Deep Generative Models: Project Report
Enhanced Image Inpainting with RePaint: Different Conditioning Approach Using Multiple Noisy Versions
Dor Kopito 208677062

RePaint employs a state-of-the-art approach to image inpainting by utilizing pre-trained diffusion models to iteratively refine and fill in missing or damaged parts of images. This method leverages an inference scheme that capitalizes on the generative capabilities of these models, producing results that seamlessly blend with the original image context without requiring additional model training or fine-tuning. Despite its effectiveness, there is potential to further enhance the model's precision and contextual coherence, especially in complex scenes. This paper explores an improved conditioning method by generating multiple noisy versions of the input image and aggregating them, aiming to achieve more robust and accurate inpainting results.

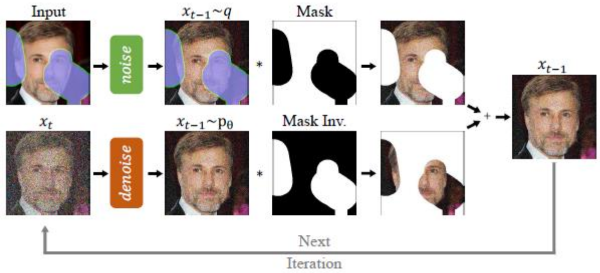

Fig1: Overview of RePaint approach. RePaint modifies the standard denoising process in order to condition on the given image content. In each step, we sample the known region (top) from the input and the inpainted part from the DDPM output (bottom).

Fig2: The output of repaint on different images and masks.

In this work, I aim to explore the potential for enhancing the RePaint approach by improving its conditioning process. Specifically, I propose generating multiple noisy versions of the input image and aggregating these versions to create a potentially more robust and precise estimate for inpainting. This technique is intended to improve the model's ability to seamlessly integrate the inpainted regions with the surrounding context, particularly in complex scenes. While the anticipated improvements are promising, this approach remains experimental and its effectiveness will be evaluated through extensive testing and analysis. If the sample image is:

$$\bar{x}_{t-1} = m \odot x_{t,known} + (1-m) \odot \tilde{x}_t$$

We create:

$$\bar{x}^k_{t-1} = m \odot x^k_{t,known} + (1-m) \odot \tilde{x}_t \,, k = 1, \dots, K$$

Where $x_{t,known}$ is the original image after t steps of noise, m is the mask, $\tilde{x}_t$ is the denoised image at timestep t and K is over how many different images we want to average.

After adjusting the code here: https://github.com/kopitodor/DGM_Final_Project and running test over the same images I obtained the results (for $K = 10$):

Fig3: The output of repaint with the modification.

The result shows that the following method as I implemented it did not work well. Perhaps increasing the value of K might obtain better results. Moreover, In the context of RePaint, utilizing Denoising Diffusion Probabilistic Models (DDPM) might be superior due to their structured approach to noise reduction and image refinement. DDPMs excel at iteratively denoising images, ensuring that each step builds upon the previous one to progressively enhance image quality. This iterative process is particularly beneficial in inpainting tasks, where maintaining the coherence and natural appearance of the image is crucial. By adhering to the probabilistic framework of DDPMs, RePaint can effectively manage the balance between preserving known regions and accurately reconstructing missing parts. This approach ensures that the inpainted areas seamlessly blend with the original content, leveraging the model's ability to handle complex textures and structures with precision.