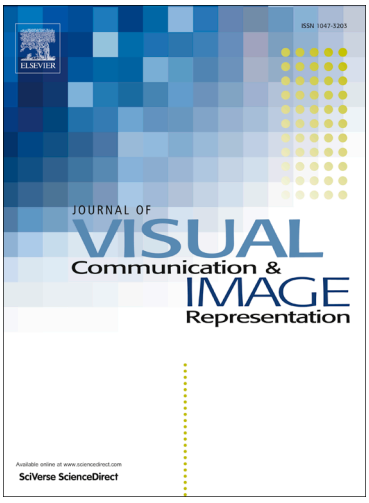

Object Detection from Dynamic Scene Using Joint Background Modeling and Fast Deep Learning Classification

Hayder Yousif, Jianhe Yuan, Roland Kays, Zhihai He

PII: S1047-3203(18)30201-3
DOI: https://doi.org/10.1016/j.jvcir.2018.08.013
Reference: YJVCI 2262

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

Received Date: 1 January 2018
Revised Date: 5 August 2018
Accepted Date: 16 August 2018

Please cite this article as: H. Yousif, J. Yuan, R. Kays, Z. He, Object Detection from Dynamic Scene Using Joint Background Modeling and Fast Deep Learning Classification, *J. Vis. Commun. Image R.* (2018), doi: https://doi.org/10.1016/j.jvcir.2018.08.013

Object Detection from Dynamic Scene Using Joint Background Modeling and Fast Deep Learning Classification

Hayder Yousif, Jianhe Yuan, Roland Kays, and Zhihai He

hyypp5@mail.missouri.edu, yuanjia@missouri.edu, rwkay@ncsu.edu, and hezhi@missouri.edu

Abstract

In this paper, we couple effective dynamic background modeling with fast deep learning classification to develop an accurate scheme for human-animal detection from camera-trap images with cluttered moving objects. We introduce a new block-wise background model, named as Minimum Feature Difference (MFD), to model the variation of the background of the camera-trap sequences and generate the foreground object proposals. We then develop a region proposals verification to reduce the number of false alarms. Finally, we perform complexity-accuracy analysis of DCNN to construct a fast deep learning classification scheme to classify these region proposals into three categories: *human*, *animals*, and *background* patches. The optimized DCNN is able to maintain high level of accuracy while reducing the computational complexity by 14 times, which allows near real-time implementation of the proposed method on CPU machines. Our experimental results demonstrate that the proposed method outperforms existing methods on our and Alexander von Humboldt Institute camera-trap datasets in both foreground segmentation and object detection.

Keywords:

Human-animal detection, camera-trap images, background subtraction, deep