

# 中華大學生物資訊學系系統開發專題報告

## 肺部聽診資料庫分析 肺部疾病判斷

### Lung auscultation data analysis and disease judgment

專題組員：林育均、潘昱伶、趙翊莖

專題編號：106001

指導老師：董其樺老師

#### 1. 摘要

本專題我們採用 kaggle 資料庫中的 Respiratory Sound 資料集，裡頭蒐集了各種呼吸道疾病病患的聽診音檔。我們將音檔透過 Librosa 讀取音頻數據，運用深度神經網路 (Deep Neural Networks, DNN) 和 Gate Recurrent Unit (GRU) 進行深度學習分析，最後訓練出一個模型可用於分析及輔助診斷呼吸道疾病。

#### 2. 簡介

聽診器是醫生判斷肺部疾病的主要工具，但傳統上的聽診器在使用上有許多的不方便，且還可能讓醫護人員陷入被感染疾病的風險，所以電子聽診器的使用為醫護人員提供了更安全的治療環境，結合數位醫療與 AI 人工智慧的電子聽診器，不直接接觸病人，就能透過無線傳輸到喇叭與紀錄系統。且在 COVID-19 疫情期間，可以

透過遠端聽診，隨時掌握病人心、肺等臟器狀況，減少醫護人員進出隔離病房次數，降低感染風險，而在中醫的理論中「初病在肺，中期在脾，末期在腎」[1]，如果能將疾病阻絕於肺之外，即可免去更嚴重的問題。我們希望利用人工智慧來解決醫生聽診的煩惱，並分辨是否患病。

我們於 Kaggle Dataset: Respiratory Sound Database[2] 蒐集 7 種很常見但很難分辨的呼吸道疾病，所蒐集的患者年齡涵蓋：兒童、成人和老年人。其中疾病包含健康 (health)、慢性肺阻塞 (COPD)、氣喘 (Asthma)、上呼吸道感染 (URTI)、下呼吸道感染 (LRTI)、支氣管擴張 (Bronchiectasis)、肺炎 (Pneumonia) [3]、細支氣管炎 (Bronchiolitis)，而慢性肺阻塞佔的比例是最高的。錄音位置分別有：氣管 (Trachea)、左前胸 (Anterior left)、右前胸



#### 4. 專題進行方式



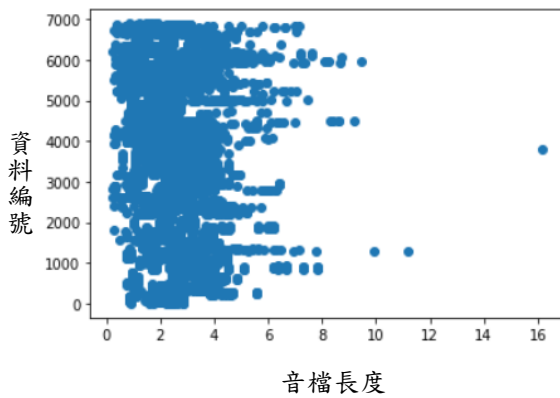
圖二: 專題進行步驟

##### 4.1. 資料來源及資料前處理

採用 kaggle 上的 Respiratory Sound Database 為訓練資料，該資料是由葡萄牙和希臘的兩個研究小組創建，資料其中包括920條帶註釋的錄音，長度由10s 至90s 不等，訓練環境為 Colaboratory。這些錄音來自126 位患者，總長度共有5.5個小時的錄音，126位患者當中一共包含8種病例:健康:26位，322筆資料、慢性肺阻塞:64位，5746筆資料、氣喘:1位，6筆資料、上呼吸道感染:14位，234筆資料、下呼吸道感染:2位，32筆資料、支氣管擴張:7

位，104筆資料、肺炎: 6位，285筆資料、細支氣管炎: 6位，160筆資料，這些病患的錄音檔當中有1864 個濕囉音、886個喘鳴聲、而506個包含濕囉音和喘鳴聲，當中包括乾淨的呼吸聲以及模擬現實生活狀況的嘈雜錄音，並將資料進行切割來擴增資料量，來降低過度擬合的問題。掛接雲端硬碟以讀取資料，並載入必要函式庫，分出八種標籤並設定字典查看病人資料，用以製作表格查看病人呼吸的狀態，當中包含每一筆呼吸的開始以及結束並將有爆裂音、喘息等高頻且連續性的不正常呼吸音，定義切聲音的函數，設定最大長度以確保音檔長度相同，檢查離群值和數據的相對最大值，因前百分之75大約在5.48秒所以取大於等於的最小整數為6秒，定義出一個函式計算切片的原始數據長度，使用 librosa 函數下載原始 wav 檔，並將原始音檔提供給定義出的函數進行切割，並檢查長度，若超過6秒，以6秒切一段，以獲取預期的音檔長度，再使用 librosa 內的 librosa.util.pad\_center 套件在不足長度的音檔數據內補0，使用 one-hot encoding 進行轉換標籤編碼並重塑陣列的型態，並將數據集分為訓練資料(Training)70%、測試資料

(Testing)30%[6-7]。



圖三:音頻長度分布圖

#### 4.2. 建立模型及訓練模型

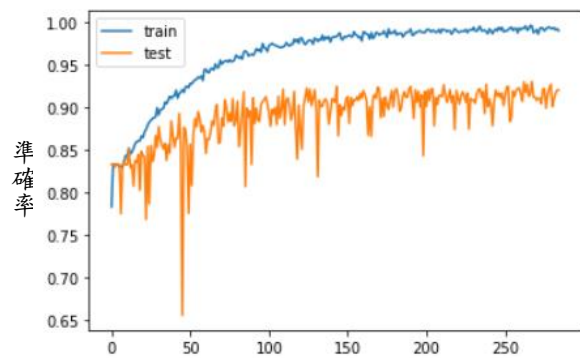
```
Model: "model"
Layer (type)                Output Shape                Param #
-----
input_1 (InputLayer)        [(None, 48000, 1)]         0
batch_normalization (Batch Normalization) (None, 48000, 1)         4
conv1d (Conv1D)              (None, 47988, 32)         448
max_pooling1d (MaxPooling1D) (None, 15996, 32)         0
dropout (Dropout)            (None, 15996, 32)         0
conv1d_1 (Conv1D)            (None, 15986, 64)         22592
max_pooling1d_1 (MaxPooling1D) (None, 5328, 64)         0
dropout_1 (Dropout)          (None, 5328, 64)         0
conv1d_2 (Conv1D)            (None, 5320, 128)         73856
max_pooling1d_2 (MaxPooling1D) (None, 1773, 128)         0
dropout_2 (Dropout)          (None, 1773, 128)         0
batch_normalization_1 (Batch Normalization) (None, 1773, 128)         512
gru (GRU)                    (None, 1773, 64)          37248
gru_1 (GRU)                  (None, 64)                 24960
batch_normalization_2 (Batch Normalization) (None, 64)                 256
dense (Dense)                (None, 64)                 4160
dense_1 (Dense)              (None, 8)                  520
-----
Total params: 164,556
Trainable params: 164,170
Non-trainable params: 386
```

圖四:網路模型圖

建立的網路模型架構總共有3層的捲積層、3層的池化層、2層的GRU及1層隱藏層(如圖四)，並使用參數標準化進行標準化，捲積層是為了

提取每一次呼吸循環的特徵值，捲積層集隱藏層的活化函數皆使用relu，輸入後先標準化，之後建立一層捲積，每一批次以32筆資料池化提取特徵值，設定隨機丟棄(Dropout)為0.3，再進行標準化，再建立一層捲積，以64筆資料池化提取特徵，隨機丟棄同為0.3，再進行標準化，再建立一層捲積，每一批次以128筆資料池化提取特徵，Dropout同為0.3，再進行標準化，並使用兩層GRU以64個神經元進行記憶神經網路學習，訓練模型設定使用compile函數進行設定，使用損失函數(loss)交叉商期訓練，optimizer優化adam使訓練更收斂，用以提高準確率，metrics設定評估模型方法以準確率方法進行，將EarlyStopping設定為如果準確率在200代以內訓練數據不再提升時就停止訓練模型，學習進行500代，每一代的學習量為32筆資料。

#### 5. 主要成果



圖五:模型訓練結果 epoch

	predict	0	1	2	3	4	5	6	7
	diagnosis								
	Asthma	4	0	0	0	1	0	1	0
	Bronchiectasis	0	98	0	6	0	0	0	0
	Bronchiolitis	0	0	131	3	14	0	0	12
	COPD	0	4	1	5703	14	1	19	4
	Healthy	0	0	2	16	294	2	2	6
將肺炎預測成健康	LRTI	0	0	1	1	9	20	0	1
	Pneumonia	0	0	1	18	14	0	247	5
將上呼吸道感染預測為健康	URTI	0	0	4	3	46	0	2	188

圖六:混淆矩陣

我們的模型學習完的準確率達 93%(如圖五)，但有過度擬合 (Overfitting)，有可能是因為資料量分布不平均以及訓練模型的設計過大造成的，混淆矩陣結果如圖六，成功辨識第0類氣喘4筆，成功辨識第1類支氣管擴張98筆、細支氣管炎131筆、慢性肺阻塞5703筆、健康294筆、下呼吸道感染20筆、肺炎247筆、上呼吸道感染188筆，從混淆矩陣中可發現肺炎及上呼吸道感染容易被誤判為健康。

## 6. 評估與展望

本專題研究在學習過程中我們有注意到各種疾病的資料個數不均，我們使用 Librosa 將音檔資料進行剪取，把每一次的呼吸循環剪取出來，來達到資料擴增，並將過多的資料進行刪減，以平衡學習，但有些資料量仍然過於缺乏，導致學習成效不如預期。未來若有可能蒐集到更多的學習資料，可將本專題進行完善並製作成網路平台，讓民眾上傳錄音檔進行分析，或是進一步結合在電子聽診器裡做為附加功能，進行即時的聲音辨識，以利

醫護人員進行診斷。

目前人工智慧在音檔學習的部分絕大部分偏向於視覺上的訓練，從我們的學習成果發現，雖然我們的網路架構設計的並不完善，導致學習成效不佳，但其實直接運用音頻的原始數據進行深度學習也是可行的，而且未必會比使用圖片分析的成效差，因為使用圖片分析可能會導致轉檔上出現誤差，所以細節上容易出現混亂，而使用原始檔案直接進行學習，不透過轉檔的步驟，就不容易出現細節上的混亂。

## 7. 結語

雖然現今已是資訊發達的時代，但對於呼吸疾病的診斷，仍然需要醫護人員與病患直接接觸來進行診斷，這項研究是希望結合 AI 人工智慧的應用，讓醫護人員在診斷上可以擁有更安全、更便利的協助也讓一般民眾可以更了解自己的健康狀況，達到及早發現、治療的預防醫學目標。

## 8. 銘謝

本專題能夠順利的完成，首先要感謝我們的專題指導老師董其樺老師，老師時時給我們寶貴的建議與指導，由於老師細心的教導，本專題才能如期完成。

最後還要謝謝組員們的配合，才能順利的完成專題製作，雖然還有很多要



學習的地方要去努力，但還是謝謝大家共同的付出，齊心完成本專題。

## 9. 參考文獻

- [1] 第15講 五臟：肺的生理功能(1)

Retrieved from:

<https://theqi.com/cmed/class/class5/c15.html>

- [2] Kaggle Dataset: Respiratory Sound

Database. Use audio recordings to detect respiratory diseases. Retrieved from:

<https://www.kaggle.com/vbookshelf/respiratory-sound-database>

- [3] 肺是人體的空調。Retrieved from:

<https://www.e-quit.org/CustomPage/HtmlEditorPage.aspx?Mid=1058&ML=3>

- [4] 梅爾倒頻譜(Mel-Frequency Cepstrum, MFC)。Retrieved from:

<https://zh.wikipedia.org/wiki/%E6%A2%85%E7%88%BE%E5%80%92%E9%A0%BB%E8%AD%9C>

- [5] TengYuan Chang：比較長短期記憶模型(LSTM)與改良後的遞歸神經網路模型：GRU。Retrieved from:

<https://medium.com/@tengyuanchang/%E6%AF%94%E8%BC%83%E9%95%B7%E7%9F%AD%E6%9C%9F%E8%A8%98%E6%86%B6%E6%A8%A1%E5%9E%8B-lstm-%E8%88%87%E6%94%B9%E8%89%AF%E5%BE%8C%E7%9A%84%E>

[9%81%9E%E6%AD%B8%E7%A5%9E%E7%B6%93%E7%B6%B2%E8%B7%AF%E6%A8%A1%E5%9E%8B-gru-813dec22ec6d](https://www.kaggle.com/danaelisanicol/cnn-part-1-create-subslices-for-each-sound)

- [6] CNN (Part 1): Create subslices for each sound. Retrieved from:

<https://www.kaggle.com/danaelisanicol/as/cnn-part-1-create-subslices-for-each-sound>

- [7] CNN (Part 2): Split to train and test. Retrieved from:

<https://www.kaggle.com/danaelisanicol/as/cnn-part-2-split-to-train-and-test>