实验记录

实验名称：激活函数对比实验

实验目的：对比不同激活函数在同一网络同一数据集上的性能差异

实验数据集：

数据集名称：cifar100，图片大小32\*32，包含100个类别，每个类别包含600张图片

数据集划分：训练图片50000， 测试图片10000

预处理方式：无

实验配置：

学习率：5e-4

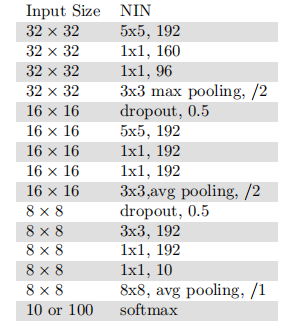
学习率衰减策略：每100轮缩小一倍

损失函数：CrossEntropy

训练轮数：400

Dropout：0.1

网络结构：



实验内容：

方案描述：在cifar100数据集上使用上述网络及网络配置测试不同的激活函数在此环境 下的性能效果。

可行性分析：新的损失函数增加了对不同程度激活神经元的不同梯度关注，使得低激活 度的神经元在训练的进程中能够对梯度做出更大的贡献，主导训练向着更 加正确的方向进行。

参考文献：Empirical Evaluation of Rectifified Activations in Convolution Network

实验结果：

训练时长：每个实验大约需要4个小时，新激活函数大约需要10小时，新激活函数训练耗时主要来自于未进行并行优化

测试时长：未单独统计，在训练过程中同步进行测试，既每轮训练结束后都会进行测试

# 训练集评测指标：样本分类准确率（Accuracy， Acc）

测试集评测指标：样本分类准确率

可视化：

一共进行了十三组对比实验，其中包含了5种已有的激活函数以及新提出激活函数的6中不同配置。（1：Relu,，2：LRelu，3：PRelu， 4: RRelu， 5: BReLU， -1：使用区间（-7，-3，-1，0，1，-3，-7）， -3：使用区间（-3，-1，0），-7：使用区间（-7， -3，-1，0），-15：使用区间（-15，-7，-3，0）以上三类初值均为1.5， -7\_0.25：使用区间（-7， -3， -1， 0）初值为0.25，-7\_0.5：使用区间（-7， -3， -1， 0）初值为0.5，-7\_1.0：使用区间（-7， -3， -1， 0）初值为1.0，-7\_2.0：使用区间（-7， -3， -1， 0）初值为2.0 ）

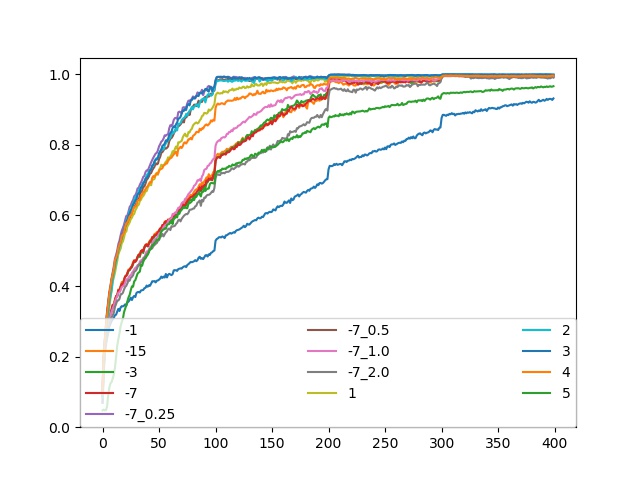


图 1 Train Acc

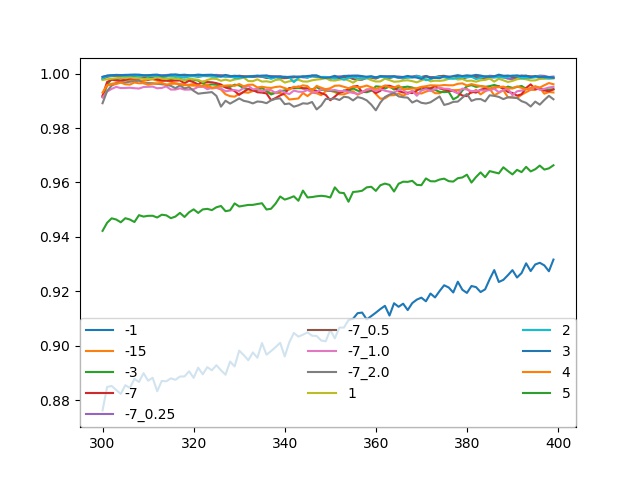


图 2 Train Acc Tail

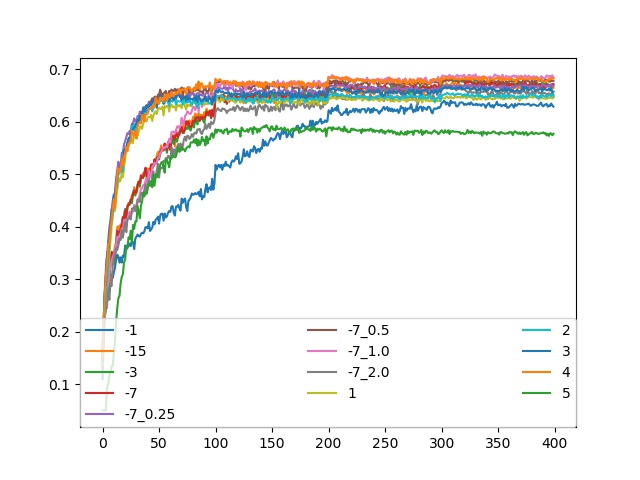


图 3 Val Acc

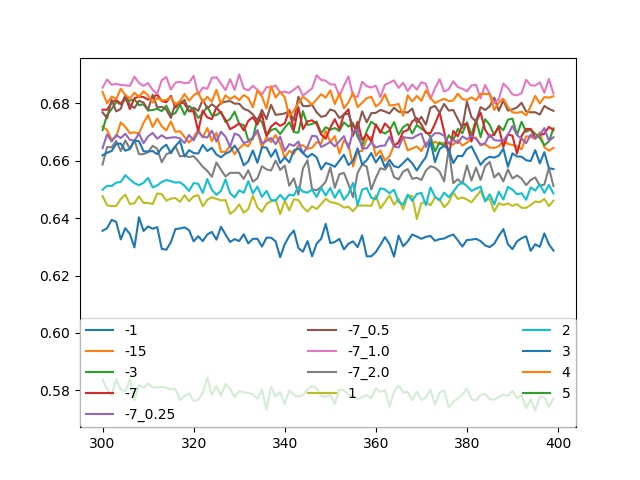


图 4 Val Acc Tail

结果分析：从上图中可以看出在训练集上不同的激活函数均达到0.99左右的准确率，但在验证集上新提出的激活函数在-7\_1.0的配置下达到了最优的性能