

秒懂真偽：以低成本證據鏈驅動的多模態 Refine-Retrieve-Reason短影音假訊息偵測

學生:蔡意嫻、施淳譯、蕭凱云

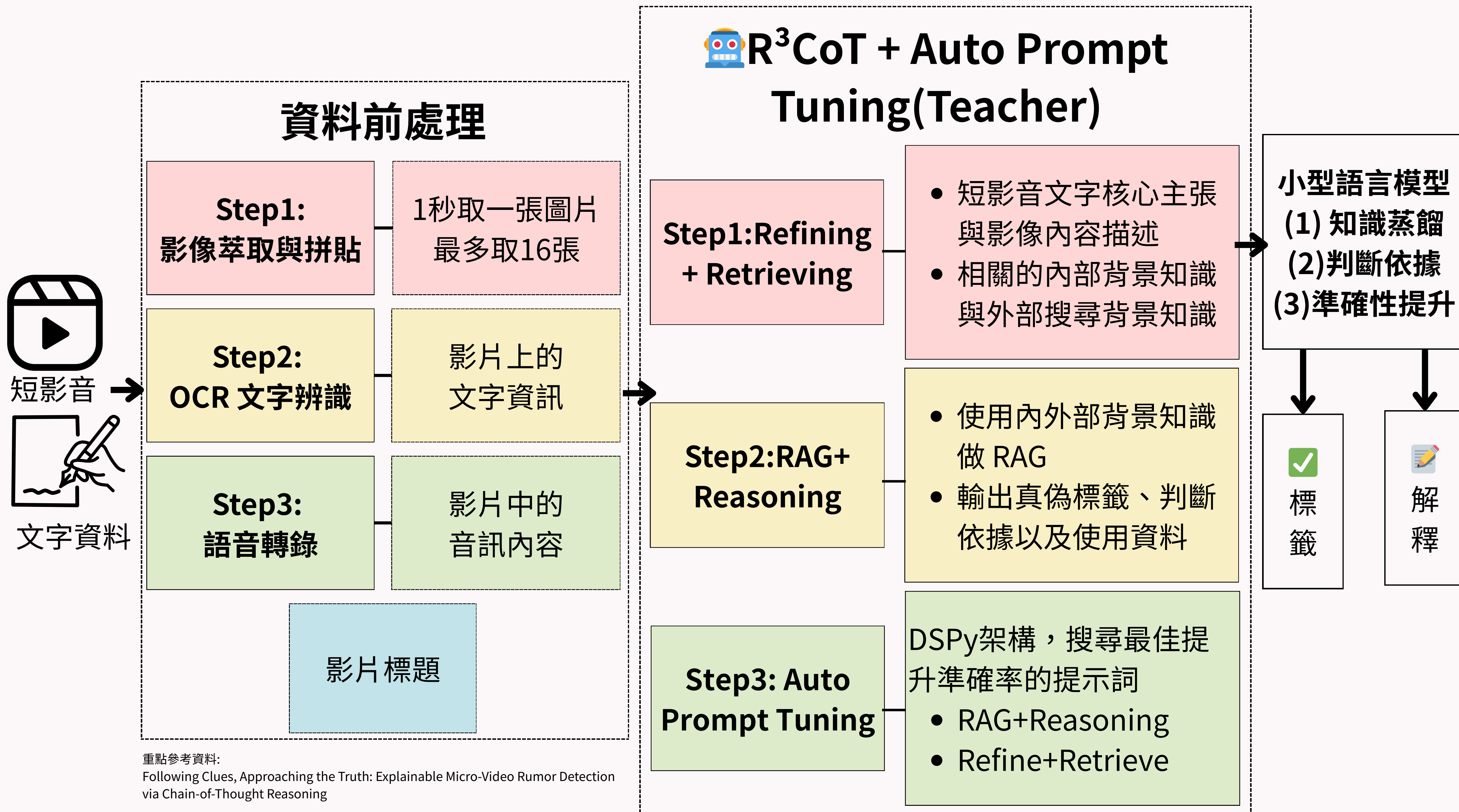
指導教授:李政德 特聘教授

研究動機與目的

- 短影音假訊息傳播快速且具多模態特性，傳統文字辨識方法效果有限。
- 結合多模態分析、外部檢索與推理、自動化提示詞優化與知識蒸餾辨識架構。降低成本同時提升準確性與可解釋性

實作技術與環境

- 多模態資料處理
- RAG 結合推理架構
- DSPy Auto Prompt Tuning
- Teacher-Student 知識蒸餾
- VS Code
- Python 3.12
- Google Colab
- Vue.js (Vite)



成果展示



結論與未來展望

- 提升準確率與 F1-score 表現、降低 API 呼叫成本
- 提供外部證據來源與推理過程，提升系統可解釋性
- 導入即時資料，提升模型對新型假訊息之泛化能力
- 增加應用於 Deepfake 與社群媒體內容審核