International Journal of Computer Vision (2023) 131:680–700
https://doi.org/10.1007/s11263-022-01728-z

Check for updates

Multi-adversarial Faster-RCNN with Paradigm Teacher for Unrestricted Object Detection

Zhenwei He[1] · Lei Zhang[1] · Xinbo Gao[2] · David Zhang[3]

Received: 23 January 2022 / Accepted: 30 November 2022 / Published online: 11 December 2022

Abstract

Recently, the cross-domain object detection task has been raised by reducing the domain disparity and learning domain invariant features. Inspired by the image-level discrepancy dominated in object detection, we introduce a **M**ulti-**A**dversarial **F**aster-RCNN (MAF). Our proposed MAF has two distinct contributions: (1) The Hierarchical Domain Feature Alignment (HDFA) module is introduced to minimize the image-level domain disparity, where Scale Reduction Module (SRM) reduces the feature map size without information loss and increases the training efficiency. (2) Aggregated Proposal Feature Alignment (APFA) module integrates the proposal feature and the detection results to enhance the semantic alignment, in which a weighted GRL (WGRL) layer highlights the hard-confused features rather than the easily-confused features. However, MAF only considers the domain disparity and neglects domain adaptability. As a result, the label-agnostic and inaccurate target distribution leads to the source error collapse, which is harmful to domain adaptation. Therefore, we further propose a Paradigm Teacher (PT) with knowledge distillation and formulated an extensive **P**aradigm **T**eacher MAF (PT-MAF), which has two new contributions: (1) The Paradigm Teacher (PT) overcomes source error collapse to improve the adaptability of the model. (2) The Dual-Discriminator HDFA (D^2-HDFA) improves the marginal distribution and achieves better alignment compared to HDFA. Extensive experiments on numerous benchmark datasets, including the Cityscapes, Foggy Cityscapes, Pascal VOC, Clipart, Watercolor, etc. demonstrate the superiority of our approach over SOTA methods.

Keywords Object detection · Transfer learning · Domain adaptation · CNN

1 Introduction

Object detection is a basic task in computer vision, which is challenging due to the diversity of illumination, viewpoint, occlusion, or other factors. Recently, object detection has attracted amounts of attention. Inspired by the success of CNN (He et al., 2016; Krizhevsky & Sutskever, 2012; Simonyan & Andrew, 2014), object detection has witnessed a great development (Girshick, 2015; Liu et al., 2016; Ren et al., 2015; Redmon & Farhadi, 2018; Lian et al., 2017; He et al., 2017).

The real-world applications require the object detectors to work in the wild, where the scenario is much different from their training environment. As a result, a detector trained with samples drawn from normal weather is not transferable to other weather conditions of different application scenarios. In fact, most of the existing datasets for object detection are domain restricted, and the trained detectors are difficult to adapt to other domains or scenarios due to the unavoidable domain disparity. Moreover, the conventional

Communicated by Wanli Ouyang.

✉ Lei Zhang
leizhang@cqu.edu.cn

Zhenwei He
hzw@cqu.edu.cn

Xinbo Gao
gaoxb@cqupt.edu.cn

David Zhang
davidzhang@cuhk.edu.cn

1 School of Microelectronics and Communication Engineering, Chongqing University, Shazheng street No.174, Chongqing 400044, China

2 Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

3 School of Science and Engineering, Chinese University of Hong Kong (Shenzhen), Shenzhen 518172, China