

以機率類神經網路為基礎的臉部名稱推薦系統

陳定宏
南台科技大學
資訊工程系/所
副教授

chendh@mail.stut.edu.tw

唐祥恩
南台科技大學
資訊工程系/所
研究生

m95g0223@webmail.stut.edu.tw

摘要

本文提出一個具有人臉偵測與輔助註釋功能的臉部名稱推薦系統，並將之實作於網路相簿之上。因為每張相片中之臉部影像品質不一，所以我們先利用光線補償與標準人臉圖樣擷取方法將臉部影像重新正規化。接下我們以改良式機率類神經網路(Modified Probabilistic Neural Network, MPNN)的來對於臉部影像進行分類與辨識。本系統具有漸進學習功能，因此當使用者註釋過一些人臉之後，就可以對於新加入的臉部建議推薦名稱。實驗結果顯示，新加入之人臉影像，可以快速及正確的得到推薦名稱，對於現今普遍使用的網路相簿，可以增加人臉註釋與搜尋的新功能。

關鍵詞：類神經網路、臉部偵測、臉部推薦

Abstract

In this paper, we propose face detection and name recommendation mechanism used in the photo albums. Before applying the name tags to the face photos, we perform light compensation to improve the image quality. Then a face detection algorithm is applied to retrieve the face region. We use modified probabilistic neural network (MPNN) to train the sample faces for each individual's photo album. After training, the new adding face photos can be recognized and a recommended name tag is assigned to this face. The experiments show a promising result for an accurate name tag annotation. The proposed method is very appropriate to provide a name or keyword based management function which applied in photo album.

Keywords: Neural Network, Face Detection, Face Recognition

1. 前言

數位相機的普及，讓影像紀錄變得很容易。根據統計，92.9%的使用者採用數位相機當作主要的影像紀錄方式，另外約有 56.9%的使用者則是以手機拍照，只有約 13.8%的使用

者採用傳統相機拍照[1]。在拍照頻率方面，「一星期至少拍照 1 次以上」的約佔 47.5%。其中更有 11.2%表示「每天或幾乎每天」都會拍照。由這些數據可以知道數位影像的應用已經非常普遍。

而相片的儲存方式，63.5%的使用者「存在電腦硬碟」，17.1%的使用者「放在網路相簿」，只有 10% 的使用者將相片「沖印出來」[2]。更因為網路相簿繼承了網路分享的特性，最多受訪者(66.5%)將照片「放在自己的網路相簿」當作與朋友分享的最主要方式，其次則是「即時通訊軟體的大頭貼」(40.8%)。由此可說明影像分享的方式，已經從舊的實體沖印變成數位瀏覽。

近幾年來，人臉偵測(face detection)、人臉辨識(face recognition)技術[11][12]已經實際被用再許多不同的領域上。像是在 2008 北京奧運上就運用了人臉辨識的技術應用在流量較大的大眾運輸工具中來做安全監控[3]。一些新款的筆記型電腦也增加了六點特徵的人臉辨識登入系統，保障了使用者使用筆記型電腦的安全性。經由以上觀察，若能夠將人臉偵測、註釋與檢索在網路相簿上結合人臉偵測，應該能增加網路相簿的實用性。

在大量的數位相片中，主要仍以人像照片為最大量，所以將能將人臉框出並且加上標註，將可以增進相簿管理與人臉檢索的效率。以手動方式讓使用者對於人臉加註，固然可以確保註釋的正確性；但重複的動作會讓使用者產生惰性，若是新增大量的相片，不僅註釋起來費時且讓系統缺乏實用性。本論文提出以改良式機率類神經網路的方法，將相簿中已註釋的臉進行訓練。若是有新的人臉出現，就會利用每一個人臉所訓練過的類神經網路結果，做為該人臉的推薦名稱。

本論文組織如下。第 2 節描述本論文所提出的人臉名稱推薦系統架構。第 3 節說明如何將取得的人臉做重複正規化以便接受機率類神經網路訓練。第 4 節則說明利用臉部推薦系統取得新進人臉的名稱排名。最後第 5 節將呈

現實的結果。第 6 節我們針對本論文所提出的方法給予一個的結論，同時提出未來可改進的方向。

2. 人臉名稱推薦系統架構

一般的網路相簿使用流程如圖 1 所示。使用者要經由加入提供服務平台網站的會員，例如：無名小站、Flickr 與 Pixnet 等。登入之後，就可以編輯新的相簿其功能包括：相簿與相片註解、設定相簿資料或是增刪相簿或相片等。經由使用者輸入相片之後經過縮圖、整理過後就可以在網路相簿上觀看。

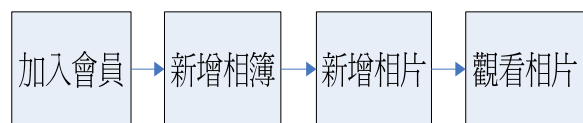


圖 1 一般網路相簿之使用流程圖

本論文的系統架構如圖 2 所示，主要分為臉部前處理系統以及臉部名稱推薦系統兩個主要的功能模組。人臉相片是分別由人臉自動偵測與使用者選取兩方面取得。然選取的相片是已經經過註釋，但是要接受訓練的相片品質參差不齊。因此臉部前處理模組主要是將所有輸入的人像，透過前處理得到一致的影像品質，同時將人臉的位置定位。人臉偵測可利用 OpenCV 函式庫或是讓使用者自己框出相片中人臉的區塊。因此本論文先以標準人臉圖樣的臉部偵測技術來進行臉部影像正規化，在利用亮度補償技術將取得的相片亮度能在同一個範圍之內，使訓練的人臉品質可以一致。

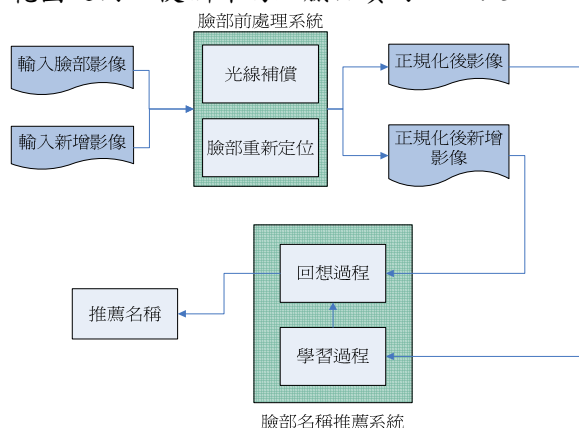


圖 2 人臉名稱推薦流程圖

經過前處理得到的臉部影像，我們將採用機率類神經網路(Probability Neural Network, PNN)作為人臉影像的訓練機制。經由機率類神經網路的學習 (learning process) 與回想

(recalling process)過程，可以對於輸入的人臉區域加以訓練，作為日後辨識新加入的人臉之用。

從網路相簿取得的臉部資料庫中，將臉部影像的像素依序排列成一個 $n*n$ 的二維陣列或 n^2 的一維向量，將此一維向量當作輸入特徵向量，輸入機率類神經網路。PNN 的學習過程就是把每一個使用者已經註釋的人臉經過前處理之後，將臉部影像的所有像素當成輸入單元與隱藏層單元之間的權重，分別訓練每一個使用者所專屬的神經網路，然後將訓練完成的權重存入網路資料庫中，如圖 3 所示。

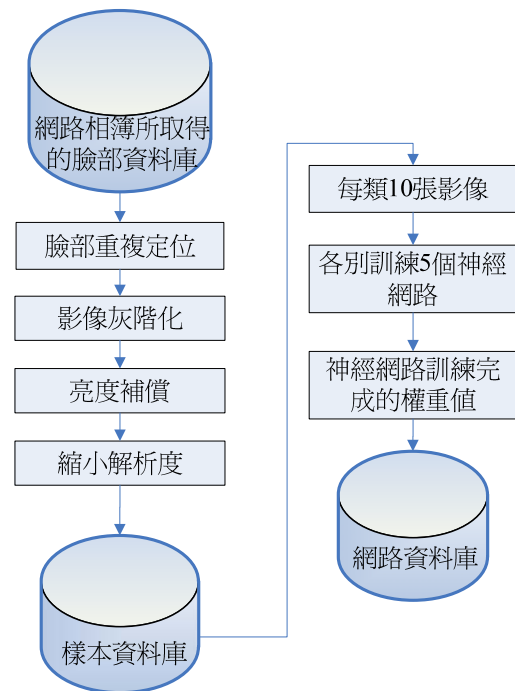


圖 3 神經網路的學習過程

在回想過程中，當有新相片需要推薦時，將新增臉部影像先進行前處理，輸入神經網路配合網路資料庫中的權重，經過訓練讓新相片得到推薦結果如圖 4。

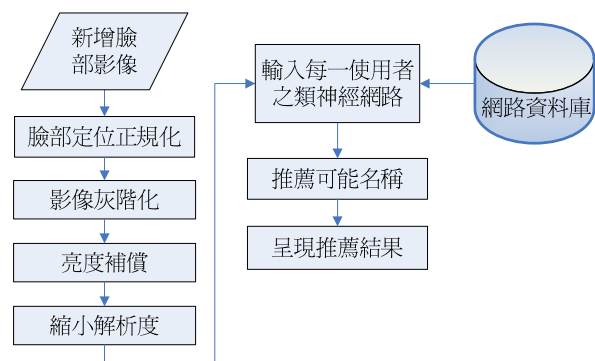


圖 4 神經網路的回想過程

3. 臉部前處理與定位正規化

從相簿中取得的臉部相片需要有一種可靠的人臉影像前處理機制，以免造成臉部影像的差異性過大，使類神經網路的學習與回想階段之結果不穩定。在此我們利用亮度補償把臉部影像的亮度調整到一定的範圍；然後使用標準臉部圖樣的臉部偵測技術重新定位出人臉的位置。

3.1 人臉定位正規化

完成亮度補償處理後，接著進行人臉的重複定位。本研究利用色彩判定技術(color-based)中的標準人臉圖樣方法來重新定位影像中的人臉。在標準人臉的製作中，需要一張五官輪廓較明顯的 64*64 像素的人臉區域，並將該人臉區域縮小為 16*16 像素即可獲得標準人臉圖樣，如圖 5 所示。製作標準人臉時應該要注意左右兩邊的對稱性，並且避免左右傾斜，必要時應該用手工略加修飾[4][6]。



圖 5 標準人臉圖樣

取得標準人臉圖樣後，從輸入影像中找尋與這標準影像最像的某幾個區域。因為標準人臉是 16*16 的影像，所以可能人臉方塊(possible block)也要縮成 16*16，然後再計算此 16*16 的方塊(block)與標準人臉的均方誤差值(mean square error, MSE)，並把誤差值最小的當作人臉區塊。如圖 6



圖 6 臉部定位正規化

使用此方法在複雜背景或是臉部亮度太過極端的影像中很容易產生錯誤。本文所使用的臉部影像均是經由 OpenCV 函式庫或是使用者所框出的影像，又經過亮度補償，不會產生複雜或亮度極端的臉部影像，所以適合用標準人臉圖樣方法來做重複定位的動作。

3.2 亮度補償

在取得的相片中每張拍攝的情況都不同，所以相片中光線的變化對於顏色的影響會相當敏感。一般在作臉部偵測時，利用膚色當作人臉的範圍是一種直覺的作法。而利用光線補償的方式則可以較準確地檢測到膚色以降低人臉偵測的誤檢率。雖然本研究不需要偵測膚色，但是若是色差相差太多將會影響實驗結果。對相片而言，不同亮度的光線會影響整張影像所呈現的結果，也就是說當光線較亮時，整體的灰階值會偏亮，反之亦然(如圖 7(a))。

為了校正臉部影像色調，對每張臉部影像的灰階值進行調整，讓每張影像的平均值為 128。假設一張影像 I 有 N 個像素， I_i 表示在影像 I 中第 i 個像素的灰階值，此影像 I 之平均值 I_{ave} 可由(1)求出。再利用(2)對影像 I 中的所有像素進行調整，已得到亮度補償過後的臉部影像[5]。

$$I_{ave} = \frac{1}{N} \sum_{i=1}^N I_i \quad (1)$$

$$I_{comp} = (128 - I_{ave}) + I_i \quad (2)$$

調整過後如圖 7 所示。



(a)



(b)

圖 7 亮度補償結果。(a)不同亮度的影像；(b)補償後的影像。

4. 臉部名稱推薦演算法

透過以上前處理，可以取得每一個使用者具有代表性的臉部影像，這些影像將會當成機率類神經網路的輸入資料進行學習。這些影像資料訓練完成後就會以使用者名稱區分為不同類別。當有新的人臉需要註釋時，新的影像進入每個機率神經網路後會產生一組機率向量，而機率向量經過比較後，就能給臉部推薦出一組適當的名稱。

4.1 機率類神經網路(probabilistic neural network, PNN)

機率神經網路是由 D.F.Specht 在 1988 年提出[8]，是一種用在分類問題上的方法。屬於

前向式的神經網路(feedforward network)架構的一種，主要的理論基礎建立在貝氏決策(Bayes decision)上。貝氏決策的理論是基於統計原理的分類觀念，主要是要找到最小的期望風險。以一個分成 k 類的情況來說明，有一 m 維的觀測值 $X = [X_1, X_2, \dots, X_m]$ ，貝氏決策的分類公式如下：

$$h_i c_i f_i(X) > h_j c_j f_j(X) \quad \forall j \neq i$$

h_k 是第 k 類的事前機率(prior probability)

c_k 代表應為第 k 類，但被誤判的損失函數

f_k 為第 k 類的機率密度函數

實際在利用貝氏決策解決分類問題時，會發現並無法事先了解該資料的機率密度函數。針對這個問題 Parzen 在 1962 年提出 Parzen 視窗法[10]來解決。Parzen 是使用一個特徵估測一個類別，對於訓練資料中的每一樣本建立一個以樣本的特徵值為中心的高斯曲線，最後把所有建立的曲線疊加成一個屬於該類別的機率密度函數。若是要用在任意維度的問題上則可以將機率密度函數表示成：

$$f_i(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sigma^d} \frac{1}{N_i} \sum_{j=1}^{N_i} \exp \left[-\frac{(x-x_{ij})^t(x-x_{ij})}{2\sigma^2} \right]$$

$f_i(x)$ = 測試向量 x 在第 i 類別的機率密度函數值

d = 訓練向量的維度

σ = 平滑參數(Soothing Parameter)

N_{ij} = 在類別 i 中的訓練向量總數

x = 測試向量

x_{ij} = 類別 i 中第 j 筆訓練資料

4.1.1 機率神經網路的設計

機率神經網路雖然是以貝氏決策為基礎，但是對機率密度函數做了下列三個假設

- 各分類的機率密度函數相同。
- 共同的機率密度函數型態為高斯分布。
- 各分類的高斯分布機率密度函數的變異矩陣維對角矩陣，且各對角元素值相同，值為 σ^2 。

因為有了這三個限制，讓機率神經網路不會有跟貝氏分類器一樣在設計上出問題。機率神經網路的網路架構為一個四層神經元的網路模型。輸入單元(input unit)、特徵單元(pattern unit)、總和單元(summation unit)及輸出單元(output unit)如圖 7 所示。

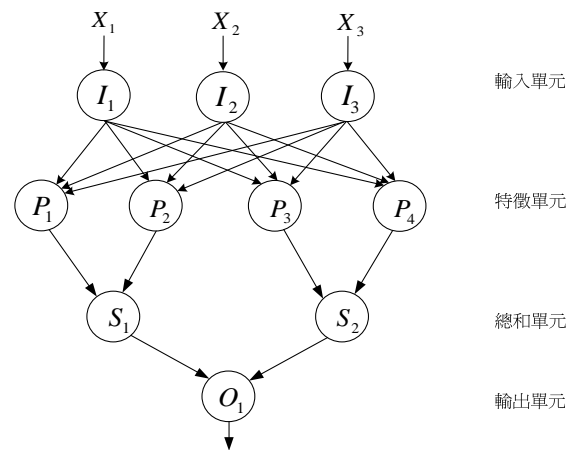


圖 7 機率神經網路架構

輸入單元可包含多個輸入影像資料，特徵單元節點各數等於訓練資料個數，輸出單元則表示分類結果。當有新影像 X 進入時，特徵單元會計算輸入影像與個別權重向量 W^{ip} 的乘積， W^{ip}

$$Z_i = X \cdot W^{ip}$$

然後對 Z_i 做非線性的轉換，雖然機率神經網路與倒傳遞網路(Back Propagation Network)一樣可以近似任何連續非線性函數，但是不同於倒傳遞類神經網路使用雙彎曲(Sigmoid)函數，在機率神經網路中是使用高斯函數(gauss activation function)。

$$P(Z_i) = \exp \left[-\frac{(Z_i - 1)^2}{\sigma^2} \right]$$

若是 X 和 W^{ip} 都已正規化到單位長度則可以簡化為

$$P_i(X) = \exp \left[-\frac{(X - W^{ip})^t(X - W^{ip})}{2\sigma^2} \right]$$

總和單元將從 N_i 個特徵單元中得到的這些向量相加取平均以產生作為一個淨輸出的機率向量。

$$S_i(X) = \frac{1}{N_i} \sum_{p=1}^{N_i} \exp \left[-\frac{(X - W^{ip})^t(X - W^{ip})}{2\sigma^2} \right]$$

最後，一個競爭(competete)轉移函數從總和單元的輸出上挑選一個或若干個最大值，令其為 1，表示屬於某一種類；其餘為 0，表示屬於其他種類。再把每個處理結果組成輸出向量，以此結果作為輸出單元的輸出 [13]。

$$O(X) = \arg \max(S_i(X)), \quad i = 1, 2, \dots, m$$

4.1.2 機率神經網路的特色

機率神經網路最重要的特色包括[7][10]：

- 快速的學習機制，可直接從訓練資料設定網路連結權重值(weight)。
- 結構上，具有平行計算架構，網路內部沒有回饋到輸入單元的鍵結，只需一次設定。
- 網路的連結個數、特徵單元節點隨訓練資料做變動。
- 學習與回想過程無關。

機率神經網路的結構可以隨分類問題任意擴充輸入及輸出的維度，當一般神經網路面對比原來更複雜的問題時，隨著神經元的增加各層之間的神經元鍵結也會跟著呈指數形式增加，若是神經網路還包含回饋機制，其計算複雜度會大幅增加。機率神經網路的網路鍵結是隨著神經元數目呈線性增加，加上分散式的結構使得計算複雜度大幅降低。

4.2 改良機率神經網路

機率神經網路可以有效地處理分類的問題，但其結果太過強勢，會使得推薦的錯誤率增加。而 Zaknick 在 1995 年提出了改良式機率神經網路架構 [14]，將機率神經網路結合重心法使神經網路的輸出是呈現一個類似圖與圖之間的相似度

在網路架構方面，改良式機率神經網路和機率神經網路的基本架構是相同的，主要的差異在於輸出的方式不同。機率神經網路的輸出方式是採取「勝者全拿」的競爭式輸出方式，其輸出結果會是所有名稱類別中的一類，所以，機率神經網路較適合處理分類的問題。但是在推薦的機制上不適合將新的相片直接指定屬於哪一類，所以去改進機率神經網路。而改良式機率神經網路的輸出方式則不同於機率神經網路競爭化的輸出方式。改良式機率神經網路架構為了解決傳統機率神經網路架構在處理函數逼近上的限制，改良式機率神經網路在輸出層利用「重心法」取代了原本的「勝者全拿的」輸出方式如圖 8。

上圖表示機率神經網路的基本架構，在特徵單元中的 $f_i(x)$ 是代表一個高斯函數，而在輸出單元中的 y_i 代表每一個由特徵單元輸出的值。採用了此種方法，使得改良式機率神經網路能夠產生新增影像與每個類別之間的推薦值。改良機率神經網路主要網路結構與傳統

機率神經網路相同，因此改良式機率神經網路保有原本機率神經網路所擁有的優點，不論是快速學習，還是可以進行增長勢的學習而不會改變已存在的網路架構。所以，此方法也適合使用在網路相簿這種需要快速反應的操作介面上，不僅是學習過程可以快速完成，在新增資料時也不會改變網路架構並且完成運算。

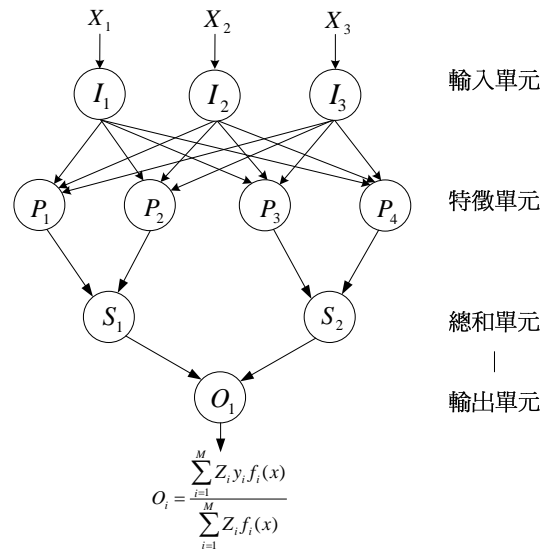


圖 8 改良機率神經網路架構

5. 實驗結果與分析

本研究的實驗樣本是由 HiAnts 網路相簿所取得[9]，對其中的使用者相簿為主要研究對象。在這本相簿中共有 5110 張自動或手動框出未註釋及前處理的臉部影像，其中有 2697 張為註釋過名稱的相片。圖 9 為使用者資料庫中所取得的相片，影像大小為 94x94 其內容包括自動及手動取得的相片。



圖 9 臉部資料庫中部份影像，由上而下的暱稱為 ccn、pili、fei、rabbit、hous。

網路相簿在註釋的時候可能會有多個使用者同時對同一張相片註釋或是一張相片有多個名稱，經過整理後剩下 1392 張真正用來做實驗的影像。在這些影像中取人臉註釋前 5 多的來當做實驗的分類樣本，之所以取 5 類是因為之後的類別臉部分布可能會少到連神經網路的訓練樣本都不夠。決定分類樣本後，分別從 5 類樣本中取出 10 張影像作為訓練樣本如圖 10，另外在從這 5 類中挑出未經註釋過名稱的相片，作為測試樣本如圖 11。



圖 10 訓練樣本類別



圖 11 測試樣本類別

實驗步驟分為影像前處理、訓練神經網路的學習過程及最後推薦臉部分名稱的回想過程三階段。前處理部分包括讀取影像、壓縮影像、將資料分為 5 回合訓練資料集。學習過程部份是將訓練資料輸入機率神經網路，讓神經網路學習如。回想過程部分是輸入測試資料至機率神經網路，紀錄推薦結果，重複學習過程直到每個訓練資料都執行過，最後再觀察結果並統計推薦率。

實驗的方法首先每一類取出 10 張影像作為類神經網路學習過程中的訓練資料集，再從這 5 類中隨機選出 100 張未經推薦的影像當作 5 個神經網路的測試資料集，每次挑出 50 張影像作測試共做 5 次觀察其結果並紀錄人臉分布及推薦成功率如表 1。

表 1 100 影像隨機取出 50 影像做推薦

實驗次數	人臉分布					錯誤張數	成功率
	ccn	pili	fei	rabbit	horse		
1	11	8	9	12	10	15	70%
2	8	11	11	10	10	14	72%
3	8	7	10	11	14	12	76%
4	9	11	6	10	14	11	78%
5	8	12	9	10	11	7	86%

由表 1 可以了解對於新資料的推薦成功率都有在 7 成以上。接下來在 5 個類別中分別對每類影像取出 20 張影像作測試，為了要觀察單一影像在不同的神經網路中被推薦的程度如何如表 2

表 2 單一類別測試結果

實驗類別	20 張影像在不同神經網路的推薦平均值					推薦成功率
	ccn	pili	fei	rabbit	horse	
ccn	0.78	0.70	0.73	0.66	0.75	75%
pili	0.68	0.86	0.72	0.75	0.76	100%
fei	0.72	0.75	0.76	0.73	0.75	30%
rabbit	0.63	0.74	0.70	0.78	0.70	85%
horse	0.73	0.74	0.73	0.69	0.79	75%

觀察表 2 可以發現在自己的相片在自己的神經網路中會表現較好，而其推薦成功率就會有一定的正確率。但是再 fei 這類中，平均值過於接近，而且推薦成功率明顯偏低，其原因可能包括臉部分布表情太多、臉部分位置不同等也因此使得此類中的臉無法得到較好的推薦。測試不同的訓練各數對於 PNN 的訓練結果所產生的影響，將影像壓縮分別為 16x16、32x32、64x64 的大小。如表 3。

表 3 不同大小推薦結果

不同影像大小	人臉訓練張數	測試張數	正確推薦張數	推薦率
16x16	50 張	50 張	41	82%
32x32	50 張	50 張	38	76%
64x64	50 張	50 張	37	74%

由表 3 可知訓練影像的各數會直接影響到推薦的正確率，因為機率神經網路的輸入單元及權重都是由整張影像輸入後所得，所以當訓練樣本數越多表示所需處理的變因就更多，所以會使神經網路的正確率有所變動。

本研究採用改良型機率神經網路的技術實做了3種不同類型的實驗，分別在隨機影像的訓練、單一影像的訓練及不同影像大小的訓練中做新影像的推薦。在隨機抽取的實驗中可以觀察到，陌生的影像在神經網路中所得到的正確的推薦率在7成以上，而且若是增加推薦數量，其推薦結果就能更正確。單一影像的實驗結果雖然有一類推薦的成功率較低，但是觀察其推薦平均值可以發現每類相差不大，就好比fei這類是一個大眾臉。雖然以肉眼來看其五官輪廓跟其他類也有很多不同之處，但是對於神經網路而言該類的臉就會被推薦成其他類的臉。在不同影像大小得實驗中雖然16x16的推薦率較高不過觀察正確推薦張數則可以比較出其實3類大小推薦結果都相當不錯，而16x16的優勢是在於其運算的速度會較快。

6. 結論

在本論文中，我們提出應用改良型機率神經網路對相簿中之人臉影像推薦名稱，讓人臉與姓名產生關連，增加網路相簿的管理與檢索功能。本系統利用標準人臉圖樣方法取出新增影像的臉部影像後，使用亮度補償產生一個標準影像。這些標準人臉影像輸入改良型機率神經網路加以訓練，可以建立起每個使用者既有的相片與其名稱的關連性。訓練完成後，若有新增影像，就可以輸入到神經網路中對新進人臉影像回想，推薦最適合的名稱。

使用機率神經網路最主要的理由為：訓練時間快速、整個網路架構適合應用在網路相簿這種需要及時反應的地方。不過會影響神經網路的推薦結果就是訓練樣本的選擇，當產生大眾臉的類別時，或許可以去搜集其他也是大眾臉的臉部影像去做錯誤辨識訓練。雖然本研究經過前處理將影像盡量保持一定的完整性，不過若是可以增加不同的特徵當作輔助一起進入神經網路訓練，則必然可以讓推薦的結果更具穩定。

7. 參考文獻

- [1] 創市際市場研究顧問 2006/09/28 『DC紀錄影像新生活網路相簿分享全世界』，http://www.insightexplorer.com/news/news_09_28_06.html123
- [2] 創市際市場研究顧問 2005/04 『你玩網路相簿咩？』，http://www.insightexplorer.com/specialtopic/ku_so_0610_05.html123

- [3] 人臉識別技術將為北京奧運會提供安全支持，<http://www.eepw.com.cn/article/62775.htm>
- [4] 陳必衷，<http://pages.cs.wisc.edu/~beechung/facenet/report.html> 123
- [5] 黃泰祥，“具備人臉追蹤與辨識功能的一個智慧型數位監視系統”，私立中原大學/電子工程學系，**全國碩士論文**，民國93年。
- [6] 孫國庭，“應用類神經網路於人臉辨識之身分識別系統”，國立高雄第一科技大學，**全國碩士論文**，民國92年。
- [7] 羅強華，類神經網路-MATLAB的應用，高立圖書有限公司，2005。
- [8] D.F.Specht，“Probabilistic Neural Network for Classification, Mapping, or Associative Memory,” *Proceeding of the IEEE International Conference on Neural Networks*, Vol. 1, pp.525-532,1988.
- [9] HiAnts創意人臉相簿，<http://hiants.tw>
- [10] Parzen, E, “On estimation of a probability density function and mode.” *Annals of Mathematical Statistics*, Vol. 33, pp. 1065-1076, 1962.
- [11] Pentland, T. Choudhury, “Face Recognition for Smart Environments,” *Computer*, Vol. 33, no. 2, pp. 50-, Feb. 2000.
- [12] Mpiperis, I., Malassiotis, S., Strintzis, M.G., “3-D Face Recognition With the Geodesic Polar Representation”, *IEEE Trans., Information Forensics and Security*, Vol. 2, pp 537 – 547, Sept. 2007.
- [13] Whei-Min Lin, Chia-Hung Lin, and Zheng-Chi Sun, “Adaptive Multiple Fault Detection and Alarm Processing for Loop System with Probabilistic Network,” *Power delivery, IEEE Transactions on*, Vol.19, Issue:1, Jan. 2000 Pages:64-69.
- [14] Zakinch, A. “Introduction to the Modified Probabilistic Neural Network for general signal processing applications” *IEEE Transactions on Signal Processing*, 1998, Vol. 46, pp. 1980-1990.