

基於模糊類神經技術建構人臉偵測與辨識系統

Face detection and recognition based on fuzzy theory and neural technique

黃佩玲
義守大學
資訊管理學系
m9622015@
isu.edu.tw

李宜靜
義守大學
資訊管理學系
m9622017@
isu.edu.tw

蔡賢亮
義守大學
資訊管理學系
助理教授
jim@isu.edu.tw

蔡佩君
義守大學
資訊管理學系
m9522005@
isu.edu.tw

摘要

本論文旨在提出一套影像中人臉偵測與辨識系統，目的在要能夠精確地定位及辨識較複雜或較多雜訊影像中的人臉位置，利用人工智慧的技術來降低影像複雜度及排除多餘的背景影像，以改善定位及辨識上遭遇雜訊干擾的問題。

我們的人臉偵測與辨識系統上可分成三大部份：一、人臉偵測系統 二、人臉定位系統 三、人臉辨識系統。人臉偵測方面，我們首先提出模糊神經人臉偵測演算法(NFFDA2)，它是個串接式人臉偵測演算法。人臉定位方面，透過精簡快速的高斯擴展法找出人臉區塊的位置。在人臉辨識上，我們結合模糊 C 均值分群法與平行式類神經網路來加速訓練及提高辨識效能。

從實驗數據顯示我們的系統確實可以降低影像的雜訊及複雜度，並縮短偵測與辨識的執行時間，更可以提昇複雜度較高的人臉影像定位及辨識效果。

關鍵字：人臉偵測、人臉辨識、高斯歸屬函數、特徵擷取類神經網路、模糊 C-均值分群法

Abstract

In this paper we present a system which can detect faces in the complex and noisy image. We reduce the complexity of the image in order to remove the redundant background image and improve interference problems that happen in the recognition process.

The face detection and recognition system has three steps: (a) human face detection (b) human face location (c) human face recognition. On the first step, we employ the NFFDA2

human face detection algorithm which consists of fuzzy Gaussian classifier and feature extracting neural network. On the second, we utilize Gaussian spread method to locate the position of the detected human face block. On the final step, we use a fuzzy c-means algorithm combined with a framework of parallel neural networks to recognize the faces.

We reduce the dimension of images so as to decrease the detection and recognition time, and improve the results of human face image recognition.

Keywords: face detection, face recognition, Gaussian membership function, feature based extract neural networks, fuzzy c-means clustering algorithm.

1. 前言

由於利用人工智慧處理圖形辨識的應用越來越蓬勃，在現今要使用電腦來辨識人臉已經不是一件難事，然而大部分的人臉辨識理論，對於輸入的人臉影像在偵測與辨識上有著許多限制，大部分以垂直正向人臉為主來做自動偵測，以提高獲得影像中人臉位置的準確率。但是人臉是一種具有相當細微複雜變化的生物結構，成為在人臉識別上面臨的重要挑戰之一，例如：(1)人臉的外貌、膚色、表情、角度等不同，增加了識別上的可變性；(2)人臉上面的附屬物(鬍鬚、眼鏡)，增加了辨識複雜性；(3)光線亮度的來源不同等考量；若再與多變的環境背景相結合，將會使其誤判率大幅的增加。本研究於人臉偵測與辨識系統可分成三大流程：一、人臉偵測系統，二、人臉定位系統，三、人臉辨識系統。在人臉辨識演算法上，本研究使用的是我們先前所提出的

NFFDA2 是改良於 Henry A. Rowley 於 1998 年提出 neural network-based face detection(NNFD)[1]類神經網路應用於人臉偵測的演算法，人臉定位方面，參考 NNFD 中的 Merging Overlapping Detections，改由透過高斯擴展法找出偵測結果來移除偵測區塊的孤立點，找出人臉的正確位置。而人臉辨識演算法方面，本研究改善由 Jianming 與 Xue 等提出的 FRFCN[7,10]，善用 FCM 分群方式來縮小類神經網路的規模與計算時間。

2. 文獻探討

本章節中，我們簡略介紹 NFFDA2 以及 Rowley 在 NNFD 中提出的合併單一網路部份重疊偵測結果的定位方法，與 Jianming 與 Xue 等提出的 FRFCN[7,10]。

2.1 Neural Fuzzy Face Detecting Algorithm 2 (NFFDA 2)

NFFDA2 的流程為當一個影像區塊經過前置處理後，將 20×20 直方圖等化後的影像轉換成一維的模糊高斯參數向量，進入模糊高斯分類器進行篩選，若初步判定為人臉區塊，便將此區塊的一維向量轉成 10×10 的二維向量，輸入至特徵擷取的類神經網路作最後判定；反之，則直接排除為非人臉區塊，如下圖 1 所示。

2.2 人臉定位演算法分析

在 NNFD 演算法中人臉影像鄰近區域常會出現相符度較低的影像區塊，影像中人臉位置的鄰近區域便容易被重複偵測，因為此現象的發生，NNFD 在經過偵測系統後，會再偵測出金字塔每一層含有人臉的影像區塊，被偵測到的人臉區塊附近亦會有許多重疊的區塊，記錄每個影像區塊的中心點位置，將中心點向 XY 方向擴展，使每個擴展後的中心點產生部分重疊，把計算重疊次數最多的中心點當作人臉區塊的中心位置，代表此區塊為所有鄰近區塊的中心，再經過金字塔多層的比對，找出人臉正確位置。本研究認為若能直接將每個偵測到的人臉區塊當成一個中心位置，與其他鄰近的中心位置做計算，同樣可以達到去除重複偵測的影像區塊的目的，並在次數的計算上改用高斯擴展方式，將偵測為人臉的區塊與其鄰近部分重疊區塊作區分，達到移除偵測區塊的孤立點，包含排除重疊區塊和偵測錯誤的區塊，找出最符合人臉在整張影像中的位置。

2.3 人臉辨識演算法分析

在過去的相關研究上 Jianming 與 Xue 等提出 FRFCN[7,10]，在 FRFCN[7,10] 研究中設計了以模糊 c-均值群集分析和平行式類神經網路為主軸的演算法，將人臉影像輸入較小規模的子網路中，這幾個子網路是由 Fuzzy c-means Clustering(FCM)初步分群後再施以合併後的結果，當平行網路的規模越小越能增進計算的效率，最後比對各平行子網路的結果與計算相似度來判定未知樣本為何類別。

首先，來源影像的大小為 384×243 ，影像需先經過三個前置處理：(1) 將整張影像經過亮度修正後[7]，(2) 利用人臉的對稱性，以鼻子作為臉部的中線，依據比例計算各部位距離以及各特徵區塊的平均灰階強度，框選出人臉，(3) 則利用 Principal Component Analysis(PCA) 計算臉部區塊得到 32×32 的二維特徵向量。再透過 FCM 分群將影像進行分群，可得知每一張影像歸屬於不同群集的程度，而歸屬程度總和為 1，影像歸屬度越大則歸屬於某群集的程度越高，歸類至同一群集的不同影像彼此的相似度較高，透過分群結果可以了解影像特性以及影像間的相似關係。影像經過 FCM 分群後，將分群結果予以合併成數個規模較小的子網路 (Subnet)，每個子網路最多能容納 6 個成員，每個平行網路為一個傳統的倒傳遞類神經網路 (back propagation neural networks, BP)，輸出層神經元皆為 6 個，將未知樣本輸入每個平行子網路進行辨識，辨識過程會經過三步驟計算：

- (1) 於各個子網路進行辨識篩選：每個樣本在每個子網路皆會有其網路輸出值，在此設定 0.5 為通過閾值，將只取大於 0.5 之最大網路輸出值留下，其餘則排除之。
- (2) 於平行式子網路進行辨識篩選：在此步驟將以比較平行式子網路之結果，即若該樣本在其他的子網路已被排除時，則該樣本在此子網路中亦會被排除。
- (3) 計算相似度來判斷之：設定相似度要大於閾值 0.97，最後將可判定此未知樣本其歸類為何人臉圖像。

而本研究採用 FCM 模糊分群方式使不同類別的人臉圖像最少可歸類至兩個群集，再結合平行式的傳統類神經網路進行網路訓練，取其網路輸出的最大值當做最後的辨識結果，便得知該人臉圖像最終的辨識是否正確，透過兩者方法的結合，有效縮小類神經網路運算上的速度，而每個子網路只針對歸屬的成員進行

辨識，縮小辨識資料庫的範圍搜尋，對於辨識效能好且精確度亦能提升。

3. 系統架構

在人臉辨識的訓練上，圖像的來源以及人臉的旋轉角度、五官的位置，甚至是光線強度皆會影響整個系統流程，所以樣本的收集是需要經過挑選的，且在前置處理上也須考量來源

示。其中人臉偵測(即演算法 NFFDA 2)又分為四小部份：一、影像前置處理；二、影像模糊高斯參數向量化；三、模糊高斯分類器；四、特徵擷取類神經網路架構。如圖 1 所示。首先藉由原影像的縮小取樣，將於每一層金字塔影像中擷取出 20×20 的圖像區塊，再經由高斯低通濾波、直方圖等化之影像前置處理後，分別計算圖像其灰階值的模糊歸屬程度分佈。當

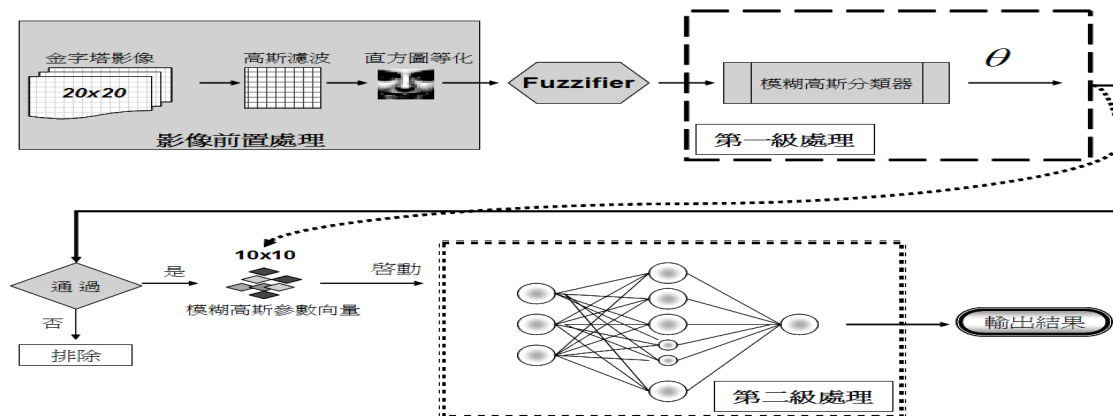


圖 1 NFFDA 2 架構圖

影像的多變性，故改變其影像處理方法。而在本研究中，利用模糊系統的優點，即在設計上不需要精確的數學模型，將此概念運用在圖像識別的數學分析，作為第一級人臉或非人臉的初步判定，以及結合第二級類神經網路訓練，我們修改 NNFD 的網路架構，並將輸入值由原先的 20×20 像素值改為 $Q-1$ 個平均值，決定 20×20 大小區塊是人臉或非人臉圖像。

表 1 系統架構比較

	RFNFD	NNFD	NFFDA
使用人臉資料庫	○	○	○
擷取影像大小	□	□	□
前置處理方式	□	□	✓
類神經網路訓練方式	□	□	□
網路架構層數	□	□	□
隱藏層神經元型態	※	※	✓
輸出神經元個數	□	□	✓
網路輸入值	※	※	✓

○ 使用 ✓ 修改 □ 相同 ※ 不同

本研究之流程包含三大部分：一、人臉偵測；二、人臉定位；三、人臉辨識。如圖 2 所

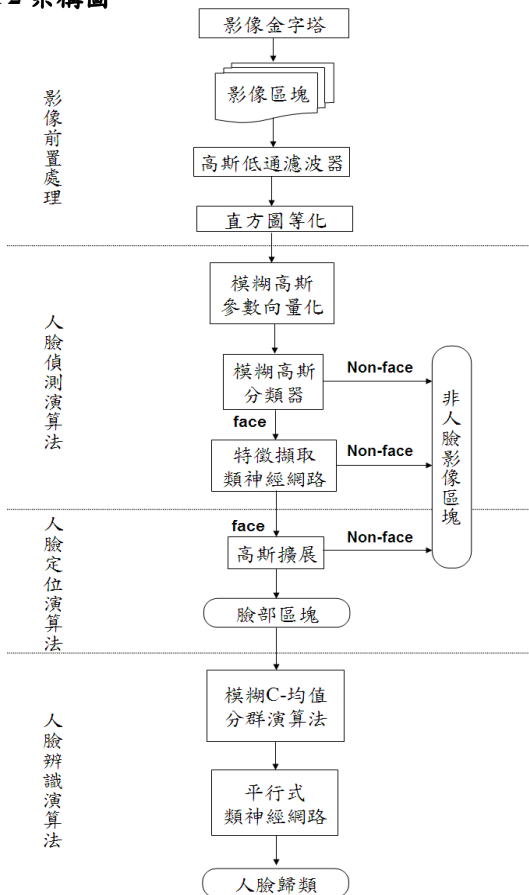


圖 2 流程圖

20x20 的人臉或非人臉影像經過模糊歸屬函數的計算，會產生數個模糊集合(fuzzy sets)，然後再與樣本圖像(prototype)的 Q 個模糊集合計算其均方誤差，並將之累加為一個總值，透過閾值(threshold)初步判定該 20x20 區塊是否包含人臉或非人臉，以上為第一級模糊高斯分

像，故將原始影像分解成數個不同大小的解析度，此即金字塔分解影像技術。金字塔分解影像的處理方式是產生許多低解析度影像，在金字塔影像最下層影像為原始影像尺寸，往上一層解析度越低，其降低解析度方式為每上升一層長寬各減為一半。如圖 4 所示，其金字塔上

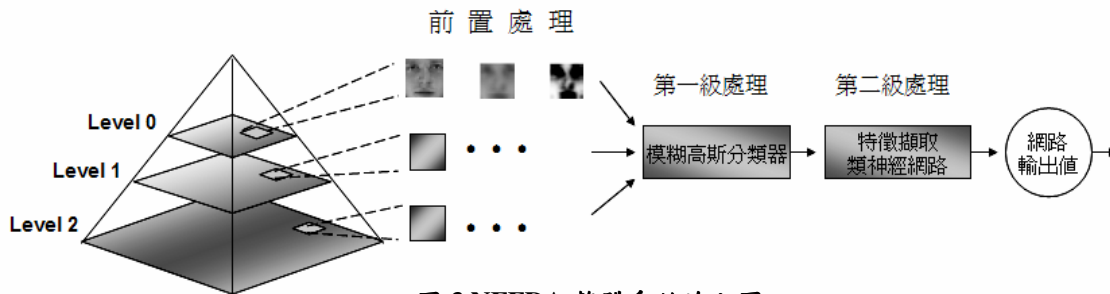


圖 3 NFFDA 整體系統流程圖

類器處理。而通過閾值的圖像區塊，將啟動第二級特徵擷取類神經網路作訓練，其最後結果的輸出值將決定 20x20 之區塊是人臉或非人臉圖像，見圖 3 NFFDA 整體系統流程圖，其中 NFFDA 1 與 NFFDA 2 之系統流程相同，並於以下分章節描述各個流程的主要細節。

層的影響越容易找出臉圖較大者，相反的金字塔下層越容易找到較小之圖像。計算方式為以每 4 個像素做平均得到一個新的像素。換句話說，Level 0 每邊的邊長將會是 Level 1 的一半，如此一來影像取樣率則為二分之一，影像縮減的意義在於，雖然影像中會損失一些較為精細的資訊，但是相對的影像中若存在有移動物體，那麼其在影像上的移動量也相對會減少。

人臉定位部份，是使用高斯擴展法來移除偵測區塊的孤立點，找出人臉正確位置。人臉辨識部份，又分為二小部份：一、FCM 分群法；二、平行式類神經網路。藉由模糊分群將不同類別之人臉影像其相似度高者作最適切的分群，並將群與群間加以合併，以期縮小資料維度，將歸類於同一群集的人臉類別視為類神經網路內的同一個(合併群)子網路(subnet)作訓練，使得當一未知樣本進入網路時，能將圖像正確判定為哪一類別的人臉影像，藉此達到辨識效果。細節將在以下分章節詳細說明。

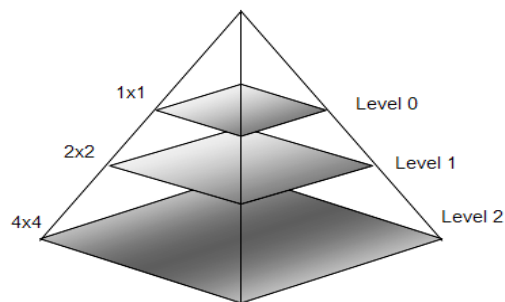


圖 4 影像金字塔

為了確定本演算法的完整性，將於其後的實驗部份比較下列三種演算法的實作結果，如表 2 演算法類型所示。

5. 影像前置處理

表 2 演算法類型

演算法	內容
NFFDA 2	ours 前置處理 → $Q_{10 \times 10}$ → 改 NNFD_BP
NFFDA 1	ours 前置處理 → $P_{20 \times 20}$ → 原 NNFD_BP
NNFD	前置處理 → $P_{20 \times 20}$ → NNFD_BP

用於樣本圖像及訓練類神經網路的人臉影像，是由 ORL 人臉資料庫中擷取出固定區塊集合而得。

4. 金字塔影像

為了能夠找到影像中不同大小的人臉圖

前置影像處理包含灰階轉換及高斯低通濾波，其目的是將原始彩色影像轉成方便後續處理的灰階影像，並加入高斯濾波器(Gaussian filter)，去除影像中較為銳利且可能造成誤判的結果。其後，便將經過高斯濾波處理後的影像作直方圖等化(histogram equalization)，以拉大其影像對比，以利後續計算。如圖 5。

5.1 高斯濾波器

在圖像經過計算之前，需將 20×20 之圖像經過高斯濾波器(Gaussian filter)處理使其影像模糊，此為一種低通濾波器，是由高斯機率分佈函數變化而來，利用相鄰數個像素的加權來產生新的像素，藉此可以產生一組新的影像，見公式(1)。將實驗之影像經過高斯濾波器處理，可以改善圖像銳利的情形，進而方便於後續計算均方誤差之步驟，以降低人臉圖像與樣本圖像(prototype)之間的誤差總值。

$$f(x, y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

$f(x, y)$ 為經過高斯函數計算之值； σ 為標準差， σ 值越大，高斯曲線越平緩，反之則越陡峭。

5.2 直方圖等化

灰階影像的直方圖是灰階度的一種統計圖，直方圖等化法是將原影像中所有像素的灰階值加以平均並分配到各個像素點，使灰階度的直方圖盡可能呈現均勻分佈，如此可提高影像的對比度，得到較清晰的影像。

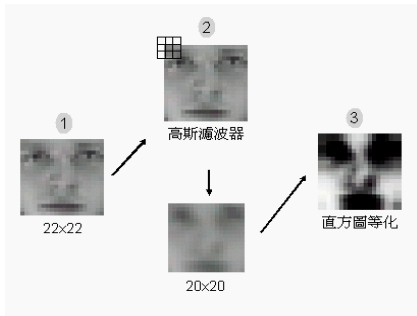


圖 5 影像前置處理

6. 影像模糊高斯參數向量化

「模糊集合」是傳統的明確集合的一種推廣，因其允許模糊集合中的元素，對此集合的歸屬程度是界於 0 到 1 之間的任意值。其模糊化的好處可以提供更佳的推廣性、容錯性、以及生活上各領域的應用。在本研究中，將影像模糊集合向量化就有助於圖像識別的量化分析，了解 20×20 的圖像像素在其模糊集合中的灰階值歸屬程度(membership degree)。

6.1 模糊集合之基本型態

明確集合 A 之特性函數 $\Phi_A(x)$ 非 1 即 0，其中 $x \in U$ ， U 為字集合。而現在我們使一個集合 A ，他的特性函數 $\Phi_A(x)$ 是介於 0 與 1 之間，表示為 $[0,1]$ ，即代表 x 屬於集合 A 之程度(degree)有大小之分。當 $\Phi_A(x_1) > \Phi_A(x_2)$ ，表示其 x_1 屬於 A 的程度比 x_2 屬於 A 之程度大。

所以這個集合 A 便是一個不明確的元素隸屬關係，這樣的集合便稱它為「模糊集合」(fuzzy sets)，表示為 A ，而其特性函數通常被稱之為「歸屬函數」(membership function)，表示法為 $\mu_A(x)$ 或 $A(x)$ ，其值定義在實軸閉區間 $[0,1]$ 之中，可為任一實數值，如下公式(2)表示。

$$\mu_A : U \rightarrow [0,1], \quad (\text{亦即 } 0 \leq \mu_A(x) \leq 1, x \in U) \quad (2)$$

6.2 高斯歸屬函數

一個模糊集合的歸屬函數的定義可以有許多方式，只要是函數值都是位於 $[0,1]$ 區間內的函數，都可成為歸屬函數[3,5]。本研究採用高斯函數的歸屬函數(Gaussian membership function)，其定義式如下公式(3)所示，形狀為吊鐘形，其中 x 為自行定義之值； m 為吊鐘形的中間點； σ 為標準差，代表吊鐘形的寬度，見圖 6 高斯歸屬函數圖形所示。

$$\mu_A(x) = \exp\left(\frac{-(x-m)^2}{2\sigma^2}\right) \quad (3)$$

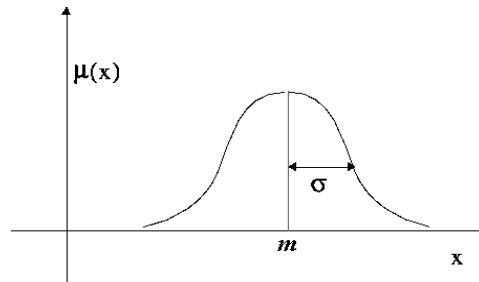


圖 6 高斯歸屬函數圖形

經過直方圖等化處理的圖像樣本(prototype)，將其等化後的灰階值作正規化(normalization)至 $[0,1]$ 之間，並將每個像素位置予以正規化，以本研究 20×20 大小的圖像，其位置個數為 $N=400$ ，並以 4 單位為一區間，予以正規化後，經由模糊歸屬函數計算可得 Q 個層級的值，其即為 Q 個模糊集合(fuzzy sets)，在此將實驗圖像利用高斯函數的歸屬函數計算，可得知在不同模糊集合中，出現於其中像素灰階值的平均值，即歸屬程度(membership degree)。模糊歸屬函數計算公式如(4)所示：

$$\mu_{(i)} = \frac{1}{N} \sum_{k=1}^N f_i(k) \cdot x_k, \quad i=1,2,\dots,Q, \quad (4)$$

其中， x_k 為樣本圖像第 k 個像素值正規化後之

值； $f_i(k)$ 為第 k 個位置之第 i 層級歸屬函數值； μ_i 為每個層級的平均值， $i=1,2,\dots,Q$ ，共 Q 個層級，其圖 7、8 為樣本圖像以及非人臉圖像經由模糊化計算後繪製出的折線圖，橫軸為 101 個 Q 層級之值；縱軸為其歸屬度。

7. 模糊高斯分類器

影像經過模糊計算後會得到 Q 個值，我們將經過模糊化計算的任意 20×20 灰階值圖像與樣本圖像作均方誤差之計算，並累加其總值來作為初步判定人臉或非人臉之依據。

7.1 均方誤差之累加

20×20 之人臉或非人臉圖像經過影像模糊計算會得到 $Q=101$ 個值，將每個區間的 Q 值與樣本圖像之 Q 值作均方誤差計算，見公式(5)並累加為一總值，如以下公式(6)所示。由其誤差總值可知其誤差越小，則有很高機率可能為人臉，然則反之。

$$e_n = (Q_n - Q_m)^2, n = 1, 2, \dots, Q, \quad (5)$$

$$E_n = \sum_1^n (e_n + e_{n-1}), \quad (6)$$

其中， Q_n 為樣本圖像之模糊集合值； Q_m 為任一圖像之模糊集合值；而 m 由 $1, 2, \dots, Q$ 。 E_n 為 e_n 個誤差累加的總值，此 E_n 值將會經由下述的閾值來作判定。

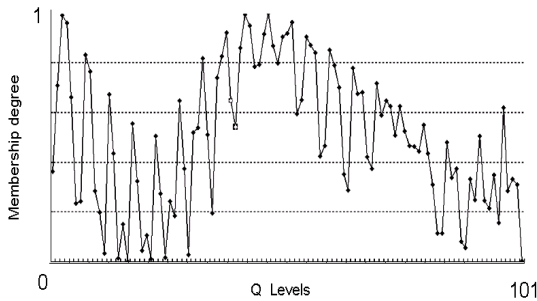


圖 7 [人臉] 影像模糊化集合分佈圖

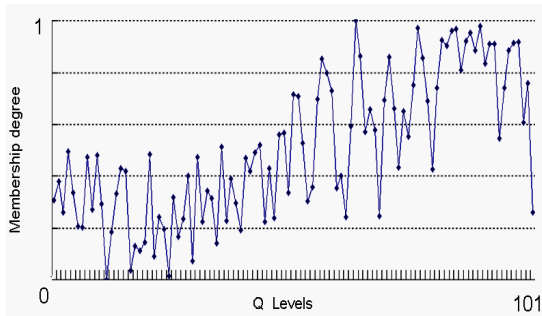


圖 8 [非人臉] 影像模糊化集合分佈圖

7.2 閾值設定

統計共 105 張人臉及非人臉圖像的差值分佈圖，見圖 9，發現其誤差值越小，是人臉的可能性越高，由分佈圖來看，人臉圖像多半集中在誤差總值為 14 以內，而少數幾張圖是落在 14 以外附近的值，為了達到人臉圖像皆通過以及非人臉圖像通過數量越少越好的目的，將閾值(threshold)設定為 14，作為初步篩選的限制值。統計表列出人臉與非人臉之誤差總值，發現其人臉圖像若有其雜訊(如戴眼鏡、鬍子、頭髮、臉頰旁的背景等)，皆能判定正確；而與人臉相似性高的非人臉圖像在此級分類便有可能歸類為人臉圖像。見表 3 模糊高斯分類器之結果。

表 3 模糊高斯分類器之結果

臉				
誤差	7.135953	7.610741	8.281150	6.159216
非臉				
誤差	10.59017	14.86791	15.94383	14.4515

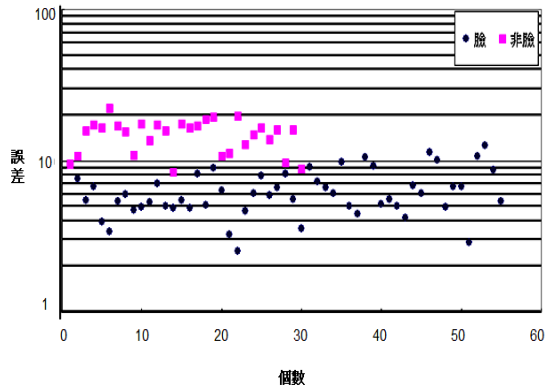


圖 9 差值分佈圖

8. 特徵擷取類神經網路架構

倒傳遞類神經網路(back propagation neural networks)為監督式學習(supervised learning)，屬於多層前饋式網路(multilayer feedforward networks)，採用誤差倒傳修正法則，不斷往回修正網路中的傳遞權重值與偏權值，以降低誤差值，達至期望目標值。

本研究針對第二級處理採用完全連結(full connected)的倒傳遞修正演算法，進行人臉與非人臉的訓練，本篇承襲 NNFD 之類神經網路架構，修正由 F. C.-H Rhee 等人[4]提出的

region based fuzzy neural networks for face detection(RFNFD)演算法為本研究適用之網路架構，改變其輸入值為模糊化後的 $Q-1$ 個平均值，並以 10×10 的二維陣列形式輸入；且針對輸入樣本修正原網路架構，使得樣本像素在結構中有部分重疊(overlap)。以串接式並用前級處理與計算，且在隱藏層的計算中，樣本像素輸入值會有些許的部分重疊，期望使辨識效果更加精確。的確，本研究在後級處理之類神經網路訓練結果，對於不同角度的人臉能有更好的辨識效果。

8.1 修改 RFNFD 架構

參考 RFNFD 處理人臉影像模糊化的方式，先將訓練資料予以平均得一張樣本圖像(prototype)，然而計算此圖像在模糊集合中的平均值，會得到 Q_n 個值，在本研究可得 101 個值，並直接依據此圖像算得的值作為第一級的誤差分類的樣本。而為了使任一人臉與人臉樣本圖像間的誤差縮小，以及使非人臉與人臉樣本圖像間的誤差拉大，我們在前置處理時會先經過高斯濾波器，再作直方圖等化，此步驟對於前後級的處理都有相當的助益；再者，運用在其後的類神經運算，我們將最後一值認定為雜訊，並且去除之，將這 100 個值以二維矩陣方式輸入至神經網路內作訓練。見表 4。

表 4 樣本圖像之像素向量表

0.486167	0.871841	0.990715	0.936828	0.723495	0.507178	0.456018	0.724629	0.660179	0.382201
0.36735	0.277661	0.526028	0.443771	0.134039	0.18959	0.202613	0.437013	0.373367	0.168435
0.214336	0.149888	0.395325	0.2864	0.088155	0.312575	0.287453	0.523198	0.413402	0.113188
0.491702	0.651517	0.741338	0.613437	0.321912	0.673282	0.904555	0.910042	0.785949	0.619623
0.801121	1	0.969163	0.888385	0.801507	0.89324	0.98505	0.938061	0.884934	0.844718
0.909359	0.928603	0.766425	0.768864	0.84936	0.891377	0.839976	0.542237	0.520927	0.74821
0.83262	0.760989	0.419063	0.387453	0.645565	0.736077	0.724463	0.48129	0.406392	0.559511
0.631123	0.661093	0.52434	0.414486	0.475385	0.519305	0.434592	0.274745	0.226358	0.347108
0.441144	0.293346	0.033907	0.033589	0.194009	0.345522	0.386074	0.152923	0.089179	0.169647
0.260504	0.487161	0.329134	0.236597	0.206121	0.16506	0.511695	0.353605	0.234623	0.149971

8.2 修改 NNFD 架構

鑑於 NNFD 架構之優點，本研究將網路架構修正為一層輸入層、一層隱藏層、一層輸出層，其中輸入值為 10×10 影像模糊化後的值；而隱藏層同原架構之神經元型態，在本架構中改以 $F1$ 、 $F2$ 、 $F3$ 代稱； $F1$ 為 4 個 5×5 的子區塊； $F2$ 為 16 個 4×4 有部分重疊的子區塊； $F3$ 為 8 個 3×10 長條狀有部分重疊的子區塊。見圖 10 特徵擷取類神經網路架構圖。

該隱藏層之設計與原架構有相同之優點，由於前級處理的方式有異，而改變其輸入值，不但減少計算量、訓練時間，也使得訓練效果

有所提升，提升辨識率。

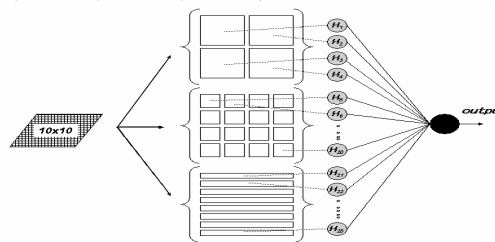


圖 10 特徵擷取類神經網路架構圖

8.3 切割區塊之決定

考量其樣本型態來決定切割區塊大小。我們使用 10×10 大小的輸入樣本，為了加強臉部特徵(眼睛、臉頰、嘴巴)的訓練及辨識度，故分割適切大小來作為隱藏層的神經元輸入值。我們將修改後的隱藏層三種型態分別稱為 $F1$ 、 $F2$ 、 $F3$ 。 $F1$ 型態為 4 個 5×5 的子區塊，訓練著重在眼睛特徵上； $F2$ 型態為 16 個 4×4 有部分重疊的子區塊，加強對於眼睛、臉頰、鼻子、嘴巴等特徵的訓練； $F1$ 及 $F2$ 型態的神經元與原架構切割的型態一致， $F3$ 型態為 8 個 3×10 有部份重疊的子區塊，在此設計與原架構一樣同為長條狀，特別針對眼睛及嘴巴的特徵部位作訓練。

下述以 $F3$ 型態之區塊，做實例說明該 10×10 樣本圖像其輸入樣本之特徵區塊分

佈：依據樣本圖像分佈圖來看，繪製每欄的走向可得 10 條折線(橫軸為欄位數，縱軸為其歸屬度)，顯示出一張正向人臉的縱向分佈大致上是相符且對稱的，並將每個落點作平均可得一張樣本圖像平均分布圖，此圖顯示了一個高峰與兩個谷峰，高峰代表臉頰及鼻樑，谷峰分別代表眼睛和嘴巴，可看到各特稱部位的欄位落點為(1~3)、(4~6)、(8~10)，皆為等距範圍，如圖 13，且考量其各臉部角度不同而造成位移或偏斜，便設計等距範圍的像素部分重疊，故 $F3$ 型態區塊的切割便為 8 塊，各有 2 像素單位的重疊，如圖 9、10 所示。 10×10 樣本請

見表 4。

9. 高斯擴展

使用高斯擴展法來做人臉定位的想法是來自於空間濾波器(filter)，而濾波器的作法是透過對指定像素以及其鄰域部份進行相對位置的加乘運算，而產生一個新的影像數值，然而，影響數值變化程度的因素便是不同的特定函數。因此，高斯函數可依據標準差參數的設定得到不同程度的效果，標準差越小其數值變化較極端，越大則變化較小；為使其高斯函數適用於本研究的定位計算，將設計一個二維高斯分佈，用於特徵類神經偵測後的網路輸出結果，如圖 11 所示，遮罩(mask)覆蓋於指定的區塊位置中由左而右、由上至下，一次平移一個位置，於遮罩下的相對位置做相加計算。

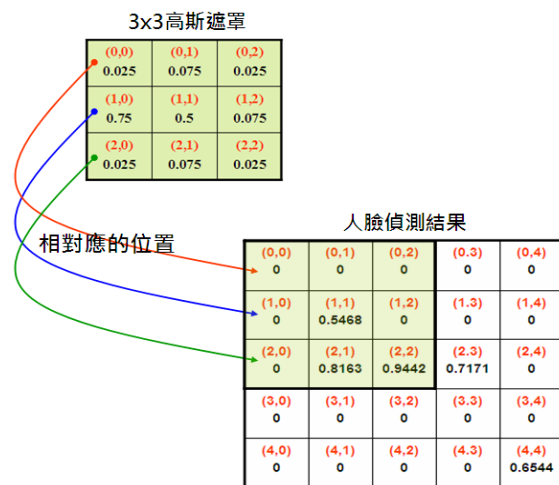


圖 11 高斯遮罩對應偵測結果表

完成計算後的偵測結果表中，每個座標位置皆有非零的數值，在本研究中設定人臉定位閾值為 1.8，所有數值超過 1.8 的便留下，表示為人臉的位置區塊所在。

10. 模糊 c-均值分群法(FCM)

10.1 模糊 c-均值分群法(FCM)

FCM 加入了模糊的概念，資料點 x 不再絕對地屬於任何群聚，而是以一個介於 0~1 之間的數字來表示 x 隸屬於某個群聚的程度。假設現在將要進行叢集分析的圖像像素數為 N ，圖像像素集合表示為 $X(x_1, x_2, \dots, x_N)$ ，其中的 k 組資料包含 P 波段點， $x_k = \{x_k^1, x_k^2, \dots, x_k^p\}^T$ ；假設將圖像像素分群為 C 個類別，則每個類別的叢集中心

$v_i = (v_i^1, v_i^2, \dots, v_i^p)$ ，叢集中心集合 $V = \{v_1, v_2, \dots, v_c\}$ ，如此，我們便可以以一個 $C \times N$ 的矩陣 U 來表示每個資料點 x_k 隸屬於以 v_i 為中心的叢集 i 的歸屬程度，定義歸屬度矩陣 U 為：

$$U = [\mu_{ik}]_{C \times N}; \quad (7)$$

矩陣 U 中每一列的元素表明所對應的像素歸屬於 C 個類別中各個叢集的歸屬度。滿足以下的限制條件：

$$\mu_{ik} \in [0, 1], \quad 1 \leq i \leq c, \quad 1 \leq k \leq N; \quad (8)$$

$$\sum_{i=1}^c \mu_{ik} = 1, \quad 1 \leq k \leq N; \quad (9)$$

$$0 < \sum_{k=1}^N \mu_{ik} < N, \quad 1 \leq i \leq c; \quad (10)$$

倘若我們針對其中每一點 x_k 而言，它隸屬於各個群聚的程度總和應該正好等於 1；且在同一個叢集 i 中，每一資料點 x_k 之程度總和大於 0 且小於 N 。

根據矩陣 U ，我們可以定義出我們的目標函數 J (objective function)：

$$J(U, C_1, C_2, \dots, C_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{k=1}^N (\mu_{ik})^m \text{dist}(v_i, x_k)^2 \quad (11)$$

m 為權重係數，可以是介於 $[1, \infty)$ 之間的任何數，此參數又稱為平滑因子，控制著模式在模糊聚類間的分享程度。 $\text{dist}(v_i, x_j)$ 是 v_i 與 x_j 之間的距離函數 (v_i 代表叢集 i 的中心， x_j 為樣本點)，一般皆直接採用歐幾里得距離。由於目標函數 J 包含兩個參數 (U, V)，故依據拉普拉斯乘數法進行最佳化計算時，針對 (U, V) 進行交替迭代最佳化計算。

FCM 演算法主要有 6 步驟，如下說明：

- (1) 確定叢集數 c ，加權指數 m ，終止誤差 ϵ 或設定目標值越趨於 0，最大迭代次數 LOOP。
- (2) 隨機初始歸屬度矩陣 U ，使此矩陣滿足所有限制條件。
- (3) 計算所有叢集中心 v_i ：

$$v_i = \left(\frac{\sum_{k=1}^N (\mu_{ik})^m \cdot x_k}{\sum_{k=1}^N (\mu_{ik})^m} \right), \quad i = 1, 2, \dots, c$$

(4) 計算目標函數 J_{new}^1 。

(5) 計算新的矩陣 U ：

$$\mu_{ik} = \left[\sum_{i=1}^c \left(\frac{d(v_i, x_k)}{d(v_j, x_k)} \right)^{\frac{2}{m-1}} \right]^{-1}, \quad x_k \neq v_i, \quad i=1, 2, \dots, c$$

(6) 重複步驟 (2) 至步驟 (5)，直到誤差 ϵ 小於一個閾值或當 $(J_{new}^{n+1} - J_{new}^n)$ 極趨近於 0 時，即停止演算法。

10.2 FCM 解模糊化

為了找出所有資料點 x_k 歸屬於叢集 i 之明確的歸屬程度值，便需將矩陣 U 進行解模糊化計算，以找出每個圖像其在群聚內最具代表性的歸屬度。將歸屬值大於閾值者留下，其餘將刪除其值並設為 0，在本研究中，人臉影像共 20 個類別，每個類別 5 張並依序編號為 NO.1~NO.100，共 100 張的人臉作為影像分群，故於解模糊化的過程中分為四步驟：

(1) 對橫向歸屬值解模糊化：

首先，希望每個資料組至少能歸屬兩個群集，故設定閾值為 $\theta = N \times \frac{2}{C} \times 100\%$ ，把小於閾值 θ 之歸屬值予以刪除並設為 0，再於每個橫向值選定前六大之歸屬值留下，將其餘值刪除並設定為 0。

(2) 對縱向歸屬值解模糊化：

完成步驟 (1) 後，於每個縱向選定前十二大之歸屬值留下，將其餘值刪除並設為 0，完成此步驟後將可得到一完整解模糊化後之明確值，便可將其值歸類至步驟 (3) 之分群結果。

(3) 產生分群結果：

將解模糊化後有其明確值的圖像歸屬至最後分群出的 C 個叢集中 (在本研究 $C = 10$ ，即 $C1, C2, \dots, C10$)，使其最大群數為六個，以及最大成員數為十二個，即表示每個資料組最多可歸屬於六個叢集，而每個叢集最大容納的成員為十二個資料組，而為了了解所有人臉影像的類別分群情況，如表 5，於後步驟

(4) 說明 (類別) 子網路如何的產生與呈現方式。

(4) 產生 (類別) 子網路：

由於將每張人臉影像依序以影像編號的方式編列，故在分群的結果中見得的分群情形殊不知為何類別的人，所以在此步驟便將各個叢集的類別歸屬計算出來，在計算後，當有相同類別編號在同一群集中出現，僅留其一類

別編號即可，由表 6 可知其各個影像編號所屬之類別的分群情況，如此便可初步的將經由 FCM 分群後的叢集 ($G1, G2, \dots, G10$) 視為類神經網路訓練的 (類別) 子網路 (Subnet1, Subnet2, ..., Subnet10)，如表 7；最後整理每個子網路是由哪些群集所組成的，如表 8 所示。

表 5 (編號)初步分群結果

群	編號	筆數
C1	26 41 42 43 45 46 47 49 50 79 83 90	12 筆
C2	11 12 13 23 32 48 49 59 68 84 96 97	12 筆
C3	20 48 56 68 72 75 86 87 89 96 97 98	12 筆
C4	1 2 3 6 7 8 9 10 16 17 39 73	12 筆
C5	15 21 24 25 44 57 58 60 63 64 65 88	12 筆
C6	3 16 19 20 32 33 51 52 53 54 55 99	12 筆
C7	4 25 29 30 35 49 76 77 81 82 85 100	12 筆
C8	16 17 18 19 71 73 74 76 80 91 92 93	12 筆
C9	1 2 3 5 36 37 38 39 40 73 94 95	12 筆
C10	26 27 28 61 62 65 66 67 68 69 70 84	12 筆

表 6 (類別子網路)初步分群結果

群	類別編號	筆數
G1	6 9 10 16 17 18	6 筆
G2	3 5 7 10 12 14 17 20	8 筆
G3	4 10 12 14 15 18 20	7 筆
G4	1 2 4 8 15	5 筆
G5	3 5 9 12 13 18	6 筆
G6	1 4 7 11 20	5 筆
G7	1 5 6 7 10 16 17 20	8 筆
G8	4 15 16 19	4 筆
G9	1 8 15 19	4 筆
G10	6 13 14 17	4 筆

表 7 最後合併子網路之結果

Subnet	成員(類別編號)	筆數
Subnet1	6 9 10 16 17 18 3 5 7 12 14 20	12 類
Subnet2	6 13 14 17 4 10 12 15 18 20 16 19	12 類
Subnet3	1 8 15 19 2 4 3 5 9 12 13 18	12 類
Subnet4	1 5 6 7 10 16 17 20 4 11	10 類

表 8 整理子網路之合併群

Subnet	群 (成員)
Subnet1	G1 G2
Subnet2	G3 G8 G10
Subnet3	G4 G5 G9
Subnet4	G6 G7

11. 平行式類神經網路

本研究在此使用的平行式類神經網路架構(pNNs)為一倒傳遞類神經網路，作為每個圖像經由 FCM 計算後進入測試的歸屬子網路。

此網路架構有三層網路層，分別為一層輸入層，其輸入值為 10×10 之模糊高斯參數向量樣本；一層隱藏層與一層輸出層，其每個子網路之輸出層的輸出神經元依據各子網路之成員數而定，期望值以 1 及 0 代表圖像辨識結果是否為該所屬之圖像類別。網路架構圖如圖 12 所示。

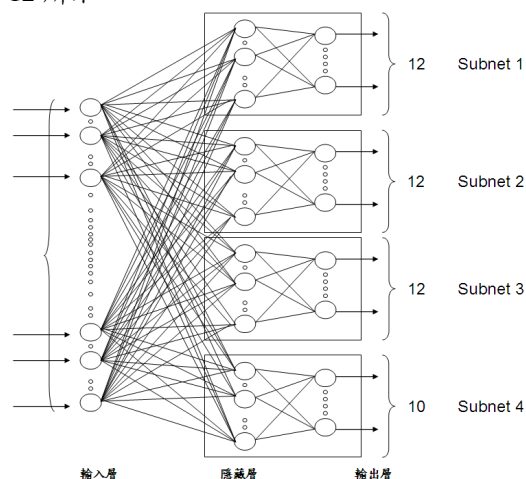


圖 12 平行式類神經網路架構(pNNs)

(1) 各子網路之訓練過程：

將其樣本由輸入層進入到子網路中作訓練，其經由 FCM 替 20 個類別的人臉影像進行計算，每一樣本會依據最終合併之分類結果進入至所屬的子網路進行訓練，而最終輸出結果。在神經訓練上為求辨識結果的公平性，於各子網路皆採用 fivefold cross validation 的方式進行訓練與測試，例如：以某一個類別說明之。類別 S1 包含 10 張不同角度及表情的影像，分別為編號 0101,0102,...,0109,,0110，則第一次訓練為 0101,0102,...,0107，第一次測試為 0109,0110；第二次訓練為 0101,0102,...,0106,0109,0110，第二次測試為 0107,0108；以此類推，直至每張圖像皆完全進行過訓練與測試階段為止。

(2) 於兩個子網路進行辨識篩選：

每個未知樣本在經過 FCM 分類後會有所歸屬之子網路，在本研究中採取二順位之子網路歸屬的方式，使得每個未知樣本皆會進入兩個平行網路做辨識，在所屬之兩個子網路中皆有其網路輸出值，在此以網路輸出之最大值作為判定結果，即可代表其辨識結果將得以瞭解該未知樣本歸類至何種類別。

12. 實驗

12.1 實驗樣本類型

本研究之實驗樣本含有人臉與非人臉影像，共 105 張，其影像隨機擷取 20×20 之影像大小擷取自 ORL 人臉影像資料庫有多種不同人臉分別各 5 張影像，及擷取自網際網路上數張，其包含有旋轉、傾斜以及略帶表情的影像；非人臉影像亦是隨機擷取 20×20 之影像大小，多為生活照片之影像。本實驗之圖像皆採用灰階像素值，作為影像前置處理以及計算、訓練測試之用。

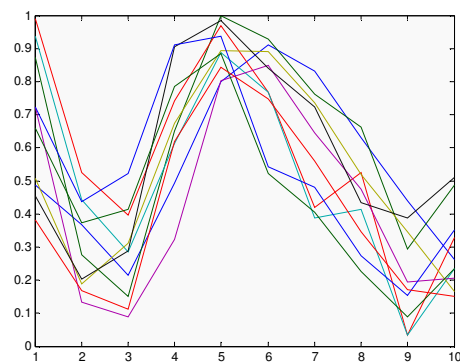


圖 13 樣本圖像分佈圖

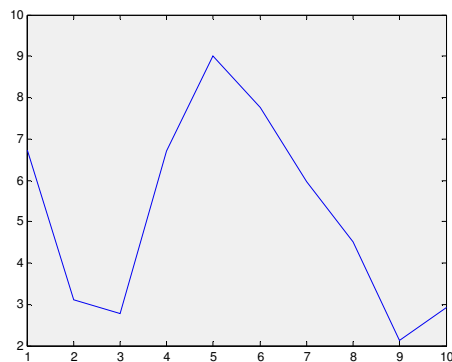


圖 14 樣本圖像平均分佈圖

12.1.1 人臉訓練樣本

由於人臉影像有其豐富性，而使得增加其在偵測與辨識上的複雜度，為了儘可能排除後

續訓練上的難度，使用在本研究上之影像包含有旋轉、傾斜以及略帶表情等影像，由人臉影像資料庫擷取 25 張人臉圖像作為訓練之用，以及擷取自生活照之人臉影像數張不等為測試之用，如圖 15 所示並說明類型如下表 9 圖像類型所示：



圖 15 人臉訓練樣本

表 9 圖像類型

性別	類別
男	無旋轉、無傾斜、無表情
	旋轉、無表情
	左傾斜、略有表情
	右傾斜、略有表情
女	無旋轉、無傾斜、無表情
	旋轉、略帶微笑
	旋轉、左傾斜、略帶微笑
	旋轉、右傾斜、略帶微笑
共 25 張	

12.1.2 非人臉訓練樣本

在有包含人臉影像的生活照片裡，亦含有許多影響人臉偵測與辨識結果的非人臉圖像，藉於此因，在倒傳遞類神經的訓練過程中必須加入非人臉影像，以增加辨識效果，提升人臉影像之辨識度。本研究以隨機方式擷取生活照中 20×20 部包含人臉之圖像大小，25 張做為訓練之用，以及數張不等測試之用圖像，如下圖 16 所例：



圖 16 非人臉訓練樣本

12.1.3 樣本圖像(prototype)

將 50 張由人臉影像資料庫擷取出的 20×20 影像區塊經加總平均而得，為一圖像樣本(prototype)，作為第一級人臉及非人臉初步篩選之依據。如下圖 17 所示：



圖 17 樣本圖像

12.1.4 人臉與非人臉之測試樣本

人臉測試樣本類型分為兩種：來自 ORL 人臉資料庫不與訓練樣本重複之影像，以及擷取自網際網路上的生活照片。非人臉測試樣本種類繁多，生活照片上任意 20×20 大小的圖

像區塊皆包括於此。本演算法採用之圖片與 NNFD 處理方式不同，如下圖 18 所示。

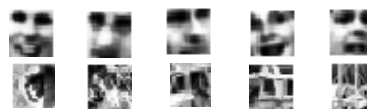


圖 18 樣本圖像

12.2 實作分析與探討

先針對不同影像處理方式以及影像模糊化中標準差 α 之設定值不同，個別作均方誤差之加總，發現其經過高斯濾波處理的圖像，確實對於人臉與非人臉的區分效果較佳。接下來將陳述三種不同演算法其實驗之結果，將以人臉測試樣本與非人臉測試樣本分別呈現，而人臉測試樣本又分為以 ORL 人臉資料庫為主，以及以生活照為主之測試樣本，最後經網路架構輸出之結果以網路輸出值為 0.5 以上，便認定為人臉圖像。此三種演算法分別為 NFFDA 1、NFFDA 2、NNFD，而 NFFDA 1 亦為串列式的系統，當經過模糊高斯分類器後才啟動類神經訓練系統，輸入樣本為直方圖等化處理後的灰階值影像；NFFDA 2 為前章節所述之系統；NNFD[9]之系統如圖 3。結果發現以本研究之 NFFDA 2 演算法有最佳的辨識效果。

本研究實驗分類如下：12.2.1 比較不同前置處理與標準差之誤差總值。12.2.2 NFFDA 1 演算法之實驗結果。12.2.3 NFFDA 2 演算法之實驗結果。12.2.4 統計與比較三種演算法之結果。

12.2.1 比較模糊高斯分類器其不同前置處理與標準差

比較人臉圖像與非人臉圖像，在是否有經過高斯濾波之影像處理下，以及在標準差為 $\alpha=1, \alpha=2, \alpha=4$ 時與樣本圖像之間，其誤差平方累加之變異情況。由實驗結果發現，當閾值 θ 設定為 14 時，在無經過高斯濾波處理的影像，與樣本圖像間的誤差比有高斯濾波處理的圖像顯得為大，表示人臉圖像在經過低通濾波處理後會更貼近樣本圖樣；在標準差的選定上，基本上標準差越高，其得出的值越高，由數據得知在 $\alpha=2$ 時，其人臉影像與樣本圖像的誤差沒有顯著增大，反而使非人臉圖像與人臉樣本圖像間的反差加大，促使在實驗結果上得以區分人臉與非人臉圖像。故在本實驗中，後續提出的演算法皆以影像須經過高斯低通濾波處理，並且使用標準差 $\alpha=2$ 計算之。實驗結果請見附表一。

在經過模糊高斯分類器的 ORL 人臉資料

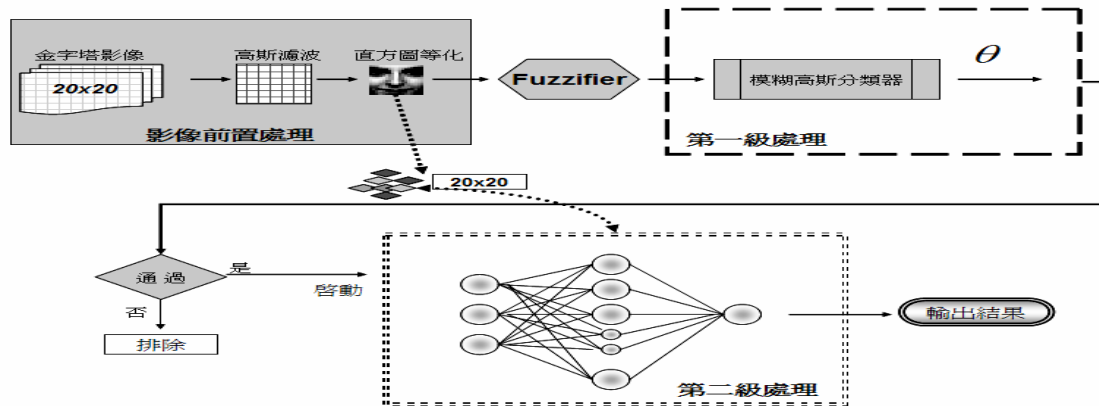


圖 19 NFFDA1 架構圖

庫與生活照為主的人臉測試樣本，其人臉圖像全數通過，而由分類器判定為近似人臉之非人臉圖像有 11 張。將通過閾值的人臉圖像及非人臉圖像，啟動下一級類神經訓練系統，並於 12.2.2 和 12.2.3 節分別呈現本研究提出的方法一(NFFDA 1)與方法二(NFFDA 2)演算法其最後測試之結果。

12.2.2 NFFDA 1 演算法實驗結果

NFFDA 1 演算法為 NNFD 的最初改良，即該演算法在經過模糊高斯分類器後，啟動類神經網路，而此網路架構便是結合 NNFD 之架構，而此訓練樣本和輸入值為直方圖等化後的影像，架構如圖 19 所示。由實驗結果發現，該網路架構有其不穩定之處，是由於訓練樣本影響的緣故，雖然已在第一級處理上刪除許多非人臉圖像，但其辨識效果仍可見於對人臉及非人臉圖像辨識的不穩定，致使提出 NFFDA 2 演算法改良之。實驗結果請見附表二。

12.2.3 NFFDA 2 演算法實驗結果

本研究之 NFFDA 2 演算法針對模糊高斯分類器處理後的圖像進行特徵擷取類神經訓練，此網路架構加重了特徵區塊的訓練，提升了臉部的辨識度，不但經由分類器初步篩選過濾大量非人臉影像外，縮小訓練樣本也可以減少訓練時間，且經過模糊計算後成為影像模糊參數向量型態的人臉影像，不易受雜訊干擾，有助於人臉在不同角度、位置上的辨識。此實驗結果請見附表三。

12.2.4 統計與比較三種演算法之結果

NNFD 其訓練結果容易受訓練樣本影響，且當加入偵測錯誤之樣本於訓練樣本中時，雖然會提升對於非人臉的辨識效果，但相對的也就容易受到近似於人臉的非人臉圖樣所干擾，影響其訓練中權重的變化，使得最後訓練的結果對於類似人臉的非人臉圖像的判定，有

模稜兩可的結果出現。

故本研究便提出演算法 NFFDA 1 演算法，加入模糊高斯分類器，可排除大量的非人臉圖像，使之減少訓練樣本數量也減少較多的樣本雜訊，再與 NNFD 之網路架構結合，有其互補的優點。雖然將非人臉暫時歸類為人臉圖像，而訓練結果可使得類似人臉之圖像多數刪除，但不難發現，由實驗附表二的圖(7,9,10,11)觀察到，近似人臉或極似人臉的非人臉圖像其訓練結果不一致，且人臉的辨識值部分介於 0.5~0.7 之間，所以本研究針對此類神經網路再做改善，提出了 NFFDA 2 演算法。

NFFDA 2 演算法有鑑於雜訊的影響之大，改變網路訓練其輸入樣本，除了結合模糊高斯分類器外，並將圖像之模糊集合以人臉的型態，即以二維的方式做為訓練樣本並測試之。如此不但提升 NFFDA 1 演算法的訓練結果，對於圖像的判定也更為明確，將類似臉的非人臉圖像幾乎排除，並提升人臉的辨識度為 0.9 左右。

最後，比較該三種演算法其訓練次數與時間，以及偵測效果。見下表 10 及表 11。

本研究使用之作業系統為 Microsoft XP Professional Version 2002 Service Pack2，處理器為 Intel(R) Pentium(R) D CPU 3.40GHz，記憶體為 1.00GB RAM，此實作系統使用 Microsoft Visual C++ 6.0 完成。

表 10 三種演算法比較

	訓練次數	訓練時間(秒)	收斂誤差
NNFD	80 萬	2825.12	0.0348673
NFFDA 1	80 萬	3248.95	0.0311592
NFFDA 2	80 萬	1779.89	0.0328714

表 11 偵測效果

	測試(張)		錯誤(張)		提升(張)	
	臉	非臉	臉	非臉	臉	非臉
NNFD	55	50	4	5		
NFFDA 1	55	50	4	2		
NFFDA 2	55	50	4	2	27	5

由以上列表整理看出，NFFDA 2 於訓練次數為七十萬次時便可達收斂，增加訓練次數為八十萬次時可使辨識度更加精確。在訓練次數相同下，其訓練時間較小，亦可達到不錯的收斂，而偵測效果以輸出結果來與 NNFD 比較，在偵測結果為正確且提升 0.01 以上者為前提，實驗得知大致上皆有提升，其精確度為三者最高。

12.3 實驗—人臉定位演算法

12.3.1 實驗樣本

本節使用 BioID 人臉資料庫影像，此為含有背景與人臉的影像，此資料庫中影像原尺寸為(寬×高)，我們縮小成適當的比例並比較不同閾值的設定，進行人臉定位實驗。

12.3.2 實作分析與探討

定位實驗將影像經過一層的縮小影像比例，依據每張影像縮小成影像中的人臉可由視窗框選的大小進行定位實驗，我們比較不同閾值的設定，分別為 1.5、1.6、1.7、1.8 共四種，不同的閾值對於篩選偵測結果的區塊會影響定位的準確度，故我們加以探討高斯擴展定位閾值的部分，來看定位後的效果有何不同，期望減少人臉的遺失率和誤判。

觀察影像經過偵測後的結果，特定人臉區塊的偵測值可達 0.9 以上，以此區塊為中心的鄰近區塊之偵測值範圍為 0.5~0.8，經過高斯擴展之後可使得特定人臉區塊在所有的偵測區塊中產生極端值，由定位實驗的閾值測試 1~測試 2，如圖 20 所示。閾值設定為 1.5 或 1.6 的定位結果，測試 1~測試 3 的臉部附近的像素平移都能找到人臉的位置，而測試 4 的臉部右上有一個偏移較多的框，所以閾值 1.5 或 1.6 人臉的定位結果較多重疊區塊，甚至可能有偏移臉部較多的區塊，容易造成誤判。而閾值設定為 1.7 或 1.8 時，臉部位置結果較準確，誤判的區塊也被移除了，留下定位區塊最少且能越精確框住臉部五官位置。



圖 20 定位實驗<1>之閾值比較

所以我們將經過高斯擴展後的結果表以閾值 1.8 作界定，僅留下大於 1.8 數值的座標區塊，而找出了人臉在影像中的位置，在閾值的限制下不但可排除重複偵測的錯誤區塊，也使人臉區塊更加集中，能框選出較佳的人臉位置。

12.4 實驗—人臉辨識演算法

12.4.1 實驗樣本

本階段實驗樣本的類別共 20 人，分別為 (S1,S2,...,S20)，其影像編號分別為 (0101,0102,...,0110); (0201,0202,...,0210);; (2001,2002,...,2010)，總共 200 張作為本階段實驗之用，如圖 21。

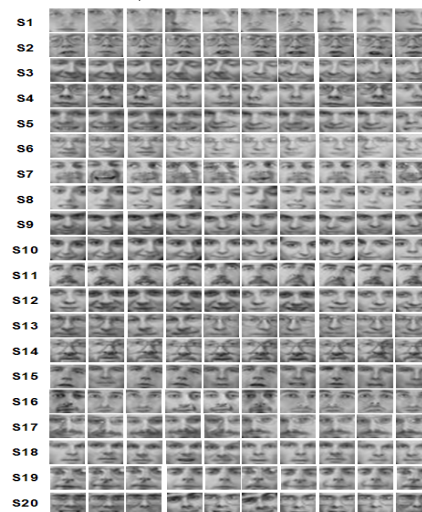


圖 21 XORL 人臉資料庫 S1~S20 共 200 張人臉影像

12.4.2 實作分析與探討

12.4.2.1 FCM 歸屬度數據分析

為了清楚了解歸屬度的特性，在此以每個類別各取一張正臉影像總共 20 張實驗之，如

附表四所示，為經過前置處理前的原灰階影像，將這 20 張 20×20 灰階值正臉 (Nf = 20) 影像分作 10 群，主要是希望每張影像儘可能的隸屬於兩群，所以設定群集數為 (C = 10)；而此例設定解模糊化之閾值 $\theta = 0.15$ 。

如表 12 所示，為這 20 筆資料組對於在 10 個群集中的歸屬程度，可以見得 Nf 2, Nf 3, Nf 5, Nf 6, Nf 7, Nf 13, Nf 16, Nf 20 在各群集的歸屬度呈現比較鬆散的分佈，代表這幾張影像在這 20 張中的分群是比較容易受到影響；相對地，歸屬度呈現較明確的數值分佈之資料組，大部分會各自佔居一群，代表這幾個資料組在全部 20 筆當中是較具代表性、較為有獨特性且不易受到其他資料組而影響分群情況的影像，然而較為鬆散的資料組便會依據當下目前的樣本組間的計算進行適當的分群。

所以，若分析同一群集內的資料組，其各資料間有其歸屬值較明確者以及鬆散者，甚至含有兩個以上較為明確的資料組於同一群中，由於歸屬值的明確性代表該影像資料不但不易受其他影像來干預分群，自身也不容易因為影像本身的特性所影響，例如：雜訊、臉部表情；總而言之，歸於同一群集內的資料組相似性較高，而群與群之間的資料則相似性較低。

如表 13 所示，為這 20 筆資料經過解模糊化計算後對於在 10 個群集中的歸屬程度，本實驗在此以橫向取前五大之歸屬值，完成後便以縱向取前六大之歸屬值，其餘未符合者便刪除之或者將其值設為 0，可以看出較為明確之資料組大致上僅剩下一個最高的歸屬值歸類於某一類，且逐一分布於各群之中，而數值鬆散之資料組便大致上分屬於兩個群集內，最後合併 10 個群集為 4 個子網路 進入至平行類神經網路作訓練，如表 14 所示。

實驗結果在同類別單張分群上，影像是否經過高斯濾波後再行直方圖等化處理，對於分群結果上的影響差異不大，其後，在分群確立後進行分群測試，抓取各類別各一張不與分群之訓練樣本重覆的影像，共 20 張，分群測試取第一順位的分群結果為有經過高斯濾波處理者優於僅經過直方圖等化之影像；如此來決定前置處理的方式。

表 14 分群結果之整理

分 群		合併後結果
C1: Nf 2 4	C6: Nf 5 7 17	Subnet 1: 2 4 5 11 19 1 Subnet 2: 3 12 20 8 15 Subnet 3: 6 9 10 13 16 18 Subnet 4: 5 7 17 13 14
C2: Nf 2 5 11 19	C7: Nf 13 14	
C3: Nf 1	C8: Nf 6 9 10 13 16	
C4: Nf 6 16 18	C9: Nf 8	
C5: Nf 15 20	C10: Nf 3 12 20	

表 12 FCM 分群後之歸屬度表

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
N1	0	0	0.999999	0	0	0	0	0	0	0
N2	0.30585	0.226198	0.058137	0.036721	0.027195	0.0627	0.076583	0.036229	0.140609	0.029779
N3	0.019096	0.029277	0.004865	0.019928	0.058557	0.020827	0.037849	0.040552	0.015637	0.753412
N4	0.999648	0.000142	0.000017	0.000014	0.000017	0.000019	0.000025	0.00002	0.000065	0.000033
N5	0.040746	0.1976	0.020732	0.099873	0.139682	0.153148	0.109544	0.083581	0.098527	0.056557
N6	0.026121	0.040187	0.037279	0.376612	0.036613	0.068853	0.060827	0.285513	0.022139	0.045856
N7	0.036229	0.139386	0.016572	0.062228	0.020076	0.526632	0.07887	0.049483	0.050142	0.020382
N8	0.000001	0.000003	0.000001	0	0	0.000001	0	0	0.999992	0
N9	0.002513	0.00578	0.001188	0.026183	0.003399	0.016605	0.009023	0.927381	0.001912	0.006014
N10	0.002634	0.006045	0.002688	0.04373	0.003517	0.012591	0.003911	0.913231	0.003162	0.008472
N11	0.012162	0.924702	0.002968	0.005831	0.005292	0.009693	0.015219	0.005438	0.012765	0.005931
N12	0.005278	0.006914	0.002328	0.005818	0.008199	0.002922	0.003552	0.009388	0.004138	0.951463
N13	0.053461	0.090525	0.021961	0.145965	0.051826	0.092303	0.181594	0.26614	0.032497	0.063727
N14	0.000002	0.000006	0.000001	0.000004	0.000002	0.000006	0.999973	0.000003	0.000002	0.000002
N15	0.000001	0.000001	0	0.000001	0.999991	0.000001	0.000001	0.000001	0.000001	0.000003
N16	0.049018	0.119322	0.138792	0.151327	0.063662	0.069272	0.04649	0.186641	0.091151	0.084325
N17	0.000386	0.001453	0.000223	0.002464	0.000334	0.98925	0.001571	0.003352	0.000556	0.000413
N18	0.000064	0.000249	0.000098	0.997095	0.000103	0.000538	0.000237	0.00137	0.00011	0.000136
N19	0.052198	0.826424	0.008483	0.011152	0.005568	0.0174	0.01021	0.011745	0.049029	0.007789
N20	0.042992	0.039612	0.015518	0.031134	0.167722	0.019447	0.028078	0.051872	0.020739	0.582257

表 13 解模糊化之歸屬度表

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
N1	0	0	0.999999	0	0	0	0	0	0	0
N2	0.30585	0.226198	0	0	0	0	0	0	0	0
N3	0	0	0	0	0	0	0	0	0	0.753412
N4	0.999648	0	0	0	0	0	0	0	0	0
N5	0	0.1976	0	0	0	0.153148	0	0	0	0
N6	0	0	0	0.376612	0	0	0	0.285513	0	0
N7	0	0	0	0	0	0.526632	0	0	0	0
N8	0	0	0	0	0	0	0	0	0.999992	0
N9	0	0	0	0	0	0	0	0.927381	0	0
N10	0	0	0	0	0	0	0	0.913231	0	0
N11	0	0.924702	0	0	0	0	0	0	0	0
N12	0	0	0	0	0	0	0	0	0	0.951463
N13	0	0	0	0	0	0	0.181594	0.26614	0	0
N14	0	0	0	0	0	0	0.999973	0	0	0
N15	0	0	0	0	0.999991	0	0	0	0	0
N16	0	0	0	0.151327	0	0	0	0.186641	0	0
N17	0	0	0	0	0	0.98925	0	0	0	0
N18	0	0	0	0.997095	0	0	0	0	0	0
N19	0	0.826424	0	0	0	0	0	0	0	0
N20	0	0	0	0	0.167722	0	0	0	0	0.582257

本實驗使用 20 個類別各取單一影像做分群訓練，以及做 20 張影像的分群測試和辨識，發現在分群測試上的結果，錯誤分群的數量逼近測試影像張數的一半，暫且不依測試影像類型論之，僅使用單張人臉正面影像做分群訓練樣本太過於單一性，使得測試上容易分群錯誤，進而使得測試影像歸類至錯誤的子網路做辨識，非但沒有達到辨識效果，反而促成該辨識整體實驗的失敗。針對於此，進行同類別多張分群的實驗。

同類別多張分群實驗為改善分群效果將

隨機抓取每個類別各 5 張影像作為分群訓練，其人臉影像包含有正向、側臉、不同表情等類型，共 100 張。

由實驗結果得之，分群測試之錯誤張數比例低於 $\frac{1}{4}$ 以下，表示說採用多張影像來分群可使得測試樣本在群集歸屬上更有彈性，不會因為臉部影像的傾斜或樣貌的些微改變(例如：同一類別戴眼鏡與否、有表情與否)而容易造成錯誤歸類。

12.4.3 採用 fivefold Cross Validation 訓練

上一節在平行式類神經網路上的訓練樣本採用固定的 100 張影像，在此實驗將進行較公平的訓練及測試，讓每張影像在辨識時都能為訓練樣本以及測試樣本。在此共有四個平行子網路 (*Subnet1* ~ *Subnet4*)，其每個子網路所屬的成員數分別為 12、12、12、10 筆，樣本設計方法如下：將 *S1* ~ *S20* 共 20 個類別的每個類別各 10 張影像分成五段，以 *Subnet1* ~ *Subnet3* 來說，成員數皆為 12 筆，並且希望每個類別抓取 8 張做為訓練樣本，剩下 2 張就為測試樣本，所以前三個子網路的訓練樣本張數皆為 96 張，測試樣本為 24 張，而第四個子網路的訓練樣本張數為 80 張，測試樣本為 20 張，分配如下：

1. training : 1 ~ 8 testing : 9,10
2. training : 1 ~ 6,9,10 testing : 7,8
3. training : 1 ~ 4,7 ~ 10 testing : 5,6
4. training : 1 ~ 2,5 ~ 10 testing : 3,4
5. training : 3 ~ 10 testing : 1,2

12.4.4 辨識演算法之結果

進行 40 萬次的訓練，隱藏層使用 15 顆神經元，進行進入單一平行網路以及進入兩個平行子網路的辨識結果，結果顯示在分群歸屬上選取兩群以及進入兩個平行網路進行訓練，最終可達平均錯誤率 $E_{ave} = 0.09$ 為較良好的辨識結果。平均錯誤率之計算方式(如公式 12)所示[8]， q 是實驗次數； n_{mis}^i 為在第 i 次實驗辨識錯誤的次數； n_{tot} 為每次實驗之總測試樣本數。

$$E_{ave} = \frac{\sum_{i=1}^q n_{mis}^i}{qn_{tot}} \quad (12)$$

辨識實驗的測試總張數為 200 張，紀錄分群歸屬取一群和取兩群的平均錯誤率，共實驗

5 次，分別記錄每次在不同分群歸屬下的辨識錯誤張數，由表 15 所示，可看出每次的實驗中，未知樣本進入兩個平行子網路所得到的辨識效果遠大於進入單一網路，減少許多辨識判別的張數，我們將 5 次實驗之錯誤張數加總與總張數相除，得到不同分群歸屬的平均錯誤率，以進入兩個平行子網路的辨識結果較好。

表 15 辨識結果之平均錯誤率

取 一 群					E_{ave}	
i (次)	1	2	3	4	5	
n_{mis}^i (張)	4	10	7	11	7	0.195
取 兩 群					E_{ave}	
i	1	2	3	4	5	
n_{mis}^i (張)	2	7	1	6	2	0.09

藉由實驗結果來探究辨識錯誤的類型有如下三種：(1)皆錯誤分群導致辨識結果錯誤；(2)訓練樣本樣貌與測試樣本大不相同所致；(3)分群正確，但因高斯模糊後與其他樣本相似。辨識結果之錯誤類型整理如附表五所示。

12.5 系統整合實作

本研究根據所提出之三大演算架構加以整合，將進行生活照片中研究所同儕們的偵測與辨識，並對實驗結果進行分析討論。

12.5.1 實驗樣本

收集研究所同學之生活照片，影像格式為 .pgm 檔(可攜式灰階圖檔)，尺寸大小為，照片中包含的人臉為 2 或 3 人，將照片經過人臉偵測、定位與辨識過程，判定照片中人臉影像的類別歸屬。

實驗 6 張生活照片 (*IMG 1* ~ *IMG 6*)，共含有 13 種不同類別的人臉影像，將每個類別影像予以編號 (*R 1* ~ *R 13*)，在實作上把照片中偵測到的人臉影像經過 FCM 分群後，進入所歸屬的子網路進行辨識，得到的網路最大輸出值則決定人臉的身分。我們隨機挑選 13 個類別各 5 張不同角度、表情或有無戴眼鏡的影像，作為分群與平行子網路的訓練，如圖 22 所示。

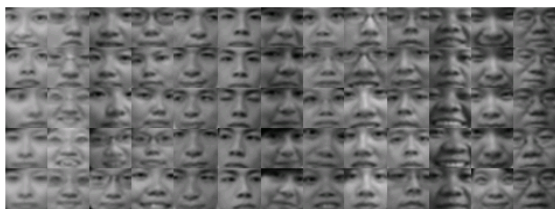


圖 22 十三種人臉類別之訓練樣本

12.5.2 實作結果分析與探討

將這 13 個類別初步分成 6 群，希望每個類別儘可能歸屬於兩群集，如表 16。分群後再將結果加以合併，縮小為 3 個合併的群集，如表 17。結果如表 18、19。

偵測、定位部份都成功擷取人臉區塊，如圖 23，而在辨識部份，絕大多數也都辨識成功，只有 IMG2 的 R4 與 IMG6 的 R3 辨識錯誤，是由於影像經過高斯濾波器後，使得模糊集合的數值相近導致誤判，為第三種錯誤類型。

表 16 初步分群結果表

G1	1	4	7	
G2	2	8	9	10
G3	3	4	9	11
G4	3	5	6	13
G5	5	12		
G6	10			

表 17 合併後之平行子網路

Subnet 1	1	4	7	3	9	11
Subnet 2	2	8	9	10		
Subnet 3	5	12	3	6	13	

表 18 平行子網路之訓練結果

Subnet	隱藏層神經元數	輸出層神經元數	訓練次數(萬)	訓練時間(秒)
Subnet1	8	6	60	568.265
Subnet2	7	4	40	266.422
Subnet3	8	5	60	492.156

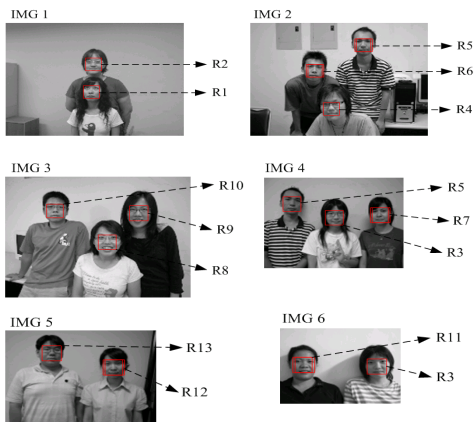


圖 23 系統整合之人臉偵測與定位結果

表 19 系統整合之人臉辨識結果

生活照片編號	人臉影像編號	歸屬子網路	網路最大輸出值	辨識結果	Y / N
IM G1	R1	Subnet 1	0.955519	R1	Y
		Subnet 1			
IM G1	R2	Subnet 2	0.992591	R2	Y
		Subnet 3			
IM G2	R4	Subnet 1	0.948085	R7	N
		Subnet 1			
	R5	Subnet 3	0.996139	R5	Y
	Subnet 2				
IM G2	R6	Subnet 3	0.996581	R6	Y
		Subnet 2			
IM G3	R8	Subnet 1	0.839562	R8	Y
		Subnet 2			
	R9	Subnet 1	0.997376	R9	Y
	Subnet 3				
IM G3	R10	Subnet 2	0.963651	R10	Y
		Subnet 3			
IM G4	R3	Subnet 1	0.428337	R3	Y
		Subnet 1			
	R5	Subnet 2	0.648196	R5	Y
	Subnet 3				
IM G4	R7	Subnet 1	0.994985	R7	Y
		Subnet 2			
IM G5	R12	Subnet 3	0.970764	R12	Y
		Subnet 3			
IM G5	R13	Subnet 3	0.99680	R13	Y
		Subnet 2			
IM G6	R3	Subnet 1	0.90233	R9	N
		Subnet 2			
	R11	Subnet 1	0.98253	R11	Y
	Subnet 2				

13. 結論

本研究提出的人臉偵測與辨識系統，有效的偵測出人臉且識別身分，可以應用於生活照片上，下列舉列本研究的重要發現：

1. 灰階影像以模糊高斯集合性態的方式來表示，有效減低雜訊干擾、降低影像複雜度，透過每個像素在集合間的高斯分佈情形的不同，提升對於人臉與非人臉區塊的區分，來排除大量的非人臉圖像。
2. 本系統對於人臉的偵測相當準確，無論是多樣貌或角度傾斜大的人臉，偵測到的重疊區塊相當集中，愈向中心集中的區塊便是人臉中五官最準確的位置。
3. 以影像集合平均值結合 FCM 分群演算法的運用，利用歸屬度找出像素在影像間以及影像在群集間的分布關係，即當人臉影像的五官在高斯分佈中越鮮明，其偵測準確度越高，辨識在 FCM 分群上的歸屬度越高，其容易成為所有類別中獨特性高的影像，當然在辨識結果上便不容易產生誤判。

隨著影像自動化時代的來臨，龐大的影像資料出現在日常生活中，人類在生活中與人互動中會判斷眼前之人為熟識者或陌生人，當面對大量群眾時便需要辨識系統的輔助，所以人臉辨識系統的存在有其重要性，期望透過本研究能推達人臉識別系統的發展與應用。

14. 致謝

本論文承蒙義守大學校內專題研究計畫 ISU98-01-12 經費補助，特此致謝。
















參考文獻

- [1] A. Henry, B. Shumeet, and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans.PAMI*, vol. 20, no.1, Jan. 1998.
- [2] B. Chen, Z.-Q. Lid, X.-H. Zhu, "Eye Location in Human Face Images using Fuzzy Internal," *Proceedings of the Second International Conference on Machine Learning and Cybernetics*, Nov 2003.
- [3] C. Lee and F. C.-H. Wee, "Face Segmentation using fuzzy memberships," *KFIS 2000*, vol. 10, no. 1, May 2000.
- [4] Frank Chung-Hoon Rhee; Changsu Lee; "Region based fuzzy neural networks for face detection," *IFSA World Congress and NAFIPS*, Vol. 2, pp.25-28, July 2001.

- [5] Fadzil, M.H., A.; Lim Cheah Choon; "Face recognition system based on neural networks and fuzzy logic," *Neural Networks, International Conference on Vol 3*, pp.9-12, June 1997.
- [6] H. R. Kanan, K. Faez, S. Mozaffari and M. Ezoji, "Face Location Using Shape Information and Genetic Algorithm," *Proceeding of the First Int. Conference on Modeling, Simulation and Applied Optimization, Sharjah, U.A.E.*, 2005.
- [7] Lu, J.; Yuan, X.; Yahagi, T., "A Method of Face Recognition Based on Fuzzy c-Means Clustering and Associated Sub-NNs," *Neural Networks, IEEE*, vol.18, Issue 1, pp.150-160, Jan. 2007.
- [8] M. J. Er and M. S. Wu., "Face recognition with radial basis function (RBF) neural networks," *IEEE TRANSACTIONS ON NEURAL NETWORKS*, vol.13, no.3, pp.697-710, May 2002.
- [9] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of Bimple features," *in Proc. IEEE CVPR*, 2001.
- [10] Xue Yuan; Jianming Lu; Yahagi, T.; "Face recognition based on neuron fuzzy systems," *Circuits and Systems, MWSCAS'04*, vol. 3, pp.25-28, July 2004.
- [11] 蔡賢亮、蔡佩君、周煜書，"Neural Fuzzy Face Detecting Algorithm," *2008 資訊科技國際研討會論文集*，2008。

附表一
















12.2.1 比較模糊高斯分類器其不同前置處理與標準差

(1) <i>ORL</i> face	1	2	3	4	5
					
$\alpha = 1$	7.299035	5.689274	4.008939	5.364057	3.257252
$\alpha = 2$	8.87189	7.57344	5.79556	7.08475	4.78995
$\alpha = 4$	9.31362	9.012087	7.619907	8.995963	6.709157
(2) 生活照 face	1	2	3	4	5
					
$\alpha = 1$	6.229736	6.775981	5.975238	5.095221	3.718123
$\alpha = 2$	7.605207	7.265848	7.045725	5.972606	4.806253
$\alpha = 4$	11.35013	8.234486	9.564441	7.935531	5.962645
(3) 非人臉	1	2	3	4	5
					
$\alpha = 1$	9.699576	9.990326	15.41411	15.50074	15.59173
$\alpha = 2$	10.83679	10.59017	14.86791	15.94383	14.4515
$\alpha = 4$	12.55267	11.92663	16.31592	19.32058	15.61206

(1) *ORL* 人臉資料庫之誤差總值，其影像無經過高斯濾波器之處理，比較其使用不同標準差之結果。列表之標準差為 $\alpha = 1, \alpha = 2, \alpha = 4$ 。由人臉資料庫擷取出的影像其因為在拍攝上有規則的調整角度與位置，故其對稱性與偏移程度皆與樣本圖像大致相符，使得誤差值很小。

(2) 生活照 face 是由網際網路擷取出的影像較為隨意，而其人臉角度與位置甚至表情都來得比人臉資料庫的影像豐富，使其與樣本圖像的誤差值增加。

(3) 擷取自生活照非人臉影像之誤差總值，大致上非人臉之圖像皆與人臉樣本大相逕庭，除了較為近似之圖像與人臉樣本圖像之誤差與生活照之人臉圖像相似外，其餘誤差值皆大於閾值限制。

(1) <i>ORL</i> face	1	2	3	4	5
					
$\alpha = 1$	7.709642	5.72482	3.949015	5.642171	3.152459
$\alpha = 2$	9.47114	7.59884	5.47457	6.70125	3.9554
$\alpha = 4$	10.4893	8.910831	6.693185	7.689151	4.676306
(2) 生活照 face	1	2	3	4	5
					
$\alpha = 1$	5.397591	6.525297	7.636761	5.167151	3.216752
$\alpha = 2$	5.927911	6.614997	8.1231881	5.589324	3.536917
$\alpha = 4$	9.789328	8.256213	10.04211	7.63039	5.422226
非臉	1	2	3	4	5
					
$\alpha = 1$	7.489663	10.20285	16.29523	16.46089	17.34961
$\alpha = 2$	9.425227	10.71811	15.67063	17.10158	16.40261
$\alpha = 4$	11.65755	11.40315	16.25726	20.05986	15.21815

(1) *ORL* 人臉資料庫之誤差總值，其影像經過高斯濾波器之處理，比較其使用不同標準差之結果。列表之標準差為 $\alpha = 1, \alpha = 2, \alpha = 4$ 。誤差值小的原因同(1)無經過高斯濾波器之處理。





















(2) 生活照 face 其誤差值原因同(2)無經過高斯濾波器之處理。

(3) 擷取自生活照非人臉影像之誤差值原因同(3)無經過高斯濾波器之處理。



附表二

12.2.2 NFFDA 1 演算法

(1) 人臉測試結果

face	1	2	3	4	5
					
NFFDA 1	0.9915	0.971523	0.946965	0.770544	0.988801
face	6	7	8	9	10
					
NFFDA 1	0.999541	0.987853	0.993853	0.936874	0.991777
face	11	12	13	14	15
					
NFFDA 1	0.985179	0.992362	0.969886	0.985718	0.993613
face	16	17	18	19	20
					
NFFDA 1	0.999237	0.990049	0.996847	0.947393	0.994982





















(2) 非人臉測試結果

非臉	1	2	3	4	5	
						
NFFDA 1	0.0908197	0.0132792	0.110299	0.0222794	0.0426323	
非臉	6	7	8	9	10	11
						
NFFDA 1	0.0159153	0.5744903	0.0933055	0.0187543	0.102614	0.818809












附表三

12.2.3 NFFDA 2 演算法

(1) 人臉測試結果

face	1	2	3	4	5
					
NFFDA 2	0.971122	0.967993	0.989715	0.942646	0.996216
face	6	7	8	9	10
					
NFFDA 2	0.999877	0.986363	0.998906	0.971062	0.998375
face	11	12	13	14	15
					
NFFDA 2	0.994327	0.999336	0.999721	0.999765	0.999831
face	16	17	18	19	20
					
NFFDA 2	0.999680	0.997968	0.999097	0.998891	0.999581

(2) 非人臉測試結果

非臉	1	2	3	4	5	
						
NFFDA 2	0.394838	0.003114	0.002240	0.000092	0.002165	
非臉	6	7	8	9	10	11
						
NFFDA 2	0.007679	0.001486	0.008823	0.003775	0.956486	0.992141

附表四

12.4.2 實作分析與探討



圖 1 二十張正面人臉影像

附表五

12.4.4 辨識演算法之結果

表 1 辨識誤判類型

錯誤類型	測試編號	測試影像 (原影像)	測試影像 (經前置處理後)	→	辨識錯誤情形
錯誤類型 1	S7.0710				
錯誤類型 1	S11.1108				
					戴眼鏡之訓練樣本
錯誤類型 2	S17.1703				
錯誤類型 2	S17.1704				
					誤判類別
錯誤類型 3	S14.1408			→ S6	
錯誤類型 3	S15.1502			→ S18	