

題目：

GAN 模擬生成指定人、物動作圖像

組員：

黃維熙 40547005S

古佳僊 40547010S

余承寰 40547034S

指導教授：

黃文吉教授

## (一)、 摘要:

本研究以一個人、物之多種動作姿態為模型，模擬生成不同的人、物卻能做出多種相同動作姿態為目標，了解並使用 CNN 建模，訓練 model 準確率，讓電腦能夠盡量準確辨認圖片影像，再利用 CNN 與反 CNN 結合實作出 GAN，來合成產生新的人、物圖形。

## (二)、 研究動機:

圖像視覺對人類來說，是不可忽略的，從眼睛所看見的東西都是圖像的一種，若能利用 GAN 生成指定人、物做指定動作，就可以使我們不需再花大量人力標記資料，利用 GAN 模型生成即可，『監督式學習』(Supervised)到『非監督式學習』，未來在許多影片或圖片的拍攝，將不需要指定的演員或模特兒，影像處理的問題也可以利用 GAN 進行合理且高辨識度的推測。

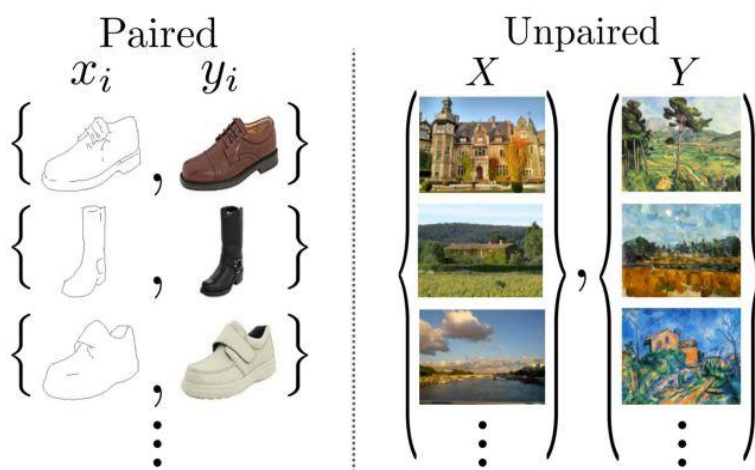


圖 2-1

圖 2-1 說明在使用 Cycle-GAN 時，不需要指定特定圖片作為輸入，即使兩張圖片差異極大，也可以直接正常的運作並利用 discriminator + generator 將其他圖片的特徵與原圖進一步結合，製作出仿真度極高的合成圖片。



圖 2-2

圖 2-2 是利用 GAN 對圖片進行影像處理。

### (三)、 研究方法及步驟:

GAN 基本架構圖(圖 3-1):

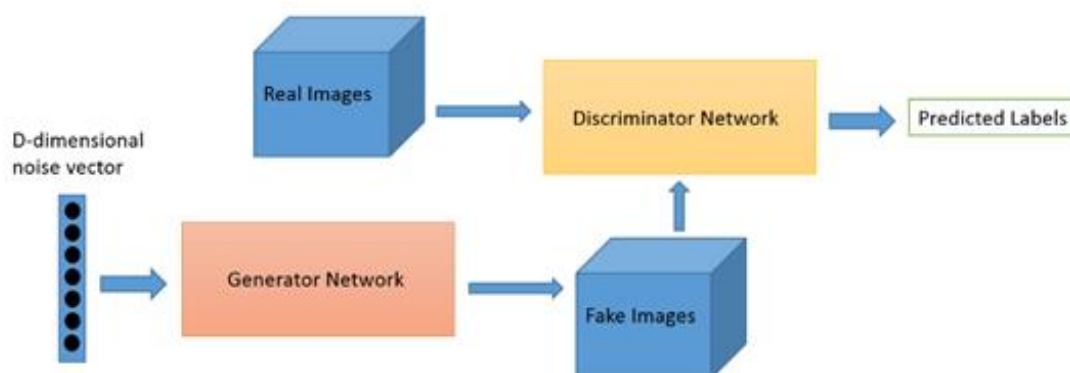


圖 3-1

GAN 最主要的兩個關鍵就是 Discriminator 與 Generator 他們都是由神經網路所組成，這兩個神經網路擁有不同目標，並相互對抗，最終透過 Discriminator 來達成目標。

Generator 的工作是負責產生擬真的圖片，而 Discriminator 的工作則是分辨程式給他的圖片中，哪些是由 Generator 產生的假影像，哪些是真正的影像。透過訓練來讓這兩個類神經網路在競爭的環境下不斷地改良，並使兩種類神經網路介於平衡之間，並能夠透過 GAN 來生成實際的圖片。

#### 1. 資料收集與整合:

我們使用爬蟲收集資料，收集完後將不符合預期的圖片移除後，將所有圖片 **resize** 成相同尺寸，並將資料分為 **test data** 跟 **train data** 兩個部分，使用資料夾分別儲存各種不同的資料圖片，並利用數字進行編號，再將圖片依照程式所需去做切割。

#### 2. GAN 模型探討與設計:

我們選擇 **Cycle-GAN** 來進行實作，一般的 GAN 在輸出時，只需要輸出一個 **true** 或 **false**，代表對整張圖的評價，但 **Cycle-GAN** 是通過疊加的卷積層輸出的是一個 **N\*N** 的矩陣，矩陣中的每個元素只有 **true** 或 **false** 兩個選擇，每個元素對應到原圖的一個 **patch**，是 **Patch GAN** 的一種，藉由此 **bool** 矩陣來判斷圖片的哪些輸入像數較為敏感，並針對那些部分加以修改或利用。

##### (i) generator:

我們使用的 **generator** 是 **U-Net CNN** 的應用(如圖 3-2)，主要有收縮與擴張兩個網路結構，收縮主要用於捕捉特徵，擴張用於將特徵定位並還原出新圖片，這種 **Generator** 較容易收斂，因此所需的訓練集較少。

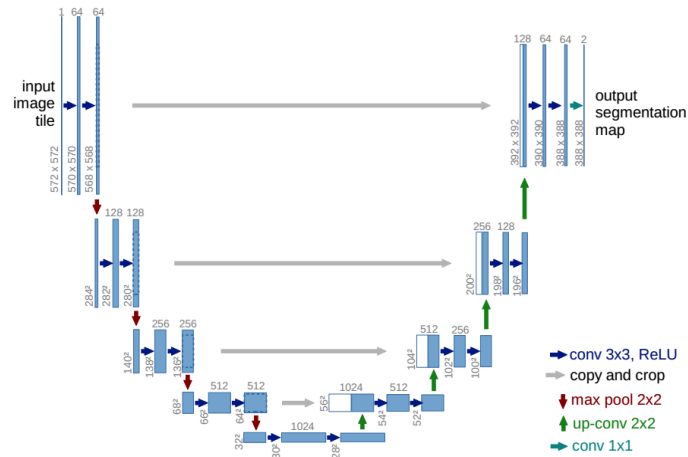


圖 3-2

(ii)discriminator:

discriminator 的部分將圖像作為輸入並辨認該圖片的真假(如圖 3-3)

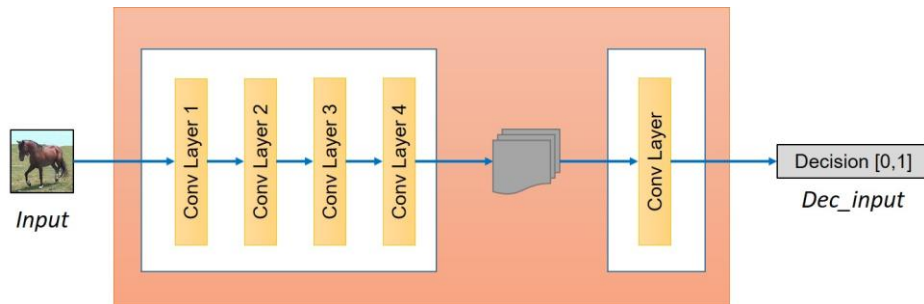


圖 3-3

(iii)train:

一次會輸入兩張圖片(A、B)，使用 generator(A)透過 keras 內建的 predict 函式產生擁有 A 特徵的 B 的新圖檔，再使用 generator(B)透過 predict 產生擁有 B 特徵的 A 的新圖檔，之後使用 discriminator 來辨識產生的圖片是否為真，並記錄該次的仿真度與 model(如圖 3-4)

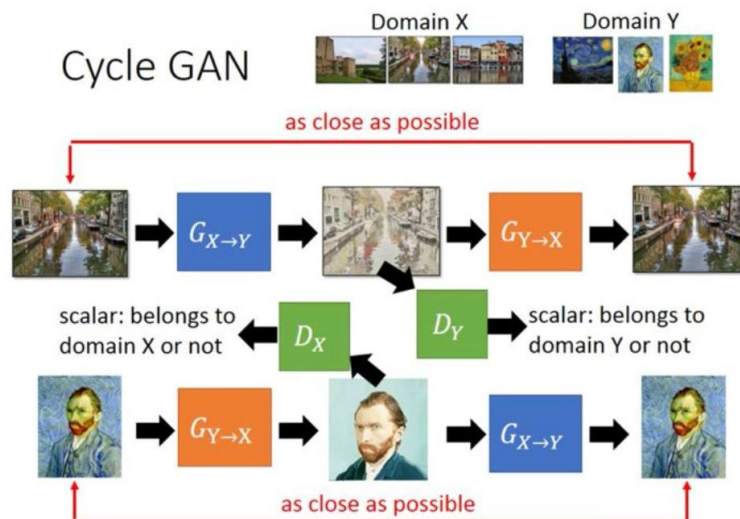


圖 3-4

## (四)、 遇到的困難:

### 1.問題:

我們一開始的目標是想要把兩張圖片中的人、物特徵都交換，進而達到讓 A 圖片的人、物出現在 B 圖片中，但當我們實作起來時，卻一直都只能將一種特徵交換，像是衣服、頭髮、膚色等等，甚至有時候根本沒交換，完全無法控制並確定 Cycle-GAN 要交換的特徵。

### 2.可能原因:

#### (i) 圖片解析度不夠:

在進行特徵交換的過程中，需要提取圖片的特徵，可能因為圖片太小，或解析度不夠，導致特徵不明顯。

#### (ii)資料集過於雜亂:

由於我們的資料集是利用爬蟲的方式，在 google 上面抓取關鍵字的圖片，再用人為的方式手動篩選過濾圖片，所以可能有些資料集的特徵會比較不明顯，導致程式在 train 的過程中，可能會有些微誤差。

#### (ii)generator 與 discriminator 強度不夠:

這個程式的做法主要是在 generator 產生圖片，然後在 discriminator 辨識仿真度並輸出，因此如果 generator 內部的 layer 不夠或是特徵抓取不夠精確，產出的圖片可能很糟。

#### (iii)使用的 GAN 可能無法達到這樣的效果:

在網路上有查到使用兩個 Condition-GAN 也有改變人的臉部特徵，但也是部分特徵改變，若想改變整個人、物，只有一種的 GAN 可能不夠，導致無法達成預期的成果。

### 3.推測可行解決辦法:

#### (i)資料與前置處理動作:

在進行轉換前，可先做一些前處理，去除圖片的 noise，並在收集 train data 時，挑選解析度較佳的圖片，且為了預防擷取的特徵太小導致不明顯的情況，盡量不要使用太小的圖片當作 train data。在挑選資料集的過程中，須更明確劃分出資料類別，減少 GAN 判斷特徵的問題(在爬蟲程式上增加一些條件，使之更明確的抓取資料)

#### (ii)GAN 修改:

首先嘗試調整 generator 的層數，並將 discriminator 認定仿真度過低的影像重新丟回 generator 重新訓練，如果無法逼近預期結果，將嘗試結合多種不同的 GAN，將輸入圖片根據不同的部位與情況使用不同種類的 GAN 做調整。

利用多種 GAN 產生多種目標輸出，最後把每個 GAN 的輸出情況加以拼湊出想要的結果，像是使用兩個以上的 Condition-GAN 進行臉部特徵的交換處理，在衣服與紋路的部分使用 Cycle-GAN 將其特徵交換，最後將兩者的輸出圖片擷取拼湊達成理想目標，但在經過這麼多處理之後可能導致最終的輸出結果一下就被 discriminator 識破，因此要讓每一次輸入的圖片重複進行多次處理直到 discriminator 認為精確度夠高，才可繼續往下進行。

## (五)、 研究結果:

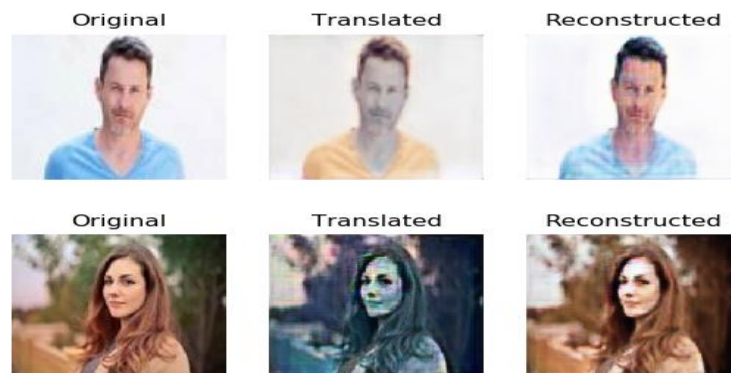
各資料集實作探討:

蚯蚓與蛇



原本設想蚯蚓和蛇的特徵類似，特徵(身上的花紋、顏色)交換的成功率會較高，但實際操作後發現，特徵交換會被背景的特徵稀釋掉，例如:蚯蚓的圖片背景土的部分比例多時，GAN 會擷取土的顏色當作特徵，套用到蛇的身上。

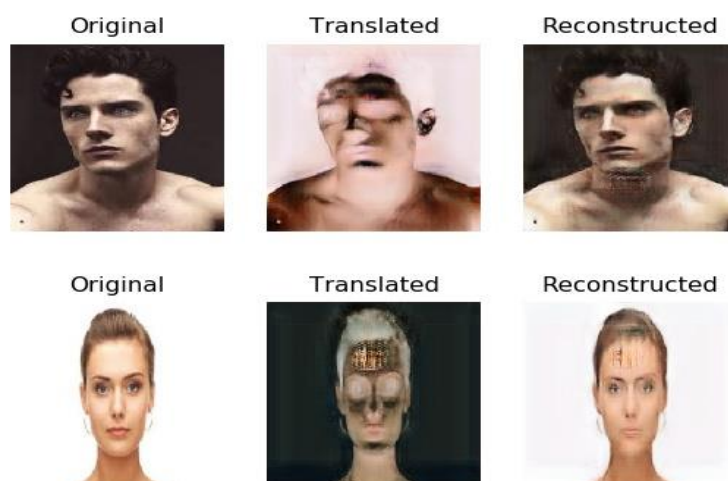
男人與女人:



實驗目標是為了將男生和女生的特徵交換，實驗失敗的原因猜測可能有兩個原因:第一個原因為和前個實驗相同，GAN 抓取的特徵為背景的特徵，而第二個原因為圖片沒有包含細部的的特徵。



男生臉部女生臉部:



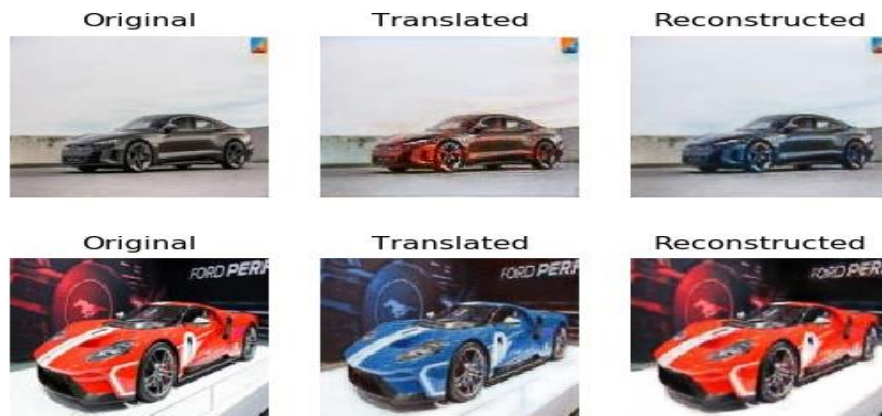
我們嘗試利用男生和女生的臉部來做為資料集，結果變成把背景特徵直接換到臉上去，猜測可能原因為目前所實作的程式，還無法精準的抓取人臉細部的特徵這類複雜特徵，這將會是我們下學期要努力實作的方向。

海綿寶寶與派大星:



接著我們嘗試蒐集特徵較簡單的卡通人物:海綿寶寶和派大星來進行實驗，實驗效果較前幾次的成功，推測原因為卡通人物的特徵簡單，同樣的訓練程式，能達到較好的效果，但是仍舊還有努力的空間，讓 **Generator** 能夠抓取除了顏色以外的特徵。

轎車和跑車的特徵交換：



我們蒐集了多種跑車、轎車的圖片，並對他們進行 CycleGAN，上圖可見，車子的顏色特徵成功轉換，但是車種類的特徵並沒有交換，推測為 GAN 擷取特徵的能力還不夠強，需增加 GAN 的層數來解決這個問題。

實驗的結論：

資料的前處理(圖片的蒐集、篩選)對訓練的結果影響深遠，在人臉部分的訓練，因為所要擷取的特徵為細部的特徵，我們需要增加 GAN 的層數，來加強抓取特徵的能力，而卡通人物和車的部分，我們目前也只能做到簡單的特徵交換(顏色交換)，可能還需要增加其他種類的 GAN 特性來解決這些問題。



## (六)、 參考文獻:

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks Jun-Yan Zhu\* Taesung Park\* Phillip Isola Alexei A. Efros Berkeley AI Research (BAIR) laboratory, UC Berkeley

<https://arxiv.org/pdf/1703.10593.pdf>

Generative Adversarial Nets Ian J. Goodfellow\* , Jean Pouget-Abadie† , Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair‡ , Aaron Courville, Yoshua Bengio§ Département d'informatique et de recherche opérationnelle Université de Montréal Montréal, QC H3C 3J7

<https://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>

Image-to-Image Translation with Conditional Adversarial Networks Phillip Isola Jun-Yan Zhu Tinghui Zhou Alexei A. Efros Berkeley AI Research (BAIR) Laboratory, UC Berkeley

StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation Yunjey Choi<sup>1,2</sup> Minje Choi<sup>1,2</sup> Munyoung Kim<sup>2,3</sup> Jung-Woo Ha<sup>2</sup> Sunghun Kim<sup>2,4</sup> Jaegul Choo<sup>1,2</sup>

<sup>1</sup>Korea University <sup>2</sup>Clova AI Research, NAVER Corp. <sup>3</sup>The College of New Jersey <sup>4</sup>Hong Kong University of Science & Technology

<https://arxiv.org/pdf/1711.09020.pdf>

Recycle-GAN: Unsupervised Video Retargeting.

Aayush Bansal<sup>1</sup> , Shugao Ma<sup>2</sup> , Deva Ramanan<sup>1</sup> , and Yaser Sheikh<sup>1,2</sup>

<sup>1</sup>Carnegie Mellon University <sup>2</sup>Facebook Reality Lab, Pittsburgh

<http://www.cs.cmu.edu/~aayushb/Recycle-GAN/>

[https://www.cs.cmu.edu/~aayushb/Recycle-GAN/recycle\\_gan.pdf](https://www.cs.cmu.edu/~aayushb/Recycle-GAN/recycle_gan.pdf)

GAN 應用與延伸:

<https://arxiv.org/pdf/1703.10593.pdf>

<http://auai.org/uai2017/media/tutorials/shakir.pdf>

GAN 生成對抗網路:

<https://mc.ai/%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-ml-notegenerative-adversarial-network-gan-%E7%94%9F%E6%88%90%E5%B0%8D%E6%8A%97%E7%B6%B2%E8%B7%AF/>

Cycle GAN 參考介紹:

<https://hardikbansal.github.io/CycleGANBlog/>