
Image-to-Image style transfer problem with DDPM

Kogut Ivan
Manuylenko Nikita
Zelenkovsky Victor

Sup.: Kolesov Aleksandr

Tasks

- ❖ Reading relevant papers
 - ❖ Run DDIB from male to female
 - ❖ Understand latent space of DDIB
 - ❖ Run UNIT-DDPM from male to female
 - ❖ Understand latent space of UNIT-DDPM
 - ❖ Conduct comparison with Cycle-GAN
 - ❖ Conduct comparison with NOT
-

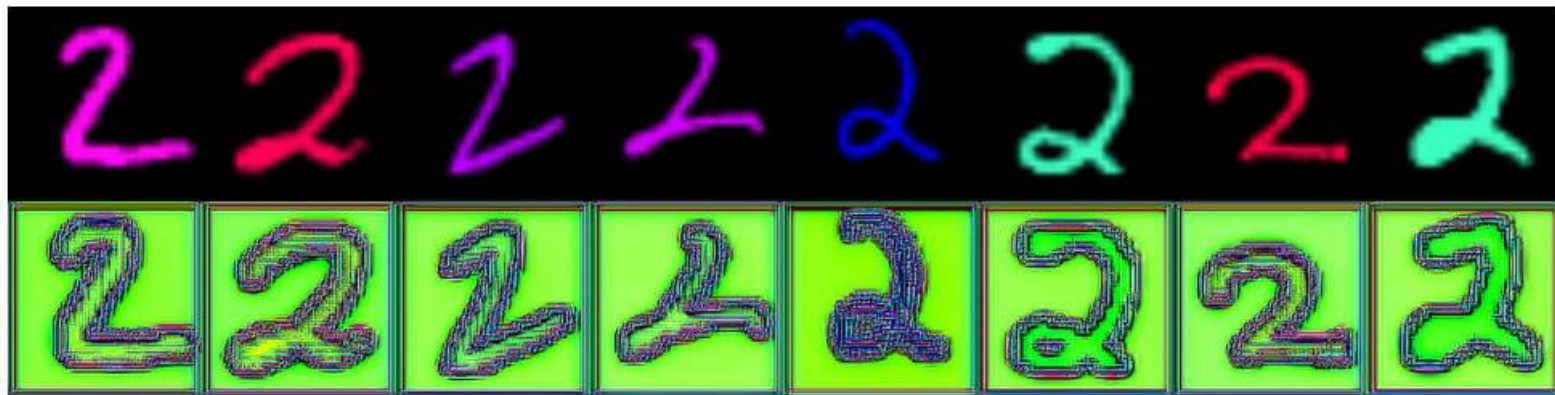
Relevant papers

- ❖ <https://arxiv.org/pdf/1703.10593.pdf> (CycleGan)
 - ❖ <https://openreview.net/forum?id=d8CBRIWNkqH>(NOT)
 - ❖ <https://arxiv.org/abs/2102.09672>(DDPM)
 - ❖ <https://openreview.net/forum?id=5HLoTvVGDe>(DDIB)
 - ❖ <https://arxiv.org/pdf/2104.05358.pdf> (UNIT-DDPM)
-

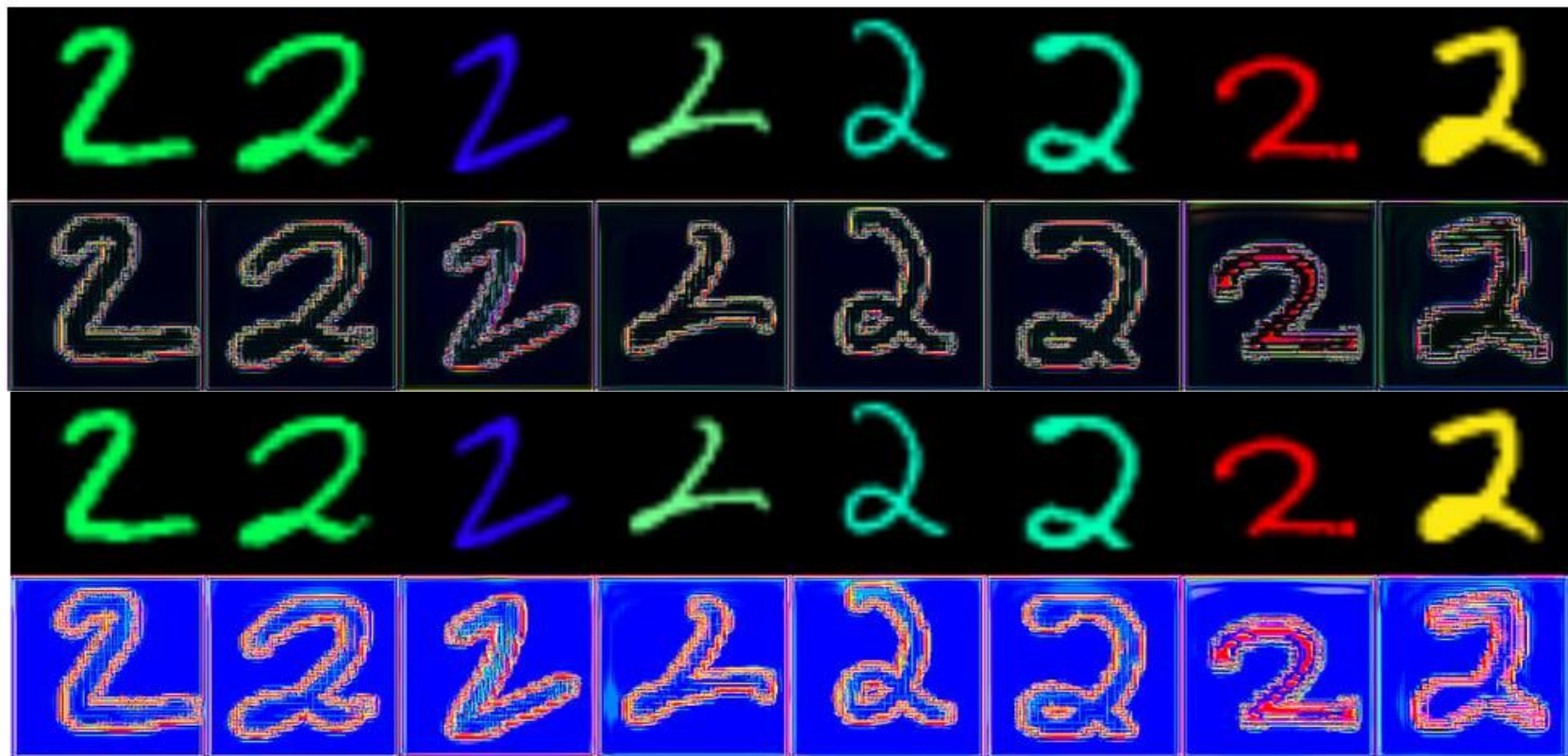
DDIB

Reference code: <https://github.com/suxuann/ddib>,
https://github.com/openai/improved-diffusion/tree/main/improved_diffusion

Dataset: Colored MNIST, class 2 -> class 3

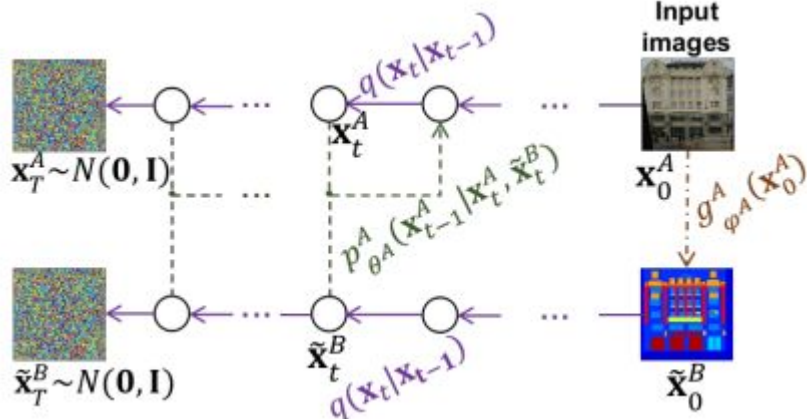


DDIB



UNIT-DDPM model training

$$\mathcal{L}_\theta(\theta^A, \theta^B) = \mathbb{E}_{t, \mathbf{x}_0^A, \epsilon} [\|\epsilon - \epsilon_{\theta^A}^A(\mathbf{x}_t(\mathbf{x}_0^A, \epsilon), \tilde{\mathbf{x}}_t^B, t)\|^2] + \mathbb{E}_{t, \mathbf{x}_0^B, \epsilon} [\|\epsilon - \epsilon_{\theta^B}^B(\mathbf{x}_t(\mathbf{x}_0^B, \epsilon), \tilde{\mathbf{x}}_t^A, t)\|^2] \quad (11)$$



training process

$$\mathcal{L}_{\epsilon^\phi}(\phi^A, \phi^B) = \mathbb{E}_{t, \mathbf{x}_0^B, \epsilon} [\|\epsilon - \epsilon_{\theta^A}^A(\mathbf{x}_t(g_{\phi^B}^B(\mathbf{x}_0^B), \epsilon), \mathbf{x}_t(\mathbf{x}_0^B, \epsilon), t)\|^2 + \|\epsilon - \epsilon_{\theta^B}^B(\mathbf{x}_t(\mathbf{x}_0^B, \epsilon), g_{\phi^B}^B(\mathbf{x}_t(\mathbf{x}_0^B, \epsilon), t)\|^2] + \mathbb{E}_{t, \mathbf{x}_0^A, \epsilon} [\|\epsilon - \epsilon_{\theta^B}^B(\mathbf{x}_t(g_{\phi^A}^A(\mathbf{x}_0^A), \epsilon), \mathbf{x}_t(\mathbf{x}_0^A, \epsilon), t)\|^2 + \|\epsilon - \epsilon_{\theta^A}^A(\mathbf{x}_t(\mathbf{x}_0^A, \epsilon), g_{\phi^A}^A(\mathbf{x}_t(\mathbf{x}_0^A, \epsilon), t)\|^2] \quad (12)$$

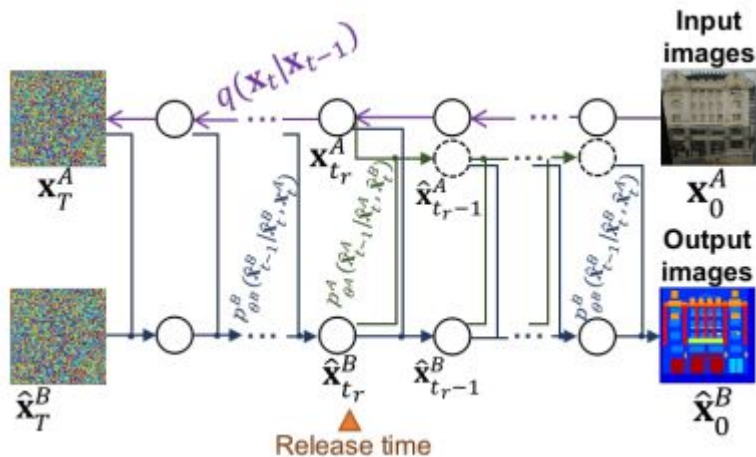
$$\mathcal{L}_{\text{cyc}^\phi}(\phi^A, \phi^B) = \mathbb{E}_{\mathbf{x}_0^B} [\|g_{\phi^A}^A(g_{\phi^B}^B(\mathbf{x}_0^B)) - \mathbf{x}_0^B\|_1] + \mathbb{E}_{\mathbf{x}_0^A} [\|g_{\phi^B}^B(g_{\phi^A}^A(\mathbf{x}_0^A)) - \mathbf{x}_0^A\|_1] \quad (13)$$

The loss function is thus described as follows:

$$\mathcal{L}_\phi(\phi^A, \phi^B) = \mathcal{L}_{\epsilon^\phi}(\phi^A, \phi^B) + \lambda_{\text{cyc}} \mathcal{L}_{\text{cyc}^\phi}(\phi^A, \phi^B), \quad (14)$$

loss functions

UNIT-DDPM model inference



inference process

$$\hat{\mathbf{x}}_{t-1}^B = \mu_{\theta^B}(\hat{\mathbf{x}}_t^B, \hat{\mathbf{x}}_t^A, t) + \Sigma_{\theta^B}(\mathbf{x}_t, t) \epsilon^B \quad (15)$$

$$\hat{\mathbf{x}}_{t-1}^A = \begin{cases} \sqrt{\bar{\alpha}_{t^A}} \mathbf{x}_0^A + \sqrt{1 - \bar{\alpha}_{t^A}} \epsilon^A & (t > t_r) \\ \mu_{\theta^A}(\hat{\mathbf{x}}_t^A, \hat{\mathbf{x}}_t^B, t) + \Sigma_{\theta^A}(\mathbf{x}_t^A, t) \epsilon^B & (t \leq t_r) \end{cases} \quad (16)$$

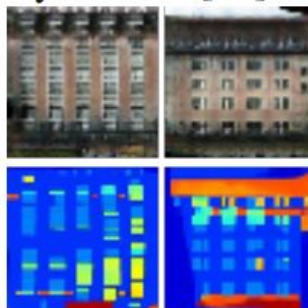
$$\hat{\mathbf{x}}_T^B, \epsilon^A, \epsilon^B \sim \mathcal{N}(0, \mathbf{I}) \quad (17)$$

UNIT-DDPM results

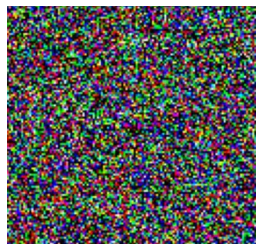
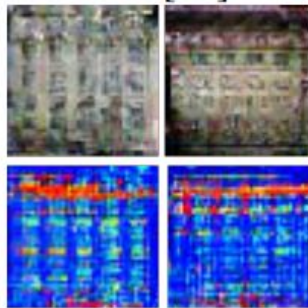
Reference code: <https://github.com/konkuad/UNIT-DDPM-Unofficial/tree/main>

Dataset: color mnist (<https://github.com/jayaneetha/colorized-MNIST>)

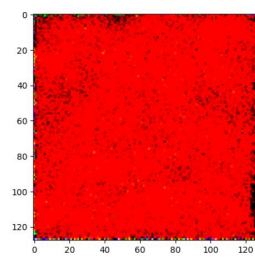
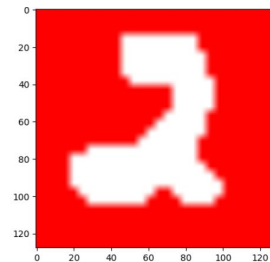
CycleGAN [44]



UNIT [25]



CelebA



colorized mnist

Original paper results
(20,000 epochs)

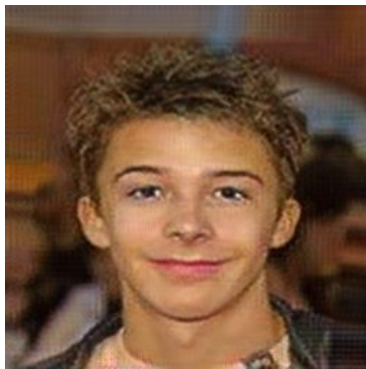
Our results
(100 epochs)

CycleGAN

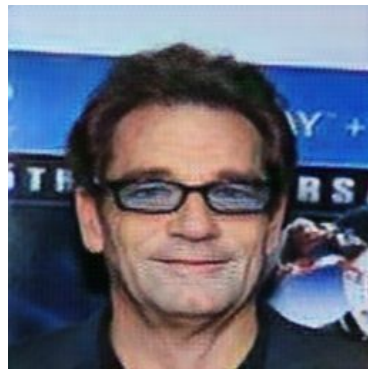
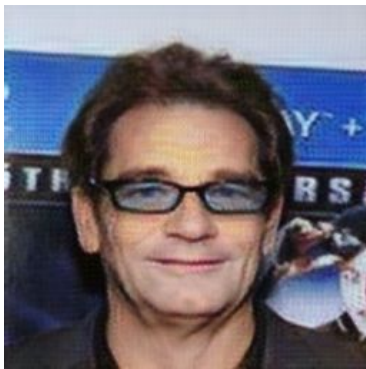
CycleGAN repository: <https://junyanz.github.io/CycleGAN/>.

Dataset: Gender Classification CelebA

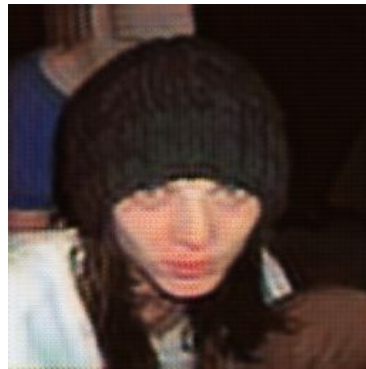
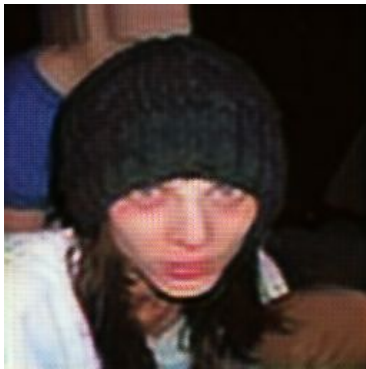
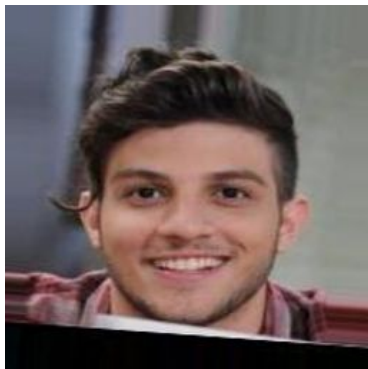
<https://www.kaggle.com/datasets/ashishjangra27/gender-recognition-200k-images-celeba?resource=download-directory>



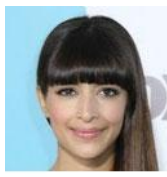
CycleGAN



CycleGAN



CycleGAN other results



NOT

- ❖ Code repository:
<https://github.com/iamalexkorotin/NeuralOptimalTransport>
 - ❖ NOT_training_strong.ipynb for one-to-one unpaired image-to-image transition.
 - ❖ No results at the moment.
-