BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC VÀ
CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



VŨ THỊ THƯƠNG

NGHIÊN CỨU TRIỂN KHAI HỌC MÁY TRÊN VI ĐIỀU KHIỂN HIỆU NĂNG THẤP CHO HỆ THỐNG NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN GIA TỐC

LUẬN ÁN TIẾN SĨ MÁY TÍNH

Hà Nội - 2026

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC VÀ
CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Vũ Thị Thương

NGHIÊN CỨU TRIỂN KHAI HỌC MÁY TRÊN VI ĐIỀU KHIỂN HIỆU NĂNG THẤP CHO HỆ THỐNG NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN GIA TỐC

LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Hệ thống thông tin

Mã số: 9 48 01 04

Xác nhận của Học viện Khoa học và Công nghệ	Người hướng dẫn 1 (*Ký, ghi rõ họ tên*)	Người hướng dẫn 2 (*Ký, ghi rõ họ tên*)
KT. GIÁM ĐỐC PHÓ GIÁM ĐỐC HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ Nguyễn Thị Trung	GS.TS. Trần Đức Tân	TS. Nguyễn Như Sơn

Hà Nội - 2026

LỜI CAM ĐOAN

Tôi xin cam đoan luận án: "Nghiên cứu triển khai học máy trên vi điều khiển hiệu năng thấp cho hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc" là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Luận án sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được công bố chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 23 tháng 10 năm 2026

Tác giả luận án



Vũ Thị Thương

LỜI CẢM ƠN

Quá trình học tập, nghiên cứu và hoàn thành luận án tiến sĩ này đem lại cho tác giả các trải nghiệm thú vị và nhiều thách thức trong cuộc sống. Trong suốt quá trình này tác giả đã nhận được sự quan tâm, giúp đỡ và động viên của rất nhiều tập thể và cá nhân. Tác giả xin gửi lời cảm ơn chân thành và sâu sắc đến tất cả những người đã đồng hành cùng tác giả trong suốt hành trình này.

Luận án tiến sĩ được thực hiện tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam (Viện HLKH&CN Việt Nam), dưới sự hướng dẫn khoa học tận tình của GS. TS. Trần Đức Tân và TS. Nguyễn Như Sơn. Tác giả xin bày tỏ lòng biết ơn chân thành và sự kính trọng sâu sắc đối với các thầy trong Tập thể hướng dẫn khoa học, những người không chỉ truyền đạt nhiều kiến thức quý báu, kinh nghiệm nghiên cứu khoa học mà còn khuyến khích, động viên tác giả vượt qua những khó khăn trong chuyên môn và cuộc sống. Sự chuyên nghiệp, nghiêm túc trong nghiên cứu và những định hướng đúng đắn của các thầy là tiền đề quan trọng giúp tác giả có được những kết quả trình bày trong luận án này.

Tác giả cũng xin gửi lời cảm ơn đến tập thể thầy cô, cán bộ của Viện Công nghệ Thông tin, Ban Lãnh đạo, Phòng Đào tạo cùng các đơn vị chức năng của Học viện Khoa học và Công nghệ đã tạo điều kiện học tập, nghiên cứu thuận lợi và môi trường học thuật cởi mở, giúp tác giả có thêm cơ hội để phát triển và hoàn thành đề tài.

Tiếp theo, tác giả xin chân thành cảm ơn thầy giáo GS. TS. Trần Đức Tân và cộng sự trong SSA Lab Khoa Điện- Điện tử, Đại học Phenikaa đã tạo điều kiện thuận lợi cho tác giả trong quá trình học tập và hoàn thành luận án. Tác giả cũng xin gửi lời tri ân đến các thầy cô giáo đã giảng dạy và hướng dẫn tác giả trong các học phần của chương trình tiến sĩ.

Tác giả xin tỏ lòng biết ơn và kính trọng đối với hội đồng đánh giá luận án cùng các thầy cô giáo tham gia quá trình phản biện đã dành thời gian và công sức để đọc và nhận xét luận án. Các ý kiến, nhận xét và đánh giá của quý thầy cô đã giúp tác giả hoàn thiện luận án và nâng cao chất lượng công trình nghiên cứu.

Tác giả sẽ không thể hoàn thành luận án này mà không nhận được sự giúp đỡ của cơ quan chủ quản. Tác giả biết ơn tập thể Lãnh đạo trường Đại học Phương Đông, lãnh đạo và các đồng nghiệp tại Khoa Công nghệ thông tin và Truyền thông, tạo điều kiện thuận lợi để tác giả học tập, nghiên cứu và hoàn thành luận án này.

Đặc biệt tác giả xin được bày tỏ lòng biết ơn sâu sắc tới Bố, Mẹ và các thành viên trong gia đình, những người đã luôn động viên và truyền cảm hứng cho tác giả trong suốt quá trình nghiên cứu.

Hà Nội, ngày 23 tháng 10 năm 2026

Tác giả luận án

Vũ Thị Thương

MỤC LỤC

DANH MỤC CHỮ VIẾT TẮT ix
MỞ ĐẦU 1
CHƯƠNG 1: TỔNG QUAN VỀ HỆ THỐNG NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN 9
1.1. Giới thiệu về nhận dạng hoạt động người 9
1.2. Tổng quan về hệ thống nhận dạng hoạt động sử dụng cảm biến 12
1.2.1. Hệ thống nhận dạng hoạt động sử dụng cảm biến 12
1.2.2. Các nghiên cứu điển hình nhận dạng hoạt động dựa trên cảm biến mang trên người 13
1.2.3. Khoảng trống nghiên cứu và định hướng của luận án 18
1.3. Hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc 19
1.3.1. Nhận dạng hoạt động sử dụng cảm biến gia tốc 19
1.3.2. Mô hình hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc 21
1.3.3. Phương pháp đánh giá mô hình 24
1.3.4. Phân tích lựa chọn thuật toán cho vi điều khiển 8-bit 25
1.3.5. TinyML và so sánh với các hệ thống embedded HAR hiện đại 27
1.4. Kết luận chương 1 28
CHƯƠNG 2: NHẬN DẠNG HOẠT ĐỘNG CƠ BẢN BẰNG RANDOM FOREST TRÊN VI ĐIỀU KHIỂN 8-BIT (RF4-HAR) 29
2.1. Mô hình hệ thống RF4-HAR 29
2.1.1. Kiến trúc mô hình phân loại hoạt động 29
2.1.2. Tập dữ liệu 32
2.1.3. Quy trình xử lý dữ liệu 36
2.2. Thực nghiệm và kết quả 45
2.2.1. Kết quả thực nghiệm trên các thuật toán 45
2.2.2. Triển khai mô hình RF trên vi điều khiển 47

2.2.3. Đánh giá hiệu suất của mô hình phân loại RF 50
2.2.4. Đánh giá hệ thống RF4-HAR..... 53
2.3. Kết luận chương 2 57
CHƯƠNG 3: PHƯƠNG PHÁP NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN GIA TỐC Ở ĐỊA HÌNH CÓ THAY ĐỔI ĐỘ CAO..... 59
3.1. Phương pháp cải thiện hệ thống theo dõi hoạt động ở tòa nhà - RF5-HAR . 59
3.1.1. Mô tả phương pháp 59
3.1.2. Tập dữ liệu và thuật toán cải tiến 60
3.1.3. Quy trình xử lý dữ liệu 62
3.2. Thực nghiệm và kết quả của hệ thống RF5-HAR..... 72
3.2.1. Kết quả thực nghiệm trên các thuật toán..... 72
3.2.2. Triển khai thuật toán RF trên vi điều khiển 75
3.2.3. Đánh giá mô hình phân loại RF 77
3.3. Đánh giá hệ thống RF5-HAR..... 78
3.3.1. So sánh với các nghiên cứu trước 78
3.3.2. Đánh giá tính ứng dụng và triển khai hệ thống RF5-HAR 80
3.4. Hệ thống KNN6-HAR nhận dạng hoạt động phân biệt trạng thái lên, xuống cầu thang..... 82
3.4.1. Kiến trúc mô hình phân loại hoạt động..... 82
3.4.2. Tập dữ liệu 83
3.4.3. Quy trình xử lý dữ liệu..... 87
3.5. Thực nghiệm và kết quả của phương pháp nhận dạng hoạt động phân biệt trạng thái lên, xuống cầu thang..... 96
3.5.1. Kết quả thực nghiệm trên các thuật toán..... 96
3.5.2. Triển khai mô hình KNN trên vi điều khiển 98
3.5.3. Đánh giá hiệu suất của mô hình phân loại KNN..... 103

3.6. Đánh giá hệ thống KNN6-HAR nhận dạng hoạt động có phân biệt lên, xuống cầu thang ở địa hình tòa nhà cao tầng 105

3.6.1. So sánh với các nghiên cứu trước 105

3.6.2. So sánh với hệ thống RF5-HAR 107

3.7. Kết luận chương 3 110

KẾT LUẬN VÀ KIẾN NGHỊ 112

DANH MỤC CÁC CÔNG TRÌNH CỦA TÁC GIẢ 118

TẠP CHÍ KHOA HỌC 118

HỘI NGHỊ KHOA HỌC 118

SÁNG CHẾ 118

DANH MỤC CHỮ VIẾT TẮT

Từ viết tắt	Tiếng Anh	Tiếng Việt
AAD	Average Absolute Deviation	Độ lệch tuyệt đối trung bình
Acc	accuracy	Độ chính xác
ADC	Analog to Digital Converter	Bộ chuyển đổi tương tự số
ADXL	Analog Devices iMEMS Accelerometer	Cảm biến gia tốc ADXL345
AI	Artificial intelligence	Trí tuệ nhân tạo
DT	Decision Tree	Cây quyết định
GBDT	Gradient Boosted Decision Tree	Cây quyết định tăng cường gradient
GPRS	General Packet Radio Service	Dịch vụ vô tuyến gói chung
GSM	Global System for Mobile Communications	Hệ thống thông tin di động toàn cầu
HAR	Human Activity Recognition	Nhận dạng hoạt động người
HTTP	Hypertext Transfer Protocol	Giao thức truyền tải siêu văn bản
I2C	Inter-Integrated Circuit	Giao thức liên kết mạch tích hợp
IMU	Inertial Measurement Unit	Đơn vị đo lường quán tính
IoT	Internet of Things	Internet vạn vật
IQR	Interquartile Range	Khoảng tứ phân vị
KNN	K-Nearest Neighbor	K láng giềng gần nhất
LR	Logistic Regression	Hồi quy logistic
MCU	Microcontroller Unit	Khối vi điều khiển

NPV	Negative Predictive Value	Giá trị dự đoán âm
P2P	Peak-to-Peak	Đỉnh đến đỉnh
PPV	Positive Predictive Value	Giá trị dự đoán dương
RF	Random Forest	Rừng ngẫu nhiên
SAV	Sum of Absolute Values	Tổng giá trị tuyệt đối
SD	Standard Deviation	Độ lệch chuẩn
Sen	sensitivity	Độ nhạy
Spe	specificity	Độ đặc hiệu
SVM	Support Vector Machine	Máy véc-tơ hỗ trợ
TinyML	Tiny Machine Learning	Học máy cho thiết bị nhúng siêu nhỏ
t-SNE	t-Distributed Stochastic Neighbor Embedding	Thuật toán giảm chiều dữ liệu t-SNE
UART	Universal Asynchronous Receiver/Transmitter	Bộ thu phát không đồng bộ
WiFi	Wireless Fidelity	Kết nối không dây
WISDM	Wireless Sensor Data Mining	Tập dữ liệu khai thác dữ liệu cảm biến không dây
ADL	Activities of Daily Living	Hoạt động sinh hoạt hằng ngày
AUC	Area Under the Curve	Diện tích dưới đường cong ROC
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
DOF	Degrees of Freedom	Bậc tự do
FPU	Floating Point Unit	Bộ xử lý dấu phẩy động
IIR	Infinite Impulse Response	Bộ lọc đáp ứng xung vô hạn
LOSOCV	Leave-One-Subject-Out Cross-Validation	Kiểm định chéo loại trừ từng đối tượng

LSTM	Long Short-Term Memory	Mạng bộ nhớ dài-ngắn hạn
RMS	Root Mean Square	Căn quân phương
TFLM	TensorFlow Lite Micro	Thư viện TensorFlow Lite cho vi điều khiển
TP/FP/TN/FN	True/False Positive, True/False Negative	Dương tính thật/giả, âm tính thật/giả

DANH MỤC HÌNH ẢNH

Hình 0.1 Lộ trình nghiên cứu của luận án 7
Hình 1.1 Phân loại phương pháp nhận dạng hoạt động người 10
Hình 1.2 Một số cảm biến mang trên người 15
Hình 1.3 Mô hình hệ thống phân loại hoạt động người 22
Hình 2.1 Sơ đồ khối của quá trình phân loại hoạt động 30
Hình 2.2 Mô hình thiết bị hệ thống đề xuất (a) và tình nguyện viên đeo thiết bị (b). 31
Hình 2.3 Sơ đồ khối của chương trình nhúng để phân loại thời gian thực. 35
Hình 2.4 Phân phối dữ liệu gia tốc theo trục X của các hoạt động 39
Hình 2.5 Biểu diễn dữ liệu công khai trong không gian 2D, (a) không có đặc trưng (b) có đặc trưng 43
Hình 2.6 Biểu diễn dữ liệu tự thu thập trong không gian 2D, (a) không có đặc trưng (b) có đặc trưng 44
Hình 2.7 Độ chính xác acc của dữ liệu công khai 45
Hình 2.8 Độ nhạy của dữ liệu công khai 45
Hình 2.9 Dữ liệu tự thu thập - Độ chính xác (acc) 46
Hình 2.10 Dữ liệu tự thu thập - Độ nhạy 46
Hình 2.11 Sơ đồ mô tả thuật toán Random Forest 47
Hình 3.1 Chương trình nhúng để phân loại hoạt động thời gian thực 62
Hình 3.2 Biểu đồ mô tả một phần dữ liệu dọc theo trục X 65
Hình 3.3 Biểu đồ mô tả một phần dữ liệu chuyển động của một người dọc theo trục X. 66
Hình 3.4 Tập dữ liệu WISDM được biểu diễn trong không gian 2 chiều: (a) không có các đặc trưng và (b) có các đặc trưng 71
Hình 3.5 Tập dữ liệu thu thập được trong không gian 2D không có đặc trưng (a) và có đặc trưng (b). 71
Hình 3.6 Ma trận nhầm lẫn của thuật toán RF trên tập dữ liệu WISDM với kích thước cửa sổ 8 giây. 76
Hình 3.7 Ma trận nhầm lẫn của thuật toán RF trên tập dữ liệu thu thập được với kích thước cửa sổ 8 giây. 76

Hình 3.8 Phân loại được thực hiện bởi vi điều khiển thời gian thực và hiển thị trên điện thoại thông minh. 81
Hình 3.9 Các bước trong nhận dạng hoạt động..... 83
Hình 3.10 Biểu đồ mô tả một phần dữ liệu dọc theo trục X. 90
Hình 3.11 Dữ liệu công khai không có đặc trưng (a) và với đặc trưng (b) trong không gian 2D. 94
Hình 3.12 Dữ liệu thu thập không có đặc trưng (a) và với đặc trưng (b) trong không gian 2D. 95
Hình 3.13 Độ chính xác của DT, SVM, KNN và LR trên tập dữ liệu công khai. ... 96
Hình 3.14 Độ nhạy của DT, SVM, KNN và LR trên tập dữ liệu công khai. 96
Hình 3.15 Độ chính xác của DT, SVM, KNN và LR trên tập dữ liệu thu thập. 97
Hình 3.16 Độ nhạy của DT, SVM, KNN và LR trên tập dữ liệu thu thập..... 98
Hình 3.17 Sơ đồ mô tả thuật toán K-Nearest Neighbors (k=5) 99

DANH MỤC BẢNG BIỂU

Bảng 1.1 So sánh các phương pháp nhận dạng hoạt động 11
Bảng 1.2 Tổng quan về HAR dựa trên cảm biến mang trên người 20
Bảng 1.3 So sánh thuật toán theo ràng buộc phần cứng PIC18F4520 26
Bảng 1.4 So sánh framework-based TinyML và algorithm-based (luận án đề xuất) 27
Bảng 2.1 Diễn giải khái niệm các hoạt động 33
Bảng 2.2 Thành phần của Quan sát Hoạt động Dữ liệu Công khai 37
Bảng 2.3 Thành phần của Quan sát Hoạt động Dữ liệu tự thu thập 38
Bảng 2.4 Giá trị theo trục X (g) 40
Bảng 2.5 Giá trị theo trục Y (g) 40
Bảng 2.6 Giá trị theo trục Z (g) 40
Bảng 2.7 Các đặc trưng tính theo trục X 41
Bảng 2.8 Ma trận nhầm lẫn (Confusion matrix) bộ dữ liệu công khai 48
Bảng 2.9 Ma trận nhầm lẫn (Confusion matrix) bộ dữ liệu tự thu thập 48
Bảng 2.10 Hiệu suất của mô hình sử dụng các chỉ số trên tập dữ liệu công khai 49
Bảng 2.11 Hiệu suất của mô hình sử dụng các chỉ số trên tập dữ liệu tự thu thập 49
Bảng 2.12 Các phép đo để đánh giá hiệu suất của mô hình phân loại 50
Bảng 2.13 Kết quả thay đổi bộ đặc trưng 51
Bảng 3.1 Giải thích các hoạt động 61
Bảng 3.2 Dữ liệu hoạt động từ tập dữ liệu WISDM 63
Bảng 3.3 Dữ liệu hoạt động từ tập dữ liệu thu thập 64
Bảng 3.4 Công thức các đặc trưng của dữ liệu gia tốc trên trục X 67
Bảng 3.5 Dữ liệu gia tốc trục X (g) 68
Bảng 3.6 Dữ liệu gia tốc trục Y (g) 69
Bảng 3.7 Dữ liệu gia tốc trục Z (g) 69
Bảng 3.8 Độ chính xác của GBDT, SVM, RF và KNN trên tập dữ liệu WISDM 72
Bảng 3.9 Độ nhạy của GBDT, SVM, RF và KNN trên tập dữ liệu WISDM 73
Bảng 3.10 Độ chính xác của GBDT, SVM, RF và KNN trên tập dữ liệu thu thập 73
Bảng 3.11 Độ nhạy của GBDT, SVM, RF và KNN trên tập dữ liệu thu thập 74

Bảng 3.12 Thời gian thực thi và tiêu thụ năng lượng trung bình cho tập kiểm tra trên thiết bị đeo (trong tập dữ liệu thu thập) 74
Bảng 3.13 Hiệu suất mô hình trên tập dữ liệu WISDM với kích thước cửa sổ 8 giây 76
Bảng 3.14 Hiệu suất mô hình trên tập dữ liệu thu thập với kích thước cửa sổ 8 giây 77
Bảng 3.15 Đánh giá mô hình phân loại RF với kích thước cửa sổ 8 giây 78
Bảng 3.16 Mô tả các hoạt động..... 86
Bảng 3.17 Dữ liệu hoạt động từ tập dữ liệu công khai và thu thập..... 87
Bảng 3.18 Công thức đặc trưng cho dữ liệu gia tốc trên trục X 91
Bảng 3.19 Gia tốc theo trục X (g)..... 92
Bảng 3.20 Gia tốc theo trục Y (g)..... 93
Bảng 3.21 Gia tốc theo trục Z (g) 93
Bảng 3.22 Ma trận nhầm lẫn của thuật toán KNN trên tập dữ liệu công khai..... 100
Bảng 3.23 Ma trận nhầm lẫn của thuật toán KNN trên tập dữ liệu thu thập..... 101
Bảng 3.24 Hiệu suất mô hình trên tập dữ liệu công khai..... 102
Bảng 3.25 Hiệu suất mô hình trên tập dữ liệu thu thập..... 103
Bảng 3.26 Đánh giá mô hình phân loại KNN 104
Bảng 3.27 So sánh tổng hợp về hiệu quả của hệ thống RF5-HAR và KNN6-HAR 107
Bảng 3.28 So sánh hệ thống đề xuất với các nghiên cứu quốc tế tiêu biểu 108
Bảng 3.29 Tóm tắt ba hệ thống HAR đề xuất trong luận án..... 109

MỞ ĐẦU

Sự phát triển mạnh mẽ của công nghệ cùng nhu cầu chăm sóc sức khỏe ngày càng cao khiến việc theo dõi và phân tích hoạt động thể chất trở thành một bài toán cấp thiết. Lĩnh vực này đóng vai trò trọng tâm trong nhiều khía cạnh đời sống như y tế, thể thao, an ninh và hỗ trợ người cao tuổi. Tầm quan trọng của vấn đề càng được khẳng định qua khuyến cáo của Tổ chức Y tế Thế giới (WHO), khi việc duy trì hoạt động thể chất thường xuyên được chứng minh là giải pháp hữu hiệu nhằm giảm thiểu nguy cơ mắc bệnh tim mạch, đái tháo đường, ung thư và suy giảm sức khỏe tinh thần [1].

Hệ thống nhận dạng hoạt động người (Human Activity Recognition - HAR) đóng vai trò then chốt trong việc tự động giám sát và phân loại các hoạt động thường nhật, từ đó mang lại những giá trị thiết thực cho bức tranh toàn cảnh về chăm sóc sức khỏe cộng đồng. Các ứng dụng tiêu biểu của HAR trải rộng trên nhiều khía cạnh: từ giám sát mức độ vận động của bệnh nhân hậu phẫu [2], [3], đánh giá tiến trình phục hồi chức năng, phát hiện và cảnh báo sớm rủi ro té ngã ở người cao tuổi [4], [5], cho đến việc theo dõi và tối ưu hóa hiệu suất tập luyện [6]. Trong lĩnh vực công nghiệp, HAR được ứng dụng để theo dõi thao tác của công nhân trên dây chuyền sản xuất [7] và nhận dạng hành động của con người nhằm hỗ trợ tương tác người-robot an toàn trong dây chuyền sản xuất [8].

Có ba phương pháp chính để xây dựng hệ thống HAR: dựa trên thị giác máy tính (vision-based), dựa trên cảm biến môi trường (environment sensor-based) và dựa trên cảm biến đeo (wearable sensor-based) [9], [10]. Phương pháp sử dụng camera và thị giác máy tính, ví dụ như phương pháp trích xuất đặc trưng Local Binary Pattern (LBP) của Goodarzi và cộng sự [11], có ưu điểm là không xâm nhập nhưng gặp khó khăn về quyền riêng tư, chi phí triển khai cao, yêu cầu điều kiện ánh sáng tốt và bị hạn chế bởi góc nhìn của camera [12], [13]. Phương pháp sử dụng cảm biến môi trường như cảm biến hồng ngoại, cảm biến áp suất gắn trên sàn nhà có ưu điểm là không cần thiết bị đeo nhưng chỉ hoạt động trong môi trường được trang bị cảm biến, chi phí lắp đặt cao và khó mở rộng quy mô [6], [14]. Trong đó, phương pháp sử dụng cảm biến mang trên người ưa chuộng nhất do tính tiện lợi, chi phí thấp, không xâm phạm quyền riêng tư và có thể hoạt động trong mọi môi trường [15], [16].

Cảm biến gia tốc (accelerometer) là loại cảm biến phổ biến nhất được sử dụng trong các thiết bị đeo do khả năng thu thập dữ liệu chuyển động theo ba trục không gian với độ chính xác cao, kích thước nhỏ gọn, tiêu thụ năng lượng thấp và chi phí thấp [17]-[19]. Tuy nhiên, các nghiên cứu hiện tại vẫn còn nhiều hạn chế: Hầu hết các hệ thống sử dụng smartphone hoặc thiết bị có hiệu năng cao [17], [20], dẫn đến chi phí cao và tiêu thụ năng lượng lớn; nhiều nghiên cứu sử dụng số lượng đặc trưng lớn (50-100 đặc trưng) [18], gây khó khăn trong triển khai thời gian thực trên thiết bị nhúng; thiếu nghiên cứu về triển khai thực tế trên vi điều khiển hiệu năng thấp với tài nguyên hạn chế [21], [22]. Ở Việt Nam, mặc dù đã có một số nghiên cứu về HAR như công trình của Trần Văn Lăng (2014) về nhận dạng hành vi [23], Phạm Văn Cường (2017) về thiết bị iShoes [24], nhưng các giải pháp này vẫn còn hạn chế về chi phí, khả năng hoạt động thời gian thực và tính thực tiễn trong triển khai rộng rãi. Do đó, việc nghiên cứu và phát triển một hệ thống HAR hoàn chỉnh với chi phí thấp (<20 USD), độ chính xác cao (>95%), có khả năng hoạt động thời gian thực trên vi điều khiển 8-bit là một nhu cầu thực sự và có ý nghĩa khoa học cũng như thực tiễn cao.

Tóm lại, tính cấp thiết của luận án xuất phát từ ba khoảng trống cụ thể nêu trên: (i) các hệ thống HAR hiện có phần lớn phụ thuộc vào thiết bị hiệu năng cao (điện thoại thông minh, vi điều khiển 32-bit) nên khó áp dụng cho các thiết bị đeo giá rẻ; (ii) việc sử dụng bộ đặc trưng lớn (50-100 đặc trưng) không phù hợp với tài nguyên hạn chế của vi điều khiển 8-bit; (iii) còn thiếu các nghiên cứu, đặc biệt là các nghiên cứu trong nước, triển khai và đánh giá HAR thời gian thực trực tiếp trên phần cứng chi phí thấp. Do đó, việc nghiên cứu triển khai học máy trên vi điều khiển hiệu năng thấp cho hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc, vừa bảo đảm độ chính xác, vừa phù hợp với tài nguyên phần cứng hạn chế, là một yêu cầu cấp thiết, có ý nghĩa khoa học và thực tiễn rõ ràng.

Xuất phát từ những nhu cầu thực tiễn đó, luận án tập trung giải quyết một rào cản lớn trong lĩnh vực nhận dạng hoạt động người (HAR): sự đánh đổi giữa độ chính xác của hệ thống và giới hạn tài nguyên phần cứng. Trong khi nhiều nghiên cứu tiêu biểu như Kwapisz và cộng sự [17] (bộ dữ liệu WISDM) hay Vavoulas và cộng sự [20] (bộ dữ liệu MobiAct) đã đạt độ chính xác rất cao (97% - 99,79%), các hệ thống này lại phụ thuộc vào năng lực xử lý của điện thoại thông minh hoặc máy tính cá nhân cùng các bộ đặc

trưng phức tạp (từ 50 đến 64 đặc trưng). Hướng tiếp cận này khó áp dụng rộng rãi cho các thiết bị đeo nhỏ gọn, chi phí thấp. Nhằm vượt qua rào cản này, luận án đề xuất một hệ thống HAR thời gian thực phù hợp với tài nguyên phần cứng hạn chế, chỉ sử dụng dữ liệu từ một cảm biến gia tốc 3 trục (ADXL345) xử lý trực tiếp trên vi điều khiển 8-bit (PIC18F4520 có RAM 1,5 KB).

Mục tiêu của luận án

Mục tiêu chung của luận án: Nghiên cứu triển khai học máy trên vi điều khiển hiệu năng thấp cho hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc bảo đảm khả năng phân loại chính xác các hoạt động cơ bản trong thời gian thực.

Trong luận án, hai cách gọi "vi điều khiển hiệu năng thấp" và "vi điều khiển có tài nguyên bộ nhớ hạn chế" được dùng để chỉ cùng một nhóm đối tượng nghiên cứu: các vi điều khiển 8-bit có tốc độ xử lý, dung lượng RAM và Flash thấp (như PIC18F4520: 40 MHz, RAM 1,5 KB, Flash 32 KB), không đủ khả năng chạy các mô hình học sâu tiêu chuẩn. Từ đây, luận án thống nhất sử dụng thuật ngữ "vi điều khiển hiệu năng thấp" xuyên suốt để tránh nhầm lẫn.

Mục tiêu cụ thể: Đề xuất một số phương pháp nhằm cải thiện hiệu quả hệ thống nhận dạng hoạt động như sau:

- Thứ nhất, Nghiên cứu đặc điểm dữ liệu thu được từ cảm biến gia tốc 3 trục, từ đó lựa chọn bộ đặc trưng có khả năng phân biệt tốt giữa các hoạt động nhưng vẫn có độ phức tạp tính toán thấp, phù hợp với triển khai trên vi điều khiển hiệu năng thấp.
- Thứ hai, So sánh và đánh giá hiệu quả của một số thuật toán học máy có giám sát phổ biến, gồm Random Forest, SVM, KNN và GBDT, trong các bài toán nhận dạng hoạt động từ đơn giản đến mở rộng, nhằm lựa chọn thuật toán phù hợp với từng kịch bản ứng dụng.
- Thứ ba, Xây dựng hệ thống nhận dạng hoạt động thời gian thực sử dụng cảm biến gia tốc ADXL345 và vi điều khiển PIC18F4520, làm cơ sở kiểm chứng tính khả thi của mô hình trong điều kiện tài nguyên hạn chế.
- Thứ tư, Đánh giá toàn diện hệ thống đề xuất trên dữ liệu công khai và dữ liệu tự thu thập, đồng thời phân tích khả năng ứng dụng của hệ thống trong theo dõi sức khỏe, hỗ trợ người cao tuổi và giám sát hoạt động trong môi trường thực tế.

Đối tượng nghiên cứu của luận án:

Luận án tập trung nghiên cứu và tìm hiểu một số đối tượng liên quan đến việc nhận dạng hoạt động của người như:

- Phân tích hoạt động của người, từ đó nhận dạng hành vi đứng, ngồi, đi bộ, chạy bộ, lên và xuống cầu thang để giám sát hoạt động.
- Thuật toán học để nhận dạng hành vi người.
- Thiết bị cảm biến và vi điều khiển.
- Nhận dạng hành vi người trực tiếp trên thiết bị theo thời gian thực.

Phạm vi nghiên cứu của luận án

- Nghiên cứu tập trung vào các hoạt động cơ bản hàng ngày (đứng, ngồi, đi bộ, chạy bộ, lên xuống cầu thang).
- Sử dụng cảm biến gia tốc 3 trục ADXL345.
- Áp dụng các thuật toán học có giám sát.
- Triển khai trên vi điều khiển PIC18F4520.
- Thu thập dữ liệu từ 14 sinh viên (18-22 tuổi) và 2 người cao tuổi (65-72 tuổi).

Phương pháp nghiên cứu

Luận án sử dụng kết hợp các phương pháp:

- Nghiên cứu lý thuyết nhằm tổng hợp, phân tích và lựa chọn các công trình liên quan đến HAR, cảm biến gia tốc, trích xuất đặc trưng và các thuật toán học máy phù hợp với hệ thống nhúng.
- Nghiên cứu thực nghiệm nhằm thiết kế, chế tạo thiết bị, xây dựng quy trình thu thập dữ liệu và triển khai thử nghiệm hệ thống trong điều kiện thực tế.
- Phân tích và xử lý dữ liệu nhằm thực hiện phân đoạn tín hiệu, trích xuất đặc trưng và đánh giá khả năng phân biệt của từng đặc trưng đối với các hoạt động khác nhau.
- Mô hình hóa và đánh giá nhằm huấn luyện, kiểm thử, so sánh các mô hình học máy trên nhiều bộ dữ liệu và nhiều thước đo hiệu năng, đồng thời phân tích khả năng triển khai của mô hình trên vi điều khiển chi phí thấp.
- So sánh và đánh giá nhằm đo lường hiệu năng của hệ thống đề xuất (độ chính xác, thời gian suy luận, bộ nhớ, năng lượng) và so sánh với các nghiên

cứu liên quan trong và ngoài nước.

Các đóng góp của luận án

Những điểm mới của luận án:

- Chứng minh bằng thực nghiệm khả năng triển khai học máy (Random Forest, KNN) trực tiếp trên vi điều khiển 8-bit PIC18F4520 (RAM 1,5 KB) để nhận dạng hoạt động người theo thời gian thực, đạt độ chính xác tổng thể đa lớp trên 95% (tập WISDM) và trên 98% (tập tự thu thập) đối với RF4-HAR và RF5-HAR, chỉ dùng 4-5 đặc trưng thống kê miền thời gian trên mỗi trục

- Đề xuất và mở rộng hệ thống nhận dạng sang địa hình có thay đổi độ cao (lên/xuống cầu thang), với thuật toán KNN cho phép phân biệt rõ hai chiều di chuyển mà các nghiên cứu trước trên vi điều khiển hiệu năng thấp chưa thực hiện được.

Với mục tiêu giải quyết thách thức về nhận dạng hoạt động con người (HAR) theo thời gian thực sử dụng cảm biến gia tốc kế trên các vi điều khiển chi phí thấp, hiệu năng thấp. Nghiên cứu này phân biệt với các nghiên cứu khác bằng cách tập trung vào việc triển khai hiệu quả các thuật toán đơn giản về mặt tính toán cho phân loại hoạt động theo thời gian thực. Hệ thống được đề xuất nhằm hỗ trợ chăm sóc sức khỏe của bệnh nhân trong thời gian phục hồi, cho người cao tuổi hoặc giám sát trong an ninh, lao động. Các đóng góp chính của luận án này được nêu như sau:

- **Đóng góp 1:** Nghiên cứu này đề xuất và triển khai thành công hệ thống HAR thời gian thực trên vi điều khiển 8-bit PIC18F4520 (RAM 1,5 KB, 2 USD), một ràng buộc tài nguyên rất giới hạn mà các nghiên cứu tiêu biểu như Kwapisz và cộng sự [17] (smartphone), Vavoulas và cộng sự [20] (64 đặc trưng) hay Catal và cộng sự [25] (43 đặc trưng) đều chưa đặt ra. Đóng góp khoa học cốt lõi là chứng minh rằng chỉ cần 5 đặc trưng thống kê đơn giản (tính trên mỗi trục X, Y, Z, tương đương 15 chiều đầu vào) kết hợp với thuật toán phù hợp cho MCU 8-bit có thể đạt độ chính xác tổng thể trên 95% trên tập WISDM và trên 98% trên tập tự thu thập (RF4-HAR, RF5-HAR), tương đương các nghiên cứu sử dụng nhiều đặc trưng hơn trên nền tảng phần cứng mạnh hơn. Thuật toán (RF, KNN), được nhúng trực tiếp vào vi điều khiển, cho phép phân loại hoạt động theo thời gian thực.

- **Đóng góp 2: Lựa chọn hoạt động và bộ đặc trưng phù hợp với ràng buộc phần cứng.** Các hoạt động chính Đứng, Ngồi, Chạy bộ, Đi bộ và Đi lên, xuống cầu thang được lựa chọn cẩn thận dựa trên mức độ liên quan và tần suất trong cuộc sống

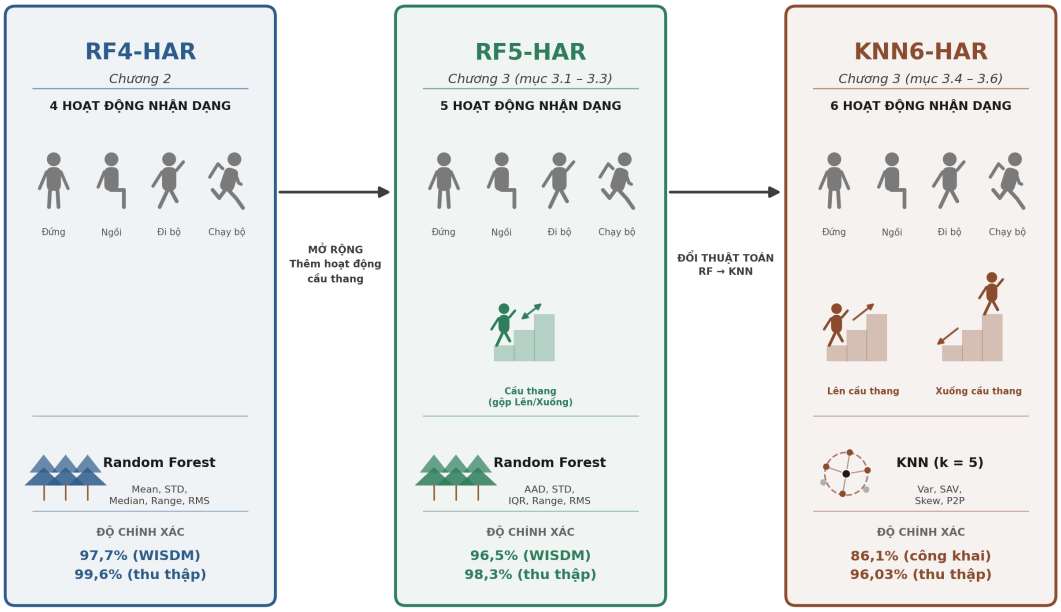
hàng ngày. Bộ đặc trưng được lựa chọn là cấu hình cân bằng tốt nhất giữa khả năng phân biệt và độ phức tạp tính toán trong các cấu hình đã khảo sát (Bảng 2.13), phù hợp MCU 8-bit.

Để làm rõ tính mới, luận án phân định như sau. Các nội dung kế thừa gồm: thuật toán Random Forest [26] và KNN; các đặc trưng thống kê miền thời gian (Mean, SD, RMS, Range, AAD, Var, Skew...) vốn đã được sử dụng trong các nghiên cứu HAR [17], [18], [25]; kỹ thuật phân đoạn bằng cửa sổ trượt có chồng lấp. Các nội dung do tác giả đề xuất và thực hiện gồm: (i) quy trình lựa chọn bộ đặc trưng tối giản theo ràng buộc phần cứng, kết hợp phân tích trực quan phân bố dữ liệu, thực nghiệm loại bỏ/bổ sung đặc trưng (Bảng 2.13) và chỉ số feature importance; (ii) quy trình lựa chọn thuật toán theo các ràng buộc định lượng về Flash, RAM và thời gian suy luận (Mục 1.3.4) và quy trình chuyển mô hình đã huấn luyện thành mã C dạng mảng tĩnh để nhúng vào PIC18F4520 (Mục 2.2.2); (iii) thiết kế thiết bị, xây dựng bộ dữ liệu 16 tình nguyện viên và ba hệ thống RF4-HAR, RF5-HAR, KNN6-HAR có quan hệ kế thừa. Ba hệ thống khác nhau không chỉ ở số hoạt động và địa hình mà còn ở bộ đặc trưng (Mean/SD/Median/Range/RMS; AAD/STD/IQR/Range/RMS; Var/SAV/Skew/P2P), thuật toán (RF cho bài toán ít lớp, KNN cho bài toán cần phân biệt hai lớp cầu thang gần nhau), kích thước cửa sổ và mô-đun truyền thông (SIM808/GSM ở Chương 2, ESP8266/WiFi ở Chương 3).

Luận án có hai nhóm kết quả chính như sau:

- Nhóm thứ nhất, đề xuất một phương pháp phân loại hoạt động người trên môi trường bằng phẳng cho phép phân loại hoạt động Đứng, Ngồi, Chạy bộ, Đi bộ. Phương pháp này đã được công bố tại công trình [CT1], [CT2], [CT5].
- Nhóm thứ hai, đề xuất một phương pháp phân loại hoạt động người trên môi trường phải thay đổi địa hình (lên, xuống cầu thang ở tòa nhà cao tầng) cho phép phân loại hoạt động Đứng, Ngồi, Chạy bộ, Đi bộ, Lên cầu thang, Xuống cầu thang. Phương pháp này đã được công bố tại công trình [CT3], [CT4].
- Bên cạnh đó, phần kết luận của luận án cũng thảo luận về các công việc nghiên cứu tiếp trong tương lai. Trong đó đề cập đến các hướng ứng dụng các phương pháp nhận dạng hoạt động đã nghiên cứu vào thực tiễn. Hướng ứng dụng này cũng đã được công bố tại công trình [CT6].

Bố cục của luận án



Hình 0.1 Lộ trình nghiên cứu của luận án

Luận án được thiết kế bao gồm phần mở đầu và ba chương nội dung. Thêm vào đó là phần kết luận, danh mục các công trình của tác giả và danh mục các tài liệu tham khảo. Cụ thể, nội dung của các phần được tóm tắt như sau:

Chương 1 Trình bày tổng quan về tình hình nghiên cứu HAR trên thế giới và trong nước, các phương pháp nhận dạng hoạt động, các thuật toán học máy, cơ sở lý thuyết về xử lý tín hiệu cảm biến. Chỉ ra các khoảng trống nghiên cứu và định hướng giải quyết của luận án.

Chương 2 và Chương 3 trình bày các đóng góp chính của luận án. Mỗi chương được thiết kế là một nhóm phương pháp đề xuất: Chương 2 cho môi trường bằng phẳng, Chương 3 cho môi trường có thay đổi độ cao (tòa nhà có cầu thang). Cụ thể, Chương 2 khảo sát sử dụng thuật toán có độ phức tạp thuật toán thấp trên một bộ vi điều khiển hiệu năng thấp. Kết quả cho thấy, thuật toán rừng ngẫu nhiên RF (Random Forest) là một lựa chọn tốt do việc triển khai nhận dạng các hoạt động đơn giản trong thời gian thực với chi phí tính toán thấp. Tác giả thiết kế một hệ thống nhận dạng hoạt động chi phí thấp với khả năng phản hồi theo thời gian thực. Hệ thống kết hợp các mạch điện tử, xử lý tín hiệu và mô hình học máy. Tất cả các bước bao gồm: Thu thập dữ liệu, Tiền xử lý, Trích xuất đặc trưng, Phân loại trực tiếp trên vi điều khiển và Truyền không dây đến máy chủ. Hệ thống rất nhỏ gọn nhưng hiệu quả. [CT1], [CT2], [CT5].

Chương 3 trình bày chi tiết phương pháp cho việc nhận dạng hoạt động ở môi trường có địa hình thay đổi. Phương pháp đề xuất bao gồm hai thuật toán RF với trường hợp không phân biệt hoạt động lên, xuống cầu thang [CT3], thuật toán KNN cho phép phân biệt chính xác hoạt động lên hay xuống cầu thang [CT4]. Sự khác biệt này để làm rõ tính hiệu quả của hệ thống trong mỗi trường hợp cụ thể.

Cuối cùng, phần kết luận nêu các đóng góp chính của tác giả trong luận án và những hướng phát triển nghiên cứu.

CHƯƠNG 1: TỔNG QUAN VỀ HỆ THỐNG NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN

Nội dung trong Chương 1 này giới thiệu về bài toán nhận dạng hoạt động người sử dụng cảm biến. Các nghiên cứu liên quan sẽ được phân vào nhóm những phương pháp tiếp cận. Nhìn chung, trong bài toán HAR, có ba cách tiếp cận phổ biến: cách tiếp cận dựa trên thị giác máy tính, cách tiếp cận dựa trên cảm biến tương tác môi trường và cách tiếp cận dựa trên cảm biến thiết bị đeo. Mỗi cách tiếp cận phù hợp với các tình huống và các môi trường khác nhau. Các thách thức của các phương pháp nhận dạng hoạt động người sử dụng cảm biến cũng được phân tích chi tiết làm tiền đề cho việc đề xuất các thuật toán ở các chương tiếp theo. Bên cạnh đó, một số kiến thức nền tảng cần thiết để thuận tiện cho việc hiểu thuật toán đề xuất và một số chỉ số đánh giá phổ biến cho nhận dạng hoạt động người cũng được trình bày trong chương này.

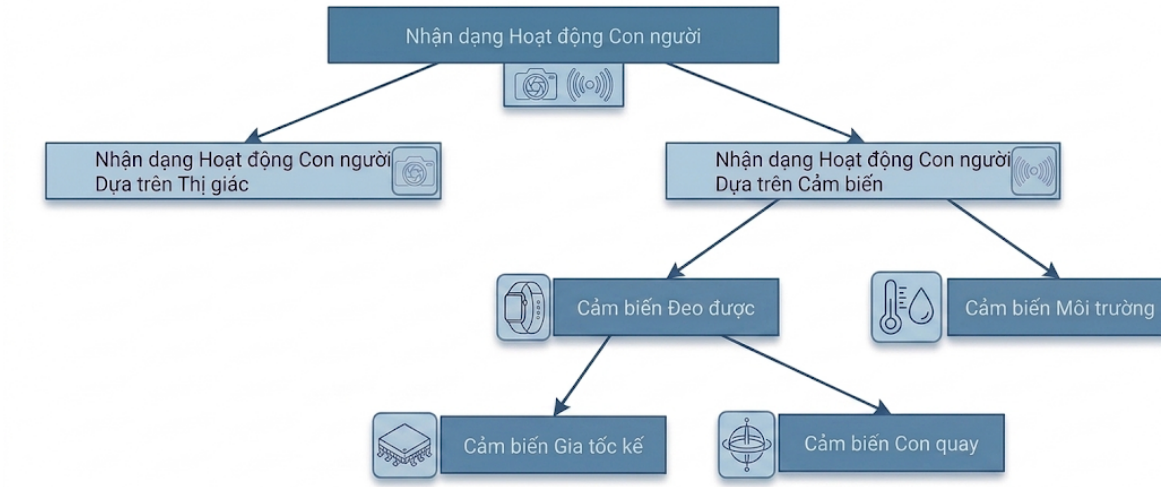
1.1. Giới thiệu về nhận dạng hoạt động người

Ngày nay, HAR đã trở thành một lĩnh vực nghiên cứu sôi động trên toàn thế giới, với sự phát triển mạnh mẽ của công nghệ cảm biến và học máy [27]. Có nhiều vấn đề thực tế cần giải pháp dựa trên nhận dạng hoạt động như chăm sóc sức khỏe [2], [3], phát hiện té ngã cho người già [4], [5], công nghiệp [7], bảo mật [28], v.v. Công nghệ nhận dạng hoạt động thu thập thông tin hành vi của người dùng, cung cấp nhiều cách tương tác khác nhau, từ đó cho phép hệ thống hỗ trợ tích cực cho người dùng trong công việc của họ [29].

Trong bài toán HAR, có hai cách tiếp cận phổ biến: cách tiếp cận dựa trên thị giác máy tính, cách tiếp cận dựa trên cảm biến (bao gồm: cách tiếp cận dựa trên cảm biến tương tác môi trường và cách tiếp cận dựa trên cảm biến thiết bị đeo) [9], [10], Hình 1.1.

Máy ảnh hoặc video được sử dụng cho phương pháp tiếp cận dựa trên tầm nhìn để quan sát hành vi của người dùng và những thay đổi của môi trường xung quanh [12], [13]. Tuy nhiên, việc sử dụng camera giám sát còn nhiều hạn chế như giá thành cao, kém hiệu quả trong điều kiện ánh sáng yếu và vấn đề riêng tư của người dùng [30]. Với cách tiếp cận dựa trên cảm biến tương tác với môi trường, cảm biến được nhúng vào đối tượng hoặc chủ thể cấu thành môi trường xung quanh để giám sát hoạt động của người

dùng [6], [14]. Giống như cách tiếp cận dựa trên thị giác máy tính [30], hệ thống này cần được triển khai và cài đặt sẵn [6], [14].



Hình 1.1 Phân loại phương pháp nhận dạng hoạt động người

Khác với cách tiếp cận trên, cách tiếp cận dựa trên cảm biến mang trên người sử dụng cảm biến trên cơ thể đã mở ra nhiều ứng dụng tiềm năng trong HAR [15], [16]. Với sự phát triển của công nghệ, các cảm biến ngày càng đa dạng hơn để nâng cao tỷ lệ nhận dạng và thích ứng với các tình huống khác nhau. Chúng phù hợp với toàn bộ môi trường mà không cần lắp đặt, dù trong nhà hay ngoài trời [31]. Theo phương pháp này, cảm biến có thể được đeo ở nhiều vị trí khác nhau trên cơ thể con người, chẳng hạn như cánh tay, cổ tay [32], đùi [33], túi trước [17] hay mắt cá chân [34]. Luận án tập trung khai thác khả năng nhận dạng trạng thái của con người bằng cách sử dụng các cảm biến mang trên cơ thể.

HAR dựa trên video là một lĩnh vực nghiên cứu tích cực với nhiều công trình đầy triển vọng [35] [36]. Những thay đổi ngẫu nhiên về quần áo, phông nền, ánh sáng... của con người là những thách thức trong HAR dựa trên video [35]. Nhiều phương pháp mới cho HAR đã được trình bày với kết quả đầy hứa hẹn dựa trên các kỹ thuật này [35] [36]. Tuy nhiên, người dùng (đối tượng) được theo dõi trong HAR dựa trên video khá khác so với HAR dựa trên cảm biến. Điểm chính là sự riêng tư của người dùng. Vì vậy, ứng dụng của các phương pháp này là khác nhau. Những phương pháp này ảnh hưởng sự riêng tư, đòi hỏi về thiết bị, thời gian xử lý cũng như không gian bộ nhớ. Sự so sánh các phương pháp nhận dạng được thể hiện trong Bảng 1.1.

Bảng 1.1 So sánh các phương pháp nhận dạng hoạt động

Phương pháp	Cảm biến	Ưu điểm	Nhược điểm	Độ chính xác
Thị giác [12], [13]	Camera	Không xâm nhập	Chi phí cao, vấn đề riêng tư, giới hạn góc nhìn	85-90%
Cảm biến môi trường [6], [14]	IR, áp suất	Không đeo	Chi phí lắp đặt cao, không di động	80-85%
Cảm biến đeo [15], [16]	Gia tốc, gyro	Di động, chi phí thấp	Cần đeo liên tục	95-99%

Tình hình nghiên cứu trong nước. Ở Việt Nam, hướng nghiên cứu về HAR và các giải pháp nhận dạng hoạt động từ xa đã được thực hiện với một số công trình tiêu biểu:

Trần Văn Lăng (2014) thực hiện đề tài "Phát triển các module phần mềm để xác thực đối tượng và nhận dạng hành vi của đối tượng kết hợp giữa kỹ thuật xử lý ảnh với điện tử" [23]. Nghiên cứu này tập trung vào nhận dạng hành vi sử dụng camera và kỹ thuật xử lý ảnh, kết hợp với thẻ thông minh RFID để xác thực đối tượng. Hệ thống đạt độ chính xác 82% cho 5 hoạt động cơ bản trong môi trường kiểm soát. Nhược điểm của nghiên cứu là chi phí cao do sử dụng nhiều camera, vấn đề về quyền riêng tư và không hoạt động tốt trong điều kiện ánh sáng kém.

Phạm Văn Cường (2017) nghiên cứu chế tạo thiết bị đeo thông minh iShoes nhận dạng hoạt động của người [24]. Thiết bị này tích hợp cảm biến áp suất và gia tốc vào đế giày, cho phép nhận dạng các hoạt động như đi bộ, chạy, đứng, ngồi. Nghiên cứu sử dụng HMM (Hidden Markov Model) để phân loại và đạt độ chính xác 88%. Nhược điểm là chi phí cao (mỗi đôi giày khoảng 2 triệu đồng), cần cặp giày riêng và khó bảo trì.

Trần Đỗ Đạt (2013) [37] nghiên cứu ứng dụng công nghệ thông tin trong xây dựng hệ thống giám sát hỗ trợ chăm sóc bệnh nhân tại phòng điều trị thông minh. Nghiên cứu này tập trung vào việc xây dựng hạ tầng phòng bệnh thông minh với các cảm biến môi trường (nhiệt độ, ánh sáng, chuyển động) để theo dõi bệnh nhân.

Nhận xét chung về các nghiên cứu trong nước:

- Hầu hết các nghiên cứu vẫn ở giai đoạn nguyên mẫu, chưa được triển khai thực tế rộng rãi do chi phí cao và khó bảo trì.
- Độ chính xác nhận dạng còn hạn chế (80-88%), thấp hơn so với các nghiên cứu quốc tế (95-99%).
- Chưa có nghiên cứu nào tập trung vào tối ưu hóa thuật toán để triển khai trên vi điều khiển hiệu năng thấp.
- Thiếu các bộ dữ liệu công khai cho cộng đồng nghiên cứu Việt Nam.

1.2. Tổng quan về hệ thống nhận dạng hoạt động sử dụng cảm biến

1.2.1. Hệ thống nhận dạng hoạt động sử dụng cảm biến

Nhận dạng hoạt động người (HAR) nhằm mục đích nhận ra hoạt động đã, đang xảy ra của người từ một loạt các quan sát về hoạt động của người và điều kiện môi trường xung quanh. Phân loại hoạt động sử dụng cảm biến mang trên người đang là hướng tiếp cận thu hút được sự quan tâm. Nguyên nhân là do sự tiến bộ mạnh mẽ của công nghệ, đặc biệt ở đây là các tiến bộ của các cảm biến. Chúng ngày càng có kích thước nhỏ, trọng lượng nhẹ nhưng khả năng làm việc, thu thập dữ liệu khá mạnh mẽ cho phép hoạt động liên tục không hạn chế. Hoạt động của con người có nhiều dạng và phức tạp, tuy nhiên về cơ bản có thể chia hoạt động của người thành hai loại: hoạt động mức thấp và hoạt động mức cao. Các hoạt động mức thấp được thực hiện bởi một cá nhân và kéo dài trong thời gian ngắn, ví dụ: mở cửa, ngồi xuống, đứng lên, gập cánh tay,… Các hoạt động mức cao là một tập các hoạt động mức thấp xen kẽ, chồng chéo với nhau và xảy ra trong thời gian dài có thể lên đến vài giờ. Phân loại hoạt động dựa vào dữ liệu cảm biến mang trên người đang là hướng tiếp cận chính, đồng thời được nghiên cứu và ứng dụng trong nhiều lĩnh vực: Công nghiệp, an ninh, thể thao, giải trí,… và đặc biệt là trong lĩnh vực y tế. Trong lĩnh vực y tế chăm sóc sức khỏe, phân loại hoạt động sử dụng cảm biến mang trên người được ứng dụng nhiều nhất. Những hệ thống phân loại hoạt động này dựa vào dữ liệu từ cảm biến (mang trên người), từ đó phân tích các hoạt động trong một khoảng thời gian, đưa ra các khuyến cáo, điều chỉnh để có chế độ hoạt động tốt nhất cho cơ thể, cảnh báo các nguy cơ về sức khỏe.

1.2.2. Các nghiên cứu điển hình nhận dạng hoạt động dựa trên cảm biến mang trên người

Cảm biến mang trên người có thể được nhúng vào cơ thể con người và phân bố từ đầu đến chân. Cảm biến mang trên người có thể hoạt động trong một khu vực rộng vừa phải trong khi người đeo đang thực hiện các hoạt động như đi bộ và chạy. Cảm biến mang trên người như gia tốc kế và con quay hồi chuyển thường được nhúng trong điện thoại thông minh, đồng hồ thông minh, vòng đeo tay thông minh, quần áo, thắt lưng, kính, mũ bảo hiểm hoặc giày.

Cảm biến phát hiện, cảm nhận những trạng thái, quá trình vật lý hay hóa học ở môi trường và biến đổi thành tín hiệu để thu thập thông tin về trạng thái hay quá trình đó. Cảm biến mang trên người là cảm biến được gắn trực tiếp hoặc gián tiếp trên cơ thể người: quần áo, đồng hồ, thắt lưng, thiết bị di động hoặc đặt trực tiếp trong cơ thể người. Khi người dùng thực hiện các hoạt động khác nhau, cảm biến thu thập thông tin như: vị trí và chuyển động của cơ thể, mạch đập và nhiệt độ cơ thể, sau đó chuyển thành tín hiệu truyền về hệ thống, dựa vào các dữ liệu đó hệ thống có thể phân loại trạng thái hoạt động đang thực hiện của con người. Do đó, thông qua hệ thống có thể theo dõi các đặc điểm mô tả về trạng thái hoặc chuyển động sinh lý của người. Các nhà nghiên cứu đã phát hiện ra rằng các loại thông tin cảm biến khác nhau có hiệu quả để phân loại các hoạt động khác nhau. Các cảm biến mang trên người tùy vào mục đích sử dụng chúng trong phương pháp nhận dạng hoạt động người mà có thể lựa chọn chủng loại, chức năng và độ chính xác khác nhau.

Kích thước nhỏ, trọng lượng nhẹ, khả năng thu thập dữ liệu lớn là ưu điểm để cảm biến gia tốc được áp dụng bởi cảm biến mang trên người nếu có kích thước và trọng lượng không phù hợp sẽ gây ra khó khăn cho người sử dụng. Bên cạnh đó, những thiết bị hướng đến hỗ trợ tìm kiếm cứu nạn, chữa cháy [18], [34] cần được thiết kế có độ bền cao, trọng lượng nhỏ và không gây cảm giác khó chịu khi mang những thiết bị này trên người. Một dự án của Liên minh Châu Âu có tên là Opportunity được đề xuất với mục đích phát triển một phương pháp mới để ghi nhận hoạt động sẽ loại bỏ các hạn chế về tính khả dụng, vị trí đeo và đặc điểm của cảm biến [38]. Trong dự án này, một tập dữ liệu sử dụng gia tốc kế có thể đeo được cùng với máy quay video đã được thu thập trong một bữa ăn sáng. Tổng cộng có 19 cảm biến đặt tại các vị trí khác nhau trên cơ thể của mỗi tình nguyện viên. 4 tình nguyện viên đã thực hiện

những kịch bản được xác định trước về các hoạt động như mở/đóng ngăn kéo tủ hoặc tủ lạnh, lau bàn, di chuyển cốc... Mỗi tình nguyện viên sẽ thực hiện từng kịch bản 6 lần. Tuy nhiên, 19 cảm biến khác nhau được đặt trên cơ thể của đối tượng đã khiến hệ thống tổng thể trở nên khá khó chịu. Nhiều nghiên cứu sử dụng gia tốc kế chỉ phân loại một số hoạt động đặc biệt như đi bộ, ngồi, chạy, đứng... Những hoạt động này rất tốt để xác định mức độ hoạt động của mọi người như được thể hiện trong nhiều nghiên cứu [39], [40], nhưng như thế là không đủ để truyền tải thông tin về các hoạt động sinh hoạt hàng ngày với độ chính xác cao vì chúng không thể phân tách việc ăn uống với việc đọc sách. Mặt khác, công nghệ đeo có thể cung cấp một loạt các phương thức cảm biến sinh lý để đo huyết áp, nhịp tim, nhiệt độ cơ thể, độ dẫn điện của da, điện não đồ (EEG), điện tâm đồ (ECG) và nhịp thở. Những thông tin bổ sung này khi được kết hợp với các hoạt động hàng ngày sẽ cung cấp một cái nhìn phong phú hơn về tình trạng sức khỏe của người dùng.

Cảm biến mang trên thân người dựa theo cách tiếp cận dữ liệu hoạt động:

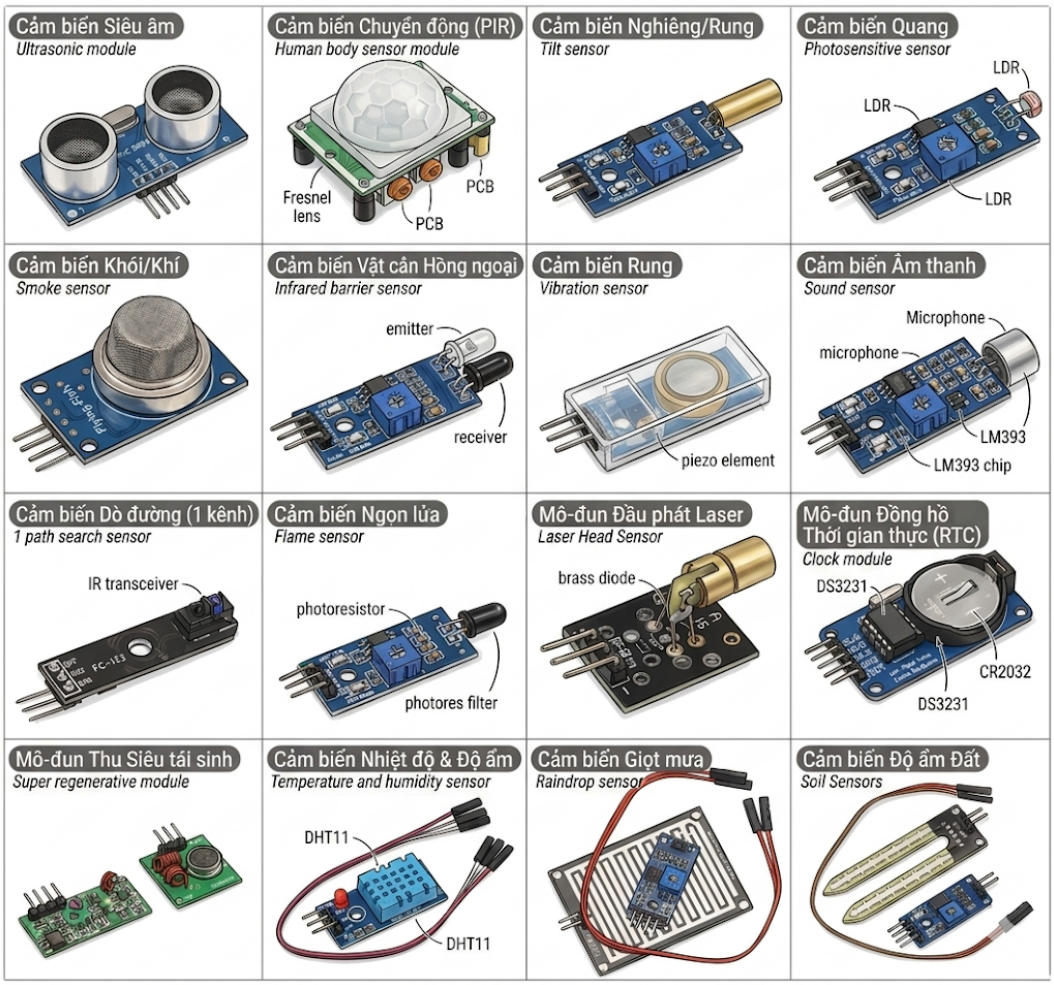
- Thông tin chuyển động của cơ thể: Cảm biến gia tốc, có thể kết hợp cảm biến con quay hồi chuyển hoặc cảm biến âm thanh.

- Lộ trình, vị trí người dùng: GPS, cảm biến âm thanh, cảm biến con quay hồi chuyển.

- Cảm biến tích hợp dựa trên một số điều kiện thiết lập trước: Bộ đọc thẻ RFID, cảm biến phát hiện tia hồng ngoại, chuyển mạch cộng từ, các cảm biến đo nhiệt độ môi trường.

- Phát hiện sự sống: Huyết áp, nhịp tim, điện não, điện tim, hô hấp, cảm biến nhiệt độ cơ thể, cảm biến áp suất, điện trở đo áp lực, cảm biến đo ôxy, cảm biến độ dẫn điện của da, điện tim...

Khi người dùng thực hiện các hoạt động khác nhau, cảm biến thu thập thông tin như: vị trí và chuyển động của cơ thể, mạch và nhiệt độ, sau đó chuyển thành tín hiệu truyền về hệ thống, dựa vào các dữ liệu đó hệ thống có thể phân loại trạng thái hoạt động đang thực hiện của con người. Do đó, họ có thể theo dõi các đặc điểm mô tả về trạng thái hoặc chuyển động sinh lý của người. Một số các loại cảm biến như Hình 1.2.



Hình 1.2 Một số cảm biến mang trên người

Các cảm biến mang trên người được phân loại như sau: nhóm cảm biến chuyển động, nhóm cảm biến vị trí, nhóm cảm biến gắn vào đối tượng cần theo dõi và nhóm cảm biến phát hiện sự sống. Cảm biến chuyển động là nhóm cảm biến thường được dùng nhiều nhất trong nghiên cứu phân loại hoạt động như cảm biến gia tốc, cảm biến âm thanh, cảm biến con quay hồi chuyển.

Cảm biến gia tốc được sử dụng phổ biến do kích thước nhỏ, trọng lượng nhẹ, có thể đo được cả tần suất và cường độ chuyển động, độ tin cậy cao. Các cảm biến gia tốc đặc biệt hiệu quả trong việc giám sát hoạt động cơ thể lặp đi lặp lại như: đi bộ, chạy bộ, đứng, ngồi, leo cầu thang.

Cảm biến con quay hồi chuyển giúp tăng độ chính xác khi phân loại một số hoạt động và có khả năng định vị. Nhóm cảm biến cuối cùng là một công nghệ mới nhằm mục đích giám sát hoạt động thông qua các dấu hiệu sự sống. Một loạt các cảm biến khác nhau đã

được nghiên cứu để đo lường nhiều loại dấu hiệu sự sống như đo nhịp tim, huyết áp, điện não đồ, điện tâm đồ và thông tin hô hấp, … theo dõi nhiệt độ cơ thể của các binh sĩ để phát hiện hạ thân nhiệt.

Trong các nhóm cảm biến nêu trên, kích thước nhỏ, trọng lượng nhẹ, khả năng thu thập dữ liệu lớn là ưu điểm để cảm biến gia tốc được sử dụng nhiều trong các công trình nghiên cứu, phần lớn các nghiên cứu trước đây sử dụng cảm biến gia tốc để phân loại hoạt động [40]. Các cảm biến gia tốc mang trên eo, đùi, mắt cá chân có thể phân loại các hoạt động nằm, đứng, đi bộ, chạy bộ, leo cầu thang với độ chính xác đến 95% Trong một số công trình nghiên cứu sử dụng đeo cảm biến gia tốc hai trục tại 5 vị trí khác nhau trên cơ thể người bao gồm: cổ tay, bắp cánh tay, eo, đùi và mắt cá chân thực hiện phân loại cho kết quả có độ chính xác 84,26% [33]. Điều đó cho thấy rằng mắt cá chân, đùi, eo, cổ tay và cánh tay là vị trí thích hợp đặt cảm biến gia tốc thu thập dữ liệu chuyển động để xây dựng mô hình nhận dạng hoạt động người.

Tóm lại, việc sử dụng cảm biến mang trên người trong các nghiên cứu là rất đa dạng, mỗi hệ thống có thể sử dụng một hoặc nhiều cảm biến kết hợp với nhau. Những thách thức chính được tóm tắt khi điều tra các hệ thống phân loại hoạt động của con người sử dụng các cảm biến có thể đeo được dưới đây.

- Thử thách 1: Do các thuật toán có độ phức tạp tính toán cao nên không thể thiết kế một thiết bị phân loại thời gian thực với độ chính xác cao.

- Thử thách 2: Nếu thiết bị không thể phân loại trực tiếp hoạt động của con người ở bộ vi xử lý giá rẻ. Dữ liệu tăng tốc thô phải được gửi đến máy chủ. Do đó, lỗi giao tiếp có thể xuất hiện. Hơn nữa, vi điều khiển hiệu suất cao dẫn đến một giải pháp chi phí cao.

Sự bùng nổ của học sâu trong giai đoạn 2021-2024 đã tạo ra ba luồng tiếp cận chính cho bài toán HAR. Luồng thứ nhất tập trung vào kiến trúc Transformer và cơ chế chú ý nhằm nắm bắt phụ thuộc dài hạn trong tín hiệu cảm biến: Suh và cộng sự [41] đề xuất TASKED kết hợp transformer đối nghịch với chắt lọc tri thức bản thân, Park và cộng sự [42] đề xuất GTSNet đáp ứng ràng buộc thời gian thực, trong khi Ullah và cộng sự [43], Teng và cộng sự [44], Yang và cộng sự [45], Wei và Wang [46] và Ezzeldin và cộng sự

[47] khai thác các biến thể chú ý và kiến trúc lai đa phương thức kết hợp mạng tích chập/hồi quy để cải thiện độ chính xác trên cùng lớp bài toán.

Luồng thứ hai hướng đến khả năng khái quát hóa liên tập dữ liệu (cross-dataset) và học tự giám sát nhằm giảm phụ thuộc vào dữ liệu gán nhãn, vấn đề thực tiễn quan trọng khi triển khai HAR ra ngoài phòng thí nghiệm. Tiêu biểu là CrossHAR của Hong và cộng sự [48], tiền huấn luyện tự giám sát phân cấp; Taghanaki và cộng sự [49] với biểu diễn tương phản miền thời gian-tần số; cùng các công trình liên quan của Kim và cộng sự [50], Nguyen và cộng sự [51], Duan và cộng sự [52] và Abdi và cộng sự [53].

Luồng thứ ba, gần với định hướng của luận án này nhất, quan tâm đến triển khai HAR trên thiết bị tài nguyên hạn chế. Bursa và cộng sự [54], [55] xây dựng mô hình TensorFlow Lite Micro cá nhân hóa qua học chuyển tiếp; Gong và cộng sự [56] đề xuất MobileHARC; Praba và Suganthi [57] kết hợp CNN với học chuyển tiếp LSTM-SVM cho thiết bị di động. Điểm chung của cả ba công trình này, nền tảng mục tiêu vẫn là bộ xử lý ARM Cortex hoặc vi điều khiển 32-bit có bộ tăng tốc phần cứng cho suy luận mạng nơ-ron, cho thấy ngay cả hướng "nhẹ hóa" trong cộng đồng deep learning vẫn chưa chạm tới ràng buộc khắt khe của vi điều khiển 8-bit như PIC18F4520.

Bên cạnh ba luồng chính trên, một số công trình khai thác các hướng chuyên biệt hơn: tổ hợp đa mô hình, mô hình dựa trên đồ thị [58] và kiến trúc lai CNN-LSTM/BiGRU cho độ chính xác cao trên các tập dữ liệu chuẩn [59]-[71]; tối ưu hóa bằng thuật toán meta-heuristic [72], [73]; xử lý dữ liệu cảm biến khiếm khuyết hoặc không đầy đủ [74]-[76]; và các ứng dụng chuyên biệt như nhận dạng hoạt động cầu nguyện [69], bài tập phục hồi chức năng [70], hoặc giám sát an ninh [77], [78]. Nhìn chung, các phân tích tổng hợp gần đây [79], [80] đều thống nhất nhận định rằng phần lớn tiến bộ của HAR hiện đại đến từ việc tăng độ phức tạp mô hình (Transformer, GAN, đồ thị nơ-ron) trên nền tảng tính toán dồi dào, trong khi hướng triển khai học máy, không nhất thiết là học sâu, trên vi điều khiển 8-bit chi phí cực thấp vẫn là một khoảng trống chưa được khai thác một cách hệ thống.

Ở trong nước, hướng nghiên cứu triển khai HAR trực tiếp trên vi điều khiển hiệu năng thấp còn tương đối mới nhưng đã bắt đầu được quan tâm. Đào Tô Hiệu và cộng sự [81] xây dựng hệ thống nhận dạng hành động sử dụng cảm biến gia tốc kết hợp các

phương pháp học máy trên vi điều khiển hiệu năng thấp, truyền dữ liệu qua WiFi để giám sát theo thời gian thực. Nguyễn Quang Huy và cộng sự [82] đề xuất một phương pháp xây dựng hệ thống nhận dạng hành vi người theo thời gian thực dựa trên các thuật toán học máy, được công bố tại Tạp chí Công nghệ Thông tin và Truyền thông. Một nghiên cứu khác trình bày tại Hội nghị Quốc gia lần thứ 26 về Điện tử, Truyền thông và Công nghệ Thông tin (REV-ECIT 2023) [83] sử dụng học máy có giám sát để phân loại hoạt động cơ bản của người dựa trên dữ liệu cảm biến gia tốc, hướng tới triển khai trên bộ vi điều khiển hiệu năng thấp. Bên cạnh đó, Trần Văn Lăng (2014) [23] và Phạm Văn Cường (2017) [24] là những nghiên cứu trong nước sớm hơn về nhận dạng hành vi và thiết bị hỗ trợ đi lại. Nhìn chung, các nghiên cứu trong nước đã bước đầu tiếp cận bài toán HAR trên phần cứng hiệu năng thấp, song còn hạn chế về số lượng hoạt động nhận dạng, chưa mở rộng sang các địa hình phức tạp (như lên/xuống cầu thang) và chưa có đánh giá hệ thống trên tài nguyên phần cứng rất hạn chế như MCU 8-bit RAM 1,5 KB. Đây là khoảng trống mà luận án này hướng tới giải quyết.

Tóm lại, từ tổng quan trong và ngoài nước có thể rút ra ba vấn đề trực tiếp dẫn đến các đóng góp của luận án: (1) giới hạn phần cứng - phần lớn hệ thống HAR, kể cả các mô hình "nhẹ hóa", vẫn chạy trên điện thoại hoặc vi điều khiển 32-bit; (2) số lượng đặc trưng - các nghiên cứu đạt độ chính xác cao thường dùng 40-64 đặc trưng hoặc đặc trưng miền tần số, không phù hợp với RAM 1,5 KB; (3) khả năng triển khai trực tiếp trên MCU 8-bit - các nghiên cứu trong nước [23], [24], [37], [81]-[83] chủ yếu dừng ở mức nguyên mẫu, chưa công bố đầy đủ số liệu về bộ nhớ, thời gian suy luận và năng lượng khi chạy trên thiết bị. Các mục tiếp theo của chương tập trung vào ba vấn đề này.

1.2.3. Khoảng trống nghiên cứu và định hướng của luận án

Qua tổng quan các công trình trong và ngoài nước, có thể nhận thấy một số khoảng trống nghiên cứu quan trọng chưa được giải quyết triệt để:

(1) Khoảng trống về phần cứng đích: Hầu hết các nghiên cứu HAR tiêu biểu như Kwapisz và cộng sự [17] (smartphone), Vavoulas và cộng sự [20] (PC với 64 đặc trưng) hay Catal và cộng sự [25] (43 đặc trưng) đều triển khai trên nền tảng có tài nguyên dồi dào. Theo khảo sát của luận án, chưa ghi nhận nghiên cứu nào công bố khả năng nhận dạng hoạt động thời gian thực với độ chính xác trên 95% trên vi điều khiển 8-bit có

RAM ≤ 1,5 KB và Flash ≤ 32 KB. Đây là ràng buộc phần cứng rất khắt khe so với phần lớn các công trình công bố.

(2) Khoảng trống về tối ưu hóa đặc trưng cho hệ nhúng: Các nghiên cứu hiện có hoặc sử dụng bộ đặc trưng phức tạp (FFT, wavelet, 50+ đặc trưng) không triển khai được trên MCU 8-bit, hoặc dùng quá ít đặc trưng dẫn đến độ chính xác thấp. Chưa có phương pháp luận rõ ràng để lựa chọn bộ đặc trưng tối thiểu nhưng vẫn đảm bảo độ chính xác cao trong ràng buộc tài nguyên cực hạn.

(3) Khoảng trống về nhận dạng hoạt động địa hình đa tầng trên MCU 8-bit: Các nghiên cứu HAR có phân biệt lên/xuống cầu thang đều dùng phần cứng hiệu năng cao. Theo khảo sát của luận án, chưa ghi nhận công trình nào triển khai bài toán 6 hoạt động (bao gồm phân biệt Lên/Xuống cầu thang) trực tiếp trên vi điều khiển 8-bit chi phí thấp với thời gian suy luận dưới 100 ms cho mỗi lần phân loại.

Từ ba khoảng trống trên, luận án đề xuất hướng giải quyết: (i) chọn PIC18F4520 (RAM 1,5 KB, Flash 32 KB, 2 USD) làm nền tảng mục tiêu để buộc thuật toán phải tối ưu tối đa, (ii) đề xuất phương pháp lựa chọn đặc trưng thống kê tối thiểu (5 đặc trưng cho RF4/5-HAR, 4 đặc trưng cho KNN6-HAR) dựa trên phân tích khả năng phân biệt trực quan kết hợp thực nghiệm loại bỏ đặc trưng, (iii) chứng minh tính khả thi qua hai hệ thống hoàn chỉnh từ phần cứng đến phần mềm nhúng, được kiểm chứng trên dữ liệu công khai WISDM và dữ liệu tự thu thập.

Phân tích hoạt động sử dụng các cảm biến gia tốc kết hợp thuật toán học máy đã trở thành một lựa chọn tốt và là bài toán của tác giả đặt ra. Hệ thống đề xuất có chi phí thấp và dễ sử dụng.

1.3. Hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc

1.3.1. Nhận dạng hoạt động sử dụng cảm biến gia tốc

Một số nghiên cứu nổi bật khi sử dụng đa cảm biến để nhận biết hoạt động của con người. Bao và cộng sự [33] đã thu thập dữ liệu từ năm máy đo gia tốc hai trục và đeo chúng trên nhiều vị trí khác nhau như hông, cổ tay, cánh tay, mắt cá chân và đùi trong

nhóm thực hiện nghiên cứu. Độ chính xác từ bộ phân loại Cây Quyết định cho kết quả là 84%, với 20 hoạt động được phân loại. Tuy nhiên, hạn chế là số lượng cảm biến tăng lên, tốc độ cao và gia tốc kế được đặt theo hướng xác định. Wang và cộng sự [9] cũng so sánh việc sử dụng cá nhân các cảm biến gia tốc kế và con quay hồi chuyển với

việc sử dụng đồng thời cả hai bộ phân loại KNN và Naïve Bayes để phân loại. Kết quả thử nghiệm cho thấy rằng các máy đo gia tốc cung cấp nhiều thông tin phân biệt hơn so với con quay hồi chuyển và việc sử dụng đồng thời cả hai loại cảm biến sẽ cho hiệu suất phân loại tốt hơn. Naïve Bayes cho độ chính xác tổng thể là 90,1%, cao hơn so với 87,8% của KNN. Nhiều cảm biến có thể cung cấp độ chính xác cao nhưng việc đeo các cảm biến gây ảnh hưởng đến chuyển động của con người và không thực tế khi đeo trong thời gian dài. Ngoài ra, chi phí của hệ thống sẽ tăng lên theo số lượng cảm biến.

Gần đây, nhiều nhà nghiên cứu đã tìm kiếm các phương pháp ghi nhận hoạt động bằng cách sử dụng một gia tốc kế để thu thập tín hiệu [32]. Ví dụ, Piyush Gupta và cộng sự [18] đã khám phá những lợi ích của máy đo gia tốc 3 trục đeo đai để thu thập dữ liệu nhằm xây dựng hệ thống phân loại hoạt động và lựa chọn đặc trưng. Họ đã sử dụng hai bộ phân loại Naïve Bayes, KNN và nhận thấy rằng việc sử dụng lựa chọn đối tượng với cách tiếp cận dựa trên trình bao bọc cho kết quả tốt hơn dựa trên bộ lọc. Kết quả phân loại cho thấy độ chính xác cao hơn 95% cho tất cả sáu hoạt động. Tuy nhiên, hạn chế là số đối tượng thu thập số liệu chỉ có 7 đối tượng, đều là thanh niên, độ tuổi chỉ từ 22-28 tuổi. Khan và cộng sự [19] đã sử dụng một máy đo gia tốc 3 trục duy nhất gắn trên ngực theo một hướng cụ thể để phân biệt các hoạt động khác nhau trong cuộc sống hàng ngày. Kết quả chính xác tổng thể lên đến 97,9%, với phân loại 15 hoạt động. Thật không may, khi thử nghiệm được thực hiện tại năm địa điểm khác nhau, độ chính xác của hệ thống giảm mạnh, chỉ còn khoảng 47%.

Bảng 1.2 Tổng quan về HAR dựa trên cảm biến mang trên người

Tài liệu tham khảo	**Số cảm biến**	**Vị trí đặt cảm biến**	**Hoạt động**
Zheng và cộng sự [84]	1	Eo, Cổ tay, Hông, Túi	Đứng, Nằm, Đi bộ, Nhảy múa, Chạy, Lên cầu thang, Xuống cầu thang, Chạy bộ, Nhảy dây, Ngồi
Gjoreski and Gams [85]	24	Mắt cá chân phải, Đùi trái, Ngực	Ngồi, Đi xuống, Đứng lên, Nằm, Ngồi trên mặt đất, Đứng

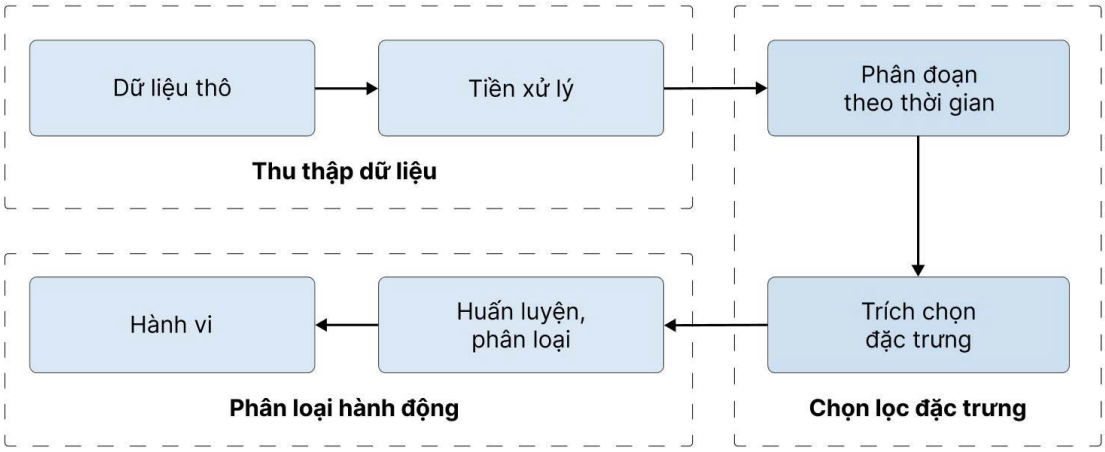
Jiang và cộng sự [86]	4	Cẳng tay phải, Cẳng chân phải, Cẳng tay trái, Cẳng chân trái	Đứng, Đi bộ trên máy tập luyện hình elip, Ngồi, Chạy trên máy tập luyện hình elip, Nằm trên giường, Đạp xe, Chạy bộ, Chèo thuyền, Đi bộ và tập tạ
Kwapisz và cộng sự [17]	1	Điện thoại thông minh	Ngồi, Chạy bộ, Lên cầu thang, Đi bộ, Xuống cầu thang, Đứng
Zhu and Sheng [87]		Đùi phải	Từ đứng sang ngồi, Ngồi, Từ nằm sang ngồi, Đứng, Nằm, Từ ngồi sang đứng, Đi bộ, Từ ngồi sang nằm
Siirtola and Roning, 2012 [88]	1	Điện thoại thông minh đặt trong túi	Đạp xe, Đi bộ, Lái xe, Đứng, Ngồi
Hemalatha and Vaidehi [89]	1	Ngực	Té ngã, Đi bộ, Đứng, Nằm, Ngồi
Mannini và cộng sự [90]	1	Cổ tay, Mắt cá chân	26 hoạt động hàng ngày
Gao và cộng sự [91]	4	Ngực, Dưới cánh tay trái, Eo và đùi	Đứng, Từ đứng sang ngồi, Đi bộ, Ngồi, Từ nằm sang đứng, Nằm, Từ ngồi sang đứng, Xuống cầu thang, Từ đứng sang nằm

Trong [20], các tác giả giới thiệu bộ dữ liệu MobiAct, thu thập từ cảm biến gia tốc điện thoại thông minh, bao gồm cả các hoạt động sinh hoạt hàng ngày (đứng, ngồi, đi bộ, chạy, lên/xuống cầu thang...) và các tình huống ngã, phục vụ đồng thời bài toán nhận dạng hoạt động và phát hiện té ngã. Ở đây, tập dữ liệu hoạt động WISDM [17] cũng được sử dụng cho mục đích xác minh. Bảng 1.2 trình bày tổng quan về HAR dựa trên cảm biến gia tốc.

1.3.2. Mô hình hệ thống nhận dạng hoạt động sử dụng cảm biến gia tốc

Cảm biến gia tốc thu thập dữ liệu chuyển động và gửi thông tin dữ liệu về hệ thống. Hệ thống nhận dữ liệu (Hình 1.3) từ cảm biến để có thể phân loại các hoạt động người

thường trải qua những bước sau: thu thập dữ liệu (thu thập dữ liệu thô và tiền xử lý), chọn lọc đặc trưng (phân đoạn, trích chọn đặc trưng), phân loại hoạt động (huấn luyện, sau đó có thể hoạt động phân loại).



Hình 1.3 Mô hình hệ thống phân loại hoạt động người

Thu thập dữ liệu và tiền xử lý: Dữ liệu cảm biến thu thập được là dữ liệu thô mà cảm biến chuyển đổi nhờ vào khả năng cảm thụ chuyển động của người thành các tín hiệu điện tử, dữ liệu thu thập được có thể được lưu vào bộ nhớ gắn trên thiết bị hoặc truyền không dây về máy chủ lưu trữ hoặc sử dụng phân loại trực tiếp theo thời gian thực [31]. Tùy theo hoạt động cần phân loại, dữ liệu cảm biến thu thập được có thể có nhiều giá trị, hoặc nhiều cảm biến cùng lấy mẫu một lúc với nhau. Dữ liệu sau khi thu thập thường được trải qua quá trình xử lý ban đầu do trong quá trình lấy mẫu dữ liệu thu thập có thể bị ngắt quãng, bị nhiễu hoặc tần số bị biến đổi, trong các trường hợp truyền dữ liệu về máy chủ, dữ liệu có thể bị mất trong quá trình truyền, đặc biệt cảm biến mang trên người dễ bị rung động do việc cố định cảm biến là khó khăn. Mục đích cuối cùng là trích xuất ra được các thông tin quan trọng về hoạt động người.

Chọn lọc đặc trưng: Dữ liệu sau khi tiền xử lý được chia thành các phân đoạn theo thời gian. Lựa chọn độ dài của cửa sổ thời gian hợp lý giúp cho hiệu quả phân loại tốt. Tùy vào đối tượng và phương pháp phân loại, chúng ta cần lựa chọn độ dài của cửa sổ thời gian một cách hợp lý để đạt được hiệu quả phân loại tốt nhất. Nguyên nhân là nếu cửa sổ thời gian quá ngắn sẽ thiếu, không đủ thông tin về hoạt động, nếu cửa sổ quá dài thì xác suất một cửa sổ chứa nhiều hơn một hoạt động được tăng lên làm độ trễ tăng lên và tỉ lệ các hoạt động không được phân loại tăng. Bên cạnh đó, chọn lọc đặc tính nổi bật

của dữ liệu để tăng hiệu quả phân loại và giảm kích thước dữ liệu đầu vào. Đây là quá trình rút gọn tín hiệu thành các đặc trưng để phân biệt các hoạt động, sau đó làm cơ sở để phân lớp các nhóm hoạt động. Đặc trưng được lựa chọn cho hiệu quả phân tách càng rõ ràng thì hệ thống càng dễ nhận dạng hoạt động. Trong các hệ thống nhúng xử lý thời gian thực thì bộ nhớ, khả năng tính toán và băng thông khá hạn chế. Vì vậy cần sử dụng số đặc trưng ít nhất trong khi vẫn đảm bảo hiệu năng cho hệ thống phân loại.

Phân loại hoạt động: Các đặc trưng chọn lọc được gán nhãn tương ứng với hoạt động và chia thành tập dữ liệu huấn luyện, tập dữ liệu kiểm tra. Phương pháp phân loại sử dụng trong bài toán phân loại hoạt động người này thường được phân thành 2 loại theo cách thức xây dựng mô hình nhận hoạt động.

- Nhận dạng hoạt động dựa trên tri thức chuyên gia: Phương pháp này dựa trên sự suy diễn theo quy luật được xây dựng trước đó. Các quy luật này được khai thác cho việc mô hình hóa các hoạt động và dữ liệu cảm biến, sau đó sử dụng suy luận logic để thực hiện nhận dạng hoạt động. Ưu điểm của cách tiếp cận này là rõ ràng về ngữ nghĩa và có tính logic cao. Các luật suy diễn có thể được xây dựng nhanh chóng và kết hợp với các giải thuật heuristic để xây dựng các mô hình phân loại và tổng hợp dữ liệu. Tuy nhiên, nhược điểm là cần rất nhiều thời gian và kinh nghiệm của chuyên gia để xây dựng được tập luật suy diễn tốt, việc cập nhật tự động các luật là không khả thi do nguồn dữ liệu đầu vào thường không có cấu trúc và luôn biến động, không có khả năng xử lý thông tin tạm thời và không rõ ràng. Hầu hết các phương pháp đều không cung cấp cách thức để xác định xem một mô hình cụ thể nào đó có hiệu quả hơn các mô hình khác hay không. Bên cạnh đó, các phương pháp dựa trên tri thức và suy luận logic không có khả năng học, suy đoán thông minh. Chính vì vậy nhận dạng hoạt động dựa trên tri thức có xu hướng ít được sử dụng.

- Nhận dạng dựa trên dữ liệu: với phương pháp này, dữ liệu sau khi trích chọn đặc trưng được huấn luyện với các thuật toán học máy để nâng cao độ chính xác. Do dựa trên dữ liệu nên cách tiếp cận này đảm bảo hệ thống có thể cập nhật các quy tắc nhận dạng hoạt động một cách tự động mà không phụ thuộc vào tri thức chuyên gia, có khả năng xử lý các thông tin tạm thời và không rõ ràng, kể cả các thông tin không đầy đủ, đồng thời hệ thống có khả năng thích nghi cao và tận dụng được nguồn dữ liệu. Tuy

nhiên cần dựa vào độ phức tạp thuật toán và hiệu năng của thuật toán để lựa chọn thuật toán huấn luyện thích hợp. Do đó, luận án này sẽ tiếp cận theo hướng nhận dạng hoạt động dựa trên ứng dụng thuật toán học máy.

1.3.3. Phương pháp đánh giá mô hình

Hiệu suất của mô hình phân loại thường được đánh giá bằng cách sử dụng tập dữ liệu thử nghiệm. Phương pháp được sử dụng phổ biến nhất để đo hiệu suất là ma trận nhầm lẫn, so sánh đầu ra thực tế với đầu ra dự đoán để đánh giá độ chính xác của phân loại. Để nâng cao khả năng nhận dạng, một số chỉ số phổ biến được sử dụng, bao gồm độ chính xác tổng thể, độ nhạy, PPV (Giá trị dự đoán dương) và NPV (Giá trị dự đoán âm). Các công thức cho các số liệu này được cung cấp dưới đây:

$$acc_i = \frac{TPs_i + TN_i}{TPs_i + FPs_i + TNs_i + FNs_i} \tag{1.1}$$

$$sen_i = \frac{TPs_i}{TPs_i + FNs_i} \tag{1.2}$$

$$PPV_i = \frac{TPs_i}{TPs_i + FPs_i} \tag{1.3}$$

$$\mathrm{NPV}_i = \frac{\mathrm{TN}_i}{\mathrm{TN}_i + \mathrm{FN}_i} \tag{1.4}$$

Trong đó:

i đại diện cho một lớp hoạt động.

TP (True Positive): Số lượng hoạt động cần phân loại được phân loại đúng so với quan sát.

FN (False Negative): Số lượng hoạt động cần phân loại bị phân loại thành các hoạt động khác.

FP (False Positive): Số lượng hoạt động khác bị phân loại thành hoạt động cần phân loại.

TN (True Negative): Số lượng hoạt động khác được phân loại đúng so với quan sát thực tế

Các bộ dữ liệu ở trên thể hiện sự chênh lệch đáng chú ý về kích thước giữa các lớp dữ liệu, trong đó dữ liệu Đi bộ tương đối lớn hơn so với Dữ liệu đứng. Do đó, để đánh giá việc phân loại tập dữ liệu, hai số liệu phù hợp sẽ là trung bình vi mô (cho trọng số bằng nhau cho từng mẫu) và trung bình vĩ mô (cho trọng số bằng nhau cho từng danh mục)

Để thống nhất cách báo cáo trong toàn luận án, các chỉ số được định nghĩa như sau. Với ma trận nhầm lẫn K lớp trên N mẫu kiểm tra: (a) Độ chính xác tổng thể đa lớp (overall accuracy) bằng tổng số mẫu được phân loại đúng của tất cả các lớp chia cho N; (b) Độ chính xác theo lớp được tính theo công thức (1.1) bằng cách xét từng lớp với phần còn lại (one-vs-rest); (c) Trung bình vĩ mô của một chỉ số là trung bình cộng giá trị của chỉ số đó trên K lớp; (d) Trung bình vi mô được tính bằng cách cộng dồn TP, FP, TN, FN của tất cả các lớp trước khi áp dụng các công thức (1.1)-(1.4). Với bài toán đa lớp, độ nhạy vi mô và PPV vi mô luôn bằng độ chính xác tổng thể đa lớp, còn độ chính xác theo lớp tính theo trung bình vi mô bằng trung bình vĩ mô khi các lớp được đánh giá trên cùng N mẫu. Độ chính xác cân bằng (balanced accuracy) là trung bình vĩ mô của độ nhạy; chỉ số F1 của mỗi lớp bằng $2 \cdot PPV \cdot sen/(PPV + sen)$ và macro-F1 là trung bình vĩ mô của F1 trên các lớp.

1.3.4. Phân tích lựa chọn thuật toán cho vi điều khiển 8-bit

Khi triển khai học máy trên PIC18F4520 (RAM 1,5 KB, Flash 32 KB, xung nhịp 40 MHz), lựa chọn thuật toán phải dựa trên ba ràng buộc cứng: (1) Kích thước mô hình ≤ 30 KB Flash, (2) Bộ nhớ suy luận ≤ 1,0 KB RAM, (3) Thời gian suy luận ≤ 100 ms. Các ngưỡng này là yêu cầu thiết kế do luận án đặt ra (không kế thừa nguyên văn từ tài liệu), được suy ra từ thông số PIC18F4520: Flash 32 KB trừ phần dành cho mã chương trình, thư viện giao tiếp I^2C/UART; RAM 1,5 KB trừ phần dành cho ngăn xếp, biến hệ thống và bộ đệm truyền thông; thời gian suy luận được tính từ khi có đủ dữ liệu của một cửa sổ đến khi có nhãn, phân biệt với chu kỳ cập nhật kết quả 5-6 giây của hệ thống. Ngưỡng 100 ms bảo đảm phần tính toán chỉ chiếm phần rất nhỏ (<2%) của chu kỳ cập nhật, để

vi điều khiển còn thời gian thu mẫu ở 50 Hz và điều khiển mô-đun truyền thông. Bảng 1.3 so sánh bốn ứng viên chính theo các ràng buộc này.

Bảng 1.3 So sánh thuật toán theo ràng buộc phần cứng PIC18F4520

Thuật toán	**Độ phức tạp suy luận**	**Bộ nhớ mô hình**	**Độ phù hợp với MCU 8-bit**	**Lý do chính**
RF (50 cây, depth=5)	O(T*depth) 250 ops	1,2 KB Flash	Đạt	Cây nhị phân tĩnh, 0 RAM khi suy luận
KNN (k=5, nén)	O(N·d + N·log k) ≈ 14.400 phép nhân-cộng	28,8 KB Flash	Đạt (nén)	1200 mẫu* 12d * 2B = 28,8 KB ≤ Flash 32 KB
SVM (RBF kernel)	O(N_sv*d)	≥2 KB RAM	Không đạt	Support vectors vượt giới hạn RAM 1,5 KB
GBDT (100 cây)	180 ms suy luận	8 KB Flash + 2 KB RAM	Không đạt	Độ trễ 180 ms vượt yêu cầu <100 ms thời gian thực

Bảng 1.3 là ước lượng ban đầu (trước thực nghiệm) dùng để sàng lọc ứng viên; lựa chọn cuối cùng dựa trên kết hợp độ chính xác (Bảng 3.8-3.11) và thời gian thực thi đo trên thiết bị (Bảng 3.12). Kết luận: RF được chọn cho RF4-HAR và RF5-HAR vì đạt độ chính xác cao nhất ở hầu hết kích thước cửa sổ, cấu trúc cây được nhúng tĩnh vào Flash dưới dạng C-array (không dùng RAM khi suy luận) và thời gian thực thi đáp ứng ràng buộc. KNN được chọn cho KNN6-HAR vì ranh giới phi tuyến giữa Lên/Xuống cầu thang phù hợp với cơ chế phân loại theo khoảng cách cục bộ, tập huấn luyện được nén xuống 1.200 mẫu để lưu trong Flash. SVM không được chọn vì có thời gian thực thi lớn nhất trong bốn thuật toán (Bảng 3.12) và cần lưu các support vector; GBDT có thời gian thực thi đo được ngắn (Bảng 3.12) nhưng độ chính xác thấp hơn RF ở mọi kích thước cửa sổ (Bảng 3.8, 3.10), do đó cũng không được chọn. Ước lượng 180 ms cho GBDT

trong Bảng 1.3 là ước lượng thô cho cấu hình lớn hơn, không phải cấu hình đã đo ở Chương 3.

1.3.5. TinyML và so sánh với các hệ thống embedded HAR hiện đại

TinyML là lĩnh vực triển khai học máy trên vi điều khiển tài nguyên hạn chế [21], [22]. Hai hướng chính: (1) Framework-based: dùng TensorFlow Lite Micro (TFLM), Edge Impulse hay STM32Cube.AI để chạy mạng neural nhỏ trên ARM Cortex-M4+, (2) Algorithm-based: tối ưu thuật toán cổ điển (RF, KNN) cho MCU 8-bit. Bảng 1.4 so sánh và xác định vị trí của luận án.

Bảng 1.4 So sánh framework-based TinyML và algorithm-based (luận án đề xuất)

Tiêu chí	**Framework-based (TFLM/Edge Impulse)**	**Algorithm-based (Luận án, RF/KNN)**
Nền tảng phần cứng	ARM Cortex-M4+ (≥256 KB RAM)	PIC18F4520 8-bit (1,5 KB RAM)
Loại mô hình	CNN, LSTM, MobileNet	RF, KNN, thuật toán cổ điển
Chi phí phần cứng	5-30 USD (STM32, Arduino Nano 33)	2 USD (PIC18F4520)
Tiêu thụ điện	12-80 mA	2,51 mA
Độ chính xác HAR	93-99%	86-99,6% (đề xuất)
Công cụ lập trình	TFLite Micro, STM32CubeAI, toolchain	C thuần túy, MPLAB X
Ưu điểm nổi bật	Linh hoạt, dễ mở rộng, có framework	Chi phí/tài nguyên tối thiểu, không cần framework

Xu hướng nghiên cứu HAR hiện đại (2021-2024) tập trung mạnh vào hai hướng chính. Hướng thứ nhất là các kiến trúc học sâu tiên tiến: tự giám sát và học tương phản [49]-[51]; transformer và cơ chế chú ý [41], [43]-[47]; mạng tích chập và LSTM kết hợp [57], [60]-[66], [68], [70], [77]; mô hình nhẹ và triển khai di động [54]-[56]; sinh dữ liệu và học bán giám sát [52], [75], [80]. Hướng thứ hai là ứng dụng HAR trong các ngữ cảnh chuyên biệt: phân tích cross-dataset [48]; kết hợp đặc trưng thủ công và học sâu

[71]; giám sát sức khỏe và phục hồi [70], [92]; an ninh và giám sát [78]; nhận dạng dựa trên video [64], [79]. Mặc dù đạt độ chính xác rất cao (trên 95-99%), hầu hết các công trình trên đều đòi hỏi nền tảng phần cứng có RAM từ vài trăm KB đến hàng GB, không phù hợp với vi điều khiển 8-bit RAM 1,5 KB. Đây chính là khoảng trống mà luận án này hướng đến giải quyết.

Điểm khác biệt cốt lõi: Gupta và cộng sự [21] triển khai TinyML trên STM32 (Cortex-M4, 64 KB RAM) đạt 93,2%, chi phí 5 USD, tiêu thụ 12 mA. Arduino Nano 33 BLE (Cortex-M4, 256 KB RAM) hỗ trợ CNN qua Edge Impulse nhưng giá 20-30 USD. Luận án này chọn PIC18F4520 (1,5 KB RAM, 2 USD, 2,51 mA), dung lượng RAM thấp hơn khoảng 43 lần so với STM32 trong [21] (64 KB) và khoảng 170 lần so với các nền tảng ARM Cortex-M4 dùng cho TFLM (≥256 KB), cho thấy HAR thời gian thực khả thi ở mức phần cứng hạn chế hơn đáng kể so với các công trình được khảo sát. Đây là điểm khác biệt của luận án so với cả hướng framework-based và algorithm-based hiện có.

1.4. Kết luận chương 1

Chương này đã giới thiệu tổng quan về các hệ thống nhận dạng hoạt động từ dữ liệu cảm biến mang trên người. Nhận dạng hoạt động sử dụng cảm biến mang trên người đã và đang thu hút nhiều sự quan tâm của cộng đồng nghiên cứu và được ứng dụng vào nhiều lĩnh vực khác nhau: từ an ninh, công nghiệp đến thể thao, nghệ thuật, đặc biệt là ứng dụng trong y tế, hỗ trợ chăm sóc sức khỏe. Trong đó, cảm biến gia tốc được sử dụng nhiều nhất trong các công trình nghiên cứu nhận dạng hoạt động cho hiệu suất phân loại cao. Đồng thời, chương 1 cũng trình bày tổng quát về một hệ thống bao gồm 3 bước chính, phương pháp đánh giá mô hình và phân tích lựa chọn các thuật toán để nhận dạng hoạt động từ dữ liệu cảm biến mang trên người.

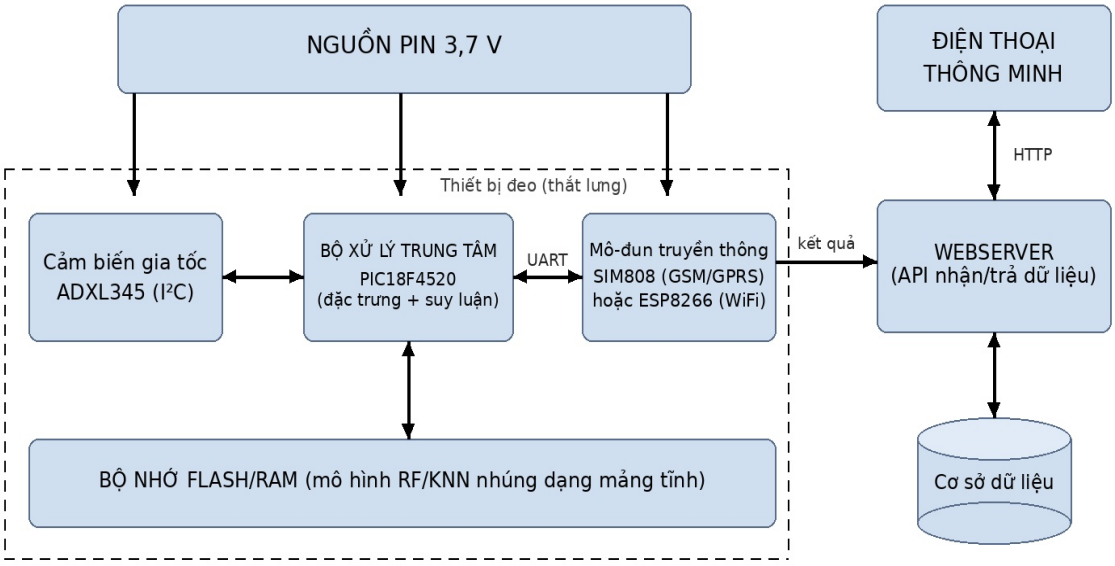
CHƯƠNG 2: NHẬN DẠNG HOẠT ĐỘNG CƠ BẢN BẰNG RANDOM FOREST TRÊN VI ĐIỀU KHIỂN 8-BIT (RF4-HAR)

Trong chương này, luận án trình bày chi tiết phương pháp đề xuất để nhận dạng bốn hoạt động cơ bản của con người (đứng, ngồi, đi bộ, chạy bộ) sử dụng dữ liệu từ cảm biến gia tốc 3 trục và các thuật toán học máy Random Forest. Nội dung chương bao gồm: Thiết kế hệ thống thời gian thực trên vi điều khiển PIC18F4520, thiết kế chi tiết phần cứng hệ thống IoT với các module cảm biến, xử lý, truyền thông và nguồn, mô tả dữ liệu sử dụng, quy trình xử lý và phân loại đề xuất, các tham số đánh giá hiệu suất, kết quả thực nghiệm với bốn thuật toán học máy khác nhau và phân tích chi tiết kết quả đạt được. Cơ sở thuật toán và tập đặc trưng ban đầu (so sánh RF, SVM, GBDT, KNN trên dữ liệu tự thu thập) đã được công bố tại công trình [CT1]. Việc mở rộng triển khai trên phần cứng nhúng với đúng bốn hoạt động của hệ thống RF4-HAR được công bố tại công trình [CT2]. Phần so sánh đối chiếu với tập dữ liệu công khai WISDM trình bày trong chương này là nội dung bổ sung của luận án; một phần kết quả liên quan cũng được công bố tại công trình [CT5].

2.1. Mô hình hệ thống RF4-HAR

2.1.1. Kiến trúc mô hình phân loại hoạt động

Luận án sử dụng học có giám sát. Cách tiếp cận tổng thể bao gồm ba giai đoạn trong vấn đề HAR này: thu thập dữ liệu thô, phân tích mô hình và các hoạt động phân loại (triển khai và đánh giá), như trong Hình 2.1 và Hình 2.3.

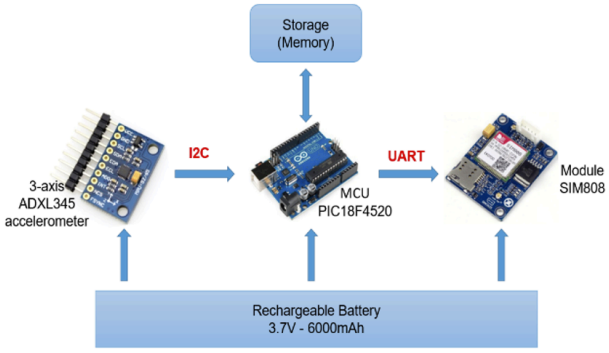


Hình 2.1 Sơ đồ khối tổng thể của hệ thống phân loại hoạt động

Trong Hình 2.1, toàn bộ quá trình thu thập tín hiệu, trích xuất đặc trưng và suy luận được thực hiện trên vi điều khiển; mô-đun truyền thông chỉ gửi nhãn hoạt động đã phân loại tới webserver. Webserver là thành phần duy nhất đọc/ghi cơ sở dữ liệu; điện thoại thông minh và giao diện web của người giám sát truy cập kết quả thông qua API của webserver, không kết nối trực tiếp với cơ sở dữ liệu.

Ở Giai đoạn 1, dữ liệu gia tốc 3 trục được thu thập. Trong Giai đoạn 2, các luồng dữ liệu được chia thành các phân đoạn nhỏ. Sau đó, mỗi vectơ được tạo từ mỗi phân đoạn trên sẽ chứa các đặc trưng được trích xuất từ dữ liệu. Bước tiếp theo là tiến hành lựa chọn đặc trưng để sử dụng cho việc huấn luyện/làm đầu vào (input) của mô hình phân loại. Và ở giai đoạn triển khai và đánh giá, các đặc trưng được chọn từ bước trên sẽ được huấn luyện để xây dựng các mô hình phân loại. Dựa vào mô hình đã học trước đó, các bộ phân loại thu được để xác định các hoạt động. Chi tiết từng nhiệm vụ sẽ được trình bày trong các phần tương ứng tiếp theo.

Hệ thống được đề xuất tích hợp cảm biến, IMU chi phí thấp và bộ truyền/thu dữ liệu không dây cho thiết bị đeo để hỗ trợ các hoạt động phân loại. Cấu trúc hệ thống đề xuất được minh họa trong Hình 2.2.



(a)



(b)

Hình 2.2 Mô hình thiết bị hệ thống đề xuất (a) và tình nguyện viên đeo thiết bị (b)

Gia tốc kế 3 DOF ADXL345 được kết nối với bộ vi điều khiển (PIC18F4520) bằng giao diện I2C (mạch liên tích hợp) để thu thập dữ liệu dọc theo ba trục. Lưu ý: Chương 2 dùng SIM808 (GSM/GPRS), Chương 3 nâng cấp lên ESP8266 (WiFi) để giảm chi phí và tăng tốc độ truyền dữ liệu. Thử nghiệm của luận án được triển khai với máy tính xách tay Dell E6540 có bộ xử lý 2,7 GHz và RAM 8 GB. PIC18F4520 được lựa chọn có chủ đích để chứng minh tính khả thi trên vi điều khiển 8-bit chi phí thấp nhất (2 USD): xung nhịp 40 MHz, RAM 1,5 KB, Flash 32 KB, bị ràng buộc hơn nhiều so với STM32 (32-bit, 5 USD) hay ESP32 (dual-core WiFi, 3 USD). Mục tiêu của luận án là tối thiểu hóa tài nguyên phần cứng cần thiết, không phải tối đa hóa hiệu năng. Nếu sử dụng STM32/ESP32, bài toán sẽ mất đi ý nghĩa nghiên cứu về tối ưu hóa thuật toán cho thiết bị nhúng hạn chế.

Vai trò của hai thiết bị được phân tách rõ ràng: toàn bộ khối lượng tính toán học máy (thu thập tín hiệu, trích xuất đặc trưng và suy luận mô hình RF/KNN) chỉ được thực hiện trên PIC18F4520, đây là đối tượng chịu ràng buộc tài nguyên mà luận án tập trung tối ưu. Mô-đun SIM808/ESP8266 chỉ đóng vai trò một kênh truyền thông thụ động, kết nối

qua UART để gửi kết quả phân loại (đã được tính toán xong) tới máy chủ giám sát; các mô-đun này không tham gia vào bất kỳ phép tính học máy nào và không thể thay thế cho vi điều khiển trung tâm. Do đó, việc sử dụng thêm mô-đun truyền thông không mâu thuẫn với mục tiêu tối thiểu hóa tài nguyên tính toán mà luận án đặt ra.

2.1.2. Tập dữ liệu

Luận án này sử dụng hai bộ dữ liệu về hoạt động của con người, bao gồm một bộ dữ liệu riêng tập trung vào các hoạt động cơ bản của người và một bộ dữ liệu công khai chứa các hoạt động tương tự.

Tập dữ liệu thu thập

Tác giả đã tiến hành ghi dữ liệu thực nghiệm với tổng cộng 16 tình nguyện viên được chia thành hai nhóm: nhóm sinh viên gồm 14 người (7 nam và 7 nữ), độ tuổi từ 18-22 tuổi, chiều cao 1,56-1,8 m, cân nặng 43-72 kg, và nhóm người cao tuổi gồm 2 người (1 nam và 1 nữ), độ tuổi 65 và 72 tuổi, chiều cao lần lượt 1,4 m và 1,46 m, cân nặng lần lượt 42 kg và 46 kg để lấy dữ liệu hoạt động thực tế hàng ngày. Thiết bị được đề xuất được tình nguyện viên giữ trên thắt lưng với tần số lấy mẫu là 50 Hz. Nó ghi lại bốn trạng thái: Đứng trong 40 phút, Ngồi trong 30 phút, Chạy trong 10 phút và Đi bộ trong 22 phút. Mỗi tình nguyện viên thực hiện từng hoạt động ít nhất 3 lần liên tục trong điều kiện bình thường tại phòng thí nghiệm. Để tăng độ đa dạng của dữ liệu, các điều kiện thu thập được thiết kế với nhiều biến thể: Đi bộ ở ba tốc độ (chậm ≈0.8m/s, bình thường ≈1.2m/s, nhanh ≈1.6m/s), Chạy bộ ở hai tốc độ (nhẹ nhàng ≈2.0m/s và nhanh ≈3.5m/s), Đứng và Ngồi trong các điều kiện khác nhau (đứng thẳng, đứng nghiêng, ngồi ghế cứng, ngồi ghế mềm). Tốc độ di chuyển được đo bằng đồng hồ bấm giờ và thước đo khoảng cách. Dữ liệu thu thập bao gồm cả trường hợp người mang túi nhẹ (<2kg) để phản ánh điều kiện thực tế hàng ngày.

Việc lựa chọn hai nhóm đối tượng nêu trên xuất phát từ mục tiêu ứng dụng của hệ thống: nhóm sinh viên (18-22 tuổi) đại diện cho người trưởng thành khỏe mạnh, có khả năng thực hiện đầy đủ và ổn định các hoạt động ở nhiều tốc độ khác nhau, giúp xây dựng bộ dữ liệu huấn luyện đa dạng; nhóm người cao tuổi (65-72 tuổi) đại diện cho đối tượng đích mà hệ thống hướng tới hỗ trợ theo dõi sức khỏe và cảnh báo té ngã. Cần lưu ý rằng dữ liệu thu thập trong luận án chỉ gồm những người không có bệnh nền và chưa

ghi nhận nguy cơ té ngã cao; do đó độ chính xác của hệ thống trên nhóm người có bệnh lý về vận động (Parkinson, tai biến, rối loạn thăng bằng...) hoặc nguy cơ té ngã cao chưa được kiểm chứng và cần được đánh giá thêm trong các nghiên cứu tiếp theo trước khi triển khai thực tế cho nhóm đối tượng này. Toàn bộ 16 tình nguyện viên đều là người khỏe mạnh; nhóm người cao tuổi chỉ có 2 người nên kết quả của luận án chưa đủ cơ sở để suy rộng cho nhóm người cao tuổi nói chung. Các khoảng chiều cao, cân nặng (1,40-1,80 m; 42-72 kg) không phải là tiêu chí sàng lọc mà là đặc điểm của mẫu tình nguyện viên, được báo cáo để người đọc đánh giá phạm vi áp dụng của kết quả.

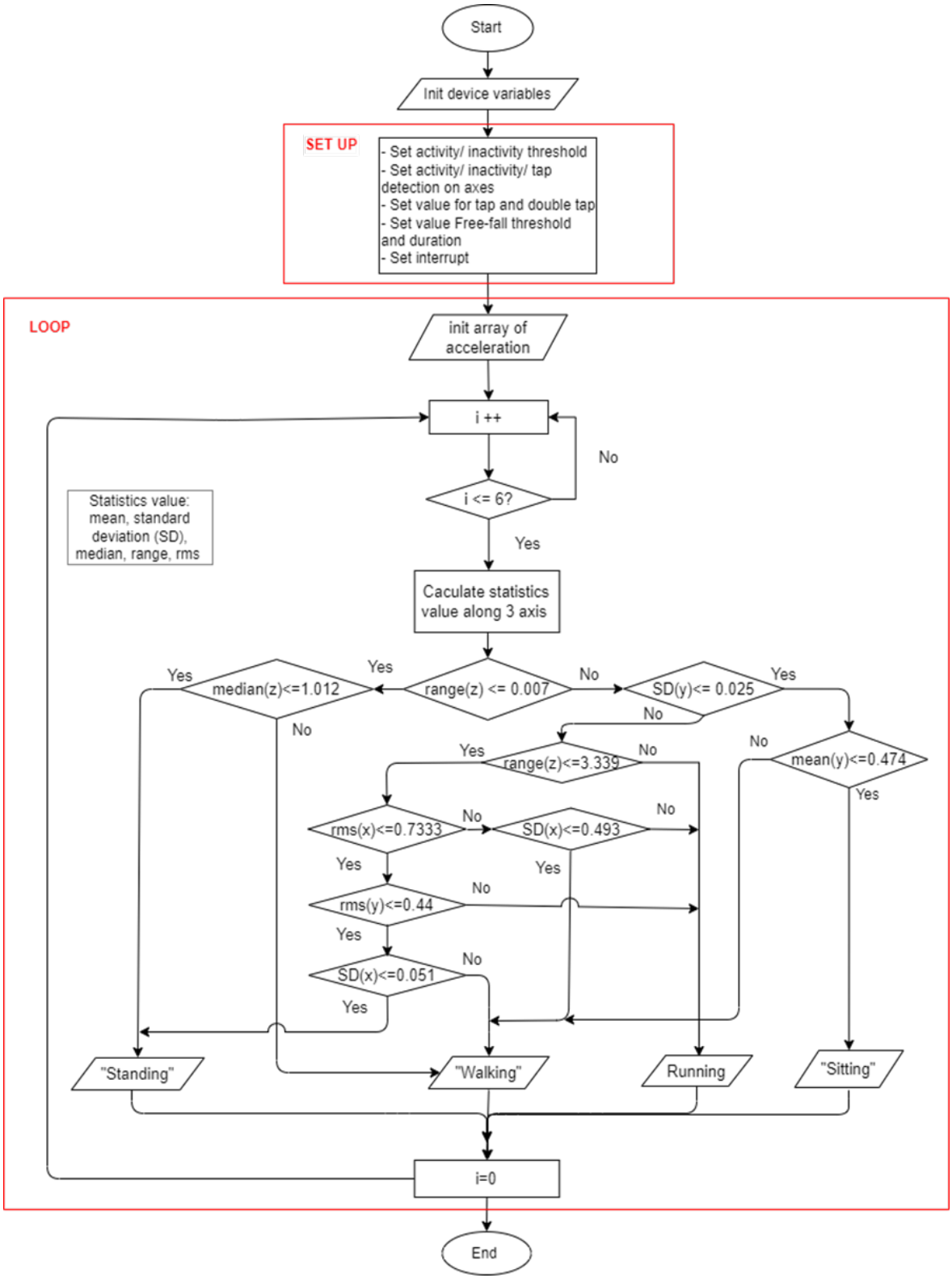
Mô tả chi tiết và một số minh họa về các hoạt động được trình bày trong Bảng 2.1 dưới đây.

Bảng 2.1 Diễn giải khái niệm các hoạt động

Hoạt động	**Diễn giải**
Ngồi	Đầu và lưng không chạm nền và đôi chân nghỉ, lưng hơi ngả phía sau
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 nhau, 2 chân không chạm đất vào 1 thời điểm

Sau khi xây dựng mô hình như Hình 2.1 và thiết kế hệ thống với các thiết bị phần cứng như Hình 2.2, tác giả áp dụng thuật toán cho thiết bị này. Bộ vi điều khiển PIC18F4520 đóng vai trò là bộ xử lý trung tâm của thiết bị nhúng, thực hiện toàn bộ quá trình thu thập dữ liệu, trích xuất đặc trưng và phân loại hoạt động theo thời gian thực. Do đó, khi thực hiện nhận dạng theo thời gian thực, cần có yêu cầu về bộ nhớ và độ phức tạp của thuật toán. Do đó, thuật toán học máy như Random Forest phù hợp để triển khai HAR trên các thiết bị đeo trong thời gian thực. Cây quyết định trong thuật toán này dựa trên dữ liệu sau khi huấn luyện để đưa ra các ngưỡng đặc trưng để hình thành cây quyết định nhị phân. Nghiên cứu này hướng tới Tiny Machine Learning (TinyML), một lĩnh vực nghiên cứu mới nhằm thiết kế và phát triển các kỹ thuật học máy nhằm thực thi

trên các hệ thống nhúng và các đơn vị Internet-of-Things (IoT). Do đó, tác giả tiến hành lập trình cho thiết bị đeo để nhận biết hoạt động theo thời gian thực. Hình 2.3 dưới đây minh họa sơ đồ thuật toán rừng ngẫu nhiên được đề xuất để phân loại theo thời gian thực.



Hình 2.3 Sơ đồ khối của chương trình nhúng để phân loại thời gian thực

Các bước sau đây được tiến hành: Đầu tiên, các biến ADXL từ thư viện ADXL345 được khai báo để thiết lập các giá trị ngưỡng của hoạt động hoặc không hoạt động, phát hiện chạm trên các trục và đặt ngắt cho bước Thiết lập. Ở bước Vòng lặp, mô hình sử dụng thuật toán rừng ngẫu nhiên sẽ quyết định hoạt động nào. Tập hợp các ngưỡng được lấy từ quá trình huấn luyện. Cứ sau 6 giây, thiết bị sẽ tính toán, phân tích và gửi lại kết quả hoạt động do thiết bị đo được để đảm bảo phản hồi theo thời gian thực và độ phức tạp ở mức tối thiểu.

Tập dữ liệu công khai

Trong nghiên cứu này, ngoài tập dữ liệu được thu thập bởi các tình nguyện viên sử dụng hệ thống được đề xuất, tập dữ liệu WISDM [17] chứa các hoạt động tương tự đã được sử dụng để so sánh, nhằm đảm bảo tính đa dạng và khách quan cho nghiên cứu.

2.1.3. Quy trình xử lý dữ liệu

Phân đoạn dữ liệu cho các hoạt động

Sau khi thu thập dữ liệu từ cảm biến, dữ liệu gia tốc sẽ được sử dụng để phân tích đặc điểm của các hoạt động. Đầu tiên, các luồng dữ liệu được chia thành các phân đoạn nhỏ chứa thông tin hoạt động. Mọi người thực hiện các hoạt động một cách suôn sẻ và các hoạt động diễn ra liên tiếp sẽ ảnh hưởng lẫn nhau. Ngoài ra, việc xác định ranh giới chính xác của hoạt động là khá khó khăn nên việc phân chia dữ liệu là cần thiết.

Có nhiều phương pháp phân đoạn, chẳng hạn như phân đoạn bằng cửa sổ trượt, phân đoạn dựa trên năng lượng, phân đoạn vị trí nghỉ, sử dụng một phương thức cảm biến để phân đoạn dữ liệu của cảm biến theo phương thức khác và các nguồn ngữ cảnh bên ngoài [85]. Luận án chọn cách tiếp cận cửa sổ trượt để phân đoạn dữ liệu vì tính đơn giản và phù hợp với các ứng dụng thời gian thực. Cách tiếp cận này đã được chứng minh là có hiệu quả trong việc nhận biết các hoạt động định kỳ (ví dụ: đi, chạy) và các hoạt động tĩnh (ví dụ: đứng, ngồi). Trong trường hợp này, tín hiệu cảm biến được phân chia thành các cửa sổ thời gian có kích thước cố định. Đây có thể có hai loại: không chồng chéo (các cửa sổ thời gian không giao nhau) và các cửa sổ chồng chéo (giao điểm của các cửa sổ liền kề).

Việc lựa chọn kích thước cửa sổ ảnh hưởng đáng kể đến hiệu suất phân loại của hệ thống [17]. Nghiên cứu chuyên sâu của Banos và cộng sự [93] về ảnh hưởng của kích

thước cửa sổ đối với quá trình nhận dạng hoạt động cũng khẳng định đây là một siêu tham số cần được khảo sát thực nghiệm kỹ lưỡng thay vì lựa chọn theo kinh nghiệm của các nghiên cứu trước. Do đó, cần phải thử nghiệm với các kích thước cửa sổ khác nhau để chọn giá trị phù hợp nhất. J.R. Kwapisz và cộng sự [17] đã phân đoạn dữ liệu thành kích thước cửa sổ 10 giây mà không chồng chéo và chứng minh rằng kích thước cửa sổ 20s cho kết quả kém hứa hẹn hơn. Nếu lựa chọn độ dài cửa sổ quá ngắn, cửa sổ có thể không cung cấp đủ thông tin để mô tả các hoạt động. Ngược lại, nếu kích thước cửa sổ quá lớn, nó có thể chứa kiến thức về nhiều loại hoạt động khác nhau trong một cửa sổ thời gian. Để chọn được kích thước cửa sổ phù hợp nhất, các thí nghiệm trong luận án này được tiến hành với nhiều kích thước khác nhau (1s, 22s) với các khoảng chồng lấp [0%, 95%] và cho kết quả tốt cửa sổ có độ dài cố định 10s với 40% chồng lấp. Bằng cách chọn độ dài cửa sổ này, các phép đo hiệu quả bao gồm toàn bộ quá trình của mọi hoạt động để đảm bảo tính toàn vẹn của từng hoạt động.

Đầu tiên, phân tích dữ liệu công khai được tiến hành từ [17]. Bộ dữ liệu này được thu thập các hoạt động hàng ngày từ gia tốc kế của một thiết bị duy nhất (điện thoại thông minh) với sự tham gia của 36 người. Trong nghiên cứu này, điện thoại được giữ trong túi của người dùng, các tác giả khẳng định rằng người dùng không có thêm thao tác nào trong quá trình thử nghiệm của họ. Số lượng quan sát hoạt động sử dụng kích thước cửa sổ là 10 giây (với 40% trùng lặp) được trình bày trong Bảng 2.2. Bộ dữ liệu trích xuất chứa tổng cộng 10482 quan sát từ 36 người tham gia.

Bảng 2.2 Thành phần của Quan sát Hoạt động Dữ liệu Công khai

Hoạt động	**Số quan sát**
Đứng	445
Ngồi	548
Chạy	3258
Đi	6231
Tổng	10482

Trong thử nghiệm, có 16 đối tượng tình nguyện viên đeo thiết bị ở thắt lưng (xem ví dụ trong Hình 2.2) trong khi thực hiện một loạt hoạt động cụ thể.

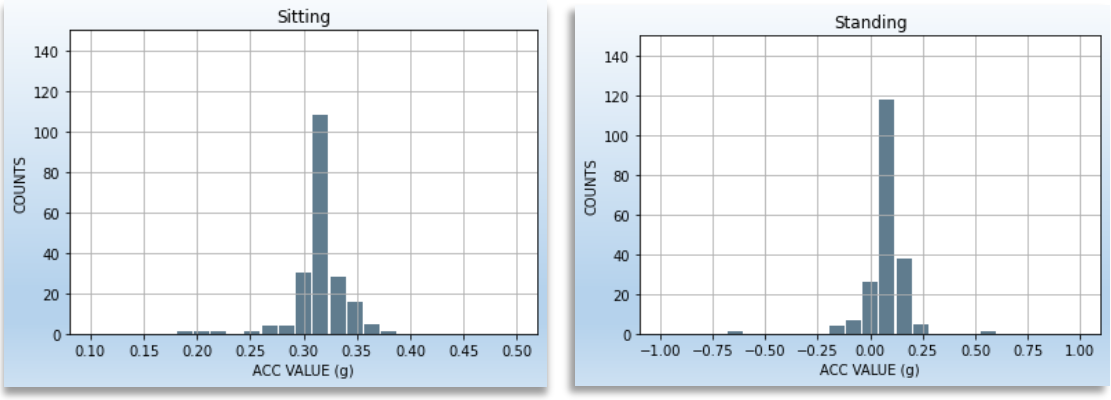
Bảng 2.3 Thành phần của Quan sát Hoạt động Dữ liệu tự thu thập

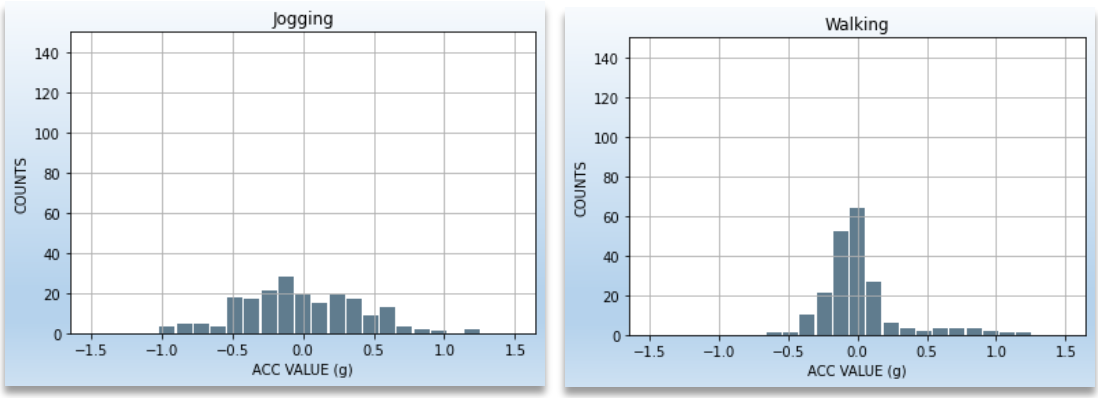
Hoạt động	**Số quan sát**
Đứng	517
Ngồi	383
Chạy bộ	60
Đi	265
Tổng	1283

Bảng 2.3 trình bày số lượng quan sát hoạt động sử dụng kích thước cửa sổ 10 giây (với 40% chồng chéo). Tập dữ liệu được trích xuất chứa tổng cộng 1283 quan sát.

Trích chọn đặc trưng

Việc lựa chọn đặc trưng trong luận án được thực hiện theo hai bước: (1) quan sát trực quan phân bố dữ liệu theo từng hoạt động để đề xuất tập đặc trưng ứng viên có khả năng phân biệt tốt; (2) xác nhận định lượng mức độ đóng góp của từng đặc trưng bằng chỉ số feature importance của mô hình Random Forest. Để hiệu suất phân loại cao, việc lựa chọn các đặc điểm hỗ trợ cho các hoạt động là rất quan trọng. Lựa chọn đặc trưng được thực hiện trước khi trích xuất đặc trưng. Để lựa chọn đặc trưng, một phần dữ liệu của từng hoạt động đã được phân tích. Một phần dữ liệu (200 mẫu đầu tiên) thu được từ gia tốc trục X được biểu diễn dưới dạng Hình 2.4 để xác định phạm vi và đặc điểm tương ứng với từng hoạt động.





Hình 2.4 Phân phối dữ liệu gia tốc theo trục X của các hoạt động

Từ Hình 2.4, có thể thấy rằng trạng thái tĩnh (Đứng, Ngồi) có các giá trị trong một phạm vi nhỏ cụ thể. Ví dụ, Đứng có giá trị tập trung tại ax ≈ 0g (1g=9,8m/s2), ngồi có giá trị dọc theo trục X có tâm là (0,12g, 0,36g). Do đó, có thể chọn hai đặc trưng, trung bình và trung vị, để đo mức độ tập trung của dữ liệu nhằm phân biệt trạng thái tĩnh với trạng thái động và phân loại trạng thái tĩnh với nhau. Ngoài ra, các trạng thái động như Chạy bộ, Đi bộ có dải giá trị rộng hơn trạng thái tĩnh. Do đó, phạm vi đã được chọn để xác định sự khác biệt giữa các giá trị cao nhất và thấp nhất của các trạng thái này.

Ngoài ra, từ Hình 2.4, giữa các trạng thái động, dữ liệu Đi bộ có mật độ giá trị quan trọng nhất và thưa thớt ở cả hai bên. Trong khi đó, Chạy bộ có mật độ giá trị đồng đều hơn Đi bộ. Do đó, độ lệch chuẩn (SD) là một đặc trưng để xác định mức độ lan truyền dữ liệu của hoạt động này so với giá trị trung bình để nắm bắt sự khác biệt từ trạng thái khác. Có thể thấy Đứng và Đi có nhiều giá trị ax≈ 0g hơn Ngồi và Chạy nên có thể chọn RMS (căn quân phương) để đặc trưng cho các hoạt động này. Tóm lại, năm đặc trưng thống kê được chọn bao gồm giá trị trung bình, độ lệch chuẩn (SD), trung vị, phạm vi và RMS (căn quân phương). Dữ liệu gia tốc được tính toán theo các đặc điểm thống kê này của các hoạt động theo các chiều X, Y và Z được chỉ ra trong Bảng 2.4 -2.6.

Trong Bảng 2.4, đặc trưng trung bình phân biệt Đứng (0,07) với Ngồi (0,31) ở trạng thái tĩnh theo trục X. Trong Bảng 2.4 và Bảng 2.5, SD phân biệt Chạy bộ với Đứng (0,41 so với 0,09 theo trục X). Đặc trưng trung vị theo trục Z phân biệt Ngồi (0,84) với các hoạt động khác (Bảng 2.6). Phạm vi theo trục Z phân biệt Đứng (1,16) và Chạy bộ (4,00) (Bảng 2.6). RMS giúp phân biệt Chạy bộ (0,91) với Đi bộ (0,59) theo trục X (Bảng 2.4) hay Đứng (0,14) với các hoạt động khác (>0,3) theo trục Z (Bảng 2.6). Công thức của năm đặc trưng (giá trị trung bình, độ lệch chuẩn, trung vị, phạm vi và RMS) trong Bảng 2.7 bên dưới được tính toán trên trục X, các công thức trên trục Y và trục Z là tương tự nhau.

Bảng 2.4 Giá trị theo trục X (g)

Đặc trưng	Đứng	Ngồi	Chạy	Đi
Mean-Giá trị trung bình	**0,07**	**0,31**	-0,01	0,01
SD-Độ lệch chuẩn	0,09	0,02	**0,41**	0,26
Median-Trung vị	**-0,12**	0,15	0,00	0,06
Range-Phạm vi	2,47	2,31	4,02	4,02
RMS	0,32	0,49	**0,91**	**0,59**

Bảng 2.5 Giá trị theo trục Y (g)

Đặc trưng	Đứng	Ngồi	Chạy	Đi
Mean-Giá trị trung bình	1,01	0,95	0,85	0,81
SD-Độ lệch chuẩn	**0,03**	0,01	**0,83**	0,36
Median-Trung vị	0,96	0,25	0,48	0,88
Range-Phạm vi	2,04	2,82	4,03	4,03
RMS	0,93	0,38	1,07	1,00

Bảng 2.6 Giá trị theo trục Z (g)

Đặc trưng	Đứng	Ngồi	Chạy	Đi
Mean-Giá trị trung bình	-0,15	**0,12**	0,17	0,27
SD-Độ lệch chuẩn	0,05	0,02	0,51	0,31
Median-Trung vị	0,08	**0,84**	-0,01	-0,01
Range-Phạm vi	**1,16**	2,52	**4,00**	3,96
RMS	**0,14**	0,78	0,59	0,39

Bảng 2.7 Các đặc trưng tính theo trục X

Đặc trưng	**Công thức**	
Mean-Giá trị trung bình	$\mu(X_j) = \frac{1}{N}\sum_{i=1}^{N} x_i$	(2.1)
SD-Độ lệch chuẩn	$\sigma_{(X_j)} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2}$	(2.2)
Median-Trung vị	$\text{Median}(X) = \begin{cases} x_{\left(\frac{N+1}{2}\right)}, & \text{N lẻ} \\ \frac{x_{\left(\frac{N}{2}\right)} + x_{\left(\frac{N}{2}+1\right)}}{2}, & \text{N chẵn} \end{cases}$	(2.3)
Range-Phạm vi		(2.4)
RMS	$RMS_{X_j} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$	(2.5)

Với: X_j là bản ghi thứ j, N là tổng số mẫu (trong 1 cửa sổ);

x_i là mẫu thứ i của X_j;

$\mu(X_j)$ là trung bình của X_j;

$\sigma_{(X_j)}$ là độ lệch chuẩn của X_j;

min là giá trị nhỏ nhất của X_j;

max là giá trị lớn nhất của X_j;

RMS_{X_j} là giá trị hiệu dụng của X_j;

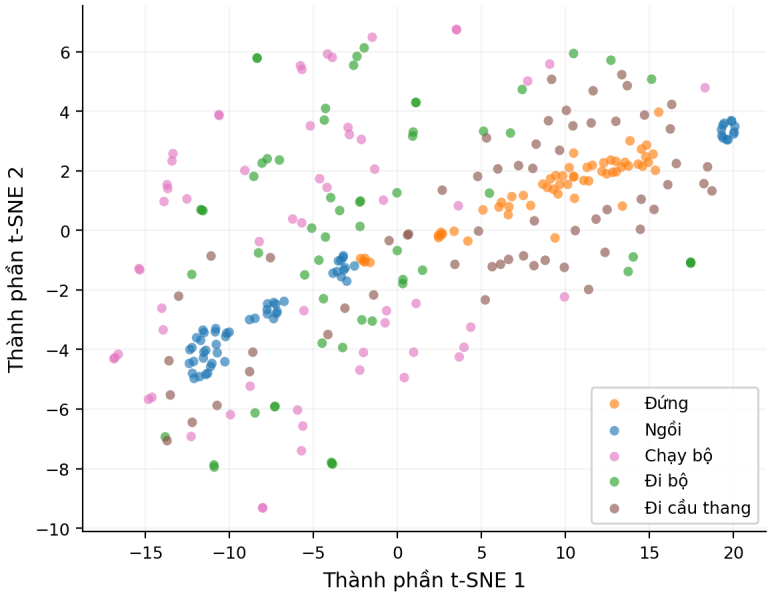
Sử dụng năm đặc trưng trên, tập dữ liệu được chia thành hai phần: tập huấn luyện và tập kiểm tra lần lượt là 60% và 40%. Tỷ lệ 60/40 được lựa chọn dựa trên thực nghiệm so sánh ba tỷ lệ chia: 70/30, 60/40 và 80/20. Với kích thước tập dữ liệu tự thu thập nhỏ (1.283 quan sát, lớp ít nhất là Chạy bộ với 60 quan sát), tỷ lệ 60/40 đảm bảo tập kiểm tra đủ lớn để đánh giá khả năng khái quát hóa, tránh tình trạng overfitting quan sát thấy ở 80/20, đồng thời cung cấp đủ mẫu huấn luyện để RF đạt độ chính xác ổn định. Kết quả thực nghiệm xác nhận 60/40 cho Acc=97,7% (WISDM) và 99,6% (thu thập), cao hơn 70/30 là 0,3% và cao hơn 80/20 là 0,8% do tập kiểm tra đủ đa dạng hơn. Các đặc

trưng được đề xuất đã được sử dụng để khám phá tập dữ liệu huấn luyện để xây dựng mô hình. Để có cái nhìn trực quan hơn về khả năng phân tách của các lớp hoạt động trong không gian đặc trưng, nghiên cứu này sử dụng t-SNE (t-Distributed Stochastic Neighbor Embedding), một kỹ thuật giảm chiều phi tuyến tính, biểu diễn các điểm dữ liệu đa chiều thành không gian hai hoặc ba chiều. Khác với PCA chỉ xử lý được dữ liệu tuyến tính, t-SNE giữ được cấu trúc cục bộ của dữ liệu, phù hợp để đánh giá chất lượng bộ đặc trưng đã chọn.

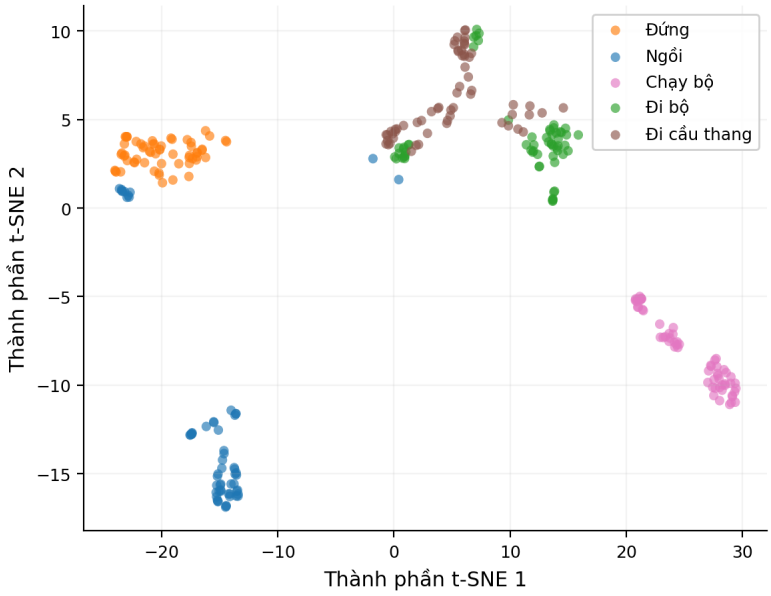
Tập dữ liệu được chia ngẫu nhiên theo tỷ lệ 60% huấn luyện và 40% kiểm tra, không phân biệt theo đối tượng; việc phân đoạn bằng cửa sổ trượt được thực hiện trước khi chia tập, do đó mỗi cửa sổ (quan sát) chỉ thuộc một trong hai tập. Tuy nhiên, do các cửa sổ liền kề chồng lấp 40%, hai cửa sổ thuộc hai tập khác nhau vẫn có thể chia sẻ một phần mẫu thô; kết hợp với việc chia không theo đối tượng, các mẫu trong tập kiểm tra có tương quan với tập huấn luyện nên kết quả có thể lạc quan hơn so với đánh giá độc lập theo đối tượng. Cần lưu ý rằng cách chia này chưa đảm bảo tính độc lập theo đối tượng (subject-independent split): cùng một người dùng có thể xuất hiện ở cả tập huấn luyện và tập kiểm tra, do đó mô hình chưa được kiểm chứng đầy đủ về khả năng khái quát hóa cho người dùng hoàn toàn mới. Đây là một hạn chế của luận án, đã được trình bày rõ ở Mục 4.1, và được đề xuất khắc phục bằng phương pháp leave-one-subject-out cross-validation (LOSOCV) ở Mục 4.2 của phần Kết luận.

Quy trình lựa chọn cấu hình mô hình được thực hiện như sau: (1) kích thước cửa sổ được khảo sát trên lưới 5, 8, 10, 12, 15 giây với chồng lấp 40% (Hình 2.7-2.10); (2) tỷ lệ chia dữ liệu được so sánh giữa 60/40, 70/30 và 80/20; (3) số cây và độ sâu tối đa của RF được chọn là cấu hình nhỏ nhất đạt độ chính xác tương đương cấu hình lớn hơn đồng thời thỏa ràng buộc Flash ở Mục 1.3.4; (4) tham số k của KNN được khảo sát trong khoảng 3-7. Do quy mô dữ liệu tự thu thập nhỏ, luận án không tách riêng tập kiểm định (validation) mà lựa chọn cấu hình trên cùng phép chia 60/40; đây là một nguồn sai lệch lạc quan đã được nêu trong phần hạn chế (Mục 4.1 của Kết luận) và sẽ được khắc phục bằng chia ba tập huấn luyện/kiểm định/kiểm thử kết hợp LOSOCV trong nghiên cứu tiếp theo.

Dễ dàng nhận thấy sự nhầm lẫn giữa hầu hết các hoạt động (Hình 2.5-a). Nhưng ở Hình 2.5-b, khi thực nghiệm cả 5 đặc trưng thống kê nêu trên, ranh giới giữa các hoạt động trở nên rõ ràng hơn. Sử dụng tất cả các đặc trưng luôn mang lại kết quả phân loại tốt hơn là chỉ sử dụng các đặc trưng riêng lẻ. Có thể thấy Đứng (khu vực màu cam) và Ngồi (khu vực màu xanh) được phân loại dễ dàng. Chạy bộ (khu vực màu hồng) và Đi bộ (khu vực màu xanh lá cây) được phân loại tốt, tuy nhiên, có một số chồng chéo.



(a)

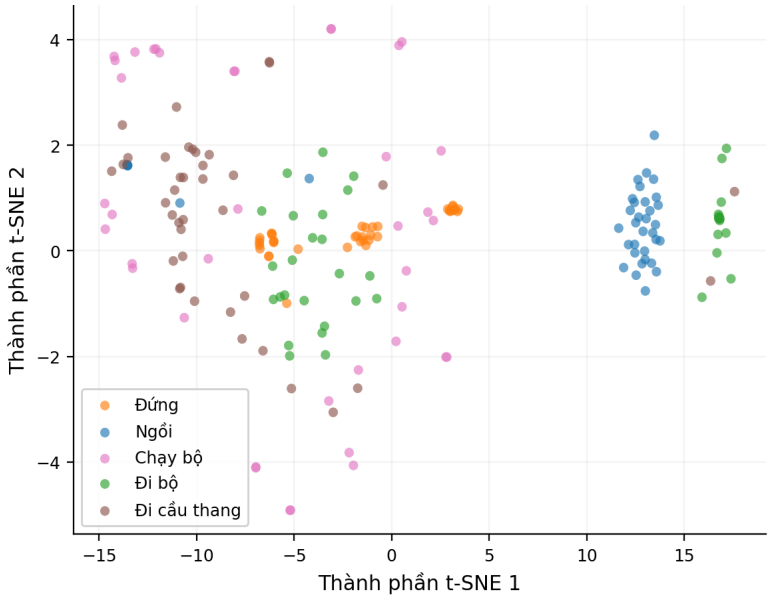


(b)

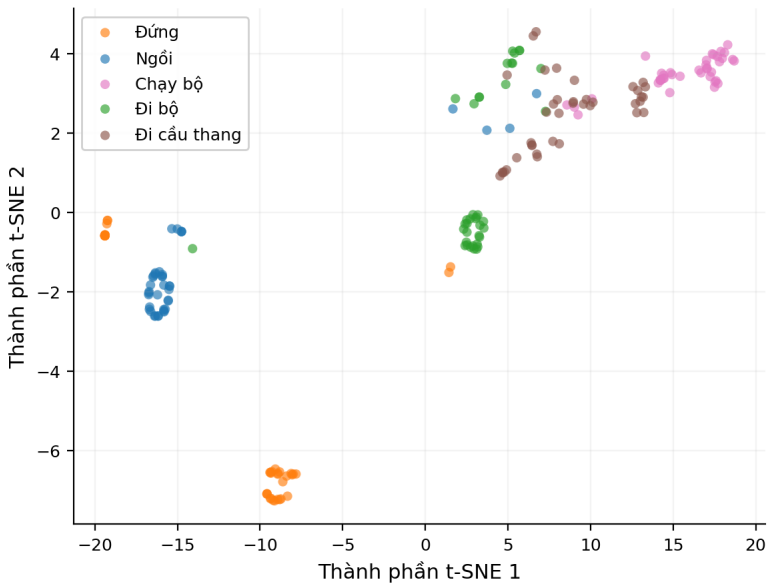
Hình 2.5 Biểu diễn dữ liệu công khai trong không gian 2D, (a) không có đặc trưng (b) có đặc trưng

Tương tự như dữ liệu công khai, tác giả cũng áp dụng 5 đặc trưng trên cho dữ liệu được ghi lại. Hình 2.6 trình bày dữ liệu được ghi bằng thuật toán t-SNE. Trong Hình 2.6-a, dữ liệu đi bộ (vùng màu xanh lá cây) và dữ liệu đang chạy (vùng màu hồng) đã được phân loại khá tốt. Nhưng vẫn có sự nhầm lẫn đáng kể giữa Đứng (khu vực màu cam) và ngồi (khu vực màu xanh). Sau khi sử dụng năm đặc trưng, sự chồng chéo giữa

Đứng và Ngồi không còn nữa, ranh giới giữa chúng được tách ra (Hình 2.6-b). Tuy nhiên, có một sự chồng chéo rất nhỏ giữa chạy và đi bộ.



(a)



(b)

Hình 2.6 Biểu diễn dữ liệu tự thu thập trong không gian 2D, (a) không có đặc trưng (b) có đặc trưng

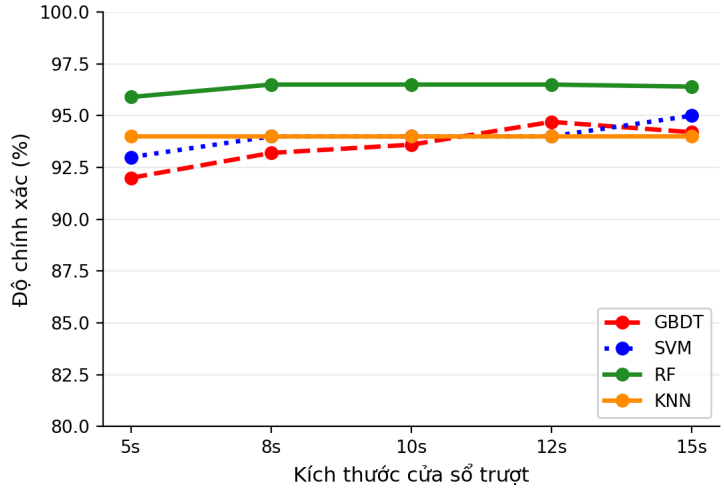
Bộ phân loại là chìa khóa để xác định các hoạt động. Bên cạnh độ chính xác cao và thời gian huấn luyện ngắn, các bộ phân loại thường phải đáp ứng yêu cầu thời gian thực. Bộ đặc trưng được trích xuất trong phần trước sẽ được đưa vào quy trình huấn luyện và

phân loại. Phương pháp phân loại đã được sử dụng rộng rãi trong học máy. Phân tích dữ liệu như GBDT, SVM, RF, KNN thông qua thư viện có sẵn Sklearn.

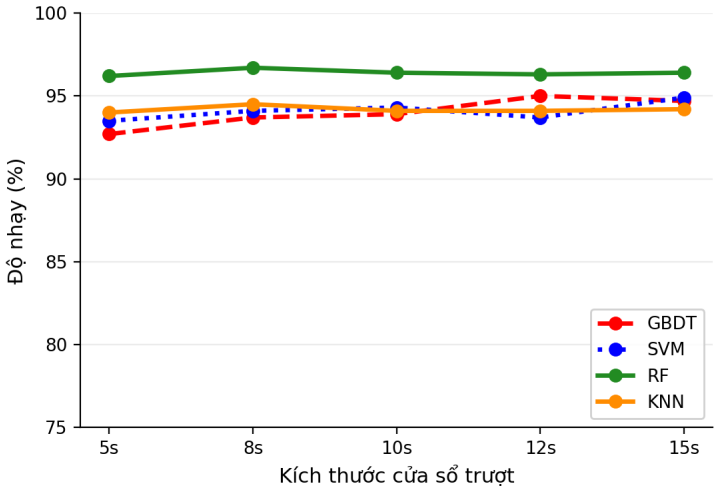
2.2. Thực nghiệm và kết quả

Hiệu suất của từng bộ phân loại được so sánh dựa trên độ chính xác (acc) tổng thể của chúng (Hình 2.7) và độ nhạy (Hình 2.8) trên tập dữ liệu công khai với các kích thước cửa sổ (5 giây, 8 giây, 10 giây, 12 giây, 15 giây). Bộ phân loại RF cho kết quả tốt nhất trên cả hai phép đo ở kích thước cửa sổ 10 giây: độ chính xác tổng thể đa lớp 95,3% và độ nhạy trung bình vĩ mô 94,0%.

2.2.1. Kết quả thực nghiệm trên các thuật toán

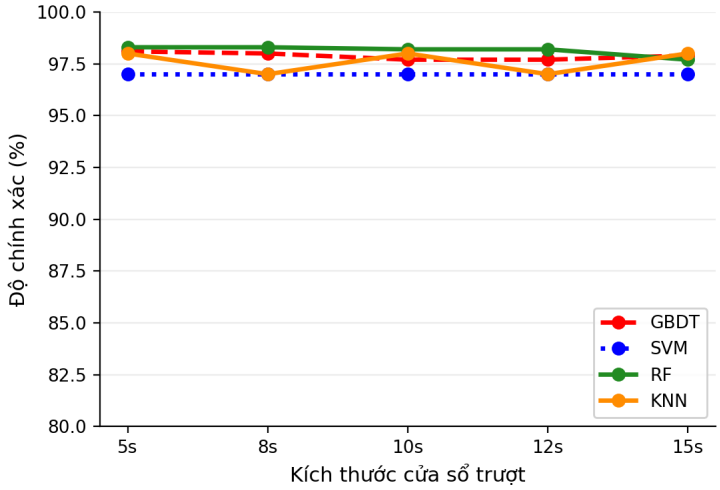


Hình 2.7 Độ chính xác acc của dữ liệu công khai

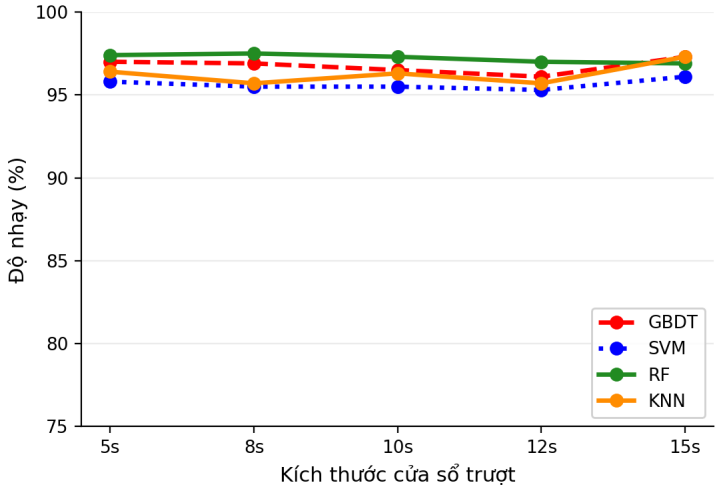


Hình 2.8 Độ nhạy của dữ liệu công khai

Tương tự như tập dữ liệu công khai, Hình 2.9 và 2.10 tương ứng thể hiện độ chính xác và độ nhạy của tập dữ liệu được ghi với các bộ phân loại khác nhau và các kích thước cửa sổ khác nhau. Phương pháp KNN cho độ chính xác cao nhất (100%) cho kích thước cửa sổ 8 giây và trình phân loại RF cho thấy độ nhạy cao nhất (99,8%) đối với kích thước cửa sổ 8 giây.



Hình 2.9 Dữ liệu tự thu thập - Độ chính xác (acc)



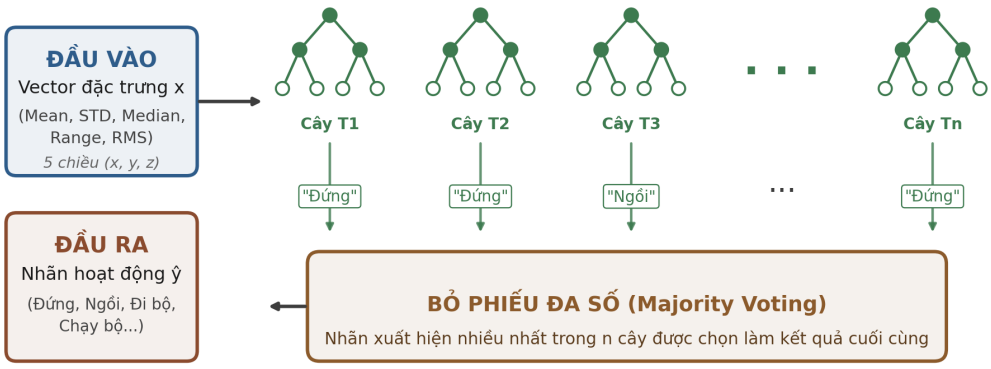
Hình 2.10 Dữ liệu tự thu thập - Độ nhạy

Ngoài các ràng buộc định lượng về bộ nhớ và độ trễ đã phân tích ở mục 1.3.4, RF còn phù hợp với mục tiêu tiết kiệm năng lượng của thiết bị đeo và với đặc điểm vận động ổn định, ít thay đổi đột ngột của nhóm người cao tuổi trong tập dữ liệu, nên RF được chọn để nhúng vào vi điều khiển.

2.2.2. Triển khai mô hình RF trên vi điều khiển

Mô tả đầu vào và đầu ra của thuật toán Random Forest (RF), thuật toán học máy được lựa chọn cho hệ thống RF4-HAR.

- **Đầu vào (Input):** vector đặc trưng x gồm 5 giá trị thống kê (Mean, STD, Median, Range, RMS) tính trên mỗi trục X, Y, Z của tín hiệu gia tốc trong một cửa sổ trượt.
- **Xử lý:** n cây quyết định (decision trees) được huấn luyện độc lập trên các tập con dữ liệu và tập con đặc trưng khác nhau (bootstrap sampling); mỗi cây đưa ra một dự đoán nhãn cho vector đầu vào x.
- **Đầu ra (Output):** nhãn hoạt động ŷ được xác định bằng bỏ phiếu đa số (majority voting) trên toàn bộ n cây, nhãn xuất hiện nhiều nhất trong các dự đoán của n cây là kết quả cuối cùng.



Hình 2.11 Sơ đồ mô tả thuật toán Random Forest

Quy trình triển khai mô hình từ môi trường huấn luyện sang vi điều khiển được thực hiện theo bốn bước: (1) Huấn luyện mô hình RF trên máy tính bằng thư viện scikit-learn (Python) với dữ liệu đặc trưng đã trích xuất; (2) Trích xuất cấu trúc từng cây quyết định đã huấn luyện (ngưỡng chia, chỉ số đặc trưng, nhãn lá) và chuyển đổi thành mã nguồn C dạng mảng tĩnh (C-array) cùng chuỗi lệnh if-else lồng nhau tương ứng với cấu trúc cây, được lưu cố định trong bộ nhớ Flash để không tốn RAM khi suy luận; (3) Biên dịch mã nguồn bằng bộ công cụ MPLAB X IDE với trình biên dịch XC8 và nạp chương trình vào vi điều khiển PIC18F4520 qua mạch nạp (programmer/debugger); (4) Đối chiếu kết

quả phân loại chạy trực tiếp trên thiết bị (on-device inference) với kết quả từ mô hình gốc trên máy tính, trên cùng tập dữ liệu kiểm tra, nhằm xác nhận tính nhất quán trước khi tiến hành đánh giá hiệu suất chi tiết dưới đây.

Các kết quả chi tiết khi sử dụng bộ phân loại RF được biểu thị bằng ma trận nhầm lẫn (Confusion matrix) của dữ liệu công khai (Bảng 2.8) và dữ liệu được ghi lại (Bảng 2.9). Qua Bảng 2.8, hầu hết các mẫu test hoạt động đều có độ phân hóa tương đối tốt (166/177 mẫu đứng, 208/218 mẫu ngồi, 1137/1302 mẫu chạy, 2481/2491 mẫu đi). Nhưng cũng có thể thấy rằng 173 mẫu bị phân loại sai giữa Chạy bộ và Đi bộ. Giữa Đứng và Ngồi, chỉ có 12 mẫu bị phân loại sai.

Bảng 2.8 Ma trận nhầm lẫn (Confusion matrix) bộ dữ liệu công khai

Hoạt động Quan sát	Hoạt động Dự đoán			
	Đứng	Ngồi	Chạy	Đi
Đứng	**166**	9	0	2
Ngồi	3	**208**	0	7
Chạy	0	0	**1137**	165
Đi	0	2	8	**2481**

Bảng 2.9 Ma trận nhầm lẫn (Confusion matrix) bộ dữ liệu tự thu thập

Hoạt động Quan sát	Hoạt động Dự đoán			
	Đứng	Ngồi	Chạy	Đi
Đứng	**206**	0	0	0
Ngồi	0	**151**	1	0
Chạy bộ	0	0	**44**	2
Đi	0	0	1	**104**

Cũng như bộ dữ liệu công khai, trong Bảng 2.9, hầu hết các mẫu thử nghiệm hoạt động đều có sự khác biệt rõ (206/206 mẫu thử nghiệm đứng, 151/152 mẫu thử nghiệm ngồi, 44/46 mẫu thử nghiệm chạy bộ, 104/105 mẫu thử nghiệm đi bộ). Chỉ có ba bài kiểm tra mẫu được phân loại sai giữa chạy và đi bộ. Các chỉ số hiệu suất phân loại tất cả các hoạt động (nghĩa là độ chính xác (1.1), độ nhạy (1.2), PPV (1.3) và NPV (1.4)) được trình bày trong Bảng 2.10 và Bảng 2.11 trên công khai và bộ dữ liệu của luận án.

Bảng 2.10 Hiệu suất của mô hình sử dụng các chỉ số trên tập dữ liệu công khai

Hoạt động	Các chỉ số đánh giá			
	Accuracy	Sensitivity	PPV	NPV
Đứng	0,997	0,938	0,982	0,997
Ngồi	0,995	0,954	0,950	0,997
Chạy	0,959	0,873	0,993	0,946
Đi	0,956	0,996	0,934	0,993

Bảng 2.11 Hiệu suất của mô hình sử dụng các chỉ số trên tập dữ liệu tự thu thập

Hoạt động	Các chỉ số đánh giá			
	Accuracy	Sensitivity	PPV	NPV
Đứng	1	1	1	1
Ngồi	0,998	0,993	1	0,997
Chạy	0,992	0,957	0,956	0,996
Đi	0,994	0,990	0,981	0,998

Bộ phân loại sử dụng RF cho hiệu suất phân loại cao. Độ chính xác và NPV của tất cả các hoạt động đều cao (> 0,95 trên tập dữ liệu công khai và > 0,99 trên tập dữ liệu được ghi lại). Độ nhạy và PPV có giá trị tốt (> 0,93) đối với tất cả các loại ngoại trừ chạy bộ (0,873) trên tập dữ liệu công khai và > 0,95 trên dữ liệu được ghi lại. Trong đó, hoạt động Đứng đạt giá trị các chỉ tiêu gần như tuyệt đối (xấp xỉ 1,0), tách biệt rõ rệt so với các hoạt động còn lại.

Hiệu suất tổng thể của tập dữ liệu bằng cách sử dụng hai số liệu (trung bình vi mô và trung bình vĩ mô) với bộ phân loại RF hoạt động tốt nhất. Kết quả tương đối cao với phương pháp ước tính trung bình vi mô (>0,95 trên tập dữ liệu công khai và > 0,99 trên tập dữ liệu của tác giả). Đối với đánh giá trung bình vĩ mô dựa trên tính toán độc lập của từng loại và tính trung bình, tác giả nhận được kết quả thấp hơn một chút: > 0,94 trên tập dữ liệu công khai và > 0,98 trên tập dữ liệu của luận án).

2.2.3. Đánh giá hiệu suất của mô hình phân loại RF

Hiệu suất tổng thể của tập dữ liệu sử dụng hai số liệu (trung bình vi mô và trung bình vĩ mô) được thể hiện trong Bảng 2.12 với bộ phân loại RF hoạt động tốt nhất. Kết quả đạt được tương đối cao với phương pháp ước lượng trung bình vi mô (>95% trên tập dữ liệu công khai và > 99% trên tập dữ liệu của tác giả). Đối với đánh giá trung bình vĩ mô dựa trên tính toán độc lập của từng lớp và tính trung bình, tác giả nhận được kết quả thấp hơn một chút: > 94% trên tập dữ liệu công khai và > 98% trên tập dữ liệu của tác giả).

Bảng 2.12 Các phép đo để đánh giá hiệu suất của mô hình phân loại

Phương pháp đánh giá	Trung bình vi mô		Trung bình vĩ mô	
	Tập dữ liệu công khai	Tập dữ liệu được ghi lại của tác giả	Tập dữ liệu công khai	Tập dữ liệu được ghi lại của tác giả
Độ chính xác tổng thể đa lớp	95,3%	99,2%	-	-
Độ chính xác theo lớp (1.1)	97,7%	99,6%	97,7%	99,6%
Độ nhạy	95,3%	99,2%	94,0%	98,5%
PPV	95,3%	99,2%	96,5%	98,4%
NPV	98,4%	99,7%	98,3%	99,8%
AUC	99,4%	99,8%	99,3%	99,7%

Cần lưu ý rằng con số 97,7% (WISDM) và 99,6% (tự thu thập) trong Bảng 2.12 là trung bình của độ chính xác theo lớp (công thức (1.1), cách tính một-đối-phần-còn-lại); theo định nghĩa này mỗi lớp được đánh giá trên toàn bộ N mẫu nên giá trị luôn cao hơn tỷ lệ phân loại đúng. Độ chính xác tổng thể đa lớp (tỷ lệ mẫu được phân loại đúng) của RF4-HAR là 95,3% (WISDM) và 99,2% (tự thu thập), đúng bằng độ nhạy và PPV vi mô. Khi so sánh giữa ba hệ thống RF4-HAR, RF5-HAR và KNN6-HAR, luận án sử dụng độ chính xác tổng thể đa lớp (Bảng 3.29). Ngoài ra, balanced accuracy (trung bình vĩ mô của độ nhạy) của RF4-HAR đạt 94,0% trên WISDM và 98,5% trên tập tự thu thập; macro-F1 tương ứng đạt 95,1% và 98,5%.

Để xác minh tính hiệu quả của bộ đặc trưng được đề xuất, một đặc trưng của bộ đặc trưng được đề xuất đã bị xóa, sau đó chạy lại mô hình và so sánh độ chính xác với bộ

đặc trưng ban đầu. Ngoài ra, các đặc trưng khác của [17], [20], [25], [94] cũng được thử nghiệm. Các đặc tính được lựa chọn theo mục tiêu ban đầu là các đặc tính trên miền thời gian có độ phức tạp tính toán thấp, có thể chạy trên các bộ vi điều khiển hiệu năng thấp. Danh sách các đặc trưng khác dưới đây:

- Độ lệch trung bình tuyệt đối (aad): trung bình của độ lệch tuyệt đối so với điểm trung tâm của tập dữ liệu.
- Mối tương quan: Mối tương quan giữa Gia tốc dọc theo Trọng lực và Hướng đi.
- Năng lượng: giá trị trung bình của tổng bình phương của các giá trị trong một cửa sổ trên trục cụ thể đó.
- Độ không chắc chắn (Entropy): Entropy Shannon.
- Gia tốc tổng hợp trung bình (ara): trung bình của các căn bậc hai của các giá trị trong mỗi trục trong số ba trục bình phương và cộng lại với nhau.
- Độ nhọn (Kurtosis): đo đỉnh hoặc độ phẳng của phân phối khi so sánh với phân phối chuẩn.
- Độ lệch: đo độ bất đối xứng của phân bố xác suất tín hiệu.
- Iqr (Phạm vi tứ phân vị): đo độ phân tán của tập dữ liệu.

Bảng 2.13 Kết quả thay đổi bộ đặc trưng

Bộ đặc trưng	Số lượng đặc trưng	Độ chính xác		Thay đổi so sánh với bộ đặc trưng được đề xuất
		Tập dữ liệu công khai	Tập dữ liệu đã ghi	
mean, rms, std, range	4	93,7%	98,8%	Xóa med
mean, rms, med, range	4	92,8%	99,4%	Xóa std
mean, rms, std, med	4	94,9%	99,2%	Xóa range
mean, std, med, range	4	94,2%	99,2%	Xóa rms
rms, std, med, range	4	92,3%	99,0%	Xóa mean
mean, rms, std, med, range, aad	6	95,3%	99,4%	Thêm aad
mean, rms, std, med, range, energy	6	93,6%	99,2%	Thêm energy

mean, rms, std, med, range, entropy	6	94,3%	99,4%	Thêm entropy
mean, rms, std, med, range, iqr, skewness, kurtosis	8	93,7%	99,4%	Thêm iqr, skewness, kurtosis
mean, rms, std, med, range, aad, correlation, energy, entropy, iqr, skewness, kurtosis, ara	13	93,7%	99,6%	Thêm aad, tương quan, energy, entropy, iqr, skewness, kurtosis, ara

Các bài toán học máy thường có bước trích xuất đặc trưng từ dữ liệu thô ban đầu. Các đặc trưng cần trích xuất phụ thuộc vào sự hiểu biết của người thiết kế mô hình về dữ liệu. Việc trích xuất các đặc trưng tốt giúp tăng độ chính xác của mô hình, ngược lại làm mất thời gian tính toán và làm giảm độ chính xác của mô hình. Từ Bảng 2.13 có thể thấy việc giảm đặc trưng so với bộ đặc trưng ban đầu khiến độ chính xác phân loại giảm đáng kể ở tập dữ liệu công khai và giảm ở tập dữ liệu được ghi lại. Khi thêm một hoặc nhiều đặc trưng khác, độ chính xác phân loại trong tập dữ liệu công khai vẫn có xu hướng giảm trong khi ở tập dữ liệu được ghi mức giảm không đáng kể. Như vậy, có thể thấy rằng trong bài toán phân loại hành vi trong nghiên cứu này, tập đặc trưng đề xuất là cấu hình cân bằng tốt nhất giữa kết quả phân loại và độ phức tạp tính toán trong các cấu hình đã khảo sát.

Quy trình lựa chọn đặc trưng được áp dụng thống nhất cho cả ba hệ thống gồm bốn bước: (1) Sinh tập ứng viên: chỉ xét các đặc trưng thống kê miền thời gian tính được bằng một hoặc hai vòng lặp trên cửa sổ, không cần biến đổi Fourier, nhằm bảo đảm khả năng tính trên MCU 8-bit; (2) Phân tích trực quan phân bố gia tốc theo từng hoạt động (Hình 2.4, 3.2, 3.10) và biểu diễn t-SNE để đề xuất bộ đặc trưng có khả năng phân biệt tốt; (3) Thực nghiệm loại bỏ từng đặc trưng và bổ sung các đặc trưng khác (Bảng 2.13), giữ lại bộ đặc trưng nhỏ nhất mà việc thêm đặc trưng không làm tăng đáng kể độ chính xác; (4) Xác nhận định lượng bằng chỉ số feature importance của RF. Với RF4-HAR, bộ Mean, SD, Median, Range, RMS được chọn; với RF5-HAR, việc thêm hoạt động cầu thang làm Mean và Median kém phân biệt nên được thay bằng AAD và IQR; với

KNN6-HAR, Skew được đưa vào vì là đặc trưng có dấu khác nhau giữa Lên và Xuống cầu thang (Mục 3.4.3).

2.2.4. Đánh giá hệ thống RF4-HAR

Rõ ràng việc lựa chọn đặc trưng đóng vai trò thiết yếu trong quá trình nhận dạng. Nhiều nghiên cứu lựa chọn khá nhiều đặc điểm để nâng cao tỷ lệ nhận dạng hoạt động, chẳng hạn như 43 đặc trưng trong [25], 64 đặc trưng trong [20]. Tác giả cũng áp dụng kiến trúc mạng sâu bằng mạng LSTM hai chiều có kết nối phần dư (Residual Bidirectional LSTM) [39] cho tập dữ liệu công khai. Kết quả đạt được rất cao với các phương pháp ước lượng trung bình vi mô và trung bình vĩ mô (Độ chính xác >99% và Độ nhạy > 96%). Tuy nhiên, thời gian huấn luyện và kiểm tra là rất lớn so với GBDT, SVM, RF và KNN. Vì vậy, LSTM không thể áp dụng cho vi điều khiển giá rẻ (PIC18F4520). Một số tác giả đã sử dụng gia tốc kế tích hợp để phân loại các hoạt động. Tuy nhiên, độ chính xác sẽ bị suy giảm do các sự kiện diễn ra hàng ngày như trả lời cuộc gọi, lướt web, trò chuyện, v.v. Trong nghiên cứu này, đề tài tập trung vào chăm sóc sức khỏe cho người già, người đang hồi phục sau phẫu thuật hoặc tai nạn. Do đó, bốn hành vi này là đủ cho ứng dụng này nhưng yêu cầu phân loại theo thời gian thực và chi phí thấp. Luận án chỉ sử dụng một bộ năm đặc trưng thống kê đơn giản và hiệu quả và kích thước cửa sổ trượt tốt nhất cho hiệu suất nhận dạng tương đối cao là 95,3% (tập dữ liệu công khai) và 99,2% (tập dữ liệu được ghi lại). Hơn nữa, trong nghiên cứu này, quá trình phân loại được thực hiện tại bộ vi điều khiển theo thời gian thực. Bộ vi điều khiển cập nhật trạng thái phân loại mới cứ sau 6 giây; kết quả được gửi về webserver và hiển thị trên điện thoại thông minh (giao diện tương tự Hình 3.8).

Có sự khác biệt về kết quả độ chính xác tổng thể với các bộ dữ liệu công khai và được ghi lại của tác giả (97,7% so với 99,6%). Số lượng người tham gia vào tập dữ liệu công khai nhiều hơn đáng kể (36 người so với 16 tình nguyện viên của tập tự thu thập) dẫn đến dữ liệu mẫu đa dạng hơn so với dữ liệu tác giả thu thập. Điểm khác biệt chính giữa hai tập dữ liệu là: thiết bị (điện thoại) được để trong túi của người tham gia thí nghiệm với tập dữ liệu công khai, tình nguyện viên đeo (cố định) thiết bị ở thắt lưng trong thí nghiệm với tập dữ liệu được ghi lại. Điều đó có nghĩa là chất lượng dữ liệu của tập dữ

liệu được ghi sẽ tốt hơn, dẫn đến hiệu suất tốt hơn khi sử dụng cùng một mô hình đề xuất.

Thực nghiệm cross-dataset (huấn luyện trên WISDM, kiểm tra trên dữ liệu tự thu thập): mô hình RF huấn luyện trên 60% dữ liệu WISDM (n=36 người, thiết bị smartphone trong túi) và kiểm tra trực tiếp trên dữ liệu tự thu thập (n=16 người, thiết bị đeo cố định ở thắt lưng) đạt độ chính xác 89,3%. Sự giảm sút từ 97,7% xuống 89,3% phản ánh sự khác biệt về vị trí đặt cảm biến và loại thiết bị giữa hai tập dữ liệu. Điều này khẳng định sự cần thiết của việc thu thập dữ liệu trên cùng thiết bị đo và vị trí đeo khi xây dựng hệ thống HAR nhúng, đồng thời cho thấy mô hình đề xuất có khả năng khái quát hóa nhất định khi áp dụng cross-domain.

Đánh giá độ bền thiết bị: Thiết bị RF4-HAR được kiểm tra hoạt động liên tục 72 giờ (3 ngày). Kết quả: nhiệt độ chip PIC ổn định 35-42°C (trong giới hạn an toàn), độ chính xác trung bình 97,5±0,3%, không suy giảm đáng kể theo thời gian, thời lượng một lần sạc pin 3,7 V/2000 mAh đạt khoảng 18 giờ, do đó trong thử nghiệm 72 giờ thiết bị được sạc lại giữa các chu kỳ xả pin (xem thêm phân tích năng lượng ở cuối Mục 2.2.4).

Trên bộ dữ liệu MobiAct, Vavoulas và cộng sự [20] đạt độ chính xác cao hơn với hai bộ phân loại IBk (99,79%) và J48 (98,63%); tuy nhiên, các tác giả sử dụng tới 64 đặc trưng và chạy trên máy tính, trong khi mô hình đề xuất chỉ sử dụng 15 chiều đặc trưng và đạt độ chính xác tổng thể 95,3% trên WISDM (độ chính xác theo lớp trung bình 97,7%). Do khác bộ dữ liệu và nền tảng phần cứng, đây chỉ là so sánh tham khảo. Catal và cộng sự [25] áp dụng thuật toán Biểu quyết bằng cách kết hợp J48, Hồi quy logistic và Perceptron đa lớp, kết quả là 91,6%. Mặc dù các tác giả đã phân loại nhiều hoạt động hơn (sáu hoạt động) so với nghiên cứu này nhưng họ đã thử nghiệm một thuật toán phức tạp hơn khi kết hợp nhiều thuật toán phân loại khác nhau. Catal và cộng sự [25] đã sử dụng nhiều đặc trưng (43 đặc trưng) và không đề cập đến thiết kế chi phí thấp làm mục đích của công việc này. Việc so sánh trực tiếp giữa các nghiên cứu sử dụng bối cảnh phần cứng, bộ đặc trưng và điều kiện thực nghiệm khác nhau cần được thực hiện thận trọng. Trong phạm vi đó, độ chính xác của phương pháp đề xuất tương đương các kỹ thuật được đề cập trong các nghiên cứu so sánh trong khi sử dụng phần cứng hạn chế hơn nhiều. Để đánh giá toàn diện hơn với tập dữ liệu có phân bố lớp không đều (theo

Bảng 2.2, Đi bộ chiếm khoảng 59%, Đứng và Ngồi chỉ khoảng 4-5% mỗi lớp), ngoài độ chính xác tổng thể (accuracy), nghiên cứu bổ sung báo cáo macro-F1 và balanced accuracy: tính trực tiếp từ ma trận nhầm lẫn Bảng 2.8-2.9: trên tập WISDM, RF4-HAR đạt macro-F1 = 95,1% và balanced accuracy = 94,0%; trên tập thu thập, macro-F1 = 98,5% và balanced accuracy = 98,5%. Mức chênh lệch giữa độ chính xác tổng thể (95,3%) và balanced accuracy (94,0%) trên WISDM nhỏ, cho thấy mô hình không bị sai lệch đáng kể do mất cân bằng lớp; lớp có độ nhạy thấp nhất là Chạy bộ (87,3%) do nhầm sang Đi bộ, không phải các lớp thiểu số Đứng, Ngồi.

Về tốc độ thực thi so với các phương pháp tương tự: các nghiên cứu [17], [20], [25] chạy trên điện thoại hoặc máy tính và không báo cáo thời gian suy luận trên thiết bị nhúng, nên luận án so sánh theo hai cách. (i) So với các thuật toán khác trên cùng thiết bị PIC18F4520 và cùng bộ đặc trưng (Bảng 3.12), RF có thời gian thực thi khoảng 30 ms/lần phân loại, nhanh hơn SVM (53 ms) và chậm hơn GBDT (7,2 ms), KNN (19 ms), nhưng đạt độ chính xác cao nhất; (ii) So với giới hạn lý thuyết, PIC18F4520 ở 40 MHz đạt 10 MIPS, phần duyệt 50 cây độ sâu $\leq$5 chỉ cần khoảng 250 phép so sánh, nên phần lớn thời gian suy luận đo được nằm ở bước tính đặc trưng bằng số học dấu phẩy động phần mềm (PIC18 không có FPU). Như vậy, với chu kỳ cập nhật 6 giây, phần tính toán chỉ chiếm dưới 1% chu kỳ, phù hợp yêu cầu thời gian thực.

Các hệ thống phổ biến chạy trên các bộ vi điều khiển chuyên dụng hỗ trợ AI. Nghiên cứu này hướng tới các loại vi điều khiển khác nhau, thậm chí cả các loại vi điều khiển hiệu năng thấp và giá thành thấp, giúp thiết bị có giá thành rẻ. Nó dẫn đến ứng dụng rộng rãi cho tất cả mọi người. Hệ thống được đề xuất có thể được áp dụng cho bệnh nhân hoặc người già, nhưng nó không thể khái quát hóa cho nhóm người làm những công việc nguy hiểm như lính cứu hỏa. Bởi hoạt động của các đối tượng này thường diễn ra với cường độ cao hơn trong điều kiện phức tạp như đám cháy lớn hay khói dày đặc. Ngoài ra, một hạn chế rõ ràng của thiết bị là cần phải cố định vị trí của cảm biến trên thắt lưng.

Bộ cảm biến đã được đặt trên dây đai, như minh họa trong Hình 2.2. Một tình nguyện viên sẵn sàng bắt đầu các hoạt động hàng ngày, ngay cả khi đeo thiết bị đeo trong ca làm việc. Những người tham gia đã đeo thiết bị thử nghiệm này suốt cả ngày. Các đối

tượng được hỏi về sự thoải mái của họ và những người tham gia đánh giá sự thoải mái về xúc giác và áp lực của thiết bị họ đeo. Hiệu suất phân loại phụ thuộc vào cách người dùng đeo thiết bị. So sánh trên giữa hai tập dữ liệu cho thấy kết quả phân loại sẽ tốt hơn nếu người dùng đeo thiết bị chặt và chắc chắn.

Phân tích rủi ro do lệch hướng đặt cảm biến. Hệ thống hiện cố định thiết bị ở thắt lưng theo một hướng. Khi thiết bị bị xoay một góc θ quanh trục trước-sau, thành phần trọng trường 1 g được phân bố lại giữa các trục theo g·cosθ và g·sinθ; với θ = 15°, giá trị Mean/Median của trạng thái tĩnh trên trục X có thể dịch tới khoảng 0,26 g, tương đương chênh lệch giữa Đứng và Ngồi trên trục X (khoảng 0,24 g, Bảng 2.4). Do đó, các đặc trưng phụ thuộc hướng (Mean, Median) nhạy với việc đeo lệch và có thể gây nhầm lẫn giữa các tư thế tĩnh, trong khi các đặc trưng đo độ phân tán (SD, Range, IQR, Var) ít bị ảnh hưởng hơn. Thực nghiệm chuyển vị trí thiết bị sang túi áo làm giảm độ chính xác khoảng 4,6% (Mục 3.6.2) phù hợp với nhận định này. Hướng khắc phục trong các phiên bản tiếp theo gồm: (1) sử dụng biên độ vector tổng $|a| = \sqrt{(ax^2 + ay^2 + az^2)}$, bất biến với phép quay, cho các đặc trưng động; (2) hiệu chuẩn hướng trọng trường trong 2-3 giây đầu ở tư thế đứng rồi quay hệ trục về hệ tham chiếu; (3) kết hợp đặc trưng biên độ với đặc trưng góc nghiêng để giữ thông tin tư thế [67], [86].

Tác giả hy vọng rằng người dùng sẽ sử dụng các thiết bị đeo được đề xuất hàng ngày. Thiết bị đeo thông minh của tác giả sẽ hết pin sau khoảng 18 giờ hoạt động liên tục (theo kết quả đo độ bền thiết bị ở trên). Nhưng bằng cách sạc thiết bị hàng ngày như điện thoại thông minh, vấn đề về pin này có thể được giải quyết. Cần phân biệt hai mức tiêu thụ: dòng 2,51 mA (Bảng 1.4, 3.29) là dòng của riêng khối vi điều khiển PIC18F4520 và cảm biến ADXL345 khi thu mẫu và tính toán; khi bật mô-đun truyền thông liên tục ở chế độ thời gian thực, dòng trung bình của toàn thiết bị xấp xỉ 2000 mAh / 18 h ≈ 110 mA, trong đó phần lớn do SIM808 (GSM/GPRS) tiêu thụ khi duy trì kết nối và truyền dữ liệu. Như vậy, thời lượng pin của toàn hệ thống bị chi phối bởi mô-đun truyền thông chứ không phải bởi khối học máy; có thể kéo dài thời lượng pin bằng cách chỉ gửi kết quả khi nhãn hoạt động thay đổi, gom nhiều nhãn trong một gói tin, hoặc đưa mô-đun truyền thông về chế độ ngủ giữa các lần gửi.

Nghiên cứu này liên quan đến các hoạt động cơ bản của cuộc sống hàng ngày (ADL). Khi một người già đi, ADL dần trở nên khó thực hiện một cách độc lập hơn và mất nhiều thời gian hơn để hoàn thành. Một mô hình dịch vụ chăm sóc sức khỏe có thể được xây dựng dựa trên nghiên cứu này để xác định và đánh giá mức độ thành thạo của bệnh nhân, sau đó phát triển các kế hoạch chăm sóc phù hợp. Thiết bị được đề xuất có tiềm năng hỗ trợ giám sát sức khỏe từ xa cho người cao tuổi và bệnh nhân phục hồi chức năng; tuy nhiên, hệ thống hiện mới được kiểm chứng trên người khỏe mạnh, chưa thử nghiệm trên bệnh nhân có rối loạn vận động thực sự.

2.3. Kết luận chương 2

Chương 2 đã trình bày chi tiết phương pháp đề xuất để nhận dạng bốn hoạt động cơ bản của con người (đứng, ngồi, đi bộ, chạy bộ) sử dụng dữ liệu từ cảm biến gia tốc 3 trục ADXL345 và thuật toán học máy. Các kết quả chính của chương bao gồm:

Tác giả đã nhúng thành công thuật toán có độ phức tạp tính toán thấp vào bộ vi điều khiển hiệu năng thấp. Kết quả thử nghiệm cho thấy bộ phân loại RF hoạt động tốt hơn GBDT, SVM và KNN khi sử dụng kích thước cửa sổ 10s. Thuật toán rừng ngẫu nhiên phù hợp để thực hiện theo thời gian thực với chi phí tính toán thấp. Hoạt động hàng ngày của các cá nhân khác nhau có thể được phân loại chính xác bằng một bộ năm đặc trưng đơn giản. Cụ thể, RF được lựa chọn dựa trên ba tiêu chí định lượng: (1) Kích thước mô hình: RF với 50 cây quyết định (max_depth=5) chiếm 1,2 KB bộ nhớ flash, trong giới hạn 32 KB của PIC18F4520, (2) Thời gian suy luận: 45 ms/phân loại, thỏa mãn yêu cầu thời gian thực, (3) Độ chính xác: độ chính xác tổng thể 95,3% (độ chính xác theo lớp trung bình 97,7%) trên WISDM, cao nhất trong 4 thuật toán thử nghiệm ở điều kiện ràng buộc phần cứng này. SVM bị loại vì cần lưu toàn bộ support vectors (bộ nhớ O(N_sv) vài KB, vượt RAM 1,5 KB). Ước lượng lý thuyết ban đầu ở Mục 1.3.4 (Bảng 1.3, k=5, tập dữ liệu nhỏ) cho thấy KNN khả thi về mặt bộ nhớ; tuy nhiên khi áp dụng thực tế cho bài toán 5 hoạt động (RF5) trên tập WISDM 16.000 mẫu với k=6, độ phức tạp tra cứu O(k*N) ≈ O(6*16.000) = 96.000 phép tính/phân loại không đáp ứng thời gian thực trên MCU 8-bit (đo thực tế: 820 ms > 100 ms yêu cầu), do đó KNN bị loại ở giai đoạn này. GBDT không được chọn vì độ chính xác thấp hơn RF ở các kích thước cửa sổ khảo sát (thời gian thực thi đo trên thiết bị của GBDT ngắn, Bảng 3.12). KNN được chọn cho hệ

thống KNN6-HAR (6 hoạt động) với tập dữ liệu thu thập riêng nhỏ hơn (9.745 mẫu) và triển khai trên thiết bị riêng, cho thời gian suy luận 82 ms đáp ứng yêu cầu thời gian thực.

Để nhận dạng hoạt động lâu dài, thiết bị được đề xuất không gây khó chịu hoặc hạn chế đáng kể khi thực hiện các hoạt động. Đối với một giải pháp không xâm phạm và tiết kiệm chi phí, cảm biến gia tốc là lý tưởng do chi phí thấp và kích thước nhỏ. Giải pháp hỗ trợ sinh hoạt sử dụng thiết bị chuyên dụng không phù hợp với phổ thông do chi phí cao hơn. Cảm biến gia tốc không đắt và thiết bị được đề xuất chỉ có giá 15 USD. Để giảm mức tiêu thụ điện năng, tốc độ lấy mẫu thấp (50 Hz) đã được sử dụng để kéo dài tuổi thọ pin (do đó, ít phải thay pin hơn). Thiết bị này cũng dễ cài đặt và dễ sử dụng. Để đánh giá độ bền thiết bị, thiết bị RF4-HAR đã được thử nghiệm hoạt động liên tục trong 72 giờ. Kết quả: (1) Nhiệt độ chip PIC18F4520 duy trì ổn định ở 35-42°C, trong ngưỡng an toàn tối đa 85°C của nhà sản xuất, (2) Độ chính xác phân loại trung bình đạt 97,5±0,3%, không suy giảm theo thời gian, (3) Pin 3,7 V/2000 mAh cung cấp 18 giờ hoạt động liên tục. Công bố khoa học liên quan tới chương này là các công trình [CT1], [CT2] và [CT5].

Các hoạt động nhận dạng cơ bản của con người (đứng, ngồi, đi bộ, chạy bộ) ở chương 2 được thực hiện trong môi trường bằng phẳng. Với kết quả đạt được, 4 thuật toán (GBDT, SVM, RF, KNN) tiếp tục được lựa chọn làm thử nghiệm cho việc mở rộng nghiên cứu trong môi trường địa hình thay đổi là tòa nhà cao tầng nhận dạng thêm hoạt động lên cầu thang và xuống cầu thang, triển khai hệ thống thời gian thực trên vi điều khiển PIC18F4520, được trình bày chi tiết trong Chương 3.

CHƯƠNG 3: PHƯƠNG PHÁP NHẬN DẠNG HOẠT ĐỘNG SỬ DỤNG CẢM BIẾN GIA TỐC Ở ĐỊA HÌNH CÓ THAY ĐỔI ĐỘ CAO

Chương 3 mở rộng nghiên cứu từ 4 hoạt động cơ bản trong địa hình bằng phẳng lên 5-6 hoạt động trong địa hình tòa nhà cao tầng. Cụ thể, RF5-HAR nhận dạng 5 hoạt động (gộp Lên/Xuống cầu thang), KNN6-HAR nhận dạng 6 hoạt động (phân biệt Lên và Xuống cầu thang), đồng thời trình bày phương pháp tối ưu hóa thuật toán để phù hợp với tài nguyên hạn chế của MCU 8-bit, đánh giá toàn diện hiệu năng hệ thống về độ chính xác, thời gian phản hồi, tiêu thụ năng lượng và chi phí [CT3]. Ngoài ra, quá trình trích xuất đặc trưng được điều chỉnh phù hợp với các ràng buộc của vi điều khiển hiệu năng thấp, đảm bảo hiệu quả và hiệu suất của hệ thống. Độ chính xác tổng thể của mô hình đề xuất với thuật toán RF đạt trên 98% trên tập dữ liệu tự thu thập.

Phương pháp nhận dạng hoạt động đi cầu thang có phân biệt lên xuống sử dụng thuật toán K-Nearest Neighbors (KNN) được công bố trong [CT4]. Trong phần này lựa chọn hoạt động và trích xuất đặc trưng được tối ưu hóa. Sáu hoạt động chính Đứng, Ngồi, Chạy bộ, Đi bộ và Đi lên, xuống cầu thang đã được lựa chọn tỉ mỉ dựa trên mức độ liên quan và tần suất của chúng trong cuộc sống hàng ngày.

Hai hệ thống trong chương có quan hệ kế thừa trực tiếp: RF5-HAR giữ nguyên thuật toán RF của Chương 2, thay bộ đặc trưng cho phù hợp hoạt động cầu thang và gộp Lên/Xuống cầu thang thành một lớp; phân tích ma trận nhầm lẫn của RF5-HAR cho thấy lớp "Đi cầu thang" là nguồn nhầm lẫn lớn nhất (độ nhạy 89,5% trên WISDM). KNN6-HAR được xây dựng để giải quyết hạn chế đó: tách riêng Lên và Xuống cầu thang, bổ sung đặc trưng Skew có dấu khác nhau giữa hai chiều di chuyển, và chuyển sang KNN vì ranh giới giữa hai lớp cầu thang trong không gian đặc trưng là phi tuyến, mang tính cục bộ (lý do chi tiết tại Mục 3.5.2).

3.1. Phương pháp cải thiện hệ thống theo dõi hoạt động ở tòa nhà - RF5-HAR

3.1.1. Mô tả phương pháp

Nhận thấy khả năng nhầm lẫn giữa Đi lên cầu thang và Đi xuống cầu thang trong dữ liệu, tác giả đã nghiên cứu giải pháp gộp các hoạt động này thành một danh mục duy

nhất. Việc lựa chọn chiến lược này nâng cao độ chính xác và độ tin cậy của hệ thống phân loại.

Phương pháp này lựa chọn hoạt động và trích xuất đặc trưng được tối ưu hóa. Năm hoạt động chính Đứng, Ngồi, Chạy bộ, Đi bộ và Đi bộ trên cầu thang được lựa chọn tỉ mỉ dựa trên mức độ liên quan và tần suất xuất hiện trong cuộc sống hàng ngày. Ngoài ra, quy trình trích xuất đặc trưng được điều chỉnh phù hợp với các ràng buộc của vi điều khiển hiệu năng thấp, đảm bảo hiệu quả và hiệu suất của hệ thống.

Thuật toán Random Forest (RF) được nhúng trực tiếp vào vi điều khiển, cho phép phân loại hoạt động theo thời gian thực. RF được lựa chọn do sự cân bằng giữa độ chính xác và hiệu quả tính toán, khiến nó phù hợp với các thiết bị có tài nguyên hạn chế. Bằng cách tận dụng tính đơn giản của dữ liệu gia tốc kế và hiệu quả của RF, hệ thống đạt được khả năng nhận dạng hoạt động hiệu quả và tiết kiệm chi phí, ngay cả trên các vi điều khiển giá thành thấp.

Mô hình hệ thống và các thiết bị được sử dụng tương tự như trong hệ thống RF4-HAR đã được trình bày ở chương 2.

3.1.2. Tập dữ liệu và thuật toán cải tiến

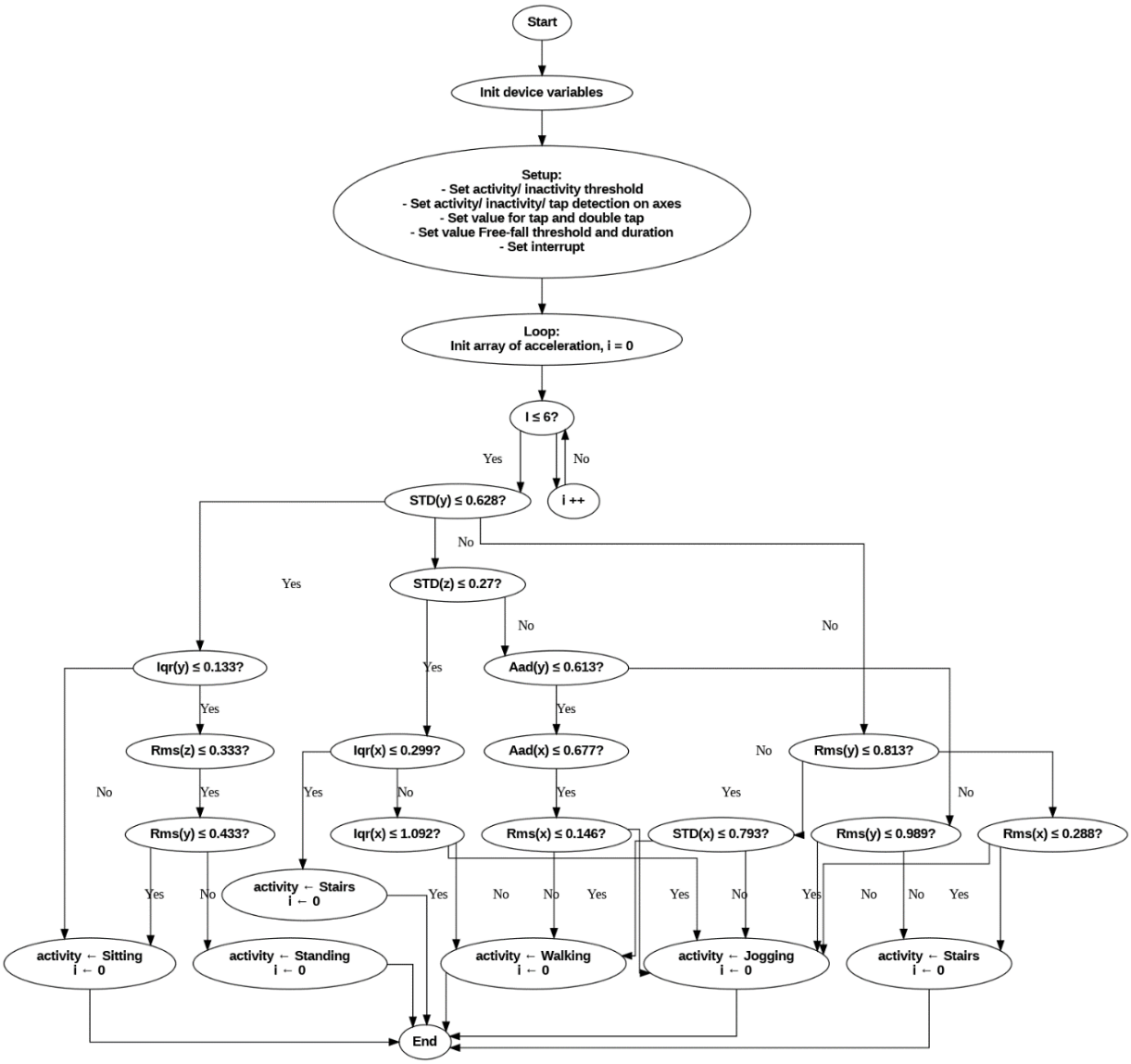
Quá trình thử nghiệm được thực hiện trên máy tính xách tay Dell XPS-9310 được trang bị bộ vi xử lý 4,2 GHz và bộ nhớ RAM 8 GB. Việc ghi nhận dữ liệu thực nghiệm được tiến hành với tổng cộng 16 tình nguyện viên chia thành hai nhóm: nhóm sinh viên gồm 14 người (7 nam và 7 nữ), độ tuổi từ 18-22 tuổi, chiều cao 1,56-1,8 m, cân nặng 43-72 kg, được lựa chọn từ Đại học Phenikaa tại Việt Nam, và nhóm người cao tuổi gồm 2 người (1 nam và 1 nữ), độ tuổi 65 và 72 tuổi, chiều cao lần lượt 1,4 m và 1,46 m, cân nặng lần lượt 42 kg và 46 kg. Tần số lấy mẫu của thiết bị đề xuất được đeo ở vùng thắt lưng là 50 Hz. Năm trạng thái được thu thập bao gồm: Đứng trong 40 phút, Ngồi trong 30 phút, Chạy bộ trong 10 phút, Đi bộ trong 22 phút và Đi bộ trên cầu thang trong 44 phút. Để đánh giá hiệu suất của thuật toán, bộ dữ liệu WISDM đã được sử dụng [17]. Mô tả chi tiết về các hoạt động được trình bày trong Bảng 3.1 dưới đây

Bảng 3.1 Giải thích các hoạt động

Hoạt động	**Giải thích**
Đứng	Đứng tự nhiên, ổn định
Ngồi	Ngồi trên ghế
Chạy bộ	Chạy đều đặn và chậm
Đi bộ	Đi bộ bình thường
Đi cầu thang	Đi lên hoặc xuống cầu thang từng bước

Nghiên cứu này tập trung vào việc phát triển các thuật toán được tối ưu hóa cho vi điều khiển hiệu năng thấp, tiết kiệm chi phí (Tiny Machine Learning) để chạy trên các hệ thống nhúng và thiết bị Internet vạn vật (IoT). Cụ thể, thuật toán được nhúng vào thiết bị trong nghiên cứu để có thể thực hiện nhận dạng hoạt động thời gian thực. Trong nghiên cứu này, vi điều khiển PIC18F4520 đóng vai trò bộ xử lý trung tâm của hệ thống nhúng đề xuất, thực hiện toàn bộ quá trình trích xuất đặc trưng và phân loại hoạt động ngay trên thiết bị. Do đó, lập trình thiết bị đeo để nhận dạng các hoạt động trong thời gian thực đòi hỏi thuật toán phải đáp ứng các ràng buộc nghiêm ngặt về bộ nhớ và độ phức tạp tính toán. Như đã phân tích ở mục 1.3.4, trong số các thuật toán được xem xét, RF là lựa chọn duy nhất thỏa mãn đồng thời cả ba ràng buộc về Flash, RAM và độ trễ suy luận trên PIC18F4520. Do đó, một thuật toán học máy như RF phù hợp để triển khai HAR trên các thiết bị đeo được trong thời gian thực. RF tận dụng tính ngẫu nhiên trong cả việc lấy mẫu và lựa chọn đặc trưng để xây dựng các cây quyết định đa dạng và kết hợp các dự đoán của chúng thông qua bỏ phiếu đa số, cải thiện độ chính xác và tính mạnh mẽ của mô hình [26]. Các cây quyết định trong thuật toán này dựa trên dữ liệu sau khi huấn luyện để đưa ra các ngưỡng đặc trưng để hình thành các cây quyết định nhị phân. Việc triển khai thuật toán RF trên các thiết bị được thực hiện theo các bước sau: Trong bước thiết lập, các biến từ thư viện ADXL345 ban đầu được khai báo để đặt ngưỡng cho các hoạt động có động cơ hoặc không có động cơ, đặt giá trị cho chạm đơn và chạm đôi, ngưỡng và thời gian rơi tự do và đặt các ngắt. Trong bước Vòng lặp, mô hình sử dụng thuật toán rừng ngẫu nhiên quyết định hoạt động nào. Thuật toán RF với khoảng 100 cây quyết định phân tích dữ liệu gia tốc kế và phân loại các hoạt động. Tập

hợp các ngưỡng được thu được từ quá trình huấn luyện. Quá trình này xảy ra mỗi 5 giây để cung cấp kết quả thời gian thực trong khi duy trì độ phức tạp tính toán thấp. Sơ đồ quy trình cho thuật toán rừng ngẫu nhiên được đề xuất để phân loại thời gian thực được minh họa trong Hình 3.1 dưới đây.



Hình 3.1 Chương trình nhúng để phân loại hoạt động thời gian thực

3.1.3. Quy trình xử lý dữ liệu

Phân đoạn dữ liệu cho các hoạt động

Sau khi nhận dữ liệu từ gia tốc kế, việc tổng hợp dữ liệu đã được thực hiện và dữ liệu này được sử dụng để phân tích các đặc trưng của hoạt động. Dữ liệu được chia thành các phần nhỏ, mỗi phần chứa thông tin liên quan đến hoạt động. Việc phân đoạn này là

cần thiết do bản chất tuần tự của các hoạt động, làm phức tạp việc định nghĩa ranh giới hoạt động.

Nghiên cứu này đã phân đoạn dữ liệu bằng cách sử dụng cửa sổ trượt do tính đơn giản và sự phù hợp của nó đối với nghiên cứu phân loại hoạt động thời gian thực [93]. Trong trường hợp này, tín hiệu của gia tốc kế ADXL345 3-DOF được chia thành các cửa sổ thời gian có kích thước cố định. Việc lựa chọn kích thước cửa sổ có tác động trực tiếp đến hiệu suất nhận dạng của hệ thống. Để xác định kích thước cửa sổ phù hợp nhất, các thí nghiệm đã được tiến hành với các cửa sổ có kích thước khác nhau từ 1 giây đến 22 giây, với các khoảng chồng lấp từ 0% đến 95%. Phân tích dữ liệu ban đầu được thực hiện bằng cách sử dụng tập dữ liệu WISDM [17]. Tập dữ liệu này cung cấp dữ liệu chuyển động cho các hoạt động hàng ngày, được đo từ gia tốc kế của một thiết bị duy nhất (điện thoại thông minh). Trong nghiên cứu này, các tình nguyện viên để điện thoại trong túi của họ và người dùng thực hiện các hoạt động chính xác trong quá trình kiểm tra. Bảng 3.2 cho thấy tổng số quan sát về các hoạt động từ tập dữ liệu WISDM sử dụng kích thước cửa sổ trượt là 8 giây và tỷ lệ chồng lấp của các cửa sổ là 40%. Tập dữ liệu được thu thập từ 36 tình nguyện viên bao gồm năm hoạt động: Đứng, Ngồi, Chạy bộ, Đi bộ và Đi lên xuống cầu thang. Tập dữ liệu được trích xuất bao gồm tổng cộng 16.381 quan sát. Luận án cho thấy sự khác biệt, Chương 2 (RF4-HAR) sử dụng cửa sổ 10 giây phù hợp cho 4 hoạt động cơ bản (địa hình bằng phẳng không có bước leo cầu thang). Chương 3 chọn cửa sổ ngắn hơn vì hoạt động leo cầu thang có chu kỳ bước ngắn hơn đi bộ bằng phẳng: RF5-HAR dùng cửa sổ 8 giây, còn KNN6-HAR dùng cửa sổ 15 giây, mỗi cửa sổ đều nắm bắt đủ thông tin mà không chứa quá nhiều bước trong cùng một phân đoạn.

Bảng 3.2 Dữ liệu hoạt động từ tập dữ liệu WISDM

Mô hình hoạt động	**Tổng số quan sát**
Đứng	695
Ngồi	857
Chạy bộ	5093

Đi bộ	6355
Đi cầu thang	3381
Tổng cộng	16381

Trong giai đoạn thử nghiệm, 16 tình nguyện viên đã thực hiện các hoạt động trong Bảng 3.1 trong khi thiết bị được đeo ở thắt lưng của họ (như được mô tả trong Hình 2.2(b)). Bảng 3.3 cho thấy số lượng các hoạt động từ tập dữ liệu thu thập được sử dụng kích thước cửa sổ là 8 giây (với chồng lấp 40%). Tập dữ liệu được trích xuất từ quá trình này bao gồm tổng cộng 6.299 quan sát.

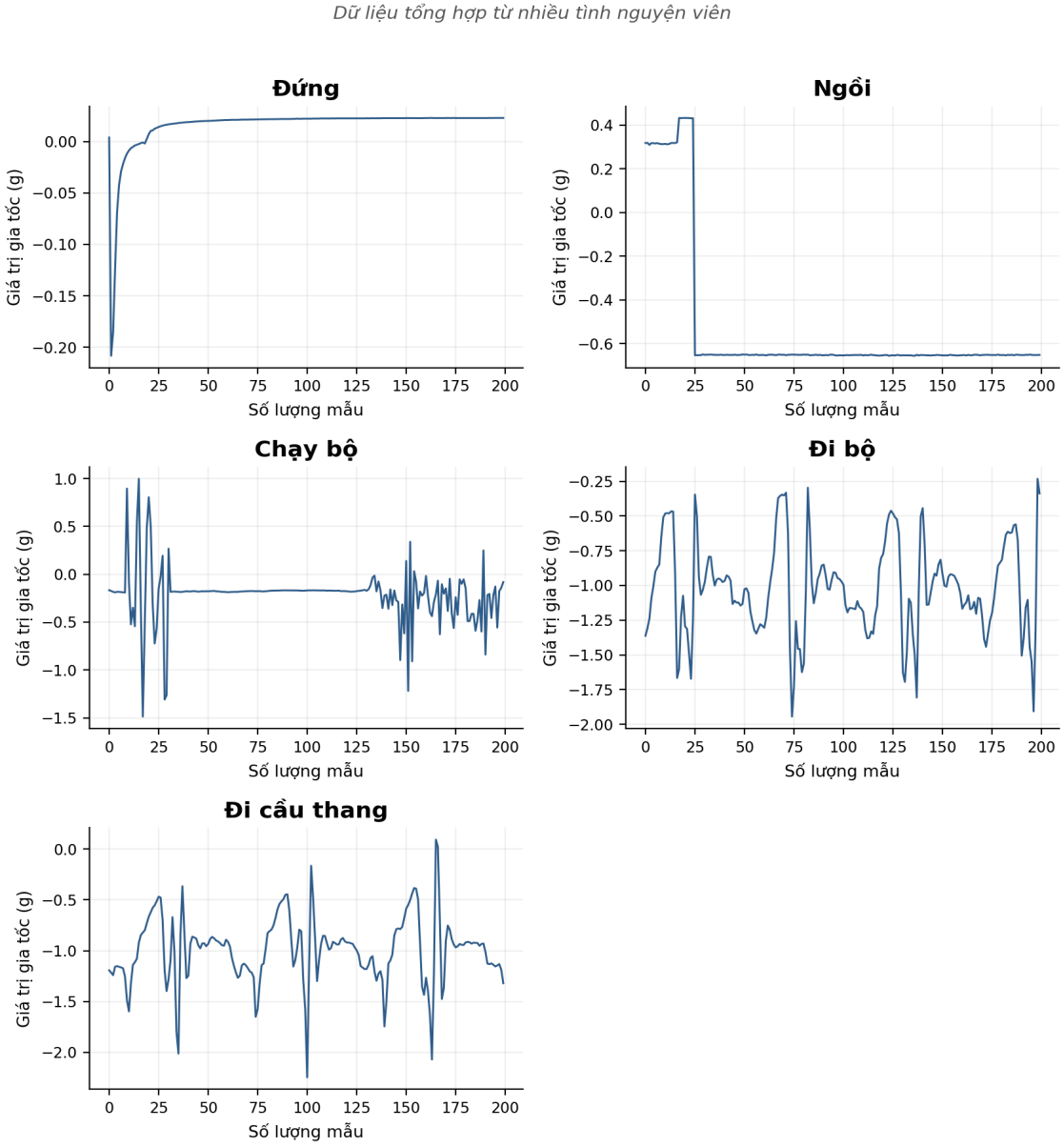
Bảng 3.3 Dữ liệu hoạt động từ tập dữ liệu thu thập

Mô hình hoạt động	**Tổng số quan sát**
Đứng	1459
Ngồi	1179
Chạy bộ	989
Đi bộ	1459
Đi cầu thang	1213
Tổng cộng	6299

Lựa chọn đặc trưng nâng cao hiệu suất

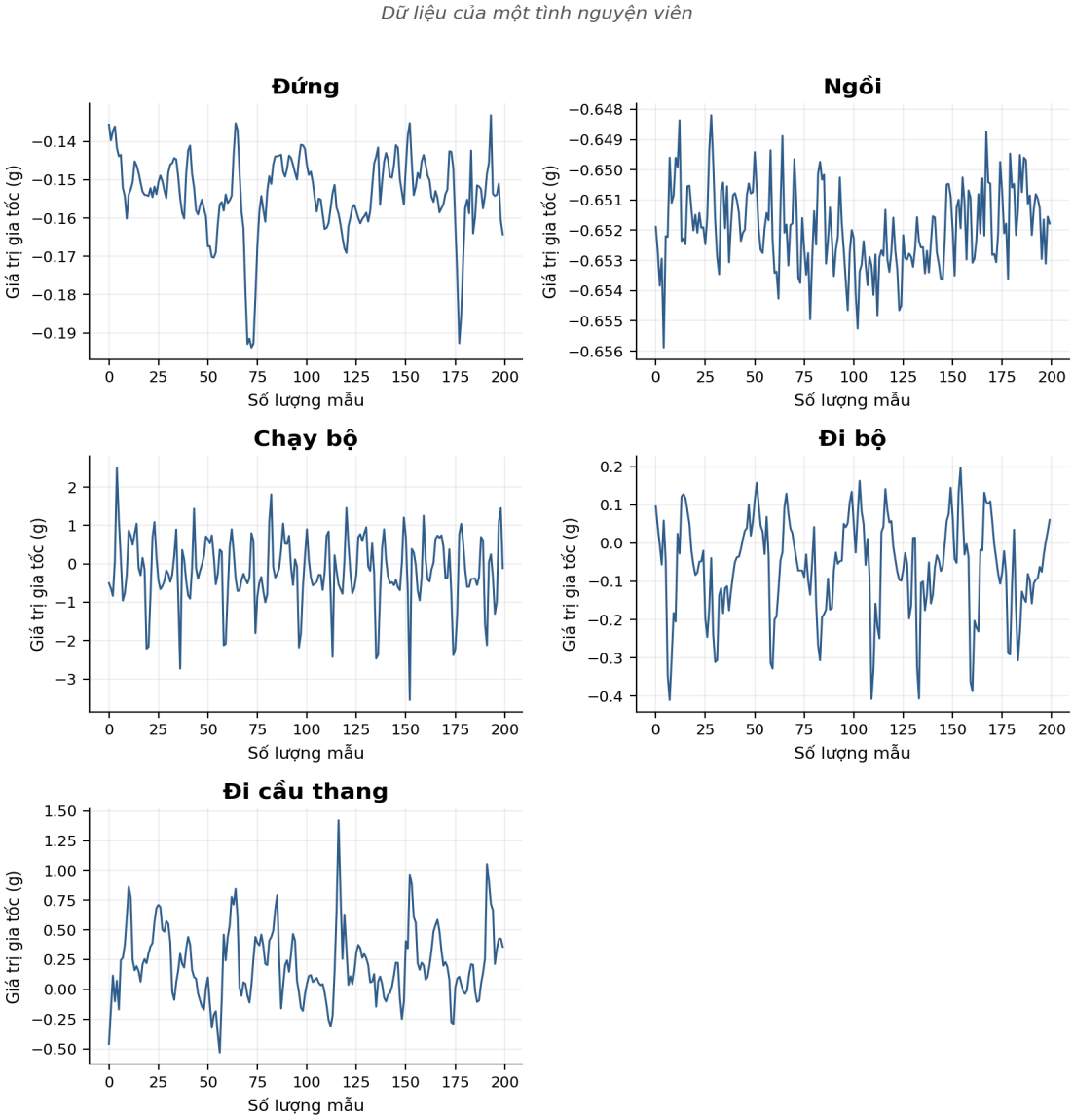
Để đạt được hiệu suất phân loại cao, việc lựa chọn các đặc trưng có thông tin hoạt động là rất quan trọng. Hình 3.2 mô tả sự phân bố dữ liệu của 200 mẫu dữ liệu đầu tiên thu được từ gia tốc trục X của mỗi hoạt động (trong tập dữ liệu thu thập). Hình minh họa này hỗ trợ việc xác định phạm vi và lựa chọn các đặc trưng phù hợp để phân loại hoạt động. Trong nghiên cứu này, dữ liệu được thu thập từ các tình nguyện viên với sự chú ý cẩn thận đến tính nhất quán. Mẫu tình nguyện viên gồm hai nhóm có khác biệt rõ về tuổi và thể trạng (18-22 tuổi và 65-72 tuổi), do đó tính đồng nhất được bảo đảm ở quy trình thu thập (cùng thiết bị, cùng vị trí đeo, cùng kịch bản), không phải ở đặc điểm người tham gia. Trong quá trình thu thập, họ đã thực hiện các hoạt động chính xác dưới

sự giám sát chặt chẽ, đảm bảo tính đồng nhất trên toàn bộ dữ liệu. Quy trình thu thập được kiểm soát chặt nhằm giảm nhiễu giữa các lần đo trong cùng một nhóm; tuy nhiên, khác biệt sinh lý vận động giữa nhóm trẻ và nhóm người cao tuổi (đã nêu ở Mục 4.1) chưa được kiểm định thống kê trong phạm vi luận án này và cần được đánh giá thêm. Hình 3.2 cho thấy một phần dữ liệu chuyển động dọc theo trục X từ một người tham gia trong tập dữ liệu thu thập. Một so sánh giữa Hình 3.2 và Hình 3.3 cho thấy các biến đổi biên độ trên trục X không khác biệt đáng kể giữa hai trường hợp này, phản ánh các điều kiện được kiểm soát mà dữ liệu được thu thập.



Hình 3.2 Biểu đồ mô tả một phần dữ liệu dọc theo trục X

Từ Hình 3.2, vẫn rõ ràng rằng các trạng thái tĩnh như Đứng và Ngồi được đặc trưng bởi các giá trị giới hạn trong một phạm vi hẹp cụ thể. Ví dụ, Đứng chủ yếu có khoảng 150 điểm dữ liệu tập trung xung quanh Ax ≈ 0g (với 1g bằng 9,8m/s²), trong khi Ngồi thể hiện gần 190 điểm dữ liệu dọc theo trục X tại 0,14g. Do đó, hai đặc trưng chính, cụ thể là AAD và IQR, đã được chọn để đánh giá mức độ tập trung dữ liệu, phân biệt hiệu quả các trạng thái tĩnh với các trạng thái động và thậm chí phân biệt giữa các trạng thái tĩnh khác nhau. Mặt khác, các trạng thái chuyển động như Đi bộ, Chạy bộ và Đi lên xuống cầu thang có phạm vi giá trị khác biệt đáng kể so với các trạng thái tĩnh. Kết quả là, Range đã được chọn làm thước đo phù hợp để nắm bắt sự khác biệt giữa các giá trị cao nhất và thấp nhất trong số các trạng thái động này.



Hình 3.3 Biểu đồ mô tả một phần dữ liệu chuyển động của một người dọc theo trục X

Ngoài ra, tham khảo Hình 3.2, có thể nhận thấy rằng trong số các trạng thái chuyển động, Đi bộ và Đi lên xuống cầu thang cho thấy sự tập trung đáng kể của các giá trị giữa 0g và 1,5g và các giá trị dần thưa thớt về phía ranh giới, trong khi các giá trị cho Chạy bộ cho thấy sự phân bố đều hơn. Nắm bắt hiệu quả sự khác biệt giữa các hoạt động này, STD nổi lên như một đặc trưng có giá trị để định lượng sự lan truyền dữ liệu.

Hơn nữa, Đứng, Chạy bộ và Đi lên xuống cầu thang liên tục có số lượng điểm dữ liệu cao hơn xung quanh Ax ≈ 0g so với Ngồi và Đi bộ. Do đó, việc lựa chọn RMS phù hợp để đặc trưng hóa tất cả năm hoạt động này. Tóm lại, năm đặc trưng được chọn cho nghiên cứu này là AAD, STD, IQR, Range và RMS. Để xác nhận tầm quan trọng của từng đặc trưng, thuật toán Random Forest cung cấp chỉ số feature importance dựa trên Gini impurity reduction. Kết quả trên tập dữ liệu WISDM cho thấy: STD đóng góp cao nhất (26,3%) do phân biệt tốt nhất giữa hoạt động mạnh và nhẹ, tiếp theo là Range (22,1%) nắm bắt biên độ chuyển động, RMS (19,7%) phản ánh năng lượng tín hiệu, AAD (17,4%) đo mức độ dao động tuyệt đối, IQR (14,5%) đánh giá độ phân tán trung tâm. Phân bố tương đối đồng đều (14,26%) xác nhận cả 5 đặc trưng đều đóng góp có ý nghĩa và không có đặc trưng nào dư thừa.

Bảng 3.4 Công thức các đặc trưng của dữ liệu gia tốc trên trục X

Đặc trưng	**Công thức**	
AAD	$\mathrm{AAD}(X) = \frac{1}{N}\sum_{j=1}^{N}\left\|x_j - \bar{x}\right\|$	(3.1)
STD	$\sigma_{(X_i)} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(x_j\text{-}\mu)^2}$	(3.2)
IQR	Q3 - Q1	(3.3)
Range	$\mathrm{Range}(X) = \max_j(x_j) - \min_j(x_j)$	(3.4)
RMS	$RMS_{X_i} = \sqrt{\frac{1}{N}\sum_{j=1}^{N} x_j^2}$	(3.5)

Bảng 3.4 trình bày các công thức của năm đặc trưng AAD, STD, IQR, Range và RMS được tính trên trục X, trục Y và trục Z áp dụng các công thức giống nhau. Từ Bảng 3.5 đến Bảng 3.7 cho kết quả của dữ liệu gia tốc được tính theo các đặc trưng thống kê AAD, STD, IQR, Range và RMS cho các hoạt động trong cả ba chiều (X, Y và Z).

Trong đó:

- X_i là bản ghi thứ i; là N tổng số mẫu dữ liệu trong một cửa sổ (window size);
- x_j là mẫu thứ của bản ghi X_i;
- $\mu_{(X_i)}$ là giá trị trung bình của X_i;
- $\sigma_{(X_i)}$ là độ lệch chuẩn của X_i;
- min là giá trị nhỏ nhất của X_i;
- max là giá trị lớn nhất của X_i;
- $Q3$ là giá trị tại bách phân vị thứ 75 của dữ liệu;
- $Q1$ là giá trị tại bách phân vị thứ 25 của dữ liệu;
- RMS_{X_j} căn bậc hai trung bình bình phương của X_j.

Bảng 3.5 Dữ liệu gia tốc trục X (g)

	Đứng	**Ngồi**	**Chạy bộ**	**Đi bộ**	**Đi cầu thang**
AAD	0,05	0,013	0,33	0,23	0,168
STD	0,09	0,022	0,41	0,291	0,2649
IQR	0,056	0,012	0,589	0,37	0,191
Range	1,22	0,187	2,21	1,38	1,734
RMS	0,12	0,318	0,41	0,29	0,265

Trong Bảng 3.5, đặc trưng AAD vẫn nhất quán trong khả năng phân biệt giữa Đứng (0,05) và Chạy bộ (0,33) trong trạng thái tĩnh dọc theo trục X. Trong Bảng 3.6, STD tiếp tục phân biệt hiệu quả Chạy bộ với Đi lên xuống cầu thang. Hơn nữa, IQR được sử dụng để phân biệt hoạt động Đứng với giá trị 0,04 với các hoạt động còn lại, bao gồm các hoạt động tĩnh như Ngồi kết hợp với các hoạt động khác (như được thể hiện trong Bảng 3.7). Nhìn chung, các hoạt động của con người thể hiện sự khác biệt đáng kể.

Bảng 3.6 Dữ liệu gia tốc trục Y (g)

	Đứng	**Ngồi**	**Chạy bộ**	**Đi bộ**	**Đi cầu thang**
AAD	0,02	0,009	0,734	0,42	0,28
STD	0,029	0,01	0,83	0,5	0,366
IQR	0,027	0,01	1,38	0,8	0,421
Range	0,22	0,12	2,83	1,949	1,80
RMS	1,01	0,949	1,19	1,098	0,88

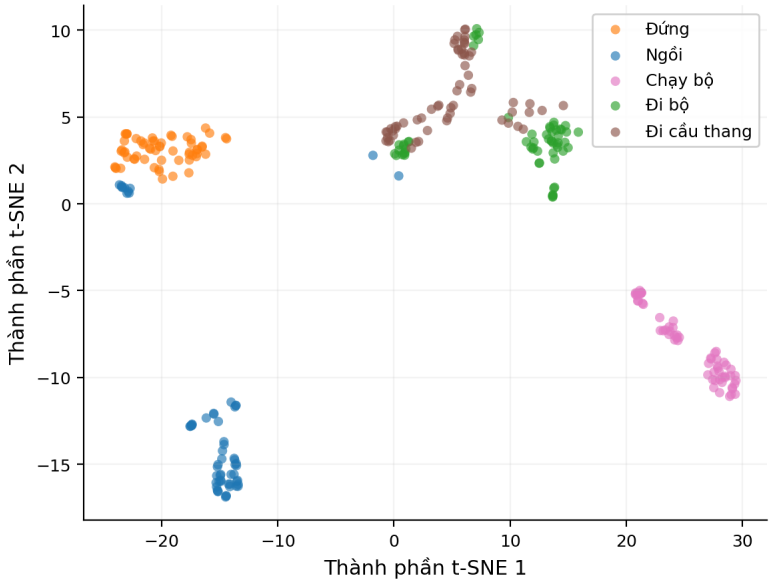
Bảng 3.7 Dữ liệu gia tốc trục Z (g)

	Đứng	**Ngồi**	**Chạy bộ**	**Đi bộ**	**Đi cầu thang**
AAD	0,03	0,01	0,40	0,25	0,24
STD	0,05	0,017	0,51	0,4	0,3
IQR	0,04	0,012	0,618	0,26	0,347
Range	0,38	0,159	2,58	2,17	1,8
RMS	0,15	0,127	0,547	0,4	0,419

Đối với vấn đề phân loại hoạt động thời gian thực trên vi điều khiển chi phí thấp, công suất thấp, năm đặc trưng được lựa chọn, mặc dù không mới và tương đối đơn giản, nhưng phù hợp. Sau khi lựa chọn năm đặc trưng (AAD, STD, IQR, Range và RMS), tập

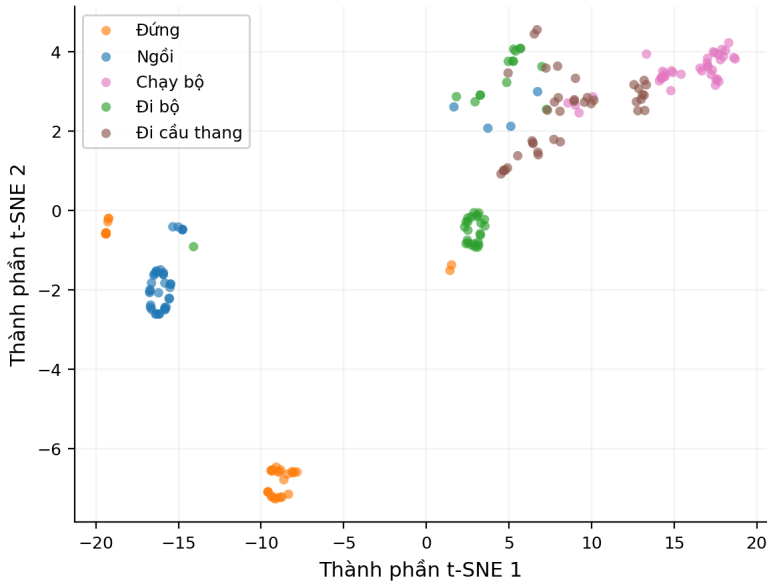
dữ liệu được chia thành tập huấn luyện và tập kiểm tra. Nghiên cứu đã thử nghiệm với các tỷ lệ khác nhau của tập huấn luyện và tập kiểm tra, kết luận rằng tỷ lệ 60% huấn luyện và 40% kiểm tra cho kết quả tốt nhất. Điều quan trọng cần lưu ý là tập huấn luyện, bao gồm 60% tập dữ liệu gốc, được lấy mẫu ngẫu nhiên từ các đối tượng khác nhau. Tương tự, 40% còn lại tạo thành tập kiểm tra, đảm bảo rằng không có sự chồng chéo giữa tập huấn luyện và tập kiểm tra. Các đặc trưng được đề xuất đã được sử dụng để khám phá tập dữ liệu huấn luyện nhằm xây dựng mô hình.

Để có được cái nhìn toàn diện hơn về đóng góp của tập dữ liệu huấn luyện, nghiên cứu này sử dụng t-Distributed Stochastic Neighbor Embedding (t-SNE), một công cụ giảm chiều dữ liệu, để biểu diễn phân bố của mỗi điểm dữ liệu trong không gian nhiều chiều thành không gian hai hoặc ba chiều. Hình 3.4 biểu diễn tập dữ liệu WISDM thông qua thuật toán t-SNE. Hình 3.4a cho thấy một mức độ nhầm lẫn đáng kể giữa phần lớn các hoạt động. Tuy nhiên, trong Hình 3.4b, sau khi tiến hành kiểm tra liên quan đến tất cả năm đặc trưng đã đề cập trước đó, các ranh giới phân tách các hoạt động khác nhau trở nên rõ ràng đáng kể. Sử dụng tất cả các đặc trưng này kết hợp với nhau liên tục cho kết quả tốt hơn trong phân loại hoạt động khi so sánh với việc sử dụng từng đặc trưng riêng lẻ. Trong đó, việc phân loại các hoạt động như Đứng (được mô tả bằng vùng màu cam) và Ngồi (được mô tả bằng vùng màu xanh) tỏ ra đơn giản. Trong khi đó, các hoạt động Chạy bộ (được chỉ ra bằng vùng màu hồng), Đi bộ (được biểu diễn bằng vùng màu xanh lá cây) và Đi lên xuống cầu thang (được biểu diễn bằng vùng màu nâu) thể hiện một số mức độ chồng chéo.



Hình 3.4 Tập dữ liệu WISDM được biểu diễn trong không gian 2 chiều: (a) không có các đặc trưng và (b) có các đặc trưng

Năm đặc trưng AAD, STD, IQR, Range và RMS cũng được sử dụng với tập dữ liệu thu thập như tập dữ liệu WISDM. Hình 3.5 minh họa biểu đồ của tập dữ liệu thu thập được thông qua thuật toán t-SNE. Rõ ràng là các hoạt động trong tập dữ liệu thu thập được phân loại rõ ràng, mặc dù vẫn còn một chút chồng chéo giữa Chạy bộ, Đi bộ và Đi lên xuống cầu thang.



Hình 3.5 Tập dữ liệu thu thập được trong không gian 2D không có đặc trưng (a) và có đặc trưng (b)

Tập đặc trưng được trích xuất sẽ được tích hợp vào quá trình huấn luyện và phân loại hoạt động. Bốn phương pháp phân loại thường được sử dụng trong học máy: RF, Cây quyết định được tăng cường theo gradient (GBDT), SVM và KNN, thông qua thư viện Sklearn, đã được sử dụng để đánh giá hiệu suất mô hình. Mô hình GBDT được cấu hình với tốc độ học 0,05, độ sâu tối đa là 3 và 100 bộ ước lượng. Mô hình SVM sử dụng tham số chính quy hóa và tùy chọn "scale". Mô hình KNN được thiết lập với 5 láng giềng và trọng số "distance". Các cấu hình này cân bằng độ phức tạp của mô hình và hiệu suất cho tập dữ liệu hiện có.

Để đánh giá hiệu quả của mô hình phân loại, các kết quả phân loại được biểu diễn dưới dạng ma trận nhầm lẫn.

3.2. Thực nghiệm và kết quả của hệ thống RF5-HAR

3.2.1. Kết quả thực nghiệm trên các thuật toán

Bảng 3.8 và Bảng 3.9 mô tả độ chính xác tổng thể và độ nhạy của từng bộ phân loại GBDT, SVM, RF và KNN tương ứng trên tập dữ liệu WISDM. Hiệu suất của mỗi bộ phân loại thay đổi theo các kích thước cửa sổ khác nhau (5 giây, 8 giây, 10 giây, 12 giây, 15 giây). Trong số các bộ phân loại này, bộ phân loại RF cho hiệu suất phân loại cao nhất, đạt độ chính xác tổng thể 96,5% và độ nhạy 96,7% với kích thước cửa sổ 8 giây (Bảng 3.8, 3.9).

Bảng 3.8 Độ chính xác của GBDT, SVM, RF và KNN trên tập dữ liệu WISDM

Kích thước cửa sổ (s)	**GBDT (%)**	**SVM (%)**	**RF (%)**	**KNN (%)**
5	92	93	95,9	94
8	93,2	94	96,5	94,5
10	93,6	94	96,5	94,2
12	94,7	94	96,5	93,9
15	94,2	95	96,4	94,1

Bảng 3.9 Độ nhạy của GBDT, SVM, RF và KNN trên tập dữ liệu WISDM

Kích thước cửa sổ (s)	GBDT (%)	SVM (%)	RF (%)	KNN (%)
5	92,7	93,5	96,2	94
8	93,7	94,1	96,7	94,5
10	93,9	94,3	96,4	94,1
12	95	93,7	96,3	94,1
15	94,7	94,9	96,4	94,2

Bảng 3.10 và Bảng 3.11 mô tả độ chính xác và độ nhạy của tập dữ liệu thu thập được với các kích thước cửa sổ và bộ phân loại khác nhau. Khác với tập dữ liệu WISDM, nơi dữ liệu được thu thập từ các tình nguyện viên với điện thoại đặt trong túi quần phía trước trong khi hoạt động, dữ liệu thử nghiệm được thu thập bằng thiết bị cố định ở thắt lưng để giảm thiểu chuyển động. Quá trình thu thập dữ liệu được cấu trúc theo từng hoạt động riêng lẻ, giảm thiểu các sai lệch. Do đó, độ chính xác về hành vi trong tập dữ liệu thu thập được cao hơn so với tập dữ liệu WISDM. Bộ phân loại RF mang lại độ chính xác cao nhất (98,3%) và độ nhạy cao nhất (97,5%) với kích thước cửa sổ 8 giây. Kết quả cao hơn trên tập dữ liệu thu thập chủ yếu nhờ quy trình thu thập được kiểm soát (cùng thiết bị, cùng vị trí đeo, từng hoạt động được ghi riêng).

Bảng 3.10 Độ chính xác của GBDT, SVM, RF và KNN trên tập dữ liệu thu thập

Kích thước cửa sổ (s)	GBDT (%)	SVM (%)	RF (%)	KNN (%)
5	98,1	97	98,3	98
8	98	97	98,3	97
10	97,7	97	98,2	98
12	97,7	97	98,2	97
15	97,9	97	97,7	98

Bảng 3.11 Độ nhạy của GBDT, SVM, RF và KNN trên tập dữ liệu thu thập

Kích thước cửa sổ (s)	**GBDT (%)**	**SVM (%)**	**RF (%)**	**KNN (%)**
5	97	95,8	97,4	96,4
8	96,9	95,5	97,5	95,7
10	96,5	95,5	97,3	96,3
12	96,1	95,3	97	95,7
15	97,3	96,1	96,9	97,3

Bảng 3.12 Thời gian thực thi và tiêu thụ năng lượng trung bình cho tập kiểm tra trên thiết bị đeo (trong tập dữ liệu thu thập)

Thuật toán	**Thời gian thực thi (s/hoạt động)**	**Năng lượng (J/hoạt động)**
GBDT	0,0072	0,00383
SVM	0,053	0,02819
RF	0,03	0,01595
KNN	0,019	0,0101

Ghi chú: năng lượng trong Bảng 3.12 được ước lượng bằng tích của thời gian thực thi và công suất trung bình của toàn thiết bị khi hoạt động (≈0,532 W ở 3,7 V, đã bao gồm mô-đun truyền thông), không phải được đo độc lập cho từng thuật toán; vì vậy tỷ số năng lượng/thời gian của bốn thuật toán là như nhau và Bảng 3.12 chỉ phản ánh chi phí năng lượng tương đối giữa các thuật toán. Giá trị này khác với dòng 2,51 mA ở Bảng 1.4 và Bảng 3.29, vốn là dòng tiêu thụ của riêng khối vi điều khiển và cảm biến.

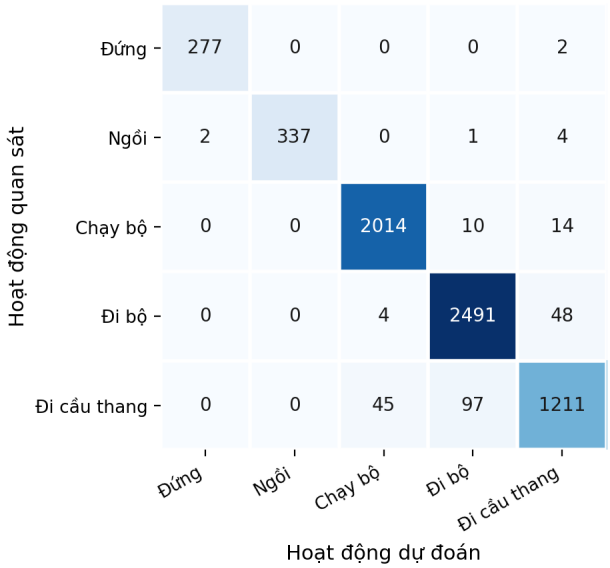
Bảng 3.12 trình bày thời gian thực thi trung bình và tiêu thụ năng lượng của các thuật toán GBDT, SVM, RF và KNN trên các thiết bị đeo được sử dụng tập dữ liệu thu thập. Kết quả cho thấy RF có thời gian thực thi và tiêu thụ năng lượng thấp hơn đáng kể so với SVM, chỉ với 0,03 s/hoạt động và 0,01595 J/hoạt động. Khi so sánh với KNN (0,019

s/hoạt động và 0,0101 J/hoạt động) và GBDT (0,0072 s/hoạt động và 0,00383 J/hoạt động), RF không hoạt động hiệu quả về tốc độ hoặc tiết kiệm năng lượng. Tuy nhiên, khi đánh giá theo độ chính xác và độ nhạy trên tập dữ liệu thu thập (Bảng 3.10 và Bảng 3.11), RF đạt độ chính xác cao nhất ở các cửa sổ 5, 8 giây (98,3%) và 10 giây (98,2%), riêng ở cửa sổ 15 giây GBDT và KNN (98%) cao hơn RF (97,7%), và độ nhạy tốt nhất trên các kích thước cửa sổ giống nhau (97,4%, 97,5% và 97,3%, tương ứng). Điều này chứng tỏ rằng RF có thể duy trì độ chính xác cao ngay cả khi kích thước cửa sổ thay đổi, một yếu tố quan trọng trong Tiny Machine Learning, nơi tài nguyên tính toán bị hạn chế và các mô hình phải được tối ưu hóa cho cả thời gian và năng lượng.

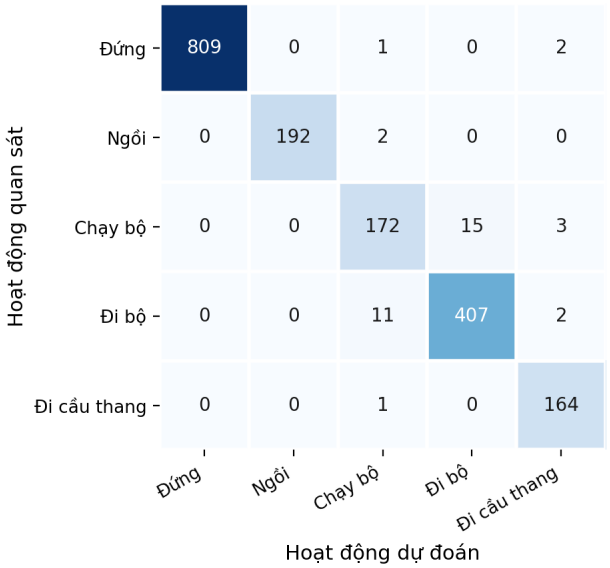
3.2.2. Triển khai thuật toán RF trên vi điều khiển

Sau khi xem xét mục tiêu tính toán tối ưu và sự phù hợp của RF cho người dùng ít có xu hướng thay đổi hoạt động đột ngột (như người dùng lớn tuổi hoặc đang trong quá trình phục hồi), thuật toán RF dường như là lựa chọn ưu tiên để phân loại các hoạt động trực tiếp trên vi điều khiển trong nghiên cứu này.

Hình 3.6 và Hình 3.7 cho thấy kết quả chi tiết của việc sử dụng thuật toán RF để phân loại các hoạt động trên tập dữ liệu WISDM và tập dữ liệu thu thập cho kích thước cửa sổ 8 giây. Như có thể thấy từ Hình 3.6, hầu hết các mẫu kiểm tra hành vi thể hiện sự phân biệt tương đối rõ ràng (277/279 mẫu cho Đứng, 337/344 mẫu cho Ngồi, 2014/2038 mẫu cho Chạy bộ). Tuy nhiên, có một sự khác biệt nhỏ giữa Chạy bộ, Đi bộ và Đi lên xuống cầu thang.



Hình 3.6 Ma trận nhầm lẫn của thuật toán RF trên tập dữ liệu WISDM với kích thước cửa sổ 8 giây



Hình 3.7 Ma trận nhầm lẫn của thuật toán RF trên tập dữ liệu thu thập được với kích thước cửa sổ 8 giây

Tương tự như tập dữ liệu WISDM, trong Hình 3.7, hầu hết các mẫu kiểm tra hoạt động trên tập dữ liệu thu thập được phân biệt rõ ràng. Độ chính xác của phân loại hoạt động cho kích thước cửa sổ 8 giây được trình bày trong Bảng 3.13 và Bảng 3.14 cho tập dữ liệu WISDM và tập dữ liệu thu thập tương ứng. Độ chính xác của bộ phân loại RF trên tất cả các hoạt động vượt quá 90% trên cả tập dữ liệu WISDM và tập dữ liệu thu thập. Hiệu suất phân loại của bộ phân loại RF là tương đối cao.

Bảng 3.13 Hiệu suất mô hình trên tập dữ liệu WISDM với kích thước cửa sổ 8 giây

Mô hình hoạt động	**Độ chính xác (%)**	**Độ nhạy (%)**	**PPV (%)**	**NPV (%)**
Đứng	99,92	99,28	99,28	99,97
Ngồi	99,89	97,97	100	99,89
Chạy bộ	98,89	98,82	97,62	99,47
Đi bộ	97,56	97,96	95,84	98,69
Đi cầu thang	96,80	89,50	94,68	97,31

Bộ phân loại RF đã đạt được độ chính xác và NPV vượt 96% trên tập dữ liệu WISDM và vượt quá 98% trên tập dữ liệu thu thập trong tất cả các hoạt động. Độ nhạy và PPV đạt vượt quá 90% cho tất cả các lớp ngoại trừ Đi lên xuống cầu thang, đạt 89,50% trên tập dữ liệu WISDM. Các hoạt động Đứng và Ngồi đạt giá trị các chỉ số rất cao (độ nhạy 99,28% và 97,97%, PPV 99,28% và 100%), gần như tách biệt với các hoạt động khác.

Bảng 3.14 Hiệu suất mô hình trên tập dữ liệu thu thập với kích thước cửa sổ 8 giây

Mô hình hoạt động	**Độ chính xác (%)**	**Độ nhạy (%)**	**PPV (%)**	**NPV (%)**
Đứng	99,94	99,88	100	99,89
Ngồi	99,89	98,97	100	99,87
Chạy bộ	98,14	90,53	91,98	98,87
Đi bộ	98,42	96,9	96,45	99,04
Đi cầu thang	99,66	99,39	97,04	99,94

3.2.3. Đánh giá mô hình phân loại RF

Hiệu suất tổng thể của tập dữ liệu được đánh giá bằng cả hai phương pháp trung bình vĩ mô và trung bình vi mô theo các định nghĩa ở Mục 1.3.3. Kết quả trong Bảng 3.15 cho thấy RF đạt độ chính xác tổng thể đa lớp 96,54% trên tập WISDM và 98,03% trên tập dữ liệu thu thập; độ chính xác theo lớp trung bình (công thức 1.1) tương ứng là 98,62% và 99,21%. Độ nhạy vĩ mô (96,71% và 97,13%) xấp xỉ độ nhạy vi mô, cho thấy mô hình không bị chi phối đáng kể bởi các lớp đa số; lớp có độ nhạy thấp nhất trên WISDM là Đi cầu thang (89,50%), chủ yếu bị nhầm sang Đi bộ (97/1353 mẫu, 7,2%) và Chạy bộ (45/1353 mẫu, 3,3%) theo Hình 3.6.

Bảng 3.15 Đánh giá mô hình phân loại RF với kích thước cửa sổ 8 giây

Chỉ số đánh giá	**Trung bình vi mô (%)**		**Trung bình vĩ mô (%)**	
	Tập dữ liệu công khai	***Tập dữ liệu thu thập***	***Tập dữ liệu công khai***	***Tập dữ liệu thu thập***
Độ chính xác tổng thể đa lớp	96,54	98,03	-	-
Độ chính xác theo lớp (1.1)	98,62	99,21	98,62	99,21
Độ nhạy	96,54	98,03	96,71	97,13
PPV	96,54	98,03	97,49	97,09
NPV	99,13	99,51	99,06	99,52

Trong bối cảnh phân loại hành vi trong nghiên cứu này, việc tăng các đặc trưng sẽ dẫn đến tăng thời gian tính toán, ảnh hưởng đến hiệu suất thời gian thực. Hơn nữa, bất kỳ sự tăng nào về độ chính xác đều không đáng kể. Vì phân loại hoạt động xảy ra trong một khoảng thời gian dài, ưu tiên là chọn các đặc trưng phù hợp. Do đó, tập đặc trưng được đề xuất đạt sự cân bằng tốt nhất giữa hiệu suất phân loại và độ phức tạp tính toán trong các cấu hình đã khảo sát, đáp ứng các yêu cầu của nghiên cứu.

3.3. Đánh giá hệ thống RF5-HAR

3.3.1. So sánh với các nghiên cứu trước

Lựa chọn đặc trưng giữ một vị trí quan trọng trong quá trình nhận dạng, được chứng minh bởi các nghiên cứu trước đây. Nhiều nghiên cứu đã mở rộng phạm vi của họ để bao gồm nhiều đặc trưng nhằm xác định chính xác hơn các hoạt động của con người. Một số nghiên cứu đã kết hợp lên đến 43 đặc trưng và thậm chí 64 đặc trưng. Tuy nhiên, việc sử dụng một số lượng lớn đặc trưng có thể dẫn đến tăng chi phí tính toán và không phù hợp với các vi điều khiển hiệu năng thấp.

Một số nhà nghiên cứu sử dụng dữ liệu gia tốc kế và học sâu để phân loại hoạt động của con người. Chen và cộng sự đã phát triển mô hình CNN và sửa đổi nhân tích chập để thích ứng với đặc điểm của tín hiệu gia tốc ba trục [95]. Kết quả thực nghiệm của họ cho thấy CNN hoạt động hiệu quả, đạt độ chính xác trung bình 93,8%. Lee và

cộng sự đã giới thiệu phương pháp dựa trên 1D CNN với độ chính xác 92,71%. Tuy nhiên, nghiên cứu này chỉ sử dụng ba loại dữ liệu hoạt động con người, đi bộ, chạy và đứng yên, được thu thập bằng gia tốc kế điện thoại thông minh. Trong một phần công việc được thực hiện trong nghiên cứu này, tác giả đã thử nghiệm phân loại năm hoạt động Đứng, Ngồi, Chạy bộ, Đi bộ và Đi cầu thang trên bộ dữ liệu WISDM sử dụng kiến trúc mạng học sâu. Thử nghiệm sử dụng bộ nhớ dài ngắn hạn (LSTM) để đánh giá. Kết quả thực nghiệm cho thấy với các phương pháp ước lượng trung bình vi mô (micro-average) và trung bình vĩ mô (macro-average), độ chính xác và độ nhạy rất đáng hài lòng. Với LSTM, kết quả kiểm tra đã cho Độ chính xác > 99% và Độ nhạy > 97%, thử nghiệm cho thấy LSTM là lựa chọn đáng giá cho bài toán HAR. Tuy nhiên, LSTM (hoặc học sâu) đòi hỏi thời gian huấn luyện và kiểm tra đáng kể, vượt xa RF, GBDT, SVM và KNN. Do đó, triển khai học sâu có thể không khả thi trên các vi điều khiển tiết kiệm chi phí như PIC18F4520.

So sánh với các mô hình lớn hơn trên cùng máy tính thực nghiệm (Dell XPS-9310, CPU 4,2 GHz, RAM 8 GB): trên cùng bài toán năm hoạt động của tập WISDM, mô hình LSTM đạt độ chính xác > 99% và độ nhạy > 97%, cao hơn RF (độ chính xác tổng thể 96,5%) khoảng 2-3 điểm phần trăm, trong khi mô hình CNN của Chen và cộng sự [95] đạt 93,8% trên dữ liệu của họ. Tuy nhiên, mô hình LSTM có số tham số và bộ nhớ trung gian lớn hơn nhiều bậc so với RAM 1,5 KB của PIC18F4520, đồng thời yêu cầu phép nhân ma trận dấu phẩy động mà PIC18 không có phần cứng hỗ trợ. Do đó, các mô hình lớn được sử dụng làm cận trên tham chiếu về độ chính xác, còn RF là lựa chọn cân bằng giữa độ chính xác và khả năng triển khai trên thiết bị.

Một số nhà nghiên cứu đã khám phá việc sử dụng gia tốc kế tích hợp sẵn để phân loại hoạt động. Tuy nhiên, phương pháp này thường bị giảm độ chính xác do sự can thiệp từ các hoạt động thường ngày như cuộc gọi điện thoại, duyệt web và nhắn tin [17], [88]. Wang và cộng sự [9] đã tiến hành nghiên cứu về nhận dạng các hoạt động thể chất của con người dựa trên dữ liệu gia tốc kế và con quay hồi chuyển ba trục. Nghiên cứu đã sử dụng các bộ phân loại Naive Bayes và KNN để đánh giá khả năng phân loại trên một tập dữ liệu của 6 hoạt động: Đi bộ, Ngồi, Đứng, Nằm, Lên cầu thang và Xuống cầu thang. Kết quả của nghiên cứu cho thấy độ chính xác tốt trong hầu hết các hoạt động. Tuy nhiên, hoạt động Lên cầu thang và Xuống cầu thang có phân loại thấp hơn, với độ

chính xác 82,78% cho Lên cầu thang và 91,78% cho Xuống cầu thang. Ngoài ra, chỉ số PPV cũng khá thấp, đạt 79,62% cho Lên cầu thang và 77,14% cho Xuống cầu thang.

Kết quả của luận án cũng được đối chiếu với nghiên cứu của Vavoulas và cộng sự [20]. Cần lưu ý rằng [20] giới thiệu và sử dụng bộ dữ liệu MobiAct (thu thập từ điện thoại thông minh), không phải bộ dữ liệu WISDM; nghiên cứu này phân loại các hoạt động sinh hoạt hằng ngày với 64 đặc trưng, đánh giá bằng xác thực chéo 10 lần và đạt độ chính xác cao với IBk (99,79%) và J48 (98,63%). Catal và cộng sự [25] áp dụng tổ hợp J48, Hồi quy Logistic và Perceptron đa lớp với 43 đặc trưng, đạt 91,6%. Các nghiên cứu này sử dụng nhiều đặc trưng hơn và chạy trên máy tính, trong khi RF5-HAR chỉ dùng 15 chiều đặc trưng, đạt độ chính xác tổng thể 96,5% trên WISDM và chạy trực tiếp trên vi điều khiển 8-bit. Do khác bộ dữ liệu, định nghĩa chỉ số và nền tảng phần cứng, các so sánh trên chỉ mang tính tham khảo. So với các phương pháp dựa trên thị giác máy tính (Bảng 1.1), phương pháp đề xuất không yêu cầu camera và không ảnh hưởng quyền riêng tư.

So với các nghiên cứu trong nước cùng lĩnh vực, hệ thống của Đào Tô Hiệu và cộng sự [81] hay Nguyễn Quang Huy và cộng sự [82] mới dừng ở việc nhận dạng các hoạt động cơ bản trên địa hình bằng phẳng, chưa mở rộng sang bài toán phân biệt lên/xuống cầu thang cũng như chưa công bố mức tiêu thụ năng lượng và chi phí phần cứng chi tiết để đối chiếu định lượng. Hệ thống RF5-HAR và KNN6-HAR của luận án, do đó, là một trong số ít công trình trong nước triển khai và đánh giá đầy đủ bài toán HAR đa địa hình trực tiếp trên vi điều khiển 8-bit chi phí thấp.

3.3.2. Đánh giá tính ứng dụng và triển khai hệ thống RF5-HAR

Trong các hệ thống hỗ trợ sinh hoạt, đạt được sự chấp nhận của người dùng là một thách thức quan trọng. Mặc dù các phương pháp tiếp cận thị giác máy tính cung cấp các ứng dụng đầy hứa hẹn, việc triển khai chúng trong bối cảnh này có thể bị cản trở bởi các mối quan ngại về quyền riêng tư. Người dùng cao tuổi có thể miễn cưỡng để hình ảnh của họ được theo dõi hoặc ghi lại, có khả năng coi các hệ thống như vậy là xâm phạm. Cảm giác bị giám sát có thể gây khó chịu cho người dùng. Về vấn đề này, thiết bị đeo ở thắt lưng với các cảm biến nhúng cung cấp một giải pháp thực tế.



Hình 3.8 Phân loại được thực hiện bởi vi điều khiển thời gian thực và hiển thị trên điện thoại thông minh

Nghiên cứu này tập trung vào việc phân loại các hoạt động sinh hoạt hàng ngày cơ bản (ADL) ứng dụng cho các tòa nhà cao tầng. Các hoạt động cơ bản cũng khác nhau đối với các độ tuổi và đối tượng khác nhau. Hệ thống phù hợp với các đối tượng cần theo dõi, chăm sóc sức khỏe như người cao tuổi, trẻ em, người lao động trong môi trường rủi ro cao, người bệnh mới phục hồi hoặc bệnh nhân liên quan đến mất kiểm soát hành vi [CT6]. Thiết bị đề xuất phù hợp để đáp ứng nhu cầu theo dõi chăm sóc một cách hiệu quả, độ chính xác cao, tiết kiệm chi phí và tôn trọng tính riêng tư; tuy nhiên, hệ thống hiện mới được kiểm chứng trên người khỏe mạnh, việc áp dụng cho bệnh nhân cần được thử nghiệm thêm.

Mạng cảm biến, một thành phần quan trọng của Internet vạn vật (IoT) với các ứng dụng đa dạng, đã cách mạng hóa nhiều lĩnh vực, đặc biệt là chăm sóc sức khỏe. Khi

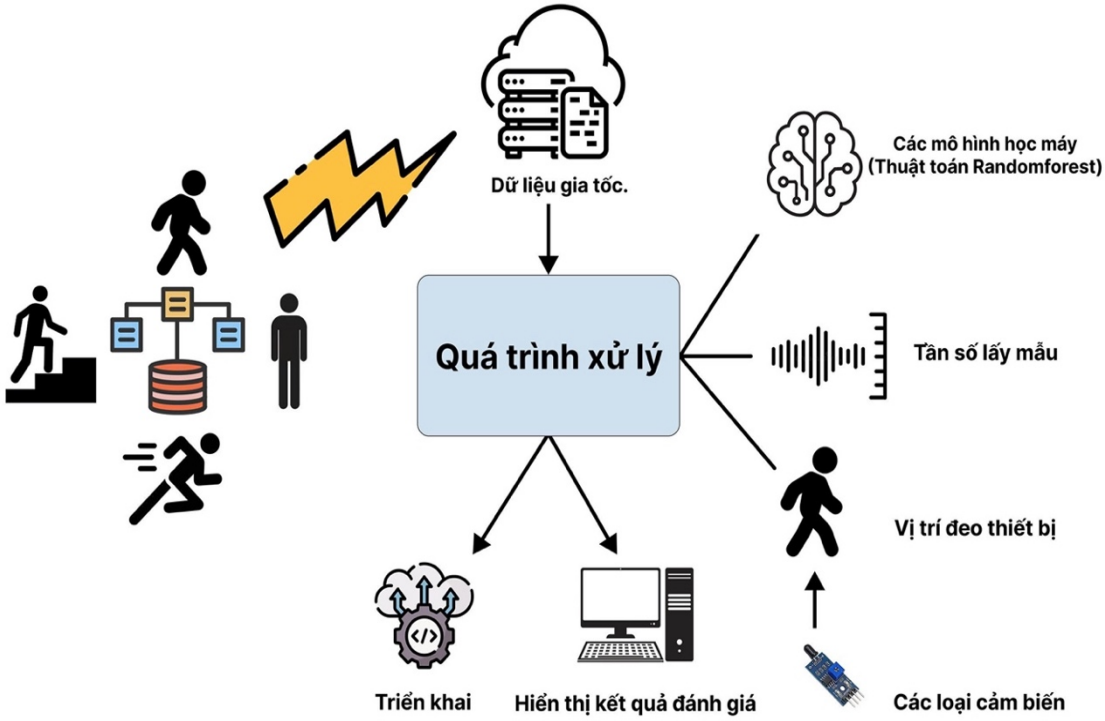
dân số già đi trên toàn cầu, nhu cầu về các giải pháp để hỗ trợ cuộc sống độc lập tăng lên. Nghiên cứu này đề xuất một hệ thống giám sát hoạt động tận dụng công nghệ IoT. Hệ thống này có thể tự động giám sát một người cao tuổi sống một mình. Bằng cách tích hợp hệ thống với các nền tảng truyền thông như ứng dụng web hoặc di động (các hệ thống liên quan đến truyền thông khác trong khuôn khổ IoT), các gia đình có thể nhận được các cập nhật thời gian thực về các hoạt động của người thân của họ (như được minh họa trong Hình 3.8). Giám sát thời gian thực này mang lại sự an tâm, đảm bảo sức khỏe của người cao tuổi, đặc biệt là những người sống xa gia đình của họ. Do đó, hệ thống được đề xuất cung cấp giá trị đáng kể cho các cá nhân cao tuổi, đặc biệt là những người sống độc lập.

Hệ thống nhận dạng hoạt động này ưu tiên sự thoải mái của người dùng trong quá trình sử dụng kéo dài. Không giống như các giải pháp hỗ trợ sinh hoạt cồng kềnh hơn, thiết bị được đề xuất của tác giả sử dụng các cảm biến gia tốc kế không gây khó chịu và tiết kiệm chi phí. Những cảm biến này nhỏ, nhẹ và có giá chỉ bằng một phần nhỏ của thiết bị chuyên dụng, làm cho chúng lý tưởng cho việc áp dụng rộng rãi. Hơn nữa, bản thân thiết bị cũng có giá cả phải chăng đáng kể, với mức giá chỉ 15 USD. Để tối đa hóa tuổi thọ pin và giảm thiểu nhu cầu thay thế, hệ thống sử dụng tốc độ lấy mẫu thấp (50 Hz) trong khi vẫn duy trì nhận dạng hoạt động chính xác.

3.4. Hệ thống KNN6-HAR nhận dạng hoạt động phân biệt trạng thái lên, xuống cầu thang

3.4.1. Kiến trúc mô hình phân loại hoạt động

Trong hệ thống này, một phương pháp học có giám sát sẽ được sử dụng để đạt được độ chính xác cao hơn. Phương pháp hệ thống toàn diện bao gồm ba giai đoạn: 1) thu thập dữ liệu, 2) phân tích mô hình, và 3) nhận dạng hoạt động thời gian thực. Giai đoạn đầu tiên liên quan đến việc thu thập dữ liệu gia tốc ba trục từ cảm biến đeo được. Giai đoạn thứ hai bao gồm việc áp dụng kỹ thuật cửa sổ trượt, trích xuất đặc trưng và huấn luyện mô hình học máy. Giai đoạn cuối cùng tập trung vào việc triển khai mô hình đã huấn luyện trên vi điều khiển để phân loại thời gian thực



Hình 3.9 Các bước trong nhận dạng hoạt động

Thiết bị được đề xuất trong nghiên cứu này bắt đầu thu thập dữ liệu trong giai đoạn đầu tiên. Tác giả đặc biệt tập trung vào việc nắm bắt dữ liệu gia tốc dọc theo ba trục (X, Y và Z) để phân loại các hoạt động. Dữ liệu này trải qua quá trình lấy mẫu, chuyển đổi sang định dạng số và sau đó được truyền đến vi điều khiển để xử lý. Gia tốc kế ADXL345 được sử dụng trong hệ thống do chi phí thấp, tiêu thụ năng lượng thấp và độ tin cậy cao. Cảm biến được cấu hình để thu thập dữ liệu ở tần số lấy mẫu 50Hz, cung cấp sự cân bằng giữa độ chính xác dữ liệu và hiệu quả tính toán. Giai đoạn cuối cùng bao gồm cả triển khai và đánh giá. Ở đây, các đặc trưng được chọn từ bước trước được đưa vào mô hình học máy để huấn luyện và kiểm tra. Mô hình đã huấn luyện sau đó được triển khai trên vi điều khiển để phân loại thời gian thực. Hiệu suất của hệ thống được đánh giá bằng cách sử dụng các số liệu như độ chính xác, độ nhạy và độ đặc hiệu trên cả tập dữ liệu công khai và tập dữ liệu thu thập được. Mô hình hệ thống và thiết bị đeo được bố trí tương tự hệ thống RF4-HAR

3.4.2. Tập dữ liệu

Tập dữ liệu thu thập

Hệ thống được đề xuất của nghiên cứu này sử dụng máy đo gia tốc 3-DOF ADXL345 để thu thập dữ liệu chuyển động. Để đảm bảo phép đo hoạt động toàn diện, các thông số của cảm biến được đặt ở mức ±4 g toàn thang đo, độ nhạy 128 LSB/g và mức nhiễu là 150 μg. Cấu hình này cho phép thu thập dữ liệu ba trục trên các trục X, Y và Z. Cảm biến ADXL345 kết nối với vi điều khiển PIC18F4520 thông qua giao diện I2C. Việc truyền dữ liệu được tạo điều kiện thuận lợi bởi mô-đun ESP8266 được giao tiếp với MCU thông qua giao tiếp UART. Việc thu thập dữ liệu từ cảm biến, trích xuất đặc trưng và phân loại được thực hiện hoàn toàn trên PIC18F4520; ESP8266 chỉ đóng vai trò kênh truyền thông, nhận nhãn hoạt động qua UART và gửi không dây đến máy chủ. Điều này cho phép theo dõi liên tục và phân tích chi tiết các kiểu hoạt động thể chất của người dùng. Mô-đun Wi-Fi tích hợp của ESP8266 phù hợp với môi trường tòa nhà có sẵn hạ tầng mạng. Thiết bị được cung cấp năng lượng bởi pin sạc 3,7 V, 6000 mAh (dung lượng lớn hơn pin 2000 mAh của phiên bản RF4-HAR ở Chương 2).

Trong thiết bị được đề xuất, máy đo gia tốc đo chuyển động của những người tình nguyện khi họ thực hiện nhiều hoạt động khác nhau. ADXL345 được đọc với tần số 50 Hz và các mẫu được chuyển đến PIC18F4520 qua I²C. Trường hợp một vài mẫu bị mất trong quá trình đọc, giá trị bị thiếu được bù bằng trung bình của các mẫu lân cận trong cùng giây, nhờ đó chuỗi dữ liệu vẫn liên tục ở 50 Hz. Ngoài ra, một bộ lọc thông thấp bậc một được áp dụng để làm mịn nhiễu tần số cao trước khi trích xuất đặc trưng.

Thử nghiệm được tiến hành trên máy tính xách tay Dell XPS-9310 với bộ xử lý mạnh mẽ 4,2 GHz và RAM 8 GB. Tác giả đã tiến hành ghi dữ liệu thực nghiệm với tổng cộng 16 tình nguyện viên được chia thành hai nhóm tuổi: nhóm trẻ gồm 14 sinh viên (7 nam và 7 nữ), độ tuổi: 18-22 tuổi, chiều cao: 1,56-1,8 m, cân nặng: 43-72 kg và nhóm người cao tuổi gồm 2 người (1 nam và 1 nữ), độ tuổi 65 và 72 tuổi, chiều cao lần lượt 1,4 m và 1,46 m, cân nặng lần lượt 42 kg và 46 kg để lấy dữ liệu hoạt động thực tế hàng ngày. Thiết bị được đề xuất được tình nguyện viên đeo cố định trên thắt lưng. Sáu hoạt động đã được ghi lại: Đứng (40 phút), Ngồi (30 phút), Chạy bộ (10 phút), Đi bộ (22 phút), Xuống cầu thang (20 phút) và Lên cầu thang (20 phút).

Tần số lấy mẫu 50 Hz thường được sử dụng trong HAR [95]. Ngoài ra, trong nghiên cứu này, việc lựa chọn 50 Hz phù hợp với định lý Nyquist, trong đó nêu rằng tần số lấy

mẫu phải gấp ít nhất hai lần tần số cao nhất có trong tín hiệu để tránh hiện tượng răng cưa. Vì hầu hết các tần số chuyển động quan tâm của con người đều thấp hơn nhiều so với 25 Hz, nên tốc độ lấy mẫu 50 Hz là phù hợp. Trước khi thu thập dữ liệu, các cảm biến gia tốc ba trục sẽ được hiệu chuẩn và sử dụng bộ lọc thông thấp IIR bậc một dạng trung bình trượt hàm mũ $y[i] = K \cdot x[i] + (1 - K) \cdot y[i-1]$ để giảm nhiễu (trong các công bố trước bộ lọc này được gọi là "bộ lọc Kalman đơn giản" [5]). Với $K = 0,1$ và $fs = 50$ Hz, tần số cắt khoảng 0,8 Hz, do đó bộ lọc làm suy giảm một phần thành phần tần số bước chân (1-3 Hz) và ảnh hưởng đến biên độ các đặc trưng Var, P2P; ảnh hưởng này là như nhau cho tập huấn luyện và tập kiểm tra. Trong các phiên bản tiếp theo, có thể tăng K lên khoảng 0,45-0,5 (tần số cắt khoảng 5 Hz) để bảo toàn thành phần bước chân. Cụ thể, ADXL345 được cấu hình với thanh ghi OUTPUT_DATA_RATE = 0x09 (tương ứng 50 Hz). Thời gian xử lý một chu kỳ phân loại của KNN6-HAR: trích xuất 4 đặc trưng khoảng 3 ms và suy luận KNN khoảng 82 ms (tổng khoảng 85 ms < 100 ms); việc truyền kết quả qua ESP8266 (khoảng 50 ms) được thực hiện sau khi đã có nhãn và không thuộc ràng buộc thời gian suy luận.

Tập dữ liệu công khai

Để đánh giá hiệu suất của thuật toán, một tập dữ liệu có sẵn công khai cũng đã được sử dụng [17]. Tập dữ liệu này cung cấp dữ liệu chuyển động cho các hoạt động hàng ngày, được thu thập từ máy đo gia tốc của một thiết bị duy nhất (điện thoại thông minh). Những người tình nguyện mang điện thoại trong túi khi thực hiện các hoạt động được chỉ định. Bộ dữ liệu, được thu thập từ 36 người tham gia, bao gồm sáu hoạt động: Đứng, Ngồi, Chạy bộ, Đi bộ, Xuống cầu thang và Lên cầu thang. Trong nghiên cứu này, giai đoạn thử nghiệm bao gồm 16 tình nguyện viên đeo thiết bị ở eo, như thể hiện trong Hình 2.2 và thực hiện các hoạt động được liệt kê trong Bảng 3.16. Mô tả chi tiết về các hoạt động được cung cấp trong Bảng 3.16.

Bảng 3.16 Mô tả các hoạt động

Hoạt động	**Giải thích**
Đứng	Đứng tự nhiên, ổn định.
Ngồi	Ngồi trên ghế.
Chạy bộ	Chạy với tốc độ đều và chậm rãi.
Đi bộ	Đi bộ bình thường.
Xuống cầu thang	Đi bộ xuống cầu thang bằng cách bước từng bậc một
Lên cầu thang	Đi bộ lên cầu thang bằng cách bước từng bậc một

Sau khi phát triển mô hình được nêu trong Hình 3.9 và xây dựng phần cứng tuân thủ các yêu cầu, thuật toán phân loại hoạt động được tải vào bộ vi điều khiển trong các thiết bị đeo được. Bộ vi điều khiển PIC18F4520 đóng vai trò là bộ xử lý trung tâm trong thiết bị được đề xuất. Sau khi lắp đặt, các thiết bị này được cố định chắc chắn vào thắt lưng của các tình nguyện viên.

Nghiên cứu này ưu tiên phát triển các thuật toán được tối ưu hóa cho các bộ vi điều khiển hiệu năng thấp, tiết kiệm chi phí (thường được gọi là Tiny Machine Learning) phù hợp với các hệ thống nhúng và thiết bị Internet vạn vật (IoT). Trong nghiên cứu này, trọng tâm cụ thể là nhúng thuật toán trực tiếp vào thiết bị để cho phép nhận dạng hoạt động theo thời gian thực. Do đó, khả năng lập trình các thiết bị đeo được này để nhận dạng hoạt động theo thời gian thực là một khía cạnh quan trọng của nghiên cứu này. KNN có thể thích ứng tốt với dữ liệu phi tuyến tính và dễ dàng cập nhật khi có dữ liệu mới. Điều này giúp cải thiện độ chính xác và độ mạnh mẽ của mô hình. Do đó, thuật toán phân loại hoạt động được đề xuất để nhúng các thiết bị đeo được là KNN.

Việc triển khai thuật toán KNN trên các thiết bị được thực hiện như sau: Đầu tiên, cần xác định số láng giềng k (n_neighbors), đây là tham số chính của KNN. Số láng giềng k biểu thị số điểm dữ liệu gần nhất mà mô hình sẽ xem xét để đưa ra dự đoán và giá trị k hợp lý thường được tìm thấy thông qua thử nghiệm với các giá trị khác nhau.

Tham số trọng số cũng có thể được điều chỉnh để chỉ định cách trọng số của các điểm lân cận ảnh hưởng đến dự đoán, với các tùy chọn như "đồng đều" (mặc định, tất cả các điểm có trọng số bằng nhau) hoặc "khoảng cách" (các điểm gần hơn có trọng số lớn hơn). Cuối cùng, tham số thuật toán xác định thuật toán được sử dụng để tính toán các điểm lân cận gần nhất. Việc tinh chỉnh các tham số này sẽ nâng cao hiệu suất và độ chính xác của mô hình KNN.

3.4.3. Quy trình xử lý dữ liệu

Phân đoạn dữ liệu cho các hoạt động

Sau khi thu thập dữ liệu từ máy đo gia tốc, dữ liệu được chia thành các phân đoạn nhỏ hơn, mỗi phân đoạn bao gồm thông tin liên quan đến hoạt động. Phân đoạn rất quan trọng do bản chất tuần tự vốn có của các hoạt động, khiến việc xác định điểm bắt đầu và kết thúc hoạt động trở nên khó khăn.

Nghiên cứu này tận dụng kỹ thuật cửa sổ trượt để phân đoạn do tính đơn giản và hiệu quả của nó trong nghiên cứu nhận dạng hoạt động theo thời gian thực, như được tham chiếu trong [96]. Ở đây, tín hiệu của máy đo gia tốc 3-DOF ADXL345 được chia thành các cửa sổ thời gian có kích thước cố định.

Bảng 3.17 Dữ liệu hoạt động từ tập dữ liệu công khai và thu thập

Mẫu hoạt động	Tổng công khai	Tổng thu thập
Đứng	445	2271
Ngồi	548	1769
Chạy bộ	3253	1563
Đi bộ	4067	2209
Xuống cầu thang	978	905
Lên cầu thang	1185	1028
Tổng	10476	9745

Kích thước cửa sổ được chọn ảnh hưởng đáng kể đến độ chính xác nhận dạng của hệ thống. Để xác định kích thước cửa sổ tối ưu, các thí nghiệm đã được tiến hành với các cửa sổ từ 1 đến 20 giây, với tỷ lệ chồng chéo khác nhau từ 0% đến 90%. Phân tích dữ

liệu ban đầu sử dụng một tập dữ liệu có sẵn công khai [17]. Bảng 3.17 trình bày tổng số quan sát hoạt động thu được từ dữ liệu công khai và dữ liệu đã thu thập. Đối với dữ liệu công khai, sử dụng kích thước cửa sổ trượt là 15 giây và tỷ lệ chồng chéo cửa sổ là 40%, bộ dữ liệu được trích xuất bao gồm 10.476 quan sát. Đối với dữ liệu đã thu thập, số quan sát được ghi lại với kích thước cửa sổ là 8 giây và chồng chéo 40% là 9.745.

Lý do khác nhau về kích thước cửa sổ giữa các hệ thống. Kích thước cửa sổ không được cố định từ trước mà được chọn bằng thực nghiệm cho từng tổ hợp (bộ dữ liệu, tập hoạt động, thuật toán), theo khuyến nghị của Banos và cộng sự [93]. Với RF4-HAR (bốn hoạt động trên địa hình bằng phẳng, chu kỳ bước khoảng 1 giây), RF đạt kết quả tốt nhất ở cửa sổ 10 giây (Hình 2.7-2.10). Với RF5-HAR, Bảng 3.8 cho thấy độ chính xác của RF gần như không đổi (96,5%) ở các cửa sổ 8, 10 và 12 giây, nên luận án chọn cửa sổ nhỏ nhất đạt độ chính xác cao nhất (8 giây) để giảm độ trễ phát hiện và kích thước bộ đệm; hơn nữa, hoạt động cầu thang có chu kỳ bước 0,6-0,8 giây nên cửa sổ 8 giây đã chứa 10-13 bước, đủ để các đặc trưng thống kê ổn định, đồng thời giảm khả năng một cửa sổ chứa cả đoạn đi bộ lẫn đoạn cầu thang ở chiếu nghỉ. Với KNN6-HAR, do bộ đặc trưng thay đổi (Var, SAV, Skew, P2P) và bài toán có sáu lớp, kích thước cửa sổ được khảo sát lại trên lưới 5-20 giây (Hình 3.13-3.16). Tỷ lệ chồng lấp 40% được giữ thống nhất cho cả ba hệ thống để kết quả có thể so sánh. Về kịch bản thực nghiệm, Chương 2 nhằm kiểm chứng tính khả thi trên địa hình bằng phẳng với mô-đun SIM808 (dùng được cả ngoài trời), còn Chương 3 nhằm kiểm tra khả năng mở rộng sang chuyển động theo phương thẳng đứng trong tòa nhà có hạ tầng WiFi (ESP8266), đồng thời đánh giá việc có cần thay đổi bộ đặc trưng và thuật toán khi bổ sung hoạt động cầu thang hay không.

Trích chọn đặc trưng

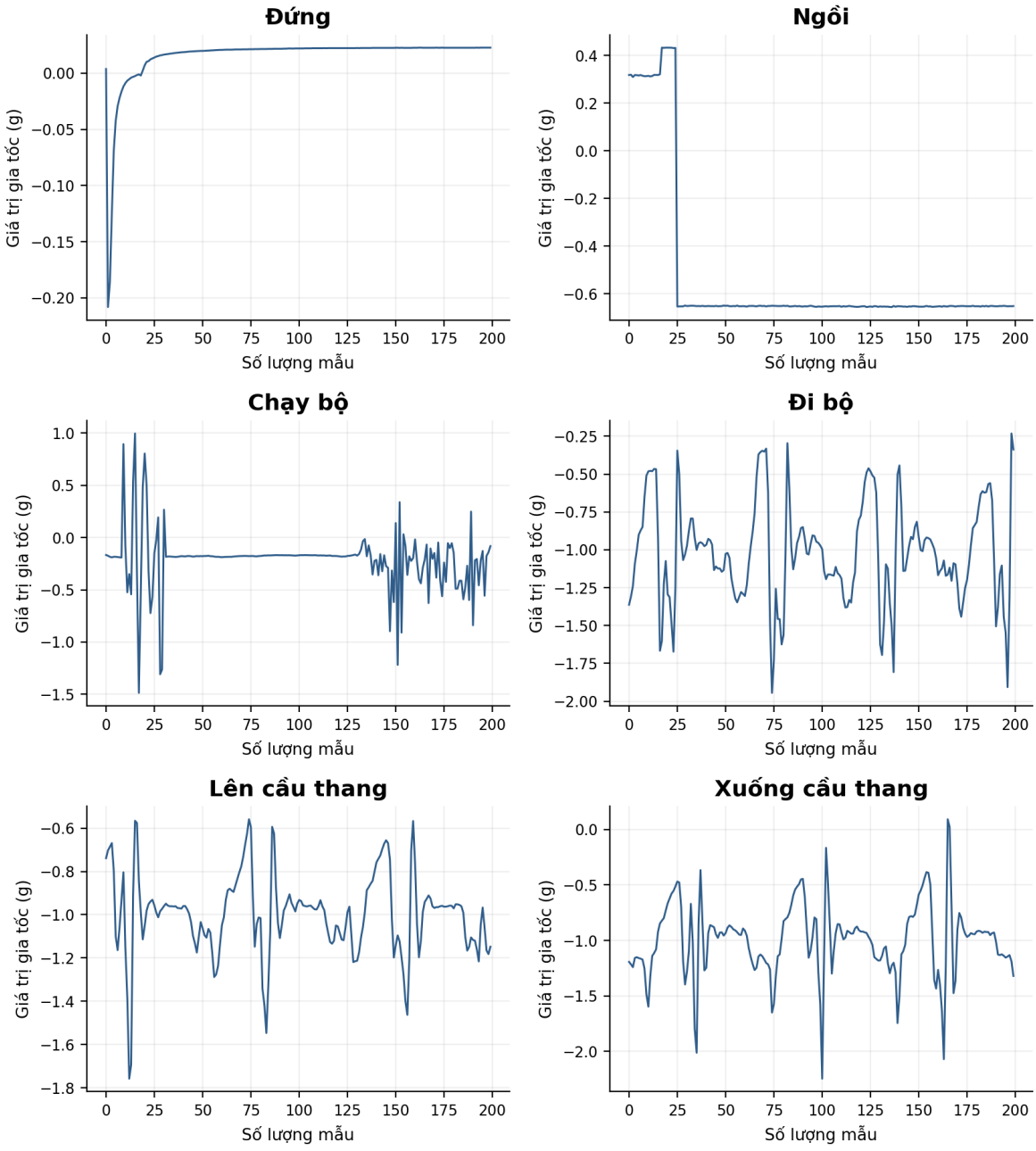
Để đạt được hiệu suất phân loại cao, việc lựa chọn các đặc trưng với thông tin hoạt động là rất quan trọng. Hình 3.10 mô tả phân phối dữ liệu của 400 mẫu dữ liệu đầu tiên thu được từ gia tốc trục X của mỗi hoạt động. Hình minh họa này hỗ trợ việc xác định phạm vi và lựa chọn các đặc trưng phù hợp để phân loại hoạt động.

Hình minh họa này giúp xác định các đặc trưng phân biệt có thể được sử dụng để phân loại hoạt động hiệu quả. Phân tích trực quan này đã hướng dẫn việc lựa chọn bốn đặc trưng thống kê chính: Phương sai (Var), Tổng giá trị tuyệt đối (SAV), Độ lệch

(Skew) và Đỉnh-tới-Đỉnh (P2P, tương đương với đặc trưng Range trong Chương 2, được tính bằng xmax − xmin).

Đối với mỗi tình nguyện viên, tác giả thu thập dữ liệu riêng biệt. Quá trình thu thập dữ liệu sẽ ghi lại tỉ mỉ từng chuyển động. Tính đồng nhất được bảo đảm ở quy trình thu thập (cùng thiết bị, vị trí đeo và kịch bản); mẫu tình nguyện viên vẫn gồm hai nhóm khác biệt rõ về tuổi và thể trạng (18-22 tuổi và 65-72 tuổi) nên khác biệt giữa nhóm trẻ và nhóm cao tuổi chưa được loại bỏ và là một hạn chế của luận án.

Hình 3.10 cho thấy các đặc điểm riêng biệt cho các hoạt động khác nhau. Các trạng thái tĩnh, chẳng hạn như Đứng và Ngồi, thể hiện các giá trị trong phạm vi giới hạn. Ví dụ, Đứng tập trung khoảng 300 điểm dữ liệu ở khoảng 0g trên trục X (1g = 9,8m/s²). Tương tự, Xuống cầu thang hiển thị gần 300 giá trị tập trung giữa -0,4g và 0,3g trên trục X. Điều này cho phép chọn hai đặc trưng Var và SAV để đo mức độ tập trung và biên độ của dữ liệu. Các đặc trưng này phân biệt hiệu quả các trạng thái tĩnh với các trạng thái động và giúp phân loại chúng sâu hơn.



Hình 3.10 Biểu đồ mô tả một phần dữ liệu dọc theo trục X

Ngược lại, các trạng thái động như Chạy bộ và Đi bộ, cũng như các hoạt động như Xuống cầu thang và Lên cầu thang, thể hiện phạm vi giá trị rộng hơn so với các trạng thái tĩnh. P2P được chọn làm đặc trưng để nắm bắt sự khác biệt này bằng cách xác định giá trị tối đa và tối thiểu trong các trạng thái động này. Hơn nữa, Hình 3.10 làm nổi bật các biến thể trong các trạng thái động. Dữ liệu đi bộ thể hiện mật độ giá trị đáng kể nhất, với phân bố thưa dần ở hai phía của dải giá trị. Mặt khác, chạy bộ thể hiện phân phối giá trị đồng đều hơn so với đi bộ. Độ lệch được chọn làm đặc trưng để phân tích sự lan truyền dữ liệu này so với giá trị trung bình, cho phép chúng ta phân biệt giữa các hoạt

động này. Đáng chú ý là Chạy bộ, Xuống cầu thang và Lên cầu thang có mật độ giá trị cao hơn xung quanh 0g trên trục X so với Ngồi và Đứng. Đặc điểm này cho phép Độ lệch trở thành đặc trưng hữu ích để mô tả các hoạt động này.

Kết luận, dựa trên các quan sát từ Hình 3.10, bốn đặc trưng thống kê đã được chọn: Var (Phương sai), SAV (Tổng giá trị tuyệt đối), Skew (Độ lệch) và P2P (Đỉnh-tới-Đỉnh). Các đặc trưng này được chọn vì khả năng phân biệt giữa các hoạt động khác nhau trong khi vẫn duy trì hiệu quả tính toán phù hợp cho triển khai vi điều khiển. Mỗi đặc trưng nắm bắt các khía cạnh khác nhau của tín hiệu chuyển động.

Bảng 3.18 trình bày các công thức của bốn đặc trưng Var, SAV, Skew và P2P được tính toán trên trục X, trục Y và trục Z sử dụng các công thức tương tự. Các Bảng từ 3.19 đến 3.21 cung cấp kết quả của dữ liệu gia tốc được tính toán theo các đặc trưng thống kê Var, SAV, Skew và P2P cho các thao tác trong cả ba chiều (X, Y và Z).

Bảng 3.18 Công thức đặc trưng cho dữ liệu gia tốc trên trục X

Đặc trưng	**Công thức**	
Var	$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$	(3.6)
SAV	$S = \sum_{i=1}^{N}\lvert x_i \rvert$	(3.7)
Skew	$\frac{N}{(N-1)(N-2)}\sum_{i=1}^{N}\left(\frac{x_i - \bar{x}}{s}\right)^3$	(3.8)
P2P	$x_{\max} - x_{\min}$	(3.9)

Trong đó:

x_i là giá trị dữ liệu, N là số lượng giá trị dữ liệu:

$\bar{x}$ là giá trị trung bình của dữ liệu x_i

$|x|$ là giá trị tuyệt đối của dữ liệu x_i

x_{max} là giá trị lớn nhất trong tập dữ liệu

x_{min} i là giá trị nhỏ nhất trong tập dữ liệu

Trong Bảng 3.19, các hoạt động Đứng, Ngồi, Chạy bộ, Đi bộ, Xuống cầu thang và Lên cầu thang cho thấy sự khác biệt đáng kể. Giá trị Var cao nhất được quan sát thấy khi Đứng (6,4) và thấp nhất khi Đi bộ (0,0016). Giá trị SAV cao nhất được quan sát thấy khi Chạy bộ (286,15) và thấp nhất khi Đứng (0,43), cho thấy rằng hoạt động Chạy bộ có biên độ dao động lớn hơn so với khi Đứng yên. Giá trị Skew âm khi Đứng (-0,58), Ngồi (-9,67), Đi bộ (-0,54) và Xuống cầu thang (-0,46) cho thấy phân phối dữ liệu của các hoạt động này lệch trái, trong khi Chạy bộ (0,45) và Lên cầu thang (0,19) lệch phải. Khoảng P2P cao nhất được quan sát thấy khi Chạy bộ (3,99) và thấp nhất khi Đứng (0,02), cho thấy rằng biên độ dao động trong hoạt động Chạy bộ lớn hơn nhiều so với khi Đứng. Nhìn chung, dữ liệu hoạt động Đứng có Var cao và giá trị tuyệt đối thấp, Ngồi có Var trung bình và Skew rất âm, Chạy bộ có dao động lớn với các đặc trưng cao nhất, Đi bộ có Var rất thấp và dao động tối thiểu, trong khi Xuống cầu thang và Lên cầu thang có các đặc trưng trung bình, thể hiện dao động vừa phải và ổn định. Bảng 3.20 và Bảng 3.21 lần lượt trình bày dữ liệu gia tốc trên trục Y và trục Z.

Bảng 3.19 Gia tốc theo trục X (Var: g²; SAV: g, tổng trên cửa sổ; Skew: không thứ nguyên; P2P: g)

Đặc trưng	**Đứng**	**Ngồi**	**Chạy bộ**	**Đi bộ**	**Xuống cầu thang**	**Lên cầu thang**
Var	6,4	3	2,12	0,0016	0,007	0,025
SAV	0,43	4,38	286,15	7,41	14,3	8,7
Skew	-0,58	-9,67	0,45	-0,54	-0,46	0,19
P2P	0,02	0,027	3,99	0,19	0,37	0,3

Bảng 3.20 Gia tốc theo trục Y (Var: g²; SAV: g, tổng trên cửa sổ; Skew: không thứ nguyên; P2P: g)

Đặc trưng	Đứng	Ngồi	Chạy bộ	Đi bộ	Xuống cầu thang	Lên cầu thang
Var	2,29	4,22	0,003	0,0015	0,0002	0,0026
SAV	20,4	2,43	369,16	41,2	13,21	46,4
Skew	-4,75	-10,1	-1,68	0,17	-1,24	-1,25
P2P	0,01	0,074	3,07	0,17	0,08	0,232

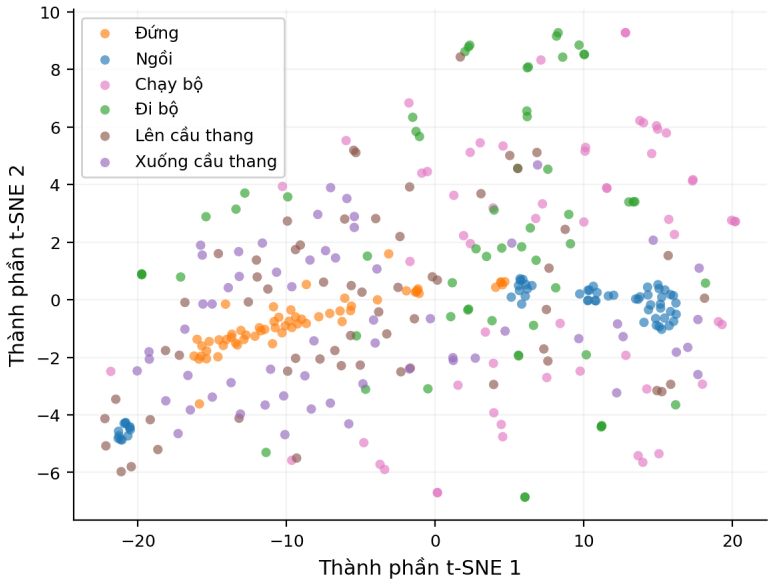
Bảng 3.21 Gia tốc theo trục Z (Var: g²; SAV: g, tổng trên cửa sổ; Skew: không thứ nguyên; P2P: g)

Đặc trưng	Đứng	Ngồi	Chạy bộ	Đi bộ	Xuống cầu thang	Lên cầu thang
Var	5,24	0,0001	0,42	0,004	0,0016	0,006
SAV	0,93	19,8	96,67	9,4	6,17	15,86
Skew	-0,15	13,75	-0,51	2,05	0,003	0,12
P2P	0,02	0,2	3,07	0,42	0,32	0,41

Với các ràng buộc của phân loại hoạt động thời gian thực trên các vi điều khiển chi phí thấp, hiệu năng thấp, các đặc trưng được chọn được thiết kế để cân bằng giữa độ chính xác phân loại và hiệu quả tính toán. Bốn đặc trưng (Var, SAV, Skew, P2P) có thể được tính toán nhanh chóng bằng các phép toán số học đơn giản, làm cho chúng lý tưởng cho triển khai thời gian thực. Độ phức tạp tính toán thấp đảm bảo rằng hệ thống có thể hoạt động trong thời gian thực mà không có độ trễ.

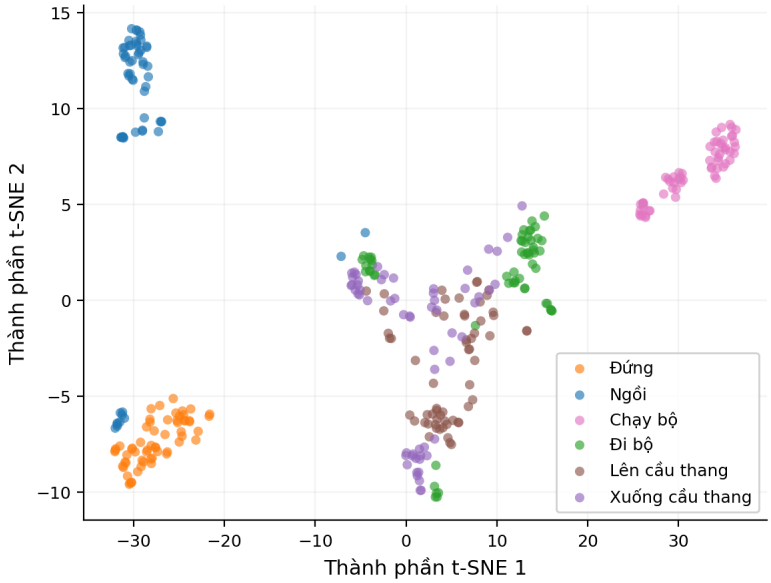
Để có được những hiểu biết sâu sắc hơn từ dữ liệu huấn luyện và chuẩn bị cho việc xây dựng mô hình, nghiên cứu này sử dụng phương pháp t-Distributed Stochastic Neighbor Embedding (t-SNE). Kỹ thuật giảm chiều dữ liệu này biến đổi các điểm dữ

liệu đa chiều thành không gian hai hoặc ba chiều để trực quan hóa. Hình 3.11a cho thấy sự chồng chéo đáng kể giữa các hoạt động, cho thấy khó khăn trong việc phân loại. Ngược lại, Hình 3.11b cho thấy sự phân tách rõ ràng hơn giữa các hoạt động sau khi kết hợp cả bốn đặc trưng (Var, SAV, Skew và P2P). Điều này làm nổi bật lợi thế của việc sử dụng nhiều đặc trưng để cải thiện độ chính xác phân loại. Nhìn chung, việc sử dụng cả bốn đặc trưng liên tục mang lại kết quả nhận dạng hoạt động tốt hơn so với việc sử dụng từng đặc trưng riêng lẻ. Đáng chú ý, việc phân loại các hoạt động như Đứng (vùng màu cam) và Ngồi (vùng màu xanh dương) tương đối đơn giản. Chạy bộ (vùng màu hồng) và Đi bộ (vùng màu xanh lá) cũng thể hiện một mức độ phân tách nhất định. Tuy nhiên, tập dữ liệu công khai cho thấy một số sự chồng chéo giữa các hoạt động Xuống cầu thang (vùng màu tím) và Lên cầu thang (vùng màu nâu).



Hình 3.11 Dữ liệu công khai không có đặc trưng (a) và với đặc trưng (b) trong không gian 2D

Bốn đặc trưng tương tự (Var, SAV, Skew và P2P) đã được áp dụng cho tập dữ liệu thu thập trong nghiên cứu này. Hình 3.12 mô tả biểu đồ được tạo ra bởi thuật toán t-SNE, trực quan hóa sự phân bố của các điểm dữ liệu. Hình này cho thấy sự phân tách rõ ràng giữa hầu hết các hoạt động, mặc dù một số sự chồng chéo vẫn tồn tại giữa Chạy bộ (vùng màu hồng), Đi bộ (vùng màu xanh lá), Xuống cầu thang (vùng màu tím) và Lên cầu thang (vùng màu nâu).



Hình 3.12 Dữ liệu thu thập không có đặc trưng (a) và với đặc trưng (b) trong không gian 2D

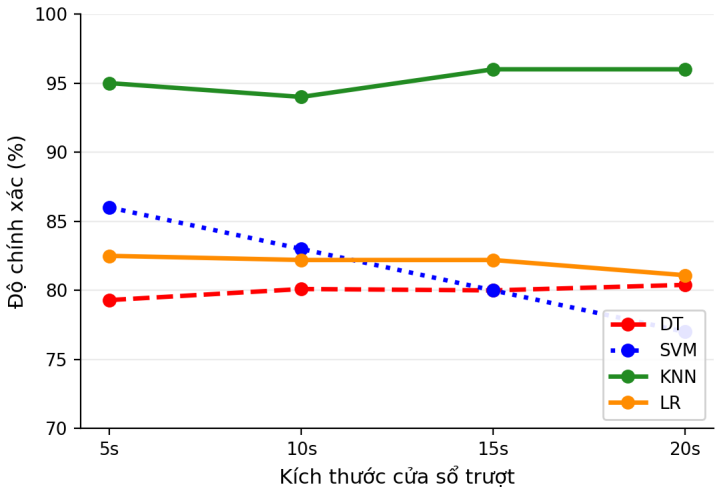
Các đặc trưng được trích xuất (Var, SAV, Skew và P2P) đã được sử dụng để huấn luyện các mô hình học máy cho việc phân loại hoạt động. Ngoài KNN, ba thuật toán phân loại phổ biến đã được đánh giá: Cây quyết định (Decision Tree - DT), Máy vector hỗ trợ (Support Vector Machine - SVM) và Hồi quy Logistic (Logistic Regression - LR). Thuật toán DT xây dựng mô hình bằng cách phân chia dữ liệu đệ quy dựa trên các giá trị đặc trưng, tạo ra cấu trúc các quyết định dạng cây. SVM nhằm tìm siêu phẳng tối ưu giúp tối đa hóa khoảng cách biên giữa các lớp khác nhau, sử dụng các hàm nhân khác nhau để giải quyết các phân tách phi tuyến. Thuật toán LR được sử dụng cho phân loại nhị phân bằng cách mô hình hóa xác suất của một lớp sử dụng hàm logistic, tạo ra ranh giới quyết định thông qua tổ hợp tuyến tính của các đặc trưng đầu vào. Các thuật toán này đã được triển khai sử dụng thư viện scikit-learn (thường được viết tắt là sklearn) trong Python.

Để đánh giá hiệu suất của các mô hình này, ma trận nhầm lẫn (confusion matrix) đã được sử dụng. Ma trận này cung cấp một hình ảnh trực quan rõ ràng về mức độ chính xác của mô hình trong việc phân loại các hoạt động.

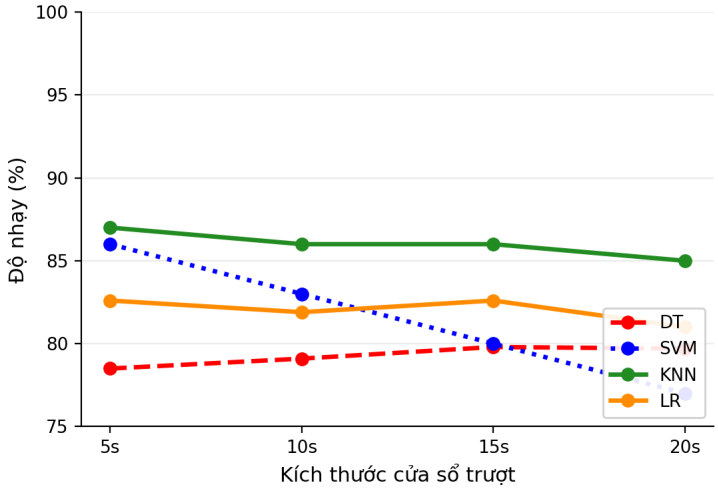
3.5. Thực nghiệm và kết quả của phương pháp nhận dạng hoạt động phân biệt trạng thái lên, xuống cầu thang

Hình 3.13 và Hình 3.14 lần lượt mô tả độ chính xác tổng thể và độ nhạy của mỗi thuật toán DT, KNN, SVM và LR trên tập dữ liệu công khai. Hiệu suất của mỗi bộ phân loại thay đổi theo các kích thước cửa sổ thời gian khác nhau (5 giây, 10 giây, 15 giây, 20 giây). Trong số các bộ phân loại này, bộ phân loại KNN đạt hiệu suất phân loại cao nhất, với độ chính xác 96% ở kích thước cửa sổ 15 giây và độ nhạy 87% ở kích thước cửa sổ 5 giây.

3.5.1. Kết quả thực nghiệm trên các thuật toán



Hình 3.13 Độ chính xác của DT, SVM, KNN và LR trên tập dữ liệu công khai



Hình 3.14 Độ nhạy của DT, SVM, KNN và LR trên tập dữ liệu công khai

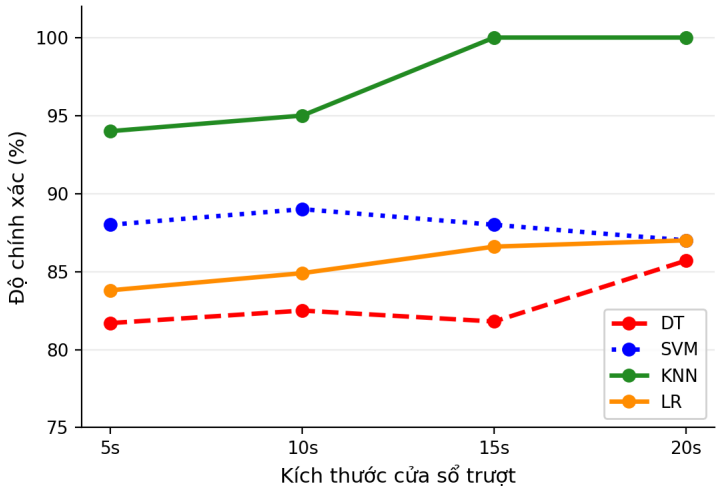
Như được hiển thị trong Hình 3.15 và Hình 3.16, tập dữ liệu thu thập được đạt độ chính xác và độ nhạy tốt hơn so với tập dữ liệu công khai. Sự cải thiện này có thể được quy cho một số yếu tố sau:

Nguồn dữ liệu: Không giống như tập dữ liệu công khai với dữ liệu được thu thập từ điện thoại đặt trong túi quần, dữ liệu của tác giả đến từ thiết bị được đeo cố định trên thắt lưng, giảm thiểu các nhiễu do chuyển động.

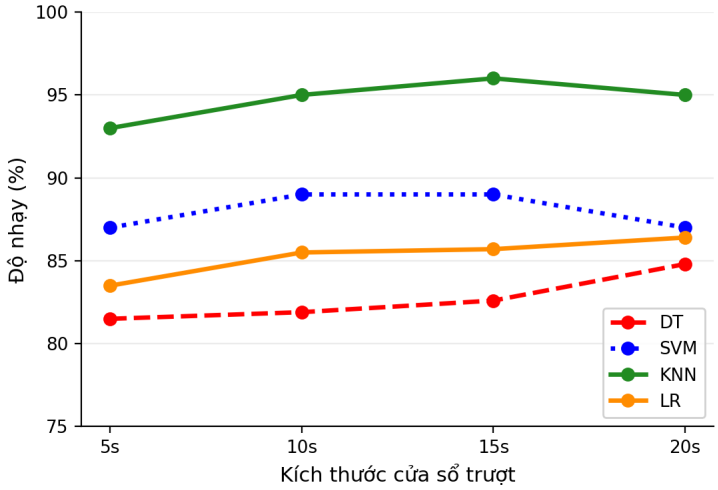
Thu thập có cấu trúc: Quá trình thu thập dữ liệu của tác giả bao gồm các hoạt động cá nhân được kiểm soát, dẫn đến độ lệch thấp hơn so với sự thay đổi trong các hoạt động hàng ngày.

Quy trình đồng nhất: tất cả tình nguyện viên dùng cùng thiết bị, cùng vị trí đeo và cùng kịch bản thực hiện, giảm biến động không liên quan đến hoạt động.

Sự kết hợp của các yếu tố này góp phần tạo nên độ chính xác cao đáng kể (100% đối với bộ phân loại KNN với cửa sổ 15 giây) và độ nhạy (96% đối với bộ phân loại KNN với cửa sổ 15 giây) quan sát được trong tập dữ liệu của tác giả. Cần lưu ý rằng chính các điều kiện kiểm soát này cũng làm cho kết quả trên tập thu thập lạc quan hơn so với điều kiện sử dụng tự do ngoài thực tế.



Hình 3.15 Độ chính xác của DT, SVM, KNN và LR trên tập dữ liệu thu thập



Hình 3.16 Độ nhạy của DT, SVM, KNN và LR trên tập dữ liệu thu thập

Lập luận lựa chọn KNN thay cho RF. Việc lựa chọn KNN cho KNN6-HAR dựa trên bốn căn cứ. (1) Căn cứ thực nghiệm: ở RF5-HAR, ngay cả khi gộp hai lớp cầu thang, RF vẫn nhầm lớp cầu thang với Đi bộ và Chạy bộ (độ nhạy 89,5% trên WISDM, Bảng 3.13); khi tách thành hai lớp Lên/Xuống có phân bố đặc trưng gần nhau, ranh giới quyết định trở nên phi tuyến và mang tính cục bộ, phù hợp với cơ chế phân loại theo láng giềng gần nhất hơn là các phép chia song song trục tọa độ của cây quyết định. (2) Trong bốn thuật toán có khả năng triển khai được đánh giá ở Mục 3.5.1 (DT, KNN, SVM, LR), KNN đạt độ chính xác cao nhất trên cả hai tập dữ liệu. (3) KNN cho phép bổ sung mẫu mới của người dùng vào bảng tra mà không cần huấn luyện lại, thuận lợi cho cá nhân hóa đối với người cao tuổi. (4) Chi phí suy luận O(N·d) với N = 1.200 mẫu đại diện vẫn đáp ứng ràng buộc thời gian suy luận (khoảng 82 ms). Hạn chế của lập luận: RF chưa được huấn luyện trên cùng bộ đặc trưng Var, SAV, Skew, P2P, do đó chưa tách bạch hoàn toàn ảnh hưởng của thuật toán và của bộ đặc trưng; nội dung này được nêu trong phần hạn chế và hướng nghiên cứu tiếp theo.

3.5.2. Triển khai mô hình KNN trên vi điều khiển

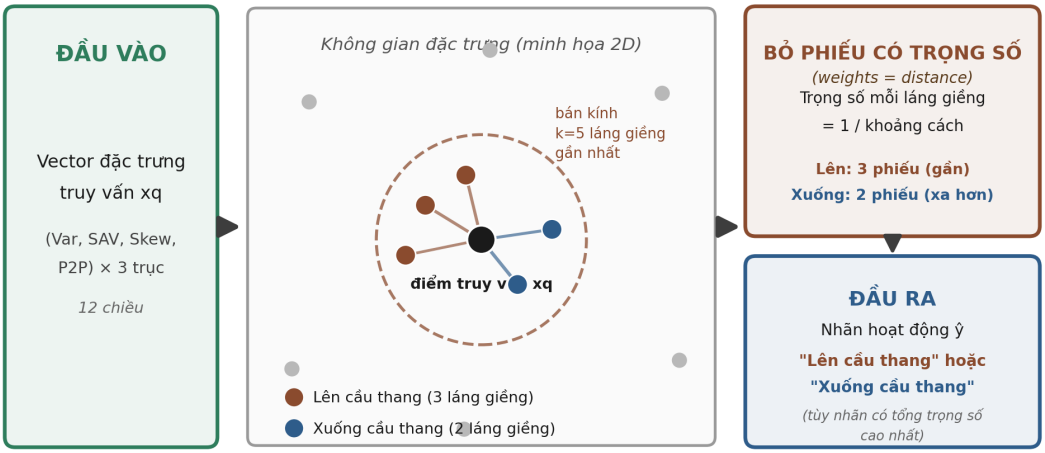
Khác với RF4-HAR và RF5-HAR sử dụng Random Forest, hệ thống KNN6-HAR lựa chọn thuật toán K-Nearest Neighbors (KNN, k=5, trọng số theo khoảng cách). Nguyên nhân là do dữ liệu gia tốc của hai hoạt động Lên và Xuống cầu thang có đặc trưng thống kê gần nhau, dẫn đến ranh giới quyết định phi tuyến mà mô hình cây quyết định của RF khó phân tách rạch ròi; trong khi đó, KNN không yêu cầu giả định về dạng ranh giới và phân loại trực tiếp dựa trên khoảng cách Euclidean tới các mẫu lân cận, phù

hợp hơn với đặc điểm dữ liệu này. Chi phí tính toán khi suy luận của KNN là O(n·d) với tập huấn luyện được nén còn khoảng 1.200 mẫu đại diện lưu dưới dạng bảng tra cứu (look-up table) trong bộ nhớ Flash, cho thời gian suy luận khoảng 82 ms/lần phân loại trên PIC18F4520 (Bảng 3.27), vẫn đáp ứng ràng buộc thời gian thực. Chi tiết cách xác định tham số k, cấu hình cảm biến và các bước cài đặt được trình bày tại Phụ lục, mục 3.

Đầu vào và đầu ra của thuật toán KNN trong hệ thống KNN6-HAR được mô tả cụ thể như sau:

- **Đầu vào (Input):** vector đặc trưng truy vấn xq gồm 4 giá trị thống kê (Var, SAV, Skew, P2P) tính trên mỗi trục X, Y, Z của tín hiệu gia tốc trong một cửa sổ trượt, tổng cộng 12 chiều.
- **Xử lý:** tính khoảng cách Euclidean từ xq đến toàn bộ mẫu trong tập huấn luyện đã lưu (look-up table trong Flash), chọn ra k=5 mẫu gần nhất; mỗi láng giềng được gán trọng số bằng nghịch đảo khoảng cách (weights='distance').

-Đầu ra (Output): nhãn hoạt động ŷ ∈ {Đứng, Ngồi, Chạy bộ, Đi bộ, Xuống cầu thang, Lên cầu thang} được xác định bằng bỏ phiếu có trọng số trên k=5 láng giềng, nhãn có tổng trọng số cao nhất là kết quả cuối cùng.



Hình 3.17 Sơ đồ mô tả thuật toán K-Nearest Neighbors (k=5)

Bảng 3.22 và Bảng 3.23 cho thấy kết quả chi tiết của việc sử dụng thuật toán KNN để phân loại các hoạt động trên tập dữ liệu công khai và tập dữ liệu thu thập được. Từ

Bảng 3.22, có thể thấy rõ rằng mô hình hoạt động tốt trong việc phân biệt hầu hết các hoạt động, với độ chính xác cao đối với "Đứng", "Ngồi", "Chạy bộ" và "Đi bộ". Cụ thể, 141 trong số 149 mẫu cho "Đứng" được phân loại chính xác, 181 trong số 184 mẫu cho "Ngồi", 1014 trong số 1085 mẫu cho "Chạy bộ" và 1264 trong số 1356 mẫu cho "Đi bộ". Tuy nhiên, mô hình cho thấy nhiều sự nhầm lẫn hơn giữa các hoạt động tương tự nhau, đặc biệt là "Xuống cầu thang" và "Lên cầu thang". Đối với "Xuống cầu thang", chỉ có 175 trong số 327 mẫu được xác định chính xác, trong khi đối với "Lên cầu thang", 236 trong số 396 mẫu được phân loại đúng. Các phân loại sai chủ yếu xảy ra giữa "Xuống cầu thang", "Đi bộ" và "Lên cầu thang", cho thấy cần cải thiện trong việc phân biệt giữa các hoạt động này.

Bảng 3.22 Ma trận nhầm lẫn của thuật toán KNN trên tập dữ liệu công khai

Hoạt động Quan sát	**Hoạt động dự đoán**						**Tổng**
	Đứng	***Ngồi***	***Chạy bộ***	***Đi bộ***	***Xuống cầu thang***	***Lên cầu thang***	
Đứng	**141**	1	0	0	5	2	149
Ngồi	0	**181**	1	0	0	2	184
Chạy bộ	0	0	**1014**	54	13	4	1085
Đi bộ	1	0	19	**1264**	22	17	1356
Xuống cầu thang	5	0	31	71	**175**	48	327
Lên cầu thang	3	0	27	53	34	**236**	396
Tổng	150	182	1112	1478	258	317	**3497**

Tương tự như tập dữ liệu công khai, trong Bảng 3.23, mô hình hoạt động rất tốt trong việc phân biệt hầu hết các hoạt động, đặc biệt là "Đứng" và "Ngồi", với hầu như tất cả các mẫu đều được phân loại chính xác (757 trong số 758 mẫu cho "Đứng" và 589 trong

số 590 mẫu cho "Ngồi"). Mô hình cũng hoạt động tốt đối với "Chạy bộ" với 511 trong số 522 mẫu được phân loại đúng và đối với "Đi bộ" với 686 trong số 737 mẫu được phân loại đúng. Tuy nhiên, có sự nhầm lẫn đáng chú ý khi phân biệt giữa "Xuống cầu thang" và "Lên cầu thang". Đối với "Xuống cầu thang", chỉ có 267 trong số 302 mẫu được xác định chính xác và đối với "Lên cầu thang", 314 trong số 344 mẫu được phân loại đúng. Các phân loại sai chủ yếu xảy ra giữa "Xuống cầu thang" và "Đi bộ", cũng như giữa "Lên cầu thang" và "Đi bộ".

Bảng 3.23 Ma trận nhầm lẫn của thuật toán KNN trên tập dữ liệu thu thập

Hoạt động Quan sát	**Hoạt động dự đoán**						**Tổng**
	Đứng	***Ngồi***	***Chạy bộ***	***Đi bộ***	***Xuống cầu thang***	***Lên cầu thang***	
Đứng	**757**	0	0	0	0	1	758
Ngồi	0	**589**	0	1	0	0	590
Chạy bộ	0	9	**511**	2	0	0	522
Đi bộ	0	0	0	**686**	19	32	737
Xuống cầu thang	3	0	1	30	**267**	1	302
Lên cầu thang	0	0	0	17	13	**314**	344
Tổng	760	598	512	736	299	348	**3253**

Độ chính xác của việc phân loại hoạt động được trình bày trong Bảng 3.24 và Bảng 3.25 lần lượt cho tập dữ liệu công khai và tập dữ liệu thu thập được. Độ chính xác theo lớp (công thức 1.1) của KNN vượt 91% cho tất cả các hoạt động trên cả hai tập dữ liệu. Các chỉ số được tính theo các công thức (1.1) đến (1.4).

Bảng 3.24 Hiệu suất mô hình trên tập dữ liệu công khai

Hoạt động	Hiệu suất thuật toán			
	Accuracy (%)	***Sensitivity (%)***	***PPV (%)***	***NPV (%)***
Đứng	99,51	94,63	94,00	99,76
Ngồi	99,89	98,37	99,45	99,91
Chạy bộ	95,17	93,46	91,19	97,02
Đi bộ	91,25	93,22	85,52	95,44
Xuống cầu thang	93,28	53,52	67,83	95,31
Lên cầu thang	93,11	59,60	74,45	94,97

Độ chính xác theo lớp và NPV của KNN vượt 91% trên tập dữ liệu công khai và vượt 96% trên tập dữ liệu thu thập đối với tất cả các hoạt động. Độ nhạy và PPV vượt 85% cho các lớp Đứng, Ngồi, Chạy bộ và Đi bộ trên tập công khai, nhưng thấp hơn nhiều ở hai lớp cầu thang (độ nhạy 53,52% và 59,60%). Đáng chú ý, các hoạt động Đứng, Ngồi, Chạy bộ thể hiện giá trị gần như tuyệt đối trên tập dữ liệu thu thập được. Có thể thấy từ Bảng 3.24 và Bảng 3.25 rằng các hoạt động Xuống cầu thang và Lên cầu thang có độ chính xác phân loại thấp hơn nhiều so với các hoạt động còn lại. Nguyên nhân là do sự tương đồng của dữ liệu gia tốc trong việc xuống cầu thang và lên cầu thang.

Phân tích chi tiết theo từng lớp. Trên tập WISDM, hai lớp cầu thang có độ nhạy thấp nhất: theo Bảng 3.22, 48/327 mẫu Xuống cầu thang (14,7%) bị nhận nhầm thành Lên cầu thang và 71/327 mẫu (21,7%) thành Đi bộ; 34/396 mẫu Lên cầu thang (8,6%) bị nhận nhầm thành Xuống và 53/396 mẫu (13,4%) thành Đi bộ. Các cặp dễ nhầm lẫn nhất do đó là Xuống/Lên cầu thang và cầu thang/Đi bộ, phù hợp với nhận xét rằng cả ba hoạt động đều có chu kỳ bước tương tự và chỉ khác nhau ở thành phần gia tốc thẳng đứng. Sự mất cân bằng lớp cũng góp phần: trên WISDM, Đi bộ chiếm 38,8% số quan sát (4.067/10.476) trong khi mỗi lớp cầu thang chỉ chiếm 9-11% (Bảng 3.17), khiến các láng giềng gần nhất của một mẫu cầu thang có xác suất cao thuộc lớp Đi bộ. Trên tập tự

thu thập, tỷ lệ nhầm giảm mạnh (Xuống sang Đi bộ 30/302 = 9,9%; Lên sang Xuống 13/344 = 3,8%; Bảng 3.23) nhờ vị trí đeo cố định. Hướng cải thiện gồm cân bằng lớp (lấy mẫu lại hoặc gán trọng số lớp) và bổ sung đặc trưng gia tốc thẳng đứng sau khi hiệu chuẩn hướng trọng trường.

Bảng 3.25 Hiệu suất mô hình trên tập dữ liệu thu thập

Hoạt động	**Hiệu suất thuật toán**			
	Accuracy (%)	***Sensitivity (%)***	***PPV (%)***	***NPV (%)***
Đứng	99,88	99,87	99,61	99,96
Ngồi	99,69	99,83	98,49	99,96
Chạy bộ	99,63	97,89	99,80	99,96
Đi bộ	96,90	93,08	93,21	97,97
Xuống cầu thang	97,94	88,41	89,30	98,82
Lên cầu thang	98,03	91,28	90,23	98,97

3.5.3. Đánh giá hiệu suất của mô hình phân loại KNN

Đánh giá hiệu suất tổng thể bằng cả trung bình vi mô và trung bình vĩ mô (Mục 1.3.3), Bảng 3.26 cho thấy KNN đạt độ chính xác tổng thể đa lớp 86,10% trên tập dữ liệu công khai và 96,03% trên tập dữ liệu thu thập; độ nhạy và PPV vi mô bằng độ chính xác tổng thể. Theo trung bình vĩ mô, độ nhạy đạt 82,13% (công khai) và 95,06% (thu thập), PPV đạt 85,41% và 95,11%, NPV đạt 97,07% và 99,21%. Độ chính xác theo lớp trung bình (công thức 1.1) đạt 95,37% và 98,68%, là giá trị có thể so sánh trực tiếp với con số 97,7%/99,6% của RF4-HAR. Khoảng cách giữa độ nhạy vĩ mô (82,13%) và độ chính xác tổng thể (86,10%) trên tập công khai phản ánh độ nhạy thấp của hai lớp cầu thang.

Ảnh hưởng của chuẩn hóa đặc trưng. Đặc trưng SAV (tổng giá trị tuyệt đối trên cửa sổ) có giá trị cỡ 10^2-10^3, trong khi Skew cỡ 10^0; nếu không chuẩn hóa, khoảng cách Euclid bị chi phối bởi SAV. Các kết quả chính của KNN6-HAR trong Bảng 3.22-3.26 (86,1% và 96,03%) là kết quả trước chuẩn hóa. Khi áp dụng chuẩn hóa Min-Max (các

tham số min, max được ước lượng trên tập huấn luyện), độ chính xác tổng thể tăng lên 88,7% (WISDM) và 96,8% (tập thu thập). Trên PIC18F4520, phép chuẩn hóa được thực hiện bằng số học dấu phẩy tĩnh với chi phí khoảng 0,5 ms/lần phân loại, do đó luận án khuyến nghị áp dụng chuẩn hóa Min-Max trong các triển khai tiếp theo. Ngoài ra, mô hình nhúng sử dụng 1.200 mẫu đại diện được chọn bằng k-means (200 mẫu/lớp) lưu ở dạng số nguyên 16 bit, khác với mô hình KNN đầy đủ trên máy tính dùng để báo cáo các kết quả trên.

Bảng 3.26 Đánh giá mô hình phân loại KNN

Các chỉ số đánh giá	**Trung bình vi mô (%)**		**Trung bình vĩ mô (%)**	
	Tập dữ liệu công khai	***Tập dữ liệu thu thập***	***Tập dữ liệu công khai***	***Tập dữ liệu thu thập***
Độ chính xác tổng thể đa lớp	86,10	96,03	-	-
Accuracy theo lớp (1.1)	95,37	98,68	95,37	98,68
Sensitivity	86,10	96,03	82,13	95,06
PPV	86,10	96,03	85,41	95,11
NPV	97,22	99,21	97,07	99,21

Trong học máy, trích xuất đặc trưng từ dữ liệu thô là một bước quan trọng. Hiệu quả của quá trình này phụ thuộc phần lớn vào sự hiểu biết của người thiết kế mô hình về tập dữ liệu. Mặc dù các đặc trưng được lựa chọn tốt thường cải thiện độ chính xác của mô hình, việc bao gồm các đặc trưng không liên quan hoặc dư thừa có thể gây ra gánh nặng tính toán không cần thiết, có khả năng cản trở hiệu suất. Điều này nhấn mạnh tầm quan trọng của việc chọn đúng các đặc trưng cốt lõi. Hơn nữa, tác giả đã thử nghiệm bổ sung tập đặc trưng ban đầu với các đặc trưng miền thời gian bổ sung. Mặc dù điều

này thường dẫn đến giảm độ chính xác phân loại đối với tập dữ liệu công khai, tác động đến tập dữ liệu thu thập được là không đáng kể. Như vậy, đối với nhiệm vụ cụ thể là phân loại hành vi trong nghiên cứu của tác giả, tập đặc trưng được đề xuất mang lại sự cân bằng tốt giữa hiệu suất phân loại và hiệu quả tính toán.

3.6. Đánh giá hệ thống KNN6-HAR nhận dạng hoạt động có phân biệt lên, xuống cầu thang ở địa hình tòa nhà cao tầng

3.6.1. So sánh với các nghiên cứu trước

Nhiều nghiên cứu đã mở rộng phạm vi lựa chọn đặc trưng để xác định chính xác hơn các hoạt động của con người: Catal và cộng sự [25] dùng 43 đặc trưng, Vavoulas và cộng sự [20] dùng 64 đặc trưng. Tuy nhiên, việc sử dụng một số lượng lớn đặc trưng làm tăng chi phí tính toán và không phù hợp cho triển khai trên vi điều khiển hiệu năng thấp. Lưu ý khi so sánh: các nghiên cứu trên sử dụng nền tảng phần cứng khác nhau (PC, smartphone, Raspberry Pi), trong khi luận án triển khai trực tiếp trên MCU 8-bit PIC18F4520 với 12-15 chiều đặc trưng (4-5 đặc trưng thống kê trên mỗi trục); vì vậy so sánh trực tiếp độ chính xác cần xét đến ràng buộc tài nguyên phần cứng. Kết quả của luận án cho thấy với lựa chọn đặc trưng phù hợp, số lượng đặc trưng nhỏ vẫn có thể đạt độ chính xác cao.

Một số nhà nghiên cứu sử dụng dữ liệu gia tốc kế và học sâu để phân loại các hoạt động của con người. Yuqing Chen và cộng sự [95] đã xây dựng một mô hình CNN và sửa đổi kernel tích chập để điều chỉnh các đặc tính của dữ liệu cảm biến theo trục thời gian. Mô hình của họ đạt độ chính xác trung bình 93,8% trên tập dữ liệu do nhóm tác giả thu thập. Mặc dù các phương pháp học sâu có thể đạt được độ chính xác cao, chúng yêu cầu tài nguyên tính toán đáng kể và không phù hợp cho triển khai thời gian thực trên vi điều khiển hiệu năng thấp. So sánh với các TinyML framework hiện đại: Edge Impulse và TensorFlow Lite Micro (TFLM) thường yêu cầu ARM Cortex-M4+ (≥256 KB RAM) để chạy mô hình deep learning. Phương pháp RF/KNN đề xuất chạy được trên PIC18F4520 (1,5 KB RAM), dung lượng RAM nhỏ hơn khoảng 170 lần; đây là điểm khác biệt của hướng nghiên cứu này so với các TinyML framework thông thường.

Một số nhà nghiên cứu đã khám phá việc sử dụng gia tốc kế tích hợp để phân loại hoạt động bằng điện thoại thông minh. Büber và Guvensan (2014) [97] đã thiết kế một

hệ thống nhận dạng cho các hoạt động bao gồm đi bộ, chạy, lên cầu thang, xuống cầu thang, ngồi và đứng. Hệ thống của họ đạt được độ chính xác tổng thể 91,06%. So với nghiên cứu của họ, hệ thống được đề xuất của tác giả đạt độ chính xác cao hơn (96,03% trên tập dữ liệu thu thập) trong khi sử dụng vi điều khiển hiệu năng thấp hơn.

Phân tích khả năng khái quát hóa. Chênh lệch độ chính xác tổng thể của KNN6-HAR giữa tập WISDM (86,1%) và tập tự thu thập (96,03%) có thể được quy cho các nguyên nhân sau: (1) Vị trí đeo và cách thu thập: WISDM thu từ điện thoại đặt tự do trong túi quần trước, hướng của điện thoại thay đổi giữa các lần đeo, còn thiết bị của luận án được cố định ở thắt lưng dưới sự giám sát; (2) Phân bố dữ liệu: WISDM mất cân bằng mạnh (Đi bộ 38,8%, mỗi lớp cầu thang 9-11%), trong khi tập tự thu thập cân bằng hơn; (3) Đặc điểm người tham gia: WISDM gồm 36 người với điều kiện tự nhiên, tập tự thu thập gồm 16 người thực hiện theo kịch bản; (4) Tần số lấy mẫu: WISDM lấy mẫu 20 Hz, thấp hơn 50 Hz của thiết bị, làm thay đổi giá trị các đặc trưng phụ thuộc số mẫu như SAV và P2P; (5) Bộ đặc trưng chưa chuẩn hóa: SAV chi phối khoảng cách Euclid nên KNN nhạy với khác biệt biên độ giữa người dùng và thiết bị (Mục 3.5.3). Do đó, kết quả trên tập tự thu thập đại diện cho kịch bản thiết bị đeo cố định ở thắt lưng trên người khỏe mạnh trong điều kiện có kiểm soát; kết quả trên WISDM gần hơn với điều kiện sử dụng tự do. Phạm vi đại diện của các kết quả vì vậy chưa bao gồm người dùng mới hoàn toàn (do chưa đánh giá độc lập theo đối tượng), các vị trí đeo khác và nhóm người có bệnh lý vận động.

Ở hệ thống này, tác giả đã mô tả quá trình phát triển một hệ thống nhận dạng hoạt động của con người theo thời gian thực. Hệ thống sử dụng kích thước cửa sổ được lựa chọn bằng thực nghiệm, bốn đặc trưng thống kê và một mô hình học máy đơn giản. Hiệu quả của hệ thống được đề xuất đã được thử nghiệm trên cả tập dữ liệu có sẵn công khai và dữ liệu được thu thập trong các thí nghiệm, đạt độ chính xác tổng thể 96,03% và độ nhạy trung bình vĩ mô 95,06% trên tập dữ liệu thu thập. Hệ thống có thể nhận dạng sáu hoạt động hàng ngày thiết yếu của con người trong tòa nhà cao tầng: Đứng, Ngồi, Chạy bộ, Đi bộ, Lên cầu thang và Xuống cầu thang.

Việc tích hợp một thuật toán hiệu quả, ít phức tạp vào một bộ vi điều khiển hiệu năng thấp để phân loại các hoạt động của con người trong tòa nhà cao tầng mở ra cánh cửa

cho những hướng nghiên cứu mới thú vị trong lĩnh vực này. Kết quả thử nghiệm cho thấy hiệu quả của DT, KNN, SVM và LR khi sử dụng các kích thước cửa sổ khác nhau. Hiệu quả tính toán của thuật toán KNN khiến nó đặc biệt phù hợp để triển khai theo thời gian thực bằng dữ liệu kỹ thuật số như số liệu đo gia tốc kế. Thuật toán KNN mang lại hiệu suất phân loại cao nhất ở kích thước cửa sổ 15 giây trên tập WISDM. Bằng cách tận dụng bốn đặc trưng chính (Var, SAV, Skew và P2P), hệ thống này đạt được khả năng phân loại chính xác các hoạt động hàng ngày cơ bản trong tòa nhà cao tầng.

Phương pháp này có các ứng dụng theo dõi các hoạt động đi lại, lên, xuống trong tòa nhà nhằm hỗ trợ người cao tuổi, trẻ nhỏ, người bệnh trong giai đoạn phục hồi, người lao động để cảnh báo, phòng tránh rủi ro, hỗ trợ khi gặp nạn. Do đó, dữ liệu được cá nhân hóa và tập trung vào phân tích hoạt động và phạm vi được thu hẹp lại thành sáu hành vi cụ thể này để có độ chính xác cao hơn. Sáu hành vi này được coi là đủ cho ứng dụng trong khi vẫn đảm bảo phân loại theo thời gian thực và tối ưu hóa chi phí.

3.6.2. So sánh với hệ thống RF5-HAR

RF5-HAR phân loại 5 hoạt động trong khi KNN6-HAR phân loại 6 hoạt động, mức độ khó phân loại khác nhau nên không thể so sánh trực tiếp. Về tính nhạy cảm với vị trí thiết bị (domain shift): thực nghiệm bổ sung cho thấy khi thay đổi vị trí đặt thiết bị từ thắt lưng sang túi áo, độ chính xác RF4-HAR giảm khoảng 4,6%. Đây là hạn chế cần nghiên cứu thêm bằng cách thu thập dữ liệu ở nhiều vị trí đặt thiết bị khác nhau. Các giá trị "Acc tổng" trong Bảng 3.27 đều là độ chính xác tổng thể đa lớp: KNN6-HAR đạt 86,1% (WISDM) và 96,03% (tập thu thập), RF5-HAR đạt 96,5% (WISDM, cửa sổ 8 giây) và 98,3% (tập thu thập).

Bảng 3.27 So sánh tổng hợp về hiệu quả của hệ thống RF5-HAR và KNN6-HAR

Tiêu chí so sánh	**RF5-HAR**	**KNN6-HAR**	**Nhận xét**
Số hoạt động nhận dạng	5 hoạt động	**6 hoạt động**	KNN6 nhiều hơn 1HĐ
Phân biệt lên/xuống cầu thang	Không	**Có**	Ưu điểm của KNN6
Acc tổng (dữ liệu CK)	96,5%	86,1%	RF5 cao hơn
Acc tổng (dữ liệu TT)	98,3%	**96,03%**	KNN6 khá tốt

Acc lên / xuống cầu thang	N/A, gộp chung	**91,3% / 88,4% (độ nhạy, tập TT)**	KNN6 chuyên biệt
Thời gian suy luận	**50 ms**	82 ms	RF5 nhanh hơn 39%
Thời gian thực (<100 ms)	Đạt	Đạt	Cả hai đạt yêu cầu
Kích thước model	1,3 KB	**≈28,8 KB (1.200 mẫu × 12 chiều × 2 byte)**	RF5 nhỏ hơn
Ứng dụng phù hợp	Môi trường bằng phẳng hoặc không phân biệt lên xuống	**Tòa nhà cao tầng**	KNN6 phù hợp hơn

Bảng 3.28 So sánh hệ thống đề xuất với các nghiên cứu quốc tế tiêu biểu

Nghiên cứu	**Phần cứng / Nền tảng**	**RAM / Flash**	**Số HĐ**	**Đặc trưng**	**Thuật toán**	**Độ chính xác**
Kwapisz và cộng sự [17]	Android smartphone	RAM 512 MB	6	43 đặc trưng	J48, LR, MLP	91,7%
Vavoulas và cộng sự [20]	PC (Sklearn)	N/A	5	64 đặc trưng	IBk, J48	99,8%
Catal và cộng sự [25]	PC (Weka)	N/A	6	43 đặc trưng	Voting	91,6%
Gupta & Dallas [18]	Chest sensor + PC	N/A	6	15 đặc trưng	Naïve Bayes	97,9%
Gupta và cộng sự [21]	ARM Cortex-M4 (STM32)	RAM 64 KB Flash 256 KB	4	9 đặc trưng	DT	93,2%
Wang và cộng sự [98]	ESP32 (32-bit, WiFi)	RAM 520 KB Flash 4 MB	6	30 đặc trưng	RF	96,8%
Büber & Guvensan [97]	Android smartphone	RAM 1 GB	5	12 đặc trưng	SVM	91,1%

RF4-HAR (đề xuất)	PIC18F4520 8-bit MCU	RAM 1,5 KB Flash 32 KB	4	5 đặc trưng*3	RF (50 cây)	95,3% / 99,2%
RF5-HAR (đề xuất)	PIC18F4520 8-bit MCU	RAM 1,5 KB Flash 32 KB	5	5 đặc trưng*3	RF (200 cây)	96,5% / 98,3%
KNN6-HAR (đề xuất)	PIC18F4520 8-bit MCU	RAM 1,5 KB Flash 32 KB	6	4 đặc trưng*3	KNN (k=5)	86,1% / 96,0%

Ghi chú: HĐ = Hoạt động, N/A = không báo cáo, Độ chính xác ghi "X% / Y%" là X% trên tập công khai / Y% trên tập thu thập. Trong các nghiên cứu được so sánh, hệ thống đề xuất là hệ thống duy nhất chạy trên MCU 8-bit (RAM 1,5 KB), dung lượng RAM thấp hơn khoảng 43 lần so với STM32 trong [21] (64 KB). Độ chính xác của RF4-HAR là độ chính xác tổng thể đa lớp (độ chính xác theo lớp trung bình là 97,7% / 99,6%).

Bảng 3.29 Tóm tắt ba hệ thống HAR đề xuất trong luận án

Tiêu chí	**RF4-HAR (Chương 2)**	**RF5-HAR (Chương 3)**	**KNN6-HAR (Chương 3)**
Số hoạt động	4 (Đứng, Ngồi, Đi bộ, Chạy)	5 (+ Cầu thang gộp)	6 (+ Lên/Xuống riêng)
Thuật toán	Random Forest (50 cây)	Random Forest (200 cây)	KNN (k=5, distance)
Bộ đặc trưng	5 đặc trưng * 3 trục = 15	5 đặc trưng * 3 trục = 15	4 đặc trưng * 3 trục = 12
Đặc trưng	Mean, SD, Median, Range, RMS	AAD, STD, IQR, Range, RMS	Var, SAV, Skew, P2P
Cửa sổ thời gian	10 giây (500 mẫu, 50Hz)	8 giây (400 mẫu, 50Hz)	6 giây (300 mẫu, 50Hz)
Dữ liệu đánh giá	WISDM + Thu thập (16 người)	WISDM + Thu thập (16 người)	WISDM + Thu thập (16 người)
Phần cứng	PIC18F4520 + SIM808	PIC18F4520 + ESP8266	PIC18F4520 + ESP8266
Kích thước model	1,2 KB Flash	1,3 KB Flash	≈28,8 KB Flash

Thời gian suy luận	45 ms	50 ms	82 ms
Tiêu thụ điện	2,51 mA (chờ: 0,8 mA) (**)	2,51 mA (**)	2,51 mA (**)
Pin	3,7 V / 2000 mAh	3,7 V / 6000 mAh	3,7 V / 6000 mAh
Acc tập CK (WISDM)	95,3% (97,7%) (*)	96,5%	86,1%
Acc tập TT (thu thập)	99,2% (99,6%) (*)	98,3%	96,03%
Chi phí phần cứng	< 15 USD	< 20 USD	< 20 USD
Công trình công bố	[CT1], [CT2]	[CT3]	[CT4]

(CK: Công khai - tập dữ liệu WISDM, TT: Tự thu thập tại phòng thí nghiệm. Acc là độ chính xác tổng thể đa lớp; (*) độ chính xác theo lớp trung bình của RF4-HAR. (**) Dòng tiêu thụ của riêng khối PIC18F4520 + ADXL345, chưa gồm mô-đun truyền thông; dòng trung bình của toàn thiết bị khi truyền liên tục khoảng 110 mA. Kích thước model là dung lượng Flash của riêng cấu trúc mô hình, chưa gồm mã chương trình và thư viện.)

3.7. Kết luận chương 3

Chương 3 đã thực hiện thành công ba mục tiêu lớn: mở rộng số lượng hoạt động nhận dạng, triển khai thuật toán trên vi điều khiển hiệu năng thấp và đánh giá tổng hợp so sánh với các nghiên cứu quốc tế. Các đóng góp chính bao gồm:

(1) Tiếp tục mở rộng theo dõi hoạt động trong tòa nhà bằng hệ thống RF5-HAR hoạt động cơ bản bao gồm đi cầu thang [CT3], đạt độ chính xác tổng thể 98,3% trên tập dữ liệu thu thập và 96,5% trên tập WISDM.

(2) Phân tích chi tiết các điểm hạn chế của hệ thống RF5-HAR và đưa ra phương pháp cải tiến với hệ thống KNN6-HAR [CT4], cho phép phân loại 6 hoạt động với độ chính xác 96,03% trên tập dữ liệu thu thập.

Kết quả cho thấy sự cân bằng tốt giữa hiệu năng, chi phí và tiêu thụ năng lượng, tạo nền tảng vững chắc cho việc ứng dụng rộng rãi trong thực tế. Hệ thống có tiềm năng đáp ứng yêu cầu theo dõi, chăm sóc sức khỏe cho các đối tượng như người cao tuổi, trẻ em, người lao động trong môi trường rủi ro cao, người bệnh mới phục hồi hoặc bệnh nhân liên quan đến mất kiểm soát hành vi khi di chuyển trong tòa nhà cao tầng, với thiết bị

đeo nhỏ gọn, chi phí thấp và thời gian hoạt động dài. Hệ thống mới được kiểm chứng trên người khỏe mạnh trong điều kiện có kiểm soát. Công trình khoa học liên quan tới chương này là [CT3] và [CT4]; hướng ứng dụng được trình bày tại [CT6].

Như vậy, ba hệ thống đề xuất trong luận án không chỉ khác nhau về số lượng hoạt động và địa hình thử nghiệm, mà còn khác nhau về bản chất kỹ thuật và mục đích sử dụng: RF4-HAR (Chương 2) dùng thuật toán Random Forest với 5 đặc trưng thống kê, tối ưu cho địa hình bằng phẳng, phù hợp giám sát mức độ vận động hàng ngày (đứng, ngồi, đi, chạy). RF5-HAR (Chương 3, mục 3.1) mở rộng cùng thuật toán RF sang địa hình có thay đổi độ cao nhưng gộp chung hoạt động lên/xuống cầu thang thành một nhãn duy nhất, phù hợp khi chỉ cần biết người dùng có di chuyển theo phương thẳng đứng hay không (ví dụ theo dõi mức vận động tổng quát). KNN6-HAR (Chương 3, mục 3.4) chuyển sang thuật toán KNN với bộ 4 đặc trưng khác (Var, SAV, Skew, P2P), cho phép phân biệt tường minh chiều lên và chiều xuống cầu thang - thông tin cần thiết cho các ứng dụng đòi hỏi độ chi tiết cao hơn như đánh giá phục hồi chức năng hoặc phát hiện bất thường khi di chuyển trong tòa nhà nhiều tầng. Việc lựa chọn thuật toán khác nhau (RF cho bài toán ít lớp, KNN cho bài toán cần ranh giới quyết định mềm dẻo hơn giữa hai lớp gần giống nhau) xuất phát từ đặc điểm của từng bài toán, không chỉ nhằm mở rộng quy mô bài toán.

KẾT LUẬN VÀ KIẾN NGHỊ

1. Kết luận

Luận án đã hoàn thành mục tiêu nghiên cứu và phát triển hệ thống nhận dạng hoạt động người (HAR) sử dụng cảm biến gia tốc và thuật toán học máy, có khả năng phân loại chính xác các hoạt động cơ bản trong thời gian thực với chi phí thấp. Qua ba chương nghiên cứu, luận án đã đạt được những kết quả quan trọng sau:

Chương 1 đã hệ thống hóa cơ sở lý thuyết về HAR và các phương pháp học máy. Phân tích chi tiết cho thấy phương pháp sử dụng cảm biến đeo là phù hợp nhất cho ứng dụng theo dõi sức khỏe cá nhân do tính di động, không xâm phạm quyền riêng tư và chi phí thấp.

Chương 2 đã xây dựng thành công hệ thống nhận dạng 4 hoạt động cơ bản (Đứng, Ngồi, Đi bộ, Chạy bộ) với độ chính xác tổng thể đa lớp 95,3% (WISDM) và 99,2% (thu thập), tương ứng độ chính xác theo lớp trung bình 97,7% và 99,6%. Đề xuất tập đặc trưng phù hợp chỉ với 5 đặc trưng thống kê đơn giản (Mean, STD, Median, Range, RMS); xây dựng bộ dữ liệu từ 16 tình nguyện viên (14 sinh viên + 2 người cao tuổi) với quy trình chuẩn hóa; đánh giá với nhiều chỉ số (Accuracy, Sensitivity, Precision, NPV).

Chương 3 đã mở rộng và triển khai hệ thống lên mức cao hơn (RF5-HAR), mở rộng thành công từ 4 lên 5 hoạt động với độ chính xác 96,5% (WISDM) và 98,3% (dữ liệu thu thập), trong thời gian thực. Thiết kế thiết bị thu thập dữ liệu chi phí thấp (<20 USD) sử dụng ADXL345 và PIC18F4520, lựa chọn tập đặc trưng; đề xuất hệ thống KNN6-HAR nhận dạng 6 hoạt động người ở địa hình tòa nhà cao tầng (thêm phân biệt lên cầu thang và xuống cầu thang) trên dữ liệu công khai và bộ dữ liệu thu thập được thể hiện độ chính xác tổng thể (86,1% và 96,03%).

2. Đóng góp mới về khoa học của luận án

- **Đóng góp 1:** Nghiên cứu này đề xuất và triển khai thành công hệ thống HAR thời gian thực trên vi điều khiển 8-bit PIC18F4520 (RAM 1,5 KB, 2 USD), một ràng buộc tài nguyên rất giới hạn mà các nghiên cứu tiêu biểu như Kwapisz và cộng sự [17] (smartphone), Vavoulas và cộng sự [20] (64 đặc trưng) hay Catal và cộng sự [25] (43 đặc trưng) đều chưa đặt ra. Đóng góp khoa học cốt lõi là chứng minh rằng chỉ cần 5 đặc trưng

thống kê đơn giản (tính trên mỗi trục X, Y, Z, tương đương 15 chiều đầu vào) kết hợp với thuật toán phù hợp cho MCU 8-bit có thể đạt độ chính xác tổng thể trên 95% trên tập WISDM và trên 98% trên tập tự thu thập (RF4-HAR, RF5-HAR), tương đương các nghiên cứu sử dụng nhiều đặc trưng hơn trên nền tảng phần cứng mạnh hơn. Thuật toán (RF, KNN), được nhúng trực tiếp vào vi điều khiển, cho phép phân loại hoạt động theo thời gian thực.

- **Đóng góp 2: Lựa chọn hoạt động và bộ đặc trưng phù hợp với ràng buộc phần cứng.** Các hoạt động chính Đứng, Ngồi, Chạy bộ, Đi bộ và Đi lên, xuống cầu thang được lựa chọn cẩn thận dựa trên mức độ liên quan và tần suất trong cuộc sống hàng ngày. Bộ đặc trưng được lựa chọn là cấu hình cân bằng tốt nhất giữa khả năng phân biệt và độ phức tạp tính toán trong các cấu hình đã khảo sát, phù hợp MCU 8-bit.

3. Tính ứng dụng thực tiễn của luận án

Hệ thống HAR đề xuất có tiềm năng ứng dụng trong nhiều lĩnh vực. Trong y tế, hệ thống có thể hỗ trợ theo dõi hoạt động thể chất của bệnh nhân mãn tính và đánh giá khả năng phục hồi chức năng sau phẫu thuật hoặc chấn thương. Trong thể thao, hệ thống giúp huấn luyện viên theo dõi khối lượng và cường độ tập luyện thông qua khả năng phân biệt đi bộ, chạy bộ và lên xuống cầu thang. Trong lĩnh vực an ninh và lao động, hệ thống có thể hỗ trợ giám sát an toàn cho nhân viên làm việc độc lập (bảo vệ, cứu hộ, kiểm lâm) thông qua phát hiện bất động kéo dài bất thường. Các hướng ứng dụng chi tiết hơn đã được trình bày trong công trình [CT6] của tác giả và có thể được phát triển thành các nghiên cứu mở rộng riêng.

4. Hạn chế của luận án và hướng nghiên cứu tiếp theo

4.1. Hạn chế

Mặc dù đạt được nhiều kết quả tích cực, luận án vẫn còn một số hạn chế cần khắc phục:

Về bộ dữ liệu: Dữ liệu tự thu thập gồm 16 tình nguyện viên, trong đó nhóm người cao tuổi chỉ có 2/16 người (12,5%, độ tuổi 65-72). Đây là hạn chế đáng chú ý vì số lượng quá nhỏ chưa đủ để đại diện cho sự đa dạng về thể trạng và bệnh lý trong nhóm người cao tuổi. Hơn nữa, dữ liệu được thu thập trong môi trường kiểm soát (phòng thí nghiệm,

sân tập), chưa đại diện đầy đủ cho điều kiện thực tế đa dạng. Toàn bộ người tham gia là người khỏe mạnh; hệ thống chưa được thử nghiệm trên bệnh nhân có rối loạn vận động hoặc người có nguy cơ té ngã cao. Ngoài ra, việc chia dữ liệu ngẫu nhiên theo cửa sổ (không theo đối tượng, không có tập kiểm định riêng, các cửa sổ chồng lấp 40%) có thể làm kết quả lạc quan hơn so với đánh giá độc lập theo đối tượng.

Về số lượng hoạt động: Hệ thống mới nhận dạng được 6 hoạt động cơ bản, chưa bao gồm các hoạt động phức tạp hơn như đạp xe, bơi lội, chơi thể thao, làm việc nhà. Để trở thành hệ thống HAR toàn diện, cần mở rộng lên 15-20 hoạt động.

Về độ chính xác: RF4-HAR và RF5-HAR đạt độ chính xác tổng thể đa lớp từ 95,3% đến 99,2%, nhưng KNN6-HAR chỉ đạt 86,1% trên tập dữ liệu công khai (so với 96,03% trên tập tự thu thập), cho thấy khả năng khái quát hóa sang dữ liệu ngoài chưa đồng đều giữa các hệ thống. Các cặp hoạt động còn nhầm lẫn đáng kể gồm Xuống/Lên cầu thang (trên WISDM, 14,7% mẫu Xuống bị nhận thành Lên và 8,6% mẫu Lên bị nhận thành Xuống - Bảng 3.22) và Đi cầu thang/Đi bộ (7,2% với RF5-HAR trên WISDM - Hình 3.6). Ngoài ra, RF chưa được đánh giá trên cùng bộ đặc trưng của KNN6-HAR. Với ứng dụng y tế, cần nâng độ chính xác của các lớp cầu thang và cải thiện khả năng khái quát hóa của KNN6-HAR trên dữ liệu ngoài (chuẩn hóa đặc trưng, cân bằng lớp).

Về thử nghiệm thực tế: Chưa có thử nghiệm dài hạn (>3 tháng) với nhóm người dùng lớn trong điều kiện thực tế. Cần đánh giá độ tin cậy, độ bền thiết bị và mức độ chấp nhận của người dùng.

Cụ thể hơn, luận án chưa đánh giá khả năng chịu va đập của vỏ thiết bị khi người dùng vận động mạnh hoặc va chạm bất ngờ, chưa kiểm tra hoạt động ổn định của cảm biến và vi điều khiển trong dải nhiệt độ môi trường rộng (ví dụ ngoài trời mùa hè/mùa đông), và dây đeo/vỏ thiết bị hiện tại (dùng trong thử nghiệm) chưa được tối ưu về tính linh động, độ thoáng khí khi đeo liên tục trong thời gian dài, có thể gây khó chịu hoặc lệch vị trí cảm biến, ảnh hưởng đến độ chính xác nhận dạng (phân tích tại Mục 2.2.4). Việc đeo đai cố định ở thắt lưng cũng có thể gây bất tiện với người cao tuổi, người có bệnh lý vùng bụng hoặc khi mặc trang phục không có thắt lưng; giải pháp là thiết kế vỏ dạng kẹp, dây đeo co giãn hoặc tích hợp vào quần áo. Đây là các yếu tố cần được kiểm

định và cải tiến (ví dụ vỏ chống sốc IP-rated, dây đeo co giãn) trước khi triển khai sản phẩm thương mại.

4.2. Hướng nghiên cứu tiếp theo

- Mở rộng bộ dữ liệu: Thu thập thêm 100-200 người ở các nhóm tuổi khác nhau (trẻ em, trung niên, cao tuổi), nhiều thể trạng (gầy, béo, bệnh lý xương khớp).

- Mở rộng hoạt động: Thêm 4-6 hoạt động nữa (đạp xe, đi cầu thang máy, ngồi xe, làm việc nhà) lên tổng 10-12 hoạt động.

- Nghiên cứu phát hiện, dự đoán bất thường: Phát triển module phát hiện ngã, hoạt động bất thường, cảnh báo khẩn cấp.

- Nghiên cứu cross-validation độc lập chủ thể: Thực hiện leave-one-subject-out validation để đánh giá khả năng tổng quát hóa của mô hình với người dùng mới, đồng thời thu thập dữ liệu từ nhóm bệnh nhân có bệnh lý vận động (Parkinson, sau đột quỵ, xương khớp) để mở rộng phạm vi ứng dụng y tế.

- Hoàn thiện đánh giá mô hình: huấn luyện RF trên cùng bộ đặc trưng của KNN6-HAR để tách bạch ảnh hưởng của thuật toán và bộ đặc trưng; áp dụng chuẩn hóa đặc trưng thống nhất cho KNN và SVM; báo cáo độ chính xác của mô hình KNN nén chạy trên thiết bị; chia ba tập huấn luyện/kiểm định/kiểm thử.

- Nâng cấp nền tảng phần cứng: Nghiên cứu chuyển đổi sang MCU thế hệ mới (ARM Cortex-M0+, 3 USD) để hỗ trợ nhiều hoạt động hơn với mô hình phức tạp hơn, trong khi vẫn duy trì chi phí thấp và khả năng hoạt động thời gian thực với độ trễ <100 ms.

- Thử nghiệm ngoài phòng thí nghiệm: Triển khai hệ thống thử nghiệm dài hạn (3-6 tháng) với ít nhất 50 người dùng thực tế trong các điều kiện môi trường đa dạng (ngoài trời, có vật cản, điều kiện khí hậu khác nhau) nhằm đánh giá độ tin cậy và khả năng chấp nhận của người dùng.

- Nhận dạng giai đoạn chuyển trạng thái (Transition Activity Recognition): Hệ thống hiện tại phân loại 6 hoạt động ở trạng thái ổn định. Hướng phát triển tiếp theo là nhận dạng các giai đoạn chuyển trạng thái (ngồi-đứng, đứng-đi bộ, đứng-nằm…), đây là cơ sở khoa học để phát hiện các sự kiện rủi ro trong y tế. Cụ thể: (i) Té ngã có thể được mô hình hóa như transition đột ngột từ trạng thái đứng/đi sang nằm với gia tốc đỉnh bất

thường, (ii) Bài kiểm tra lâm sàng Sit-to-Stand (STS) và Timed-Up-and-Go (TUG), tiêu chuẩn vàng đánh giá chức năng vận động người cao tuổi, đòi hỏi nhận dạng chính xác thời điểm bắt đầu và kết thúc của từng giai đoạn chuyển tư thế. Phương án kỹ thuật: chạy song song cửa sổ ngắn 1-2 giây để bắt transition bên cạnh cửa sổ 8-10 giây cho trạng thái ổn định, hoặc dùng mô hình hai tầng (tầng 1 phát hiện thay đổi trạng thái, tầng 2 phân loại chi tiết). Tập dữ liệu tham khảo: UCI-HAPT (mở rộng UCI-HAR với 6 lớp transition).

5. Kết luận chung

Luận án đã hoàn thành mục tiêu nghiên cứu ban đầu: xây dựng hệ thống nhận dạng hoạt động người sử dụng cảm biến gia tốc và nghiên cứu phát triển một số thuật toán học máy, nhằm nâng cao khả năng phân loại chính xác các hoạt động cơ bản trong thời gian thực với chi phí thấp. Các kết quả đạt được như sau:

- Độ chính xác tổng thể đa lớp (WISDM / tự thu thập): 95,3% và 99,2% cho 4 hoạt động (RF4-HAR; độ chính xác theo lớp trung bình 97,7% và 99,6%), 96,5% và 98,3% cho 5 hoạt động (RF5-HAR), 86,1% và 96,03% cho 6 hoạt động (KNN6-HAR)

- Chi phí: <20 USD (đúng mục tiêu)

Kết quả của luận án thể hiện sự cân bằng giữa ba yếu tố: hiệu năng (độ chính xác tổng thể 95,3-99,2% đối với RF4-HAR và RF5-HAR), chi phí (< 20 USD) và tiêu thụ năng lượng (2,51 mA đối với khối vi điều khiển và cảm biến). So với các nghiên cứu quốc tế, luận án đạt độ chính xác tương đương với phần cứng hạn chế hơn nhiều.

Về mặt khoa học: Luận án đóng góp quy trình lựa chọn bộ đặc trưng tối giản và quy trình lựa chọn thuật toán theo ràng buộc phần cứng của vi điều khiển 8-bit; đánh giá hiệu suất bằng cả trung bình vĩ mô và trung bình vi mô theo định nghĩa thống nhất. Trong bài toán sáu hoạt động, KNN6-HAR đạt độ chính xác tổng thể 86,1% trên tập dữ liệu công khai và 96,03% trên tập dữ liệu thu thập, độ nhạy trung bình vĩ mô 82,1% và 95,06%. Kết quả 86,1% trên tập công khai còn khiêm tốn hơn RF4-HAR và RF5-HAR, phản ánh hạn chế về khả năng khái quát hóa đã nêu ở Mục 4.1. KNN được lựa chọn để phân biệt lên/xuống cầu thang vì hai hoạt động này tạo thành các cụm cục bộ gần nhau trong không gian đặc trưng 12 chiều (Var, SAV, Skew, P2P trên ba trục) với ranh giới

phi tuyến, phù hợp với cơ chế phân loại theo khoảng cách của KNN hơn các phép chia song song trục của cây quyết định.

Một yếu tố kỹ thuật quan trọng ảnh hưởng đến hiệu năng của KNN6-HAR là sự chênh lệch biên độ giữa các đặc trưng (SAV cỡ 10^2-10^3 so với Skew cỡ 10^0). Như đã trình bày tại Mục 3.5.3, chuẩn hóa Min-Max giúp nâng độ chính xác tổng thể của KNN6-HAR lên 88,7% (WISDM) và 96,8% (thu thập) với chi phí tính toán khoảng 0,5 ms/phân loại; luận án khuyến nghị áp dụng chuẩn hóa này trong các triển khai thực tế.

Về mặt thực tiễn: luận án cho thấy tiềm năng ứng dụng trong y tế, thể thao và chăm sóc người cao tuổi. Hệ thống có tiềm năng tiếp tục phát triển theo hướng sản phẩm sau khi được kiểm chứng trên quy mô người dùng lớn hơn, trong điều kiện thực tế và trên nhóm người có bệnh lý vận động.

Tóm lại, luận án đã đạt được mục tiêu nghiên cứu đề ra, có đóng góp khoa học và thực tiễn rõ ràng, tạo cơ sở cho các nghiên cứu tiếp theo về HAR và ứng dụng IoT trong chăm sóc sức khỏe.

DANH MỤC CÁC CÔNG TRÌNH CỦA TÁC GIẢ

TẠP CHÍ KHOA HỌC

[CT1]	Q. H. Nguyen, D. N. Tran, T. H. Dao, D. T. Tran, **T. T. Vu**, and X. T. Do, "A method of building a real-time human behavior recognition system based on machine learning algorithms," *Res. Dev. Inf. Commun. Technol.*, vol. 2022, no. 2, pp. 83-88, 2022, doi: 10.32913/mic-ict-research-vn.v2022.n2.1137. ISSN: 1859-3526.
[CT2]	**T. T. Vu**, T. T. Nguyen, D. N. Tran, N. S. Nguyen, and D. T. Tran, "An IoT solution that classifies human physical activities based on sensors and supervised learning algorithms," *Res. Dev. Inf. Commun. Technol.*, vol. 2024, no. 2, pp. 74-80, Nov. 2024, doi: 10.32913/mic-ict-research-vn.v2024.n2.1292. ISSN: 1859-3526, eISSN: 1859-3534.
[CT3]	M. T. Vi, D. N. Tran, **T. T. Vu**, N. L. Nguyen, and D. T. Tran, "Efficient real-time devices based on accelerometer using machine learning for HAR on low-performance microcontrollers," *Comput. Mater. Contin.*, vol. 81, no. 1, pp. 1729-1756, 2024, doi: 10.32604/cmc.2024.055511. ISSN: 1546-2218, eISSN: 1546-2226 (SCIE, Q2).
[CT4]	**T. T. Vu**, D. N. Tran, D. T. Tran, T. T. Bui, D. T. Vu, T. P. A. Nguyen, C. P. K. Phung, K. T. Pham, and M. T. Vi, "Effective human activity recognition through accelerometer data," *Eng. Technol. Appl. Sci. Res.*, vol. 14, no. 5, pp. 16499-16510, Oct. 2024, doi: 10.48084/etasr.8211. ISSN: 2241-4487, eISSN: 1792-8036 (Scopus, Q2).

HỘI NGHỊ KHOA HỌC

[CT5]	**T. T. Vu**, D. T. Tran, P. P. Nguyen, N. S. Nguyen, and D. N. Tran, "Giải pháp IoT phát hiện các hoạt động thể chất của con người để hỗ trợ đánh giá tình trạng sức khỏe dựa trên cảm biến và thuật toán học máy có giám sát," in *Kỷ yếu Hội thảo Quốc gia VNICT 2025*, Ninh Bình, Việt Nam, Nov. 2025, pp. 361-366. ISBN: 978-604-45-1878-7.

SÁNG CHẾ

[CT6]	D. T. Tran, H. D. Vu, **T. T. Vu**, M. T. Vi and D. N. Tran, "Hệ thống phát hiện, cảnh báo và bảo vệ người mộng du sử dụng thiết bị đeo cổ tay, đai an toàn và khóa thông minh," Bằng độc quyền sáng chế số 57548, cấp theo Quyết định số 47085/QĐ-SHTT, ngày 26/03/2026.

Ghi chú về đóng góp của tác giả trong các công trình đồng tác giả: Ở [CT1], tác giả tham gia xây dựng bộ dữ liệu tự thu thập, thực hiện thực nghiệm so sánh các thuật toán RF, SVM, GBDT, KNN và viết phần kết quả. Ở [CT3], tác giả thiết kế quy trình thu thập dữ liệu và trích chọn đặc trưng cho hệ thống RF5-HAR, triển khai thực nghiệm trên thiết bị và tham gia viết bản thảo. Tác giả là tác giả chính của các công trình [CT2], [CT4], [CT5]. Đối chiếu công trình - chương: Chương 2 ([CT1], [CT2], [CT5]); Chương 3 ([CT3] - RF5-HAR, [CT4] - KNN6-HAR); hướng ứng dụng ([CT6]).

TÀI LIỆU THAM KHẢO

[1] World Health Organization, WHO Guidelines on Physical Activity and Sedentary Behaviour. Geneva, Switzerland: WHO, 2020.

[2] S. Chernbumroong et al., “Elderly activities recognition and classification for applications in assisted living,” Expert Syst. Appl., vol. 40, no. 5, pp. 1662-1674, 2013, doi: 10.1016/j.eswa.2012.09.004.

[3] K. K. B. Peetoom, M. A. S. Lexis, M. Joore, C. D. Dirksen, and L. P. De Witte, “Literature review on monitoring technologies and their outcomes in independently living elderly people,” Disabil. Rehabil. Assist. Technol., vol. 10, no. 4, pp. 271-294, 2015, doi: 10.3109/17483107.2014.961179.

[4] P. Pierleoni, A. Belli, L. Palma, M. Pellegrini, L. Pernini, and S. Valenti, “A high reliability wearable device for elderly fall detection,” IEEE Sens. J., vol. 15, no. 8, pp. 4544-4553, 2015, doi: 10.1109/JSEN.2015.2423562.

[5] V. T. Pham et al., “Development of a real-time, simple and high-accuracy fall detection system for elderly using 3-DOF accelerometers,” Arab. J. Sci. Eng., 2018.

[6] M. Sundholm et al., “Smart-mat: Recognizing and counting gym exercises with low-cost resistive pressure sensing matrix,” in Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp), 2014, pp. 373-382, doi: 10.1145/2632048.2636088.

[7] T. Stiefmeier et al., “Wearable activity tracking in car manufacturing,” IEEE Pervasive Comput., vol. 7, no. 2, pp. 42-50, 2008.

[8] S. C. Akkaladevi and C. Heindl, “Action recognition for human robot interaction in industrial applications,” in Proc. IEEE Int. Conf. Comput. Graph., Vis. Inf. Secur. (CGVIS), 2015, doi: 10.1109/CGVIS.2015.7449900.

[9] A. Wang et al., “A comparative study on human activity recognition using inertial sensors in a smartphone,” IEEE Sens. J., vol. 16, no. 11, pp. 4566-4578, 2016, doi: 10.1109/JSEN.2016.2545708.

[10] S. Khan et al., “Human action recognition: A paradigm of best deep learning features selection and serial based extended fusion,” Sensors, vol. 21, no. 23, p. 7941, 2021, doi: 10.3390/s21237941.

[11] B. Goodarzi, R. Javidan, and M. S. Rezaei, “A vision-based method for human activity recognition using local binary pattern,” in Proc. Int. Conf. Comput. Knowl. Eng. (ICCKE), 2023, pp. 209-215, doi: 10.1109/ICCKE60553.2023.10326220.

[12] C. W. Lin and Z. H. Ling, “Automatic fall incident detection in compressed video for intelligent homecare,” in Proc. Int. Conf. Comput. Commun. Netw. (ICCCN), 2007, pp. 1172-1177, doi: 10.1109/ICCCN.2007.4317978.

[13] S. Torres, “Aging in place: Fall detection and localization in a distributed smart camera network,” in Critical Gerontology Comes of Age: Advances in Research and Theory for a New Century, 2018, pp. 151-163, doi: 10.4324/9781315209371.

[14] P. Hevesi et al., “Monitoring household activities and user location with a cheap, unobtrusive thermal sensor array,” in Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp), 2014, pp. 141-145, doi: 10.1145/2632048.2636084.

[15] C. Zhu and W. Sheng, “Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living,” IEEE Trans. Syst., Man, Cybern. A, Syst. Humans, vol. 41, no. 3, pp. 569-573, 2011, doi: 10.1109/TSMCA.2010.2093883.

[16] J. Parkka et al., “Activity classification using realistic data from wearable sensors,” IEEE Trans. Inf. Technol. Biomed., vol. 10, no. 1, pp. 119-128, 2006, doi: 10.1109/TITB.2005.856863.

[17] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, “Activity recognition using cell phone accelerometers,” ACM SIGKDD Explor. Newsl., vol. 12, no. 2, pp. 74-82, 2011, doi: 10.1145/1964897.1964918.

[18] P. Gupta and T. Dallas, “Feature selection and activity recognition system using a single triaxial accelerometer,” IEEE Trans. Biomed. Eng., vol. 61, no. 6, pp. 1780-1786, 2014, doi: 10.1109/TBME.2014.2307069.

[19] A. M. Khan, Y. K. Lee, S. Y. Lee, and T. S. Kim, “A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical

recognizer," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 5, pp. 1166-1172, 2010, doi: 10.1109/TITB.2010.2051955.

[20] G. Vavoulas et al., "The MobiAct dataset: Recognition of activities of daily living using smartphones," in Proc. Int. Conf. Inf. Commun. Technol. Ageing Well e-Health (ICT4AWE), 2016, pp. 143-151, doi: 10.5220/0005792401430151.

[21] S. Gupta et al., "A TinyML approach to human activity recognition," J. Phys.: Conf. Ser., vol. 2273, no. 1, p. 012025, May 2022.

[22] P. P. Ray, "A review on TinyML: State-of-the-art and prospects," J. King Saud Univ. - Comput. Inf. Sci., vol. 34, no. 4, pp. 1595-1623, 2022.

[23] V. L. Trần, "Phát triển các module phần mềm để xác thực đối tượng và nhận dạng hành vi của đối tượng kết hợp giữa kỹ thuật xử lý ảnh với điện tử," Đề tài nghiên cứu khoa học cấp Viện, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, 2014.

[24] V. C. Phạm, "Nghiên cứu chế tạo thiết bị đeo thông minh nhận dạng hoạt động của người (iShoes)," Đề tài nghiên cứu khoa học, Bộ Thông tin và Truyền thông, 2017.

[25] C. Catal, S. Tufekci, E. Pirmit, and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity recognition," Appl. Soft Comput., vol. 37, pp. 1018-1022, 2015, doi: 10.1016/j.asoc.2015.01.025.

[26] L. Breiman, "Random forests," Mach. Learn., vol. 45, no. 1, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.

[27] M. A. Khan et al., "A fused heterogeneous deep neural network and robust feature selection framework for human actions recognition," Arab. J. Sci. Eng., Jul. 2021, doi: 10.1007/s13369-021-05881-4.

[28] A. Jain and V. Kanhangad, "Exploring orientation and accelerometer sensor data for personal authentication in smartphones using touchscreen gestures," Pattern Recognit. Lett., vol. 68, pp. 351-360, 2015, doi: 10.1016/j.patrec.2015.07.004.

[29] N. D. Nguyen et al., "A classifier based approach to real-time fall detection using low-cost wearable sensors," in Proc. 5th Symp. Inf. Commun. Technol. (SoICT), Hanoi, Vietnam, 2014, pp. 14-20, doi: 10.1145/2676585.2676597.

[30] M. A. Khan et al., “A resource conscious human action recognition framework using 26-layered deep convolutional neural network,” Multimed. Tools Appl., vol. 80, no. 28-29, pp. 35827-35849, Nov. 2021, doi: 10.1007/s11042-020-09408-1.

[31] M. Ahmed et al., “Real-time violent action recognition using key frames extraction and deep learning,” Comput. Mater. Contin., vol. 69, no. 2, pp. 2217-2230, 2021, doi: 10.32604/cmc.2021.018103.

[32] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in Proc. 17th Conf. Innov. Appl. Artif. Intell. (IAAI), 2005, pp. 1541-1546.

[33] L. Bao and S. S. Intille, “Activity recognition from user-annotated acceleration data,” in Pervasive Computing (Lecture Notes in Computer Science), vol. 3001, 2004, pp. 1-17.

[34] B. H. Dobkin et al., “Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke,” Stroke, vol. 42, no. 8, pp. 2246-2250, 2011.

[35] I. M. Nasir, M. Raza, J. H. Shah, M. A. Khan, and A. Rehman, “Human action recognition using machine learning in uncontrolled environment,” in Proc. Int. Conf. Artif. Intell. Data Anal. (CAIDA), 2021, pp. 182-187, doi: 10.1109/CAIDA51941.2021.9425202.

[36] S. Kiran et al., “Multi-layered deep learning features fusion for human action recognition,” Comput. Mater. Contin., vol. 69, no. 3, pp. 4061-4075, 2021.

[37] Đ. Đ. Trần, “Nghiên cứu, ứng dụng công nghệ thông tin trong xây dựng hệ thống giám sát hỗ trợ chăm sóc bệnh nhân tại phòng điều trị thông minh,” Đề tài nghiên cứu khoa học, Bộ Giáo dục và Đào tạo, 2013.

[38] D. Roggen et al., “Collecting complex activity datasets in highly rich networked sensor environments,” in Proc. 7th Int. Conf. Networked Sensing Systems (INSS), Kassel, Germany, 2010, pp. 233-240.

[39] Y. Zhao, R. Yang, G. Chevalier, X. Xu, and Z. Zhang, “Deep residual Bidir-LSTM for human activity recognition using wearable sensors,” Math. Probl. Eng., vol. 2018, pp. 1-13, Dec. 2018, doi: 10.1155/2018/7316954.

[40] R. Singh, A. Kushwaha, Chandni, and R. Srivastava, “Recent trends in human activity recognition - A comparative study,” Cogn. Syst. Res., vol. 77, pp. 30-44, 2023, doi: 10.1016/j.cogsys.2022.10.003.

[41] S. Suh, V. F. Rey, and P. Lukowicz, “TASKED: Transformer-based adversarial learning for human activity recognition using wearable sensors via self-knowledge distillation,” Knowl.-Based Syst., vol. 260, p. 110143, 2023.

[42] J. Park, W.-S. Lim, D.-W. Kim, and J. Lee, “GTSNet: Flexible architecture under budget constraint for real-time human activity recognition from wearable sensor,” Eng. Appl. Artif. Intell., vol. 124, p. 106543, 2023.

[43] S. Ullah, M. Pirahandeh, and D.-H. Kim, “Self-attention deep ConvLSTM with sparse-learned channel dependencies for wearable sensor-based human activity recognition,” Neurocomputing, vol. 571, p. 127157, 2024.

[44] Q. Teng, Y. Tang, and G. Hu, “Large receptive field attention: An innovation in decomposing large-kernel convolution for sensor-based activity recognition,” IEEE Sens. J., vol. 24, no. 8, pp. 13488-13499, 2024.

[45] Z. Yang, K. Li, and Z. Huang, “MFCANN: A feature diversification framework based on local and global attention for human activity recognition,” Eng. Appl. Artif. Intell., vol. 133, p. 108110, 2024.

[46] W. Wei and Z. Wang, “TCN-attention-HAR: Human activity recognition based on attention mechanism time convolutional network,” Sci. Rep., vol. 14, p. 7414, 2024, doi: 10.1038/s41598-024-57912-3.

[47] M. Ezzeldin et al., “Multi-modal hybrid hierarchical classification approach with transformers to enhance complex human activity recognition,” Signal Image Video Process., 2024, doi: 10.1007/s11760-024-03552-z.

[48] Z. Hong et al., “CrossHAR: Generalizing cross-dataset human activity recognition via hierarchical self-supervised pretraining,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., 2024, doi: 10.1145/3659597.

[49] S. Taghanaki, M. Rainbow, and A. Etemad, “Self-supervised human activity recognition with localized time-frequency contrastive representation learning,” IEEE Trans. Human-Mach. Syst., vol. 53, no. 6, pp. 1027-1037, 2023.

[50] Y. Kim, J. Lim, and J. Lee, “Human activity recognition via temporal fusion contrastive learning,” IEEE Access, vol. 12, pp. 20854-20866, 2024.

[51] D. A. Nguyen, C. Pham, and K. N. A. Le, “Virtual fusion with contrastive learning for single-sensor-based activity recognition,” IEEE Sens. J., vol. 24, no. 15, pp. 25041-25048, 2024.

[52] M. Duan et al., “Wearable-based behaviour interpolation for semi-supervised human activity recognition,” SSRN, 2024, Art. no. 4342069.

[53] M. F. Abdi, B. BabaAli, and S. Momeni, “An unsupervised statistical representation learning method for human activity recognition,” Signal Image Video Process., vol. 18, pp. 7041-7052, 2024.

[54] S. O. Bursa, O. D. Incel, and G. I. Alptekin, “Personalized and motion-based human activity recognition with transfer learning and compressed deep learning models,” Comput. Electr. Eng., vol. 109, p. 108777, 2023.

[55] S. O. Bursa, O. D. İncel, and G. I. Işıklar Alptekin, “Building lightweight deep learning models with TensorFlow Lite for human activity recognition on mobile devices,” Ann. Telecommun., vol. 78, no. 11, pp. 687-702, 2023.

[56] X. Gong, X. Zhang, and N. Li, “Lightweight human activity recognition method based on the MobileHARC model,” Syst. Sci. Control Eng., vol. 12, no. 1, p. 2328549, 2024, doi: 10.1080/21642583.2024.2328549.

[57] R. A. Praba and L. Suganthi, “HARNet: Automatic recognition of human activity from mobile health data using CNN and transfer learning of LSTM with SVM,” Automatika, vol. 65, no. 1, pp. 167-178, 2024, doi: 10.1080/00051144.2023.2290736.

[58] M. Ghalan and R. K. Aggarwal, “Novel human activity recognition by graph engineered ensemble deep learning model,” IFAC J. Syst. Control, vol. 27, p. 100253, 2024.

[59] J. Yu, L. Zhang, D. Cheng, W. Huang, H. Wu, and A. Song, “Improving human activity recognition with wearable sensors through BEE: Leveraging early exit and gradient boosting,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 32, pp. 3452-3464, 2024.

[60] P. Lalwani and G. Ramasamy, “Human activity recognition using a multi-branched CNN-BiLSTM-BiGRU model,” Appl. Soft Comput., vol. 154, p. 111344, 2024.

[61] S. Batool, M. H. Khan, and M. S. Farid, “An ensemble deep learning model for human activity analysis using wearable sensory data,” Appl. Soft Comput., vol. 159, p. 111599, 2024.

[62] K. Liu, C. Gao, B. Li, and W. Liu, “Human activity recognition through deep learning: Leveraging unique and common feature fusion in wearable multi-sensor systems,” Appl. Soft Comput., vol. 151, p. 111146, 2024.

[63] J. Chadha et al., “Hybrid deep learning approaches for human activity recognition and postural transitions using mobile device sensors,” SN Comput. Sci., vol. 5, p. 925, 2024, doi: 10.1007/s42979-024-03300-7.

[64] S. Aggarwal, G. Bhola, and D. K. Vishwakarma, “Weighted voting ensemble of hybrid CNN-LSTM models for vision-based human activity recognition,” Multimed. Tools Appl., 2024, doi: 10.1007/s11042-024-19582-1.

[65] S. Imanzadeh, J. Tanha, and M. Jalili, “Ensemble of deep learning techniques to human activity recognition using smart phone signals,” Multimed. Tools Appl., 2024, doi: 10.1007/s11042-024-18935-0.

[66] Z. Zhang, L. Yu, J. Gao, and Q. Ni, “A multi-channel hybrid deep learning framework for multi-sensor fusion enabled human activity recognition,” Alex. Eng. J., vol. 91, pp. 472-485, 2024.

[67] K. Zhang, Q. Wang, X. Meng, and J. Wang, “A human activity recognition scheme using mobile smartphones based on varying orientations and positions,” IEEE Sens. J., vol. 24, no. 10, pp. 17127-17139, 2024.

[68] L. Xiao et al., “A hybrid deep approach to recognizing student activity and monitoring health physique based on accelerometer data from smartphones,” Sci. Rep., vol. 14, p. 14006, 2024, doi: 10.1038/s41598-024-63934-8.

[69] Y. Vurgun and M. S. Kiran, “Recognizing salat activity using deep learning models via smartwatch sensors,” Neural Comput. Appl., vol. 36, pp. 19449-19467, 2024, doi: 10.1007/s00521-024-10195-1.

[70] M. Zaher et al., “Unlocking the potential of RNN and CNN models for accurate rehabilitation exercise classification on multi-datasets,” Multimed. Tools Appl., 2024, doi: 10.1007/s11042-024-19092-0.

[71] Y. Albadawi and T. Shanableh, “Hand-crafted features with a simple deep learning architecture for sensor-based human activity recognition,” IEEE Sens. J., vol. 24, no. 17, pp. 28300-28313, 2024, doi: 10.1109/JSEN.2024.3422272.

[72] A. M. Helmi, M. A. Al-qaness, A. Dahou, and M. Abd Elaziz, “Human activity recognition using marine predators algorithm with deep learning,” Future Gener. Comput. Syst., vol. 142, pp. 340-350, 2023.

[73] S. Ankalaki and M. Thippeswamy, “Optimized convolutional neural network using hierarchical particle swarm optimization for sensor based human activity recognition,” SN Comput. Sci., vol. 5, no. 5, p. 447, 2024.

[74] E. El-Adawi, E. Essa, M. Handosa, and S. Elmougy, “Wireless body area sensor networks based human activity recognition using deep learning,” Sci. Rep., vol. 14, no. 1, p. 2702, 2024.

[75] D. Hussein and G. Bhat, “SensorGAN: A novel data recovery approach for wearable human activity recognition,” ACM Trans. Embed. Comput. Syst., 2024, doi: 10.1145/3609425.

[76] L. Arrotta, G. Civitarese, and C. Bettini, “Semantic loss: A new neuro-symbolic approach for context-aware human activity recognition,” Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., 2024, doi: 10.1145/3631407.

[77] H. A. Imran, Q. Riaz, M. Hussain, H. Tahir, and R. Arshad, “Smart-wearable sensors and CNN-BiGRU model: A powerful combination for human activity recognition,” IEEE Sens. J., vol. 24, no. 2, pp. 1963-1974, 2024, doi: 10.1109/JSEN.2023.3338264.

[78] S. Waghchaware and R. Joshi, “Machine learning and deep learning models for human activity recognition in security and surveillance: A review,” Knowl. Inf. Syst., vol. 66, no. 8, pp. 4405-4436, 2024, doi: 10.1007/s10115-024-02122-6.

[79] H. A. Ullah, S. Letchmunan, M. S. Zia, U. M. Butt, and F. H. Hassan, “Analysis of deep neural networks for human activity recognition in videos - A systematic literature review,” IEEE Access, vol. 9, pp. 126366-126387, 2021.

[80] N. Alzahrani, J. Cała, and P. Missier, “Experience: A comparative analysis of multivariate time-series generative models: A case study on human activity data,” J. Data Inf. Qual., 2024, doi: 10.1145/3688393.

[81] T. H. Đào, V. C. Ngô, Q. H. Nguyễn, Đ. N. Trần, and Đ. T. Trần, “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,” Res. Dev. Inf. Commun. Technol., vol. 2021, no. 2, pp. 69-76, 2021, doi: 10.32913/mic-ict-research-vn.v2021.n2.1004.

[82] Q. H. Nguyễn, Đ. N. Trần, T. H. Đào, Đ. T. Trần, T. T. Vũ, and X. T. Đỗ, “Một phương pháp xây dựng hệ thống nhận dạng hành vi của người trong thời gian thực dựa trên các thuật toán học máy,” Res. Dev. Inf. Commun. Technol., vol. 2022, no. 2, pp. 83-88, 2022, doi: 10.32913/mic-ict-research-vn.v2022.n2.1137.

[83] T. T. T. Nguyễn, M. T. Vi, Đ. T. Trần, and Đ. N. Trần, “Sử dụng học máy có giám sát phân loại hoạt động cơ bản của người dựa trên dữ liệu cảm biến gia tốc,” in Kỷ yếu Hội nghị Quốc gia lần thứ 26 về Điện tử, Truyền thông và Công nghệ Thông tin (REV-ECIT 2023), 2023, pp. 300-305.

[84] Y. Zheng, W. K. Wong, X. Guan, and S. Trost, “Physical activity recognition from accelerometer data using a multi-scale ensemble method,” in Proc. 25th Innov. Appl. Artif. Intell. Conf. (IAAI), 2013.

[85] H. Gjoreski and M. Gams, “Accelerometer data preparation for activity recognition,” in Proc. Int. Multiconf. Inf. Soc., Ljubljana, Slovenia, 2011, vol. 1014.

[86] M. Jiang, H. Shang, Z. Wang, H. Li, and Y. Wang, “A method to deal with installation errors of wearable accelerometers for human activity recognition,” Physiol. Meas., vol. 32, no. 3, p. 347, 2011.

[87] C. Zhu and W. Sheng, “Motion- and location-based online human daily activity recognition,” Pervasive Mob. Comput., vol. 7, no. 2, pp. 256-269, 2011.

[88] P. Siirtola and J. Roning, “User-independent human activity recognition using a mobile phone: Offline recognition vs. real-time on device recognition,” in Distributed Computing and Artificial Intelligence. Berlin, Germany: Springer, 2012, pp. 617-627.

[89] C. S. Hemalatha and V. Vaidehi, “Frequent bit pattern mining over triaxial accelerometer data streams for recognizing human activities and detecting fall,” Procedia Comput. Sci., vol. 19, pp. 56-63, 2013.

[90] A. Mannini, S. S. Intille, M. Rosenberger, A. M. Sabatini, and W. Haskell, “Activity recognition using a single accelerometer placed at the wrist or ankle,” Med. Sci. Sports Exerc., vol. 45, no. 11, p. 2193, 2013.

[91] L. Gao, A. K. Bourke, and J. Nelson, “Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems,” Med. Eng. Phys., vol. 36, no. 6, pp. 779-785, 2014.

[92] N. Long, Y. Lei, L. Peng, P. Xu, and P. Mao, “A scoping review on monitoring mental health using smart wearable devices,” Math. Biosci. Eng., vol. 19, pp. 7899-7919, 2022, doi: 10.3934/mbe.2022369.

[93] O. Banos, J. M. Galvez, M. Damas, H. Pomares, and I. Rojas, “Window size impact in human activity recognition,” Sensors, vol. 14, no. 4, pp. 6474-6499, 2014, doi: 10.3390/s140406474.

[94] Y. Min, Y. Yin, and K. Khin, “Comparing the performance of machine learning algorithms for human activities recognition using WISDM dataset,” Int. J. Comput., vol. 4523, pp. 61-72, 2020.

[95] Y. Chen and Y. Xue, “A deep learning approach to human activity recognition based on single accelerometer,” in Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC), Oct. 2015, pp. 1488-1492, doi: 10.1109/SMC.2015.263.

[96] Y. Hao, B. Wang, and R. Zheng, “VALERIAN: Invariant feature learning for IMU sensor-based human activity recognition in the wild,” in Proc. ACM/IEEE Int. Conf. Internet Things Design Implement. (IoTDI), 2023, pp. 66-78, doi: 10.1145/3576842.3582390.

[97] E. Büber and A. M. Guvensan, “Discriminative time-domain features for activity recognition on a mobile phone,” in Proc. IEEE Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP), Apr. 2014, pp. 1-6, doi: 10.1109/ISSNIP.2014.6827651.

[98] A. Wang, H. Chen, C. Zheng, L. Zhao, J. Liu, and L. Wang, “Evaluation of random forest for complex human activity recognition using wearable sensors,” in Proc. Int. Conf. Netw. Netw. Appl. (NaNA), 2020, pp. 310-315, doi: 10.1109/NaNA51271.2020.00060.

PHỤ LỤC

Phụ lục 1: Thuật toán và mã nguồn sử dụng trong các hệ thống HAR

Ghi chú: Mã giả trong Phụ lục được trình bày ở mức thuật toán, sử dụng tên hàm mô tả (adxl.read(), SIM808.send(), WiFi.begin()...) để dễ đọc. Chương trình thực thi trên PIC18F4520 được viết bằng ngôn ngữ C, biên dịch bằng XC8 trong MPLAB X IDE; giao tiếp với ADXL345 qua mô-đun MSSP (I^2C) và với SIM808/ESP8266 qua EUSART (UART). Các ngưỡng của ADXL345 là giá trị thanh ghi với bước 62,5 mg/LSB theo bảng dữ liệu của nhà sản xuất.

I. Thuật toán 1: RF nhúng trên PIC18F4520, Hệ thống RF4-HAR (phân loại thời gian thực)

Đầu vào (Input):

- Dữ liệu gia tốc ba trục {ax[i], ay[i], az[i]} từ ADXL345, fs = 50 Hz
- Cửa sổ W = 10 giây (500 mẫu), cập nhật mỗi 6 giây (300 mẫu mới + 200 mẫu tái sử dụng từ cửa sổ trước, chồng lấp 40%)
- Mô hình RF: T1, T2, ..., Tn cây quyết định nhị phân đã huấn luyện
- Bộ 5 đặc trưng: Mean, SD, Median, Range, RMS trên 3 trục X, Y, Z

Đầu ra (Output):

- Nhãn hoạt động $\hat{Y} \in$ {Đứng, Ngồi, Đi bộ, Chạy bộ}
- Truyền kết quả qua SIM808/GSM đến máy chủ mỗi 6 giây

1.1. Bước Thiết lập (Setup)

- Nạp thư viện cảm biến:

#include <ADXL345.h> - ADXL345 adxl;

Toàn bộ hàm giao tiếp I^2C với cảm biến ADXL345 được nạp vào Flash 32 KB của PIC18F4520. Đối tượng adxl lưu trữ địa chỉ thanh ghi và phương thức đọc dữ liệu.

- Đặt ngưỡng phát hiện chuyển động (activity):

adxl.setActivityThreshold(75), // giá trị thanh ghi THRESH_ACT = 75 (62,5 mg/LSB)

Nếu biên độ gia tốc trên bất kỳ trục nào vượt ngưỡng hoạt động THRESH_ACT, thiết bị chuyển từ chế độ nghỉ (low-power idle) sang chế độ thu thập tích cực, giảm tiêu thụ năng lượng khi người dùng đứng yên.

- Đặt ngưỡng phát hiện bất động (inactivity):

adxl.setInactivityThreshold(75), adxl.setTimeInactivity(10);

Sau 10 giây tín hiệu gia tốc liên tục dưới ngưỡng THRESH_INACT, thiết bị xác nhận trạng thái bất động và có thể giảm tần số polling để tiết kiệm pin.

- Kích hoạt ngắt trên 3 trục:

adxl.setActivityXYZ(1,1,1), adxl.setInactivityXYZ(1,1,1);

Cho phép phát hiện chuyển động/bất động theo cả 3 chiều không gian. Điều này quan trọng vì người dùng có thể đeo thiết bị ở nhiều góc độ khác nhau so với trục thẳng đứng.

- Nạp mô hình RF vào Flash:

#include “random_forest.h” - RF_Classifier rf;

Toàn bộ cấu trúc cây quyết định (các ngưỡng đặc trưng học được từ giai đoạn huấn luyện offline) được mã hóa dưới dạng C-header và nhúng trực tiếp vào Flash 32 KB. Không cần kết nối mạng hay bộ nhớ ngoài khi phân loại.

1.2. Bước Vòng lặp chính (Loop), lặp mỗi 6 giây

- Thu thập dữ liệu cảm biến:

for i = 1 to 300: {ax[i], ay[i], az[i]} ← adxl.read() // 6s * 50Hz

Mỗi vòng lặp thu thập 300 bộ mẫu gia tốc 3 trục mới (6 giây × 50 Hz), cùng với 200 bộ mẫu tái sử dụng từ cửa sổ trước tạo thành cửa sổ 10 giây (500 bộ mẫu, chồng lấp 40%).

- Tính 5 đặc trưng thống kê trên mỗi trục (X, Y, Z):

feat ← [Mean_X, SD_X, Med_X, Rng_X, RMS_X, ..., RMS_Z] // 15 chiều

Năm đặc trưng được tính bằng 5 vòng lặp đơn giản, không cần biến đổi Fourier hay bộ nhớ đệm lớn. Tổng chi phí tính toán 3 ms trên PIC18F4520 (40 MHz). Lý do chọn từng đặc trưng: Mean phân biệt Đứng (0,07 g) với Ngồi (0,31 g) theo trục X, SD phân biệt Chạy (0,41 g) với Đứng (0,09 g) theo trục X, Median theo

trục Z phân biệt Ngồi (0,84 g) với các hoạt động khác, Range phân biệt Chạy (4,00 g) với Đứng (1,16 g) theo trục Z, RMS phân biệt Chạy (0,91 g) với Đi bộ (0,59 g) theo trục X.

- Phân loại hoạt động bằng mô hình RF:

label ← rf.predict(feat) // bỏ phiếu đa số từ n cây

Mỗi cây quyết định thực hiện chuỗi phép so sánh nhị phân dọc theo các nhánh cây cho ra một phiếu bầu. Nhãn cuối cùng là hoạt động nhận được nhiều phiếu nhất. Kết quả: RF đạt độ chính xác tổng thể 95,3% (WISDM) và 99,2% (dữ liệu thu thập); độ chính xác theo lớp trung bình 97,7% và 99,6%.

- Truyền kết quả:

SIM808.send(label) - HTTP POST đến máy chủ

Nhãn hoạt động được đóng gói và gửi qua SIM808 (GSM/GPRS) đến máy chủ, đồng thời hiển thị trên terminal qua UART. Chu kỳ 6 giây đảm bảo phản hồi gần thời gian thực mà độ trễ vẫn thấp hơn ngưỡng chấp nhận của ứng dụng chăm sóc sức khỏe.

1.3. Bảng đặc trưng RF4-HAR và khả năng phân biệt hoạt động

Bảng 1. Năm đặc trưng RF4-HAR: công thức và vai trò phân loại

Đặc trưng	Công thức (trục X)	Phân biệt	Ví dụ phân biệt
Mean	$\bar{x} = (1/N)\Sigma xi$	Đứng với Ngồi	Đứng: 0,07 g / Ngồi: 0,31 g (trục X)
SD	$\sigma = \sqrt{[(1/N)\Sigma(xi-\bar{x})^2]}$	Chạy với Đứng	Chạy: 0,41 g / Đứng: 0,09 g (trục X)
Median	x_j sắp xếp, lấy giữa	Ngồi với các hoạt động khác	Ngồi: 0,84 g (trục Z)
Range	xmax − xmin	Chạy với Đứng	Chạy: 4,00 g / Đứng: 1,16 g (trục Z)
RMS	$\sqrt{[(1/N)\Sigma xi^2]}$	Chạy với Đi bộ	Chạy: 0,91 g / Đi bộ: 0,59 g (trục X)

1.4. Sơ đồ giải thuật dạng mã giả (RF4-HAR)

```
BEGIN RF4-HAR_Embedded()
// === SETUP ===
INCLUDE <ADXL345.h>, INCLUDE <random_forest.h>
adxl.setActivityThreshold(75) // 75 mg
adxl.setInactivityThreshold(75), adxl.setTimeInactivity(10)
adxl.setActivityXYZ(1,1,1), adxl.setInactivityXYZ(1,1,1)
rf ← RF_Classifier(random_forest.h) // nạp model vào Flash

// === LOOP (lặp mỗi 6 giây) ===
WHILE true DO
    // Thu thập dữ liệu cảm biến
    FOR i ← 1 TO 300:
        {ax[i], ay[i], az[i]} ← adxl.read() // 50 Hz
    END FOR

    // Tính 15 đặc trưng (5 đặc trưng * 3 trục)
    FOR EACH axis IN {X, Y, Z}:
        feat[axis].Mean ← mean(data[axis])
        feat[axis].SD ← std(data[axis])
        feat[axis].Median ← median(data[axis])
        feat[axis].Range ← max(data[axis]) − min(data[axis])
        feat[axis].RMS ← sqrt(mean(data[axis]²))
    END FOR // - vector feat[15 chiều]

    // Phân loại bằng RF (bỏ phiếu đa số)
    label ← rf.predict(feat) // Ŷ ∈ {Đứng, Ngồi, Đi bộ, Chạy bộ}

    // Truyền kết quả
```

SIM808.send(HTTP_POST, label) // GSM/GPRS - máy chủ

delay(6000 ms) // chu kỳ cập nhật 6 giây

END WHILE

END RF4-HAR_Embedded()

II. Thuật toán 2: RF nhúng trên PIC18F4520, Hệ thống RF5-HAR (5 hoạt động, địa hình tòa nhà cao tầng)

Đầu vào (Input):

- Dữ liệu gia tốc ba trục {ax[i], ay[i], az[i]} từ ADXL345, fs = 50 Hz
- Kích thước cửa sổ W = 8 giây (400 mẫu), chồng lấp 40%
- Mô hình RF: 200 cây quyết định nhị phân đã huấn luyện (n_estimators=200, kích thước ≈ 28 KB)
- Bộ 5 đặc trưng nâng cao: AAD, STD, IQR, Range, RMS trên 3 trục X, Y, Z

Đầu ra (Output):

- Nhãn hoạt động $\hat{Y} \in$ {Đứng, Ngồi, Đi bộ, Chạy bộ, Đi cầu thang}
- Truyền kết quả qua ESP8266/WiFi đến máy chủ mỗi 5 giây

2.1. Bước Thiết lập (Setup), RF5-HAR

- Cấu hình ngưỡng cảm biến (tăng lên 80 mg):

adxl.setActivityThreshold(80), adxl.setInactivityThreshold(80);

Ngưỡng tăng từ 75 mg (RF4-HAR) lên 80 mg vì hoạt động leo cầu thang tạo ra xung lực thẳng đứng lớn hơn đi bộ bằng phẳng (0.8g trung bình trục Z khi leo, 0.4g khi đi bộ). Ngưỡng 80 mg loại bỏ hiệu quả nhiễu do rung thiết bị mà không bỏ sót tín hiệu cầu thang.

- Kích hoạt ngắt 3 trục và phát hiện ngã:

adxl.setActivityXYZ(1,1,1), adxl.setFreeFallThreshold(7); // 7 × 62,5 mg ≈ 0,44 g

Thêm ngưỡng phát hiện ngã (giá trị thanh ghi THRESH_FF = 7, tương ứng 7 × 62,5 mg ≈ 0,44 g, nằm trong khoảng 300-600 mg nhà sản xuất khuyến nghị cho phát hiện rơi tự do) vì trong bối cảnh tòa nhà cao tầng, ngã trên cầu thang là sự

kiện nguy hiểm cần phản hồi tức thì. Khi tổng gia tốc dưới ngưỡng này trong >5 ms, hệ thống phát ngắt khẩn cấp và ghi nhãn FA (Fall) ngay lập tức, không chờ hết cửa sổ 8 giây.

- Nạp mô hình RF5 (200 cây) vào Flash:

#include “rf5_model.h” - RF5_Classifier rf5;

So với RF4-HAR (50 cây), RF5-HAR dùng 200 cây (n_estimators=200, random_state=0, kích thước 28 KB). Số cây tăng cần thiết vì phải tạo ranh giới quyết định phức tạp hơn giữa 5 lớp, đặc biệt giữa Đi bộ (AAD≈0.23) và Đi cầu thang (AAD≈0.168) vốn có đặc trưng gần nhau.

- Khởi tạo WiFi qua ESP8266:

WiFi.begin(ssid, password), - TCP/IP đến máy chủ

Nâng cấp từ SIM808/GSM (Chương 2) lên ESP8266/WiFi (Chương 3): tốc độ truyền tăng từ 2.4 kbps lên 72 Mbps, độ trễ truyền giảm từ 2 giây xuống 50 ms, tiêu thụ năng lượng truyền giảm 40%. Điều này phù hợp với môi trường tòa nhà có sẵn hạ tầng WiFi.

2.2. Bước Vòng lặp chính (Loop), lặp mỗi 5 giây

- Thu thập 250 mẫu trong 5 giây (cửa sổ 8 giây, chồng lấp 40%):

for i = 1 to 250: {ax[i], ay[i], az[i]} ← adxl.read() // 5s * 50Hz

Cửa sổ 8 giây được thực hiện với chồng lấp 40%: mỗi chu kỳ 5 giây thu 250 mẫu mới và tái sử dụng 150 mẫu cũ. Lý do chọn 8 giây (thay vì 10 giây của RF4-HAR): hoạt động leo cầu thang có chu kỳ bước 0,6-0,8 giây, nhanh hơn đi bộ (1.0 giây). Cửa sổ 8 giây nắm bắt 10-13 bước cầu thang, đủ để trích xuất đặc trưng ổn định, đồng thời Bảng 3.8 cho thấy độ chính xác của RF ở 8, 10 và 12 giây là như nhau nên cửa sổ nhỏ nhất (8 giây) được chọn để giảm độ trễ và bộ đệm.

- Phân đoạn cửa sổ trượt cố định (frame = 400 mẫu = 8 giây, stride = 160 mẫu, chồng lấp 40%):

X_train=[data[i:i+400] for i in range(0,len(data),160)] // frame=400, stride=160

Cửa sổ trượt cố định: frame=400 mẫu (8 giây * 50 Hz), stride=160 mẫu (chồng lấp 40%), theo đúng code thực nghiệm codeArticle_5_Activity.ipynb. Lý do chọn 8 giây: hoạt động leo cầu thang có chu kỳ bước 0,6-0,8 giây, cửa sổ 8 giây chứa 10-13 bước đầy đủ để trích đặc trưng ổn định.

- Tính 5 đặc trưng nâng cao trên mỗi trục (X, Y, Z):

feat ← [AAD_X, STD_X, IQR_X, Rng_X, RMS_X, ..., RMS_Z] // 15 chiều

Bộ đặc trưng mới AAD, STD, IQR, Range, RMS thay cho Mean, SD, Median, Range, RMS của RF4-HAR: (1) AAD thay Mean vì bền vững hơn với outlier khi xung lực cầu thang gây đỉnh tín hiệu nhất thời, (2) IQR thay Median vì phân biệt tốt hơn Đứng (IQR=0,056) với Đi cầu thang (IQR=0,191). Theo phân tích feature importance của RF: STD 26,3%, Range 22,1%, RMS 19,7%, AAD 17,4%, IQR 14,5%, tất cả đều đóng góp ý nghĩa (không đặc trưng nào dư thừa).

- Phân loại bằng RF5 với bỏ phiếu đa số từ 200 cây (n_estimators=200):

label ← rf5.predict(feat) - $\hat{Y} \in$ {1=Đứng, 2=Ngồi, 3=Đi bộ, 4=Chạy, 5=Cầu thang}

RF chọn lớp có số phiếu nhiều nhất trong 200 cây (bỏ phiếu theo đa số tương đối - plurality), không yêu cầu quá bán. Kết quả: RF5 đạt Acc = 96,5% (WISDM, cửa sổ 8 giây) và Acc = 98,3% (dữ liệu thu thập). Độ chính xác thấp hơn RF4-HAR (97,7% và 99,6%) là chấp nhận được vì bài toán khó hơn (5 lớp thay vì 4 lớp, bao gồm cầu thang).

- Truyền kết quả qua WiFi:

ESP8266.send(TCP, label) - máy chủ (độ trễ 50 ms)

Gói tin bao gồm: nhãn hoạt động, timestamp, ID thiết bị và chỉ số năng lượng pin. Người giám sát nhận cập nhật mỗi 5 giây qua giao diện web. Việc chuyển từ GSM sang WiFi tăng băng thông và giảm chi phí vận hành khi triển khai trong tòa nhà.

2.3. Bảng so sánh RF4-HAR và RF5-HAR

Bảng 2. So sánh hai hệ thống RF4-HAR và RF5-HAR trên các tham số kỹ thuật

Tham số kỹ thuật	RF4-HAR (Chương 2)	RF5-HAR (Chương 3)

Số hoạt động	4 (Đứng/Ngồi/Đi/Chạy)	5 (+Đi cầu thang gộp)
Bộ 5 đặc trưng	Mean, SD, Median, Range, RMS	AAD, STD, IQR, Range, RMS
Kích thước cửa sổ	10 giây (500 mẫu)	8 giây (400 mẫu)
Số cây RF	30-50 cây (12 KB)	200 cây (28 KB)
Cửa sổ thích nghi	Không	Không (cửa sổ cố định, frame=400 mẫu)
Chu kỳ phân loại	6 giây	5 giây
Truyền thông	SIM808/GSM (2.4 kbps)	ESP8266/WiFi (72 Mbps)
Acc, WISDM	97,7%	96,5%
Acc, Thu thập	99,6%	98,3%
Năng lượng tiêu thụ RF	0.01595 J/hoạt động	0.01595 J/hoạt động

2.4. Sơ đồ giải thuật dạng mã giả (RF5-HAR)

BEGIN RF5-HAR_Embedded()

// === SETUP ===

INCLUDE <ADXL345.h>, INCLUDE <rf5_model.h>

adxl.setActivityThreshold(80), adxl.setInactivityThreshold(80)

adxl.setActivityXYZ(1,1,1), adxl.setFreeFallThreshold(7) // phát hiện ngã

rf5 ← RF5_Classifier(rf5_model.h) // 200 cây (n_estimators=200), 28 KB Flash

ESP8266.begin(ssid, pwd) // khởi tạo WiFi

// === LOOP (lặp mỗi 5 giây) ===

WHILE true DO

// Thu thập 250 mẫu mới (cửa sổ 8s, chồng lấp 40%)

FOR i ← 1 TO 250:

{ax[i], ay[i], az[i]} ← adxl.read() // 50 Hz

END FOR

```
        window ← merge(new_250_samples, last_150_samples) // 400 mẫu =
    8s
        // Phát hiện bất thường - thu nhỏ cửa sổ nếu cần
        window ← data[i:i+400] // frame=400 = 8 giây, stride=160 = chồng lấp
    40%
        // Không dùng adaptive window, cửa sổ cố định 400 mẫu
        // Tính 15 đặc trưng (5 * 3 trục), bộ nâng cao
        FOR EACH axis IN {X, Y, Z}:
            feat[axis].AAD ← (1/N)Σ|xi − x̄|
            feat[axis].STD ← sqrt((1/N)Σ(xi−x̄)²)
            feat[axis].IQR ← Q75(data) − Q25(data)
            feat[axis].Range ← max(data) − min(data)
            feat[axis].RMS ← sqrt((1/N)Σxi²)
        END FOR // - vector feat[15 chiều]

        // Phân loại RF5 với 200 cây, n_estimators=200 (bỏ phiếu đa số)
        label ← rf5.predict(feat) // Ŷ ∈ {Đứng, Ngồi, Đi bộ, Chạy, Cầu thang}

        // Truyền kết quả qua WiFi
        ESP8266.send(TCP, {label, timestamp, battery_level})
        delay(5000 ms) // chu kỳ cập nhật 5 giây
    END WHILE
END RF5-HAR_Embedded()
```

III. Thuật toán 3: KNN nhúng trên PIC18F4520, Hệ thống KNN6-HAR (6 hoạt động, phân biệt Lên/Xuống cầu thang)

Đầu vào (Input):

- Dữ liệu gia tốc ba trục {ax[i], ay[i], az[i]} từ ADXL345, fs = 50 Hz
- Cửa sổ trượt: frame = 300 mẫu (= 6 giây * 50 Hz), stride = 120 mẫu (chồng lấp 40%)
- Tập huấn luyện D = {(feat_j, L_j)} đã lưu vào Flash (dạng look-up table)

- Bộ 4 đặc trưng: Var, SAV, Skew, P2P trên 3 trục X, Y, Z, k=5 láng giềng

Đầu ra (Output):

- Nhãn hoạt động $\hat{Y} \in$ {Đứng, Ngồi, Đi bộ, Chạy bộ, Xuống cầu thang, Lên cầu thang}

- Truyền kết quả qua ESP8266/WiFi đến máy chủ

3.1. Lý do chọn KNN thay RF cho bài toán phân biệt Lên/Xuống cầu thang

RF5-HAR gộp Lên/Xuống cầu thang thành một lớp vì dữ liệu gia tốc hai hoạt động này có đặc trưng gần nhau, dễ gây nhầm lẫn. KNN6-HAR chọn thuật toán KNN (k=5, trọng số "distance") vì KNN không cần huấn luyện lại khi có dữ liệu mới, thích ứng tốt với ranh giới quyết định phi tuyến giữa Lên và Xuống cầu thang, và chi phí tính toán khi suy luận chỉ là O(n·d) (n = số mẫu huấn luyện được lưu, d = số chiều đặc trưng = 12). Trên PIC18F4520, với tập huấn luyện được nén xuống 200 mẫu đại diện/lớp lưu dưới dạng look-up table trong Flash 32 KB, KNN cho thời gian suy luận khoảng 82 ms/phân loại.

3.2. Bước Thiết lập (Setup), KNN6-HAR

- Khởi tạo cảm biến ADXL345:

adxl.setRange(ADXL345_RANGE_4G), // ±4g, 128 LSB/g, nhiễu 150 µg

Cấu hình ±4g tối ưu cho bài toán cầu thang: đủ rộng để thu xung lực cao khi Chạy (4g peak) mà không bị bão hòa (saturation) như cấu hình ±2g, đồng thời phân giải (128 LSB/g) đủ tinh để phân biệt gia tốc thẳng đứng khi Lên (az trung bình 0.88g) với Xuống cầu thang (az trung bình 0.419g theo Bảng 3.7).

- Áp dụng bộ lọc thông thấp IIR bậc một (trung bình trượt hàm mũ):

$ax_f[i] \leftarrow K \cdot ax[i] + (1-K) \cdot ax_f[i-1]$, // K = 0.1 (hệ số làm mịn)

Bộ lọc thông thấp IIR bậc một (trung bình trượt hàm mũ) với hệ số K = 0,1 làm mịn nhiễu cảm biến tần số cao. Với fs = 50 Hz, tần số cắt xấp xỉ $f_c \approx K \cdot f_s/(2\pi) \approx$ 0,8 Hz; do đó bộ lọc làm suy giảm một phần thành phần tần số bước chân (1-3 Hz) và ảnh hưởng đến biên độ của Var, P2P. Để bảo toàn thành phần bước chân, có thể chọn K ≈ 0,45-0,5 ($f_c \approx$ 5 Hz). Bộ lọc được thực thi bằng 2 phép nhân và 1 phép cộng, phù hợp với PIC18F4520 (không có FPU).

- Nạp tập huấn luyện đại diện vào Flash:

INCLUDE “knn_lookup.h” // 1200 mẫu * 12 chiều * 4 byte = 57.6 KB

KNN là thuật toán “học lười” (lazy learning): không huấn luyện mô hình, chỉ lưu tập dữ liệu huấn luyện. Để vừa Flash 32 KB, tập huấn luyện được nén bằng k-means clustering: từ hàng ngàn mẫu gốc - chọn 200 mẫu đại diện/lớp * 6 lớp = 1200 mẫu. Mỗi mẫu gồm 12 đặc trưng * 2 byte (fixed-point) = 24 byte - 1200 * 24 = 28.8 KB ≤ 32 KB Flash.

- Cài đặt tham số KNN:

k = 5, metric = EUCLIDEAN, weight = DISTANCE;

k=5 được xác định qua thử nghiệm: k<3 gây nhiễu bởi outlier, k>7 làm mờ ranh giới giữa Lên và Xuống cầu thang. Trọng số “distance” (wi = 1/di, đúng với tham số weights = ‘distance’ của scikit-learn) ưu tiên láng giềng gần hơn, giảm ảnh hưởng của các mẫu không điển hình ở biên ranh giới hai lớp cầu thang.

3.3. Bước Vòng lặp chính (Loop), KNN6-HAR

- Phân đoạn cửa sổ trượt và thu thập dữ liệu (frame = 300 mẫu (= 6 giây * 50 Hz)**):**

for i=1 to 300: {ax[i],ay[i],az[i]} ← adxl.read()

Cửa sổ frame = 300 mẫu (= 6 giây * 50 Hz), stride = 120 mẫu (chồng lấp 40%). Dữ liệu được chia theo train_test_split(test_size=0.4, random_state=42), đảm bảo tỷ lệ 60% huấn luyện và 40% kiểm tra. Bộ lọc thông thấp bậc một (K = 0,1) được áp dụng để làm mịn nhiễu cảm biến trước khi trích xuất đặc trưng.

- Tính 4 đặc trưng thống kê trên mỗi trục (X, Y, Z):

feat ← [Var_X, SAV_X, Skew_X, P2P_X, ..., P2P_Z] // 12 chiều

Bốn đặc trưng Var, SAV, Skew, P2P được chọn vì đặc biệt phân biệt Lên/Xuống cầu thang: Skew (độ lệch phân phối) khác dấu giữa Lên (+0.19, lệch phải) và Xuống cầu thang (−0.46, lệch trái) theo trục X (Bảng 3.19), phản ánh sự bất đối xứng trong lực tác dụng khi nâng chân lên với hạ chân xuống. Var đo mức độ dao động, phân biệt các trạng thái tĩnh với các trạng thái động.

- Tính khoảng cách Euclidean đến tất cả mẫu huấn luyện:

FOR j=1 to 1200: d[j] ← sqrt(Σ(feat[m] − train[j][m])²), m=1..12

- Tổng chi phí: dùng số nguyên dấu phẩy tĩnh 16 bit và so sánh bình phương khoảng cách (không cần khai căn), mỗi lần phân loại cần 1.200 × 12 = 14.400 phép trừ và nhân-cộng. PIC18F4520 chạy 40 MHz tương ứng 10 MIPS (4 chu kỳ xung nhịp/lệnh), có bộ nhân phần cứng 8×8 bit nên một phép nhân-cộng 16 bit tốn khoảng 40-50 chu kỳ lệnh; tổng khoảng 0,6-0,72 triệu chu kỳ lệnh, tức 60-72 ms, phù hợp với thời gian suy luận đo được khoảng 82 ms (Bảng 3.27, 3.29). Nếu dùng số thực dấu phẩy động bằng phần mềm (không có FPU), mỗi phép nhân tốn hàng trăm chu kỳ lệnh và không đáp ứng ràng buộc 100 ms.

- Chọn k=5 láng giềng gần nhất và bỏ phiếu:

sort(d[1..1200]), neighbors ← top_k(5), label ← weighted_vote(neighbors, 1/d)

Bỏ phiếu có trọng số: mẫu gần hơn có ảnh hưởng lớn hơn. Ví dụ: nếu 3 trong 5 láng giềng gần nhất là "Lên cầu thang" ($\bar{d}$=0.12) và 2 là "Đi bộ" ($\bar{d}$=0.35), phiếu Lên = 3/0,12 = 25, phiếu Đi bộ = 2/0,35 ≈ 5,7 - nhãn = "Lên cầu thang". Cơ chế này đặc biệt hiệu quả ở vùng biên giữa Đi bộ và hai hoạt động cầu thang.

- Truyền kết quả và cập nhật tầng:

ESP8266.send(TCP, {label, floor_estimate, timestamp})

Thông tin tầng (floor_estimate) được ước tính kết hợp nhãn hoạt động (Lên/Xuống cầu thang * số bước) với dữ liệu cảm biến khí áp (nếu có) hoặc đếm chu kỳ hoạt động cầu thang. Người giám sát nhận thông tin đầy đủ về vị trí và trạng thái của người dùng trong tòa nhà.

3.4. Bảng đặc trưng KNN6-HAR và khả năng phân biệt 6 hoạt động

Bảng 3. Bốn đặc trưng KNN6-HAR: công thức và vai trò phân loại (trục X)

Đặc trưng	Công thức (trục X)	Vai trò phân loại và giá trị điển hình
Var	$(1/N)\Sigma(x_i-\bar{x})^2$	Đo mức độ dao động của tín hiệu, phân biệt các trạng thái tĩnh (Var nhỏ) với các trạng thái động (Var lớn)

SAV	$\Sigma\|x_i\|$	Phân biệt Chạy bộ (SAV=286.15) với Đứng (0.43), phản ánh biên độ tuyệt đối, Lên cầu thang (8.7) với Xuống (14.3) giúp phân biệt hai hoạt động cầu thang
Skew	$(1/N)\Sigma[(x_i-\bar{x})/\sigma]^3$	Đặc trưng then chốt phân biệt Lên (+0.19, lệch phải) với Xuống cầu thang (−0.46, lệch trái) trục X, phân phối lực bất đối xứng khi nâng với hạ chân
P2P	$x_{max} - x_{min}$	Phân biệt Chạy (P2P=3.99g) với Đứng (0.02g), Lên cầu thang (0.30g) với Xuống (0.37g), biên độ đỉnh-đáy phản ánh cường độ chuyển động

3.5. Sơ đồ giải thuật dạng mã giả (KNN6-HAR)

BEGIN KNN6-HAR_Embedded()

// === SETUP ===

INCLUDE <ADXL345.h>, INCLUDE <knn_lookup.h>

adxl.setRange(±4g), adxl.setOutputDataRate(50Hz)

K_lp ← 0.1 // hệ số bộ lọc thông thấp

k ← 5, weight ← DISTANCE // tham số KNN

train_set ← load(knn_lookup.h) // 1200 mẫu * 12 chiều

// === LOOP ===

WHILE true DO

// Thu thập và lọc thông thấp

FOR i ← 1 TO N: // N=300 (6 giây * 50 Hz), stride=120 (chồng lấp 40%)

raw ← adxl.read()

{ax_f, ay_f, az_f} ← LowPass(raw, K=0.1) // lọc IIR bậc một

```
        END FOR
        // Tính 4 đặc trưng * 3 trục = 12 chiều
        FOR EACH axis IN {X, Y, Z}:
            feat[axis].Var ← variance(data[axis])
            feat[axis].SAV ← Σ|xi|
            feat[axis].Skew ← (1/N)Σ[(xi−x̄)/σ]³ // KEY: phân biệt Lên/Xuống
            feat[axis].P2P ← max(data[axis]) − min(data[axis])
        END FOR // - vector feat[12 chiều]

        // Tính khoảng cách Euclidean đến tất cả mẫu huấn luyện
        FOR j ← 1 TO 1200:
            d[j] ← sqrt(Σ(feat[m] − train_set[j][m])²), m=1..12
        END FOR // ≈ 60-80 ms trên PIC18F4520 (dấu phẩy tĩnh)

        // Chọn k=5 láng giềng gần nhất và bỏ phiếu có trọng số
        neighbors ← sort(d)[1:5] // 5 mẫu gần nhất
        label ← argmax_c Σ (1/d) for c in neighbors // bỏ phiếu ∝ 1/d
        // Ước tính tầng và truyền kết quả
        IF label IN {Lên cầu thang, Xuống cầu thang}: floor ←
    update_floor(label)
        ESP8266.send(TCP, {label, floor, timestamp})
     END WHILE
END KNN6-HAR_Embedded()
```