

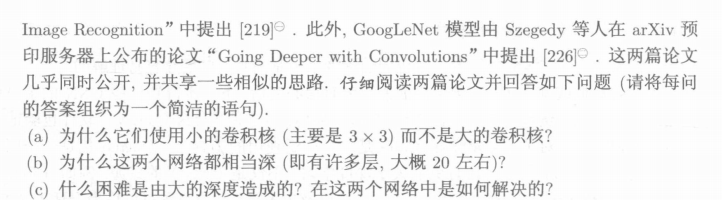
（a）在训练过程中，Dropout的实现是让神经元以超参数 (p) 的概率停止工作或者激活被置为0

（b）在测试过程中，不进行随机失活，而是将Dropout的参数p乘以输出。

（c）随机失活通常降低了神经网络的计算开销。

（d）因为归零操作可以得到稀疏矩阵。



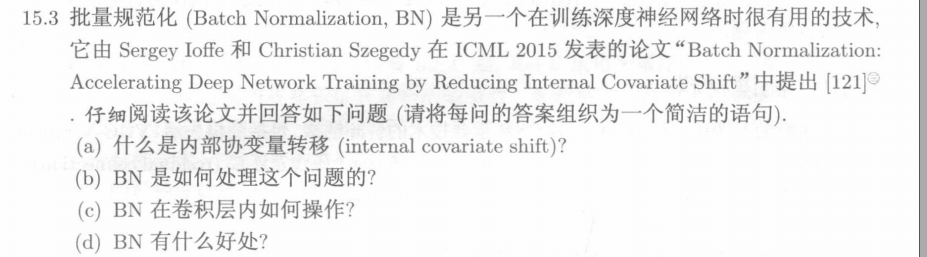


1. 因为这样可以减少参数个数。
2. 因为增加深度可以提高神经网络的性能。

（c）（1）容易过拟合；（2）计算需要的资源增加，消耗内存，算力等资源。

VGG16的网络的解决办法为加深结构全部使用ReLU激活函数，以此来提升非线性变化的能力。

GoogLeNet团队提出了Inception的新型网络结构，来搭建一个稀疏性、高计算性能的网络结构。



1. 每层输入的分布在训练过程中会发生变化，因为前面的层的参数会发生变化。通过要求较低的学习率和仔细的参数初始化减慢了训练，并且使具有饱和非线性的模型训练起来非常困难。我们将这种现象称为内部协变量转移。
2. 它通过对神经网络的中间层进行逐层归一化，让每一个中间层输入的分布保持稳定，即保持同一分布。

（c）（1）求数据均值

（2）求数据方差

（3）数据进行标准化

（4）训练参数γ，β

（5）输出y通过γ与β的线性变换得到新的值

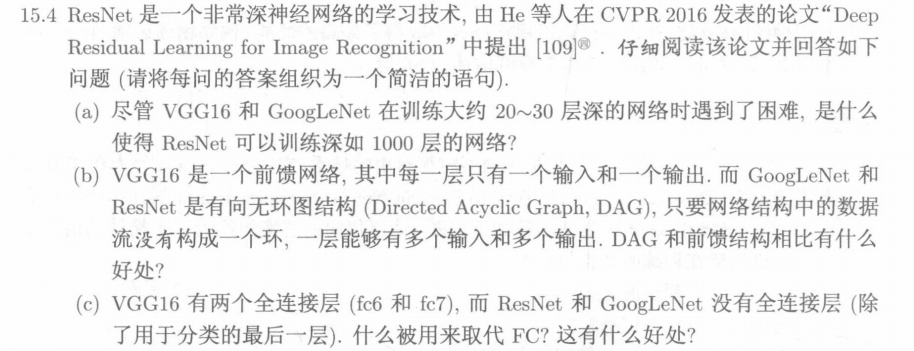
（d）(1) 可以使用更高的学习率。

　　 (2) 移除或使用较低的Dropout。

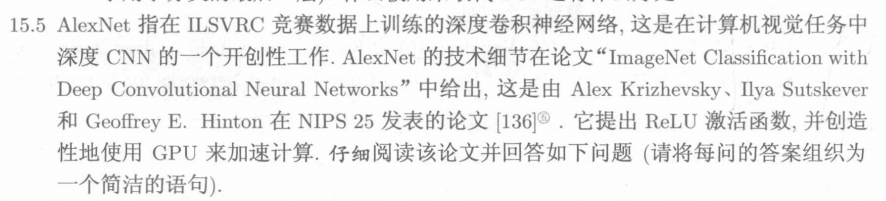
　　 (3) 降低L2权重衰减系数。

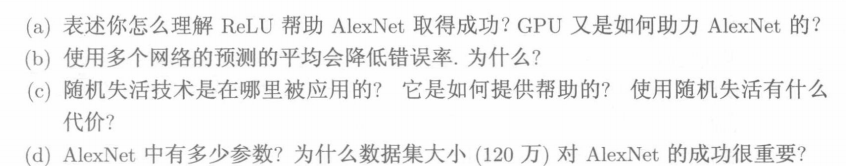
　　 (4) 取消Local Response Normalization层。

　　 (5) 减少图像扭曲的使用。



1. 让神经网络某些层跳过下一层神经元的连接，弱化每层之间的强联系。
2. 能更快的向前传播数据或向后传播梯度。
3. 使用了全局平均池化层，可以减少大量参数，提高效率。





（a）成功使用ReLU作为CNN的激活函数，并验证其效果在较深的网络超过了Sigmoid，成功解决了Sigmoid。

阵运算

（b）避免模型发生过拟合。

（c）在训练时使用。能够防止过拟合。

（d）约6000万，因为数据量越多，越能避免过拟合的发生。