

CNN 图片分类任务实验报告

58121124 张博彦

2023 年 5 月 27 日

1 CNN 模型原理

卷积神经网络是一种广泛应用于计算机视觉和图像处理领域的深度学习模型。CNN 模型通过在图像上应用卷积操作来提取图像特征，并通过多层卷积和池化层的堆叠来实现高级特征的学习和表示。CNN 模型主要包括卷积层、激活函数、池化层、全连接层这几个关键部分。CNN 模型通过反向传播算法进行训练，以最小化损失函数，并通过梯度下降法更新模型参数。在训练过程中，CNN 模型可以自动学习到图像中的特征表示，从而具备了在分类等任务中出色的性能。

2 CIFAR-10 数据集的内容和数据预处理

本实验使用 10% 的 CIFAR-10 数据集，共包含 10 个类别的 6000 张图片。这里使用 `torchvision.datasets` 中的 `ImageFolder` 函数来处理数据集，将数据集处理为 `tensor` 的形式，并将像素值归一化到 `[0, 1]` 范围内。

同时由于在测试集上样本 `feature` 与 `label` 并不对应，故采取从官网上下载原文件，之后解压两次成为能直接被 `pickle.load` 函数处理的文件，之后将图片与 `label` 一一对应进行储存。此时处理过后的数据便能够被 `ImageFolder` 函数直接处理，并且此时样本的 `feature` 与 `label` 相互匹配。

3 CNN 模型构建

1. 输入层：32x32 彩色图像
2. 卷积层 1：输出通道数 32，过滤器大小 3x3，激活函数：ReLU

3. 卷积层 2: 输出通道数 64, 过滤器大小 3x3, 激活函数: ReLU
4. 最大池化层 1: 池化大小 2x2
5. 卷积层 3: 输出通道数 128, 过滤器大小 3x3, 激活函数: ReLU
6. 卷积层 4: 输出通道数 128, 过滤器大小 3x3, 激活函数: ReLU
7. 最大池化层 2: 池化大小 2x2
8. 卷积层 5: 输出通道数 256, 过滤器大小 3x3, 激活函数: ReLU
9. 卷积层 6: 输出通道数 256, 过滤器大小 3x3, 激活函数: ReLU
10. 最大池化层 3: 池化大小 2x2
11. 展平输出
12. 全连接层 1: 1024 个神经元, 激活函数: ReLU
13. 全连接层 2: 512 个神经元, 激活函数: ReLU
14. 输出层: 10 个神经元, 激活函数: Softmax
15. 损失函数: 分类交叉熵

4 实验设置

1. 运行环境: GPU
2. 参数设置: 训练批次大小设置为 64, 训练周期设置为 18, 学习率设置为 0.001
3. 评估方法: 根据训练后模型在训练集、验证集上的损失和在验证集上的准确率进行评估

5 实验结果

1. 训练过程中验证集上的准确率和训练轮数的关系如图 1
2. 训练过程中训练集和验证集上的分类交叉熵损失和训练轮数的关系如图 2

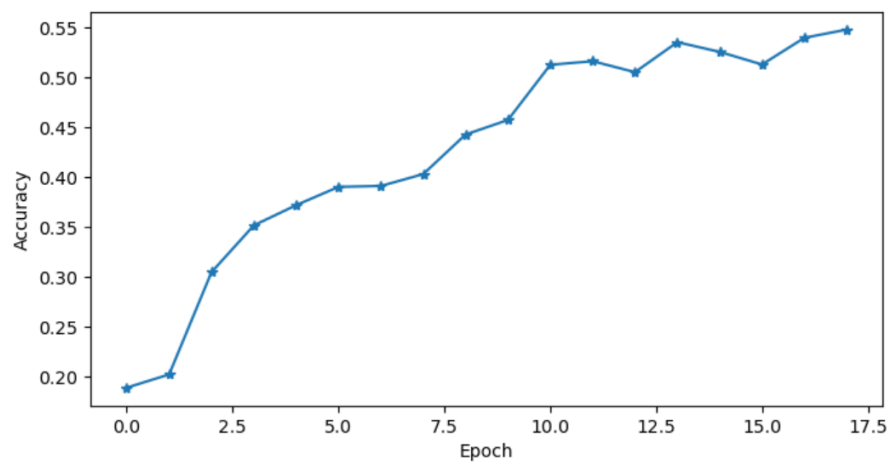


图 1: 验证集上的准确率

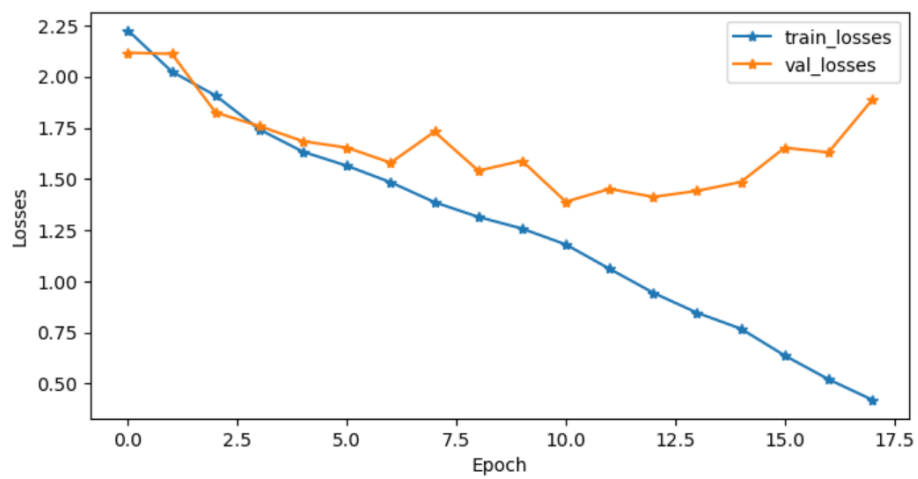


图 2: 训练集和验证集上的分类交叉熵损失

6 参数分析实验

6.1 参数设置

1. 批次大小 (batch size): 一般默认设置为 64, 故不做改动
2. 训练周期 (epoch): 当训练周期低于 16 时, 会出现欠拟合情况导致, 在训练集和测试集上的性能都较差。当训练周期高于 20 时, 会出现过拟合情况, 导致在训练集上有较好的性能, 但在验证集上表现较差。故选择训练周期为 18, 使得模型能够较好的拟合
3. 学习率 (learning rate): 当学习率低于 0.001 时, 会使模型收敛过慢, 导致在 18 个轮次之后仍旧无法达到局部最小处, 而当学习率高于 0.005 时, 会使模型不收敛, 导致模型无论如何训练都无法达到局部最小值。故选择学习率为 0.001, 使得模型能够以较快且合适的速度收敛至局部最小处附近

6.2 模型性能评估

该训练出的模型在测试集上的准确率约在 0.52, 该值并不算高, 但经过多次参数调整, 发现准确率并无显著提高。故可能需要优化图像的处理方式以及神经网络的构建才能够提高该模型的性能。