

人民币冠名号识别方案分享

---Titanik Data

---2019.07

团队介绍

队长：自强

队员：夏小满

王三岁

龙哥

健成



项目分析

人民币冠名号的识别

难点：

- 图片整体太大，需要识别的区域所占整体图片比例很小
- 没有标注的坐标位置(若以检测切片+识别)
- 算法的搭建(检测+识别/直接识别)



难点解决方案

- 方案一: 检测加识别
 1. 手动标注的框的位置, 用于训练检测网络(labelimg)
 2. 使用mmdetection库用于进行训练检测网络,参考知乎链接
(<https://zhuanlan.zhihu.com/p/56223200>)
(<https://www.zhihu.com/question/294578141>)
 3. 将检测下来的图片放入识别网络进行识别
(<https://github.com/clovaai/deep-text-recognition-benchmark>)
(<https://arxiv.org/abs/1904.01906>)

方案一：检测加识别

- 检测

我们采用的是faster rcnn，效果如下，当然可以选用专用的文字检测网络(CTPN/EAST)



faster rcnn

SE73411824



faster rcnn

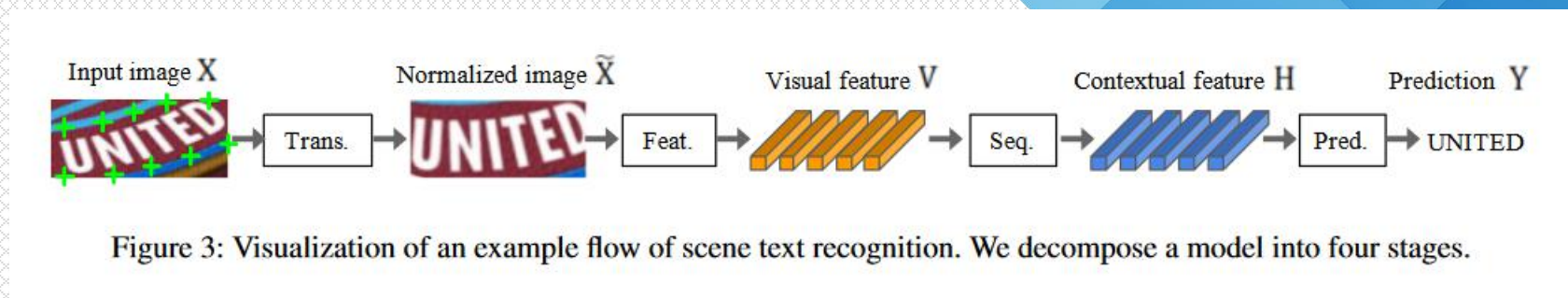
EM89439643

方案一：检测加识别

- 识别

我们采用的是deep-text-recognition-benchmark,论文主要针对的是弯曲的文字进行矫正后的识别，多了10个点的参数，相当于进行仿射变换后再进行识别。

论文：<https://arxiv.org/abs/1904.01906>



1. Trans: 进行仿射变换

2. Feat: 提取feature(Resnet)

3. seq: LSTM进行编码

4. Pred: CTC或Attn进行解码

方案一：检测加识别

- 识别

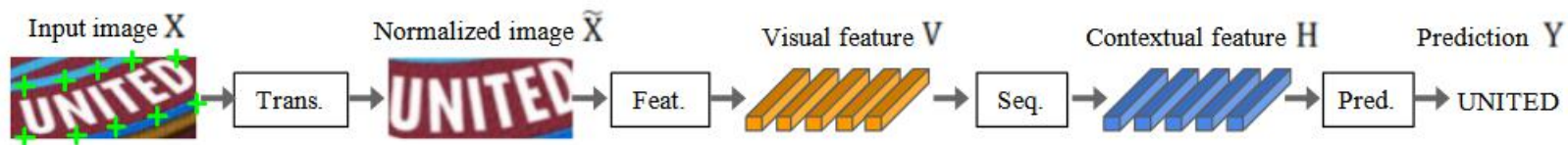


Figure 3: Visualization of an example flow of scene text recognition. We decompose a model into four stages.

训练了两个模型: CTC以及Attention, 均可以达到99.*%

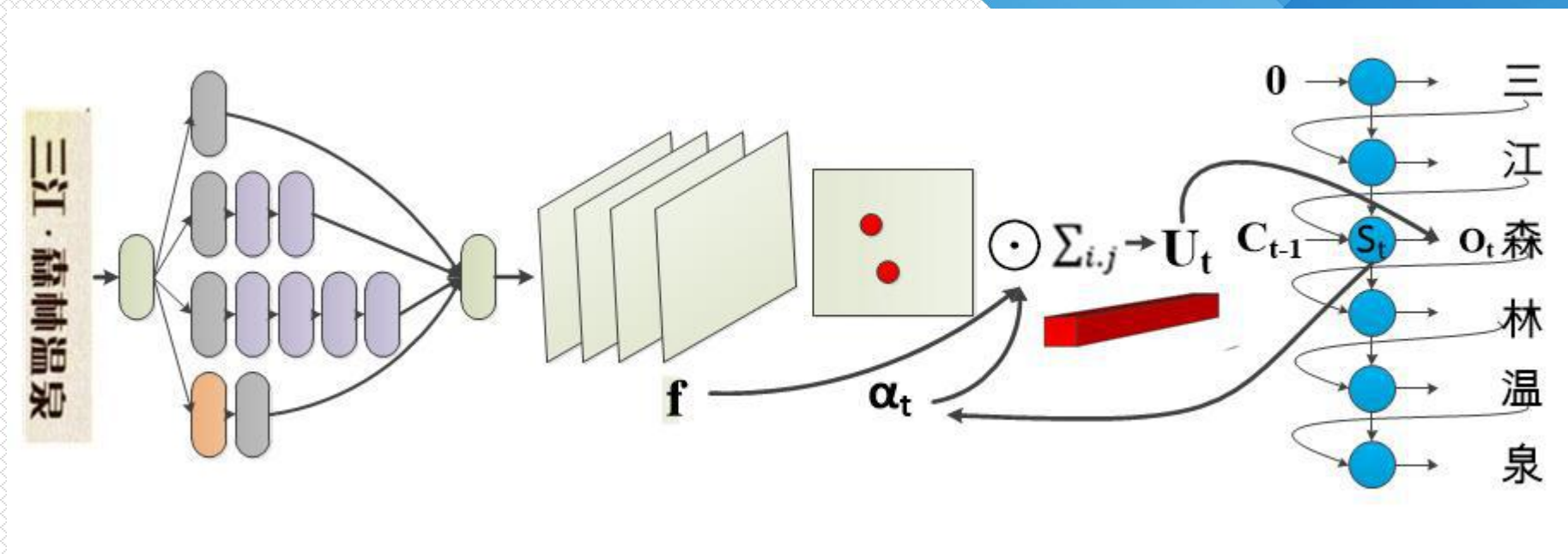
方案二：空间(Spatial)ATTN

- 空间attention(google)

https://github.com/tensorflow/models/tree/master/research/attention_ocr

<https://python.ctolib.com/A-bone1-Attention-ocr-Chinese-Version.html>

<https://arxiv.org/abs/1704.03549>

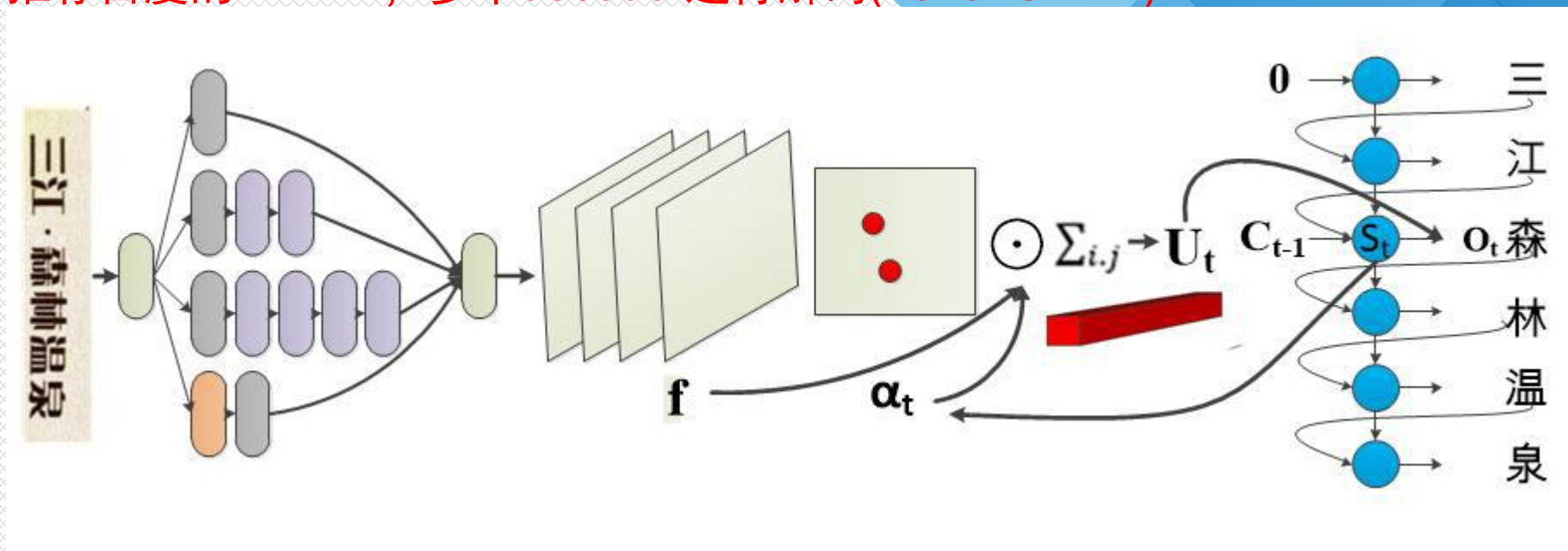


方案二：空间(Spatial)ATTN

优点: 不需要标注位置

缺点: 准确率相对来说更低(98.*%)

推荐百度的EATAN，多个decoder进行解码(2019 ICDAR)



我们的方案

1. 先用Faster RCNN进行检测切割，进行切片
2. 使用切片的数据进行CNN + LSTM + CTC/ATTN的训练两个模型
3. 使用Spatial ATTN进行训练
4. 得到三个模型，进行预测，在之前2w个图片的基础上进行了投票，类似其他队伍的做法，尽可能的挑出正确答案。
5. 将投票出来的数据进行扩充，三个模型中最好的模型是CTC模型，将数据放到CTC中进行数据的扩充(trick)，可以在扩充的数据上在训练三个模型进行投票(待试验)

致谢

1. 谢谢主办方的有序举办比赛，阿花的答疑。
2. 谢谢队友，几个队友帮忙标注数据。
3. 谢谢老婆





Q&A