Multi-Source Domain Adaptation for Object Detection with Prototype-based Mean Teacher - Supplementary Material

Atif Belal[1], Akhil Meethal[1], Francisco Perdigon Romero[2], Marco Pedersoli[1], Eric Granger[1]
[1] LIVIA, Dept. of Systems Engineering, ETS Montreal, Canada
[2] GAIA Montreal, Ericsson Canada
atif.belal.1@ens.etsmtl.ca, akhilpm135@gmail.com, francisco.perdigon.romero@ericsson.com, {marco.pedersoli, eric.granger}@etsmtl.ca

0.1. Impact of γ on different settings

The main component responsible for simplifying the MSDA problem with our approach is prototype-based learning. In this study, we investigate the importance of contrastive loss on prototypes by studying the impact of γ in different settings. For experiments, we considered the two cases; Cross Time Adaptation and Mixed Domain Adaptation settings. The Cross Time Adaptation setting is much simpler compared to the Mixed Domain Adaptation setting. By choosing these settings, we get an idea of the effect of γ on simple and complex MSDA problems. The analysis result is reported in Table 1. It can be observed that for the Cross Time Adaptation setting, the performance of our method increased with the increase in γ, while it decreased for the Mixed Domain Adaptation setting. With the increase in value of γ, the model is forced to learn a class-conditioned aligned feature space. Learning this feature space is much easier if the domain shift among the datasets is not much. But with the increase in domain shift, learning the aligned feature space becomes difficult.

γ	Source Domain	Target Domain	mAP
0.1	D+N	Dusk/Dawn	43.6
0.5			43.9
0.9			44.6
1.2			45.3
1.5			45.1
0.1	C+M+S	Daytime	39.7
0.5			39.1
0.9			38.7
1.2			38.3
1.5			38.2

Table 1. Effect of γ under different domain shift conditions.

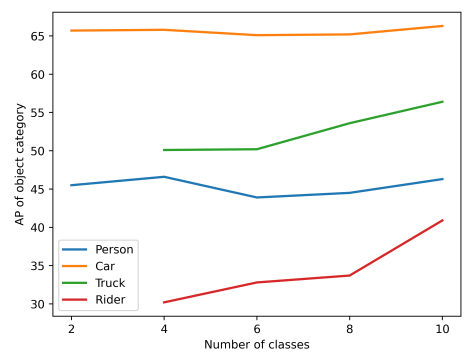

Figure 1. AP of PMT with a growing number of object categories.

0.2. Impact of the Number of Classes

In this ablation, we consider the effect of a different number of object categories for training and evaluation on our PMT. We analyze the performance of a given class when trained with a varying number of other classes and how this affects the final detection performance. As the prototypes are trained with a contrastive loss, we expect that an increased number of classes could help to further improve the overall adaptation. In Fig. 1, we present the AP for the classes Person and Car when trained with 0 to 8 additional classes and Truck and Raider when trained with 2 to 8 extra classes. While for some of the classes (Person and Car), the effect is not clear, for others (Truck and Rider) there is a clear trend in which more classes are used and better the results. We hypothesize that this different behavior could be due to the number of training samples per class. For very common classes with many samples, the model is already quite general, and adding more classes does not help. Instead for classes with fewer data points, contrasting the prototypes with other classes helps to improve detection.

Setting	Source	Method	Bike	Bus	Car	Motor	Person	Rider	Light	Sign	Train	Truck	mAP
	D		35.1	51.7	52.6	9.9	31.9	17.8	21.6	36.3	-	47.1	30.4
Source Only	N	FRCNN [6]	27.9	32.5	49.4	15.0	28.7	21.8	14.0	30.5	-	30.7	25.0
	D+N		31.5	46.9	52.9	8.4	29.5	21.6	21.7	34.3	-	42.2	28.9
		Strong-Weak [7]	29.7	50.0	52.9	11.0	31.4	21.1	23.3	35.1	-	44.9	29.9
		Graph Prototype [1]	31.7	48.8	53.9	20.8	32.0	21.6	20.5	33.7	-	43.1	30.6
UDA Blending	D+N	Cat. Regularization [9]	25.3	51.3	52.1	17.0	33.4	18.9	20.7	34.8	-	47.9	30.2
		UMT [2]	42.3	48.1	56.4	13.5	35.3	26.9	31.1	41.7	-	40.1	33.5
		Adaptive Teacher [3]	43.1	48.9	56.9	14.7	36.0	27.1	32.7	43.8	-	42.7	34.6
		MDAN [11]	37.1	29.9	52.8	15.8	35.1	21.6	24.7	38.8	-	20.1	27.6
		M^3SDA [5]	36.9	25.9	51.9	15.1	35.7	20.5	24.7	38.1	-	15.9	26.5
MSDA	D+N	DMSN [10]	36.5	54.3	55.5	20.4	36.9	27.7	26.4	41.6	-	50.8	35.0
		TRKP [8]	48.4	56.3	61.4	22.5	41.5	27.0	41.1	47.9	-	51.9	39.8
		PMT(ours)	**55.3**	**59.8**	**67.6**	**29.9**	**47.6**	**32.7**	**46.3**	**56.0**	-	**57.7**	**45.3**
Oracle	D+N	Target Only	27.2	39.6	51.9	12.7	29.0	15.2	20.0	33.1	-	37.5	26.6
		All-Combined	56.4	59.9	67.3	30.8	47.9	33.9	47.2	57.8	-	54.8	45.3

Table 2. Class-wise AP of PMT compared against the baseline, UDA, MSDA, and oracle methods on BDD100K. Source domains are daytime (D) and night (N) subsets and the target is always Dusk/Dawn of BDD100K

Setting	Source	Method	Person	Car	Rider	Truck	Motor	Bicycle	Bike	mAP
Source Only	C	FRCNN [6]	26.9	44.7	22.1	17.4	17.1	18.8	16.7	23.4
Source Only		FRCNN [6]	35.2	49.5	26.1	25.8	18.9	26.1	26.5	29.7
UDA Blending		Unbiased Teach. [4]	30.7	28.0	3.9	11.2	19.2	17.8	18.7	18.5
UDA Blending	C+M	Adaptive Teacher [3]	31.2	31.7	15.1	16.4	17.1	20.9	27.9	22.9
MSDA		TRKP [8]	39.2	53.2	**32.4**	28.7	25.5	31.1	37.4	35.3
MSDA		**PMT(ours)**	**41.1**	**53.5**	31.2	**31.9**	**33.7**	**34.9**	**44.6**	**38.7**
Source Only		FRCNN [6]	36.6	49.0	22.8	24.9	26.9	28.4	27.7	30.9
UDA Blending		Unbiased Teach. [4]	32.7	39.6	6.6	21.2	21.3	25.7	28.5	25.1
UDA Blending	C+M+S	Adaptive Teacher [3]	36.3	42.6	19.7	23.4	24.8	27.1	33.2	29.6
MSDA		TRKP [8]	40.2	53.9	31.0	30.8	30.4	34.0	39.3	37.1
MSDA		**PMT(ours)**	**43.3**	**54.1**	**32.0**	**32.6**	**35.1**	**36.1**	**44.8**	**39.7**
	C+M	Target Only	35.3	53.9	33.2	46.3	25.6	29.3	46.7	38.6
Oracle	C+M	All-Combined	40.2	60.1	47.1	60.0	29.2	36.3	56.9	47.1
	C+M+S	All-Combined	41.7	63.9	49.5	58.1	31.6	39.1	53.5	48.2

Table 3. Class-wise AP of PMT compared against the baselines on Daytime domain of BDD100K. C, M, and S refer to Cityscapes, MS COCO, and Synscapes datasets.

References

[1] Qi Cai, Yingwei Pan, Chong-Wah Ngo, Xinmei Tian, Lingyu Duan, and Ting Yao. Exploring object relation in mean teacher for cross-domain detection, 2019. 2

[2] Jinhong Deng, Wen Li, Yuhua Chen, and Lixin Duan. Unbiased mean teacher for cross-domain object detection, 2021. 2

[3] Yu-Jhe Li, Xiaoliang Dai, Chih-Yao Ma, Yen-Cheng Liu, Kan Chen, Bichen Wu, Zijian He, Kris Kitani, and Peter Vajda. Cross-domain adaptive teacher for object detection. In *CVPR*, 2022. 2

[4] Yen-Cheng Liu, Chih-Yao Ma, Zijian He, Chia-Wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira, and Peter Vajda. Unbiased teacher for semi-supervised object detection, 2021. 2

[5] Xingchao Peng, Qinxun Bai, Xide Xia, Zijun Huang, Kate Saenko, and Bo Wang. Moment matching for multi-source domain adaptation, 2019. 2

[6] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In *NeurIPS*, 2015. 2

[7] Kuniaki Saito, Yoshitaka Ushiku, Tatsuya Harada, and Kate Saenko. Strong-weak distribution alignment for adaptive object detection, 2019. 2

[8] Jiaxi Wu, Jiaxin Chen, Mengzhe He, Yiru Wang, Bo Li, Bingqi Ma, Weihao Gan, Wei Wu, Yali Wang, and Di Huang. Target-relevant knowledge preservation for multi-source domain adaptive object detection, 2022. 2

[9] Chang-Dong Xu, Xing-Ran Zhao, Xin Jin, and Xiu-Shen Wei. Exploring categorical regularization for domain adaptive object detection, 2020. 2

[10] Xingxu Yao, Sicheng Zhao, Pengfei Xu, and Jufeng Yang. Multi-source domain adaptation for object detection, 2021. 2

[11] Han Zhao, Shanghang Zhang, Guanhang Wu, José M. F. Moura, Joao P Costeira, and Geoffrey J Gordon. Adversarial multiple source domain adaptation. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. 2