

Sanyuan

目录视图

摘要视图

RSS 订阅

个人资料



SanyuanLiu

关注

发私信

访问：1784次

积分：57

等级：BLOG > 1

排名：千里之外

原创：3篇

转载：0篇

译文：0篇

评论：1条

异步赠书：9月重磅新书升级，本本经典

程序员9月书讯

每周荐书：ES6、虚拟现实、物联网（评论送...）

Tensorflow 中Tfrecords的使用心得

标签：Tensorflow tfrecord 数据读取

2017-03-30 20:10

607人阅读

评论(0)

收藏

举报

分类：

Tensorflow学习心得

版权声明：本文为博主原创文章，未经博主允许不得转载。

这篇文章主要讲一下如何用Tensorflow中的标准数据读取方式简单的实现对自己数据的读取操作，主要分为以下两个步骤：(1)将自己的数据集转化为 xx.tfrecords的形式；(2):在自己的程序中读取并使用.tfrecords进行操作。

数据集转换：为了便于讲解，我们简单制作了一个数据，如下图所示：

文章搜索

文章分类

- 深度学习I (1)
- 深度学习理解 (0)
- Tensorflow学习心得 (1)
- 阅读笔记 (1)
- 环境安装 (1)

文章存档

- 2017年09月 (1)
- 2017年07月 (1)
- 2017年04月 (1)
- 2017年03月 (1)

阅读排行

- DualGAN: Unsupervised Dual L... (1098)
- Tensorflow 中Tfrecords的使用... (607)
- Ubuntu 下修改键盘排列 (50)
- caffe 绘制accracy曲线 IndexEr... (23)

评论排行

- caffe 绘制accracy曲线 IndexEr... (1)



程序：

```
[python]
01. import tensorflow as tf
02. import numpy as np
03. import os
04. from PIL import Image
05. def _int64_feature(value):
06.     return tf.train.Feature(int64_list=tf.train.Int64List(value=[value]))
07.
08. def _bytes_feature(value):
09.     return tf.train.Feature(bytes_list=tf.train.BytesList(value=[value]))
10.
11.
12. def img_to_tfrecord(data_path):
13.     rows = 256
14.     cols = 256
15.     depth = 3
16.     writer = tf.python_io.TFRecordWriter('test.tfrecords')
17.     labelfile=open("random.txt")
18.     lines=labelfile.readlines()
19.     for line in lines:
20.         #print line
21.         img_name = line.split(" ")[0]#name
```

关闭

- Tensorflow 中Tfrecords的使用... (0)
- DualGAN: Unsupervised Dual L... (0)
- Ubuntu下修改键盘排列 (0)

推荐文章

- * CSDN新版博客feed流内测用户征集令
- * Android检查更新下载安装
- * 动手打造史上最简单的 Recycleview 侧滑菜单
- * TCP网络通讯如何解决分包粘包问题
- * SDCC 2017之大数据技术实战线上峰会
- * 快速集成一个视频直播功能

最新评论

- caffe 绘制accracy曲线 IndexError: list inde...
- cdw_FstLst : 写的什么鬼，代码都没对齐

Tensorflow 中Tfrecords的使用心得 - Sanyuan - CSDN博客

```

22.         label = line.split(" ")[1]#label
23.         img_path = data_path+img_name
24.         img = Image.open(img_path)
25.         img = img.resize((rows,cols))
26.         #img_raw = img.tostring()
27.         img_raw = img.tobytes()
28.         example = tf.train.Example(features = tf.train.Features(feature = {
29.             'height': _int64_feature(rows),
30.             'weight': _int64_feature(cols),
31.             'depth': _int64_feature(depth),
32.             'image_raw': _bytes_feature(img_raw),
33.             'label': _bytes_feature(label)}))
34.
35.         writer.write(example.SerializeToString())
36.     writer.close()
37.
38.
39.
40. if __name__ == '__main__':
41.     current_dir = os.getcwd()
42.     data_path = current_dir + '/data/'
43.     #name = current_dir + '/data'
44.     print('Convert start')
45.     img_to_tfrecord(data_path)
46.     print('done!')
```

运行该段程序可以看到在dataset_tfrecord文件夹下面有test.tfrecord文件生成。

在TF的Session中调用这个生成的文件：

```

[python]
01. #encoding=utf-8
02. # 设置utf-8编码，方便在程序中加入中文注释。
03. import os
```

关闭

```

04. import scipy.misc
05. import tensorflow as tf
06. import numpy as np
07. from test import *
08. import matplotlib.pyplot as plt
09.
10. def read_and_decode(filename_queue):
11.
12.     reader = tf.TFRecordReader()
13.     _, serialized_example = reader.read(filename_queue)
14.
15.     features = tf.parse_single_example(serialized_example, features = {
16.         'image_raw':tf.FixedLenFeature([], tf.string)})
17.     image = tf.decode_raw(features['image_raw'], tf.uint8)
18.     image = tf.reshape(image, [OUTPUT_SIZE, OUTPUT_SIZE, 3])
19.     image = tf.cast(image, tf.float32)
20.     #image = image / 255.0
21.
22.     return image
23.
24. data_dir = '/home/sanyuan/dataset_animal/dataset_tfrecords/'
25.
26. filenames = [os.path.join(data_dir, 'train%d.tfrecords' % ii) for ii in range(1, 10)]
    果有多个文件，直接更改这里即可
27. filename_queue = tf.train.string_input_producer(filenames)
28. image = read_and_decode(filename_queue)
29. with tf.Session() as sess:
30.     coord = tf.train.Coordinator()
31.     threads = tf.train.start_queue_runners(coord=coord)
32.     for i in xrange(2):
33.         img = sess.run([image])
34.         print(img[0].shape) # 设置batch_size等于1 . 每次读出来只有一张图
35.         plt.imshow(img[0])
36.         plt.show()
37.     coord.request_stop()
38.     coord.join(threads)
39.

```

关闭

程序到这里就已经处理完成了，当然在decode的过程中也是可以进行一些预处理操作的，不过建议还是在制作数据集的时候进行，TFRecord使用的是队列的方式进行读取数据，这个对于多线程操作来说还是很方便的，只需要设置好格式，每次直接读取就可以了。

01. [python]

01. [python]

顶 踩
0 0

- [下一篇](#) [DualGAN: Unsupervised Dual Learning for Image-to-Image Translation](#) 阅读笔记

相关文章推荐

- TensorFlow学习（十一）：保存TFRecord文件
- TensorFlow高效读取数据——TFRecord
- Presto的服务治理与架构在京东的实践与应用--王哲...
- Retrofit 从入门封装到源码解析
- 学习笔记TF016:CNN实现、数据集、TFRecord、加...
- TensorFlow学习记录-- 7.TensorFlow高效读取数据...
- 深入掌握Kubernetes应用实践--王渊命
- 自然语言处理工具Word2Vec

- TFRecords 文件的生成和读取
- Python基础知识汇总
- TFRecord 的使用
- Android核心技术详解
- 如何使用tfrecord
- tensorflow实战 猫狗大战（三）制作tfrecords数据集...
- Tensorflow之构建自己的图片数据集TFrecords(精)
- Tensorflow 中Tfrecords的使用心得

查看评论

暂无评论

您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

关闭

[公司简介](#) | [招贤纳士](#) | [广告服务](#) | [联系方式](#) | [版权声明](#) | [法律顾问](#) | [问题报告](#) | [合作伙伴](#) | [论坛反馈](#)

网站客服

杂志客服

微博客服

webmaster@csdn.net

400-660-0108

| 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 | 江苏乐知网络技术有限公司

京 ICP 证 09002463 号 | Copyright © 1999-2017, CSDN.NET, All Rights Reserved

