一种改进的图像中的文本检测模型

刘　春 1，田倬韬 2，刘绍辉 2，安　源 1

（1 大庆师范学院，黑龙江 大庆 163712；2 哈尔滨工业大学，黑龙江 哈尔滨 163712）

摘　要： 本文提出了一种基于深度神经网络的高效自然场景文本检测模型 IEAST (Improved EAST).该模型基于 EAST (An Efficient and Accurate Scene Text Detector)进行了一系列改进，同时结合了物体检测以及实例分割的思想，通过用 Inception Module 替换原本特征融合的方式、加入 Path Aggregation 的思想以及采用 Fuzzy Mask 来提升模型的整体性能.IEAST 是一段式模型，其不仅检测准度上不落后于绝大多数二段式 State-of-the-art 的文本检测模型，还能保证极快的检测速度.

关键词： 深度神经网络；文本检测；实例分割

An improved text detection model in image

LIU Chun1，TIAN Zhuo-tao2，LIU Shao-hui2，AN Yuan1

(1 Daqing Normal University, Daqing 163712， China; 2 Harbin Institute of Technology, Harbin 150001, China)

Abstract： Image Text Detection is the process of accurately locating text area in given images, which is the important prerequisite of OCR (Optical Character Recognition). This paper introduces Improved IEAST (Improved EAST) which is an efficient and accurate scene text detection model based on deep neural network. This model which combines the knowledge of both object detection and instance segmentation is a variant of EAST (An Efficient and Accurate Scene Text Detector) with several improvements. By replacing conv3x3 feature merging with inception module, adding the idea of Path Aggregation and adopting Fuzzy Mask during training, IEAST gains improvements in performance with little compensation. Though IEAST is a one-stage model, it outperforms most of state-of-the-art two-stage text detection models by not only maintaining comparable accuracy but also inferencing faster. IEAST shows good result on ICDAR (International Conference on Document Analysis and Recognition) 2015 dataset

Key words： deep Learning; text detection; image processing

作者简介：

刘　春　女，（1970-），硕士，教授.研究方向为数据库、云计算、软件工程.

田倬韬　男，(1996-)，硕士研究生.研究方向为人工智能，深度学习，计算机视觉.

刘绍辉 （通讯作者） 男，（1977-），博士，副教授，硕士生导师.研究方向为计算机视觉、图像处理、目标跟踪.

E-mail:shliu@hit.edu.cn.

安　源　男，（1982-），硕士，讲师.研究方向为图像处理，软件工程.