Abstract

Variational Autoencoders (VAEs) have been widely used in the fields of representation learning and generative modeling. However, VAEs often fall short in generating high-quality samples when compared to other generative models like Generative Adversarial Networks (GANs) and Denoising Diffusion Probabilistic Models (DDPMs). In this paper, we propose a method that integrates the strengths of DDPMs into the VAE framework. Traditional VAEs constrain the encoder to learn encodings that follow a certain prior distribution, however, our model uses DDPMs to learn a flexible distribution over compressed representations of data samples. The advantage of our approach lies in its efficiency compared with GANs and DDPMs—our model learns distributions over a low-dimensional latent space instead of the full data space. Our model’s efficacy is demonstrated through experiments conducted on three benchmark datasets: MNIST, Fashion MNIST, and CIFAR-10. Our model outperforms several existing autoencoder-based methods based on FID scores.