

# GAI-project 4

F74102030 資訊114 練智剛

## Theoretical Justification

### Proposed Solution

#### DIP

先使用DIP對於原始資料集生成初始化的圖像，DIP的優點是不需要大量數據或者預訓練，且訓練速度快，是透過CNN的方式捕捉圖像結構。透過DIP生成初始化的圖像，能夠加強重點，提供了DDPM一個已經具備基本結構和內容的起點。DDPM在後續的去噪過程中，可以專注於細節的改進，而不是從頭開始生成圖像。

#### DDPM

然後使用DDPM來改進這一初始圖像。先在圖像中逐步加入noise，再逐步去噪生成逼真的圖像。從基於DIP的初始先驗開始，使得DDPM需要更少的去噪步驟來達到高品質圖像，從而加速生成過程。

## Design choices and Assumptions

### 結合DIP和DDPM：

- **DIP的優勢**：DIP利用CNN的內在結構從隨機noise生成圖像，無需大量訓練數據，這使其成為生成強初始先驗的理想選擇。
- **DDPM的優勢**：DDPM通過隨機過程生成高保真圖像，但由於需要大量的去噪步驟，計算成本較高。
- **結合**：通過用DIP生成的圖像初始化DDPM，減少了DDPM的負擔，使其專注於改進已經品質不錯的圖像，而不是從頭開始生成。兩種方法的互補優勢，DIP的快速結構捕捉能力和DDPM的詳細改進能力。

## Potential Benefits and Limitations

### Benefits

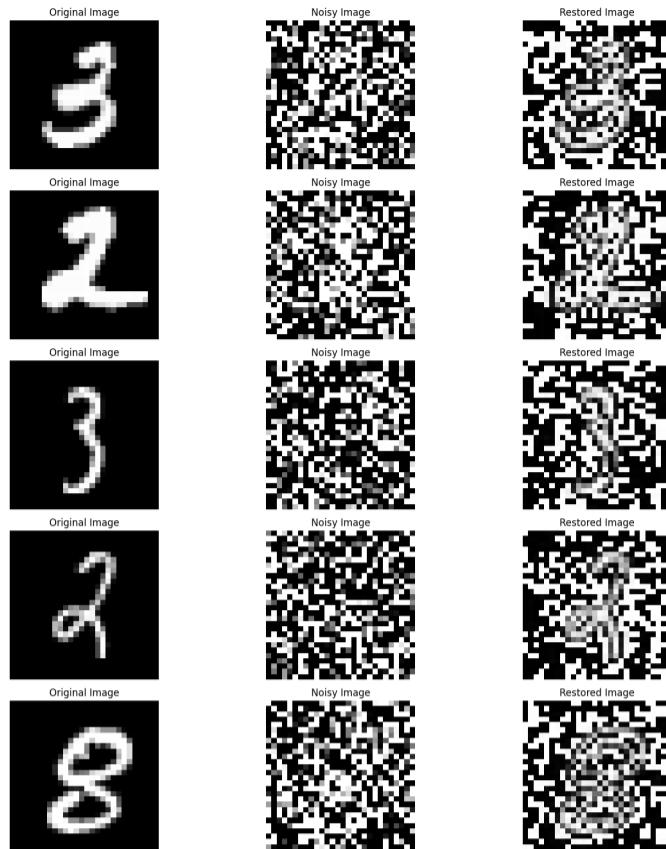
- **Speed:** 結合的方法顯著減少了DDPM所需的步驟數，從而加快圖像生成速度
- **Quality:** 從基於DIP的先驗開始，確保生成的圖像具有強大的結構基礎，DDPM可以進一步細化以達到高品質

### Limitations

- **Dependency on Initial Prior:** 最終圖像的質量在很大程度上取決於DIP生成的初始先驗。如果DIP先驗與目標分佈不一致，DDPM的性能可能會受到影響
- **Complexity:** 結合的方法在生成管道中引入了額外的複雜性，需要對DIP和DDPM組件進行仔細調整

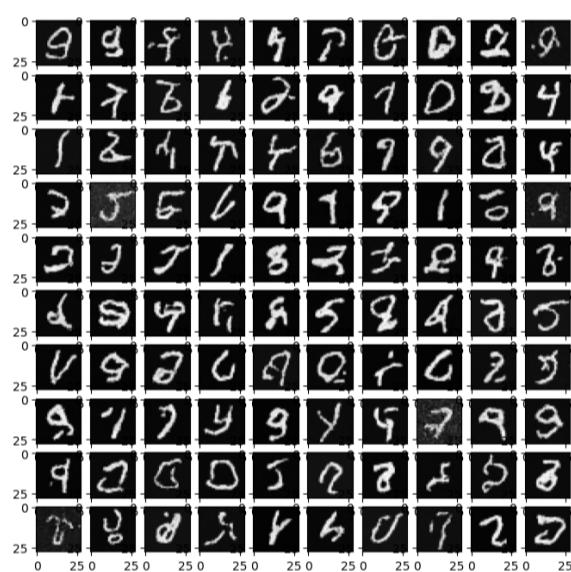
## Experimental Verification

### DIP alone



## DDPM alone

Final result

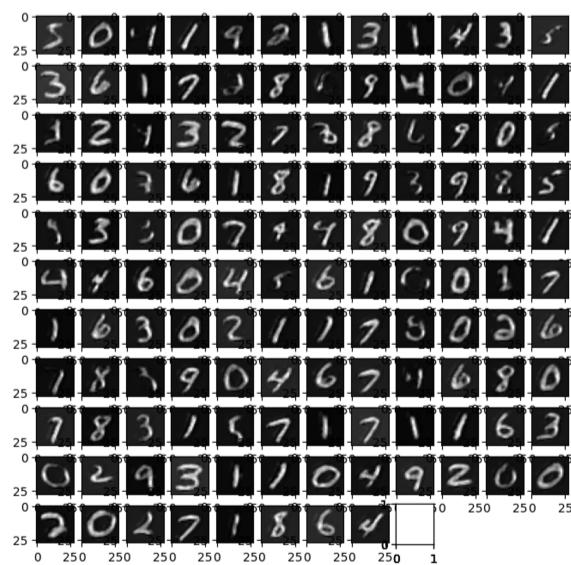


Training progress: 100%

20/20 [07:40<00:00, 22.93s/it]

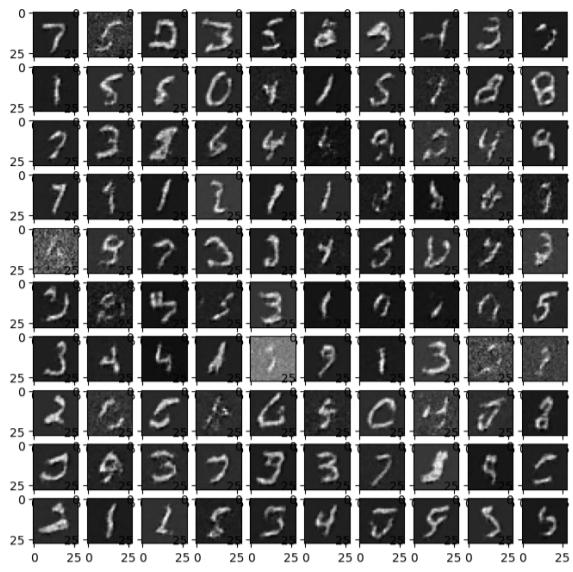
## Combine DIP and DDPM

經過DIP處理的圖片，可以看出重點部分被加強了



放到DDPM做訓練後的結果

## Final result



Training progress: 100%  20/20 [05:10<00:00, 15.49s/it]

可以發現時間縮短了，從原本的7分多鐘縮短成5分多鐘，達到原本預期的速度增快，生成出來的數字形體比單獨使用ddpm來的好，少掉了一些不規則的邊邊，但是缺點是生成出來的數字顏色比較灰階，相較之下對比比較沒有那麼高。

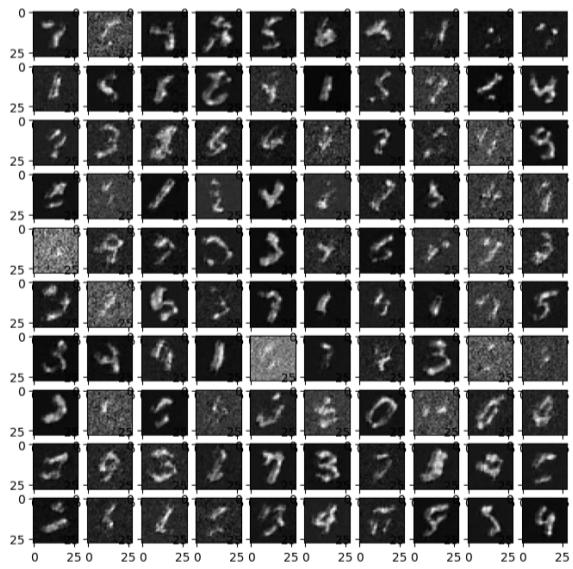
## Ablation Studies and Analysis

比較不同 `dip_lr` 和 `ddpm_lr` 的參數，以及在使用高斯加噪的過程中 `std` (標準差)的對於生成速度以及生成圖片品質的影響

**`dip_lr=0.001 ddpm_lr=0.0005 std=1.2`**

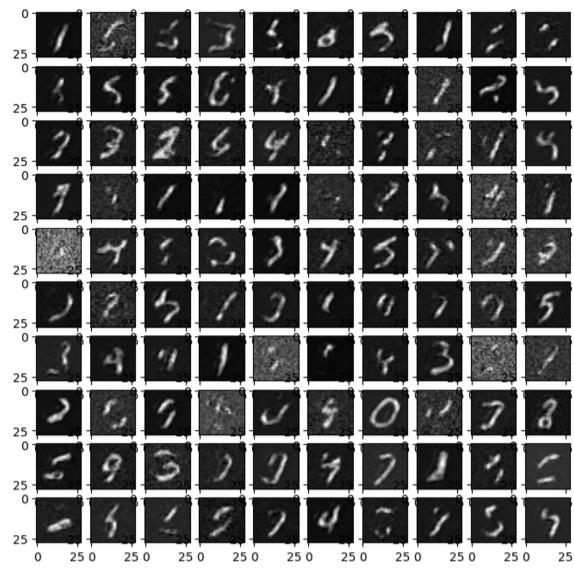
Training progress: 100%  20/20 [05:07<00:00, 15.28s/it]

## Final result



**`dip_lr=0.001 ddpm_lr=0.0005 std = 1`**

Training progress: 100%  20/20 [05:10<00:00, 15.49s/it]

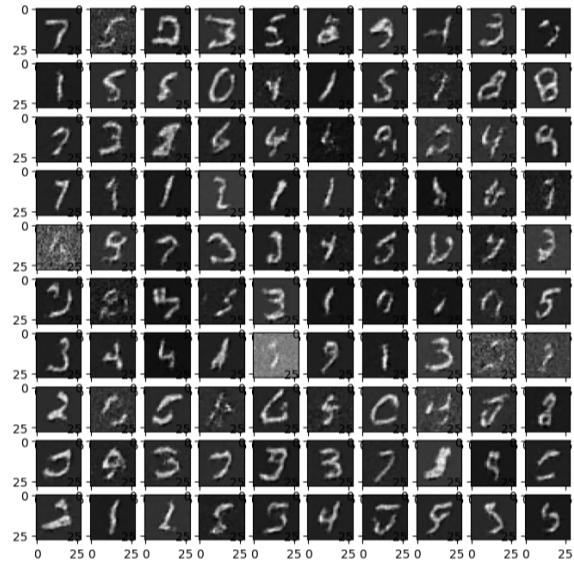


**dip\_lr=0.001 ddpm\_lr=0.0005 std=0.8**

Training progress: 100%

20/20 [05:03<00:00, 15.19s/it]

Final result



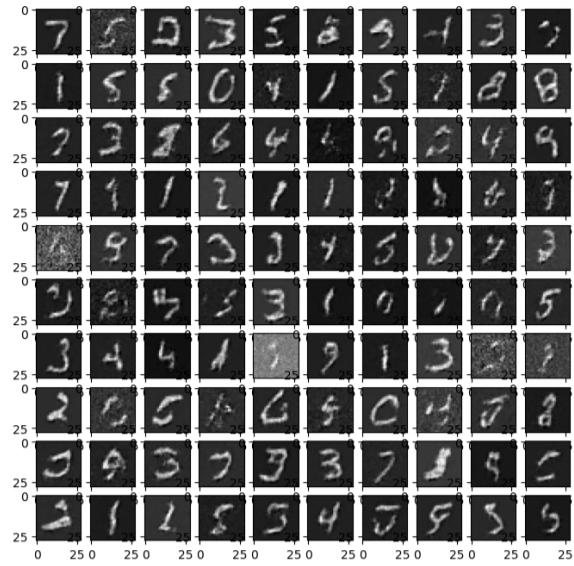
從以上三組結果可以推測，std越高生成出來的圖片越不精準，因為std越大代表所加的noise偏差越大，在還原上有比較大的難度，而在training的時間上都差不多

**dip\_lr=0.0005 ddpm\_lr=0.001 std=1.2**

Training progress: 100%

20/20 [05:05<00:00, 15.32s/it]

Final result

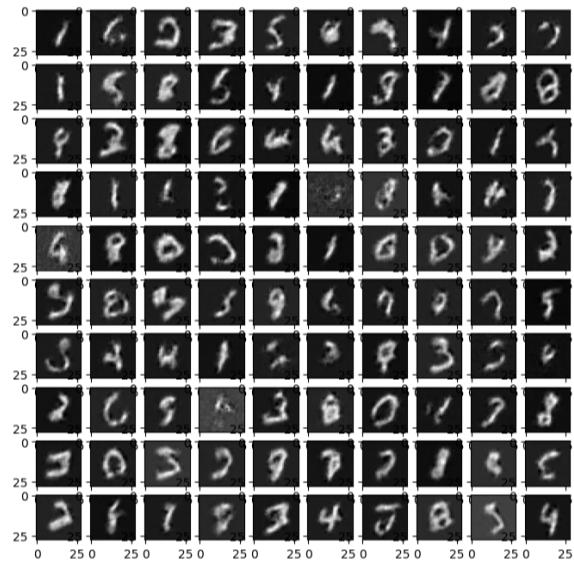


dip\_lr=0.002 ddpm\_lr=0.001 std=1.2

Training progress: 100%

20/20 [05:05<00:00, 15.27s/it]

Final result

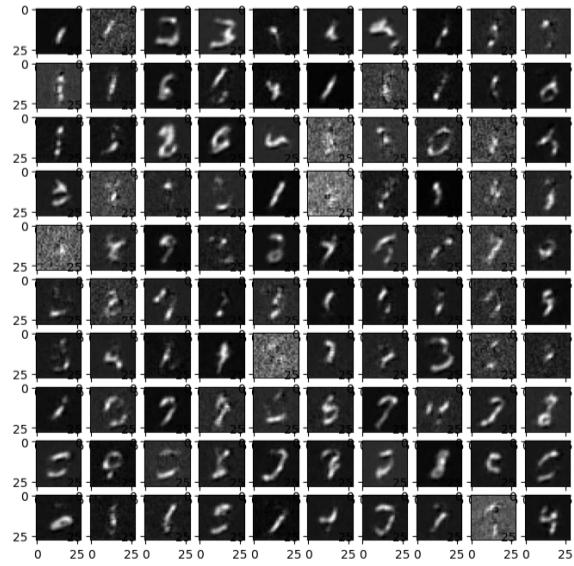


dip\_lr=0.005 ddpm\_lr=0.001 std=1.2

Training progress: 100%

20/20 [05:08<00:00, 15.41s/it]

Final result



改變dip\_lr可以發現當learning rate越大的時候生成出來的圖像越不清楚