

利用遷移學習方法提升語音辨識HIV患者情緒之成效

Using transfer learning methods to improve the effectiveness of speech recognition in HIV patients' emotions

指導老師：林瑞豐 教授

學生：劉紫瑄 游騰捷 洪語佳

研究動機

近年情緒議題日益受到關注，世界衛生組織於2020年指出憂鬱症為需重視的疾病，而HIV患者更屬於**高風險族群**。為協助醫療人員了解患者情緒狀況，本研究以**語音情緒辨識技術**進行初步判定，期望能降低憂鬱與自殺風險。本研究結合桃園醫院HIV感染者語音資料與大型語料庫，並運用**遷移學習(Transfer Learning)**建構AI模型，以改善情緒辨識的準確性與實用性。



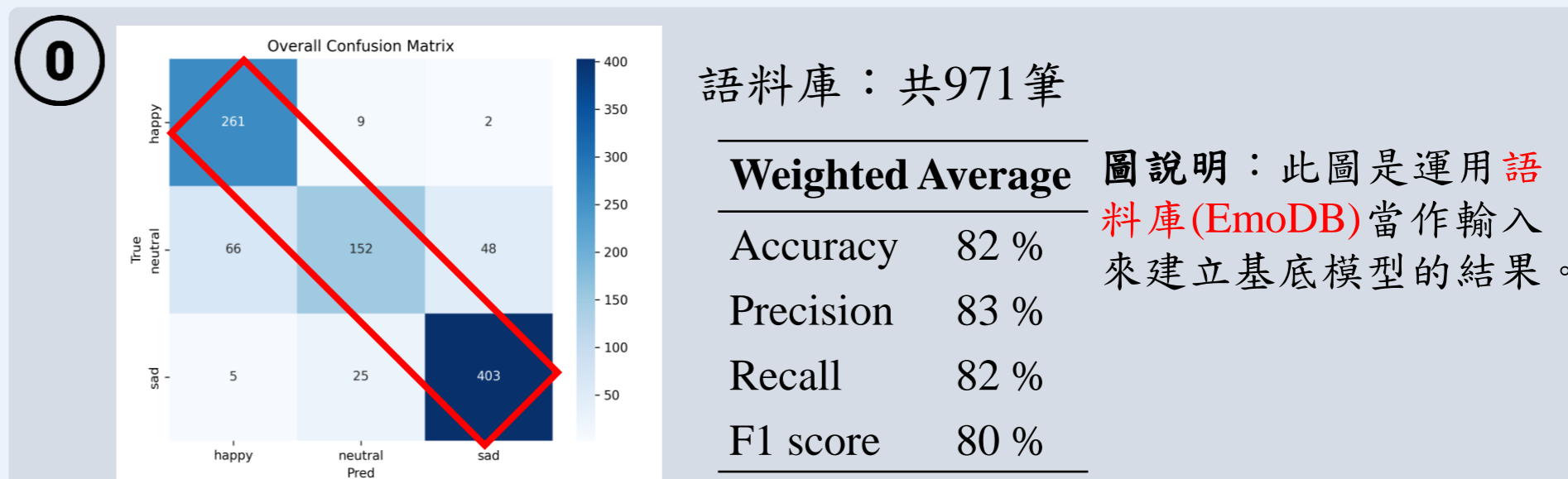
研究目的

- ▶ 利用聲音發展**辨別情緒狀況之工具**
- ▶ 以**遷移學習**方法提升語音情緒辨識之模型成效
- ▶ 提升語音情緒辨識HIV患者情緒之成效
- ▶ 能更快速幫助醫師初步**了解患者情緒狀況**

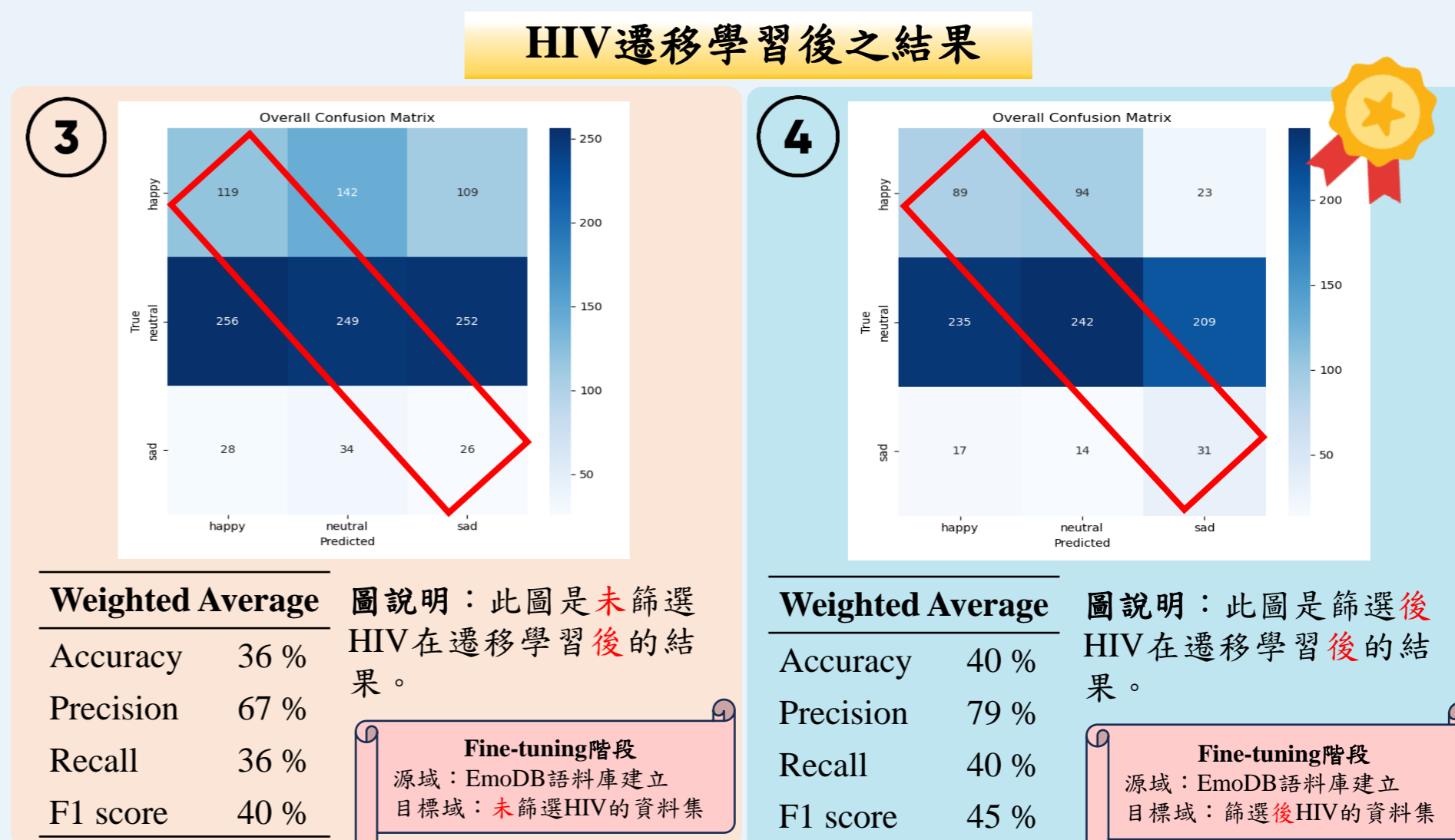
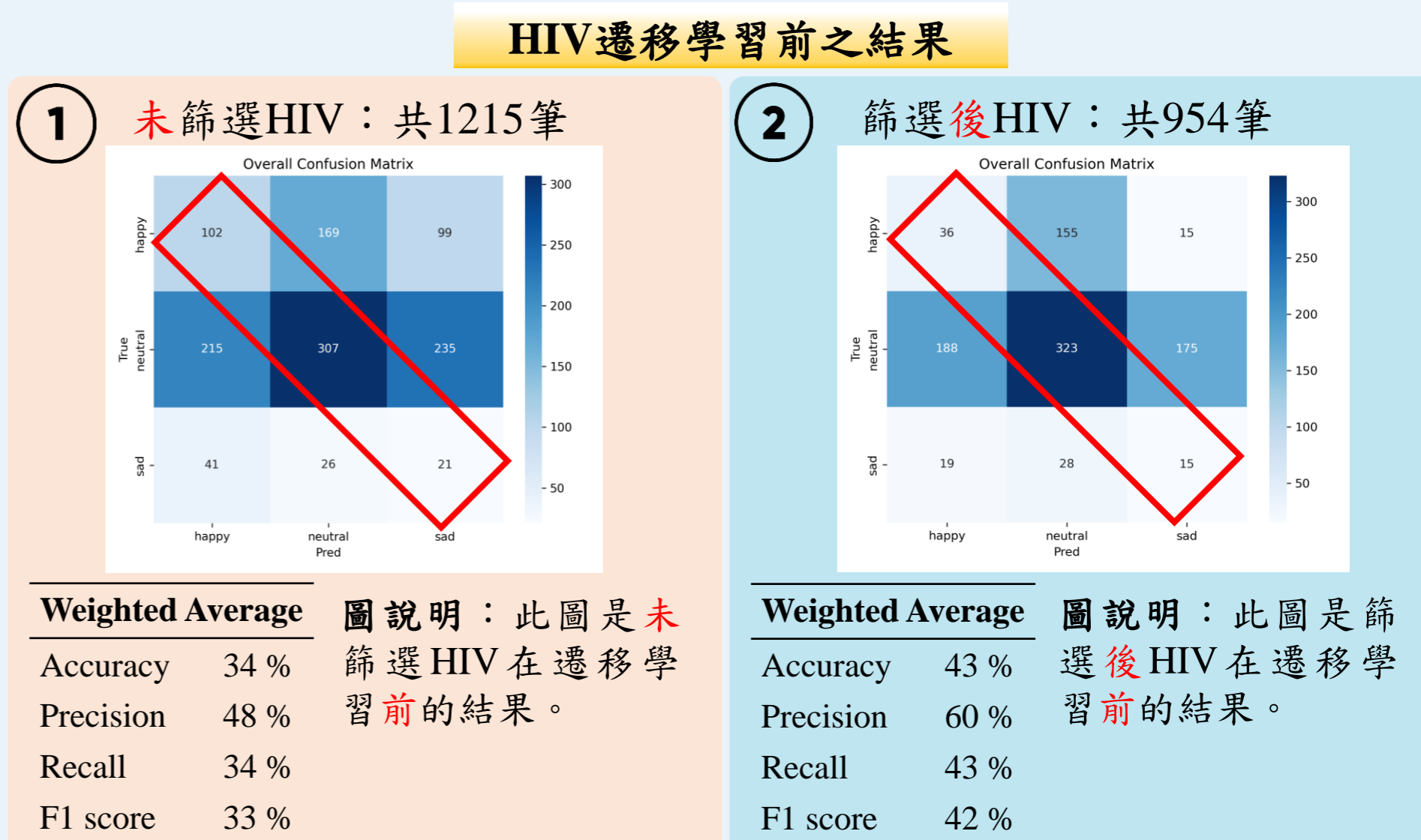


研究結果

✓ 語料庫(EmoDB)建立基底模型之結果



✓ HIV資料建模之結果



1. 結果比較—資料篩選之成效

資料篩選對模型表現有**明顯改善效果**，由上述可知，模型的績效指標較原始資料皆有略為提升，顯示**排除不良或異常樣本**能降低模型混淆，提升模型分類情緒的穩定度。

2. 結果比較—遷移學習之成效

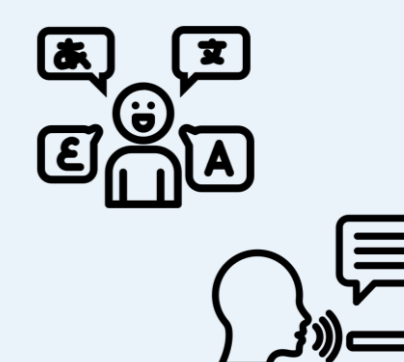
遷移學習能進一步**提升模型表現**，尤其是在**Precision及F1 score**的增幅最為明顯，代表遷移學習有助於讓模型更準確辨識各分類情緒。

綜合比較四種條件(原始資料/篩選資料×遷移/未遷移)，「**篩選後+遷移學習**」取得最高的分類績效，說明**良好的資料與適當的模型訓練策略**對提升最終結果都很重要。

結論與未來展望

最終結論：「**篩選後+遷移學習**」的成效最佳

- 未來可嘗試**增加語料庫的變異**
 - ✓ 找尋更多語言的語料庫，如中文的語料庫
- 將梅爾頻譜圖裁切的**overlap程度提升**
- 未來**增加受測者資料**
 - ✓ 蒐集更多HIV受測者的語音與情緒資訊，提升訓練的準確率

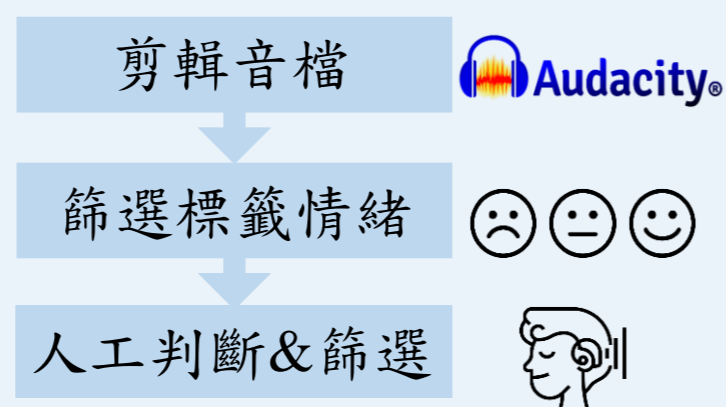


資料處理

✓ 資料說明

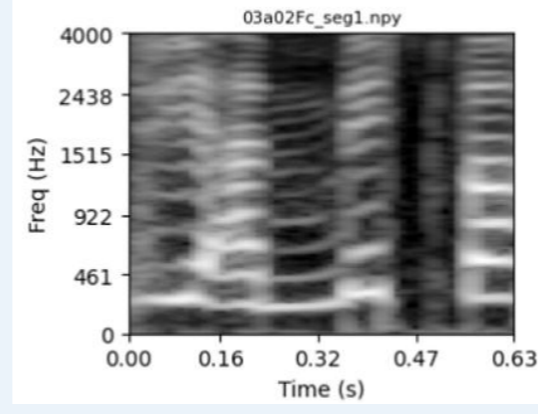
資料來源	桃園醫院	大型語料庫
資料名稱	HIV(中文)	EmoDB(德文)
音檔筆數	1090筆	535筆
參與人數	36位	10位

✓ 資料處理

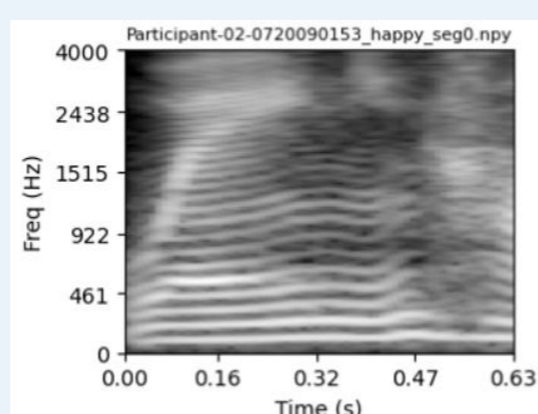
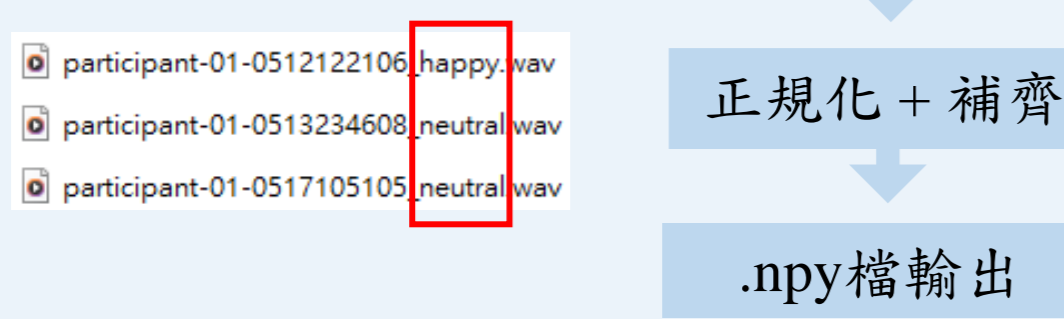


梅爾頻譜圖

✓ EmoDB音檔

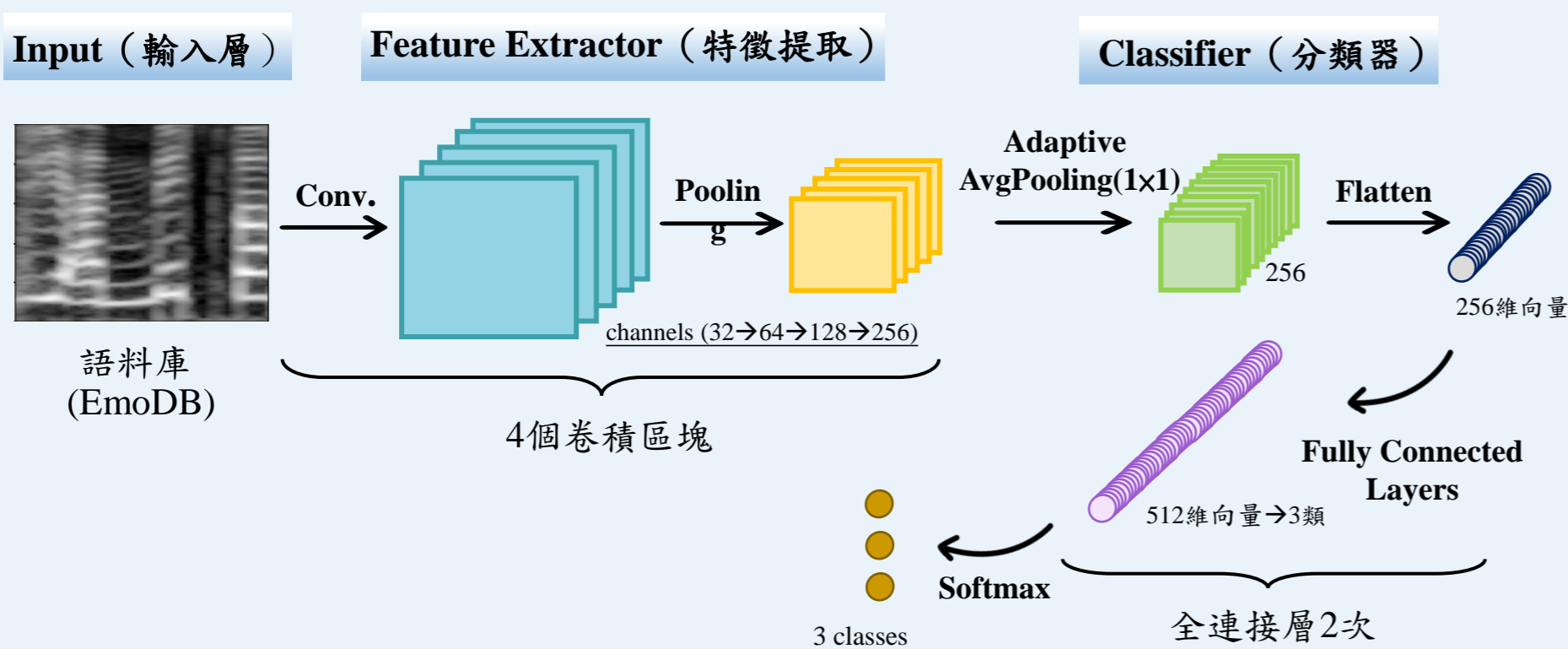


✓ HIV音檔



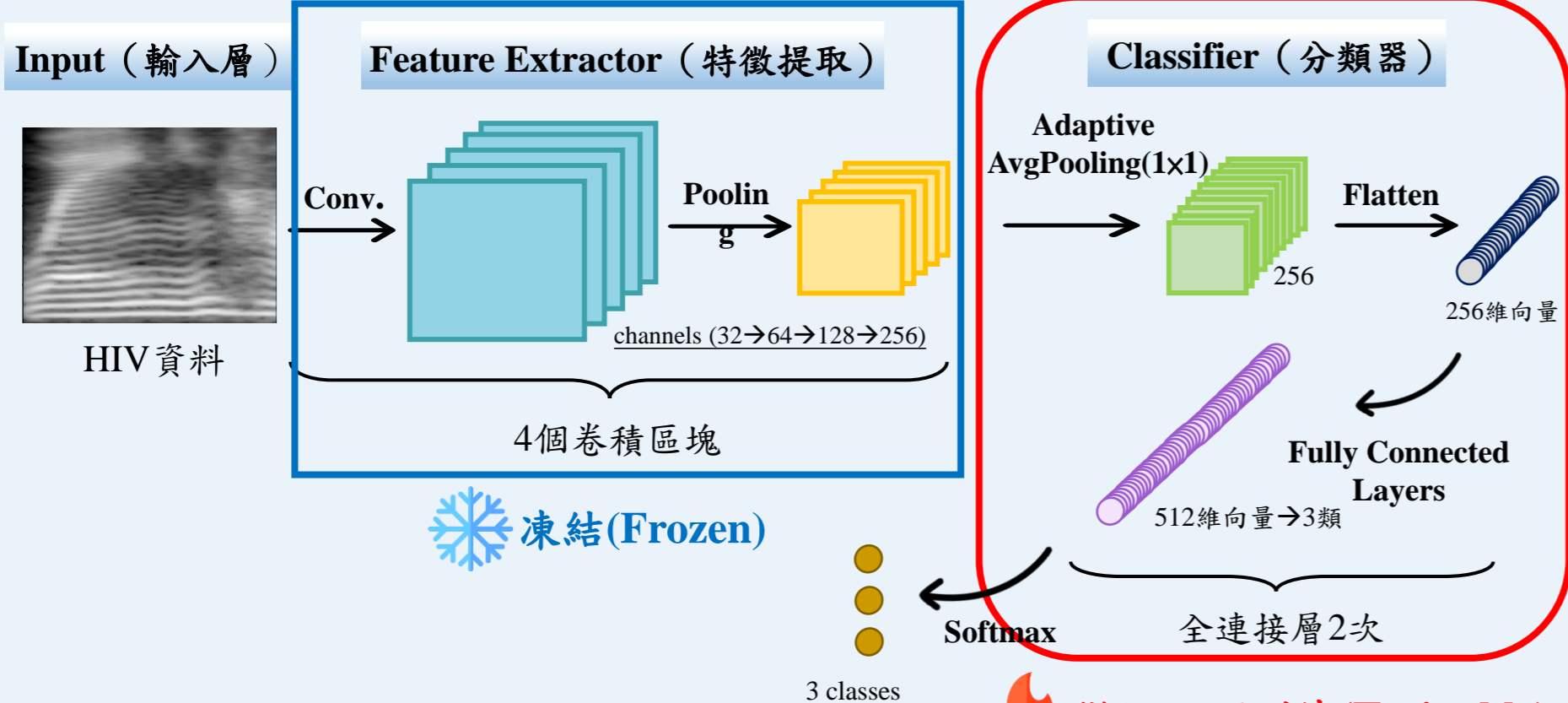
建基底模型

✓ 卷積神經網路(Convolutional Neural Network)之架構



遷移學習

✓ CNN架構—遷移學習(Classifier)模式



成效比較

✓ 比較資料篩選與遷移學習之成效

- All HIV音檔建模 HIV未篩選建模之結果 $\frac{1}{5}$
- Select HIV音檔建模 HIV篩選後建模之結果 $\frac{4}{5}$
- Before Transfer Learning HIV遷移學習前之結果 $\frac{1}{5}$
- After Transfer Learning HIV遷移學習後之結果 $\frac{4}{5}$