

ResNet-18 与 Transformer

王逸群 19307110397

2022.5

GitHub repo 链接: <https://github.com/quniLcs/cv-final>

网盘链接:

1 数据集

本项目使用 CIFAR-100 数据集, 其中包含 60000 张 32×32 的彩色图片, 其中训练集 50000 张, 测试集 10000 张, 被平均分为 100 类。

2 网络结构

2.1 ResNet

本项目使用的第一种网络结构是 ResNet-18, 其中激活函数为 ReLU, 最大的特征为残差连接。后者包括两种单元结构如图 1 和图 2所示。

对于输入的图像, 先进行步长为 1 的 $3 \times 64 \times 3 \times 3$ 卷积操作, 并进行批归一化和激活, 维度变为 $64 \times 32 \times 32$; 接着通过两次第一种单元结构, 维度不变; 再通过第二种单元结构, 维度变为 $128 \times 16 \times 16$; 再通过第一种单元结构, 维度不变; 再通过第二种单元结构, 维度变为 $256 \times 8 \times 8$; 再通过第一种单元结构, 维度不变; 再通过第二种单元结构, 维度变为 $512 \times 4 \times 4$; 再通过第一种单元结构, 维度不变; 最后通过全连接得到输出。

2.2 Transformer

3 超参数设置

数据增强: 裁剪、水平翻转、CutOut;

参数初始化: MSRA;

学习率: 由 0.1 开始每 10 个回合阶梯下降一个数量级;

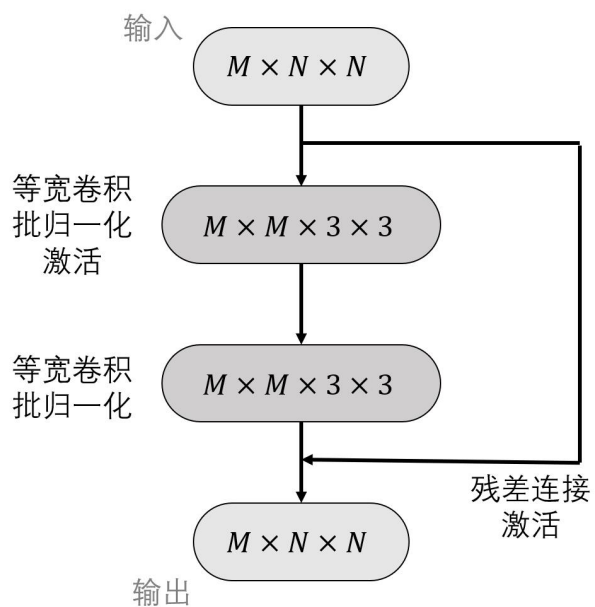


图 1: 残差连接第一种单元结构

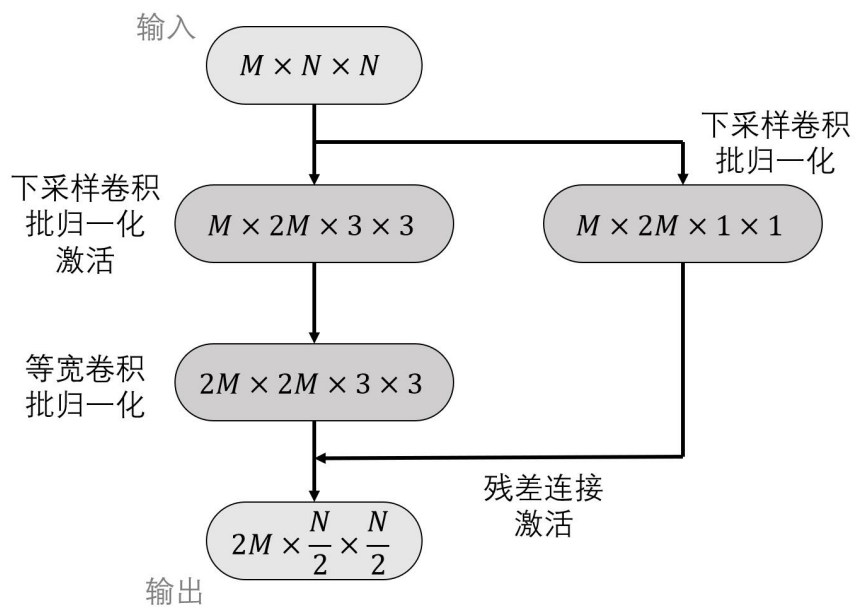


图 2: 残差连接第二种单元结构

优化器：带有 0.9 动量的随机梯度下降算法；

正则化参数：0.0005；

回合数：40；

批量大小：128；

每回合循环数：391；

总循环数： $40 \times 391 = 15640$ ；

损失函数：交叉熵损失函数；

评价指标：精确度。

4 实验结果