

# MẠNG NƠN HỌC TĂNG CƯỜNG DỰA TRÊN QUYẾT ĐỊNH ĐỂ NHẬN DẠNG CHÂN DUNG

BÙI CÔNG CƯỜNG<sup>1</sup>, PHÍ ANH QUÂN<sup>2</sup>

<sup>1</sup>Viện Toán học, Viện Khoa học và Công nghệ Việt Nam

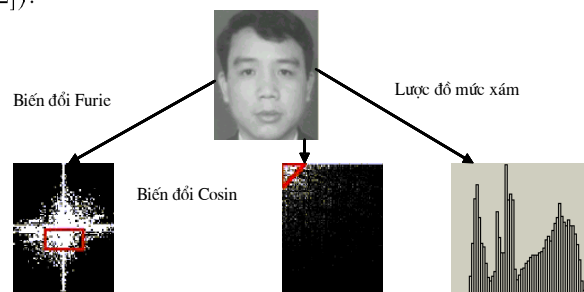
<sup>2</sup>Cục Công nghệ Tin học nghiệp vụ - Bộ Công An

**Abstract.** Decision-based neural networks (DBNNs) were originally proposed by Kung and Taur [1] for robust pattern classification. One unique feature of DBNNs is that they adopt a modular network structure. Thus, a DBNN is composed of a number of small sub-networks, with each sub-network representing one class. In this paper we proposed a modified learning algorithm with an unknown number of classes. The modified algorithm showed its effectiveness in a face recognition with simply extracted facial features. Some experimental results performed with the ORL face database and BioID Face Database should be presented.

**Tóm tắt.** Mạng nơon dựa trên quyết định được phát triển bởi Kung và Taur [1] cho bài toán nhận dạng mẫu. Một đặc trưng quan trọng của mạng này là chúng kế thừa cùng một cấu trúc mạng, tức là mạng nơon học tăng cường dựa trên quyết định bao gồm một số mạng con, mà mỗi mạng con này tương ứng với một mẫu cần phân loại. Trong bài báo này chúng tôi giới thiệu một giải thuật luyện cải tiến với số lớp không được xác định trước. Giải thuật cải tiến này chỉ ra rằng nó hiệu quả trong bài toán nhận dạng chân dung với các đặc trưng được trích chọn một cách đơn giản. Bài báo cũng giới thiệu một số kết quả thực nghiệm với Cơ sở dữ liệu là tập cơ sở dữ liệu của ORL và BioID. và cơ sở dữ liệu 18000 ảnh chân dung thu nhận từ Cục xuất nhập cảnh Bộ Công an.

## 1. CÁC PHƯƠNG PHÁP TRÍCH CHỌN ĐẶC TRƯNG KHUÔN MẶT

Trích chọn đặc trưng là một phần không thể thiếu của mỗi hệ nhận dạng với mục đích làm giảm thiểu số chiều của không gian đối tượng. Trong bài báo này, chúng tôi sử dụng một vài phương pháp thường được sử dụng trong xử lý ảnh để áp dụng cho việc trích chọn đặc trưng khuôn mặt (xem [2]).



Hình 1. Các phương pháp trích chọn đặc trưng

Giả sử trên ảnh khuôn mặt thể hiện nổi bật và to nhất kích thước  $M \times N$ . Nếu coi mỗi điểm ảnh là một đặc trưng tương thích kích thước vector đặc trưng cho toàn bộ ảnh gốc sẽ rất lớn làm cho tốc độ xử lý bị chậm đi rất nhiều. Các phương pháp trích chọn đặc trưng được minh họa như Hình 1.

### 1.1. Phương pháp biến đổi Fourier rời rạc (DFT - Discrete Fourier Transform) [3]

Sử dụng biến đổi Fourier cho phép nhận được các đặc trưng có tính chất bất biến so với sự dịch chuyển theo chu kỳ của vùng chứa khuôn mặt trong tổng thể ảnh gốc. Công thức xác định vector đặc trưng từ các biến đổi Fourier được tính toán như sau:

$$C(p, r) = \left| \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(m, n) e^{-j2\pi \frac{rn}{N}} e^{-j2\pi \frac{pm}{M}} \right|.$$

Kết quả chỉ ra rằng vector đặc trưng trở nên nhạy cảm với hình dáng khuôn mặt hơn là màu sắc của khuôn mặt. Như trong Hình 1 chúng ta thấy năng lượng quang phổ tập trung ở chính giữa ảnh và có tính chất đối xứng, vì vậy để thu được kết quả chính xác thì chúng ta cần đủ 20 giá trị quang phổ trong kết quả biến đổi Fourier, như vậy kích thước của vector đặc trưng là  $20 \times 20/2 = 200$ .

### 1.2. Phương pháp biến đổi Cosin ([4])

Ý nghĩa của sử dụng phương pháp biến đổi Cosin ở chỗ từ hệ thống các hàm biến đổi Cosin có thể xây dựng được một hàm xấp xỉ với các hàm đặc biệt trong biến đổi KLT (Karhunen Loeve). Phương pháp này cho phép thay thế các biến đổi thành phần chính (PCA - Principle Component Analysis) hay phép biến đổi KLT ([10,11]) bằng những biến đổi đơn giản hơn và thể hiện ảnh gốc bằng vector đặc trưng có số chiều nhỏ hơn rất nhiều. Như trong Hình 1 chúng ta thấy năng lượng quang phổ tập trung ở góc trái trên phía trái trong một khung có kích thước  $20 \times 20$  do vậy chiều của vector đặc trưng là 200. Công thức biến đổi Cosin như sau:

$$C(p, r) = a_p a_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X(m, n) \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)r}{2N}.$$

Cũng như phương pháp biến đổi Fourier, vector đặc trưng sẽ nhạy cảm với khung hình khuôn mặt hơn là màu sắc của da.

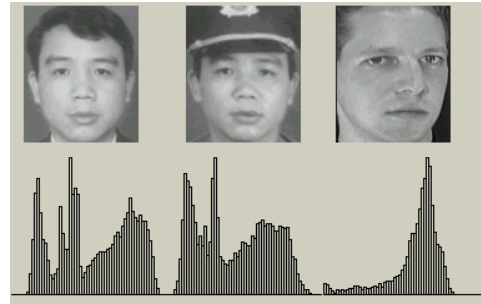
Nếu ảnh gốc thể hiện dưới dạng ma trận  $X_{M \times N}$  khi đó đặc trưng của phổ chính là kết quả của phép nhân ma trận  $T_{P \times M} X_{M \times N} T_{N \times P}$ , trong đó  $T_{P \times M}$  và  $T_{N \times P}$  là các ma trận trong biến đổi Cosin:

$$T_{P \times M} = [t_M^{(p,m)}], \quad t_M^{(p,m)} = \begin{cases} \frac{1}{\sqrt{M}}, & \text{với } p = 0, m = 0, 1, \dots, M-1, \\ \sqrt{\frac{2}{M}} \cos \frac{\pi(2m+1)p}{2M}, & \text{với } p = 1, \dots, P-1, m = 0, 1, \dots, M-1. \end{cases}$$

$$T_{N \times P} = [t_N^{(n,p)}], \quad t_N^{(n,p)} = \begin{cases} \frac{1}{\sqrt{N}}, & \text{với } p = 0, n = 0, 1, \dots, N-1, \\ \sqrt{\frac{2}{N}} \cos \frac{\pi(2n+1)p}{2N}, & \text{với } p = 1, \dots, P-1, n = 0, 1, \dots, N-1. \end{cases}$$

### 1.3. Phương pháp sử dụng biểu đồ mức xám ([5])

Trong phương pháp này giá trị các cột của biểu đồ đặc trưng cho sự phân bố theo tần số các lớp độ sáng của ảnh. Trong công việc so sánh các ảnh với nhau thì hình dạng của biểu đồ đóng vai trò rất quan trọng. Trong Hình 2 cho ta thấy sự khác nhau về biểu đồ mức sáng đối với 2 người khác nhau và sự tương đương của biểu đồ đối với cùng một người.



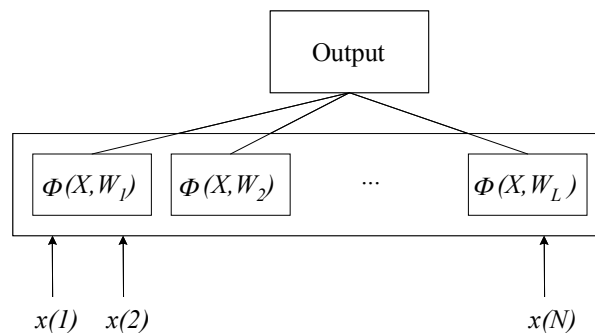
Hình 2

Tuy nhiên cần lưu ý rằng biểu đồ mức sáng chỉ minh họa cho tần số xuất hiện của giá trị độ sáng, chính vì vậy nếu như 2 bức ảnh như nhau về các vật thể trong nó nhưng được chụp ở những độ sáng khác nhau thì sẽ có hình dáng khác nhau. Chính vì vậy mà phương pháp này thường được dùng trong trường hợp ánh sáng chiếu và đối tượng có thể kiểm soát được.

## 2. MÔ HÌNH MẠNG NƠN HỌC TĂNG CƯỜNG DỰA TRÊN QUYẾT ĐỊNH

Với các đặc trưng có được, hiện nay có rất nhiều phương pháp hiệu quả được sử dụng như nhận dạng theo cấu trúc, nhận dạng theo mô hình Markov, và nhận dạng theo mô hình mạng nơron. Trong đó mô hình mạng nơron đã được sử dụng khá hiệu quả để nhận dạng khuôn mặt với nhiều mô hình mạng được đề xuất. Mô hình mạng nơron học tăng cường dựa trên quyết định với cơ chế kế thừa cùng một cấu trúc mạng kết hợp với những cải tiến về giải thuật luyện cho phép mở rộng số lớp cần phân loại đã cho những kết quả tốt đối với bài toán nhận dạng chân dung.

Mạng nơron học tăng cường dựa trên quyết định (DBNN) nguyên thủy được phát triển bởi Kung và Taur cho bài toán phân loại mẫu thô. Một đặc trưng cơ bản của các mạng nơron dựa trên quyết định là chúng kế thừa cùng một cấu trúc mạng, tức là mạng nơron học tăng cường dựa trên quyết định bao gồm một số mạng con, mà mỗi mạng con này tương ứng với một mẫu cần phân loại. Cấu trúc của mạng DBNN được minh họa trong Hình 3, trong đó mỗi mạng con được đặc trưng bằng một hàm toán học (basic function) của vectơ tín hiệu vào  $X$  và ma trận trọng số liên kết  $W_1$  được ký hiệu là  $Net$ ,  $Net = \Phi(X, W_1)$ .



Hình 3

## 2.1. Một số dạng hàm Net

### a. Hàm tổng (Linear Basic Function)

Một cách biểu diễn Net đơn giản nhất là tổng các tích của tín hiệu vào  $x_i$  và các trọng số  $w_i$  tương ứng:

$$Net = \Phi(X, W_l) = \sum_{i=1}^N x_i w_{li},$$

trong đó  $N$  là số chiều của tín hiệu vào và  $l$  là chỉ số của mạng con thứ  $l$ . Đây là cách biểu diễn đơn giản nhất và cũng hữu ích khi chúng ta xây dựng một mạng với nhiều nơron cho bài toán khả tách tuyến tính.

### b. Hàm cơ sở xuyên tâm (Radial Basic Function - RBF)

Net được biểu diễn là một hàm của khoảng cách giữa tín hiệu vào và nơron trung tâm (trong đó nơron trung tâm được hiểu theo nghĩa gần với tín hiệu vào nhất):

$$Net = \Phi(x, w_l) = -\frac{\|x - w_l\|^2}{2}.$$

Tương tự như giải thuật luyện Kohonen[6], hàm RBF đặc biệt tỏ ra hữu ích đối với các bài toán phân loại với các đường ranh giới phi tuyến.

### c. Hàm cơ sở elip (Elliptic Basis Function - EBF)

Cơ sở của việc áp dụng hàm RBF là giả thiết rằng không gian đặc trưng là độc lập tuyến tính và đồng nhất theo tất cả chiều của không gian đặc trưng. Tuy nhiên trong thực tế sự khác nhau giữa các tập đặc trưng được có nhiều mức độ phụ thuộc vào khái niệm độ đo khoảng cách giữa 2 vectơ nào được áp dụng. Vì vậy hàm EBF cho mỗi lớp mạng con thứ  $l$  được biểu diễn:

$$Net = \Phi(x, w_l) = \sum_{k=1}^N \alpha_{lk} (x_k - w_{lk})^2 + \theta_l,$$

trong đó  $N$  là số chiều của mẫu vào, ma trận trọng số  $w_l$  bao gồm tất cả các tham số  $\alpha_{lk}, w_{lk}, \theta_l$ .

## 2.2. Luật học tăng cường dựa trên quyết định

Mạng nơron học tăng cường dựa trên quyết định thuộc lớp mạng học có giám sát. Thông tin chỉ đạo trong quá trình học chỉ cho biết sự phân loại đúng đối với mỗi mẫu luyện. Thông tin chỉ đạo là một tập các ký hiệu  $T = \{t_i\}$  được gán nhãn cho mỗi mẫu luyện. Mục tiêu của quá trình luyện mạng là tìm kiếm ma trận trọng số của mạng sao cho đáp ứng được với bài toán cần phân loại. Chiến lược học của mạng được minh họa như Hình 4, với mỗi mẫu luyện đi qua mạng, với việc lan truyền tín hiệu của mạng và dựa vào thông tin chỉ đạo sẽ quyết định các vectơ trọng số của mạng có cần hiệu chỉnh hay không theo 3 nguyên tắc luyện đối với luật học tăng cường dựa trên quyết định:

- Ma trận trọng số chỉ được cập nhật khi xảy ra trường hợp phân loại sai.
- Không phải toàn bộ các nơron đều được cập nhật lại trọng số mà chỉ có các nơron trong mạng con tương ứng với mẫu luyện được luyện tăng cường và các nơron trong mạng con tương ứng với việc phân loại nhằm được luyện ngược tăng cường.

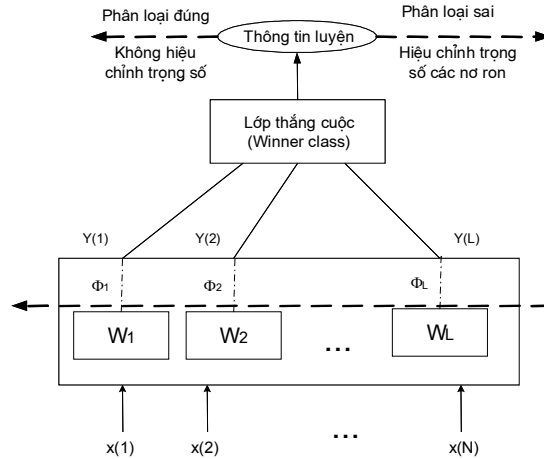
- Do đường ranh giới giữa các mẫu phụ thuộc vào hàm Net nên ma trận trọng số  $W$  được cập nhật theo hướng gradient đối với luyện tăng cường và ngược gradient đối với luyện ngược tăng cường. Có nghĩa là :

$$\Delta W = \pm \eta \nabla \phi(X, W),$$

trong đó  $\eta$  là hằng số luyện.

$$\nabla \phi(x, w) = \frac{\partial \phi(x, w)}{\partial w} = \left[ \frac{\partial \phi}{\partial w_1}, \frac{\partial \phi}{\partial w_2}, \dots, \frac{\partial \phi}{\partial w_v} \right]^T,$$

trong đó  $v$  là tổng số mẫu luyện.



Hình 4

**Thuật toán luyện**

Cho trước tập mẫu luyện  $S = \{x^{(1)}, \dots, x^{(M)}\}$  trong đó mỗi phần tử  $x^{(m)} \in R^N$  được phân loại vào một trong  $L$  lớp cho trước  $\{\Omega_i, i = 1, \dots, L\}$ , mỗi một lớp được mô hình hóa bởi một mạng con (subnet) với  $Net = \phi(X, W_i), i = 1, \dots, L$ .

Giả sử rằng mẫu luyện thứ  $m : x^{(m)}$  được biết là thuộc về lớp  $\Omega_i$  và ta có:

$$\phi(x^{(m)}, w_j^{(m)}) > \phi(x^{(m)}, w_l^{(m)}), \forall l \neq j,$$

điều đó có nghĩa là lớp  $j$  được gọi là lớp thắng cuộc. Khi đó xảy ra 2 trường hợp:

- Nếu  $i = j$ , tức là mẫu luyện  $x^{(m)}$  đã được phân loại đúng và khi đó không cần phải cập nhật lại ma trận trọng số.
- Nếu  $i \neq j$ , tức là mẫu luyện  $x^{(m)}$  bị phân loại sai, khi đó ma trận trọng số  $W$  được cập nhật lại theo qui tắc:

+ Học tăng cường: 
$$W_i^{(m+1)} = W_i^{(m)} + \eta \nabla \phi(X, W_i).$$

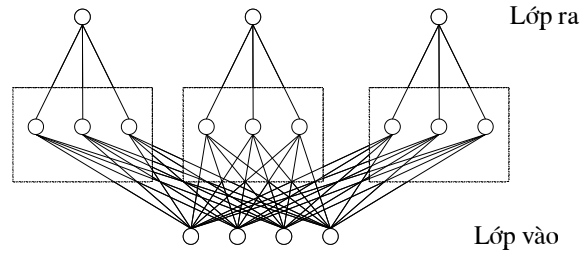
+ Ngược tăng cường: 
$$W_j^{(m+1)} = W_j^{(m)} - \eta \nabla \phi(X, W_j).$$

Chú ý rằng với mọi  $k \neq i$  và  $k \neq j$  thì  $W_k^{(m+1)} = W_k^{(m)}$ .

Tốc độ hội tụ: Theo [7], giải thuật luyện tăng cường dựa trên quyết định luôn hội tụ đến kết quả tối ưu sau một số lần lặp hữu hạn.

### 2.3. Cấu trúc mạng

Chúng ta đã quen với cấu trúc mạng lớn được thiết kế cho tất cả các lớp cần phân loại được gọi là cấu trúc “Tất cả các lớp trên cùng một mạng” (All-Class-in-One-Net: ACON). Trong cấu trúc này, nơron trên lớp ra được nối đầy đủ với các nơron trên lớp liền kề, vì vậy đối với cấu trúc này thường bùng nổ kết nối khi số lớp cần phân loại lớn dẫn đến tình trạng kéo dài tốc độ ổn định của mạng trong quá trình luyện mạng.



Hình 5

Một đặc trưng của cấu trúc mạng học tăng cường dựa trên quyết định là cấu trúc “mỗi lớp một mạng” (One-Class-One-Net: OCON). Trong Hình 5, chúng ta có thể thấy sự giảm rõ rệt các kết nối của cấu trúc OCON so với cấu trúc ACON. Bước tiếp theo, mỗi mạng con được xây dựng từ nhiều nơron, có 2 cấu trúc đối với mạng nơron học tăng cường dựa trên quyết định được nghiên cứu là cấu trúc nút ẩn (Hide-node) như Hình 6a và cấu trúc subcluster như Hình 6b. Để thuận tiện cho việc đánh chỉ số đối với cấu trúc mạng chúng ta gán chỉ số 1 cho mỗi mạng con và chỉ số  $k_l$  cho các nơron trong cùng một lớp  $l$ . Khi đó hàm mô hình của mạng cho mạng con thứ  $l$   $Net_l = \phi(x, w_l)$  và đối với mỗi nơron  $k_l$  trong mạng con thứ  $l$   $Node_{k_l} = \varphi(x, w_{k_l})$  trong đó  $k_l = 1, \dots, K_l$  và  $K_l$  là số nơron trong mạng con thứ  $l$ .

Trong cấu trúc nút ẩn, hàm net được xác định bằng tổng các net của các nơron được trọng số:

$$\phi(x, w_l) = \sum_{k_l=1}^{K_l} c_{k_l} \varphi(x, w_{k_l}),$$

trong đó  $\{c_{k_l}\}$  được xác định là các trọng số của nơron của lớp ra đối với mỗi mạng con. Tín hiệu ra của mỗi mạng con được xác định nhờ các hàm kích hoạt:  $Out = f(net)$  với các hàm thường được sử dụng như hàm dấu, hàm sigmoid và hàm Gaussian.

Hàm chữ S (Sigmoid function),

$$f(u_i) = \frac{1}{1 + e^{-u_i/\sigma}}.$$

Hàm Gaussian,

$$f(u_i) = e^{-\frac{u_i^2}{\sigma^2}}.$$

Trong cấu trúc subcluster, thay vì sử dụng các trọng số của nơron trên lớp ra, chúng ta sử dụng giải pháp “người thắng được tất” (winner-take-all), tức là chỉ có nơron thắng cuộc được gán trọng số liên kết là 1 còn lại đều có giá trị 0. Trong giải pháp này, sử dụng 2 khái niệm Local winner và Global winner, trong đó Local winner là nơron thắng cuộc trong cùng một mạng con. Local winner của mạng con thứ  $l$  được đánh chỉ số  $s_l$

$$s_l = \text{Arg max}_{k_l} \varphi(x, w_{k_l}).$$

Global winner là mạng con thắng giữa tất cả các mạng con khác. Mạng con thứ  $j$  sẽ được gọi là Global winner nếu như nó là Local winner và lớn hơn tất cả các local winner khác, có nghĩa là:

$$\varphi_j(x, w_{s_j}) > \varphi_l(x, w_{s_l}), \forall l \neq j.$$

Tín hiệu vào sẽ được phân loại là thuộc vào lớp thứ  $j$  nếu như mạng con thứ  $j$  là Global winner.

Đối với cấu trúc subcluster, giải thuật luyện tăng cường được áp dụng với việc thay thế

$$\phi(x, w_i) \Leftrightarrow \varphi_i(x, w_{s_i}).$$

**Giải thuật:**

Một mô hình mạng nhiều mạng con với các  $s_i, s_j \dots$  là các local winner tương ứng. Giả sử rằng mẫu luyện  $x^{(m)}$  được biết là thuộc về lớp  $\Omega_i$  nhưng mạng con thứ  $j$  lại được chọn là global winner. Nếu như  $i \neq j$ , tức là mẫu  $x^{(m)}$  bị phân loại sai, khi đó các trọng số được hiệu chỉnh theo qui tắc:

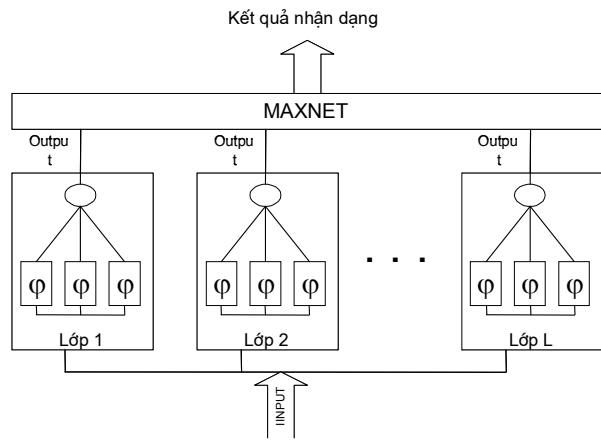
$$W_{s_i}^{(m+1)} = W_{s_i}^{(m)} + \eta \nabla \phi(X, W_{s_i}).$$

$$W_{s_j}^{(m+1)} = W_{s_j}^{(m)} + \eta \nabla \phi(X, W_{s_j}).$$

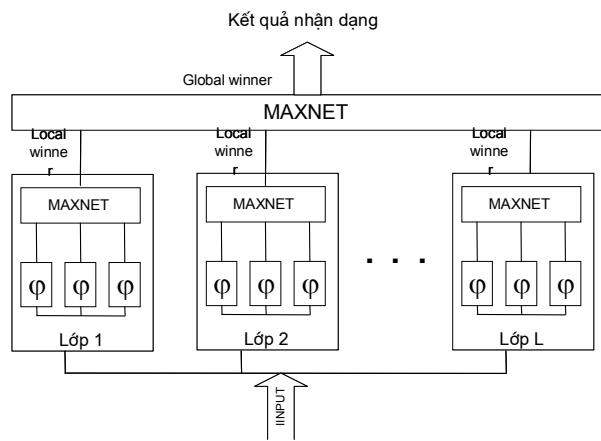
Nói cách khác, luyện ngược tăng cường được áp dụng cho mạng con global winner và luyện tăng cường được áp dụng cho mạng con local winner đúng với lớp cần phân loại.

*Hằng số luyện  $\eta$*

Ngoài việc các giá trị trọng số được khởi tạo ban đầu một cách ngẫu nhiên có ảnh hưởng tới quá trình luyện mạng thì một nhân tố khác ảnh hưởng tới hiệu lực và độ hội tụ của giải thuật luyện tăng cường dựa trên quyết định là hằng số luyện  $\eta$ . Không có một giá trị duy nhất phù hợp với các bài toán khác nhau. Hằng số luyện thường được chọn bằng thực nghiệm cho mỗi bài toán ứng dụng cụ thể bằng phương pháp thử sai. Hằng số luyện  $\eta$  lớn làm tăng tốc quá trình hội tụ. Điều này không có lợi vì thủ tục học sẽ kết thúc nhanh tại một cực tiểu cục bộ gần nhất. Ngược lại nếu hằng số luyện quá nhỏ thì tốc độ hội tụ của giải thuật trở nên rất chậm.



Hình 6a



Hình 6b

Trong bài này chúng tôi thử nghiệm cài đặt với hằng số luyện thích nghi tương ứng với các tình trạng luyện của mạng. Hằng số luyện sẽ tăng hoặc giảm tùy theo tốc độ hội tụ và hiệu quả của quá trình luyện. Tiêu chuẩn để đánh giá hiệu quả của mạng là số mẫu nhận đúng tăng lên hay giảm đi sau một số lần luyện nhất định. Nếu như số mẫu luyện được phân loại đúng giảm đi thì hằng số luyện cần tăng lên, ngược lại nếu như số mẫu luyện được phân loại đúng tăng lên một cách đột biến thì hằng số luyện phải được giảm đi, nhằm giảm tốc độ hội tụ của mạng. Để đảm bảo tránh tình trạng hằng số luyện dao động xung quanh giá trị được chọn, việc thay đổi hằng số luyện được thử nghiệm như sau:

$$\eta = \begin{cases} \alpha\eta & \text{khi số mẫu sai tăng lên với } 0,4 \leq \alpha \leq 0,6 \\ \beta\eta & \text{khi số mẫu sai giảm nhiều với } 2,8 \leq \beta \leq 3 \\ \eta & \text{trong các trường hợp còn lại.} \end{cases}$$

### 3. NHỮNG CẢI TIẾN TRONG GIẢI THUẬT LUYỆN VỚI SỐ LỚP KHÔNG XÁC ĐỊNH

Khi mặt phẳng phân cách giữa các lớp là rõ ràng và số lớp đã được xác định trước, chúng ta có thể dễ dàng xác định được ma trận trong số tối ưu cho bài toán phân loại. Để tăng bờ phân cách giữa các lớp, chúng tôi đề xuất sử dụng khái niệm khoảng cách trung bình giữa các lớp  $\varepsilon$  để đánh giá việc lựa chọn lớp thắng cuộc

$$\phi(x^{(m)}, w_j^{(m)}) > \phi(x^{(m)}, w_i^{(m)}) + \varepsilon, \quad \forall i \neq j, \quad (*)$$

khi đó  $j$  mới được chọn là lớp thắng cuộc. Trong đó  $\varepsilon$  được xác định

$$\varepsilon = \frac{2}{N(N-1)} \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} \|x_i - x_j\|,$$

với  $x_i, x_j$  là các vectơ đặc trưng đại diện cho mỗi mẫu thử tiêu biểu cho mỗi lớp, trong nhận dạng chân dung đó là các ảnh chân dung chụp chính diện tiêu biểu cho mỗi đối tượng. Trong trường hợp bài toán số lớp không xác định trước, giải thuật được cải tiến như sau:

B1. Xác định một số lớp nào đó cần thiết để cấu trúc xây dựng mạng. Tiến hành luyện mạng với số lớp đối tượng cố định.

B2. Tính khoảng cách trung bình  $\varepsilon$  giữa các lớp đối tượng đã chọn. Tiếp tục thử tục luyện tăng cường và ngược tăng cường với mục đích cực đại hóa bờ phân cách giữa các lớp. Tại vòng lặp thứ  $k$ , nếu (\*) không được thỏa mãn giảm  $\varepsilon = 0,6\varepsilon$ .

B3. Bổ sung một lớp mới bằng cách thêm một mạng con có cấu trúc như các mạng con khác đã cài đặt cho các lớp trước. Áp dụng giải thuật luyện tăng cường cho mạng con với mỗi mẫu phân loại sai

$$W_{s_i}^{(m+1)} = W_{s_i}^{(m)} + \eta \nabla \phi(X, W_{s_i}).$$

Giải thuật luyện dừng khi với mẫu vào thuộc lớp mới, tiêu chí chọn lớp thắng thỏa (\*).

### 4. CƠ SỞ DỮ LIỆU THỬ NGHIỆM

Ngoài một số ảnh chân dung trực tiếp thu nhận, các kết quả được thử nghiệm trên tập ảnh của hãng ORL và BioID với ảnh chân dung của 40 người, mỗi người 10 kiểu ảnh chụp



khác nhau, trong đó có 3 ảnh được sử dụng minh họa cho phương pháp trích chọn đặc trưng sử dụng lược đồ mức xám ([8,9]).

Bảng 1. Kích cỡ vector đặc trưng

Phương pháp/ Kích cỡ	Biến đổi Fourier	Biến đổi Cosine	Lược đồ mức xám
Tham số	20×10	20×20	256
Kích cỡ	200	200	256

Tất cả các ảnh đều được lưu dưới dạng ảnh 256 mức xám chuẩn BMP có kích thước 92×112. Kích thước của vector đặc trưng đối với từng phương pháp trích chọn đặc trưng được cho trong Bảng 1.

Đối với mỗi phương pháp trích chọn đặc trưng chúng tôi đánh giá kết quả thử nghiệm số lượng mẫu học và số lượng mẫu thử trong tổng số 10 ảnh của mỗi đối tượng. Kết quả thử nghiệm được cho trong Bảng 2 với việc chọn ngẫu nhiên một số lượng ảnh làm mẫu luyện và mẫu thử.

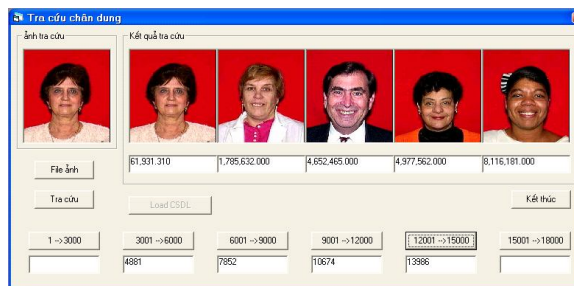


Hình 7

Bảng 2. Kết quả nhận dạng

Phương pháp	2/8	3/7	4/6	5/5	6/4	7/3	8/2
Biến đổi Furier	65%	82%	85%	89%	91%	96%	98%
Biến đổi Cosine	68%	83%	87%	90%	91%	98%	100%
Lược đồ mức xám	68%	84%	88%	90%	95%	97%	100%

Trong kết quả trên ta thấy phương pháp trích chọn đặc trưng dựa trên lược đồ mức xám cho kết quả khá cao, tuy nhiên kết quả trên chỉ phù hợp với cơ sở dữ liệu của ORL, đối với những bức ảnh tự thu nhận, do sự đa dạng về độ sáng tối nên kết quả không tốt bằng. Một phương pháp thường được sử dụng với phương pháp này là trước khi xử lý, ảnh được cân bằng sáng, kết quả đạt được cũng gần tương đương.



Bảng 3

Mô hình	Tỷ lệ lỗi	T.gian phân loại	T.gian luyện
Mạng nơ ron dựa trên quyết định xác suất	4%	0,8 giây	35 phút
Mạng nơ ron ánh xạ tự tổ chức	3,8 %	0,5 giây	4 giờ
Mạng nơ ron dựa trên quyết định	4%	0,8 giây	1 giờ

Đối với cơ sở dữ liệu 18000 ảnh chân dung thu nhận được từ Cục xuất nhập cảnh Bộ

Công an, giao diện chương trình được thiết kế như hình trên với ảnh bên trái là ảnh tra cứu và bên phải là danh sách các ảnh tra cứu được sắp xếp theo thứ tự giống với ảnh tra cứu nhất dựa trên ngưỡng sai khác với ảnh tra cứu. Các kết quả thử nghiệm so với một số mô hình khác được đánh giá bởi Bảng 3.

Các kết quả trên được thực hiện với cùng một phương pháp biến đổi cosine trên cùng một cơ sở dữ liệu 18.000 ảnh

## 5. KẾT LUẬN

Từ các kết quả đạt được bằng các phương pháp trích chọn đặc trưng đơn giản và mô hình mạng nơron học tăng cường dựa trên quyết định cho thấy mô hình đơn giản và tốc độ xử lý nhanh, hiệu quả (Phương pháp biến đổi Fourier có tốc độ xử lý chậm nhất). Do cấu trúc mạng đơn giản nên việc sử dụng tài nguyên của máy tính tăng không quá lớn nên có thể mở rộng kích cỡ của cơ sở dữ liệu. Một tính chất quan trọng khác là khi bổ sung một đối tượng mới vào trong cơ sở dữ liệu, hệ thống không phải tổ chức lại cấu trúc của mạng mà chỉ phát triển thêm một mạng con mới do vậy mạng phải không học lại nên số lượng đối tượng không bắt buộc phải xác định trước, phù hợp với bài toán nhận dạng chân dung trong thực tế. Tuy nhiên thời gian luyện mạng vẫn chưa được tối ưu so với mạng cùng loại được phát triển dựa trên xác suất, vì vậy hướng nghiên cứu tiếp theo của chúng tôi là mô hình mạng nơron dựa trên quyết định xác suất cho bài toán nhận dạng chân dung.

## TÀI LIỆU THAM KHẢO

- [1] S. Y. Kung and J. S. Taur, Decision-based neural networks with signal/image classification applications, *IEEE Trans. on Neural Networks* **6** (1) (1995) 170–181.
- [2] Geory Kakharev and Adam Nowoseilski, “Simple Facial Features Extrator”, Machine Vision and Applications manuscript, 2005.
- [3] Nguyễn Thanh Thủy, *Nhập môn xử lý ảnh*, Nhà xuất bản Khoa học và Kỹ thuật, Hà nội, 2004.
- [4] Z. M. Hafed, M. D. Lewine, Face recognition using the Discrete Cosine Transform, *Inter Journal of Computer Vision* **43** (3) (2001) 167–188.
- [5] M. Tistarelli, E. Grosso, Active face recognition with a hybrid approach, *Pattern Recognition Letters* **18** (1997) 933–946.
- [6] T. Kohonen, Self-Organization and Associative Memory, *Series in Information Science* **8** Springer-Verlag, New York, 1984.
- [7] N. J. Nilsson, *Learning Machines*, New York, McGaraw-Hill, 1965.
- [8] The ORL Database of Faces <http://www.uk.research.att.com>
- [9] The BioID Face Database <http://www.bioid.com/research/index.htm>
- [10] Bùi Công Cường, Một số lớp mạng nơron nhân tạo và ứng dụng vào nhận dạng, trường thu, “Hệ mờ và ứng dụng”, lần thứ 3, 24-26/8/2003, Viện Toán học, Hà Nội.
- [11] S. Theodoridis and K. Koutrumbas, *Pattern Recognition*, Academic press, London, 1999.

Nhận bài ngày 9 - 5 - 2006