

卷积神经网络的训练过程



元沫

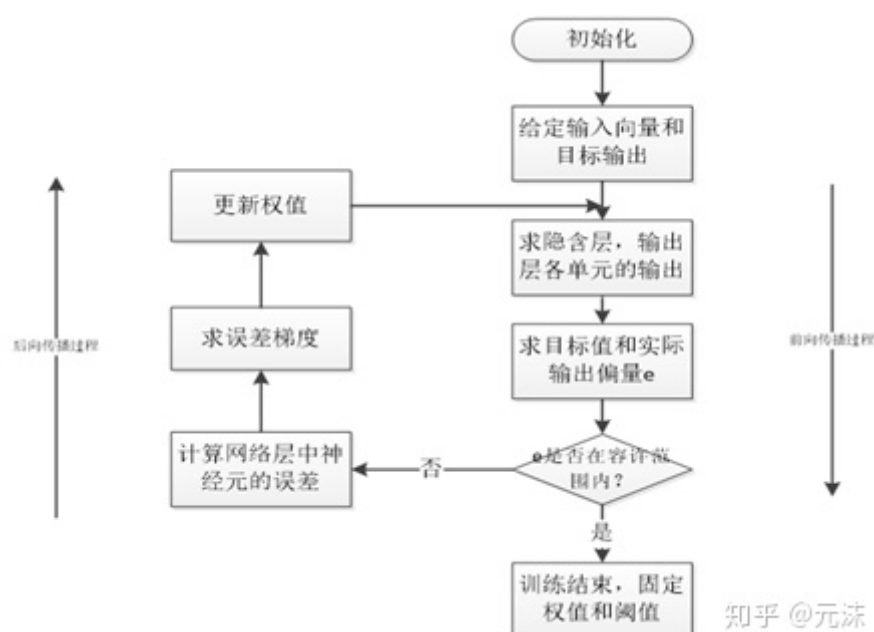
[关注她](#)

60 人赞同了该文章

卷积神经网络的训练过程

卷积神经网络的训练过程分为两个阶段。第一个阶段是数据由低层次向高层次传播的阶段，即前向传播阶段。另外一个阶段是，当前向传播得出的结果与预期不相符时，将误差从高层次向底层次进行传播训练的阶段，即反向传播阶段。训练过程如图4-1所示。训练过程为：

- 1、网络进行权值的初始化；
- 2、输入数据经过卷积层、下采样层、全连接层的向前传播得到输出值；
- 3、求出网络的输出值与目标值之间的误差；
- 4、当误差大于我们的期望值时，将误差传回网络中，依次求得全连接层，下采样层，卷积层的误差。各层的误差可以理解为对于网络的总误差，网络应承担多少；当误差等于或小于我们的期望值时，结束训练。
- 5、根据求得误差进行权值更新。然后在进入到第二步。



知乎 @元沫

1.1 卷积神经网络的前向传播过程

在前向传播过程中，输入的图形数据经过多层卷积层的卷积和池化处理，提出特征向量，将特征向量传入全连接层中，得出分类识别的结果。当输出的结果与我们的期望值相符时，输出结果。

1.1.1 卷积层的向前传播过程

卷积层的向前传播过程是，通过卷积核对输入数据进行卷积操作得到卷积操作。数据在实际的网络中的计算过程，我们以图3-4为例，介绍卷积层的向前传播过程。其中一个输入为15个神经元的图片，卷积核为 $2 \times 2 \times 1$ 的网络，即卷积核的权值为 W_1, W_2, W_3, W_4 。那么卷积核对于输入数据的卷积过程，如下图4-2所示。卷积核采用步长为1的卷积方式，卷积整个输入图片，形成了局部感受野，然后与其进行卷积算法，即权值矩阵与图片的特征值进行加权和（再加上一个偏置量），然后通过激活函数得到输出。

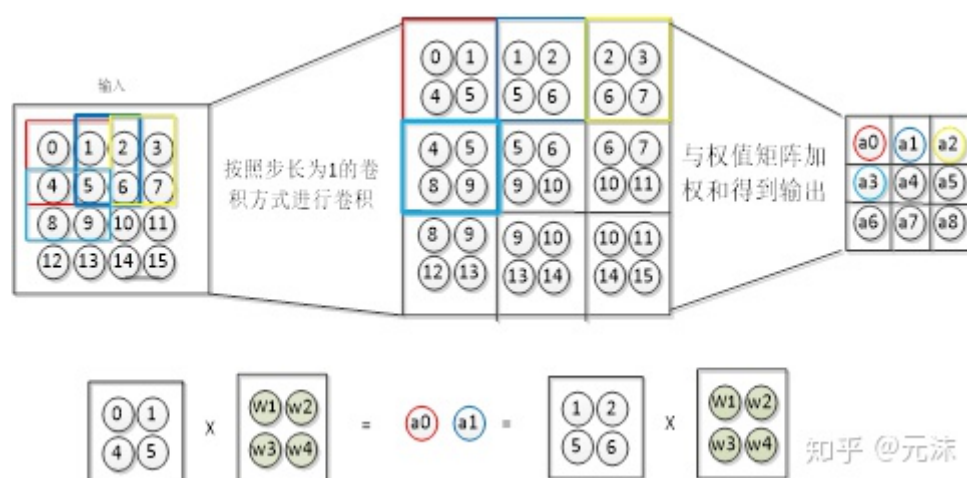


图4-2 图片深度为1，卷积层的向前传播过程

而在图3-4中，图片深度为2时，卷积层的向前传播过程如图4-3所示。输入的图片的深度为 $4 \times 4 \times 2$ ，卷积核为 $2 \times 2 \times 2$ ，向前传播过程为，求得第一层的数据与卷积核的第一层的权值的加权和，然后再求得第二层的数据与卷积核的第二层的权值的加权和，两层的加权和相加得到网络的输出。

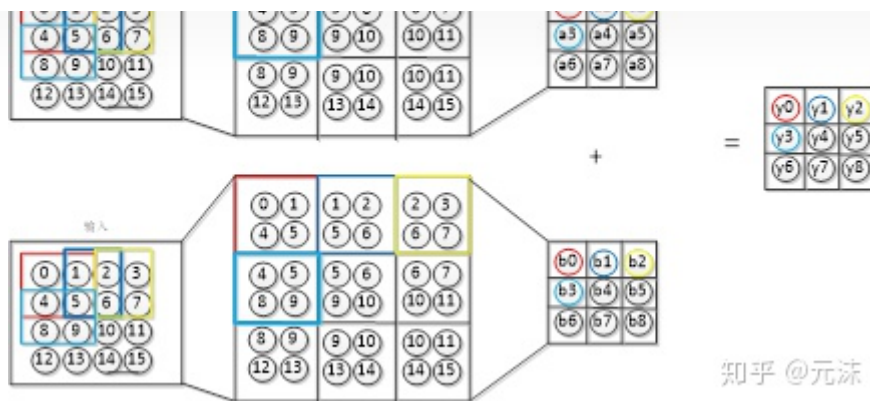


图4-3图片深度为2，卷积层的向前传播过程

1.1.2下采样层的向前传播过程

上一层（卷积层）提取的特征作为输入传到下采样层，通过下采样层的池化操作，降低数据的维度，可以避免过拟合。如图4-4中为常见的池化方式示意。最大池化方法也就是选取特征图中的最大值。均值池化则是求出特征图的平均值。随机池化方法则是先求出所有的特征值出现在该特征图中的概率，然后在来随机选取其中的一个概率作为该特征图的特征值，其中概率越大的选择的几率越大。

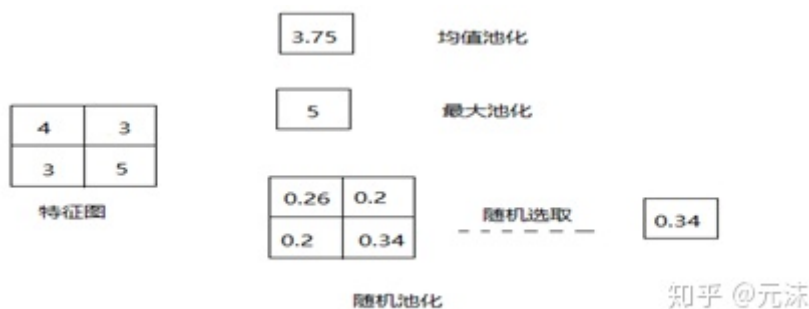


图4-4池化操作示意图

1.1.3全连接层的向前传播过程

特征图进过卷积层和下采样层的特征提取之后，将提取出来的特征传到全连接层中，通过全连接层，进行分类，获得分类模型，得到最后的结果。图4-5为一个三层的全连接层。假设卷积神经网络中，传入全连接层的特征为 x_1, x_2 。则其在全连接层中的向前传播过程如图4-5所示。第一层全连

下一层。

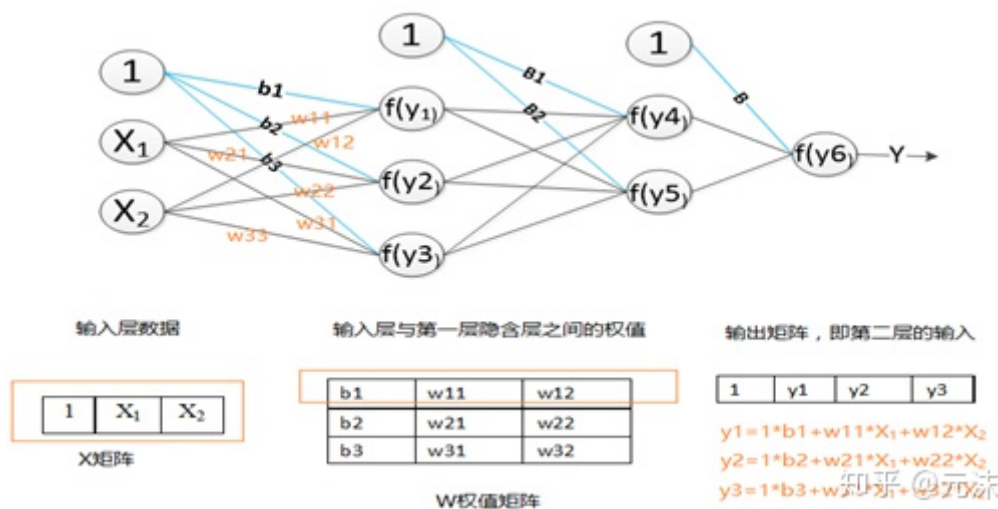


图4-5全连接层的向前传播过程

1.2卷积神经网络的反向传播过程

当卷积神经网络输出的结果与我们的期望值不相符时，则进行反向传播过程。求出结果与期望值的误差，再将误差一层一层的返回，计算出每一层的误差，然后进行权值更新。该过程的主要目的是通过训练样本和期望值来调整网络权值。误差的传递过程可以这样来理解，首先，数据从输入层到输出层，期间经过了卷积层，下采样层，全连接层，而数据在各层之间传递的过程中难免会造成数据的损失，则也就导致了误差的产生。而每一层造成的误差值是不一样的，所以当我们求出网络的总误差之后，需要将误差传入网络中，求得该各层对于总的误差应该承担多少比重。

反向传播的训练过程的第一步为计算出网络总的误差：求出输出层n的输出a(n)与目标值y之间为误差。计算公式为：

$$\delta(n) = -(y - a(n)) * f'(z(n))$$

其中，为激励函数的导函数的值。

1.2.1全连接层之间的误差传递

求出网络的总误差之后，进行反向传播过程，将误差传入输出层的上一层全连接层，求出在该层由

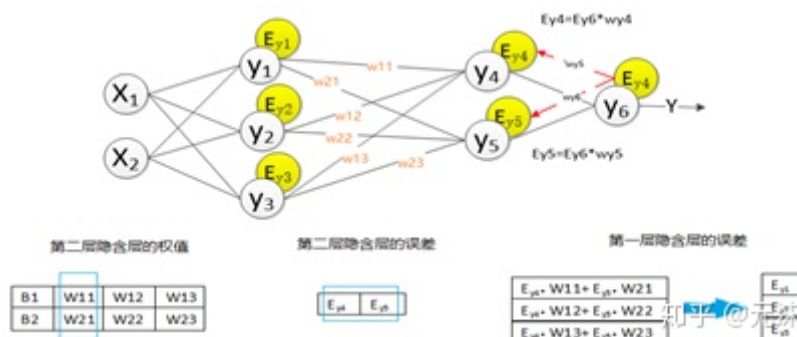


图4-6 全连接层中误差的传递过程

1.2.2 当前层为下采样层，求上一层的误差

在下采样层中，根据采用的池化方法，把误差传入到上一层。下采样层如果采用的是最大池化（max-pooling）的方法，则直接把误差传到上一层连接的节点中。果采用的是均值池化（mean pooling）的方法，误差则是均匀的分布到上一层的网络中。另外在下采样层中，是不需要进行权值更新的，只需要正确的传递所有的误差到上一层。

1.2.3 当前层为卷积层，求上一层的误差

卷积层中采用的是局部连接的方式，和全连接层的误差传递方式不同，在卷积层中，误差的传递也是依靠卷积核进行传递的。在误差传递的过程，我们需要通过卷积核找到卷积层和上一层的连接节点。求卷积层的上一层的误差的过程为：先对卷积层误差进行一层全零填充，然后将卷积层进行一百八十度旋转，再用旋转后的卷积核卷积填充过程的误差矩阵，并得到了上一层的误差。如图4-7为卷积层的误差传递过程。图右上方为卷积层的向前卷积过程，而右下方为卷积层的误差传递过程。从图中可以看出，误差的卷积过程正好是沿着向前传播的过程，将误差传到了上一层。

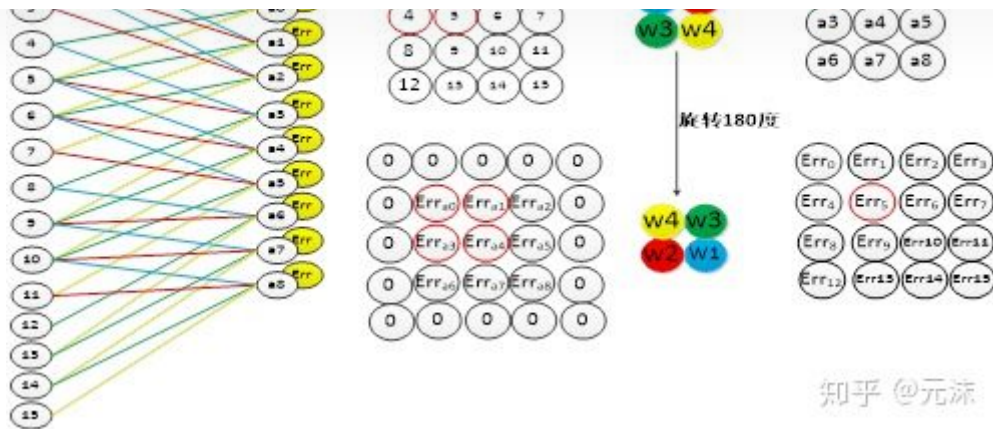


图4-7卷积层的误差传递过程

1.3 卷积神经网络的权值更新

1.3.1 卷积层的权值更新

卷积层的误差更新过程为：将误差矩阵当做卷积核，卷积输入的特征图，并得到了权值的偏差矩阵，然后与原先的卷积核的权值相加，并得到了更新后的卷积核。如图4-8，图中可以看出，该卷积方式的权值连接正好和向前传播中权值的连接是一致的。

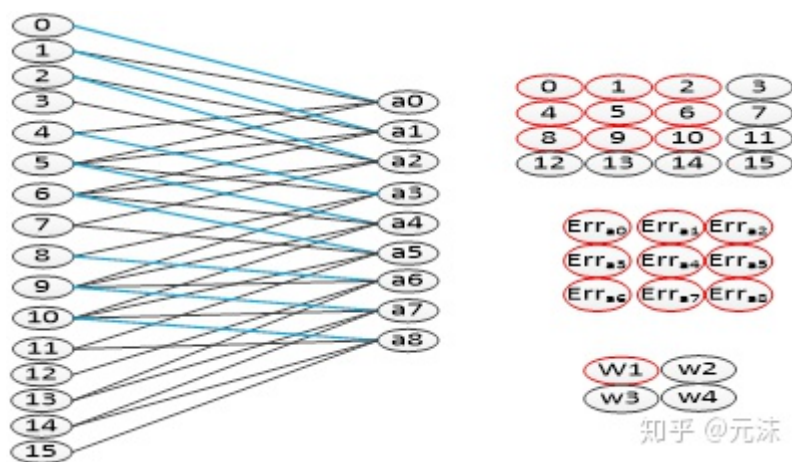


图4-8卷积核的权值更新过程

1.3.2 全连接层的权值更新过程

2、原先的权值加上偏导值，得到新的权值矩阵。具体的过程如图4-9所示（图中的激活函数为Sigmoid函数）。

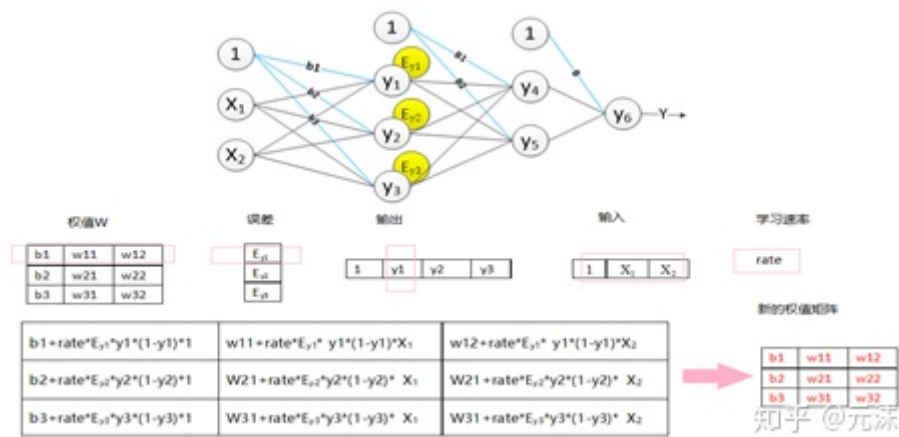


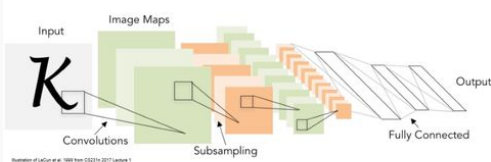
图4-9全连接层的权值更新过程

发布于 2018-05-09 11:53

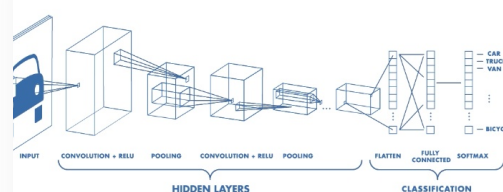
卷积神经网络 (CNN) 深度学习 (Deep Learning) 神经网络 卷积

推荐阅读

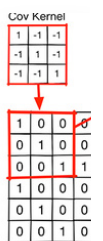
Next: Convolutional Neural Networks



卷积神经网络



卷积神经网络学习路线（一） | 卷积神经网络的组件以及卷积...



卷积神经网络