

[首页](#)[专栏](#)[专家](#)[热文](#)[thesby的专栏](#)

[原]Caffe学习之自建数据

2015-2-5 阅读534 评论0

平台: Ubuntu14.04LTS

在编译成功后的 `./build/tools/` 目录下有一个 `convert_imageset` 或者 `convert_imageset.bin`, 这个工具就是用来将自己的图片数据转换为Caffe所需要的数据格式。但是关于这个工具如何使用, 它的源程序中也语焉不详。好吧, 我摸索了一天, 网上查了各种资料, 都无果而终, 因为大家都说的是同一个不知道哪位先驱者给出的例子, 但是在我这个版本下已经完全无法使用, 因为里面没有设置图片的宽度和高度, 而默认的是0, 所以这样创建的数据库文件是无法使用的。正在绝望的时刻, 突然想到了Mnist的这个示例中不是有转换数据的脚本吗。对, 就是 `create_mnist.sh`。

```
#!/usr/bin/env sh
```

```
# This script converts the mnist data into lmdb/leveldb format,
```

```
# depending on the value assigned to $BACKEND.
```

```
EXAMPLE=examples/mnist
```

```
#定义路径
```

```
DATA=data/mnist
```

```
BUILD=build/examples/mnist
```

```
BACKEND="lmdb"echo "Creating ${BACKEND}..."
```

```
rm -rf $EXAMPLE/mnist_train_${BACKEND}
```

```
#如果生成过数据, 则删除后重新创建, 否则会创建失败
```

```
rm -rf $EXAMPLE/mnist_test_${BACKEND}
```

```
$BUILD/convert_mnist_data.bin $DATA/train-images-idx3-ubyte \
```

```
    $DATA/train-labels-idx1-ubyte $EXAMPLE/mnist_train_${BACKEND} --  
backend=${BACKEND}
```

```
#上面这句话意思是使用$BUILD路径下的convert_mnist_data.bin程序, 来创建数据库。第一个参数$DATA/train-
```

#images-idx3-ubyte是原始数据的路径，第二个参数就是你创建的**label**文件了，第三个参数是指定创建的数据库

#路径及文件名。似乎已经没问题了，然后我们看看这个程序的源代码，里面有一行说如果参数不是四个的话，会

#报错，看来后面的参数都属于第四个参数的了。（我是Linux的小白，如果你是大神，看到这样的推断别笑话。）

#后面的参数有很多，包括必填项 **--resize_width**和**--resize_height**。还有**--shuffle**表示是否随机读取，**--backend**表示

#保存的数据库类型

```
$BUILD/convert_mnist_data.bin $DATA/t10k-images-idx3-ubyte \  
$DATA/t10k-labels-idx1-ubyte $EXAMPLE/mnist_test_${BACKEND} --  
backend=${BACKEND}  
  
echo "Done."
```

哈哈，相信我写到这儿，加了那么多注释，应该没问题了吧。

然后就仔细讲讲第二个参数。我们自己写的时候就写成**txt**的文件，每一行第一个写文件名，第二个写类型标号。比如：

[上一篇](#)[下一篇](#)

请先登录后，再发表评论！

[查看评论](#)

[更多评论 \(0\)](#)

[回顶部](#)

[移动版](#) | [桌面版](#)

©1999-2012, CSDN.NET, All Rights Reserved