

实验 1 读取并显示 VOC 数据集数据

实验时长：4 小时

实验难度：一般

实验摘要：MNIST 数据集.idx3-ubyte 这种使用二进制字节的形式存储的，无法直接进行读取显示，需要通过编写程序来打开查看。

实验建议：了解二进制文件的读取方法。

实验目标：能够通过编写程序来读取并显示 MNIST 数据集数据。

1、读取并显示 voc 数据集数据

1.1、导入相关库

以训练集图片文件读取为例，代码如下：

```
import os

from matplotlib import pyplot as plt # plt 用于显示图片

import matplotlib.image as mpimg # mpimg 用于读取图片

import math
```

1.2、切换工作目录

读取文件数据在 PNGImages 目录下，切换到该目录下，代码如下所示：

```
ImgPath = '/home/data/voc2005/PNGImages/INRIA_graz-person-train'
file_list = os.listdir(ImgPath)
os.chdir(ImgPath)
```

1.3、查看图片文件

调用读取文件以及图像展示的函数，示例代码如下：

```
n = math.sqrt(len(file_list))  
plt.figure()  
for index,elemt in enumerate(file_list):  
    plt.subplot(n,n,index+1)  
    img = mpimg.imread(elemt)  
    plt.imshow(img)  
plt.show()
```

运行程序即可生成数据集的 16 张图片文件。



