

Review

Domain Generalization Through Data Augmentation: A Survey of Methods, Applications, and Challenges

Junjie Mai, Chongzhi Gao *  and Jun Bao

School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou 510006, China; 2112306052@e.gzhu.edu.cn (J.M.); 2112306144@e.gzhu.edu.cn (J.B.)

* Correspondence: czgao@gzhu.edu.cn

Abstract: Domain generalization (DG) has become a pivotal research area in machine learning, focusing on equipping models with the ability to generalize effectively to unseen test domains that differ from the training distribution. This capability is crucial, as real-world data frequently exhibit domain shifts that violate the assumption of independent and identically distributed (i.i.d.) data, resulting in significant declines in model performance. Among the various strategies to address domain generalization, data augmentation has garnered substantial attention as an effective approach for mitigating domain shifts and improving model robustness. In this survey, we examine the role of data augmentation in domain generalization, offering a comprehensive overview of its methods, applications, and challenges. We present a detailed taxonomy of data augmentation techniques, categorized along three dimensions: scope, nature, and training dependency. Additionally, we provide a comparative analysis of key methods, highlighting their strengths and limitations. Finally, we explore the domain-specific applications of data augmentation and analyze their effectiveness in enhancing generalization across various real-world tasks, including computer vision, NLP, speech, and robotics. We conclude by examining key challenges—such as computational cost and augmentation overfitting—and outline promising research directions, with a focus on advancing cross-modal augmentation techniques and developing standardized evaluation benchmarks.

Keywords: domain generalization; data augmentation; robust

MSC: 68T07



Academic Editor: Giovanni Stabile

Received: 5 February 2025

Revised: 24 February 2025

Accepted: 27 February 2025

Published: 28 February 2025

Citation: Mai, J.; Gao, C.; Bao, J. Domain Generalization Through Data Augmentation: A Survey of Methods, Applications, and Challenges. *Mathematics* **2025**, *13*, 824. <https://doi.org/10.3390/math13050824>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The success of machine learning (ML) algorithms is largely predicated on the assumption of independent and identically distributed (i.i.d.) data, where both training and testing samples are drawn from the same underlying distribution. However, in real-world applications, distributional discrepancies between training and testing data are ubiquitous, often leading to significant declines in model performance [1,2]. For example, in medical image segmentation, variations in imaging equipment and patient demographics can reduce segmentation accuracy, thereby adversely affecting disease diagnosis and treatment planning [3]. Similarly, in autonomous driving, differences in geographic environments, weather conditions, and sensing equipment can impair the accurate identification of road boundaries or pedestrians, increasing the risk of traffic accidents [4]. Collecting data that comprehensively capture all possible distributions is not only prohibitively expensive but also practically infeasible.

These distributional shifts expose a fundamental limitation of ML models: their inability to generalize effectively beyond the training domains. This limitation represents a significant barrier to the deployment of ML systems in real-world scenarios. Consequently, ensuring the robustness and reliability of ML algorithms under cross-domain settings has become a critical challenge. To address this issue, the concept of Domain Generalization (DG) has emerged and garnered significant attention in recent years. DG seeks to develop models capable of performing well on unseen target domains without requiring access to target domain data during training. Various techniques have been proposed to address the DG challenge, which can be broadly categorized into representation learning, optimization strategies, and data augmentation. Among these, data augmentation is particularly notable for its effectiveness in improving model robustness and generalization by transforming and enriching existing datasets [5]. By introducing variations—such as rotations, scaling, or color adjustments—into the training data, these techniques enable models to learn more generalized and domain-invariant representations [6].

While studies like [7] and [8] provide systematic reviews of advancements in DG research, discussing its definition, theoretical underpinnings, and primary solutions, they also catalog relevant datasets and practical applications, offering a broad overview of the field. Similarly, [9] examines common data augmentation methods, highlighting their strengths and limitations. However, these reviews lack a focused exploration of the specific role of data augmentation in addressing DG challenges. In particular, they do not offer a comprehensive taxonomy of DG data augmentation methods or an in-depth analysis of their underlying principles. In contrast, this survey seeks to fill this gap by systematically categorizing data augmentation methods for DG based on their key characteristics. Through the development of an improved taxonomy framework, this work aims to provide researchers with clearer guidance for selecting or designing data augmentation techniques tailored to specific DG challenges.

The remainder of this article is organized as follows. Section 2 formally defines the domain generalization (DG) problem and introduces related research areas. In Section 3, we present a detailed taxonomy of data augmentation methods used in DG. Section 4 explores the applications of data augmentation in addressing DG challenges across various domains. Section 5 discusses the limitations of current data augmentation techniques in tackling DG problems, while Section 6 examines emerging research trends and future directions. Finally, we conclude the survey in Section 7.

2. Background

2.1. Formalization of Domain Generalization

Let the input space be denoted as $\mathcal{X}$ and the output space as $\mathcal{Y}$. The training dataset consists of M domains, $D_S = \{SD^1, SD^2, \dots, SD^M\}$, where each domain is represented as $SD^m = \{(x_i^m, y_i^m)\}_{i=1}^{N_m}$ for $m \in \{1, \dots, M\}$. The test dataset is $D_T = \{TD^1, TD^2, \dots, TD^J\}$ such that $D_S \cap D_T = \emptyset$. In domain generalization, it is assumed that both the source domains D_S and the target domains D_T share the same label space $\mathcal{Y}$. Formally, this is expressed as

$$\mathcal{Y}_S = \mathcal{Y}_T = \mathcal{Y},$$

where $\mathcal{Y}_S$ and $\mathcal{Y}_T$ represent the label spaces of the source and target domains, respectively.

The goal of domain generalization is to train a model $f : \mathcal{X} \rightarrow \mathcal{Y}$ on the training domains (source domains) D_S such that the model achieves minimal prediction error on unseen test domains (target domains) D_T :

$$\min_f E_{(x,y) \in D_T} \mathcal{L}(f(x), y),$$

where $\mathcal{L}$ is a predefined loss function.

When incorporating data augmentation into domain generalization, let the transformation function be denoted as $\mathcal{T} : \mathcal{X} \rightarrow \mathcal{X}$. For a given sample $(x_i^m, y_i^m) \in SD^m$, the transformation is applied to obtain an augmented sample $\hat{x}_i^m = \mathcal{T}(x_i^m)$. The resulting augmented dataset is $\hat{D}_S = \cup_{m=1}^M \{(\hat{x}_i^m, y_i^m)\}_{i=1}^{N_m}$. The augmented training dataset is then represented as $D'_S = D_S \cup \hat{D}_S$. Domain generalization will train the model f on the augmented training dataset D'_S .

2.2. Related Research Areas

Domain Adaptation (DA). Domain Adaptation (DA) is closely related to Domain Generalization (DG), as both aim to address the challenges posed by distributional discrepancies between training and testing data [10]. However, the two differ in their assumptions: DA assumes access to target domain data during training, whereas DG assumes no access to target domain data, making DG more challenging and applicable to real-world scenarios. DA is further categorized based on the extent of access to target domain data during training: **unsupervised DA** assumes access only to the input data x from the target domain, without corresponding labels y [11], while **semi-supervised DA** assumes access to all inputs x and a limited number of corresponding labels y [12]. By leveraging the target domain data, DA methods aim to learn domain-aligned representations that reduce the distribution gap between source and target domains. Common techniques include **adversarial training**, which uses adversarial processes to learn domain-invariant feature representations [13], and **distribution matching methods**, such as Maximum Mean Discrepancy (MMD), which minimize domain discrepancies in the feature space [14]. Interestingly, some of these techniques have been adapted for DG. By minimizing feature distribution differences across multiple source domains, DG models can learn cross-domain representations that generalize well to unseen target domains [15–18].

Transfer Learning. Transfer Learning focuses on leveraging knowledge from a source domain to improve performance on a related target domain or task, typically through a pretrain–fine-tune paradigm. In the pretraining phase, a base model is trained on data from source domains. During fine-tuning, the base model is adapted to the target domains using data from the target domains [19,20]. Unlike DG, Transfer Learning explicitly assumes access to target domain data for adaptation.

Multi-Task Learning (MTL). Multi-Task Learning is a framework in which multiple related tasks are optimized simultaneously, enabling knowledge sharing to improve learning efficiency. By leveraging task interrelations, MTL enhances performance across tasks through shared representations. Parameter sharing occurs either in the shallow layers of a single model or via regularization constraints across separate models [21,22]. Unlike DG, MTL focuses on improving performance across multiple tasks rather than generalizing to unseen domains.

Continual Learning (CL). Continual Learning, also known as lifelong learning, enables models to learn new tasks incrementally while retaining knowledge of previous tasks. Its primary objective is to address catastrophic forgetting, where models lose knowledge of earlier tasks when learning new ones [23]. In CL, models are trained on a sequence of tasks and are expected to perform well across all tasks in the sequence [24,25]. Unlike DG, CL assumes sequential access to target domain data over time.

Meta-Learning. Often described as “learning to learn”, meta-learning focuses on acquiring knowledge from multiple tasks to enable rapid adaptation to new tasks with limited data. Meta-learning typically optimizes a model’s initial parameters for quick adaptation to unseen tasks using only a few samples and gradient updates [26,27]. In DG, meta-learning has been applied by simulating “meta-training” and “meta-testing” tasks

through splitting multiple source domains. These approaches help extract domain-invariant features across source domains, enhancing generalization to unseen target domains and improving DG performance [28].

Other Approaches in DG. Beyond data augmentation, DG research has explored techniques such as **representation learning** and **learning strategies** to address distribution shifts. Representation learning focuses on constraining sample features to enable the model to learn domain-invariant representations, thereby improving robustness. Conversely, data augmentation modifies sample features, either in the input or feature space, to simulate potential variations in the target domain data. Examples of representation learning techniques include reducing feature redundancy [29,30], feature disentanglement [31–36], and distribution alignment [17,18,37–41].

Learning strategy methods aim to enhance the training process and optimization strategies to improve generalization. These include **ensemble learning** [42–45], **meta-learning techniques** [46–51], **flat minima optimization** [52–56], and **distillation techniques** [57–60].

Importantly, data augmentation can be combined with many of these approaches, as it is not mutually exclusive. However, certain techniques, such as adversarial networks with domain discriminators [61], face integration challenges. Adversarial networks rely on domain discriminators to classify the domain of each sample accurately, while the encoder must extract features that “fool” the discriminator [16], ensuring domain invariance. Since augmented samples cannot always be explicitly assigned to a specific domain, incorporating data augmentation into adversarial processes may require further adaptation. For instance, [62] proposes using unsupervised clustering to assign domain labels to augmented samples, enabling the integration of data augmentation with adversarial techniques. This approach illustrates the potential for combining data augmentation with adversarial strategies to improve DG performance.

3. Taxonomy of Data Augmentation Techniques for Domain Generalization

Taxonomy Overview. Data augmentation is pivotal in enhancing a model’s ability to generalize to unseen domains within the DG context. First, it simulates variations in the potential data distributions of unseen domains, allowing models to train on a broader range of domain scenarios. This exposure improves the model’s generalization performance on unknown domains. Second, data augmentation introduces perturbations to domain-specific features, compelling the model to focus on domain-invariant features. This focus strengthens the robustness of the extracted features across different domains.

To systematically categorize and analyze data augmentation methods in DG, we propose a taxonomy framework based on three critical dimensions. This framework provides a comprehensive perspective for understanding the similarities, differences, and applications of various data augmentation methods. The three dimensions are as follows:

- **Scope of Data Augmentation.** This dimension identifies where data augmentation is applied, distinguishing between methods that operate in the image space and those in the feature space. Methods in the image space modify raw input data to indirectly influence the high-level features extracted by the model, while feature space methods directly alter high-level features to improve generalization.
- **Nature of Data Augmentation.** This dimension reflects the principles underlying the augmentation methods. It includes gradient-based methods that utilize backpropagation, generative methods that generate synthetic data by altering latent representations, and rule-based methods that apply predefined transformations to modify the data.
- **Training Dependency.** This dimension evaluates the reliance of data augmentation methods on the training process within the DG pipeline. It categorizes methods into

those that can be integrated without requiring additional training, methods leveraging pretrained models directly for augmentation, and methods requiring the simultaneous training of augmentation-related parameters alongside the DG model.

This taxonomy framework not only provides a structured lens through which data augmentation methods in DG can be analyzed but also serves as a guide for effectively combining and leveraging their synergistic effects. Importantly, these three dimensions are not mutually exclusive. Each data augmentation method can be classified along all three dimensions and may span multiple subcategories within a single dimension.

3.1. Scope of Data Augmentation

We classify data augmentation methods into input-level augmentation and feature-level augmentation, based on whether modifications are applied in the image space or the model's feature space.

Input-Level Augmentation. Input-level augmentation refers to techniques that directly modify raw input images at the pixel level before model processing, expanding training data diversity while preserving semantic information. These methods can be categorized into **geometric transformations**, **appearance-based augmentations**, **frequency-based augmentations**, and **texture-based augmentations**. Geometric transformations modify the spatial properties of images to simulate variations in sensor perspectives, camera angles, and object sizes, thereby enhancing robustness to positional shifts; common techniques include cropping, flipping, rotation, scaling, shearing, translation, and warping. Appearance-based augmentations alter color, texture, and intensity variations to mimic real-world domain shifts, using techniques such as color jittering, noise injection, and style transfer. AugMix [63] and RandAugment [64] integrate multiple transformations to enhance domain diversity, while StyleAugment [65] leverages AdaIN-based feature statistics manipulation to introduce diverse style variations. GeomTex [66] further extends this by combining both texture-based and geometric style transfer using WarpST [67]. Frequency-based augmentations manipulate spectral properties to alter structural details as seen in Fourier-based methods like FACT [68] and CIRCLE [30], which interpolate amplitude components while preserving phase information, forcing the model to focus on phase-based domain-invariant features. VIPAug [69] perturbs both amplitude and phase, while MLRT [47] introduces Gaussian noise to low-frequency components, simulating realistic domain variations. Texture-based augmentations introduce controlled perturbations in texture and structure, such as RandConv [70], which applies random convolution kernels to generate diverse textures, though Pro-RandConv [71] improves upon this by using deeper kernels to retain meaningful features. Additionally, gradient-guided augmentation approaches such as [72–75] optimize augmentation parameters through adversarial loss maximization, producing more effective domain-agnostic samples. Input-level augmentation methods are widely adopted due to their simplicity, computational efficiency, and compatibility with various architectures, requiring no model modifications and providing visually interpretable transformations; however, if not carefully tuned, geometric- and appearance-based augmentations may introduce distortions that mislead the model, reducing its ability to learn domain-invariant representations and potentially hindering generalization.

Feature-Level Augmentation. Feature-level augmentation modifies intermediate representations within the model's processing pipeline rather than raw input images, directly influencing how features are extracted and learned. These methods are broadly categorized into **feature perturbation** and **adversarial feature perturbation**. Feature perturbation techniques introduce controlled variations in feature space to improve model robustness against distribution shifts. Methods such as SFA [76], ISDA [77], and CDSA [78] inject Gaussian noise by sampling scaling and bias values of feature activations from

Gaussian distributions, simulating real-world variations in intermediate representations. MixStyle [79], inspired by AdaIN, applies feature-wise style transfer directly to shallow-layer activations, generating domain-diverse representations without requiring decoders. Style Neophile [80] extends this concept by leveraging MMD distance and encoding rates to iteratively discover new style prototypes and performing style transfer on intermediate features, thereby increasing domain variability. DSU [81] models feature statistics with Gaussian distributions, enabling stochastic style transfer by sampling diverse domain variations, while CSU [82] captures inter-channel correlations, constructing multivariate Gaussian distributions across channels to generate more comprehensive feature transformations. Additionally, feature-level augmentation can be applied in the frequency domain, where models typically transition from capturing low-frequency features to high-frequency domain-specific variations [83]. In contrast, [84] highlights that low-frequency components encode critical style information. DFF [33] selectively suppresses high-frequency components using learnable filters, promoting the extraction of domain-invariant features, whereas ALOFT [85] perturbs low-frequency components with Gaussian noise, disrupting global texture information to encourage reliance on structural details. In adversarial feature perturbation, augmentation is guided by adversarial optimization to enforce domain invariance. Methods like DFP [86] introduce adversarial noise into feature activations, making it harder for domain discriminators to distinguish between domains, thereby encouraging the model to learn domain-agnostic representations. Similarly, RASP [87] estimates adversarial style parameters, such as feature-wise mean and variance, using a class discriminator and applies style transfer in feature space to generate adversarially perturbed feature representations, reducing the model's reliance on domain-specific cues. Feature-level augmentation provides greater flexibility than input-level augmentation by directly modifying internal representations, enabling stronger generalization to unseen domains; however, its effectiveness depends on the model's representation learning capability, and implementation often requires modifying the model architecture, making it more complex to integrate into existing pipelines.

3.2. Nature of Data Augmentation

The nature of a data augmentation method refers to the fundamental principle governing how the augmentation is applied. Based on these principles, data augmentation methods can be categorized into three main types: **gradient-based**, **generative**, and **rule-based** approaches. Each category represents a distinct mechanism for modifying data to enhance model generalization across diverse domains.

Gradient-Based Augmentation. Gradient-based augmentation methods generate augmented samples by leveraging gradient information obtained through backpropagation. By computing gradients with respect to model loss functions, these approaches modify input or feature representations in a direction that enhances generalization. These methods typically create perturbations by maximizing task-specific classification loss or domain classification loss, followed by an adversarial optimization process to improve the model's ability to generalize to unseen domains. CrossGrad [72] generates adversarial samples in the input space by maximizing the loss of both a category classifier and a domain classifier, where adversarial samples targeting the category classifier are assumed to contain minimal category-specific information and are used to train the domain classifier, while those targeting the domain classifier are assumed to lack domain-specific information and are employed to train the category classifier, enabling the category classifier to extract domain-invariant features. Other methods, such as TeacherAugment [73] and MODE [74], apply geometric transformations and appearance-based augmentations in the input space, where task classification loss is maximized to learn augmentation parameters that control

transformation intensity, effectively simulating potential domain distributions that similarly increase classification loss. Unlike RandConv [70] and Pro-RandConv [71], which apply randomly initialized convolution kernels to image pixels, ASRConv [88] operates in the frequency domain, updating convolution kernel parameters in the direction that maximizes task loss. Gradient-based methods also extend to feature space perturbations, such as DFP [86], which introduces adversarial perturbations in the feature space of the target model to increase the loss of a domain discriminator, forcing the model to focus on domain-invariant features. Similarly, [89] minimizes mutual information between the generated samples and their original counterparts to capture the styles of potential domains, while RASP [87] performs style transfer in the feature space using styles that maximize the loss of a class discriminator, simulating style variations encountered in unseen domains. DFDG [90] employs SmoothGrad saliency maps [91] to identify class-discriminative regions within images, adding Gaussian noise to areas outside these regions, thereby forcing the model to focus on domain-invariant, class-discriminative features. Gradient-based data augmentation methods provide a strong theoretical foundation by explicitly aligning the augmentation direction with the target task's objectives; however, leveraging gradient information to guide augmentation introduces additional computational costs due to the backpropagation operations required for gradient computation and optimization.

Generative Augmentation. Generative augmentation methods leverage pretrained or concurrently trained generative models to produce augmented data, introducing diversity that enables models to learn features better aligned with unseen domains. These methods can be categorized into **GAN-based, autoencoder-based, and diffusion-based** approaches, depending on the underlying generative model. Generative Adversarial Networks (GANs) [92] employ adversarial training between a generator and a discriminator to learn the distribution of input images and generate synthetic data, with advancements such as CycleGAN [93] and BigGAN [94] refining this process to improve sample quality and diversity. For DG, [95] applies GANs by fine-tuning separate StyleGAN2 models on each source domain and blending their parameters to generate data from diverse potential domains. Autoencoders [96] are self-supervised generative models that map input data into a lower-dimensional latent space before reconstructing them through a decoder. Methods such as L2A-OT [97] and GINet [98] maximize the Wasserstein distance [99] between generated samples and original domain data while minimizing classification loss, enabling the generator to map source domain data to potential unknown domains. VDN [100] enhances this by combining category-specific and domain-specific features to create novel samples with randomly selected domain-specific characteristics. Style transfer methods such as StyleAugment [65], EFDM [101], and DAI [102] perform style transformations on encoded features within autoencoder frameworks, generating images with diverse styles, while [103] applies feature-level style transfer to simulate potential domain shifts. Diffusion models generate high-quality images by iteratively denoising random noise [104]; ED-SAM [105] extends this by introducing perturbations to latent vectors after the final noise addition step, producing images with diverse stylistic semantics. CDGA [106] guides a pretrained diffusion model using images or text prompts from different source domains to generate new data, while DINDEX [107] employs a diffusion model to create images from potential domains while aligning their features with domain distributions through unsupervised domain adaptation. Additionally, FDS [108] fine-tunes pretrained diffusion models for specific tasks, interpolating text prompt encodings from various source domains to guide the generation of diverse domain samples. Generative augmentation methods offer significant advantages over gradient-based and rule-based approaches by producing diverse samples that help models learn richer and more domain-invariant representations. However, these methods introduce additional computational and storage costs during both

training and inference. Furthermore, generative models are challenging to train effectively, and suboptimal training can compromise augmentation effectiveness.

Rule-Based Augmentation. Rule-based augmentation methods apply predefined, non-parametric transformations to modify data at either the input or feature level, introducing perturbations that enhance model robustness against domain-specific variations. Unlike gradient-based and generative approaches, these methods operate independently of backpropagation and learning mechanisms, relying instead on explicit transformation rules to determine intensity and direction factors. Common techniques include **geometric transformations**, such as cropping, rotation, and flipping; **appearance-based transformations**, such as noise injection, frequency augmentation, and style transfer; and **convolutional perturbations** that alter texture characteristics. In frequency-based augmentation, predefined rules modify frequency components of images to simulate domain shifts. For instance, APR [109] swaps the amplitudes of images within a batch while preserving phase information, whereas FACT [68] interpolates amplitudes across batch samples. VIPAug [69] extends APR by adding Gaussian noise to the phases of low-amplitude frequencies or randomly substituting them with phases from other images, further increasing diversity. In style transfer-based augmentation, MixStyle [79] applies AdaIN transformations to intermediate features, blending styles from different batch samples to simulate novel domain styles. Style Neophile [80] refines this idea by identifying new styles using MMD distance and coding rate metrics, constructing Gaussian distributions over styles, and sampling from these distributions to perform style transfer on intermediate features. Similarly, DSU [81] and CSU [82] model the distribution of style statistics within a batch and generate diverse samples by sampling new style variations. In contrast, CPerb [110] introduces a localized augmentation strategy by performing patch-based transformations, where patches within the same image randomly adopt style characteristics from different feature channels, effectively simulating potential domain shifts. Rule-based augmentation offers a computationally efficient approach to enhancing data diversity without the need for complex model training. However, compared to gradient-based and generative augmentation methods, rule-based approaches are inherently limited in their ability to generate highly diverse and realistic domain variations.

3.3. Training Dependency

Data augmentation methods can be categorized based on their dependency on training into three main types: **training-free transformations**, **pretrained transformations**, and **training-simultaneous transformations**.

Training-Free Transformations. These methods do not require additional parameters or models specifically trained for data augmentation. They modify input- or feature-level data using predefined rules to generate diverse samples. Common techniques include standard transformations such as cropping, rotation, flipping, and color jitter. More advanced techniques combine multiple transformations as demonstrated by AugMix [63] and RandAugment [64]. Other approaches involve perturbations in the frequency domain, such as APR [109], FACT [68], CIRCLE [30], MLRT [47], and VIPAug [69], or feature-space modifications using AdaIN-based transformations like MixStyle [79], Style Neophile [80], DSU [81], CPerb [110], and CSU [82]. Additionally, image perturbations via random convolutions, as in RandConv [70] and Pro-RandConv [71], generate texture variations. Since these methods require no additional training, they incur no computational cost during training and can be seamlessly integrated into existing pipelines. However, they are inherently limited in generating diverse or task-specific domain shifts.

Pretrained Transformations. These data augmentation methods utilize off-the-shelf generative models to produce diverse augmented samples. They leverage generative mod-

els trained on external datasets, introducing diversity aligned with the target task without requiring additional model training. Common examples include pretrained style transfer models applied to input-level augmentation as demonstrated by [111], StyleAugment [65], GeomTex [66], and EFDm [101], as well as pretrained diffusion models in methods like ED-SAM [105], CDGA [106], and DIDE [107]. By directly leveraging pretrained models, these methods bypass the need for additional training while offering relatively high diversity. However, their effectiveness depends on the quality of the pretrained generative model and its alignment with the domain characteristics of the target task.

Training-Simultaneous Transformations. These transformations involve learning domain-specific data augmentation models during the training of the target task. They often require training additional models, fine-tuning generative models, or optimizing augmentation parameters dynamically using gradient information. For instance, methods like that in [95] and FDS [108] fine-tune pretrained generative models on the target task to align generated samples with potential domain shifts, while [103] trains a style-mapping function that transforms style statistics in feature space using AdaIN. Similarly, L2A-OT [97], VDN [100], GINet [98], and DAI [102] train generative models alongside the primary model to produce domain-adaptive augmented samples. UniFreqSDG [112] introduces low-frequency perturbations using Gaussian noise to simulate realistic distribution shifts, while gradient-based methods such as CrossGrad [72], DFDG [90], TeacherAugment [73], ASRConv [88], MODE [74], ALT [75], DFP [86], and RASP [87] use loss gradients to guide augmentation strategies. These methods allow augmentation to adapt dynamically to the model's learning process, leading to more realistic domain variations. However, they introduce higher computational costs due to backpropagation and optimization overhead, and they may cause model instability if augmentation parameters are not well regularized.

3.4. Comparative Summary of Taxonomy

This subsection provides a comprehensive comparison of data augmentation methods for domain generalization, classified according to the taxonomy dimensions discussed in previous sections. The table summarizes the scope (input-level vs. feature-level), nature (rule-based, gradient-based, or generative), and training dependency (training-free, pre-trained, or training-simultaneous) of each method. This summary enables readers to better understand the strengths, limitations, and applications of these methods. See Tables 1 and 2 for a detailed overview.

Table 3 summarizes the performance improvements of various data augmentation techniques, presenting the average accuracy of ResNet-18-based domain generalization methods on the PACS [113] and Office-Home [114] datasets. These datasets are selected due to their widespread use as standard benchmarks for evaluating cross-domain robustness: PACS encompasses distinct artistic styles (photo, art, cartoon, and sketch), while Office-Home includes diverse real-world environments (art, clipart, product, and real). The use of ResNet-18 ensures fair comparisons by maintaining consistent model capacity, reflecting its common adoption as a baseline architecture in DG studies. Among input-level methods, CDGA [106] achieves top accuracy on both PACS (88.4%) and Office-Home (70.2%) by leveraging diffusion models to improve cross-domain generalization. In feature-level augmentation, CDSA [78] records the highest scores on PACS (89.3%) and Office-Home (73.0%) through a hybrid strategy that integrates spectral perturbations, inter-class semantic augmentation, and contrastive learning. Both approaches significantly outperform the baseline ResNet-18, which achieves 78.3% and 63.9% accuracy on PACS and Office-Home, respectively, highlighting the effectiveness of data augmentation in enhancing domain generalization.

Table 1. Classification of data augmentation methods for domain generalization across taxonomy dimensions.

Scope	Methods	Nature	Td	Methods	Nature	Td
Input	VIPAug [69]	R	F	CrossGrad [72]	Gr	S
	RCT [115]	R	F	ED-SAM [105]	Ge	P
	Pro-RandConv [71]	R	F	CDGA [106]	Ge	P
	APR [109]	R	F	DIDEX [107]	Ge	P
	FACT [68]	R	F	GeomTex [66]	Ge	P
	RandAugment [64]	R	F	EFDM [101]	Ge	P
	RandConv [70]	R	F	DAI [102]	Ge	S
	AugMix [63]	R	F	FDS [108]	Ge	S
	ALT [75]	Gr	S	[95]	Ge	S
	MODE [74]	Gr	S	GINet [98]	Ge	S
	TeacherAugment [73]	Gr	S	VDN [100]	Ge	S
	AdvStyle [116]	Gr	S	DIVA [117]	Ge	S
	[89]	Gr	S	L2A-OT [97]	Ge	S
	DFDG [90]	Gr	S	DDAIG [118]	Ge	S
Feature	TFS-ViT [108]	R	F	DSU [81]	R	F
	START [119]	R	F	SFA [76]	R	F
	CPerb [110]	R	F	MixStyle [79]	R	F
	CSU [82]	R	F	DFP [86]	Gr	S
	MRFP [120]	R	F	RASP [87]	Gr	S
	CDSA [78]	R	F	ASRConv [88]	Gr	S
	ALOFT [85]	R	F	DFF [33]	Gr	S
	Style Neophile [80]	R	F	[103]	Ge	S

Abbreviations : Td: Training dependency; R: Rule Based; Gr: Gradient Based; Ge: Generative; P: Pretrained; F: Training Free; S: Training Simultaneous.

Table 2. Strengths, weaknesses, and use cases of input-level vs. feature-level augmentation.

Scope	Strengths	Weakness	Use Cases
Input	Simple to implement, no model modification needed, visually interpretable	Limited control over feature variations, potential distortions	Image-based tasks (e.g., object classification and medical imaging)
Feature	Directly manipulates learned representations, better generalization	Requires model modifications, harder to visualize, dependency on model's representation learning capability	Feature-driven tasks (e.g., NLP, speech and ViT-based models)

Table 3. Leave-one-domain-out generalization performance of ResNet-18 on the PACS and Office-Home datasets.

Scope	Methods	PACS avg.	Office-Home avg.
BaseLine	ERM	78.3	63.9
Input	CrossGrad [72]	80.7	64.4
	DDAIG [118]	83.1	65.5
	L2A-OT [97]	82.8	65.6
	FACT [68]	84.5	66.5
	CIRL [30]	86.3	67.1
	MODE [74]	86.9	66.9
	GINet [98]	83.7	66.9
	GeomTex [66]	86.6	66.8
	Pro-RandConv [71]	84.3	64.6
	CDGA [106]	88.4	70.2
Feature	MixStyle [79]	83.7	65.5
	RASP [87]	84.7	67.3
	Style Neophile [80]	85.5	65.9
	CDSA [78]	89.3	73.0
	DSU [81]	84.1	66.1
	CSU [82]	85.2	66.8
	[103]	86.0	66.6
	DFP [86]	83.0	63.7

3.5. Discussion of The Taxonomy Design

The proposed taxonomy framework offers a comprehensive and adaptable categorization scheme for data augmentation methods in domain generalization. By introducing three key dimensions—scope of augmentation, nature of augmentation, and training dependency—it provides a structured lens through which to explore the similarities, differences, and interactions among various techniques.

Input-level augmentation operates directly on raw pixel data, making it easy to implement across various models without requiring architectural changes. These methods are computationally efficient and provide visually interpretable transformations. However, they are often limited in addressing higher-level feature variations and may introduce visual distortions that do not accurately represent domain shifts. In contrast, feature-level augmentation manipulates internal model representations, allowing more targeted adjustments to domain-specific features. While this approach generally yields superior generalization performance, it necessitates modifications to the model architecture and is often more complex to implement.

The **scope of augmentation** dimension emphasizes how augmentation methods integrate into the training pipeline, distinguishing between input-level and feature-level augmentations. This clarification highlights the stages of the pipeline impacted by different methods. The **nature of augmentation** dimension categorizes techniques into rule-based, gradient-based, and generative approaches, emphasizing the diversity in augmentation mechanisms. Lastly, the **training dependency** dimension examines the relationship between augmentation and training processes, classifying methods as training-free transformations, pretrained transformations, or training-simultaneous transformations. This dimension reveals the varying degrees of interaction between augmentation strategies, model optimization, and task datasets.

Certain methods, such as style transfer and frequency augmentation, naturally transition between categories within a single dimension. For instance, in the scope of augmentation, style transfer may occur at the input level [65] or the feature level [79,80]. Similarly,

within the nature of augmentation, style-related parameters can be predefined [65,81] or derived from gradient information [87,89]. Although these methods may span multiple categories within a single dimension, their shared objective is expanding data diversity. Techniques such as [75,78,110] further maximize data perturbation by incorporating strategies from multiple categories, effectively simulating potential unseen domain distributions.

Training-simultaneous transformations exemplify the dynamic interaction between data augmentation and model optimization. Here, augmentation parameters are iteratively adjusted during training to align with the model's current learning state. Unlike static transformations, which uniformly apply changes to the original data, this adaptive mechanism generates samples tailored to the model's representations. This dynamic interaction is particularly valuable for domain generalization, as it produces samples with enhanced inter-domain variability, enabling models to generalize effectively to unseen test domains.

Our future work will focus on extending augmentation methods across the proposed taxonomy dimensions and designing hybrid strategies that combine rule-based, generative, and gradient-based mechanisms. Such hybrid approaches aim to harness the strengths of each method, enhancing their efficiency, adaptability, and robustness in addressing domain shift challenges across diverse scenarios.

4. Applications of Data Augmentation in Domain Generalization

In this section, we explore the application of data augmentation methods in various research areas to address domain generalization challenges.

Computer Vision. In computer vision tasks such as object classification, object detection, image segmentation, and person re-identification, domain shifts often arise due to variations in the lighting, viewpoints, and background conditions of input images. In medical imaging, differences in imaging devices and patient conditions can lead to changes in image attributes such as brightness, contrast, and cell morphology. To address domain-specific challenges in various computer vision tasks, researchers have developed diverse data augmentation strategies tailored to the unique requirements of each application. In object detection, studies such as [47,121] randomize domain-specific information in the frequency space, transforming single or multiple source domains into diverse new domains to enhance model performance. For image segmentation, [122] treats style features as learnable parameters and updates them through adversarial training, enabling the generation of adversarial images that improve model robustness. Similarly, [116] perturbs fine-grained features using random convolutions and applies AdaIN for style transfer on coarse features to create diverse augmented data. Building on this approach, [120] introduces AdaIN-based feature perturbation to generate style-diverse samples, further enhancing segmentation performance. In person re-identification, [33] improves generalization by employing learnable filters to suppress high-frequency features, reducing reliance on domain-specific information. For medical imaging, [88] utilizes gradient-guided convolution layers to enrich image features, increasing feature diversity and mitigating overfitting to specific domains. Additionally, works like [123,124] leverage GANs to generate varied synthetic data, thereby strengthening model robustness across diverse medical datasets. Figure 1 visually illustrates the impact of geometric transformations on the PACS object recognition dataset. The visualization shows that data augmentation facilitates clearer inter-class separations in the model's feature representations, ultimately enhancing its generalization ability to unseen domains.

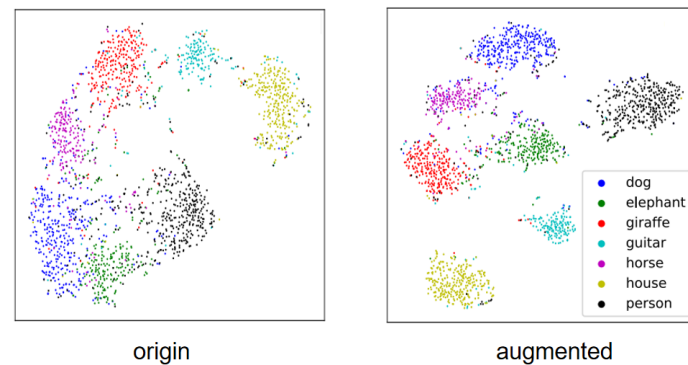


Figure 1. Comparison of model representations before and after image augmentation using geometric transformations in GeomTex.

Natural Language Processing. In natural language processing (NLP), domain generalization challenges impact tasks such as sentiment analysis, semantic parsing, and logical reasoning. These challenges often stem from domain shifts caused by variations in the distribution of training data, influenced by factors such as the medium or subject matter of the text. For instance, texts collected from books typically use formal language, while social media texts frequently incorporate emojis and colloquial expressions. Such distributional differences can significantly affect model performance. To mitigate these challenges, researchers have developed various data augmentation techniques tailored to text-based tasks. Common methods include synonym replacement, random deletion, word insertion, spelling and grammar perturbations, and style transformations. For instance, [125] fine-tunes a generative model to produce in-domain, fine-grained sentences, thereby enhancing model accuracy. Similarly, [126] increases data diversity by incorporating special character insertion and random entity replacement during back-translation. Furthermore, [127] leverages large language models (LLMs), such as GPT-4, to generate synthetic dialogue data, expanding the training set and improving model robustness.

Speech Analysis. Domain generalization challenges are prevalent in speech analysis, impacting tasks such as speech recognition, emotion classification, and speech synthesis. These challenges arise from distributional shifts in training data caused by factors such as environmental noise, device characteristics, application scenarios, speaker tone, and speech rate. Such shifts significantly affect the model's performance in target domains. To mitigate these issues, common data augmentation techniques for speech data include adding background noise, adjusting volume, applying pitch shifts, and performing time scaling. Such methods are widely employed to enhance generalization in speech analysis tasks. For instance, [128] leverages prosodic features to improve pronunciation modeling, enabling more adaptive representations of diverse accents. Similarly, [129] employs speed perturbation and trains a GAN-based model to augment the dataset, thereby enhancing generalization in dysarthric speech recognition. In another approach, [130] integrates reinforcement learning, where a policy controller selects optimal temporal and spectral masking operations for data augmentation, resulting in improved model robustness and performance.

Robotics and Autonomous Systems. In robotics and autonomous systems, tasks such as path planning, object recognition, semantic navigation, and decision-making in dynamic environments are significantly impacted by domain generalization challenges. These challenges arise due to variations in data collection environments, hardware differences, and sensor characteristics, which can hinder model performance during deployment. Common data augmentation methods in this field include adding sensor noise, applying geometric transformations, introducing random occlusions, and simulating various lighting conditions. For instance, [131] employs a simulation environment with randomized,

non-realistic textures to generate diverse training data, enhancing the model's ability to generalize to varying visual conditions. Similarly, [132] introduces noise into source trajectories and applies a pretrained diffusion model for denoising, effectively aligning the trajectories with target domain dynamics. This approach enables models to better adapt to diverse and unpredictable environmental conditions encountered during real-world deployment.

5. Challenges and Open Problems

In this section, we discuss the challenges and open problems in applying data augmentation for domain generalization. These challenges hinder the widespread adoption and effectiveness of data augmentation techniques across various fields.

Computational Cost. While data augmentation methods enhance model robustness in domain generalization tasks, they often come with increased computational and storage requirements. The associated costs vary depending on the type of data augmentation method used. Generative augmentation methods are computationally expensive due to both the training and inference phases of generative models. Gradient-based methods require additional computational resources due to the backpropagation process needed to optimize augmentation parameters. Even some rule-based methods, while simpler, demand computations for heuristic metrics. Additionally, augmented data increase storage requirements, further exacerbating resource demands. These factors contribute to longer training times and higher hardware requirements, presenting a significant barrier to the practical deployment of data augmentation methods, particularly in resource-constrained environments. To address the high computational costs associated with data augmentation, recent research has focused on developing parameter-efficient augmentation strategies. AutoAugment [133] introduces an automated search framework for augmentation policies but is hindered by substantial computational demands. In contrast, RandAugment [64] improves both efficiency and generalizability by simplifying the search space and employing randomized policy selection, making it a more practical solution for real-world applications. Additionally, [134] leverages latent space diffusion to accelerate data generation processes, while [135] enhances model fine-tuning efficiency through the use of low-rank matrix updates. These approaches collectively contribute to reducing computational overhead, making data augmentation more accessible for deployment in environments with limited resources.

Optimal Augmentation Strategies. Despite the advantages of data augmentation for domain generalization, identifying the optimal strategy for a given task remains an open problem. Current domain generalization research lacks a unified metric to assess the effectiveness of different augmentation strategies. Moreover, augmentation techniques that enhance generalization for one task may introduce irrelevant or overly complex latent domain features in another, potentially degrading performance. Another crucial challenge is determining the appropriate augmentation intensity. The distribution of augmented data must align with the original task domain's distribution. Insufficient augmentation intensity may fail to capture potential domain shifts, while excessive augmentation could introduce synthetic artifacts that diverge too far from the target domain's true distribution. Developing adaptive augmentation strategies that dynamically adjust based on task characteristics and domain shifts is a key research direction. Recent efforts have explored various methods to achieve optimal augmentation strategies. Beyond gradient-guided approaches [73,74] and strategy filtering methods, [108,136] introduces reinforcement learning to dynamically select augmentation policies. This approach enables the model to adapt augmentation strategies based on task-specific characteristics and observed domain shifts, thereby improving generalization while avoiding unnecessary complexity or synthetic artifacts.

Overfitting to Augmentation. Poorly designed data augmentation strategies can lead to overfitting, ultimately compromising generalization performance on unseen domains. A major challenge in domain generalization is reducing the model's dependence on domain-specific features. If augmented data deviate significantly from the true distribution, they may introduce noise or irrelevant characteristics. This can cause models to learn spurious features rather than task-relevant ones, thereby reducing performance on unseen target domains [137]. Achieving a balance between data diversity and distribution alignment is critical. Future research should focus on developing robust training strategies that mitigate overfitting to augmented data while preserving the model's ability to generalize effectively.

6. Emerging Trends and Future Directions

Augmentation for Specific Architectural Models. In recent years, advanced architectures such as Vision Transformers (ViT) [138] and Vision Mamba (ViM) [139] have demonstrated remarkable success in computer vision tasks, sparking a growing interest in their application to domain generalization (DG) challenges. This has led to the development of data augmentation strategies specifically tailored to these unique architectures.

Both ViT and ViM process input images by dividing them into patches, encoding these patches into tokens, and using mechanisms such as Transformers [140] or Mambas [141] to capture global information and generate feature-level representations. Recent studies, such as [108,119], have explored style transfer techniques applied at the feature-level token stage, enabling these models to capture domain-invariant features and enhancing their generalization to unseen domains.

Augmentation strategies designed for advanced architectures like ViT and ViM capitalize on their strengths in modeling long-range dependencies and expressing domain-agnostic global information. By leveraging these capabilities, these methods improve the robustness and generalization performance of state-of-the-art models.

Cross-Modality Augmentation. With the growing adoption of advanced architectures like Transformers and Mamba, the use of multimodal models in data augmentation has attracted significant attention. A prominent example is the Stable Diffusion model [134], an extension of diffusion models that integrates text prompts into the image generation process. By utilizing text prompts to semantically guide the generation, Stable Diffusion achieves a high degree of alignment between the generated images and their corresponding textual descriptions. Additionally, diffusion models can incorporate reference images as inputs, further enhancing semantic guidance during the generation of target images.

CDGA [106] employs training domain names as text prompts and uses different domain images to generate new data. Similarly, [108] encodes training domain names, mixes the resulting textual encodings, and guides the diffusion model to create data for potential domains through text-to-visual augmentation. Furthermore, [107] constructs a library of text prompts to describe potential domains and utilizes these prompts to generate corresponding domain images via diffusion models.

Cross-modal augmentation methods hold immense promise by integrating information from different modalities to create diverse data tailored to unseen domains. Approaches that synchronize text and image generation introduce innovative possibilities for data augmentation. In the future, the development of robust cross-modal data augmentation mechanisms leveraging multimodal models will be a pivotal area of research.

7. Conclusions

Data augmentation has emerged as a crucial tool in domain generalization, enabling models to mitigate domain shifts and learn domain-invariant features, thereby enhancing robustness and performance in unseen environments. This survey provides a compre-

hensive review of data augmentation methods for DG in computer vision, introducing a taxonomy framework based on scope, nature, and training dependency to guide the design of effective strategies.

While data augmentation has achieved notable success, several challenges remain, including high computational costs, the absence of unified evaluation metrics, and the risk of overfitting. Addressing these issues is essential to further improving the applicability and efficiency of augmentation techniques. At the same time, the emergence of novel architectures such as Transformers and Mamba, along with advancements in cross-modal models, presents exciting opportunities to develop more flexible and robust augmentation strategies. A key direction for future research lies in developing standardized benchmarks to evaluate the effectiveness of cross-modality augmentation methods. Such benchmarks are crucial for assessing multimodal generalization and ensuring fair comparisons across different approaches. Establishing comprehensive evaluation frameworks will not only advance data augmentation strategies but also facilitate the development of next-generation AI systems capable of robust performance across diverse domains and modalities.

We hope this survey provides valuable insights and inspires further innovation and progress in the field of domain generalization.

Author Contributions: Resources, J.B.; Writing—original draft, J.M.; Writing—review and editing, C.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Natural Science Foundation of China (12126609, U1936116).

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Quionero-Candela, J.; Sugiyama, M.; Schwaighofer, A.; Lawrence, N. *Dataset Shift in Machine Learning*; MIT Press: Cambridge, MA, USA, 2008. <https://doi.org/10.7551/mitpress/9780262170055.001.0001>.
2. Khosla, A.; Zhou, T.; Malisiewicz, T.; Efros, A.A.; Torralba, A. Undoing the damage of dataset bias—Proceedings, Part I 12. In Proceedings of the Computer Vision—ECCV 2012: 12th European Conference on Computer Vision, Florence, Italy, October 7–13, 2012; pp. 158–171.
3. Xie, Q.; Li, Y.; He, N.; Ning, M.; Ma, K.; Wang, G.; Lian, Y.; Zheng, Y. Unsupervised domain adaptation for medical image segmentation by disentanglement learning and self-training. *IEEE Trans. Med. Imaging* **2022**, *43*, 4–14.
4. Schwonberg, M.; Niemeijer, J.; Termöhlen, J.A.; Schmidt, N.M.; Gottschalk, H.; Fingscheidt, T.; et al. Survey on unsupervised domain adaptation for semantic segmentation for visual perception in automated driving. *IEEE Access* **2023**, *11*, 54296–54336.
5. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 1–48.
6. Wang, J.; Perez, L. The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Netw. Vis. Recognit* **2017**, *11*, 1–8.
7. Zhou, K.; Liu, Z.; Qiao, Y.; Xiang, T.; Loy, C.C. Domain generalization: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 4396–4415.
8. Wang, J.; Lan, C.; Liu, C.; Ouyang, Y.; Qin, T.; Lu, W.; Chen, Y.; Zeng, W.; Philip, S.Y. Generalizing to unseen domains: A survey on domain generalization. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 8052–8072.
9. Kumar, T.; Brennan, R.; Mileo, A.; Bendeche, M. Image data augmentation approaches: A comprehensive survey and future directions. *IEEE Access* **2024** <https://doi.org/10.1109/ACCESS.2024.3470122>.
10. Wang, M.; Deng, W. Deep visual domain adaptation: A survey. *Neurocomputing* **2018**, *312*, 135–153.
11. Liu, X.; Yoo, C.; Xing, F.; Oh, H.; El Fakhri, G.; Kang, J.W.; Woo, J. Deep unsupervised domain adaptation: A review of recent advances and perspectives. *APSIPA Trans. Signal Inf. Process.* **2022**, *11*, 1.
12. Li, B.; Wang, Y.; Zhang, S.; Li, D.; Keutzer, K.; Darrell, T.; Zhao, H. Learning invariant representations and risks for semi-supervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 1104–1113.
13. Ganin, Y.; Lempitsky, V. Unsupervised domain adaptation by backpropagation. In Proceedings of the International Conference on Machine Learning. PMLR, Lille, France, 7–9 Jul 2015; pp. 1180–1189.

14. Long, M.; Cao, Y.; Wang, J.; Jordan, M. Learning transferable features with deep adaptation networks. In Proceedings of the International Conference on Machine Learning, PMLR, Lille, France, 7–9 July 2015; pp. 97–105.
15. Li, H.; Pan, S.J.; Wang, S.; Kot, A.C. Domain generalization with adversarial feature learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5400–5409.
16. Albuquerque, I.; Monteiro, J.; Darvishi, M.; Falk, T.H.; Mitliagkas, I. Generalizing to unseen domains via distribution matching. *arXiv preprint* **2019**, arXiv:1911.00804.
17. Zhu, W.; Lu, L.; Xiao, J.; Han, M.; Luo, J.; Harrison, A.P. Localized adversarial domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–24 June 2022; pp. 7108–7118.
18. Nguyen, T.; Do, K.; Duong, B.; Nguyen, T. Domain Generalisation via Risk Distribution Matching. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2024; pp. 2790–2799.
19. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A comprehensive survey on transfer learning. *Proc. IEEE* **2020**, *109*, 43–76.
20. Agarwal, N.; Sondhi, A.; Chopra, K.; Singh, G. Transfer learning: Survey and classification. In *Smart Innovations in Communication and Computational Sciences: Proceedings of International Conference on Smart Innovations in Communication and Computational Sciences*, online, 7–9 April 2020; Springer: Berlin/Heidelberg, Germany, 2021; pp. 145–155.
21. Zhang, Y.; Yang, Q. A survey on multi-task learning. *IEEE Trans. Knowl. Data Eng.* **2021**, *34*, 5586–5609.
22. Vandenhende, S.; Georgoulis, S.; Proesmans, M.; Dai, D.; Van Gool, L. Revisiting multi-task learning in the deep learning era. *arXiv preprints* **2020**, arXiv:2004.13379.
23. Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A.A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. Overcoming catastrophic forgetting in neural networks. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3521–3526.
24. Zhang, Z.; Ning, G.; Cen, Y.; Li, Y.; Zhao, Z.; Sun, H.; He, Z. Progressive neural networks for image classification. *arXiv preprint* **2018**, arXiv:1804.09803.
25. Parisi, G.I.; Kemker, R.; Part, J.L.; Kanan, C.; Wermter, S. Continual lifelong learning with neural networks: A review. *Neural Netw.* **2019**, *113*, 54–71.
26. Finn, C.; Abbeel, P.; Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. In Proceedings of the International Conference on Machine Learning PMLR, Sydney, Australia, 6–11 August 2017; pp. 1126–1135.
27. Baik, S.; Choi, J.; Kim, H.; Cho, D.; Min, J.; Lee, K.M. Meta-learning with task-adaptive loss function for few-shot learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Online, 11–17 October 2021; pp. 9465–9474.
28. Khoei, A.G.; Yu, Y.; Feldt, R. Domain generalization through meta-learning: A survey. *Artif. Intell. Rev.* **2024**, *57*, 285.
29. Zhang, X.; Cui, P.; Xu, R.; Zhou, L.; He, Y.; Shen, Z. Deep Stable Learning for Out-Of-Distribution Generalization. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 5368–5378. <https://doi.org/10.1109/cvpr46437.2021.00533>.
30. 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, New Orleans, LO, USA, 21–24 June 2022; pp. 8046–8056.
31. Nam, H.; Lee, H.; Park, J.; Yoon, W.; Yoo, D. Reducing Domain Gap by Reducing Style Bias. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 8686–8695. <https://doi.org/10.1109/cvpr46437.2021.00858>.
32. Zhang, H.; Zhang, Y.; Liu, W.; Weller, A.; Schölkopf, B.; Xing, E. Towards Principled Disentanglement for Domain Generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.
33. Lin, S.; Zhang, Z.; Huang, Z.; Lu, Y.; Lan, C.; Chu, P.; You, Q.; Wang, J.; Liu, Z.; Parulkar, A.; et al. Deep frequency filtering for domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 11797–11807.
34. Chen, Y.; Wang, Y.; Pan, Y.; Yao, T.; Tian, X.; Mei, T. A Style and Semantic Memory Mechanism for Domain Generalization. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021.
35. Jin, X.; Lan, C.; Zeng, W.; Chen, Z. Style Normalization and Restitution for Domain Generalization and Adaptation. *IEEE Trans. Multimed.* **2022**, *24*, 3636–3651. <https://doi.org/10.1109/tmm.2021.3104379>.
36. Chen, H.; Zhang, Q.; Huang, Z.; Wang, H.; Zhao, J. Towards domain-specific features disentanglement for domain generalization. *arXiv* **2023**, arXiv:2310.03007.
37. Pei, S.; Sun, J.; Da Xu, R.Y.; Xiang, S.; Meng, G. Domain decorrelation with potential energy ranking. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 9–14 February 2023; Volume 37, pp. 2020–2028.
38. Li, T.; Qiao, F.; Ma, M.; Peng, X. Are Data-driven Explanations Robust against Out-of-distribution Data? In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 18–22 June 2023; pp. 3821–3831.

39. Chen, L.; Zhang, Y.; Song, Y.; Van Den Hengel, A.; Liu, L. Domain generalization via rationale invariance. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 1751–1760.
40. Huang, Z.; Wang, H.; Zhao, J.; Zheng, N. IDAG Invariant dag searching for domain generalization. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 19169–19179.
41. Zhang, X.; Su, H.; Liu, X. Graph convolutional network for adversarial domain generalization. *IEEE Trans. Comput. Soc. Syst.* **2024**. <https://doi.org/10.1109/TCSS.2024.3367972>.
42. Huang, K.; Ren, Z.; Zhu, L.; Lin, T.; Zhu, Y.; Zeng, L.; Wan, J. Intra-domain self generalization network for intelligent fault diagnosis of bearings under unseen working conditions. *Adv. Eng. Inform.* **2025**, *64*, 102997.
43. Qu, S.; Pan, Y.; Chen, G.; Yao, T.; Jiang, C.; Mei, T. Modality-agnostic debiasing for single domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 24142–24151.
44. Yu, G.; Hwang, H. A2XP: Towards Private Domain Generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 23544–23553.
45. Li, B.; Shen, Y.; Yang, J.; Wang, Y.; Ren, J.; Che, T.; Zhang, J.; Liu, Z. Sparse mixture-of-experts are domain generalizable learners. *arXiv* **2022**, arXiv:2206.04046.
46. Zhang, L.; Liu, Z.; Zhang, W.; Zhang, D. Style uncertainty based self-paced meta learning for generalizable person re-identification. *IEEE Trans. Image Process.* **2023**, *32*, 2107–2119.
47. Zhang, L.; Qin, L.; Xu, M.; Chen, W.; Pu, S.; Zhang, W. Randomized Spectrum Transformations for Adapting Object Detector in Unseen Domains. *IEEE Trans. Image Process.* **2023**, *32*, 4868–4879.
48. Li, M.; Wang, Z.; Hu, X. Restoration towards decomposition: A simple approach for domain generalization. *Inf. Sci.* **2024**, *679*, 121053.
49. Lv, F.; Liang, J.; Li, S.; Zhang, J.; Liu, D. Improving generalization with domain convex game. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 24315–24324.
50. Zhang, J.; Qi, L.; Shi, Y.; Gao, Y. Mvdg: A unified multi-view framework for domain generalization. In Proceedings of the European Conference on Computer Vision. Springer, Tel Aviv, Israel, 23–27 October 2022; pp. 161–177.
51. Huang, C.; Cao, Z.; Wang, Y.; Wang, J.; Long, M. Metasets: Meta-learning on point sets for generalizable representations. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 8863–8872.
52. Cha, J.; Chun, S.; Lee, K.J.; Cho, H.C.; Park, S.H.; Lee, Y.; Park, S. SWAD: Domain Generalization by Seeking Flat Minima. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 22405–22418.
53. Zhang, X.; Xu, R.; Yu, H.; Dong, Y.; Tian, P.; Cui, P. Flatness-aware minimization for domain generalization. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 5166–5179.
54. Wang, P.; Zhang, Z.; Lei, Z.; Zhang, L. Sharpness-aware gradient matching for domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 3769–3778.
55. Zhang, R.; Fan, Z.; Yao, J.; Zhang, Y.; Wang, Y. Domain-Inspired Sharpness-Aware Minimization Under Domain Shifts. *arXiv* **2024**, arXiv:2405.18861.
56. Zhang, J.; Qi, L.; Shi, Y.; Gao, Y. Exploring Flat Minima for Domain Generalization with Large Learning Rates. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 6145–6185.
57. Lee, K.; Kim, S.; Kwak, S. Cross-domain ensemble distillation for domain generalization. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 1–20.
58. Cha, J.; Lee, K.; Park, S.; Chun, S. Domain generalization by mutual-information regularization with pre-trained models. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 440–457.
59. Wang, Y.; Wu, X.; Liu, X.; Chu, F.; Liu, H.; Han, Z. Label smoothing regularization-based no hyperparameter domain generalization. *Knowl. Based Syst.* **2024**, *309*, 112877.
60. Chen, Z.; Wang, W.; Zhao, Z.; Su, F.; Men, A.; Meng, H. PracticalDG: Perturbation Distillation on Vision-Language Models for Hybrid Domain Generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 23501–23511.
61. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144.
62. Matsuura, T.; Harada, T. Domain generalization using a mixture of multiple latent domains. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34; pp. 11749–11756.
63. Hendrycks, D.; Mu, N.; Cubuk, E.D.; Zoph, B.; Gilmer, J.; Lakshminarayanan, B. Augmix: A simple data processing method to improve robustness and uncertainty. *arXiv* **2019**, arXiv:1912.02781.

64. Cubuk, E.D.; Zoph, B.; Shlens, J.; Le, Q.V. Randaugment: Practical automated data augmentation with a reduced search space. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 702–703.
65. Chun, S.; Park, S. Styleaugment: Learning texture de-biased representations by style augmentation without pre-defined textures. *arXiv* **2021**, arXiv:2108.10549.
66. Liu, X.C.; Yang, Y.L.; Hall, P. Geometric and textural augmentation for domain gap reduction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–25 June 2022; pp. 14340–14350.
67. Liu, X.C.; Yang, Y.L.; Hall, P. Learning to warp for style transfer. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3702–3711.
68. 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, Nashville, TN, USA, 20–25 June 2021; pp. 14383–14392.
69. Lee, I.; Lee, W.; Myung, H. Domain Generalization with Vital Phase Augmentation. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 23–28 February 2024; Volume 38; pp. 2892–2900.
70. Xu, Z.; Liu, D.; Yang, J.; Raffel, C.; Niethammer, M. Robust and generalizable visual representation learning via random convolutions. *arXiv* **2020**, arXiv:2007.13003.
71. Choi, S.; Das, D.; Choi, S.; Yang, S.; Park, H.; Yun, S. Progressive random convolutions for single domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 10312–10322.
72. Shankar, S.; Piratla, V.; Chakrabarti, S.; Chaudhuri, S.; Jyothi, P.; Sarawagi, S. Generalizing Across Domains via Cross-Gradient Training. In Proceedings of the International Conference on Learning Representations, International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
73. Suzuki, T. Techaugment: Data augmentation optimization using teacher knowledge. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–25 June 2022; pp. 10904–10914.
74. Dai, R.; Zhang, Y.; Fang, Z.; Han, B.; Tian, X. Moderately Distributional Exploration for Domain Generalization. In Proceedings of the International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; pp. 6786–6817.
75. Gokhale, T.; Anirudh, R.; Thiagarajan, J.J.; Kailkhura, B.; Baral, C.; Yang, Y. Improving Diversity with Adversarially Learned Transformations for Domain Generalization. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 2–7 January 2023; pp. 434–443.
76. 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, Montreal, QC, Canada, 10–17 October 2021; pp. 8886–8895.
77. Wang, Y.; Huang, G.; Song, S.; Pan, X.; Xia, Y.; Wu, C. Regularizing deep networks with semantic data augmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3733–3748.
78. Wang, M.; Liu, Y.; Yuan, J.; Wang, S.; Wang, Z.; Wang, W. Inter-class and inter-domain semantic augmentation for domain generalization. *IEEE Trans. Image Process.* **2024**, *33*, 1338–1347.
79. Zhou, K.; Yang, Y.; Qiao, Y.; Xiang, T. Domain generalization with mixstyle. *arXiv* **2021**, arXiv:2104.02008.
80. Kang, J.; Lee, S.; Kim, N.; Kwak, S. Style neophile: Constantly seeking novel styles for domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 7130–7140.
81. Li, X.; Dai, Y.; Ge, Y.; Liu, J.; Shan, Y.; Duan, L. Uncertainty Modeling for Out-of-Distribution Generalization. In Proceedings of the International Conference on Learning Representations, Online, 25–29 April 2022; pp. 2000–2009.
82. Zhang, Z.; Wang, B.; Jha, D.; Demir, U.; Bagci, U. Domain generalization with correlated style uncertainty. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2024; pp. 2000–2009.
83. Wang, H.; Wu, X.; Huang, Z.; Xing, E.P. High-frequency component helps explain the generalization of convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8681–8691.
84. Wang, J.; Du, R.; Chang, D.; Liang, K.; Ma, Z. Domain generalization via frequency-domain-based feature disentanglement and interaction. In Proceedings of the 30th ACM International Conference on Multimedia, Lisbon, Portugal, 10–14 October 2022; pp. 4821–4829.
85. Guo, J.; Wang, N.; Qi, L.; Shi, Y. Aloft: A lightweight mlp-like architecture with dynamic low-frequency transform for domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 18–24 June 2023, Vancouver, BC, Canada; pp. 24132–24141.
86. Wang, C.; Zhang, Z.; Zhou, Z. Domain Feature Perturbation for Domain Generalization. In *ECAI 2024*; IOS Press: Amsterdam, The Netherlands, 2024; pp. 2532–2539.
87. Kim, T.; Han, B. Randomized adversarial style perturbations for domain generalization. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 3–8 January 2024, Waikoloa, HI, USA; pp. 2317–2325.

88. Zhang, Z.; Li, Y.; Shin, B.S. Learning generalizable visual representation via adaptive spectral random convolution for medical image segmentation. *Comput. Biol. Med.* **2023**, *167*, 107580.
89. Wang, Z.; Luo, Y.; Qiu, R.; Huang, Z.; Baktashmotlagh, M. Learning to diversify for single domain generalization. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 11–17 October 2021; pp. 834–843.
90. Zhang, W.; Ragab, M.; Sagarna, R. Robust domain-free domain generalization with class-aware alignment. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, Canada, 6–11 June 2021; IEEE: New York, NY, USA, 2021, pp. 2870–2874.
91. Smilkov, D.; Thorat, N.; Kim, B.; Viégas, F.; Wattenberg, M. Smoothgrad: Removing noise by adding noise. *arXiv* **2017**, arXiv:1706.03825.
92. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, *27*.
93. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2242–2251.
94. Brock, A.; Donahue, J.; Simonyan, K. Large Scale GAN Training for High Fidelity Natural Image Synthesis. In Proceedings of the International Conference on Learning Representations, Vancouver, BC, Canada, 30 April–3 May 2018.
95. Bai, H.; Yang, C.; Xu, Y.; Chan, S.H.G.; Zhou, B. Improving out-of-distribution robustness of classifiers via generative interpolation. *arXiv* **2023**, arXiv:2307.12219.
96. Rumelhart, D.E.; McClelland, J.L.; Group, P.R. *Parallel Distributed Processing, Volume 1: Explorations in the Microstructure of Cognition: Foundations*; The MIT Press: Cambridge, MA, USA, 1986.
97. Zhou, K.; Yang, Y.; Hospedales, T.; Xiang, T. Learning to generate novel domains for domain generalization. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Proceedings, Part XVI 16. Springer: Berlin/Heidelberg, Germany, 2020, pp. 561–578.
98. Xia, H.; Jing, T.; Ding, Z. Generative inference network for imbalanced domain generalization. *IEEE Trans. Image Process.* **2023**, *32*, 1694–1704.
99. Cuturi, M. Sinkhorn distances: Lightspeed computation of optimal transport. *Adv. Neural Inf. Process. Syst.* **2013**, *26*.
100. Wang, Y.; Li, H.; Cheng, H.; Wen, B.; Chau, L.P.; Kot, A.C. Variational disentanglement for domain generalization. *arXiv* **2021**, arXiv:2109.05826.
101. Zhang, Y.; Li, M.; Li, R.; Jia, K.; Zhang, L. Exact feature distribution matching for arbitrary style transfer and domain generalization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 8025–8035.
102. Zhang, Z.; Yang, S.; Dang, Q.; Jiang, T.; Liu, Q.; Wang, C.; Gu, L. Improving diversity and invariance for single domain generalization. *Inf. Sci.* **2025**, *692*, 121656.
103. Wang, Y.; Qi, L.; Shi, Y.; Gao, Y. Feature-based style randomization for domain generalization. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 5495–5509.
104. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 6840–6851.
105. Truong, T.D.; Li, X.; Raj, B.; Cothren, J.; Luu, K. ED-SAM: An Efficient Diffusion Sampling Approach to Domain Generalization in Vision-Language Foundation Models. *arXiv* **2024**, arXiv:2406.01432.
106. Hemati, S.; Beitollahi, M.; Estiri, A.H.; Omari, B.A.; Lamghari, S.; Khalil, Y.H.; Chen, X.; Zhang, G. Beyond Loss Functions: Exploring Data-Centric Approaches with Diffusion Model for Domain Generalization. *Trans. Mach. Learn. Res.* **2024**.
107. Niemeijer, J.; Schwonberg, M.; Termóhlen, J.A.; Schmidt, N.M.; Fingscheidt, T. Generalization by adaptation: Diffusion-based domain extension for domain-generalized semantic segmentation. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2024; pp. 2830–2840.
108. Noori, M.; Cheraghalikhani, M.; Bahri, A.; Hakim, G.A.V.; Osowiechi, D.; Ayed, I.B.; Desrosiers, C. TFS-ViT: Token-level feature stylization for domain generalization. *Pattern Recognit.* **2024**, *149*, 110213.
109. Chen, G.; Peng, P.; Ma, L.; Li, J.; Du, L.; Tian, Y. Amplitude-phase recombination: Rethinking robustness of convolutional neural networks in frequency domain. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 458–467.
110. Zhao, D.; Qi, L.; Shi, X.; Shi, Y.; Geng, X. A Novel Cross-Perturbation for Single Domain Generalization. *IEEE Trans. Circuits Syst. Video Technol.* **2024**, *34*, 10903–10916.
111. Jackson, P.T.; Abarghouei, A.A.; Bonner, S.; Breckon, T.P.; Obara, B. Style augmentation: Data augmentation via style randomization. In Proceedings of the CVPR Workshops, Long Beach, CA, USA, 16–20 June 2019, Volume 6; pp. 10–11.
112. Liu, C.; Cao, Y.; Su, X.; Zhu, H. Universal Frequency Domain Perturbation for Single-Source Domain Generalization. In Proceedings of the 32nd ACM International Conference on Multimedia, Melbourne, VIC, Australia, 28 October–1 November 2024; pp. 6250–6259.

113. Li, D.; Yang, Y.; Song, Y.Z.; Hospedales, T.M. Deeper, broader and artier domain generalization. In Proceedings of the IEEE International Conference on COMPUTER Vision, Venice, Italy, 22–29 October 2017; pp. 5542–5550.
114. Venkateswara, H.; Eusebio, J.; Chakraborty, S.; Panchanathan, S. Deep hashing network for unsupervised domain adaptation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5018–5027.
115. Guo, S.; Ji, K. Random color transformation for single domain generalized retinal image segmentation. *Eng. Appl. Artif. Intell.* **2024**, *136*, 108907.
116. Zhong, Z.; Zhao, Y.; Lee, G.H.; Sebe, N. Adversarial style augmentation for domain generalized urban-scene segmentation. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 338–350.
117. Ilse, M.; Tomczak, J.M.; Louizos, C.; Welling, M. Diva: Domain invariant variational autoencoders. In Proceedings of the Medical Imaging with Deep Learning, Montreal, QC, Canada, 6–8 July 2020; pp. 322–348.
118. Zhou, K.; Yang, Y.; Hospedales, T.; Xiang, T. Deep domain-adversarial image generation for domain generalisation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34; pp. 13025–13032.
119. Guo, J.; Qi, L.; Shi, Y.; Gao, Y. START: A Generalized State Space Model with Saliency-Driven Token-Aware Transformation. *arXiv* **2024**, arXiv:2410.16020.
120. Udupa, S.; Gurunath, P.; Sikdar, A.; Sundaram, S. MRFP: Learning Generalizable Semantic Segmentation from Sim-2-Real with Multi-Resolution Feature Perturbation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 5904–5914.
121. Zhang, S.; Zhang, L.; Liu, Z.Y. Frequency-based pseudo-domain generation for domain generalizable object detection. *Neurocomputing* **2023**, *542*, 126265.
122. Lee, S.; Seong, H.; Lee, S.; Kim, E. Wildnet: Learning domain generalized semantic segmentation from the wild. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 9926–9936.
123. Zhang, Z.; Li, Y.; Shin, B.S. Generalizable Polyp Segmentation via Randomized Global Illumination Augmentation. *IEEE J. Biomed. Health Inform.* **2024**, *28*, 2138–2151.
124. Xu, Z.; Tang, J.; Qi, C.; Yao, D.; Liu, C.; Zhan, Y.; Lukasiewicz, T. Cross-domain attention-guided generative data augmentation for medical image analysis with limited data. *Comput. Biol. Med.* **2024**, *168*, 107744.
125. Xue, J.; Li, Y.; Li, Z.; Cui, Y.; Zhang, S.; Wang, S. A Cross-Domain Generative Data Augmentation Framework for Aspect-Based Sentiment Analysis. *Electronics* **2023**, *12*, 2949. <https://doi.org/10.3390/electronics12132949>.
126. Taheri, A.; Zamanifar, A.; Farhadi, A. Enhancing aspect-based sentiment analysis using data augmentation based on back-translation. *Int. J. Data Sci. Anal.* **2024**, *2024*, 1–26.
127. Niu, C.; Wang, X.; Cheng, X.; Song, J.; Zhang, T. Enhancing Dialogue State Tracking Models through LLM-backed User-Agents Simulation. *arXiv* **2024**, arXiv:2405.13037.
128. Fang, Y.; Wang, W.; Dong, L.; Gao, S.; Lai, H.; Yu, Z. Decoupling-Enhanced Vietnamese Speech Recognition Accent Adaptation Supervised by Prosodic Domain Information. In Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 30 June–5 July 2024; pp. 1–8.
129. Wang, H.; Jin, Z.; Geng, M.; Hu, S.; Li, G.; Wang, T.; Xu, H.; Liu, X. Enhancing Pre-trained ASR System Fine-tuning for Dysarthric Speech Recognition using Adversarial Data Augmentation. In Proceedings of the ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Republic of Korea, 14–19 April 2024; pp. 12311–12315.
130. Jin, Z.; Xie, X.; Wang, T.; Geng, M.; Deng, J.; Li, G.; Hu, S.; Liu, X. Towards Automatic Data Augmentation for Disordered Speech Recognition. In Proceedings of the ICASSP 2024–2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Republic of Korea 14–19 April 2024; pp. 10626–10630.
131. Tobin, J.; Fong, R.; Ray, A.; Schneider, J.; Zaremba, W.; Abbeel, P. Domain randomization for transferring deep neural networks from simulation to the real world. In Proceedings of the 2017 IEEE/RSJ international conference on intelligent robots and systems (IROS), Vancouver, BC, Canada, 24–18 September 2017; pp. 23–30.
132. Niu, H.; Chen, Q.; Liu, T.; Li, J.; Zhou, G.; Zhang, Y.; Hu, J.; Zhan, X. xTED: Cross-Domain Adaptation via Diffusion-Based Trajectory Editing. *arXiv* **2024**, arXiv:2409.08687.
133. Cubuk, E.D.; Zoph, B.; Mane, D.; Vasudevan, V.; Le, Q.V. Autoaugment: Learning augmentation strategies from data. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 113–123.
134. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10684–10695.
135. Hu, E.J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. Lora: Low-rank adaptation of large language models. *ICLR* **2022**, *1*, 3.

136. Zhu, R.; Zhang, Z.; Liang, S.; Liu, Z.; Xu, C. Learning to transform dynamically for better adversarial transferability. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 16–22 June 2024; pp. 24273–24283.
137. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
138. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. In Proceedings of the International Conference on Learning Representations, Vienna, Austria, 3–7 May 2021.
139. Zhu, L.; Liao, B.; Zhang, Q.; Wang, X.; Liu, W.; Wang, X. Vision mamba: Efficient visual representation learning with bidirectional state space model. *arXiv* **2024**, arXiv:2401.09417.
140. Vaswani, A. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
141. Gu, A.; Dao, T. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv* **2023**, arXiv:2312.00752.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.