

# 中華民國第 59 屆中小學科學展覽會 作品說明書

---

高級中等學校組 電腦與資訊學科

佳作

052510

幻想濾鏡：人工神經網路實現手機拍照增強

學校名稱：新北市立中和高級中學

作者： 高二 鍾天睿 高一 柯勁廷	指導老師： 王一哲
-------------------------	--------------

關鍵詞：人工智慧、神經網路、影像處理

# 摘要

自動影像處理一直是個熱門的話題，目前已提出許多提升手機拍攝照片品質的方法。在這份研究中，我們建立 GAN 模型，嘗試消除手機於夜間拍照時照片中的噪點，並且提升照片畫質。我們先用非成對數據訓練 CycleGAN，利用 CycleGAN 由高品質圖片產生對應的低品質圖片，再使用此數據訓練生成器，藉此達到降噪以及提升畫質的效果。藉由我們的模型，使用者可以透過簡易的方式，有效地去除手機拍攝夜景照片中，惱人且難以消除的噪點。據我們所知，我們的研究是目前最方便且有效的方式，為自動影像處理開創嶄新的道路。

## 壹、研究動機

現代人常常使用手機拍照，記錄生活中的點滴。但受限於手機感光元件大小，特別是在夜間或低亮度的環境拍攝，時常無法獲得較好的畫質。因此我們想建構可以用於自動提升畫質、修正色調的軟體，讓使用者可以方便地處理照片。

## 貳、研究目的

- 一、探討使用神經網路處理圖片之可行性
- 二、探討使用非成對手機拍攝照片及高品質照片訓練神經網路之可行性
- 三、訓練增強畫質的模型
- 四、嘗試建構圖形化使用者介面工具

## 參、研究設備及器材

### 一、訓練數據集

- (一) 由 Unsplash 下載的夜景照片（不包含星空，約 1000 張）
- (二) 自行拍攝的高品質夜景照片（使用數位單眼相機拍攝，約 160 張）
- (三) 自行拍攝的低品質夜景照片（使用 HTC U11/U11+ 拍攝，約 700 張）
- (四) DPED Dataset [1]

### 二、訓練設備及框架

由於訓練生成對抗網路 (Generative Adversarial Network, GAN) 及其相關神經網路需要占用大量顯示記憶體，我們的模型均在以下設備訓練：

#### (一) 工作站電腦

1. 顯示卡：GEFORCE® RTX 2080 Ti
2. 記憶體：32GB DDR4

#### (二) 作業環境

为了更好的相容性並善用相關工具，我們採用 Ubuntu 18.04 LTS 為作業系統，並使用 ssh 進行遠端連線觀察訓練過程。

#### (三) 神經網路框架、工具

1. Python 3.6 （主要程式語言）
2. TensorFlow-GPU 1.13 （人工神經網路框架）

## 肆、研究過程或方法

### 一、名詞解釋

表 1 特殊名詞

中文	英文	解釋
機器學習	machine learning	人工智慧的其中一個研究分支
激勵函數	activation function	用以增加神經網路的非線性。
損失函數	loss function	衡量神經網路產生內容與實際內容差距的函數。
卷積神經網路	convolutional neural networks, CNN	經常用於圖形辨識，利用卷積層及池化層提取圖形的特徵。
生成對抗網路	generative adversarial network, GAN	人工神經網路的其中一種，常用於產生內容，運作方式為透過讓兩個神經網路相互對抗的方式進行學習。
迭代	iteration	將資料輸入至人工神經網路訓練的過程。
批量	batch	指固定數量的數筆資料。
批量大小	batch size	一次輸入多少筆資料。
全局步數	global step	不計 epoch，一共訓練了幾次批量。
端對端訓練	end-to-end training	使用端對端的方式進行訓練。
假影	artifact	電腦重組的圖像與原圖像的誤差部分。例如電腦斷層掃描所截取的圖像，可能由於放射線的不穩定、人體的移動或是散射干擾所產生。

## 二、文獻探討

### (一) 影像降噪

影像降噪的目的是去除圖像中的噪點，同時保留圖片中較為重要的細節。即使是良好的數位感光元件或傳統底片，拍攝的照片仍然有可能產生雜訊，因此降噪的技術被廣泛利用。

#### 1. 訊號雜訊比 (Signal to Noise Ratio, SNR)

訊號雜訊比的定義如下：

$$SNR = \frac{\sigma(u)}{\sigma(n)}$$

其中  $\sigma(u)$  所代表的為訊號本身的標準差，而  $\sigma(n)$  則是雜訊的標準差。

#### 2. 降噪方法

##### (1) 高斯平滑 (高斯模糊)

高斯平滑的原理在於將影像和高斯濾波器進行卷積，藉此使影像模糊以去移除雜訊以及細節。雖然可以有效去除噪點，但會使影像變得模糊。

##### (2) 非局部平均平滑

實作此方法時，對各個目標像素周圍定義一個區域，並且對整個影像的所有像素，依照該像素周圍區塊的區塊與目標像素區塊的相似度給予權重並進行平均。這種算法可保留較多的細節，清晰度較高，但時間複雜度較高，需要較多的時間完成計算。

## (二) 人工神經網路 (Artificial Neural Network, ANN)

這是機器學習領域當中一種常用的方法，原理為藉由電腦程式模擬生物的神經網路系統，藉此處理一些傳統程式設計上難以解決的問題，例如圖形辨識和語音辨識。

人工神經網路通常由許多層所組成，常見的層有以下幾種：

### 1. 全連接層 (Fully Connected Layer)

全連接層又名稠密層 (dense layer)，是最簡單的網路結構之一。其中有數個權重 (weight) 以及偏移值 (bias)，用以控制各個神經元連接的方式。

### 2. 卷積層 (Convolutional Layer)

卷積層由許多卷積單位所構成，每個卷積單位的參數會由反向傳播最佳化。卷積運算的目標是提取的不同層級特徵，第一層卷積層可能只提取一些低階的特徵，例如邊緣、線條……等層級，而更深的結構能夠提取更複雜的特徵。

## (三) 高階人工神經網路結構

下方列出目前常見的人工神經網路結構：

### 1. 卷積神經網路 (Convolutional Neural Networks)

卷積神經網路一般包含數個卷基層，以及少數全連接層，常用在圖形分類問題中。卷基層會將圖形轉換成特徵圖，並經由全連接層輸出結果。

### 2. 全卷積網路 (Fully Convolutional Networks, FCN)

所有的網路結構都由卷積層及池化層 (pooling layer) 組成，不包含全連接層，或是將轉換後的特徵圖經由上採樣解碼為圖片，早期被應用於圖像語意分割。

### 3. U-Net [2]

U-Net 可以理解為全卷積網路的改良版，其原理為將較淺層的特徵圖連接至上採樣解碼的卷積層。利用這種方式可以有效地轉換高、低層級的特徵圖，並且具有訓練較為迅速、在較深的網路結構中容易訓練等特性。在圖像語意分割問題中，使用這種方式可以獲得比較好的輸出結果。

### 4. 生成對抗網路 (Generative Adversarial Network, GAN)

同時訓練生成器以及鑑別器，鑑別器結構一般為全卷積網路。鑑別器必須區分資料為生成器所產生或是來自於數據集，而生成器則必須生成令鑑別器無法區分的內容，藉此使生成器生成的內容或風格接近數據集。生成對抗網路可應用於許多方面，例如生成內容的 DCGAN [3]、實現風格轉換的 CycleGAN [4]、超解析度處理的 SRGAN [5] 等。若用於圖像端對端訓練，一般網路結構會包含編碼（下採樣）、風格轉換（多為殘差模塊）、解碼（上採樣）。

## 三、尋找解決方法

### (一) 建構整體實驗流程

基於目前所習得資訊及研究目標，必須依照數據類型，設計適合的神經網路。以下為本研究的實驗流程概覽：

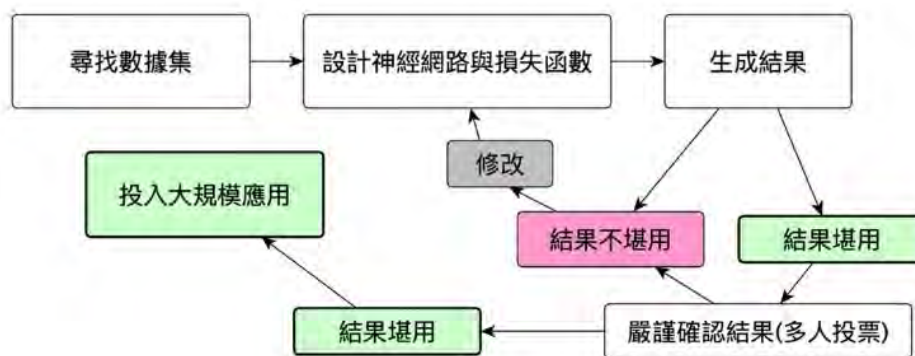


圖 1 實驗流程

## (二) 探討可行之訓練數據類型

訓練數據可簡單分為成對 (paired) 與非成對 (unpaired) 圖像，其中成對數據為拍攝同樣場景且經過對齊的數據集；非成對則為拍攝不同場景或未經過對齊。由於蒐集經過對齊的成對圖片較有難度，因此我們採用手機及單眼相機自行拍攝非成對的數據集。

## (三) 尋找合適之深度學習框架

以下為目前常見之深度學習框架：

### 1. TensorFlow

於 2015 年由 Google 發表，以 Python、C++ 與 CUDA 開發而成，編程方式極具特色，需於建立計算圖 (graph) 及存放資料用的佔位符 (placeholder) 後，再開啟運算需要用的對話 (session)。雖然對於剛接觸的新手可能較為複雜，但熟練後即可快速應用，而且自由度較高。使用 TensorFlow 提供 TensorBoard 工具，可即時觀看訓練的過程。

### 2. Keras

這是一種高級神經網路的應用程式介面 (application programming interface, API)，以 Python 開發而成，可以使用 TensorFlow、Theano 與 CNTK 作為後端運作的框架，讓使用者可以用簡單且少數的程式碼架設神經網路模型。

### 3. PyTorch

前身為 Torch，由於 Torch 是以 Lua 開發而成的，使用者較少。後來開發適用於 Python 的版本，並於 2016 年底發表，是目前使用者人數成長相當快速的深度學習框架。



## 四、準備及整理訓練數據

### (一) Unsplash 夜景照片數據集

Unsplash 夜景照片中，包含大量比例的星空照片。由於手機拍攝數據集中不包含星空照片，為了平衡資料比例，我們手動從 Unsplash 下載的照片不包含星空及散景。以下簡稱本數據集為 Unsplash 數據集。

### (二) 自行拍攝之手機夜景數據集

由於網路上難以獲得大量同一手機拍攝的夜間戶外照片，我們自行使用手機拍攝夜間照片供神經網路訓練。使用設備包含 HTC U11 以及 HTC U11+，拍攝模式為自動，並且啟用自動 HDR 模式。以下簡稱本數據集為手機數據集。

### (三) 自行拍攝之單眼夜景數據集

由於 Unsplash 數據集中的照片調色風格較為激進，而且有些本身也帶有噪點，可能對訓練的模型造成負面的影響，因此我們使用單眼相機進行長時間曝光（約 10~30 秒），並使用較小的光圈以獲得較深的景深。本數據共包含 160 張照片，照片內容包含河濱公園以及都市街景。以下簡稱本數據集為單眼數據集。

## 五、建構神經網路

由於近期的研究中，已可以有效地利用成對數據集訓練用以提升畫質的模型[1][6]，因此本專案專注於研究利用非成對數據訓練提升畫質的模型。

### (一) 結構概覽

我們首先利用 CycleGAN [4] 訓練由高品質照片轉換為低品質照片的模型，再利用此模型從高品質圖片生成低品質圖片，最後利用生成的低品質圖片與對應的高品質圖片（成對資料）訓練我們的“Fantastic Filter”模型。

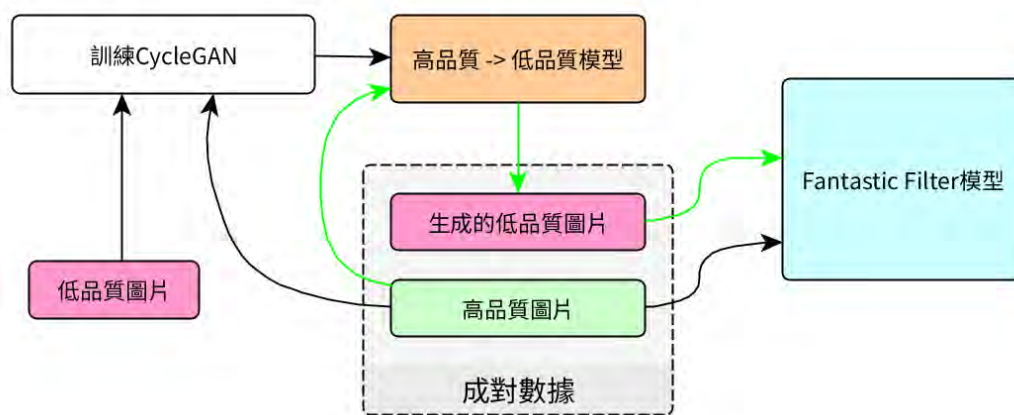


圖 2 訓練流程圖

### (二) 生成成對數據集

我們利用 CycleGAN 由高品質圖片生成對應的低品質圖片。CycleGAN 的結構如下圖， $x$  代表手機拍攝的圖片， $y$  代表單眼拍攝的圖片， $F$  代表由單眼圖片生成手機圖片的生成器， $G$  代表由手機圖片生成單眼圖片的生成器， $D$  代表鑑別器。此網路中一共使用 4 個優化器，分別對於：生成器  $G$ 、生成器  $F$ 、鑑別器  $D_x$ 、以及鑑別器  $D_y$ 。其中  $D_x$  應盡可能區別  $F$  所產生的圖片及  $x$  的圖片， $D_y$  應盡可能區分  $G$  所生成的圖片及  $y$  的圖片。而  $F$  及  $G$  則應盡力使鑑別器無法區分產生的照片以及真實照片。

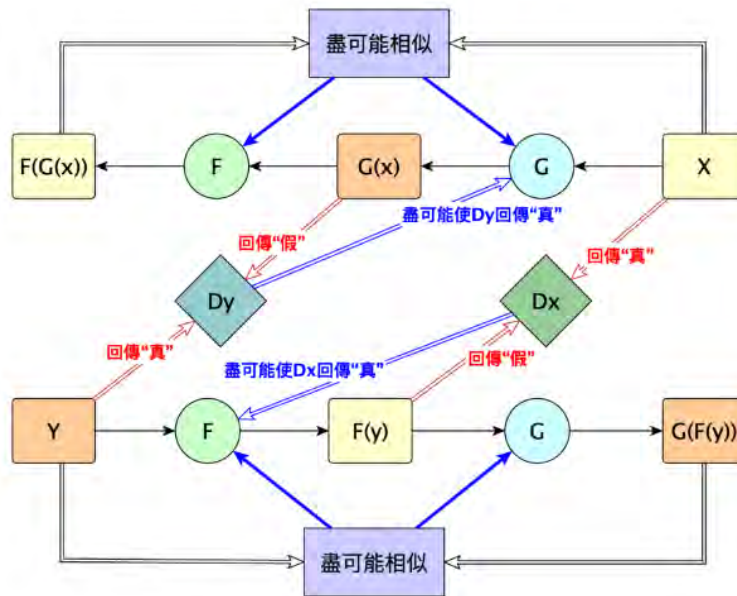


圖 3 CycleGAN 結構示意圖

相較於 CycleGAN 原文，我們加上了**特徵圖的層級上的循環一致損失**（將來源圖片與重構圖片輸入預訓練的 VGG-19 模型，取出特徵圖），如此可以獲得較為平滑的訓練過程。另外我們也使用**最臨近點插值**代替卷積轉置，以避免產生明顯的假影。我們分別使用一次以及兩次降低畫質的步驟，即  $F(y)$  以及  $F(F(y))$ ，探討對於最終訓練結果的影響，詳細的對比請見[討論](#)。下表為原圖以及使用 CycleGAN 生成對應的低品質圖片：

表 2 經由 CycleGAN 降低畫質的圖片

原圖	經 CycleGAN 降低畫質的圖片

### (三) 使用成對數據集訓練生成器

我們利用由 CycleGAN 產生的成對數據集訓練生成器。生成器的結構如下：

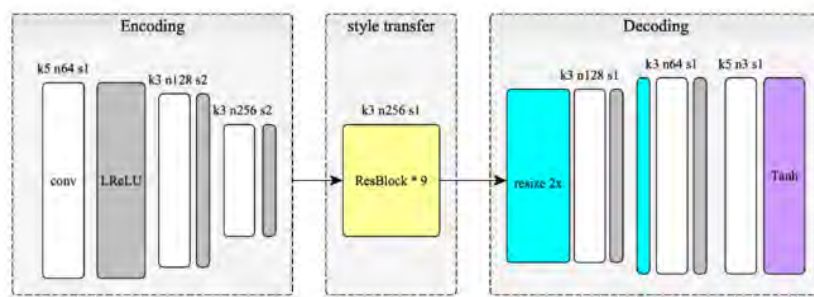


圖 4 生成器結構

其中 ResBlock 代表殘差模塊，結構如下圖：

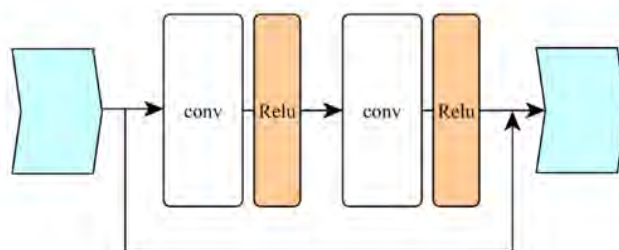


圖 5 殘差模塊

另外我們加入了鑑別器，鑑別器的結構如下圖：

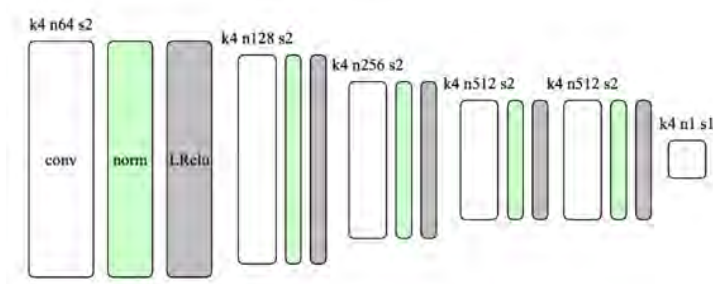


圖 6 鑑別器結構

#### (四) 損失函數

##### 1. 顏色損失

我們使用 L1 損失來計算像素層級上顏色的差。

$$L_{color} = \|y - f(x)\|_1$$

##### 2. 內容損失

我們將圖像輸入預訓練的 VGG-19 [7] 模型，並計算生成圖片與目標圖片的特徵圖損失。經過多次實驗，我們採用“relu\_5\_4”層上的特徵圖。

$$L_{content} = \|VGG(y) - VGG(f(x))\|_1$$

##### 3. 對抗損失

根據 pix2pix [8] 文中所述，單純地使用 L1 損失會造成生成結果模糊，因此我們加入了對抗損失。損失函數如下：

$$L_{adv} = \|D(f(x)) - 1\|_2^2$$

##### 4. 總損失

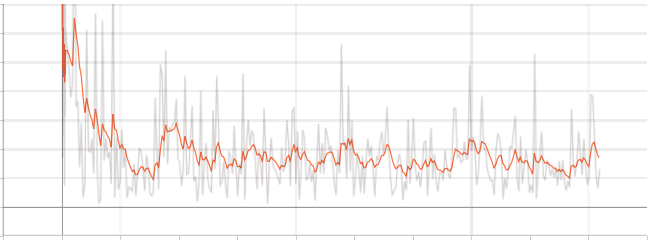
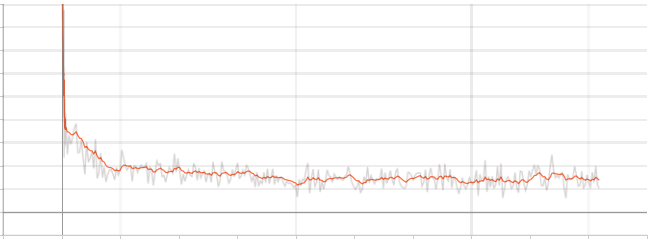
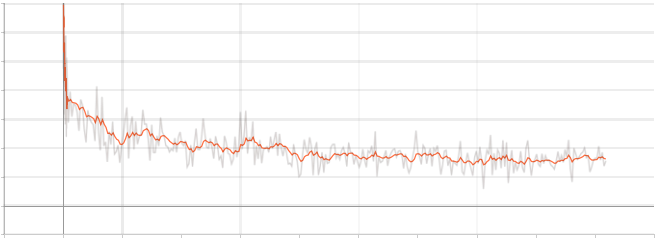
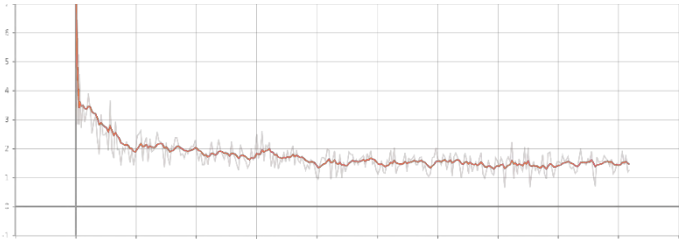
計算總損失時，我們簡單地將顏色損失、內容損失、以及對抗損失乘以比例並且相加。關於比例對於最終結果的影響請見[討論](#)。

$$L_{sum} = \alpha L_{color} + \beta L_{content} + \gamma L_{adv}$$

## 六、訓練神經網路

我們嘗試過許多超參數的組合，最後使用 RMSProp 優化器，初始學習率為  $2 \times 10^{-4}$ ，並於全局步數 5000 至 25000 時，學習率由  $2 \times 10^{-4}$  線性下降至  $1 \times 10^{-5}$ ，批量大小為 8，並且透過滑動平均獲得較為穩定的訓練過程。我們嘗試過不同損失函數比例的組合，均有類似的收斂過程。訓練時圖片均裁切為  $256 \times 256$  的 patch。

表 3 損失趨勢圖

損失名稱	損失—全局步數圖。縱軸：損失；橫軸：全局步數。
對抗損失	
顏色損失	
內容損失	
總損失	

## 七、建構圖形化使用者介面 (Graphical User Interface, GUI)

由於一般使用者可能不會編寫 Python 語言，我們使用 Python 內建的函數庫 Tkinter 建構使用者介面，讓一般使用者可以方便地使用本專案的模型提升照片畫質。下圖為我們建構的圖形化使用者介面：

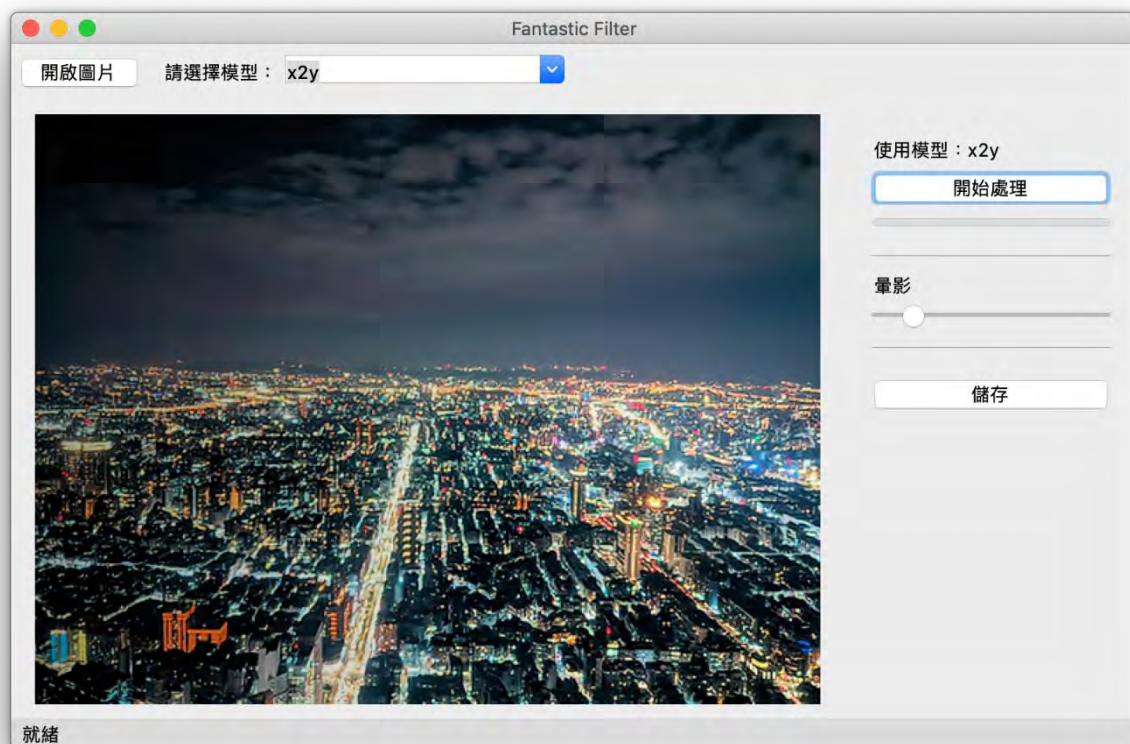


圖 7 圖形化使用者介面工具

## 八、製作手機 APP

為了方便用戶即時地增強所拍攝的圖像，我們嘗試將模型遷移至 iOS 以及 Android 平台，在 Android 平台我們使用 TensorFlow-Android，於 iOS 使用 CoreML 框架。經過實測，在 HTC U11/U11+、Samsung Note 8、小米 8 lite/SE 均可在 10 秒內處理 1280x720 大小的圖片。

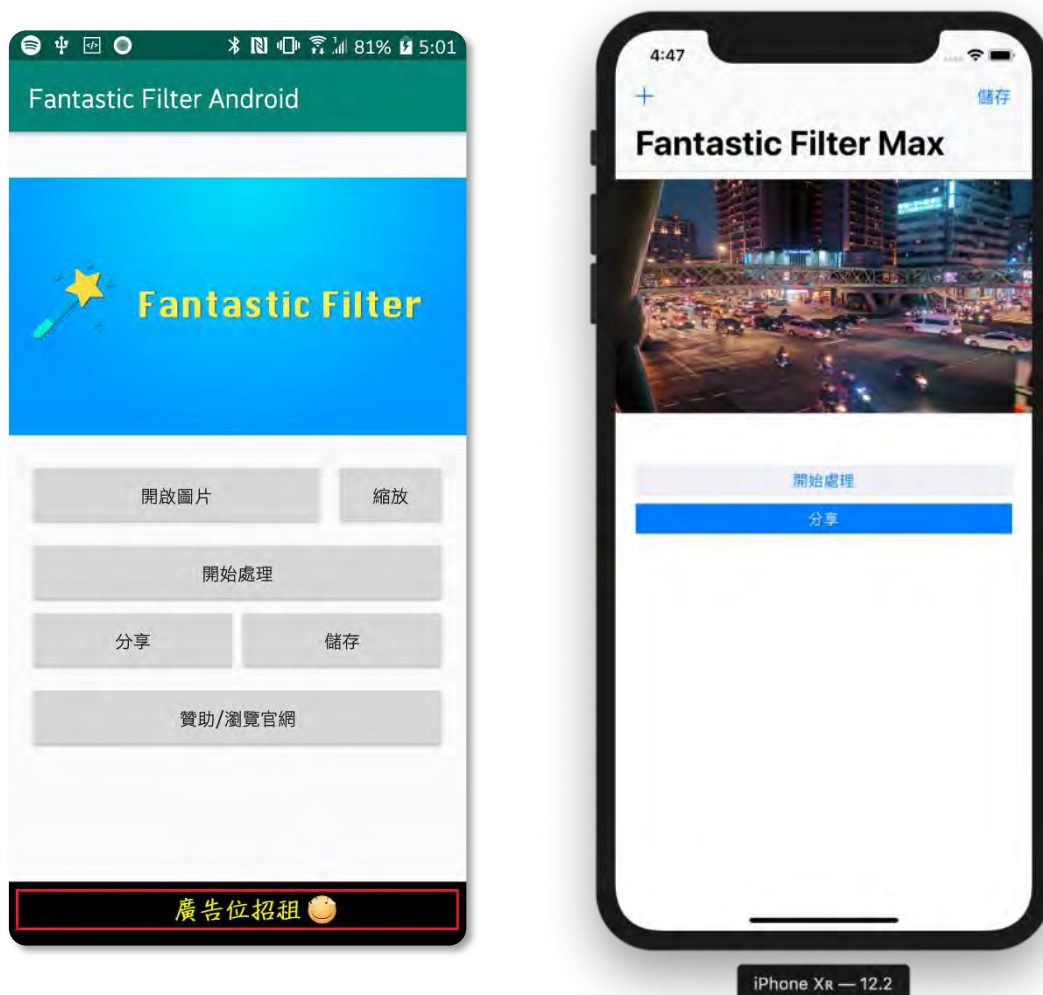


圖 8 “Fantastic Filter” 在 HTC U11+中執行，以及在 Xcode 虛擬環境中執行的截圖



## 伍、研究結果

### 一、生成照片與真實高品質照片對比

我們發現透過我們的模型可以有效消除低品質夜景手機中的噪點，並且提升部分細節的清晰度。下圖為處理前、後，以及單眼相機所拍攝相同場景的對比圖：

表 4 處理前後對比圖（一）



表 5 處理前後對比圖（二）

處理前



處理後



高品質



可以發現經過處理的圖片，噪點被有效地消除。由於圖片難以完全對齊，無法直接計算生成圖片與目標之間的差異，如結構相似性 (structural similarity index, SSIM index)。

## 二、我們的方法“Fantastic Filter”與 CycleGAN 結果對比

下表為使用“Fantastic Filter”處理圖片與單純使用 CycleGAN 所產生的圖片對比：

表 6 “Fantastic Filter” 生成結果與 CycleGAN 對比







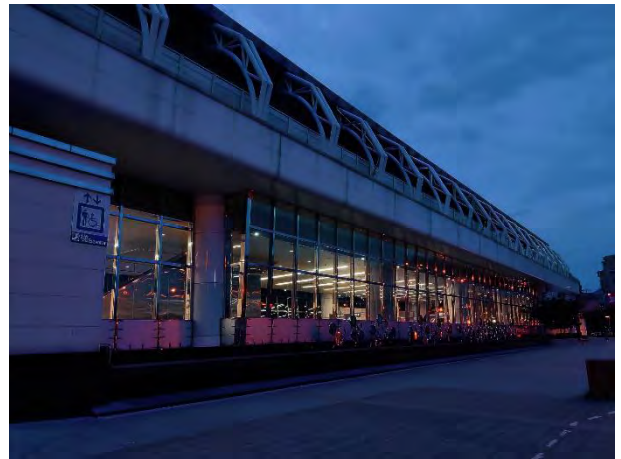
CycleGAN	Fantastic Filter (我們的)
	
	
	

表 7 “Fantastic Filter” 生成結果與 CycleGAN 對比（續）

CycleGAN

Fantastic Filter（我們的）

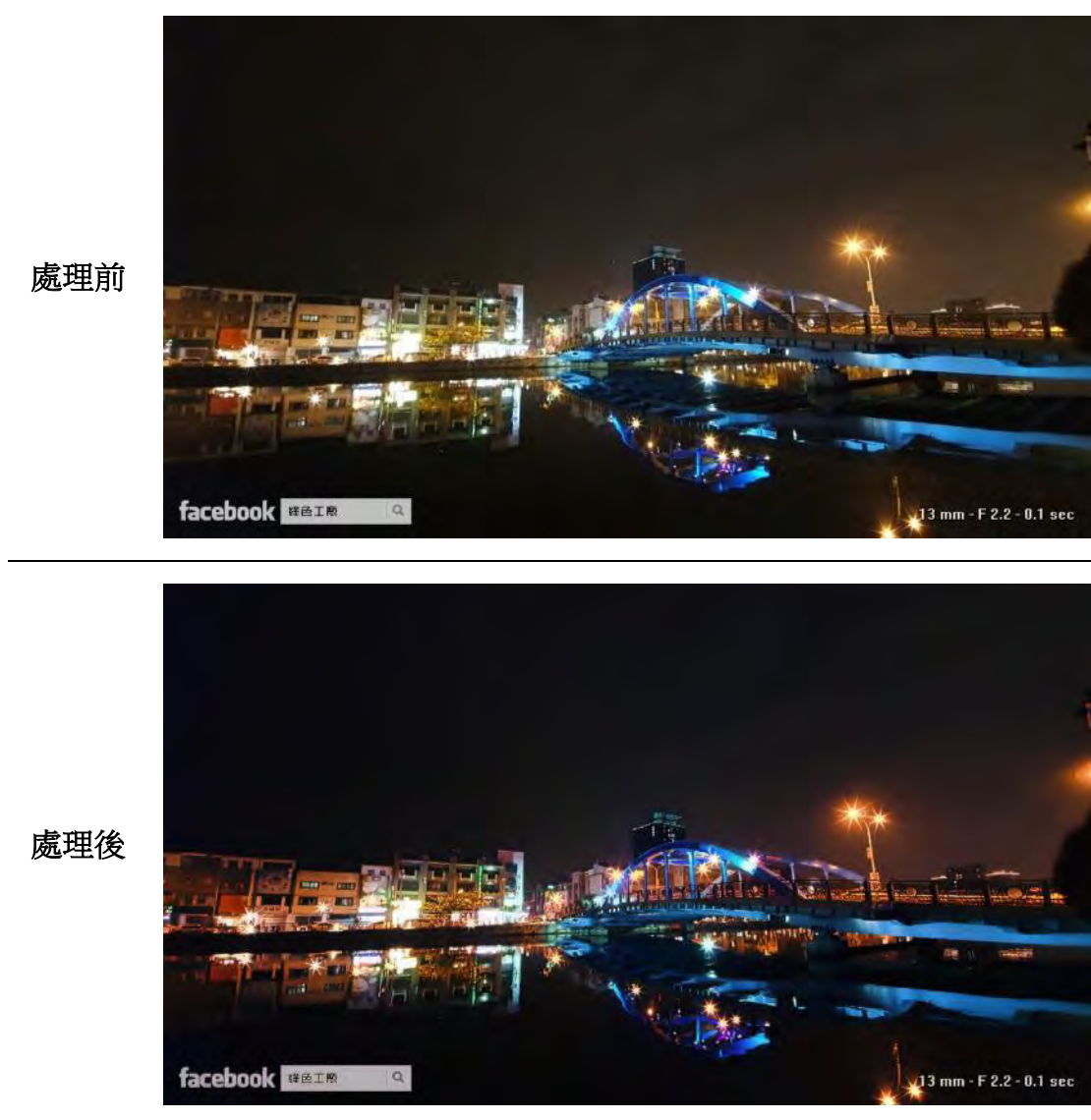


由上表可發現，相較於 CycleGAN，由本專案產生的圖片邊緣較為自然，細節較為豐富，且較可以有效的處理噪點。

### 三、模型泛用性

由於我們所蒐集的高品質圖片多為河濱公園的街景，不包含散景、星空等場景，且均由相同或類似的手機拍攝。下方列出模型對於不同手機型號場景處理的能力。

表 8 模型泛用性測試結果（手機型號）



上圖原圖取自 <https://ez3c.tw/7189>，使用 Samsung S10+ 拍攝。原圖經過網站壓縮，並且使用最鄰近點縮放至 1760 x 990 後輸入模型增強。可以發現即使輸入不同感光元件的手機所拍攝的照片，模型依然可以有效地處理照片。

表 9 模型泛用性測試結果（星空）

處理前



處理後



上圖取自 <https://m.eprice.com.tw/mobile/talk/102/5122715/1/rv/huawei-p20-pro-review/>，使用 Huawei P20 Pro 拍攝。參數為 F/1.8、30 s、ISO 500。可以看到在原圖畫面中充滿噪點，在經過模型處理後，噪點可以有效地被消除，且仍然保留星星等細節。

表 10 模型泛用性測試結果（場景）

處理前



處理後



在訓練資料中不包含的夕陽場景中，模型依然可以有相當好的表現，有效地去除原圖中的噪點。受限於版面空間，更多處理前後對比的樣張請見[附錄](#)。

## 陸、討論

### 一、討論

#### (一) CycleGAN 訓練過程

在 CycleGAN 循環中，由低品質生成高品質圖片的生成器所產生的圖片，畫面損失的細節較多，且畫質較為粗糙；但由高品質轉換成低品質的生成器所產生的圖片，則可以有效的生成類似於手機在夜拍中產生的噪點，因此我們嘗試利用 CycleGAN 由高品質圖片產生低品質圖片，再利用生成的低品質圖片作為輸入、對應的原始高品質圖片作為目標圖片的方式訓練另一個模型，藉此達到較好的效果。或許在其它圖像翻譯任務中也可以利用類似的方式達到較好的生成品質。

#### (二) 數據集對於生成結果

在先前的實驗中，我們分別嘗試使用 Unsplash 包含與不包含星空的照片，以及利用單眼相機所拍攝的照片訓練 CycleGAN 模型。藉由觀察訓練過程，我們發現利用帶有星空的 Unsplash 與手機數據集訓練模型時，生成器常於暗處生成類似星星的紋理；在利用 Unsplash 不包含星星的圖像訓練模型時，由於 Unsplash 照片已經過影像處理程序，雖然在訓練時生成的 patch 接近目標圖片，但在實際測試時發現生成的圖片較為模糊。因此，我們最終使用由自行拍攝的高品質照片訓練的模型。

#### (三) 損失函數比例對於結果的影響





我們分別嘗試調整損失函數中各項損失函數的比例，在我們的實驗中，無論加上對抗損失以及特徵圖層級的 L1 損失與否，實驗結果均相同，僅使用像素層級的 L1 損失（顏色損失）也不會導致生成圖片模糊，與 pix2pix [8] 所述有所出入。我們實驗所使用的損失比例組合請見附錄。



#### (四) 影響生成結果的要素

我們分別利用處理一次及兩次的低品質數據訓練模型，以及分別利用 U-Net 與 ResNet 作為模型結構，詳細生成的結果如下，表中 F 為降低畫質的生成器：

表 11 不同的網路結構以及不同的數據集對於結果的影響

	ResNet	U-Net
$F(y) - y$		
$F(F(y)) - y$		

可以發現模型結構對於生成結果的影響較小，而訓練資料則為影響生成結果的主要原因。在經過兩次降低畫質的圖片訓練的模型所處理的結果顏色偏離原圖較多，但是降噪的效果較好。

## 二、未來展望

### (一) 設法取得更多訓練資料，提升成像品質。

受限於時間，本專案用以訓練型的數據集，僅包含 160 張數位單眼所拍攝的高品質照片以及 700 張手機所拍攝的照片。雖然模型已經足以應付多數場景，但仍然會有少數場景出現過度飽和、亮度過低的情況，我們認為這是由於訓練資料太少所導致。若未來可以取得更多訓練資料，應可有效提升模型品質。

### (二) 蒐集多種風格高品質照片，優化特定場景。

在自行利用數位單眼相機蒐集高品質數據集時，受限於硬體性能，我們必須使用極低的 ISO 值 ( $\leq 200$ ) 搭配較長的快門時間 (20 s ~ 30 s) 以避免產生噪點，故較難拍攝人像，且透過觀察，人像照片的風格與風景略有不同，若未來可取得較好的設備，在學習人像拍攝技巧後，或許可以另外蒐集帶人像的照片數據集，訓練適合處理夜間人像的模型。

### (三) 將卷積過程嵌入較通用的 GPU 框架，方便一般使用者快速使用。

目前如欲使用我們所製作的 GUI 工具透過 GPU 快速地處理圖片，須於電腦安裝 NVIDIA CUDA Toolkit，對於一般使用者而言安裝過程非常麻煩；如果沒有安裝 CUDA 則必須等待 CPU 緩慢地運算。因此我們希望可以將程式嵌入較為通用的函數庫，例如 OpenGL，藉此獲得較佳的使用體驗。

## 柒、結論

一、在我們的任務中，pix2pix 生成品質較 CycleGAN 高。

在我們的實驗中，先利用 CycleGAN 生成成對數據，再利用此數據訓練 pix2pix 模型的方式，生成的圖片品質比 CycleGAN 所生成的品質高。這可能是因為 CycleGAN 中需同時處理較多生成器，無法專注於目標任務，而且受限於某些損失函數，如顏色正規化、循環一致損失，生成器可能無法完全的轉換風格，並由於生成結果僅受鑑別器判斷，畫面中時常帶有雜訊；但在 pix2pix 中由於任務單一，且有真實的目標圖片作為參照，生成的結果較無雜訊、假影等造成畫質下降的要素，因此生成的結果較好。

二、透過 GUI 工具，使用者可以方便地使用本專案的模型提升照片畫質。

經由我們釋出 GUI 工具，使用者僅需開啟欲處理的圖片，並且點擊開始處理，即可自動地移除圖片中的噪點，並且有效地提升畫質、增強影像。

## 捌、參考資料及其他

- [1] Ignatov, Andrey, et al. "DSLR-quality photos on mobile devices with deep convolutional networks." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.
- [2] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
- [3] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." *arXiv preprint arXiv:1511.06434* (2015).
- [4] Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *Proceedings of the IEEE international conference on computer vision*. 2017.
- [5] Ledig, Christian, et al. "Photo-realistic single image super-resolution using a generative adversarial network." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [6] Chen, Chen, et al. "Learning to see in the dark." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [7] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- [8] Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

## 玖、附錄

### 一、處理前後樣張對比

以下圖片均由手機拍攝，並經由本程式處理而成。其中前四組樣張參與 CycleGAN 訓練但不參與後續成對訓練，其餘則完全未參與訓練。在螢幕上放大可以更好地觀察細節。

表 12 處理前後樣張對比


處理前	處理後
	
	
	

表 13 處理前後樣張對比 (續)

處理前	處理後
	
	
	
	

## 二、損失函數比例實驗細節

在我們的實驗中，下列超參數組合皆可有相同的生成結果。

表 14 損失函數比例實驗結果

編號	$\alpha$ (顏色損失)	$\beta$ (內容損失)	$\gamma$ (對抗損失)	附註
[0]	20	5	1	對照組
[1]	20	0	1	不使用內容損失
[2]	20	10	1	加大內容損失比例
[3]	20	10	0	不使用對抗損失
[4]	20	10	10	加大對抗損失比例
[5]	20	10	30	加大對抗損失比例
[6]	20	0	0	僅使用像素層級 L1 loss

## 【評語】 052510

此作品提出一個提升手機拍攝照片品質的方法，所採用方法是建立 GAN 模型來消除手機於夜間拍照時照片中的噪點（雜訊）。此作品採用 GAN 的作法由高品質圖片產生對應的低品質圖片再使用此數據來訓練生成器，藉此達到降噪以及提升畫質的效果。

此作品具有一些技術深度，但降噪的成果應有更科學（統計）的方式進行比較，而非僅以圖例讓讀者用肉眼來進行影像品質的比較。此外，研究成果也應與相關研究成效進行比較。



# 摘要

在這份研究中，我們建立 GAN 模型，嘗試消除手機夜拍中的噪點以及提升畫質。我們先利用 CycleGAN 產生成對數據，在使用此數據訓練生成器，藉此達到降噪及提升畫質的效果。藉由我們的模型，用戶可以透過簡易的方式，有效地去除手機拍攝夜景圖像中惱人且難以消除的噪點。據我們所知，我們的研究是目前最方便且有效的方式，為自動影像處理開創嶄新的道路。

## 研究動機

現代人常常使用手機拍照，記錄生活中的點滴。但受限於手機感光元件大小，時常無法獲得較好的畫質，特別是在夜間或低亮度的環境拍攝。因此我們想建構可以用於自動提升畫質、修正色調的軟體，讓使用者可以方便地使用。

## 環境 & 工具

- Arch Linux
- GeForce® RTX 2080 Ti
- Python 3.6
- TensorFlow 1.13.1
- 低品質照片數據集 (由 HTC U11/U11+ 自行拍攝)
- 高品質照片數據集 (由數位單眼相機自行拍攝)



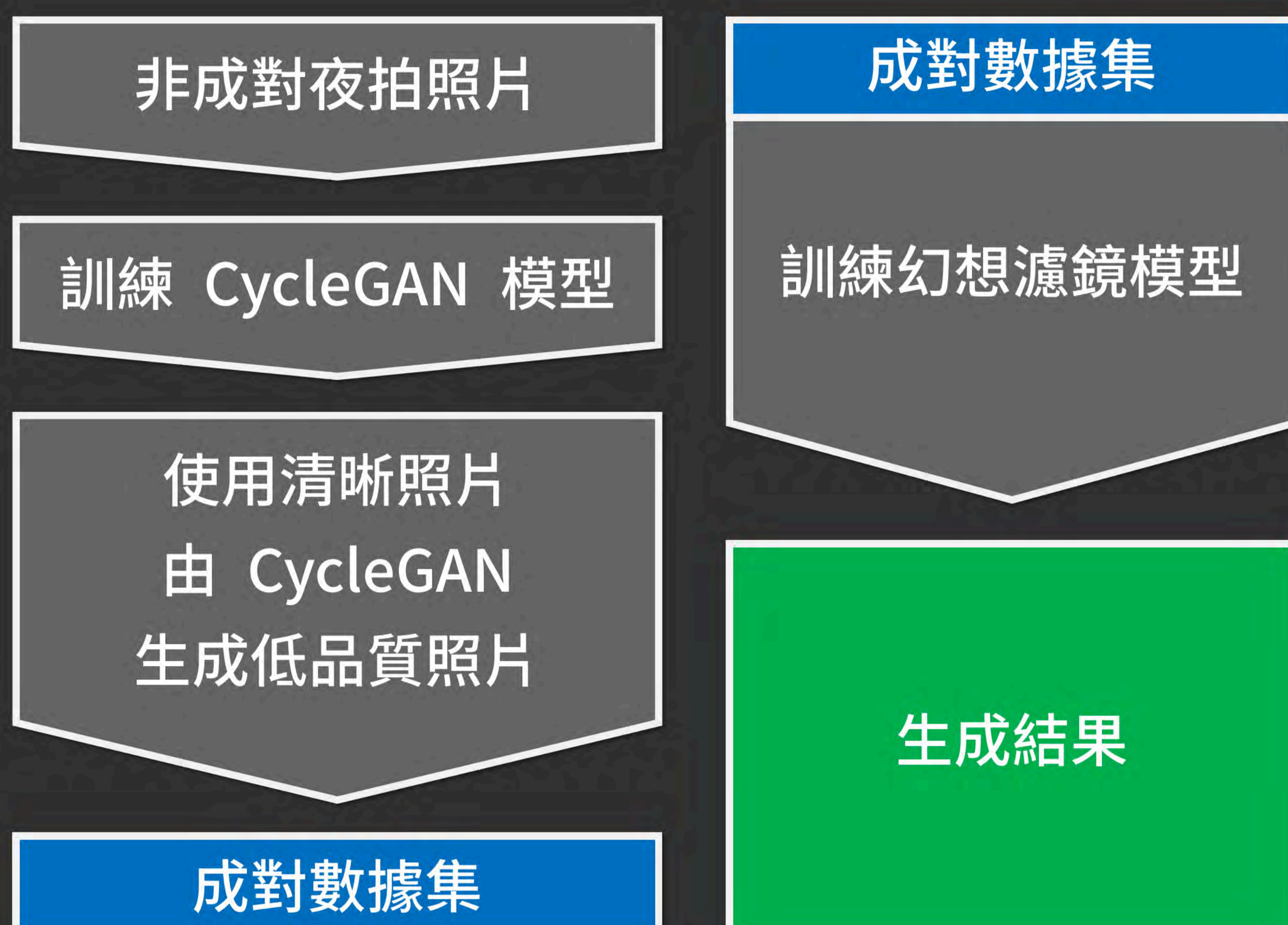
## 研究目的

- 探討使用神經網路處理圖片之可行性
- 探討使用非成對手機拍攝圖像及高品質照片訓練神經網路之可行性
- 訓練增強畫質的模型
- 製作圖形化使用者介面工具

# 研究方法 & 過程

## 訓練流程

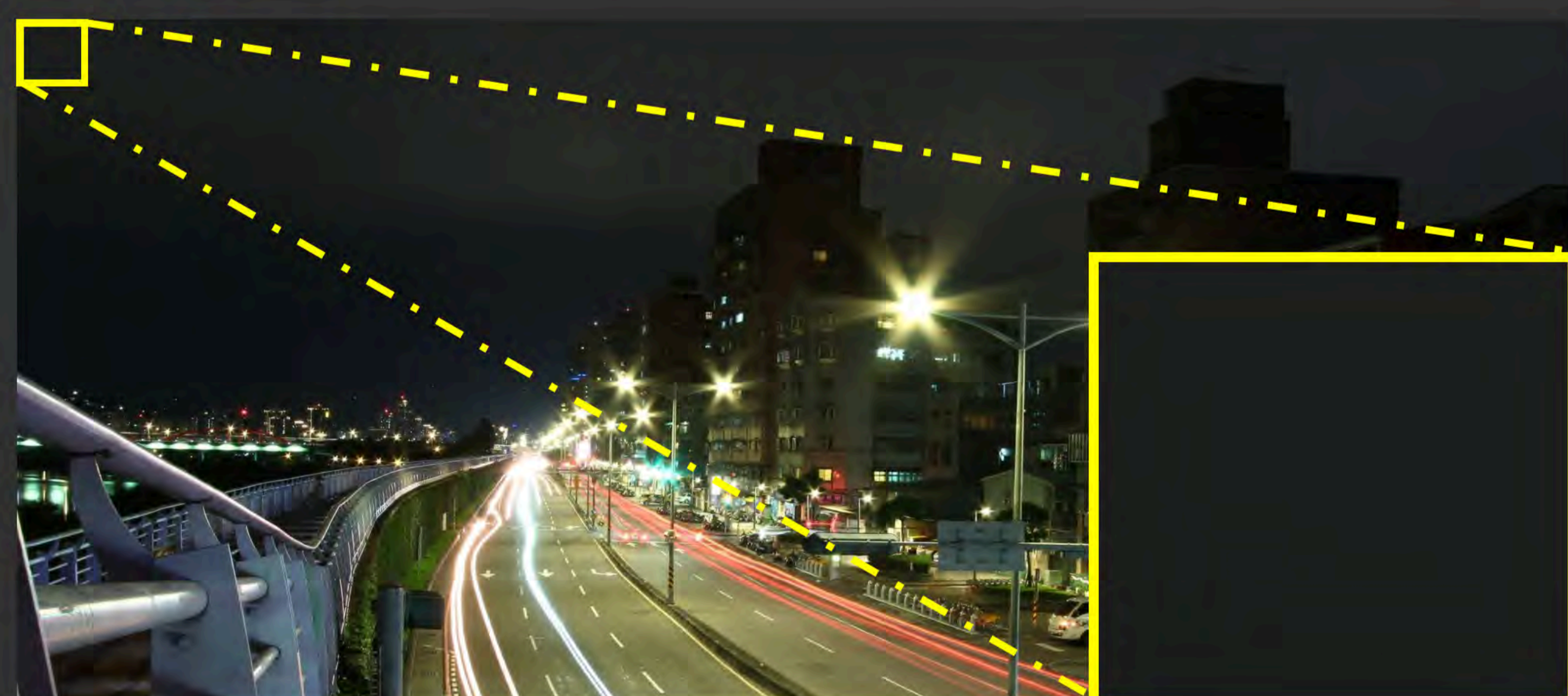
我們將高品質圖片及低品質圖片輸入 CycleGAN 模型，嘗試利用 CycleGAN 模型產生成對的高、低品質數據集，並利用此成對數據集訓練另一個生成器，藉此達到提升畫質、消除噪點的效果。



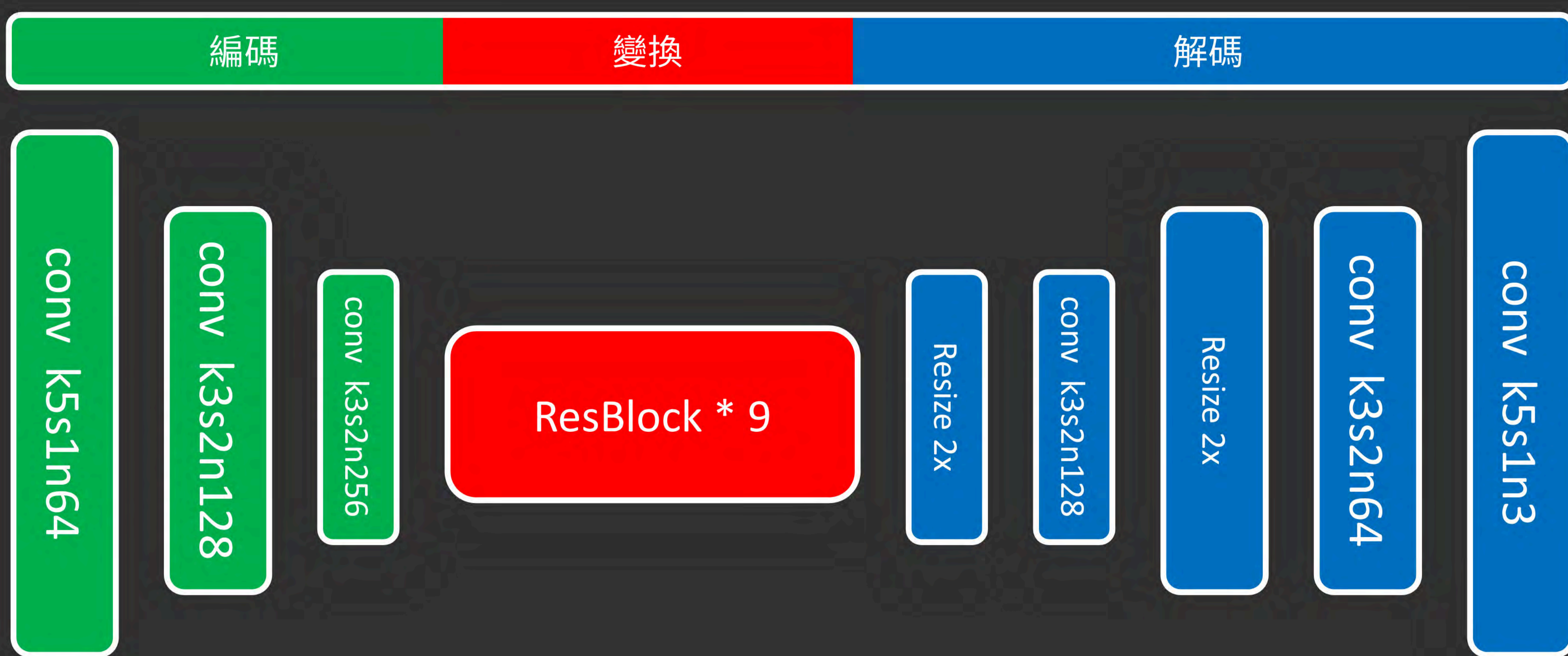
## 由非成對數據產生成對數據集



我們利用非成對高、低品質數據訓練 CycleGAN 模型，並將高品質圖片輸入此模型，以獲得成對的高 (上圖中的  $y$ )、低品質數據集 (上圖中的  $F(y)$ )。



# 由成對數據訓練模型

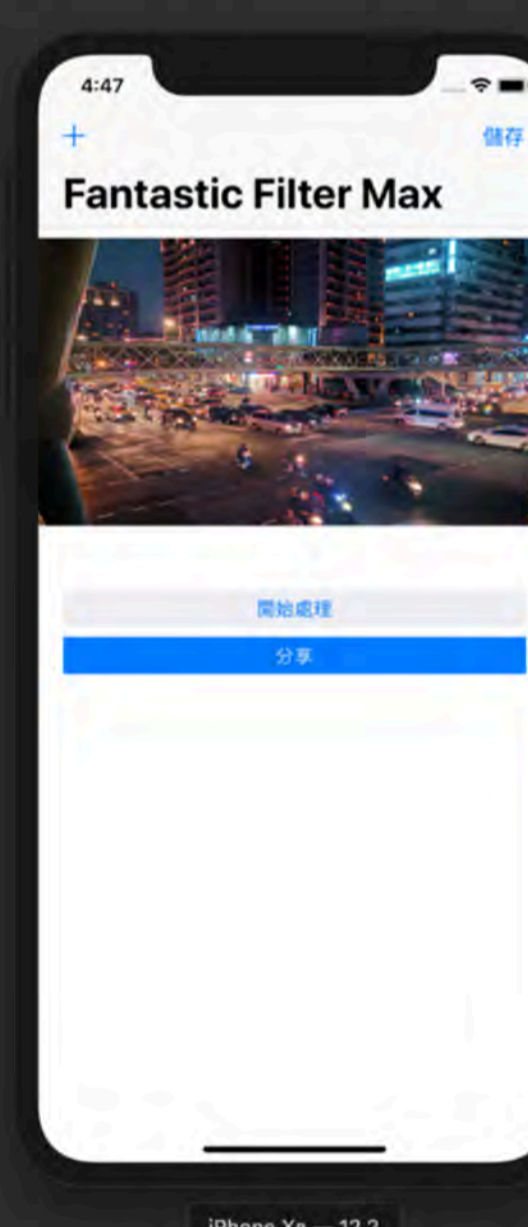
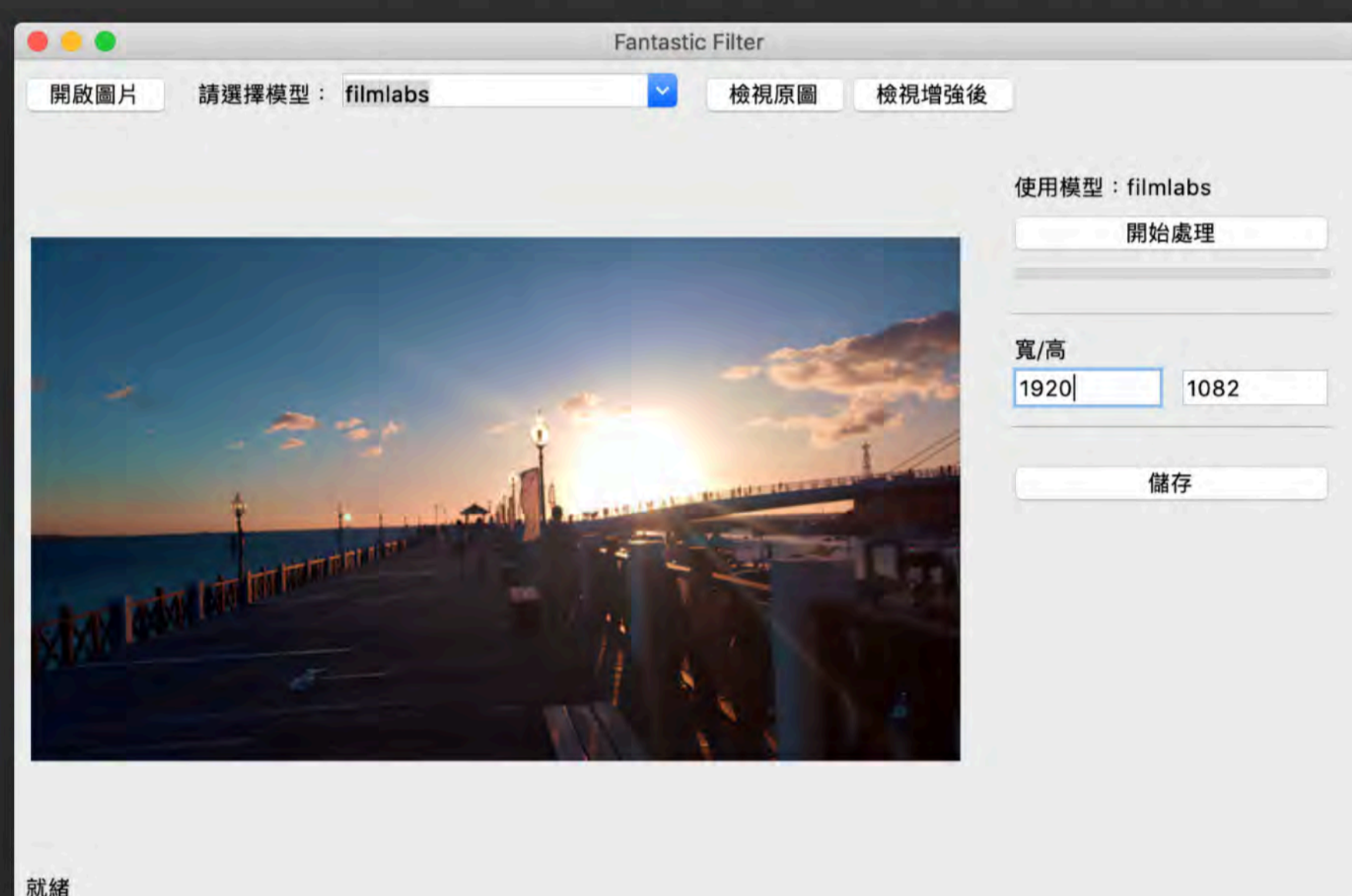


## 研究結果



兒童新樂園。處理前（左）處理後（右）

## 圖形化使用者介面工具



## 結論 & 展望

- 使用 CycleGAN 生成低品質數據的表現優異。
- 透過本專案所訓練的模型，可以有效提升影像畫質及照片純淨度。
- 目前所使用的模型對於室內等特定場景處理能力較差，未來可望訓練多個模型，適應不同場景，以達到更好的處理結果。

## 參考資料

- Isola, Phillip, et al. "Image-to-image translation with conditional adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- Ignatov, Andrey, et al. "DSLR-quality photos on mobile devices with deep convolutional networks." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.
- Zhu, Jun-Yan, et al. "Unpaired image-to-image translation using cycle-consistent adversarial networks." *Proceedings of the IEEE international conference on computer vision*. 2017.
- Chen, Chen, et al. "Learning to see in the dark." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.