Weakly Supervised Object Localization with Progressive Domain Adaptation

Dong Li[1], Jia-Bin Huang[2], Yali Li[1], Shengjin Wang[1], Ming-Hsuan Yang[3]

[1]Tsinghua University, [2]University of Illinois, Urbana-Champaign, [3]University of California, Merced

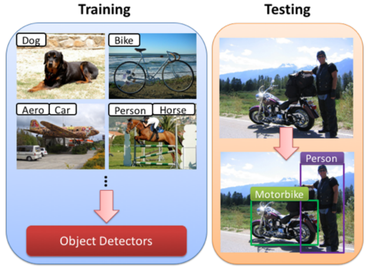
CVPR 2016

Highlights

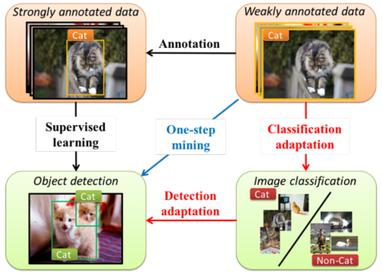
- Weakly supervised learning of object detectors
- State-of-the-art results on the PASCAL 2007, 2010, 2012 and ILSVRC 2013 datasets
- Code and models available: *http://bit.ly/wsl2016*

Introduction

Problem setting: Training object detectors with image-level annotations



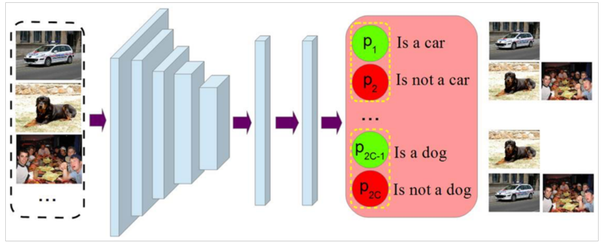
Key observation: Large gap between source domain and target task



Proposed Approach - progressive domain adaptation

I. Classification adaptation:

Goal: ImageNet single-label classification → multi-label classification

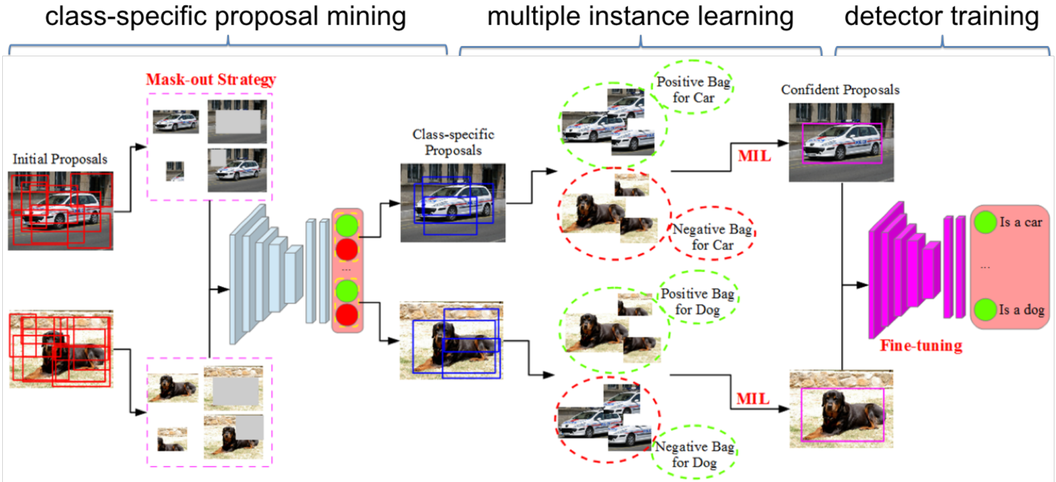


Proposed multi-label loss

- Softmax loss for binary classification of each class
- Summation of softmax losses over all classes

II. Detection adaptation:

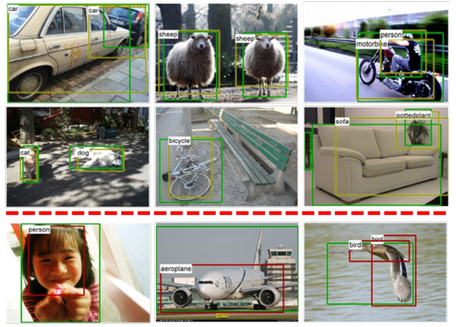
Goal: multi-label classification → object detection



Experiments

Comparisons to the state-of-the-arts:

- OM: class-specific proposal mining ➢ **VOC'07**
- MIL: multiple instance learning
- FT: detector training



Methods	mAP
Cinbis et al. CVPR'14	22.4
Song et al. ICML'14	22.7
Song et al. NIPS'14	24.6
Bilen et al. BMVC'14	26.4
Bilen et al. CVPR'15	27.7
Wang et al. ECCV'14	31.6
OM + MIL	23.4
OM + FT-AlexNet	19.5
MIL + FT-AlexNet	23.0
OM + MIL+ FT-AlexNet	31.0
OM + FT-VGGNet	20.5
MIL + FT-VGGNet	26.2
OM + MIL+ FT-VGGNet	**39.5**

➢ **VOC'10, '12 and ILSVRC'13**

Methods	VOC 2010	VOC 2012	ILSVRC 2013
Cinbis et al. CVPR'14	18.5	-	-
Wang et al. ECCV'14	-	-	6.0
OM + MIL+ FT-AlexNet	23.3	21.7	7.7
OM + MIL+ FT-VGGNet	**34.6**	**32.6**	**10.8**

Error analysis: