

2022「中技社AI創意競賽」

2022 CTCI Foundation AI Innovation Competition



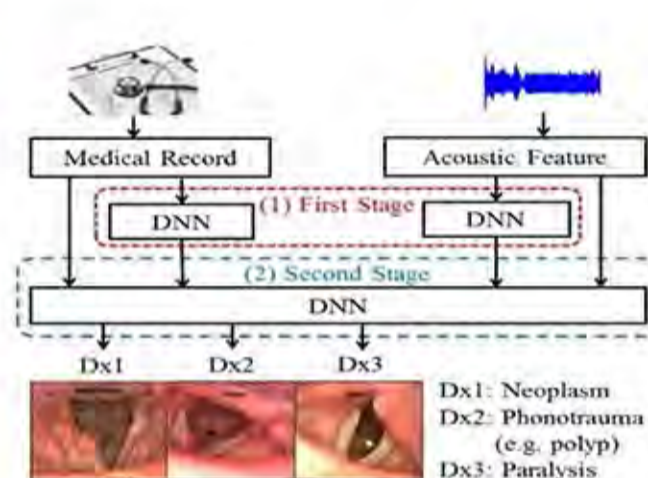
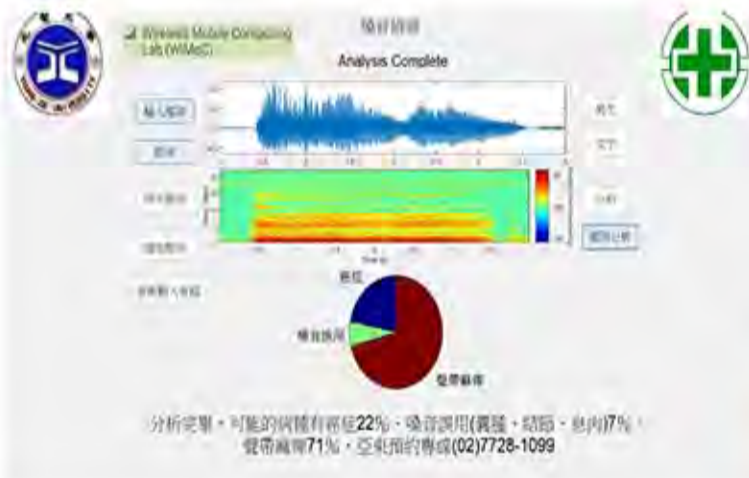
基於可解釋AI之病理嗓音偵測

元智大學 電機工程學系

方士豪教授、王緒翔教授、洪照祥博二、江宥緯碩二
王淳賢碩一、王怡婷大四、張雅瑜大四

作品概述

本作品的目的是透過嗓音訊號進行喉嚨疾病偵測之可解釋性，使用資料，母音「阿」，進行可解釋性分析，使醫生能夠提升診斷效率和節省門診時間。除準確度之外，AI 模型本身可視為“黑盒子”，對結果的解釋往往被忽略。模型的可解釋性能幫助人類從大量歷史資料中發現新現象，萃取有用的新知識，於生醫領域當中，可解釋性至為關鍵，亦較能夠獲得使用者的信賴。

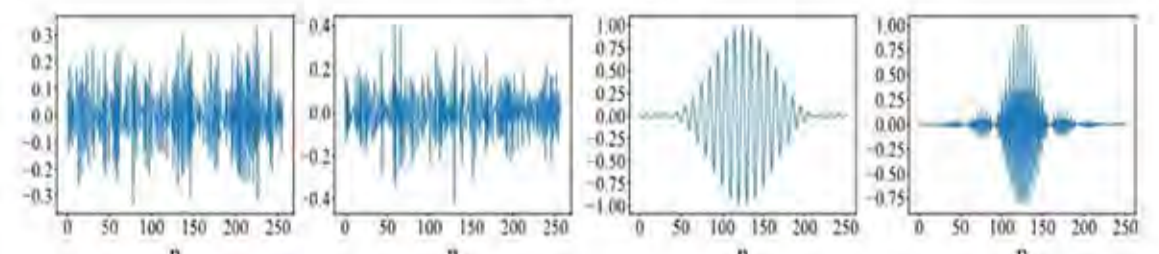
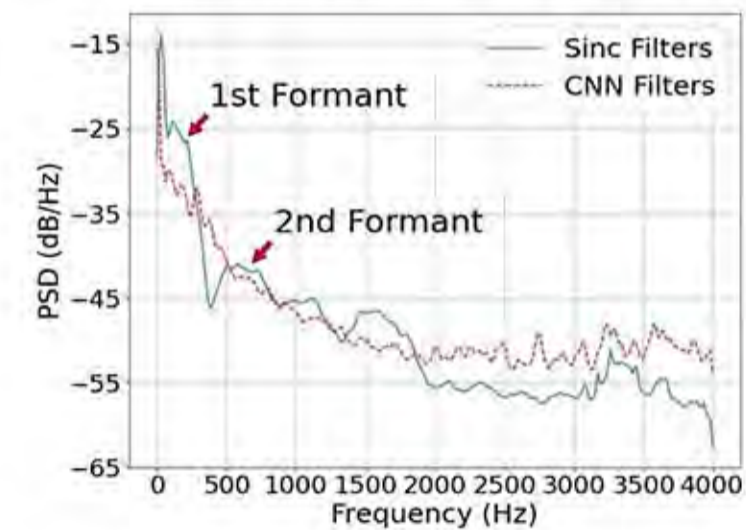


產業及社會貢獻

1. 本作品使用一系列可學習的sinc函數來開發一個可解釋的SincNet系統來學習嗓音疾病的聲音。與傳統的CNN不同，作為SincNet的前端信號處理器，所應用的sinc濾波器可以構建有意義的層，並直接用於提取聲學特徵。最後，我們打算根據我們的評估結果，在系統輸出和第一層提取的語音特徵之間提供可能的解釋。
2. 通過教育部所發起的人工智慧技術及應用人才培育計畫中的其中一項子計畫(AI CUP 2023)，透過訂定競賽議題(多模態病理嗓音分類)，希望能透過競賽，培養解決AI問題的經驗與能力，並導入人工智慧技術與觀念至教育體系，期望將來在投入實習場域進行實戰時，能學以致用。同時舉辦巡迴課程致力於推廣到跨領域的學生，希望能讓非本科系的學生淺顯易懂的接觸人工智慧這塊領域，培養跨領域學生，做到真正的落實人才培育。

創意及核心技術

- 本作品使用新革網路(SincNet)進行病理嗓音偵測與分類。新革網路限制卷積神經網路(convolutional neural network, CNN)的第一層為新革函數(sinc function)，旨在學習更有意義的帶通濾波器，提取聲音特徵，並提昇模型的可解釋性。本文基於亞東醫院所收集的1,061筆阿語音數據進行實驗。實驗結果顯示，與傳統卷積神經網路相比，新革網路不但有較高的準確率和靈敏度，更快的收斂速度，和更好的模型可解釋性。SincNet模型的準確率與CNN模型的準確率相比有顯著的提升(53.15%-61.12%)。
- 本團隊於2019年發表了全球第一篇運用深度學習偵測病理嗓音之論文，並獲得國際期刊Journal of Voice三年內最高被引用次數之殊榮。2021到2022年內發表了4篇期刊與1篇會議的相關論文。



(a) CNN Filters_{35,72}

(b) Sinc Filters_{35,72}