

機器辨識

大家想到大數據總是又愛又恨，愛的是他的功能強大，只要有足夠的數據，能在幾毫秒得到所需答案，比過去數萬人、同時作業數年的效果更快更精準，相對的也可能涉及到隱私問題。機器視覺中的人臉辨識已是最基礎的，各種生物、物品的辨識所造成的人力精簡及時間的省節，都可藉由深度學習來達成，且準確度更不會有人為的疏失！

(一) MNIST、Cifar-10 辨識

MNIST 是一種黑白的手寫數字資料集，分類為 10 類 (0~9)，在踏入CNN這領域時，大多數人都會使用這資料集當作練習，可以將解決 MNIST 視為深度學習的 "Hello World!"。

Cifar-10 資料集內含 10 種類別的圖片，分別是飛機、汽車、鳥、貓、鹿、狗、青蛙、馬、船、卡車，且屬於彩色圖片。

1. 實際執行

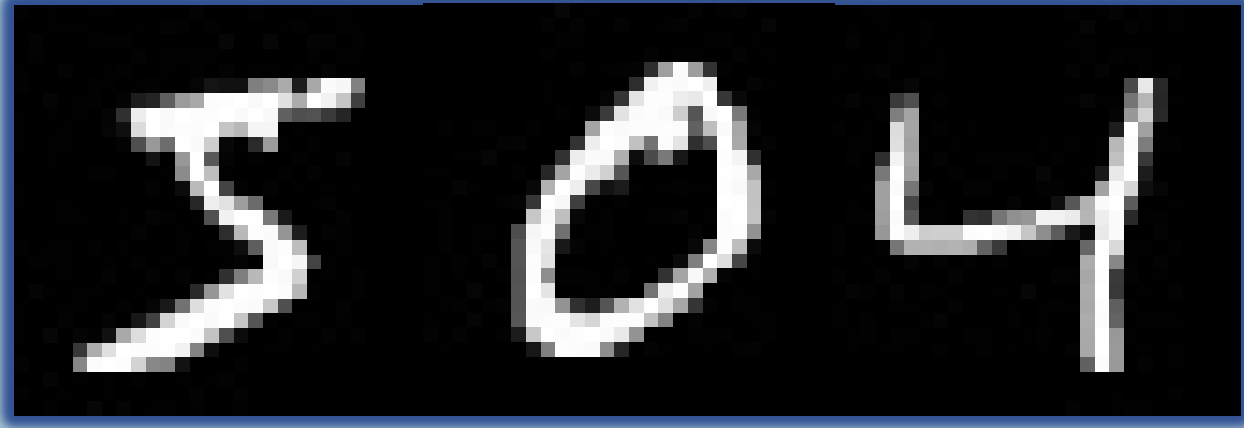


圖3 偵測為5、0、4

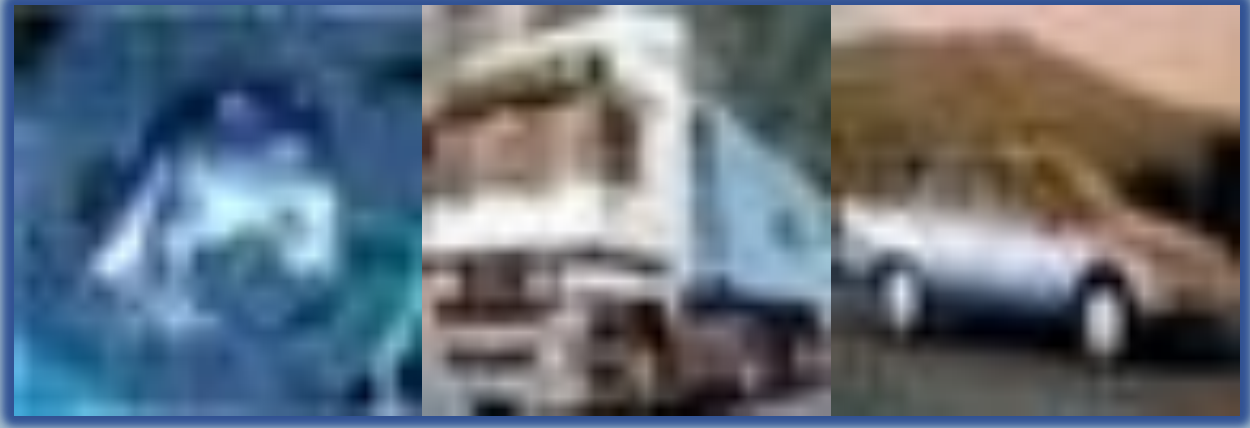


圖4 偵測為青蛙、卡車、汽車

2. 結論

第一次辨識成功後，我非常驚訝只用數學運算就可辨識物件，難怪現在深度學習如此有人氣，想不到未來這技術能有多強大，真的很佩服想到深度學習的人。

(二) 貓狗偵測

試驗 YOLOv8 技術，採用 yolov8l-seg 訓練模型，較講求精準。

1. 實際執行



圖5 貓狗

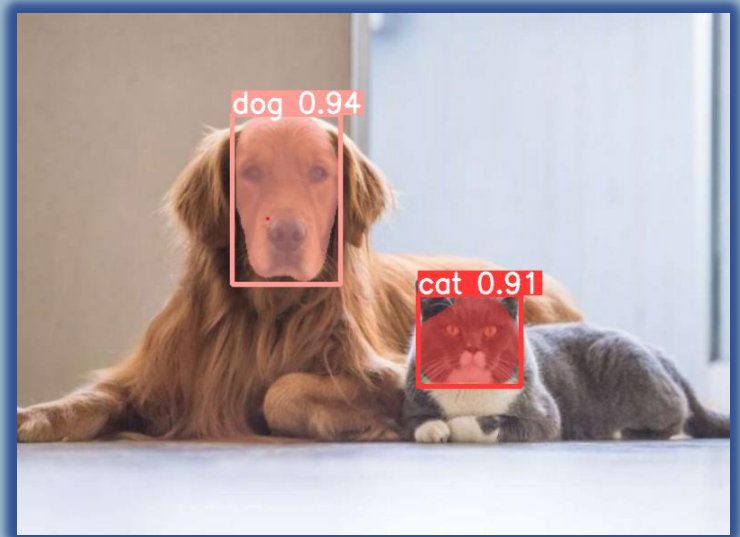


圖6 貓狗偵測

2. 結論

製作資料集時手真的差點抽筋，非常痛苦的三小時，其餘都相對簡單。意外的是，只需120張圖片就可做出正確率如此高的辨識模型，讚嘆 YOLO。

(三) CSGO 自動瞄準

在逛 Roboflow 網站時，意外找到 CSGO(射擊遊戲) 人物偵測的資料集，馬上意識到這資料集是為了外掛而生。使用 YOLO 預測人物位置，再移動鼠標至人物，加點隨機數，這不就是個無法防範的外掛(一般外掛以破解遊戲內部資料為主)。

1. 實際執行

準備就緒後，開一場電腦場(只有我一個玩家)，不出意外的話就要出意外了。偵測不給力、時間差、容易偵測到隊友、鼠標移動，種種問題浮現，主要還是偵測太慢且不准。輸入到偵測完成，敵人早已移動到下個位置了，如果能一直偵測到，這問題就可以用預測的方式解決，但這偵測模型只要距離稍微遠一些，就完全偵測不到了。



圖7 SCGO畫面



圖8 SCGO偵測

2. 結論

可說是完全的失敗，不過我相信只要優化過偵測模型，還是能成為很好的外掛。但為什麼這種外掛模式不怎麼聽過，很明顯效益不高，與傳統外掛相比，傳統外掛直接竊取內部資料簡單、快速、準確、更強。相反這種外掛還需要看到敵人才能有所反應，各項技能皆差距傳統外掛一大截，唯一好處是難被抓包，不過遊戲端應該也可反用AI抓這種外掛。

總結

透過最基礎的卷積神經網路技術的測試以及進一步的應用，可以發現其在辨識與定位方面的應用潛力非常大。除了手寫數字辨識、貓狗辨識等基礎應用外，還可以應用在許多重要的場景中，例如自駕車、軍事、失蹤人口協尋等等。然而，應用卷積神經網路技術時也必須注意遵守相關法律法規以及道德規範，確保技術應用的合法性與合理性。總體而言，卷積神經網路技術的應用前景廣闊，將對未來的生活、工作、社會等方面產生深遠的影響。

藝術

透過AI的協助更可擴大藝術家、非藝術家的創意發想，只要想得到的，不只都能創作出來，且可發揮比想到的更大更精準的效果，並引領創作者達到另一個層次與想像！換句話說，也不再需要藝術家了，人人都可以透過此設計就比藝術家更厲害了。

(一) 神經風格轉換 (Neural Style Transfer, NST)

過去幾個月流行將現實圖片轉換成動漫風，而我也被這技術驚豔到，因此學習神經風格轉換。

1. 實際執行

選擇本人照片當內容圖，而使用 梵谷《星夜》(The Starry Night)、康丁斯基《第七號構圖》(Composition VII)、葛飾北齋《神奈川沖浪裏》(The Great Wave off Kanagawa) 為風格圖。



圖9 風格轉換

2. 結論

神經風格轉換為神經網路影像風格轉換的開山始祖，但每次要生成一張新的風格化影像需要重新訓練一次 VGG-19，效率頗低，且完全不能即時轉換。

(二) DCGAN

發現神經風格轉換無法實現即時轉換後，嘗試找出何種方式可即時轉換。雖然當時沒找到，但找到生成對抗網絡這有趣技術，只要給無標註的圖片，就可隨機生成出同類別圖片。

1. 實際執行

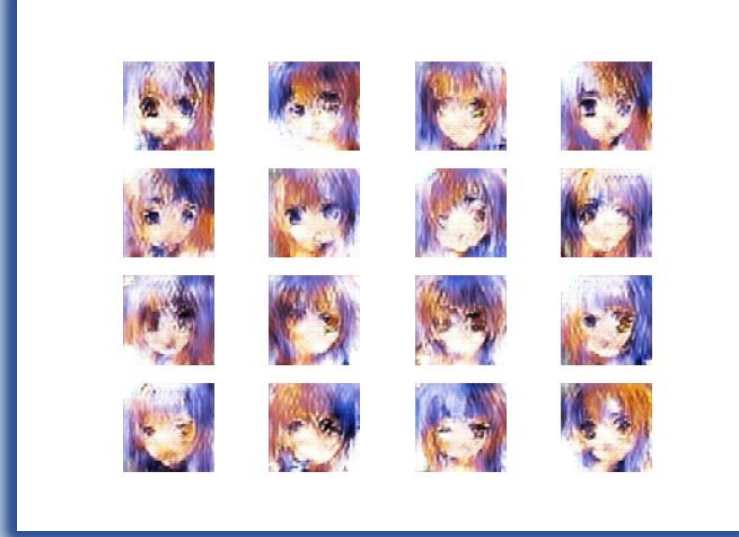


圖10 版本一 大小64*64

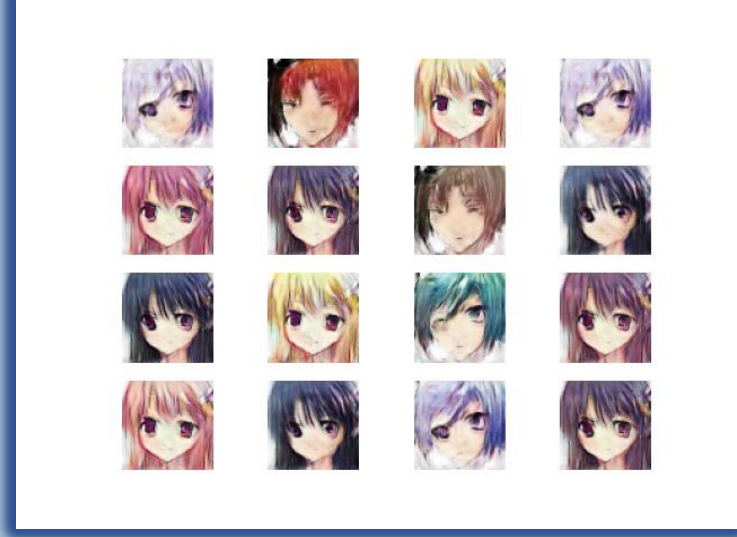


圖11 版本二 大小64*64

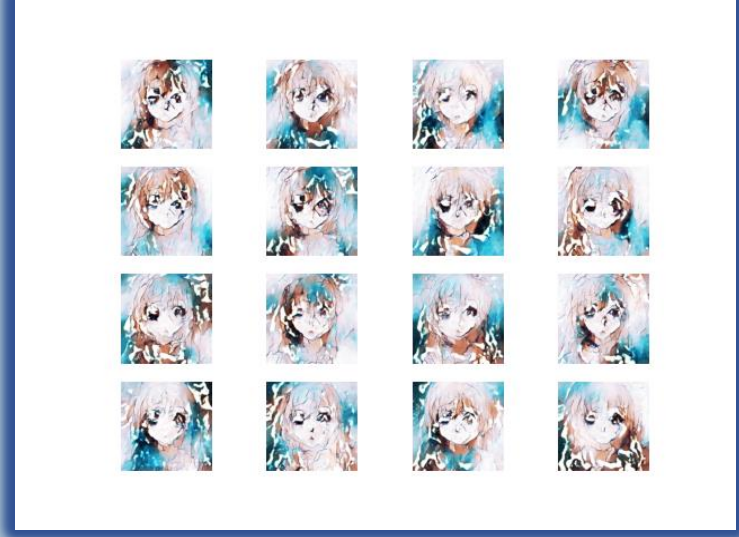


圖12 版本一 大小256*256

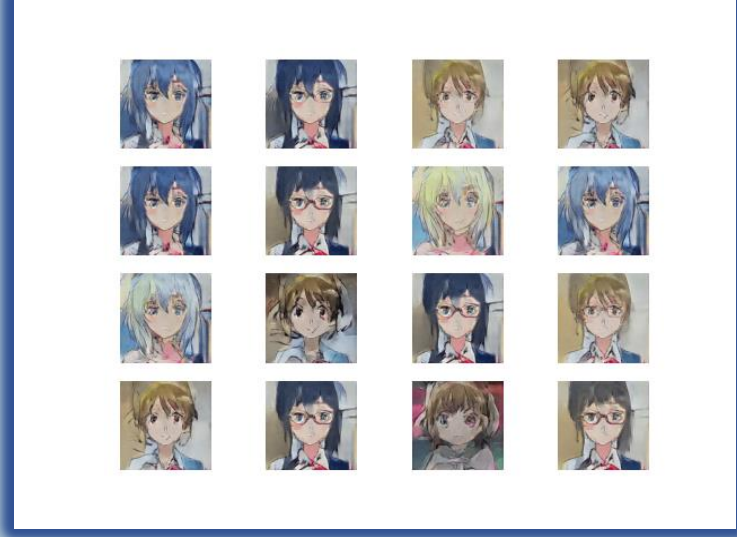


圖13 版本二 大小256*256

2. 結論

只能說待加強，感覺模型還可再增強、學習率也可再調，生成對抗網絡是由兩個模型互相比較而來，兩個模型要一起成長、進步，不能其中一個的能力強過多於另一個，因此微調其中參數也是很重要的一環。

(三) CycleGAN

生成對抗網絡有非常多變體，多到以英文字頭+GAN取名的變體都快佔滿了，當中我無意看中 CycleGAN，希望藉此達成即時轉換功能。

1. 實際執行



圖14 現實轉莫內圖片

2. 結論

DCGAN 都如此難訓練了，更別說 CycleGAN 了。原理是很有趣、巧妙，但產生的圖形缺乏多樣化，訓練時損失震盪巨大，成效差強人意，並不是太有效的風格轉換方式。

總結

使用 GAN 技術進行影像風格轉換是一種很有前景的應用，它可以讓我們將一張圖片的風格轉換成另外一種風格，例如將印象派畫作的風格應用到一張照片中。儘管成果目前並不十分樂觀，但這個技術仍有很大的發展空間，特別是隨著AI技術的不斷進步，未來我們可能能夠更好地掌握影像的細節，進一步提高影像風格轉換的效果。

這項技術對美化世界有很大的潛力，它可以幫助藝術家以更快、更有效的方式創作出令人讚嘆的作品，同時也可以創造出更多具有價值的商品。除此之外，影像風格轉換還可以應用在影視、遊戲等領域，讓觀眾享受到更多美感和豐富的視覺體驗。總之，這項技術有很大的潛力，值得我們繼續關注和發展。