

Enhancing Generative Modeling with Hybrid VAE-Diffusion Architectures

Project Overview and Instructions

Deep Generative Models

October 30, 2024

Project Overview

- Explore hybrid VAE-Diffusion models.
- Combine VAE's latent representations with diffusion model's refinement.
- Replicate DiffuseVAE baseline; analyze extensions (joint training, adaptive conditioning).

Stage 1: Baseline Replication

Objective: Replicate DiffuseVAE to understand hybrid modeling.

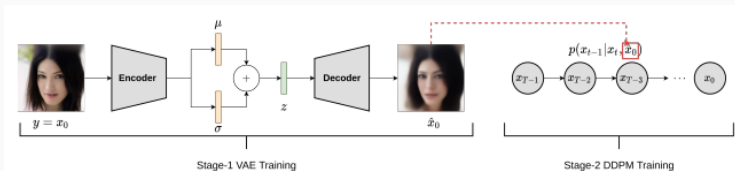


Figure 2: Proposed DiffuseVAE generative process under the simplifying design choices discussed in Section 3.2. DiffuseVAE is trained in a two-stage manner: The VAE encoder takes the original image x_0 as input and generates a reconstruction \hat{x}_0 which is used to condition the second stage DDPM.

■ Tasks:

- ▶ Implement VAE and diffusion models.
- ▶ Datasets: CIFAR-10 and CelebA-64.

■ Evaluation:

- ▶ Quantitative: FID and IS metrics.
- ▶ Qualitative: Visual inspection.

Stage 2: Joint End-to-End Training

Objective: Experiment with joint training for better alignment.

■ Tasks:

- ▶ Implement joint training with a weighted loss.
- ▶ Experiment with complex hybrid losses.

■ Evaluation:

- ▶ Compare FID, IS, and visualize latent space.

Stage 3: Adaptive Conditioning

Objective: Implement adaptive conditioning based on VAE output.

■ **Tasks:**

- ▶ Use latent space variance for conditioning.
- ▶ Implement flow-based prior in VAE.

■ **Evaluation:**

- ▶ Compare FID, IS, and sampling efficiency.

Objective: Systematically evaluate each extension.

■ **Guiding Questions:**

- ▶ **Quality:** Does joint training improve FID?
- ▶ **Efficiency:** How does adaptive conditioning impact sampling time?
- ▶ **Robustness:** How do models respond to noisy outputs?

- Follow specified format: Introduction, Methodology, Results, Discussion, Conclusion.
- Include a reflective section discussing lessons learned and future improvements.

Questions?

- Any questions about the project instructions or requirements?