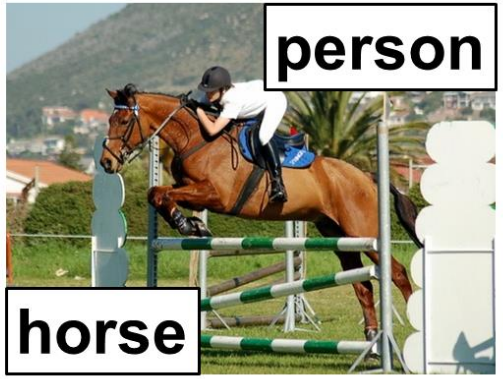
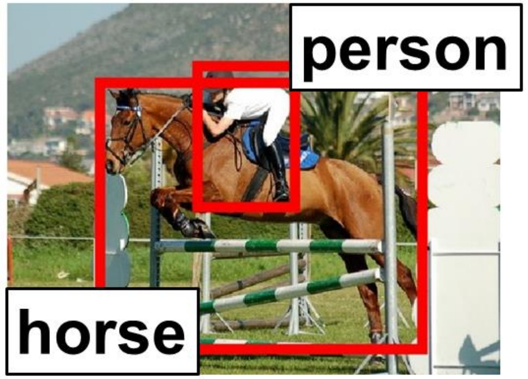
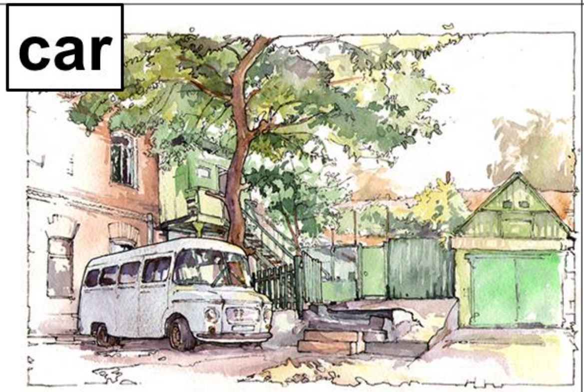
浙江大学 信息与电子工程学院
COLLEGE OF INFORMATION SCIENCE & ELECTRONIC ENGINEERING
ZHEJIANG UNIVERSITY

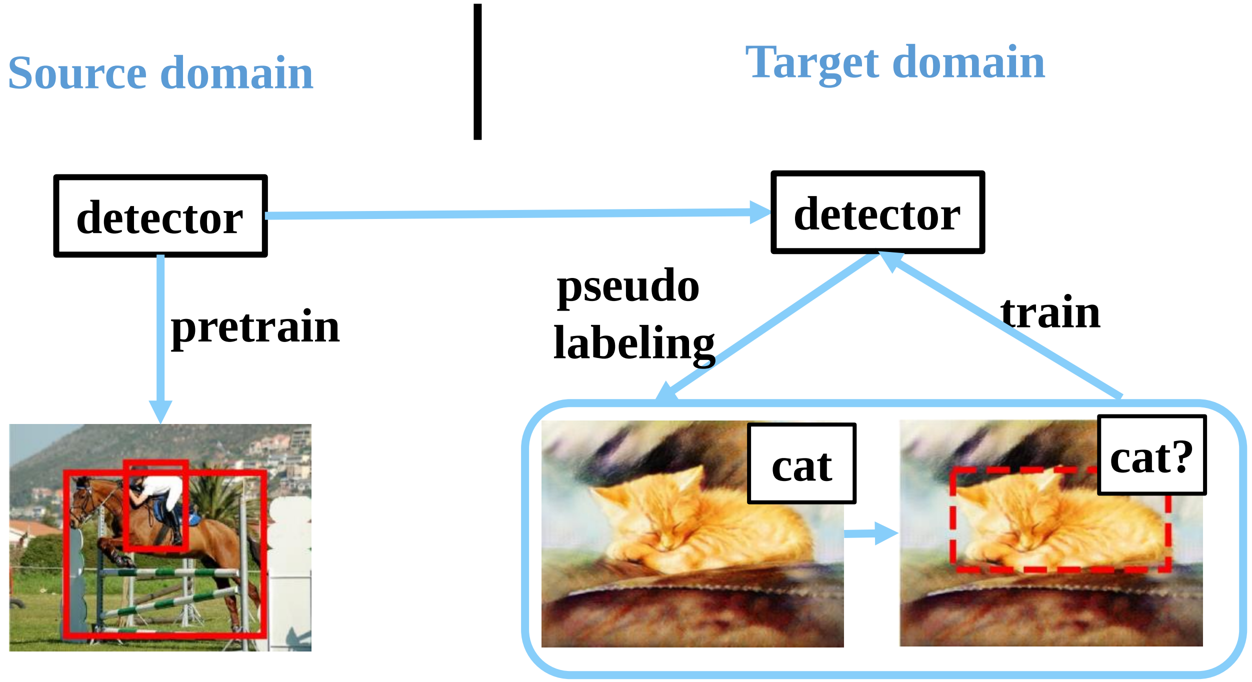
PSEUDO-LABEL GENERATION-EVALUATION FRAMEWORK FOR CROSS DOMAIN WEAKLY SUPERVISED OBJECT DETECTION

Shengxiong Ouyang, Xinglu Wang, Kejie Lyu, Yingming Li*
Zhejiang University

Background

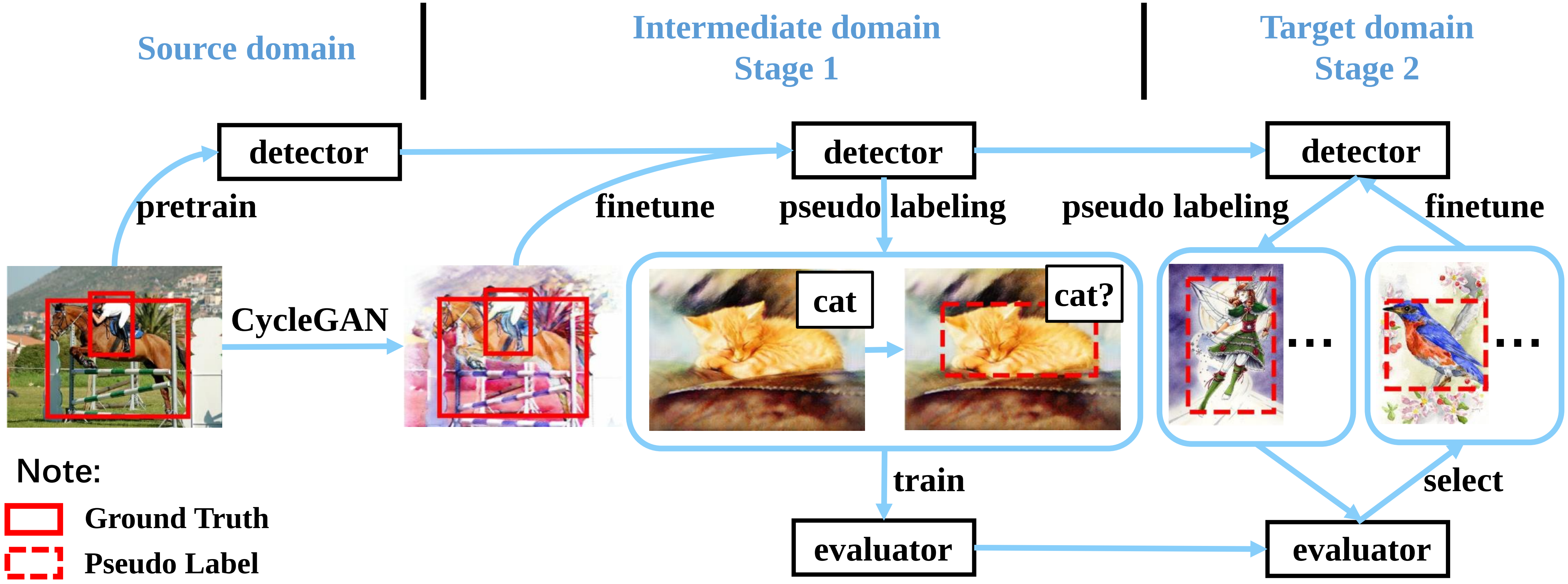
	Level of annotations	
	Image	Instance
Source domain		
Target domain		—

Challenges



Mislabeled or inaccurate samples may hurt the performance of the detector and thus be less valuable for training.

Methods



Source domain
Intermediate domain
Stage 1
Target domain
Stage 2
detector
pretrain
finetune
pseudo labeling
CycleGAN
cat
cat?
train
select
evaluator
Note:
Ground Truth
Pseudo Label

Methods

Algorithm 1 Overall Procedure

Input: source domain data $\mathbb{S}$, target domain data $\mathbb{T}$, the detector with parameter θ_d, the evaluator with parameter θ, selection ration r and maximum round number n.

▷ Pretrain the detector on the source domain data $\mathbb{S}$.

▷ **Stage1**

1. Transfer $\mathbb{S}$ to the intermediate domain and produce data $\mathbb{F}$ via CycleGAN. Divide $\mathbb{F}$ into $\mathbb{F}_1$ and $\mathbb{F}_2$.
2. Finetune the detector on $\mathbb{F}_1$ and generate pseudo labels on $\mathbb{F}_2$. Train the evaluator on $\mathbb{F}_2$.

▷ **Stage2**

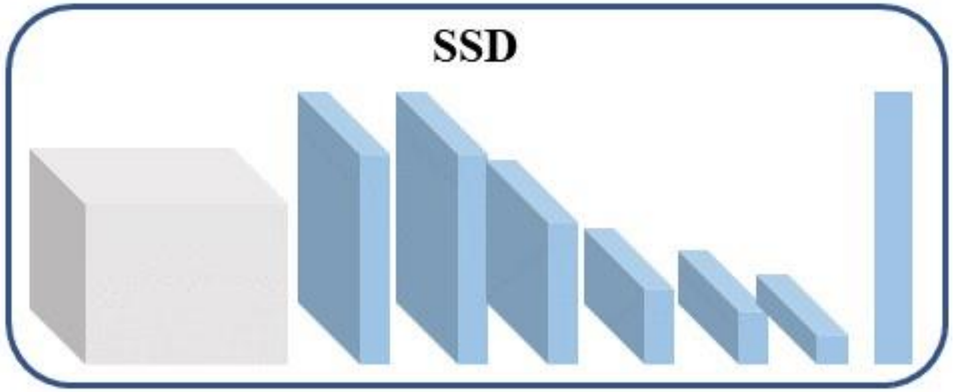
for $i = 1; i \leq n; i++$ **do**

1. Generate pseudo labels on $\mathbb{T}$ and compute quality scalar q of each sample via evaluator.
2. Select top $r \times |\mathbb{T}|$ samples as the subdataset. Finetune the detector on it.

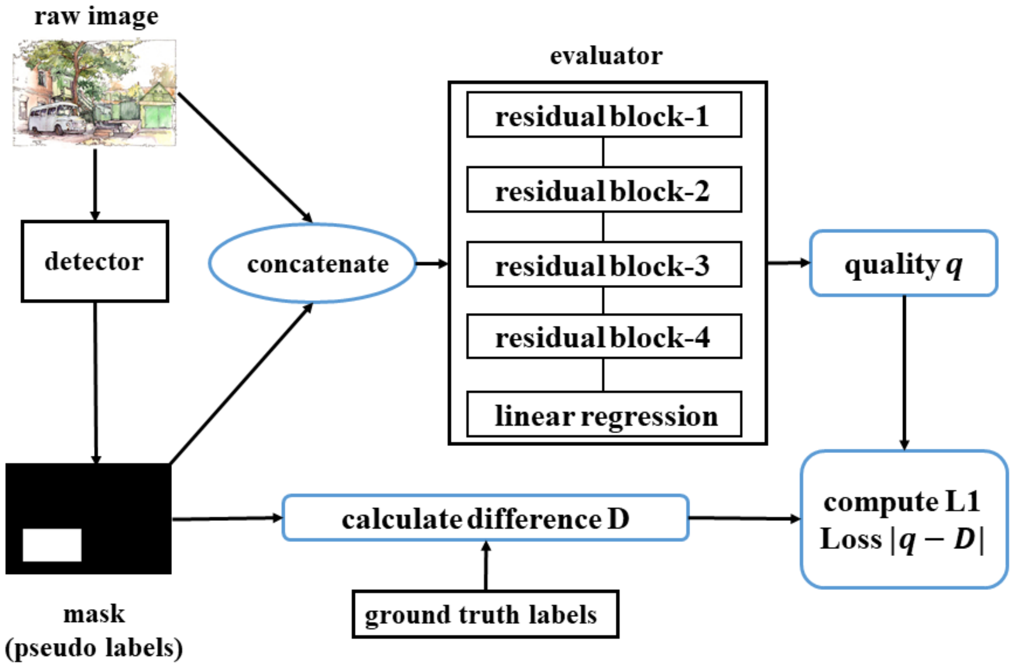
end for

Output: Detector with optimized parameter θ_d.

Detector



Evaluator



Experiments

Methods	Brief introduction
SSD300	A direct baseline
Ideal case	Instance-level annotations are available in the target domain
DT+PL	The SOTA CDWSOD method
DC	The SOTA unsupervised domain adaptation algorithm
MI-max	The SOTA weakly-supervised object detector
SSD-joint	To validate the effectiveness of the evaluator
Ours-type1	Sort the quality scores of all samples
Ours-type2	Sort the quality score within each class at the first round

Experiments

Pascal VOC→ Clipart1k																					
Model	aero	bike	bird	boat	bttle	bus	car	cat	chair	cow	table	dog	horse	mbike	prsn	plnt	sheep	sofa	train	tv	**mAP**
baseline	19.8	49.5	20.1	23.0	11.3	38.6	34.2	2.5	39.1	21.6	27.3	10.8	32.5	54.1	45.3	31.2	19.0	19.5	19.1	17.9	26.8
Compared																					
DT+PL	35.7	**61.9**	26.2	**45.9**	29.9	74.0	48.7	2.8	**53.0**	72.7	50.2	19.3	40.9	83.3	62.4	42.4	22.8	**38.5**	49.3	59.5	46.0
MI-max	42.4	46.4	25.0	45.6	45.6	52.6	43.7	**24.0**	45.5	42.4	29.1	5.9	35.5	52.3	55.5	50.0	2.1	15.7	**60.3**	47.9	38.4
DC	**47.1**	53.2	**38.8**	37.0	**46.6**	45.8	52.6	14.5	39.1	48.4	31.7	23.7	34.9	87.0	**67.8**	**54.0**	22.8	23.8	44.9	51.0	43.2
SSD-joint	36.1	57.2	27.6	44.2	29.7	**74.3**	50.0	4.6	52.2	**74.4**	48.9	**24.0**	47.1	87.0	62.8	42.7	29.5	35.7	44.3	59.1	46.5
Ours-type1	41.2	57.3	30.1	44.3	28.1	73.8	53.8	2.5	51.7	73.7	**54.6**	22.5	**55.9**	84.5	65.0	39.7	27.9	36.0	49.5	**62.0**	**47.7**
Ours-type2	43.4	52.5	29.4	40.1	30.4	71.9	**54.9**	3.6	52.4	73.8	53.5	**24.0**	54.8	**89.1**	65.1	40.5	**32.3**	33.8	45.4	61.0	47.6
Ideal case	50.5	60.3	40.1	55.9	34.8	79.7	61.9	13.5	56.2	76.1	57.7	36.8	63.5	92.3	76.2	49.8	40.2	28.1	60.3	74.4	55.4

Table 1. Comparison of all the methods in terms of AP [%] in Clipart1k.

Experiments

浙江大学 信息与电子工程学院
COLLEGE OF INFORMATION SCIENCE & ELECTRONIC ENGINEERING
ZHEJIANG UNIVERSITY

	Pascal VOC→ Watercolor2k							Pascal VOC→ Comic2k						
Model	bike	bird	car	cat	dog	prsn	**mAP**	bike	bird	car	cat	dog	prsn	**mAP**
baseline	79.8	49.5	38.1	35.1	30.4	65.1	49.6	43.9	10.0	19.4	12.9	20.3	42.6	24.9
Compared														
DT+PL	76.5	54.9	46.0	37.4	38.5	72.3	54.3	**55.2**	18.5	38.2	22.9	34.1	54.5	37.2
MI-max	**84.1**	47.4	48.2	30.9	27.9	58.2	49.5	45.3	9.7	33.7	14.4	21.6	37.0	27.0
DC	76.7	53.2	45.3	41.6	35.5	70.0	53.7	52.7	17.4	**43.4**	23.3	25.9	58.7	36.9
SSD-joint	79.0	54.9	47.8	38.2	37.5	72.8	55.0	54.7	16.1	39.6	28.1	35.3	56.4	38.4
Ours-type1	76.4	**56.2**	49.0	41.9	40.4	74.4	56.4	54.4	20.5	41.1	**35.2**	37.1	**61.4**	41.6
Ours-type2	73.7	56.1	**50.6**	**42.5**	**41.8**	**74.6**	**56.5**	55.0	**21.2**	40.0	35.1	**37.9**	60.9	**41.7**
Ideal case	76.0	60.0	52.7	41.0	43.8	77.3	58.4	55.9	26.8	40.4	42.3	43.0	70.1	46.4

Table 2. Comparison in terms of AP [%] in Watercolor2k and Comic2k.

Thank you!

References

[1] Naoto Inoue, Ryosuke Furuta, Toshihiko Yamasaki, and Kiyoharu Aizawa, “Cross-domain weakly-supervised object detection through progressive domain adaptation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 5001–5009.

[2] Xuanyi Dong and Yi Yang, “Teacher supervises students how to learn from partially labeled images for facial landmark detection,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp.783–792.

[3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg, “Ssd: Single shot multibox detector,” in European conference on computer vision. Springer, 2016, pp. 21–37.

[4] Feng Liu, Xiaoxong Zhang, Fang Wan, Xiangyang Ji, and Qixiang Ye, “Domain contrast for domain adaptive object detection,” arXiv preprint arXiv:2006.14863, 2020.

[5] Nicolas Gonthier, Sa¨ıd Ladjal, and Yann Gousseau, “Multiple instance learning on deep features for weakly supervised object detection with extreme domain shifts,” arXiv preprint arXiv:2008.01178, 2020.