JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

WED-AM-340

Deep Frequency Filtering for Domain Generalization

Shiqi Lin[1] , Zhizheng Zhang[2] , Zhipeng Huang[1] , Yan Lu[2] , Cuiling Lan[2] , Peng Chu[2] , Quanzeng You[2] , Jiang Wang[2] , Zicheng Liu[2] , Amey Parulkar[2] , Viraj Navkal[2] , Zhibo Chen[1]

[1]*University of Science and Technology of China* [2]*Microsoft*

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Summary

- In this work, we seek to learn **generalizable features** from **a frequency perspective for domain generalization**.

- We propose **Deep Frequency Filtering** (DFF) for learning domain-generalizable features, which is the first endeavour to **explicitly modulate the frequency components of different transfer difficulties** across domains **in the latent space** during training.

- To achieve this, we perform **Fast Fourier Transform** (FFT) for **the feature maps** at different layers, then **adopt a light-weight module to learn instance-adaptive attention masks** from the frequency representations after FFT to **enhance transferable components while suppressing the components not conducive to generalization.**

Background

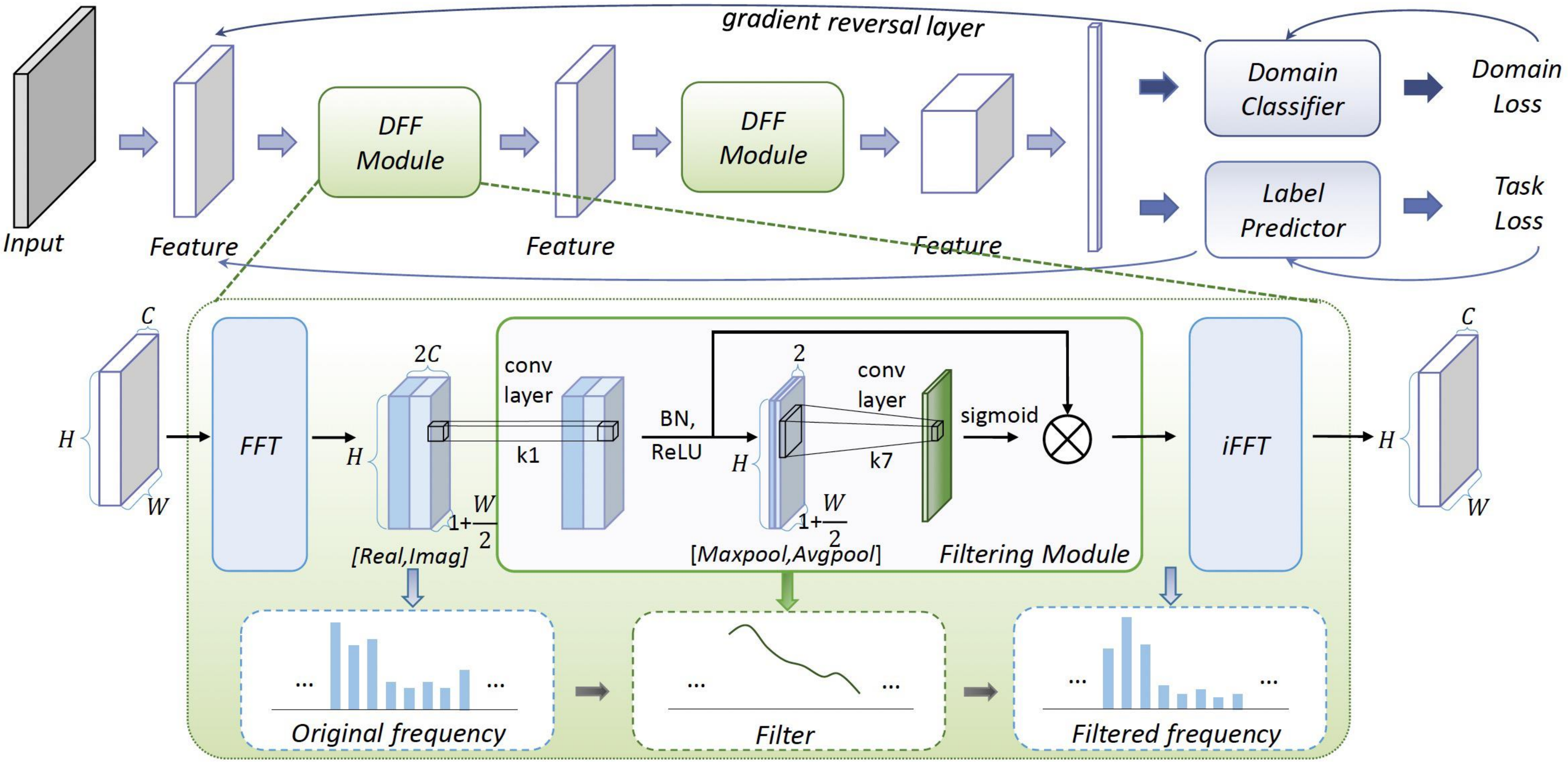
- **Domain Generalization** (DG) seeks to break through the i.i.d. assumption.
- Learning **generalizable feature** representations is of high practical value for both industry and academia.
- **Frequency analysis** has been widely used in conventional digital image processing for decades.
- Recently, frequency-based operations,e.g., Fourier transform, set forth to be incorporated into deep learning methods for different purposes.

Motivation

DNNs have **preferences for some frequency components** in the learning process and indicated that this may affect the robustness of learned features.

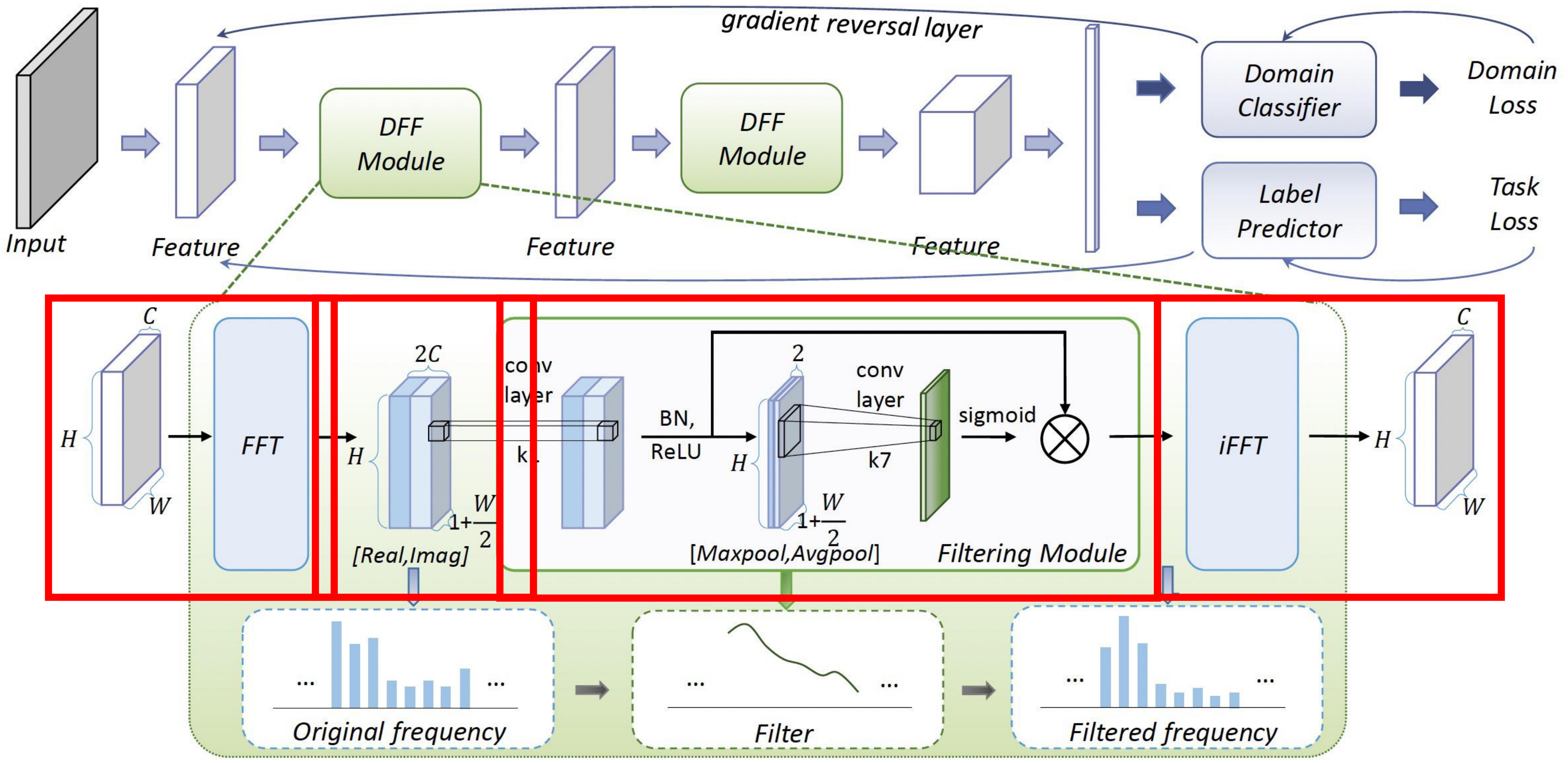
The frequency components of different cross-domain transferability are dynamically modulated in an end-to-end manner during training.

Overall framework



JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

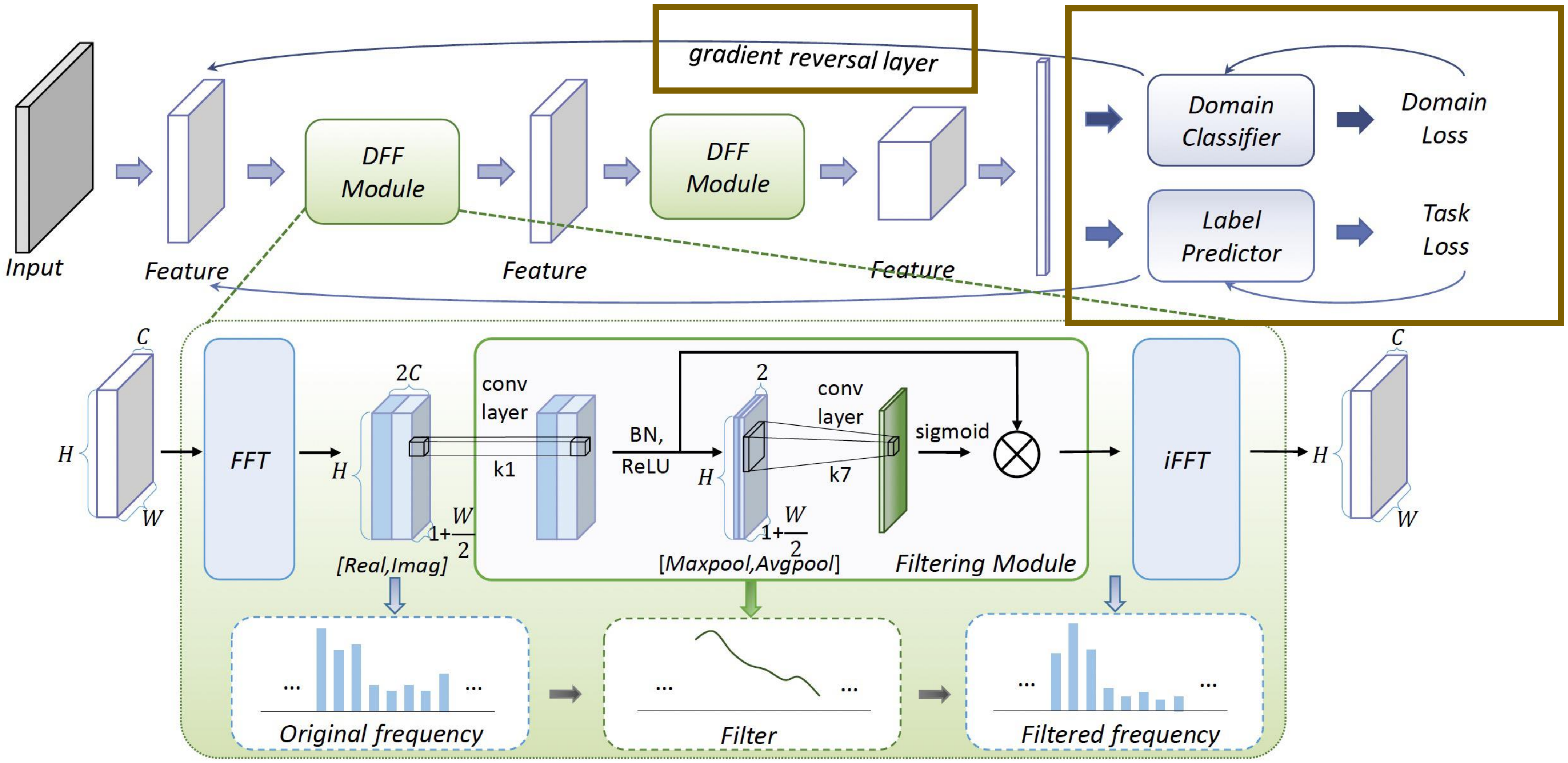
Overall framework



$$\mathbf{X}'_F = \mathbf{X}_F \otimes \sigma(\mathrm{Conv}_{7\times 7}([\mathrm{AvgPool}(\mathbf{X}_F), \mathrm{MaxPool}(\mathbf{X}_F)]))$$

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

Overall framework

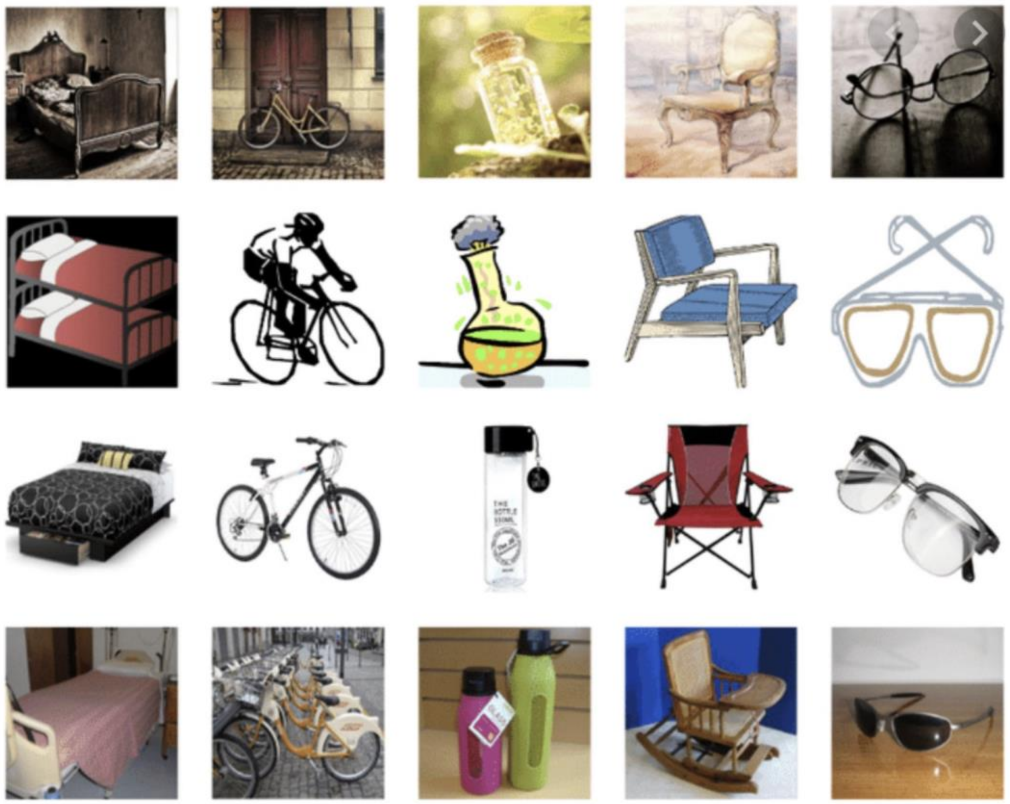


Experiments

Task-1: the close-set classification task

Train

Test



Task-2: the open-set retrieval task, i.e., person re-identification (ReID).

Train

Test

Ablation Studies

- *Frequency-domain v.s. Original-domain filtering*

Method	Source→Target					
	MS+CS+C3→MA		MS+MA+CS→C3		MA+CS+C3→MS	
	mAP	R1	mAP	R1	mAP	R1
Base	59.4	83.1	30.3	29.1	18.0	41.9
SBase (FFC)	66.2	84.7	35.8	35.4	19.4	44.8
Ori-F in f_l	66.9	85.0	36.2	35.9	19.8	45.1
Ori-F in f_g	61.9	83.5	32.7	31.9	18.4	42.8
Fre-F (Ours)	**71.1**	**87.1**	**41.3**	**41.1**	**25.1**	**50.5**

Ablation Studies

- *The importance of instance-adaptive attention*

Method	Source→Target					
	MS+CS+C3→MA		MS+MA+CS→C3		MA+CS+C3→MS	
	mAP	R1	mAP	R1	mAP	R1
Base	59.4	83.1	30.3	29.1	18.0	41.9
Task.(C)	62.7	80.0	32.1	31.4	19.5	44.9
Task.(S)	68.6	85.8	37.0	36.3	20.8	45.4
Ins.(C)	69.8	86.2	36.4	35.9	21.0	45.7
Ins.(S) (Ours)	**71.1**	**87.1**	**41.3**	**41.1**	**25.1**	**50.5**

Experiments

- *On close-set classification task*

Method	Source→Target				Avg
	Cl,Pr,Rw→Ar	Ar,Pr,Rw→Cl	Ar,Cl,Rw→P	Ar,Cl,Pr→Rw	
Baseline	52.2	45.9	70.9	73.2	60.5
CCSA [51]	59.9	49.9	74.1	75.7	64.9
D-SAM [18]	58.0	44.4	69.2	71.5	60.8
MMD-AAE [35]	56.5	47.3	72.1	74.8	62.7
CrossGrad [61]	58.4	49.4	73.9	75.8	64.4
JiGen [8]	53.0	47.5	71.5	72.8	61.2
RSC [29]	58.4	47.9	71.6	74.5	63.1
MixStyle [88]	58.7	53.4	74.2	75.9	65.5
Ours	**65.4**	**53.7**	**74.9**	**76.5**	**67.6**

Experiments

- *On open-set person re-id task*

Method	Source	Target:VIPeR(V)		Target:PRID(P)		Target:GRID(G)		Target:iLIDS(I)		Mean of V,P,G,I	
		R1	mAP	R1	mAP	R1	mAP	R1	mAP	R1	mAP
CDEL [43]	All	38.5	-	57.6	-	33.0	-	62.3	-	47.9	-
DIMN [62]	All	51.2	60.1	39.2	52.0	29.3	41.1	70.2	78.4	47.5	57.9
DDAN [9]	All	56.5	60.8	62.9	67.5	46.2	50.9	78.0	81.2	60.9	65.1
RaMoE [15]	All	56.6	64.6	57.7	67.3	46.8	54.2	85.0	90.2	61.5	69.1
SNR [31]	All	49.2	58.0	47.3	60.4	39.4	49.0	77.3	84.0	53.3	62.9
CBN [90]	All	49.0	59.2	61.3	65.7	43.3	47.8	75.3	79.4	57.2	63.0
Person30K [2]	All	53.9	60.4	60.6	68.4	50.9	56.6	79.3	83.9	61.1	67.3
DIR-ReID [80]	All	58.3	62.9	71.1	75.6	47.8	52.1	74.4	78.6	62.9	67.3
MetaBIN [13]	All	56.2	66.0	72.5	79.8	49.7	58.1	79.7	85.5	64.5	72.4
QAConv$_{50}$ [42]	All w/o D	57.0	66.3	52.3	62.2	48.6	57.4	75.0	81.9	58.2	67.0
M^{3}L [84]	All w/o D	60.8	68.2	55.0	65.3	40.0	50.5	65.0	74.3	55.2	64.6
MetaBIN [13]	All w/o D	55.9	64.3	61.2	70.8	50.2	57.9	74.7	82.7	60.5	68.9
Ours	All w/o D	**65.7**	**74.2**	**71.8**	**78.6**	**56.4**	**65.5**	**83.6**	**88.3**	**69.4**	**76.7**

Experiments

- *On open-set person re-id task*

Method	Source	Market-1501		Source	CUHK03		Source	MSMT17	
		mAP	R1		mAP	R1		mAP	R1
SNR [31]	MS+CS+C3	34.6	62.7	MS+CS+MA	8.9	8.9	CS+MA+C3	6.8	19.9
M^3L [84]		58.4	79.9		20.9	31.9		15.9	36.9
M^3L* [84]		61.5	82.3		34.2	34.4		16.7	37.5
$QAConv_{50}$ [42]		63.1	83.7		25.4	24.8		16.4	45.3
MetaBIN [13]		57.9	80.0		28.8	28.1		17.8	40.2
Ours		**71.1**	**87.1**		**41.3**	**41.1**		**25.1**	**50.5**
SNR [31]	Com-(MS+CS+C3)	52.4	77.8	Com-(MS+CS+MA)	17.5	17.1	Com-(CS+MA+C3)	7.7	22.0
M^3L [84]		61.2	81.2		32.3	33.8		16.2	36.9
M^3L* [84]		62.4	82.7		35.7	36.5		17.4	38.6
$QAConv_{50}$ [42]		66.5	85.0		32.9	33.3		17.6	46.6
MetaBIN [13]		67.2	84.5		43.0	43.1		18.8	41.2
Ours		**81.0**	**92.3**		**51.5**	**51.2**		**25.3**	**51.8**

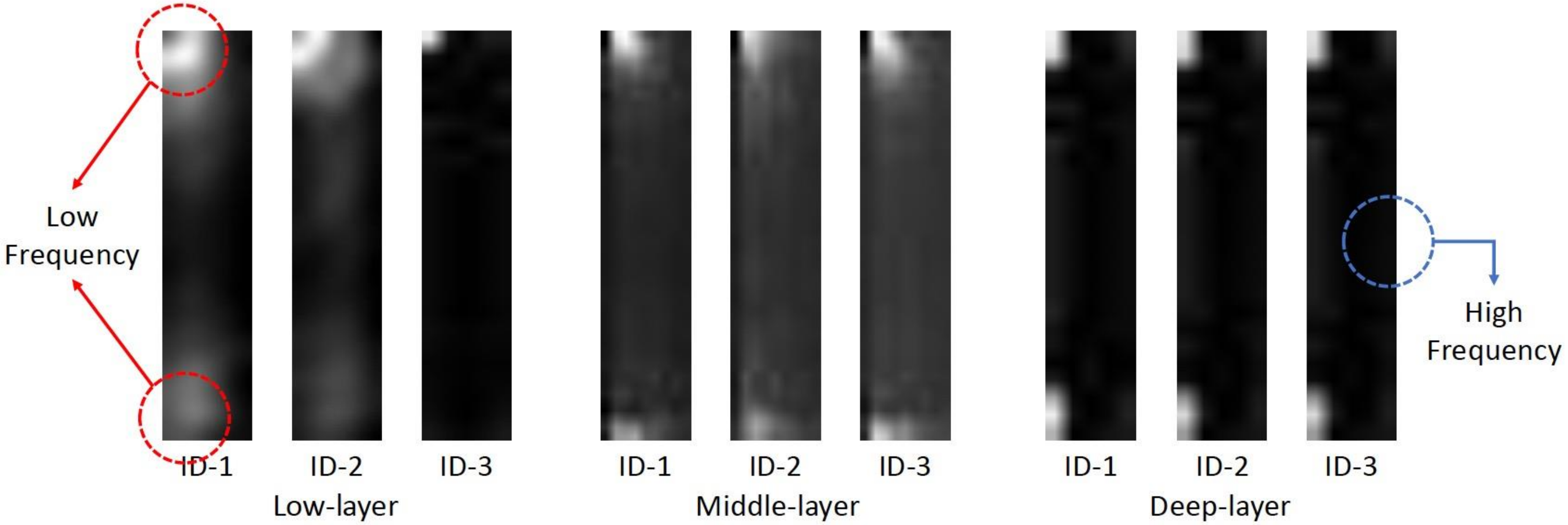
Experiments

- *Complexity Analysis*

Model	#Para	GFLOPs	Acc.	Model	#Para	GFLOPs	Acc.
ResNet18	11.7M	1.8	69.8	ResNet50	25.6M	4.1	76.3
DFF-R18	12.2M	2.0	72.3	DFF-R50	27.7M	4.5	77.9

Visualization

- *Visualization of Learned Masks*



Visualization

- *Visualization of Learned Feature Maps*

Conclusion

- **Contribution 1:** We conceptualize ***Deep Frequency Filtering (DFF)***, which is a new technique capable of **enhancing the transferable frequency components** and **suppressing the ones not conducive** to generalization **in the latent space**.

- **Contribution 2:** We conduct an empirical study for the comparison of different design choices on implementing DFF and find that the **instance-level adaptability** is required when learning frequency-space filtering for domain generalization.

- **Broad Impact:** We leave the exploration on more instantiations of our conceptualized DFF in future work and encourage more combinations and interplay between **conventional signal processing and deep learning technologies**.

Thanks for watching!