

# Pix2Pix 與 SinGAN 的訓練應用

走進經典電影裡的世界

N76101012 謝慈芯

Pix2Pix: Image-to-image translation with a conditional GAN

--Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros

SinGAN: Learning a Generative Model from a Single Natural Image

--Tamar Rott Shaham, Tali Dekel, Tomer Michaeli

2021/01/12

# 目錄

---

- 01 研究背景
- 02 研究動機
- 03 研究目的
- 04 理論與原理
- 05 實驗設計
- 06 實驗結果與討論
- 07 結論與心得感想

AI人工智慧世界下的電影

---

# 研究背景 01

# AI人工智慧世界下的電影

- 前些陣子在疫情之下，許多電影院關門，因為疫情影響所以電影業處於蕭條期。前陣子透過網路看了2021年萊恩雷諾斯（Ryan Reynolds）主演的《脫稿玩家》（Free Guy），**講述關於不斷變化的遊戲和 AI 人工智慧世界。**
- 電影設定提出了對 AI 的想像，已經有許多電影像這樣**用電影探討 AI 的可能性**，例如《駭客任務》、《全面啟動》、《一級玩家》、《機械公敵》等等，很多都是經典。
- 讓我不禁想像，那麼如果利用現在的 AI 技術，**創造出我進入電影裡的影像**會是甚麼樣？



Pix2Pix / SinGAN

---

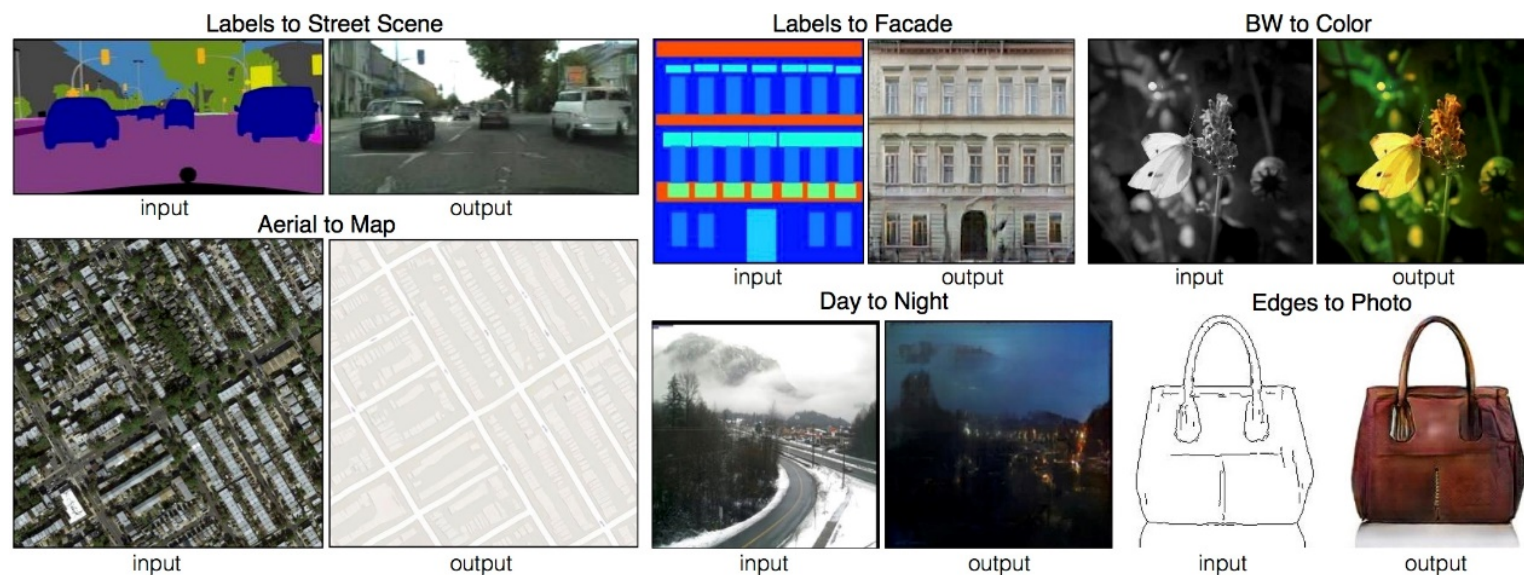
# 研究動機 02

# Pix2Pix

Pix2Pix: Image-to-image translation with a conditional GAN

--Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros

- 本研究從這篇論文啟發，這篇論文旨在研究如何構建和訓練一個稱為 **pix2pix** 的條件生成對抗網絡 (cGAN)，該網絡學習從**輸入圖像**到**輸出圖像**的圖像轉換訓練。
- pix2pix 它可以應用於廣泛的任務，包括從標籤**地圖合成照片**、從**黑白圖像生成彩色照片**、將**谷歌地圖照片轉換為衛星航拍圖像**、或是**將草圖轉換為照片**。



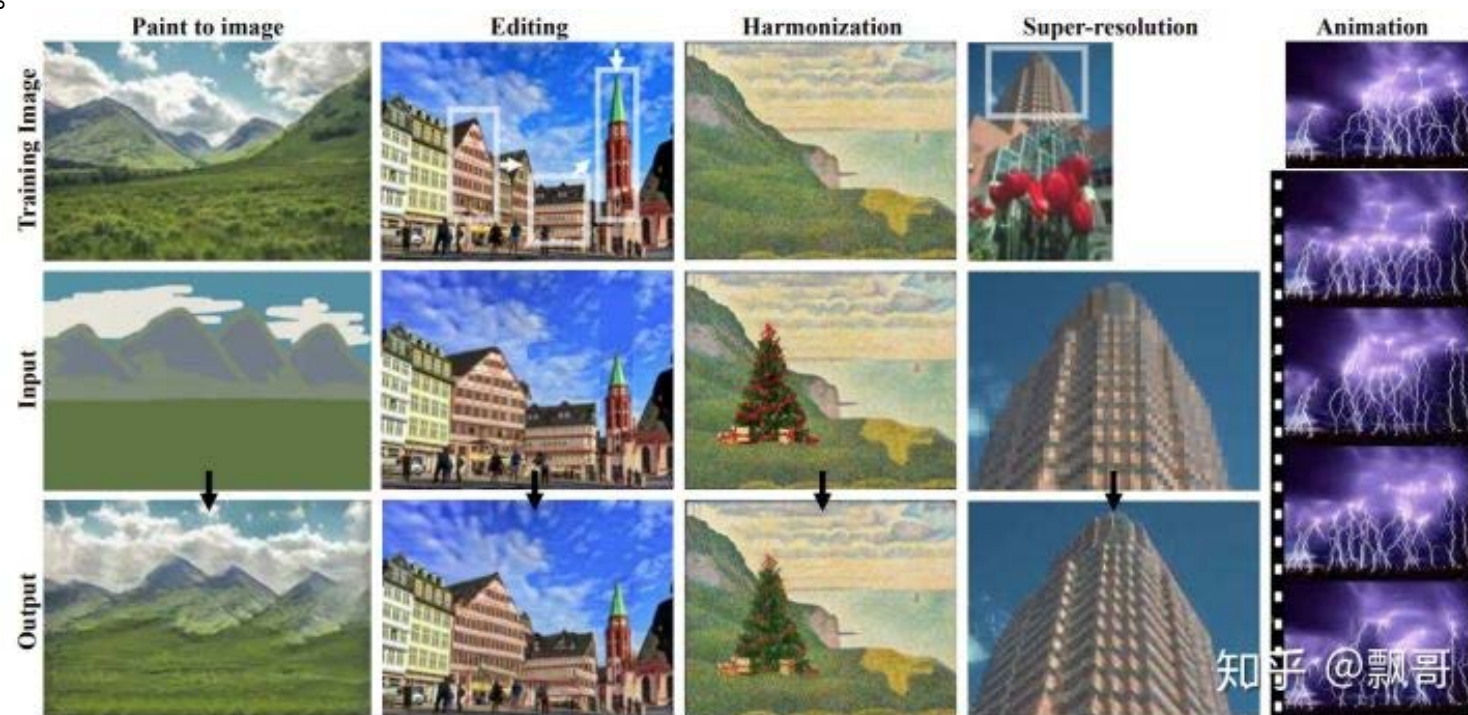


# SinGAN

## SinGAN: Learning a Generative Model from a Single Natural Image

--Tamar Rott Shaham, Tali Dekel, Tomer Michaeli

- 本研究第二個運用到 SinGAN 做影像融合，這篇論文作者來自以色列理工學院 (Technion) 和谷歌 (Google)。
- 一般來說訓練 GAN 來生成圖像，至少需要有成千上萬的訓練數據才可能有不錯的效果。SinGAN 只需要單張自然圖像，就可以訓練一個生成式模型。



情境 / 步驟

---

# 研究目的 03



# 目的

現在的電影因為科技的影響，可以做到像《機械公敵》那樣科幻的景象，讓我想到了，既然現在電影都會出現 AI 了，那麼如果讓我們**利用 AI 進到電影裡面跟電影融合**呢？

- 情境： 本計畫設定了兩個情境
  1. 一是透過訓練進入**過去**的電影中。
  2. 二是透過訓練進入帶有**未來**感的彩色電影中。
- 步驟： 本計畫根據兩步驟
  1. 一是透過 **Pix2Pix** 訓練將黑白的經典電影片段重新製作成**彩現**。
  2. 二是將事先拍攝好的我的照片經過 **SinGAN** 訓練後**融合**進電影裡。

Pix2Pix 原理 / SinGAN 原理

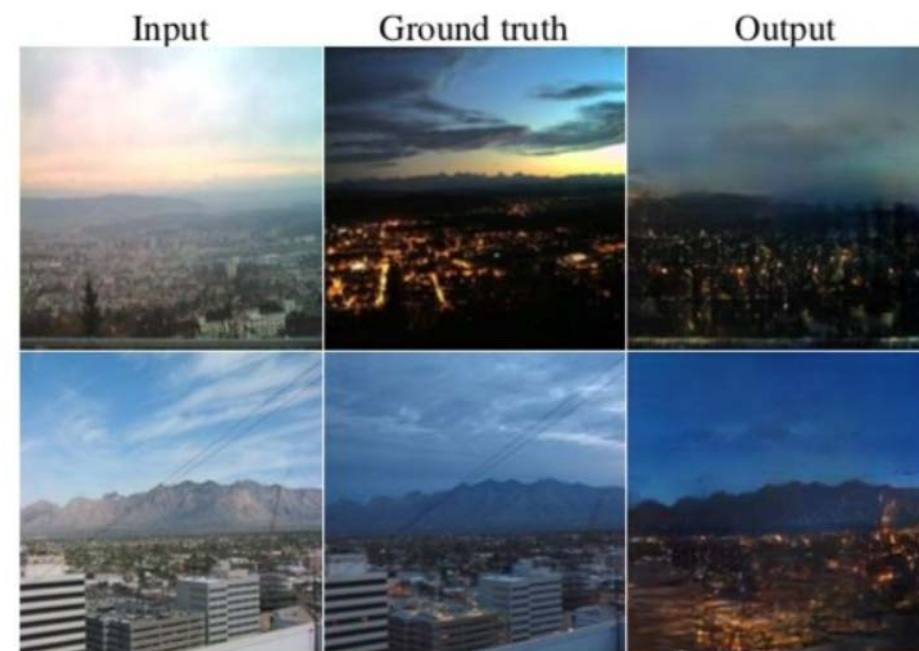
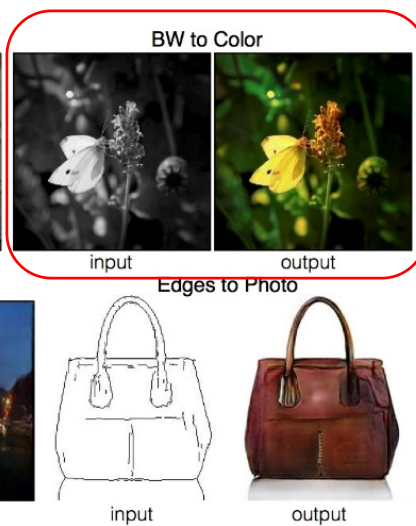
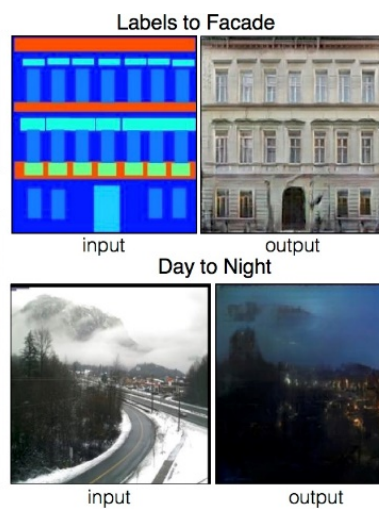
---

# 理論與原理 04

# 理論與原理

## 步驟一：Pix2Pix 原理

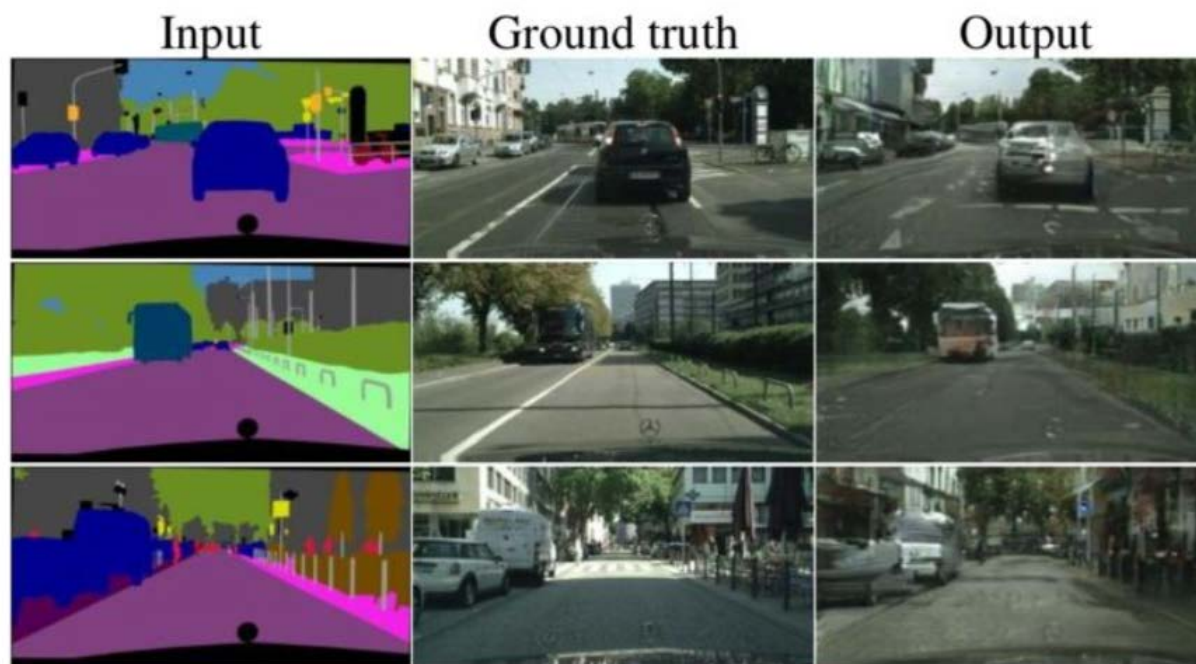
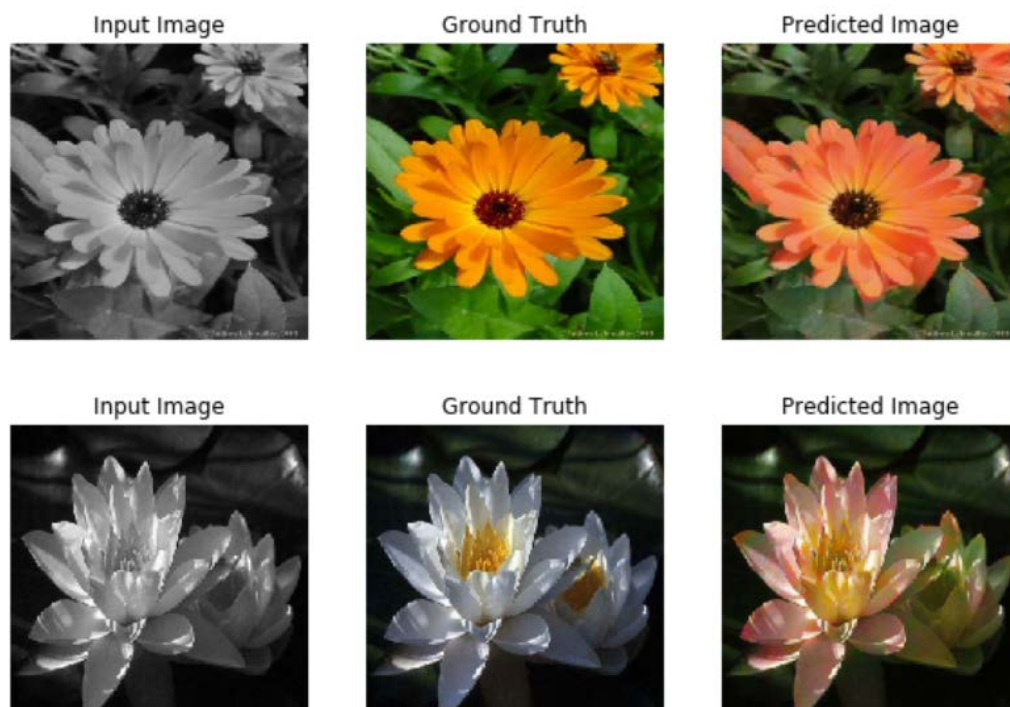
- 本企劃將運用到 Pix2Pix 其中的一個訓練功能---**黑白影像採現(BW to color)**
- Pix2pix 是用 cGAN 做 image translation，其原理是以輸入圖像為條件並生成相應的輸出圖像。



# 理論與原理

## 步驟一：Pix2Pix 原理

- 為了方便可以使用由 pix2pix 作者創建的數據集去做訓練處理。

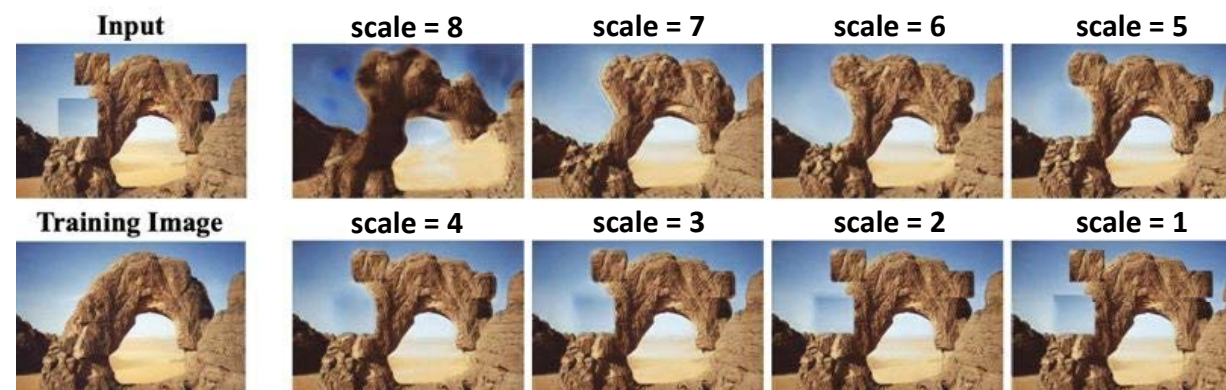
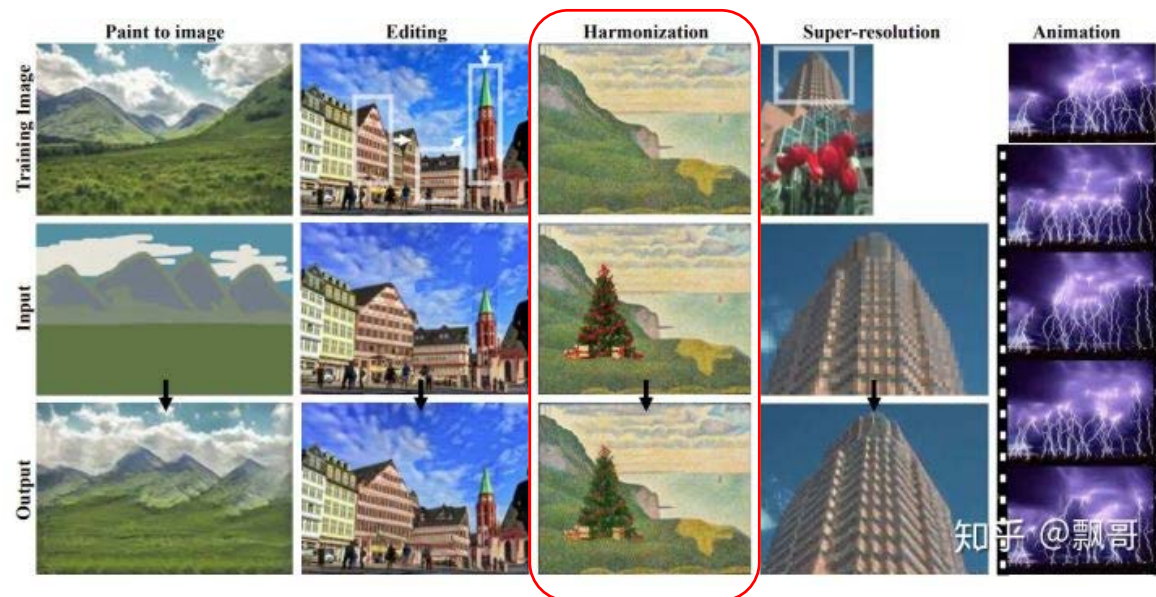




# 理論與原理

## 步驟二：SinGAN 原理

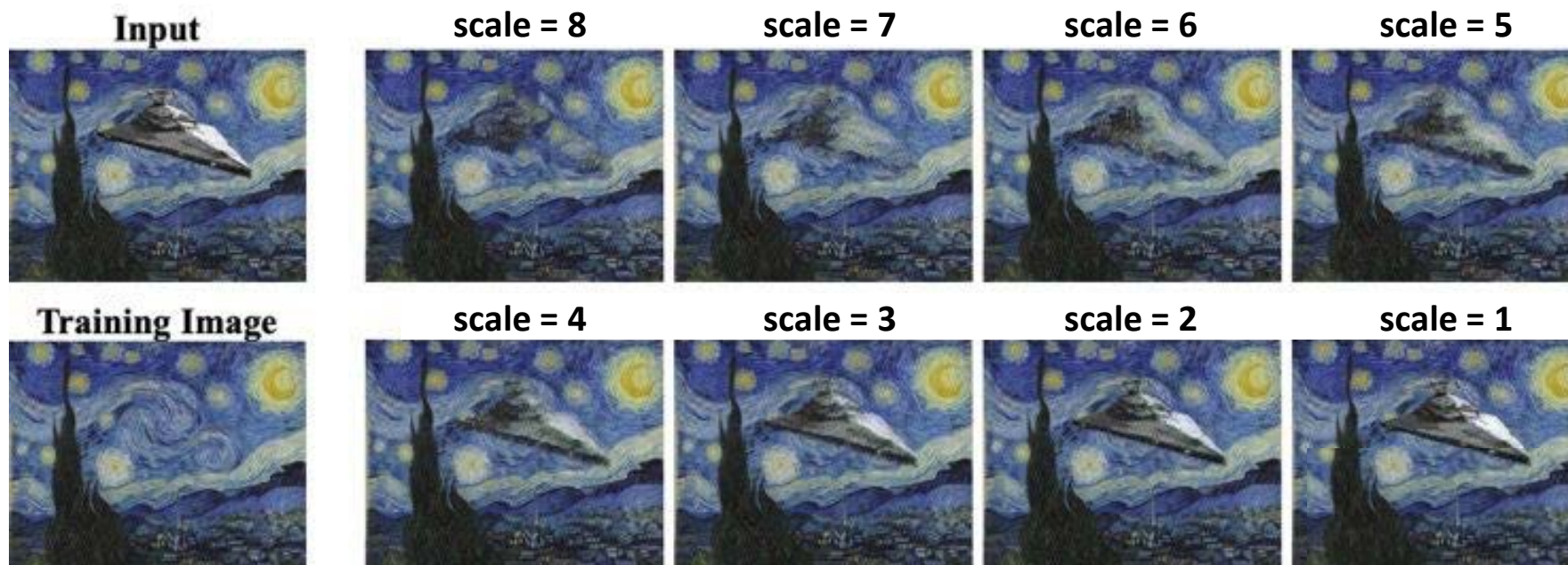
- 本企劃將運用到 SinGAN 的其中一個訓練功能---**圖像和諧化 (Image Harmonization)**
- 其原理是在一張圖像上訓練 SinGAN，將經過編輯（下例中為圖像方塊的移動）的圖像輸入生成器，輸出的結果**恢復了圖像的真實感**。



# 理論與原理

## 步驟二：SinGAN 原理

- **圖像和諧化 (Image Harmonization)** 是在圖像 INPUT 中加入 “不和諧” 因素（如一個 “不和諧” 的飛船），輸入某中間尺度生成器後，輸出的結果實現了圖像整體風格的和諧化。
- 下圖比較了輸入不同層所產生輸出的區別。





實驗方法 / 系統架構 / 實驗步驟

---

# 實驗設計 05

# 實驗方法

情境一：透過訓練走進**過去電影場景**

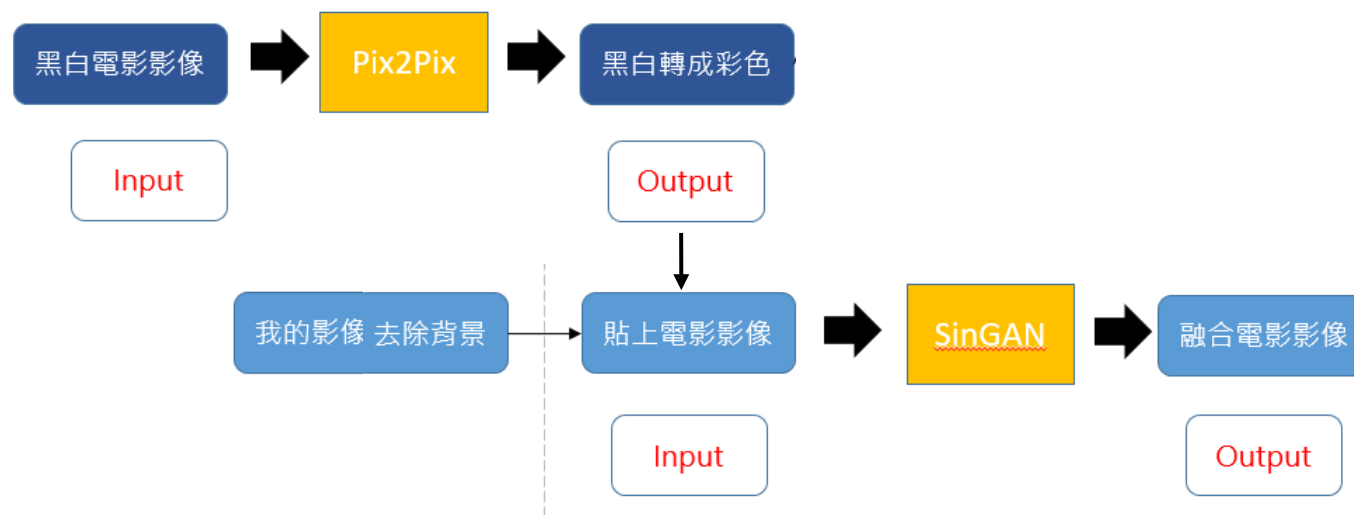
本企劃預先利用 **Pix2Pix** 透過訓練**將黑白電影場景轉成彩色**，再將事先拍攝好的我的照片去除背景後貼上電影場景影像，再利用 **SinGAN** 的**融合訓練模型**將我融入電影當中。

情境二：透過訓練走進**未來電影場景**

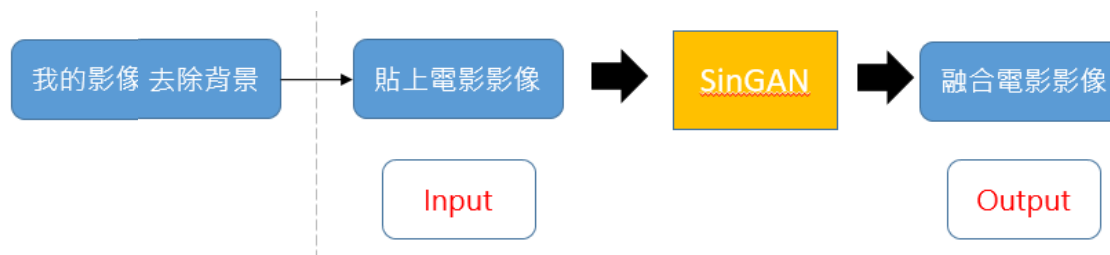
本企劃預將事先拍攝好的我的照片去除背景後，貼上具有未來想像的**AI**電影片段，再利用 **SinGAN** 的**融合訓練模型**將我融入電影當中。

# 系統架構

- 情境一：透過訓練走進過去電影場景



- 情境二：透過訓練走進未來電影場景



# 實驗步驟

## 步驟一：Pix2Pix 訓練過程

- Model Training

本訓練使用 **Pix2Pix** 原論文作者的數據集，設定訓練 50 個 epoch，利用 colab 進行訓練。

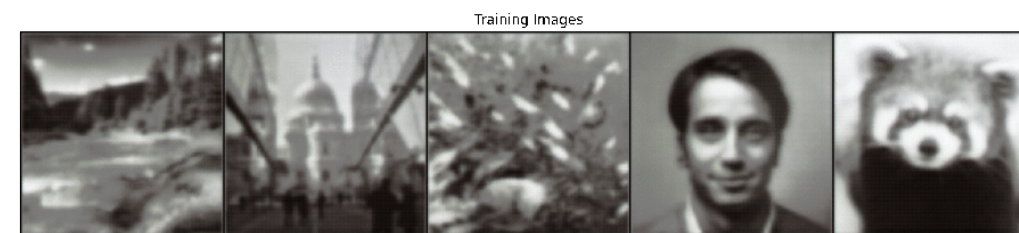
```
for epoch in range(initial_epoch, 50):  
    running_loss_D = 0.0  
    running_loss_G = 0.0  
    for i, data in tqdm(enumerate(generator)):  
  
        L, ab = data[0]['L'].to(device), data[0]['ab'].to(device)  
        fake_color = net_G(L)  
        real_image = torch.cat([L, ab], dim=1)  
        if nb_of_output_channels == 2:  
            fake_image = torch.cat([L, fake_color], dim=1)  
        elif nb_of_output_channels == 3:  
            fake_image = fake_color  
        rgb = data[1].to(device)
```



Input



processing



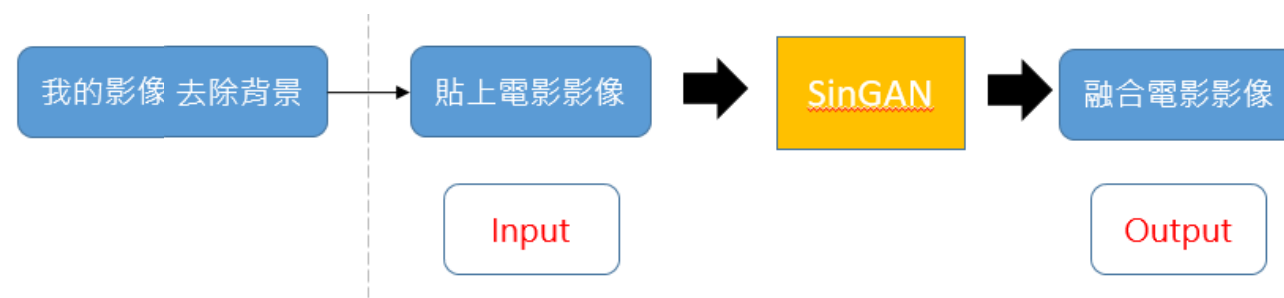
# 實驗步驟

## 步驟二：SinGAN 訓練過程

- Model Training

本訓練使用 SinGAN 原論文的程式，因此只需要一張原始自然圖像、一張合成(未融合)的目標圖像、一張合成地方的遮罩。

由於 colab 的 GPU 已過量，超過個人用量，因此本階段則是使用本機 GPU 進行訓練。



Input



Source



Target



Mask

processing



訓練過程與成果

---

# 實驗結果與討論 06



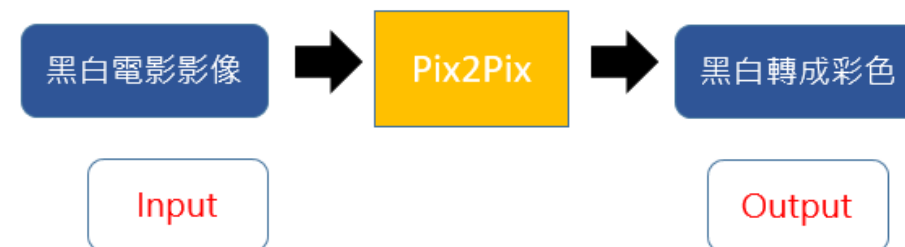
# 訓練過程

## 步驟一：Pix2Pix 訓練過程

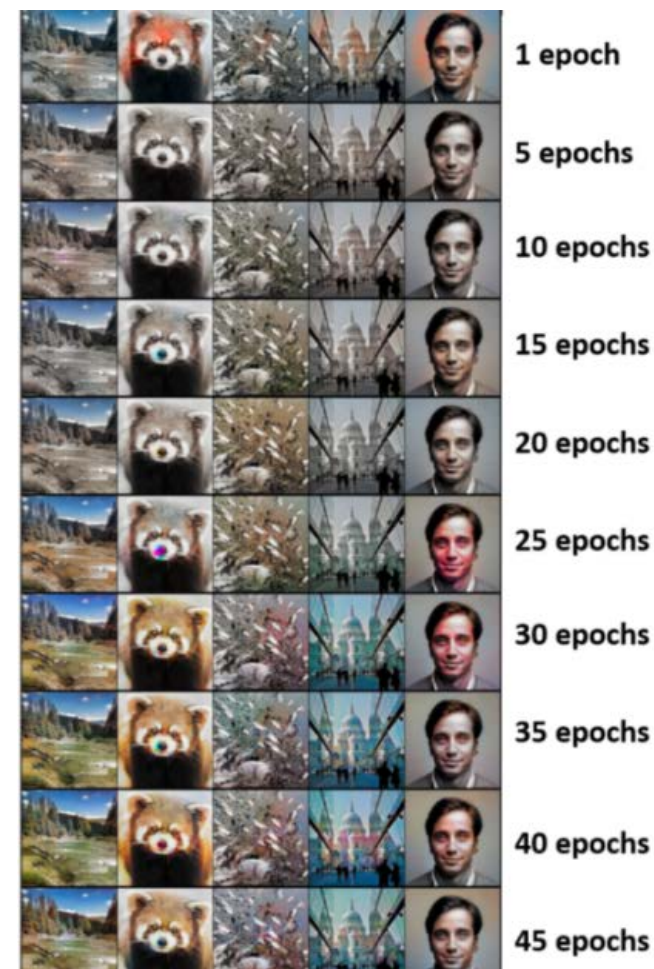
### 1. Model Training

對於 Pix2Pix 訓練出來的模型，一開始圖像以棕色為主。即是繪畫中的三種原色（紅、黃、藍）的混合物變成棕色。

在 45 個 epoch 之後，就可以看到圖像沒有太大的變化。在訓練學習過程，有些圖片明顯模糊。



processing



# 訓練成果

## 步驟一：Pix2Pix 訓練成果

### 2. Model Testing

```
!python pix2pix.py --mode test --input_dir /Users/Gene/myImages/ --output_dir /Users/Gene/myGeneratedImages --checkpoint /Users/Gene/myModel
```

Input



Output



電影：《卓別林》(charlie chaplin)

# 訓練過程

## 步驟一：SinGAN 訓練過程

### 1. Model Training

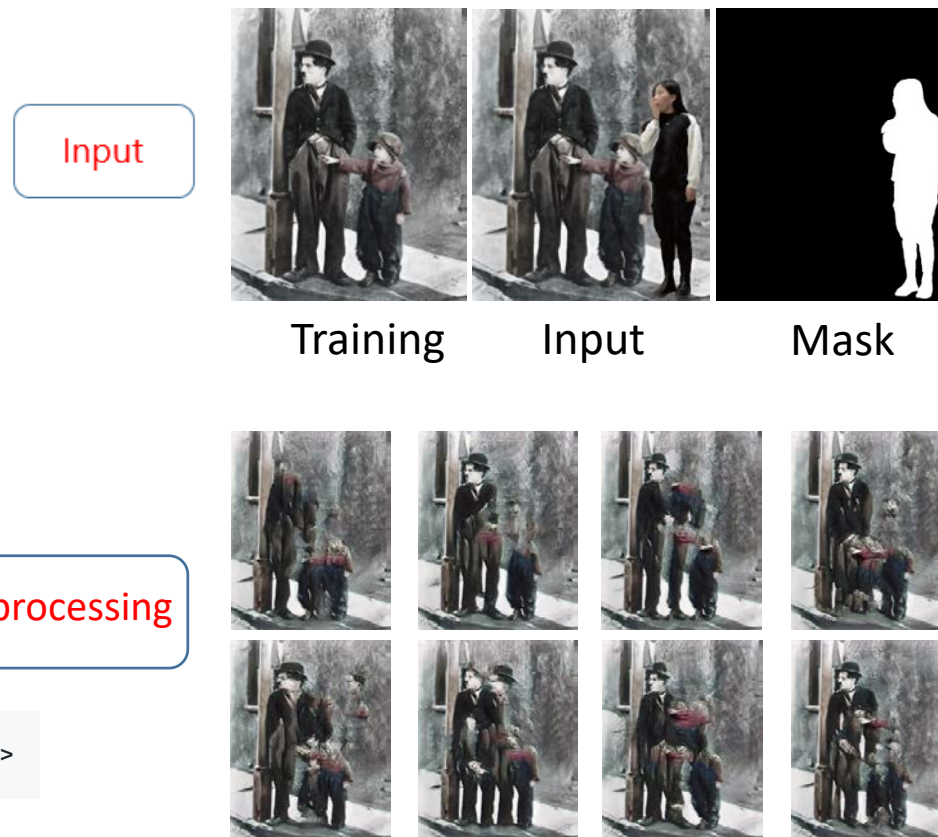
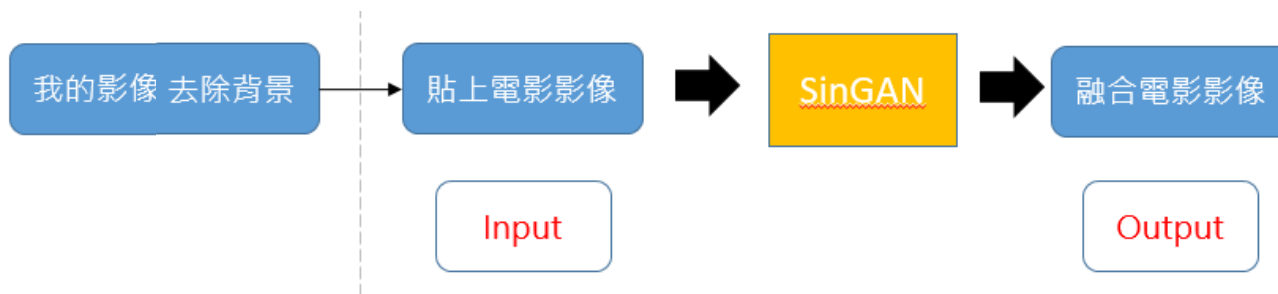
對於 SinGAN 訓練出來的模型，使用前需要先將 Input Training、Input 以及 Mask 的圖片各一張。

要從任何起始生成需先運行隨機樣本 RandomSamples：

```
python random_samples.py --input_name <training_image> --mode random_samples --gen_start_scale <start scale>
```

會 Output 出數張 RandomSamples 圖像，再接著運行 Harmonization 並填上 scale 值

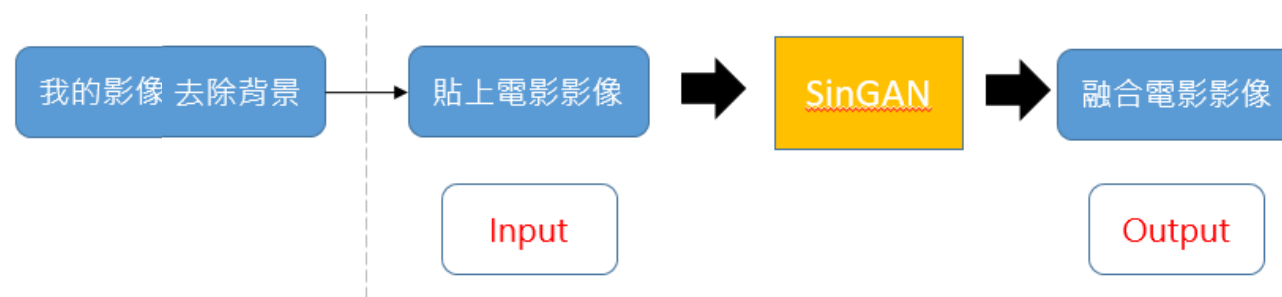
```
python harmonization.py --input_name <training_image> --ref_name <pasted_reference_image> --harmonization_start_scale <scale>
```



processing

# 訓練成果

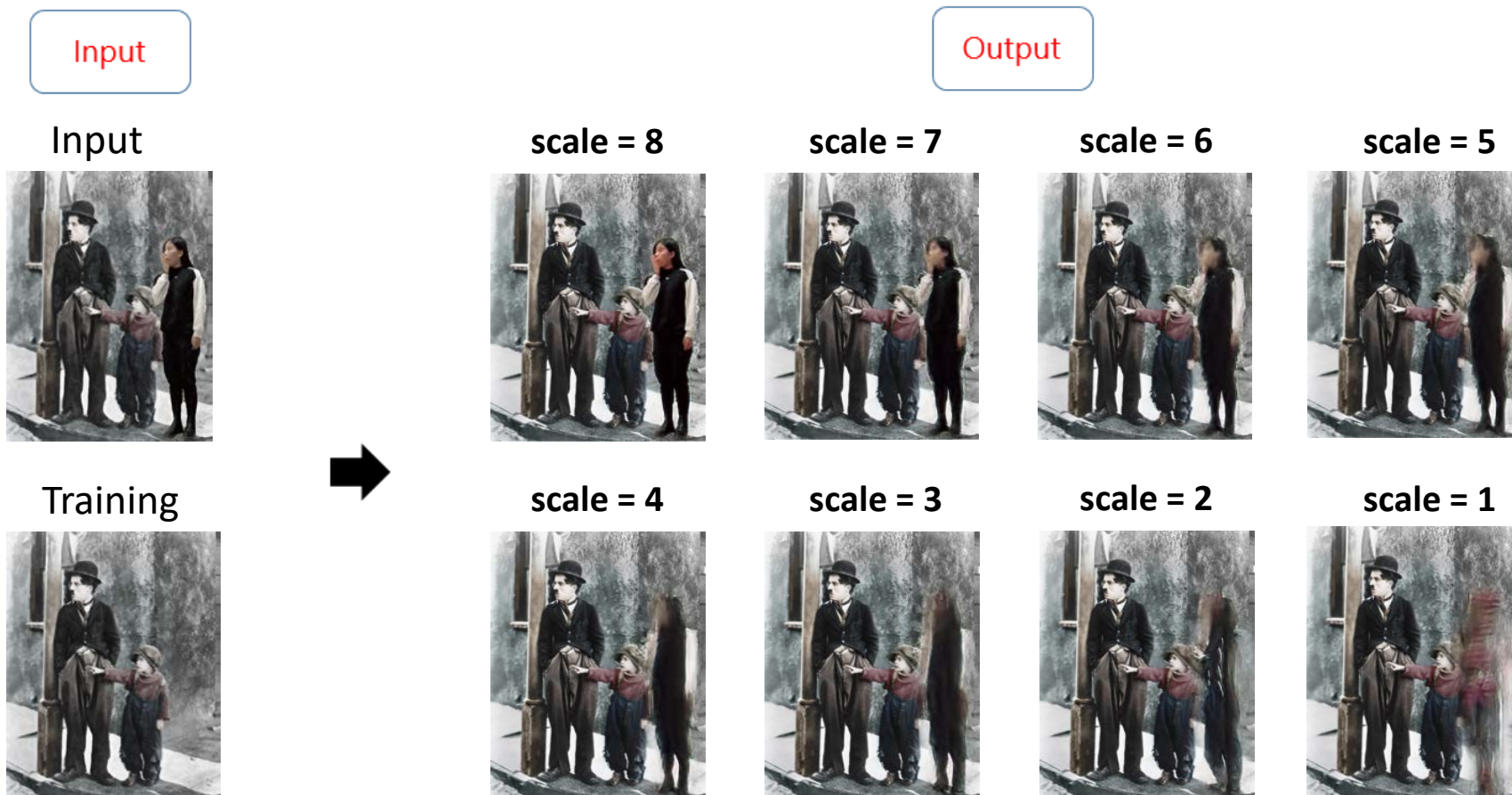
## 步驟一：SinGAN 訓練成果



### 2. Output

過去電影片段:

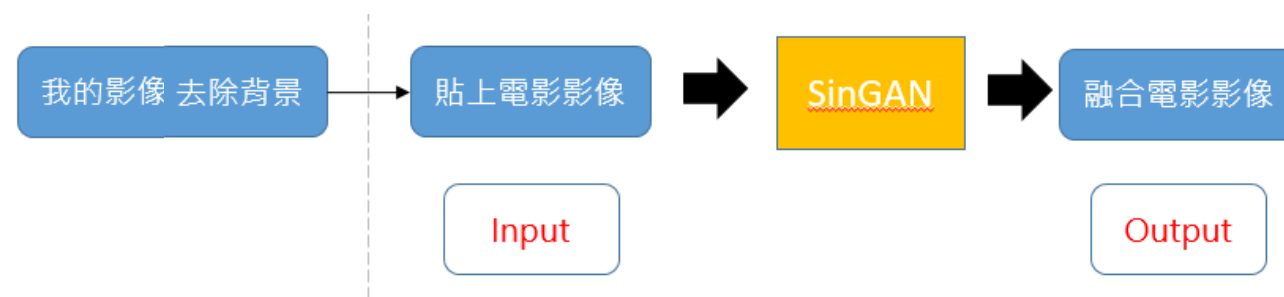
《卓別林》  
(Charlie Chaplin)





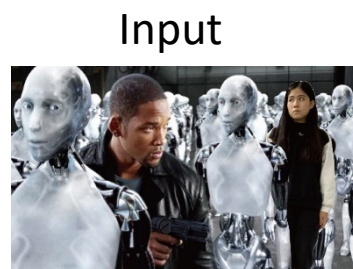
# 訓練成果

## 步驟一：SinGAN 訓練成果



### 2. Output

Input



Output

scale = 8



scale = 7



scale = 6



scale = 5



Training



scale = 4



scale = 3



scale = 2



scale = 1



未來電影片段:

《機械公敵》  
(I, Robot)

執行過程問題與未來展望

---

# 結論與心得感想 07



# 執行過程問題與未來展望

## 過程問題

### Google Colab

一開始透過 Colab 訓練是因為可省去安裝大量函式庫的步驟，它提供免費GPU使用，也可以與 Google Drive 連結，直接存取/修改 Google Drive 裡的檔案比較方便。

然而過程遇到執行時需保持瀏覽器窗口開啟，與GPU有「限量」的問題，且須定期確認是否還留在線上，使得換了兩個帳號之後才訓練完第一階段。因此第二階段的 SinGAN 單張圖像訓練便使用了本機電腦的 GPU 訓練。

## 未來展望

目前實驗成果是做成影像，也有嘗試將訓練過程製成 gif 檔，未來也許可以取電影片段每一幀都擷取形成動畫檔。我認為在 AI 的發展與想像中，使電影加入真實的人像透過 AI 模型訓練可以產生更多想像與有趣的反應來做討論。

資料來源

---

# Reference 08

# Reference

pix2pix: Image-to-image translation with a conditional GAN

<https://www.tensorflow.org/tutorials/generative/pix2pix>

Image-to-Image Translation with Conditional Adversarial Networks

<https://arxiv.org/abs/1611.07004>

生成對抗網絡系列(4)——pix2pix

<https://zhuanlan.zhihu.com/p/38411618>

SinGAN: Learning a Generative Model from a Single Natural Image

[https://openaccess.thecvf.com/content\\_ICCV\\_2019/papers/Shaham\\_SinGAN\\_Learning\\_a\\_Generative\\_Model\\_From\\_a\\_Single\\_Natural\\_Image\\_ICCV\\_2019\\_paper.pdf](https://openaccess.thecvf.com/content_ICCV_2019/papers/Shaham_SinGAN_Learning_a_Generative_Model_From_a_Single_Natural_Image_ICCV_2019_paper.pdf)

ICCV2019最佳論文SinGAN全面解讀

[https://www.sohu.com/a/362210741\\_610300](https://www.sohu.com/a/362210741_610300)

# THANKS FOR LISTENING

N76101012 謝慈芯

Pix2Pix 與 SinGAN 的訓練應用

--走進經典電影裡的世界

2021/01/12