

An Smart Automatic Door Lock and Visitor Management System Using Deep Learning

人工智慧門鎖迎賓管理系統

專題組員：吳思葶，黃潔金，柳譯筑

專題編號：NTPUCSIE-108-009

執行時間：2019 年 7 月至 2020 年 6 月

國立臺北大學 資訊工程學系

1. 摘要

隨著近年來智慧行動裝置的普及，物聯網在各領域的應用上迅速發展。藉著物聯網裝置，不但可以促成人類日常生活的自動化，更能提升人們的生活品質。

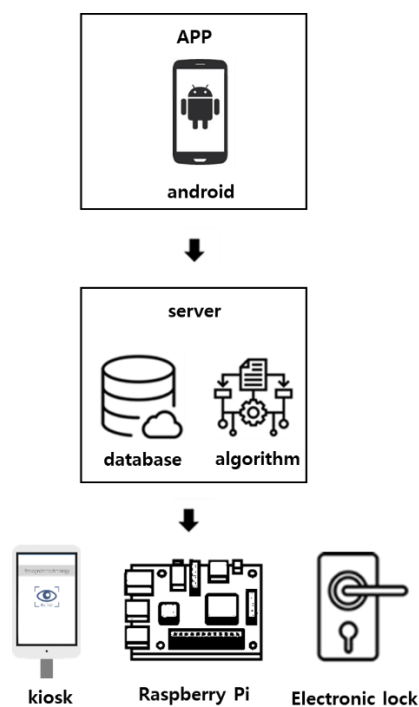
觀察到居家安全一直是民眾的重要考量，我們若是可以透過面部快速辨識使用者，不但可以達到居家安全的訴求，也可以讓生活更為便利，此外對於家中沒有保全系統的一般家庭，此系統可以避免昂貴的費用及額外的施工來達到智慧家庭安全的訴求。

藉此發想，本研究計畫結合居家安全與透過 AI 人臉辨識服務，整合手機 app 端、用樹莓派控制的智慧門禁系統、展示用的 kiosk 機台。在手機 app 端可以上傳照片、新增並儲存使用者資料，透過在 kiosk 前的攝影機拍照，利用 AI 人臉辨識是否為登記過的使用者，藉此決定能否打開門鎖，並提供 AI 明星臉偵測服務：迎賓小遊戲。藉由這個研究項目，希望可以將電子門鎖、KIOSK、手機做一個結合，實現『智慧城市』的理念。

2. 簡介

本專題計畫設計了一套系統，結合了手機應用程式與利用樹莓派控制的門鎖和用來呈現的 KIOSK 介面。加入以下三種重要的元素。(一)門禁管理系統：使用者加入系統會員後，在進門前拍照辨識，成功便將門開啟。(二)人臉辨識系統：使用 google 的

facenet 模型，達成 one shot learning 就能進行身分的辨識。(三)明星臉偵測：為了讓門禁系統增添趣味性，使用者在辨識人臉時，會順便依據當下拍的照片，呈現出長相最為相似的明星資訊，因此每次拍攝皆會有不同的成果。



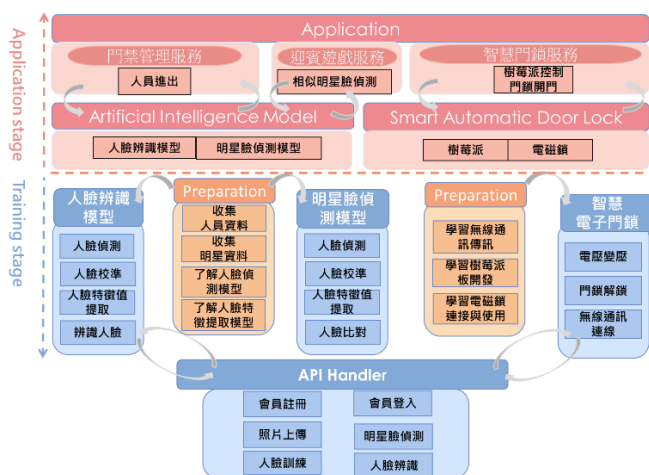
圖一 系統應用情境說明

根據以上四個重要元素，圖一說明本系統應用的情境，app 上傳照片傳送至伺服器，伺服器進行 AI 人臉訓練增加的新照片。接著，使用者便能到 KIOSK 前拍照，並將資料傳送到伺服器，以進行 AI 人臉辨識的運算。接收伺服器傳送的辨識成果，若成功，則呈現出使用者的名字以及相似的明星臉照片，還有相似度，同時用 socket 連線

樹梅派，控制電磁鎖開啟，若失敗，則會請使用者再次點選辨識的按鈕，重新辨識。

3. 專題進行方式

根據第二節的情境描述，系統實作如圖二所示可以分為人臉辨識系統、明星臉偵測系統、伺服器，以下將分別詳細介紹：



圖二 本系統之系統架構

3.1. 人臉辨識系統



圖三 人臉辨識步驟

操作步驟講解：

3.1.1 使用 opencv2 執行從攝影機讀取照片

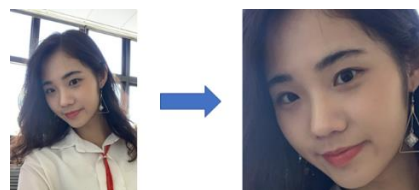
本系統主要會用到 opencv 函式庫，VideoCapture，執行從攝影機讀取照片的動作，並使用 cv2.COLOR_BGR2GRAY，將讀取的照片轉化為灰階的圖像，以降低運算的複雜度。

3.1.2 將讀取的照片運用 dlib 的人臉偵測演算法，偵測出人臉

再分辨人臉之前，我們需要先從照片中找出人臉的位置為何，因此我們將會使用 Dlib 中的人臉偵測演算法 detector，此演算的流程使用方向梯度直方圖（HOG）的特

徵加上線性分類器（linear classifier）、影像金字塔（image pyramid）與滑動窗格（sliding window）等技術來實作。

執行完演算法後，透過 detector.run() 函式得到一個分數，此分數代表偵測到的圖形是否與人臉接近，若分數越高，代表愈接近人臉，分數越低則代表誤判，而 detector 回傳的人臉資訊為 list 的結構，如[(76, 119)(461, 504)]，分別代表著左上方的座標和右下方的座標，將偵測出的人臉位置儲存下來，裁切出人臉的圖片如圖四，並存成新的圖片，以供人臉訓練的模型使用。



圖四 裁切出人臉的圖片

3.1.3 裁切好的人臉照片做角度調整

由於拍攝到的人臉照片不一定皆為正臉，如此一來，會降低人臉辨識的精準度，所以在執行人臉辨識前，我們需先將偵測出的人臉照片，調整到正臉的角度，以改善此狀況。而進行調整之前，需要先擁有人臉特徵點(facial landmarks)作為校準的依據，此步驟使用面部特徵點估算演算法，以下介紹此種方法：



圖五 68 個特徵點

面部特徵點估算 (face landmark estimation) 演算法，在人臉上找到普遍存在的特定點(landmarks)--包含眼睛、鼻子、嘴巴及臉的輪廓，共 68 個特徵點如圖五。因此，我們使用 Dlib 提供的面部特徵模型，使用 Dlib 的 dlib.shape_predictor 取得每個人臉的 landmark，以得到人臉照片中 68 個特定點。面部特徵點估算 (face landmark estimation) 成果如圖下：



圖六 畫出人臉照片的 68 個特徵值

在取得 68 個特徵點後，就能依據此進行臉部校準(face alignment)，我們使用 Adrian Rosebrock 的 alignment 方法，其做法為：

1. 移動臉部到相片中點
2. 旋轉臉部讓雙眼位置在相同的水平線上
3. 調整臉部至適合的大小

由於 Adrian Rosebrock 有將 align 方法寫成 module 放到開源 imutils 套件中，因此我們 import imutils，將人臉資訊，輸入到 FaceAligner 得到校準後的臉部圖片，結果如下表示：



圖七 校準人臉的示意圖

3.1.4 在新增使用者時，只上傳一張照片就能讓神經網絡去學習辨識

辨識人臉這一步驟，由於考量到辨識人臉所需要存入資料庫的照片不少，因此為了能有效率的方式，在龐大的資料中找到與照片相符的人臉，所以我們在第二節所提及

解決方式，就是使用 GOOGLE 研發的人臉辨識系統，facenet，該系統為透過訓練深度卷積神經網路，從人臉圖像中，embedding 成 128 維的向量如下圖：

$$f(\text{face image}) = \begin{pmatrix} 0.112 \\ 0.067 \\ 0.091 \\ 0.129 \\ 0.002 \\ 0.012 \\ 0.175 \\ \vdots \\ 0.023 \end{pmatrix} \quad \left. \vphantom{\begin{pmatrix} 0.112 \\ 0.067 \\ 0.091 \\ 0.129 \\ 0.002 \\ 0.012 \\ 0.175 \\ \vdots \\ 0.023 \end{pmatrix}} \right\} \begin{array}{l} \text{128} \\ \text{feature vector} \end{array}$$

圖八 人臉 embedding 出 128 維特徵向量

再將向量投射到歐式空間(式-1)，計算兩人之間的距離：

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} = \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}} \quad \dots \text{式-1}$$

此模型確保相同的人，在不同照片仍然得到距離接近的結果，不同的人，則會得到很遠的距離。如此一來，只要透過人臉比對規則(式-2)計算需辨識的人臉照片，再從資料庫中找到距離最接近的，就能找到其身分。

If $x^{(i)}, x^{(j)}$ are the same person, $\|f(x^{(i)}) - f(x^{(j)})\|^2$ is small.
If $x^{(i)}, x^{(j)}$ are different persons, $\|f(x^{(i)}) - f(x^{(j)})\|^2$ is large.

...式-2

因此，我們將需要訓練的照片，輸入到他們預訓練好的模型，就能得到該相片 128 維度的特徵值，再存到 python 中的 pickle 檔，因此下次再辨識時，就不需要再重新計算了。

3.1.5 若是同一時間，多人使用上傳或辨識的功能時，需同時處理大量資訊

當一個系統的使用普及範圍擴大時，運用本功能的會員隨之增加，因此，若是同一時間，出現許多使用者需要進行訓練或是辨

識時，會造成系統負擔加重，為了解決此問題，本系統選擇以非同步程序

(Asynchronous Programming) 去設計，它是並行處理的方式為主，各自程序分開執行，當呼叫非同步方法後，執行完會回傳呼叫端，因此不需要等待這個函式執行完就能繼續執行下一行敘述。

3.2 明星臉偵測系統

家中或公司不免會有訪客或貴賓前來拜訪，因此我們想提供一個趣味性的迎賓遊戲，在辨識人臉的同時，偵測使用者與哪位明星長的最為相似，並將結果呈現在 KIOSK 上。

3.2.1 以爬蟲獲取照片，篩選明星照片

先從網站上，使用 BeautifulSoup 將明星的名字讀取下來，存取完畢後，再依據每個名字爬取其照片下來，目前明星的蒐集資料為 300 位。

然而，有些照片因為拍攝關係，會產生兩種狀況：(1)用裝飾品遮住臉的一部分，或是拍攝角度為背面 (2)照片出現兩位以上的人物

為了解決此問題，我們採取的篩選機制

3.2.2 明星照片篩選

將照片套入到人臉偵測模型，並計算偵測出的人臉數量。

- 若數量等於 1，即為我們所要的照片，將其存下來，
- 若偵測出的數量大於 2 以上，代表照片中出現的人物不只目標的明星本人，還有其他人，所以是不行的。
- 若偵測出的數量為 0，代表遮蓋面積太大，例如只剩下兩隻眼睛，或

是轉到背面，因此無法偵測出人臉位置，所需刪除原本的照片。

3.2.3 套入明星臉偵測模型

將篩選好的照片，套入到人臉校準、特徵點提取模型中做人臉訓練，並將這些資料分別存成 pickle 檔，因此要做明星臉比對時，讀取該 pickle 檔的資料即可。接下來，若有使用者在 KIOSK 前拍攝照片時，將讀取的照片進行人臉偵測、人臉校準、人臉特徵點提取後，再與資料庫中的特徵點向量做計算，若距離越近的，代表他們的特徵值相似的程度越大，因此選取最接近距離的明星照片，作為結果。

3.3 伺服器架設

伺服器底層以 Window 為作業系統，使用 asp.net mvc 架構製作：

controller 的部分負責扮演整個系統中資料的分析者，透過 4G 網路、http 與手機、KIOSK 傳送資料，以及使用 socket 連結樹莓派，提供三種 AI 模型放在伺服器運算：

- Face_training：接收來自手機上傳的照片，套入人臉訓練模型後，將結果存到資料庫中。
- Face_recognition：接收來自 KIOSK 的攝影機所傳送的資料，套入人臉辨識模型後，將結果回傳給 KIOSK 和樹莓派。
- Star_detect：接收來自 KIOSK 的攝影機所傳送的資料，套入明星臉偵測模型後，將結果回傳給 KIOSK。

model 的部分使用開源資料庫 Microsoft SQL server 管理系統資料，儲存會員資料：

- User_Info：儲存使用者資料包含姓名、電話、email、帳號、密碼。

在專案開發完成後，本系統將其上架到 IIS 上，讓 client 端可以透過網路向 server access API。

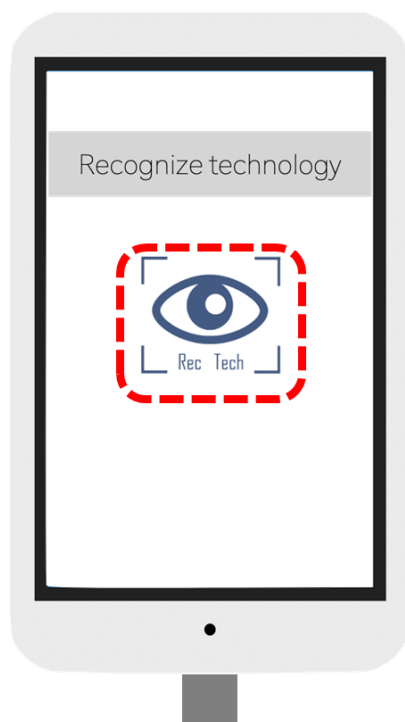
4. 主要成果

4.1 互動式資訊服務站 (KIOSK)

本系統使用的是 POSIFLEX TK-3200 KIOSK，利用官方提供的 Win10 版本，可以透過 4G 網路完成 KIOSK 與伺服器的連線、傳輸資料。以下分別介紹 KIOSK 的首頁拍照、人臉辨識和明星臉偵測成果頁面和分析。

4.1.1 首頁拍照

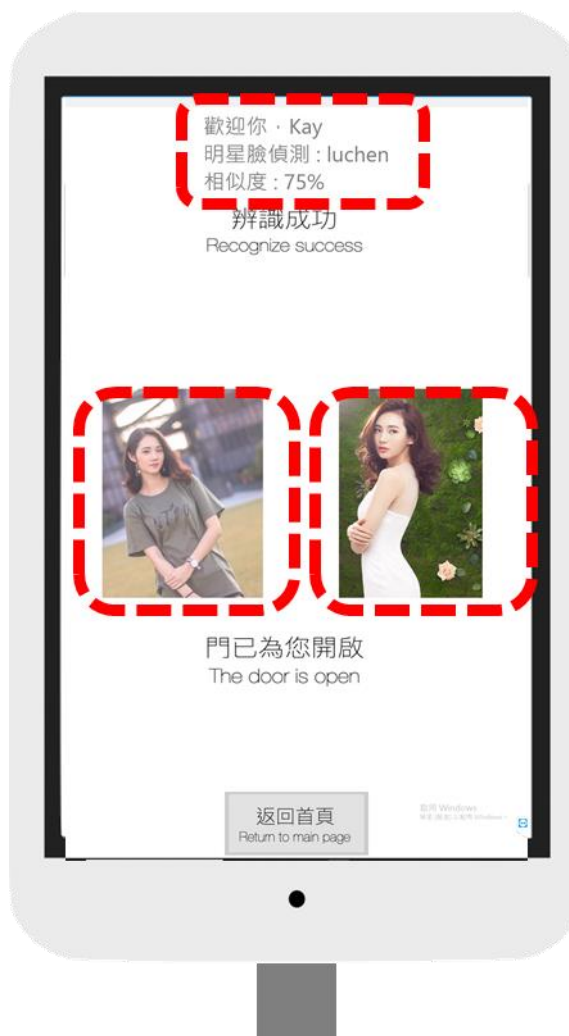
- (1) 使用者按下圖九中間的 logo 圖片，進行拍照，程式碼接受到指令會呼叫 KIOSK 上的攝影機，拍下須辨識的使用者的照片，並利用 4G 網路將資料傳送到伺服器，以進行 AI 人臉辨識的運算。



圖九 KIOSK 首頁

4.1.2 人臉辨識和明星臉偵測成果頁面

- (2) 接收伺服器傳送的辨識成果，若成功，會呈現圖十中使用者和相似明星臉的姓名、照片，以及明星臉偵測的相似度，若失敗，則呈現圖十一，並請使用者按下返回首頁按鈕，以重新辨識。



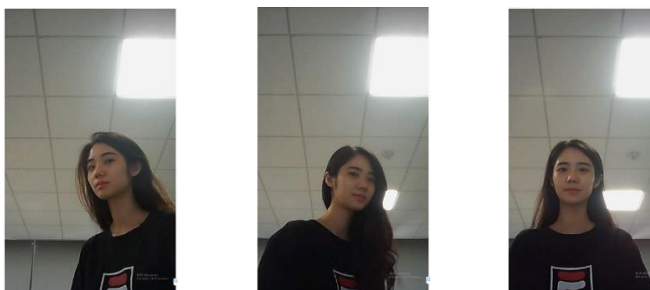
圖十 人臉辨識成功畫面



圖十一 人臉辨識失敗畫面

4.1.3 分析

如圖十二所示，無論使用者拍攝角度為正臉、側臉、或是微上揚的角度，本系統皆能順利辨識出是同一個人。



圖十二 各個拍攝角度

4.2 手機 App (Mobile Application)

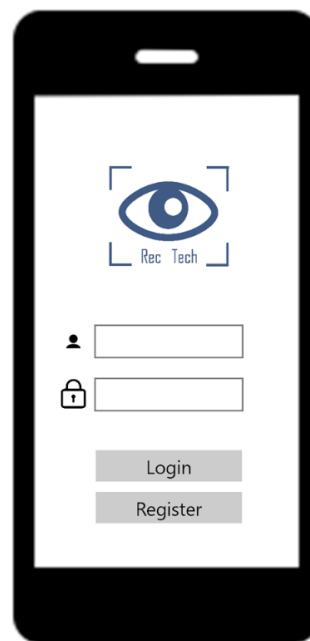
手機應用程式使用 Microsoft 的 Xamarin 撰寫，使用 .NET 建立適用於 iOS、Android 應用程式。以下分別介紹 KIOSK 的註冊/登入、照片上傳頁面。

4.2.1 註冊/登入

新會員需要透過註冊圖十三，將基本資料輸入到系統中，以擁有會員資格。成為本系統會員後，登入圖十四畫面使用本系統提供的服務。



圖十三 註冊會員



圖十四 登入會員

4.2.2 上傳照片

為了讓人臉訓練模型認識使用者的長相，因此需要使用者在圖十五畫面先上傳自己的照片，由於上述提及本系統使用 one

shot learning 模型，所以使用者只要上傳一張照片即可。

- 左下方的按鈕是從相簿中挑選照片
- 中間的按鈕是開啟相機拍攝
- 選擇的照片呈現在中間方形畫面中。
- 若選擇好滿意的照片，就按下右下角的按鈕，app 即會透過網路將照片上傳到 server 進行人臉訓練。



圖十五 上傳照片

4.3 樹莓派和電磁鎖

本系統使用樹莓派接收人臉辨識的結果，若成功，則會由樹莓派控制電磁鎖打開。以下分別介紹樹莓派開發板模組、電源模組、電磁感應門鎖磁鐵等相關硬體設備，其各自功能。

- 主控模組：主要是集中分配的功能，需要提供測驅動晶片等工作信號。



圖十六 樹莓派模組

- KIPO-門禁開關電源供應模組：主要負責提供符合要求的電流與電壓，由於樹莓派使用 5V，然而電磁鎖需要 12V，所以此模組將輸入 12V 電壓開門，專用於直接輸出電壓控制電鎖的門禁控制器。



圖十七 KIPO-門禁開關電源供應模組

- 電磁感應門鎖磁鐵：主要負責開關門的動作，本系統所使用的電磁閥為鎖扣式的，利用通 12V 的電壓把鎖頭中心的軸利用電磁鐵相吸，若沒有通電時，鎖頭會呈現突起的狀態，以勾住孔洞把門卡住。



圖十八 電磁感應門鎖磁鐵

5. 結語與展望

本專題研究成功結合 kiosk 機台、手機 APP、樹莓派、電子門鎖與伺服器，透過物聯網技術建構出一個可以透過『人臉』做人員管理、開門的系統。藉由這項研究，使我們的日常生活更為便利，期望未來這項系統可以更為普及，讓使用者體會如何能快速又方便的開門、管理者也更簡單做人員的控制。

6. 銘謝

感謝指導教授從我們大二下就開始給我們銜接大三專題的方向，給我們許多學習空間與資源去利用，並在每次的會議上帶領我們給予指導。也感謝實驗室的學長姐給予建議，使我們的專題如虎添翼。最後感謝組員與協助我們的同學，雖然常常要待在實驗室待到很晚，但最後看到我們成功做出了這個專題結果真的很開心！

7. 參考文獻

- [1] Biggs, P., L. Srivastava, and I.T. Union, "ITU Internet Reports: The Internet of Things: International Telecommunication Union", 2005.
- [2] Ashton, K., "That 'internet of things' thing. RfID Journal", vol 22, p. 97-114. 2009.
- [3] Interactive kiosk - Wikipedia (https://en.wikipedia.org/wiki/Interactive_kiosk)
- [4] 石宗哲，2004，〈論未來 KIOSK 之發展與設計-以戶外公共資訊站為例〉，台灣師範大學設計研究所碩士論文
- [5] "What is Facial Recognition? - Definition from Techopedia". Techopedia.com. Retrieved 2018-08-27.
- [6] ahid Kazemi and Josephine Sullivan. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1867 - 1874, 2014.
- [7] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 815 - 823, 2015.