

Lab5 实验报告

姓名：陈翎玺 学号：523030910039 班级：电院 2302

1 实验概览

本次实验学习了简单的机器学习和深度学习知识。通过 pytorch 构建了自己的网络实现图像分类和图像检索。

机器学习的本质是找到一个函数，使得对给定的输入有一个令人满意的输出。

由于我们求解的函数通常有着多维输入和输出，求解这个函数的过程通常采用神经网络方法。

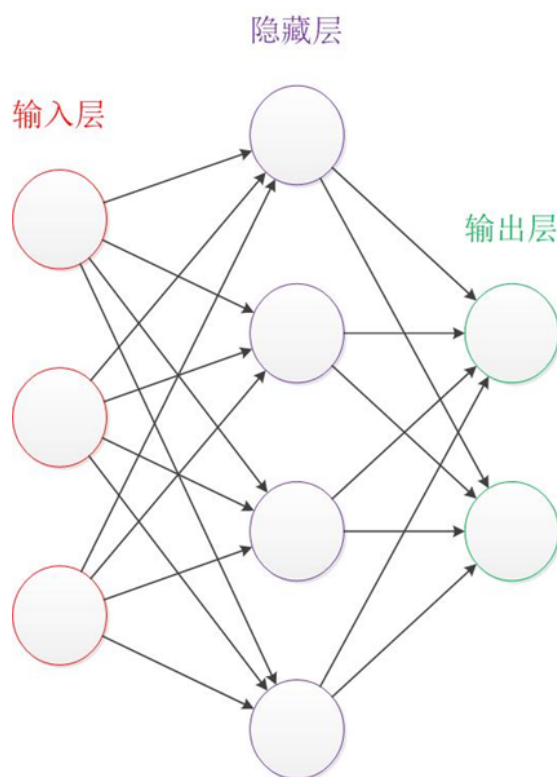


图 1: 神经网络

这是一个包含三个层次的神经网络。红色的是输入层，绿色的是输出层，紫色的是中间层（也叫隐藏层）。输入层有 3 个输入单元，隐藏层有 4 个单元，输出层有 2 个单元。后文中，我们统一使用这种颜色来表达神经网络的结构。

通常情况下，输入和输出的单元是固定的，而隐藏层的单元数可以手动调节。求解神经网络的过程就是求解层与层之间单元的函数关系，进而拟合出总体的输入和输出的关系。可以视为一种加权求和。

当然，可以有多个隐藏层。现代神经网络通常有几十层或者更多，对于这种多层神经网络的训练，可以认为是深度学习。

神经网络的训练通常包括前向传播（得到输出结果），计算损失（输出结果和期望结果的误差），反向传播（根据误差，通过梯度下降等方法改变网络中间的权值，以期得到更好的结果），更新权重等过程。

2 练习题的解决思路

2.1 1.CIFAR-10 图片分类

exp2.py 中的 test 函数需要我们完善。注意到这里是 test 部分，我们不需要更新权重，只需要通过前向传播得到输出结果即可。通过统计每个 batch 中匹配的数量和 batch 的大小即可计算正确率，代码如下：

```
test_loss = 0
correct = 0
total = 0
for batch_idx, (inputs, targets) in enumerate(testloader):
    outputs = model(inputs)
    loss = criterion(outputs, targets)
    test_loss += loss.item()
    _, predicted = outputs.max(1)
    total += targets.size(0)
    correct += predicted.eq(targets).sum().item()
acc = 100. * correct / total
```

图 2: 代码片段 1

对于自动训练（即前五个 epoch 的 lr 设置为 0.1，后五个设置为 0.01）只需要在第六个 epoch 训练开始前将 lr 更新，并修改 optimizer 中的 lr 信息即可。参考如下：

对于可视化部分，主要使用了 tensorboard 库自带的训练可视化功能，通过将训练过程中的 loss 和 acc 等数据写入指定的文件夹，再通过 tensorboard

```

for epoch in range(start_epoch, end_epoch + 1):
    if epoch >= 5:
        lr = 0.01
        for param_group in optimizer.param_groups:
            param_group['lr'] = lr
    train(epoch)
    test(epoch)

```

图 3: 代码片段 1

`-logdir=dir` (`dir` 为指定的目录) 在本机生成对应的 web 网页，将数据可视化。

2.2 2. 图像特征提取

类似于 lab4 中的 LSH 操作，我们需要从图像中提取特征信息。只不过与 LSH 操作不同的是，我们将整个图像的每个像素点都直接作为输入，通过一系列操作，得到远多于 LSH 操作的特征信息。

本次实验仍采用了推荐的 ResNet50 作为神经网络的模型，包含一系列的卷积、降采样、池化等操作，可以将 $1 \times 3 \times 224 \times 224$ 的输入映射到 $1 \times 2048 \times 1 \times 1$ 的特征向量上。具体细节不过多赘述。

将特征向量 `resize` 成 `size` 为 2048 的向量，再通过除以向量的模长 (L2-norm) 进行归一化，即可通过简单的点积操作计算向量之间的相似性。

实验中直接采取了 Pytorch 自带的 ResNet50 预训练数据，并未在数据集上进一步训练，故效果明显偏差。`dataset` 和 `target` 分别从网络上随机选取的十类物体的各 5/1 张图片组成，`target` 中图像的第一位代表组号，与 `dataset` 中的连续五张图片相对应。

代码见 `extract_feature`

3 代码运行结果及实验结果分析

3.1 1.CIFAR-10 图片分类

训练结果如下图，来源于 tensorboard 的自动生成结果。浅色线为实际值，深色线为平滑结果。

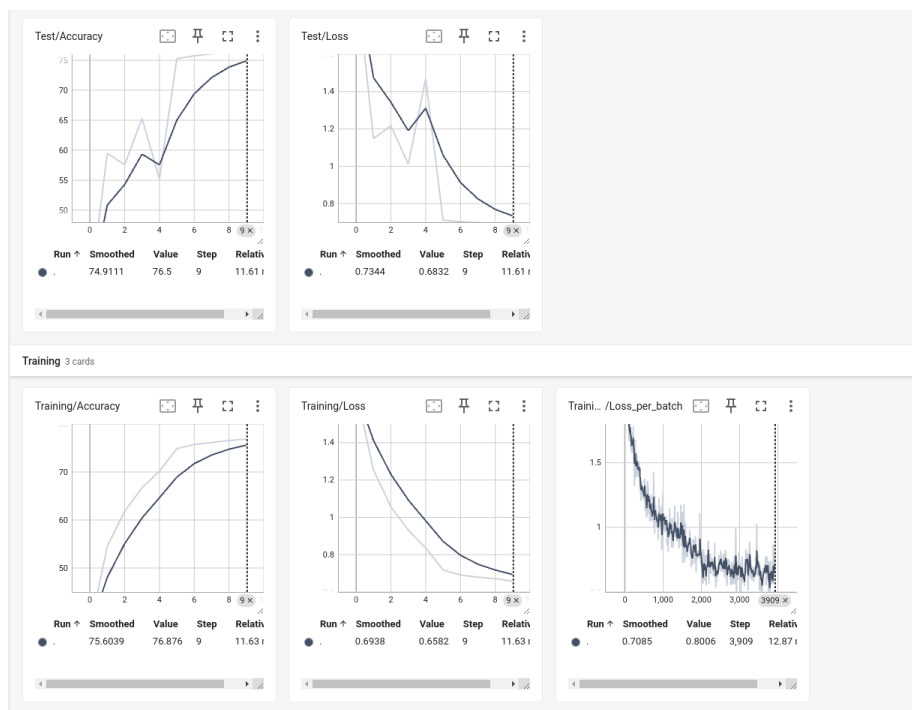


图 4: 训练结果

思考题解答:

Q1: Train acc 和 Test acc 有什么关联和不同?

A: Train acc 是指模型在训练集上的训练正确率, Test acc 是指模型在测试集上的正确率。通常情况下测试正确率与训练正确率正相关。但当训练次数过多时,可能出现模型在训练集上出现过拟合的现象,即在训练集上表现良好,测试集上表现一般。一旦更换同类型的数据,表现会变差很多,即缺乏可迁移性。

Q2: 在 lr 从 0.1 变到 0.01 后, acc 发生了什么变化? 为什么?

A: 训练集上的 acc 波动变小。对于较大的 lr, 单次对参数的改变过大, 可能出现较理想值略大变为较理想值远小的情况, 进而导致随训练次数增加, 准确率反而下降的情况 (如左上图 1-2, 3-4)

3.2 2. 图像特征提取

输出日志如下:

```

1 Extract features!
2 Time for extracting features: 2.65
3 The matching result for image 0panda.png
4 Top 1 similar image : 2.jpeg, with the similarity
   99.61307644844055%
5 Top 2 similar image : 4.jpeg, with the similarity
   99.58519339561462%
6 Top 3 similar image : 5.jpeg, with the similarity
   99.58409070968628%
7 Top 4 similar image : 3.jpeg, with the similarity
   99.57888126373291%
8 Top 5 similar image : 34.jpeg, with the similarity
   99.56896305084229%
9
10 The matching result for image 1plane.jpeg
11 Top 1 similar image : 30.jpeg, with the similarity
   99.59495067596436%
12 Top 2 similar image : 8.jpeg, with the similarity
   99.5837926864624%
13 Top 3 similar image : 10.jpeg, with the similarity
   99.5837926864624%
14 Top 4 similar image : 27.jpeg, with the similarity
   99.58122968673706%
15 Top 5 similar image : 35.jpeg, with the similarity
   99.57005977630615%
16
17 The matching result for image 2lion.jpeg
18 Top 1 similar image : 5.jpeg, with the similarity
   99.56948757171631%
19 Top 2 similar image : 13.jpeg, with the similarity
   99.56185817718506%
20 Top 3 similar image : 35.jpeg, with the similarity
   99.56108331680298%

```

21 Top 4 similar image : 50.jpeg, with the similarity
99.55136775970459%

22 Top 5 similar image : 33.jpeg, with the similarity
99.54777359962463%

23

24 The matching result for image 3car.jpeg

25 Top 1 similar image : 42.jpeg, with the similarity
99.59943294525146%

26 Top 2 similar image : 18.jpeg, with the similarity
99.57766532897949%

27 Top 3 similar image : 17.jpeg, with the similarity
99.57575798034668%

28 Top 4 similar image : 16.jpeg, with the similarity
99.57316517829895%

29 Top 5 similar image : 44.jpeg, with the similarity
99.56287145614624%

30

31 The matching result for image 4train.jpeg

32 Top 1 similar image : 18.jpeg, with the similarity
99.57295060157776%

33 Top 2 similar image : 24.jpeg, with the similarity
99.5633602142334%

34 Top 3 similar image : 35.jpeg, with the similarity
99.5570182800293%

35 Top 4 similar image : 42.jpeg, with the similarity
99.55503940582275%

36 Top 5 similar image : 43.jpeg, with the similarity
99.55331683158875%

37

38 The matching result for image 5ship.jpeg

39 Top 1 similar image : 8.jpeg, with the similarity
99.56227540969849%

40 Top 2 similar image : 10.jpeg, with the similarity

99.56227540969849%

41 Top 3 similar image : 34.jpeg, with the similarity
99.56071376800537%

42 Top 4 similar image : 31.jpeg, with the similarity
99.56071376800537%

43 Top 5 similar image : 30.jpeg, with the similarity
99.55897331237793%

44

45 The matching result for image 6chrysanthemum.jpeg

46 Top 1 similar image : 34.jpeg, with the similarity
99.62146282196045%

47 Top 2 similar image : 31.jpeg, with the similarity
99.62146282196045%

48 Top 3 similar image : 32.jpeg, with the similarity
99.59045648574829%

49 Top 4 similar image : 30.jpeg, with the similarity
99.5827853679657%

50 Top 5 similar image : 35.jpeg, with the similarity
99.58176612854004%

51

52 The matching result for image 7city.jpeg

53 Top 1 similar image : 39.jpeg, with the similarity
99.6126651763916%

54 Top 2 similar image : 36.jpeg, with the similarity
99.59376454353333%

55 Top 3 similar image : 32.jpeg, with the similarity
99.59142208099365%

56 Top 4 similar image : 33.jpeg, with the similarity
99.56973791122437%

57 Top 5 similar image : 30.jpeg, with the similarity
99.56700205802917%

58

59 The matching result for image 8camera.jpeg

```

60 Top 1 similar image : 41.jpeg, with the similarity
    99.61915612220764%
61 Top 2 similar image : 44.jpeg, with the similarity
    99.58404302597046%
62 Top 3 similar image : 45.jpeg, with the similarity
    99.56462383270264%
63 Top 4 similar image : 43.jpeg, with the similarity
    99.56437349319458%
64 Top 5 similar image : 42.jpeg, with the similarity
    99.56368207931519%
65
66 The matching result for image 9violin.jpeg
67 Top 1 similar image : 47.jpeg, with the similarity
    99.64690208435059%
68 Top 2 similar image : 48.jpeg, with the similarity
    99.63594675064087%
69 Top 3 similar image : 50.jpeg, with the similarity
    99.59607124328613%
70 Top 4 similar image : 49.jpeg, with the similarity
    99.58329796791077%
71 Top 5 similar image : 46.jpeg, with the similarity
    99.57795143127441%
72
73 Total Accuracy is 56.0%

```

这里简单认为如果识别到同类别图片则识别正确，否则错误，得到识别正确率为 56%。仅使用默认训练参数的情况下，虽然 56% 的正确率并不高，但足以接受这个结果。为进一步提升匹配正确率，应模仿 CIFAR-10 构建相应的训练和测试数据进行预训练，再进行匹配。考虑数据由手动从网络上下载，数据量不足，不便于实现相关实验。

通过特征向量的点积结果作为相似度的百分数（和得分），并以此为排序依据，得到每张图和最相似的五张图的结果。观察可以发现，照相机（8camera）和小提琴（9violin）组中，全部五张同类图像被找到，而在狮子（2lion）、火车（4train）和轮船（5ship）组中，仅有一种同类图像被找到。

其可能原因在于 8、9 两组图像特征明显且位姿接近，而 2、4、5 三组存在特征模糊等情况。下面以 8 组和 2 组及匹配认为最相近的图像为例分析可能的原因（按日志顺序从左至右、从上至下给出，第一张为目标图像）。

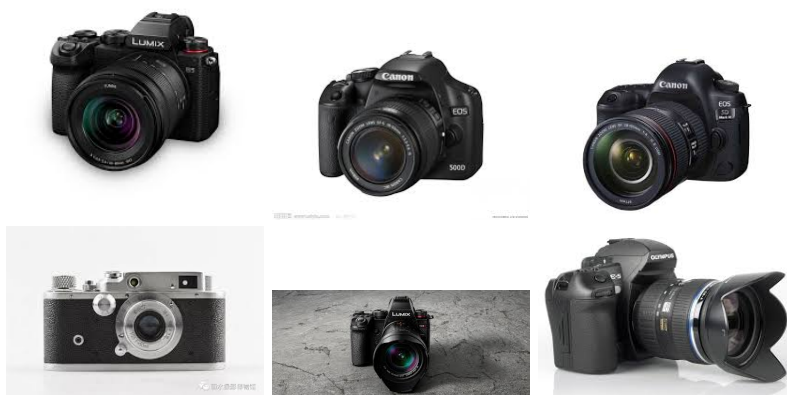


图 5: 第 8 组结果

第 8 组中，相机整体由黑色构成，背景以白色居多。在给定数据集中几乎没有其他类有如此单调的图像构成。其识别正确率高可以理解。

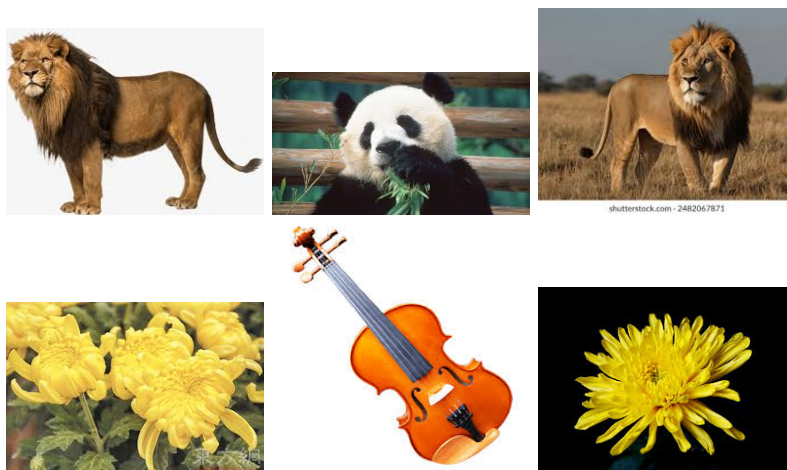


图 6: 第 2 组结果

第二组中，熊猫背后的栏杆颜色和狮子毛发颜色接近，且有相当的白色构成；菊花和小提琴的整体色调也接近狮子，可能导致模型的误判。

4 实验感想

本次实验学习了神经网络检测图像特征的方法，相较于传统方法准确率高且代码实现容易。不断调整参数和网络结构也有助于我们对神经网络不同部分的作用有更深入的理解。