看图说话机器人项目

项目组成员

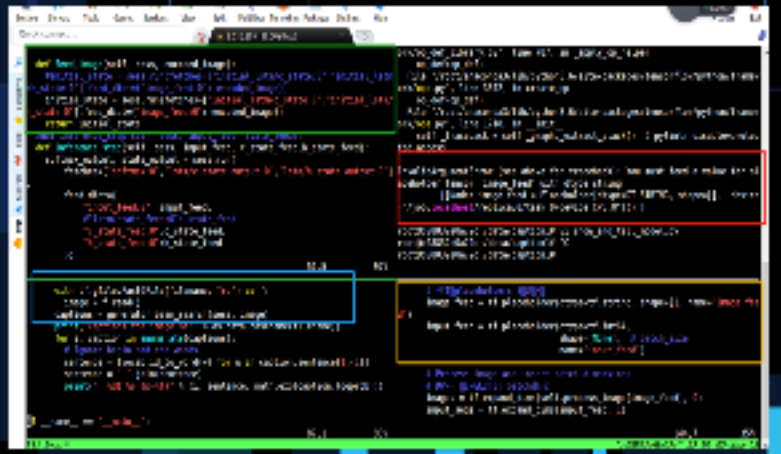
余杨 冯宇琨 王磊

本周解决的问题

问题：增加attention机制后，train过程中loss过大

原因：batch loss直接等于的loss，没有除以sum weight。修改后解决

```
INFO:tensorflow:global step 120220: loss = 3.1565 (0.255 sec/step)
INFO:tensorflow:global step 120240: loss = 2.9335 (0.264 sec/step)
INFO:tensorflow:global step 120540: loss = 3.1861 (0.315 sec/step)
INFO:tensorflow:global step 120550: loss = 3.2105 (0.282 sec/step)
INFO:tensorflow:global step 120560: loss = 3.1300 (0.244 sec/step)
INFO:tensorflow:global step 120570: loss = 2.8171 (0.285 sec/step)
INFO:tensorflow:global step 120580: loss = 2.8540 (0.269 sec/step)
INFO:tensorflow:global step 120590: loss = 2.8826 (0.260 sec/step)
INFO:tensorflow:global step 120700: loss = 2.7818 (0.259 sec/step)
INFO:tensorflow:global step 120710: loss = 2.8450 (0.269 sec/step)
INFO:tensorflow:global step 120720: loss = 2.8503 (0.256 sec/step)
INFO:tensorflow:global step 120730: loss = 2.9163 (0.255 sec/step)
INFO:tensorflow:global step 120740: loss = 3.1060 (0.326 sec/step)
```



Thanks !

前期回顾

1. im2txt框架运行调试
2.运行界面及后台搭建
3.encoder模块替换
4.attention机制引入
5.denscap搭建运行

本周工作

1.im2txt框架下attention的调试

包括train\evaluate\inference

2.densecap下joint inference和context fusion部分的调试

joint inference和context fusion是类似attention的机制

3.整理文档和代码

本周解决的问题

问题：增加attention机制后，train过程中loss过大

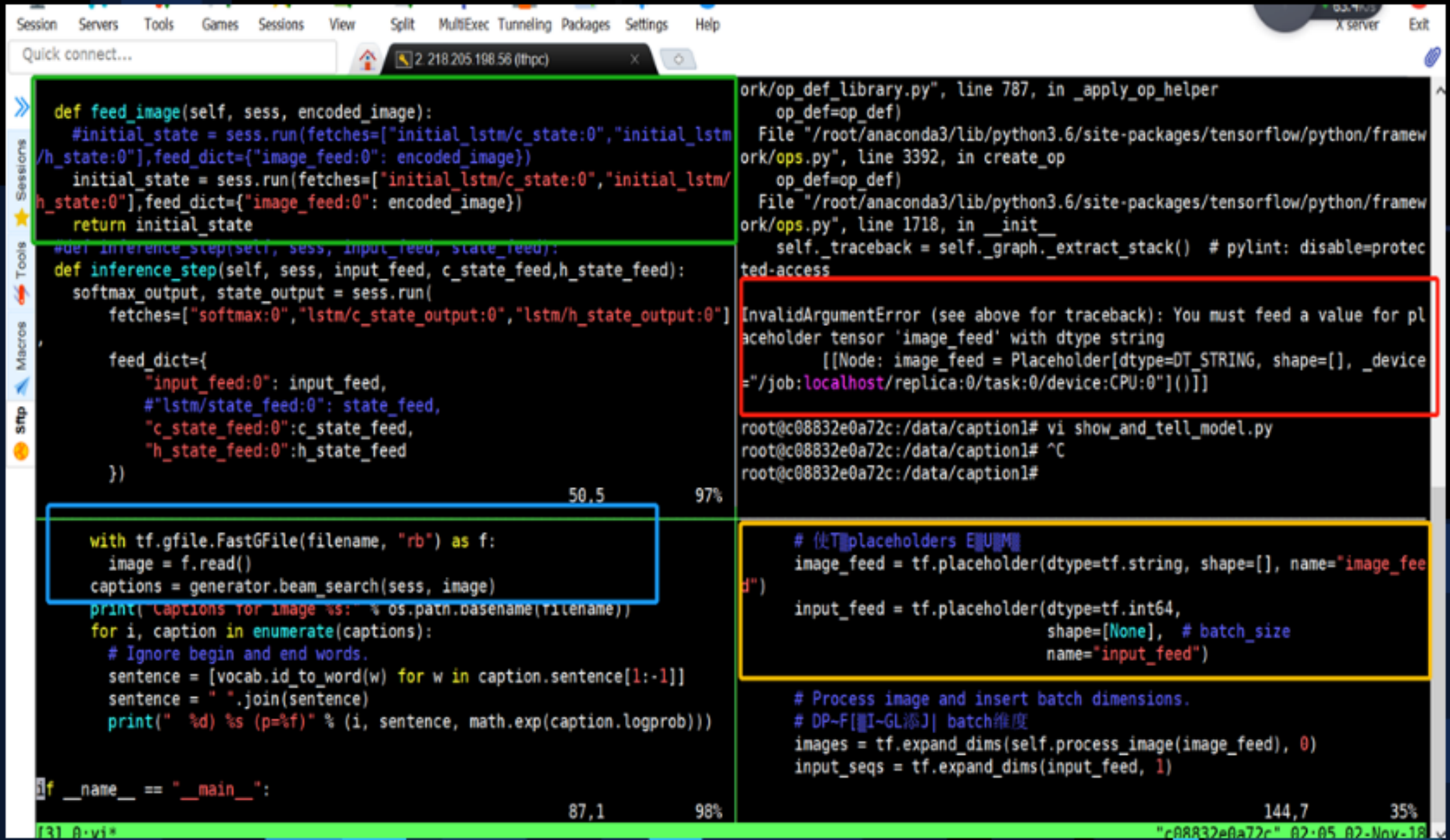
原因：batch loss直接等于的loss，没有除以sum weight，修改后解决

```
INFO:tensorflow:global step 120620: loss = 3.1545 (0.255 sec/step)
INFO:tensorflow:global step 120630: loss = 2.9335 (0.360 sec/step)
INFO:tensorflow:global step 120640: loss = 3.1681 (0.315 sec/step)
INFO:tensorflow:global step 120650: loss = 3.0103 (0.282 sec/step)
INFO:tensorflow:global step 120660: loss = 3.1390 (0.244 sec/step)
INFO:tensorflow:global step 120670: loss = 2.8171 (0.265 sec/step)
INFO:tensorflow:global step 120680: loss = 2.9540 (0.269 sec/step)
INFO:tensorflow:global step 120690: loss = 2.8826 (0.260 sec/step)
INFO:tensorflow:global step 120700: loss = 2.7818 (0.256 sec/step)
INFO:tensorflow:global step 120710: loss = 2.6459 (0.288 sec/step)
INFO:tensorflow:global step 120720: loss = 2.6503 (0.256 sec/step)
INFO:tensorflow:global step 120730: loss = 2.9163 (0.255 sec/step)
INFO:tensorflow:global step 120740: loss = 3.1060 (0.326 sec/step)
```

杨 鸿 宇

遇到的问题

1.im2txt运行evaluate和inference时，有bug在调试
2.使用densecap，将句子连接成段落阶段，loss未收敛，在调试


Session Servers Tools Games Sessions View Split MultiExec Tunneling Packages Settings Help
X server Exit
Quick connect...
2. 218.205.198.56 (thpc)
Sessions Tools Macros Sftp
def feed_image(self, sess, encoded_image):
#initial_state = sess.run(fetches=["initial_lstm/c_state:0","initial_lstm/h_state:0"],feed_dict={"image_feed:0": encoded_image})
initial_state = sess.run(fetches=["initial_lstm/c_state:0","initial_lstm/h_state:0"],feed_dict={"image_feed:0": encoded_image})
return initial_state
def inference_step(self, sess, input_feed, c_state_feed,h_state_feed):
softmax_output, state_output = sess.run(
fetches=["softmax:0","lstm/c_state_output:0","lstm/h_state_output:0"],
feed_dict={
"input_feed:0": input_feed,
#"lstm/state_feed:0": state_feed,
"c_state_feed:0":c_state_feed,
"h_state_feed:0":h_state_feed
})
50,5 97%
with tf.gfile.FastGFile(filename, "rb") as f:
image = f.read()
captions = generator.beam_search(sess, image)
print("Captions for image %s:" % os.path.basename(filename))
for i, caption in enumerate(captions):
# Ignore begin and end words.
sentence = [vocab.id_to_word(w) for w in caption.sentence[1:-1]]
sentence = " ".join(sentence)
print(" %d) %s (p=%f)" % (i, sentence, math.exp(caption.logprob)))
if __name__ == "__main__":
87,1 98%
[3] 0:vi*
ork/op_def_library.py", line 787, in _apply_op_helper
op_def=op_def)
File "/root/anaconda3/lib/python3.6/site-packages/tensorflow/python/framework/ops.py", line 3392, in create_op
op_def=op_def)
File "/root/anaconda3/lib/python3.6/site-packages/tensorflow/python/framework/ops.py", line 1718, in __init__
self._traceback = self._graph._extract_stack() # pylint: disable=protected-access
InvalidArgumentError (see above for traceback): You must feed a value for placeholder tensor 'image_feed' with dtype string
[[Node: image_feed = Placeholder[dtype=DT_STRING, shape=[], _device="/job:localhost/replica:0/task:0/device:CPU:0"]()]]
root@c08832e0a72c:/data/caption1# vi show_and_tell_model.py
root@c08832e0a72c:/data/caption1# ^C
root@c08832e0a72c:/data/caption1#
# 使T placeholders E U M
image_feed = tf.placeholder(dtype=tf.string, shape=[], name="image_feed")
input_feed = tf.placeholder(dtype=tf.int64,
shape=[None], # batch_size
name="input_feed")
# Process image and insert batch dimensions.
# DP~F[I~GL添J| batch维度
images = tf.expand_dims(self.process_image(image_feed), 0)
input_seqs = tf.expand_dims(input_feed, 1)
144,7 35%
"c08832e0a72c" 02:05 02-Nov-18


Exit
ramew
ramew
rotec
or pl
evice
e_fee
35%
Thanks !