

智慧賣場服務機器人之開發與應用

(一)摘要

近年來由於少子化問題嚴重，人口逐漸老化，青壯年及幼兒人口大量減少，因此許多高科技國家積極發展機器人相關研究，使得各行各業在工作中逐漸引入機器人來協助人類，以機器取代人力的方式來增加工作的產值與效能。

本計畫著重於開發服務型智慧機器人之賣場服務功能，此機器人之硬體主要結合了 Raspberry Pi 3 Model B、L298N 馬達驅動模組、3.5mm 迷你揚聲喇叭、USB logic 攝像頭來製作，對話方面使用 Google API 來做語音辨識讓機器人能夠和消費者有對話的互動，了解目前所需服務的項目，並利用 tensorflow 和 keras 等深度學習工具訓練影像分類模型，讓機器人可以透過攝像頭對賣場產品做影像辨識判斷產品，做商品總額估價；帶位的移動控制方式則使用 L298N 馬達驅動模組結合直流馬達做驅動，同時也利用紅外線感測器做循線穩定行走方向，為了避免在吵雜環境中語音辨識不佳，無法使用服務功能，並再加以開發一個圖畫介面的 Android APP，能與機器人透過藍芽連線做控制，增加與機器人的溝通方式，進而開發成一台人性化的賣場服務機器人。

(二)研究動機與研究問題

近年來由於 AI 人工智慧快速崛起，在現今社會中機器人已經深入到人們的生活中，譬如近幾年熱門的掃地機器人、陪伴長者小孩的陪伴型機器人，又或者是麵店的刀切麵機器人，隨著機器人的導入，許多業者藉由機器人可以無止盡的運作也不會喊累，可以大大的提升生產率，並且能有效的減少生產過程中的不良率，而當在面對客人時，也能保持良好的服務態度。

每逢國定假日、周年慶，百貨公司及各大賣場總是會被擠得水洩不通，許多消費者會找不到服務人員幫忙，但如果在平時就增加人手，又會增加人事成本，而在服務的過程，有時由於客人過多，服務人員難免會有不耐煩的情緒，進而影響消費者的購物意願，如果此時有一個賣場服務機器人，不但能解決以上的問題，還可以利用機器人互動的噱頭吸引消費者上門購物。此專題的目的就是想要利用 Raspberry Pi 去開發一套能夠帶位、預先估價、對話的賣場服務機器人，由機器人身兼多職，進而減少賣場人士成本並提高顧客的消費品質。而本計畫預計開發之服務功能如下：

- 1.圖形使用者介面 APP
- 2.預先估價
- 3.詢問商品位置
- 4.產品帶位

當消費者在購物需要服務時，可以至服務台尋找機器人做以上之服務，若使用語音時，使用機器人前需先述說喚醒語【比飛】將機器人喚醒，再述說服務的項目，若述說完商品帶位或詢問商品位置後則僅需要再述說欲尋找的產品即可進行服務；若述說預先估價，述說完後，即可將商品在攝像頭前方做掃描，商品掃描完成後再將櫃台圖片讓機器人掃描，即可讓機器人進行計算完成預先估價的服務。

(三)文獻回顧與探討

目前市面上已經出現不少以賣場為導向的功能型機器人，其中最為相仿的機器人是數位所開發的 Shoait 即時影音商務平台「莎莎」如(圖一)與特力屋的「XYZrobot 整合型服務機器人」如(圖二)。

「XYZrobot 整合型服務機器人」的服務功能有:導航帶位、即時線上客戶服務等。此機器人的溝通方式相當單一，僅能透過螢幕點選所需服務內容下達指令，若顧客須商品帶位之服務的話，需由顧客點擊螢幕當中的文字下達指令，若螢幕發生髒污時容易造成觸控不良導致無法點擊亦或是顧客語言不通，不了解螢幕上的文字內容時，則會無法與機器人溝通而無法服務。

Shoait 即時影音商務平台「莎莎」的服務功能有:詢問賣場商品價格、商品位置、叫計程車等，溝通方式亦相當單一，僅能透過語音與屏幕內的機器人做詢問。若賣場環境相當吵雜，則無法與屏幕做溝通啟用服務功能，且此機器人屬於虛擬機器人，限於出現在螢幕中，若消費者尋找不到商品時僅能透過語音與機器人詢問了解商品位置，無法透過實體帶位具體了解商品所在地點。

可發現上述之機器人，互動與溝通的方式皆相當單一，有鑑於此，本計畫旨在開發之機器人將整合並改善以上優缺點，增加多種溝通方式做服務，使此趟購物更加有趣，藉此更加吸引消費者前來。



圖一 Shoait 即時影音商務平台「莎莎」[1] 圖二 XYZrobot 整合型服務機器人[2]

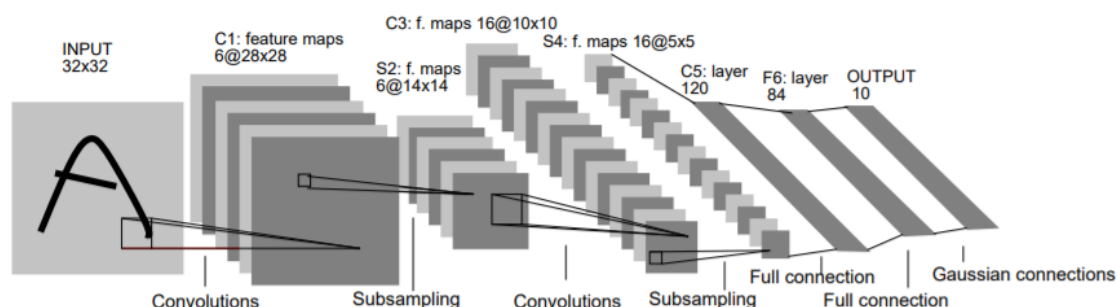
CNN 卷機神經網路:

為了使電腦能分辨圖片不同，可以利用卷機神經網路(Convolutional Neural Network)簡稱 CNN[3]來進行訓練，它是目前深度神經網路(Deep Neural Network)[4]領域發展的主力，在圖片辨別甚至可以做到比人類還精準的程度。

卷積神經網路的搭建結構，主要分為卷積層 1、池化層 1、卷積層 2、池化層 2、以及一個全連結層所構成(圖三)。

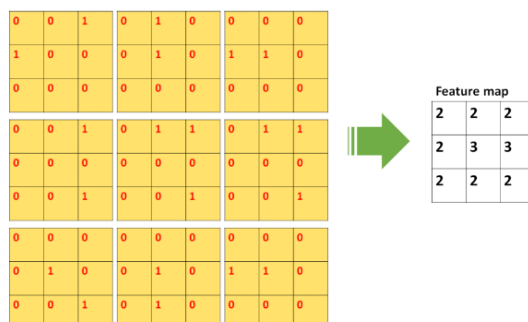
PROC. OF THE IEEE, NOVEMBER 1998

7



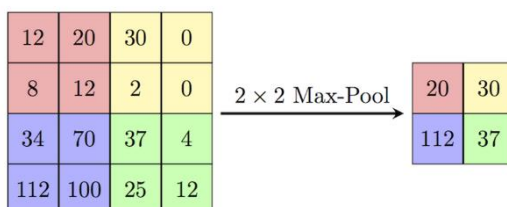
圖三 卷積神經網路 [3]

我們所輸入的圖片其實是一堆數值的矩陣，而將圖片一層一層掃過進行運算，就是所謂的卷積(圖四)，卷積其實就是將圖片做特徵擷取，卷積過後的矩陣即是特徵值。



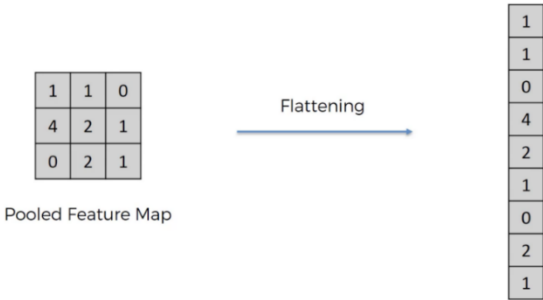
圖四 卷基運作 [5]

當我們的圖片大小相當大時，會十分占用記憶體空間而池化層即能有效地縮減記憶體空間，我們目前採用的方法是 Max Pooling(圖五)只挑選出陣列的最大值，相當於挑選圖片中最明顯的特徵，這樣即能縮減卷積運算時所產生的圖片數量。



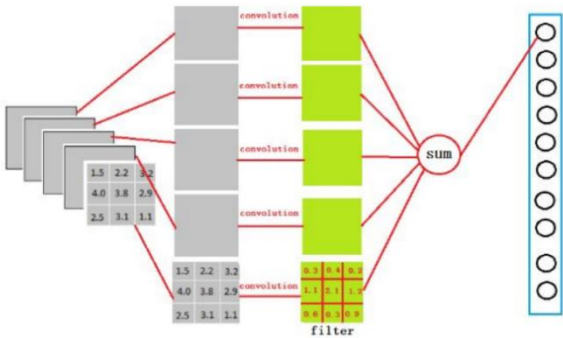
圖五 Max Pooling [5]

為了增強訓練成果，通常會使用兩次的卷積、池化，也就是卷積層 2，池化層 2 的處理，前面幾層得到許多特徵值，全連結層就是將這些圖整合，全連結層大致可以分為三層，平坦層是將前面所得轉輸出為一維矩陣，如(圖六)。



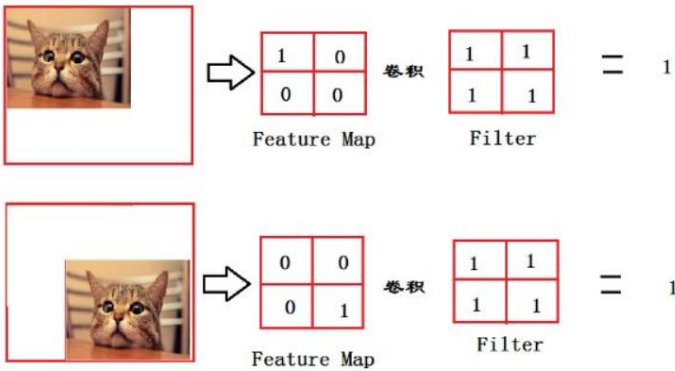
圖六 平坦層 [6]

隱藏層就是在做運算，是將一維陣列透過激活函數[7]得到一個輸出，這個輸出就是一個值(圖七)，輸出層就是隱藏層所得的每個預測值，如(圖七)。



圖七 全連結層 [6]

簡單來說，我們可以得知貓咪的特徵值相同，但是位置不同，原本對於電腦來就算特徵值相同，但位置不同所分類結果也不相同，這時全連結層的作用在於，不管貓在哪裡，我只要貓(圖八)。



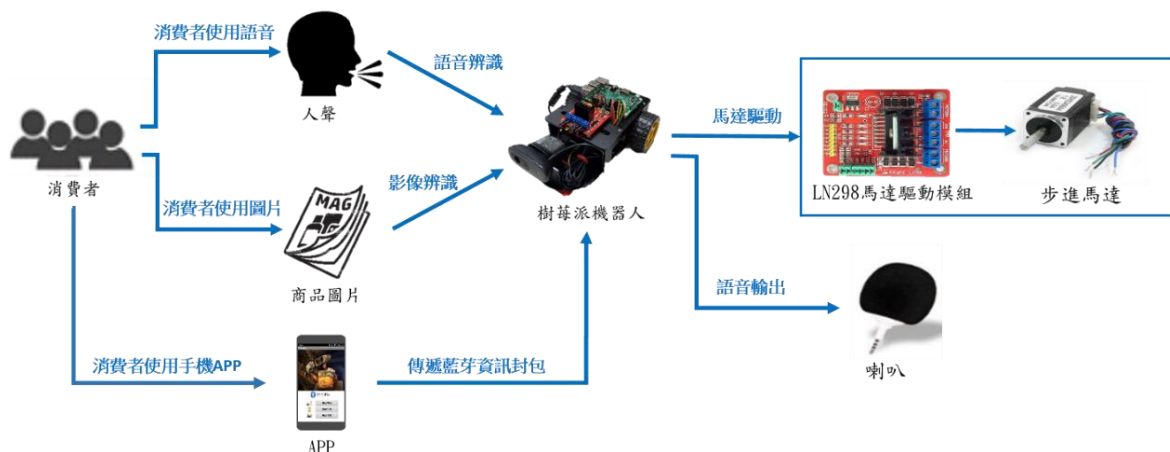
圖八 實用說明 [6]

(四)研究方法及步驟

當消費者在購物需要服務時，可以至服務台尋找機器人做服務，本計畫為了應對不同環境設定多種互動情境，將設計語音與 APP 兩種溝通方式向機器人下達指令。若使用語音時，使用機器人前需先述說喚醒語【比飛】將機器人喚醒，再述說服務的項目，而欲設計的服務項目有商品帶位、預先估價、詢問商品位置三種，若述說完商品帶位或詢問商品位置後則僅需要再述說欲尋找的產品即可進行服務；若述說預先估價，述說完後，即可將商品在攝像頭前方做掃描，商品掃描完成後再將櫃台圖片讓機器人掃描，即可讓機器人進行計算完成預先估價的服務。而環境吵雜無法使用語音時，則可使用 APP 對機器人做藍芽連線，點選所需服務的項目與商品按鈕進行服務，增加多種溝通方式，使互動性更多樣更有趣。

4.1 系統架構圖

圖九為賣場服務機器人之系統架構圖，使用者可使用三種方法控制機器人，語音方式由機器人攝像頭內的麥克風感測顧客聲音，接著透過 API 做語音辨識；APP 控制方式來透過藍芽通訊，向機器人傳遞封包；辨識商品圖片方式透過攝像頭進行影像辨識；若需要回話則透過 3.5mm 迷你喇叭輸出音效；帶位則需要透過 LN298 馬達驅動模組來控制。



圖九 系統架構圖

4.2 機器人功能模組

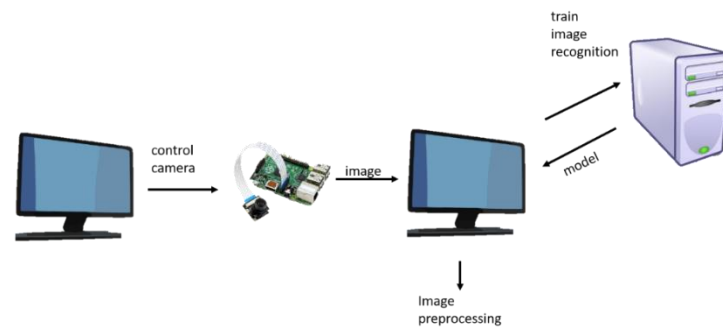
4.2.1 影像辨識

● 蒐集資料

本階段中為使機器人能在商場服務，本計畫預計收集會在賣場內看到的各種商品圖片，控制攝像頭拍攝遠近角度約一千張之照片再加以標註類別，形成訓練模組所需的資料集合。

● 訓練模組

為使電腦可以辨識圖片的不同，本計畫預計使用 Python 來實踐卷積神經網路，將圖片進行分類後，以 Open CV 輸入圖片並加以處理，利用 Keras 套件來建立神經網路，以 Tensorflow 進行運算得出我們想要的訓練模組，並傳給電腦來進行影像辨識(圖十)。



圖十 訓練系統架構圖

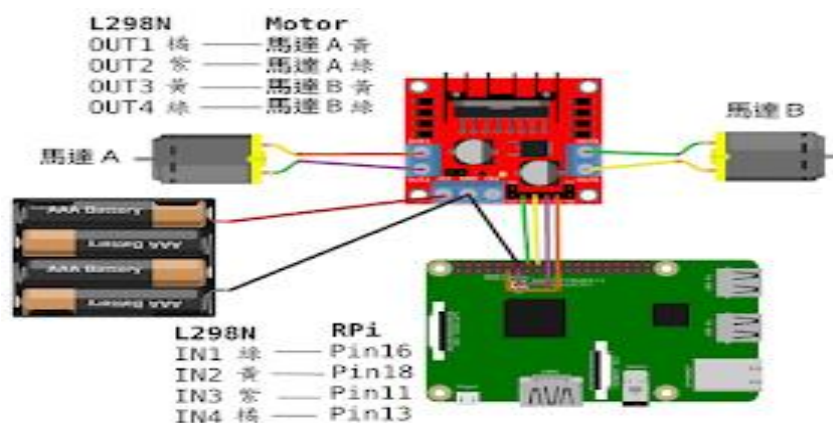
4.2.2 對話控制

機器人之音源相關硬體有:USB logic 攝像頭、3.5mm 迷你喇叭。語音辨識流程為先透過 USB logic 攝像頭內的收音麥克風錄製聲音，將其聲音透過語音辨識之相關套件轉進行判斷語意，來決定回覆之話語，再將其欲回應之語句透過 3.5mm 迷你喇叭發出。

在 Python 套件中一些語音識別服務可通過 API 在線使用，而在這些套件中 SpeechRecognition 套件功能相當完整，SpeechRecognition 庫可滿足多種主流語音 API。而本計畫利用此套建中的 Google API 來做語音辨識，並利用 Python 中字串裡尋找字元的語法來取得關鍵字做判斷。至於語音輸出將使用 gTTS 套件，gTTS 可以將文字轉成語音進行保存為 MP3 格式，然後再讀取語音實現發音，它是通過調用 Google 提供的 TTS 服務來實現將文字轉語音。

4.2.3 馬達控制

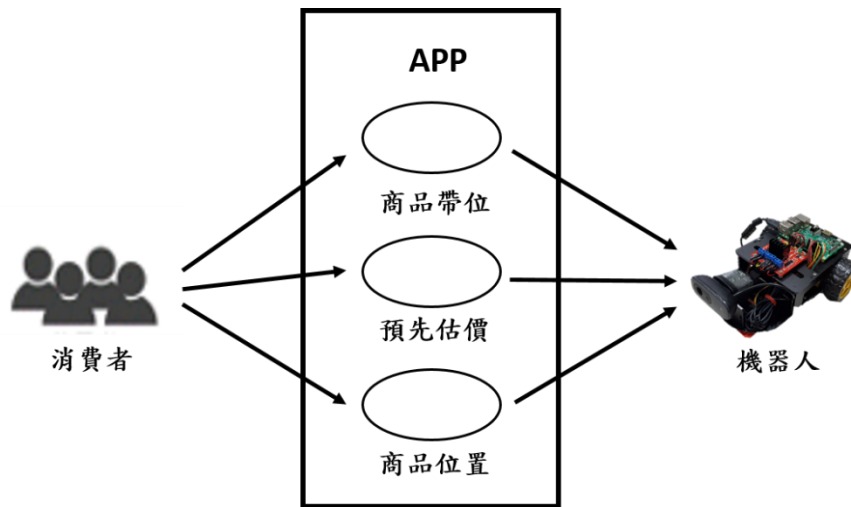
本計畫之馬達控制方法是透過 L298N 馬達驅動模組來使用[8]，大部分在控制 DC 馬達驅動時，若僅是單一方向轉動，只要控制其電壓正負極的接腳及可控制轉向。若需要有控制正反轉能力的話，就必須透過電路設計將馬達電壓反向，故需先如(圖十二)電路圖將 L298N 馬達驅動模組與 Raspberry Pi 和馬達相接並外接 5V、2A 之電源供電，再將 Raspberry Pi 中 16 腳、11 腳設為前進，18 腳、13 腳設為後退，並透過 python 來控制腳位輸出電壓，來決定馬達之正轉、反轉及轉速。



圖十二 馬達控制電路圖[8]

4.3 圖形使用者介面 APP

我們將設計一款 APP 以利讓進入賣場的消費者除了使用語音與機器人互動，能有多一種控制機器人的方式(圖十三)。



圖十三 APP 使用個案圖

4.3.1 介面設計

此 APP 包含首頁和三大主要功能：商品帶位、預先估價和商品位置。首先，進入 App 時會呈現建立連線、三大主要功能鍵及多個商品按鈕(圖十四)，消費者需藉由【建立連線】按鈕先與機器人進行連線，連線完成後，即可點選【商品帶位】【預先估價】【商品位置】按鈕做需服務的選擇。若點選【商品帶位】或【商品位置】按鈕則須再往下滑動，在多種商品中點選所需服務之商品即可進行服務；若點選【預先估價】按鈕則同樣須往下滑動，連續點選預購買之商品，當點選完畢後再點選估價按鈕即可讓機器人進行商品總額估價。



圖十四 APP 介面圖

4.3.2 藍芽通訊

在 Raspberry Pi 物聯網的方式中，TCP/IP 協定與藍芽協定屬於比較主要的方式。有線網路與 WiFi 無線網路使用的是 TCP/IP 協定，而藍芽當然使用的是藍芽協定。兩者皆有優缺點，在 TCP/IP 中，若兩個設備想要順利的溝通往往還需要其他額外的設備，例如：IP 分享器、交換器、路由器等。藍芽協定要求相對較低，只要兩方設備皆支援藍芽就可以直接進行連線。

● 藍芽 Client 端

接收藍芽訊號的 Client 端為消費者的手機，本計畫預計使用 App Inventor 中的 BluetoothClient 元件[9]來開發藍芽相關功能，使手機先與機器人做連線，再利用 SendByteNumber(text number)函式，對已連接的藍牙裝置發送 1 位元組長度之字串，若點選商品圖片則傳送商品名稱之字串，按下不同之按鈕則傳送如(表一)之字串封包。

表一 按鈕所傳送之封包範例

	按鈕	封包字串
1	商品帶位	a
2	預先估價	b
3	商品位置	c
4	商品 1	(product name)
5	商品 2	(product name)
6	商品 3	(product name)
7	估價	count

● 藍芽 Server 端

接收藍芽訊號的 Server 端為賣場之機器人。Raspberry Pi 中級有內建的藍芽系統，版本為 Bluetooth4.1 (Bluetooth Low Energy, BLE)，減少頻寬降低功耗低功號，可直接與手機連線與配對，若要接收到藍芽訊號與資料封包需先安裝 pybluez 等藍芽相關套件，在安裝完套件後，再利用 pybluez 套件來開發，接收資料封包資料，藉由接收到不同的封包來控制機器人之行為。

4.4 預先估價

當消費者選擇完即將採買的商品後，可透過此功能預先了解自己的購物車價錢總額，若要使用此功能需先透過語音，使用喚醒語「比飛」喚醒機器人，再述說「預先估價」之關鍵語句，讓機器人透過語音辨識判斷預服務的功能，當功能啟用成功後，則機器人將述說「請將商品給我看喔」做回覆，此時消費者即可將商品放置攝像頭前方做掃描，掃描成功後機器人將使用影像辨識做判斷並在 python 套件中的 sqlite 資料庫搜尋商品價格，再以商品價格做回覆，當消費者已掃描完預購物之商品後，再將結帳台圖片放置攝像頭前方，則機器人會立即計算並述說掃描之總額做回覆。

4.5 詢問商品位置與商品帶位

當消費者找不到預採買之商品時，可透過此功能來尋找此商品，若要使用此功能需先透過語音，使用喚醒語「比飛」喚醒機器人，再述說「商品位置」或「商品帶位」等關鍵語句，讓機器人透過語音辨識判斷預服務的功能，當功能啟用成功後，則機器人將述說「你想找什麼呢」做回覆，則消費者需再述說預尋找之商品，若使用詢問商品位置之服務機器人則會使用語音回覆並描述商品的所在地點與區域；若使用商品帶位之服務，機器人會透過 LN298 馬達驅動模組來做移動與控制，並在行走同時使用紅外線感測器做循線穩定行走方向。

(五)預期結果

本計畫之預期產出兩項主要成果：

1. Raspberry Pi 機器人：語音對話、影像辨識、馬達控制。
2. Android APP：可與 Raspberry Pi 機器人藍芽連線做控制。

計畫執行期間 109 年 7 月至 110 年 2 月之進度規畫甘特圖如(表二)所示。

表二 進度規劃甘特圖

任務名稱	7 月	8 月	9 月	10 月	11 月	12 月	1 月	2 月
Raspberry Pi 機器人								
蒐集 data								
訓練模組								
語音對話								
馬達控制								
藍牙 serve								
Android APP								
GUI 設計								
系統開發								
系統整合								
系統整合								

(六)參考文獻

- [1] 機器人莎莎
<https://www.cna.com.tw/postwrite/Detail/225441.aspx#.Xh8nNMgzaUk>
- [2] XYZrobot 整合型服務機器人
<https://www.techbang.com/posts/41001-xyzrobot-integrated-service-robot-into-the-hola-taichung-mall-with-indoor-positioning-and-navigation-real-time-customer-service>
- [3] 卷機神經網路(Convolutional Neural Network)
Gradient-Based Learning Applied to Document Recognition
(Yann LeCun, L è on Bottou, Yoshua Bengio, and Patrick Haffner)
- [4] 深度神經網路(Deep Neural Network)
<https://ithelp.ithome.com.tw/articles/10217967?sc=iThelpR>
- [5] 卷機神經網路(Convolutional Neural Network)
<https://chtseng.wordpress.com/2017/09/12/%E5%88%9D%E6%8E%A2%E5%8D%B7%E7%A9%8D%E7%A5%9E%E7%B6%93%E7%B6%B2%E8%B7%AF/>
- [6] 卷機神經網路(Convolutional Neural Network)
<https://zhuanlan.zhihu.com/p/33841176>
- [7] 激活函數
<https://zhuanlan.zhihu.com/p/32824193>
- [8] 馬達驅動模組 L298N 控制
<http://yehnan.blogspot.com/2016/11/raspberry-pi.html?m=1>
- [9] App Inventor 2 指令中文化 BluetoothClient 藍牙用戶端
http://www.appinventor.tw/ai2_connectivity_bluetoothclient

(七)需要指導教授指導內容

- 系統架構規劃與設計
- 機器學習
- Android APP
- 影像辨識
- 語音對話