

基於深度網路之人臉情感檢測系統

開放平台期末專題

組長

1051447 梁雅錡

組員

1051536 林淳瑄

1051539 李昇禎

1053336 陳妙樺

1053347 陳吟瑄

June 14, 2019

Contents

1	產品介紹	4
1.1	目的	4
1.2	預期讀者和閱讀建議	5
1.3	計畫範圍	5
2	總體描述	6
2.1	產品透視	6
2.2	產品功能	7
2.3	用戶類別和特質	8
2.4	操作環境	8
2.4.1	軟體需求	8
2.4.2	硬體需求	9
2.5	設計和實作上的限制	9
2.6	系統假設和依賴	10
3	介面規格設計	11
3.1	使用者介面	11
3.2	硬體介面	12
3.3	軟體介面	12
4	系統特色	14
4.1	重點描述(Description and Priority)	14
4.2	刺激/反應(Stimulus/Response Sequences)	14
4.3	功能性需求	19

5	其他非功能性需求	20
5.1	性能需求	20
5.2	Safety需求	20
5.3	Security需求	20
6	附錄	21
6.1	量化成果	21

1 產品介紹

1.1 目的

隨著世界經濟的快速發展，產業轉型已成為全球化趨勢下的經濟發展策略。為了同時滿足客戶的各種不同需求情境，未來台灣產業的轉型，自然也面臨一些挑戰，包括：產業轉型腳步落後、人口紅利的消失、創新集中特定領域且效益不足、以及區域經濟整合等議題。

其中，產業轉型腳步落後的問題，以台灣服務產業為例，趨勢已邁向行動科技，且朝智慧生活應用發展，例如：智慧建築、智慧工廠、智慧醫療等，但服務產業在自動化管理的導入與整合能力皆有待強化，包括：服務態度、商品品質、以及倉儲管理等管理手段。

在服務態度的自動化管理上，充滿笑容的服務態度已成為服務產業的互動過程中，最不可或缺的必備條件之一。傳統管理從業人員服務態度的方式，是經由問卷、電話調查，獲得客戶對從業人員服務態度的反應後，進而檢討從業人員服務態度的缺失。

然而，上述方法是透過人工且被動的方式，得知從業人員的服務狀況，服務產業的管理者，無法主動獲得從業人員當下的服務態度表現。由此可知，服務產業在管理端的自動整合能力確實有待加強。

因此，我們開發基於深度網路之人臉情感檢測系統，如圖1.1所示，將人臉的情感做明確的量化，檢測從業人員的情緒反應，透過檢測結果可隨時針對服務態度做調整，提供未來的服務產業在服務態度的品質上，擁有更優質的自動化管理系統，以提高服務產業的競爭力與優勢條件，加速台灣服務產業成功轉型升級。

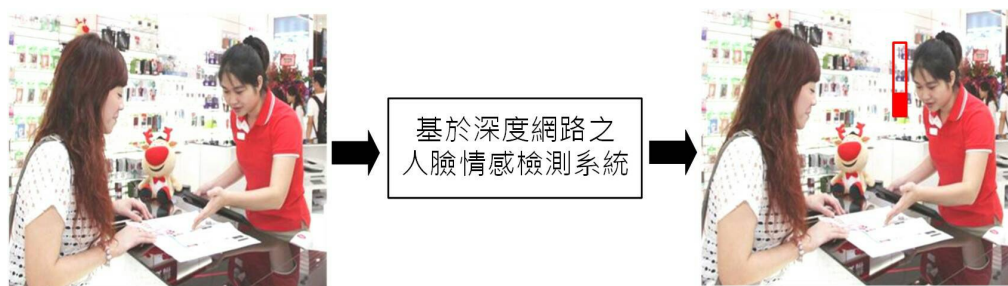


圖 1.1: 基於深度網路之人臉情感檢測系統，讀取從業人員影像，檢測從業人員的情緒反應。

1.2 預期讀者和閱讀建議

本SRS文件提供給使用者、開發團隊與維護人員閱讀，使用者透過閱讀本文件，了解產品以及操作說明，開發團隊進行產品開發時的依據，與維護人員日後維護系統時的參考文件。

1.3 計畫範圍

與其他章節有所重複，請參考文件1.1。

2 總體描述

2.1 產品透視

爲了能有效的檢測從業人員的服務態度，並針對個人進行情感檢測，我們開發基於深度網路之人臉情感檢測系統，如圖2.1所示，本系統分爲兩部分實施，第一部分是人臉偵測模組，第二部分是情緒量化模組。

第一部分：人臉偵測模組，我們爲了針對個人進行情感檢測，利用服務場所的監視器或攝影機，讀取拍攝到的從業人員影像，並且輸入至人臉偵測模組，進行人臉偵測，偵測各個從業人員的人臉特徵。

第二部分：情緒量化模組，我們導入人臉偵測模組偵測到的人臉特徵，進行情緒量化，計算該人臉的笑容強度，以視覺化方式呈現於監控畫面，具體呈現從業人員在服務態度上的情緒表現。

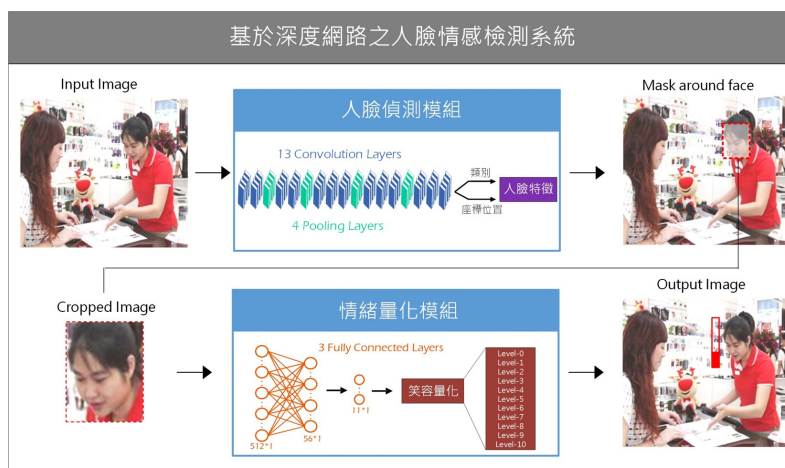


圖 2.1: 基於深度網路之人臉情感檢測系統之架構圖

2.2 產品功能

為提供未來的服務產業在服務態度的品質上，擁有更優質的自動化管理系統，我們提出基於深度網路之人臉情感檢測系統，透過監視器或攝影機，時刻擷取個別從業人員的影像，檢測從業人員在服務態度上的情緒表現，讓管理者可以即時得知從業人員的服務態度，並針對檢測結果進行矯正，因此，本產品列舉有下列功能：

● 偵測從業人員人臉特徵，針對個人進行情感檢測

我們將透過工作場所的監視器與攝影機，擷取從業人員影像，將擷取到的影像輸入本系統，經由本團隊設計的人臉偵測模組即時分析後，便可獲得符合該從業人員的人臉特徵，以便針對個人進行情感檢測，如圖2.2所示。

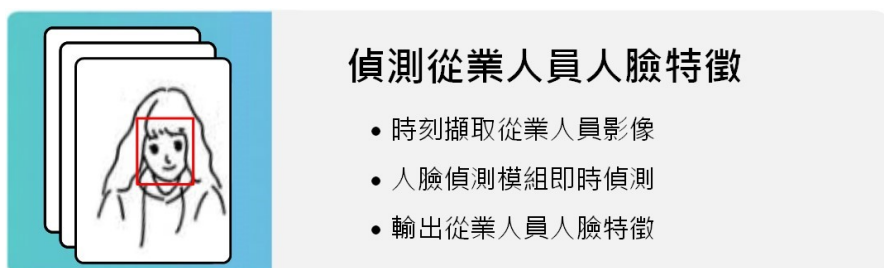


圖 2.2: 偵測從業人員人臉特徵示意圖

● 計算從業人員笑容強度，監視從業人員服務態度

讀取人臉偵測模組輸出的人臉特徵，透過我們設計的情緒量化模組，針對從業人員的人臉特徵，計算笑容量化數值，將人臉的情感做明確的量化，並以視覺化的方式，顯示笑容量化數值於監控畫面，讓管理者可以隨時針對檢測結果，進行服務態度品質上的檢討與控管，如圖2.3所示。

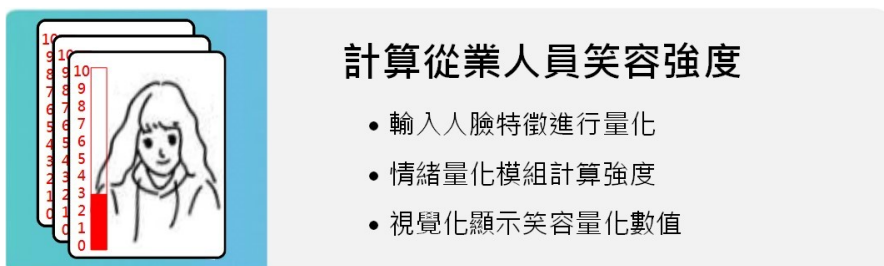


圖 2.3: 偵測從業人員人臉特徵示意圖

2.3 用戶類別和特質

在服務態度的自動化管理上，充滿笑容的服務態度已成為服務產業的互動過程中，最不可或缺的必備條件之一。因此，我們基於深度網路之人臉情感檢測系統，可量化從業人員的情緒，將本產品導入服務產業的自動化管理系統。因此，本產品的用戶類別可以分為(1)從業人員檢測自身情緒反應，(2)管理者監視從業人員服務狀態，本產品的用戶類別如下：

(1) 從業人員檢測自身情緒反應

本產品可以即時偵測個別從業人員的臉部情緒，將從業人員的情緒分成笑容量化數值0~10，共11種笑容量化數值，情緒越開心笑容量化數值也越大。其中，這11種笑容量化數值可粗略分為4種情景，分別為無笑容(笑容量化數值0~1)、微笑(笑容量化數值2~4)、露齒笑(笑容量化數值5~7)、張嘴笑(笑容量化數值8~10)。透過將從業人員的臉部情緒做清楚的量化，可以檢測各個從業人員在服務態度上的情緒反應。

(2) 管理者監視從業人員服務態度

我們為了具體呈現從業人員在服務態度上的情緒表現，設計將從業人員的檢測結果，以視覺化的方式，即時輸出於監控畫面上，管理者透過監視從業人員的笑容量化數值，可以針對從業人員的服務態度做矯正，達到有效的控管從業人員服務態度。

2.4 操作環境

2.4.1 軟體需求

- (1) 作業系統：Windows 10
- (2) 主要開發程式語言：Python 3.6
- (3) 程式開發軟體：Pycharm

2.4.2 硬體需求

- (1) 攝影鏡頭：具有720、1080p的高解析度，以及自動對焦的功能
- (2) 螢幕：1080p/60fps的最大解析度
- (3) NVIDIA Jetson TX2：JetPack 4.2

2.5 設計和實作上的限制

由於本產品需安裝在服務場所，因此需要攝影機或監視器來拍攝從業人員影像。我們預計於室內之小型服務場所(空間小於15m)，安裝HD網路攝影機；於室內之大型服務場所(空間大於15m)或室外任何場所，安裝夜視半球監視器，如圖2.4所示。其中，攝影機與監視器分別具有720、1080p的高解析度，以及自動對焦的功能，因此，均能在服務場所有效拍攝各個從業人員影像。

	<table><tr><th>HD 網路攝影機</th><th>技術規格</th></tr><tr><td>具有適用於 HD 視訊通話自動對焦功能的可折疊攜帶網路攝影機</td><td>最大解析度：720p/30fps 對焦類型：自動對焦；視野: 69° 適用：電腦螢幕通用腳架 範圍：15m²</td></tr></table>	HD 網路攝影機	技術規格	具有適用於 HD 視訊通話自動對焦功能的可折疊攜帶網路攝影機	最大解析度：720p/30fps 對焦類型：自動對焦；視野: 69° 適用：電腦螢幕通用腳架 範圍：15m ²
HD 網路攝影機	技術規格				
具有適用於 HD 視訊通話自動對焦功能的可折疊攜帶網路攝影機	最大解析度：720p/30fps 對焦類型：自動對焦；視野: 69° 適用：電腦螢幕通用腳架 範圍：15m ²				
	<table><tr><th>夜視半球監視器</th><th>技術規格</th></tr><tr><td>具有1080p高解析度以及感紅外線IR鍍膜處理鏡頭的夜視半球監視器</td><td>最大解析度：1080p/60fps 對焦類型：自動對焦；視野: 360° 適用：室內、外牆角皆可安裝 範圍：25m²</td></tr></table>	夜視半球監視器	技術規格	具有1080p高解析度以及感紅外線IR鍍膜處理鏡頭的夜視半球監視器	最大解析度：1080p/60fps 對焦類型：自動對焦；視野: 360° 適用：室內、外牆角皆可安裝 範圍：25m ²
夜視半球監視器	技術規格				
具有1080p高解析度以及感紅外線IR鍍膜處理鏡頭的夜視半球監視器	最大解析度：1080p/60fps 對焦類型：自動對焦；視野: 360° 適用：室內、外牆角皆可安裝 範圍：25m ²				

圖 2.4: 本系統所需之環境需求示意圖

2.6 系統假設和依賴

本產品做出以下假設：

1. 使用的套件包含：

套件	版本
OpenCV	3.4.3
Tensorflow	1.9.0
Keras	2.2.2

2. 使用者的電腦有內建鏡頭或能外接鏡頭

3. 使用者所在的環境能夠連上網路。

若以上假設不正確、不一致或被更改，就會使系統受到影響。

3 介面規格設計

3.1 使用者介面

本產品在計算使用者之笑容量化數值時，我們利用電腦螢幕作為監控畫面的形式，顯示使用者影像以及其笑容量化數值，如圖3.1所示。其中，圖3.1(a)為從業人員以微笑表情服務客人時的情形，笑容量化數值檢測結果為3；圖3.1(b)為從業人員以露齒笑表情服務客人時的情形，笑容量化數值檢測結果為6。透過即時顯示從業人員的笑容量化數值，管理者可從監控畫面得知各個從業人員的笑容數值，並依管理者需求，矯正從業人員服務態度。

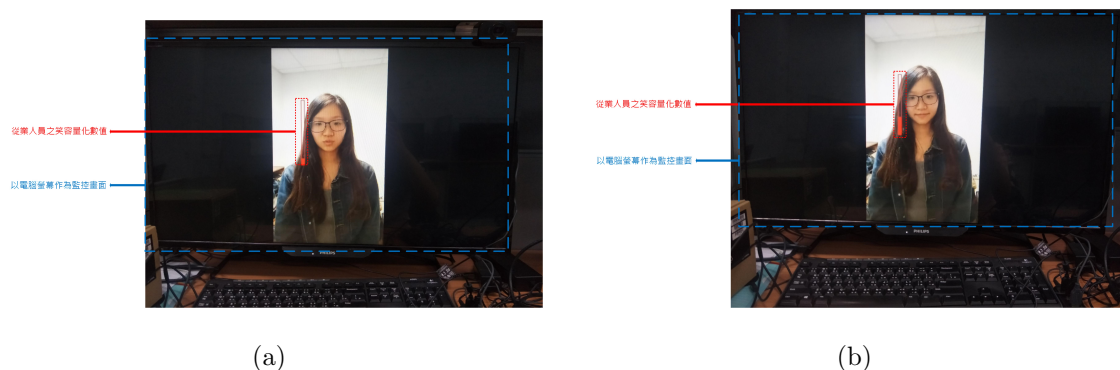


圖 3.1: 使用者介面呈現結果。(a)使用者笑容量化數值為3；(b)使用者笑容量化數值為6。

3.2 硬體介面

硬體設計方面，我們使用NVIDIA Jetson TX2，如圖3.2(a)所示，來執行本系統的程式，且搭載了最新版本的SDK — JetPack 4.2，可提供快速又精準的人工智慧推論機制，因此本團隊選擇使用NVIDIA Jetson TX2來展示本系統的執行成果。並且，我們利用HD網路攝影機連接NVIDIA Jetson TX2，以拍攝從業人員影像。HD網路攝影機如圖3.2(b)所示，是具有適用於HD自動對焦功能的可折疊攜帶網路攝影機。透過硬體介面的設計，使本系統得以在服務場所有效拍攝各個從業人員影像。



(a)



(b)

圖 3.2: 本產品之硬體設備。(a)NVIDIA Jetson TX2實體圖;(b)HD網路攝影機實體圖。

3.3 軟體介面

軟體設計方面，本產品主要分成四個步驟，分別為輸入圖片(Input image)、人臉位置(Bounding box)、裁切人臉位置(Cropped image)、以及輸出圖片(Output image)，如圖3.3所示。首先，在輸入圖片時，我們以HD網路攝影機為例，解析度為720p，即大小為1280*720的圖片；接著透過本系統設計之人臉偵測模組，獲得該從業人員的臉部位置，且以bounding box的形式儲存，其具有4個bit，分別為座標x、座標y、寬(w)、高(h)；獲得人臉位置後，依據此位置裁切此範圍的圖片，得到寬為w、高為h的裁切圖片；最後，情緒量化模組會根據此裁切圖片進行笑容量化數值的計

算，計算後的笑容量化數值將以視覺化的方式與原輸入圖片，共同輸出於監控畫面上，輸出圖片大小仍為1280*720。

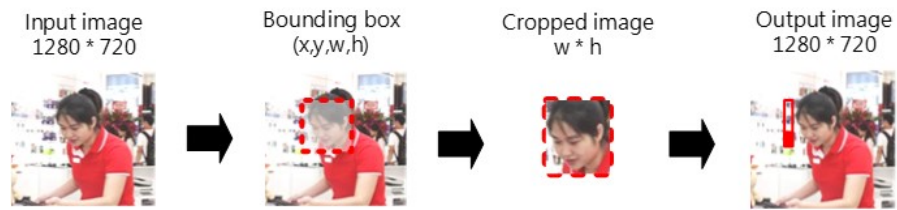


圖 3.3: 本產品之軟體設計分成四個步驟實行，分別為輸入圖片(Input image)、人臉位置(Bounding box)、裁切人臉位置(Cropped image)、以及輸出圖片(Output image)。

4 系統特色

4.1 重點描述(Description and Priority)

本系統的主要目標，是提出一套基於深度網路之人臉情感檢測系統，將拍攝後的從業人員影像，運用人工智慧(Artificial Intelligence, AI)中的深度學習(Deep Learning)，配合卷積神經網路(Convolution Neural Networks, CNNs)進行情感檢測，提供服務產業在服務態度的自動化管理系統。

本系統分為兩部分實施，第一部分是人臉偵測模組，第二部分是情緒量化模組。首先，將攝影機或監視器拍攝的圖片，輸入至人臉偵測模組，獲得該從業人員的臉部位置，並將它裁切。之後將裁切的圖片輸入至情緒量化模組，進行笑容量化數值的計算，並視覺化顯示於監控畫面，具體呈現從業人員在服務態度上的情緒表現。

4.2 刺激/反應(Stimulus/Response Sequences)

- 人臉偵測模組

為了針對個人進行情感檢測，我們設計一套人臉偵測模組，偵測各個從業人員的人臉特徵。首先，我們利用服務場所的監視器或攝影機，讀取拍攝到的從業人員影像，並且輸入至人臉偵測模組，如圖4.1所示，進行人臉偵測。

首先，讀取從業人員的影像後，我們採用Google提供的TensorFlow模組，來設計卷積神經網路架構模型。我們預計提出的卷積神經網路架構，是由1個輸入層、13個卷積層(Convolutional Layer)、4個池化層(Pooling Layer)、以及1個輸出層所構成。

其中，在卷積層中，將輸入圖片與濾波器(Filter)進行卷積運算，並使用線性整

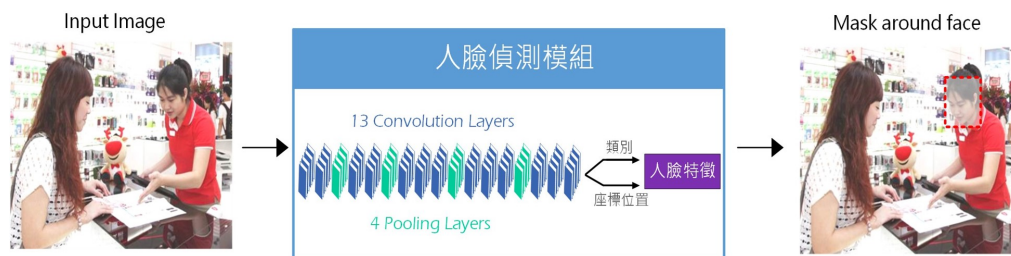


圖 4.1: 人臉偵測模組架構圖

流函數 (Rectified Linear Unit, ReLU)，作為活化函數(Activation Function)後，輸出特徵圖(Feature Maps)，取得從業人員影像的局部特徵。在池化層中，採用最大池化層(Max Pooling Layer)，並選出每個特徵圖的局部範圍中最大的數值，成為該局部的數值，以減少特徵圖像素，並保留重要特徵。

並且，為了使偵測更準確，我們藉由調整卷積層、池化層的特徵圖數量，以及設計模型的損失函數(Loss Function)調整模型參數，使得訓練出來的係數以及誤差達到最佳的準確率(Accuracy)。

當從業人員影像經過卷積神經網路計算後，輸出層將輸出兩個資訊，分別為座標位置、類別。其中，座標位置是經由卷積層跟池化層的運算後，所得之人臉位置；類別則是判斷該人臉位置，是屬於有人臉的前景(Foreground)類別、或是無人臉的背景(Background)類別。

最後，將這兩個資訊合併成人臉特徵，共同作為人臉偵測模組的輸出，便可得到個人的臉部特徵，再透過情緒量化模組，針對個人進行情緒檢測。

● 情緒量化模組

透過人臉偵測模組，獲得從業人員的人臉特徵後，為了量化人臉特徵的情緒，我們設計一個情緒量化模組。首先，導入人臉偵測模組輸出的人臉特徵，如圖4.2所示，進行情緒量化。

讀入人臉特徵後，將屬於前景的人臉座標位置，經由2個全連接層(Fully Connected Layer)，並使用線性整流函數，配合大量數據庫的訓練，計算該人臉的笑容量化數值。

其中，笑容量化數值的計算，透過獲得屬於前景的人臉座標位置，計算每個人

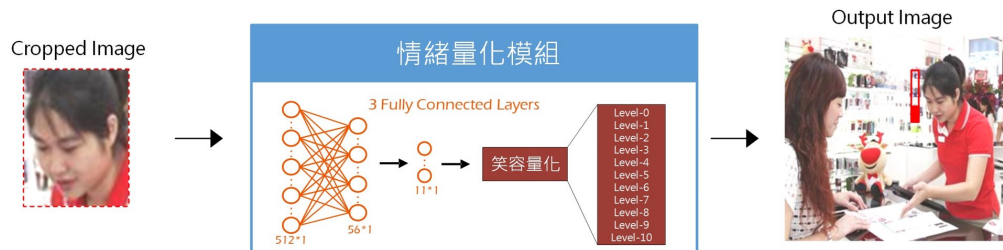


圖 4.2: 情緒量化模組架構圖

臉與各個笑容等級的符合機率，且選出符合機率最大的笑容等級，成為個別從業人員的笑容量化數值。再透過一個輸出層，利用全連接的方式，輸出從業人員的笑容量化數值。

最後，為了讓管理者可以即時查看從業人員的情緒表現，我們將笑容量化數值即時顯示於監控畫面，並以視覺化方式，具體呈現每一位從業人員在服務態度上的情緒表現。

此外，在我們這組所設計的情緒量化數據庫中，如圖4.3所示，包含0到10共11個笑容等級，等級越高代表開心程度越高，其劃分等級的標準，主要是以眼部及嘴部，來做為量化笑容的主要依據。因此，我們運用情緒量化數據庫，來訓練上述的模型，使情緒量化模組能夠準確地量化笑容等級。

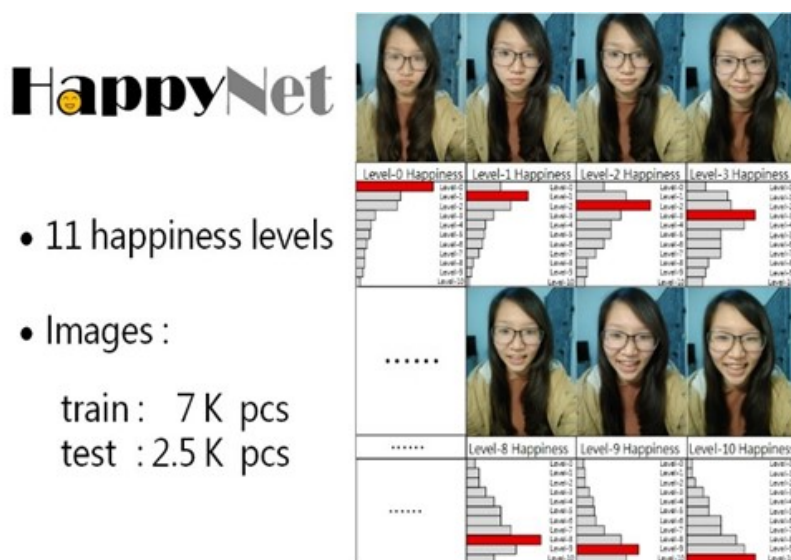


圖 4.3: 情緒量化數據庫

- 反應(response)

本系統的測試案例可粗略分為兩種：包含人臉的案例、不包含人臉的案例，兩種案例分別呈現如下：

在包含人臉的案例中，我們進行單人與多人的測試案例。首先，在單人的測試案例，如圖4.4所示，透過本系統的模組計算後，圖4.4(a)、4.4(b)、4.4(c)、以及4.4(d)分別被檢測為笑容量化數值0、3、6、9的笑容等級，因此，本系統在有包含人臉的測試上，均能夠成功偵測人臉，並計算出符合該人臉特徵的笑容量化數值。情緒。

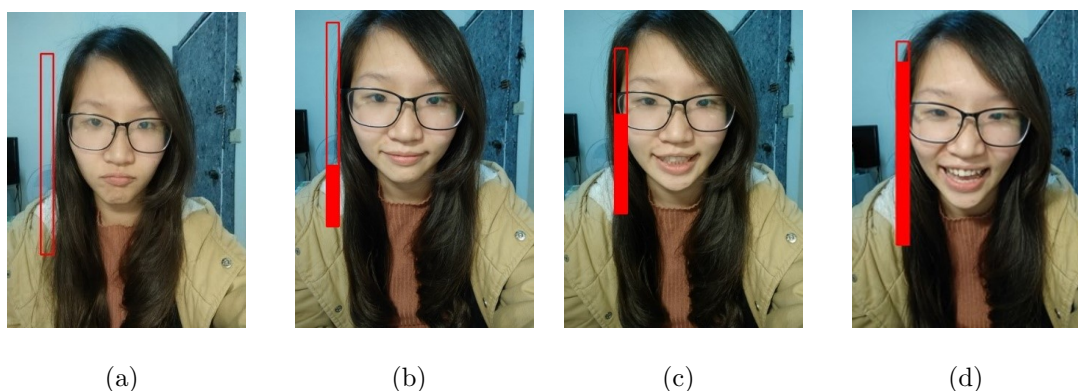


圖 4.4: 包含人臉的測試案例成果圖。(a)檢測出笑容量化數值為0；(b) 檢測出笑容量化數值為3；(c) 檢測出笑容量化數值為6；(d) 檢測出笑容量化數值為9。

在多人的測試案例中，我們分成4種情境進行測試，分別會拍攝兩位、三位、四位、以及五位從業人員，如圖4.5所示，我們在每種情況都能準確檢測到每一位從業人員的情緒，且笑容量化數值均符合個別從業人員的笑容強度，由此可知，本系統在同一個拍攝視角內，可以一次檢測多位從業人員的服務態度，以滿足服務產業的需求。



(a)



(b)



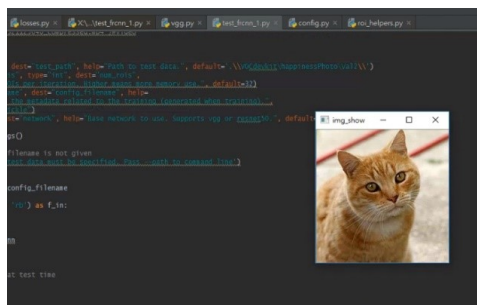
(c)



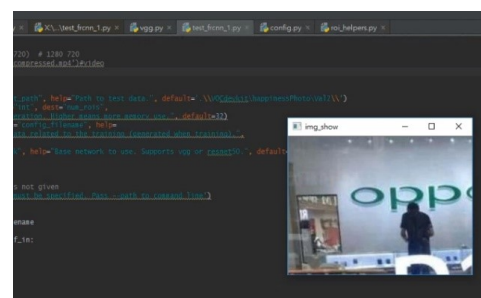
(d)

圖 4.5: 包含人臉的多人測試案例成果圖。(a)兩位；(b) 三位；(c) 四位；(d) 五位。(a)兩位；(b) 三位；(c) 四位；(d) 五位。

在不包含人臉的案例中，如圖4.6所示，均為不包含人臉的影像，透過本系統的模組計算後，圖4.6(a)、4.6(b)都沒有偵測到符合人臉的特徵，因此不做笑容數值的計算，也沒有進行額外的影像處理，會直接輸出原始讀取的輸入圖片。



(a)



(b)

圖 4.6: 不包含人臉的測試案例成果圖。(a)沒有偵測到人臉，故沒有輸出笑容量化數值；(b) 沒有偵測到人臉，故沒有輸出笑容量化數值。

4.3 功能性需求

功能性需求	功能性需求描述
即時讀取從業人員影像	利用監視器或攝影機，將從業人員的影像，即時輸入本系統。
偵測從業人員人臉特徵	偵測符合該從業人員的人臉特徵，以便針對個人進行情感檢測。
計算從業人員笑容強度	本系統爲了量化從業人員的情緒表現，計算其人臉特徵的笑容量化數值。
監視從業人員服務態度	透過視覺化的方式顯示笑容量化數值，讓管理者得知，該從業人員在服務態度上的情緒表現。

5 其他非功能性需求

5.1 性能需求

非功能性需求	非功能性需求描述
蒐集與整理數據集	本系統設計的情緒量化數據庫，包含0到10共11個笑容等級，等級越高代表開心程度越高，劃分等級的標準，主要是以眼部及嘴部，做為量化笑容的主要依據。因此我們這組蒐集了每個笑容等級的圖片，使11種笑容等級的圖片能蒐集完整，並建構成可以用來訓練的格式。
設計人臉偵測模組	本系統設計人臉偵測模組，使模組可以偵測輸入影像中，包含人臉的前景位置，作為該從業人員的人臉特徵，且準確率達九成以上。
設計情緒量化模組	本系統設計情緒量化模組，針對從業人員的人臉特徵，計算笑容量化數值，且準確率達九成以上。
檢測結果視覺化	本系統讓管理者可以即時查看從業人員的情緒表現，得到笑容量化數值後，將以視覺化的方式，呈現於監控畫面。

5.2 Safety需求

None.

5.3 Security需求

None.

6 附錄

6.1 量化成果

在量化成果的測試上，爲了證明本系統的實用性，需準確量化個別從業人員的情緒表現，因此，我們總共使用7000張包含人臉的影像，並且利用Precision(P)作爲評估指標，公式如下：

$$Precision = \frac{TP}{TP + FP}$$

其中，TP是指True Positive，FP是指False Positive。透過此公式的計算，將正確量化的結果除以預測量化的結果，得到Precision的數值，如表6.1第二列及第四列所示，Precision的數值越大，表示量化笑容的準確度越高，因此從量化成果得知，本系統能達到良好的預期效果。

表 6.1: 本系統情緒量化準確度的量化成果

笑容量化數值	1	2	3	4	5
Precision	92.64%	90.07%	94.18%	89.80%	92.13%
笑容量化數值	6	7	8	9	10
Precision	88.79%	90.92%	95.48%	93.25%	91.88%