Literature Review of Denoising Diffusion Probabilistic Models

Anagha H.C. Sachin Prasanna Abhayjit Singh

October 2023

1 Denoising Diffusion Probabilistic Model

Diffusion Probabilistic Models, often abbreviated as DDPMs, are a noteworthy category within the realm of generative models. They do so by iteratively introducing controlled noise into an initial input signal. The underlying concept is to acquire a deep understanding of the noise removal process, enabling the generation of entirely fresh and coherent data samples.

One notable achievement of DDPMs is their ability to produce high-quality images through a process inspired by nonequilibrium thermodynamics. This process involves iteratively applying noise and denoising to create new samples with impressive image synthesis results [1]. Further, Improved Denoising Diffusion Probabilistic Models have shown promising results in various applications, demonstrating enhanced capabilities in generating high-quality images with reduced noise levels [3].

This methodology stands in contrast to traditional generative models, which typically focus on modeling data distribution and sampling from it directly. In DDPMs, the noise injection and subsequent denoising steps are central to the creative process. This approach draws inspiration from the principles of diffusion, mimicking how noise dissipates in physical systems, such as in nonequilibrium thermodynamics.

2 Generative Models

Generative models are a class of artificial intelligence algorithms that focus on creating data rather than making predictions. They learn the underlying patterns and structures within datasets, enabling them to generate new, synthetic data samples. These models have diverse applications, including generating realistic images, natural language text and audio. The ability to generate data with high fidelity and diversity has significant implications for fields like computer vision, natural language processing, and data augmentation. DDPMs also have a similar aim of learning the distribution of a training data sample and then generating a new sample that closely resembles it.

3 Generative Adversarial Networks

Generative Adversarial Networks (GANs), introduced by Ian Goodfellow and colleagues in 2014, comprise two crucial neural networks: a generator and a discriminator. The generator and discriminator (aka critic).The generator produces a sample, such as an image, from a latent code. Ideally, the distribution of these images should be indistinguishable from the training distribution[2], while the discriminator's role is to distinguish real data from generated data. They engage in a competitive game, with the generator refining its output to resemble genuine data and the discriminator enhancing its ability to differentiate. Typically, a GAN consists of two networks: GAN training strikes a balance with a dynamic feedback loop; as the generator improves, the discriminator adapts, fostering ongoing competition. GANs excel in producing highly realistic data for computer vision, art, and data augmentation. In a related context, adversarial nets establish a competitive framework, pitting a generative model against a discriminative model. The generative model aims to craft "counterfeit" samples indistinguishable from genuine data, while the discriminative model detects discrepancies [5]. This approach, trainable using methods like backpropagation and dropout algorithms, holds promise for deep generative modeling, aligning with our exploration of Diffusion Models in this literature review.

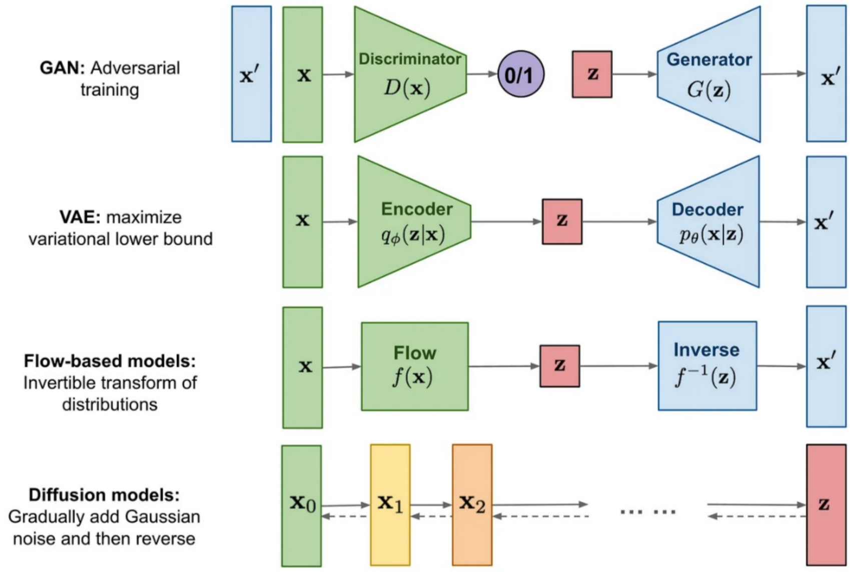

Figure 1: A Pictorial Comparison between Generative Models

4 Comparing Generative Adversarial Networks (GANs) and DDPMs

In recent years, generative models have made remarkable progress in producing realistic natural language, high-quality synthetic images, and diverse human speech and music. These models have found applications across various domains, from generating images based on text prompts to acquiring valuable feature representations. However, GANs are notorious for their training challenges, often demanding careful selection of hyperparameters and regularizers. Despite their cutting-edge performance, GANs exhibit limitations that hinder their scalability and applicability in new domains. Consequently, significant efforts have been devoted to achieving GAN-like sample quality with likelihood-based models. These models, such as diffusion models, have recently demonstrated the ability to generate high-quality images while also providing advantageous characteristics like distribution coverage, a stationary training objective, and easy scalability. Diffusion models generate samples by gradually reducing noise from a signal, with their training objective expressed as a reweighted variational lower-bound [4].

5 Generative Modelling in Medical Image Generation

Kebaili et. al highlighted the promising role of deep generative models in overcoming the challenge of limited training data in medical image analysis. They found each model type - VAEs, GANs, and DMs - presents distinct strengths and limitations [18]. Khader et. al demonstrated the potential of diffusion probabilistic models, known for their success in image generation, in synthesizing high-quality medical imaging data, specifically MRI and CT images [12]. Bieder et. al introduced PatchDDM, a memory-efficient patch-based diffusion model, which enables the application of denoising diffusion models to large 3D medical imaging datasets, demonstrating its effectiveness in generating meaningful segmentations for tumor analysis in the BraTS2020 dataset [9].

Zhou et. al introduced a novel framework utilizing vector-quantization GAN and a transformer with masked token modeling to generate diverse and high-resolution 3D brain tumor regions of interest (ROIs), effectively addressing imbalanced data challenges in brain tumor classification tasks [10]. Frisch et. al proposed a novel approach utilizing a conditional generative model based on DDIM and CFG to address imbalances in cataract surgery data, successfully synthesizing high-quality examples and improving tool classification performance by up to 10% for rare cases [11]. Kim and Ye introduced a novel diffusion deformable model (DDM) for generating intermediate temporal volumes in 4D medical imaging, specifically focusing on cardiac MR images, showcasing superior performance compared to existing deformation methods [16]. Fu et. al introduced a novel recycling method in denoising diffusion models for image segmentation, aligning training with inference and outperforming standard diffusion training across multiple medical imaging datasets, ultimately achieving on-par performance with non-diffusion-based supervised training [18].

6 Generative Modelling in Retinal Image Generation

Liskowski et. al presented a supervised segmentation technique for retinal vascular network analysis, leveraging deep neural networks trained on a large dataset with preprocessing techniques [6]. Costa et. al presented a novel approach to directly synthesize eye fundus images from vessel trees, demonstrating visually distinct results while maintaining high image quality compared to true images [13]. Guo et. al introduces RetiGAN, a GAN-based retinal image generation model addressing data scarcity and class imbalance in medical images. It employs a dual-scale discriminator and content loss embedding for improved semantic retention, and incorporates edge enhancement techniques for clearer detail [15]. Zhang et. al put forth AG-Net, an Attention Guided Network, which effectively preserves structural information in retinal image segmentation tasks, demonstrating notable improvements in blood vessel, optic disc, and cup segmentation compared to conventional CNN approaches [14].

7 Other Notable Applications of DDPM

Nair et. al provided a cutting-edge solution for Thermal-to-Visible (T2V) image translation in low-light conditions, utilizing a DDPM to achieve state-of-the-art results across various datasets [7]. They also use DDPMs for atmospheric turbulence mitigation in long-range imaging, offering a stable training process and impressive results, particularly in reconstructing facial features under high turbulence conditions [8].

8 Individual Contribution

Student Name	**Paper Referred**
Sachin Prasanna	1, 2, 8, 12, 13, 16
Anagha H C	3, 4, 7, 11, 14, 17
Abhayjit Singh Gulati	5, 6, 9, 10, 15, 18

References

1. https://arxiv.org/pdf/2006.11239.pdf - Denoising Diffusion Probabilistic Models

2. https://arxiv.org/pdf/1710.10196.pdf - Progressive Growing of GANS for Improved Quality, Stability and VariationN

3. https://arxiv.org/pdf/2102.09672.pdf - Improved Denoising Diffusion Probabilistic Models

4. https://arxiv.org/pdf/2105.05233.pdf - Diffusion Models Beat GANs on Image Synthesis

5. https://arxiv.org/pdf/1406.2661.pdf - Generative Adversarial Nets

6. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=arnumber=7440871 - Segmenting Retinal Blood Vessels With Deep Neural Networks

7. https://arxiv.org/pdf/2209.08814.pdf - T2V-DDPM: Thermal to Visible Face Translation using Denoising Diffusion Probabilistic Models

8. https://arxiv.org/pdf/2208.11284.pdf - AT-DDPM: Restoring Faces degraded by Atmospheric Turbulence using Denoising Diffusion Probabilistic Models

9. https://arxiv.org/pdf/2303.15288.pdf - Diffusion Models for Memory-efficient Processing of 3D Medical Images

10. https://arxiv.org/pdf/2310.01251.pdf - Generating 3D Brain Tumor Regions in MRI using Vector-Quantization Generative Adversarial Networks
11. https://arxiv.org/pdf/2308.02587.pdf - Synthesising Rare Cataract Surgery Samples with Guided Diffusion Models
12. https://arxiv.org/pdf/2211.03364.pdf - Medical Diffusion: Denoising Diffusion Probabilistic Models for 3D Medical Image Generation
13. https://arxiv.org/abs/1701.08974 - Towards Adversarial Retinal Image Synthesis
14. https://www.researchgate.net/publication/336396983 - Attention Guided Network for Retinal ImageSegmentation
15. https://www.sciencedirect.com/science/article/pii/S1746809422004748 - A novel retinal image generation model with the preservation of structural similarity and high resolutionc
16. https://arxiv.org/pdf/2206.13295.pdf - Diffusion Deformable Model for 4D Temporal Medical Image Generation
17. https://arxiv.org/pdf/2307.13125.pdf - Deep Learning Approaches for Data Augmentation in Medical Imaging: A Review
18. https://arxiv.org/pdf/2308.16355.pdf - A recycling training strategy for medical image segmentation with diffusion denoising models