Project III - Diffusion

Deep Learning 2024

Grzegorz Zakrzewski, 313555

Description of the research problem

Figure: Sample bedroom images from LSUN Bedrooms dataset.

Architecture

Utilized models:

- Denoising Diffusion Probabilistic Models (DDPM)
- Denoising Diffusion Implicit Models (DDIM)

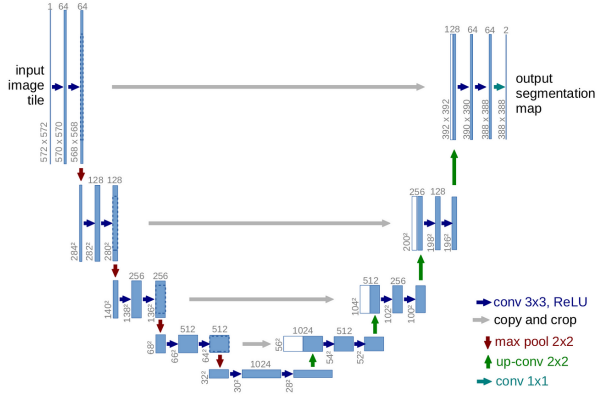

Figure: U-NET architecture example (image source: [1])

Experiment - DDPM

No.	Epochs	Samples	Learning rate	Image size	Block depth	Widths
1	100	3200	2e-4	64x64	2	`[64, 128, 256, 512]`

Table: Details of the DDPM experiment.

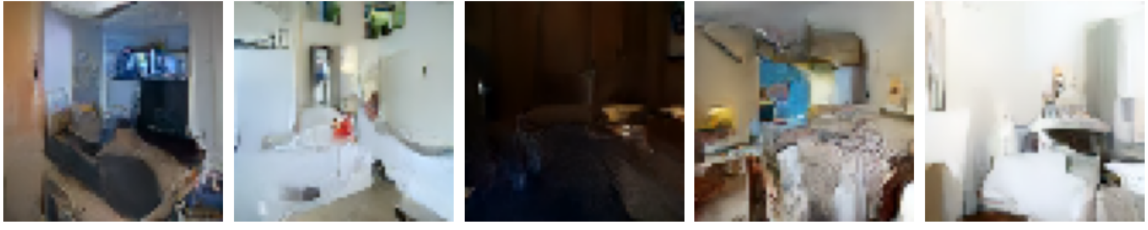

Figure: Images generated in DDPM experiment 1.

Experiments - DDIMs

No.	Epochs	Samples	Learning rate	Weight decay	Image size	Block depth	Widths	FID
1	100	1600	1e-3	1e-4	128x128	2	`[32, 64,`	10.32
2	250	12800					`96, 128]`	10.45
3	200	16000	5e-4	5e-5				18.33
4	300	8000	1e-3	1e-4	80x80			9.75
5	200	8000 (x2 aug.)						9.77
6	400	8000 (x2 aug.)	5e-4	5e-5				7.93
7	300	12000 (x2 aug.)	1e-4	1e-5	80x80	3		5.73
8	300	16000 (x2 aug.)	5e-5	5e-6	80x80	4		5.01
9	400	3200 (x2 aug.)	5e-4	1e-5	112x112	4	`[32, 64, 96,`	42.72
10	400	6400 (x2 aug.)					`128, 256]`	45.87
11	200	12800 (x2 aug.)						56.45

Table: Details of the DDIM experiments.

Experiment DDIM 1.

No.	Epochs	Samples	Learning rate	Weight decay	Image size	Block depth	Widths	FID
1	100	1600	1e-3	1e-4	128x128	2	`[32, 64, 96, 128]`	10.32

Table: Details of the DDIM experiment 1.

Figure: Images generated in DDIM experiment 1 - random collage of many different images.

Experiment DDIM 3.

No.	Epochs	Samples	Learning rate	Weight decay	Image size	Block depth	Widths	FID
3	200	16000	5e-4	5e-5	128x128	2	`[32, 64, 96, 128]`	18.33

Table: Details of the DDIM experiment 3.

Figure: Images generated in DDIM experiment 3 - low quality, blurred.

Experiment DDIM 8.

No.	Epochs	Samples	Learning rate	Weight decay	Image size	Block depth	Widths	FID
8	300	16000 (x2 aug.)	5e-5	5e-6	80x80	4	`[32, 64, 96, 128]`	5.01

Table: Details of the DDIM experiment 8.

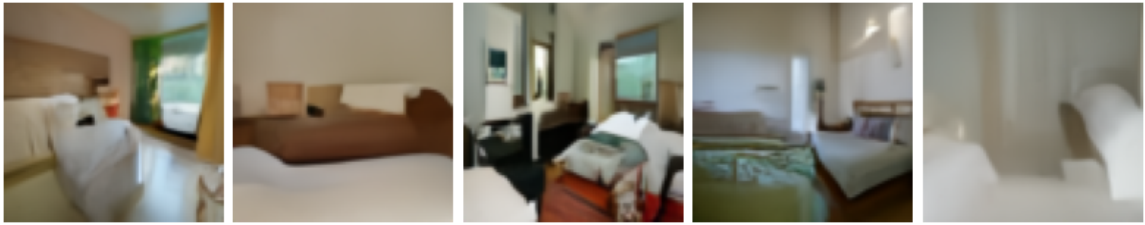

Figure: Images generated in DDIM experiment 8 - lowest learning rate, best FID.

Experiment DDIM 9.

No.	Epochs	Samples	Learning rate	Weight decay	Image size	Block depth	Widths	FID
9	400	3200 (aug.)	5e-4	1e-5	112x112	4	`[32, 64, 96, 128, 256]`	42.72

Table: Details of the DDIM experiment 9.

Figure: Images generated in DDIM experiment 9 - complex architecture, small sample size.

Experiment - latent matrix interpolation

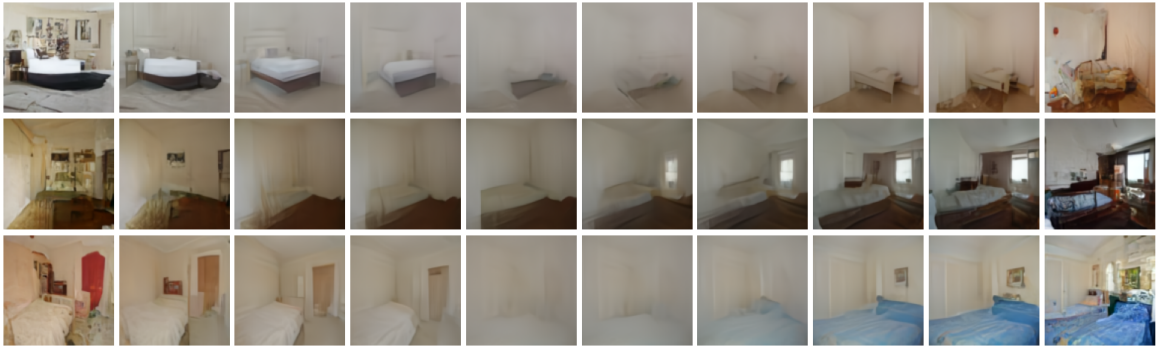

Figure: Results from the experiment on interpolation between latent matrices.

Conclusions

- The project was prepared according to the instructions.
- Used DDPM and DDIM models.
- Tested various hyper-parameters and architectures.
- Generated images often unsharp, blurred or unrealistic.
- These may be the best results achievable with such models and hardware constraints.

References

Olaf Ronneberger, Philipp Fischer, and Thomas Brox.
U-net: Convolutional networks for biomedical image segmentation.
In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.