

1.全连接层的作用

2.全局平均池化的理解

2.1 传统的池化的作用

2.2 全连接层

2.3 全局平均池化

1.全连接层的作用

- 全连接层也是一种卷积层。
- 它的参数基本和卷积层的参数一样，只是它的卷积核大小和原数据大小一致。
- 起到将学到的“分布式特征表示”映射到样本标记空间的作用，“分类器”的作用
- 用 global average pooling **取代** FC，已经成为了**大势所趋**。

目前由于全连接层参数冗余（仅全连接层参数就可占整个网络参数80%左右），近期一些性能优异的网络模型如ResNet和GoogLeNet等**均用全局平均池化（global average pooling, GAP）取代FC**来融合学到的深度特征，最后仍用softmax等损失函数作为网络目标函数来指导学习过程。需要指出的是，用GAP替代FC的网络通常有较好的预测性能。

那么为什么 **全连接层参数冗余（仅全连接层参数就可占整个网络参数80%左右）** 呢？

因为全连接层的卷积核横截面做得和 输入的 feature map 一样大。而

常规卷积层的卷积核横截面只有一个小滑窗那么大。很明显，二者的参数数量级根本就不在一个level上。

2.全局平均池化的理解

2.1 传统的池化的作用

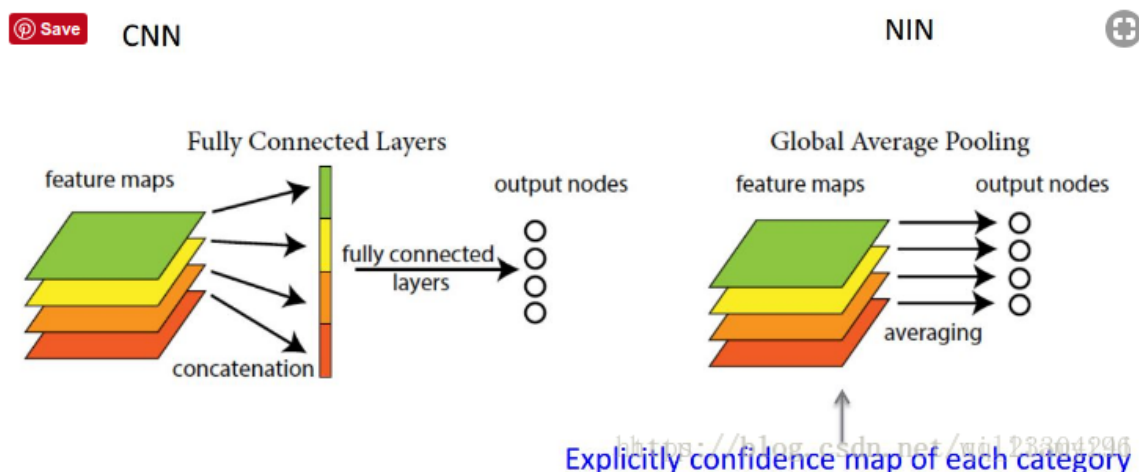
众所周知CNN网络中常见结构是：卷积、池化和激活。卷积层是CNN网络的核心，激活函数帮助网络获得非线性特征，而池化的作用则体现在降采样：保留显著特征、降低特征维度，增大kernel的感受野。深度网络越往后面越能捕捉到物体的语义信息，这种语义信息是建立在较大的感受野基础上。因此对feature map降维，进而增大后面各层kernel的感受野是一件很重要的事情。另外一点值得注意：pooling也可以提供一些旋转不变性。

2.2 全连接层

很长一段时间以来，全连接网络一直是CNN分类网络的标配结构。一般在全连接后会有激活函数来做分类，假设这个激活函数是一个多分类softmax，那么全连接网络的作用就是将最后一层卷积得到的feature map stretch成向量，对这个向量做乘法，最终降低其维度，然后输入到softmax层中得到对应的每个类别的得分。

全连接层如此的重要，以至于全连接层过多的参数重要到会造成过拟合，所以也会有一些方法专门用来解决过拟合，比如dropout。

2.3 全局平均池化



GAP的真正意义是：对整个网络在结构上做正则化防止过拟合。既要参数少避免全连接带来的过拟合风险，又要能达到全连接一样的转换功能，怎么做呢？直接从feature map的通道上下手，如果我们最终有1000类，那么最后一层卷积输出的feature map就只有1000个channel，然后对这个feature map应用全局池化，输出长度为1000的向量，直接剔除了全连接层中黑箱的特征，直接赋予了每个channel实际的类别意义。

global average pooling 的最后输出结果仍然是 10 个 feature map，而不是一个，只不过每个 feature map 只剩下一个像素罢了。这个像素就是求得的平均值。

实践证明其效果还是比较可观的，同时GAP可以实现任意图像大小的输入。但是值得我们注意的是，使用gap可能会造成收敛速度减慢。