iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** X **Month of publication:** October 2025

DOI:

Enhancing AI Imagery: The Rise of Denoising Diffusion Models in Real-life Applications

Shipra Bhanja
Dr. D.Y. Patil College of Engineering, Akurdi

Abstract: Denoising Diffusion Probabilistic Models (DDPMs) have quickly become a reliable and effective technique for generating, and enriching images. The key difference with DDPMs, compared with adversarial models such as GANs, is that DDPMs relies on an incremental process for adding, and in-turn, removing noise. The incremental process builds in training stability, and produces high-quality, and diverse images. The structure and flexibility of DDPMs, compared with GANs, are a beneficial advantage of DDPMs for application in real-world scenarios including biometrics, medical imaging, and design for creativity, or design for industry. Moreover, recent models such as eDffIQA have improved computational speed and efficiency which is useful in time-resource limited environments when applying diffusion-based systems. This paper will describe the key underlying principles, mathematical concepts, and overall observable applications of DDPMs, including the speed-up in computational efficiency and relevant advances, as diffusion-based systems now have a growing impact in areas of applied AI imagery.

Keywords: Diffusion Models, Image Generation, Biometrics, Medical Imaging, Neural Networks, Image Enhancement

I. INTRODUCTION

Artificial Intelligence has revamped methodologies for digital image processing in terms of how we create, interpret, and alter visual data. Recent deep learning methods – like convolutional neural networks (CNNs) - have shown significant improvement on tasks such as classification and segmentation. Generative Adversarial Networks (GANs) have also advanced the field by enabling image synthesizing and creative image generation - however, in practice, GANs are often unstable in training, suffer from issues such as vanishing gradients, and sometimes limit variation in the synthesizing outputs due to their mode collapse problems.

Denoising Diffusion Probabilistic Models (DDPMs) demonstrate a stable alternative and exploit ideas from stochastic processes and statistical physics. DDPMs operate on a two-step Markov process approach: a forward process which incrementally adds noise to the data and a reverse process that utilizes the sampled noise to reconstruct the image. The procedures of incremental refinement illustrated in DDPMs avoids qualities that occasionally manifest in adversarial based methods, while doing so allows for generation of highly realistic images with variability. Diffusion models have various applications in different fields. In biometric systems, such models aid in enhancing face image quality and eliminating low-quality samples that reduce identification performance. In the medical field, they are used to enhance MRI and CT scans by denoising and recovering details, which can potentially assist with medical diagnoses. In creative fields, diffusion-based methods support tasks such as text-to-image translation, concept art, and film production. The recent introduction of eDifFIQA illustrates that knowledge distillation can improve the speed and practicality of these models' developing applications for real-time contexts. The purpose of this study is to review the theoretical and applied aspects of diffusion models. The structure of the models as well as improvements to efficiency, and applications of the models will be discussed in this chapter. The authors seek to review the progress of DDPMs and their influence on image generation and enhancement in both academic and industry settings.

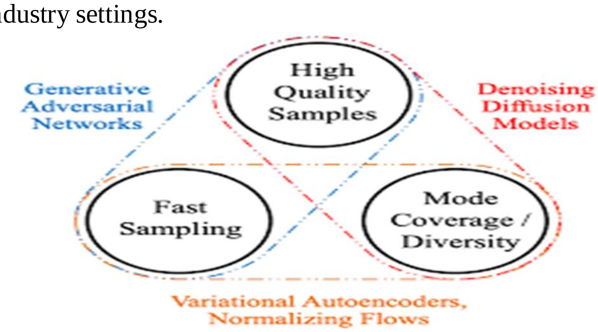

Fig. 1 GAN vs VAE vs DDM (Source: Nvidia Developer Blog)

II. METHODS

A. Conceptual Overview

At the heart of a diffusion model is teaching a system to reconstruct data that has been incrementally corrupted. By training to learn the opposite of the corruption data process, the model learns the underlying distribution of the data itself. These approaches dispute the competitive nature in these GANs and promotes more reliable and stable learning.

B. Forward Diffusion Process

In the forward process, an uncorrupted image x_0 is gradually corrupted over a number of time steps. At timestep t, Gaussian noise is added according to a schedule β_t. The process is given as:

$$q(x_t|x_{t-1}) = N(x_t; \sqrt{1-\beta_t}\, x_{t-1}, \beta_t I)$$

After many iterations, the image becomes nearly indistinguishable from random noise. This process is given inclosed form by:

$$q(x_t|x_0) = N(x_t; \sqrt{\bar{\alpha}_t}\, x_0, (1-\bar{\alpha}_t) I)$$

here $\bar{\alpha}_t = \prod_{s=1}^{t} (1-\beta_s)$.
This formulation allows direct sampling at any point in the process.

C. Reverse Denoising Process

The reverse process is where generation actually occurs. A neural network predicts and removes the noise step-by-step by slowly reconstructing the clean image. The reverse transition can be modelled as:

$$p_\theta (x_{t-1}|x_t) = N(x_{t-1};$$
$$\mu_\theta (x_t, t), \sigma_t^2 I)$$

U-Net architectures, often used alongside attention mechanisms, are effective because they capture both local textures and global structures within an image.

D. Training Objective

The model is optimized to minimize the error between the actual noise and the noise predicted by the neural network. This mean-squared error objective guarantees stable optimization, resolving adversarial training's instability and generated diversity. The loss function can be written as:

$$L_{DDPM} = E_{(x_0, t, \epsilon)} [\|\epsilon - \epsilon_\theta (x_t, t)\|^2]$$

E. Improving Efficiency

Diffusion models generate very realistic images but are generally expensive in computational cost, producing an image in hundreds of denoising steps. There are several approaches to make them more usable:

1) Knowledge Distillation: A smaller "student" network distils the full diffusion "teacher" model while maintaining performance and at a significantly reduced computation cost. An example of this is the eDifFIQA model.
2) Latent Diffusion Models (LDMs): These models do not operate directly on the pixel data but rather they run the diffusion process in a compressed latent space. This method, which is used in Stable Diffusion, greatly reduces computational resources needed while yielding images of similar quality.
3) Accelerated sampling: Sampling techniques like Denoising Diffusion Implicit Models (DDIM) and SDE-based solvers are able to reduce the sampling steps from thousands to a few dozen, which drastically increases the image generation speed to real-time for interactive systems.

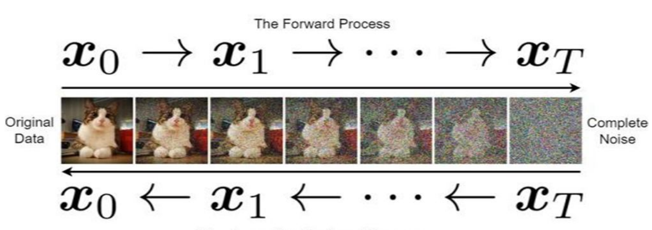

Fig. 2 Forward and Reverse Diffusion Process

III. APPLICATIONS

A. Biometrics

In biometric identification using face, fingerprint, or iris images, diffusion models enhance reliability and fairness in identification systems. For example, eDifFIQA, using principles of diffusion, identifies and filters low-quality face images to improve recognition accuracy and demographic fairness.

B. Medical Imaging

Medical imaging needs to be clear and precise. DDPMs have also been demonstrated to denoise MRI and CT scans, enabling the clinician to visualize critical diagnostic characteristics more clearly, since the images become more defined. DDPMs are also able to produce synthetic medical images for the purpose of training machine learning models, while maintaining patient confidentiality.

C. Creative and Artistic Design

Diffusion models are transforming digital art and design practices. They allow artists and designers to create photorealistic images, conceptual sketches, and digital effects based on written descriptions or prompts. High-end generative capabilities, like those used by Stable Diffusion, are now accessible to most designers and developers.

D. Image Restoration and Enhancement

Methods that are based on diffusion show performance benefits over conventional approaches when recovering visual clarity from images that have been blurred, taken in low-light conditions, or are noisy. These methods are exploited in satellite imaging problems, the preservation of cultural heritage, and the enhancement of surveillance footage. Diffusion-based methods generate visual outputs that appear more natural and exhibit additional visual detail.

E. Video Processing

Diffusion-based methods demonstrate improved performance over traditional methods in recovering visual clarity from images that are blurred, captured in low-light conditions, and are noisy. Diffusion-based methods have been used in satellite imaging applications, cultural heritage preservation, and improving surveillance footage. Diffusion methods create visual outputs that look more natural and have more detailed visual information.

F. Interactivity with Human-AI

Diffusion models are useful in new forms of interaction to produce customized avatars, immersive environments, and virtual learning tools. They combine creativity and ease of use to provide intuitive user control over AI content, thereby increasing user agency.

IV. DEPLOYMENT

Most diffusion models leverage either U-Net or transformer backbones to extract and handle visual data. Deployments of diffusion models most often leverage various combinations of a small number of optimization strategies to balance quality and efficiency:

1) *Fast Sampling:* Sampling strategies such as DDIM make use of fewer diffusion steps.
2) *Model Districting:* This usually refers to reducing larger diffusion systems at a cost of performance.
3) *Fast Hardware:* This is where more contemporary hardware, like GPUs, or a distributed computing infrastructure as standard for large-scale training.
4) *Edge deployment:* Efficient implementations and attention mechanisms for real-time use in mobile and embedded devices.

An example like eDifFIQA has demonstrated that in certain scenarios more efficient architectures can drastically reduce compute time without any disadvantage to healthful performance (in regard to image quality), which makes a huge difference in real export situations.

V. CONCLUSION

Denoising diffusion models have rapidly emerged as one of the most enticing contemporary pipelines targeted towards image generation. Images are generated with high quality, consistency, and diversity that make them suitable for applications in various fields, such as healthcare, security, art, and design. New developments in model distillation, latent-space representations, and multimodal representations are rapidly increasing the speed, interpretability, and efficiency of diffusion systems.

Moving forward, greater attention is expected to be placed on transparency, ethical assurances, and explainability in order to initiate responsible use of these technologies for making decisions in the foreseeable future. It is also anticipated that diffusion will continue to spearhead innovation in intelligent visual technologies.

REFERENCES

[1] Ž. Bubnik, P. Peer, and V. Struk, "eDifFIQA: Towards Efficient Face Image Quality Assessment Based on Denoising Diffusion Probabilistic Models," IEEE Transactions on Biometrics, Behavior, and Identity Science, vol. 6, no. 4, 2024.
[2] Abdullah, T. Huang, I. Lee, and E. Ahn, "High-resolution histopathology whole slide image generation using wavelet diffusion model," IEEE ISBI 2025.
[3] H. Chen et al., "Comprehensive exploration of diffusion models in image generation: a survey," Springer, 2025.
[4] A. Iuliano et al., "Probabilistic diffusion models for synthetic medical image generation," IEEE MetroLiveNV, 2024.
[5] M. K. Song et al., "Principal Component Diffusion-based Image Generation Model," IEEE Access, 2024.
[6] Q. Gao et al., "Relevant Error-Modulated Diffusion Model for Low-Dose CT Denoising," IEEE Transactions on Medical Imaging, vol. 43, no. 2, 2024.
[7] J. Yang, D. Zhou, J. Feng, and X. Wang, "Diffusion Probabilistic Models Made Slim," arXiv preprint, 2024.