表CM03

**三、研究計畫內容（以中文或英文撰寫）：**

1. 研究計畫之背景。請詳述本研究計畫所要探討或解決的問題、研究原創性、重要性、預期影響性及國內外有關本計畫之研究情況、重要參考文獻之評述等。如為連續性計畫應說明上年度研究進度。
2. 研究方法、進行步驟及執行進度。請分年列述：1.本計畫採用之研究方法與原因及其創新性。2.預計可能遭遇之困難及解決途徑。3.重要儀器之配合使用情形。4.如為須赴國外或大陸地區研究，請詳述其必要性以及預期效益等。
3. 預期完成之工作項目及成果。請分年列述：1.預期完成之工作項目。2.對於參與之工作人員，預期可獲之訓練。3.預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）。4.學術研究、國家發展及其他應用方面預期之貢獻。
4. 整合型研究計畫說明。如為整合型研究計畫請就以上各點分別說明與其他子計畫之相關性。

當前的人工智慧技術 (特別是所謂的「深度學習」) 在許多應用上（如：影像辨識、語音辨識、自然語言處理等) 有極其亮眼之表現，往往能提供高辨識率之精準預測結果。這樣具備高辨識率之預測模型有非常廣泛之應用價值，但在某些應用情境下 (如：醫療診斷、金融決策、詐欺評估、法律判決、安全決策、軍事武器等)，它卻有個致命的缺點：缺乏解釋性。雖然有些人認為只要預測模型具備高準確率就可以安心地直接採用，沒必要還需要去提供解釋原因。但有些領域的決策是需要有適當的解釋和因果關係的推論來支撐決策，才能讓人們可以對相關的決策覺得具合理性和感到安心。模型若缺乏解釋性就會導致部分使用者的信心不足，而具備可解釋性的模型則能大大提高使用者的信心度。

越來越多的國家都制定車輛需要加裝先進車輛輔助系統 (Advanced Driver Assistance Systems, ADAS) 來為駕駛人提供車輛的工作狀況與車外的行駛環境變化等資訊以提高行車安全。在此，我們就以先進車輛輔助系統 (Advanced Driver Assistance Systems, ADAS) 和無人駕駛車為例子，來說明高準確率有可能導致行車危險的事件發生。雖然深度學習的模型在有足夠多的交通號誌代表樣本的訓練之下，能得到令人十分滿意的高辨識率，但有一些極端的例子，透過某種數位攻擊（digital attack）的方式讓原本正常的交通號誌加入特定的雜訊後，雖然對人類肉眼來看仍是正常的交通號誌，但就會讓深度學習的預測模型做出錯誤的判斷，如圖一所示 [1]。原本應該是「停止」的標誌卻會被錯誤地分類為「最大速限100」！這樣子的錯誤判斷會導致先進車輛輔助系統和無人駕駛車產生極大之危險且有時是具致命性的錯誤判斷。而我們知道交通號誌經常會有汙損會某種程度的破壞 (如：彈孔等) ，如圖二所示，我們無法保證這種汙損不會造成類似之錯誤判斷。上述之範例要說明的是，高準確率之深度學習的預測模型有時會做出令人訝異的錯誤判斷，而這些判斷對人類的肉眼來說是極為不可思議的低級錯誤判斷。

美國國防部的DARPA單位在2016年就提出20億美元的新一代AI計畫，其中特別強調 Explainable Artificial Intelligence（XAI）的重要性，DARPA 的資料 [1]，我們可以用圖5來說明可解釋性可以用在那些應用領域和必要性何在。

圖5-6顯示了XAI 的需求和發展方向，指出 XAI可以替人們回答那些問題： (1) 為何AI決定這麼做？(2) 為何不那麼做？(3) 為何可以成功？(4) 為何會失敗？(5) 為何人們可以信任AI？(6) 人們該如何修正錯誤？這些問題，也正點出了可解釋的AI科技需要解決的挑戰。這個Explainable AI （XAI）計劃旨在創建一套機器學習技術，這些技術可以：(1) 產生更多可解釋的模型，同時保持高水準的學習表現（預測準確性）和(2) 使人類使用者能夠理解、適當地信任並有效地管理新興的人工智慧合作夥伴。

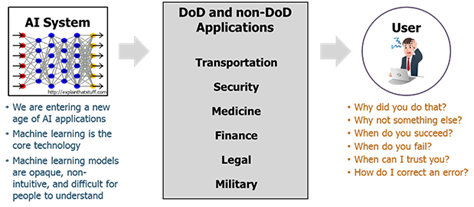


圖5、可解釋的AI (XAI) 的應用領域 [44]。

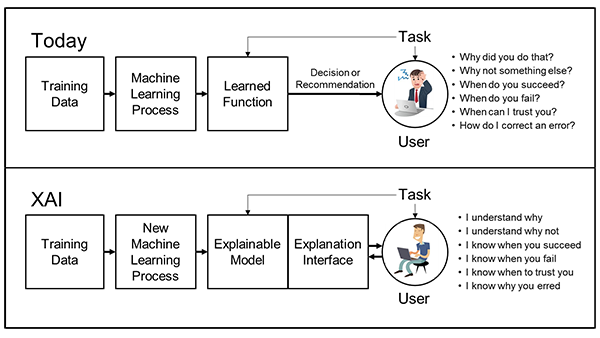


圖6、可解釋的AI 的整體概念 [44]。

既然 XAI 是如此地重要，我們該如何達成這個研究目標呢？我們首先回顧整個人工智慧的發展歷史，可以發現人工智慧原先有兩大不同的發展方向：(1) 符號人工智慧（Symbolic AI）和 (2) 類神經網路。早年符號人工智慧（Symbolic AI）的發展主軸之一就是專家系統，專家系統本身就是透過大量規則來模擬專家的知識和判斷。所以，專家系統的可解釋性很高。至於類神經網路則是透過層層複雜的矩陣運算後，才得到輸出結果，這樣的運算方式雖然可以從訓練資料迴歸出高準確率的輸入/輸出的關係，但對人類來說是缺乏可解釋性的。特別是近幾年的深度學習模型快速崛起之後，層數越來越深，導致不可解釋的問題變得越來越嚴重。

通常結構越簡單 (如：線性迴歸) 越具備可解釋性，但也由於結構簡單則會導致難以處理比較複雜的問題；反之，越複雜的模型則可以得到較高的準確度，但模型結構越複雜則越容易導致人類難以理解其特徵與結果的關係。所以，有時在高準確率和可解釋性之間必須做一定程度上的取捨。近幾年來，陸續有許多研究者投入 XAI 的研究 [45]-[49]，大致上，可以分為兩種不同的解決方案：(1) 發展同時具備可解釋性與高準確性的全新模型框架，模型本身就具備可解釋性來提升模型的透明度（Transparency）和 (2) 發展模型的事後可解釋方法，以「事後歸因」（Post-hoc Explanations）的方式來解讀模型如何得出網路輸出結果。第一種解決方案其實在1990 年代就已有一些可行的做法，如：the nested generalized example (NGE) algorithm [50], the fuzzy adaptive learning control network (FALCON) [51], the fuzzy min–max neural network classifier (FMMC) [52], adaptive-network-based fuzzy inference systems (ANFIS) [53], HyperRectangular Composite Neural Networks (HRCNNs) [54]。但上述的解決方案須搭配有效之特徵擷取方法才能提供高準確度之輸出結果，對於現今講求所謂的 End-to-End 的深度學習的解決方式，則仍有大幅改良之空間。至於第二種的解決方式則有 LIME [55]、SHAP [56]、Saliency Map [57] 等方法。LIME 的做法是在訓練出DNN 模型後，再針對一筆訓練資料 (如：一張圖) 中的局部資料進行干擾，以便觀察該局度資料是否會影響預測結果，而來判斷訓練資料中的那個部分對預測結果具有更關鍵的作用。而繪製顯著圖（Saliency Map）則是另一種廣為被使用的方法，它可以讓我們了解一張圖片中有哪些特定部位對模型結果預測的影響較大。

基於上述的文獻分析和探討，我們預計在這個子計畫中會開發出具可解讀性之深度學習的模型，並將其應用於自動化人臉表情辨識系統及其於學生專注力之偵測應用，以提高資料可信度和安全性。

參考文獻

[1] M. Turek, “Explainable Artificial Intelligence (XAI).” <https://www.darpa.mil/program/explainable-artificial-intelligence>.

[45] C. Molnar, [Interpretable Machine Learning: A Guide for Making Black Box Models Explainable](https://christophm.github.io/interpretable-ml-book/), lulu.com, 2021.

[46] T. Miller, “[Explanation in artificial intelligence: Insights from the social sciences](https://arxiv.org/abs/1706.07269),” *Artificial Intelligence,*2018.

[47] [F. K. Došilović](https://ieeexplore.ieee.org/author/37086411745), [M. Brčić](https://ieeexplore.ieee.org/author/37086117933), and [N. Hlupić](https://ieeexplore.ieee.org/author/37550775400), “Explainable artificial intelligence: a survey,” 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2018.

[48] S. R. Islam, W. Eberle, S. K. Ghafoor, and, M. Ahmed, “Explainable Artificial Intelligence Approaches: A Survey,” [arXiv:2101.09429](https://arxiv.org/abs/2101.09429), 2021.

[49] E. Tjoa, and C. Guan, “A Survey on Explainable Artificial Intelligence (XAI): towards Medical XAI,” JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, pp. 1-22, 2015.

[50] Salzberg SL. Learning with nested generalized exemplars. Academic: Kluwer; 1990.

[51] Lin C-T, Lee CSG. Neural-network-based fuzzy logic control and decision system. IEEE Trans Comput 1991;40(12):1320–36.

[52] Simpson PK. Fuzzy min-max neural networks-Part 1: Classification. IEEE Trans. on Neural Network 3, pp. 776–86, 1992.

[53] Roger Jang JS. ANFIS: adaptive-network-based fuzzy inference systems. IEEE Trans. on Syst Man Cybern;23(3), pp. 665–85, 1993.

[54] Su MC. Use of neural networks as medical diagnosis expert systems. Comput. Biol Med, 24(6), pp. 419–29, 1994.

[55] M. T. Ribeiro, S. Singh, and C. Guestrin,“Why Should I Trust You?”: Explaining the Predictions of Any Classifier,” [Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining](https://dl.acm.org/doi/proceedings/10.1145/2939672), pp. 1135–1144[,](https://doi.org/10.1145/2939672.2939778) 2016.

[56] S. M. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,”[Advances in Neural Information Processing Systems 30 (NIPS 2017)](https://proceedings.neurips.cc/paper/2017), pp. 4765--4774, 2017.

[57] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps,” [arXiv:1312.6034](https://arxiv.org/abs/1312.6034), 2013.

(二) 研究方法、進行步驟及執行進度。

**第一年: 研發基於多層自我組織映射圖之可解讀性之深度學習模型**

本計畫的第一年預計要開發基於多層自我組織映射圖 [58] 之可解讀性之深度學習模型，所使用的網路架構是結合非監督式和監督式的類神經網路 (如圖7所示)。整體構想是基於人類的視覺神經系統的運作機制，當外界的物件透過視網膜的成像後，會將視覺刺激不斷地傳到下一層皮質區來處理，最後傳達大腦時，由大腦再做綜合判斷，並才指令來因應外界的刺激。

基本架構將採用自我組織映射，將特徵相似度高的放在一起；再透過適應共振理論，對於整個架構的輸出結果做儲存；最後，利用學習向量量化網路，對適應共振理論所儲存的特徵映射響應圖進行微調，而非使用一般常見的梯度修正。另外，我們透過可視覺化，將每層的特徵轉換成具有解讀性特徵影像，使人可以有效的解讀特徵所賦予的意義。

|  |
| --- |
|  |
| 圖7、多層自我組織映射之架構圖。 |

網路訓練：訓練架構又可細分為三部分。第一部份：對影像作切割；第二部份：將切割影像經過層層的處裡；第三部份：將最後的結果，記錄下來，以便下次辨識。一開始會對輸入影像做切割，形成最小單位的影像大小稱為基礎影像，基礎影像是人眼的視網膜能夠接收的最小單位。再將這些基礎影像送入實驗架構中。

在這步驟相當於每一層的皮質層正在學習一些新物體，將其特徵記錄下來，並且產生出下一層能夠辨識的資訊。到最後一層的RM生成時，會將生成的RM輸入至適應共振理論（Adaptive Resonance Theory，簡稱ART）中，讓ART記錄下最後的RM，並且做加標的動作。這步驟相當於大腦正在學習新的物體，因為每一種物體都有屬於他們的特徵，而大腦的功用就是將他們每一種特徵給紀錄下來，方便下一次做判斷時使用。

在第一年的計畫規畫中，我們會先用現有的臉部表情資料集來驗證此可視覺化深度學習模型的有效性。

**預計可能之困難與解決途徑：**

**困難一：需要幾層的特徵映射圖？**

解決途徑：目前需針對不同的問題試煉不同大小的特徵映射圖。

**困難二：每一層特徵映射圖的大小如何設定？**

解決途徑：層數越大，所需要的尺寸也會隨著變大。

**第二年:**

**(1) 持續改良第一年所發展之基於多層自我組織映射圖之可視覺化深度學習模型**

**(2) 開發高齡人士的情緒辨識系統(如圖8所示)。**

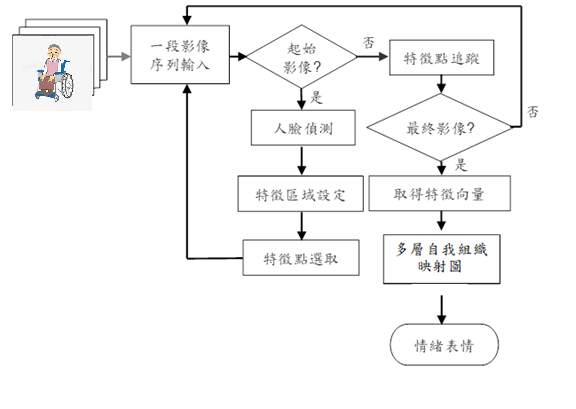


圖8、高齡人士的情緒辨識系統辨識系統。

台灣在 2017 年之後，幼年人口就開始與高齡人口比例形成交叉之走勢，高齡者比例一路快速爬升，導致社會勞動力人口的萎縮。在 2061 年後高齡人口比例會高達全台 40% 之上。相關資料顯示在 2065 年後，整體社會勞動人口比例將跌破 50%，這數字傳達出：(1) 社會需要被照護的高齡人口急速上升和 (2) 加上少子化的因素，更會導致可以投入照顧高齡人士的青年勞動人力逐年減少。因此，未來高齡照護人力會嚴重不足。上述現象是台灣高齡照護產業共同遭遇到的社會難解議題。因此，本子計畫第二年主要是想針對高齡人士的情緒辨識系統的開發和應用，希望此情緒辨識系統的開發和應用可以結合其他子計畫的研究成果，一同合作來減輕高齡照護人力不足的問題。

情緒辨識是未來提供老人陪伴、娛樂、撫慰之情感智能機器人、智能物聯網、AI語音助理的第一步。當系統偵測到老人情緒低落時，情感智能機器人或AI語音助理可以適時地給予即時的情感撫慰的服務，進而改善老人的低落或孤寂的情緒，達到提升老人心理健康與生活品質之目的。

這一年基本上是會與採用子計畫一所建立之自然情境下所蒐集之表情資料庫來精進前一年度之表情辨識系統，來解決高齡人士的情緒辨識的問題。高齡人士的活動樣態包含：正常走動、坐在椅子或沙發上、坐輪椅、臥床休息等，就會因為攝影機所拍攝的角度的不同而導致姿態差異（如：俯角、仰角、側臉等）、照明度、距離遠近、背景複雜度和遮擋干擾因素等。此外，也可能會因為年齡、性別、種族背景和情緒表達方式等因素，導致辨視率下降。

因此，今年有許多待克服的技術瓶頸需要克服，會參酌或改良現有有效之臉部校正的工具 (如：Active Appearance Model (AAM) [33]、the mixtures of trees (MoT) [34]、the supervised descent method (SDM) [35]、Cascaded CNN [36]、Multi-task CNN (MTCNN) [37] 等)；照明和對比度也會參考採用discrete cosine transform (DCT)-based normalization [38] 等方法來降低照明和對比度造成的干擾。此外，姿態正規化 (pose normalization) 也是前處理需要解決的問題，預計會參考採用以下方法來進行 (如： Hassner et al. [59]、Sagonas et al. [60]、FF-GAN [61], TP-GAN [62]等)。

此外，也會參考3D CNN 的做法 [63]，增加動態串流的資訊，來提高辨識率。

**預計可能之困難與解決途徑：**

**困難一：**姿態差異（如：俯角、仰角、側臉等）和背景複雜度和遮擋干擾因素導致辨識率下降。

解決途徑：提出更有效之前處理的方法。此外，也會考慮採用其他深度學習模型 （如：RNN、LSTM、Transformer等）來增進辨識率。

**困難二：數據資料不足的問題。**

解決途徑：考慮採用資料增量的方式（如：GAN [64]等）和來增加數據量。

**第三年: 完成學生專注度分析系統**

根據微軟2015年對2000位參與者所做的調查，受到3C時代影響，現代人專注於一件事情的時間越來越短，從2000年的12秒，降至2015年的8秒。近年來，老師和家長也發現學生專注於一件事情的能力下降，會不斷分心做其他事情，像是在上課時每幾分鐘會瞄一下手機螢幕或是滑手機，甚至目光會被手機的聲光刺激吸引住，然而學生的學習成效往往決定在專注於課堂的時間上，若是分心於使用電子產品或是其他因素上，則無法學習到完整的上課內容，長期下來會造成學習進度的落後。除了分心於非課堂事物以外，精神不濟也是造成專注力不足的原因之一，人在疲倦狀態時無法有效的接收眼前的資訊，學生在下課之餘會需要花更多的時間來補齊漏掉的課堂進度。因此老師專心於上課教學之餘，可能需要時時關心學生的學習狀況，但是基於課堂進度的安排，往往只能倚靠小考或是問答的方式來確認學生的學習效果。

因此，本研究的目的為希望能透過影像分析的方式來偵測學生的上課專注度，幫助老師了解學生們的學習狀況。希望在課堂環境中架設彩色攝影機，建構一套學生專注度分析系統。使用者將攝影機設置於高處之後，透過人臉偵測和特徵擷取演算法來找到人臉特徵，藉此來判斷學生的臉部朝向角和疲勞程度；此外，也透過姿態評估系統找到骨架特徵，藉此判斷學生的學習姿態。之後，綜合以上的結果來分析每位學生的專心程度。

根據以上，本研究提出學生專注度分析的應用，利用深度學習判定臉部資訊和動作辨識以評估學生的專注度，老師可以透過這些資訊調整上課的進度或是教學方式，以拉回學生的注意力，提升學生學習品質。

學生專注度分析的

系統 (如圖 9 所示) 預計達到下列之功能：

1. 偵測人臉和找到人臉特徵，判斷人臉資訊和朝向角。
2. 透過姿態評估系統找到骨架特徵，利用類神經網路和物品辨識判斷姿態。

根據以上兩點的結果分析學生的專注度。

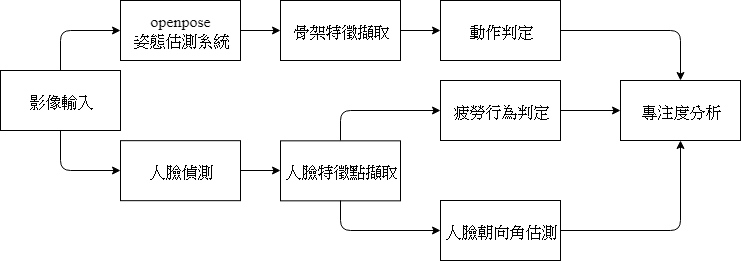


圖9、學生專注度分析系統。

根據以上，本研究提出學生專注度分析的應用，利用深度學習判定臉部資訊和動作辨識以評估學生的專注度，老師可以透過這些資訊調整上課的進度或是教學方式，以拉回學生的注意力，提升學生學習品質。

預計會使用Dlib函式庫中已經訓練好的CNN模型 [65] 進行人臉偵測，以及Bulat及Tzimiropoulos [66] 提出的Face Alignment Network (FAN) 網路架構和其開源程式碼 [67] 進行人臉特徵擷取，FAN結合4個Hourglass (HG) 網路 [68]，主要架構如圖10所示。人臉特徵擷取主要會取得68個人臉特徵點的座標 [69]，如圖11所示。

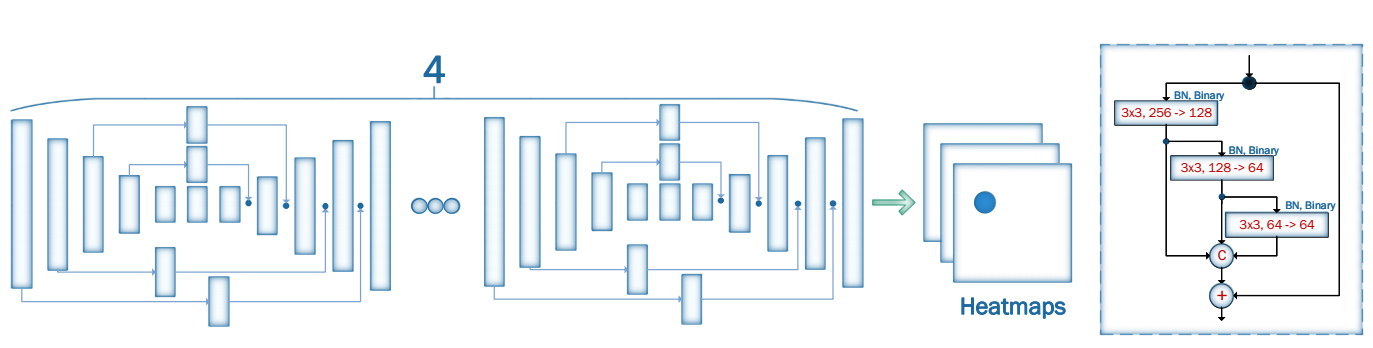


圖10、FAN網路架構圖（取自 [68]）。

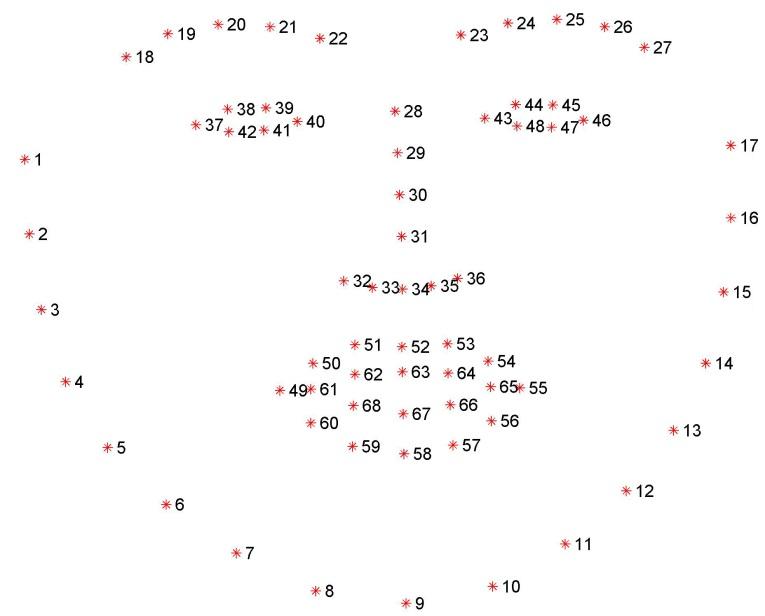
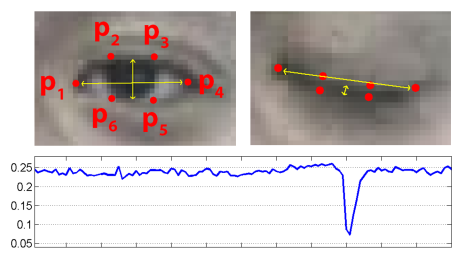
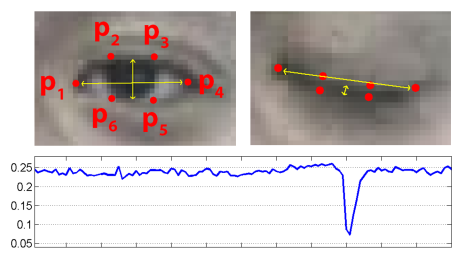


圖11、人臉的68個特徵點示意圖（取自 [68]）。

然後，會參考Soukupová及Čech [70] 提出的演算法來判定閉眼行為的發生，利用眼睛的6個特徵點座標計算出eye aspect ratio (EAR)，如圖12 (a) 所示，EAR的數值在閉眼時有明顯下降的趨勢，如圖11 (b) 所示。圖13則為哈欠行為判定的流程，使用連續時間的計數器，若影像的嘴巴特徵超過距離閥值則計數器加一，沒超過則讀取下一個幀像，並且計數器歸零，若計數器的數值超過時間閥值則判定有哈欠行為發生。接著， 我們會使用人臉的5個特徵點（左眼角、右眼角、鼻尖、左嘴角、右嘴角）來計算特徵，分別為特徵點間的距離變化和角度變化，再透過類神經網路來判斷臉的朝向角度。綜合上述的處理流程和功能，我們就可以完成： (1) 疲勞度判斷和 (2) 人臉朝向角估測。



(a) (b)

圖12、閉眼時，連續幀像EAR的變化量（取自 [69]）

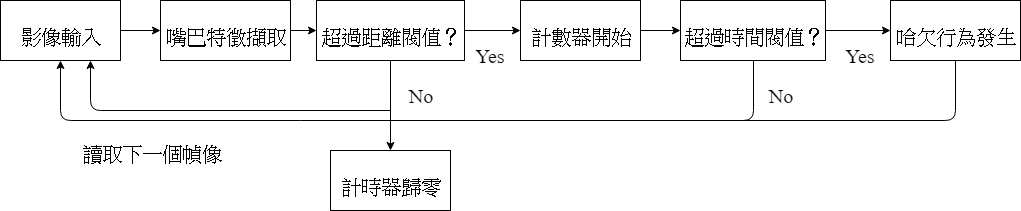


圖13、哈欠行為判定流程。

專注力也可以從一個人的行為來判定，像是手撐頭的動作可能表示對上課內容感到厭煩和枯燥，因此我們利用openpose的姿態估測系統取得骨架資訊，再從骨架資訊中取得骨架特徵，最後經由類神經網路訓練以判定學生的行為 (如：教室情境下常見的八個動作，分別為寫筆記、使用筆電、滑手機、舉單手、歪頭、單手撐頭、趴睡和接電話等)。若是使用原始影像當作輸入用來訓練類神經網路，可能會因為背景干擾或其他物件遮蔽而誤判，因此本研究預計取得影像的骨架資料之後，再將經正規化後的骨架資料繪製出骨架影像當作輸入，用以提升類神經網路的辨識能力。由於本研究的學習情境下都是以坐著為主，因此移除髖關節以下骨架的資訊，避免影響辨識效果。然後，會搭配第一年所開發之模型或卷積類神經網路來加以作棟動作辨識 (如圖 14 所示)。

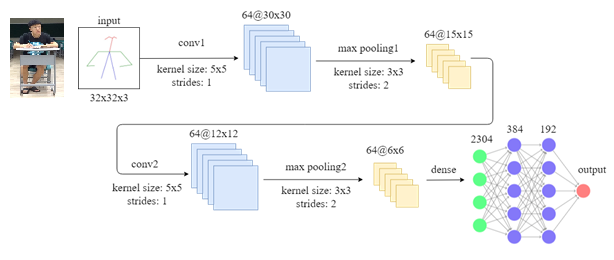


圖14、動作辨識架構。

**預計可能之困難與解決途徑：**

**困難一：**寫筆記、使用筆電、滑手機有彼此辨錯的情形，這些動作的骨架影像是比較類似的，如何增加辨識率？

**解決途徑：**需要加入物品辨識 (如：書本、筆電和手機等)的方式來加強辨識效果。

**困難二：如何在實際讀書情境下精準專注力？**

**解決途徑：**可以考慮結合時間域的特徵，以增加辨識效果。

**參考文獻：**

[59] T. Kohonen, *Self-Organization and Associative Memory*, 3*rd* ed. New York, Berlin: Springer-Verlag, 1989.

[59] T. Hassner, S. Harel, E. Paz, and R. Enbar, “Effective face frontalization in unconstrained images,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4295–4304, 2015.

[60] C. Sagonas, Y. Panagakis, S. Zafeiriou, and M. Pantic, “Robust statistical face frontalization,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 3871–3879, 2015.

[61] X. Yin, X. Yu, K. Sohn, X. Liu, and M. Chandraker, “Towards largepose face frontalization in the wild,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3990–3999, 2017.

[62] R. Huang, S. Zhang, T. Li, and R. He, “Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2439–2448, 2017.

[63] S. Ji, W. Xu, M. Yang, and K. Yu, “3d convolutional neural networks for human action recognition,” IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 221–231, 2013.

[64] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in Advances in neural information processing systems, pp. 2672–2680, 2014.

[65] Dlib CNN face detector example. Available:http://dlib.net/cnn\_face\_detector.py.html. [Accessed: 12-Jun-2018].

[66] A. Bulat and G. Tzimiropoulos, “How far are we from solving the 2D & 3D Face Alignment problem? (and a dataset of 230,000 3D facial landmarks),” *International Conference on Computer Vision*, pp. 1021-1030, 2017.

[67] 2D and 3D FAN GitHub code. [Online]. Available: https://github.com/1adrianb/2D-and-3D-face-alignment. [Accessed: 12-Jun-2018].

[68]A. Newell, K. Yang, and J. Deng, “Stacked hourglass networks for human pose estimation,” *European Conference on Computer Vision*, pp. 483-499, 2016.

[69] 68 facial landmarks. [Online]. Available:https://ibug.doc.ic.ac.uk/resources/300-W/. [Accessed: 12-Jun-2018].

[70] T. Soukupová and J. Čech, “Real-Time Eye Blink Detection using Facial Landmarks,” *21st Computer Vision Winter Workshop*, pp. 1-8, 2016.

(三) 預期完成之工作項目及成果。請分年列述：1.預期完成之工作項目。2.對於參與之工作人員，預期可獲之訓練。3.預期完成之研究成果（如實務應用績效、期刊論文、研討會論文、專書、技術報告、專利或技術移轉等質與量之預期成果）。4.學術研究、國家發展及其他應用方面預期之貢獻。

**第一年:**

預期完成之工作項目：

1. 完成基於多層自我組織映射圖之可視覺化深度學習模型的學習演算法。
2. 完成表情辨識雛型系統。
3. 論文撰寫。

參與之工作人員，預期可獲之訓練：

1. 學會基於多層自我組織映射圖之可視覺化深度學習模型的設計。
2. 可提高計劃參加者之程式語言撰寫能力。
3. 學會深度學習和類神經網路的理論與技術。
4. 懂得跨領域之整合及溝通協調能力。

**第二年:**

預期完成之工作項目：

1. 完成基於多層自我組織映射圖之可視覺化深度學習模型於自動化人臉表情辨識系統。
2. 完成高齡人士之表情辨識系統。
3. 論文撰寫。

參與之工作人員，預期可獲之訓練：

1. 可提高計劃參加者之程式語言撰寫能力。
2. 學會表情辨識系統各種前處理的理論與技術。
3. 學會深度學習和類神經網路的理論與技術。
4. 懂得跨領域之整合及溝通協調能力。

**第三年:**

預期完成之工作項目：

1. 完成疲勞度和臉部朝向角估測模組功能。
2. 完成動作辨識模組功能。
3. 完成學生專注度分析系統。
4. 論文撰寫。

參與之工作人員，預期可獲之訓練：

1. 學會學生專注度分析系統的設計。

1. 學會深度學習和動作辨識的理論與技術。
2. 懂得跨領域之整合及溝通協調能力。