

# 一套結合 AR 多模態互動與生成式 AI 的 IoT 實作學習輔助系統

## An IoT Implementation Learning Support System Integrating AR Multimodal

### Interaction and Generative AI

黃宇晨<sup>1</sup>, 金凱儀<sup>1\*</sup>

<sup>1</sup>台灣 東吳大學 資料科學系

\* yuchen.cookie@gmail.com, kychin.scholar@gmail.com

**【摘要】** 物聯網 (IoT) 實作課程涉及軟硬體整合, 學生常因頻繁切換視線於教材與電路之間而產生外在認知負荷。雖然擴增實境 (AR) 能提供視覺化指引, 但現有手持式載具常導致「雙手佔用」問題, 且缺乏針對程式邏輯的即時語意回饋。為此, 本研究提出一套整合 AR 智慧眼鏡與 Google Gemini 大型語言模型的智慧輔助系統架構。本研究設計透過 AR 眼鏡將電路資訊疊加於實體元件, 釋放學習者雙手, 並結合語音辨識與生成式 AI, 旨在提供即時的錯誤偵錯與原理引導。本系統預期能透過虛實整合與多模態互動機制, 降低實作認知負荷並提升問題解決效率。

**【關鍵字】** 擴增實境; 生成式人工智慧; 物聯網教育; 認知負荷; 智慧眼鏡

**Abstract:** In IoT implementation courses, students often face high cognitive load due to switching attention between materials and circuits. While Augmented Reality (AR) offers visual guidance, handheld devices cause "hands-occupied" issues and lack real-time semantic feedback. This study proposes the design of an intelligent support system integrating AR smart glasses with the Google Gemini Large Language Model. The proposed system architecture overlays circuit information onto physical components via AR glasses, enabling a hands-free environment. By combining voice recognition with Generative AI, it is designed to provide real-time debugging assistance and theoretical scaffolding. This multimodal interaction approach is expected to reduce extraneous cognitive load and enhance problem-solving efficiency in IoT learning contexts.

**Keywords:** Augmented reality, Generative AI, IoT education, Cognitive load, Smart glasses

## 1. 前言

隨著物聯網 (IoT) 技術普及, 實體運算 (Tangible Computing) 已成為培養運算思維的關鍵途徑 (Blikstein, 2013; Papert, 1980)。然而, IoT 教學涉及複雜的軟硬體整合, 學生在進行電路組裝與程式撰寫時, 常需在「二維平面教材」與「三維實體電路」間反覆切換視線。這種教材與實物在空間上的分離, 不可避免地產生了注意力分散效應 (Split-attention Effect), 進而增加學習者的外在認知負荷 (Sweller et al., 2011)。

雖然擴增實境 (AR) 技術能透過虛實疊加 (Superimposition) 將抽象訊號 (如電流、腳位) 視覺化, 解決空間對應問題 (Wu et al., 2013), 但目前主流教學應用仍多依賴手持式裝置 (如平板或手機)。在高度依賴雙手協作 (Bimanual Operation) 的 IoT 實作場域中, 手持裝置導致了「雙手佔用 (Hands-occupied)」困境, 迫使學生頻繁中斷操作去滑動螢幕, 嚴重干擾實

作的連續性 (Billingshurst et al., 2001; Wang et al., 2014)。因此，亟需導入穿戴式 AR 智慧眼鏡，以實現「釋放雙手 (Hands-free)」的操作環境，降低物理操作負擔。

此外，現有的 AR 輔助系統多僅提供預設的靜態視覺指引，缺乏對程式邏輯錯誤的即時語意理解與個別化回饋能力 (Akçayır & Akçayır, 2017; Radu, 2014)。當學生遭遇非預期的程式錯誤時，單純的視覺疊加往往無法提供足夠的解釋，導致深層知識建構受阻。為彌補此缺口，引入具備推論能力的生成式人工智慧 (Generative AI) 至關重要。GenAI 能突破規則式系統的限制，根據當下脈絡生成個別化的鷹架 (Scaffolding) 引導，從「資訊展示」升級為「智慧家教」 (Baidoo-Anu & Ansah, 2023; Prather et al., 2023)。

基於上述背景，本研究目前處於系統設計與雛型開發階段，提出一套整合「AR 智慧眼鏡」與「Google Gemini 大型語言模型」的智慧輔助系統。本設計對應認知負荷理論如下：AR 免持操作與空間疊加旨在減少物理干擾與視線分離，以降低外在認知負荷；多模態互動與 AI 回饋則提供適時鷹架，促使深層思考，以提升有助於學習的增生認知負荷 (Germane Cognitive Load)。

## 2. 系統架構設計

本研究提出之系統架構如圖 1 所示，採雲端協作模式，整合前端穿戴式互動載具與後端 AI 推論引擎。

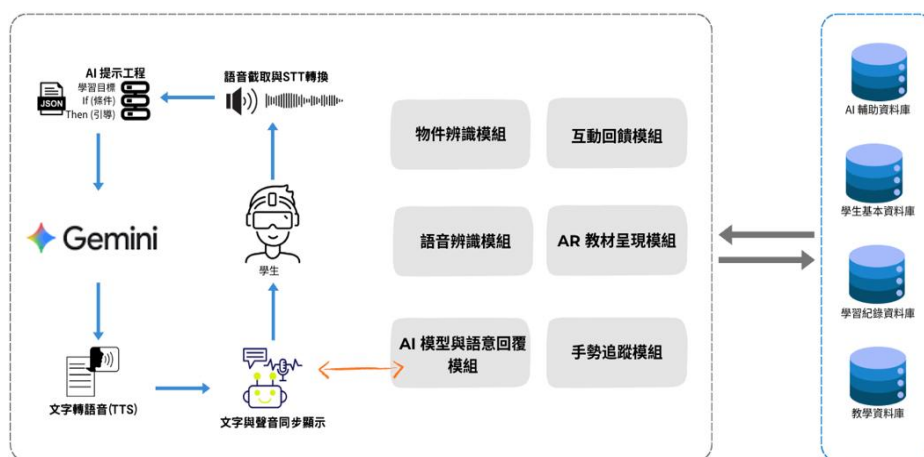


圖 1 系統架構圖

甲、**硬體架構與教材配置** 系統規劃選用佐臻 (Jorjin) J7EF Plus 智慧眼鏡作為顯示終端，利用其雙眼透視技術與前置鏡頭，實現第一人稱視角的虛擬疊加。教學場域設計採用 Arduino 與 ESP32 開發板搭配標準電子模組 (如圖 2 所示)，作為物聯網實作之標的物件。



圖 2 佐臻 J7EF Plus 智慧眼鏡與 IoT 實作教材範例

乙、**系統模組功能設計** 系統軟體包含三大核心模組：

- i. **物件辨識與 AR 呈現：**透過鏡頭即時識別教材元件，並動態渲染 3D 虛擬資訊（如腳位定義、電流動畫），實現虛實整合。
- ii. **手勢追蹤與免持操作：**為解決雙手佔用問題，系統導入直觀手勢操作設計（如圖 3）。使用者可透過左右揮動切換教材，下滑進入 AR 模式，上滑觸發 AI 語音問答，全程無需觸碰實體控制器。

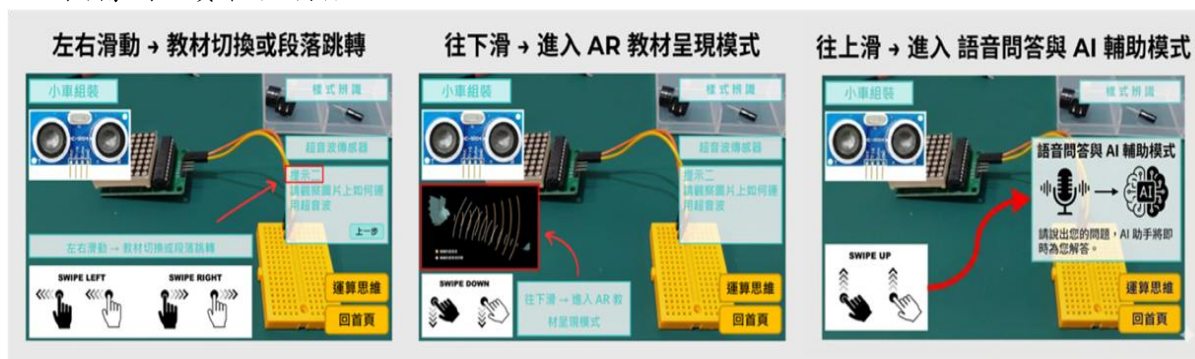


圖 3 手勢追蹤模組功能示意圖

- iii. **Gemini AI 語意回饋：**當觸發 AI 模式時，系統利用語音辨識（STT）將學習者問題轉為文本，針對實體實驗室嘈雜且多名學生同時操作的環境，本系統採用智慧眼鏡內建的指向性麥克風並結合背景降噪演算法（Noise Cancellation），以過濾環境雜音與語音干擾。隨後傳送至 Google Gemini 模型進行分析，並將生成的個別化指導透過文字轉語音（TTS）即時回饋。

### 3. 系統運作流程與應用情境

為驗證系統設計之可行性，本研究規劃以下教學應用情境，模擬學生在進行 Arduino 超音波測距專題時的學習歷程：

- 甲、**虛實整合的組裝引導** 當學生佩戴智慧眼鏡觀看實體超音波感測器時，系統透過物件辨識自動偵測元件，並於鏡片上疊加虛擬接線圖與腳位說明（VCC, Trig, Echo, GND）。學生無需手持平板或講義，雙手可直接進行麵包板插線，並透過下滑手勢切換至下一步驟，保持操作的連續性。
- 乙、**多模態 AI 除錯輔助** 在程式撰寫階段，若學生遭遇如「感測器數值始終為 0」等異常狀況，可透過上滑手勢啟動語音助理提問。為避免系統演變為「直接給答案」的代工工具，後端 Gemini 模型導入了嚴格的系統提示詞（System Prompt）作為護欄。AI 被限制僅能透過蘇格拉底式提問提供鷹架引導（例如提示：「請檢查 Trig 腳位是否已正確定義為輸出模式」），嚴禁直接生成完整程式碼，以此促進學生的主動反思與除錯能力，而非單純的答案複製。
- 丙、**實體課堂的系統擴展性 (Scalability)** 針對多人同時上線的實體教室情境，本系統架構採「邊緣與雲端協同運算」設計，讓輕量級的物件辨識與手勢追蹤直接於 AR 眼鏡端（邊緣運算）即時處理，大幅降低網路頻寬消耗；僅在觸發 AI 問答時，才透過區域網路將文字資料傳送至雲端 Gemini API 進行語意推論。此資源分流機制能有效減輕網路負載，確保多設備連線時的低延遲與系統穩定性。

### 4. 結論與未來工作

本研究提出了一種結合 AR 智慧眼鏡與生成式 AI 的新型態 IoT 學習支援系統架構。相較於傳統教學與手持式 AR 應用，本設計具備以下具體特徵：第一，透過穿戴式載具解決了手持裝置的操作干擾，實現了釋放雙手（Hands-free）的操作環境；第二，利用空間臨近原則（Mayer, 2014）將資訊直接疊加於實物，降低了視線切換帶來的認知負荷；第三，整合 GenAI 的語意推論能力，達成了即時且具備脈絡理解的家教式回饋。本系統展示了虛實整合技術在工程實作教育中的應用潛力。

為進一步驗證本系統之成效，本研究已規劃於後續的 ESP32 物聯網基礎課程中導入準實驗研究。實驗預計招募多名大學生，隨機分派為「AR 眼鏡結合 AI 實驗組」、「純 AR 眼鏡實驗組」與「手持式 AR 對照組」進行比較。研究工具將綜合使用學習成效測驗、三維度認知負荷量表（內在、外在與增生負荷）、科技接受度量表（TAM）及半結構式訪談。期望透過此嚴謹的混合研究設計（Mixed-methods Design），具體且量化地實證本多模態輔助系統在降低外在認知負荷與提升 IoT 實作效能上之實質貢獻。

## 參考文獻

- Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1-11.
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62.
- Billinghurst, M., Kato, H., & Poupyrev, I. (2001). The MagicBook: A transitional AR interface. *Computers & Graphics*, 25(5), 745-753.
- Blikstein, P. (2013). Digital fabrication and 'making' in education: The democratization of invention. *FabLabs: Of machines, makers, and inventors*, 203-221.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.
- Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. Basic Books.
- Prather, J., Denny, P., Leinonen, J., Becker, B. A., Albluwi, I., Craig, M., ... & Smith, J. (2023). The robotic teaching assistant: A systematic review of AI-powered feedback in programming education. *ACM Transactions on Computing Education*, 23(4), 1-36.
- Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Computers & Education*, 68, 150-163.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer Science & Business Media.
- Wang, M., Callaghan, V., Bernhardt, J., White, K., & Peña-Rios, A. (2014). Augmented reality in education and training: Pedagogical approaches and illustrative case studies. *Journal of Ambient Intelligence and Humanized Computing*, 5(6), 1391-1403.
- Wu, H. K., Lee, S. W. Y., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62, 41-49.