

图像恢复和增强的扩散模型——综合综述

李鑫、任玉林、金鑫、兰翠玲、王星瑞、曾文军、王新超、陈志博

摘要——图像恢复 (IR) 一直是低级视觉领域中不可或缺且具有挑战性的任务，它致力于提高因各种形式的退化而扭曲的图像的主观质量。最近，扩散模型在 AIGC 的视觉生成方面取得了重大进展，从而提出了一个直观的问题，“扩散模型是否可以促进图像恢复”。为了回答这个问题，一些开创性的研究试图将扩散模型集成到图像恢复任务中，从而获得了比以前基于 GAN 的方法更出色的性能。尽管如此，对基于扩散模型的图像恢复的全面和启发性的调查仍然很少。在本文中，我们首次全面回顾了最近基于扩散模型的图像恢复方法，涵盖了学习范式、条件策略、框架设计、建模策略和评估。具体来说，我们首先简要介绍扩散模型的背景，然后介绍两种利用扩散模型进行图像恢复的流行工作流程。随后，我们对使用扩散模型进行 IR 和盲/现实世界 IR 的创新设计进行分类和强调，旨在激发未来的发展。为了全面评估现有方法，我们总结了常用的数据集、实现细节和评估指标。此外，我们还对三个任务中的开源方法进行了客观比较，包括图像超分辨率、去模糊和修复。最后，根据现有研究的局限性，我们提出了基于扩散模型的 IR 未来研究的五个潜在且具有挑战性的方向，包括采样效率、模型压缩、失真模拟和估计、失真不变学习和框架设计。存储库将在 <https://github.com/lixinustc/Awesome-diffusion-model-for-image-processing/> 发布

索引词——扩散模型、图像恢复、图像增强、图像超分辨率

1 引言

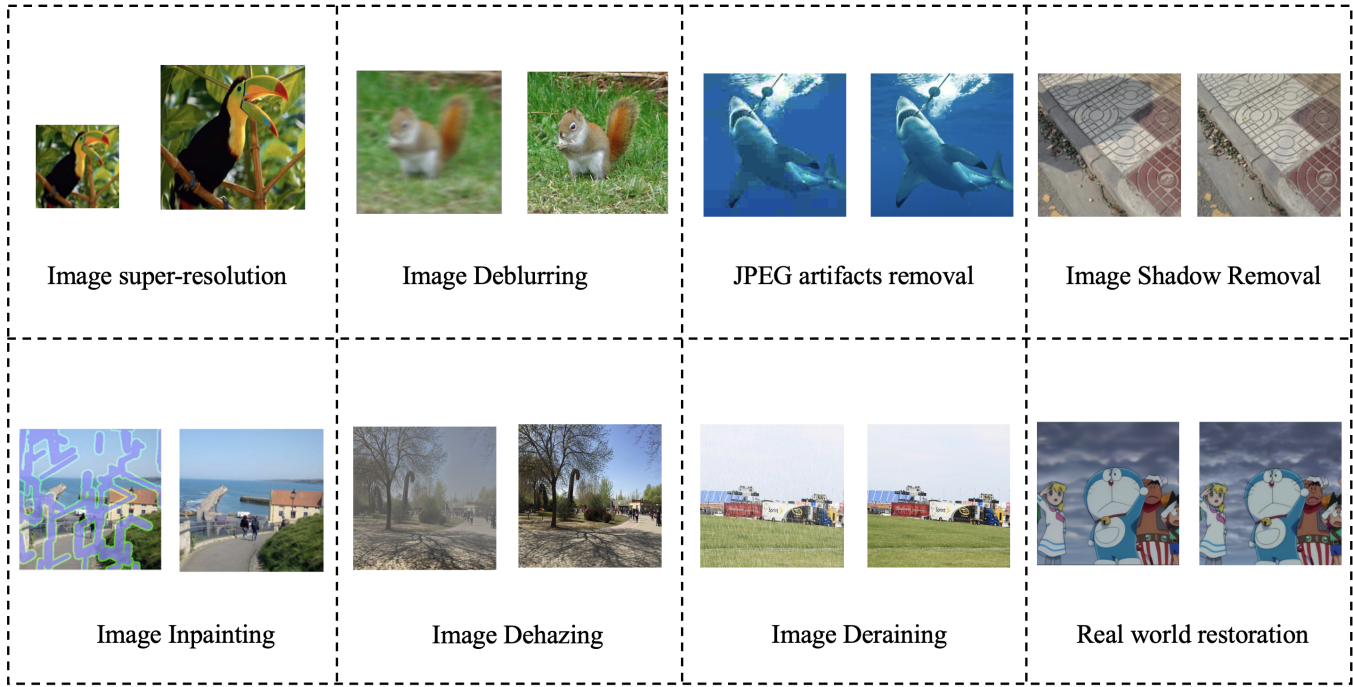
图像复原 (IR) 一直是低级视觉任务中的长期研究课题，在提升图像主观质量方面发挥着不可替代的作用。流行的 IR 任务包括图像超分辨率 (SR) [1–10]、去模糊 [11–17]、去噪 [18–25]、修复 [26–31] 和压缩伪影消除 [32–38] 等。图 1 为部分 IR 任务的直观说明。为了恢复失真的图像，传统的 IR 方法将恢复视为信号处理，并从空间或频率角度使用手工算法减少伪影 [18, 39–44]。随着深度学习的发展，许多 IR 作品收集了一系列针对各种 IR 任务定制的数据集，如用于 SR 的 *e.g.*、DIV2K [45]、Set5 [46] 和 Set14 [47]，用于 SR 的 Rain800 [48]、Rain200 [?]、Raindrop [49] 和 DID-MDN [50] 用于排水，REDS [51] 和 Gopro [52] 用于运动去模糊等。利用这些数据集，最近的大多数工作 [1–3, 7–11, 13, 16, 19, 21–23, 32–34, 53–55] 致力于通过基于卷积神经网络 (CNN) [56] 或 Transformer [57] 精心设计的主干网络来提高 IR 网络对复杂退化的表示能力。虽然这些工作在客观质量 (*e.g.*、PSNR 和 SSIM) 方面取得了突出的进步，但恢复的图像仍然存在纹理生成不令人满意的问题，阻碍了 IR 方法在实际场景中的应用。

得益于生成模型 [58–66] 的发展，尤其是生成对抗网络 (GAN) [64]，一些开创性的 IR 研究 [5, 6, 67–70] 指出，以前的逐像素损失、*e.g.*、MSE 损失和 L1 损失易受模糊纹理的影响，并将 GAN 的对抗性损失引入到 IR 网络的优化中，从而增强其纹理生成能力。例如，SRGAN [5] 和 DeblurGAN [12] 分别利用逐像素损失和对抗性损失的组合来实现面向感知的 SR 网络和去模糊网络。在此之后，改进基于 GAN 的 IR 的两个主要方向来自增强生成器 (*i.e.*，恢复网络) [5, 6, 71–73] 和鉴别器 [74–77]。尤其是 ESRGAN [6] 引入了强大的 RRDB [6] 作为基于 GAN 的 SR 任务的生成器。三种流行的判别器，包括逐像素判别器 (U-Net 形状) [74]、逐块判别器 [75, 78–80] 和逐图像判别器 [76, 77] (*i.e.*，类似 VGG 的架构)，旨在关注不同粒度级别的主观质量 (*i.e.*，从局部到全局)。尽管上述工作取得了进展，但大多数基于 GAN 的 IR 研究仍然面临两个不可避免但至关重要的问题：1) 基于 GAN 的 IR 的训练容易受到模式损坏和不稳定优化的影响；2) 大多数生成图像的纹理看起来是假的和反事实的。

近年来，扩散模型作为生成模型的一个新分支而出现，并在视觉生成任务中取得了一系列突破。扩散模型的原型可以追溯到[81]的工作，并由DDPM[82]、NCSN[83]和SDE[84]发展起来。一般来说，扩散模型由正向/扩散过程和逆向过程组成，其中

Xin Li and Yulin Ren contributed equally to this work. (Corresponding authors: Zhibo Chen and Xinchao Wang.)

Xin Li is with the University of Science and Technology of China (USTC) and the National University of Singapore (NUS). Yulin Ren, Xingrui Wang and Zhibo Chen are with the USTC. Xinchao Wang is with NUS. Xin Jin and Wenjun Zeng are with Eastern Institute for Advanced Study (EIT). Cuiling Lan is with Microsoft Research Asia (MSRA). (e-mail: lixin666@mail.ustc.edu.cn)



如图1：不同图像恢复任务的示例（左图和右图分别是扭曲的图像及其对应的干净的原始图像）。

前向过程逐步增加图像的像素噪声，直到满足高斯噪声；反向过程旨在通过分数估计[83]或噪声预测[82]去噪来重建图像。与GAN相比，扩散模型可以产生高保真和多样化的生成结果，从而在视觉生成[82–86]和条件视觉生成[86–97]等一系列领域成功取代GAN。随着视觉语言模型的进步，扩散模型已经扩展到跨模态生成，例如StableDiffusion [98]和DALLE-2 [99]。这极大地促进了人工智能生成内容（AIGC）的发展。我们在图2中按照时间轴列出了扩散模型的代表作品。

受扩散模型卓越生成能力的启发，大量研究探索了其在图像恢复任务中的应用，旨在促进纹理恢复。根据训练策略，这些工作大致可分为两类：1）第一类[100–109]致力于通过监督学习从头开始优化IR的扩散模型；2）第二类（*i.e.*，零样本模型）[110–117]致力于利用预先训练的IR扩散模型中的生成先验。通常，基于监督学习的方法需要收集大规模的失真/干净图像对，而基于零样本的方法主要依赖于已知的退化模式。这些限制阻碍了这些基于扩散模型的方法在现实世界场景中的应用，因为其中的失真通常是多种多样且未知的。为了进一步解决上述问题，一些研究[118–123]通过结合真实世界的失真模拟、核估计、域

翻译和扭曲不变学习。

尽管扩散模型在图像恢复方面表现出显著的效果，但相关技术和基准表现出相当大的多样性和复杂性，使得它们难以遵循和改进。此外，缺乏对基于扩散模型的IR的全面评论，进一步限制了其发展。在本文中，我们首次回顾和总结了基于扩散模型的图像恢复方法的工作，旨在提供一个结构良好且深入的知识库，并促进其在图像恢复社区中的发展。

在本综述中，我们首先在第2部分介绍扩散模型的背景，重点介绍三种基础建模方法 *i.e.*、NCSN [83]、DDPM [82] 和 SDE [84]，以及从优化策略、采样效率、模型架构和条件策略的角度对扩散模型的进一步改进。基于这些准备工作，我们在第3部分从两个不同的方向阐明了扩散模型在图像恢复中的进展：1）基于监督扩散模型的IR，和2）基于零样本扩散模型的IR。在第4部分中，我们总结了在更实际和更具挑战性的场景 *i.e.*、盲/现实世界退化下的基于扩散模型的IR。这旨在进一步增强基于扩散模型的IR方法满足实际应用需求的能力。为了便于进行合理和详尽的比较，在第4部分中，我们总结了基于扩散模型的IR 5，我们阐明了不同基于扩散模型的IR任务中常用的数据集和实验设置。此外，还提供了不同任务之间的基准测试的全面比较。在第6节中，我们深入分析了基于扩散模型的IR的主要挑战和潜在方向。最后

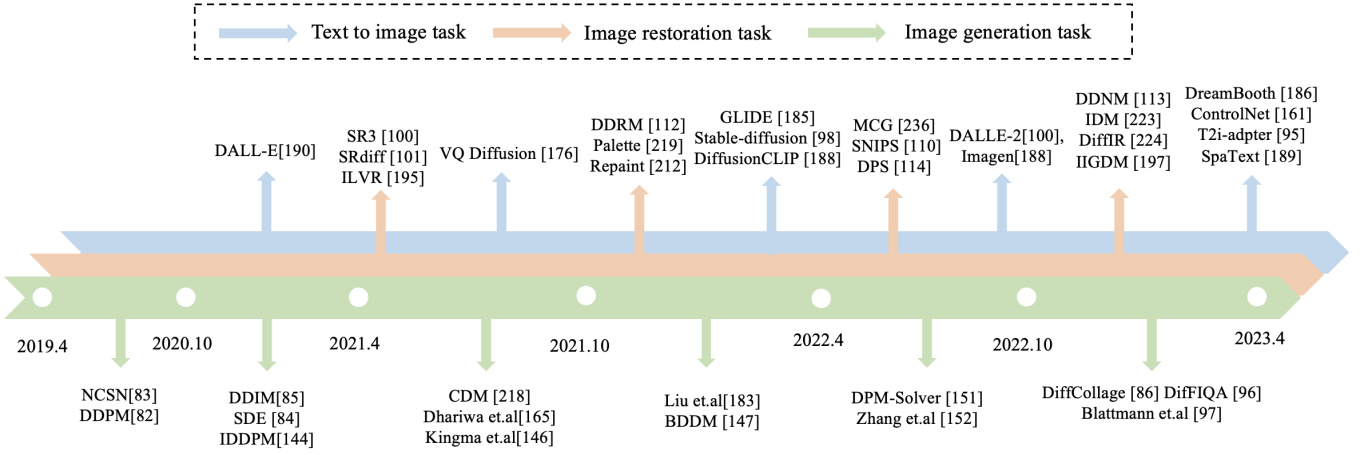


图2：不同应用的扩散模型的代表性作品（绿色表示图像生成任务，橙色表示图像恢复任务，蓝色表示文本转图像任务）

本次审查的结论总结在第 7 节中。

2 扩散模型（DM）的背景

扩散概率模型（*i.e.*，扩散模型）带来了生成模型领域的一场革命，它通过马尔可夫链建模将复杂不稳定的生成过程转化为几个独立稳定的逆过程。有三种基础扩散模型被广泛应用，包括DDPM[82]、NCSNs[83]和SDE[84]。其中，NCSNs[83]试图通过从一系列递减的噪声尺度序列中采样，利用退火的朗之万动力学来建模数据分布。相反，DDPM[82]通过添加高斯噪声的固定过程来建模正向过程，这将扩散模型的逆过程简化为变分界限目标的求解过程。这两个基本扩散模型实际是基于分数的生成模型[84]的特例。SDE[84]作为统一形式，用随机微分方程（SDE）对连续扩散和逆向扩散进行建模。证明了NCSN和DDPM只是SDE的两个独立离散化。我们将在后续章节中阐明这三个基本扩散模型的建模策略。

2.1 噪声条件分数网络（NCSN）

生成模型的目的是学习目标数据的概率分布。与以前基于似然法的[124 – 129]和基于GAN的[130 – 138]方法不同，NCSN旨在从对数密度函数（*i.e.*，得分函数 $\nabla \log p(x)$ ）的梯度估计数据分布，从而引导采样逐步向数据分布中心的方向进行。具体而言，NCSN使用由 θ 参数化的神经网络 s_θ 预测原始数据的得分函数。为了避免得到的分布塌缩到低维流形以及低密度区域的得分估计不准确，针对基于得分的生成模型设计了退火朗之万动力学[139, 140]，其中预定义的噪声具有单调递减的水平

$\sigma_{i=1}^L$ 来扰动数据。原始的朗之万动力学采样过程可以表示为：

$$\tilde{x}_t = \tilde{x}_{t-1} + \frac{\epsilon}{2} \nabla_{\tilde{x}} \log p(\tilde{x}_{t-1}) + \sqrt{\epsilon} z_t, \quad (1)$$

其中 z_t 是时间步长 t 处的随机正态高斯噪声， ϵ 是固定步长。当时间步长 $T \rightarrow \infty$ 和 $\epsilon \rightarrow 0$ 时，分布 $p(\tilde{x}_T)$ 将等于原始数据分布 $p(x)$ 。在添加级别为 σ 的噪声后，扰动分布将为 $q_\sigma(\tilde{x}) \triangleq \int p(x) \mathcal{N}(\tilde{x}|x, \sigma^2 I) dx$ 。噪声条件得分网络（NCSN）可以针对 $s_\theta(\tilde{x}, \sigma) = -\nabla_{\tilde{x}} \log q_\sigma(\tilde{x})$ 进行优化，噪声得分匹配目标为：

$$\mathcal{L}(\theta, \sigma) = \frac{1}{2} E_{p(x)} E_{\tilde{x} \sim \mathcal{N}(x, \sigma^2 I)} [\|s_\theta(\tilde{x}, \sigma) + \frac{\tilde{x} - x}{\sigma^2}\|_2^2] \quad (2)$$

2.2 去噪扩散概率模型

DDPM（去噪扩散概率模型）[82] 源自扩散模型 [141]，它通过将方差 β_t 设置为固定值，为扩散模型引入了简单的变分边界目标。扩散模型中有两个关键过程，*i.e.*，正向过程和反向过程。具体来说，正向过程（*i.e.*，DDPM 中的扩散过程）旨在逐步将训练数据破坏为高斯噪声，这是一个参数化的马尔可夫链，如下所示：

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} \cdot x_{t-1}, \beta_t \mathbf{I}), \quad (3)$$

其中 $x_0, x_1, \dots, x_T$ 是噪声隐变量，通过逐步将噪声添加到训练数据点 $x_0 \sim p_{data}(x)$ 中，噪声计划为 $\beta_1, \dots, \beta_T \in (0, 1)$ ，步骤为 T 。我们可以计算给定 x_0 的 x_t 的概率分布，如下所示：

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\hat{\alpha}_t} x_0, \sqrt{1 - \hat{\alpha}_t} \mathbf{I}), \quad (4)$$

其中 $\alpha_t = 1 - \beta_t$ 和 $\hat{\alpha}_t = \prod_{i=1}^t \alpha_i$ 。当时间步长 $t \rightarrow T$ 足够大时，由于 $\hat{\alpha}_t \rightarrow 0$ ，因此 $\hat{\alpha}_t$ 的分布将为标准高斯分布 $\pi(x_T) \sim \mathcal{N}(0, \mathbf{I})$ 。

扩散模型的逆过程旨在通过近似后验分布 $q(x_{t-1}|x_t, x_0)$ 从高斯噪声中恢复数据分布：

$$q(x_{t-1}|x_t, x_0) = \mathcal{N}(x_{t-1}; \tilde{\mu}_t(x_t, x_0), \tilde{\beta}_t \mathbf{I}), \quad (5)$$

其中, $\tilde{\mu}_t(x_t, x_0) = \frac{\sqrt{\hat{\alpha}_{t-1}\beta_t}}{1-\hat{\alpha}_t}x_0 + \frac{\sqrt{\hat{\alpha}_t(1-\hat{\alpha}_{t-1})}}{1-\hat{\alpha}_t}x_t = \frac{1}{\sqrt{\hat{\alpha}_t}}(x_t - \frac{\beta_t}{\sqrt{1-\hat{\alpha}_t}})\epsilon$ 对应于 $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 和 $\tilde{\beta}_t = \frac{1-\hat{\alpha}_{t-1}}{1-\hat{\alpha}_t}$ 。如公式 5 所示, 方差表 β_t 是预定义的, 因此, 只需要通过去噪网络 $\epsilon_\theta(x_t, t)$ 近似平均值 $\mu_\theta(x_t, t) = \tilde{\mu}_t(x_t, x_0)$ 。去噪网络的优化目标 [82] 可以写成：

$$\mathcal{L}_{simple} = \mathbb{E}_{t, x_0, \epsilon} [\|\epsilon - \epsilon_\theta(\sqrt{\hat{\alpha}_t}x_0 + \epsilon\sqrt{1-\hat{\alpha}_t}, t)\|_2^2] \quad (6)$$

去噪扩散概率模型 (DDPM) 的直观图示如图3所示

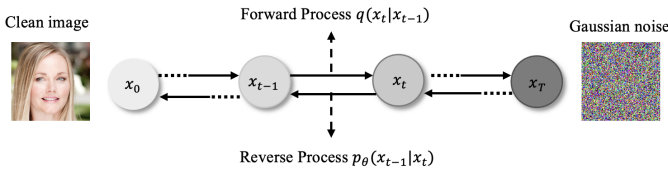


图 3：去噪扩散概率模型。

2.3 随机微分方程 (SDE)

为了将这些方法与基于分数的生成模型和扩散概率模型统一起来, SDE [84] 通过随机可微分方程 (SDE) 利用连续扩散过程, 如下所示：

$$dx = \mathbf{f}(x, t)dt + g(t)d\mathbf{w} \quad (7)$$

其中, $\mathbf{w}$ 为标准维纳过程, $\mathbf{f}(\cdot, t)$ 称为 $x(t)$ 的漂移系数, $g(\cdot)$ 称为 $x(t)$ 的扩散系数。这里, 扩散系数可以理解受随机噪声扰动的程度, 漂移系数可以设计为保证高斯分布, 例如 DDPM [82] 和 NCSN [83]。上述连续扩散过程的逆过程 (*i.e.*, 从噪声中采样数据) 也是扩散过程, 可以用逆时间 SDE 建模：

$$d\mathbf{x} = [\mathbf{f}(x, t) - g(t)^2 \nabla_x \log p_t(x)]dt + g(t)d\hat{\mathbf{w}} \quad (8)$$

这里 dt 是无穷小负时间步长, $\hat{\mathbf{w}}$ 是时间从 T 逆流到 0 时的标准维纳过程。逆时间 SDE 的核心是利用神经网络估计得分函数, 然后通过得分匹配求解式 8 [142, 143]。

DDPM 与 NCSN 可以看作是两个不同的 SDE 的离散化, 当时间变量趋于无穷大时, DDPM 的前向过程将收敛为：

$$d\mathbf{x} = -\frac{1}{2}(1 - \hat{\alpha}_t)\mathbf{x}dt + \sqrt{1 - \hat{\alpha}_t}d\mathbf{w}. \quad (9)$$

NCSN 的 SDE 形式如下：

$$d\mathbf{x} = \sqrt{\frac{d[\sigma^2(t)]}{dt}}d\mathbf{w}. \quad (10)$$

这里, 公式 9 和公式 10 分别称为方差保持 (VP) SDE 和方差爆炸 (VE) SDE。

2.4 扩散模型的改进

基于上述基础扩散模型, 人们从优化策略、采样效率、模型架构、条件策略等角度开展了大量工作来进一步增强扩散模型。

优化策略。为了增强扩散模型的稳定性和性能, 一些研究 [144–148] 探索了正向和反向过程中方差/噪声时间表的优化。值得注意的是, 正向过程中的噪声时间表控制每一步的扰动程度, 这对于逆向过程尤为重要。作为代表性研究, DDPM [82] 对扩散过程采用了简单的线性噪声时间表。但这种方法往往会导致次优结果, 尤其是对于低分辨率图像生成 [144]。为了缓解这种情况, IDDPM [144] 引入了余弦噪声时间表, 以消除早期扰动阶段快速噪声积累的负面影响。同时, Diederik *et al.* [146] 使用单调神经网络参数化噪声时间表, 并与扩散模型联合优化它。一般来说, 逆向过程中的方差时间表是固定的, 并与正向过程中的噪声时间表一起计算。然而, IDDPM [144] 发现学习方差可以进一步提高对数似然性, 因此在正向和反向过程中都采用方差表作为方差的可学习线性插值。相比之下, Analytic-DPM [149] 通过解析估计得出最优方差轨迹, 从而提高各种 DPM 的对数似然性。与上述方法不同, Jolicœur-Martineau *et al.* [150] 提出了一种全新的采样程序 *i.e.*, 即一致退火采样, 对于扩散模型, 它比退火朗之万方法更稳定。

采样效率。扩散模型的生成质量在很大程度上取决于大量的采样步骤, 因此对其在实际应用中的效率提出了挑战。为了缓解这个问题, 已经提出了四条主要的工作路线来加速采样过程。1) 第一条路线涉及与 ODE 相关的手工采样策略 [85, 151–154]。例如, DDIM [85] 在前向过程中引入了非马尔可夫链, 使扩散模型能够实现任意步长的采样。相比之下, DPM 求解器 [151] 通过分析计算 ODE 解的线性部分而不是使用黑箱 ODE 求解器来寻求快速的 ODE 求解器。因此, 该方法大大缩小了生成高质量图像所需的采样步骤, 将它们限制在 10 到 20 的可接受范围内。2) 第二个范例是修改扩散过程 [155, 156]。作为代表作品, Lou *et al.* [155] 提出使用早期停止机制截断扩散过程, 并从非高斯分布中启动采样, 该分布由预先训练的 VAE/GAN 模型生成。3) 在第三个

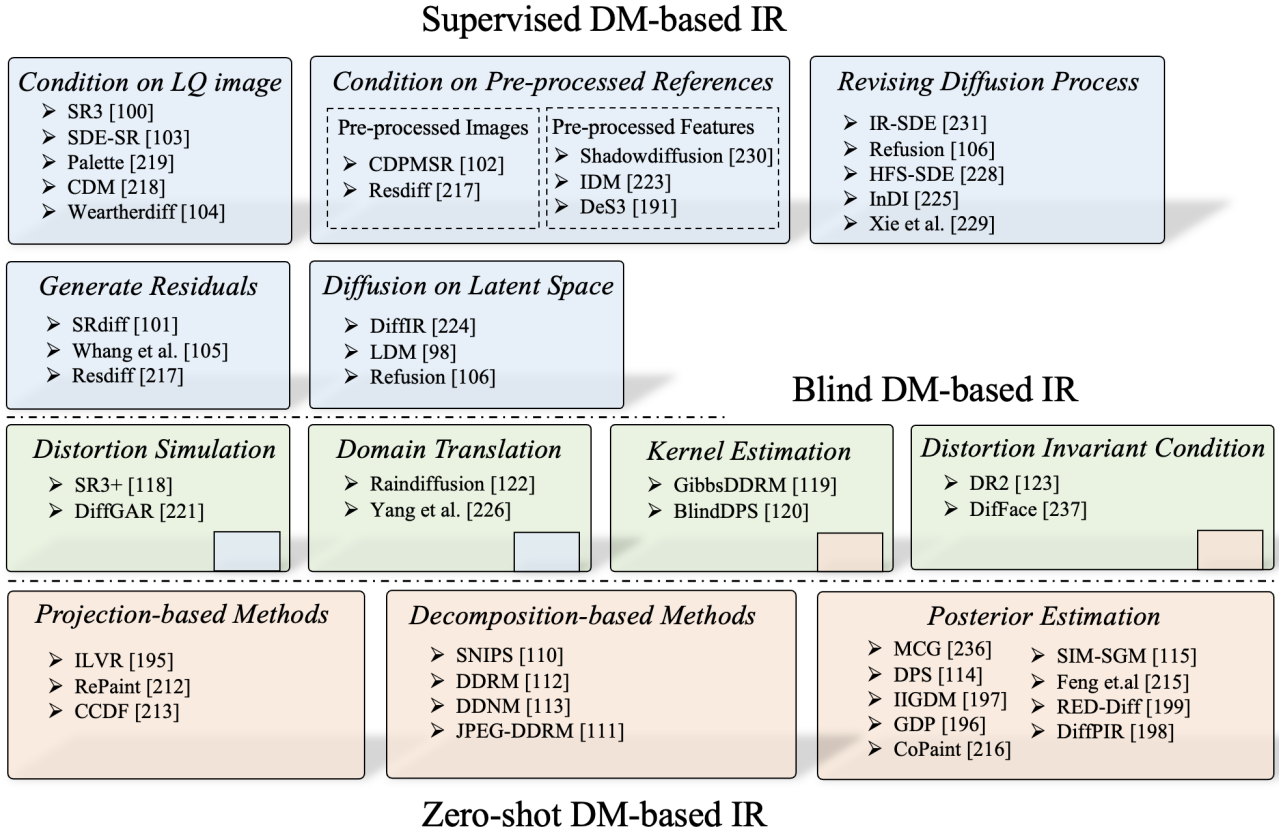


图 4：基于扩散模型的图像恢复模型概览。该图根据训练方法将扩散模型分为两类，即基于监督的模型（用蓝色背景表示）和基于零样本的模型（用橙色背景表示）。此外，该图根据条件的结合方式对这两类模型进行了更详细的分类。具有绿色背景背景的模型专门用于真实图像恢复。

策略，采用知识蒸馏将生成能力从多个采样步骤转移到几个采样步骤[157–160]。4) 最后一种方法利用条件策略[98, 161–164]来嵌入生成先验，从而优化采样效率。

模型架构。扩散模型主要采用两种架构，*i.e.*、基于 CNN 的 U-Net 和基于 Transformer 的模型。值得注意的是，U-Net 架构在早期研究中更适合用于噪声/分数预测 [82, 83]，这得益于其分辨率保持能力以及通过多粒度下采样特征空间消除资源成本。在此之后，后续工作中进行了一系列努力来改进 U-Net 架构，包括结合交叉注意模块 [98]、组规范化 [82, 165]、多头注意 [84, 144, 145] 和位置编码 [82]。最近，Transformer 已经证明了其在建模长距离依赖关系和统一不同模态方面的能力 [166–173]。因此，一些文本转图像的研究 [174–181] 探索使用 Transformer 的主干网络，如 ViT [182]、Swinv2 [179]，来取代原始的基于 CNN 的 U-Net，以预测逆过程中的噪声，其中时间步长 t 和其他条件通过自适应层归一化 [175、177、178] 或交叉注意 [174] 输入到 Transformer。

条件策略。一种有效的条件策略对于扩散模型在条件生成中的功能至关重要。这促使了大量的工作去探索有效且有效的条件机制。Nicol *et al.* [165] 创新地训练了一个辅助分类器来指导扩散模型，利用其梯度将图像生成引导至特定语义。另一种流行的方法 [183, 184] 以无分类器的方式将条件引入到分数估计/噪声预测模型中，例如 GLIDE [185] 和 DALLE-2 [99]。为了保留其无条件生成能力，利用空标记 $\emptyset$ 替换扩散模型中的条件 $\theta(x, c)$ ，从而启用 $\theta(x) = \theta(x, \emptyset)$ ，其中 c 表示条件，例如文本特征。由于其优势，一系列工作将无分类器条件引入到文本到图像任务中 [185–190]。此外，除了类别标签和文本提示之外，扩散模型还可以整合其他模态条件，*e.g.*、图像、分割图、潜在特征，这极大地方便了其在各种需求中的应用，如稳定扩散[98]、ControlNet[161]。

3 基于扩散模型的图像复原方法

根据扩散模型 (DM) 是否无需训练即可用于 IR，我们可以初步将基于 DM 的 IR 方法分为两类，*i.e.*、基于监督 DM 的方法 [100、105、107、108、121、191–194] 和基于零样本 DM 的方法 [112、114、115、195–200]。具体而言，基于监督 DM 的 IR 方法需要使用 IR 数据集的成对扭曲/干净图像从头开始训练扩散模型。与之前基于 GAN 的方法 [201–209] 直接将扭曲图像作为输入不同，基于 DM 的 IR 采用精心设计的条件机制，在逆过程中将扭曲图像作为指导。尽管其纹理生成结果很有希望，但也存在两个显著的局限性：1) 从头开始训练扩散模型依赖于大量成对的训练数据。2) 在现实世界中收集成对的扭曲/干净图像具有挑战性。相比之下，基于零样本 DM 的方法提供了一种有吸引力的替代方案，因为它只需要扭曲图像，无需重新训练扩散模型。它不是从 IR 的训练数据集中获取恢复能力，而是从预先训练的扩散模型中挖掘和利用结构和纹理先验来恢复图像。其核心思想源于这样一种直觉，即将预训练的生成模型看作是结构和纹理存储库，使用大量现实世界数据集构建，例如 ImageNet [210] 和 FFHQ [211]。因此，基于零样本 DM 的 IR 方法面临的一个基本挑战是：*how*

to extract the corresponding perceptual priors while preserving the data structure from distorted images. 在随后的小节中，我们首先简要回顾一下具有代表性的监督 DM 的 IR 方法：SR3 [100] 和基于零样本 DM 的 IR 方法：ILVR [195]。然后我们从条件策略、扩散建模和框架的角度对这两类方法进一步分类，分别总结在表 1 和表 2 中。此外，扩散模型的总体分类如图 4 所示。

3.1 SR3 – IR 的代表性监督 DM

与从噪声中合成图像的纯图像生成任务不同，图像恢复旨在从相应的退化/低质量图像中生成高质量图像。因此，基于监督 DM 的 IR 的关键挑战在于

how to effectively incorporate the degraded/low-quality image into the diffusion model as the condition. 我们将退化图像表示为 y 。IR 的扩散模型 (DM) 的基本目标是学习时间步骤 t 的后验分布 $p_\theta(x_{t-1}|y, x_t)$ ，使得 $x_0 \sim q(x|y)$ 和 x 表示相应的高质量图像。为了实现这一目标，引入了一种具有直接条件策略的开创性的基于 DM 的监督方法 SR3。具体而言，它直接将退化图像与时间步骤 t 的生成图像 x_t 连接起来，从而有效地实现了 SR 的条件图像生成。

如图 5 所示，SR3 遵循典型的 DDPM [82] 框架，并利用 U-Net 模型作为噪声预

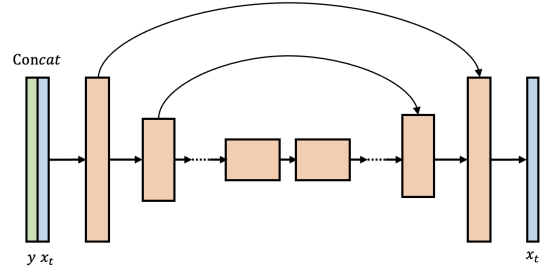


图 5：SR3 网络的主干

dictor。给定低分辨率 (LR) 图像 y ，SR3 首先使用双三次插值将其上采样到所需分辨率。随后，它将超分辨率 LR 图像 y 与 t 时间步骤的去噪输出 x_t 连接起来，作为扩散模型的输入，以预测 $t-1$ 步骤的噪声。当达到 $t=0$ 时，扩散模型可以提供 y 的上采样高质量图像 x_0 作为 $x_0 \approx x$ 。

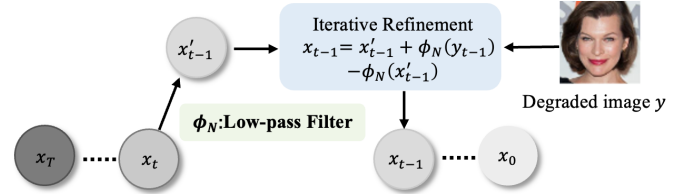


图 6：ILVR 网络的主干

3.2 ILVR – IR 的代表性零样本 DM

尽管基于监督 DM 的 IR 方法表现出了卓越的性能，但训练过程需要大量的计算成本和大规模配对数据集，这可能会让一些研究人员望而却步。为了解决这个问题，提出了基于零样本 DM 的 IR [110, 195, 212–214]，以利用预训练扩散模型中的内在知识。具体而言，据观察，使用大量自然图像训练的用于图像生成的预训练扩散模型包含了大量关于丰富纹理的先验知识。因此，这些预训练的扩散模型可以被视为纹理信息的存储库。探索重新利用此类先验知识进行免训练图像恢复是低级视觉领域一个新兴且有前途的方向。

作为最初的工作，Choi *et al.* [195] 引入了迭代潜变量细化 (*i.e.*，ILVR) 方法，该方法利用无条件扩散模型实现图像 SR 和图像翻译的无训练条件生成。ILVR 的关键创新在于用参考图像中的低频分量替换去噪输出中的低频分量。如图 6 所示，此替换过程确保了生成图像和参考图像之间的结构和语义一致性，从而促进了条件生成。具体而言，给定参考图像 y (*e.g.*，失真图像

在 IR 中)，在时间步骤 t ，ILVR 使用以下公式预测时间 $t-1$ 的去噪结果：

$$\mathbf{x}'_{t-1} = \sigma_t \mathbf{z} + \frac{1}{\sqrt{\alpha_t}} (\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \alpha_t}} \epsilon_\theta(\mathbf{x}_t, \mathbf{t})) \quad (11)$$

其中 $\epsilon_\theta(x_t, t)$ 表示降噪网络预测的噪声， $\mathbf{z}$ 表示控制生成图像随机性的标准随机高斯噪声。然而，这种采样过程不可避免地会在 x_t 中产生不一致的结构/纹理，这需要通过低频替换进行细化以与参考图 y 中的结构/纹理对齐：

$$x_{t-1} = x'_{t-1} + \Phi_N(y_{t-1}) - \Phi_N(x'_{t-1}). \quad (12)$$

这里， Φ_N 表示用于绕过输入中的低频分量的低通滤波器， $y_{t-1} \sim q(y_{t-1}|y)$ 是 y 在 $t-1$ 步的扩散状态。遵循 ILVR，大多数基于零样本 DM 的 IR 方法 [112–114, 197, 200, 215, 216] 主要侧重于增强采样过程中的细化策略，因此无需训练。

3.3 基于监督 DM 的 IR。

受 SR3 [100] 的启发，许多研究致力于优化基于监督 DM 的 IR 框架，重点是增强条件策略和探索潜在且更高效的生成空间。关于条件策略，我们根据条件将这些研究分为三类：1) 低质量参考图像，2) 预处理参考，3) 修改扩散过程。关于生成空间，基于监督 DM 的 IR 方法可分为三个不同的组：图像空间、残差空间和潜在空间。如果没有提到，大多数研究都在图像空间内生成恢复的图像，其中需要直接生成结构和纹理。相比之下，残差空间扩散模型侧重于重建低质量图像与其对应的高质量图像之间的残差，这简化了生成整个图像的复杂性。基于隐含层的方法利用精心设计的编码器将图像转换为紧凑的潜在空间进行生成，从而提高生成效率。本节将从上述三种条件策略和最后两种生成空间的角度阐述现有的基于监督 DM 的 IR 研究。

3.3.1 Condition with Low-quality Reference Image

如第 3.1 节所强调的，将失真图像作为条件对于基于监督 DM 的 IR 来说是必不可少的，也是至关重要的。SR3 方法表明，通过简单的连接操作就可以实现显著的性能。它利用低质量参考图像与 $t-1$ 步骤的去噪结果的直接连接作为 t 步骤噪声预测的条件。利用相同的条件策略，Saharal *et al.* [219] 提出了一个统一的扩散模型，称为 Palette，用于图像到图像的转换任务，该模型在图像着色、修复、反裁剪和 JPEG 伪影去除方面取得了优异的性能。此外，他们研究了不同的优化目标对样本多样性和

强调了 U-Net 中自注意力机制对于扩散模型的关键作用。尽管上述方法非常有效，但它们也存在局限性，因为一旦训练完成，它们仅支持固定分辨率的 IR。为了使扩散模型适应任意大小的真实世界 IR，Özdenizci *et al.* [104] 将退化图像和相应的采样结果 x_t 分成几个重叠的块，然后利用逐块连接作为扩散模型的输入进行噪声预测。此外，为了解决重叠区域中不同采样块造成的不一致问题，本研究引入了重叠区域内每个像素的平均估计噪声。为了提高生成图像的质量，Ho *et al.* [218] 引入了基于 SR3 主干的三层级联扩散模型。第一个扩散模型用于实现类条件低分辨率图像生成，另外两个扩散模型级联用于对低分辨率生成图像进行超分辨率处理，从而生成更高分辨率、更逼真的图像。除了基于 DDPM 的方法外，还有另一项研究 [103] 探索了连续扩散模型 SDE [84] 的各种变体，使用预测校正采样实现人脸超分辨率。

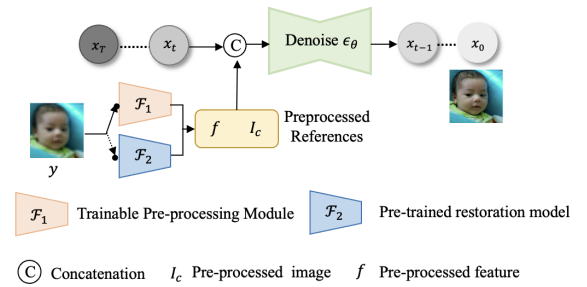


图 7：预处理参考的架构，其中低质量图像首先由预训练的恢复网络或可训练的预处理模块处理。输出参考可以是特征或干净的图像

3.3.2 Condition with Pre-processed References

尽管直接与低质量图像连接表现出良好的性能，但低质量图像中的伪影不可避免地会对扩散模型的生成造成有害影响，尤其是对于严重且多样的失真。为了缓解这个问题，一些研究努力通过使用联合训练模块或预训练恢复网络对低质量图像进行预处理来增强条件。如图 7 所示，这些工作可以根据预处理策略分为两类，*i.e.*、具有预处理参考图像和特征的条件。

预处理后的参考图像。为了减轻低质量图像中伪影的副作用，CDPMSR [102] 利用现有的超分辨率模型 *e.g.*、RCAN [232]、SwinIR [53]、EDSR [233] 来增强低质量图像，从而为扩散模型提供高质量和更可靠的条件。此外，它避免在逆过程中进行随机采样，而采用确定性去噪过程，从而产生卓越的图像质量和更快的推理。值得注意的是，预处理

表 1：基于监督扩散模型的图像恢复模型

Models(Papers)	Conference&Year	Methods&Insights	Target Tasks
SR3 [100]	TPAMI 2022	Channel concatenation in DDPM with LR input	SR
SRdiff [101]	Neurocomputing 2022	DDPM with LR encoder information	SR
CDPMSR [102]	Arxiv2023	DDPM with Pretrained SR model	SR
Res-Diff [217]	Arxiv2023	DDPM with CNN model	SR
SDE-SR [103]	SIBGRAPI 2022	SDE with concat LR input	SR
LDM [98]	CVPR 2022	DDPM in latent space , inpainting	SR, Inpainting
CDM [218]	J.Mach.Learn.Res.2022	Cascaded DDPM with condition augmentation	SR
WeatherDiff [104]	TPAMI 2023	DDPM with patch-based sampling for size-agnostic tasks	Deraining, Desnowing, Dehazing
DeblurDPM [105]	CVPR 2022	Simple DDPM with predict and refine strategy	Deblurring
Palette [219]	SIGGRAPH 2022	DDPM with Concatenation of input, achieving multiple I2I tasks	Colorization, Inpainting, Uncropping, JPEG
SR3+ [118]	Arxiv2023	Improved SR3 architecture, data augmentation, real-world SR	SR
DriftRec [220]	Arxiv2022	SDE with adaptive strategy, Blind JPEG restoration	JPEG Artifacts Removal
DiffGAR [221]	Arxiv2022	DDPM with encoder in Latent space	Generative Artifacts Removal
DG-DPM [222]	Arxiv2022	DDPM with domain generalization for real-world Deblurring	Deblurring
IDM [223]	CVPR 2023	Neural representation for continuous SR	SR
DiffIR [224]	Arxiv2023	DDPM with transformer-structure .	SR, Deblurring
RainDiffusion [122]	Arxiv2023	Circular DDPM for real-world deraining	Real World Deraining
InDI [225]	TMLR 2023	Alternative methods proceed by iteratively restoring the input image of Diffusion model	Deblurring, SR, JPEG Restoration
Yang <i>et al.</i> [226]	Arxiv2022	DDPM for data pairs synthesis for real-world tasks training	Training Data Synthesis
Adadiff [227]	Arxiv2022	Accelerated DDPM with adaptive diffusion priors	MRI Reconstruction
HFS-SDE [228]	TMI 2022	SDE in high-frequency space	MRI Reconstruction
Xie <i>et al.</i> [229]	Arxiv2023	DDPM with three added noise distributions	Image Denoising
ShadowDiffusion [230]	CVPR 2023	DDPM with unrolling-inspired diffusive sampling strategy	Image Shadowing Removal
Refusion [106]	CVPRW 2023	Latent DDPM with compression strategies, large-scale restoration	Image Shadow Removal, Dehazing
IR-SDE [231]	ICML 2023	Mean-reverting SDE	Deraining, Dehazing, Denoising
DDS2M [109]	Arxiv2023	DDPM with Variational Spatio-Spectral Module	Hyperspectral Image Restoration
SUD ² [108]	Arxiv2023	DDPM with few training pairs	Inpainting, Dehazing
DeS3 [191]	Arxiv2022	Classifier-driven attention guided DDPM	Image Shadow Removal
DiffBFR [192]	Arxiv2023	DDPM with few training pairs	Blind Face Restoration
Stable-SR [121]	Arxiv2023	Finetune Text-to-image diffusion with time-awareness and feature wrapping module	Real World Super-resolution
PyDiff [107]	Arxiv2023	Pyramid diffusion with pyramid resolution style	Low-light Image Enhancement
DiffLL [193]	Arxiv2023	Wavelet-based diffusion model with high efficiency	Low-light Image Enhancement

参考图不仅可以作为增强条件，还可以作为初始恢复良好的图像。因此，ResDiff [217] 利用预训练的 CNN 生成低频内容丰富的图像作为初始恢复图像，并利用条件扩散模型进一步生成预处理后的扭曲图像与其对应的干净图像之间的残差。

预处理参考特征。另一种流行的做法是使用参考图的特征作为扩散模型的条件。IDM [223] 致力于实现连续图像超分辨率，首先使用 EDSR [233] 提取低分辨率图像的初始特征。然后，将初始特征下采样到多个尺度，这些尺度

用作扩散模型中不同上采样层的条件，目的是细化隐式表示。相比之下，ShadowDiffusion [230] 利用预先训练的 Transformer 主干从失真的参考图中提取退化先验（*i.e.*，退化相关特征）。提取的退化先验被用作辅助手段来细化生成的阴影掩模，并作为无阴影图像生成的条件。

3.3.3 Condition by Revising Diffusion Process

值得注意的是，上述基于监督 DM 的 IR 方法通过修改网络引入了条件

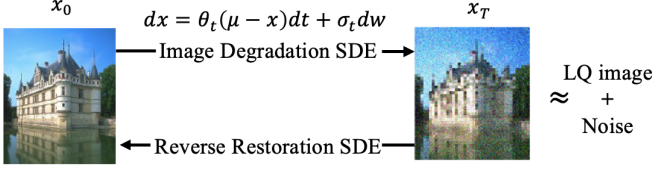


图 8：IR-SDE 的架构 [231]，其中前向扩散过程近似模拟了图像退化过程。前向过程的最终输出相当于一幅带有随机噪声的低质量图像

同时保留了 DDPM 中的扩散过程 [82]。然而，这要求生成过程从噪声开始。为了避免这个问题，一些研究通过修改扩散过程来调整扩散模型，使得扩散输出 x_T (i.e., 逆过程的起点) 近似于受少量高斯噪声破坏的低质量图像。如图 8 所示，Luo et al. [231] 使用均值回归 SDE 修改了正向过程，将 IR 的统一退化过程建模为：

$$dx = \theta_t(\mu - x)dt + \sigma_t dw, \quad (13)$$

其中 θ_t 和 σ_t 是与时间相关的参数。 μ 表示失真图像， x 表示其对应的高质量图像。利用均值回归 SDE，这项工作成功地分别通过修改的前向和反向过程对图像退化和恢复过程进行了建模。这避免了纯噪声的产生并实现了更好的恢复性能。在 IR-SDE [231] 的基础上，同一团队引入了 Refusion [106]，通过优化网络架构、噪声水平和去噪步骤等方面进一步细化了 IR-SDE [231]。为了降低计算成本，Refusion 引入了 U-Net 压缩策略，从而能够在潜在空间中实现高效采样。

出于同样的目的，谢等人 [229] 重新定义了扩散过程，使得采样从噪声图像开始。考虑到噪声的多样性，他们分别针对高斯、伽马和泊松噪声推导出三个独立的扩散过程。相比之下，InDI [225] 引入了连续扩散过程：

$$x_t = (1 - t)x + ty, \quad (14)$$

其中 x 和 y 分别为高质量图像及其对应的低质量图像。它可以解释为在时间步长 t 对高质量图像和低质量图像进行逐步插值，将原来监督图像复原的单步预测分解为几个小步骤，有效地避免了传统监督图像复原中经常出现的回归均值效应。与空域中的扩散过程不同，HFS-SDE [228] 在频率空间中重新表述了磁共振 (MR) 重建的扩散过程。在这种方法中，前向过程逐渐将噪声添加到高频空间，从而得到由高频噪声和低频数据组成的最终 x_T 。在反向过程中，HFS-SDE 采用预测-校正 (PC) 方法 [84] 进行采样。

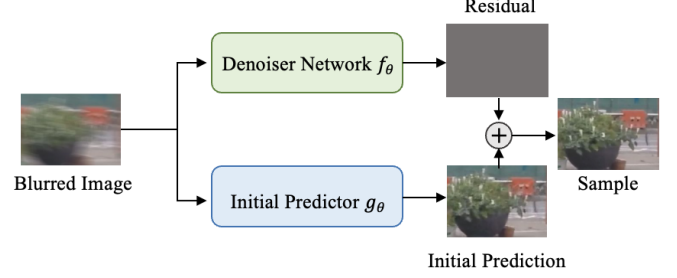


图 9：预测和优化模型的架构 [105]。预测器 g_θ 生成干净图像的初始预测，而残差信息则通过扩散过程建模

3.3.4 Generate Residuals

在基于监督 DM 的 IR 中，大多数研究直接从噪声中生成高质量图像，这需要同时生成结构和纹理。然而，重新生成低质量图像中已经存在的结构/纹理会不必要地增加扩散模型的负担并增加额外的资源成本。受此启发，一些代表性研究 [101, 105, 217, 234, 235] 试图将生成过程从图像空间转移到残差空间。目标是生成成对的高质量 and 低质量图像之间的残差。作为开创性的工作，SRDiff [101] 首次利用扩散模型来预测 SR 中的残差。相比之下，Whang [105] 为图像去模糊任务引入了一种预测和细化策略。如图 9 所示，本研究首先用确定性去模糊网络预测初始去模糊图像，然后通过随机扩散模型生成残差。上面提到的 ResDiff [217] 也对 SR 采用了这种策略。

3.3.5 Diffusion on Latent Space

为了减轻扩散模型的训练和采样成本，StableDiffusion [98] 是第一个在潜在空间中实现基于 DM 的生成的工作。具体来说，它预训练了一个自动编码模型 (i.e., 一种编码器-解码器架构) 来学习感知空间，这能够在降低计算复杂度的同时保持重建图像的感知质量。利用预训练的自动编码器，StableDiffusion 将图像级扩散过程转换为感知空间，然后使用交叉注意力机制将各种条件 (e.g., 文本、分割图和图像) 引入扩散模型。受此启发，Refusion [106] 引入了潜在扩散模型用于图像恢复，以加速训练和采样，如图 10 所示。与上述通过压缩原始图像获得潜在空间的工作相比，DiffIR [224] 利用潜在扩散模型生成紧凑的 IR 先验，从而指导基于动态变压器的恢复网络 (DIRformer) 实现更好的恢复。

3.4 基于零样本 DM 的 IR。

与有监督的基于 DM 的 IR 不同，基于零样本 DM 的 IR 致力于实现无训练和无数据的图像

表 2：基于零样本扩散模型的图像恢复模型

Models(Papers)	Conference&Year	Methods&Insights	Target Tasks
ILVR [195]	ICCV 2021	Adds projection into sampling	Image-translation, SR
Repaint [212]	CVPR 2022	Impose projection into sampling with resampling strategy.	Inpainting
CCDF [213]	CVPR 2022	Adds projection into sampling	Image-translation, SR, Inpainting, MRI reconstruction
SNIPS [110]	NeurIPS 2021	DDPM in Spectral Space with SVD decomposition	Deblurring, SR, Compressive sensing
DDRM [112]	ICLR 2022	DDPM in Spectral Space with SVD decomposition	Deblurring, SR
JPEG-DDRM [111]	NeurIPS 2022	JPEG restoration based on DDRM framework	JPEG Restoration
DDNM [113]	ICLR 2023	Pretrained DDPM with Null Space, Time travel strategy	SR, Inpainting, Deblurring, Colorization
MCG [236],	NeurIPS 2022	Pretrained DDPM with manifold constraints	Colorization, reconstruction CT
DPS [114]	CVPR2023	Pretrained DDPM with Posterior sampling	SR, Inpainting, Deblurring
BlindDPS [120]	ICLR 2023	Parallel diffusions for images and kernels, achieving blind Deblurring	Deblurring
GibbsDDRM [119]	Arxiv2023	Pretrained DDPM with joint posterior sampling for blind deblurring	Deblurring Vocal dereverberation
Dif-Face [237]	ICLR 2023	Pretrained DDPM with initialization reverse image for real-world task	Blind face Restoration
DR2 [123]	CVPR 2023	Pretrained DDPM degradation removal module with enhancement module	Real-world tasks
DiracDiffusion [200]	Arxiv2023	Incremental reconstruction with ensured data consistency	Deblurring, Inpainting
IIGDM [197]	ICLR 2023	Pseudoinverse-guided Diffusion Models, non-linear tasks	SR, Inpainting, JPEG Restoration
GDP [196]	CVPR 2023	DDPM with Generative diffusion priors	SR, Deblurring, Inpainting, Colorization
COPAINT [216]	Arxiv2023	DDPM with Bayes framework, Enhance Coherence	Inpainting
SIM-SGM [115]	ICLR 2022	Pretrained NCSN with matrix decomposition	MRI Reconstruction
ADIR [117]	Arxiv2022	Adaptive diffusion model with CLIP	SR, Deblurring, Text-Guided Editing
Feng <i>et al.</i> [215]	Arxiv2023	Posterior Sampling with Score-Based Priors	Denoising, Deblurring
Liu <i>et al.</i> [116]	Arxiv2023	A latent diffusion guider with a lightness-aware VQVAE model.	Colorization
DiffPIR [198]	CVPR 2023	Integrate plug-and-play method into diffusion	SR, Inpainting, Deblurring
RED-Diff [199]	Arxiv2023	Integrate regularization-by-denoising, Improve posterior score estimation	Inpainting, SR

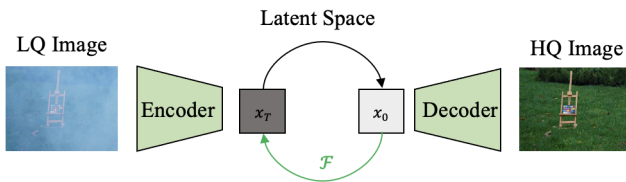


图 10：Refu-sion [106] 中使用的扩散模型概览，低质量图像首先通过编码器压缩为潜在信息，然后通过扩散过程对潜在信息进行建模。

恢复。它通常依赖于为生成任务设计的预训练扩散模型，并在采样过程中纳入低质量图像的情况。该任务的核心挑战来自 i) 如何保持低质量图像和恢复之间的数据一致性

生成的图像，因为预训练的扩散模型致力于保持数据分布而不是像素级数据一致性。ii) 如何挖掘与低质量图像一致的感知知识，这对条件的设计提出了更高的要求。在本文中，我们粗略地将基于零样本 DM 的 IR 方法总结为三类，*i.e.*，投影，分解和后验估计。

3.4.1 Projection-based Methods.

为了缓解零样本 DM IR 的主要挑战，一些研究引入了基于投影的方法 [195, 212, 213]。该方法旨在从低质量图像中提取固有结构/纹理，作为每一步生成图像的补充，从而确保数据一致性。例如，图像修复任务仅涉及生成掩码区域的内容。低质量图像的未掩码区域

可以在第 $t - 1$ 步替换去噪图像的相应部分，从而建立采样过程中数据一致性的条件。与此相符，RePaint [212] 利用一个简单的投影进行图像修复任务：

$$x_{t-1} = m \odot x_{t-1}^{known} + (1 - m) \odot x_{t-1}^{unknown}, \quad (15)$$

其中 $x_{t-1}^{known} \sim \mathcal{N}(\sqrt{\alpha_t}y, 1 - \alpha_t\mathbf{I})$ 是在时间步骤 $t - 1$ 处对蒙版图像 y 添加噪声得到的扩散结果。 $x_{t-1}^{unknown}$ 是从扩散模型的去噪预测中采样得到的。相比之下，ILVR [195] 采用低频投影进行图像超分辨率。理论上，在时间步骤 $t - 1$ 处，预测的潜在变量 x_{t-1} 和在扩散过程) 的第 $t - 1$ 步骤中对低分辨率图像 y 添加噪声的 y_{t-1} (i.e. 应该共享相同的低频分量。因此，它用来自 y_{t-1} 的低频分量替换 x_{t-1} 的低频分量，确保数据一致性并为扩散模型建立改进条件。作为一种先进的解决方案，CCDF [213] 引入了一种统一的投影方法：

$$x_{t-1} = Ax'_{t-1} + b, \quad (16)$$

其中 A 和 b 的设置是为了实现数据一致性。例如，在 SR 任务中，上述投影可以实例化为：

$$x_{t-1} = (\mathbf{I} - \mathbf{P})x'_{t-1} + y_{t-1}, \quad (17)$$

其中 $\mathbf{P}$ 是低分辨率图像的退化过程。此外，这项工作证明了从更好的初始化开始的生成可以提高逆向过程的速度。

3.4.2 Decomposition-based methods.

值得注意的是，大多数图像恢复问题可以看作是线性逆问题，可以表示为：

$$y = Hx + z, \quad (18)$$

其中， H 为线性退化算子， z 为污染噪声。在这种情况下，由于存在噪声 z ，因此无法直接估计条件概率 $p(x|y)$ 。为了消除噪声 z ，SNIPS [110] 和 DDRM [112] 在退化算子 H 上进行奇异值分解 (SVD)，并在谱域中运行扩散过程。具体而言，SNIPS [110] 基于退火朗之维动力学，并在谱空间上推导出条件得分函数，这在图像去模糊、超分辨率和压缩感知任务上取得了出色的性能。继 SNIPS 之后，DDRM [112] 进一步将 SVD 分解扩展到线性逆问题的变分目标，这表明预训练的 DDPM [82]/DDIM [85] 可以成为该问题的最优解。值得注意的是，上述工作仅关注线性逆问题。相比之下，Kawar 等人 [111] 研究了基于 DDRM 特例 (i.e. 逆问题中无噪声 z) 的非线性逆问题，并扩展了伪逆概念以实现 JPEG 伪影校正。对于 MRI 重建，SVD 分解并不合适。为了解决这个问题，Song *et al.* 从头开始训练医学图像的非条件生成模型

然后利用采样过程中的矩阵分解来解决线性逆问题，这对于未知测量过程具有普遍性。

出于不同的目的，DDNM [113] 引入了另一种分解策略，即范围零空间分解，以进一步改善零样本图像恢复，其中范围空间负责数据一致性，零空间用于改善真实性 (i.e. 感知质量)。给定无噪声逆 $y = Hx$ ，它可以分解为：

$$y = HH^\dagger Hx + H(I - H^\dagger H)x, \quad (19)$$

其中 $H^\dagger$ 是退化操作 H 的伪逆。我们可以看到范围空间 $HH^\dagger Hx = Hx = y$ 可以确保数据一致性，而零空间 $H(I - H^\dagger H)x$ 对数据一致性没有影响，因为 $H(I - H^\dagger H)x = 0$ 。如图 11 所示，基于此，DDNM [113] 将时间步 t 处 x_0 的预测校正为：

$\hat{x}_{0|t} = H^\dagger y + (I - H^\dagger H)x_{0|t}$ ，其中 $x_{0|t}$ 可以通过时间步 t 处的噪声预测来估计。利用校正后的 $\hat{x}_{0|t}$ ，我们可以计算时间步 $t - 1$ 处的去噪输出 x_{t-1} ，这确保了数据的一致性并为下一次噪声预测提供了更好的条件。此外，DDNM 还利用 SVD 来解决带噪声的线性逆问题，称为 DDNM+。

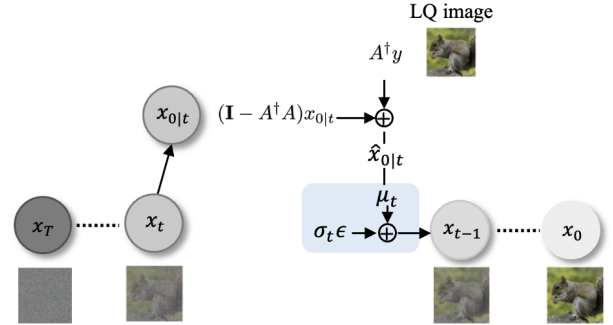


图 11：DDNM [113] 的体系结构，将前向退化算子 A 分解到范围-零空间以满足真实性和数据一致性。

3.4.3 Posterior estimation.

基于投影的方法在图像恢复的逆问题中表现出色，其中在扩散模型的反向采样步骤之后添加了基于投影的测量一致性校正。然而，大多数基于投影的工作都致力于无噪声逆问题，并且通常存在不令人满意的数据一致性问题，因为投影会将样本路径抛出数据流形[236]。为了解决一般的噪声线性逆问题，一些工作[114,196,197,200,216,236]旨在利用基于贝叶斯定理的非条件扩散模型来估计后验分布 $p(x|y)$ 。这等同于在逆过程的每一步估计条件后验 $p(x_t|y)$ 。基于贝叶斯定理，可以推导出：

$$p(x_t|y) = p(y|x_t)p(x_t)/p(y). \quad (20)$$

And the corresponding score function can be estimated as:

$$\nabla_{x_t} \log p_t(x_t|y) = \nabla_{x_t} \log p_t(y|x_t) + s_\theta(x, t), \quad (21)$$

其中, $s_\theta(x, t)$ 可以从预训练模型中提取, 而项 $p_t(y|x_t)$ 则难以处理。从上式可以看出, 获得更好的图像恢复逆问题求解器的关键因素是准确估计 $p(y|x_t)$ 。

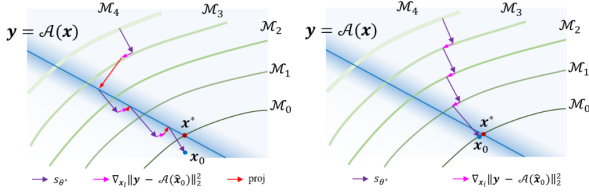


图 12: MCG [236] 和 DPS [114] 采样过程的几何图示。左图: MCG 方法在每个校正步骤 (粉红线) 后添加额外的投影步骤 (与 $y = \mathcal{A}(x)$ 线正交的红线), 这可能会在最后的去噪步骤中从流形 $\mathcal{M}_0$ 中脱落。右图: DPS 方法仅在每个扩散采样步骤后添加校正步骤 (紫色)。

作为先驱工作, MCG [236] 和 DPS [114] 用 $p(y|\hat{x}_o)$ 近似后验 $p(y|x_t)$, $\hat{x}_o$ 是利用 Tweedie 公式将 x_t 定为 $\hat{x}_o = E[x_o|x_t]$ [114] 时的期望。具体而言, MCG [236] 从数据流形的角度考虑数据一致性, 提出流形约束梯度, 使校正位于数据流形上。然而, 如图 12 所示, DPS [114] 指出 MCG 中的投影操作有损数据一致性, 因为它可能导致采样路径偏离数据流形。基于此, DPS [114] 在逆过程中舍弃了投影步骤, 将后验估计为:

$$\begin{aligned} \nabla_{x_t} \log p_t(y|x_t) &\approx \nabla_{x_t} \log p(y|\hat{x}_o) \\ &\approx -\frac{1}{\sigma^2} \nabla_{x_t} \|y - H(\hat{x}_o(x_t))\|_2^2 \end{aligned} \quad (22)$$

继上述工作之后, Π GDM [197] 进一步将公式 22 扩展为线性、非线性、可微逆问题的统一形式, 其中退化函数 h 的 Moore-Penrose 伪逆 $h^\dagger$ 为:

$$\nabla_{x_t} \log p_t(y|x_t) \approx r_t^{-2} ((h^\dagger(y) - h^\dagger(h(\hat{x}_o)))^T \frac{\partial \hat{x}_o}{\partial x_t})^T \quad (23)$$

其中 r_t^{-2} 设置为 $\sqrt{\frac{\sigma_t^2}{\sigma_t^2 + 1}}$ 。根据该公式, Π GDM 开发了其管道, 如图 13 所示。

与上述研究不同, 一些研究 [196, 216] 尝试使用其他策略对 $p(y|x_t)$ 进行建模。值得注意的是, 更高的条件概率 $p(y|x_t)$ 相当于 $D(x_t)$ 和 y 之间的距离更小 [196]。因此, GP [196] 提出了分布 $p(y|x_t)$ 的启发式近似, 如下所示:

$$p(y|x_t) \approx \frac{1}{Z} \exp(-[s\mathcal{L}(D(x_t), y) + \lambda Q(x_t)]), \quad (24)$$

其中 $\mathcal{L}$ 和 Q 分别表示距离度量和质量损失。 Z 是归一化因子, s 是控制指导权重的缩放因子。然而, 距离 $\mathcal{L}$ 很难定义, 因为 x_t 和 y 中的噪声幅度不同。因此, 它们代替

x_t 在距离测量中, 其干净的估计值 $\hat{x}_o$ 。出于同样的目的, Co-paint [216] 尝试通过神经网络的一步估计来预测 $\hat{x}_o$ 。

Feng *et al.* [215] 没有对难以处理的分布 $p(y|x_t)$ 进行建模, 而是直接从变分的角度估计后验 $p(x_t|y)$ 。遵循 DPI [238, 239], 他们通过 RealNVP [240] 定义了一个分布族 q_θ , 该分布族使用参数 θ 对流进行归一化, 并通过真实后验和估计分布 q_θ 之间的最小 KL 散度进行优化。

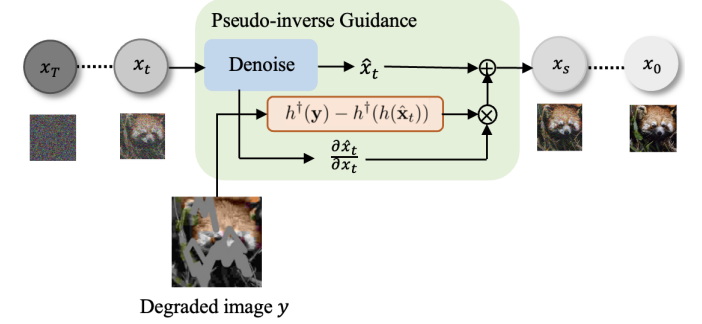


图 13: Π GDM 的架构 [197]。在每个去噪步骤中, 采用伪逆引导来保证去噪结果和退化图像 y 之间的数据一致性。

4 盲/真实世界图像恢复的扩散模型

尽管第 3 节中的方法在图像恢复方面取得了重大突破, 但其中大多数 [100、101、104、112–114、197、218、219] 侧重于解决合成失真, 而这些失真通常在分布外 (OOD) 真实世界/盲降级中表现不佳。原因源于真实世界 IR 的固有挑战: 1) 未知的降级模式难以识别。2) 收集失真/干净图像对并非易事, 甚至在现实世界中无法获得。为了解决这个问题, 先前的研究 [241–248] 试图通过模拟现实世界的退化 [72、241–244、246] 和无监督学习 [245、247、248] 等来解决这个问题。受这些研究的启发, 一些先驱性的研究 [117、118、120、123、221] 开始探索如何利用扩散模型来解决现实世界的退化问题。在本文中, 我们将基于 DM 的盲/真实世界 IR [108、109、118–121、123、220–222、226] 分为四类, 即 *i.e.*、失真模拟 [118、226]、核估计 [119、120]、域平移 [122、226] 和失真不变扩散模型 [123、222、237]。

4.1 失真模拟

值得注意的是, 现实世界的扭曲通常是盲目的/未知的, 其分布与简单的合成扭曲不同。对于基于监督学习的 IR, 这需要恢复网络具有强大的泛化能力, 或者合成数据集可以覆盖现实世界的扭曲。从因果关系的角度来看 [249], 这两个目的都

依赖于模拟与现实世界失真相似的各种失真，我们称之为失真模拟/增强。有几种基于 DM 的代表性 IR 方法 [118, 221] 利用失真模拟来提高其方法对现实世界退化的鲁棒性。代表性方法是 SR3+ [118]，它基于 SR3 [100] 的扩散模型，并引入了 RealESRGAN [72] 的二阶退化模拟进行训练。类似地，为了模拟现实世界的退化，Yang *et al.* [226] 提出使用扩散模型合成现实世界的失真/干净训练对，其中失真图像在 RealESRGAN [72] 中使用二阶退化初始化。

4.2 核估计

核估计最早是在盲图像恢复中提出的[250–255]，其中退化可以建模为 $y = (x * k) \downarrow_s + n$ 。这里， k 是退化核， n 是加性噪声。在这种设置下，可以估计核 k 作为指导，以提高恢复网络的适应性。受此启发，BlindDPS[120] 和 GibbsDDRM[119] 尝试通过估计采样过程中未知的退化核来解决盲逆问题。具体而言，BlindDPS[120] 利用 DPS[114] 架构并采用一个并行扩散模型进行退化核估计。用于核估计的扩散模型是在合成核上预先训练的。与 BlindDPS 不同的是，GibbsDDRM [119] 通过部分折叠的 Gibbs 采样器 [256] 实现采样过程，该采样器从联合后验 $p(x_t|k, y)$ 对核参数和图像一起进行采样。

4.3 域名翻译

在现实世界中，很难收集扭曲/干净的图像对。尽管一些工作试图模拟真实扭曲图像的退化过程，但合成扭曲的分布与真实世界扭曲的分布仍然相差甚远。为了进一步解决真实世界的红外问题，一系列工作探索了图像恢复的领域翻译技术。领域翻译[257–261]旨在将图像从一个领域翻译到另一个领域。从领域翻译的角度来看，合成扭曲图像、真实世界扭曲图像和高质量图像可以看作是三个不同的领域，它们共享相同的内容。

基于 DM 的 IR 的领域转换工作大致可分为两类：1) 第一类 [226] 旨在通过将低质量图像从合成域转换到真实世界域，模拟更可靠的真实世界扭曲/干净图像对。通过这种方式，模拟的数据集可以使恢复网络具有更好的真实世界退化恢复能力。例如，Yang *et al.* [226] 首次利用预先训练的扩散模型来合成真实世界的训练对，其中扩散模型用真实世界的低质量图像进行预训练，通过将合成的低质量图像扭曲到噪声空间（*i.e.*，生成反转）来实现转换。2) 另一类流行方法 [122] 利用无监督学习，

其中域转换是通过循环一致性约束实现的。具体来说，两个生成器构成一个循环路径，其中一个生成器旨在将失真图像转换为无失真图像，另一个生成器用于将干净图像转换为失真图像。这使得使用不成对的真实世界失真和高质量图像进行无监督训练成为可能。作为代表性工作，RainDiffusion [122] 提出使用两个协作分支去除雨水，其中非扩散转换分支旨在利用预先训练的循环一致性生成器来生成初始配对的干净/有雨图像，而扩散转换分支利用多尺度扩散模型来细化结果。

4.4 畸变不变扩散模型

由于盲畸变通常多样且复杂，因此要求扩散模型具备现实世界中这些畸变的生成能力（*i.e.*，畸变不变能力）。为了实现畸变不变的扩散模型，DiffFace *et al.* [237] 引入了预先训练的恢复网络 *e.g.*、SRCNN [1] 或 SwinIR [53]，以获取初始干净图像作为采样起点 x_N ，其中恢复网络使用来自 RealESRGAN [72] 的二阶退化进行训练，从而表现出良好的泛化能力并为扩散模型生成畸变不变的初始干净图像。Ren *et al.* [222] 提出实现具有多尺度退化不变指导信息的畸变不变扩散模型。他们在退化图像上采用畸变增强策略，以获得以结构信息为指导的不变表示。相比之下，Wang [123] 利用低通滤波器滤除低质量图像中的失真不变分量，因为不同的现实世界失真图像通常具有相同的结构信息。如图 14 所示，他们在采样阶段采用了类似于 ILVR 的简单迭代细化。在获得退化不变 $\hat{x}_0$ 后，他们使用增强模块（强大的基于 CNN 或基于变压器的恢复方法）进一步提高图像质量。

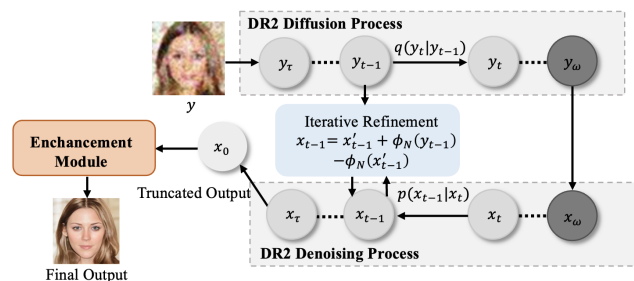


图 14：DR2 的架构 [123]。它由退化消除模块和增强模块组成。输入图像首先被送入扩散过程，以逐渐消除退化信息。输出的无退化图像将通过预先训练的面部恢复网络进行增强。

5 实验

为了确保高效、彻底地比较不同的基于扩散模型的 IR 方法，我们首先总结了不同任务的流行数据集、实验配置和评估指标。然后，我们对几种典型图像恢复任务中的现有基准进行了比较，包括图像超分辨率、修复、去模糊和 JPEG 伪影去除。

5.1 数据集和实施细节

数据集。值得注意的是，不同 IR 任务的数据集中的内容和退化模式存在显著差异。因此，我们在表 8 中总结了基于 IR 任务的常用数据集，包括 SR、图像去模糊、图像修复、阴影去除、去雪、排水和去雾。

对于传统图像 SR (*i.e.*, 双三次下采样)，标准训练数据通常由 DIV2K [45] 和 Flickr2K [262] 组成。然而，扩散模型的性能本质上受数据集大小的限制。因此，SR3 [118] 使用 ImageNet 训练扩散模型进行自然图像 SR，使用 FFHQ [211] 进行人脸 SR。在测试过程中，它利用 ImageNet 1K [263] 评估自然图像 SR，使用 CelebA-HQ 进行人脸 SR。除此之外，一系列工作还引入了常用的 SR 测试数据集进行评估，例如 Set5 [46]、Set14 [47]、BSD100 [264]、Manga109 [265]、Urban100 [266]。对于现实世界的 SR，SR3+ [118] 提供了两个版本的训练数据，其中第一个版本由 DF2K 和 OST [267] (*i.e.*, DIV2K、Flickr2K 和 OST300) 组成，第二个版本包含额外的 61M 内部图像和 DF2K+OST。对于评估，测试数据由 RealSR [268] 和 DRealSR [269] 组成，由两台配备不同镜头的 DSLR 相机获得。

对于图像去模糊，基于扩散模型的方法通常使用 GoPro 训练数据集 [52] 进行训练，并在 Gopro 测试数据集、RealBlur-J [270]、REDS [51] 和 HIDE [271] 上进行验证。在阴影去除任务中，ISTD [272] 和 SRD [273] 用于训练和评估，其中 ISTD 包含 135 个带有阴影掩模的场景。对于图像去雾任务，使用三个典型数据集进行评估，包括 Haze-4K [274]、Dense-Haze [275] 和 RESIDE [276]。其中，Haze-4K 包含 4,000 幅有雾图像，Dense-Haze [275] 由 33 对户外有雾和无雾图像组成，RESIDE [276] 收集了现实世界的 443950 幅训练图像和 5342 幅测试图像。Image Desnowing 利用三个数据集，*i.e.*, CSD [277]，Snow100k [278] 和 SRRS [279]。图像去雨中的数据包含多种雨水类型。例如，Rain100H [?]，Rain100L [?]，Rain800 [48]，DDN-Data [280] 包含大量合成雨纹。RainDrop [49] 收集了 1119 对有雨/无雨图像，具有各种背景和雨滴。Outdoor-Rain [281] 考虑了雨水的累积，为大雨图像提供了更合理的建模。SPA-data [282] 通过从多张连续雨天图像中扭曲干净的图像来构建大规模真实世界雨纹。有关上述数据集的更多详细信息，请参见表 8。

实现细节

我们分别在表 9 和表 10 中总结了监督算法和零样本算法的实现细节和数据集。对于监督算法，我们描述了训练过程和测试过程中的配置，包括批量大小、训练迭代、学习率、训练过程中的采样步骤以及推理过程中的采样步骤。对于基于零样本的方法，我们阐明了预训练的扩散模型、评估的数据集和推理过程中的采样步骤。常用的数据增强策略由旋转和翻转操作组成。

5.2 评估指标

客观和主观指标在衡量和比较基于 DM 的 IR 的不同算法之间的性能方面起着至关重要的作用。在本节中，我们详细说明了图像恢复中常用的指标，*i.e.*, PSNR、SSIM [283]、LPIPS [284]、DISTS [285]、FID [286]、KID [287]、NIQE [288] 和 PI [289]。

- PSNR 作为图像恢复中最流行的度量，旨在通过计算均方误差 (MSE) 来测量失真图像与其对应的干净图像之间的像素距离。
- SSIM [283] 也是传统的图像质量评估 (IQA) 指标，旨在满足人类视觉感知系统的要求。与 PSNR 相比，它从对比度、亮度和结构三个角度比较失真图像和清晰图像之间的相似性。为了进一步改进它，将多尺度信息引入 SSIM，称为 MS-SSIM [290]。与基于学习的 IQA 指标相比，SSIM 具有快速的计算速度，但与人类感知仍相差甚远。
- LPIPS [284] 是一种基于全参考学习的 IQA 度量，广泛应用于感知导向的图像恢复任务。它不是利用逐图统计量进行质量测量，而是利用预先训练的 AlexNet 作为特征提取器，并针对人类感知优化线性层。LPIPS 值越低，表示两幅图像在感知空间中越相似。
- DISTS [285] 观察到，两幅图像的纹理相似性和结构相似性可以分别通过 VGG [291] 中特征的均值和相关性来衡量。基于这一发现，本文在特征空间中对纹理和结构相似性进行了类似 SSIM 的距离测量。
- FID [286] (Fréchet 初始距离) 被广泛用于测量生成图像的保真度和多样性，它是初始分数 [292] 的改进。与缺乏真实世界参考图像的 IS 相比，FID 利用初始模型编码层的特征来对采样图像的多元高斯分布进行建模，并计算生成图像分布与参考图像之间的 Fréchet 距离。
- KID [287] 和 FID 利用初始模型中的相同特征进行质量评估，但拥有不同的距离测量策略 (*i.e.*,

TABLE 3: Quantitative Results of Supervised Models on Super-resolution Task

Models	DIV2K [45]				Urban100 [266]				Time(s/image)	Parameters	Flops (G)
	PSNR	SSIM	LPIPS	FID	PSNR	SSIM	LPIPS	FID			
Bicubic	25.36	0.643	0.31	108.2	24.26	0.628	0.34	134.3	-	-	-
SR3 [100]	26.17	0.682	0.24	111.23	25.18	0.693	0.21	61.15	50.4	155.3M	155.2G
SRDiff [101]	26.87	0.694	0.21	110.33	26.49	0.791	0.17	51.37	4.5	13.2M	84.22G
ResDiff [217]	27.94	0.723	0.23	107.69	27.43	0.824	0.14	42.35	-	-	-
CDPMSR [102]	27.43	0.712	0.19	101.23	26.98	0.801	0.16	39.87	-	-	-
IDM [223]	27.13	0.703	0.18	106.32	26.76	0.657	0.13	43.98	59.5	116.6M	-

表 4：超分辨率任务中零样本模型的定量结果。粗体数据代表最佳性能。这里我们还将一种监督方法 StableSR [121] 与零样本方法进行了比较，因为它具有现实世界的恢复能力

Models	ImageNet 1K				CelebA 1K				Time(s/image)	Flops (G)
	PSNR	SSIM	LPIPS	FID	PSNR	SSIM	LPIPS	FID		
Bicubic	25.36	0.643	0.27	108.20	24.26	0.628	0.34	134.29	-	-
ILVR [195]	27.40	0.871	0.21	43.76	31.59	0.878	0.22	32.24	41.3	1113.75
SNIPS [110]	24.31	0.684	0.21	124.54	27.34	0.675	0.27	105.19	31.4	-
DDRM [112]	27.38	0.869	0.22	40.75	31.64	0.946	0.19	31.04	10.1	1113.75
DPS [114]	25.88	0.814	0.15	39.24	29.65	0.878	0.18	29.97	141.2	1113.75
DDNM [113]	27.46	0.871	0.15	40.15	31.64	0.945	0.16	28.27	15.5	1113.75
GDP [196]	26.51	0.832	0.14	38.45	28.65	0.876	0.17	27.51	3.1	1113.76
StableSR* [121]	25.76	0.842	0.09	25.29	25.34	0.823	0.08	16.12	18.15	-

使用多项式核的最大均值差异 (MMD) 来计算。具体而言，即使样本很少，KID 也比 FID 更稳定。

- NIQE [288] 是一种早期的无参考/盲图像质量评估指标，其中质量得分是通过采用多元高斯模型 (MGM) 计算失真图像的自然场景静态 (NSS) 与自然图像之间的距离来计算的。
- PI 是在 PIRM 感知 SR 挑战赛 [289] 中提出的，旨在评估超分辨图像的感知质量。它定义为 $PI = 0.5((10 - Ma) + NIQE)$ ，其中 Ma [293] 是 SR 的无参考 IQA 指标。

5.3 实验结果

为了证明不同扩散模型的优越性，我们在多个任务上对它们进行了客观的质量比较。具体来说，我们选择了三个常见的 IR 任务，包括图像超分辨率、图像去模糊和图像修复。评估指标包括 PSNR、SSIM [283]、FID [286] 和 LPIPS [284]。为了比较计算成本和网络复杂度，我们还测量了基于扩散模型的 IR 方法的运行时间、参数和触发器。一些扩散模型的定性结果如图 15、图 16 和图 17 所示。

图像超分辨率结果。基于监督扩散模型的 IR 模型在 4 倍图像超分辨率上的实验结果列于表 3，其中包括

在 DIV2K [45] 和 Urban100 [266] 数据集上进行测试。我们发现 Resdiff [217] 在 PSNR 和 SSIM 方面表现非常出色，与其他扩散模型相比，PSNR 提高了大约 0.5dB。这是因为 Resdiff 利用扩散模型生成残差信息，并使用预处理图像进行条件生成，从而确保恢复图像与像素级高分辨率图像的一致性。相比之下，IDM [223] 和 CDPMSR [102] 在主观指标上表现良好。它们利用预处理图像或预处理特征作为条件输入，这被证明对感知质量有有益的影响。在模型参数、运行时间和计算复杂度方面，SRdiff 的表现远远优于 IDM 和 SR3。SRdiff 生成一张图像耗时 43.5 秒，耗费 84.22 GFlops 和 13.2M 个参数。这是因为 SRDiff 将低分辨率图像编码到潜在空间，从而降低了处理的维度。此外，它有 100 个采样步骤，因此采样速度比其他两个模型更快。

对于基于零样本的扩散 IR 模型，定量比较如表 4 所示。我们在两个数据集上测试了六个开源扩散模型，以实现 4 倍超分辨率：ImageNet 1K [263] 和 CelebA 1K [294]。从表中可以看出，DDRM [112] 和 DDNM [113] 在各种指标上表现良好，其次是 ILVR [195]。这是因为 DDRM 和 DDNM 从分解的角度考虑了与低分辨率图像的数据一致性，而 ILVR 仅确保了低频一致性。



图 15：扩散模型对图像超分辨率的定性结果



Fig. 16：图像 Deb 上的扩散模型的定性结果

劝诱

DPS [114] 和 GDP [196] 更注重生成图像的感知质量。GDP 在两个数据集上的感知指标都表现良好，因为它使用经验公式而不是 Tweedie 公式进行后验估计，因此相对于 DPS 而言性能有所提升。GDP 还表现出最快的图像生成速度，反向采样步数为 25。原因是它不需要 DDRM 中复杂的 SVD 分解和计算，从而加速了图像生成。另一方面，DPS 在每个采样步骤（大约 1000 步）之后进行校正，因此无法利用基于 DDIM 的采样方法进行加速。因此，使用 DPS 生成单张图像大约需要 141.2 秒。

一种即插即用的采样方法，并合并了 DDIM 采样策略，在加快采样过程的同时确保生成图像的保真度和真实感。Diracdiff 包括感知优化（PO）和失真优化（DO）模型，并采用增量重建和早期停止方法来实现感知-失真权衡。因此，这两个模型在失真和感知指标上都表现非常出色。在所有使用的模型中，DDRM 的采样时间最短，平均每张图像不到 10 秒，因为它仅使用 20 个采样步骤。在模型参数方面，所有基于零样本 DM 的 IR 方法都使用预训练模型，对于 ImageNet 数据集有 552.8M 个参数，对于 CelebA 数据集有 126M 个参数。

图像去模糊结果。我们还使用 ImageNet 1K [263] 和 CelebA 1K [294] 数据集评估了高斯去模糊任务中五种基于零样本 DM 的 IR 方法。实验结果如表 2 所示。我们可以发现 DiffPIR [198] 和 Dirac-DO [200] 在 PSNR 和 SSIM 上实现了具有竞争力的性能，平均比 DDRM [112] 和 DDNM [113] 提高了 1.0 dB 到 1.4 dB。此外，Dirac-PO [200] 和 DiffPIR [198] 在感知指标上表现出色。DPS [114] 在感知指标（包括 LIPS 和 FID）上表现良好，但每幅图像的生成时间较长。DiffPIR 利用

图像修复结果。我们验证了五个零样本扩散模型在图像修复（窄掩模）任务上的性能，如表 6 所示。除了三个多任务模型 DPS [114]、DDRM [112] 和 DDNM [113] 之外，我们还添加了两个专门为图像修复设计的模型，Repaint [212] 和 Copaint [216]。在失真度量方面，与超分辨率的情况类似，DDRM 和 DDNM 实现了更好的性能。在感知质量方面，DPS 在 CelebA-HQ 上表现出比 ImageNet 好得多的感知性能，而 Repaint 和 Copaint 优于其他



图 17：扩散模型在图像修复中的定性结果

表 5：扩散模型在图像去模糊任务上的定量结果

Target	Models	ImageNet 1K				CelebA 1K				Time(s/image)	Flops (G)
		PSNR	SSIM	LPIPS	FID	PSNR	SSIM	LPIPS	FID		
Gaussian Deblurring	Deblurred imaged	19.56	0.553	0.61	168.20	19.95	0.578	0.53	164.31	-	-
	DPS [114]	21.51	0.516	0.41	52.61	25.56	0.688	0.25	65.67	130.1	1113.75
	DDRM [112]	24.54	0.668	0.39	71.65	27.21	0.767	0.29	87.32	8.9	1113.75
	DDNM [113]	25.37	0.731	0.33	56.54	27.25	0.782	0.25	62.65	16.2	1113.75
	Dirac-DO [200]	25.18	0.719	0.43	75.29	28.46	0.848	0.31	91.32	-	-
	Dirac-PO [200]	24.23	0.673	0.34	53.91	26.76	0.742	0.24	59.45	-	-
	DiffPIR [198]	25.26	0.721	0.32	53.21	28.35	0.832	0.25	58.56	23.5	1113.78

模型。Copaint 优于 Repaint 模型，在 CelebA-HQ 和 ImageNet 数据集上分别将 FID 指标降低了 0.08dB 和 1.33dB。这是因为 Copaint 从贝叶斯后验估计的角度考虑了未显示区域和显示区域之间的一致性，这比 Repaint 中使用的重采样策略更有理论支持。同时，Copaint 还采用了 DDNM [113] 中的时间旅行方法以获得更好的恢复质量，但它也增加了计算复杂度。在运行时间方面，由于 DDRM 的采样步数少（运行时间与 N FE 几乎线性相关），因此仍然是生成单幅图像的最快模型。虽然 CoPaint 的步长比 DPS 短，为 250 步，但它采用了时间旅行，这将生成单幅图像的采样运行时间增加到大约 298 秒。同样，Repaint 也需要类似的时间来生成图像，但 CoPaint 的采样时间比 Repaint 略快。

推广到看不见的失真。在本节中，我们比较了基于扩散模型的 IR 方法与现有的基于 CNN 和基于 Transformer 的 IR 方法在图像超分辨率上的泛化能力。这些方法的框架使用合成数据集 DIV2K [45] 进行训练，并在看不见的真实世界数据集 *i.e.*, RealSR [268] 上进行评估。如表 7 所示，在可见的退化上，基于 CNN 和基于 Transformer 的 IR 方法可以在

PSNR/SSIM，而基于扩散的 IR 方法可以实现令人满意的主观质量。然而，它们在未见场景中的表现都不佳，包括 PSNR、SSIM 和感知度量 LPIPS。相反，StableSR 方法在未见场景中表现出卓越的泛化能力。原因是它们利用了 RealESRGAN 的失真合成策略，旨在模拟现实世界的退化。

任意尺寸的图像恢复。通常，扩散模型生成的图像的分辨率需要与优化过程一致。这种限制阻碍了基于扩散模型的图像恢复处理任意尺寸的失真图像，尤其是对于高分辨率图像，例如 2K 和 4K。一个直观的解决方案是，我们可以生成整个图像的每个部分，然后将它们拼接成一张图像。然而，这将导致严重的不匹配问题和每个部分边缘的不一致，这是由于扩散模型固有的随机性造成的。最近，几篇论文 [104, 121, 196, 222] 提出了针对此问题的非常有效的解决方案。特别是，DG-DPM [222] 使用全卷积层来处理任意大小的输入，但由于网络结构庞大，该方法的计算成本很高。相比之下，Weather-diff [104] 和 GDP [196] 都采用基于块的恢复方法。它们从输入图像中提取重叠的块，并将每个块输入到扩散过程中

TABLE 6: Quantitative Results of Diffusion Models on Image Inpainting Task

Target	Models	ImageNet 1K				CelebA-HQ				Time(s/image)	Parameters	Flops (G)
		PSNR	SSIM	LPIPS	FID	PSNR	SSIM	LPIPS	FID			
Inpainting	Masked Imaged	14.86	0.712	0.37	79.83	15.67	0.794	0.36	181.60	-	-	-
	Repaint [212]	31.87	0.967	0.07	4.31	35.23	0.982	0.04	6.19	353.4	552.7M(ImageNet)	10736.60
	DDRM [112]	31.72	0.966	0.18	8.82	35.73	0.967	0.16	10.82	8.8		1113.75
	DPS [114]	30.87	0.929	0.23	28.32	33.48	0.943	0.08	5.37	136.3		1113.75
	DDNM [113]	32.06	0.969	0.11	7.89	35.64	0.982	0.11	7.54	16.4		1113.75
	Copaint [216]	31.93	0.976	0.08	4.23	34.97	0.975	0.04	4.86	293.8		6083.27

表 7：基于 CNN 的（RRDB [295]）、基于 transformer 的（SwinIR [53]）和基于 DM 的（SRdiff [101]、StableSR [121]）在分布数据集 RealSR [268] 上的比较

Models	DIV2K			RealSR		
	PSNR(HWC/Y)	SSIM(HWC/Y)	LPIPS	PSNR(HWC/Y)	SSIM(HWC/Y)	LPIPS
Bicubic	(26.93,28.45)	(0.756,0.778)	0.395	(26.29, 28.03)	(0.768,0.807)	0.437
SRdiff [101]	(27.70/29.22)	(0.764/0.786)	0.116	(26.28/28.01)	(0.768,0.807)	0.428
SwinIR [53]	(30.37/31.85)	(0.835/0.853)	0.241	(26.30,27.96)	(0.770,0.808)	0.449
RRDB [295]	(29.68/31.25)	(0.822/0.841)	0.26	(26.28/28.01)	(0.769,0.808)	0.445
StableSR [121]	(21.88/23.26)	(0.559/0.573)	0.26	(21.14,22.69)	(0.627,0.663)	0.095

去噪。对块的重叠部分在噪声维度上进行平均，以保持不同块之间的一致性。Stable-SR [121] 也采用了基于块的方法，同时使用高斯滤波器来平滑块重叠部分的噪声。我们在图 18 中展示了它们的有效性，其中不同块之间的重叠区域几乎无法区分。

6 挑战和未来方向

尽管最近的研究在基于扩散模型的 IR 方面取得了显著进展，但由于其有限的鲁棒性，模型复杂性，运行效率和恢复能力，将其扩展到实际应用仍然存在一些挑战。为了进一步促进图像恢复的发展，我们在本节中总结了主要挑战并提出了解决它们的潜在方向。

6.1 采样效率

值得注意的是，采样效率是扩散模型面临的一个典型挑战，因为较少的采样步骤会导致生成保真度受限。扩散模型的这一固有问题损害了图像恢复的训练和推理速度。如表 3 所示，SR3 [100] 需要大约 50 秒来恢复一幅大小为 224×224 的图像，这比现有的 IR 方法慢得多。先前关于扩散模型的研究试图从四个角度提高采样效率：1) 使用非马尔可夫链对扩散过程进行建模，例如 DDIM [85]。2) 设计高效的 ODE 求解器、e.g.、DPM 求解器。3) 利用知识蒸馏减少采样步骤 [157–160]，以及 4) 引入具有条件的跨模态先验

机制 [116, 117, 121]。在上述改进下，扩散模型的采样步骤大幅减少到 $10 \sim 20$ 步，这也有助于更快速地恢复图像。特别是，DDRM [112] 使用 DDIM [85] 的采样策略将一幅 224×224 图像的推理速度降低到 8 秒。

尽管如此，上述策略并不特定于图像恢复任务。不同的是，考虑到 IR 中的低质量图像包含丰富的结构和文本信息，一些工作 [106, 220, 231] 通过从低质量图像而不是纯噪声中采样来实现图像恢复，这避免了原始 DDPM 中的额外采样步骤。尽管过程很大，但它与实时应用仍有很大的差距，迫切需要解决。通过提高采样效率来加速基于扩散模型的 IR 将是一个潜在的方向。

6.2 模型压缩

模型大小也是影响计算成本的重要因素，限制了基于扩散模型的图像恢复（IR）的实时应用，例如移动设备。特别是，DDPM [82] 和 SR3 [100] 分别具有 113.7M 和 155.3M 个参数，大大超过了之前基于 CNN [34、35、232、296–300] 或基于 Transformer 的 IR 主干 [10、16、22、53–55、301]。为了缓解这种情况，扩散模型压缩是实现高效 IR 的一个潜在但尚未得到充分探索的研究方向。模型压缩 [302] 的目标是在保持任务性能的同时降低计算成本，已从四个方面取得了重大突破：1) 模型剪枝，旨在通过估计每个参数的重要性得分来去除不重要的参数，2) 模型量化目标



Stable-SR (512x512)

图 18：任意



Weatherdiff(958x1438)

尺寸

降低浮点参数的位深度以方便存储或计算；3) 提出知识蒸馏，将知识从复杂的教师模型转移到简单高效的学生模型；4) 低秩分解用于将参数张量分解为多个低秩张量。在此基础上，一些工作在研究扩散模型的模型压缩方面迈出了一步。Kim *et al.* [303] 为扩散模型引入了块去除知识蒸馏，通过从 U-Net 架构中移除一些残差和注意力块来构建学生模型。Fang *et al.* [304] 指出并非所有扩散步骤都对生成过程有贡献，然后利用部分扩散步骤来估计重要和不重要的权重，并使用泰勒展开式进行参数剪枝。此外，还有一些开创性的工作 [305–308] 对扩散模型进行模型量化，以加速采样过程。尽管取得了这样的进展，但是很少有研究探索如何设计基于扩散模型的 IR 的模型压缩，以期开发实时应用。

6.3 失真模拟与估计

真实世界/盲 IR 是一项具有挑战性但意义重大的任务，其目标是解决现实世界中遇到的未知且复杂的退化问题。与合成退化不同，合成退化中失真预定义的，并且有成对的训练样本可用，而收集成对的真实世界失真/干净对并非易事，从而阻碍了使用监督学习进行训练。为了解决这一限制，引入了无监督学习来利用不成对的真实世界失真/干净图像。然而，这种学习范式通常会产生不令人满意的结果

确保恢复图像和低质量图像之间的纹理一致性。相反，失真模拟是另一种通过模拟现实世界的退化来维持监督学习的有效策略。通常，RealESGRAN [72] 和 SR3+ [118] 是探索真实世界 IR 手工制作的二阶退化的代表性工作。尽管如此，手工制作的失真模拟很难涵盖现实世界中的所有退化。为了缓解这个问题，一些工作受到领域翻译的启发，并引入 GAN / 扩散模型将合成的失真图像转换为真实世界图像或将真实世界的失真图像转换为合成图像。前者旨在模拟真实世界的训练对以进行监督学习 [226]，而后者旨在直接利用合成图像训练的 IR 网络 [122]。

现实世界/盲 IR 的另一个关键挑战来自失真估计，这涉及显式/隐式识别失真类型或水平。在本文中，我们从两个角度总结了失真估计的利用：1) 失真自适应学习和 2) 解决 IR 中的逆问题。从第一个角度来看，一个值得注意的例子是盲 IR 的核预测 [120]，其中估计的核/表示用于指导预训练的 IR 模型对未知退化的适应。受此启发，如果我们能够以显式/隐式的方式估计失真类型或程度，我们就可以实现基于失真自适应学习的统一 IR 框架。对于第二个角度，如第 3.4 节所述，许多基于零样本扩散模型的 IR 方法都是基于线性逆问题的建模。这对识别退化模式提出了要求，这对于一致性是必要的

扩散模型中的约束。因此，由于现实世界中的失真模式难以识别，因此大多数都致力于合成失真。迫切需要开发一种失真估计技术，以将基于零样本扩散模型的IR扩展到现实世界的应用。

6.4 失真不变学习

近期，我们目睹了基于扩散模型的IR在特定退化方面的快速发展。然而，当应用于未知的失真类型和程度时，它不可避免地会遇到鲁棒性不令人满意的问题。这就提出了一个基本问题：如何在不同的失真类型和程度上实现一致的图像恢复？为了实现这一目标，我们提出了一个称为失真不变学习（DIL）[309]的方向，旨在使IR模型能够推广到未知和多样的退化。DIL的原理是学习在各种退化模式下不变的表示，并保留足够的结构和文本信息以进行重建。

受域泛化（DG）[310–314]的启发，我们可以通过将每个失真模式视为一个域来实现IR的DIL。在DG领域，有三种典型的方法来学习域不变特征，包括域对齐[315–319]、数据增强[320–323]和元学习[324–327]。具体而言，域对齐旨在通过最小化对比损失[328]、最大平均差异（MMD）或对抗学习等来对齐源域和目标域的表示。利用数据增强来扩展域多样性和一致性，使模型获得域不变能力。元学习旨在通过对齐不同域之间的梯度来学习域不变表示，这是从优化的角度出发的。通过将失真模式视为一个特定领域，我们可以获得几种实现失真不变学习的策略：1）我们可以利用IR的编码器-解码器架构，并在解码器之前对齐来自不同失真图像的表示。2）第二种策略可以从失真增强*i.e.*中学习失真不变表示，尽可能模拟现实世界中的各种失真。3）使用元学习（如[309]）优化IR中的经验风险最小化。

对于基于扩散模型的图像恢复，模型通常由两部分组成：噪声预测器和条件模块。因此，我们可以从两个角度实现失真不变学习：1）学习失真不变噪声预测器；2）失真不变条件。显然，一旦我们实现了失真不变条件，我们就可以在监督IR中保持噪声预测器不变，或者在零样本IR中利用预训练的扩散模型。基于此，一些先驱工作尝试重新设计条件模块以实现失真不变条件，*e.g.*、DiffFace[237]和DR2[123]。值得注意的是，失真不变条件也依赖于失真不变学习来获得更好的条件，这仍然需要在未来的工作中付出大量努力来改进。

6.5 框架设计

作为图像恢复的基础，如何设计一个有效而强大的IR框架是一个持续且重要的问题。我们可以注意到，最新的基于扩散模型的IR方法[100, 101, 112–114, 196, 197, 219]都是基于DDPM[82]的U-Net架构设计的，并从三个角度追求更好的框架，即*i.e.*、条件策略[10, 112, 114, 212, 231]、生成空间[98, 106, 224]和噪声预测器[224]。IR中扩散模型的条件旨在从低质量图像中引入结构和文本信息。在早期工作中，SR3直接通过连接选择低质量图像作为条件。为了改善这种状况，一些研究[102, 191, 217, 223, 230]通过设计预处理网络（如特征提取器和预训练的恢复网络）来改善这种状况。从生成空间来看，框架通常从四个空间设计，包括图像空间、残差空间、潜在空间和频率空间。其中，逐像素空间可以保留更多的空间结构和文本信息，可以生成高质量的图像[102, 217, 223]或残差[101, 217]，同时计算成本和参数较高。相比之下，潜在空间生成需要的计算成本较少。然而，一个精心设计的编码器和解码器对于潜在空间生成至关重要，需要在效率和保真度之间做出权衡。频率空间在图像恢复中被广泛应用，包括小波变换、傅里叶变换等。与图像空间相比，频率空间更善于捕捉全局上下文信息，其中低频指的是结构信息，高频表示纹理和风格信息。继DDPM[82]之后，在大多数作品中，噪声预测器都基于U-Net架构。对于基于监督扩散模型的IR，噪声预测器的修改通常是通过增加U-Net中的残差块数量或调整不同分辨率的通道乘数来实现的，比如SR3。很少有作品致力于研究如何为基于扩散模型的IR中的噪声预测器设计基于Transformer的全新架构。此外，如何为基于扩散模型的IR任务设计像Painter[329]那样统一的基础架构也迫切需要探索。

7 结论

本研究全面回顾了近期流行的IR扩散模型，挖掘了它们强大的生成能力以增强结构和纹理恢复。首先，我们说明了扩散模型的定义和演变。随后，我们从训练策略和退化场景的角度对现有工作进行了系统的分类。具体来说，我们将现有工作分为三个主要流程：基于监督DM的IR、基于零样本DM的IR和基于DM的盲/真实世界IR。对于每个流程，我们根据技术提供细粒度的分类法，并详细描述它们的优点和缺点。对于评估，我们总结了基于DM的IR的常用数据集和评估指标。并且我们比较了开放式

在三个典型任务（包括图像 SR、去模糊和修复）上采用具有失真和感知度量的 SOTA 方法。为了克服基于 DM 的 IR 中的潜在挑战，我们重点介绍了未来有望探索的五个潜在方向。

8 致谢

我们衷心感谢 MMLab@NTU 的 Chen Chang Loy 教授对本文提供的宝贵建议和帮助，这极大地提高了我们的工作质量。

参考

[1] C. Dong、C. C. Loy、K. He 和 X. Tang, “学习深度卷积网络用于图像超分辨率”, 载于 *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings, Part IV* 13. Springer, 2014 年, 第 184–199 页。[2] J. Kim、J. K. Lee 和 K. M. Lee, “使用非常深的卷积网络实现准确的图像超分辨率”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016 年, 第 1646–1654 页。[3] B. Lim、S. Son、H. Kim、S. Nah 和 K. Mu Lee, “增强深度残差网络用于单图像超分辨率”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017 年, 第 136–144 页。[4] T. Dai、J. Cai、Y. Zhang、S.-T. Xia 和 L. Zhang, “二阶注意力网络用于单图像超分辨率”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019 年, 第 11 065–11 074 页。[5] C. Ledig、L. Theis、F. Huszár、J. Caballero、A. Cun-ningham、A. Acosta、A. Aitken、A. Tejani、J. Totz 和 Z. Wang *et al.*, “使用生成对抗网络实现照片级逼真的单图像超分辨率”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 4681–4690 页。[6] X. Wang、K. Yu、S. Wu、J. Gu、Y. Liu、C. Dong、Y. Qiao 和 C. Change Loy, “Esrgan: 增强型超分辨率生成对抗网络”, 载于 *Proceedings of the European conference on computer vision (ECCV) workshops*, 2018 年, 第 0–0 页。[7] Z. Lu、J. Li、H. Liu、C. Huang、L. Zhang 和 T. Zeng, “用于单图像超分辨率的变换器”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 457–466 页。[8] Z. Lu、H. Liu、J. Li 和 L. Zhang, “用于单图像超分辨率的高效变换器”, 载于 *arXiv preprint arXiv:2108.11084*, 2021 年。[9] X. Chen、X. Wang、J. Zhou 和 C. Dong, “在图像超分辨率变换器中激活更多像素。” *arxiv* 2022, “*arXiv preprint arXiv:2205.04437*。[10] D. Zhang、F. Huang、S. Liu、X. Wang 和 Z. Jin, “Swinfir: 通过快速傅里叶卷积和改进的图像超分辨率训练重新审视 Swinir”, *arXiv preprint arXiv:2208.11247*, 2022 年。[11] S. Nah、T. H yun Kim 和 K. Mu Lee, “用于动态场景的深度多尺度卷积神经网络

去模糊”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 3883–38 91 页。[12] O. Kupyn、V. Budzan、M. Mykhailych、D. Mishk in 和 J. Matas, “Deblurgan: 使用条件对抗网络进行盲运动去模糊”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 8183–8192 页。[13] T. M. Nimisha、A. Kumar Singh 和 A. N. Rajagopalan, “用于盲去模糊的模糊不变深度学习”, 载于 *Proceedings of the IEEE international conference on computer vision*, 2017 年, 第 4752–4760 页。[14] S. Zhao、Z. Zhang、R. Hong、M. Xu、Y. Yang 和 M. Wang, “Fcl-gan: 一种轻量级实时无监督盲图像去模糊基线”, 载于 *Proceedings of the 30th ACM International Conference on Multimedia*, 2022 年, 第 6220–6229 页。[15] K. Kim、S. Lee 和 S. Cho, “Mssnet: 用于单图像去模糊的多尺度阶段网络”, 载于 *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part II*. Springer, 2023 年, 第 524–539 页。[16] S. W. Zamir、A. Arora、S. Khan、M. Hayat、F. S. Khan 和 M.-H. Yang, “Restormer: 用于高分辨率图像恢复的高效变换器”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 5728–5739 页。[17] F.-J. Tsai、Y.-T. Peng、Y.-Y. Lin、C.-C. Tsai 和 C.-W. Lin, “Stripformer: 用于快速图像去模糊的条带变换器”, 载于 *Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XIX*. Springer, 2022 年, 第 146–162 页。[18] K. Dabov、A. Foi、V. Katkovnik 和 K. Egiazarian, “通过稀疏 3-d 变换域协同过滤进行图像去噪”, *IEEE Transactions on image processing*, 第 16 卷, 第 8, 第 2 080–2095 页, 2007 年。[19] K. Zhang、W. Zuo、Y. Chen、D. Meng 和 L. Zhang, “超越高斯去噪器: 深度 cnn 的残差学习用于图像去噪”, *IEEE transactions on image processing*, 第 26 卷, 第 7 期, 第 3142–3155 页, 2017 年。[20] Y. Chen 和 T. Pock, “可训练非线性反应扩散: 一种快速有效图像恢复的灵活框架”, *IEEE transactions on pattern analysis and machine intelligence*, 第 39 卷, 第 6, 第 1256–1272 页, 2016 年。[21] D. Valsesia、G. Fracastoro 和 E. Magli, “深度图卷积图像去噪”, *IEEE Transactions on Image Processing*, 第 29 卷, 第 8226–8237 页, 2020 年。[22] C.-M. Fan、T.-J. Liu 和 K.-H. Liu, “Sunet: 用于图像去噪的 swin 变换 unet”, 载于 *2022 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2022 年, 第 2333–2337 页。[23] X. Liu、Y. Hong、Q. Yin 和 S. Zhang, “Dnt: 从单个噪声图像中学习无监督去噪变换器”, *Proceedings of the 4th International Conference on Image Processing and Machine Vision*, 2022 年, 第 50–56 页。[24] C. Tian、M. Zheng、W. Zuo、B. Zhang、Y. Zhang 和 D. Zhang, “使用小波变换进行多阶段图像去噪”, *Pattern Recognition*, 第 134 卷, 第 109050 页, 2023 年。

- [25] D. Zhang, F. Zhou, Y. Jiang 和 Z. Fu, “Mm-bsn: 基于盲点网络的多掩模真实世界自监督图像去噪”, *arXiv preprint arXiv:2304.01598*, 2023 年.
- [26] J. Xie、L. Xu 和 E. Chen, “使用深度神经网络进行图像去噪和修复”, *Advances in neural information processing systems*, 卷. 25, 2012.
- [27] D. Pathak、P. Krahenbuhl、J. Donahue、T. Darrell 和 A. A. Efros, “上下文编码器: 通过修复进行特征学习”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016 年, 第 2536-2544 页.
- [28] J. Yu、Z. Lin、J. Yang、X. Shen、X. Lu 和 T. S. Huang, “具有上下文注意力的生成性图像修复”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 5505-5514 页.
- [29] L. Zhao, Q. Mo, S. Lin, Z. Wang, Z. Zuo, H. Chen, W. Xing 和 D. Lu, “Uctgan: 基于无监督跨空间翻译的多样化图像修复”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020 年, 第 5741-5750 页.
- [30] T. Yu, R. Feng, R. Feng, J. Liu, X. Jin, W. Zeng 和 Z. Chen, “修复一切: 当分割一切遇到图像修复”, *arXiv preprint arXiv:2304.06790*, 2023 年.
- [31] W. Li, Z. Lin, K. Zhou, L. Qi, Y. Wang 和 J. Jia, “Mat: 用于大孔图像修复的掩模感知变换器”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022 年, 第 10758-10768 页.
- [32] C. Dong, Y. Deng, C. C. Loy 和 X. Tang, “通过深度卷积网络减少压缩伪影”, *Proceedings of the IEEE international conference on computer vision*, 2015 年, 第 576-584 页.
- [33] P. Svoboda、M. Hradis、D. Barina 和 P. Zemcik, “使用卷积神经网络去除压缩伪影”, *arXiv preprint arXiv:1605.00366*, 2016 年.
- [34] M. Ehrlich、L. Davis、S.-N. Lim 和 A. Shrivastava, “量化引导的 jpeg 伪影校正”, 载于 *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII* 16. Springer, 2020 年, 第 293-309 页.
- [35] J. Jiang、K. Zhang 和 R. Timofte, “实现灵活的盲 jpeg 伪影去除”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 4997-5006 页.
- [36] X. Jiang, W. Tan, R. Cheng, S. Zhou, 和 B. Yan, “学习视差变换网络用于立体图像 JPEG 伪影去除”, *Proceedings of the 30th ACM International Conference on Multimedia*, 2022 年, 第 6072-6082 页.
- [37] X. Wang, X. Fu, Y. Zhu, 和 Z.-J. Zha, “通过对比表示学习去除 JPEG 伪影”, *Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVII*. Springer, 2022 年, 第 615-631 页.
- [38] X. Jiang, W. Tan, Q. Lin, C. Ma, B. Yan, 和 L. Shen, “多模态深度网络用于 JPEG 伪影减少”, *arXiv preprint arXiv:2305.02760*, 2023 年.
- [39] G. Chantzas, N. P. Galatsanos, R. Molina, 和 A. K. Katsaggelos, “使用空间加权总变分图像先验乘积的变分贝叶斯图像恢复”, *IEEE transactions on image processing*, 第 19 卷, 第 2 期, 第 351-362 页, 2009 年.
- [40] J. Mairal、G. Sapiro 和 M. Elad, “学习多尺度稀疏表示以进行图像和视频恢复”, *Multiscale Modeling & Simulation*, 第 7 卷, 第 1 期, 第 214-241 页, 2008 年.
- [41] Z. Xu 和 J. Sun, “利用块稀疏性通过块传播进行图像修复”, *IEEE transactions on image processing*, 第 19 卷, 第 2 期, 第 351-362 页, 2009 年.
- [42] K. I. Kim 和 Y. Kwon, “使用稀疏回归和自然图像先验实现单图像超分辨率”, *IEEE transactions on pattern analysis and machine intelligence*, 第 32 卷, 第 6 期, 第 1127-1133 页, 2010 年.
- [43] W. Dong、L. Zhang、G. Shi 和 X. Li, “用于图像恢复的非局部中心化稀疏表示”, *IEEE transactions on Image Processing*, 第 22 卷, 第 4 期, 第 1620-1630 页, 2012 年.
- [44] U. Schmidt 和 S. Roth, “用于有效图像恢复的收缩场”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014 年, 第 2774-2781 页.
- [45] E. Agustsson 和 R. Timofte, “Ntire 2017 单图像超分辨率挑战: 数据集和研究”, 载于 *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2017 年 7 月.
- [46] M. Bevilacqua、A. Roumy、C. Guillemot 和 M. L. Alberi-Morel, “基于非负邻域嵌入的低复杂度单图像超分辨率”, 2012 年.
- [47] R. Zeyde、M. Elad 和 M. Protter, “使用稀疏表示进行单图像放大”, 载于 *Curves and Surfaces: 7th International Conference, Avignon, France, June 24–30, 2010, Revised Selected Papers* 7. Springer, 2012 年, 第 711-730 页.
- [48] H. Zhang、V. Sindagi 和 V. M. Patel, “使用条件生成对抗网络进行图像去雨”, *IEEE transactions on circuits and systems for video technology*, 第 30 卷, 第 11 期, 第 3943-3956 页, 2019 年.
- [49] R. Qian、R. T. Tan、W. Yang、J. Su 和 J. Liu, “用于从单幅图像中去除雨滴的细心生成对抗网络”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 2482-2491 页.
- [50] H. Zhang 和 V. M. Patel, “使用多流密集网络进行密度感知的单幅图像去雨”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 695-704 页.
- [51] S. Nah、S. Baik、S. Hong、G. Moon、S. Son、R. Timofte 和 K. M. Lee, “Ntire 2019 视频去模糊和超分辨率挑战: 数据集和研究”, 载于 *CVPR Workshops*, 2019 年 6 月.
- [52] S. Nah、T. H. Yun Kim 和 K. Mu Lee, “用于动态场景去模糊的深度多尺度卷积神经网络”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 3883-3891 页.
- [53] Z. Liu、Y. Lin、Y. Cao、H. Hu、Y. Wei、Z. Zhang、S. Lin 和 B. Guo, “Swin transformer: 使用移位窗口的分层视觉变换器”, 载于 *Proceedings of*

- the IEEE/CVF international conference on computer vision, 2021 年, 第 10 012-10 022 页。[54] Z. Liu、H. Hu、Y. Lin、Z. Yao、Z. Xie、Y. Wei、J. Ning、Y. Cao、Z. Zhang、L. Dong et al., “Swin transformer v2: 扩大容量和分辨率”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022 年, 第 12 009-12 019 页。[55] Z. Wang、X. Cun、J. Bao、W. Zhou、J. Liu 和 H. Li, “Uformer: 用于图像恢复的通用 u 形 transformer”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 17 683-17 693 页。[56] A. Krizhevsky、I. Sutskever 和 G. E. Hinton, “使用深度卷积神经网络”, *Communications of the ACM*, 第 60 卷, 第 6 期, 第 84-90 页, 2017 年。[57] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、Ł. Kaiser 和 I. Polosukhin, “Attention is all you need”, *Advances in neural information processing systems*, 第 30 卷, 2017 年。[58] D. P. Kingma 和 M. Welling, “自动编码变分贝叶斯”, *arXiv preprint arXiv:1312.6114*, 2013 年。[59] A. Van Den Oord、N. Kalchbrenner 和 K. Kavukcuoglu, “像素循环神经网络”, 载于 *International conference on machine learning*. PMLR, 2016 年, 第 1747-1756 页。[60] A. Van den Oord、N. Kalchbrenner、L. Espeholt、O. Vinyals、A. Graves et al., “使用 pixelcnn 解码器进行条件图像生成”, *Advances in neural information processing systems*, 第 29 卷, 2016 年。[61] T. Salimans、A. Karpathy、X. Chen 和 D. P. Kingma, “Pixelcnn++: 使用离散化逻辑混合似然和其他修改改进 pixelcnn”, *arXiv preprint arXiv:1701.05517*, 2017 年。[62] D. Rezende 和 S. Mohamed, “使用正则化流进行变分推理”, 载于 *International conference on machine learning*. PMLR, 2015 年, 第 1530-1538 页。[63] G. Papamakarios、T. Pavlakou 和 I. Murray, “用于密度估计的掩蔽自回归流”, *Advances in neural information processing systems*, 第 30 卷, 2017 年。[64] A. Creswell、T. White、V. Dumoulin、K. Arulkumaran、B. Sengupta 和 A. A. Bharath, “生成对抗网络: 概述”, *IEEE signal processing magazine*, 第 35 卷, 第 1 期, 2 017 年, 第 53–65 页, 2018 年。[65] T. Karras、T. Aila、S. Laine 和 J. Lehtinen, “渐进式增长 gans 以提高质量、稳定性和变化”, *arXiv preprint arXiv:1710.10196*, 2017 年。[66] J. Zhao、M. Mathieu 和 Y. LeCun, “基于能量的生成对抗网络”, *arXiv preprint arXiv:1609.03126*, 2016 年。[67] A. Dosovitskiy 和 T. Brox, “基于深度网络的具有感知相似性指标的图像生成”, *Advances in neural information processing systems*, 第 29 卷, 2016 年。[68] X. Yu 和 F. Porikli, “通过判别生成网络对人脸图像进行超分辨”, *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V*. Springer, 2016 年, 第 318-333 页。[69] J. Johnson、A. Alahi 和 L. Fei-Fei, “感知损失用于实时风格转换和超分辨率”, 载于 *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*. Springer, 2016 年, 第 694-711 页。[70] J. Bruna、P. Sprechmann 和 Y. LeCun, “具有深度卷积充分统计的超分辨率”, *arXiv preprint arXiv:1511.05666*, 2015 年。[71] W. Zhang、Y. Liu、C. Dong 和 Y. Qiao, “Ranksrgan: 带有排序器的生成对抗网络, 用于图像超分辨率”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019 年, 第 96-3105 页。[72] X. Wang、L. Xie、C. Dong 和 Y. Shan, “Real-esrgan: 使用纯合成数据训练真实世界的盲超分辨率”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 1905-1914 页。[73] K. Zhang、J. Liang、L. Van Gool 和 R. Timofte, “设计深度盲图像超分辨率的实用退化模型”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 4791-4800 页。[74] E. Schonfeld、B. Schiele 和 A. Khoreva, “基于 u-net 的生成对抗网络鉴别器”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020 年, 第 8207-8216 页。[75] P. Isola、J.-Y. Zhu、T. Zhou 和 A. A. Efros, “使用条件对抗网络进行图像到图像翻译”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 1125-1134 页。[76] K. Simonyan 和 A. Zisserman, “用于大规模图像识别的超深卷积网络”, *arXiv preprint arXiv:1409.1556*, 2014 年。[77] T. Karras、S. Laine、M. Aittala、J. Hellsten、J. Lehtinen 和 T. Aila, “分析和提高 stylegan 的图像质量”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020 年, 第 8110-8119 页。[78] Z. Luo、Y. Huang、S. Li、L. Wang 和 T. Tan, “学习盲图像超分辨率的退化分布”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 6063-6072 页。[79] X. Xu、P. Wei、W. Chen、Y. Liu、M. Mao、L. Lin 和 G. Li, “跨设备真实世界图像超分辨率的双重对抗适应”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 5667-5676 页。[80] Y. Wang、F. P. Erizzi、B. McWilliams、A. Sorkine-Hornung、O. Sorkine-Hornung 和 C. Schroers, “一种完全渐进式的单图像超分辨率方法”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018 年, 第 864-873 页。[81] J. Sohl-Dickstein、E. Weiss、N. Maheswaranathan 和 S. Ganguli, “使用非平衡热力学的深度无监督学习”, 载于 *International Conference on Machine Learning*. PMLR, 2015 年, 第 2256-2265 页。[82] J. Ho、A. Jain 和 P. A. Beel, “去噪扩散

- 概率模型”, *Advances in Neural Information Processing Systems*, 第 33 卷, 第 6840-6851 页, 2020 年。[83] Y. Song 和 S. Ermon, “通过估计数据分布梯度进行生成建模”, *Advances in neural information processing systems*, 第 32 卷, 2019 年。[84] Y. Song、J. Sohl-Dickstein、D. P. Kingma、A. Kumar、S. Ermon 和 B. Poole, “通过随机微分方程进行基于分数的生成建模”, *arXiv preprint arXiv:2011.13456*, 2020 年。[85] J. Song、C. Meng 和 S. Ermon, “去噪扩散隐式模型”, *arXiv preprint arXiv:2010.02502*, 2020 年。[86] Q. Zhang、J. Song、X. Huang、Y. Chen 和 M.-Y. Liu, “Diff collage: 使用扩散模型并行生成大内容”, *arXiv preprint arXiv:2303.17076*, 2023 年。[87] V. Singh、S. Jandial、A. Chopra、S. Ramesh、B. Krishna-murthy 和 V. N. Balasubramanian, “关于使用扩散模型调节输入噪声以生成受控图像”, *arXiv preprint arXiv:2205.03859*, 2022 年。[88] V. Popov、I. Vovk、V. Gogoryan、T. Sadekova 和 M. Kudinov, “Grad-tts: 用于文本转语音的扩散概率模型”, 载于 *International Conference on Machine Learning*. PMLR, 2021 年, 第 8599-8608 页。[89] A. Levkovitch、E. Nachmani 和 L. Wolf, “零样本语音条件用于去噪扩散 tts 模型”, *arXiv preprint arXiv:2206.02246*, 2022 年。[90] U. Singer、A. Polyak、T. Hayes、X. Yin、J. An、S. Zhang、Q. Hu、H. Yang、O. Ashual 和 O. Gafni et al., “制作视频: 无需文本视频数据的文本转视频生成”, *arXiv preprint arXiv:2209.14792*, 2022 年。[91] C.-H. Lin、J. Gao、L. Tang、T. Takikawa、X. Zeng、X. Huang、K. Kreis、S. Fidler、M.-Y. Liu 和 T.-Y. Lin, “Magic3d: 高分辨率文本到 3D 内容创建”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 300-309 页。[92] L. Zhang 和 M. Agrawala, “为文本到图像扩散模型添加条件控制”, *arXiv preprint arXiv:2302.05543*, 2023 年。[93] X. Xu、Z. Wang、E. Zhang、K. Wang 和 H. Shi, “多功能扩散: 文本、图像和变化尽在一个扩散模型中”, *arXiv preprint arXiv:2211.08332*, 2022 年。[94] S. Gu、D. Chen、J. Bao、F. Wen、B. Zhang、D. Chen、L. Yuan 和 B. Guo, “用于文本到图像合成的矢量量化扩散模型”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 10 页 696–10 706。[95] C. Mou、X. Wang、L. Xie、J. Zhang、Z. Qi、Y. Shan 和 X. Qie, “T2i 适配器: 学习适配器以挖掘文本到图像扩散模型的更多可控能力”, *arXiv preprint arXiv:2302.08453*, 2023 年。[96] Ž. B. abnik、P. Peer 和 V. Štruc, “Diffiqa: 使用去噪扩散概率模型进行人脸图像质量评估”, *arXiv preprint arXiv:2305.05768*, 2023 年。[97] A. Blattmann、R. Rombach、H. Ling、T. Dockhorn、S. W. Kim、S. Fidler 和 K. Kreis, “对齐你的潜在因素: 使用潜在扩散模型进行高分辨率视频合成”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 22 563-22 575 页。[98] R. Rombach、A. Blattmann、D. Lorenz、P. Esser 和 B. Ommer, “使用潜在扩散模型的高分辨率图像合成”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 10 684-10 695 页。[99] A. Ramesh、P. Dhariwal、A. Nichol、C. Chu 和 M. Chen, “使用剪辑潜在模型的分层文本条件图像生成”, *arXiv preprint arXiv:2204.06125*, 2022 年。[100] C. Saharia、J. Ho、W. Chan、T. Salimans、D. J. Fleet 和 M. Norouzi, “通过迭代细化实现图像超分辨率”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022 年。[101] H. Li、Y. Yang、M. Chang、S. Chen、H. Feng、Z. Xu、Q. Li 和 Y. Chen, “Srdiff: 具有扩散概率模型的单图像超分辨率”, *Neurocomputing*, 卷。479, 第 47-59 页, 2022 年。[102] A. Niu、K. Zhang、T. X. Pham、J. Sun、Y. Zhu、I. S. Kweon 和 Y. Zhang, “Cdpmsr: 单图像超分辨率的条件扩散概率模型”, *arXiv preprint arXiv:2302.12831*, 2023 年。[103] M. Dos Santos、R. Laroca、R. O. Ribeiro、J. Neves、H. Proença 和 D. Menotti, “使用随机微分方程实现人脸超分辨率”, 载于 *2022 35th SIBGRAP Conference on Graphics, Patterns and Images (SIBGRAP)*, 第 1 卷。IEEE, 2022 年, 第 216-221 页。[104] O. Özdenizci 和 R. Legenstein, “使用基于块的去噪扩散模型在恶劣天气条件下恢复视觉”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023 年。[105] J. Whang、M. Delbracio、H. Talebi、C. Saharia、A. G. Dimakis 和 P. Milafar, “通过随机细化去模糊”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 16 293-16 303 页。[106] Z. Luo、F. K. Gustafsson、Z. Zhao、J. Sjölund 和 T. B. Schön, “融合: 使用潜在空间扩散模型实现大尺寸逼真图像恢复”, *arXiv preprint arXiv:2304.08291*, 2023 年。[107] D. Zhou、Z. Yang 和 Y. Yang, “用于低光图像增强的金字塔扩散模型”, *arXiv preprint arXiv:2305.10028*, 2023 年。[108] M. A. Chan、S. I. Young 和 C. A. Metzler, “Sud²: 用于图像重建的去噪扩散模型监督”, *arXiv preprint arXiv:2303.09642*, 2023 年。[109] Y. Miao、L. Zhang、L. Zhang 和 D. Tao, “Dds2m: 用于高光谱图像恢复的自监督去噪扩散空谱模型”, *arXiv preprint arXiv:2303.06682*, 2023 年。[110] B. Kavar、G. Vaksmann 和 M. Elad, “Snips: 随机求解噪声逆问题”, *Advances in Neural Information Processing Systems*, 卷。34, 第 21 757–21 769 页, 2021 年。[111] B. Kavar、J. Song、S. Ermon 和 M. Elad, “使用去噪扩散恢复模型进行 JPEG 伪影校正”, *arXiv preprint arXiv:2209.11888*, 2022 年。[112] B. Kavar、M. Elad、S. Ermon 和 J. Song, “去噪扩散恢复模型”, *arXiv preprint arXiv:2201.11793*, 2022 年。

- [113] Y. Wang、J. Yu 和 J. Zhang, “使用去噪扩散零空间模型进行零样本图像恢复”, *arXiv preprint arXiv:2212.00490*, 2022 年。[114] H. Chung、J. Kim、M. T. Mccann、M. L. Klasky 和 J. C. Ye, “用于一般噪声逆问题的扩散后验采样”, *arXiv preprint arXiv:2209.14687*, 2022 年。[115] Y. Song、L. Shen、L. Xing 和 S. Ermon, “使用基于分数的生成模型解决医学成像中的逆问题”, *arXiv preprint arXiv:2111.08005*, 2021 年。[116] H. Liu、J. Xing、M. Xie、C. Li 和 T.-T. Wong, “通过搭载模型改进基于扩散的图像着色”, *arXiv preprint arXiv:2304.11105*, 2023 年。[117] S. Abu-Hussein、T. Tirer 和 R. Giryes, “Adir: 用于图像重建的自适应扩散”, *arXiv preprint arXiv:2212.03221*, 2022 年。[118] H. Sahak、D. Watson、C. Saharia 和 D. Fleet, “用于野外稳健图像超分辨率的去噪扩散概率模型”, *arXiv preprint arXiv:2302.07864*, 2023 年。[119] N. Murata、K. Saito、C.-H. Lai、Y. Takida、T. Uesaka、Y. Mitsufuji 和 S. Ermon, “Gibbsddrm: 一种用于解决盲逆问题与去噪扩散恢复的部分折叠吉布斯采样器”, *arXiv preprint arXiv:2301.12686*, 2023 年。[120] H. Chung、J. Kim、S. Kim 和 J. C. Ye, “用于盲逆问题的算子和图像的并行扩散模型”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 6059-6069 页。[121] J. Wang、Z. Yue、S. Zhou、K. C. Chan 和 C. C. Loy, “利用扩散先验实现真实世界图像超分辨率”, *arXiv preprint arXiv:2305.07015*, 2023 年。[122] M. Wei、Y. Shen、Y. Wang、H. Xie 和 F. L. Wang, “Raindiffusion: 当无监督学习遇到扩散模型时, 实现真实世界图像去雨”, *arXiv preprint arXiv:2301.09430*, 2023 年。[123] Z. Wang、Z. Zhang、X. Zhang、H. Zheng、M. Zhou、Y. Zhang 和 Y. Wang, “Dr2: 基于扩散的稳健退化去除器, 用于盲人脸修复”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 1704-1713 页。[124] L. Dinh、D. Krueger 和 Y. Bengio, “Nice: 非线性独立成分估计”, *arXiv preprint arXiv:1410.8516*, 2014 年。[125] J. Ho、X. Chen、A. Srinivas、Y. Duan 和 P. Abbeel, “Flow++: 通过变分反量化和架构设计改进基于流的生成模型”, 载于 *International Conference on Machine Learning*. PMLR, 2019 年, 第 2722-2730 页。[126] R. T. Chen、Y. Rubanova、J. Bettencourt 和 D. K. Duvenaud, “神经常微分方程”, *Advances in neural information processing systems*, 第 31, 2018 年。[127] W. Grathwohl、R. T. Chen、J. Bettencourt、I. Sutskever 和 D. Duvenaud, “Ffjord: 可扩展可逆生成模型的自由形式连续动力学”, *arXiv preprint arXiv:1810.01367*, 2018 年。[128] L. Maaløe、M. Fraccaro、V. Liévin 和 O. Winther, “Biva: 用于机器学习的非常深层的潜在变量层次结构”, 2018 年。
- 生成建模”, *Advances in neural information processing systems*, 第 32 卷, 2019 年。[129] D. Lee、C. Kim、S. Kim、M. Cho 和 W.-S. Han, “使用残差量化的自回归图像生成”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 11523-11532 页。[130] T. Karras、S. Laine 和 T. Aila, “基于风格的生成对抗网络生成器架构”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019 年, 第 4401-4410 页。[131] D. Berthelot、T. Schumm 和 L. Metz, “开始: 边界平衡生成对抗网络”, *arXiv preprint arXiv:1703.10717*, 2017 年。[132] J. Li、X. Liang、Y. Wei、T. Xu、J. Feng 和 S. Yan, “用于小物体检测的感知生成对抗网络”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 1222-1230 页。[133] Y. Bai、Y. Zhang、M. Ding 和 B. Ghanem, “Sod-mtgan: 通过多任务生成对抗网络进行小物体检测”, *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018 年, 第 206-221 页。[134] H. Zhang、I. Goodfellow、D. Metaxas 和 A. Odena, “自注意生成对抗网络”, 载于 *International conference on machine learning*. PMLR, 2019 年, 第 7354-7363 页。[135] J.-Y. Zhu、T. Park、P. Isola 和 A. Efros, “使用循环一致对抗网络进行非配对图像到图像转换”, 载于 *Proceedings of the IEEE international conference on computer vision*, 2017 年, 第 2223-2232 页。[136] J. Kim、M. Kim、H. Kang 和 K. Lee, “U-gat-it: 用于图像到图像转换的具有自适应层实例规范化的无监督生成注意力网络”, *arXiv preprint arXiv:1907.10830*, 2019 年。[137] B. Zhang、S. Gu、B. Zhang、J. Bao、D. Chen、F. Wen、Y. Wang 和 B. Guo, “Stylewin: 用于高分辨率图像生成的基于 Transformer 的 GAN”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022 年, 第 11 页 304-11 314。[138] G. Huang 和 A. H. Jafari, “增强平衡 GAN: 少数类图像生成”, *Neural computing and applications*, 第 35 卷, 第 7 期, 第 5145-5154 页, 2023 年。[139] G. Parisi, “相关函数和计算机模拟”, *Nuclear Physics B*, 第 180 卷, 第 3 期, 第 378-384 页, 1981 年。[140] U. Grenander 和 M. I. Miller, “复杂系统中的知识表示”, *Journal of the Royal Statistical Society: Series B (Methodological)*, 第 56 卷, 第 4 期, 第 549-581 页, 1994 年。[141] J. Sohl-Dickstein、E. Weiss、N. Maheswaranathan 和 S. Ganguli, “使用非平衡热力学的深度无监督学习”, 载于 *International Conference on Machine Learning*. PMLR, 2015 年, 第 2256-2265 页。[142] A. Hyvärinen, “分数匹配的一些扩展”, *Computational statistics & data analysis*, 第 51 卷, 第 5 期, 第 2499-2512 页, 2007 年。[143] A. Hyvärinen 和 P. Dayan, “非平衡热力学的估计”, 载于 *Conference on Machine Learning*.

通过分数匹配实现标准化统计模型。”

Journal of Machine Learning Research, 第 6 卷, 第 4 期, 2005 年。[144] A.Q. Nichol 和 P. Dhariwal, “改进的去噪扩散概率模型”, 载于 *International Conference on Machine Learning*. PMLR, 2021 年, 第 8162-8171 页。[145] Y. Song 和 S. Ermon, “改进的基于分数的生成模型训练技术”, *Advances in neural information processing systems*, 第 33 卷, 第 12 438-12 448 页, 2020 年。[146] D. Kingma、T. Salimans、B. Poole 和 J. Ho, “变分扩散模型”, *Advances in neural information processing systems*, 第 34 卷, 第 21 696–21 707 页, 2021 年。[147] M. W. Lam、J. Wang、R. Huang、D. Su 和 D. Yu, “双边去噪扩散模型”, *arXiv preprint arXiv:2108.11514*, 2021 年。[148] T. Chen, “论噪声调度对扩散模型的重要性”, *arXiv preprint arXiv:2301.10972*, 2023 年。[149] F. Bao、C. Li、J. Zhu 和 B. Zhang, “Analytic-dpm: 扩散概率模型中最优逆方差的分析估计”, *arXiv preprint arXiv:2201.06503*, 2022 年。[150] A. Jolicœur-Martineau、R. Piché-Taillefer、R. T. d. Combes 和 I. Mitliagkas, “对抗性分数匹配和改进的图像生成采样”, *arXiv preprint arXiv:2009.05475*, 2020 年。[151] C. Lu、Y. Zhou、F. Bao、J. Chen、C. Li 和 J. Zhu, “Dpm-solver: 一种用于扩散概率模型采样的快速求解器, 大约需要 10 个步骤”, *arXiv preprint arXiv:2206.00927*, 2022 年。[152] Q. Zhang 和 Y. Chen, “使用指数积分器对扩散模型进行快速采样”, *arXiv preprint arXiv:2204.13902*, 2022 年。[153] C. Lu、Y. Zhou、F. Bao、J. Chen、C. Li 和 J. Zhu, “Dpm-solver++: 用于扩散概率模型引导采样的快速求解器”, *arXiv preprint arXiv:2211.01095*, 2022 年。[154] T. Karras、M. Aittala、T. Aila 和 S. Laine, “阐明基于扩散的生成模型的设计空间”, *arXiv preprint arXiv:2206.00364*, 2022 年。[155] Z. Lyu、X. Xu、C. Yang、D. Lin 和 B. Dai, “通过提前停止扩散过程加速扩散模型”, *arXiv preprint arXiv:2205.12524*, 2022 年。[156] H. Zheng、P. He、W. Chen 和 M. Zhou, “截断扩散概率模型”, *stat*, 第 1050 卷, 第 7, 2022 年。[157] E. Luhman 和 T. Luhman, “迭代生成模型中的知识提炼以提高采样速度”, *arXiv preprint arXiv:2101.02388*, 2021 年。[158] C. Meng、R. Rombach、R. Gao、D. Kingma、S. Ermon、J. Ho 和 T. Salimans, “论引导扩散模型的提炼”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 14 297-14 306 页。[159] T. Salimans 和 J. Ho, “用于快速采样扩散模型的渐进式提炼”, *arXiv preprint arXiv:2202.00512*, 2022 年。[160] Z. Zhou 和 S. Tulsiani, “S parsefusion: 提炼视图条件扩散以进行 3d 重建”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 12 588–12 597。

[161] L. Zhang 和 M. Agrawala, “为文本到图像的扩散模型添加条件控制”, *arXiv preprint arXiv:2302.05543*, 2023 年。[162] H. Zheng、W. Nie、A. Vahdat、K. Azizzadenesheli 和 A. Anandkumar, “通过算子学习快速采样扩散模型”, *arXiv preprint arXiv:2211.13449*, 2022 年。[163] R. Huang、M. W. Lam、J. Wang、D. Su、D. Yu、Y. Ren 和 Z. Zhao, “Fastdiff: 一种用于高质量语音合成的快速条件扩散模型”, *arXiv preprint arXiv:2204.09934*, 2022 年。[164] Z. Wang、Y. Jiang、H. Zheng、P. Wang、P. He、Z. Wang、W. Chen 和 M. Zhou, “块扩散: 更快、更数据高效的扩散模型训练”, *arXiv preprint arXiv:2304.12526*, 2023 年。[165] P. Dhariwal 和 A. Nichol, “扩散模型在图像合成方面击败了 GAN”, *Advances in Neural Information Processing Systems*, 第 34 卷, 第 8780-8794 页, 2021 年。[166] H. Luo、L. Ji、B. Shi、H. Huang、N. Duan、T. Li、J. Li、T. Bharti 和 M. Zhou, “Univ1: 用于多模态理解和生成的统一视频和语言预训练模型”, *arXiv preprint arXiv:2002.06353*, 2020 年。[167] C. Sun、A. Myers、C. Vondrick、K. Murphy 和 C. Schmid, “Videobert: 用于视频和语言表示学习的联合模型”, *Proceedings of the IEEE/CVF international conference on computer vision*, 2019 年, 第 7464-7473 页。[168] J. Lu、D. Batra、D. Parikh 和 S. Lee, “Vilbert: 用于视觉和语言任务的预训练任务无关的视觉语言学表征”, *Advances in neural information processing systems*, 第 32 卷, 2019 年。[169] J. Yang、J. Duan、S. Tran、Y. Xu、S. Chanda、L. Chen、B. Zeng、T. Chilimbi 和 J. Huang, “基于三重对比学习的视觉语言预训练”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 15 671-15 680 页。[170] B. Shi、W.-N. Hsu、K. Lakhota 和 A. Mohamed, “通过掩蔽多模态聚类预测学习视听语音表示”, *arXiv preprint arXiv:2201.02184*, 2022 年。[171] C. Sun、F. Baradel、K. Murphy 和 C. Schmid, “使用对比双向变压器学习视频表示”, *arXiv preprint arXiv:1906.05743*, 2019 年。[172] H. Zhang、P. Zhang、X. Hu、Y.-C. Chen、L. Li、X. Dai、L. Wang、L. Yuan、J.-N. Hwang 和 J. Gao, “Glipv2: 统一定位和视觉语言理解”, *Advances in Neural Information Processing Systems*, 卷. 35, 第 36 067–36 080 页, 2022 年。[173] A. Nagrani、S. Yang、A. Arnab、A. Jansen、C. Schmid 和 C. Sun, “多模态融合的注意力瓶颈”, *Advances in Neural Information Processing Systems*, 卷 34, 第 14 200–14 213 页, 2021 年。[174] F. Bao、S. Nie、K. Xu、Y. Cao、C. Li、H. Su 和 J. Zhu, “一切都值得一说: 扩散模型的 vit 主干”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 22 669–22 679 页。[175] S. Sheynin、O. Ashual、A. Polyak、U. Singer、O. Gafni、E. Nachmani 和 Y. Taigman, “Knn-diffusion: 图像

- 通过大规模检索生成”, *arXiv preprint arXiv:2204.02849*, 2022 年。[176] Z. Tang、S. Gu、J. Bao、D. Chen 和 F. Wen, “改进的矢量量化扩散模型”, *arXiv preprint arXiv:2205.16007*, 2022 年。[177] X. Yang, S.-M. Shih、Y. Fu、X. Zhao 和 S. Ji, “您的 vit 秘密地是一种混合判别-生成扩散模型”, *arXiv preprint arXiv:2208.07791*, 2022 年。[178] S. Gu、D. Chen、J. Bao、F. Wen、B. Zhang、D. Chen、L. Yuan 和 B. Guo, “用于文本到图像合成的矢量量化扩散模型”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 10 696-10 706 页。[179] R. Li、W. Li、Y. Yang、H. Wei、J. Jiang 和 Q. Bai, “Swinv2-imagen: 用于文本到图像生成的分层视觉变换器扩散模型”, *arXiv preprint arXiv:2210.09549*, 2022 年。[180] S. Chai、L. Zhuang 和 F. Yan, “Layoutdm: 基于 Transformer 的布局生成扩散模型”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 18 349-18 358 页。[181] S. Pan、T. Wang、R. L. Qiu、M. Axente、C.-W. Chang、J. Peng、A. B. Patel、J. Shelton、S. A. Patel 和 J. Roper *et al.*, “使用基于 Transformer 的去噪扩散概率模型进行二维医学图像合成”, *Physics in Medicine & Biology*, 第 68 卷, 第 10 期, 第 105004, 2023 年。[182] A. Dosovitskiy、L. Beyer、A. Kolesnikov、D. Weisenborn、X. Zhai、T. Unterthiner、M. Dehghani、M. Minderer、G. Heigold、S. Gelly *et al.*, “一张图像价值 16x16 个单词: 用于大规模图像识别的 Transformers”, *arXiv preprint arXiv:2010.11929*, 2020 年。[183] X. Liu、D. H. Park、S. Azadi、G. Zhang、A. Chopikyan、Y. Hu、H. Shi、A. Rohrbach 和 T. Darrell, “更多免费控制! 具有语义扩散指导的图像合成”, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023 年, 第 2 89-299 页。[184] J. Ho 和 T. Salimans, “无分类器扩散指导”, *arXiv preprint arXiv:2207.12598*, 2022 年。[185] A. Nichol、P. Dhariwal、A. Ramesh、P. Shyam、P. Mishkin、B. McGrew、I. Sutskever 和 M. Chen, “Glide: 通过文本引导的扩散模型实现逼真的图像生成和编辑”, *arXiv preprint arXiv:2112.10741*, 2021 年。[186] N. Ruiz、Y. Li、V. Jampani、Y. Pritch、M. Rubinstein 和 K. Aberman, “Dream booth: 微调文本到图像扩散模型以实现主体驱动的生成”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 22 500-22 510 页。[187] C. Saharia、W. Chan、S. Saxena、L. Li、J. Whang、E. L. Denton、K. Ghasemipour、R. Gontijo Lopes、B. Karagol Ayan 和 T. Salimans *et al.*, “具有深度语言理解的真实感文本到图像扩散模型”, *Advances in Neural Information Processing Systems*, 第 35 卷, 第 36 479–36 494 页, 2022 年。[188] G. Kim、T. Kwon 和 J. C. Ye, “Diffusionclip: 用于稳健图像处理的文本引导扩散模型”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 2426–2435。[189] O. Avrahami、T. Hayes、O. Gafni、S. Gupta、Y. Taigman、D. Parikh、D. Lischinski、O. Fried 和 X. Yin, “Spatext: 可控图像生成的空间文本表示”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 18 370–18 380 页。[190] A. Ramesh、M. Pavlov、G. Goh、S. Gray、C. Voss、A. Radford、M. Chen 和 I. Sutskever, “零样本文本到图像生成”, 载于 *International Conference on Machine Learning*. PMLR, 2021 年, 第 8821–8831 页。[191] Y. Jin、W. Yang、W. Ye、Y. Yuan 和 R. T. Tan, “Shadowdiffusion: 使用分类器驱动的注意力和结构保留进行基于扩散的阴影去除”, *arXiv preprint arXiv:2211.08089*, 2022 年。[192] “Diffbfr: 面向盲人脸修复的引导扩散模型”, *arXiv preprint arXiv:2305.04517*, 2023 年。[193] H. Jiang、A. Luo、S. Han、H. Fan 和 S. Liu, “使用基于小波的扩散模型进行低光图像增强”, *arXiv preprint arXiv:2306.00306*, 2023 年。[194] Z. Chen、Y. Zhang、D. Liu、B. Xia、J. Gu、L. Kong 和 X. Yuan, “用于逼真图像去模糊的分层集成扩散模型”, *arXiv preprint arXiv:2305.12966*, 2023 年。[195] J. Choi、S. Kim、Y. Jeong、Y. Gwon 和 S. Yoon, “Ilvr: 用于去噪扩散概率模型的条件方法”, *arXiv preprint arXiv:2108.02938*, 2021 年。[196] B. Fei、Z. Lyu、L. Pan、J. Zhang、W. Yang、T. Luo、B. Zhang 和 B. Dai, “用于统一图像恢复和增强的生成扩散先验”, *arXiv preprint arXiv:2304.01247*, 2023 年。[197] J. Song、A. Vahdat、M. Mardani 和 J. Kautz, “用于逆问题的伪逆引导扩散模型”, 载于 *International Conference on Learning Representations*, 2023 年。[198] Y. Zhu、K. Zhang、J. Liang、J. Cao、B. Wen、R. Timofte 和 L. Van Gool, “用于即插即用图像恢复的去噪扩散模型”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 1219-1229 页。[199] M. Mardani、J. Song、J. Kautz 和 A. Vahdat, “用扩散模型解决逆问题的变分视角”, *arXiv preprint arXiv:2305.04391*, 2023 年。[200] Z. Fabian、B. Tinaz 和 M. Soltanolkotabi, “Diracdiffusion: 确保数据一致性的去噪和增量重建”, *arXiv preprint arXiv:2303.14353*, 2023 年。[201] Y. Yuan、S. Liu、J. Zhang、Y. Zhang、C. Dong 和 L. Lin, “使用循环生成对抗网络的无监督图像超分辨率”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018 年, 第 701-710 页。[202] P. Li、Z. Li、X. Pang、H. Wang、W. Lin 和 W. Wu, “用于生成超分辨率 cta 图像的多尺度残差去噪 GAN 模型”, *Journal of Ambient Intelligence and Humanized Computing*, 第 1-10 页, 2022 年。[203] T. Yang、P. Ren、X. Xie 和 L. Zhang, “用于盲人脸修复的 GAN 先验嵌入网络

- wild”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021 年, 第 672-681 页。[204] K. C. Chan, X. Wang, X. Xu, J. Gu 和 C. C. Loy, “Glean: 用于大因子图像超分辨率的生成潜在库”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021 年, 第 14 245-14 254 页。[205] D. Roich, R. Mokady, A. H. Bermato 和 D. Cohen-Or, “基于潜在技术的真实图像编辑的关键调整”, *ACM Transactions on Graphics (TOG)*, 第 42 卷, 第 1, 第 1-13 页, 2022 年。[206] X. Pan, X. Zhan, B. Dai, D. Lin, C. C. Loy 和 P. Luo, “利用深度生成先验进行多功能图像恢复和处理”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 第 44 卷, 第 11 期, 第 7474-7489 页, 2021 年。[207] H. Tampubolon, A. Setyoko 和 F. Purnamasari, “Snpe-srgan: 使用 snpe 框架在移动设备上实现单图像超分辨率的轻量级生成对抗网络”, 载于 *Journal of Physics: Conference Series*, 第 1898 卷, 第 1 期。IOP Publishing, 2021 年, 第 012038。[208] T.M.Dinh, A.T.Tran, R.Nguyen 和 B.-S.Hua, “Hyperinverter: 通过超网络改进 stylegan 反演”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 11 389-11 398 页。[209] J.Wang, W.Zhou, G.-J.Qi, Z.Fu, Q.Tian 和 H.Li, “用于无监督图像合成和表示学习的转换 gan”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020 年, 第 472-481 页。[210] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla 和 M. Bernstein *et al.*, “Imagenet 大规模视觉识别挑战赛”, *International journal of computer vision*, 第 115 卷, 第 211-252 页, 2015 年。[211] T. Karras, S. Laine 和 T. Aila, “基于风格的生成对抗网络生成器架构”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019 年, 第 4401-4410 页。[212] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte 和 L. Van Gool, “重绘: 使用去噪扩散概率模型进行修复”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022 年 6 月, 第 11 461-11 471 页。[213] H. Chung, B. Sim 和 J. C. Ye, “更接近扩散更快: 通过随机收缩加速逆问题的条件扩散模型”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 12 413-12 422 页。[214] Z. Zhang, Z. Zhao 和 Z. Lin, “从预训练扩散概率模型进行无监督表示学习”, *Advances in Neural Information Processing Systems*, 第 35 卷, 第 22 117–22 130 页, 2022 年。[215] B. T. Feng, J. Smith, M. Rubinstein, H. Chang, K. L. Bouman 和 W. T. Freeman, “基于评分的扩散模型作为逆成像的原则先验”, *arXiv preprint arXiv:2304.11751*, 2023 年。[216] G. Zhang, J. Ji, Y. Zhang, M. Yu, T. Jaakkola 和 S. Chang, “使用去噪扩散隐式模型实现相干图像修复”, *arXiv preprint arXiv:2304.03322*, 2023 年。[217] S. Shang, Z. Shan, G. Liu 和 J. Zhang, “Resdiff: 结合 cnn 和扩散模型实现图像超分辨率”, *arXiv preprint arXiv:2303.08714*, 2023 年。[218] J. Ho, C. Saharia, W. Chan, D. J. Fleet, M. Norouzi 和 T. Salimans, “用于高保真图像生成的级联扩散模型”, *J. Mach. Learn. Res.*, 第 23 卷, 第 47, 第 1-33 页, 2022 年。[219] C. Saharia, W. Chan, H. Chang, C. Lee, J. Ho, T. Salimans, D. Fleet 和 M. Norouzi, “调色板: 图像到图像扩散模型”, *ACM SIGGRAPH 2022 Conference Proceedings*, 2022 年, 第 1-10 页。[220] S. Welker, H. N. Chapman 和 T. Gerkmann, “Driftrec: 将扩散模型应用于盲图像恢复任务”, *arXiv preprint arXiv:2211.06757*, 2022 年。[221] Y. Yin, L. Huang, Y. Liu 和 K. Huang, “Diff-gar: 使用图像到图像扩散模型从生成伪影中进行与模型无关的恢复”, *arXiv preprint arXiv:2210.08573*, 2022 年。[222] M. Ren, M. Delbracio, H. Talebi, G. Gerig 和 P. Milanfar, “使用领域可泛化的扩散模型进行图像去模糊”, *arXiv preprint arXiv:2212.01789*, 2022 年。[223] S. Gao, X. Liu, B. Zeng, S. Xu, Y. Li, X. Luo, J. Liu, X. Zhen 和 B. Zhang, “用于连续超分辨率的隐式扩散模型”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 10 021-10 030 页。[224] B. Xia, Y. Zhang, S. Wang, Y. Wang, X. Wu, Y. Tian, W. Yang 和 L. Van Gool, “Diffir: 用于图像恢复的有效扩散模型”, *arXiv preprint arXiv:2303.09472*, 2023 年。[225] M. Delbracio 和 P. Milanfar, “直接迭代反演: 一种用于图像恢复的去噪扩散替代方法”, *arXiv preprint arXiv:2303.11435*, 2023 年。[226] T. Yang, P. Ren, L. Zhang *et al.*, “合成真实的图像恢复训练对: 一种扩散方法”, *arXiv preprint arXiv:2303.06994*, 2023 年。[227] S. U. Dar, S. Öztürk, Y. Korkmaz, G. Elmas, M. Özbey, A. Güngör 和 T. Çukur, “用于加速 MRI 重建的自适应扩散先验”, *arXiv preprint arXiv:2207.05876*, 2022 年。[228] C. Cao, Z.-X. Cui, S. Liu, D. Liang 和 Y. Zhu, “加速 MRI 的高频空间扩散模型”, *arXiv preprint arXiv:2208.05481*, 2022 年。[229] Y. Xie, M. Yuan, B. Dong 和 Q. Li, “生成图像去噪的扩散模型”, *arXiv preprint arXiv:2302.02398*, 2023 年。[230] L. Guo, C. Wang, W. Yang, S. Huang, Y. Wang, H. Pfister 和 B. Wen, “阴影扩散: 当降级先验遇到扩散模型进行阴影去除时”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 14 049-14 058 页。[231] Z. Luo, F. K. Gustafsson, Z. Zhao, J. Sjölund 和 T. B. Schön, “具有均值恢复随机微分方程”, *arXiv preprint arXiv:2301.11699*, 2023 年。

- [232] Y. Zhang、K. Li、K. Li、L. Wang、B. Zhong 和 Y. Fu, “使用非常深的残差通道注意力网络实现图像超分辨率”, 载于 *Proceedings of the European conference on computer vision (ECCV)*, 2018 年, 第 286-301 页。[233] B. Lim、S. Son、H. Kim、S. Nah 和 K. Mu Lee, “用于单图像超分辨率的增强深度残差网络”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2017 年, 第 136-144 页。[234] Q. Gao 和 H. Shan, “Cocodiff: 一种用于低剂量 CT 图像去噪的上下文条件扩散模型”, 载于 *Developments in X-Ray Tomography XIV*, 第 12242. SPIE, 2022 年。[235] K. Karchev、N. A. Montel、A. Coogan 和 C. Weniger, “具有去噪扩散恢复模型的强透镜源重建”, *arXiv preprint arXiv:2211.04365*, 2022 年。[236] H. Chung、B. Sim、D. Ryu 和 J. C. Ye, “使用流形约束改进逆问题的扩散模型”, *arXiv preprint arXiv:2206.00941*, 2022 年。[237] Z. Yue 和 C. C. Loy, “Difface: 具有扩散误差收缩的盲人脸恢复”, *arXiv preprint arXiv:2212.06512*, 2022 年。[238] H. Sun 和 K. L. Bouman, “深度概率成像: 计算成像的不确定性量化和多模态解决方案表征”, *Proceedings of the AAAI Conference on Artificial Intelligence*, 第 35 卷, 第 3 期, 2021 年, 第 2628-2637 页。[239] H. Sun、K. L. Bouman、P. Tiede、J. J. Wang、S. Blunt 和 D. Mawet, “ α -深度概率推理 (α -dpi): 从系外行星天体测量到黑洞特征提取的有效不确定性量化”, *The Astrophysical Journal*, 第 932 卷, 第 2 期, 第 99, 2022 年。[240] L. Dinh、J. Sohl-Dickstein 和 S. Bengio, “使用真实 nvp 进行密度估计”, *arXiv preprint arXiv:1605.08803*, 2016 年。[241] Y. Wang、J. Ming、X. Jia、J. H. Elder 和 H. Lu, “具有退化感知自适应的盲图像超分辨率”, *Proceedings of the Asian Conference on Computer Vision*, 2022 年, 第 894-910 页。[242] J. Zhang、T. Xu、J. Li、S. Jiang 和 Y. Zhang, “具有真实世界退化建模的遥感图像单图像超分辨率”, *Remote Sensing*, 第 14 卷, 第 12 期, 第 2895, 2022 年。[243] J. Liang、H. Zeng 和 L. Zhang, “用于真实世界图像超分辨率的高效退化自适应网络”, 载于 *Computer Vision—ECCV 2022: 17th European Conference, Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part XVIII*. Springer, 2022 年, 第 574-591 页。[244] K. Purohit、M. Suin、A. Rajagopalan 和 V. N. Bodde ti, “使用失真引导网络进行空间自适应图像恢复”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 2309-2319 页。[245] W. Wang、H. Zhang、Z. Yuan 和 C. Wang, “无监督的真实世界超分辨率: 领域适应视角”, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 4318-4327 页。[246] K. Zhang、J. Liang、L. Van Gool 和 R. Timofte, “设计深度盲图像超分辨率的实用退化模型”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 4791-4800 页。[247] A. Lugmayr、M. Danelljan 和 R. Timofte, “现实世界超分辨率的无监督学习”, 载于 *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, 2019 年, 第 3408-3416 页。[248] Y. Wang、X. Yan、F. L. Wang、H. Xie、W. Yang、M. Wei 和 J. Qin, “Ucl-dehaze: 通过无监督对比学习实现真实世界图像去雾”, *arXiv preprint arXiv:2205.01871*, 2022 年。[249] X. Li、B. Li、X. Jin、C. Lan 和 Z. Chen, “从因果关系角度学习用于图像恢复的畸变不变表示”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 1714-1724 页。[250] D. G. Tzikas、A. C. Likas 和 N. P. Galatsanos, “基于变分贝叶斯稀疏核的盲图像反卷积与学生 t 先验”, *IEEE transactions on image processing*, 第 18 卷, 第 4, 第 753–764 页, 2009 年。[251] L. Huang 和 Y. Xia, “联合模糊核估计和 cnn 用于盲图像恢复”, *Neurocomputing*, 第 396 卷, 第 324–345 页, 2020 年。[252] S. Vasu、V. R. Maligireddy 和 A. Rajagopalan, “非盲去模糊: 使用 cnns 处理内核不确定性”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018 年, 第 3272–3281 页。[253] J. P. Oliveira、M. A. Figueiredo 和 J. M. Bioucas-Dias, “参数模糊估计用于自然图像盲恢复: 线性运动和失焦”, *IEEE Transactions on Image Processing*, 第 23 卷, 第 1, 第 466–477 页, 2013 年。[254] J. Jiang、K. Zhang 和 R. Timofte, “实现灵活的盲 jpeg 伪影去除”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021 年, 第 4997–5006 页。[255] R. Zhou 和 S. Susstrunk, “在真实低分辨率图像上进行核建模超分辨率”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019 年, 第 2433–2443 页。[256] D. A. Van Dyk 和 T. Park, “部分折叠的吉布斯采样器: 理论与方法”, *Journal of the American Statistical Association*, 第 103 卷, 第 482, 第 790-796 页, 2008 年。[257] A. Anoosheh、E. Agustsson、R. Timofte 和 L. Van Gool, “Combogan: 图像域翻译的无约束可扩展性”, *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018 年, 第 783-790 页。[258] J. Lin、Y. Wang、T. He 和 Z. Chen, “学习迁移: 无监督元域翻译”, *arXiv preprint arXiv:1906.00181*, 2019 年。[259] C. Chu 和 R. Wang, “神经机器翻译领域适应调查”, *arXiv preprint arXiv:1806.00258*, 2018 年。[260] D. Saunders, “神经机器翻译的领域适应和多领域适应: 调查”, *Journal of Artificial Intelligence Research*, 第 75 卷, 第 351-424 页, 2022 年。

- [261] A. Lugmayr、M. Danelljan 和 R. Timofte, “现实世界超分辨率的无监督学习”, 参见 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). IEEE, 2019 年, 第 3408–3416 页。[262] Y. Wang、L. Wang、J. Yang、W. An 和 Y. Guo, “Flickr1024: 立体图像超分辨率的大型数据集”, *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019 年, 第 0–0 页。[263] O. Russakovsky、J. Deng、H. Su、J. Krause、S. Satheesh、S. Ma、Z. Huang、A. Karpathy、A. Khosla、M. Bernstein、A. C. Berg 和 L 飞飞, “ImageNet 大规模视觉识别挑战”, *International Journal of Computer Vision (IJCV)*, 卷. 115, 第 3 期, 第 211-252 页, 2015 年。[264] D. Martin、C. Fowlkes、D. Tal 和 J. Malik, “人体分割自然图像数据库及其在评估分割算法和测量生态统计数据”, 载于 *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 第 2 卷。IEEE, 2001 年, 第 416-423 页。[265] Y. Matsui、K. Ito、Y. Aramaki、A. Fujimoto、T. Ogawa、T. Yamasaki 和 K. Aizawa, “使用 manga109 数据集进行基于草图的漫画检索”, *Multimedia Tools and Applications*, 第 76 卷, 第 21 811-21 838 页, 2017 年。[266] J.-B. Huang、A. Singh 和 N. Ahuja, “通过变换后的自我样本实现单幅图像超分辨率”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015 年, 第 5197-5206 页。[267] X. Wang、K. Yu、C. Dong 和 C. C. Loy, “通过深度空间特征变换恢复图像超分辨率中的真实纹理”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 606-615 页。[268] J. Cai、H. Zeng、H. Yong、Z. Cao 和 L. Zhang, “走向现实世界的单图像超分辨率: 新的基准和新模型”, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019 年, 第 37 页。3086–3095。[269] P. Wei、Z. Xi、H. Lu、Z. Zhan、Q. Ye、W. Zuo、以及 L. Lin, “真实世界图像超分辨率的分量分治法”, *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII* 16. Springer, 2020 年, 第 101-117 页。[270] J. Rim、H. Lee、J. Won 和 S. Cho, “用于学习和基准测试的真实世界模糊数据集去模糊算法”, 位于 *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV* 16. Springer, 2020 年, 第 184-201 页。[271] Z. Shen、W. Wang、X. Lu、J. Shen、H. Ling、T. Xu 和 L. Shao, “基于人机感知的运动去模糊”, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019 年, 第 5572–5581 页。[272] J. Wang、X. Li 和 J. Yang, “用于联合学习阴影检测和阴影消除的堆叠条件生成对抗网络”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018 年, 第 1788-1797 页。[273] L. Qu、J. Tian、S. He、Y. Tang 和 R. W. Lau, “Deshad-ownet: 一种用于阴影去除的多上下文嵌入深度网络”, 载于 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017 年, 第 4067-4075 页。[274] Y. Liu、L. Zhu、S. Pei、H. Fu、J. Qin、Q. Zhang、L. Wan 和 W. Feng, “从合成到真实: 利用未标记的真实数据进行图像去雾”, 载于 *Proceedings of the 29th ACM international conference on multimedia*, 2021 年, 第 50-58 页。[275] C. O. Ancuti、C. Ancuti、M. Sbert 和 R. Timofte, “浓雾: 使用浓雾和无雾图像进行图像去雾的基准”, 载于 2019 IEEE international conference on image processing (ICIP). IEEE, 2019 年, 第 1014-1018 页。[276] B. Li、W. Ren、D. Fu、D. Tao、D. Feng、W. Zeng 和 Z. Wang, “单图像去雾及其他技术的基准测试”, *IEEE Transactions on Image Processing*, 第 28 卷, 第 1 期, 第 492-505 页, 2019 年。[277] S. Golodetz、T. Cavallari、N. A. Lord、V. A. Prisacariu、D. W. Murray 和 P. H. Torr, “协同大规模密集 3D 重建与在线跨代理姿势优化”, *IEEE transactions on visualization and computer graphics*, 第 24 卷, 第 11 期, 第 2895-2905 页, 2018 年。[278] Y.-F. Liu、D.-W. Jaw、S.-C. Huang 和 J.-N. Hwang, “Desnownet: 用于除雪的上下文感知深度网络”, *IEEE Transactions on Image Processing*, 第 27 卷, 第 6 期, 第 3064-3073 页, 2018 年。[279] W.-T. Chen、H.-Y. Fang、J.-J. Ding、C.-C. Tsai 和 S.-Y. Kuo, “Jstasr: 基于改进的部分卷积和遮蔽效应去除的联合尺寸和透明度感知除雪算法”, 载于 *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI* 16. Springer, 2020 年, 第 754-770 页。[280] X. Fu、J. Huang、D. Zeng、Y. Huang、X. Ding 和 J. Paisley, “通过深度细节网络从单幅图像中去除雨水”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017 年, 第 3855-3863 页。[281] R. Li、L.-F. Cheong 和 R. T. Tan, “暴雨图像修复: 结合物理模型和条件对抗学习”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019 年, 第 1633-1642 页。[282] T. Wang、X. Yang、K. Xu、S. Chen、Q. Zhang 和 R. W. Lau, “使用高质量真实降雨数据集进行空间感知型单图像去雨”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019 年, 第 12 270–12 279 页。[283] Z. Wang、A. C. Bovik、H. R. Sheikh 和 E. P. Simoncelli, “图像质量评估: 从错误可见性到结构相似性”, *IEEE transactions on image processing*, 第 13 卷, 第 13 期。4, 第 600–612 页, 2004 年。[284] R. Zhang、P. Isola、A. A. Efros、E. Shechtman 和 O. Wang, “深度特征作为感知指标的不合理有效性”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 586–595 页。[285] K. Ding、K. Ma、S. Wang 和 E. P. Simoncelli, “图像质量评估: 统一结构和纹理

- 相似性”, *IEEE transactions on pattern analysis and machine intelligence*, 第 44 卷, 第 5 期, 第 2567-2581 页, 2020 年。[286] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler 和 S. Hochreiter, “通过两个时间尺度更新规则训练的 Gans 收敛到局部纳什均衡”, *Advances in neural information processing systems*, 第 44 卷, 第 5 期, 第 2567-2581 页, 2020 年。30, 2017 年。[287] M. Bińkowski, D. J. Sutherland, M. Arbel 和 A. Gretton, “揭开 md gans 的神秘面纱”, *arXiv preprint arXiv:1801.01401*, 2018 年。[288] A. Mittal, R. Soundararajan 和 A. C. Bovik, “制作“完全盲目的”图像质量分析仪”, *IEEE Signal processing letters*, 第 20 卷, 第 3 期, 第 209-212 页, 2012 年。[289] Y. Blau, R. Mechrez, R. Timofte, T. Michaeli 和 L. Zelnik-Manor, “2018 年 pirm 感知图像超分辨率挑战赛”, *Proceedings of the European Conference on Computer Vision (ECCV) Work-shops*, 2018 年, 第 0-0 页。[290] Z. Wang, E. P. Simoncelli 和 A. C. Bovik, “多尺度结构相似性用于图像质量评估”, *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers*, 2003, 第 2 卷。Ieee, 2003 年, 第 1398-1402 页。[291] K. Simonyan 和 A. Zisserman, “用于大规模图像识别的超深卷积网络”, *arXiv preprint arXiv:1409.1556*, 2014 年。[292] S. Barratt 和 R. Sharma, “关于初始分数的注释”, *arXiv preprint arXiv:1801.01973*, 2018 年。[293] C. Ma, C.-Y. Yang, X. Yang 和 M.-H. Yang, “学习单图像超分辨率的无参考质量度量”, *Computer Vision and Image Understanding*, 第 2 卷。158, 第 1-16 页, 2017 年。[294] Z. Liu, P. Luo, X. Wang 和 X. Tang, “深度学习自然人脸属性”, *Proceedings of the IEEE international conference on computer vision*, 2015 年, 第 3730-3738 页。[295] Y. Zhang, Y. Tian, Y. Kong, B. Zhong 和 Y. Fu, “用于图像超分辨率的残差密集网络”, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018 年, 第 2472-2481 页。[296] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv 和 X. Peng, “针对未知损坏的一体化图像恢复”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022 年, 第 17452-17462 页。[297] X. Li, S. Sun, Z. Zhang 和 Z. Chen, “用于 vvc 帧内编码的多尺度分组密集网络”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020 年, 第 158-159 页。[298] X. Li, X. Jin, J. Lin, S. Liu, Y. Wu, T. Yu, W. Zhou 和 Z. Chen, “学习解缠绕特征表示以实现混合失真图像恢复”, 载于 *Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XXIX 16*. Springer, 2020 年, 第 313-329 页。[299] X. Li, X. Jin, T. Yu, S. Sun, Y. Pang, Z. Zhang 和 Z. Chen, “学习全频区域自适应表示以实现真实图像超分辨率”, 载于 *Proceedings of the AAAI Conference on Artificial Intelligence*, 第 35 卷, 第 3 期, 2021 年, 第 1975-1983 页。[300] X. Li, X. Jin, J. Fu, X. Yu, B. Tong 和 Z. Chen, “通过失真关系引导的迁移学习实现小样本真实图像恢复”, *arXiv preprint arXiv:2111.13078*, 2021 年。[301] B. Li, X. Li, Y. Lu, S. Liu, R. Feng 和 Z. Chen, “Hst: 用于压缩图像超分辨率的分层 swin 变换器”, 载于 *European Conference on Computer Vision*. Springer, 2022 年, 第 651-668 页。[302] Z. Li, H. Li 和 L. Meng, “深度神经网络的模型压缩: 一项调查”, *Computers*, 第 12 卷, 第 3 期, 第 60 页, 2023 年。[303] B.-K. Kim, H.-K. Song, T. Castells 和 S. Choi, “论文本到图像扩散模型的架构压缩”, *arXiv preprint arXiv:2305.15798*, 2023 年。[304] G. Fang, X. Ma 和 X. Wang, “扩散模型的结构修剪”, *arXiv preprint arXiv:2305.10924*, 2023 年。[305] X. Li, L. Lian, Y. Liu, H. Yang, Z. Dong, D. Kang, S. Zhang 和 K. Keutner, “Q 扩散: 量化扩散模型”, *arXiv preprint arXiv:2302.04304*, 2023 年。[306] Y. He, L. Liu, J. Liu, W. Wu, H. Zhou 和 B. Zhuang, “Ptqd: 扩散模型的精确训练后量化”, *arXiv preprint arXiv:2305.10657*, 2023 年。[307] Y. Shang, Z. Yuan, B. Xie, B. Wu 和 Y. Yan, “训练后扩散模型的无数据量化”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 1972-1981 页。[308] C. Wang, Z. Wang, X. Xu, Y. Tang, J. Zhou 和 J. Lu, “面向扩散模型的精确无数据量化”, *arXiv preprint arXiv:2305.18723*, 2023 年。[309] X. Li, B. Li, X. Jin, C. Lan 和 Z. Chen, “从因果关系角度学习用于图像恢复的畸变不变表示”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023 年, 第 1714-1724 页。[310] K. Zhou, Z. Liu, Y. Qiao, T. Xiang 和 C. C. Loy, “领域泛化: 一项调查”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022 年。[311] I. Gulrajani 和 D. Lopez-Paz, “寻找丢失的领域泛化”, *arXiv preprint arXiv:2007.01434*, 2020 年。[312] D. Li, Y. Yang, Y.-Z. Song 和 T. Hospedales, “学习泛化: 领域泛化的元学习”, *Proceedings of the AAAI conference on artificial intelligence*, 第 32 卷, 第 1 期, 第 177-182 页。1, 2018 年。[313] J. Wang, C. Lan, C. Liu, Y. Ouyang, T. Qin, W. Lu, Y. Chen, W. Zeng 和 P. Yu, “推广到未知领域: 领域泛化调查”, *IEEE Transactions on Knowledge and Data Engineering*, 2022 年。[314] Q. Dou, D. Coelho de Castro, K. Kamnitsas 和 B. Glocker, “通过与模型无关的语义特征学习实现领域泛化”, *Advances in neural information processing systems*, 第 32 卷, 2019 年。[315] Z. Wang, M. Loog 和 J. van Gemert, “尊重领域关系: 领域泛化的假设不变性”, *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021 年, 第 9756-9763 页。[316] Y. Li, X. Tian, M. Gong, Y. Liu, T. Liu, K. Zhang, 和 D. Tao, “通过条件不变对抗网络实现深度领域泛化”, 载于 *Proceedings of the European Conference on Computer Vision (ECCV)*,

- 2018年9月。[317] R. Shao、X. Lan、J. Li 和 P. C. Yuen, “用于人脸呈现攻击检测的多对抗性判别性深度域泛化”, 载于 *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019年6月。[318] S. Motiian、M. Piccirilli、D. A. Adjeroh 和 G. Doretto, “统一深度监督域自适应和泛化”, 载于 *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017年10月。[319] S. Zhao、M. Gong、T. Liu、H. Fu 和 D. Tao, “通过熵正则化实现域泛化”, *Advances in Neural Information Processing Systems*, 第33, 第16 096–16 107 页, 2020年。[320] Z. Xu、D. Liu、J. Yang、C. Raffel 和 M. Niethammer, “通过随机卷积进行稳健且可泛化的视觉表征学习”, *arXiv preprint arXiv:2007.13003*, 2020年。[321] K. Zhou、Y. Yang、T. Hospedales 和 T. Xiang, “用于领域泛化的深度领域对抗性图像生成”, 载于 *Proceedings of the AAAI conference on artificial intelligence*, 第34卷, 第07, 2020年, 第13 025-13 032 页。[322] K. Zhou、Y. Yang、Y. Qiao 和 T. Xiang, “使用 mixstyle 进行领域泛化”, *arXiv preprint arXiv:2104.02008*, 2021年。[323] M. Mancini、Z. Akata、E. Ricci 和 B. Caputo, “在未知领域中识别未知类别”, 载于 *European Conference on Computer Vision*. Springer, 2020年, 第466-483 页。[324] B. Wang、M. Lapata 和 I. Titov, “语义解析中领域泛化的元学习”, *arXiv preprint arXiv:2010.11988*, 2020年。[325] Y. Jia、J. Zhang、S. Shan 和 X. Chen, “人脸反欺骗的单边领域泛化”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020年, 第8484-8493 页。[326] M. M. Rahman、C. Fookes、M. Baktashmotlagh 和 S. Sridharan, “关联感知对抗领域自适应与泛化”, *Pattern Recognition*, 第100卷, 第107124, 2020年。[327] I. Albuquerque、J. Monteiro、M. Darvishi、T. H. Falk 和 I. Mitliagkas, “通过分布匹配推广到看不见的域”, *arXiv preprint arXiv:1911.00804*, 2019年。[328] F. Wang 和 H. Liu, “理解对比损失的行为”, *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021年, 第2495-2504 页。[329] X. Wang、W. Wang、Y. Cao、C. Shen 和 T. Huang, “图像用图像说话: 用于情境内视觉学习的通才画家”, *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023年, 第6830-6839 页。[330] Y. Zhang、F. Yu、S. Song、P. Xu、A. Seff、和 J. Xiao, “大规模场景理解挑战: 房间布局估计”, *CVPR Workshop*, 2015年。[331] S. Gu、A. Lugmayr、M. Danelljan、M. Fritsche、J. Lamour、和 R. Timofte, “Div8k: 多样化 8k 分辨率图像数据集”, *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, 2019年, 第3512-3516 页。[332] R. Franzen, “Kodak 无损真彩色图像套件”, *source: http://r0k.us/graphics/kodak*, 第4卷, 第2期, 第9页, 1999年。[333] D. Martin、C. Fowlkes、D. Tal 和 J. Malik, “人类分割自然图像数据库及其在评估分割算法和测量生态统计数据中的应用”, *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 第2卷. IEEE, 2001年, 第416-423 页。[334] L. Zhang、X. Wu、A. Buades 和 X. Li, “通过局部方向插值和非局部自适应阈值进行色彩去马赛克”, *Journal of Electronic imaging*, 第20卷, 第2, 第023 016-023 016 页, 2011年。[335] F. Yu、A. Seff、Y. Zhang、S. Song、T. Funkhouser 和 J. Xiao, “Lsun: 在人类参与的情况下使用深度学习构建大规模图像数据集”, *arXiv preprint arXiv:1506.03365*, 2015年。[336] A. L. López-Ci-fuentes、M. Escudero-Vinolo、J. Bescós 和 Á. García-Martín, “基于语义感知的场景识别”, *Pattern Recognition*, 第102卷, 第107256, 2020年。[337] G. B. Huang、M. Mattar、T. Berg 和 E. Learned-Miller, “野生标记人脸: 用于研究无约束环境下人脸识别的数据库”, 载于 *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*, 2008年。[338] T. Karras、T. Aila、S. Laine 和 J. Lehtinen, “逐步发展 GAN 以提高质量、稳定性和变化性”, *arXiv preprint arXiv:1710.10196*, 2017年。[339] Y. Choi、Y. Uh、J. Yoo 和 J.-W. Ha, “Stargan v2: 适用于多个领域的多样化图像合成”, 载于 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020年, 第8188-8197 页。[340] X. Fu、J. Huang、D. Zeng、Y. Huang、X. Ding 和 J. Paisley, “通过深度细节网络从单幅图像中去除雨水”, 载于 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017年7月。[341] W. Yang、R. T. Tan、J. Feng、J. Liu、Z. Guo 和 S. Yan, “从单幅图像中进行深度联合雨水检测与去除”, 载于 *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017年, 第1357-1366 页。[342] B. Zhou、A. Lapedriza、A. Khosla、A. Oliva 和 A. Torralba, “地点: 用于场景识别的千万级图像数据库”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017年。[343] H. Le 和 D. Samaras, “通过阴影图像分解去除阴影”, 载于 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019年, 第8578-8587 页。[344] J. Zbontar、F. Knoll、A. Sriram、T. Murrell、Z. Huang、M. J. Muckley、A. Defazio、R. Stern、P. Johnson 和 M. Bruno *et al.*, “fastmri: 加速 MRI 的开放数据集和基准”, *arXiv preprint arXiv:1811.08839*, 2018年。

表 8：不同基于扩散模型的图像恢复任务中广泛使用的数据集总结

Task	Dataset	Year	Training	Testing	Short Description
Image Super-resolution	DIV2K [45]	2017	800	100	2K resolutions
	Flickr2K [262]	2017	2650	-	2K resolutions
	Set5 [46]	2012	-	5	Classic 5 images
	Set14 [47]	2012	-	14	Classic 14 images
	BSD100 [264]	2001	-	100	Objects, Natural images
	Manga109 [265]	2015	-	109	109 manga volumes
	Urban100 [266]	2015	-	100	100 urban scenes
	OST300 [330]	2018	-	300	Outdoor scenes
	DIV8K [331]	2019	1304	100	8k resolution for large scale factor
	RealSR [268]	2019	565	30	Real world images for SR
	DRealSR [269]	2020	884	83	Large-scale dataset
Image Deblurring	GoPro [52]	2017	2103	1111	Blurred images at 1280x720
	HIDE [271]	2019	6397	2025	Blurry and sharp image pairs
	RealBlur [270]	2020	3758	980	182 different scenes
Image Denoising	Kodak [332]	1999	-	24	Resolution=768x512
	CBSD68 [333]	2001	-	68	Images with different noisy levels
	McMaster [334]	2011	-	18	Crop size=500x500
Image Classification	ImageNet [210]	2010	1,281,167	100,000	1000 object classes
	ImageNet1K [263]	2020	-	1000	1000 image subset of ImageNet
	LSUN [335]	2015	120,000 to 3,000,000(each category)	1000(each category)	10 scene categories(Classification)
	Places365 [336]	2019	1.8 million	36000	434 scene classes
Face Generation(Recognition)	LFW [337]	2008	13233	-	Images from web with 1080 people
	FFHQ [211]	2019	70000	-	Various faces at 1024x1024
	Celeba-HQ [338]	2018	30000	-	High quality faces at 1024x1024
	AFHQ [339]	2020	15000	-	Animal faces at 512x512
	CelebA [294]	2015	202599	-	Resolution at 178x218
Shadow Removal	ISTD [272]	2018	1330	540	135 scenes with shadow mask images
	SRD [273]	2017	2680	408	Large scale dataset for shadow removal
Image Desnowing	CSD [277]	2021	8000	2000	Large scale dataset for desnowing
	Snow100k [278]	2017	50000	50000	Own 1369 realistic snowy images
	SRRS [279]	2020	15000(paired)+1000(unpaired)	-	Real scenarios from Internet
Image Deraining	RainDrop [49]	2018	1119	-	Various scenes and raindrops
	Outdoor-Rain [281]	2019	9000	1500	Outdoor rainy images
	DDN-data [340]	2017	9100	4900	Real-world clean/rainy image pairs
	SPA-data [282]	2019	295000	1000	Various natural rain scenes
	Rain-100H [341]	2017	1800	100	Five rain streaks
	Rain-100L [341]	2017	200	100	Five rain streaks
Image Dehazing	Haze-4K [274]	2021	4000	-	Attached with associated images
	Dense-Haze [275]	2019	33	-	Out-door hazy scenes
	RESIDE [276]	2019	443950	5342	Real-world hazy images

附录 A 数据集

表 8 总结了用于不同 IR 任务的数据集，包括 SR、图像修复、去模糊、去噪、阴影去除、图像去雪、图像排水和图像去雾。它由发布年份、训练样本和测试样本的数量以及简短描述组成。

附录 B 实施细节

我们在表 9 中总结了监督和基于零样本 DM 的 IR 方法的实现细节，Ta-

图 10 分别。对于基于监督 DM 的 IR，我们阐明了训练数据集、测试数据集和一些关键的实施细节，包括批次大小、迭代、学习率以及训练和推理阶段的采样步骤数。对于基于零样本 DM 的 IR，我们从测试数据集、预训练模型和采样步骤的角度总结了实施细节。

表 9：基于监督扩散模型的方法的实现细节。BS、Iters、LR、Step-T 和 Steps-I 分别表示训练和推理阶段的批次大小、训练迭代、学习率和采样步数。

Target Tasks	Methods	Training Datasets	Testing Datasets	Implementaion Details				
				BS	Iters	LR	Steps-T	Steps-I
Super-resolution	SR3 [100]	FFHQ [211], ImageNet [210]	CelebA-HQ [338], ImageNet 1k [263]	32	1M	1e-4	2000	2000
	SRDiff [101]	CelebA [294], DIV2K [45]+Flickr2K [262]	Dev split, DIV2K(100 test) [45]	16	400K	2e-4	100	100
	CDPMSR [102]	DIV2K [45]	Set5 [46], Set14 [47], Urban100 [266], BSD100 [264], Manga109 [265]	16	300K	1e-4	1000	100
	SR3+ [118]	DIV2K [45], Flickr2K [262], OST300 [330] Collected 61M images	RealSR [268], DRealSR [269]	256	1.5M	-	2000	256
	Resdiff [217]	FFHQ, DIV2K	CelebA(5000) [294] DIV2K(100) [45], Urban100(20) [266]	4	250k	1e-4	1000	-
	SDE-SR [103]	FFHQ [211]	CelebA [294]	-	1M	2e-4	2000	2000
	CDM [218]	LSUN [335]	LSUN [335]	1024	40k	1e-4	2000	100
	IDM [223]	FFHQ [211], DIV2K [45]	DIV2K(100) [45], LSUN [335]	64	1.5M	1e-4	2000	2000
IR	LDM [98]	ImageNet [210]	ImageNet 1K [263]	32	1M	1e-4	2000	2000
	Palette [219]	ImageNet [210], Places2 [342]	imagenet-ctest10k [210] places10k [342]	512	500K	1e-4	2000	1000
	DiffIR [224]	Places-Standard [336], DIV2K [45] CelebA-HQ [338], Flickr2K [262]	CelebA [294], Set5 [46], Set14 [47], Urban100 [266]	30	-	2e-4	4	4
Shadow Removal	Shadowdiffusion [230]	ISTD [272], ISTD+ [343], SRD [273]	ISTD-test [272], ISTD+test [343], SRD-test [273]	-	-	3e-5	1000	25
	DeS3 [191]	SRD [273]	SRD [273]	-	-	-	1000	25
	Refusion [106]	Flickr1024 [262] (Random 800 images)	Flickr1024 [262] (Select 112 images)	8	500K	3e-4	100	100
Image Deblurring	Deblur-DPM [105]	GoPro [52]	HIDE [271]	32	-	1e-4	2000	10-500
	DG-DPM [222]	GoPro [52]	RealBlur-J [270], REDS [51], HIDE [271]	256	60k	1e-4	1000	20-1000
Image Deraining	WeatherDiff [104]	Snow100k [278], Rain-drop [49] Outdoor-rain [281]	Outdoor-rain(test) [281] Snow100k(test) [278], RainDrop-A [49]	64	2000k	2e-5	-	10,50
	RainDiffusion [122]	Rain200H [?] , Rain200L [?] , Rain800 [48], DDN-Data [340]	Rain200H [?] , Rain200L [?] , Rain800 [48], DDN-Data [280]	4	-	2e-5	-	-

表 10：基于零样本扩散模型的方法中使用的实现细节

Methods	Training Datasets	Datasets	Inference Steps
RED-Diff [199]	Pretrained on ImageNet [210]	ImageNet 1K [263]	1000
Rapaint [212]	Pretrained on CelebA-HQ [338] and ImageNet [210]	CelebA-HQ [338], ImageNet [210]	250
CoPaint [216]	Pretrained on ImageNet [210], CelebA-HQ [338]	CelebA-HQ [338], ImageNet [210]	250
ILVR [195]	Pretrained on FFHQ [211], AFHQ [339], LSUN[335], Places365 [336]	Images from the website	250
CCDF [213]	Pretrained on FFHQ [211], AFHQ [339] , fastMRI knee [344]	FFHQ [211], AFHQ [339], fastMRI knee [344]	20
SNIPS [110]	Pretrained NSCN on CelebA [294],	CelebA [294], LSUN [335]	500
DDRM [112]	Pretrained on CelebA-HQ [338], LSUN [335], ImageNet [210]	FFHQ [211], ImageNet 1K [263], LSUN [335]	20
DDNM [113]	Pretrained on ImageNet [210], CelebA-HQ [338]	ImageNet 1K [263] CelebA [294]	100
Bahjat Kavar et.al [111]	Pretrained on ImageNet [210]	ImageNet 1K [263]	20
MCG [236]	Pretrained on FFHQ [211], ImageNet [210]	FFHQ [211], LSUN [335]	1000
DPS [114]	Pretrained on ImageNet [210]	FFHQ [211], ImageNet 1K [263]	1000
GibbsDDRM [119]	Pretrained on FFHQ [211], AFHQ [339]	FFHQ [211], AFHQ [339]	50
DiffFace [237]	Pretrained on FFHQ [211], ImageNet [210]	CelebA-HQ [338]	250
Dirac-Diffusion [200]	SDE-VP trained on CelebA [294] , ImageNet [210]	CelebA [294], ImageNet [263]	100
IIGDM [197]	Diffusion model trained on ImageNet [210]	ImageNet [263]	100
ADIR [117]	Pretrained on ImageNet [210]	DIV2K [45]	1,000