DEEP LEARNING

o

GENERATIVE MODELS

Natalia Safiejko | Wojciech Grabias

PREPARING THE DATASET

- Resizing (64 x 64)
- Center cropping
- Normalization
- Conversion to tensor

ARCHITECTURES

DCGAN

VAE

DIFFUSION MODEL

DCGAN VAE Diffusion model

Figure 1: Three sample images generated by three different architectures

HYPERPARAMETERS

Evaluation: FID Score

TRAINING PROCESS

Learning Rate
Optimizer
Beta value
Loss function

ARCHITECTURE-SPECIFIC PARAMETERS

Latent dimension
Filter counts in GAN
Timesteps in UNet diffusion model

BATCH SIZE

Table 2: Performance comparison across different batch sizes

Architecture	**batch=64**	**batch=128**	**batch=256**
DCGAN	346.018	357.736	370.288
VAE	120.807	133.822	152.050
Diffusion model	196.418	164.738	182.891

BATCH SIZE

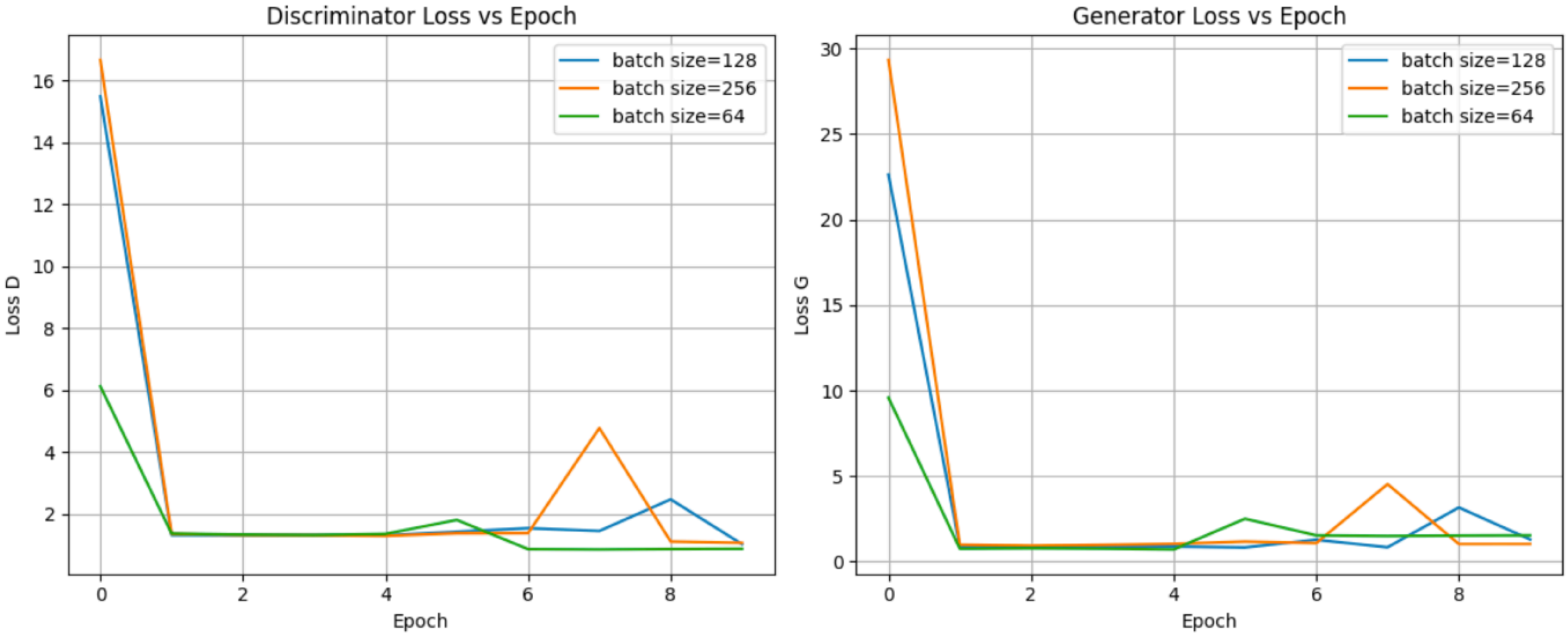


Figure 4: Discriminator and generator loss curves for different batch sizes during DCGAN training.

LEARNING RATE

Table 3: Performance comparison across different learning rates

Architecture	**0.0001**	**0.0002**	**0.0005**	**0.001**	**0.005**	**0.01**
DCGAN	38.589	75.355	83.213	79.890	50.029	412.715
VAE	109.137	120.807	306.088	274.463	347.216	347.190
Diffusion model	147.226	261.701	230.554	298.246	340.721	NaN

LEARNING RATE

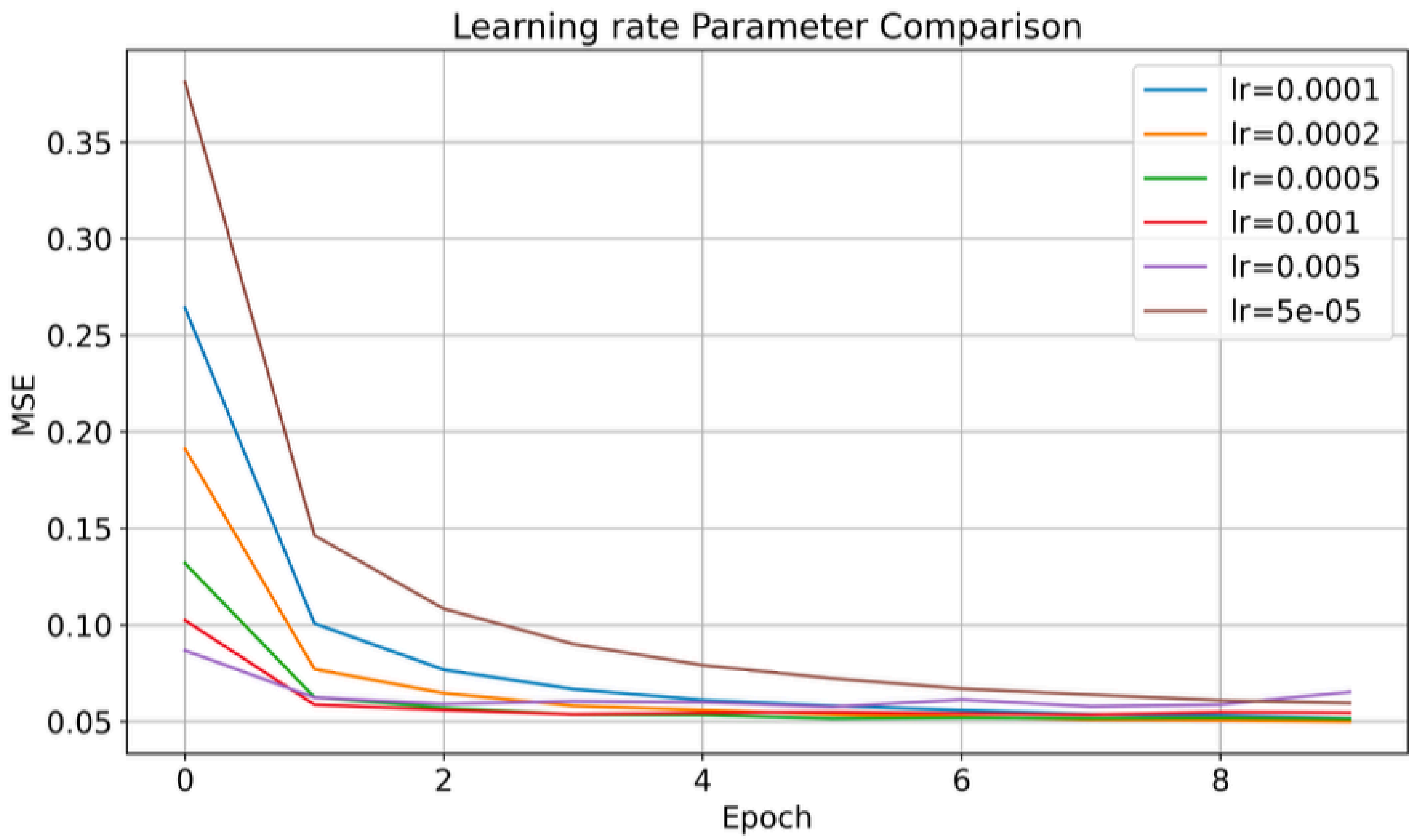


Figure 6: Total loss curve for different learning rates during diffusion model training.

BETA PARAMETER

Table 4: Performance comparison across different β_1 values

Architecture	$\beta_1 = 0$	$\beta_1 = 0.3$	$\beta_1 = 0.5$	$\beta_1 = 0.7$
DCGAN	38.09	48.81	38.59	55.51
VAE	111.22	123.15	109.14	127.23
UNET	278.665	187.994	237.30	258.43

NDF NGF

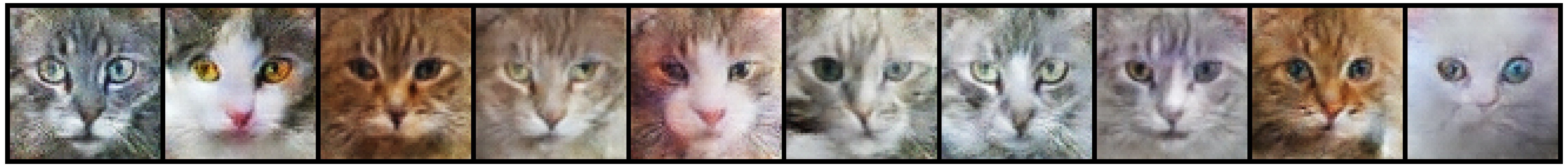
Table 7: Performance metrics for different filter counts in DCGAN architecture

Parameter	**32 filters**	**64 filters**	**128 filters**
ndf (Discriminator)	48.83	38.59	39.19
ngf (Generator)	53.02	38.59	53.66

DIFFERENT LOSS FUNCTIONS

Table 5: Performance comparison across different loss functions

DCGAN	**LSGAN**	**SNGAN**
44.180	35.731	30.976



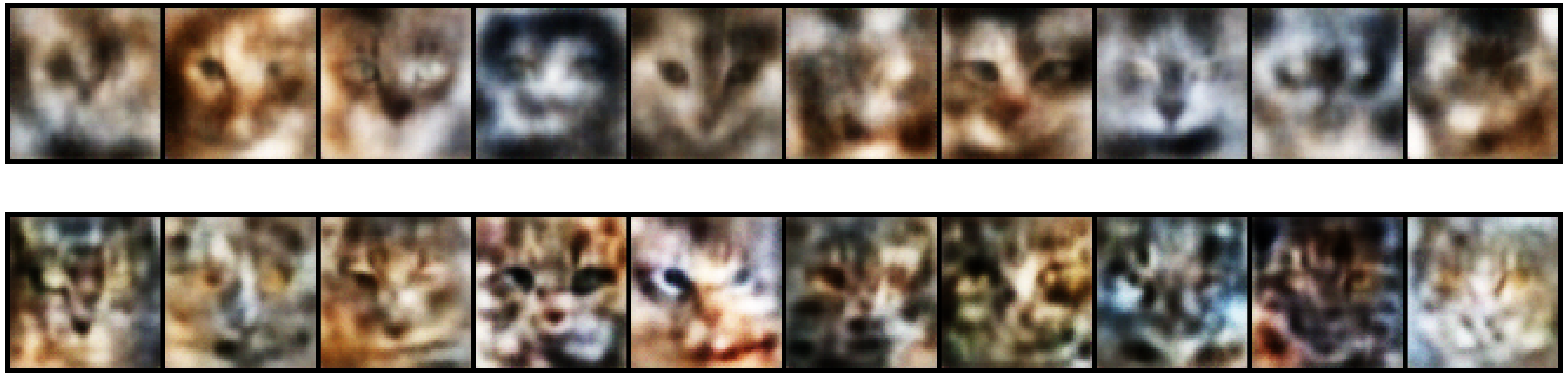
LATENT DIM

Table 6: Performance comparison across different latent dimension sizes

Architecture	**50-dim**	**100-dim**	**200-dim**
DCGAN	39.454	38.589	41.928
VAE	113.689	109.137	104.949

DIM=50

DIM=200

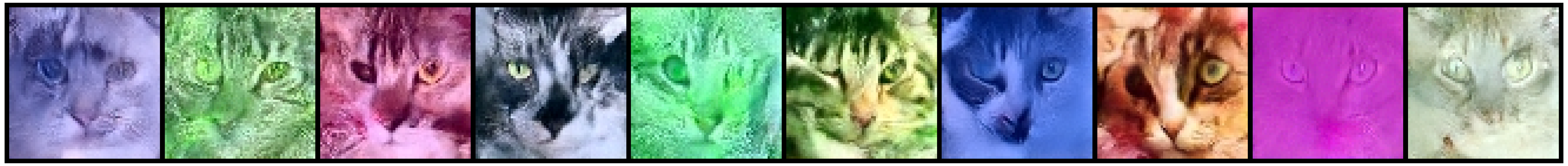


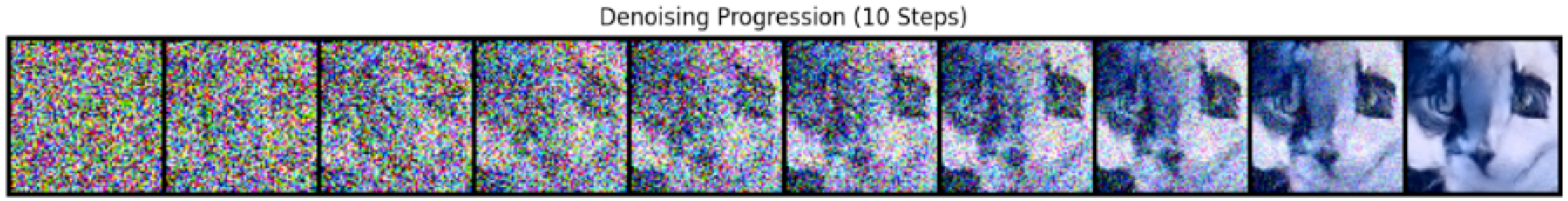
VAE

TIMESTEPS IN UNET DIFFUSION MODEL

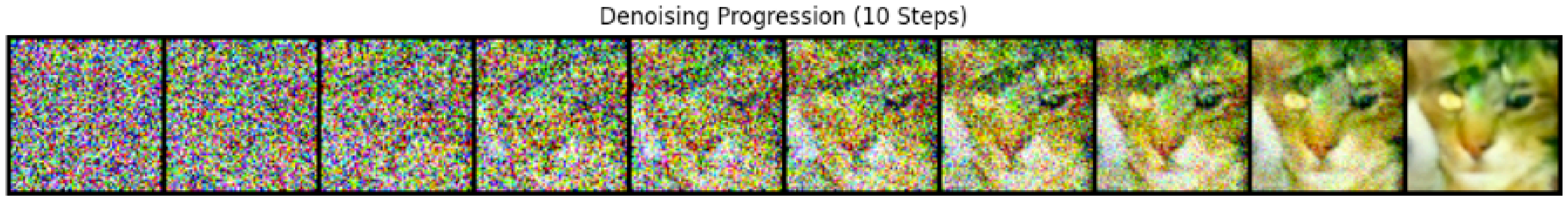
Table 8: Performance scores at different timesteps count

Timesteps	**100**	**250**	**500**	**750**	**1000**	**1500**
Score	83.648	171 280	140.617	231.182	200.919	177.716

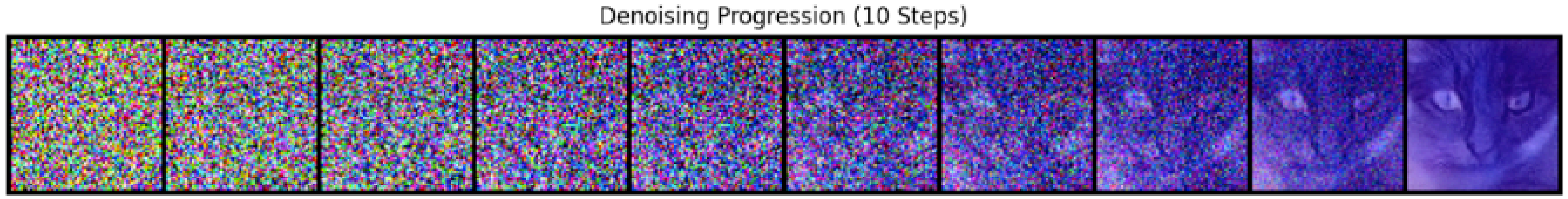




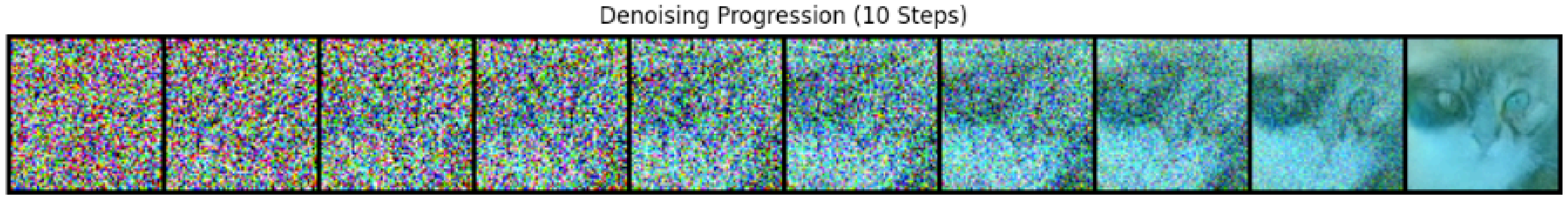
(a) 100 timesteps



(b) 500 timesteps



(c) 1000 timesteps



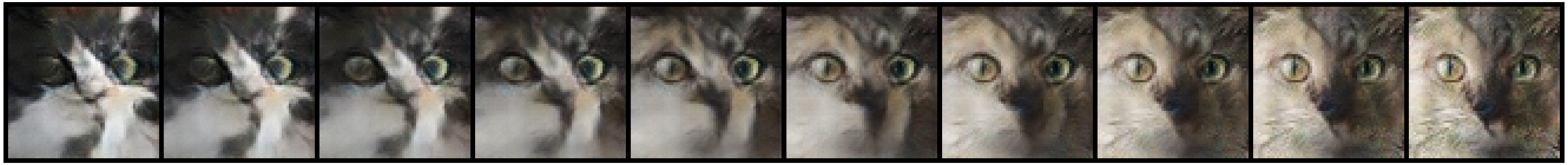
(d) 1500 timesteps

Figure 10: Comparison of images' denoising process for different timestep counts in the diffusion model

CATS AND DOGS

LATENT SPACE EXPLORATION AND INTERPOLATION

$$\mathbf{z}_i = (1 - \alpha_i) \cdot \mathbf{z}_1 + \alpha_i \cdot \mathbf{z}_2, \quad \alpha_i \in \left\{ \frac{1}{9}, \frac{2}{9}, \ldots, \frac{8}{9} \right\}$$



BIBLIOGRAPHY

[1] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. Advances in neural information processing systems, 27, 2014.

[2] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. arXiv preprint arXiv:2006.11239, 2020.

[3] Diederik P Kingma, Max Welling, et al. An introduction to variational autoencoders. Foundations and Trends® in Machine Learning, 12(4):307–392, 2019.

[4] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. Advances in Neural Information Processing Systems, 2019.

THANK YOU