

NHẬN DẠNG MẪU HÌNH ẢNH SỬ DỤNG MÔ-MEN HU

HU'S MOMENTS FOR VISUAL PATTERN RECOGNITION

Hoàng Lê Uyên Thực, Phạm Văn Tuấn

Trường Đại học Bách khoa, Đại học Đà Nẵng; hluthuc@dut.udn.vn, pvtuan@dut.udn.vn

Tóm tắt - Nhận dạng mẫu hình ảnh hiện đang nhận được rất nhiều sự quan tâm từ các nhà nghiên cứu do tính ứng dụng sâu rộng của nó trong nhiều lĩnh vực khác nhau. Trong bài báo này, chúng tôi tập trung tìm hiểu bài toán nhận dạng mẫu hình ảnh bao gồm ảnh tĩnh và video bằng cách dùng mô-men Hu để mô tả hình dạng đối tượng trong khung hình. Trước tiên, đối tượng quan tâm được trích ra khỏi phần còn lại của khung hình rồi được chuyển đổi thành một vec-tơ đặc trưng 7 chiều, trong đó mỗi thành phần của vec-tơ chính là một trong 7 giá trị mô-men Hu. Tiếp đến, tùy theo dữ liệu xem xét là ảnh tĩnh hay video để chọn phương pháp nhận dạng tương ứng dùng mạng nơ-ron nhân tạo hay mô hình Markov ẩn. Các thí nghiệm trên hai ví dụ nhận dạng cây đối với dữ liệu ảnh tĩnh và phát hiện dáng đi bệnh lý đối với dữ liệu video cho kết quả rất khả quan xét theo tiêu chí tỷ lệ nhận dạng đúng trung bình.

Từ khóa - mô-men Hu; nhận dạng mẫu hình ảnh; nhận dạng cây; phát hiện dáng đi bệnh lý; đặc trưng hình dạng

1. Đặt vấn đề

Nhận dạng mẫu hình ảnh (visual pattern recognition), là một lĩnh vực nghiên cứu theo hướng “hiểu” nội dung bức ảnh hoặc nội dung của đoạn video. Trong những năm gần đây, nhận dạng mẫu hình ảnh đã thu hút sự quan tâm nghiên cứu rất lớn do khả năng ứng dụng phong phú của nó trong nhiều lĩnh vực khác nhau. Đối với dữ liệu ảnh tĩnh, nhận dạng mẫu đã được ứng dụng trong nhận dạng ký tự quang học chẳng hạn như chữ viết tay, nhận dạng khuôn mặt, nhận dạng dấu vân tay, nhận dạng lá cây, v.v... Đối với dữ liệu video, nhận dạng mẫu được ứng dụng trong nhận dạng hành động con người, phát hiện các sự kiện bất thường như hành vi bạo lực, tai nạn té ngã, v.v...

Nhìn chung, một hệ thống nhận dạng mẫu hình ảnh gồm có ba bước xử lý chính là trích đối tượng, trích đặc trưng và nhận dạng [1].

Ở bước thứ nhất, đối tượng quan tâm được trích ra khỏi phần nền bằng các thuật toán trích đối tượng. Tùy theo từng ứng dụng cụ thể mà đối tượng quan tâm sẽ khác nhau, chẳng hạn như với ứng dụng nhận dạng khuôn mặt thì đối tượng quan tâm chính là khuôn mặt trong ảnh, với ứng dụng nhận dạng hành động thì đối tượng quan tâm chính là con người trong đoạn video. Đối với dữ liệu là ảnh tĩnh thì tùy vào đặc điểm của đối tượng, việc trích đối tượng có thể dựa vào mức ngưỡng, dựa vào đường viền hoặc là dựa vào cấu trúc ảnh, v.v... [2]. Đối với dữ liệu là video, phương pháp trích đối tượng phổ biến nhất là trừ nền, dựa theo nguyên lý cơ bản là trừ khung hình hiện tại cho khung hình nền tham chiếu [3].

Tiếp theo, các đặc điểm của đối tượng như hình dạng, màu sắc, sự chuyển động, v.v... trong ảnh hoặc đoạn video được trích ra và biểu diễn dưới dạng vec-tơ đặc trưng đa chiều. Để nhận dạng chính xác, các đặc trưng phải chứa đựng được các đặc tính hữu hiệu nhất và duy nhất của đối tượng. Có rất nhiều bộ mô tả đặc trưng đã được công bố,

Abstract - Visual pattern recognition has attracted great attention from researchers due to its far-reaching applications in many different fields such as optical character recognition, action recognition, abnormal behavior detection, etc. In this paper, we concentrate on the recognition of visual pattern including image and video patterns by using set of Hu's moments to describe the shape of interested objects in an image frame. Initially, we extract the object from the rest of image frame, then we transfer the extracted object into a 7-dimension feature vector, each vector component of which is one of the 7 values of Hu's moments. Next, we choose artificial neural network and hidden Markov model to recognize image patterns and video patterns, respectively. Experiments on two examples which are plant identification for image data and pathological gait detection for video data show promising results in terms of total average recognition rate.

Key words - Hu's moments; visual pattern recognition; plant identification; pathological gait detection; shape-based feature

chẳng hạn như mô-men Hu [4], ảnh lịch sử chuyển động MHI (Motion History Image) [5], ảnh cường độ chuyển động MII (Motion Intensity Image) [6], v.v...

Cuối cùng, một thuật toán phân loại được áp dụng vào các vec-tơ đặc trưng để nhận dạng các đối tượng khác nhau. Để thực hiện được nhiệm vụ này, hệ thống phải trải qua quá trình “huấn luyện”, tức là phân tích dữ liệu huấn luyện để nắm bắt các đặc điểm của các phân bố xác suất chưa biết của dữ liệu. Từ đó đưa ra các suy luận đối với dữ liệu kiểm tra chưa biết. Có nhiều phương pháp nhận dạng đã được đề xuất, tiêu biểu như mạng nơ-ron nhân tạo ANN (Artificial Neural Network) [7], mô hình Markov ẩn (HMM) [8], v.v...

Trong bài báo này, chúng tôi tập trung xét trường hợp nhận dạng mẫu hình ảnh đơn giản nhất là phân loại 2 lớp. Hai ứng dụng được chọn ở đây là nhận dạng cây trong ảnh tĩnh và phát hiện dáng đi bệnh lý trong đoạn video. Phần tiếp theo của bài báo gồm các nội dung chính như sau: mục 2 giới thiệu bài toán nhận dạng, mục 3 trình bày chi tiết về hệ thống nhận dạng mẫu hình ảnh đề xuất, mục 4 nêu các thí nghiệm kiểm tra, đánh giá hệ thống đề xuất và cuối cùng là kết luận ở mục 5.

2. Giới thiệu bài toán nhận dạng

Như đã nêu trên, mục này giới thiệu hai bài toán nhận dạng mẫu hình ảnh được xét trong bài báo.

2.1. Bài toán nhận dạng cây trong ảnh tĩnh

Nhận dạng cây (plant identification) là một phần việc quan trọng trong các nghiên cứu về đa dạng sinh học, nhằm nhận dạng các loài cây quý hiếm, phát hiện loài cây mới, sắp xếp phân loại các loài cây cũng như xác định phân bố địa lý của chúng [9]. Ngày nay, thay vì thực hiện nhận dạng cây theo kiểu truyền thống thì việc nhận dạng tự động nhờ vào các công cụ phần mềm bắt đầu phát triển. Phương pháp nhận dạng mới được thực hiện hoàn toàn tự động nên mang tính

khách quan và góp phần đáng kể vào việc tiết kiệm thời gian và nhân lực ngành phân loại học hiện đang bị thiếu hụt.

Để thực hiện nhận dạng cây trong ảnh tĩnh, chúng tôi thử nghiệm với hoa hồng và hoa cúc. Đây là hai loài hoa rất phổ biến, đa dạng, nhiều màu sắc và đặc biệt là ngay cả trong cùng họ hoa hồng hoặc hoa cúc thì các loài khác nhau cũng có vẻ bề ngoài rất khác nhau. Hình 1 là một số mẫu hoa hồng và hoa cúc được sử dụng trong bài toán.



Hình 1. Mẫu hoa hồng (hàng trên) và hoa cúc (hàng dưới)

2.2. Bài toán phát hiện dáng đi bệnh lý trong đoạn video

Dáng đi bộ là một dấu hiệu quan trọng biểu hiện sức khỏe con người. Một số vấn đề về sức khỏe như đau khớp, tổn thương mô, tổn thương thần kinh, v.v... đều gây ra sự bất thường trong dáng đi. Như vậy, việc phát hiện sớm dáng đi bệnh lý là rất quan trọng, giúp can thiệp y khoa kịp thời. Ngày nay, cùng với thăm khám lâm sàng thì phương pháp phát hiện dáng đi bệnh lý tự động đã bắt đầu nhận được nhiều quan tâm nghiên cứu [10]. Phương pháp mới được thực hiện tự động dựa vào tín hiệu video liên tục ghi hình bệnh nhân đi bộ nên giúp phát hiện kịp thời và khách quan các dáng đi bệnh lý.

Để thực hiện phát hiện dáng đi bệnh lý trong đoạn video, chúng tôi thử nghiệm với trường hợp phát hiện dáng đi bệnh Parkinson với góc quay ngang. Parkinson là căn bệnh rất phổ biến ở người cao tuổi, thể hiện ở dáng đi thiếu hoặc ít đánh tay tự nhiên, tay run, bước đi ngắn, lưng gập chúi người về trước [11]. Hình 2 là một số khung hình trích trong đoạn video dáng đi Parkinson.



Hình 2. Mẫu dáng đi bệnh Parkinson trong đoạn video

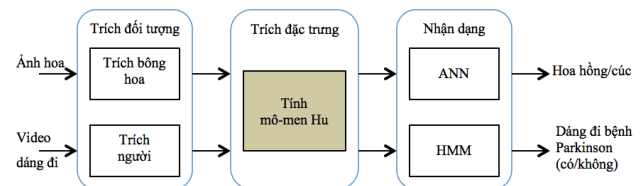
Hai bài toán nêu trên, mặc dù thoát nhìn khác hẳn nhau nhưng đều có một điểm chung là có thể sử dụng đặc trưng hình dạng (shape-based feature) để mô tả các đặc điểm về hình dạng của đối tượng cần nhận dạng. Đối với bài toán nhận dạng hoa, qua quan sát thấy hầu hết các bông hoa trong cùng họ hồng hoặc cúc, dù thuộc các loài khác nhau thì vẫn có những đặc điểm riêng về hình dạng khác với các bông hoa thuộc họ khác. Đối với bài toán phát hiện dáng đi bệnh Parkinson ở góc quay ngang, qua quan sát thấy dù mức độ bệnh nặng nhẹ khác nhau, dù đối tượng quan sát là khác nhau thì về bên ngoài của dáng đi bệnh Parkinson vẫn

có những đặc điểm riêng khác với dáng đi khỏe mạnh. Vì thế, một bộ mô tả đặc trưng hình dạng tốt có thể giúp cho quá trình nhận dạng được thành công. Qua xem xét các bộ mô tả đặc trưng hình dạng sẵn có, mô-men Hu được chọn dùng cho bài toán đặt ra, do mô-men Hu có tính chất bất biến đối với phép dịch chuyển, co giãn và phép quay [4].

Trong phần nhận dạng, mạng nơ-ron nhân tạo [7] được chọn để nhận dạng hoa hồng và hoa cúc trong ảnh tĩnh và mô hình Markov ẩn [8] được chọn để phát hiện dáng đi bệnh Parkinson từ tín hiệu video, do các ưu điểm nổi bật như hiệu quả nhận dạng cao và có sẵn nhiều công cụ phần mềm hữu hiệu để hỗ trợ cho huấn luyện/kiểm tra.

3. Hệ thống nhận dạng mẫu hình ảnh đề xuất

Hình 3 mô tả tổng quan hệ thống nhận dạng đề xuất, bao gồm 3 khối chức năng chính là trích đối tượng, trích đặc trưng và nhận dạng sẽ được trình bày lần lượt trong các mục 3.1, 3.2 và 3.3.



Hình 3. Sơ đồ khối tổng quát hệ thống đề xuất

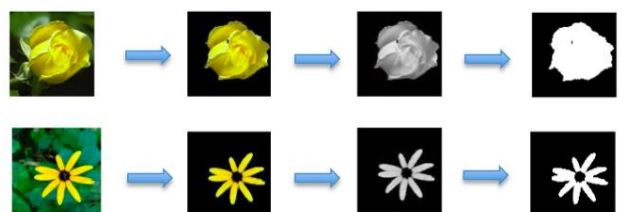
3.1. Trích đối tượng

Bước này nhằm trích đối tượng quan tâm ra khỏi phần còn lại của khung hình. Phương pháp trích đối tượng được thực hiện khác nhau tùy vào đối tượng.

3.1.1. Trích bông hoa trong ảnh

Nhìn chung, trong một bức ảnh hoa ta thấy có 3 vùng màu cơ bản tương ứng với màu hoa, lá và phần còn lại. Vì thế có thể thực hiện trích bông hoa dựa vào màu sắc, bằng cách phân nhóm ảnh ra thành 3 nhóm là nhóm có màu hoa, nhóm có màu lá và nhóm màu còn lại. Thuật toán phân nhóm được chọn ở đây là K-means với nguyên tắc là nhóm các điểm ảnh có màu gần nhau lại thành một nhóm [12]. Ở đây, số nhóm được chọn là $K = 3$ do có 3 nhóm màu trong một bức ảnh, trọng tâm ban đầu của 3 nhóm được chọn ngẫu nhiên.

Sau khi phân bức ảnh ra làm 3 vùng tương ứng với 3 nhóm màu, ta trích lấy vùng màu của cánh hoa để được ảnh bông hoa trên nền đen. Tiếp đó thực hiện chuyển ảnh bông hoa trên nền đen thành ảnh xám và cuối cùng là ảnh nhị phân với điểm trắng thuộc về bông hoa và điểm đen thuộc về nền. Kết quả các bước xử lý trích bông hoa trong ảnh được biểu diễn trên Hình 4.



Hình 4. Kết quả trích bông hoa trong ảnh

3.1.2. Trích con người trong đoạn video

Như hầu hết các hệ thống nhận dạng hành động khác, đối tượng con người được trích ra khỏi nền bằng phương pháp trừ nền, với mô hình nền được xây dựng bằng mô hình hợp Gauss GMM (Gaussian Mixture Model) [3]. Trong phương pháp này, mỗi điểm ảnh được biểu diễn bằng một tổng có trọng số các phân bố Gauss và được cập nhật qua từng khung với các phân bố Gauss mới.

Sau khi trừ nền, ảnh trích đối tượng được xử lý qua một số phép hình thái toán học [2] để được chất lượng tốt hơn. Cuối cùng, để giảm kích thước ảnh, ta chỉ trích lấy vùng quan sát ROI (Region Of Interest) - là vùng hình chữ nhật có tâm trùng với trọng tâm của đối tượng và bao quanh đối tượng. Hình 5 là một ví dụ về trích đối tượng con người và ROI tương ứng.



Hình 5. Kết quả trích đối tượng con người trong đoạn video

3.2. Trích đặc trưng hình dạng

Trích đặc trưng chính là chuyển đổi đối tượng trích được thành một vec-tơ đặc trưng đa chiều sao cho vec-tơ này chứa đựng các đặc điểm hữu hiệu và riêng biệt, giúp phân biệt đối tượng này với đối tượng khác mà không cần phải dùng toàn bộ khung hình.

Như đã trình bày ở mục 2.2, cả đối tượng bông hoa và con người đều có thể được mô tả bằng đặc trưng hình dạng dùng mô-men Hu.

Quá trình tính toán các giá trị mô-men Hu làm đặc trưng hình dạng được tóm tắt trong 4 bước sau đây [4]:

- **Bước 1:** Tính các mô-men 2 chiều:

$$m_{pq} = \iint x^p y^q r(x, y) \quad (1)$$

ở đây: (x, y) là tọa độ điểm ảnh, $p(x, y)$ là hàm ảnh nhị phân, là 1 hoặc là 0 tùy theo điểm ảnh (x, y) thuộc vùng đối tượng hoặc vùng nền tương ứng.

- **Bước 2:** Tính các mô-men trung tâm nhằm làm cho các mô-men 2 chiều ở (1) trở nên bất biến đối với sự dịch chuyển của ảnh nhị phân trong khung hình:

$$m_{pq} = \iint (x - \bar{x})^p (y - \bar{y})^q r(x, y) \quad (2)$$

ở đây điểm $(\bar{x}, \bar{y})$ là trọng tâm của ảnh nhị phân: $\bar{x} = \frac{m_{10}}{m_{00}}$ và

$$\bar{y} = \frac{m_{01}}{m_{00}}.$$

- **Bước 3:** Chuẩn hóa các mô-men trung tâm ở (2) để chúng bất biến đối với sự co giãn của ảnh nhị phân:

$$h_{pq} = \frac{m_{pq}}{m_{00}^{p+q+1}} \quad (3)$$

- **Bước 4:** Tính 7 mô-men Hu dựa vào các mô-men trung tâm chuẩn hóa ở (3) theo công thức (4) sau:

$$\begin{aligned} S_1 &= h_{20} + h_{02} \\ S_2 &= (h_{20} - h_{02})(h_{20} - h_{02}) + 4h_{11}h_{11} \\ S_3 &= (h_{30} - 3h_{12})(h_{30} - 3h_{12}) + (h_{30} - 3h_{21})(h_{30} - 3h_{21}) \\ S_4 &= (h_{30} + h_{12})(h_{30} + h_{12}) + (h_{03} + h_{21})(h_{03} + h_{21}) \\ S_5 &= (h_{30} - 3h_{12})(h_{30} + h_{12})[(h_{30} + h_{12})(h_{30} + h_{12}) - 3(h_{03} + h_{21})(h_{03} + h_{21})] \\ &\quad + (3h_{21} - h_{03})(h_{03} + h_{21})[3(h_{30} + h_{12})(h_{30} + h_{12}) - (h_{03} + h_{21})(h_{03} + h_{21})] \\ S_6 &= (h_{20} - h_{02})[(h_{30} + h_{12})(h_{30} + h_{12}) - (h_{03} + h_{21})(h_{03} + h_{21})] \\ &\quad + 4h_{11}(h_{30} + h_{12})(h_{03} + h_{21}) \\ S_7 &= (3h_{21} - h_{03})(h_{30} + h_{12})[(h_{30} + h_{12})(h_{30} + h_{12}) - 3(h_{21} + h_{03})(h_{21} + h_{03})] \\ &\quad - (h_{30} - 3h_{12})(h_{21} + h_{02})[3(h_{30} + h_{12})(h_{30} + h_{12}) - (h_{21} + h_{03})(h_{21} + h_{03})] \end{aligned} \quad (4)$$

Khi tính toán các mô-men Hu, ta nhận thấy giá trị của các mô-men này rất nhỏ đến nỗi gần như bằng 0 (một số còn mang giá trị âm). Do đó cần phải ánh xạ các điểm vec-tơ đặc trưng gốc là tập gồm 7 mô-men rất gần nhau sang một không gian mới, ở đó chúng cách nhau đủ xa để xử lý ở bước tiếp theo. Điều này được thực hiện đơn giản thông qua tính logarit của trị tuyệt đối của các mô-men. Bằng thực nghiệm, chúng tôi cũng đã chứng minh được việc chuyển đổi mô-men Hu sang không gian mới nhờ phép logarit, giúp tăng tỷ lệ nhận dạng của toàn hệ thống lên đáng kể. Bảng 1 đưa ra kết quả tính mô-men Hu của ảnh hoa cúc thứ 25 trong cơ sở dữ liệu hoa.

Bảng 1. Tập đặc trưng mô-men Hu của một mẫu dữ liệu ảnh tĩnh trước và sau khi tính logarit

Mô-men	Trước logarit	Sau logarit
S ₁	0,273778883916765	-1,2954
S ₂	0,003420124059901	-5,6781
S ₃	0,001017048934577	-6,8909
S ₄	0,000025592197673	-10,5732
S ₅	0,000000004114001	-19,3089
S ₆	-0,000000312461600	-14,9788
S ₇	-0,000000102712582	-16,0913

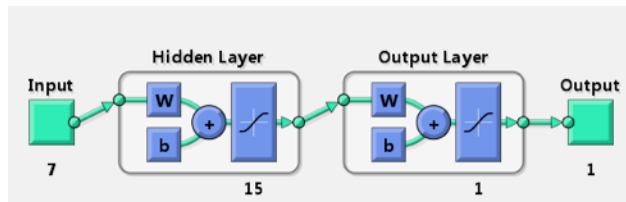
3.3. Nhận dạng

Nhiệm vụ cuối cùng của hệ thống chính là xác định vec-tơ đặc trưng trích từ bức ảnh (hoặc chuỗi vec-tơ đặc trưng trích từ chuỗi khung video) thuộc vào loại mẫu nào trong các loại mẫu quan tâm đã huấn luyện cho hệ thống.

3.3.1. Nhận dạng hoa

Nhận dạng hoa được thực hiện nhờ mạng nơ-ron nhân tạo ANN [7]. Trong số các loại mạng ANN thì loại perceptron nhiều lớp MLP (Multi Layer Perceptron) là nổi tiếng và phổ biến nhất [7]. Do đó MLP được chọn dùng trong hệ thống đề xuất. MLP là mạng ANN được tổ chức theo kiểu phân lớp với 3 lớp chính là lớp vào, lớp ra và các lớp ẩn. Trong hệ thống, số nơ-ron lớp vào là 7, số lớp ẩn là 1, số nơ-ron lớp ẩn là 15, hàm kích hoạt là hàm sigmoid, số nơ-ron lớp ra là 1 như trên Hình 6.

Hình 6 cho thấy cấu trúc của mạng nơ-ron MLP dùng trong hệ thống đề xuất.

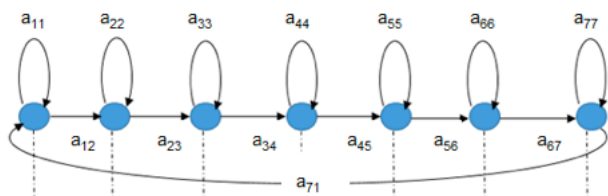


Hình 6. Mô hình mạng nơ-ron dùng trong hệ thống
(w: trọng số, b: ngưỡng kích thích của nơ-ron)

3.3.2. Nhận dạng dáng đi bệnh lý

Nhận dạng dáng đi là một trường hợp riêng của nhận dạng hành động nên có thể áp dụng mô hình nhận dạng hành động phổ biến nhất là mô hình Markov ẩn HMM vào bài toán. Một mô hình HMM được xác định bởi các tham số: số trạng thái ẩn N , số ký hiệu quan sát phân biệt M , ma trận chuyển đổi trạng thái A , ma trận quan sát B và phân bố trạng thái khởi đầu π [8].

Trong ứng dụng phát hiện dáng đi bệnh lý, đối tượng cần được mô hình hóa bởi HMM chính là dáng đi của con người. Dáng đi có đặc điểm là lặp đi lặp lại theo chu kỳ với số lần lặp tùy ý và sự chuyển động của cơ thể bên trong mỗi chu kỳ không hoàn toàn giống hệt nhau. Để mô hình hóa loại tín hiệu như thế này, hệ thống đề xuất sử dụng mô hình HMM tuần hoàn CHMM (Cyclic HMM). Đây là loại mô hình HMM trái-phải có một chuyển tiếp từ trạng thái cuối cùng về trạng thái đầu tiên [13]. Bằng thực nghiệm đã chọn được số trạng thái ẩn là $N = 7$ và số ký hiệu quan sát là $M = 64$ như trên Hình 7.



Hình 7. Mô hình CHMM dùng trong hệ thống

4. Thực nghiệm kiểm tra đánh giá hệ thống đề xuất

Phần này trình bày quá trình thí nghiệm và các kết quả thí nghiệm đối với hai bài toán đã nêu.

4.1. Đánh giá hệ thống nhận dạng cây trong ảnh tĩnh

Để đánh giá hệ thống, trước hết ta xây dựng cơ sở dữ liệu ảnh hoa hồng và hoa cúc. Các mẫu hoa được sưu tầm từ internet, mỗi bức ảnh có trọn vẹn một bông hoa, với nhiều loài hoa khác nhau trong cùng họ, góc chụp khác nhau, kích thước, vị trí và màu sắc hoa khác nhau nhằm đảm bảo tính đa dạng cho cơ sở dữ liệu. Tổng cộng có 120 bức ảnh hoa gồm 60 ảnh hoa hồng và 60 ảnh hoa cúc được thu thập trong cơ sở dữ liệu. Các ảnh được cắt thủ công sao cho có cùng kích thước là 100x100.

4.1.1. Quá trình thí nghiệm nhận dạng hoa

Quá trình thí nghiệm gồm hai pha là huấn luyện và kiểm tra, được tiến hành theo các bước sau:

- **Bước 1:** trích bông hoa trong ảnh theo các bước như trình bày ở mục 3.1.1.
- **Bước 2:** chuyển mỗi bức ảnh nhị phân thành một vec-tơ 7 chiều, theo các công thức từ (1) đến (4).
- **Bước 3:** tạo dữ liệu chuẩn bị cho huấn luyện và kiểm

tra, bằng cách lập 2 ma trận: ma trận thứ nhất kích thước 120x7 để chứa 120 vec-tơ đặc trưng của 120 bức ảnh, với 7 mô-men Hu trong mỗi vec-tơ, ma trận thứ hai có kích thước 120x1 biểu diễn 120 đầu ra mong muốn, mỗi đầu ra có giá trị là 0 (ứng với hoa cúc) hoặc 1 (ứng với hoa hồng).

- **Bước 4:** thiết lập mạng ANN MLP như Hình 6.
- **Bước 5:** chia 120 bức ảnh ra làm 3 phần: 84 bức ảnh (70%) dành cho huấn luyện, 18 bức ảnh (15%) cho thẩm định (validation) và 18 bức ảnh còn lại (15%) cho kiểm tra. Thông thường thì càng kéo dài thời gian huấn luyện thì lỗi huấn luyện càng thấp, tuy nhiên nếu huấn luyện quá lâu thì mô hình sẽ rơi vào tình trạng “học thuộc mẫu” (overtraining) và sẽ cho lỗi rất lớn đối với các mẫu kiểm tra. Để tránh tình trạng này, ta dành 15% ảnh cho thẩm định và tiến hành huấn luyện và thẩm định cùng lúc và dừng ngay quá trình huấn luyện khi thấy lỗi thẩm định bắt đầu đảo chiều đi lên.

4.1.2. Kết quả thí nghiệm nhận dạng hoa

Để đánh giá hệ thống, tất cả các mẫu hoa có trong dữ liệu đều được lần lượt kiểm tra. Ở đây có tổng cộng 10 đợt huấn luyện - kiểm tra, với một đợt huấn luyện - kiểm tra được thực hiện trên một tập mẫu kiểm tra khác nhau và được thực hiện nhiều lần (5 lần). Thực tế cho thấy, đôi khi có sự chênh lệch khá lớn về tỷ lệ nhận dạng đúng giữa các lần kiểm tra trong một đợt, ví dụ ở lần kiểm tra đầu trong đợt thứ 6 thì tỷ lệ nhận dạng đúng là 83,3% nhưng ở lần kiểm tra sau đó thì tỷ lệ đó chỉ còn có 77,8%. Theo phân tích thì có 2 nguyên nhân chính gây ra hiện tượng này là do thay đổi tập huấn luyện và thẩm định và do lỗi MSE (Mean Square Error) huấn luyện bị phụ thuộc vào giá trị trọng số khởi đầu.

Kết quả cuối cùng để đánh giá hệ thống là trung bình cộng của các kết quả tốt nhất của 10 đợt kiểm tra nói trên và được biểu diễn dưới dạng ma trận nhầm lẫn (confusion matrix) như trên Bảng 2. Từ đây ta tính được tỷ lệ nhận dạng đúng trung bình là 87,88%. Đây là một kết quả chấp nhận được nếu so với chi phí thời gian (mỗi lần huấn luyện chỉ mất chưa đầy 0,5s).

Bảng 2. Ma trận nhầm lẫn (%) của hệ thống nhận dạng cây

Kết quả nhận dạng	Mẫu hoa	
	Hoa cúc	Hoa hồng
Hoa cúc	87,95	12,19
Hoa hồng	12,05	87,81

4.2. Đánh giá hệ thống phát hiện dáng đi bệnh lý trong đoạn video

Để đánh giá hệ thống, trước tiên ta xây dựng cơ sở dữ liệu dáng đi gồm dáng đi bộ bình thường và dáng đi bệnh Parkinson. Môi trường quay là trong nhà, dọc theo hành lang vào ban ngày. Camera thuộc loại 2D thông dụng và được đặt ở vị trí sao cho có thể ghi hình được trọn vẹn cơ thể người đi bộ với góc quay ngang. Tất cả 10 tình nguyện viên tham gia đều được huấn luyện kỹ và được tham khảo tài liệu y khoa [11] trước khi thực hiện ghi hình, nhằm đảm bảo dáng đi bệnh Parkinson được mô phỏng giống thực nhất có thể. Mỗi tình nguyện viên thực hiện một kiểu dáng đi 10 lần với các

tốc độ khác nhau. Như vậy tổng cộng có 200 đoạn video trong cơ sở dữ liệu, tất cả được định dạng .avi, kích thước khung hình 180x144, tốc độ 25 khung hình/giây.

4.2.1. Quá trình thí nghiệm phát hiện dáng đi bệnh lý

Tương tự như ở mục 4.1.1, quá trình thí nghiệm gồm hai pha là huấn luyện và kiểm tra, được tiến hành theo các bước sau:

- **Bước 1:** trích đối tượng là người đi bộ theo các bước đã trình bày ở 3.1.2.
- **Bước 2:** chuyển chuỗi khung hình nhị phân thành chuỗi vec-tơ đặc trưng mô-men Hu 7 chiều tính theo các công thức từ (1) đến (4).
- **Bước 3:** chia các vec-tơ đặc trưng ra thành tập huấn luyện trích từ các đoạn video huấn luyện và tập kiểm tra trích từ các đoạn video kiểm tra.
- **Bước 4:** rời rạc hóa chuỗi vec-tơ huấn luyện bằng cách lượng tử hóa vec-tơ dựa vào phương pháp K-means, với số nhóm $K = 64$ (hay cũng chính là số ký hiệu quan sát M), rồi huấn luyện mô hình CHMM ứng với dáng đi bình thường và CHMM ứng với dáng đi bệnh Parkinson. Kết quả lượng tử hóa vec-tơ còn là một bảng mã có 64 từ mã là trọng tâm của 64 nhóm.
- **Bước 5:** rời rạc hóa chuỗi vec-tơ kiểm tra bằng cách tính khoảng cách của từng vec-tơ đến tất cả các từ mã rồi gán cho mỗi vec-tơ bằng một số nguyên là chỉ số của từ mã gần nó nhất. Sau đó tính xác suất mà mỗi mô hình CHMM có thể tạo ra chuỗi vec-tơ kiểm tra, so sánh với nhau rồi quyết định chuỗi vec-tơ kiểm tra sẽ ứng với mô hình nào cho xác suất cao nhất.

4.2.2. Kết quả thí nghiệm phát hiện dáng đi bệnh lý

Để đánh giá hệ thống, tất cả các đoạn video có trong dữ liệu đều được lần lượt kiểm tra. Ở đây thực hiện 10 đợt huấn luyện - kiểm tra với mỗi đợt là để kiểm tra 10 đoạn video ghi hình từ cùng một người.

Kết quả cuối cùng để đánh giá hệ thống được tổng hợp từ tất cả 10 đợt kiểm tra nói trên và được biểu diễn dưới dạng ma trận nhầm lẫn như trên Bảng 3. Từ đây ta tính được tỷ lệ phát hiện đúng dáng đi bệnh Parkinson là 99%, tỷ lệ bỏ sót bệnh là 1%, tỷ lệ phát hiện nhầm là 0%, tỷ lệ nhận dạng trung bình (còn gọi là độ chính xác) là 99,5%. Đây là một tỷ lệ khả thành công trong bối cảnh kích thước tập dữ liệu huấn luyện không quá lớn.

Bảng 3. Ma trận nhầm lẫn của hệ thống phát hiện dáng đi bệnh lý

Kết quả nhận dạng	Mẫu dáng đi	
	Bệnh Parkinson	Bình thường
Bệnh Parkinson	99	0
Bình thường	1	100

5. Kết luận

Trong bài báo này, chúng tôi đã xây dựng hệ thống nhận

dạng mẫu hình ảnh với hai ứng dụng là nhận dạng hoa trong ảnh tĩnh và phát hiện dáng đi bệnh lý trong đoạn video. Các giá trị của mô-men Hu đã được sử dụng để miêu tả đặc trưng hình dạng của các đối tượng và cho kết quả rất khả quan. Cụ thể là hai mẫu hoa hồng và hoa cúc đã được nhận dạng đúng với tỷ lệ 87,88% và dáng đi bệnh lý Parkinson đã được phát hiện với độ chính xác 99,5%. Kết quả này là nhờ sự đóng góp của tất cả các khâu xử lý trong hệ thống, từ trích đối tượng, trích đặc trưng cho đến nhận dạng, đặc biệt là mô-men Hu đã mô tả tốt đối tượng trong ảnh nhị phân với khả năng bất biến đối với sự co giãn, vị trí đối tượng và sự xoay ảnh.

Lời cảm ơn

Các tác giả chân thành cảm ơn các em sinh viên chương trình tiên tiến AP-ECE, Trung tâm Xuất sắc và sinh viên khoa Điện tử - Viễn thông, Trường Đại học Bách khoa, Đại học Đà Nẵng đã nhiệt tình tham gia xây dựng cơ sở dữ liệu dáng đi bệnh lý.

TÀI LIỆU THAM KHẢO

- [1] Shian-Ru Ke, Hoang Le Uyen Thuc, Yong-Jin Lee, Jenq-Neng Hwang, Jang-Hee Yoo, and Kyoung-Ho Choi, "A Review on Video-Based Human Activity Recognition", *MDPI Computers*, vol. 2(2), 2013, pp. 88-131.
- [2] Rafael C. Gonzalez and Richard E. Woods, "Digital Image Processing", 3rd edition, Prentice-Hall, USA, 2006.
- [3] C. Stauffer and W. E. L. Grimson, "Adaptive Background Mixture Models for Real-time Tracking", *IEEE ICCV 1999*, Jun 1999, pp. 246-252.
- [4] Zhihu Huang and Jinsong Leng, "Analysis of Hu's Moment Invariants on Image Scaling and Rotation", *2nd Int. Conf. on Computer Engineering and Technology (ICCET)*, 2010, pp. 476-480.
- [5] Aaron F. Bobick and James W. Davis, "The Recognition of Human Movement Using Temporal Templates", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 23(3), 2001, pp. 257-267.
- [6] Abdunnaser Diaf and Rachid Benlamri, "An Effective View-based Motion Representation for Human Motion Recognition", *Int. Symposium on Modeling and Implementation of Complex Systems*, 2010.
- [7] Yu Hen Hu and Jenq-Neng Hwang, "Handbook of Neural Network Signal Processing", *CRC Press*, 2002.
- [8] Lawrence R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proc. IEEE*, vol. 77(2), 1989, pp. 257-286.
- [9] James S. Cope, David Corney, Jonathan Y. Clark, Paolo Remagnino, and Paul Wilkin, "Plant species identification using digital morphometrics: A review", *Expert Systems with Applications*, vol. 39, 2012, pp. 7562-7573.
- [10] E. Auvinet, F. Multon and J. Meunier, "New lower-limb gait asymmetry indices based on a depth camera", *Mdpi Sensors*, vol. 15(3), pp. 4506-4623, 2015.
- [11] Joseph H. Friedman, "Gait Disorders in the Elderly", *Medicine & Health*, vol. 95(3), 2012, pp. 84-85.
- [12] T. Graepel, "Statistical Physics of Clustering Algorithms", Technical Report 171822, FB Physic, Institute for Theoretical Physics, 1998.
- [13] Hoang Le Uyen Thuc, Shian-Ru Ke, Jenq-Neng Hwang, Pham Van Tuan, and Truong Ngoc Chau, "Quasi-periodic Action Recognition from Monocular Videos via 3D Human Models and Cyclic HMMs", *Int. Conf. on ATC*, 2012, pp. 110-113.