**題目:**

GAN模擬生成指定人、物動作圖像

**組員:**

黃維熙 40547005S

古佳儫 40547010S

余承寰 40547034S

**指導教授:**

黃文吉教授

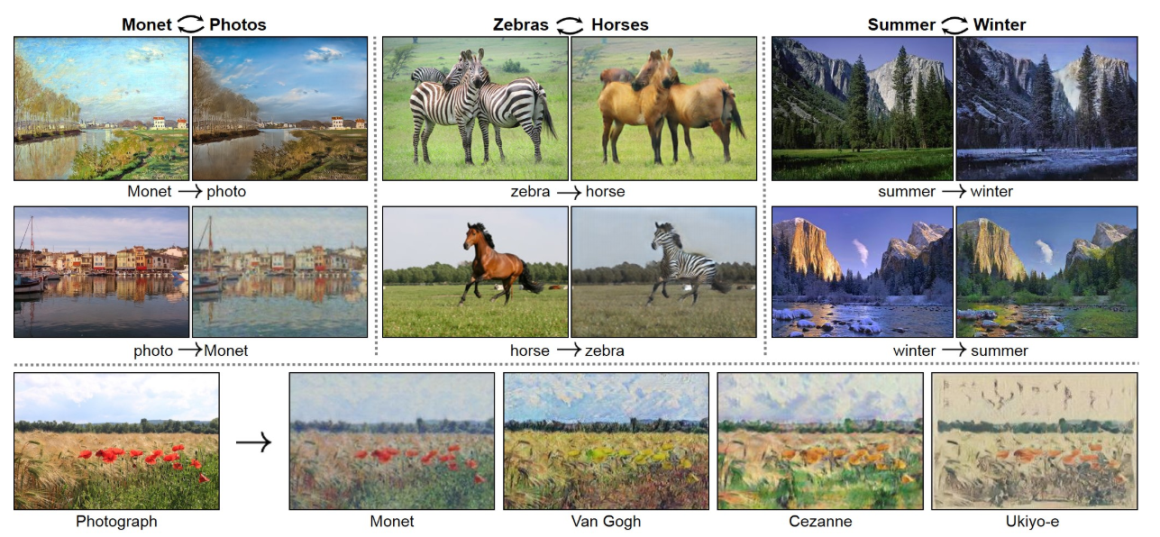
**(一)、 摘要:**

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，前期使用CNN建模並訓練model準確率，讓電腦能夠盡量準確判斷圖片影像，後期利用GAN的方式來實現目標。

**(二)、 研究動機:**

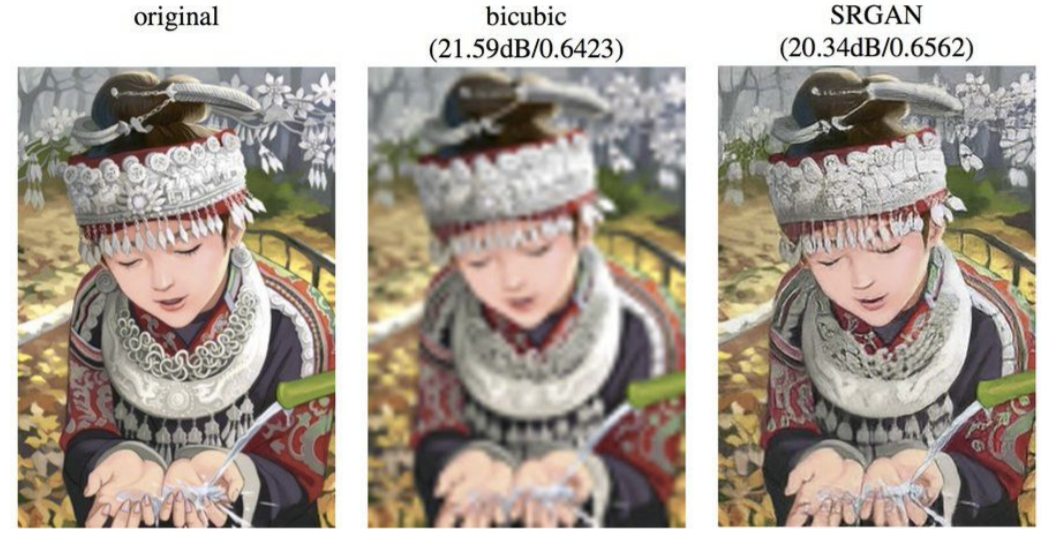
圖像視覺對人類來說是不可忽略的，從眼睛所看見的東西都是圖像的一種，若能利用GAN生成指定人、物做指定動作，就可以使我們不需再花大量人力標註資料，利用GAN模型生成即可，『監督式學習』(Supervised)就變成『非監督式學習』，未來在許多影片或圖片的拍攝，將不需要指定的演員或模特兒，也可以使圖片的風格轉換。

例:圖(一)



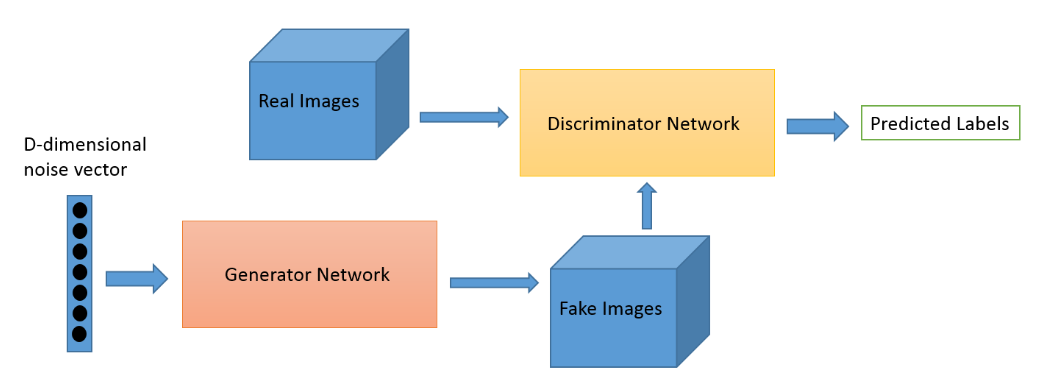
再者若照片解析度不佳的情況下，也能重新合成出一幅高解析度的照片

例:圖(二)



還能將此方法運用在預測物體的變化上。

**(三)、 研究方法及步驟:**

GAN基本架構圖:

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。

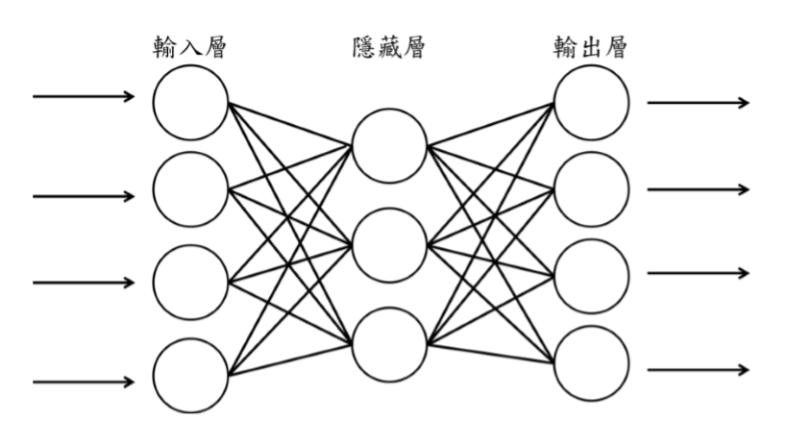
1. 各明星、演員圖像收集與整理:

收集來的資料分為test data跟train data兩個部分，裡面有各種不同人的資料圖片並利用英文字母進行編號分類，例如有三個人，在train data中就會有3個資料夾a、b、c分別存三個人的各種圖片，依此類推不斷增加類別，並將這些資料利用python函式依照設定比例切分成train data與test data。

1. 神經網路模型探討與設計:

本研究預計以類神經網路（neural network）作為模型來進行機器學習。由多種人工神經元連接，形成一個類似生物神經網路。每個人工神經元都是一個激發函數，以另一個人 工神經元的輸出作為輸入，其輸出最為再另一個人工神經元的輸入。

類神經網路基礎架構(圖四)分為三層，輸入層、隱藏層與輸出層，輸入層用來表現輸入的變數，也就是特徵，其神經元數目會依據輸入的狀況而定，隱藏層用來處理輸入層的資料， 並將結果傳遞至輸出層，輸出層會將來自隱藏層的資料輸出，用來表現輸出的變數，依此方式反覆訓練並調整出最好的結果。

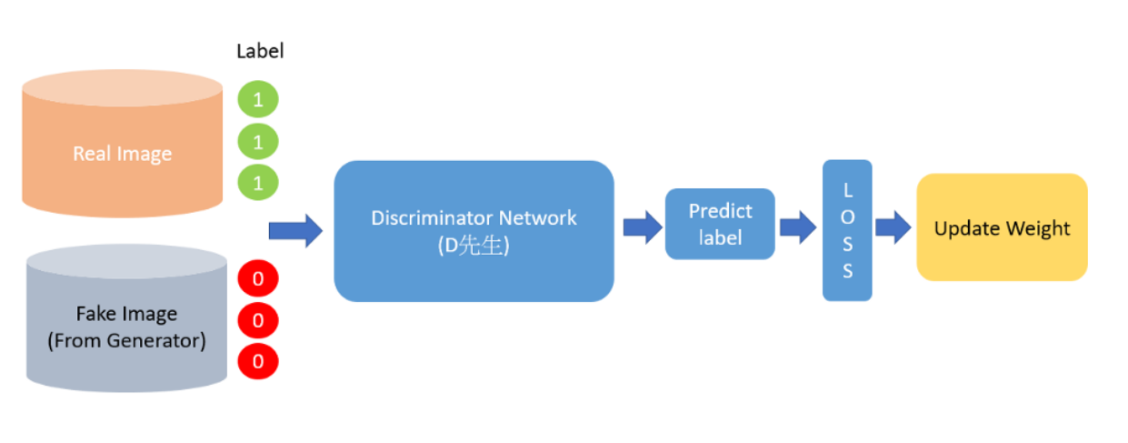
(圖4)

1. 模型訓練（機器學習）:

首先要把收集來的資料與參數初始化，將圖片大小與channels進行調整，接著分別搭建Discriminator Network與Generator Network

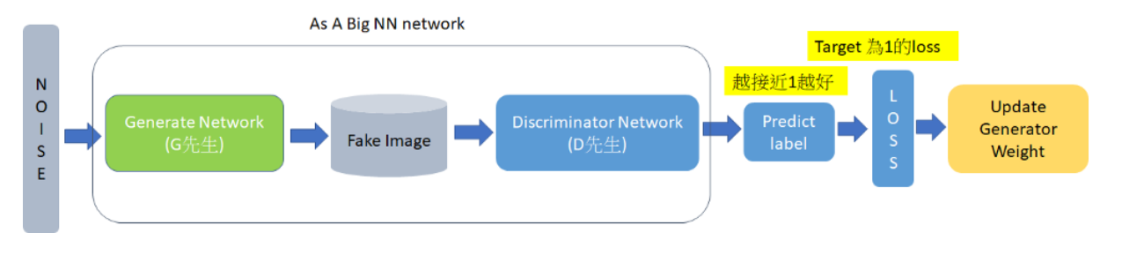
Discriminator Network:

訓練出一個Neural Network可以分辨偽造出來的圖與真實的圖，將透過Generator產生的圖像標記為0(fake image)，並把真實的圖像標記為1

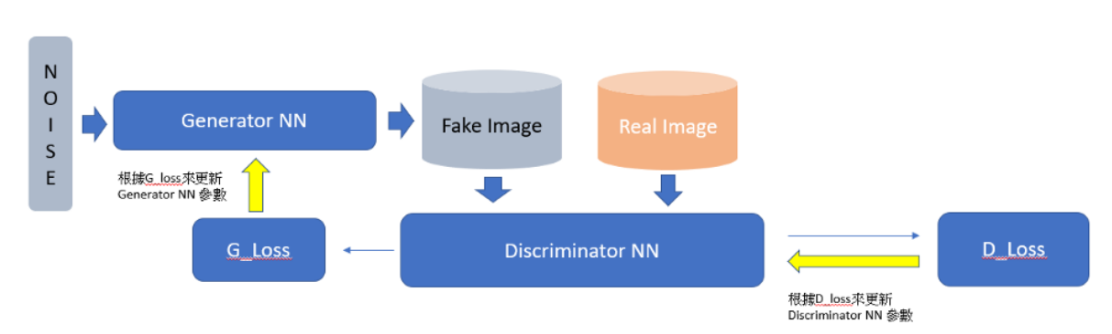


Generator Network:

訓練出一個Neural Network可以讓Discriminator分辨出來的結果越接近真實(1)的結果越好



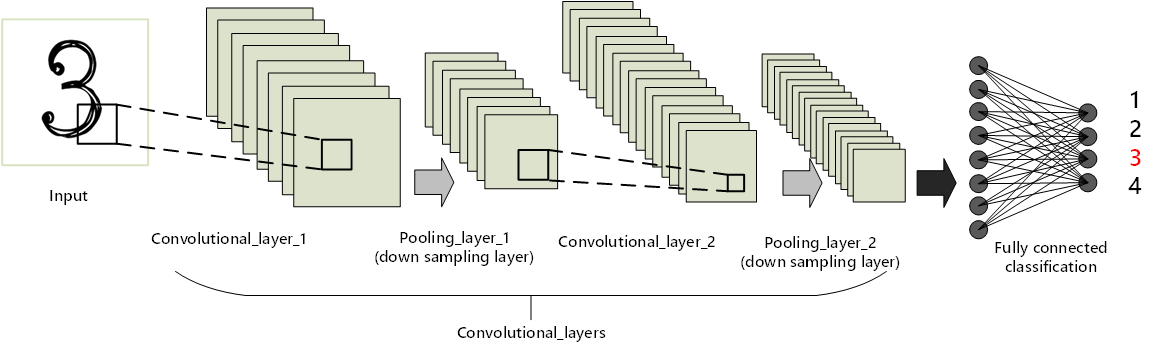
將這兩個神經網路串接起來並進行訓練，在訓練的過程中，將不斷計算Generator輸入進Discriminator後的輸出跟1(真實圖片)的loss(G\_loss)、Generator輸入進Discriminator後輸出跟0(fake image)的loss+真實圖片輸入進Discriminator跟1的loss(D\_loss)，並調整Generator的權重，使得調整過後的Generator，讓Discriminator接近真實(1)的結果



最後將生成的圖像儲存，並輸出結果。

**(四)、 目前進度:**

我們以python3.6作為編譯環境，並使用keras套件來進行訓練與模擬，我們先從實作CNN著手，一開始練習使用keras套件中的mnist數據庫，裏面包含了各種數字的灰階圖像，接著利用 model = Sequential() 建立一個線性堆疊的模型，再用 model.add 建立卷積層與池化層模擬出基礎的CNN訓練模型(如下圖範例)。



有了一些基本概念之後我們改用keras套件中的cifar10數據庫，從數字的辨識進入到各類圖像的辨識，在cifar10的訓練中，我們有嘗試在不同的時機調整learning rate，後來發現在 model.fit() 函式中有一個叫callback的回調函數，我們可以建一個list，讓我們可以在進行訓練的同時根據狀況來微調我們的learning rate、存取最佳權種及網路等等......成功完成cifar10訓練後，我們開始練習使用自己的資料進行訓練，目前在蒐集資料的階段，接下來我們打算要對女明星、女演員進行人臉辨認的訓練。

**(五)、 預期結果:**

本研究預期產出如下:

1.利用蒐集的資料進行前處理:

蒐集、拍攝約1萬張的照片，來當作Generator Network與Discriminator Network的訓練基礎，接著利用Generator依據圖片類別自主產生預估5000張模仿原圖的影像。

2.進行GAN實作:

利用整理好的圖片數據，來產生新的圖片，例如:改變拍照的姿勢、改變頭髮的顏色、改變性別以及模擬出不同人做相同的動作等等，利用GAN的實作，能夠產生出擬真的照片。

3.調整與分析:

根據上述1、2實作的結果取得各數據並進行改善。

**(六)、 參考文獻:**

keras mmist實作參考:

<https://chtseng.wordpress.com/2017/09/23/%E5%AD%B8%E7%BF%92%E4%BD%BF%E7%94%A8keras%E5%BB%BA%E7%AB%8B%E5%8D%B7%E7%A9%8D%E7%A5%9E%E7%B6%93%E7%B6%B2%E8%B7%AF/>

<https://github.com/jonbruner/generative-adversarial-networks/blob/master/gan-notebook.ipynb>

GAN應用與延伸:

<https://junyanz.github.io/CycleGAN/>

<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/>