

國立臺北大學資訊工程學系114學年度專題 基於超聲波之蝙蝠物種辨識系統

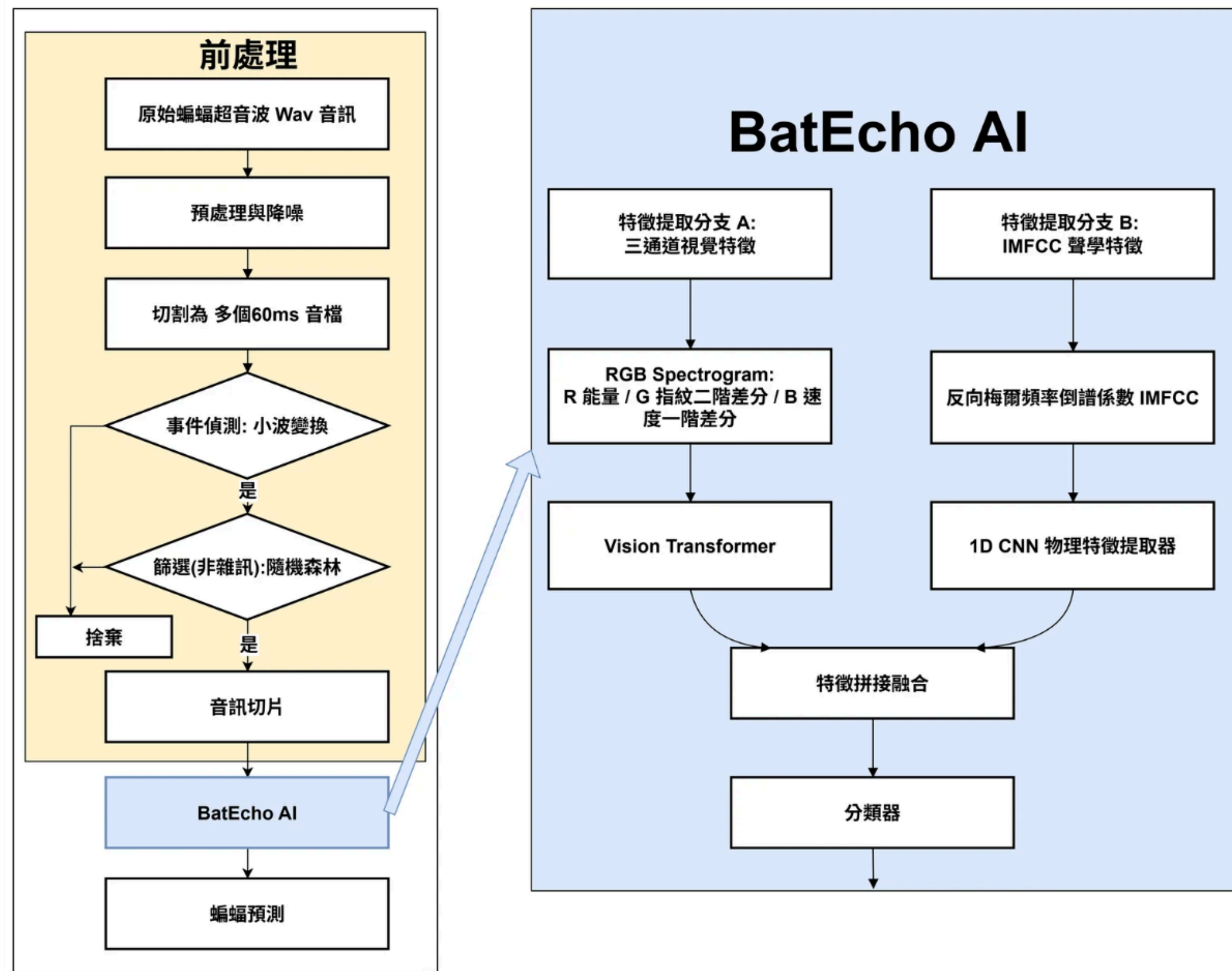
專題編號: PRJ-NTPUCSIE-114-14
成員: 陳立亞、蔡承祐、林子欽、曾柏碩

1. 簡介

蝙蝠在台灣生態系中占有重要地位與保育價值，然而因其夜行性及棲地隱密特性，使蝙蝠的相關調查不易進行。近年來，研究人員逐漸利用蝙蝠回聲定位叫聲進行物種辨識與生態調查，此方法能補足傳統捕捉方式難以記錄的物種資料，但由於傳統聲學監測方法高度依賴人工判讀、耗時，使大規模蝙蝠監測與多樣性研究面臨挑戰。

目前台灣尚未有一套完整的蝙蝠音訊監測系統，因此本專題以蝙蝠回聲定位聲訊為研究對象，開發一套基於超聲波音訊之蝙蝠物種自動辨識系統。系統以蝙蝠回聲定位聲訊為主要分析對象，整合音訊前處理、事件偵測、雜訊篩選與深度學習分類等流程，實現自動物種辨識，期望降低聲學資料分析門檻。

2. 系統框架



3. 前處理

在野外環境中錄製蝙蝠回聲定位聲訊時，音檔常混有背景音、低頻干擾與錄音設備雜訊。若直接進行切片與分類，模型可能受到非目標聲音影響。因此，本系統先以頻譜減法進行降噪，透過短時傅立葉轉換將音訊轉至頻率域，並利用音檔前段估計背景雜訊模板，再自原始頻譜中扣除，以降低穩定背景雜訊干擾。其概念可表示為：

$$\hat{S}(f, t) = \max(|X(f, t)| - N(f), 0)$$

$|X(f, t)|$ 表示原始音訊在頻率 f 與時間 t 下的頻譜幅度， $N(f)$ 表示估計出的背景雜訊頻譜， $\hat{S}(f, t)$ 則表示降噪後保留的頻譜成分。此外，因蝙蝠回聲定位聲訊主要位於高頻區域，系統移除 35 kHz 以下低頻成分，並抑制 35-50 kHz 區間背景干擾。

降噪後，系統以滑動視窗將音訊切割為 60 ms 片段，並以 20 ms 位移，使相鄰片段約有 66.7% 重疊，以降低叫聲落在邊界而被截斷的可能。其切片起始位置可表示為：

$$s_i = i \times L_{\text{shift}}$$

s_i 表示第 i 個片段的起始位置， L_{shift} 表示每次視窗位移長度。本研究設定片段長度為 60 ms，位移長度為 20 ms。完成切片後，系統使用 db4 小波轉換進行事件偵測，透過多尺度係數變化判斷片段是否可能包含蝙蝠叫聲。事件門檻可表示為：

$$T_j = \mu_j + k\sigma_j$$

T_j 表示第 j 個尺度的事件判斷門檻， μ_j 與 σ_j 分別表示該尺度小波係數的平均值與標準差， k 則為門檻調整參數。

為進一步排除昆蟲聲與環境雜訊，本研究再以 Random Forest 作為第二階段篩選方法，其多數決概念可表示為：

$$\hat{y} = \text{majority}(h_1(x), h_2(x), \dots, h_T(x))$$

x 表示音訊片段的聲學特徵向量， $h_T(x)$ 表示第 T 棵決策樹的分類結果， \hat{y} 表示多數決後的最終判斷結果。最終僅保留有效片段進入物種分類流程，以提升模型輸入品質與辨識穩定性。

4. 訓練模型

本專題主要模型框架分為兩條特徵分支，分別為三通道視覺特徵(簡稱分支A)、IMFCC聲學特徵(簡稱分支B)，並透過兩條分支特徵結合作為最終模型。

- 分支A: 為Vision Transformer 其中三通道分別為:
 - R通道: (強度能量): 代表聲紋強度。
 - G通道: (二階差分): 代表音頻的突變輪廓。
 - B通道: (一階導數): 捕捉聲紋隨時間滑動或漸變的動態行為。
- 分支B: 對輸入訊號進行短時傅立葉轉換計算功率譜，通過一組反梅爾濾波器組後取得各頻帶能量[1]，再取對數進行離散餘弦變換，最終輸出到一維CNN產出聲學特徵向量[2]。

$$E_m(t) = \sum_{f=1}^F H_m^{IMel}(f) |X(f, t)|^2 \quad [1]$$

$$C_k(t) = DCT(\log(E_m(t) + \epsilon)), \quad k = 1, \dots, 10 \quad [2]$$

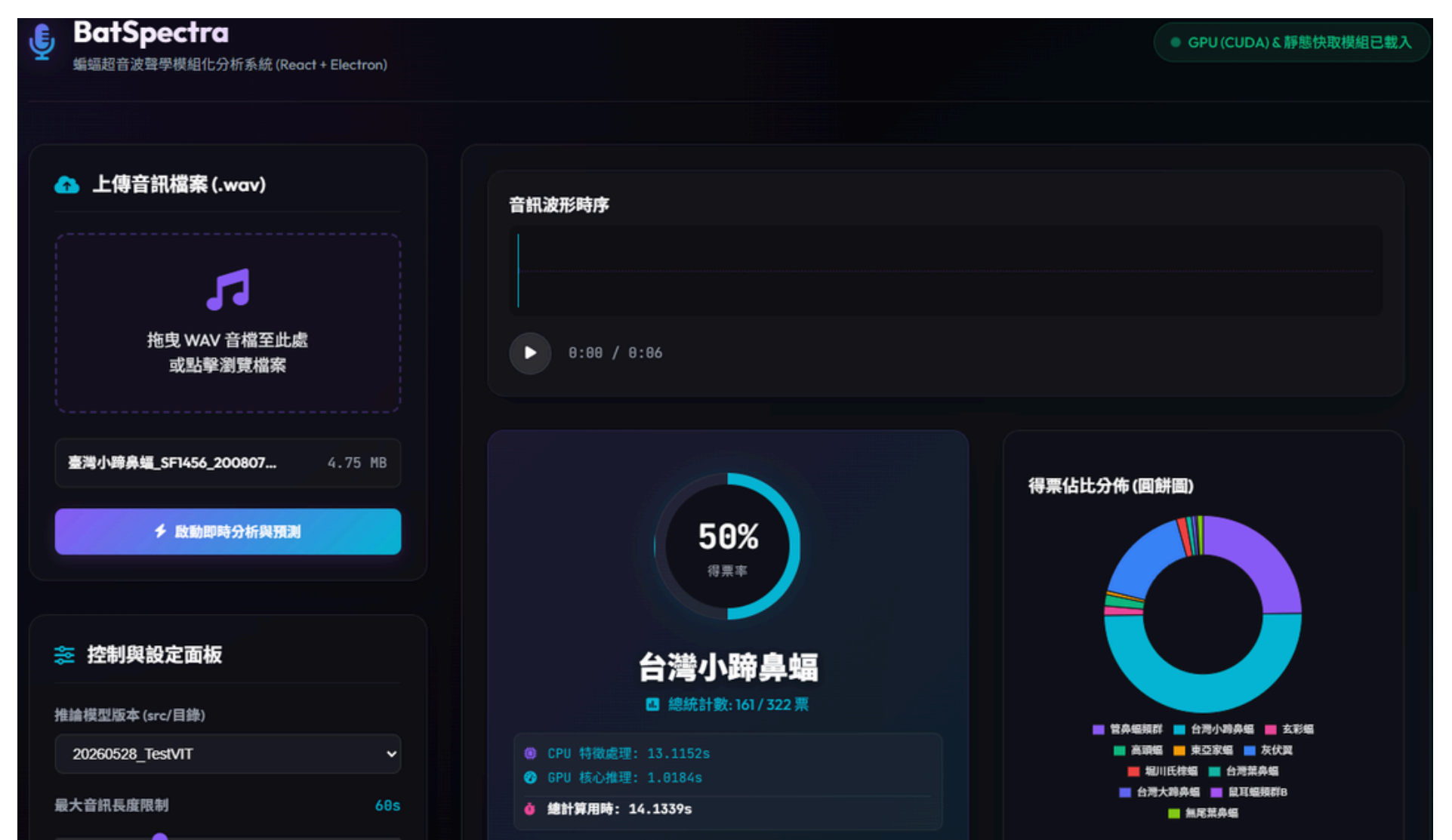
5. 結果

各蝙蝠類別之分類效能評估結果

Class	Precision	F1 Score
台灣大蹄鼻蝠	1.0000	1.0000
台灣小蹄鼻蝠	0.5714	0.7273
台灣葉鼻蝠	1.0000	0.8750
東亞摺翅蝠	1.0000	0.4444
玄彩蝠	1.0000	1.0000
管鼻蝠類群	0.7500	0.8182
鼠耳蝠類群A	1.0000	0.4000
鼠耳蝠類群B	0.2222	0.3077
Accuracy	--	0.7174
Weighted Average	0.8408	0.7154

盲測結果
平均準確度為

84%



本專題的使用者介面

6. 結語

本專題成功開發出一套基於超聲波音訊之蝙蝠物種自動辨識系統，解決了傳統仰賴人工判讀且耗時的問題。系統整合了頻譜減法、滑動視窗切片與小波轉換事件偵測等前處理技術，並結合 Random Forest 進行二階段雜訊過濾，大幅提升了音訊資料的有效性。本專題採用之雙分支架構，證實了多模態特徵在超聲波辨識上的研究價值，期許在未來能為台灣生物多樣性檢測與蝙蝠生態保育做出實質貢獻。