

ĐẠI HỌC THÁI NGUYÊN  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

**Đông Thanh Vân**

**NGHIÊN CỨU MỘT SỐ PHƯƠNG PHÁP CƠ BẢN  
VỀ NHẬN DẠNG MẶT NGƯỜI TRONG ẢNH  
VÀ ỨNG DỤNG**

**LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH**

Giáo viên hướng dẫn: **TS. Vũ Việt Vũ**

*Thái Nguyên, tháng 07 năm 2015*

## LỜI CAM ĐOAN

Tôi xin cam đoan đề tài "**Nghiên cứu một số phương pháp cơ bản về nhận dạng mặt người trong ảnh và ứng dụng**" là công trình nghiên cứu được tôi thực hiện dưới sự hướng dẫn của giáo viên hướng dẫn khoa học.

Các kết quả nghiên cứu và kết quả thử nghiệm nêu trong luận văn là trung thực và chưa từng được công bố trong bất kỳ tài liệu nào khác. Trong phần kiến thức chung, nghiên cứu giải thuật áp dụng tôi có tham khảo ở một số tài liệu và đã có trích dẫn đúng và đầy đủ.

**Học viên**

## MỤC LỤC

<b>LỜI CAM ĐOAN</b> .....	1
<b>LỜI MỞ ĐẦU</b> .....	4
<b>CHƯƠNG 1. TỔNG QUAN</b> .....	8
1.1 Giới thiệu về trí tuệ nhân tạo và học máy .....	8
1.1.1 Giới thiệu.....	8
1.1.2 Ứng dụng của lĩnh vực học máy .....	10
1.2 Tổng quan về xử lý ảnh số và ứng dụng .....	11
1.2.1 Giới thiệu.....	11
1.2.2 Những khái niệm cơ bản trong xử lý ảnh .....	13
1.3 Một số bài toán trong xử lý ảnh .....	18
1.3.1 Tăng cường ảnh.....	18
1.3.2 Phân vùng ảnh .....	19
1.3.3 Trích chọn đặc tính.....	20
1.3.4 Nén ảnh .....	21
1.3.5 Phát hiện biên ảnh .....	22
1.3.6 Khôi phục ảnh .....	24
<b>CHƯƠNG 2. NGHIÊN CỨU PHƯƠNG PHÁP PCA VÀ PHƯƠNG PHÁP SVM TRONG BÀI TOÁN NHẬN DẠNG MẶT NGƯỜI</b> .....	28
2.1 Tổng quan.....	28
2.1.1 Giới thiệu.....	28

2.1.2 Bài toán nhận dạng đối tượng .....	29
2.2 Bài toán học có giám sát .....	33
2.3 Phương pháp PCA và ứng dụng cho bài toán nhận dạng mặt người	34
2.3.1 Giới thiệu về phương pháp PCA.....	34
2.3.2 Ví dụ về phương pháp PCA.....	36
2.3.3 PCA cho bài toán nhận dạng mặt người .....	37
2.4 Phương pháp SVM.....	41
2.4.1 Trường hợp dữ liệu có thể phân tách được .....	42
2.4.2 Trường hợp dữ liệu huấn luyện không phân tách tuyến tính.....	46
2.4.3 Phương pháp SVM cho bài toán nhiều lớp.....	48
<b>CHƯƠNG 3. KẾT QUẢ THỰC NGHIỆM</b> .....	<b>50</b>
3.1 Giới thiệu bài toán.....	50
3.2 Lựa chọn giải thuật và lập trình .....	51
3.2.1 Giải thuật.....	51
3.2.2 Công cụ phát triển .....	51
3.3 Kết quả thực nghiệm .....	52
<b>CHƯƠNG 4. KẾT LUẬN</b> .....	<b>57</b>
<b>Tài liệu tham khảo</b> .....	<b>59</b>

## LỜI MỞ ĐẦU

### 1. Tính cấp thiết của đề tài

Ngày nay, công nghệ thông tin là một trong những lĩnh vực quan trọng góp phần thúc đẩy sự phát triển kinh tế của đất nước. Ứng dụng của công nghệ thông tin (CNTT) xuất hiện ở rất nhiều lĩnh vực, nó là nền tảng của các nghiên cứu về công nghệ cao như khoa học vũ trụ, công nghệ điện tử và tự động hóa, các hệ thống thông minh,... Khi Internet ra đời vào những năm 90 đã kết nối thông tin trên toàn thế giới lại với nhau. Chúng ta chỉ cần ngồi bất cứ nơi đâu với máy tính và mạng là có thể truy nhập thông tin về bất cứ lĩnh vực gì trên thế giới.

Tại Việt Nam, từ những năm 70 của thế kỷ 20 đã bắt đầu có các nghiên cứu về máy tính cũng như những ứng dụng về CNTT. Hiện nay Công nghệ thông tin cũng luôn được xác định là lĩnh vực then chốt được đầu tư không những hiện đại hóa cơ sở hạ tầng mà còn là việc ứng dụng sâu rộng vào các lĩnh vực của đời sống xã hội.

Nghị quyết 36 ngày 1/7/2014 của Bộ chính trị tiếp tục khẳng định sự quan tâm chỉ đạo sát sao của Đảng đối với công tác ứng dụng, phát triển CNTT, khẳng định vị trí, vai trò của CNTT trong sự nghiệp xây dựng và bảo vệ Tổ quốc thời kỳ mới. Những định hướng, tư tưởng, giải pháp lớn trong Nghị quyết 36 là tiền đề quan trọng để CNTT Việt Nam phát triển mạnh mẽ, sâu rộng hơn, đề ra những định hướng lớn cho sự phát triển ngành CNTT nước ta trong khoảng 10 - 20 năm tới.

Nghị quyết 36 đã chỉ rõ mục tiêu phải đạt được thời gian tới, đó là: CNTT phải được ứng dụng rộng rãi và trở thành một ngành kinh tế có tác động lan tỏa trong phát triển kinh tế - xã hội, bảo đảm quốc phòng, an ninh, góp phần nâng cao năng suất lao động, năng lực cạnh tranh quốc gia, chất

lượng cuộc sống, chỉ số phát triển con người Việt Nam và nâng cao khả năng phòng thủ quốc gia trong chiến tranh mạng. Đến năm 2030, đưa năng lực nghiên cứu, ứng dụng, phát triển, sản xuất và cung ứng sản phẩm, dịch vụ CNTT đạt trình độ tiên tiến thế giới; Việt Nam trở thành quốc gia mạnh về CNTT trong khu vực.

Trí tuệ nhân tạo là một hướng nghiên cứu của lĩnh vực CNTT và Khoa học máy tính nhằm phát triển các hệ thống thông minh nhằm giải quyết các bài toán trong thực tế giống như hoạt động của bộ não người. Trí tuệ nhân tạo được bắt đầu nghiên cứu từ những năm 50 của thế kỷ 20 và trong khoảng 30 năm trở lại đây đã được cộng đồng các nhà khoa học quan tâm mạnh mẽ. Rất nhiều các hội thảo lớn về lĩnh vực này được tổ chức hàng năm trên thế giới. Các ứng dụng tiêu biểu của trí tuệ nhân tạo vào đời sống xã hội bao gồm: người máy, robot, xử lý ngôn ngữ tự nhiên, nhận dạng, phát hiện dị thường, an ninh quốc phòng, tin sinh học, khoa học vũ trụ và trái đất,...

Trong khuôn khổ luận văn Thạc sỹ của mình, qua việc được trang bị các môn học lý thuyết như thuật toán, xử lý ảnh, trí tuệ nhân tạo,... tôi đã lựa chọn đề tài *Nghiên cứu một số phương pháp cơ bản về nhận dạng mặt người trong ảnh và ứng dụng*. Chủ đề phân lớp dữ liệu là một nhánh quan trọng trong lĩnh vực học máy (machine learning) của trí tuệ nhân tạo. Chúng ta có thể kể các ứng dụng của học máy trong thực tế như người máy, robot, nhận dạng mặt người, nhận dạng tiếng nói và chữ viết,...

## **2. Đối tượng và phạm vi nghiên cứu.**

### *2.1. Đối tượng nghiên cứu*

Bài toán phân lớp có giám sát (classification) đặc biệt là hai phương pháp phân lớp Support Vector Machine và phương pháp PCA.

### *2.2. Phạm vi nghiên cứu*

- **Lý thuyết:** Nghiên cứu lý thuyết về xử lý ảnh, phương pháp SVM và phương pháp PCA.
- **Thực nghiệm:** Lập trình trên ngôn ngữ C# cho ứng dụng nhận dạng mặt người trong ảnh.

### 3. Phương pháp nghiên cứu.

- Thu thập, phân tích các tài liệu và thông tin liên quan đến đề tài từ đó đưa ra cái nhìn tổng quan, các khó khăn gặp phải, các ràng buộc của bài toán...
- Tiến hành phân tích, xây dựng giải pháp nhận dạng gồm có : Tiền xử lý, trích chọn đặc trưng, huấn luyện mô hình, hậu xử lý.
- Xây dựng và kiểm thử việc đánh giá hiệu quả phương pháp nhận dạng bằng ngôn ngữ C#.

### 4. Ý nghĩa khoa học và thực tiễn của đề tài.

**Về khoa học:** Giúp đưa ra một cái nhìn tổng quan về bài toán nhận dạng mặt người. Đánh giá về hiệu quả của phương pháp PCA và SVM cũng như khả năng ứng dụng của chúng.

**Về thực tiễn:** Nhận dạng mặt người là phương pháp có những ứng dụng thực tiễn mà những phương pháp khác không có được. Trong tương lai nhận dạng mặt người hứa hẹn sẽ được sử dụng phổ biến trong nhiều lĩnh vực như an ninh, phòng chống tội phạm, quản lý nhân sự...

### 5. Cấu trúc của luận văn.

Cấu trúc của luận văn *Nghiên cứu một số phương pháp cơ bản về nhận dạng mặt người trong ảnh và ứng dụng* bao gồm những chương như sau:

**Chương 1:** Tổng quan về lĩnh vực xử lý ảnh và học máy.

**Chương 2:** Nghiên cứu tìm hiểu phương pháp PCA và phương pháp SVM.

**Chương 3:** Chương này sẽ đi sâu vào việc cài đặt ứng dụng cụ thể cho bài toán nhận dạng mặt người trong ảnh bằng phương pháp PCA.

**Kết luận:** Tóm tắt các vấn đề được tìm hiểu trong luận văn và các vấn đề liên quan trong luận văn, đưa ra một số hướng nghiên cứu tiếp theo.



## CHƯƠNG 1. TỔNG QUAN

### 1.1 Giới thiệu về trí tuệ nhân tạo và học máy

#### 1.1.1 Giới thiệu

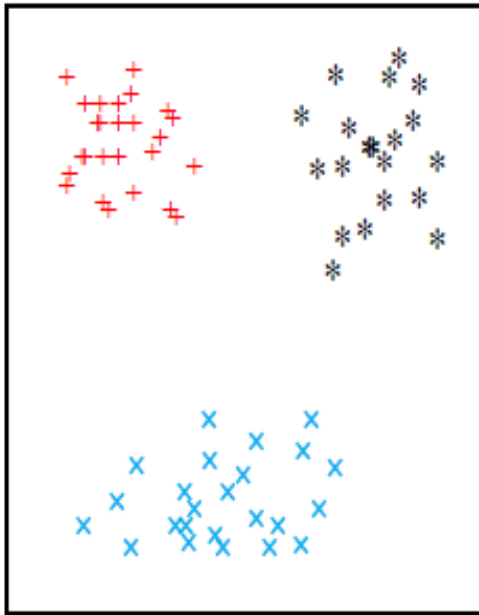
Trí tuệ nhân tạo là lĩnh vực nhằm nghiên cứu và phát triển các hệ thống (phần mềm, phần mềm nhúng) có khả năng giải quyết các bài toán giống như cách thức giải quyết của con người trong một ngữ cảnh nào đó. Theo [13], một hệ thống trí tuệ nhân tạo có thể có một hoặc một số khả năng như: suy nghĩ, lập luận, cảm nhận, quyết định, học, điều khiển, nhận dạng, ... Các hệ thống trí tuệ nhân tạo bắt đầu được nghiên cứu từ những năm 1956 của thế kỷ 20 với các nghiên cứu tại Stanford về mạng nơ ron, lý thuyết tự động của John MacCarthy [13].

Học máy (machine learning) là một lĩnh vực của trí tuệ nhân tạo nghiên cứu phát triển các phần mềm cho dùng cho *máy tính* hoặc *hệ thống máy tính* có thể giải quyết các tình huống cụ thể hoặc nhận dạng ra các mẫu giống như con người (xem hình 1.1). *Máy tính* hoặc *hệ thống máy tính* ở đây hiểu rằng là bất kỳ hệ thống nào mà có thể nạp và sử dụng phần mềm để thực hiện trên nó.

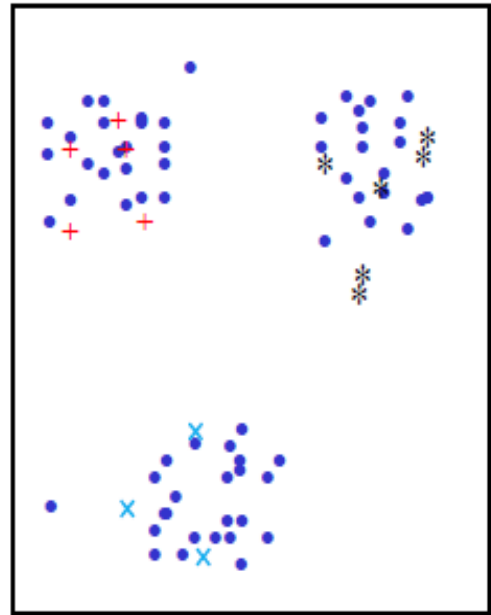
Trong lĩnh vực học máy hiện nay có ba phương pháp học cơ bản bao gồm: học có giám sát, học nửa giám sát và học không giám sát.

- Ý tưởng cơ bản của học có giám sát có thể hiểu như chúng ta cung cấp một số mẫu (ví dụ dữ liệu, hình ảnh, đồ vật đã gán nhãn) cho hệ thống học và sau đó thiết kế phát triển các hệ thống có thể suy diễn hay nhận biết mẫu mới nằm trong phạm vi nó đã được học.
- Học nửa giám sát khác với học có giám sát là các thuật toán dạng này chỉ sử dụng một lượng nhỏ các mẫu (các dữ liệu đã gán nhãn) để học và suy luận ra các dữ liệu chưa gán nhãn.

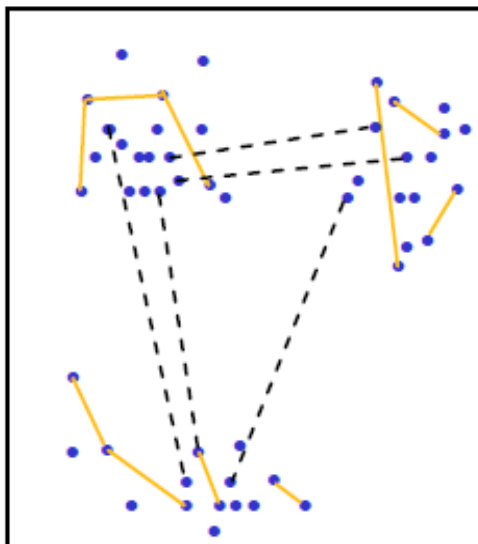
- Học không giám sát không dùng bất kỳ dữ liệu gán nhãn nào mà chỉ sử dụng các dữ liệu không có nhãn để thực hiện yêu cầu nào đó chẳng hạn như phân cụm các dữ liệu hay phát hiện các dị thường trong dữ liệu hay ngoại suy.



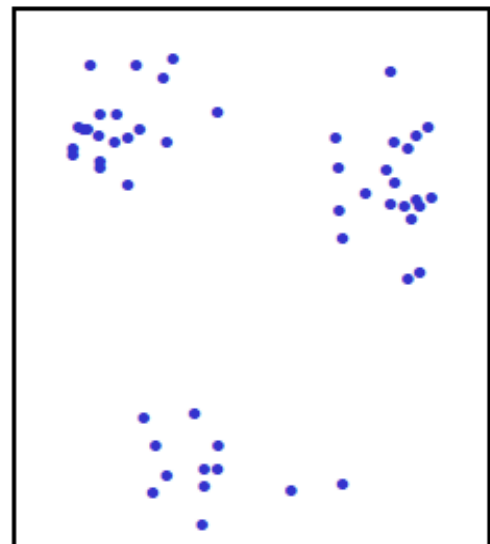
(a) – học có giám sát



(b) – học nửa giám sát



(c)- học nửa giám sát



(d)- học không giám sát

**Hình 1.1** Các mô hình của học máy

### 1.1.2 Ứng dụng của lĩnh vực học máy

- Xử lý ngôn ngữ tự nhiên (Natural Language Processing): xử lý văn bản, giao tiếp người - máy.
- Nhận dạng (Pattern recognition): nhận dạng tiếng nói, chữ viết tay, vân tay, thị giác máy, thiên văn học
- Máy tìm kiếm như Google, Yahoo, You tube: các hệ thống này sử dụng các công cụ của học máy để phát triển hệ thống
- Chẩn đoán trong y tế: trợ giúp phân tích ảnh X - quang, các hệ chuyên gia chẩn đoán tự động
- Tin sinh học: phân loại và dự đoán chuỗi gene, dự đoán tính chất của thuốc mới
- Phát hiện gian lận tài chính, gian lận thẻ tín dụng, phát hiện dị thường (hình 1.3)
- Phân tích thị trường chứng khoán (stock market analysis)
- Trò chơi: chơi cờ (Deep blue, IBM, 1998),
- Người máy (robot): là tổng hợp của rất nhiều ngành khoa học, trong đó học máy tạo nên hệ thần kinh/bộ não của người máy (xem hình 1.2).



*Hình 1.2 Người máy ASIMO của hãng HONDA, một trong những ứng dụng của lĩnh vực trí tuệ nhân tạo*



Hình 1.3 Phát hiện vùng bị hỏng (dị thường) trên vỏ cam

Trong nội dung của luận văn này, chúng tôi chọn bài toán phân lớp (một dạng của phương pháp học có giám sát) để nghiên cứu và tìm hiểu cũng như thử nghiệm các ứng dụng thực tế. Trong phần tiếp theo, chúng tôi sẽ trình bày các thuật ngữ, các định nghĩa và khái niệm cơ bản, tiếp đó chúng tôi nghiên cứu hai phương pháp phân lớp cơ bản là PCA và SVM trong chương 2, phần thực nghiệm và đánh giá kết quả là nội dung của chương 3 sẽ tổng kết các kết quả đã làm được và hướng phát triển tiếp theo.

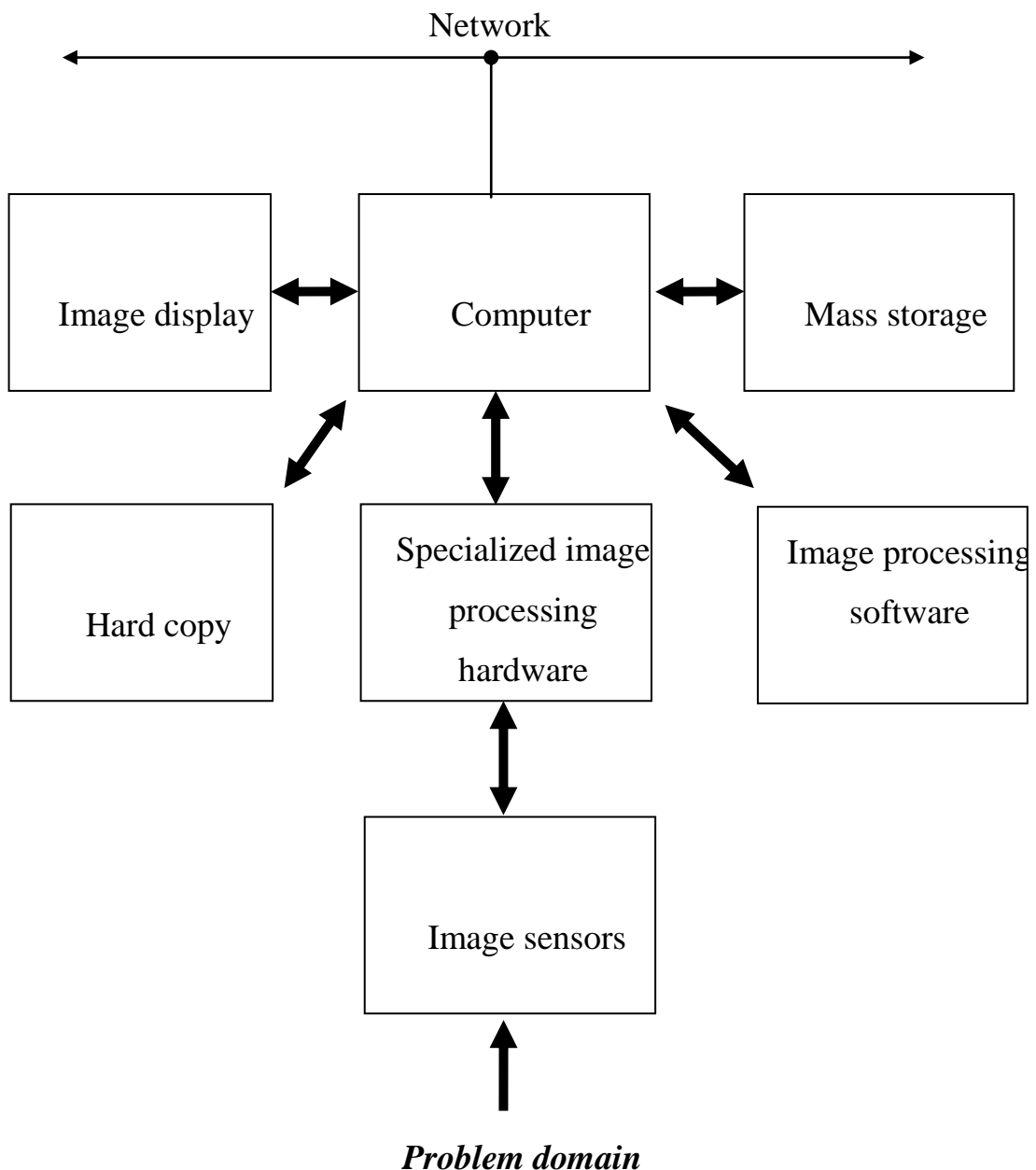
## 1.2 Tổng quan về xử lý ảnh số và ứng dụng

### 1.2.1 Giới thiệu

Lĩnh vực xử lý ảnh số là việc thực hiện xử lý ảnh dựa trên một thuật toán nào đó trên máy tính [1, 7]. Trong những năm gần đây lĩnh vực xử lý ảnh và ứng dụng phát triển mạnh mẽ trên thế giới, mỗi năm luôn có các hội thảo quy mô lớn về lĩnh vực này. Một hệ thống xử lý ảnh tổng quát có thể minh họa trên hình 1.4.

- *Problem domain*: Vấn đề cần giải quyết: chẳng hạn bài toán nhận dạng mặt người, bài toán phân cụm ảnh, bài toán phát hiện đối tượng trong ảnh,...

- *Specialized image processing hardware*: là bộ chuyển đổi ảnh số, đôi khi là những công cụ chuyên dụng cho phép xử lý nhanh và song song các phép toán.



**Hình 1.4.** Tổng quan về hệ thống xử lý ảnh trên máy tính [4]

- *Computer*: Là thành phần không thể thiếu trong hệ thống xử lý ảnh, computer ở đây có thể bao gồm nhiều loại từ thông thường đến các hệ thống siêu máy tính.

- *Software*: là các phần mềm chuyên cho từng mục đích cụ thể

- *Massstorage*: Vùng lưu trữ, với dữ liệu là các ảnh yêu cầu một không gian lưu trữ lớn.

- *Image display*: có thể là màn hình máy tính, LCD, TV,...

- *Hardcopy*: máy in lase, camera, máy in kim,...

- *Networking*: hệ thống mạng dùng cho mục đích truyền thông

Các ứng dụng của xử lý ảnh hiện nay rất rộng, bao gồm: y học, khoa học trái đất, vũ trụ, nhận dạng, phân nhóm, giám sát...

## 1.2.2 Những khái niệm cơ bản trong xử lý ảnh

### 1.2.2.1 Điểm ảnh (pixel)

Gốc của ảnh là ảnh liên tục về không gian và độ sáng. Để xử lý bằng máy tính (hệ thống máy tính), ảnh cần phải được số hoá. Số hoá ảnh được hiểu là sự biến đổi gần đúng một ảnh liên tục thành một tập điểm phù hợp với ảnh thật về vị trí (không gian) và độ sáng (mức xám). Khoảng cách giữa các điểm ảnh được thiết lập sao cho mắt người không phân biệt được ranh giới giữa chúng. Mỗi một điểm như vậy gọi là điểm ảnh hay *pixel*. Đối với ảnh hai chiều, mỗi *pixel* tương ứng với cặp tọa độ  $(x, y)$ .

### 1.2.2.2 Độ phân giải của ảnh

Độ phân giải (resolution) của ảnh là số lượng điểm ảnh trên một ảnh (vùng ảnh) cụ thể. Theo định nghĩa, khoảng cách giữa các điểm ảnh phải được chọn sao cho mắt người vẫn thấy được sự liên tục của ảnh. Việc lựa chọn khoảng

cách thích hợp tạo nên một mật độ phân bố, đó chính là độ phân giải và được phân bố theo trục  $x$  và  $y$  trong không gian hai chiều. Thực tế độ phân giải càng lớn thì ảnh càng *thật* ngược lại độ phân giải thấp ảnh sẽ mờ đi hoặc thậm chí không rõ.

### **1.2.2.3 Mức xám của ảnh**

Là kết quả của sự biến đổi tương ứng một giá trị độ sáng của một điểm ảnh với một giá trị nguyên dương. Thông thường nó xác định trong khoảng  $[0...255]$ . Tùy thuộc vào giá trị xám mà mỗi điểm ảnh được biểu diễn trên 1, 4, 8, 24 hay 32 bit trong máy tính.

### **1.2.2.4 Ảnh số**

Ảnh số bao gồm tập hợp hữu hạn các điểm ảnh, thường được biểu diễn bằng một mảng hai chiều  $I(m,n)$  với  $m$  là số hàng,  $n$  là số cột. Ta ký hiệu  $P(x,y)$  là một điểm ảnh tại vị trí  $(x, y)$ . Số lượng điểm ảnh trên mỗi hàng hoặc các hàng xác định độ phân giải của ảnh. Ảnh số được chia làm 3 loại:

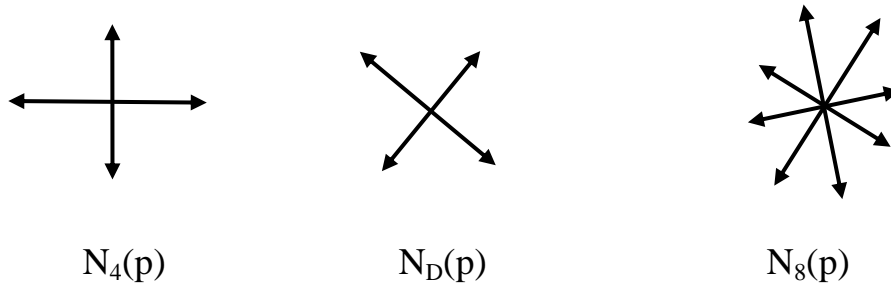
- Ảnh nhị phân
- Ảnh xám
- Ảnh màu

### **1.2.2.5 Các mối quan hệ cơ bản giữa các điểm ảnh**

Lân cận của một điểm ảnh: một điểm ảnh  $p$  tại tọa độ  $(x, y)$  có các lân cận theo chiều ngang và chiều dọc là:  $(x+1, y)$ ,  $(x-1, y)$ ,  $(x,y+1)$ ,  $(x, y-1)$ .

Tập hợp các điểm ảnh trên được gọi là lân cận 4 của  $p$ , ký hiệu  $N_4(p)$ . Mỗi điểm ảnh có khoảng cách đơn vị đến  $(x, y)$ , và nếu  $(x, y)$  nằm trên biên của ảnh thì lân cận của nó có thể nằm ngoài ảnh.

Các lân cận chéo của  $p$  có tọa độ:  $(x+1, y+1)$ ,  $(x+1, y-1)$ ,  $(x-1, y+1)$ ,  $(x-1, y-1)$ . Tập lân cận chéo được ký hiệu  $ND(p)$ . Tập lân cận chéo cùng với lân cận 4 tạo thành lân cận 8 của  $p$ , ký hiệu  $N_8(p)$  (xem hình 1.5).



Hình 1.5 Lân cận của 1 điểm ảnh

Các môi liên kết được sử dụng để xác định giới hạn (*Boundaries*) của đối tượng vật thể hoặc xác định vùng trong một ảnh. Một liên kết được đặc trưng bởi tính liên kề giữa các điểm và mức xám của chúng.

Giả sử  $V$  là tập các giá trị mức xám. Một ảnh có các giá trị cường độ sáng từ thang mức xám từ 32 đến 64 được mô tả như sau:

$$V = \{32, 33, \dots, 63, 64\}$$

Có 3 loại liên kết:

- Liên kết 4: Hai điểm ảnh  $p$  và  $q$  với các giá trị cường độ sáng  $V$  được nói là liên kết 4 nếu  $q$  nằm trong tập  $N_4(p)$ .
- Liên kết 8: Hai điểm ảnh  $p$  và  $q$  với các giá trị cường độ sáng  $V$  được gọi là liên kết 8 nếu  $q$  nằm trong tập  $N_8(p)$ .
- Liên kết  $m$  (liên kết hỗn hợp): Hai điểm ảnh  $p$  và  $q$  với các giá trị cường độ sáng  $V$  được nói là liên kết  $m$  nếu:  $q$  thuộc  $N_4(p)$  hoặc  $q$  thuộc  $ND(p)$ .

### 1.2.2.6 Biểu diễn ảnh



Ảnh sau khi số hoá sẽ được lưu vào bộ nhớ, hoặc chuyển sang các khâu tiếp theo để phân tích. Nếu lưu trữ ảnh trực tiếp từ các ảnh thô, đòi hỏi dung lượng bộ nhớ cực lớn và không hiệu quả theo quan điểm ứng dụng và công nghệ. Thông thường, các ảnh thô đó được đặc tả (biểu diễn) lại (hay đơn giản là mã hoá) theo các đặc điểm của ảnh được gọi là các đặc trưng ảnh (*Image Features*) như: biên ảnh (*Boundary*), vùng ảnh (*Region*). Một số phương pháp biểu diễn thường dùng:

- Biểu diễn bằng mã chạy (Run-Length Code)
- Biểu diễn bằng mã xích (Chain -Code)
- Biểu diễn bằng mã tứ phân (Quad-Tree Code)

### **Biểu diễn bằng mã chạy**

Phương pháp này thường biểu diễn cho vùng ảnh và áp dụng cho ảnh nhị phân. Một vùng ảnh  $R$  có thể mã hoá đơn giản nhờ một ma trận nhị phân:

$$U(m,n) = 1 \text{ nếu } (m,n) \text{ thuộc } R$$

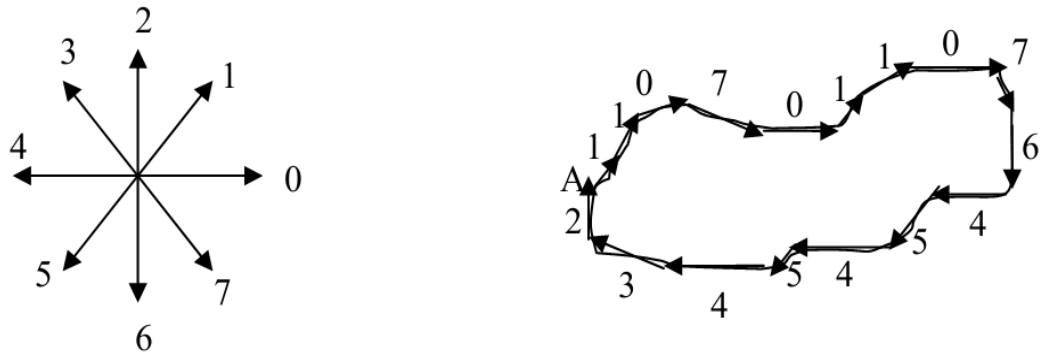
$$U(m,n) = 0 \text{ nếu } (m,n) \text{ không thuộc } R$$

Trong đó:  $U(m, n)$  là hàm mô tả mức xám ảnh tại tọa độ  $(m, n)$ . Với cách biểu diễn trên, một vùng ảnh được mô tả bằng một tập các chuỗi số 0 hoặc 1. Giả sử chúng ta mô tả ảnh nhị phân của một vùng ảnh được thể hiện theo tọa độ  $(x, y)$  theo các chiều và đặc tả chỉ đối với giá trị “1”. Khi đó dạng mô tả có thể là:  $(x, y)r$ ; trong đó  $(x, y)$  là tọa độ,  $r$  là số lượng các bit có giá trị “1” liên tục theo chiều ngang hoặc dọc.

### **Biểu diễn bằng mã xích (Chain -Code)**

Phương pháp này thường dùng để biểu diễn đường biên ảnh. Một đường bất kỳ được chia thành các đoạn nhỏ. Nối các điểm chia, ta có các đoạn thẳng

kế tiếp được gán hướng cho đoạn thẳng đó tạo thành một dây xích gồm các đoạn. Các hướng có thể chọn 4, 8, 12, 24,... mỗi hướng được mã hoá theo số thập phân hoặc số nhị phân thành mã của hướng (xem hình 1.6).



Hình 1.6 Hướng các điểm biên và mã tương ứng

### Biểu diễn bằng mã tứ phân (Quad-Tree Code)

Phương pháp mã tứ phân được dùng để mã hoá cho vùng ảnh. Vùng ảnh đầu tiên được chia làm bốn phần thường là bằng nhau. Nếu mỗi vùng đã đồng nhất (chứa toàn điểm đen hoặc trắng), thì gán cho vùng đó một mã và không chia tiếp. Các vùng không đồng nhất được chia tiếp làm bốn phần theo thủ tục trên cho đến khi tất cả các vùng đều đồng nhất. Các mã phân chia thành các vùng con tạo thành một cây phân chia các vùng đồng nhất.

#### 1.2.2.7 Các định dạng ảnh cơ bản

Ảnh thu được sau quá trình số hóa thường được lưu lại cho các quá trình xử lý tiếp theo hay truyền đi. Trong quá trình phát triển của kỹ thuật xử lý ảnh, tồn tại nhiều định dạng ảnh khác nhau từ ảnh đen trắng (với định dạng IMG), ảnh đa cấp xám cho đến ảnh màu: (BMP, GIF, JPE...). Tuy các định dạng này khác nhau, song chúng đều tuân thủ theo một cấu trúc chung nhất. Nhìn chung, một tệp ảnh bất kỳ thường bao gồm 3 phần:

- Mào đầu tệp (Header)

- Dữ liệu nén (Data Compression)
- Bảng màu (Palette Color)

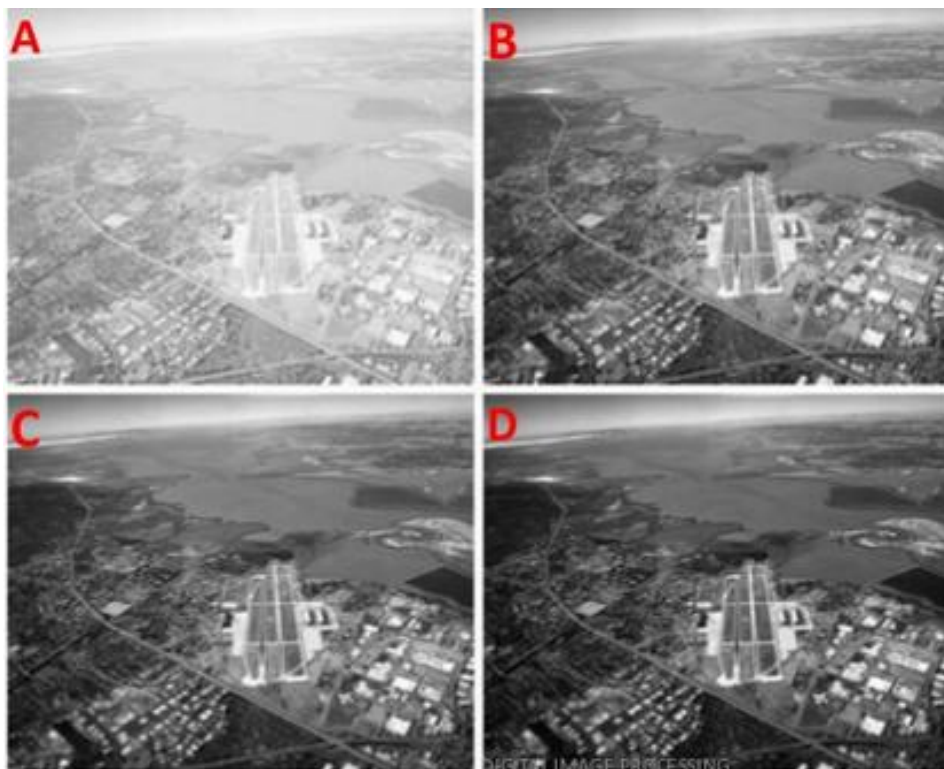
### 1.3 Một số bài toán trong xử lý ảnh

#### 1.3.1 Tăng cường ảnh

Tăng cường ảnh là bước quan trọng, tạo tiền đề cho toàn bộ quá trình xử lý ảnh. Tăng cường ảnh nhằm hoàn thiện các đặc tính của ảnh như:

- Lọc nhiễu, hay làm trơn ảnh.
- Tăng độ tương phản, điều chỉnh mức xám của ảnh.
- Làm nổi biên ảnh.

Các thuật toán triển khai việc nâng cao chất lượng ảnh hầu hết dựa trên các kỹ thuật trong miền điểm, không gian và tần số. Hình 1.7 minh họa ví dụ về tăng cường ảnh.



*Hình 1.7 Ví dụ về tăng cường ảnh: Hình B và D tương ứng là kết quả của sự tăng cường ảnh của hình A và C*

### 1.3.2 Phân vùng ảnh

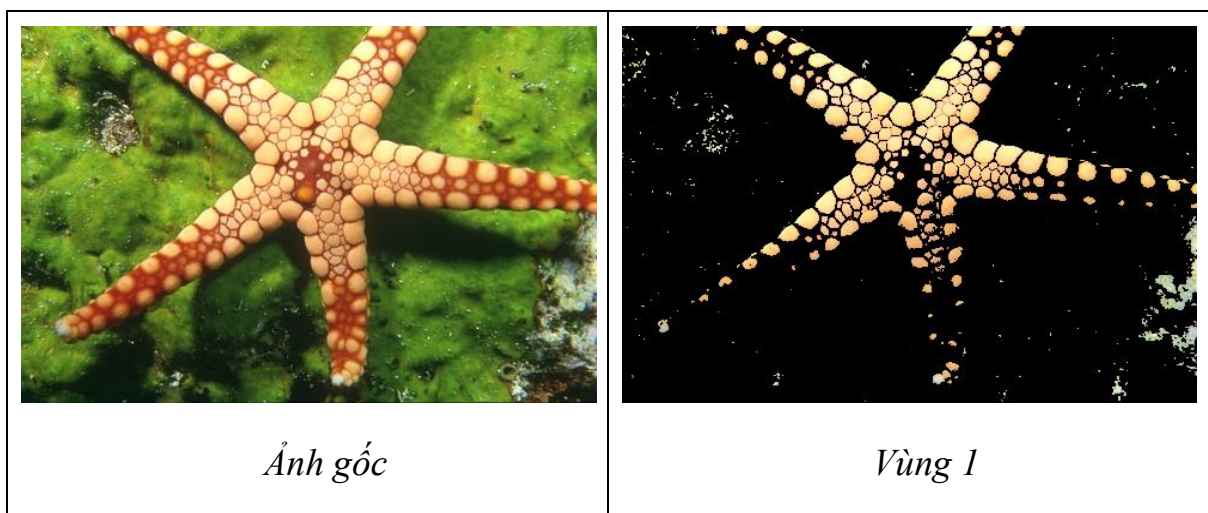
Để phân tích các đối tượng trong ảnh, chúng ta cần phải phân biệt được các đối tượng cần quan tâm với phần còn lại của ảnh. Những đối tượng này có thể tìm ra được nhờ các kỹ thuật phân vùng ảnh.

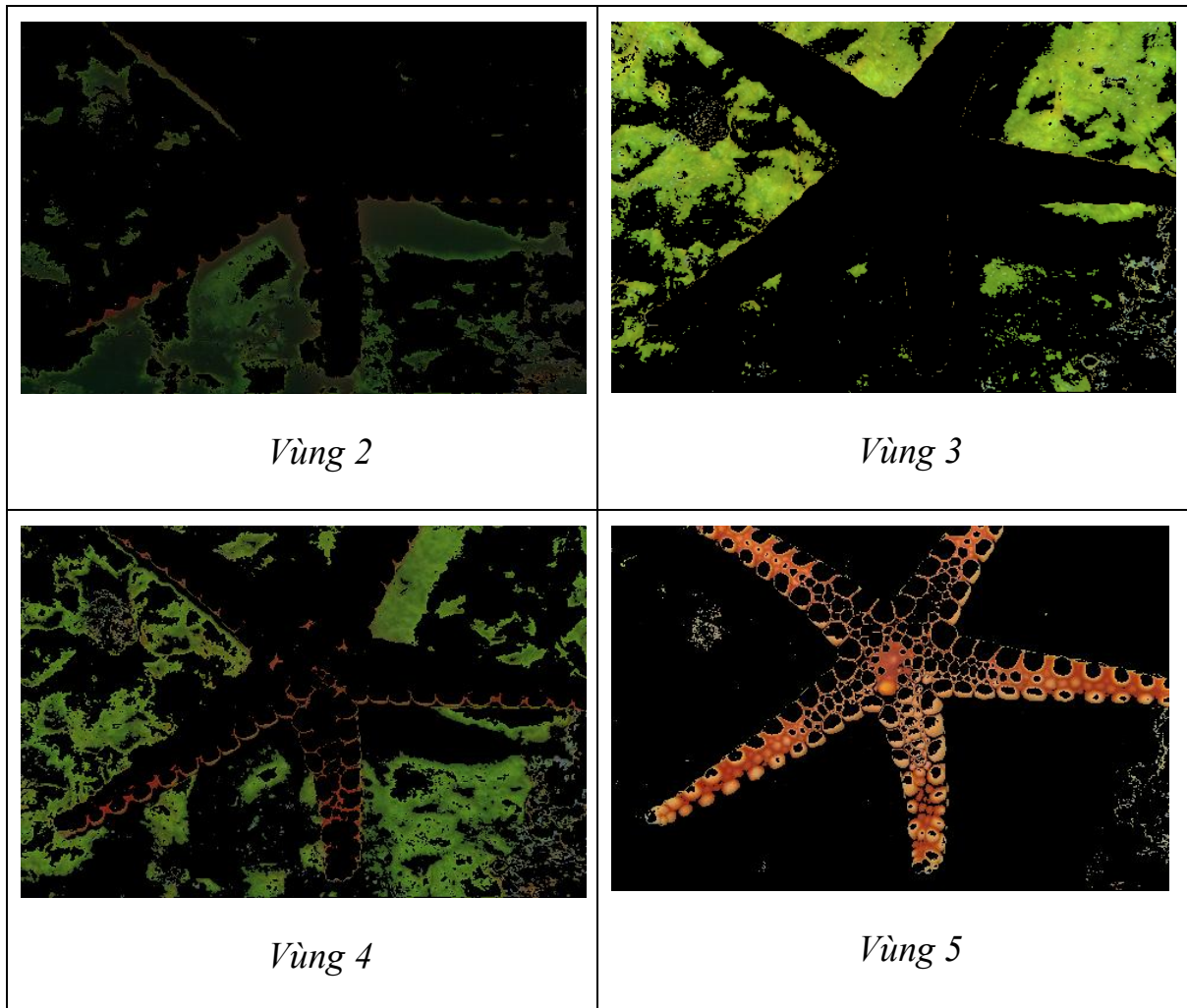
Vùng ảnh là một chi tiết, một thực thể trong toàn cảnh. Nói đến vùng ảnh là nói đến tính chất bề mặt của ảnh. Nó là một tập hợp các điểm có cùng hoặc gần cùng một tính chất nào đó: mức xám, màu sắc... Đường bao quanh một vùng ảnh (Boundary) là biên ảnh. Các điểm trong một vùng ảnh có độ biến thiên giá trị mức xám tương đối đồng đều hay tính kết cấu tương đồng.

Một phương pháp phân vùng ảnh là sử dụng một ngưỡng giá trị xám để phân tách ảnh thành đối tượng và nền (những điểm dưới ngưỡng xám thuộc về nền, ngược lại thuộc về đối tượng).

Ứng dụng của phân vùng ảnh trong các lĩnh vực như nhận dạng đối tượng, phát hiện dị thường,...

Hình 1.8 minh họa về bài toán phân vùng ảnh.





*Hình 1.8 Ví dụ về phân vùng ảnh (màu đen tương ứng là màu nền): Ảnh gốc và 5 vùng tương ứng của nó*

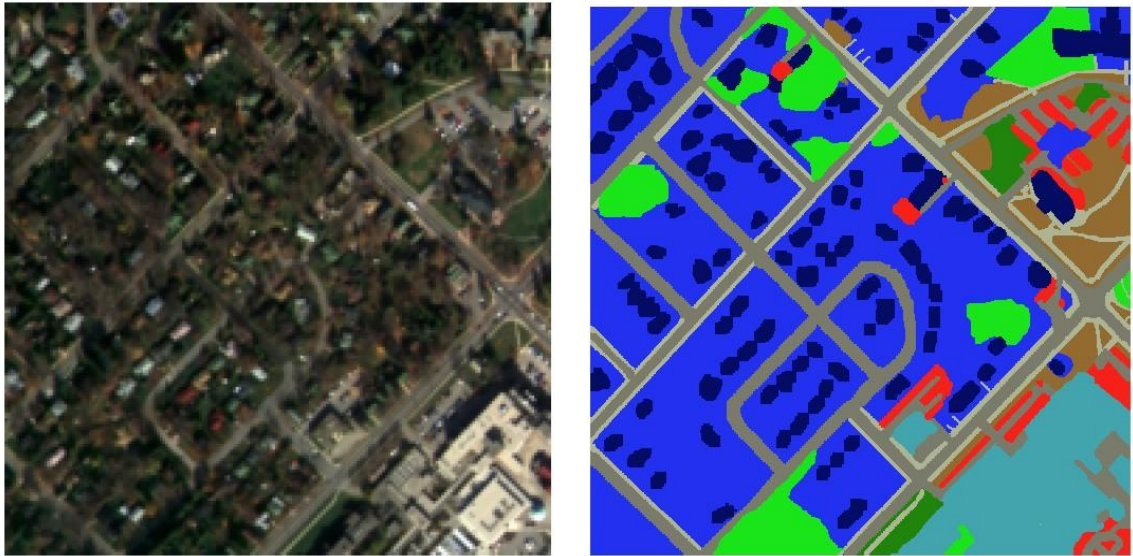
### 1.3.3 Trích chọn đặc tính

Mục đích của việc trích chọn các đặc tính là làm giảm kích thước của ảnh và đồng thời vẫn bảo toàn thông tin cần thiết về bức ảnh. Việc trích chọn hiệu quả các đặc điểm giúp cho việc nhận dạng các đối tượng ảnh chính xác, với tốc độ tính toán cao và dung lượng nhớ lưu trữ giảm xuống.

Dựa trên các thông tin thu nhận được qua quá trình phân vùng, kết hợp với các kỹ thuật xử lý để đưa ra các đặc trưng về đối tượng ảnh cũng như các thông

tin cần thiết trong quá trình xử lý. Trích chọn đặc trưng phụ thuộc vào bản chất của bài toán cần xử lý.

Hình 1.9 minh họa về việc trích chọn đặc trưng của ảnh.



*Hình 1.9 Ví dụ về xác định các đặc trưng của ảnh dựa trên phân vùng ảnh*

### 1.3.4 Nén ảnh

Nhằm giảm thiểu không gian lưu trữ của ảnh, nén ảnh thường được tiến hành theo cả hai khuynh hướng là nén có bảo toàn và không bảo toàn thông tin.

Nén không bảo toàn thì thường có khả năng nén cao hơn nhưng khả năng phục hồi thì kém hơn. Trên cơ sở hai khuynh hướng, có 4 cách tiếp cận cơ bản trong nén ảnh:

- Nén ảnh thống kê: Kỹ thuật nén này dựa vào việc thống kê tần suất xuất hiện của giá trị các điểm ảnh, trên cơ sở đó mà có chiến lược mã hóa thích hợp. Một ví dụ điển hình cho kỹ thuật mã hóa này là \*.TIF
- Nén ảnh không gian: Kỹ thuật này dựa vào vị trí không gian của các điểm ảnh để tiến hành mã hóa. Kỹ thuật lợi dụng sự giống nhau của các

điểm ảnh trong các vùng gần nhau. Ví dụ cho kỹ thuật này là mã nén \*.PCX

- Nén ảnh sử dụng phép biến đổi: Đây là kỹ thuật tiếp cận theo hướng nén không bảo toàn và do vậy, kỹ thuật thường nén hiệu quả hơn. \*.JPG chính là tiếp cận theo kỹ thuật nén này.
- Nén ảnh Fractal: Sử dụng tính chất Fractal của các đối tượng ảnh, thể hiện sự lặp lại của các chi tiết. Kỹ thuật nén sẽ tính toán để chỉ cần lưu trữ phần gốc ảnh và quy luật sinh ra ảnh theo nguyên lý Fractal.

### 1.3.5 Phát hiện biên ảnh

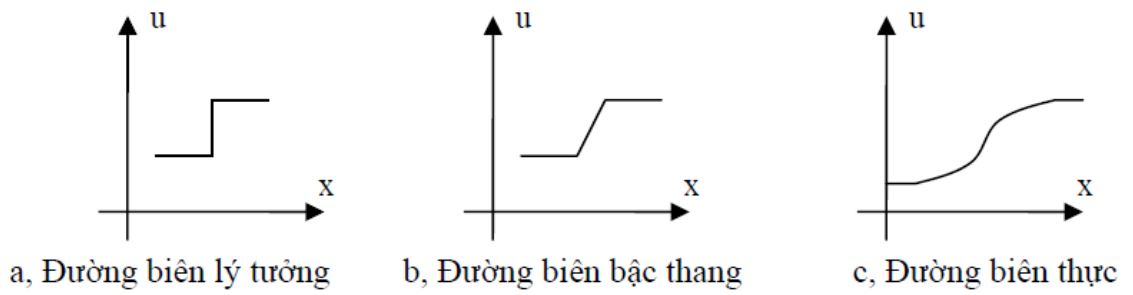
*Điểm Biên:* Một điểm ảnh được coi là điểm biên nếu có sự thay đổi nhanh hoặc đột ngột về mức xám (hoặc màu). Ví dụ trong ảnh nhị phân, điểm đen gọi là điểm biên nếu lân cận nó có ít nhất một điểm trắng.

*Đường biên (đường bao: boundary):* tập hợp các điểm biên liên tiếp tạo thành một đường biên hay đường bao.

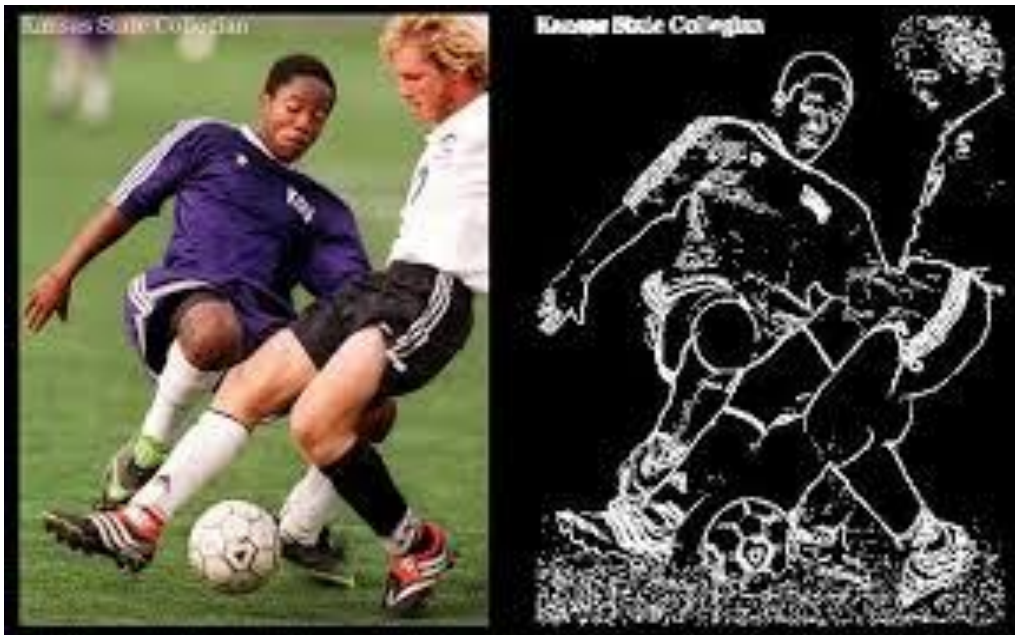
*Ý nghĩa của đường biên trong xử lý:* ý nghĩa đầu tiên: đường biên là một loại đặc trưng cục bộ tiêu biểu trong phân tích, nhận dạng ảnh. Thứ hai, người ta sử dụng biên làm phân cách các vùng xám (màu) cách biệt. Ngược lại, người ta cũng sử dụng các vùng ảnh để tìm đường phân cách.

*Tầm quan trọng của biên:* để thấy rõ tầm quan trọng của biên, xét ví dụ sau: khi người họa sỹ muốn vẽ một danh nhân, họa sỹ chỉ cần vẽ vài đường nét tốc họa mà không cần vẽ một cách đầy đủ.

*Mô hình biểu diễn đường biên,* theo toán học: điểm ảnh có sự biến đổi mức xám  $u(x)$  một cách đột ngột (hình 1.10, 1.11).



Hình 1.10 Ví dụ về đường biên ảnh



Hình 1.11 Ví dụ về xác định đường biên trong ảnh



### **Phân loại các kỹ thuật phát hiện biên**

Từ định nghĩa toán học của biên người ta sử dụng hai phương pháp phát hiện biên như sau:

*a, Phương pháp phát hiện biên trực tiếp:* phương pháp này chủ yếu dựa vào sự biến thiên độ sáng của điểm ảnh để làm nổi biên bằng kỹ thuật đạo hàm.

- Nếu lấy đạo hàm bậc nhất của ảnh: ta có phương pháp Gradient
- Nếu lấy đạo hàm bậc hai của ảnh: ta có phương pháp Laplace.

Hai phương pháp này được gọi chung là phương pháp dò biên cục bộ. Ngoài ra, người ta còn sử dụng phương pháp “đi theo đường bao” dựa vào công cụ toán học là nguyên lý quy hoạch động và được gọi là phương pháp dò biên tổng thể. Phương pháp dò biên trực tiếp có hiệu quả và ít bị tác động của nhiễu.

*b, Phương pháp phát hiện biên gián tiếp:* Nếu bằng cách nào đó, chúng ta thu được các vùng ảnh khác nhau thì đường phân cách giữa các vùng đó chính là biên. Nói cách khác, việc xác định đường bao của ảnh được thực hiện từ ảnh đã được phân vùng. Phương pháp dò biên gián tiếp khó cài đặt nhưng áp dụng tốt khi sự biến thiên độ sáng nhỏ.

### **1.3.6 Khôi phục ảnh**

Khôi phục ảnh là phục hồi lại ảnh gốc so với ảnh ghi được đã bị biến dạng. Nói cách khác, khôi phục ảnh là các kỹ thuật cải thiện chất lượng những ảnh ghi đảm bảo gần được như ảnh thật khi ảnh bị méo.

Để khôi phục được ảnh có kết quả, điều cần thiết là phải biết được các nguyên nhân, các hàm (hay dạng) gây ra biến dạng ảnh. Các nguyên nhân biến dạng thường do:

- Do camera, đầu thu ảnh chất lượng kém.
- Do môi trường, ánh sáng, hiện trường (scene), khí quyển, nhiễu xạ.

- Do chất lượng.

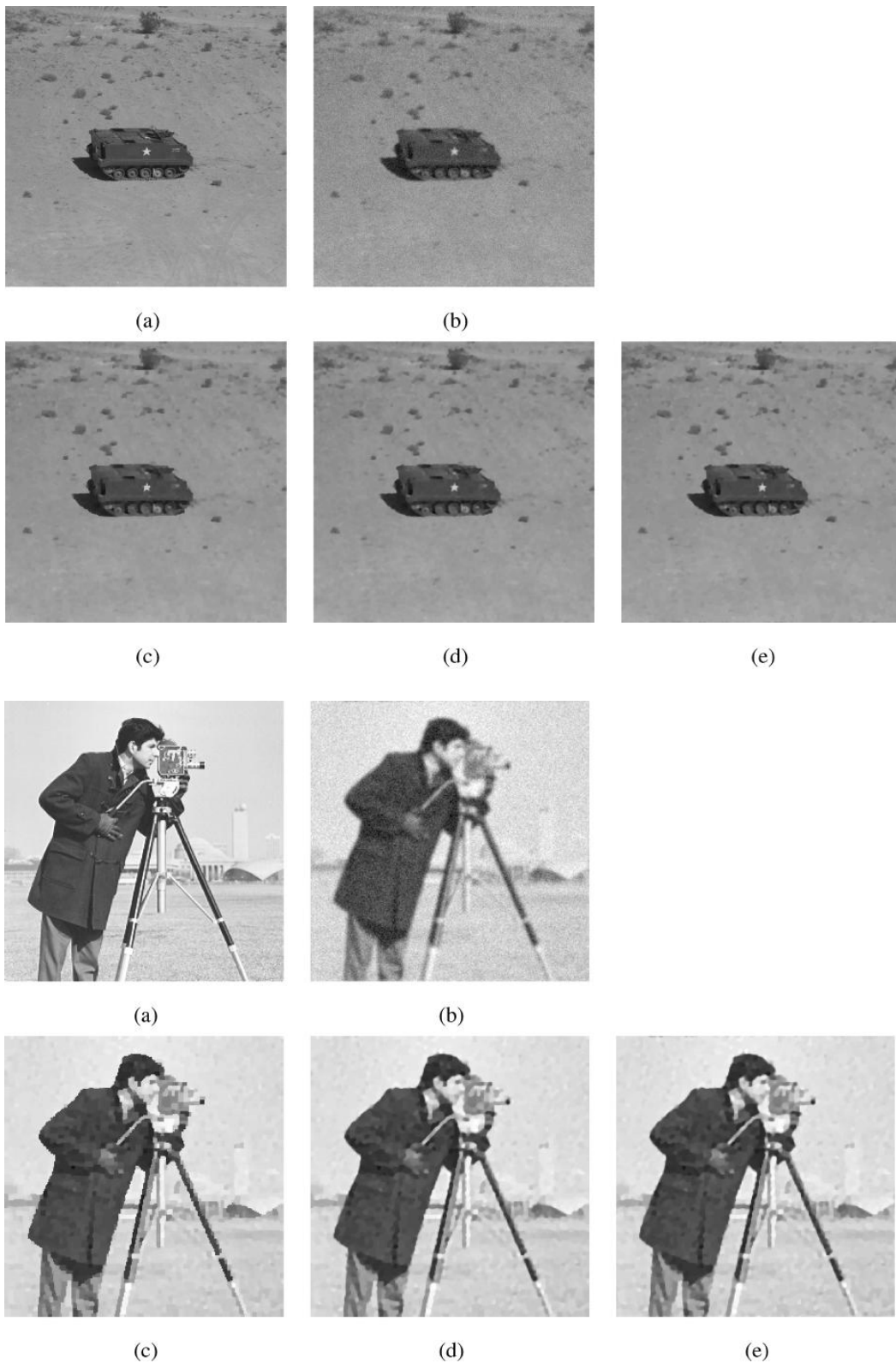
Kỹ thuật khôi phục ảnh có thể được xác định như việc ước lượng lại ảnh gốc hay ảnh lý tưởng từ ảnh quan sát được bằng cách đo ngược lại những hiện tượng gây biến dạng, qua đó ảnh được chụp. Như vậy, kỹ thuật khôi phục ảnh đòi hỏi kiến thức về các hiện tượng gây biến dạng ảnh.

Về mặt phương pháp, các nguồn tạo biến dạng có thể nhóm lại theo các xử lý như sau:

- Biến dạng điểm.
- Biến dạng không gian
- Biến dạng thường (Temporal)
- Biến dạng màu sắc (Chromatic)

Do quá trình tạo ảnh ghi được liên quan đến điểm ảnh xung quanh, chúng ta tập trung xét các biến dạng trong không gian. Một số ví dụ điển hình của biến dạng không gian được xem xét như sau.

- Nhiễu loạn của khí quyển (thiên văn) giữa các ống kính thu và đối tượng trong quá trình chụp ảnh. Do sai số hệ thống (hệ phi tuyến).
- Sai lệch hệ thống có thể biểu diễn bằng sai lệch hàm truyền (ví dụ: sự dịch pha hàm truyền coherent trong quang học...)



Hình 1.12 Ví dụ về khôi phục ảnh

**Các kỹ thuật khôi phục ảnh:**

- Mô hình khôi phục ảnh có: mô hình tạo ảnh, mô hình gây nhiễu, mô hình quan sát.
- Lọc tuyến tính có: lọc ngược, đáp ứng xung, lọc hữu hạn FIR.
- Các kỹ thuật khác: Entropy cực đại, mô hình Bayes, giải chập.

Hình 1.12 minh họa về việc khôi phục ảnh.

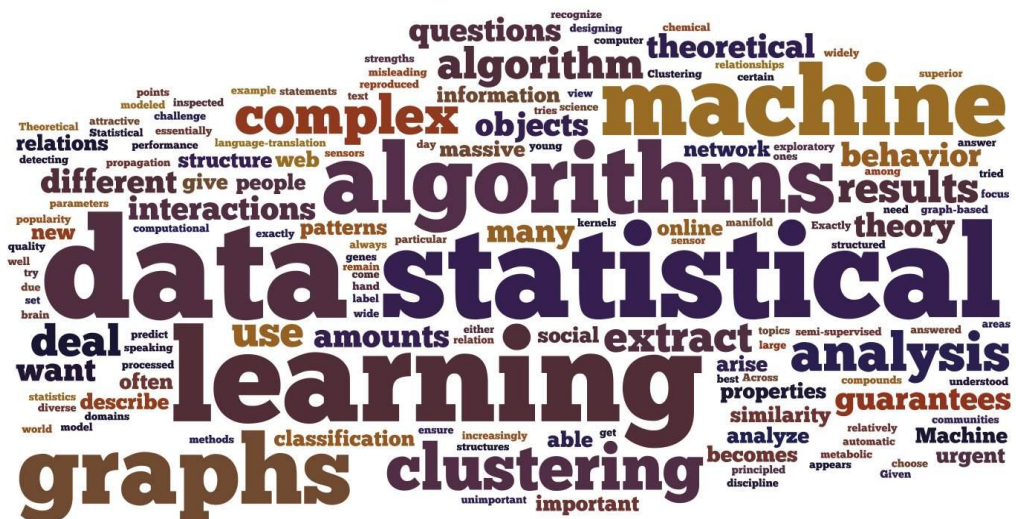
Trong các chương tiếp theo chúng tôi sẽ trình bày các phương pháp nhận dạng mặt người trong ảnh như SVM, PCA và từ đó xây dựng hệ thống thử nghiệm cho bài toán nhận dạng. Các kiến thức nền tảng về xử lý ảnh, đại số tuyến tính, ma trận,... cũng sẽ được sử dụng vào các phần tiếp theo.

## CHƯƠNG 2. NGHIÊN CỨU PHƯƠNG PHÁP PCA VÀ PHƯƠNG PHÁP SVM TRONG BÀI TOÁN NHẬN DẠNG MẶT NGƯỜI

### 2.1 Tổng quan

#### 2.1.1 Giới thiệu

Học máy (machine learning) là một lĩnh vực của trí tuệ nhân tạo nghiên cứu phát triển các thuật toán dùng cho *máy tính* hoặc *hệ thống máy tính* mà nó có thể giải quyết các tình huống cụ thể hoặc nhận dạng ra các mẫu giống như con người. Học máy kế thừa thành tựu của nhiều lĩnh vực khoa học (xem hình 2.1).



Hình 2.1 Các lĩnh vực liên quan đến học máy

*Lý thuyết về trí tuệ nhân tạo:* Là một lĩnh vực của trí tuệ nhân tạo nên học máy dùng các khái niệm như: bài toán tìm kiếm, biểu diễn và xử lý tri thức, sử dụng tri thức và suy diễn,...

*Lý thuyết độ phức tạp tính toán:* tính độ phức tạp của các nhiệm vụ học đo qua các ví dụ đào tạo, số lỗi và các tính toán...

*Lý thuyết điều khiển:* Các thủ tục học để điều khiển quá trình nhằm tối ưu hoá mục đích định trước hay học cách đoán các trạng thái tiếp theo của quá trình điều khiển.

*Lý thuyết thông tin:* các độ đo của nội dung thông tin và entropy, mã tối ưu và quan hệ của chúng tới dãy đào tạo tối ưu để mã hoá một giả thuyết.

*Triết học:* những nguyên lý như Occam's razor (cho rằng giả thuyết đơn giản nhất là tốt nhất, các phân tích luận chứng để tổng quát hoá các dữ liệu quan sát được.

*Tâm lý học và thần kinh học:* các đáp ứng thực tế của con người, các mô hình neural.

*Thống kê:* đặc trưng lỗi, lý thuyết lấy mẫu, khoảng tin cậy

### **2.1.2 Bài toán nhận dạng đối tượng**

Một mẫu (pattern) là một sự sắp xếp của các yếu tố mô tả, *feature* được sử dụng trong lĩnh vực nhận dạng đối tượng để chỉ đặc trưng của một mô tả. Một lớp các mẫu biểu thị các mẫu có cùng các tính chất nào đó. Nhận dạng mẫu là kỹ thuật để gán các mẫu vào lớp tương ứng của nó – gán tự động với một số ít nhất sự trợ giúp của con người. Một mẫu trong thực tế sử dụng ba dạng biểu diễn cơ bản là vector, xâu kí tự và cây (tree). Trên thực tế có hai chiến lược nhận dạng cơ bản bao gồm nhận dạng theo lý thuyết quyết định (mô hình toán học), nhận dạng theo cấu trúc của đối tượng.

*Nhận dạng dựa trên lý thuyết quyết định*

- Phương pháp đối sánh

Trong phương pháp này, mỗi lớp biểu diễn bằng một vector mẫu. Một mẫu chưa biết sẽ được gán vào lớp gần nó nhất. Chẳng hạn cách đơn giản nhất là tính khoảng cách từ mẫu đó đến các lớp rồi chọn lớp có khoảng cách ngắn nhất để

quyết định.

- Phân lớp dựa trên lý thuyết tối ưu và thống kê

Với phương pháp này, lý thuyết xác suất sẽ được sử dụng và một trong những phương pháp phổ biến là lý thuyết Bayes cho các mẫu có phân bố Gaussian.

- Phân lớp dựa trên mạng Nơ ron

Phương pháp sử dụng mạng Nơ ron được nghiên cứu từ những năm 80 của thế kỷ 20 và đã có nhiều ưu điểm cũng như các ứng dụng sử dụng. Ý tưởng của phương pháp này là sử dụng các nơ ron để xây dựng cấu trúc mạng và giải quyết bài toán tối ưu để xác định hàm phân lớp. Hiện nay mạng Nơ ron vẫn là một công cụ hiệu quả cho các bài toán nhận dạng trong các lĩnh vực khác nhau.

*Các phương pháp dựa trên cấu trúc*

- Phương pháp so sánh hình dạng

Ý tưởng của phương pháp này là đối sánh các đường biên theo, độ tương tự giữa hai vùng được định nghĩa như là số các cạnh trùng nhau lớn nhất.

- Phương pháp so sánh xâu kí tự

Giả sử rằng cần đối sánh hai vùng biên a và b, mỗi vùng biểu diễn bởi  $a_1a_2\dots a_n$  và  $b_1b_2\dots b_n$ . Đặt  $\alpha$  là số lượng các kí tự trùng nhau đôi một của các xâu a và b. Số lượng các điểm không thỏa mãn là:

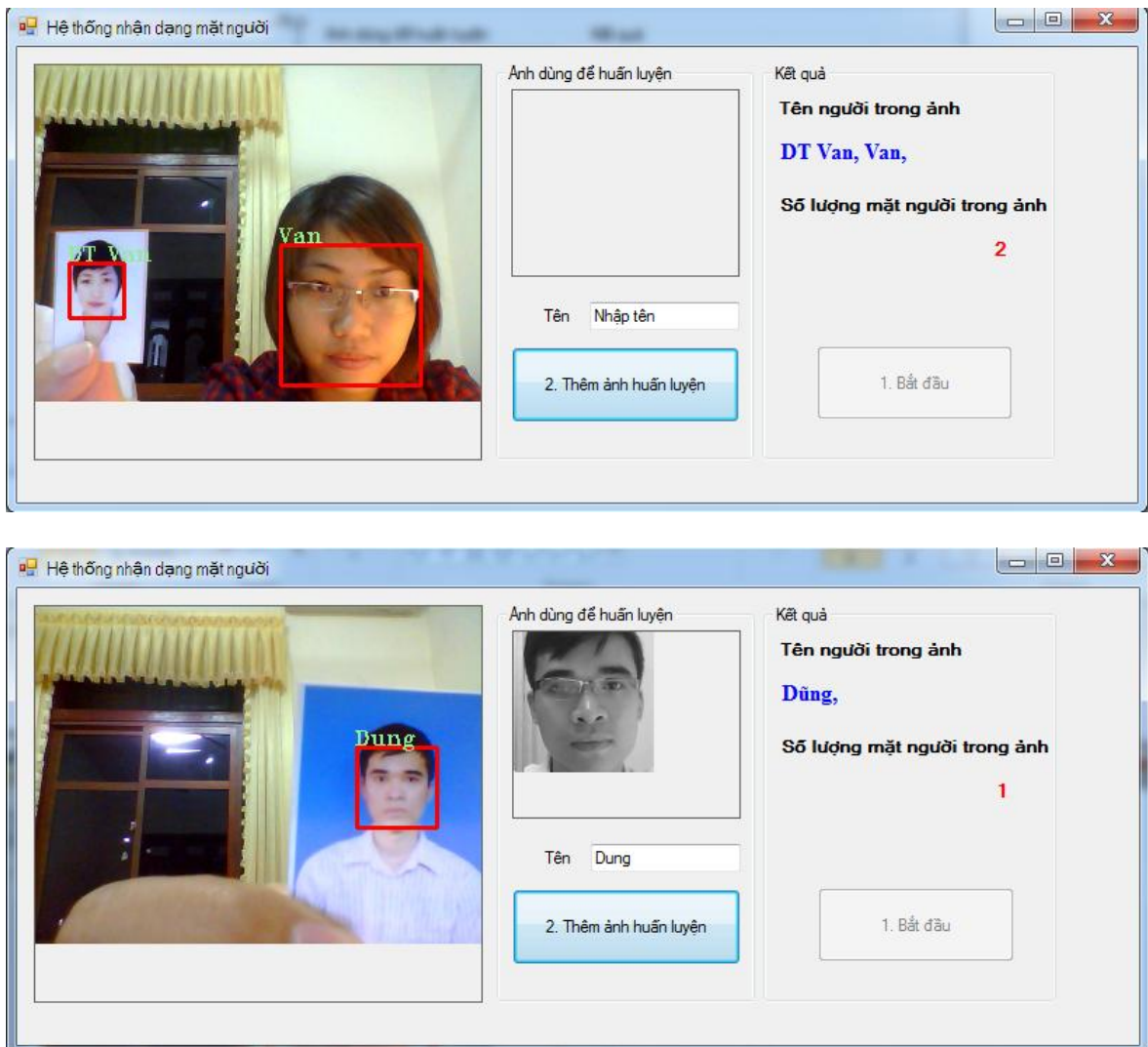
$$\beta = \max(|a|, |b|) - \alpha$$

Độ đo tương tự giữa a và b là:

$$R = \frac{\alpha}{\beta} = \frac{\alpha}{\max(|a|, |b|) - \alpha}$$

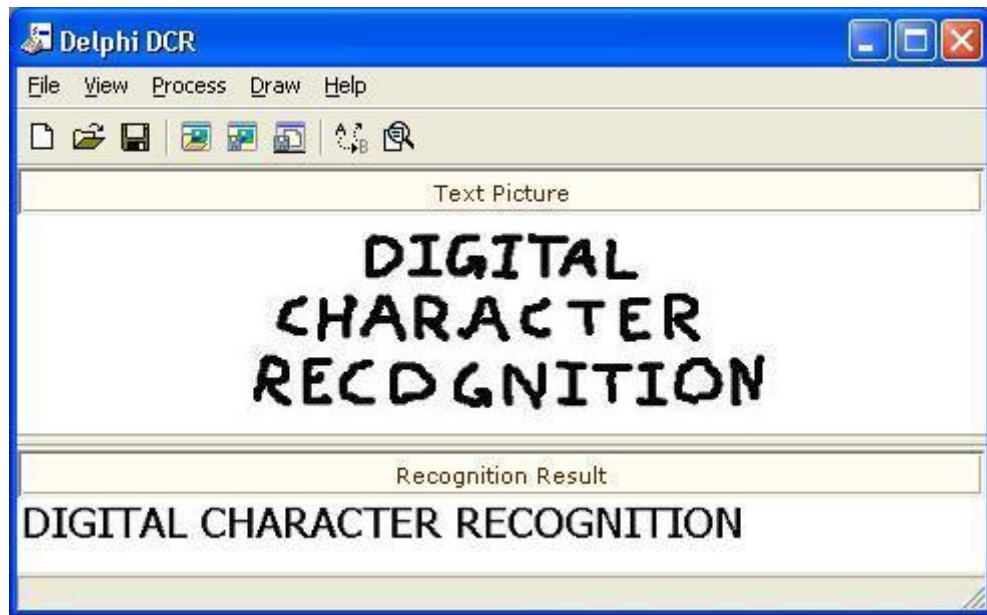
**Ứng dụng của bài toán nhận dạng**

Một số bài toán nhận dạng khá phổ biến hiện nay đang được áp dụng trong khoa học và công nghệ là: nhận dạng ký tự (chữ in, chữ viết tay, chữ ký điện tử), nhận dạng văn bản (Text), nhận dạng vân tay, nhận dạng mã vạch, nhận dạng mặt người (xem hình 2.2 - 2.4). Chúng ta có thể tìm thấy nhiều ứng dụng của các bài toán trên trong các bài toán thực tế. Các kỹ thuật để nhận dạng ảnh được nghiên cứu nhiều trong trí tuệ nhân tạo.

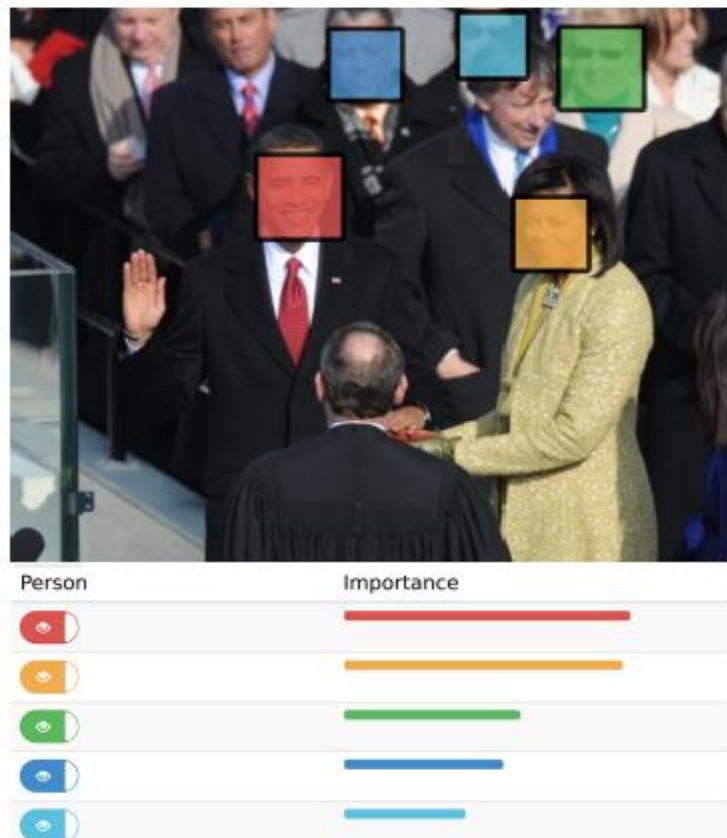


Hình 2.2 Ví dụ về hệ thống nhận dạng mặt người





Hình 2.3 Ví dụ về hệ thống nhận dạng chữ viết tay



Hình 2.4 Ví dụ về tìm kiếm người quan trọng trong ảnh

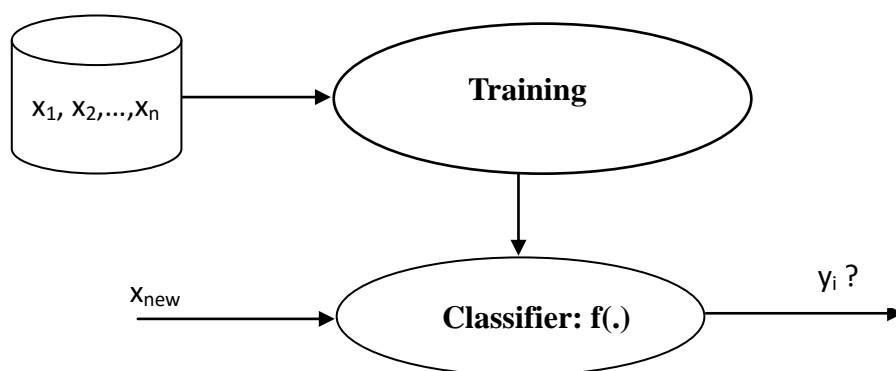
Trong phần tiếp theo của chương chúng tôi sẽ trình bày hai phương pháp nhận dạng cơ bản là phương pháp PCA (Principle Component Analysis) và phương pháp SVM (Support Vector Machine).

## 2.2 Bài toán học có giám sát

Học có giám sát (supervised learning) là một trong những chủ đề quan trọng trong lĩnh vực trí tuệ nhân tạo. Bài toán học có giám sát được phát biểu như sau:

*Cho tập  $X = \{x_i\}$ ,  $i = 1, 2, \dots, N$  với  $x_i$  diễn tả một đối tượng trong không gian  $d$  chiều. Với mỗi  $x_i$  tương ứng có nhãn là  $y_i$  với  $y_i \in \{C_1, C_2, \dots, C_M\}$  hoặc  $y_i \in R$ . Mục tiêu của bài toán học có giám sát là đi xây dựng hàm  $f(\cdot)$  dựa trên các bộ  $(x_i, y_i)$  sao cho hàm  $f$  có khả năng xác định nhãn cho một đối tượng  $x_{new}$  nào đó.*

Bài toán học có giám sát được minh họa bởi hình 2.5. Học có giám sát tức là học từ các mẫu thu thập được sau đó nhận dạng hay phân lớp cho mẫu mới, nó khác với bài toán học không giám sát – không có mẫu học ban đầu mà chỉ có dữ liệu cần giải quyết.



Hình 2.5 Mô hình học có giám sát

Chúng ta cũng lưu ý rằng khái niệm *học máy* ở đây được hiểu là phát triển các thuật toán cài đặt trên máy tính để dạy cho máy tính học và giải

quyết các bài toán giống như cách giải quyết của con người trong một khía cạnh nào đó và trong một lĩnh vực cụ thể nào đó.

Như vậy từ các dữ liệu ban đầu, chúng ta phải phát triển thuật toán (pha training: pha học) nhằm tạo ra một bộ phân lớp (classifier) để có thể xác định nhãn cho các đối tượng đầu vào mới ( $x_{\text{new}}$ ).

Một số ví dụ về bài toán học có giám sát:

- Bài toán nhận dạng biển số xe: Hệ thống sẽ được học các ký tự chữ và số, sau đó sẽ tự động phát hiện các biển số xe và chuyển về dạng văn bản.
- Bài toán nhận dạng mặt người: Hệ thống sẽ được huấn luyện với các mẫu ảnh mặt người kèm với nhãn của nó, sao đó có thể định danh cho một khuôn mặt mới.

Các thuật toán học có giám sát tiêu biểu trong thực tế hiện nay có thể kể đến như: phương pháp Support vector machine (SVM), phương pháp mạng Nơ ron, phương pháp cây quyết định, phương pháp Principle Component Analysis (PCA),... Mỗi phương pháp có những ưu và nhược điểm riêng cũng như phù hợp cho các ứng dụng theo lĩnh vực. Trong phần tiếp theo của chương này, tôi lựa chọn hai phương pháp là SVM và PCA cho việc nghiên cứu và tìm hiểu cũng như phân tích ưu nhược điểm của nó.

## **2.3 Phương pháp PCA và ứng dụng cho bài toán nhận dạng mặt người**

### **2.3.1 Giới thiệu về phương pháp PCA**

Ý tưởng cơ bản của phương pháp PCA là giảm số chiều dữ liệu đầu vào bằng cách chuyển không gian đầu vào sang không gian đích với số chiều ít hơn mà vẫn đảm bảo thông tin về các đối tượng được bảo toàn cơ bản. Phương pháp này được giới thiệu bởi năm 1901 và được biết đến như phương pháp phân tích thành phần cơ bản.

Trong không gian  $n$  chiều, với mỗi điểm  $x$  ta có biểu diễn sau:

$$x = a_1v_1 + a_2v_2 + \dots + a_nv_n$$

chúng ta sẽ tìm một phép biến đổi  $x \rightarrow \hat{x}$  ( $\hat{x}$  trong không gian  $K$  chiều  $K \ll n$ ):

$$\hat{x} = b_1u_1 + b_2u_2 + \dots + b_ku_k$$

sao cho  $\|x - \hat{x}\|$  là nhỏ nhất.

Cho  $M$  vector cột trong không gian  $N$  chiều  $x_1, x_2, \dots, x_M$ , phương pháp PCA bao gồm 6 bước như sau:

- **Bước 1:**  $\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i$

- **Bước 2:** Tính  $\Phi_i = x_i - \bar{x}$

- **Bước 3:** Đặt ma trận  $A = [\Phi_1 \ \Phi_2 \ \dots \ \Phi_M]$  (ma trận  $A$  có  $N$  hàng và  $M$  cột), tiếp đó tính:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = \frac{1}{M} AA^T$$

- **Bước 4:** Tính các giá trị riêng của ma trận  $C$ :  $\lambda_1 > \lambda_2 > \dots > \lambda_N$ .

- **Bước 5:** Tính các vec tơ riêng của  $C$ :  $u_1, u_2, \dots, u_N$ , ta có:

$$x - \bar{x} = b_1u_1 + b_2u_2 + \dots + b_Nu_N = \sum_{i=1}^N b_iu_i$$

- **Bước 6:** Giữ lại  $K$  giá trị riêng lớn nhất:

$$\hat{x} - \bar{x} = \sum_{i=1}^K b_iu_i, \quad K \ll N.$$

Để chọn  $K$  ta sử dụng công thức sau:

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^N \lambda_i} > \varepsilon$$

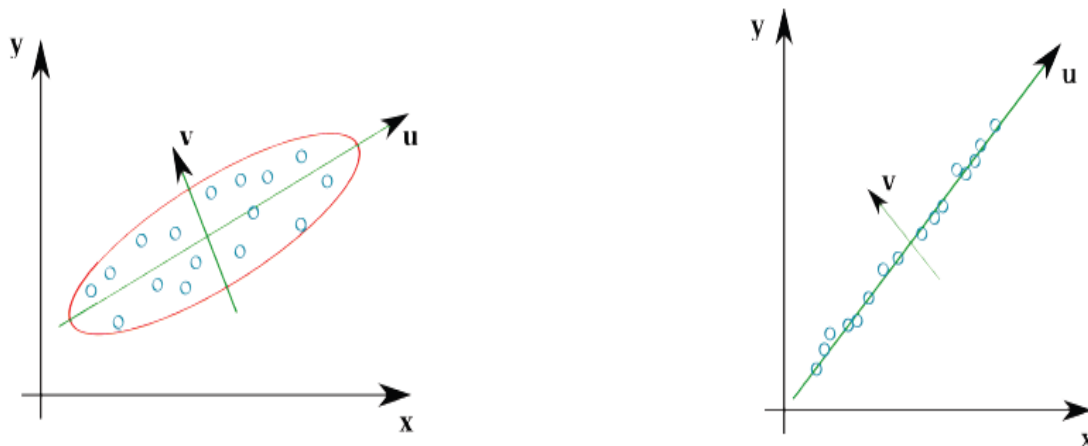
trong đó  $\varepsilon$  là ngưỡng do người dùng chọn (ví dụ 0.9, 0.95,...).

Bằng phương pháp trên, chúng ta đã chuyển dữ trong không gian  $N$  chiều về không gian  $K$  chiều trong đó  $K$  bé hơn rất nhiều. Chúng ta sẽ ứng dụng phương pháp này vào trường hợp cụ thể của bài toán nhận dạng mặt người trong phần tiếp sau đây.

### 2.3.2 Ví dụ về phương pháp PCA

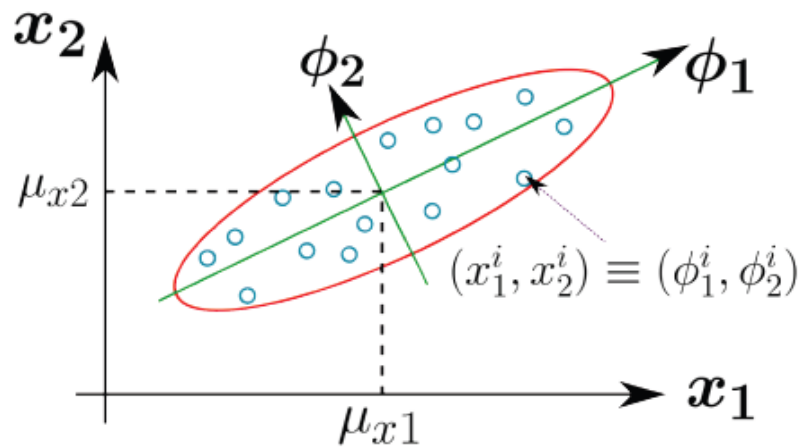
Hình 2.6 (bên trái) minh họa ví dụ về việc sử dụng PCA trong biểu diễn dữ liệu. Dữ liệu trong không gian  $X$ - $Y$  được chuyển về không gian  $U$ - $V$  cho chúng ta một cách biểu diễn dữ liệu *chắc chắn*. Hướng  $U$  và  $V$  được gọi là các thành phần cơ bản.

Hình 2.6 (bên phải), nếu dữ liệu chuyển về không gian  $U$ - $V$ , giá trị các điểm theo trục  $V$  gần như bằng không khi đó chúng ta hoàn toàn có thể bỏ qua chiều  $V$  mà chỉ giữ lại giá trị nằm trong thành phần  $U$ , ví dụ này cho chúng ta một ví dụ về giảm số chiều của dữ liệu.



*Hình 2.6 Ví dụ về PCA cho biểu diễn dữ liệu và giảm số chiều*

Hình 2.7 cho chúng ta một ví dụ về phương pháp PCA trong biến đổi dữ liệu từ không gian  $(x_1, x_2)$  thành không gian  $(\Phi_1, \Phi_2)$ .



Hình 2.7 Ví dụ về biến đổi dữ liệu trong PCA

### 2.3.3 PCA cho bài toán nhận dạng mặt người

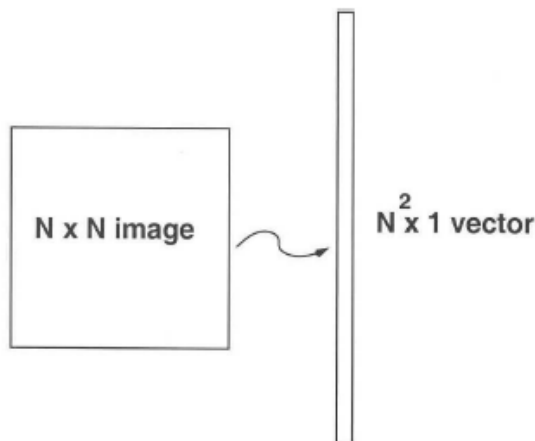
#### 2.3.2.1 Pha huấn luyện

Giả sử có  $M$  ảnh dùng chop ha huấn luyện  $I_1, I_2, \dots, I_M$ . Các ảnh ở đây đều là ảnh mặt người và chúng có mặt ở tâm ảnh và các ảnh có cùng kích thước với nhau (xem hình 2.8).



Hình 2.8. Ví dụ về tập dữ liệu mẫu

Với mỗi ảnh  $I_i$  (kích thước  $N \times N$ ) sẽ được biểu diễn bởi một vector  $\Gamma_i$  có  $N^2$  hàng và 1 cột như hình 2.9.



Hình 2.9 Biểu diễn ảnh

Chúng ta thấy rằng thông thường số chiều của mỗi ảnh là rất lớn (ví dụ  $128^2$ ,  $256^2$ , ...) trong phân tiếp theo, chúng ta sẽ áp dụng phương pháp PCA cho bài toán nhận dạng mặt người. Ý tưởng sử dụng phương pháp PCA trong bài toán nhận dạng mặt người được giới thiệu bởi Sirovich và Kirby năm 1987 [14].

Sử dụng phương pháp PCA với  $M$  ảnh mẫu làm dữ liệu đầu vào ta thực hiện các bước như sau:

- **Bước 1:** Đọc  $M$  ảnh vào hệ thống
- **Bước 2:** Với mỗi ảnh  $I_i$  sẽ được biểu diễn bởi một vector cột  $\Gamma_i$  có  $N^2$  giá trị
- **Bước 3:** Tính giá trị trung bình của các vector:  $\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i$
- **Bước 4:** Tính  $\Phi_i = \Gamma_i - \Psi$
- **Bước 5:** Đặt ma trận  $A = [\Phi_1 \ \Phi_1 \ \dots \ \Phi_M]$  (ma trận  $A$  có  $N^2$  hàng và  $M$  cột), tiếp đó tính:

$$C = \frac{1}{M} \sum_{i=1}^M \Phi_i \Phi_i^T = \frac{1}{M} AA^T \quad (\text{ma trận kích thước } N^2 \times N^2)$$

- **Bước 6:** Tính các vector riêng của ma trận C. Chú ý rằng kích thước của ma trận C rất lớn nên không thể tính trực tiếp các vector riêng ở đây, ta thực hiện các bước sau:

*Bước 6.1:* Giả sử ta có vector riêng  $v_i$  của ma trận  $A^T A$  thỏa mãn:

$$A^T A v_i = \mu_i v_i,$$

nhân hai vế với A ta được:

$$AA^T A v_i = \mu_i A v_i$$

Ta nhận thấy  $A v_i$  là vector riêng của ma trận  $AA^T$ . Như vậy thay vì tính vector riêng của ma trận  $AA^T$  (kích thước  $N^2 \times N^2$ ) ta sẽ đi tính các vector riêng của ma trận  $A^T A$  kích thước  $M \times M$  nhỏ hơn rất nhiều tại bước 6.2 sau đây.

*Bước 6.2:* Tính M vector riêng của  $AA^T$ :  $u_i = A v_i$  và chuẩn hóa các  $u_i$

- **Bước 7:** Giữ lại K vector riêng  $u_1, u_2, \dots, u_K$  tương ứng với K giá trị riêng lớn nhất.

Mỗi ảnh trong tập dữ liệu vào được biểu diễn bởi (mỗi ảnh trong tập dữ liệu đầu vào sẽ chuyển thành biểu diễn của tổ hợp tuyến tính của các vector mặt riêng):

$$\hat{\Phi}_i - \Psi = \sum_{j=1}^K w_j u_j, \quad (w_j = u_j^T (\Gamma_i - \Psi)), \quad K \ll N.$$

$u_j$  ( $j = 1, \dots, K$ ) được gọi là các mặt riêng (*eigenfaces*)

#### 2.2.2.2 Pha nhận dạng

Cũng giống như bước xử lý đầu ở trên, trước tiên ta đọc ảnh I vào từ tập thử nghiệm rồi chuyển ảnh về mức xám và vector 1 chiều  $\Gamma$ . Sau đó tiến hành



chuẩn hóa cho bức ảnh, tức là trừ đi giá trị vector trung bình mặt hay nói cách khác là tính sai số của ảnh so với vector mặt trung bình.

**Bước 1:** Chuẩn hóa  $\Gamma$ :  $\Phi = \Gamma - \Psi$

**Bước 2:** Đưa vào không gian đặc trưng

Sau khi đã chuẩn hóa, chúng ta chiếu ảnh thử nghiệm vào không gian mặt riêng như sau:

$$\hat{\Phi} = \sum_{i=1}^K w_i u_i, (w_i = u_i^T \Phi), (\|u_i\| = 1)$$

Trong đó  $K$  là số vector đặc trưng của không gian ảnh  $u_i$  ( $i = 1..K$ ) là  $K$  vector đặc trưng lớn nhất  $w_i$  là các trọng số tương ứng với các vector đặc trưng

**Bước 3:** Ảnh sau khi được đưa vào không gian trị riêng thì nó sẽ được tương đương với một vector  $\Omega$  mà mỗi phần tử của  $\Omega$  là hệ số tương ứng của ảnh thử nghiệm so với các vector đặc trưng tương ứng trong không gian trị riêng.

$$\Omega = \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_K \end{bmatrix}$$

**Bước 4:** Tính khoảng cách  $e_r$

Bước này chúng ta sẽ tính khoảng cách của ma trận  $\Omega$  so với không gian mặt. Tức là ta sẽ đi tính khoảng cách tới từng bức ảnh trong tập ảnh luyện và tìm ra khoảng cách tới bức ảnh luyện gần nhất trong không gian mặt. Khoảng cách này được gọi là khoảng cách trong không gian mặt (distance within the face space).

$$e_r = \min_l \|\Omega - \Omega^\ell\|$$

$$\|\Omega - \Omega^k\| = \sum_{i=1}^K (w_i - w_i^k)^2, \text{ tính theo khoảng cách O cơ lit}$$

$$\|\Omega - \Omega^k\| = \sum_{i=1}^K \frac{1}{\lambda_i} (w_i - w_i^k)^2, \text{ tính theo khoảng cách Mahalanobis}$$

**Bước 5:** So sánh khoảng cách ngưỡng và rút ra kết luận

Sau khi tính được khoảng cách, ta so sánh với khoảng cách ngưỡng  $T_r$  (khoảng cách ngưỡng này được xác định thực nghiệm tùy thuộc vào từng tập ảnh luyện và tùy thuộc vào số vector đặc trưng được giữ lại). Trên thực tế ta sử dụng hàm khoảng cách theo Mahalanobis sẽ cho kết quả tốt hơn.

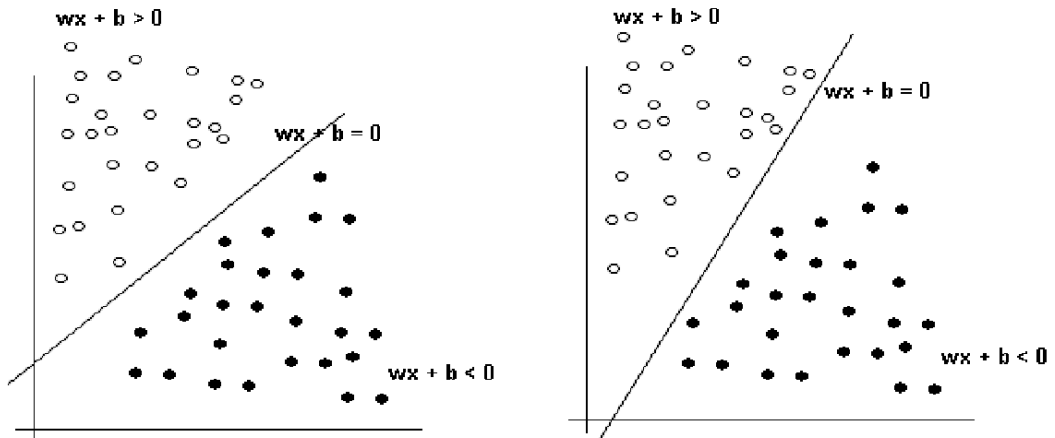
Nếu  $e_r < T_r$  thì  $\Gamma$  được nhận diện là khuôn mặt thứ  $\ell$  trong tập luyện còn không thì  $\Gamma$  không là khuôn mặt nào trong tập luyện đã cho.

## 2.4 Phương pháp SVM

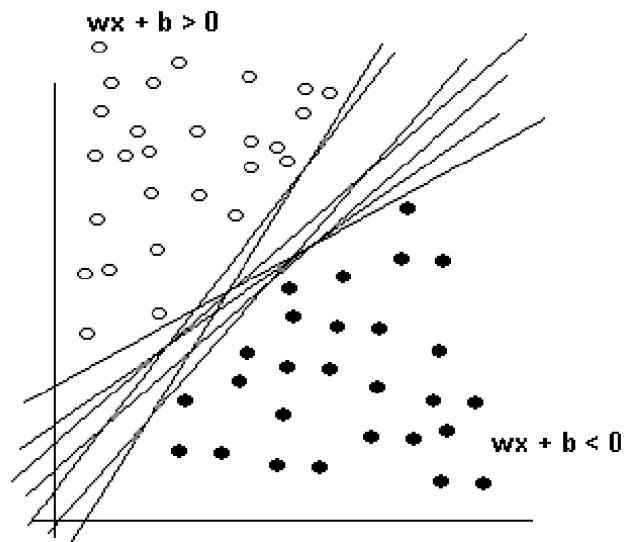
Support Vector Machine (SVM) là một phương pháp phân lớp dựa trên lý thuyết thống kê, được đề xuất bởi Vapnik [10]. Ý tưởng chính của thuật toán này là cho trước một tập huấn luyện được biểu diễn trong không gian vector (giả sử xét bài toán 2 lớp), phương pháp này tìm ra một mặt phẳng quyết định tốt nhất có thể chia các điểm trên không gian này thành hai lớp riêng biệt tương ứng lớp + và lớp -. Chất lượng của siêu mặt phẳng này được quyết định bởi khoảng cách (gọi là biên) của điểm dữ liệu gần nhất của mỗi lớp đến mặt phẳng này. Khoảng cách biên càng lớn thì mặt phẳng quyết định càng tốt đồng thời việc phân loại càng chính xác. Mục đích thuật toán SVM tìm ra được khoảng cách biên lớn nhất để tạo kết quả phân lớp tốt.

### 2.4.1 Trường hợp dữ liệu có thể phân tách được

Ta sẽ bắt đầu bằng bài toán phân lớp nhị phân, dữ liệu dùng chop ha huấn luyện có thể phân tách tuyến tính. Hình 2.10 mô tả dữ liệu trong không gian hai chiều phân tách tuyến tính.



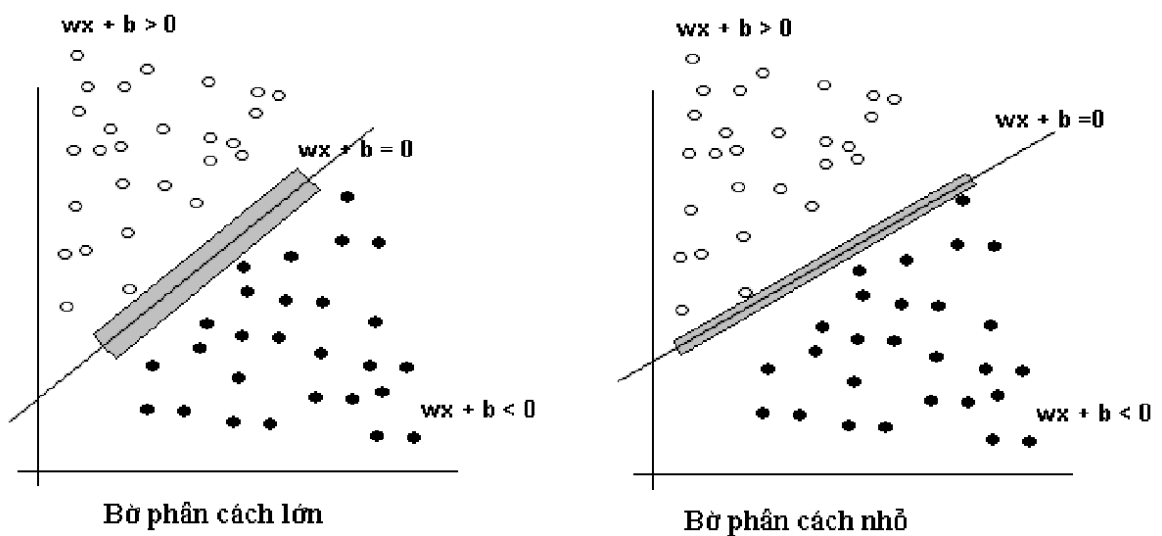
Hình 2.10 Hai cách phân lớp trên một tập dữ liệu



Hình 2.11 Nhiều cách phân lớp khác nhau

Cho tập dữ liệu huấn luyện  $\{x_i, y_i\}$ ,  $x_i \in \mathbb{R}^d$ ,  $i = 1, \dots, n$  và  $y_i \in \{-1, 1\}$ . Giả sử tồn tại một siêu mặt phẳng phân loại các dữ liệu mẫu thành hai loại mẫu dương với mẫu âm. Điểm  $x$  (trong không gian  $d$  chiều) nằm trên siêu mặt thỏa

$w \cdot x + b = 0$ , trong đó  $w$  là pháp tuyến của siêu mặt,  $|b|/\|w\|$  là khoảng cách từ siêu mặt đến gốc tọa độ, và  $\|w\|$  là độ lớn (theo khoảng cách Euclide) của  $w$ . Đặt  $(d_+)$  và  $(d_-)$  là khoảng cách ngắn nhất tương ứng từ siêu mặt phân cách đến điểm mẫu dương và âm gần nhất. Định nghĩa *bờ* (margin) của siêu mặt phân cách (kí hiệu  $r$ ) là  $(d_+)(d_-)$ . Rõ ràng với một tập dữ liệu phân tách tuyến tính sẽ tồn tại nhiều siêu mặt thỏa mãn yêu cầu, mục tiêu của thuật toán SVM sẽ tìm siêu mặt có khoảng cách bờ  $r$  là cực đại (hình 2.12).



Hình 2.12 Ví dụ về lề phân cách

Ta có thể mô hình hóa ý tưởng trên đây bằng toán học như sau: Giả sử mọi điểm trong tập mẫu thỏa các ràng buộc:

$$x_i w + b \geq +1 \quad \text{với } y_i = +1 \quad (2.1)$$

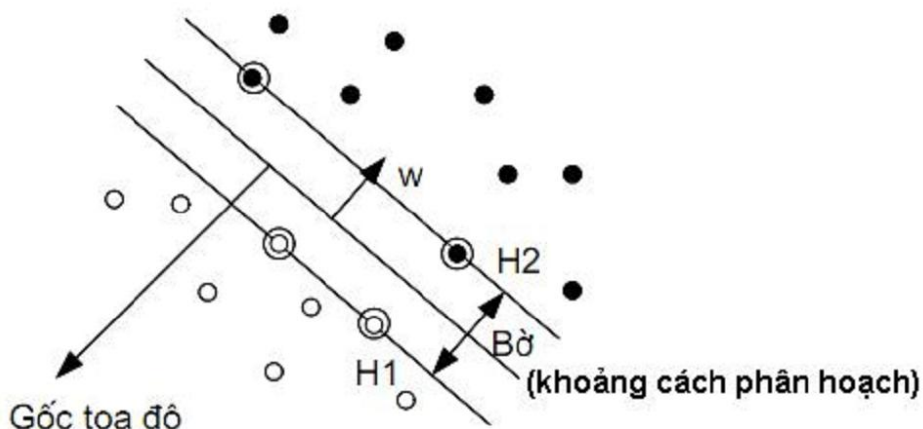
$$x_i w + b \leq -1 \quad \text{với } y_i = -1 \quad (2.2)$$

Kết hợp hai bất đẳng thức trên thành một bất đẳng thức ràng buộc như sau:

$$y_i(x_i w + b) - 1 \geq 0 \quad \forall i \quad (2.3)$$

Các mẫu dữ liệu thỏa công thức (2.1) nằm trên siêu mặt  $H_1$   $x_i w + b = 1$  có vector pháp tuyến là  $w$ , và khoảng cách đến gốc tọa độ là  $|1 - b|/\|w\|$ . Tương tự, các

mẫu thỏa công thức (2.2) nằm trên siêu mặt  $H_2: x_i w + b = -1$ , có pháp tuyến là  $w$  là khoảng cách đến gốc tọa độ là  $|-1 - b| / \|w\|$ . Khi đó,  $d_+ = d_- = 1 / \|w\|$  và bờ  $r = 2 / \|w\|$ . Lưu ý rằng  $H_1$  và  $H_2$  song song với nhau và không có điểm dữ liệu nào nằm giữa chúng. Vì vậy, ta có thể tìm cặp siêu mặt có bờ là cực đại, bằng việc cực tiểu  $\|w\|$  với ràng buộc (2.3). Ta mong muốn lời giải cho trường hợp không gian hai chiều có dạng như trong hình 2.13. Những điểm huấn luyện thỏa phương trình (2.3) (tức những điểm nằm trên một trong hai siêu mặt  $H_1, H_2$ ), và việc loại bỏ chúng sẽ làm thay đổi lời giải, được gọi là các vector hỗ trợ, đó là các điểm được bao bằng các hình tròn trong hình 2.13 và kí hiệu các support vector chính là các điểm được bao bằng viền tròn.



Hình 2.13 Siêu mặt phân cách tuyến tính cho trường hợp dữ liệu tuyến tính

Như vậy để tìm siêu phẳng tối ưu tương ứng với việc giải bài toán sau:

$$\text{cực tiểu } \frac{1}{2} \|w\|^2 \text{ với } y_i(w \cdot x_i + b) \geq 1$$

Sử dụng phương pháp Lagrange, bài toán trên chuyển về việc tìm các điểm yên ngựa của hàm sau:

$$L(w, b, \alpha) = \frac{1}{2} w^2 - \sum_{i=1}^n \alpha_i [y_i (w \cdot x_i + b) - 1]$$

với  $\alpha_i$  là các hệ số Lagrange. Phương trình trên trở thành cực tiểu  $L(w, b, \alpha)$  với  $\alpha_i \geq 0$  với mọi  $i$ .

Lấy vi phân từng phần của  $L$  ta có:

$$\frac{\partial L}{\partial w}(w, b, \alpha) = w - \sum_{i=1}^M y_i \alpha_i x_i = 0$$

$$\frac{\partial L}{\partial w}(w, b, \alpha) = \sum_{i=1}^M y_i \alpha_i = 0$$

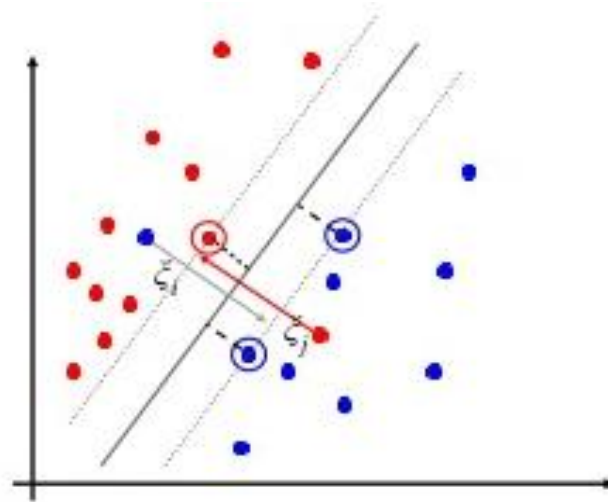
Thay vào công thức trên, bài toán tối ưu chuyển về bài toán cực đại hàm sau:

$$W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j$$

Bài toán cực đại hóa hàm mục tiêu trên là giải được theo phương pháp quy hoạch toàn phương phổ biến. Và hàm phân lớp sẽ có dạng sau:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i^0 y_i x_i \cdot x + b \right)$$

Đối với trường hợp dữ liệu có nhiễu, chẳng hạn như hình:

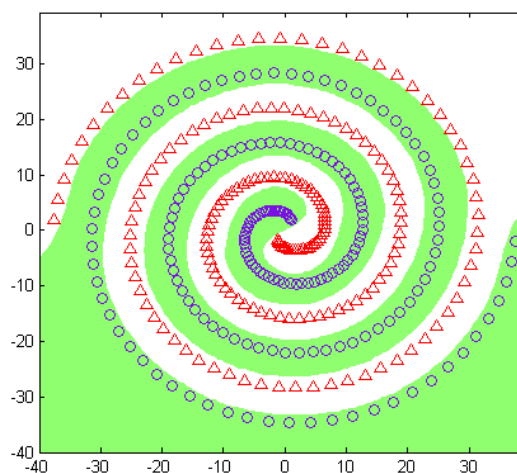
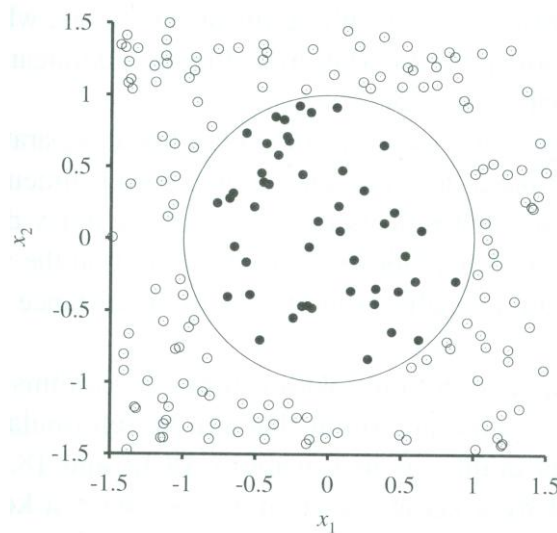


Hình 2.14 Ví dụ về dữ liệu có nhiễu

Trong hình 2.14 trên điểm màu xanh (chẳng hạn thuộc lớp +1) lại nằm vào lớp điểm màu đỏ (thuộc lớp -1) và ngược lại. Trong trường hợp này thuật toán SVM cố gắng tìm siêu phẳng phân tách và thỏa mãn lỗi phát sinh là nhỏ nhất có thể.

### 2.4.2 Trường hợp dữ liệu huấn luyện không phân tách tuyến tính

Thuật toán trên chỉ phù hợp cho trường hợp tập dữ liệu huấn luyện có thể phân tách tuyến tính được, trong phần này chúng ta sẽ đi nghiên cứu trường hợp dữ liệu huấn luyện là không phân tách tuyến tính được.



Hình 2.15 Ví dụ về dữ liệu huấn luyện không phân tách tuyến tính

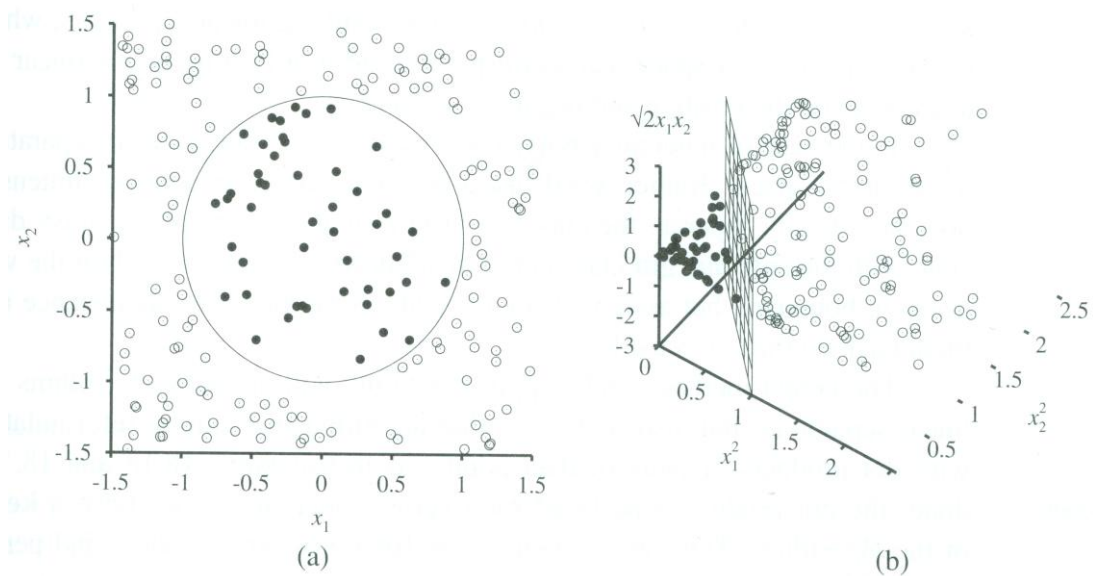
Hình 2.15 cho ta một ví dụ về hai lớp phân tách được nhưng biên phân tách không phải là tuyến tính.

Ý tưởng cơ bản để giải bài toán trên là đi tìm một phép biến đổi dữ liệu đầu vào sang không gian có số chiều lớn hơn sao cho trong không gian này dữ liệu là phân tách tuyến tính bằng cách sử dụng một hàm hạt nhân K (Kernel function).

Giả sử rằng có một phép ánh xạ  $\Phi: x \rightarrow \Phi(x)$ , ta sẽ có:

$$W(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \Phi(x_i) \cdot \Phi(x_j)$$

Thuật toán huấn luyện sẽ chỉ phụ thuộc duy nhất vào tích vô hướng trong không gian đặc trưng – tức là hàm  $\Phi(x_i) \cdot \Phi(x_j)$ . Giả sử rằng chúng ta có một hàm đối xứng  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$ , khi đó thuật toán huấn luyện chỉ cần duy nhất K mà không cần sử dụng đến ánh xạ  $\Phi$  (xem hình 2.16).



Hình 2.16 Ví dụ về việc ánh xạ dữ liệu



Cho một ánh xạ  $\Phi$ , ta sẽ tính được  $K$ , tuy nhiên cho hàm  $K$ , câu hỏi đặt ra là điều kiện gì để tồn tại ánh xạ  $\Phi$ ?. Điều này được xác định bằng điều kiện Mercer của định lý sau đây:

**Định lý 2.1 (Mercer 1909).** Cho  $K(x,y)$  là một hàm đối xứng liên tục trong  $L_2(\mathcal{X}^2)$ . Điều kiện tồn tại ánh xạ  $\Phi$  và một khai triển:

$$K(x, y) = \sum_{i=1}^{\infty} \Phi(x)_i \cdot \Phi(y)_i$$

Nếu và chỉ nếu, với mọi tập đóng  $C$  và  $g \in L_2(C)$ ,

$$\int_{C \times C} K(x, y) g(x) g(y) dx dy \geq 0$$

Khi có hàm  $K$  thỏa mãn điều kiện trên được sử dụng, hàm quyết định phân lớp sẽ là:

$$f(x) = \text{sgn} \left( \sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right)$$

Một số hàm hạt nhân được nghiên cứu và giới thiệu bởi Vapnik như sau:

*Hàm Polynomial:*  $K(x,y) = (x \cdot y + 1)^p$

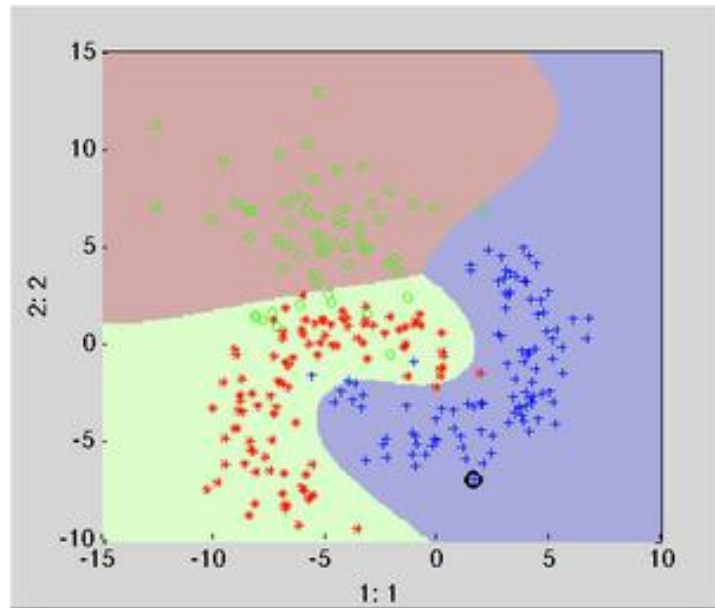
*Hàm Radial:*  $K(x,y) = \exp(-\|x-y\|^2/2\sigma^2)$

*Hàm Nơ ron:*  $K(x,y) = \text{tanh}(a \cdot x \cdot y - b)$

### 2.4.3 Phương pháp SVM cho bài toán nhiều lớp

Các nghiên cứu ở trên ứng với trường hợp bài toán phân lớp nhị phân, đối với trường hợp  $k$  lớp, chúng ta sử dụng phương án phân tách từng lớp với các lớp còn lại. Như vậy chúng ta sẽ thực hiện  $k(k-1)/2$  lần phân lớp nhị phân trong trường hợp này.

Hình 2.17 minh họa SVM trong trường hợp nhiều lớp.



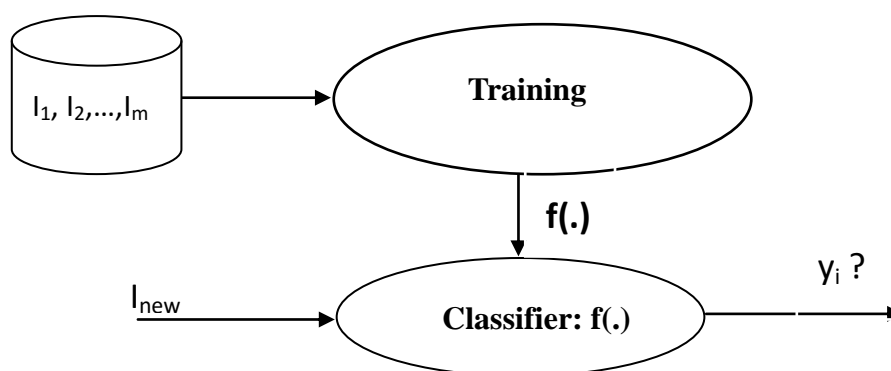
*Hình 2.17 Ví dụ về dữ liệu nhiều lớp*

## CHƯƠNG 3. KẾT QUẢ THỰC NGHIỆM

### 3.1 Giới thiệu bài toán

Mục đích của bài toán nhận dạng mặt người qua ảnh (camera) là đi xác định ảnh của một người được thu nhận có trong cơ sở dữ liệu ảnh hay không. Trong phần thực nghiệm của luận văn này, chúng tôi xây dựng một hệ thống nhận dạng mặt người qua ảnh hoặc qua thu trực tiếp từ camera. Kết quả trả ra là định danh hay tên của người đó.

Sơ đồ tổng quan của hệ thống như sau:



Hình 3.1 Hệ thống nhận dạng mặt người

Cho  $m$  ảnh mặt người kèm theo nhãn của nó (nhãn ở đây là định danh của mỗi khuôn mặt – chẳng hạn tên), hãy xây dựng hệ thống cho phép xác định khuôn mặt mới  $I_{\text{new}}$  là nhãn bao nhiêu?

Bài toán trên có nhiều ứng dụng, chẳng hạn như việc xác thực ảnh cho bài toán tuyển sinh: Hàng năm các trường Đại học đều tuyển các thí sinh mới, việc xác nhận các thí sinh này so với ảnh đã đăng ký thi đại học (hay tại kỳ thi quốc gia chung) là rất cần thiết. Trên cơ sở dữ liệu ảnh thí sinh đã sử dụng để tham dự kỳ thi, sau khi vào trường, thí sinh sẽ được yêu cầu chụp ảnh mới để làm thẻ sinh viên, thẻ thư viện,... Hệ thống nhận dạng mặt người sẽ kiểm chứng ảnh của

thí sinh lúc vào trường có trùng với ảnh của thí sinh lúc đăng ký dự thi hay không ?.

## **3.2 Lựa chọn giải thuật và lập trình**

### **3.2.1 Giải thuật**

Như đã trình bày trong chương 2, có nhiều phương pháp để giải quyết bài toán nhận dạng, trong khuôn khổ luận văn này tôi chọn phương pháp PCA - một phương pháp đơn giản dễ hiểu, dễ sử dụng.

Hệ thống nhận dạng mặt người gồm các bước sau:

- Đọc dữ liệu huấn luyện
- Trích chọn đặc trưng (ở đây là trích ra các khuôn mặt)
- Sử dụng thuật toán PCA
- Kiểm thử với dữ liệu mới

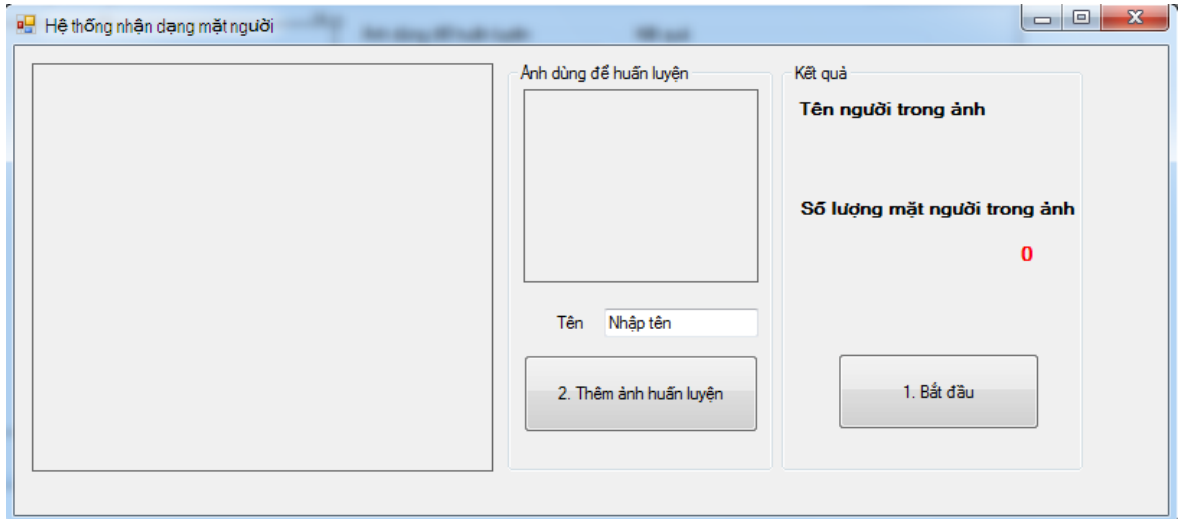
Sau khi có tập dữ liệu dùng cho pha huấn luyện thì vấn đề khó nhất là việc phát hiện khuôn mặt (face detection), có một số thuật toán phát hiện khuôn mặt đã được nghiên cứu và sử dụng. Trong phần thực nghiệm này tôi sử dụng thuật toán Haarcascade, một trong những thuật toán nổi tiếng nhất về phát hiện khuôn mặt. Toàn bộ hàm thực hiện điều này có trong thư viện Emgu CV của cộng đồng nghiên cứu và chúng ta hoàn toàn có thể sử dụng nó.

### **3.2.2 Công cụ phát triển**

Ngôn ngữ lập trình chúng tôi sử dụng là C#, một trong những môi trường lập trình mạnh của bộ Visual Studio. Từ các công cụ đã có, chúng tôi thiết kế giao diện cũng như xử lý việc ghi dữ liệu cũng như sử dụng hàm trích chọn đặc trưng theo phương pháp PCA.

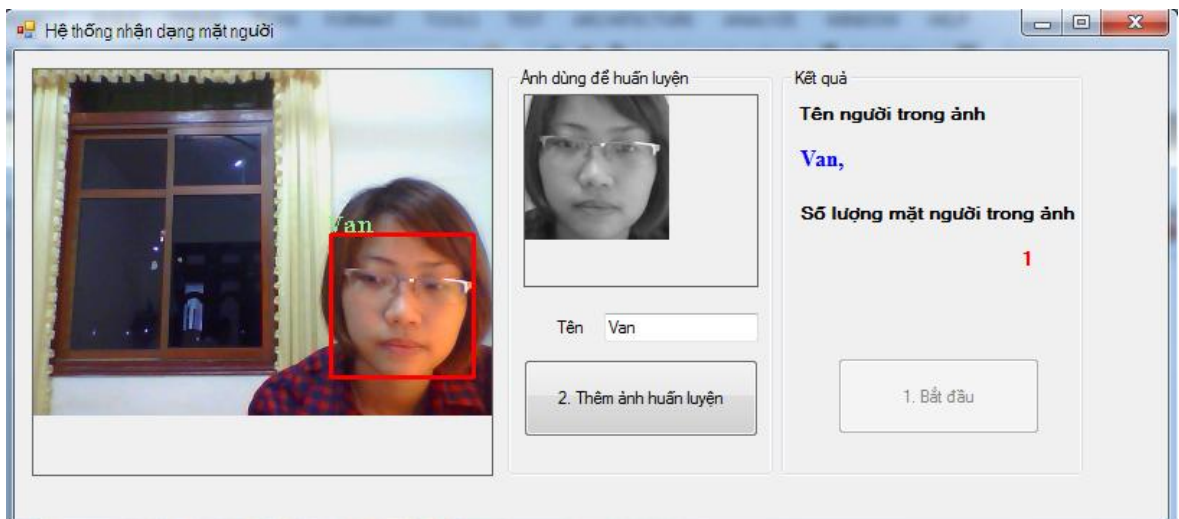
### 3.3 Kết quả thực nghiệm

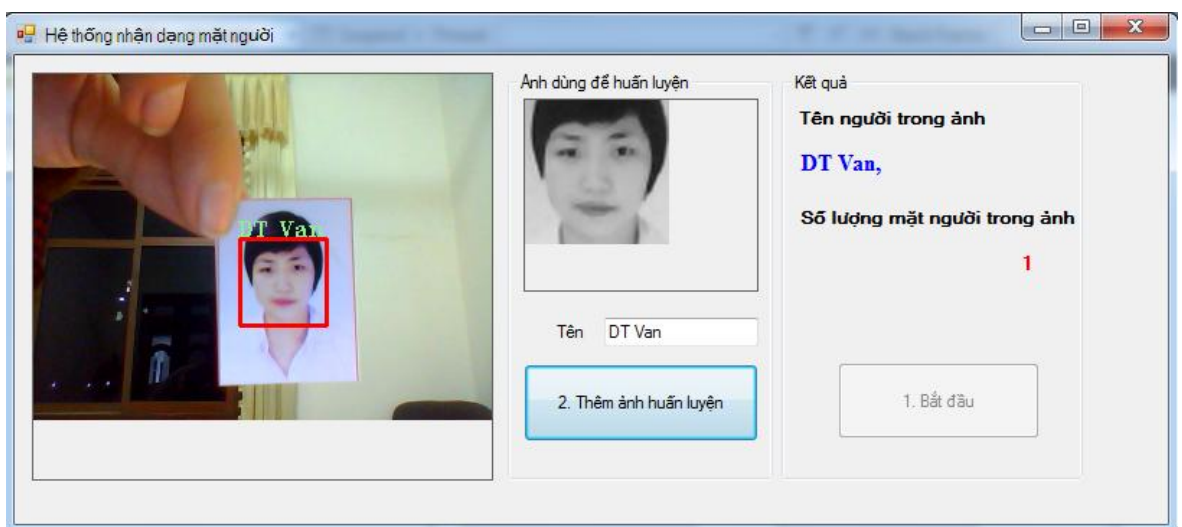
Giao diện của hệ thống như hình vẽ



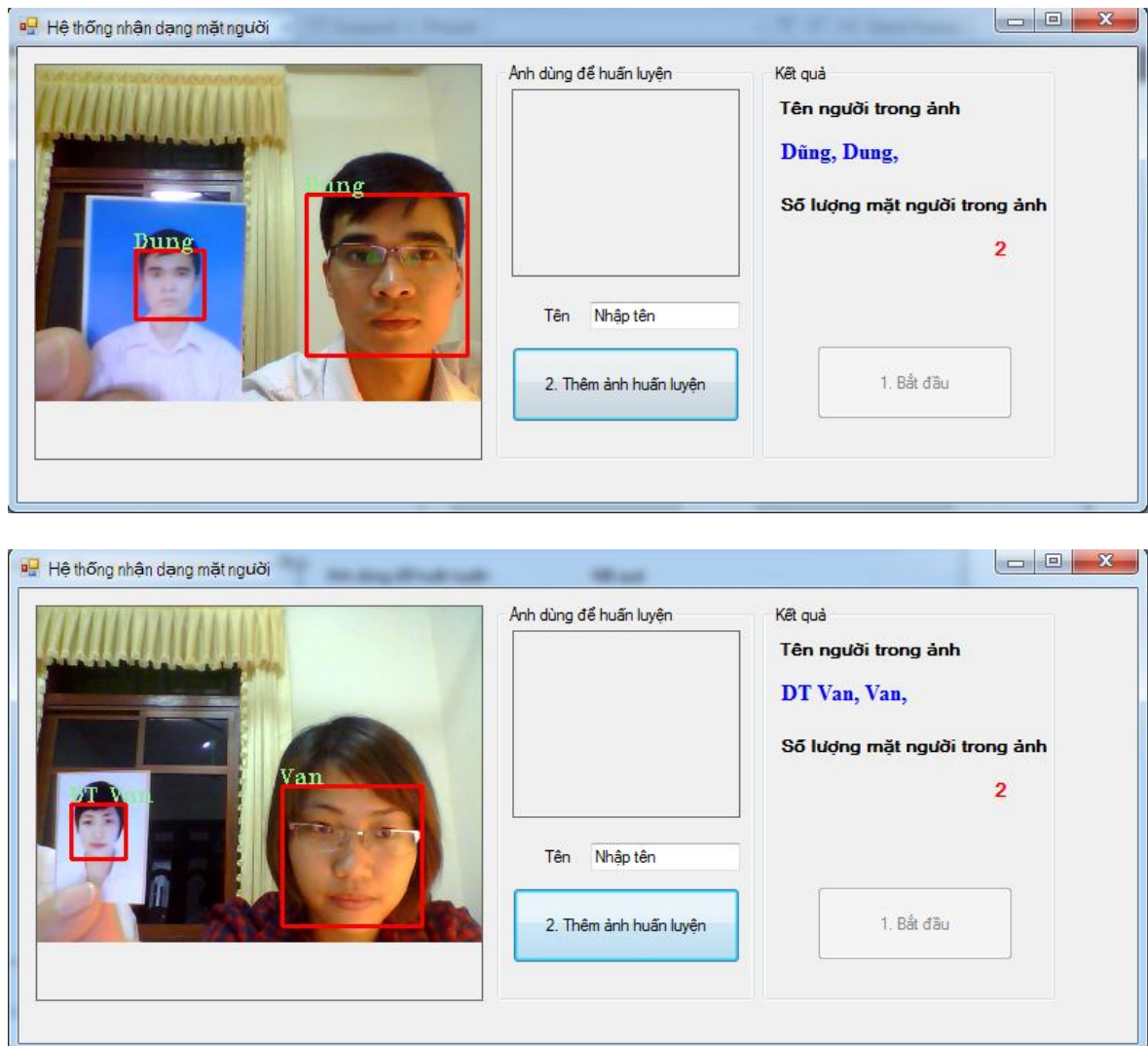
Hình 3.2 Giao diện của hệ thống

Từ giao diện của chương trình, chúng ta tạo các tập huấn luyện bằng việc sử dụng chức năng Thêm ảnh huấn luyện. Ảnh này sẽ được cập nhật vào tập huấn luyện cùng với tên của nó.



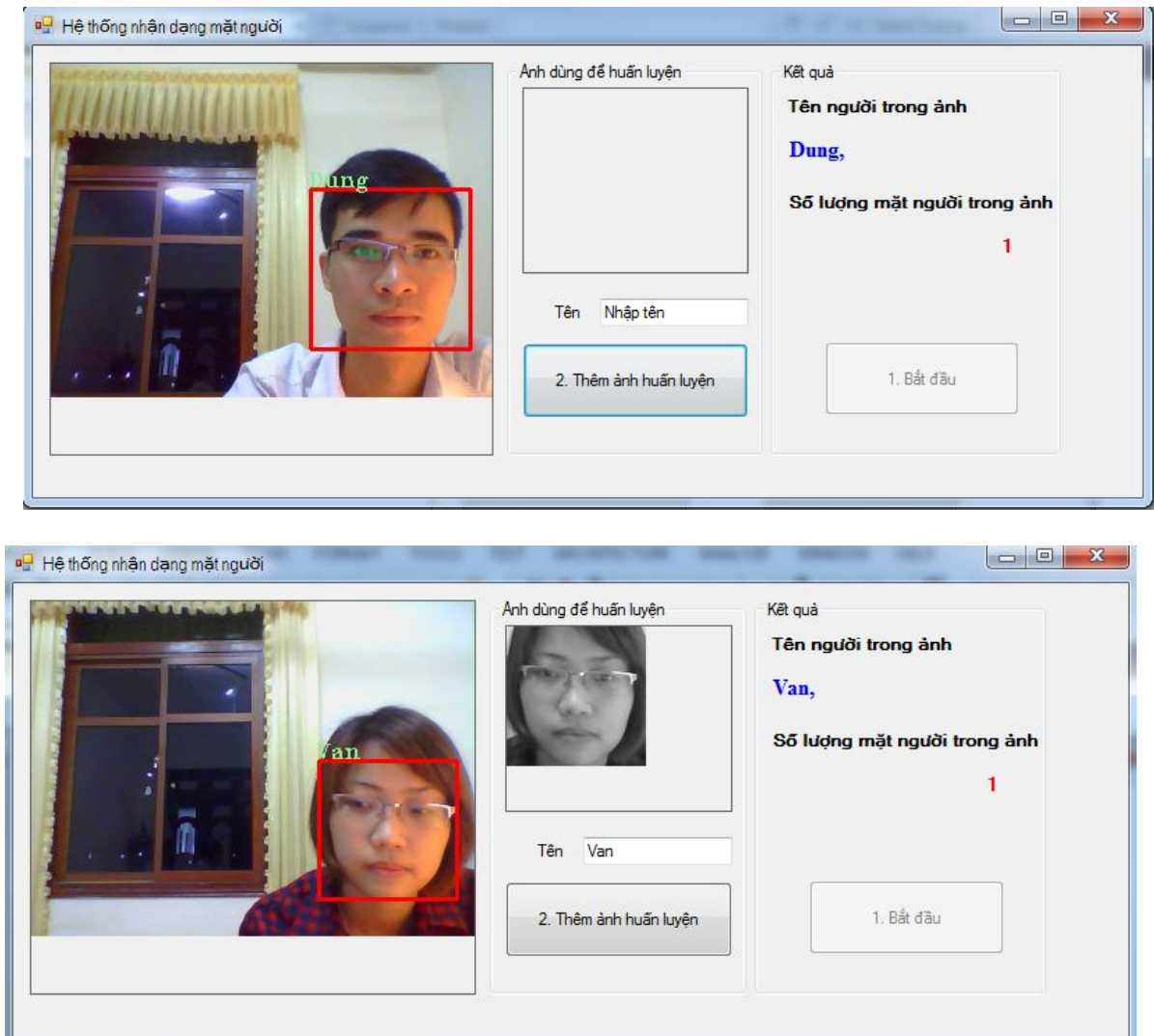


Hình 3.3 Quá trình huấn luyện của hệ thống



Hình 3.4 Pha nhận dạng từ ảnh thẻ

Để kiểm tra xem nhãn của khuôn mặt mới là gì chúng ta có thể sử dụng came ra hoặc sử dụng ảnh thẻ qua hệ thống camera hoặc máy quét để định danh khuôn mặt trong ảnh. Kết quả thử nghiệm với nhiều trường hợp khác nhau (hình 3.4-3.5) thuật toán cho kết quả tương đối tốt và hoàn toàn có tính ứng dụng thực tế. Hệ thống này hoàn toàn đáp ứng cho bài toán xác thực ảnh trong tuyển sinh của các trường Đại học tại Việt Nam.



Hình 3.5. Giao diện khi nhận dạng mặt từ camera

Theo [14], chất lượng của bài toán nhận dạng phụ thuộc vào tập dữ liệu huấn luyện cũng như các dữ liệu thử nghiệm. Nếu hệ thống các ảnh luyện cũng như là các ảnh kiểm thử là tương đồng nhau, chỉ khác nhau về độ sáng của ảnh thì kết quả đạt được là 96%, tuy nhiên nếu thay đổi về hướng của mặt thì kết quả đạt khoảng 85%.

Kết quả trên được giả thích là nếu thay đổi độ sáng thì các điểm kề nhau vẫn giữ được sự tương quan của nó, trong khi nếu chúng ta thay đổi hướng của khuôn mặt sẽ làm mất đi thông tin giữa các điểm ảnh nên chất lượng nhận dạng



ảnh sẽ giảm đi rõ rệt. Rõ ràng trong bài toán nhận dạng ảnh thì quy định về ảnh như các ảnh phải có cùng kích thước, mặt người phải ở tâm ảnh,... là quan trọng vì để giải quyết bài toán dạng này chúng ta cần biểu diễn dạng vector các số nên nếu ảnh đầu vào có sai khác thì giá trị các phép toán cũng sai khác theo.

## CHƯƠNG 4. KẾT LUẬN

### Những kết quả đã đạt được

Sau khi thực hiện luận văn với chủ đề nghiên cứu phương pháp phân phân lớp PCA và SVM và ứng dụng cho bài toán nhận dạng mặt người, tôi đã thực hiện được các công việc sau đây:

- Đã nghiên cứu và tìm hiểu hai phương pháp PCA và SVM cũng như các lý thuyết tổng quát về trí tuệ nhân tạo và học máy
- Đã tìm hiểu và nắm bắt lý thuyết về xử lý ảnh, các kỹ thuật cơ bản dùng trong xử lý ảnh cũng như ứng dụng của xử lý ảnh trong thực tế,
- Đã lập trình trên ngôn ngữ C# cho bài toán nhận dạng mặt người trong ảnh ứng dụng cho bài toán điểm danh. Các kết quả cho thấy đề tài có khả năng ứng dụng vào thực tiễn trong các ứng dụng về xử lý ảnh. Bước đầu hiểu và biết cách sử dụng thư viện Emgu CV, một thư viện nổi tiếng về thị giác máy của cộng đồng các nhà nghiên cứu về lĩnh vực này trên thế giới.

### Hướng phát triển của đề tài

Do thời gian và kiến thức còn hạn chế, trong khuôn khổ của luận văn tôi không thể nghiên cứu kỹ và toàn diện bài toán phân cụm. Trong tương lai, một số hướng nghiên cứu mà tôi dự kiến tiếp tục như sau:

- Trong thời gian tới tôi mong muốn tìm hiểu kỹ hơn về các phương pháp phân lớp khác như mạng Nơ ron hoặc phương pháp cây quyết định.
- So sánh các phương pháp phân lớp hiện tại với các tập dữ liệu khác nhau trong các không gian nhiều chiều, dữ liệu phức tạp và có thể không phải là dữ liệu số.
- Tôi cũng mong muốn tìm hiểu sâu hơn về các bài toán học máy khác như phân cụm, phương pháp học nửa giám sát cũng như so sánh

phương pháp phân lớp có giám sát và phân cụm nửa giám sát.

- Cuối cùng tôi mong muốn có thể áp dụng được các lý thuyết về học máy cũng như trí tuệ nhân tạo vào các bài toán trong thực tế như sinh học, y học, ngân hàng,...

## Tài liệu tham khảo

### Tiếng Việt

- [1]. Phạm Việt Bình, Đỗ Năng Toàn (2008), *Xử lý ảnh*, Nhà xuất bản Khoa học và Kỹ thuật, 2008.
- [2]. Vũ Mạnh Hùng (2013), *Nhận dạng mặt người sử dụng đặc trưng PCA*, Luận Văn Thạc sỹ Công nghệ Thông tin, Học Viện Công nghệ Bưu Chính Viễn thông (Hà Nội).
- [3]. Nguyễn Thi Len (2014), *Đề xuất phương pháp nhận dạng khuôn mặt người và ứng dụng chống gian lận trong thi cử*, Luận Văn Thạc sỹ Công nghệ Thông tin, Học Viện Công nghệ Bưu Chính Viễn thông (Hà Nội).
- [4]. Nguyễn Thị Lan Anh (2013), *Nghiên cứu thuật toán học máy SVM và ứng dụng trong bài toán khai phá ý kiến phản hồi của khách hàng trên website*, Luận Văn Thạc sỹ ngành Hệ thống Thông tin, Học Viện Công nghệ Bưu Chính Viễn thông (Hà Nội).
- [5]. Trần Minh Tân (2012), *Nghiên cứu kỹ thuật SVM trong kiểm soát nội dung hình ảnh*, Luận Văn Thạc sỹ Công nghệ Thông tin, Trường Đại học Lạc Hồng.
- [6]. Phan Thị Thu Hồng, Đoàn Thị Thu Hà, Nguyễn Thị Thủy (2013), *Ứng dụng phân lớp ảnh chụp lá cây bằng phương pháp máy vector hỗ trợ*, Tạp chí Khoa học và Phát triển, tập 11, số 7, trang: 1045-1052.

### Tiếng Anh

- [7]. Rafael C. Gonzalez and Richard E. Woods (2007), *Digital Image Processing*, Prentice Hall; 3 edition.
- [8]. <http://www.emgu.com/>
- [9]. Paul Viola, Michael J Jones, *Robust Real-Time Face Detection*: International Journal of Computer Vision 57, pp. 137-154, Netherlands, 2004.
- [10] Vladimir N.Vapnik, *The Nature of Statistical Learning Theory*, AT&Labs-Research, London University.
- [11] Jeffrey Huang, David Li, Xuhui Shao, và Harry Wechsler, *Pose Discrimination và Eye Detection Using Support Vector Machines (SVM)*, Department of Computer Science George Mason University

- [12] Christopher J.C. Burges, A Tutorial on Support Vector Machines for Patter Recognition, (<http://citeseer.nj.nec.com/burges98tutorial.html>)
- [13] Stuart Russell, Peter Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2009.
- [14] L. Sirovich and M. Kirby, *Low-dimentional procedure for the characterization of human face*. Journal of the optical Society of America A, 4(3), 519-524, 1987.