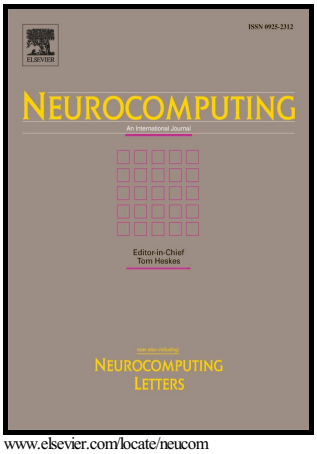
Author's Accepted Manuscript



www.elsevier.com/locate/neucom

Boundary-Aware Box Refinement for Object Proposal Generation

Xiaozhi Chen, Huimin Ma, Chenzhuo Zhu, Xiang Wang, Zhichen Zhao

PII: S0925-2312(16)31073-6
DOI: http://dx.doi.org/10.1016/j.neucom.2016.09.045
Reference: NEUCOM17574

To appear in: *Neurocomputing*

Received date: 29 March 2016
Revised date: 18 July 2016
Accepted date: 18 September 2016

Cite this article as: Xiaozhi Chen, Huimin Ma, Chenzhuo Zhu, Xiang Wang and Zhichen Zhao, Boundary-Aware Box Refinement for Object Proposal Generation, *Neurocomputing,* http://dx.doi.org/10.1016/j.neucom.2016.09.045

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting galley proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain

ACCEPTED MANUSCRIPT

Boundary-Aware Box Refinement for Object Proposal Generation

Xiaozhi Chen, Huimin Ma*, Chenzhuo Zhu, Xiang Wang, Zhichen Zhao

Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Abstract

Object proposals have been widely used in object detection to speed up object searching. However, many of existing object proposal generators have pool localization quality, which weakens the performance of object detectors. In this paper, we present an effective approach to improve the localization quality of object proposals. We leverage the boundary-preserving property of superpixels and design an efficient algorithm for object proposal refinement. Our approach first performs bounding box alignment to adapt proposals to potential object boundaries, and then diversifies the proposals via multi-thresholding superpixel merging. The algorithm only takes 0.15s and can be applied to any existing proposal methods to improve their localization quality. Extensive experiments on PASCAL VOC 2007 and ILSVRC 2013 datasets show our approach significantly and consistently improves the recall, localization accuracy, and detection performance of existing proposal methods. When combining with Region Proposal Network, our approach outperforms the state-of-the-art object detectors by a large margin.

Keywords: object proposals, object detection, R-CNN

Accepted manuscript

1. Introduction

Object proposal generation has become a crucial technique for many vision recognition tasks, such as class-specific object detection and instance segmentation. The goad of object proposal generation is to select a small set of object candidates that cover most of the objects in an image. The advantages of object proposals over traditional sliding windows [1] lie in two aspects: reducing computations with fewer regions of interest, and improving accuracy by using more sophisticated features and classifiers. Recent object detectors [2, 3, 4] utilizing object proposals have achieved state-of-the-art performance on challenging PASCAL [5], ImageNet [6] and MS COCO [7] datasets.

Two distinct pipelines emerge for object proposal generation: bottom-up approaches based on low-level cues, and data-driven approaches based on Convolutional Neural Networks (CNN). Most bottom-up approaches generate proposals by window scoring or region grouping. Existing models belonging to this pipeline struggle to achieve a good balance between localization accuracy and computational efficiency. In particular, window scoring based methods, such as BING [8], are computationally efficient, but they suffer from pool localization quality, i.e., low recall under strict intersection over union (IoU) overlap criteria (e.g., IoU > 0.7). Region grouping based methods such as Selective Search [9] and MCG [10] are computationally expensive, while they usually achieve higher localization quality. Data-driven approaches [4, 11] leverage rich convolutional neural network features to directly predict the regions of interest. Due to the powerful discrimination ability of CNN features, these models typically achieve higher recall than bottom-up proposals under loose overlap criteria (e.g., IoU < 0.7). However, similar to window scoring based methods, they usually have much lower recall at high overlap threshold.

Motivation. Most window scoring based and CNN-based methods fail to achieve high recall under strict overlap criteria. While region grouping based methods have better localization accuracy, they are usually computationally expensive. The goal of this paper is to improve localization accuracy of existing object proposals while preserving computational efficiency. Our work is inspired by the observation that superpixels have the good property of preserving object boundaries, which could benefit object localization. Prior superpixel-based methods typically exploit low-level features like color, texture, to compute region similarity for image parti-

*Corresponding author
Email address: mhmpub@tsinghua.edu.cn (Huimin Ma)

Download English Version:

https://daneshyari.com/en/article/4948243

Download Persian Version:

https://daneshyari.com/article/4948243

Daneshyari.com