

# 深度学习快速入门

01Pytorch 加载数据

POET

2024 年 2 月 11 日

## 1 torchvision 数据集介绍

- torchvision 中有很多数据集，当我们写代码时指定相应的数据集指定一些参数，它就可以自行下载。
- CIFAR-10 数据集包含 60000 张  $32 \times 32$  的彩色图片，一共 10 个类别，其中 50000 张训练图片，10000 张测试图片。

## 2 查看 CIFAR10 数据集内容

```
1 import torchvision
2 train_set = torchvision.datasets.CIFAR10(root="./
    dataset", train=True, download=True) # root为存放
    数据集的相对路线
3 test_set = torchvision.datasets.CIFAR10(root="./
    dataset", train=False, download=True) # train=
    True是训练集，train=False是测试集
4
5 print(test_set[0])          # 输出的3是target
6 print(test_set.classes)    # 测试数据集中有多少种
7
8 img, target = test_set[0] # 分别获得图片、target
9 print(img)
10 print(target)
11
12 print(test_set.classes[target]) # 3号target对应的
    种类
13 img.show()
```

ps: 如果下载过慢可以复制 python 控制台中的链接到迅雷下载，将文件复制粘贴到 root 目录下即可

### 3 Tensorboard 查看内容

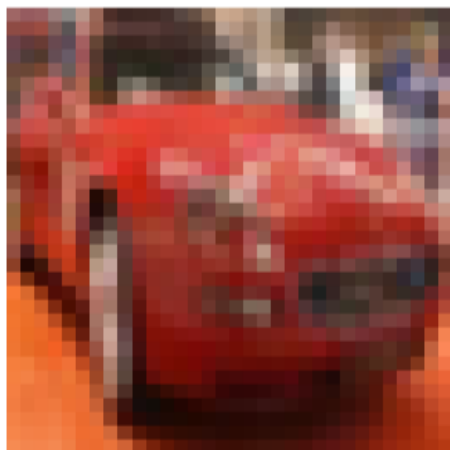
```
1 import torchvision
2 from torch.utils.tensorboard import SummaryWriter
3
4 dataset_transform = torchvision.transforms.Compose
5     ([torchvision.transforms.ToTensor()])
6 train_set = torchvision.datasets.CIFAR10(root="./
7     dataset", train=True, transform=dataset_transform
8     , download=True) # 将ToTensor应用到数据集中的每
9     一张图片，每一张图片转为Tensor数据类型
10 test_set = torchvision.datasets.CIFAR10(root="./
11     dataset", train=False, transform=
12     dataset_transform, download=True)
13
14 writer = SummaryWriter("logs")
15 for i in range(11):
16     img, target = test_set[i]
17     writer.add_image("test_set", img, i)
18     print(img.size())
19
20 writer.close() # 一定要把读写关闭，否则显示不出来
21     图片
```

---

test\_set



Step 5



CSDN @Asa13opp

---