**Week 7 Thinking**

郭想

参数共享指的是什么？

参数共享（权值共享）意思就是整张图片（输入）在使用同一个卷积核内的参数，比如一个3\*3\*1的卷积核，**这个卷积核内9个的参数被整张图共享，而不会因为图像内位置的不同而改变卷积核内的权系数**。或者说，就是用一个卷积核不改变其内权系数的情况下卷积处理整张图片。

我们让输出层学习的权重和偏差在所有输入层的小块之间共享。要注意的是此时随着筛选器深度的增加，权重和偏置的数量仍然会增加，因为权重不会再输出通道之间共享。

我在网上看到参数共享的额外好处是带来较强的灵活性，如果我们没有在所有小块中共享相同的权重和偏差，我们就必须让每个小块和隐藏层神经元学习新的参数。如果输入是大分辨率的图像，学习效率将会大打折扣。因此,共享参数不仅有助于我们实现平移不变性,而且还为我们提供了一个更小、更可扩展的模型。

为什么会用到batch normalization ?

通过一篇康奈尔大学的文章，我们了解到batch normalization在CNN当中起到的作用。对于一个卷积神经网络的每一层来说，由于前一层的参数更新，所以这一层的输入（前一层的输出）的分布会发生变化，这种现象被称之为ICS（**Internal Covariate Shift**）。

所以，BN就是通过一定的规范化手段，把每层神经网络任意神经元　这个输入值的分布强行拉回到均值为0方差为1的标准正态分布。因此，每一层的输入（上一层输出经过BN后）分布的稳定性都提高了，故而整体减小了网络的ICS。与此同时，BN也使得非线性变换函数的输入值落入对输入比较敏感的区域，从而避免梯度消失问题。这样输入的小变化就会导致损失函数较大的变化（使得梯度变大，避免梯度消失问题产生），同时也让收敛速度更快，加快训练速度。

使用dropout可以解决什么问题？

dropout指的是在神经网络的训练过程中提出的一种**防止过拟合**的策略。

策略旨在训练过程中按照一定的概率（一般情况下：隐藏层采样概率为0.5，输入层采样概率为0.8）随机删除网络中的神经元（输出层除外）。

大致的原理是因为Dropout技术使得某两个神经元不一定每次都在一个子网络结构中出现。基于此权值的更新不在依赖于固定关系的隐含节点的共同作用，阻止了某些特征仅仅在其他特征下才能有效的情况。迫使网络去学习更加鲁棒的特征（更加具有通适性）。