Exploring Diffusion Models: A Survey in Super-Resolution and In-Painting Applications

Stelios Zarifis

National Technical University of Athens

el20435@mail.ntua.com

Advanced AI Student Conference
January 14, 2024

Overview

What are Diffusion Models (DMs)?

- DMs are a class of generative AI models.

How DMs work?

- They gradually add Gaussian noise to the original data in a forward diffusion process and then learn to remove it in a reverse process.

Tasks where DMs work well:

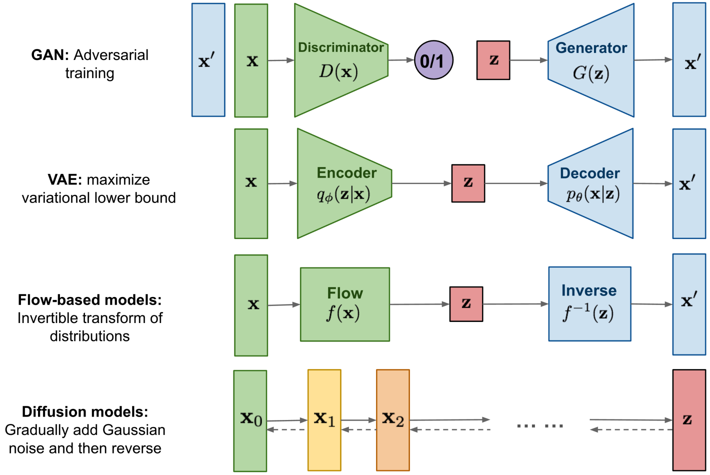
- Image In-painting
  - RePaint
  - CoPaint
- Super-Resolution
  - Latent Diffusion Models
  - Partial Diffusion Models
  - DiffPIR

Deep Generative Models (DGMs)

DGMs:

- generate new data similar to the training data.
- capture the input data distribution through unsupervised learning.
- use deep neural networks to learn a complex mapping from a single latent vector z to the observed data x.

Generative Models Architectures



[1]

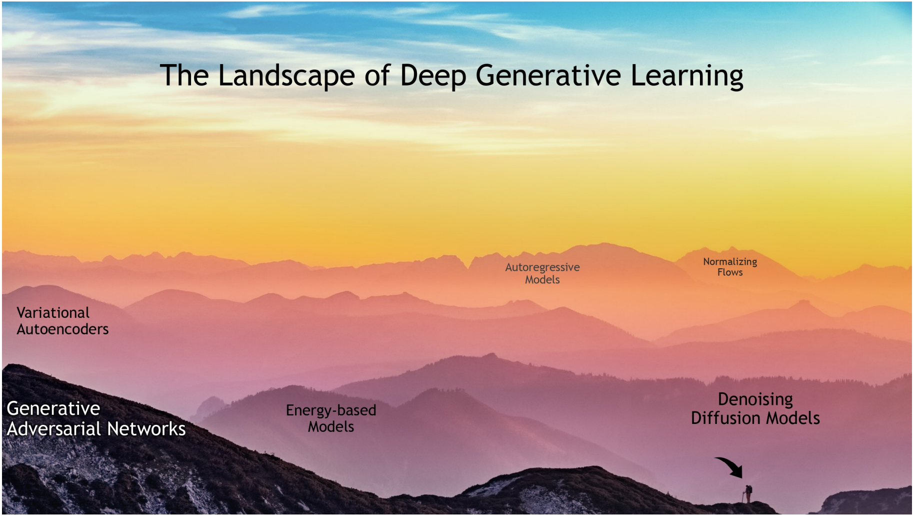
[1] [Wen] Lilian Weng. "What are diffusion models?"

Timeline of Deep Generative Models

Table Timeline of Deep Generative Models

1920s	Autoregressive models
2000s	Normalizing Flows
2003	Energy-Based Models
2006	Restricted Boltzmann Machines (RBM)
2013	Variational Autoencoders (VAEs)
2014	Generative Adversarial Networks (GANs)
2015	Diffusion Models
2020+	Diffusion Models gaining widespread recognition and impact

Generative Models Landscape



[2]

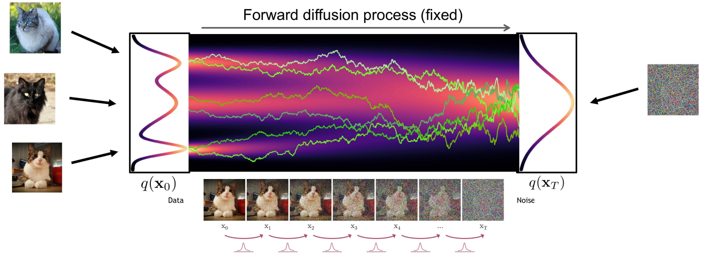
[2] [KGV] K. Kreis, R. Gao, and A. Vahdat. Denoising diffusion-based generative modeling: foundations and applications. CVPR Tutorial. 2022.

How Various DMs work

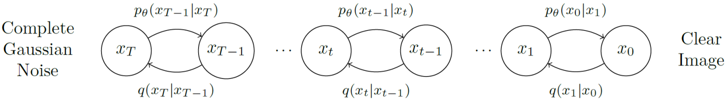
Inspired by non-equilibrium thermodynamics.
They learn by gradually adding noise to the input data and then reconstructing it.

- Denoising Diffusion Probabilistic Models (DDPMs):
  - function as a Markov chain with latent states (x_t).
- Score-Based Generative Models (SGMs):
  - follow the formalism of stochastic differential equations
  - directly learn the score function of the data
  - equivalent to DDPMs

Forward Diffusion Process as Stochastic Differential Equation



[3]



Forward Process: Gradual destruction of the images using Gausian kernels allows the model to learn complex patterns, from fine features to more generic ones.
Reverse Process: The inverse learned kernels provide a way to sample from the simple prior distribution and produce an image that almost lies on the initial data distribution.

[3] [KGV] K. Kreis, R. Gao, and A. Vahdat. Denoising diffusion-based generative modeling: foundations and applications. CVPR Tutorial. 2022.

Image In-painting Applications

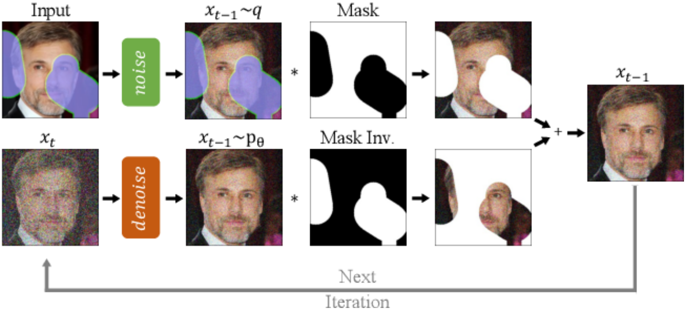
The process of creating a natural and cohesive image from a partially disclosed reference.

(a) Original images (b) Masked images (c) Inpainted images

RePaint Model

Based on DDPMs, consist of two main components:

- Conditioning on Known Region
  Merging known and unknown pixels guides the reconstruction of missing regions.
- Resampling
  Diffusing the output back to the input with a defined jump length.



[4]

[4] [LUGMAYR] Image source: Lugmayr et al., "RePaint: Inpainting using Denoising Diffusion Probabilistic Models," 2022, arXiv:2201.09865 [cs.CV]

CoPaint Model

Developed to address and overcome the incoherence challenges faced by In-painting models.

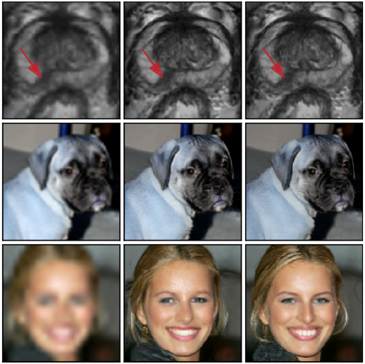
- An additional condition enforces a perfect match between the inpainted image and the revealed part of the reference.
- The "Time Travel" feature revisites previous denoising steps to improve internal consistency.

(a) Input (b) BLENDED (c) DDRM (d) COPAINT [5]

[5] [ZHANG] Zhang et al.. Towards Coherent Image Inpainting Using Denoising Diffusion Implicit Models. 2023. arXiv:2304.03322 [cs.CV]

Super-Resolution Applications

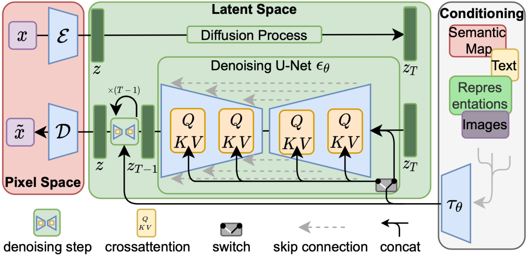
The task of enhancing the spatial resolution of an image, from lower-resolution observations.

 [6]

[6] [ZHAO] K. Zhao, A.L.Y. Hung, K. Pang, H. Zheng, K. Sung. PartDiff: Image Super-resolution with Partial Diffusion Models. 2023. arXiv:2307.11926 [eess.IV]

To address challenges with AutoRegressive Transformers i.e. complex architecture resulting to substantial energy consumption for rendering subtle image details.
Two-phase solution:

- Autoencoder learns an efficient, lower-dimensional latent space.
- DMs operate in the latent space.



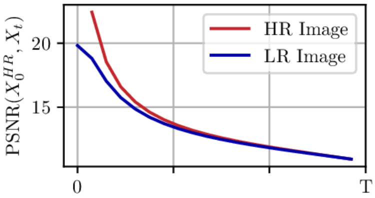
[7]

[7] [ROMBACH] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer. High-Resolution Image Synthesis with Latent Diffusion Models. 2022. arXiv:2112.10752 [cs.CV]

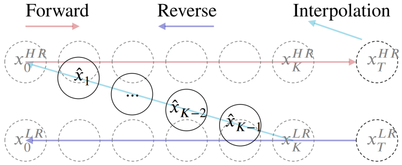
Partial Diffusion Models

To address challenges with computational demands i.e. too many diffusion steps resulting in excessive training times.

- It is assumed that latent states of low and high resolution images converge.
- Low-resolution image can serve as a proxy for its high-resolution counterpart thus, denoising is applied to a subset of steps.
- Latent alignment is applied to mitigate approximation errors



(a) PartDiff Approximation



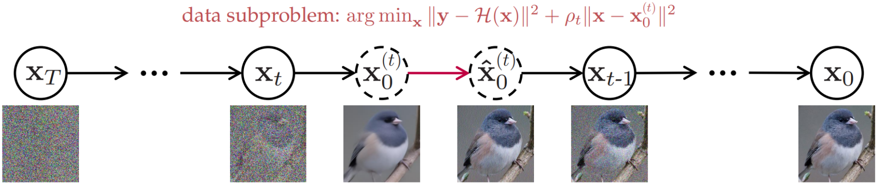
(b) PartDiff Alignment [8]

[8] [ZHAO] K. Zhao, A.L.Y. Hung, K. Pang, H. Zheng, K. Sung. PartDiff: Image Super-resolution with Partial Diffusion Models. 2023. arXiv:2307.11926 [eess.IV]

DiffPIR Model

To address the limitation of task-specific existing methods for image enhancing applications.

- Exploiting Bayes’ Theorem.
- Separating data and prior term from the conditional diffusion process.
- A pre-trained unconditional DM along with a classifier can solve the conditional problem.



[9]

[9] [ZHU] Y. Zhu, K. Zhang, J. Liang, J. Cao, B. Wen, R. Timofte, and L. Van Gool, "Denoising Diffusion Models for Plug-and-Play Image Restoration," 2023, arXiv:2305.08995 [cs.CV]

References I

- J. Sohl-Dickstein, E. A. Weiss, N. Maheswaranathan, and S. Ganguli, *Deep Unsupervised Learning using Nonequilibrium Thermodynamics*, arXiv preprint arXiv:1503.03585 (2015).
- J. Ho, A. Jain, and P. Abbeel, *Denoising Diffusion Probabilistic Models*, arXiv preprint arXiv:2006.11239 (2020).
- Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, *Score-Based Generative Modeling through Stochastic Differential Equations*, arXiv preprint arXiv:2011.13456 (2021).
- Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, Luc Van Gool, *RePaint: Inpainting using Denoising Diffusion Probabilistic Models*, arXiv preprint arXiv:2201.09865 (2022).
- G. Zhang, J. Ji, Y. Zhang, M. Yu, T. Jaakkola, and S. Chang, *Towards Coherent Image Inpainting Using Denoising Diffusion Implicit Models*, arXiv preprint arXiv:2304.03322 (2023).
- R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, *High-Resolution Image Synthesis with Latent Diffusion Models*, arXiv preprint arXiv:2112.10752 (2022).

References II

- K. Zhao, A. L. Y. Hung, K. Pang, H. Zheng, and K. Sung, *PartDiff: Image Super-resolution with Partial Diffusion Models*, arXiv preprint arXiv:2307.11926 (2023).
- Y. Zhu, K. Zhang, J. Liang, J. Cao, B. Wen, R. Timofte, and L. Van Gool, *Denoising Diffusion Models for Plug-and-Play Image Restoration*, arXiv preprint arXiv:2305.08995 (2023).
- L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, W. Zhang, B. Cui, and M.-H. Yang, *Diffusion Models: A Comprehensive Survey of Methods and Applications*, ACM Transactions on Computing, Vol. 1, No. 1, pp. 1-49, October 2023.

The End

Questions? Comments?