VGG16

1. Introduction

VGGNet divides the network into 5 segments, each segment connects multiple 3 \* 3 convolutional networks in series, each segment is followed by a maximum pooling layer, and finally there are 3 fully connected layers and a softmax layer.

The model explores the relationship between the depth of the convolutional neural network and its performance. By repeatedly stacking 3 \* 3 small convolution kernels and 2 \* 2 maximum pooling layers, 16 ~ 19 deep convolutions are successfully constructed. Product neural network. VGGNet all uses 3 \* 3 convolution kernels and 2 \* 2 pooling kernels to improve performance by continuously deepening the network structure. The increase in the number of network layers will not bring about an explosion in the amount of parameters, because the amount of parameters is mainly concentrated in the last three fully connected layers.

According to the size of the convolution kernel and the number of convolution layers, VGG can be divided into 6 ConvNet Configurations: A, A-LRN, B, C, D, E Among them, D and E are more commonly used and are called VGG16 and VGG19, respectively.

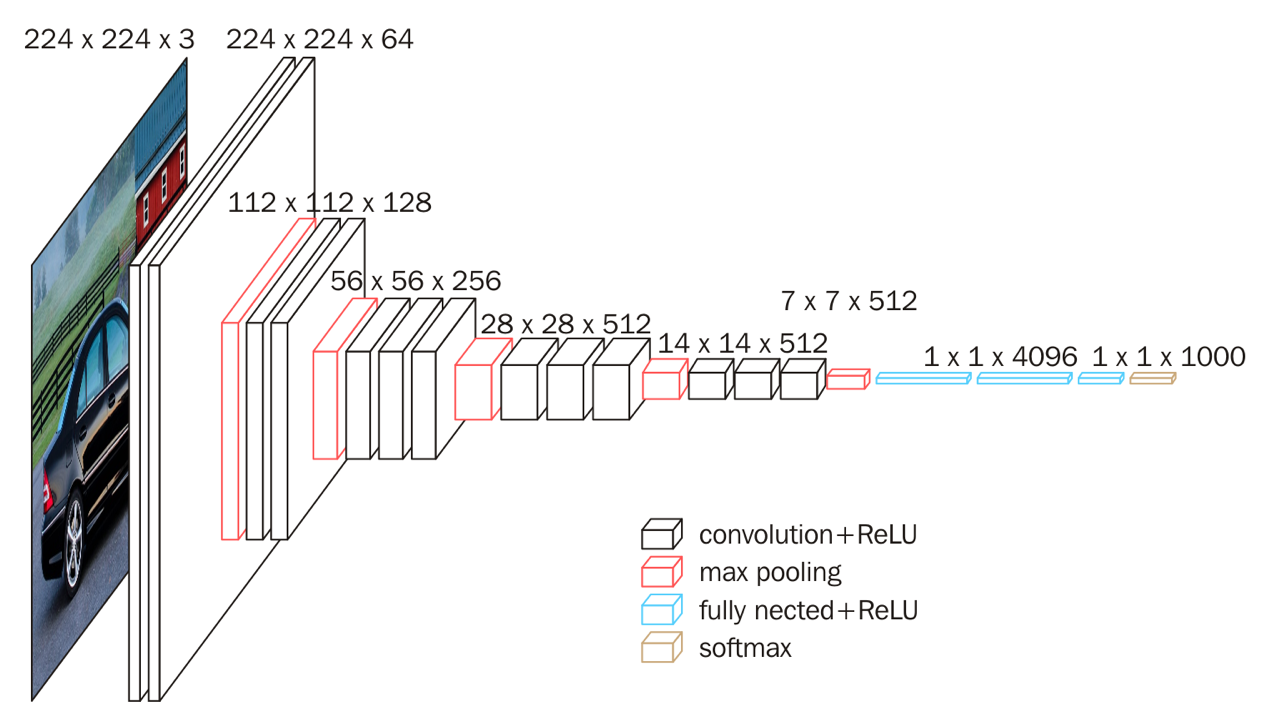


Fig1. Specific structure diagram of VGG16.[1]

The outstanding feature of VGG16 is simplicity, which is reflected in: 1. The convolutional layers all use the same convolution kernel parameters. 2. The pooling layer adopts the same pooling nuclear parameters. 3. The model is composed of a stack of convolutional layers and pooling layers, and it is relatively easy to form a deep network structure. Based on the above analysis, the advantages of VGG can be summarized as: Small filters, Deeper networks.

2. Result

2.1 Result in Dataset MNIST

The results of training the MNIST dataset and the curves of Train Loss and Train Accuracy with VGGNnet model are shown in Fig. 2 and Fig. 3.

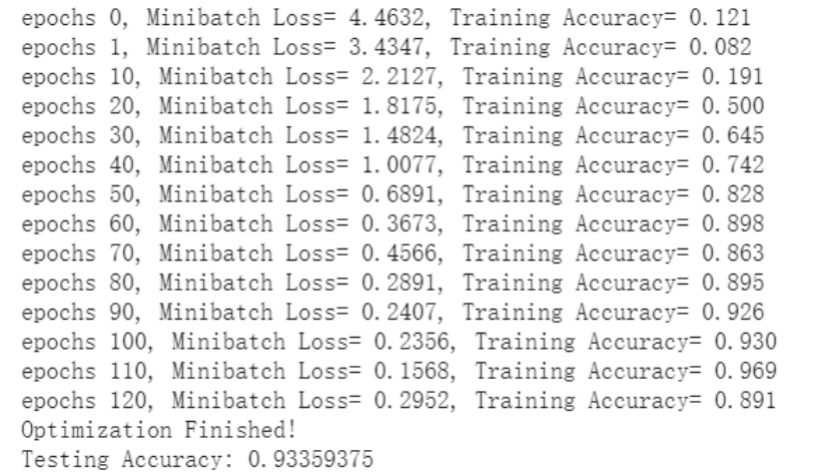


Fig. 2. Accuracy in MNIST with VGGNet

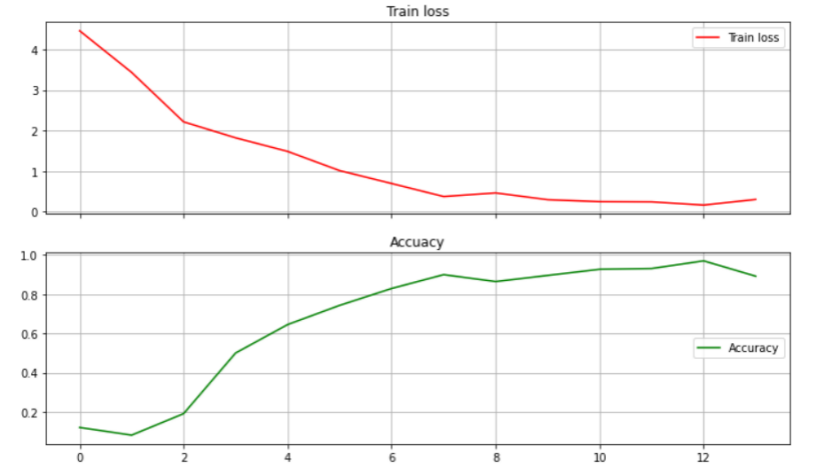


Fig. 3. Train Cost and Accuracy in MNIST with VGGNet

2.2 Result in Dataset CIFAR-10

The results of training the CIFAR-10 dataset with AlexNet model are as follows. As shown in Fig. 4, the accuracy is 89.790%. Besides, the curves of Train Loss and Train Accuracy are shown in Fig. 5.

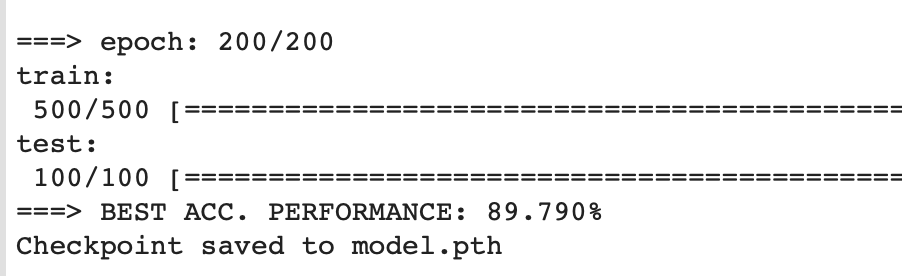


Fig. 4. Accuracy in CIFAR-10 with VGGNet

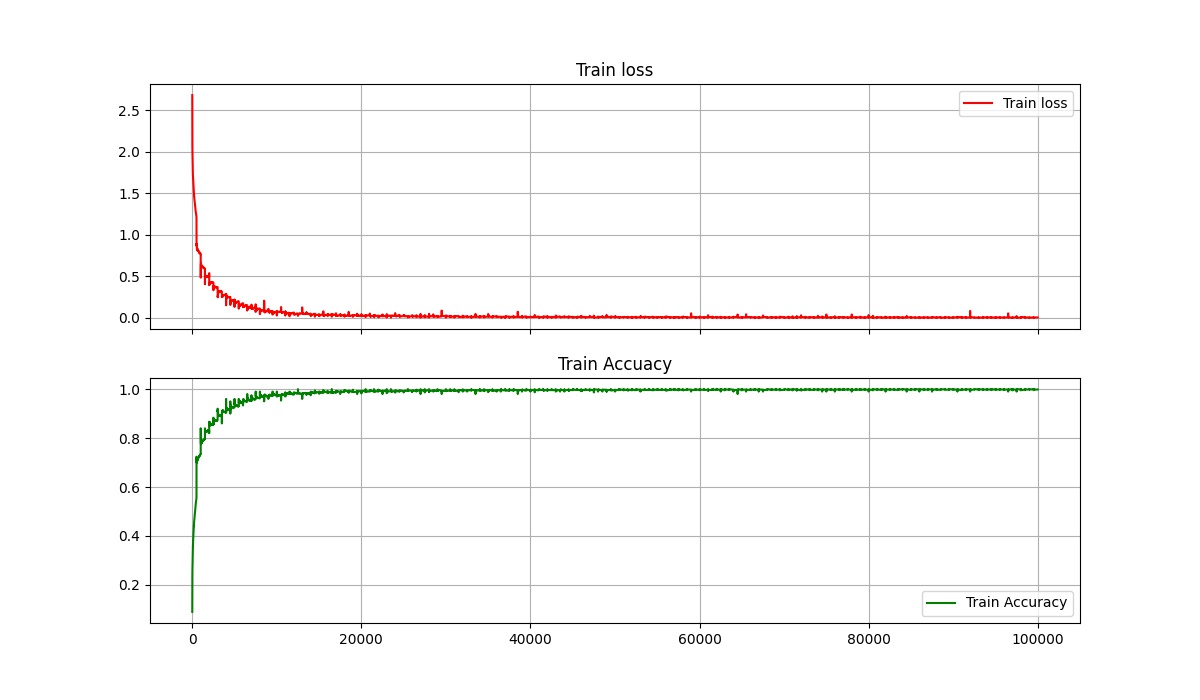


Fig. 5. Train Loss and Train Accuracy in CIFAR-10 with VGG16

Reference

[1] Hussam Qassim, Abhishek Verma, David Feinzimer. Compressed residual-VGG16 CNN model for big data places image recognition[C]// IEEE Computing & Communication Workshop & Conference. IEEE, 2018.