

Categorization of Multiple Objects in a Scene Using a Biased Sampling Strategy

Lei Yang · Nanning Zheng · Mei Chen · Yang Yang · Jie Yang

Received: 28 March 2010 / Accepted: 2 May 2013 / Published online: 28 June 2013
© Springer Science+Business Media New York 2013

Abstract Recently, various bag-of-features (BoF) methods show their good resistance to within-class variations and occlusions in object categorization. In this paper, we present a novel approach for multi-object categorization within the BoF framework. The approach addresses two issues in BoF related methods simultaneously: how to avoid scene modeling and how to predict labels of an image when multiple categories of objects are co-existing. We employ a biased sampling strategy which combines the bottom-up, biologically inspired saliency information and loose, top-down class prior information for object class modeling. Then this biased sampling component is further integrated with a multi-instance multi-label learning and classification algorithm. With the proposed biased sampling strategy, we can perform multi-object categorization within an image without semantic segmentation. The experimental results on PASCAL VOC2007

and SUN09 show that the proposed method significantly improves the discriminative ability of BoF methods and achieves good performance in multi-object categorization tasks.

Keywords Object categorization · Bag-of-features method · Biased sampling strategy · Multi-instance multi-label learning

1 Introduction

Object categorization is challenging due to the need for generalization across large intra-class appearance variations. Recently, the problem of appearance variations in object categorization has been successfully addressed by the discriminative bag-of-features (BoF) methods (Csurka et al. 2004; Lazebnik et al. 2006; Zhang et al. 2007; Shotton and Cipolla 2008; Parikh et al. 2008), which represents an image as an un-ordered collection of local features. However, when there are multiple co-existing categories or background clutters in the scene, the current BoF methods would often fail because they are applied to the entire scene rather than the isolated object of interest specifically. Figure 1 shows an example that the typical BoF representation mixes information from different categories of objects as well as the background in an image. It is clear that the characteristics of each object are obscured by others. Thus, for a realistic multi-object categorization problem, the classical BoF methods can not be utilized directly.

To avoid scene modeling and solve the multi-category co-existence problem, a straightforward solution is to seek help from the object detection (localization) (Pantofaru et al. 2006; Ferrari et al. 2008; Lampert et al. 2009; Harzallah et al. 2009) or semantic annotation (Rabinovich et al. 2007;

This paper was originally submitted to the special issue featuring extended papers from ACCV.

L. Yang (✉)
China Mobile Research Institute,
Xi'an Jiaotong University, Beijing 100053, China
e-mail: ylangei@hotmail.com

N. Zheng · Y. Yang
Institute of Artificial Intelligence and Robotics,
Xi'an Jiaotong University, Xi'an 710049, China
e-mail: nnzheng@mail.xjtu.edu.cn

Y. Yang
e-mail: yyang_993@hotmail.com

M. Chen
Intel Labs Pittsburgh, Pittsburgh, PA 15213, USA
e-mail: mei.chen@intel.com

J. Yang
Carnegie Mellon University, Pittsburgh, PA 15213, USA
e-mail: jie.yang@cs.cmu.edu