

resnet18尝试第一天

convert_fp16.py

test_onnx.py

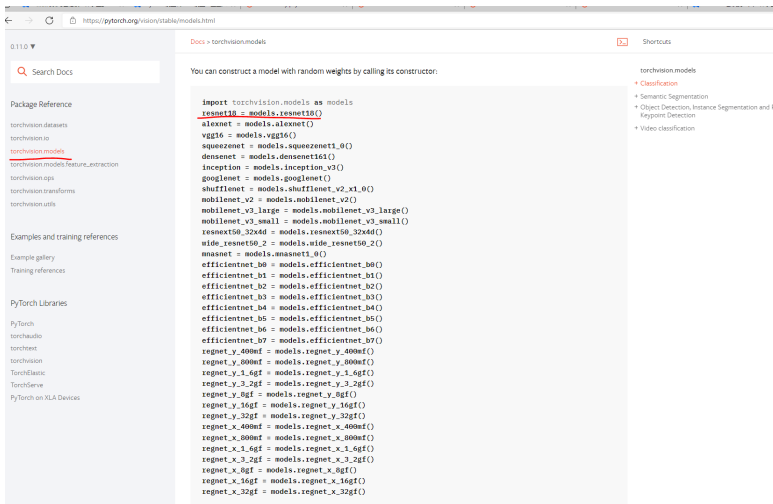
获得resnet18. onnx模型

1. 安装torch和torchvision,

在<https://pytorch.org/get-started/locally/>中安装自己torch, 例如:

```
conda install pytorch torchvision torchaudio cudatoolkit=10.2 -c pytorch
```

2. 在pytorch官网中找到你想要模型，一般路径在Docs/torchvision/models，然后看看有没有你想要的模型（这里我将尝试resnet18模型），如下图：



3. 开始下载resnet18模型

```
import torchvision
resnet18 = torchvision.models.resnet50(pretrained=True)
```

4. 安装onnx, 利用onnx把下载好的pytorch模型转换成onnx模型

```
pip install onnx      #onnx
```

下载模型pytorch模型并把模型转换成onnx模型

```
resnet18.onnx()
```

```
import torchvision
import torch
import torchvision.models as models
resnet18 = models.resnet18(pretrained=True).cuda() #pytorchgpu,cuda()

input_name = ['input']
output_name = ['output']
input = torch.randn(1, 3, 224, 224).cuda() #shapegpu,cuda()
onnx_resnet18=torch.onnx.export(resnet18, input, 'resnet18.onnx',
input_names=input_name, output_names=output_name, verbose=True) #pytorch
onnx.save(onnx_resnet18, 'resnet18.onnx')
```

5. check一下转换好的resnet18.onnx模型

```
import torchvision
import torch
import torchvision.models as models
import onnx
import onnxruntime

test = onnx.load('resnet18.onnx')
onnx.checker.check_model(test)
print("==> Passed")
# "=="> Passed"
```

注：如果你在check的时候报“RuntimeError: module compiled against API version 0xf but this version of numpy is 0xe”，这表示你的numpy版本不配对，把你的numpy卸载了然后重新安装就可以啦

验证resnet18.onnx精度

1. 环境要求：

```
onnx == 1.9.0
onnxruntime == 1.7.0
numpy == 1.19.5
torch == 1.7.1
torchvision == 0.8.2
onnxruntime-gpu == 1.4.0
heapq_max == 0.21
```

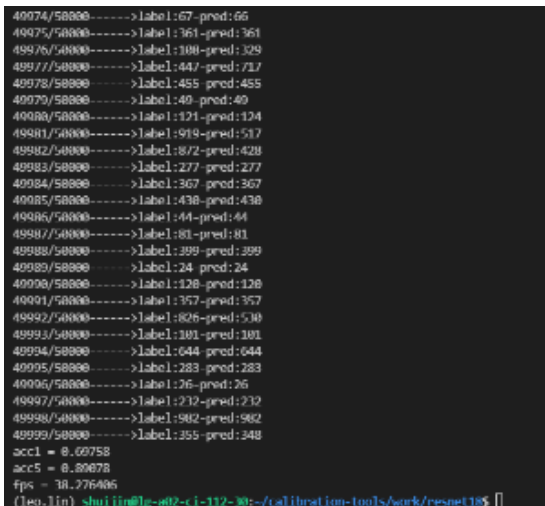
2. 准备数据集 (<http://image-net.org/challenges/LSVRC/2012/>)

```
mkdir data #
cd data
wget http://dl.caffe.berkeleyvision.org/caffe_ilsvrc12.tar.gz #
tar -zxvf caffe_ilsvrc12.tar.gz #
cp -r val.txt ../
```

3. 开始验证fp32精度

```
python test_onnx.py --model=resnet18.onnx --input_width=299 --
input_height=299
```

验证结果如下图：



```
49974/50000----->label:67-pred:66
49975/50000----->label:361-pred:361
49976/50000----->label:100-pred:129
49977/50000----->label:447-pred:717
49978/50000----->label:455-pred:455
49979/50000----->label:40-pred:40
49980/50000----->label:121-pred:124
49981/50000----->label:910-pred:517
49982/50000----->label:872-pred:428
49983/50000----->label:277-pred:277
49984/50000----->label:367-pred:367
49985/50000----->label:438-pred:438
49986/50000----->label:44-pred:44
49987/50000----->label:81-pred:81
49988/50000----->label:399-pred:399
49989/50000----->label:24-pred:24
49990/50000----->label:128-pred:128
49991/50000----->label:357-pred:357
49992/50000----->label:826-pred:510
49993/50000----->label:101-pred:101
49994/50000----->label:644-pred:644
49995/50000----->label:283-pred:283
49996/50000----->label:26-pred:26
49997/50000----->label:212-pred:212
49998/50000----->label:982-pred:982
49999/50000----->label:355-pred:348
acc1 = 0.68758
acc5 = 0.88878
fps = 38.77686
(160.1in) shuifan@ec2-54-132-30:/calibration-tools/caffe/resnet18$
```

4. 把fp32 onnx模型转成fp16 onnx模型

```
python convert_fp16.py --model ==resnet18
```

5. 验证fp16 onnx模型精度

```
python test_onnx.py --model=inception_v3_fp16.onnx --input_width=299 --
input_height=299 --test-fp16
```

验证结果如下图：