Course Project

Code of Honor. All external resources used in the project, including research papers, open-source repositories, datasets, and any content or code generated using AI tools, e.g., ChatGPT, GitHub Copilot, Claude, Gemini, must be *clearly cited* in the final submission. The final report must also include *a clear breakdown of individual group member contributions.* Any lack of transparency in the use of external resources or in reporting group contributions will be considered academic dishonesty and will significantly impact the final evaluation.

Topic	Tiny Diffusion Model with Alternative Core for Image Generation
Category	Tiny AI Products
Supervisor	Likun Cai

Objective Design and implement a **tiny diffusion model** for low-resolution image generation, with a focus on architectural simplification, ablation analysis, and experimentation. Rather than using existing *denoising diffusion probabilistic models (DDPMs)* implementations, students are expected to *build a minimal functional prototype from scratch*, inspired by the original DDPM paper and recent simplifications.

Motivation Diffusion models are state-of-the-art generative methods. This project focuses on a small scale DDPM [1] with minimal functional prototypes to help understanding what goes on behind the scene of these powerful models.

Anti-Plagiarism Note The final model *must* be implemented *from scratch* and not adapted from existing DDPM repositories. Using such code as reference is permitted, but copying core training, sampling, or noise-scheduling code is considered academic dishonesty. The groups may optionally compare their design against these baselines, but all components *must be written and documented by the group.*

Requirements The final submission should fulfill the following requirements. All components must be implemented by the group members from scratch.

1. Implement a minimal DDPM using a small dataset such as CIFAR-10, MNIST, or Fashion-MNIST. The model must be implemented from scratch, including the forward and reverse processes.
2. Explore and compare at least two noise schedules, e.g., linear, cosine, and report their effect on generated sample quality.
3. Replace the UNet in the standard DDPM implementation with an alternative architecture, such as a ResNet, CNN or MLP-based encoder-decoder. Justify your design and analyze potential trade-offs.
4. Include a simple quantitative evaluation, e.g., FID or CLIP similarity, and visual inspection of generated outputs.
5. All results, design choices, and ablation experiments should be documented clearly in the final report.

Milestones

1. *Literature review and dataset selection.* Review the DDPM paper [1] and at least one simplified diffusion implementation. Select a dataset.

2. *Minimal DDPM.* Implement the full DDPM pipeline, including noise schedule and forward-reverse processes.

3. *Core model replacement.* Design and implement a custom architecture to replace UNet. Compare training stability and sample quality with the original design.

4. *Schedule comparison.* Train your model under different noise schedules and evaluate their effects.

5. *Final tuning and reporting.* Produce your final samples, document the performance, and elaborate your observations.

Submission Guidelines The main body of work is submitted through Git. In addition, each group submits a final paper and gives a presentation. In this respect, please follow these steps.

- Each group must maintain a Git repository, e.g., GitHub or GitLab, for the project. By the time of final submission, the repository should have
  - Well-documented codebase
  - Clear `README.md` with setup and usage instructions
  - A `requirements.txt` file listing all required packages or an `environment.yaml` file with a reproducible environment setup
  - Demo script or notebook showing sample input-output
  - *If applicable,* a `/doc` folder with extended documentation

- A final report (maximum *5 pages*) must be submitted in a PDF format. The report should be written in the provided formal style, including an abstract, introduction, method, experiments, results, and conclusion.
  Important: Submissions that do not use template are considered *incomplete.*

- A 5-minute presentation (maximum *5 slides including the title slide*) is given on the internal seminar on Week 14, i.e., *Aug 4 to Aug 8,* by the group. For presentation, any template can be used.

Final Notes While planning for the milestones please consider the following points.

1. You are encouraged to explore innovative approaches to conditioning or generation as long as the core objectives are met.

2. While computational resources are limited, carefully chosen datasets and training setups can make even diffusion models feasible. Trade-offs, e.g., resolution, training steps, are expected and should be justified.

3. Teams are expected to manage their computing needs and are advised to perform early tests to estimate runtime and training feasibility. As graduate students, team members can use facilities provided by the university, e.g., ECE Facility. Teams are expected to inform themselves about the limitations of the available computing resources and design the model accordingly.

References

[1] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 33:6840–6851, 2020.