

賴穎暉老師實驗室研究成果簡介

以深度學習技術提升「聽覺」與「溝通」輔具之效益

一、實驗室研究摘要簡介:

語音是人類重要的溝通方式。當「聽」與「說」的能力受損時，將直接影響我們與他人的順暢溝通，進而影響生活品質。有鑒於此，Speech and Bio-signal Processing(SBP)實驗室專注於「聽覺」與「溝通」輔具相關研究工作，並整合「人工智慧」及「半導體」等技術來發展相關醫材，以幫助患者能夠應用這些提出的方法來改善日常生活中的溝通效率，進而提升使用者的生活品質。詳細的實驗室介紹可參考「[SBP Lab 網頁](#)」。未來也歡迎對於聽語科學感興趣的同學們加入，為人類的溝通輔助科技共同努力！

二、SBP Lab 近年於「聽、語輔具」研究成果詳細說明:

A. 「聽覺」輔具研究成果部份:

近年隨著社會高齡化的加劇，聽力損失問題也受到越來越多的重視。根據 2022 年世界衛生組織的統計資料顯示[1]，目前全球約有 5% 的人口 (約 4.3 億人) 受到聽力損失的影響，並預估到 2050 年，這一數字將達到 25 億，且認為至少有 7 億人需要聽力輔具的協助。這些數據直接凸顯了未來需要投入更多研究資源以提升聽覺輔具產品效益的必要性。

隨著美國 over-the-counter (OTC) 法案的通過，近年來已經出現了一些新型態概念的助聽器產品，試圖提升助聽器使用者的聆聽效益。OTC 助聽器產品主要的目的是希望能透過「科技」的輔助下讓使用者自行完成選配，以克服目前市售專業選配人員不足之挑戰，並同時能夠為使用者降低購買助聽器之負擔。而對於 OTC 助聽器產品所需要的關鍵「科技」技術中，SBP Lab 於過去期間著重於「APP 整合主動抗噪技術進行聽力檢測」、「環境噪音情況下的語音聆聽效益」及「以深度學習技術優化助聽器晶片效益」等三部份進行深入的研究，其研究成果分別說明如下：

1. 以 APP 整合主動抗噪技術於非隔音室情況下之聽力檢測準確性:

對於 OTC HA 來說，讓使用者自行進行聽力檢測將會是未來重要的發展趨勢，而克服環境噪音的挑戰將成為此項技術發展的重要議題。過去的研究指出，環境噪音會導致聽檢者的聽力閾值偏移，進而影響聽力檢測的準確性。目前，使用者進行聽力檢測必須在相對安靜的環境中進行，但這對使用者造成不便且限制了檢測的場所。因此，SBP Lab 在過去的研究中深入探討了主動式降噪技術

(Active Noise Cancellation, ANC) [2-4]的應用，並進行多項研究以評估其在聽力檢測中的可行性。

圖 1 展示了本研究團隊於 2022 年開發的聽力篩檢 APP，並結合市售的 ANC 耳機 (如 BOSE、Apple AirPods Pro 等) 進行臨床測試可行性評估計畫。結果顯示，在噪音環境下採用 ANC 技術能顯著改善檢測結果，並與專業聽檢室得到的結果有高度相關性表現 (論文成果請參考著作 2)。然而，在 ANC 產品開發過程中，常受到不同耳機機構特性影響，導致 ANC 能力下降。目前尚缺乏客觀的聲電評估方法來量化 ANC 降噪能力與聽力篩檢準確性之間的關係。因此，SBP Lab 團隊隨後於 2022 年進一步提出了從物理性聲電特性分析方法來建立客觀性評估指標 (圖 2)，並透過臨床驗證得到了 ANC 降噪能力與聽力測試結果之間的對應表 (圖 3)。這提出的驗證方法與臨床測試結果的對應關係有助於研發人員快速了解演算法改良對使用者帶來的潛在效益，進而「加速與降低 ANC 耳機開發成本」(論文成果請參考著作 6)。

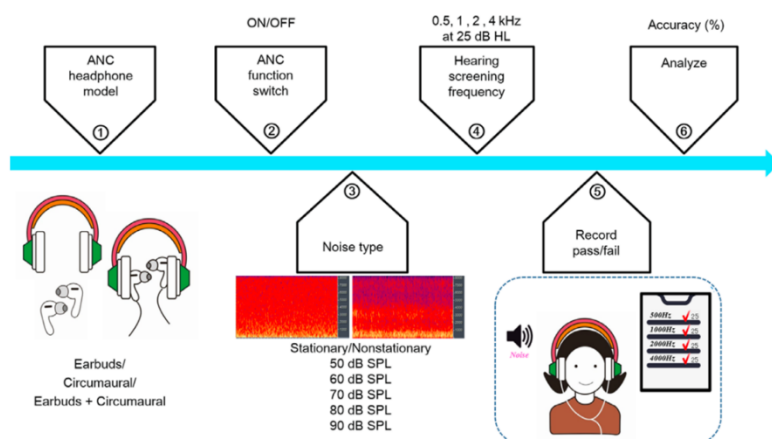


圖 1: SBP Lab 團隊在噪音環境下進行的 ANC 降噪技術聽力篩檢準確性的臨床實驗設計。(參考著作 2)

圖 2: 團隊提出之 ANC 降噪聽檢客觀性評估方法。(參考著作 6)

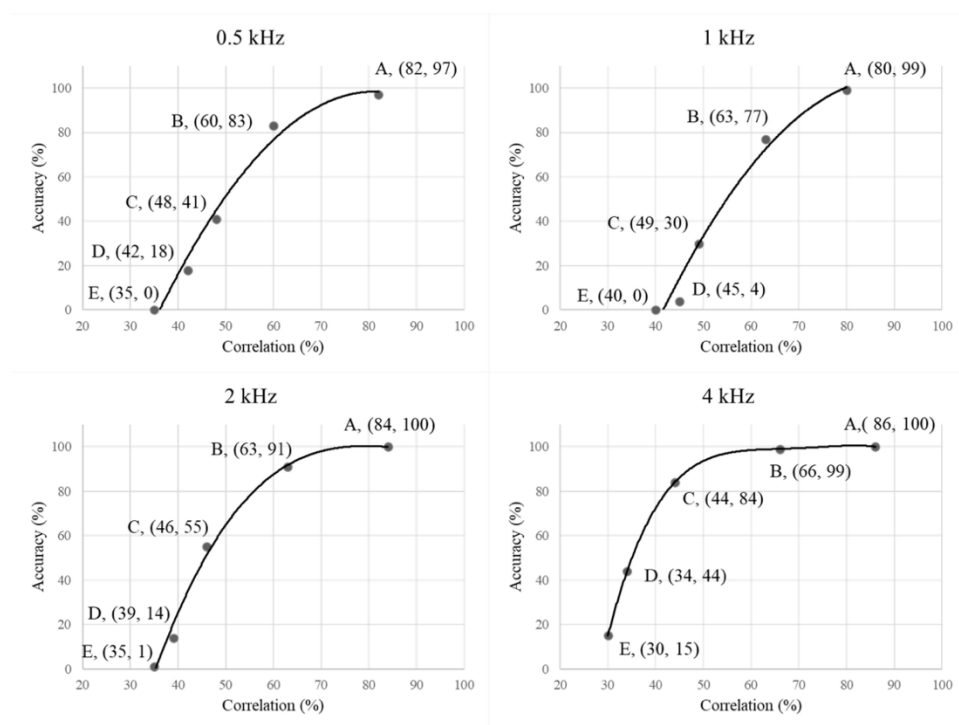
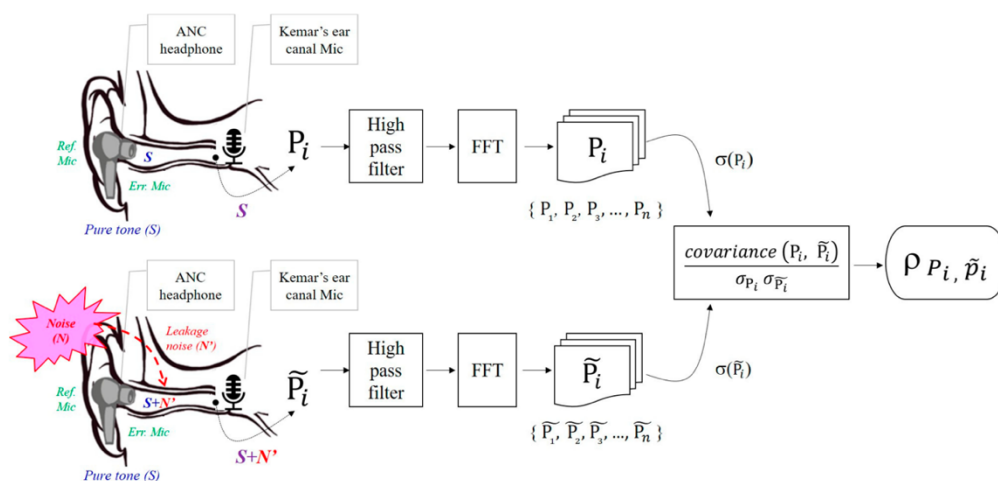


圖 3: 團隊提出的 ANC 降噪聽檢客觀性評估方法所得之分數與臨床效益驗證結果之間的轉換關係。每張子圖的 X 軸代表提出方法量化的分數，Y 軸則表示臨床聽測結果的準確性。(參考著作 6)

上述研發的「主動降噪技術於聽力檢測」研究成果受到產業界 (弘憶國際與瑞昱半導體) 的重視並持續提供後續研究資源支持，促進了上述系統的實際產品



雛型的開發和臨床效益驗證。此外，這些研究成果在 2022 年 4 月份已進行技術轉移至產業端進行產品開發(參考著作 17)，詳細成果請參考官方網站

(<https://rehearaudio.com/>)。同時，這主題部份技術成果也與臨床端（臺北榮總耳科）有密切的臨床研究合作，並於近期也受到「2023 未來科技獎」和「2023 國家新創獎」的肯定和鼓勵。

2. 環境噪音情況下的語音與音樂聆聽效益:

對於聽損者而言，助聽器（或人工電子耳）是目前最常見的解決方法。在過去的幾十年裡，助聽器的信號處理方面已經取得了很多進展，包括補償處方和結構、噪音消除演算法、回饋音消除等。儘管如此，仍然存在許多需要持續改進的部分，例如在噪音環境中的語音表現。為了改善助聽器使用者在噪音環境下的聆聽體驗，各種噪音消除方法被廣泛研究和應用，希望提高語音的訊噪比（signal-to-noise ratio, SNR）、增強語音的理解度（speech intelligibility）和減輕聆聽的負擔（listening effort）[5]。為實現這一目標，SBP Lab 過去嘗試了兩種不同的方法，分別是「深度學習結合麥克風」和「深度學習結合光學感測」等兩種噪音消除方法。此外，SBP Lab 近年再進一步地將過去所熟悉的音訊分離技術應用於幫助聽損者提升音樂辨識能力之研究，其方法為「以深度學習為基礎的音樂音訊分離技術」。而上述三項技術的詳細說明如下：

✓ 深度學習結合麥克風噪音消除技術:

深度學習技術近年來已被證明對噪音消除具有顯著效益，然而，如何有效地將模型縮小並部署到嵌入式系統（或晶片）以實現邊緣運算仍然是一個需要努力的研究方向。基於這一考量，SBP Lab 近年來通過轉移學習技術對過去提出的「以情境感知為基礎之噪音消除架構（NC+DDAE）」進行了優化，提出了稱為 NC+DDAE_T 的方法（見圖 4）。由客觀性分析和臨床聽測結果證實，NC+DDAE_T 能夠提供與 NC+DDAE 系統相似（有時甚至更好）的語音理解效果。此外，NC+DDAE_T 方法的另一優勢是能夠減少近 76.5% 的系統參數量，即用更少的存儲參數和記憶體空間實現系統，從而直接降低了系統實施的成本。我們還將此架構應用於 iOS 平台的應用程式中，完成了原型系統並進行了測試。初步測試顯示，該方法確實能夠使系統在 iOS 平台上實現即時運算，未來可用作人工電子耳裝置的前端處理單元。有關這部分的論文發表成果，請參閱(參考著作 10)。

An environment-aware-based NR for CI users

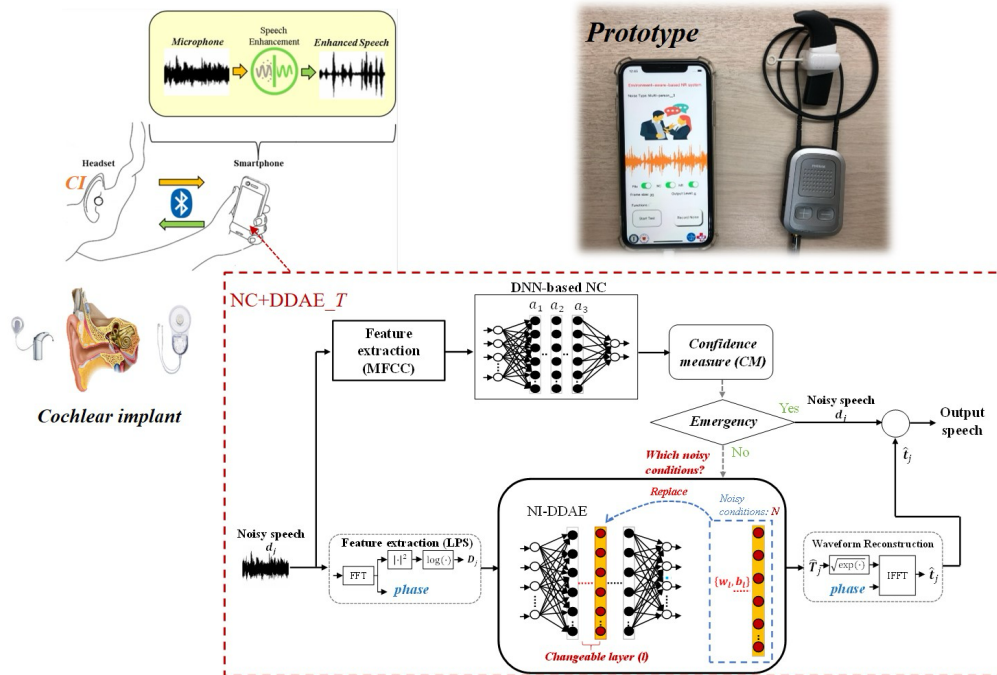


圖4: 以情境感知為基礎之噪音消除架構 (NC+DDAE_T)概念圖。此架構主要延續 NC+DDAE 架構持續進行精進，並引入遷移式學習(transfer learning)進行模型改良及加入客製化情境噪音及緊急音源提醒功能。(參考著作10)

✓ 深度學習結合麥克風噪音消除技術(見光聞聲)

雖然以麥克風整合深度學習技術所開發的噪音消除系統對於聽覺輔具使用者在噪音環境下已有改善，但在面對「語者溝通距離較遠」和「低 SNR 聆聽環境」時仍然存在挑戰和待努力的方向。基於這一考慮，SBP Lab 近年提出了「見光聞聲」收音技術，旨在克服上述挑戰，並利用深度學習技術來提升系統的語音信號還原品質。這項技術主要通過非接觸式的光學感測器(例如雷射都普勒儀[6-8])來獲取聲音源(或其附近)產生的震動信號，使系統不受環境噪音和收音距離的影響。SBP Lab 團隊的最新研究結果表明，透過這種光學感測技術確實能夠比傳統的麥克風技術在收音距離方面獲得更大的優勢，並提升使用者在「低 SNR 聆聽環境」和「遠距交談情況」下的聆聽效益。在這項研究中(圖 5)，我們首先探討了通過雷射都普勒儀來感測語者說話時引起的口罩振動信號，然後通過語音辨識系統試圖辨識這些信號。實驗結果表明，在噪音環境下，所採用的光學感測方法能夠比現有的麥克風裝置有顯著的語音辨識提升。這些成果已於 2022 年 5 月發表在《JASA express letter》期刊(參考著作 7)和 IEEE APSIPA 研討會論文¹

¹ 發表之Enhancing the Performance of Automatic Speech Recognition with Optical Microphone Technology Through Data Augmentation Approach: A Pilot Study論文: <https://reurl.cc/V4aQ5n>。

中，同時也受到美國物理聯合會《科學之光》²的邀請進行專題報導。此外，這些初步成果已部份完成技術轉移至產業進行相關產品發展(參考著作 12)。

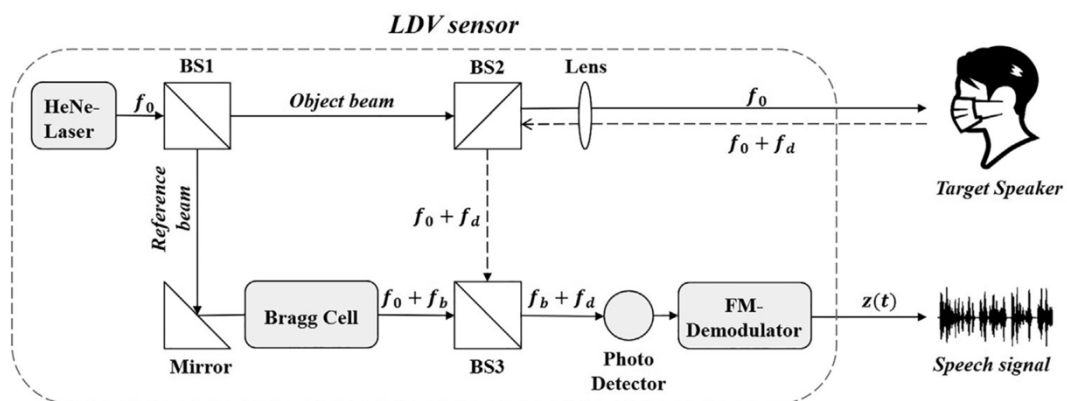


圖 5: 以光學感知技術提升噪音環境下語音擷取能力驗證之系統架構圖(參考著作 7)。

儘管先前研究已經顯示「見光聞聲」技術在噪音情況下能有效提升聽損者的語音聆聽效益，但仍然存在一些需要進一步努力的議題。例如，我們發現聲波在不同物體表面引起的振動受到材料的自然頻率影響[9]，導致擷取到的聲音有所差異，使得系統無法達到最佳的信號處理效果，進而影響聽覺輔具使用者聽到的聲音品質。基於這一問題，SBP Lab 團隊提出了一個語音增強單元(圖 6 中的 speech enhancement unit)，利用深度學習技術來處理不同材料特性對聲音品質的影響，從而提升使用者的聆聽效益。實驗結果顯示，這個新提出的「見光聞聲」系統在噪音和遠距離應用情況下比傳統的麥克風結合深度學習架構(圖 4 概念)表現更佳，能提供更好的語音理解力和聲音品質，有助於聽損者在噪音環境中更清晰地聽到聲音。這項研究成果也受到 IEEE Transactions on Biomedical Engineering 期刊的鼓勵予以刊登(主代表著作 2)。與此同時，另一項技術突破是提出了雙聲道語音增強模型，成功通過單點光學感測技術還原了遠端聲源的立體音訊特性，目前這一研究成果已提交至 IEEE/ACM Transactions on Audio, Speech, and Language Processing 期刊進行審查(目前已經進行了第二次編修，進入小修程度的修訂階段)。

² 美國物理聯合會《科學之光》(AIP Scilight)之報導: Optical microphones measure speech through masks: <https://reurl.cc/Do1LyO>。

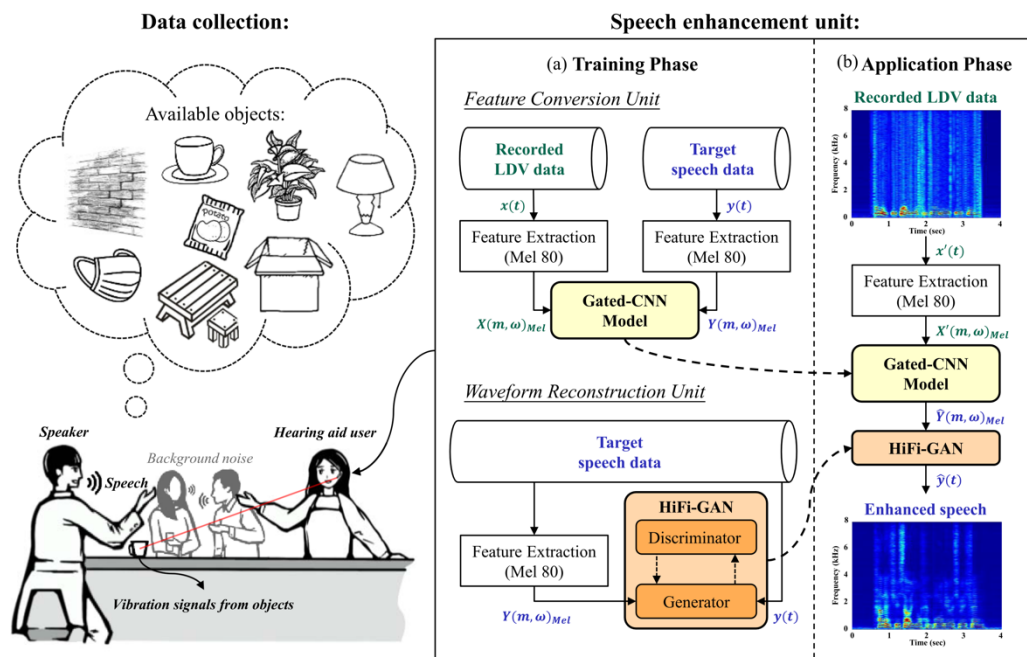
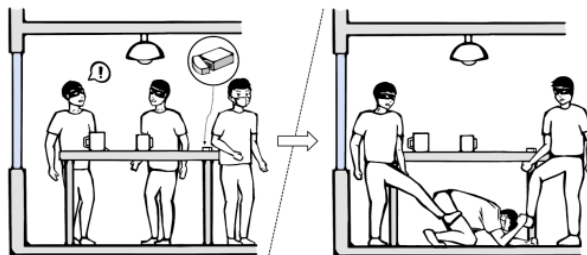


圖 6: SBP Lab 團隊近期提出之「見光聞聲」音訊處理系統(主代表著作 2)。

除了在助聽器應用外，SBP Lab 團隊所開發出之「見光聞聲」技術近期也應用於調查局的研究委託案(2024.01.01至2024.12.16)·透過進一步的改良來運用於法務調查中的合法監聽應用情境中(應用概念如圖 7B 所示)。舉例來說，調查人員在過去為了執行合法監聽時，常採需近距離安裝竊聽器(如偽裝的打火機)進行監聽。然而，竊聽器必須需要調查人員事先進行布置(例如放置或是安裝竊聽裝置)，而這也常使得調查人員可能會有人身安全的風險。而當調查人員在放置竊

(A) 傳統方法



(B) 本計畫提出之方法

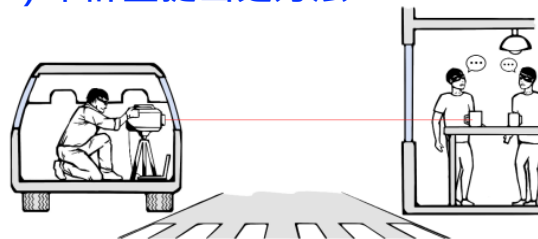


圖 7: 見光聞聲技術未來能予以延伸之應用方向。

聽裝置時，若不幸被嫌犯發現時常會讓調查人員受到傷害(如圖 7A)，進而增加調查人員在執行勤務時的風險。此外，當嫌犯在進行對話時，預先設置好的裝置收音位置、方向不理想時，調查人員將無法立即根據當下的情況(例如：嫌犯遠離原先設置之收音裝置、環境噪音上升...等)進行調整，而導致先前所付出的努力以及

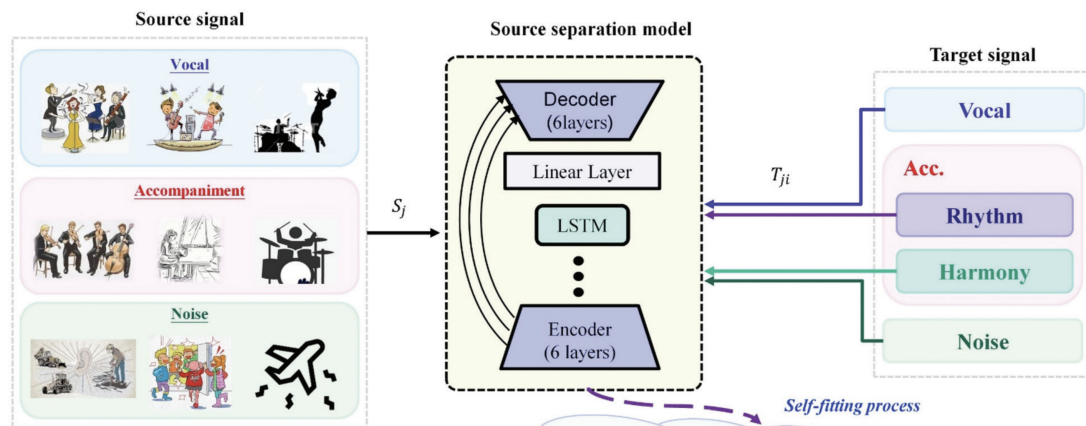
所冒的風險全部功虧一簣。於此應用情況下，若採用本計畫開發之系統來克服「環境噪音」及「遠距交談」問題將對於執法人員在進行監聽任務時十分重要。換言之，若能透過此提出的新形態的聲音感測系統來完成一個長距離收音之功能，其將有助於讓調查人員可以在安全的地方進行監聽行動，進而提升執法人員之安全。

綜合而言，SBP Lab 團隊在這個研究範疇中整合了「語音處理」、「人工智慧」、「光學感測」以及「嵌入式系統」等技術，旨在開發一種創新的收音系統，以解決過去音訊處理中的兩大困難問題。我們將透過一系列客觀分析和臨床驗證，驗證提出的方法所帶來的優勢，以幫助聽力損失者（或其他應用情況下的使用者）獲得最佳的語音聆聽效果，從而提升溝通品質。

✓ 以深度學習為基礎之音樂音訊分離技術

聆賞音樂為何是聽覺輔具使用者臨床應用上的另一聲學難題。以流行樂曲為例：當聽損者聆聽流行樂曲時受到眾多樂器同時產生的聲音時，欲聚焦於某一樂器樂音(或歌手音訊)時，其情境就如同噪音環境下的語音辨識。由於音樂是世界上最複雜的聲音訊息，電子耳執行樂音辨識勢必更為困難。且電子耳內建編碼之限制，音高、音色，與音質感知皆將失真；悅耳的音樂，恐將成為電子耳使用者難以聆賞的噪音。而近期的研究指出，對於電子耳使用者來說，在聆聽音訊時每個人對不同音訊(例如:歌手聲音、背景鼓聲、鋼琴音、吉他...等)的喜好度不同，應能夠獨立調整才能獲得適當的聆聽效益。有鑑於此，SBP Lab 團隊為提升電子耳使用者精準聆聽音樂之效益，我們提出了在時域訊號中以端對端 U-Net 模型，發展基於深度學習且具有即時運算能力之音訊分離系統(圖 8)，讓使用者聆聽音樂時，能分離不同類型樂音之訊號(例如：人聲與樂器)，再透過軟體介面，進行客製化之聆聽模式設定(例如：調整人聲與樂器響度強弱之比例)，以『循序漸進，累加音訊』的方式，逐漸習得人類社會多聲道音樂之常模。於本研究的客觀評估結果看來，提出系統於帶噪情境聆聽現場演唱會下之效益均比 baseline 系統來得優異;此外，於 vocoder 模擬的聽力測試比較下，其提出系統在 MUSHRA 測試中也分別高於兩個 baseline 系統。這項研究成果也受到 Journal of the Acoustical Society of America (JASA)期刊的鼓勵予以刊登(參考著作 1)。

(A) Training stage



(B) Testing stage

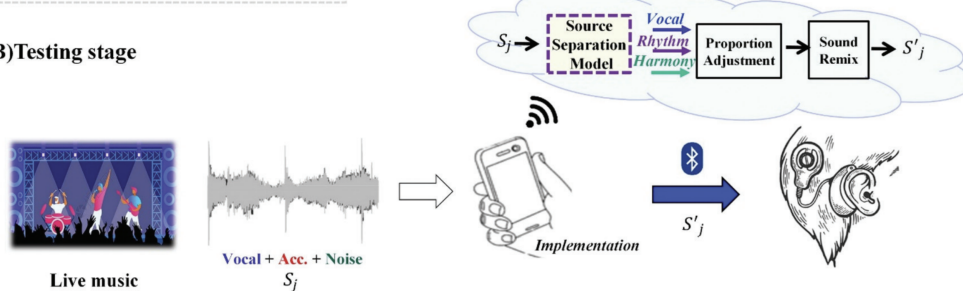


圖 8: SBP Lab 團隊提出之「音樂音訊分離技術」(參考著作 1)。

3. 以深度學習技術優化助聽器晶片效益:

SBP Lab 近年於助聽器研究中，除了在演算法、臨床效益驗證及聲學特性...等理論的研究外，也實際與半導體產業進行密切的合作，進而將「人工智慧」、「聽語科學」、「物理性聲電量測」及「藍芽音訊晶片」整合來完成新式的非處方型助聽器(over-the-counter hearing aid, OTC HA)產品。再在此節研究主軸前期，SBP Lab 與台北榮總團隊於 2022 年初即對當時具備 OTC HA 潛力之聽覺輔具(例如:Apple AirPods Pro or AirPods 2)進行相關的研究(參考著作 5 與 8)，其包含專業型的聲電特性量測及語音聽辨力測試，並與專業型助聽器進行比較(實驗設計如圖 9 所示)。註:參考著作 8 於 2023 年 7 月獲選為「第 19 屆榮總台灣聯合大學合作研究計畫優良論文獎-第二名」。而於實驗結果觀察到，新型態耳機裝置(例如:真無線藍芽耳機)的確有潛力且多項聲電特性均能符合專業型助聽器

之要求;換言之,近年新
 型態的藍芽晶片技術
 已有機會提供聽損者
 改善聽覺感知的能力。
 而此部份的成果也於
 2022 年受到 iScience
 國際期刊肯定予以刊
 登(參考著作 5)。然而,
 在這項研究中我們也
 更深入的發現當時的
 藍芽晶片技術要實際
 做為 OTC HA 產品仍還
 有許多待研究及解決
 的問題,才能讓使用者

獲得 OTC HA 產品的核
 心優點。有鑑於此,SBP
 Lab 也於 2022 年將當
 時看到的待克服之問題
 進行系統化的研究規劃
 (圖 10),並透過產業端
 的合作將研究成果商品
 化,其主要有三項關鍵
 技術(APP 聽檢、AI-
 based 聽覺補償調整及
 聽見原音);於此產品中,
 SBP Lab 至目前已與產
 學合作的 ICT 產業(弘憶
 國際、瑞音生醫及瑞昱
 半導體)有三年期的產
 學合作計畫,並已完成
 三件研發成果技術轉移

(參考著作 11、13 及 17)及四件專利(參考著作 18~21)等成果。而圖 10 部份的
 相關技術成果,也可以透過以下二個影片的介紹讓評審委員能夠更深入的了解
 (影片連結 1、影片連結 2)。

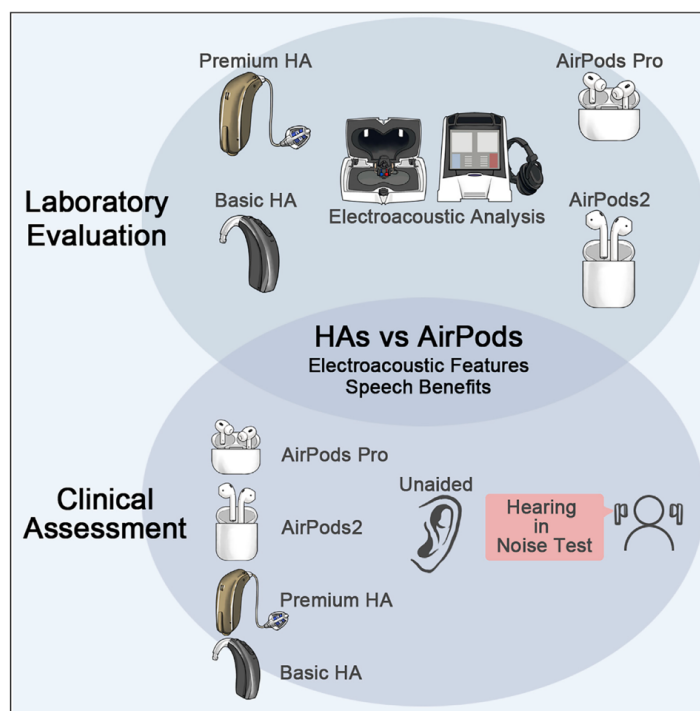


圖 9: SBP Lab 團隊於 2022 年與北榮耳科團隊合作,對於當時真無線藍芽耳機產品做為助聽器可行性之研究設計架構圖。(參考著作 5)

三項關鍵產品技術:

1. APP 聽檢 (聽檢走出聽力室)。
2. AI-based 聽覺補償參數調整。
3. 「聽見原音」補償技術 (Rehear mode):
 → 面對面、串流音訊 & 接聽電話三種模式。



圖 10: SBP Lab 團隊過去將研發成果實踐於藍芽音訊晶片之關鍵流程架構圖,並進一步的與產業合作進行商品化的實踐。(參考著作 11、13 及 17~21)。

總結於 SBP Lab 第一部份「聽覺輔具」的研究中，不僅注重聲學、演算法和軟體優化等理論研究，也強調將這些研發出的理論技術實際應用並進行產品化的實踐。最終目標是讓更多聽力損失者受益於高品質的聽覺輔具，從而改善助聽器使用者的溝通效益。

B. 「溝通」輔具研究成果部份:

人類在溝通過程中，除了聽覺扮演著重要角色外，語者清晰的語音也是至關重要的。然而，若發聲系統出現問題，將直接影響語音的清晰度，進而影響生活品質[10, 11]。人類的發聲過程主要涉及肺部呼出的空氣，這股氣流通過狹窄的聲門，使聲帶上的黏膜產生波動，最終透過擾動空氣產生物理性的聲波。接下來，通過一系列器官（例如咽、口腔、鼻腔及鼻竇）的共鳴進行音量的放大，最終受到嘴唇、牙齒等器官的影響，形成語音。這個複雜的發聲過程受到人類大腦控制。因此，當大腦神經系統或周圍神經系統受損，例如中風[12]、腦性麻痺[12]、帕金森氏症[13]、漸凍人[14]、腦瘤[15]、多發性硬化症[15]，患者常因一個或多個發聲器官出現控制失常問題而產生構音異常，進而直接影響生活品質。根據美國語言聲學學會(ASHA)的研究，構音障礙主要是由於上述疾病患者無法靈活控制發聲所使用的肌肉，進而影響語音的清晰度與自然度，使得他人難以確切理解患者的語意。這種問題直接導致患者的社交障礙，並引發憂鬱、對人群的害怕等心理問題，甚至引發各種社會問題。

對於構音障礙患者而言，目前常見的治療方式是透過專業醫療人員對其症狀和程度進行診斷。根據診斷結果，臨床治療人員整合經驗，進行復健治療。例如，語言治療師根據構音障礙患者的程度設計適當的復健課程，包括改善說話頻率、強化語言肌肉力量、呼吸訓練等[16]。然而，這些治療方式需要患者家屬的耐心學習和配合溝通技巧，且往往費心、費力且費時。這常常使得患者、家屬和醫療人員都面臨巨大的壓力，且由於病情的持續變化，治療效果可能不佳而令人沮喪。此外，對於較嚴重的構音異常患者，可能需要依賴視覺暗示、手語、字卡或者更高階的語音發聲板、眼動追蹤儀等輔助溝通系統（Augmentative and Alternative Communication, AAC）才能有效與他人進行溝通。然而，這些 AAC 對比患者直接使用語音的溝通速率仍有很大的差距[17, 18]。鑒於上述困難，以語音驅動的溝通輔具將有助於仍具有發聲能力的輕度到中度構音異常患者快速提升語音清晰度與溝通效率，改善患者的生活品質。

基於上述臨床問題，SBP Lab 於八年前即展開了一系列的研究工作，並與宇康生科（奧地利公司）長期合作，進行相關技術開發研究工作至今。其主要研究重點包括語料收集、語音轉換系統、語音辨識系統、語音合成器、資料擴增技術、APP 開發、雲端系統建立等演算法開發工作。隨後，SBP Lab 的這一研究主題也

獲得了國科會 107 至 110

年間「射月計畫」的支持，進一步完成了一套以深度

學習技術為基礎的語音轉換系統，旨在幫助構音障礙

患者說出清晰的語音。此外，SBP Lab 過去與中正大

學晶片中心（王進賢教授、葉經緯教授、林泰吉教授）

合作開發的語音轉換系統晶片及硬體裝置（圖 11），透

過軟硬體之團隊合作，完成了第一代構音障礙語音轉

換晶片的實踐，以實現邊緣運算落地應用的可能。

在過去的研究中，SBP Lab 團隊首先採用了「信號轉換」的概念來開發一個構音障礙語音轉換系統，名為 dysarthria voice conversion 3.1 (DVC 3.1)(見圖 12)。該架構主要是利用預先訓練的聲學模型來提取語音後驗概率(Phonetic Posterior Grams, PPG)，將構音異常患者的語音轉換為音素機率與時間的特徵矩陣。接著，使用門控卷積神經網路(gated Convolutional Neural Network, GCNN)將 PPG 轉換成正常語者的語音。這個架構首先利用 PPG 的特性，有效補足患者缺失的輔音、去除語速不一致等語音變異性的問題。同時，結合具有注意力機制的 GCNN 架構，能夠有效地捕捉音素之間的前後關係，以產生自然而清晰易懂的語音。此外，經過臨床實驗證明，DVC 3.1 技術能有效幫助構音障礙患者提升語音清晰度，並且比當時的基準系統有更佳的轉換效果。該成果也於 2022 年在《Computer Methods and Programs in Biomedicine》國際期刊上發表(主代表著作 5)。

雖然 DVC 3.1 系統已經能夠有效地協助患者提升語音清晰度和溝通效率，但是該系統仍還有進步的空間。例如，該系統所使用的 PPG 特徵高度依賴預訓練的聲學模型的準確率，而建立一個強健的聲學模型需要大量的語音數據。另一方面，基於深度學習技術的轉換模型同樣需要大量的患者語音以及其成對的目標語者語音進行訓練，才能使模型具有一定程度的轉換能力。這使得 DVC3.1 系統需要收錄大量的患者語音以確保其轉換效益。有鑑於此挑戰，SBP Lab 更進一步的發展以語音辨識系統為基礎之 DVC 4.0(圖 13)系統，其概念主要為透過語音辨識系統的「語言模型」的優勢來提升系統面臨患者語音變異度與不清晰之影響，試圖提升系統在臨床上的應用效益。於研究結果也發現，DVC 4.0 系統透過「語言模型」的幫助下，的確於較為挑戰的應用情況下能更顯著的提供語音轉換效益，進而幫助患者在應用此系統下有更佳的效果。而此成果也於 2023 年受到

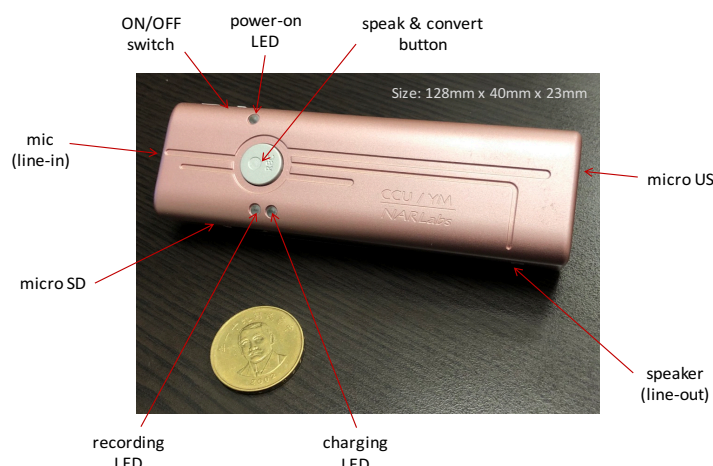


圖 11:本團隊與中正大學晶片中心合作開發之語音轉換系統晶片及硬體裝置。

Biomedical Signal Processing and Control 國際期刊肯定予以刊登(主代表著作 4)。

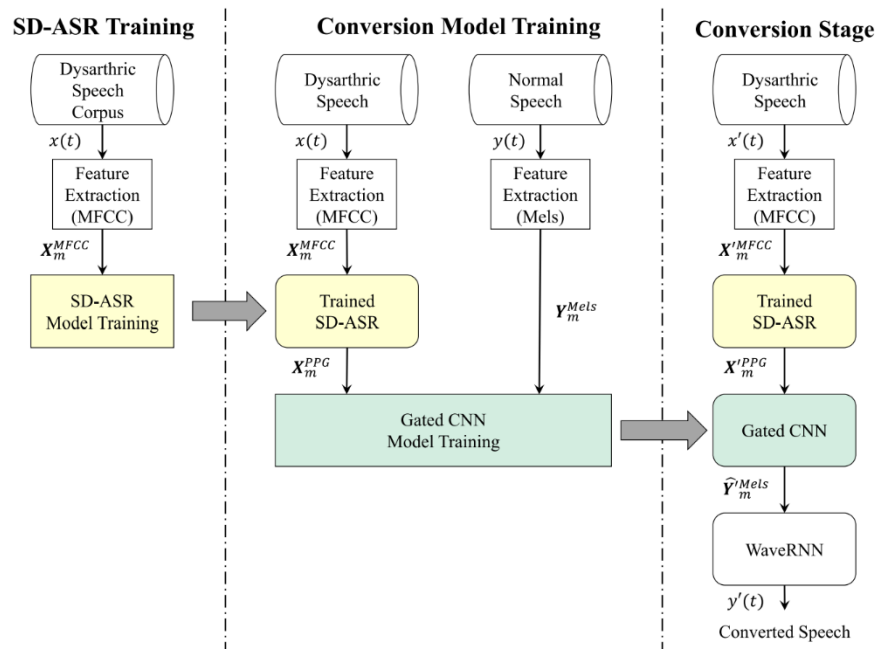


圖 12: SBP Lab 團隊提出之 DVC 3.1 音訊處理系統(主代表著作 5)。

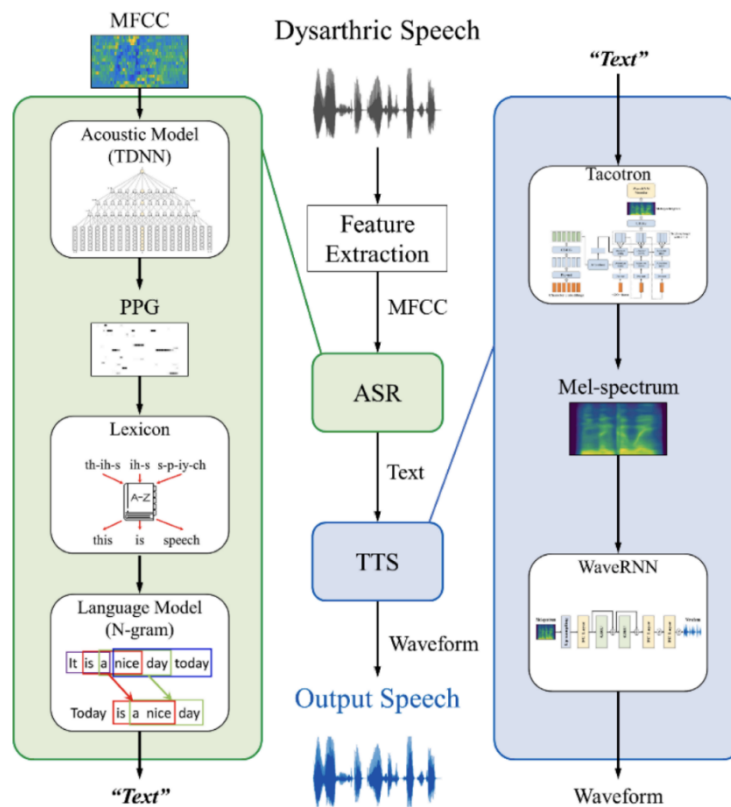


圖 13: SBP Lab 團隊提出之 DVC 4.0 音訊處理系統(主代表著作 4)。

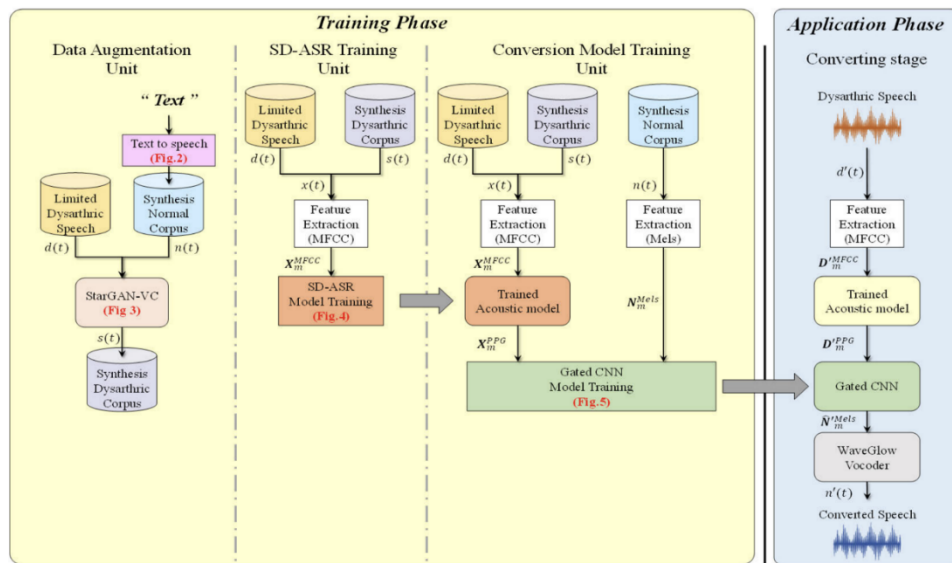


圖 14: SBP Lab 團隊提出之 DVC 3.1 plus 音訊處理系統(主代表著作 1)。

雖然 DVC 4.0 在更貼近臨床應用時已有最佳的轉換效益，但此 DVC 4.0 系統仍需要大量患者的語音才能不斷提升系統的辨識效率。然而，構音異常患者更容易且更頻繁地感到語音疲勞，且越嚴重的患者越容易有此類現象[19]，這使得構音異常患者在實際應用此套系統時面臨巨大的錄音負擔，並可能導致使用意願降低。有鑑於此，SBP Lab 團隊近期再提出了一個以 Text-to-Speech (TTS)和 Star-GAN 多對多轉換 Vocoder 的「資料擴增系統」來克服訓練資料不足之問題，稱 DVC 3.1 plus(圖 14)。該子系統可以使用少量的患者語句作為樣本進行訓練，生成大量的類構音異常患者語音，以幫助 DVC 3.1 系統有效建立構音異常患者語音轉換模型，並進一步的提升轉換準確性。此外，我們也透過一系列的研究證實了該架構相較於原系統能夠節省 93%的資料量(約節省 18.8 小時的語音)，並實證了類患者資料擴增技術生成類患者語音代替真實患者語音的可行性，而此成果也於近期發表在 IEEE Transactions on Neural Systems and Rehabilitation Engineering 國際期刊中(主代表著作 1)。近期，SBP Lab 也更進一步的與合作多年的宇康生科合作，將提出的 DVC 3.1 plus 系統中的「資料擴增系統」整合至宇康生科實踐出商化的 AiSpeak 產品中(<https://www.facebook.com/reel/847127216791719>)，以幫助構音障礙患者未來不需要大量的錄製語音即可以讓 AiSpeak 有良好的語音辨識能力。而上述多年期的研究成果除了在論文的發表、與產業互動並商品化實踐之成績外，SBP Lab 也透過這些主題來幫助培育陽明交大醫工系學生透過不同方式的學習，以了解智慧醫材與轉具的研究方法，而這些過程也的確看到培育的學生們已在多次的國際競賽上獲得許多的學習經驗並取得好成績，舉例來說，這些 DVC 系統就在過去分別獲得：2020 年美律電聲論文獎(銅質獎;獎金 10 萬元)、2021 年美律電聲論文獎(銅質獎;獎金 15 萬元)、2022 年全國醫學工程創意競賽(新創獎-博而美

使命獎(獎金 3 萬元)、2022 年 gSIC-TW 金牌獎，且於近期於泰國所辦的「2023 i-CREATE 全球學生創新輔具競賽」中獲得金牌獎(獎金美金 1,500 元)³等成績。而這些成果的例子也是 SBP Lab 過去不斷的努力的方法，期望能透過不同的競賽挑戰機會把「研究」與「培育人才」整合，來不斷提升台灣於此主題上的研發能量。

除了針對構音障礙患者說出不清晰的語音進行處理外，SBP Lab 過去與亞東醫院喉科合作，透過藍芽耳機與深度學習技術研發出了「發聲監控系統(見圖 15)」，用於對嗓音異常(如哨聲)患者進行即時性嗓音用量偵測。當系統偵測到患者在日常生活中說話的語速及音量過高時，將即時透過提示音提醒患者放慢語速及減少語音音量進行溝通，以協助患者提升治療效果。此系統經過臨床效益驗證後，在實際應用中能夠準確評估嗓音發聲量，這一成果也於 2023 年獲得了 IEEE Systems Journal 國際期刊的肯定並予以刊登(主代表著作 3)。此外，該技術也於近年獲得了中華民國專利，並已進行技術轉移以進行產品開發(參考著作 14)。

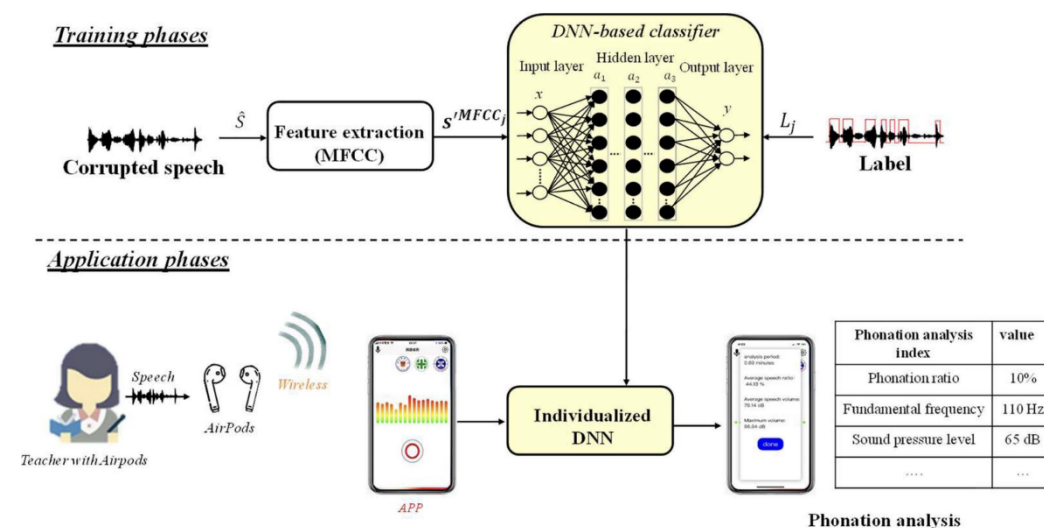


圖 15: SBP Lab 團隊提出之「發聲監控系統」(主代表著作 3)。

基於上述技術之基礎，SBP Lab 再進一步的改良過去之系統並透過以 self-attention 機制之深度學習技術(圖 16)來進一步的透過患者發出的嗓音資訊來辨識其是否有聲帶上的疾病，藉此檢測患者的聲帶健康情況。於研究結果發現，此方法在世界公開之嗓音辨識資料庫中可超過當時 baseline 系統之表現，並於 2023 年受到 Journal of Voice 國際期刊肯定予以刊登(參考著作 3)。隨後，SBP Lab 再依據上述基礎及過去應用深度學習技術於聲音事件偵測之研究經驗，進一步的與易祿達公司進行一年期之產學合作案，並於一年後將研發成果再次轉譯成實際的產品(圖 17)，且成立一家新創公司「聲智科技股份有限公司」。圖 17 所採用之技術主要是應用聲音特徵與深度學習技術之整合，進而透過深度學習模型的

³ <https://www.nycu.edu.tw/nycu/ch/app/news/view?module=headnews&id=2994&serno=9076c90b-2bc2-4881-93d5-aaae8c6931ec>

辨識優勢來進行異音偵測(例如:工具機生產良率控制、馬達生產品質偵控...等)，以幫助工業界於智慧製造上提供一個快速且低設備成本之方法，進而提升台灣智慧製造的優勢。此外，我們近期也將這項技術應用於生理信號異音偵測的情況，例如:(1)血管氣泡異音辨識及(2)洗腎瘻管狹窄度判斷等部份。而透過初步的研究成果看來，SBP Lab 團隊所提出之方法於上述二項醫學應用之測試資料測試下，其均能達到高於 98%以上的辨識效益，此也表示 SBP Lab 未來此項技術應有潛力能更進一步的導入醫學信號上進行應用。

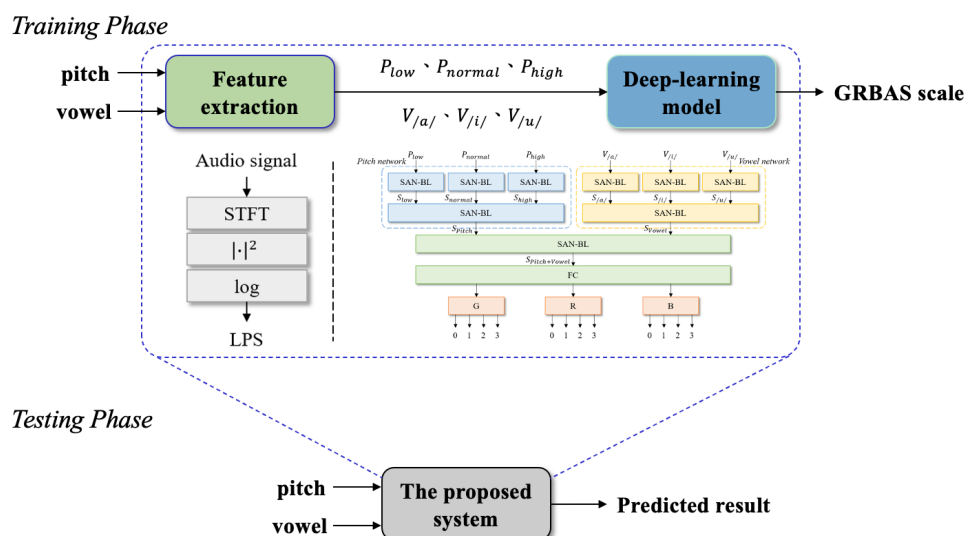


圖 16: SBP Lab 團隊提出之具備 self-attention 機制之聲帶疾病檢測系統。(參考著作 3)

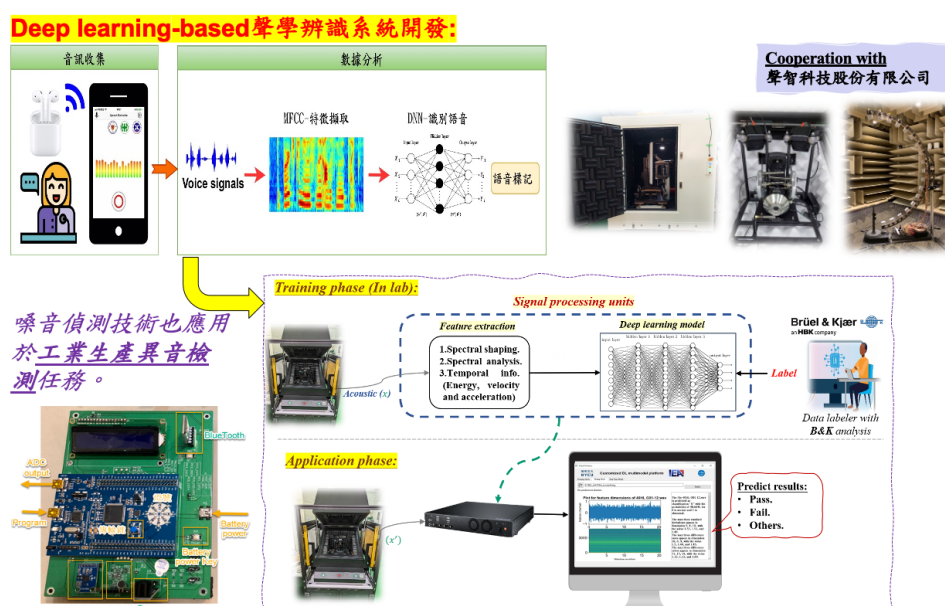


圖 17: SBP Lab 團隊基於過去聲音事件偵測演算法及電子電路設計經驗之研究成果，進一步與產業合作提出的「基於深度學習的聲學辨識系統」。該技術促成

了台灣一家新創公司的成立，名為「聲智科技股份有限公司」。

總結近年的論文和研究方向，我們主要利用人工智慧和各種感測技術(包括麥克風和光學)開發語音信號相關的演算法，以進一步提升「聽覺」和「溝通」輔具在臨床上的應用效益。在過去的研究中，我們已成功將這些核心技術應用於App或硬體晶片中，為患者提供更便捷的使用介面，從而提高了他們的生活品質。基於這些研究的基礎，我們將繼續不斷改進相關技術，並將其應用於更多醫療器材的設計中，特別是在與聲學相關的信號處理方面。基於我們的研究專長，我們將持續致力於智慧醫療器材的研發，以進一步提升患者的治療效果，造福更多需要幫助的人們。

A. 主代表著作				
1	Improving the Efficiency of Dysarthria Voice Conversion System based on Data Augmentation	2023. 11	IEEE Transactions on Neural Systems & Rehabilitation Engineering, 2023:31:4613-4623 (IF= 4.9; 排名:4/68)	Wei-Zhong Zheng, Ji-Yan Han, Chen-Yu Chen, Yuh-Jer Chang, and <u>Ying-Hui Lai*</u>
2	Optical Microphone-based Speech Reconstruction System with Deep Learning for Individuals with Hearing Loss	2023. 11	IEEE Transactions on Biomedical Engineering 70(12):3330-3341 (IF= 4.6; 排名: 34/96)	Yu-Min Lin, Ji-Yan Han, Cheng-Hung Lin and <u>Ying-Hui Lai*</u>
3	Ambulatory phonation monitoring using wireless headphones with deep learning technology	2023. 08	IEEE Systems Journal,17,3, 4752 - 4762 (IF=4.4;排名:26/86)	Ji-Yan Han, Chi-Te Wang*, Jia-Hui Li, and <u>Ying-Hui Lai*</u> (Co-corresponding author)
4	Comparing the performance of classic voice-driven assistive systems for dysarthric speech	2023. 03	Biomedical Signal Processing and Control,81, 104447 (IF= 5.1; 排名:- 26/96)	Wei-Zhong Zheng, Ji-Yan Han, Hsiu-Lien Cheng, Wei-Chung Chu, Ko-Chiang Chen and <u>Ying-Hui Lai*</u>

5	Phonetic posteriorgram-based voice conversion system to improve speech intelligibility of dysarthric patients	2022.03	Computer Methods and Programs in Biomedicine, 215, 106602. (IF= 6.1; 排名:15/111)	Wei-Zhong Zheng, Ji-Yan Han, Chen-Kai Lee, Yu-Yi Lin, Shu-Han Chang, and <u>Ying-Hui Lai*</u>
B. 參考著作				
1	Enhancing music recognition using deep learning-powered source separation technology for cochlear implant users	2024.02	The Journal of the Acoustical Society of America, 155(3):1694-1703 (IF=2.4;排名:8/27)	Yuh-Jer Chang, Ji-Yan Han, Wei-Chung Chu, Lieber Po-Hung Li, and <u>Ying-Hui Lai*</u>
2	Evaluating the hearing screening effectiveness of active noise cancellation technology for young adults: A pilot study	2023.01	Journal of the Chinese Medical Association, 86(1):105-112 (IF=3.0; 排名:78/169)	Hsiu-Lien Cheng, Ji-Yan Han, Yuan-Chia Chuc, Yen-Fu Cheng, Chia-Mei Lin, Ming-Chang Chiang, Shang-Liang Wu, <u>Ying-Hui Lai*</u> , Wen-Huei Liao* (Co-corresponding author)
3	Enhancing the performance of pathological voice quality assessment system through the attention-mechanism based neural network	2023.01	Journal of voice (IF=2.20; 排名:12/27)	Ji-Yan Han, Ching-Ju Hsiao, Wei-Zhong Zheng, Ko-Cheng Weng, Guan-Min Ho, Chia-Yuan Chang, Chi-Te Wang, Shih-Hau Fang, and <u>Ying-Hui Lai*</u>
4	Prediction of Visual Impairment in Epiretinal Membrane and Feature Analysis - A Deep Learning Approach Using Optical Coherence Tomography	2023.01	Asia-Pacific Journal of Ophthalmology, 12(1):21-28 (IF= 4.3; 排名:8/62)	Yun Hsia, Yu-Yi Lin, Bo-Sin Wang, Chung-Yen Su, <u>Ying-Hui Lai*</u> , and Yi-Ting Hsieh* (Co-corresponding author)
5	Smartphone-Bundled Earphones as Personal Sound Amplification Products: Laboratory and Clinical Performance Assessments among Individuals with Hearing Loss	2022.12	iScience,25,12, 105436 (IF=5.8; 排名: 15/73)	Heng-Yu Haley Lin [#] , Hoi-Shan Lai [#] , Chii-Yuan Huang, Chih-Hao Chen, Shang-Liang Wu, Yuan-Chia Chu, Yu-Fu Chen,

				<u>Ying-Hui Lai</u> [#] , Yen-Fu Cheng* (Co-first author, 共同第一作者)
6	Objective signal analysis approach to investigate the feasibility of active noise cancellation technology for hearing screening application	2022.09	Sensors (2022): 22(19), 7329 (IF= 3.9; 排名:19/63)	Hsiu-Lien Cheng, Ji-Yan Han, Wei-Zhong Zheng, Yen-Fu Cheng, Yuan-Chia Chu, Chia-Mei Lin, Ming-Chang Chiang, Wen-Huei Liao *, <u>Ying-Hui Lai</u> * (Co-corresponding author)
7	Study of optical-based speech acquisition system using vibration signals from speakers' medical masks	2022.05	JASA Express Letters, 055202 (2022). (IF= 1.0; 排名: While journals indexed in AHCI and ESCI are receiving a JIF for the first time in June 2023, they will not receive ranks, quartiles, or percentiles until the release of 2023 data in June 2024)	Yu-Min Lin, Ji-Yan Han, Wei-Zhong Zheng, Yi-Chieh Lin, Cheng-Hung Lin, and <u>Ying-Hui Lai</u> *
8	Comparison of personal sound amplification products and conventional hearing aids for patients with hearing loss: A systematic review with meta-analysis	2022.04	EClinicalMedicine, 46:101378 (IF=15.1; 排名: 14/169)	Chih-Hao Chen, Chii-Yuan Huang, Hsiu-Lien Cheng, Heng-Yu Haley Lin, Yuan-Chia Chu, Chun-Yu Chang, <u>Ying-Hui Lai</u> , Mao-Che Wang, Yen-Fu Cheng*
9	Human Motion Tracking Using 3D Image Features with a Long Short-Term Memory Mechanism Model—An Example of Forward Reaching	2021.12	Sensors 22.1 (2022): 292 (IF= 3.9; 排名:19/63)	Kai-Yu Chen, Li-Wei Chou, Hui-Min Lee, Shuenn-Tsong Young, Cheng-Hung Lin, Yi-Shu Zhou, Shih-Tsang Tang, and <u>Ying-Hui Lai</u> *
10	Improved Environment-Aware—	2021.	Journal of	Lieber Po-Hung Li, Ji-

	Based Noise Reduction System for Cochlear Implant Users Based on a Knowledge Transfer Approach: Development and Usability Study	10	Medical Internet Research, 23(10), e25460, 2021 (IF= 7.4; 排名:3/106)	Yan Han, Wei-Zhong Zheng, Ren-Jie Huang, <u>Ying-Hui Lai*</u>
11	以 AI 技術精進藍芽音訊晶片於「聽覺補償」與「主動降噪」應用效益研究	2024.0 2	技轉案	賴穎暉
12	光學麥克風音訊即時信號分析與聆聽軟體系統 (本校合約號: YC-Y23T008)	2023.0 6	技轉案	賴穎暉
13	以超低延遲補償技術提升聽覺輔具效益-App 平台實踐與效益驗證 (本校合約號:YC-Y23C003)	2023.0 4	技轉案	賴穎暉
14	發聲監控方法及系統 (本校合約號: YC-Y22T017)	2022.1 2	專利技轉 -中華民國專利 (I749663)	賴穎暉、王榮德
15	構音異常語料擴增方法及系統、語音辨識平台及構音異常輔助裝置	2022.1 0	中華民國專利 (I780738)	賴穎暉、何冠旻、張嘉原
16	提升構音患者語音轉換效益之系統及方法	2022.0 6	中華民國專利 (I766575)	賴穎暉、李沛群、李振愷
17	聽覺補償 App 平台實踐與聲電特性驗證 (1)(計畫代碼「111A40073Y」)	2022.0 4	技轉案	賴穎暉
18	聽覺補償裝置及具有該聽覺補償裝置之聽力設備	2022.1 2	中華民國新型專利 (M635174)	葉明翰、賴穎暉
19	具有真耳測量的自調式聽覺補償裝置及具有自調式聽覺補償裝置之聽力設備	2023.0 1	中華民國新型專利 (M636702)	葉明翰、賴穎暉
20	耳機裝置及聽力設備	2023.0 9	中華民國新型專利(M645890)	葉明翰、賴穎暉
21	聽覺補償裝置及聽覺補償方法	2023/1 2	中華民國專利 (I827009)	葉明翰、賴穎暉

Key references:

- [1] WHO. "Deafness and hearing loss." <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> (accessed.

- [2] C. N. Hansen, *Understanding active noise cancellation*. CRC Press, 1999.
- [3] E. O. Lopez-Caudana, "Active Noise Cancellation: The Unwanted Signal and the Hybrid Solution," *Adaptive Filtering Applications, Dr. Lino Garcia (Ed.)*, pp. 49-84, 2011.
- [4] S. M. Kuo, S. Mitra, and W.-S. Gan, "Active noise control system for headphone applications," *Control Systems Technology, IEEE Transactions on*, vol. 14, no. 2, pp. 331-335, 2006.
- [5] I. Brons, R. Houben, and W. A. Dreschler, "Effects of noise reduction on speech intelligibility, perceived listening effort, and personal preference in hearing-impaired listeners," *Trends in hearing*, vol. 18, p. 2331216514553924, 2014.
- [6] P. Lutzmann, B. Göhler, C. A. Hill, and F. D. van Putten, "Laser vibration sensing at Fraunhofer IOSB: review and applications," *Optical Engineering*, vol. 56, no. 3, p. 031215, 2016.
- [7] Y. Li, S. Meersman, and R. Baets, "Realization of fiber-based laser Doppler vibrometer with serrodyne frequency shifting," *Applied optics*, vol. 50, no. 17, pp. 2809-2814, 2011.
- [8] S. Rothberg *et al.*, "An international review of laser Doppler vibrometry: Making light work of vibration measurement," *Optics and Lasers in Engineering*, vol. 99, pp. 11-22, 2017.
- [9] T. Fukumori *et al.*, "Optical laser microphone for human-robot interaction: speech recognition in extremely noisy service environments," *Advanced Robotics*, vol. 36, no. 5-6, pp. 304-317, 2022.
- [10] S.-A. Selouani, M. S. Yakoub, and D. O'Shaughnessy, "Alternative speech communication system for persons with severe speech disorders," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, p. 6, 2009.
- [11] S. J. Stoeckli, M. Guidicelli, A. Schneider, A. Huber, and S. Schmid, "Quality of life after treatment for early laryngeal carcinoma," *European archives of oto-rhino-laryngology*, vol. 258, no. 2, pp. 96-99, 2001.
- [12] J. R. Duffy, *Motor Speech Disorders-E-Book: Substrates, Differential Diagnosis, and Management*. Elsevier Health Sciences, 2013.
- [13] S. Pinto *et al.*, "Dysarthria in individuals with Parkinson's disease: a protocol for a binational, cross-sectional, case-controlled study in French and European Portuguese (FraLusoPark)," *BMJ open*, vol. 6,

no. 11, p. e012885, 2016.

- [14] S. Körner *et al.*, "Speech therapy and communication device: impact on quality of life and mood in patients with amyotrophic lateral sclerosis," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 14, no. 1, pp. 20-25, 2013.
- [15] D. Crystal, *Clinical linguistics*. Springer Science & Business Media, 2013.
- [16] K. M. Yorkston. "Treatment Efficacy Summary." ASHA. <https://www.asha.org/uploadedFiles/public/TESDysarthria.pdf> (accessed).
- [17] H. C. Shane, S. Blackstone, G. Vanderheiden, M. Williams, and F. DeRuyter, "Using AAC technology to access the world," *Assistive technology*, vol. 24, no. 1, pp. 3-13, 2012.
- [18] A. Waller, "Telling tales: unlocking the potential of AAC technologies," *International journal of language & communication disorders*, vol. 54, no. 2, pp. 159-169, 2019.
- [19] V. Young and A. Mihailidis, "Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: A literature review," *Assistive Technology*, vol. 22, no. 2, pp. 99-112, 2010.