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：imagenet，图片大小224\*224，包含100个类别

数据集划分：采用数据集自带的划分方案

预处理方式：将输入图片先等长宽比缩放到256\*256，然后随机裁剪到224\*224

实验配置：

学习率：0.1

学习率衰减策略：无

损失函数：CrossEntropy

训练轮数：60

Dropout：无

网络结构：ResNet18

实验内容：

方案描述：在imagenet数据子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12小时，新激活函数大约需要12小时

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：

一共进行了十四组对比实验，其中包含了八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， 6：ELU，7：SELU，8：Swish，-1：使用区间（-7，-3，-1，0，1，3，7）， -2：使用区间（-7，-3，-1，0），-3：使用区间（-3，-2，-1，0），-4：使用区间（-15，-7，-3，0），-11：使用区间（-15，-7，-3，0，3，7，15），-12：使用区间（-3，-2，-1，0，1，2，3）以上六类初值均为(0, 2)上的均匀分布。

表 1 在ResNet18网络上的分类准确度

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ResNet (%) | ReLU | LReLU | PReLU | RReLU | BReLU | ELU | SELU | Swish | GReLU | MReLU |
| CIFAR-10 | 88.19 | 87.86 | 88.32 | 88.22 | 88.12 | 87.35 | 88.03 | 88.48 | 88.42 | **88.56** |
| CIFAR-100 | 72.22 | 72.71 | 73.11 | 72.63 | 72.31 | 72.75 | 73.52 | 72.01 | 72.58 | **73.54** |
| ImageNet  (TOP 1) | 83.8 | 83.025 | 82.625 | 82.925 | 82.65 | 83.875 | 83.875 | 83.275 | 83.65 | **83.95** |

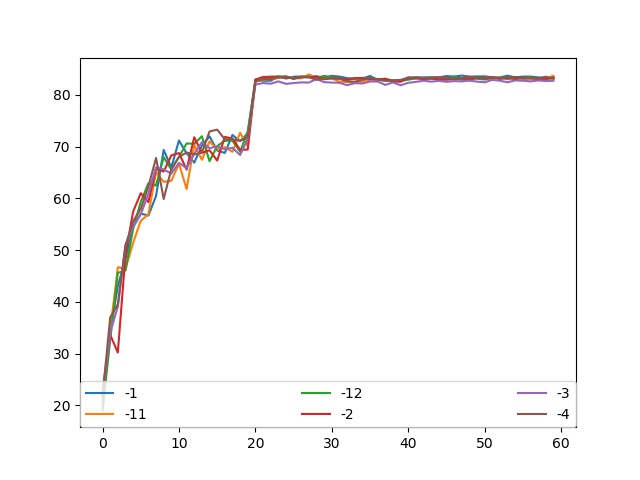


图 1 Val TOP1 ACC

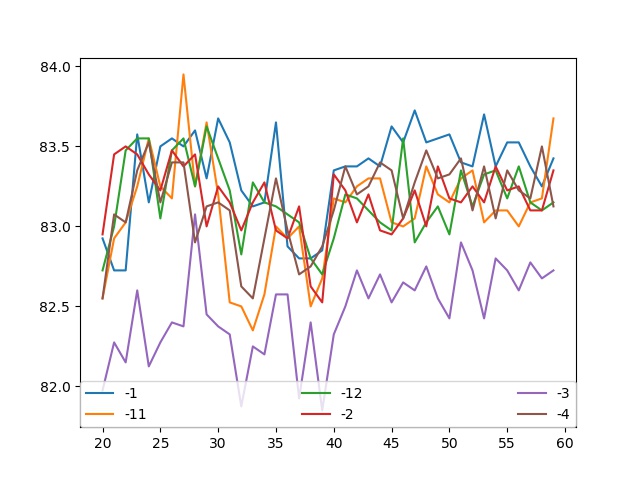


图 2 Val TOP 1 ACC Tail

结果分析：结合表1和图2所示，新提出的激活函数在-11的配置下取得了最优的性能，在验证集上分类准确率达到83.95%。