

期末專題

上課請睜眼

第 4 組

機械三 B06502028 莊立楷

機械三 B06502048 陳亭瑋

機械三 B06502155 陳冠綸

心理所博二 D07222008 蔡秉叡

壹、主題

我們的主題是「上課請睜眼」，這是一台學習專注神器，利用相機對學生的臉部進行辨識，分析學生的上課專注程度，並適當給予回饋和反應，不但可以幫助老師知道學生的上課情形，同時也能提醒學生要專心上課，如此功能強大的機器將來也期待能再改良，甚至進一步商業化，為教育界開闢出一條新的康莊大道。

貳、動機與背景介紹

我們偶然在一次討論中提到上課睡覺的議題，發現老師要在上課時一次注意底下那麼多學生的上課狀況是相當困難的事，因此我們決定製作一台學習專注神器——上課請睜眼，利用相機拍攝學生臉部，不但可以即時觀察學生的上課專注度，也可以對學生不同的上課情況作出相對應的反應，像是給學生能打起精神的一拳，這樣一來就可以幫助老師減輕負擔，學生也能夠自主學習。

參、研究方法

一、軟體：CNN 網路辨識睜閉眼以及輸出執行訊號

此部分分為兩個部分，第一，我們架構**辨識與實驗目的**的關係，以及如何利用此關係，給出流程(如圖一)，描述如何將有用的訊號，傳遞給**執行端**。第二，探討執行端如何將**辨識結果**輸出給**機器**做出動作。



圖一：流程

1. 第一部分

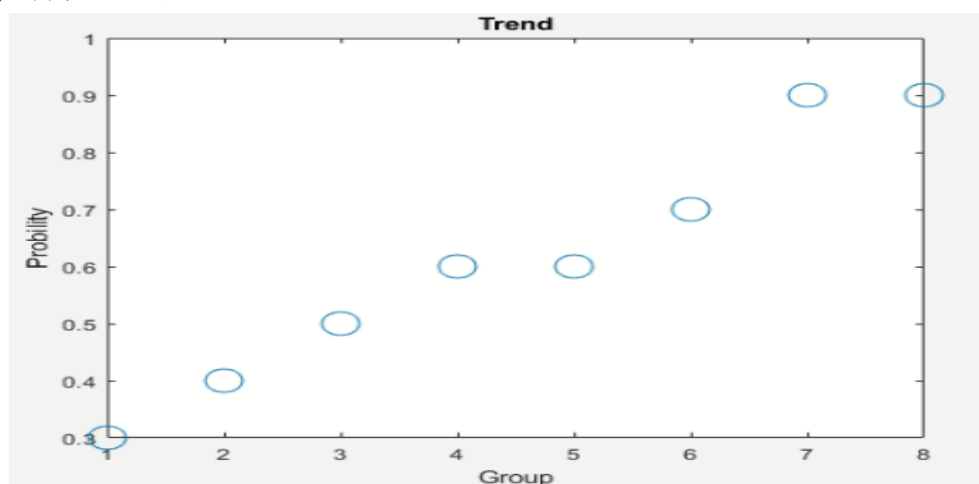
為了得知學生的學習效果，我們著重學生的專注力的測量，來瞭解上課品質，為了達此目的我們測量眼睛的開閉眼，我們使用 CNN 來偵測學生在每個時間點(frame)的睜閉眼狀況。在此同時，我們設立一個偵測時間 T，在這個時間 T 內，我們會盡可能蒐集 CNN 所判定每一個 frame 的狀況。

當我們在一個偵測時間 T 內，蒐集了 N 張 frame 的狀態結果，我們可以取一質因數(方便用途)，來達到將 N 分割成數個小區段，每個小區段會計算睜眼的比例，來代表小區段的結果。我們對此我們將此分割小區段以時間軸排列，可以得一睜眼的趨勢圖，如圖二。

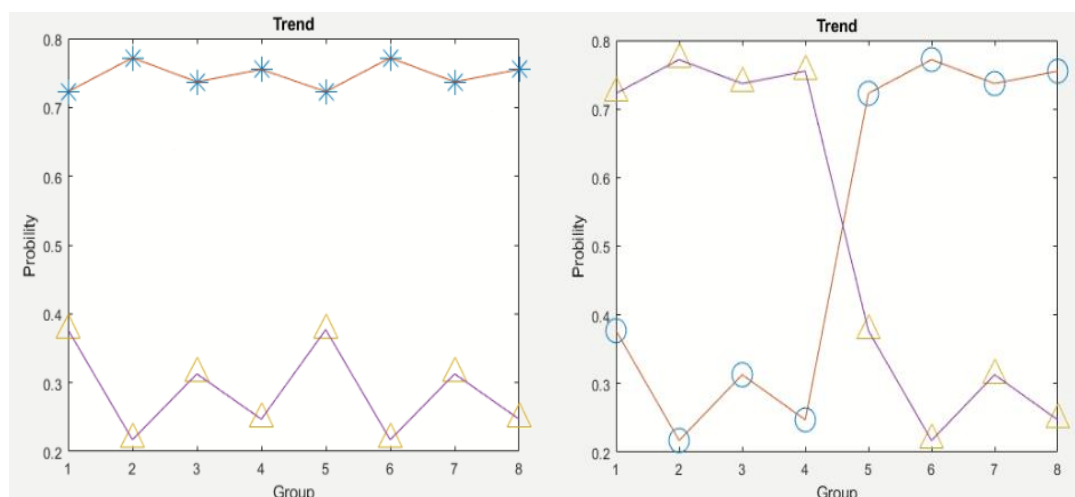
睜眼趨勢可以給我們四種狀況的輸出結果，第一種為睜眼-睜眼，此階段如圖三左的*標號曲線表示，第二種為閉眼-閉眼，如圖三左的三角形，第三種與第四種分別為睜眼-閉眼以及閉眼-睜眼，如圖三右的曲線。

真實傾向會與這四種狀況做匹配，將時間 T 所偵測的真實趨勢使用 Pearson correlation 做比對，得知個案的狀況。趨勢會輸出給機器手臂，完成分

析到輸出的階段。



圖二：睜眼趨勢圖



圖三：四種輸出結果

2. 第二部分

當我們將預測出的序列結果對事先預設好的「模板」進行 correlation 的比對完後，便可得到四種可能的結果(因我們只預設四種模板)，分別為 OO(持續睜眼)、OC(先睜後閉)、CO(先閉後睜)、CC(持續閉眼)，指出測試者在這段時間內的眼睛開合狀態，並給予機械手臂對應的行為。

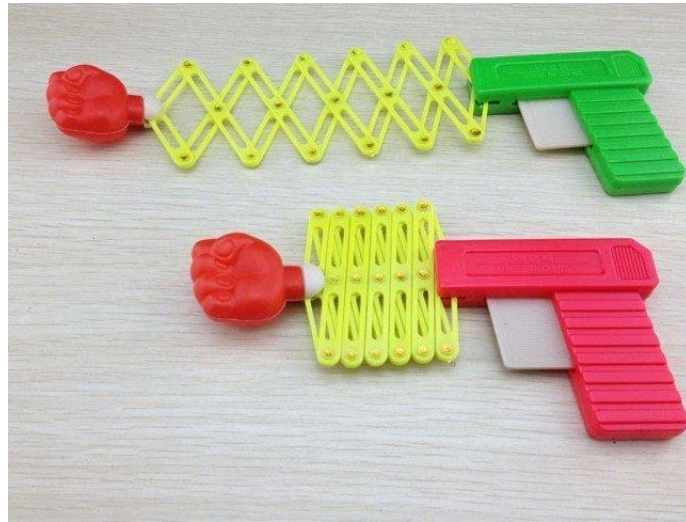


圖五：樹莓派

由於我們不希望整個裝置過於笨重，故不使用個人電腦上的相機，而用較輕巧的樹莓派，同時作為影像擷取與訊號傳輸(對伺服馬達 PWM 輸出)的媒介。我們只需遠端對樹莓派運行程式，輕巧的裝置便可做出對應的行為，這樣的設計對於硬體的要求不高，有助於裝置安裝的廣設與普及(畢竟書桌也就這麼大小)。

二、硬體

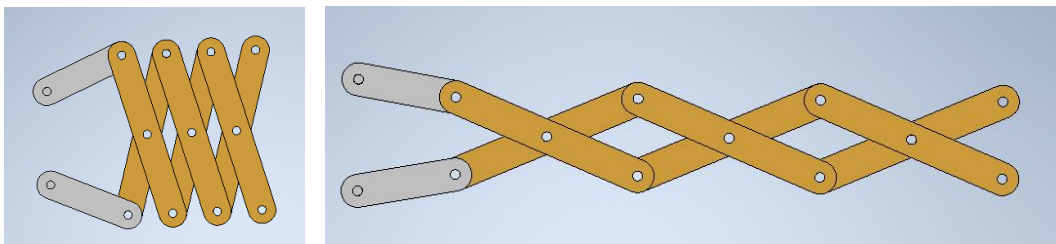
硬體部分我們需要能夠根據辨識結果做出反應的機構，且此機構為伸縮機構，能夠伸出拳頭對學生進行擊打，因此我們想到了一種玩具：拳頭槍（如下圖），此機構利用多根連桿交叉連接，能夠自由伸縮，且伸出長度也有一定的距離，此機構雖然簡單，但是能夠滿足我們的需求且效果十足。



圖六：拳頭槍

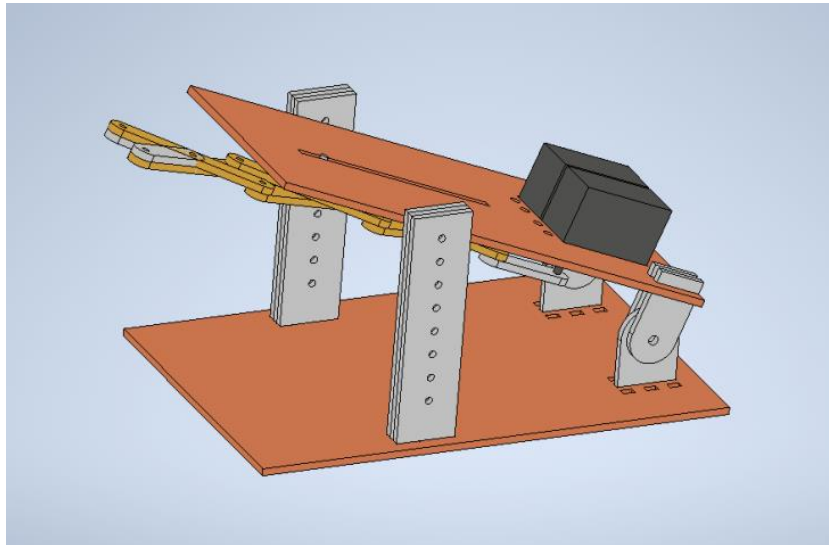
肆、模擬實驗

由於機構要能夠根據不同狀況伸出與收回，所以我們在機構的根部連接兩顆伺服馬達（如下圖，灰色部分為伺服馬達的舵盤），如此一來我們就能透過訊號來讓拳頭伸縮，當伺服馬達同時向內旋轉一定角度，機構便會受到兩個力矩使整體快速伸出，達到擊打的效果；伺服馬達同時向外旋轉即可將拳頭快速收回，等待下一次出擊。另外我們也製作了一個支架，能夠調整高度讓機構有一個仰角，使拳頭可以擊中受試者的臉部。



圖七：收縮情形

圖八：伸出情形

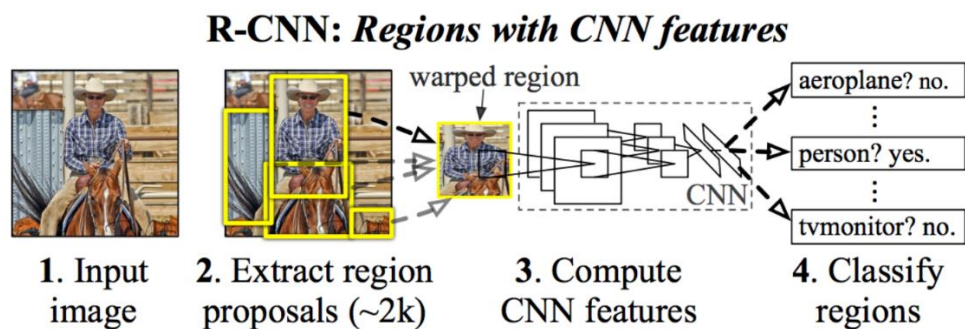


圖九：整體機構示意圖

伍、分析與結論

一、軟體

在進行影像辨識的模型訓練時，我們遭遇許多困難，例如資料庫多元性不足、攝影機拍出的影像範圍和訓練資料差異大等等問題，故輸入整張影像進入 CNN 是不足的，我們必須引入一些較潮的 object detection 的技術，其在模型訓練上的資料的 label 不僅有分類，還有「物件在影像」的位置訊息。這樣的模型也是基於 CNN 的模型做延伸，但其網路架構特殊，在一個網路裡安插了多項任務，比如「先隨機產生多個位置」、「辨識在此位置上的物件」、以及最後會加入一個 regression 層去對物件位置做最佳化。完成了這些任務後，只要輸入一張影像，模型便可 output 出物件的位置及分類，達成 object detection 的效果。



圖十：CNN 模型

然而，我們做出的 object detection 的強度不足，推測可能是訓練資料多元性不足，我們常常將「眉毛」辨識為「閉著的眼睛」，這將大幅影響我們評分的準確性。關於這個問題，我們只能盡量多增加訓練資料增加模型泛化能力，但處理這樣的辨識問題就花了我們大部分的時間，最後僅能訓練出一個還能接受的模型，

對我們四個組員的預測相當準確，但對其他受試者，僅有 70~80% 的準確率。

有了辨識的模型後，我們一直在思考如何能將評分做得更公平而有效，其中遇到最大的問題是：「閉眼需要較大的權重」，因一個學生在 10 秒測試內若有 2 秒鐘的時間都在閉眼，客觀上來講其實學生應該是要被判斷為想睡覺的，也就是閉眼情形一旦出現，其需要造成較大的懲罰項，於是我們不斷更換評分標準，最後才決定使用 correlation 的方式，去將測量結果與模板做相關性評判，我們發現這樣判斷的效果意外地好，也更能在學生「應該被懲罰時」做懲罰。

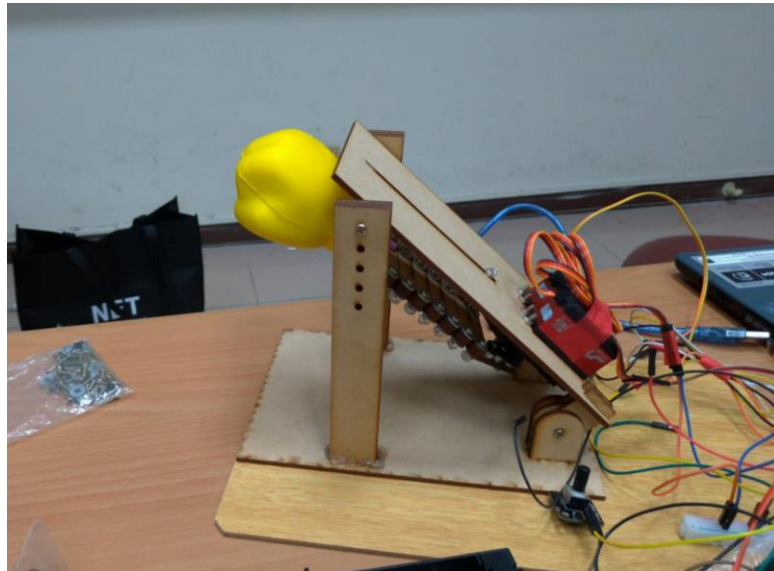
$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2} \cdot \sqrt{E(Y^2) - E(Y)^2}}$$

圖十一：Correlation 公式

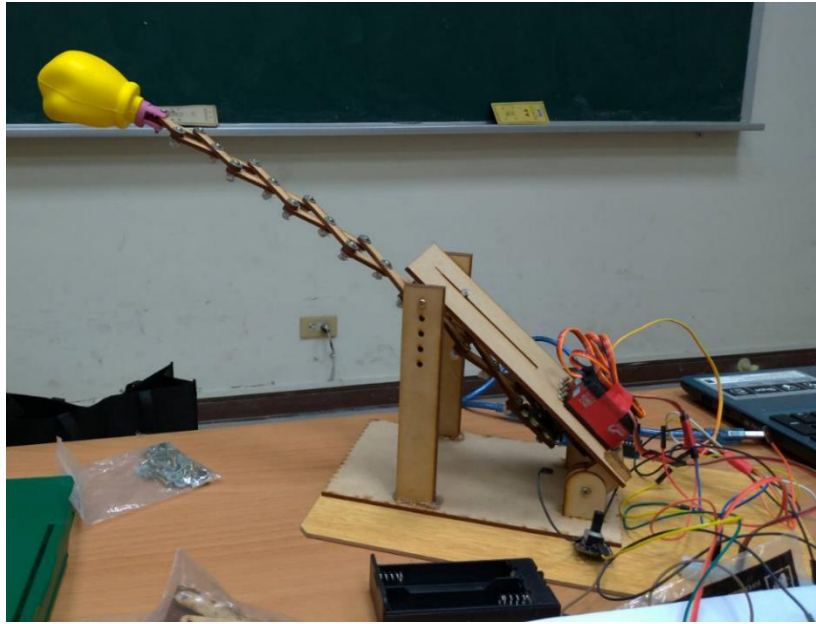
二、硬體

測試過程中，我們主要遇到兩個問題，首先是機構伸不出去，由於我們原先使用 Arduino 板 5V 供電，電壓不足會讓馬達轉動不穩定，進而使機構無法如預期伸縮。為了解決供電問題，我們改用兩顆 18650 電池供電，可到 8.4V，電壓問題即可改善，機構也能如預期伸縮。

第二個問題為機構在伸出去後，由於長度變長，且裝在前面的拳頭也有些重量，會使作用在馬達舵盤的力矩增大，導致機構整體下垂，無法達到我們預期的高度。經思考後我們決定在機構的第二個交點位置增加支撐，透過長螺絲穿過上層板的軌道並用螺帽卡住，減輕馬達舵盤的負擔，測試後也發現機構下垂的情況也有改善，能符合我們預期的高度。



圖：機構收起情形



圖：機構伸出情形

陸、未來改善方向：

一、效果更好

1. 精進 model，使能辨識更多學生姿勢：趴下、滑手機、睜著眼睛但卻不是在專心等等更多情形，model 若能辨識出更複雜的學生行為，對學生專注度的評分將能更多元而廣泛。精進的方式可能為增加訓練資料的多元性、使用更潮的 object detection 的技術等等。
2. 增加更多拍攝相機，以記錄學生更多的特徵資料，從不同角度進行拍攝驗證可降低模型預測的錯誤率。
3. 硬體方面，可以增加更多互動式機構，像是噴水、吹氣等，多元化的「懲罰」也會不容易學生感到疲勞，適時提醒學生專注學習。

二、應用更多

1. 將評分轉換為學習專注曲線，供授課老師參考。拉長時間序列，藉由機器長時間觀察學生行為(例如：一節課)，將評分對時間作圖，授課老師便可覆盤，觀察學生專注程度在時間軸上的變化，藉以修正上課內容與方向，達到更好的教學品質。
2. 此機器可同時作為學習自律神器，當學生在家自學時可開啟「上課請睜眼」評分，如此即可督促自我學習，適合那些想要好好學習但自律能力不好的人，「上課請睜眼」如同成為了「隱形的家教」，提供了陪讀功能。
3. 整台機器也可一體成形，將不必要的線路裝置包裝起來，更有商品感，未來或許能進一步商業化，為寒窗苦讀的莘莘學子提供適時的幫助。

柒、分工

一、分工情形

立楷：嵌入式裝置(樹莓派)安裝與軟硬體整合

亭瑋：硬體機構設計與製作

冠綸：硬體機構設計與製作

秉叡：模型訓練與主程式撰寫

二、感想

立楷：我認為這次的專注度辨識與分析是一個簡單又困難的任務，因網路上已有眾多人眼開闔的預訓練模型，但真正評估一個學生是否為專注上課應有更多更多需要考量的因素，比如是否在滑手機啦，使否只是睜開眼睛在發呆呢，諸如此類的問題，我們本來期待能輸入大量資料，讓機器替我們做這些行為上的判斷，但後來發現這是一個大工程，比如如果鏡頭沒有照到學生的主要特徵該怎麼辦？如果照的角度不對，模型是否就無法辨識？我們明白一個強大的模型應該要能克服這些問題，但受限於時間與能力，我們只能訓練出「需指定受試者在鏡頭的範圍」的模型，這是我們將就妥協的結果，然而若有時間，我們一定會做出泛化能力更好的模型，真正地完成這個「上課請睜眼」的學習神器。

亭瑋：這次專題是個很有趣的題目，不但有效果，難度也頗大，很值得挑戰。過程中遇到很多問題，花了很多時間在做重複的事，最後還是有完成基本的辨識與攻擊反應，真的是很有成就感！

冠綸：在這次的專題中，我們試著利用機器學習來進行影像辨識，同時也自行設計了機構來與訓練出的模型結合，儘管遇到了很多困難，同時也有其他課程的期中期末壓力，我們仍然花費了許多時間在這個專題上，包括小組討論、研究模型、設計機構等等，這樣一學期下來真的是很感謝組員們，一起努力地將這個專題完成。

秉叡：CNN 真的很難處理，還要考慮頭痛以及各種背景雜訊的問題，還好這堂課終於結束了，謝謝各位夥伴，這一切真是不容易啊。

捌、參考文獻

[1] 基於 CNN 的人眼定位與狀態分類

<http://gb.oversea.cnki.net/KCMS/detail/detail.aspx?filename=1019204958.nh&dbcode=CMFD&dbname=CMFDTEMP>

[2] Raspberry Pi Camera + Python + OpenCV

<https://ricelee.com/raspberry-pi-camera-python-opencv-2days-workshop-at-nfu/>

[3] Object detection

<https://medium.com/cubo-ai/%E7%89%A9%E9%AB%94%E5%81%B5%E6%B8%AC-object-detection-740096ec4540>

[4] Correlation and dependence

https://en.wikipedia.org/wiki/Correlation_and_dependence

[5] CNN 入門，什麼是微調？

<https://zhuanlan.zhihu.com/p/35890660>

[6] 一文讀懂目標檢測：R-CNN、Fast R-CNN、Faster R-CNN、YOLO、SSD

<https://kknews.cc/zh-tw/code/k2yqmvb.html>

[7] Python+樹莓派+YOLO 打造一款人工智慧照相機

<https://www.itread01.com/article/1514880543.html>

[8] Support Vector Machine

<https://scikit-learn.org/stable/modules/svm.html>