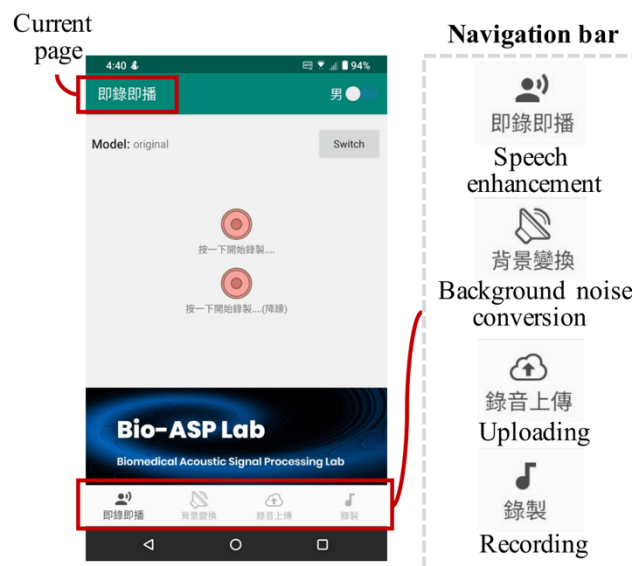


## 簡介

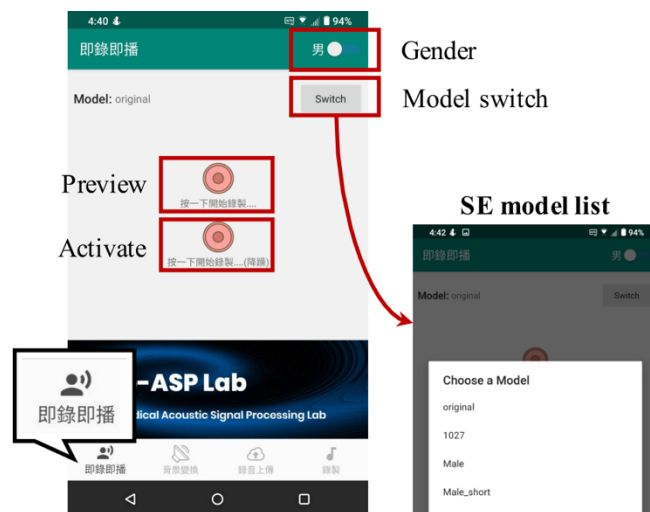
我們提出了一種基於深度學習的語音信號處理應用程式，CITISEN。CITISEN 可以實現語音增強(Speech Enhancement, SE)、模型適應(Model Adaptation, MA) 和背景噪聲轉換 (Background Noise Conversion, BNC) 三種功能。這三種功能使得 CITISEN 可以作為語音增強模型的利用和評估平台、靈活擴展使用在各種語音增強模型、並且調整模型以適應各種噪音環境和語者。為運行語音增強的功能，CITISEN 在雲端伺服器上下載已訓練好的語音增強模型，然後再使用這些模型有效地減少使用者所提供音訊的噪聲。當遇到包含未知語者或噪聲類型的嘈雜語音信號時，模型適應功能可以讓 CITISEN 有效地提高語音增強性能。為使用模型適應功能，使用者需上傳一些包含新語者或新噪聲的音頻至雲端伺服器。然後這些音頻會被用於調整已訓練好的語音增強模型。最後，為實現背景噪聲轉換功能，CITISEN 使用語音增強模型去除原始背景噪聲，然後將處理後的語音信號與新的背景噪聲混合。新穎的背景噪聲轉換功能可用於評估特定條件下的語音增強性能、掩蓋人們的足跡並作為娛樂用途。

## 用戶界面和操作方式



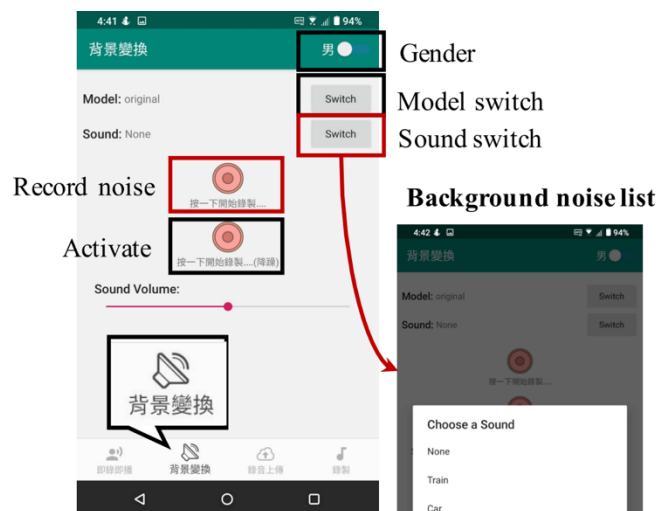
圖一、CITISEN 的四個主要頁面。CITISEN 應用程式有四個頁面，“即錄即播”(Speech enhancement)、“背景轉換”(Background noise conversion)、“錄音上傳”(uploading)、“錄製”(Recording)。每個頁面的當前頁面名稱和導航按鈕分別列在應用程式的左上角和底部。

## (1) 即錄即播 (Speech enhancement) 頁面



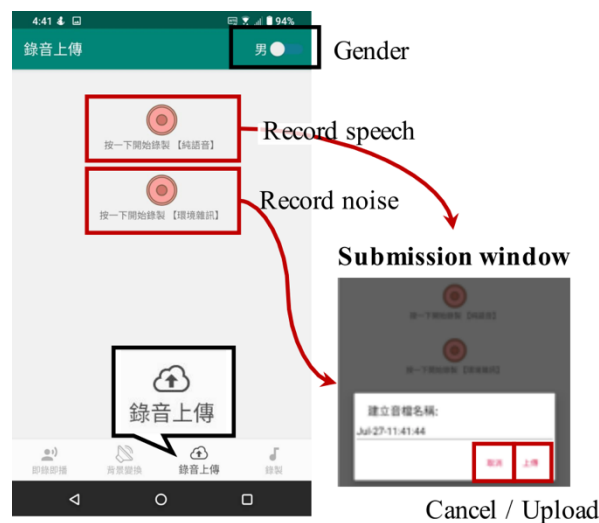
圖二、即錄即播頁面。右上角的“Gender”（性別）按鈕用於選擇模型適合處理的語音性別。按下“Model switch”（模型切換）按鈕，會彈出語音增強模型列表，用戶可以更欲使用之語音增強模型。按下“Preview”（預覽）按鈕後，用戶將聽到原本的即時錄音，按下“Activate”（運行）按鈕後，用戶將聽到增強後的錄音。

## (2) “背景轉換” (Background noise conversion) 頁面



圖三、背景轉換頁面。按下“Sound switch”（聲音開關）按鈕，會彈出背景噪音列表。按下“Record noise”（記錄噪音）按鈕後，用戶可以記錄並保存新的噪音信號。按下“Activate”（運行）按鈕後，用戶將聽到帶有指定背景噪音的增強即時錄音。其中“Gender”（性別）按鈕和“Model switch”（模型轉換）按鈕的功能與即錄即播頁面中的功能相同。

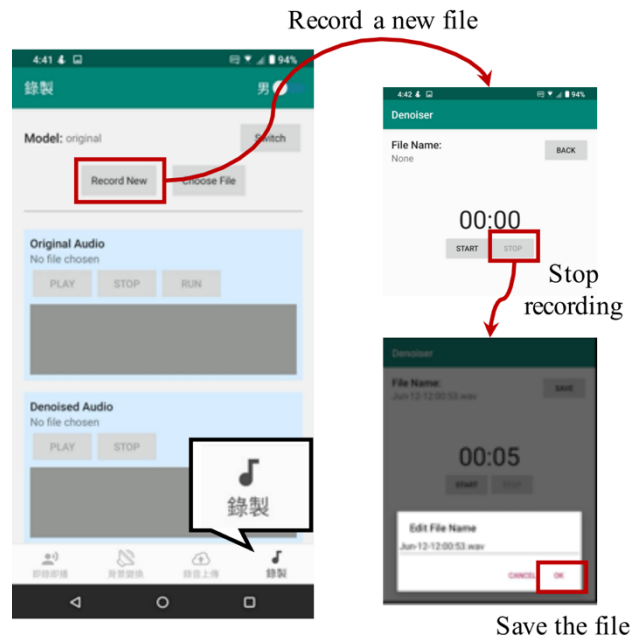
### (3)錄音上傳頁面 (Uploading)



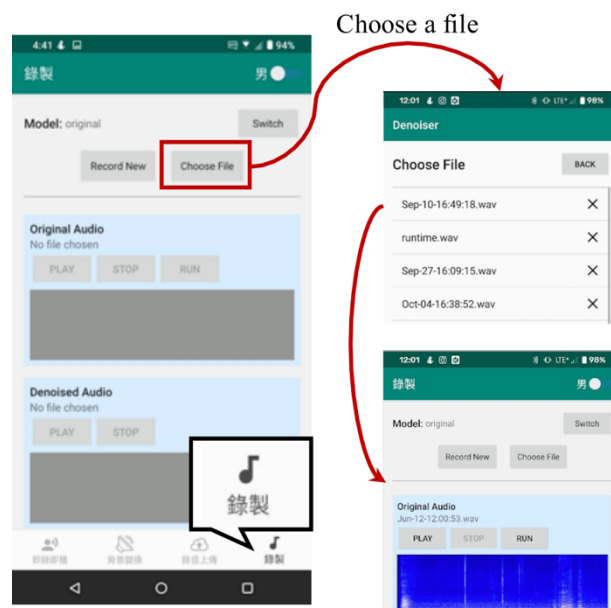
圖四、錄音上傳頁面。上傳頁面用於上傳模型適配功能所需要的數據音檔。由於 CITISEN 同時提供針對未知環境和未知語者做模型調整，因此有兩個文件上傳按鈕：“Record speech”（錄製語音）和“Record noise”（錄製噪聲）。使用者只需按相對應的按鈕即可開始錄製。再次按下按鈕完成錄製後，CITISEN 將彈出上傳音檔的頁面。接著使用者即可命名音檔並將檔案上傳到伺服器。

### (4) 錄製頁面 (Recording)

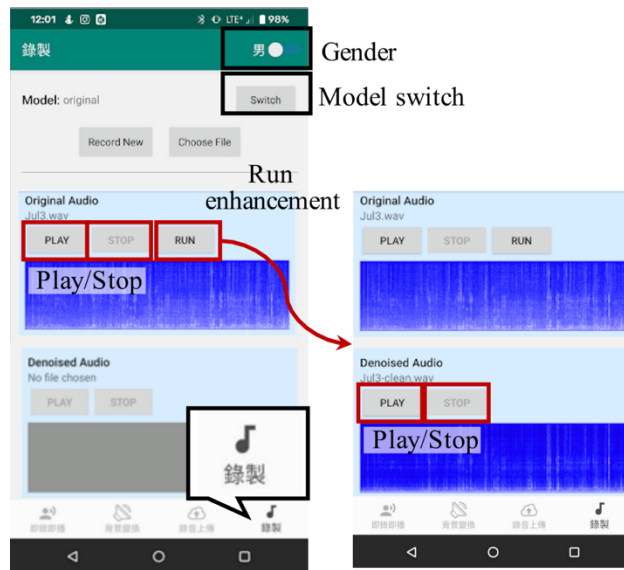
錄製頁面提供一般錄音和語音增強模型評估的功能。具體來說，在錄製頁面中，使用者可以錄製、回放語音信號並觀測使用語音增強後的結果。首先，使用者可以點擊“Record New”（錄製新檔案）前往處理頁面。再按下“Stop”（停止）按鈕完成錄音後，使用者可以命名並保存音檔。然後，使用者可以分別使用“Choose File”（選擇文件）、“Gender”（性別）和“Switch”（模型切換）按鈕選擇音檔和預使用的語音增強模型模式。最後，通過按下“Activate”（運行）按鈕生成增強的語音信號。使用者可藉由原始音檔和增強音檔的頻譜圖或直接播放音檔評估語音增強的效果。



圖五、錄音頁面一。按下“Record New”(錄製新檔案)按鈕後即可開始錄製。錄製完成後可在彈出的頁面命名音檔並保存該文件。



圖六、錄音頁面二。通過按下“Choose File”(選擇文件)按鈕，使用者可以在彈出的頁面中選擇想要測試的音檔。



圖七、錄音頁面三。使用者可以通過“Gender”（性別）和“Model switch”（模型切換）按鈕選擇語音增強模型的類型。此外，使用者可以利用此頁面評估語音增強的效果。