

國立彰化師範大學113年百萬專題探索申請表

申請編號：

一、專題基本資料

申請類別	<input type="checkbox"/> 產學合作類 <input checked="" type="checkbox"/> 跨域整合類 <input type="checkbox"/> 競賽類 <input type="checkbox"/> 主題式專題類 <input type="checkbox"/> 大專學生研究計畫				
課程名稱	(有配合課程請填寫，無配合課程可不填)				
類別 非主題式專題免填	<input type="checkbox"/> 專業領域精進 <input type="checkbox"/> 跨領域學習 <input type="checkbox"/> 在地文化探索 <input type="checkbox"/> 服務學習 <input type="checkbox"/> 其他：_____				
專題題目	AI 手語通：整合人工智慧技術的手語翻譯及互動平台				
指導老師	教師姓名		服務單位		聯絡電話
	陳仁德		資訊工程學系		8439
	林似霖		車輛科技研究所		7059
職稱	學生姓名	系級	學號	E-mail	電話
專題聯絡人	季正偉	工二甲	S1131034	4902505521621a@gmail.com	0983392669
專題成員	季正偉	工二甲	S1131034	4902505521621a@gmail.com	0983392669
	楊惟麟	資管一	S1161016	paiyang066@gmail.com	0972919478
	楊敦傑	資工一	S1254013	grasonjas@gmail.com	0978852141
	施丞澤	資工一	S1254005	s1254005@gm.ncue.edu.tw	0968818757
	吳翊愷	輔三乙	S1011137	S1011137@gm.ncue.edu.tw	0977385814
擬參加之競賽 (優先審核條件)	競賽性質		<input type="checkbox"/> 國際性 <input checked="" type="checkbox"/> 全國性 <input type="checkbox"/> 區域性 <input type="checkbox"/> 校內 <input type="checkbox"/> 院系內		
	競賽名稱		全國大專院校產學創新實作競賽		
	每年大約參賽人數		100		
	競賽網址(若無競賽網址，請檢附相關資料證明)		http://eedept.ncue.edu.tw/innovation2023/intro.html		

產學合作類填寫

合作廠商名稱	
優先核准條件	1. <input type="checkbox"/> 已與廠商簽訂產學合作計畫，計畫金額 _____。 2. <input type="checkbox"/> 預定協助廠商申請產學合作相關計畫(如 SBIR、CITD等)。

審查結果	(此欄由教學發展中心填寫) <input type="checkbox"/> 通過 <input type="checkbox"/> 修改後通過，修改項目： <input type="checkbox"/> 不通過
------	--

二、專題內容說明(以不超過3頁為原則)

(一)專題簡介(主題式專題類：請描述執行專題動機、目標、執行策略、學習進度及自評方式)

手語是殘障人士溝通的主要方式，然而由於手語系統在不同國家和地區存在差異，來自不同語言背景的手語使用者之間存在溝通障礙。本計畫旨在開發一種系統，透過手語辨識和機器翻譯技術，幫助英語手語使用者與中文使用者實現無障礙交流。

目前，英語手語使用者與中文使用者之間的交流存在巨大挑戰。由於語言和手語系統的差異，雙方無法直接交流，只能依賴有限的手語翻譯人員，這極大限制了他們的生活和工作。發展一種自動手語翻譯系統，能夠打破此語言障礙，促進不同語言背景者之間的無障礙溝通。

成功發展此系統，不僅能促進英語和中文手語使用者之間的交流，也將推動殘障人群的社會融合，提高他們的教育、就業和生活品質。同時，該系統在旅遊景點、醫療機構等公共場所的應用，也將惠及更廣泛的用戶群體。

MediaPipe 是一個由 Google 開發的開源跨平台框架，旨在使機器學習解決方案能夠在異質環境中高效運作。它支援多種感知任務，包括手語辨識。MediaPipe 的核心是用於構建機器學習管道的框架，提供了數據流模組化設計、跨平台部署模型和高性能低延遲推理的功能。開發者可以輕鬆地將自訂機器學習模型整合到 MediaPipe 中。

在手語識別和翻譯應用中，MediaPipe 能夠檢測視頻/圖像中的手部區域，跟踪和估計手部關鍵點位置，對手語動作序列進行識別，並將識別出的手語翻譯為相應的文字或語音。MediaPipe 為建構手語翻譯系統提供了強大支援，能在不同硬體平台上高效運行機器學習模型，實現即時手語檢測、追蹤和翻譯。

OpenCV 是一個跨平台的開源電腦視覺和機器學習庫，由一系列 C++ 函數組成，可運行在 Windows、Linux、Android 和 macOS 等系統上。它提供了圖像和視頻處理的各種功能，在計算機視覺領域得到廣泛應用。

在手語辨識和翻譯研究中，OpenCV 可以發揮重要作用。它為視訊/影像擷取、處理和分析提供基礎功能。OpenCV 可用於視訊讀取和影像擷取、影像預處理如降噪、增強對比度，以及手部區域檢測與分割、手勢特徵提取如輪廓、HOG 等，還可用於機器學習模型的訓練和測試。OpenCV 庫提供了大量影像處理演算法和工具，開發者可充分利用來建立高效手語辨識管道。例如，結合 Haar 級聯分類器進行手部檢測，使用自帶 HOG 描述符提取手勢特徵等。

TensorFlow 是一個由 Google 主導開發的開源人工智慧庫，它可以被廣泛應用於構建和訓練機器學習模型，尤其在計算機視覺和自然語言處理等領域。在手語辨識和中文翻譯的研究項目中 TensorFlow 可以發揮重要作用。

TensorFlow 擁有強大的電腦視覺能力，可以處理諸如影像分類、物件偵測、影像分割等任務。在手語辨識，TensorFlow 可用於建構卷積神經網路模型，對手部影像進行特徵提取，辨識出特定的手語動作。

TensorFlow 也擅長序列資料建模，可用於建構循環神經網路等模型，對手語連續動作序列進行端到端的辨識與翻譯。這對於實現流暢的手語到中文的即時翻譯至關重要。

再者，TensorFlow 支援多種平台部署，包括伺服器、行動端、嵌入式設備等，這使得交付能夠針對不同場景優化，提高系統的可用性。

除了核心庫之外，TensorFlow 生態系統還包括 TensorFlow 擴展支援各種模型架構，如用於手語識別的姿態估計模型；TensorFlow Hub 提供大量預訓練模型以及可遷移的功能模組；TensorFlow Serving 用於高效部署生產模型等，這些功能可以加速模型開發和最佳化流程。TensorFlow 是一個全面且領先的機器學習開發平台，涵蓋了從數據處理、模型構建到模型優化部署的完整流程，將有助於高效開發出性能優異的手語識別及翻譯系統。

Transformer 是一種全新的基於自注意力機制的序列模型架構，由 Google 的 Vaswani 等人在 2017 年提出，主要應用於機器翻譯等自然語言處理任務。相較於傳統的基於循環神經網路 (RNN) 的序列模型，Transformer 不存在循環計算，而是完全依賴自注意力機制對輸入序列進行建模。

在本專案的英語到中文神經機器翻譯系統中，我們採用了編碼器-解碼器結構，其中編碼器利用了 Transformer 的自注意力機制對源語言(英語)進行建模，解碼器則融合了 Transformer 解碼器和注意力機制對目標語言(中文)進行產生。

Transformer 的自注意力機制能夠同時捕捉序列中任意兩個位置的關係，更好地學習長距離依賴資訊，這正是在處理手語等長序列資料時所需要的，與 RNN 相比，Transformer 結構還具有更高的平行運算能力，在訓練和推理時可以獲得更快的運行速度。

除了在翻譯模型中的應用，Transformer 架構也可用於手語動作辨識模型。透過自注意力層對手部關鍵點座標序列建模，能夠更好地學習手勢動作之間的時序相關性。

總體來說，Transformer 模型憑藉其強大的長距離建模能力和高效的平行計算結構，在本項目的手語識別和英漢互譯任務中發揮著關鍵作用，將有助於提高系統的準確性和即時性，為使用者帶來更好的無障礙交流體驗。

執行策略：

1. 使用 OpenCV 讀取視訊串流或影像序列作為輸入來源，負責取得視訊畫面並進行預處理，降噪、增強對比度等，以提高影像品質。
2. 將預處理後的影像影格輸入到 MediaPipe 手部檢測模型中。MediaPipe 將檢測影像/影片中的手部區域，並利用其先進的手部關鍵點模型估計出每一幀中手部的 21 個 3D 關鍵點座標。

3. 將MediaPipe輸出的手部關鍵點座標序列作為輸入，送入基於TensorFlow的手語動作辨識模型中，此模型結合循環神經網路等技術，能夠對連續的手部關鍵點座標序列進行端到端建模，辨識出對應的手語動作。手語動作辨識模型的輸出為辨識出的單字或字母序列。對於字母手語，我們將利用TensorFlow對序列進行拼接，拼接出完整的單字，並組合成句子。
4. 手語動作辨識模型由TensorFlow訓練產生，利用了Kaggle上的公開手語資料集以及自行收集的手語資料進行訓練，此模型輸出為辨識出的單字或字母序列。
5. 將辨識出的英文單字送入融合了Transformer模型的基於TensorFlow訓練的英文到中文神經網路翻譯模型中，輸出對應的中文譯文。此翻譯模型採用編碼器-解碼器架構，編碼器利用Transformer的自注意力機制對源語言進行建模，解碼器則融合了Transformer解碼器和注意力機制對目標語言進行生成。引入Transformer能夠更好地捕捉長距離依賴，提高翻譯品質。翻譯模型使用了公開平行語料庫和自建語料進行訓練。
6. 最後，將模型輸出的中文譯文呈現在螢幕或語音設備上，以便中文使用者理解手語使用者的輸入內容，實現無障礙溝通。
7. 除了對手語進行辨識與翻譯，此系統還需具備辨識數字手語的能力。
8. 為提高系統的即時性和穩定性，我們將使用TensorFlow Serving進行模型部署，並結合NumPy等庫進行高效的資料預處理和特徵提取。
9. 對於即時視訊輸入，系統需要在每一幀上執行上述步驟。透過高效的管線設計和模型最佳化，系統將能夠達到即時處理和低延遲的目標。

透過結合公開資料集和自建資料集，我們能夠為手語識別和機器翻譯模型提供更加多樣化和充足的訓練資料，從而提高這兩個關鍵模型的性能和泛化能力。公開資料集為模型奠定了基礎，而自建資料集則有助於針對實際應用場景進行客製化最佳化，使系統更加貼合最終用戶的需求。

此外，在訓練階段,我們還可以利用 TensorFlow 提供的各種訓練策略和優化器，資料增強、遷移學習、模型整合等，以充分挖掘資料潛力，提升模型的準確性。

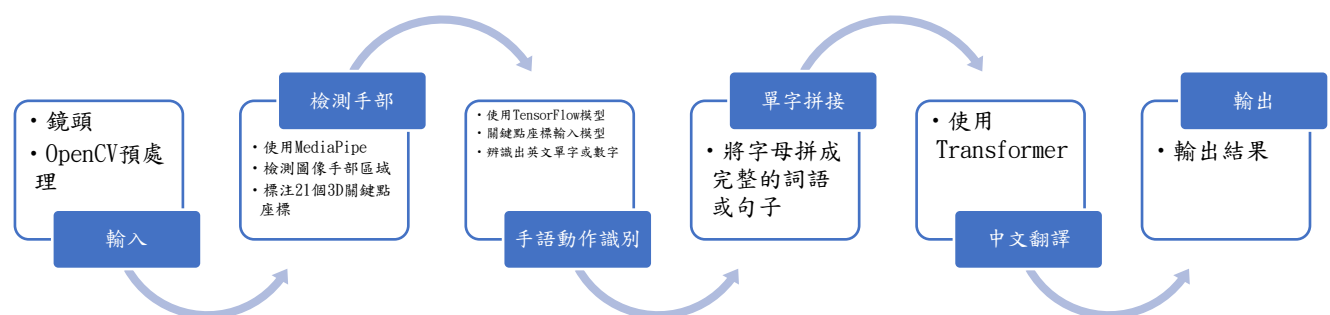


圖1.執行策略圖

(二)預期成果

本計畫綜合運用了電腦視覺、自然語言處理和機器學習等人工智慧技術，旨在為英語手語使用者和中文使用者建構無障礙溝通管道。系統將能夠即時識別英語手語動作序列，並將其精準地翻譯為相應的中文文本，有效打破不同語言背景者之間的溝通障礙。

透過合理利用公開資料集和自建資料集的策略，為核心的手語識別和機器翻譯模型提供了高品質的訓練資料支援。公開資料為模型奠定基礎，而自建資料則使系統能針對實際應用場景進行客製化最佳化，更好地滿足最終使用者需求。同時，借助 TensorFlow 等先進工具進行模型訓練、優化和高效部署，確保了系統的高效能和實用性。

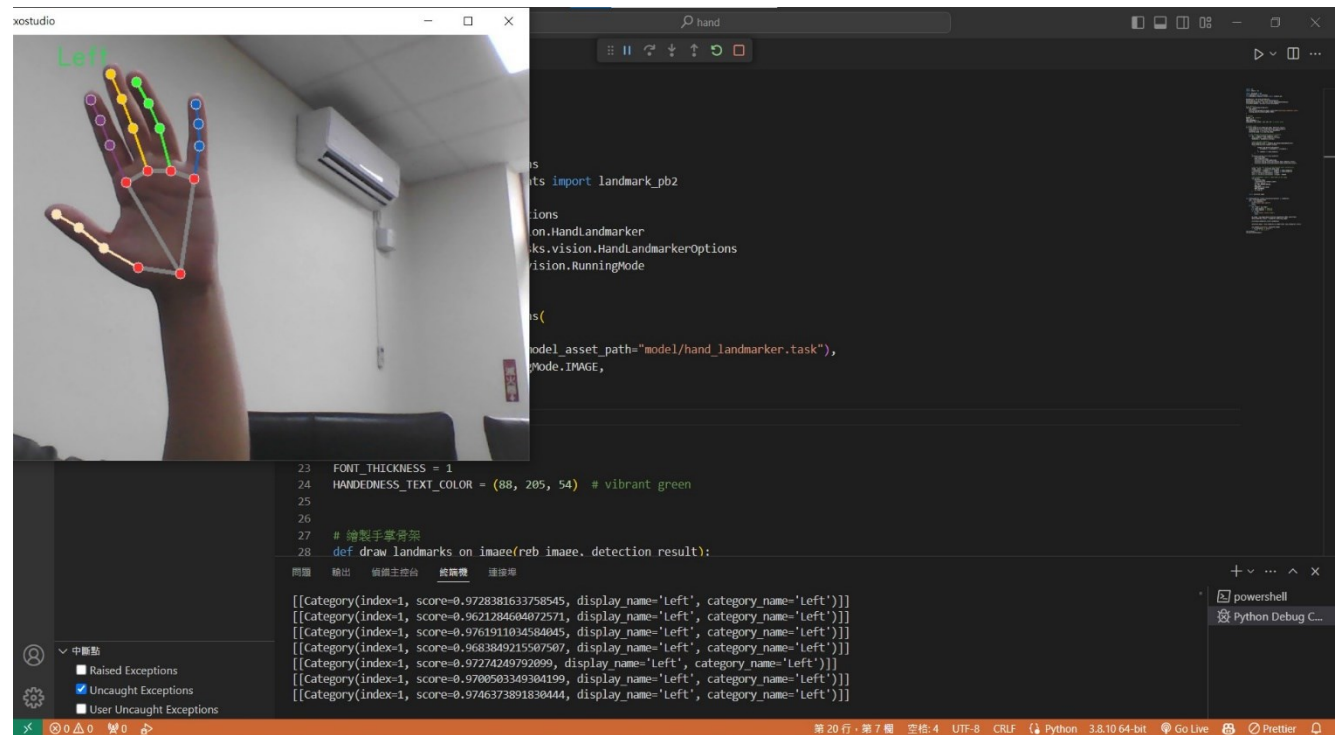


圖2. MediaPipe 辨識並標出手部座標

本計畫將會帶來以下成果：

1. 突破語言障礙，為英語手語使用者和中文使用者建構全新的無障礙交流橋樑。
2. 顯著提高殘障人群的社會融合度，改善其教育、就業和生活品質。
3. 在旅遊景點、醫療機構等公共場所推廣，惠及更廣泛的使用者群體。
4. 推動人工智慧技術在助殘領域的創新應用，為語音視覺障礙人士帶來全新的獲取資訊和表達需求的方式。
5. 系統模組可擴展至其他手語和語言,促進不同文化背景群體的相互理解與包容。
6. 整合成便於使用的應用程式。

總之，透過整合多種人工智慧技術，並在資源配置上精心部署，本計畫將為不同語言背景的手語使用者提供高效、無障礙且高性能、實用的手語識別和英中互譯系統，真正實現多元文化的包容與融合，為建構和諧社會貢獻重要力量。

三、經費編列

經費編列及使用必須與專題題目直接相關，經費核銷依教育部及本校相關規定辦理。

補助經費項目包括物品、印刷費、雜支、交通費等，可依實際需求調整，補助詳細內容請詳閱本試辦計畫。

經費項目		預估經費總額	請說明預購買之物品項目 (不需逐項詳列)
業務 費	物品費	11000	鏡頭、電子零件、墨水
	印刷費	500	影印資料及成果報告書
	雜支	1000	資料夾、影印紙等
	交通費	0	
總計		12500	

經費核銷方式

皆由**教學發展中心協助報支**，請同學將購買物品之單據交至進德校區教學大樓5樓教學發展中心，由教學發展中心同仁 KEY 單後進行核銷流程。收據/發票開立需有本校統編或抬頭，如下：

收據抬頭：國立彰化師範大學

本校統一編號：58815502

四、申請申明(請務必勾選)

請教師確認以下內容並簽名確認：

■本人瞭解113年百萬專題探索試辦計畫及申請表所列之注意事項，並誠實填寫申請表內容。

■申請表所列經費項目均為專題相關所用，亦無重複申請或報支其他計畫經費。

☐產學合作類，已檢附與產企業合作之相關證明(格式不拘)。

申請教師 簽章	
------------	--

備註：

1.表格不敷使用，請自行延伸；其他相關佐證資料，請與本文件一併裝訂於後。

2.申請方式：填妥申請單後，將申請表電子檔寄至 cup900324@cc.ncue.edu.tw，信件主旨請註明「百萬專題探索申請表-○○○(專題題目)」。

3.相關問題請洽教學發展中心承辦人員，林怡均 04-7232105 #1612。