the input layer, if it is a gray image, there is only one feature image; if it is a color image, there are generally three feature images. The features of the upper layer will be convolved with the corresponding convolution kernel to output the new features. We assume that the input layer is the L-1\(^{th}\) layer, its input feature graph is \( \chi^{(l-1)} (m \times m) \), and the convolution kernel corresponding to the feature is \( K^{(l)} (n \times n) \). We add a bias term \( B^{(l)} \) to each output, so the output \( Z^{(l)} ((m-n+1) \times (m-n+1)) \) of the convolutional layer is:

\[
Z_{u,v}^{(l)} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} x_{i+u,j+v}^{(l-1)} \cdot k_{rot,i,j}^{(l)} \cdot \chi(i,j) + b^{(l)}
\]  

(3)

In type 3, \( \chi(i,j) = 1, 0 \leq i, j \leq n \\
0, \text{ others} \) , \( a_{u,v}^{(l)} = f(Z_{u,v}^{(l)}) \)

3. Experimental results and analysis

3.1. Data loading and visualization

The fashion MNIST data set in Keras was used for the experiment, and the first sample was taken, as shown in Figure 1.
In the training set, the shape of the image sample is (60000 28 28) and the shape of the label is (600001). As you can see from figure 1, the range of pixels is [0,255], so the normalization is completed by dividing all images by 255.0.

25 processed images are extracted, as shown in Figure 2.

![Normalized data example](image)

Figure 2. Example of normalized data

After counting all the labels, there are 10 categories, namely: T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, Ankle boot.

### 3.2. Model structure and visualization

In this paper, two neural networks with different structures are constructed, in which model 1 is a serial structure and Model 2 is a parallel structure. Both models are convolutional neural networks, which contain the same input layer, output layer and full connection layer. The difference is that the two convolutional layers of Model 1 are serial structures, that is, the input of the latter layer is the output of the former layer. The two convolutional layers of Model 2 are parallel structures, that is, the input of the previous layer is simultaneously obtained and output to the next layer.

The function of convolutional layer is to extract features. When the serial structure model is used to scan an image, the input accepted by the latter layer is the result of the convolution kernel scanning of the former layer, which contains the image features scanned. When the second convolutional layer scans the results of the first convolutional layer, features will be further screened to select valid features.

When the model with parallel structure is used to scan an image, two convolutional layers process an image at the same time, and then output and fuse the scanning results at the same time.

The method emphasizes intuition, that is, the characteristics obtained from the initial scan.

FIG. 2 and FIG. 3 are network structure diagrams of serial structure CNN and parallel structure CNN respectively.
3.3. Model performance and visualization

The two models adopt the same initialization method, the optimizer adopts Adam algorithm uniformly, and the learning rate is 1E-3. Train each model for 5 rounds respectively, and count the time of each round. Then the training effect of Model 1 is:

![Figure 5. Training effect of Model 1](image)

The training effect of Model 2 is as follows:

![Figure 6. Training effect of Model 2](image)

After 5 rounds of training, the minimum loss and maximum accuracy of Model 1 were 1.67 and 0.87 respectively, and the average time of each round was 25 seconds.

After 5 rounds of training, the minimum loss and maximum accuracy of Model 2 were 1.57 and 0.89, and the average time of each round was 16 seconds.

4. Conclusion
The parallel structure is more efficient for the recognition of simple graphics. The advantage of serial structure is that it can dig the features deeply, so it is suitable for complex cases. If time is not taken into account, and accuracy is guaranteed, then the serial structure may perform better at the end, because it
can dig out deeper features.

REFERENCES
[1] McCulloch W S, Pitts W H. A logical calculus of the ideas immanent in nervous activity[J]. The Bulletin of Mathematical Biophysics, 1943, 5:115-133.
[2] Hebb D O. The Organization of Behavior: A Neuropsychological Theory[M]. John Wiley, Chapman & Hall, 2013.
[3] Malsburg C V D. Frank Rosenblatt: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms[J]. 1986.
[4] Geoffrey Hinton. Experiments on Learning by Back Propagation[J]. 1986
[5] Barbosa D, Coury D V, Oleskovicz M. New approach for power transformer protection based on intelligent hybrid systems[J]. IET Generation, Transmission & Distribution, 2012, 6(10): 1009-1018
[6] Jazebi S, Vahidi B, Jannati M. A novel application of wavelet based SVM to transient phenomena identification of power transformers[J]. Energy Conversion and Management, 2011, 52(2): 1354-1363
[7] Wu W C, Ji T Y, Li M S, et al. Using mathematical morphology to discriminate between internal fault and inrush current of transformers[J]. IET Generation, Transmission & Distribution, 2009, 29(7): 98-105.