









理. 图 7(a)、图 7(b) 分别为两种算法的初步分割图, 相较于本文算法, 对照组初步分割的结果较粗糙, 有较多孤立的小片区域, 且文字区域不够明显. 最终分割结果如

图 7(c)、图 7(d), 对照组能够分割出文字, 但是出现一些漏分割现象, 如图中矩形框所示, 且精细程度不如本文提出的算法.

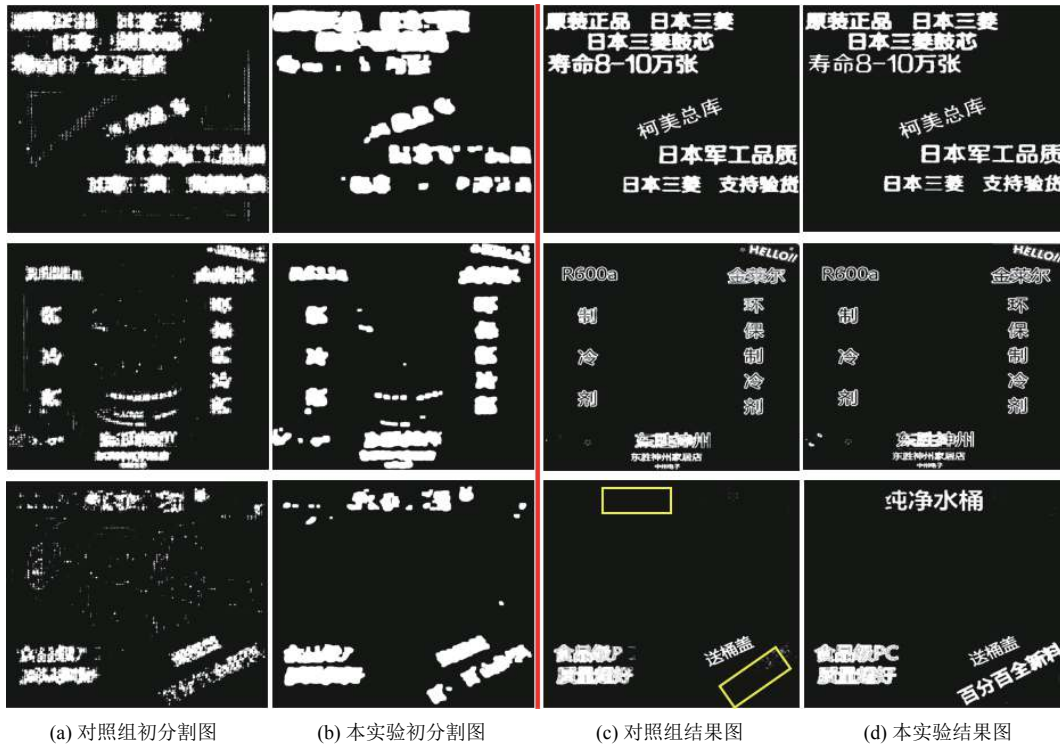


图 7 对比结果

由于验证集和测试集的标签类型如图 4(b) 所示, 无法使用传统的图像分割的评判标准, 如像素准确度、分割交占比 (IOU) 等, 本文提出一种新的评判标准, 评判步骤如下:

(1) 先用 OpenCV 对文字分割结果用最小矩形框集合, 如图 8(a) 所示, 矩形框内部即为目标文字区域.

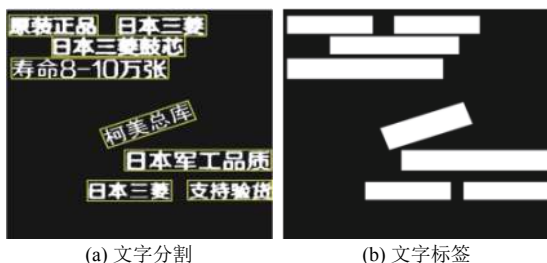


图 8 评价结果图

(2) 对真实标签的每个文字区域 (如图 8(b)), 按照式 (8) 计算其对预测结果中每个文字区域的  $C_i$ , 选取最

大  $C_{max}$ , 并将取最大值时两个文字区域的索引值记录到集合  $I$  中. 若  $C_{max} \geq 60\%$  表示预测正确; 若  $10\% \leq C_{max} < 60\%$  预测错误; 若  $C_{max} < 10\%$  则表示漏检; 若预测结果中文字区域的索引值未出现在  $I$  中, 则此区域为误检.

$$C = \frac{S_p \cap S_t}{S_t \cup S_t} \quad (8)$$

其中,  $S_p$  表示预测的文字区域,  $S_t$  表示真实文字区域.

(3) 错误总数 ( $N_{err}$ ) 为预测错误个数 ( $N_e$ ) 与漏检个数 ( $N_m$ ) 之和, 准确率计算如下:

$$acc = \frac{N_c}{N_a} \quad (9)$$

其中,  $N_c$  表示预测正确的总个数,  $N_a$  为验证集/测试集文字区域的总个数

根据此项判断标准, 实验结果如表 1 所示, 可以看出整体的实验效果较好, 没有出现拟合的情况; 同时, 可以发现结果中误检数量较漏检数量多, 可知本次实

验提出的算法容易将类似文字的图案(背景)识别为文字(目标). 本文算法的速度为 0.181 s/幅, 即 5.5 fps, 运行速度较快.

表 1 图像目标文字分割准确率分析

参数	验证集	测试集
图像数	1000	1000
真实区域个数 ( $N_a$ )	19 514	18 571
预测正确个数 ( $N_c$ )	17 036	15 915
预测错误个数 ( $N_e$ )	1641	1529
漏检个数 ( $N_m$ )	837	1127
误检个数 ( $N_f$ )	733	975
准确率 ( $acc$ )(%)	87.3	85.7

### 3 结论

本文提出了一种基于改进全卷积神经网络的图像目标分割算法, 此算法使用 FCN 模型进行初步分割, 再利用大津法进行二值化, 最后使用全连接条件随机场进行修正. 此算法在准确度和速度上都取得了较好的效果, 在测试集上可以达到 85.7% 的准确度以及 0.181 s/幅的速度.

#### 参考文献

- 林孜阳, 穆雪, 吴凯锋, 等. 基于连通域的快速文字图像分割算法. 计算机光盘软件与应用, 2014, (22): 89–90.
- 易小波. 身份证图像识别系统中文字分割的研究. 企业技术开发, 2003, (11): 19–21.
- 郑泽鸿, 黄成泉, 梁毅, 等. 基于 AP 聚类的中文字符分割. 智能计算机与应用, 2018, 8(1): 65–67, 71.
- Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 1–9.
- He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.

- Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Amsterdam, The Netherlands. 2016. 21–37.
- Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 779–788.
- Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 4293–4302.
- Chen LC, Papandreou G, Kokkinos I, *et al.* Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834–848. [doi: 10.1109/TPAMI.2017.2699184]
- Zhang Z, Zhang CQ, Shen W, *et al.* Multi-oriented text detection with fully convolutional networks. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 4159–4167.
- Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640–651. [doi: 10.1109/TPAMI.2016.2572683]
- Otsu N. A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62–66. [doi: 10.1109/TSMC.1979.4310076]
- Krähenbühl P, Koltun V. Efficient inference in fully connected CRFs with gaussian edge potentials. <http://arxiv.org/abs/1210.5644>.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. <https://arxiv.org/abs/1409.1556>.
- Abadi M, Agarwal A, Barham P, *et al.* TensorFlow: Large-scale machine learning on heterogeneous distributed systems. arXiv: 1603.04467, 2016.