摘要

这个问题可以被视为一个过参数化（）网络的网络压缩问题。论文使用了朴素的贝叶斯学习算法通过对架构参数建模。训练了一轮过度参数化的网络，然后更新他的架构。同时可以用来压缩CNN

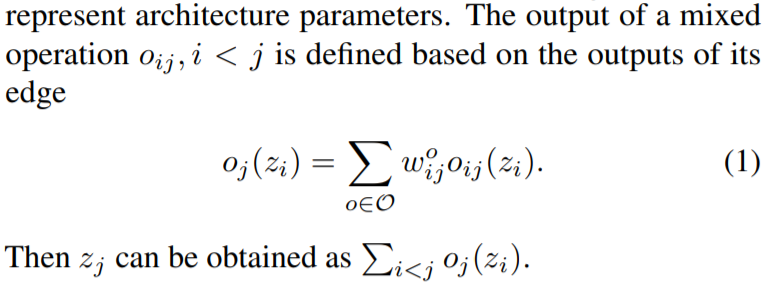
动机

1.节点对和其前驱和后继之间的依赖关系经常被忽略，这会导致对0操作的不当处理。

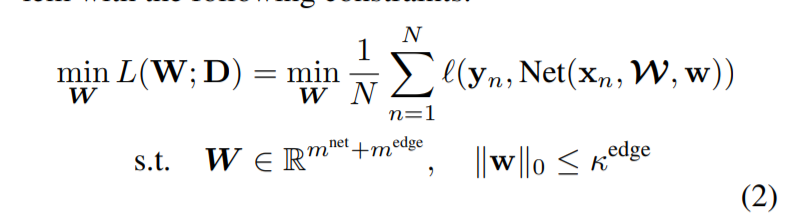
2.架构参数的修改基于他们的规模是不可靠的。

搜索空间设计

搜索空间可以被视为一个DAG图（）。假设eij是一个从点i到点j的一条边，操作oij（）是和eij相关的。在所有的操作上添加了缩放标量来控制信息的流动。设其为woij代表了架构的参数。

（）

问题形式化的定义



Net（xn，W(斜)，w） = output

L(yn, output) 距离差（范数）

(ΣL)/N 方差

目标最小化方差

减小w的方式可以提升设计灵活性。为了减小基于依赖性和量级的影响，对于每一个woij我们设置了一个开关soij。首先开-关功能不是独自得被其大小决定的。第二，独立性是被考虑在内的。前驱的节点有高于后继节点的控制。第三，soij是一个辅助的变量它并不会被反向传播所修改，但是会被直接用于修改边的开关。

定理1

j到k在o'的操作下有信息流通过，当且仅当j的至少有一个前驱的一个操作是非0的。