

Group 6 資料科學導論 Final Project Report

How Bad Are You?

106502504 倪滙渝 、 106502520 賴浩維

動機

由於 Instagram 的「Which Disney Are You」濾鏡很受到大家的歡迎，因此我們想運用這個概念來當作我們 Final Project 的發想。因為我們很喜歡看天之驕女，所以就利用了裡面的角色當作素材，並選擇了八個最具代表性的壞人。

資料蒐集

我們使用 PotPlayer 的軟體來輔助我們從 Youtube 影片中截圖，總共截取了 30 集的照片當作訓練資料。由於有一些人物並不是在我們選的八個人裡面，所以後來也有透過人工的方式來篩選掉，並只選擇單個人物的畫面。接著我們再 Preprocess 使用 MTCNN 來偵測人臉的範圍，並將角色分類，以 0 到 7 標記，以便後續的訓練。



方法與結果

有了前述處理好的資料後，我們使用 VGGFace，並使用 retrain Resnet50，加上一層 Flatten Layer 以及兩層 Dense Layer。雖然說標示的 Label 為 0 到 7，但是在訓練時以 categorical type 來訓練，cross entropy 的部分也使用 categorical cross entropy，以避免數值上的差別影響到最後判斷的結果。而在訓練階段有設定 early stop，以防止 over fitting。下圖為訓練結果示意圖。

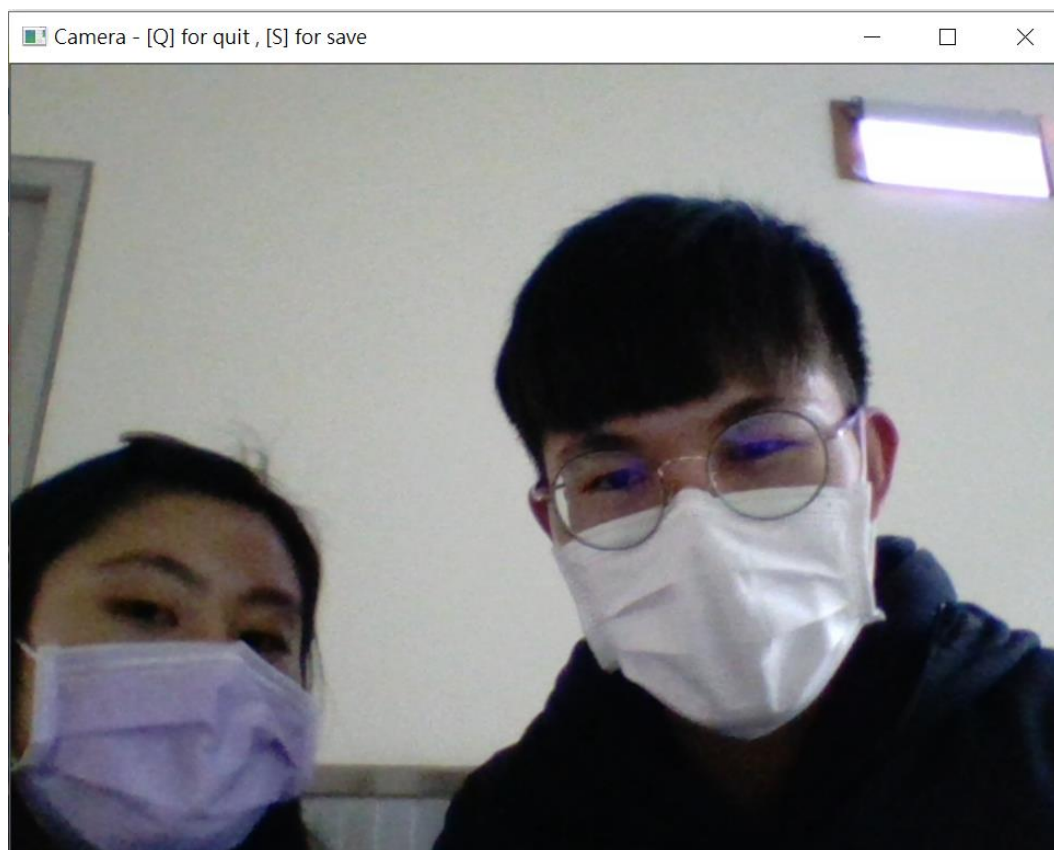
```
Epoch 00011: saving model to C:\Users\mcu\Desktop\model_best.h5
Epoch 12/20
18/18 [=====] - 166s 9s/step - loss: 1.0027e-05 - accuracy: 1.0000 - val_loss: 0.3247 - val_accuracy: 0.9621

Epoch 00012: saving model to C:\Users\mcu\Desktop\model_best.h5
Epoch 00012: early stopping
```

介面介紹



上圖為我們 Final Project 的介面示意圖，**左邊**照片顯示選擇的圖像，**右邊**照片則顯示判斷的角色及其名字，除此之外，還可以利用電腦照相機拍照選擇，照相機的部分則利用 **S** 鍵拍攝儲存，**Q** 鍵關閉。



檔案說明

final.ipynb :

用來訓練與 preprocess 的程式檔，會輸出訓練好的 model (model_best.h5)。

Demo.ipynb :

為此次 Demo 的 UI 介面，運用 final.ipynb 輸出的 model 來偵測。

若需要執行程式，僅需將 train 好的 model 與 Demo.ipynb 放在一起並下載所需的 package 即可。

心得

雖然這樣的概念已經被執行過了，但是換個主題，並運用上課所學的概念，從實做的過程中，可以更深刻的學習到資料科學以及機器學習的強大，也期許我們在未來的道路上，持續跟隨 AI 的潮流，並發想出更好的方法。