IMAGE INPAINTING USING LATENT DIFFUSION MODELS

Bhoumik Patidar Guntas Singh Saran Vamsi Srivathsa

Prof. Shanmuganathan Raman

Indian Institute of Technology Gandhinagar

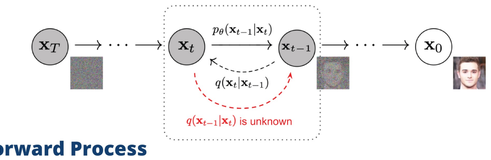
Computer Vision, Imaging, and Graphics (CVIG) Lab

Problem Statement

Editing masked patches by generating areas in the image masked out by the patch from a learnt latent diffusion model without specifically training it for this inpainting task.

Original Image Mask-removed Image Inpainted Image from LDM

Diffusion Process



Forward Process

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbb{I})$$

Reverse Process

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \left(\frac{(1-\bar{\alpha}_{t-1})\sqrt{\alpha_t}}{1-\bar{\alpha}_t}\mathbf{x}_t + \frac{(1-\alpha_t)\sqrt{\bar{\alpha}_{t-1}}}{1-\bar{\alpha}_t}\mathbf{x}_0\right), \frac{(1-\alpha_t)\cdot(1-\bar{\alpha}_{t-1})}{(1-\bar{\alpha}_t)}\mathbb{I})$$

Loss Function

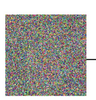
$$L_t^{\text{Simple}} = \mathbb{E}_{t\sim[1,T],\mathbf{x}_0,\boldsymbol{\epsilon}_t}\left[\|\boldsymbol{\epsilon}_\theta((\sqrt{\bar{\alpha}_t})\mathbf{x}_0 + (\sqrt{1-\bar{\alpha}_t})\boldsymbol{\epsilon}_t, t) - \boldsymbol{\epsilon}_t\|^2\right]$$

By simply learning to **predict the noise** in the image, the model learns the parameters of the reverse distribution

$\boldsymbol{\epsilon}_t$

Denoising Diffusion

Probabilistic Model



$\boldsymbol{\epsilon}_\theta$

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

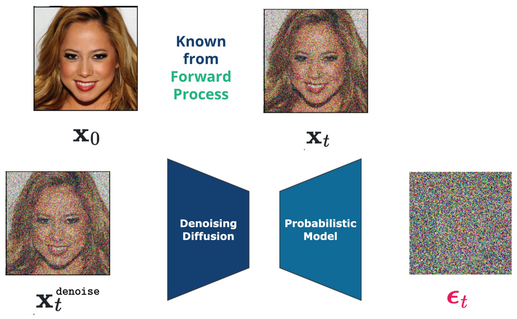
Sample from the reverse distribution

$\mathbf{x}_{t-1}$

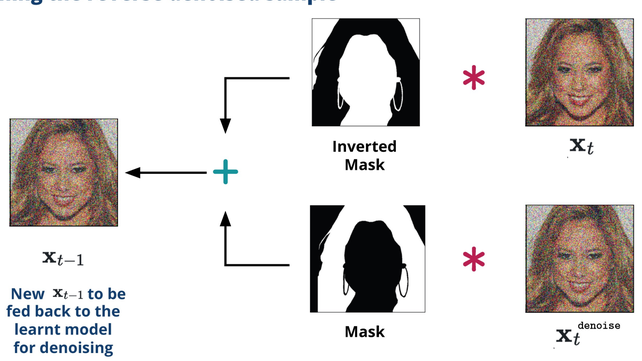
$$\mathbf{x}_{t-1} = \boldsymbol{\mu}_\theta + \sigma_t\mathbf{z} \quad \mathbf{z}\sim\mathcal{N}(\mathbf{0}, \mathbb{I})$$

Methodology

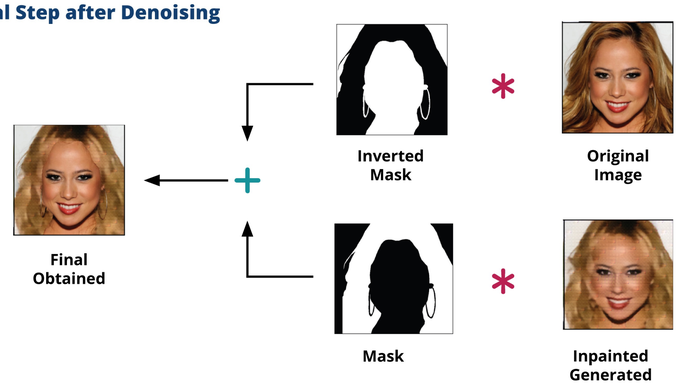
During Inference/Reverse Sampling



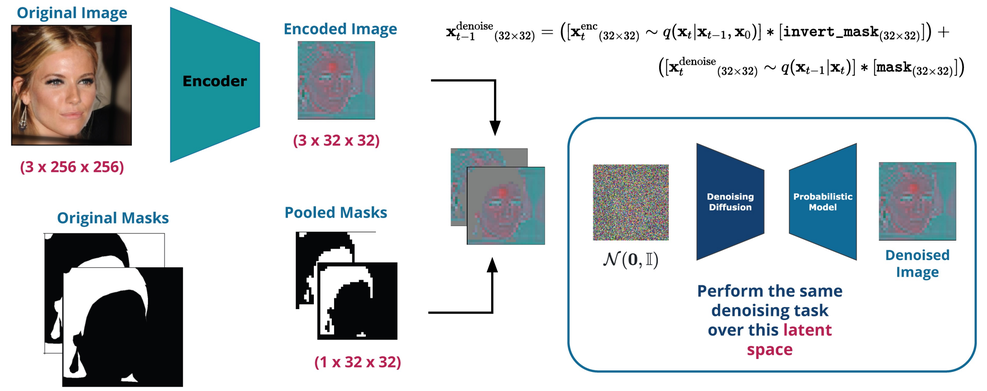
Obtaining the reverse denoised sample

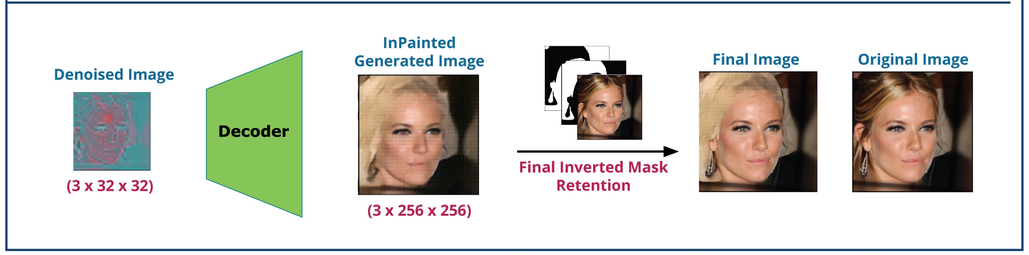


Final Step after Denoising



The Latent Space





Results