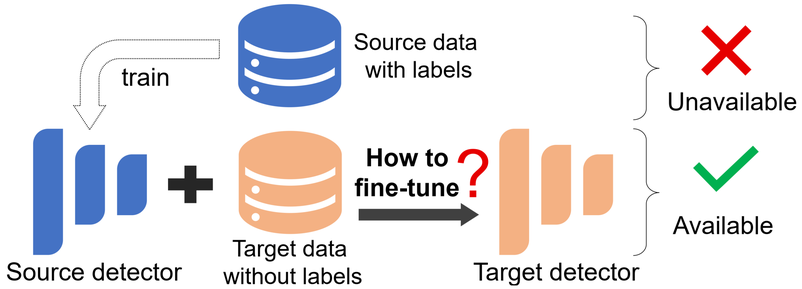
Source-Free Object Detection by Learning to Overlook Domain Style

Shuaifeng Li[1], Mao Ye[1], Xiatian Zhu[2], Lihua Zhou[1], Lin Xiong[1]

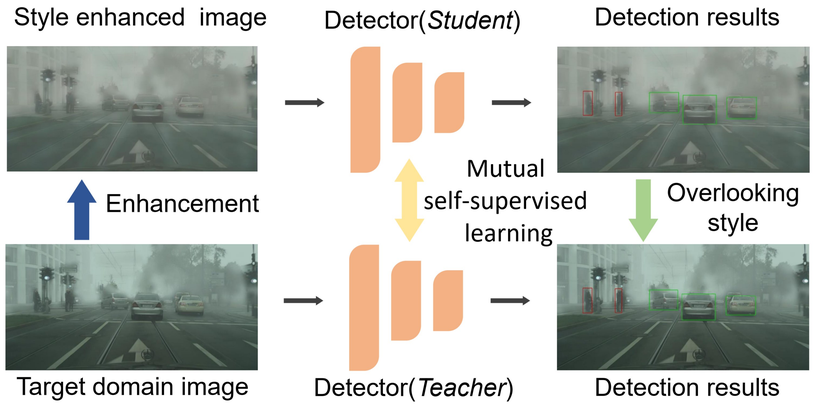
[1]University of Electronic Science and Technology of China & [2]University of Surrey

What is Source-Free Object Detection?



Overlook domain style strategy

1. ***Enhance*** the target domain style for each target image
2. ***Overlook*** the enhanced style by directional alignments from student to teacher

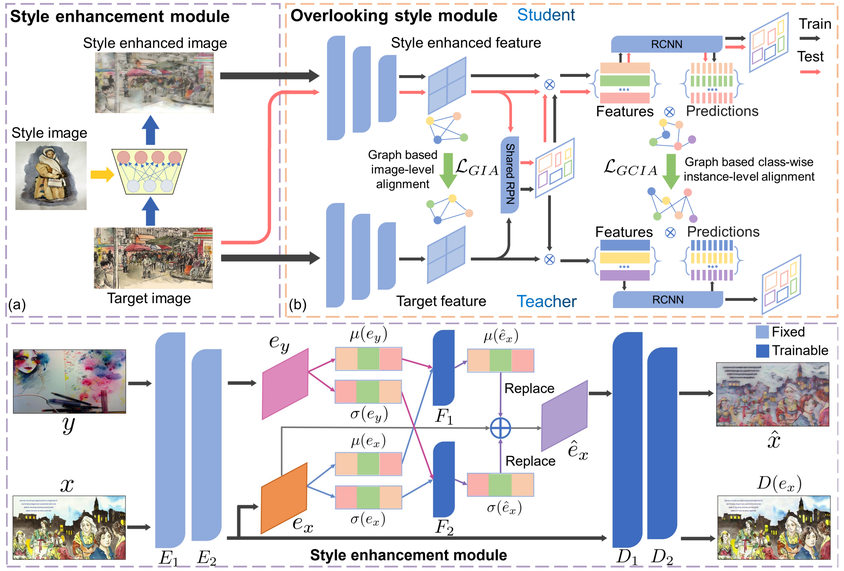


Our contributions:

- Propose a novel learning to overlook domain style strategy.
- Propose a style enhancement method.
- Propose a new Mean-Teacher framework variant which achieves a two-way knowledge distillation.

Learning to overlook domain style method

1. ***Style enhancement module*** adds the style of the style image (a random target image) onto another target image
2. ***Overlooking style module*** forces both detectors to overlook target domain style by two directional alignments



Quantitative results

LODS achieves best performance on Pascal to Watercolor

Methods	Bike	Bird	Car	Cat	Dog	Person	mAP
Source Only	85.6	46.8	43.1	24.5	21.9	54.8	46.1
SWDA [28]	82.3	55.9	46.5	32.7	35.5	66.7	53.3
DBGL [3]	83.1	49.3	50.6	39.8	38.7	61.3	53.8
SAPNet [20]	81.1	51.1	**53.6**	34.3	39.8	**71.3**	55.2
VDD [34]	90.0	56.6	49.2	39.5	38.8	65.3	56.6
SOAP [36]	79.3	44.3	41.4	**45.7**	39.3	55.9	51.0
Our method	95.2	53.1	46.9	37.2	**47.6**	69.3	**58.2**

LODS achieves best performance on Cityscapes to Foggy-Cityscapes

Methods	Person	Rider	Car	Tuck	Bus	Train	Mcycl	Bicycle	mAP
Source Only	25.8	33.3	35.2	13.0	26.4	9.1	19.0	32.3	24.3
DA-Faster [5]	25.0	31.0	40.5	22.1	35.3	20.2	20.0	27.1	27.6
SWDA [28]	29.9	42.3	43.5	24.5	36.2	32.6	35.3	30.0	34.3
MTOR [2]	30.6	41.4	44.0	21.9	38.6	**40.6**	28.3	35.6	35.1
iFAN [45]	32.6	40.0	48.5	27.9	**45.5**	31.7	22.8	33.0	35.3
SED [23]	21.7	44.0	40.4	32.6	11.8	25.3	34.5	34.3	30.6
SED(Mosaic) [23]	25.5	44.5	40.7	**33.2**	22.2	28.4	34.1	39.0	33.5
HCL [13]	26.9	**46.0**	41.3	33.0	25.0	28.1	**35.9**	**40.7**	34.6
SOAP [36]	**35.9**	45.0	48.4	23.9	37.2	24.3	31.8	37.9	35.5
Our method	34.0	45.7	**48.8**	27.3	39.7	19.6	33.2	37.8	**35.8**

Detection results

Validate the overlooking domain style ability

The SE, RT and RSE images stand for style enhanced image, the reconstructed image with target feature and the reconstructed image with style enhanced feature

Target image vs RT image and SE image vs RSE image

Paper and code are available: https://github.com/Flashkong/Source-Free-Object-Detection-by-Learning-to-Overlook-Domain-Style