(CVPR2021)

FSDR: Frequency Space Domain Randomization for Domain Generalization

Jiaxing Huang, Dayan Guan, Aoran Xiao, Shijian Lu*
School of Computer Science Engineering, Nanyang Technological University
{Jiaxing.Huang, Dayan.Guan, Aoran.Xiao, Shijian.Lu}@ntu.edu.sg

Introduction

- Limitation of Existing DR Methods

  Most existing DR methods randomize the whole spectrum of images in the spatial space which tends to modify domain invariant features undesirably

- FSDR

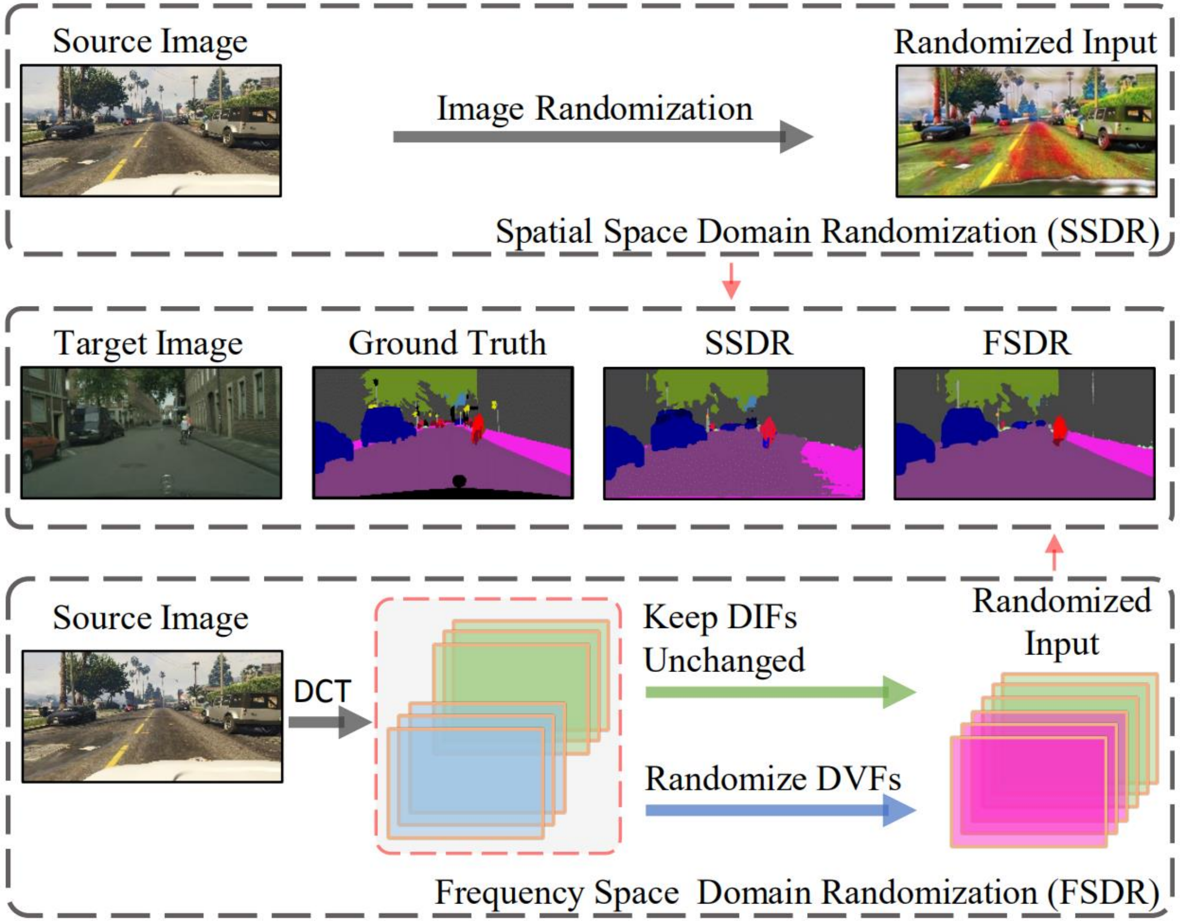
  Transforms images into frequency space and performs domain generalization by identifying and randomizing domain-variant frequency components (DVFs) while keeping domain-invariant frequency components (DIFs) unchanged

- Contribution

  - Proposed an innovative frequency space domain randomization technique that transforms images into frequency space and achieves domain randomization by changing DVFs only while keeping DIFs unchanged.
  - Designed two randomization approaches in the frequency space that identify DVFs and DIFs effectively through empirical experiments and dynamic learning, respectively.
  - Experiments over multiple domain generalization tasks show that our proposed frequency space domain randomization technique achieves superior semantic segmentation consistently.

Introduction

- FSDR



Methods

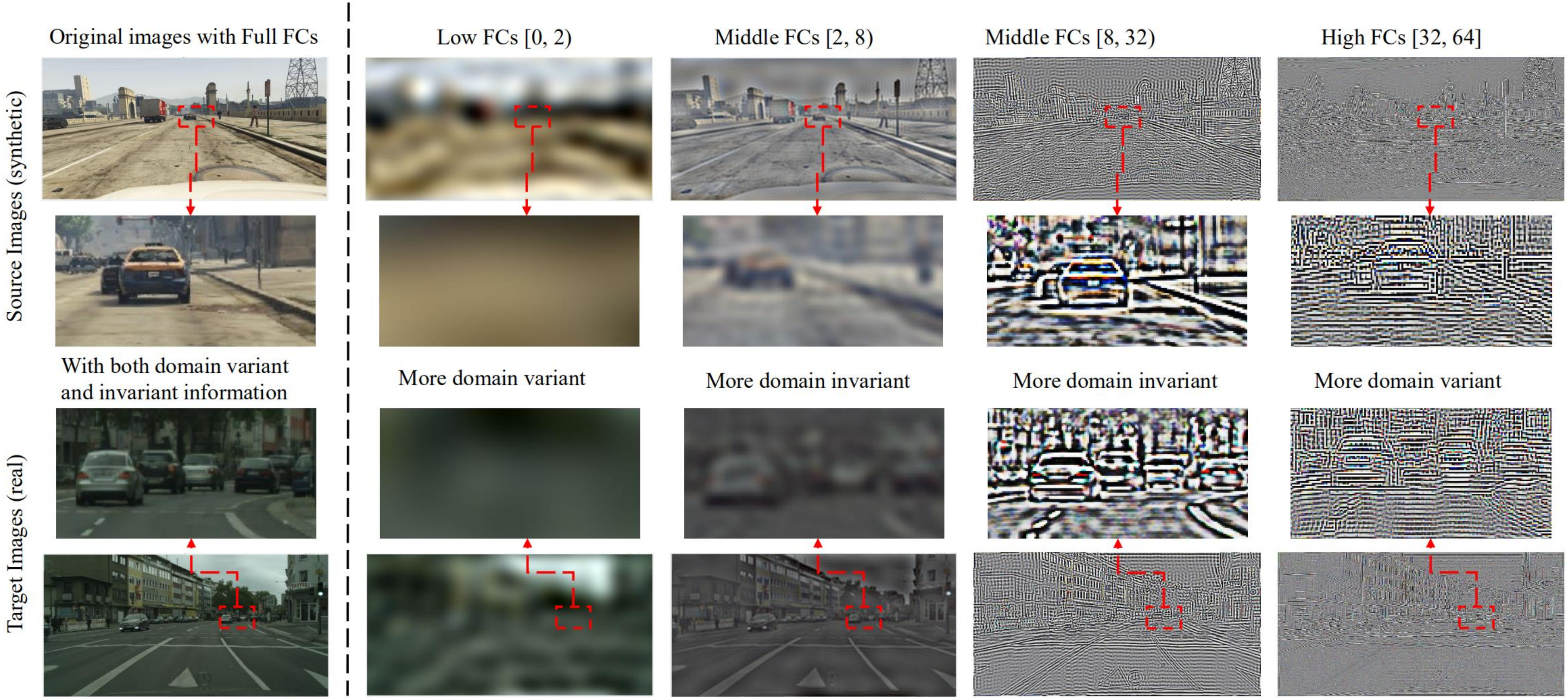
- Spectrum Analysis

$$f_s = \mathcal{B}_p(\mathcal{D}(x_s)) = \{f_s(0), f_s(1), f_s(3), ..., f_s(63)\}$$

Band-reject Spectrum analysis		
Rejected band	Source acc.	Target acc.
Null (with full FCs)	95.5	65.2
[0, 1)	95.1	68.6
[1, 2)	95.3	67.1
[2, 4)	95.4	62.3
[4, 8)	95.4	62.7
[8, 16)	95.5	64.6
[16, 32)	95.6	64.9
[32, 64]	95.9	67.4

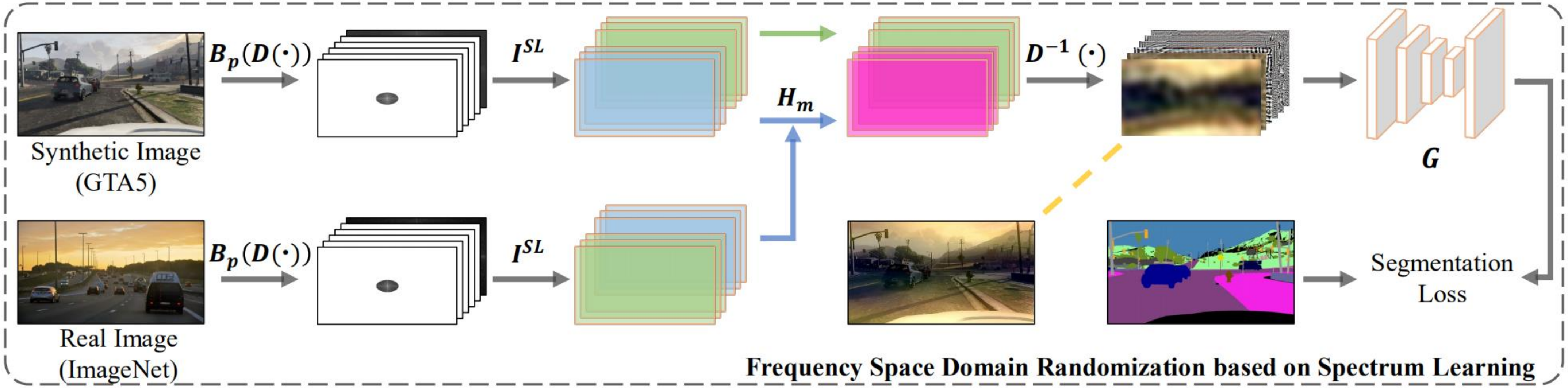
Methods

- Spectrum Analysis



Methods

- FSDR-SA

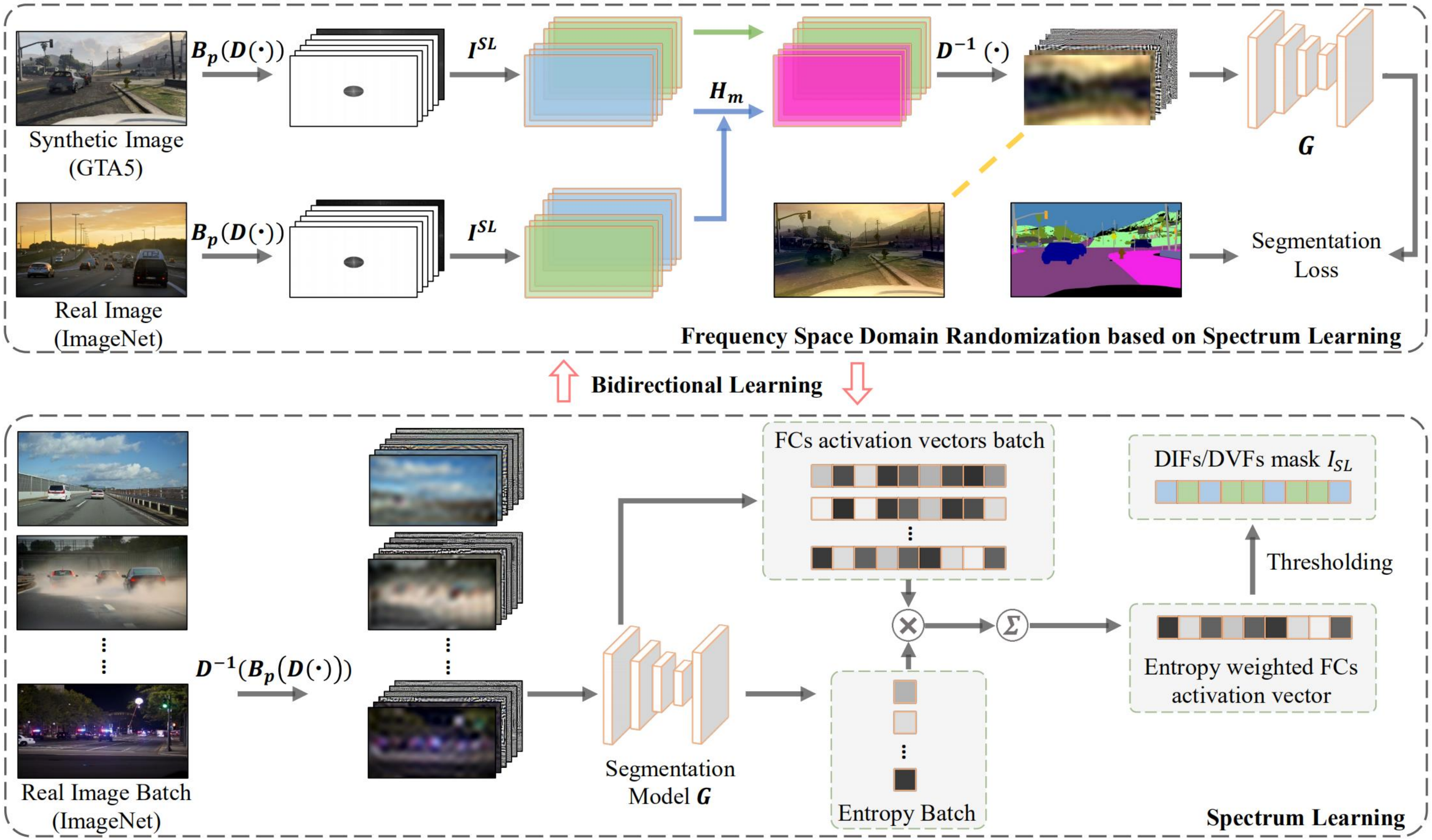


$$\mathcal{R}^{SA}(x_s; x_{img}) = \mathcal{D}^{-1}(\mathcal{H}_m(f_s(1-I^{SA}), f_{img}(1-I^{SA})))$$

$$\mathcal{L}_{SA} = l(G(\mathcal{R}^{SA}(x_s; x_{img})), y_s)$$

Methods

- FSDR-SL



Experiments

- Ablation Study

Table 2. Ablation study for the domain generalization task GTA $\rightarrow$ {Cityscapes, Mapillary and BDD} (using ResNet-101 as backbone) in mIoU. Losses L_{orig}, L_{SA}, and L_{SL} are defined in Eq. 1, 5, and 7, respectively.

Method	$\mathcal{L}_{orig}$	$\mathcal{L}_{SA}$	$\mathcal{L}_{SL}$	**mIoU**		
				City.	Mapi.	BDD
Baseline	✓			33.4	27.9	27.3
FSDR-SA	✓	✓		42.1	39.2	37.8
FSDR-SL	✓		✓	43.6	42.1	40.1
FSDR	✓	✓	✓	44.8	43.4	41.2

Experiments

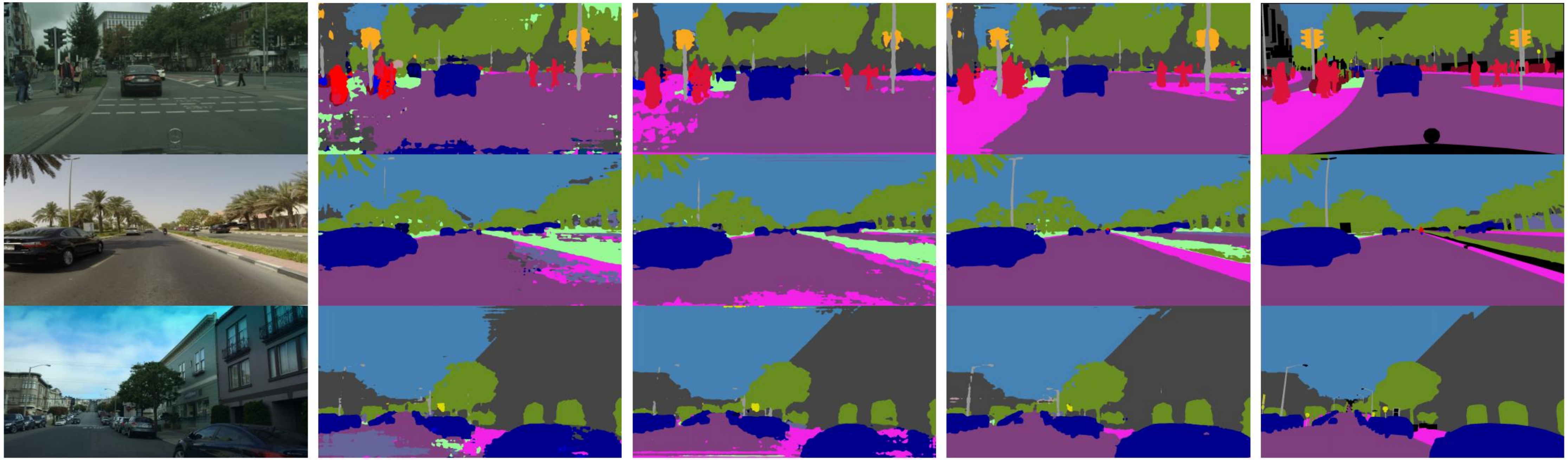
- Comparison with SOTA

Net.	Method	w/ Tgt	C	M	B	Mean
ResNet101	CBST [88]	✓	44.9	40.3	40.5	41.9
	AdaSeg. [66]	✓	41.4	38.3	36.2	38.6
	MinEnt [72]	✓	42.3	38.5	34.4	38.4
	FDA [76]	✓	45.0	39.6	38.1	40.9
	IDA [46]	✓	**45.1**	39.4	37.6	40.7
	CrCDA [26]	✓	43.7	39.3	37.3	40.1
	IBN-Net [47]	✗	40.3	35.9	35.6	37.2
	DRPC [77]	✗	42.5	38.0	38.7	39.8
	Ours (FSDR)	✗	44.8	**43.4**	**41.2**	**43.1**
VGG16	CBST [88]	✓	38.1	34.6	33.9	35.5
	AdaSeg. [66]	✓	35.0	32.6	31.3	33.0
	MinEnt [72]	✓	32.8	30.7	29.5	31.0
	FDA [76]	✓	37.9	33.8	32.1	34.6
	IDA [46]	✓	**38.5**	34.2	32.3	35.0
	CrCDA [26]	✓	36.1	32.6	31.8	33.5
	IBN-Net [47]	✗	34.8	31.0	30.4	32.0
	DRPC [77]	✗	36.1	32.3	31.6	33.3
	Ours (FSDR)	✗	38.3	**37.6**	**34.4**	**37.1**

Net.	Method	w/ Tgt	C	M	B	Mean
ResNet101	CBST [88]	✓	41.4	37.1	**37.6**	38.7
	AdaSeg. [66]	✓	38.2	36.1	35.3	36.5
	MinEnt [72]	✓	38.1	35.8	35.5	36.4
	FDA [76]	✓	41.2	36.1	36.4	37.9
	IDA [46]	✓	**41.7**	36.5	37.0	38.4
	CrCDA [26]	✓	39.0	36.4	36.7	37.4
	IBN-Net [47]	✗	37.5	33.7	33.0	34.7
	DRPC [77]	✗	37.6	34.1	34.3	35.4
	Ours (FSDR)	✗	40.8	**39.6**	37.4	**39.3**
VGG16	CBST [88]	✓	38.2	33.5	32.2	34.6
	AdaSeg. [66]	✓	32.6	30.3	29.4	30.8
	MinEnt [72]	✓	31.4	29.8	28.9	30.0
	FDA [76]	✓	37.9	33.1	31.8	34.2
	IDA [46]	✓	**38.6**	34.2	32.7	35.1
	CrCDA [26]	✓	36.4	32.8	31.9	33.7
	IBN-Net [47]	✗	33.9	31.1	30.4	31.8
	DRPC [77]	✗	35.5	32.2	29.5	32.4
	Ours (FSDR)	✗	37.9	**37.2**	**34.1**	**36.4**

Experiments

- Comparison with SOTA



Target Image Baseline DRPC **Ours (FSDR)** Ground Truth

Figure 4. Qualitative illustration of domain generalizable semantic segmentation for GTA5 to Cityscapes (1st row), Mapillary (2nd row), and BDD (3rd row). FSDR preserves domain invariant information during domain randomization which produces better semantic segmentation especially around edge/class-transition area. As a comparison, DRPC [77] does not isolate and preserve domain invariant feature which leads to sub-optimal segmentation.

Experiments

- Discussion

Table 5. FSDR is complementary with existing domain adaptation and generalization methods. For the task GTA $\rightarrow$ {Cityscapes, Mapillary, and BDD}, including FSDR (Ours*) improves the domain adaptation and generalization performance consistently. "w/ Tgt" labels methods that train models with (✓ i.e. domain adaptation) or without (✗ *i.e.*, domain generalization) accessing target data in Cityscapes.

	w/ Tgt	Cityscapes		Mapillary		BDD	
		Base	+Ours*	Base	+Ours*	Base	+Ours*
Adapt-SegMap [66]	✓	41.4	**45.6**	38.3	**43.9**	36.2	**41.9**
MinEnt [72]	✓	42.3	**45.7**	38.5	**43.7**	34.4	**41.7**
CBST [88]	✓	44.9	**46.8**	40.3	**44.3**	40.5	**42.8**
FDA [76]	✓	45.0	**46.1**	39.6	**44.2**	38.1	**42.1**
IBN-Net [47]	✗	40.3	**45.3**	35.9	**44.0**	35.6	**42.1**
DRPC [77]	✗	42.5	**45.8**	38.0	**44.2**	38.7	**42.6**

Table 6. The sensitivity of parameter p affects domain generalization: For the task GTA $\rightarrow$ Cityscapes, the domain generalization performance varies with p as evaluated in mIoU.

	Proportion of preserved FCs						
Method	1	5/6	4/6	3/6	2/6	1/6	0
FSDR	33.4	43.8	44.3	44.8	44.2	41.0	38.3

Experiments

- Discussion

Table 7. FSDR is generic and can work for other tasks like object detection: The detection task is SYNTHIA → {Cityscapes, Mapillary, and BDDS} as evaluated using metric mAP. "w/ Tgt" labels methods that train models with (✓ *i.e.*, domain adaptation) and without (✗ *i.e.*, domain generalization) accessing target-domain data in Cityscapes.

Net.	Method	w/ Tgt	C	M	B	Mean
ResNet101	Faster-RCNN	✗	24.3	20.8	20.1	21.7
	DA [6]	✓	30.2	21.2	21.8	24.4
	MinEnt [72]	✓	30.2	21.7	22.4	24.7
	CBST [88]	✓	32.7	23.5	23.9	26.7
	FDA [76]	✓	32.4	23.3	23.8	26.5
	IBN-Net [47]	✗	30.1	22.3	23.1	25.1
	Ours (FSDR)	✗	**33.5**	**24.9**	**25.2**	**27.8**

(CVPR2021)

A Fourier-based Framework for Domain Generalization

Qinwei Xu[1] Ruipeng Zhang[1] Ya Zhang[1,2] ✉ Yanfeng Wang[1,2] Qi Tian[3]

[1] Cooperative Medianet Innovation Center, Shanghai Jiao Tong University

[2] Shanghai AI Laboratory [3] Huawei Cloud & AI

{qinweixu, zhangrp, ya_zhang, wangyanfeng}@sjtu.edu.cn, tian.qi1@huawei.com

Introduction

- Property of Fourier Transform

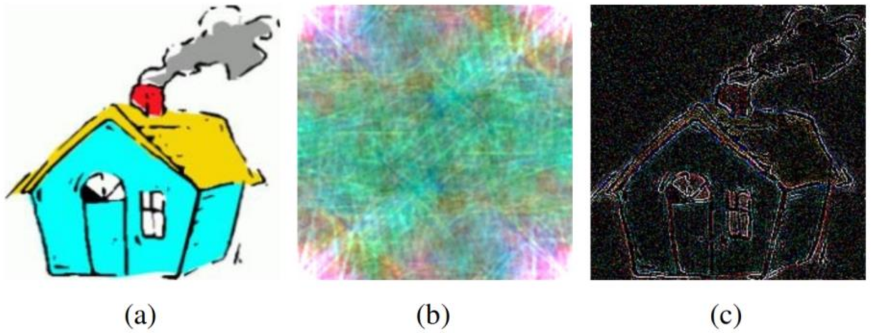
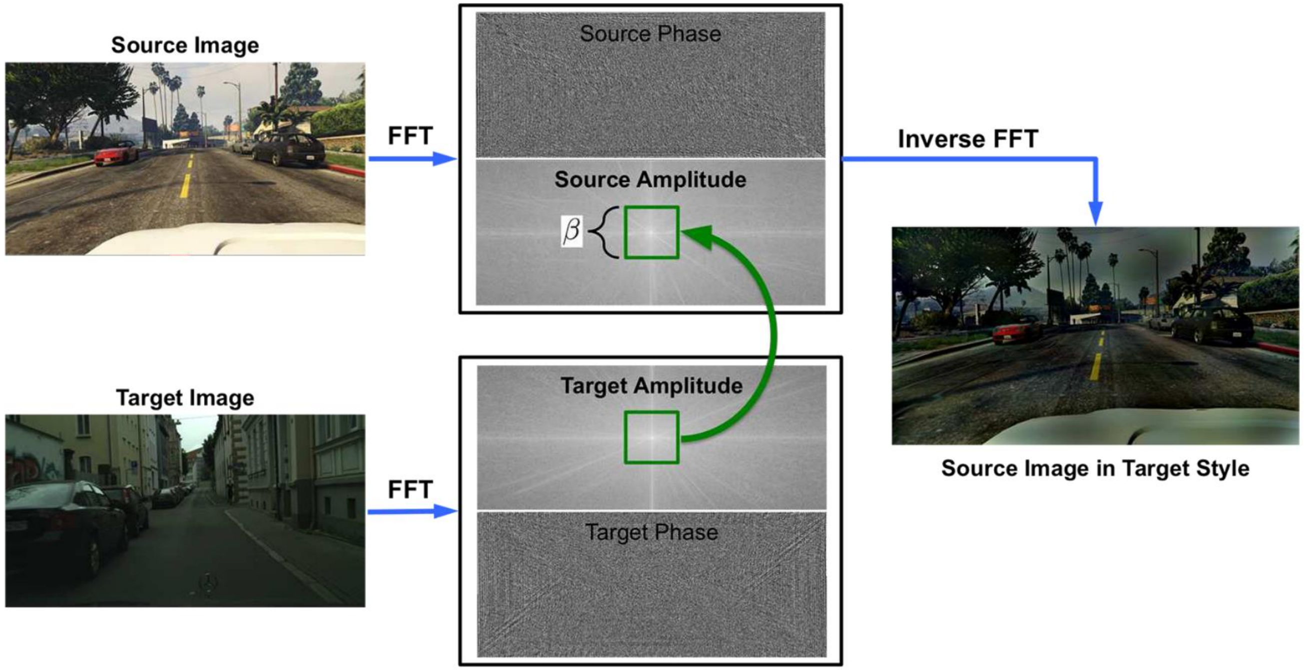


Figure 1. Examples of the amplitude-only and phase-only reconstruction: (a) original image; (b) reconstructed image with amplitude information only by setting the phase component to a constant; (c) reconstructed image with phase information only by setting the amplitude component to a constant.



* FDA: Fourier Domain Adaptation for Semantic Segmentation.

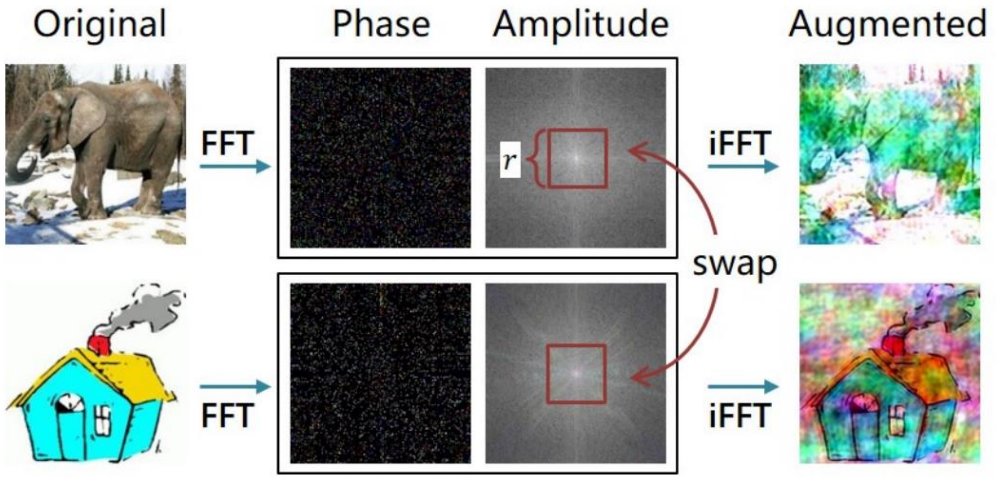
- Fourier-based Augmented Co-Teacher (FACT)

  Models highlight the phase information have better generalization ability across domains

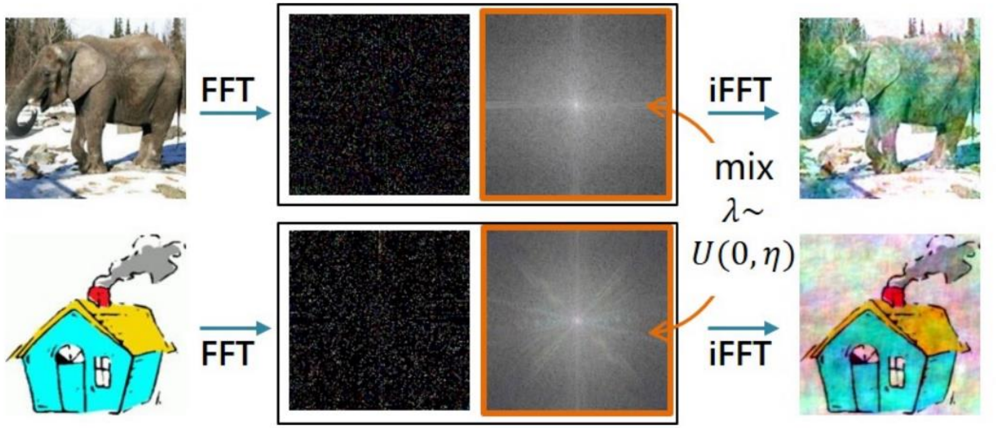
  - Fourier-based data augmentation
  - Co-teacher regularization

Methods

- Fourier-based data augmentation



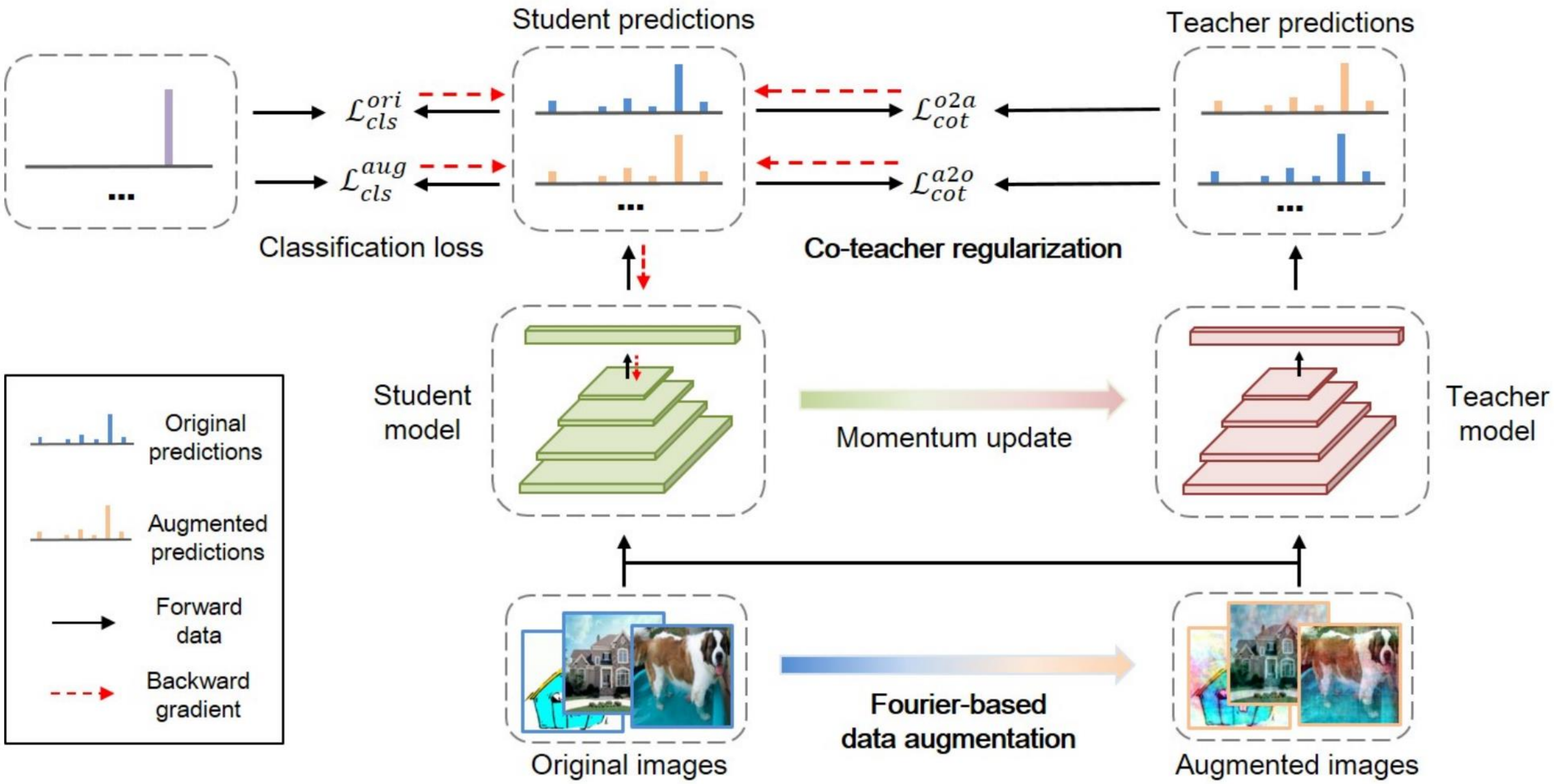
(a)



(b)

Methods

- Fourier-based Augmented Co-Teacher (FACT)



Experiments

- Comparison with SOTA

Methods	MNIST	MNIST-M	SVHN	SYN	Avg.
DeepAll [50]	95.8	58.8	61.7	78.6	73.7
CCSA [28]	95.2	58.2	65.5	79.1	74.5
MMD-AAE [23]	96.5	58.4	65.0	78.4	74.6
CrossGrad [37]	96.7	61.1	65.3	80.2	75.8
DDAIG [50]	96.6	<u>64.1</u>	68.6	81.0	77.6
Jigen [2]	96.5	61.4	63.7	74.0	73.9
L2A-OT [51]	<u>96.7</u>	63.9	<u>68.6</u>	<u>83.2</u>	<u>78.1</u>
FACT (*ours*)	**97.9**	**65.6**	**72.4**	**90.3**	**81.5**

Methods	Art	Clipart	Product	Real	Avg.
DeepAll	57.88	<u>52.72</u>	73.50	74.80	64.72
CCSA [28]	59.90	49.90	74.10	75.70	64.90
MMD-AAE [23]	56.50	47.30	72.10	74.80	62.70
CrossGrad [37]	58.40	49.40	73.90	75.80	64.40
DDAIG [50]	59.20	52.30	<u>74.60</u>	76.00	65.50
L2A-OT [51]	**60.60**	50.10	**74.80**	**77.00**	<u>65.60</u>
Jigen [2]	53.04	47.51	71.47	72.79	61.20
RSC [16]	58.42	47.90	71.63	74.54	63.12
Jigen (*our imple.*)	57.95	49.21	72.61	74.90	63.67
RSC (*our imple.*)	57.67	48.48	72.62	74.16	63.23
FACT (*ours*)	<u>60.34</u>	**54.85**	74.48	<u>76.55</u>	**66.56**

Methods	Art	Cartoon	Photo	Sketch	Avg.
ResNet18					
DeepAll	77.63	76.77	95.85	69.50	79.94
MetaReg [1]	83.70	77.20	95.50	70.30	81.70
JiGen [2]	79.42	75.25	96.03	71.35	80.51
Epi-FCR [22]	82.10	77.00	93.90	73.00	81.50
MMLD [26]	81.28	77.16	96.09	72.29	81.83
DDAIG [50]	<u>84.20</u>	78.10	95.30	74.70	<u>83.10</u>
CSD [36]	78.90	75.80	94.10	<u>76.70</u>	81.40
InfoDrop [39]	80.27	76.54	<u>96.11</u>	76.38	82.33
MASF [4]†	80.29	77.17	94.99	71.69	81.04
L2A-OT [51]	83.30	78.20	**96.20**	73.60	82.80
EISNet [45]	81.89	76.44	95.93	74.33	82.15
RSC [16]	83.43	80.31	95.99	80.85	85.15
RSC (*our imple.*)	80.55	**78.60**	94.43	76.02	82.40
FACT (*ours*)	**85.37**	<u>78.38</u>	95.15	**79.15**	**84.51**
ResNet50					
DeepAll	84.94	76.98	**97.64**	76.75	84.08
MetaReg [1]	<u>87.20</u>	79.20	<u>97.60</u>	70.30	83.60
MASF [4]†	82.89	80.49	95.01	72.29	82.67
EISNet [45]	86.64	<u>81.53</u>	97.11	78.07	<u>85.84</u>
RSC [16]	87.89	82.16	97.92	83.35	87.83
RSC (*our imple.*)	83.92	79.52	95.15	<u>82.20</u>	85.20
FACT (*ours*)	**89.63**	**81.77**	96.75	**84.46**	**88.15**

Experiments

- Ablation Study

Method	AM	$\mathcal{L}_{cot}^{a2o}$	$\mathcal{L}_{cot}^{o2a}$	Teacher	Art	Cartoon	Photo	Sketch	Avg.
Baseline	-	-	-	-	77.63±0.84	76.77±0.33	**95.85±0.20**	69.50±1.26	79.94
Model A	✓	-	-	-	83.90±0.50	76.95±0.45	95.55±0.12	77.36±0.71	83.44
Model B	✓	✓	✓	-	83.71±0.30	77.84±0.49	94.73±0.12	78.55±0.46	83.71
Model C	-	✓	✓	✓	82.68±0.44	78.06±0.39	95.35±0.44	74.76±0.67	82.71
Model D	✓	✓	-	✓	83.97±0.77	77.04±0.86	94.59±0.03	79.08±0.56	83.67
Model E	✓	-	✓	✓	84.07±0.43	77.70±0.65	95.28±0.34	78.29±0.61	83.84
FACT	✓	✓	✓	✓	**85.37±0.29**	**78.38±0.29**	95.15±0.26	**79.15±0.69**	**84.51**

Experiments

- Ablation Study

Table 5. Ablation studies of different choices of the Fourier data augmentation on the PACS dataset with ResNet18.

Methods	Art	Cartoon	Photo	Sketch	Avg.
		DeepAll with			
AS-partial	82.00	76.19	93.89	77.27	82.34
AS-full	83.50	76.07	94.49	77.13	82.80
AM	**83.90**	**76.95**	**95.55**	**77.36**	**83.44**
		FACT with			
AS-partial	81.61	76.95	93.83	78.30	82.67
AS-full	83.46	77.37	94.10	78.63	83.39
AM	**85.37**	**78.38**	**95.15**	**79.15**	**84.51**