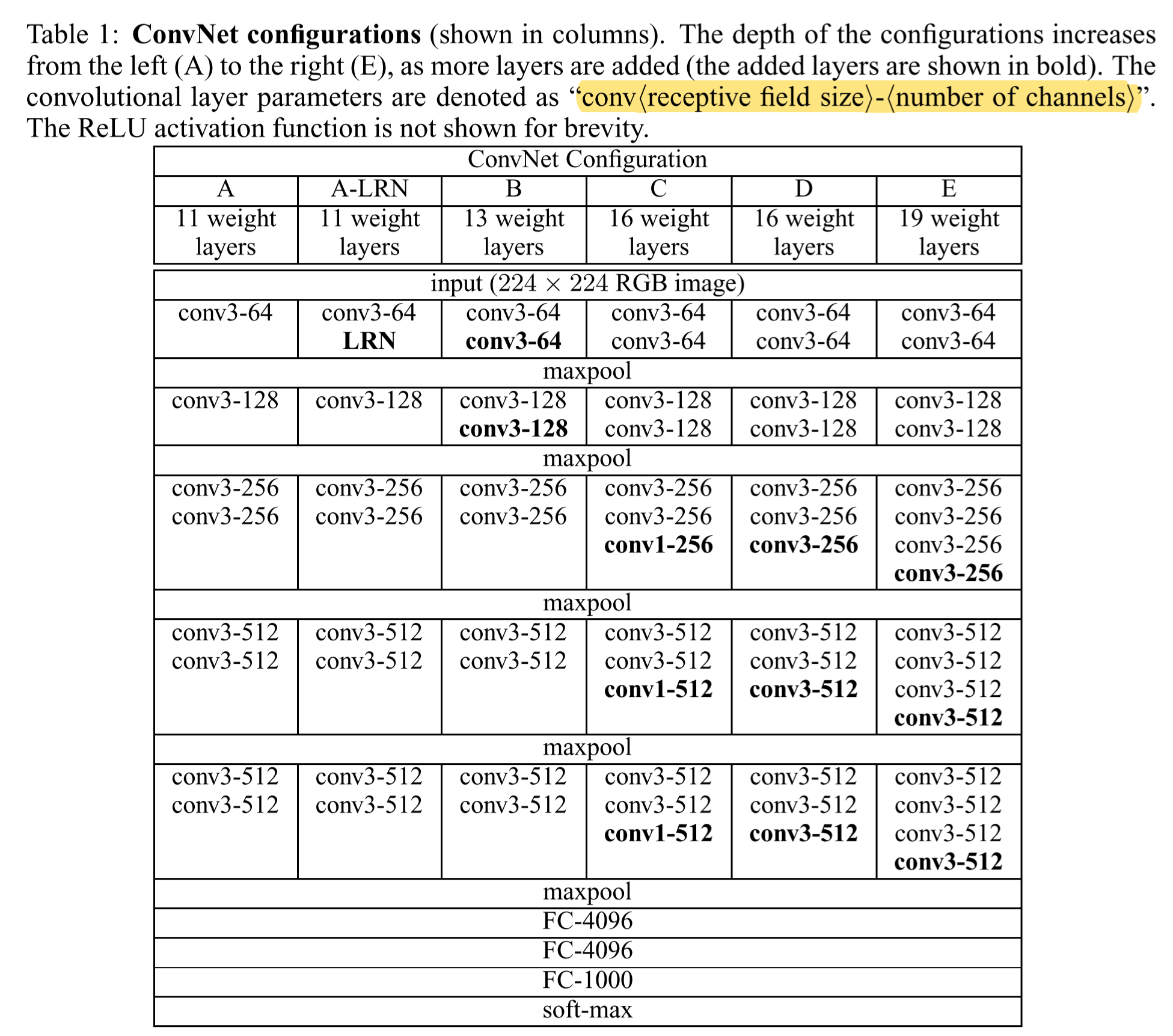
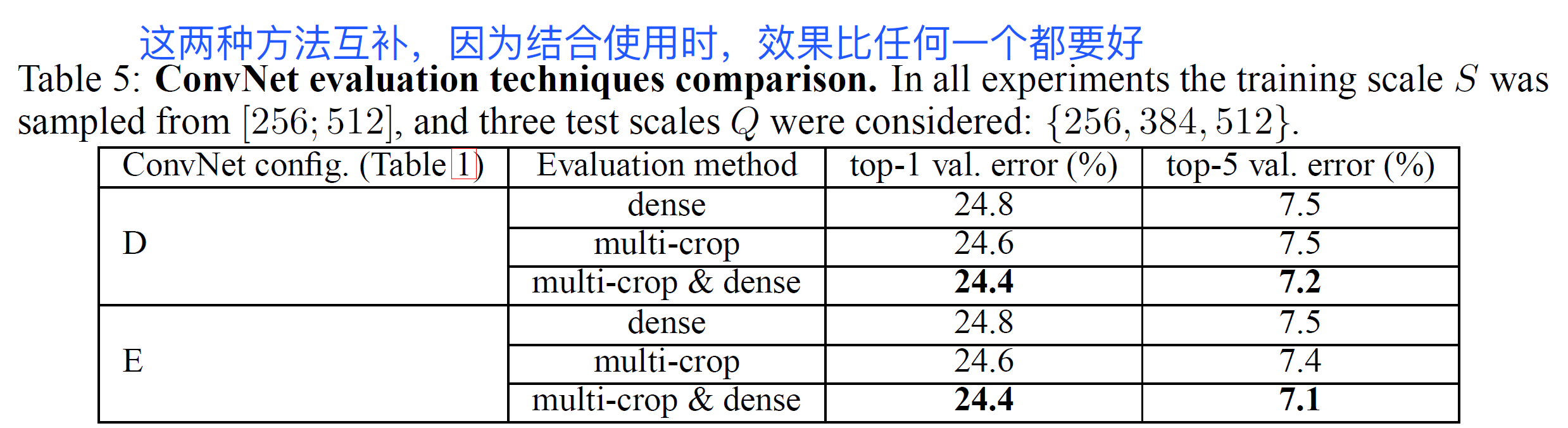
VGGNet Notes

1. **Architecture**
   1. 这篇论文主要关注CNN的深度对分类效果的影响。
   2. 训练图片的输入：224\*224 RGB图片，数据预处理只用了减去RGB平均值。
   3. 网络中负责提取特征的卷积核都是3\*3大小的，这个大小的卷积核是能捕捉到上下左右概念的最小卷积核。Stride=1，Padding=1，这样可以使卷积操作之后空间分辨率不变。Max Pooling 2\*2，stride=2。
   4. 局部响应归一化没什么用。
2. **Network Configuration**



* 1. 两个3\*3的卷积核堆叠，感受野和一个5\*5卷积核相同，但是增加了一个非线性函数，三个3\*3卷积核堆叠相当于一个7\*7卷积核。
  2. 使用较小的卷积核，可以显著降低参数量，如：假设输入、输出通道数量都是C，使用三个3\*3卷积核的参数量是3\*32­\*C2（输入层C个通道，输出层C个特征图），而使用一个7\*7卷积核的参数量是72C2。
  3. 这里效仿NIN，使用1\*1的卷积核增加非线性。

1. **训练过程**
   1. batch size 256，momentum 0.9，L­2 weight decay 5\*10-4，前两个全连接层dropout 0.5
   2. 训练图片的获取。令**S**等于原始图片resize后较短边的长度，训练所用图片是从resize的图片中裁剪224\*224大小的图片。有两种设置S的方法，第一种：固定S，论文中使用了S=256、384两种尺寸分别训练了模型；第二种：随机从[Smin, Smax]中采样S，Smin=256，Smax=512。
2. **测试过程**
   1. 这个要参考OverFeat这篇论文，作者说的dense什么意思没懂。
3. **实验**
   1. 两种测试方法：dense、multi-crop



* 1. dense这个方法出自OverFeat，先把图片resize成Q，然后应用FCN获得最后的feature map，Q不必等于S。multi-crop出自GoogLeNet，看到再说