Shadow Removal via Diffusion Model

CSE 527: Introduction to Computer Vision

Neelesh Verma (115233265)
Bhavesh Kumar Vasnani (115119433)

Motivation

- Diffusion models have showcased the benefits over traditional Generative Adversarial Networks (GANs) for several image generation scenarios.

- Shadow removal is however different from regular generation, we need to take care of
  - Information in the shadow region
  - Influencing diffusion appropriately with information retrieved

Milestones

- Validated results of RePaint on Places2 and CelebA-HQ
- Validated results of ILVR for faces
- Trained Guided Diffusion model on augmented ISTD Dataset
- Validated ILVR for Shadow Removal
- Extended RePaint for Shadow Removal task
- Experimented merits of DDIM

ILVR Architecture

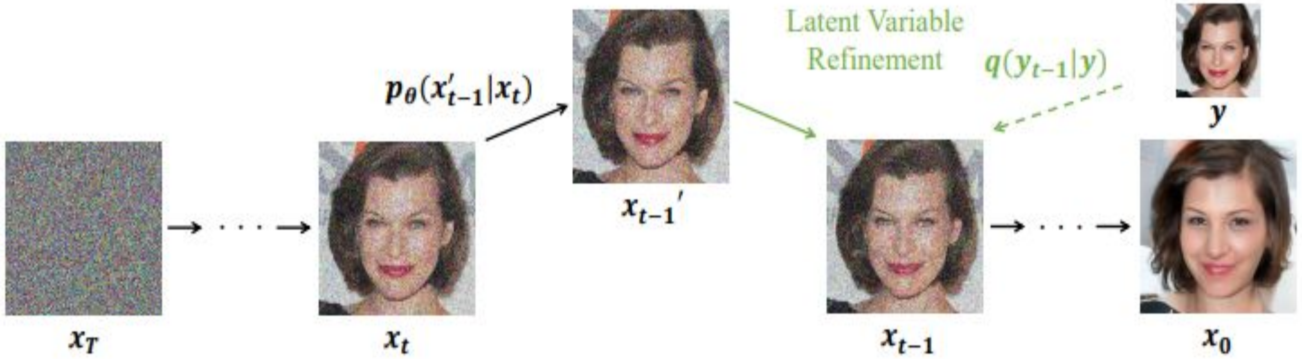


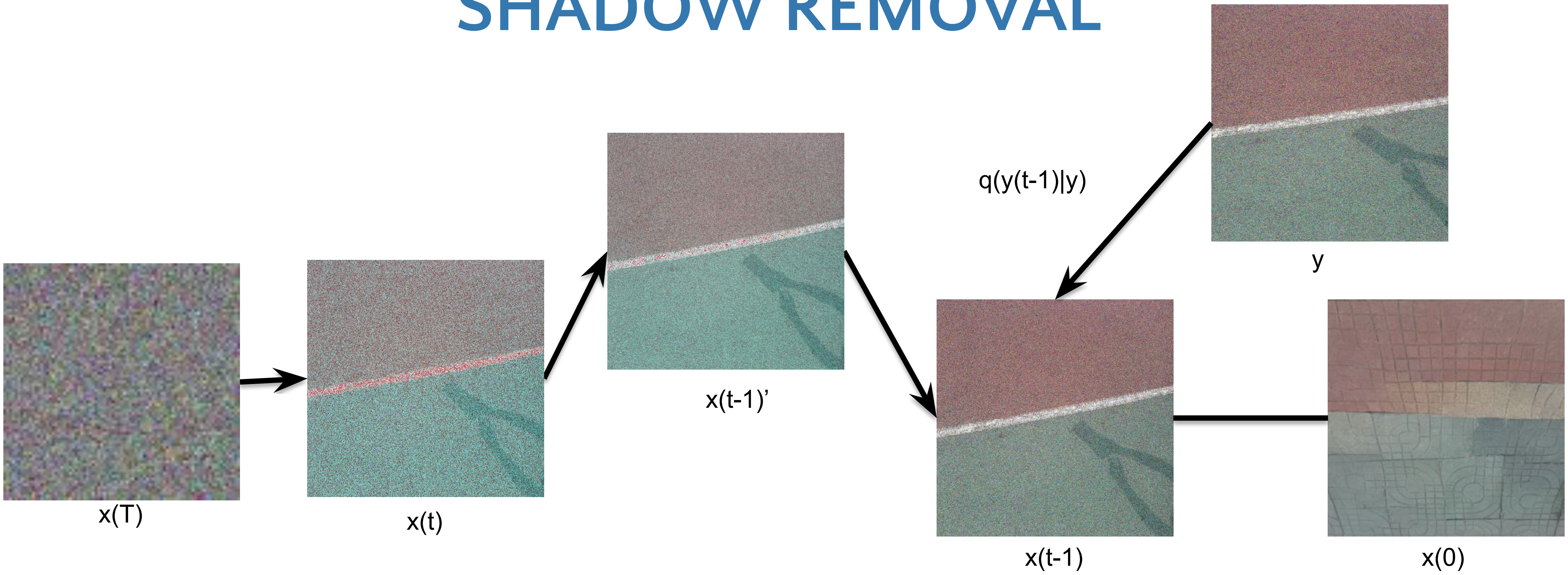
Figure 2: **Graphical model of Iterative Latent Variable Refinement.** From state x_t, we first sample unconditional proposal x_{t-1} according to Eq. 5. Then, we match latent variable with encoded condition y_{t-1} according to Eq. 8.

Algorithm 1 Iterative Latent Variable Refinement

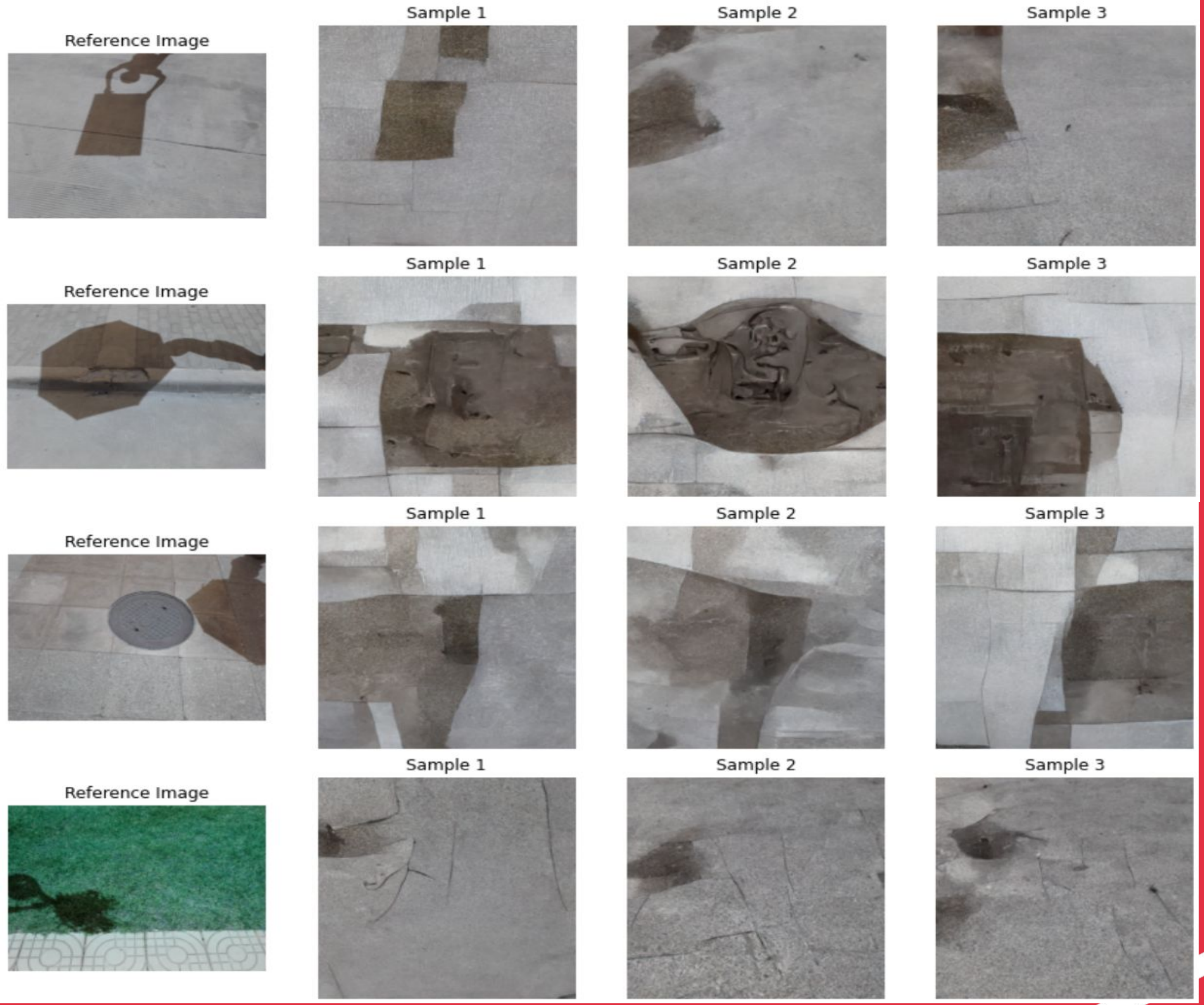
Input: Reference image y
Output: Generated image x
$\phi_N(\cdot)$: low-pass filter with scale N
Sample $x_T \sim N(\mathbf{0}, \mathbf{I})$
for $t = T, ..., 1$ **do**
    $\mathbf{z} \sim N(\mathbf{0}, \mathbf{I})$
    $x'_{t-1} \sim p_\theta(x'_{t-1}|x_t)$ ▷ unconditional proposal
    $y_{t-1} \sim q(y_{t-1}|y)$ ▷ condition encoding
    $x_{t-1} \leftarrow \phi_N(y_{t-1}) + x'_{t-1} - \phi_N(x'_{t-1})$
end for
return x_0

Jooyoung Choi et al. ILVR: Conditioning Method for Denoising Diffusion Probabilistic Models. 2021. doi: 10.48550/ARXIV.2108.02938. url: https://arxiv.org/abs/2108.02938.

ILVR PIPELINE FOR SHADOW REMOVAL



ILVR Detailed Results



RePaint Architecture

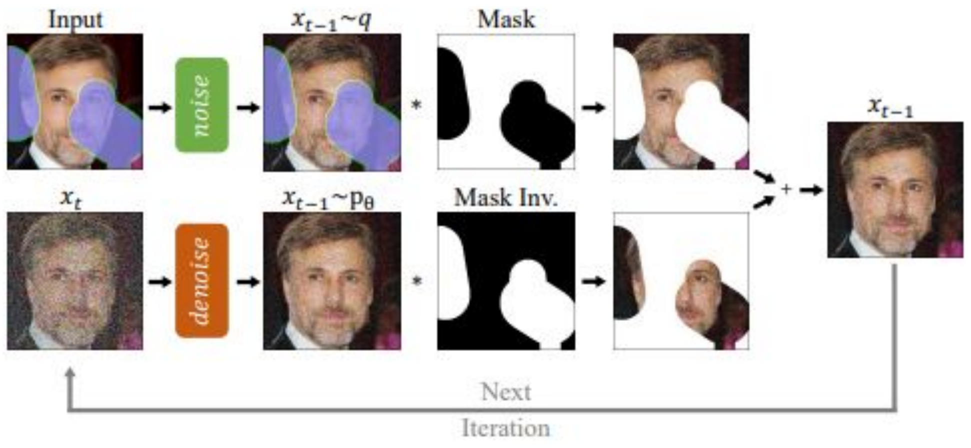


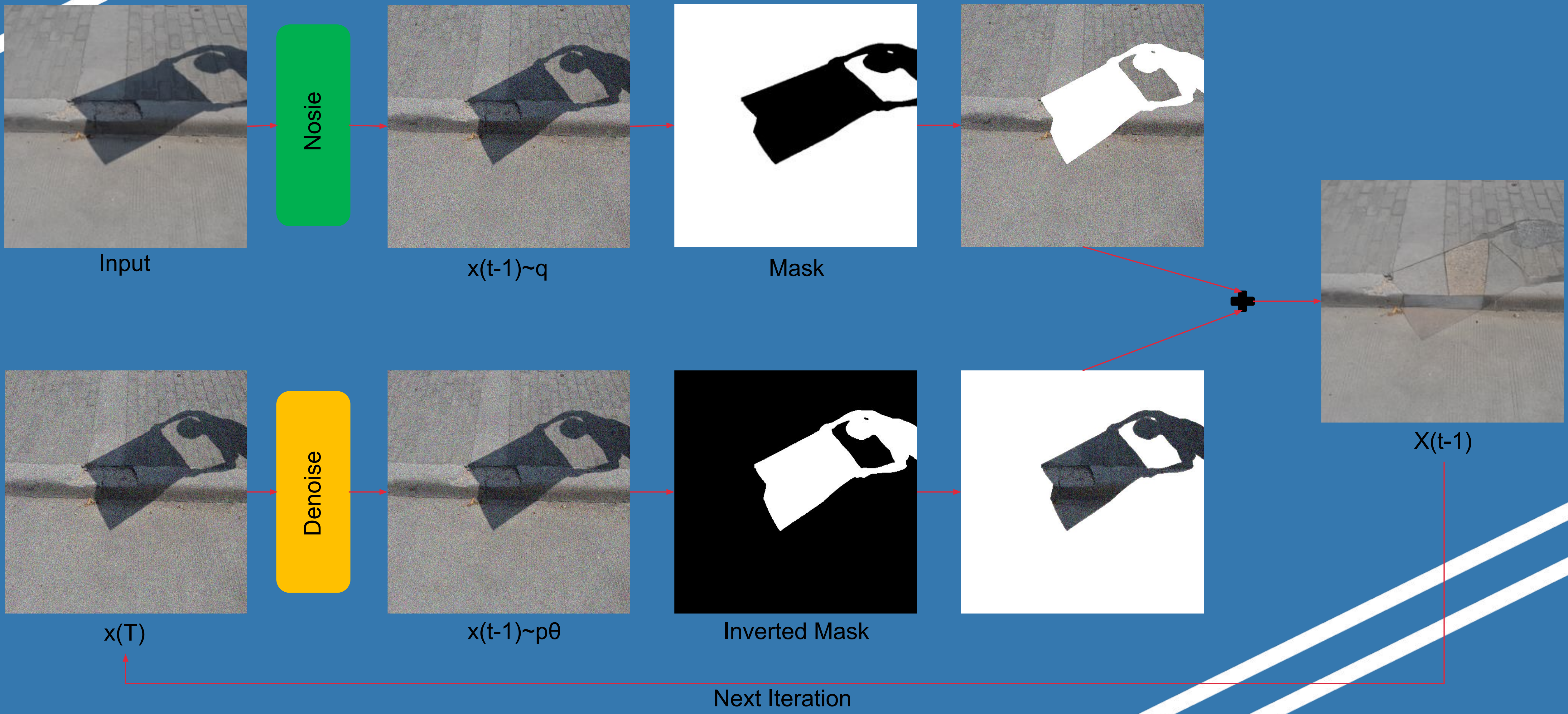
Figure 2. **Overview of our approach.** RePaint modifies the standard denoising process in order to condition on the given image content. In each step, we sample the known region (*top*) from the input and the inpainted part from the DDPM output (*bottom*).

Algorithm 1 Inpainting using our RePaint approach.

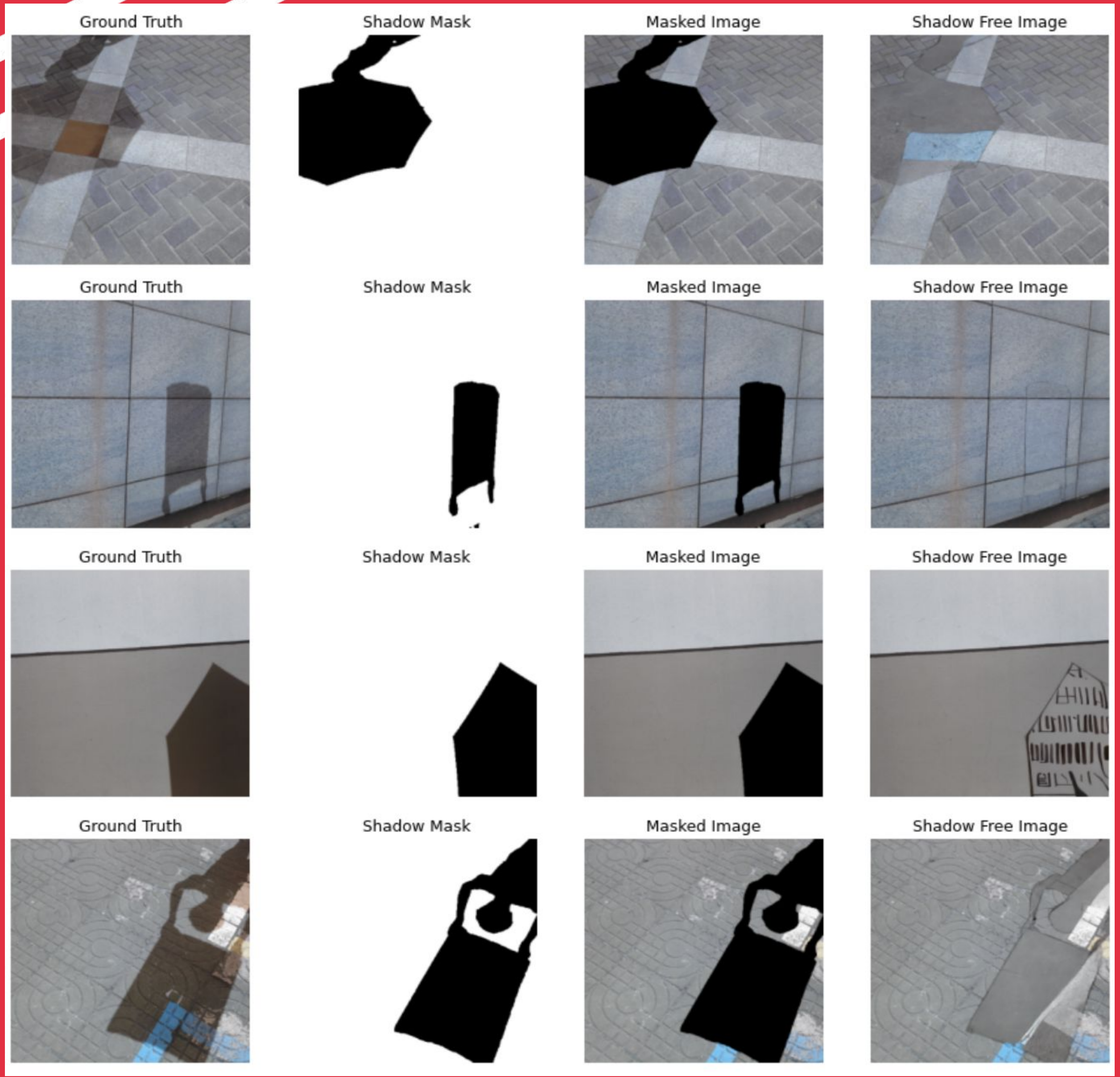
$x_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
for $t = T, \ldots, 1$ **do**
  for $u = 1, \ldots, U$ **do**
    $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\epsilon = \mathbf{0}$
    $x_{t-1}^{\text{known}} = \sqrt{\bar{\alpha}_t} x_0 + (1 - \bar{\alpha}_t)\epsilon$
    $z \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$, else $\mathbf{z} = \mathbf{0}$
    $x_{t-1}^{\text{unknown}} = \frac{1}{\sqrt{\alpha_t}} \left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_\theta(x_t, t) \right) + \sigma_t z$
    $\boxed{x_{t-1} = m \odot x_{t-1}^{\text{known}} + (1 - m) \odot x_{t-1}^{\text{unknown}}}$
    if $u < U$ and $t > 1$ **then**
      $x_t \sim \mathcal{N}(\sqrt{1 - \beta_{t-1}} x_{t-1}, \beta_{t-1}\mathbf{I})$
    end if
  end for
end for
return x_0

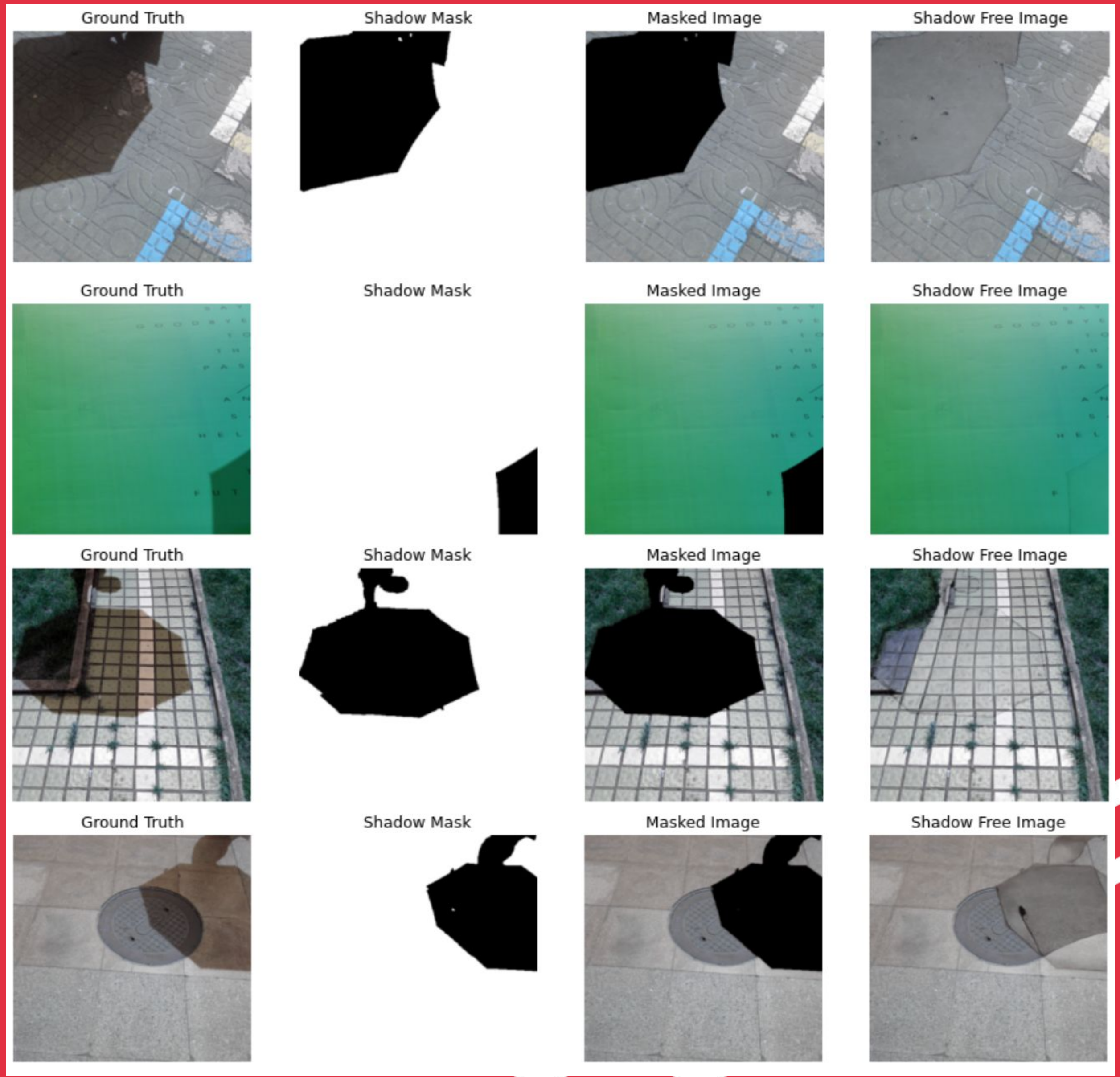
Andreas Lugmayr et al. RePaint: Inpainting using Denoising Diffusion Probabilistic Models. 2022. doi: 10.48550/ARXIV.2201.09865. url: https://arxiv.org/abs/2201.09865.

RePaint Pipeline for Shadow Removal

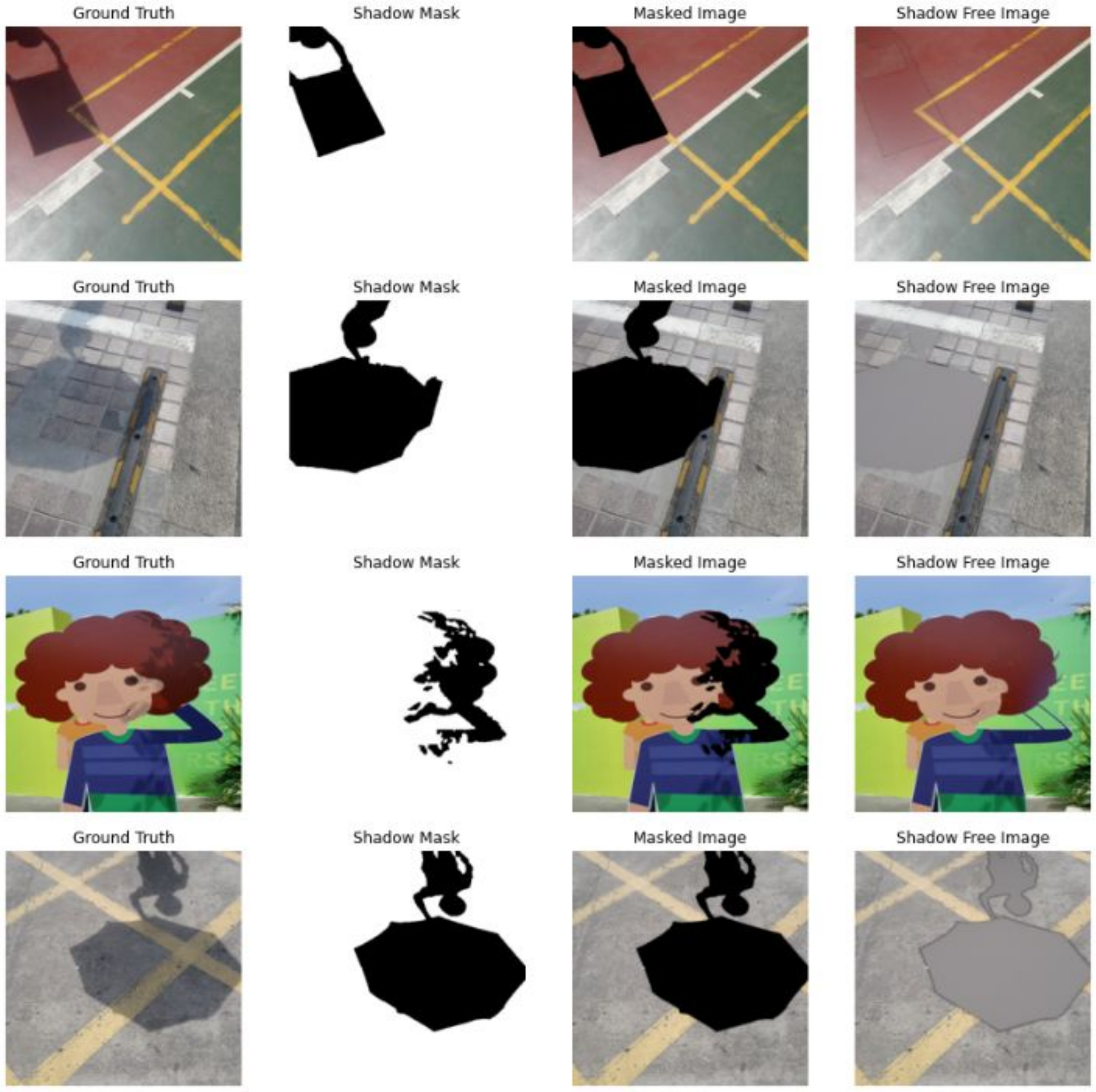


RePaint Results





RePaint (Using Places2 pre-trained model)



Modification to RePaint: Passing shadow information

- Passed information from Ground Truth during inference
- Used weighted decay rate to gradually guide the reverse diffusion
- Results showed improvement and patterns were appearing
- Couldn't successfully reconstruct the entire shadow region

$$x_t = (\text{mask} * (\text{WeightedGT}) + (1 - \text{mask}) * ((1.0 - \text{decay}) * x_{t-1} + (\text{decay} * \text{WeightedGT})))$$

Modification to RePaint: Passing shadow information

Shadow Information

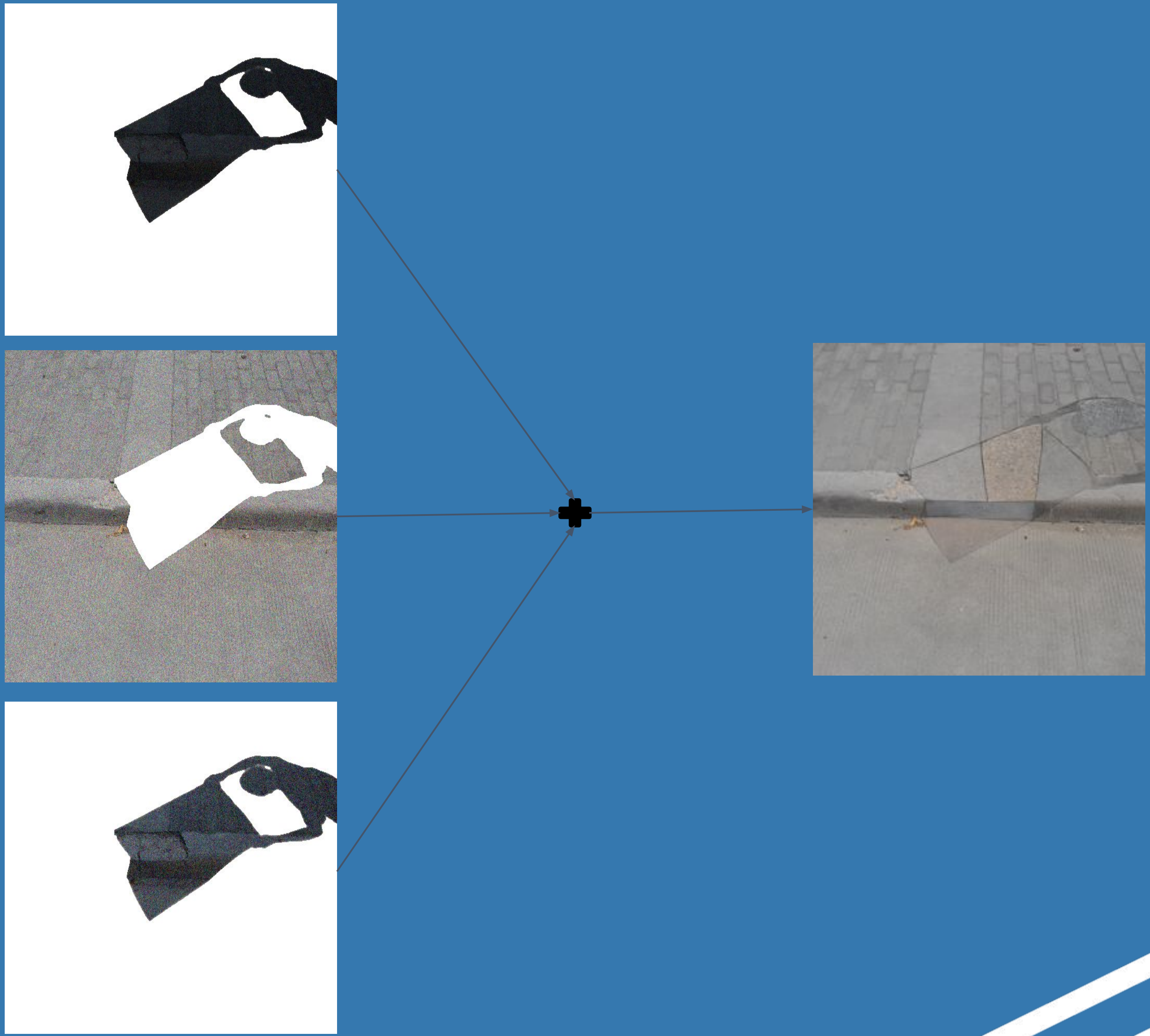
(1 - mask)(weightedGT)

Input

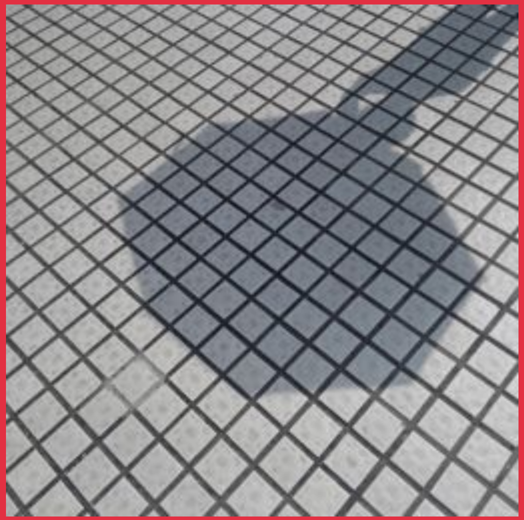
mask * weightedGT

Generated Image

$\mathbf{x_{t-1}}$

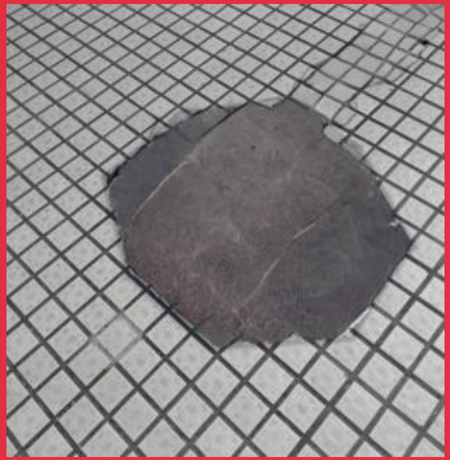


Modified RePaint Results for Varying Decay Rate

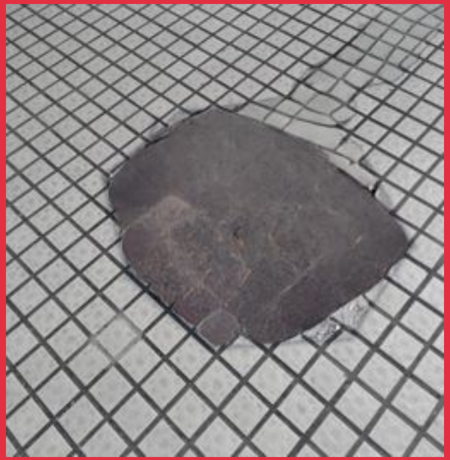


Ground Truth

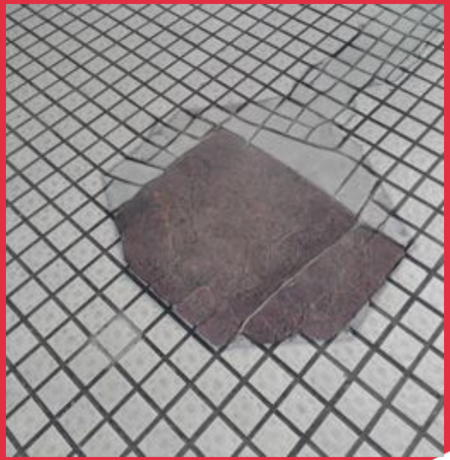
.15



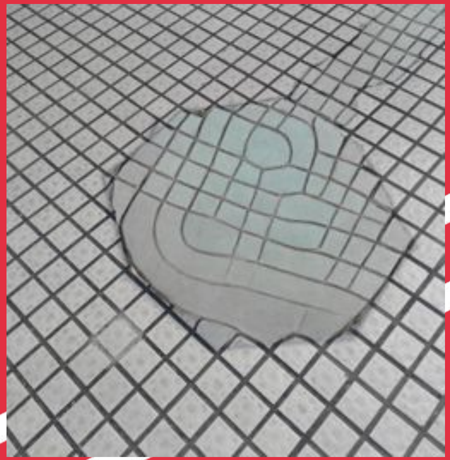
.20



.30

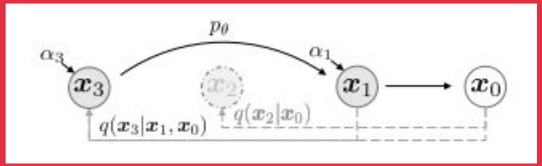


.50



.95

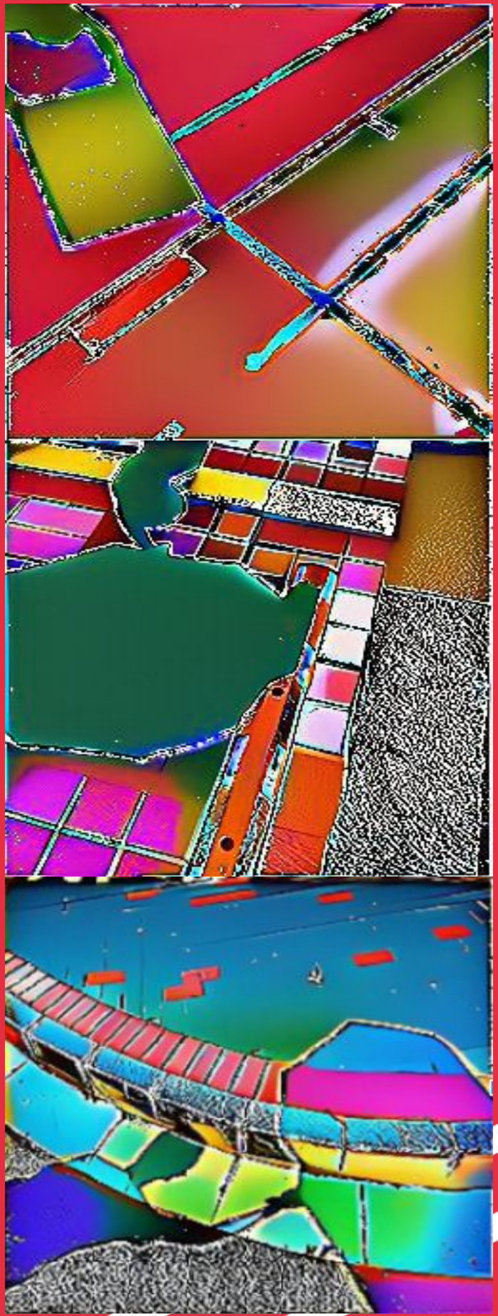
DDIM Results



Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising Diffusion Implicit Models. 2020. doi: 10.48550/ARXIV.2010.02502. url: https://arxiv.org/abs/2010.02502PAPER

Ground Truth

Generated Sample



Learnings and Conclusion

- Generating capability of our model is low

- Current state of diffusion models not suitable for shadow removal
  - Diffusion models require huge dataset to learn
  - Difficult to condition diffusion models with shadow information

- Ideal would be to introduce modification into training phase
  - Use of custom priors or cold diffusion techniques to directly generate from shadow image instead of noisy image

- More shadow guided loss functions (Chromaticity, Pattern) instead of MSE based loss functions used in Diffusion model

- How to handle soft and hard shadows within one model framework?