

ABSTRACT

Scene text detection has been developing in recent years. It is due to its numerous practical applications, such as automatic text translation, assist visually impaired people, content-based image retrieval and so on. This study proposes a quadtree-based candidate text regions extraction (CTR) to localize texts in the scene image. The CTR takes the advantage of color consistency of text. The result of CTR is a CTR map, extracted by dividing the image into homogeneous regions with quadtree and performing a dilation operation. To perform text localization and classification, a convolutional neural network (CNN) is adopted. The final result is taken from the bounding box of CNN which is compared with the CTR map. The experiments on MSRA-TD500 and ICDAR 2013 show that the proposed method outperforms the previous ones, and it achieves a competitive result on ICDAR 2015. The proposed method achieves F-score of 78.69%, 91.59%, 82.15% on MSRA-TD500, ICDAR 2013, and ICDAR 2015 datasets respectively.

Keywords: Scene text detection, quadtree, convolutional neural networks