

GAI_Project4_Generative Models for Visual Signals

系級:資訊 114 學號:E34106010 姓名:黃偉峰

Github link: https://github.com/ukp66482/GAI_Project4

1. Theoretical Justification

此次作業我挑選了 Example1 中提到的 Accelerating DDPM with DIP-based initial Priors 來實作，透過 DIP，並且以 DDPM-only 以及 DDPM with DIP 來作比較。

結合這兩種方法，DDPM with DIP-Based Initial Prior 的原理如下：

DIP 訓練：

首先訓練 DIP 模型，使其能夠重建目標圖像的高級特徵。DIP 模型在短時間內訓練，以捕捉圖像的主要結構和模式，而不過度擬合噪聲。

DDPM 初始化：

使用訓練好的 DIP 模型生成的初始圖像作為 DDPM 的起點，而不是隨機噪聲。這樣 DDPM 模型可以從一個已經包含圖像高級結構的初始圖像開始訓練。

Advantages:

- 加速訓練收斂過程:由於 DIP 提供的初始圖像已經包含了目標圖像的高級結構，DDPM 模型不需要從完全隨機的噪聲開始學習。因此，DDPM 模型可以更快地收斂，減少所需的訓練時間。
- 更好的初始點：DIP 生成的初始圖像為 DDPM 模型提供了一個更好的起點，使得生成的樣本質量在早期階段就能達到較高水平。
- 減少模糊：DIP 模型在初始階段捕捉到的高級結構有助於減少 DDPM 生成圖像中的模糊部分，提高圖像的清晰度和細節。
- 更穩定的訓練：從一個有意義的初始圖像開始訓練可以減少訓練過程中的不穩定性，使 DDPM 模型的訓練更加平穩和可靠

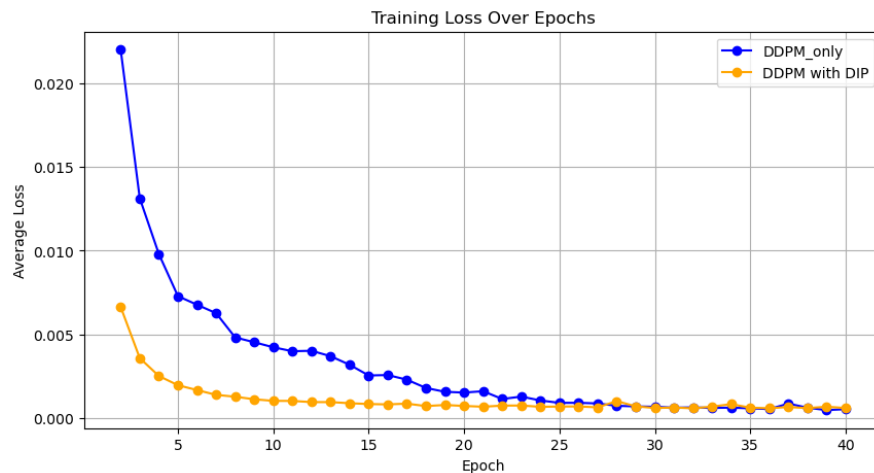
Potential limitation:

- 複雜度增加：將兩種方法結合將會增加模型的複雜性，需要更多的計算資源。
- 生成速度：將兩種方法結合可以提高生成圖片品質但在某些情況下，生成速度可能仍然是一個瓶頸。特別是對於實時應用，生成速度仍需優化。

在某些任務中，單獨使用一種方法可能已經很好，結合的方法優勢可能並不明顯，甚至可能帶來額外的計算負擔。

2. Experimental Verification

• Loss



上圖為訓練過程中 Loss 隨著 Epoch 改變的摺線圖，可以看到在初始的 loss 在 DDPM with DIP 的部分初始就較快且訓練過程中也較 DDPM_only 的部分較穩定，DDPM with DIP 的 Loss 程度也都低於 DDPM_only。

- Image Quality

DDPM-only:

	Image	PSNR	SSIM	MSE	MAE	Delta E
---	-----	-----	-----	-----	-----	-----
0	1	35.7304	0.983977	0.000267279	0.0131437	3.70649
1	2	28.778	0.984231	0.00132495	0.0336992	4.38799
2	3	27.667	0.972272	0.0017112	0.0349826	4.54233
3	4	26.7933	0.879174	0.00209252	0.0363661	7.82913
4	5	27.748	0.970887	0.00167956	0.0339818	5.1482

Average PSNR: 29.3433
 Average SSIM: 0.9581
 Average MSE: 0.0014
 Average MAE: 0.0304
 Average Delta E: 5.1228

DDPM with DIP:

	Image	PSNR	SSIM	MSE	MAE	Delta E
---	-----	-----	-----	-----	-----	-----
0	1	30.7499	0.970973	0.000841421	0.0228	4.00871
1	2	30.9817	0.984668	0.000797682	0.023292	3.22095
2	3	31.5878	0.984511	0.000693773	0.0214486	2.58559
3	4	26.6016	0.917525	0.00218698	0.0416067	5.17127
4	5	26.8675	0.979828	0.00205708	0.0358885	4.15306

Average PSNR: 29.3577
 Average SSIM: 0.9675
 Average MSE: 0.0013
 Average MAE: 0.0290
 Average Delta E: 3.8279

上面兩張圖中也可以看到我們抽取 5 張圖片，DDPM with DIP 相較 DDPM_only 在各項指標上有較好的水準。

3. Ablation Studies and Analysis

減少 DIP conv 通道數(64->16):

Conv channel 64:

	Image	PSNR	SSIM	MSE	MAE	Delta E
0	1	30.7499	0.970973	0.000841421	0.0228	4.00871
1	2	30.9817	0.984668	0.000797682	0.023292	3.22095
2	3	31.5878	0.984511	0.000693773	0.0214486	2.58559
3	4	26.6016	0.917525	0.00218698	0.0416067	5.17127
4	5	26.8675	0.979828	0.00205708	0.0358885	4.15306

Average PSNR: 29.3577
Average SSIM: 0.9675
Average MSE: 0.0013
Average MAE: 0.0290
Average Delta E: 3.8279

Conv channel 16:

	Image	PSNR	SSIM	MSE	MAE	Delta E
0	1	23.3959	0.801714	0.00457521	0.0530708	12.1531
1	2	21.5194	0.866331	0.00704784	0.0709071	11.799
2	3	19.6634	0.794672	0.010806	0.0851904	10.6776
3	4	14.7761	0.470916	0.0332962	0.157478	24.4508
4	5	25.2126	0.881641	0.00301118	0.0430393	9.17196

Average PSNR: 20.9135
Average SSIM: 0.7631
Average MSE: 0.0117
Average MAE: 0.0819
Average Delta E: 13.6505

從實驗結果可以看出，減少 conv 通道數可以降低模型的計算複雜度和訓練時間，但這會導致圖片品質的下降。通道數為 64 的模型在所有指標上均優於通道數為 16 的模型，具體表現為更高的 PSNR 和 SSIM，更低的 MSE、MAE 和 Delta E。因此，在實際應用中，需要在計算資源和生成影像品質之間找到平衡點，根據具體需求選擇合適的通道數。