Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol. 31 No. 7s (2024)

Comprehensive Survey of Noise Strategies in Diffusion Model Frameworks

Sanjay Kumar[1], **Adhisha Sahoo**[2], **Hari Pratap**[3*]

[1]SAG, DRDO, New Delhi, Delhi-110054, India
[2]Manipal Institute of Technology, Manipal, Udupi, Karnataka-576104, India
[3]Department of Mathematics, PGDAV(E) College, University of Delhi, Delhi-110065, India *Corresponding Author: pratap.pgdavcollege.hari@gmail.com

Received: 02-11-2024

Revised: 07-11-2024

Accepted: 11-11-2024

Published: 19-11-2024

Abstract:

Diffusion models are swiftly evolving the state-of-the-art for image enhancement applications, resulting in a strong generalizable framework for restoration, super-resolution, inpainting, deblurring, low-light imaging, etc. This survey provides a novel and inclusive noise/initialization-based taxonomy with diffusion models placed in relationship to the type of noise mechanisms exploited: Additive Gaussian Noise (AGN), Conditional Noise Injection (CNI), Learned Noise (LN), and Poisson/Signal Dependent Noise (PSD). For each category, the main methods and benefits are discussed, focusing on how dimensional or hierarchical control of noise results in increased perceptual quality, improved restoration accuracy, and greater flexibility to compensate for real-world degradations. Such approaches include recent advances that combine scheduling, adaptive conditioning, and physical signal beliefs, such that the latter models generalise better across tasks and domains. By structurally regrouping diffusion models around noise strategy, we hope to provide enough clarity to rethink the future practicality of developing next-generation image enhancement solutions.

Keyword: Additive Gaussian Noise, Conditional Noise Injection, Gaussian Noise, Diffusion Model, Noise Strategies

1 Introduction

Diffusion Models recently became one of the most powerful generative and restoration paradigms in computer vision. Diffusion models can be thought of as a two-stage Markovian process: in the forward process (or the forward diffusion process), noise is added to an image iteratively over a series of time steps until the time-dependent image is indistinguishable from white noise; in the reverse set of processes (or reverse denoising process), a neural network (typically a U-Net) is trained to learn the reverse denoising evolution that restores the image to the moment before all additive noise, using each progressively noisier image to reconstruct the clean image over the series of discrete time steps. This probabilistic modeling framework is generally parameterized as a known (multivariate) Gaussian distribution or characterized as another sampling process by stochastic differential equations [25], allowing the diffusion model to act as a probabilistic data prior and allow for any other type of image restoration process with a very few constraints across a wide range of image enhancement tasks. Diffusion for image restoration fundamentally derives from classical

PDE-based diffusion methods. The development of anisotropic diffusion and reaction–diffusion models was foundational in demonstrating how diffusion can be guided to remove noise while simultaneously preserving edges [7] [2]; Because of this, anisotropic diffusion and reaction–diffusion models have become prevalent in, for instance, medical imaging and radiography. The introduction of fractional and nonlinear higher-order diffusion processes improved stability in the diffusion process while maintaining a more descriptive framework to model the more complex dynamics of images [3]. Hybrid processes that deployed isotropic diffusion with Perona–Malik style delivered superior denoising while maintaining structural integrity and fidelity [40]. Collectively, the classical models demonstrated the capacity for mathematical diffusion to assist in noise removal, contrast enhancement, and defect detection.

The emergence of deep generative models paved the way for the creation of denoising diffusion probabilistic models (DDPMs) and their derivatives. DDRM and SDEdit showed how it was possible to apply Gaussian noise schedules to guided denoising and editing; score-based generative models reframed diffusion in terms of stochastic differential equations [25]. Works like DiffIR exhibited the utility of these approaches in image restoration, and Guided Diffusion for adversarial purification provided an appropriate level of confidence regarding safety and robustness. Related frameworks such as Plug-and-Play diffusion models [56] and ResShift further improved computational efficiency by using priors along with early stopping strategies, while residual noise priors [35] [50] explicitly modelled noise structure to enhance restoration.

As tasks grew more varied, different applications of diffusion were developed in a specialized manner. DDFM applied noise-dependent blending for multimodal fusion (e.g., visible and infrared images) [54]. JPEG restoration models learned forward degradation while treating compression artifacts as statistical noise [15]. RePaint generalized non-uniform corruption to inpainting and more contextually faithful inpainting [22]. Adverse weather restoration models used patch-based denoising for fog and haze removal [28]. Cold Diffusion generalized the forward process to replace stochastic noise with deterministic forms of degradation, e.g., blur, indicating the universality of diffusion mechanisms [4].

The field quickly moved toward conditional and context-aware diffusion. InpDiffusion introduced mask-guided inpainting [39], and GradPaint employed gradient-based conditioning, giving sharper contexts for inpainted images [9]. NoiseCollage was a method that combined structured noise maps and multi-scale noise for layout-aware image generation [36]. Furthermore, Palette, PAIR Diffusion, Latent Paint, and Personalized Face Inpainting were created, extending the field with multimodal priors and semantic priors, giving rise to coherent object-aware editing and robust inpainting/outpainting [8] [6] [32] [43]. Super-resolution (SR) has

emerged as another significant application area. CDPMSR [27] and ACDMSR [26] demonstrated conditional SR pipelines, while ECDP utilized the more efficient probability flow sampling method [47]. ResDiff combined CNN residuals of the model with diffusion priors, SeeSR integrated semantic priors for real world SR, and SuperResDiff GAN integrated adversarial objectives with the idea of diffusion noise adaptation to attain higher frequency detail preservation [17] [41]. Multi-scale methods like PartDiff and Multi-scale adversarial diffusion networks allowed for adaptive noise distributed across a variation of image scales to ensure not only the proper global consistency but also facilitated fine detail recovery [53]. Each of these works demonstrated that learning noise and scheduling noise adaptively improves the perceptual fidelity of the output images in SR tasks. With evidence of sensor-aware and physics-grounded environments, researchers proposed Poisson

and signal-dependent diffusion models for low-light imaging and photon-limited imaging. Poisson-guided decomposition networks and Photon-counting flow models explicitly modelled discrete photon statistics, enabling more natural denoising in extremely dark environments. DiffRetinex [46], for example, DiffLLE [44], Pyramid Diffusion [55], RW-DM [18], etc, provided low-light image enhancement through the combination of physical priors and diffusion structure, while also respecting edges and the original structure of the image. Noise Synthesis for Low-Light [20], Global Structure Aware Diffusion [12], and LightenDiffusion (among others) [14], all showed how signal-dependent noise can be modelled in the diffusion method adaptively; preparing for actual sensor signal distributions produced more realistic reconstructions than Gaussian-based models. Each of these works demonstrated how the presence of signal-dependent noise found in many imaging applications makes diffusion models appropriate for cases like medical data, CT, and photon-limited imaging. Together, these contributions establish that the noise injection method has a significant impact on

the overall performance of diffusion models in various tasks. However, existing surveys tend to categorize diffusion models based on their application domain (restoration, SR, inpainting, fusion) or their architectural variant (DDPM, latent diffusion, score-based). In this paper, we present a noise-centred taxonomy that categorizes diffusion models into:

- **Additive Gaussian Noise (AGN)** based models, which are the most common approach, typically used for restoration and denoising;
- **Conditional Noise Injection (CNI)** based models, in which conditions from external sources such as masks and/or priors guide the noise schedule;
- **Learned Noise** based models, which can modify the noise process or learn to train it, to enable semantic- or uncertainty-based restoration; and
- **Poisson/Signal-Dependent** which can be understood as being physically embedded in sensor degradation processes such as photon counting.

This classification underscores that noise represents the unifying design axis that distinguishes diffusion models. By re-conceptualizing diffusion research around noise mechanisms, this taxonomy elucidates theoretical foundations, indicates why certain models perform best for specific types of degradations, and provides a principled guide for choosing or developing new models for existing tasks in appearance enhancement.

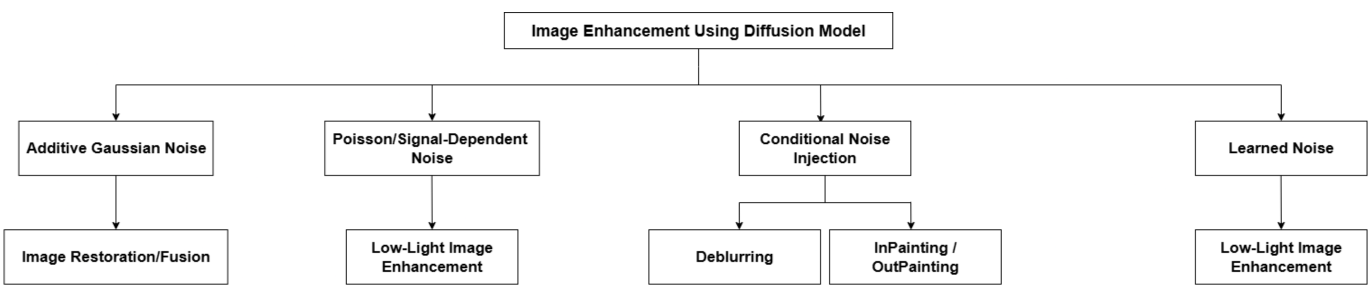

Figure 1: Proposed Classification

2 Proposed Classification

Diffusion Models have been substantially investigated over the last decade to achieve optimum results in the case of Image Enhancement. In this review, we have classified Diffusion Models used for Image Enhancement into - Additional Gaussian Noise, Learned Noise, Poisson/Signal Dependent Noise, and Conditional Noise Injection - based on the type of noise that is implemented by the model on the image, during the forward diffusion process. The details of classification is depicted in Fig.1

1 Additional Gaussian Noise

Modern generative and image restoration frameworks rely on a carefully structured stochastic mechanism for further Gaussian noise addition in diffusion models. Fundamentally, the forward diffusion process adds Gaussian noise iteratively at each discrete time step, usually indexed as t ranging from 1 to T, in order to convert a clean image into a series of increasingly noisier versions. This is controlled by a variance schedule, represented by $\{\beta_t\}_{t=1}^{T}$, which accurately regulates the amount of noise added at each stage. This process can be expressed mathematically as

$$x_t = \sqrt{1-\beta_t}\, x_{t-1} + \sqrt{\beta_t}\, \epsilon_t$$

where x_0 is the original image, ϵ_t is a separately sampled value from a standard multivariate normal distribution $\mathcal{N}(0, I)$ and β_t is a small, positive scalar that increases over time.

Gaussian noise is used owing to its statistical features, which include the characteristics of being memoryless, additive, and, after sufficient iterations, generating a final state x_T, that is nearly identical to pure white noise, i.e. $x_T \quad \mathcal{N}(0, I)$. This property is vital because it makes the reverse process mathematically tractable, as well as making it feasible to learn the denoising trajectory using powerful neural approximators [11]. In the opposite procedure, a neural network—typically a U-Net or similar architecture—is trained to determine the clean image x_0 straight from the noisy observation x, or the noise component ϵ_t added at each step. High-quality reconstructions and generative samples are empirically generated when the training objective intends to minimize the mean squared error between the predicted and actual noise. The broad application of this forward-noising and reverse-denoising Markov chain structure in restoration of pictures, fusion, artifact correction, and enhancement tasks displays both its theoretical sophistication and empirical reliability. [42] For instance, by using the noise process to blend and then denoise the fused representation, the diffusion process in image fusion models such as DDFM enables the smooth integration of multi-modal data (such as visible and infrared images) [54]. Similarly, the model learns to reverse complex degradations by simulating them as a series of Gaussian noise additions in restoration tasks like contrast enhancement or JPEG artifact correction [15]. This process is further refined by sophisticated diffusion model variations. While some use mixtures of Gaussians or even hybridize with reaction-diffusion or anisotropic diffusion models to capture more nuanced image statistics [2] [7], others introduce non-isotropic or correlated noise (such as blue noise) to better match the frequency characteristics of natural images. To significantly speed up inference without affecting restoration quality, early stopping strategies have also been proposed [23], in which the forward process is truncated before reaching full noise and the model is trained to denoise from these intermediary states. Another extension is score-based diffusion models [25], in which the neural network uses stochastic differential equations to guide the denoising process after learning the gradient (score) of the data distribution.

1.1 Image Restoration/Fusion

Image restoration is a key reference topic in both computer vision and image processing. Image restoration aims to obtain a high-quality image from a low-quality observation or reconstruction. Degradations can result from noise, blurring, compression artifacts, missing data, or other forms of distortion while an image is obtained, communicated, and stored. The goal in restoration is to create an image that appears and/or measures as close to that of the original or unaltered scene as possible. Typical applications of image restoration include noise removal, blur correction, haze removal, artifact removal, super-resolution/post-recovery super-resolution, inpainting, and other similar enhancements.

Additive Gaussian Noise (AGN) is considered to be the most suitable noise type for Image Restoration using Diffusion Models, due to several mathematical, statistical and practical reasons. It is isotropic in nature, which means that the noise uniformly affects the data space, making sure that the noise addition process does not introduce any bias or artifacts in the original image. This property is crucial for maintaining the structural integrity of the images as they are corrupted stepwise through a Markov Chain.

Analytically, Gaussian noise is tractable. Stability under addition and the central limit theorem make AGN a go-to choice for modeling the set of many small, independent noise sources [4]. This allows the

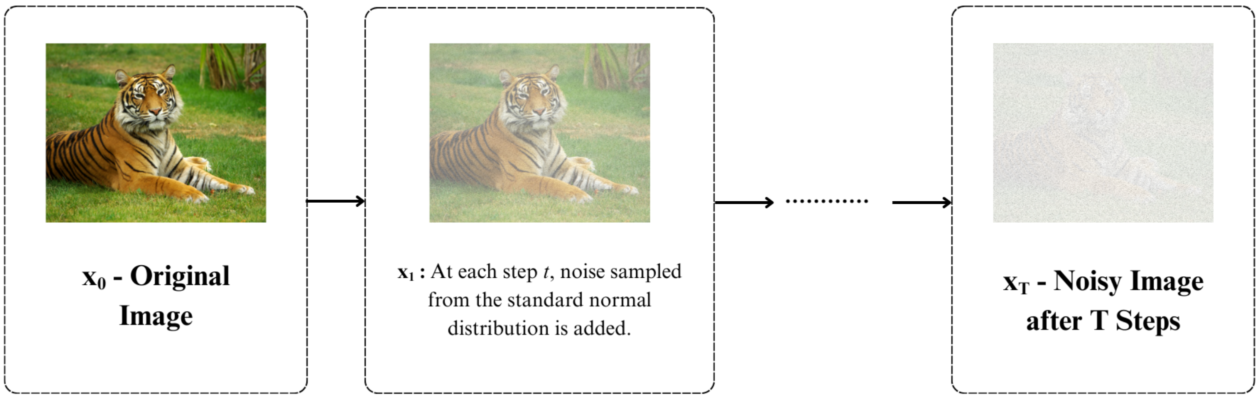

Figure 2: Forward Diffusion Process in Additive Gaussian Noise Type Diffusion Model

forward diffusion process to be carried out in the form of a simple Markov Chain that finally results in a standard normal distribution. The reverse diffusion process, which is basically denoising the noisy image, step by step, becomes efficiently parametrized and optimized.

A wide variety of diffusion models, using AGN specifically, have been developed for image restoration. Some of the most notable types include:

- **Denoising Diffusion Probabilistic Model (DDPM):** These form the foundational framework of diffusion models used for image enhancement. For image restorations, DDPMs are well-suited as they can robustly model the complex distribution of natural images and systematically remove the noise added in the forward diffusion process. [25]
- **Score-Based Generative Models (SBGM):** Instead of directly modeling the denoising process, these models learn the gradient/score of the data distribution. They guide the reverse diffusion process into generating a high-quality restoration by estimating how the probability density of clear images changes in the presence of noise. This approach is extremely powerful as it leverages the underlying geometry of the data manifold. [42] [38]
- **Plug - and - Play Model:** These models achieve a natural balance of flexibility and performance by using traditional image restoration models or priors in conjunction with the diffusion sampling process, to provide a more modular way to adapt to different types of image corruption without having to retrain the entire system from scratch. This is especially versatile because this provides the flexibility to adapt in situations where the form of the image corruption is not well-defined. Therefore, the diffusion models deliver restoration and manipulation solutions with improved performance when combined with task-specific tools. [56] [23] [51]

- **Accelerated and Efficient Diffusion Models:** These models combat the inherent computation demands of diffusion methods, which can take hundreds, if not thousands, of iterations to compute. There have been methods of optimization, such as residual shifting, that have reduced the inference time, but not at the expense of restoration capabilities. These are important methods in situations where rapid and efficient results are required, without affecting the fidelity of restored images. [50] [24]
- **Hybrid/Specialized Models:** These models take advantage of diffusion processes in conjunction with a prior or mathematical framework. For example, models that incorporate residual noise priors, Gaussian mixture models, or reaction-diffusion equations for specific restoration tasks (e.g., multi-modality fusion, adversarial purification, or artifact correction). The goal of these hybrid methods is to improve restoration effectiveness in special cases that leverage both domain knowledge and associated techniques to tackle complex or non-standard degradations. [54] [15]
- **Zero-Shot/Blind Models:** These methods, based on stochastic differential equations or dual priors, are trained to produce unseen degradations and generalizations for new tasks without the need for task-specific training. Zero-shot restoration models leverage the generalization ability of learned diffusion priors to apply robust restoration even when little is known about the image degradation. This promotes versatility and flexibility for real-world and open-source applications. [25]

In Sum.

Additive Gaussian noise (AGN) is the primary underlying mechanism in diffusion approaches to image restoration and enhancement. It maintains statistical tractability while uniformly and unbiasedly corrupting images, making denoising a mathematically convenient process with neural networks, such as UNet [42]. Classic DDPM models effectively express complex noise properties throughout their restoration. In contrast, score-based models help improve restoration quality, indirectly 'gradually' learning the probability gradients that exist in the data manifold [11] [25] [38]. Although plug-and-play and hybrid models permit greater flexibility by combining external priors or residual shifting, these also promote intensive domain specificity and customisation — able to flexibly handle tasks such as fusion, artifact or damage correction, and adversarial purification [56] [54] [15] [24] [51] [50]. Accelerated models (including those that leverage early stopping) satisfy an increasing and widespread demand for expedient results without sacrificing faithfulness or realism, while zero-shot/blind models using stochastic differential equations, when trained, create a strong confidence to generalize restoration to unseen image degradations. While AGN consistently outperforms other methods like impulse as a standard for restoration across a diversity of imaging tasks, its implications are varied. AGN possesses stability, flexibility, and superior restoration performance across diverse imaging tasks, such as denoising of raw data or fused data, bistatic, artificially included noise removal from images, generative noise removal, and deleting or transforming the image scene itself [25].

2 Conditional Noise Injection (CNI)

The core idea of CNI is a complex one in diffusion models, wherein the noise that is usually injected during the forward diffusion process comes under the influence of extrinsic conditions, such as degraded images, a semantic mask, or other ancillary information. This approach is the generalized form of the usual diffusion process; instead of a crude and pure random corruption, it is now an organised, information-preserving process linked to the task [48].

In traditional diffusion models, the forward process corrupts the input image by adding isotropic Gaussian noise at each timestep until it is completely noise, completely uninformed and regards all pixels equally. This is a practical approach when the goal is unconditional generation, but suboptimal for tasks that require guidance, such as picture inpainting. Thus, conditional noise injection completely alters this paradigm by permitting noise addition to depend on a given condition. Theoretically, this would indicate that the forward process is not only destroying information but also encoding into the noisy representation those cues that are relevant for the task at hand. [39].For example, a violent corruption in the masked areas (missing content) could be combined with a gentle noise injection in the known ones — this would lead the reverse process-generative power to be focused wherever it is most needed. For layout-aware synthesis, the noise map could then be spatially structured to conform to a desired layout so that spatial information is directly embedded into the corrupted input. [36]

This noise addition method allows the reverse (denoising) strategy to capitalize on both the noisy input itself, as well as the structure within the noise, to produce outputs that are better aligned with the condition and more suited to the restoration or generation task.

The standard forward process in a diffusion model is a Markov chain defined as:

$$q(\mathbf{x}_t \mid \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\,\mathbf{x}_{t-1}, \beta_t \mathbf{I}),$$

where $\mathbf{x}_0$ is the original image, $\mathbf{x}_t$ is the image at time step t, β_t is the variance schedule, and $\mathcal{N}$ denotes a Gaussian distribution. This process can be expressed in closed form as:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\mathbf{z},$$

where $\bar{\alpha}_t = \prod_{s=1}^{t}(1-\beta_s)$ and $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$.

In conditional noise injection, the noise addition is parameterized by a condition $\mathbf{c}$:

$$q(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{c}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t(\mathbf{c})}\,\mathbf{x}_{t-1}, \beta_t(\mathbf{c})\mathbf{I}) \tag{1}$$

or more generally:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t(\mathbf{c})}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t(\mathbf{c})}\,\mathbf{z}(\mathbf{c}) \tag{2}$$

Here, the condition $\mathbf{c}$ can affect:

- the variance schedule $\beta_t(\mathbf{c})$, making the noise addition adaptive to the condition;
- the noise sample $\mathbf{z}(\mathbf{c})$, allowing for spatially or semantically structured noise;
- or the noise distribution itself, which may be replaced by a non-Gaussian form parameterized by $\mathbf{c}$. [48]

Implementation variants of conditioned augmentation fall into three primary categories:

- **Adaptive variance**, wherein β_t is set as a function of the conditioning $\mathbf{c}$—for example, applying higher noise to masked regions in inpainting;
- **Conditioned noise sampling**, where one generates a noise sample that reflects spatial or semantic patterns from the condition, such as cropping and merging noise maps according to a layout;
- **Distribution substitution**, in which the noise distribution is replaced with a mixture of Gaussians or another form, with parameters determined by $\mathbf{c}$.

The mathematical consequence of these approaches is that the forward process remains a Markov chain, while the transition kernels become parameterized by the conditioning variable. If the noise remains Gaussian, the marginal distribution at each step becomes:

$$q(\mathbf{x}_t \mid \mathbf{x}_0, \mathbf{c}) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t(\mathbf{c})}\,\mathbf{x}_0, (1-\bar{\alpha}_t(\mathbf{c}))\mathbf{I}),$$

or otherwise, a more general distribution depending on the noise formulation.

The reverse process must then learn to denoise with respect to the condition, which is often implemented by either concatenating the condition $\mathbf{c}$ with $\mathbf{x}_t$ as input to the denoising network, or by incorporating attention mechanisms to modulate the denoising process based on $\mathbf{c}$.

A conceptual example can be used to illustrate this approach: Suppose the task is image inpainting, with the mask being the condition. In standard diffusion processes, noise is generated uniformly everywhere, sometimes with the consequence of destroying useful information even in known areas. With conditional noise injection, the model injects more noise in masked (unknown) regions and less in unmasked (known) regions. During denoising, the model learns to dedicate its generative capacity to reconstructing missing regions while keeping intact the areas treated as known originally. Hence, the noise map encodes, from the onset, where restoration is required and guides the reverse process.

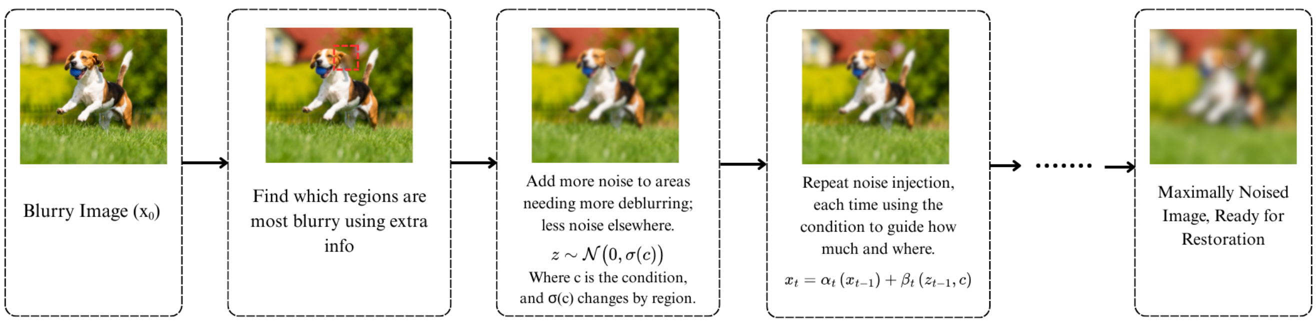

Figure 3: Forward Diffusion Process in Conditional Noise Injection Type Diffusion Model

2.1 Image Inpainting/Outpainting

CNI is the most favorable noise addition mode (when image corruption is performed), for image inpainting/outpainting, as it closely links the forward corruption process to the image's structure, mask, or context. Such combined advantage brings great enhancement in visual consistency, semantic consistency and task-agnostic property. [39] [36] [48] [8]

The principal theoretical benefit of CNI is the ability to schedule the noise region-wise or in context:

- In inpainting, CNI adds noise to the masked or missing pixels while preserving detail and limiting distortion in non-spanned regions. This strategy avoids the model wasting its capacity and resources trying to reconstruct regions that do not need to change, and instead focuses the generative effort and diversity into the inpainted areas. [39] [9]
- In multimodal and layout-aware tasks, noise maps are divided, cropped or combined based on object masks or layout commands, resulting in the model being provided with explicit per-object or per-region stochasticity. This effectively disentangles object attributes, eliminates blending artifacts, and enables spatially selective generation. [36]
- Joint conditioning with degradation maps, reference guidance and user input facilitates the setting of the magnitude and the structure of the noise, making the model more optimal for occlusion handling in the real-world mixed degradation or multi-modal cases. [48] [31]

Mathematically, the diffusion process is represented as:

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\,\mathbf{z},$$

where classical models assume $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I})$ as standard Gaussian noise.

In the case of Conditional Noise Injection (CNI), both the noise sample $\mathbf{z}$ and the variance schedule β_t become functions of a conditional signal $\mathbf{c}$, which may represent a mask, the observed context, or semantic maps. This allows for spatially-local or structure-aware corruption that effectively primes the reverse process to generate content that is contextually plausible and consistent across boundaries. [36] [39]colour

In practice, Conditional Noise Injection (CNI) yields substantial benefits:

- **More accurate semantic blending with fewer artifacts:** CNI couples noise patterns to the editable regions, helping the model generate color, gradients, and textures that are coherent with the surrounding context. [9]
- **Adaptable and flexible editing:** Our multimodal editors can demonstrate precise object-level control by biasing noise between shape and appearance. This enables robust strategies for outpainting, recoloring, and inpainting — all within a single, generalizable model. [8] [32]
- **Better Boundary Integrity:** Edge- and context-sensitive noise assignment helps diffusion models not to "bleed" information across mask/boundaries, leading to more realistic inpainted and outpainted regions. [39] [9]

2.1.1 Evidence from Recent Research

Several recent works have explored the use of Conditional Noise Injection (CNI):

- **InpDiffusion** proposed an adaptive conditional network to extract semantic edge features, and utilizes conditional noise to selectively guide the denoising inferred by the masked regions of the inputs, thus achieving better mask refinement and boundary fidelity. [39]
- **NoiseCollage** extends crop-and-merge and demonstrates that noise patches for object and background can be composed separately while maintaining parts of stochasticity optimal for a flexible layout-aware image synthesis. [36]
- **GradPaint** has shown that the conventional uniform noise degrades the harmony between the inpainted masks and their context while CNI constrains the global and local consistencies. [9]
- **The Joint Conditional Diffusion Model** outperformed single-degradation or unconditional techniques for mixed degradations by conditioning both the forward and reverse processes on degraded images and masks. This allows the model to adaptively adjust noise levels. [48]
- In order to ensure that the injected noise respects high-level structure and appearance, **PAIR Diffusion** and **Palette** both made use of multimodal conditions (such as text, panoptic maps, and reference images). This greatly increased controllability and realism for editing and outpainting jobs. [8] [32]
- According to **LatentPaint**, mask-independent inpainting is made possible by combining unconditional latent diffusion with mask-based and semantic CNI, producing outcomes that closely resemble global statistics and textures. [6]

In Sum.

CNI greatly improves diffusion models for image inpainting and outpainting by adding a condition to noise injection based on outside information, such as masks, semantic maps, or context, allowing different amounts of noise to be injected across different image regions [48]. As an example, InpDiffusion allows for refinement of mask boundaries by only injecting noise for conditions on masked images while keeping detail in important edge areas [39], and Gradpaint improves on the limitations of inpainting by non-uniform noise injection better consistency across inpainted and known regions [9]. NoiseCollage also builds upon many of these ideas, but allows the noise to be composed modularly as patches for objects and the background, allowing for layout-aware synthesis that includes some degree of spatial stochasticity [36]. PAIR Diffusion also builds upon these concepts by introducing multimodal conditions (e.g., panoptic maps, text), potentially improving control and realism in the generated content [8]. Finally, models like the Joint Conditional Diffusion Model demonstrate strong adaptability to cases of mixed degradation, conditioning both forward and reverse diffusion on the degraded inputs [48], and models like LatentPaint even successfully integrated unconditional latent diffusion with semantic CNI to enable mask-independent inpainting, successfully preserving both global textures and local structures [6]. In all of these examples of diffusion models conditioned on how to add noise, the noise schedules, which allow for or tailor in a region-specific manner, show considerable improvements in the integrity of the boundary, semantic coherence, and flexibility of the image editing process, securing CNI as the preferred noise process for inpainting and outpainting.

2.2 Deblurring

The process of restoring a crisp, clear image from a blurry input is known as image deblurring. Motion during image capture, camera shake, defocusing, optical flaws, and environmental factors are some of the causes of blurring. To restore the original image f while reducing noise n, the mathematical model commonly represented as g=Hf+n deconvolutes the blurred image g using a point spread function (PSF) H. Effectively reversing the blur's effect (often represented as a convolution) while preventing noise amplification and artifacts is the primary issue in deblurring, particularly when the blur's properties are unknown or differ between images.

Recent research in image restoration has shown that making the deblurring process data-dependent can greatly improve the results. In the case of diffusion models, this data dependency means the focus of the restoration process can vary based on input edge information, the blur statistics, and the granularity of the semantic context, rather than deblurring every part of the image the same. This intelligent approach makes for better reconstructions, particularly for hard and uneven blurs.

Conditional noise injection (CNI) uses input-specific clues, like estimated blur kernels or degradation maps, by directing noise from the generative capacity of a diffusion model to image regions impacted most by blur or uncertainty. In turn, allowing restoration to occur only in the corrupted areas of the image, it allows the model to effectively reconstruct sharp content wherever the image was previously degraded and maintain fidelity in additional regions of the image where fidelity was already high.

The adaptive accessibility of CNI virtues is important for the protection of necessary structural information and texture. By controlling where to inject noise and limiting the amount added, CNI can minimise the unnecessary corruption of non-blurry, sharp or clear areas of the image. This allows the model to exert less effort in trying to 'fix' more well-preserved content, and allows more of its 'effort' to be directed towards fixing the blur and reconstructing detail, only in the places where it was actually needed. In general, there is a trade-off, or compromise made by the model, and the resulting reconstructions are sharper and artifact-free with structures and natural textures protected to the maximum extent possible and remained intact through the overall restoration process.

Another advantage of this noise adding method is the ability to generalize to a larger range of blur patterns. The conditional noise process quickly adapts to unknown and diverse blur conditions for each image, including uniform, spatially varying, or unusual artifact causes. As a result, models with CNI recover better for real-world, non-uniform, and spatially varying blur situations for which more inflexible or unconditional methods can fail.

CNI is also important for strong performance in blind deblurring situations where the blur kernel is not known a priori or the degradation function has some other form of uncertainty. By adding noise degradation-aware—using information that is derived from the input itself—CNI helps orient the restoration towards plausible sharp solutions even when the blur itself is manifestly different or completely uninformed. Overall, this accommodation enables diffusion-based models to perform exceedingly well even when dealing with complicated or compound degradations while imposing more reliable and improved quality in arduous situations.

Finally, implementing Conditional Noise Injection in hierarchical or plug-and-play designs optimizes the deblurring procedures and also improves the efficiency and adaptivity. By allowing models to seamlessly combine and include various types of external priors, adaptively fuse information in a range of scales, and fasten the search to achieve convergence, CNI-based frameworks decrease computational cost and make operationalizing deblurring methods into downstream applications more feasible. This flexibility ensures that state-of-the-art restoration performance can be reasonably achieved across a diverse set of images and degradation forms.

2.2.1 Evidence from Recent Research

- **Cold Diffusion** developed a formalism which reverses arbitrary image transforms, not only noise, sometimes substituting noise with deterministic degradations (e.g., blur). This work emphasizes that the forward process of diffusion can be generalized, and conditioning on, in this case, image-specific, degradations can lead to high fidelity restoration, even when the complications of noise are omitted.

- **DiffBlur** utilizes a Latent Kernel Prediction Network (LKPN) to adaptively model pixelwise blur kernels for each input, and uses this information in the diffusion process. The noise injection then adapts accordingly, preserving sharp structures to target severely blurry areas for restoration, allowing for better detail and fidelity across both synthetic and real-world benchmarks. [16]

- **DifPIR** uses a plug-and-play denoising process in the diffusion framework where the model can use prior knowledge (e.g., degradation maps/blur) flexibly during both forward noise scheduling and reverse inference. This results in a lossless way to enhance robustness and perceptual quality of the source, while keeping inference efficient. [57]

- **DifFace** addresses blind face restoration by minimizing restoration backbone errors through a well-designed posterior transition (conditioned on the degraded input), resulting in a reliable restoration process even for unseen or severe degradations that highlight the power of condition-aware diffusion. [49]

- **DiffBIR** adopts a two-phased approach: first, the image-specific degradation must be removed (based on the input); second, the missing information is generated through a latent diffusion model and reconstructed over a diffusion chain based degradation-path controlled via noise and restoration modulation functions to allow flexible region/adaptive denoising and high quality synthesis for blind restoration tasks. [19]

- **Hierarchical Integration Diffusion Model (HI-Diff)** executes diffusion in a highly compressed latent representation, where the deblurring phase integrates condition-adaptive priors at various scales. While other merging integrations will generate with adaptation to appropriate denoising priors, the hierarchical integration facilitates a more effective generalization and efficiency across complex blurry scenarios since the priors and thus noise process are conditioned explicitly on the input blur characteristics. [5]

In Sum.

Conditional Noise Injection for image deblurring can use specific information about the input through one or more estimated blur kernels, degradation maps, or edge information to dynamically bypass, modulate, or control noise injection and carry out restoration on image zones with the worst blur. Cold Diffusion is a more generalized version of the standard forward diffusion process, where deterministic degradations like blur are considered for better restoration of image-specific degradations. DiffBlur performs pixelwise blur modeling variably through a Latent Kernel Prediction Network to introduce noise imperfectly under a certain condition while preserving sharp structure and restoring fine details [16]. The Posteriors, like DiffPIR, can use degradation priors from the initial noise scheduling step up to the denoising step to enhance robustness and perceptual quality [8]. DifFace is aimed toward blind-aware restoration of face images by degraded input information and has been demonstrated to generalize well to unseen or hard blur in testing [49] [5]. An elaborate situation having DiffBIR and HI-Diff comprises the adaptive prior condition with multi-scale or hierarchical latent-diffusion set-ups for an efficient, scalable image restoration method for many categories of blur, even those spatially varying [19]. All taken together, such work suggests that such a dependence on data and spatial awareness to noise injection via CNI is central to successful deblurring, more so toward a sharp, artifact-free, and more faithful and generalized reconstruction of the blur introduced by the world in real-life scenarios.

3 Learned Noise

A learnt noise diffusion model is a subset of the diffusion model family in which the noise process is parameterized and modified—learned directly from data and task context during model training—instead of being fixed (e.g., conventional isotropic Gaussian at every timestep) [35] [51]. Diffusion models can now inject, modulate, and remove noise in a way that is dynamically matched to the content and context of the image being restored, synthesized, or upsampled [17] [44]. This approach theoretically and mathematically removes the strong simplifying assumption that noise must be static and content-agnostic. The forward process in classical diffusion models applies predetermined, typically white Gaussian noise incrementally at each timestep to turn a data sample (such an image) into pure noise. Given that it treats all pixels, spatial regions, and image contexts as equally uncertain and challenging to reconstruct, this standardization may be less-than-ideal, particularly for tasks like super-resolution, inpainting, or restoration. However, it does make training and theoretical analysis easier. This limitation is addressed by the idea of learned noise, which parameterizes the forward noise process as a learnable function as opposed to treating it as immutable [35]. The input itself, uncertainty maps, semantic labels, structural cues, degradation levels, or any other domain-relevant signal may all influence this function [41] [44]. As a result, the model can "inject" more randomness where the mapping from degraded to high-fidelity content is unclear or underdetermined, and less where context constraints are tight or features are well-predicted [17]. Noise can also vary in space, time, and context. As the denoising procedure is adjusted to the precise geometry and statistics of the data, this results in increased model flexibility, better uncertainty modeling, and enhanced restoration performance.

The standard diffusion forward process is defined as a Markov chain, as shown in eq. (1) where β_t is the (typically hand-designed) noise variance schedule, and $\mathbf{I}$ is the identity matrix, yielding isotropic Gaussian noise. The cumulative process over T steps can be re-expressed in closed form shown in eq. (2) with $\bar{\alpha}_t = \prod_{s=1}^{t}(1-\beta_s)$.

Learned noise diffusion models generalize this framework by making both the variance schedule and/or the noise sampling process themselves functions of the input, conditioning context, or even learnable parameters. Mathematically, the forward process becomes:

$$q(\mathbf{x}_t \mid \mathbf{x}_{t-1}, \mathbf{c}) = \mathcal{N}\left(\mathbf{x}_t; \sqrt{1-\beta_t(\mathbf{c})}\,\mathbf{x}_{t-1}, \beta_t(\mathbf{c})\,\boldsymbol{\Sigma}(\mathbf{c})\right) \tag{3}$$

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol. 31 No. 7s (2024)

Or, more flexibly,

$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t(\mathbf{c})}\,\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t(\mathbf{c})}\,\mathbf{z}(\mathbf{c}) \tag{4}$$

where:

- $\mathbf{c}$ is any conditioning variable (e.g., LR image for SR, uncertainty maps, semantic masks).
- $\beta_t(\mathbf{c})$, $\mathbf{\Sigma}(\mathbf{c})$, and even $\mathbf{z}(\mathbf{c})$ are now functions learned from data, possibly via neural networks.
- In the most general case, the noise $\mathbf{z}$ itself may not be Gaussian, but sampled from a learned distribution tailored to the application's requirements.

In structure-aware super-resolution, for instance, the variance β_t can be kept higher for flat or ambiguous regions to inject diversity and/or lower for well-constrained ones; in semantic-aware models, the noise can be scheduled independently for any image region to further optimize content recovery. Some models go beyond this by learning mixture-of-Gaussian or even non-Gaussian noise distributions for greater generative expressiveness.

Backwards (or reverse) diffusion process, learns to filter out this learned context-dependent noise, thus obtaining high-fidelity or plausible candidates that are consistent with the context (e.g., LR image, semantic segmentation, mask) and statistically matched with the trained data distribution.

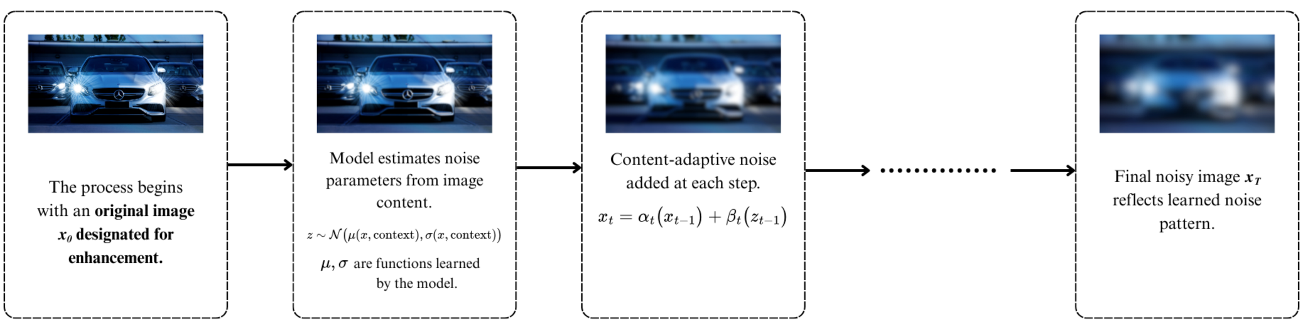

Figure 4: Forward Diffusion Process in Learned Noise Type Diffusion Model

3.1 Super Resolution (SR)

Image super-resolution (SR) is the recovery of a high-resolution (HR) image from a low-resolution (LR) counterpart. SR is an important operation in many areas, including medical imaging, remote sensing, scientific visualization, security video forensics, and consumer photo upscaling. The most challenging part is restoring high-frequency details that are lost from downsampling or sensor resolution, with the guarantee of both sharpness and perceptual coherence in the constructed HR image.

Diffusion models have become a preferred architecture for SR, capitalizing on their success in generative modeling and image restoration. They iteratively modify the noisy HR candidate, often guided by the LR image, so as to induce textures, structures, and plausible details that are not present in the input image. Early diffusion models performed a fixed-noise schedule for every image and timestep—often just isotropic Gaussian noise. However, conditioning the noise process on the LR input, on uncertainty estimates, or on semantic cue has proven to greatly improve quality and efficiency.

Learned noise refers to the approach in which the quantity, distribution, or style of noise injected during the diffusion process itself, is learned from data, usually made a function of the particular LR guidance, region content, or semantic priors. Applying noise can serve for restoration or editing functions. Rather than applying generic or uniform schedules, the noise processes can allocate selective model capacity - it adds less noise to regions that it can reliably predict (e.g., smooth surfaces) and more noise to areas that remain ambiguous or underdetermined (e.g., textured edges, fine details). This selective addition of noise helps to boost the diversity and realism in high-frequency areas, gain consistent restoration behaviour on semantic boundaries, and allow faster convergence during inference. For example, in ECDP [47], the noise schedule is modulated depending on the probability flow given by the LR structure, thereby avoiding unnecessary updates and finishing with high-quality HR results in fewer denoising steps. In the uncertainty-guided perturbation framework, the noise scale is, in fact, inversely proportional to uncertainty maps generated through the LR-to-HR mapping so that only the hard-to-predict regions are perturbed. This results in both more convincing and sharper HR reconstructions.

3.1.1 Evidence from Recent Research

Recent works provide compelling empirical and architectural support for the superiority of learned noise in diffusion-based super-resolution.

- **Efficient Conditional Diffusion Model with Probability Flow Sampling for Image Super-Resolution (ECDP)** introduced a conditional SR framework that tries to dynamically adjust the noise schedule on the LR so that the noise schedule depends on it and on structural cues. Hence, very little noise needs to be injected into those areas that can already be predicted with a high degree of confidence, and maximum generative effort is applied where LR evidence is weak, resulting in sharper HR outputs via fewer sampling steps. [47]

- **CDPMSR (Conditional Diffusion Probabilistic Models for Single Image Super-Resolution) and ACDMSR (Accelerated Conditional Diffusion Models for Super-Resolution)**, further demonstrated that conditioning both the forward diffusion process and reverse diffusion process on the LR image improves restoration performance. Especially in the real world, a restoration algorithm is needed to condition on the LR image, where the appropriate noise level is different for each semantic region and texture in the image. [27] [24]

- **Latent Space Super-Resolution for Higher-Resolution Image Generation with Diffusion Models** operates in learned feature space, where the injected noise is modulated by the semantic structure, allowing for fast SR inference and high-resolution restorations at minimal memory cost. [13]

- **ResDiff: Combining CNN and Diffusion Model for Image Super-Resolution** stresses the benefits of learned noise processes with CNN-based residuals, where noise is injected guided by the LR context and the residual feature maps. This allows for efficient HR reconstruction using the well-positioned residuals with highly contextual noise injected. [33]

- **Uncertainty-guided Perturbation for Image Super-Resolution Diffusion Model and SeeSR: Towards Semantics-Aware Real-World Image Super-Resolution** introduced models that provide estimates of region-wise uncertainty, or semantic segmentation maps, and would then modify the noise injection to incorporate the region-wise uncertainty. These approaches can produce highly photorealistic and structure-preserving SR results, which are crucial for medical, remote sensing or out-of-distribution images. [41] [52]

- **SupResDiffGAN** fuses diffusion and adversarial learning through adaptive, task-driven noise modulations used to stabilize training and enhance both sharpness and realism in the output generation of HR, with more speed and quality than GAN- or Diffusion-only baselines. [17]

- Multi-stage or cascaded diffusion is utilized in **RELAY DIFFUSION** and **Multi-scale Adversarial Diffusion Network**, so noise and denoising are learned and distributed hierarchically across resolutions. This allows for fine-detail recovery and a global consistency which is important for very high resolution and real-world SR tasks. [37] [34]

- **PartDiff** partitions the noise process across meaningful segments, focusing generative capacity on ambiguous or underdefined regions, while **XPSR (Cross-modal Priors for Diffusion-based SR)** uses external priors and semantic information to drive noise scheduling and enhance cross-domain SR fidelity. [53] [29]

In Sum.

Learned noise diffusion models stand out by utilizing a trainable noise process that is informed by not just the diffusion time step, but also the content of the underlying image, its semantics, and the uncertainty in each region. This enables the model to apply varying amounts of noise in different regions and contexts, allowing it to utilize its generative ability where uncertainty or ambiguity is greatest, and preserve structure and detail where the context is most defined. The benefits of learned noise are highly adaptive restoration, which allows the model to restore textures, fine details, and semantic boundaries in any region by varying the noise schedule it applies to each part of the image, depending on the requirements of that part of the image. This adaptivity is particularly useful in complex tasks such as image super-resolution and when degradation is spatially inconsistent. This flexibility also allows the learned noise model to extract and leverage uncertainty maps, semantic priors, or degradation cues, allowing for sharper and more perceptually realistic reconstructions, while avoiding the smoothing and artifacts common in static-noise models. Nevertheless, it is crucial to acknowledge the limitations associated with this power; learned noise is substantially more complex, more computationally intensive, and requires a significant amount of high-quality data and careful fine-tuning of the conditioning pathways. The specific conditioning pathways and network architecture also influence model performance and generalization, making the overall performance of learned noise more sensitive to these design choices. Taking everything into account, learned noise adds flexibility and strength to diffusion models, but introduces complexities during training and inference that may require significant engineering and computational efforts.

4 Poison/Signal Dependent Noise

Models for diffusion arising out of Poisson mechanics/signal noise form the primary stage of advancement over classical diffusion approaches for image scenarios dominated by signal-dependent, discrete noise, such as low-light photography, medical imaging via photon counting, or defect detection in industrial applications. Each diffusion step, instead, is mathematically defined with a Poisson kernel as its forward corruption process, rather than adding Gaussian noise of fixed variance. That is, for every pixel or signal element, the corrupted observation z_γ at diffusion level γ is sampled from a Poisson distribution with mean γx, directly mirroring the stochastic arrival of photons or counts observable in real sensors:

$$P(z_\gamma \mid x) = \frac{(\gamma x)^{z_\gamma} e^{-\gamma x}}{z_\gamma!}, \quad z_\gamma \in \{0, 1, 2, \ldots\} \tag{5}$$

[10] [30]

Here, noise mean and variance increase with the signal intensity, mirroring the empirical scenario of photon noise where brighter regions contain more information but are also more variable. With an increasing diffusion-step sequence γ any initial image transforms into a thoroughly randomized, Poisson-distributed array. Contrarily, in the context of Gaussian corruption, noise is uniform and agnostic of the signal, whereas Poisson diffusion ensures each location's uncertainty is aware of and physically meaningful concerning the signal.

From a mathematical perspective, the forward diffusion kernel at an intermediate step is given by $q(z_t \mid z_{t-1}, x) = \text{Poisson}(\lambda_t(x))$ where the schedule $\lambda_t(x)$ (x) increases monotonically so as to randomize the image. Note the main aspect that distinguishes this process, the noise is not additive and is data-dependent: image regions with weak signal vanish in the noise (just like in actual photon-starved sensors), whereas bright, more textured regions gather more noise energy-mimicking a more realistic model of degraded image.

Reverse diffusion in these models thus involves learning a generative mapping $p(x \mid z_t)$ that reconstructs the original, clean image x from noisy observations z_t. This is often operationalized as a neural network trained to minimize a Poisson negative log-likelihood loss

$$\mathcal{L}_{\text{Poisson}} = -\sum_i \left(z_i \log \hat{x}_i - \hat{x}_i - \log(z_i!)\right)$$

which is both physically grounded and provides an exact likelihood term for image counts or intensities [10]. Some Poisson/signal diffusion models are placed in continuous time or high-dimensional space (as with Poisson Flow Generative Models), wherein data evolves along the vector field solution of the Poisson equation-just like how physical fields propagate. The forward trajectory follows stochastic, data-dependent progressions prescribed on the image intensity, while the reverse tries to bring data to the original manifold. [21]

The advantages are quite clear: signal-based systems are inherently able to adapt to scenarios where noise statistics are signal-dependent and discrete, unlike Gaussian-based models, which have the tendency of oversmoothing textures or creating artifacts in the presence of quantization, degradation, or spartan photon counts [30] [46]. Poisson-guided models preserve edges and fine details mainly because the noise adaptation is strongest in flat or low-signal areas, while being numerically more stable and robust regarding training and inference, precisely due to an exact match with the expected data likelihoods. These models are thus the best fit towards reconstructing images from photon-limited sensors, denoising in medical and scientific spaces, as well as enhancing highly-degraded images towards defect detection with severely low SNR. [55] [44]

Innovations like Information-Theoretic Discrete Poisson Diffusion (ItDPDM) eliminate the need for variational approximations. They use exact Poisson-based likelihoods to reinforce statistics during training and inference, which enhances reliability. Signal diffusion formulations, particularly those drawn from Poisson field equations, provide both physical consistency and generative flexibility. The forward process is no longer uniform or random; it is controlled by field theory, leading to better convergence and realism. [18] [45]

In practical terms, this means that Poisson/signal diffusion frameworks work better than Gaussian models when imaging is limited by photon count or sensor physics. They offer sharper restoration, introduce fewer artifacts, recover more faithfully in low-light, medical, or quantized situations, and are more resilient to the modeling and dequantization artifacts that often affect classical models. The solid math and real sensor data backing these techniques make them the best option for the next generation of image improvement, restoration, and scientific imaging systems.

4.1 Low Light Image Enhancement

Poisson/signal-dependent noise diffusion models are excellent for low-light enhancement as they "naturally" model the heteroscedastic, signal-dependent nature of photon noise under these conditions. As a result, they

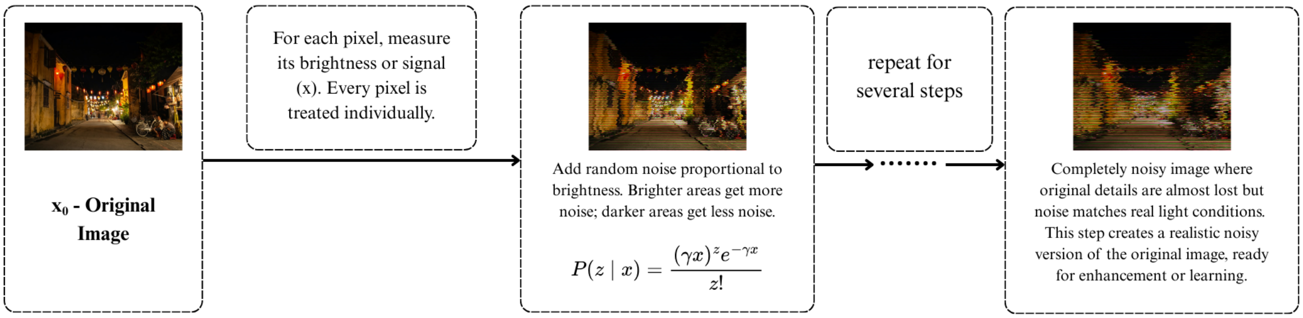

Figure 5: Forward Diffusion Process in Poison/Signal Dependent Noise Type Diffusion Model

afford accurate noise modeling with appropriately low noise variance for darker, lower photon regions, and with proportionately more noise for brighter, higher photon count regions, giving restoration of both light highlights and subtle shadows [30].

Furthermore, unlike Gaussian-based diffusion, which frequently calls for approximations or variance stabilisation transforms, the precise Poisson likelihood formulations used to train these models allow for stable, theoretically-grounded optimisation that is more appropriate for discrete photon count data [30]. This makes models more resilient to domain shifts and improves generalisation to real-world low-light photos, which differ greatly in sensor characteristics and illumination levels.

By using reaction-diffusion dynamics in conjunction with Poisson noise modeling, models that employ signal-dependent noise can also improve underlying image structures while suppressing noise. This combined effect smoothes noise in flat areas and enhances the recovery of high-frequency details [1] [12]. Thus, in addition to having physical significance, the noise process serves as a learnt prior that aids in adaptively separating noise from the real signal.

4.1.1 Evidence From Research

- **Poisson-Guided Decomposition Network for Extreme Low-Light Image Enhancement** showed that by directly modeling signal-dependent Poisson noise, the method could better separate real image structure from noise, even in very low light. By adjusting the diffusion and enhancement process to fit the actual noise distribution in low-light sensor readings, the model significantly improved the preservation of fine edges and complex textures that traditional Gaussian-based enhancement would often smooth out or lose. This approach created visual results with clear boundaries and detailed surfaces, especially in difficult, dark environments. [30]

- **Photon-Counting CT Denoising using Unsupervised Poisson Flow Generative Models** advanced the field by demonstrating that generative models designed for Poisson noise can deliver top-notch denoising results in photon-limited CT scans. In these scans, each pixel count directly reflects the sparse measurement of incoming photons. Importantly, their method functioned well even without paired clean and noisy data for guidance. This unsupervised approach is especially relevant for low-light photography, as it shows that effective signal-aware noise modeling allows for strong restoration in data-scarce or challenging imaging conditions. This research connects simulation with real-world sensor data. [10]

- **Noise Synthesis for Low-Light Image Denoising with Diffusion Models** showed the benefits of simulating Poisson noise during the training of diffusion models for denoising. By creating artificial noisy images that closely resembled actual sensor noise from low-light cameras, they trained denoisers that yielded more realistic, artifact-free results. Their research demonstrated that models adjusted with signal-dependent noise synthesis perform better than standard diffusion models, which often cause over-smoothing, leave residual artifacts, or struggle to adapt in low-light conditions. [20]

- **Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model** and **DiffLLE: Diffusion-guided Domain Calibration for Unsupervised Low-light Image Enhancement** both showed that Poisson-based diffusion modeling is essential for overall low-light recovery. **Diff-Retinex** proved that using Poisson noise not only improved the reconstruction of lost brightness and detail but also allowed for a more accurate restoration of dynamic range. This ensured that highlights and shadows stayed balanced and visually realistic. **DiffLLE** built on these ideas, demonstrating that domain calibration techniques based on signal-dependent diffusion models could effectively transfer enhancement skills across very different domains and device types. This led to images with better perceptual quality, natural light balance, and adaptability to various lighting conditions. These results are hard to achieve with static, Gaussian-noise-based methods. [46] [44]

In Sum.

Poisson and signal-dependent Diffusion Models provide major advantages for low-light image enhancement and for tasks that are also sensitive to the photon statistics, such as in medical and scientific imaging applications. Diffusion models leverage noise as a function of local signal intensity, rather than fixed or Gaussian noise processes; this helps to reflect the physical reality of the photon-limited sensor, and allows for custom denoising for recovery of shadows and highlights from the image, corrupted by noise. Using a signal-aware strategy is a good method to enhance the textural content of low-light imagery while maintaining local edge fidelity and a natural balance of global lightness, where one has fewer artifacts or saturation and a higher dynamic range than the existing methods. In addition, exact Poisson-likelihood training allows for handling discrete count data in a robust mathematical framework, which helps with resilience to different types of sensors or lighting in the scenes. This option is not without some limitations, as all these models depend on accurate signal statistics, and they are also less generalizable to all tasks if the source of noise in images is not intrinsically referred to as signal-dependent. They may also require some aspects of domain knowledge when implemented for optimal denoising performance, as this may be less extensible for other types of image restoration or enhancement applications.

Conclusion

This review provides a noise-based framework for classifying image enhancement diffusion models. A taxonomy of diffusion models based on noise mechanisms enables the classification of models into four categories: Additive Gaussian Noise (AGN), Conditional Noise Injection (CNI), Learned Noise, and Poisson/Signal-Dependent. The AGN model remains the state-of-the-art for restoration since AGN is easily scalable across dimensions while remaining robust and tractable. The CNI models introduce critical advancements toward flexible and region-aware editing, improving boundary fidelity across a variety of image enhancement tasks. The learned noise models prioritize flexible applicability and perceptual quality in difficult tasks such as super-resolution, using noise characteristics that are tailored based on the means of uncertainty and task context. The Poisson/signal-dependent models abstract significant benefits over low-light and sensor-limited images by utilizing noise that fits real-world optical sensor patterns. Taken independently but in concert, the field continues to advance in all classes of noise methods whereby scheduling of noise function and conditioning provides improvements in quality, speed, and generalization outlines the progress made and the importance

Noise Type	Typical Applications	Key Advantages	Main Limitations	Example Models/ References
Additive Gaussian Noise (AGN)	Restoration, Fusion, Denoising	Statistically tractable, unbiased, simple reverse	Less adaptive for structured degradation	DDPM, DiffIR, DiffPIR, ResShift
Conditional Noise Injection (CNI)	Inpainting, Outpainting, Deblurring	Region/task-specific control, better semantic blending	Requires suitable prior/condition, complex design	InpDiffusion, GradPaint, PAIR
Learned Noise	Super-resolution, Semantic Tasks	Data- and region-adaptive, uncertainty modeling	Training complexity, needs rich data	ECDP, ResDiff, SeeSR
Poisson/ Signal Dependent Noise	Low-light Enhancement, Medical/ Scientific Imaging	Matches sensor statistics, edge/texture preservation, robust to low SNR	Not ideal for tasks with uniform/ global noise, may require domain-specific tuning	Poisson Flow, DiffRetinex, DiffLLE

Figure 6: Comparison Table

of continuing to develop task-relevant noise methods to improve practical and robust image enhancement outcomes [23]

References

[1] A Abirami, P Prakash, and Y-K Ma. Variable-order fractional diffusion model-based medical image denoising. *Math. Probl. Eng.*, 2021:1–10, December 2021.

[2] Noureddine Alaa. Generic reaction-diffusion model with application to image restoration and enhancement. *Unknown*, 2018.

[3] Author1FirstName Author1LastName and Author2FirstName Author2LastName. Finite difference and legendre spectral method for a time-fractional diffusion-convection equation for image restoration. *Journal Name*, Volume Number(Issue Number):Page Range, 2025.

[4] Arpit Bansal, Eitan Borgnia, Hong-Min Chu, Jie S Li, Hamid Kazemi, Furong Huang, Micah Goldblum, Jonas Geiping, and Tom Goldstein. Cold diffusion: Inverting arbitrary image transforms without noise. August 2022.

[5] Zheng Chen, Yulun Zhang, Ding Liu, Bin Xia, Jinjin Gu, Linghe Kong, and Xin Yuan. Hierarchical integration diffusion model for realistic image deblurring. 2023.

[6] Ciprian Corneanu, Raghudeep Gadde, and Aleix M Martinez. LatentPaint: Image inpainting in latent space with diffusion models. In *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 4322–4331. IEEE, January 2024.

[7] M. B. Gharsallah and E. B. Braiek. New anisotropic diffusion method to improve radiographic image quality. *Kuwait Journal of Science*, 2017.

[8] Vidit Goel, Elia Peruzzo, Yifan Jiang, Dejia Xu, Nicu Sebe, Trevor Darrell, Zhangyang Wang, and Humphrey Shi. Pair-diffusion: Object-level image editing with structure-and-appearance paired diffusion models. *arXiv preprint arXiv:2303.17546*, 2023.

[9] Asya Grechka, Guillaume Couairon, and Matthieu Cord. Gradpaint: Gradient-guided inpainting with diffusion models. 2023.

[10] Dennis Hein, Staffan Holmin, Timothy P. Szczykutowicz, Mats Danielsson, and Mats Persson. Noise suppression in photon-counting computed tomography using unsupervised poisson flow generative models. *Visual Computing for Industry, Biomedicine and Art*, 7(1):24, 2024.

[11] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. 2020.

[12] Jinhui Hou, Zhiyu Zhu, Junhui Hou, Hui Liu, Huanqiang Zeng, and Hui Yuan. Global structure-aware diffusion process for low-light image enhancement. October 2023.

[13] Jinho Jeong, Sangmin Han, Jinwoo Kim, and Seon Joo Kim. Latent space super-resolution for higher-resolution image generation with diffusion models. March 2025.

[14] Hai Jiang, Ao Luo, Xiaohong Liu, Songchen Han, and Shuaicheng Liu. LightenDiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. July 2024.

[15] Bahjat Kawar, Jiaming Song, Stefano Ermon, and Michael Elad. JPEG artifact correction using denoising diffusion restoration models. 2022.

[16] Lingshun Kong, Jiawei Zhang, Dongqing Zou, Jimmy Ren, Xiaohe Wu, Jiangxin Dong, and Jinshan Pan. Deblurdiff: Real-world image deblurring with generative diffusion models. *arXiv preprint arXiv:2502.03810*, 2025.

[17] Dawid Kopeć, Wojciech Kozłowski, Maciej Wizerkaniuk, Dawid Krutul, Jan Kocoń, and Maciej Zięba. SupResDiffGAN a new approach for the super-resolution task. In *Lecture Notes in Computer Science*, Lecture notes in computer science, pages 66–80. Springer Nature Switzerland, Cham, 2025.

[18] Yadong Li, Jun Tian, Yang Chen, Hongdong Wang, Hui Yan, Yang Peng, and Tianjiao Wang. Rw-Dm: Retinex and wavelet-based diffusion model for low-light image enhancement in underground coal mines. *Complex Intell. Syst.*, 11(8), August 2025.

[19] Xinqi Lin, Jingwen He, Ziyan Chen, Zhaoyang Lyu, Bo Dai, Fanghua Yu, Yu Qiao, Wanli Ouyang, and Chao Dong. DiffBIR: Toward blind image restoration with generative diffusion prior. In *Lecture Notes in Computer Science*, Lecture notes in computer science, pages 430–448. Springer Nature Switzerland, Cham, 2025.

[20] Liying Lu, Raphael Achddou, and Sabine Susstrunk. Dark noise diffusion: Noise synthesis for low-light image denoising. *IEEE Trans. Pattern Anal. Mach. Intell.*, PP, August 2025.

[21] Liying Lu, Raphael Achddou, and Sabine Susstrunk. Dark noise diffusion: Noise synthesis for low-light image denoising. *IEEE Trans. Pattern Anal. Mach. Intell.*, PP, August 2025.

[22] Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. RePaint: Inpainting using denoising diffusion probabilistic models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11451–11461. IEEE, June 2022.

[23] Zhaoyang Lyu, Xudong Xu, Ceyuan Yang, Dahua Lin, and Bo Dai. Accelerating diffusion models via early stop of the diffusion process. 2022.

[24] Zhaoyang Lyu, Xudong Xu, Ceyuan Yang, Dahua Lin, and Bo Dai. Accelerating diffusion models via early stop of the diffusion process. 2022.

[25] Chenlin Meng, Yutong He, Yang Song, Jiaming Song, Jiajun Wu, Jun-Yan Zhu, and Stefano Ermon. SDEdit: Guided image synthesis and editing with stochastic differential equations. August 2021.

[26] Axi Niu, Trung X Pham, Kang Zhang, Jinqiu Sun, Yu Zhu, Qingsen Yan, In So Kweon, and Yanning Zhang. ACDMSR: Accelerated conditional diffusion models for single image super-resolution. *IEEE Trans. On Broadcast.*, 70(2):492–504, June 2024.

[27] Axi Niu, Kang Zhang, Trung X Pham, Jinqiu Sun, Yu Zhu, In So Kweon, and Yanning Zhang. CDPMSR: Conditional diffusion probabilistic models for single image super-resolution. February 2023.

[28] Ozan Ozdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Trans. Pattern Anal. Mach. Intell.*, 45(8):10346–10357, August 2023.

[29] Yunpeng Qu, Kun Yuan, Kai Zhao, Qizhi Xie, Jinhua Hao, Ming Sun, and Chao Zhou. XPSR: Cross-modal priors for diffusion-based image super-resolution. 2024.

[30] Isha Rao and Sanjay Ghosh. A poisson-guided decomposition network for extreme low-light image enhancement. June 2025.

[31] Chitwan Saharia, William Chan, Huiwen Chang, Chris A Lee, Jonathan Ho, Tim Salimans, David J Fleet, and Mohammad Norouzi. Palette: Image-to-Image diffusion models. November 2021.

[32] Chitwan Saharia, William Fedus, Saurabh Khanna, Shariq Iqbal, Ben Wang, Dmitry Kalenichenko, Mohammad Norouzi, and William T. Freeman. Palette: Image-to-Image Diffusion Models. *arXiv preprint arXiv:2111.05826*, 2021.

[33] Shuyao Shang, Zhengyang Shan, Guangxing Liu, Lunqian Wang, Xinghua Wang, Zekai Zhang, and Jinglin Zhang. ResDiff: Combining CNN and diffusion model for image super-resolution. March 2023.

[34] Yanli Shi, Xianhe Zhang, Yi Jia, and Jinxing Zhao. Multi-scale adversarial diffusion network for image super-resolution. *Sci. Rep.*, 15(1):11690, April 2025.

[35] Zhenning Shi, Haoshuai Zheng, Chen Xu, Changsheng Dong, Bin Pan, Xueshuo Xie, Along He, Tao Li, and Huazhu Fu. Resfusion: Denoising diffusion probabilistic models for image restoration based on prior residual noise. November 2023.

[36] Takahiro Shirakawa and Seiichi Uchida. NoiseCollage: A layout-aware text-to-image diffusion model based on noise cropping and merging. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8921–8930. IEEE, June 2024.

[37] Jiayan Teng, Wendi Zheng, Ming Ding, Wenyi Hong, Jianqiao Wangni, Zhuoyi Yang, and Jie Tang. Relay diffusion: Unifying diffusion process across resolutions for image synthesis. 2023.

[38] Jinyi Wang, Zhaoyang Lyu, Dahua Lin, Bo Dai, and Hongfei Fu. Guided diffusion model for adversarial purification. 2022.

[39] Kai Wang, Shaozhang Niu, Qixian Hao, and Jiwei Zhang. InpDiffusion: Image inpainting localization via conditional diffusion models. January 2025.

[40] Na Wang, Yu Shang, Yang Chen, Min Yang, Quan Zhang, Yi Liu, and Zhiguo Gui. A hybrid model for image denoising combining modified isotropic diffusion model and modified Perona-Malik model. *IEEE Access*, 6:33568–33582, 2018.

[41] Rongyuan Wu, Tao Yang, Lingchen Sun, Zhengqiang Zhang, Shuai Li, and Lei Zhang. SeeSR: Towards semantics-aware real-world image super-resolution. November 2023.

[42] Bin Xia, Yulun Zhang, Shiyin Wang, Yitong Wang, Xinglong Wu, Yapeng Tian, Wenming Yang, and Luc Van Gool. Diffir: Efficient diffusion model for image restoration. *ICCV*, 2023.

[43] J. Xu, S. Motamed, P. Vaddamanu, C. Wu, C. Häne, J. Bazin, and F. la Torre. Personalized face inpainting with diffusion models by parallel visual attention. *CoRR*, abs/2312.03556, 2023.

[44] Shuzhou Yang, Xuanyu Zhang, Yinhuai Wang, Jiwen Yu, Yuhan Wang, and Jian Zhang. DiffLLE: Diffusion-guided domain calibration for unsupervised low-light image enhancement. August 2023.

[45] Chia-Hung Yeh and Cheng-Yue Liou. Diffusion-based low-light image enhancement with Kolmogorov-Arnold networks (KANs). *Array (N. Y.)*, 27(100431):100431, September 2025.

[46] Xunpeng Yi, Han Xu, Hao Zhang, Linfeng Tang, and Jiayi Ma. Diff-retinex: Rethinking low-light image enhancement with a generative diffusion model. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12268–12277. IEEE, October 2023.

[47] Yutao Yuan and Chun Yuan. Efficient conditional diffusion model with probability flow sampling for image super-resolution. April 2024.

[48] Yufeng Yue, Meng Yu, Luojie Yang, and Tong Liu. Joint conditional diffusion model for image restoration with mixed degradations. *Neurocomputing*, 626(129512):129512, April 2025.

[49] Zongsheng Yue and Chen Change Loy. DifFace: Blind face restoration with diffused error contraction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 46(12):9991–10004, December 2024.

[50] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. ResShift: Efficient diffusion model for image super-resolution by residual shifting. 2023.

[51] Zongsheng Yue, Jianyi Wang, and Chen Change Loy. Efficient diffusion model for image restoration by residual shifting. *IEEE Trans. Pattern Anal. Mach. Intell.*, 47(1):116–130, January 2025.

[52] Leheng Zhang, Weiyi You, Kexuan Shi, and Shuhang Gu. Uncertainty-guided perturbation for image super-resolution diffusion model. March 2025.

[53] Kai Zhao, Alex Ling Yu Hung, Kaifeng Pang, Haoxin Zheng, and Kyunghyun Sung. PartDiff: Image super-resolution with partial diffusion models. 2023.

[54] Zixiang Zhao, Haowen Bai, Yuanzhi Zhu, Jiangshe Zhang, Shuang Xu, Yulun Zhang, Kai Zhang, Deyu Meng, Radu Timofte, and Luc Van Gool. DDFM: Denoising diffusion model for Multi-Modality image fusion. 2023.

[55] Dewei Zhou, Zongxin Yang, and Yi Yang. Pyramid diffusion models for low-light image enhancement. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 1795–1803, California, August 2023. International Joint Conferences on Artificial Intelligence Organization.

[56] Yuanzhi Zhu, Yutong Li, Xing Guo, Song He, and Shuhang Gu. Denoising diffusion models for plug-and-play image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1220–1229, 2023.

[57] Yuanzhi Zhu, Kai Zhang, Jingyun Liang, Jiezhang Cao, Bihan Wen, Radu Timofte, and Luc Van Gool. Denoising diffusion models for plug-and-play image restoration. 2023.