Old Photo Restoration using Diffusion

Judhajit Roy
University of Illinois at Chicago
jroy20@uic.edu

1 Introduction

Photos are taken to capture the happy moments and looking at them revives one's memories of the past. Old photo prints deteriorate over time when they are not properly preserved, causing the photograph to get permanently damaged. Fortunately, we can now digitalize the photos and invite a skilled specialist for restoration. However, manually retouching and editing these old photos is very laborious and takes a lot of time. Hence, it would be quite handy to have a model that can restore old photos to their former glory. There are several types of degradations that occur on the photos like film grain, sepia effect and scratches. these can be further categorized into structured and unstructured degradations. In this experiment we try to remove such degradations from images in CIFAR10 dataset using a Diffusion Model.

2 Background

Image Restoration typically includes super-resolution , de-noising , de-blurring and compression removal. While in old photo restoration problem multiple such degradations exist. Earlier work done to restore old photos uses Variational AutoEncoders in shared latent space with real old photos and synthetic versions of the original images sharing the saming latent space, while the ground truth is projected to different latent space. After translating the synthetic image to a latent space it maps this to the latent space of ground truth using a parameterized model. Thus, by learning this translation, it can replicated for the real old image as well. Another such work uses Generative Facial Prior restore realistic facial details. This method restores photo containing faces particularly. Diffusion models show promising results in areas like super-resolution and inpainting. Thus, it is quite possible that such Diffusion models can be useful for the old photo restoration task. For solving the super-resolution problem the paper Denoising Diffusion Restoration Models builds up on the Denoising Diffusion Probabilistic Model. In this approach, they mention using a reference image as an additional context to condition on in the diffusion process. By conditioning on this reference image they were able to generate High quality versions of the image. Thus, we use this technique to condtion the inputs in the experiments but condition on images with blur and film grain noise.

3 Method

In a Diffusion Modelling, we need to build a model that is able to learn the systematic decay of information by addition of noise. After learning the systematic decay the model should be able to reverse the process and therefore, recover the information back from the noise. This concept is similar to what VAEs do when optimizing an objective function by initially projecting the data to a latent space and then recovering back the initial state. However, a diffusion model instead learns to model a chain of noise distributions in a Markov Chain and decodes/denoises the data in a hierarchical fashion.

Thus, Denoising diffusion modeling is a two stage process: containing the forward diffusion process and the reverse process. In the forward diffusion process, gaussian noise is introduced to the data successively in steps until the data is fully converted to noise. The reverse process undoes the noise

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

by learning the conditional probability density at an earlier time step given the current state of the system.

The forward diffusion has a closed form because at any arbitrary time step "t" we can sample x_t using the reparameterization trick. Whereas, we cannot easily estimate $q(x_{t-1}|x_t)$ because it requires the entire dataset and therefore we need to learn a model p_θ to approximate these conditional probabilities for the reverse diffusion process.

Diffusion model is given by the equation of the joint distribution,

$$p_\theta(x_{0:\mathrm{T}}) = {p_\theta}^{(\mathrm{T})}(x_T) \prod_{t=0}^{T-1} {p_\theta}^{(\mathrm{t})}(x_t|x_{\mathrm{t+1}})$$

After estimating $x_{0:T}$, we only sample x_0 for the generative model and train the reverse process using the variational inference distribution,

$$q(x_{1:\mathrm{T}}|x_0) = q^{(\mathrm{T})}(x_T|x_0) \prod_{t=0}^{T-1} q^{(\mathrm{t})}(x_t|x_{\mathrm{t+1}}, x_0)$$

However, in order to remove a degradation this diffusion model is not sufficient. It can only regenerate the image x. But, if x is additionally provided a context y of a degraded image when d, then it is possible to reverse the diffusion process and additionally removing the degradation. This can be done by constructing a Markov chain x_T to x_0 conditioned on y,

$$p_\theta(x_{0:\mathrm{T}}|y) = {p_\theta}^{(\mathrm{T})}(x_T|y) \prod_{t=0}^{T-1} {p_\theta}^{(\mathrm{t})}(x_t|x_{\mathrm{t+1}}, y)$$

While in the reverse diffusion process to retrieve the image we can condition on y,

$$q(x_{1:\mathrm{T}}|x_0, y) = q^{(\mathrm{T})}(x_T|x_0, y) \prod_{t=0}^{T-1} q^{(\mathrm{t})}(x_t|x_{\mathrm{t+1}}, x_0, y)$$

Thus, apart from diffusing the original image it is conditioned on a reference image to generate an image without the degradation.

4 Experiments and Results

For Old Photo restoration using Diffusion models, apart form the original image we provide a reference image (which is an image with a certain type of degradation) as context to be conditioned on, so that the model learns to remove such degradation. Two types of experiments, in the first the reference image is a blurred version of the original image and in the second experiment the reference image is made noisy to replicate film grain degradation. The CIFAR10 Dataset images are originally of 32x32 dimensions. For generating the blur effect we resize the image 8x8 and concatenate it in the diffusion process to the objective image. Similarly, the film grain effect was created by adding Gaussian of intensity 0.1.

In the training process, the model calculates the perturbed image obtained by forward diffusion at a time step and then it predicts the actual noise added that particular time step. Further, the Mean Squared loss is used to measure the difference between the predicted and actual noise. As this loss decreases with iterations the model correctly predicts the noise and is able to generate better images. Also, the image is downsampled and upsampled using a U-Net architecture with residual blocks. The residual connections are added to upsampling layers occasionally.

These results were obtained after training the diffusion model on two of the classes of the CIFAR20 dataset (horse and dog) for 20 epochs with a learning rate of 1e-4. As per observations, the diffusion model is effectively able to generate images without the film grain noise but for the blurred images the model find it difficult.

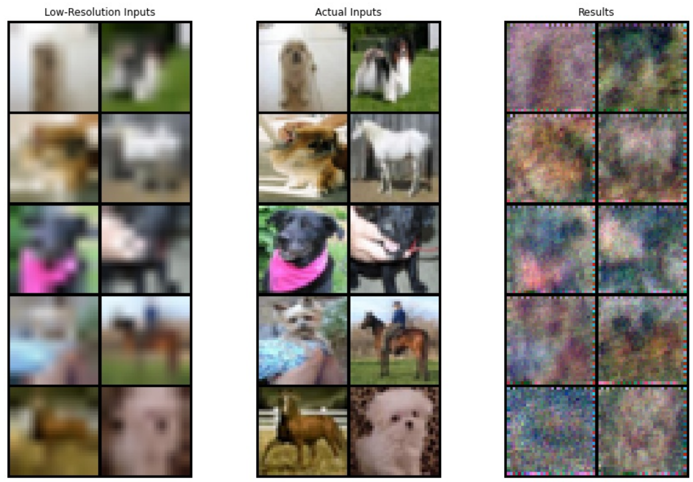

Figure 1: Performance of Diffusion Model on Photo with Blur

Figure 2: Performance of Diffusion Model on Photo with Film Grain

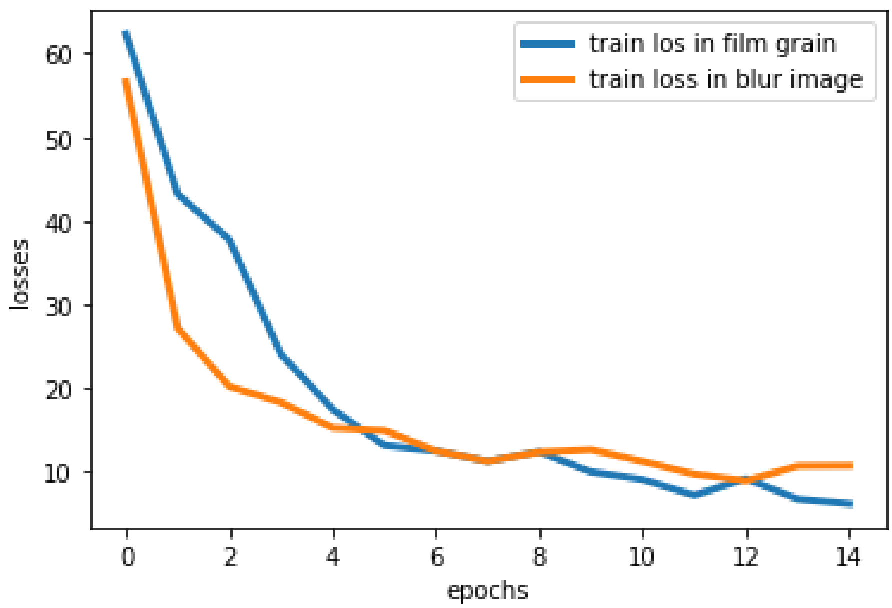

Figure 3: MSE Loss plot for the models

5 Discussion

Due to computation limitations we were only able to eperiment on two classes of CIFAR10 dataset. We are able to somewhat restore the degradation from the images. But this is done independently, whereas Old Photos contain multiple such degradations. A possible way that was discovered was that, multiple diffusion models can be cascaded one after the other to progressively remove multiple degradations. Here, each diffusion is trained to remove one such degradation. When these models are cascaded it is possible that all the degradations are removed. Apart from another experiment could be to project the images into latent space and create a latent diffusion model.

References

[1] Kawar, B., Elad, M., Ermon, S., Song, J. (2022). Denoising diffusion restoration models. arXiv preprint arXiv:2201.11793.

[2] Ho, J., Jain, A., Abbeel, P. (2020). Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems, 33, 6840-6851.

[3] Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D. J., Norouzi, M. (2022). Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] https://github.com/KiUngSong/Generative-Models/tree/main/SR3