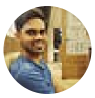
A Very Short Introduction to Diffusion Models



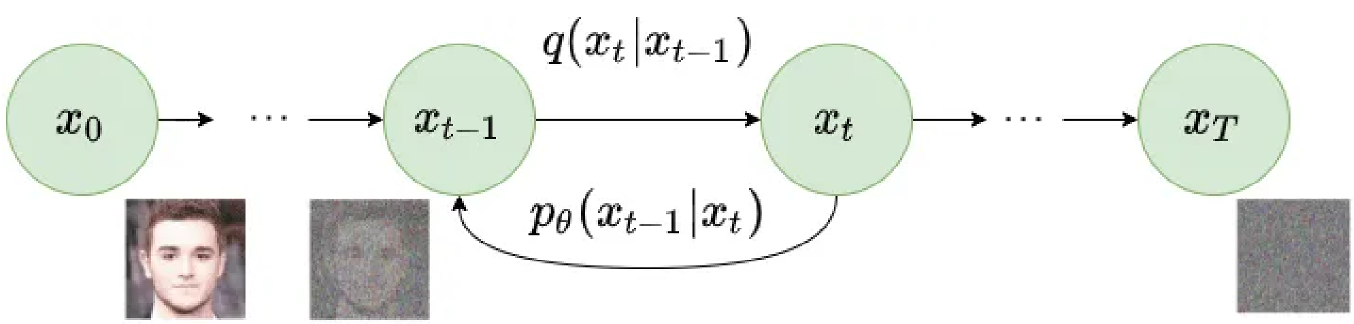
Kailash Ahirwar · Follow

4 min read · Sep 26, 2023

Artificial Intelligence is constantly evolving to solve hard and complex problems. Image generation is one such complex problem for AI models. GANs, VAEs, and Flow Models have been good but struggled to generate high-resolution images with high fidelity. On the other hand, D**iffusion odels** are very good at producing high-resolution images of diverse quality with great accuracy. Currently, they are at the forefront of the generative AI (GenAI) revolution that we see everywhere. Models like GLIDE, DALL.E-3 by OpenAI, Imagen by Google, and Stable Diffusion are some trendy diffusion models. Let's have a look at Diffusion Models.

What are Diffusion Models?

Diffusion models are a class of generative AI models that generate high-resolution images of varying quality. They work by gradually adding Gaussian noise to the original data in the **forward diffusion process** and then learning to remove the noise in the **reverse diffusion process**. They are latent variable models referring to a hidden continuous feature space, look similar to VAEs(Variational Autoencoders), and are loosely based on **non-equilibrium thermodynamics**.



Diffusion Models: Credits: https://theaisummer.com/diffusion-models/

Problems with the existing models

Existing deep learning models like GANs and VAEs are good at generating images but they struggle with some problems. GANs still struggle with training instability and generating diverse image problems, due to their adversarial training nature. The surrogate loss in VAEs creates issues.

Let's understand Diffusion Models in detail

A denoising diffusion modeling is a two step process:

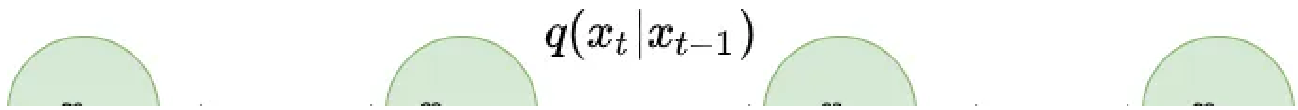
1. **Forward diffusion process** — The forward diffusion process is the Markov chain of diffusion steps in which we slowly and randomly add noise to the original data.
2. **Reverse diffusion process** — The reverse diffusion process tries to reverse the diffusion process to generate original data from the noise.

Forward diffusion process

In the forward diffusion process, we slowly and gradually add Gaussian noise to the input image x_0 through a series of T steps. We start with sampling a data point x_0 from the real data distribution $q(x)$ like ($x_0 \sim q(x)$) and then adding some Gaussian noise with variance β_t to x_{t-1}, producing a new latent variable x_t with distribution $q(x_t|x_{t-1})$

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \quad q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{t=1}^{T} q(\mathbf{x}_t|\mathbf{x}_{t-1})$$

Forward diffusion process. Credits: https://lilianweng.github.io/posts/2021-07-11-diffusion-models/



Forward diffusion process. Credits: https://theaisummer.com/diffusion-models/

Here, $q(x_t|x_{t-1})$ is defined by the mean μ as

$$\boldsymbol{\mu}_t = \sqrt{1-\beta_t}\mathbf{x}_{t-1}$$

and ∑ as $\Sigma_t = \beta_t \mathbf{I}$. FYI, I is the identity matrix and Σ will always be a diagonal matrix of variances. As T approaches ∞, x_{T} becomes isotropic Gaussian distribution.

The reparameterization trick

Applying $q(x_t | x_{t-1})$ and calculating x_t for an arbitrary time step can get very costly for a large number of steps. The reparameterization trick solves this problem and allows us to sample x_t at any arbitrary time step from the following distribution:

$$\mathbf{x}_t \sim q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I})$$

After parameterization trick

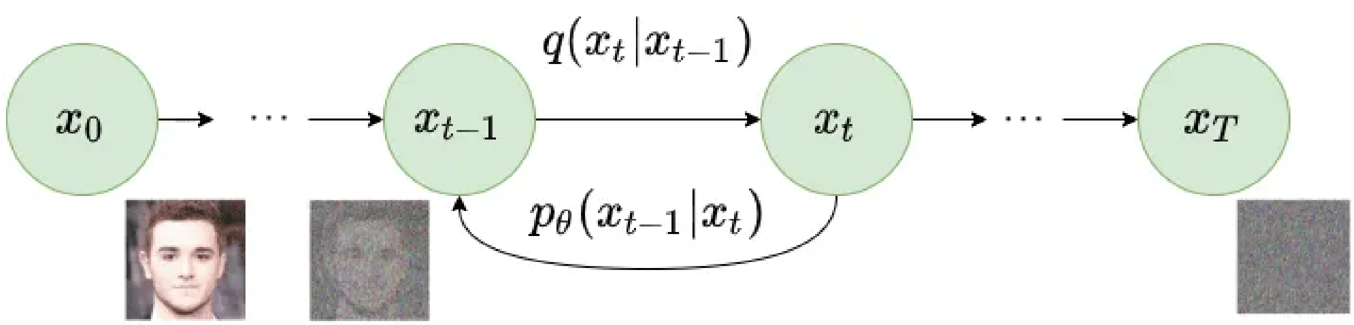
You can learn more about the reparameterization trick here.

Reverse diffusion process

It is the process of training a neural network to recover the original data by reversing the noising process applied in the forward pass. Estimating $q(x_{t-1}|x_t)$ is difficult as it can require the whole dataset. That's why a parameterized model p_θ(Neural Network) can be used to learn the parameters. For small enough β_t, it will be a Gaussian and can be obtained by just parameterizing the mean and variance.

$$p_\theta(\mathbf{x}_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) \quad p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

Reverse diffusion process. Credits: https://lilianweng.github.io/posts/2021-07-11-diffusion-models/



Reverse diffusion process. Credits: https://theaisummer.com/diffusion-models/

We train the network to predict the mean and variance for each time step. Here μ_θ(x_t,t) is the mean, and ∑_θ(x_t,t) is the covariance matrix.

Top Diffusion Models

1. Diffusion probabilistic models (DPM; Sohl-Dickstein et al., 2015)
2. Denoising diffusion probabilistic models (DDPM; Ho et al. 2020)
3. Cascading Diffusion Models (CDM; https://cascaded-diffusion.github.io)

4. Latent Diffusion Models (LDM; https://arxiv.org/abs/2112.10752)

This was a very short introduction to Diffusion Models. I wanted to give a basic understanding of Diffusion Models and how they work. If you want to understand Diffusion Models in detail and learn the mathematics behind them, read the following articles:

1. What are Diffusion Models? — https://lilianweng.github.io/posts/2021-07-11-diffusion-models
2. How diffusion models work: the math from scratch — https://theaisummer.com/diffusion-models
3. Introduction to Diffusion Models for Machine Learning — https://www.assemblyai.com/blog/diffusion-models-for-machine-learning-introduction
4. Diffusion Models Made Easy — https://towardsdatascience.com/diffusion-models-made-easy-8414298ce4da
5. Diffusion Models: A Comprehensive Survey of Methods and Applications — https://arxiv.org/pdf/2209.00796.pdf

Artificial Intelligence | Deep Learning | Computer Science | Computer Vision | Diffusion Models

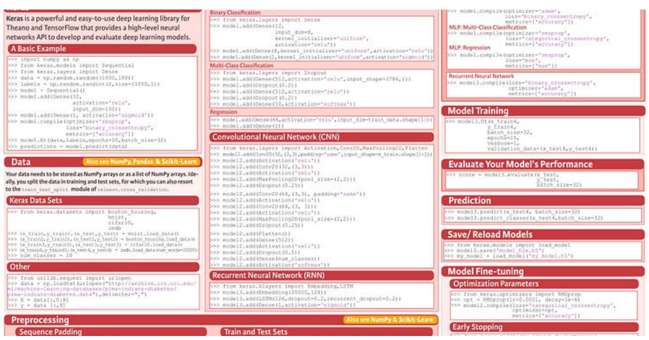
Written by Kailash Ahirwar

Follow

5.9K Followers

Artificial Intelligence Research | Author - Generative Adversarial Networks Projects | Co-founder - Raven Protocol | Mate Labs

More from Kailash Ahirwar



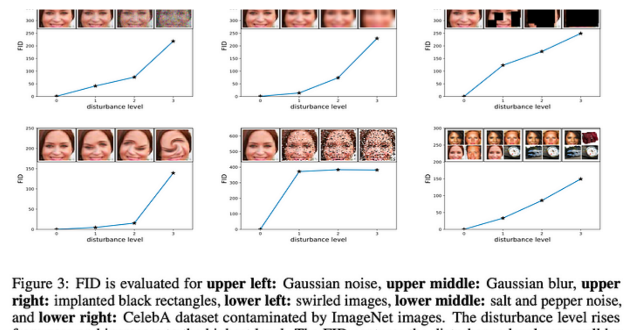
 Kailash Ahirwar in Startups & Venture Capital

Essential Cheat Sheets for Machine Learning and Deep Learning...

Machine learning is complex. For newbies, starting to learn machine learning can be...

4 min read · May 28, 2017

25K 107



 Kailash Ahirwar in DataDrivenInvestor

A Very Short Introduction to Frechlet Inception Distance(FID)

Generative Adversarial Networks(GANs) are very difficult to evaluate as compared to oth...

4 min read · Sep 4, 2020

152

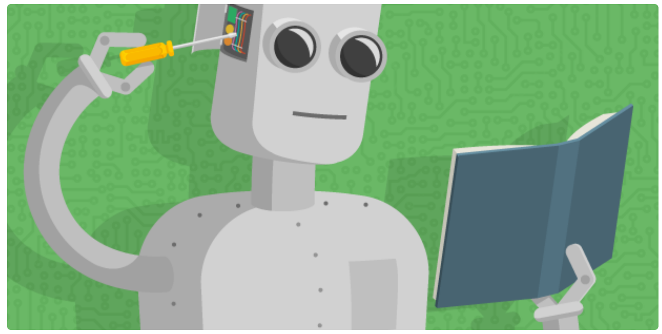
Kailash Ahirwar

A Very Short Introduction to Inception Score(IS)

Generative Adversarial Networks or GANs for short were successful in generating high-...

4 min read · Feb 24, 2021

86



Kailash Ahirwar in Towards Data Science

Why we need a better learning algorithm than Backpropagation i...

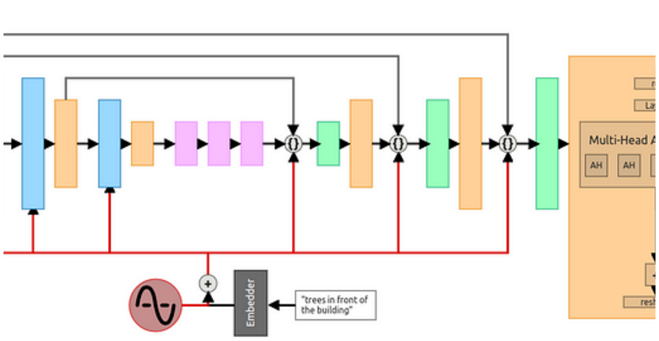
We all agree on one thing that Backpropagation is a revolutionary learning...

3 min read · Jan 30, 2018

859 13

See all from Kailash Ahirwar

Recommended from Medium



Kemal Erdem (burnpiro)

Step by Step visual introduction to Diffusion Models.

How the diffusion models works under the hood? Visual guide to diffusion process and...

15 min read · Nov 9, 2023

205 1

Jonathan Kernes in Towards Data Science

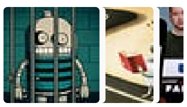
Diffusion Models

What are they, how do they work, and why now?

23 min read · Dec 13, 2022

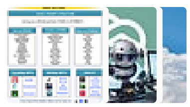
343 4

Lists



AI Regulation

6 stories · 323 saves

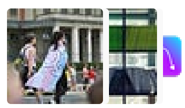


ChatGPT

21 stories · 474 saves

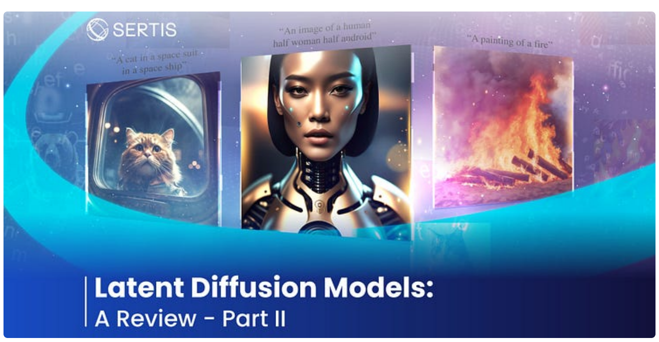
Natural Language Processing

1215 stories · 690 saves



Generative AI Recommended Reading

52 stories · 739 saves

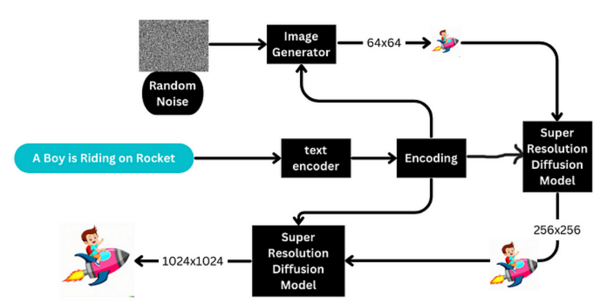


Sertis

Latent Diffusion Models: A Review —Part II

This blog-post series aims to explain the mechanism of Latent Diffusion Models...

8 min read · Oct 30, 2023

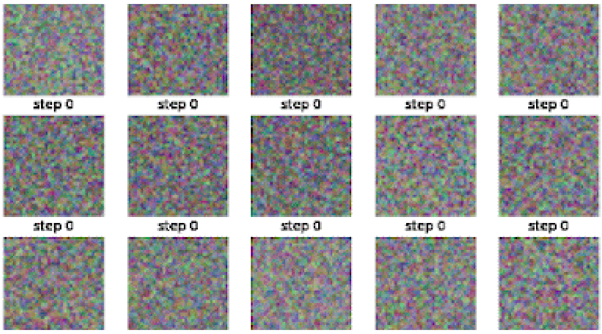


Shyam Patel

Introduction to Diffusion Models and IMAGEN: The Magic Behind...

Introduction to Diffusion Model

4 min read · Sep 26, 2023

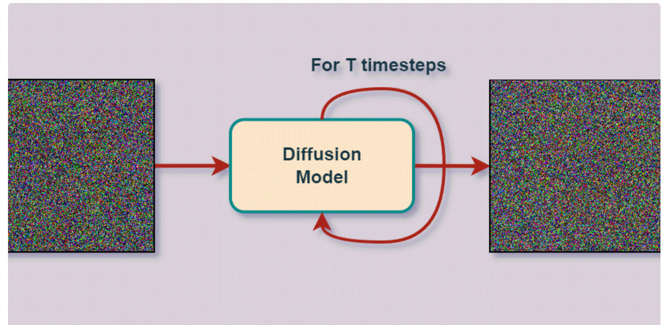


Abdulkader Helwan

Diffusion models from scratch

Diffusion models are a type of generative deep learning model that can generate new...

· 4 min read · Oct 4, 2023



Gabriel Mongaras in Better Programming

Diffusion Models — DDPMs, DDIMs, and Classifier Free Guidance

A guide to the evolution of diffusion models from DDPMs to Classifier Free guidance

28 min read · Mar 13, 2023

See more recommendations