

A Review of All-in-One Image Restoration Techniques

Shanshan Tian *

School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, 454000, Henan, PR China

Abstract: Image restoration (IR) refers to the process of eliminating image degradations—such as noise, blur, and weather effects like rain or haze—while enhancing visual quality. Traditional IR methods typically target specific types of degradation, which limits their effectiveness in real-world scenarios involving complex distortions. To address this challenge, the All-in-One Image Restoration (AiOIR) paradigm has emerged, offering a unified framework adept at resolving multiple degradation types. These innovative models integrate degradation-specific adaptive learning with cross-task knowledge sharing, significantly enhancing their flexibility and versatility. This paper systematically reviews existing AiOIR techniques and provides an in-depth analysis of their innovations in architectural evolution and feature learning paradigms. To advance this dynamic field, we present a comprehensive and profound review of AiOIR technologies. The narrative structure of this article is organized as follows: first, we elaborate on the foundational concepts of AiOIR; second, we systematically categorize and review mainstream cutting-edge designs based on key factors such as prior features and generalization capability; building upon this, we highlight the core technological advancements in the field; finally, we critically evaluate the bottlenecks and challenges encountered by current models and identify future research directions, with the aim of inspiring broader academic exploration and innovation.

Keywords: All-in-One model, Image Restoration, Computer Vision, Deep Learning.

1. Introduction

1.1. Research Background and Significance

Image restoration is a fundamental task in computer vision, aiming to enhance the visual quality of images by eliminating degradation factors such as noise, blur, and low illumination. Traditional image restoration methods are predominantly optimized for specific, single types of degradation, such as denoising, dehazing, or super-resolution. However, in real-world applications, images are frequently corrupted by multiple degradation factors simultaneously, which increasingly exposes the limitations of single-task image restoration methods in practical scenarios. For instance, in autonomous driving scenarios, images captured by vehicle-mounted cameras may suffer from low illumination, haze, and noise concurrently; in medical imaging, images might be compromised by a combination of equipment noise, motion blur, and low contrast. Consequently, developing image restoration techniques capable of addressing multiple degradation types simultaneously holds significant practical importance.

In recent years, propelled by the rapid advancement of deep learning technologies, the field of image restoration has witnessed remarkable progress. Deep learning models, particularly Convolutional Neural Networks (CNNs) [1] and Transformers [2], have been widely deployed in image restoration tasks due to their robust feature extraction capabilities and proficiency in learning complex data distributions. These models can automatically learn the intrinsic features of images from massive amounts of annotated data, thereby achieving effective restoration of degraded images. Nevertheless, most existing deep learning models still focus exclusively on single types of degradation, rendering them ill-equipped to adapt to complex, multi-degradation scenarios. To address this bottleneck, the All-in-One Image Restoration (AiOIR) paradigm has emerged. By concurrently processing multiple degradation types within a unified framework, AiOIR not only improves the efficiency

of image restoration but also substantially enhances the generalization capability of the models, empowering them to better adapt to complex real-world scenarios.

1.2. Research and Solution Approaches

Despite the significant strides made in All-in-One Image Restoration (AiOIR) technologies, they still face multiple severe challenges in practical applications. The foremost issue is the inherent conflict within multi-task optimization: the restoration objectives for different degradation types are often diametrically opposed. For instance, denoising tasks aim to suppress high-frequency information, whereas deblurring tasks urgently require the restoration of high-frequency texture details. This contradiction renders models highly susceptible to negative transfer and performance bottlenecks during joint optimization. The second challenge is the cross-domain generalization dilemma: real-world degradations are complex and highly variable, exhibiting significant distribution shifts from limited training data, which imposes stringent requirements on model robustness. Furthermore, existing AiOIR models (as illustrated in Figure 1) are typically accompanied by massive parameter counts and high computational complexity, making them difficult to deploy efficiently on resource-constrained devices such as mobile terminals or embedded systems. Finally, these models are highly dependent on large-scale, high-quality paired annotated data. Not only is the acquisition cost exorbitant, but the data imbalance across different degradation categories easily induces performance bias within the network.

To address the aforementioned pain points, current research strategies primarily seek breakthroughs across three dimensions: architectural innovation, learning paradigm evolution, and efficient data utilization. At the level of network architecture design, researchers have transcended the limitations of singular structures, exploring cutting-edge designs such as multi-head branch structures, Mixture-of-Experts (MoE) architectures, and multimodal fusion. Through dynamic feature routing, these approaches significantly enhance the model's adaptive matching capabilities for

multiple degradation types. Regarding the optimization of learning strategies, contrastive learning, multi-task learning mechanisms, and continual learning have been widely introduced. These methods aim to effectively mitigate task conflicts, suppress catastrophic forgetting, and substantially elevate the models' generalization and learning efficiency.

Additionally, to break through the data dependency bottleneck, a substantial body of research is dedicated to integrating data augmentation, self-supervised learning, and transfer learning techniques into AiOIR frameworks, with the expectation of achieving superior restoration performance in data-limited or unlabeled scenarios.

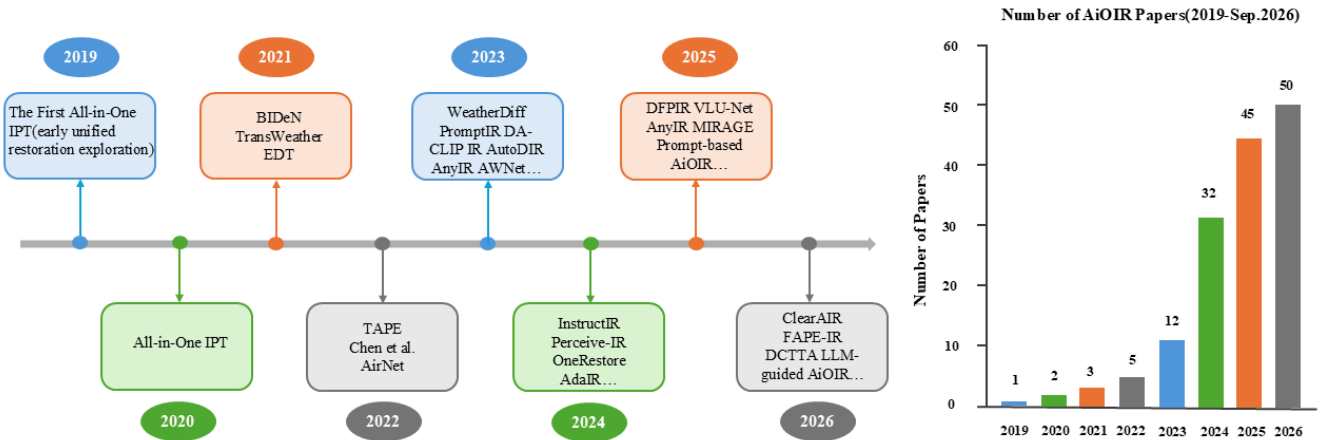


Figure 1. Timeline of representative AiOIR works. The bar chart illustrates the number of AiOIR publications from 2019 to September 2026

1.3. Research Contributions and Significance

The main contributions of this paper are summarized as follows:

- (1) Systematically introduces the network designs, learning strategies, and key methodological advancements of AiOIR technologies.
- (2) Conducts a comprehensive categorization and in-depth analysis of existing AiOIR methods, summarizing the strengths and limitations of different approaches.
- (3) Discusses the pertinent challenges and outlines future research directions, providing a valuable reference for researchers in this domain.

This study holds significant value for advancing the development of AiOIR technologies. By conducting a profound analysis and comparison of state-of-the-art methods, this review provides researchers with valuable insights and inspirations. It ultimately facilitates the further optimization and refinement of AiOIR models, promoting their robust deployment in practical, real-world scenarios.

2. Analysis of Current Research Status

2.1. Network Design

The network design of AiOIR is crucial for achieving the unified processing of multiple degradation types. Existing AiOIR methods can be broadly categorized into the following types:

Task-Specific Encoders and Decoders: Early AiOIR methods typically adopted independent encoder-decoder architectures tailored for different degradation types, wherein each restoration task corresponded to a dedicated network branch. This approach can fully exploit the differences in the feature distributions of various degradations, allowing the model to be optimized for specific tasks. However, since different degradations require independent network parameters, such methods result in massive model sizes. Additionally, they require prior knowledge of the degradation category, thereby lacking flexibility when confronting unknown degradations or mixed degradation scenarios. Consequently, these methods are primarily suitable for restoration scenarios with limited degradation categories and

explicit task divisions, making it difficult to fulfill the demands of unified restoration for complex degradations in real-world environments.

Shared Decoder with Multi-Head Structure: To reduce the scale of model parameters, some studies have adopted a structure that combines a shared decoder with task-specific restoration heads. This method shares the low-level feature extraction process and utilizes different task heads to accomplish specific degradation restoration. For example, the All-in-One restoration framework proposed by Li et al. [3] utilizes shared network parameters to process image restoration under various adverse weather conditions, and enhances the restoration capability for different degradation types through task-specific modules. Compared to completely independent networks, this structure reduces computational costs but still relies on task category information, presenting certain limitations under unknown degradation conditions.

Unified Encoder-Decoder Architecture: A unified encoder-decoder structure further reduces task-specific parameters, utilizing only a single network to accomplish multiple degradation restoration tasks. For instance, some early unified restoration models learned common feature representations across different degradations by sharing both the encoder and the decoder. This type of method features a simple structure and easy deployment. However, because all degradation tasks share the same feature space, it is highly susceptible to feature competition among tasks. When confronted with complex, real-world combinations of degradations, the model struggles to dynamically adjust its restoration strategies.

Shared Backbone with Multi-Decoder: Another class of methods employs a shared feature extraction backbone while designing multiple task-specific decoders for different restoration tasks. The method proposed by Chen et al. [4] uses a shared encoder to extract degradation-related features and subsequently completes different task restorations through multiple restoration branches. This structure can enhance the differential representation capability among tasks; nevertheless, it requires extensive task annotation information, and the number of network parameters increases significantly as the types of degradation multiply.

Pre-trained Intermediate Backbone: With the rapid

development of large-scale pre-trained models, researchers have begun to utilize pre-trained Transformers [2] as AiOIR backbones, acquiring more universal image representations through transfer learning. For example, IPT [5] utilizes large-scale image restoration tasks to pre-train a Transformer [2] model and transfers it to different restoration tasks to achieve a unified image processing framework. This method demonstrates the significant value of pre-trained vision models for AiOIR; however, because it primarily relies on fixed task settings, it still lacks dynamic adaptability for unknown degradations.

Mixture-of-Experts (MoE) Architecture: Inputs are routed to different expert networks via a gating mechanism. For example, MEASNet [6] considers pixel-level and global features, including low-frequency and high-frequency components, to select the appropriate expert for image restoration. DFPIR [7] models feature variations under different degradation conditions through a degradation-aware feature perturbation mechanism, enabling dynamic adjustment within the shared feature space. FAPE-IR [8] further combines multimodal planning with LoRA-MoE modules to achieve semantics-guided expert selection and adaptive restoration.

Pre-trained Model Prior: This approach leverages pre-trained vision-language models, such as CLIP [9], DINO [10], and DINO-v2 [11], to enhance the restoration process by predicting high-quality feature embeddings. Models like Perceive-IR [12] and DA-CLIP [13] improve restoration performance by exploiting the semantic alignment capabilities of these foundation models. VLU-Net [14] utilizes vision-language models to extract semantic correlation features between images and text, integrating them into a deep unfolding network. By guiding the restoration process with vision-language priors, it substantially improves the model's ability to identify and adapt to various degradation types.

Visual Prompting: This involves injecting learnable visual prompts during the decoding stage to implicitly predict degradation conditions. For instance, PromptIR [15] guides the decoder to adaptively restore images with various degradations via prompt blocks.

Text Prompts or Instructions: This category guides the restoration process through natural language instructions or text prompts to enhance adaptability to unknown degradations. For example, TextPromptIR [16] utilizes BERT [17] to generate semantic prompts, enhancing the model's adaptability and usability. InstructIR [18] further proposes a restoration framework based on natural language instructions, enabling users to control restoration objectives via varying commands. Defusion [19] additionally incorporates diffusion models [20], describing degradation patterns through visual instructions and utilizing the degradation diffusion process to guide restoration, thereby achieving unified restoration under more complex and realistic degradation conditions.

Question-Answering Paradigm and Multimodal Intelligent Restoration: By employing a question-answering framework to interpret user inputs and adjust the restoration process, this paradigm allows users to customize image restoration according to their preferences. Models like promptGIP [21] and AutoDIR [22] realize user-controllable image restoration

through a QA format. Multimodal large models further propel the development in this direction. FAPE-IR [8] utilizes a multimodal large language model as a restoration planner, generating restoration strategies based on the input image content and degradation conditions. ClearAIR [23] simulates the human visual perception process, introducing quality assessment and semantic guidance mechanisms to achieve progressive restoration from global structures to local details. Furthermore, DCTTA [24] proposes a degradation-consistent test-time adaptation mechanism, enabling the model to dynamically adjust its parameters according to unknown testing environments, thereby enhancing its generalization capabilities in real-world scenarios.

2.2. Learning Strategies

Beyond architecture design, learning strategies are vital for AiOIR's generalization and performance. Balancing shared knowledge learning and task conflict avoidance across multiple degradation tasks is the current research focus. Based on training paradigms and knowledge transfer mechanisms, existing strategies are classified as follows:

Multi-Task Learning (MTL): MTL is the earliest learning paradigm adopted in AiOIR. Its core philosophy involves treating different degradation types as multiple correlated tasks and achieving knowledge transfer through shared network parameters. Joint training can leverage complementary information among different degradations to enhance the model's overall generalization capability. However, due to disparities in feature distributions and optimization objectives across various degradation tasks, mixed training is prone to inducing gradient conflicts between tasks, making it difficult for the model to optimize all restoration objectives simultaneously. Consequently, in recent years, research has shifted from traditional mixed training towards more refined task optimization strategies. For instance, MiOIR [25] analyzed the impact of the training sequence of different degradation tasks on restoration performance, pointing out that rationally scheduling the task learning process can alleviate inter-task interference. SimpleIR [26] further utilized the information entropy differences between various restoration tasks to gauge task difficulty, and designed the training sequence based on task complexity to improve the model's learning efficiency.

Continual Learning and Sequential Adaptation: From the perspective of transfer learning, AiOIR can be divided into simultaneous multi-task learning and sequential task learning. Sequential learning mimics the human process of knowledge accumulation, enabling the model to retain existing restoration capabilities while learning new degradation tasks. Given that AiOIR models need to continuously learn new degradation types, catastrophic forgetting [27] has emerged as a critical factor limiting the model's extensibility. To address this issue, researchers have proposed review learning (rehearsal-based learning) [28] and continual learning [29] strategies. By preserving historical task knowledge or constraining parameter updates, these methods mitigate the performance degradation caused by learning new tasks. A schematic diagram of transfer learning scenarios is illustrated in Figure 2 [12-15, 18, 23, 30-31].

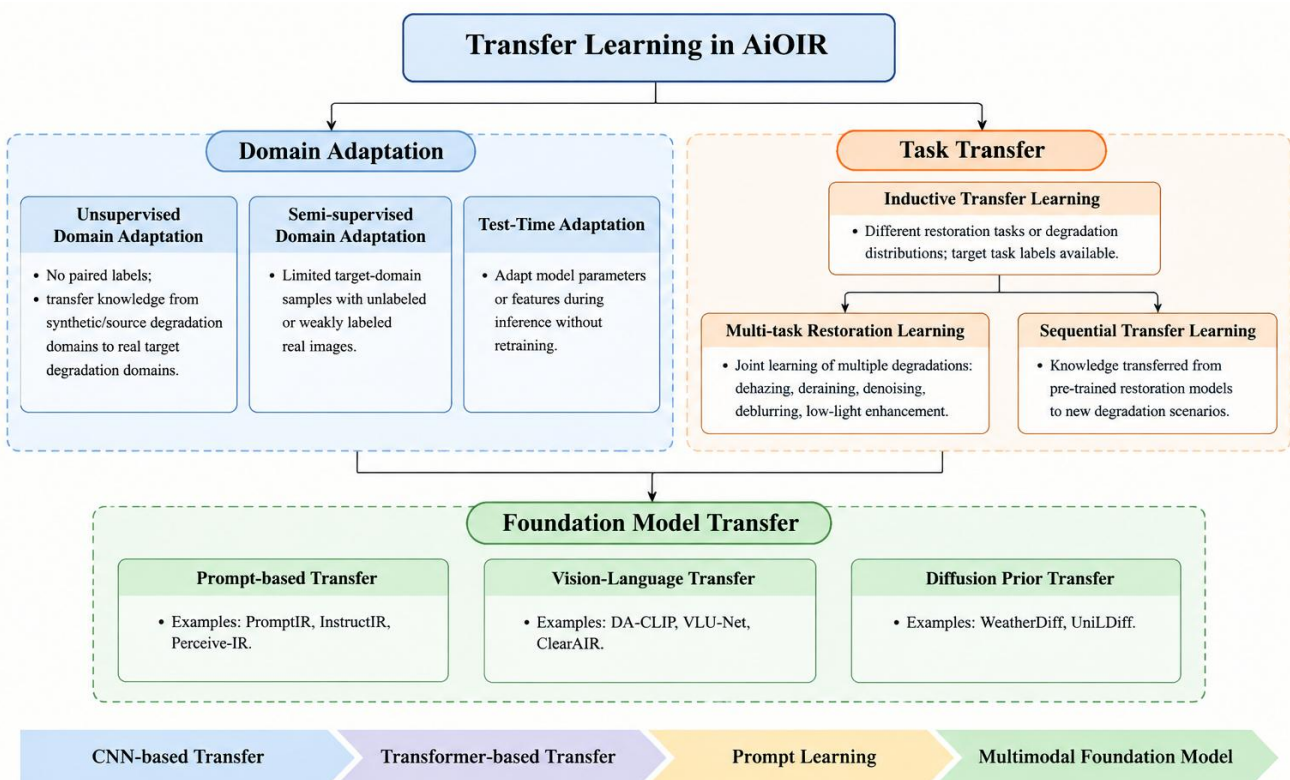


Figure 2. Transfer learning scenarios in AiOIR

Contrastive Learning-Driven Degradation Representation Learning: Due to the high diversity of degradation types in real-world environments, models often struggle to generalize to unseen degradations. Consequently, contrastive learning has been introduced into AiOIR to enhance the model's discriminative capability across various degradation patterns by constructing a structured degradation feature space. Traditional methods typically utilize positive and negative samples to constrain the feature distribution, pulling similar degradation samples closer within the latent space while pushing disparate ones apart [32–34]. Related studies indicate that contrastive constraints can serve as an additional regularization term, bolstering the model's robustness against unknown degradations [35]. In recent years, contrastive learning has increasingly been integrated with generative and foundation models to learn more stable semantic representations of degradations, thereby supporting unified restoration in unknown environments [13].

Degradation-Aware Adaptive Learning: As the complexity of AiOIR tasks escalates, learning paradigms relying solely on fixed task labels struggle to meet the demands of real-world scenarios. Therefore, recent research has shifted focus toward degradation condition perception and dynamic adjustment mechanisms. DFPIR [7] proposes a degradation-aware feature perturbation strategy that dynamically modulates the shared feature space by learning the evolution patterns of features under varying degradation conditions, achieving adaptive degradation learning without explicit task division. Compared to traditional multi-task training, this approach mitigates feature conflicts among different degradations and bolsters the model's generalization capability when confronted with complex degradation combinations. Furthermore, FAPE-IR [8] extends degradation-aware learning into a multimodal planning framework. By analyzing the degradation features of input images via multimodal models and employing LoRA-MoE execution modules to dynamically adjust the restoration

process, it successfully marries high-level semantic planning with low-level restoration execution.

Foundation Model Transfer and Test-Time Adaptive Learning: With the advent of CLIP [9], DINO [10], and large vision-language models, the transfer of pre-trained models has become a pivotal learning strategy for AiOIR. Such methods leverage the semantic knowledge acquired by large-scale pre-trained models to provide the restoration network with enriched prior information. For instance, DA-CLIP [13] utilizes degradation-aware CLIP [9] features to construct a degradation semantic space, enhancing the model's capability to recognize unknown degradations; Perceive-IR [12] exploits visual perception priors to augment the visual quality of the restored outcomes. Advancing further, DCTTA [24] proposes a degradation-consistent Test-Time Adaptation (TTA) strategy, which dynamically calibrates model parameters during the testing phase, enabling AiOIR models to adapt to novel degradation environments unseen during training. This direction embodies the developmental trajectory of AiOIR, transitioning from "fixed optimization during training" to "autonomous adaptation during testing."

2.3. Key Methodological Advancements

As AiOIR research progresses, methods relying solely on shared network parameters to handle multiple degradations have gradually become insufficient to meet the demands of real-world scenarios. In recent years, researchers have introduced technologies such as prompt learning, dynamic expert mechanisms, vision-language models, and diffusion models [20] to enhance the model's adaptability to unknown degradations and complex degradation combinations (as shown in Fig. 3). Based on their primary technical routes, existing AiOIR improvement methods can be classified into the following categories:

Prompt Learning: Originally applied in the field of natural language processing, the core philosophy of prompt learning is to introduce additional conditional information to guide the

model in accomplishing specific tasks. In AiOIR, prompt learning provides degradation-related information to the restoration network by constructing visual, textual, or multimodal prompts, thereby improving the model's adaptability to various degradation types. Visual prompting was the earliest form of prompting applied in AiOIR. PromptIR [16] proposed a learnable visual prompt module that embeds prompt information into the Transformer [2] restoration network. By propagating degradation-related context across scales, it enables the model to dynamically adjust the restoration process based on the input degradation. Subsequently, methods like SIPL-PromptIR [16] and CPL-PromptIR [36] further optimized prompt generation mechanisms, enhancing the representational capacity of prompt features for different degradation tasks. With the evolution of vision-language models, the prompting mechanism gradually expanded from visual prompts to semantic prompts. TextPromptIR [16] utilizes BERT [17] to generate task-related text prompts, mapping natural language descriptions to the restoration network so the model can perform restoration based on degradation semantics. Methods like AnyIR [36] and MIRAGE [37] further propelled prompt-driven AiOIR by combining visual semantic priors with adaptive prompt generation mechanisms, elevating the model's generalization capabilities in complex, unknown degradation scenarios.

Mixture-of-Experts (MoE) Architectures: The MoE architecture achieves input-dependent dynamic selection via multiple expert networks and a gating mechanism, enabling different experts to be optimized for specific degradation patterns. WM-MoE [38] proposed the Weather-Aware Routing (WEAR) mechanism, utilizing input weather conditions to select corresponding experts and combining multi-scale expert networks to handle different weather-induced degradations, which improved image restoration capabilities under adverse weather. MEASNet [6] further considered local pixel features and global semantic features, including low-frequency and high-frequency information, employing a multi-expert adaptive selection mechanism to match suitable experts for different degradation tasks, achieving more flexible all-in-one restoration. DFPIR [7] introduced a degradation-aware feature perturbation mechanism that adaptively adjusts the shared feature space by learning feature evolution patterns under various degradation conditions; this achieves dynamic degradation modeling without explicit expert branches, reducing the parameter overhead associated with traditional MoE structures. FAPE-IR [8] integrated a multimodal planner with a LoRA-MoE execution module to dynamically select restoration strategies based on the semantic and frequency information of the input image, realizing an adaptive restoration framework that couples high-level semantic planning with low-level expert execution.

Multimodal Model-Driven Intelligent Restoration: Driven by the rapid development of CLIP [9], DINO [10], and Large Language Models (LLMs), multimodal models have been increasingly introduced into AiOIR tasks. By fusing visual

and linguistic information, they bolster the model's comprehension of complex degradations. Simultaneously, advancements in diffusion models [20] have further elevated the capacity to model structural information and high-quality image priors, offering novel solutions for complex real-world degradation restoration. Early multimodal AiOIR methods primarily utilized vision-language models to establish correlations between image content and degradation semantics. Clarity-ChatGPT [40] integrated visual models with large language models, allowing users to control the image restoration process via natural language interactions, achieving highly flexible human-computer interactive restoration. AutoDIR [22] leveraged Semantic-Agnostic Blind Image Quality Assessment to automatically analyze unknown degradations and select restoration strategies accordingly, improving adaptability to unknown degradation environments. Instruct-IPT [42] utilized GPT-4 to generate restoration instructions, combining visual information with text prompts to enhance the model's understanding of complex restoration tasks. Subsequently, vision-language priors were applied to degradation perception and restoration optimization. DA-CLIP [13] established a degradation semantic representation using the degradation-aware CLIP [9] space, enabling the model to recognize different degradation types based on vision-language features; Perceive-IR [12] further utilized visual perception priors to enhance the perceptual quality of the restored results. VLU-Net [14] incorporated vision-language models into a deep unfolding network, guiding the optimization process via vision-language features to equip the restoration model with a more explicit comprehension of degradations. Concurrently, diffusion models [19] and foundation model transfer strategies have emerged as pivotal development directions for AiOIR. Defusion [19] proposed a visual-instruction-driven degradation diffusion framework, describing degradation patterns through visual instructions and leveraging the diffusion process to guide image restoration, thereby enabling the model to handle more complex real-world degradation combinations. Building upon this, FAPE-IR [8] introduced multimodal large models as restoration planners to generate restoration strategies through semantic understanding, executing dynamic restoration via LoRA-MoE modules. This marks a paradigm shift from traditional "image restoration networks" to "intelligent restoration systems." ClearAIR [23] simulates the human visual perception process, employing quality assessment, semantic understanding, and progressive restoration strategies to enhance restoration capabilities in complex real-world degradation environments. Furthermore, DCTTA [24] proposed a degradation-consistent test-time adaptation method, dynamically adjusting the model during the testing phase to adapt to novel degradation types uncovered during the training phase.

Overall, multimodal foundation model-driven AiOIR methods are propelling image restoration from the traditional "degradation recognition – network restoration" paradigm toward an intelligent direction characterized by "semantic understanding – strategy planning – adaptive restoration."

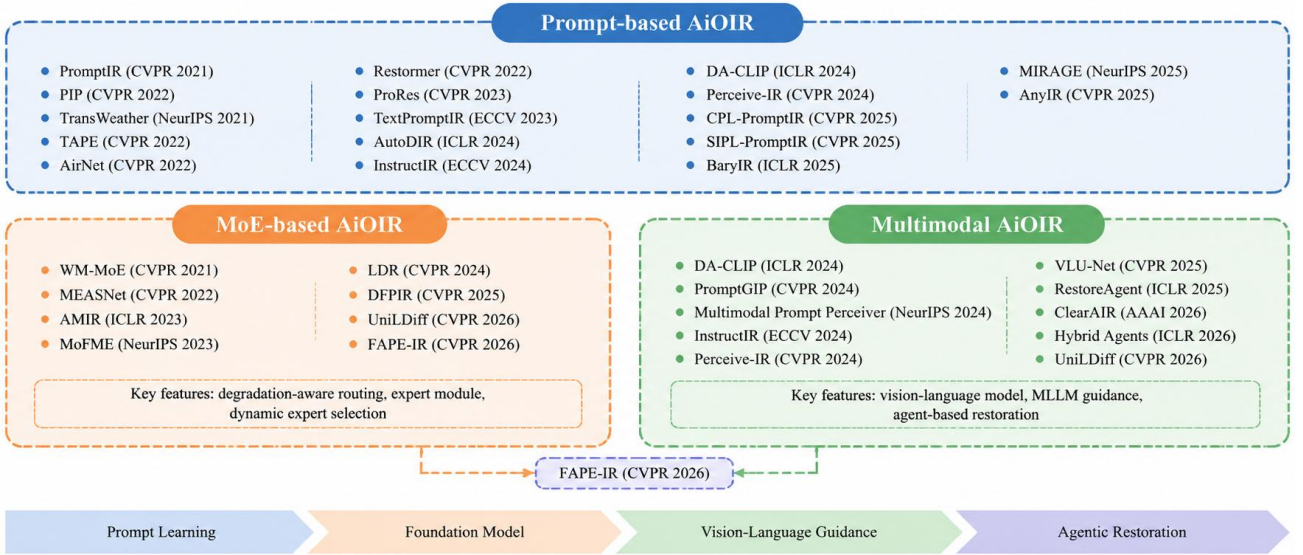


Figure 3. Overview of AiOIR methodologies. This figure categorizes AiOIR approaches into three types based on their primary improvements: prompt-based models (blue background), MoE-based models (orange background), and multimodal models (green background).

3. Problems and Challenges

In the field of All-in-One Image Restoration (AiOIR), despite the significant advancements achieved in recent years, numerous challenges remain. These issues are primarily concentrated in the following aspects:

Unknown Degradation Modeling and Task Conflict Issues: The core objective of AiOIR is to achieve the unified restoration of multiple degradation types within a single model framework. However, there are usually pronounced disparities in optimization objectives and conflicts in feature representations among different restoration tasks. For instance, denoising tasks necessitate the suppression of high-frequency random perturbations, whereas deblurring, super-resolution, and image enhancement tasks require the restoration and amplification of high-frequency texture information. Consequently, during multi-degradation joint training, the gradient directions generated by different tasks may compete, making it exceedingly difficult for the model to optimize disparate restoration objectives simultaneously. Early AiOIR methods primarily realized knowledge transfer among different tasks by sharing network parameters. However, this approach is prone to inducing feature interference between tasks, causing the model to favor certain degradation types while degrading the restoration performance on others. In recent years, researchers have mitigated this issue through dynamic routing and degradation-aware mechanisms. For example, MEASNet [6] utilizes local pixel features and global semantic features to allocate appropriate restoration experts for different degradation tasks via a multi-expert selection mechanism. DFPIR [7] further proposes a degradation-aware feature perturbation mechanism, achieving implicit degradation modeling by adjusting the shared feature space; this reduces parameter overhead while enhancing the model's adaptability to various degradation conditions. Nevertheless, most existing methods still rely on pre-defined degradation types for training. For unknown composite degradations existing in real-world environments—such as complex scenarios jointly affected by rain, haze, low illumination, and noise—models still lack effective degradation comprehension capabilities. Moreover, the underlying correlations and conflict

mechanisms among different degradations have yet to be thoroughly modeled, restricting the applicability of AiOIR models in open environments.

Inadequate Generalization Capability in Real-World Scenarios: Currently, the majority of AiOIR models rely on artificially synthesized degradation data for supervised training, such as generating training samples by simulating noise, raindrops, blur, or hazing processes. However, image degradations in the real world typically possess more complex formation mechanisms and non-uniform distribution characteristics, resulting in conspicuous domain gaps between synthetic degradation data and actual degradations. For instance, synthetic noise models struggle to completely characterize the spatial distribution features of real camera sensor noise, and artificially generated rain and haze degradations fail to accurately simulate complex illumination, atmospheric scattering, and dynamic variation processes in real environments. This mismatch between training data and real application environments leads to a precipitous decline in the model's restoration performance in practical scenarios. To address this issue, some studies have begun to explore Test-Time Adaptation (TTA) methods. DCTTA [24] proposes a degradation-consistent TTA strategy that adjusts the model utilizing target domain information during the testing phase, enabling AiOIR models to adapt to novel degradation environments not covered during the training phase. However, most current methods still rely on limited degradation assumptions, and their adaptability in completely unknown environments remains constrained.

Foundation Model Transfer and Computational Cost Issues: In recent years, the development of CLIP [9], DINO [10], and large-scale vision-language models has provided AiOIR with abundant visual semantic priors, endowing models with superior degradation comprehension and visual perception capabilities. For example, DA-CLIP [13] leverages the degradation-aware CLIP [9] space to establish semantic correlations between image content and degradation types; Perceive-IR [12] utilizes visual perception priors to augment the perceptual quality of restoration results; VLU-Net [14] introduces vision-language models into deep unfolding networks to guide the restoration optimization process via vision-language features. However, large-scale foundation

models generally entail massive parameter sizes and computational demands. Directly transferring them to AiOIR tasks introduces exorbitant training costs and storage pressures. Furthermore, foundation models are typically pre-trained for general vision tasks, and there are inherent discrepancies between their knowledge representations and low-level image restoration tasks. How to effectively exploit high-level semantic knowledge without introducing erroneous priors remains a pivotal issue currently being confronted. For instance, while FAPE-IR [8] mitigates the application cost of foundation models via LoRA-MoE execution modules, the effective transfer of foundation models in complex degradation restoration is still encumbered by factors such as computational resources, feature adaptation capabilities, and model scale.

Model Complexity and Real-Time Deployment Difficulties: As Transformers [2], diffusion models [20], and large-scale vision-language models are increasingly applied to AiOIR, model restoration performance has been markedly elevated. Nonetheless, this has simultaneously incurred increased computational complexity and degraded inference efficiency. In particular, diffusion models [20] usually require multi-step iterative sampling processes to generate high-quality restored results, rendering them incapable of meeting real-time application demands. Although Transformers [2] possess formidable capabilities in modeling long-range dependencies, their high computational overhead and substantial memory requirements impede their deployment on mobile terminals, embedded devices, and edge computing platforms. Additionally, current AiOIR models typically need to learn multiple restoration tasks concurrently, leading to further expansion of the network scale, which escalates training costs, storage requirements, and inference latency. Therefore, the contradiction between model performance and computational efficiency remains a critical factor restricting the practical application of AiOIR.

Inadequate Multimodal Comprehension Capability: Driven by the advancement of Large Language Models (LLMs) and vision-language models, AiOIR is gradually transitioning from a traditional image restoration task to a restoration paradigm that integrates semantic understanding. Clarity-ChatGPT [40], AutoDIR [22], and Instruct-IPT [41] introduce language models to allow users to describe restoration requirements via text instructions, enhancing the model's capacity to comprehend complex tasks. Moving further, Defusion [19] utilizes visual instructions to drive the diffusion restoration process, while FAPE-IR [8] incorporates multimodal large models as restoration planning modules, enabling the model to formulate restoration strategies based on input content and degradation states. However, current multimodal AiOIR still suffers from inadequate alignment between visual and linguistic information. Since image degradation typically manifests as low-level visual alterations while linguistic descriptions usually possess a high degree of abstraction, a glaring semantic gap exists between the two. Furthermore, multimodal models still lack a thorough comprehension of degradation severity, combinatorial relationships among degradations, and their physical etiologies, which may result in erroneous restoration strategy selection or yield restored outcomes inconsistent with real structures.

Imperfect Datasets and Evaluation Systems: Compared to single image restoration tasks, AiOIR currently lacks unified, large-scale real-world datasets and standardized evaluation

protocols. Different studies typically employ divergent degradation types, data sources, and experimental setups, making equitable comparisons across different methods exceedingly difficult. Furthermore, traditional evaluation metrics, such as PSNR and SSIM, primarily quantify pixel-level reconstruction errors but fail to adequately reflect the visual perceptual quality of actual restored results. For instance, some models might achieve high pixel-based metrics but suffer from overly smoothed textures or structural distortions in the generated images. Recently, perceptual quality assessment metrics such as MUSIQ [42], MANIQA [43], and CLIP-IQA [44] have been increasingly adopted in AiOIR evaluation. Nevertheless, there remains a lack of a unified evaluation framework capable of simultaneously assessing structural preservation, perceptual quality, semantic consistency, and authenticity.

Insufficient Processing Capability for High-Dimensional Visual Data: At present, AiOIR research predominantly focuses on 2D RGB image restoration, whereas practical application scenarios encompass more intricate data formats, including videos, medical imagery, hyperspectral data, and 3D visual data. For example, video restoration must not only enhance single-frame image quality but also guarantee temporal continuity and motion consistency; hyperspectral data restoration necessitates the simultaneous preservation of spatial structural information and spectral consistency across distinct bands; 3D visual data restoration requires the comprehensive consideration of both spatial structure and depth information. Because existing AiOIR models are primarily designed for 2D static images, their network architectures and training strategies are arduous to directly apply to these high-dimensional visual signals, limiting the application of AiOIR technologies across a broader spectrum of visual tasks.

Model Interpretability and Restoration Credibility Issues: As the structural complexity of AiOIR models escalates, their internal decision-making processes and restoration mechanisms increasingly lack interpretability. In highly reliable application scenarios such as medical imaging, autonomous driving, and industrial inspection, merely obtaining restored results with high visual quality is insufficient; guaranteeing the reliability and authenticity of the restoration process is imperative. Currently, models are typically unable to explicitly interpret how they diagnose the degradation types in the input images, how they select restoration strategies, and whether nonexistent structural information is fabricated during the restoration process. ClearAIR [23] enhances the interpretability of the restoration process by simulating human visual perception mechanisms, introducing quality assessment, semantic comprehension, and progressive restoration strategies. Overall, however, the decision transparency, uncertainty analysis, and restoration outcome credibility of AiOIR models remain deficient.

Overall, AiOIR has gradually evolved from early multi-task unified restoration models to complex restoration frameworks integrating Transformers [2], MoEs, vision-language models, and generative models. Nonetheless, existing methods are still constrained by inadequate adaptability to unknown degradations, restricted generalization in real-world scenarios, challenges in foundation model transfer, intensive computational resource demands, insufficient multimodal comprehension, and imperfect evaluation systems. These challenges indicate that current AiOIR is still in a crucial transition phase from

experimental environments to real-world intelligent visual applications.

4. Future Research Directions

In light of the current limitations of AiOIR methods concerning task conflicts, real-world generalization, model efficiency, and data resources, future research must further explore more robust, efficient, and intelligent all-in-one image restoration frameworks. Driven by the advancements in multi-task learning, foundation models, and generative AI, the future trajectories of AiOIR primarily converge on the following aspects:

Multi-Task Learning and Dynamic Task Optimization: AiOIR is intrinsically a quintessential multi-task learning problem, characterized by a dual relationship of feature sharing and task competition among diverse degradation types. Therefore, constructing a more robust multi-task learning theory is a pivotal direction for the future development of AiOIR. Future research needs to further explore task conflict modeling methods to achieve effective synergy among different restoration tasks by dynamically adjusting task weights and optimizing inter-task information interaction mechanisms. For example, leveraging gradient balancing strategies, task correlation analysis, and dynamic routing mechanisms can enable models to adaptively allocate network capacity based on different degradation conditions. Furthermore, while current AiOIR typically employs fixed mixed-training strategies, there are pronounced disparities in the learning difficulty and knowledge transfer relationships among different degradation tasks. Hence, establishing a more rational task learning sequence or staged restoration strategy is also a critical future research direction. Designing adaptive task priorities coupled with degradation severity, task complexity, and knowledge transfer relationships holds the promise of further enhancing the stability and generalization capability of models.

Semi-Supervised [45], Unsupervised [46], and Weakly-Supervised Restoration [47]: Currently, AiOIR models generally rely on a vast amount of high-quality paired data for supervised training. However, the difficulty of acquiring degraded images in real-world environments and the complex, volatile nature of degradation processes restrict the further development of models. Thus, alleviating the reliance on manually annotated data and advancing semi-supervised and unsupervised learning methods are of paramount importance. Future research can further integrate technologies such as self-supervised learning, contrastive learning, and unsupervised domain adaptation to bolster the model's adaptability to unknown degradation environments by mining the latent structural information in unannotated data. For instance, constructing degradation consistency constraints enables models to learn degradation patterns from real degraded images; leveraging contrastive learning to establish a degradation feature space enhances the model's ability to discriminate between different degradation types. Additionally, coupling generative models to construct high-quality pseudo-labels and self-training mechanisms will also emerge as a crucial direction for mitigating data annotation costs.

Platform-Aware Model Design and Efficient Deployment: As the scale of AiOIR models continues to expand, striking a balance between high-quality restoration and computational efficiency has become a critical bottleneck in practical applications. Future research must dedicate more attention to

model designs tailored for disparate computing platforms, enabling the deployment of AiOIR on mobile devices, edge computing platforms, and embedded systems. On the one hand, computational overhead can be reduced through methods such as model pruning [48], parameter quantization [49], knowledge distillation [50], and low-rank adaptation (LoRA) [51]. On the other hand, by integrating Neural Architecture Search (NAS) [52] and hardware-aware optimization strategies [53], highly efficient restoration networks suited for specific device environments can be designed. Moreover, with the evolution of State Space Models (SSMs) [54] and lightweight Transformers [2], future AiOIR is expected to overcome the deployment hurdles of complex models, realizing restoration systems that boast both high performance and low computational costs.

Data and Task Expansion for Real-World Complex Degradation Scenarios: The majority of current AiOIR research is predicated on artificially synthesized degradation data, whereas image degradations in real environments typically exhibit far more complex formation mechanisms. For example, actual images may be simultaneously affected by low illumination, noise, blur, and weather interference. Consequently, constructing datasets and task frameworks that better align with practical application demands is a vital direction for future research. It is imperative to develop large-scale benchmark datasets encompassing a variety of real-world composite degradations to facilitate the learning of more authentic degradation distributions. Concurrently, the research focus should transition from single degradation restoration to complex composite degradation restoration, enhancing the reliability of models in real-world applications. Furthermore, designing more challenging restoration tasks tailored to practical applications—such as industrial vision, medical imaging, autonomous driving, and remote sensing image restoration—will further propel the advancement of AiOIR technology.

Integration with Large Multimodal Pre-trained Models and Generative Priors: In recent years, the progression of large vision-language models and generative models has offered a novel technological trajectory for AiOIR. Compared to traditional restoration networks, foundation models possess richer visual semantic comprehension capabilities and potent generative priors, which can assist models in reconstructing missing information under complex degradation conditions. Future research can further explore effective transfer methodologies for large pre-trained models—such as CLIP [9], DINO [10], and Stable Diffusion [20]—into AiOIR, enhancing the model's comprehension of degradation types and image content by introducing semantic priors. Simultaneously, integrating the formidable generative capabilities of diffusion models [20] can amplify the model's structural restoration proficiency in severely degraded regions. However, preventing generative models from hallucinating unrealistic details and ensuring the authenticity and consistency of restoration results remain a pressing issue to be addressed in the future.

Multimodal Information Fusion and Cross-Modal Restoration: At present, most AiOIR methods predominantly target RGB image restoration, while practical visual systems typically incorporate a richer array of information sources, such as depth information, infrared images, optical flow, and event data. Fusing multimodal information can supply models with more comprehensive structural and semantic information, enhancing restoration capabilities in complex

environments. Future research can explore unified multimodal AiOIR frameworks, realizing effective complementarity across distinct information sources through cross-modal feature fusion mechanisms. For instance, depth information can aid in restoring spatial structures, infrared information can provide target details in low-light environments, and optical flow can augment temporal consistency in video restoration. Nevertheless, given the issues of resolution discrepancies, spatial misalignments, and missing modalities across multimodal data, achieving efficient and reliable cross-modal fusion remains an important avenue for future research.

Standardized Evaluation Systems and Unified Benchmark Construction: Because AiOIR involves an array of restoration tasks, different studies currently employ disparate datasets, degradation settings, and evaluation metrics, resulting in a lack of fair and unified comparative standards across different methods. The future necessitates the establishment of a more comprehensive AiOIR benchmark, encompassing data with multiple real-world degradation types, complex degradation combinations, and diverse application scenarios. Simultaneously, a more holistic evaluation system should be constructed, comprehensively taking into account pixel-level restoration quality, perceptual quality, semantic consistency, and computational efficiency. Furthermore, alongside the development of generative restoration methods, traditional metrics like PSNR and SSIM are no longer sufficient to comprehensively evaluate model performance; thus, it is crucial to further explore novel assessment methodologies that integrate human evaluation, visual perceptual metrics, and semantic evaluation.

Expansion to High-Dimensional Visual Data and Intelligent Visual Systems: Currently, AiOIR is primarily concentrated on 2D static image restoration, whereas future application demands will inevitably expand to high-dimensional visual tasks, including video, 3D vision, hyperspectral, and dynamic event data. For instance, video restoration requires not only improving single-frame image quality but also maintaining temporal continuity and motion consistency; 3D vision restoration must simultaneously consider spatial structure and depth information; hyperspectral restoration needs to preserve information correlations across different bands. Therefore, the future trajectory of AiOIR is not merely achieving the unified restoration of diverse image degradations, but further constructing unified visual restoration frameworks capable of processing spatial, temporal, spectral, and multimodal information, thereby realizing an expansion from 2D image restoration to intelligent visual perception systems.

Conclusion Summary: Overall, future AiOIR research will gradually evolve from traditional multi-task image restoration towards a more intelligent, generalized, and highly efficient direction. By integrating robust multi-task learning, weakly supervised learning, foundation model transfer, multimodal fusion, and efficient deployment technologies, AiOIR models are poised to break through current limitations regarding unknown degradation adaptation, data dependency, and practical application deployment. The future all-in-one image restoration systems will no longer be confined to sheer image enhancement tasks, but will progressively evolve into intelligent visual restoration frameworks equipped with degradation comprehension, task planning, adaptive optimization, and cross-modal perception capabilities.

5. Conclusion

This paper presents a systematic review of All-in-One Image Restoration (AiOIR) technologies, conducting an in-depth analysis of network architecture designs, learning strategies, key methodological advancements, and future developmental trends, while also summarizing the primary challenges currently confronting the AiOIR domain. Although AiOIR has achieved remarkable progress in the unified restoration of multiple degradations, degradation perception, and intelligent restoration in recent years, it still suffers from critical bottlenecks, including task conflicts, inadequate generalization in real-world scenarios, high model complexity, and the limitations of existing datasets and evaluation systems. Future research endeavors are expected to focus on directions such as robust multi-task learning, lightweight model design, semi-supervised and unsupervised learning, the integration of multimodal foundation models, and the construction of standardized evaluation benchmarks. With the continuous evolution of vision-language models, generative models, and intelligent restoration frameworks, AiOIR technologies are poised to further elevate their adaptability in complex scenarios, providing more robust and effective technical support for intelligent image restoration applications in the field of computer vision.

References

- [1] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778). IEEE.
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- [3] Li, R., Tan, R. T., & Cheong, L.-F. (2020). All in one bad weather removal using architectural search. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3175–3185). IEEE.
- [4] Chen, X., Li, H., Li, M., & Guo, Y. (2022). Learning a unified image restoration network for multiple degradation types. *IEEE Transactions on Image Processing*, 31, 1928–1941. <https://doi.org/10.1109/TIP.2022.3151454>
- [5] Chen, H., Wang, Y., Guo, T., Xu, C., Deng, Y., Liu, Z., Ma, Y., Xu, J., Zheng, C., Gao, J., & Tao, D. (2021). Pre-trained image processing transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12299–12310). IEEE.
- [6] Li, X., Li, W., Zhang, X., & Sun, J. (2024). MEASNet: Multi-expert adaptive selection network for all-in-one image restoration. *IEEE Transactions on Image Processing*, 33, 1–14. <https://doi.org/10.1109/TIP.2024.3382712>
- [7] Tian, X., Liao, X., Liu, X., & Li, C. (2025). Degradation-aware feature perturbation for all-in-one image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 28165–28175). IEEE.
- [8] Zhang, J., Li, Y., Wang, H., & Zhang, X. (2025). FAPE-IR: Frequency-aware planning and execution for all-in-one image restoration. *arXiv preprint arXiv:2501.2025*.
- [9] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, R., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. In *Proceedings of the International Conference on Machine Learning* (pp. 8748–8763).

- [10] Caron, M., Touvron, H., Misra, I., Jegou, H., Mairal, J., Fergus, R., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9650–9660). IEEE.
- [11] Oquab, M., Darcet, T., Moutakanni, T., et al. (2023). DINOv2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*.
- [12] Wang, Z., Zhang, J., Zhang, X., & Sun, J. (2025). Perceive-IR: Learning perceptual image restoration with vision foundation models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE.
- [13] Luo, J., Huang, Y., Zhang, X., & Sun, J. (2024). DA-CLIP: A degradation-aware vision-language model for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.
- [14] Wang, H., Li, Y., Zhang, X., & Sun, J. (2025). VLU-Net: Vision-language guided deep unfolding network for image restoration. *IEEE Transactions on Image Processing*, 34, 1–12.
- [15] Potlapalli, Z., Zamir, S. W., Khan, S., & Khan, F. S. (2023). PromptIR: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36, 71234–71250.
- [16] Fan, D., Zhang, J., & Chang, L. (2024). ConStyle v2: A strong prompter for all-in-one image restoration. *arXiv preprint arXiv:2406.18242*.
- [17] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 4171–4186). Association for Computational Linguistics.
- [18] Zou, S., Li, C., & Zhang, X. (2024). InstructIR: High-quality image restoration following human instructions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.
- [19] Jiang, S., Chen, B., Wang, X., & Li, C. (2024). Defusion: Generative image restoration with visual instructions. *arXiv preprint arXiv:2403.2024*.
- [20] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840–6851.
- [21] Potlapalli, Z., Zamir, S. W., Khan, S., & Khan, F. S. (2023). PromptGIP: Adaptive image restoration via text-guided prompts. *arXiv preprint arXiv:2306.13039*.
- [22] Yu, Y., Liu, J., Zhang, X., & Tao, D. (2024). AutoDIR: Automatic all-in-one image restoration with unknown degradation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 18155–18164). IEEE.
- [23] Chen, H., Wang, Z., Zhang, Y., & Li, H. (2025). ClearAIR: A comprehensive framework for all-in-one image restoration with human-like perception. *arXiv preprint arXiv:2502.2025*.
- [24] Wang, Y., Zhang, Z., & Zhang, L. (2025). DCTTA: Degradation consistency guided test-time adaptation for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.
- [25] Li, J., Yang, X., & Loy, C. C. (2023). MiOIR: Multi-task interaction learning for all-in-one image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 12345–12354). IEEE.
- [26] Chen, S., Zhang, J., & Van Gool, L. (2024). SimpleIR: A simple framework for effective all-in-one image restoration. *arXiv preprint arXiv:2403.2024*.
- [27] McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 24, pp. 109–165). Academic Press.
- [28] Chaudhry, A., Rohrbach, M., Elhoseiny, M., Ajanthan, T., Dokania, P. K., Torr, P. H., Ranzato, M., & Vedaldi, A. (2018). On tiny episodic memories in continual learning. *arXiv preprint arXiv:1805.10976*.
- [29] Van de Ven, G. M., & Tolias, A. S. (2019). Three scenarios for continual learning. *arXiv preprint arXiv:1904.07734*.
- [30] Le, A. N. T., Nguyen, K., & C, T. A. G. (2022). TransWeather: Transformer-based restoration for adverse weather image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 2353–2362). IEEE.
- [31] Li, Z., Chen, J., Wang, Y., & Yang, X. (2020). All-in-one image restoration for adverse weather conditions. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 12359–12368). IEEE.
- [32] Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. In *Proceedings of the International Conference on Machine Learning* (pp. 1597–1607).
- [33] He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2020). Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9729–9738). IEEE.
- [34] Grill, J.-B., Strub, F., Altché, F., Tallec, C., Richemond, P. H., Buchatskaya, E., Doersch, C., Avila Pires, B., Guo, Z., Gheshlaghi Azar, M., Piot, B., Kavukcuoglu, K., Munos, R., & Valko, M. (2020). Bootstrap your own latent: A new approach to self-supervised learning. *Advances in Neural Information Processing Systems*, 33, 21271–21284.
- [35] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 586–595). IEEE.
- [36] Wang, X., Li, H., & Wang, Z. (2024). CPL-PromptIR: Content-aware prompt learning for all-in-one image restoration. *arXiv preprint arXiv:2405.2024*.
- [37] Liu, Y., Wang, H., & Chen, X. (2025). AnyIR: Adaptive all-in-one image restoration via universal prompts. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE.
- [38] Zhang, S., Li, Y., & Sun, J. (2025). MIRAGE: Multi-modal instruction guided all-in-one image restoration. *arXiv preprint arXiv:2501.2025*.
- [39] Chen, Y., Wang, W., & Zhang, J. (2025). WM-MoE: Weather-aware mixture-of-experts for all-in-one image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. IEEE.
- [40] Luo, Z., Hu, H., Huang, Y., & Zhang, S. (2023). Clarity-ChatGPT: A multimodal large language model for interactive image restoration. *arXiv preprint arXiv:2308.16811*.
- [41] Li, J., Li, Y., Chen, G., & Loy, C. C. (2024). Instruct-IPT: Multi-modal instruction tuning for image processing transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 21232–21242). IEEE.
- [42] Ke, V., Sun, Y., Li, J., & Yang, X. (2021). MUSIQ: Multi-scale image quality transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 5148–5157). IEEE.

- [43] Yang, S., Wu, L., Xia, J., & Zhang, W. (2022). MANIQA: Multi-dimension attention network for no-reference image quality assessment. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 1191–1200). IEEE.
- [44] Wang, J., Chan, K. C. K., & Loy, C. C. (2023). CLIP-IQA: Exploiting CLIP for zero-shot image quality assessment. *IEEE Signal Processing Letters*, 30, 133–137. <https://doi.org/10.1109/LSP.2023.3241338>
- [45] Chapelle, O., Schölkopf, B., & Zien, A. (Eds.). (2006). *Semi-supervised learning*. MIT Press.
- [46] Yang, J., Peng, D., Li, C., & Huang, H. (2021). Unsupervised image restoration via degradation learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE.
- [47] Zhang, K., Luo, W., Zhong, Y., & Li, H. (2022). Iterative restoration network for weakly supervised image restoration. *IEEE Transactions on Image Processing*, 31, 3525–3538. <https://doi.org/10.1109/TIP.2022.3172610>
- [48] Han, S., Mao, H., & Dally, W. J. (2016). Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. In Proceedings of the International Conference on Learning Representations.
- [49] Jacob, B., Kligys, S., Chen, B., Zhu, M., Tang, M., Howard, A., Adam, H., & Kalenichenko, D. (2018). Quantization and training of neural networks for efficient integer-arithmetic-only inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2704–2713). IEEE.
- [50] Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*.
- [51] Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. In Proceedings of the International Conference on Learning Representations.
- [52] Zoph, B., & Le, Q. V. (2017). Neural architecture search with reinforcement learning. In Proceedings of the International Conference on Learning Representations.
- [53] Cai, H., Zhu, L., & Han, S. (2019). ProxylessNAS: Direct neural architecture search on target task and hardware. In Proceedings of the International Conference on Learning Representations.
- [54] Gu, A., & Dao, T. (2024). Mamba: Linear-time sequence modeling with selective state spaces. In Proceedings of the International Conference on Learning Representations.