Enfusion

CS-726 Project

Team Con-Fusion - Abhinav Garg, Dhiraj Kumar Sah, Tejomay Padole, Prathamesh Yeole

Overview

- Introduction
- Background & Motivation
- Proposed Methodology
- Results & Analysis
- Conclusion
- References

Introduction

- Denoising Diffusion Probabilistic models
- Noise Schedulers
- Applications

Background & motivation

Can one scheduler learn everything well?

Objectives

- Can we leverage on some weaker Denoisers that learn only specific noise-level-based steps?
- How can we create an ensemble of such methods?
- Can we use something like this to speed up inference?

Proposed Method

- Have multiple UNets to learn different stages
- Ensure that 2 models are more focused on the extreme noise levels
- Train a single model for some steps and then use its weights as a starting point for others.

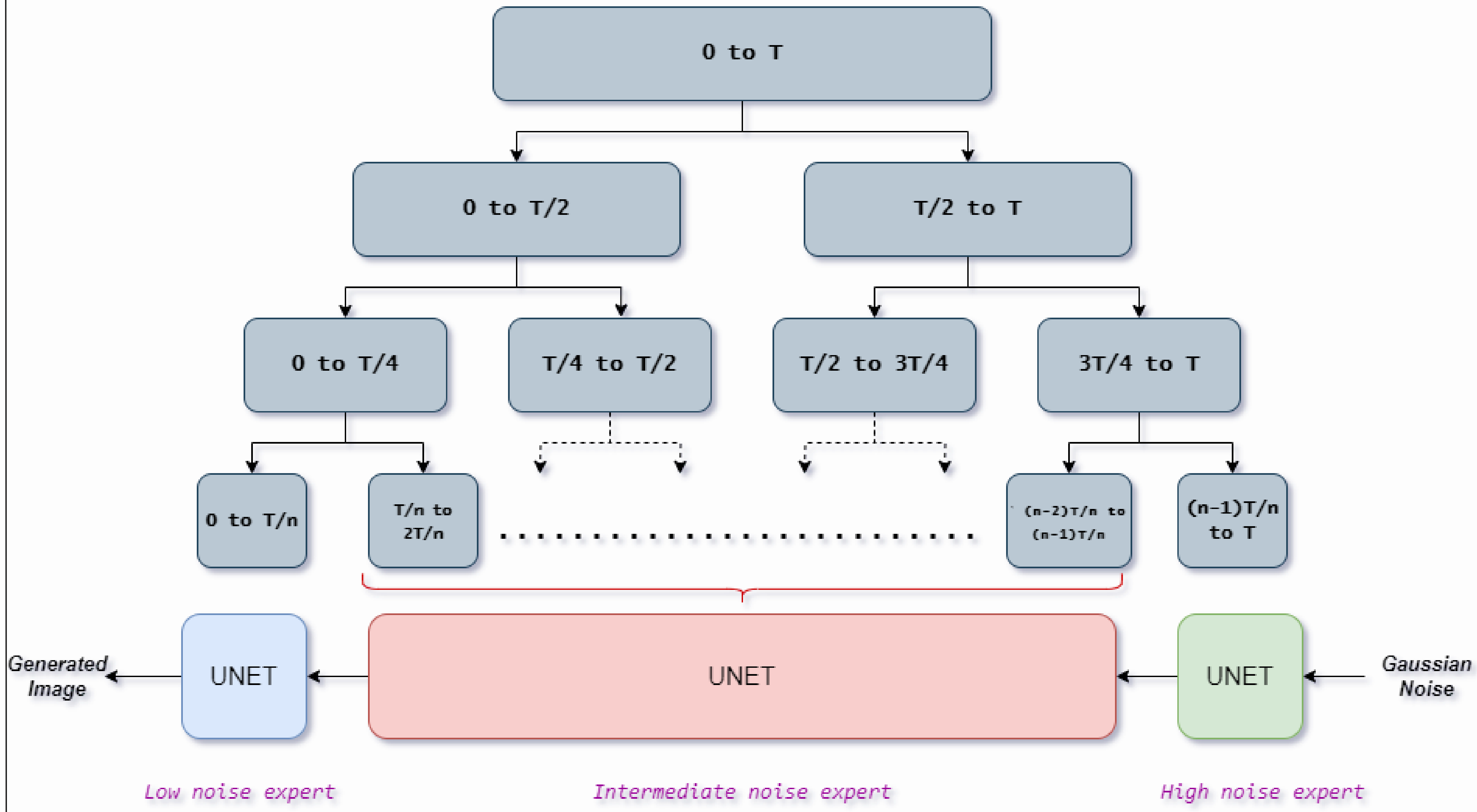


Fig. Proposed Ensemble Model

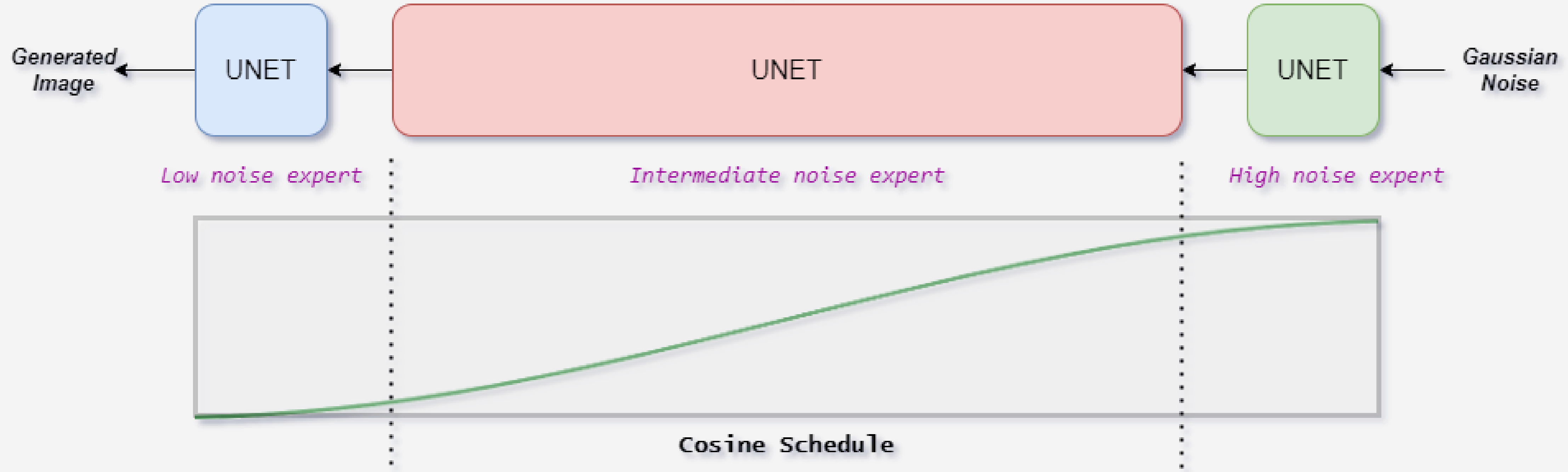
Results

Samples from vanilla DDPM

Samples from EnFusion DDPM

Model	Earth Mover's Distance	Chamfer's Distance
Vanilla DDPM	2666.7165	60613.0382
EnFusion DDPM	2571.7633	59792.7957

Analysis

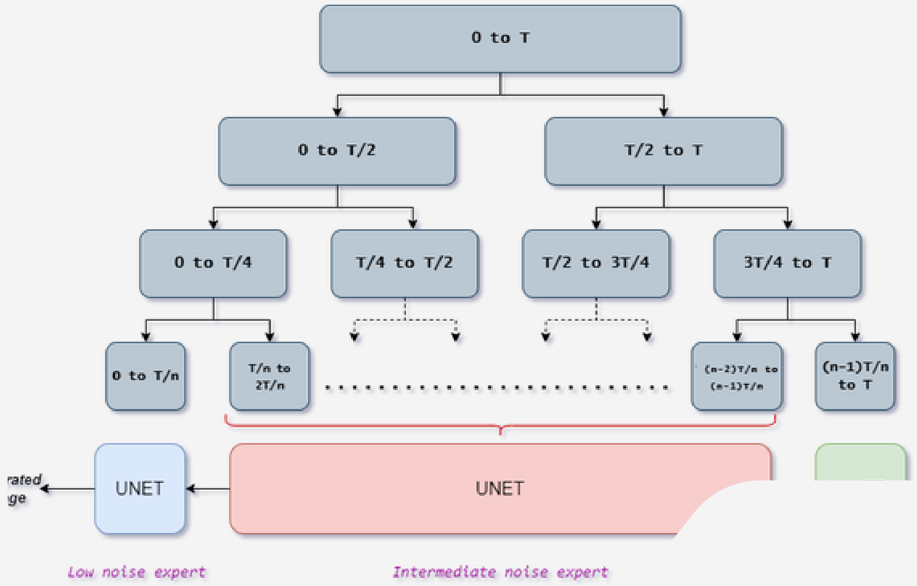


- We introduce dedicated Denoisers for a specific range of timesteps so the model is better trained on a specific task accordingly.
- Using stage-wise learning to help the denoisers get familiar with a common representation space.
- Theoretically, we can implement a pipeline for inference and thus speed it up for a huge amount of sample generation.

Conclusion

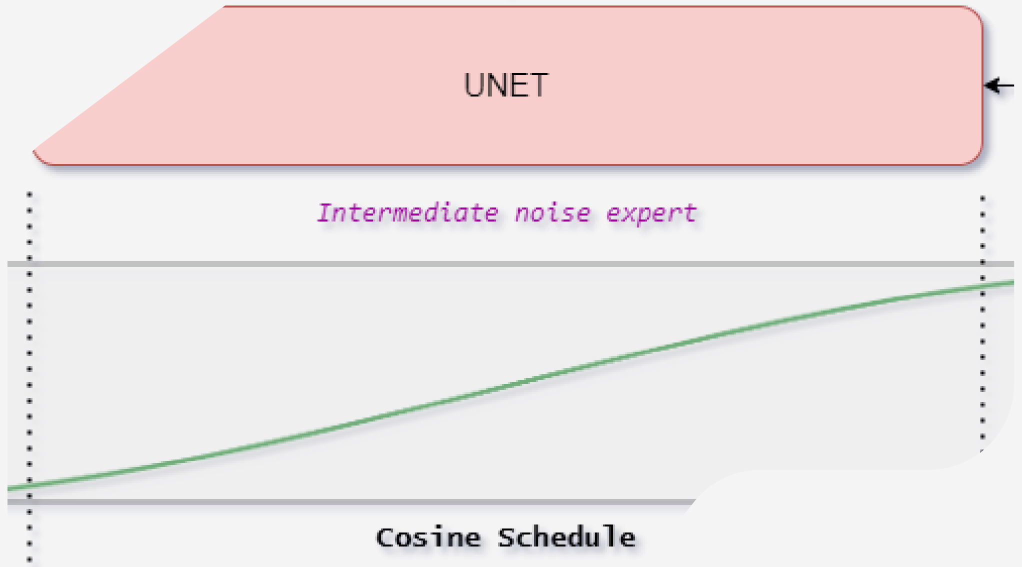
DDPMs use a single Denoiser

We expect a huge model to learn denoising patterns for the entire noise range. However, it'd be easier to get specific models to learn the denoising stages.



Using different denoisers for diffent noise levels

Using an ensemble of denoisers to learn specific noise level based model.



Can use pipelining to speed up the inference process for many samples.

As the 3 UNets here are working or different noise ranges, we can run these in parallel to get some speed up.

References

- Yogesh Balaji, Seungjun Nah, Xun Huang, Arash Vahdat, Jiaming Song, Qinsheng Zhang, Karsten Kreis, Miika Aittala, Timo Aila, Samuli Laine, Bryan Catanzaro, Tero Karras, and Ming-Yu Liu. ediff-i: Text-to-image diffusion models with an ensemble of expert denoisers, 2023

- Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In Marina Meila and Tong Zhang, editors, Proceedings of the 38th International Conference on Machine Learning, volume 139 of Proceedings of Machine Learning Research, pages 8162–8171. PMLR, 18–24 Jul 2021.

- Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS'20, Red Hook, NY, USA, 2020. Curran Associates Inc.