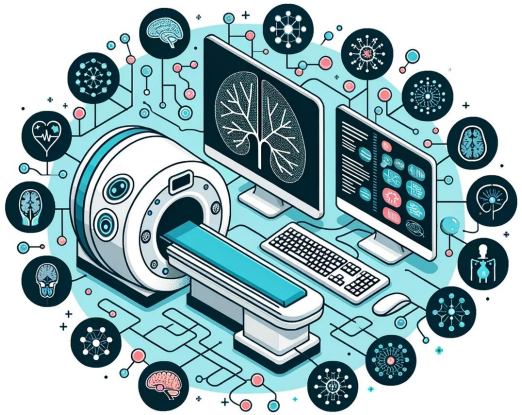
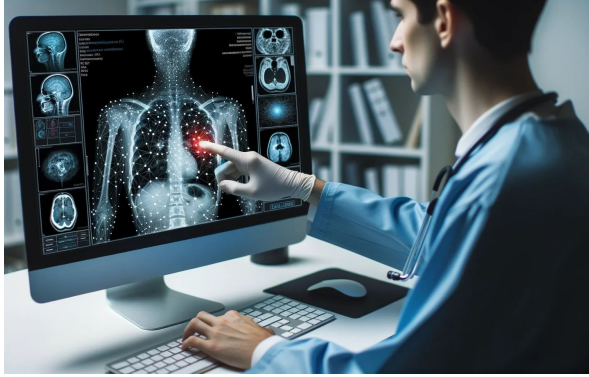
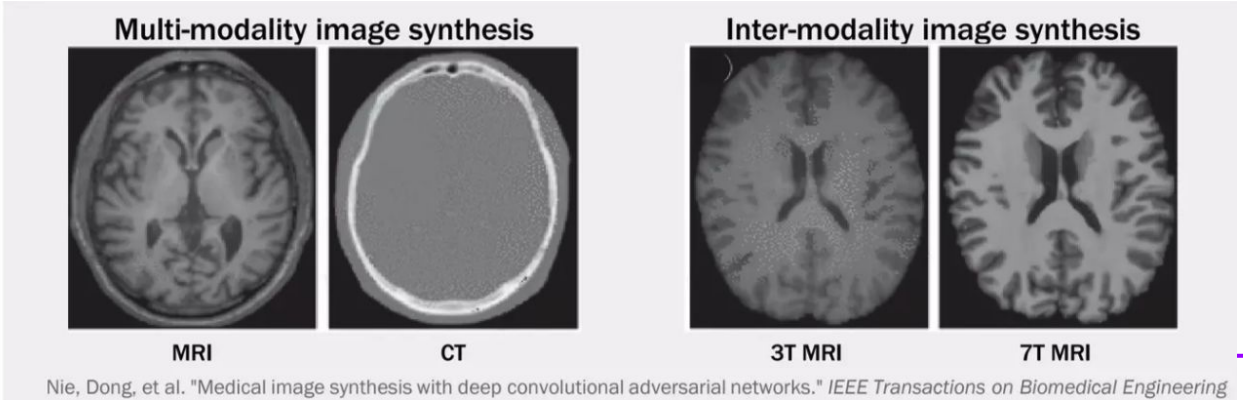




Deep Generative Learning
Session 05
Oct 04 2025

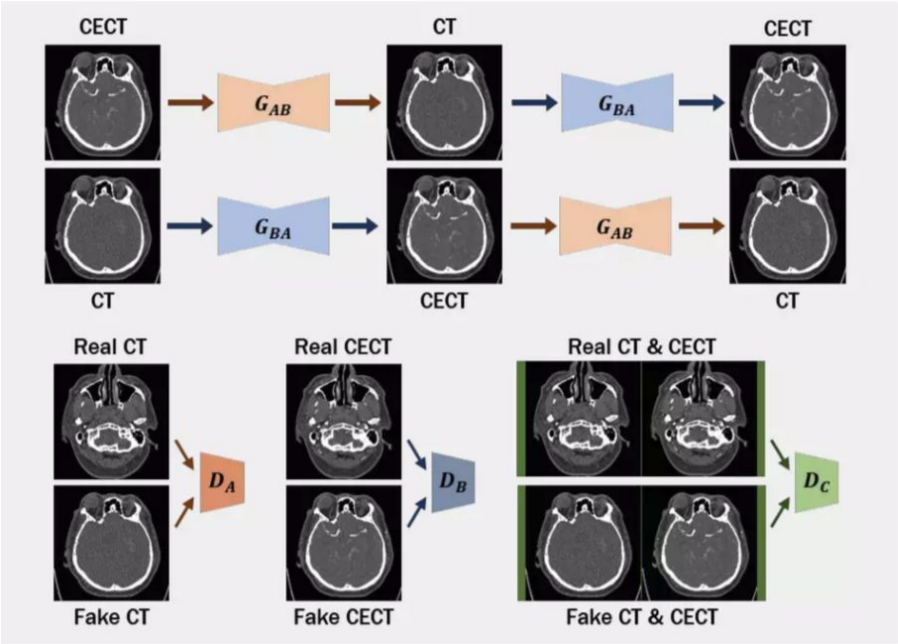
Medical image synthesis

- An approach to modeling a mapping from given images to unknown images
- Necessity
  - Potential risk of radiation exposure for multiple acquisition of medical images (ec. CT, PET)
  - Not always accessible modalities for every patient
  - Alignment issue on the analysis of multi-images



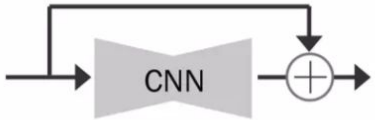
Nie, Dong, et al. "Medical image synthesis with deep convolutional adversarial networks." *IEEE Transactions on Biomedical Engineering*

CycleGAN framework



- **Residual learning** for generators

  Prevent loss of medical information



- **Two Generators**

  G_{AB} : CECT → CT (generate synthetic CT)

  G_{BA} : CT → CECT (generate synthetic CECT)

- **Three Discriminators**

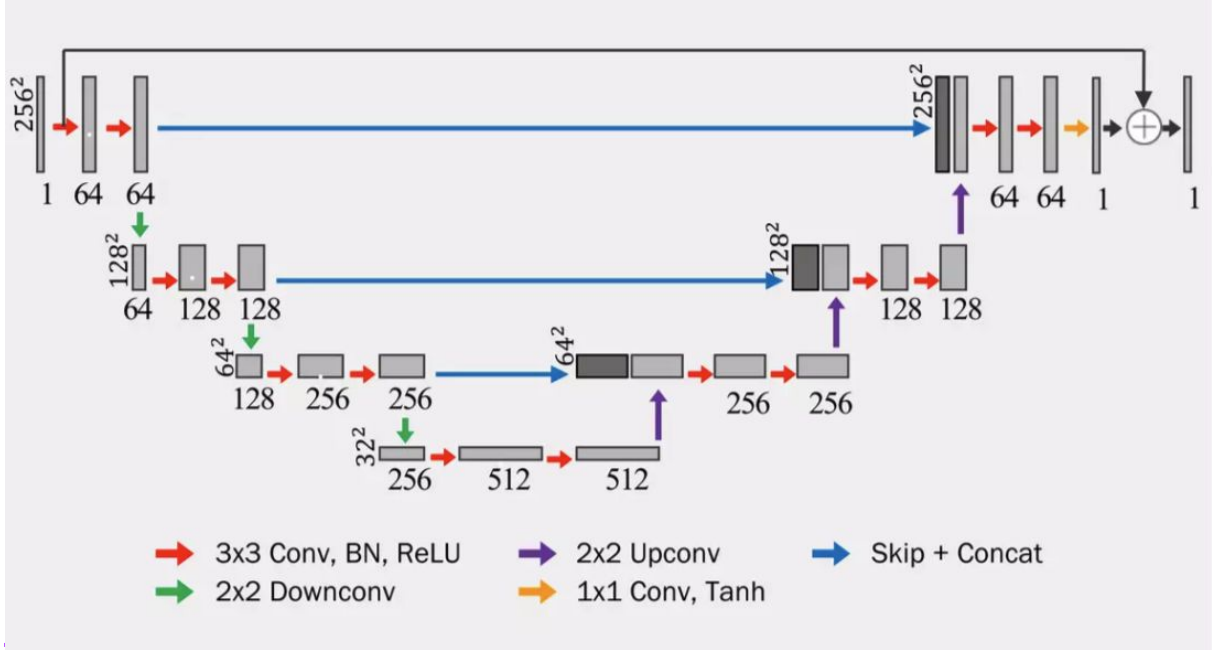
  D_A : Distinguish real and fake CT

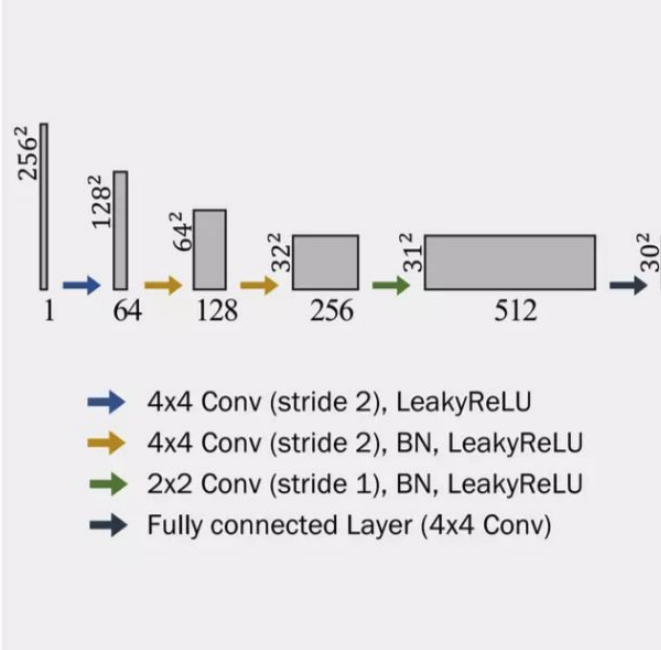
  D_B : Distinguish real and fake CECT

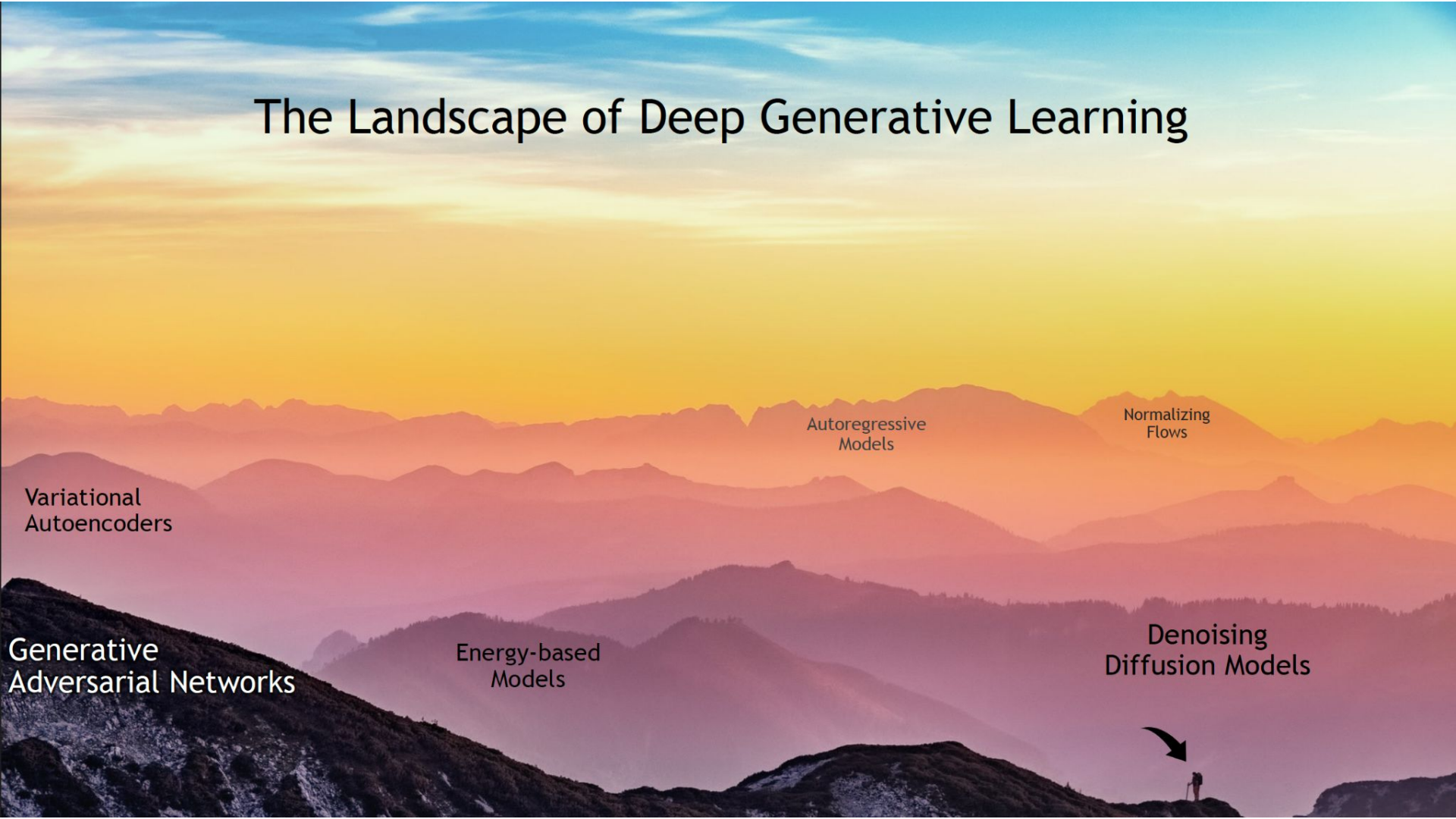
  $\boldsymbol{D_C}$: Distinguish real and fake pair of CT & CECT

Network architecture

- Generator (UNet)
- Discriminator






The Landscape of Deep Generative Learning
Autoregressive Models
Normalizing Flows
Variational Autoencoders
Generative Adversarial Networks
Energy-based Models
Denoising Diffusion Models

Denoising Diffusion Models



"Diffusion Models Beat GANs on Image Synthesis"
Dhariwal & Nichol, OpenAI, 2021



"Cascaded Diffusion Models for High Fidelity Image Generation"
Ho et al., Google, 2021

Text-to-Image Generation

DALL·E 2

"a teddy bear on a skateboard in times square"



"Hierarchical Text-Conditional Image Generation with CLIP Latents" Ramesh et al., 2022

Imagen

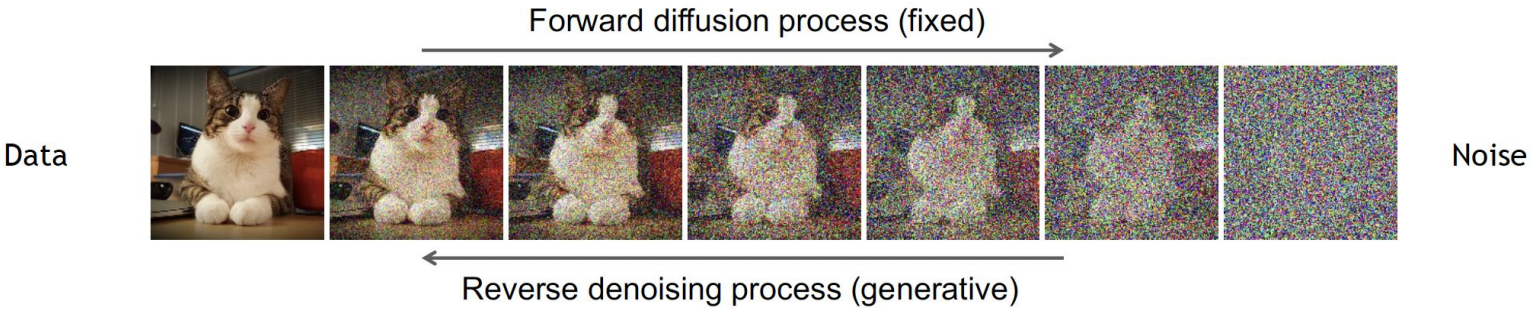
A group of teddy bears in suit in a corporate office celebrating the birthday of their friend. There is a pizza cake on the desk.



"Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding", Saharia et al., 2022

Learning to Generate by Denoising

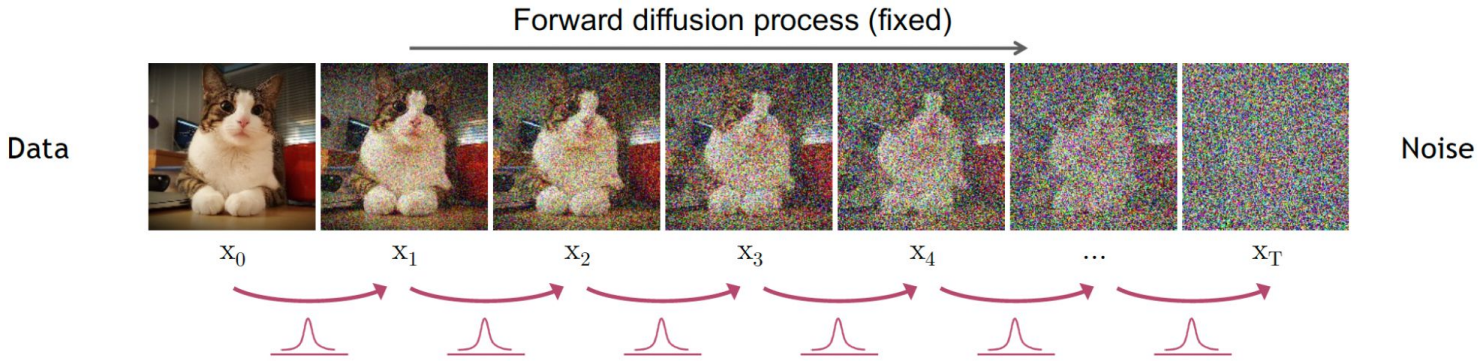
- Denoising diffusion consists of two processes:
  - Forward diffusion to gradually add noise to input
  - Reverse denoising that learns to generate data by denoising



Sohl-Dickstein et al., Deep Unsupervised Learning using Nonequilibrium Thermodynamics, ICML 2015
Ho et al., Denoising Diffusion Probabilistic Models, NeurIPS 2020
Song et al., Score-Based Generative Modeling through Stochastic Differential Equations, ICLR 2021

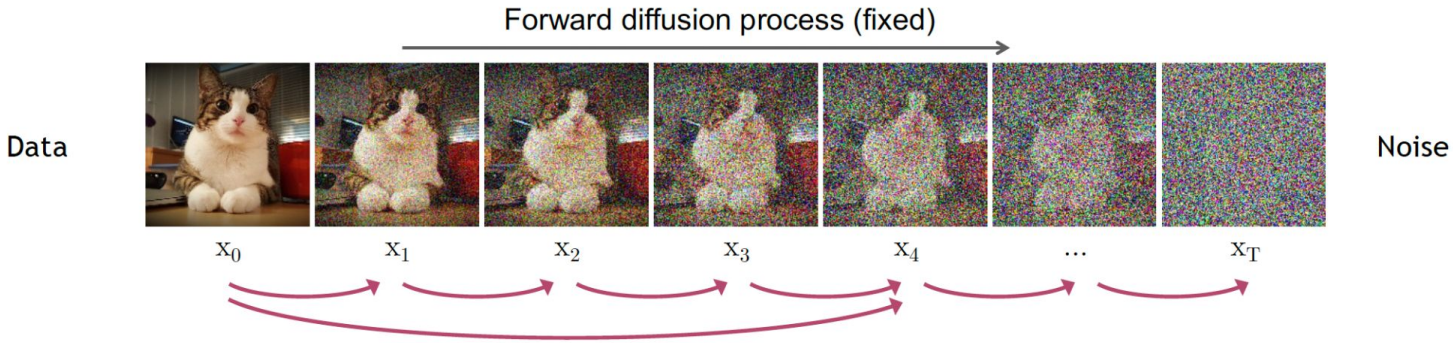
Forward Diffusion Process

- The formal definition of the forward process in T steps:



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

Diffusion Kernel



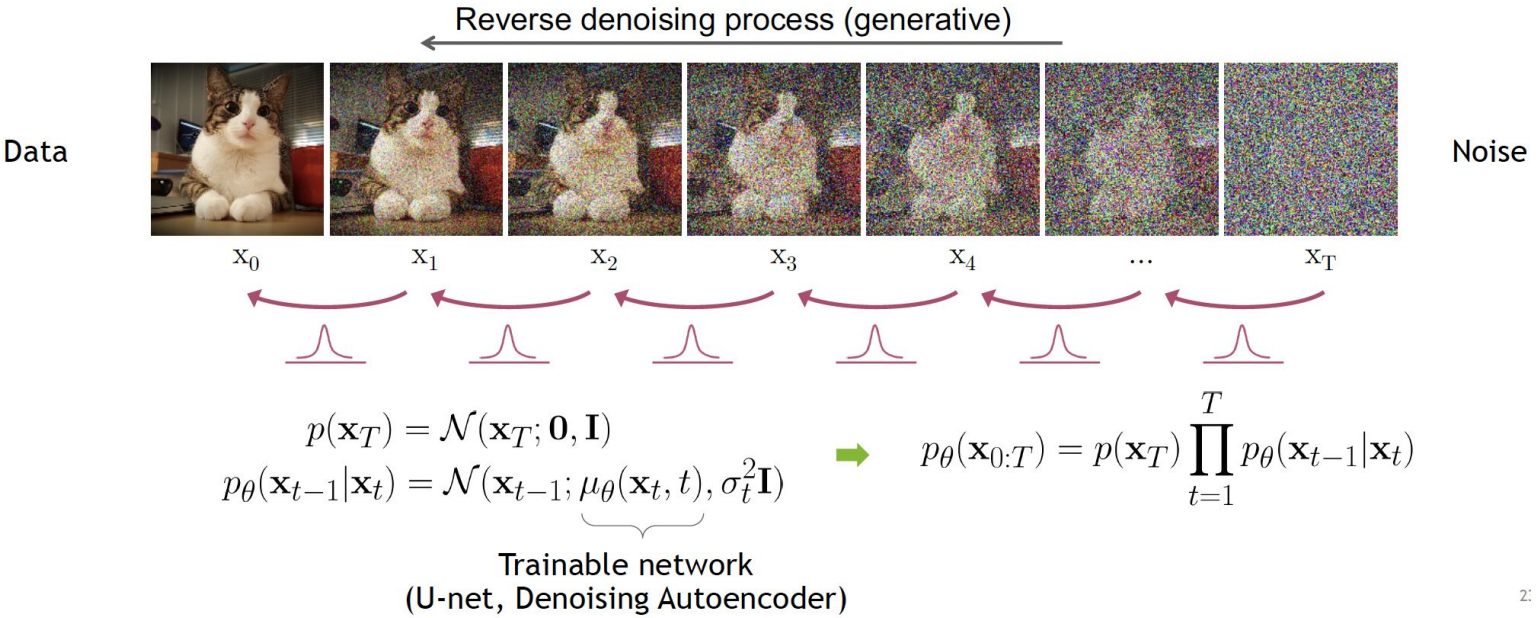
Define $\bar{\alpha}_t = \prod_{s=1}^{t}(1-\beta_s)$ ➡ $q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}))$ (Diffusion Kernel)

For sampling: $\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{(1-\bar{\alpha}_t)}\,\epsilon$ where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

β_t values schedule (i.e., the noise schedule) is designed such that $\bar{\alpha}_T \to 0$ and $q(\mathbf{x}_T|\mathbf{x}_0) \approx \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I}))$

Reverse Denoising Process

- Formal definition of reverse process in T steps:



Training and Sampling Algorithms

Algorithm 1 Training

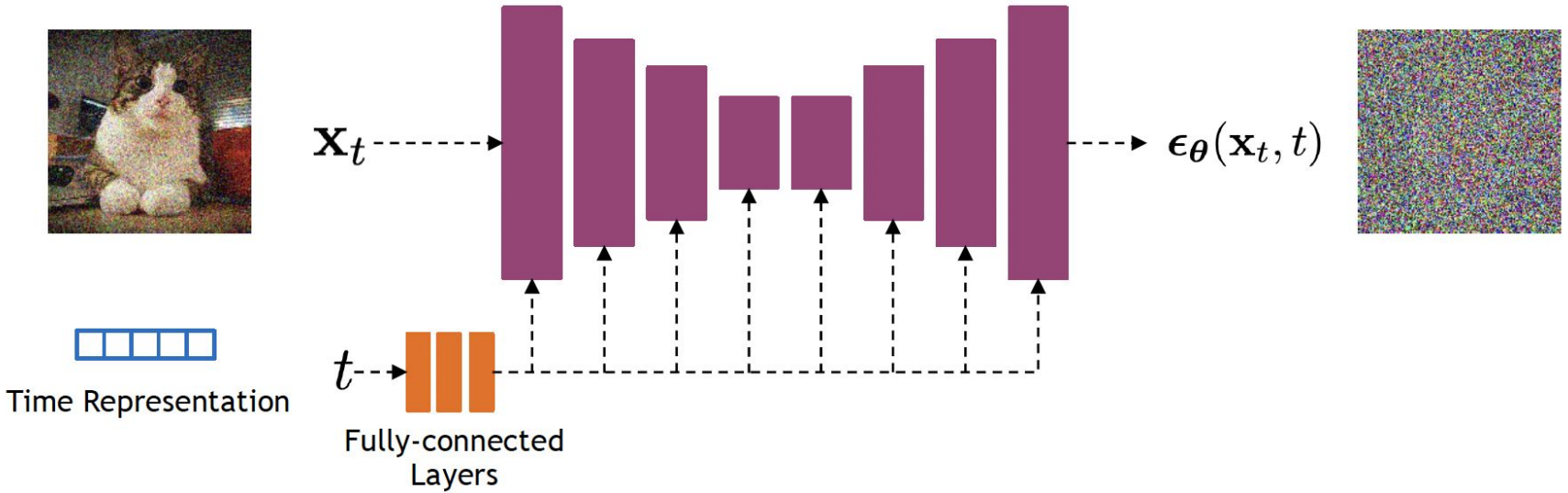
1: **repeat**
2: $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
3: $t \sim \text{Uniform}(\{1, \dots, T\})$
4: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
5: Take gradient descent step on
$$\nabla_\theta \left\| \boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\sqrt{\bar{\alpha}_t}\mathbf{x}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t) \right\|^2$$
6: **until** converged

Algorithm 2 Sampling

1: $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
2: **for** $t = T, \dots, 1$ **do**
3: $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
4: $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}$
5: **end for**
6: **return** $\mathbf{x}_0$

Network Architectures

- Diffusion models often use U-Net architectures with ResNet blocks and self-attention layers



Time representation: sinusoidal positional embeddings or random Fourier features.

Diffusion Parameters

- The noise schedule

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

Data  Noise

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \sigma_t^2\mathbf{I})$$

Above, β_t and σ_t^2 control the variance of the forward diffusion and reverse denoising processes respectively.

Often a linear schedule is used for β_t, and σ_t^2 is set equal to β_t.

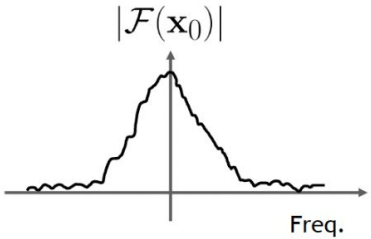
Fourier Analysis of the Forward Process

Recall that sampling from $q(\mathbf{x}_t|\mathbf{x}_0)$ is done using $\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{(1-\bar{\alpha}_t)}\,\epsilon$ where $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

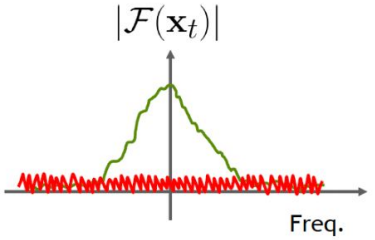
$$\mathbf{x}_t = \sqrt{\bar{\alpha}_t}\,\mathbf{x}_0 + \sqrt{(1-\bar{\alpha}_t)}\,\epsilon$$

Fourier Transform

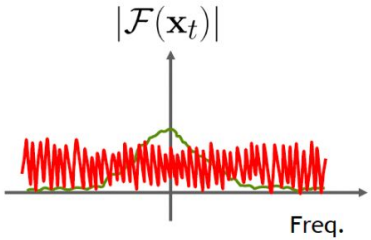
$$\mathcal{F}(\mathbf{x}_t) = \sqrt{\bar{\alpha}_t}\,\mathcal{F}(\mathbf{x}_0) + \sqrt{(1-\bar{\alpha}_t)}\,\mathcal{F}(\epsilon)$$



Small t
$\bar{\alpha}_t \sim 1$



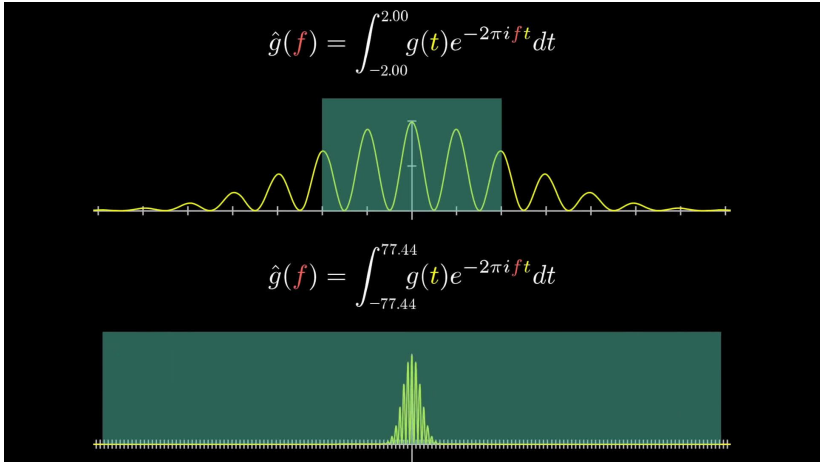
Large t
$\bar{\alpha}_t \sim 0$

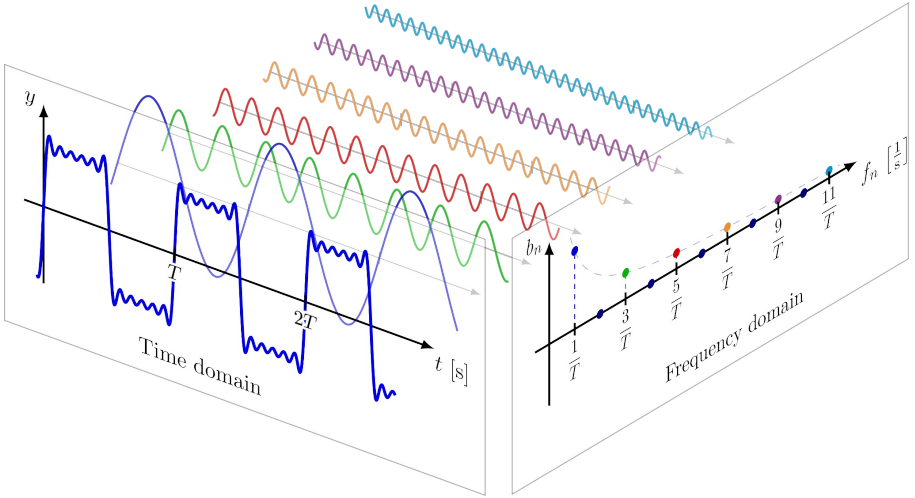


In the forward diffusion, the high frequency content is perturbed faster.

Fourier Analysis

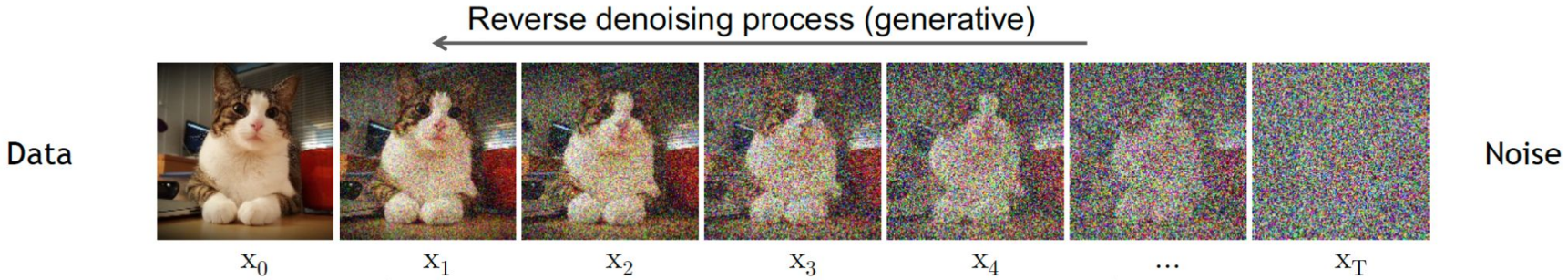
- Decomposition of signal into frequency components



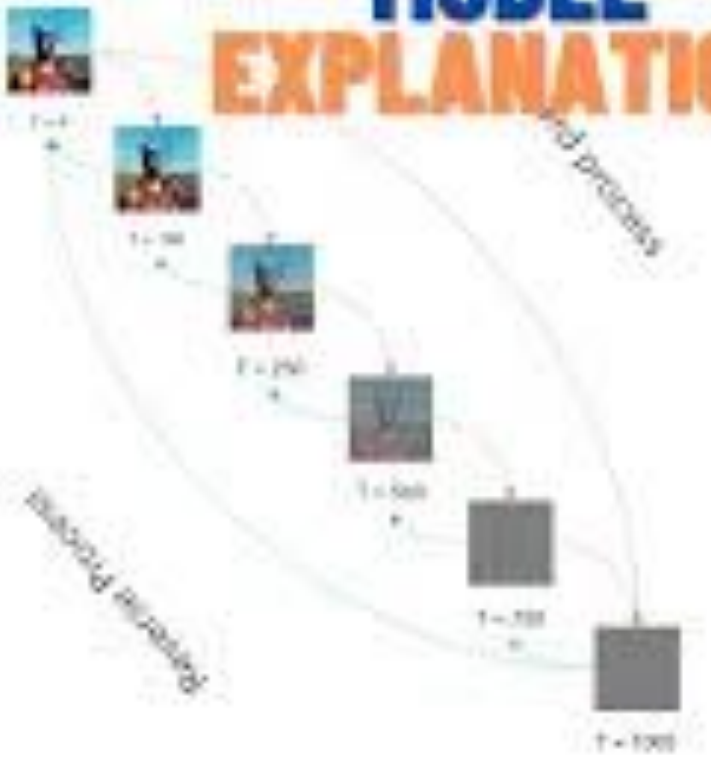


https://dibsmethodsmeetings.github.io/fourier-transforms/

Content-Detail Tradeoff



The weighting of the training objective for different timesteps is important!





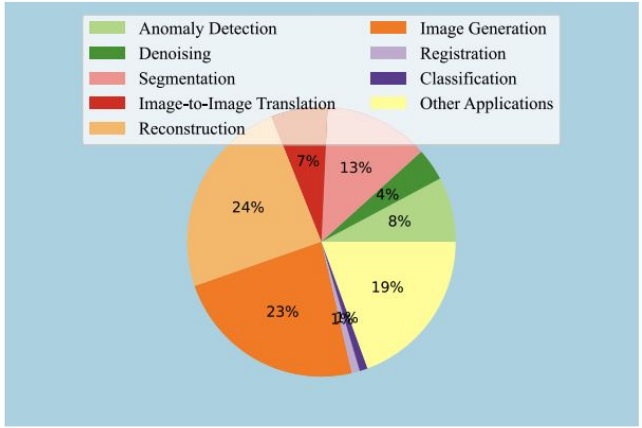
Denoising Diffusion Probabilistic Models (DDPM)

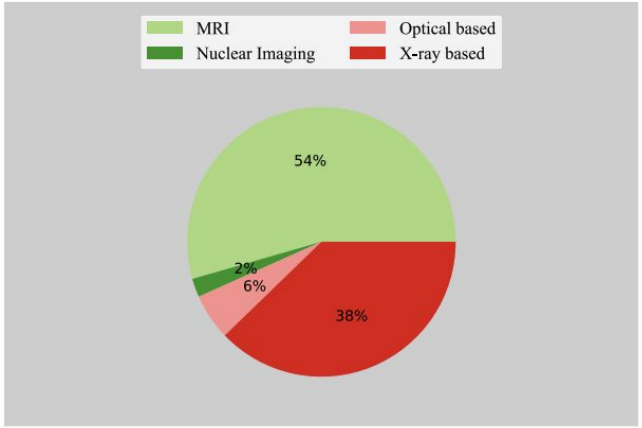
Umar Jamil

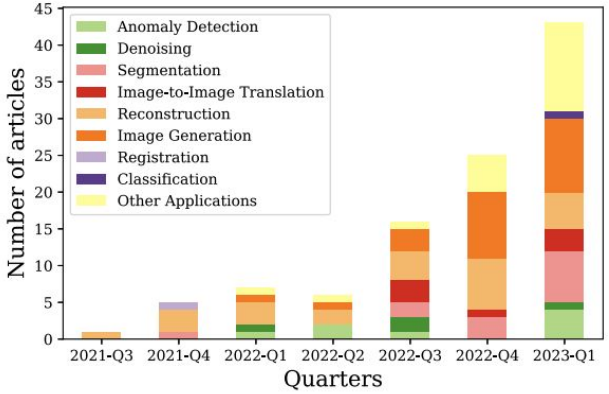
FREE TRAINING CODE

Diffusion Models in Medical Imaging

- Diffusion models have emerged as powerful generative models in medical imaging
  - improved sample quality, mode coverage, and versatility across various applications.

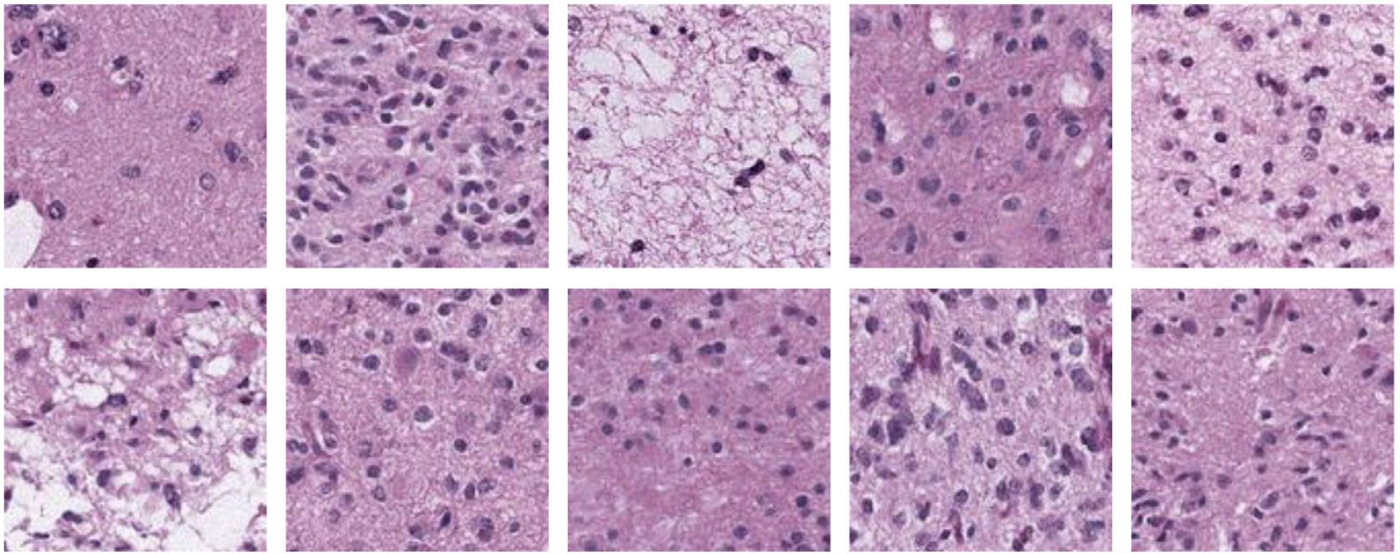






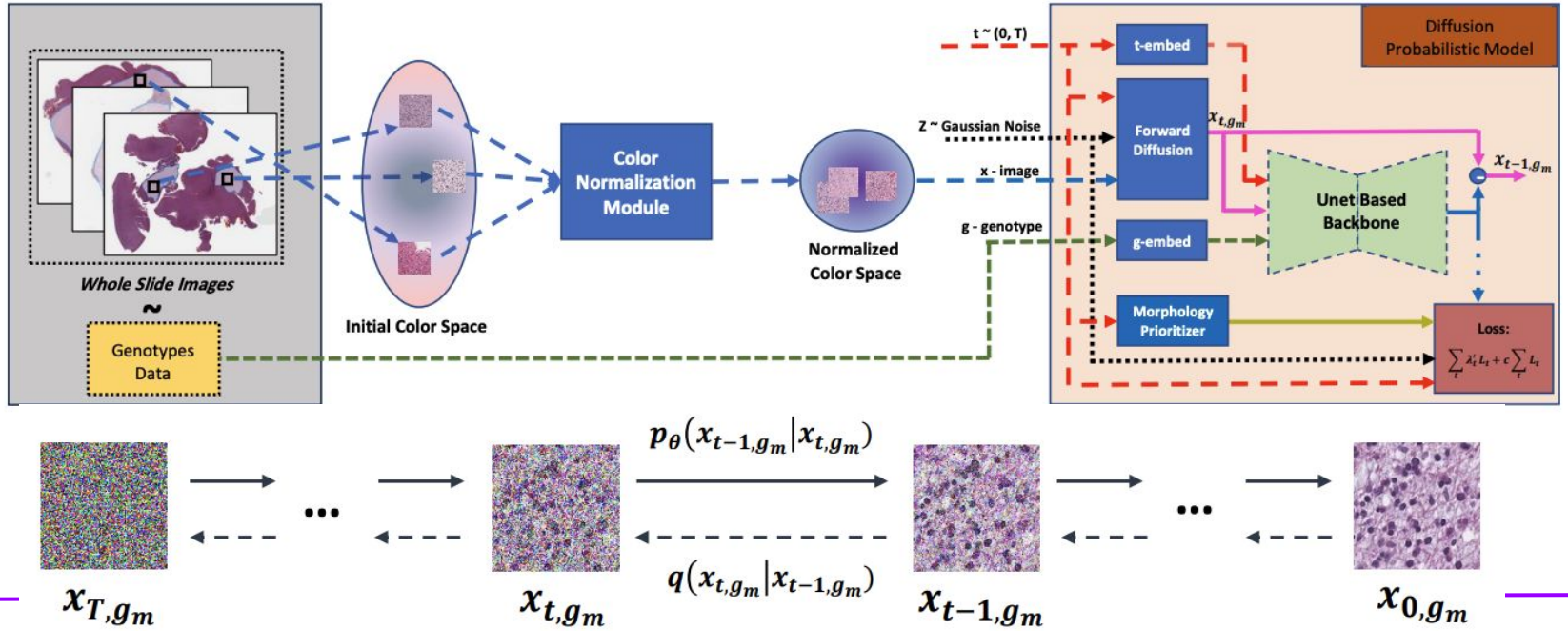
Applications in medical image generation

- Diffusion models have remarkable performance in generating synthetic medical images
  - aiding data augmentation and rare disease representation



Applications in medical image generation

- Generating histopathology images with genotype guidance



Genotype-conditioned Image Generation

How Are IDH Mutation and 1p/19q Co-Deletion Extracted for Training Data?

The training data for the diffusion model includes paired histopathology images and genotype information. The genotype labels (IDH mutation status and 1p/19q co-deletion) are extracted through molecular and genetic tests, including:

1. IDH Mutation Detection

- Immunohistochemistry (IHC): Uses antibodies to detect mutant IDH1 protein in tumor samples.
- Next-Generation Sequencing (NGS): Directly sequences the IDH1/IDH2 genes to detect mutations.

2. 1p/19q Co-Deletion Detection

- Fluorescence In Situ Hybridization (FISH): A cytogenetic test that detects the loss of chromosome arms 1p and 19q in tumor cells.
- Comparative Genomic Hybridization (CGH): Identifies chromosomal deletions.
- PCR-based methods: Detects loss of heterozygosity (LOH) in 1p and 19q regions.

Once the genetic data is obtained, it is paired with corresponding histopathology images to create a genotype-labeled dataset for training.

Genotype-conditioned Image Generation

- What is an IDH Mutation?
  - IDH (Isocitrate Dehydrogenase) Mutation refers to genetic alterations in the IDH1 or IDH2 genes.
  - These mutations are commonly found in gliomas (brain tumors) and are important for tumor classification, prognosis, and treatment decisions.
  - IDH-mutant gliomas tend to have better survival rates compared to IDH-wildtype gliomas, which are more aggressive.
- What is 1p/19q Co-Deletion?
  - The 1p/19q co-deletion is a chromosomal alteration where parts of chromosomes 1p (short arm of chromosome 1) and 19q (long arm of chromosome 19) are missing.
  - This is a key molecular marker used to classify gliomas.
  - Gliomas with 1p/19q co-deletion are almost always oligodendrogliomas, which respond well to chemotherapy and radiation therapy.
  - If a glioma has IDH mutation but no 1p/19q co-deletion, it is classified as an astrocytoma instead of an oligodendroglioma.

Inception score

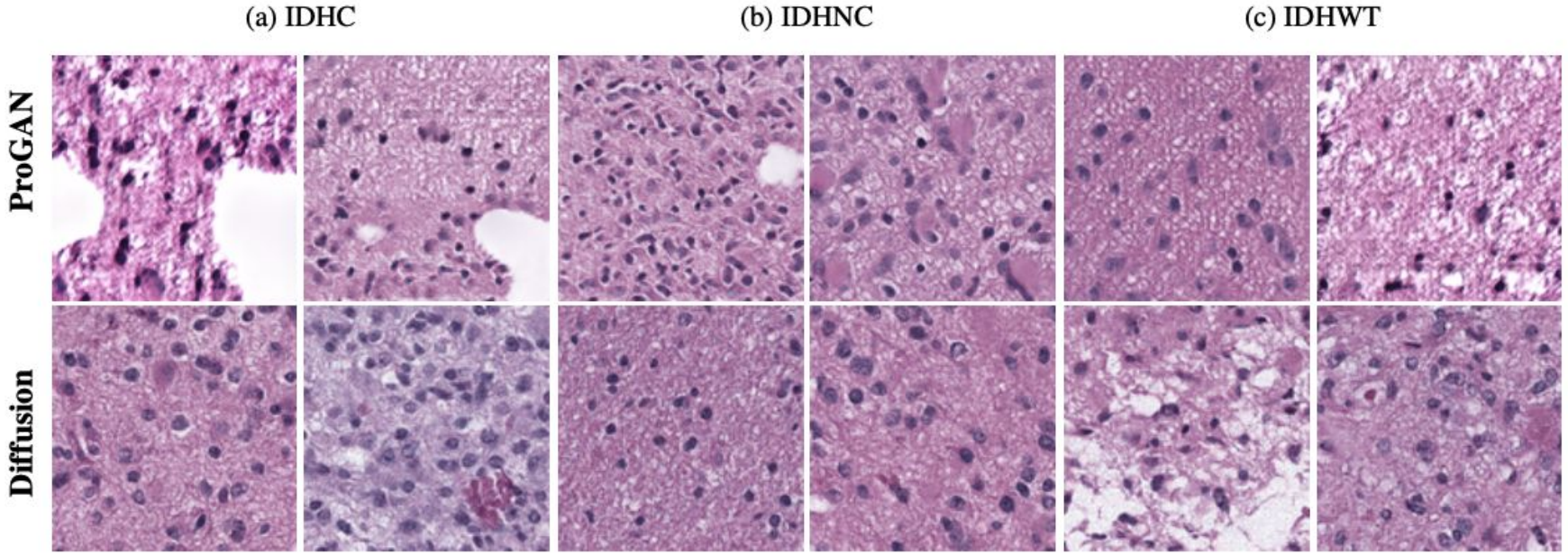
What it measures: The quality and diversity of generated images.

How it works:

- Uses a pre-trained Inception network (e.g., Inception v3) to classify generated images.
- Measures:
  - **Quality**: If a generated image is highly classifiable (i.e., strong class predictions).
  - **Diversity**: If the generated images cover multiple classes.
- Requires a well-trained classifier on medical data, unlike natural images (e.g., CIFAR, ImageNet).



Applications in medical image generation



	ProGAN	Diffusion Model
Improved Recall	0.4816	**0.8528**
Improved Precision	0.0078	**0.2573**

	ProGAN	Diffusion Model
Inception Score	1.67	**2.08**
FID	53.85	**20.11**
sFID	24.37	**6.32**

MedAI Group

Session #96: Denoising Diffusion Models for Medical Image Analysis

Julia Wolleb

