

**簡報大綱：Leveraging Allophony in Self-Supervised Speech Models for Atypical Pronunciation Assessment**

**一、研究背景與動機**

* 音位異音（Allophony）現象與其在語音評估中的重要性
* 傳統自動發音評估方法的侷限
  + 單一分布假設
  + 無法有效處理異常發音（如語音障礙、非母語者）
* 自監督語音模型（S3M）於語音特徵表徵的突破

**二、相關研究回顧**

* Goodness of Pronunciation (GoP) 傳統方法及其演進
* DNN、S3M於發音評估的應用現況
* 異常語音評估與分布外（OOD）偵測的研究趨勢

**三、研究目標**

* 提出能有效建模音位異音的新方法
* 提升異常發音自動評估的準確性與魯棒性
* 驗證S3M特徵於異音建模與發音評估的效益

**四、方法與工作流程**

* 傳統GoP方法的數學模型與侷限
* MixGoP方法核心設計
  + 使用高斯混合模型（GMM）建模每個音位的多子分布（異音）
  + 結合S3M（如WavLM、XLS-R）特徵
  + 以對數似然分數評估發音異常程度
* 實驗流程圖/步驟
  + 特徵萃取 → 音位分割 → 聚類初始化 → GMM訓練 → 分數計算 → 評估

**五、主要數學公式解析與比喻**

* GoP分數（語音教師信心分數）
* 傳統分類器logits（座標距離比喻）
* GMM建模（水果籃與品種隱喻，子分布意義）
* MixGoP分數（屬於該音位的機率對數分數）
* Mahalanobis距離（多維距離與變異考量）
* 全句發音分數（平均分數，整體表現）

**六、實驗設計與結果**

* 資料集介紹（語音障礙、非母語共五組資料）
* 特徵比較（MFCC、Mel spectrogram、TDNN-F、S3M）
* 基線方法與MixGoP表現比較
* 結果摘要：MixGoP於四個資料集達最佳表現
* S3M特徵異音建模能力分析

**七、深入分析**

* S3M特徵異音資訊可視化與量化
* 異音建模能力與下游任務表現的關聯
* MixGoP樣本效率與子分布數量敏感性

**八、討論與未來展望**

* 方法優勢與侷限（語言泛化性、時間對齊品質等）
* S3M於語音異常評估的應用潛力
* 未來可延伸的研究方向

**九、結論**

* MixGoP方法突破傳統侷限，顯著提升異常發音評估效能
* S3M特徵在異音建模與實際應用上的優越性
* 對自動語音評估與語音特徵分析領域的貢獻

**十、Q&A**

**備註：**

* 每一部分可搭配流程圖、公式說明、數據圖表等輔助說明
* 公式部分建議以生活化比喻協助理解
* 實驗結果與分析可重點強調MixGoP的優勢與創新

⁂