Proceedings of Machine Learning Research vol 278:1–7, 2025 2025 2nd International Conference on Machine Learning and Intelligent Computing

A Spectrum Filtering Framework for Domain Generalization

Fuchao Li* LFC717817@GMAIL.COM
Northeastern University, School of Computer Science and Engineering, Shenyang China

Kun Li
Northeastern University, School of Software, Shenyang China
**Corresponding author*
Fuchao Li and Kun Li contributed equally to this work

Editors: Nianyin Zeng, Ram Bilas Pachori and Dongshu Wang

Abstract

Domain generalization aims to address the distribution shift problem inherent in neural networks, wherein a misalignment between test data distribution and training data distribution leads to significant performance degradation. This paper introduces Fourier Style Restitution (FSR), a Fourier-transform-driven method for domain generalization. FSR integrates the principles of Fourier augmentation and style disentanglement with feature reconstruction, enhancing model generalizability to unseen domains. The framework implements a cross-domain filtering enhancement strategy based on Fourier transform, leveraging frequency domain filtering to bolster model robustness against distributional variations. Through this paradigm, each sample transcends source domain constraints to derive optimized domain-invariant feature representations tailored to its intrinsic characteristics. The framework further incorporates style regularization to distill consistency signals from stylized images and employs prototype compensation to recover lost domain-invariant features. Extensive experiments demonstrate state-of-the-art performance on benchmark datasets. The method's efficacy stems from feature enhancement and style reconstruction through Fourier-based operations for robust domain generalization.

Keywords: Domain generalization, Fourier transform, data augmentation, image classification, style transfer.

1. Introduction

Image classification has advanced significantly due to deep neural networks, yet practical deployment faces major challenges—particularly domain shift. Traditional models assume identical training-testing data distributions, yielding high performance when met. However, real-world scenarios often violate this assumption (Li et al., 2024b). Distribution discrepancies arise from factors like weather variations, lighting changes, and shooting angles. Consequently, models excelling in training suffer substantial accuracy drops during real-world testing, severely limiting practical applications of these algorithms.

Domain adaptation techniques are commonly employed to address domain shift issues in image classification (Li et al., 2025). These methods operate under the premise that source and target domains exhibit certain relational affinities, permitting transformation through specific mapping functions to diminish inter-domain discrepancies. This enhances model generalization within target domains. A key advantage lies in achieving competent target-domain performance with limited samples, circumventing the need for extensive target data collection. Furthermore, such approaches mitigate domain shift induced by distributional differences, thereby augmenting model generalization capabilities. However, standard domain adaptation presupposes accessibility of target domain imagery—an assumption often violated in real-world contexts where target data may be unavailable, prohibitively complex, or costly to acquire. This limitation constrains adaptation applicability, motivating the emergence of domain generalization (DG) as an alternative research focus.

We propose Fourier Style Restitution (FSR), a domain restitution model that integrates Frequency-Domain Feature Augmentation, style removal and prototype restitution. Experimental results on benchmark dataset demonstrate the efficacy of our framework. Through extensive evaluations, We establish that FSR outperforms several advanced Domain Generalization (DG) (Jahanifar et al., 2025) methods in terms of generalization capability to unseen target domains. This finding indicates that learning enhanced domain-consistent representations from frequency-domain information with feature compensation indeed facilitates superior cross-domain generalization. The effectiveness of the proposed modules is further validated via rigorously designed ablation studies. Through in-depth theoretical analysis of the underlying hypotheses and rationale, we demonstrate that phase information inherently contains domain-invariant features within visual data. Thus, FSR effectively extracts these consistent representations to improve generalization performance.

2. Related Work

Domain Generalization (DG) fundamentally aims to train models with robust generalization capabilities to unseen target domains, intrinsically addressing the Distribution Shift problem in deep neural networks.Conventional DG approaches primarily extend domain alignment techniques from domain adaptation (Li et al., 2024b), employing kernel methods (Li et al., 2025) to project source domain features into higher-order spaces with minimized distribution divergence. Through meta-learning frameworks (Lin et al., 2023), source domains are partitioned into training and validation subsets, simulating domain generalization scenarios for model optimization. Adversarial learning strategies further enhance this process, where generators synthesize images to fool discriminators, thereby compelling the model to extract domain-invariant representations. Inspired by human visual systems that filter irrelevant information, domain-specific information elimination mechanisms have been proposed. Additionally, self-supervised jigsaw puzzle tasks (Chen et al., 2023b) enforce domain-invariant learning via designed pretext objectives, while other methods integrate low-quality signal decomposition with adversarial learning. Crucially and in contrast to these paradigms, our proposed Fourier Style Restitution (FSR) diverges fundamentally. FSR dynamically generates diverse source domains with heterogeneous distributions through Fourier transformations, extracts cross-domain consistent features via structured style elimination, and reconstructs discriminative information from the eliminated components. Empirical evaluations confirm the superior performance of this approach.

3. Method

Given a training set $D_s = \{D_1, \dots, D_s\}$ comprising s source domains, where the $k-\mathrm{th}$ domain D_k contains N_k labeled samples $\left\{\left(x_i^k, y_i^k\right)\right\}_{i=1}^{N_k}$. Here x_i^k denotes an input samples and $y_i^k \in \{1, \dots, C\}$ its discrete label. The goal of domain generalization is to learn a model $f(\cdot, \theta)$ capable of generalizing well to unseen domains.

Motivated by the phase spectrum's domain-invariant properties, we propose a Fourier-based augmentation strategy centered on adaptive frequency filtering. This approach generates style diversity through filter-constrained amplitude operations, preserving critical spectral bands while simulating real-world domain variations. Additionally, we introduce a style removal mechanism that eliminates nuisance styles while preserving discriminative features. Subsequently, prototype restitution reintroduces the beneficial information back into the de-stylized features, enhancing the model's perception of key characteristics. We employ a Residual Network (ResNet (Xu et al., 2023)) as the feature extractor. Figure 1 illustrates the FSR framework. The following sections describe its various components.

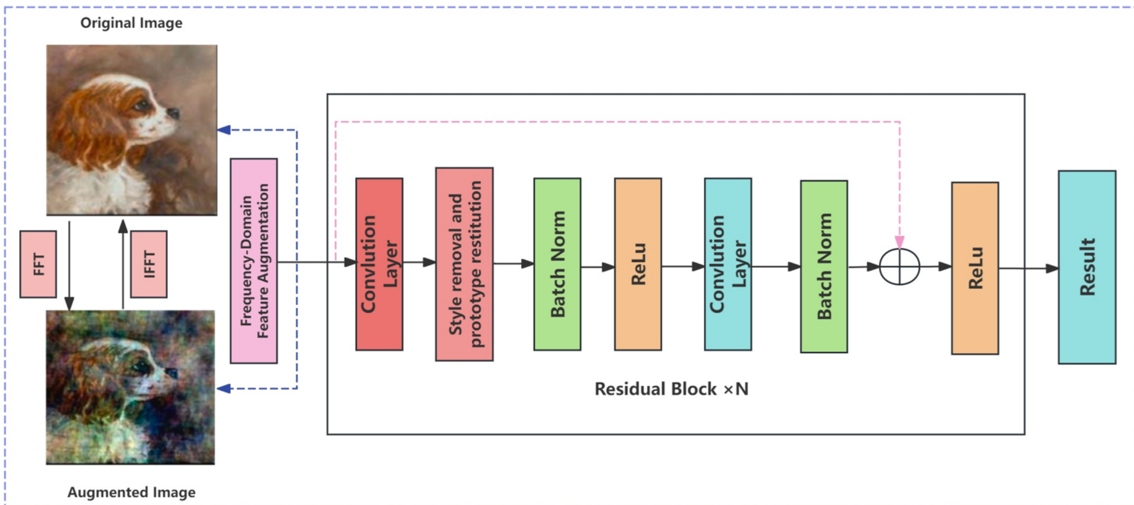

Figure 1: The framework of the proposed FSR

3.1. Frequency-Domain Feature Augmentation

In this section, we implement a frequency-domain Feature Augmentation method by combining spectral filtering and Fourier transform. Specifically, we adopt the classic discrete Fourier transform to implement a cross-domain filtering enhancement strategy.

In the Fourier transform of images, the amplitude spectrum characterizes domain-specific stylistic information (e.g., illumination, texture), while the phase spectrum encodes domain-invariant structural information (e.g., object contours). We decompose randomly sampled image pairs from the training set into their frequency-domain representations, extracting amplitude and phase components. This procedure operates as follows:

$$\mathcal{F}_i = \mathcal{FFT}(x_i^k) \to (\mathcal{A}_i, \mathcal{P}_i) \tag{1}$$

$$\mathcal{F}_j = \mathcal{FFT}\left(x_j^l\right) \to (\mathcal{A}_j, \mathcal{P}_j) \tag{2}$$

$$AMP\left(x\right) = \sqrt{Re^2 + Im^2} \tag{3}$$

$$Phase\left(x\right) = arctan(Im/Re) \tag{4}$$

Subsequently, the amplitude is filtered and a blending coefficient is introduced to modulate the perturbation magnitude of the amplitude spectrum (Jin et al., 2025), thereby simulating unknown domain style variations, while the phase spectrum remains unaltered to preserve structural information from the source image.

$$\mathcal{A}\left(x_i^k\right) = (1-\alpha)\cdot \text{ filtered } \left(\text{Amp}\left(x_i^k\right)\right) + \alpha\cdot \text{ filtered } \left(\text{Amp}\left(x_j^l\right)\right) \tag{5}$$

Specifically, let $H(u,v)$ be a frequency-domain filter, then the filtered amplitude spectrum is shown in Formula (6) below to achieve band-filtered amplitude mixing.

$$filtered\,(A)\,(u,v) = H\,(u,v)\cdot A\,(u,v) \tag{6}$$

This approach mathematically formulates cross-domain data augmentation through convex combinations of Fourier amplitude spectra between image pairs, while retaining the structural integrity encoded in phase components.

$$\hat{x}_i^k = \mathcal{IFFT}(\mathcal{A}(x_i^k)\cdot e^{jPhase}) \tag{7}$$

The resultant Fourier representation is then transformed back into the spatial domain using the inverse Fourier transform to produce the augmented image $\hat{x}_i^k$. This particular Fourier-based augmentation technique is referred to as Frequency-Domain Feature Augmentation.

3.2. Style Removal and Prototype Restitution

In this section, we present a style removal and reconstruction framework that integrates a local cross-channel attention mechanism and a feature compensation strategy to achieve orthogonal residual decomposition. As shown in Figure 1, the feature maps output by the convolutional layer undergo instance normalization to eliminate style variations across different samples, resulting in style-removed features $\widetilde{Y}$. Subsequently, we compute the feature residual R between the normalized and original features to represent the filtered style information and discriminative information.

$$\tilde{Y} = \text{IN}(Y) = \gamma\left(\frac{Y-\mu(Y)}{\sigma(F)}\right) + \beta, R = Y - \tilde{Y} \tag{8}$$

Local cross-channel attention mechanism that decomposes the residual R into two orthogonal components: a task-relevant component R^+ and a task-irrelevant component R^-. First, global average pooling is applied to the residual R to generate channel-wise statistics z. Next, a 1D convolution with an adaptively determined kernel size k (operating on the channel dimension) is employed to capture local cross-channel interactions. Finally, a sigmoid activation function is used to generate weights w, which are then utilized to separate task-relevant and task-agnostic components. Ultimately, the task-relevant component is re-injected into the normalized features to enhance their discriminative power. This process can be expressed by the following formula:

$$z = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} R_c(i,j) \tag{9}$$

$$\hat{z} = \text{Conv1D}(z), \omega = \sigma(\hat{z}) \tag{10}$$

$$R^+ = \omega\odot R, R^- = (1-\omega)\odot R \tag{11}$$

$$Z = \tilde{Y} + R^+ \tag{12}$$

4. Experiments

This section presents Fourier Style Restitution (FSR)'s state-of-the-art domain generalization performance on PACS (Tan et al., 2024) and provides ablation analyses.

4.1. Implementation details

We used ResNet as the backbone network, SGD (Gu et al., 2023) as the optimizer, batch size 128, initial learning rate set to 0.005e-4 weight decay, and trained for 50 epochs, with a decay of 0.1 every 20 epochs. Our experiments mainly show advanced performance on PACS.

Table 1: Leave-one-domain-out results on PACS.

Methods	Art	Cartoon	Photo	Sketch	Average Accuracy
DeepAll (Guo et al., 2023)	77.63	76.77	95.85	69.50	79.94
MetaReg (Li et al., 2024a)	83.70	77.20	95.50	70.30	81.70
Epi-FCR (Chen et al., 2023a)	82.10	77.00	93.90	73.00	81.50
L2A-OT (Zhang et al., 2022)	83.30	**78.20**	**96.20**	73.60	82.80
FSR(ours)	**83.74**	76.87	95.07	**76.87**	**83.14**

Through experiments on PACS datasets, our model achieves good results on Art, Cartoon and Sketch, and achieves an accuracy of 83.14 on average, which is 3.2 percentage points higher than the baseline DeepAll, while outperforming the other three baseline models with the highest mean accuracy under the FSR framework.

4.2. Ablation studies

We evaluate component contributions on the PACS benchmark (Table 2). We take DeepAll as the Baseline, and respectively overlay the Baseline with a simple variant StyleV1(Add instance normalization (IN) layers after each convolutional block) of style removal and prototype restitution, as well as the complete form StyleV2, to verify the effectiveness of style reconstruction. Fourier refers to adding the Frequency-Domain Feature Augmentation module to the Baseline, while FSR simultaneously incorporates both the style reconstruction and Frequency-Domain Feature Augmentation modules. As shown in the experimental results of Table 2, the modules proposed in this paper can improve the performance of the baseline model to a certain extent.

Table 2: Ablation studies on PACS.

Methods	Art	Cartoon	Photo	Sketch	Average Accuracy
Baseline	77.63	76.77	**95.85**	69.50	79.94
StyleV1	78.65	76.21	95.44	70.34	80.16
StyleV2	79.42	75.95	95.62	73.13	81.03
Fourier	83.18	76.35	95.22	76.45	82.80
FSR(ours)	**83.74**	**76.87**	95.07	**76.87**	**83.14**

5. Conclusion

In this paper, we propose a useful generalization framework FSR. Leveraging Fourier transform and its Spectrum Filtering properties to generate images in multiple styles and reconstruct effective information through style manipulation, as confirmed by extensive experiments, our model demonstrates promising performance.

References

Jin Chen, Zhi Gao, Xinxiao Wu, and Jiebo Luo. Meta-causal learning for single domain generalization, 2023a.

Yingyi Chen, Xi Shen, Yahui Liu, Qinghua Tao, and Johan A. K. Suykens. Jigsaw-vit: Learning jigsaw puzzles in vision transformer, 2023b.

Xinran Gu, Kaifeng Lyu, Longbo Huang, and Sanjeev Arora. Why (and when) does local sgd generalize better than sgd?, 2023.

Jintao Guo, Lei Qi, and Yinghuan Shi. Domaindrop: Suppressing domain-sensitive channels for domain generalization, 2023.

Mostafa Jahanifar, Manahil Raza, Kesi Xu, Trinh Thi Le Vuong, Robert Jewsbury, Adam Shephard, Neda Zamanitajeddin, Jin Tae Kwak, Shan E Ahmed Raza, Fayyaz Minhas, and Nasir Rajpoot. Domain generalization in computational pathology: Survey and guidelines. *ACM Comput. Surv.*, 57(11), 2025. doi: 10.1145/3724391.

Xin Jin, Hongyu Zhu, Siyuan Li, Zedong Wang, Zicheng Liu, Juanxi Tian, Chang Yu, Huafeng Qin, and Stan Z. Li. A survey on mixup augmentations and beyond, 2025.

Deng Li, Aming Wu, Yaowei Wang, and Yahong Han. Prompt-driven dynamic object-centric learning for single domain generalization, 2024a.

Jindong Li, Yongguang Li, Yali Fu, Jiahong Liu, Yixin Liu, Menglin Yang, and Irwin King. Clip-powered domain generalization and domain adaptation: A comprehensive survey. 04 2025. doi: 10.48550/arXiv.2504.14280.

Jingjing Li, Zhiqi Yu, Zhekai Du, Lei Zhu, and Heng Tao Shen. A comprehensive survey on source-free domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8): 5743–5762, 2024b. doi: 10.1109/TPAMI.2024.3370978.

Mingkai Lin, Wenzhong Li, Ding Li, Yizhou Chen, Guohao Li, and Sanglu Lu. Multi-domain generalized graph meta learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37:4479–4487, 06 2023. doi: 10.1609/aaai.v37i4.25569.

Zhaorui Tan, Xi Yang, and Kaizhu Huang. Rethinking multi-domain generalization with a general learning objective, 2024.

Wanni Xu, You-Lei Fu, and Dongmei Zhu. Resnet and its application to medical image processing: Research progress and challenges. *Computer Methods and Programs in Biomedicine*, 240: 107660, 2023. doi: https://doi.org/10.1016/j.cmpb.2023.107660.

Yabin Zhang, Minghan Li, Ruihuang Li, Kui Jia, and Lei Zhang. Exact feature distribution matching for arbitrary style transfer and domain generalization, 2022.