

心肺音聽診學習與語音掛號系統

The cardiopulmonary sounds auscultative learning and voice registered system

劉郁弦、吳萃秀、陳冠謀、洪資竣、許欣能

指導教授:韓欽銓教授

國立聯合大學 資訊工程學系

苗栗市南勢里聯大 2 號

{U0624012,U0624004,U0624034,U0624038,U0624042}@gm.nuu.edu.tw

摘要

人流密集的醫院容易使有傳染性的病毒、流感傳播，尤其是今年爆發了新型肺炎，人人都需要戴口罩，為了減少醫院人流，所以我們決定開發醫療技術系統。心肺音在醫學方面佔有很重要的角色，為了讓醫學生可以有更多的機會練習聽診，我們利用電子聽診器結合手機錄音與教學平台來讓醫學生可以做練習。心肺音辨識用來協助醫生聽診，因為人耳能聽見的聲音有限，透過電子聽診器錄音，再由電腦將聲音轉成數位訊號來辨識，進而避免在診查時沒聽到細微的聲音，而導致誤診。語音掛號可以讓民眾方便掛號，避免親自到醫院卻無法掛號的情形。

關鍵詞：教學平台、電子聽診器、語音掛號、心肺音、辨識。

Abstract

The virus of infectious and flu are easy to infect in any hospital with lots of people. Especially, COVID-19 has appear in this year. Everyone wear mask on their face. To reduce foot traffic, we decided to develop some technology systems of medical. Cardiopulmonary sounds stands in an important position on medical. To let the student of medical do more and more practices. We combine the Electronic stethoscope and cellphone on the learning platform. The recognition of cardiopulmonary sounds is using at helping doctors doing auscultations. Human ear can't hear soft sounds easily. We used the Electronic stethoscope record cardiopulmonary sounds and change it into digital signal. To avoid the situation of ignoring weak sounds during the diagnosis. People can easily register with using the voice registered. To prevent wasting time in the hospital and can't register.

Keyword: learning platform, Electronic stethoscope, voice registered, cardiopulmonary sounds,

recognition.

一、緒論

聽診是一種聽取聲音的診斷技術。透過聲音判斷有無異常，比如聽到心臟有雜音時就表示心臟可能出現某些問題，肺部聽診聽取呼吸音相對於心音較難去判斷，基本上呼吸音室空氣進入我們的呼吸道和肺泡等器官所產生的氣流聲響較難判斷，讓醫生、研究生，難以教學與學習聽診。

心肺音是看診時最重要的部份，配合儀器，能更精確地去掌握病人的病因，並對症下藥，但醫生每天要面諸多病人，每個人的狀況也都不盡相同，醫生必須仔細聆聽，十分消耗精力，這是醫院面臨的一大問題。

手機的使用日漸頻繁，隨著科技的日新月異，除了熟悉的鍵盤打字，還發展出語音呼叫的功能，人們逐漸跳脫出打字的框架，由於每一個地方口音不同，辨識出來的文字也相同，要能準確的辨識出正確字句，還需要情境的配合，畢竟發音一樣但意思不一樣字詞太多，這是目前語音辨識上的一大難題。

二、系統架構

2.1 系統架構

(一)心肺音學習平台

心肺音聽診教學平台架構(圖 1)分為兩大項:教學端、學生端，兩端的帳戶是共用的，一個帳戶可以當教學方也可以當學習方。

錄音 APP 與電子聽診器連接，選擇聽診位置，畫面顯示出目前要聽診的位置名稱，錄音的檔案名稱是利用聽診位置、日期與時間組合，設置好聽診位置，可以勾選是否進行錄音，如果勾選，按下開始聽診，會一邊聽診一邊錄音，沒有勾選，就只會

聽診而已。錄音結束，點選前往 Studio 的按鈕平台教學端，登入後點選新增好的課程，上傳音檔到課程的檔案儲存區。

課程內容中新增聽診練習的組件，點選編輯把課程儲存區的檔案複製的檔案位置，貼到編輯中的檔案位置後，按下確認後，系統把檔案位置傳送到組件中透過 JavaScript 的 wavesurfer.js 載入音檔後產生音檔波形。之後進入編輯時就可以編輯音檔異常區域，編輯時按下 X 鍵標示異常區域，標示完成後，按下設定答案，透過 ajax post 方式把轉成 JSON 格式的正确答案傳送至 Python 檔案中儲存，按下確認後頁面刷新，發布後就更新課程內容。

學生聽診練習按下 Z 鍵可以標示異常區域，標示完成後可以與卻答案比對重疊率，重疊率是透過記錄下來的正确答案與學生答案，比對重疊率來算出百分比，讓學生了解與正确答案差異。觀看正确答案的按鈕，按下後，教師的正确答案會以標示好的波形方式顯示，方便學生進行對照。

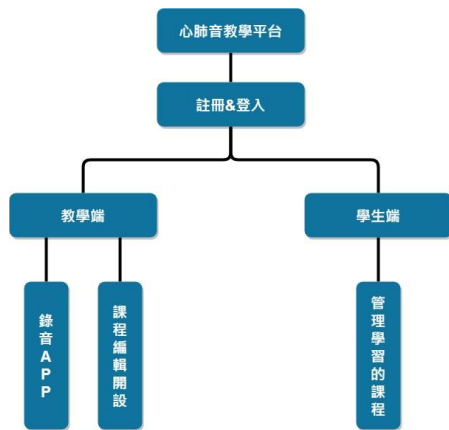


圖 1 心肺音教學平台架構圖

(二)心肺音辨識

心肺音辨識的架構(圖 2)我們分成三個部份，分類、訓練及辨識。

讀取檔案提取特徵並儲存，掃描第一層有哪些子資料夾，並取得完整路徑，接著使用 SVM 分類器進行分類，分為正常與異常兩類，並將其特徵值參數設好訓練，訓練完後存為特定檔案，再讀取要辨識的音檔，套入訓練好的模組，將要辨識的所有音檔進行辨識，然後輸出音檔名稱、預測結果、真實答案，辨識結果與真實答案不同的就輸出檔案名稱、正確結果與預測的結果，等到全部辨識完後將準確率計算後輸出，並繪製混淆矩陣。最後將辨識結果回傳至網頁，按下指定按鈕後顯示於網頁上。

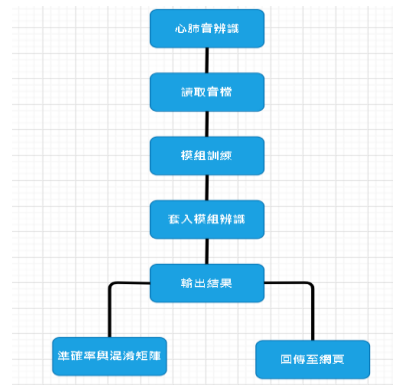


圖 2 心肺音辨識架構圖

(三)語音掛號 APP

語音掛號 APP 的系統架構(圖 3)主要目的為掛號系統，而在進行掛號與使用其他功能之前需要使用者先輸入個人資料，輸入姓名、電話、身分證後方可繼續使用本 APP。

第一個功能為語音掛號，透過連接聊天機器人 (Dialogflow) 進行對話，完成掛號程序，對話中獲取掛號必要資訊，例如:掛號科別、時間(星期)、時段(上午、下午、晚上)、醫師名稱，等資訊；若使用者對自身病情不了解，或不清楚科別，可以詢問、說出病情，將會回復建議的科別，例如:說出肚子不舒服會建議您去掛膽肝腸胃科，在確認到已獲取掛號所需資料後，系統將個人資料與掛號資料合併，然後上傳至後端資料庫儲存，完成掛號程序。

若是使用者對醫院掛號時間不熟悉，可使用科別查詢系統，透過語音輸入欲查詢的科別，系統呼叫後端 PHP 程式並核對資料庫查詢符合的科別名稱，並抓取符合查詢目標傳回 APP 並顯示於畫面上，使用者能直接點選螢幕上的資料進行掛號，點選確定後，系統將合併個人資料然後上傳至後端資料庫。

系統首頁除了語音掛號與科別查詢功能外，增設了跳轉按鈕，點選後將跳轉至心肺音學習網站。

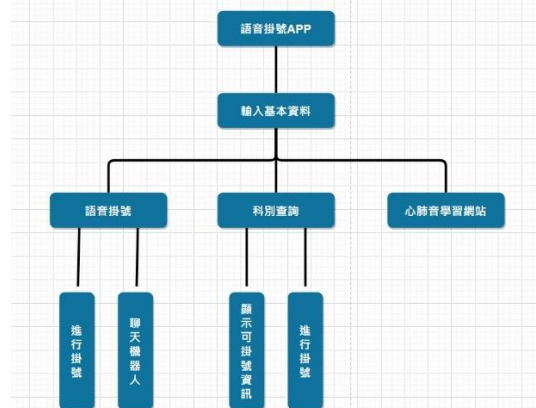


圖 3 語音掛號 APP 架構圖

2.2 開發環境與工具

- A. Open edx:開源的教學平台系統，架設在作業系統為 ubuntu 16.04 上，安裝的版本: Native Open edX platform Ubuntu 16.04 64 bit, version: ironwood.master。
- B. Wavesurfer.js:是一個可自定義的視覺化聲音波形利用 Web Audio API 和 HTML5 Canvas 來實現的。[7]
- C. XBlock 組件:是 edx Mooc 平台的 SDK (Software development kit) ，使用 Python 、 Html 、JavaScript、CSS 編寫的可擴展系統。做為課程內容的編寫方式。[5]
- D. Android Studio: Android 平台開發程式的整合式開發環境。程式語言:Java。
- E. 心肺音辨識:使用 Python 及其相關套件，系統的主要功能分為三部分，分類、訓練及辨識。
- F. 語音掛號系統: 使用 Dialogflow 平台建構聊天機器人，並選擇以 Android Studio 作為 APP 開發環境與機器人連接，後端我們採用 Python 抓取醫院掛號資訊、PHP 做為中繼站與 APP 傳輸資料，使用 Mysql 作為資料庫進行資料存取。

2.3 硬體設備

電子聽診器:不需要呼叫 API 或安裝任何他們產品軟體就可與程式結合，當成一般的麥克風就可使用，可以與行動裝置或是 PC 連接使用，聽診可以使用耳機、2 母 1 公的線與聽診器連接就可使用。製造商: 淄博艾吉電器有限公司。



圖 4 ETZ-1A(C)電子聽診器

三、技術探討與執行

3.1 研究方法

(一) 音軌標示

利用 wavesurfer.js 套件來產生聲音的波形，與產生有顏色的區域，透過 JavaScript 的事件監聽中的 Keydown(按鍵事件)，來獲取使用者按下什麼按鍵，學生端設定為按下 Z 鍵而教師端設定為按下 X 鍵，按下的按鍵是否為 Z、X 鍵，按下後會把當前播放的時間儲存至陣列中，存入陣列之前會先判斷與前一個是否一致，以防使用者連續按按鍵，每按一下也會判斷是否為按下第二次，因按下第一次為記錄標示開始的時間，按下第二次為記錄標示結束的時間，有了開始與結束的時間便可使用 wavesurfer.js 中的 addRegion 來新增標示的區域。設置一個 clear 按鈕，點擊按鈕會把使用者所有標示的區域給清除掉，給使用者可以重新標示異常區域。



圖 5 音軌標示後

(二) 音軌標示嵌入教學網站

音軌標示新增至教學網站中作為課程內容的功能，課程內容是由許多塊不同功能的 XBlock 組件所組合起來，分別使用 XBlock 裡面的 XBlock API、Fields API、Runtime API、Fragment API。[2]

Fragment API 的 Fragment 內容包含在網頁中顯示 XBlock 所需要的所有資料，包含 HTML 內容、JavaScript 和 CSS 資源，每一個只包含內容。所以把先前撰寫好的音軌標示照著 XBlock 的架構去做修改。

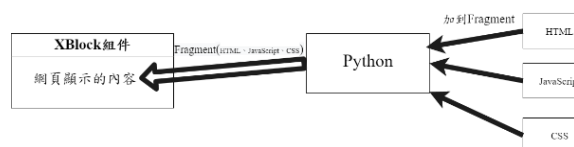


圖 6 XBlock 的網頁資料如何使用

使用 Fragment API 可以讓 Html 使用 python 檔案中的變數，來存取值。Fields API 的 scope.setting 把變數設定為課程設定值，設定檔案位置為課程設定值，編輯課程時可以自由設定以達到可以更換為自己教學的音檔，而 python 分為 student、studio 函式，分別為作用在學習端與教學端。python 中寫了儲存答案的函式 save，儲存答案時用 XBlock API 裡

面的 runtime.handlerUrl 去呼叫 python 中的 save 函式，事先把要傳輸的資料轉換成 JSON 格式，再利用 ajax 中的 post 把資料傳送到 save 函式儲存資料，利用這種方式作為 JavaScript 與 python 的資料傳遞方式。

(三) 錄音 APP

利用電子聽診器不需要額外 API 的優勢，撰寫簡單的錄音 APP。Android 的錄音有分為 Media Recorder、AudioRecord，兩種方式都可以做錄音，而我們需要可以實現與傳統聽診一樣的功能，即時聽診，然而 Media Recorder 的錄音方式會直接把我們錄音下來的音訊資料進行編碼與壓縮，存成一般我們所知道的音訊格式，這種辦法就無法對音訊即時處理，而 AudioRecord 可以對音訊即時處理，缺點輸出是 PCM 的語音資料，不能被播放器播放的，必須額外做處理，而我把 PCM 語音資料額外處理成 wav 格式的音訊格式。

把 Media Recorder 與 AudioRecord 一起使用，前者用來錄音後者用來做即時處理的話，Android 10 之前有麥克風權限問題，擁有 GUI 介面的程式優先，或是先使用的優先使用，結束後其他的才能使用；Android 10 後加入優先權，在切換後，先前的會繼續執行但是會保持靜音，而讓聽診輸出的聲音靜音，也因為這麥克風權的問題，而改為只使用 AudioRecord 做錄音與聽診的實現，錄音與聽診同時，音訊資料寫入緩衝區的同時一邊利用 AudioTrack 讀出資料播放，一邊把緩衝區的資料寫入檔案內，直到結束後，會把 PCM 的語音資料轉成 wav 格式檔案儲存。[8]

錄音與聽診都不是寫在主執行緒中，是開一個新的執行緒去執行錄音與聽診，因為錄音與聽診需要做較多的處理，而主執行緒執行一些簡單的處理，例如：按鈕的觸發、顯示文字。

(四) 心肺音分類

讀檔案提取特徵並儲存，使用 os 套件掃描第一層有哪些子資料夾，並取得完整子資料夾路徑，後讀取子資料夾中的 REFERENCE.csv，此為儲存音檔是正常或異常的檔案，答案來源由 matlab 網站所提供[6]，REFERENCE.csv 中的音檔名作為 key、label 作為其 value，而 value 的值分為 1 及 -1，1 代表此音檔為正常，-1 則代表音檔為異常。再取得子資料夾中的所有檔名，拆解檔案名稱與副檔名，然後只保留讀取到的 wav 檔，讀取來訓練的 wav 音檔總共有 3240 個。再來取得完整的音檔路徑，後讀取音檔，並將音檔切段提取每段特徵值，及最後我們用來訓練的預設 MFCC，利用 python_speech_features 模組中提供的函數 mfcc，因為它有自帶預加重參數只須設定 winlen 音框長為 0.512 秒，winstep 步長為 0.256 秒，默認的 nfft 為 512，但是我們設定的音框長為 512，因此將 nfft 設為 1024，最後將檔名、路

徑、MFCC 與 label 值用 pickle 套件儲存起來。

(五) 心肺音訓練

提取儲存好的 pickle 資料，讀取要訓練的資料並設定每一個 clip 要有多少音框，目前設定的 winlen 音框長 0.512 秒、winstep 步長 0.256 秒，20 個音框為一個 clip，一個 clip 大概是 5.12 秒，將音檔切割成很多小的 clip，再來計算一個音檔能切成幾個 clip，索引值大概如 0~20、20~40、40~60、.....以此類推，換算下來等於 0~5.12 秒、5.12~10.24 秒、10.24~15.36 秒，每個音框有 13 個 MFCC 係數，前 12 個 MFCC 係數加訊框能量疊加得 13 維的係數，我們使用預設的 MFCC 值。由於辨識的特徵只有正常與異常，所以我們使用線性的 SVM，簡單的將特徵分類，SVM 屬於一維，所以要拉平 20*13 變成 260*1，另外修改 label，normal 為 1，abnormal 為 0，將這些分類好的模型以 pickle 套件儲存起來並且做訓練。

(六) 心肺音辨識

將錄音好的檔案，用 excel 標示正常或異常並存為 csv 檔，匯入程式中，即可顯示出標示的正確與預測結果。

讀取資料與模型進行辨識，使用 pickle 套件讀取要辨識資料，總共 301 個音檔，對每一個讀取的音檔進行辨識，使用陣列儲存並辨識出 abnormal 及 normal 的數量，每一個音檔中的 clip 都執行完成後，統計音檔中，不同 clip 的辨識結果是 0 還是 1，計算 abnormal 或 normal 誰比較多，並記錄結果。最後輸出音檔名稱、預測結果、真實答案，辨識錯誤的就輸出檔案名稱、正確結果與預測的結果，並繪製混淆矩陣。例如：301 筆音檔的辨識準確率大約為 99.667%，錯誤的一筆為檔名 b0033(圖 7)。

Fail: b0033, Prediction: abnormal, Truth: normal
Accuracy: 0.9966777408637874

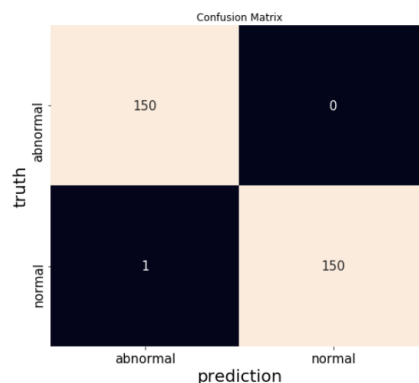


圖 7 辨識結果混淆矩陣

(七) 聊天機器人

我們使用 Dialogflow 作為聊天機器人的建構環境，透過設定 Entity(關鍵字)與 Intents(意圖)，完成

與對話，在 Dialogflow 建立掛號程序，依對話輸入科別、時間(星期)、時段、醫生，再結合 APP 系統完成掛號。

我們依據身體各的部位的病情與對應的科別建立了詢問病情的語境，當使用者提出身體不適時，接著輸入不適的部位，輸入病情，機器人將回復您符合您病情的建議科別，並帶領您回到掛號程序，或是使用者也可以直接結束對話，不進行掛號。

(八) 語音辨識

透過手機麥克風接收到聲波，將其轉為數位音訊。音波會先透過聲學模型將特徵擷取出來後，再透過語言模型進行比對，找出最有可能的字詞，進而達到語音辨識之效果。

(九) 羅馬拼音轉換

在語音輸入方面，可能會因為人名或科別因為有諧音字造成辨識錯誤的問題，而在注音ㄣ以及注音ㄥ的錯誤更是明顯，所以在這個問題上我們使用 TinyPinyin，這是一個 Android 漢字轉拼音開源庫，根據開放詞典網數據，對照漢語的各種拼音系統之間的轉換關係，提供多元化的拼音轉換功能，透過呼叫這個 API 可以把文字轉為羅馬拼音，再經過我們所設定的人名拼音進行比對，即可解決無法以語音辨識辨識出正確名稱的問題。

(十) 醫院 API 與資料庫

醫院資料庫中包含許多資料，其中最敏感的是病患的個人資料不能隨意洩漏、更改，所以醫院方另外建構了一個 API，可以透過特定金鑰讓使用者可以對資料庫進行查詢，透過連接醫院的 API 可以獲取 JSON 格式的醫院掛號資料，呼叫醫院 API 時，需要將醫院給予的查詢碼加上金鑰並進行 MD5 轉換，透過 Python 程式碼進行金鑰 MD5 轉換並連接醫院 API 得到掛號資料的 JSON 檔案，經過 Python 格式轉換成中文，定時進行查詢並存進我們建立的后端資料庫(圖)，其中的掛號資訊有: deptID 是醫院科別 ID 代號; deptName 是醫院科別的名稱; opdData 是日期，包含年、月、日; weekDay 代表星期; opdTimeID 是時段包含上午、下午、晚上在醫院端是使用 1、2、3 代表各時段; roomID 是診間代號; roomLocation 是診間位置提供樓層與大樓位置; doctorID 是醫生代號; doctorName 是醫生名稱; canReg 是顯示可否掛號，以 Y/N 代表; memo 是掛號明細，會顯示當前掛號人數或是其他資訊依醫院方為準。

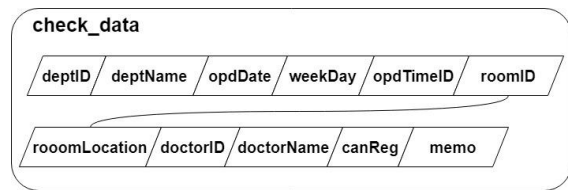


圖 8 check_data(資料表)

(十一) PHP 後端與 MYSQL 資料庫

語音掛號系統需要時常與資料庫進行連接，而 APP 程式無法直接與 MYSQL 資料庫進行連接，所以需要一個資料的中繼站，作為 APP 與資料庫連接的媒介，透過相互 POST 資料的方式讓 PHP 與 APP 溝通，相互傳輸資料庫的資料，MYSQL 是一個功能十分強大的資料庫系統，透過多樣化的 SQL 指令對資料庫的資料進行查詢、統整，PHP 傳輸 SQL 至資料庫，再將回傳的資料進行處理後，POST 整合的資料給 APP，APP 在將資料顯示於畫面上，如此一來就完成與資料庫的連結。

四、結論

心肺音聽診學習平台對於學生，讓他們能增加接觸的病例又不用特定到其他的地方，與近期因疫情關係而漸漸浮出水面的遠端這個詞，平台雖然無法做到即時的互動，但是可以提供許多的內容來教學，聽診部分的教學，因聽診是醫學相關的專業知識，所以在標示的部分，只能專業的醫生來做標示，會花醫生一部分的標示的時間，希望在未來可以有幫助醫生標示異常部位的功能，會減少很多因為標示而花掉的時間。錄音 APP 雖然可以錄音不過因為錄音的時候也要實現聽診，所以錄下來的聲音會有一些雜音，而影響到聽診判斷，有想過要透過降噪的方式降低噪音，不過因降噪有可能會把異常所發出的聲音也一起當成噪音可降低，也會大大的影響聽診判斷，這是錄音 APP 目前無法解決的問題。

目前心肺音辨識的部分，因訓練所使用的音檔是一整段的，所以訓練出來的模組只能用來辨識整段音檔，沒辦法做更細節的辨識，像是單純看音檔哪部分的片段正常或異常，是沒辦法做辨識的，但如果之後有足夠的片段音檔可以拿來訓練，那就可以辨識小片段心肺音的正常與異常。

語音掛號 APP 對於病人而言，讓他們可以進行科別查詢以及了解病症對應於那些科別，讓病人不用到醫院即可快速進行掛號;而對於學生而言，在通勤或者是不在電腦前時，可以利用 APP 裡面的連結至心肺音聽診學習平台上進行進一步的學習。在語音掛號系統上還未與醫院進行連接，所存取的病患掛號資訊目前還只存放於後端的資料庫，若是未來需要與醫院對接的時候需要把資料庫的資料進行處理，傳至醫院端進行完整的掛號程序。

參考文獻

- [1] 余新荃，以聊天機器人為主之長期照顧行動應用開發，大同大學資訊經營研究所碩士論文，民國一〇七年。
<http://www.mis.ttu.edu.tw/ezfiles/72/1072/img/322/199982194.pdf>
- [2] <https://edx.readthedocs.io/projects/xblock/en/latest/fields.html>
- [3] <https://open.edx.org/get-started/get-started-self-managed/>
- [4] <https://zh.wikipedia.org/wiki/EdX>
- [5] <https://en.wikipedia.org/wiki/XBlock>
- [6] MATLAB Heart Sound Classifier
<https://reurl.cc/WL7NGe>
- [7] <https://wavesurfer-js.org/>
- [8] Sharing audio input | Android Developers
<https://developer.android.com/guide/topics/media/sharing-audio-input>