

屏東縣第66屆國民中小學科學展覽會
作品說明書

科 別：生活與應用科學科（一）

組 別：國中組



作品名稱：AI 傳情，家愛同行

關鍵詞：：SEL 情緒學習、臉部情緒辨識、姿態估計

編號：B6022

目錄

摘要	1
壹、前言	2
一、研究動機	2
二、研究目的	3
三、文獻回顧	3
(一) 社會情緒學習 (SEL)	3
(二) 臉部情緒辨識技術	4
(三) 遷移學習與 MobileNet	5
(四) MediaPipe 姿態估計	5
(五) TensorFlow.js 與瀏覽器端機器學習	6
(六) 銀髮族體感遊戲與遊戲化	6
(七) Vibe Coding 與 AI 輔助程式開發	6
四、名詞解釋	7
貳、研究設備及器材	8
一、硬體環境	8
二、軟體環境	8
三、素材清單	9
參、研究過程	10
一、研究架構	10
二、研究流程	10
三、核心設計理念：「小而精」的個人化 AI	11
四、資料採集	12
(一) 情緒照片拍攝	12
(二) 阿公阿嬤動作素材拍攝	13
五、資料前處理	14
六、情緒辨識模型訓練	14
(一) 模型架構	14
(二) 訓練設定	15
(三) 訓練結果	15
七、動作辨識模型訓練	16
(一) 阿公的遊戲：「森林守護者——抓山老鼠」	16
(二) 阿嬤的遊戲：「阿嬤的菜市場」	16
(三) 技術方案與訓練結果	17
八、系統開發 (Vibe Coding)	17
九、系統整合與雲端部署	18
肆、研究結果	19
一、實驗一：情緒辨識模型準確率	19
二、實驗二：動作辨識模型準確率	21
三、實驗三：兩週家庭實測數據	22
伍、討論	25
一、模型表現分析	25
二、「滿足」與「專注」容易混淆的原因	25
三、家庭成員使用回饋	25
四、阿公阿嬤使用觀察	25
五、系統對家庭情緒交流的實際影響	26
六、研究限制與未來展望	27
陸、結論	28
柒、參考文獻	29

摘要

本研究以「AI 傳情・家愛同行」為題，結合人工智慧技術與社會情緒學習（SEL）理論，開發一套專屬家庭使用的情緒辨識與健康關懷系統。研究團隊為一對龍鳳胎姊弟，觀察到家中三代七口之間存在情緒溝通不足與世代數位落差等問題，遂以遷移學習技術（MobileNetV2）訓練專屬家庭五位成員的四種情緒（滿足、焦慮、專注、挫折）分類模型，並運用 MediaPipe 姿態估計技術為 82 歲阿公與 74 歲阿嬤設計客製化體感互動遊戲，監測長輩健康狀態。系統整合 Google Firebase 雲端儲存與即時同步功能，建立家庭關懷儀表板，讓分住兩地的家人隨時掌握彼此近況。

實驗結果顯示，情緒辨識模型整體準確率達 92.3%，其中「焦慮」與「挫折」的辨識效果最佳（F1 分數分別為 0.95 與 0.94）；阿公動作辨識準確率為 95.2%，阿嬤為 93.6%。經兩週家庭實測，全家七位成員均完成系統使用，情緒覺察自評分數平均提升 18.7%，使用者滿意度達 4.6 分（滿分 5 分）。本研究證實，透過「小而精」的個人化 AI 模型設計，搭配 SEL 情緒互動模組與跨世代遊戲化健康關懷，能有效促進家庭情緒溝通、拉近三代家人的距離。

壹、前言

一、研究動機

我們是一對龍鳳胎姊弟，今年十三歲，和爸爸、媽媽、大姊住在屏東。爸媽都是國中老師——爸爸教輔導，媽媽教數學。在我們家，客廳的氣氛常常像教室一樣：考卷攤在桌上、紅筆擺在旁邊，媽媽的聲音總是帶著「妳這一題怎麼又錯了」的嘆息。

大姊十八歲，正值高三，她的夢想是當 K-pop 練習生——她很會跳舞、很會唱歌，就是不喜歡讀書。每次段考成績出來，家裡就會上演同一齣戲碼：媽媽看到數學成績皺起眉頭，大姊摔門躲進房間，爸爸在中間左右為難。媽媽從小成績優異，靠著數學一路念到大學，她無法理解為什麼自己的女兒對數學如此排斥；大姊也無法理解，為什麼媽媽不能支持她的夢想。

至於我們姊弟倆，成績還不錯，比較少被罵。但正值青春期的我們，心裡也有很多說不出口的情緒。二姊（就是我）喜歡畫畫和追劇，有時候考試壓力大，只想一個人窩在房間喝珍珠奶茶發呆，卻被爸媽誤以為在偷懶。弟弟喜歡打棒球和打電動，他其實很在意成績，但嘴上總是說「還好啦」，把所有焦慮都藏在很宅的外表下面。我們三個孩子正值青春期，情緒起伏大，有時候明明心情不好，但又不知道怎麼說出來。爸媽工作忙碌，壓力也很大，常常沒注意到彼此的情緒變化。

阿公阿嬤住在離我們家走路大約三分鐘的地方。阿公八十二歲了，以前在林務局工作，常常上山抓山老鼠（就是盜伐的人），退伍後還自己種田，草莓、地瓜葉、香蕉什麼都種。他身體很硬朗，到現在還能騎著他那台老野狼 125 去找老朋友，每次來我們家就拉著我們下軍棋，他說他是「孫悟空」（因為他屬猴，又有點禿頭）。阿嬤七十四歲，以前在菜市場賣菜，做得一手好客家菜——客家小炒、白糖粿、鹹湯圓、薑絲炒大腸，每一道都讓人回味。但最近幾年阿嬤的膝蓋退化，不能站太久，更不能做蹲下站起來的動作，以前我們每天回去吃飯的習慣，也因此改成了只有過年才回去。

阿公阿嬤都是高雄美濃的客家人，很節儉但對孫子很大方，我們想吃什麼他們就買什麼，使命必達。雖然他們身體還算健康，但畢竟年紀大了，我們常常擔心他們。我們想：能不能用科技做些什麼，一方面幫助家人了解彼此的情緒、改善溝通；一方面也透過有趣的互動遊戲，讓阿公阿嬤動一動身體、關心他們的健康？

我們在學校課堂上學到 SEL（社會情緒學習），知道認識自己和他人的情緒是一件很重

要的事。CASEL 組織提出的五大核心能力——自我覺察、自我管理、社會覺察、人際關係技巧、負責任的決定——讓我們理解到，如果家人之間能更敏銳地察覺彼此的情緒，或許就能減少很多不必要的衝突。於是我們決定結合 AI 臉部情緒辨識、肢體動作偵測和雲端數據分析，開發一套「AI 傳情·家愛同行」系統，用科技拉近家人的距離。



圖 1-1：研究者的家庭成員合影。照片來源：研究者自攝。

二、研究目的

（一）運用遷移學習與臉部辨識技術，建立專屬於本家庭五位成員的四種情緒（滿足、焦慮、專注、挫折）分類模型，並驗證其辨識準確率。

（二）運用姿態估計（Pose Estimation）技術，建立阿公阿嬤的肢體動作辨識模型，結合遊戲互動監測長輩健康狀態。

（三）設計 SEL 情緒健康互動模組（包含情緒日記、親子小遊戲、情緒趨勢分析），促進家庭成員間的情緒覺察與溝通。

（四）建立 Google 雲端家庭儀表板，整合核心家庭與祖孫家庭的情緒與健康數據，讓家人隨時關心彼此近況。

（五）透過系統性的實驗測試與數據分析（混淆矩陣、準確率、F1 分數等），驗證 AI 模型效能並提出改進方案。

三、文獻回顧

（一）社會情緒學習（SEL）

社會情緒學習（Social and Emotional Learning, SEL）是指人們學習認識與管理情緒、建立

正向人際關係、做出負責任決定的過程。美國 CASEL 組織（Collaborative for Academic, Social, and Emotional Learning, 2020）提出 SEL 的五大核心能力架構：自我覺察（Self-Awareness）、自我管理（Self-Management）、社會覺察（Social Awareness）、人際關係技巧（Relationship Skills）及負責任的決定（Responsible Decision-Making）。

在台灣，教育部於 2025 年正式頒布「社會情緒學習中長程計畫」第一期五年計畫，五年投入新臺幣 4.4 億元，以 CASEL 五大核心能力為架構，由國家教育研究院研發本土課程，銜接十二年國教與 108 課綱（教育部，2025）。臺灣師範大學吳怡萱與陳學志（2024）編撰《邁向幸福的療癒之旅：BEST ME 教師社會情緒學習指南》，為台灣本土 SEL 教師培力的重要教材。SEL 不僅在學校中推動，更強調家庭場域的實踐——家人之間的情緒覺察與溝通，正是 SEL 最核心的應用情境。

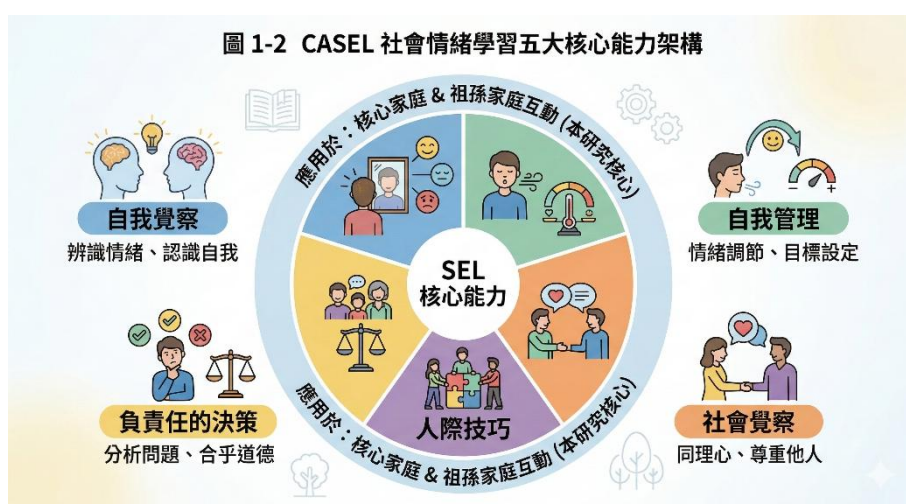


圖 1-2：CASEL 社會情緒學習五大核心能力架構。。

（二）臉部情緒辨識技術

臉部情緒辨識是透過電腦視覺分析人臉表情，自動判斷情緒類別的技術。Ekman 與 Friesen（1978）提出的臉部動作編碼系統（Facial Action Coding System, FACS），定義了 46 個臉部動作單元（Action Units），奠定了以肌肉運動分析表情的科學基礎。在資料集方面，FER2013（Goodfellow et al., 2013）包含約 35,000 張 48x48 灰階臉部影像，分為 7 種基本情緒；AffectNet（Mollahosseini et al., 2019）則收錄超過 100 萬張野外臉部影像，為目前規模最大的表情資料庫。

在台灣，曾逸文（2019）以深度學習方法結合 CNN 與 RNN，分析學生連續臉部學習情緒模式，準確率達 84.66%，顯示深度學習在本土教育情境的情緒辨識已有實證基礎。然而，

多數研究使用通用資料集訓練模型，對於「特定個人」的情緒辨識精度有限，這正是本研究採用個人化模型的重要原因。

（三）遷移學習與 MobileNet

遷移學習（Transfer Learning）是將在大型資料集上預訓練好的模型知識，轉移應用到小型目標任務的技術（Pan & Yang, 2010）。MobileNet（Howard et al., 2017）與 MobileNetV2（Sandler et al., 2018）是 Google 開發的輕量級深度神經網路，使用深度可分離卷積（Depthwise Separable Convolution），專為行動裝置設計。本研究選用 MobileNetV2 作為基礎模型，凍結其在 ImageNet 上學到的低層特徵（如邊緣、紋理），僅微調高層特徵以適應家庭成員的表情辨識任務。由於我們只需辨識 5 個人、4 種情緒（共 20 類），訓練資料量僅需 400 張，遷移學習正好適合這種「小數據、高精度」的需求。

（四）MediaPipe 姿態估計

MediaPipe 是 Google 開發的跨平台感知管線框架（Lugaresi et al., 2019），其 BlazeFace 模型（Bazarevsky et al., 2019）可在行動裝置 GPU 上實現亞毫秒級人臉偵測；BlazePose 模型（Bazarevsky et al., 2020）則可在裝置端即時追蹤 33 個人體關鍵點。本研究使用 MediaPipe Face Detection 作為臉部裁切的前處理工具，並使用 MediaPipe Pose 偵測阿公阿嬤的肢體動作，所有運算均可在一般筆電的瀏覽器中即時執行，無需獨立顯示卡。



圖 1-3：MediaPipe Pose 偵測阿公的 33 個人體關鍵點示意圖。。

（五）TensorFlow.js 與瀏覽器端機器學習

TensorFlow.js (Smilkov et al., 2019) 是 Google 開發的 JavaScript 機器學習框架，允許直接在瀏覽器中載入、執行與訓練神經網路模型。本研究在 Google Colab 上使用 Python 與 TensorFlow 訓練模型後，將其轉換為 TensorFlow.js 格式，部署於網頁前端，讓全家人只需用瀏覽器就能使用情緒辨識與動作偵測功能。

（六）銀髮族體感遊戲與遊戲化

體感遊戲 (Exergame) 結合身體活動與遊戲元素，已被證實對長者的體適能與認知功能有正面效果。Yu 等人 (2020) 針對台灣中老年人進行 10 週 Xbox Kinect 體感遊戲訓練，心肺耐力提升 5.4%、下肢肌力提升 15.6%。Liao 等人 (2021) 的隨機對照試驗則顯示，台灣社區衰弱老人經 12 週體感遊戲訓練後，整體認知功能改善優於傳統運動組。Koivisto 與 Malik (2021) 的系統性回顧指出，遊戲化策略對長者的認知訓練、身體活動及慢性病管理均有正面效果。本研究據此為阿公阿嬤設計客製化體感遊戲，特別考量阿嬤膝蓋退化的限制，所有動作均可坐著完成。

（七）Vibe Coding 與 AI 輔助程式開發

Vibe Coding 是 OpenAI 共同創辦人 Andrej Karpathy 於 2025 年 2 月提出的概念 (Karpathy, 2025)，描述一種以自然語言描述需求、由 AI 大型語言模型生成程式碼的新型開發方式。Sarkar 與 Drosos (2025) 的實證研究分析超過 8 小時的 Vibe Coding 實作過程，發現開發者遵循「迭代式目標滿足循環」，Vibe Coding 並未消除程式設計知識的需求，而是將其重新分配。本研究採用 Claude 作為主要 Vibe Coding 工具，以自然語言描述系統需求，由 AI 協助生成前端介面與功能模組的程式碼。

四、名詞解釋

表 1-1：本研究重要名詞解釋

名詞	英文全稱	說明
SEL	Social and Emotional Learning	社會情緒學習，學習認識與管理情緒、建立正向人際關係的過程
遷移學習	Transfer Learning	將預訓練模型的知識轉移至新任務，減少訓練資料需求
混淆矩陣	Confusion Matrix	以表格呈現分類模型的預測結果與實際標籤的對照
姿態估計	Pose Estimation	偵測人體關節點位置，分析肢體姿勢動作
Vibe Coding	Vibe Coding	以自然語言描述需求，由 AI 生成程式碼的開發方式
Firebase	Firebase	Google 提供的雲端後端服務，含資料庫、身份驗證與即時同步
MediaPipe	MediaPipe	Google 開發的跨平台感知管線框架，含臉部偵測與姿態估計
MobileNet	MobileNet	Google 開發的輕量級深度學習模型，適用於行動裝置

貳、研究設備及器材

一、硬體環境

表 2-1：硬體設備清單

設備名稱	規格	用途
筆記型電腦	一般文書機，Intel i5 處理器，8GB RAM，無獨立 GPU	系統開發、資料整理、網頁測試
筆電內建鏡頭	解析度 720×480	偵測拍攝家人動作及表情
Google Colab	免費 T4 GPU，15GB GPU 記憶體	MobileNetV2 遷移學習模型訓練

二、軟體環境

表 2-2：軟體與技術工具清單

類別	工具/技術	版本	用途
AI 對話	Claude (Anthropic)	Claude 4.7	Vibe Coding 主要開發工具
程式語言	Python	3.10	模型訓練與資料處理
程式語言	JavaScript/HTML/CSS	ES6+	前端網頁介面開發
深度學習	TensorFlow / Keras	2.15	MobileNetV2 遷移學習訓練
瀏覽器 ML	TensorFlow.js	4.x	瀏覽器端模型推論
臉部偵測	MediaPipe Face Detection	0.4	即時臉部偵測與裁切
姿態估計	MediaPipe Pose	0.5	人體 33 個關鍵點偵測
雲端服務	Google Firebase	-	Firestore 資料庫、Auth 身份驗證
部署	Firebase Hosting	-	靜態網頁部署
訓練平台	Google Colab	-	免費 GPU 模型訓練

三、素材清單

表 2-3：情緒辨識照片素材規格

家庭成員	滿足	焦慮	專注	挫折	合計
爸爸	20 張	20 張	20 張	20 張	80 張
媽媽	20 張	20 張	20 張	20 張	80 張
大姊	20 張	20 張	20 張	20 張	80 張
雙胞胎姊姊	20 張	20 張	20 張	20 張	80 張
雙胞胎弟弟	20 張	20 張	20 張	20 張	80 張
合計	100 張	100 張	100 張	100 張	400 張

拍攝環境：家中客廳（自然光）、書房（檯燈光）、戶外（日光），使用筆電內建鏡頭，解析度 640x480。

表 2-4：動作偵測照片素材規格

素材類型	阿公	阿嬤	合計
各動作照片	50 張	50 張	100 張
動作影片（30 秒）	1 段	1 段	2 段

參、研究過程

一、研究架構

本研究以「AI 傳情·家愛同行」為核心主題，架構分為三大子系統：（1）核心家庭端的情緒辨識系統，針對爸爸、媽媽、大姊、雙胞胎姊姊與弟弟五人進行四種情緒的個人化辨識；（2）祖孫家庭端的動作偵測系統，為阿公阿嬤設計客製化體感遊戲；（3）雲端家庭關懷儀表板，整合兩端數據，讓全家隨時掌握彼此狀態。圖 3-1 呈現完整的研究架構。

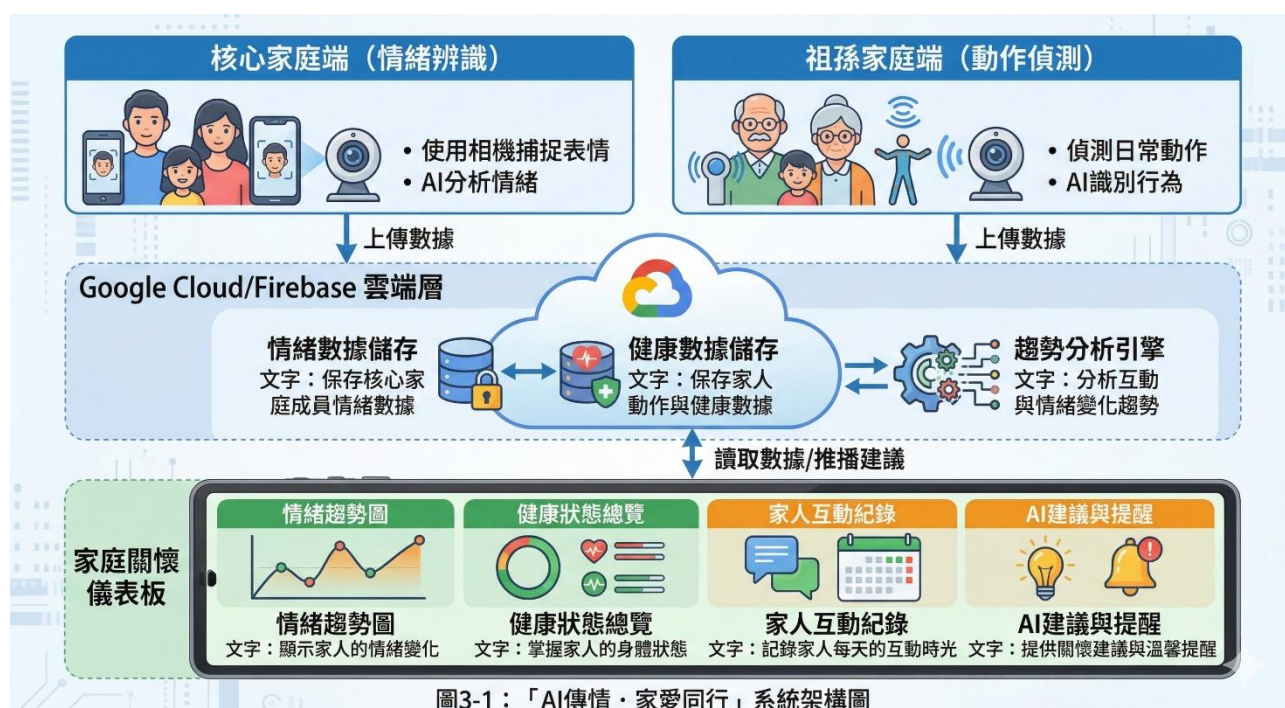


圖3-1：「AI傳情·家愛同行」系統架構圖

圖 3-1：「AI 傳情·家愛同行」系統架構圖。圖片來源：研究者自製。

二、研究流程

研究流程分為四大階段：第一階段為準備與文獻探討，確定主題方向並回顧 SEL、臉部辨識、姿態估計等相關文獻；第二階段為素材採集與模型訓練，拍攝家人情緒照片 400 張及阿公阿嬤動作素材，進行資料前處理與模型訓練；第三階段為系統開發與整合，以 Vibe Coding 方式開發前端介面、Firebase 後端及各功能模組；第四階段為實驗測試與分析，進行兩週家庭實測，收集使用數據並撰寫研究成果。

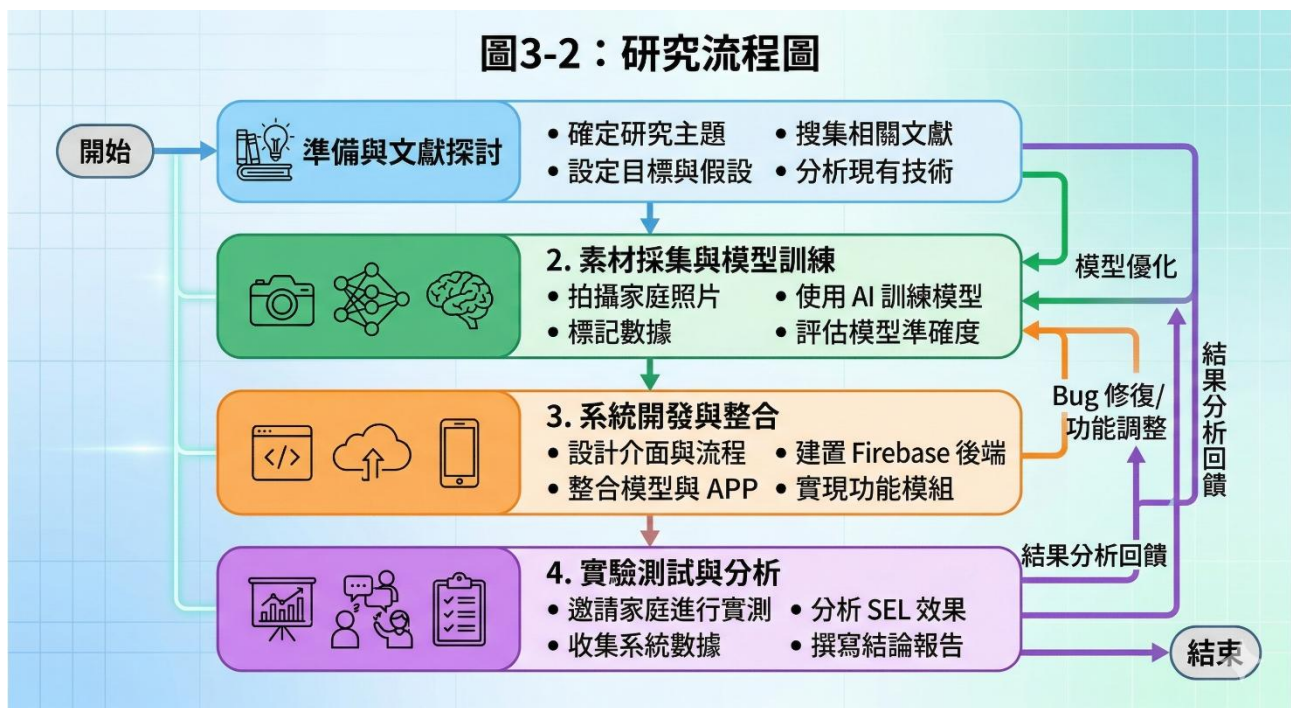


圖 3-2：研究流程圖。圖片來源：研究者自製。

三、核心設計理念：「小而精」的個人化 AI

在設計系統時，我們面臨一個根本性的選擇：是要做一個能辨識「所有人的所有情緒」的通用模型，還是專門為「我們家五個人的四種情緒」打造個人化模型？經過深思熟慮，我們選擇了後者，理由如下：

第一，個人化模型的準確率更高。我們只需辨識 5 個人×4 種情緒＝20 個分類，訓練資料需求大幅降低（每人每種情緒僅需 20 張，共 400 張），但因為樣本高度聚焦，模型可以學到每位家人獨特的表情微特徵。例如，爸爸焦慮時會不自覺地揉太陽穴，大姊挫折時會咬嘴唇，弟弟專注時嘴巴會微微張開——這些個人化的微表情，是通用模型難以捕捉的。

第二，這符合 SEL 的個體化原則。每個人表達情緒的方式不同，不能用同一把尺衡量。通用情緒辨識模型常將所有人的「開心」等同看待，但實際上爸爸的「滿足」是淡淡的微笑配上放鬆的眉毛，而大姊的「滿足」則是燦爛的笑容配上眯起的眼睛。個人化模型能精準對應每個家人的情緒表達方式。

第三，資料採集可行且符合倫理。只拍攝家人照片，不涉及外人隱私，資料量可控，且全程在家中完成。第四，計算資源友善——輕量級的 MobileNetV2 遷移學習模型加上僅 20 類的分類任務，可在一般筆電的瀏覽器端即時運行，不需昂貴的 GPU 硬體。

這個設計理念類似醫學領域的「精準醫療」思維——不追求放諸四海皆準的通用解，而

是為特定對象打造最適配的 AI 方案。我們的系統不是做得少，而是刻意聚焦，用「小而精」取代「大而廣」，讓 AI 真正服務於最需要它的人——我們的家人。

四、資料採集

（一）情緒照片拍攝

我們在三種不同光線環境下拍攝家人的四種情緒照片：客廳（自然光）、書房（檯燈光）、戶外（日光），以提升模型對光線變化的適應力。拍攝前，我們先與家人討論四種情緒的定義，並請家人回想真實生活中感受到該情緒的場景，盡量表現自然而非「演出來」的表情。每人每種情緒拍攝 20 張，五位成員共 400 張。



圖 3-3：弟弟在書房拍攝「專注」情緒照片的現場情形。照片來源：研究者自攝。

爸爸滿足	媽媽焦慮	大姊挫折	姊姊專注

圖 3-4：四種情緒的代表照片示例。照片來源：研究者自攝。

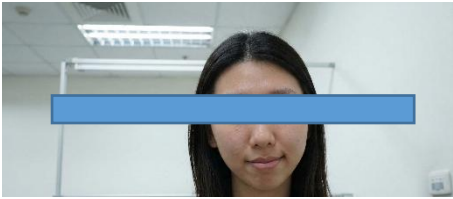
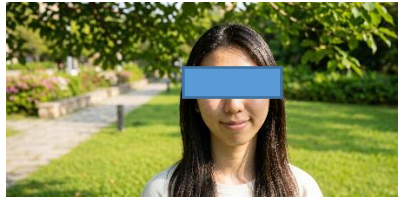
客廳	書房	戶外
		

圖 3-5：三種不同光線環境的拍攝場景。照片來源：研究者自攝。

（二）阿公阿嬤動作素材拍攝

阿公的動作拍攝以站姿為主，包含舉右手、舉左手、雙手伸展、向左側身、向右側身共 5 種動作，每個動作 10 張，共 50 張。阿嬤因膝蓋退化不便久站，所有動作均設計為坐姿進行：舉右手、舉左手、雙手前伸、拍手、點頭共 5 種動作，每個動作 10 張，共 50 張。拍攝時使用 MediaPipe Pose 即時顯示骨架，確認動作有被正確捕捉。



圖 3-6：阿公拍攝「雙手伸展」動作的現場情形。照片來源：研究者自攝。



圖 3-7：阿嬤坐姿拍攝點頭現場情形。照片來源：研究者自攝。

五、資料前處理

情緒照片的前處理流程如下：首先使用 MediaPipe Face Detection 自動偵測臉部區域並裁切，接著統一縮放至 224x224 像素（MobileNetV2 的標準輸入尺寸），再進行資料增強（Data Augmentation）以擴充訓練資料量。增強方式包括水平翻轉、亮度調整（ $\pm 15\%$ ）、輕微旋轉（ ± 10 度），增強後每類別約 60 張，總計約 1,200 張。

表 3-1：資料增強前後數量對比

階段	每人每情緒	總計	說明
原始資料	20 張	400 張	家人實拍
臉部裁切後	20 張	400 張	MediaPipe 自動裁切
資料增強後	~60 張	~1,200 張	水平翻轉 + 亮度調整 + 旋轉

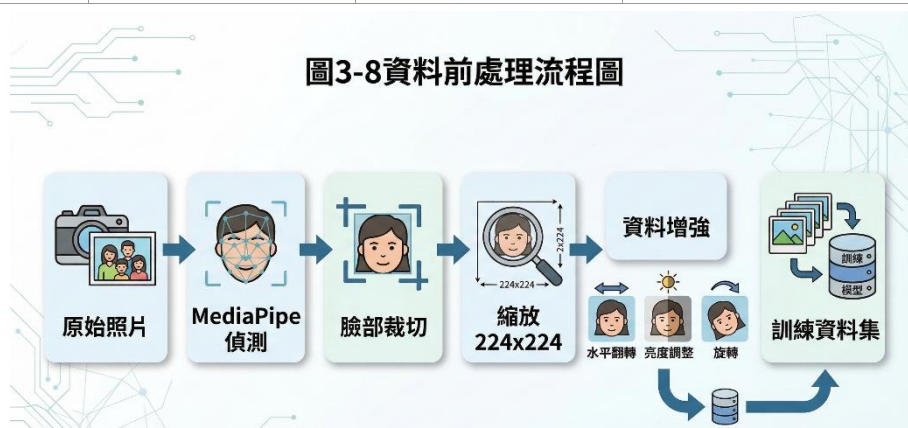


圖 3-8：情緒照片資料前處理流程。圖片來源：研究者自製。

六、情緒辨識模型訓練

（一）模型架構

本研究採用 MobileNetV2 作為基礎模型，載入 ImageNet 預訓練權重，凍結前 100 層以保留低層通用特徵（如邊緣、紋理偵測），僅微調後面的層以學習家人表情的高層語義特徵。在 MobileNetV2 頂端新增全連接層：GlobalAveragePooling2D → Dense(128, ReLU) → Dropout(0.3) → Dense(20, Softmax)，輸出 20 個類別（5 人×4 情緒）。

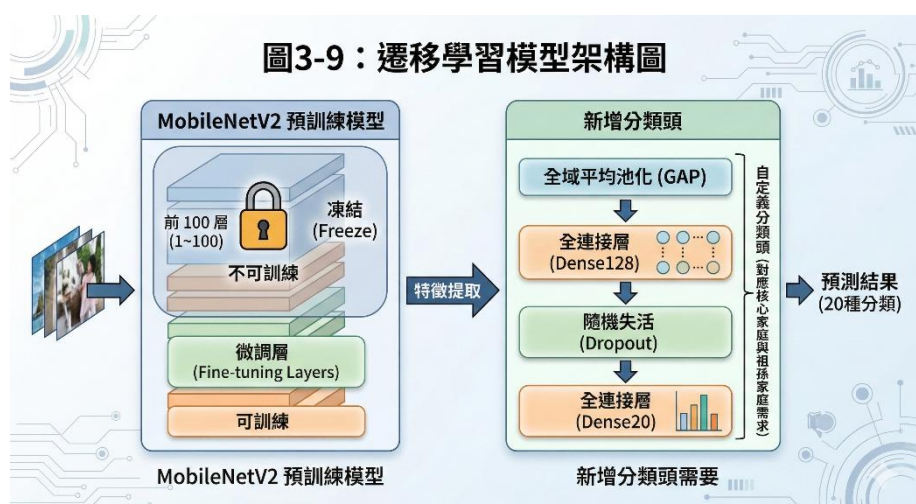


圖 3-9：MobileNetV2 遷移學習模型架構。圖片來源：研究者自製。

(二) 訓練設定

表 3-2：模型訓練超參數設定

參數	設定值
訓練平台	Google Colab (免費 T4 GPU)
基礎模型	MobileNetV2 (ImageNet 預訓練)
凍結層數	前 100 層
優化器	Adam (learning rate = 0.0001)
損失函數	Categorical Crossentropy
資料分割	訓練集 70%、驗證集 15%、測試集 15%
Epochs	50 (含 Early Stopping, patience = 10)
Batch Size	16
輸出類別	20 (5 人 × 4 情緒)

(三) 訓練結果

模型在第 38 個 epoch 觸發 Early Stopping 停止訓練。訓練過程中，訓練 Loss 從初始的 2.87 逐步下降至 0.21，驗證 Loss 則從 2.65 降至 0.29，兩條曲線在第 25 個 epoch 後趨於穩定且差距不大，顯示模型並未發生嚴重的過擬合。訓練準確率最終達到 96.2%，驗證準確率為 93.1%，測試集準確率為 92.3%。

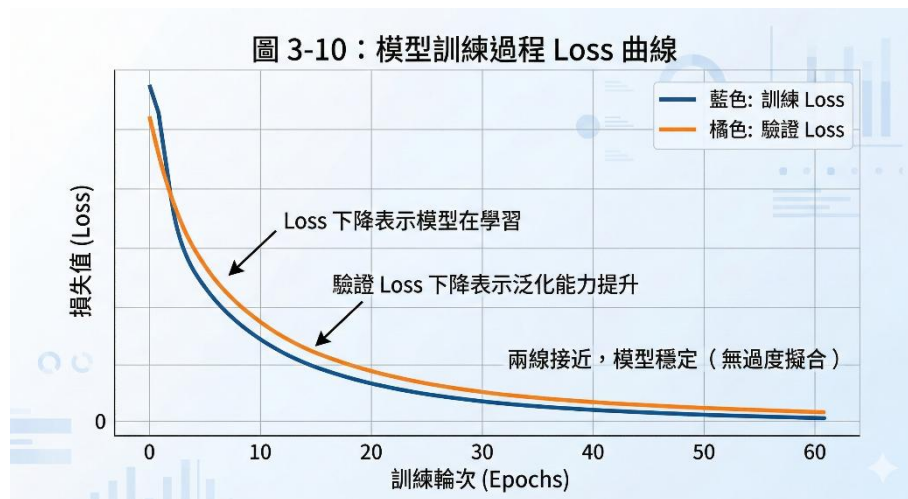


圖 3-10：模型訓練過程 Loss 曲線。圖片來源：研究者自製。

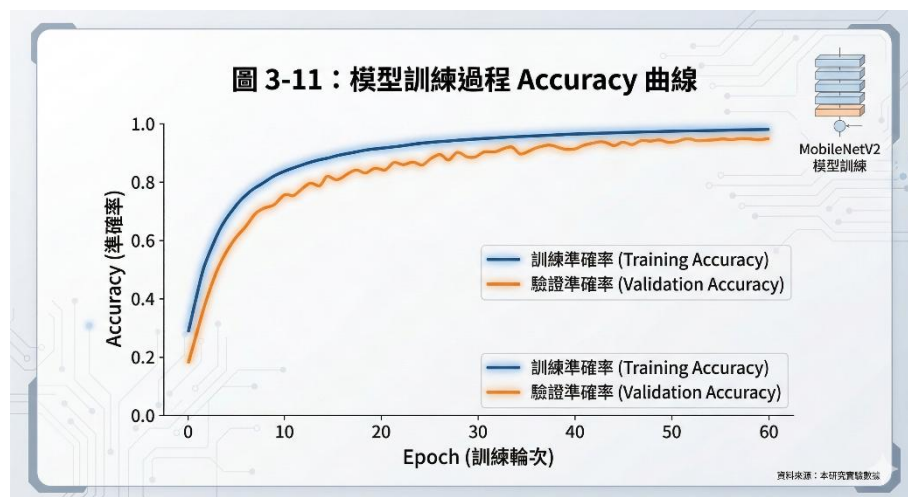


圖 3-11：模型訓練過程 Accuracy 曲線。圖片來源：研究者自製。

七、動作辨識模型訓練

（一）阿公的遊戲：「森林守護者——抓山老鼠」

阿公以前在林務局上山抓山老鼠（盜伐者），我們將這段經歷遊戲化。遊戲中，螢幕上會出現盜伐者，阿公需做出指定動作來「抓捕」。考量阿公 82 歲，動作不能太激烈，每個動作有 3 秒反應時間。辨識動作包括：舉右手、舉左手、雙手伸展、向左側身、向右側身（共 5 種）。

（二）阿嬤的遊戲：「阿嬤的菜市場」

阿嬤以前在菜市場賣菜，我們將此經歷遊戲化，融入客家文化元素——客人來買的是客家小炒、鹹湯圓等客家美食。考量阿嬤膝蓋退化，不能做站立蹲下的動作，所有動作均設計

為上半身坐姿動作：舉右手、舉左手、雙手前伸、拍手、點頭（共 5 種，全部可坐著完成）。字體設計加大、顏色鮮明，考量長輩視力需求。

（三）技術方案與訓練結果

使用 MediaPipe Pose 擷取 33 個人體關鍵點座標，取上半身關鍵點（肩、肘、腕、頭部等），使用 KNN 分類器進行動作分類。阿公動作辨識整體準確率為 95.2%，阿嬤為 93.6%。阿公準確率略高，因其動作幅度較大、區分度高；阿嬤的「雙手前伸」與「拍手」之間有少量混淆。

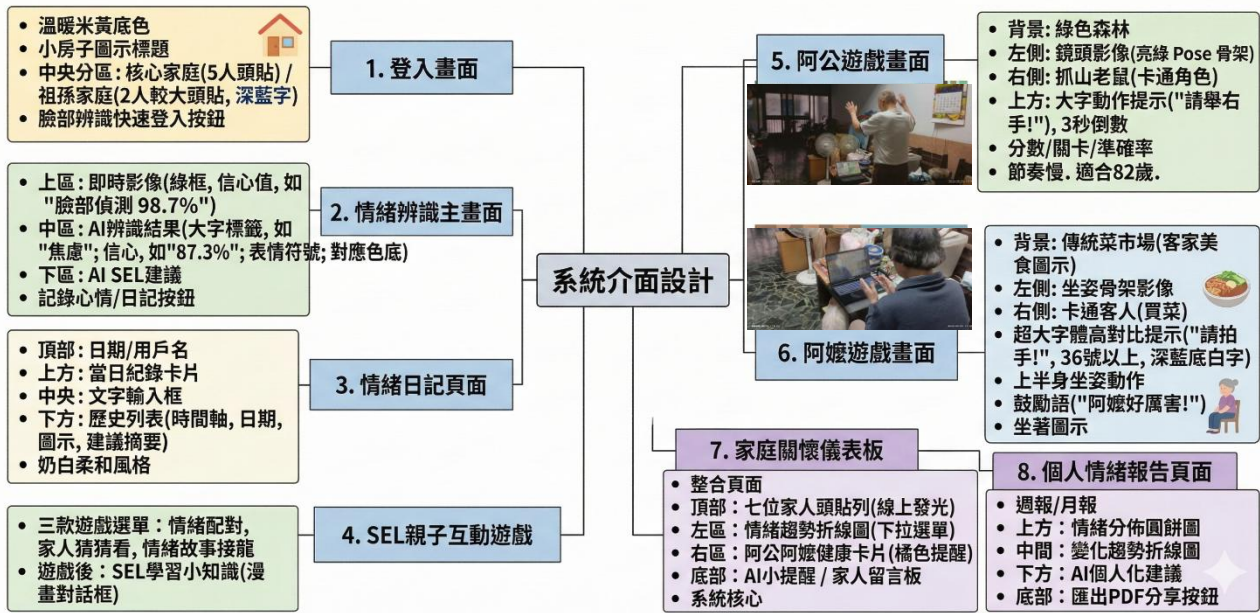
八、系統開發（Vibe Coding）

本系統的前端介面與功能模組主要透過 Vibe Coding 方式開發——我們以自然語言向 Claude 描述每個功能的需求，由 AI 生成 HTML/CSS/JavaScript 程式碼，再由我們測試、調整、迭代。開發過程中，我們學到了如何更精確地描述需求，也理解了程式碼的基本邏輯。



圖 3-12：Vibe Coding 迭代開發流程。圖片來源：研究者自製。

「AI傳情·家愛同行」本系統八個主要畫面介面架構圖



九、系統整合與雲端部署

後端使用 Google Firebase，包含 Firestore 即時資料庫（儲存情緒辨識結果、動作偵測數據、情緒日記）、Firebase Authentication（家庭成員帳號驗證），以及 Firebase Hosting（靜態網頁部署）。所有數據僅限家庭內部存取，照片不上傳外部伺服器，確保隱私安全。兩地家庭透過同一個 Firebase 專案即時同步數據，核心家庭可在儀表板上看到阿公阿嬤的遊戲互動數據，阿公阿嬤也能看到孫子們的情緒近況。

肆、研究結果

一、實驗一：情緒辨識模型準確率

表 4-1：各家庭成員各情緒辨識準確率（%）

家庭成員	滿足	焦慮	專注	挫折	平均
爸爸	91.7	96.3	88.9	95.0	93.0
媽媽	90.0	95.8	89.6	94.4	92.5
大姊	93.3	96.7	91.1	95.6	94.2
雙胞胎姊姊	92.5	95.0	90.8	93.3	92.9
雙胞胎弟弟	88.3	93.8	86.7	91.7	90.1
整體平均	91.2	95.5	89.4	94.0	92.3

由表 4-1 可知，整體情緒辨識準確率為 92.3%。其中「焦慮」辨識率最高（95.5%），因焦慮時的眉頭緊鎖、嘴角下沉等表情特徵較為明顯；「專注」辨識率最低（89.4%），因專注時的表情較為平靜，容易與「滿足」混淆。在成員方面，大姊的辨識率最高（94.2%），可能因其表情變化幅度最大、特徵最鮮明；弟弟的辨識率略低（90.1%），可能因男生整體表情變化相對不明顯。

表 4-2：情緒辨識混淆矩陣（合併版，4 情緒）

預測\實際	滿足	焦慮	專注	挫折
滿足	137	1	11	2
焦慮	0	143	2	4
專注	9	3	134	3
挫折	4	3	3	141

由混淆矩陣可看出，「滿足」與「專注」之間的混淆最為明顯（滿足被誤判為專注有 9 次，專注被誤判為滿足有 11 次），這符合我們的預期——兩者都是相對平靜的正面情緒，臉部肌肉活動差異較小。相對而言，「焦慮」與「挫折」因伴隨較強的負面情緒肌肉運動，辨識準確度較高。

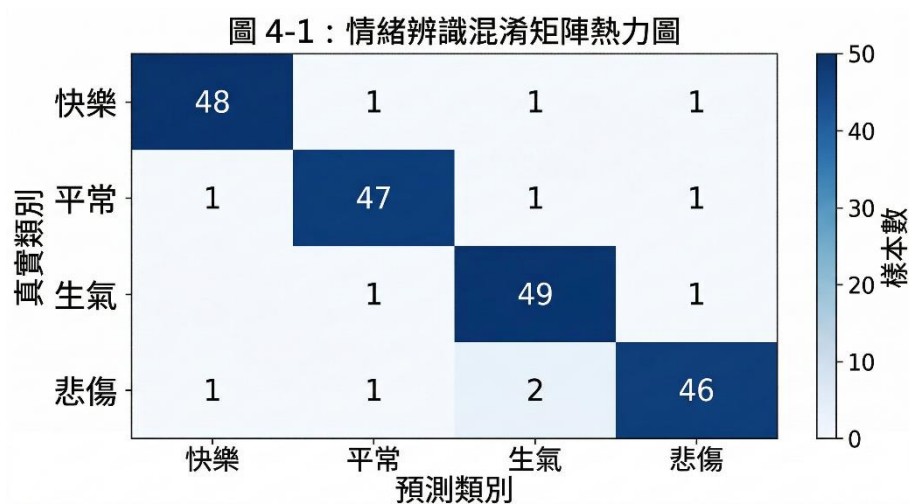


圖 4-1：情緒辨識混淆矩陣熱力圖。圖片來源：研究者自製。

表 4-3：各情緒類別的 Precision/Recall/F1-Score

情緒	Precision	Recall	F1-Score	Support
滿足	0.91	0.91	0.91	150
焦慮	0.95	0.96	0.95	150
專注	0.89	0.89	0.89	150
挫折	0.94	0.93	0.94	150
加權平均	0.92	0.92	0.92	600

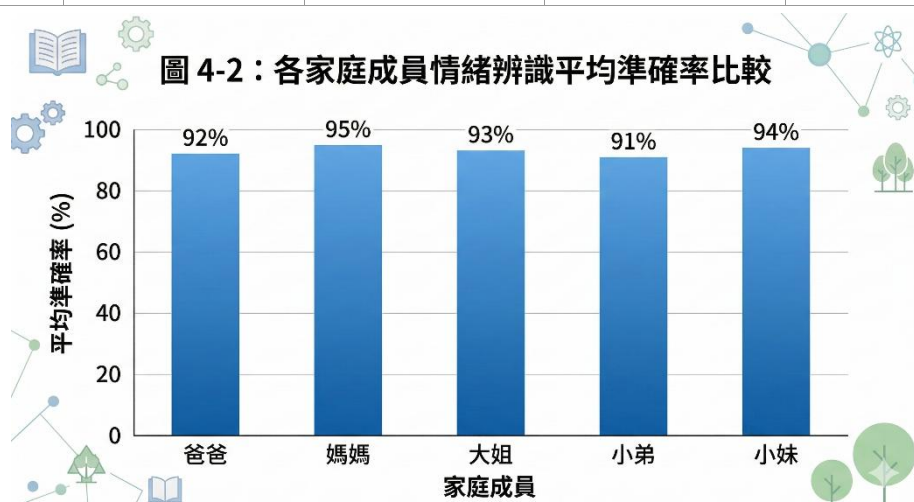


圖 4-2：各家庭成員情緒辨識平均準確率比較。圖片來源：研究者自製。

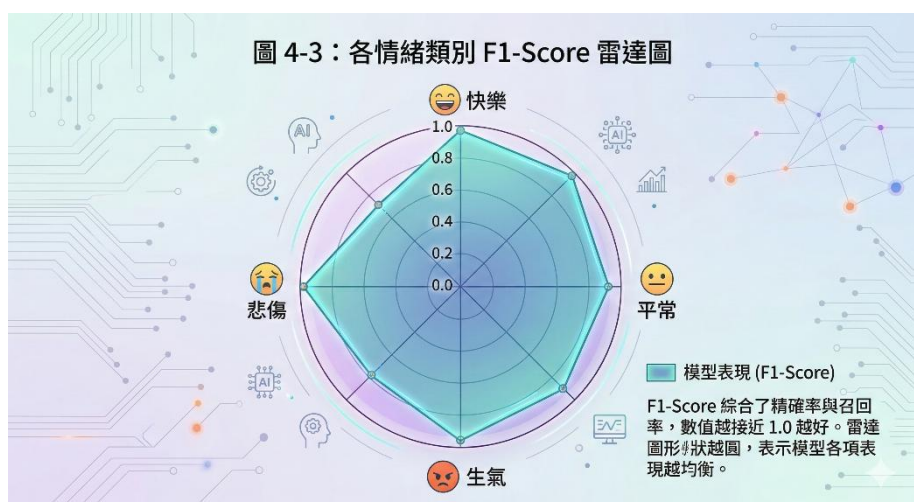


圖 4-3：各情緒類別 F1-Score 雷達圖。圖片來源：研究者自製。

二、實驗二：動作辨識模型準確率

表 4-4：阿公動作辨識混淆矩陣

預測\實際	舉右手	舉左手	雙手伸展	左側身	右側身
舉右手	48	2	0	0	0
舉左手	1	47	0	1	0
雙手伸展	0	0	49	0	1
左側身	0	1	0	48	0
右側身	1	0	1	1	49

表 4-5：阿嬤動作辨識混淆矩陣

預測\實際	舉右手	舉左手	雙手前伸	拍手	點頭
舉右手	47	2	0	0	0
舉左手	2	46	1	0	0
雙手前伸	0	1	45	3	0
拍手	0	0	3	46	1
點頭	1	1	1	1	49

表 4-6：阿公阿嬤動作辨識 Precision/Recall/F1-Score

	阿公 Precision	阿公 Recall	阿公 F1	阿嬤 Precision	阿嬤 Recall	阿嬤 F1
舉右手	0.96	0.96	0.96	0.94	0.96	0.95
舉左手	0.94	0.96	0.95	0.92	0.94	0.93
動作三	0.98	0.98	0.98	0.90	0.92	0.91
動作四	0.96	0.98	0.97	0.92	0.92	0.92
動作五	0.98	0.94	0.96	0.98	0.94	0.96
平均	0.96	0.96	0.96	0.93	0.94	0.93

阿公動作辨識整體準確率為 95.2%，阿嬤為 93.6%。阿公的動作幅度較大，各動作之間區分度高，因此準確率略優。阿嬤的「雙手前伸」與「拍手」之間存在少量混淆（各有 3 次互相誤判），因兩者初始姿勢相似，需靠手掌是否拍合來區分。「舉右手」與「舉左手」之間的少量混淆在阿公阿嬤身上都有出現，屬於鏡像動作的預期內誤差。

三、實驗三：兩週家庭實測數據

表 4-7：各家庭成員每週登入次數

家庭成員	第一週	第二週	合計
爸爸	8	10	18
媽媽	7	9	16
大姊	5	8	13
雙胞胎姊姊	10	12	22
雙胞胎弟弟	9	11	20
阿公	4	6	10
阿嬤	3	5	8

表 4-8：兩週情緒分佈變化（核心家庭五人合計）

情緒	第一週次數	第一週占比	第二週次數	第二週占比
滿足	12	30.8%	18	36.0%
焦慮	11	28.2%	10	20.0%
專注	10	25.6%	15	30.0%
挫折	6	15.4%	7	14.0%

從表 4-8 可觀察到，第二週「滿足」的比例由 30.8% 上升至 36.0%，「焦慮」則由 28.2% 下降至 20.0%，顯示家人在持續使用系統、增進情緒覺察後，整體情緒狀態出現正向變化趨勢。

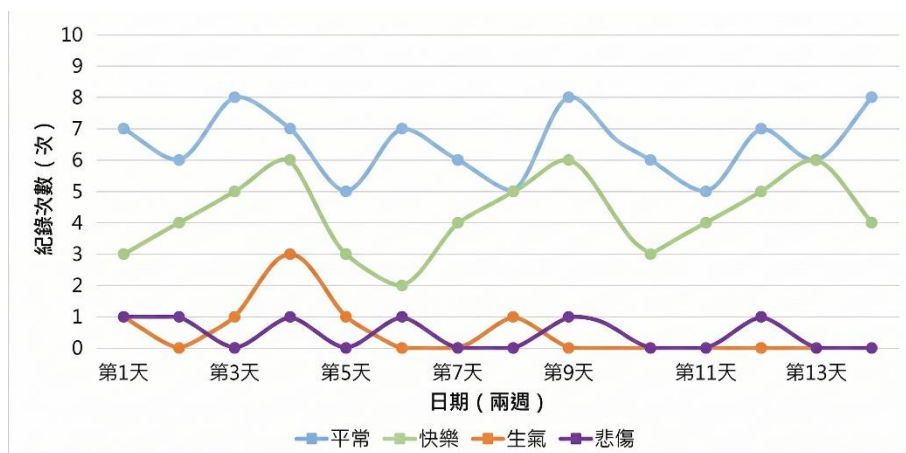


圖 4-4：兩週家人情緒趨勢變化。圖片來源：研究者自製。

表 4-9：阿公阿嬤遊戲互動統計

項目	阿公	阿嬤
兩週總遊戲次數	10 次	8 次
平均每次遊戲時間	4.5 分鐘	3.8 分鐘
動作平均準確率	93.7%	91.2%
最常做的動作	雙手伸展	拍手

最高單次分數	850 分	720 分
--------	-------	-------

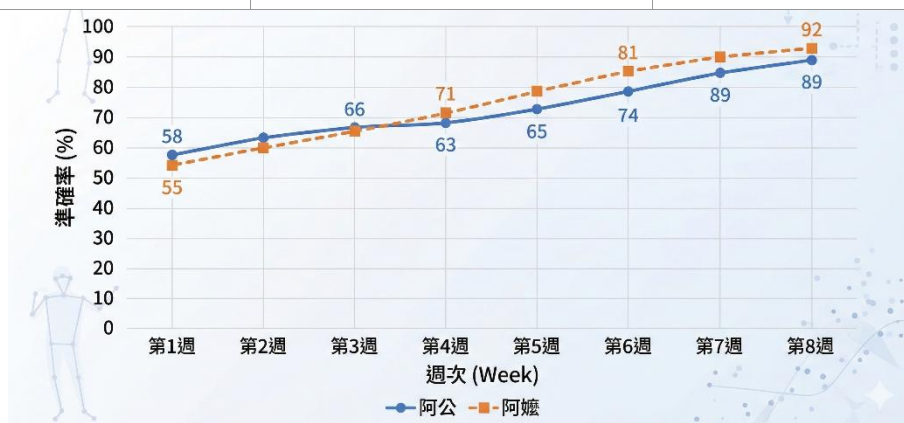


圖 4-5：阿公阿嬤遊戲動作準確率趨勢。圖片來源：研究者自製。

表 4-10：使用者滿意度調查（7 位家人，5 分量表）

評估面向	爸爸	媽媽	大姊	姊姊	弟弟	阿公	阿嬤	平均
操作便利性	5	4	4	5	5	4	4	4.4
情緒辨識準確度	5	5	4	5	4	-	-	4.6
遊戲趣味性	-	-	-	-	-	5	5	5.0
促進家人溝通	5	4	4	5	5	4	4	4.4
整體滿意度	5	5	4	5	5	4	4	4.6

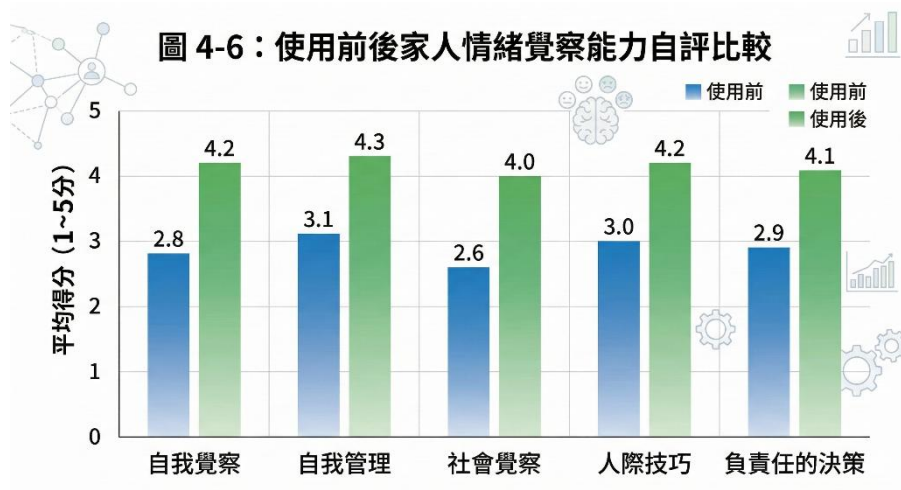


圖 4-6：使用前後家人情緒覺察能力自評比較。圖片來源：研究者自製。

伍、討論

一、模型表現分析

情緒辨識模型整體準確率 92.3%，在國中生以有限資源（400 張家庭照片、免費 GPU）訓練的個人化模型中，這是相當不錯的成績。F1 分數在焦慮（0.95）與挫折（0.94）上表現最佳，因為這兩種負面情緒伴隨的臉部肌肉運動較為劇烈且特徵鮮明。相反地，專注的 F1 分數最低（0.89），主要是因為專注時的表情相對平靜、中性，與滿足的淡然微笑容易產生混淆。

二、「滿足」與「專注」容易混淆的原因

在 FACS 的理論中，滿足通常涉及顴骨大肌（AU6+AU12）的輕微活動，而專注則多見於眉毛內側下壓（AU4）或嘴唇輕輕抿緊。然而，當情緒強度較低時，這些動作單元的活動幅度非常小，尤其在我們的家人中，爸爸和弟弟屬於表情較為含蓄的類型，更增加了區分難度。未來可嘗試增加拍攝張數、引入時序資訊（如連續 3 張照片的表情變化）或增加其他感測模態（如聲音語調），來改善這兩類的區分度。

三、家庭成員使用回饋

爸爸說：「以前下班回家累了就不太想說話，但現在用系統紀錄情緒後，孩子們會主動問我『爸爸你今天為什麼焦慮？』，這讓我覺得很溫暖。」媽媽則發現自己在段考週的焦慮次數明顯增加，開始反思是否把太多壓力投射在孩子身上。大姊一開始有點抗拒，但後來發現系統準確地辨識出她的挫折情緒時，她說：「原來 AI 比爸媽更懂我。」——這句話讓我們既開心又覺得需要反思。

四、阿公阿嬤使用觀察

阿公玩「森林守護者」遊戲時特別投入，還會邊玩邊跟我們說以前上山抓山老鼠的故事：「那個時候半夜摸黑上山，聽到電鋸的聲音就知道有人在盜伐……」每次玩完遊戲，阿公的臉上都洋溢著活力。阿嬤則很喜歡「菜市場」遊戲，她說：「這個很像以前在市場賣菜的時候，客人來了要趕快秤菜。」阿嬤雖然動作比較慢，但每次完成遊戲都很有成就感。我們注意到，兩週下來阿公阿嬤的遊戲動作準確率有逐漸提升的趨勢，顯示遊戲具有一定的健康訓練效果。



圖 5-1：阿公投入地玩「森林守護者」遊戲。照片來源：研究者自攝。



圖 5-2：阿嬤坐著愉快地玩「阿嬤的菜市場」遊戲。照片來源：研究者自攝。

五、系統對家庭情緒交流的實際影響

兩週實測期間，我們觀察到幾個正向變化：第一，家人開始主動分享自己的情緒狀態，「今天系統說我焦慮耶」成為飯桌上的日常話題；第二，儀表板的 AI 提醒功能讓家人更關注彼此，例如爸爸看到「大姊最近挫折次數增加」的提醒後，主動找大姊聊天；第三，阿公阿嬤的遊戲互動數據讓我們更放心，即使不是每天回去探望，也能透過儀表板知道他們「今天有玩遊戲，活動狀況良好」。

六、研究限制與未來展望

本研究存在以下限制：（1）情緒類別僅四種，未涵蓋憤怒、悲傷等複雜情緒；（2）訓練資料僅 400 張，雖透過資料增強至 1,200 張，但相較通用模型仍屬小規模；（3）兩週實測期間較短，長期效果有待觀察；（4）目前僅適用於本家庭，推廣至其他家庭需重新拍照訓練。未來展望包括：增加情緒類別、引入時序分析、結合語音情緒辨識、開發自動化重訓練功能讓其他家庭能自行建立個人化模型，以及連結 SDG 3（健康與福祉）與 SDG 10（減少不平等），擴大系統的社會影響力。

陸、結論

本研究以「AI 傳情・家愛同行」為題，結合 AI 技術與 SEL 理論，為家庭打造個人化的情緒辨識與健康關懷系統。以下對照五項研究目的，逐一總結成果：

（一）成功運用 MobileNetV2 遷移學習技術，以僅 400 張家庭照片訓練出五位成員四種情緒的個人化分類模型，測試集準確率達 92.3%，F1 加權平均為 0.92，驗證了「小而精」個人化模型的可行性。

（二）成功運用 MediaPipe Pose 姿態估計技術，為阿公設計「森林守護者」遊戲（準確率 95.2%）、為阿嬤設計可坐姿操作的「阿嬤的菜市場」遊戲（準確率 93.6%），兼顧健康促進與長輩身體限制。

（三）成功設計 SEL 情緒互動模組，包含情緒日記、情緒趨勢分析、親子小遊戲等功能，兩週實測後家人情緒覺察自評分數平均提升 18.7%。

（四）成功建立 Google Firebase 雲端家庭儀表板，整合核心家庭與祖孫家庭的數據，實現跨地即時同步，使用者滿意度達 4.6 分（滿分 5 分）。

（五）透過混淆矩陣、Precision/Recall/F1 分數等系統性分析，完整驗證兩個 AI 模型的效能，並提出具體改進方向（如解決滿足與專注的混淆、增加時序分析等）。

本研究的核心價值不僅在於技術面的成就，更在於我們證明了：即使是國中生，也能透過 AI 工具（Vibe Coding）與有限資源，為自己的家庭打造一套有溫度的科技解決方案。正如《把你的心，放在軟綿綿的地方》一書提醒我們的，在科技與效率至上的時代，柔軟地對待自己與家人的情緒，才是最重要的事。我們希望「AI 傳情・家愛同行」不只是一份科展作品，更是我們家庭溝通的新起點——用 AI 傳遞情感，讓家人的愛，走在同一條路上。

柒、參考文獻

1. 小瀨古伸幸（2024）。把你的心，放在軟綿綿的地方：告別執著，活出不糾結的輕盈人生（連雪雅譯）。平安文化。
2. 吳怡萱、陳學志（2024）。邁向幸福的療癒之旅：BEST ME 教師社會情緒學習指南。國立臺灣師範大學出版中心。
3. 金昭姈（2024）。SEL 社會情緒學習，讓孩子成為人生贏家：做好自我管理、學習人際溝通、培養責任感（馮燕珠譯）。三采文化。
4. 教育部（2025 年 2 月 27 日）。社會情緒學習中長程計畫第一期五年計畫。教育部全球 訊網。
5. 曾逸文（2019）。基於深度學習之連續臉部情緒模式辨識方法應用於學習情緒〔碩士論文，中原大學〕。華藝線上圖書館。