Week 4 Report

Weitian Li weitian.li@rutgers.edu

1. Summarized optimization problems and methods for quantization neural networks
完成情况：总结目前发现的问题和相应的方法。

问题：神经网络的计算和运行需要大量的计算和存贮空间，部署到移动设备上还有点距离。
1. 模型参数数量多。
2. 计算比较复杂。（ 乘法卷积运算 ）

现在大多数的网络加速通过压缩或者加速运算。
1. Song Han：神经网络剪枝。
先训练一个全链接网络，模型的误差收敛到一定程度的时候，用一个阈值判别函数判断权重对网络的重要性，低于阈值直接弃掉不用该权重，实现一个剪枝的效果。然后重新训练网络到收敛的程度，直到网络参数变成一个高度稀疏的矩阵，不断调参增加压缩率，增大阈值压缩网络。文章还设计了的一个基于准确率损失和压缩率上升的公式，最后这是一个稀疏的参数矩阵，文章提出哈夫曼编码来存贮参数矩阵。

2. m 值化网络。
二值化网络，XNOR-NET，
问题：精度为 32 位单精度浮点数，计算过程比较耗费内存。
二值化网络采取二值化权重和用异或来代替正常的计算，并且对卷积核进行一个二值的操作，大幅减少内存的使用，尽量使精度贴近原来的网络。目前还有三值化网络的操作，三值化为+1，0，-1，通过先验阈值来三值。三值化被认为这种分布更符合一种正态分布或者均匀分布的组合，甚至还有五值化网络，大概原理和二值化的网络差不多。
问题：acc 问题比较严重。

3. 设计结构化矩阵
通过结构化矩阵，用少于 mxn 哥参数来描述 mxn 阶矩阵。
问题：找一个合适的结构矩阵很困难，acc 问题。

4. 低秩分解和稀疏性
通过使用压缩卷积层的典型低秩方法，使用低阶滤波器加速卷积层的处理。“按照这个方向，Lebedev 提出了核张量的典型多项式（CP）分解，使用非线性最小二乘法来计算。Tai 提出了一种新的从头开始训练低秩约束 CNN 的低秩张量分解算法。它使用批量标准化（BN）来转换内部隐藏单元的激活。一般来说， CP 和 BN 分解方案都可以用来从头开始训练 CNN。

低秩方法很适合模型压缩和加速，但是低秩方法的实现并不容易，因为它涉及计算成本高昂的分解操作。另一个问题是目前的方法都是逐层执行低秩近似，无法执行全局参数压缩，因为不同的层具备不同的信息。最后，分解需要大量的重新训练来达到收敛。”

5. 知识蒸馏（Hinton 的 2014 的论文，最近出了一篇在线蒸馏的文章）

通过使用 softmax 的概率分布以及 hard label 来保持高精度的压缩网络。通过实现高 T 的教师大型网络的训练，把教师大型网络学习到的 softmax 概率分布输入到学生小型网络来预测教师大型网络的输出，学生的 T 为正常 T。
问题：适用于 softmax 的网络任务，模型假设非常严格。

http://yanjoy.win
总体问题：1. 需要较好的原模型，在复杂任务不好实现。
2. 现在基本都是在比较完善的 CNN 网络上实现，很少别的网络。
3. 有些优化需要不断的训练调试，需要人类的先验知识，这些时间也需要考虑在内。
4. 二值化等剪枝方法加速效果非常不错，但是 acc 问题也很严重。

2. Read code of HWGQ
完成情况：大概理解数学的公式和回传过程，代码理解中。核心思想是从激活函数上下手，通过在前传时对 ReLU 函数进行分段，8Bit（0～255），这个分段对应是依靠半波高斯量化器对这些激活行为进行近似估计，根据高斯分布的情况来优化量化器，在误差回传的时候，因为采取的分段化的激活函数，所以梯度回传的时候会为 0，所以采用 ReLU 来进行回传，论文中采取了好几种回传的 ReLU，Vanilla ReLU，Clipped ReLU，Log-trailed ReLU 来比较，但是效果都差不太多。