

研究目的

由於可見光與電磁波在水中傳遞距離有限，聲學技術成為監測船舶聲紋的主要方法，尤其在偵測水下載具時更是唯一可行手段。掌握水下聲學環境即能保障港域水下安全。然而，港域內多船混響使聲學監測面臨挑戰。本研究建立一套適用於多聲源環境的船舶聲音分離與分類方法，先以非負矩陣分解(NMF)方法在不同噪音與訊噪比條件下分離船舶聲音，再利用AI進行船種分類。研究成果可提升港域智慧監測與安全，並增進臺灣於國際海事技術與國防自主的競爭力。

研究方法

$$V_{ft} \approx (WH)_{ft} = \sum_{a=1}^r W_{fa} H_{at}$$

NMF將觀測到的音檔，作為非負數據矩陣V，分解為兩個非負矩陣：
 - 基底矩陣W (BASIS MATRIX)：基底分量在頻率F上的特徵強度
 - 編碼矩陣H (ENCODER MATRIX)：基底分量在時間上的特徵強度
 以實現訊號分解與特徵學習的目的 (LIN ET AL., 2017)

聲源分離步驟與成果

資料整理

NMF

提取特徵 訓練模型

建立兩船舶種類的聲紋模型，其頻率特性與時間分布特性皆不同

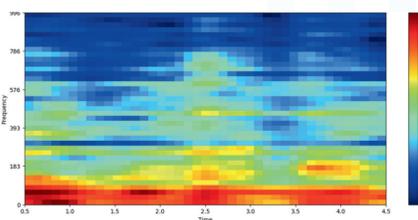


圖1. 待訓練的音檔特徵範例

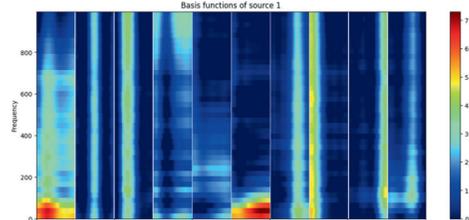


圖2. NMF將聲源特徵分離

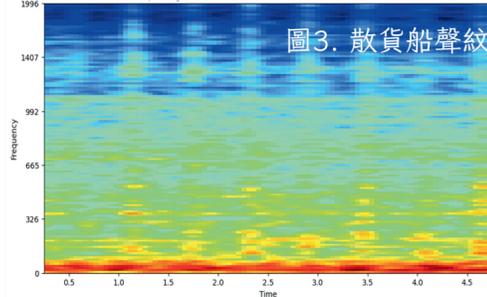


圖3. 散貨船聲紋

散貨船的聲紋特徵為一個頻帶
(圖上一條紅線)

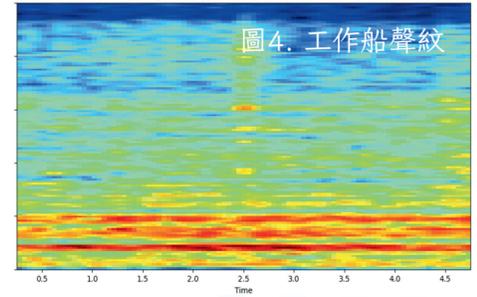


圖4. 工作船聲紋

工作船的聲紋特徵為三個頻帶
(圖上三條紅線)

待訓練資料樣本數為69筆單一船舶音檔資料，因此有69張頻譜圖，以其中一艘船舶為例

利用NMF提取69張頻譜圖上的船舶聲紋特徵，再將特徵納入AI模型進行訓練

預測船舶種類相似度比較

本研究使用RESNET18 VIA TORCHVISION掃描NMF分離結果圖，提取時頻圖上特徵，並轉換為數字呈現。接著以PYTORCH COSINE 相似度，比較兩組時頻圖之特徵向量，計算結果為當角度越小(數值越接近 1)，表示圖片越相似

利用模型 預測混合音訊

模型預測結果判定為工作船，與圖4.工作船的聲紋特徵比對，皆是圖上三條紅線，結果相符。下一步進行相似度比較。

相似度比較表

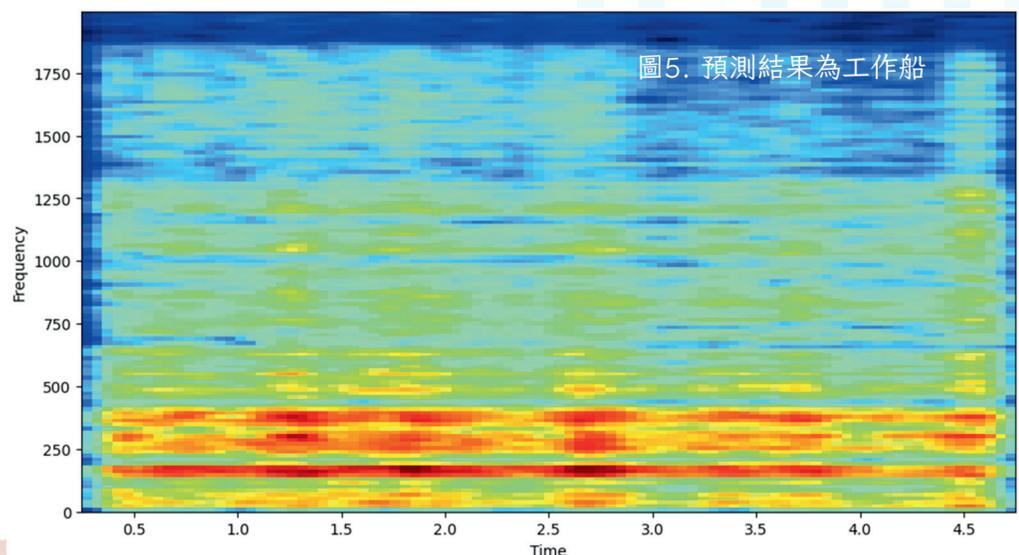
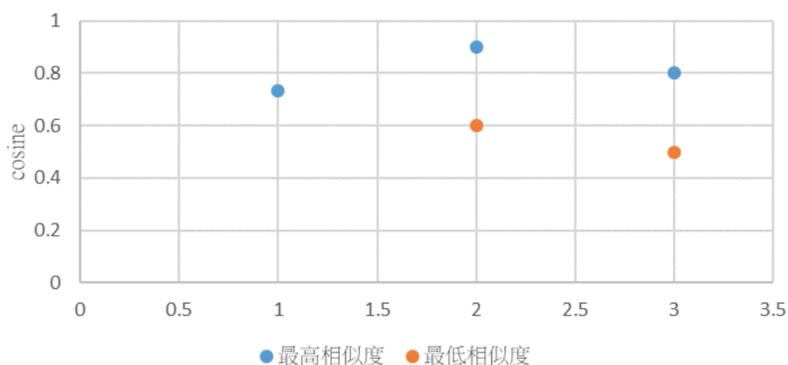


圖5. 預測結果為工作船

成果與討論

1. 原始音頻與分離後時頻圖的 COSINE 相似度為 0.735 (73.5%)，顯示分離後的音頻保留大部分原始頻譜特徵
2. 與相關文獻比較，本研究在船舶音頻分離任務中的效能處於合理且可接受的範圍
3. 約26.5%的差異可能來自分離過程中的雜訊或細微失真，可透過大數據訓練來提高相似度

本研究透過LISTEN聲學技術了解與辨別水下目標物種類，守護台灣水下環境安全

致謝

感謝中央研究院林子皓副研究員提供 SOUNDSCAPE_IR 平台 NMF 演算法，作為本研究分析之基礎工具。感謝國立成功大學水利及海洋工程學系劉冠汶助理教授於專題研究與海報製作上的指導，並感謝劉姿瑩碩士生在程式操作與海報製作方面提供協助與建議。

參考資料

