Liu, Zichuan; Lin, Guosheng; Goh, Wang Ling


Summary: A novel detection framework named Markov Clustering Network (MCN) is proposed for fast and robust scene text detection. Different from the traditional top-down scene text detection approaches that inherit from the classic object detection, MCN detects scene text objects in a bottom-up manner. MCN predicts instance-level bounding boxes by firstly converting an image into a stochastic flow graph where Markov Clustering is performed based on the predicted stochastic flows. The stochastic flows encode the local correlation and semantic information of scene text objects. An object is modeled as strongly connected nodes by flows, which allows flexible and bottom-up detection for scale-varying and rotated text objects without prior knowledge of object size. The flow prediction is supported by the advanced Convolutional Neural Networks architectures and Position-aware spatial attention mechanism, which provides enhanced flow prediction by adaptively fusing spatial representations. The experimental evaluation on public benchmarks shows that our MCN method achieves the state-of-art performance on public benchmarks, especially in retrieving long and oriented texts.

MSC:
68T45 Machine vision and scene understanding
62H30 Classification and discrimination; cluster analysis (statistical aspects)
68T07 Artificial neural networks and deep learning

Keywords: neural networks; Markov clustering; self-attention; text detection and segmentation

Software:
EAST; CUDA; R2CNN; TextBoxes++; TextBoxes; Faster R-CNN; SSD

Full Text: DOI

References:


[55] Xue, Chuhui; Lu, Shijian; Zhan, Fangneng, Accurate Scene Text Detection Through Border Semantics Awareness and Bootstrapping, Computer Vision - ECCV 2018, 370-387 (2018), Cham: Springer International Publishing, Cham


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.