DTU Compute
Department of Applied Mathematics and Computer Science

DDPM-based reparation of overexposure

Jakub Wiśniewski, Tobias Munch
s230212@dtu.dk, s183478@dtu.dk

1 Introduction

Photography is a skill that is easy to learn but hard to master. One of the potential mistakes that one can make is to let too much light into the camera, overexposing the camera sensors as seen in Figure 1. These overexposed areas are considered lost information, but in this project, we will train a Diffusion model to recreate this information based on the rest of the image.

Figure 1: Example of overexposure. Notice that the skies are completely white

2 Data

- Modified MIT-Adobe FiveK dataset with 5,000 images featuring various exposure levels, brightness and contrast further increase to emulate exposure damage
- Mask based on pixels at the maximum value of 255 across all RGB channels.
- Image is separated into known regions (reference for DDPM) and unknown regions (DDPM inference target).

3 Model

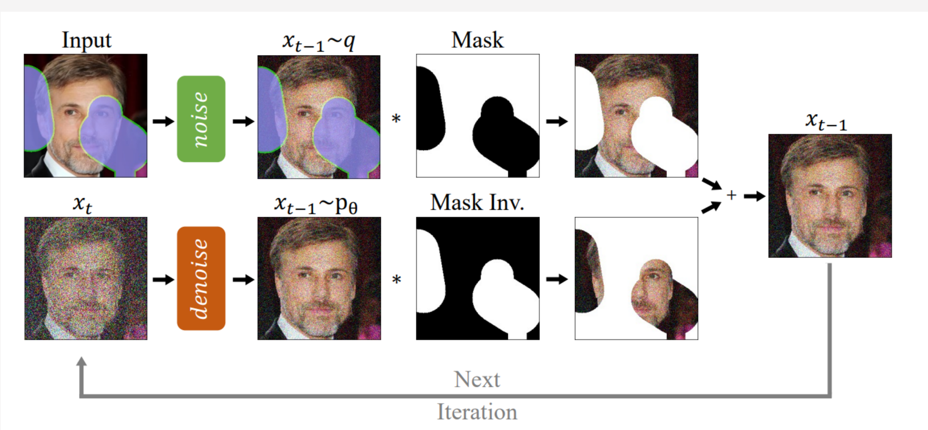


Figure 2: The procedure for masked inference as used in the paper by *Lugmayr et al.*

- DDPM model based on UNet structure, reverse process expanded with mask functionality from *Lugmayr et al.*'s work.
- Trained on both randomly initialized and CIFAR-10 pre-trained weights.
- *Histogram stretching* used to fix non-damaged parts.

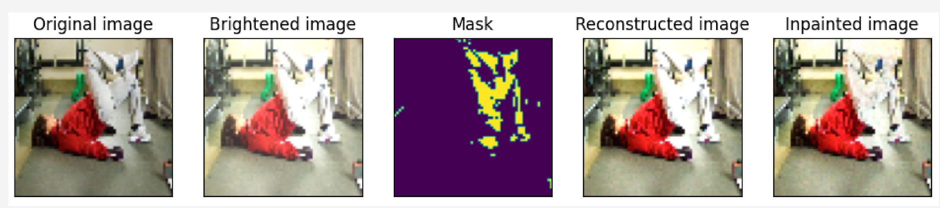


Figure 3: Image reconstruction pipeline

4 Results

- Training parameters: 30 epochs, MSE loss, 64x64 image size, batch size of 4.
- Training curves were compared between a pretrained model and a model trained from the ground up, loss given in figure 4

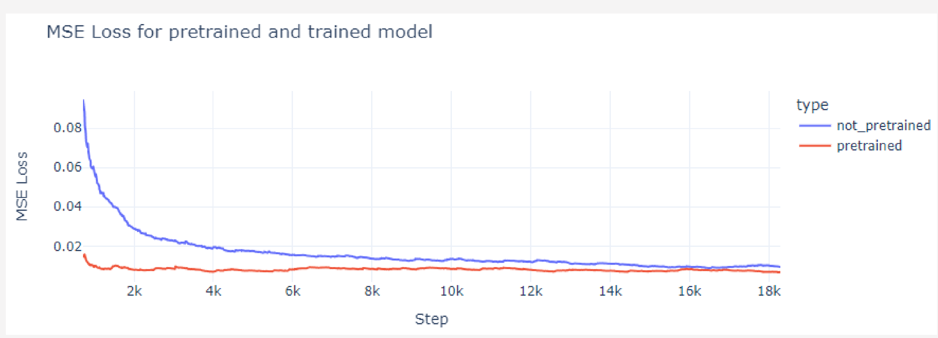


Figure 4: Training loss for trained and pretrained model

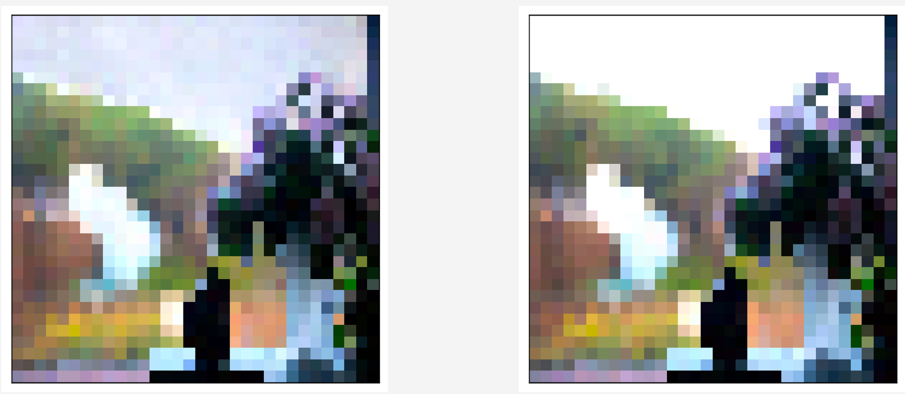
Figure 5: Results of training. The upper image is at the start of training (epoch 1), while the bottom image is after training (epoch 30).

5 Ablative study

	Pretrained	Not pretrained
Trained with MSE	0.026	0.032
Trained with LPIPS	0.006	0.273

Table 1: Comparison of different training setups. The metric used for comparison is the perceptual LPIPS metric

- Trained four distinct models on 32x32 images for 30 epochs, results shown in Table 1.
- Pretrained model with LPIPS has lowest loss, but assumed to have overfit (See figure 5)
- Pretrained model with MSE is preferred



(a) Trained with MSE

(b) Trained with LPIPS

Figure 6: Comparison of 2 pretrained models

6 Discussion

- Model is able to generate new content in most overexposed images, but struggles with certain images
- Input image size limitations or CIFAR-10 preconditioning seems to be most plausible.
- Potential in the small model, and could benefit from more computational power.

7 Conclusions

- We have trained a DDPM model that is able to generate new content in images damaged by overexposure
- Model works well on small images using a pretrained network and MSE loss, and there is potential for working on bigger images.
- Another idea to improve the model would be to use a combination of MSE and perceptual loss to increase the quality of the generation