THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

CVPR Highlight

Learning Diffusion Texture Priors for Image Restoration

Tian Ye[1], Sixiang Chen[1], Wenhao Chai[2], Zhaohu Xing[1], Jing Qin[3], Ge Lin[1], Lei Zhu[1,4]

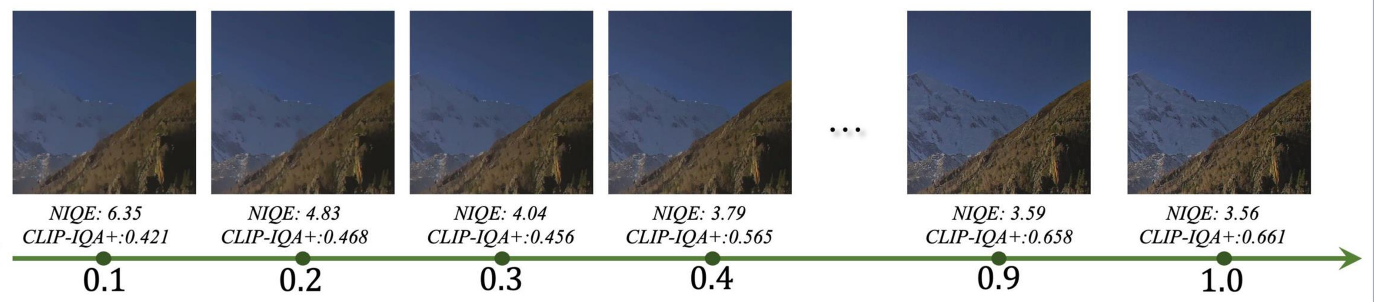
[1]The Hong Kong University of Science and Technology (Guangzhou)

[2]University of Washington [3]The Hong Kong Polytechnic University [4]The Hong Kong University of Science and Technology

CVPR SEATTLE, WA JUNE 17-21, 2024

Motivation&Chanllenge&Rethinking

Motivations



The quality of the image's texture details have a significant impact on the visual perceptual quality of the image.

We use a coefficient α ranging from 0.1 to 1 to control the intensity of texture removal. As α increases, leading to richer textures from left to right in the images, we can observe a corresponding improvement in the scores of no-reference image quality metrics.

Challenges

(a)Directly applying diffusion image generation techniques to image restoration tasks is often impractical. The high content fidelity required in image restoration clashes with the stochastic nature of diffusion models.

(b)While several previous studies have successfully adopted physical models with neural networks or crafting hand-designed priors to mitigate the inherent randomness of the diffusion paradigm, they lack versatility and show limited generalization capability on real-world scenes.

(c) The iterative nature and complexity of the denoising process in diffusion models neces- sitate extensive data and lengthy training cycles for effective learning.

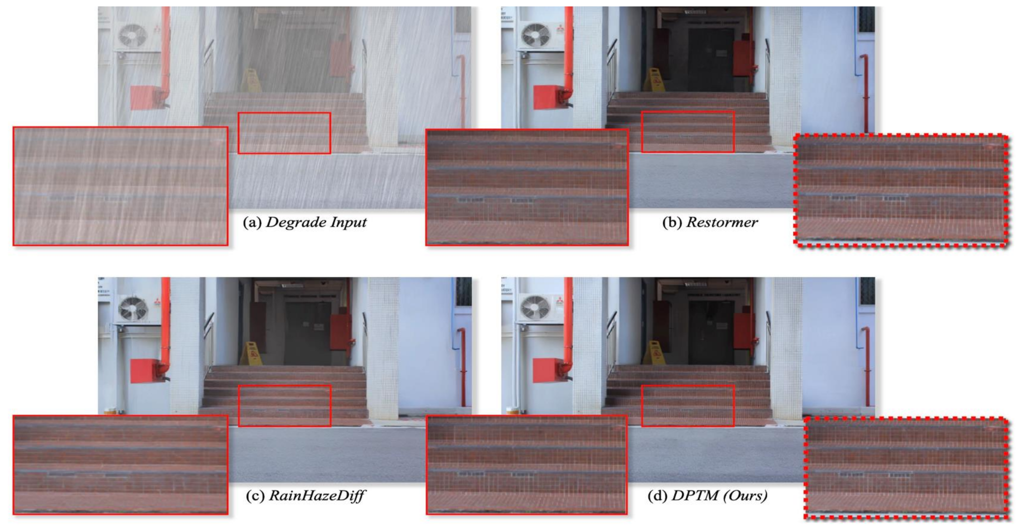
Rethinking Diffusion Model for image restoration

(i) To preserve high-level fidelity in restoration, we propose using a diffusion model to recover only texture layers. This approach emphasizes and highlights the importance of fine textures in visual perception (See the Figure). By focusing diffusion on texture recovery, we minimize randomness of diffusion process and leverage its strength in generating realistic details.

(ii) Recognizing the diffusion paradigm as a potent representation learning tool that re- quires extensive data for convergence, we pre-train our model, named the Diffusion Texture Priors Model (DTPM), on a large-scale dataset of high-quality images.

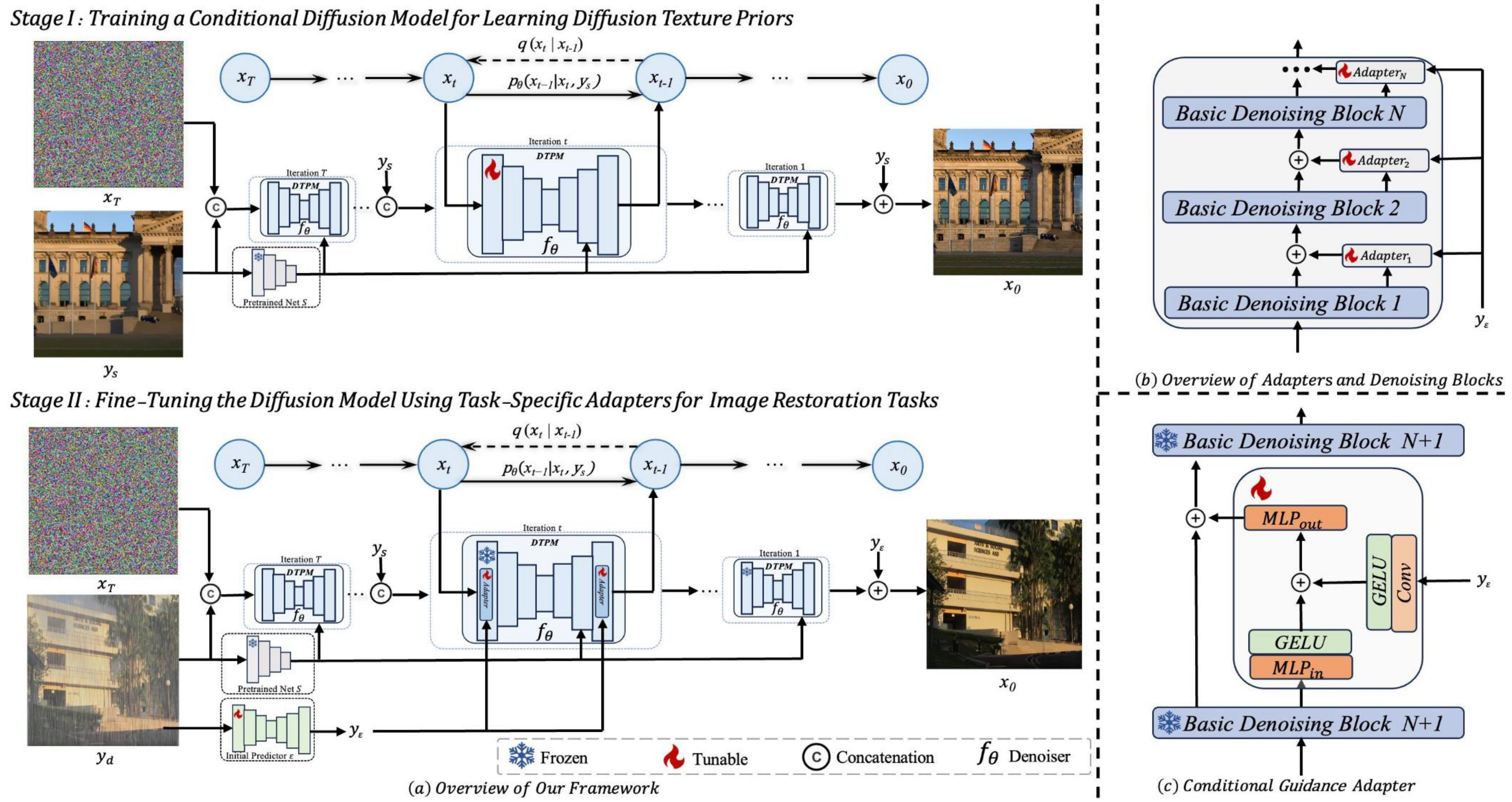
Our Aim

Our aim is to create a unified framework for image restoration tasks based on our Diffusion Texture Priors. This framework comprises two stages: In Stage 1, we train a conditional diffusion model with semantic latent feature constraints using a large dataset of high-resolution, high-quality natural images. During Stage 2, we incorporate conditional guidance adapters into the conditional diffusion model to facilitate the model's adaptation to downstream image restoration tasks.



It's worth noting that our approach excels in generating finer details, thanks to the incorporation of Diffusion Texture Priors we proposed.

Our Framework



In the stage I, under the guiding constraints of semantic code, the diffusion model learns texture layers through residual learning from a large amount of high-quality data, which allows us to encapsulate diverse and rich texture knowledge into the diffusion model. In the stage II, we fix most of the parameters of the trained diffusion model and insert Conditional Guidance Adapters between each layer for efficient fine-tuning and conditional guidance on image restoration tasks.

Results