

GAI 2024 Project 4

資訊三甲 F64106058 紀宣如

Theoretical Justification :

這次作業我挑選的是 example2 : "By adding noise to the target image at different levels and using these noisy versions as intermediate targets, you can guide the DIP model to learn a hierarchical representation of the image." 實作。一般來說，DIP 模型訓練是使用完全高斯分布的噪聲圖像作為輸入，而 example2 的做法就是在 DIP 訓練過程中參考 DDPM 的訓練過程，引入漸進的去噪步驟，通過給目標圖像添加不同程度的噪聲，並用這些有噪聲的圖像作為中間目標，引導 DIP 模型學習圖像的不同層次。

- potential benefits :

- 分層學習：引入漸進去噪步驟能幫助 DIP 模型更有效地學習圖像的分層表示，可能會提高最終重建圖像的質量。
- 更好的停止點：通過監控每個去噪階段的重建質量，可以更準確地確定最佳停止點，避免過度訓練或訓練不足。
- 更穩定的訓練：逐步去噪的過程可能會使訓練更穩定，減少因直接處理高噪聲圖像而帶來的不穩定性。

-potential limitations :

- 計算成本增加：引入多個去噪階段會增加計算成本和訓練時間。
- 複雜度增加：訓練過程變得更複雜，需要調整更多的參數和配置。
- Over fitting 風險：如果噪聲水平和去噪階段設置不當，可能會導致模型過擬合於有噪聲的圖像，而不能很好地重建乾淨的目標圖像。

Experimental Verification :

DDPM 會不斷在影像上反覆添加雜訊(Noise)達到訓練的效果，參考以下公式：

$q(x_t|x_{t-1}) := \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I)$ 定義 $\alpha = 1-\beta$ 會變為：

$q(x_t|x_0) := \mathcal{N}(x_t; \sqrt{\alpha_t}x_0, (1-\alpha_t)I)$

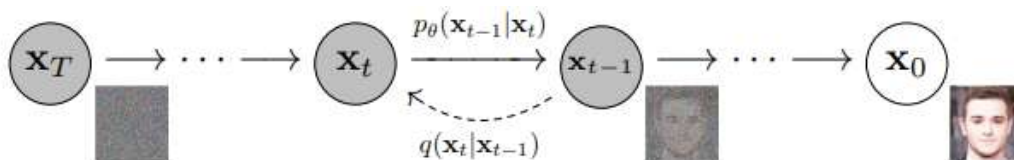


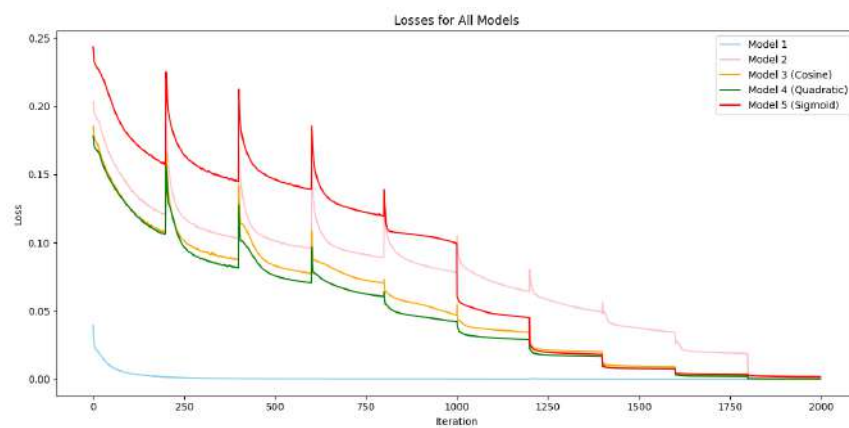
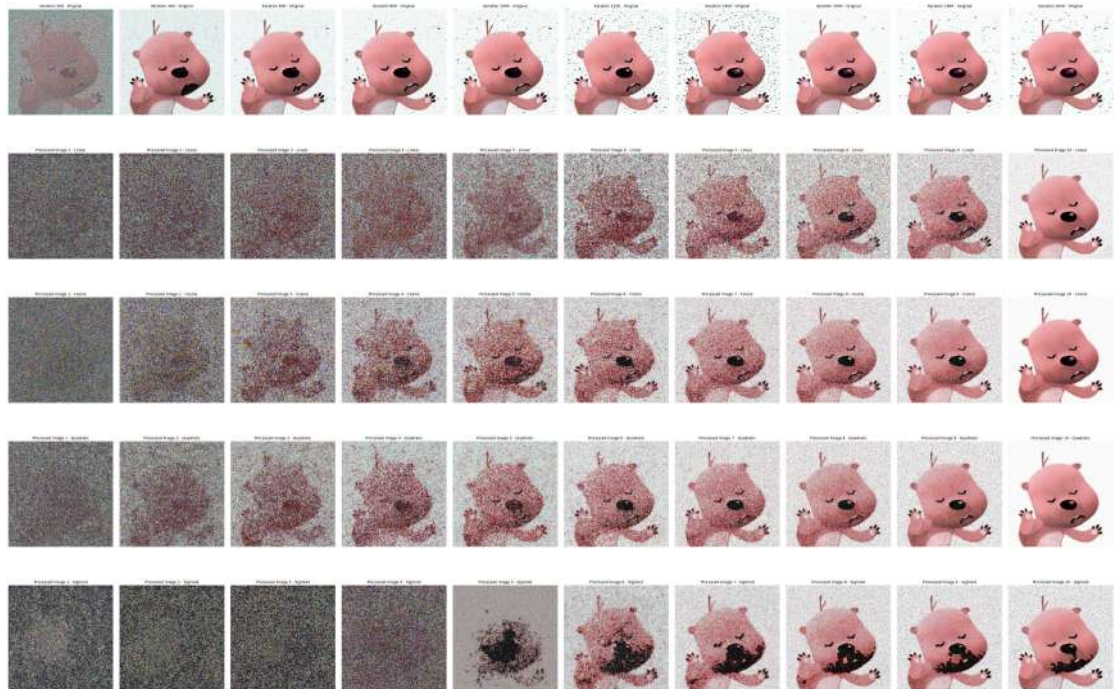
Figure 2: The directed graphical model considered in this work.

$\mathcal{N}()$ 代表 Normal Distribution，裡面的三個 parameters 分別是 $\mathcal{N}(\text{output image}, \text{mean}, \text{variance})$ ， β 是表示一個 schedule，決定圖片加上雜訊的過程快慢。在最初的論文中使用的是 linear schedule(等差數列)，而在 2021 有論文提出了 cosine schedule，改善了圖片資訊破壞過快的問題。在這次實作中，我又新增了

quadratic, sigmoid shedule 兩種方式來增加雜訊：



訓練結果如圖（由上到下分別為原始 DIP 以及上面四種增加雜訊方式）：



	original	linear	cosine	quadratic	sigmoid
SSIM	0.9959204	0.9864079	0.9874548	0.9909497	0.9542108

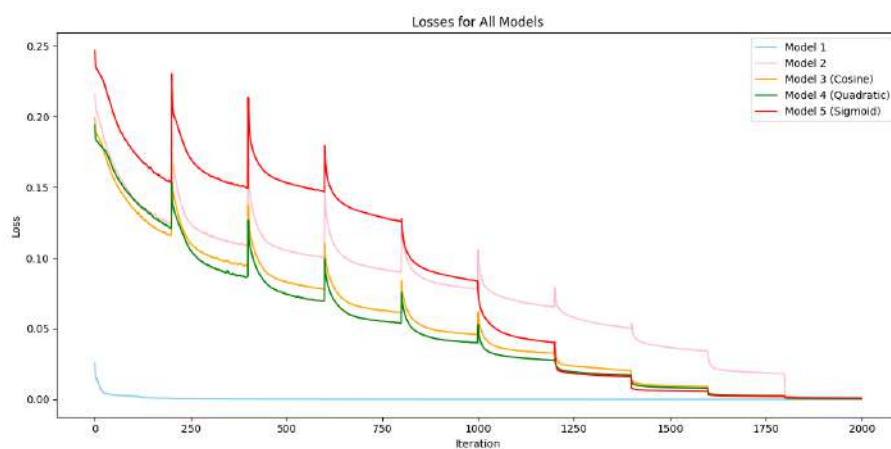
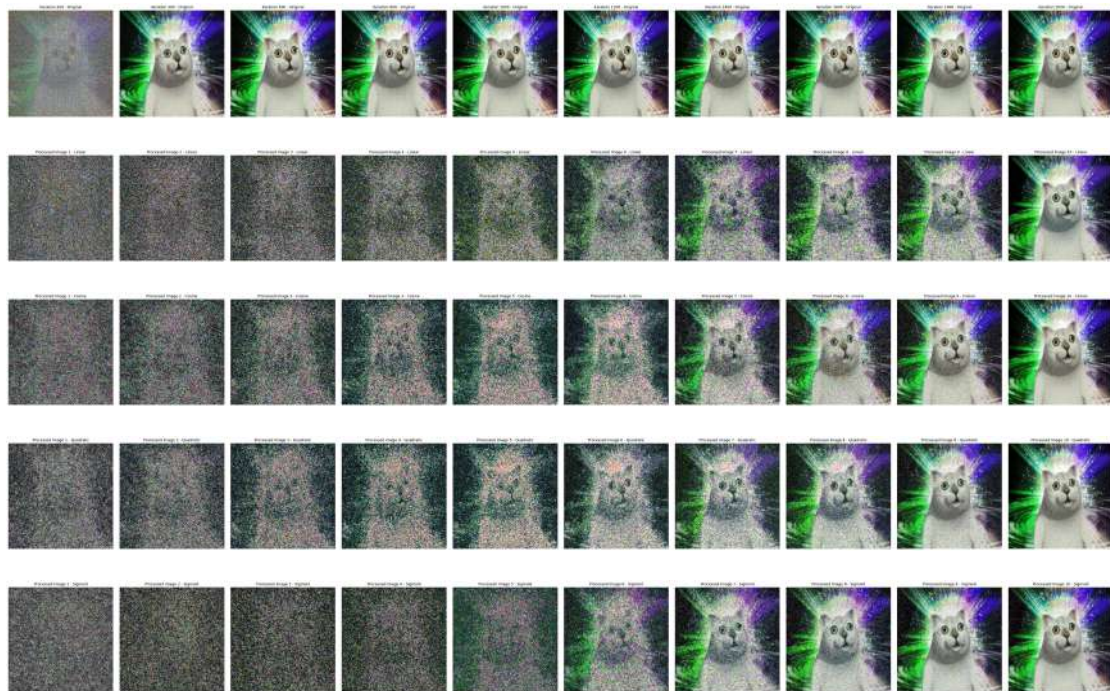
觀察發現 Loss 都有收斂，而 SSIM 則是以 quadratic 表現最佳，但就圖片來看，沒有 time step noise 的 DIP 在處理白色背景上表現較差，而有 time step noise 的 DIP 基本上都能還原出白色背景。

Ablation Studies and Analysis :

除了使用四種不同的噪聲調度來查看對模型重建質量的影響，我還調整了不同 learning rate 以及使用不同圖片

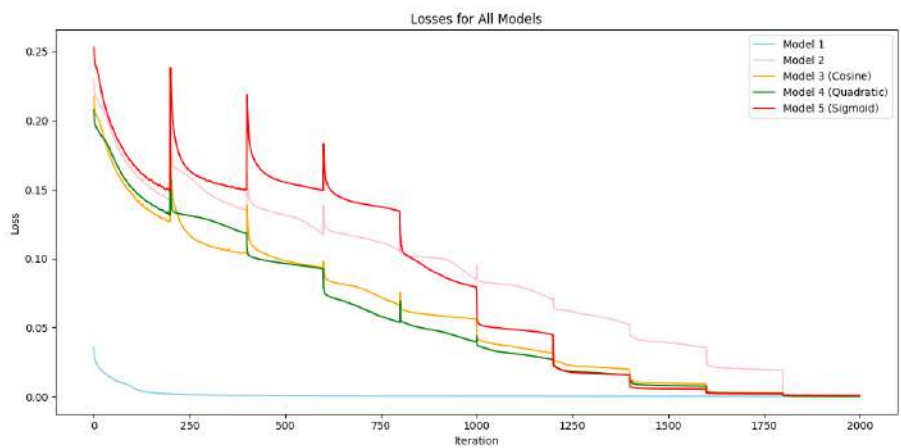
learning rate : $1e-2$ 、 $1e-1$ 、 $1e-3$ 輸出結果差不多。

色彩較豐富的圖：



	original	linear	cosine	quadratic	sigmoid
SSIM	0.9952752	0.9276231	0.9769276	0.9712541	0.9166557

較暗且較真實的照片：



	original	linear	cosine	quadratic	sigmoid
SSIM	0.9870823	0.9339840	0.9469224	0.9478220	0.8401645

觀察：在色彩豐富且複雜的圖中，consine 表現較好，而較暗且較真實的照片中，quadratic 表現則較佳，但兩者都不如 original DIP 的表現，但肉眼觀察的話，可以看出有 time step noise 的模型去雜訊後的圖片比較和諧，沒有突兀的噪點，在實際應用中，還需視情況選擇適合圖片的 model。