PatchDDM-Based Inpainting Pipeline for CelebA

7/14/2025

UNIVERSITY of ROCHESTER

1. Dataset Preparation

- CelebA faces, resized to 64×64.
- Custom PatchMaskDataset applies square patch masking.
- Returns (masked_image, original_image) pairs for training.

2. Denoising Model: UNet with Time Embedding

- Encoder–decoder structure using residual blocks.
- Downsampling and upsampling
- Time embedding encodes timestep t for conditional denoising.

3. Diffusion Framework (DDPM)

- Forward process: adds Gaussian noise to image (x_0 → x_t).
- Reverse process: model predicts noise to denoise x_t.
- Loss: MSE between predicted noise and true noise.

4. Training

- Input: noisy image x_t from q_sample(original image, t)
- Model output: predicted noise $\varepsilon_t(x_t, t)$.
- Loss: MSE(ε, $\varepsilon_t(x_t, t)$).
- Trained on masked images to enable inpainting.

5. RePaint / Sampling

Free-form image inpainting using an unconditional DDPM (no mask required in training)

Input:

- Image with missing pixels (binary mask m)
- Pretrained unconditional DDPM

Standard DDPM Sampling:

- Trained to reverse a Gaussian noise process
- Each reverse step:

$$x_{t-1} = \mu_\theta(x_t, t) + \sigma_t z$$