

數字圖像處理

干皓丞，2101212850, 信息工程學院; 鄭翰濃，2101212849, 信息工程學院

2022 年 5 月 14 日

1 作業目標與章節摘要

面向網路教學的聽課狀態視頻分析系統作業，此作業目標為完成該期末作業中所規劃的疲勞檢測部分。該作業為原期末報告的改進部分。

2 作業內容概述

作業可以從 GitHub 下的 kancheng/2022-dip-final-report-for-deeplearning-system 專案找到，作業程式碼目錄為 2022-dip-final-report-for-deeplearning-system/sys/。實際執行的環境與實驗設備為 Google 的 Colab、MacBook Pro (Retina, 15-inch, Mid 2014)、Acer Aspire R7 與 HP Victus (Nvidia GeForce RTX 3060)。



Fig. 1. 作業專案位置

3 作業概念規劃與架構

根據原本的期末架構規劃，其功能為一個教學系統，包含了學習檢測、代碼提交、簽到功能、課業搶答、考試功能、會員功能、視頻觀看、作業繳交、教學直播，而當中判斷學生的上課注意力部分，則是在教學直播與視頻觀看的部分。當中判斷學生的狀況，則從有沒有人在電腦攝像頭前、正臉有無、眼睛視線有無與人的狀況是否疲勞。而疲勞的部分，從原期末報告可知，可以從人臉的特徵來判斷是否是如此。而此次報告則根據原本的可行規劃的基礎，繼續進行工作。

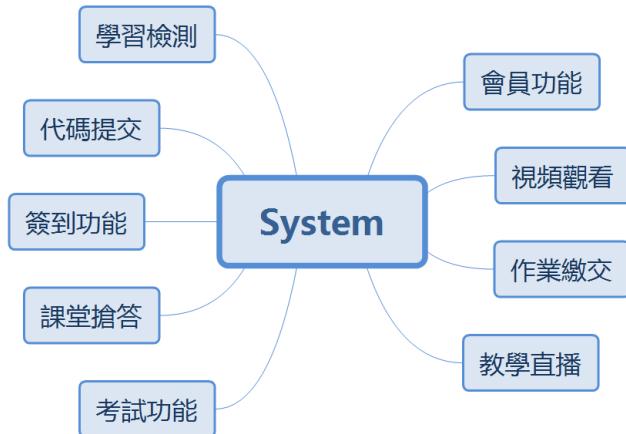


Fig. 2. 期末作業系統架構規劃

原系統設計是用 Python Flask 進行處理，當中整合當下前端框架 Vue 和深度學習框架 Pytorch，整個功能較為複雜，所以此次作業只是單純的驗證，人類臉部的狀態，也就是人的睜眼與眨眼狀況來進行判斷。一但有眨眼便會發起警告。

```

1 if (由沒有人在鏡頭) {
2     if (在鏡頭前的人有沒有可檢測的正臉) {
3         if (眼睛的視線) {
4             回傳
5         }
6         if (判斷人的表情是否為疲憊狀態) {
7             回傳
8         }
9         回傳
10    }
11    回傳
12 }
  
```

J Cech 等人提出了一種實時檢測標準攝像機視頻序列中眨眼的算法，近期在野外數據集上訓練的地標檢測器對相機的頭部方向、不同的照明和面部表情表現出出色的魯棒性，該研究表明，地標的檢測足夠精確，可以可靠地估計眼睛張開的水平。因此，所提出的算法估計地標位置，提取單個標量 - 眼睛縱橫比 EAR - 表徵每幀中的眼睛張開。最後，SVM 分類器將眨眼檢測為短時間窗口中的 EAR 值模式。其簡單的算法在兩個標準數據集上優於最先進的結果。

根據該篇論文，人臉關鍵點檢測中人眼共有 6 個關鍵點，睜眼時與閉眼時的關鍵點狀態如下圖，該篇研究提出了這個公式：

$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \|p_1 - p_4\|} \quad (1)$$

通過該公式的歐氏距離計算，我們可以得到某一幀中眼睛是睜開還是閉著的狀態。計算左眼和右眼的平均 EAR 值，若 EAR 值小於某一閾值，則表明了這個人在某一幀中是睜眼還是閉眼的狀態。設定閾值 n，連續 n 幀中若眼睛都是閉著的狀態，那麼代表這個人眨了一次眼。

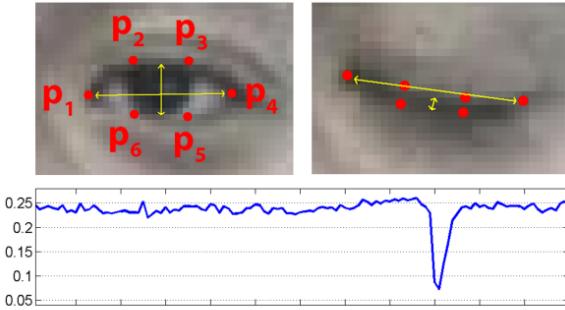


Figure 1: Open and closed eyes with landmarks p_i automatically detected by [1]. The eye aspect ratio EAR in Eq. (1) plotted for several frames of a video sequence. A single blink is present.

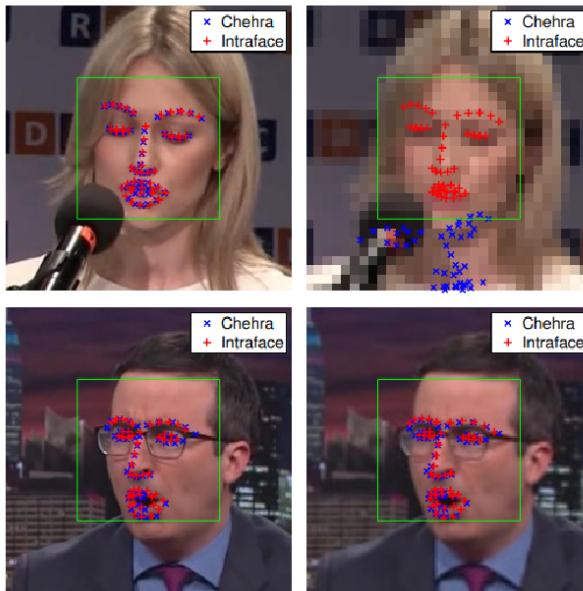


Figure 3: Example images from the 300-VW dataset with landmarks obtained by Chehra [1] and Intraface [16]. Original images (left) with inter-ocular distance (IOD) equal to 63 (top) and 53 (bottom) pixels. Images subsampled (right) to IOD equal to 6.3 (top) and 17 (bottom).

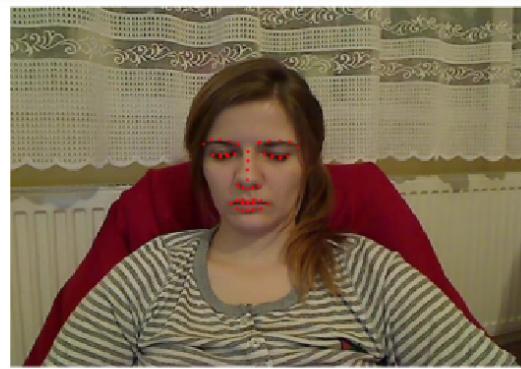


Figure 2: Example of detected blinks. The plots of the eye aspect ratio EAR in Eq. (1), results of the EAR thresholding (threshold set to 0.2), the blinks detected by EAR SVM and the ground-truth labels over the video sequence. Input image with detected landmarks (depicted frame is marked by a red line).

Fig. 3. 期末作業文獻支持

4 成果展示

從下面的測試資料中可以看到，鏡頭前的人臉在閉上眼睛時畫面會跳出警告的訊息，而睜開眼睛則能夠看到當下是安全的文字訊息。當中綠色方框為人的臉部，而紅點所圈起的結構則是人的五官結構。

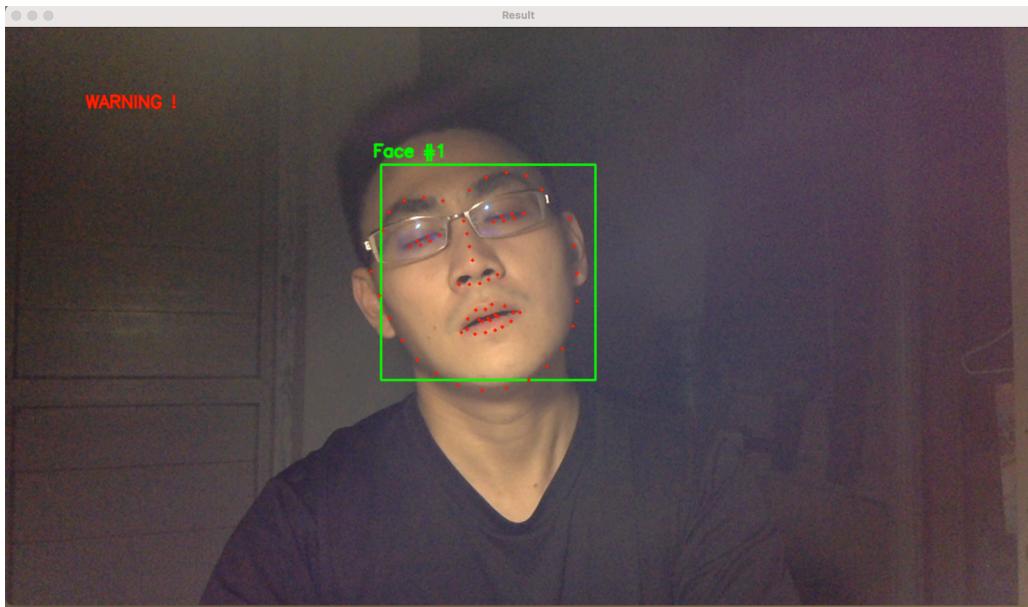


Fig. 4. 人臉閉眼狀態

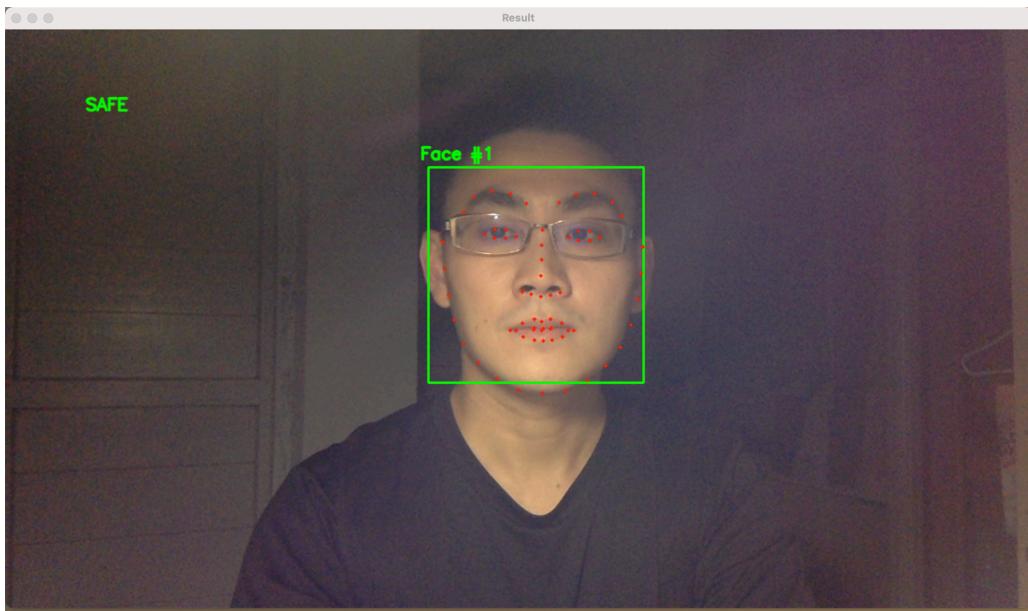


Fig. 5. 人臉開眼狀態