

# 第一次报告

丁逸凡 2018202112

通过对相关论文的学习, 我们组确定了, 用卷积神经网络提取图像中的可视化特征, 使用LSTM循环神经网络将这些特征编码成为句子的大体思路. 同时, 加入注意力机制以改善句子的质量.

## 用到的库

```
1  Tensorflow
2  Numpy
3  OpenCV
4  Natural Language Toolkit
5  Pandas
6  Matplotlib
7  tqdm
```

## 用到的数据集

The COCO train2014 and val2014 data

## 实验结果

用BLEU标准打分

- BLEU-1 = 70.3%
- BLEU-2 = 53.6%
- BLEU-3 = 39.8%
- BLEU-4 = 29.5%

下面有一些比较成功的案例可供展示

**a train traveling down a track  
next to a forest.**



**a man wearing a suit and tie  
talking on a cell phone.**



## 参考资料

[Show, Attend and Tell: Neural Image Caption Generation with Visual Attention]

[Choosing\_Linguistic\_over\_Vision\_to\_Describe\_Image]