



中国科学院大学
University of Chinese Academy of Sciences

Residual Denoising Diffusion Models

Jiawei Liu^{1,4}, **Qiang Wang**^{1,3}, **Huijie Fan**^{1*}, **Yinong Wang**²,
Yandong Tang¹, **Liangqiong Qu**^{2*}

¹State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences ²The University of Hong Kong

³Shenyang University ⁴University of Chinese Academy of Sciences

Code: <https://github.com/nachifur>

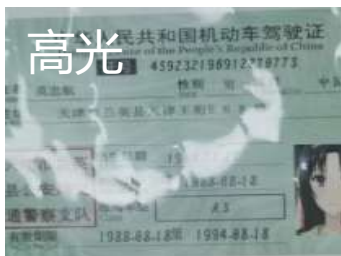
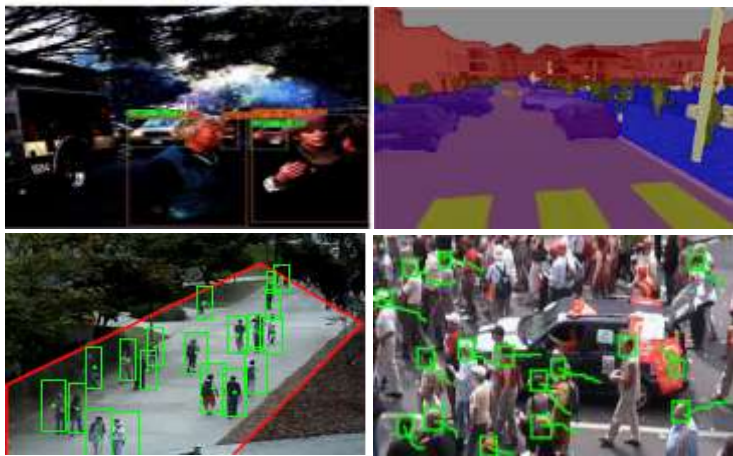


1 研究背景

➤ 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？

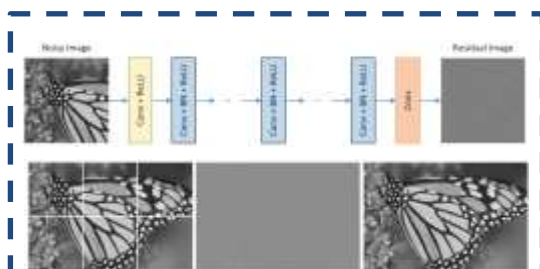


2 研究现状

➤ 图像恢复方法研究现状

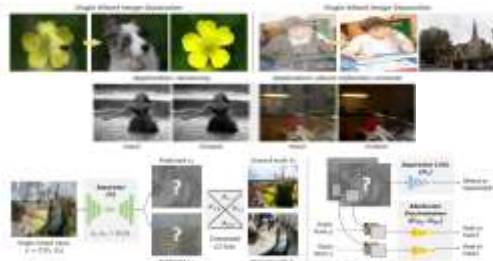
- **技术思路:** 基于CNN、基于GAN、基于Transformer恢复退化图像
- **存在问题:** 不同任务需设计不同的网络结构、损失函数和训练策略
- **代表方法:**

基于CNN[1]



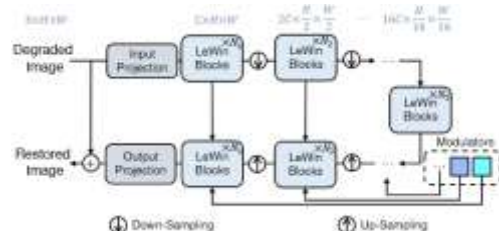
- ✓ 哈尔滨工业大学
- ✓ 利用残差学习实现了对高斯去噪、超分辨率和JPEG去块等去噪任务

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



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）



目标: $I_0 \sim q(I_0)$ ← $p_\theta(I_{0:T}) := p_\theta(I_T) \prod_{t=1}^T p_\theta(I_{t-1}|I_t)$ ——— 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: $\sqrt{\alpha_t} I_{t-1} + \sqrt{1 - \alpha_t} \epsilon_{t-1} \xrightarrow{\text{add noise}} I_t \longrightarrow \epsilon_\theta(I_t) \xrightarrow{\text{Bayes}} I_0^\theta \xrightarrow{p_\theta(I_{t-1}|I_t)} \xrightarrow{\text{sample}} I_{t-1}$

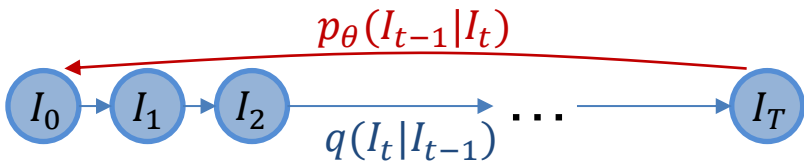


4 残差去噪扩散模型

- 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

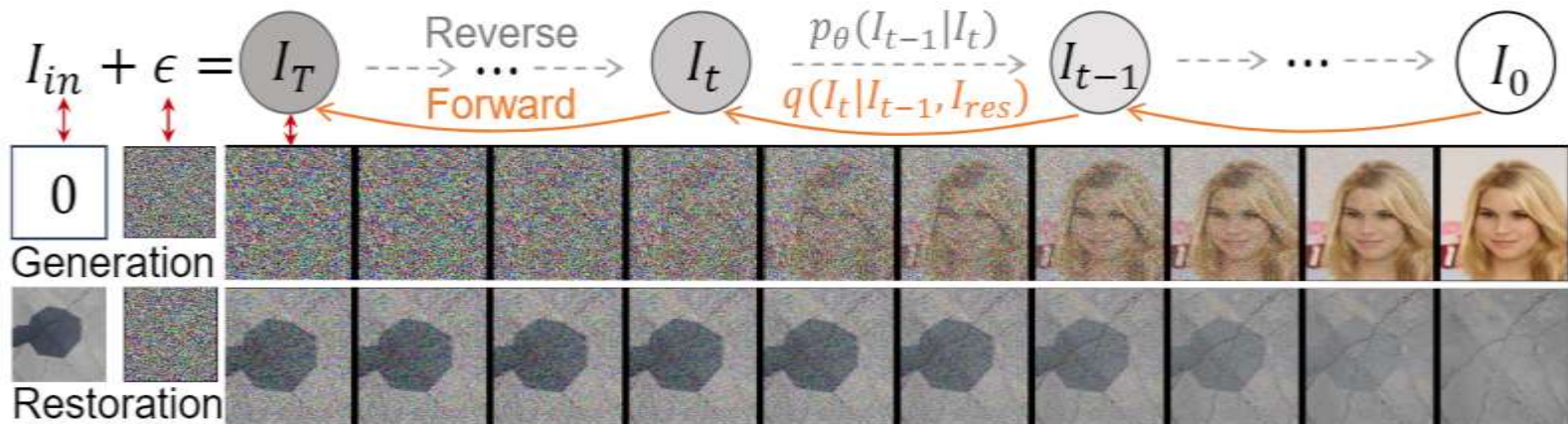
$$\begin{aligned} I_t &= \sqrt{\alpha_t} I_{t-1} + \sqrt{1 - \alpha_t} \epsilon_{t-1} \\ &= \sqrt{\alpha_t} (\sqrt{\alpha_{t-1}} I_{t-2} + \sqrt{1 - \alpha_{t-1}} \epsilon_{t-2}) + \sqrt{1 - \alpha_t} \epsilon_{t-1} \\ &= \dots \\ &= \sqrt{\bar{\alpha}_t} I_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon \end{aligned}$$



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：**

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



➤ **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**

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

➤ **扩散系数时间表转化：**

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

➤ **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	
	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 $\hat{\alpha}$ 和 $\hat{\beta}$, 初始可学习参数 $\lambda_{res}^0 = 0.5$, 深度网络 G , 参数为 θ , 初始学习率为 1, 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}^0 .

```

1  $\theta \leftarrow \text{InitWight}(G)$  ▷ 初始化网络参数
2 for  $i \leftarrow 1$  to  $n + m$  do
3    $t \leftarrow \text{Uniform}(\{1, 2, \dots, T\})$ ,  $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ ,  $I_{res} \leftarrow I_{in} - I_0$ 
4    $I_t \leftarrow I_0 + \hat{\alpha}_t I_{res} + \hat{\beta}_t \epsilon$  ▷ 通过公式(5-8)合成  $I_t$ 
5    $I_{out} \leftarrow G(I_t, t, I_{res})$ 
6    $I_{res}^0 \leftarrow \lambda_{res}^0 \times I_{out} + (1 - \lambda_{res}^0) \times f_{t \rightarrow res}(I_{out})$   $\triangleright f_{t \rightarrow res}(\cdot)$  表示从  $\epsilon$  到  $I_{res}$ 
   的转换通过公式(5-11)
7    $\epsilon_t \leftarrow \lambda_{res}^0 \times f_{res \rightarrow t}(I_{out}) + (1 - \lambda_{res}^0) \times I_{out}$   $\triangleright f_{res \rightarrow t}(\cdot)$  表示从  $I_{res}$  到  $\epsilon$ 
   的转换通过公式(5-11)
8    $\mathcal{L}_{adv} \leftarrow \text{Loss}(I_{res}^0, I_{res}, \epsilon_t, \epsilon)$  ▷ 基于公式(5-34)
9    $\theta, \lambda_{res}^0 \leftarrow -\nabla_{\theta, \lambda_{res}^0}(\mathcal{L}_{adv}, t)$  ▷ 梯度更新
10  if  $\text{abs}(\lambda_{res}^0 - 0.5) < \delta$  then ▷ 类似对抗的训练
11    guess ▷ 类似对抗的训练
12  else ▷ 停止梯度更新
13     $\lambda_{res}^0 \leftarrow \text{Detach}(\lambda_{res}^0)$  ▷ 重新初始化网络参数
14     $\theta \leftarrow \text{InitWight}(G)$  ▷ 重新初始化网络参数
15    if  $\lambda_{res}^0 > 0.5$  then ▷ SM-Res
16       $\lambda_{res}^0 \leftarrow 1$ 
17    else ▷ SM-N
18       $\lambda_{res}^0 \leftarrow 0$ 
19  end
20 end

```



4 残差去噪扩散模型

➤ 模型性质：求和约束的方差时间表

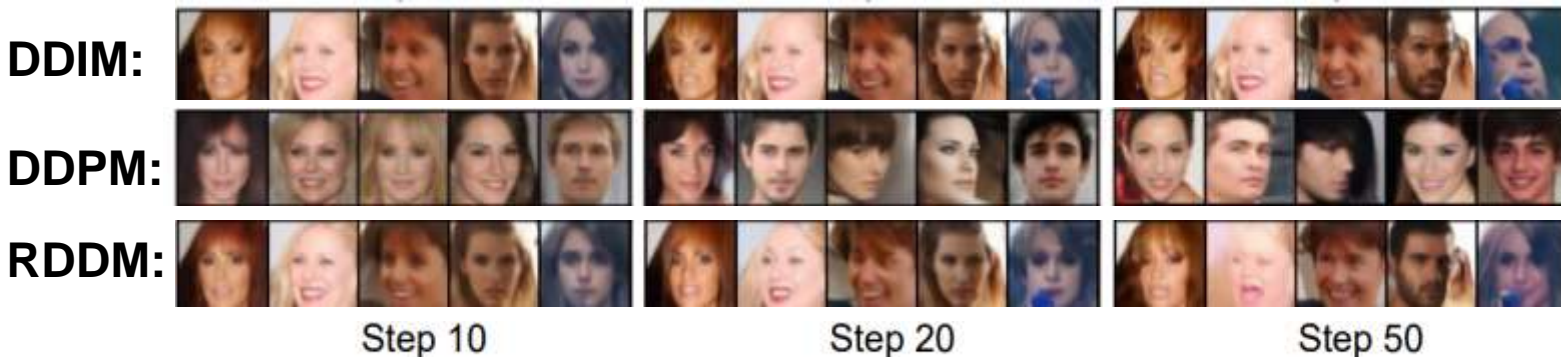
- DDPM每一步的方差为1；DDIM方差为0；
- RDDM是求和约束的方差，10步采样时每步方差=1/10，100步采样=1/100。

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

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

$$\sum_{i=1}^T \sigma_i^2(RDDM) = \sum_{i=1}^T \eta \beta_i^2 \frac{\beta_{i-1}^2}{\beta_i^2} \leq \sum_{i=1}^T \beta_i^2 \leq 1,$$

定性比较





4 残差去噪扩散模型

➤ 模型性质：解耦的前向扩散过程

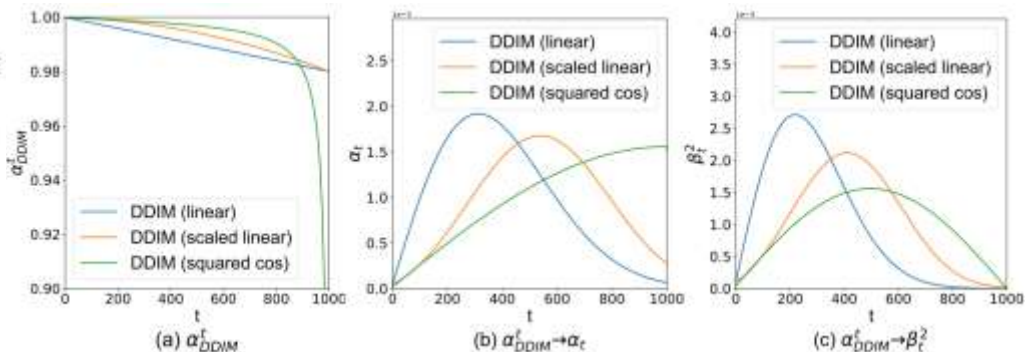
- α_t 控制残差扩散的速度； β_t 控制噪声扩散的速度；
- 更好的系数时间表： α_t 线性递减， β_t 线性递增。

扩散速度曲线

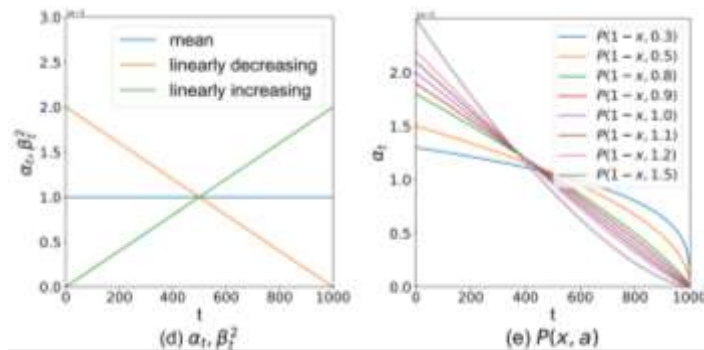
✓更清晰的物理解释

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

✓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)	<u>27.82</u>	2.26
α_t (linearly decreasing), β_t^2 (linearly increasing)	23.25	2.05





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_{DDIM}^t** 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_{DDIM}^t 导致**去噪失败**。

重新调整系数时间表

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



(a) DDIM (linear) Score:9.4 (b) $\alpha_{DDIM}^t \rightarrow \alpha_t, \beta_t^2$ Score:9.4 (c) $\alpha_{DDIM}^t \rightarrow$ scaled linear (d) $\alpha_{DDIM}^t \rightarrow$ squared cosine (e) $\alpha_t \rightarrow \alpha_t$
 $\beta_t^2 \rightarrow P(1-x, 1)$



$P(1-x, 0.3)$ Score:9.8 $P(1-x, 0.5)$ **Score:9.8** $P(1-x, 0.8)$ Score:9.7 $P(1-x, 1.0)$ Score:9.1 $P(1-x, 1.2)$ Score:9.3 $P(1-x, 1.5)$ Score:8.2
 (f) convert α_{DDIM}^t to α_t, β_t^2 and readjust the converted α_t without touching the β_t^2



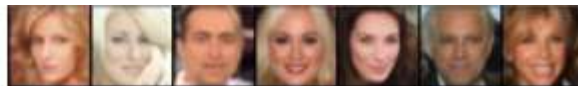
4 残差去噪扩散模型

➤ 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 $I_{res}^\theta(I_t, t, 0) \rightarrow I_{res}^\theta(I_t, \bar{\alpha}_t \cdot T, 0), \epsilon_\theta(I_t, t, 0) \rightarrow \epsilon_\theta(I_t, \bar{\beta}_t \cdot T, 0)$

部分路径独立的生成过程

(a) Training: DDIM (linear)



(b) Test: $\alpha_t \rightarrow \alpha_t, \beta_t^2 \rightarrow P(1-x, 1)$



(c) Test: $\beta_t^2 \rightarrow \beta_t^2, \alpha_t \rightarrow P(x, 0)$



Denoising ($\epsilon_\theta(I_t, \bar{\beta}_t \cdot T)$)

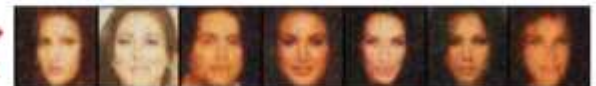
(d) Training: DDIM (linear)



(e) Test: $\alpha_t, \beta_t^2 \rightarrow P(x, 0)$



(f) Test: $\alpha_{DDIM}^t \rightarrow$ squared cosine



Denoising ($\epsilon_\theta(I_t, t)$) + Deresidual ($I_{res}^\theta(I_t, t)$)

(g) Training: DDIM (linear)



(h) Test: $\alpha_t, \beta_t^2 \rightarrow P(x, 0)$



(i) Test: $\alpha_t, \beta_t^2 \rightarrow P(1-x, 1)$



Path Independence Generation Process (Denoising ($\epsilon_\theta(I_t, \bar{\beta}_t \cdot T)$) + Deresidual ($I_{res}^\theta(I_t, \bar{\alpha}_t \cdot T)$)

(j) Test: $\alpha_t, \beta_t^2 \rightarrow P(1-x, 1.5)$



(k) Test: $\alpha_{DDIM}^t \rightarrow$ scaled linear



(l) Test: $\alpha_{DDIM}^t \rightarrow$ squared cosine



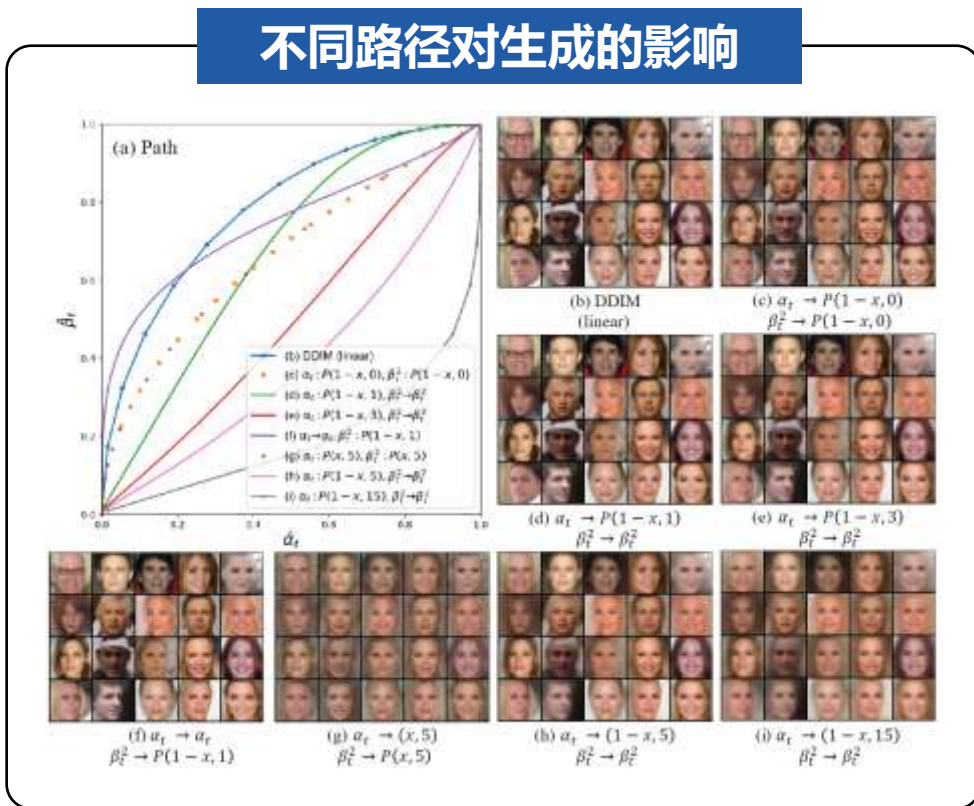


4 残差去噪扩散模型

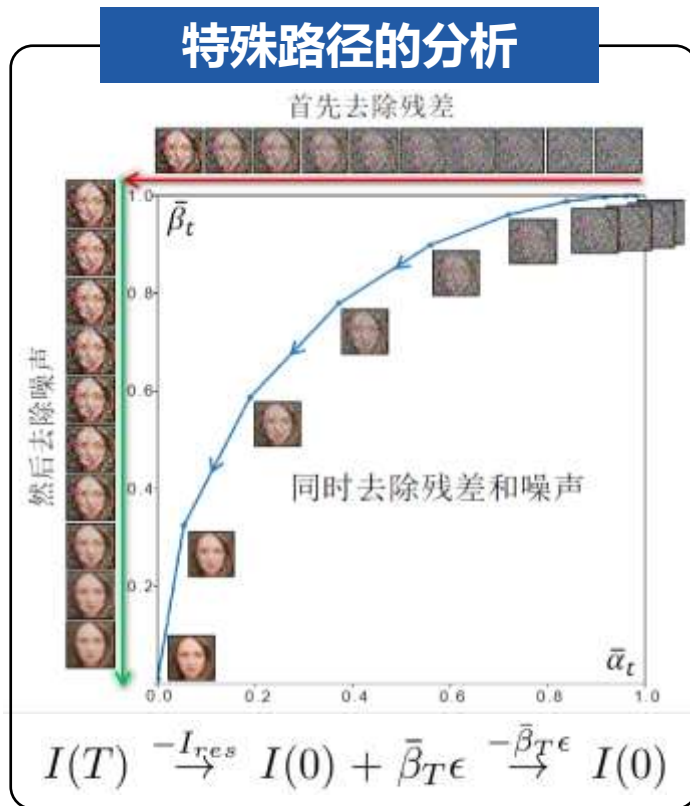
➤ 模型性质：部分路径独立的生成过程

- 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，
$$\frac{\partial I_{res}^\theta(I(t), \bar{\alpha}(t) \cdot T)}{\partial \bar{\beta}(t)} \approx 0, \quad \frac{\partial \epsilon_\theta(I(t), \bar{\beta}(t) \cdot T)}{\partial \bar{\alpha}(t)} \approx 0.$$

不同路径对生成的影响



特殊路径的分析





4 残差去噪扩散模型

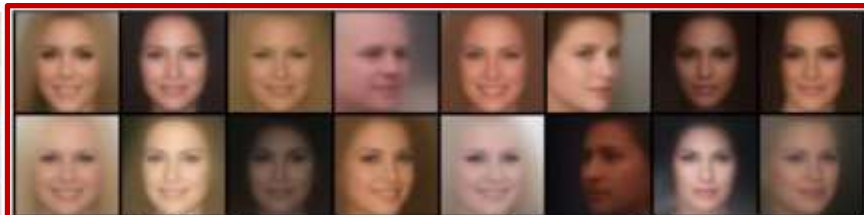
➤ 模型性质：部分路径独立的生成过程

- 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， $\frac{\partial I_{res}^\theta(I(t), \bar{\alpha}(t) \cdot T)}{\partial \bar{\beta}(t)} \approx 0, \quad \frac{\partial \epsilon_\theta(I(t), \bar{\beta}(t) \cdot T)}{\partial \bar{\alpha}(t)} \approx 0.$

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]	DSLRL [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 ($\alpha_{DDIM}^t \rightarrow \alpha_t, \beta_t^2$)	(b) Shadow Removal	MSE(↓)			SSIM(↑)			PSNR(↑)		
				S	NS	ALL	S	NS	ALL	S	NS	ALL
5 steps	69.60	69.60	DSC [39]	9.48	6.14	6.67	0.967	-	-	33.45	-	-
10 steps	40.45	40.41	FusionNet [40]	7.77	5.56	5.92	0.975	0.880	0.945	34.71	28.61	27.19
15 steps	32.67	32.71	BMNet [41]	7.60	4.59	5.02	0.988	0.976	0.959	35.61	32.80	30.28
20 steps	30.61	30.77	DMTN [5]	7.00	4.28	4.72	0.990	0.979	0.965	35.83	33.01	30.42
100 steps	23.66	24.92	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 (↓)	(d) Deraining	PSNR(↑)	SSIM(↑)	(e) Deblurring	PSNR(↑)	SSIM(↑)
Retinex-Net [42]	16.774	0.462	0.474	pix2pix [43]	28.02	0.8547	Nah <i>et al.</i> [44]	29.08	0.914
KinD [45]	17.648	0.779	0.175	DuRN [46]	31.24	0.9259	Zhang <i>et al.</i> [47]	29.19	0.931
KinD++ [48]	17.752	0.760	0.198	RainAttn [49]	31.44	0.9263	DeblurganV2 [50]	29.55	0.934
RUAS [51]	18.230	0.720	0.350	AttnGAN [52]	31.59	0.9170	Gao <i>et al.</i> [53]	30.90	0.935
KinD++-SKF [54]	20.363	0.805	0.201	IDT [55]	31.87	0.9313	Suin <i>et al.</i> [56]	31.85	0.948
DCC-Net [57]	22.72	0.81	-	RainDiff64 [28]	32.29	0.9422	MPRNet [58]	32.66	0.959
SNR-Aware [59]	24.608	0.840	0.151	RainDiff128 [28]	32.43	0.9334	Uformer-B [60]	32.97	0.967
Ours (RDDM)	25.392	0.937	0.134	Ours (RDDM)	32.51	0.9563	Ours (RDDM)	32.40	0.963

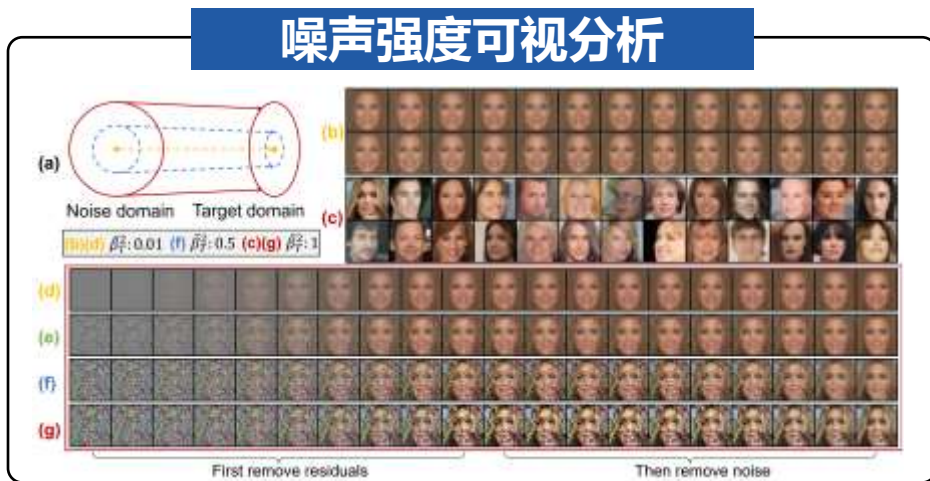


4 残差去噪扩散模型

➤ 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
$\beta_T = 0.01$	MAE-ALL (\downarrow)	4.83	4.69	4.67	4.72	4.90
	PSNR-S (\uparrow)	36.83	36.98	36.74	36.59	36.41
	LPIPS (\downarrow)	0.0344	0.0308	0.0305	0.0314	0.0334
$\beta_T = 1$	MAE-ALL (\downarrow)	5.07	4.94	4.90	4.87	4.99
	PSNR-S (\uparrow)	36.93	37.20	37.07	37.01	36.62
	LPIPS (\downarrow)	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，RDDDM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RDDDM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RDDDM仅需要 4.8G GPU 内存进行训练

运行效率结果

(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 (RDDDM)	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

更快训练
更快推理
更少参数
更高性能

- ✓ 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 未来展望

未来的研究思路

- 对于未知新任务，**自适应学习噪声强度**，采用**可学习系数时间表**减少采样步骤，提高图像生成质量；
- 使用**一组预训练参数**处理多个不同的图像恢复任务(中山大学-CVPR24 DiffUIR)；
- 将残差扩散模型用于**文生图和文生视频**，通过**蒸馏**获得残差预测网络；
- 构建**统一多任务表征学习和多模态感知**系统，利用高层语义指导图像恢复，通过底高层视觉协同学习鲁棒特征表示（高/底层视觉协同）。

多模态（文本语音）、多任务、大模型、统一、通用、mamba等

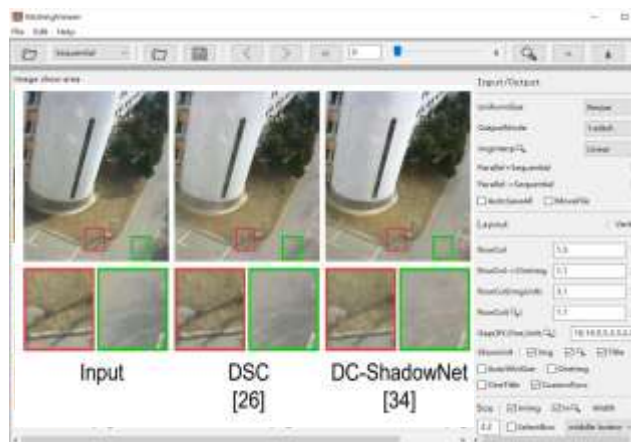


6 开源软件-MulimgViewer

➤ 项目网站: <https://github.com/nachifur/MulimgViewer>

- 并行放大比较、跨平台、中英文档
- 1.1k+Star|6.9k+ 下载|200+人社区|7 人参与贡献
- 已被国内外很多**顶级大学、公司、研究机构的研究人员使用**，如：中
科大、清北、南理、UC Berkeley、纽约大学石溪分校、南洋理工、帝
国理工、字节、腾讯、旷世、沈自所、光电所等

多图像浏览器



多张图像并行显示、放大、拼接



1.1k+stars|6.9k+ 下载

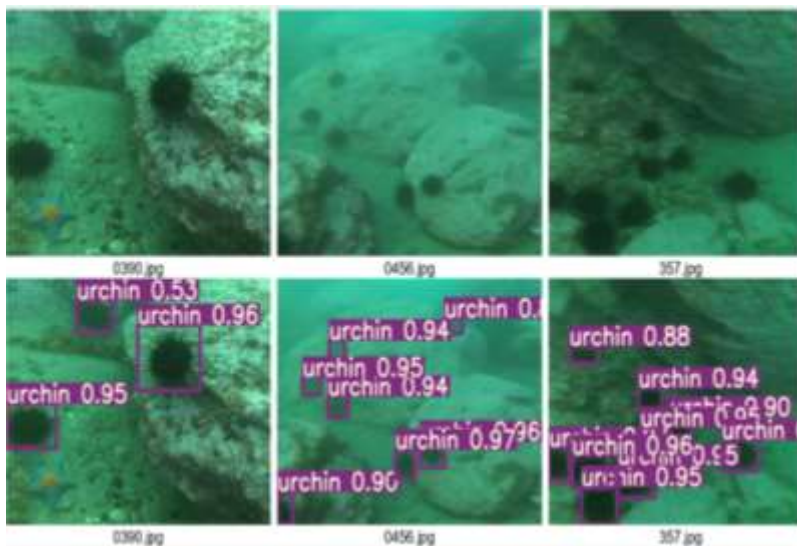


创建并维护开源项目社区



6 开源软件-MulimgViewer

图像显示前端



- ✓ 支持用户任何定制的python图像处理算法，如：使用CPU的传统图像处理算法，或者使用GPU的深度神经网络算法（检测、识别、分割等）

多维数据的可视化



- ✓ 为欧洲海洋地质项目Geoswim Project提供多维数据的可视化解决方案（水上图像、水下图像、经纬度、温度、压强等）。

- ✓ <https://github.com/nachifur/MulimgViewer/issues/74>
- ✓ <https://github.com/nachifur/MulimgViewer/issues/57>
- ✓ <https://youtu.be/-3oI1gSRG3c>

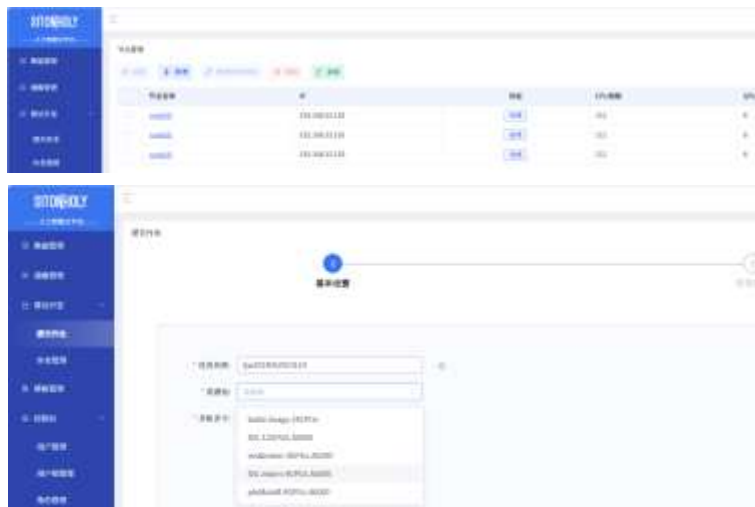


7 保研、联培、实习生、论文合作等

➤ SIA-IVG智能视觉研究小组

- **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、沈化工等高校联合培养
- **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



中国科学院大学
University of Chinese Academy of Sciences

Residual Denoising Diffusion Models

汇报人：刘佳伟 (liujiawei@sia.cn)

导师：唐延东 研究员

中国科学院沈阳自动化研究所
机器人学国家重点实验室
机器人视觉课题组



微信: nachifur



欢迎大家交流!

