

## 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、毫米波雷達感測器。

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，以及毫米波雷達感測進行活體辨識的人臉檢測，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。

**關鍵詞：**kiosk、人臉識別、深度學習、智慧門鎖、毫米波雷達。

## Abstract

Since mobile devices has been getting more and more popular, there are lots of developments on IoT ( Internet of Things)[1] in many fields in recent years. With IoT device, it can not only facilitate automation of the daily life of human kinds, it can also improve the quality of our life.

We observe that home safety has always been an important concern of people, if we can quickly detect people's identity with the feature of every individual's face, not only can we satisfy the need for security, but we can also make our daily life become more convenient. What's more, for those whose house doesn't have security system, this system can also avoid expensive cost and extra construction to reach the demands of the security of smart home.

Taking this opportunity, we had made a system which combines home safety and AI detection and recognition on facial feature. This system integrates app on mobile phone, smart door lock system controlled by raspberry pi, kiosk machine for exhibition and taking pictures, and Millimeter Wave (mmWave) radar sensors.

We can upload photos, add and store user information with the app on mobile phone, take picture through the camera in front of kiosk. Taking advantage of the AI detection model to recognize whether the person who tries to open the door has registered and the mmWave radar sensing on liveness detection of face detection as an user, so the system could decide whether it should open the door and provide a game - AI celebrity face detection as welcoming to the guests.

We hope that we can combine the electric door lock, kiosk, and smart phone, to achieve the demand of " smart city".

**Keywords :** Kiosk, Face Recognition, Deep Learning, Smart Door Lock, Millimeter Wave (mmWave) radar sensors

# 目錄

摘要 .....	I
ABSTRACT .....	II
目錄 .....	III
壹、 緒論.....	1
一、 研究動機 .....	1
二、 研究內容 .....	1
貳、 相關研究 .....	4
一、 KIOSK 機台 .....	4
二、 GOOGLE FACENET .....	5
三、 毫米波雷達 .....	5
參、 AI 健康管理系統實作 .....	7
一、 門禁管理系統 .....	8
二、 伺服器架設 .....	10
三、 人臉辨識系統 .....	11
四、 明星臉偵測系統 .....	15
五、 活體辨識偵測 .....	17
肆、 門禁管理系統實驗結果及評析 .....	20
一、 系統開發工具 .....	20
二、 系統實驗結果 .....	20
伍、 結論與未來研究方向 .....	24
參考文獻.....	25

## 圖目錄

圖 1-1 系統應用情境說明 .....	3
圖 2-1 系統應用情境說明 .....	7
圖 2-2 FMCW 雷達波模 .....	7
圖 3-1 系統架構圖 .....	7
圖 3-2 資料在系統中的傳遞與各設備間的通訊方法 .....	8
圖 3-3 尚無使用者再 KIOSK 前 .....	9
圖 3-4 人正在靠近門禁系統 .....	10
圖 3-5 人正遠離門禁系統 .....	10
圖 3-6 伺服器環境說明 .....	11
圖 3-7 人臉辨識步驟 .....	12
圖 3-8 裁切出人臉的圖片 .....	12
圖 3-9 人臉 68 個特徵點 .....	13
圖 3-10 畫出人臉照片的 68 個特徵值 .....	13
圖 3-11 校準人臉的示意圖 .....	13
圖 3-12 人臉 embedding 出 128 維特徵向量 .....	14
圖 3-13 Synchronization Server 控制 UI 介面 .....	15
圖 3-14 明星照片篩選機制 .....	16

圖 3-16 臉部 4 個位置的座標示意圖 .....	17
圖 3-17 十字滑軌 .....	18
圖 3-18 Arduino CNC Shield v3 擴展板 .....	18
圖 3-19 臉部 4 個位置被平面物體擋住 .....	19
圖 3-20 臉部 4 個位置沒被擋住 .....	19
圖 4-1 註冊會員 .....	21
圖 4-2 登入會員 .....	21
圖 4-3 上傳照片 .....	21
圖 4-4 人臉辨識成功畫面 .....	22
圖 4-5 人臉辨識失敗畫面 .....	22
圖 4-6 各個拍攝角度 .....	22
圖 4-7 樹莓派模組 .....	23
圖 4-8 KIPO-門禁開關電源供應模組.....	23
圖 4-9 電磁感應門鎖磁鐵 .....	23

## 壹、 緒論

在正式進入專題介紹前，我們先大致介紹本系統的研究動機與研究內容，其中包含人臉辨識系統、明星臉偵測系統、KIOSK 機台和 app 的使用者介面設計、樹莓派控制的電磁鎖、搭配十字滑軌的毫米波雷達。接下來將一一來做介紹。

### 一、 研究動機

由於近年來的趨勢趨向『萬物皆聯網』，再加上看見近期最火紅的『A I 人臉辨識』，在發想專題主題的時候我們想到將物聯網結合 A I 應用-人臉辨識智慧門鎖，解決日常中常常困擾我們的小事情-忘記帶鑰匙出門、每次開門都需要記得鎖門、還有我們總是需要花時間在包包裡找鑰匙。

另外，kiosk 機台在日常生活中其實非常常見，但我們都不知道他已經悄悄融入我們的生活中了-速食店的自動點餐機與大賣場裡的自動結帳省去店家雇用店員的人本費用、吸睛的廣告看板，還有停車場的自動繳費機、台灣城市街道上的 U Bike 借還站...等。KIOSK 跟我們的生活變得越來越密切相關，所以我們配合用來展示使用者介面的機台-KIOSK，結合智慧門鎖。

『AI 始終來自於人性』，在這個系統上又加入了迎賓小遊戲，讓使用者擁有更好的體驗，在人機互動中更顯智慧城市與我們市民的親近，也讓單純的開門鎖、大門進出人員管理這件事變得更為有趣。

總合以上，本專題計畫設計了一套系統，結合了手機應用程式、利用樹莓派控制的門鎖和用來呈現的 KIOSK 介面三個主要層面，讓我們可以快速的通行、又能夠實現居家安全與生活樂趣。

### 二、 研究內容

本系統整合 KIOSK、深度學習、手機 app 與樹莓派，使門禁管理系統具有更完整的服務，其中將分為六個部分做簡單的介紹：

#### 1. Kiosk 機台

使用者介面：在 kiosk 機台上主要有兩部份畫面需要顯示：一開始的首頁拍照跟完成人臉辨識後的歡迎使用者資訊和明星臉偵測成果頁面。

## **2. 人臉辨識模型**

我們需要做到拍攝一張照片就可以正確辨識出此人身份，此 AI 模型要能偵測出人臉位置、提取人臉特徵，並計算相似度。

## **3. 伺服器架設**

伺服器負責扮演整個系統中資料的分析者。server 透過 4G 網路、http 與手機、KIOSK 傳送資料，以及使用 socket 連結樹莓派。並且提供三種 AI 模型放在伺服器運算，其中為人臉辨識 (Face\_recognition)、人臉訓練 (Face\_training)、與明星臉偵測 (Star\_detect)。另外伺服器上用來儲存會員資料的 Database 我們選用 Microsoft SQL server 管理系統資料。

## **4. 手機 app 端**

隨著智慧型手機的普及，透過手機裝置上 app 註冊會員，並上傳照片至 server 做訓練。

## **5. 樹莓派和電磁鎖**

使用樹莓派接收人臉辨識的結果，若辨識成功，會由樹莓派控制電磁鎖打開。本專題為了控制門鎖所用到的主要硬體設備有：主控模組-用以處理工作信號、KIPO-門禁開關電源供應模組-用以輸出電壓控制電鎖與電磁感應門鎖磁鐵-在通電時可以勾住門鎖，利用三項硬體與 socket 跟 server 做的連結，實現智慧門鎖的需求。

## **6. 毫米波雷達感測**

使用毫米波雷達感測判斷 KIOSK 前是否有使用者需要使用，若偵測到則通知伺服器進行門禁管理系統的服務，並且搭配十字滑軌偵測人臉等 4 個位置，以判別前來偵測的人沒有使用照片、手機上的影像、影片等蒙騙行為，達到活體辨識的目的。

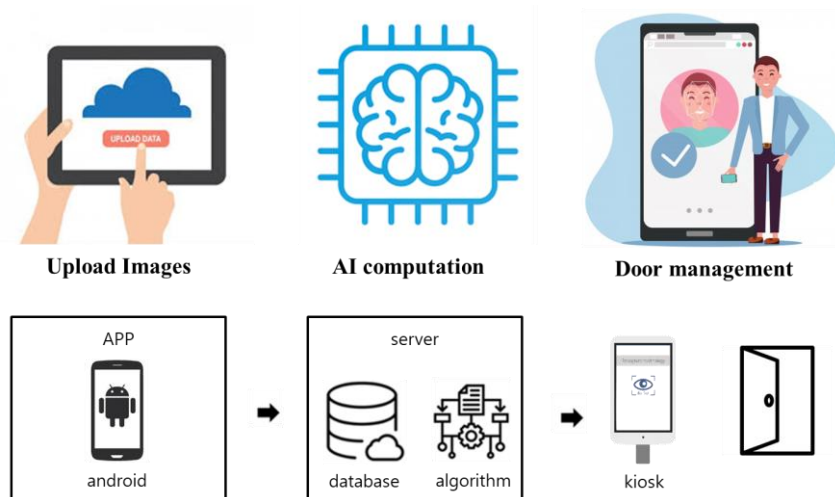


圖 1-1 系統應用情境說明

上圖 1-1 說明本系統應用的情境，使用者在手機 app 上登入會員，並上傳一張自己的照片，手機在連線 server 做新會員的資料的人臉訓練，並將訓練成功的資訊存到資料庫中，等到使用者至 KIOSK 前進行人臉辨識，毫米波雷達感測會偵測到人，並主動通知伺服器，伺服器會傳送指令讓 KIOSK 拍下使用者當前的照片，上傳至伺服器進行人臉辨識，同時使用毫米波進行活體辨識，若辨識結果成功，則連線樹莓派控制電磁鎖開啟，並將明星臉偵測出的結果呈現在 KIOSK 上，供使用者欣賞。



## 貳、 相關研究

### 一、 Kiosk 機台

早期的 Kiosk 通常被拿來當作較小型的銷售亭或零售攤，Kiosk 外型設計都與一般實體銷售亭無異，只是加上了螢幕，播放產品廣告或資訊性的短片而已。

到了現在，隨著科技日新月異，觸控螢幕開始普遍被使用。因此，早期電子化的電子銷售亭現在已變成擁有大面積觸控螢幕，提供人機互動的直立機台。

Kiosk 的用途有：速食店點餐及付款的銷售顯示屏、購物商場內的多媒體資訊機、自助設施預約系統、自動販賣機等等[2]。在科技化的現代，他逐漸省去了人力、可以取代重複性高的工作。

由亞馬遜推出無人商店「Amazon Go」[3]，在 2018 年 1 月 22 日於美國西雅圖正式開幕。Amazon Go 結合了人工智慧、機器學習、雲端運算，利用大量感測器、人工智慧攝錄系統，顧客先下載店家的 app，從顧客進入店內開始，顧客的每個動作移動路線、瀏覽或拿取、放回商品、檢視標籤這些動作都會透過深度學習演算法替顧客置入 app 內的虛擬購物車。最後再透過虛擬購物車，直接結合會員帳戶從中扣款，免去了結帳這個動作。

2018 年，在台北市信義區，開啟了台灣第一間無人商店[4]。在入店以前先以人臉識別身份，結帳時將商品放上 pos 機台上，點選確認結帳再經過人臉識別完成結帳。

「善用新科技，不是為了取代人力，而是釋放人力去做更具價值的事情，簡化他們繁瑣的例行公事。」透過這些例子可見，kiosk 越來越融入生活是人類必定會走向的未來，因此我們決定根據現在所見趨勢，結合人臉識別將應用套用在 kiosk 上。

## 二、 Google Facenet

在過去我們使用傳統的 LBP、HOG 搭配支援向量機 SVM 進行人臉辨識，但隨著近年來深度學習、卷積神經網路 CNN(Convolutional Neural Network)的興起，使用複雜的深度學習模型不但辨識率非常高、透過單樣本學習(one shot learning)，能夠識別的人臉數量數以萬計。

CNN 在影像識別上固然可以說是最佳的方法之一，但考量到若需要將每個人的照片及身分作為標籤，透過 CNN 及 softmax 函數訓練參數，會遇到的問題有：訓練樣本不足-在影像識別中，我們用大量照片以增加樣本數去訓練模型，但往往在人臉識別的場景我們需要一張照片就可以辨識出身份;擴充性差-要新增辨識的人員時，會改變 CNN 輸出矩陣大小，因此需要重新 label，這是相當費力且不實際的;拍攝角度不盡相同-側臉和正臉的五官位置便會產生差異，因此在進行人臉辨識前，需要先進行人臉校準[5]。

因此 facenet[6]他的架構包含了卷積層、池化層及全連接層，與 CNN 架構相似，差異在於輸出的向量是代表該影像的向量，而非代表影像分類的向量，架構中在預訓練模型 (pretrained model) 跟人臉識別 (recognition) 時皆採用孿生神經網路，兩個 convet 網路擁有相同的參數與權重，經由 Convolution(卷積)、Pooling(池化)、Fully connected layers(全連接層)後得到 128 維的特徵向量(feature vector)，這個過程稱之為 encoding(編碼)。而孿生神經網路就是希望學習出一種最準確的 encoding 方法。

透過三重損失函數(Triple loss function)[7]對代表兩個影像的向量相減取範數，得到兩張影像的差距，進而判斷是否為同一個人。

## 三、 毫米波雷達

毫米波雷達，是一種在毫米波波段 (millimeter wave) 的電磁波雷達，藉由線性調頻連續波 (Frequency Modulated Continuous Wave, FMCW) 雷達技術，其頻率隨時間呈線性變化，透過比對發送和接收之間的頻率隨時間而改變的信號變化，產生都卜勒效應[7]，如圖 2-1 所示，利用所偵測的物體距離、速度以及角度等獲得時間上 (Temporal) 和空間上 (Spatial) 的

資訊，因此能夠提供動作辨識中更多資訊的來源依據。

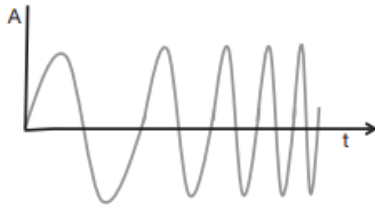


圖 2-1 系統應用情境說明

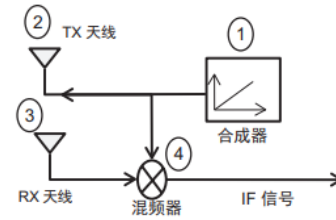


圖 2-2 FMCW 雷達波模

雷達波系統包含發送(Tx)、接收(Rx)、射頻(RF)等模擬模組，還有訊號處理的模樹轉換器(ADC)、微控制器(MCU)和數字信號處理器(DSP)等模組，如圖 2-2 所示。[8]

對於 mmWave 應用上，過去大多專注於自動車產業[9]，最近則是在較小範圍的行為辨識上，例如肢體的辨識[10]、手勢的辨識[11]-[15]、臉部的辨識。目前已有需多廠商生產毫米波感測器，例如 Texas Instruments 的毫米波感測器[16]，以及 Google ATAP's Project Soli [17]，使得毫米波感測的使用上難度下降許多，也使得其應用變得更加多元。

## 參、 AI 健康管理系統實作

本系統欲藉由跨平台的串聯，將打造出完善的門禁系統給使用者。下圖 3-1 為系統主要開發的四大部分：(一) 門禁管理系統：使用者使用手機 app 加入系統會員後，在進門前到 KIOSK 拍攝照片進行人體辨識以及活體辨識的判斷，若辨識成功，由樹莓派控制電磁鎖開啟，並將最終的辨識成果顯示在 KIOSK。(二) 伺服器架設：存入系統會員的資料，提供人臉特徵訓練、辨識的運算、明星臉偵測、樹莓派控制等 API 指令。(三) 人臉辨識系統：採用 google 的 facenet 模型，達成只要上傳一張照片便能辨識的 one shot learning 技術。(四) 明星臉偵測系統：使用者在辨識人臉時，同時依據當下拍的照片，尋找出長相最為相似的明星資訊，因此每次辨識時拍攝皆會有不同的成果。(五) 活體偵測系統：使用毫米波雷達偵測搭配十字滑軌，對使用者的臉進行感測，利用 FFT 分析距離，進而判斷是否為立體的，而不是用平面的照片作為冒充。接下來，將針對以上五步驟進行詳細的說明：

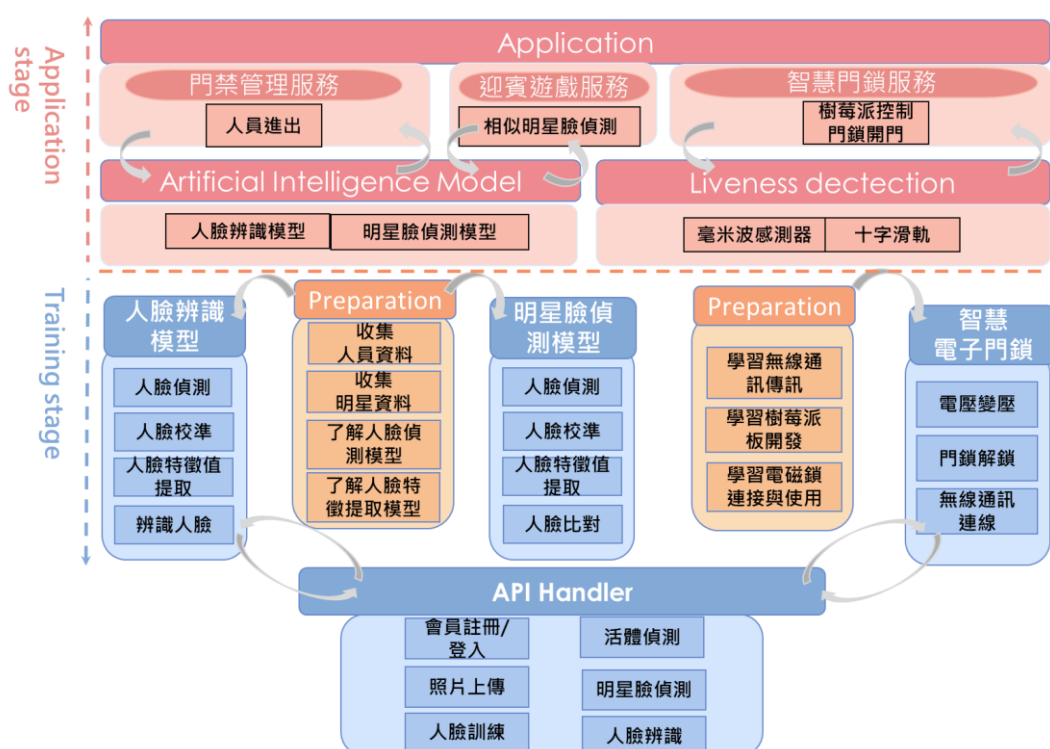


圖 3-1 系統架構圖

## 一、門禁管理系統

### (1) 資料傳輸架構

本系統利用無線網路的傳輸技術做開發，分別為 4G 無線網路。手機端上註冊會員以及上傳照片，再透過 4G 網路將照片傳送到伺服器，套入人臉訓練的模型，將結果存入資料庫，接著，KIOSK 會抓取使用者的拍攝照片，透過 4G 網路將照片傳送到伺服器進行處理，套入人臉辨識、明星臉偵測模型，再取到結果後透過 4G 網路分別回傳送到 KIOSK 端做呈現、樹莓派控制電子鎖解開使門開啟。

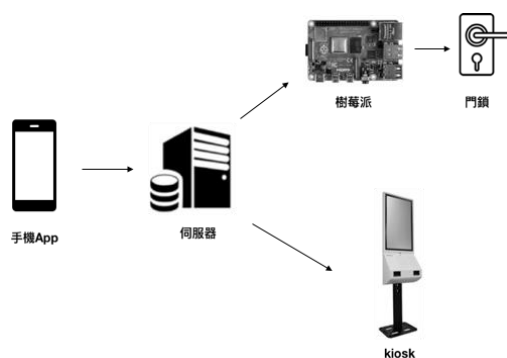


圖 3-2 資料在系統中的傳遞與各設備間的通訊方法

### (2) 訪客管理

本系統能夠預先登錄來訪的貴賓，只要使用者在手機 App 上傳貴賓的照片、姓名即可，因此當貴賓到現場時，可以直接走到 KIOSK 前進行辨識，而且辨識成功後，KIOSK 會呈現迎賓的訊息以及迎賓小遊戲。

### (3) 使用毫米波雷達判別門前是否有人

由於一直開啟著相機拍攝，再傳到伺服器進行人臉辨識，在長時間閒置的情況下，效率是不佳，而且也很浪費電腦運算的功耗。原本，我們一開始規劃使用者需要到 KIOSK 前，按下按鈕後，再開起相機拍攝，然而，在尚未按下按鈕前，這段時間是閒置著，為了將人臉偵測這個動作從被動化為主動，優化使用者體驗的流程，我們選擇使用毫米波雷達進行偵測是否有人正走向 KIOSK。

有鑑於毫米波雷達使用 FMCW 技術，其發射的頻率會隨時間線性遞增，能夠從接收的信號計算成距離多普勒雷達圖（Range Doppler Map, RDM）。在 RDM 圖形的表示上，本系統將

Y 軸定義為發射偵測距離(Distance)，X 軸可透過多普勒偏移 (Doppler Shift) 運用數學方式換算為有向性的速度 (Velocity)，透過此特徵，將偵測到的狀況分為兩種:

#### I. KIOSK 前沒有人

毫米波雷達無感測到人，RDM 圖形上會形成如圖 3-3 的狀況，偵測到的物體會靜止不動的，由此可以知道目前狀況尚無使用者在 KIOSK 前，因此傳送"NOBODY"訊息到伺服器，伺服器接受到後就繼續保持無動作的狀態。

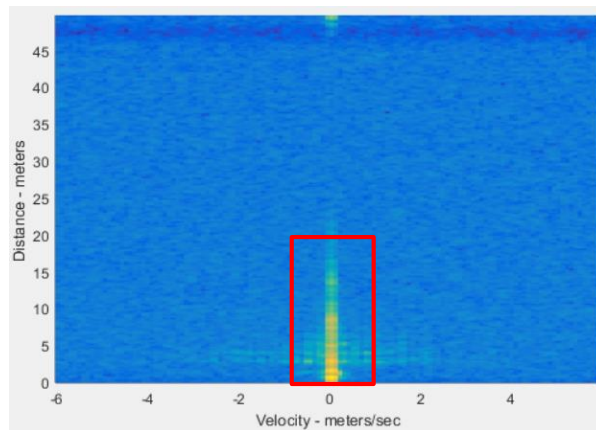


圖 3-3 尚無使用者在 KIOSK 前

#### II. KIOSK 前有人

毫米波雷達感測到人時，這時 RDM 圖形上會有偵測到物體移動的速度，由於速度是有向性的特徵，所以可以判斷此物體是靠近或是遠離毫米波雷達，若物體為遠離毫米波雷達時，代表物體的和雷達波的距離變長，速度呈現正的狀況，因此計算訊號中速度為正和負的點數量，若速度為正的較多時，如圖 3-4 所示，因此可以判斷使用者沒有要進入門禁系統，則傳送"NOBODY"訊息到伺服器，以保持無動作的狀態。若是偵測到有使用者要進行人臉辨識，其距離會與雷達波變短，速度呈現負的狀況，計算出速度為負的點較多時，如圖 3-5 所示，因此傳送"PEOPLE DETECTED"訊息到伺服器，伺服器接受到傳送開啟相機指令至 KIOSK，開啟首頁的狀態，等待使用者站在攝影機前拍攝照片，並傳送照片至伺服器進行人臉辨識。

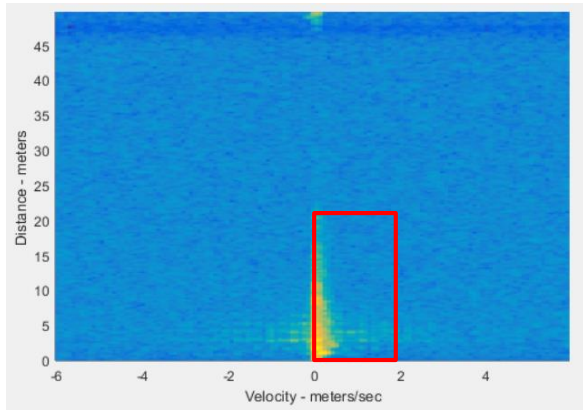


圖 3-4 人正在靠近門禁系統

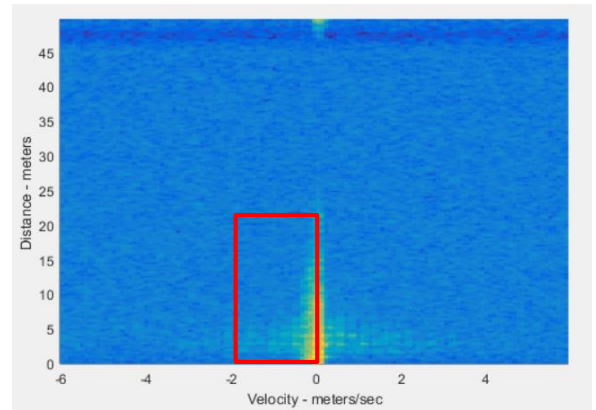


圖 3-5 人正遠離門禁系統

## 二、 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，保留 asp.net 快速開發、層級明確的優點，並且改善頁面加載的缺點，提升頁面加載的效率。

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 協定與手機、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，將結果存至資料庫中。
- 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。

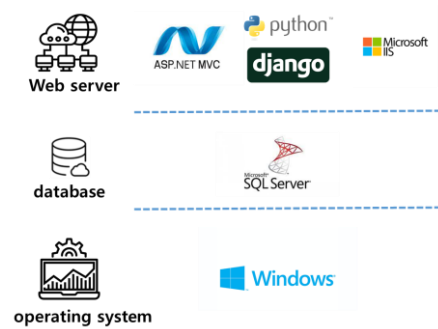


圖 3-6 伺服器環境說明

另外，人臉辨識和明星臉偵測等程式碼由 python 程式碼寫成，由於 asp.net mvc 架構是由 C#打造成的，為了提升跨程式語言的效能，使用 django 架構做為兩者之間的橋樑。

此外，有鑑於資訊安全的考量上，門禁管理系統最注重的是資料外洩的風險，避免資料在傳輸時受到有心人士的擷取，因此採用 Token 驗證系統，在系統中，在 Token 訊息後加上專屬的亂數值，接著透過雜湊函數得到固定長度的雜湊值，以驗證資料的完整性，確保資料沒有被他人竄改過，因此，當系統接收 Token 和其搭配的雜湊值時，就能依據這些資料驗證 Token 的來源。本系統的 Token 設計技術運用 RSA 加密技術，以及 SHA-256 雜湊函數演算法完成。

### 三、 人臉辨識系統

本系統將利用使用者上傳的照片資料，作為門禁辨識系統的偵測，因此，必須有合適的流程做出正確的判斷，在這判別的過程中分為:拍攝照片、人臉偵測、人臉校準、擷取 128 為特徵向量、辨識人臉五個步驟，以下將針對這五個步驟做說明:



圖 3-7 人臉辨識步驟

#### 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。

## 2. 將讀取的照片運用 dlib 的人臉偵測演算法，偵測出人臉

在分辨人臉之前，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 3-8，並存成新的圖片，以供人臉訓練的模型使用。

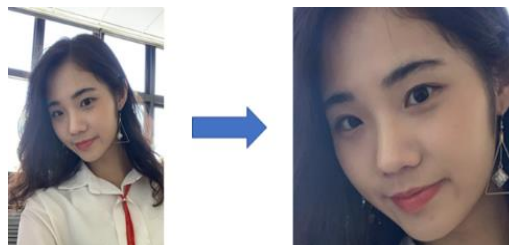


圖 3-8 裁切出人臉的圖片

## 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：

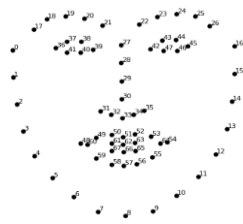


圖 3-9 68 個特徵點

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 3-9。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 3-10 下：



圖 3-10 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

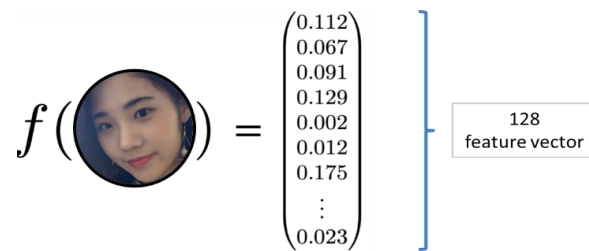
由於 Adrian Rosebrock 有將 align 方法寫成 module 放到開源 imutils 套件中，因此我們 `import imutils`，將人臉資訊，輸入到 `FaceAligner` 得到校準後的臉部圖片，結果如下表示：



圖 3-11 校準人臉的示意圖

#### 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 3-12:



The diagram illustrates the process of converting a face image into a 128-dimensional feature vector. On the left, a circular face image is shown. An arrow points from the image to a large right square bracket. To the left of the bracket is the function notation  $f(\text{image})$ . To the right of the bracket is a vertical list of numerical values representing the feature vector: 0.112, 0.067, 0.091, 0.129, 0.002, 0.012, 0.175, followed by a vertical ellipsis, and finally 0.023. To the right of the entire bracketed structure is a box labeled "128 feature vector".

圖 3-12 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:

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

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。

$$\begin{aligned} &\text{If } x^{(i)}, x^{(j)} \text{ are the same person, } \|f(x^{(i)}) - f(x^{(j)})\|^2 \text{ is small.} \\ &\text{If } x^{(i)}, x^{(j)} \text{ are different persons, } \|f(x^{(i)}) - f(x^{(j)})\|^2 \text{ is large.} \quad \dots \text{式-2} \end{aligned}$$

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。

## 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，本系統選擇在伺服器上使用同步化機制，在 GPU 上創設多個 Virtual Device，以 Semaphore 機制去控管在同一時間進行訓練或是辨識的數量。

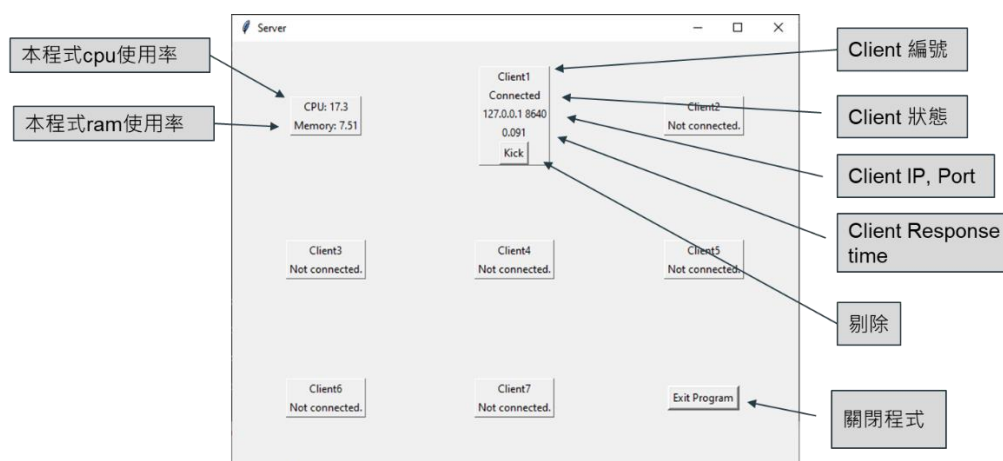


圖 3-13 Synchronization Server 控制 UI 介面

## 四、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。

### 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，下一項將做詳細說明。

## 2. 明星照片篩選

(1) 計算人臉數量：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，所以是不及格的照片，將不列入考量。
- 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，需刪除原本的照片。



Face number = 1



Face number  $\geq 2$



Face number = 0



圖 3-14 明星照片篩選機制

(2) 以性別分類明星照片

為了避免一位男性使用者偵測成女明星，或是女性偵測成男明星的窘境，因此我們會將男明星的照片和女明星的照片加以分類，存成兩個資料夾，以便之後訓練時能夠清楚了解其內容

## 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 pkl 檔，因此要做明星臉比對時，讀取該 pkl 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。

## 五、活體辨識偵測

本系統使用毫米波雷達感測進行偵測，以 TI 所開發的 IWR1443 進行偵測，如圖 3-15 所示，是一款基於 FMCW 雷達技術的單片 mmWave 感測器，其雷達的頻率為 76-81GHz。由於毫米波雷達能夠偵測出物體的距離，藉由十字滑軌的搭配感測臉部 4 個位置：左上、左下、右上、右下，分析這 4 張與感測器的距離分布，以判斷使用者的臉部是否為立體的，而不是以照片冒充其他人。

### (1) 得到使用者臉部位置

當伺服器收到使用者在 KIOSK 前所拍的照片時，運用上述人臉辨識系統的人臉偵測得出臉的長與寬，並將臉部位置分為左上、左下、右上、右下四個部分，座標資訊如圖 3-16 所示，以頭頂為原點，分別計算出臉部四個部份的座標資訊，再藉由使用者在註冊時輸入的身高，將身高減去臉部長度以得到相對的臉部位置的計算，並將此結果回傳至 Arduino 控制十字滑軌進行臉部的感測。

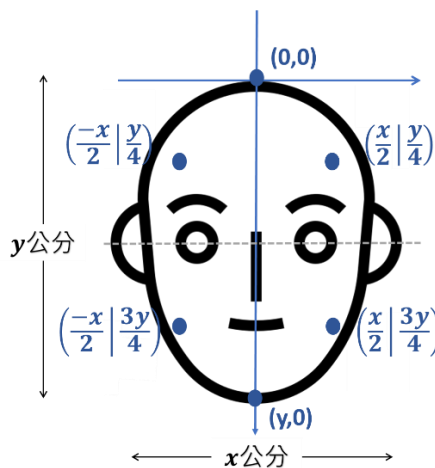


圖 3-16 臉部 4 個位置的座標示意圖

### (2) 使用十字滑軌將毫米波雷達到指定位置進行感測

為了讓毫米波雷達能夠感測臉部四個位置，因此我們使用圖 3-17 十字滑軌進行，十字滑



軌運用兩個步進馬達的轉動，長度皆為 50 公分，以圖 3-18 中 Arduino Uno 開發板與 Arduino CNC Shield v3 擴展板搭配 2 個 A4988 驅動板，以達到載物在 x、y 軸的平行上移動。因此當 Arduino 接收到臉部四個部份的位置，即能控制十字滑軌至指定的位置，由於臉部平均的長度為 20 公分，而十字滑軌的 Y 軸總長為 50 公分，因此目前設定使用者身高為 150-180 公分，其能掃描的高度為 130-180 公分。

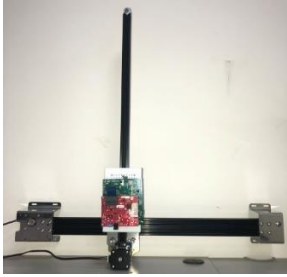


圖 3-17 十字滑軌



圖 3-18 Arduino CNC Shield v3 擴展板

### (3)分析臉部 4 個位置接收的訊號，判斷其是平面或立體

將接收到的 4 個訊號，經過距離上 FFT 的處理，以取得臉部距離感測器的距離，x 軸代表距離毫米波雷達的距離，y 軸表示接收的反射訊號強度，會隨著感測物體的材質而有不同的變化，當物體越趨向實體時，所接收的訊號強度越弱。

接著，分析波形的呈現為平穩或起伏較大，以判斷其為平面或是立體，因此我們使用動態分析策略，針對每一張距離分析圖，運算出其閾值 (threshold)，套入演算法分辨需要的目標物資訊。 $S_{\Delta} = \frac{\sum_{i=1}^n S(r, v) - \omega}{n}$ ， $S_{\Delta}$  為閾值的變化量評估，使用截頭去尾平均值 (Mid-Minimum Spacing) 作為運算方式，按照比例去掉最兩端的極值  $\omega$ ，以避免受到極值影響。

動態選填策略方式為： $threshold(r) = S(r, v = 0) - S_{\Delta}$

$$S(r, v) = \begin{cases} 0, & S(r, v) < threshold(r) \\ S(r, v), & S(r, v) \geq threshold(r) \end{cases}$$

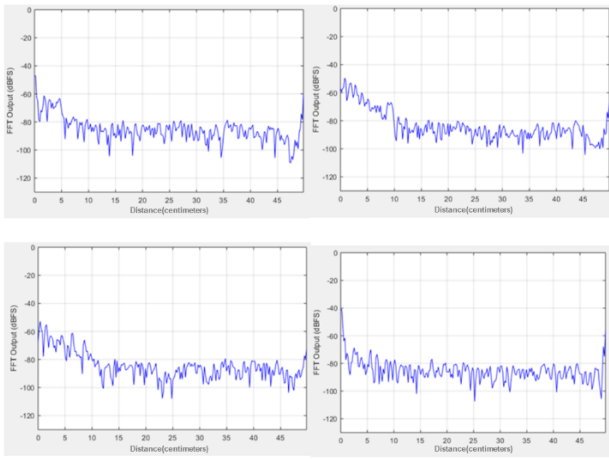


圖 3-19 臉部 4 個位置被平面物體擋住

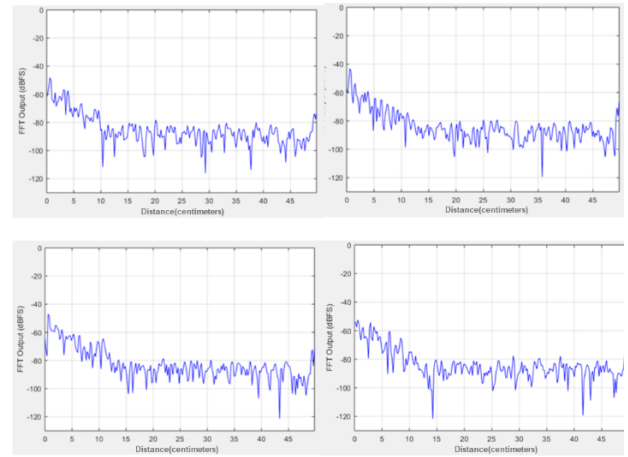


圖 3-20 臉部 4 個位置沒被擋住





# 門禁管理系統實驗結果及評析

## 一、系統開發工具

### 1. 程式語言開發工具

#### a. Python

人臉辨識和明星臉偵測系統皆由 python 撰寫。利用 Python Django 框架建立 python 語言的網路伺服器架構。另外，控制電磁鎖開關的樹梅派，在 Linux 開放系統的支援下，也能運用 python 開發。

#### b. C#

使用 asp.net mvc 架構撰寫伺服器的環境。

使用 Microsoft 的 Xamarin 撰寫手機 App、KIOSK 呈現畫面，並運用 .NET 建立適用於 iOS、Android 應用程式。

## 二、系統實驗結果

### 1. 手機 APP

手機 APP 主要用於會員註冊、登入、上傳照片等應用，再透過 4G 網路傳輸到伺服器。

#### (1) 註冊/登入

新會員需要透過註冊圖 4-1，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，登入圖畫面使用本系統提供的服務。



圖 4-1 註冊會員

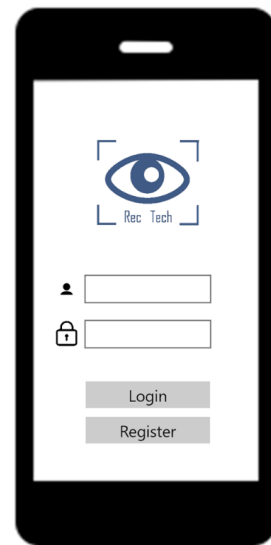


圖 4-2 登入會員

## (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 4-3 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
- 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。

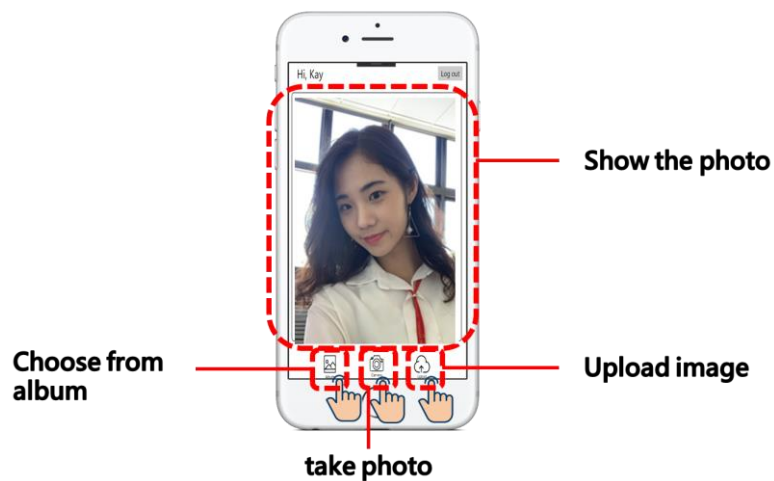


圖 4-3 上傳照片

## 2. 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。

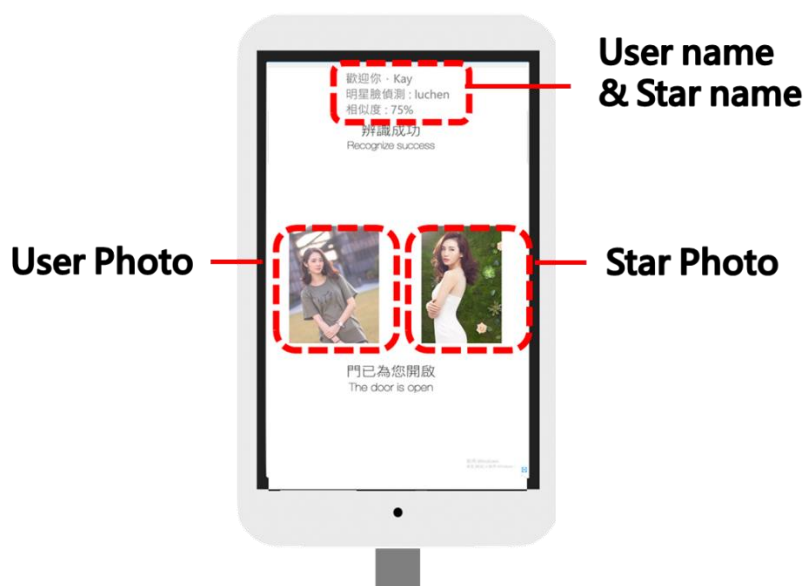


圖 4-4 人臉辨識成功畫面



圖 4-5 人臉辨識失敗畫面

### (1) 人臉辨識和明星臉偵測成果頁面

接收伺服器傳送的辨識成果，若成功，會呈現圖 4-4 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 4-5，系統將於 5 秒後回到首頁以重新進行辨識。

### (2) 分析

如圖 4-6 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。

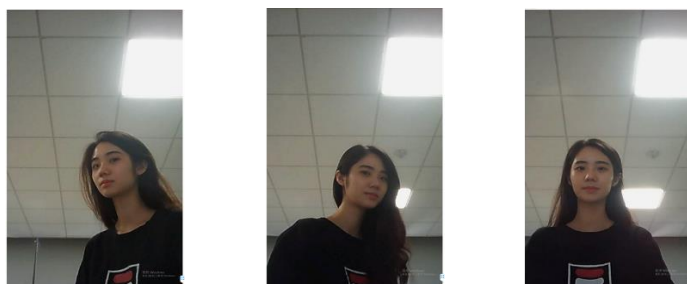


圖 4-6 各個拍攝角度

### 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。
- 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。
- 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，以勾住孔洞使門卡住無法開啟。



圖 4-7 樹莓派模組



圖 4-8 KIPO-門禁開關電源供應模組圖



4-9 電磁感應門鎖磁鐵

## 肆、 結論與未來研究方向

由於新冠肺炎(Covid-19)在全球造成大混亂，不但造成許多城市相繼實施封城禁令，同時也宣導保持社交距離，畢竟「人與人接觸」成為最高風險的行為，因此這些措施皆是希望能降低感染的機率，以避疫情的擴增，而全世界也因為疫情造成各種經濟活動的停擺，使經濟與產業皆是負面的影響，面臨前所未有的危機，甚至到今日能無法預測其結束的一天。

然而「危機就是轉機」，舉動，疫情間接改變了人類過往的生活型態，扭轉人類互動的方式，以「零接觸」為主的概念成為現今最為火紅的產業，因為許多企業與學校希望運用科技來減少人與人接觸的機會，因此開發出許多零接觸的技術，例如:電子商務、遠距應用、視訊會議...等。

本系統的門禁管理系統，原本研究的初衷為優化門禁管理的服務，也因為剛好遇到「零接觸」的技術正備受關注，因此在系統的設計上，對於需展示於各個使用者面前的 KIOSK 機台格外的注重零接觸的議題，例如在 KIOSK 的設計上希望能不用到按鈕，就能進入下一個流程，所以在使用者登入辨識畫面時，不以按鈕開起辨識服務，以毫米波偵測其移動方向，作為判斷的代替，在辨識成功後，使用者離開 KIOSK 機台，也透過毫米波感測其離開的方向，將畫面自動返回至首頁以等待下一個使用者的使用。

由於本系統使用到多項硬體設備，包含樹莓派、arduino 和毫米波雷達，由於每一項產品皆需要插頭提供電力的來源，也因此電線的長度造成平台擺放的限制，亦或者需要延長線才能完成，也造成視覺上有許多電線的窘境，因此，若是未來能將這些由無線供電的方式達成，不但能夠應用在更多的環境下，也能增加公司、住家裝設的方便性，使得生活能夠更加便利，進而邁向智慧生活的發展。

## 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] Tim Chiang (2001) A Study on Taiwan KIOSK Industrial
- [3] 林國均(2017)。Amazon Go 重新定義「無人商店」。能力雜誌，2月號。
- [4] 7-ELEVEN 企業情報(2018)。台灣無人商店的想像、從 7-ELEVEN「X-STORE」開始以智慧型超商概念策劃、展現台灣零售服務業的科技創新力 ([https://www.7-11.com.tw/company/news\\_page.asp?dId=630](https://www.7-11.com.tw/company/news_page.asp?dId=630))
- [5] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees", *CVPR*, no. 2, 2014.
- [6] Florian Schroff, Dmitry Kalenichenko and James Philbin, "Facenet: A unified embedding for face recognition and clustering", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 815-823, 2015.
- [7] G. E. Smith, K. Woodbridge, and C. J. Baker, "Radar micro-Doppler signature classification using dynamic time warping," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1078–1096, Jul. 2010.
- [8] C. Iovescu and S. Rao, "The fundamentals of millimeter wave sensors", *Texas Instruments*, 2017.
- [9] S. M. Patole, M. Torlak, D. Wang and M. Ali, "Automotive radars: A review of signal processing techniques", *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 22-35, Mar. 2017.

- [10] R. Zhang and S. Cao, “Real-time human motion behavior detection via cnn using mmwave radar,” *IEEE Sensors Letters*, vol. 3, no. 2, pp. 1–4, Feb 2019.
- [11] J. S. Suh, S. Ryu, B. Han, J. Choi, J. Kim, and S. Hong, “24 GHz FMCW Radar System for Real-Time Hand Gesture Recognition Using LSTM,” in *2018 Asia-Pacific Microwave Conference (APMC)*, 2018, pp. 860–862
- [12] B. Dekker, S. Jacobs, A. S. Kossen, M. C. Kruithof, A. G. Huizing, and M. Geurts, “Gesture recognition with a low power fmcw radar and a deep convolutional neural network,” *2017 European Radar Conference (EURAD)*, pp. 163–166, 2017.
- [13] Z. Zhang, Z. Tian, and M. Zhou, “Latarn: Dynamic continuous hand gesture recognition using fmcw radar sensor,” *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3278–3289, 2018.
- [14] M. Ritchie, A. Jones, J. Brown, and H. D. Griffiths, “Hand Gesture Classification using 24 GHz FMCW Dual Polarised Radar,” in *International Conference on Radar Systems (Radar 2017)*, 2017, pp. 1–6.
- [15] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, “Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radiofrequency spectrum,” *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, pp. 851–860, 2016.
- [16] Iovescu, Cesar, and Sandeep Rao. ”The fundamentals of millimeter wave sensors.” *Texas Instruments*, SPYY005 (2017).
- [17] Lien, Jaime, et al. ”Soli: Ubiquitous gesture sensing with millimeter wave radar.” *ACM Transactions on Graphics (TOG)* 35.4 (2016): 142.