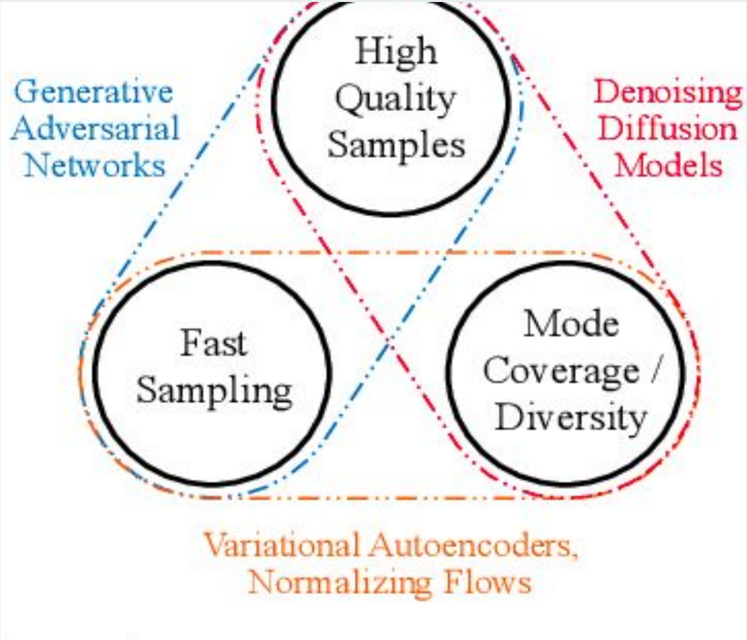
Exploring techniques to speed up the generation of high-dimensional objects in Diffusion Model

Team Members :

Meghana 200050135
Namitha 200050062
Ruthvika 200050018
Sindhuja 200050047



Tackle the trilemma by accelerating diffusion models

Variational diffusion models

https://openreview.net/pdf?id=2LdBqxc1Yv

Parametrization of Noise Schedule

SOTA likelihood

Denoising Diffusion GANs

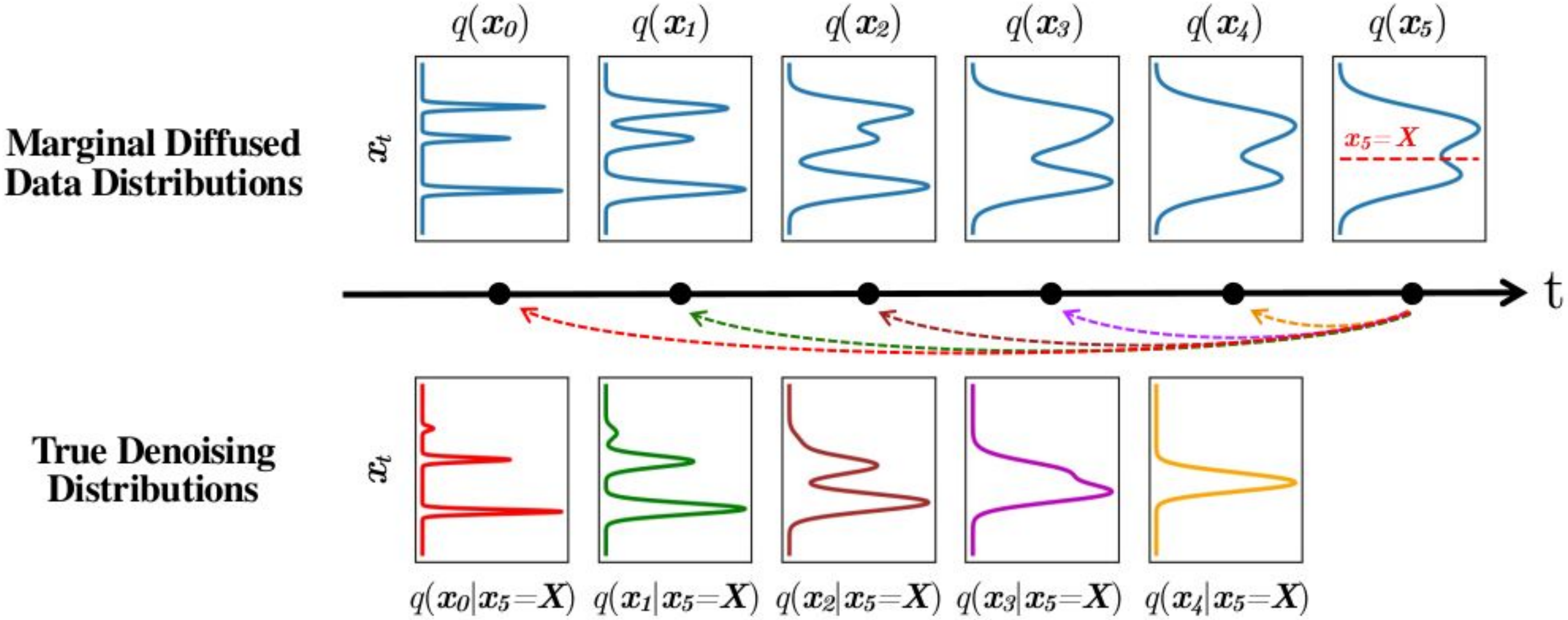
Diffusion models are slow due to modelling the reverse process as a Guassian distribution which only holds for small β_t requiring a large number of steps . DDGANs model the reverse process as a multi-modal distribution to produce samples in a small number of steps

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

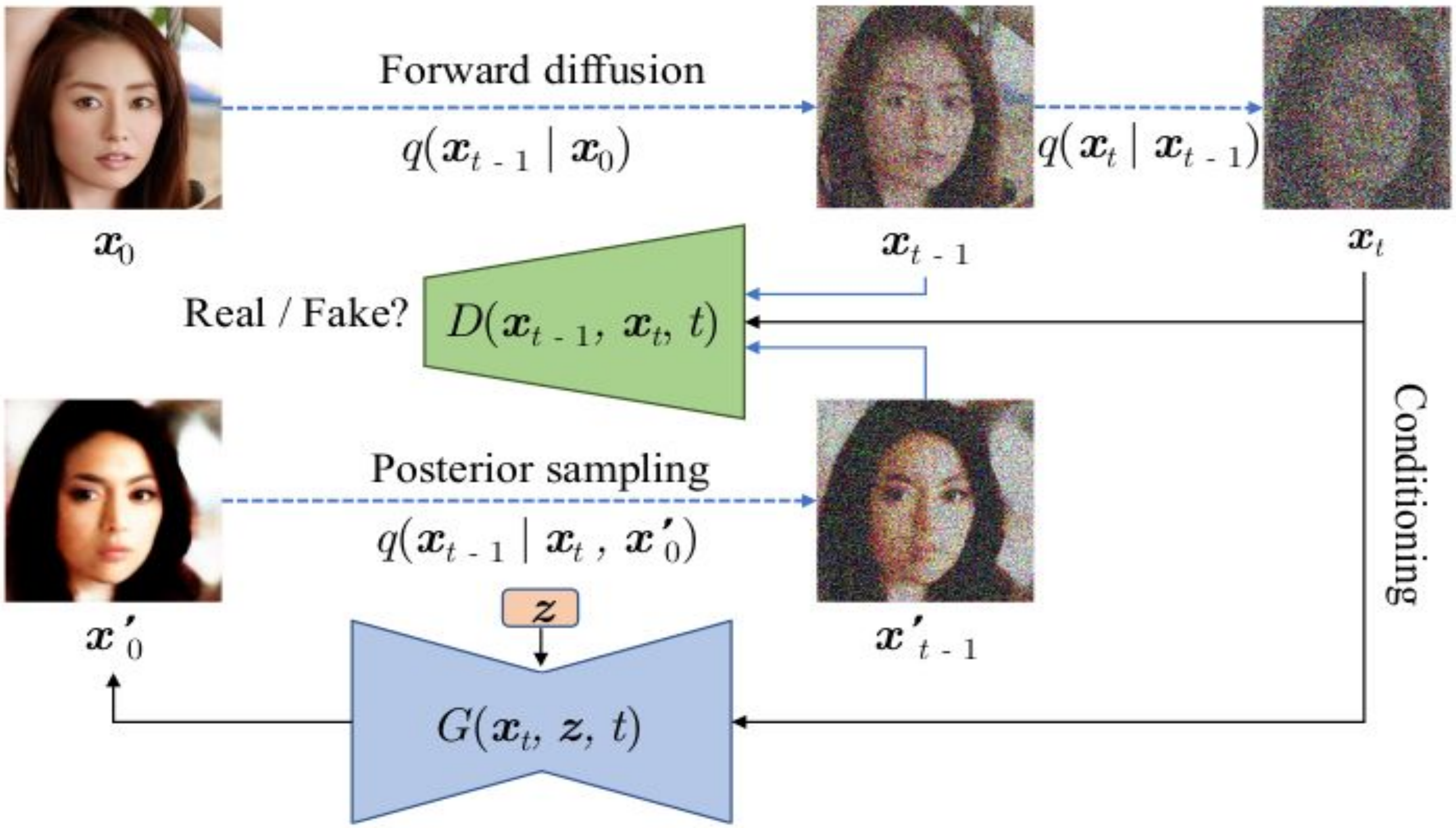
$$\min_{\phi} \sum_{t\geq 1} \mathbb{E}_{q(\mathbf{x}_t)} \left[\mathbb{E}_{q(\mathbf{x}_{t-1}|\mathbf{x}_t)}[-\log(D_\phi(\mathbf{x}_{t-1}, \mathbf{x}_t, t)] + \mathbb{E}_{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}[-\log(1-D_\phi(\mathbf{x}_{t-1}, \mathbf{x}_t, t))]\right]$$

$$\max_\theta \textstyle\sum_{t\geq 1} \mathbb{E}_{q(\mathbf{x}_t)}\mathbb{E}_{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}[\log(D_\phi(\mathbf{x}_{t-1}, \mathbf{x}_t, t))],$$

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := \int p_\theta(\mathbf{x}_0|\mathbf{x}_t) q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) d\mathbf{x}_0 = \int p(\mathbf{z}) q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0 = G_\theta(\mathbf{x}_t, \mathbf{z}, t)) d\mathbf{z}$$



Marginal Diffused Data Distributions
True Denoising Distributions
x_t
$q(\boldsymbol{x}_0)$
$q(\boldsymbol{x}_1)$
$q(\boldsymbol{x}_2)$
$q(\boldsymbol{x}_3)$
$q(\boldsymbol{x}_4)$
$q(\boldsymbol{x}_5)$
$\boldsymbol{x}_5 = \boldsymbol{X}$
t
$q(\boldsymbol{x}_0|\boldsymbol{x}_5 = \boldsymbol{X})$
$q(\boldsymbol{x}_1|\boldsymbol{x}_5 = \boldsymbol{X})$
$q(\boldsymbol{x}_2|\boldsymbol{x}_5 = \boldsymbol{X})$
$q(\boldsymbol{x}_3|\boldsymbol{x}_5 = \boldsymbol{X})$
$q(\boldsymbol{x}_4|\boldsymbol{x}_5 = \boldsymbol{X})$


Forward diffusion
$q(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_0)$
$q(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})$
$\boldsymbol{x}_0$
$\boldsymbol{x}_{t-1}$
$\boldsymbol{x}_t$
Real / Fake?
$D(\boldsymbol{x}_{t-1}, \boldsymbol{x}_t, t)$
Conditioning
Posterior sampling
$q(\boldsymbol{x}_{t-1} \mid \boldsymbol{x}_t, \boldsymbol{x}'_0)$
$\boldsymbol{x}'_0$
$\boldsymbol{x}'_{t-1}$
$\boldsymbol{z}$
$G(\boldsymbol{x}_t, \boldsymbol{z}, t)$

Denoising Diffusion Implicit models

Latent Space Diffusion Model