

2025年臺灣國際科學展覽會

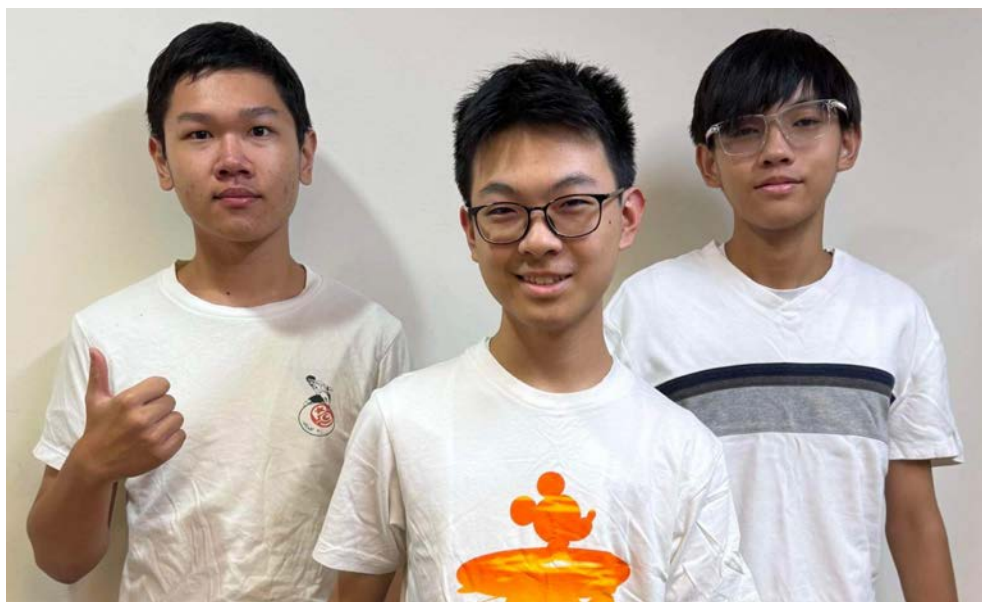
優勝作品專輯

作品編號	100026
參展科別	工程學
作品名稱	為視障者開發之學習輔助平台：結合Image-to-3D AI 模型之可觸式三維擴增實境顯示器與個人化之檢索增 強生成（RAG）自然文字系統
得獎獎項	三等獎 國際科學博覽會(ESI)

就讀學校	國立臺南第一高級中學
指導教師	莊坤達 顏永進
作者姓名	林澈 饒凱富 廖紘德

關鍵詞 視覺障礙輔助、陣列作動器、大型語言模型

作者簡介



我們是台南一中高二的饒凱富（右）、林澈（中）、與廖紘德（右），熱愛資訊科技與工程實作。本次科展聚焦於開發基於觸覺反饋的視障者學習助手，透過科技協助視障人士突破學習障礙，推進教育平等性。林澈曾獲全國高中英語辯論與演講冠軍，展現了卓越的語言與邏輯能力，並在團隊中負責整體架構規劃與專案管理；廖紘德與林澈對地球科學充滿熱情，專注於模組設計與系統整合，深入探討科學知識對社會發展的影響；而饒凱富則專注於電子工程，負責三維設計與硬體開發，對電路穩定性與效能優化有深入研究。我們在過程中結合創新設計與程式開發，探索觸覺介面的應用與影響力。除了專案開發，我們還與視障人士密切交流，根據實際需求調整設計，提升系統實用性與人性化。未來，我們將持續深耕輔助科技，整合跨領域知識，致力於開發改變生活的技術，為社會帶來創新。

2025 年臺灣國際科學展覽會

研究報告

區別：

科別：工程學科

作品名稱：

為視障者開發之學習輔助平台：結合 Image-to-3D AI 模型之可觸式三維擴增實境顯示器與個人化之檢索增強生成（RAG）自然文字系統

關鍵詞：視覺障礙輔助、陣列作動器、大型語言模型
編號：

（編號由國立臺灣科學教育館統一填列）

目錄

中文摘要	1
English abstract.....	2
內文.....	3
壹、前言.....	3
一、研究動機.....	3
二、研究目的.....	4
三、文獻回顧.....	5
(一) 視覺障礙者之學習特點及所面臨之困境	5
1、視覺障礙的定義	5
2、視障者的感官知覺主要來源.....	5
3、視障者觸摸立體物件時的認知歷程	5
4、視障學生之學習困境	6
(二) 現有視障者輔具未能有效協助視障者理解複雜空間圖形	7
1、市面上提供之視覺型輔具類型	9
2、立體視覺輔具及動態資訊輔具獲取不易.....	9
(三) 大型語言模型 (LLM) 在教育上的使用	10
1、LLM 之強化方式探討.....	10
2、RAG (Retrieval Augmented Generation) 搜索式強化生成	11
3、大型語言模型 (LLM) 在教育上使用之優點	11
貳、研究方法與過程	12
一、研究方法.....	12
(一) 進行文獻探討及訪談視障者和相關人士以確定研究方向	12
(二) 設計實驗架構與規劃流程	13
(三) 確立「視障者針對空間概念學習使用之智慧輔具系統 VDE (Visual Disabilities Education)Hybrid Transformer」涵蓋之功能	13
1、圖形感知強化之一：研究與開發互動式觸覺輔具—Tactile Pad.....	13
2、圖形感知強化之二：探討結合 Tripo-SR 模型與 Unity AR 系統.....	13
3、文字理解強化—基於 RAG 的 LLM 互動系統的實施.....	13
(四) 視障者使用 VDE Hybrid Transformer 多功能輔助系統效能評估	14
1、圖形感知強化方面：視障學生使用 Tactile Pad 之效能評估	14
2、文字理解強化方面：文字理解強化效果之評估	14

二、研究過程.....	14
圖形感知強化系統.....	14
文字理解強化系統.....	14
(一) Tactile Pad 觸覺反饋系統的探討與開發過程.....	15
1、觸覺反饋 (Tactile Feedback) 裝置的技術探討.....	15
2、繪製整體設計圖稿.....	18
3、設計電路.....	20
4、硬體設計與製作.....	23
5、第二代-使用者交互介面 Control Panel 開發.....	24
6、第三代-光學雷達系統校正和數據反饋.....	25
(二) AR Unity System 與 2D 轉 3D 生成模型及實現.....	31
1、使用 TripoSR 生成 3D 物件.....	31
2、使用 Unity 3D 渲染並透過觸控板探索 3D 物件.....	33
(三) 研究與開發強化視障生文字理解方法—LLM-based text optimizer for VD Education.....	33
1、研究 LLM-based Personalization & RAG 輔助系統.....	33
2、嵌入視障生輔助相關資料到檢索增強生成 (RAG) 系統中.....	37
(四) 驗證「文字強化系統」之成效.....	41
1、使用 VDEA Benchmark 驗證輔助成效.....	41
2、評估新式輔具之成效.....	42
三、研究使用之材料、硬體設備、器材與軟體.....	43
(一) 使用之材料.....	43
(二) 使用之硬體設備與器材.....	43
(三) 使用之軟體.....	44
參、研究結果與討論.....	44
一、觸覺反饋強化與文字強化對視障者實際理解與認知能力的影響.....	44
二、視障生對本研究開發之雙系統之評價.....	49
三、利用假說檢定進行更嚴謹的統計分析.....	50
四、Visual Disabilities Education Hybrid Transformer 與常用視覺輔具之差異.....	53
三、未來展望.....	55
(一)往 VisionLLM 進行探討應用與最佳化.....	55
(二)採用多重 (Multimodel)觸覺重現系統.....	55

肆、	結論與應用	57
一、	本研究提出了視障者多元輔助系統	58
二、	本研究提出了全新的視障生學習輔助系統成果驗證方式—VDEA benchmark	58
三、	本研究開發的 AR 特殊多點輔助板	58
四、	本系統開發過程中發掘許多未經深入探索的領域	58
伍、	參考文獻	59

中文題目

為視障者開發之學習輔助平臺：結合 Image-to-3D AI 模型之可觸式三維擴增實境顯示器與個人化之檢索增強生成（RAG）自然文字系統

中文摘要

全球約有 9 千萬的兒童是視障者，他們的學習依賴著點字器材。點字書難以傳達 3D (三維)圖形的概念，讓他們在理解 3D 圖形有許多挑戰，而點字書過長的文字描述加大了他們與正常同儕之間的差距。本研究旨在開發一個學習輔助平臺，同時強化視障者的觸覺認知和文字圖形理解能力。

在強化觸覺認知方面，開發之系統能將傳統的图片，利用影像轉 3D 之人工智慧建模(Image-to-3D AI model)技術，轉換圖片轉為 3D，呈現在開發的擴增實境顯示器上，讓視障者能夠親身透過觸摸立體顯示器之 3D 模型，瞭解圖片表達的空間結構。在文字優化方面，利用微調大型語言模型與搜索式強化生成等方法，優化視障者閱讀之內容。系統開發過程中，二位視障者進行體驗，持續以修正系統設計的便利性、友善性與有效性。

本研究成果為視障者帶來新式個人化的學習輔助工具，增強視障者對立體圖形學習能力與文字理解。

English title

Developing an assistant platform for visual disabilities' learning: a 3D-AR tactile pad integrated Image-to-3d-AI-model and a personalized text generation system strengthen with Retrieval-Augmented Generation (RAG)

English abstract

Background and Study Aim:

The International Agency for the Prevention of Blindness (IAPB) estimates, globally, at least 90 million children live with some form of visual disabilities. The learning of those children with visual disabilities depends on traditional braille books or braille display. However, braille books fail to effectively convey complex 3D graphics. Braille display merely present the long contents from textbooks. Those impede the comprehension of visual disabilities' learning and widening the learning gap between the children with visual disabilities and their normal peers. The study aimed to develop a novel leaning assistant system that enhances tactile perception and textual image understanding for children with visual disabilities.

Method and Results:

For to better tactile perception, this study utilized the image-to-3D artificial intelligence (AI) model to transform images into interactive 3D data and then presented on the developed 3D display, namely tactile pad, after an Augmented Reality (AR) Unity system processing. Visual disabilities could physically engage with the 3D shape and understand the spatial structures concepts presented in images of textbooks. For text optimization, the use of large language models (LLMs) with Retrieve Augmented Generation (RAG) based reinforcement generation techniques was explored to assist visual disabilities in understanding complex textual information. Two visual disabilities were invited to experience during the device development and shared their opinions in convenience, friendliness, and effectiveness of the developed system.

Conclusion:

This study shows a new platform for visual disability as a personal learning assistant on metrics such as spatial understanding and text recognition.

內文

壹、前言

一、研究動機

聯合國在 2015 年提出了「永續發展目標」(Sustainable Development Goals, SDGs)，期望透過全球各國的共同努力，於 2030 年實現這些目標 (United_Nations, 2024)。其中，SDGs 的第四項目標明確指出：到 2030 年前，必須消除教育上的性別差距，並確保弱勢族群，包括身心障礙者、原住民以及弱勢孩童，能平等接受各層級的教育與職業訓練。SDGs 第四項目標的核心在於保證所有人都能獲得公平和優質的教育，並促進終身學習的機會。對於視障學生而言，SDGs 第四項目標的實現至關重要，因為這關係到視障者的權益保障和受教育機會的均等 (J. H. Zhang et al., 2022)。全球約有 9 千萬的兒童是視障者 (Bourne et al., 2021)，是不小的族群；為他們實現 SDGs 的理念，是非常重要的。

本研究團隊就讀之高中校園裡有一位使用白手杖的視障學生，時常引起同學們的注意。由於校內沒有特殊教育班級，這位視障學生必須在普通班級上課。教師的授課內容對於明眼的學生來說都感到吃力，研究團隊難以理解這位視障學生是如何應對如此複雜的學習內容。

儘管近年來，因特殊教育的介入以及法律的保障，視力障礙學生的教育已取得很大進展，但視障學生在學習現場仍然面臨許多挑戰。一次偶然的機會，本研究團隊鼓起勇氣與校內的視障學生攀談，發現他確實遇到許多困難。本研究團隊查閱了一份由高中數學教師撰寫的碩士論文，發現視障學生在數理相關課程中感到最為困難，尤其是圖形的辨識。對於先天全盲或缺乏立體圖形經驗的重度視障學生來說，數理科面臨的挑戰尤為明顯 (簡群恩, 2014)

以下圖 1 為例，明眼學生之課本插圖，能利用不同的顏色輕易區分每一線條代表不同的意義，透過色彩暗示能瞭解地球為球體，星體運行軌跡是橢圓形呈現出立體感，且解析度高，圖片上的資較充足，黃色太陽的黃道也一目了然。反觀視障學生要明白圖 2 描述的全貌，要逐一摸遍觸覺圖上所有線條，並在腦中拼湊這些線條結合起來的樣貌，實非易事。

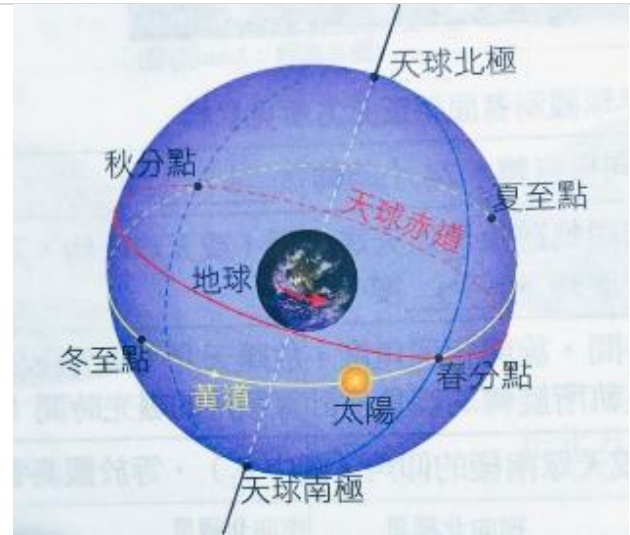


圖 1：太陽一年在黃道上通過春分、夏至、秋分、冬至點之循環。(圖取自龍騰高中地球科學課本)

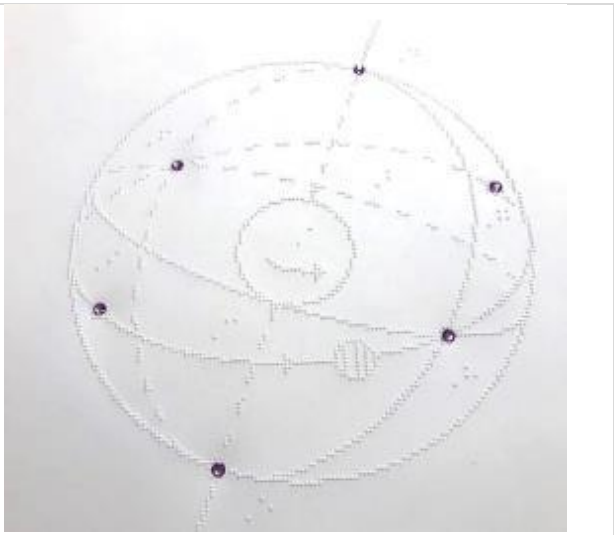


圖 2：盲人觸覺圖—太陽軌跡循環圖(圖取自中華民國無障礙科技發展協會高中地球科學課本)

本研究團隊經過文獻探討及與視障者訪談後發現，教材文字的強化對視障者極為重要。傳統教科書通常缺乏圖形方面的詳細解釋，導致視障學生在理解文字時難以與相應的圖形進行聯想和思考。另一方面，諮詢特殊教育老師後發現，個人化（personalization）以及解釋時的技巧在輔助和教導視障學生上至關重要。視障學生通常需要將新知識與既有經驗相連結，以便更好地理解新概念。在解釋新知識時，常需考慮個人的經驗、年齡以及熟悉的事物。舉例來說，在教導視障學生「水分子結構」時，特教老師提到，可以用生活中相似的物品作比喻，例如中間一顆較大的球，旁邊兩顆較小的球以筷子連接。這種圖形化的解釋方式能使學生更快速、有效地明白新概念和圖形。然而，傳統視覺輔具難以做到這點，只能依靠教師的額外補充來加速理解。此外將明視學生的教材，直譯在點字機上，讓視障學生閱讀會花上很長的時間。嘗試運用大型語言模型來增強文字與圖形的理解速度及效果，可能是加速視障學生理解文字的一種方法。

因此，研究團隊假設：研究團隊設法開發出一個綜合性的平臺，來加速視障學生學習。

二、研究目的

基於對公平正義的追求，為視障者提供更多元化的教學資源，確保他們能夠獲得與普通學生相同的教育質量和學習機會：本研究團隊期望能為視障者開發出一個智能平臺，同時以**立體輔具**幫助視障者更好地理解**空間概念**，並重新為文字教材通過**智慧改寫**，以加速他們對文字之理解。

本研究之 graphic abstract 如圖 3 與圖 4，

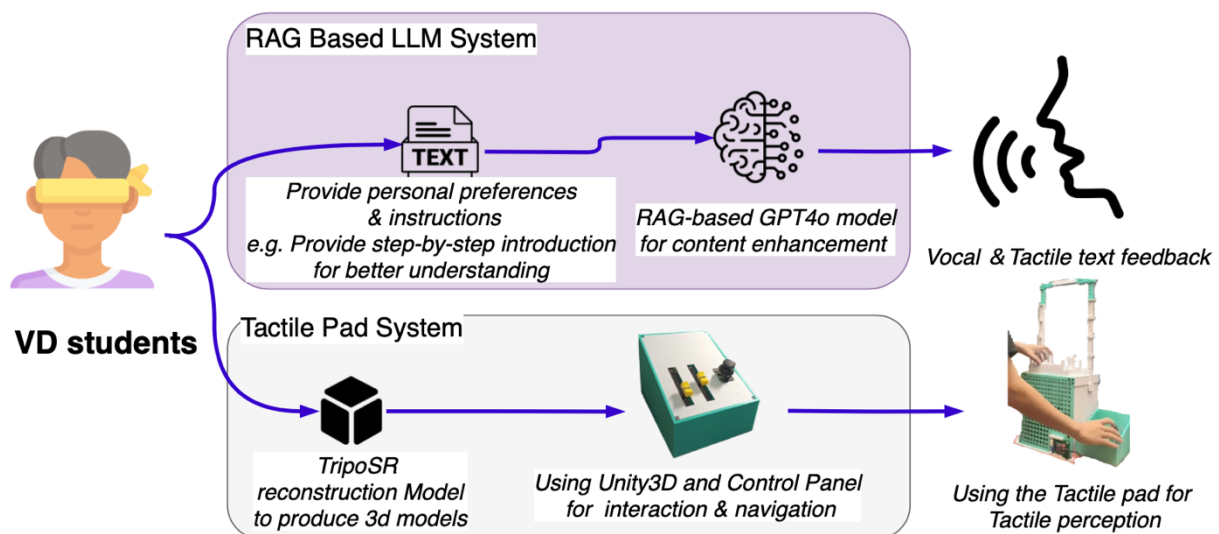


圖 3：本研究團隊開發之視障多元輔助系統（圖片來源：第一作者繪製），VD: visual disabilities (視障者)；RAG (Retrieval Augmented Generation) 搜索式強化生成語言系統；TripoSR 是一款 VAST-Reasearch 基於架構的 3D 重建模型工具；Unity3D 是由 Unity Technologies 開發的創建三維視頻之工具；GPT4o 是由 OpenAI 推出之 ChatGPT 優化版本。

為完成本研究目的，本研究訂有**四大研究目標 (Study aims)** 如下

- (一) 探討利用作動器陣列製作，將 2D 圖形以 3D 觸覺顯示之輔具呈現，協助視障學生提高空間概念的理解成效。

- (二) 研究利用擴充實境（AR）與文字增強之方式，使此系統對視障生更友善
- (三) 研究利用多元測驗方式測試視障輔具之成效
- (四) 比較此兩種方法對視障生學習與理解能力提升的實際成效 (如圖 4)

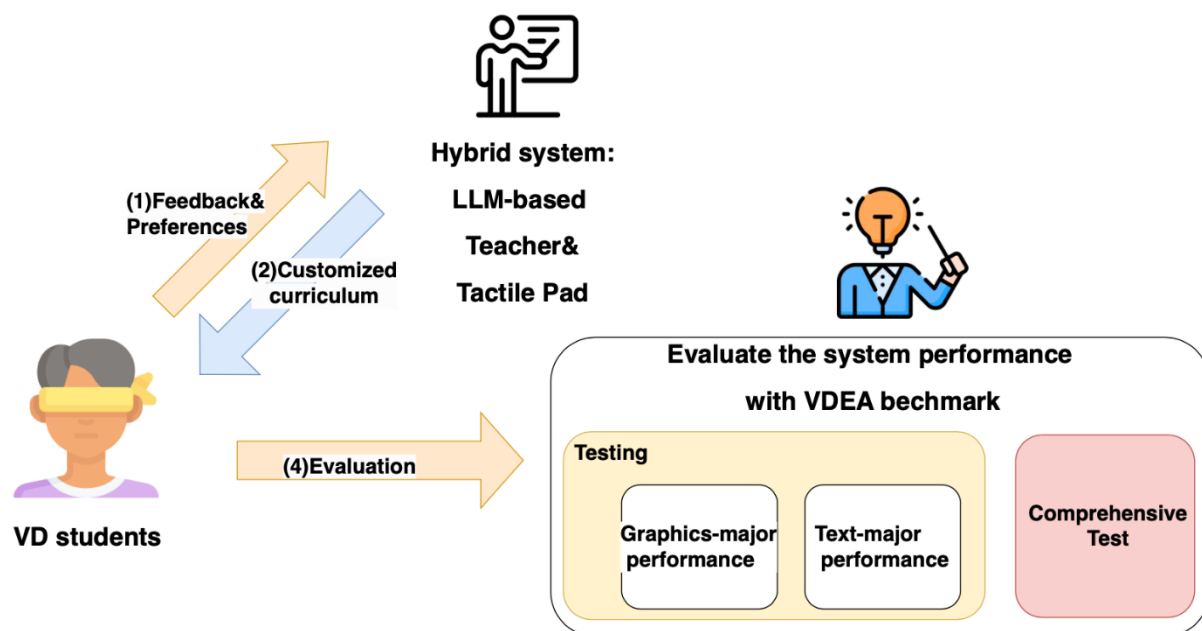


圖 4：視障者如何使用研究者提供的輔具之系統概述（圖片來源：第一作者繪製）

三、文獻回顧

（一）視覺障礙者之學習特點及所面臨之困境

1、視覺障礙的定義

依據《特殊教育學生及幼兒鑑定辦法》(全國法規資料庫, 2024)第四條，視覺障礙指經矯正後其視覺辨認仍有困難，致影響參與學習活動，鑑定基準依下列各款規定之一：

- (1) 遠距離或近距離視力經最佳矯正後，優眼視力未達 0.4。
- (2) 兩眼視野各為二十度以內。
- (3) 無法以一般標準化工具測定時，以其他醫學專業採認之檢查，綜合研判之。

2、視障者的感官知覺主要來源

Geruschat & Smith (1997) 指出，人類的感官中唯有視覺能快速與大量地獲取遠距離的訊息 (Geruschat, 1977)。徐鉅昌 (1987) 也提到人類各種感官知覺獲取訊息中，以視覺得到的訊息比例最高，其次依序為聽覺與觸覺及其他(徐鉅昌, 1987)。由此推知，若視覺功能不足或喪失後，訊息最主要的來源便是聽覺及觸覺。

視障者觸覺靈敏度雖然優於看得見的人，然而觸覺無法全面替代視覺的功能，乃因觸覺感受是由點到線到面；由局部再到整體，而視覺訊息的獲得是一目了然的，直接、快速、全面、且完整的進入眼簾。(萬明美, 2001)

3、視障者觸摸立體物件時的認知歷程

人類透過觸摸 3D 物體的感知，依賴於探索過程和觸覺提示的結合。該領域的研究表明，雖然個體擅長僅使用觸摸來識別物體，但理解 3D 模型內的空間關係需要更高級的交互作用(Coates & Bollegala, 2018; Lawson, 2014)。視障學生能否成功地透過觸摸而理解事物的樣貌，是視障學生學習複雜科目的關鍵所在。視障學生運用觸覺與聽覺來探索環境，再用心理想像統整，此種概念化過程較抽象，也容易感受不完整或有誤差。(方華罄, 2022)

4、視障學生之學習困境

在其他學生視力皆正常的班級內，教師非專門教授特殊學生之特教教師，講授對象也大多都是明眼學生，教師無法只以視障學生適合之方式授課。而視障學生感官功能亦無法全面接收教師提供的教學內容，導致學習出現落差。

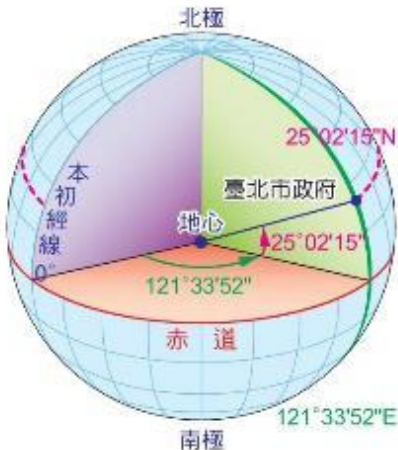
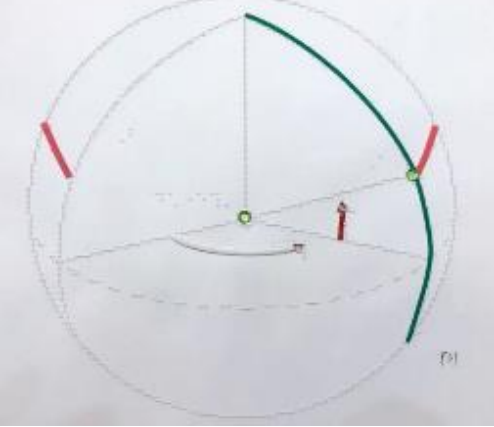
(1) 學習材料的可近性不同造成資源不平等

研究顯示學習材料的可近性與視障學生的知識獲取量有高度相關。方華罄(2022)提到，全國高中職以下的視障學生用書都是由教育部提供點字書，然而其他明眼學生使用的補充講義或練習卷，視障學生卻無點字版本可用。視障學生未能與明眼生一樣獲得補充資料或使用練習卷，導致學習資源不平等的現象。

(2) 學習與認知負荷的影響

視力障礙伴隨著知識獲取過程不易。視障學生經常面臨閱讀時間延長、瀏覽文本的能力有限、不易記筆記等等，導致視障者認知負荷增加且在課堂上落後。

舉例來說，圖 5 是高中地理課本中呈現臺北市政府的經緯度標記法，明眼人能一目了然地瞭解圖 5 表達的經緯度概念。對視障者而言，上課會使用搭配學校用書的觸覺圖冊，臺北市政府的經緯度標記法如圖 6 所示，由於缺乏明暗與光影的提示，經緯度也沒有全部呈現，視障者無法辨認圖片上的圓形輪廓是一個球體或是一個扁平的圓形，點字書考量到解析度問題，無法提供精細的經緯度以及文字資訊，因此很難理解經緯度的含意，相同的概念，視障者卻不適用與明眼人一樣的插圖來學習。

	
圖 5：臺北市政府經緯度的說明圖（圖取自南一版高中地理課本第一冊）	圖 6：臺北市政府經緯度的說明圖之觸覺圖（圖取自中華民國無障礙科技發展協會。南一版高中地理課本第一冊）

（二）現有視障者輔具未能有效協助視障者理解複雜空間圖形

保障視障學生的受教權益應包括全方位的支持系統，以確保他們能平等參與學習。專業人力的不足以及教材和教具的缺乏或不適用，都是視障教育長期以來面臨的挑戰（萬明美，2001）。

透過訪談及文獻得知視障學生在課堂上，面臨「理解立體形狀」和「空間推理」等與空間有關的概念，很難透過單純的說明或是使用傳統輔具便理解相關知識。這些困境也包括理解複雜的主題，例如幾何數學、化學中的原子模型和地質結構等等，這些空間圖像對於知識掌握至關重要。然而，學校目前難以提供有效的輔具協助教學，導致視障學生無法獲得公平的教育機會，成績與普通學生落差甚大。以高中化學課本中圖 7 表示的《元素分析裝置》圖為例，明眼人能立即辨識每項器材的材質是玻璃或是金屬，實驗流程也能在短時間內明瞭。然而視障學生使用解析度極低的觸覺圖（圖 8），並無額外的材質說明，視障學生以手指小面積的觸摸，再慢慢移動到其他部位，在腦海裡慢慢想像與拼湊全貌，整個過程如同患有低度近視的明眼人以管窺天，看不清楚且無法直接看到全貌如圖 9 所示。研究者將明眼學生所見之《元素分析裝置》圖與視障學生上課所用之觸覺圖做詳細比較，整理成表 1。

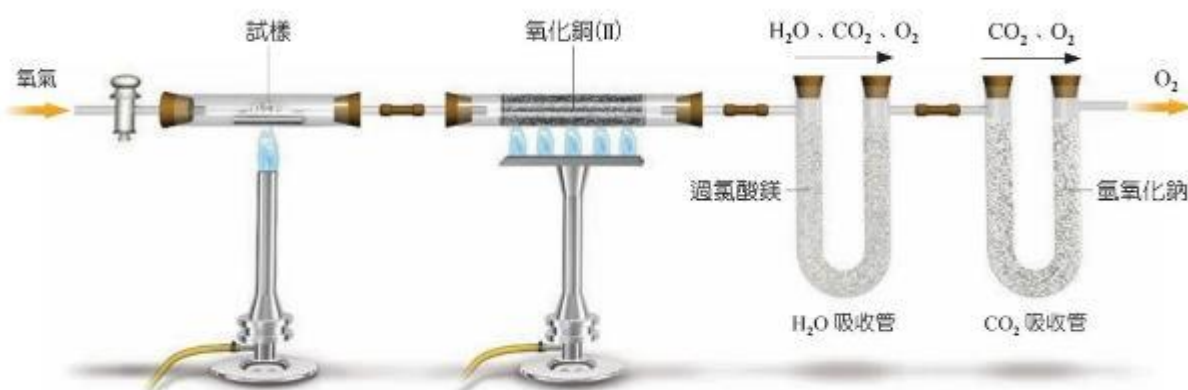


圖 7：元素分析裝置簡圖（圖取自翰林版高中化學課本）

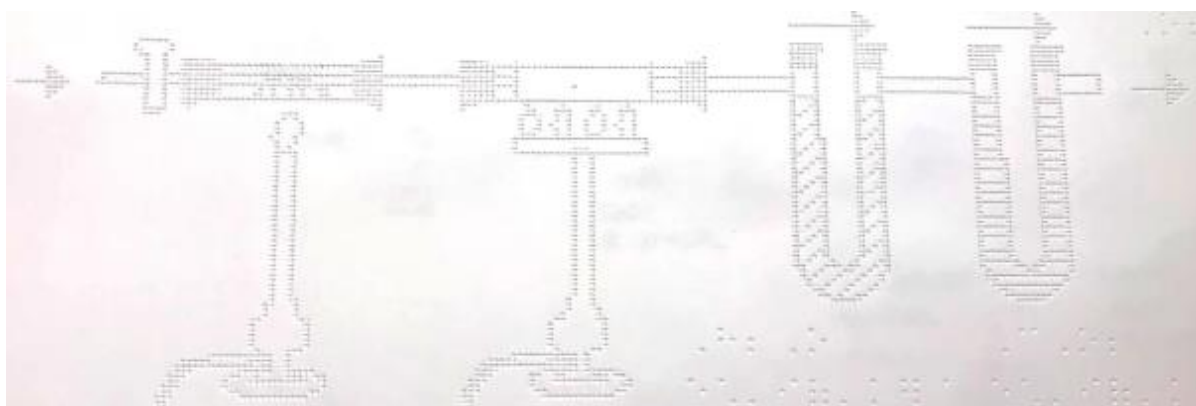


圖 8：元素分析裝置簡圖之觸覺圖（圖取自中華民國無障礙科技發展協會高中化學課本）

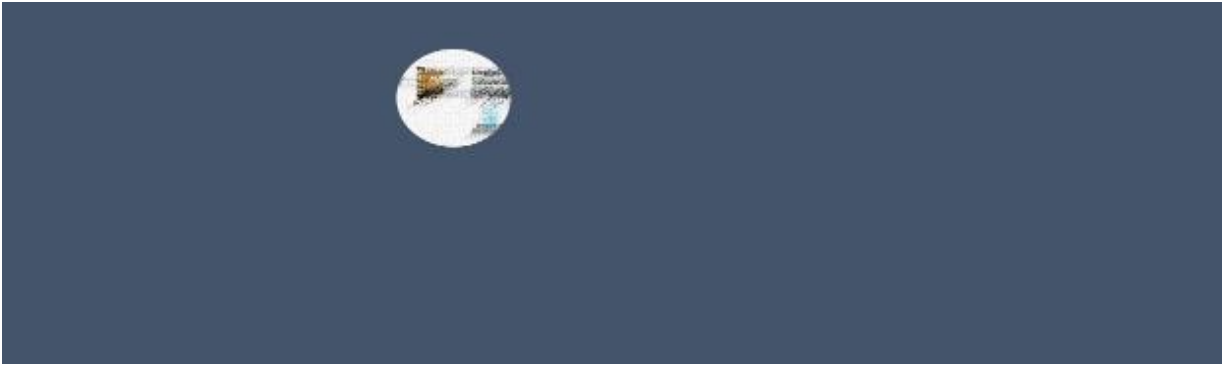


圖 9：盲人以手指觸摸解析度很低的圖，如同患有低度近視的明眼人以管窺天，看不清楚且無法直接看到全貌（圖片來源：第一作者繪製）

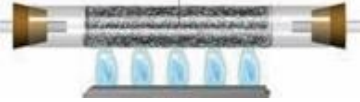
	明眼學生課本插圖	視障學生觸覺圖
		
圖片解析度	很高	非常低
細節資訊	很清楚	不清楚
認知過程	一目了然	以手指觸摸，每次只能感知很小的區塊，只能從點到線，再用心理想像統整全部樣貌。
認知時間	較短暫	較明眼人花費較長時間
立體感知	透過明暗、光影暗示，立即可知是圓筒狀	無，摸起來是扁平長方形
辨識火	立即	無法透過觸摸辨識，且點字圖上並無說明。
感知材質類別，如：玻璃、金屬、空氣	靠視覺經驗，立即辨識	無法透過觸摸辨識，點字圖上並無說明。
圖 7 的物體擺放概念流程圖	一目了然	以心理想像統整全部樣貌，易有誤差。
是否能學到完整概念	是	否，點字書提供資訊少很多，無法領會明眼人學到的知識內容

表 1：明眼學生課本之插圖與視障學生之觸覺圖比較表（表來源：第一作者整理）

根據衛生福利部社會及家庭署（2024）所提供的輔具資源，關於視覺相關的輔具類別僅提供點字相關設備，或針對尚存些微視力者使用的擴視機、放大機，或是有語音協助的手機、收錄音機等。（衛生福利部社會及家庭署，2024）

根據衛生福利部統計處（2024）「111 學年度全國特殊教育學校高中職學生學習輔具服務量統計概況」111 年度全台只有 8 位視覺障礙高中生，總共申請 10 次視覺輔具，而同一份統計

表中，「行動與擺位輔具」使用頻率最高，有 695 位高中生申請 974 次。視覺輔具的申請率不高，顯示目前專為視障者設計的輔助設備與實用性存在顯著差距。(衛生福利部統計處, 2024)

1、市面上提供之視覺型輔具類型

全國特殊教育資訊網（2024）資料顯示，目前學校使用的輔具，除數學用的盲用量角器、算盤以及有聲籃球與有聲排球等球類之外，大多與各科目專業學習較無相關，特殊教育所提供的輔具多侷限於將文字或視覺轉化為點字或語音輸出。衛生福利部社會及家庭署「輔具資源入口網」提供了兩種較先進的視覺輔具資訊，如表 2，一種是將檔放大的類型，一種是具有攝影辨識功能的輔具。

擴視機		
 圖片來源：衛生福利部社會及家庭署	主要功能	放大
	使用對象	低視力者獲取一般人日常的視覺資訊
	配備	將視覺資訊透過語音傳達給使用者
	機型舉例	桌上型擴視機-Clover Book XL
攝影辨識機		
  圖片來源：衛生福利部社會及家庭署	主要功能	辨識影像
	辨識內容	文字、顏色、環境、鈔票、通話、對話、在人群中也能找到認識的親友。
	使用對象	低視力者或全盲視障者
	配備	將視覺資訊透過語音傳達給使用者 「呼叫親友」功能可視訊通話，並透過即時影像，讓親友為您定位、尋人、尋物。
	機型舉例	救世主 AI 眼鏡

表 2：新型輔具介紹（資料來源：第一作者整理）

2、立體視覺輔具及動態資訊輔具獲取不易

傳統的觸覺圖形製作技術，如熱成型、壓模和 3D 列印輪廓等，對於基本形狀和紋理的傳達很有幫助。然而，對於複雜的 3D 空間概念的呈現，如「分子結構」、「幾何圖形」和「地球科學模型」，現有輔具往往呈現效果不佳。然而涉及動態資訊的觸覺輔具不容易製作，且通常需要複雜且昂貴的設備，限制了它們在典型教育環境中的可用性。

綜觀上述文獻，即便已有許多單位製作出許多不同的產品，但具有空間與立體感的教育用智慧觸覺輔具，目前還尚未在教育場域出現，激勵研究者早日開發出產品，保障視障學生的受教權。

（三）大型語言模型（LLM）在教育上的使用

將明視學生的教材，直譯在點字機上，視障學生閱讀會花上很長的時間。重新改寫明視學生的教材，可加速視障學生理解文字。目前，許多人工智慧的模型，可以加速視障學生的學習。

1、LLM 之強化方式探討

大語言模型（Large Language Model, LLM）或（Foundational Model）是指使用大量資料集和龐大參數進行訓練的模型。自從 ChatGPT 問世以來，LLM 的應用快速普及。然而，通用型 LLM 在特定領域的表現仍然有限。雖然它可以通過少數例子（Few-shot）進行情境學習（In-context Learning），但使用微調（Fine-Tuning）來修正模型的效果通常更佳。

直接使用 Fine-Tuning 對所有參數進行微調的成本相當高。因此，近年來研究者開始專注於更高效的微調方法，稱為 Parameter-Efficient Fine-Tuning, PEFT。由 Microsoft 團隊提出的 Low-Rank Adaptation, LoRA 方法，通過凍結原本的預訓練模型（例如 GPT-3）的權重，並搭配較小的模型進行微調，能夠實現良好的微調效果。這一方法類似於適配器

（Adapter）的概念：在凍結的 LLM 上，通過微調新增的小型網絡作為外掛程式。整體思路如圖 10 所示：在特定層中插入小型 LoRA 網絡，使模型能夠適應不同的任務。

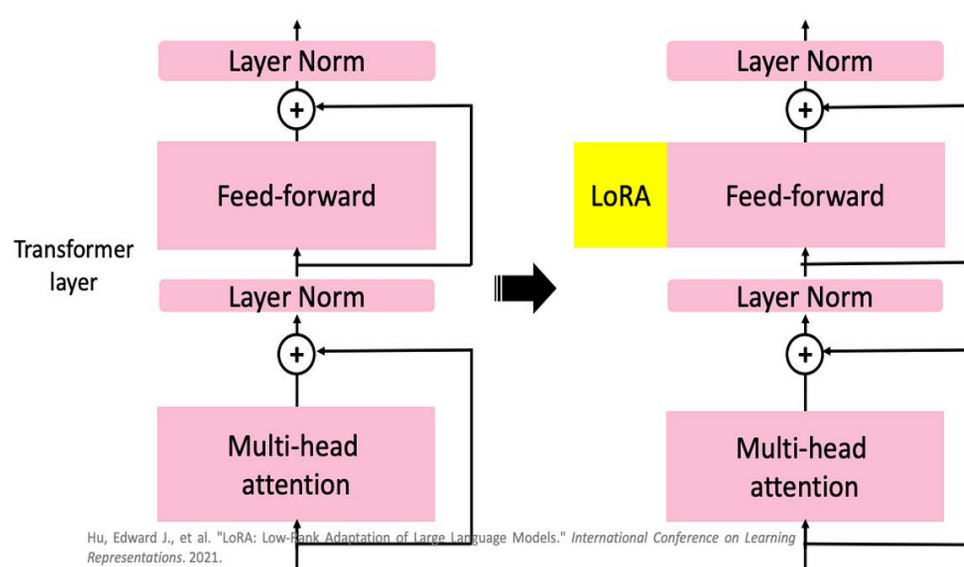


圖 10：利用 Lora 進行微調的架構示意圖；圖取自 (Hu et al., 2021)

2、RAG (Retrieval Augmented Generation) 搜索式強化生成

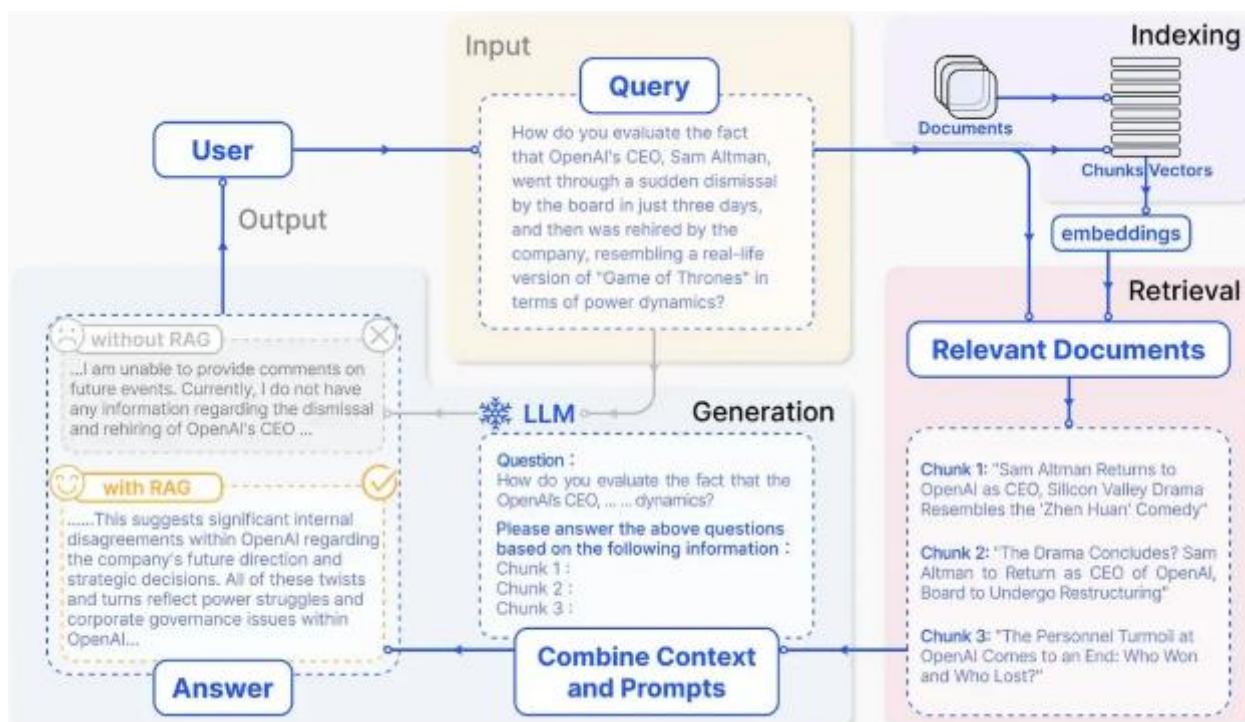


圖 11：搜索式強化生成的架構圖（圖取自 ChiChieh Huang，2024）

RAG (Retrieval Augmented Generation) 即「檢索增強生成」，RAG 是從外部來源獲取事實，以提高生成人工智慧模型準確性和可靠度的一種技術。

LLM 能以極快的速度回應一般提示，但無法滿足一些想要更深入研究特定主題的使用者。於是 Patrick Lewis 等人在 2020 年開發了 RAG，將生成式人工智慧服務與外部資源相連接，尤其是那些包含最新技術細節的資源。(Lewis et al., 2020)

RAG 有以下幾項特點：

- (1) 生成的文本更具可靠性：

RAG 提供可引用的來源，使用者可以核查資訊出處，使生成的文本更具可信度。

- (2) 具備擷取功能：

RAG 能從大量知識庫中擷取出最相關的檔，資訊來文可以是文章、報告或其他文本。

- (3) 具備生成能力：

生成器是 RAG 的關鍵之一，它能根據擷取到的文字生成內容，使生成的內容與擷取到的資訊保持一致，並能將這些資訊轉換為自然流暢的文字。

3、大型語言模型 (LLM) 在教育上使用之優點

GPT-4 和 Claude 等 LLM 的出現，引起了人們對其應用在教育領域的興趣，LLM 能產生適應性並且針對具體情況的解釋和評估方等面具有顯著的能力，亦可用於個人化學習體

驗，有助於學生取得更好的學習成果。此外，LLM 在自動評分和績效評估方面的有效性也得到了驗證，顯示出 LLM 可協助人類評估者的工作，同時具備可擴展性和適應性（蔣正漢和李鴻毅，2023）。

貳、研究方法與過程

一、研究方法

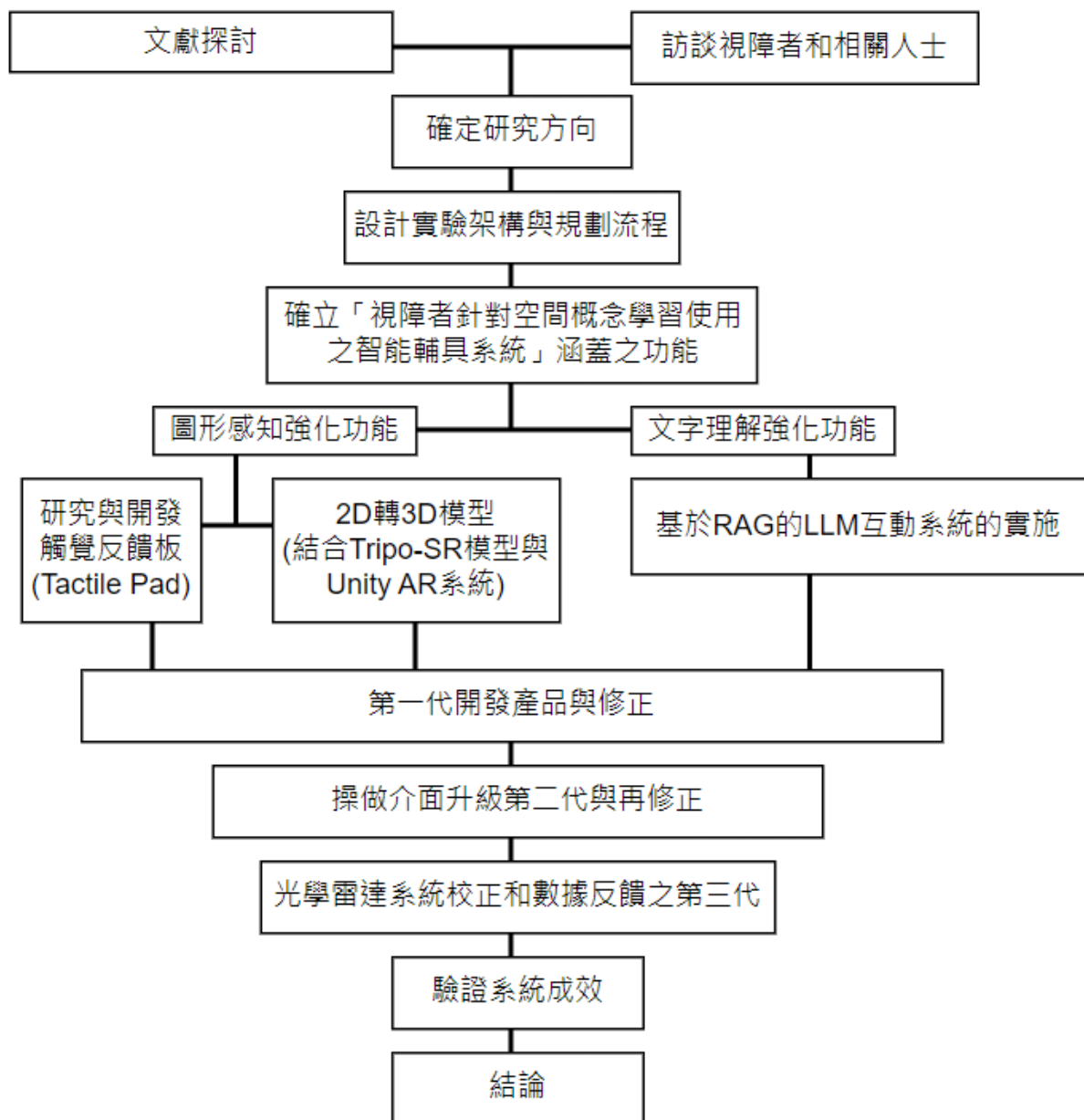


圖 12：研究流程圖（圖片來源：第二作者繪製）

本研究之研究過程如下：

（一）進行文獻探討及訪談視障者和相關人士以確定研究方向

有了初步構想後，第一步是進行文獻探討，以瞭解目前在視障教育中觸覺學習輔助工具所能提供的協助面向，以及 LLM 在特殊教育中的應用。在觸覺學習輔助工具的方面，本研究團隊探討了視障學生如何透過觸覺來理解抽象概念，並調查台灣現有視覺輔具是否能有效協助視障學生理解課程內容，例如點字、聲音播放機和觸覺圖形等輔具。本研究團隊調查市面上可用的輔助科技和觸覺顯示設備，如點字顯示器和觸覺反饋設備，以分析這些設備的優缺點，作為開發互動式輔具的參考。針對人工智慧輔助視障學生學習的方面，本研究團隊回顧了相關 AI 技術，例如語言模型和檢索增強生成（RAG）在教育中的應用方式，探討如何搭配 AI 技術提供視障學生有效的支援。為了更全面地瞭解實際教學中現有視覺輔具是否有效協助教師達成教學目標，以及滿足視障學生的需求，本研究團隊訪談了視障學生及特教老師教師，並對視障學生及家長進行了問卷調查及訪談，以分析視障學生在學習過程中所面臨的具體需求與挑戰。透過上述訪談結果與文獻資料，構思了研究目的，提出解決視障學生困境之方案。

（二）設計實驗架構與規劃流程

在完成文獻探討和及視障者學習需求分析後，針對研究目的進行設計實驗架構與規劃流程。此過程涵蓋了多項變數，包括實驗組與對照組的劃分、測試方案的設計，以及如何有效收集和分析數據。

本研究團隊欲提出**解決視障者需求之輔助設備**，實驗設計的核心鎖定能夠準確反映本研究團隊開發之學習輔助系統的效果，因此以「使用傳統的觸覺輔具」與「使用客製化的觸覺輔具 Tactile Pad 結合基於 RAG 的互動系統」進行比較。同時，實驗流程也規劃了多階段的數據收集，從學生的即時反應到最終的學習成果評估，並納入多元化的評量方法，包括多選題、開放性問答和實際操作等任務。

（三）確立「視障者針對空間概念學習使用之智慧輔具系統 VDE (Visual Disabilities Education)Hybrid Transformer」涵蓋之功能

VDE Hybrid Transformer 具有多功能輔助系統，包括圖形感知強化與文字理解強化兩部分，茲分述如下：

1、圖形感知強化之一：研究與開發互動式觸覺輔具—Tactile Pad

本研究團隊透過特教老師、班級教師與視障生的訪談結果，設定觸覺回饋裝置之開發重點為提供精確、反應靈敏的觸摸體驗，以滿足視障生對立體輔具的需求。進行相關技術評估與研究後，著手設計與開發觸覺反饋裝置：Tactile Pad。（Tactile Pad 在概念，也是一種立體顯示器。）

2、圖形感知強化之二：探討結合 Tripo-SR 模型與 Unity AR 系統

本研究團隊將 2D 重建為 3D 模型並與 Unity AR 互動，將結果回饋在 Tactile Pad，視障者透過觸摸 Tactile Pad 感知原來的 2D 圖形的 3D 模型，利用 Tripo-SR 模型，將課本內的 2D 圖像重建為 3D 模型，並在 Unity 3D 上處理與顯示。為了使視障生在操作上更便利，與視障生合作探討如何增進操作方式的便利性，例如使用滑桿、搖桿讓視障學生透過控制系統與 Tactile Pad 探索模型。系統使用 Unity 獲取高度矩陣，並將其傳送至觸覺顯示器，學生得以使用觸覺感知立體模型。

3、文字理解強化—基於 RAG 的 LLM 互動系統的實施

本研究團隊探討基於檢索增強生成（RAG）的 LLM，增強 GPT4o 在輔助視障生之能力，以提供視障學生更完善的文字學習內容。該系統可藉 Agent 由相關教育文獻取得回饋，

並從 prompt 瞭解學生之喜好與需求，從而調整教育內容以匹配個別學生的偏好和學習風格。關鍵的設計面向包括符合視障生個人化需求、使文字更易理解，幫助視障學生更容易瞭解課本的重要內涵。

(四) 視障者使用 VDE Hybrid Transformer 多功能輔助系統效能評估

1、圖形感知強化方面：視障學生使用 Tactile Pad 之效能評估

將視障學生分為實驗組和對照組。實驗組使用本研究團隊開發的學習輔助系統整合之 VDE Transformer 系統，包括 Tactile Pad 與 RAG-based GPT-4o，而對照組使用傳統的觸覺輔具，例如點字書籍和觸覺圖。課程進行後測驗學習成效，根據測驗結果進行比較分析。

2、文字理解強化方面：文字理解強化效果之評估

本研究團隊欲評估 Tactile Pad 是否能協助視障學生增強文字理解，對視障學生進行兩個不同版本的測驗：傳統版本和使用 RAG 技術強化與優化的視障生專用版本。視障學生不會被告知採用何種版本，以確保公正的比較。考試將包括多項選擇題和需要觸覺識別和 3D 概念理解的觸摸任務。施測的成果將衡量 Tactile Pad 在各種學習任務中相對於傳統方法的有效性。

另一方面，為驗證視障生在開放式學習與自學之能力，本研究團隊也探討了視障生在學習新知識的能力提升，分別利用 RAG 進行文字強化，其中包含利用相關視障輔助檔所提供之修飾方式對描述方式進行修正並利用 Step-by-step 的解釋方式使文本更容易理解。並探討結合觸覺板 Tactile Pad 之理解成效。

視障學生評估結果將呈現系統可用性、學習有效性和整體使用者體驗。這些評估將與考試結果相結合，以進行全面的數據分析，其中包括使用 GPT-4o 進行口述答案評分。

本研究使用 T-Test 比較實驗組和對照組的表現。

二、研究過程

本系統的開發旨在為視障學生提供適應性與互動式的學習體驗。系統的核心架構結合了開發**強化圖形感知**的觸覺反饋板（Tactile Pad）與**強化文字理解**的檢索增強生成（Retrieval-Augmented Generation, RAG）的大型語言模型（Large Language Model, LLM），並整合了將 2D 圖片重現為 3D 模型的 Unity AR 系統，幫助視障學生藉由動態反饋來探索教材中的圖形內容。

● 圖形感知強化系統

Tactile Pad 作為學生感知三維空間概念的主要介面，具備 256 個可動的圖元點，每個圖元點均能動態調整高度，呈現各種立體形狀與結構。當學生與觸控板進行互動時，系統會根據即時的回饋動態調整觸覺呈現的內容，從而提供學生多方面的學習體驗，幫助學生理解並探索抽象或複雜的概念。該觸覺強化技術具體利用 Triopo-SR 演算法，將 2D 平面圖形重建為可觸覺感知的 3D 模型，這些模型隨後動態顯示於觸控板上，確保學生能夠以直觀的方式理解立體結構。

● 文字理解強化系統

該系統基於 RAG 架構的 LLM 系統，當視障學生進行學習時，系統會檢索相關視障輔助資料，並生成符合學生需求的詳細回應。透過與 RAG 的結合，系統能即時提供精確且個性化的文字解釋，強化學生對學習內容的理解與掌握。

在系統的開發過程中，研究者根據視障學生及其教師的回饋進行多次反覆運算與優化。觸覺強化系統與文字強化系統的協同運作，確保視障學生能夠通過個性化的觸覺與文字反饋，有效地學習與理解教材內容，從而滿足其多樣化的學習需求。

（一）Tactile Pad 觸覺反饋系統的探討與開發過程

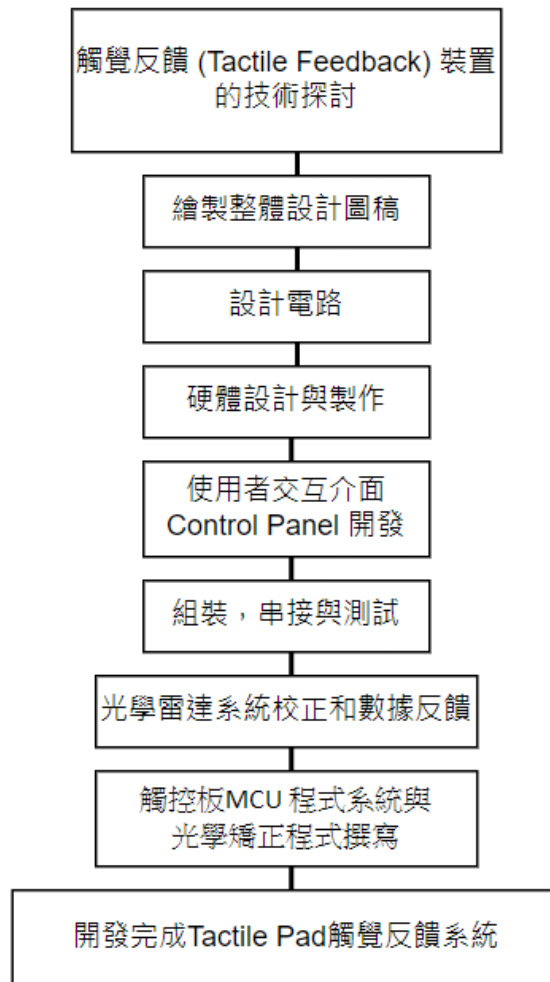


圖 13：Tactile Pad 開發流程圖（圖片來源：第三作者繪製）

Tactile Pad 是研究者系統中的關鍵組件，專為增強視障者的觸覺學習而設計。觸控板的设计重點在於通過觸覺回饋準確地傳達三維空間的概念。在開發過程中，研究者仔細考慮了可用性、精度和適應性，以充分滿足教育需求。

本研究團隊設計了一個具有 256 個圖元之觸控反饋板，能夠使視障生獲得圖形化觸覺反饋。

1、觸覺反饋 (Tactile Feedback) 裝置的技術探討

為了能夠重建擁有多點的三維模型，本研究團隊針對目前較有代表性的幾項技術進行分析，並比較各裝置的優缺點。

首先，現今大多數觸覺致動器使用偏心旋轉馬達 (ERM)、線性諧振致動器 (LRA) 或音圈

執行器 (VCA)，適合提供 100 至 300 Hz 的振動。因此，大多數觸覺設備透過不同頻率與幅度的振動傳遞訊息。這些致動器由電磁力驅動，使用剛性材料並需要多個移動部件，這不僅增加了複雜性，也使它們難以整合到穿戴式裝置中。此外，剛性材料與人體組織之間的機械阻抗差異導致能量傳輸效率低。

Artimus Robotics 公司提出了一種軟觸覺致動器如圖 14，可提供可控且自然的運動，實現廣泛頻率範圍內的觸覺回饋，並可根據需求定制尺寸與形狀(Acome, 2023)。

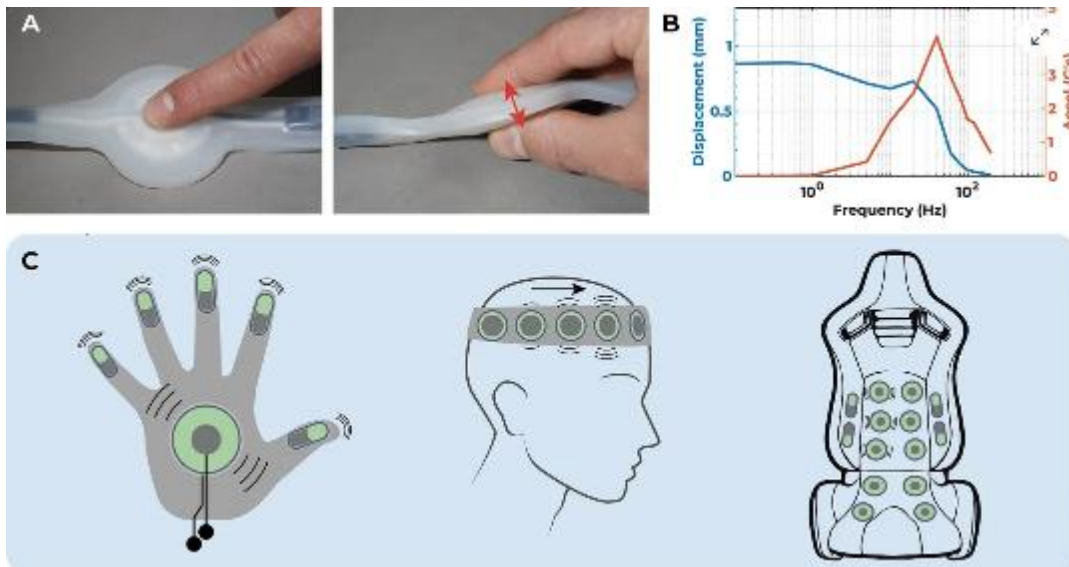


圖 14：A：相容的膨脹執行器為使用者提供觸覺回饋。B：Artimus 觸覺致動器的獨特優勢包括大位移和寬頻寬加速度。C：相容的致動器可以輕鬆整合到各種觸覺設備中，包括穿戴式裝置、座椅和其他提供觸覺的裝置。圖片取自 (Acome, 2023)

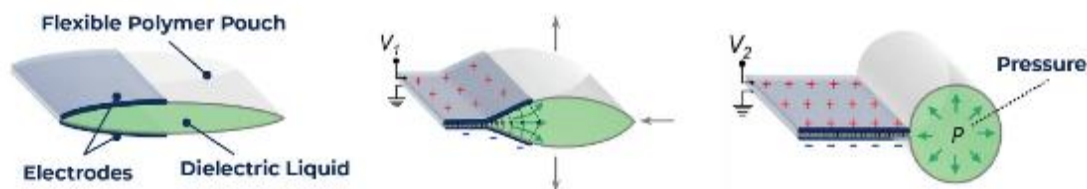


圖 15：：Artimus Robotics 的 HASEL 致動器的基本橫截面。此執行器由三個主要部件組成：柔性聚合物袋、介電液體和電極。電極覆蓋袋子的一部分，當電壓 V_1 施加到電極上時，袋子開始拉在一起。這會取代液體電介質，從而導致袋形狀的整體變化和袋內壓力的增加。透過改變施加的電壓來控制驅動位移和輸出力。在此所示的理想化示意圖中，在電壓 V_2 下，電極已完全拉在一起，並且袋子的一半呈圓形橫截面。（圖片取自 (Acome, 2023)）

另一種高強度、超出平面運動的致動器技術是結合微電機系統 (MEMS) 和微型剪刀運動放大器的裝置，稱為「MEMS-enabled microscissor motion amplifier」。此技術通常使用壓電材料 (PZT)，能在小型化裝置中實現大幅度的運動放大，並產生較大的力量，適用於精密控制和微機械系統。其優勢在於能在有限空間內提供強大的驅動力，並具備高度的靈活性和精度。(Xie & Livermore, 2015)

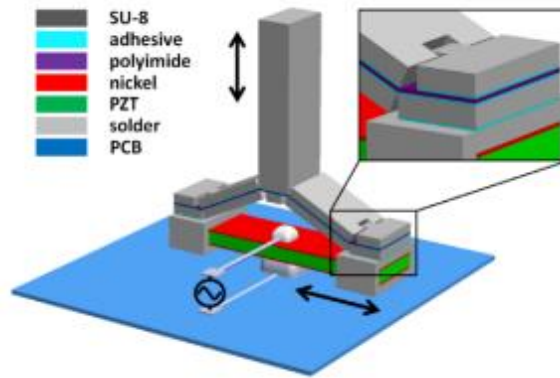


圖 16：觸覺陣列中單個致動器的示意圖。該致動器由一個面內壓電（PZT）致動器和一個層壓的微製造剪刀放大器組成，用於將小的面內振動轉換為 10 倍更大的超出平面的振動。圖片來源：(Xie & Livermore, 2015)。

還有一種名為「靜電黏附制動器」的技術（如圖 17、18），專門應用於高空間解析度的可刷新 2.5D 觸覺形狀顯示器。該技術利用靜電力來精確控制顯示器表面金屬針的位置，從而實現動態的形狀變化，為用戶提供逼真的觸覺反饋和立體感。透過快速刷新功能，顯示器能根據用戶的觸摸即時改變形狀，從而帶來細緻而真實的觸覺體驗。這項技術廣泛適用於虛擬現實、增強現實以及人機交互等領域，其高解析度特性使得裝置能夠呈現精細的觸覺細節。（K. Zhang, González, Guo, & Follmer, 2019）

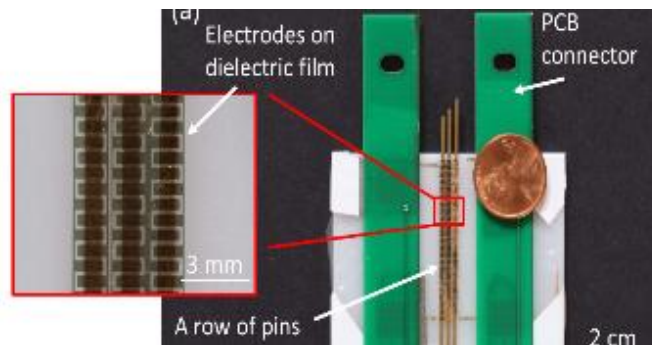


圖 17：(a) 一排針腳，間距為 1.7 毫米。兩個薄的印刷電路板（PCB）用作指狀電極和主控制 PCB 之間的連接器。插圖顯示了模式化的電靜態粘附制動器的特寫。圖片來源：(K. Zhang et al., 2019)

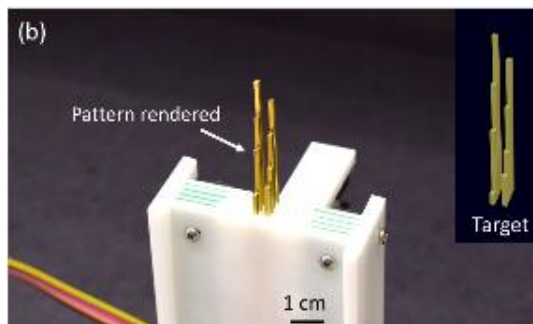


圖 18：(b) 由形狀顯示原型渲染的形狀圖案。四排金屬針腳被緊密排列在一起，形成一個 4x2 的形狀顯示單元。每個針腳都由一個電靜態粘附制動器獨立控制。圖片來源：(K. Zhang et al., 2019)

	螺旋裝置（螺旋柱與螺帽）	軟觸覺致動器（Artimus Robotics 所開發）(Acome, 2023)	A high-force, out-of-plane actuator with a MEMS-enabled microscissor motion amplifier（P Z T）(Xie & Livermore, 2015)	Electrostatic Adhesive Brakes for High Spatial Resolution Shape Displays(K. Zhang et al., 2019)
Force	強	中	強	弱
速度	慢	快	快	中
作動範圍	大	小	非常小	大
Price	中	中	中	中
結構與驅動複雜度	小	中，需要薄膜結構	高，需要高電壓控制	中，需要電容薄膜裝置
單裝置大小	小	大	小	小
危險性	極小	極小	大，需要高電壓	極小

表 3：觸覺反饋 (Tactile Feedback) 裝置技術之比較表（表格出處：第一作者整理）

在決定使用何種作動器方面，首先研究者排除了使用壓電陶瓷之作動器，因為其作動範圍極小。研究者亦排除了使用液壓系統的軟性觸覺裝置，因為考慮到單一圖元的體積與面積較大以及結構較複雜，難以呈現出較高的解析度。

最後考慮到實際實施與設計方面，**研究者最終選擇採用螺旋裝置**，搭配**微型馬達**進行設計。

2、繪製整體設計圖稿

Tactile Pad 的初始設計強調對視障者友善的介面，其表面能夠依據視障生的操作即時更新呈現之形狀，使視障學生能夠觸摸形狀和高度的細微變化，增強視障者理解課本裡複雜的三維概念及圖形之能力。

本 Tactile Pad 的設計受到 Pin Art (圖 19) 的啟發，視障者能夠透過觸摸觸覺圖形來感受三維物體。本研究團隊因此設想出一套能讓使用者互動並探索三維圖形的輔助裝置，進一步提高視障學生學習複雜概念的效率。

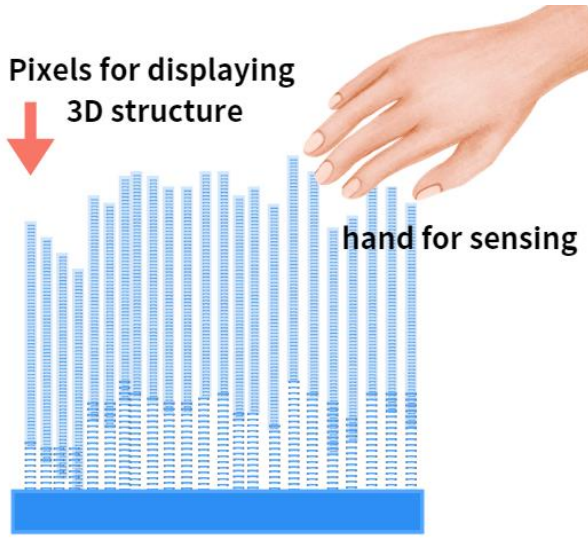
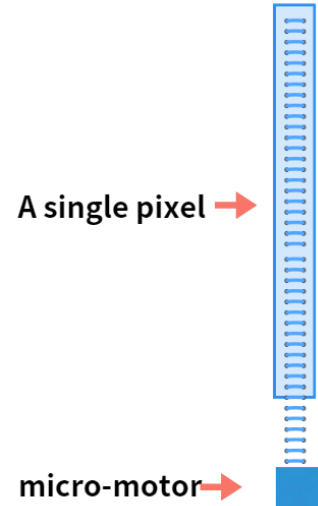


圖 19：研究發想來自 Pin Art。圖片取自 DottorGadget 網站。

透過結合不同的凸起的表面和可以模擬 3D 物件感覺的動態輪廓。觸控板會產生與不同幾何形狀相對應的獨特圖案，視障生能夠透過觸摸來概念化 3D 結構，使空間概念更容易理解。為此研究者在這方面進行了初步的構思以及設計。

研究者預期視障生與 Tactile Pad 的互動方式如圖 20。單一作動器的設計 (Pixel) 部分如圖 21 表示。本開發項目預計使用微型馬達，並且透過螺栓將旋轉運動轉換成垂直運動如圖 22。圖 23 展示了如何將作動器與電路板結合。而圖 24 表示將此電路板整合到觸覺顯示器的擺放情況以及連結方式。

研究者在與視障生討論後，研究者受到 PCIE (圖 25) 介面的啟發，決定電路板採用模組化設計。因此可以依據視障者學生的需求進行圖元擴充，以利用更多的作動模組擴充觸覺面積。

 Tactile Pad with movable pins for feedback	 A single pixel micro-motor
圖 20：視障生與 Tactile Pad 的互動方式預想圖 (圖片來源：第一作者繪製)	圖 21：單一作動器 (Pixel) 的設計 (圖片來源：第一作者繪製)

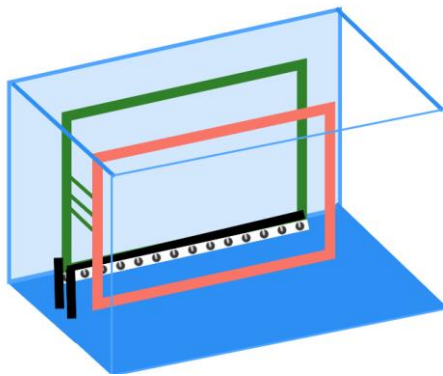


圖 22：微型馬達能透過螺栓將旋轉運動轉換成垂直運動（圖片來源：第一作者繪製）

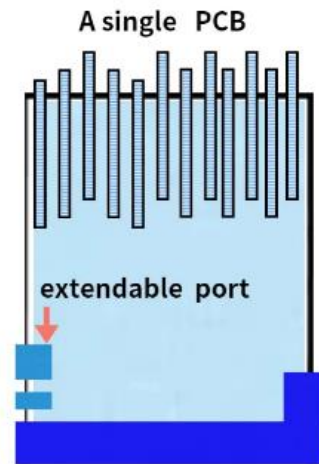


圖 23：作動器與電路板結合之構圖（圖片來源：第一作者繪製）

Take inspiration from PCIe slots

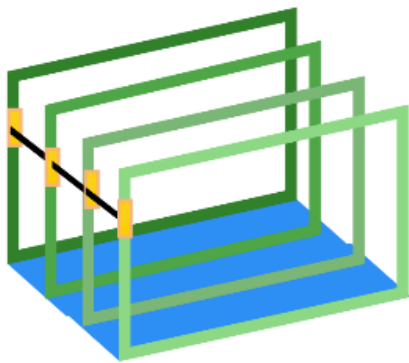


圖 24：電路板整合到觸覺顯示器的擺放情況以及連結方式由, PCIe 的模組化排列啟發（圖片來源：第一作者繪製）



圖 25：PCI Express 介面允許將主機板擴展到其預設的 GPU、網路和存儲配置之外。（圖片取自凌威科技網站）

3、設計電路

在 PCB 設計方面，研究者採用了主從架構（Master-slave architecture）以及模組化電路板。這種設計為未來視障生擴充顯示面積提供了可能性。並且研究者使用為移位暫存器 ShiftRegister 進行訊號延伸以及馬達控制。

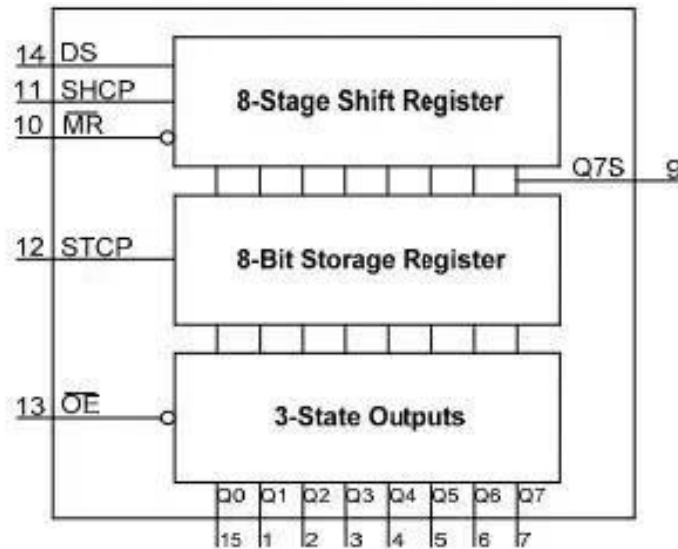


圖 26：移位暫存器 74HC595 的正確功能框圖，SHCP、主複位（MR）和輸入 DS 連接到 8 級移位暫存器。（圖片取自 The engineering projects 網站）

74HC595 是一種 8 位元串列進並列出的移位暫存器（Shift Register），常用於擴展微控制器或 Arduino 的輸出腳位數量。它能將微控制器的少量 I/O 腳位轉換為多個並列輸出，適合用於控制多個 LED、七段顯示器或其他需要多個數位元訊號的元件。

利用 74HC595 巧妙的特性，將數個 74HC595 進行串接構成一個觸摸控制模組，每塊觸摸模組有 8 個 74HC595 進行串接，並且在每個 74HC595 下都裝載 4 個馬達驅動模組（L9110S），達到每個觸摸模組(One PCB) 有 32 個可動圖元圖元。在設計完成後，使用麵包版進行電路 Prototyping，在完成雛形的電路實測之後本研究團隊進行電路設計，將雛形的電子回路與晶片繪製成電子版本如圖 27，最後將其轉換為 PCB，並進行 PCB layout 完成最終 PCB 設計，以及 3D 預覽圖。

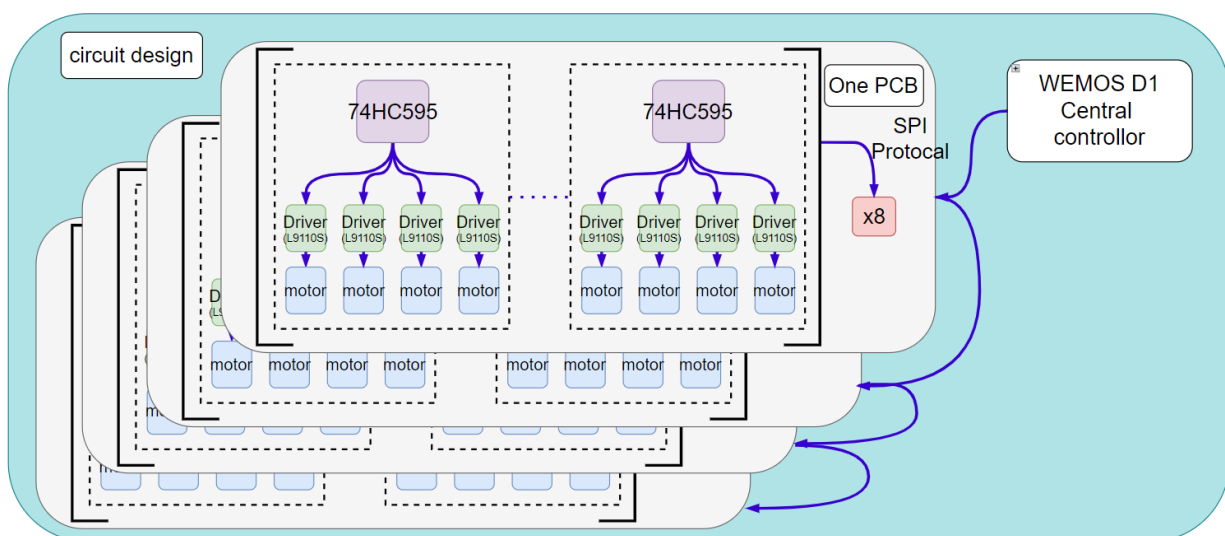


圖 27：利用 74HC595 與 L9110S 所建構之模組化控制模組系統示意圖。圖片來源：第一作者繪製

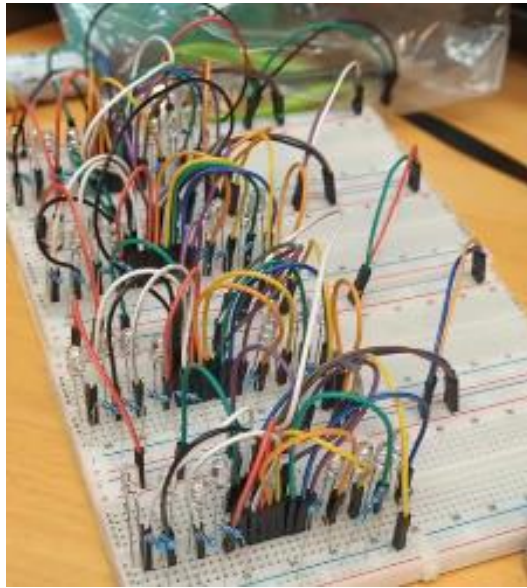


圖 28：本研究團隊使用麵包版進行電路 Prototyping（圖片來源：第一作者拍攝）

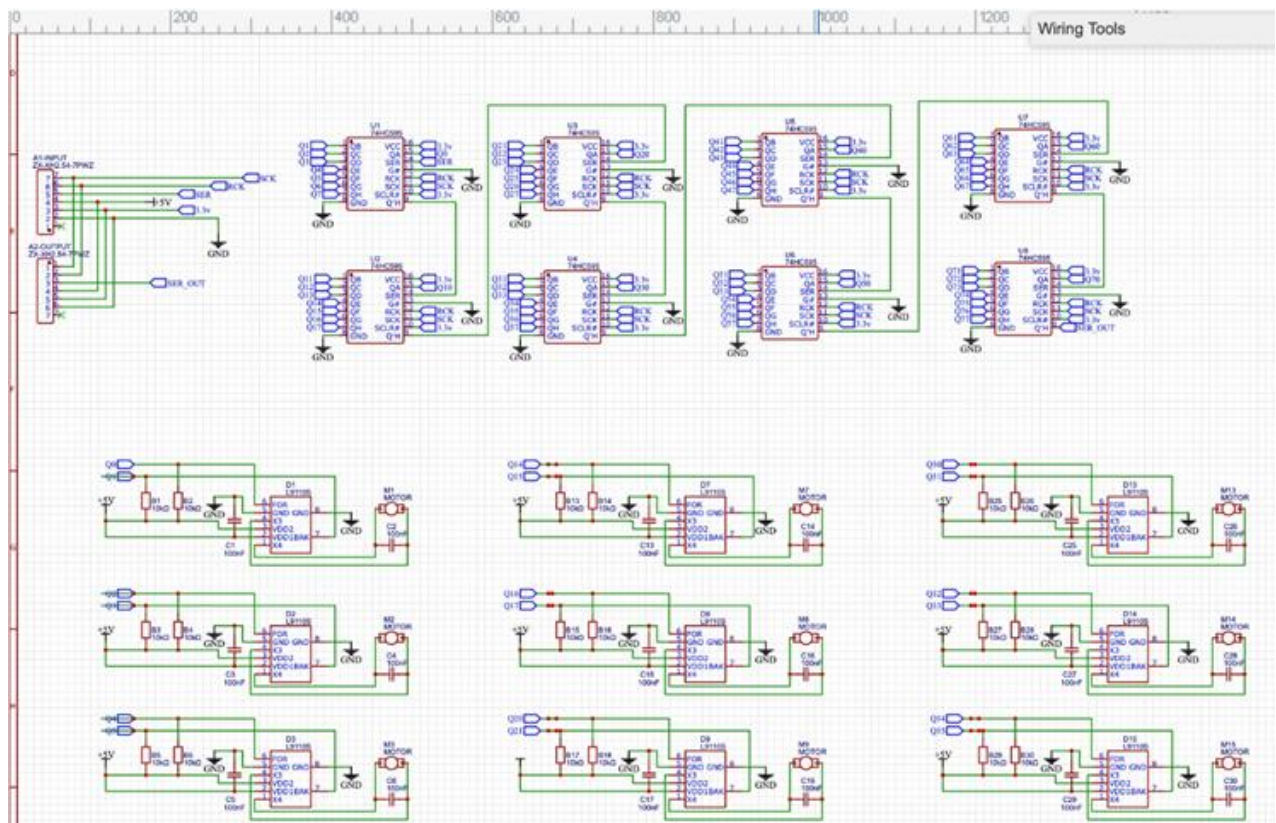


圖 29：本研究團隊使用 KiCAD 繪製電路電子版本（圖片來源：第一作者繪製）

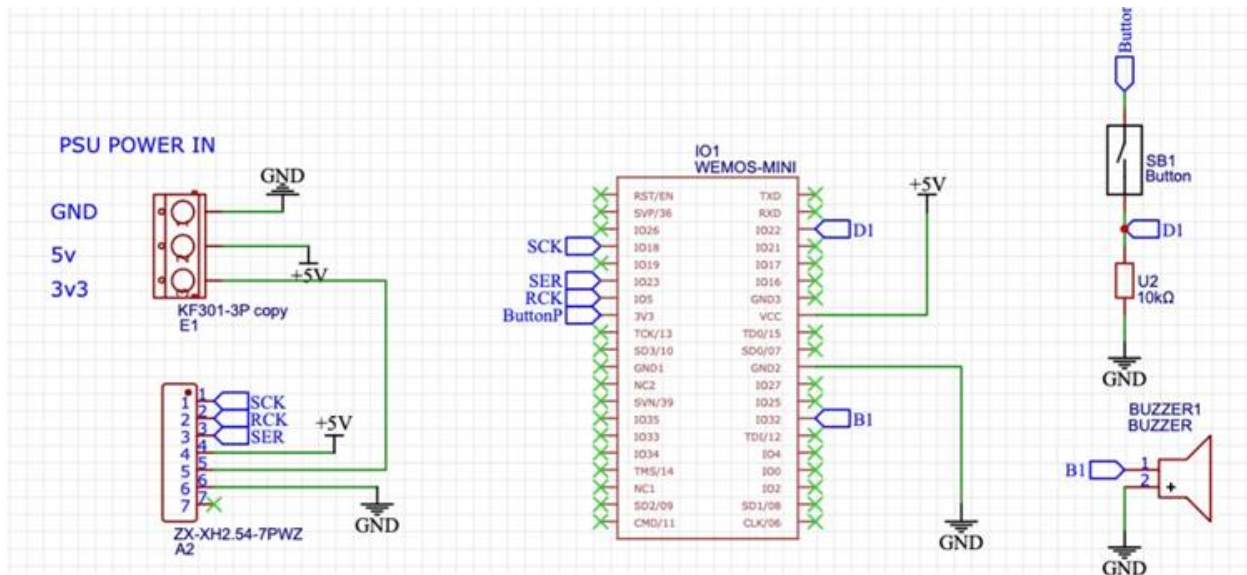


圖 30：研究者使用 KiCAD 繪製電路電子版本（圖片來源：第一作者繪製）

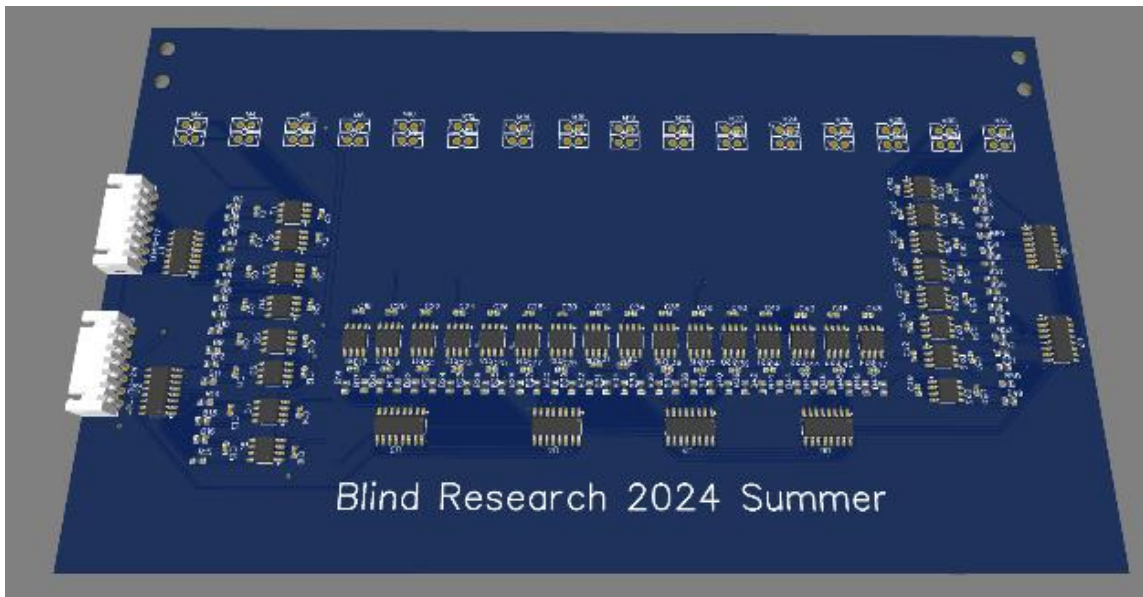


圖 31：研究者繪製之 3D 預覽圖（圖片來源：第一作者繪製）

4、硬體設計與製作

Tactile Pad 觸覺反饋系統的硬體製作，本研究團隊使用 Fusion 360 設計，並使用 3D 列印機列印。

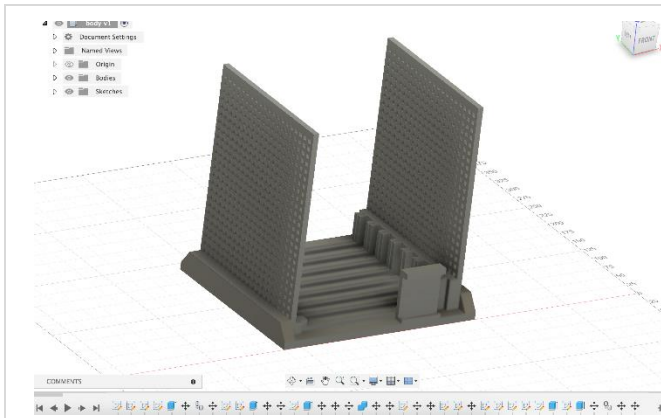


圖 32：Tactile Pad 之外殼設計
(圖片來源：第二作者繪製)

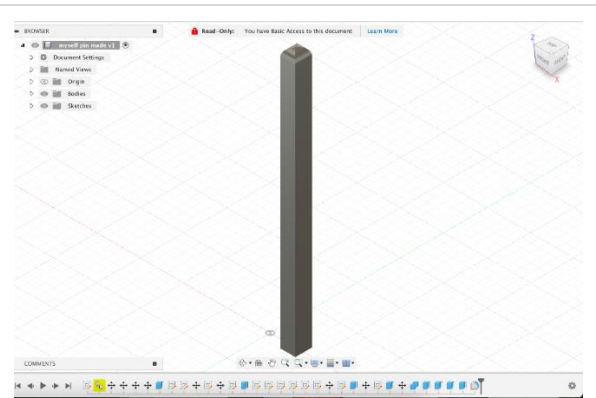


圖 33：Tactile Pad 基本單元元件
(圖片來源：第二作者繪製)

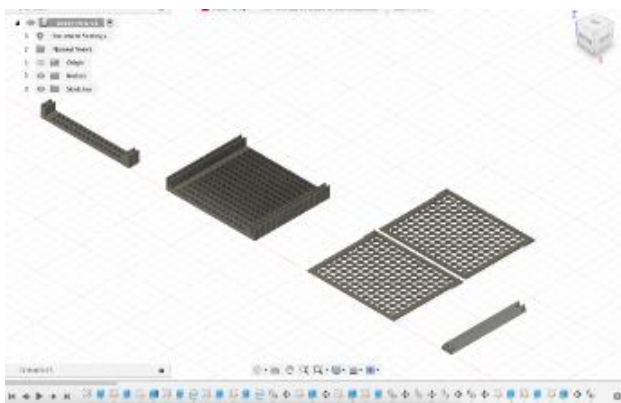


圖 34：Tactile Pad 基本單元之插槽與定位欄框 (圖片來源：第二作者繪製)

5、第二代-使用者交互介面 Control Panel 開發

本研究團隊起初設想視障生可以用電腦鍵盤控制觸覺版，但是與視障生持續討論後發現使用滑桿與搖桿能夠更好地多角度探索模型。為了使視障者生能夠更好的操作 Tactile Pad，決定改變操作方式，於是開發第二代 Tactile Pad。研究者利用 MQTT Protocol 與 PC 端進行溝通，並與 Tactile pad 互動。本研究團隊以觸控面板為開發主軸，設計了符合使用者直覺的操作系統和交互介面。關於此系統的控制邏輯以及程式也會在第二部分進行詳細紀錄。

(1) Control panel

本研究團隊欲使用滑桿 joystick 和搖桿 slider 透過 unity 操作模型，將互動數據透過 mqtt 由操作面板發送到機器，以達到控制物件旋轉縮放的效果。

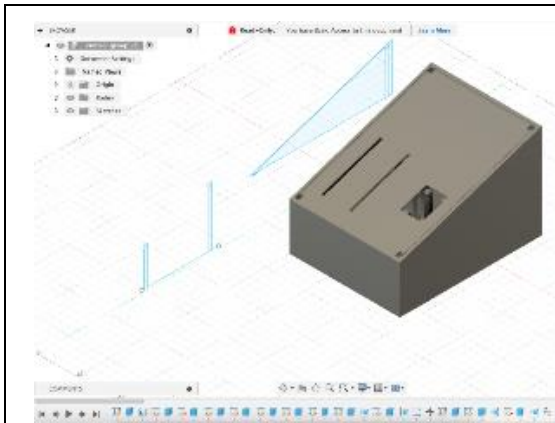


圖 35：滑桿和搖桿控制器設計圖
(圖片來源：第二作者繪製)

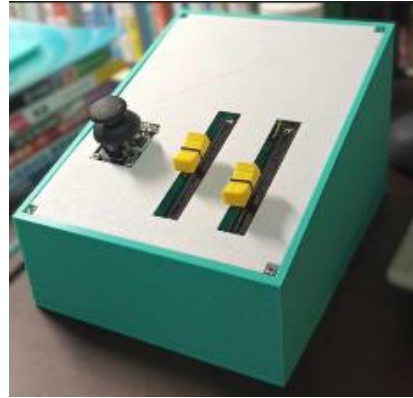


圖 36：滑桿和搖桿完成品
(圖片來源：第二作者拍攝)

(2) 組裝，串接與測試

所有硬體都製作完成後，使用者進行組裝，串接與測試。



圖 37 進行組裝，串接與測試
圖片來源：第二作者拍攝



圖 38 由 PCIe 插槽啟發的
電路板結構設計
圖片來源：第二作者拍攝

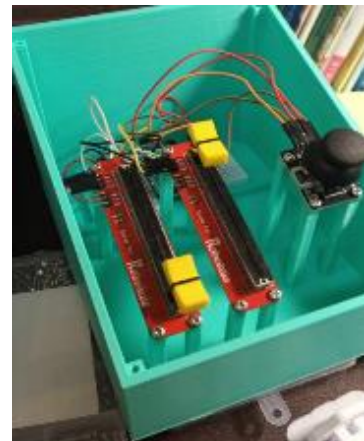


圖 39 進行組裝，串接與測試
圖片來源：第二作者拍攝

6、第三代-光學雷達系統校正和數據反饋

(1) 光學校正系統研發及實現

考慮到每一顆微型馬達的個體差異以及隨著時間而產生的誤差，本研究團隊採用了深度相機進行高度矯正。使用此陣列深度相機的優勢在於未來研究者或視障學生想要對圖元的排列方式進行改變時，不需要增加矯正所需的成本。在顯示器運行的過程中，若沒有外部偵測器對機器的顯示做讀取，並且將結果回饋給機器，讓機器分析結果並執行校正，即控制電路，那麼誤差將會隨著時間而成長。分析和權衡了多種偵測器及偵測類型後，決定以深度相機的立體視差和特色偵測(Feature Detecting)作為校正系統主軸。

藉由立體視差(Stereo Depth)深度相機，濾除環境噪訊後，可以藉由三角性質和視角變換(perspective transformation)，由特色偵測(Feature Detecting)定位點後，運算得到深度相機和物

體之間的距離，而重建並輸出機器所顯示的的高度矩陣，並分析偏移量和執行校正。下圖是立體視差測距法的運作方式：

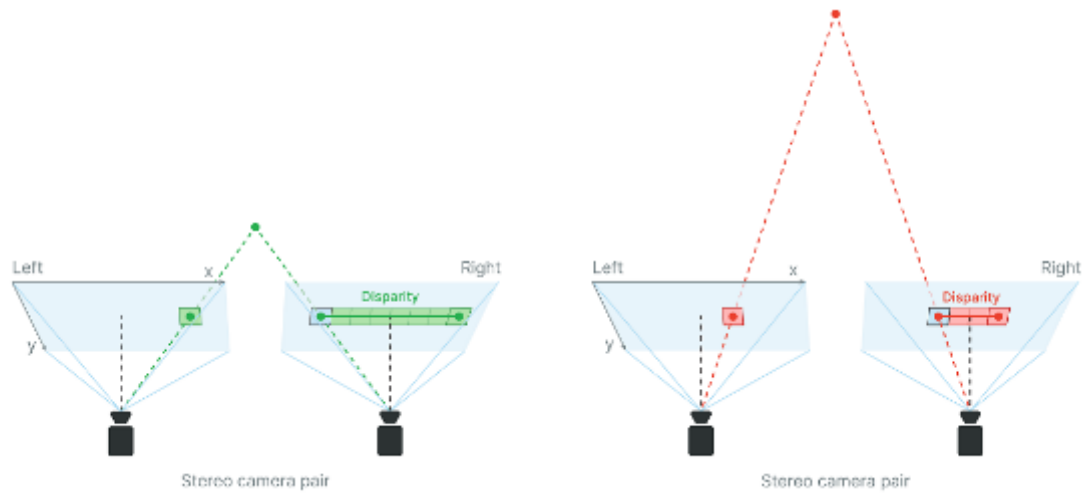


圖 40：視差深度說明圖

其中深度的計算公式如下：

$$\text{depth [mm]} = f_x[\text{px}] \cdot \frac{\text{baseline [mm]}}{\text{disparity[px]}}$$

depth_cm：深度（以公分為單位）

fx_px：焦距（以圖元為單位元）

baseline_cm：立體相機對中兩台相機之間的距離

disparity_px：圖元差異

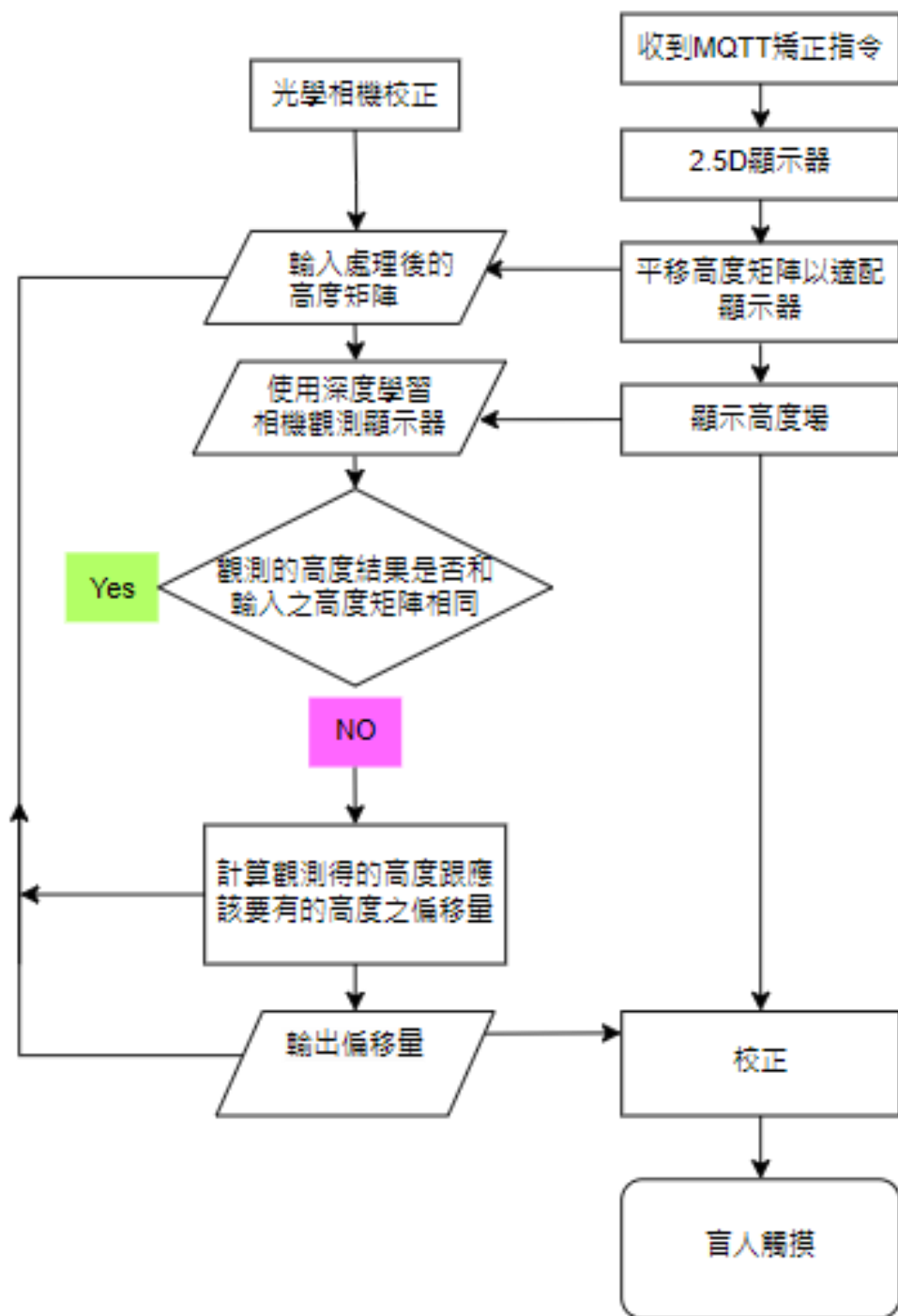


圖 41：矯正系統控制邏輯（圖片來源：第三作者繪製）

在確定使用的矯正方式後，研究者對矯正系統之架構進行了規劃，如圖 41，並完成了硬體架構圖，以及研究者完成系統的整體配置如圖 42。

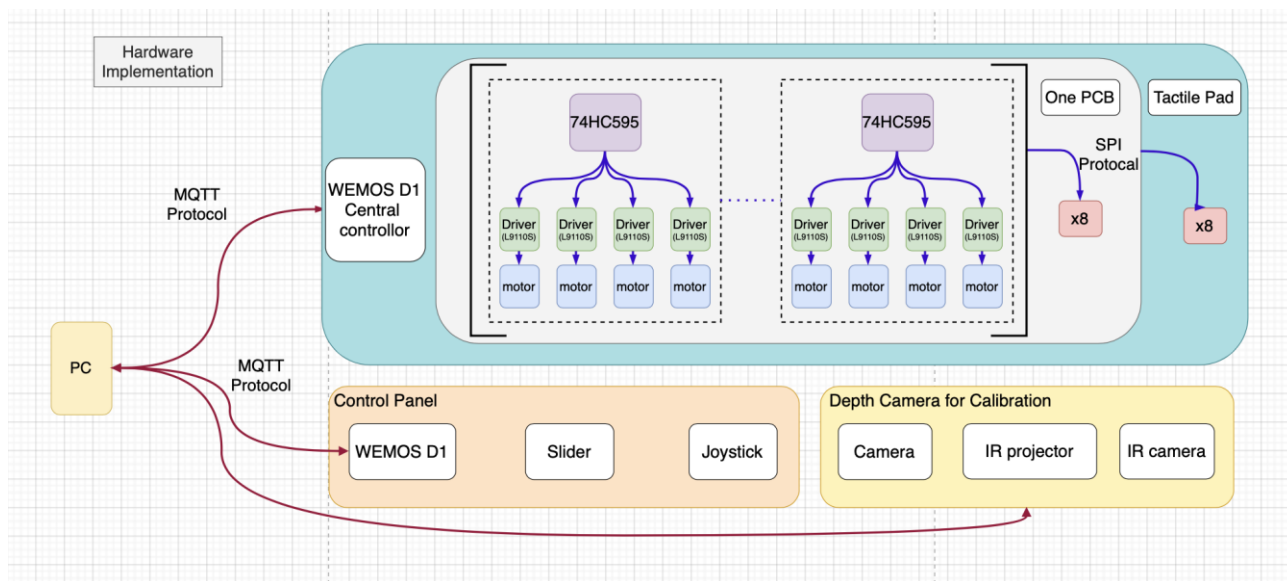


圖 42：最終的硬體配置圖（圖片來源：第三作者繪製及拍攝）

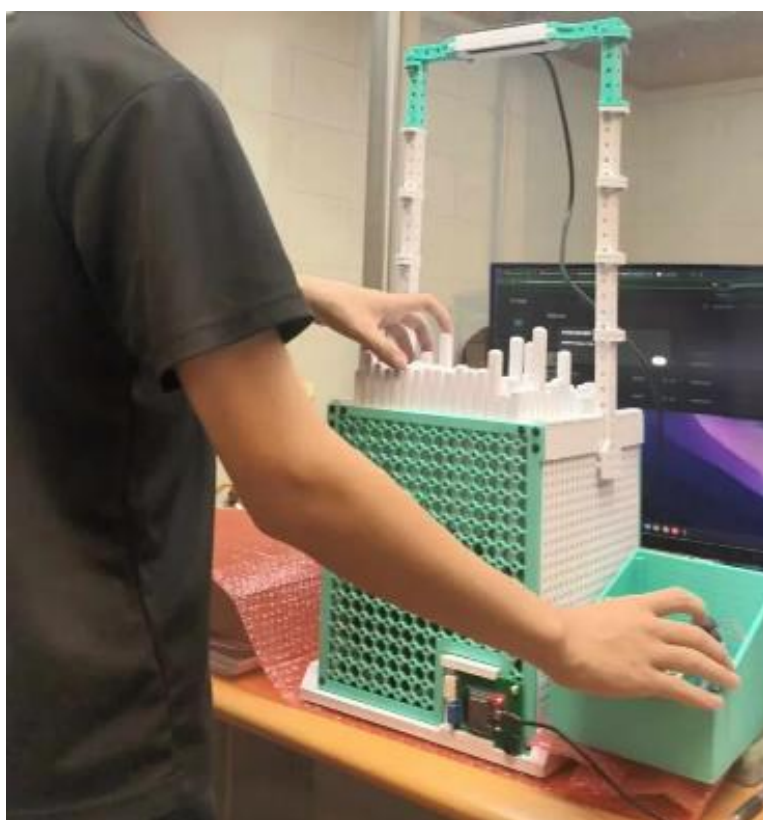


圖 43：將矯正系統與先前探討之系統結合之完成圖（圖片來源：第三作者拍攝）

(2) 觸控板 MCU 程式系統與光學矯正程式撰寫

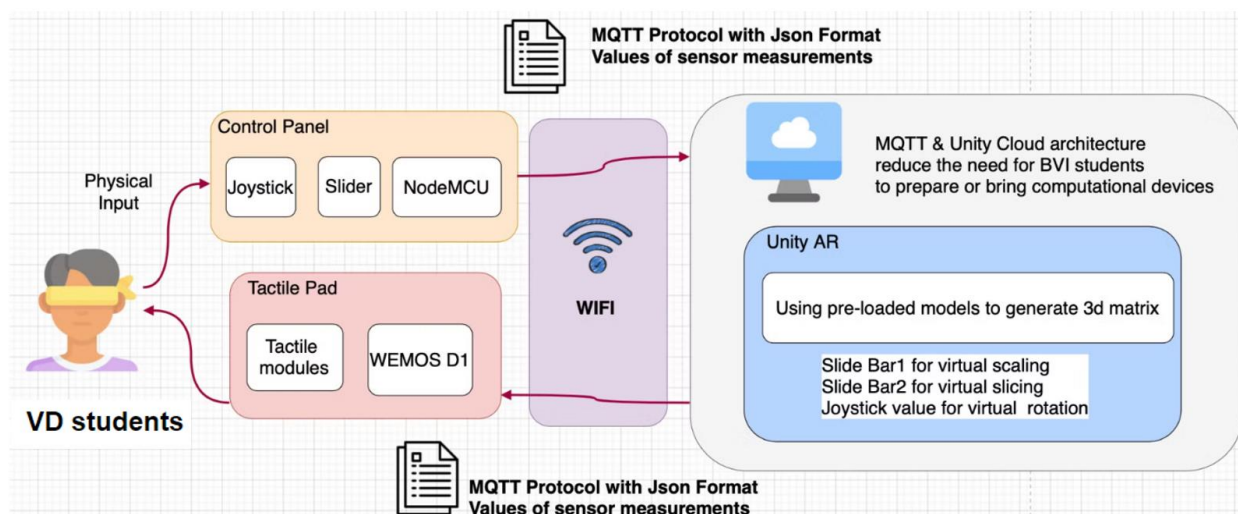


圖 44：本系統通信圖與軟硬體架構（圖片來源：第一作者繪製）

在軟體部署的部分，第三代的硬體下研究者探討了使用雲端運算架構的可行性。利用雲端架構，視障生在使用本系統時能減少使用本地運算資源，例如使用筆記型電腦與 GPU 時產生的額外成本與不便性，提高了實際在學校應用的可行性。

Control Panel 的設計利用一系列可變電阻，預處理後將資料以 Json 格式搭配 MQTT protocol 傳輸至 Unity 端。其中，Unity 會將預先用 TripoSR 模型重建的 3D models 如下圖的方式進行距離取值並生成高度矩陣（Height Matrix）。本研究團隊將在下一個段落詳細探討 TripoSR 的使用。

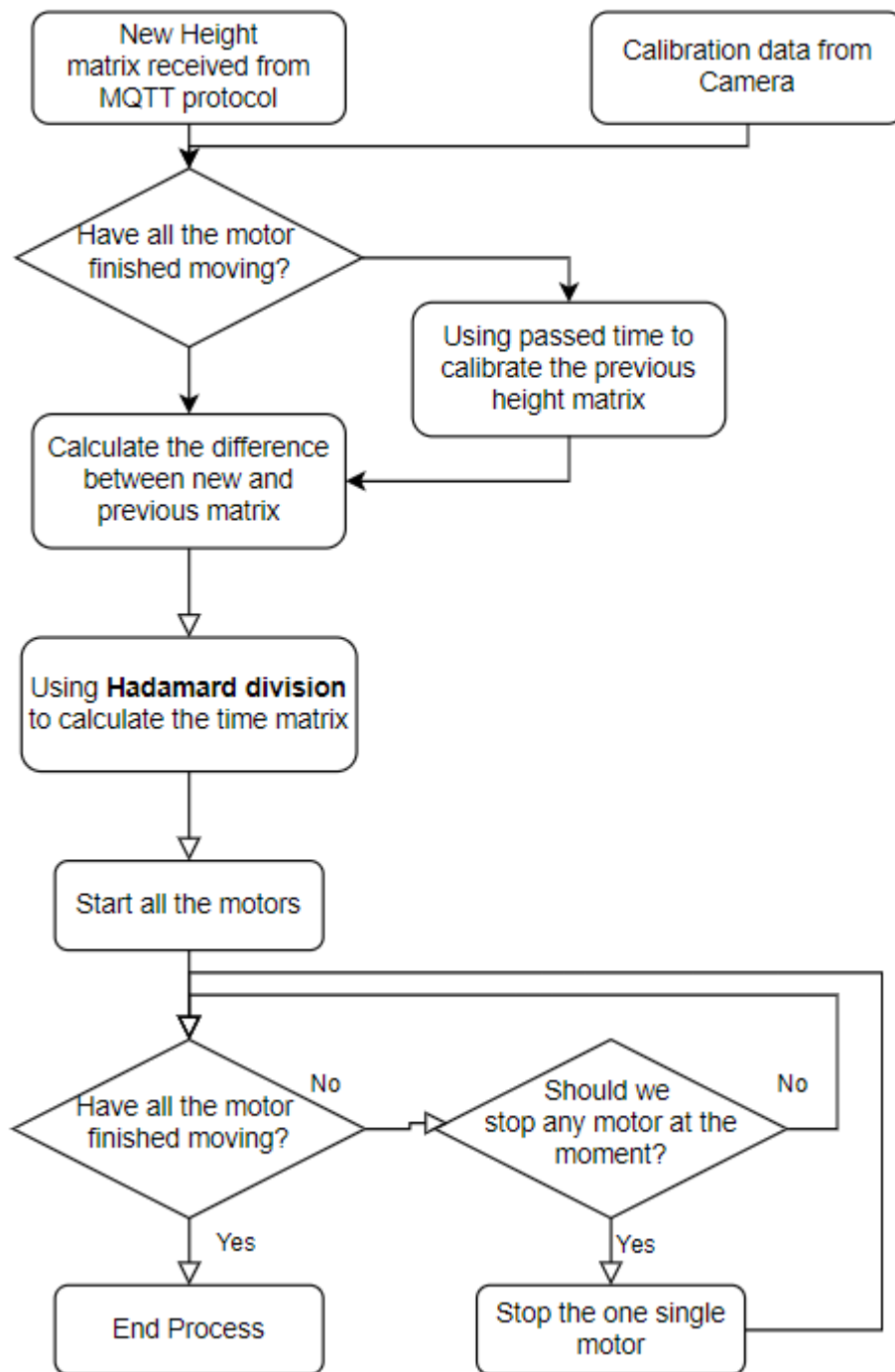


圖 45：MCU 與馬達系統系統運作流程（圖片來源：第二作者繪製）

在運作過程中，MCU 會建立以下的矩陣:TargetMatrix, OldMatrix, TimeMatrix StatusMatrix. 其中，StausMatrix 是記錄圖元移動以及步進移動關鍵，提供多個狀態指示，包括移動完成，移動中與移動準備層級。其中，研究者將以上規劃應用到系統中，如圖 46 所示。



圖 46：進行硬體測試以及程式編寫(圖片來源：第三作者拍攝)

在程式控制過程中，本研究團隊意識到當互動顯示器圖元變換數量超過 41 個時，圖元移動速度會明顯降低，因此對電路的狀況進行了分析。透過電源供應器，當一片顯示模組在全速運轉且無負載時，能使用高達 6.2w 左右，且經過評估本顯示器電源配置僅能供給約 15.7w，並且後端顯示圖元有顯著降壓至約 3.1v 左右情況，在權衡速度與穩定性下，採用了步進調控(stepping-control)的方式，限制單批動作圖元在 16 個圖元。

(二) AR Unity System 與 2D 轉 3D 生成模型及實現

1、使用 TripoSR 生成 3D 物件

為視障學生提供無障礙教育之挑戰在於將書本上複雜的插圖，轉換為「可觸摸之 3D 模型」，使視障生更容易瞭解課本傳達的概念，本研究利用重建 (Reconstrucion) 模型對課本圖片進行立體還原，並且使用 Unity 引擎處理後投射到觸控板上。

本研究團隊首先使用 TripoSR 生成 3D 物件，可以快速將 2D 圖片自動建立成 3D 模型，特色為不容易失真且速度快。圖 47 是 2D 原始圖片，圖 48~50 是本研究團隊利用 TripoSR 生成之 3D 物件圖，並擷取不同視角的圖片。圖 51 是教科書呈現之胃部構造與利用 TripoSR 重建之模型檢視圖。



圖 47：樹根圖（圖取自：CLEANPNG 網站）

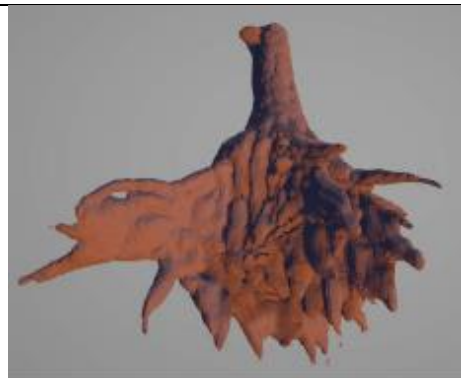


圖 48：研究者使用 TripoSR 將圖 47 生成 3D 物件之仰視圖(第一作者拍攝)

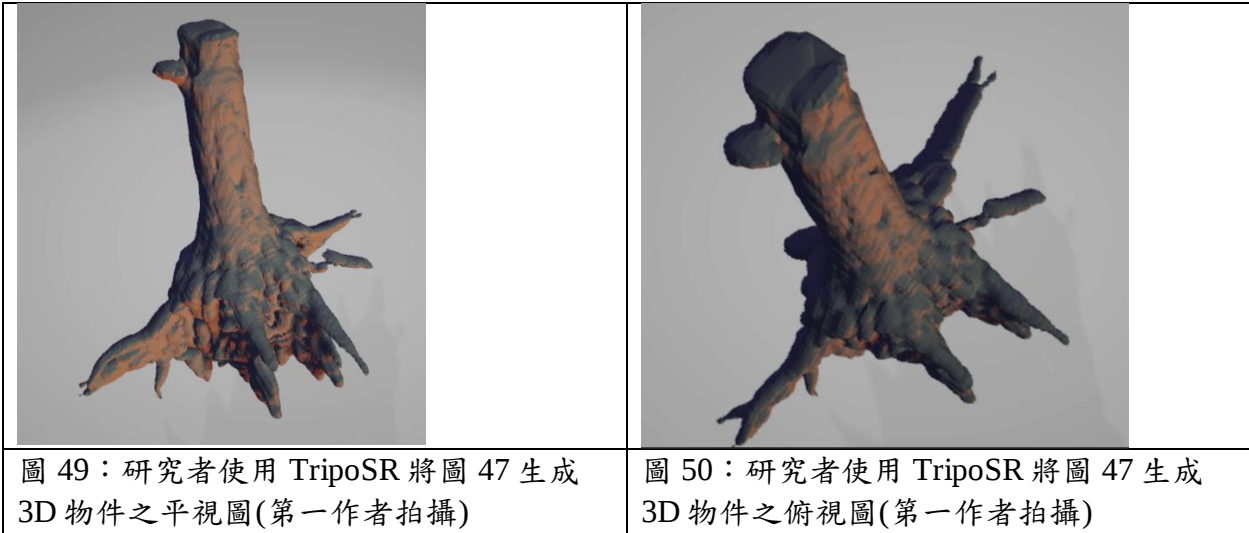


圖 51：教科書呈現之胃部構造與利用 TripoSR 重建之模型檢視圖(圖片來源：第一作者拍攝)

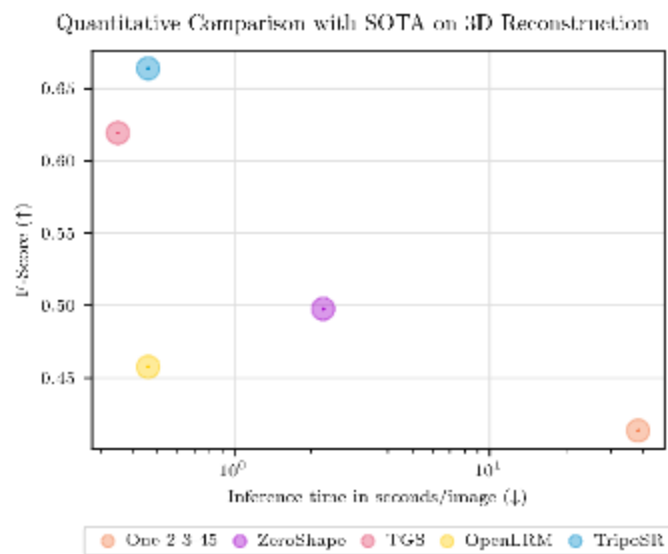


圖 52：TripoSR 與其他領先模型的性能之視覺比較和指標(圖取自(Tochilkin et al., 2024))

2、使用 Unity 3D 渲染並透過觸控板探索 3D 物件

使用 TripoSR 生成 3D 物件後，將在 Unity 3D 中渲染，Unity 3D 提供與觸控板系統互動的介面。Unity 3D 引擎有助於彌合虛擬 3D 環境和實體觸控板之間的差距，使視障學生能夠探索和操作 3D 物件。為了促進這一點，研究者整合了多種輸入機制：

(1) 用於縮放和橫斷面視圖的滑桿輸入的 control panel：使用者可以使用滑桿控制縮放等級並調整 3D 模型的橫斷面視圖。此功能可詳細探索內部結構，放大局部結構，增強複雜概念的學習深度。

(2) 用於旋轉和變換視角的搖桿輸入：將 Unity 連接 Control Panel，可以精確控制 3D 模型的平移和旋轉。使用者可以直觀地移動物體，從不同角度檢查它，並全面瞭解其幾何形狀。

研究者使用 unity 做為人機交互平臺，利用平面分割點發射射線獲取高度矩陣，再以 mqtt 方式投射到觸覺顯示器上。

(三) 研究與開發強化視障生文字理解方法—LLM-based text optimizer for VD Education

研究團隊嘗試運用大型語言模型來增強文字與圖形的理解速度及效果，團隊也瞭解：依視障學生理解文字能力個人化也很重要。

相關方法流程及架構如下圖所示：

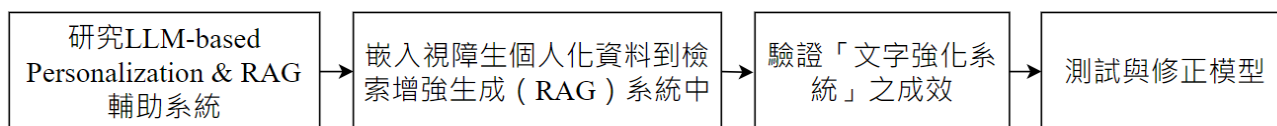


圖 52：研究文字強化流程。由第一作者繪製

1、研究 LLM-based Personalization & RAG 輔助系統

(1) 模型的選擇

為了使本系統產生可靠的回應以及解釋，本研究團隊分析了現有的模型，並且選擇了其中具有代表性的模型進行深入研究：

大參數模型：GPT-4o

雖然 ChatGPT-3.5 (OpenAI, 2022), GPT-4, 和 Claude 等模型已經被證實在文本處理上有良好的效能(Cooper, 2023; Zhai, 2023; Latif et al., 2023; Ahmad et al., 2023; Yan et al., 2024; Xu et al., 2024)，在多次檢閱測試的回應後本研究者認為 GPT-4o 在產生中文教學類文章上的解釋比較流暢自然，因此最終選擇了 GPT4o 作為大參數模型的代表。GPT-4o 擅長處理複雜的文字內容，是 OpenAI 目前頂尖的模型之一。

Llama3-TAIDE-LX-8B-Chat-Alpha1

為了探討本系統在本地輔助裝置運行的可能性以及效能，本研究團隊選擇了以下的小模

型進行測試：

本部分以 Llama3 TAIDE 模型為基礎，該模型由中華民國國家科學及技術委員會開發，並導入適合台灣多領域應用的文本與訓練資料，提升模型在繁體中文回應和特定任務上的表現。Llama3-TAIDE-LX-8B-Chat-Alpha1 基於 Meta 所開源的 LLaMA3-8b，使用繁體中文資料進行持續預訓練，並透過指令微調增強其在辦公室常見任務與多輪對話中的表現，研究者以 RTX3060 12GB 進行測試實驗，Llama3-TAIDE-LX-8B-Chat-Alpha1 可以在家用電腦上運行，視障學生不需要將個人資料上傳到雲端進行運算，因此這部分實驗可探討本系統在保障視障者個人資料與喜好時之效能。

將此系統部署於本地環境中具有顯著的優勢。在生成強化文字時不需將資料上至雲端，降低了視障生資料洩露的風險，確保了學生隱私的保護。隨著資訊安全的重要性日益增加，採用本地操作模式能夠保證視障者個人資料的安全性，有望提升系統的信任度和隱私保護。

此外，本地運行可以提供更快的反應時間，因為它不依賴於網際網路連接速度，所有的處理和計算均在本地設備上完成。這對於需要即時互動的教育應用尤為重要。

特點	GPT-4o	Llama3-TAIDE-LX-8B-Chat-Alpha1
模型架構	高度複雜的大型語言模型	中等規模語言模型 (約 80 億參數)
性能	語言生成和理解能力強，適合高精度任務	提供良好效率，用於資源有限的環境
應用場景	高級文本創作、專業翻譯、智能對話系統	基礎文本分析、簡單聊天機器人
資源需求	需要較高的計算資源	資源需求較少，易於部署
運行成本	較高，適用於需要高性能輸出的場景	較低，在低計算能力的硬體上也可運行
擴展性	對大規模數據集處理優勢明顯	更適合快速部署和運行效能優化
幻覺頻率	雖然較先進，但仍可能在某些狀況下生成幻覺	可能在幻覺生成上表現與 GPT-4o 相似或更頻繁
影響	在處理複雜任務時，幻覺可能影響內容的可靠性	在資源有限應用中，幻覺可能影響用戶體驗
隱私性	需要將資料上傳雲端	可在本地運行，保障使用者個人資料安全

表 4：GPT-4o 與 Llama3-TAIDE-LX-8B-Chat-Alpha1 之比較表

本研究團隊在實際部署後發現 Llama3-TAIDE-LX-8B-Chat-Alpha1 在產生視障生輔助教學文章的方面的效能相對 GPT-4o 較弱，在訊息量不足的情況下，Llama3-TAIDE-LX-8B-Chat-Alpha1 模型相對 GPT-4o 模型更容易產生幻覺，例如某些回答使用編造虛無的資訊。為此，本研究團隊額外嘗試使用 fine-tuning 的方式對 Llama3-TAIDE-LX-8B-Chat-Alpha1 進行強化。

(2) 利用資料集對模型進行微調之實驗

本研究團隊嘗試使用 Fine-tuning 強化模型對視障生輔助的能力與對話順暢度，Fine Tuning 過程中，使用 Hugging Face sidfeels/visually-impaired-llm-assistance-dataset 之訓練資料

集，包含約 45k 個視障輔助個視障輔助問答，用以增強模型在輔助上的知識。

微調架構上，本研究團隊使用 Unsloth.ai 之微調框架以及 Low-Rank Adaptation (LoRA) 進行 Fine tuning。在 Fine-tuning 中，研究者使用 Google Colab 環境進行微調：

Intel Xeon CPU with 2 vCPUs (virtual CPUs) and 13GB of RAM Nvidia Tesla T4 with 2560 CUDA Cores

Python version : 3.10.12

以下是微調的過程：

GPU = Tesla T4. Max memory = 14.748 GB.
5.594 GB of memory reserved.

Loading...

```
trainer_stats = trainer.train()
```

```
==((====))==  Unsloth - 2x faster free finetuning | Num GPUs = 1
  \  / \  /   Num examples = 51,760 | Num Epochs = 1
O^O/ \_/ \   Batch size per device = 2 | Gradient Accumulation steps = 4
 \___/      Total batch size = 8 | Total steps = 60
  "-__-"      Number of trainable parameters = 41,943,040
[60/60 07:41, Epoch 0/1]
```

Step	Training Loss
------	---------------

1	1.819400
---	----------

2	2.292800
---	----------

3	1.691000
---	----------

4	1.946300
---	----------

5	1.642700
---	----------

6	1.601700
---	----------

7	1.193500
---	----------

圖 53：研究者利用 Unsloth 對 Llama3-TAIDE-LX-8B-Chat-Alpha1 進行微調的過程。

圖片來源: 由第一作者紀錄

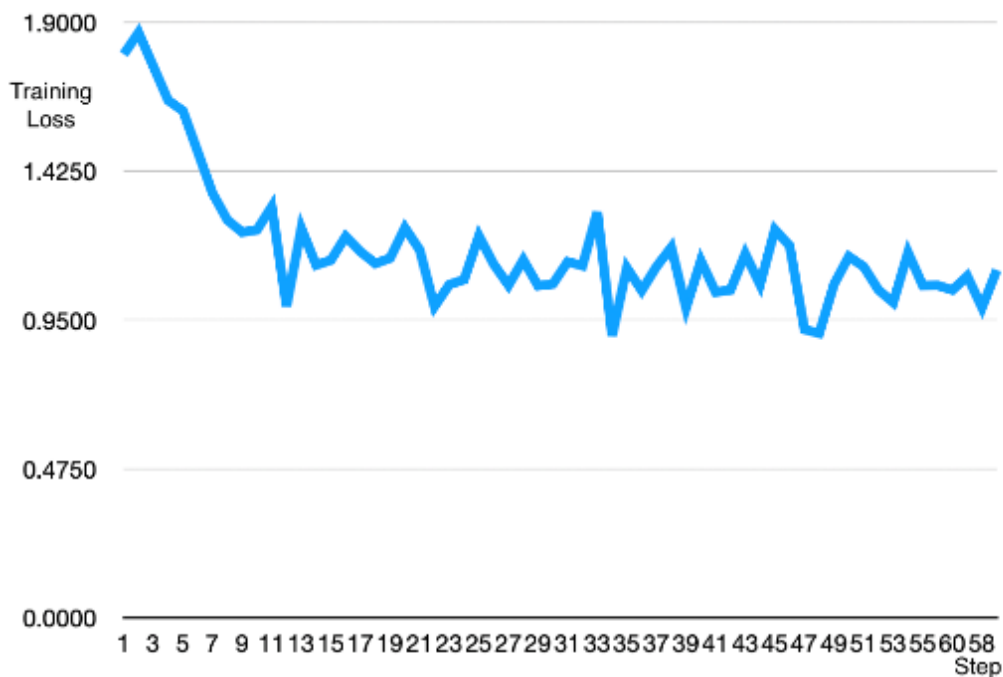


圖 54：研究者對 Llama3-TAIDE-LX-8B-Chat-Alpha1 微調過程的 Training Loss 變化
 圖片來源：第一作者拍攝

以下是分別使用 Llama3-TAIDE-LX-8B-Chat-Alpha1 以及 GPT-4o 進行文字增強的成果：

維基百科	使用 Llama3-TAIDE-LX-8B-Chat-Alpha1 進行文字增強	使用 GPT-4o 進行文字增強
山是地面上被平地所圍繞的具有較大的絕對高度和相對高度而凸起的地貌區。山離地面高度通常在海拔 600 公尺以上，包括低山、中山與高山，是否被稱作山取決於當地人。海拔低於 600 公尺的地形則稱為丘陵。 山一般是因板塊碰撞或是火山作用而產生。山會因河流、氣候作用或是冰河而慢慢侵蝕。有些山會形成單獨的頂峰，不過大部份的山會連在一起形成山脈。高山上因為高海拔而出現高山氣候，氣溫會比	山是一種地貌區，具有高度的特點。高度的山通常在海拔 600 公尺以上，包括低山、中山和高山。山的高度會因為當地人對山的定義而有所不同。低於 600 公尺的高度則稱為丘陵。 山的形成可以是因為板塊碰撞或火山作用而產生。山會因為河流、氣候和冰河的作用而慢慢侵蝕。有些山會形成單獨的頂峰，而大部份的山會連在一起形成山脈。 高山的氣候不同於平地，氣溫會較低。這種氣候會影響高山的生態系。不同的高度會有不同的動植	山是指突出於周圍平地，具有明顯高度的地理特徵。這樣的地形區域一般與平地相比有較大的絕對高度和相對高度。一般來說，海拔高度超過 600 公尺的地形被稱為山，包含低山、中山和高山。這一判斷標準可能會因地區文化而異，不同地方的人對於山的定義可能有所不同。若海拔低於 600 公尺，則稱之為丘陵。 山的形成大多為板塊碰撞或火山活動所致。這些形成於地球表面的隆起隨著時間會因河流流動、水和空氣的氣候作用，以及冰川的活動而慢慢被侵蝕。一些山呈現獨立的尖峰，但大部分山往往聯合成為山脈。

海平面低，因此會影響高山的生態系。不同的高度會有不同的動植物，由於其氣候及環境較不適宜居住，高山上的農業較平地少。有些休閒娛樂和高山有關，例如爬山及滑雪。 地球上最高的山是在亞洲喜馬拉雅山上的珠穆朗瑪峰，距海平面的高度為 8848 公尺[2]。太陽系中最高的山是火星上的奧林帕斯山，高度是 21229 公尺[3]。 世界各大陸山地分布面積近 3,600 萬平方公里，約占全球陸地面積的 24%[4]。 成因 山的形成，是經歷過火山作用或大陸漂移（板塊碰撞）而產生，而引起褶皺、斷裂、隆起、沉積及侵蝕等地理因素。如喜馬拉雅山脈，就是歐亞大陸板塊及印度-澳洲板塊兩個大陸板塊之間互相碰撞對方而產生的板塊擠壓運動；而北美東海岸的阿帕拉契山脈就屬於古大陸地形，後遭侵蝕作用後形成現貌。	物，且高山的農業較少。高山也會有休閒娛樂活動，例如爬山和滑雪。 地球上最高的山是喜馬拉雅山上的珠穆朗瑪峰，高度為 8848 公尺。太陽系中最高的山是火星上的奧林帕斯山，高度為 21229 公尺。 世界各大陸山地分布面積近 3,600 萬平方公里，約占全球陸地面積的 24%。 山的形成是由於火山作用或大陸漂移（板塊碰撞）而產生。這些過程會引起褶皺、斷裂、隆起、沉積和侵蝕等地理因素。例如喜馬拉雅山脈是由於歐亞大陸板塊和印度-澳洲板塊之間互相碰撞對方而產生的板塊擠壓運動；而北美東海岸的阿帕拉契山脈則是古大陸地形，後遭侵蝕作用後形成現貌。	高海拔的地區會展現出特有的高山氣候，這裡的溫度通常比在海平面更低，這種氣候特徵進而影響了棲息於高山的生態系統。不同的高度擁有各自適應環境的動植物群。由於氣候嚴峻且環境不適合人類居住，高山上的農業活動比平地更為稀少。不過，一些戶外活動，如登山和滑雪，常選擇在這樣的環境中進行。 在地球上，最高的山峰是位於亞洲喜馬拉雅山脈中的珠穆朗瑪峰，其海拔高度達到 8848 公尺。此外，在太陽系中，高度最高的山是位於火星的奧林帕斯山，高度達到 21229 公尺。 從全球範圍來看，世界各大大陸的山區面積合計大約為 3,600 萬平方公里，這些山區約佔全球陸地總面積的 24%。 成因方面，山的形成是長期地質作用的結果，諸如火山活動或是由於大陸板塊漂移及碰撞所導致的地層緊壓運動，隨後經歷褶皺、斷裂、隆起、沉積及侵蝕等地理過程。例如，喜馬拉雅山脈的形成是由於歐亞大陸板塊和印度-澳洲板塊兩大板塊之間的碰撞所致。另一方面，北美洲東岸的阿帕拉契山脈可追溯至古老大陸地形，經過長期侵蝕演化而形成現今的山貌。」
---	--	---

表 5：使用 Llama3-TAIDE-LX-8B-Chat-Alpha1 以及 GPT-4o 進行文字增強的成果

表格出處：第一作者整理

2、嵌入視障生輔助相關資料到檢索增強生成（RAG）系統中

經由查閱相關文獻，並與指導老師及特教老師討論後，本研究團隊發現，將視障學生輔

助相關的研究資料，包括相關論文和視障學生輔助教學要領，以及視障學生的個人資料和喜好嵌入到檢索增強生成（RAG）系統中，有望顯著提升視障學生的理解能力。

因此，本研究團隊計劃建構一個 RAG 系統，以進一步探討該系統的效能。利用實際建構出此 RAG 系統以評估這一系統在實際應用中如何改善視障學生的學習體驗和成效。透過收集和分析實驗數據，研究者將驗證這種強化生成方法的有效性，並探索該系統在教育領域更廣泛的應用潛力。

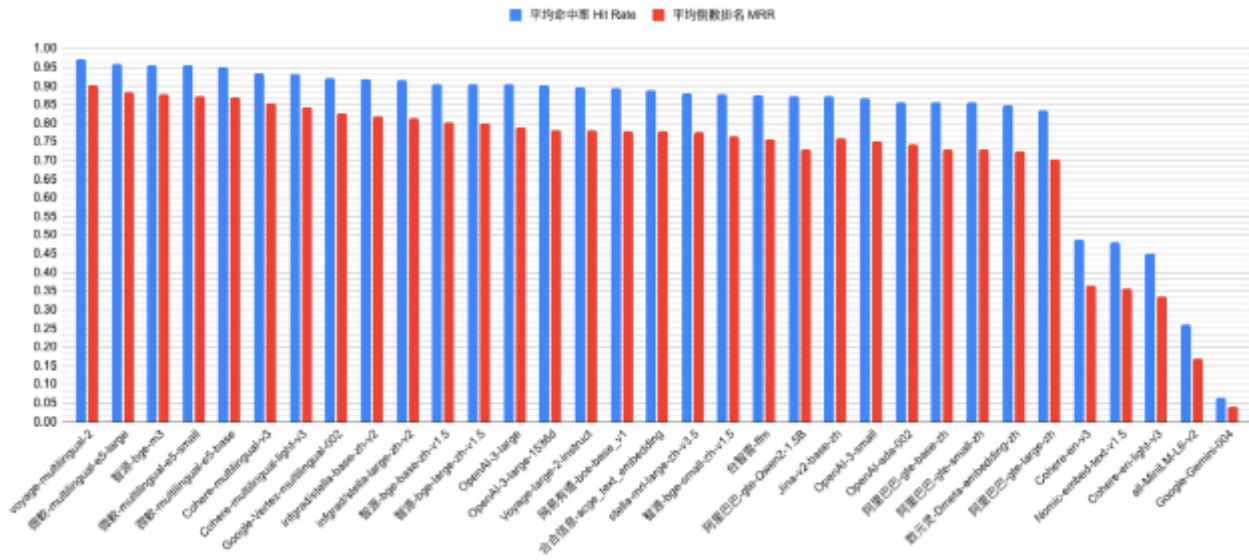


圖 55：使用繁體中文評測各家 Embedding 模型的檢索能力。圖取自 (ihower, 2024)。

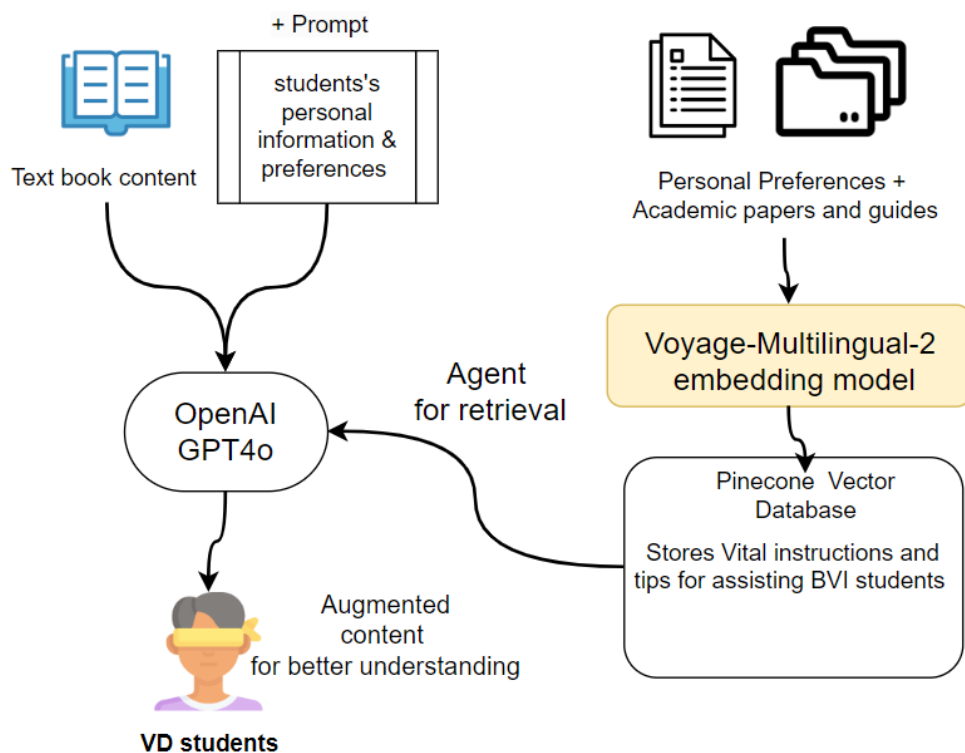


圖 56：最終的 RAG 系統架構圖與視障學生的對話模式（圖片來源：第一作者繪製）

本系統利用檢索增強生成（RAG）方法，實現專為視障學生量身定制的個性化大型語言模型（LLM）。基於先前的實驗，本研究團隊採用 GPT-4o 作為 RAG 系統的 LLM。為了增強視障生的輔助能力，研究者整合了 Embedding 模型 Voyage multilingual，使 GPT4o 能夠藉由 Agent 從向量資料庫取得幫助視障者學生理解的解說方法，並透過 prompt engineering 適應每位學生的偏好、學習行為和交流風格。

ihower(2024)之研究顯示，Voyage-multilingual-2 嵌入模型在繁體中文中的平均 Hit Rate (命中率) 與平均 Mean Reciprocal Rank 都明顯優於其他模型 (ihower, 2024)，因此本研究採用此模型進行 Embedding 視障生個人化的過程始於收集特定數據。研究者使用 Google 表單與視障生進行訪談與紀錄，例如喜好的講解方式與需要注意的地方等。這些數據經過研究者處理並加入到本 RAG 系統之 Prompt 中，使得 GPT-4o 能夠更好理解學生喜好。

此方法的核心在於使用 prompt engineering 來強化 LLM 的回覆。研究者的系統優先考慮能夠更好地符合每位視障者學生的解說風格細微差別，而不僅僅依賴於 Retrieval 過程中的文獻與方法檢索。這使得 LLM 能夠以更貼近學生學習風格的方式生成解釋和說明。

此外，為了使 GPT4o 在指導視障學生之能力有進一步之提升，在搜尋相關輔助文獻後，本研究團隊將 25 篇國內外之特教輔助論文、視障學生教導手冊與相關視障生輔助文獻，利用 Embedding 模型嵌入至向量資料庫中，Agent 可以依據 GPT-4o 的指令進行文獻查找。

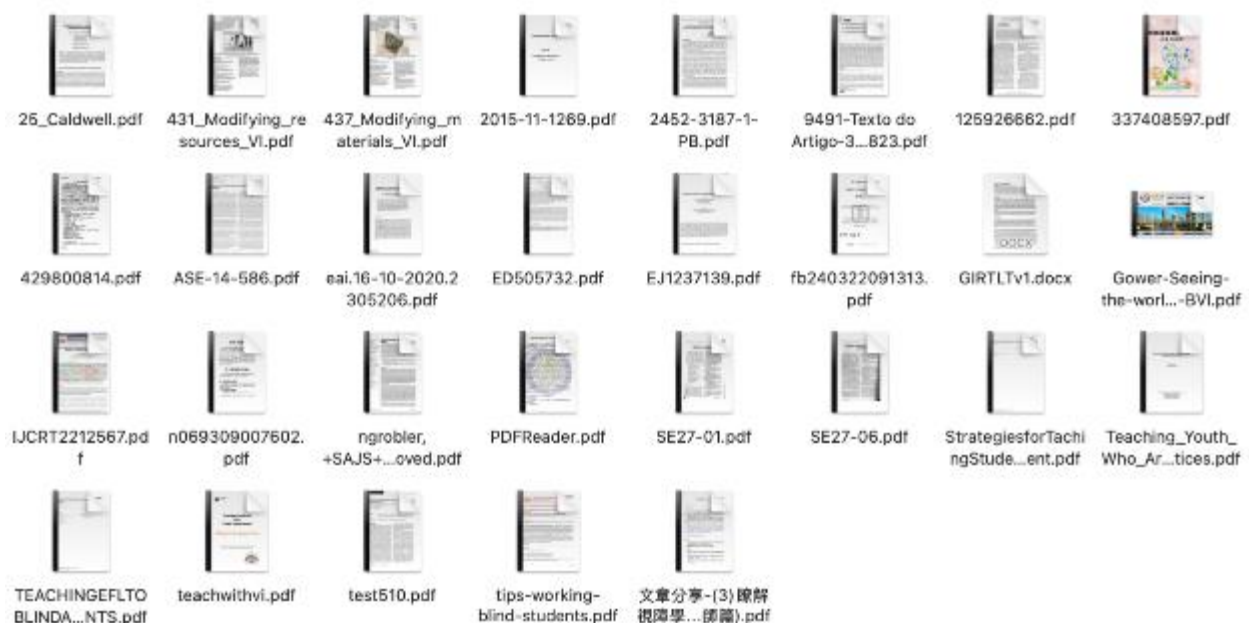


圖 57：本研究團隊建構 vector-database 使用之文章範例（圖片來源：第一作者拍攝）

RAG 架構基於 prompt 從向量資料庫中檢索相關視障輔導資料，並結合 GPT-4o 的生成能力來運行。結果是，系統可以提供量身定制的、語境相關的響應。例如，其中一位學生強調在回答問題時專注在一步一步引導與說明，另一位學生提到，希望再有圖片的時候能夠加強圖片的解釋，這種個性化對於視障學生特別有利，因為它確保**內容以可訪問?**且符合他們個人學習需求的方式呈現。

本研究分別探討了此強化 LLM 在增強視障者在理解如維基百科頁面以及地理課本的新知識時之成效，以及在強化視障生在確保考試公平的情況下增強圖形理解能力。

第一部分是強化視障生在閱讀以及理解新概念時之探討，本研究團隊搜集包含維基百科

及地理課本等相關新知並利用此系統進行強化理解。

以下是結合 Voyage Multilingual , Pinecone vector database 與 GPT4o 之文字強化之結果範例，此範例以維基百科之文獻為參考進行強化以利視障生理解：

首先，GPT4o 理解透過 system_prompt 理解視障者的要求，並且利用本系統建構之輔助文獻與手冊資料庫尋找相關文獻與技巧 (observation)。下圖是研究者使用的 prompt:

```
#system

You are a specialized assistant for reinforcing textbook content for blind students.
Your role is to convert visual information into clear, detailed, and easy-to-understand descriptions, replacing visual elements with step-by-step explanations.

Be aware that you should respond in accordance with the student's personal preference. Use tools to search for methods and strategies to enhance the text by simplifying complex ideas and making them more accessible. Your goal is to ensure the generated text is suitable for direct printing for blind students.

Use the tool to search for academic methods to improve the clarity and accessibility of the following text, adding additional information as needed to enhance understanding. Please respond in Traditional Chinese or Taiwanese, focusing on clear descriptions and explanations

You should add more details directly in the context or below.

If some words or phenomenon is considered difficult for blind student to understand , you should specifically detail and explain them.

.

#query
```

圖 58：研究者使用的 prompt:（圖片來源：第一作者拍攝）

Query 1

Chat with knowledge and tools:

Thought: To provide a detailed step-by-step explanation and use detailed text to describe the phenomenon, I need to search for methods and strategies for teaching and assisting blind students in learning. This will help me understand how to modify the text effectively.

Action: arxiv_search

Action Input: "methods and strategies for teaching and assisting blind students in learning step-by-step explanations and detailed descriptions"

Observation: non-verbal messages and responses, introducing lessons, transitions, closures to all activities, announcing assignments and naming speakers. Provide advance notice to the student and the teacher assistant of materials required for the next class so that everyone is prepared.

Real-life examples and concrete material can assist in establishing relationships between abstract learning and the student's experience. Consider the use of "hands-on" material where possible, to provide opportunities for kinesthetic and tactile learning.

Individual explanation: You may find the student with blindness needs additional explanation. When the class begins a task it may be useful to check with the student to ensure comprehension.

non-verbal messages and responses, introducing lessons, transitions, closures to all activities, announcing assignments and naming speakers. Provide advance notice to the student and the teacher assistant of materials required for the next class so that everyone is prepared.

Real-life examples and concrete material can assist in establishing relationships between abstract learning and the student's experience. Consider the use of "hands-on" material where possible, to provide opportunities for kinesthetic and tactile learning.

Individual explanation: You may find the student with blindness needs additional explanation. When the class begins a task it may be useful to check with the student to ensure comprehension.

圖 59：RAG 系統在參考相關手冊與資料後，GPT4o 基於搜索到的文獻整理所發現的注意事項

項與技巧，並加以生成最終強化後的文字（圖片來源：第一作者拍攝）

山是地面上被平地所圍繞的具有較大的絕對高度和相對高度而凸起的地貌區。山離地面高度通常在海拔600公尺以上，包括低山、中山與高山，是否被稱作山取決於當地人。海拔低於600公尺的地形則稱為丘陵。

山的形成

山一般是因板塊碰撞或是火山作用而產生。以下是山形成的詳細步驟：

1. **板塊碰撞**：地球表面由多個板塊組成，這些板塊會移動並相互碰撞。當兩個板塊碰撞時，地殼會受到擠壓，形成褶皺和斷裂，最終導致山的形成。例如，喜馬拉雅山脈就是由歐亞大陸板塊和印度-澳洲板塊碰撞而形成的。
2. **火山作用**：火山噴發時，熔岩從地球內部噴出，冷卻後堆積在地表，逐漸形成山。例如，夏威夷的火山就是這樣形成的。

山的侵蝕

山會因河流、氣候作用或是冰河而慢慢侵蝕。以下是侵蝕的詳細步驟：

1. **河流侵蝕**：河流流經山區時，會攜帶泥沙和石塊，這些物質會磨損山體，逐漸削減山的高度。
2. **氣候作用**：風、雨、雪等氣候因素會對山體表面產生侵蝕作用，導致山體逐漸變低。
3. **冰河侵蝕**：冰河在移動過程中會攜帶大量冰塊和石塊，這些物質會對山體產生強烈的磨損作用。

山的氣候和生態系

高山上因為高海拔而出現高山氣候，氣溫會比海平面低，因此會影響高山的生態系。以下是高山氣候對生態系的影響：

1. **氣溫**：高山氣溫低，植物和動物需要適應寒冷的環境。例如，高山上的植物通常矮小且耐寒。
2. **生態系**：不同的高度會有不同的動植物。例如，海拔較低的山區可能有茂密的森林，而海拔較高的山區則可能只有苔原植物。
3. **農業**：由於高山氣候和環境較不適宜居住，高山上的農業較平地少。

世界著名的山

地球上最高的山是在亞洲喜馬拉雅山上的珠穆朗瑪峰，距海平面的高度為8848公尺。太陽系中最高的山是火星上的奧林帕斯山，高度是21229公尺。

山地分布

世界各大陸山地分布面積近3,600萬平方公里，約占全球陸地面積的24%。

成因

山的形成，是經歷過火山作用或大陸漂移（板塊碰撞）而產生，而引起褶皺、斷裂、隆起、沉積及侵蝕等地理因素。例如，喜馬拉雅山脈是由歐亞大陸板塊及印度-澳洲板塊之間互相碰撞對面而產生的板塊擠壓運動；而北美東海岸的阿巴拉契亞山脈則屬於古大陸地形，後經侵蝕作用後形成現貌。

這樣的修改提供了更詳細的步驟和描述，幫助盲學生更好地理解山的形成和特徵。

圖 60 :最終強化後的文字（圖片來源：第一作者拍攝）

這種方法借鑒了先前研究的觀點 (Neelakanteswara, Chaudhari, & Zamani, 2024)，其強調了基於利用 Embedding 已實現個人化的方法的優勢。本研究中視障生個人化以 prompt engineering 的方式傳遞給 GPT-4o，免去額外再 Retrieve 的過程。利用以上的方法，有望幫助學生掌握可能通過傳統學習材料難以理解的複雜概念。

（四）驗證「文字強化系統」之成效

為了測試本文字強化系統是否能夠提升視障學生的理解能力與文字表達能力，本研究團隊探討相關文獻後，原本選擇採用 TrueLens、資料集及大型語言模型（LLM）進行系統的初步評估。然而，經過實際運作後，發現這些方法無法充分反映研究者開發之系統對於視障學生理解狀況的效果。因此，研究者決定使用下列方法，以便更直接地評估系統的有效性和實際應用情境中的表現：

1、使用 VDEA Benchmark 驗證輔助成效

本研究團隊與視障學生之相關教師討論後，設計出輔助系統驗證方法 VDEA (visual disabilities Education Assistant) Benchmark 用於驗證本系統之輔助效力。

本研究團隊整合超過 100 題涵蓋化學生物、數學、生物、綜合長篇閱讀，其中的項目分成三類：圖形理解、文字理解與混合理解能力。

在問題反饋的方面，設計了選擇題之正確率作為計分標準，同時本研究團隊也會取得視障生之理解感受想法。

在視障生的主觀理解感受方面研究者使用李克特量表（Likert Scale）進行量化評分，此量表用來衡量受試者的主觀感受。李克特量表的分數從 1 到 5，其中 1 代表非常不同意，5 代表非常同意。這種量表廣泛應用於心理學、社會學以及教育研究等領域。在此次實驗中研究者將 1 分到 5 分，分為非常不理解，稍微不理解，中立，稍微理解與非常理解。並且，也會測試受試者的主觀系統滿意程度，評分為 1 到 5 分。

2、評估新式輔具之成效

本研究中研究者邀請到 2 位視障生參與研究，其分為受試者 A 與受試者 B。其中，受試者 A 為先天性全盲，受試者 B 為嚴重視覺喪失，幾乎無法看清。研究者將控制組設為未使用舊教材，即僅使用點字書以及傳統點圖，並且將介紹實驗組的詳細方法：

(1)實驗組別與研究設計

本研究將實驗組分為三個組別，具體如下：

- 無觸覺反饋系統 (Tactile Pad) 且無使用 RAG 強化系統
- 有使用 Tactile Pad 但無使用 RAG 強化系統
- 有使用 Tactile Pad 且有使用 RAG 強化系統

透過前兩組的比較，可以探討觸覺反饋系統對學習效果的影響；而 2、3 組的比較則分析 RAG 強化系統是否能進一步提升觸覺反饋系統的效果。如下圖所示：



圖 61：視障生使用本研究開發之雙系統（Tactile Pad + RAG 系統）進行學習。

圖片來源：由第三作者拍攝。

(2)硬體與 RAG 技術的比較分析

此部分研究的目的是比較有無硬體設備（Tactile Pad）及有無 RAG（Retrieval-

Augmented Generation) 技術對學習成效的影響。本研究團隊訪談多位視障學生，瞭解他們在學習過程中面臨的主要困難。根據視障學生的回饋，化學、數學與生物被認為是最具挑戰性的科目，因此本研究選擇這三個學科作為實驗範圍，並聚焦於高一課程內容。

每個學科設計 10 道單選題，共 20 題。題目難度保持一致，呈現方式則根據不同組別進行調整。此外，本研究團隊從網路篩選合適的文章，並使用 GPT-4 設計兩道額外單選題，使總題數達到 20 題，供受試者作答。

(3)RAG 技術的效果評估

本研究進一步評估 RAG 技術對不同學科學習的影響。本研究團隊選擇高二範圍內的地理、歷史、公民、物理、化學、生物與地球科學等七個科目，每個科目挑選兩篇文章，國文則選取六篇文章。總計 20 篇文章，每篇包含單選題和簡答題，共 40 題。地球科學部分取自高一課程，國文科則選自課外讀物。

考量到英文科長篇閱讀題會因受試者英語能力差異而影響理解，因此未將其納入實驗；數學科因包含大量公式與圖形，視障受試者理解會有難度，也未被納入。最終，本研究以國文科作為英文科與數學科的替代學科。

三、研究使用之材料、硬體設備、器材與軟體

(一) 使用之材料

名稱	數量	名稱	數量
3D 線材(實白)	4 捲	3D 線材(海琉璃)	2 捲
螺絲 M2*10mm	16 顆	螺絲 M3*5mm	16 顆
螺絲 M3*10mm	32 顆	螺絲 M3*12mm	30 顆
螺絲 M4*18mm	4 顆	螺柱 M3*150mm	256 根
螺帽 M2	16 顆	螺帽 M3	334 顆
螺帽 M4	4 顆	排線*100mm	7 排
排線*200mm	1 排	微型馬達	256 顆
滑桿式電組	2 個	三軸按鈕	1 個
杜邦線	約 200 條	麵包版	4 片

表 6：研究使用之材料（表格出處：第三作者整理）

(二) 使用之硬體設備與器材

名稱	數量	名稱	數量
3D 列印機	1 台	電源供應器 1000W	2 捲
直流電源供應器 1000W	1 顆	OAK- Pro Depth camera	1 顆
i7 6700k with 16GB RAM	1 組	數位電錶	1 台
Wemos D1 mini 控制板	1 片	客製化 PCB —主控板	1 片
客製化 PCB-從型	8 片	NodeMCU-32S WIFI	1 片
5V 5cm*5cm*2cm 風扇	1 顆	RTX3060 12GB	1 顆
iPad Pro 2021 (5th	1 台	Macbook Pro 2019 16	1 台

Generation)		inch	
-------------	--	------	--

表 7：研究使用之硬體設備與器材（表格出處：第三作者整理）

（三）使用之軟體

名稱	名稱
Fusion 360	Arduino IDE
Microsoft Visual Studio	Unity 3D
Jupyter notebook	anaconda navigator
notion 實驗紀錄與管理文件	github
Google drawio	Voyage Embedding api
OpenAI GPT4o api	TAIDE Llambda 3
pincone database api	Anaconda

表 8：研究使用之軟體（表格出處：第三作者整理）

參、研究結果與討論

一、觸覺反饋強化與文字強化對視障者實際理解與認知能力的影響

本研究團隊經過考量後，使用了百分比與李克特量表（Likert scale）進行統計與分析。其一是百分比：透過計算參與者在不同變項下的行為或反應，以比例形式呈現，方便直觀比較不同群體或條件的差異。百分比統計能有效顯示結果的分佈狀況，並有助於快速檢視實驗效益與影響範圍。例如，某一群體在特定條件下的反應占比可反映該條件對行為的影響程度。其二是李克特量表：則被用來評估受試者對實驗中某些因素的態度、看法或理解度。研究者通常設計一個包含多個題項的問卷，並將受試者的回應依照李克特量表的五點或七點等級來量化，從「非常不滿意」到「非常滿意」或從「非常不理解」到「非常理解」等，進而分析受試者的平均反應趨勢及變異性。

本研究團隊將兩種方式結合使用，為後續的研究或應用提供依據。

(一)答題正確率(百分比)數據

正確率(百分比)		純文字+ 傳統點圖		純文字+ Tactile Pad		RAG + Tactile Pad		正確答案
		受試者A	受試者B	受試者A	受試者B	受試者A	受試者B	
化學科	1	C	A	C	A	C	C	C
	2	D	C	A	C	A	C	C
	3	D	D	D	B	C	B	C
	4	A	A	A	A	A	A	D
	5	C	A	C	A	C	A	A
	6	B	D	B	D	B	B	B
	7	C	C	C	C	C	C	A
	8	B	D	B	D	B	D	D
	9	C	C	C	C	C	C	C
	10	B	D	B	D	B	D	D
數學科	1	A	D	B	D	D	D	D
	2	A	A	A	B	A	B	C
	3	C	A	C	A	C	A	C
	4	A	A	A	A	A	A	A
	5	D	C	D	C	D	C	C
	6	A	D	A	A	A	A	A
	7	A	D	B	B	B	B	B
	8	C	A	C	C	C	C	C
	9	C	C	A	C	A	C	A
	10	B	C	B	B	B	B	B
生物科	1	A	B	A	B	A	B	B
	2	D	A	D	D	D	D	D
	3	A	C	D	C	B	C	C
	4	A	A	A	A	A	A	A
	5	B	C	A	D	C	D	D
	6	B	A	B	A	A	A	A
	7	C	C	C	C	C	C	C
	8	C	B	B	B	B	B	B
	9	C	C	C	C	C	C	C
	10	B	B	B	B	B	B	B
答對題數		12	15	16	22	19	24	
答對率		0.4	0.5	0.53	0.73	0.63	0.8	

表 9：實驗答題正確率統計圖(綠色為答對之題目)（表格出處：第二作者整理）

(二)理解度(五等級李克特量表) 數據

	理解度(李克特量表)		純文字+ 傳統點圖		純文字+ Tactile Pad		RAG + Tactile Pad	
			受試者A	受試者B	受試者A	受試者B	受試者A	受試者B
1		1	3	2	4	3	5	5
2		2	3	3	4	3	4	4
3		3	3	3	4	3	5	4
4		4	2	2	2	2	2	2
5		5	2	3	2	3	2	3
	化學科	6	3	2	4	4	4	4
		7	2	1	3	3	3	3
		8	2	3	4	5	4	5
		9	4	5	5	5	5	5
		10	2	3	2	5	4	5
		1	3	3	4	5	5	5
		2	1	1	2	2	2	2
		3	3	2	5	3	5	4
		4	5	5	5	5	5	5
		5	2	2	4	4	4	4
	數學科	6	3	3	4	5	4	5
		7	2	2	5	5	5	5
		8	3	3	5	5	5	5
		9	3	2	5	4	5	4
		10	3	2	3	3	3	4
		1	3	3	4	5	5	5
		2	4	3	4	4	4	5
		3	3	3	4	4	5	4
		4	4	3	4	4	5	4
		5	3	3	4	3	4	4
	生物科	6	3	3	4	3	4	4
		7	5	4	4	4	4	4
		8	3	5	4	5	4	5
		9	5	5	5	5	5	5
		10	4	4	4	4	4	5
	平均值		3.033333333	2.933333333	3.9	3.933333333	4.166666667	4.266666667
	標準差		0.9643054793	1.080655399	0.9228890171	0.9802650357	0.949894126	0.8683449709

表 10：實驗理解度統計圖(表格)（表格出處：第二作者整理）

在本研究中，本研究團隊針對 Tactile Pad 與 RAG 兩種不同的實驗條件，進行了答對率與理解度的量化比較。結果顯示，Tactile Pad 的導入在提升實驗成效上表現顯著，其帶來的答對率增幅達到平均 18%，並且用戶的滿意度也略微提升，平均增加了 0.9 分。這表明觸控板在提升學習或操作精度方面具有較高的有效性，參與者對其使用體驗的評價也相對正面。

有無導入Tactile Pad 的正確率影響

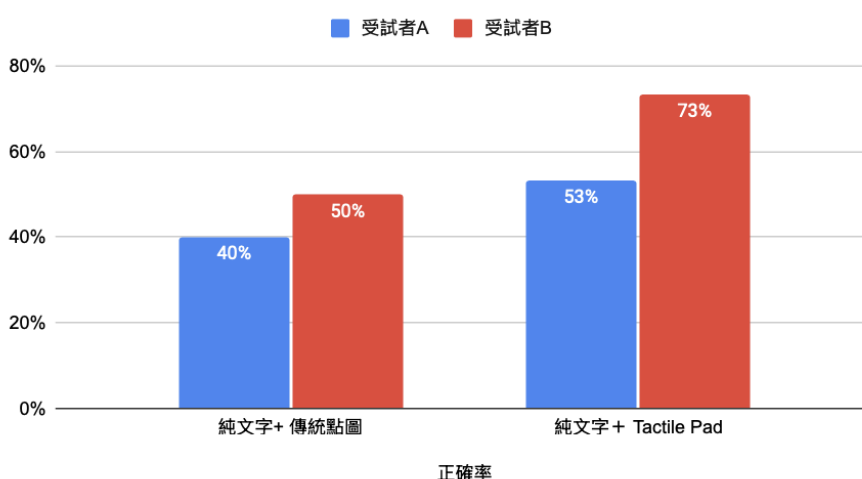


圖 62：比較有無 Tactile Pad 之答題正確率柱狀圖，平均相對傳統點字系統有約 18% 的提升。
圖片來源：第二作者拍攝

有無導入Tactile Pad 的理解感受影響

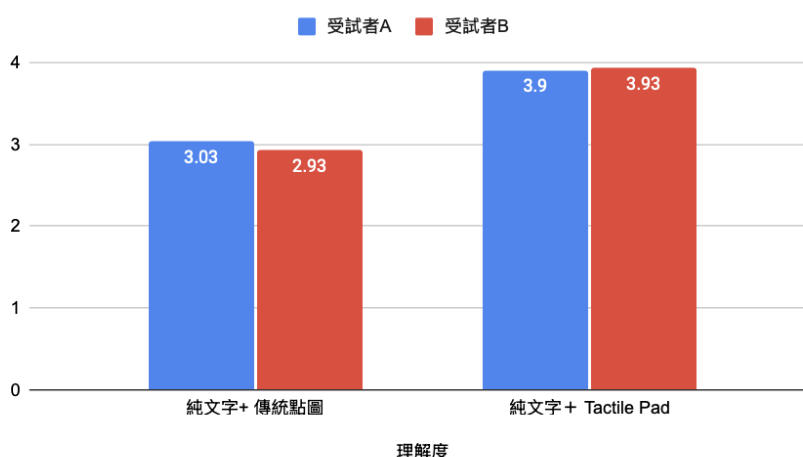


圖 63：比較有無 Tactile Pad 之理解度感受柱狀圖，平均相對傳統點字系統有約 0.9 點的提升，表示視障生對理解好感度約有 1/5 的提升（圖片來源：第二作者拍攝）

相較之下，RAG 系統的導入在提升答對率上表現較為溫和，答對率僅增長約在原本的基礎上增加 9%。儘管其增幅不如 Tactile Pad 顯著，但仍然有一定的正向影響。同時，滿意度的提升幅度也較小，僅增加了 0.3 分。由數據可知，RAG 在提升學習或操作成效上，雖具備一定效果，但其在有 Tactile Pad 的前提下對使用者體驗或操作便捷性方面的貢獻相對有限。

有無導入RAG技術的正確率影響

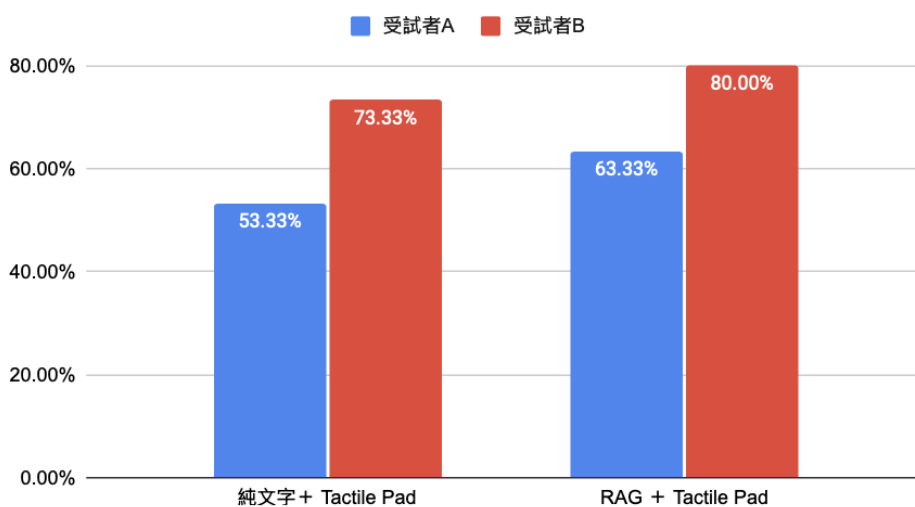


圖 64：比較有無 RAG 技術之答題正確率柱狀圖，平均在有本研究開發之 Tactile Pad 之基礎上有約 9%的提升（圖片來源：第二作者拍攝）

有無導入RAG的理解感受影響

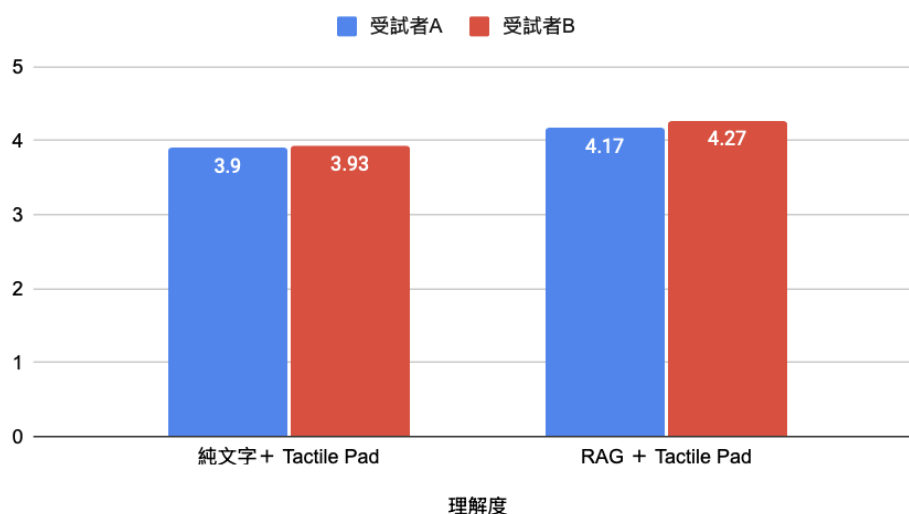


圖 65：比較有無 RAG 技術之理解度感受柱狀圖，平均相對傳統點字系統有約 0.3 點的提升，在原有 Tactile Pad 的基礎上之提升相對溫和（圖片來源：第二作者拍攝）

答題正確率

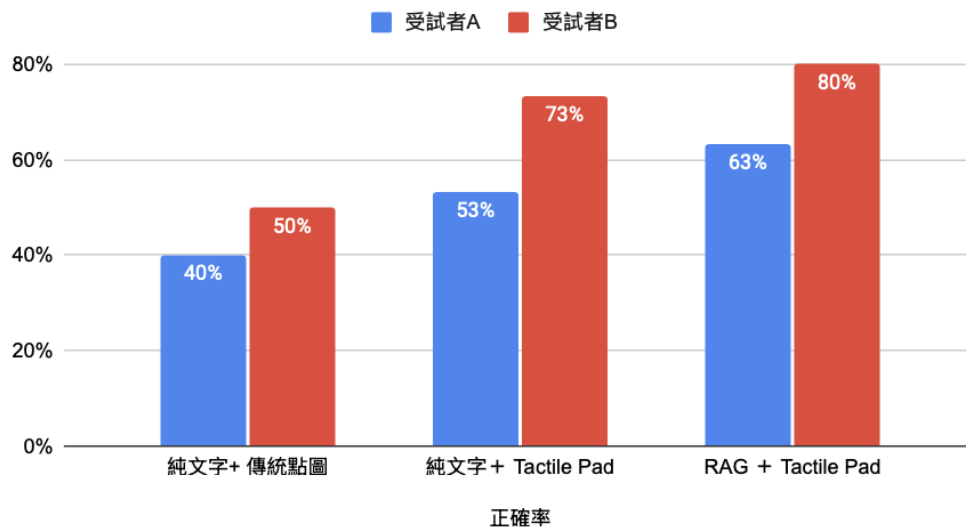


圖 66：舊有教材與本研究開發之雙系統對視障生答題正確率的影響
(圖片來源：第二作者拍攝)

理解感受度

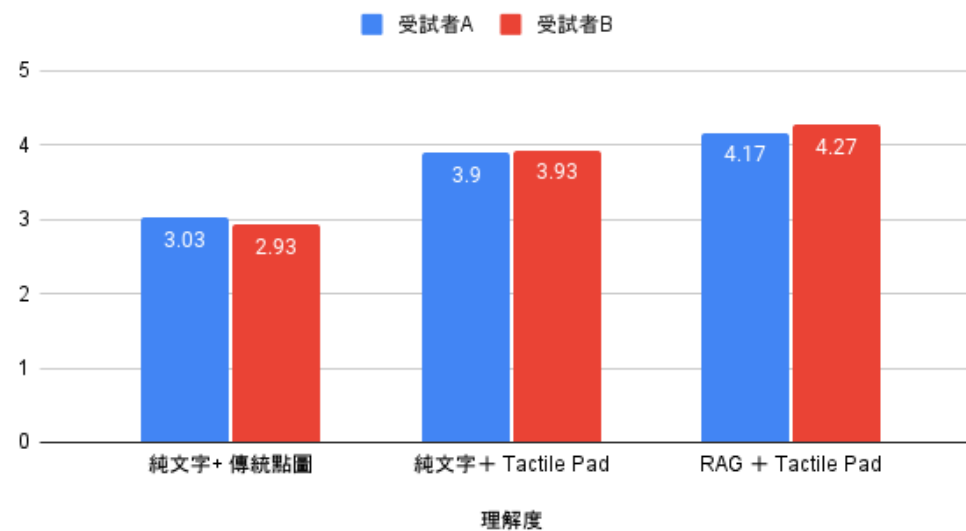


圖 67：舊有教材與本研究開發之雙系統對視障生理解感受度的影響
(圖片來源：第二作者拍攝)

二、視障生對本研究開發之雙系統之評價

透過質性訪談，研究者取得了視障者對本系統之主觀評價，並對其進行了分析。

受試者 A:

“還不錯，可以明顯分出不同器官或物質的差異，不過原子模型希望可以更細緻化一點，可以更能分出不同的原子模型。”

在受試者 A 的回應可以瞭解到，視障生對本系統的整體感受良好，能清楚辨別不同器官或物質的差異。然而，對於原子模型的呈現，期望能提高細緻度，使不同的原子模型能更加清晰區分。系統的主要限制在於像素密度的不足。

受試者 B:

“我認為大部份的感受都很好，只是 RAG 系統比較沒有幫助我這麼多”

在受試者 A 的回應可以瞭解到，視障生覺得大部分的體驗都很好，整體感受正面，但特別提到「RAG」這個部分對他/她幫助不大。指出 RAG 系統對每個人的效果不同

三、利用假說檢定進行更嚴謹的統計分析

為了更加深入地分析樣本之間的差異，本研究團隊選擇採用假說檢定（hypothesis testing）來比較樣本，進行更嚴謹的統計分析。假說檢定是一種用來檢驗數據是否符合特定假設的方法，目的是判斷樣本數據是否足以反駁原假說。透過此方法，可以有效檢驗不同變數之間的關聯性和差異性，從而得出更具信效度的結論。

首先，本研究團隊針對不同數據樣本的規模及性質，選擇了合適的統計檢定方式。包含卡方檢定（Chi-Square Test）、Z 檢定（Z-Test）和 T 檢定（T-Test）。檢定方法各自針對不同的樣本情況，適用於大樣本或小樣本，並在不同場景下幫助研究者得出結論。以下是對每種檢定方式的詳細分析和應用場景：

統計檢定	卡方檢定	Z 檢定	T 檢定
適用數據	大樣本($n > 30$)	大樣本($n > 30$)	小樣本($n \leq 30$)
使用場景	檢驗觀察頻數與預期頻數之間的差異	檢驗兩組樣本平均數間的差異	檢驗樣本平均數與母體平均數間的差異
分佈	卡方分佈	常態分佈	T 分佈
例子	檢驗不同地區的顧客滿意度是否有差異	檢驗一個產品的平均重量是否符合標準	比較兩個不同教學方法的考試成績

表 11：統計檢定比較（表格出處：第二作者整理）

在進行進一步的分析時，研究者選擇使用「配對樣本單尾 T 檢定（Paired Sample One-Tailed T-Test）」來比較相同個體在不同條件下的結果。配對樣本 T 檢定適合用於測試在兩個相關的條件下，樣本的平均數是否有顯著差異。例如，研究者可以比較同一組受試者在不同實驗環境下的表現，以評估幹預措施的效果。由於研究的假設方向性明確，研究者採用了單尾檢定，假設某一條件下的結果顯著優於另一條件。

單尾檢定或雙尾檢定的選擇取決於研究假設的方向性：

- 單尾檢定（One-Tailed Test）：用於檢驗樣本在一個特定方向上是否顯著優於或劣於另一組樣本，假設只有一個方向會發生變化。如果研究者假設某一方案或條件會導致更高的成績或更低的錯誤率，便會使用單尾檢定。
- 雙尾檢定（Two-Tailed Test）：用於檢驗樣本之間是否存在任何方向的顯著差異，無論結果是更好或更差。當研究者無法預測結果的變化方向，或對任一方向的變化均感興趣時，會選擇雙尾檢定。

在進行統計檢定時，T 值（T-value）和 P 值（P-value）是用來判斷結果是否具有顯著性的兩個重要指標。

- T 值（T-value）：T 值是 T 檢定計算過程中的一個統計量，用來衡量兩組樣本平均數之間的差異。T 值的大小與樣本平均數之間的差異有關，T 值越大，表示兩組樣本之間的差異越明顯。T 值也受到樣本變異性和樣本量的影響。如果樣本間的變異性較小，或樣本量較大，T 值會增加。
- P 值（P-value）：P 值是用來衡量檢定結果的顯著性水準。具體定義為 T 分佈中，以 T 值為以 T 值為下限之機率密度函數曲線下方的面積(下圖)。具體來說，P 值代表在零假設（即假設兩組樣本之間沒有差異）的前提下，觀察到現有數據或更極端數據的機率。通常，P 值小於設定的顯著性水準（例如 0.05 或 0.01），表示可以拒絕零假設，說明樣本之間的差異具有統計上的顯著性。P 值越小，意味著結果具有更高的顯著性。

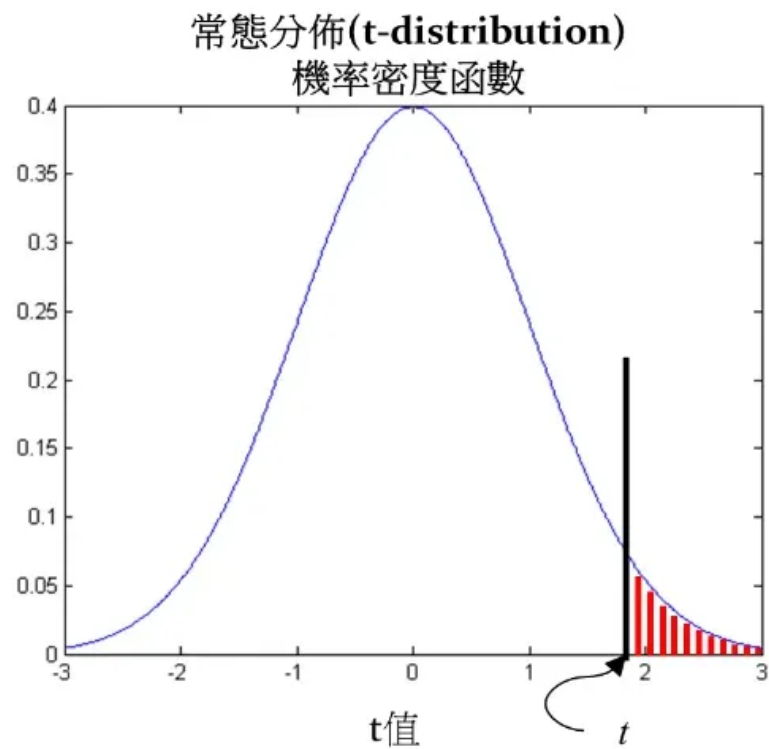


圖 68：P 值和 T 值對應關係圖（圖片來源：第二作者拍攝）

以下是分析結果：

	受試者 A	受試者 B
觸控板	$p=0$ $t=5.517$	$p=0$ $t=5.784$

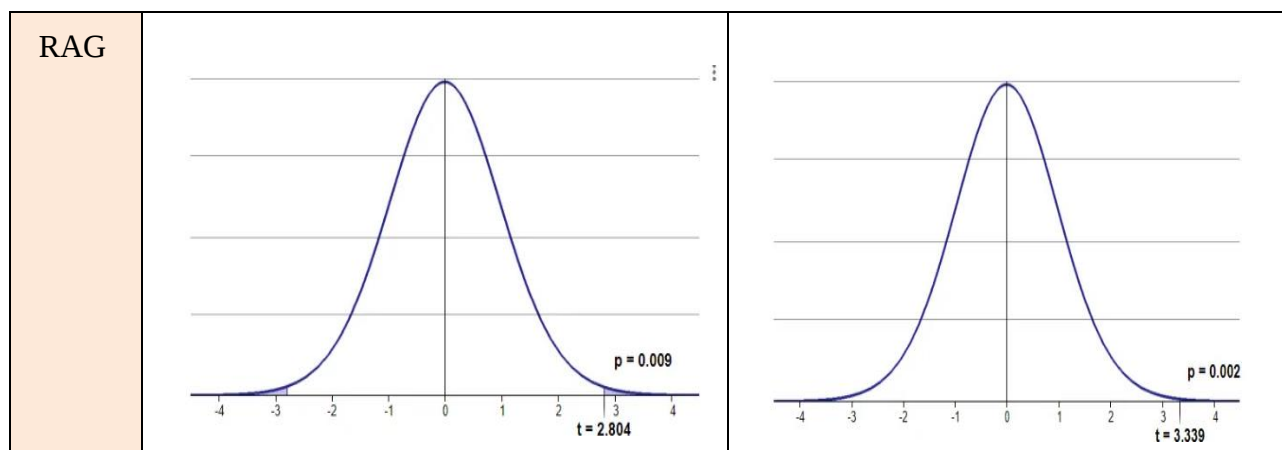


表 12：配對樣本 T 檢定結果分析(單尾檢定)（圖片來源：第二作者拍攝）

首先，根據實驗數據，本研究團隊發現 RAG 與觸控板這兩種不同的互動模式在提升答對率和使用者的理解度上，均有顯著效果。然而，透過進一步的統計檢定發現，兩者在提升效果上的影響性有所不同。具體來說，RAG 實驗的 T 值相對於觸控板來說較小，表明 RAG 在樣本中所造成的效果差異較小；但即便如此，RAG 實驗的 P 值依然趨近於 0，小於顯著水準，這意味著該結果在統計上依然具有顯著性。本研究團隊便可以推翻虛無假設，認為 RAG 的操作模式確實對使用者產生了積極影響。同樣地，觸控板時實驗的 T 檢定結果表現出更大的 T 值，顯示其在提升答對率和滿意度方面具有更明顯的效果，且 P 值也顯示出極為顯著的統計學證據。因此，不論是 RAG 還是觸控板，在不同情境下均能有效提升實驗結果，並且其差異均通過統計檢定確認了具有顯著性。

進一步整合長條圖所呈現的數據，可以直觀地看到兩者對於提升答對率和使用者的理解度的差異：觸控板影響下的答對率和理解度漲幅均高於 RAG。從這些數據可見，儘管兩者都能有效提升使用者的表現，但觸控板在提升效能上表現得更加顯著。這些結果與之前的配對樣本 T 檢定結果相符，進一步鞏固了實驗結論。

總結來看，本研究通過詳細的實驗數據與統計檢定，驗證了觸控板與 RAG 均能顯著提升使用者的答對率與滿意度。本研究團隊在假說檢定中成功證明對立假設成立，證實了兩種模式在學習成效與使用體驗上均具有統計顯著性。

四、Visual Disabilities Education Hybrid Transformer 與常用視覺輔具之差異

本研究團隊經由文獻探討，發現視障者能取得之輔具侷限於點字與近乎平面之觸覺圖形，文獻中並未出現視障者使用個人化立體視覺輔具之資料，因此本研究團隊僅將本研究開發之輔助平臺 VD Education Hybrid Transformer 與視障者常用的點字顯示器作比較，比較比較結果如下圖所示：

	VD Education Hybrid transformer 第一代	VD Education Hybrid Transformer 第三代(具備深度陣列相機)	Focus 14 Blue - 5th Generation	Brailliant BI 40X braille display
功能	利用重建模型與擴增實境系統，	利用重建模型與擴增實境系統，可顯示空間資訊，允許使用者探索圖	只可顯示文字	只可顯示文字 Speech to

	可顯示空間資訊，允許使用者探索重圖片重建的模型 此外，透過基於搜索是增強的 GPT-4o 可以強化課本的文字，以便更好的理解	片重建的模型 未來研究者可以搭配相關的軟體研究，允許使用者利用手勢與 VisionLLM 互動 此外，透過基於搜索式增強的 GPT-4o 可以強化課本的文字，以便視障生更好的理解		text
顯示點數	透過出色的主從架構 電路設計，可顯示高達 256 個可動空間圖元點	透過出色的主從架構 電路設計，可顯示高達 256 可動空間圖元點	14 個盲文	40 個盲文
可擴充性	藉由 ShiftRegister 的特性，理論上可擴展至 521 pixel, 1024 pixel 或甚至更多 視障學生擴充顯示面積時，不需要添購新顯示器	藉由 ShiftRegister 的特性，理論上可擴展至 521 pixel, 1024 pixel 或甚至更多，亦可以進行縮小 視障學生擴充顯示面積時，不需要添購新顯示器 此外透過深度 camera, 未來擴充的部分不需要增加額外的矯正系統 視障者學生可以 尋找適合的面積與顯示大小	不可	不可

表 13:本研究開發之系統與市售點字機之功能比較（表格出處：第二作者整理）



圖 69：Brailiant BI 40X braille display 市場上最好的 40 單元點字顯示器之一（圖取自：Instruloge 官方網站）

三、未來展望

(一)往 VisionLLM 進行探討應用與最佳化

先前的研究較少出現 RAG 與 VisionLLM 的搭配，因此在本研究中並未採用此架構進行模型強化，但本研究團隊建議未來的研究者可以深入往這個方向繼續探討 並嘗試利用相關的方法，包括 Fine-tuning 進行強化以更好的理解並輔導視障學生。

另一方面，本研究因為單純使用 GPT 4o，並未加入視覺模型，因此無法偵測視障生在探索 Tactile Pad 時之過程。研究者期待未來研究者可以結合 VisionLLM，學生能夠指出模型的位置並提出疑問，且 Vision LLM 可即時回答。

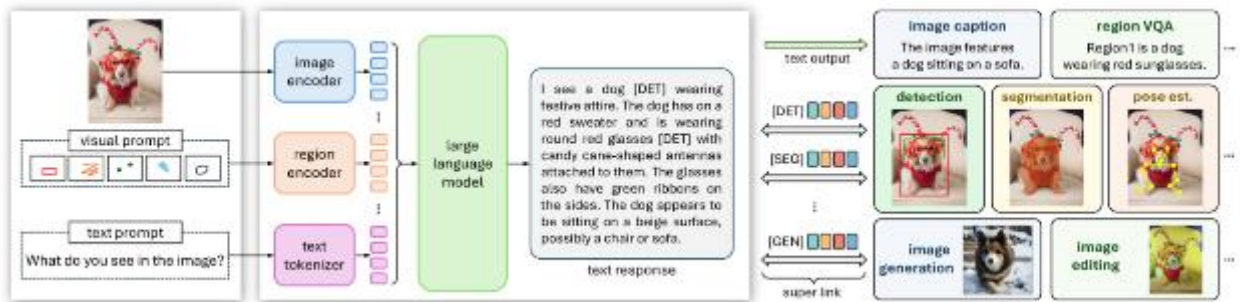


圖 70：VisionLLM v2 的總體架構。它接收圖像和文字及視覺提示作為輸入。LLM 解析使用者指令並產生文字回應。除了輸出純文字之外，LLM 還可以在需要時輸出特殊的路由標記，例如[DET]。超級連結查詢將自動附加在路由權杖嵌入之後，並由 LLM 進一步處理。它們充當連接 LLM 和特定任務解碼器的橋樑。這樣，我們的通才模型就可以支援數百個視覺任務。(Wu et al., 2024)

此系統亦可能對其他特殊族群產生顯著作用。例如，對自閉症患者而言，透過將表情辨識與聲音分析模型結合進 RAG 系統，能夠為患者提供更個性化的互動方式，幫助他們更好地理解情感線索與社會互動。同時，這些技術還可以作為患者訓練社交技能的輔助工具，增強他們與他人溝通的能力，從而提升生活品質。此外，該系統的靈活性也允許針對不同需求進行調整，使其應用範圍能涵蓋更多其他具有特殊需求的群體。

(二) 採用多重 (Multimodel)觸覺重現系統

本研究團隊當前的系統主要通過語音和觸覺形狀方式提供反饋。然而，未來的研究可以探索整合額外的感官反饋方法，以增強視障學生的學習體驗。例如，使用壓電材料可以通過更準確地模擬紋理來增強觸覺反饋。液體基底機制可用於模擬彈性流體，提供更細微的觸覺感受。此外，引入基於振動的反饋可以模擬阻力，幫助學生以更沉浸式的方式更好地理解物理力或材料特性。如圖 71~76 所示，人類可透過不同的觸覺反饋進行探索。

通過超越基本的觸覺和聽覺反饋，這些先進技術可以提供更豐富的多感官學習體驗，使複雜概念對視障學生來說更容易理解和掌握。

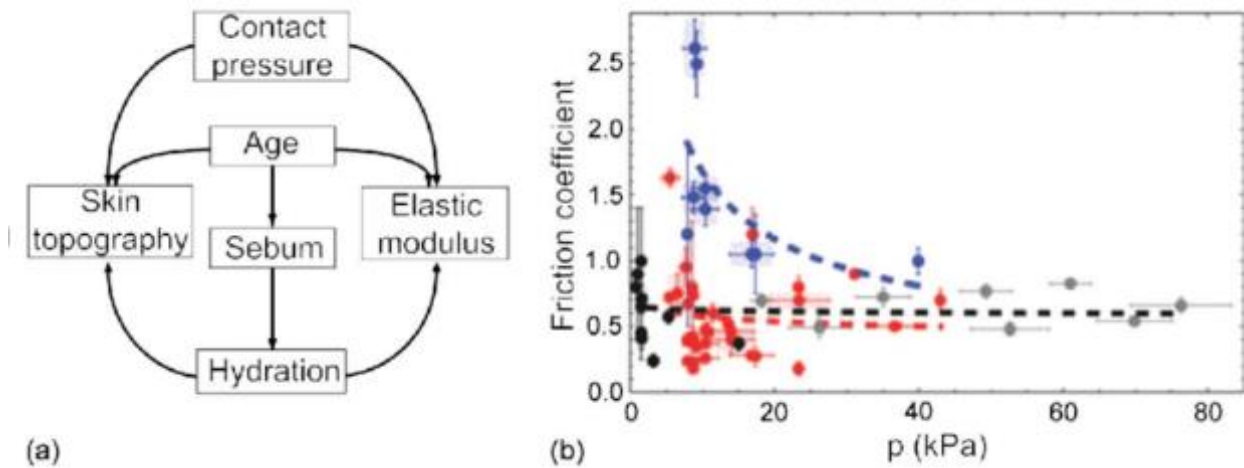


圖 71：(a) 影響人體皮膚摩擦的重要因素的相互作用（皮脂是皮膚腺體分泌的油性/蠟狀物質）。(b) 使用線性移動探針測量的皮膚滑動摩擦係數作為表觀接觸壓力的函數。上面的曲線（藍點）對應於濕皮膚上的球形探針；數據擬合顯示這是摩擦係數取決於接觸壓力的唯一條件。其餘數據顯示接觸壓力的不變性，對應於乾燥皮膚上的探針。(Derler & Gerhardt, 2012)

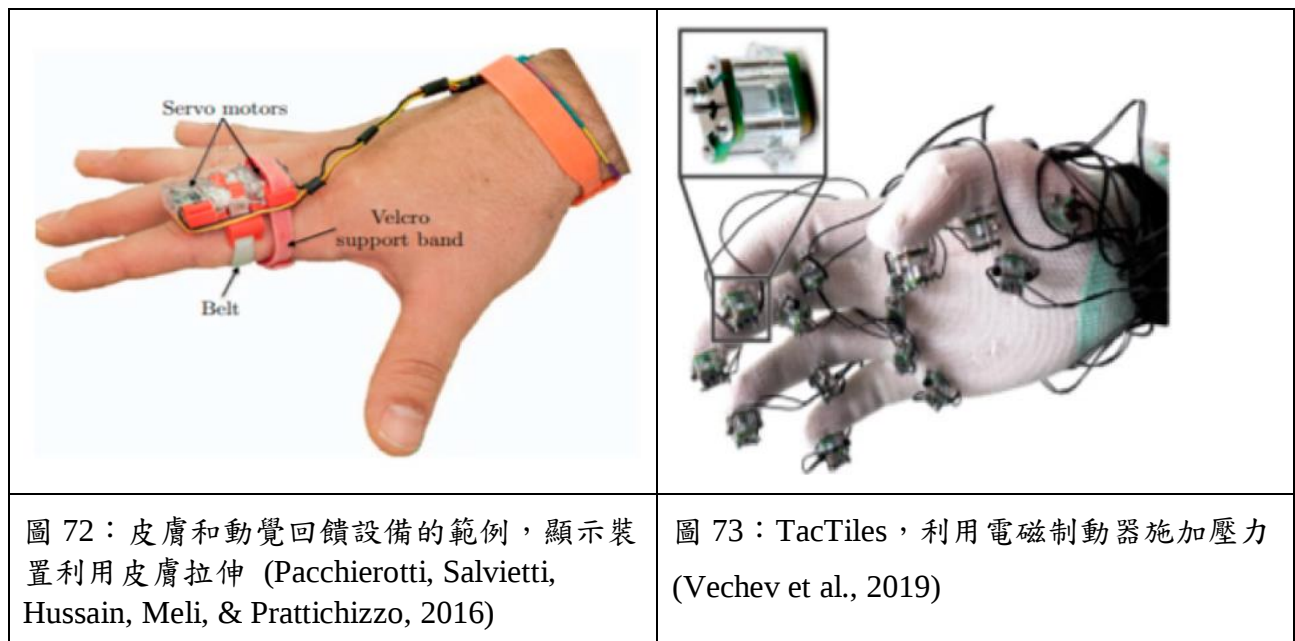




圖 74：LucidVR，基於字串的執行器
(LucidVR, LucidGloves) (Burbach, Steiger, & Giesser, 2022)

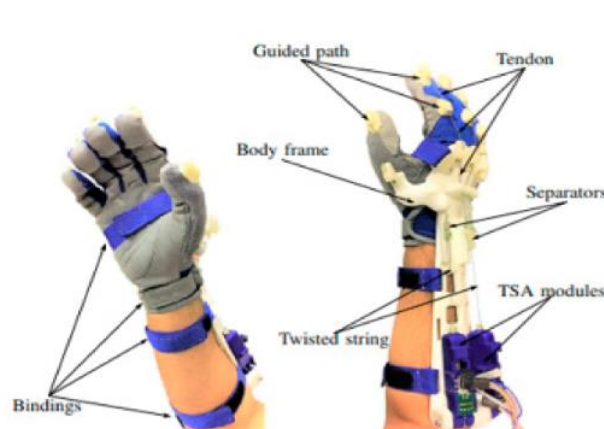


圖 75：ExoTen，基於 McKibben 的人造肌肉 (Hosseini, Sengül, Pane, De Schutter, & Bruyninckx, 2018)

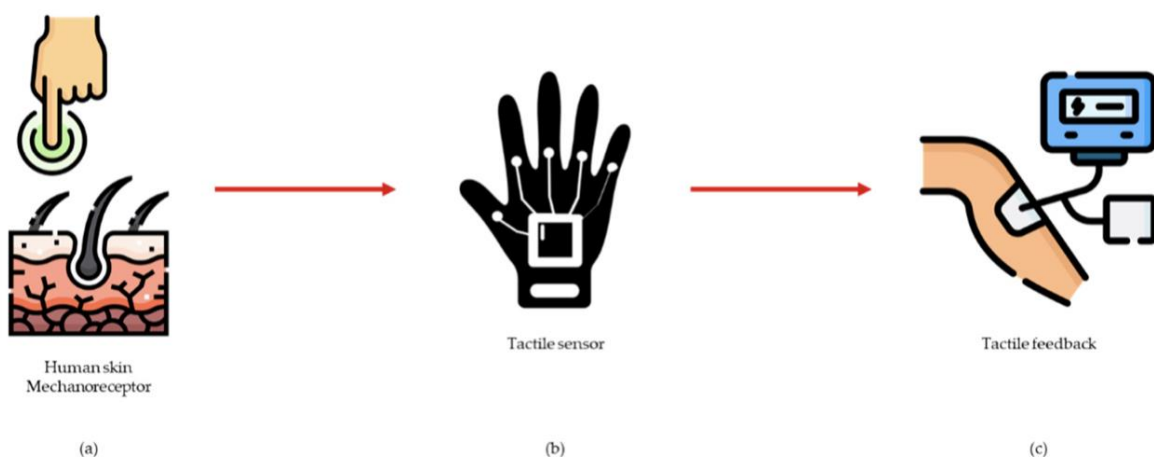


圖 76 未來研究者可能採用的觸覺感知手套範例圖片 (See, Choco, & Chandramohan, 2022)

上圖（圖 76）展示了一款未來研究者可能採用的觸覺感知手套範例圖片。此手套內建有多種感測器，能夠捕捉並放大物體的觸覺資訊，提供視障學生更精確的材質與形狀感知。在未來的研究中，研究者可以探討將這款手套與本研究所提出的混合教學系統相結合，進行更深入的實驗，研究手套在此系統中的應用對視障學生探索周遭環境以及辨別不同材料之能力的提升效果。透過手套感知反饋的強化，學生將能夠更加有效地感知複雜的表面紋理及物體的結構，有助於改進學習效果並提高他們在日常生活及學習中的獨立性。

肆、結論與應用

本研究之結論與應用，可以總結於以下四點

一、本研究提出了視障者多元輔助系統

研究結果顯示，AB 二位視障學生在導入在雙系統（Tactile Pad + RAG 系統）的組合下達到約 27% 的平均正確率提升。此外視障生平均在理解度上有約 1.2 層級的理解感受度提升。本研究驗證了該系統在輔助台灣二位視障學生學習方面的顯著效果。此系統結合了創新多點觸覺反饋裝置和專業個人化的語言模型，使視障學生能夠更有效地掌握抽象概念和空間關係。這種多感官的學習方式不僅提高了學習效率，也增強了學生的自主學習能力。隨著科技與視障者輔助的進展，未來視障者輔具之發展方向極有可能不只強化單一的感官能力，而是如本研究能夠多方面的強化視障生的理解能力，因此本研究為未來多模態視障者輔助系統的發展方向奠定了紮實的基礎。

二、本研究提出了全新的視障生學習輔助系統成果驗證方式—VDEA benchmark

本研究經文獻探討和與特教及資訊老師討論後，提出了全新的視障生學習輔助系統成果驗證方式——VDEA benchmark。這為後續視障生研究者提供了可參考的驗證模式與指標。本研究設計的 VDEA benchmark 不僅涵蓋了傳統正確率評估，還包括了視障學生在使用輔助系統時的理解感受度與主觀評論等多個維度。這種全面的評估方法使得研究結果更具說服力和實用性。此外，VDEA benchmark 的標準化特性也為不同研究間的比較提供了可能，促進了視障教育研究領域的整體發展。

三、本研究開發的 AR 特殊多點輔助板

本研究開發之 Tactile Pad 採用主從（master-slave）架構，為未來擴充與面積可變性提供了可能。其中，研究者提出了使用深度相機為觸覺輔助板提供校正模式。此外，本研究還探索了將觸覺反饋與擴增實境技術結合的可能性。透過 Unity 引擎，我們成功地將虛擬 3D 模型投射到現實環境中，為視障學生提供了更加直觀和互動的學習體驗。這種創新的 AR 輔助模式不僅增強了學習的趣味性，還為視障學生提供了一種全新的方式來理解和探索周遭世界。這些創新可供未來輔具研究者參考與借鑑。

四、本系統開發過程中發掘許多未經深入探索的領域

包括利用視障生輔助文獻與視障生個人化資料來強化語言模型，以及探討專為視障生教育最佳化的二維轉三維模型，例如本研究使用之 Tripo-SR 模型。這些發現為後續視障生研究與發展開啟了全新的篇章。透過整合視障生的個人化學習需求和偏好，RAG 系統能夠提供更加精準和有效的學習支援。未來研究者可結合 VisionLLM 與深度相機進行進一步系統改進。此外，優化的二維轉三維模型有望提高視障生對空間概念的理解，更好地重建課本中的圖片內容，還有望為視障生提供更豐富、更直觀的學習與回饋。

伍、參考文獻

- Acome, E. (2023). Soft Haptic Actuators from Artimus Robotics. Retrieved from <https://www.artimusrobotics.com/post/soft-haptic-actuators-from-artimus-robotics>
- Bourne, R., Steinmetz, J. D., Flaxman, S., Briant, P. S., Taylor, H. R., Resnikoff, S., . . . Vos, T. (2021). Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *The Lancet Global Health*, 9(2), e130-e143. doi:10.1016/S2214-109X(20)30425-3
- Burbach, S., Steiger, A., & Giesser, C. (2022). Suitability Testing of the LucidGloves Prototype 4 for Extended Reality Medical Teaching. *Current Directions in Biomedical Engineering*, 8, 419-422. doi:10.1515/cdbme-2022-1107
- Coates, J., & Bollegala, D. (2018). Frustratingly Easy Meta-Embedding -- Computing Meta-Embeddings by Averaging Source Word Embeddings. doi:10.48550/arXiv.1804.05262
- Derler, S., & Gerhardt, L.-C. (2012). Tribology of Skin: Review and Analysis of Experimental Results for the Friction Coefficient of Human Skin. *Tribology Letters - TRIBOL LETT*, 45. doi:10.1007/s11249-011-9854-y
- Geruschat, D. R., and Smith, A. . (1977). *Low vision and mobility*. In B.B. Blasch, editor; , W.R. Wiener, editor; , and R.L. Welsh, editor. (eds.), *Foundations of orientation and mobility* (2nd ed., pp. 60-103). New York, NY: AFB Press.
- Hosseini, M., Sengül, A., Pane, Y., De Schutter, J., & Bruyninckx, H. (2018). *ExoTen-Glove: A Force-Feedback Haptic Glove Based on Twisted String Actuation System*.
- Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., & Chen, W. (2021). *LoRA: Low-Rank Adaptation of Large Language Models*.
- ihower. (2024). 使用繁體中文評測各家 Embedding 模型的檢索能力. Retrieved from <https://ihower.tw/blog/archives/12167>
- Lawson, R. (2014). Recognizing familiar objects by hand and foot: Haptic shape perception generalizes to inputs from unusual locations and untrained body parts. *Attention, Perception, & Psychophysics*, 76(2), 541-558. doi:10.3758/s13414-013-0559-1
- Lewis, P. S. H., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., . . . Kiela, D. (2020). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*.
- Neelakanteswara, A., Chaudhari, S., & Zamani, H. (2024). *RAGs to Style: Personalizing LLMs with Style Embeddings*.
- Pacchierotti, C., Salvietti, G., Hussain, I., Meli, L., & Prattichizzo, D. (2016). *The hRing: A wearable haptic device to avoid occlusions in hand tracking*.
- See, A. R., Choco, J. A. G., & Chandramohan, K. (2022). Touch, Texture and Haptic Feedback: A Review on How We Feel the World around Us. *Applied Sciences*, 12(9), 4686. Retrieved from <https://www.mdpi.com/2076-3417/12/9/4686>
- Tochilkin, D., Pankratz, D., Liu, Z., Huang, Z., Letts, A., Li, Y., . . . Cao, Y.-P. (2024). TripoSR: Fast 3D Object Reconstruction from a Single Image. *ArXiv*, abs/2403.02151.
- United_Nations. (2024). The 17 Sustainable Development Goals (SDGs). Retrieved from <https://sdgs.un.org/goals>
- Vechev, V., Zarate, J., Lindlbauer, D., Hinchet, R., Shea, H., & Hilliges, O. (2019). *TacTiles: Dual-Mode Low-Power Electromagnetic Actuators for Rendering Continuous Contact and Spatial Haptic Patterns in VR*.
- Wu, J., Zhong, M., Xing, S., Lai, Z., Liu, Z., Wang, W., . . . Dai, J. (2024). *VisionLLM v2: An End-to-End Generalist Multimodal Large Language Model for Hundreds of Vision-Language Tasks*.
- Xie, X., & Livermore, C. (2015). A high-force, out-of-plane actuator with a MEMS-enabled microscissor motion amplifier. *Journal of Physics: Conference Series*, 660, 012026.

doi:10.1088/1742-6596/660/1/012026

- Zhang, J. H., Ramke, J., Jan, C., Bascaran, C., Mwangi, N., Furtado, J. M., . . . Burton, M. J. (2022). Advancing the Sustainable Development Goals through improving eye health: a scoping review. *The Lancet Planetary Health*, 6(3), e270-e280. doi:[https://doi.org/10.1016/S2542-5196\(21\)00351-X](https://doi.org/10.1016/S2542-5196(21)00351-X)
- Zhang, K., González, E., Guo, J., & Follmer, S. (2019). Design and Analysis of High-Resolution Electrostatic Adhesive Brakes Towards Static Refreshable 2.5D Tactile Shape Display. *IEEE Transactions on Haptics*, PP, 1-1. doi:10.1109/TOH.2019.2940219
- 方華馨. (2022). 一位教師對其視障兒巡迴輔導教學之經驗敘說. (碩士). 國立臺東大學, Available from Airiti AiritiLibrary database. (2022)
- 全國法規資料庫. (2024). 2024. Retrieved from <https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=H0080065>
- 徐鉅昌. (1987). 空中教學的根源：視聽教育。 . 視聽教育雙月刊, 29(1), 1 - 19。 .
- 萬明美. (2001). 視障教育(第一版)。臺北市：五南。 .
- 衛生福利部社會及家庭署. (2024). 輔具資源. Retrieved from <https://newrepat.sfaa.gov.tw/>
- 衛生福利部統計處. (2024). 111 學年度全國特殊教育學校高中職學生學習輔具服務量統計概況. Retrieved from <https://newrepat.sfaa.gov.tw/home/download-file/4bc1e2b48929d3940189fd17f33b0b57.pdf>
- 簡群恩. (2014). 外加口述影像之評量調整措施對重度視障生在數學科圖題理解之成效. (碩士). 國立臺北科技大學, Available from Airiti AiritiLibrary database. (2014 年)

【評語】100026

1. 此一作品擬應用三維擴增實境顯示與生成式人工智慧技術，整合到觸摸立體顯示器上，開發有助於視障者之學習輔助平台，值得鼓勵！
2. 建議可加強量化數據之呈現與說明。
3. 建議 3D 立體顯示的設計與製作應有多一些討論，這將決定三維空間重建品質。