**第一次大作业**

1. **实验目的**

1）在Cifar10上完成图片分类任务

2）探究模型效果的影响因素并改进

1. **实验原理**
2. CNN在图片识别中的作用

在图片分类的问题中，最为自然的想法是计算两张图片之间的距离，由最近的1个（或者k个, knn算法）来决定该图片的类别。这种做法从相似度的角度来看有一定道理，但是问题在于这种计算忽略掉了图片的空间结构，所以正确率偏低(个人测试最高50%左右)。

那么CNN就是解决了这个问题，通过设置卷积核，让卷积核在原来的二维图片上滑动，计算每个像素周围的加权值，这种操作更好的保持了原来图片的空间结构。而且随着卷积层的增多，越靠后的卷积核的感受野越大，越能学习到整张图片的深层的语义信息，效果更好。

1. ResNet的优势

对于CNN来说，增加越多的层数，应该会学习到越多的特征，效果应该越好，但是实验中发现，更深的CNN难以训练，常常无法收敛。一方面我们可以增加BN层等，来避免梯度消失或爆炸，加速收敛。另一方面就是改进网络结构，本次实验就是采用了KaimingHe等提出的ResNet的网络结构[[1]](#footnote-1)，来进一步提升模型效果。

ResNet的出发点是，更深层次的神经网络应该不会比更浅的网络效果差，因为后面的深层可以全部设置为恒同映射。所以我们可以为网络增加一个恒同映射的连接，对于网络而言，学习一个零映射比恒同映射更为简单，所以训练模型应该更加容易收敛。借助这个想法，ResNet在网络上增加一些恒同映射，并成功训练了152层的ResNet，效果优异。本实验将尝试搭建ResNet，并在其基础上训练cifar10。

1. **实验步骤（数据预处理，核心模型代码，模型训练，超参设置）**
2. **实验结果**
3. **模型效果评估**
4. **分析影响因素**
5. **实验总结**

1. https://arxiv.org/abs/1512.03385 [↑](#footnote-ref-1)