

Accepted Manuscript

Temporally Enhanced Image Object Proposals for Online Video Object and Action Detections

Jiong Yang, Yuan Junsong

PII: S1047-3203(18)30075-0
DOI: <https://doi.org/10.1016/j.jvcir.2018.03.018>
Reference: YJVC I 2165

To appear in: *J. Vis. Commun. Image R.*

Received Date: 26 September 2017
Accepted Date: 28 March 2018

Please cite this article as: J. Yang, Y. Junsong, Temporally Enhanced Image Object Proposals for Online Video Object and Action Detections, *J. Vis. Commun. Image R.* (2018), doi: <https://doi.org/10.1016/j.jvcir.2018.03.018>

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 proof before it is published in its final 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.



Temporally Enhanced Image Object Proposals for Online Video Object and Action Detections

Jiong Yang^{a,*}, Yuan Junsong^b

^aInterdisciplinary Graduate School, Nanyang Technological University, Singapore

^bSchool of Electrical Electronic Engineering, Nanyang Technological University, Singapore

Abstract

Despite the recent advances of image object proposals (IOPs) and video object proposals (VOPs), it still remains a challenge to apply them to online video object/action detection. To address this problem, we propose a novel form of image object proposals, Temporally Enhanced Image Object Proposals (TE-IOPs), for online video object/action detection. The proposed TE-IOPs augment the existing IOPs at every frame by their temporal dynamics in the past few frames. We develop a dynamic programming scheme to efficiently search for such TE-IOPs in an online manner. Compared with existing VOPs that cannot run online, our TE-IOPs can be used for online detection. Compared with IOPs, our TE-IOPs bring rich temporal dynamics with minor computational cost. Experiments on benchmark datasets validate the superior performance of the proposed TE-IOPs over existing IOPs and VOPs, in terms of both the proposal re-ranking and the application of online action detection.

Keywords: video, proposal, online, detection, temporal

1. Introduction

In recent years, image object proposals (IOPs) [1, 2] and video object proposals (VOPs) [3, 4, 5] have been actively studied. IOP methods generate image bounding boxes or segments to capture candidate object locations in each frame, while VOP methods generate spatiotemporal tubes to capture the spatiotemporal extent of the candidate objects throughout the video. Using these proposals can dramatically reduce the search space of the traditional sliding window based

*Corresponding author

Email addresses: yang0374@e.ntu.edu.sg (Jiong Yang), jsyuan@ntu.edu.sg (Yuan Junsong)

Download English Version:

<https://daneshyari.com/en/article/6938246>

Download Persian Version:

<https://daneshyari.com/article/6938246>

[Daneshyari.com](https://daneshyari.com)