Vgg16 structure

Abel - 2022年1月21日



Structure intro

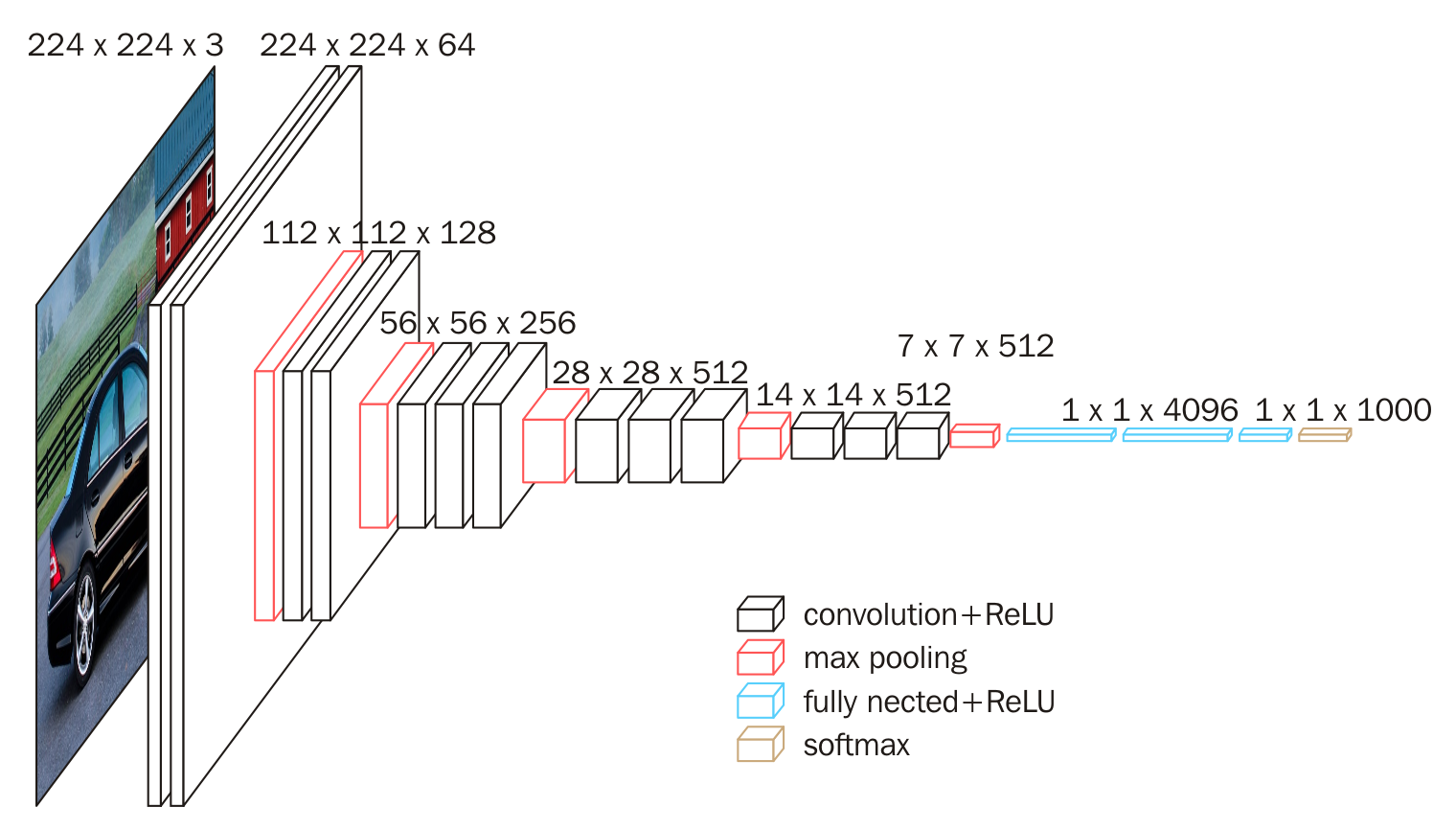
注释：VGG16 名字就显示出来：13个卷积层及3个全连接层。

VGG是由Simonyan 和Zisserman在文章《Very Deep Convolutional Networks for Large Scale Image Recognition》中提出卷积神经网络模型。

在这篇论文中分别使用了A、A-LRN、B、C、D、E这6种网络结构进行测试，这6种网络结构相似，都是由5层卷积层、3层全连接层组成，其中区别在于每个卷积层的子层数量不同，从A至E依次增加（子层数量从1到4），总的网络深度从11层到19层（添加的层以粗体显示）.

VGG16是13个卷积层及3个全连接层的特定版本，还有19层的版本。图像

原来论文图引用说明



vgg16网络结构说明

Structure and workflow

Working with vgg16: The 16 in VGG-16 means that the number of layers with trainable parameters in the entire network is 16. In VGG-16, the filter width and height of all convolutional layers are 3, the stride is 1, and the padding is the same convolution; the filter width and height of all pooling layers are 2, and the stride is 2.

VGG-16 structure:

input layer

convolutional layer

convolutional layer

pooling layer

convolutional layer

convolutional layer

pooling layer

convolutional layer

convolutional layer

convolutional layer

pooling layer

convolutional layer

convolutional layer

convolutional layer

pooling layer

convolutional layer

convolutional layer

convolutional layer

pooling layer

fully connected layer

fully connected layer

Fully connected layer, output layer

In VGG, 3 3x3 convolution kernels are used instead of 7x7 convolution kernels, and 2 3x3 convolution kernels are used instead of 5\*5 convolution kernels. The main purpose of this is to ensure the conditions of the same perception field. , the depth of the network is improved, and the effect of the neural network is improved to a certain extent. Its working flow:

1. Input a 224x224x3 image, and do two convolution + ReLU through 64 3x3 convolution kernels, and the size after convolution becomes 224x224x64

2. For max pooling (maximum pooling), the size of the pooling unit is 2x2 (the effect is that the image size is halved), and the size after pooling becomes 112x112x64

3. After 128 3x3 convolution kernels for twice convolution + ReLU, the size becomes 112x112x128

4. Do 2x2 max pooling pooling, and the size becomes 56x56x128

5. After 256 3x3 convolution kernels for three convolution + ReLU, the size becomes 56x56x256

6. Do 2x2 max pooling pooling, and the size becomes 28x28x256

7. After 512 3x3 convolution kernels for three convolution + ReLU, the size becomes 28x28x512

8. Make 2x2 max pooling pooling, and the size becomes 14x14x512

9. After 512 3x3 convolution kernels for three convolution + ReLU, the size becomes 14x14x512

10. Do 2x2 max pooling pooling, and the size becomes 7x7x512

11. Full connection + ReLU with two layers of 1x1x4096 and one layer of 1x1x1000 (three layers in total)

12. Output 1000 prediction results through softmax