（一）摘要

為了因應因科技發展，越來越多課程改以線上學習的方式來進行，本次的計畫主要為透過影片或視訊鏡頭來即時判斷其專注度。首先針對影像或視訊鏡頭擷取人臉部分，手動針對專注度評分後並利用深度學習模型加以訓練，再以訓練成果為基準交由系統自動偵測人臉並分析其專注度，同時也能讓使用者自行輸入產生資料及產生客製化的評分標準。

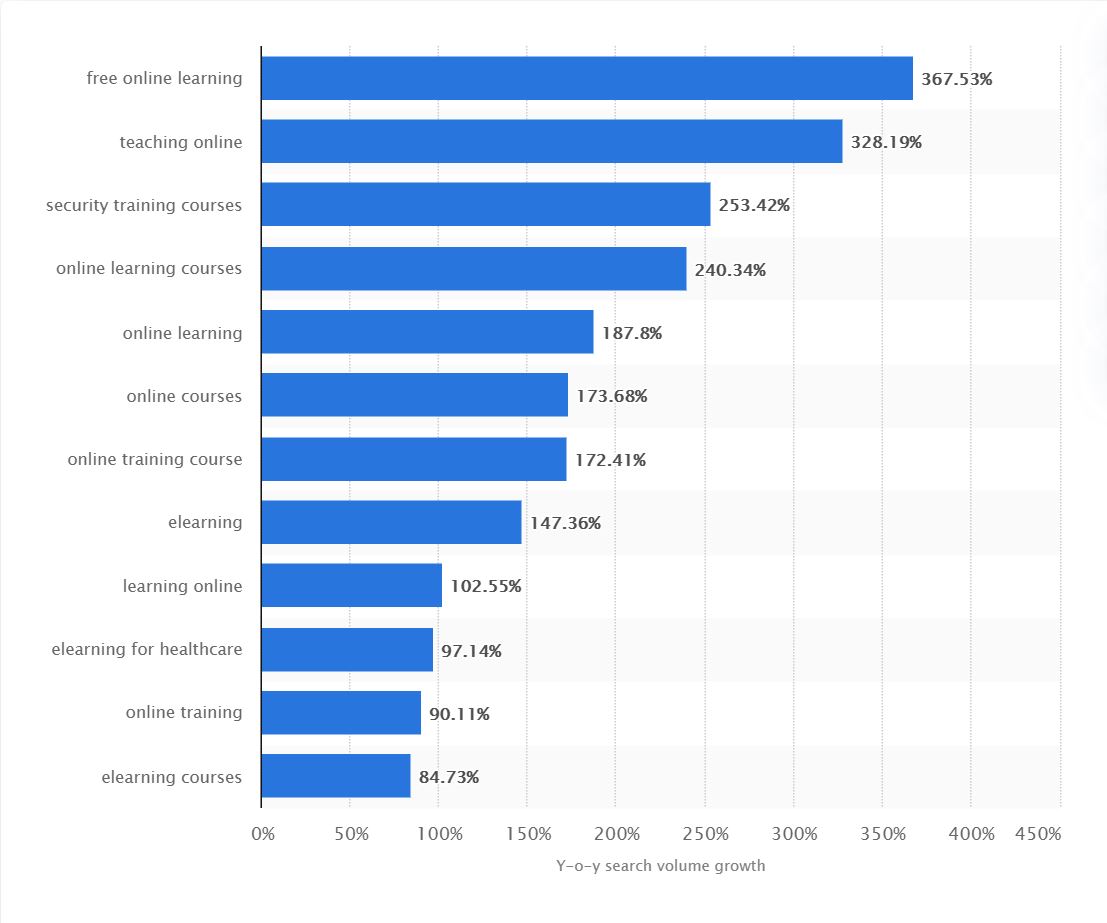
關鍵字：專注度分析（attention analysis）、OpenCV、人臉辨識、深度學習

（二）研究動機與研究問題

**研究動機：**

近年來網路通訊的發展使得授課的方式越來越多元，遠距學習(或稱數位學習)的上課方式也逐漸被廣泛的使用。過往數位學習的方式大多像是教師預錄教學影片，而學生則利用時間自行點閱學習。2020年全球因COVID-19的肆虐造成許多採用傳統教學的方式被迫改以線上教學來進行，根據Google搜尋引擎的統計，在英國與數位學習相關詞彙的搜索量皆有大幅度的成長[1]，這顯示了線上學習將是未來的趨勢所在。

而傳統學校的課程除了作業、考試的成績外，教師們往往透過觀察學生的出缺席以及上課的專注度來給予所謂的平常分數，若以線上即時教學的方式授課，教師較難以針對學生課堂上的專注度來督促學生，而影像處理加上深度學習便是一個能夠客觀評分專注度且可調整的評比方式。



圖(一)：2020英國地區Google數位學習相關詞彙的搜索成長率[1]

雖台灣在國人的配合下疫情相全球已是相當的安全，各階段的教育機構也都能正常的讓學生到校上課，但也一度延後開學的時間，甚至有幾間學校都已做好遠距教學的準備。雖然國內疫情在那之後漸趨和緩，但目前來看這場與病毒的戰爭勢必會陷入持久戰，若能對於遠距教學的專注度評分事先研究，在最糟糕的情況到來時即可確保教學評分的公正性以及維護學生的權益。

**研究問題：**

本計畫目標是針對視訊鏡頭或影片存檔來做到即時偵測人臉並評分專注度分數，也同時讓使用者能夠自行評分並輸入模型訓練。首先調整OpenCV的分類器，調整出可以準確框出人臉的參數，並決定是否重新評分並產生新的訓練集；後端訓練的部分則是首先先輸入經過手動評分專注度的資料，依此資料為基礎產生出系統初步評判的可靠標準，若使用者對於評判標準有所疑慮可再重新輸入樣本已訓練出更精準的資料。詳細細項如下：

**1.擷取人臉**

1.1 正確的偵測到人臉並截圖，並捨去其餘圖像或增強偵測的參數設定

1.2 偵測人臉後提供客觀的專注度評分

1.3 確認是否將截圖的圖片加入新的訓練集

**2.後端訓練**

2.1 依客觀的評判標準手動評分截圖

2.2 依2.1之資料進行訓練，得到可用度高之樣本

2.3 系統即時截圖並即時評分專注度

2.4 可讓系統加入新訓練集持續訓練，增加可靠度

**3.介面整合**

3.1 監控桌面，並即時將截圖呈現於視窗中預覽

3.2 能對截圖在使用者介面下決定是否手動評分

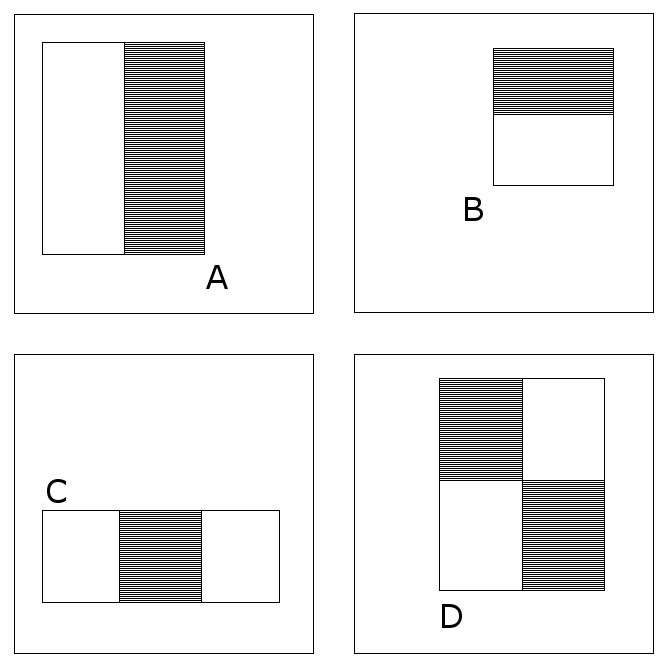
3.3 能在使用者介面下決定是否將該圖加入新訓練集

（三）文獻回顧與整理

近年來，機器學習研究蓬勃發展，被廣泛運用在資訊產業的各個範疇，伴隨著越來越多學者研究這項議題，電腦視覺也成為機器學習能應用的領域之一，倒不如說，正是機器學習的發展帶領著電腦視覺領域邁向了更大範圍的應用。

**前端偵測：**

首先是初步的人臉偵測部分，P.Viola及M.Jones於2001年提出《Rapid Object Detection using a Boosted Cascade of Simple Features》[2]和《Robust Real-Time Face Detection》[3]，是第一種可以即時處理的物體檢測的方法。這種方法可被訓練來尋找多樣物體，而主要應用還是在解決人臉偵測方面。檢測框架使用的[特徵](https://zh.wikipedia.org/w/index.php?title=%E7%89%B9%E5%BE%81_(%E8%AE%A1%E7%AE%97%E6%9C%BA%E8%A7%86%E8%A7%89)&action=edit&redlink=1)涉及到圖像上矩形區域的像素和，就是Haar cascade，這些特徵以前多用於基於圖片的物體檢測。然而，由於Viola以及Jones使用的特徵包含不止一個矩形區域，就顯得更為複雜。圖(二)的圖片是兩者使用的四種不同的特徵[4]。每一個特徵的值就是白色矩形的像素值加總與深色矩形像素值加總的差值。相比其他複雜的steerable filters，這種矩形特徵是十分原始的。雖然這種矩形特徵對水平和垂直方向比較敏感，它們的回饋是比較粗略的。但也因為如此，矩形特徵的計算可以在常數時間內完成，也就使得它們具有速度上的優勢。



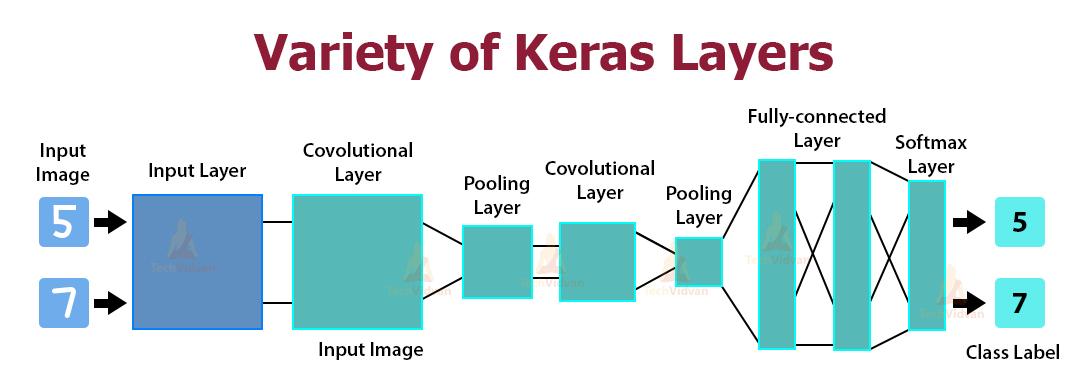
圖(二)Viola及Jones使用的特徵[4]

目標檢測框架使用了一個稱為[AdaBoost](https://zh.wikipedia.org/wiki/AdaBoost)[5]的機器學習演算法來選擇特徵並訓練分類器。AdaBoost，是英文"Adaptive Boosting"的縮寫，是一種[機器學習](https://zh.wikipedia.org/wiki/%E6%9C%BA%E5%99%A8%E5%AD%A6%E4%B9%A0)方法，由Yoav Freund和Robert Schapire提出。AdaBoost所謂的自適應在於：前一個分類器分錯的樣本會被用來訓練下一個分類器。AdaBoost方法是一種疊代算法，在每一輪中加入一個新的弱分類器，直到達到某個預設的夠小的錯誤率。每一個訓練樣本都被賦予一個權重，表示它被某個分類器選入訓練集的機率。如果某個樣本點已經被準確地分類，那麼在產生下一個訓練集中，它被選中的機率就被降低；反之，如果某個樣本點沒有被準確地分類，那麼它的權重就得到提高。通過這樣的方式，AdaBoost方法能「聚焦於」那些較難分（更多資訊）的樣本上。同時，AdaBoost方法對於雜訊數據和異常數據很敏感。

在AdaBoost演算法的基礎上，使用Haar-like小波特徵方法進行人臉檢測，它們兩者不是最早使用提出小波特徵的，但是他們設計了針對人臉檢測更有效的特徵，並對AdaBoost訓練出的強分類器進行相關聯。也因此當時提出的這個演算法被稱為Viola-Jones檢測器(Viola Jones object detection)[6]。又過了一段時間，Rainer Lienhart和Jochen Maydt將這個檢測器進行了擴充套件[7]，最終形成了OpenCV現在的Haar分類器。而本計畫預計利用OpenCV底下的Haar分類器(Haar Classfiers)來做初步的人臉偵測。Haar分類器可以幫助我們辨識人臉、眼睛、五官、交通號誌，甚至是貓或狗，等等......依據Haar特徵檔而定。

**後端訓練：**

而訓練模型則是採用Keras[8]，Keras是由純python編寫的基於theano/tensorflow的深度學習框架，能夠在TensorFlow上執行。Keras能夠快速實現[深度神經網路](https://zh.wikipedia.org/wiki/%E6%B7%B1%E5%BA%A6%E5%AD%A6%E4%B9%A0)，專注於使用者友善、模組化和可延伸性，是ONEIROS（開放式神經電子智慧機器人作業系統）專案研究工作的部分產物。

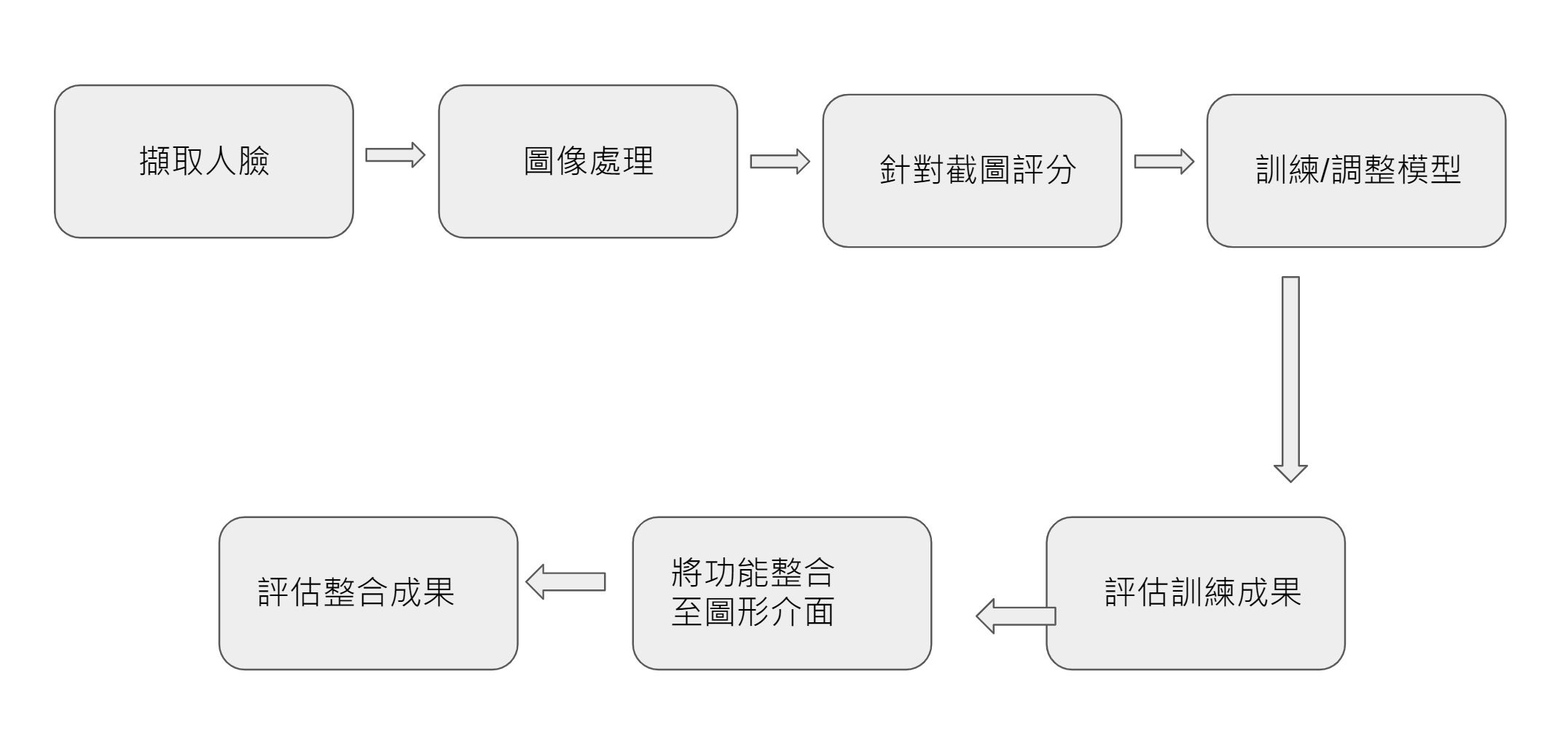


圖(二)：Keras架構圖[8]

Keras以及PyTorch是目前較熱門的深度學習的開源框架，兩者主要的差別在於頂層以及底層差別[9]，還有易用性和靈活性。Keras將常用的深度學習曾以及操作包裝成高階的API，使使用者能快速上手並不用全盤聊解深度學系背後的複雜原理，其程式碼相較起來也較少較精簡。而PyTorch則是提供一個較為底層的操作環境，使用者可以自由地更改底層的應用，在上手之後可以開發更複雜的結構，當然代價及是冗長複雜的上手過程。基於上述理由，本計畫目的是以專注度為基礎來訓練模型，因此選擇了較高層且容易上手的Keras。

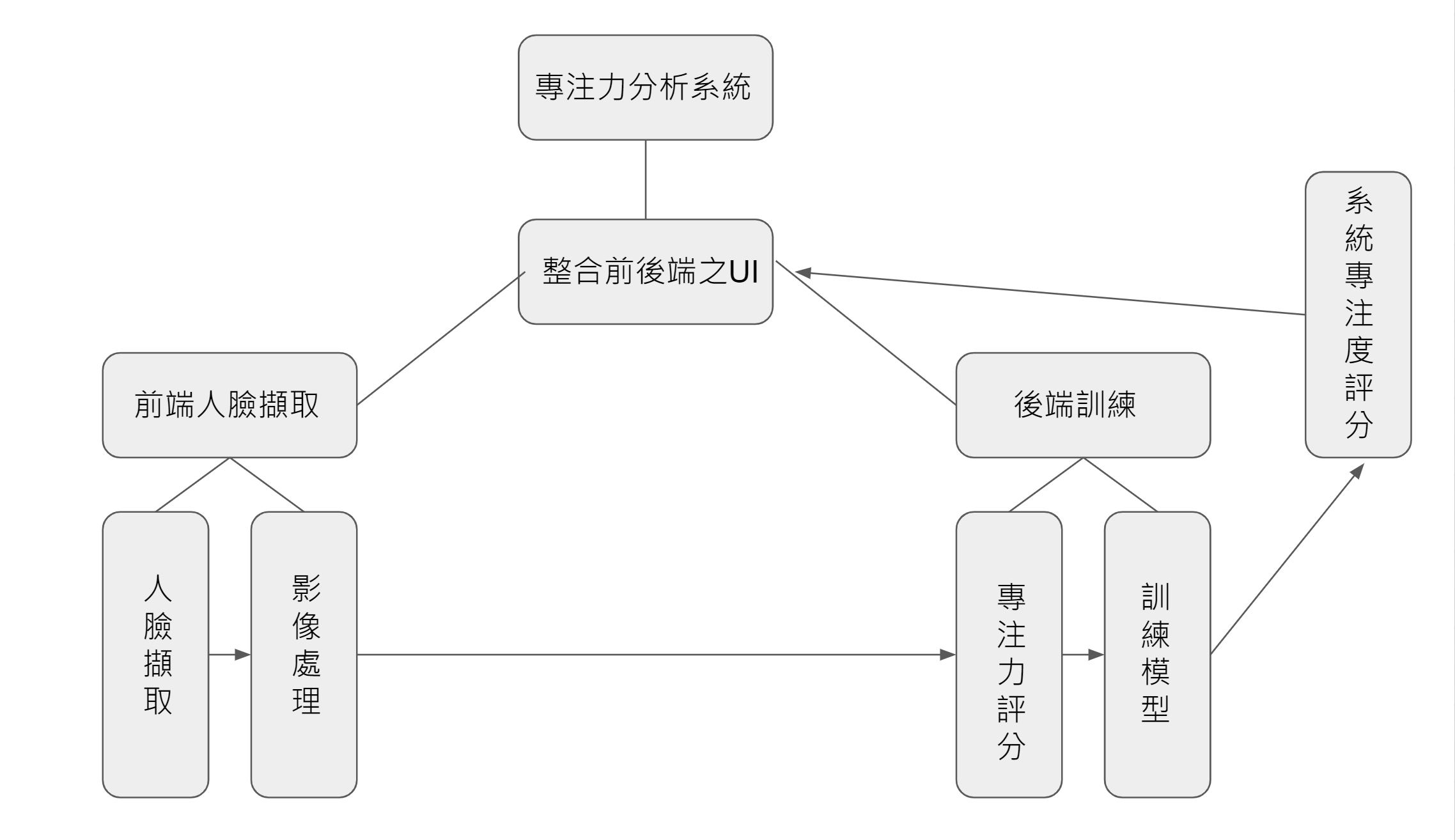
「注意力」是認知心理學中非常受重視的主題，注意力與學習的關係是非常密切的，沒有注意就沒有辨識、學習與記憶[10]。然而近年來教學方法日新月異，很多的學習已不是老師在課堂上實際授課，而是透過各種教學媒體，由學生主動去學習。或是有些老師會運用合作學習的模式，進行教學，以期讓所有的學生在這樣的學習模式下，發揮最大的學習表現。但是，這樣的學習模式，讓老師無法掌握學生的學習狀況及參與程度。因此專注力的量測就變得格外重要。從過去的文獻中，專注力的量測可分為生理內部、外在行為及利用輔助工具等方式[11]。外在行為的表徵，則是測量受測者在專注時，對聽覺的反射、臉部表情、眼睛運動的狀態，以判斷其專注程度[12]。也有方法是利用偵測眼睛閉合來判別專心程度[13]。計畫的一部份是將收集到的數據評分並加以訓練，預計的評分標準在後面部分會再探討。

（四）研究方法與步驟

　　預計開發流程：

圖(三)：預計開發流程

　　預計系統架構：



圖(四)：專注力分析系統架構

**1.擷取人臉：**

1.1 了解Python下OpenCV各函式使用方法。

1.2 了解OpenCV下各分類器擷取成果。

1.3 選擇適當分類器擷取臉部部分。

**2.圖像處理：**

2.1 將截取下的圖片轉換成灰階圖，以使後續計算量大幅減少。

2.2 將灰階圖統一尺寸，使後續訓練能減少時間以及增加正確度。

**3.針對截圖評分：**

3.1 將照片前處理結束後先行手動評分

　　(I) 觀察雙眼是否正對鏡頭(無向下看、向旁看)。

　　(II) 觀察其眼睛開闔。

　　(III) 觀察其面部表情動作，如打哈欠等等......

**4.訓練、調整模型：**

4.1 將完成評分的截圖輸入模型訓練。

**5.評估訓練成果：**

　　5.1 觀察評分狀況，若評分狀況沒問題則至步驟６。

　　5.2 觀察評分狀況，若評分狀況仍有問題則：

　(I) 重新評分再輸入訓練。

　(II) 加入更多的訓練樣本。

**6.將功能整合至圖形介面：**

6.1 了解、練習Tkinter使用。

　　6.2 製作出人性化的圖形介面編排。

　　6.3 使圖形介面能夠即時預覽評分成果，並決定是否重新評分以加入訓練集。

以下為計劃現階段草擬之圖形介面：



圖(五)：圖形介面預估樣式

（五）預期結果：

本計畫期望能夠：

1. 針對視訊影像及影片成功偵測出正確的區域(人臉部分)以及實地在不同光源、 不同地點錄製視訊影像來增加訓練集的豐富度。

2. 找出評斷各指標的適當權重以得出客觀專注度(如打哈欠扣1分、睡著扣5分)。

3. 能夠讓使用者自行輸入樣本以得到客觀之分數，輸入方式不限自行輸入一系列 臉部截圖或者重新評分的部分。

4.讓使用者能夠成功評分課堂參與者之專注成績。

（六）文獻參考

1. Biggest year-on-year growth of E-learning related Google search terms during the COVID-19 outbreak in the United Kingdom (UK) in 2020.<https://www.statista.com/statistics/1126825/e-learning-search-term-growth-during-covid-19-uk/>
2. Paul Viola, Michael Jones. "Rapid Object Detection using a Boosted Cascade of Simple Features", [Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001](https://ieeexplore.ieee.org/xpl/conhome/7768/proceeding)
3. Paul Viola, Michael Jones. "Robust Real-Time Face Detection", International Journal of computer,2004
4. Viola–Jones object detection framework. <https://zh.wikipedia.org/wiki/%E7%BB%B4%E5%A5%A5%E6%8B%89-%E7%90%BC%E6%96%AF%E7%9B%AE%E6%A0%87%E6%A3%80%E6%B5%8B%E6%A1%86%E6%9E%B6>
5. Robert E. Schapire, Yoav Freund "Using AdaBoost to Minimize Training Error". MIT Press, 2012
6. [Dmytro Peleshko](https://ieeexplore.ieee.org/author/37321352000), [Kateryna Soroka](https://ieeexplore.ieee.org/author/37070370100). "Research of usage of Haar-like features and AdaBoost algorithm in Viola-Jones method of object detection", [International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)](https://ieeexplore.ieee.org/xpl/conhome/6526842/proceeding),2013
7. Rainer Lienhart, Jochen Maydt. "An Extended Set of Haar-like Features for Rapid Object Detection", Image Processing. Proceedings. 2002 International Conference on volume 1, 2002
8. varieties layers of Keras. https://techvidvan.com/tutorials/keras-layers

# Keras vs TensorFlow vs PyTorch : Comparison of the Deep Learning Frameworks. https://www.edureka.co/blog/keras-vs-tensorflow-vs-pytorch

1. 鄭昭明，"普通心理學 "，2006
2. Lin-Huang, Cheng, 2010
3. Frank A. Drews, Monisha Pasupathi, and David L. Strayer"Passenger and Cell Phone Conversations in Simulated Driving". 2008
4. 黃俊源，"學習專注力即時偵測回饋系統",國立成功大學工程科學研究所碩士論文. July 2010

（七）需要指導教授指導內容

進行此專案研究的過程中，我們將針對以下需求，不定時與指導教授進行討論以增進相關領域的基礎、實作部分。

1.影像處理相關理論、基礎

2.培養深度學習的基礎理論，如其演算法，關鍵字等等

3.系統設計與相關實作

4.研究專案的方向，可行性評估、與進度報告。