Xây dựng hệ thống nhận dạng hành động sử dụng cảm biến gia tốc và các phương pháp học máy trên vi điều khiển hiệu năng thấp

Đào Tô Hiệu[1,2], Ngô Văn Công[2,3], Nguyễn Quang Huy[2], Trần Đức Nghĩa[2], Trần Đức Tân[4,*]

[1] Trường Đại học Công nghệ thông tin và Truyền thông, Đại học Thái Nguyên, Thái Nguyên

[2] Học Viện Khoa học và Công nghệ (VAST), Hà Nội

[3] Tổng cục Tiêu chuẩn Đo lường Chất lượng, Hà Nội

[4] Trường Đại học Phenikaa, Hà Nội

Tác giả liên hệ: Trần Đức Tân, tan.tranduc@phenikaa-uni.edu.vn

Ngày nhận bài: 19/08/2021, ngày sửa chữa: 15/11/2021, ngày duyệt đăng: 21/11/2021

Định danh DOI: 10.32913/mic-ict-research-vn.v2021.n2.1004

Tóm tắt: Hệ thống ghi nhận hoạt động của con người sử dụng cảm biến giúp hỗ trợ đánh giá tình trạng sức khỏe. Yêu cầu đặt ra là hiệu suất hoạt động cao, giá cả rẻ và hoạt động theo thời gian thực ngay trên các vi điều khiển có hiệu năng thấp. Từ dữ liệu thu thập của các hành động cần trích xuất tính chất đặc trưng điển hình. Những đặc trưng này được huấn luyện để phân loại hành động. Một thiết bị gắn trên eo người được lập trình kết nối hệ thống thông qua WiFi và cho phép phân loại hành động người theo thời gian thực, đồng thời người dùng có thể giám sát hoạt động và quản lý dữ liệu trên máy tính hoặc điện thoại thông minh.

Từ khóa: *Phân loại, hành động, học máy, web server, cảm biến gia tốc, ứng dụng, thu thập dữ liệu.*

Title: **Building Human Activity Recognition System using Accelerometers and Machine Learning Methods on Low-Performance Microcontrollers**

Abstract: The system recognizes human activities using sensors to help assess health status. The requirements are high performance, low price and real-time operation even on low-performance microcontrollers. From the collected data of the activities, it is necessary to extract the typical feature. These features are trained to classify activities. A programmed waist-mounted device connects to the system via WiFi and allows real-time classification of human activities, and users can monitor activities and manage data on smart devices.

Keywords: *Classification, activity, machine learning, web server, acceleration sensor, application, data collection.*

I. GIỚI THIỆU

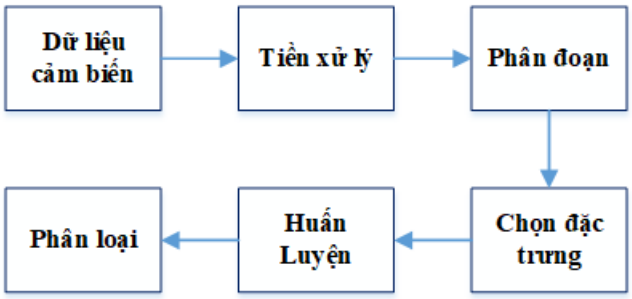
Trong vài thập kỷ qua, những nghiên cứu về hành động của con người phát triển rất mạnh mẽ và có nhiều ứng dụng vào cuộc sống. Những nhu cầu ứng dụng dựa trên phân loại hành động như chăm sóc sức khỏe [1], [2], phát hiện tai nạn cho người già [3], [4], công nghiệp [5], an ninh [6], ... Mô hình phân loại hành động cần những thông tin về hành vi của người và các hoạt động tương tác, từ đó cho phép xây dựng hệ thống hỗ trợ người dùng một cách chủ động trong công việc [7].

Với cách tiếp cận dựa trên cảm biến mang trên người đã mở ra nhiều ứng dụng tiềm năng trong phân loại hành động [8-13]. Cách tiếp cận này không cần thiết lập trước các điều kiện và không phụ thuộc vào không gian, thời gian sử dụng. Do đó, nhóm nghiên cứu hướng đến ứng dụng kỹ thuật phân loại hành vi người để hướng đến xây dựng một hệ thống chăm sóc khỏe cho người già, người bệnh mới hồi phục sau tai nạn, ... thông qua hệ thống theo dõi hành vi được gắn trực tiếp lên người dùng. Bởi vậy, nghiên cứu này sẽ tập trung vào việc khai thác khả năng phân loại theo thời gian thực các hành động cơ bản của con người giúp cho các bác sĩ có thể giám sát hoạt động của người bệnh và từ đó có cơ sở dữ liệu để chẩn đoán, hỗ trợ chăm sóc sức khỏe từ xa. Bên cạnh đó, nhóm nghiên cứu đề xuất sử dụng phân đoạn dữ liệu theo thời gian (cửa sổ trượt) có chiều rộng cố định n-giây. Các đặc trưng sẽ được trích xuất trên mỗi cửa sổ trượt (Sliding Window) và dùng làm dữ liệu đầu vào cho quá trình huấn luyện, phân loại hành động. Yêu cầu đặt ra với hệ thống là hiệu suất hoạt động

cao, giá rẻ và hoạt động theo thời gian thực ngay trên các vi điều khiển có hiệu năng thấp. Để thực hiện được mục tiêu như vậy, thiết bị cần được thiết kế gọn nhẹ, tất cả các bước gồm: 1/ thu thập dữ liệu, 2/ tiền xử lý, 3/ trích chọn đặc trưng, 4/ phân loại trực tiếp trên vi điều khiển, 5/ truyền không dây tới server phải thật gọn nhẹ và hiệu quả để có thể đáp ứng yêu cầu thời gian thực ngay trên các dòng vi điều khiển hiệu năng thấp (MCU ESP8266).

II. PHƯƠNG PHÁP NGHIÊN CỨU

Thiết bị giám sát được gắn trực tiếp trên eo người và thu thập dữ liệu gia tốc theo ba trục Ax, Ay và Az. Dữ liệu được phân đoạn theo phương pháp cửa sổ trượt. Sau đó, mỗi vectơ đặc trưng được trích xuất từ mỗi cửa sổ dữ liệu và được huấn luyện bởi thuật toán học máy là cây quyết định [14] để nâng cao hiệu suất phân loại. Để xây dựng mô hình phân loại hành động, bài báo nghiên cứu 2 đặc trưng của dữ liệu gia tốc: giá trị trung bình (mean) và độ lệch chuẩn (SD). Cấu trúc mô hình phân loại được thể hiện trong Hình 1.



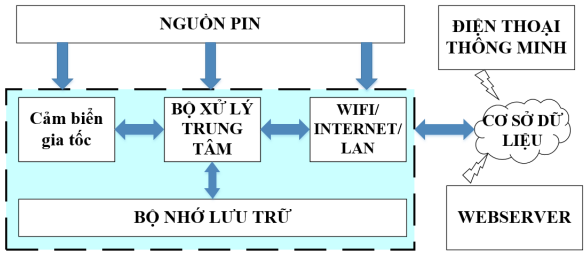
Hình 1. Cấu trúc mô hình phân loại hành động

1. Hệ thống thu thập dữ liệu

Cảm biến gia tốc được kết nối với vi điều khiển ESP8266 thông qua giao tiếp I2C để thu nhận dữ liệu gia tốc theo ba trục Ax, Ay, Az. Dữ liệu này được ghi log trên thẻ nhớ (bộ nhớ ngoài), đồng thời được truyền về Server thông qua WIFI (INTERNET/LAN) theo giao thức HTTP (Hypertext Transfer Protocol) và phương thức GET. Trong thiết kế này, thiết bị thu thập sử dụng pin lithium 3.7V-850mAh. Cấu trúc hệ thống được minh họa trong Hình 2.

2. Dữ liệu

Thiết bị được đeo trên người như trong Hình 3. Sau khi thu thập, dữ liệu này được tiền xử lý để loại bỏ các tín hiệu bị mất trong quá trình truyền, loại bỏ nhiễu và được gán nhãn theo năm lớp, bao gồm các hành động: nằm, ngồi, đứng, đi bộ và chạy bộ (Bảng I). Tần số lấy mẫu dữ liệu cảm biến được sử dụng là 1Hz.



Hình 2. Sơ đồ khối hệ thống phân loại hành vi

Bảng I
ĐỊNH NGHĨA CÁC HÀNH ĐỘNG

HÀNH ĐỘNG	ĐỊNH NGHĨA
Nằm	Toàn thân người trên mặt nền trong tư thế nghỉ.
Ngồi	Đầu và lưng không chạm nền và đôi chân nghỉ.
Đứng	Người thẳng, đôi chân không di chuyển .
Đi bộ	Người thẳng và di chuyển chậm đến vị trí khác nhau, có ít nhất 1 chân chạm đất.
Chạy bộ	Người thẳng và di chuyển nhanh đến vị trí khác, 2 chân không chạm đất khi chạy tại 1 thời điểm.



Hình 3. Vị trí đặt thiết bị giám sát

Gia tốc trên từng trục được tính theo:

$$A_i = \frac{\left(\frac{Sam_i}{1024} * R - O_i\right)}{S_i} \quad (1)$$

A_i: Giá trị gia tốc theo trục i (X, Y, Z);

Sam_i: Giá trị sau khi lấy mẫu trục i;

R: Tham chiếu điện áp;

O_i: Phần bù;

S_i: Độ nhạy của cảm biến gia tốc trên trục i.

Dữ liệu đo được từ các hành động được phân đoạn để tăng hiệu suất nhận dạng hành động. Trong nghiên cứu này, nhóm nghiên cứu sử dụng phương pháp cửa sổ trượt có kích thước cố định theo thời gian là 6 giây. Nếu kích thước cửa sổ quá ngắn, tỷ lệ trùng lặp trong dữ liệu thu thập tăng lên. Tuy nhiên nếu chiều rộng cửa sổ thời gian quá dài thì sẽ tăng xác suất có nhiều hành động xuất hiện trong một cửa sổ trượt [13].

Bảng II
SỐ LƯỢNG CỬA SỔ DỮ LIỆU

HÀNH ĐỘNG	CỬA SỔ TRƯỢT
Nằm	540
Ngồi	572
Đứng	700
Đi bộ	728
Chạy bộ	505
Tổng	3045

Do vậy, bộ dữ liệu thu thập được phân đoạn thành 3045 cửa sổ (Bảng II) với kích thước mỗi cửa sổ là 6 giây (s). Mẫu bản ghi đoạn dữ liệu cho mỗi cửa sổ có độ dài 6s được thể hiện trong Bảng III. Giá trị gia tốc sử dụng đơn vị là g với ($g = 9.8m/s^2$).

Bảng III
SỐ QUAN SÁT HÀNH ĐỘNG THU THẬP TỪ EO

X(g)	Y(g)	Z(g)	THỜI GIAN
0,14	0,03	1,03	2021-08-03 09:44:30
0,14	0,04	1,02	2021-08-03 09:44:31
0,15	0,03	1,03	2021-08-03 09:44:32
0,14	0,03	1,03	2021-08-03 09:44:33
0,14	0,04	1,02	2021-08-03 09:44:34
0,14	0,04	1,03	2021-08-03 09:44:35

3. Trích chọn đặc trưng

Để phân loại hành động đạt hiệu suất cao, việc lựa chọn đặc trưng từ dữ liệu các hoạt động rất quan trọng. Các đặc trưng được sử dụng trong nghiên cứu này: Mean và SD, Các đặc trưng này được trích xuất theo công thức (2), (3) với đầu vào là giá trị cảm biến thu thập tương ứng trên từng trục X, Y và Z với đơn vị là g.

$$Mean \qquad \mu(X_j) = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{2}$$

$$SD \qquad \sigma(X_j) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{3}$$

Trong đó: X_j là bản ghi thứ j và x_i là mẫu thứ i trong bản ghi X_j; N là số bản ghi; $\mu(X_j)$ là giá trị Mean của X_j.

Dữ liệu tổng hợp tại Bảng IV cho thấy đặc trưng Mean của 3 hành động đứng, đi bộ và chạy bộ có xu hướng dao động quanh một giá trị trung tâm theo trục X là -0,93g. Bên cạnh đó, hành động nằm và ngồi (trạng thái nghỉ) có giá trị đặc trưng Mean tập trung trên trục Y, Z khác biệt so với đứng, đi bộ chạy bộ (trạng thái vận động). Ví dụ: hành động nằm và ngồi có đặc trưng Mean trên trục Y nhỏ nhất là 0,37g trong khi với hành động đứng, đi bộ và chạy bộ có đặc trưng Mean trên trục Y lớn nhất là 0,085g. Tiếp theo, giữa hành động nằm và ngôi có giá trị các trục khác biệt trông thấy. Ví dụ: nằm có đặc trưng Mean trên trục Z là 0,501g còn đối với ngồi là -0,105g. Do đó, Mean có thể phân biệt người đó ở trạng thái nằm hay ngồi và phân biệt với những trạng thái còn lại.

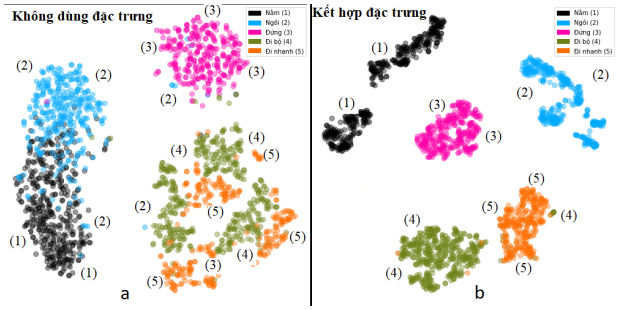
Bảng IV
CÁC ĐẶC TRƯNG DỮ LIỆU GIA TỐC THEO TRỤC X, Y VÀ Z

Đặc trưng	Trục	nằm	Ngồi	Đứng	Đi bộ	Chạy bộ
Mean	X(g)	-0,064	-0,874	-0,934	-0,928	-0,932
	Y(g)	0,823	0,37	0,085	0,072	0,004
	Z(g)	0,501	-0,105	-0,024	0,03	-0,032
SD	X(g)	0,018	0,011	0,0085	0,184	0,488
	Y(g)	0,022	0,023	0,025	0,206	0,42
	Z(g)	0,021	0,045	0,03	0,111	0,209

Khác với đặc trưng Mean, trạng thái động (đi bộ, chạy bộ) có giá trị đặc trưng SD khác biệt rõ ràng so với trạng thái tĩnh (đứng, ngồi, nằm). Ví dụ: giá trị đặc trưng của người khi đang di chuyển $SD_{min} = 0,111g$ trong khi ở trạng thái không di chuyển thì đặc trưng $SD_{max} = 0,045g$. Bên cạnh đó, đặc trưng SD có chênh lệch rõ rệt giữa hai hành động đi bộ và chạy bộ. Ví dụ: Theo trục X, đi bộ có $SD = 0,184g$ so với chạy bộ có $SD = 0,488g$. Do đó, đặc trưng có thể dùng để phân biệt một người ở trạng thái tĩnh hay động và người đó đi bộ hay chạy bộ.

4. Phân loại hành động

Để giảm số lượng kích thước dữ liệu giúp cho việc quan sát dữ liệu trực quan hơn, nhóm nghiên cứu sử dụng phương pháp t-SNE (t-Distributed Stochastic Neighbor Embedding) [15]. Thuật toán này sẽ ánh xạ từng đặc trưng đã được trích xuất trong không gian 3 chiều (X, Y, Z) sang không gian 2 chiều mà không sai lệch tính chất đặc trưng. Nếu không dùng đặc trưng, các hành động trùng lặp gần như không thể phân biệt (Hình 4).

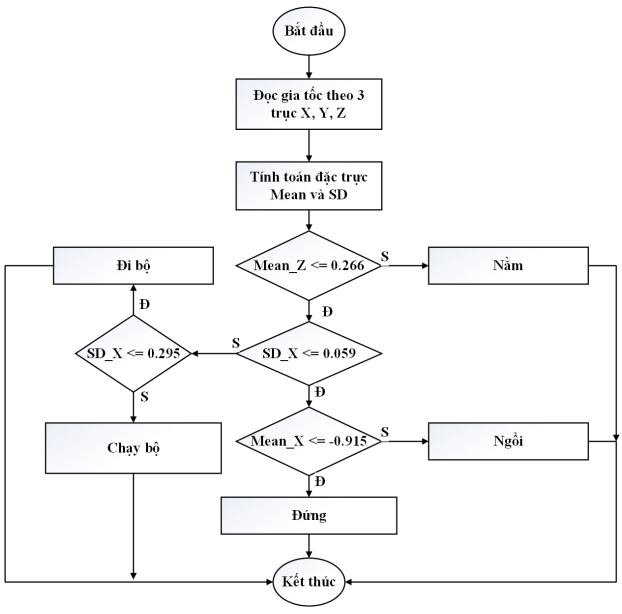


Hình 4. Ánh xạ bộ dữ liệu trước và sau khi sử dụng đặc trưng

Với việc kết hợp hai đặc trưng SD và Mean, việc phân loại hành động đã cho đạt kết quả tốt. Trong Hình 4 - a, sự trùng lặp nhỏ xuất hiện giữa hành động đi bộ (khu vực màu xanh lá cây) và chạy bộ (khu vực màu hồng). Tuy nhiên ranh giới giữa các hành động được xác định khá rõ rệt so với khi chưa dùng đặc trưng.

5. Thuật toán phân loại

Đặc trưng Mean và SD sau khi được trích xuất từ tập dữ liệu sẽ được huấn luyện bởi thuật toán học máy cây quyết định và từ đó xây dựng thuật toán phân loại hành động (Hình 5).

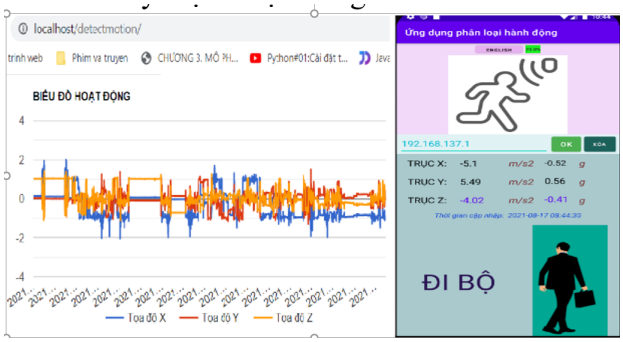


Hình 5. Thuật toán phân loại hành động người

Thiết bị sau khi được nhúng thuật toán sẽ thực hiện quá trình phân loại hành động dựa theo dữ liệu gia tốc theo thời gian. Cứ sau mỗi 6 giây dữ liệu, thiết bị sẽ phân loại hành động.

6. WebServer và ứng dụng phân loại hành động

Việc lưu trữ dữ liệu trên bộ nhớ mở rộng của thiết bị gây khó khăn trong việc quan sát dữ liệu thu thập và truy cập dữ liệu. Do đó, thiết bị được kết nối đến hệ thống WebServer (hình 6), dữ liệu sẽ được lưu trữ trên Server và có thể quan sát trên Website hay điện thoại thông minh.



Hình 6. Giao diện WebServer và phần mềm giám sát

Giao diện của ứng dụng trên điện thoại thông minh được xây dựng giúp người sử dụng dễ dàng quan sát tham số thu thập từ cảm biến theo đơn vị gia tốc g hoặc m/s^2 và thời gian cập nhập dữ liệu từ cảm biến. Bên cạnh đó, chức năng hiển thị dung lượng pin giúp người dùng biết khi nào thiết bị cần sạc pin. Người dùng có thể nhấn vào nút chức năng English/Tiếng việt để chuyển đổi giao diện ngôn ngữ hiển thị.

III. KẾT QUẢ VÀ ĐÁNH GIÁ

1. Phương pháp đánh giá

Hiệu năng của mô hình sử dụng được đánh giá theo ba thông số sau: độ chính xác (Accuracy), độ nhạy (Sensitivity) và độ chỉ rõ (Specify). Các giá trị trên được xác định dựa trên ma trận nhầm lẫn (Confusion Matrix) [16] cho từng hành động:

$$acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

$$sen = \frac{TP}{TP + FN} \tag{5}$$

$$spe = \frac{TN}{TN + FP} \tag{6}$$

Trong đó, TP (True Positive): Số lượng hành động cần phân loại được phân loại đúng so với quan sát. FN (False Negative): Số lượng hành động cần phân loại bị phân loại thành các hành động khác. FP (False Positive): Số lượng độ hành động khác bị phân loại thành hành động cần phân loại. TN (True Negative): Số lượng hành động khác được phân loại đúng so với quan sát thực tế.

2. Đánh giá mô hình phân loại

Trong nghiên cứu này, chúng tôi đánh giá mô hình đề xuất với những mô hình phân loại như: Gradient Boosted Decision Tree (GBDT) [17], Support Vector Machine [18], Random Forest [19], k-Nearest Neighbor [20].

Bảng V
SO SÁNH KHI ÁP DỤNG MỘT SỐ THUẬT TOÁN THÔNG DỤNG

Đặc trưng	**DT**	**GBDT**	**SVM**	**RF**	**KNN**
Mean	85%	92,7%	91%	93,4%	92%
SD	80,3%	84,6%	88%	84%	83%
Kết hợp	99,3%	99,6%	100%	99,6%	100%

Đặc trưng dữ liệu được huấn luyện và đánh giá trong Bảng V. Hiệu suất cao nhất khi sử dụng mô hình phân loại GBDT với đặc trưng Mean. Độ chính xác thấp nhất của các mô hình phân loại là 85% với mô hình cây quyết định (DT) và đặc trưng SD. Tuy nhiên, sự kết hợp giữa 2 đặc trưng SD và Mean đã cho kết quả tăng đến 99.3%.

Bảng VI
MA TRẬN NHẦM LẪN TRÊN DỮ LIỆU THU THẬP

Hành Động	Hành động dự đoán				
	Nằm	Ngồi	Đứng	Đi bộ	Chạy bộ
Nằm	278	0	0	0	0
Ngồi	0	228	0	0	0
Đứng	0	0	215	1	0
Đi bộ	0	0	1	286	4
Chạy bộ	0	0	0	4	198

Kết quả được thể hiện chi tiết của mô hình cây quyết định được trình bày trong Bảng VI. Thống kê số trường hợp hành động được phân loại chính xác và số trường hợp nhầm lẫn với các hành động khác cho thấy các mẫu thử nghiệm về hành động được phân loại tốt. Tuy có 4/202 mẫu chạy bộ, 5/291 mẫu đi bộ và 1/216 mẫu đứng phân loại sai nhưng tỷ lệ sai trong khoảng 0,5% - 2% là chấp nhận được.

Bảng VII thống kê độ chính xác (acc), độ nhạy (sen), độ chỉ rõ (spe) khi phân loại hành động nằm và ngồi đạt 100%. Những hành động còn lại có độ chính xác, độ chỉ rõ và độ nhạy lớn hơn 98%.

Nhìn chung với bộ dữ liệu thu thập được, các chỉ số đạt được trên 98% với các hành động. Điều này cho thấy mô hình phân loại cây quyết định thực sự có hiệu quả rất cao do dữ liệu các hành động được thực hiện lấy mẫu là thuần nhất, ít có sự chồng chéo giữa các hành động được lấy mẫu và không có hành động khác xen vào gây nhiễu.

Bảng VII
CHỈ SỐ CHÍNH XÁC MÔ HÌNH CÂY QUYẾT ĐỊNH

Chỉ số	**Nằm**	**Ngồi**	**Đứng**	**Đi bộ**	**Chạy bộ**
acc	100%	100%	99,83%	99,18%	99,34%
sen	100%	100%	99,54%	98,28%	98,02%
spe	100%	100%	99,90%	99,46%	99,60%

3. Kết quả thực nghiệm

Quá trình thực nghiệm phân loại hành động trên thiết bị thực được tiến hành như sau: mỗi hành động được kiểm tra trong khoảng thời gian 5-10 phút, quá trình này diễn ra liên tục và có sự thay đổi hành động sát với thực tế trong quá trình thực nghiệm. Ví dụ chuyển từ ngồi sang đi bộ, chuyển từ đi bộ sang nằm.

Bảng VIII
MA TRẬN NHẦM LẪN KHI THỰC NGHIỆM

Hành Động	Hành động dự đoán				
	Nằm	Ngồi	Đứng	Đi bộ	Chạy bộ
nằm	65	0	0	0	0
Ngồi	3	60	5	0	0
Đứng	0	5	54	2	0
Đi bộ	0	3	7	60	3
Chạy bộ	0	0	0	9	55

Kết quả thực nghiệm phân loại các hành động trên thiết bị gắn tại eo người được tổng hợp trên ma trận chồng chéo (Bảng VIII). Hành động nằm được phân loại với độ chính xác rất cao đạt 96,71%, tương đối ít hành động bị phân loại nhầm với độ nhạy thấp nhất là 82,19%. Hiệu suất khi phân loại từng hành động trên thiết bị theo thời gian thực được mô tả tại Bảng IX. Hiệu suất tổng thể khi phân loại trên thiết bị có độ chính xác đạt trên 91%. Trong đó, độ chính xác rất cao đạt được 98,94% với hành động nằm, tiếp theo là hành động chạy bộ với 96,08%. Hành động đứng và ngồi có độ chính xác thấp hơn một chút nhưng cao hơn so với hành động đi bộ.

Bảng IX
HIỆU SUẤT MÔ HÌNH KHI THỰC NGHIỆM

Chỉ số	**Nằm**	**Ngồi**	**Đứng**	**Đi bộ**	**Chạy bộ**
acc	96,71%	94,23%	94,53%	91,59%	96,08%
sen	100%	85,71%	84,38%	82,19%	85,94%
spe	95,82%	96,69%	97,17%	94,35%	98,76%

IV. THẢO LUẬN

Thực nghiệm phân loại theo thời gian thực có độ chính xác giảm xuống do dữ liệu gia tốc thu thập không thuần

nhất. Trong một phân đoạn dữ liệu theo thời gian có thể bị lẫn vài giây của hành động trước đó hoặc kế tiếp gây khó khăn trong việc phân loại. Ví dụ: thời gian diễn ra hành động đứng có thể tồn tại thời gian chuyển từ hành động đi bộ sang hành động đứng và ngược lại nhưng điều này không bao giờ xảy ra với bộ dữ liệu được huấn luyện trên máy tính do đã được nhóm theo các hành động. Bên cạnh đó nhiễu tích lũy có thể xảy ra khi chuyển trạng thái tĩnh sang động với tốc độ nhanh. So sánh kết quả nghiên cứu của chúng tôi với kết quả nghiên cứu của Nishkam Ravi [8] và Ling Bao [9] được trình bày trong Bảng X.

Bảng X
SO SÁNH VỚI CÁC MÔ HÌNH PHÂN LOẠI NAVI VÀ BAO

Hành động	**L. Bao**	**N. Ravi**	**Nghiên cứu này**
Nằm	95%		99,00%
Ngồi	94,8%	96,6%	94,89%
Đứng	95,7%	91,8%	93,99%
Đi bộ	89,7%	99,5%	92,52%
Chạy bộ	87,7%	84,8%	96,12%
Tổng thể	92,58%	94,54%	95,3%

Bên cạnh đó, nhiều nghiên cứu lựa chọn khá nhiều đặc điểm để nâng cao tỷ lệ phân loại hoạt động. Ví dụ: Nghiên cứu sử dụng 43 đặc trưng của C. Y. Min [10], Catal [12], Van der Maaten [15]; nghiên cứu sử dụng đặc trưng của Vavoulas [11]. Trong tài liệu [12], Catal đã cách kết hợp rất nhiều đặc trưng và cho kết quả là 91,6%. Mặc dù họ đã phân loại nhiều hoạt động hơn (sáu hoạt động) so với nghiên cứu của chúng tôi, nhưng họ đã thử nghiệm một thuật toán phức tạp hơn khi kết hợp nhiều thuật toán phân loại khác nhau. Bên cạnh đó, việc sử dụng 43 đặc trưng khác nhau khiến giá thành và chi phí thiết bị tăng cao hơn rất nhiều và có thể khó áp dụng đại trà. Tuy nhiên với kết quả thực nghiệm thu được mô hình phân loại có khả năng ứng dụng tốt bởi nhiễu luôn xảy ra trong thực tế.

Để kiểm chứng hiệu quả của mô hình đề xuất, nhóm nghiên cứu áp dụng mô hình phân loại trên tập dữ liệu WISDM [23], [24] được thu thập từ 36 người tình nguyện viên với các hoạt động thông thường gồm đi bộ (đi bộ bình thường, lên/xuống cầu thang), chạy bộ, ngồi và đứng. Bộ dữ liệu công khai này được lấy mẫu với tần số 20Hz. Tỷ lệ dữ liệu đào tạo và dữ liệu kiểm tra là 6/4. Số mẫu trên một phân đoạn dữ liệu là 120 mẫu tương ứng với tần số lấy mẫu 20Hz (Bảng XI)

Kết quả phân loại hành động trên phần dữ liệu kiểm tra được tổng hợp trong Bảng XII và XIII. Kết quả này có sự tương đồng với kết quả thực nghiệm (Bảng VIII) về sự nhầm lẫn giữa các hành vi với nhau, phù hợp với phân tích phía trên của chúng tôi. Độ chính xác tổng thể khi phân loại trên tập dữ liệu WISDM (98.4%) giảm xuống so với khi

Bảng XI
KẾT QUẢ PHÂN LOẠI TRÊN TẬP DỮ LIỆU WISDM

Hành động	Hành động dự đoán			
	Ngồi	Đứng	Đi bộ	Chạy bộ
Ngồi	169	9	4	0
Đứng	1	142	5	0
Đi bộ	0	1	2029	47
Chạy bộ	0	0	47	1039

phân loại trên tập dữ liệu thu thập (99,3%) nhưng cao hơn khi tiến hành thực nghiệm (95,26%). Cụ thể, hành động ngồi có 9 mẫu bị phân loại thành hành động đứng và 9 mẫu bị phân loại thành đi bộ; hành động đi bộ có số lượng mẫu nhiều nhất (2076 mẫu) nhưng chỉ có 47 mẫu bị phân loại thành chạy bộ. Bên cạnh đó tỷ lệ hành động ở trạng thái di chuyển bị phân loại thành hành động ở trạng thái di chuyển là nhỏ với 1 mẫu hành động đi bộ bị phân loại thành đứng.

Bảng XII
KẾT QUẢ PHÂN LOẠI TẬP DỮ LIỆU WISDM

Hành động	**Số lượng mẫu**	**Cửa sổ dữ liệu**	**Dữ liệu đào tạo**	**Dữ liệu kiểm tra**
Ngồi	59939	457	275	182
Đứng	48395	371	223	148
Đi bộ	647696	5193	3116	2007
Chạy bộ	342177	2716	1630	1086
Tổng	1098207	8737	5244	3493

Các chỉ số đạt được khi áp dụng mô hình đề xuất trên tập dữ liệu WISDM là khá tốt trên 92%. Độ nhạy thấp nhất với hành động ngồi (92,86%) và cao nhất với hành động đi bộ (97,69%). Bên cạnh đó, hành động đứng và chạy bộ có độ nhạy tương đương với 96%. Độ chỉ rõ trên 98% với các hành động ngồi, đứng và chạy bộ trong khi hành động đi bộ có độ chỉ rõ đạt 96%.

Bảng XIII
HIỆU SUẤT PHÂN LOẠI TRÊN TẬP WISDM

Chỉ số	**Ngồi**	**Đứng**	**Đi bộ**	**Chạy bộ**	**Tổng thể**
acc	99,59%	99,53%	97,01%	97,29%	98,36%
sen	92,86%	95,95%	97,69%	95,67%	95,54%
spe	99,97%	99,69%	96,02%	98,03%	98,43%

Có thể thấy rằng, việc giám sát hành vi người có thể chia thành hai giải pháp là dùng 1/ camera và 2/ cảm biến gắn người. Mỗi giải pháp này đều có những ưu và nhược điểm khác nhau. Việc lựa chọn giải pháp sẽ tùy thuộc vào mục đích sử dụng. Những nghiên cứu hiện nay về xử lý ảnh

đã đạt được nhiều thành tựu rất lớn khi có thể phân loại rất nhiều hành vi với độ chính xác mà một trong những bộ dữ liệu mà các nhóm nghiên cứu thường sử dụng đó là HMDB51 dataset gồm 6766 video clips và 51 lớp hành động. Cụ thể, trong một nghiên cứu mới năm 2021 [25], độ chính xác trung bình đạt được là 95,4% cho 51 lớp hành động. Giải pháp dùng camera sẽ thể hiện nhiều bất lợi về môi trường hoạt động (ví dụ: bị gián đoạn, che khuất khi người cần giám sát di chuyển từ phòng tới khu vệ sinh, hành lang không trang bị camera,..), độ phức tạp cao khi định danh nếu giám sát nhiều người cùng lúc (điều không gặp phải nếu dùng thiết bị gắn người), sự riêng tư của người sử dụng... Ngoài ra, đây sẽ còn là là một hệ thống phức tạp và khó thành công về mặt thương mại, giá cao hơn nhiều so với giá của thiết bị sử dụng vi điều khiển và cảm biến gia tốc giá thành rẻ. Nghiên cứu của chúng tôi hướng đến chăm sóc sức khỏe cho người già, người hồi phục sau phẫu thuật hay tai nạn. Do đó, số lớp hành vi liên quan tới vận động cơ bản là đủ cho loại ứng dụng này nhưng lại cần phân loại được theo thời gian thực và giá thành rẻ. Do đó, giải pháp dùng xử lý ảnh chưa phù hợp theo mục tiêu của ứng dụng này.

Mô hình của chúng tôi để có thể được áp dụng cho những đối tượng thực hiện công việc đặc biệt nguy hiểm như lính cứu hỏa thì sẽ phát triển thêm để nhận một số dạng hành động khác như: bò, đi lên, đi xuống, trạng thái sinh tồn, định vị [21,22]. Bởi lẽ, hoạt động của các đối tượng này thường diễn ra với cường độ cao hơn trong điều kiện phức tạp như cháy lớn, khói dày đặc.

V. KẾT LUẬN

Trong nghiên cứu này, nhóm nghiên cứu đã xây dựng một hệ thống phân loại hành động với chi phí thấp, thời gian tính toán nhỏ thời gian thực. Kết quả thử nghiệm với cửa sổ trượt có kích thước 6 giây cho thấy khả năng phân loại theo mô hình cây quyết định có độ chính xác cao hơn GBT, SVM, KNN. Đối với nghiên cứu sâu hơn, chúng tôi tiếp tục phát triển và cải thiện những kết quả này thông qua các thí nghiệm trên các hoạt động phức tạp hơn để kiểm tra độ chính xác của mô hình đề xuất. Bên cạnh đó, ngoài các đặc trưng thống kê cửa sổ thời gian, nghiên cứu có thể mở rộng các đặc trưng khác trong miền tần số hoặc kết hợp thông tin trên cả hai miền thời gian và tần số trong quá trình trích chọn đặc trưng. Hiện nay, xu hướng thực thi học máy trên các bộ vi điều khiển hiệu năng thấp (TinyML [26]) đang dần trở nên thịnh hành trên thế giới và nghiên cứu này cũng góp phần thực hiện TinyML cho các ứng dụng giám sát hành vi phục vụ chăm sóc sức khỏe từ xa.

LỜI CẢM ƠN

Trần Đức Nghĩa được hỗ trợ bởi Quỹ Đổi mới sáng tạo Vingroup (VinIF), Viện Nghiên cứu Dữ liệu lớn (VINBIGDATA) trong khuôn khổ chương trình học bổng Sau Tiến sĩ trong nước (mã số học bổng VINIF.2021.STS.21).

TÀI LIỆU THAM KHẢO

[1] Pham, Minh, Yehenew Mengistu, Ha Manh Do, and Weihua Sheng. "Cloud-based smart home environment (CoSHE) for home healthcare.", *IEEE International Conference on Automation Science and Engineering*, pp. 483-488, 2016.

[2] Peetoom, Kirsten KB, et al. "Literature review on monitoring technologies and their outcomes in independently living elderly people.", *Disability and Rehabilitation: Assistive Technology 10.4*, 271-294, 2015.

[3] Tan, Tran Duc, and Nguyen Van Tinh. "Reliable fall detection system using an 3-DOF accelerometer and cascade posture recognitions.", *Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pp. 1-6, 2014.

[4] Van Thanh, Pham, Duc-Tan Tran, Dinh-Chinh Nguyen, Nguyen Duc Anh, Dang Nhu Dinh, S. El-Rabaie, and Kumbesan Sandrasegaran. "Development of a real-time, simple and high-accuracy fall detection system for elderly using 3-DOF accelerometers.", *Arabian Journal for Science and Engineering*, vol 44, no. 4, pp. 3329-3342, 2019

[5] Stiefmeier, Thomas, Daniel Roggen, Georg Ogris, Paul Lukowicz, and Gerhard Tröster. "Wearable activity tracking in car manufacturing.",*IEEE Pervasive Computing*, vol 8 , no. 2, pp.42-50, 2018.

[6] Taha, Ahmed, Hala H. Zayed, M. E. Khalifa, and El-Sayed M. El-Horbaty. "Human activity recognition for surveillance applications." In Proceedings of the 7th International Conference on Information Technology, pp. 577-586. 2015.

[7] Putra, I. P. E. S., James Brusey, Elena Gaura, and Rein Vesilo. "An event-triggered machine learning approach for accelerometer-based fall detection.", *Sensors*, vol 18, no. 1, 2018.

[8] Ravi, Nishkam, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman. "Activity recognition from accelerometer data.", *Aaai*, vol. 5, pp. 1541-1546, 2005.

[9] Bao, Ling, and Stephen S. Intille. "Activity recognition from user-annotated acceleration data.", *International conference on pervasive computing*, pp. 1-17, 2004.

[10] Min, Ya, Yin Yin Htay, and Khin Khin Oo. "Comparing the performance of machine learning algorithms for human activities recognition using wisdm dataset.", *International Journal of Computer (IJC)*, vol 38, no. 1, pp. 61-72, 2020.

[11] Vavoulas, George, Charikleia Chatzaki, Thodoris Malliotakis, Matthew Pediaditis, and Manolis Tsiknakis. "The mobiact dataset: Recognition of activities of daily living using smartphones.", *International Conference on Information and Communication Technologies for Ageing Well and e-Health*, vol. 2, pp. 143-151. 2016.

[12] Catal, Cagatay, Selin Tufekci, Elif Pirmit, and Guner Kocabag. "On the use of ensemble of classifiers for accelerometer-based activity recognition.", *Applied Soft Computing*, vol 37, pp. 1018-1022, 2015.

[13] Pham, Van Thanh, Duc Anh Nguyen, Nhu Dinh Dang, Hong Hai Pham, Van An Tran, Kumbesan Sandrasegaran, and Duc-Tan Tran. "Highly accurate step counting at various walking states using low-cost inertial measurement unit support indoor positioning system.", *Sensors*, vol 18, no. 10, 2018.

[14] Kamiński, Bogumił, Michał Jakubczyk, and Przemysław Szufel. "A framework for sensitivity analysis of decision

trees.", *Central European journal of operations research*, vol 26, no. 1, pp. 135-159, 2018.
[15] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE.", *Journal of machine learning research*, vol 9, no. 11, 2008.
[16] Wang, Jun, Zhitao He, Guoqiang Zheng, Song Gao, and Kaixuan Zhao. "Development and validation of an ensemble classifier for real-time recognition of cow behavior patterns from accelerometer data and location data.", *PloS one*, vol 13, no. 9, 2018.
[17] Alcolea, Adrián, and Javier Resano. "FPGA Accelerator for Gradient Boosting Decision Trees.", *Electronics*, vol 10, no. 3, 2021.
[18] Cortes, Corinna, and Vladimir Vapnik. "Support-vector networks." Machine learning 20, no. 3 (1995): 273-297.
[19] Shiri, Jalal. "Improving the performance of the mass transfer-based reference evapotranspiration estimation approaches through a coupled wavelet-random forest methodology.", *Journal of Hydrology 561*, pp. 737-750, 2018.
[20] Hattori, Kazuo, and Masahito Takahashi. "A new edited k-nearest neighbor rule in the pattern classification problem.", *Pattern Recognition*, vol 33, no. 3, pp. 521-528, 2000.
[21] Pham, Van Thanh, Quang Bon Le, Duc Anh Nguyen, Nhu Dinh Dang, Huu Tue Huynh, and Duc Tan Tran. "Multi-sensor data fusion in a real-time support system for on-duty firefighters.", *Sensors*, vol 19, no. 21, 2019.
[22] Tan, Tran Duc, L. M. Ha, N. T. Long, N. D. Duc, and N. P. Thuy. "Integration of inertial navigation system and global positioning system: Performance analysis and measurements.", *International Conference on Intelligent and Advanced Systems*, pp. 1047-1050, 2007.
[23] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, “Activity recognition using cell phone accelerometers,”, *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.
[24] Weiss, Gary Mitchell, and Jeffrey Lockhart. "The impact of personalization on smartphone-based activity recognition.", *Workshops at the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.
[25] Varshney, Neeraj, and Brijesh Bakariya. "Deep convolutional neural model for human activities recognition in a sequence of video by combining multiple CNN streams.", *Multimedia Tools and Applications*, pp. 1-13, 2021.
[26] Dutta, Lachit, and Swapna Bharali. "TinyML Meets IoT: A Comprehensive Survey.", *Internet of Things*, 2021.

SƠ LƯỢC VỀ TÁC GIẢ

Đào Tô Hiệu



Tốt nghiệp thạc sĩ chuyên ngành Kỹ thuật điện tử tại Đại học Bách Khoa Hà Nội. Đang làm nghiên cứu sinh tại Viện Hàn lâm Khoa học và Công nghệ Việt Nam (VAST).
Hiện là giảng viên trường Đại học Công nghệ Thông tin và Truyền thông, Đại học Thái Nguyên.
Email: dthieu@ictu.edu.vn

Ngô Văn Công



Tốt nghiệp trường Đại học kỹ thuật Lê Quý Đôn năm 2012, nhận bằng thạc sĩ năm 2015.
Hiện là nghiên cứu sinh năm nhất tại Học viện Khoa học và Công nghệ.
Lĩnh vực nghiên cứu: Điều khiển và tự động hóa; Tiêu chuẩn - Đo lường - Chất lượng.
Email: ngovancong@tcvn.gov.vn

Nguyễn Quang Huy



Sinh năm 1990, tốt nghiệp thạc sĩ ngành Công nghệ thông tin tại Học viện Công nghệ Bưu chính Viễn thông.
Hiện công tác tại Viện Công nghệ Thông tin (IOIT), Viện Hàn lâm Khoa học và Công nghệ Việt Nam (VAST).
Lĩnh vực nghiên cứu: Phân tích dữ liệu, mạng cảm biến không dây và các ứng dụng.
Email: nqhuy@ioit.ac.vn

Trần Đức Nghĩa



Nhận bằng tiến sĩ tại Université Paris Descartes (Paris, Pháp) năm 2018.
Lĩnh vực nghiên cứu: Phổ EPR, Internet vạn vật, ứng dụng học máy, cảm biến gia tốc.
Email: nghiatd@ioit.ac.vn

Trần Đức Tân



Sinh năm 1980 tại Bắc Ninh. Được phong PGS năm 2013.
Hiện là Phó Trưởng khoa Điện – Điện tử, Trường Đại học PHENIKAA.
Lĩnh vực nghiên cứu: Kỹ thuật điện tử - xử lý tín hiệu.
Email: tan.tranduc@phenikaa-uni.edu.vn