

[CSDN首页 \(http://www.csdn.net?ref=toolbar\)](http://www.csdn.net?ref=toolbar)[学院 \(http://edu.csdn.net?ref=toolbar\)](http://edu.csdn.net?ref=toolbar)[下载 \(http://download.csdn.net?ref=toolbar\)](http://download.csdn.net?ref=toolbar)

更多 ▾

[下载 CSDN APP \(http://www.csdn.net/app?ref=toolbar\)](http://www.csdn.net/app?ref=toolbar)[写博客 \(http://write.blog.csdn.net/postedit?ref=toolbar\)](http://write.blog.csdn.net/postedit?ref=toolbar)[登录 \(https://passport.csdn.net/account/login?ref=toolbar\)](http://passport.csdn.net/account/login?ref=toolbar)[首页 \(http://blog.csdn.net/\)](http://blog.csdn.net/)[移动开发 \(http://blog.csdn.net/mobile/newarticle.html\)](http://blog.csdn.net/mobile/newarticle.html)

全部 □

**CSDN** (<http://www.csdn.net>)

## 目录 Tensorflow 训练自己的数据集（二）（TFRecord）



原创 2017年04月12日 20:31:57

3897

6

4

喜欢



Tensorflow 提供了一种统一的格式来存储数据，这个格式就是TFRecord,上一篇博客中所提到的方法当数据的来源更复杂，每个样例中的信息更丰富的时候就很难有效的记录输入数据中的信息了，于是Tensorflow提供了TFRecord来统一存储数据，接下来我们就来介绍如何使用TFRecord来同意输入数据的格式。



评论



## 1. TFRecord格式介绍

分享

TFRecord文件中的数据是通过tf.train.Example Protocol Buffer的格式存储的，下面是tf.train.Example的定义

ruiweicas (<http://blog.cs...>)

+ 关注

[http://blog.csdn.net/Best\\_Coder](http://blog.csdn.net/Best_Coder)

码云

原创

粉丝

喜欢

未开通

38

32

0

[https://github.com/Best\\_Coder/Best\\_Coder](https://github.com/Best_Coder/Best_Coder)

### 他的最新文章

[更多文章 \(http://blog.csdn.net/Best\\_Coder\)](http://blog.csdn.net/Best_Coder)

Tensorflow框架下Faster-RCNN实践  
（二）——用自己制作的数据训练  
Faster-RCNN网络（附代码）  
(/best\_coder/article/details/76615000)

Tensorflow框架下Faster-RCNN实践  
（一）——Faster-RCNN所需数据集制  
作（附代码）  
(/best\_coder/article/details/76577544)

window10 编译安装opencv3.2+contrib  
(/best\_coder/article/details/71515722)

编辑推荐

最热专栏

  
返回顶部

```

1  message Example {
2    Features features = 1;
3  };
4
5  message Features{
6    map<string,Feature> featrue = 1;
7  };
8
9  message Feature{
10   oneof kind{
11     BytesList bytes_list = 1;
12     FloatList float_list = 2;
13     Int64List int64_list = 3;
14   }
15 };

```



目录



喜欢



收藏



评论



分享

从上述代码可以看到，tf.train.Example 的数据结构相对简洁。tf.train.Example中包含了一个从属性名称到取值的字典，其中属性名称为一个字符串，属性的取值可以为字符串（BytesList），实数列表（FloatList）或整数列表（Int64List）。例如我们可以将解码前的图片作为字符串，图像对应的类别标号作为整数列表。

## 2. 将自己的数据转化为TFRecord格式

### 准备数据

在上一篇博客中，我们为了像伟大的MNIST致敬，所以选择图像的前缀来进行不同类别的分类依据，但是大多数的情况下，在进行分类任务的过程中，不同的类别都会放在不同的文件夹下，而且类别的个数往往浮动性又很大，所以针对这样的情况，我们现在利用不同类别在不同文件夹中的图像来生成TFRecord。

我们在Iris&Contact这个文件夹下有两个文件夹，分别为iris,contact。对于每个文件夹中存放的是对应的图片

Notes on tensorflow（七）将数据集转...

Tensorflow构建自己的图片数据集TFrec...

学习笔记TF016:CNN实现、数据集、T...

TensorFlow学习记录-- 7.TensorFlow高...

### 在线课程



【免费】深入理解Docker

([http://edu.csdn.net/course/detail/563?](http://edu.csdn.net/course/detail/563?utm_source=blog9)

内部原理及网络配置

([http://edu.csdn.net/course/detail/563?](http://edu.csdn.net/course/detail/563?utm_source=blog9)

实战线上峰会

([http://edu.csdn.net/course/detail/563?](http://edu.csdn.net/course/detail/563?utm_source=blog9)

SDCC 2017之区块链技术

([http://edu.csdn.net/course/series\\_detail/66?](http://edu.csdn.net/course/series_detail/66?utm_source=blog9)

实战线上峰会

([http://edu.csdn.net/course/series\\_detail/66?](http://edu.csdn.net/course/series_detail/66?utm_source=blog9)

实战线上峰会

([http://edu.csdn.net/course/series\\_detail/66?](http://edu.csdn.net/course/series_detail/66?utm_source=blog9)

实战线上峰会



返回顶部

## 转换数据

数据准备好以后，就开始准备生成TFRecord,具体代码如下：

```
1 import os
2 import tensorflow as tf
3 from PIL import Image
4 import matplotlib.pyplot as plt
5
6 cwd='/home/ruyiwei/Documents/Iris&Contact/'
7 classes={'iris','contact'}
8 writer= tf.python_io.TFRecordWriter("iris_contact.tfrecords")
9
10 for index,name in enumerate(classes):
11     class_path=cwd+name+'/'
12     for img_name in os.listdir(class_path):
13         img_path=class_path+img_name
14         img=Image.open(img_path)
15         img= img.resize((512,80))
16         img_raw=img.tobytes()
17         #plt.imshow(img) # if you want to check you image,please delete '#'
18         #plt.show()
19         example = tf.train.Example(features=tf.train.Features(feature={
20             "label": tf.train.Feature(int64_list=tf.train.Int64List(value=[index])),
21             'img_raw': tf.train.Feature(bytes_list=tf.train.BytesList(value=[img_raw]))
22         }))
23         writer.write(example.SerializeToString())
24
25 writer.close()
```

[目录](#)[喜欢](#)[收藏](#)[评论](#)[分享](#)

## 3. Tensorflow从TFRecord中读取数据

[返回顶部](#)

```
1 def read_and_decode(filename): # read iris_contact.tfrecords
2     filename_queue = tf.train.string_input_producer([filename])# create a queue
3
4     reader = tf.TFRecordReader()
5     _, serialized_example = reader.read(filename_queue)#return file_name and file
6     features = tf.parse_single_example(serialized_example,
7                                       features={
8       'label': tf.FixedLenFeature([], tf.int64),
9       'img_raw' : tf.FixedLenFeature([], tf.string),
10    })#return image and label
11
12    img = tf.decode_raw(features['img_raw'], tf.uint8)
13    img = tf.reshape(img, [512, 80, 3]) #reshape image to 512*80*3
14    img = tf.cast(img, tf.float32) * (1. / 255) - 0.5 #throw img tensor
15    label = tf.cast(features['label'], tf.int32) #throw label tensor
16    return img, label
```



目录



喜欢



收藏



评论



分享

## 4. 将TFRecord中的数据保存为图片



返回顶部



目录



喜欢



收藏



评论



分享

```
1 filename_queue = tf.train.string_input_producer(["iris_contact.tfrecords"])
2 reader = tf.TFRecordReader()
3 _, serialized_example = reader.read(filename_queue) #return file and file_name
4 features = tf.parse_single_example(serialized_example,
5                                   features={
6                                       'label': tf.FixedLenFeature([], tf.int64),
7                                       'img_raw' : tf.FixedLenFeature([], tf.string),
8                                   })
9 image = tf.decode_raw(features['img_raw'], tf.uint8)
10 image = tf.reshape(image, [512, 80, 3])
11 label = tf.cast(features['label'], tf.int32)
12 with tf.Session() as sess:
13     init_op = tf.initialize_all_variables()
14     sess.run(init_op)
15     coord=tf.train.Coordinator()
16     threads= tf.train.start_queue_runners(coord=coord)
17     for i in range(20):
18         example, l = sess.run([image,label])#take out image and label
19         img=Image.fromarray(example, 'RGB')
20         img.save(cwd+str(i)+'_'+"Label_"+str(l)+'.jpg')#save image
21         print(example, l)
22     coord.request_stop()
23     coord.join(threads)
```

版权声明：本文为博主原创文章，如果特别强烈的想转载，那就转载吧，有问题联系ruyiwei.cas@gmail.com

举报

标签：TFRecord (<http://so.csdn.net/so/search/s.do?q=TFRecord&t=blog>) /

tensorflow (<http://so.csdn.net/so/search/s.do?q=tensorflow&t=blog>) /

制作数据集 (<http://so.csdn.net/so/search/s.do?q=制作数据集&t=blog>) /



返回顶部



[youni215 \(/u011463646\)](#)

2017-06-21 21:46

6楼

博主你好，我在这里，img=Image.fromarray(example, &#39;RGB&#39;);，一直显示错误，显示Can not convert a ndarray into a Tensor or Operation.不知道是什么原因，求解答。



回复

喜欢



[u011463646 \(/u011463646\)](#) 2017-04-14 16:18

5楼

用tf.nn.sampled\_softmax\_ops获取标签，每次生成的标签会不一样，它会随机选取一个的？。。。。。



收藏

回复



评论



[u011463646 \(/u011463646\)](#) 2017-04-14 16:16

4楼

用tf.nn.sampled\_softmax\_ops获取标签，会出现每次生成的标签不一样，它会随机选取一个的。。。。？



回复

分享

查看 6 条热评



## 相关文章推荐

### Tensorflow中使用tfrecord方式读取数据 ([/u010358677/article/details/70544241](#))


前言本博客默认读者对神经网络与Tensorflow有一定了解，对其中的一些术语不再做具体解释。并且本博客主要以图片数据为例进行介绍，如有错误，敬请斧正。使用Tensorflow训练神经网络时，我们可以...



返回顶部

 u010358677 (<http://blog.csdn.net/u010358677>) 2017-04-23 20:11  4377

## Notes on tensorflow（七）将数据集转换为TFRecord (/weixin\_35653315/article/details/71015845)

 TFRecord是tensorflow使用的数据格式，类似于caffe的imdb，mxnet的recordio。使用框架定义的数据格式好处是有强大的框架支持，例如封装了数据解析、多线程等操作，使用...

目录


 weixin\_35653315 ([http://blog.csdn.net/weixin\\_35653315](http://blog.csdn.net/weixin_35653315)) 2017-04-30 13:46  2071

喜欢



收藏

## 精选：深入理解 Docker 内部原理及网络配置 ([http://edu.csdn.net/huiyiCourse/detail/563?utm\\_source=blog10](http://edu.csdn.net/huiyiCourse/detail/563?utm_source=blog10))



 多令人称道的功能，如 Docker 的镜像管理。然而，Docker的网络一直以来都比较薄弱，所以我们有必要深入了解Docker的网络知识，以满足更高的网络需求。

评论



## 分享Tensorflow构建自己的图片数据集TFrecords (/tengxing007/article/details/54428262)

: C/C++ 图像二进制存储与读取中，已经讲解了如何利用C/C++的方法存储与读取二进制图像文件，本文继续讲述如何根据CIFAR-10的格式制作自己的数据集。所述博文与代码均已同步至GitHub ...



 tengxing007 (<http://blog.csdn.net/tengxing007>) 2017-01-14 19:24  3684

## 学习笔记TF016:CNN实现、数据集、TFRecord、加载图像、模型、训练、调试 (/weixin\_38776853/article/details/72811405)


AlexNet(Alex Krizhevsky,ILSVRC2012冠军)适合做图像分类。层自左向右、自上向下读取，关联层分为一组，高度、宽度减小，深度增加。深度增加减少网络计算量。训练模型数据集 S...




返回顶部

 weixin\_38776853 ([http://blog.csdn.net/weixin\\_38776853](http://blog.csdn.net/weixin_38776853)) 2017-05-30 20:00  532

## TensorFlow学习记录-- 7.TensorFlow高效读取数据之tfrecord详细解读 (/qq\_16949707/article/details/53483493)


 why tfrecord?对于数据量较小而言,可能一般选择直接将数据加载进内存,然后再分batch输入网络进行训练 (tip:使用这种方法时,结合yield 使用更为简洁,大家自己尝试一下吧,我就不...

目录  qq\_16949707 ([http://blog.csdn.net/qq\\_16949707](http://blog.csdn.net/qq_16949707)) 2016-12-06 10:06  3686



喜欢

## Ubuntu启动问题以及Grub Rescue修复方法 (/xixinyan/article/details/6969469)

 Ubuntu启动问题以及Grub Rescue修复方法 问题：之前系统是Windows7 64bit（C盘）+ D，E盘（都是NTFS）+ Ubuntu。今天，在...

收藏



评论



 xixinyan (<http://blog.csdn.net/xixinyan>) 2011-11-14 17:22  0



## SqueezeNet (/u014568921/article/details/72565976)

分享

深度学习方法（七）：最新SqueezeNet 模型详解，CNN模型参数降低50倍，压缩461倍！


 u014568921 (<http://blog.csdn.net/u014568921>) 2017-05-19 21:00  239

## Tensorflow 训练自己的数据集（一）（数据直接导入到内存） (/best\_coder/article/details/70141075)

制作自己的训练集下图是我们数据的存放格式，在data目录下有验证集与测试集分别对应iris\_test，iris\_train 为了向伟大的MNIST致敬，我们采用的数据名称格式和MNIST类似 ...

  
返回顶部




 Best\_Coder ([http://blog.csdn.net/Best\\_Coder](http://blog.csdn.net/Best_Coder)) 2017-04-12 11:32 2377

## Tensorflow 训练自己的数据集（二）（TFRecord） (/best\_coder/article/details/70146441)

1. TFRecord输入数据格式2. TFRecord格式介绍3. 将自己的数据转化为TFRecord格式准备数据转换数据import os import tensorflow as tf fr...



 Best\_Coder ([http://blog.csdn.net/Best\\_Coder](http://blog.csdn.net/Best_Coder)) 2017-04-12 20:31 3898



## TensorFlow 制作自己的TFRecord数据集 (/bbzz2/article/details/60581006)

喜欢



官网的mnist和cifar10数据之后，笔者尝试着制作自己的数据集，并保存，读入，显示。TensorFlow可以支持cifar10的数据格式，也提供了标准的TFRecord 格式，而关于 ten...

收藏



 BBZZ2 (<http://blog.csdn.net/BBZZ2>) 2017-03-06 16:30 785

评论




## caffe中训练自己的数据集的步骤 (/wfei101/article/details/75142626)

caffe中训练自己的数据集的步骤

 wfei101 (<http://blog.csdn.net/wfei101>) 2017-07-15 07:55 406

## Fast RCNN训练自己的数据集 (/jxl\_1207/article/details/72900315)

转载请注明出处，楼赅(yi)航的blog，<http://www.cnblogs.com/louyihang-loves-baiyan/> [https://github.com/YihangLou/...](https://github.com/YihangLou/)

 JXL\_1207 ([http://blog.csdn.net/JXL\\_1207](http://blog.csdn.net/JXL_1207)) 2017-06-07 15:32 205

## Faster-RCNN训练自己的数据集 (/gvfdbdf/article/details/52214008)

就目前我在了解到的资料来看，训练RCNN系列，一般有两种思路：修改CNN的工程代码，适应自己的数据集格式 将自己的数据格式做成VOC2007形式的数据集 从工作量来看，无疑后者更容易一些，本文的思路...

 GVFDBDF (<http://blog.csdn.net/GVFDBDF>) 2016-08-15 19:34 5313


### darknet yolov2 训练自己的数据集 (/qq\_27840681/article/details/63682694)



ubuntu14.04+opencv2.4.10+cuda7.5+cudnn5.0 训练集：20张图片 00001.jpg-00020.jpg 测试集：5张图片 00021.jpg  
目录...



喜欢

 qq\_27840681 ([http://blog.csdn.net/qq\\_27840681](http://blog.csdn.net/qq_27840681)) 2017-03-19 10:53 629



### caffe训练自己的数据集 (/fahaiL/article/details/74935756)

收藏

利用python生成txt文件后，就可以利用caffe自带工具将图片转换为lmdb格式，其代码如下#!/usr/bin/env sh # Create the image and label data for caffe  
genet lmdb inputs...



评论


 fahaiL (<http://blog.csdn.net/fahaiL>) 2017-07-10 19:31 171



分享

### 用YOLOv2训练自己的数据集 (/zhy8623080/article/details/73188542)

一. 系统初始环境系统:Ubuntu16.04: ubuntu-16.04-desktop-amd64.iso cuda安装文件: cuda-repo-ubuntu1604-8-0-local\_8.0...

 zhy8623080 (<http://blog.csdn.net/zhy8623080>) 2017-06-13 17:08 602

### TensorFlow 训练 MNIST 数据（二） (/willduan1/article/details/52036831)

上一篇博客讲了一个简单的基于 SoftMax 回归的学习模型，准确率大概在91%左右，这篇构建一个深度卷积神经网络。主要的教程还是来自于极客学院，但是讲的很琐碎，我把自己整理的思路和最后写的完整的代码...



返回顶部



willduan1 (<http://blog.csdn.net/willduan1>) 2016-07-26 19:46 9098

## Faster RCNN 训练自己的数据集 (/apple\_\_fly/article/details/51712402)

Faster RCNN 训练自己的数据集



Apple\_\_fly ([http://blog.csdn.net/Apple\\_\\_fly](http://blog.csdn.net/Apple__fly)) 2016-06-19 14:55 1952

目录



喜欢



收藏



评论



分享



返回顶部