

經濟部 110 年度
《數位信任鑑識技術先導研發計畫
—不實資訊鑑識分析與決策處理 (1/1)》
合作研究計畫

《時域語言定位技術演算法研究與雛形開發》

建議書徵求文件

財團法人資訊工業策進會

中華民國 110 年 05 月 26 日

110年度合作研究計畫建議書徵求文件

一、 簡介

近年來民眾新聞資訊來源，已逐漸透過FB、LINE等社群頻道取得新知。但若是屬於不實訊息，藉由網路傳播與社群媒體，可更快速讓一小群人影響、干擾或操弄輿論，不論意圖為何，皆有可能對政府社會或機構造成不信任與傷害，無論從技術發展、社會觀感以及法律規範上，均有不同層次之影響。隨著AI偽造技術的進步，影像類不實訊息占比約30%，較前一年上升約40%以上。現階段Google等國際大廠雖也有提供如以圖找圖等查核工具給查核單位。但查核人員多需要一一到各個工具進行操作，缺乏有效整合機制，且影像類查核的工具並不完整。另一方面，資安業者或輿情業者，現階段僅針對社群資料進行統計，部分透過文字探勘技術，發展情緒分析等模組，但對於影像資料也缺乏分析的工具。

有鑑於影片類型的不實訊息逐漸增多，需耗費大量人力時間、且最難查核分析，本計畫旨在開發可應用於輔助影像類事實查核的相關技術，以提高事實查核機構的查核效率。

二、 計畫目標

影片類型的不實訊息逐漸增多，需耗費大量人力時間、且最難查核分析。FB、Youtube受到影響最大，目前都有計畫向全球查核機構尋求支援。不實影片以變造的技術來區分，可分為：(1) 演算法偽生成的影像訊號；(2) 人為編輯、錯置及重組的影像及文字；(3) 將文字為主的不實訊息製作成有大量字幕及/或配音的內容農場型影片。由於Deepfake成本高，所以現階段不實訊息還是以前述的第二類及第三類為主，尤其第三種影像類不實訊息的影片長度偏長，為台灣不實訊息的一大特色。

以變造類影片（第二類）來說，近期較知名的是社群媒體上瘋傳宣稱「民主黨候選人拜登在節目上睡覺」的影片，根據台灣事實查核中心現階段所施行的人工查核方法，主要透過下列幾種方式檢驗是否為不實訊息：a. 核對影片標題與影片中相關人的發言；b. 檢視影片是否有剪輯痕跡；c. 將影片分成局部畫面，再以圖反搜；d. 檢索新聞媒體是否有相關報導。

以內容農場類影片（第三類）來說，此類影片的不實訊息主要來自影片中的大量字幕，影像反而不是捏造重點，但這類影片長度偏長，查核人員需看完影片並謄錄內容，才能往下進行內容查核，既耗時也耗工。

基於以上影像類查核實務常用手法，本計畫擬開發以下兩種技術，從兩方面來輔助查核人員提升查核效率：

- (1) 應用時域語言定位技術 (Temporally Language Grounding in Untrimmed Videos)，提供查核人員可透過自然語言、圖片，來針對影片內容進行檢索功能，找出符合描述的影像段落，使查核人員快速比對不实訊息是由哪些影片中的片段所組合而成，或檢視影像是否有人為編輯、錯置及重組的問題，加速影片查核過程。
- (2) 影像文字自動辨識技術，自動將影像中的字幕/影像內文字如招牌、路牌、後製文字等等自動轉出，並區分為字幕文字與影像文字，且搭配文字出現時間。

三、計畫範圍

- (1) 進行文獻閱讀與研究，評估現階段技術缺口，發展新型態演算法。
- (2) 提供以下兩項技術模組雛形，與計畫分項系統整合。
 - 時域語言定位技術
 - 影像文字自動辨識技術

四、預期成果

查核點	時間	產出物	內容概述
M1	2021/11/01	(1)期中報告 (2)「影像文字自動辨識技術」程式雛形	<ul style="list-style-type: none"> ■ 期中報告(不拘形式) ■ 影像文字定位與辨識文獻探討 ■ 問題定義 ■ 模型設計 ■ 資料集蒐集 ■ 雛形軟體成果 (python code / docker) 與操作說明 (含測試結果 /IN/OUT put說明等)
M2	2022/03/15	(1)期末報告 (2)「時域語言定位技術」程式雛形	<ul style="list-style-type: none"> ■ 期末研究報告書 ■ 文字影像跨域研究文獻探討 ■ 問題定義 ■ 模型設計

			<ul style="list-style-type: none"> ■ 資料集蒐集 ■ 雛形軟體成果 (python code / docker) 與操作說明 (含測試結果 /IN/OUT put說明等)
--	--	--	---

(1) 時域語言定位技術：

- 根據輸入文字，找出資料庫中影片相關片段時間，回傳相關影片與對應時間。
- 輸入文字除了能查詢一般動作外(例如：一個小女孩在玩球)，希望能夠支援名人影片搜尋(例如：川普在玩球)。
- 輸入文字為英語或繁體中文。

(2) 影像文字自動辨識技術：

- 自動辨識影像中的字幕(含固定位置與跑馬燈式字幕)/影像內文字(如招牌、路牌、商店名稱、浮水印等)並轉換成文字。
- 轉換文字可區分為字幕與影像內文字，並標記出現文字於影片中出現時間。
- 可識別語系包含中文繁簡體與英語語系。

※前述成果如有專利構想或專利申請產出時，需注意專利申請之新穎性(novelty)。因凡經公開發表之研發成果，如擬申請專利，須於公開發表後6個月內完成，前述成果如是以論文方式公開發表，將無法取得大陸與歐盟等國之專利。

五、 執行方式

- (1) 計畫時程：110年06月25日日至111年03月31日
- (2) 合作研究單位經審核通過執行本計畫後，指派承辦人員一名作為本計畫之聯繫窗口。
- (3) 合作研究單位審核通過執行本計畫後，每月以實體或線上會議方式說明執行進度與相關執行內容並進行議題討論。
- (4) 有關本計畫工作之進行，應依照本會之相關管理程序及程式管理方法進行管理並累積相關品質紀錄文件。

六、 計畫期程及預估計畫總經費

計畫執行區間：110年06月25日至111年03月31日

總經費：800,000元

七、 驗收標準(含教育訓練)

- (1) 「影像文字自動辨識技術」演算法實作程式
- (2) 「時域語言定位技術」演算法實作程式
- (3) 期末研究報告書乙份（包含以上兩項技術之操作說明）

八、技術能力需求

- (1) 熟悉深度學習、自然語言處理、影像處理相關技術
- (2) 具備時域語言定位技術與影像文字自動辨識技術演算法設計經驗與研究分析能力
- (3) 具備程式撰寫與軟體架構開發能力，實作演算法原型，演算法效能優化，並進行實驗量化驗證