

数字图像处理 PA_2

——实验报告

闫力敏*

1 实验要求

根据提供的 50 类图片的 10-shot，实现一个较为精确的图像分类功能。

2 实验过程

我们小组采用机器学习的方法来实现分类，采用的网络框架为 Alexnet，这一经典的分类网络框架，我们在此基础上实现了 4 个实验，实验 1 基于 TensorFlow，但是没有采用 finetune，实验 2 基于 TensorFlow，采用 finetune，实验 3 基于 keras，没有采用 finetune。实验 4 基于 keras，没有采用 finetune，但是将图像转为灰度图。

每个实验的具体过程分为：训练数据处理，管理输入数据，设计并实现项目框架，统计训练实验结果。

3 实验工作

我的具体工作是设计并实现项目框架，分别实现了基于 TensorFlow 以及 keras 的框架，其他人负责完成每个框架下的训练数据处理，管理输入数据，以及统计训练实验结果。主要完成的实验是基于 TensorFlow，采用 finetune

4 程序清单

main.cpp 基于 TensorFlow 采用 finetune。run.py 基于 keras 实现。

5 运行结果

为了较好的体现各个实验的公平性，我们采用的所有参数都是统一的，我们把训练数据在训练前，将他们的预测结果准确率作为准确率，下文统一称：训练准确率。

我 keras 采用的参数为：

*清华大学计算机系，学号：2015011391，邮编：100084

```
{
    loss='categorical_crossentropy',
    optimizer='sgd',
    batch_size=125,
    epochs=300,
    verbose=1,
    'num_classes' = 50
}
```

我们最开始实现的是基于 keras 实现的 Alexnet（因为简单），在此框架下我们并没有采用 finetune（.npy 貌似仅支持 TensorFlow），在此基础上我们所得到的训练准确率只有达到 5%，当时以为分类有 50 个，和 2% 差异不大。

我们 TensorFlow 采用的参数为：

```
{
    'learning_rate' = 0.03
    'num_epochs' = 300
    'batch_size' = 125
    'display_step' = 4;
    'dropout' = 0.5
    'num_classes' = 50
}
```

随后我们决定采用 TensorFlow，在此框架下我们并没有采用 finetune（方便于 keras 对比），结果发现准确率一直持续在 4% 左右，可以说结果相当差。

我们猜想可能是数据处理没有采用归一化，我们在 keras，和 TensorFlow 下同时实现了数据的归一化处理，发现在 keras 下训练准确率迅速爬升到了 10% 20%，但是 TensorFlow 下的训练准确率依然持续在 4% 左右，我们猜想可能是 TensorFlow 的学习速率是定值，而可能 keras 内部学习速率取得比较适合，所以导致了训练结果的巨大差异。

但是在与助教的交流中我们发现，TensorFlow 下采用 finetune，的准确率可以达到 40% 所以我们将 finetune fc7 以前的网络值固定只训练最后一层全连接。结果出乎意料的好，我们训练准确率非常容易的达到了 80%，在 epochs=200 左右的时候就有 98%，显然不应该有这么高，应为我们的训练数据毕竟有限，我们将预测的结果取出来，并比对了一部分结果，发现在测试集上的结果并不是非常精确，比如：



图 1:

我们看到这两张图片都被分到了 CD，但是第一张毕竟有一个保龄球和 CD 一样都是圆形，而且这张图既有保龄球又有保龄瓶，说实在的我也不知道给那个类（50 类中包括保龄球和保龄瓶）。

当然还有其他错例，出现这样的原因很大程度在于训练集过小以至于非常容易出现过拟合现象，通过加入一个预测集应该会有所提高。

6 实验数据处理

我们的做法是将一张图片旋转 90 度 4 次，翻转后再旋转 4 次得到 8 张图片进行训练，但是根据实验的真实结果，loss 值稍微会偏小，但是对预测准确率的影响没有想象中的大。

关于实验 4，因为在 keras 下非常容易实现，只需要将网络的输入图像通道设置为 1 即可，我们发

现最后得出的训练准确率可以到达 30%，同比 5%，以及 10%-20% 要高出不少，猜想可能是一些图像虽有边缘但是图像的不同通道体现的效果不是很好，因此在灰度图的意义下边缘部分就保存的比较好。因为人在分辨物体的时候更主要的还是看轮廓而不是颜色。当然在 TensorFlow 下由于网络结构固定，无法实现。

7 体会展望

此次试验，虽然比较简单，但是我们明显的看出了采用 finetune 的训练效果，短时间内就可以达到较好的训练结果。最后感谢助教提供实验基础。