Modality-Agnostic Debiasing for Single Domain Generalization

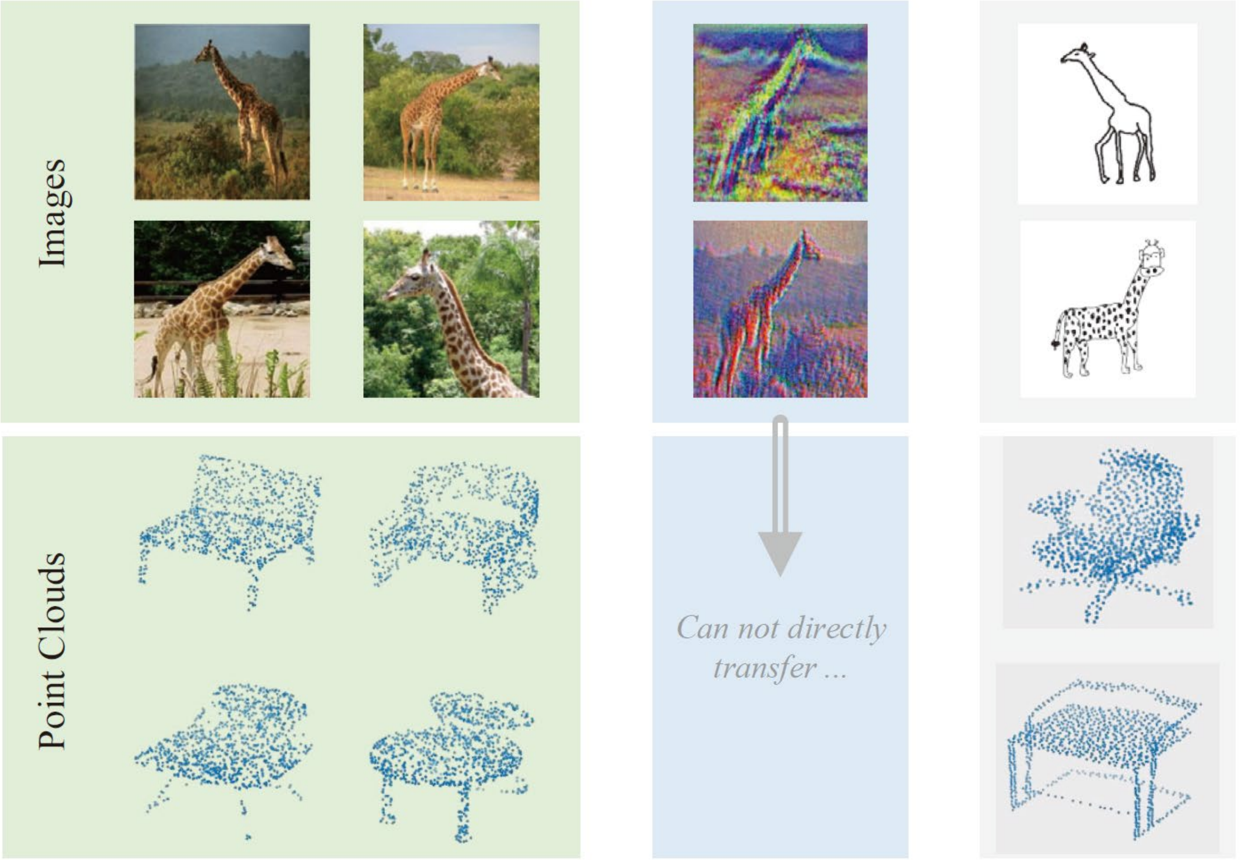
Sanqing Qu, Yingwei Pan, Guang Chen, Ting Yao, Changjun Jiang, Tao Mei

1Tongji Univesity, 2 HiDream.ai Inc.

TONGJI UNIVERSITY

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Introduction



Motivation:

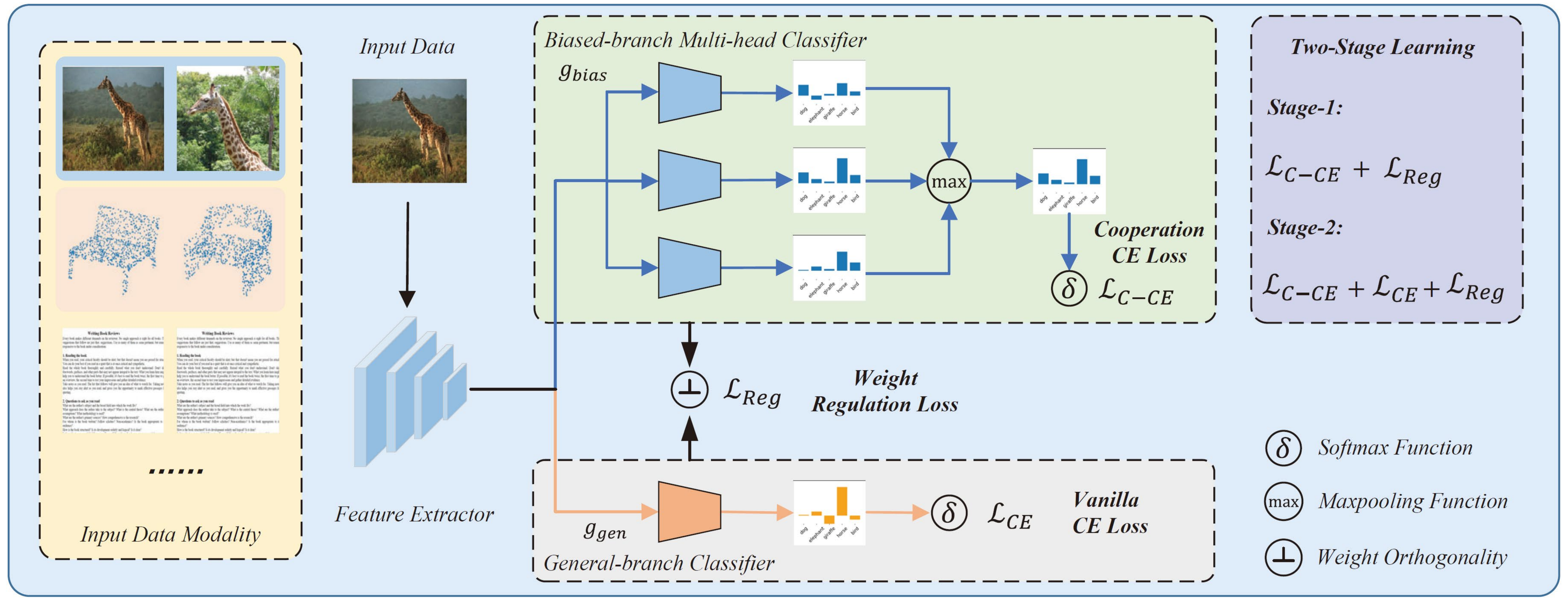
- Deep Neural Networks (DNNs) fail to generalize well to outside of distribution (OOD) data due to the notorious short-cut learning.
- Existing single domain generalization (single-DG) commonly devise various data-augmentation algorithms, which are typically modality-specific.

Contribution:

- We target for a versatile Modality-Agnostic Debiasing (MAD) framework for single-DG, that enables generalization for different modalities.
- MAD is pluggable to most single-DG models.

Methodology

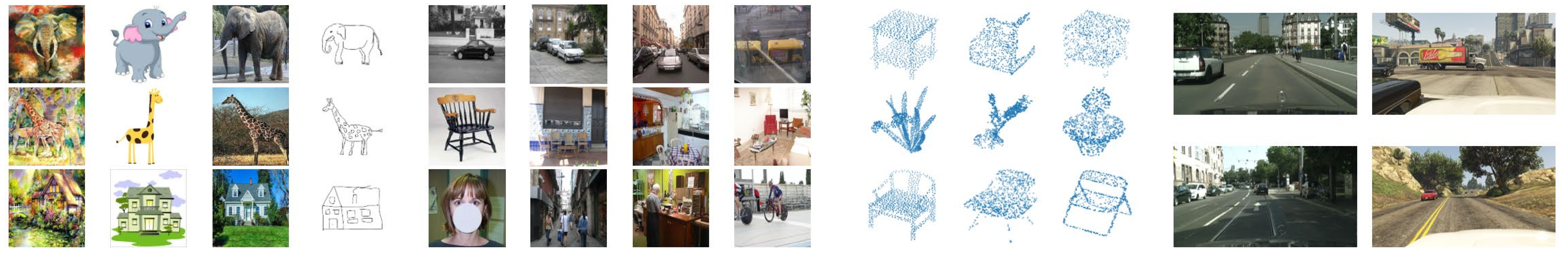
- Introducing a novel two-branch classifier, a biased-branch encourages the classifier to identify domain-specific (superficial) features, a general-branch captures domain-generalized (semantic) features.



Experiments

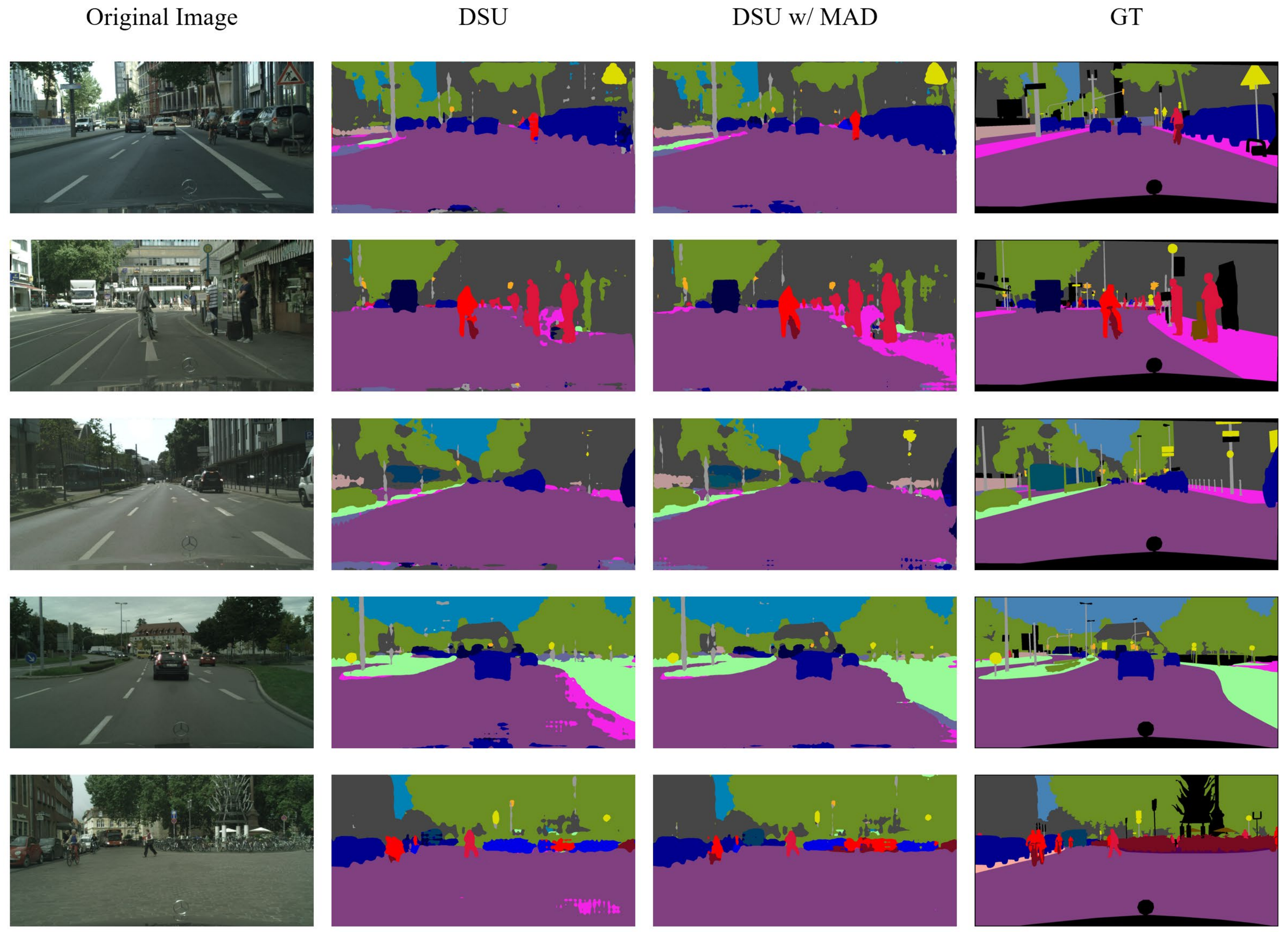
Datasets Configuration:

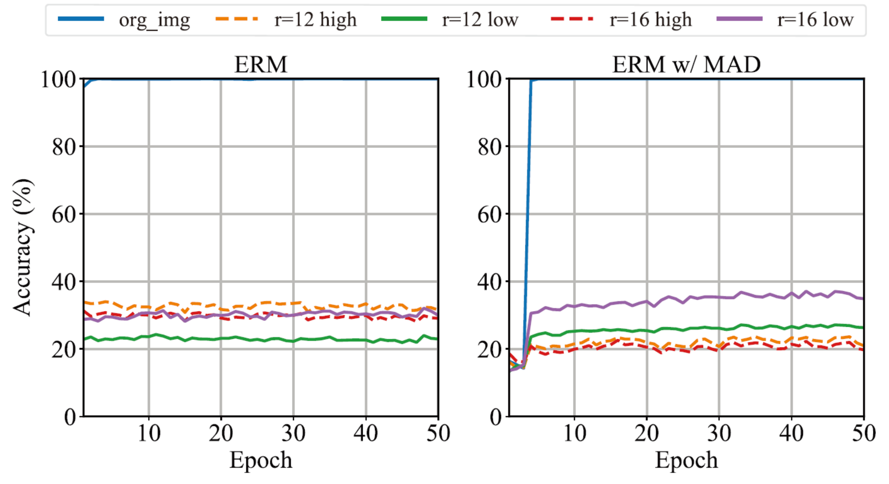
We validate the superiority of MAD in a variety of single-DG scenarios with different modalities, including recognition on 1D texts, 2D images, 3D point clouds, and semantic segmentation on 2D images. We inject our MAD to several baseline methods, including *ERM*, *pAdaIn*, *Mixstyle*, *DSU*, *etc.*



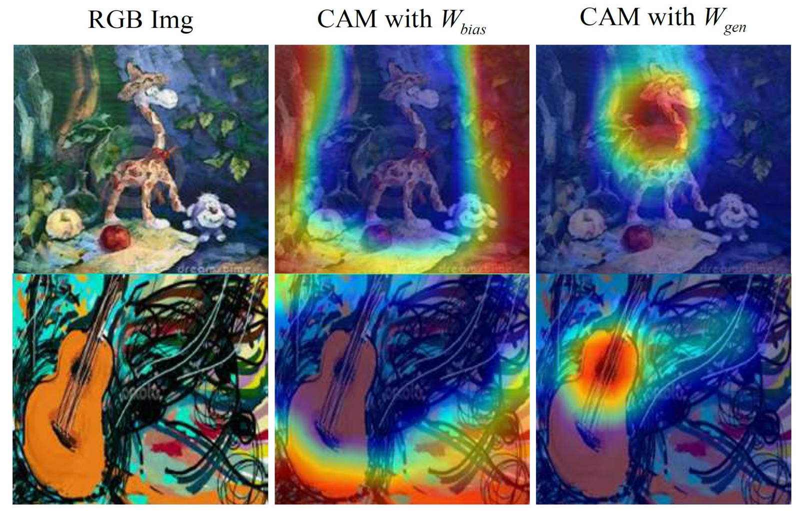
(a) PACS (b) VLCS (c) PointDA-10 (d) GTA-5 → Cityscapes

Qualitative Results:





(a) LFC vs. HFC in "Photo" domain on PACS



Quantitively Results:

Single-DG results on 1D texts (Amazon-Review dataset).

Methods	Venue	D	E	K	B	Avg
ERM	-	74.17	73.17	73.67	71.58	73.15
ERM w/ MAD		76.08	74.33	73.33	74.67	**74.60**
Mixup [61]	ICLR 18	74.83	72.17	73.58	72.67	73.31
Mixup w/ MAD		75.33	73.58	74.33	73.75	**74.25**
Mixstyle [65]	ICLR 21	74.75	73.17	74.33	72.33	73.65
Mixstyle w/ MAD		75.17	72.75	75.00	75.25	**74.54**
DSU [30]	ICLR 22	75.00	73.45	75.25	73.08	74.20
DSU w/ MAD		76.42	74.33	76.50	75.17	**75.60**

Single-DG results on 2D images (PACS dataset).

Methods	Venue	P	A	C	S	Avg
ERM		33.65	65.38	64.20	34.15	49.34
ERM w/ MAD		32.32	66.47	69.80	34.54	**50.78**
Augmix [24]	ICLR 19	38.30	66.54	70.16	52.48	56.87
Augmixw/ MAD		36.19	68.04	73.11	54.44	**57.94**
pAdaIn [34]	CVPR 21	33.66	64.96	65.24	32.04	48.98
pAdaIn w/ MAD		34.66	65.64	70.10	42.85	**53.31**
Mixstyle [65]	ICLR 21	37.44	67.60	70.38	34.57	52.50
Mixstyle w/ MAD		41.57	69.88	71.61	41.58	**56.16**
ACVC [15]	CVPR 22	48.05	73.68	77.39	55.30	63.61
ACVC w/ MAD		52.95	75.51	77.25	57.75	**65.87**

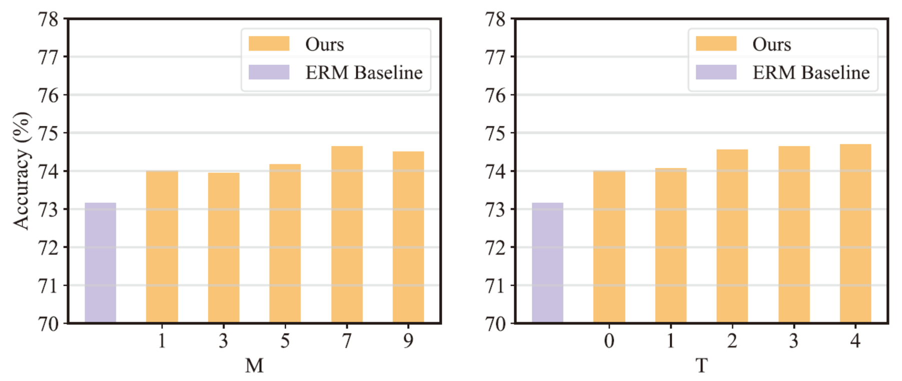
Single-DG results on 3D point cloud (PointDA-10 dataset).

Methods	Venue	SH	SC	M	Avg
ERM	-	25.69	45.09	32.94	34.57
ERM w/ MAD		31.11	48.07	34.69	**37.91**
Mixstyle [65]	ICLR 21	27.18	46.25	27.93	33.78
Mixstyle w/ MAD		29.89	51.01	33.57	**38.16**
DSU [30]	ICLR 22	25.74	43.53	31.61	33.63
DSU w/ MAD		28.92	47.69	32.72	**36.45**

Single-DG results on 2D images segmentation (GTA-5 → Cityscapes).

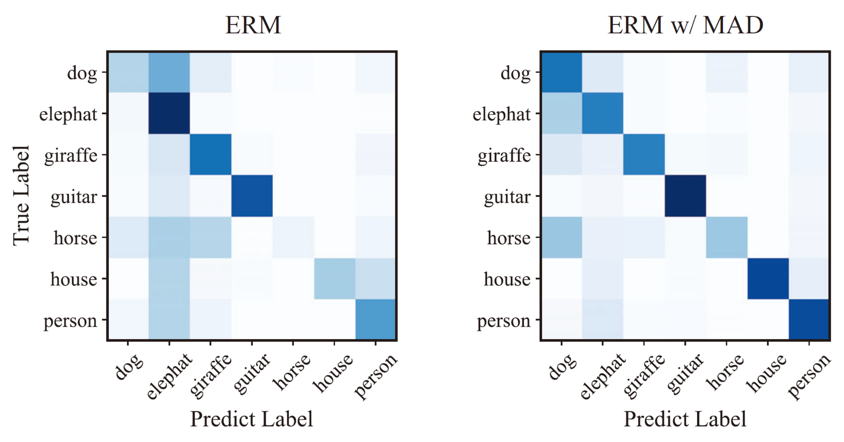
Methods	Venue	mIOU(%)	mACC(%)
ERM	-	37.0	51.5
pAdaIN [34]	CVPR 21	38.3	52.1
Mixstyle [65]	ICLR 21	40.3	53.8
DSU [30]	ICLR 22	42.3	54.7
ERM w/ MAD	-	**38.9**	**52.2**
DSU w/ MAD	-	**43.8**	**57.2**

Analysis Results:



(a) sensitivity to M (T = 3) (b) sensitivity to T (M = 7)

Hyper-parameter Sensitivity Analysis on 1D texts.



Confusion Matrix Analysis on 2D images (PACS).

Acknowledgement: This work is supported by Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01), ZJ Lab, and Shanghai Center for Brain Science and Brain-Inspired Technology, and the Shanghai Rising Star Program (No.21QC1400900).