

# ICPR MTWI 2018 挑战赛二比赛报告

队名: nelslip(iflytek&ustc)

团队成员: 朱意星, 张建树, 杜俊, 戴礼荣, 吴嘉嘉,

陈明军, 胡金水

## 一、 算法介绍

本次挑战赛我们使用的方法基于 FPN [1], 并且我们使用了 PANet [2] 的方法扩展了网络, 由于文本行有可能是任意形状的四边形, 针对文本行的特殊性, 我们使用了 SLPR [3], 具体而言我们利用水平和垂直的滑动线条去等距地拟合文本行的边缘, 这样可以解决任意四边形框四个顶点标注的歧义性, 之后我们使用文本行的轮廓去生成旋转矩形, 事实上类似于 Cascade R-CNN [4], 我们使用了两次 R-CNN, 第一步我们使用 RPN 生成水平矩形框, 第二步我们使用 SLPR 生成的轮廓得到旋转矩形框, 最后再次拟合四个顶点坐标。

## 二、 模型结构

我们增加了一些策略来使得模型获得更好的性能, 包括数据增强 (旋转, 缩放), 多尺度测试, 我们使用了在 ImageNet 上预训练的 ResNeXt-101 (32\*8d) [5] 作为网络的基础, 并且只使用了单一模型。

## 三、 开发环境

我们使用了 caffe2 作为实验平台, 所有实验都是在 1080Ti GPU 上进行的。

## 四、 数据集

本次比赛官方提供了 1 万张图片, 我们将这 1 万张图片分出 9 千张作为训练集, 另外 1 千张作为开发集。本次比赛我们没有使用额外数据

## 五、 团队介绍

本次参赛团队是由中国科学技术大学语音及语言信息处理国家工程实验室和科大讯飞合作组成, 其中张建树, 朱意星, 杜俊, 戴礼荣为中国科学技术大学语音及语言信息处理国家工程实验室成员, 陈明军, 吴嘉嘉, 胡金水为科大讯飞成员。我们希望通过参加本次比赛让外界关注到我们在文本检测领域的技术。

本团队的队长为张建树, 其邮箱地址为 [xysszjs@mail.ustc.edu.cn](mailto:xysszjs@mail.ustc.edu.cn), 手机号码为 15856910468, 现住址为中国安徽省合肥市蜀山区黄山路 443 号中国科学技术大学西校区 8 号宿舍楼 640 室。

## 六、 参考文献

[1] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature

pyramid networks for object detection,” in CVPR, vol. 1, no. 2, 2017, p. 4.

[2] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path aggregation network for instance segmentation,” arXiv preprint arXiv:1803.01534, 2018.

[3] Y. Zhu and J. Du, “Sliding line point regression for shape robust scene text detection,” arXiv preprint arXiv:1801.09969, 2018.

[4] Z. Cai and N. Vasconcelos, “Cascade r-cnn: Delving into high quality object detection,” arXiv preprint arXiv:1712.00726, 2017.

[5] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, “Aggregated residual transformations for deep neural networks,” in Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on. IEEE, 2017, pp. 5987–5995.