Supplementary Material for Representation Enhancement-Stabilization: Reducing Bias-Variance of Domain Generalization

Wei Huang[1], Yilei Shi[1], Zhitong Xiong[1], and Xiao Xiang Zhu[1,2]

1. Technical University of Munich, Arcisstraße 21, 80333 München, Germany
2. Munich Center for Machine Learning, 80333 Munich
`{w2wei.huang, yilei.shi, zhitong.xiong, xiaoxiang.zhu}@tum.de`

Abstract. In the supplementary document, we first compare the effects of data augmentation on model generalization from two perspectives: (1) image-level vs. feature-level frequency augmentation, and (2) normal feature augmentation vs. feature frequency augmentation. The comparison results demonstrate the superiority of our feature frequency augmentation in the Representation Enhancement (RE) module. Then we provide the detailed domain-wise train-validation results of the proposed Representation Enhancement-Stabilization (RES) on five used domain datasets.

1 Supplementary Experiments

1.1 Image-level vs. Feature-level Frequency Augmentation

Frequency domain-based data augmentation for DG usually operates at image level [2,3]. Different from them, the proposed RE module conducts the frequency data augmentation at the feature level. We compare their performance using the same augmentation strategies, including `random noise` (ns), `random dropout` (dp), and `mixup` (mix), on the phase spectrum extracted from an RGB-channel image or a multi-layer feature map. The experimental results of image-level feature augmentation, denoted as RE_I, and feature-level frequency augmentation, denoted as RE, are provided in Table 1 and Table 2, respectively.

Frequency data augmentations at both levels contribute to improvements over the baseline Empirical Risk Minimization (ERM) method. However, our feature-level frequency augmentation exhibits a higher increase, boosting performance by 3.3 percentage points from 85.5% to 88.8% (RE^{ns}+RE^{dp}+RE^{mix}). In contrast, image-level frequency augmentation yields a smaller gain of 1.5 percentage points, elevating performance from 85.5% to 87.0% (RE_I^{dp}). Besides, our feature-level frequency augmentation has a smoother training process with a standard deviation of 0.6, which is smaller than that of image-level counterpart. The comparison results highlight the superior performance of our method. This advantage is likely due to the phase spectrum of an image remaining constant, whereas the phase spectrum of the corresponding feature map changes at different training stages. Consequently, frequency augmentation at the feature level offers greater representational diversity, enhancing the model's generalization capabilities.

Table 1: Image-level frequency augmentation on PACS. The best results are **bolded**.

ERM	RE_I^{ns}	RE_I^{dp}	RE_I^{mix}	PACS Art	PACS Clipart	PACS Painting	PACS Sketch	**Avg**
✓				84.4±0.6	81.1±0.7	96.3±0.5	80.2±2.5	85.5±1.1
✓	✓			86.6±1.1	80.1±0.8	96.5±0.4	79.9±1.9	85.8±1.1
✓		✓		87.2±1.5	80.9±2.7	**96.9±0.2**	**82.9±1.2**	**87.0±1.5**
✓			✓	87.5±0.5	**81.3±1.0**	96.0±0.3	82.6±1.4	86.9±0.8
✓	✓	✓	✓	**88.4±0.9**	80.0±0.9	96.4±0.6	82.1±0.3	86.7±1.4

Table 2: Feature-level frequency augmentation (our RE module) on PACS. The best results are **bolded**.

ERM	RE^{ns}	RE^{dp}	RE^{mix}	PACS Art	PACS Clipart	PACS Painting	PACS Sketch	**Avg**
✓				84.4±0.6	81.1±0.7	96.3±0.5	80.2±2.5	85.5±1.1
✓	✓			88.8±1.2	82.2±0.5	97.0±0.6	81.2±0.6	87.3±0.7
✓		✓		87.6±0.3	82.7±1.0	96.8±0.8	82.1±1.6	87.3±0.9
✓			✓	**89.4±0.2**	81.4±1.1	96.6±0.4	85.1±0.9	88.5±0.7
✓	✓	✓	✓	89.0±0.9	**83.0±0.4**	**97.0±0.4**	**85.1±0.7**	**88.8±0.6**

1.2 Normal Feature Augmentation vs. Feature Frequency Augmentation

To confirm the impact of augmenting the phase spectrum of features, we conduct a comparative experiments between normal feature augmentation and our feature frequency augmentation. In the normal feature augmentation, augmentation strategies are applied directly to the original feature map, labeled as FE^{ns}, FE^{dp}, and FE^{mix}, and its results are given in Table 3. Here FE^{ns} equals [1] from the perspective of augmentation strategy.

It is observed that augmenting the original features directly can also enhance DG performance. Nevertheless, as detailed in Table 2, our feature frequency augmentation method maintains a performance edge, achieving an advantage of 1.8 percentage points, from 87.0 to 88.8. The findings underscore the importance of enhancing features while preserving their semantics for improving DG.

Table 3: Normal feature augmentation on PACS. The best results are **bolded**.

ERM	FE^{ns}	FE^{dp}	FE^{mix}	PACS				Avg
				Art	Clipart	Painting	Sketch	
✓				84.4±0.6	81.1±0.7	96.3±0.5	80.2±2.5	85.5±1.1
✓	✓			87.5±1.3	82.1±0.6	**97.1±0.6**	77.5±1.7	86.0±1.1
✓		✓		86.5±1.7	80.2±0.9	96.3±0.3	**81.1±0.9**	86.0±1.0
✓			✓	87.0±0.9	81.6±1.9	96.8±0.3	81.1±1.2	86.6±1.1
✓	✓	✓	✓	**87.8**±**1.0**	**82.9**±**1.9**	97.0±0.6	80.2±2.1	**87.0**±**1.4**

1.3 Domain-wise Performance of RES

Here we provide the detailed domain-wise performance of all the five used dataset based on in-domain train-val model selection, in Tables 4 and 5.

Table 4: Performance comparison of PACS between SOTA methods and the proposed RES. The bests are **bolded**.

	PACS				Avg	VLCS				Avg
	Art	Clipart	Painting	Sketch		Caltech101	LabelMe	SUN09	VOC2007	
RES	91.6±0.4	84.2±0.3	98.1±0.1	86.1±0.4	90.0±0.3	99.3±0.1	63.7±0.1	80.6±0.5	75.5±0.5	79.8±0.3
	OfficeHome				Avg	TerraIncognita				Avg
	Art	Clipart	Product	RealWorld		L100	L38	L43	L46	
RES	67.0±0.2	60.7±0.4	78.9±0.3	80.5±0.4	71.8±0.3	60.9±0.5	45.8±0.7	58.1±0.7	40.7±0.5	51.4±0.6

Table 5: Performance comparison of DomainNet between SOTA methods and the proposed RES. The bests are **bolded**.

Method	DomainNet						Avg
	Clipart	Info.	Painting	Quick.	Real	Sketch	
RES	66.3±0.2	22.6±0.2	55.0±0.1	15.7±0.1	63.8±0.4	57.0±0.2	46.7±0.2

References

1. Li, P., Li, D., Li, W., Gong, S., Fu, Y., Hospedales, T.M.: A simple feature augmentation for domain generalization. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 8886–8895 (2021)
2. Lv, F., Liang, J., Li, S., Zang, B., Liu, C.H., Wang, Z., Liu, D.: Causality inspired representation learning for domain generalization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 8046–8056 (2022)
3. Xu, Q., Zhang, R., Zhang, Y., Wang, Y., Tian, Q.: A fourier-based framework for domain generalization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14383–14392 (2021)