Diffusion Pobabilistic Models

Adrien G. ,Oumaïma B., Yann T.

Dauphine - PSL

8th November 2022

Introduction

Figure: Images generated by DALL.E 2 (Open AI) from text descriptions

Introduction

Inspired by non-equilibrium thermodynamics

The forward and reverse processes

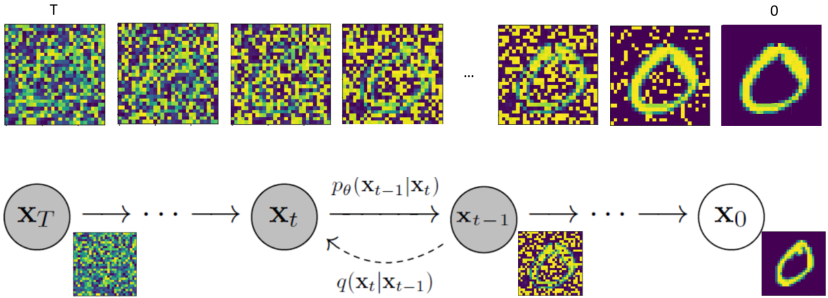


Figure: The Markov chain

$$q(x_t|x_{t-1}) := \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$$
$$p_\theta(\mathrm{x}_{t-1}|\mathrm{x}_t) := \mathcal{N}(\mathrm{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathrm{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathrm{x}_t, t))$$

The forward process

$$q(x_t|x_0) := \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I)$$
$$x_t = \sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon \text{ where } \epsilon \sim \mathcal{N}(0, I)$$

▶ The noise schedule : linear, cosine, quadratic, sigmoid

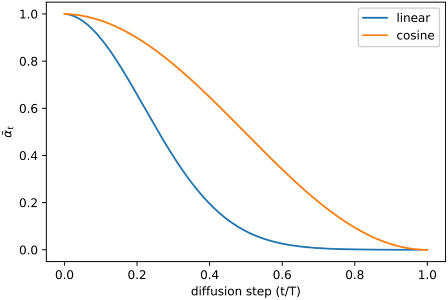


Figure: Comparison of the Linear and Cosine noise schedule (Nichol & Dhariwal,2021)

The forward process

- The choice of the noise schedule

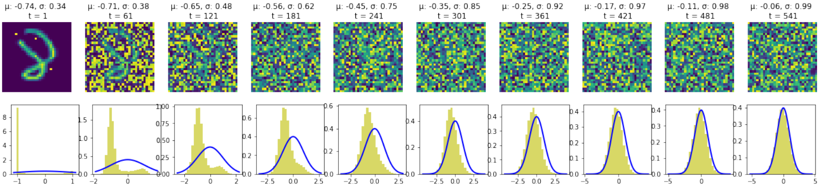


Figure: Linear noise schedule

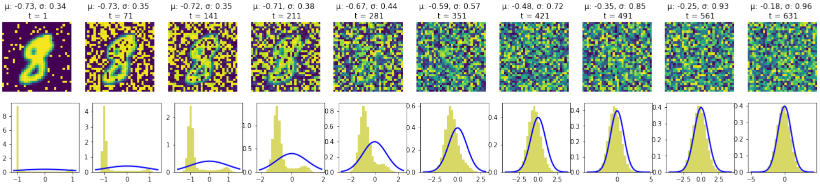


Figure: Cosine noise schedule (optimal according to the papers)

The reverse process

approximated model distribution $p_\theta(\mathrm{x}_{t-1}|\mathrm{x}_t) := \mathcal{N}(\mathrm{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathrm{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathrm{x}_t, t))$

$\theta^* = \arg\max_\theta \sum_{i=1}^{n} logp_\theta(x^{(i)})$

The computed loss function :

$L_{simple}(\theta) := \mathbb{E}_{t,x_0,\epsilon}[\left\|\epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}_t}x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon, t)\right\|^2]$

Algorithm 1 Training

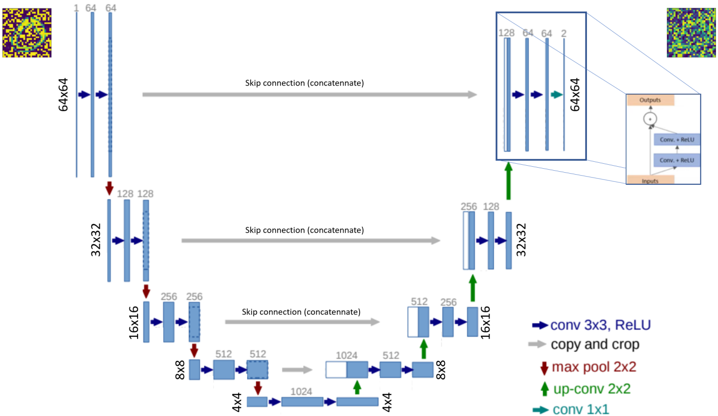
```
repeat
  x_0 ~ q(x_0)
  t ~ Uniform({1, ..., T})
  ε ~ N(0, I)
  Take gradient descent step on
      ∇_θ ‖ε − ε_θ(√(ᾱ_t) x_0 + √(1 − ᾱ_t) ε, t)‖²
until converged
```

Algorithm 2 Sampling

```
x_T ~ N(0, I)
for t = T, ..., 1 do
  z ~ N(0, I) if t > 1, else z = 0
  x_{t−1} = (1/√(α_t)) (x_t − ((1−α_t)/√(1−ᾱ_t)) ε_θ(x_t, t)) + σ_t z
end for
return x_0
```

Figure: Pseudo Code

- The Neural Network architecture : the U-net



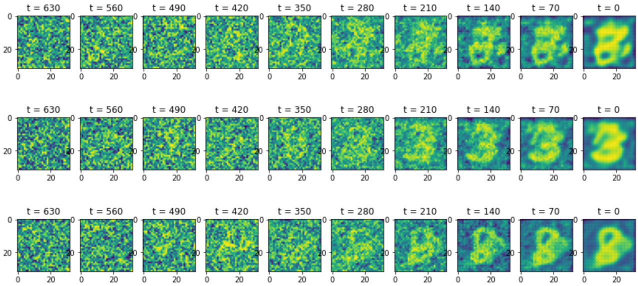


Figure: Preliminary results

▶ Parameters :

Epochs = 10

Image size = 32x32

Variance schedule : sigmoid

T = 700

NN : Simple Unet

Embedding : Sinusoidal Position

Areas for improvement

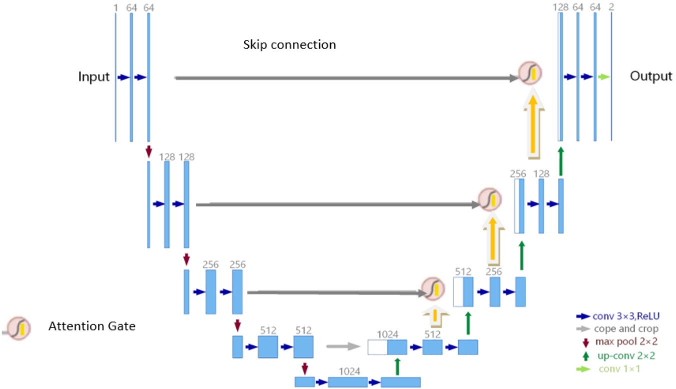


Figure: Attention Unet Structure

References

[1] Denoising Diffusion Probabilistic Models [Jonathan Ho, Ajay Jain , Pieter Abbeel]

[2] file:///C:/Users/bendr/OneDrive/Bureau/M2%20IASD/Data%20Science%20Pro
%20Denoising%20diffusion%20implicit%20models.pdf

[3] https://www.youtube.com/watch?v=GAYJ81M58y8

[4] https://lilianweng.github.io/posts/2021-07-11-diffusion-models/

[5] https://www.youtube.com/watch?v=a4Yfz2FxXiYt=13s