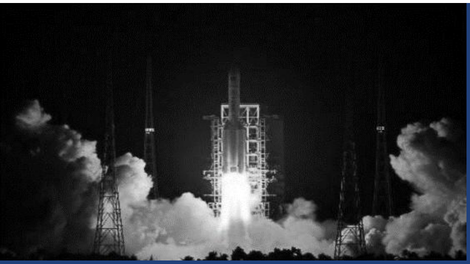
探索统一的图像到图像分布变换框架

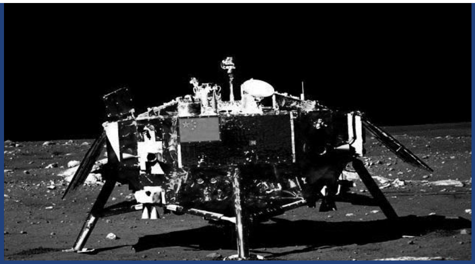
--- 残差去噪扩散模型

汇报人：刘佳伟 博士 助理研究员

单　位：中国科学院沈阳自动化研究所

邮　箱：liujiawei@sia.cn

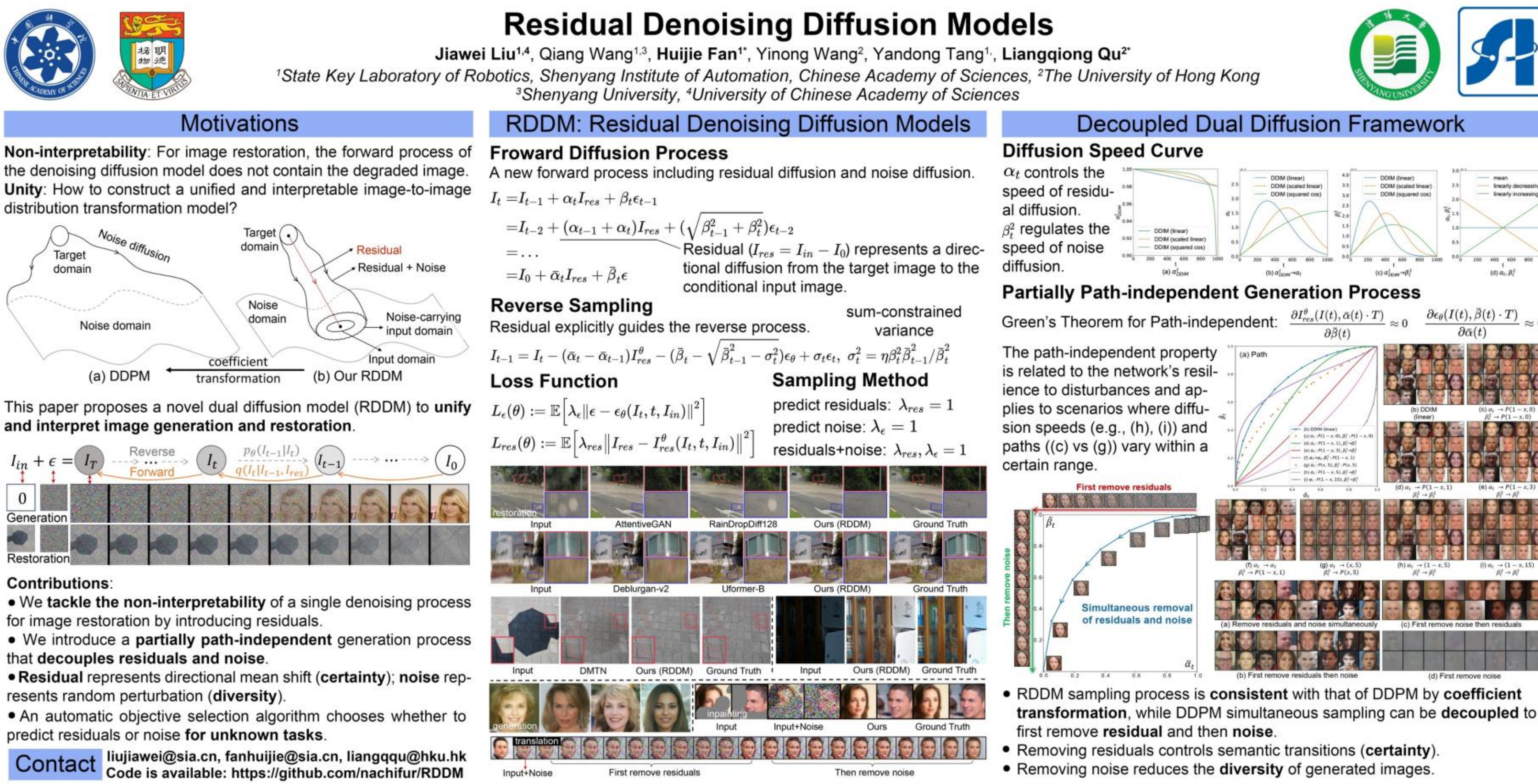




0 概述

➢ 残差去噪扩散模型 (Residual Denoising Diffusion Models)

- **成果：IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2024|人工智能国际顶级学术会议|CCF A类会议)**

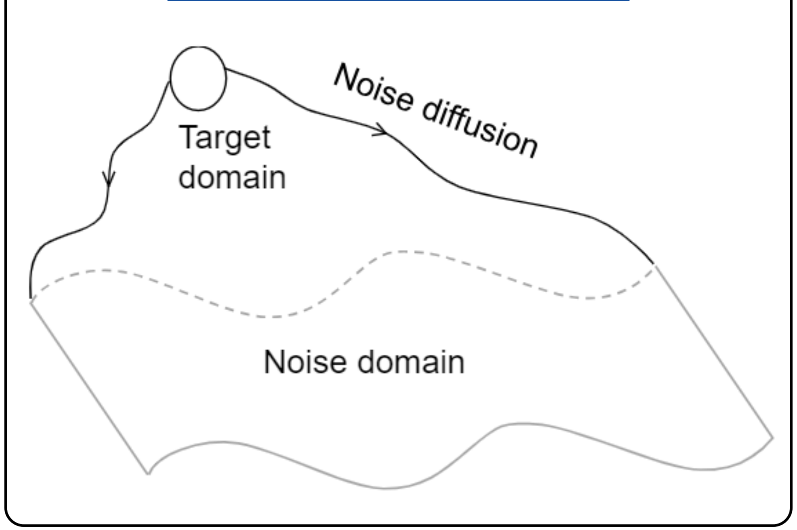


0 概述

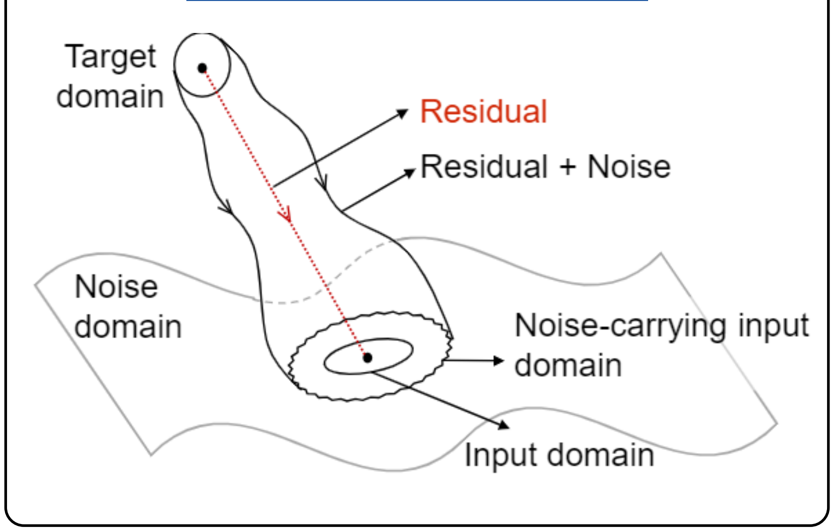
➢ 残差去噪扩散模型是统一的图像到图像分布变换框架

- **统一性：统一建模**图像生成/补全、成对图像（恢复）、不成对图像（翻译）
- **解释性：**残差强调确定性，而噪声反应多样性；**解决了去噪扩散模型中随机性和确定性相矛盾**的问题、**解决了去噪模型不能解释图像恢复**任务的问题；
- **兼容性：**转换系数的残差去噪模型与去噪扩散模型**采样过程一致性证明**。

去噪扩散模型



残差去噪扩散模型

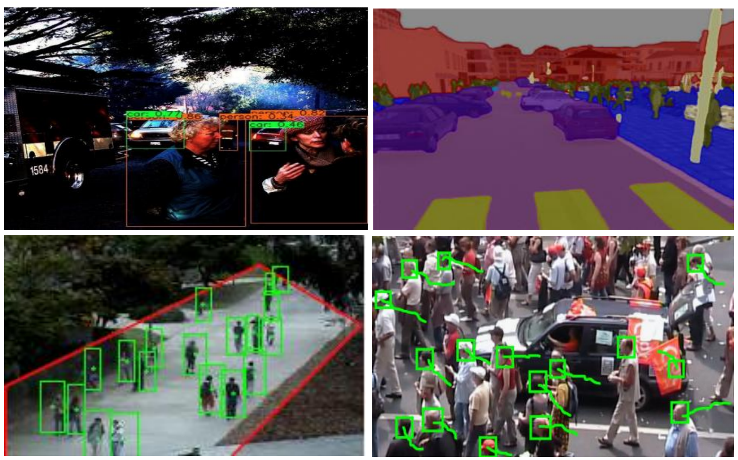


1 研究背景

➢ 复杂光照/恶劣天气

- **阴影**覆盖背景物体**降低质量**；
- **低光**照图像**对比度低噪声多**；
- **高光**严重丢失纹理细节；
- **雨雪遮挡和模糊**背景物体；
- **雾和运动模糊损失大量细节**；
- ……。

复杂环境严重影响高层视觉性能

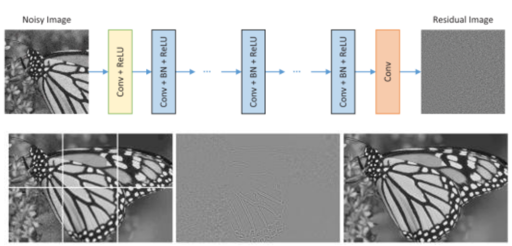


如何构建统一光照处理和恢复模型，提高视觉系统鲁棒性？

➢ 图像恢复方法研究现状

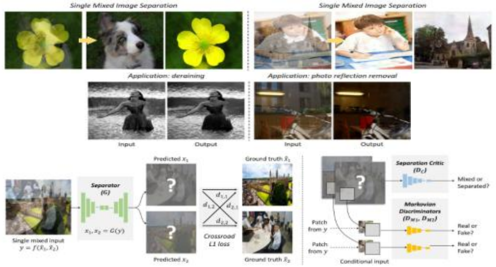
- **技术思路：**基于CNN、基于GAN、基于Transformer恢复退化图像
- **存在问题：**不同任务需设计不同的网络结构、损失函数和训练策略
- **代表方法：**

基于CNN[1]



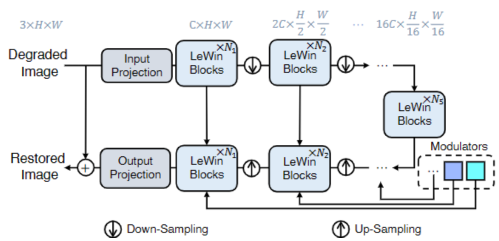
- ✓ 哈尔滨工业大学
- ✓ 利用残差学习实现了对高斯去噪、超分辨率和JPEG 去块等去噪任务

基于GAN[2]



- ✓ 美国密歇根大学
- ✓ 使用对抗训练，将去雨、去反光看成叠加图像分离问题

基于Transformer[3]



- ✓ 中国科学技术大学
- ✓ 通过设计局部窗口增强Transformer块捕获局部上下文，可去除噪声和模糊

[1] Zhang K, Zuo W, Chen Y, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising [J]. IEEE TIP, 2017, 26(7): 3142-3155.
[2] Zou Z, Lei S, Shi T, et al. Deep Adversarial Decomposition: A Unified Framework for Separating Superimposed Images [C/OL]//Proc. CVPR. 2020: 12803-12813.
[3] Wang Z, Cun X, Bao J, et al. Uformer: A general u-shaped transformer for image restoration [C]//Proc. CVPR. 2022: 17683-17693.

3 去噪扩散模型

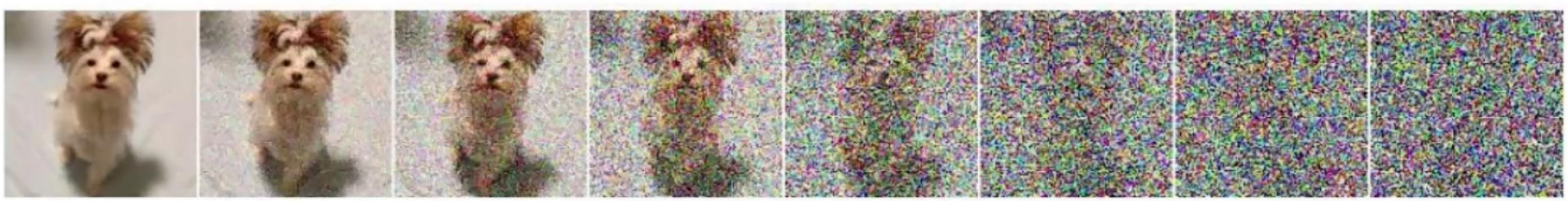
➢ 前向过程

- **将噪声添加到清晰图像中**

前向过程

✓ 与VAE不同，扩散模型的前向过程是不可训练的、固定的马尔科夫链

$$I_0 \rightarrow I_1 \rightarrow I_2 \rightarrow \cdots \rightarrow I_T$$



目标: $I_0 \sim q(I_0) \xrightarrow{q(I_{1:T}|I_0) := \prod_{t=1}^{T} q(I_t|I_{t-1})}$ 噪声: $\mathcal{N}(I_T; \mathbf{0}, \mathbf{I})$

加噪: $q(I_t|I_{t-1}) := \mathcal{N}(I_t; \sqrt{\alpha_t} I_{t-1}, (1-\alpha_t)\mathbf{I})$ $I_t = \sqrt{\alpha_t} I_{t-1} + \sqrt{1-\alpha_t}\epsilon_{t-1}$

3 去噪扩散模型

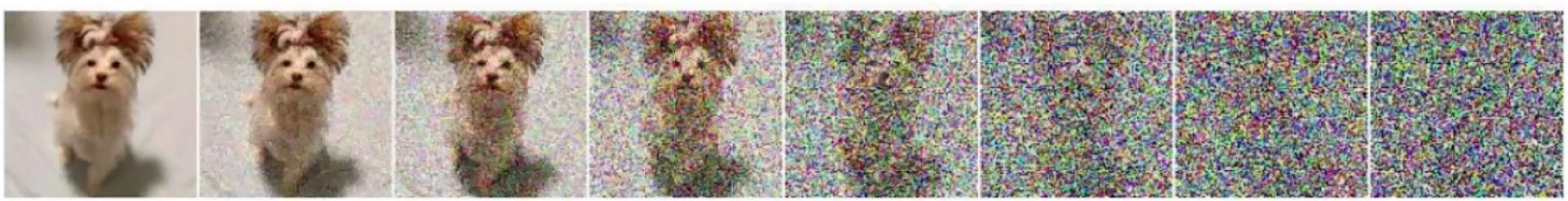
➢ 逆向过程

- **去除在前向过程中添加的噪声，然后生成清晰图像**

逆向过程

✓ 具可学习高斯转移概率的马尔可夫链（从高斯噪声开始）

$$I_0 \leftarrow I_1 \leftarrow I_2 \leftarrow \cdots \leftarrow I_T$$



$$p_\theta(I_{0:T}) := p_\theta(I_T)\prod_{t=1}^{T}\boxed{p_\theta}(I_{t-1}|I_t)$$

目标: $I_0 \sim q(I_0) \longleftarrow$ 噪声: $\mathcal{N}(I_T;\mathbf{0},\mathbf{I})$

采样: $p_\theta(I_{t-1}|I_t) := \mathcal{N}(I_{t-1};\mu_\theta(I_t,t),\Sigma_t\mathbf{I})$ $\quad I_{t-1} = \mu_\theta(I_t,t) + \sqrt{\Sigma_t}\cdot\epsilon_t$

流程: $\sqrt{\alpha_t}I_{t-1} + \sqrt{1-\alpha_t}\epsilon_{t-1} \xrightarrow[\text{noise}]{\text{add}} I_t \longrightarrow \epsilon_\theta(I_t) \longrightarrow I_0^\theta \xrightarrow{\text{Bayes}} p_\theta(I_{t-1}|I_t) \xrightarrow{\text{sample}} I_{t-1}$

4 残差去噪扩散模型

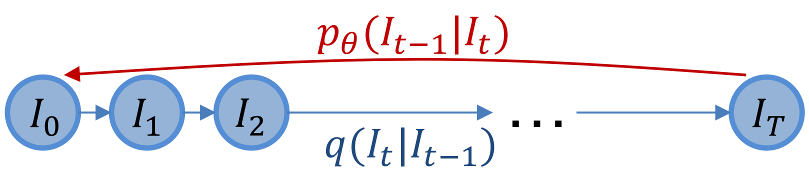
➢ 统一可解释的图到图分布变换模型的设计动机

- **现有基于去噪扩散模型的图像恢复方法缺乏解释性**

基于去噪扩散的图像恢复方法

✓ 逆向过程：去除噪声，生成清晰图像

✓ 前向过程：将噪声添加到清晰图像中



不可解释性

✓逆向过程没必要从纯噪声开始，因为图像恢复中退化图像是已知的

$$\mathcal{N}(I_T;\mathbf{0},\mathbf{I}) \xrightarrow{\epsilon_\theta(I_t,t)} I_0$$

Conditions: I_{in}

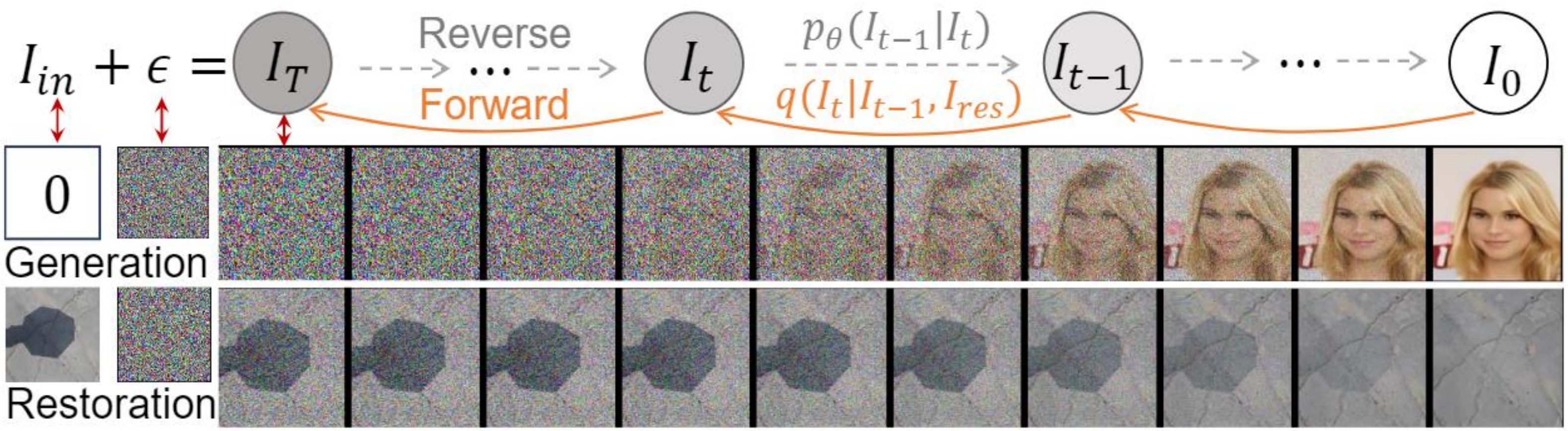
✓前向过程不能解释图像恢复，因为不包含有关退化图像的任何信息

$$\begin{aligned}
I_t &= \sqrt{\alpha_t} I_{t-1} + \sqrt{1-\alpha_t}\epsilon_{t-1} \\
&= \sqrt{\alpha_t}\left(\sqrt{\alpha_{t-1}} I_{t-2} + \sqrt{1-\alpha_{t-1}}\epsilon_{t-2}\right) + \sqrt{1-\alpha_t}\epsilon_{t-1} \\
&= \cdots \\
&= \sqrt{\bar{\alpha}_t} I_0 + \sqrt{1-\bar{\alpha}_t}\epsilon
\end{aligned}$$

4 残差去噪扩散模型

➢ 核心贡献：

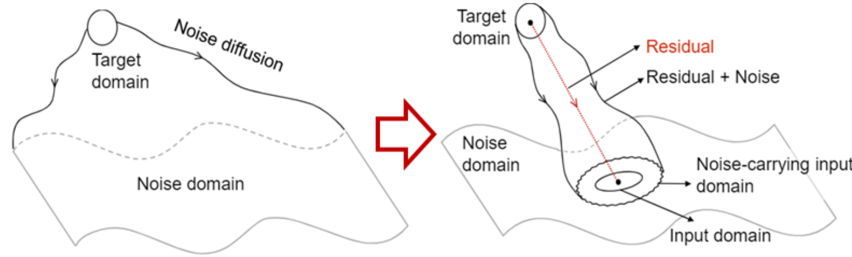
- 提出了**残差去噪扩散模型**，可**统一和解释**图像**生成**和图像**恢复**；
- 提出了**部分路径无关**的生成过程，**残差**控制定向偏移（**确定性**），而**噪声**控制随机扰动（**多样性**）；
- 设计了一种**优化目标自动选择**算法，对于未知新任务，可以**自动地选择残差预测还是噪声预测**；
- 适用**不同确定性或多样性需求的任务**，如生成、恢复、补全和翻译。



4 残差去噪扩散模型

- **重新定义新的前向过程：**

$$
\begin{aligned}
I_t &= I_{t-1} + \alpha_t I_{res} + \beta_t \epsilon_{t-1} \\
&= I_{t-2} + (\alpha_{t-1} + \alpha_t) I_{res} + (\sqrt{\beta_{t-1}^2 + \beta_t^2}) \epsilon_{t-2} \\
&= \ldots \\
&= I_0 + \bar{\alpha}_t I_{res} + \bar{\beta}_t \epsilon,
\end{aligned}
$$



噪声扩散　　残差扩散+噪声扩散

- **逆向过程：受扰动生成($\eta = 1$)和确定采样($\eta = 0$)：**

$$
q_\sigma(I_{t-1}|I_t, I_0, I_{res}) := \mathcal{N}(I_{t-1}; I_0 + \bar{\alpha}_{t-1} I_{res} + \sqrt{\bar{\beta}_{t-1}^2 - \sigma_t^2} \frac{I_t - (I_0 + \bar{\alpha}_t I_{res})}{\bar{\beta}_t}, \sigma_t^2 \mathbf{I}),
$$

$$
I_{t-1} = I_t - (\bar{\alpha}_t - \bar{\alpha}_{t-1}) I_{res}^\theta - (\bar{\beta}_t - \sqrt{\bar{\beta}_{t-1}^2 - \sigma_t^2}) \epsilon_\theta + \sigma_t \epsilon_t, \text{where } \epsilon_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I}).
$$

- **目标函数：残差预测和噪声预测**

$$
L_{res}(\theta) := \mathbb{E}\left[\lambda_{res} \left\| I_{res} - I_{res}^\theta(I_t, t) \right\|^2\right], \qquad L_\epsilon(\theta) := \mathbb{E}\left[\lambda_\epsilon \left\| \epsilon - \epsilon_\theta(I_t, t) \right\|^2\right],
$$

- **扩散系数时间表转化：**

$$
\bar{\alpha}_t = 1 - \sqrt{\bar{\alpha}_{DDIM}^t}^{1}, \qquad \bar{\beta}_t = \sqrt{1 - \bar{\alpha}_{DDIM}^t}, \qquad \sigma_t^2 = \sigma_t^2(DDIM).
$$

- **理论证明：扩散系数转化后，RDDM与DDPM/DDIM的采样一致**

4 残差去噪扩散模型

➢ 如何选择最佳采样方法?

- 残差侧重确定性，因此**残差预测适用图像恢复**；
- 噪声强调多样性，因此**噪声预测适用图像生成**；
- 提出了一种**优化目标自动选择**算法，自动地选择残差预测还是噪声预测。

采样方法分析

采样方法	图像生成（CelebA[169]）		阴影去除（ISTD[22]）			低光照增强（LOL[170]）		去雨（RainDrop[35]）	
	FID (↓)	IS (↑)	MAE(↓)	PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)
SM-Res	31.47	1.73	4.76	30.72	0.959	**25.39**	**0.937**	31.96	0.9509
SM-N	**23.25**	**2.05**	81.01	11.34	0.175	16.30	0.649	19.15	0.7179
SM-Res-N	28.90	1.78	**4.67**	**30.91**	**0.962**	23.90	0.931	**32.51**	**0.9563**

采样方法	网络结构	MAE(↓)	SSIM(↑)	PSNR(↑)
SM-Res	Residual network	4.76	0.959	30.72
SM-Res-N-2Net	Residual network+noise network	4.67	0.962	30.91
SM-Res-N-1Net	One network, only shared encoder	4.72	0.959	30.73
SM-Res-N-1Net	One network	**4.57**	**0.963**	**31.10**

采样方法	阴影去除	去雨	去模糊	图像生成	
Method	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	FID(↓)	IS(↑)
SM-Res	30.72/0.959	31.96/0.9509	32.32/0.957	31.47	1.73
SM-N	11.34/0.175	19.15/0.7179	9.49/0.087	**23.25**	**2.05**
SM-Res-N	30.91/0.962	**32.51/0.9563**	**32.40/0.963**	28.90	1.78
SM-Res-N-1Net	**31.10/0.963**	31.79/0.9504	31.69/0.951	28.57	1.81

优化目标自动选择算法

算法 5-1: 带有 AOSA 的训练管道。

输入: 一张退化输入图像 I_{in}，及其对应的标签图像 I_0，高斯噪声 ϵ，时间条件 t，系数时间表 $\bar{\alpha}$ 和 $\bar{\beta}_t$，初始可学习参数 $\lambda_{res}^{\theta}=0.5$。深度网络 G，参数为 θ，初始学习率为 l，训练迭代次数 n，AOSA 的迭代次数 m。转化训练的阈值 $\delta=0.01$。

输出: 训练好的参数 θ 和 λ_{res}^{θ}。

$\theta \leftarrow$ InitWight (G) ▷ 初始化网络参数
for $i \leftarrow 1$ **to** $n+m$ **do**
  $t \sim$ Uniform $(\{1,2,...,T\})$, $\epsilon \sim \mathcal{N}(\mathbf{0},\mathbf{I})$, $I_{res} \leftarrow I_{in} - I_0$
  $I_t \leftarrow I_0 + \bar{\alpha}_t I_{res} + \bar{\beta}_t \epsilon$ ▷ 通过公式(5-8)合成 I_t
  $I_{out} \leftarrow G(I_t, t, I_{in})$
  $I_{res}^{\theta} \leftarrow \lambda_{res}^{\theta} \times I_{out} + (1-\lambda_{res}^{\theta}) \times f_{\epsilon \to res}(I_{out})$ ▷ $f_{\epsilon \to res}(\cdot)$ 表示从 ϵ 到 I_{res} 的转换通过公式(5-11)
  $\epsilon_{\theta} \leftarrow \lambda_{res}^{\theta} \times f_{res \to \epsilon}(I_{out}) + (1-\lambda_{res}^{\theta}) \times I_{out}$ ▷ $f_{res \to \epsilon}(\cdot)$ 表示从 I_{res} 到 ϵ 的转换通过公式(5-11)
  $L_{auto} \leftarrow$ Loss $(I_{res}^{\theta}, I_{res}, \epsilon_{\theta}, \epsilon)$ ▷ 基于公式(5-34)
  $\theta, \lambda_{res}^{\theta} \xleftarrow{+} -\nabla_{\theta,\lambda_{res}^{\theta}}(\mathcal{L}_{auto}, l)$ ▷ 梯度更新
  if abs $(\lambda_{res}^{\theta} - 0.5) < \delta$ **then**
    pass ▷ 类似对抗的训练
  else
    $\lambda_{res}^{\theta} \leftarrow$ Detach (λ_{res}^{θ}) ▷ 停止梯度更新
    $\theta \leftarrow$ InitWight (G) ▷ 重新初始化网络参数 **if** $\lambda_{res}^{\theta} > 0.5$ **then**
      $\lambda_{res}^{\theta} \leftarrow 1$ ▷ SM-Res
    else
      $\lambda_{res}^{\theta} \leftarrow 0$ ▷ SM-N
    end
  end
end

4 残差去噪扩散模型

➢ 模型性质：求和约束的方差时间表

- DDPM每一步的方差为1；DDIM方差为0；
- RDDM是求和约束的方差，10步采样时每步方差=1/10，100步采样=1/100。

DDPM: $$\sigma_t^2(DDIM) = \eta \frac{(1-\bar{\alpha}_{DDIM}^{t-1})}{1-\bar{\alpha}_{DDIM}^{t}}(1-\frac{\bar{\alpha}_{DDIM}^{t}}{\bar{\alpha}_{DDIM}^{t-1}}) \approx 1$$

RDDM: $$\sigma_t^2(RDDM) = \eta\bar{\alpha}_{DDIM}^{t-1} \frac{(1-\bar{\alpha}_{DDIM}^{t-1})}{1-\bar{\alpha}_{DDIM}^{t}}(1-\frac{\bar{\alpha}_{DDIM}^{t}}{\bar{\alpha}_{DDIM}^{t-1}}),$$

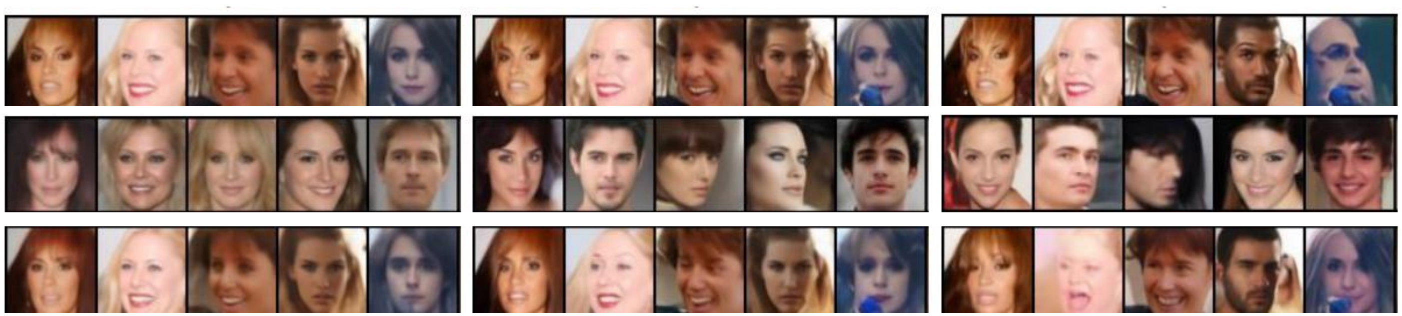
$$\sum_{i=1}^{T} \sigma_t^2(RDDM) = \sum_{i=1}^{T} \eta\beta_t^2 \frac{\bar{\beta}_{t-1}^2}{\bar{\beta}_t^2} \leq \sum_{i=1}^{T} \beta_t^2 \leq 1,$$

定性比较

DDIM:

DDPM:

RDDM:



Step 10　Step 20　Step 50

4 残差去噪扩散模型

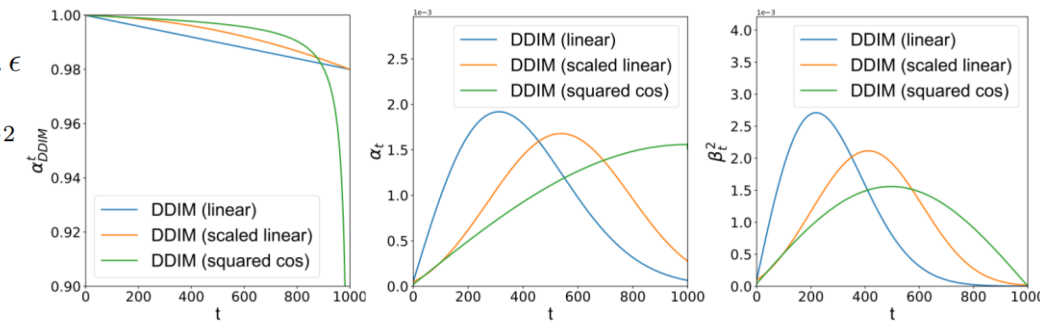
➢ 模型性质：解耦的前向扩散过程

- α_t控制残差扩散的速度；β_t控制噪声扩散的速度；
- 更好的系数时间表：α_t线性递减，β_t线性递增。

扩散速度曲线

✓更清晰的物理解释

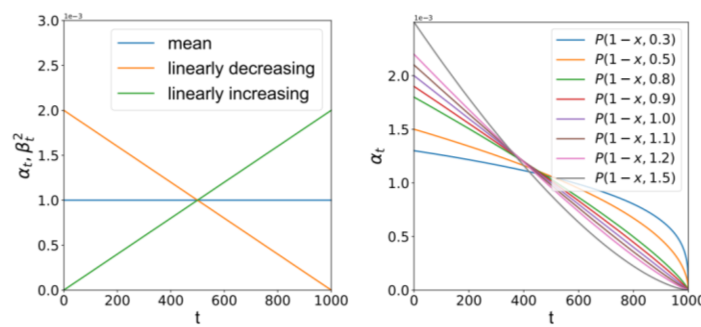
$$
\begin{aligned}
I_t =& I_{t-1} + \alpha_t I_{res} + \beta_t \epsilon_{t-1}, \text{where } \epsilon_{t-1}, \epsilon_{t-2} \ldots \epsilon \\
=& I_{t-2} + (\alpha_{t-1} + \alpha_t) I_{res} + (\sqrt{\beta_{t-1}^2 + \beta_t^2}) \epsilon_{t-2} \\
=& \ldots \\
=& I_0 + \bar{\alpha}_t I_{res} + \bar{\beta}_t \epsilon,
\end{aligned}
$$



(a) α_{DDIM}^t (b) $\alpha_{DDIM}^t \to \alpha_t$ (c) $\alpha_{DDIM}^t \to \beta_t^2$

✓更灵活的系数时间表设计

Schedules	FID (↓)	IS (↑)
Linear (DDIM [51])	28.39[4]	2.05
Scaled linear [48]	28.15	2.00
Squared cosine [44]	47.21	**2.64**
α_t (mean), β_t^2 (mean)	38.35	2.22
α_t (linearly increasing), β_t^2 (linearly increasing)	40.03	2.45
α_t (linearly decreasing), β_t^2 (linearly decreasing)	27.82	2.26
α_t (linearly decreasing), β_t^2 (linearly increasing)	**23.25**	2.05



(d) α_t, β_t^2 (e) $P(x, a)$

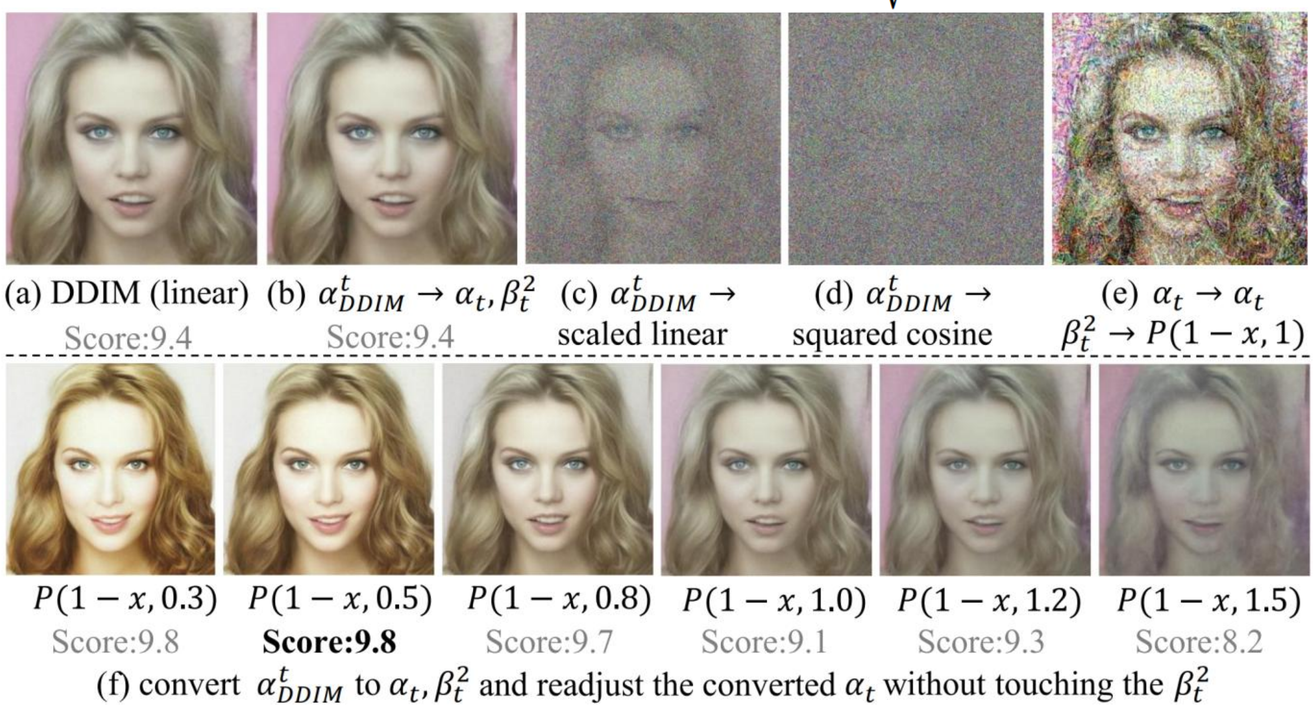
4 残差去噪扩散模型

➢ 解耦的前向过程->生成过程是否解耦？

- 在DDPM或DDIM中，当**系数时间表α^t_{DDIM}改变**时，扩散过程也发生改变，因此需重新训练去噪网络。如(c-d)，直接改变α^t_{DDIM}导致**去噪失败**。

重新调整系数时间表

✓采样公式：$I_{t-1} = I_t - (\bar{\alpha}_t - \bar{\alpha}_{t-1}) I^{\theta}_{res} - (\bar{\beta}_t - \sqrt{\bar{\beta}^2_{t-1} - \sigma^2_t}) \epsilon_\theta + \sigma_t \epsilon_t$

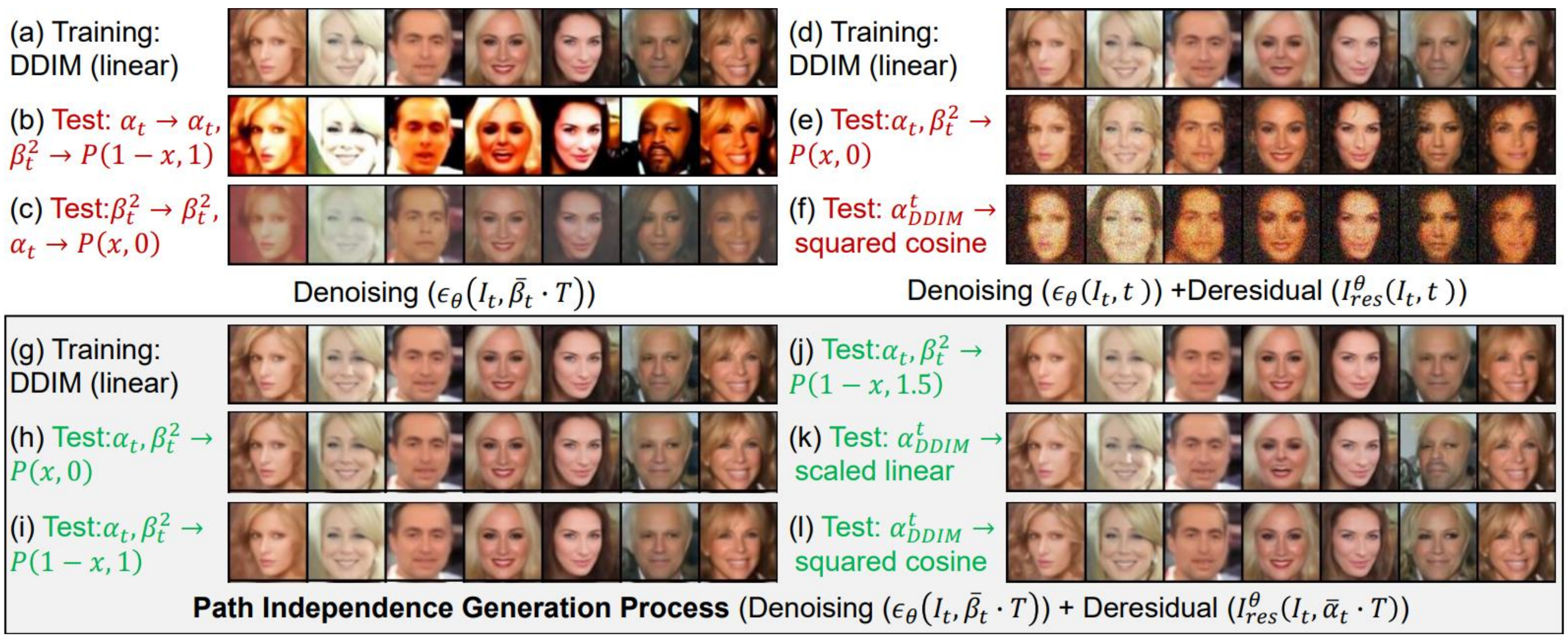


4 残差去噪扩散模型

➢ 两点简单改进->部分路径独立的生成过程

- 使用两个网络分别估计I_{res}^{θ}和ϵ_θ，避免互相表示 $I_t = I_{in} + (\bar{\alpha}_t - 1)I_{res} + \bar{\beta}_t\epsilon$
- $\bar{\alpha}_t \cdot T, \bar{\beta}_t \cdot T$作为时间条件 $I_{res}^{\theta}(I_t, t, 0) \to I_{res}^{\theta}(I_t, \bar{\alpha}_t \cdot T, 0)$，$\epsilon_\theta(I_t, t, 0) \to \epsilon_\theta(I_t, \bar{\beta}_t \cdot T, 0)$

部分路径独立的生成过程

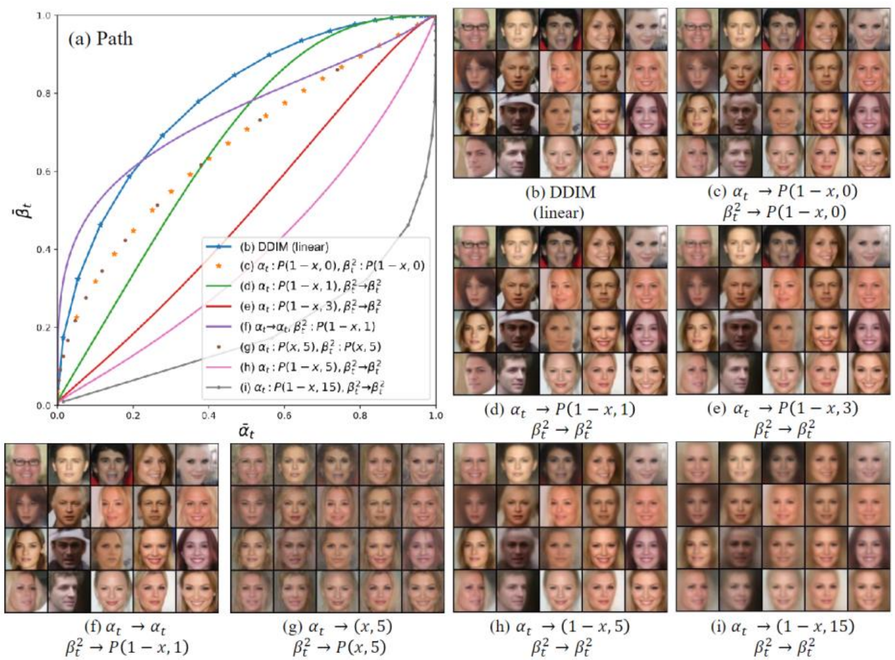


4 残差去噪扩散模型

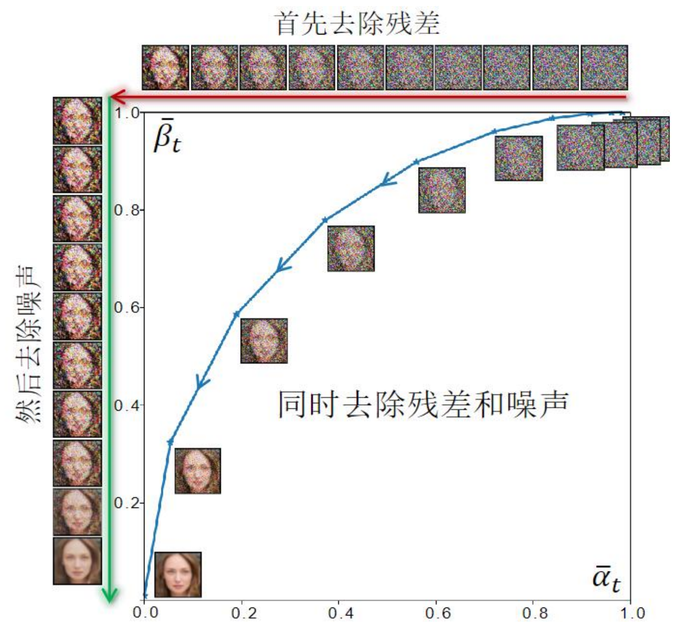
➢ 模型性质：部分路径独立的生成过程

- 格林公式与路径独立：当**扩散速度和路径在一定范围内**变化时，生成过程呈现出**路径独立**的性质，$\frac{\partial I^{\theta}_{res}(I(t),\bar{\alpha}(t)\cdot T)}{\partial \bar{\beta}(t)} \approx 0, \quad \frac{\partial \epsilon_{\theta}(I(t),\bar{\beta}(t)\cdot T)}{\partial \bar{\alpha}(t)} \approx 0.$

不同路径对生成的影响



特殊路径的分析



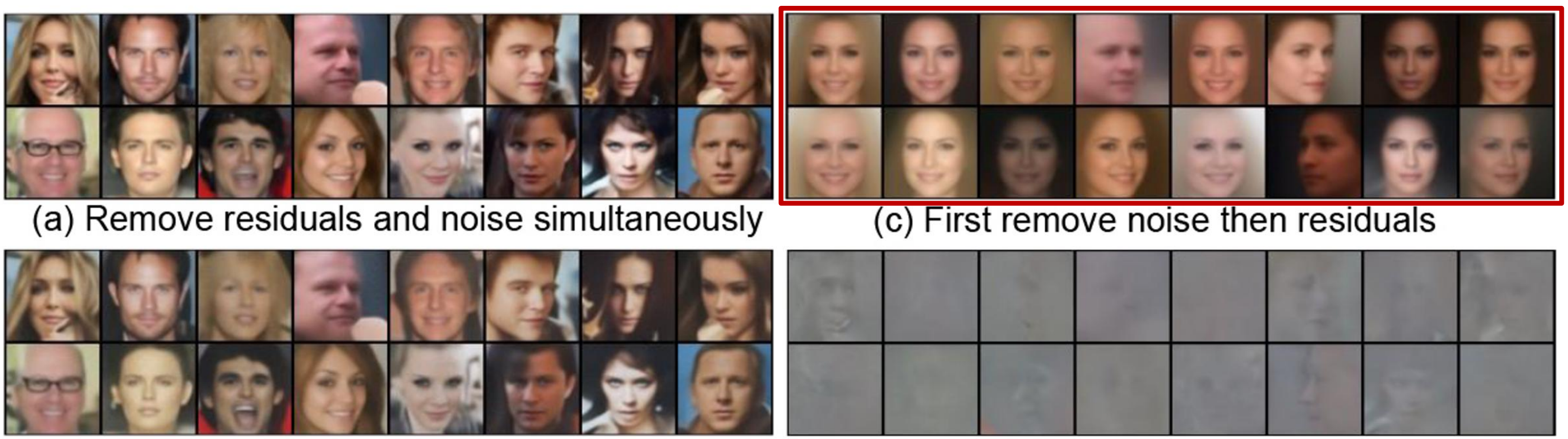
$$I(T) \xrightarrow{-I_{res}} I(0)+\bar{\beta}_T\epsilon \xrightarrow{-\bar{\beta}_T\epsilon} I(0)$$

4 残差去噪扩散模型

➢ 模型性质：部分路径独立的生成过程

- 格林公式与路径独立：当**扩散速度和路径在一定范围内**变化时，生成过程呈现出**路径独立**的性质，$\frac{\partial I^{\theta}_{res}(I(t),\bar{\alpha}(t)\cdot T)}{\partial \bar{\beta}(t)} \approx 0, \quad \frac{\partial \epsilon_{\theta}(I(t),\bar{\beta}(t)\cdot T)}{\partial \bar{\alpha}(t)} \approx 0.$

特殊路径的分析



(a) Remove residuals and noise simultaneously

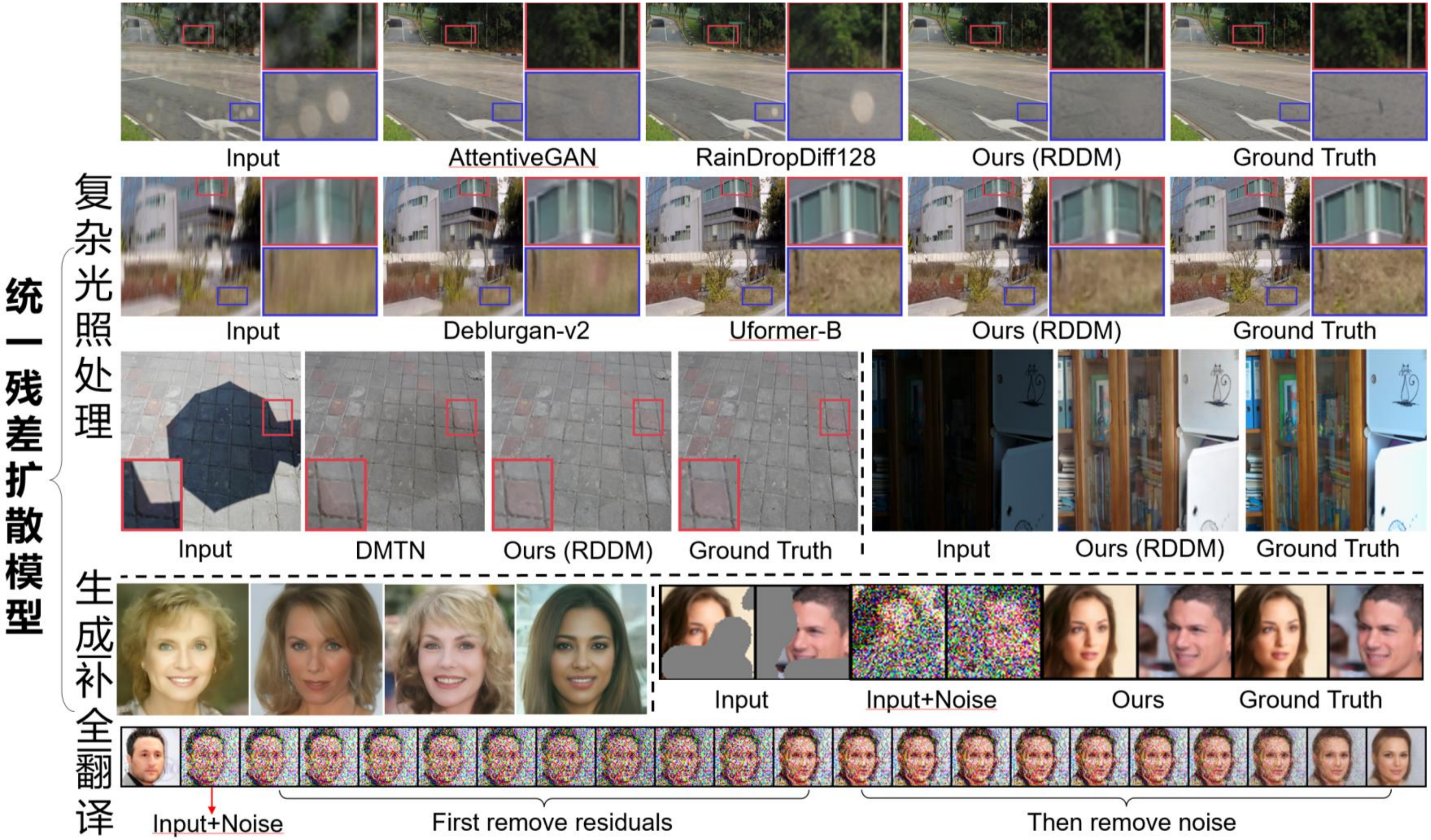
(c) First remove noise then residuals

(b) First remove residuals then noise

(d) First remove noise

✓ 先去除噪声后去噪残差：$I(T) \xrightarrow{-\bar{\beta}_T \epsilon} I_{in} \xrightarrow{-I_{res}} I(0)$，由于对于图像生成任务$I_{in}=0$，从$I_{in}$到$I(0)$失败。

4 残差去噪扩散模型



统一建模：成对图像（恢复）、生成和补全、不成对图像（翻译）

4 残差去噪扩散模型

➢ 实验结论：

- 图像生成：RDDM**兼容**DDPM/DDIM图像生成模型;
- 图像恢复：RDDM在**不超过5个采样步数**下实现SOTA;
- 适用**图像补全和图像翻译，** 且不用额外的条件编码。

定量实验结果

图像生成、低光照、阴影去除、去雨、去模糊

Method	LIME [86]	DSLR [87]	SID [88]	D&E [65]	MIR-Net [89, 66]	UTVNet [67]	SNR-Aware [59]	Ours (RDDM)
PSNR(↑)	17.76	17.25	21.16	22.13	22.34	22.69	22.87	**23.97**
SSIM(↑)	0.3506	0.4229	0.6398	0.7172	0.7031	0.7179	0.625	**0.8392**

(a) CelebA [37](FID)	DDIM [17]	Our RDDM ($\alpha^t_{DDIM} \to \alpha_t, \beta_t^2$)
5 steps	69.60	69.60
10 steps	40.45	40.41
15 steps	32.67	32.71
20 steps	30.61	30.77
100 steps	23.66	24.92

(b) Shadow Removal	MSE(↓)			SSIM(↑)			PSNR(↑)		
	S	NS	ALL	S	NS	ALL	S	NS	ALL
DSC [39] ¶	9.48	6.14	6.67	0.967	-	-	33.45	-	-
FusionNet [40]	7.77	5.56	5.92	0.975	0.880	0.945	34.71	28.61	27.19
BMNet [41]	7.60	4.59	5.02	0.988	0.976	0.959	35.61	32.80	30.28
DMTN [5]	7.00	4.28	4.72	**0.990**	**0.979**	**0.965**	35.83	33.01	30.42
Ours (RDDM)	**6.67**	**4.27**	**4.67**	0.988	**0.979**	0.962	**36.74**	**33.18**	**30.91**

(c) Low-light	PSNR(↑)	SSIM(↑)	LPIPS (↓)
Retinex-Net [42]	16.774	0.462	0.474
KinD [45]	17.648	0.779	0.175
KinD++ [48]	17.752	0.760	0.198
RUAS [51]	18.230	0.720	0.350
KinD++-SKF [54]	20.363	0.805	0.201
DCC-Net [57]	22.72	0.81	-
SNR-Aware [59]	24.608	0.840	0.151
Ours (RDDM)	**25.392**	**0.937**	**0.134**

(d) Deraining	PSNR(↑)	SSIM(↑)
pix2pix [43]	28.02	0.8547
DuRN [46]	31.24	0.9259
RainAttn [49]	31.44	0.9263
AttnGAN [52]	31.59	0.9170
IDT [55]	31.87	0.9313
RainDiff64 [28]	32.29	0.9422
RainDiff128 [28]	32.43	0.9334
Ours (RDDM)	**32.51**	**0.9563**

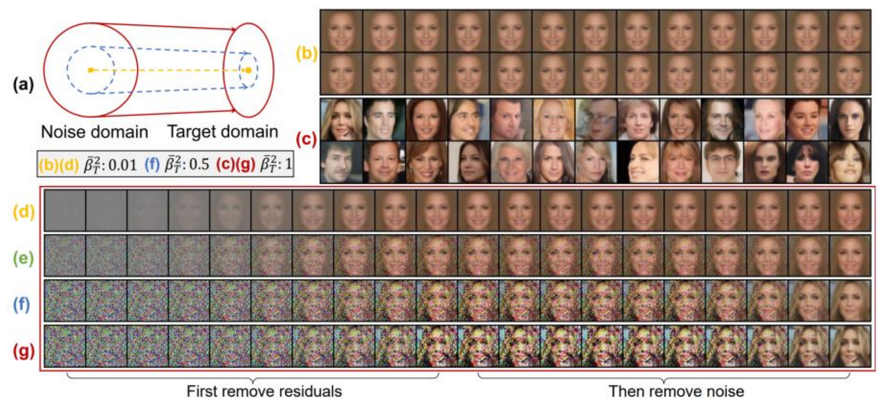
(e) Deblurring	PSNR(↑)	SSIM(↑)
Nah *et al.* [44]	29.08	0.914
Zhang *et al.* [47]	29.19	0.931
DeblurganV2 [50]	29.55	0.934
Gao *et al.* [53]	30.90	0.935
Suin *et al.* [56]	31.85	0.948
MPRNet [58]	32.66	0.959
Uformer-B [60]	**32.97**	**0.967**
Ours (RDDM)	32.40	0.963

4 残差去噪扩散模型

➢ 消融实验分析：

- **噪声控制生成图像的多样性**，且利于恢复图像细节和增强感知质量；
- **输入图像作为**去残差和去噪网络**条件**时，确定性被增强，**多样性减少**。

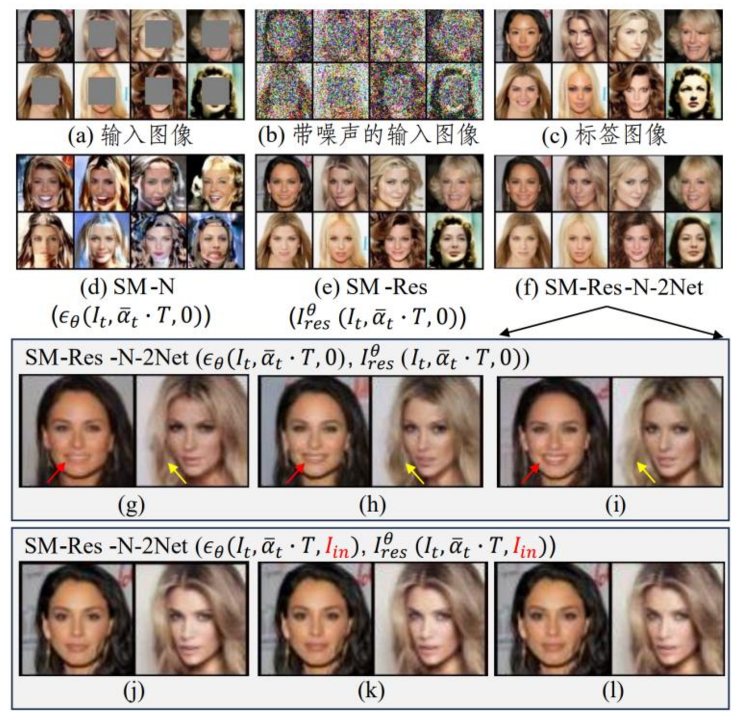
噪声强度可视分析



适量噪声利于恢复图像

RDDM (SM-Res-N)	metric	1 step	2 step	5 step	10 step	100 step
$\bar{\beta}_T^2 = 0.01$	MAE-ALL (↓)	4.83	4.69	**4.67**	4.72	4.90
	PSNR-S (↑)	36.83	**36.98**	36.74	36.59	36.41
	LPIPS (↓)	0.0344	0.0308	**0.0305**	0.0314	0.0334
$\bar{\beta}_T^2 = 1$	MAE-ALL (↓)	5.07	4.94	4.90	**4.87**	4.99
	PSNR-S (↑)	36.93	**37.20**	37.07	37.01	36.62
	LPIPS (↓)	0.0346	0.0314	**0.0298**	0.0300	0.0319

采样方法的消融



4 残差去噪扩散模型

➢ 运行效率消融分析：

- 由于采样步数较少，RDDM的推理时间可与CNN方法相竞争；
- 与基于去噪扩散的SR3(TPAMI 2022) 相比，**基于残差的RDDM训练次数减少 10 倍，推理时间加快10 倍，参数减少10倍，性能提升10%**；
- 采用残差预测的RDDM仅需要 4.8G GPU 内存进行训练

运行效率结果

更快训练
更快推理
更少参数
更高性能

（a）低光照	PSNR (↑)	SSIM (↑)	LPIPS (↓)	Params (M)	MAC (G)×Steps	Inference Time(s)
LLformer	23.649	0.816	0.169	24.51	**22.0×1 = 22.0**	0.09×1 = 0.09
LLFlow	25.19	0.93	**0.11**	17.42	286.33×1 = 286.3	0.18×1 = 0.18
Ours (RDDM)	**25.392**	**0.937**	0.116	**7.73**	32.9×2 = 65.8	**0.03×2 = 0.06**

（b）阴影去除	MAE (↓)	PSNR (↑)	SSIM (↑)	Params (M)	MAC (G) × Steps	Inference Time (s)
Shadow Diffusion [102]	**4.12**	**32.33**	**0.969**	-	-	-
SR3 [103] (80k)	14.22	25.33	0.780	155.29	155.3×100=15530.0	0.02×100 = 2.00
SR3 [103] (500K)	13.38	26.03	0.820	155.29	155.3×100=15530.0	0.02×100 = 2.00
SR3 [103] (1000K)	11.61	27.49	0.871	155.29	155.3×100=15530.0	0.02×100 = 2.00
Ours (only res, 80k)	4.76	30.72	0.959	**7.74**	**33.5×5 = 167.7**	**0.03×5 = 0.16**
Ours (80k)	4.67	30.91	0.962	15.49	67.1×5 = 335.5	0.06×5 = 0.32

（c）去雨	PSNR (↑)	SSIM (↑)	Params (M)	MAC (G) × Steps	Inference Time (s)
RainDiff64[28]	32.29	0.9422	109.68	252.4×10 = 2524.2	0.03×10 = 0.38
RainDiff128[28]	32.43	0.9334	109.68	248.4×50 = 12420.0	0.038×50 = 1.91
Ours (only res)	31.96	0.9509	**7.73**	**32.9×5 = 164.7**	**0.032×5 = 0.16**
Ours	**32.51**	**0.9563**	15.47	65.8×5 = 329.3	0.07×5 = 0.35

✓ **J. Liu**, Q. Wang, H. Fan, Y. Wang, Y. Tang, L. Qu. Residual Denoising Diffusion Models[C]. The IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024. **(顶级学术会议|CCF-A|CVPR 2024)**

5 学术影响力

- **成果：IEEE/CVF Conference on Computer Vision and Pattern Recognition (人工智能国际顶级学术会议|CCF A 类会议)**
- **谷歌学术引用140次**
  - 被国内外一些知名学者发表于**顶级学术期刊和会议论文（IEEE TPAMI、IEEE CVPR、NeurIPS、AAAI等）引用**且正面评价。
  - 引用本成果的论文作者分别来自**剑桥大学、牛津大学、新加坡国立大学、帝国理工学院、佐治亚理工学院、北京大学、上海交通大学、中国科学技术大学、自动化所、微软亚洲研究院、IBM研究院、腾讯优图实验室**等。
- **Github Star 561次**
  - 训练、评估、模型代码全部开源
  - Github：https://github.com/nachifur/RDDM
  - YouTube：https://www.youtube.com/watch?v=E-ObZs32fEU

5 学术影响力

➢ 重要学术评价

IEEE TPAMI 2025

Diffusion Models in Low-Level Vision: A Survey

Chunming He, Yuqi Shen, Chengyu Fang, Fengyang Xiao, Longxiang Tang, Yulun Zhang, Wangmeng Zuo, *Senior Member, IEEE,* Zhenhua Guo, Xiu Li

ditional guidance based on Residual Denoising Diffusion Models [211] to improve image restoration performance. Improving the internal mechanisms of deep learning to better learn the distribution of multi-task degradations represents a promising direction for future DM-based explorations.

[211] J. Liu, Q. Wang, H. Fan, Y. Wang, Y. Tang, and L. Qu, "Residual denoising diffusion models," in *CVPR*, 2024.

本成果**被IEEE TPAMI 2025**的综述文章（Diffusion Models in Low-Level Vision: A Survey）**收录**，肯定了基于残差去噪扩散模型（RDDM 模型）的共享分布映射与条件引导方法，**为未来基于扩散模型学习多任务退化分布的探索指明前景方向。**

NeurIPS 2024

Resfusion: Denoising Diffusion Probabilistic Models for Image Restoration Based on Prior Residual Noise

Zhenning Shi [1] Haoshuai Zheng [1] Chen Xu [1] Changsheng Dong [1] Bin Pan [2] Xueshuo Xie [3] Along He [1*] Tao Li [1,3*] Huazhu Fu [4]

recovery in the latent space. Liu et al. [17] introduced the Residual Denoising Diffusion Models (RDDM), generalizing the diffusion process of InDI and I^2SB. RDDM points out that co-learning the residual term and the noise term can effectively improve the model performance. However, RDDM

including shadow removal, low-light enhancement and deraining. For fair comparisons, we use an U-net [28] structure which is the **same** as RDDM as the backbone. We simply concatenate x_t and $\hat{x}_0$

7, 17, 19, 45] whenever possible. For all image restoration tasks, we used an identical U-net as the backbone, which is the same as RDDM [17]. We take the shadow images and shadow masks together

term (I^2SB [15]). Similar to RDDM [17], Resfusion simultaneously predicts both the residual term and the noise term, and provides the quantitive relationship between them.

RainDropDiff [5], and RDDM(SM-Res-N)[17]. We conduct experiments on two settings for the Raindrop dataset for fairness, following the methods employed in RDDM [17] and WeatherDiff [5]:

U-net to predict resnoise, outperforming RDDM with two U-nets to predict the residual term and the noise separately in terms of PSNR and SSIM. Resfusion use **half** the number of parameters of RDDM and achieved better quantitative evaluation metrics. For the LOL dataset, under the **same** parameters, Resfusion outperforms RDDM in terms of PSNR (+18%) and LPIPS (-40%) significantly. Furthermore, for all datasets, we employed a simple truncated linear schedule, while RDDM utilized

发表在NeurIPS 2024的论文将本成果作为**对比基准**，进行了**全面的对比**。

5 学术影响力

➢ 重要学术评价

CVPR 2024

Selective Hourglass Mapping for Universal Image Restoration Based on Diffusion Model

Dian Zheng[1] Xiao-Ming Wu[1] Shuzhou Yang[2] Jian Zhang[2] Jian-Fang Hu[1] Wei-Shi Zheng[1,3*]

3.1. Revist the condition mechanism of RDDM

ated way (*i.e.*, concatenation); RDDM [28] explicitly fuses the condition into the diffusing algorithm, achieving high image quality. However, the problem of multi-partite map-

age) into the forward process and achieves impressive performance on several image restoration benchmarks.

WeiShi Zheng|IAPR Fellow|中山大学教授

受本成果的启发，提出了DiffUIR模型，评价本文的工作“自然地将退化图像整合到扩散过程中，在几个图像恢复基准上**取得了令人印象深刻的性能**。”

NeurIPS 2024

DF40: Toward Next-Generation Deepfake Detection

Zhiyuan Yan[1], Taiping Yao[2†], Shen Chen[2], Yandan Zhao[2], Xinghe Fu[2], Junwei Zhu[2], Donghao Luo[2], Chengjie Wang[2], Shouhong Ding[2], Yunsheng Wu[2], Li Yuan[1†]

30. RDDM [45] RDDM, or Residual Denoising Diffusion Model, is a very recent advanced diffusion model that was published in CVPR 2024. It aims to improve the robustness and stability of the image generation process. By separating the conventional single denoising diffusion process

into residual and noise diffusion components, RDDM establishes a dual diffusion framework. This framework employs residual diffusion to model directed degradation from a target image to a degraded input, guiding restoration processes, while noise diffusion accounts for stochastic perturbations. RDDM's innovation lies in its ability to balance certainty (via residuals) and diversity (through noise),

[45] Jiawei Liu, Qiang Wang, Huijie Fan, Yinong Wang, Yandong Tang, and Liangqiong Qu. Residual denoising diffusion models. *arXiv preprint arXiv:2308.13712*, 2023.

发表于NeurIPS 2024的论文DF40: Toward Next-Generation Deepfake Detection评价本成果为“**a very recent advanced diffusion model**”，“能够**平衡确定性和多样性**”取得了一定的学术影响力。

6 应用案例与未来展望

➢ 国内外同行学者开展针对本成果的跟进、改进或拓展应用工作。

全色图像锐化

Accelerated Diffusion via High-Low Frequency Decomposition for Pan-Sharpening

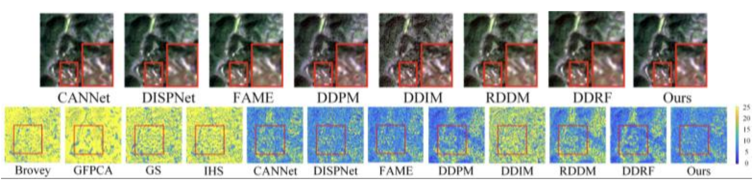
Ge Meng[1,2], Jingjia Huang[1,2], Jingyan Tu[1,2], Yingying Wang[1,3], Yunlong Lin[1,2], Xiaotong Tu[1,2], Yue Huang[1,2,3], Xinghao Ding[1,2,3*]

ity models. RDDM (Liu et al. 2023) introduces residuals to guide the reverse diffusion process. DDRF (Cao et al.

Meng, and Ermon 2020), RDDM (Liu et al. 2023), and DDRF (Cao et al. 2023). For fairness, we also consider the PAN image as the input condition for other diffusion-based

RDDM	40.8172	0.9702	0.0277	1.0795	30.1007	0.9225	0.0841	3.2946	44.6996	0.9855	0.0163	0.7751
DDRF	41.2331	0.9659	0.0240	0.9926	30.0801	0.9105	0.0825	3.2783	45.3347	0.9790	0.0144	0.6799
Ours	**42.5780**	**0.9735**	**0.0203**	**0.8432**	**31.0295**	**0.9298**	**0.0683**	**2.9475**	**48.7757**	**0.9901**	**0.0094**	**0.4632**

Metrics	GS	Brovey	IHS	GFPCA	CANNet	DISPNet	FAME	DDPM	DDIM	RDDM	DDRF	Ours
$D_\lambda\downarrow$	0.0696	0.1378	0.0770	0.0914	0.0861	0.0671	0.0674	0.0628	0.0714	0.0685	0.0804	**0.0611**
$D_S\downarrow$	0.2456	0.2605	0.2985	0.1635	0.1144	0.1826	0.1121	0.1136	0.2407	0.1055	0.1100	**0.1018**
QNR↑	0.7025	0.6390	0.6485	0.7615	0.7884	0.7638	0.8291	0.8319	0.6896	0.8345	0.8200	**0.8462**



发表于AAAI 2025的论文**将本成果用于遥感领域**全色图像的多光谱图像融合，作为**对比基准**，进行**全面的对比**。

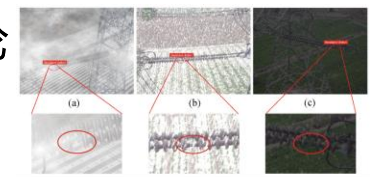
绝缘子缺陷检测

Insulator Defect Detection via a Residual Denoising Diffusion Mechanism

Li Zhang[1,2,3], Mengyang Song[2], Huaping Guo[2,*], Yange Sun[2] and Xinxia Wang[2]

detection and highly noisy environments. To overcome this limitation, we introduce the Residual Denoising Diffusion Mechanism (RDDM) [11], which is specifically designed for insulator defect detection under challenging conditions. The RDDM dynamically adjusts

发表于Materials的论文**基于本成果**进行绝缘子缺陷检测。



水下图像增强

WEDM: Wavelet-Enhanced Diffusion with Multi-Stage Frequency Learning for Underwater Image Enhancement

Junhao Chen[1], Sichao Ye[2], Xiong Ouyang[3] and Jiayan Zhuang[2,*]

To address these challenges and enhance underwater image quality, we propose a novel UIE framework based on the Conditional Residual Denoising Diffusion Model (RDDM), named the WEDM. This framework combines frequency domain information with

基于残差去噪扩散模型进行水下图像增强。

6 应用案例与未来展望

➢ 国内外同行学者开展针对本成果的跟进、改进或拓展应用工作。

心电图测量

EDDM: A Novel ECG Denoising Method Using Dual-Path Diffusion Model

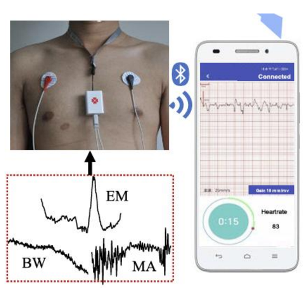
Zhiyuan Li, Yuanyuan Tian, Yanrui Jin, Xiaoyang Wei, Mengxiao Wang, Jinlei Liu, and Chengliang Liu

with e_θ, ε_θ. Refer to [35] and [36], we derive the relationship between x_{t-1} and x_t. The detailed derivation process can be

2) *Implementation Details:* In the EDDM training phase, we generated the ECG noise e diffusion coefficient sequence α_t with a "linear decrease" strategy, and the white noise ε diffusion coefficient sequence with a "linear decrease" strategy [35]. We set the number of diffusion steps to 50,

[35] J. Liu, Q. Wang, H. Fan, Y. Wang, Y. Tang, and L. Qu, "Residual denoising diffusion models," 2023, *arXiv:2308.13712*.

发表于IEEE TIM 2025的论文**基于本成果**提出了心电图去噪扩散模型，具有更稳定和更高性能的降噪结果，**在可穿戴心电图测量系统中的巨大潜力**。



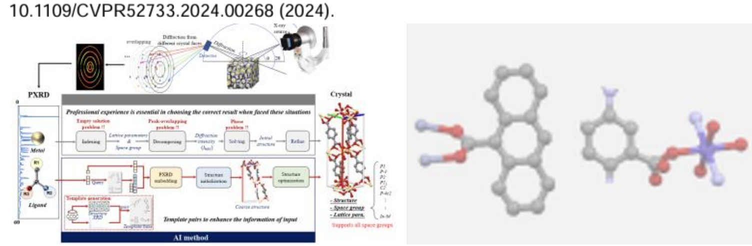
材料科学-晶体结构

A Powder Diffraction-AI Solution for Crystalline Structure

Di Wu[1,2,3#], Pengkun Wang[1,2,3#], Shiming Zhou[3,4], Bochun Zhang[1,4], Liheng Yu[3], Xi Chen[3], Xu Wang[3], Zhengyang Zhou[3], Yang Wang[1,2,3✉], Sujing Wang[1,3,4,5✉], Jiangfeng Du[1,4,5]

matched templates. The second stage applies a diffusion-based residual network[24,25] to refine the structure at the atomic level, incorporating crystal geometry and chemical constraints.

25 Liu, J., et al. Residual denoising diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, doi:2773-2783, 10.1109/CVPR52733.2024.00268 (2024).



精准智能化学全国重点实验室的中科大汪苏靖团队**基于本成果的残差去噪扩散模型**，在结合了晶体几何和化学约束下完善原子水平的结构，确定了有机和无机杂化**晶体固体结构**。

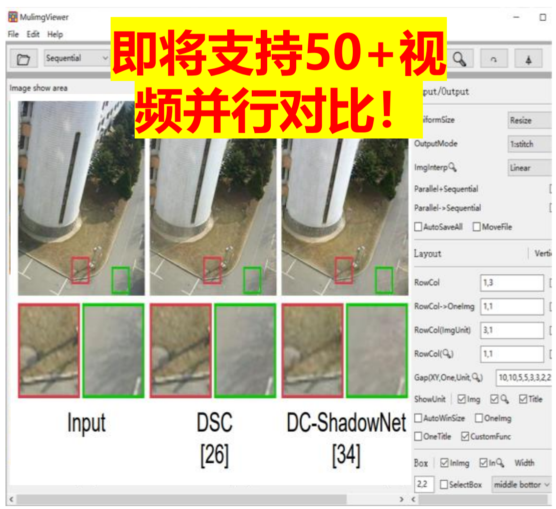
7 开源软件

➢ 重塑多图像浏览体验下一代浏览器-MulimgViewer

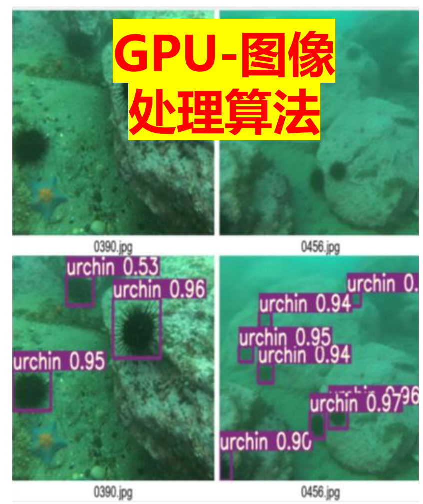
- 独立研发的多图像浏览器，被国内外**顶级机构的研究人员广泛使用**，如清华大学、北京大学、UC Berkeley、帝国理工、Meta、字节、腾讯、旷世、华为
- 多图像并行显示/放大/比较(**1000+**)、超分率图像无损对比（**4K/8K**）
- 即将推出：多视频并行对比（**50+**）、**多算法并行显示**

多图像浏览器

1万+下载



多图像并行显示|放大|拼接



图像显示前端

多维数据可视化

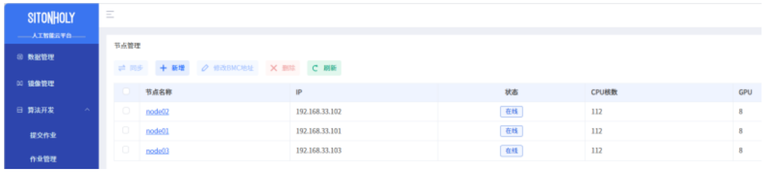
Github **1.3k Star，1万+下载**| 网站：https://github.com/nachifur/MulimgViewer

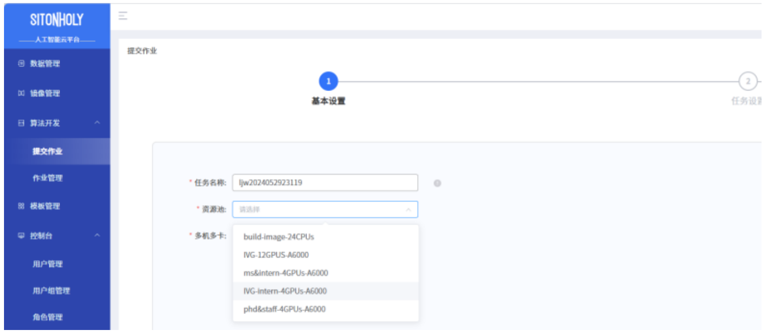
8 科研合作

➢ SIA-AIG人工智能研究小组

- **福利**：充足的GPU资源|专业论文写作指导|idea交流
- 本科生：211/985保研->中国科学院大学-沈阳自动化研究所
- 研究生：与东北大学、沈理工、沈工业等高校联合培养
- 实习生：**有激情！** 6月以上|编程强|有论文发表和科研经历优先

研究小组集群软硬件





✓ 新购置3台8卡A6000服务器

Intelligence & Robotics 期刊首届青年学术论坛

探索统一的图像到图像分布变换框架

--- 残差去噪扩散模型

汇报人： 刘佳伟 博士 助理研究员

单　位： 中国科学院沈阳自动化研究所

邮　箱： liujiawei@sia.cn

衷心感谢各位专家！