北京大學
PEKING UNIVERSITY

Employing Diffusion Models for Low-Level Vision

Jiaying Liu

Wangxuan Institute of Computer Technology
Peking University

Visual Degradation

Data Acquisition

Degradation before Data Acquisition

- Heavy Rain/Snow
- Underwater
- Low Light
- Haze/Sandstorm

Degradation in Data Acquisition

- Downsample
- Motion Blur
- System Noise
- Optical Distortion

Degradation after Data Acquisition

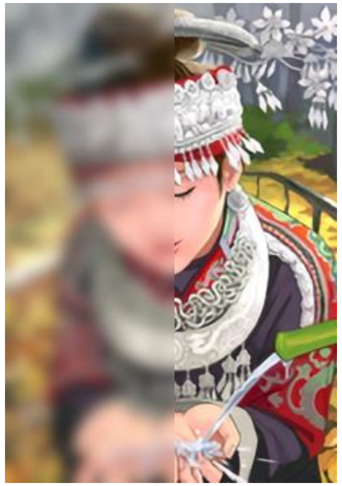
- Scratches
- Watermark
- Package Loss
- Compression Loss

Image Restoration and Enhancement

Underwater Enhancement

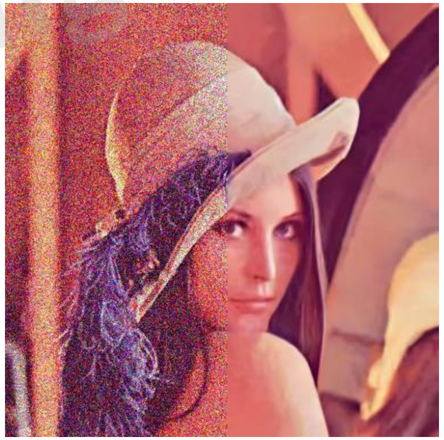
Dehazing

Text Removal



Super Resolution

Rain Streak Removal



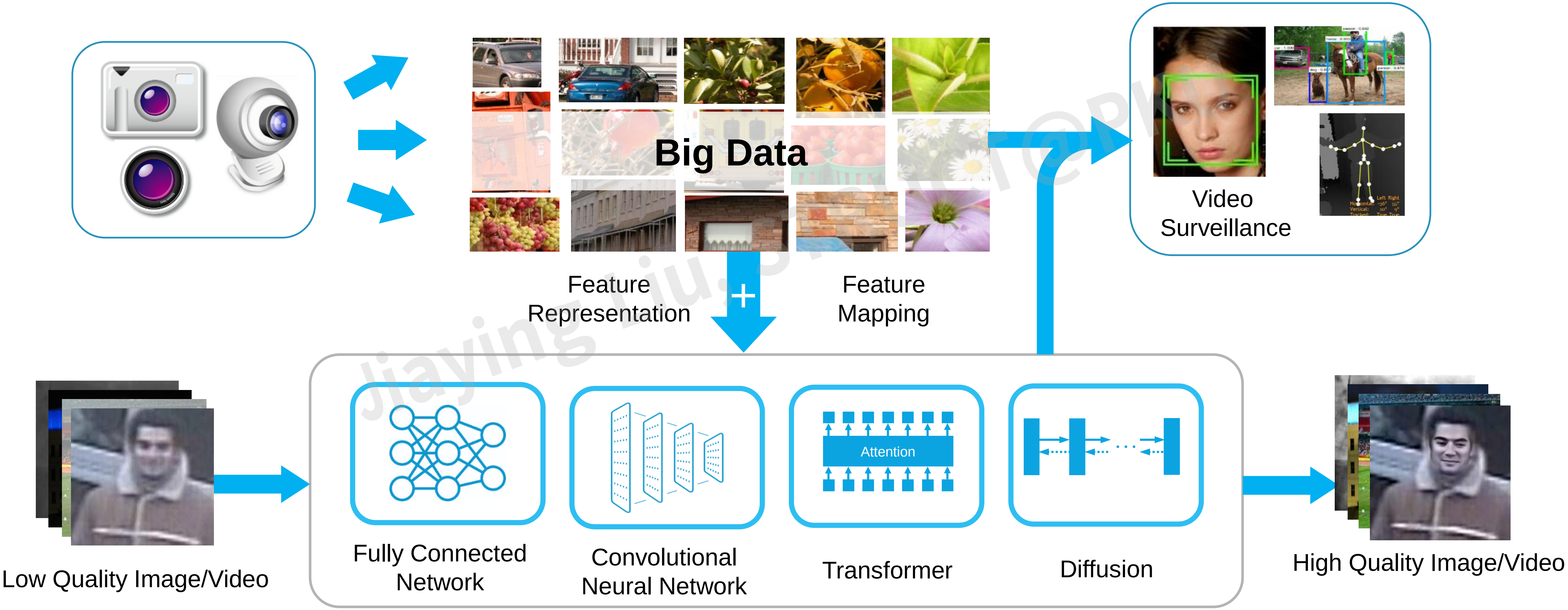
Denoising

Rain Drop Removal

Low Light Enhancement

Image Restoration and Enhancement

- Data-Driven Solution



Data-Driven Solution

Diffusion for Image Restoration

Diffusion for Image Restoration (IR)

Diffusion which creates visual details with high quality can play a role.

Text:	A gray sketch on paper of a Ferrari car, full car, pencil art

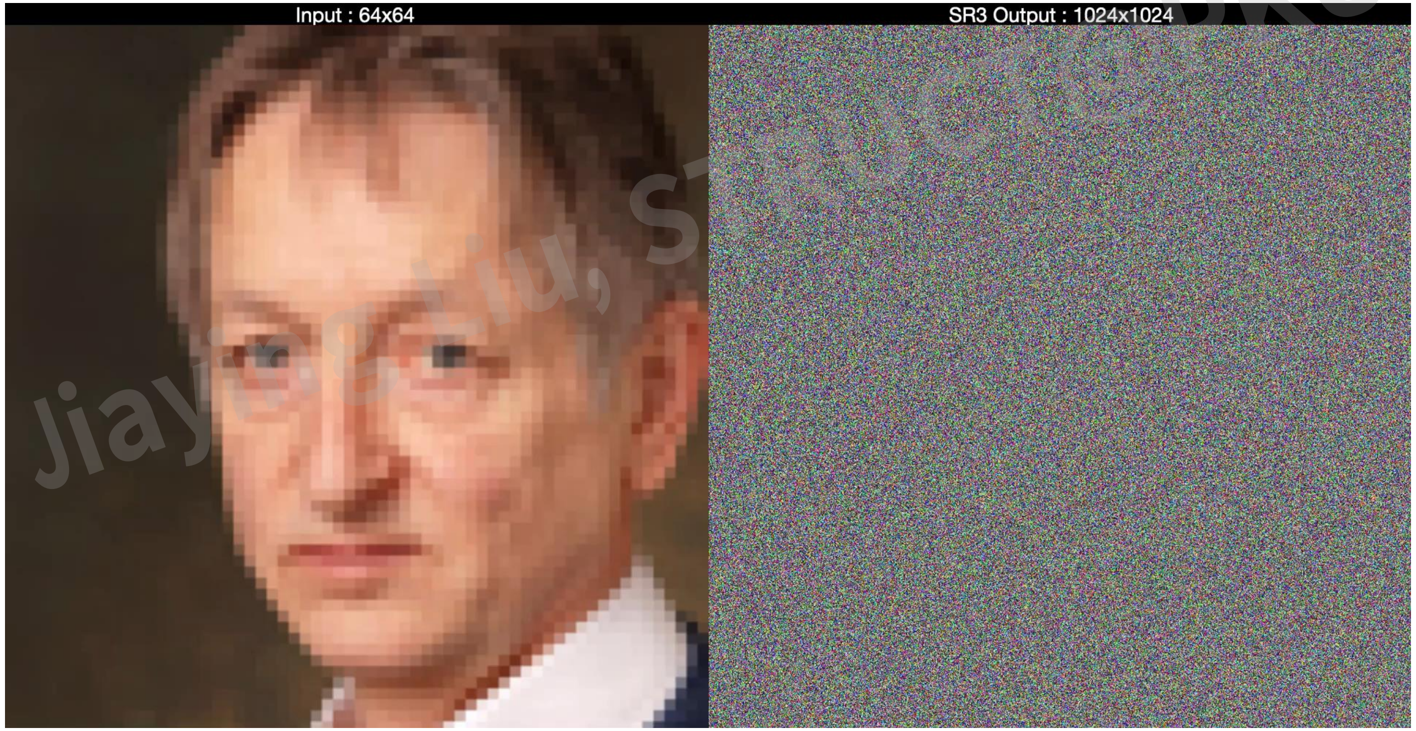
Image synthesis

[1] https://prog.world/sherudim-under-the-hood-of-stable-diffusion/

Diffusion for Image Restoration

■ Diffusion for Image Restoration (IR)

Diffusion which creates visual details with high quality can play a role.



Restoration

[2] Chitwan Saharia, et al. Image Super-Resolution via Iterative Refinement, IEEE TPAMI, 2022

01

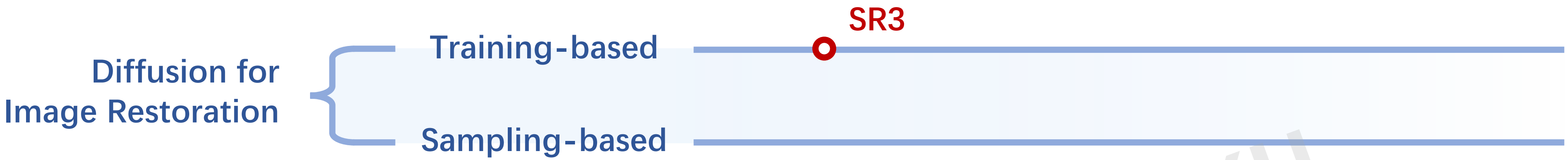
Diffusion for Image Restoration

Diffusion for Image Restoration – SISR

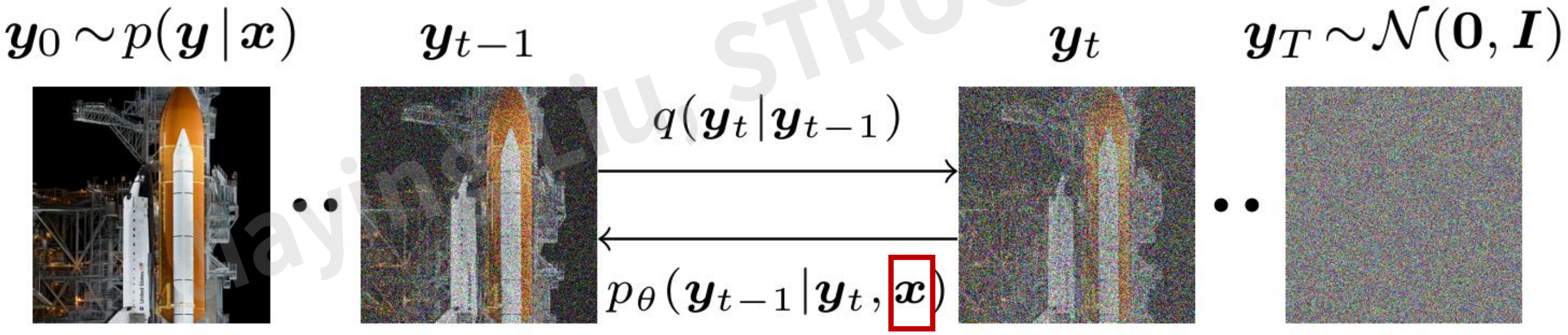
Diffusion for
Image Restoration

- Training-based
- Sampling-based

Diffusion for Image Restoration – SISR



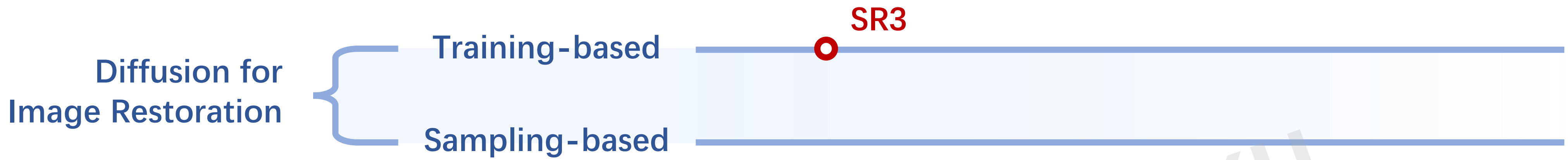
- Image Super-Resolution via Repeated Refinement (SR3)



  - **Conditional** denoising diffusion model
  - Forward diffusion process $\boldsymbol{q}$, reverse inference process $\boldsymbol{p}$
  - Source image $\boldsymbol{x}$, target image $\boldsymbol{y_0}$

[2] Chitwan Saharia, et al. Image Super-Resolution via Iterative Refinement, IEEE TPAMI, 2022

Diffusion for Image Restoration – SISR



- Image Super-Resolution via Repeated Refinement (SR3)

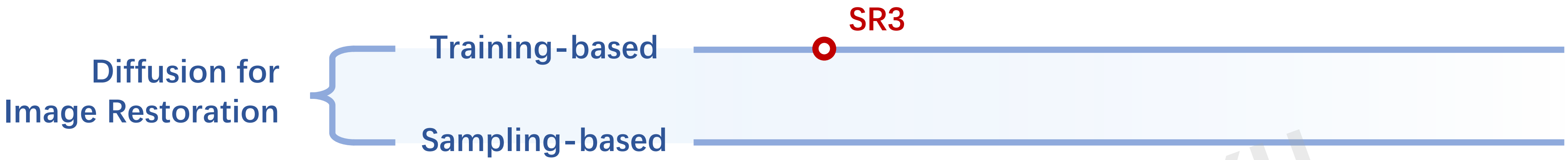
Conditioned reverse inference process

$$
\begin{aligned}
p_\theta(\boldsymbol{y}_{0:T}|\boxed{\boldsymbol{x}}) &= p(\boldsymbol{y}_T)\prod_{t=1}^{T} p_\theta(\boldsymbol{y}_{t-1}|\boldsymbol{y}_t,\boldsymbol{x}) \\
p(\boldsymbol{y}_T) &= \mathcal{N}(\boldsymbol{y}_T \mid \mathbf{0}, \boldsymbol{I}) \\
p_\theta(\boldsymbol{y}_{t-1}|\boldsymbol{y}_t,\boxed{\boldsymbol{x}}) &= \mathcal{N}(\boldsymbol{y}_{t-1} \mid \mu_\theta(\boldsymbol{x},\boldsymbol{y}_t,\gamma_t), \sigma_t^2\boldsymbol{I})
\end{aligned}
$$

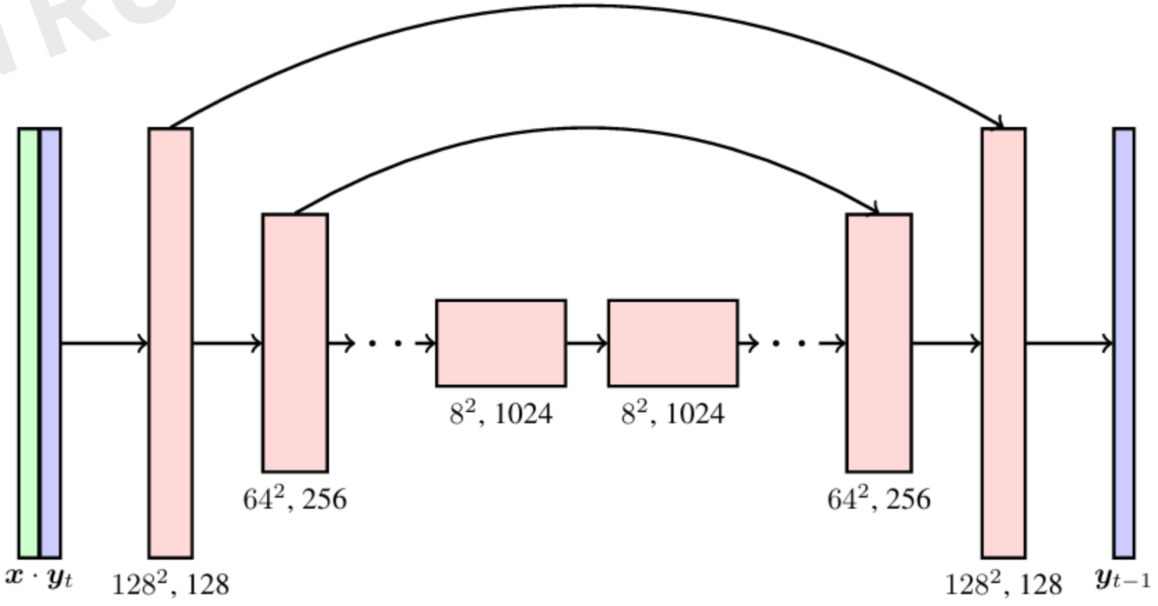
  - The reverse inference process conditioned on source image $\boldsymbol{x}$
  - $p_\theta(\boldsymbol{y}_{t-1}|\boldsymbol{y}_t,\boldsymbol{x})$ learned by the neural network

[2] Chitwan Saharia, et al. Image Super-Resolution via Iterative Refinement, IEEE TPAMI, 2022

Diffusion for Image Restoration – SISR

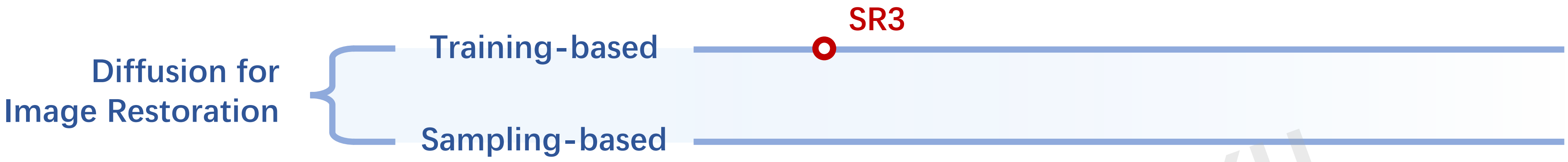


- Image Super-Resolution via Repeated Refinement (SR3)
  - U-Net architecture
  - $\boldsymbol{x}$ interpolated to target high resolution
  - $\boldsymbol{x}$ concatenate with noisy image $\boldsymbol{y_t}$

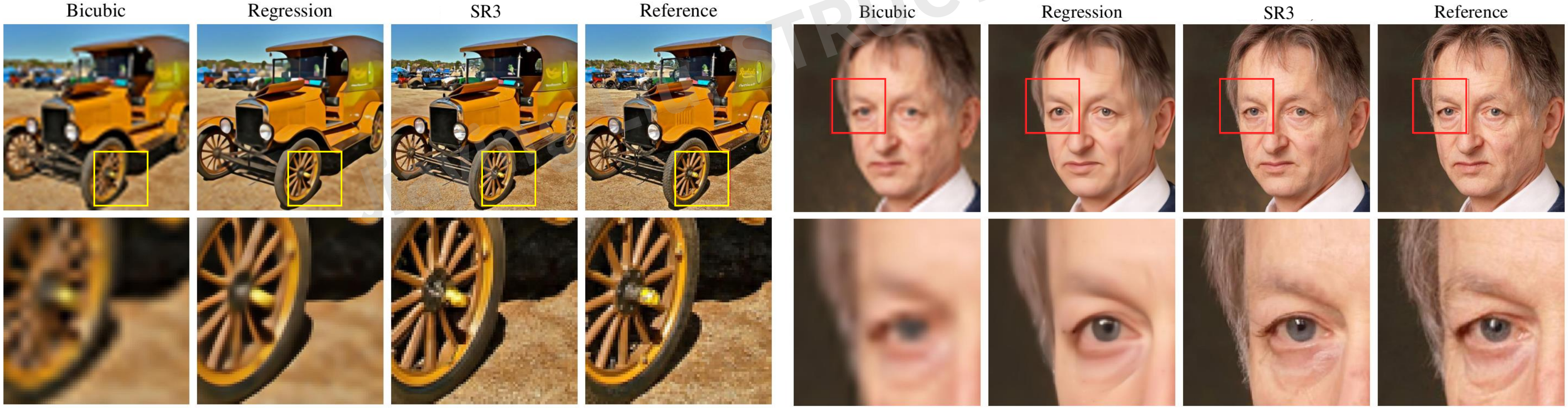


[2] Chitwan Saharia, et al. Image Super-Resolution via Iterative Refinement, IEEE TPAMI, 2022

Diffusion for Image Restoration – SISR

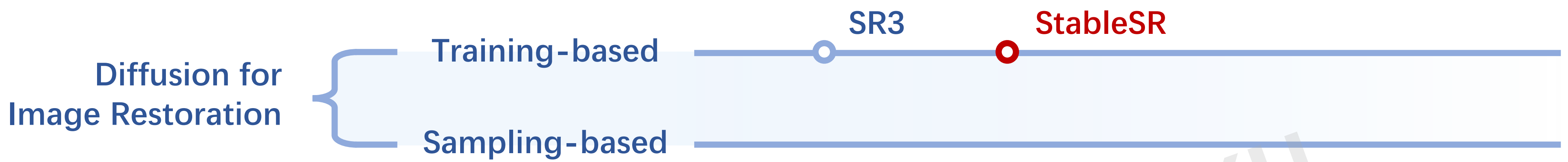


- Image Super-Resolution via Repeated Refinement (SR3)

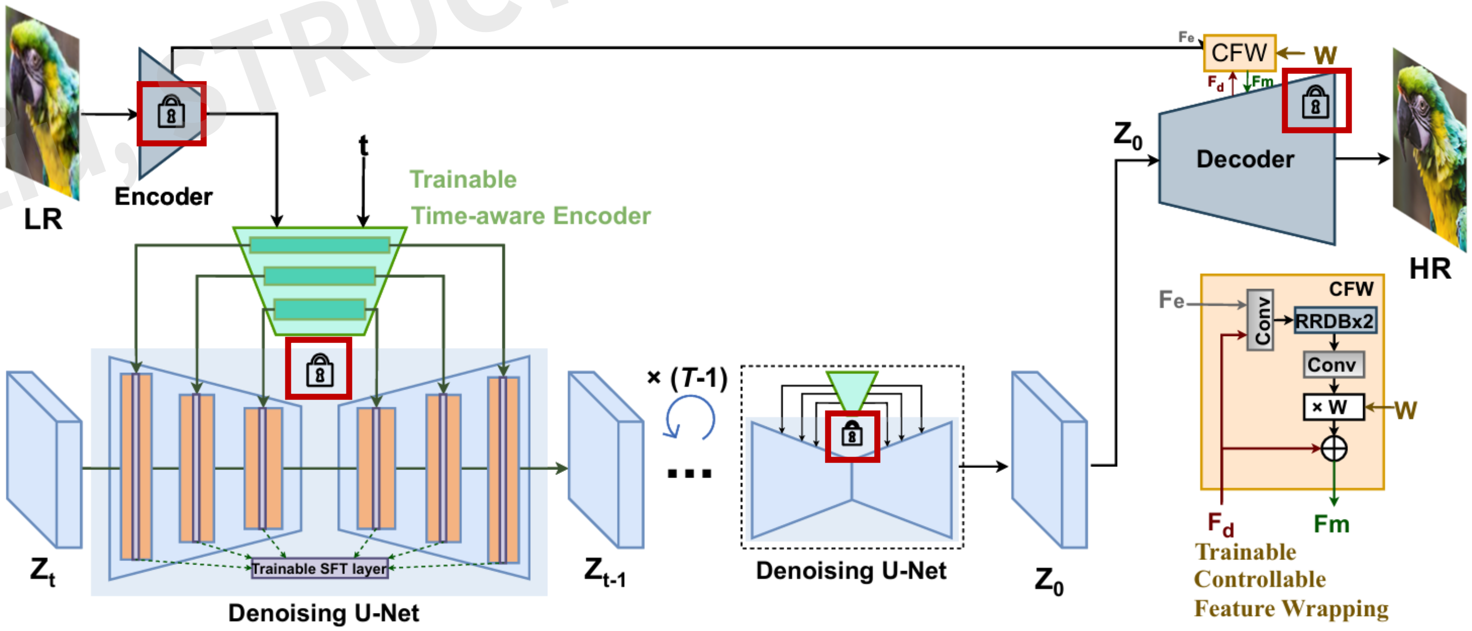


[2] Chitwan Saharia, et al. Image Super-Resolution via Iterative Refinement, IEEE TPAMI, 2022

Diffusion for Image Restoration – SISR

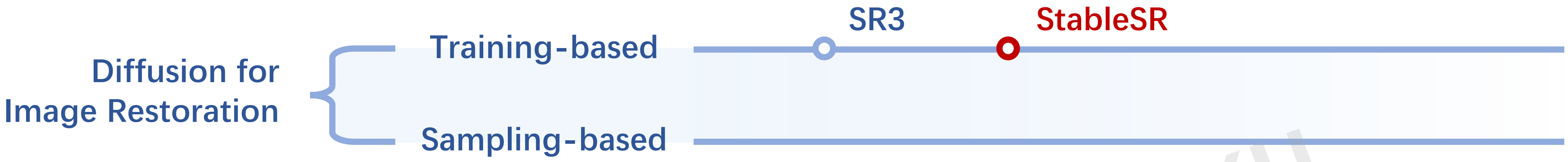


- StableSR -- Image SR with diffusion prior
  - Pretrained latent diffusion model for image synthesis



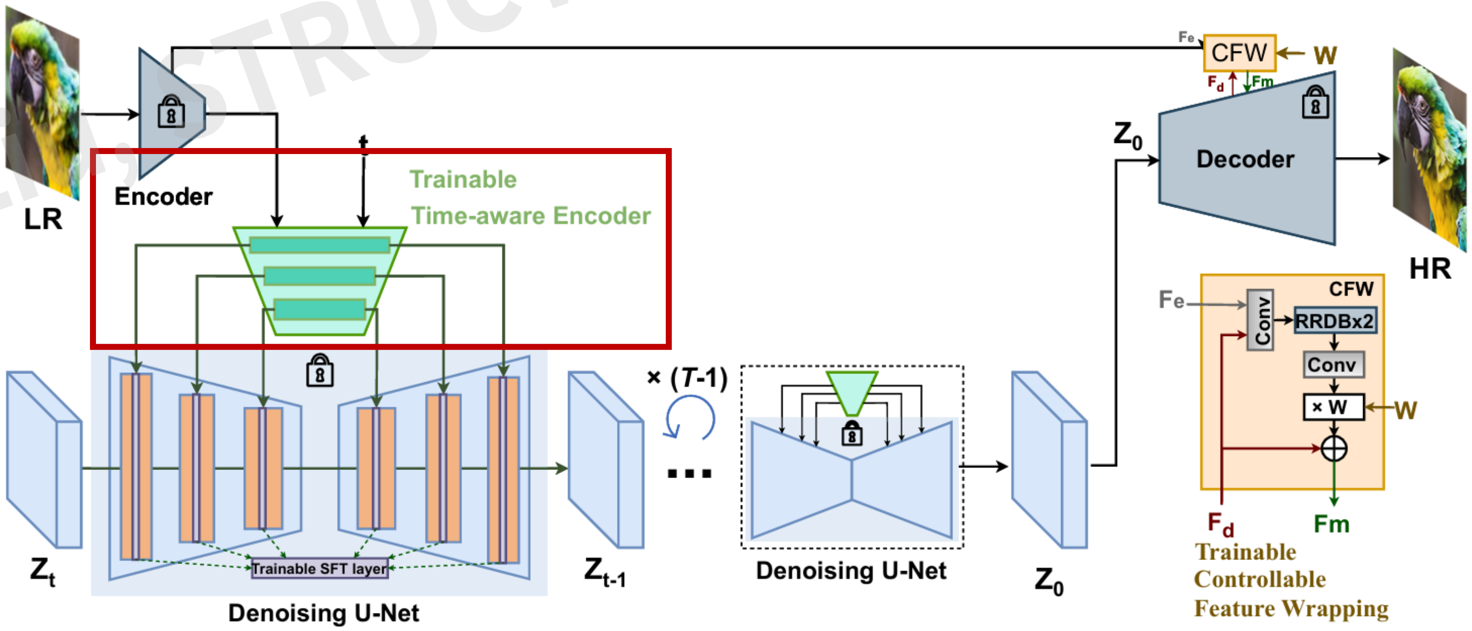
[3] Jianyi Wang, et al. Exploiting diffusion prior for real-world image super-resolution. arXiv, 2023

Diffusion for Image Restoration – SISR



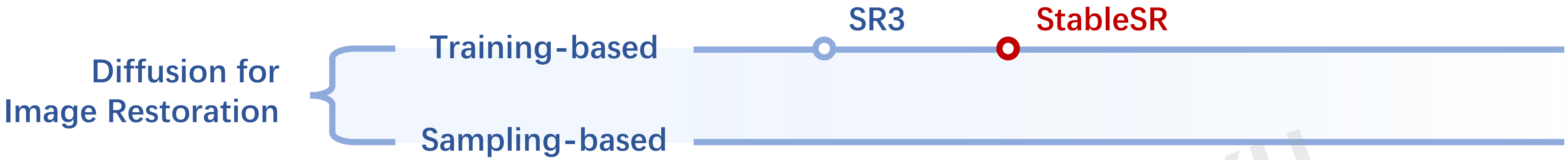
■ StableSR -- Image SR with diffusion prior

- Pretrained latent diffusion model for image synthesis
- Train time-aware encoder for SR



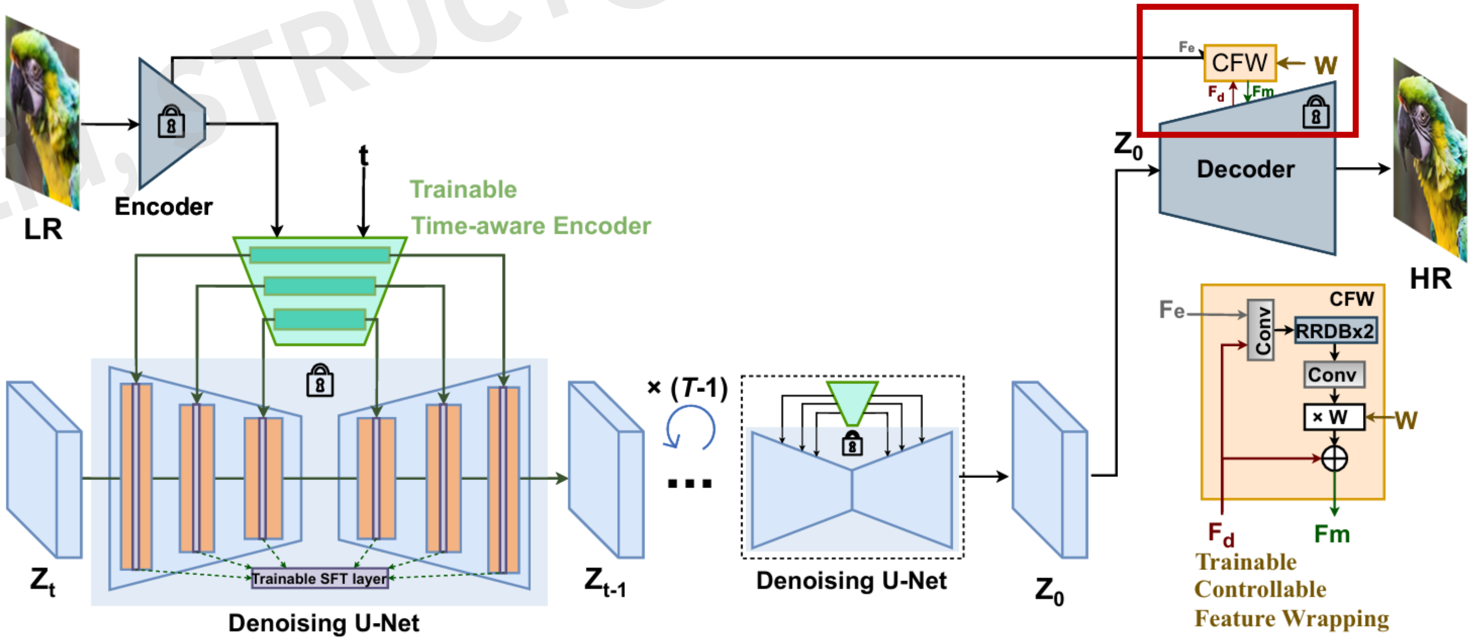
[3] Jianyi Wang, et al. Exploiting diffusion prior for real-world image super-resolution. arXiv, 2023

Diffusion for Image Restoration – SISR



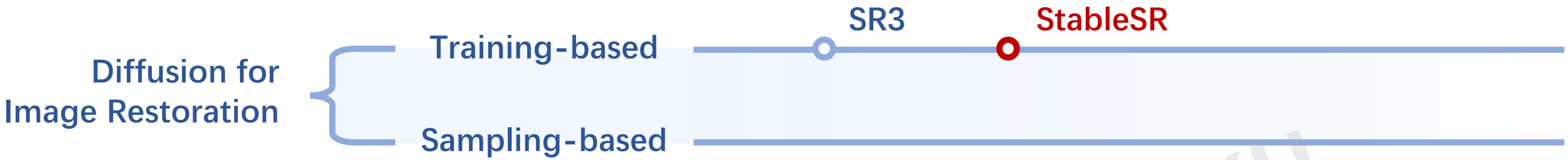
StableSR -- Image SR with diffusion prior

- Pretrained latent diffusion model for image synthesis
- Train time-aware encoder for SR
- Controllable feature wrapping to balance fidelity and perceptual quality

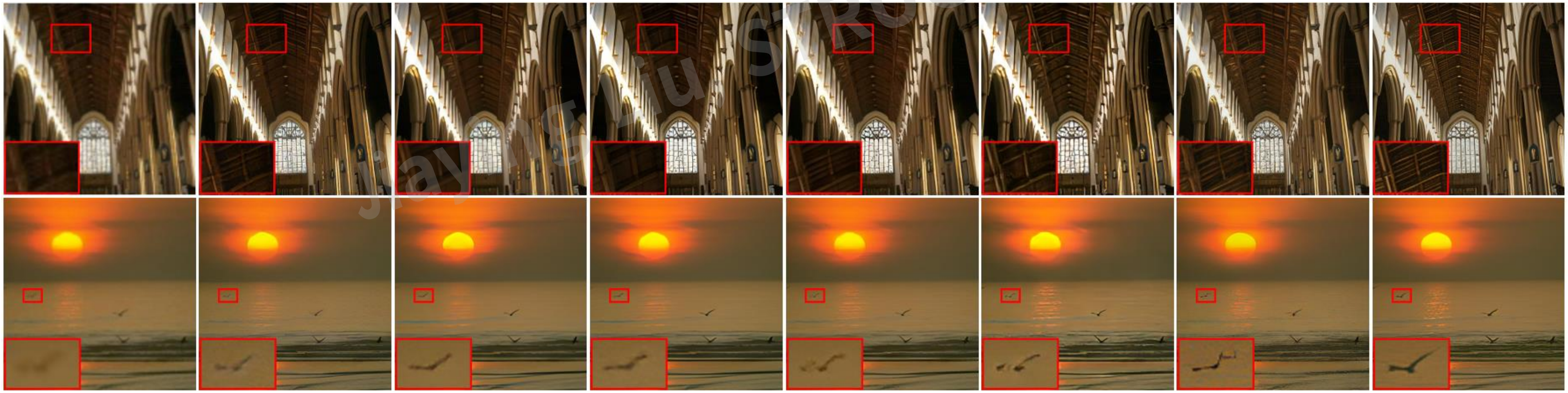


[3] Jianyi Wang, et al. Exploiting diffusion prior for real-world image super-resolution. arXiv, 2023

Diffusion for Image Restoration – SISR



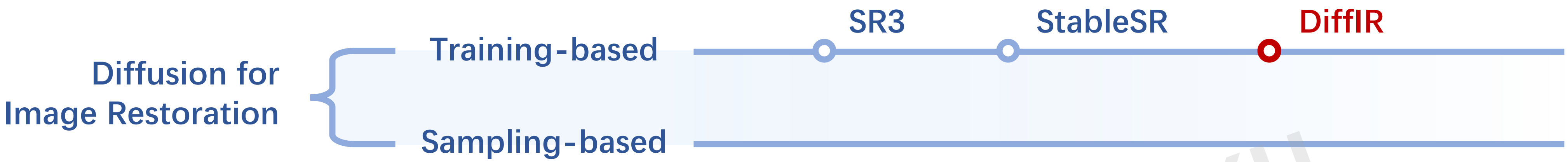
- StableSR -- Image SR with diffusion prior



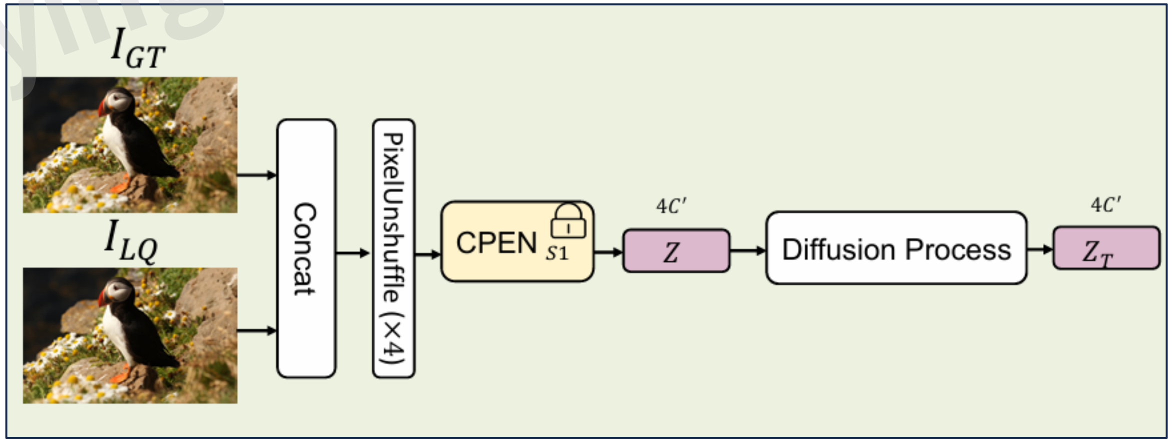
(a) Zoomed LR (b) RealSR (c) BSRGAN (d) Real-ESRGAN+ (e) DASR (f) FeMaSR (g) LDM (h) StableSR

[3] Jianyi Wang, et al. Exploiting diffusion prior for real-world image super-resolution. arXiv, 2023

Diffusion for Image Restoration – SISR

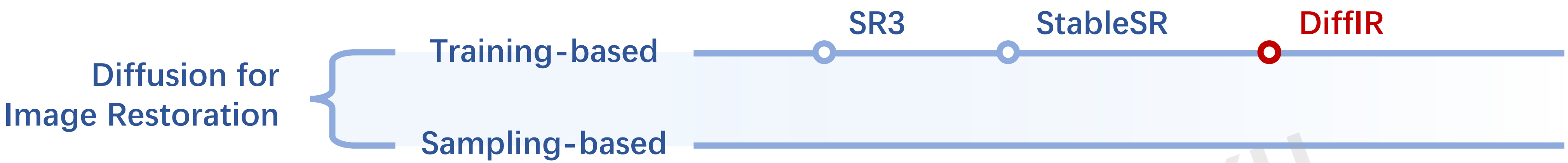


- Diffusion Model for Image Restoration (DiffIR)
  - Instead of generate target HR image with diffusion model, generate compact IR prior
  - In **Stage 1**, IR prior Z is generated from ground truth and LR input, to guide restoration network

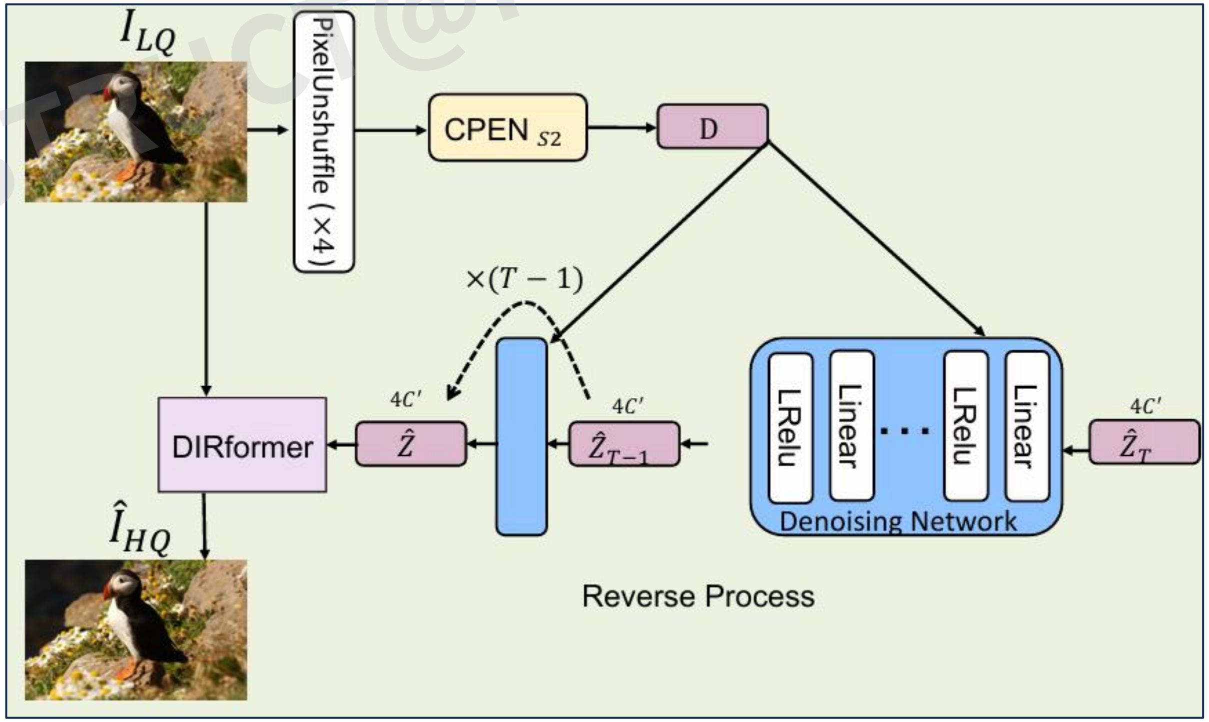


[4] Bin Xia, et al. DiffIR: Efficient Diffusion Model for Image Restoration, ICCV, 2023

Diffusion for Image Restoration – SISR

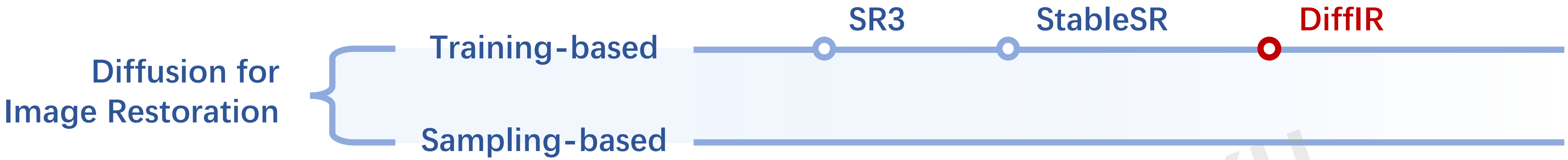


- DiffIR
  - In **Stage 2**, $\hat{Z}$ is generated through conditional denoising diffusion model
  - Guide restoration network DIRformer
  - The generation of compact IR prior is faster and stabler

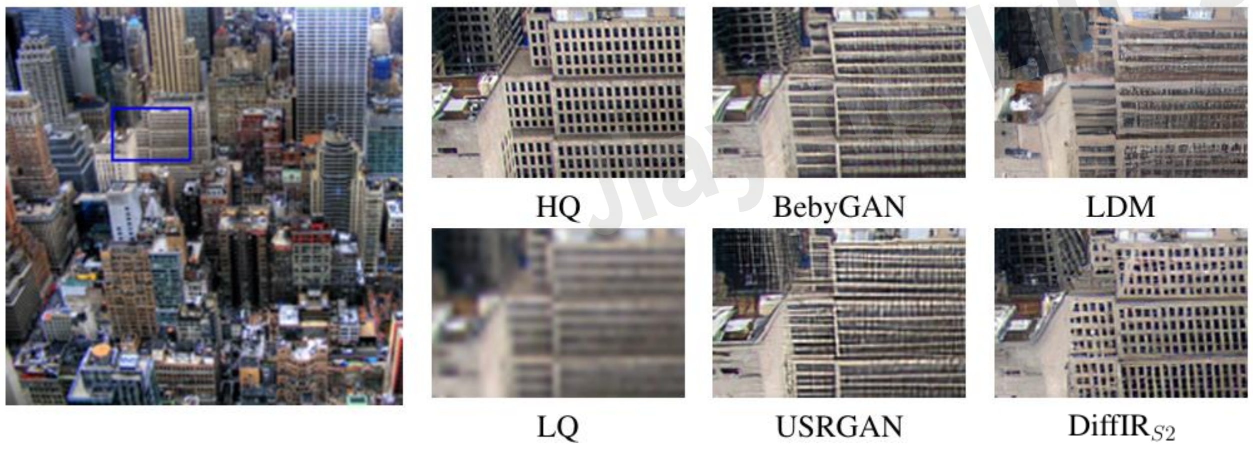


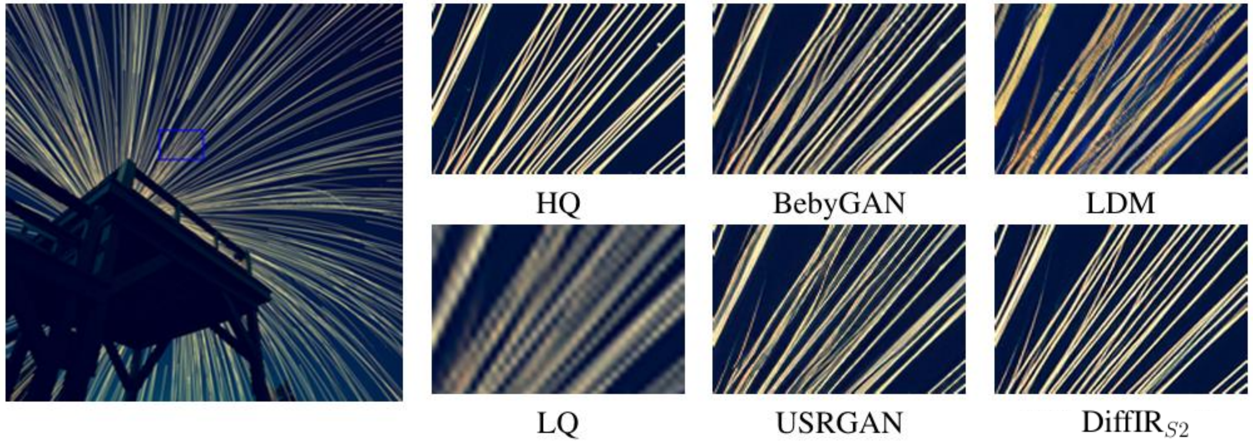
[4] Bin Xia, et al. DiffIR: Efficient Diffusion Model for Image Restoration, ICCV, 2023

Diffusion for Image Restoration – SISR



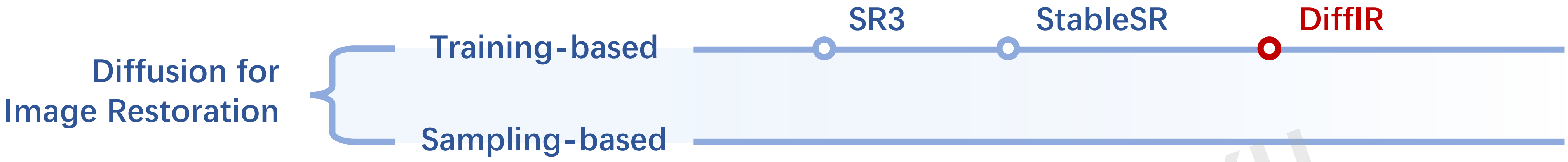
- Diffusion model for image restoration (DiffIR)
  - Image super-resolution (4x)



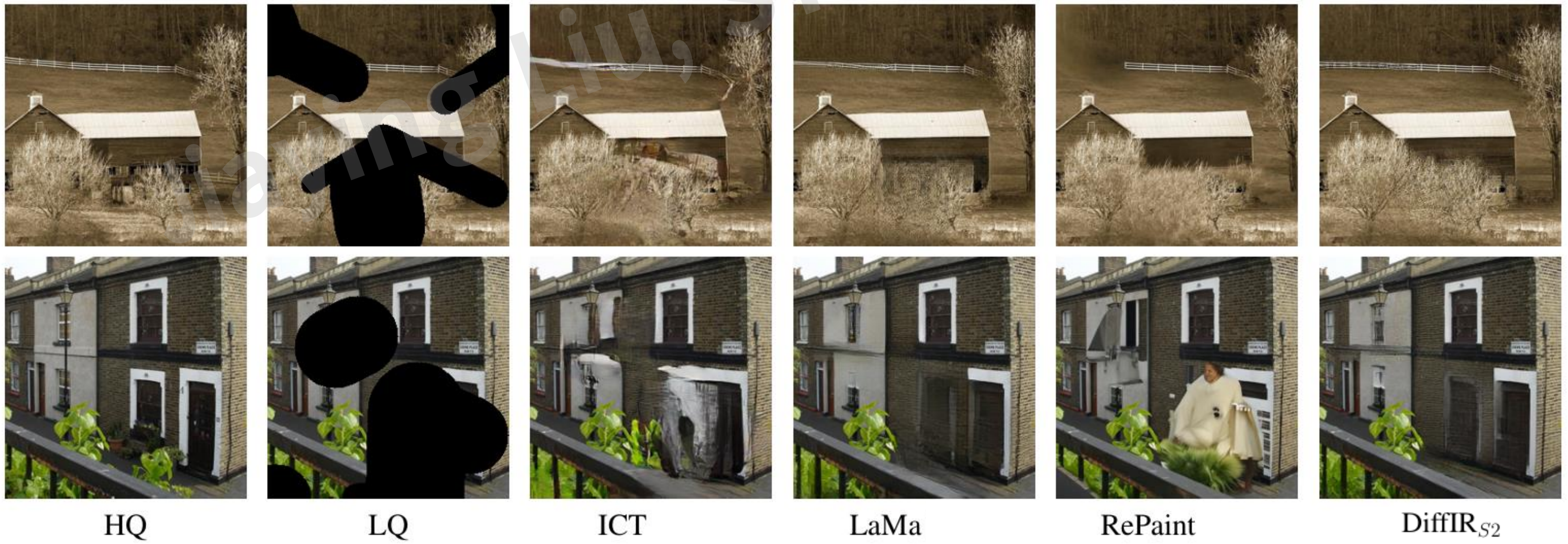


[4] Bin Xia, et al. DiffIR: Efficient Diffusion Model for Image Restoration, ICCV, 2023

Diffusion for Image Restoration – SISR

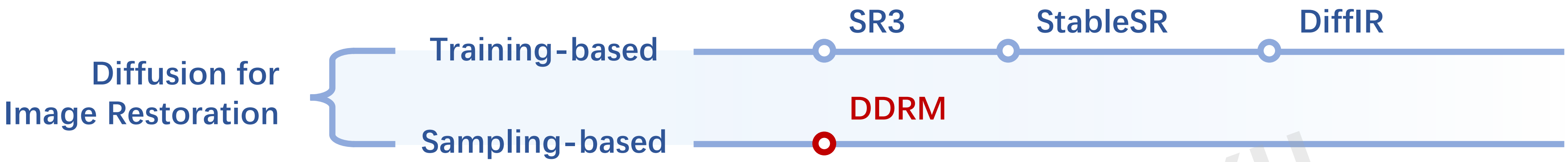


- Diffusion model for image restoration (DiffIR)
  - Inpainting task results



[4] Bin Xia, et al. DiffIR: Efficient Diffusion Model for Image Restoration, ICCV, 2023

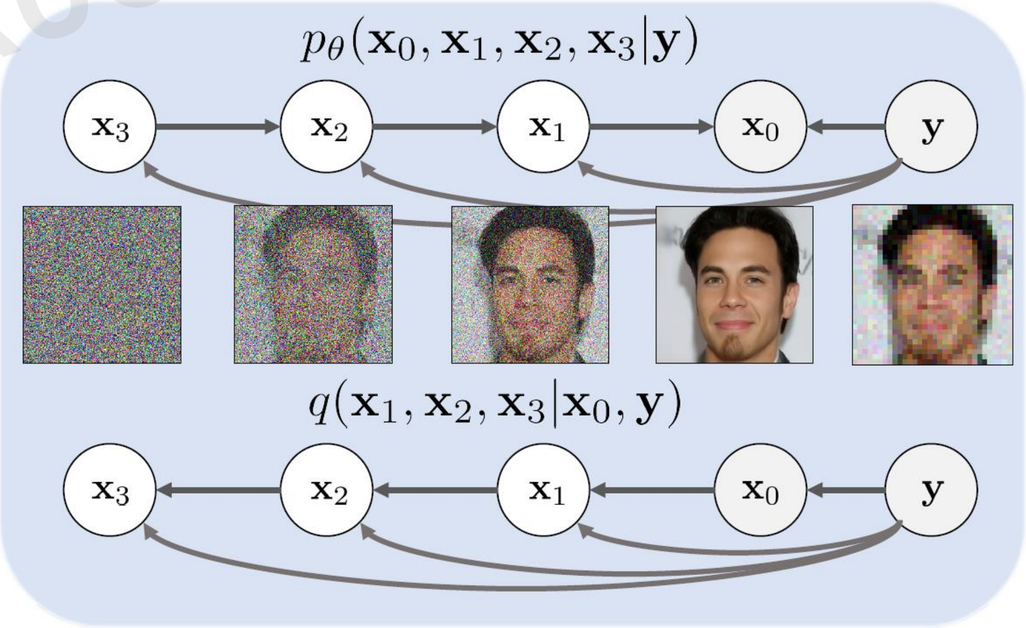
Diffusion for Image Restoration – SISR



- Sampling restored images directly from pre-trained models
  - Take the degraded image as the guidance

- Example:

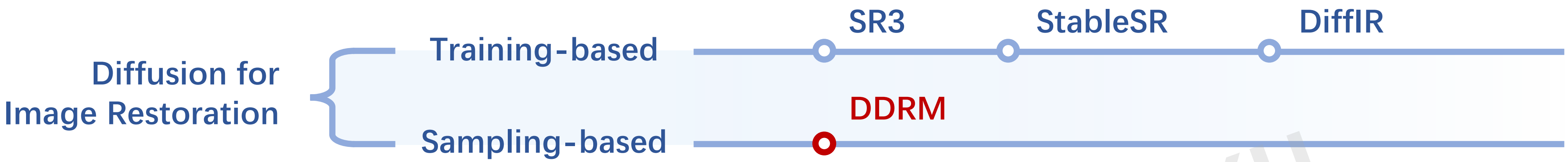
Denoising Diffusion Restoration Models (DDRM)



Denoising Diffusion Restoration Models

[5] Bahjat Kawar, et al. Denoising Diffusion Restoration Models, NIPS, 2022

Diffusion for Image Restoration – SISR



- Denoising Diffusion Restoration Models
  - A simplified degradation model in general:

$$\mathbf{y} = \boldsymbol{H}\mathbf{x} + \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \sigma_{\mathbf{y}}^2 \boldsymbol{I})$$

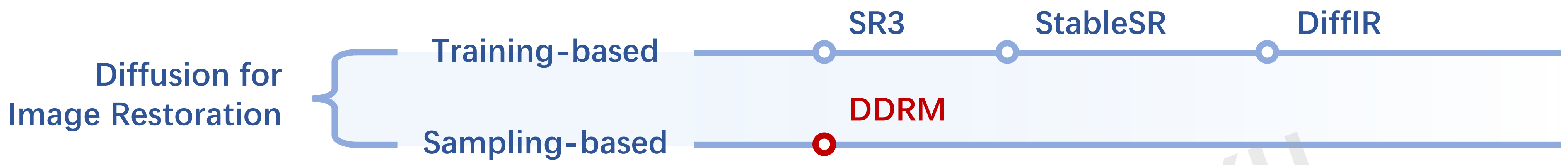
  - The target is to find:

$$p_\theta(\mathbf{x}_{0:T}|\mathbf{y}) = p_\theta^{(T)}(\mathbf{x}_T|\mathbf{y}) \prod_{t=0}^{T-1} p_\theta^{(t)}(\mathbf{x}_t|\mathbf{x}_{t+1}, \mathbf{y})$$

where θ is the parameter set of a pre-trained diffusion model.

[5] Bahjat Kawar, et al. Denoising Diffusion Restoration Models, NIPS, 2022

Diffusion for Image Restoration – SISR



- Denoising Diffusion Restoration Models
  - The conditional distributions are demonstrated as:

Perform SVD $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$

Correct Sampling Distribution

Distribution of the initial noise

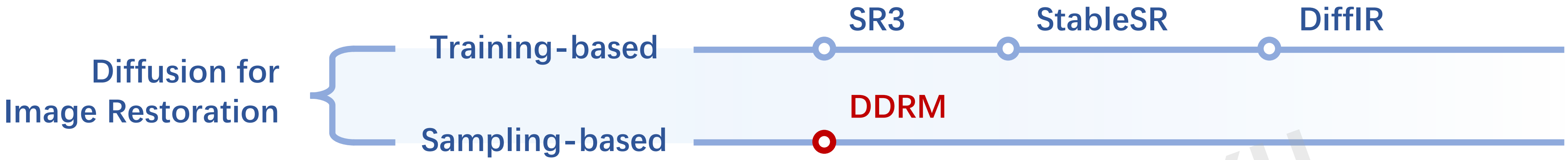
$$\boxed{p_\theta^{(T)}(\bar{\mathbf{x}}_T^{(i)}|\mathbf{y})} = \begin{cases} \mathcal{N}(\bar{\mathbf{y}}^{(i)}, \sigma_T^2 - \frac{\sigma_{\mathbf{y}}^2}{s_i^2}) & \text{if } s_i > 0 \\ \mathcal{N}(0, \sigma_T^2) & \text{if } s_i = 0 \end{cases}$$

$$p_\theta^{(t)}(\bar{\mathbf{x}}_t^{(i)}|\mathbf{x}_{t+1}, \mathbf{y}) = \begin{cases} \mathcal{N}(\bar{\mathbf{x}}_{\theta,t}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{\mathbf{x}}_{t+1}^{(i)} - \bar{\mathbf{x}}_{\theta,t}^{(i)}}{\sigma_{t+1}}, \eta^2\sigma_t^2) & \text{if } s_i = 0 \\ \mathcal{N}(\bar{\mathbf{x}}_{\theta,t}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{\mathbf{y}}^{(i)} - \bar{\mathbf{x}}_{\theta,t}^{(i)}}{\sigma_{\mathbf{y}}/s_i}, \eta^2\sigma_t^2) & \text{if } \sigma_t < \frac{\sigma_{\mathbf{y}}}{s_i} \\ \mathcal{N}((1-\eta_b)\bar{\mathbf{x}}_{\theta,t}^{(i)} + \eta_b\bar{\mathbf{y}}^{(i)}, \sigma_t^2 - \frac{\sigma_{\mathbf{y}}^2}{s_i^2}\eta_b^2) & \text{if } \sigma_t \geq \frac{\sigma_{\mathbf{y}}}{s_i} \end{cases}$$

where $\bar{\mathbf{x}}_t$ is the set of vectors in spectral space: $\bar{\mathbf{x}}_t = \boldsymbol{V}^\top \mathbf{x}_t$. Requires known degradation.

[5] Bahjat Kawar, et al. Denoising Diffusion Restoration Models, NIPS, 2022

Diffusion for Image Restoration – SISR



- Denoising Diffusion Restoration Models
  - The conditional distributions are demonstrated as:

Perform SVD $\boldsymbol{H} = \boldsymbol{U\Sigma V}^\top$

Correct Sampling Distribution

$$p_\theta^{(T)}(\bar{\mathbf{x}}_T^{(i)}|\mathbf{y}) = \begin{cases} \mathcal{N}(\bar{\mathbf{y}}^{(i)}, \sigma_T^2 - \frac{\sigma_\mathbf{y}^2}{s_i^2}) & \text{if } s_i > 0 \\ \mathcal{N}(0, \sigma_T^2) & \text{if } s_i = 0 \end{cases}$$

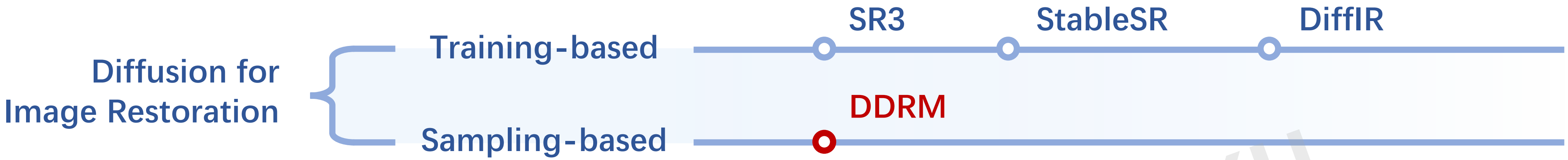
Distribution of one single denoise step

$$\boxed{p_\theta^{(t)}(\bar{\mathbf{x}}_t^{(i)}|\mathbf{x}_{t+1}, \mathbf{y})} = \begin{cases} \mathcal{N}(\bar{\mathbf{x}}_{\theta,t}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{\mathbf{x}}_{t+1}^{(i)} - \bar{\mathbf{x}}_{\theta,t}^{(i)}}{\sigma_{t+1}}, \eta^2\sigma_t^2) & \text{if } s_i = 0 \\ \mathcal{N}(\bar{\mathbf{x}}_{\theta,t}^{(i)} + \sqrt{1-\eta^2}\sigma_t \frac{\bar{\mathbf{y}}^{(i)} - \bar{\mathbf{x}}_{\theta,t}^{(i)}}{\sigma_\mathbf{y}/s_i}, \eta^2\sigma_t^2) & \text{if } \sigma_t < \frac{\sigma_\mathbf{y}}{s_i} \\ \mathcal{N}((1-\eta_b)\bar{\mathbf{x}}_{\theta,t}^{(i)} + \eta_b\bar{\mathbf{y}}^{(i)}, \sigma_t^2 - \frac{\sigma_\mathbf{y}^2}{s_i^2}\eta_b^2) & \text{if } \sigma_t \geq \frac{\sigma_\mathbf{y}}{s_i} \end{cases}$$

where $\bar{\mathbf{x}}_t$ is the set of vectors in spectral space: $\bar{\mathbf{x}}_t = \boldsymbol{V}^\top \mathbf{x}_t$. Requires known degradation.

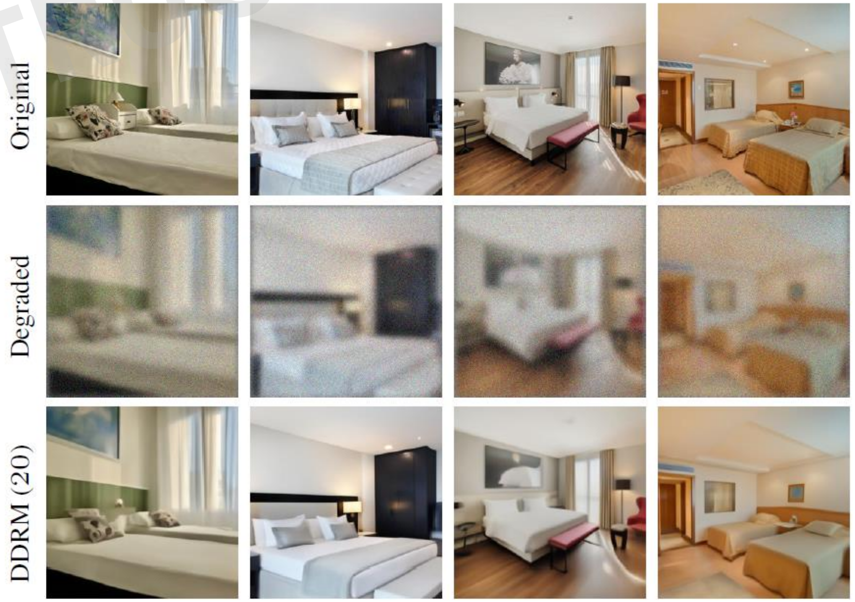
[5] Bahjat Kawar, et al. Denoising Diffusion Restoration Models, NIPS, 2022

Diffusion for Image Restoration – SISR



■ Denoising Diffusion Restoration Models

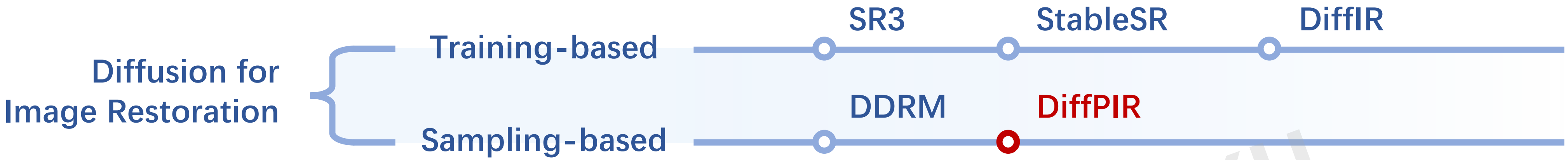
(a) Inpainting results on cat images.



(b) Deblurring results ($\sigma_{\mathbf{y}} = 0.05$) on bedroom images.

[5] Bahjat Kawar, et al. Denoising Diffusion Restoration Models, NIPS, 2022

Diffusion for Image Restoration – SISR



- Denoising Diffusion Models for Plug-and-Play Image Restoration (DiffPIR)

Employ a refinement during sampling process:

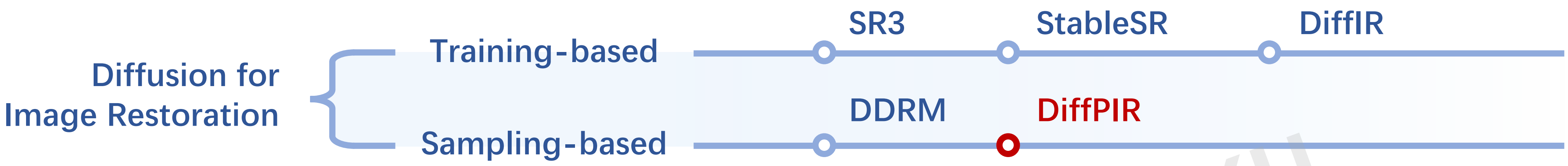
$$\hat{\mathbf{x}}_0^{(t)} = \arg\min_{\mathbf{x}} \underbrace{\boxed{\|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2}}_{\text{Fidelity to degraded image}} + \rho_t \|\mathbf{x} - \mathbf{x}_0^{(t)}\|^2$$

Fidelity to degraded image

where $\mathbf{x}_0^{(t)}$ is the predicted noise-free image at timestep t. Requires known degradation.

[6] Yuanzhi Zhu, et al. Denoising Diffusion Models for Plug-and-Play Image Restoration, CVPR, 2023

Diffusion for Image Restoration – SISR



- Denoising Diffusion Models for Plug-and-Play Image Restoration

Employ a refinement during sampling process:

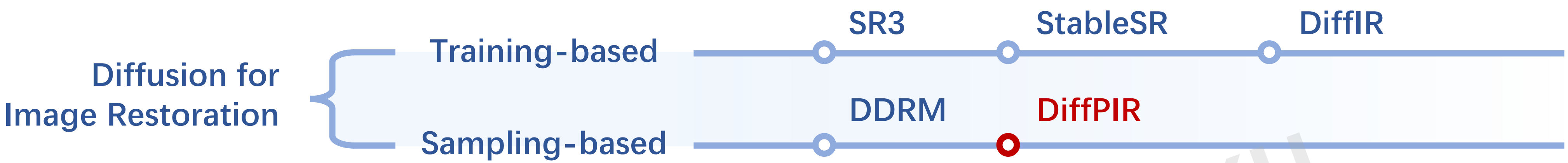
$$\hat{\mathbf{x}}_0^{(t)} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathcal{H}(\mathbf{x})\|^2 + \boxed{\rho_t \|\mathbf{x} - \mathbf{x}_0^{(t)}\|^2}$$

Regularization of generated image

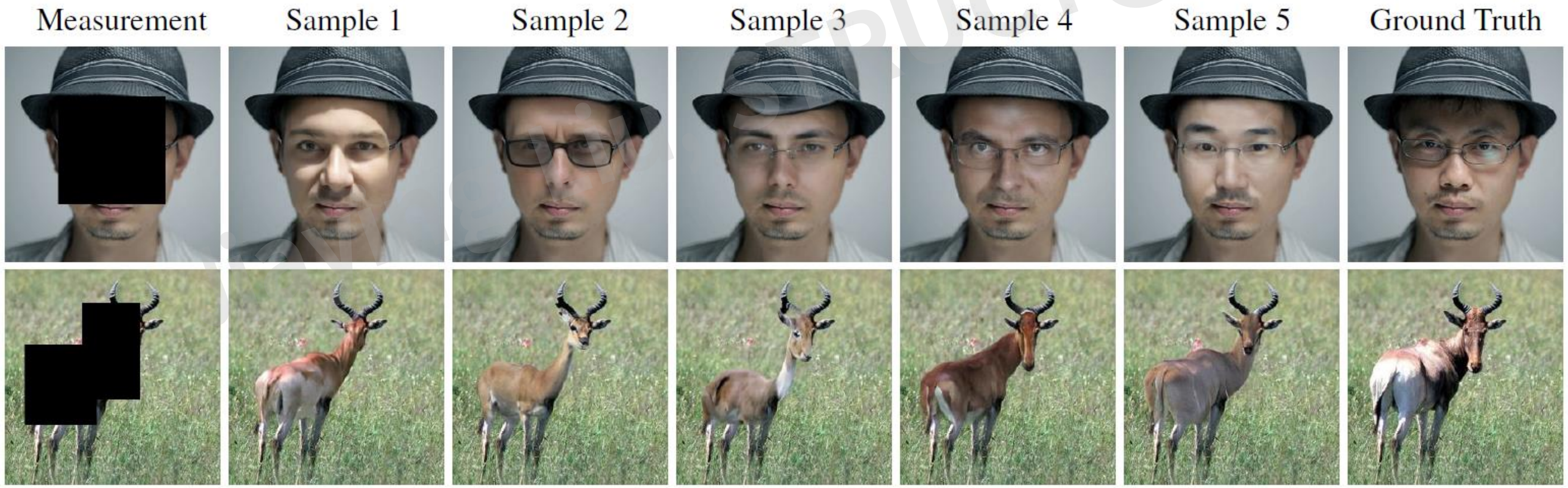
where $\mathbf{x}_0^{(t)}$ is the predicted noise-free image at timestep t. Requires known degradation.

[6] Yuanzhi Zhu, et al. Denoising Diffusion Models for Plug-and-Play Image Restoration, CVPR, 2023

Diffusion for Image Restoration – SISR



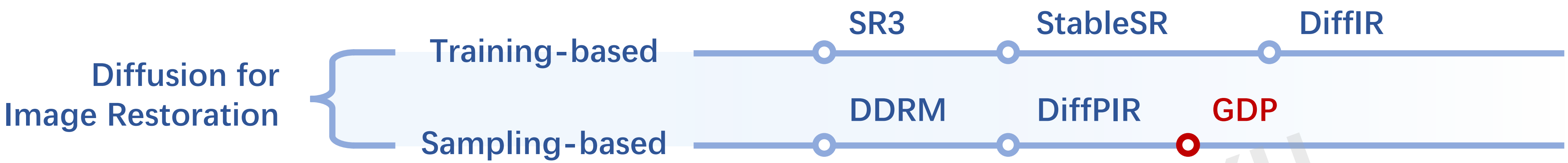
- Denoising Diffusion Models for Plug-and-Play Image Restoration



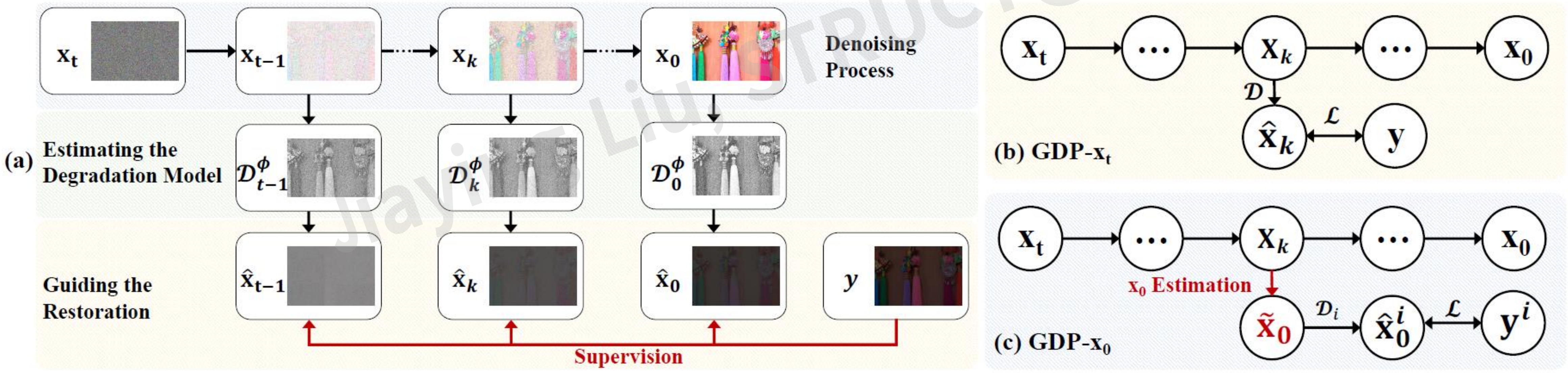
Diverse and reasonable results.

[6] Yuanzhi Zhu, et al. Denoising Diffusion Models for Plug-and-Play Image Restoration, CVPR, 2023

Diffusion for Image Restoration – SISR



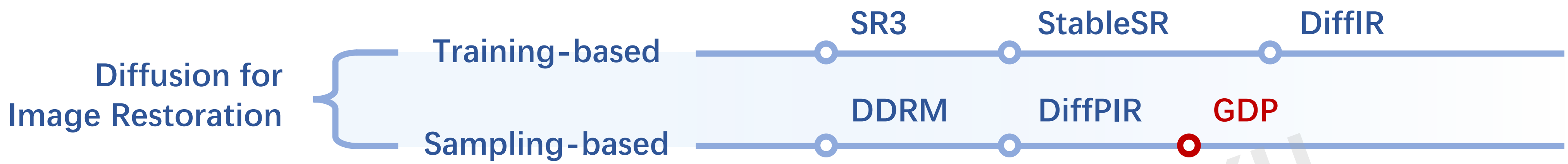
- Generative Diffusion Prior (GDP)



Learn the degradation model during the sampling process.

[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

Diffusion for Image Restoration – SISR



■ Generative Diffusion Prior (GDP)

At the t-th step, the degradation model will be updated via:

$$\mathcal{L}_{\phi,\tilde{\boldsymbol{x}}_0}^{total} = \mathcal{L}(\boldsymbol{y}, \boxed{\mathcal{D}_\phi}(\tilde{\boldsymbol{x}}_0)) + \mathcal{Q}(\tilde{\boldsymbol{x}}_0)$$

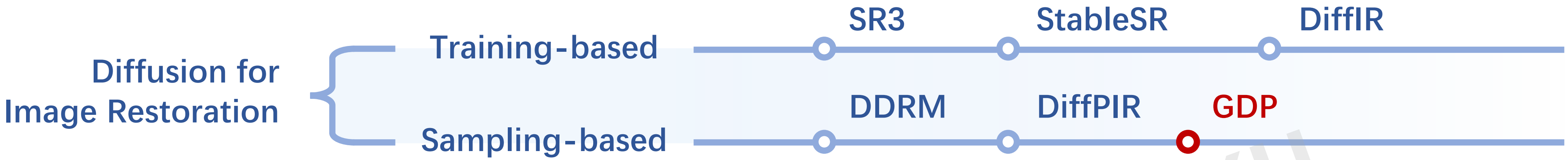
$$\phi \leftarrow \phi - l \nabla_\phi \mathcal{L}_{\phi,\tilde{\boldsymbol{x}}_0}^{total}$$

The noisy image at (t-1)-th step will be sampled with guidance provided by the loss:

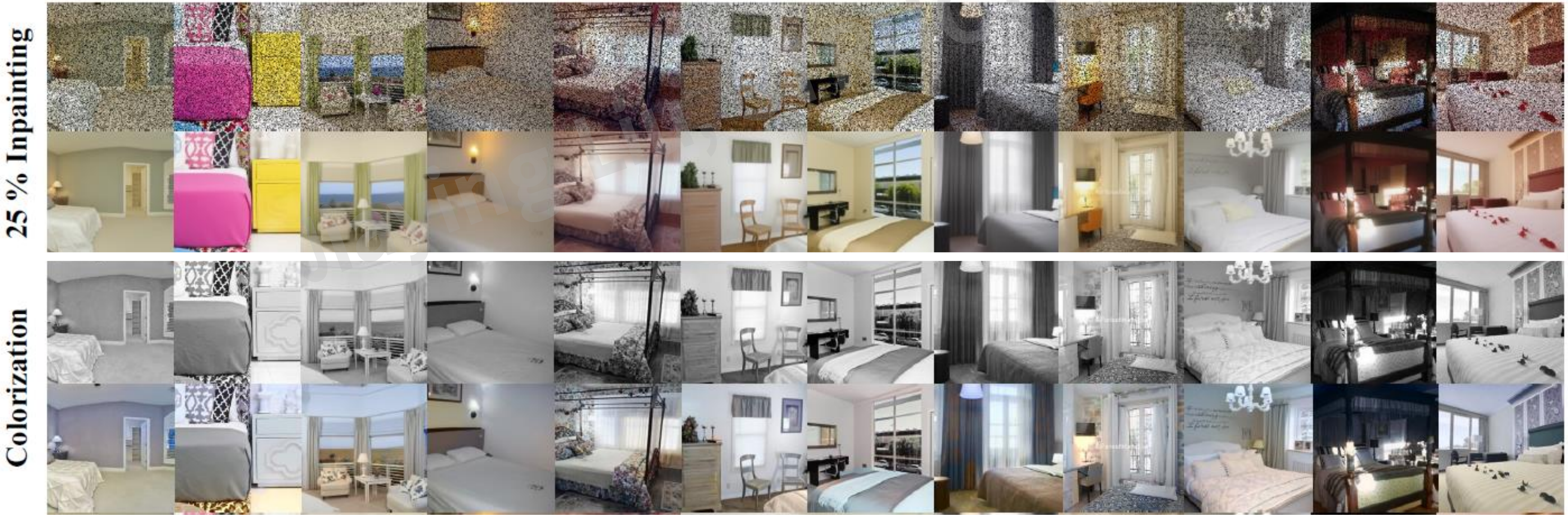
$$\text{Sample } \boldsymbol{x}_{t-1} \text{ by } \mathcal{N}\left(\mu + s\nabla_{\tilde{\boldsymbol{x}}_0} \mathcal{L}_{\phi,\tilde{\boldsymbol{x}}_0}^{total}, \Sigma\right)$$

[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

Diffusion for Image Restoration – SISR

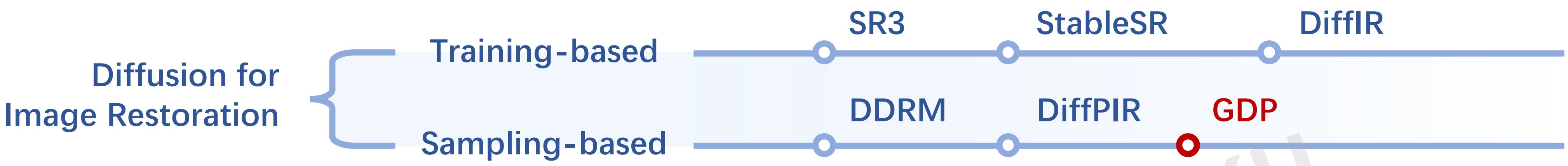


- Generative Diffusion Prior (GDP)



[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

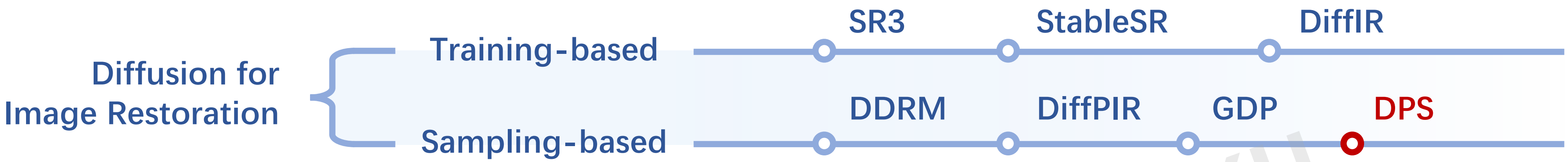
Diffusion for Image Restoration – SISR



- Generative Diffusion Prior (GDP)

[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

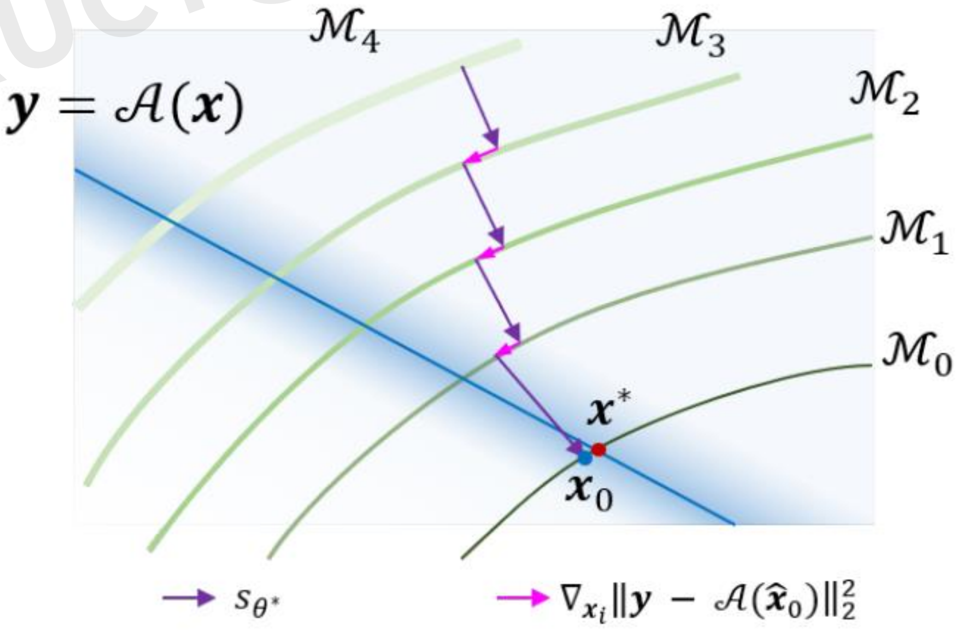
Diffusion for Image Restoration – SISR



- Diffusion Posterior Sampling (DPS)

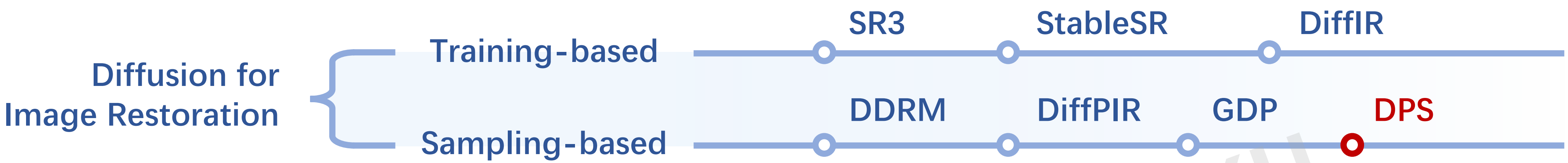
refine the noisy state at sampling process via:

$$\boldsymbol{x}_{i-1} \leftarrow \boldsymbol{x}'_{i-1} - \zeta_i \nabla_{\boldsymbol{x}_i} \|\boldsymbol{y} - \mathcal{A}(\hat{\boldsymbol{x}}_0)\|_2^2$$



[8] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

Diffusion for Image Restoration – SISR

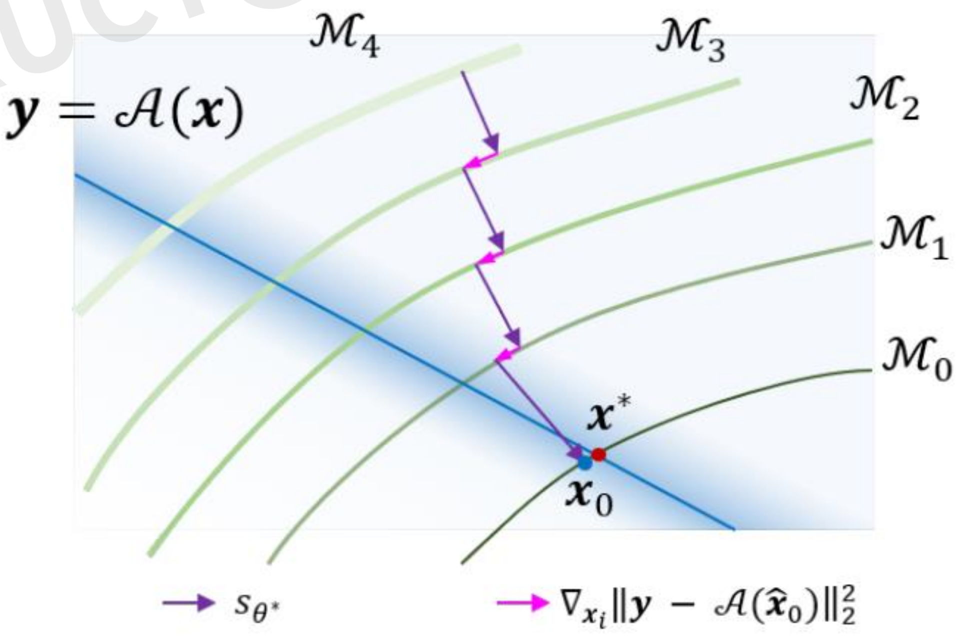


- Diffusion Posterior Sampling (DPS)

refine the noisy state at sampling process via:

$$\boldsymbol{x}_{i-1} \leftarrow \boldsymbol{x}'_{i-1} - \zeta_i \boxed{\nabla_{\boldsymbol{x}_i} \|\boldsymbol{y} - \mathcal{A}(\hat{\boldsymbol{x}}_0)\|_2^2}$$

Correction Item



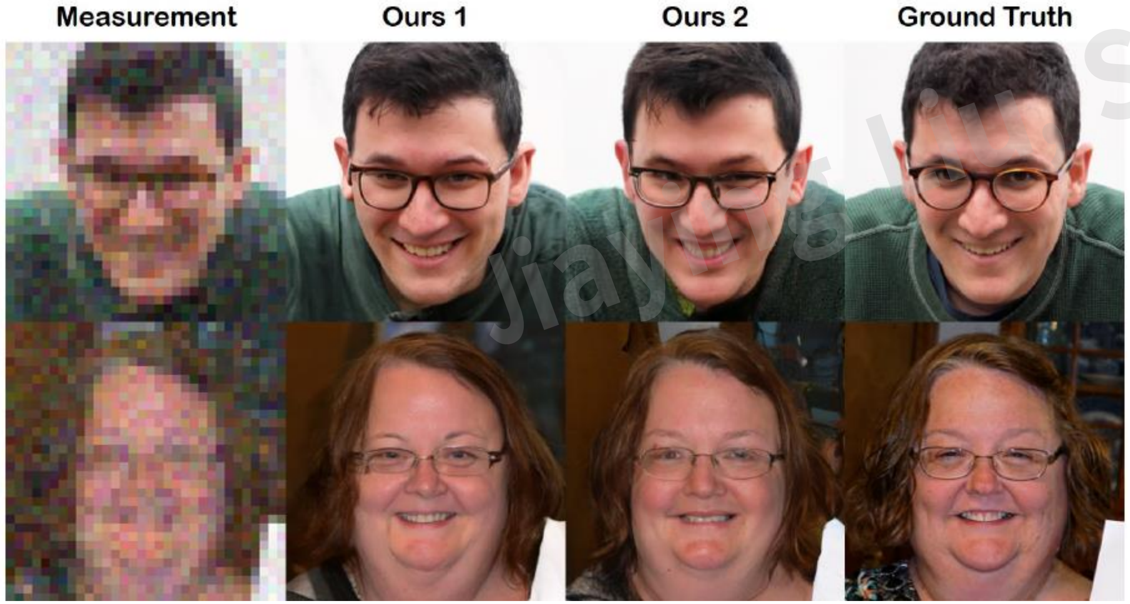
[8] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

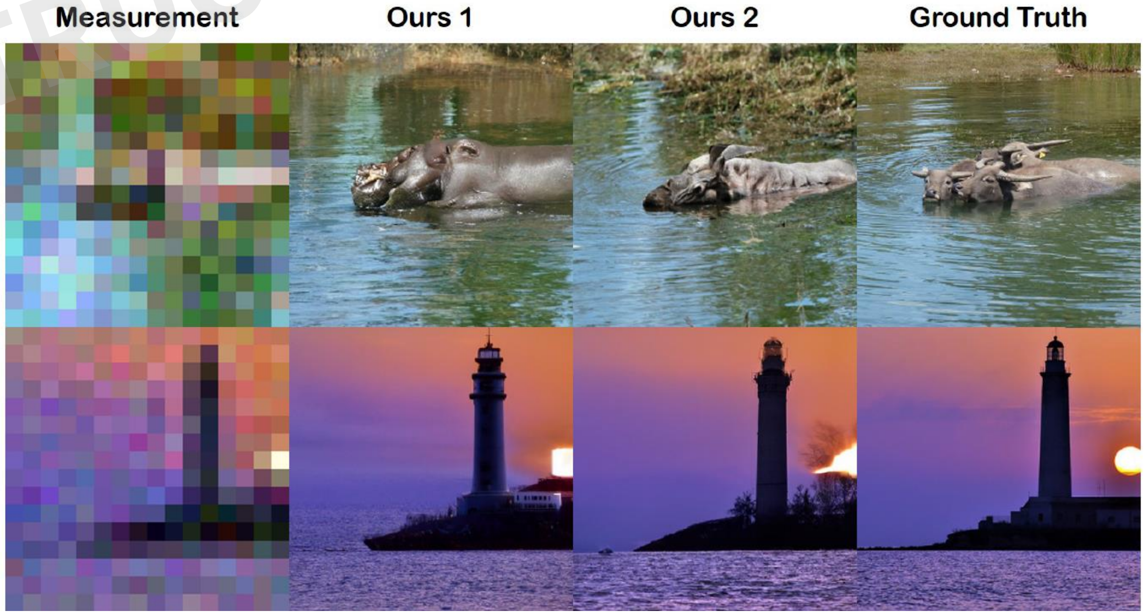
Diffusion for Image Restoration – SISR

Diffusion for Image Restoration

- Training-based: SR3, StableSR, DiffIR
- Sampling-based: DDRM, DiffPIR, GDP, DPS

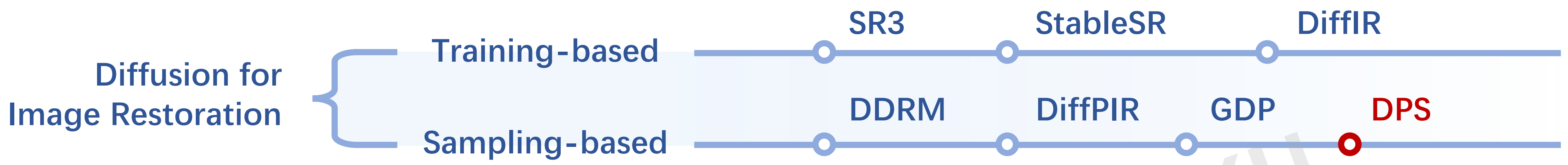
Diffusion Posterior Sampling (DPS)



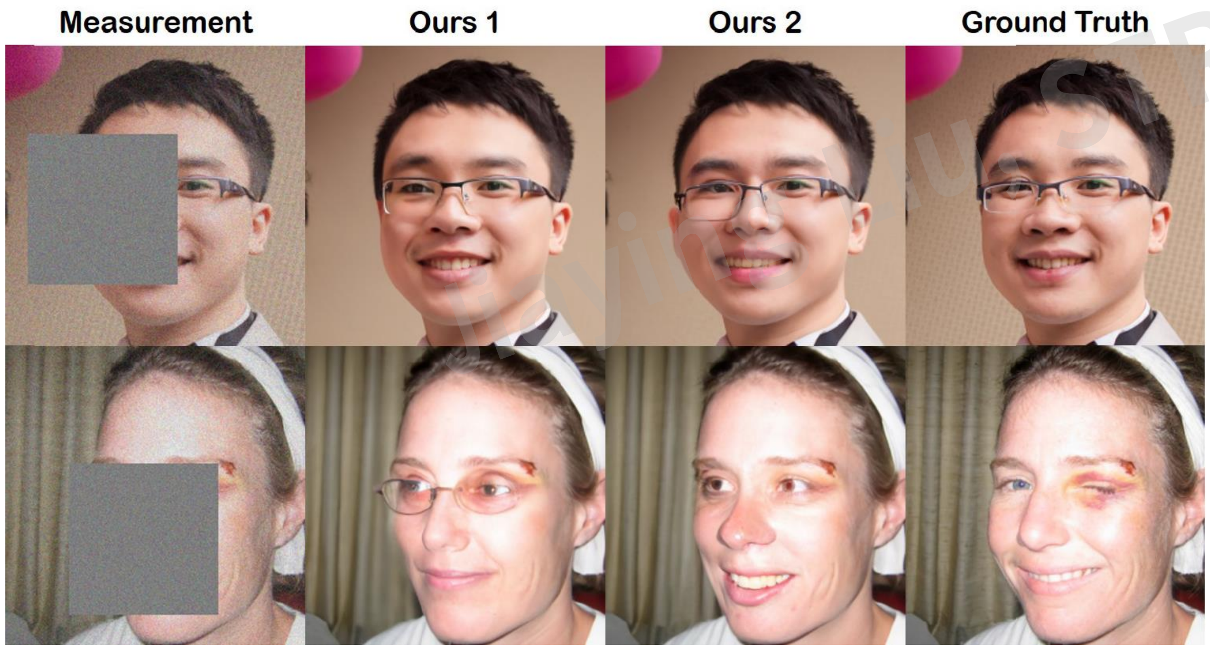


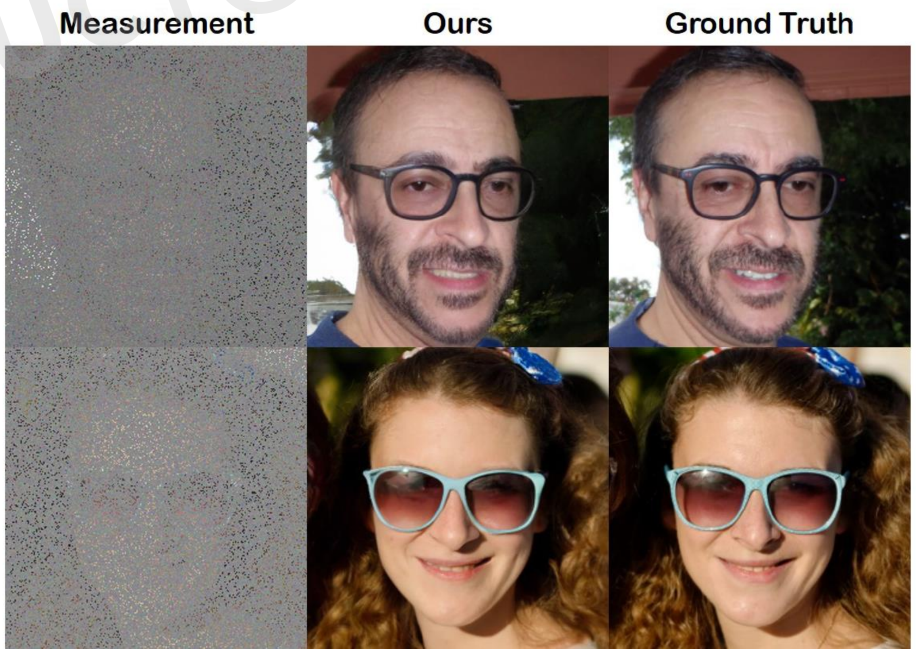
[8] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

Diffusion for Image Restoration – SISR



- Diffusion Posterior Sampling (DPS)





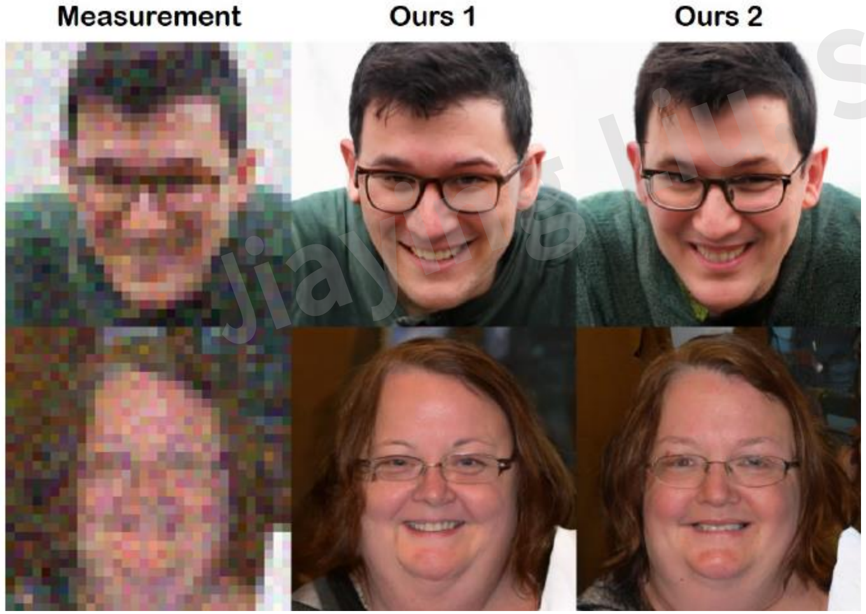
[8] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

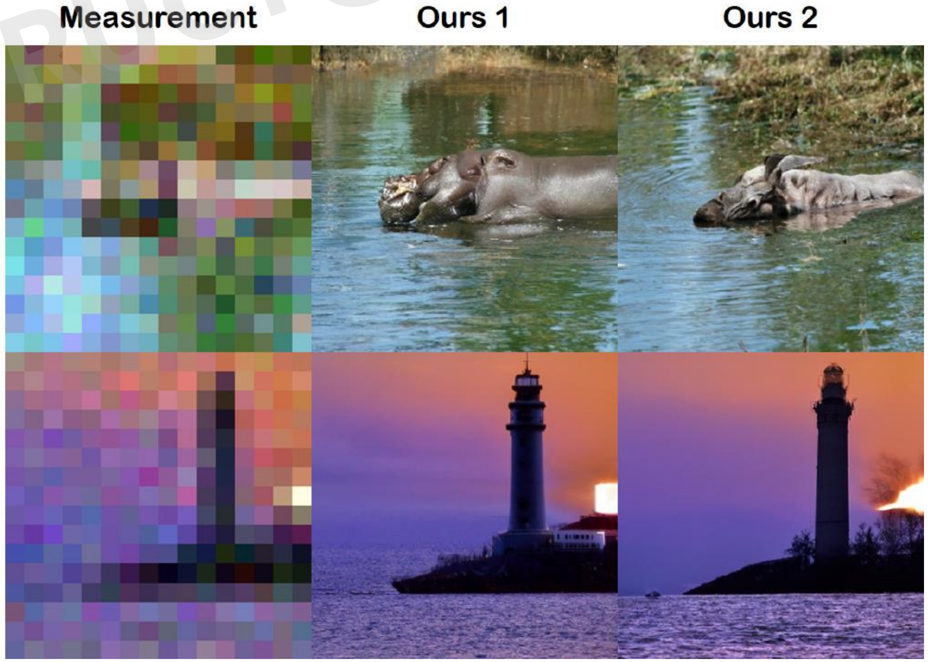
Diffusion for Image Restoration – SISR

Problem of existing works:

There are two characteristics of diffusion-based IR models:

- Sampling **instability** due to the randomness of diffusion models





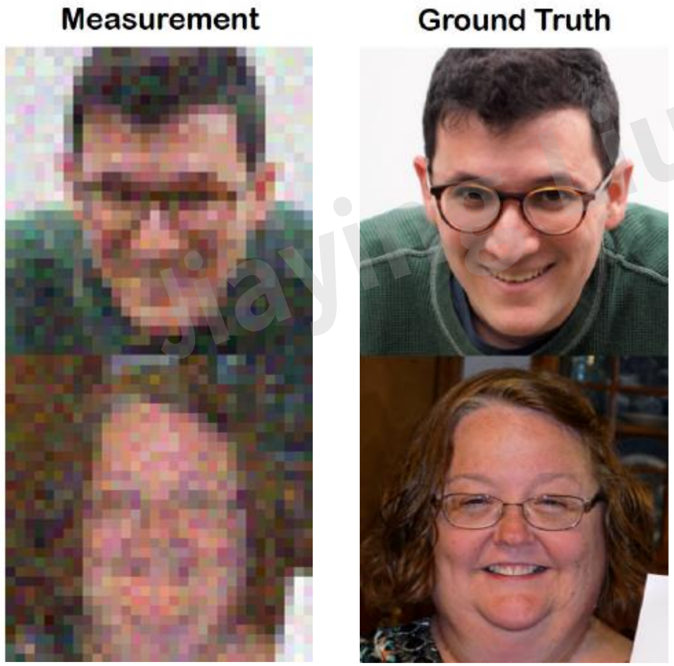
[10] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

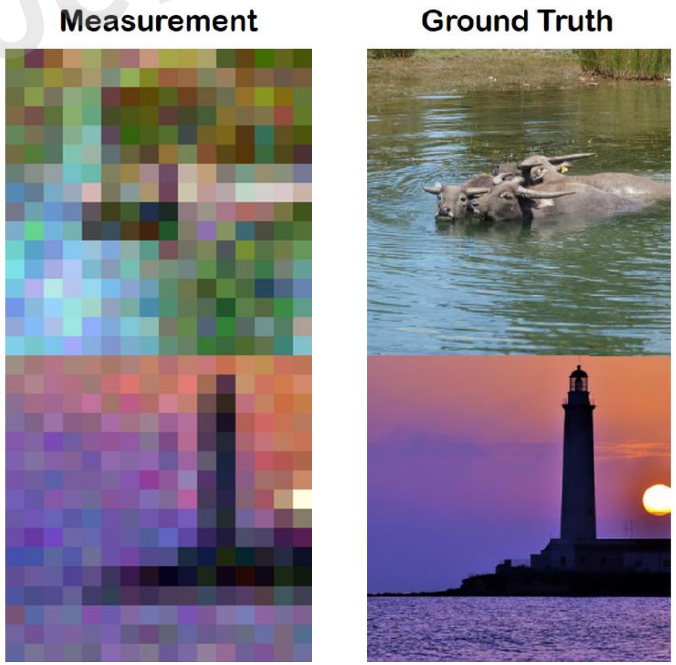
Diffusion for Image Restoration – SISR

Problem of existing works:

There are two characteristics of diffusion-based IR models:

- **Unique ground truths** of the task of image IR





[10] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

Diffusion for Image Restoration – SISR

- **Problem of existing works:**

  There are two characteristics of diffusion-based IR models:

  - Sampling **instability** due to the randomness of diffusion models

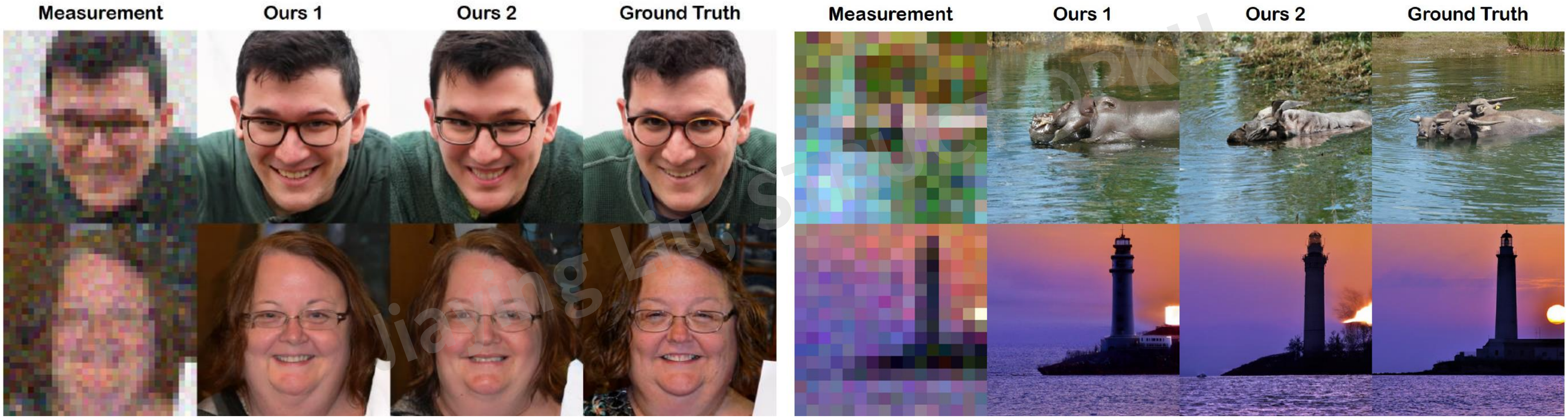
  Vs. **Unique ground truths** of the task of image IR

  It leads to the **problem**:

  - The quality of each sample **cannot be guaranteed**

Diffusion for Image Restoration – SISR

- Problem:



Diverse and reasonable, but there is **only one GT**!

Which one is **better**? How to **ensure the quality of the next sample**?

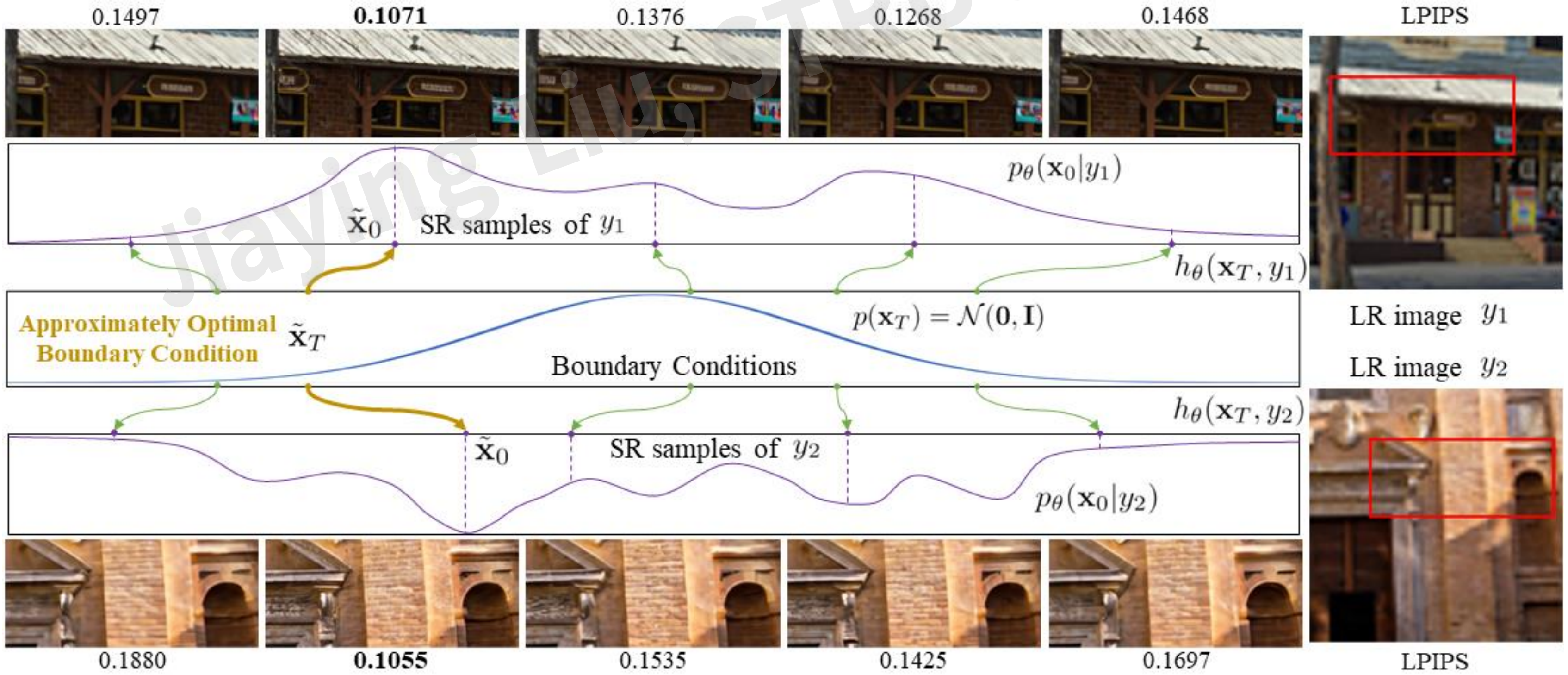
[8] Hyungjin Chung, et al. Diffusion Posterior Sampling for General Noisy Inverse Problems, ICLR, 2023

Diffusion for Image Restoration – SISR

- Our solution:

Solving Diffusion ODEs with Optimal Boundary Conditions for Better Image Super-Resolution

Yiyang Ma, *Huan Yang*, *Wenhan Yang*, *Jianlong Fu*, *Jiaying Liu* **ICLR 2024**



Diffusion for Image Restoration – SISR

- Existing Works:

  Does not manage to solve the problem of instability.

- Our Method:

A **stable sampling method** of diffusion-based SR models, **boosting all the diffusion-based SR models**.

- Analysis on the Approximately Optimal BC of Diffusion ODEs
- Obtaining the $\tilde{\mathbf{x}}_T$ with a refence set of image pairs
- Applying the $\tilde{\mathbf{x}}_T$ to all the LRs, achieving better results stably

Diffusion for Image Restoration – SISR

Preliminary: Diffusion ODE sampler

The results are determined by the boundary condition $\mathbf{x}_T$.

(*i.e.,* the Gaussian noise at the start of the sampling process).

e.g., DDIM [3], DPM-Solver [11]

Thus, the SR images can be a function of boundary condition (BC) and LR image:

$$\mathbf{x}_0 = h_\theta(\mathbf{x}_T, \mathbf{y})$$

[9] Jiaming Song, et al. Denoising Diffusion Implicit Models, ICLR, 2021
[10] Cheng Lu, et al. DPM-Solver: A Fast ODE Solver for Diffusion Probabilistic Model Sampling in Around 10 Steps, 2022

Diffusion for Image Restoration – SISR

- Our Solution: Optimal Boundary Condition $\mathbf{x}_T^*$

The BC that can generate the SR image which is the most close to the HR image:

$$\mathbf{x}_0^* = h_\theta(\mathbf{x}_T^*, \mathbf{y}) \text{ , where } \quad \mathbf{x}_0^* = \arg\max_{\mathbf{x}_0} p_\theta(\mathbf{x}_0|\mathbf{y})$$

We prove that the $\mathbf{x}_T^*$ is approximately consistent to different LRs:

$$\mathbf{x}_T^* = \underset{\mathbf{x}_T \sim \mathcal{N}(\mathbf{0},\mathbf{I})}{\arg\max} \; p_\theta(h_\theta(\mathbf{x}_T, \mathbf{y})) \approx \underset{\mathbf{x}_T \sim \mathcal{N}(\mathbf{0},\mathbf{I})}{\arg\max} \; p_\theta(h_\theta(\mathbf{x}_T, \mathbf{y}_i)), \forall \mathbf{y}_i \in \mathcal{C}$$

Diffusion for Image Restoration – SISR

- Find an Approximately Optimal BC $\tilde{\mathbf{x}}_T$

We employ a perceptual distance between SR and HR as an approximated implementation of the likelihood.

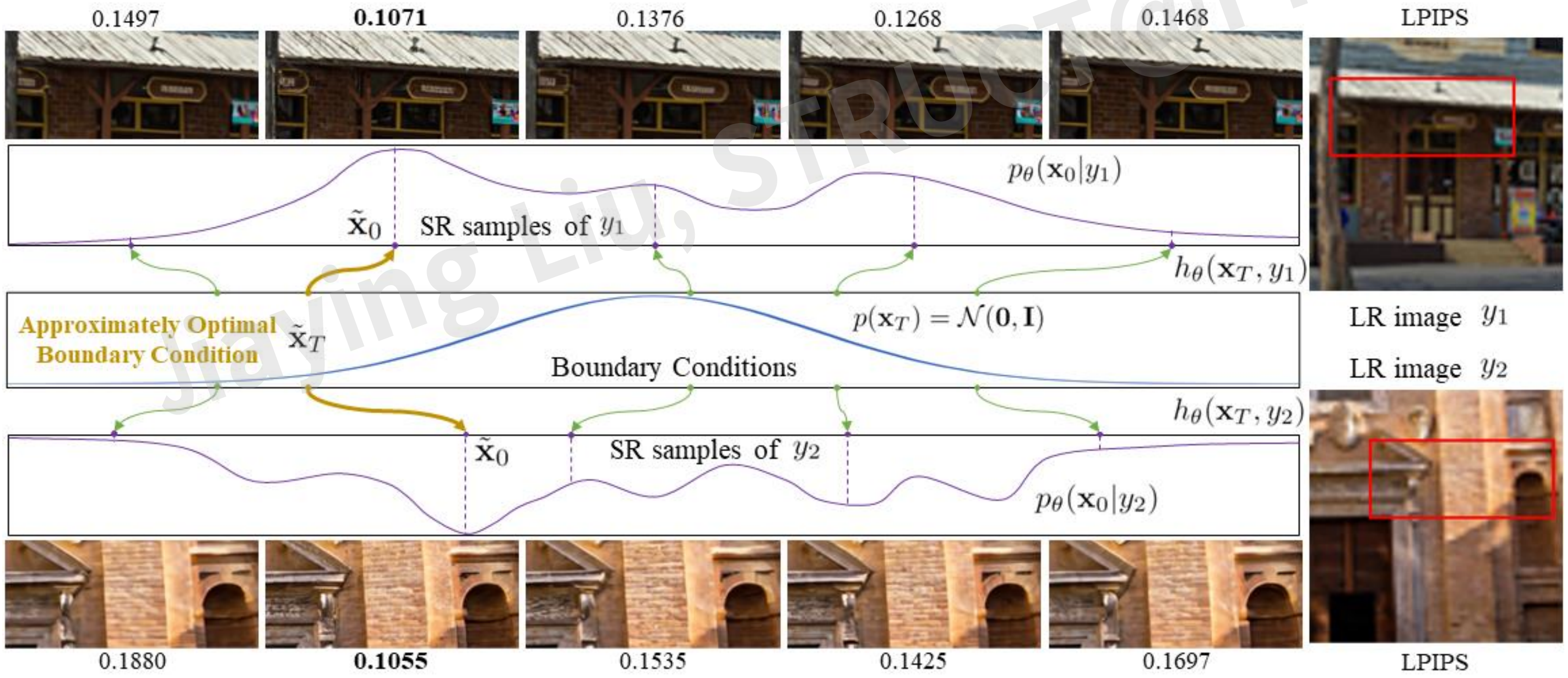
We build a reference set and calculate the $\tilde{\mathbf{x}}_T$ on it:

$$\tilde{\mathbf{x}}_T \approx \underset{\mathbf{x}_T \sim \mathcal{N}(\mathbf{0},\mathbf{I})}{\arg\min} \sum_{i=1}^{R} M(h_\theta(\mathbf{x}_T, \mathbf{y}_i), \mathbf{z}_i)$$

Diffusion for Image Restoration – SISR

- Characteristic of Approximately Optimal BC $\tilde{\mathbf{x}}_T$

As we have discussed before, the $\tilde{\mathbf{x}}_T$ is approximately shared:



The one which can be used to sample the best SR image.

- Characteristic of Approximately Optimal BC $\tilde{\mathbf{x}}_T$

As we have discussed before, the $\tilde{\mathbf{x}}_T$ is approximately shared:

The one which can be used to sample the best SR image.

Diffusion for Image Restoration – SISR

Experiments

- The proposed method can be implement on different SR models.

The results of SR3 on bicubic-SR.

Model (& sampling method)		DIV2k-test LPIPS ↓	DIV2k-test PSNR ↑	Urban100 LPIPS ↓	Urban100 PSNR ↑	BSD100 LPIPS ↓	BSD100 PSNR ↑
ESRGAN		0.1082	28.18	0.1226	23.04	0.1579	23.65
RankSRGAN		0.1171	27.98	0.1403	23.16	0.1714	23.80
SRDiff		0.1286	28.96	0.1391	23.88	0.2046	24.17
SR3	DDPM-1000	0.1075	28.75	0.1165	24.33	0.1555	23.86
	DDPM-250	0.1142	28.95	0.1181	24.41	0.1621	24.00
	DDPM-100	0.1257	29.16	0.1232	24.51	0.1703	24.15
	DPMS-20	0.1653	27.25	0.1413	23.46	0.2037	22.79
	DDIM-50	0.1483	28.55	0.1333	24.16	0.1823	23.75
	DDIM-100	0.1571	28.16	0.1335	24.05	0.1950	23.55
	DPMS-20 + $\tilde{\mathbf{x}}_T$	0.1210	27.45	0.1179	23.57	0.1687	22.81
	DDIM-50 + $\tilde{\mathbf{x}}_T$	0.1053	28.65	0.1164	24.26	0.1552	23.99
	DDIM-100 + $\tilde{\mathbf{x}}_T$	0.1032	28.48	0.1136	24.12	0.1505	23.67

Diffusion for Image Restoration – SISR

Experiments

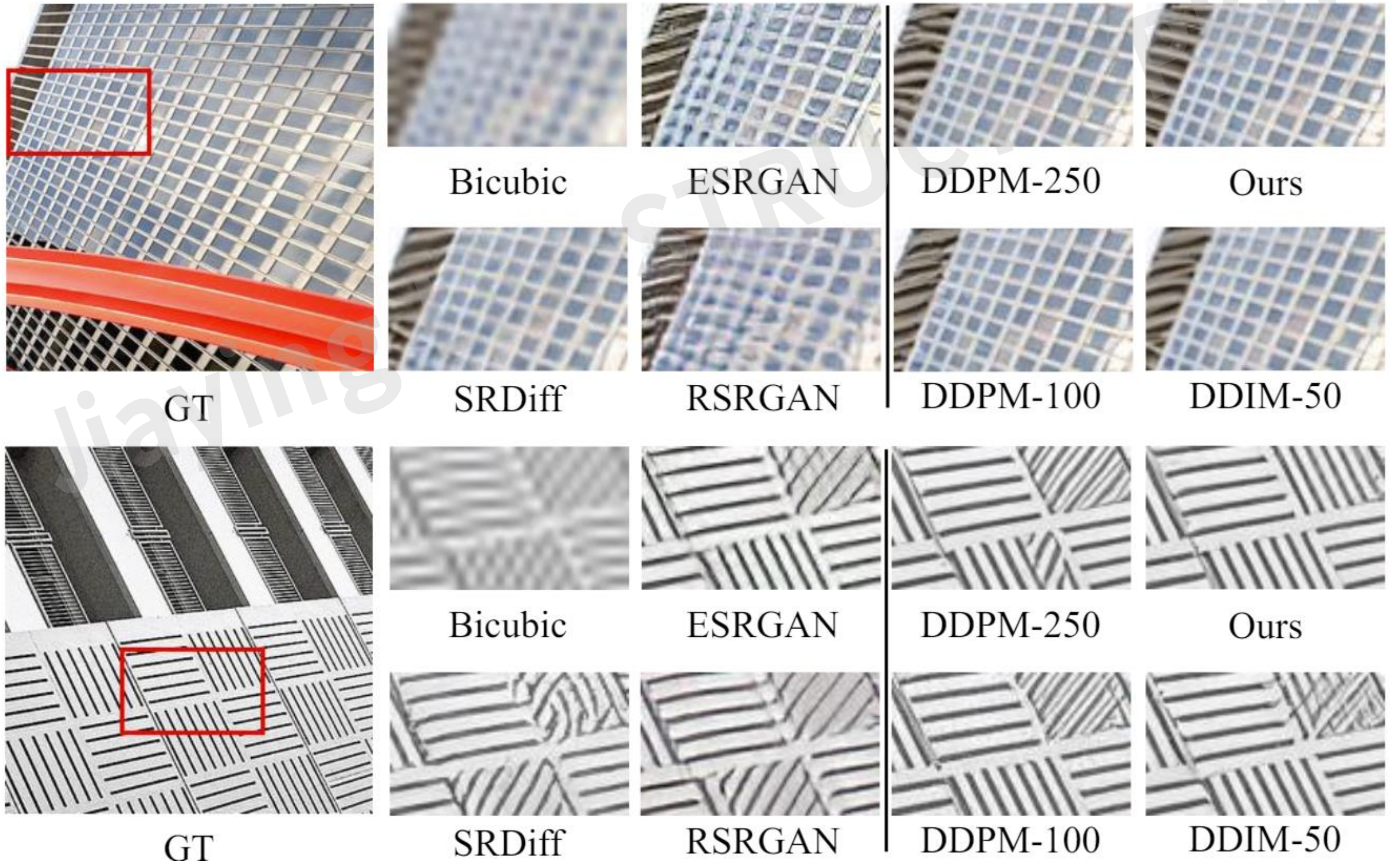
- The proposed method can be implement on different SR models.

The results of StableSR and DiffIR on real world-SR.

Model (& sampling method)		DIV2k-test DISTS ↓	DIV2k-test LPIPS ↓	DIV2k-test PSNR ↑	RealSR DISTS ↓	RealSR LPIPS ↓	RealSR PSNR ↑
RealSR		0.3051	0.5148	22.52	0.2532	0.3673	26.30
BSRGAN		0.2253	0.3416	22.13	0.2057	0.2582	25.52
DASR		0.2340	0.3444	22.02	0.2113	0.3014	26.32
Real-ESRGAN		0.2108	0.3109	22.36	0.2020	0.2511	25.12
KDSR-GAN		0.2022	0.2840	22.92	0.2006	0.2425	26.09
StableSR	DDPM-200	0.2010	0.3189	19.42	0.2210	0.3065	21.37
	DDIM-50	0.2217	0.3629	18.82	0.2336	0.3536	21.24
	DDIM-50 + $\tilde{\mathbf{x}}_T$	0.2046	0.3169	19.55	0.2164	0.2999	22.13
DiffIR	D-4	0.1773	0.2360	22.94	0.2076	0.2604	25.33
	D-4 + $\tilde{\mathbf{x}}_T$	0.1772	0.2357	22.95	0.1993	0.2419	25.82

Diffusion for Image Restoration – SISR

- Experiments
  - Subjective results comparisons:



Diffusion for Image Restoration – SISR

Summary of the method

We propose a **stable sampling method** of diffusion-based SR models, **boosting all the diffusion-based SR models**.

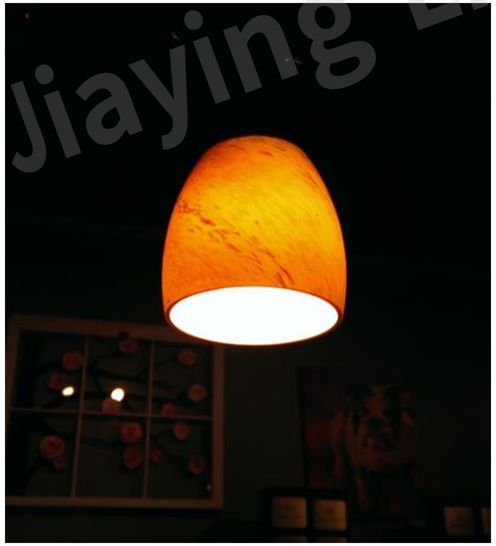
- We analyze the characteristics of BCs.
- We propose a method of approximating an optimal BC.
- The proposed method can be utilized in different diffusion-based SR models without any external training.

02

Diffusion for Low-light Enhancement

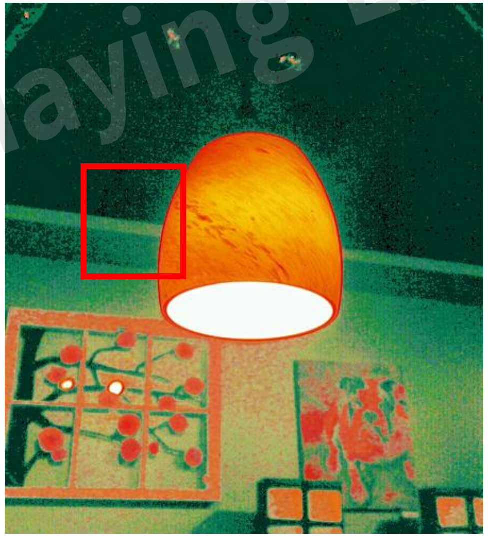
Diffusion for Low-Light Enhancement

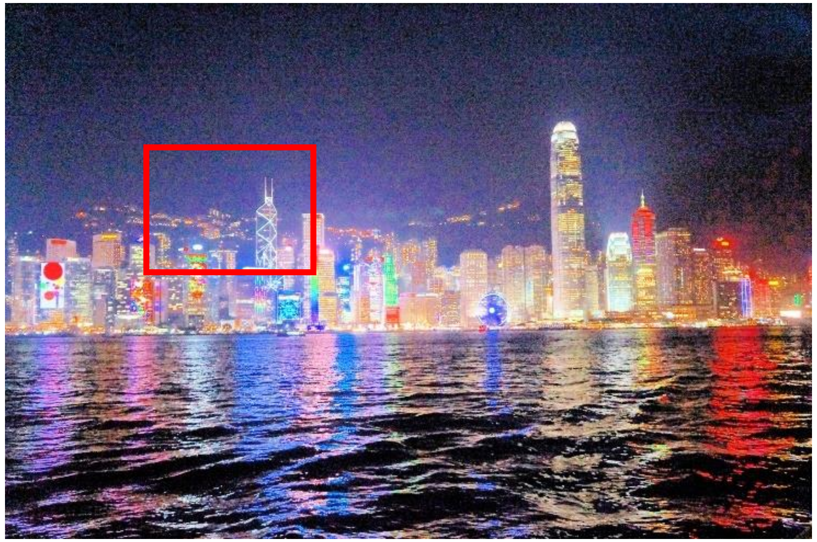
- Low visibility
  - Details are buried due to degraded contrast and low illumination

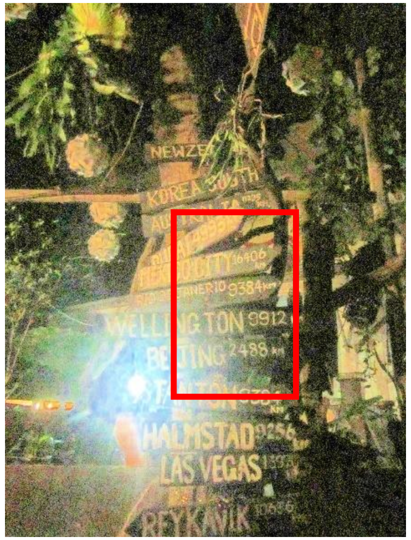


Diffusion for Low-Light Enhancement

- Intensive noises
  - After simple operations,
    e.g. histogram equalization,
    noises become visible





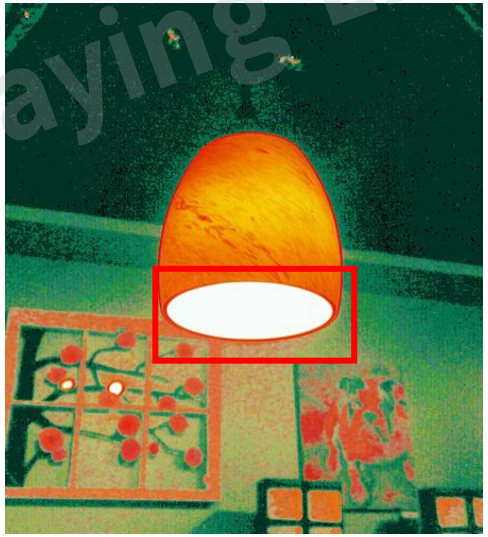


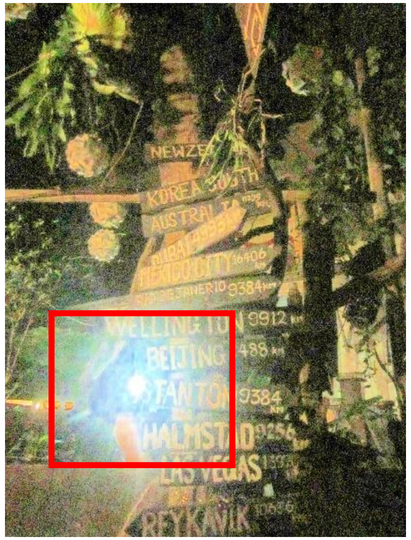
Diffusion for Low-Light Enhancement

- Intensive noises
  - After simple operations,
    e.g. histogram equalization,
    noises become visible

Diffusion for Low-Light Enhancement

- Non-uniform illumination
  - Under-exposures
  - Over-exposures



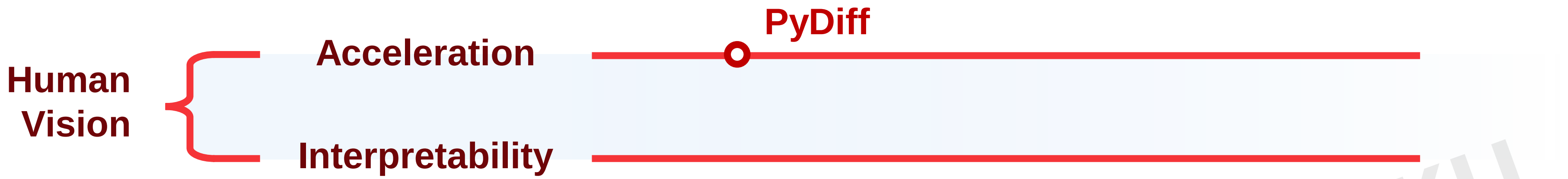


Diffusion for Low-Light Enhancement

Human Vision

- Acceleration
- Interpretability

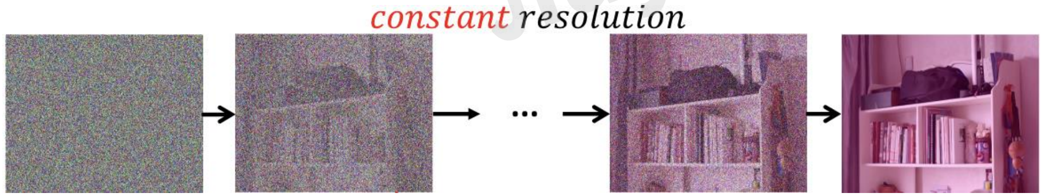
Diffusion for Low-Light Enhancement



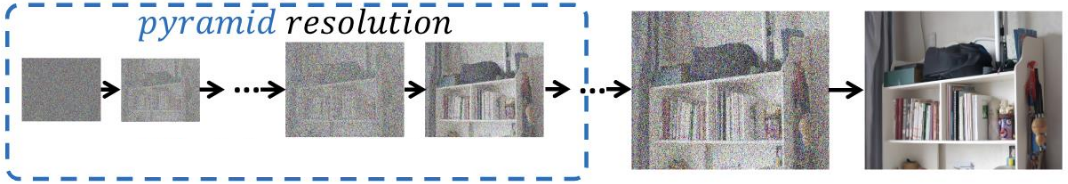
■ Pyramid Diffusion Model (PyDiff)

Problems of diffusion-based low-light enhancement

- Slow inference due to high resolution → Sampling in a **pyramid** resolution style



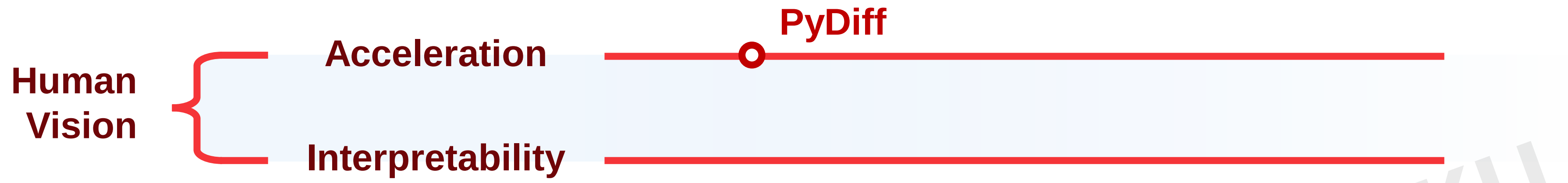
Vanilla diffusion



PyDiff

[1] Dewei Zhou, et al. Pyramid Diffusion Models For Low-light Image Enhancement, IJCAI 2023.

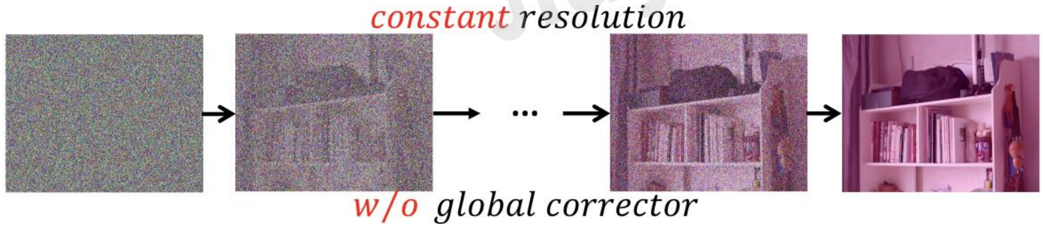
Diffusion for Low-Light Enhancement



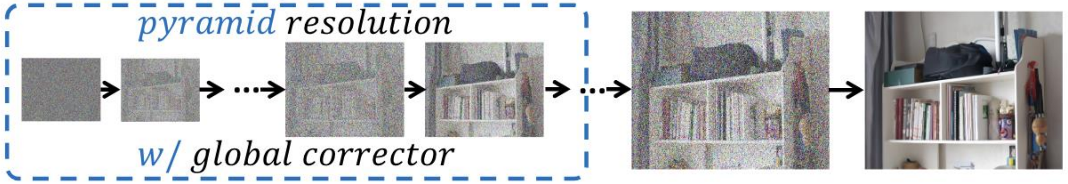
Pyramid Diffusion Model (PyDiff)

Problems of diffusion-based low-light enhancement

- Slow inference due to high resolution → Sampling in a **pyramid** resolution style
- Global degradation (*e.g.*, RGB shift) → **A global corrector**



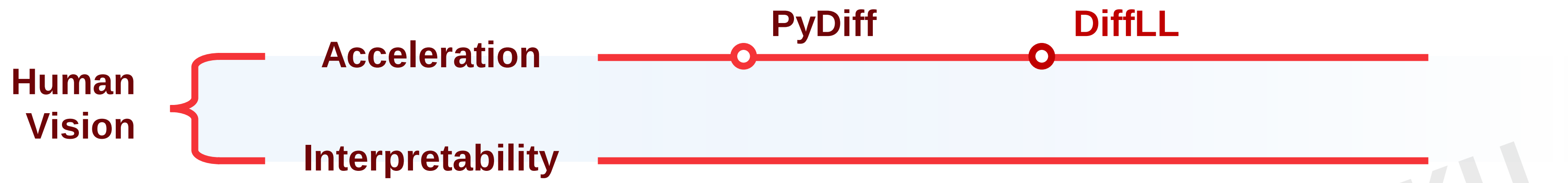
Vanilla diffusion



PyDiff

[1] Dewei Zhou, et al. Pyramid Diffusion Models For Low-light Image Enhancement, IJCAI 2023.

Diffusion for Low-Light Enhancement



DiffLL

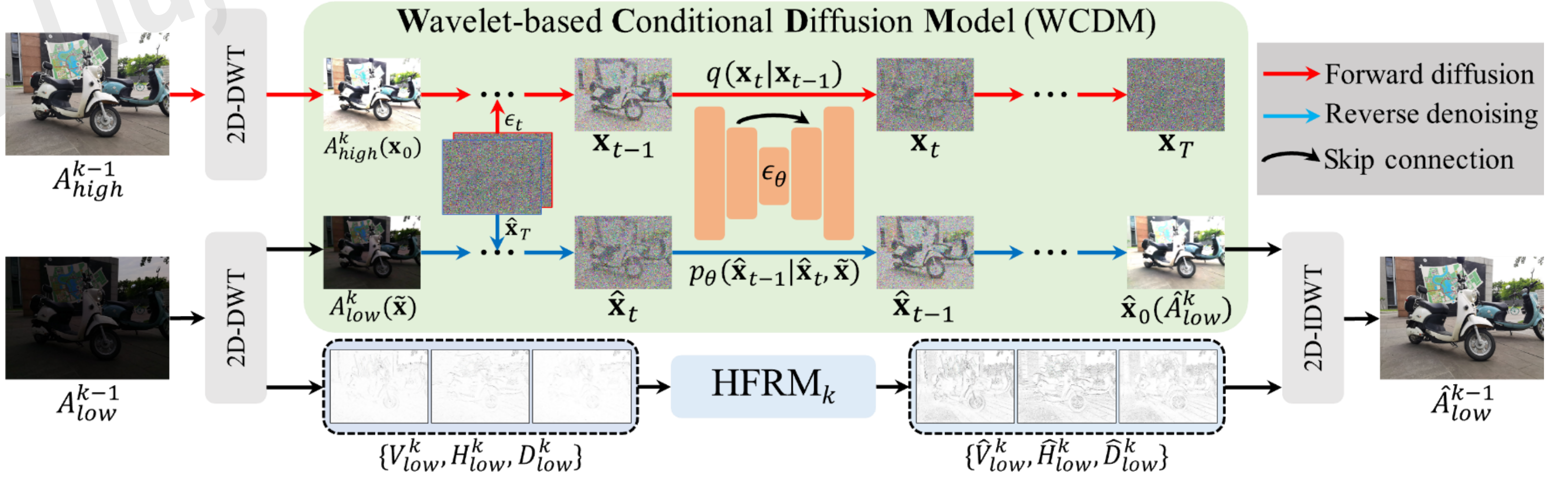
Wavelet transformation: accelerate inference

- *Diffusion-based* WCDM

  restores **the average coefficient** A_{low}^{K}

- *Conv-based* HFRM

  restores **high-frequency coefficients** $\{V_{low}^{K}, H_{low}^{K}, D_{low}^{K}\}$



[1] Hai Jiang, et al. Low-light Image Enhancement with Wavelet-based Diffusion Models, TOG 2023.

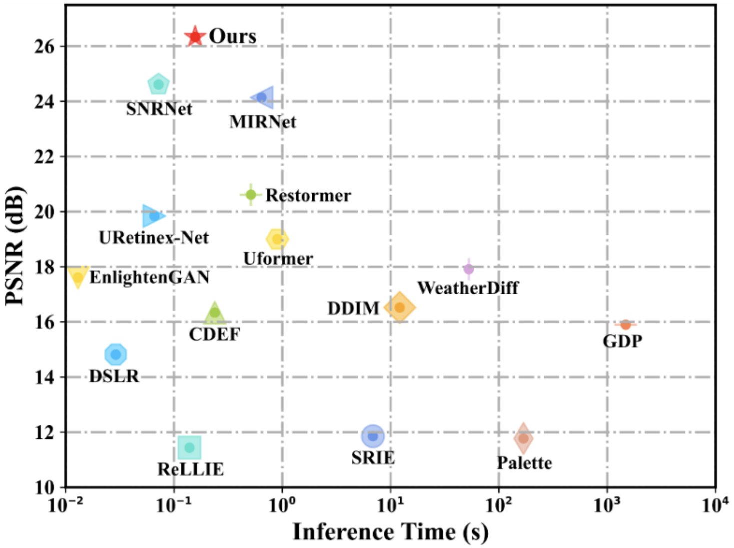
Diffusion for Low-Light Enhancement

Human Vision

Acceleration — PyDiff — DiffLL

Interpretability

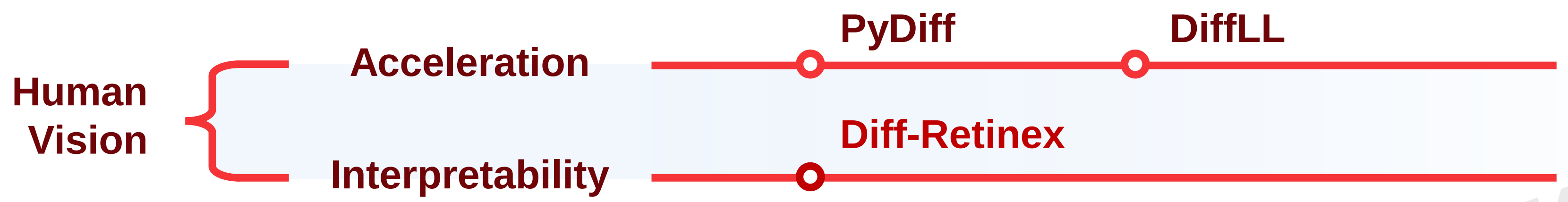
- DiffLL



(a) Performance vs. Efficiency

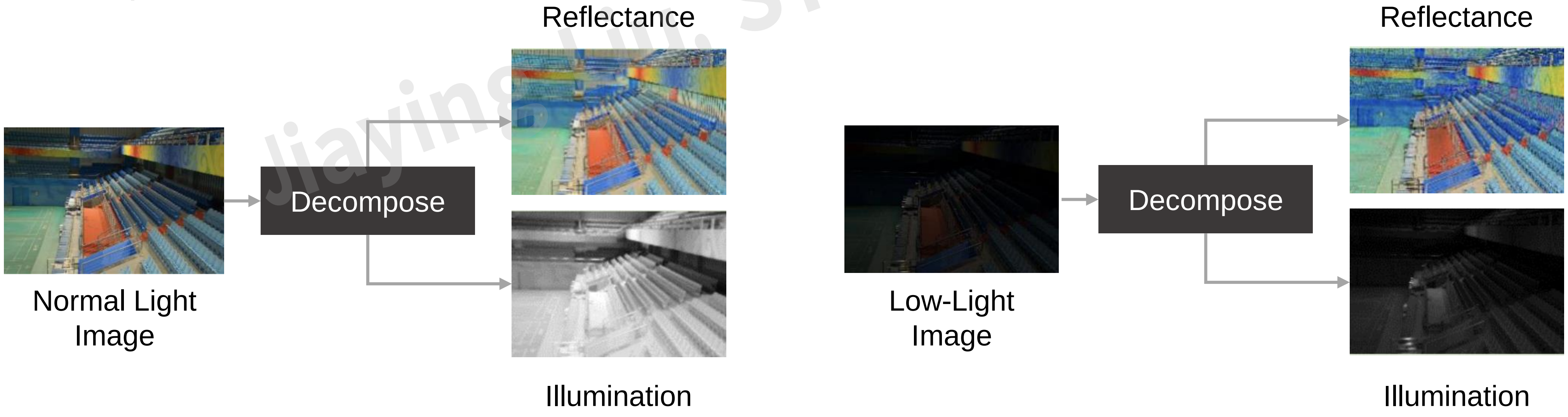
[1] Hai Jiang, et al. Low-light Image Enhancement with Wavelet-based Diffusion Models, TOG 2023.

Diffusion for Low-Light Enhancement



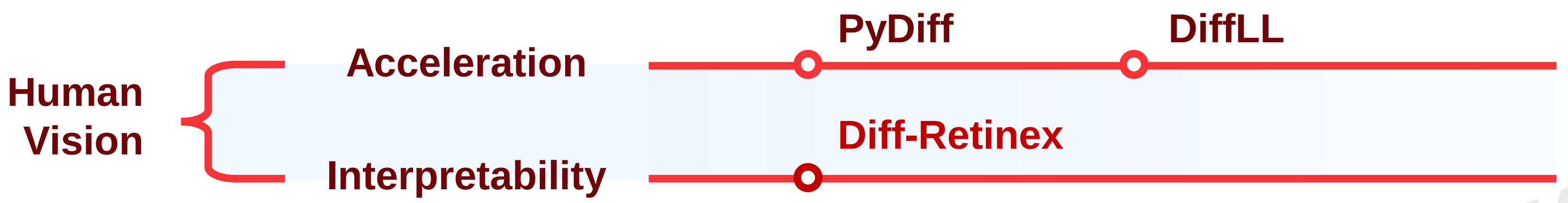
- Diff-Retinex

Retinex decomposition



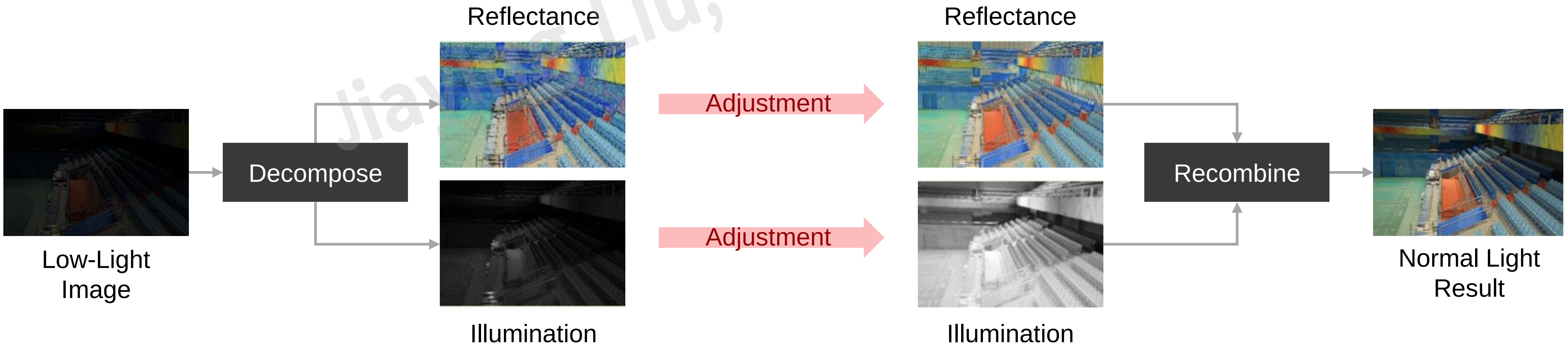
[1] Xunpeng Yi, et al. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model, ICCV 2023.

Diffusion for Low-Light Enhancement



Diff-Retinex

Retinex-based low-light enhancement



[1] Xunpeng Yi, et al. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model, ICCV 2023.

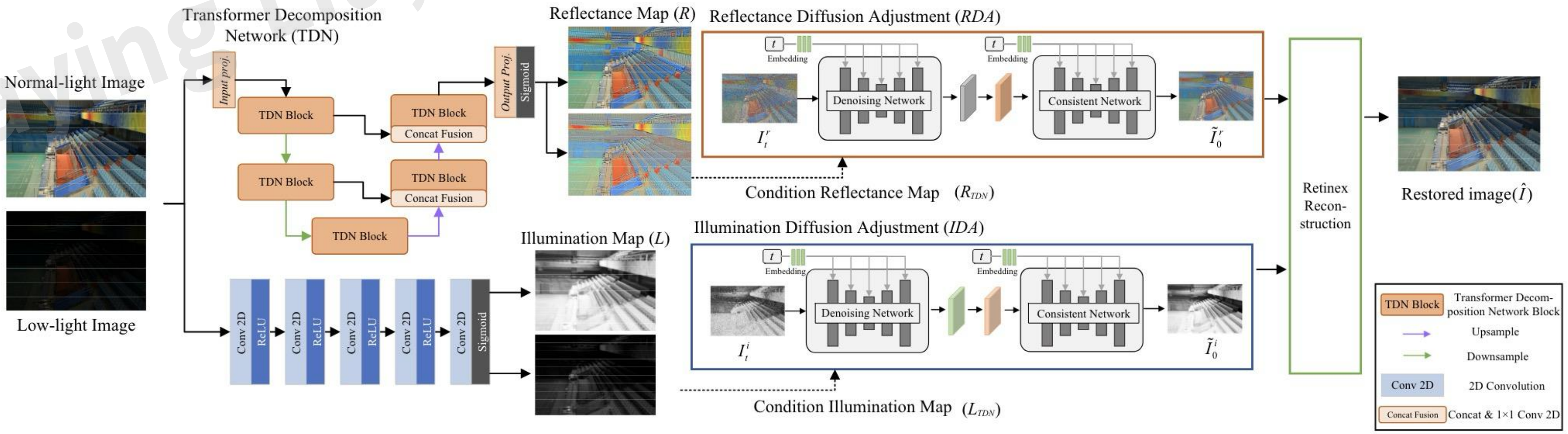
Diffusion for Low-Light Enhancement

Human Vision
- Acceleration: PyDiff, DiffLL
- Interpretability: Diff-Retinex

Diff-Retinex

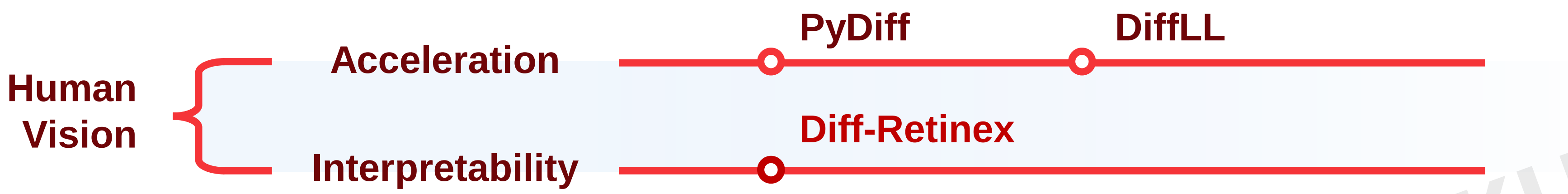
Physically explainable (**Retinex decomposition**) + Diffusion

- Diffusion for adjusting reflectance and illumination

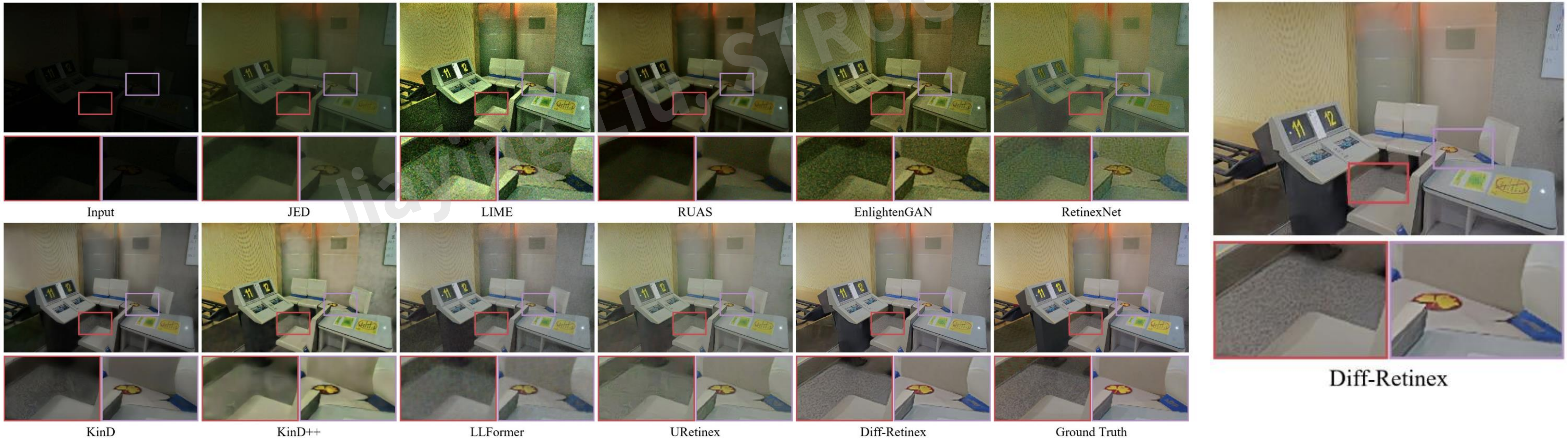


[1] Xunpeng Yi, et al. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model, ICCV 2023.

Diffusion for Low-Light Enhancement

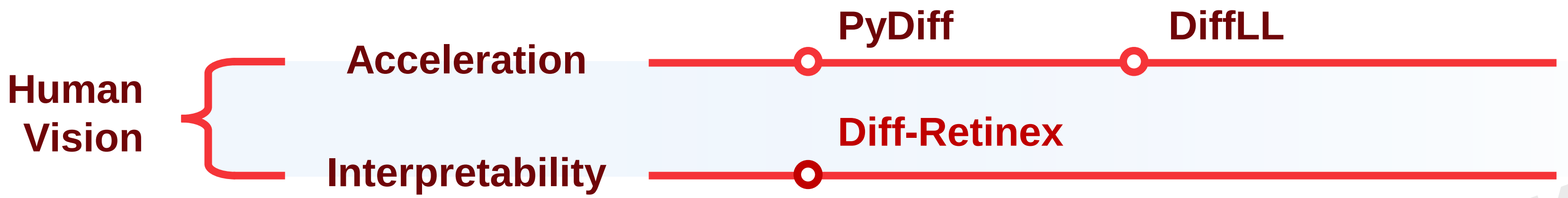


Diff-Retinex



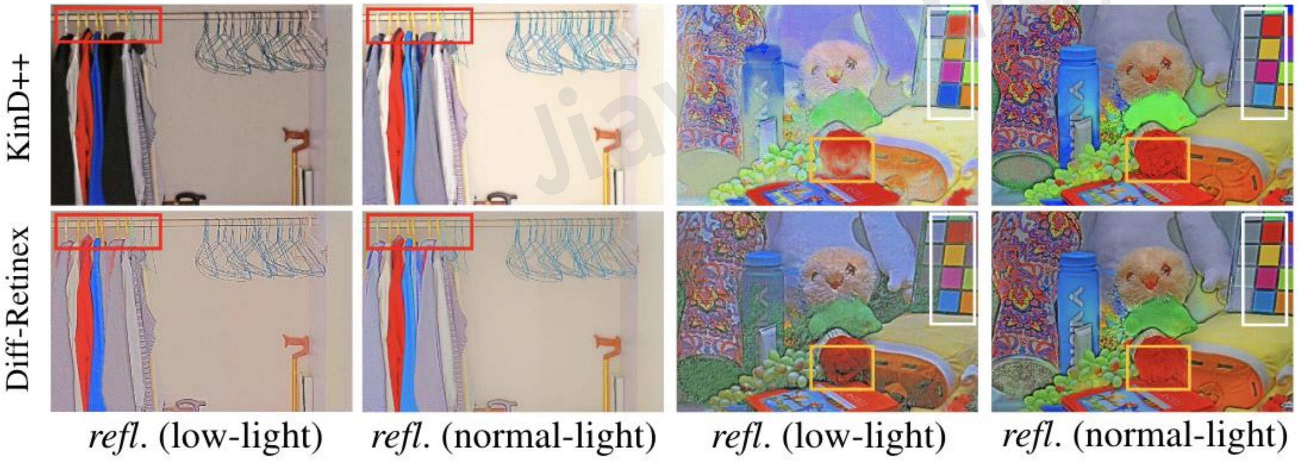
[1] Xunpeng Yi, et al. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model, ICCV 2023.

Diffusion for Low-Light Enhancement

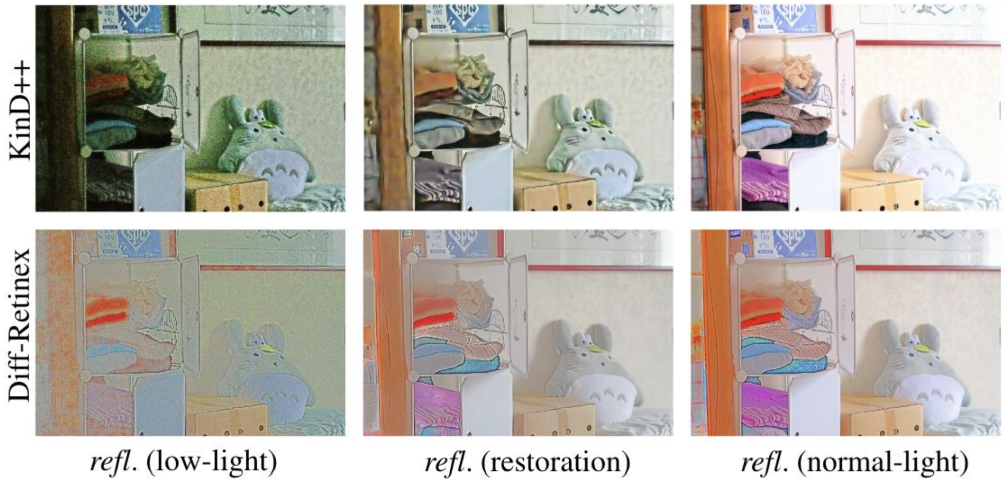


■ Diff-Retinex

Results of Retinex-decomposition

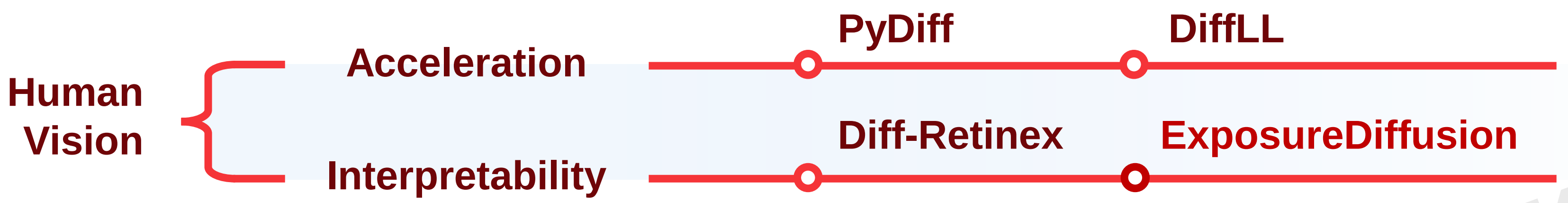


Visualization of restoring the reflectance map



[1] Xunpeng Yi, et al. Diff-Retinex: Rethinking Low-light Image Enhancement with A Generative Diffusion Model, ICCV 2023.

Diffusion for Low-Light Enhancement



■ ExposureDiffusion

Physically explainable (**exposure process**) + Diffusion

[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

Diffusion for Low-Light Enhancement

Human Vision

- Acceleration: PyDiff, DiffLL
- Interpretability: Diff-Retinex, ExposureDiffusion

ExposureDiffusion

Physically explainable (**exposure process**) + Diffusion

- A raw image can be formulated as

$$X_t = \lambda_t K I + K N_p + N_{ind}$$

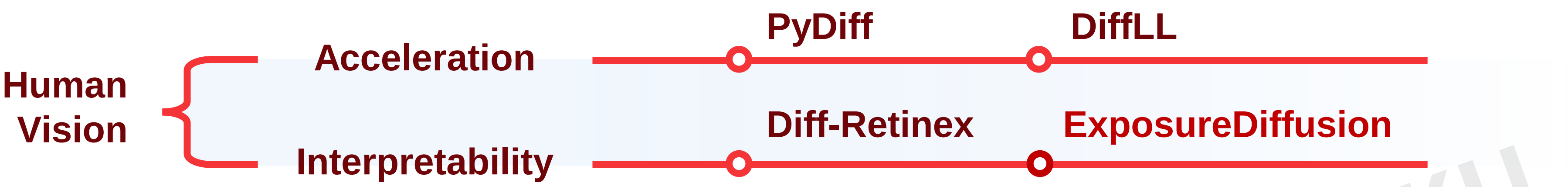
$$(\lambda_t I + N_p) \sim \mathcal{P}(\lambda_t I)$$

Poisson distribution

- λ_t exposure time
- K overall system gain
- I rate of the photoelectrons
- N_p photon shot noise
- N_{ind} signal independent noise

[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

Diffusion for Low-Light Enhancement



ExposureDiffusion

Physically explainable (**exposure process**) + Diffusion

- A raw image can be formulated as

$$X_t = \underbrace{\lambda_t K I + K N_p}_{} + N_{ind}$$

$$(\lambda_t I + N_p) \sim \mathcal{P}(\lambda_t I)$$

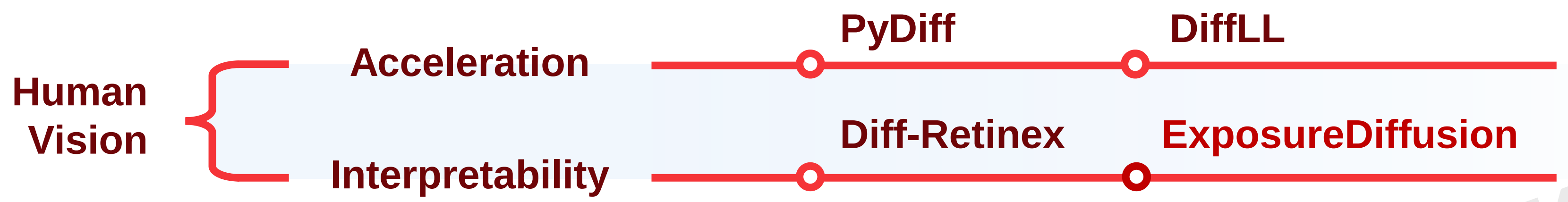
Poisson distribution

- λ_t exposure time
- K overall system gain
- I rate of the photoelectrons
- N_p photon shot noise
- N_{ind} signal independent noise

	Conditional diffusion	Exposure diffusion
Objective	maximizing $p_\Theta(X\|Y)$	minimizing KL divergence with the real exposure process
Initial state X_T	$X_T \sim \mathcal{N}(0,1)$	$X_T \sim q(X_T)$
Assumption	$q(X_t\|X_{t-1}) := \mathcal{N}(X_t; \sqrt{1-\beta_t}X_{t-1}, \beta_t\mathbf{I})$	$q(X_{t-1}\|X_t, X_{ref}) := \mathcal{P}(\frac{X_{t-1}-X_t}{K}; \frac{(\lambda_{t-1}-\lambda_t)X_{ref}}{\lambda_{ref}K})$
Reverse process	$p_\Theta(X_{t-1}\|X_t, Y) := \mathcal{N}(X_{t-1}; \mu_\Theta(X_t, Y, t), \sigma_t^2\mathbf{I})$	$p_\Theta(X_{t-1}\|X_t) := \mathcal{P}(\frac{X_{t-1}-X_t}{K}; \frac{(\lambda_{t-1}-\lambda_t)F_\Theta(X_t)}{\lambda_{ref}K})$
Training	The expectation over $q(X_t\|X_0)$	The expectation over $p_\Theta(X_t\|X_T)$

[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

Diffusion for Low-Light Enhancement

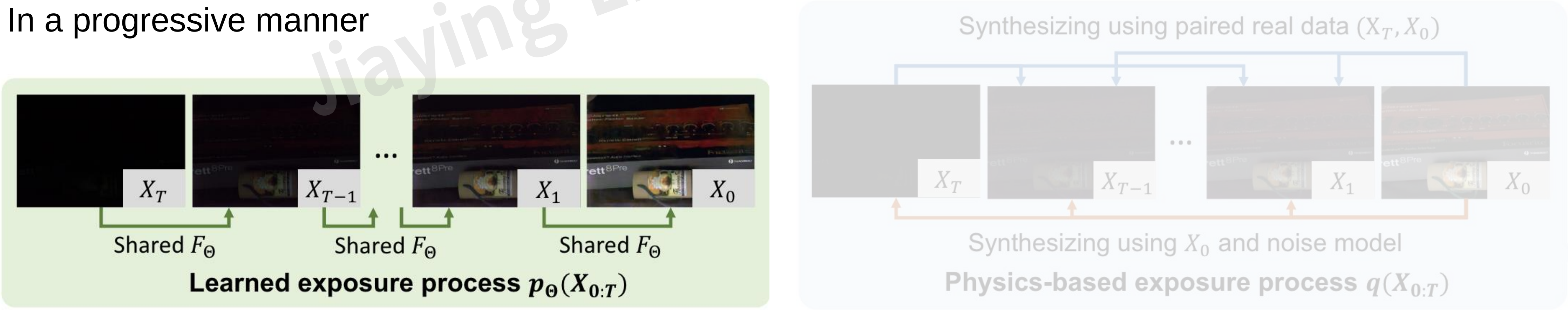


ExposureDiffusion

Physically explainable (**exposure process**) + Diffusion

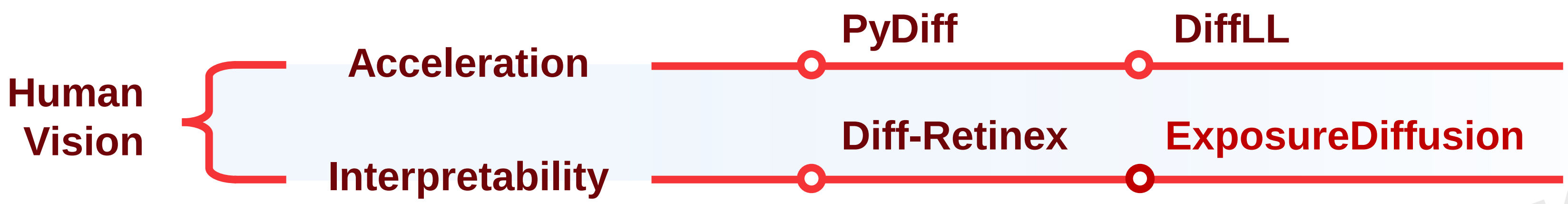
- Simulate using a **shared** neural network
- In a progressive manner

Applicable to both real-captured and synthetic noise models



[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

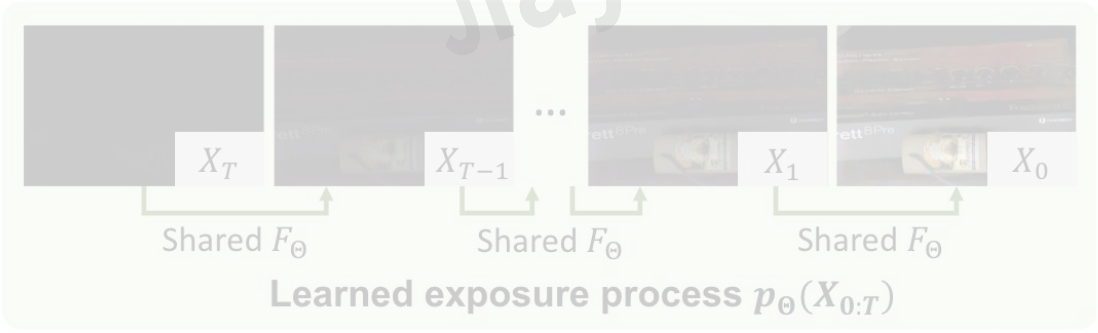
Diffusion for Low-Light Enhancement



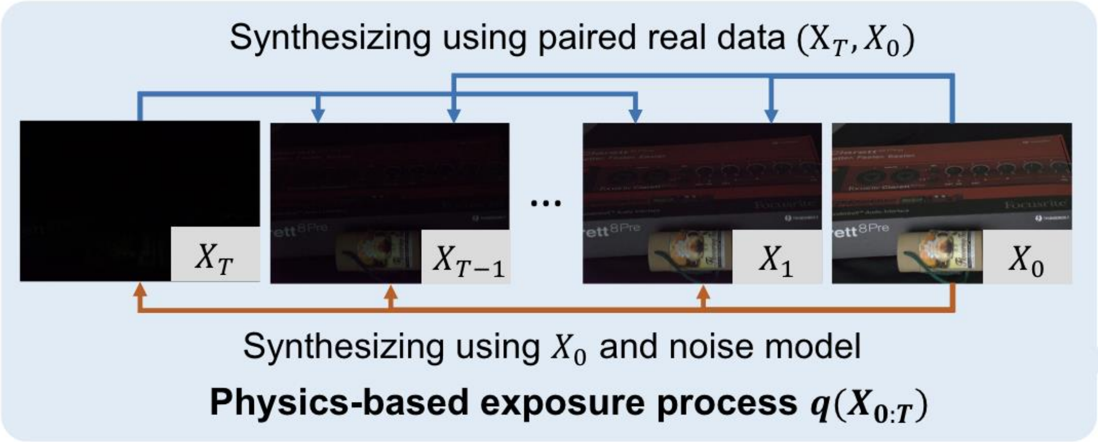
ExposureDiffusion

Physically explainable (**exposure process**) + Diffusion

- Simulate using a **shared** neural network
- In a progressive manner



Applicable to both real-captured and synthetic noise models



[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

Diffusion for Low-Light Enhancement

Human Vision

- Acceleration: PyDiff → DiffLL
- Interpretability: Diff-Retinex → ExposureDiffusion

■ ExposureDiffusion

Input BM3D N2N P+G Ours Reference

[1] Yufei Wang, et al. ExposureDiffusion: Learning to Expose for Low-light Image Enhancement, ICCV 2023.

Diffusion for Low-Light Enhancement

Human Vision

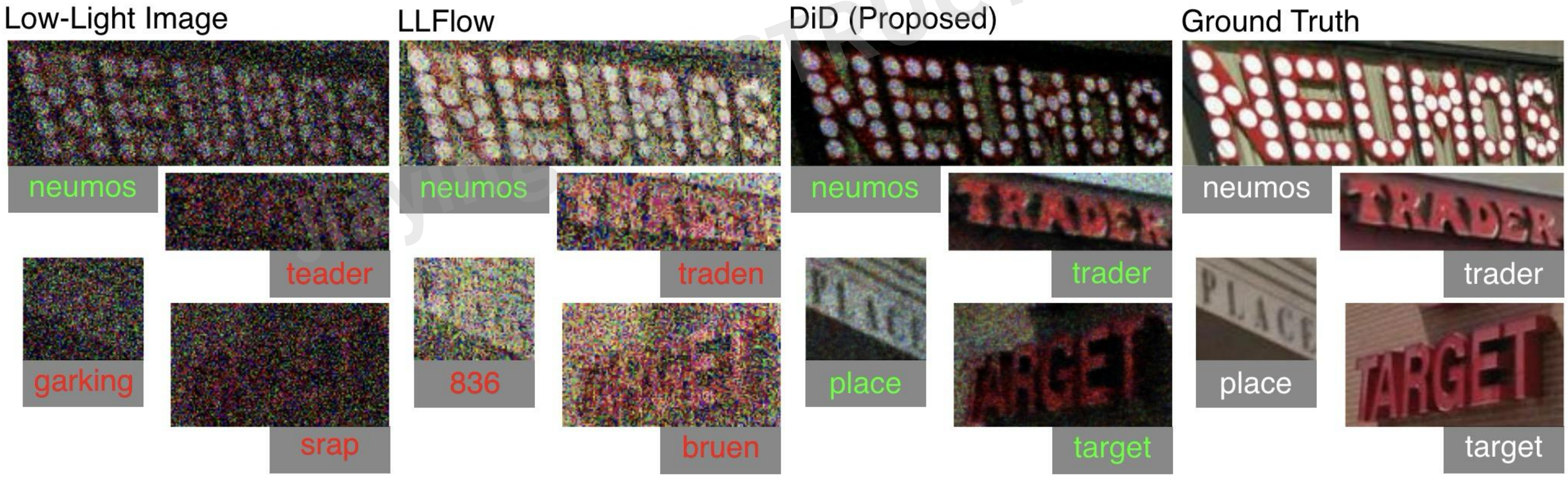
Acceleration — PyDiff — DiffLL

Interpretability — Diff-Retinex — ExposureDiffusion

Machine Vision

Diffusion for Low-Light Enhancement

■ Diffusion in the Dark (DiD) for low-light text recognition



[1] Cindy M. Nguyen, et al. Diffusion in the Dark: A Diffusion Model for Low-Light Text Recognition, WAIC 2024.

Diffusion for Low-Light Enhancement

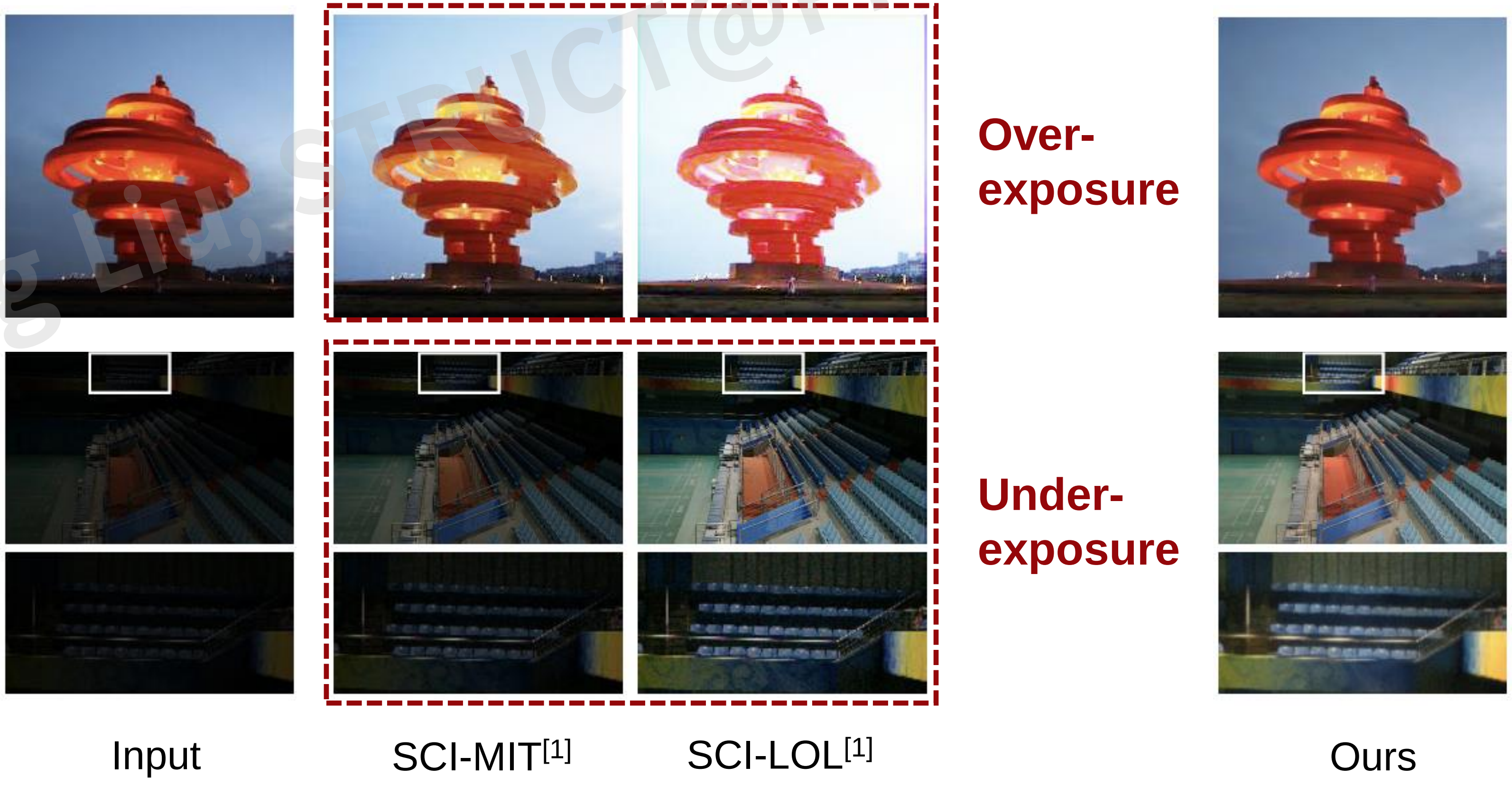
■ Problem of existing works

Cannot maintain a consistent brightening effect in dark and moderately dark images

Essential problem:

Lack a genuine

concept of illumination



[1] Long Ma, et al. Toward fast, flexible, and robust low-light image enhancement. In CVPR, 2022.

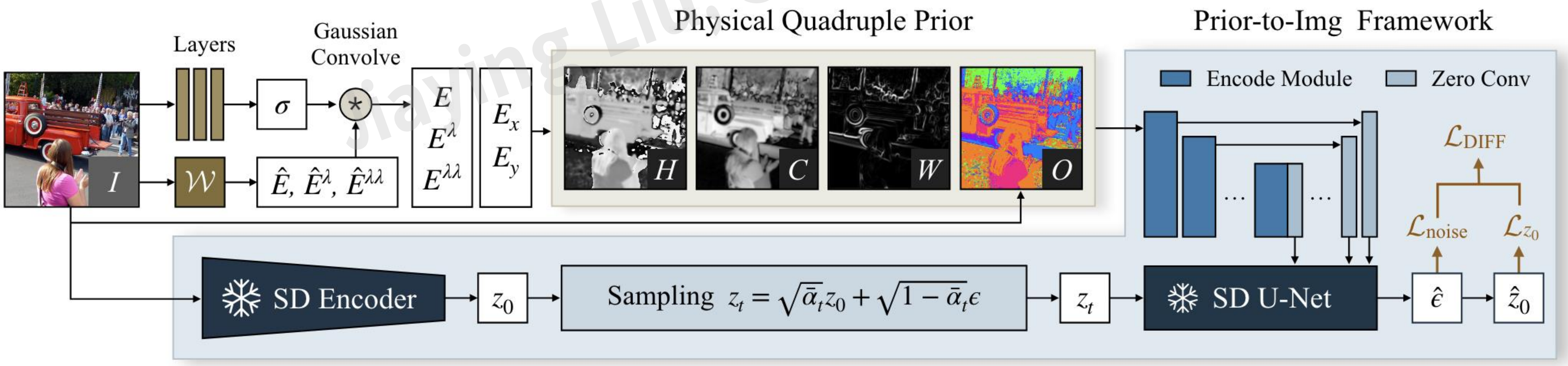
Diffusion for Low-Light Enhancement

- Our solution:

Zero-Reference Low-Light Enhancement via Physical Quadruple Priors

Wenjing Wang, Huan Yang, Jianlong Fu, Jiaying Liu

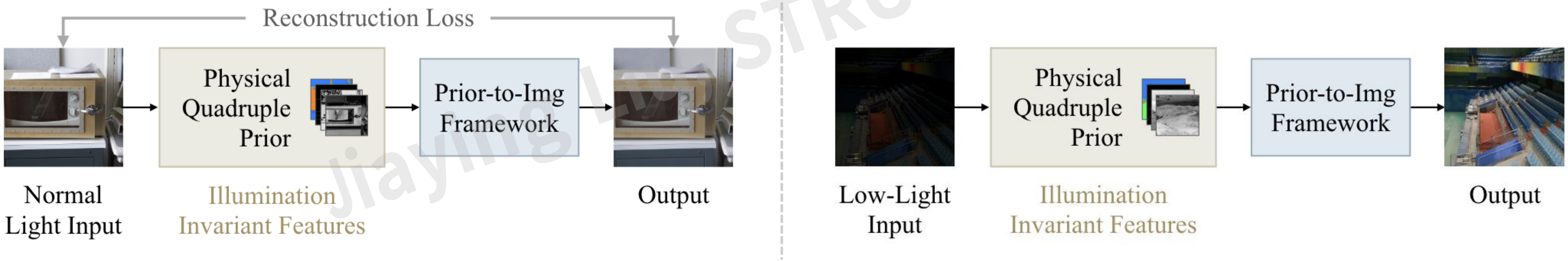
CVPR 2024



Diffusion for Low-Light Enhancement

Our solution:

- Develop an illumination-invariant prior (Physical Quadruple Prior)
- Employ it as an intermediary between low-light and normal light images

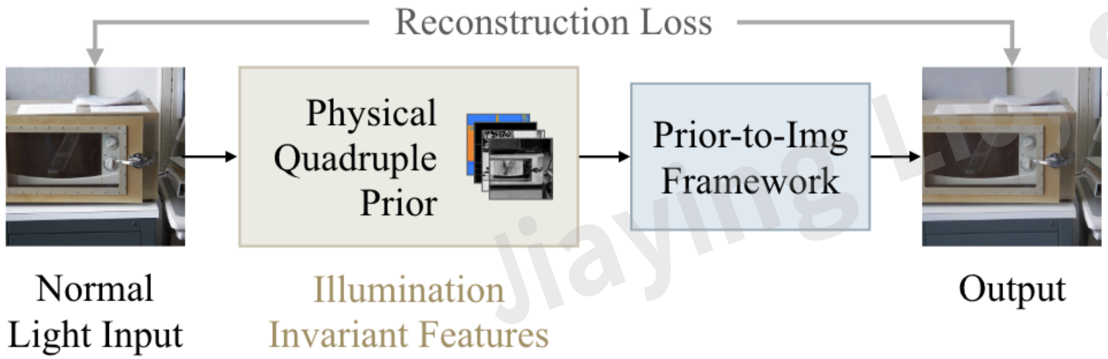


Training on normal light images

Inference on low-light images

Diffusion for Low-Light Enhancement

- Our solution:
  - Develop an illumination-invariant prior (Physical Quadruple Prior)
  - Employ it as an intermediary between low-light and normal light images



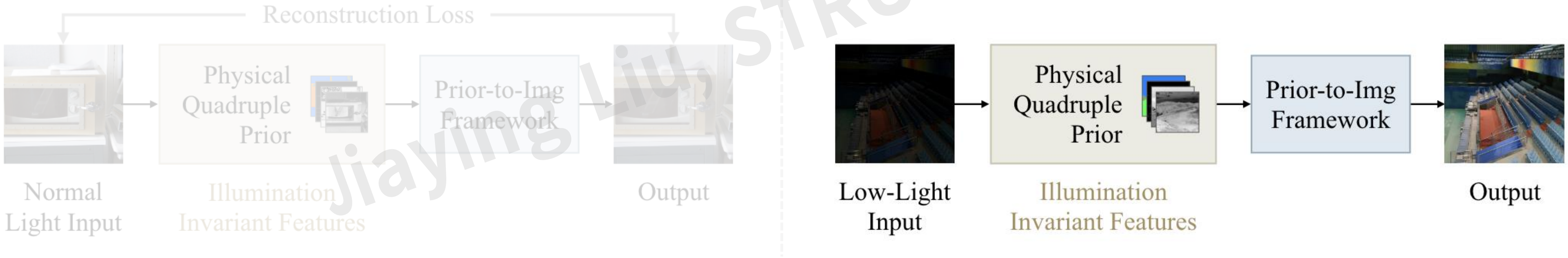
Training on normal light images

Inference on low-light images

Diffusion for Low-Light Enhancement

Our solution:

- Develop an illumination-invariant prior (Physical Quadruple Prior)
- Employ it as an intermediary between low-light and normal light images



Training on normal light images

Inference on low-light images

Diffusion for Low-Light Enhancement

- The Kubelka-Munk theory of light transfer

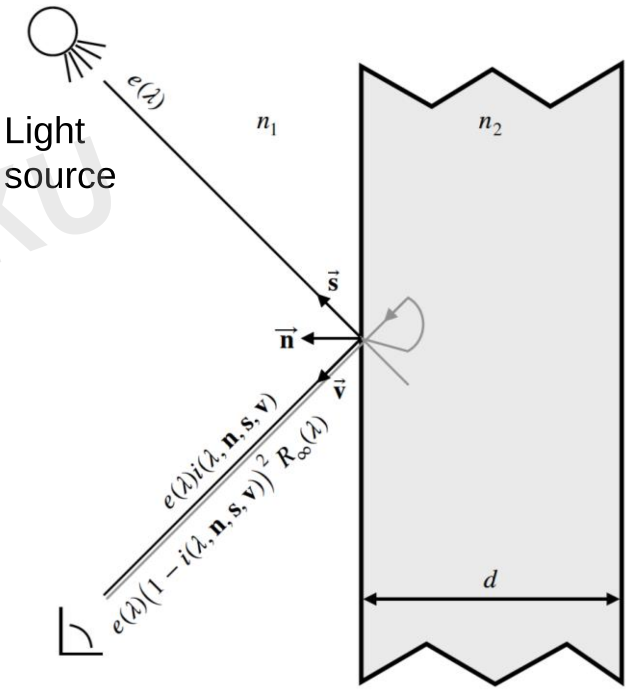
$$E(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x}) \left((1 - i(\mathbf{x}))^2 R_\infty(\lambda, \mathbf{x}) + i(\mathbf{x}) \right),$$

$e(\lambda, \mathbf{x})$: the spectrum of the light source

$i(\mathbf{x})$: the specular reflection

Illumination dependent

$R_\infty(\lambda, \mathbf{x})$: the material reflectivity → Illumination invariant



- when the object is matte, i.e., $i(\mathbf{x}) \approx 0$

$$E(\lambda, \mathbf{x}) = e(\lambda, \mathbf{x}) R_\infty(\lambda, \mathbf{x}),$$

the Retinex theory is a special case

[1] Theo Gevers, et al. Color in computer vision: Fundamentals and applications. 2012.

Diffusion for Low-Light Enhancement

- Through simplifying assumptions, obtain a series of invariants
  - Assuming equal energy illumination

$$E(\lambda, \mathbf{x}) = \tilde{e}(\mathbf{x})\left((1 - i(\mathbf{x}))^2 R_\infty(\lambda, \mathbf{x}) + i(\mathbf{x})\right)$$

- Then we can obtain

$$\frac{E^\lambda}{E^{\lambda\lambda}} = \frac{\tilde{e}(\mathbf{x})(1 - i(\mathbf{x}))^2 R_\infty^\lambda}{\tilde{e}(\mathbf{x})(1 - i(\mathbf{x}))^2 R_\infty^{\lambda\lambda}} = \boxed{\frac{R_\infty^\lambda}{R_\infty^{\lambda\lambda}}},$$

Illumination Invariant

- Set $H = \arctan\left(E^\lambda / E^{\lambda\lambda}\right)$, we can find that H is illumination invariant

[1] Jan-Mark Geusebroek, et al. Color invariance. IEEE TPAMI, 2001.

Diffusion for Low-Light Enhancement

- Through simplifying assumptions, obtain a series of invariants
  - Assuming equal energy illumination

$$E(\lambda, \mathbf{x}) = \tilde{e}(\mathbf{x}) \left((1 - i(\mathbf{x}))^2 R_\infty(\lambda, \mathbf{x}) + i(\mathbf{x}) \right) \Rightarrow H = \arctan \left(E^\lambda / E^{\lambda\lambda} \right)$$

[1] Jan-Mark Geusebroek, et al. Color invariance. IEEE TPAMI, 2001.

Diffusion for Low-Light Enhancement

- Through simplifying assumptions, obtain a series of invariants
  - Assuming equal energy illumination

$$E(\lambda, \mathbf{x}) = \tilde{e}(\mathbf{x}) \left((1 - i(\mathbf{x}))^2 R_\infty(\lambda, \mathbf{x}) + i(\mathbf{x}) \right) \Rightarrow H = \arctan\left(E^\lambda / E^{\lambda\lambda} \right)$$

  - Further assuming that the surface is matte

$$E(\lambda, \mathbf{x}) = \tilde{e}(\mathbf{x}) R_\infty(\lambda, \mathbf{x}), \Rightarrow C = \log\left(\frac{(E^\lambda)^2 + (E^{\lambda\lambda})^2}{E(\lambda, \mathbf{x})^2} \right)$$

  - Further assuming uniform illumination

$$E(\lambda, \mathbf{x}) = \bar{e} R_\infty(\lambda, \mathbf{x}), \Rightarrow W = \tan\left(\left| \frac{\partial E(\lambda, \mathbf{x})}{\partial \mathbf{x}} \frac{1}{E(\lambda, \mathbf{x})} \right| \right)$$

[1] Jan-Mark Geusebroek, et al. Color invariance. IEEE TPAMI, 2001.

Diffusion for Low-Light Enhancement

- Through simplifying assumptions, obtain a series of invariants

$$H = \arctan\left(E^{\lambda}/E^{\lambda\lambda}\right) \qquad C = \log\left(\frac{(E^{\lambda})^2 + (E^{\lambda\lambda})^2}{E(\lambda, \mathbf{x})^2}\right) \qquad W = \tan\left(\left|\frac{\partial E(\lambda, \mathbf{x})}{\partial \mathbf{x}} \frac{1}{E(\lambda, \mathbf{x})}\right|\right)$$

- Assuming that illumination maintains the order of colors, we propose the order of the RGB three channels as a fundamental illumination-invariant feature

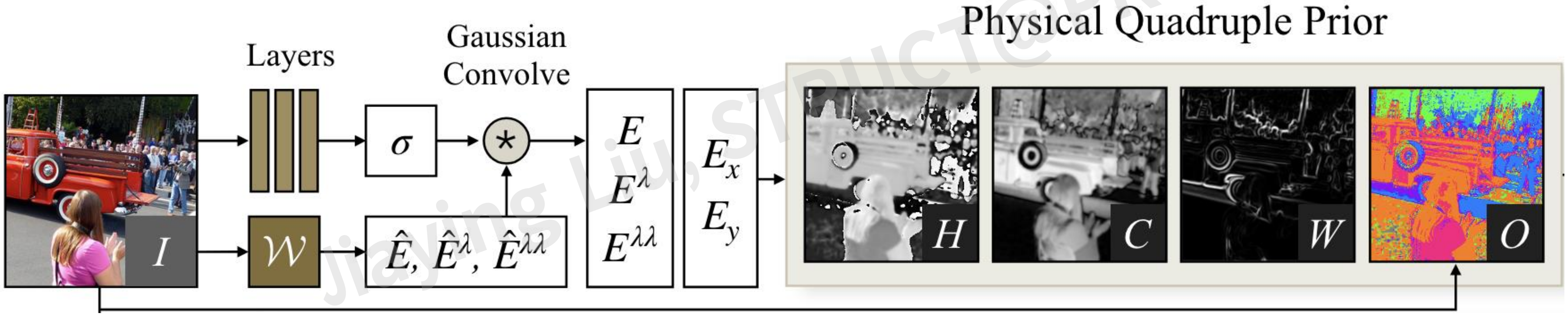
$$O(x, y) = [O_R(x, y), O_G(x, y), O_B(x, y)]$$

- *H*, *C*, *W*, and *O* are concatenated in the channel dimension to form our **physical quadruple prior**

Diffusion for Low-Light Enhancement

- Learnable Illumination-Invariant Prior

  Learn within deep neural networks

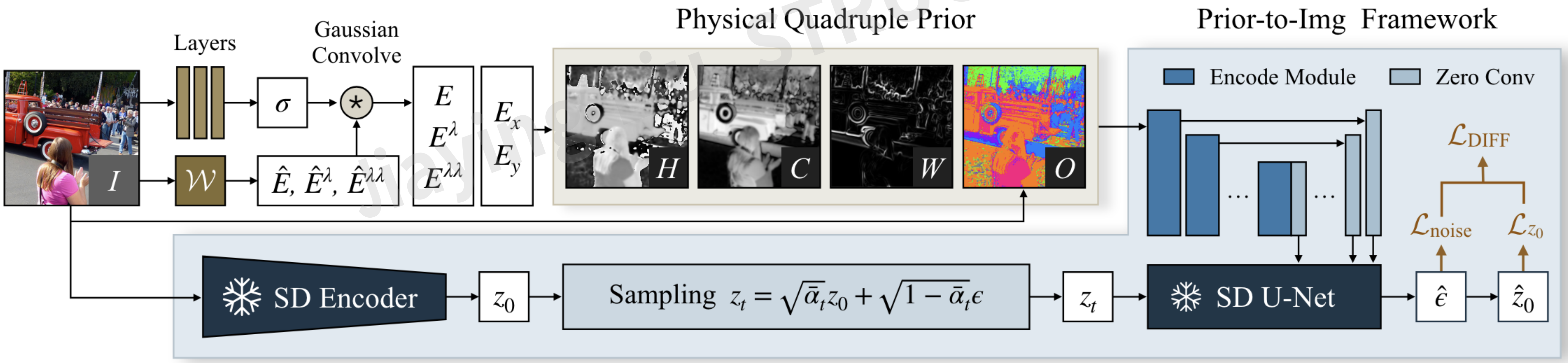


Diffusion for Low-Light Enhancement

Prior-to-Image via Diffusion Models

Leverage the knowledge of a pretrained image generative model

Use **Physical Quadruple Prior** as the condition of generation



Diffusion for Low-Light Enhancement

- Prior-to-Image via Diffusion Models

Issues of generative models: *slow convergence, detail degradation*

Diffusion for Low-Light Enhancement

- Prior-to-Image via Diffusion Models

  Issues of generative models: *slow convergence, detail degradation*

- Improvements to make it more suitable for image restoration task

  1. Training objective

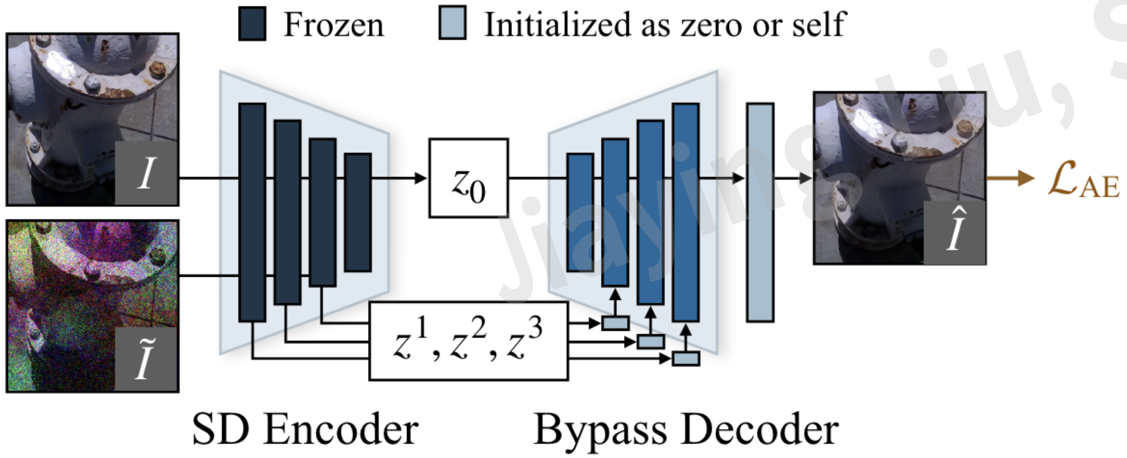
$$\boxed{\mathcal{L}_{z_0} = ||z_0 - \hat{z}_0||_2^2 = ||z_0 - \frac{z_t - \sqrt{1-\bar{\alpha}_t}\hat{\epsilon}}{\sqrt{\bar{\alpha}_t}}||_2^2}$$

$$\boxed{\mathcal{L}_{\text{noise}} = ||\epsilon - \hat{\epsilon}||_2^2}$$

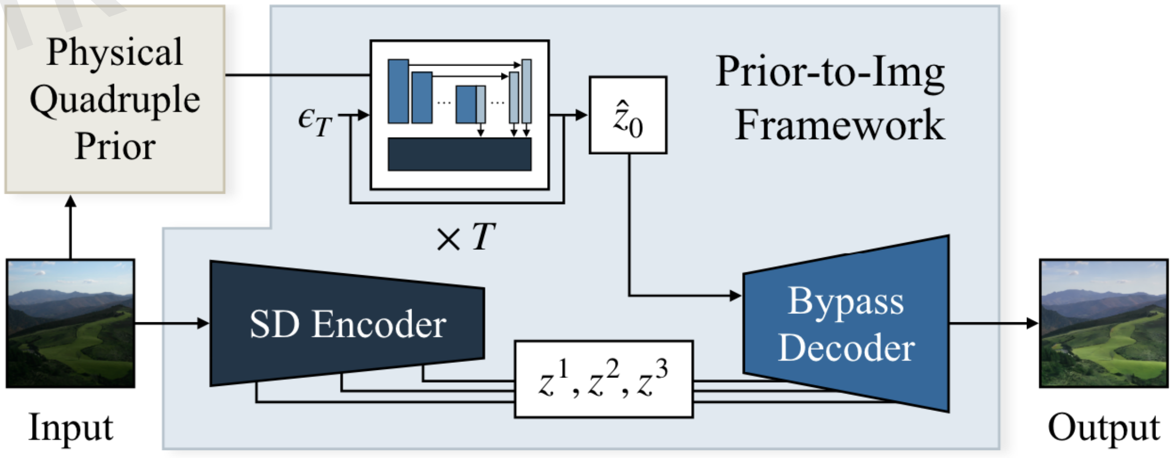
$$\mathcal{L}_{\text{DIFF}} = \boxed{\mathcal{L}_{z_0}} + \boxed{\mathcal{L}_{\text{noise}}}$$

Diffusion for Low-Light Enhancement

- Improvements to make it more suitable for image restoration task
  2. Bypass decoder
     - Handle detail distortion through a new decoder



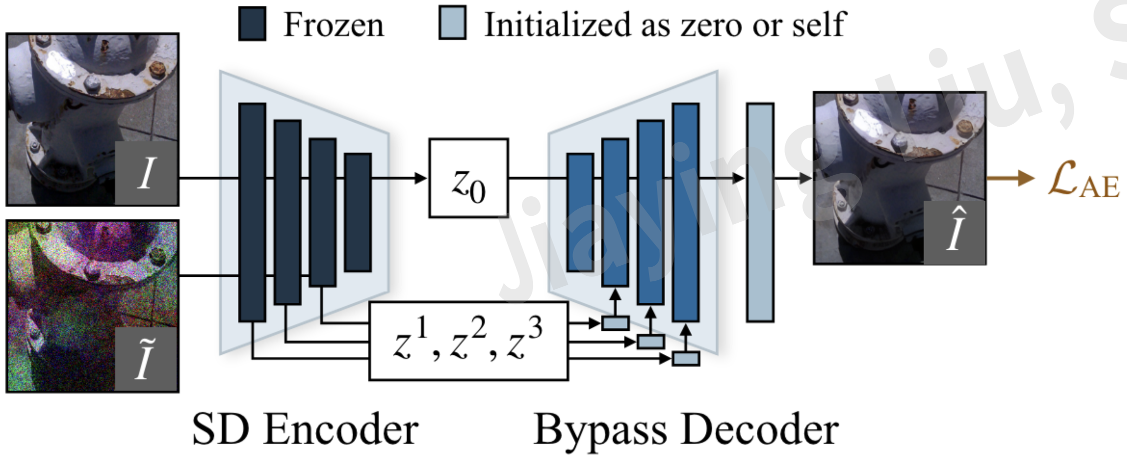
Training



Inference

Diffusion for Low-Light Enhancement

- Improvements to make it more suitable for image restoration task
  2. Bypass decoder
     - Handle detail distortion through a new decoder

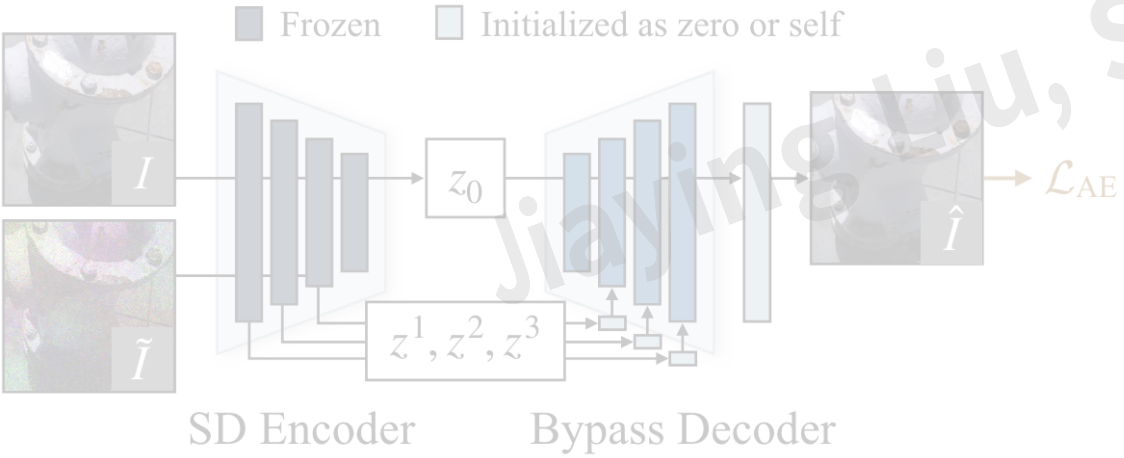


Training

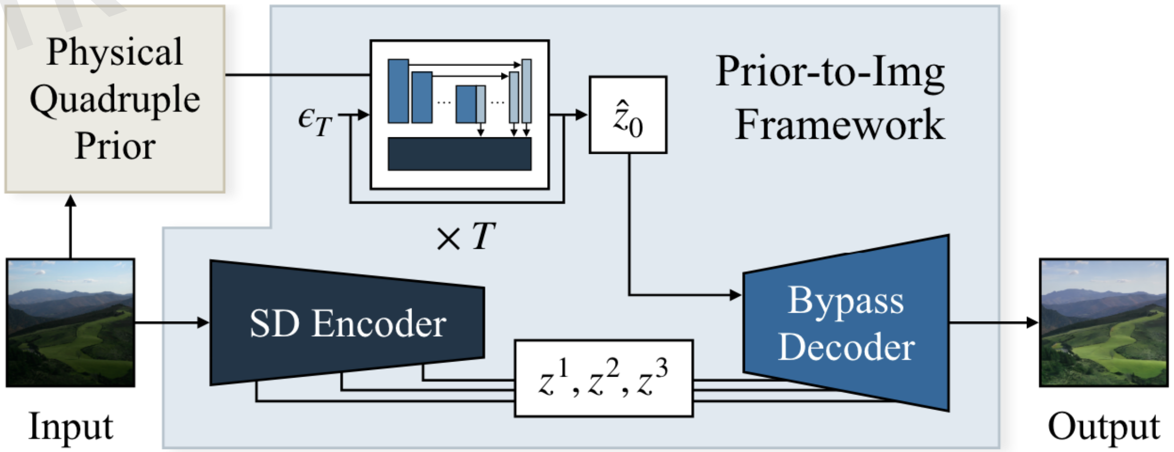
Inference

Diffusion for Low-Light Enhancement

- Improvements to make it more suitable for image restoration task
  2. Bypass decoder
     - Handle detail distortion through a new decoder



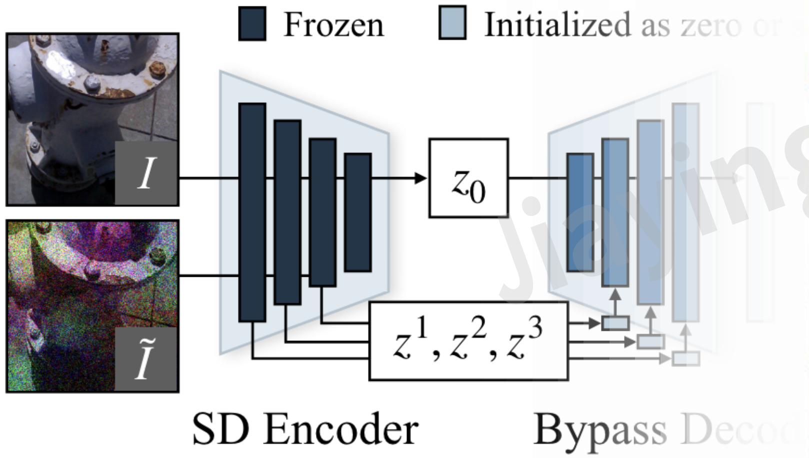
Training



Inference

Diffusion for Low-Light Enhancement

- Improvements to make it more suitable for image restoration task
  - Bypass decoder
    - Handle detail distortion through a new decoder

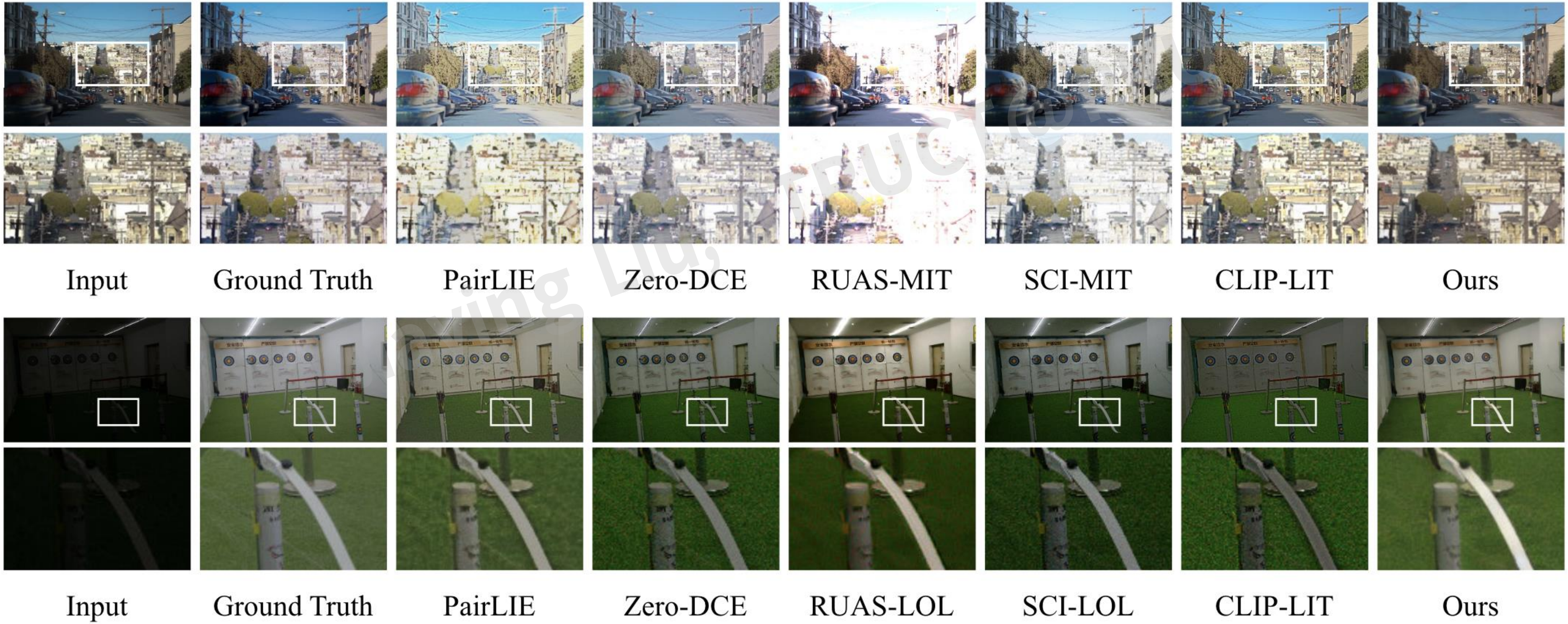


Training

(a) I (b) SD Decoded (c) $\tilde{I}$ (d) Ours

Diffusion for Low-Light Enhancement

Experiments



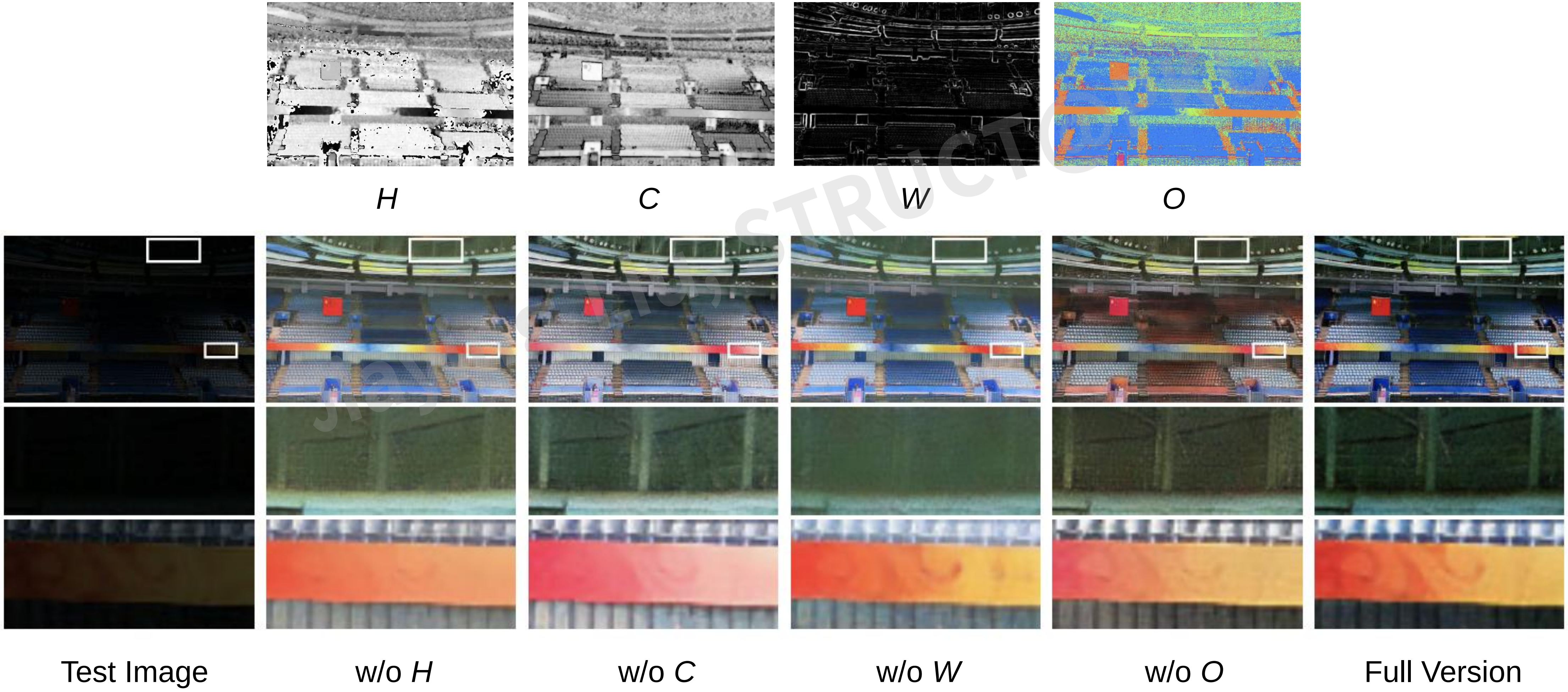
Diffusion for Low-Light Enhancement

Experiments

Datasets		Train Set	LOL [48, 55]				MIT-Adobe FiveK [2]				Unpaired Sets	
Metrics			PSNR↑	SSIM↑	LPIPS↓	LOE↓	PSNR↑	SSIM↑	LPIPS↓	LOE↓	BRISQUE↓	NL↓
Supervised	Retinex-Net [48]	LOL	16.19	0.403	0.534	0.346	12.30	0.687	0.258	0.244	27.10	3.254
	KinD [63]	LOL	20.21	0.814	0.147	0.245	14.71	0.756	0.176	0.174	26.89	0.700
	KinD++ [64]	LOL	16.64	0.662	0.410	0.288	15.76	0.650	0.319	0.176	26.16	0.431
	URetinex-Net [49]	LOL	20.93	0.854	0.104	0.245	14.10	0.734	0.182	0.187	23.80	1.319
	Retinexformer [3]	LOL	28.48	0.877	0.117	0.256	13.87	0.692	0.222	0.224	14.77	1.064
	Retinexformer [3]	MIT	13.02	0.426	0.365	0.280	24.93	0.907	0.063	0.162	24.13	0.684
	DiffLL [18]	LOL+	28.54	0.870	0.102	0.253	15.81	0.719	0.244	0.213	14.96	0.888
Unsupervised	ExCNet [61]	test images	16.29	0.455	0.380	0.295	14.21	0.719	0.197	0.197	19.03	1.563
	EnlightenGAN [19]	own data	18.57	0.700	0.302	0.291	13.28	0.738	0.203	0.199	20.65	0.779
	PairLIE [7]	LOL+	19.70	0.774	0.235	0.278	10.55	0.642	0.273	0.225	29.84	1.471
	NeRCo [54]	LSRW [12]	19.67	0.720	0.266	0.310	17.33	0.767	0.208	0.213	22.81	0.603
	CLIP-LIT [27]	own data	14.82	0.524	0.371	0.320	17.00	0.781	0.159	0.194	23.44	1.962
	ZeroDCE [10]	own data	17.64	0.572	0.316	0.296	13.53	0.725	0.201	0.191	21.76	1.569
	ZeroDCE++ [23]	own data	17.03	0.445	0.314	0.391	12.33	0.408	0.280	0.417	19.34	1.150
	RUAS [30]	MIT	13.62	0.462	0.346	0.292	9.53	0.610	0.301	0.272	29.91	2.091
	RUAS [30]	LOL	15.47	0.490	0.305	0.330	5.15	0.373	0.669	0.399	44.70	3.312
	RUAS [30]	FACE [56]	15.05	0.456	0.371	0.292	5.00	0.366	0.685	0.398	46.21	3.633
	SCI [34]	MIT	11.67	0.395	0.361	0.286	16.29	0.795	0.143	0.165	16.73	0.853
	SCI [34]	LOL+	16.97	0.532	0.312	0.289	7.83	0.573	0.360	0.187	24.46	1.893
	SCI [34]	FACE [56]	16.80	0.543	0.322	0.297	10.95	0.684	0.272	0.205	18.33	1.335
	Ours	COCO [28]	20.31	0.808	0.202	0.281	18.51	0.785	0.163	0.188	14.64	0.423

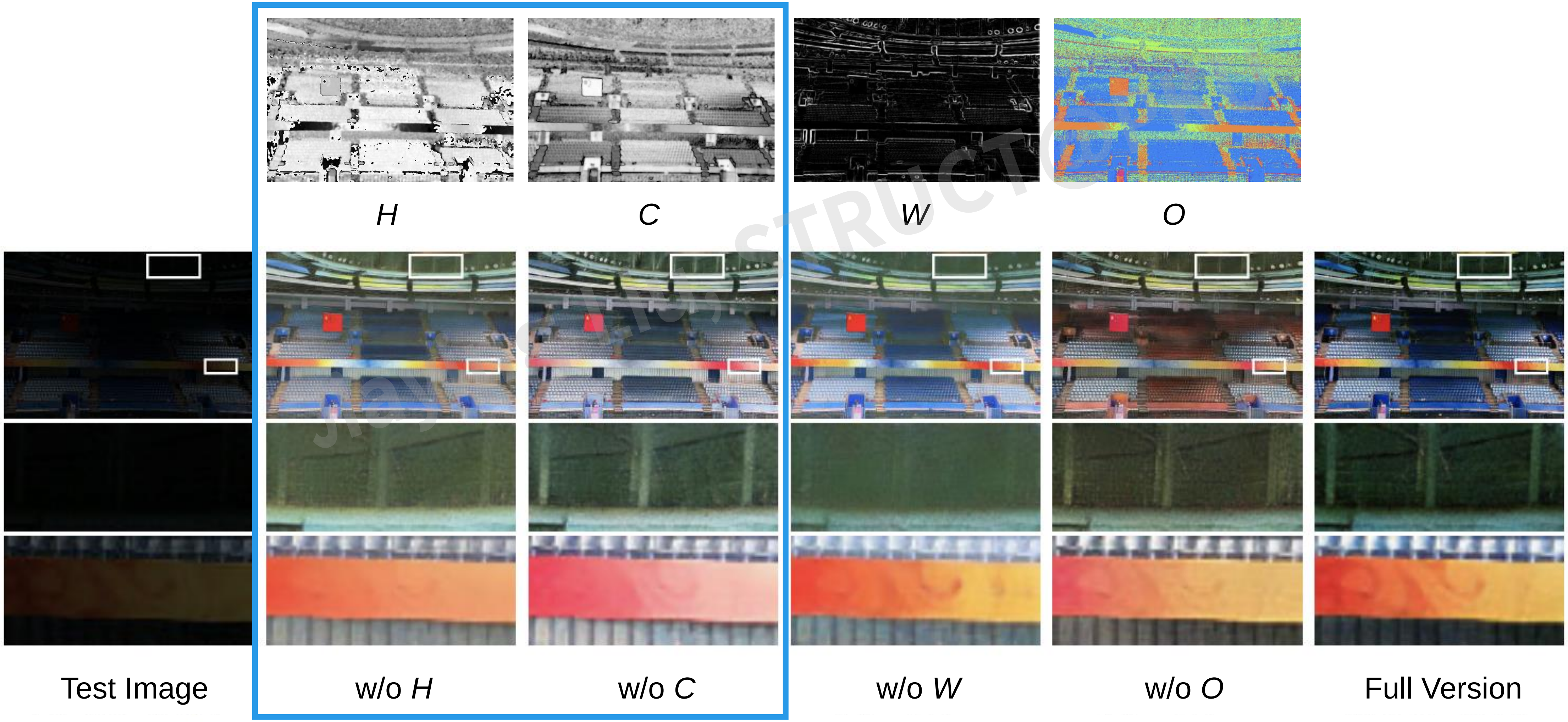
Diffusion for Low-Light Enhancement

- Ablation study on physical quadruple prior



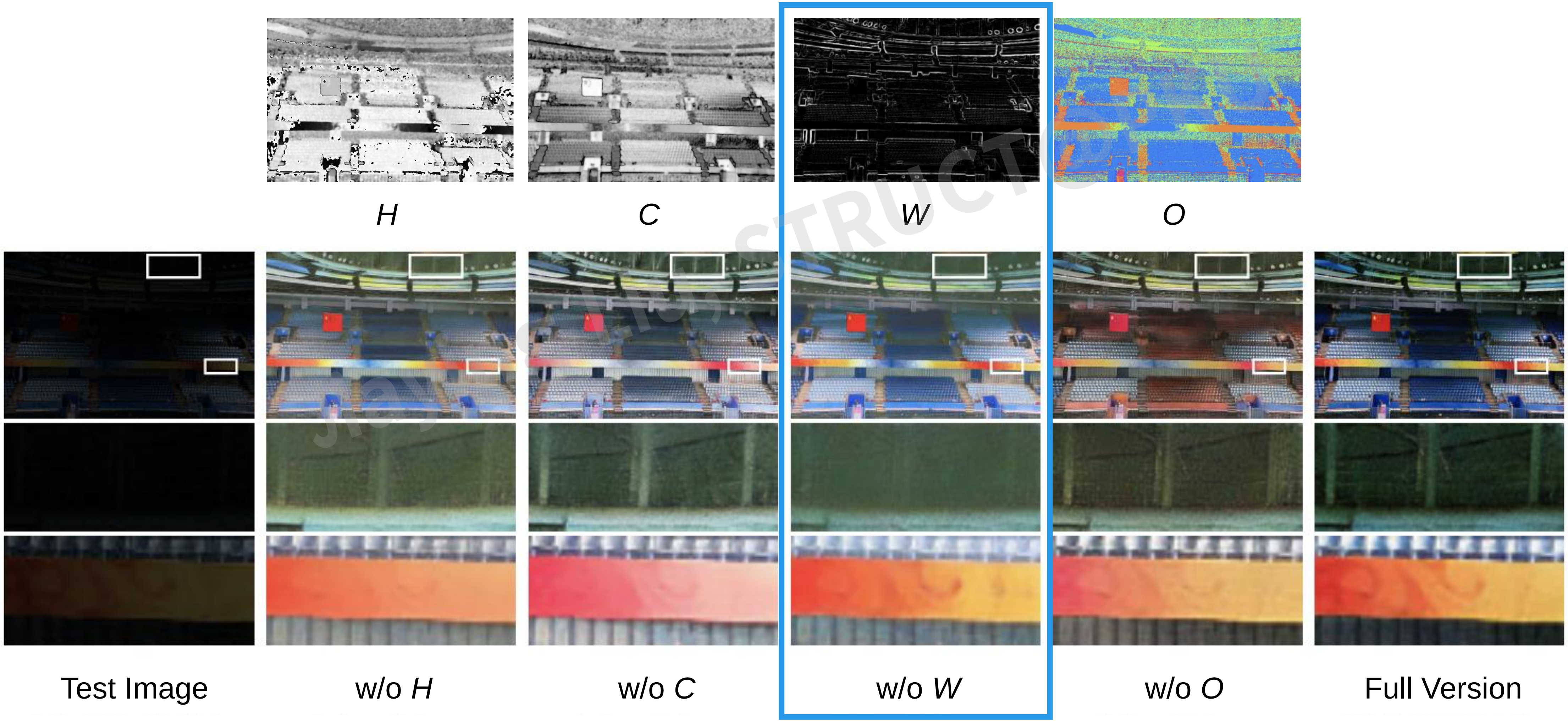
Diffusion for Low-Light Enhancement

- Ablation study on physical quadruple prior



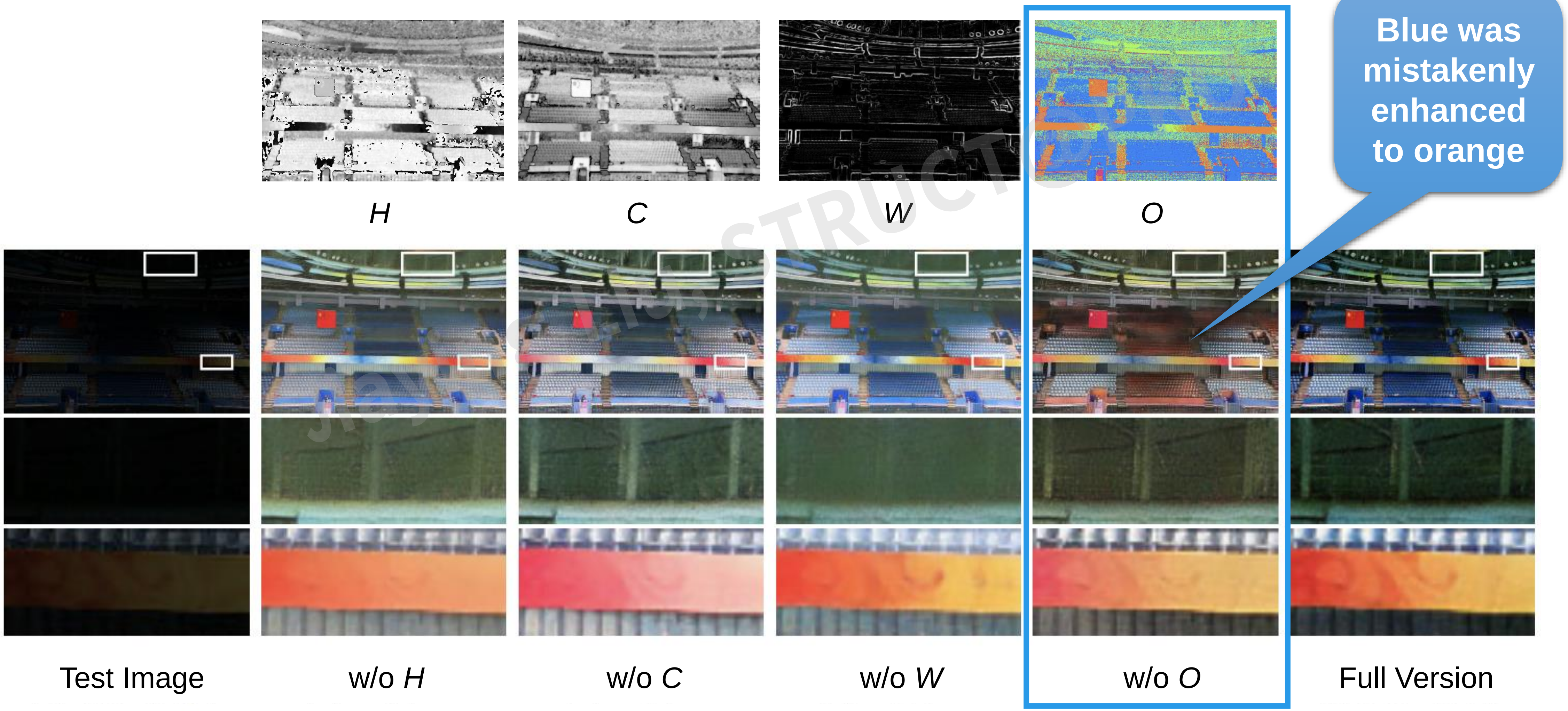
Diffusion for Low-Light Enhancement

- Ablation study on physical quadruple prior

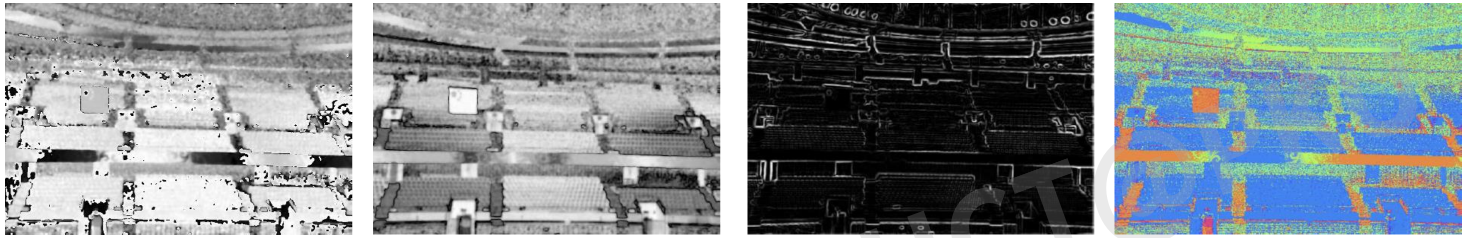


Diffusion for Low-Light Enhancement

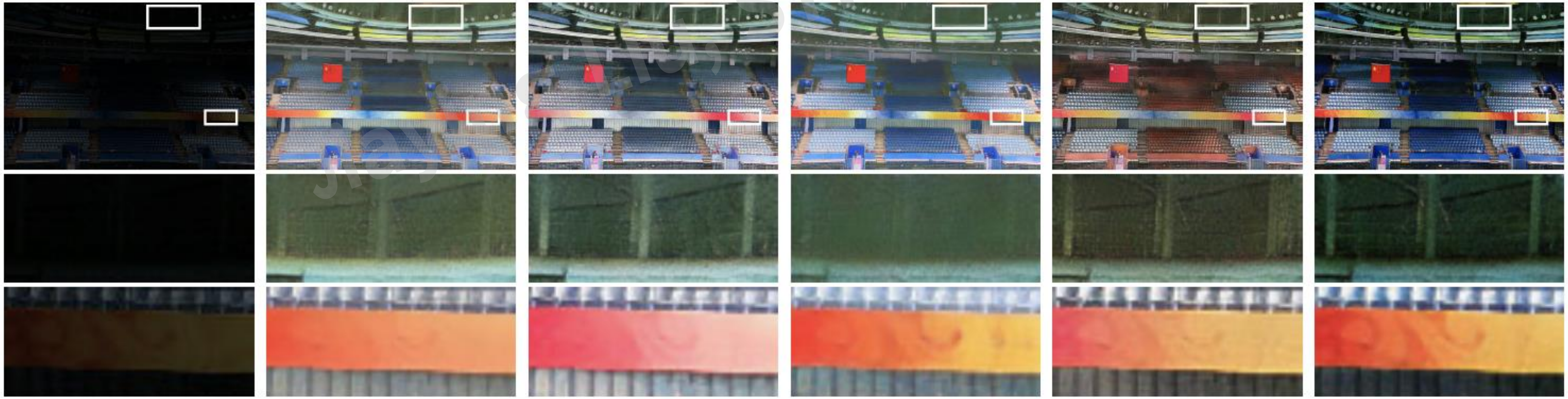
- Ablation study on physical quadruple prior



■ Ablation study on physical quadruple prior



H *C* *W* *O*



Test Image | w/o *H* | w/o *C* | w/o *W* | w/o *O* | Full Version

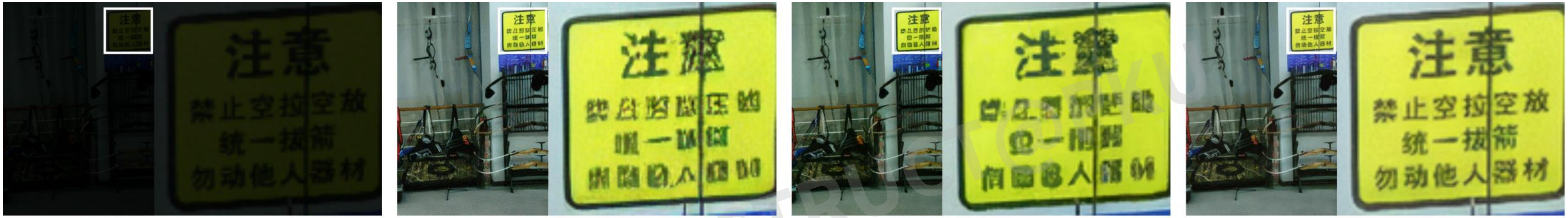
Diffusion for Low-Light Enhancement

- Ablation study on prior-to-image mapping

Test Image | SR3 as prior-to-image | Our prior-to-image

Diffusion for Low-Light Enhancement

- Effects of different decoders



Test Image | Stable Diffusion Decoder | Consistency Decoder | Our Bypass Decoder

- The proposed decoder can also be applied on the colorization task

Test Image | Stable Diffusion Decoder | Consistency Decoder | Our Bypass Decoder

- Summary of the method

We propose a **zero-reference** low-light enhancement framework trainable solely with normal light images.

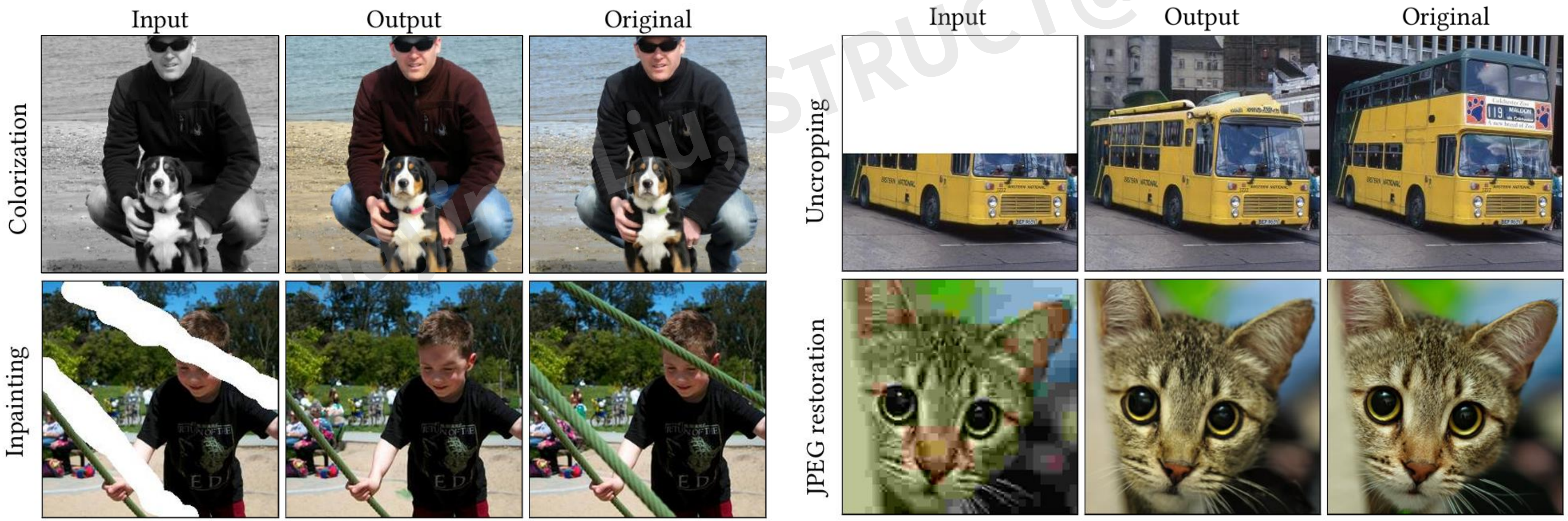
  - An **illumination-invariant prior** inspired by physical light transfer.
  - A **prior-to-image** framework trained without low-light data
    - Use the prior as a condition to control a pretrained large-scale generative diffusion model
    - Bypass decoder to address the distortion issue

03

Unified Image Restoration and Enhancement

Unified Image Restoration and Enhancement

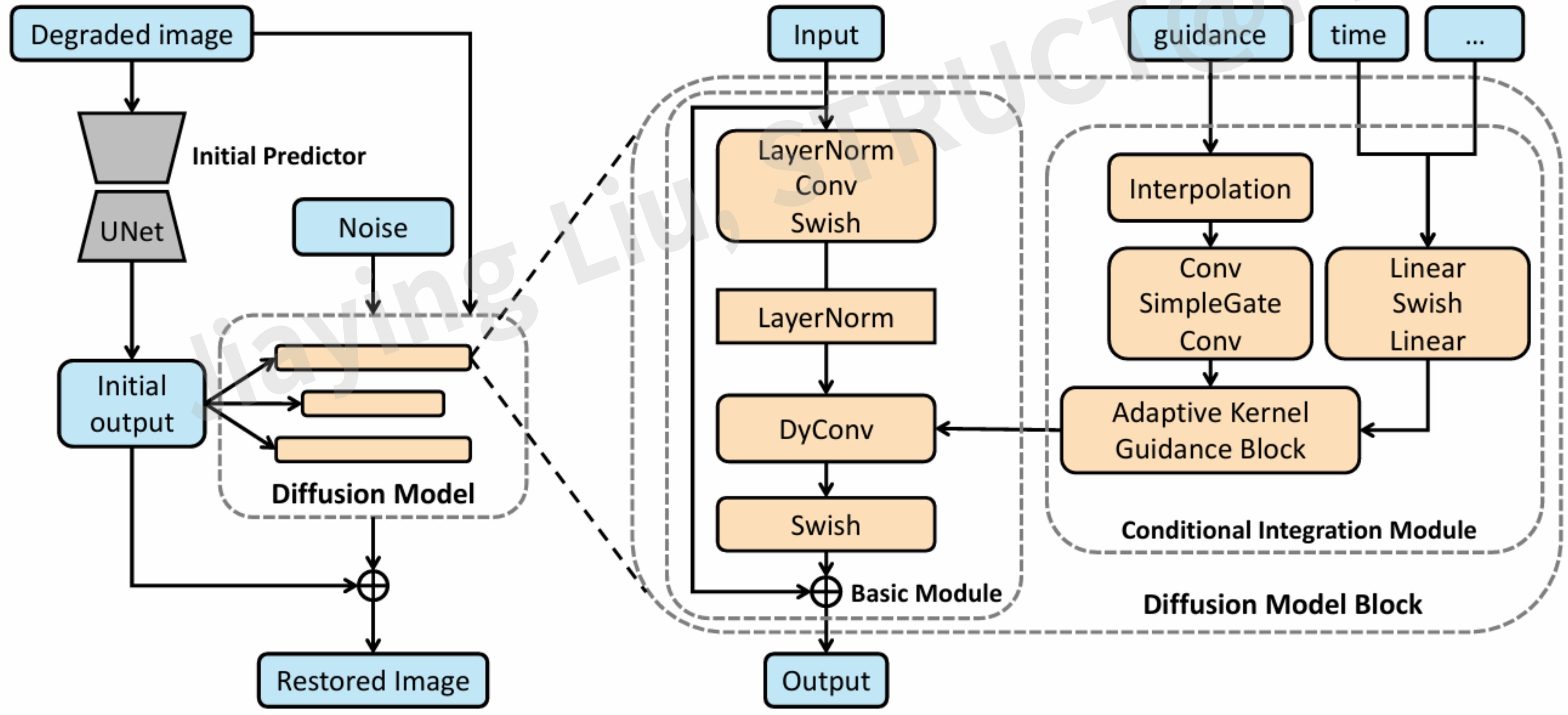
- Palette: Image-to-image diffusion models
  - Based on the SR3 structure: U-Net architecture
  - Different type of conditional images



[23] Chitwan Saharia, et al. Palette: Image-to-image diffusion models. ACM SIGGRAPH, 2022.

Unified Image Restoration and Enhancement

- A Unified Conditional Framework for Diffusion-based
  - Same model structure and training hyper-parameters
  - Take different conditional guidance and train on different datasets



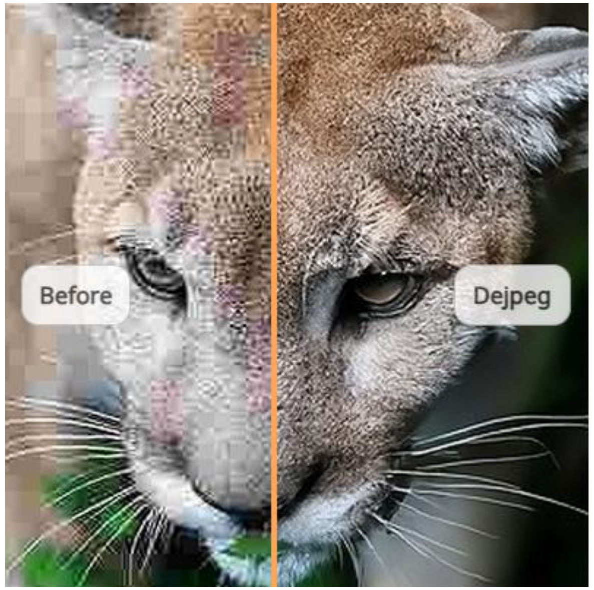
[24] Yi Zhang, et al. A unified conditional framework for diffusion-based image restoration. NIPS, 2023

Unified Image Restoration and Enhancement

- A Unified Conditional Framework for Diffusion-based
  - Same model structure and training hyper-parameters
  - Take different conditional guidance and train on different datasets

Low-light denoise

Deblur

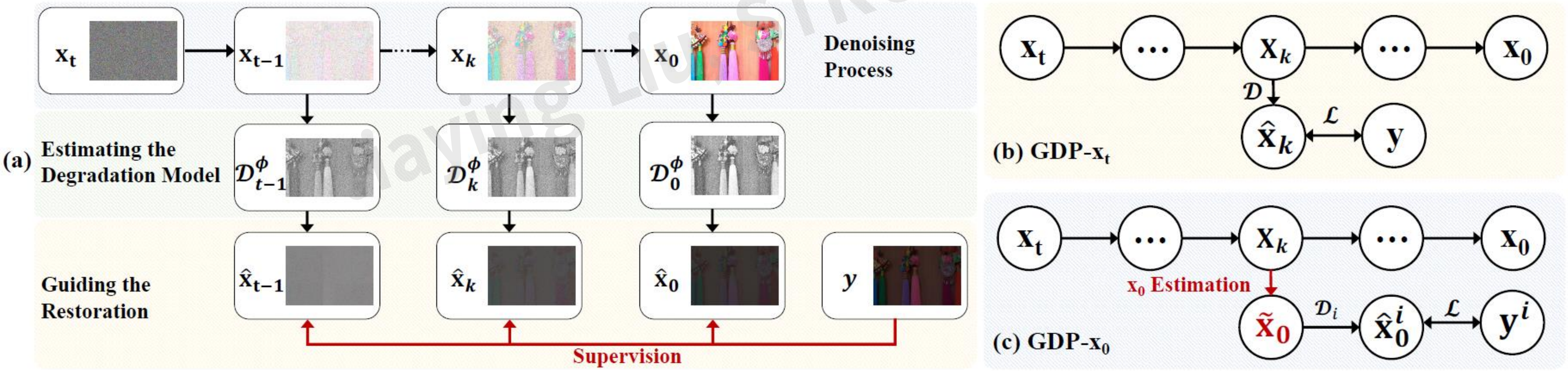


JPEG restoration

[24] Yi Zhang, et al. A unified conditional framework for diffusion-based image restoration. NIPS, 2023

Unified Image Restoration and Enhancement

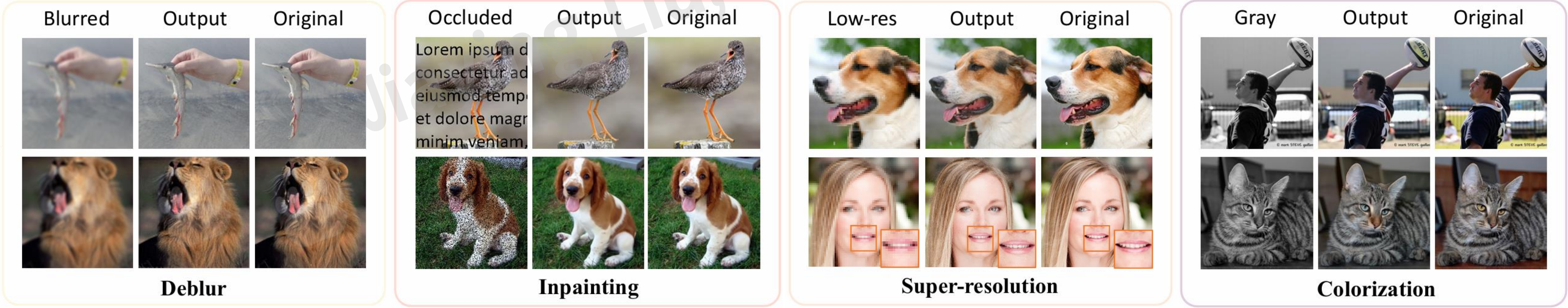
- Generative Diffusion Prior (GDP)
  - A unified framework for multiple restoration and enhancement tasks.
  - Use a pretrained unconditional image synthesis diffusion model as prior.



[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

Unified Image Restoration and Enhancement

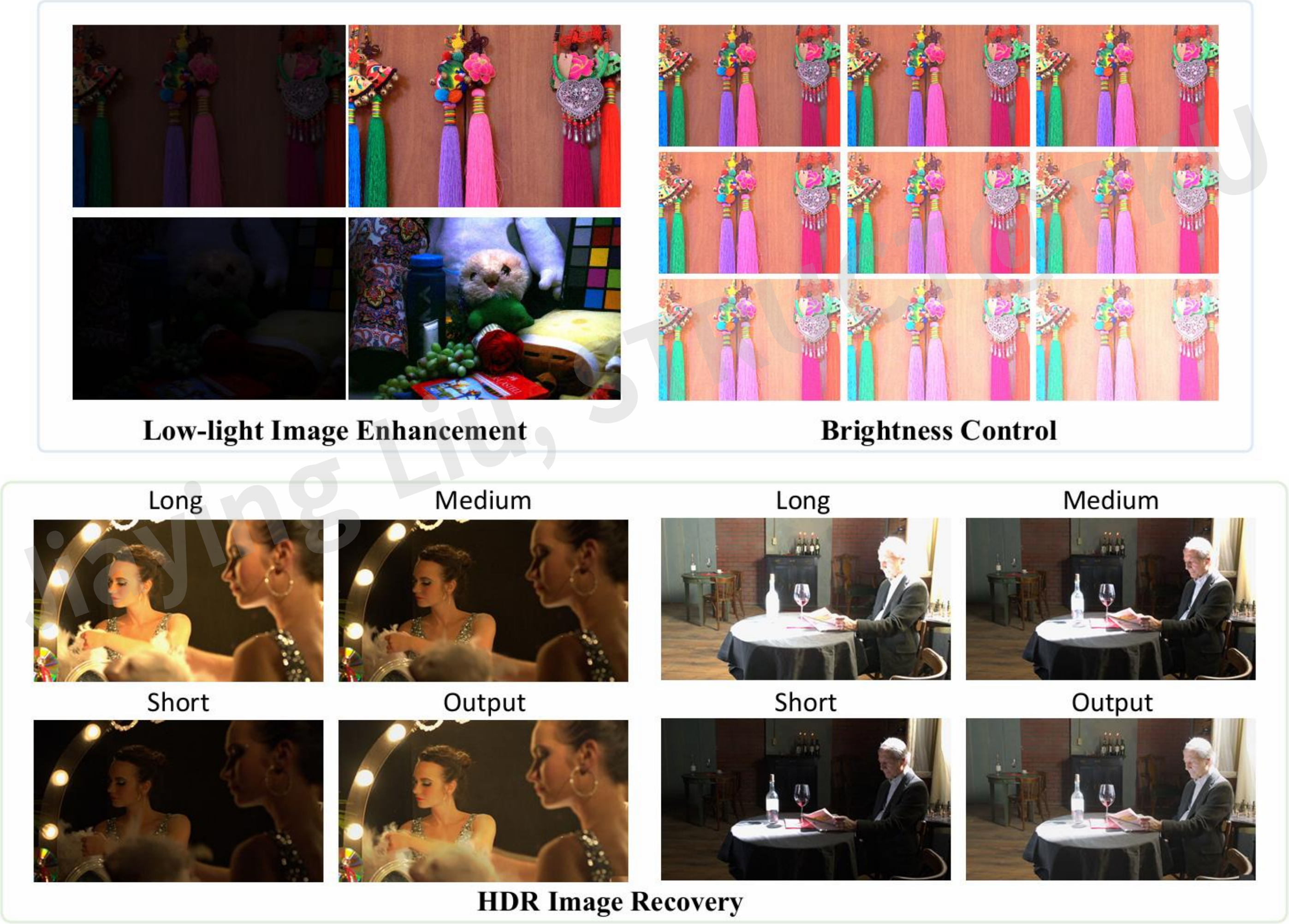
- Generative Diffusion Prior (GDP)
  - A unified framework for multiple restoration and enhancement tasks.
  - Use a pretrained unconditional image synthesis diffusion model as prior.
  - Different degradation models learned during the sampling process.



[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

Unified Image Restoration and Enhancement

- Generative Diffusion Prior (GDP)



[7] Ben Fei, et al. Generative Diffusion Prior for Unified Image Restoration and Enhancement, CVPR, 2023

liujiaying@pku.edu.cn

ISCAS 2024

STRUCT Group

Intelligent image computing

Wangxuan Institute of Computer Technology, Peking University

Spatial and Temporal **Restoration**,
Understanding and **Compression** Team

PI: **Jiaying Liu**

- **Email:** liujiaying@pku.edu.cn
- **Website:** http://www.wict.pku.edu.cn/struct/