

國立清華大學 電機工程學系
實作專題研究摘要

**Korotkoff Sound-Based
Augmentation Strategy for Enhanced
Heart Failure Detection**
基於柯式音增強策略的心臟衰竭檢測
優化方法

專題領域：資工領域

組別：A284

指導教授：李祈均

組員姓名：黃宇妍

研究期間：2022 年 9 月 1 日至 2023 年 4 月 29 日止，共8個月

摘要

心臟衰竭是心臟肌肉受損或負荷過度，以致無法輸送足夠的血液供給身體之需要而產生的嚴重症狀，他的高致病率、致死率以及罹患人數日益增升的狀況讓及早發現心臟衰竭這件變得愈來愈重要。在過去，許多研究利用心音來做判別，也證明了音訊這種媒介在心臟病診斷上的可行性。此項研究預計將結合柯氏音、訊號處理及機器學習的技術，提出一種非侵入性而能自動偵測心臟衰竭的診斷方法。在其中，將透過音訊重新取樣、取出連續小波轉換 (CWT)、梅爾濾波器特徵(FBank)和梅爾頻率倒譜係數 (MFCC)這三種特徵，並透過 SpecAugment 將 FBank 特徵資料集做資料強化，最後利用機器學習中的隨機森林(Random Forest)演算法模型來做交叉驗證，訓練後得出正確率(Accuracy)、靈敏度(sensitivity)、特異度(specificity)、F1 分數(F1 score)以及未加權平均召回率(UAR)指標，並在比對過後得出最好的分類組合。研究結果顯示，在四種資料集中表現最好的是經資料強化過後的 FBank 特徵資料集，在正確率(Accuracy)、靈敏度(sensitivity)、特異度(specificity)、F1 分數(F1 score)以及未加權平均召回率(UAR)的表現分別是 91.71%、75.18%、82.76%、83.44%和 82.77%，可以發現特徵種類選擇和資料強化的重要性。本研究顯示了報告中提出的基於柯氏音增強策略的心臟衰竭判斷方法能夠正確判斷心臟衰竭，提供一個非侵入性而便捷的心臟衰竭偵測方法。

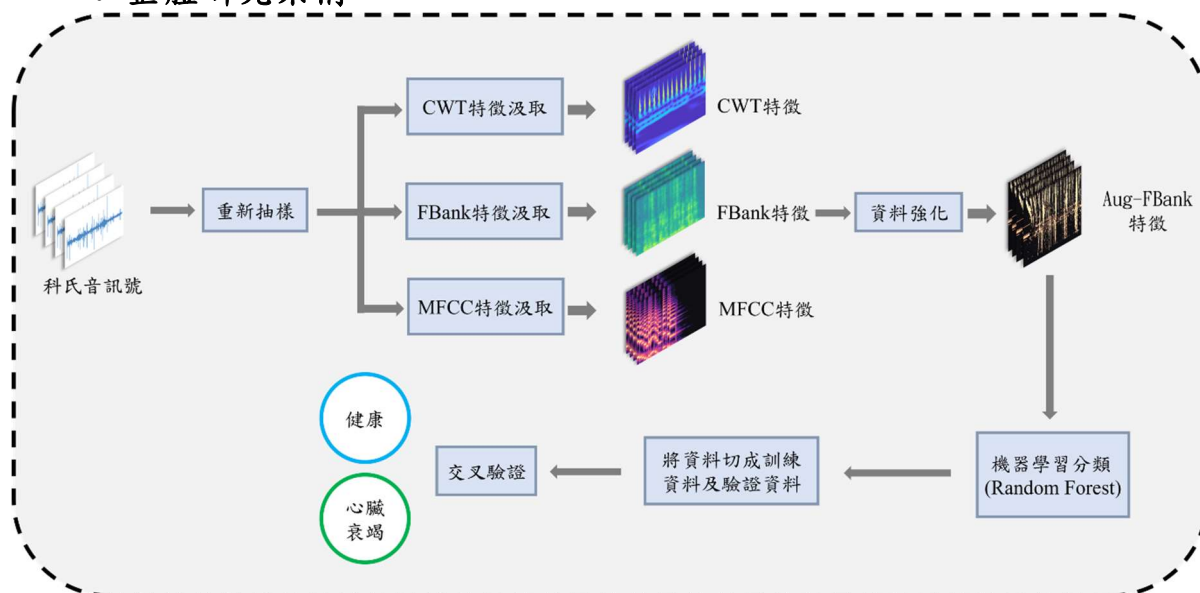
內容

一、 研究動機與目的

根據世界衛生組織 2022 的資料指出，心血管疾病為全球排名第一的死因，且其占比正逐年攀升。近年來，隨著科學技術的進步，大數據及人工智慧被廣泛地運用在各個領域的研究。在其中，機器學習在的訊號應用上有相當不錯的成果，但醫療資料取得不易，使得原先的模型在一些少見的異常心音分類上準確率低下。然而，只要透過強化訓練資料的方式，就能有效改善訓練成果針對各種類型心音的正確率。Covid-19 疫情席捲全世界，大量的醫療資源缺乏使得人們更加難以監測並追蹤心血管疾病。為此，此項研究希望能透過上述兩項技術的結合，利用柯氏音的語音資料做機器學習，將心血管疾病偵測自動化，減輕人力及金錢需求，加速並擴大心血管疾病的偵測。此研究希望能利用柯氏音這種非侵入、測量簡單且易於取得的血壓訊號，運用機器學習的方式，將心臟衰竭的偵測自動化。

二、 研究方法

1. 整體研究架構



(圖一) 整體研究架構

2. 資料庫(Database)

由心音分類挑戰提供，包含了 351 個正常心音以及 194 個不正常心音共 545 個音檔，而其中不正常心音又分為心雜音、多餘心音、心律不整以及帶有噪音的心雜音四種類型。

3. 特徵值(Feature)

在擁有的訊號資料中做連續小波轉換(CWT)的時頻分析、汲取梅爾濾波器組(Fbank)和梅爾頻率倒譜係數(MFCC)的特徵值，而 CWT、MFCC 和 FBank 能將原本的音訊資料轉換為二維的頻譜圖特徵。

● 連續小波轉換 (CWT)

CWT 是一種用來分解連續時間的函數，在分析時間和頻率之間的關聯上非常有效。此研究中，CWT 濾波器組會用(3, 60)的莫萊小波分析，每八度有 48 個小波的帶通濾波器，並透過 L1 標準化來確保在連續小波轉換的過程振幅在不同的尺度下都能維持相同的強度。

● 梅爾濾波器特徵(FBank)

梅爾單位非常適合用來模擬人類語音行為，因此被廣泛的應用於許多人類語音的領域，例如語音辨識和講者辨識等。要得到 FBank 需要經過預加重、分幀處理、切分視窗、短時傅立葉轉換、梅爾濾波並將最後結果取對數。而其中的梅爾濾波步驟是將一組由 20 至 40 個三角濾波器組成的梅爾濾波器組應用在經過短時傅立葉轉換後的結果。此研究使用 26 個梅爾濾波器，框架長度設為 25，框架平移值設為 10，並用每 10ms 的 25ms 漢明窗函數，得到 $128 \times 100t$ 的 FBank 頻譜圖特徵。

● 梅爾頻率倒譜係數 (MFCC)

汲取 MFCC 的過程包含預加重、分幀處理、切分視窗、快速傅立葉轉換、梅爾濾波、離散餘弦轉換。離散餘弦轉換常被用於將對數梅爾頻譜轉換回時域，也就是將其轉換為梅爾頻率倒譜係數，在汲取柯氏音的 MFCC 上，使用 26 個梅爾濾波器，框架長度設為 256，框架平移值設為 128，並用漢明窗函數來防止頻譜洩漏。

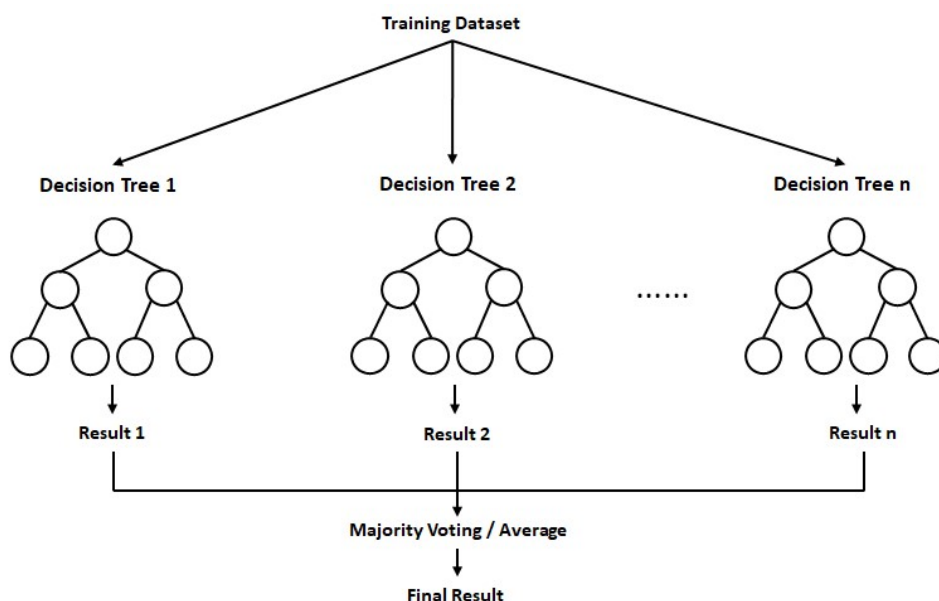
4. 資料強化(Augmentation)

在資料強化部分，針對原先因資料量少而分類結果較差的類別，利用 SpecAugment 的方式將表現最好的 FBank 資料集做資料量的平衡和強化。

SpecAugment 透過時間規整(Time Warping)、時間掩蔽(Time Masking)和頻率掩蔽(Frequency Masking)的方式使模型在時間形變、部分資訊丟失的狀態下也能正確判斷並分類。

5. 機器學習模型(Machine-learning Model)

隨機森林演算法模型是一個由多顆獨立的決策樹(Decision Tree)集合而成，以投票決定最後結果的分類模型，其中心原理為引導聚集算法，也就是從訓練資料中隨機取出幾種特徵值，再以這些取出的特徵值訓練出決策樹，而每次取出的特徵值會再被放回母池中，因此每次取出的特徵值種類會有部分重疊，但因種類並非完全相同且每棵決策樹的訓練過程各自獨立，所以最後訓練完成的決策樹之間仍具差異性。圖二為隨機森林的整體架構。



(圖二) 隨機森林演算法架構

三、 研究結果

經過特徵汲取及資料強化後得到 CWT 特徵、FBank 特徵、MFCC 特徵和 Aug-FBank 特徵資料集，表一顯示不同資料集透過隨機森林模型分類的結果比較，從表中

可以看到 Aug-FBank 資料集有最佳表現，在正確率、靈敏度、特異度、F1 分數和未加權平均召回率分別是 91.71%、75.18%、82.76%、83.44%和 82.77%。表現次佳的是 FBank 資料集，在正確率、靈敏度、特異度、F1 分數和未加權平均召回率分別是 70.73%、74.64%、73.76%、73.76%和 71.90%。第三為 CWT 資料集，在正確率、靈敏度、特異度、F1 分數和未加權平均召回率分別是 67.02%、70.95%、70.28%、68.99%和 67.00%。最後是 MFCC 資料集，在正確率、靈敏度、特異度、F1 分數和未加權平均召回率分別是 61.22%、70.02%、68.44%、65.62%和 65.09%。

(表一) 資料集分類結果

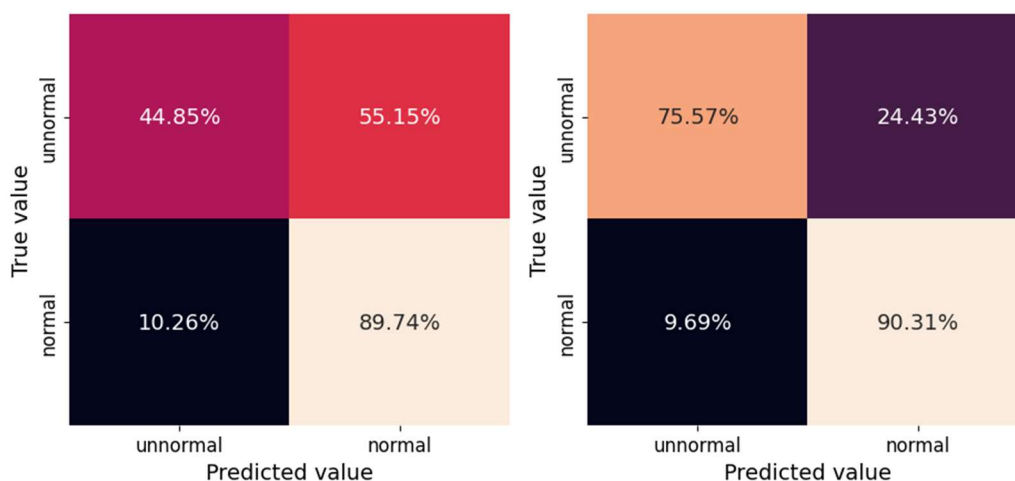
	Sensitivity	Specificity	Accuracy	UAR	F1 Score
MFCC	61.22%	70.02%	68.44%	65.62%	65.09%
CWT	67.02%	70.95%	70.28%	68.99%	67.00%
FBank	70.73%	74.64%	73.76%	73.76%	71.90%
Aug—FBank	91.71%	75.18%	82.76%	83.44%	82.77%

表二顯示不同資料集在非正常心音中每個子類別的 F1 分數。其中各子類最佳表現，extrahs (多餘心音)類別為 FBank 資料集的 89.33%，心雜音(murmur)類別為 Aug-FBank 資料集的 77.86%，心律不整(extrasystole)類別為 Aug-FBank 資料集的 87.87%，帶有噪音的心雜音(noisymurmur)類別則為 Aug-FBank 資料集的 92.12%。

(表二) 資料集的非正常心音子類別 F1 分數

	extrahs	murmur	extrasystole	noisymurmur
MFCC	52.00%	0.62%	0.17%	19.05%
CWT	53.81%	63.11%	13.78%	20.99%
FBank	89.33%	75.17%	0.00%	47.84%
Aug—FBank	79.32%	77.86%	87.87%	92.12%

圖三的(a)為 FBank 資料集分類結果的混淆矩陣，其中非正常心音中被正確分類和誤分至正常心音的比例分別是 44.85%和 55.15%，正常心音中被正確分類和誤分至非正常心音的比例分別是 89.74%和 10.26%。圖三的(b)為 Aug-FBank 資料集分類結果的混淆矩陣，其中非正常心音中被正確分類和誤分至正常心音的比例分別是 75.57%和 24.43%，正常心音中被正確分類和誤分至非正常心音的比例分別是 90.31%和 9.69%。



(圖三) (a) FBank 資料集分類結果混淆矩陣, (b) Aug-FBank 資料集分類結果混淆矩陣

此外，我們觀察非正常心音的各個子分類中被正確分類和錯分成正常心音的音檔間的區別。多餘心音(extrahs)中被正確分類的音檔特色為於某一段有一連串不規律的訊號，在中間間隔後又有一連串不規律訊號。典型的心雜音(murmur)訊號會有明顯一大段起伏卻沒有明確峰值，而一段段起伏之間有相似の間隔。被正確分類的心律不整(extrasystole)音檔多具有明顯但大小不一的峰值，而峰對峰之間の間隔距離也不相同。

正確分類的帶有噪音的心雜音會有明顯一大段起伏卻沒有明確峰值，且和 murmur 相比帶有更多雜音，而不管是在任何非正常心音子類，會被錯分成正常心音的音檔和正常心音音檔的特色相似，大多有明顯的峰值而峰值之間の間隔距離有規律。

四、 結論

在此份報告中提出了一項基於柯氏音增強策略的心臟衰竭判斷方法，透過將資料集取特徵值並分成 CWT、FBank 和 MFCC 三種不同資料集來觀察不同的特徵值對分類結果的影響，並將三類中表現最好的 FBank 做資料強化產生 Aug-FBank 資料集，比較兩者在分類結果的表現。分類方法採用機器學習中的隨機森林演算法做五折交叉驗證，最後使用正確率(Accuracy)、靈敏度(sensitivity)、特異度(specificity)、F1 分數(F1 score)以及未加權平均召回率(UAR)來評估分類表現。

根據研究結果，經資料強化後的資料集在表現上均優於未強化前的特徵資料集。透過比較強化前和強化後的 FBank 資料集，可以發現在每種分類都擁有一定資料量對於分類表現有相當大的影響。此外，在未強化的特徵資料集中，FBank 資料集的表現優於 CWT 資料集和 MFCC 資料集，說明了 FBank 特徵在柯氏音辨識中的優勢地位。此項研究展現了利用柯氏音這種非侵入性、自動偵測心臟衰竭的方法，顯示了他在作為偵測心臟衰竭工具的未來可能性。

截至目前，報告中提出的方法仍在初步的研究階段。在未來，我們會深入研究了解柯氏音的獲取原理，並改善非正常心音中和正常心音有相似特徵的音檔辨識準確率，嘗試應用其他機器學習模型進行優化與擴大訓練及驗證資料集的樣本數，以得到更優異的分類表現。

心得

在專題研究期間也同時修習了機器學習相關的課程，而透過專題研究對語音訊號的資料處理像是切分、降噪籍資料強化等方法有更多的了解，前期透過閱讀 paper 累積和研究主題有關的知識，藉此決定特徵和訓練模型的選擇，中途也繞了不少路，嘗試過很多種不同方法及資料處理方式後才得出報告中的組合，也希望之後能將此項研究更加精進，非常感謝實驗室中博班學姐和教授的指導。