FastDiff: A Lightweight and Accelerated Diffusion Framework for Efficient Image-to-Image Generation

Authors: Jincheng Rao, Zihan Xiao
Course: ECE 661 – Machine Learning & Deep Neural Networks
Instructor: Prof. Hai "Helen" Li

1. Project Overview

Diffusion models have emerged as a dominant paradigm for high-quality image generation, capable of synthesizing realistic and diverse samples by learning to reverse a gradual noising process. This project aims to develop FastDiff, a lightweight and accelerated diffusion framework designed for efficient image-to-image generation. By optimizing the model architecture and sampling strategy, FastDiff seeks to achieve competitive visual quality with significantly reduced computational cost.

2. Objectives

- Build a DDPM training pipeline for efficient image denoising and generation.
- Implement DDPM and DDIM samplers to compare stochastic vs. deterministic inference.
- Evaluate the trade-offs between generation quality and efficiency.
- Explore optional extensions such as classifier-free guidance or image inpainting.

3. Methodology

The proposed approach consists of three main components:

(1) Lightweight DDPM Training Pipeline: Implement forward and reverse diffusion using Gaussian noise schedules (linear and cosine). Use sinusoidal timestep embeddings and a compact U-Net architecture with reduced depth and channel width. Maintain an exponential moving average (EMA) of model weights for stability.

(2) Accelerated Sampling with DDIM: Integrate deterministic DDIM sampling with variable step budgets (e.g., 50, 100, 200 steps) to achieve faster generation while maintaining quality.

(3) Evaluation on Image-to-Image Tasks: Apply trained models to image translation and denoising tasks (CIFAR-10, MNIST) and explore classifier-free guidance for controllable generation.

4. Risks and Mitigation

Risk 1: Potential degradation of image fidelity when reducing steps or network size. Mitigation: Gradually reduce sampling steps and fine-tune β-schedules.
Risk 2: Training instability due to reduced model capacity. Mitigation: Apply gradient clipping, smaller learning rates, and cosine decay.
Risk 3: Computational constraints. Mitigation: Use reduced datasets (32×32 CIFAR-10) and smaller batch sizes.

5. Project Timeline

Week	Date Range	Milestone
1	Oct 21–27	Literature review and baseline DDPM implementation
2	Oct 28–Nov 3	Train baseline DDPM and analyze results
3	Nov 4–10	Implement and test DDIM accelerated sampling
4	Nov 11–17	Midterm check-in: evaluate speed–quality tradeoff
5	Nov 18–Dec 1	Extend to image-to-image translation (denoising)
6	Dec 2–9	Poster preparation, final experiments, visualization

6. Evaluation and Expected Outcome

Quantitative Metrics:
- Fréchet Inception Distance (FID) and Inception Score (IS) for generated images.
- Sampling time per image vs. quality (efficiency–quality curve).

Qualitative Metrics:
- Visual comparison between DDPM and DDIM.
- Demonstration of fast image-to-image generation tasks.

Expected Outcome:
By the end of the project, we expect to deliver a functional and optimized FastDiff framework capable of generating visually appealing images with significantly fewer sampling steps, demonstrating both efficiency and expressiveness.

7. References

1. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. NeurIPS 33.
2. Nichol, A. Q., & Dhariwal, P. (2021). Improved Denoising Diffusion Probabilistic Models. ICML 2021.
3. Dhariwal, P., & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. NeurIPS 34.