

Boosting 與人臉性別辨識問題

Learning Gender with Boosting

專題組員:謝文哲

指導老師:楊致芳老師

專題編號:PRJ-XXXX-XX-XXX

執行期間:100 年 02 月 至 101 年 01 月

1. 摘要

本專題的第一部份是從 Boosting 的角度比較兩種機器學習的演算法 AdaBoost 和 LogitBoost，討論人臉性別樣本的特性如何影響兩種演算法在性別辨識問題上的表現。我從誤差函數的論點判斷 LogitBoost 比較適合用於性別辨識。以 Color FERET Database 為樣本，建構實驗系統以驗證理論上的探討，我的 LogitBoost 性別分類器獲得了 89.31% 的正確率，高於 AdaBoost 性別分類器的 73.89%。專題的第二部分是結合第一部分在性別辨識問題的經驗，整合人臉辨識，實作一個支援網路攝影機和靜態影像的人臉性別辨識系統。

關鍵詞：Face Detection、Gender Classification、Boosting、Feature Extraction、Machine Learning

2. 簡介

2.1 人臉辨識

人臉辨識是讓電腦在一張任意的影像中判斷人臉可能的位置和大小，並且忽略其它不是人臉的東西像是身體、建築物等。人臉辨識是物件檢測的特例，最著名的演算法是 Viola 等人提出的物件檢測架構[1]，它包含了三個主要的步驟：特徵擷取，用特徵從影像中擷取數據；特徵選取，利用機率統計的方法從數據中自動分析獲得規律，找出對檢測有幫助的數據；Cascade 架構，可以大幅加

速辨識的速度。目前人臉辨識已經廣泛的應用在生物特徵識別，像是數位相機透過人臉辨識以自動對焦，Sony NEX-7、Nikon D7000、Olympus X-43 等機種皆有此功能。

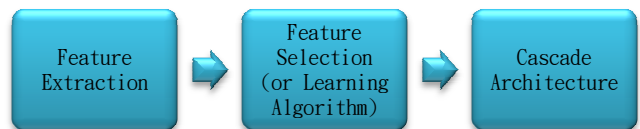


圖 1：Viola-Jones 物件檢測架構

2.2 性別辨識

性別辨識是給定一張人臉，透過過去的經驗計算藉以判斷人臉是男性還是女性。我們解決問題的策略與上述的 Viola-Jones 物件檢測架構類似，但少了 Cascade 架構，因為樣本數量的差異沒有人臉辨識問題來得懸殊。我們採用 Haar-like 特徵做為在特徵擷取階段的工具，我們認為使用 Haar-like 特徵比直接計算像素來得快速，擷取的特徵數目也相對較少。在特徵選取階段我選了兩種機器學習的演算法 AdaBoost 和 LogitBoost，並且比較兩種演算法在性別辨識的表現。



圖 2：性別辨識的問題解決策略

解決性別辨識問題需要兩種類別的樣本：男性和女性，將大部分的樣本做為 Training Set，經過特徵擷取和特徵選取兩個訓練階段可獲得分類器，分類器蒐集了對性別辨識有幫助的資訊，利用剩下的樣本做為 Test Set，

用以評斷分類器的好壞。

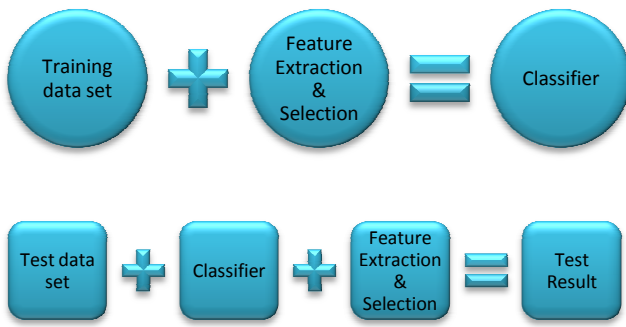


圖 3：訓練程序和測試程序

3. 專題進行方式

3.1 Boosting

Boosting 是一種監督式學習法，在二元分類問題，例如性別辨識，將男性標示為 1，女性為 -1，給定一個樣本集合 $\{(x_1, y_1), \dots, (x_N, y_N)\}$ ，其中 $x_n \in X$, $y_n \in Y = \{-1, 1\}$, $n=1, \dots, N$ ，Boosting 的任務就是建構一個強分類器 $F(x): x \rightarrow \{-1, 1\}$ 使得誤差最小化。強分類器是由一個個循序挑選的弱分類器所組成，只要弱分類器的預測能力大於隨機猜測，增加弱分類器的就可以有效的降低誤差。每組樣本皆有權重 w_n ，由弱分類器的表現調整每一輪的樣本權重，以表示過去的學習經驗。常見的做法是降低分類正確樣本的權重，例如 AdaBoost，使得下一輪在挑選弱分類器時，能夠專注於過去分類錯誤的樣本，然而也有一些 Boosting 提高分類正確的樣本權重，像是 BrownBoost。

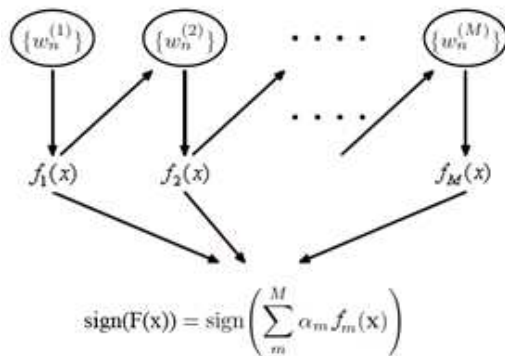


圖 4：Boosting 的基礎架構[4]

3.2 AdaBoost

AdaBoost 的全名為 Adaptive Boosting，強分類器為一群弱分類器的組合， $F(x) = \sum_{m=1}^M \alpha_m f_m(x)$ ，藉由逐次增加弱分類器 $f_m(x)$ ，AdaBoost 能最小化誤差函數 $E(e^{-yF(x)})$ ，這種函數對於樣本分類正確與否呈現指數成長，使得 AdaBoost 對於難以分類的樣本較為敏感，它大幅提高分類錯誤樣本的權重，這使得下一輪挑選的弱分類器能夠學習前幾輪弱分類器難以分類的樣本，這說明了 AdaBoost 的學習方式是專注於經常分類錯誤的樣本，有文獻[5]顯示 AdaBoost 處理 Noisy data 時會產生 Overfitting。

表 1：AdaBoost Algorithm[4]

1. Training set, N samples,
 $S = \{(x_1, y_1), \dots, (x_N, y_N)\}$,
 $x_n \in X, y_n \in Y = \{-1, 1\}, n=1, \dots, N$
 Number of Iteration M
2. Initialize data weighting coefficients $\{w_n\}$ by setting
 $w_n^{(m)} = 1/N, n=1, \dots, N$
3. Construct weak classifier $f \in \{-1, 1\}$
4. for $m = 1$ to M
 - a. Fit a classifier f_m to the training set by minimizing the weighted error function

$$f_m = \operatorname{argmin}_{f_m} \sum_{n=1}^N w_n^{(m)} I(f_m(x_n) \neq y_n)$$

$$I(f_m(x_n) \neq y_n) = \begin{cases} 1 & \text{when } f_m(x_n) \neq y_n \\ 0 & \text{otherwise} \end{cases}$$
 Evaluate the quantities

$$\varepsilon_m = \frac{\sum_{n=1}^N w_n^{(m)} I(f_m(x_n) \neq y_n)}{\sum_{n=1}^N w_n^{(m)}}$$

$$\alpha_m = \ln\left(\frac{1 - \varepsilon_m}{\varepsilon_m}\right)$$
 - b. Update data weighting coefficients

$$w_n^{(m+1)} = w_n^{(m)} \exp(\alpha_m I(f_m(x_n) \neq y_n))$$
5. Make predictions using model

$$\operatorname{sign}[F(x)] = \operatorname{sign}\left[\sum_{m=1}^M \alpha_m f_m(x)\right]$$

3.3 LogitBoost

為了解決 AdaBoost 有 Overfitting 的問題，Friedman 等人提出了 LogitBoost[2]，以統計學的觀點將 AdaBoost 的強分類器看作一個 Additive Logistic Model，
$$F(x) = \sum_{m=1}^M f_m(x) = \frac{1}{2} \log \frac{P(y=1|x)}{P(y=-1|x)}$$
，把 Logistic Regression 的方法套用到 Boosting。相較於 AdaBoost 最小化誤差函數 $E(e^{-yF(x)})$ 以達到優化，LogitBoost 最大化 Bernoulli Log-likelihood $E(l(p(x), y)) = -E \log(1 + e^{-2yF(x)})$ ，此種函數對於樣本的分類正確與否產生線性的變化，對於樣本的分類錯誤較不敏感，所以我認為 LogitBoost 處理 Noisy data 時比較不容易發生 Overfitting。

表 2：LogitBoost Algorithm

1. Training set, N samples,
 $S = \{(x_1, y_1), \dots, (x_N, y_N)\}$,
 $x_n \in X, y_n \in Y = \{-1, 1\}, n=1, \dots, N$
 Number of Iteration M
2. Initialize data weighting
 coefficients $W_n^{(m)} = 1/N, n=1, \dots, N$
 Initialize strong classifier $F(x)=0$
 and probability $P_n = p(y_n^*=1|x_n)=1,$
 $n=1, \dots, N$
3. for $m = 1$ to M
 - a. Compute the weights and working response
 $W_n = P_n[1 - P_n]$
 $Z_n = \frac{y_n^* - P_n}{P_n(1 - P_n)}, \text{ where } y_n^* = \frac{y_n + 1}{2}$
 - b. Fit f_m by a weighted least-squares regression of Z_n to X_n using weights W_n . In our study we use Classification And Regression Tree (CART) to fit the data $\{(x_1, y_1), \dots, (x_N, y_N)\}$ using weights W_n .

c. Update

$$F(x) = F(x) + \frac{1}{2} f_m(x) \text{ and } P_n = \frac{e^{F(x_n)}}{e^{F(x_n)} + e^{-F(x_n)}}$$

4. Make predictions using model

$$\text{sign}[F(x)] = \text{sign}[\sum_{m=1}^M \frac{1}{2} f_m(x)]$$

3.4 特徵擷取

Haar-like 特徵[1]是一種影像遮罩，用以擷取特定區塊的影像資訊，其計算方法是將黑色區的像素和減去白色區的像素和。基本的 Haar-like 特徵有以下五種，若是將黑白的區域對調又可以得到一倍的特徵。影像遮罩可以利用縮放和平移的方式從一張影像上擷取大量的灰階資訊，以一張大小為 20×20 的人臉為例，利用 5 種 Haar-like 特徵可以取得 153120 個不同的數值。Haar-like 特徵可搭配積分影像[1]大幅加速計算的速度。

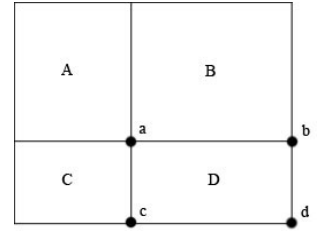
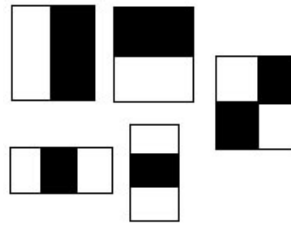


圖 5：五種 Haar-like 特徵

圖 6：區塊 D 的積分影像為 $d - b - c + a$

3.5 AdaBoost 的弱分類器

我採用 Haar-like 特徵做為 AdaBoost 在特徵擷取中的弱分類器，單一的 Haar-like 特徵對許多張不同的影像回答相異的數值。在分類問題中，可以選定一個門檻值將數值分為 1 和 -1 兩種類別，我採用取樣的方法，先將最大與最小的數值拿來當作最大與最小的門檻值，設定一個偏移量從最小的門檻值開始不連續取樣，最後選定一個誤差最小的門檻值。如此一來就能夠使 Haar-like 特徵轉化為弱分類器

$$h(x, f, p, \theta) = \begin{cases} 1 & \text{if } pf(x) < p\theta \\ -1 & \text{otherwise} \end{cases}, \text{ 其中 } F(x)$$

表示特徵值， p (Polarity) 表示黑白是否對

3.6 LogitBoost 的弱分類器

建立一個 CART 最常見的做法是採用

3.7 訓練樣本



3.8 人臉性別辨識系統設計

系統共分為人臉辨識和性別辨識兩個部分：我用 OpenCV 處理人臉辨識和影像處理的工作，性別辨識的工作由 C/C++ 自行寫成，使用 Wxwidgets 建立圖形使用者介面。專案全程採取靜態編譯，我用 CMake 自動 Makefile 產生器編譯整個專案，並由 MinGW 提供 GNU

GCC 與 G++。由於 CMake、Wxwidgets、OpenCV 都支援 Cross Platform，稍微修改程式碼即可在 Linux 下運行。

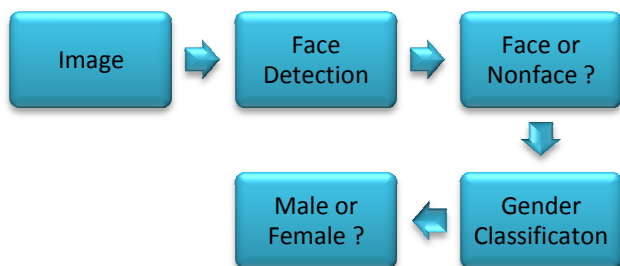


圖 12：系統流程圖

我可以透過一些簡易的演算法加強 OpenCV 人臉辨識系統的辨識率：

- 適當的旋轉影像並蒐集偵測結果，刪除重複偵測的區域，這樣就能偵測一些原本辨識系統抓不到之傾斜度過大的人臉。
- 用人臉膚色刪除一些辨識錯誤的區域，像是看起來像人臉的足球，首先將影像的色彩空間從 RGB 轉換至 YCrCb，膚色定義[7]為 $Cr \in [136, 156]$ 和 $Cb \in [110, 123]$ ，統計膚色所占比例淘汰比例過低的區域。

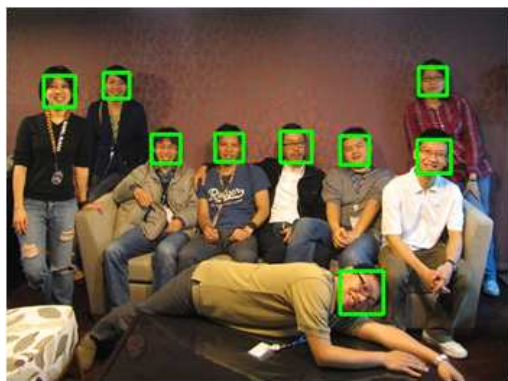


圖 13：偵測傾斜過大的人臉，如右下角的臉

4. 主要成果與評估

4.1 實驗與討論

我做了若干實驗比較 LogitBoost 和 AdaBoost 兩種 Boosting，希望能從實驗中驗證上述理論性的探討。實驗採用 Cross Validation 把實驗樣本分成五組子樣本 A、B、C、D 與 E，任選四組子樣本做為 Training Set 並且將一組剩下的子樣本拿來當作 Test Set，

總共需要做⁽⁵⁾次實驗，並且將得到的五組數據平均。實驗的男女樣本大小為 20*20 的人臉，以 Haar-like 特徵擷取影像數據，使用 153120 個 Haar-like 特徵。AdaBoost 使用單個 Haar-like 特徵做為弱分類器，LogitBoost 採用 Decision stump 做為弱分類器，組成強分類器的弱分類器數目都以 1000 個為上限。由於 Color FERET 資料庫中標記為 fa 與 fb 的 1494 個男性與 914 個女性，其數目的比例並不是 1:1，這裡將男女的比例調至 1:1，平均每個子樣本有 182 個男性和 182 個女性。實驗主要是透過分類器在 Training Set 的表現評估實驗的正確性，藉由在 Test Set 的表現判斷何種 Boosting 所訓練的分類器適合用於性別辨識問題。實驗系統用 C 語言自行寫成。

圖 14 為 LogitBoost 弱分類器數目與辨識正確率之關係的折線圖，Y 軸為辨識正確率，X 軸為弱分類器組成強分類器的數目。上方的曲線為 LogitBoost 分類器對於 Training Set 的實驗結果，正確率隨著弱分類器的數目增加而上升，在分類器數目 100 之後可達到 9 成以上的正確率，這印證了 LogitBoost 能夠有效降低誤差函數值。下方的曲線為分類器對 Test Set 之結果，弱分類器數目為 400 之後趨於平穩，也可達到接近 9 成的正確率。

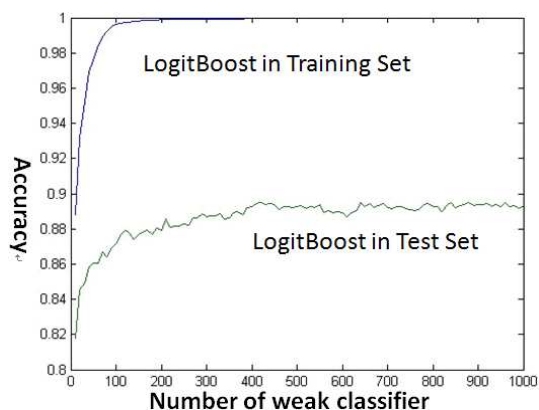


圖 14：LogitBoost 的弱分類器數目與正確率之關係

圖 15 為 AdaBoost 的曲線圖，圖表的形式與 LogitBoost 的曲線圖相同，上方與下方的曲線分別是 AdaBoost 分類器對於 Training Set 和 Test Set 的結果。正確率皆隨著弱分類器的數目增加而上升，表示 AdaBoost 也能有效降低誤差函數值。

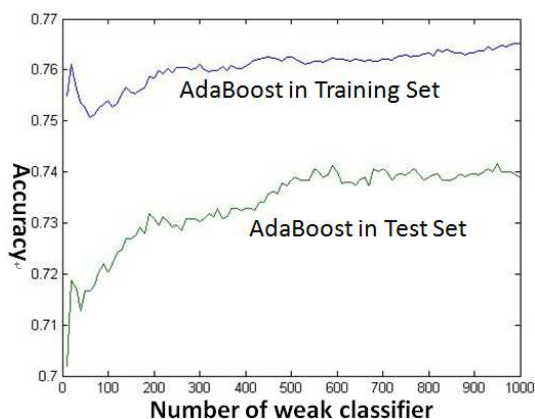


圖 15：AdaBoost 的弱分類器數目與正確率之關係

本專題檢驗了 AdaBoost 的 ϵ_m 和 α_m 值以進一步確認實驗的正確性，圖 16 為 AdaBoost 具有權重的誤差 ϵ_m 與弱分類器數目之關係，呈現上升的趨勢，此數值未超過 0.5，表示弱分類器的辨識效果好於隨機猜測。圖 17 為 AdaBoost 的 α_m 值與弱分類器之數目關係，隨著弱分類器數目的增加有著下降的趨勢， α_m 值的範圍大約在 0.72 至 1.06 間，表示後來選取的弱分類器佔整體的重要性越來越低。

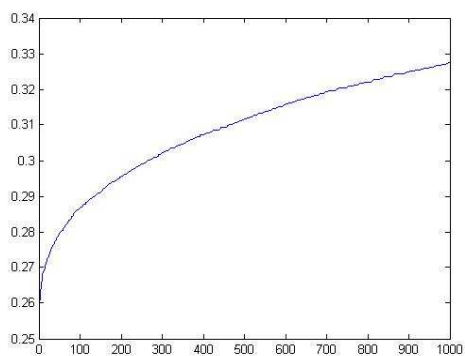


圖 16： ϵ_m 值與弱分類器數目之關係

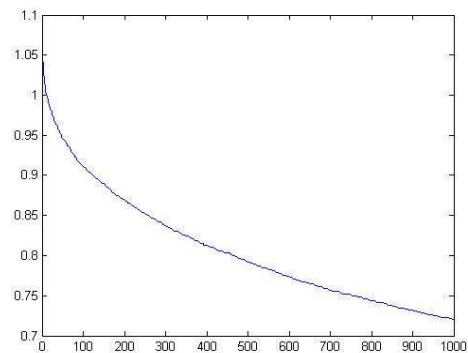


圖 17： α_m 值與弱分類器數目之關係

最後比較 LogitBoost 與 AdaBoost 兩者在 Test Set 的表現，由圖 18 可知 LogitBoost 的辨識正確的比率勝於 AdaBoost，這驗證了前述理論探討中 LogitBoost 比 AdaBoost 更適合於性別辨識問題的論點，我認為區分男性和女性是很容易搞混的，不像人臉辨識問題中人臉和非人臉有著懸殊的差距。最後將實驗的結果整理至表 3。

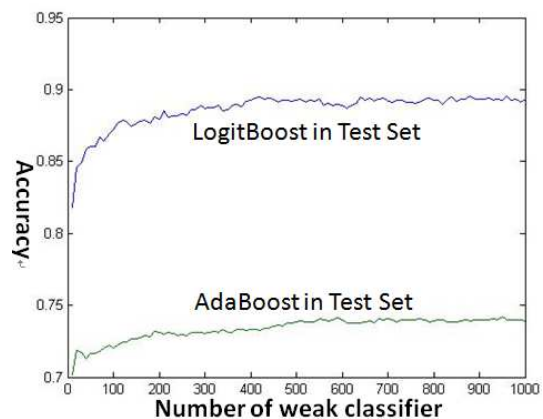


圖 18：LogitBoost 的弱分類器數目與正確率之關係

Classifier	Best Accuracy in Test Set
AdaBoost	73.89%
LogitBoost	89.31%
Sample size: 20 by 20 pixel “Male / Female” is close to 1	

表 3：AdaBoost 和 LogitBoost 在 Test Set 的表現

4.2 人臉性別辨識成果

人臉性別辨識系統的使用者圖形介面共有以下幾個功能：瀏覽檔案、開啟影像、儲存影像、打開攝影機、進行人臉性別辨識、調整參數等，可調製的參數包含膚色比率和偵測重疊比率的門檻，旋轉影像的角度和儲存檔案的格式等。

我用兩種顏色的方框表示區域被判定為人臉，其中，藍色被認定是男性，紅色則是女性。

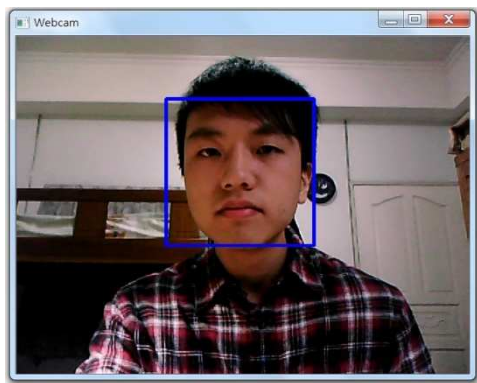


圖 19: 打開網路攝影機作即時人臉性別辨識

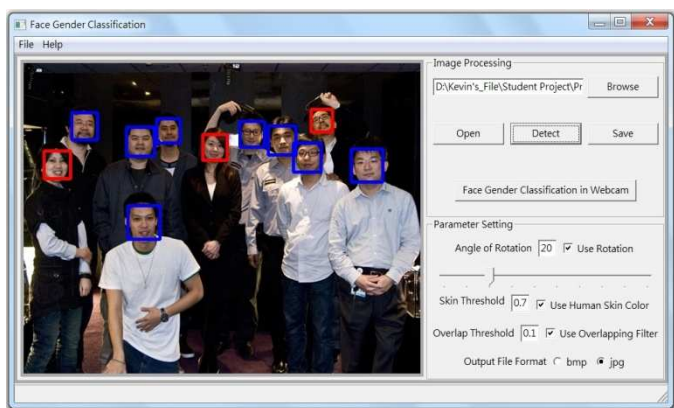


圖 20：系統使用者圖形介面設計

最後，我把現實生活中的團體照和肖像照做一個測試。在實作上有許多因素可能導致性別辨識錯誤，例如人臉辨識的步驟可能抓取一些側臉或是傾斜角度過大的人臉，在性別辨識之前需要對擷取的人臉進行前處理；其次是亮度，這些圖片不像人臉影像資料庫，在拍攝的過程中保持一定的亮度，但我認為直接對擷取的人臉進行 Histogram

Equalization 會破壞影像的資訊，經過實驗沒有進行亮度調整的性別辨識結果較好。

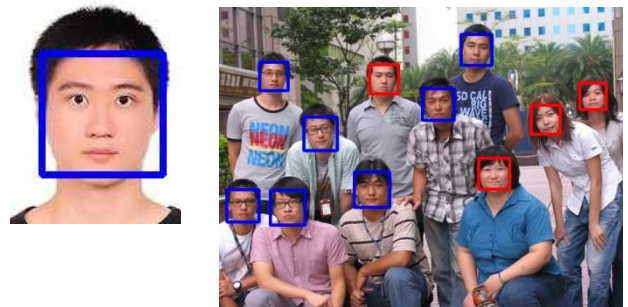
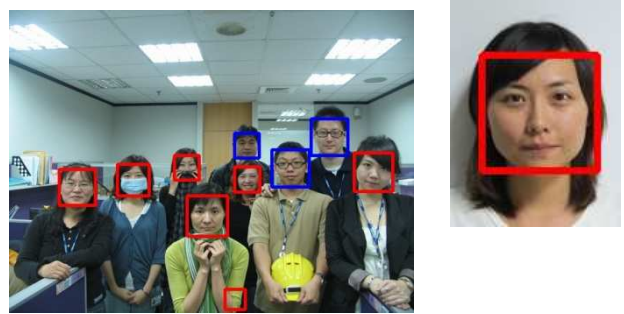


圖 21：人臉性別辨識成果

5. 結語與展望

我從理論和實驗兩個面向提出 LogitBoost 比 AdaBoost 更適合用於性別辨識的結論，就理論的層面，LogitBoost 降低誤差的方式使得它適合用於性別辨識，就實驗的結果，我在 Color FERET Database 獲得了 LogitBoost 的性別辨識正確率高於 AdaBoost 的結論。最後，我實作了人臉性別辨識系統，此系統可以正確的辨識影像或是網路攝影機中的人臉，並且辨認其性別。

物件檢測是一種融合影像處理和電腦視

覺的一種技術，它讓電腦學習去辨識一些特定的物件類別，例如人臉和行人等。物件檢測在圖片檢索和視訊監控有著相當多的研究成果，人臉辨識的技術從十年前 Paul Viola 和 Michael Jones 發表 Viola-Jones 物件檢測架構之後已經逐漸成熟，人臉辨識的蓬勃發展也使得性別辨識成為一個值得研究的議題。

6. 銘謝

本專題能夠順利完成，首先要感謝我的專題指導老師楊致芳老師，給我一個不錯的研究題目好好發揮，在研究過程中不厭其煩的提出許多的建議。也感謝實驗室的學長與同學們，給了我不同的意見。最後感謝系上給我們專題製作的機會，讓我獲得研究學習和程式設計的寶貴經驗。

Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office.

7. 參考文獻

- [1] Viola, Paul and Jones, Michael J. "Robust real-time object detection". Proceedings of the IEEE Workshop on Statistical and Computational Theories of Vision, 2001
- [2] J. Friedman, T. Hastie, R. Tibshirani, "Additive logistic regression: a statistical view of boosting", Technical Report, Stanford University, 1998.
- [3] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, "The FERET Evaluation Methodology for Face Recognition Algorithms, "IEEE Trans. Pattern

Analysis and Machine Intelligence, Vol. 22, pp. 1090-1104, 2000.

[4] Bishop, C.M., Pattern Recognition and Machine Learning. Springer: Singapore, 2006

[5] G RaTsch and T Onoda. "Soft margin for adaboost". Technical Report 021, NeuroCOLT, Rudowe Chaussee 5, 12489 Berlin, Germany, August 1998.

[6] Baluja, S., Rowley, H.: Boosting sex identification performance. In: AAAI-IAAI 2005

[7] H. Li and R. Forchheimer, "Location of face using color cues", in Proc. Picture Coding Symp., Lausanne, Switzerland, Mar. 1993, paper 2.4