基于改进型卷积神经网络和行特征的文本检测

方承志，樊梦雅

（南京邮电大学 电子与光学工程学院，江苏 南京 210023）

摘　要：针对现有的自然场景文本检测算法准确率尚未理想的问题，提出了一种基于改进型卷积神经网络和行特征的文本检测方法.首先，采用增强的最大稳定极值区域(MSER)提取图像的连通分量，并应用剪枝方法来获取孤立的连通区域；其次，应用改进型卷积神经网络(CNN)对非字符区域进行消除，获得候选字符区域；然后，提出基于行特征构建多方向候选文本行的算法，用于检测任意定向和弯曲的场景文本；最后，应用 C4.5 决策树算法对候选文本行进行分类.该算法在 ICDAR2013、ICDAR2015 和 MSER-TD500 数据集上进行实验, 实验结果表明，该算法能显著提高自然场景文本检测的准确率和召回率，且适用于任意方向、语言和字体的文本.

关键词：文本检测；最大稳定极值区域；卷积神经网络；行特征；C4.5 决策树算法

Text detection based on improved convolutional neural network and the feature of text lines

FANG Cheng-zhi，FAN Meng-ya

(College of Electronic and Optical Engineering，Nanjing University of Posts and Telecommunications，Nanjing 210023，China）

Abstract：In order to solve the problem of low precision in existing natural scene text detection algorithm, a text detection algorithm based on improved convolutional neural network and the feature of text lines was proposed. Firstly, the enhanced maximally stable extreme region (MSER) was used to extract connected components of the image and isolated connected regions were obtained by pruning method. Secondly, an improved convolutional neural network (CNN) was used to eliminate non-character regions and obtain candidate character regions. Thirdly, an algorithm based on the feature of text lines for constructing multi-oriented candidate text lines was proposed to detect arbitrary oriented and curved scene text. Finally, C4.5 decision tree algorithm was applied to classify candidate text lines. Experiments were carried out on ICDAR2013, ICDAR2015 and MSRA-TD500 datasets. Experimental results show that the algorithm can significantly improve the precision and recall of text detection in natural scenes, and is suitable for text in any direction, language and font.

Key words：text detection; maximally stable extremal region (MSER); convolutional neural network (CNN); the feature of text lines; C4.5 decision tree algorithm

作者简介：

方承志　男，（1976-），博士，副教授.研究方向为模式识别、嵌入式系统和信号处理.

樊梦雅（通讯作者）　女，（1994-），硕士研究生.研究方向为图像处理与模式识别.

E-mail:956594122@qq.com.