

Intelligent Integration of Enterprise

Individual Project

Real Time Face Mask Wearing Detection

109034534 梁芷蘋

指導教授：邱銘傳 教授

目錄

大綱	5
Abstract.....	6
1. 前言	7
2. 文獻回顧	7
3. 方法論	8
3.1 CNN 模型架構	8
3.2 即時鏡頭辨識	8
3.3 系統架構	9
4. 個案研究	9
4.1 資料蒐集	9
4.2 資料前處理	10
4.2.1 資料標準化	10
4.2.2 資料擴增	11
4.3 模型實作結果	11
4.4 導入模組的情境前後比較	13
5. 結論與未來展望	13
參考資料	15

圖目錄

圖 一、CNN 模型架構	8
圖 二、即時鏡頭程式碼	9
圖 三、系統架構圖	9
圖 四、原始資料集圖片資料	10
圖 五、資料色彩與大小標準化	10
圖 六、圖片資料擴增	11
圖 七、Model A 模型架構圖	12
圖 八、模型實作成果	13

表目錄

表一、5W1H.....	7
表二、超參數調整	11
表三、模型績效比較	11
表四、As-Is 與 To-Be 比較表	13

大綱

在疫情嚴峻的現在，口罩成為人人不可離身的必備防疫工具。各機關單位開始規定，進出特定場所如車站、百貨公司等皆須配戴口罩。有些地方政府更增加一些罰則來強制民眾配戴口罩。為此，執法人員或是單位工作人員對於是否配戴口罩的檢查就顯的非常重要。然而，人力不足或是疲勞而遭成疏失的情況卻容易出現在現在的口罩檢查站中。因此，為提升口罩檢查站的效率，本計劃將使用 CNN 卷積神經網路來建置一個即時口罩檢查模組，使攝影鏡頭在偵測到人臉後，可以即時標示出有戴口罩以及沒有戴口罩的人們。

關鍵字：口罩、CNN 卷積神經網路、即時影像偵測

Abstract

The epidemic is severe these days, masks have become an indispensable preventive tool for everyone. Various agencies have begun to stipulate that masks must be worn when entering and exiting specific places such as stations and department stores. Some local governments have increased penalties to force people to wear masks. For this reason, law enforcement officers or unit staff need to check whether they wear masks. However, the lack of human resources or fatigue caused by negligence is easy to appear in the current mask checkpoints. Therefore, to improve the efficiency of mask inspection stations, this plan will use CNN convolutional neural networks to build a real-time mask inspection module, so that after the camera lens detects a human face, it can instantly indicate whether a mask is worn or not.

Keywords: mask, convolutional neural networks, real-time image recognition

1. 前言

在人人自危的疫情下，有沒有配戴口罩成為人人必做的國民風潮。沒有配戴口罩的人，極有可能成為防疫漏洞，人人皆害怕這樣沒有配戴口罩的人是否會出現在自己身邊，將病毒傳染給自己。因此，能夠即時偵測人們是否有配戴口罩，將可有效的馬上進行勸阻或是罰則，以確定人人都有配戴口罩，降低病毒散播的可能性以及健康的人被感染的機會。

為釐清整個情境，其 5W1H 表列如下：

表一、5W1H

WHO	執法人員、口罩檢查站人員
WHAT	檢測是否配戴口罩
WHEN	政府或單位強制人員配戴口罩時
WHERE	有可能造成群聚感染的環境
WHY	降低檢查站人員誤判的情況發生
HOW	利用 CNN 卷積網路結合即時影像偵測

有了這些問題定義，將整個情境釐清，可有助於系統模組的開發進展與過程。

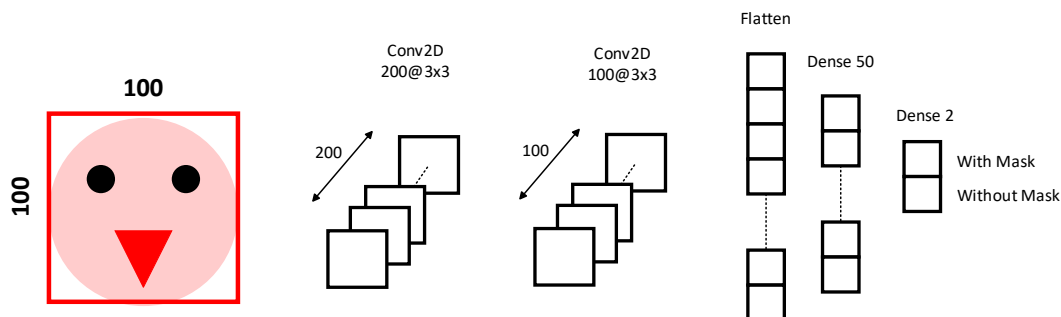
2. 文獻回顧

目前針對偵測口罩配戴的方法有很多種，近年來多主要使用深度學習進行有無配戴口罩的分類。深度學習的方法主要有 CNN(包含 R-CNN、Fast R-CNN 等)與 YOLO。其中，針對 Optimizer 的部分，使用 Adam 的效果最佳，辨識率可以達到九成 (Mohamed & Gunasekaran, 2020)。

CNN(卷積神經網路 Convolutional Neural Networks)一種深度學習的電腦視覺演算法，同時也是使用有標籤的資料的一種監督式學習法，經常被使用在影像分類問題上，卷積神經網路可以藉由擷取與影像形狀相關的特徵進行影像辨識，架構主要可分為卷積層、線性整流函數層、池化層與全聯階層(Girshick, Donahue, Darrell, & Malik, 2014)。

3. 方法論

3.1 CNN 模型架構



圖一、CNN 模型架構

本計畫使用的 CNN 模型為 Sequential CNN model，使用 Keras 即可達成。此模型適用於簡單的層堆棧，其中每一層正好具有一個輸入量和一個輸出量。標準化後的圖片資料經由兩層捲積層依序處理後，再將結果平坦化後得到神經網絡，模型會回歸到 Dense 進行二元分類，即可得到有無戴口罩的分類。

3.2 即時鏡頭辨識

```
face_clsfr=cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
source=cv2.VideoCapture(0)

labels_dict={0:'MASK',1:'NO MASK'}
color_dict={0:(0,255,0),1:(0,0,255)}

while(True):
    ret,img=source.read()
    gray=cv2.cvtColor(img,cv2.COLOR_BGR2GRAY)
    faces=face_clsfr.detectMultiScale(gray,1.3,5)

    for x,y,w,h in faces:
        face_img=gray[y:y+w,x:x+w]
        resized=cv2.resize(face_img,(100,100))
        normalized=resized/255.0
        reshaped=np.reshape(normalized,(1,100,100,1))
        result=model.predict(reshaped)

        label=np.argmax(result,axis=1)[0]

        cv2.rectangle(img,(x,y),(x+w,y+h),color_dict[label],2)
        cv2.rectangle(img,(x,y-40),(x+w,y),color_dict[label],-1)
        cv2.putText(img, labels_dict[label], (x, y-10),cv2.FONT_HERSHEY_SIMPLEX,0.8,(255,255,255),2)

    cv2.imshow('LIVE',img)
    key=cv2.waitKey(1)

    if(key==27):
        break

cv2.destroyAllWindows()
source.release()
```


圖 二、即時鏡頭程式碼

利用 OpenCV 的 Cascade Classifier 來導入臉部辨識工具，並使用 OpenCV 的 Video Capture 影片讀取功能來串流攝影機的影像，並且讀取圖像數據。為提升整個模組的效率，會將影像圖片由彩色轉為灰階，並且將該影像圖片與 CNN 模組進行比對，並將有無戴口罩的臉部影像套用到即時影像畫面以不同顏色框住。

3.3 系統架構

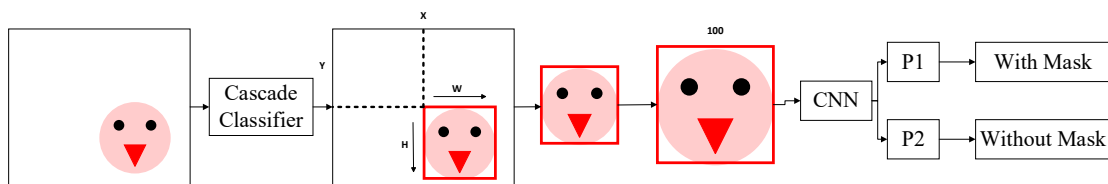


圖 三、系統架構圖

系統首先經由攝影鏡頭拍攝動態影像，再將這些影像轉換成一張張的圖片，經由運算判斷出圖片中的人臉，並將人臉的大小標準化。之後，透過 CNN 將標準化後的圖片進行分析，並套至即時影像畫面，若有戴口罩則將面部以綠色外框框住；若無，則以紅色外框框住。

4. 個案研究

4.1 資料蒐集

本系統總共收集了兩種資料：一是有戴口罩的面部照片，另一則是沒有戴口罩的面部照片。原始的資料集為 Github 上由 Prajna Bhandary 所提供，無口罩面部資料共 480 筆，而有配戴口罩的照片數量為 482 筆。訓練集的資料是來自原始資料集的 90%，剩下的 10% 則用做測試集資料，驗證集的資料則是來自訓練集資料中的 20%，也就是佔原始資料集的 18%。



圖 四、原始資料集圖片資料

4.2 資料前處理

4.2.1 資料標準化

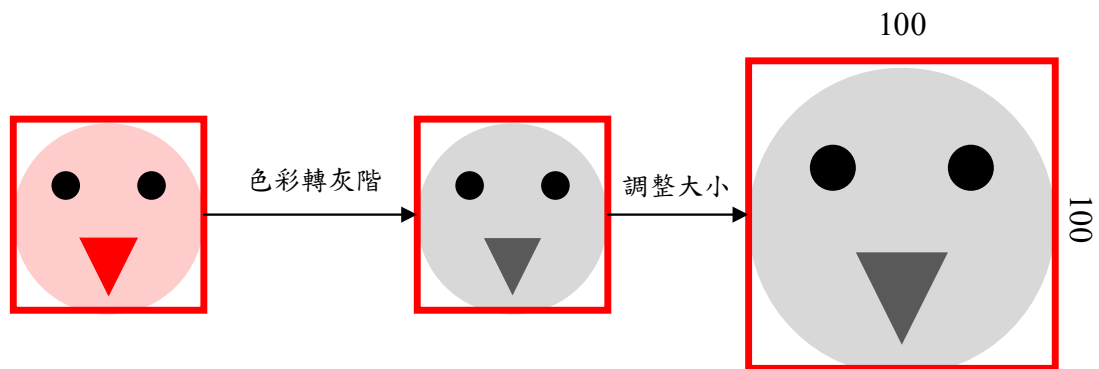


圖 五、資料色彩與大小標準化

圖片在做前處理時，除了資料清理（包含去除重複之資料、錯誤標示的資料、缺失或空的資料值與異常值）外，還需要考量圖片大小，來降低訓練模型所會造成的誤差，因此本模組會將這些人臉圖片的大小轉換為同一大小。除此之外，為了使即時影像偵測能夠具體實現而降低出現延遲的機會，本系統會將原本一個像素佔 24bit 的 RGB 彩色像素轉化為僅佔 8 bit 的灰階圖片，如此一來可大幅增加運算效率與速度。

4.2.2 資料擴增

為增加模型辨識的精準度，除了原先的資料 962 筆，再從原本的資料庫隨機抽選增加 316 張擴增的圖片資料。同時，增加 Kaggle 上的 Face Mask Detection 資料庫 825 張，共 2,102 筆資料來跨大整個資料集。透過擴增厚的資料集，來設計出最佳績效的模型。



圖 六、圖片資料擴增

4.3 模型實作結果

本計畫模型主要使用的參數調整如下表：

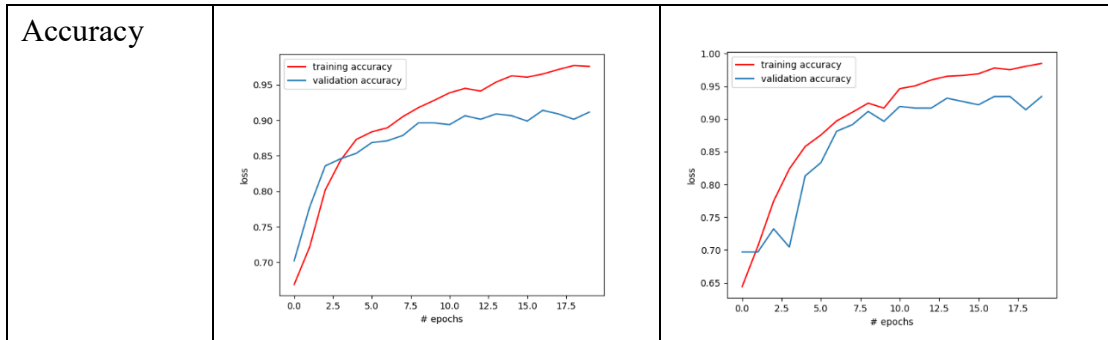
表 二、超參數調整

	Model A	Model B
Conv	100	150
Dense	50	100
Activation	ReLu	ELU
Optimizer	Adam	Adagrad

接著，進行測試，可以得到兩個模型的績效如下表：

表 三、模型績效比較

	Model A	Model B
Loss		



由表可知，以 accuracy 與 loss 來看，Model B 模型的表現較佳。而 Model B 的模型架構如下圖。

```

Model: "sequential_1"
-----
Layer (type)                Output Shape              Param #
-----
conv2d_1 (Conv2D)           (None, 98, 98, 200)      2000
activation_1 (Activation)    (None, 98, 98, 200)      0
max_pooling2d_1 (MaxPooling2 (None, 49, 49, 200)      0
conv2d_2 (Conv2D)           (None, 47, 47, 150)      270150
activation_2 (Activation)    (None, 47, 47, 150)      0
max_pooling2d_2 (MaxPooling2 (None, 23, 23, 150)      0
flatten_1 (Flatten)         (None, 79350)            0
dropout_1 (Dropout)         (None, 79350)            0
dense_1 (Dense)             (None, 50)                3967550
dense_2 (Dense)             (None, 2)                 102
-----
Total params: 4,239,802
Trainable params: 4,239,802
Non-trainable params: 0

```

圖 七、Model B 模型架構圖

選出最佳的模型後，將模型打包成 model 檔，匯入攝影機即時偵測模組，即可進行即時的口罩偵測識別。由模型實作成果圖可以看到，有戴口罩者即被綠色框框住，並標註「MASK」；而無戴口罩者則被紅色框框住，並標註「NO MASK」。一旦原本有戴口罩被綠色框住的對象脫下口罩，系統即時自動將綠色框轉為紅色框；反之，當沒有戴口罩的對象戴上口罩，紅色框會自動即時將紅色框轉為綠色框。



圖 八、模型實作成果

4.4 導入模組的情境前後比較

表 四、As-Is 與 To-Be 比較表

	As-Is	To-Be
人力資源	多	少
是否因時間降低辨識率	是	否
工作時長限制	有	無

透過上表可以得知，模組導入後，可有效降低人力資源應用以及因疲乏可能造成的失誤，同時，模組不會有換班交接問題與工作時長限制，可以有效的遵守勞動基準法，避免過長或是不合理的工作時間出現，可 24 小時不間斷的運行。

5. 結論與未來展望

透過 CNN 卷積神經網路可有效的判別人員是否有配戴口罩，結合影像鏡頭可即時偵測人員配戴口罩與否。因應口罩商業客製化的興起，口罩的顏色與形狀已不限於平面、N95、藍色、白色、綠色與粉色，增加了更多五花八門款式，將

增加辨識的困難度。因此在未來若能增加更多口罩種類的訓練集將可提升此系統的準確度，若還能檢查口罩是否配戴正確（例如，是否有將口鼻確實遮好），並結合訊號傳輸技術，將提醒配戴簡訊以及柵門裝置結合偵測系統，可以在偵測到無配戴口罩的人後，將提醒簡訊直接傳入該人的手機；在柵欄偵測到無配戴口罩的人，可即時放下柵欄，俾至該人員的通行。如此一來，無人化科技防疫的效果將能有效降低人力資源運用以及增加防疫的效果。

參考資料

1. 訓練集 1 : <https://github.com/aieml/face-mask-detection-keras>
2. 訓練集 2 : https://www.kaggle.com/andrewmvd/face-mask-detection?fbclid=IwAR2TDmsGNBsdXZB6uqZrW5vwwwz15bSza3-vyTi7J5j_b1SYMdbdcOH91Hw
3. Face Mask Detection , <https://www.youtube.com/watch?v=d3DJqucOq4g&list=LL&index=3&t=101s>
4. Visual C++ Clr 彩色圖片轉灰階 , <https://medium.com/@s12121296simon/%E5%BD%B1%E5%83%8F%E8%99%95%E7%90%86-%E4%BA%8C-visual-c-clr-%E5%BD%A9%E8%89%B2%E5%9C%96%E7%89%87%E8%BD%89%E7%81%B0%E9%9A%8E-f6f9208277cb>
5. 利用 OpenCV 抓取相片中的臉部數據 , <https://hardliver.blogspot.com/2017/07/opencv-opencv.html>
6. OpenCV 擷取網路攝影機串流影像 , 處理並寫入影片檔案教學 , <https://blog.gtwang.org/programming/opencv-webcam-video-capture-and-file-write-tutorial/>
7. Loey M, Manogaran G, Taha MHN, Khalifa NEM. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection. *Sustain Cities Soc.* 2020 Nov 12:102600. doi: 10.1016/j.scs.2020.102600.
8. R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.
9. The Sequential model , https://keras.io/guides/sequential_model/
10. CNN 模型設計 , <https://ithelp.ithome.com.tw/articles/10192028>
11. 卷積神經網絡介紹(Convolutional Neural Network) , <https://medium.com/jameslearningnote/%E8%B3%87%E6%96%99%E5%88%86%E6%9E%90-%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-%E7%AC%AC5-1%E8%AC%9B-%E5%8D%B7%E7%A9%8D%E7%A5%9E%E7%B6%93%E7%B6%B2%E7%B5%A1%E4%BB%8B%E7%B4%B9-convolutional-neural-network-4f8249d65d4f>