

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



VŨ HOÀNG DIỆU

**NÂNG CAO HIỆU QUẢ PHÂN LOẠI TỪ THỂ NGỮ TỪ
DỮ LIỆU CẢM BIẾN GIA TỐC BA TRỤC THEO
CÁCH TIẾP CẬN KẾT HỢP CÁC MÔ HÌNH HỌC SÂU**

LUẬN ÁN TIẾN SĨ MÁY TÍNH

VŨ HOÀNG DIỆU

2026

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 HOÀNG DIỆU

**NÂNG CAO HIỆU QUẢ PHÂN LOẠI TƯ THỂ NGŨ TỪ DỮ
LIỆU CẢM BIẾN GIA TỐC BA TRỤC THEO CÁCH TIẾP CẬN
KẾT HỢP CÁC MÔ HÌNH HỌC SÂU**

LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Xác nhận của Học viện
Khoa học và Công nghệ



Người hướng dẫn 1
(Ký, ghi rõ họ tên)

TS. TRẦN ĐỨC NGHĨA

Người hướng dẫn 2
(Ký, ghi rõ họ tên)

GS. TS. TRẦN ĐỨC TÂN

Trần Thị Phương Anh

Hà Nội - 2026

LỜI CAM ĐOAN

Tôi xin cam đoan luận án này 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 03 tháng 09 năm 2026

Tác giả luận án

(Ký và ghi rõ họ tên)



Vũ Hoàng Diệu

LỜI CẢM ƠN

Luận án này được hoàn thành với sự nỗ lực không ngừng của tác giả và sự giúp đỡ hết mình từ các thầy giáo hướng dẫn, bạn bè và người thân.

Đầu tiên, tác giả xin bày tỏ lời tri ân tới TS Trần Đức Nghĩa và GS. TS. Trần Đức Tân, cảm ơn các thầy đã tận tình hướng dẫn tác giả hoàn thành luận án này.

Tác giả xin gửi lời cảm ơn tới các thầy, cô giáo và cán bộ của Viện Công nghệ thông tin, Học viện Khoa học và Công nghệ (Viện Hàn lâm Khoa học và Công nghệ Việt Nam) đã nhiệt tình giúp đỡ và tạo ra môi trường nghiên cứu tốt để tác giả hoàn thành công trình của mình; cảm ơn các thầy, cô và các đồng nghiệp ở các nơi mà tác giả tham gia viết bài đã có những góp ý chính xác để tác giả có được những công bố như ngày hôm nay. Tác giả xin cảm ơn tới Ban Giám hiệu Đại học Phenikaa, các đồng nghiệp khoa Điện - Điện tử nơi tác giả công tác đã ủng hộ và hỗ trợ để luận án được hoàn thành.

Quá trình hoàn thành luận án này đối với tác giả không thể thiếu sự đồng hành của gia đình, bạn bè, người thân, đặc biệt là vợ Cấn Khánh Ly đã luôn ở bên động viên, ủng hộ bằng tất cả tâm huyết và chân thành. Tác giả xin cảm ơn bằng tất cả sự yêu thương, chân thành sâu sắc nhất của mình.

Tác giả cũng xin gửi lời cảm ơn tới các bạn sinh viên Phạm Đình Đạt, Phạm Quang Tú và các bạn sinh viên khác đã đồng hành cùng tác giả trong các nghiên cứu liên quan luận án này. Tác giả xin gửi lời cảm ơn tới Quỹ Đổi mới sáng tạo của tập đoàn Vingroup (VinIF), quỹ Nafosted đã tài trợ cho nghiên cứu này.

Hà Nội, ngày 03 tháng 09 năm 2026

Tác giả luận án
(*Ký và ghi rõ họ tên*)



Vũ Hoàng Diệu

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN	ii
MỤC LỤC	iii
DANH MỤC CÁC CHỮ VIẾT TẮT	vi
DANH MỤC HÌNH VẼ	viii
DANH MỤC BẢNG	x
MỞ ĐẦU	1
1 TỔNG QUAN VỀ PHÂN LOẠI TƯ THỂ NGỦ VÀ CÁC TIẾP CẬN DƯA TRÊN CẢM BIẾN	11
1.1 Bài toán phân loại tư thể ngủ	11
1.1.1 Tầm quan trọng về mặt lâm sàng của theo dõi tư thể ngủ . . .	11
1.1.2 Tư thể độ phân giải cao và tư thể trung gian trong nghiên cứu lâm sàng	14
1.2 Tổng quan về các hệ thống phân loại tư thể ngủ	18
1.3 Tổng quan về các phương pháp học máy cho phân loại tư thể ngủ . . .	23
1.3.1 Các tiếp cận học máy truyền thống	23
1.3.2 Học sâu cho dữ liệu cảm biến chuỗi thời gian	26
1.3.3 Mạng nơ-ron tích chập cho tín hiệu cảm biến	27
1.3.4 Mạng nơ-ron hồi quy cho mô hình hóa chuỗi thời gian	29
1.3.5 Kiến trúc Lai CNN-RNN cho dữ liệu cảm biến	31
1.4 Thử nghiệm sơ bộ: xác thực tính khả thi	32
1.5 Các chỉ số đánh giá	35
1.5.1 Confusion Matrix (ma trận nhầm lẫn)	35

1.5.2 Accuracy	36
1.5.3 Precision	37
1.5.4 Recall	37
1.5.5 F1-score	38
1.6 Kết luận chương 1	38
2 ĐỀ XUẤT MÔ HÌNH ANPONET VỚI SỐ LƯỢNG THAM SỐ	
RÀNG BUỘC CHO BÀI TOÁN PHÂN LOẠI TƯ THỂ NGŨ	40
2.1 Phương pháp luận cho việc thiết kế bộ dữ liệu	40
2.1.1 Mục tiêu và phạm vi của bộ dữ liệu	40
2.1.2 Cấu hình cảm biến và thiết lập phần cứng	41
2.1.3 Giao thức thu thập dữ liệu được chuẩn hóa	47
2.1.4 Quy trình gán nhãn và tiền xử lý	52
2.1.5 Các bộ dữ liệu	54
2.1.6 Thiết lập thực nghiệm	59
2.1.7 Xử lý dữ liệu đầu vào và tổ chức tập dữ liệu	60
2.2 Thiết kế mô hình học sâu hướng tối thực thi trên thiết bị biên	62
2.2.1 Phương pháp tiếp cận của AnpoNet	63
2.2.2 Mô hình hóa bài toán	64
2.2.3 Thiết kế khung thuật toán AnpoNet	66
2.2.4 Bài toán huấn luyện và hàm mất mát	71
2.2.5 Phân tích độ phức tạp và đặc tính tính toán của AnpoNet	72
2.3 Thực nghiệm và đánh giá	76
2.3.1 Các mô hình so sánh	76
2.3.2 Thiết kế thực nghiệm	78
2.4 Kết quả thực nghiệm và thảo luận	81
2.4.1 Phân tích hiệu suất	81
2.4.2 Đánh giá dưới các hạt giống ngẫu nhiên khác nhau	84
2.4.3 Đánh giá dựa trên phương pháp K-Folds	86
2.4.4 Ảnh hưởng của độ dài chuỗi, tỷ lệ chồng lấp đến hiệu suất mô hình	88
2.4.5 Phân tích quá trình huấn luyện	90
2.4.6 Phân tích độ phức tạp mô hình và hiệu quả tính toán	91
2.4.7 So sánh hiệu suất với các mô hình tương đương về kiến trúc	92
2.4.8 Tác động của nhiễu trắng đến độ bền vững của mô hình	93

2.4.9	Hiệu suất theo từng lớp	94
2.5	Kết luận chương 2	95
3	ĐỀ XUẤT MÔ HÌNH CGB-NET NHẪM TĂNG CƯỜNG ĐỘ CHÍNH XÁC TRONG PHÂN LOẠI TƯ THỂ NGƯỜI	97
3.1	Bài toán và vấn đề nghiên cứu	97
3.2	Phương pháp tiếp cận mô hình đề xuất	99
3.3	Khung thuật toán đề xuất CGB-Net	101
3.4	Thực nghiệm	107
3.4.1	Các mô hình so sánh	107
3.4.2	Huấn luyện và tối ưu hóa mô hình	109
3.5	Kết quả thực nghiệm và thảo luận	111
3.5.1	Phân tích hiệu suất của mô hình	111
3.5.2	Đánh giá với Seed ngẫu nhiên	116
3.5.3	Phân tích ma trận nhầm lẫn	117
3.5.4	Quá trình huấn luyện	121
3.5.5	Ảnh hưởng của độ dài chuỗi và tỷ lệ chồng lấp	122
3.6	Kết luận chương 3	124
	KẾT LUẬN VÀ KIẾN NGHỊ	125
	DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ	130
	TÀI LIỆU THAM KHẢO	131

DANH MỤC CÁC CHỮ VIẾT TẮT

Chữ viết tắt	Ý nghĩa tiếng Anh	Ý nghĩa tiếng Việt
1D CNN	One-Dimensional Convolutional Neural Network	Mạng nơ-ron tích chập một chiều
2D CNN	Two-Dimensional Convolutional Neural Network	Mạng nơ-ron tích chập hai chiều
3D CNN	Three-Dimensional Convolutional Neural Network	Mạng nơ-ron tích chập ba chiều
ACC	Accelerometer	Gia tốc kế
AE	Autoencoder	Mạng tự mã hóa
ANN	Artificial Neural Network	Mạng nơ-ron nhân tạo
ATT	Attention Mechanism	Cơ chế chú ý
AUC	Area Under the Receiver Operating Characteristic Curve (ROC)	Diện tích dưới đường cong đặc trưng hoạt động của bộ thu
BiGRU	Bidirectional Gated Recurrent Unit	Mạng GRU hai chiều
BiLSTM	Bidirectional Long Short-Term Memory	Mạng LSTM hai chiều
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
CRNN	Convolutional Recurrent Neural Network	Mạng nơ-ron tích chập hồi quy
DCNN	Deep Convolutional Neural Network	Mạng nơ-ron tích chập sâu
DNN	Deep Neural Network	Mạng nơ-ron sâu
EEG	Electroencephalography	Điện não đồ
EOG	Electrooculography	Điện nhãn cầu

Chữ viết tắt	Ý nghĩa tiếng Anh	Ý nghĩa tiếng Việt
F1	F1-Score (Harmonic Mean of Precision and Recall)	Điểm F1 (Trung bình điều hòa của độ chính xác và độ phủ)
FCN	Fully Convolutional Network	Mạng tích chập toàn phần
FMCW	Frequency-Modulated Continuous Wave (Radar)	Radar sóng liên tục điều tần
FN	False Negative	Âm tính giả
FP	False Positive	Dương tính giả
GRU	Gated Recurrent Unit	Đơn vị hồi quy có cổng
HAR	Human Activity Recognition	Nhận dạng hoạt động người
IMU	Inertial Measurement Unit	Bộ đo quán tính
LSTM	Long Short-Term Memory	Bộ nhớ dài hạn
MLP	Multilayer Perceptron	Mạng nơ-ron đa lớp
PPG	Photoplethysmography	Kỹ thuật ghi biến thiên thể tích quang học
Recall	Proportion of True Positives Identified Correctly	Độ phủ / Độ thu hồi (Tỷ lệ dương tính thật được xác định đúng)
ReLU	Rectified Linear Unit	Đơn vị chỉnh lưu tuyến tính
RNN	Recurrent Neural Network	Mạng nơ-ron hồi quy
RMSE	Root Mean Square Error	Căn bậc hai trung bình bình phương sai số
SMOTE	Synthetic Minority Oversampling Technique	Kỹ thuật lấy mẫu quá mức thiểu số tổng hợp
SVM	Support Vector Machine	Máy vector hỗ trợ
TP	True Positive	Dương tính thật
TN	True Negative	Âm tính thật
UWB	Ultra-Wideband (Radar)	Radar băng thông siêu rộng

DANH MỤC HÌNH VẼ

1.1	Biểu đồ tròn để trực quan hóa của tư thế ngủ cho thấy thời gian dành cho mỗi góc thân mình ($\pm 15^\circ$) và chỉ số ngừng thở-giảm thở (AHI) tương ứng theo từng vị trí. [1]	17
1.2	Một cấu hình 1D CNN mẫu với 3 lớp CNN và 2 lớp MLP. [2]	29
1.3	Hệ thống phân loại tư thế ngủ	32
1.4	Kiến trúc của mô hình phân loại LSTM tuần tự	33
2.1	Hệ thống hai mô-đun cho phân loại 12 Tư thế Ngủ (12-SPSC)	46
2.2	Sơ đồ hệ thống thu thập dữ liệu gia tốc: ngoại tuyến và trực tuyến với ứng dụng di động.	46
2.3	Giao diện người dùng ứng dụng di động cho mô-đun thu thập dữ liệu.	47
2.4	Các hình ảnh minh họa của 12 tư thế ngủ.	49
2.5	Quá trình gán nhãn tư thế	52
2.6	Tổng thể luồng xử lý của AnpoNet	67
2.7	Độ chính xác và F1-Score của ba đối tượng trong tập dữ liệu kiểm tra qua các hạt giống ngẫu nhiên khác nhau.	84
2.8	Độ chính xác và F1-Score của AnpoNet trong tập dữ liệu kiểm tra qua các fold khác nhau.	86
2.9	A. Độ chính xác của AnpoNet với tỷ lệ chồng lấp thay đổi; B. Độ chính xác của AnpoNet với độ dài chuỗi thay đổi	88
2.10	Các đường cong huấn luyện: độ chính xác và giá trị hàm mất mát trên các tập dữ liệu huấn luyện (Training) và xác thực (Validation).	90
2.11	Ma trận nhầm lẫn của mô hình AnpoNet đề xuất	94
3.1	Tổng thể luồng xử lý của CGB-Net	102
3.2	Độ chính xác và F1-Score của CGB-Net qua các lần khởi tạo seed ngẫu nhiên.	116

3.3	Ma trận nhầm lẫn (Confusion Matrix) của CGB-Net, 1D CNN + GRU	
	và AnpoNet	117
3.4	Đường cong hàm mất mát (loss) và độ chính xác của CGB-Net	121
3.5	Hiệu suất CGB-Net theo kích thước cửa sổ và tỷ lệ chồng lấp.	123

DANH MỤC BẢNG

1.1	Các công trình liên quan sử dụng gia tốc kế để phân loại tư thế ngủ . .	22
1.2	So sánh với các mô hình cơ sở và học sâu cho nhiệm vụ phân loại 12 tư thế.	33
2.1	Thông số kỹ thuật phần cứng của thiết bị đeo sử dụng trong nghiên cứu này.	43
2.2	Tóm tắt các tư thế ngủ dựa trên góc giữa hướng của bụng và trục x của hệ tọa độ gốc	50
2.3	Thống kê đặc điểm hình thể của các đối tượng tham gia trong Dataset A	55
2.4	Ví dụ minh họa cấu trúc dữ liệu gia tốc ba trục cho một tư thế ngủ . .	61
2.5	Kích thước của các tập dữ liệu huấn luyện, xác thực và kiểm tra được xây dựng với độ dài của số cố định là 100 mẫu và tỷ lệ chồng lấp là 0.4.	79
2.6	So sánh hiệu suất của các mô hình trên tập dữ liệu kiểm tra	81
2.7	Độ phức tạp mô hình và hiệu quả tính toán	91
2.8	So sánh AnpoNet và LSTM với số lượng tham số tương tự (gọi tắt là Small LSTM)	93
2.9	So sánh AnpoNet và AnpoNet với Nhiều Trắng	94
2.10	Precision và Recall cho mỗi Lớp	95
3.1	Độ chính xác (%) của các mô hình theo bốn thiết lập thí nghiệm	112
3.2	Chỉ số F1 (%) của các mô hình theo bốn thiết lập thí nghiệm	112
3.3	So sánh hiệu suất trung bình và chi phí tính toán của các mô hình trên tập kiểm thử.	114

MỞ ĐẦU

1. Tính cấp thiết của luận án

Những tiến bộ gần đây trong công nghệ cảm biến cùng với việc các thiết bị đeo được sử dụng rộng rãi đã mở ra những khả năng mới trong các ứng dụng chăm sóc sức khỏe thông minh [3,4]. Đặc biệt, cảm biến gia tốc được sử dụng phổ biến nhờ kích thước nhỏ gọn, mức tiêu thụ năng lượng thấp, chi phí hợp lý và khả năng ghi lại chuyển động của cơ thể theo thời gian thực mà không gây ảnh hưởng đến sinh hoạt hàng ngày [3].

Trong các ứng dụng đó, theo dõi tư thế ngủ đã thu hút sự quan tâm ngày càng lớn bởi vì tư thế ngủ có mối liên hệ chặt chẽ với bệnh lý như ngưng thở khi ngủ (OSA), bệnh trào ngược dạ dày (GERD), đau cơ xương khớp và sự hình thành loét do tỳ đè. Dù đã có nhiều nghiên cứu trước đây [5,6] tiến hành phân loại tư thế ngủ dựa vào thiết bị đeo, nhưng phần lớn chỉ tập trung vào ba hoặc bốn tư thế cơ bản (nằm ngửa, nằm sấp, nằm nghiêng trái, nằm nghiêng phải). Sự hạn chế về mức độ chi tiết trong phân loại này làm giảm khả năng ứng dụng trong lâm sàng của các hệ thống, đặc biệt trong những ứng dụng mà các sai lệch góc nhỏ có ý nghĩa sinh lý quan trọng [7]. Các bằng chứng lâm sàng gần đây chỉ ra rằng các nhóm tư thế ở mức độ cơ bản là không đủ để mang lại những hiểu biết có ý nghĩa trong chăm sóc sức khỏe. Năm 2017, nhóm của Pickenbrock và cộng sự đã chứng minh rằng phân bố áp lực khác biệt đáng kể giữa sáu tư thế ngủ riêng biệt, bao gồm các tư thế nằm nghiêng ở góc 30° và 90° [8]. Kết quả của họ cho thấy ngay cả những sai lệch góc tương đối nhỏ cũng có thể làm thay đổi đáng kể tác động sinh lý và nguy cơ lâm sàng. Điều này nhấn mạnh nhu cầu về các hệ thống giám sát có khả năng phân biệt tư thế ở mức độ chi tiết hơn.

Một hạn chế quan trọng khác của các nghiên cứu hiện nay là việc sử dụng các phương thức đánh giá phụ thuộc đối tượng [7,9], trong đó dữ liệu từ cùng một người tham gia xuất hiện đồng thời trong cả tập huấn luyện và tập kiểm tra. Các thiết lập như vậy có xu hướng tạo ra các kết quả về hiệu năng của mô hình nhưng không phản ánh đúng với điều kiện triển khai trong thực tế. Trong thực tế, mô hình phải có khả

năng làm việc tốt với những người dùng chưa từng xuất hiện trước đó. Ngược lại, đánh giá độc lập theo đối tượng đặt ra những thách thức lớn do sự khác biệt rất nhỏ giữa các cá nhân về hình thái cơ thể, hành vi vận động và sự thay đổi của vị trí gắn cảm biến, khiến việc đạt được khả năng tổng quát trở nên khó khăn hơn.

Bên cạnh đó, để việc theo dõi tư thế ngủ có tiềm năng ứng dụng trong các hệ thống cảm biến đeo và thiết bị tài nguyên hạn chế, mô hình phân loại cần được thiết kế theo hướng có độ phức tạp tính toán thấp, số lượng tham số hợp lý và thời gian suy luận ngắn. Trong phạm vi luận án này, khả năng hướng tới triển khai trên thiết bị biên được xem xét thông qua các chỉ số gián tiếp như số lượng tham số, chi phí tính toán, bộ nhớ sử dụng và thời gian suy luận, thay vì triển khai hoàn chỉnh mô hình trên phần cứng nhúng trong môi trường thực tế.

Từ những khoảng trống nghiên cứu trên, luận án tập trung vào hai thách thức trung tâm:

- Phát triển mô hình có khả năng phân loại chi tiết 12 tư thế ngủ khác nhau, nắm bắt các biến thiên góc có ý nghĩa, đồng thời kiểm soát độ phức tạp mô hình theo định hướng triển khai trên các thiết bị cảm biến đeo hoặc thiết bị tài nguyên hạn chế.
- Đạt được khả năng khái quát hóa, độ chính xác cao với kịch bản độc lập theo đối tượng một cách vững chắc chỉ với một cảm biến gia tốc đeo ở vùng bụng, một vị trí gắn cảm biến thực tiễn và có ý nghĩa lâm sàng, phản ánh đáng tin cậy định hướng toàn thân đồng thời vẫn đảm bảo sự thoải mái trong khi ngủ.

Hai bộ dữ liệu được thu thập theo cùng một giao thức được sử dụng xuyên suốt luận án này. Bộ dữ liệu thứ nhất làm cơ sở cho việc phát triển và đánh giá mô hình cơ sở được trình bày trong Chương 2, trong khi bộ dữ liệu thứ hai mở rộng tập đối tượng phục vụ cho phân tích trong Chương 3.

Tất cả các thí nghiệm đều tuân theo đánh giá độc lập theo đối tượng một cách nghiêm ngặt nhằm phản ánh tốt hơn các kịch bản triển khai thực tế, nơi các mô hình phải hoạt động đáng tin cậy trên những cá nhân chưa từng được quan sát trước đó. Thiết lập này làm tăng độ khó của bài toán và làm nổi bật nhu cầu về các kiến trúc có khả năng mô hình hóa theo thời gian và khái quát hóa mạnh mẽ hơn.

2. Mục tiêu nghiên cứu của luận án

Mục tiêu tổng quát của luận án là phát triển và đánh giá các mô hình học sâu

cho bài toán phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục đeo tại vùng bụng. Luận án tập trung vào việc nâng cao hiệu quả nhận dạng các tư thế ngủ chi tiết, tăng khả năng tổng quát hóa trong điều kiện độc lập theo đối tượng, đồng thời kiểm soát độ phức tạp của mô hình theo định hướng triển khai trên các hệ thống cảm biến đeo có tài nguyên tính toán hạn chế.

Mục tiêu tổng quát này được cụ thể hóa thành các mục tiêu sau:

1. Xây dựng bài toán phân loại 12 tư thế ngủ từ tín hiệu gia tốc ba trục dưới dạng bài toán phân loại chuỗi thời gian đa lớp, trong đó dữ liệu được thu thập từ một cảm biến gia tốc đeo tại vùng bụng và được đánh giá theo thiết lập độc lập theo đối tượng nhằm phản ánh yêu cầu tổng quát hóa đối với người dùng mới.
2. Đề xuất mô hình AnpoNet, một kiến trúc học sâu lai kết hợp mạng nơ-ron tích chập một chiều và mạng LSTM, nhằm khai thác đồng thời các đặc trưng cục bộ và phụ thuộc thời gian của tín hiệu gia tốc, đồng thời kiểm soát số lượng tham số và chi phí tính toán theo định hướng triển khai trên thiết bị tài nguyên hạn chế.
3. Đề xuất mô hình CGB-Net, một kiến trúc học sâu lai kết hợp mạng nơ-ron tích chập một chiều, GRU và BiLSTM, nhằm tăng cường năng lực biểu diễn đặc trưng không gian-thời gian của tín hiệu gia tốc và cải thiện hiệu quả phân loại trong điều kiện dữ liệu mở rộng.
4. Xây dựng quy trình thực nghiệm để đánh giá hiệu năng của các mô hình đề xuất trong bài toán phân loại 12 tư thế ngủ, bao gồm đánh giá độc lập theo đối tượng, so sánh với các mô hình cơ sở, phân tích ảnh hưởng của cấu hình dữ liệu đầu vào và phân tích lỗi theo từng lớp tư thế.
5. Làm rõ ưu điểm, hạn chế và phạm vi ứng dụng của các mô hình đề xuất, từ đó xác định các hướng phát triển tiếp theo cho hệ thống giám sát tư thế ngủ dựa trên cảm biến đeo.

3. Đối tượng và phạm vi nghiên cứu của luận án

Đối tượng nghiên cứu của luận án này là bài toán phân loại tư thế ngủ của con người từ dữ liệu cảm biến gia tốc ba trục. Cụ thể, luận án khảo sát cách thức áp dụng các mô hình học sâu nhằm phân loại các tư thế ngủ chi tiết dựa trên dữ liệu chuyển động thu thập từ một cảm biến gia tốc kế ba trục duy nhất được đeo tại vùng

bụng. Trọng tâm của nghiên cứu là phát triển các mô hình có khả năng nhận dạng 12 tư thế ngủ tĩnh, tăng khả năng tổng quát hóa trong điều kiện đánh giá độc lập theo đối tượng, đồng thời kiểm soát độ phức tạp mô hình theo định hướng triển khai trên các hệ thống cảm biến đeo hoặc thiết bị tài nguyên hạn chế.

Phạm vi nghiên cứu của luận án này được giới hạn như sau:

- **Vị trí cảm biến:** Luận án chỉ sử dụng một cảm biến gia tốc kế ba trục được gắn tại vùng bụng của người tham gia. Các cảm biến bổ sung như con quay hồi chuyển, cảm biến áp lực, cảm biến hô hấp hoặc các kỹ thuật kết hợp đa cảm biến không được xem xét trong phạm vi nghiên cứu này.
- **Tập tư thế:** Bài toán phân loại được giới hạn trong 12 tư thế ngủ tĩnh được xác định trước và được thực hiện trong điều kiện thu thập dữ liệu có kiểm soát. Các tư thế chuyển tiếp, tư thế không rõ ràng, chuyển động trong khi ngủ và dữ liệu ngủ tự nhiên qua đêm chưa được xem xét trong luận án.
- **Kỹ thuật mô hình hóa:** Nghiên cứu tập trung vào các mô hình học sâu dựa trên các kiến trúc 1D CNN, LSTM, GRU và BiLSTM. Các phương pháp học máy truyền thống và các kiến trúc học sâu khác chỉ được sử dụng với mục đích làm cơ sở so sánh nhằm đánh giá hiệu quả của các mô hình đề xuất.
- **Giao thức đánh giá:** Đánh giá độc lập theo đối tượng là giao thức chính được sử dụng trong luận án. Trong thiết lập này, dữ liệu của các đối tượng dùng để kiểm tra không xuất hiện trong quá trình huấn luyện mô hình. Việc chia dữ liệu được thực hiện theo cấp độ người tham gia thay vì chia ngẫu nhiên theo cửa sổ tín hiệu, nhằm hạn chế nguy cơ rò rỉ thông tin giữa tập huấn luyện và tập kiểm tra, đồng thời phản ánh tốt hơn yêu cầu tổng quát hóa của mô hình đối với người dùng mới.
- **Triển khai trên thiết bị biên:** Luận án định hướng thiết kế các mô hình có độ phức tạp thấp và có tiềm năng triển khai trên thiết bị tài nguyên hạn chế. Tuy nhiên, trong phạm vi nghiên cứu này, các mô hình đề xuất chưa được triển khai hoàn chỉnh trên vi điều khiển, thiết bị đeo công suất thấp hoặc hệ thống nhúng trong môi trường sử dụng thực tế. Khả năng hướng tới thiết bị biên được đánh giá thông qua các chỉ số như số lượng tham số, chi phí tính toán, bộ nhớ sử dụng và thời gian suy luận. Việc triển khai và đánh giá trực tiếp trên phần cứng nhúng được xác định là một hướng nghiên cứu tiếp theo.

- **Bối cảnh ứng dụng:** Phạm vi của nghiên cứu được giới hạn ở phân tích ngoại tuyến trên dữ liệu cảm biến đã được thu thập và phân đoạn. Luận án chưa thực hiện đánh giá chất lượng giấc ngủ, chưa kiểm chứng trên dữ liệu ngủ tự nhiên qua đêm và chưa tiến hành xác nhận lâm sàng trên các nhóm bệnh nhân. Các bằng chứng lâm sàng được sử dụng trong luận án nhằm làm rõ động cơ khoa học cho việc phân loại tư thế ngủ chi tiết, thay vì đưa ra kết luận trực tiếp về chẩn đoán hoặc điều trị.

4. Nội dung của luận án

Dựa trên các mục tiêu nghiên cứu đã được trình bày ở trên, luận án này xác định các nội dung nghiên cứu chính sau đây, sẽ được trình bày chi tiết trong các chương tiếp theo:

- Khảo sát và phân tích các nghiên cứu liên quan đến bài toán phân loại tư thế ngủ sử dụng cảm biến đeo, đặc biệt là các phương pháp dựa trên cảm biến gia tốc, học máy truyền thống, học sâu và các giao thức đánh giá độc lập theo đối tượng. Trên cơ sở đó, luận án xác định các khoảng trống nghiên cứu liên quan đến phân loại tư thế ngủ chi tiết, khả năng tổng quát hóa cho người dùng mới và yêu cầu kiểm soát độ phức tạp mô hình.
- Xây dựng và mô tả bộ dữ liệu tư thế ngủ sử dụng một cảm biến gia tốc kế ba trục gắn tại vùng bụng. Bộ dữ liệu bao gồm 12 tư thế ngủ tĩnh được xác định trước và được thu thập trong điều kiện có kiểm soát. Dữ liệu sau khi thu thập được tiền xử lý, phân đoạn và tổ chức phục vụ cho các thí nghiệm phân loại chuỗi thời gian đa lớp.
- Đề xuất mô hình AnpoNet, một kiến trúc học sâu lai kết hợp 1D CNN và LSTM cho bài toán phân loại 12 tư thế ngủ. Mô hình được thiết kế nhằm khai thác các đặc trưng cục bộ và phụ thuộc thời gian của tín hiệu gia tốc, đồng thời kiểm soát số lượng tham số và chi phí tính toán theo định hướng triển khai trên thiết bị tài nguyên hạn chế.
- Đề xuất mô hình CGB-Net, một kiến trúc học sâu lai kết hợp 1D CNN, GRU và BiLSTM nhằm tăng cường năng lực biểu diễn đặc trưng không gian-thời gian của tín hiệu gia tốc. Mô hình này được xây dựng nhằm cải thiện hiệu quả phân loại trong điều kiện dữ liệu mở rộng và được so sánh với các mô hình cơ sở cũng như mô hình AnpoNet.

- Thực hiện đánh giá thực nghiệm các mô hình đề xuất trong điều kiện độc lập theo đối tượng. Các kết quả được phân tích thông qua các chỉ số đánh giá như accuracy, precision, recall và F1-score, đồng thời kết hợp với ma trận nhầm lẫn, phân tích theo từng lớp tư thế, đánh giá ảnh hưởng của độ dài cửa sổ, tỷ lệ chồng lấp và độ ổn định của mô hình qua các lần khởi tạo khác nhau.
- Tổng hợp kết quả nghiên cứu, thảo luận các ưu điểm và hạn chế của các mô hình đề xuất, từ đó xác định các hướng phát triển tiếp theo như mở rộng dữ liệu ngủ tự nhiên qua đêm, kiểm chứng trên nhóm đối tượng đa dạng hơn và triển khai thực nghiệm trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.

5. Phương pháp nghiên cứu

Phương pháp nghiên cứu của luận án được xây dựng theo hướng kết hợp giữa nghiên cứu tổng quan, thiết kế dữ liệu thực nghiệm, mô hình hóa bằng học sâu và đánh giá định lượng. Các phương pháp chính được sử dụng trong luận án bao gồm:

1. Phương pháp tổng quan tài liệu và xác lập khoảng trống nghiên cứu:

Luận án tiến hành khảo sát các nghiên cứu liên quan đến theo dõi tư thế ngủ, các hệ thống cảm biến đeo, các phương pháp học máy truyền thống và học sâu cho dữ liệu chuỗi thời gian cảm biến. Trên cơ sở tổng quan tài liệu, luận án xác định các khoảng trống nghiên cứu chính, bao gồm: sự hạn chế của các nghiên cứu phân loại tư thế ngủ ở mức 3–4 lớp cơ bản, thiếu đánh giá độc lập theo đối tượng trong nhiều công trình trước đây, và nhu cầu thiết kế mô hình học sâu có độ phức tạp được kiểm soát cho các hệ thống cảm biến đeo.

2. Phương pháp thiết kế và xây dựng bộ dữ liệu:

Luận án xây dựng bộ dữ liệu tư thế ngủ dựa trên tín hiệu gia tốc ba trục thu thập từ một cảm biến đeo tại vùng bụng. Các tư thế được xác định trước theo 12 nhãn rời rạc, bao gồm các tư thế cơ bản và các tư thế có độ lệch góc chi tiết hơn. Quá trình thu thập dữ liệu được thực hiện trong điều kiện có kiểm soát nhằm đảm bảo sự nhất quán về vị trí cảm biến, thời lượng ghi, tần số lấy mẫu và cách thực hiện tư thế giữa các đối tượng tham gia.

3. Phương pháp tiền xử lý và tổ chức dữ liệu:

Dữ liệu gia tốc sau khi thu thập được tiền xử lý, chuẩn hóa và phân đoạn thành các cửa sổ tín hiệu có độ dài xác định. Các cửa sổ dữ liệu được gán nhãn theo tư thế tương ứng và được

tổ chức phục vụ cho bài toán phân loại chuỗi thời gian đa lớp. Việc tổ chức dữ liệu được thực hiện theo nguyên tắc độc lập theo đối tượng, trong đó dữ liệu của đối tượng kiểm tra không xuất hiện trong quá trình huấn luyện mô hình.

4. **Phương pháp mô hình hóa bằng học sâu:** Luận án đề xuất hai mô hình học sâu lai cho bài toán phân loại 12 tư thế ngủ. Mô hình AnpoNet kết hợp 1D CNN và LSTM nhằm khai thác đặc trưng cục bộ và phụ thuộc thời gian của tín hiệu gia tốc với số lượng tham số được kiểm soát. Mô hình CGB-Net kết hợp 1D CNN, GRU và BiLSTM nhằm tăng cường năng lực biểu diễn đặc trưng không gian-thời gian trong điều kiện dữ liệu mở rộng. Các mô hình được huấn luyện theo hướng học biểu diễn trực tiếp từ tín hiệu gia tốc đã phân đoạn, thay vì phụ thuộc hoàn toàn vào các đặc trưng thủ công.
5. **Phương pháp đánh giá thực nghiệm:** Hiệu năng của các mô hình được đánh giá thông qua các chỉ số phổ biến trong bài toán phân loại đa lớp, bao gồm accuracy, precision, recall và F1-score. Bên cạnh đó, luận án sử dụng ma trận nhầm lẫn để phân tích lỗi theo từng lớp tư thế, so sánh với các mô hình cơ sở, đánh giá ảnh hưởng của độ dài cửa sổ và tỷ lệ chồng lấp, cũng như kiểm tra độ ổn định của mô hình qua các lần khởi tạo khác nhau. Các thí nghiệm chính được thiết kế theo nguyên tắc độc lập theo đối tượng nhằm đánh giá khả năng tổng quát hóa của mô hình đối với người dùng mới.
6. **Phương pháp phân tích và thảo luận kết quả:** Các kết quả thực nghiệm được phân tích nhằm làm rõ ưu điểm, hạn chế và phạm vi áp dụng của từng mô hình đề xuất. Việc thảo luận không chỉ tập trung vào giá trị trung bình của các chỉ số đánh giá mà còn xem xét sự khác biệt giữa các lớp tư thế, các trường hợp nhầm lẫn phổ biến, độ phức tạp mô hình và khả năng hướng tới triển khai trên thiết bị tài nguyên hạn chế.

6. Đóng góp chính của luận án

Với mục tiêu nâng cao hiệu quả phân loại tư thế ngủ chi tiết từ dữ liệu cảm biến gia tốc ba trục, luận án có hai nhóm đóng góp chính như sau:

- Nhóm thứ nhất, đề xuất mô hình AnpoNet cho bài toán phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục đeo tại vùng bụng. Mô hình đề xuất được xây dựng nhằm giải quyết đồng thời hai yêu cầu: nâng cao hiệu quả nhận dạng các tư thế ngủ chi tiết và kiểm soát độ phức tạp mô hình theo định hướng triển khai trên thiết

bị cảm biến đeo hoặc thiết bị có tài nguyên hạn chế. AnpoNet kết hợp khả năng trích xuất đặc trưng cục bộ của mạng nơ-ron tích chập một chiều với khả năng mô hình hóa phụ thuộc thời gian của mạng hồi quy, qua đó khai thác được đặc trưng không gian-thời gian của tín hiệu gia tốc. Kết quả thực nghiệm cho thấy mô hình AnpoNet đạt hiệu quả phân loại cao trong bài toán 12 tư thế ngủ, với độ chính xác 94,67% và F1-score 92,94% trên tập kiểm tra. Bên cạnh đó, mô hình chỉ sử dụng 1.916 tham số, yêu cầu bộ nhớ khoảng 34 kB, có chi phí tính toán 0,1603M FLOPs, thời gian suy luận 0,081 giây cho mỗi lần suy luận. Các kết quả này cho thấy AnpoNet đạt được sự cân bằng giữa độ chính xác phân loại và độ gọn nhẹ của mô hình, phù hợp với các hệ thống theo dõi tư thế ngủ sử dụng thiết bị đeo. Phương pháp đề xuất này đã được công bố tại công trình [CT1]. Hướng tiếp cận sử dụng cảm biến gia tốc kết hợp mô hình LSTM cho phân loại tư thế ngủ cũng đã được công bố tại công trình [CT4].

- Nhóm thứ hai, đề xuất mô hình CGB-Net nhằm tăng cường độ chính xác và khả năng tổng quát hóa cho bài toán phân loại 12 tư thế ngủ trong điều kiện dữ liệu mở rộng. Mô hình CGB-Net được thiết kế trên cơ sở kết hợp ba thành phần gồm mạng nơ-ron tích chập một chiều, GRU và BiLSTM. Trong đó, mạng tích chập một chiều được sử dụng để trích xuất đặc trưng cục bộ từ tín hiệu gia tốc, GRU hỗ trợ mô hình hóa quan hệ tuần tự với chi phí tham số hợp lý, còn BiLSTM khai thác ngữ cảnh thời gian hai chiều nhằm tăng cường khả năng phân biệt giữa các tư thế có đặc trưng gần nhau. Kết quả thực nghiệm cho thấy CGB-Net đạt hiệu quả tốt trong các thiết lập đánh giá khác nhau, đặc biệt ở thiết lập độc lập theo đối tượng. Trong thiết lập thí nghiệm thứ nhất, CGB-Net đạt độ chính xác 89,34% và F1-score 86,24%, cao hơn AnpoNet khoảng 4,14 điểm phần trăm về độ chính xác và hơn 10 điểm phần trăm về F1-score. Sự cải thiện rõ rệt của F1-score cho thấy mô hình đề xuất không chỉ nâng cao tỷ lệ dự đoán đúng tổng thể mà còn cải thiện khả năng phân loại cân bằng giữa các lớp tư thế ngủ. Các phân tích về ma trận nhầm lẫn, độ ổn định qua các lần khởi tạo ngẫu nhiên, ảnh hưởng của độ dài chuỗi và tỷ lệ chồng lấp cũng cho thấy CGB-Net duy trì hiệu quả ổn định trong điều kiện dữ liệu có sự khác biệt giữa các cá nhân. Phương pháp đề xuất này đã được công bố tại công trình [CT2].

7. Cấu trúc của luận án

Ngoài phần Mở đầu, Kết luận và Kiến nghị, luận án được tổ chức thành ba chương chính. Cấu trúc các chương được xây dựng theo trình tự từ tổng quan cơ sở lý thuyết, xây dựng bài toán và dữ liệu thực nghiệm, đến đề xuất và đánh giá các mô

hình học sâu cho phân loại tư thế ngủ.

Chương 1 trình bày tổng quan về bài toán phân loại tư thế ngủ và các tiếp cận dựa trên cảm biến. Chương này phân tích ý nghĩa của việc theo dõi tư thế ngủ, đặc biệt là nhu cầu phân biệt các tư thế ở mức độ chi tiết hơn so với bốn tư thế cơ bản. Các hệ thống phân loại tư thế ngủ dựa trên thiết bị đeo, thị giác, cảm biến áp lực và các phương thức khác được khảo sát nhằm làm rõ ưu điểm, hạn chế và phạm vi ứng dụng của từng nhóm phương pháp. Bên cạnh đó, chương này trình bày cơ sở lý thuyết về các phương pháp học máy và học sâu cho dữ liệu cảm biến chuỗi thời gian, bao gồm 1D CNN, LSTM, GRU, BiLSTM và các kiến trúc lai CNN–RNN. Từ các phân tích này, chương xác định các khoảng trống nghiên cứu liên quan đến phân loại 12 tư thế ngủ, đánh giá độc lập theo đối tượng và yêu cầu kiểm soát độ phức tạp mô hình.

Chương 2 trình bày quá trình xây dựng dữ liệu thực nghiệm và đề xuất mô hình AnpoNet cho bài toán phân loại 12 tư thế ngủ. Phần đầu của chương mô tả cấu hình cảm biến, thiết lập phần cứng, giao thức thu thập dữ liệu, quy trình gán nhãn, tiền xử lý và tổ chức dữ liệu. Trên cơ sở đó, chương đề xuất AnpoNet, một kiến trúc học sâu lai kết hợp 1D CNN và LSTM nhằm khai thác các đặc trưng cục bộ và phụ thuộc thời gian của tín hiệu gia tốc ba trục. Mô hình được thiết kế theo hướng gọn nhẹ, kiểm soát số lượng tham số và chi phí tính toán, phù hợp với định hướng triển khai trên thiết bị tài nguyên hạn chế. Chương này cũng trình bày các thí nghiệm so sánh, phân tích độ phức tạp, đánh giá theo từng lớp tư thế, khảo sát ảnh hưởng của độ dài cửa sổ, tỷ lệ chồng lấp và độ ổn định của mô hình trong điều kiện đánh giá độc lập theo đối tượng.

Chương 3 trình bày mô hình CGB-Net nhằm tăng cường năng lực biểu diễn đặc trưng không gian–thời gian cho bài toán phân loại 12 tư thế ngủ. Chương này trước hết phân tích vấn đề nghiên cứu và các hạn chế còn tồn tại khi chỉ sử dụng các kiến trúc gọn nhẹ trong điều kiện dữ liệu mở rộng. Từ đó, mô hình CGB-Net được đề xuất trên cơ sở kết hợp 1D CNN, GRU và BiLSTM nhằm khai thác đồng thời đặc trưng cục bộ, quan hệ tuần tự và ngữ cảnh hai chiều của tín hiệu gia tốc. Chương này trình bày kiến trúc mô hình, quy trình huấn luyện, các thiết lập thực nghiệm, kết quả so sánh với các mô hình cơ sở và AnpoNet, cũng như các phân tích về ma trận nhầm lẫn, độ ổn định theo các lần khởi tạo ngẫu nhiên, ảnh hưởng của độ dài cửa sổ và tỷ lệ chồng lấp.

Phần Kết luận và Kiến nghị tổng hợp các kết quả chính đạt được của luận

án, làm rõ các đóng góp về bài toán, mô hình và phân tích thực nghiệm. Đồng thời, phần này thảo luận các hạn chế còn tồn tại, bao gồm phạm vi dữ liệu còn giới hạn ở các tư thế tĩnh được thu thập trong điều kiện có kiểm soát, việc chưa kiểm chứng trên dữ liệu ngẫu nhiên qua đêm và chưa triển khai hoàn chỉnh trên phần cứng nhúng. Trên cơ sở đó, luận án đề xuất một số hướng nghiên cứu tiếp theo như mở rộng dữ liệu, tích hợp thêm cảm biến, đánh giá trên nhóm đối tượng đa dạng hơn và triển khai thực nghiệm trên thiết bị đeo công suất thấp.

CHƯƠNG 1. TỔNG QUAN VỀ PHÂN LOẠI TƯ THẾ NGỦ VÀ CÁC TIẾP CẬN DỰA TRÊN CẢM BIẾN

Nội dung trong Chương 1 này trình bày tổng quan về bài toán phân loại tư thế ngủ và các hướng tiếp cận dựa trên cảm biến. Trước hết, chương phân tích ý nghĩa lâm sàng của việc theo dõi tư thế ngủ, đặc biệt trong các bệnh lý có liên quan đến tư thế như ngưng thở khi ngủ do tắc nghẽn, trào ngược dạ dày thực quản, đau cơ xương khớp và nguy cơ loét do tỳ đè. Bên cạnh các tư thế ngủ cơ bản, chương cũng làm rõ nhu cầu phân loại tư thế ở mức độ chi tiết hơn, bao gồm các tư thế trung gian và các sai lệch góc nhỏ có ý nghĩa trong theo dõi lâm sàng. Tiếp theo, các hệ thống phân loại tư thế ngủ hiện có được khảo sát và phân tích theo các nhóm tiếp cận khác nhau, bao gồm cảm biến đeo, cảm biến không tiếp xúc, cảm biến áp lực và các phương thức liên quan. Ngoài ra, chương trình bày các phương pháp học máy và học sâu thường được sử dụng cho dữ liệu cảm biến chuỗi thời gian, như CNN, RNN, LSTM, GRU và các kiến trúc lai CNN-RNN. Những phân tích trong chương này làm cơ sở để xác định khoảng trống nghiên cứu và định hướng đề xuất các mô hình phân loại tư thế ngủ ở các chương tiếp theo.

1.1 Bài toán phân loại tư thế ngủ

1.1.1 Tầm quan trọng về mặt lâm sàng của theo dõi tư thế ngủ

Tư thế ngủ đã được chứng minh trong lâm sàng thể hiện ảnh hưởng đến các tình trạng sức khỏe khác nhau trong các cơ quan như hô hấp, tim mạch, tiêu hóa và cơ xương khớp [10,11]. Đặc biệt, nhiều nghiên cứu đã chứng minh rằng tư thế ngủ ảnh hưởng đáng kể đến chức năng hô hấp và đặc biệt ở bệnh nhân mắc chứng ngưng thở khi ngủ do tắc nghẽn (OSA) [12].

Nhiều nghiên cứu xác nhận rằng tư thế nằm ngửa (supine) có liên quan đến việc tăng khả năng xẹp đường hô hấp trên, chỉ số ngưng thở (AHI) cao hơn và tình trạng khử bão hòa oxy nghiêm trọng hơn so với các tư thế không nằm ngửa ở bệnh nhân OSA [13]. Các phép đo khách quan cho thấy áp lực tối hạn vùng hầu họng (Pcrit), một chỉ số về khả năng xẹp đường thở, cao hơn đáng kể ở tư thế nằm ngửa so với tư thế nằm nghiêng (lateral) trong tất cả các giai đoạn ngủ, cho thấy đường thở dễ bị xẹp hơn khi nằm ngửa [14]. Các nghiên cứu khác còn chứng minh rằng cả tần suất và thời gian

ngưng thở, cũng như độ sâu và thời gian của các đợt khử bão hòa oxy, đều lớn hơn ở tư thế nằm ngửa mà không phụ thuộc vào mức độ nghiêm trọng của OSA [14,15]. Các nghiên cứu giải phẫu cho thấy rằng tư thế nằm ngửa dẫn đến sự dịch chuyển về phía sau của lưỡi và vòm miệng, làm giảm không gian đường hô hấp trên và tăng khả năng tắc nghẽn [15,16]. Những thay đổi sinh lý này dẫn đến AHI cao hơn và tình trạng khử bão hòa oxy rõ rệt hơn trong khi ngủ nằm ngửa, những phát hiện này nhất quán ở các nhóm tuổi và cả hai giới [15,17]. Do đó, tư thế nằm ngửa là một yếu tố nguy cơ chính có thể điều chỉnh được, làm trầm trọng thêm mức độ nghiêm trọng của OSA thông qua việc tăng khả năng xẹp đường thở.

Các nghiên cứu đã chứng minh rằng liệu pháp tư thế, khuyến khích tránh tư thế nằm ngửa, có thể giảm đáng kể chỉ số ngưng thở-giảm thở (AHI) ở bệnh nhân mắc OSA phụ thuộc tư thế. Mặc dù liệu pháp tư thế thường kém hiệu quả hơn so với máy thở áp lực dương liên tục (CPAP) [18] trong việc giảm AHI tổng thể và cải thiện quá trình oxy hóa, nhưng nó mang lại những ưu điểm đáng chú ý như ít biến chứng liên quan đến thiết bị hơn và sự tuân thủ của bệnh nhân cao hơn, khiến nó trở thành một lựa chọn thay thế khả thi cho những người không dung nạp CPAP [18-20]. Về mặt cơ chế, tư thế nằm ngửa làm trầm trọng thêm OSA do tác động của trọng lực làm giảm dung tích cặn chức năng (FRC), thúc đẩy sự dịch chuyển về phía sau của lưỡi và vòm miệng, và tăng áp lực đóng hết vùng hầu họng (Pcrit), tất cả đều góp phần làm tăng khả năng xẹp và tắc nghẽn đường hô hấp trên trong khi ngủ [21-23]. Những thay đổi sinh lý này giải thích tại sao liệu pháp tư thế đặc biệt hiệu quả ở những bệnh nhân có tình trạng OSA trở nên tồi tệ hơn khi nằm ngửa, vì việc chuyển sang ngủ nghiêng giúp cải thiện sự thông thoáng của đường thở và giảm mức độ nghiêm trọng của các biến cố hô hấp [21,23].

Trong một bệnh lý khác, nhiều nghiên cứu đã chứng minh rằng tư thế ngủ ảnh hưởng sâu sắc đến sinh lý bệnh và các biểu hiện lâm sàng của bệnh trào ngược dạ dày thực quản (GERD), đặc biệt là trong giai đoạn ban đêm. Trong số các tư thế khác nhau, tư thế nằm nghiêng trái (LLD) liên tục được báo cáo là có lợi nhất cho bệnh nhân mắc GERD [24-26].

Một số bài nghiên cứu tổng quan hệ thống cho thấy rằng so với tư thế nằm nghiêng phải (RLD) và nằm ngửa, tư thế LLD làm giảm đáng kể thời gian tiếp xúc với axit thực quản (AET) và thời gian làm sạch axit (ACT) vào ban đêm. Cụ thể, hai nghiên cứu đã chứng minh rằng AET và ACT ở tư thế LLD giảm lần lượt 2,03% và

81,84 s so với RLD, và giảm 2,71% và 74,47 s so với tư thế nằm ngửa [24,25,27]. Hơn nữa, một thử nghiệm kiểm chứng ngẫu nhiên sử dụng thiết bị liệu pháp tư thế điện tử cho thấy việc tăng thời gian ngủ ở tư thế LLD đã cải thiện rõ rệt các triệu chứng trào ngược ban đêm, tăng số đêm không trào ngược và nâng cao chất lượng cuộc sống liên quan đến GERD [26].

Cơ chế sinh lý đằng sau lợi ích này dựa trên mối quan hệ trọng lực và giải phẫu giữa dạ dày và thực quản. Ở tư thế LLD, dạ dày nằm thấp hơn vị trí của thực quản, điều này tạo điều kiện thuận lợi cho việc thoát dịch dạ dày nhờ trọng lực và giảm áp lực trong thực quản, từ đó làm giảm khả năng xảy ra các đợt trào ngược axit. Ngược lại, tư thế RLD và nằm ngửa làm dạ dày ở vị trí cao hơn so với thực quản hoặc đặt cả hai ở mức tương đương, tạo điều kiện cho các biến cố trào ngược [24-26]. Các nghiên cứu sử dụng giám sát pH-trở kháng kết hợp với cảm biến tư thế đã xác nhận rằng AET và ACT ở tư thế LLD thấp hơn đáng kể so với các tư thế khác, đồng thời tần suất và thời gian của các đợt trào ngược cũng giảm rõ rệt [25,26].

Ngoài ra, các can thiệp hỗ trợ tư thế như thiết bị phản hồi rung xúc giác đeo trên người hoặc gối chuyên dụng thúc đẩy tư thế LLD trong khi ngủ đã được chứng minh là làm tăng tỷ lệ thành công trong điều trị, giảm các triệu chứng ban đêm và cải thiện chất lượng cuộc sống ở bệnh nhân GERD, bao gồm cả những người đáp ứng kém với liệu pháp ức chế bơm proton (PPI) [26,28,29]. Đáng chú ý, việc kết hợp tư thế LLD với nâng cao đầu giường (khoảng 30°) giúp giảm thêm điểm số trào ngược ban đêm so với việc chỉ thực hiện một trong hai can thiệp [26].

Tóm lại, các bằng chứng hiện tại chỉ ra rằng việc duy trì tư thế ngủ nghiêng trái là một can thiệp phi dược lý đơn giản, an toàn và hiệu quả để giảm tiếp xúc với axit thực quản ban đêm, giảm bớt các triệu chứng trào ngược và cải thiện chất lượng cuộc sống ở bệnh nhân GERD. Do đó, các hướng dẫn lâm sàng nên xem xét khuyến nghị tư thế LLD, đặc biệt đối với bệnh nhân có triệu chứng GERD ban đêm, kết hợp với các chiến lược hỗ trợ tư thế thích hợp [24-26].

Tổng hợp lại, những phát hiện này nhấn mạnh rằng tư thế ngủ không chỉ đơn thuần là một biến số mô tả mà là một biomarker có thể điều chỉnh được, kết nối giữa sinh lý, bệnh lý và can thiệp hành vi. Việc theo dõi liên tục và chính xác tư thế ngủ hỗ trợ các mục tiêu lâm sàng sau:

- Phân tầng bệnh nhân cho liệu pháp tư thế, ví dụ: xác định những người mắc OSA chủ yếu khi nằm ngửa, những người có thể hưởng lợi từ các thiết bị điều

chỉnh tư thế.

- Định lượng thời gian dành cho các tư thế "nguy cơ cao" như ngủ nằm ngửa vượt quá 50% tổng thời gian ngủ.
- Đánh giá dài hạn tại nhà mà không cần sử dụng đo đa ký giấc ngủ (PSG) một đêm tại phòng thí nghiệm, nắm bắt tư thế thói quen và sự tuân thủ điều trị.

Những lý do này cho thấy sự cần thiết của việc phát triển các hệ thống đeo trên người đáng tin cậy để nhận dạng tư thế ngủ, có thể cho phép theo dõi chính xác trong cả môi trường lâm sàng và tại nhà.

1.1.2 Tư thế độ phân giải cao và tư thế trung gian trong nghiên cứu lâm sàng

Nghiên cứu gần đây công nhận rằng tư thế ngủ không chỉ giới hạn ở bốn tư thế cơ bản như nằm ngửa, nằm sấp, nghiêng trái và nghiêng phải mà thay vào đó tư thế ngủ cũng có một dải liên tục. Những biến đổi nhỏ, chẳng hạn như xoay thân một phần, nâng cao đầu cũng có thể gây ra những tác động sinh lý và lâm sàng đáng kể. Các nghiên cứu đã chỉ ra rằng những thay đổi tư thế nhỏ này có thể ảnh hưởng đến chức năng hô hấp, chất lượng giấc ngủ và thậm chí cả hoạt động của não. Ví dụ, tư thế nằm ngửa có liên quan đến tần suất và mức độ nghiêm trọng của các biến cố ngưng thở khi ngủ do tắc nghẽn (OSA) cao hơn so với các tư thế nằm nghiêng, và hơn một nửa số bệnh nhân OSA được coi là phụ thuộc tư thế, trải qua số lần ngưng thở hoặc giảm thở khi nằm ngửa nhiều gấp ít nhất hai lần so với khi nằm nghiêng [30,31]. Việc nâng cao đầu và thân đã được chứng minh là làm giảm xẹp đường hô hấp trên và cải thiện chỉ số ngưng thở-giảm thở (AHI), đặc biệt là ở những người thừa cân hoặc béo phì [31,32].

Ngoài các tác động về hô hấp, tư thế ngủ còn ảnh hưởng đến sức khỏe cơ xương khớp và hoạt động thần kinh. Các tư thế nằm nghiêng có liên quan đến sự khác biệt về thời gian bắt đầu giấc ngủ và các mẫu kích hoạt não, gợi ý rằng sở thích tư thế và các tư thế ngủ theo thói quen có thể định hình cả cấu trúc giấc ngủ và chức năng vỏ não [33-35]. Hơn nữa, chỉ riêng vị trí đầu—chẳng hạn như xoay đầu trong khi thân vẫn nằm ngửa đã được chứng minh là có thể làm giảm mức độ nghiêm trọng của OSA một cách độc lập, đặc biệt là ở những cá nhân không béo phì [32,36,37]. Ngược lại, việc gập đầu có thể làm trầm trọng thêm mức độ nghiêm trọng của OSA [37]. Những

phát hiện này nhấn mạnh tầm quan trọng của việc xem xét các tư thế ngủ phức tạp hơn có những tác động lớn đối với quản lý bệnh, các can thiệp trị liệu và chất lượng giấc ngủ tổng thể [31,33,36,37].

Bằng chứng từ nghiên cứu [38] cung cấp giải pháp với các việc sử dụng liệu pháp tư thế của độ lệch góc hợp lý ở các bệnh nhân sau đột quỵ. Trong nghiên cứu thực nghiệm này với sự tham gia của 112 người cao tuổi, các tác giả đã kiểm tra hiệu quả của việc duy trì tư thế ngủ bán nghiêng phải hoặc trái với một góc 30° bằng hệ thống gối kiểm soát tư thế. Những người tham gia trong nhóm can thiệp cho thấy chất lượng giấc ngủ tốt hơn rõ rệt trong đó 85,7% đạt được "giấc ngủ ngon" theo Chỉ số Chất lượng Giấc ngủ Pittsburgh (PSQI), so với 66,1% ở nhóm đối chứng. Kích thước hiệu ứng lớn (Cohen's $d \approx 2.07$) đã xác nhận những cải thiện có ý nghĩa thống kê và lâm sàng trên nhiều lĩnh vực phụ của PSQI, bao gồm thời gian của giấc ngủ, hiệu quả và sự rối loạn ban đêm. Những kết quả này gợi ý rằng độ góc bán nghiêng trái hoặc phải ổn định (30°) giúp tăng cường đáng kể cả khả năng giúp cơ thể nghỉ ngơi và phục hồi thần kinh, nhấn mạnh tiềm năng lâm sàng của việc định vị góc của cơ thể khi ngủ có kiểm soát trong phục hồi chức năng thần kinh.

Năm 2017, Pickenbrock và cộng sự [8] đã nghiên cứu các tác động cơ sinh học của độ nghiêng của thân mình ở những bệnh nhân bất động hoặc nằm liệt giường. Sử dụng hệ thống bản đồ áp suất độ phân giải cao, họ đã chứng minh rằng nằm bán nghiêng 30° làm giảm đáng kể áp lực vùng xương cụt, một trong những yếu tố nguy cơ quan trọng nhất đối với việc hình thành vết loét do tì đè trong khi vẫn giữ cho cột sống thẳng và sự thoải mái của người bệnh. Tuy nhiên, các góc lớn hơn 30° có liên quan đến việc giảm khả năng dung nạp, cho thấy sự cân bằng giữa việc giảm tải hiệu quả và sự thoải mái của bệnh nhân. Nghiên cứu này đã xác định 30° là một sự lựa chọn tối ưu giữa lợi ích sinh lý và sự ổn định cơ học, nhấn mạnh nhu cầu kiểm soát góc chính xác trong chăm sóc dài hạn và định vị sau phẫu thuật.

Phân tích định lượng sâu hơn đã được trình bày trong [39], so sánh một cách hệ thống sự phân phối áp suất qua các góc xoay 0° , 15° , 30° và 45° . Các tác giả nhận thấy rằng cả áp lực trung bình và cao đều giảm dần khi tăng độ nghiêng bên, đặc biệt là ở các vùng xương cụt, xương bả vai và gót chân, những vị trí phổ biến của chấn thương do áp lực. Vị trí 30° mang lại sự cân bằng thuận lợi nhất giữa việc giảm áp lực và sự thoải mái, trong khi độ nghiêng 45° , mặc dù hiệu quả hơn về mặt cơ học, lại dẫn đến sự khó chịu gia tăng. Những kết quả này làm nổi bật rằng độ lệch toàn thân, chứ

không chỉ đơn thuần là hướng của thân, là một yếu tố quyết định then chốt đối với các chỉ số sinh lý, củng cố lý do cho việc phát hiện góc độ phân giải cao trong giám sát lâm sàng đối với các tư thế ngủ.

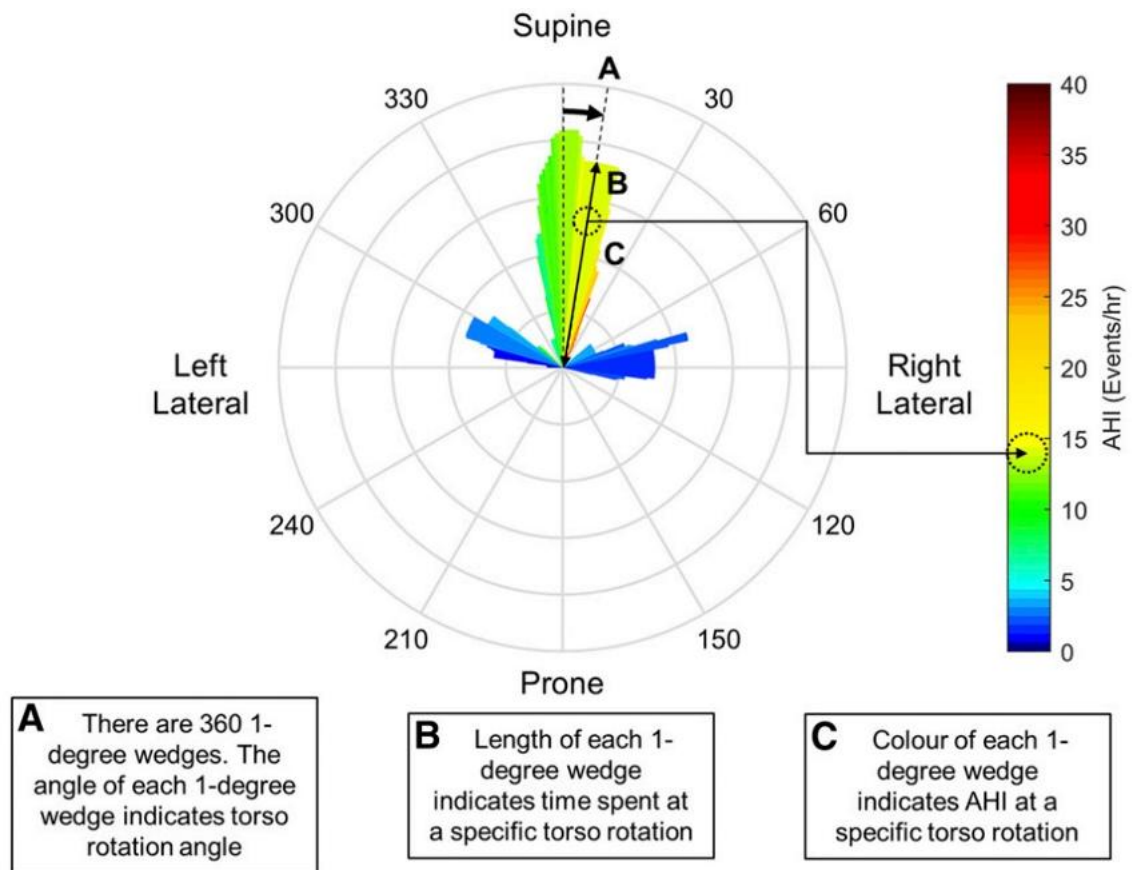
Định lượng độ phân giải cao các góc tư thế chi tiết trong OSA phụ thuộc tư thế. Nghiên cứu được trình bày trong [1] về cơ bản định nghĩa lại sử dụng lâm sàng để đánh giá tư thế ngủ bằng cách chứng minh rằng các độ lệch góc liên tục. Các tác giả lập luận rằng phân loại sử dụng góc 90 độ truyền thống (nằm ngửa, nằm sấp, nghiêng trái, nghiêng phải) quá đơn giản và không đầy đủ để hỗ trợ về mặt chẩn đoán, vì nó không thể hiện được hết các mối quan hệ quan trọng giữa việc xoay thân một phần và khả năng xẹp đường thở.

Để khắc phục hạn chế này, các nhà nghiên cứu đã phát triển một cảm biến tư thế dựa trên gia tốc kế ba trục tùy chỉnh được gắn chặt trên xương ức, có khả năng định lượng sự xoay thân như một biến góc liên tục tính bằng độ so với trục nằm ngửa. Cảm biến đạt được sai số đo góc trung bình chỉ $2,6^\circ$, trên toàn dải 180° , xác nhận tính phù hợp của nó đối với việc giám sát độ phân giải cao.

Để trực quan hóa, nghiên cứu đã giới thiệu một biểu diễn biểu đồ hình tròn mới, trong đó 360 độ góc rời rạc được sử dụng để hiển thị các phân phối tư thế như trong Hình 1.1. Mỗi rãnh quạt của hình tròn đại diện cho một phạm vi 30° (góc trung tâm $\pm 15^\circ$), cho phép trực quan hóa trực quan mối quan hệ thời gian–góc và sự thay đổi Chỉ số ngưng thở (AHI) tương ứng của nó. Cách tiếp cận này không chỉ bảo toàn tính liên tục của dữ liệu mà còn cho phép phân tích thống kê về sự kết hợp giữa tư thế và bệnh lý trên một dải góc liên tục thay vì thông qua các tư thế cơ bản.

Phân tích các dữ liệu độ phân giải cao này cho thấy 4 tư thế cơ bản chỉ nắm bắt được một phần nhỏ hành vi ngủ trong thế giới thực. Khi định nghĩa các tư thế cơ bản là trong vòng $\pm 15^\circ$ so với các trục giải phẫu trung tâm của chúng (ví dụ: nằm ngửa = $0^\circ \pm 15^\circ$), tỷ lệ phần trăm trung bình của tổng thời gian ngủ nằm trong các ranh giới này chỉ là $59,7 \pm 26,0\%$. Ngược lại, $40,3 \pm 26,0\%$ thời gian ngủ được dành cho các tư thế trung gian, các cấu hình hoàn toàn bị loại trừ khỏi phân loại thông thường. Những góc trung gian này, trước đây bị bỏ qua trong cả đánh giá lâm sàng và các thiết bị đeo, đã được chứng minh là có tác động sâu sắc đến sinh lý bệnh OSA và đáp ứng điều trị.

Cách tiếp cận này nghiên cứu tư thế ngủ theo các góc liên tục đã cho thấy mối quan hệ giữa sự xoay thân và mức độ nghiêm trọng của OSA. Ngay cả những độ lệch nhẹ so với trục nằm ngửa cũng mang lại lợi ích trị liệu đáng kể ở một số cá nhân nhất



Hình 1.1: Biểu đồ tròn để trực quan hóa của tư thế ngủ cho thấy thời gian dành cho mỗi góc thân mình ($\pm 15^\circ$) và chỉ số ngưng thở-giảm thở (AHI) tương ứng theo từng vị trí. 1

định. Ví dụ, một bệnh nhân có AHI ban đầu là 18 lần ngưng thở/giờ ở tư thế nằm ngửa đã trải qua sự sụt giảm nhanh chóng xuống dưới 5 lần ngưng thở/giờ chỉ với một góc xoay 15° sang trái. Những "góc xoay hiệu quả tối thiểu" như vậy làm nổi bật rằng ngay cả những thay đổi tư thế nhỏ, đặc thù cho từng bệnh nhân cũng có thể bình thường hóa các mẫu hô hấp, một hiểu biết hoàn toàn bị mất đi dưới cách đánh giá dựa trên góc 90 độ với 4 tư thế cơ bản.

Nghiên cứu trên cũng đã xác định các góc trị liệu được cá nhân hóa cho liệu pháp tư thế. Bằng cách định lượng mối quan hệ giữa AHI và hướng góc liên tục, các bác sĩ lâm sàng có thể chỉ định các vùng mục tiêu chính xác (ví dụ: tránh các tư thế trong vòng $\pm 60^\circ$ so với nằm ngửa) và thiết kế các can thiệp để duy trì bệnh nhân trong các dải nguy cơ thấp trong khi ngủ. Độ chi tiết về góc này cho phép phát triển các hệ thống đeo trên người tùy chỉnh, dễ dung nạp và hiệu quả để quản lý OSA phụ thuộc tư thế.

Về bản chất, nghiên cứu cung cấp bằng chứng thực nghiệm thuyết phục rằng

mức độ nghiêm trọng của OSA thay đổi liên tục theo góc thân, và các ngưỡng trị liệu quan trọng thường xảy ra ở độ phân giải góc chi tiết (15° – 60°). Do đó, nghiên cứu thiết lập cả nền tảng khái niệm và phương pháp luận cho việc theo dõi tư thế ngủ độ phân giải góc cao.

Như vậy, tư thế ngủ của con người không chỉ tồn tại dưới dạng 4 tư thế cơ bản mà còn theo một dải góc liên tục, ngay cả những độ lệch nhỏ 15° – 45° cũng mang lại các hiệu ứng sinh lý, thần kinh và trị liệu có thể đo lường được. Quan trọng hơn, bằng chứng về độ phân giải cao được trình bày trong [1] chứng minh rằng mức độ nghiêm trọng của chứng ngưng thở khi ngủ do tắc nghẽn (OSA) thay đổi liên tục theo góc thân mình, và các ngưỡng trị liệu quan trọng thường xảy ra ở độ phân giải góc chi tiết (15° – 60°). Điều này thách thức cơ bản mô hình bốn tư thế truyền thống, tạo ra động lực về lâm sàng và sinh lý mạnh mẽ cho việc theo dõi tư thế ngủ với nhiều góc lệch chi tiết. Tuy nhiên nghiên cứu của [1] chưa đề cập đến độ chính xác của thiết bị sử dụng cho việc theo dõi các tư thế chi tiết như trên.

Do đó, nhu cầu cấp thiết là một hệ thống có khả năng phát hiện và diễn giải những độ lệch góc chi tiết với độ chính xác và độ tin cậy cao. Một hệ thống như vậy phải kết hợp phần cứng có gắn cảm biến đeo trên người, có khả năng phát hiện các tư thế tinh vi trong môi trường ngủ thực tế. Việc phát triển một khung phân loại tư thế chi tiết dựa trên học sâu do đó không chỉ đơn thuần là một nỗ lực kỹ thuật, mà là xuất phát trực tiếp từ các nhu cầu lâm sàng và chẩn đoán chưa được đáp ứng. Luận án này xây dựng trên nền tảng đó, đề xuất một hệ thống cảm biến và mô hình học sâu được thiết kế phù hợp để cho phép giám sát tư thế ngủ với độ chính xác cao và được cá nhân hóa.

1.2 Tổng quan về các hệ thống phân loại tư thế ngủ

Có nhiều cách khác nhau để phân loại các phương pháp phát hiện tư thế ngủ. Một số phương pháp phân biệt các hệ thống dựa trên việc chúng là xâm lấn [40] hay không xâm lấn [41], trong khi những phương pháp khác phân loại chúng thành các kỹ thuật tiếp xúc hoặc không tiếp xúc. Luận án này phân loại dựa trên phương thức cảm biến chính được sử dụng để thu thập thông tin tư thế, nhất trí với cách tiếp cận được đề xuất bởi Li và cộng sự (2024) [42]. Theo đó, các phương pháp phân loại tư thế ngủ được nhóm thành ba loại chính: các phương pháp phát hiện dựa trên thiết bị đeo [43], các phương pháp phát hiện dựa trên thị giác, và các phương pháp dựa trên cảm biến

áp lực [17,42]. Phân loại dựa trên phương thức này cung cấp một cơ sở rõ ràng và có ý nghĩa về mặt chức năng để so sánh các hệ thống theo cách dữ liệu tư thế được thu thập, xử lý và diễn giải.

Phân loại dựa trên phương thức này đặc biệt phù hợp với các mục tiêu của luận án này, vốn tập trung chủ yếu vào các cách tiếp cận dựa trên thiết bị đeo để phát hiện tư thế ngủ. Các hệ thống đeo trên người thường tích hợp gia tốc kế, con quay hồi chuyển hoặc các đơn vị đo lường quán tính cho phép theo dõi tư thế liên tục và chi tiết với sự can thiệp tối thiểu vào hành vi ngủ tự nhiên. Do đó, việc áp dụng phương thức cảm biến không chỉ phản ánh các phân biệt công nghệ rộng hơn trong lĩnh vực mà còn phù hợp chặt chẽ với trọng tâm phương pháp luận của luận án này, cho phép phân tích mạch lạc và có cấu trúc về các kỹ thuật phát hiện bằng thiết bị đeo trong suốt luận án.

Trong những năm gần đây, công nghệ sóng điện từ (RF) đã được áp dụng rộng rãi trong các lĩnh vực như nhận dạng hoạt động [44], giám sát giấc ngủ [45], theo dõi tín hiệu sinh lý [46] và các lĩnh vực liên quan. Các cảm biến thường được sử dụng để nhận dạng tư thế ngủ dựa trên RF bao gồm radar, RFID [47] và các thiết bị WiFi [48,49]. Liu và cộng sự [47] đã sử dụng ba thẻ RFID đặt gần người ngủ để nắm bắt các mẫu nhịp thở. Bằng cách phân tích độ rõ nét của các tín hiệu hô hấp được trích xuất, các nghiên cứu có thể phân loại các tư thế ngủ. Tuy nhiên, cách tiếp cận này phụ thuộc nhiều vào việc đặt thẻ đúng cách và yêu cầu thêm giá đỡ ba chân khi không có tường gần giường. Hơn nữa, các tác giả đã báo cáo sự sụt giảm đáng kể về hiệu suất mô hình khi áp dụng trực tiếp cho người dùng mới.

Các phương pháp dựa trên WiFi nhận dạng tư thế ngủ bằng cách phân tích thông tin trạng thái kênh (CSI), vì các đặc điểm CSI thay đổi dưới các tư thế cơ thể khác nhau. Ví dụ, Liu và cộng sự [49] đã trích xuất các đặc trưng từ chuỗi thời gian CSI và sử dụng các phương pháp học máy như SVM để phân loại các tư thế ngủ. Bằng cách kiểm tra các tín hiệu CSI trong miền tần số, họ có thể trích xuất các mẫu nhịp thở của hai cá nhân cùng một lúc, vì các nhịp thở khác nhau tạo ra các đỉnh riêng biệt trong phổ. Tuy nhiên, phương pháp của họ không thể mở rộng để nhận dạng tư thế cho hai người dùng cùng một lúc, vì các đặc trưng không thể tách rời cho từng cá nhân.

Các phương pháp dựa trên radar tận dụng khả năng nắm bắt khoảng cách và hướng của người dùng. Nhận dạng tư thế có thể đạt được bằng cách phân tích các bản

đồ nhiệt vị trí do radar tạo ra hoặc các biến đổi khoảng cách gây ra bởi các chuyển động cơ thể như lật người [46, 50–52]. Kiriazi và cộng sự [52] và Islam và cộng sự [51] đã áp dụng radar tần số kép để phân loại các tư thế ngủ bằng cách thiết lập các ngưỡng trên các tham số phản xạ radar hoặc sử dụng các bộ phân loại học máy. Mặc dù họ đạt được độ chính xác cao, các mô hình của họ đã được huấn luyện và xác thực bằng dữ liệu từ cùng một đối tượng, làm hạn chế khả năng tổng quát hóa và giảm hiệu quả khi áp dụng cho người dùng mới. Piriyaajitakonkij và cộng sự [46] đã sử dụng radar băng thông siêu rộng xung vô tuyến (IR-UWB) để nắm bắt các tín hiệu thời gian–khoảng cách của các chuyển đổi tư thế. Họ đã sử dụng mạng nơ-ron để phân loại các chuyển đổi giữa bốn tư thế ngủ dựa trên cả các đặc trưng miền thời gian và miền tần số. Tuy nhiên, việc nhận dạng tư thế toàn diện yêu cầu xác định ít nhất sáu chuyển đổi (ngủ–ngiêng, ngủ–sấp, nghiêng–ngủ, nghiêng–sấp, sấp–ngiêng, sấp–ngủ). Việc nhận dạng các tư thế nghiêng trái và nghiêng phải còn làm tăng gấp đôi số lượng chuyển đổi cần thiết lên mười hai. Lai và cộng sự [32] cũng sử dụng radar IR-UWB và so sánh các cấu hình cảm biến khác nhau: radar đơn, radar kép và radar ba. Phương pháp của họ đạt được độ chính xác lần lượt là 0,586, 0,808 và 0,651. Mặc dù nhiều radar giúp cải thiện độ chính xác, chi phí cho người dùng cũng tăng lên. Yue và cộng sự [16] đã áp dụng mạng nơ-ron cho các bản đồ nhiệt vị trí để nhận dạng tư thế của mỗi người, nhưng mô hình của họ không thể tổng quát hóa cho các bối cảnh hai người. Hơn nữa, việc hiệu chuẩn với dữ liệu người dùng mới là cần thiết để duy trì hiệu suất giữa các đối tượng. Luo và cộng sự [33] đã đề xuất một phương pháp sử dụng radar FMCW đầu vào đơn, đầu ra đơn đặt phía trên giường. Bằng cách xây dựng các bản đồ khoảng cách–Doppler, họ có thể phân loại các tư thế ngủ. Tuy nhiên, trong các trường hợp nhiều người dùng, các tín hiệu từ hai cá nhân bị chồng lấp trong cả miền khoảng cách và Doppler, khiến phương pháp này không phù hợp cho các kịch bản nhiều người dùng.

Các hệ thống giám sát tư thế ngủ dựa trên thị giác sử dụng camera RGB [53], camera hồng ngoại [54] hoặc camera chiều sâu [55] để thu nhận hình ảnh hoặc video của người dùng, sau đó áp dụng các thuật toán xử lý ảnh và mô hình học sâu để nhận dạng tư thế ngủ. Camera RGB hoạt động trong phổ ánh sáng khả kiến và có khả năng cung cấp thông tin hình dạng cơ thể với độ phân giải cao trong điều kiện quan sát thuận lợi. Camera hồng ngoại có thể hoạt động trong điều kiện ánh sáng yếu hoặc ban đêm, trong khi camera chiều sâu bổ sung thông tin không gian ba chiều, qua đó hỗ trợ

phân biệt hình dạng và hướng cơ thể. Nhìn chung, ưu điểm nổi bật của các phương pháp thị giác là không yêu cầu gắn cảm biến trực tiếp lên cơ thể người dùng, nhờ đó giảm sự can thiệp vật lý trong quá trình theo dõi giấc ngủ.

Một nghiên cứu gần đây [43] đã thực hiện so sánh trực tiếp giữa phương pháp dựa trên ảnh và phương pháp sử dụng cảm biến gia tốc cho bài toán phân loại năm trạng thái/tư thế ngủ. Trong thiết lập này, mô hình thị giác sử dụng VGG16 đạt độ chính xác 93,49%, trong khi phương pháp dựa trên cảm biến gia tốc đạt độ chính xác trên 99,8%. Bên cạnh khác biệt về hiệu suất phân loại, dữ liệu đầu vào của mô hình ảnh chiếm khoảng 343,15 MB so với 22,3 MB của dữ liệu gia tốc, và thời gian huấn luyện lần lượt là khoảng 1680 s và 18 s. Các kết quả này cho thấy, trong thiết lập thực nghiệm cụ thể, cảm biến gia tốc có thể đạt hiệu quả phân loại cao với khối lượng dữ liệu và chi phí tính toán thấp hơn đáng kể so với phương pháp dựa trên ảnh. Tuy nhiên, các giá trị này không nên được xem là phép so sánh tuyệt đối giữa hai phương thức do sự khác biệt về dữ liệu, kiến trúc mô hình và giao thức thực nghiệm.

Ở phạm vi rộng hơn, các nghiên cứu tổng quan về nhận dạng hoạt động người cũng chỉ ra rằng phương pháp dựa trên thị giác và phương pháp dựa trên cảm biến đeo tồn tại những đặc tính bổ sung nhau. Các hệ thống thị giác có ưu điểm về khả năng thu nhận thông tin hình dạng và ngữ cảnh mà không cần gắn thiết bị lên cơ thể, nhưng thường xử lý dữ liệu ảnh hoặc video có kích thước lớn và có thể bị ảnh hưởng bởi điều kiện quan sát. Ngược lại, các cảm biến quán tính tạo ra dữ liệu chuỗi thời gian có số chiều thấp hơn, phù hợp hơn với xử lý liên tục trên các thiết bị có tài nguyên hạn chế.

Đối với giám sát tư thế ngủ, các phương pháp dựa trên thị giác vẫn tồn tại một số hạn chế thực tế. Camera RGB phụ thuộc vào điều kiện ánh sáng; khả năng nhận dạng có thể suy giảm khi cơ thể bị che khuất bởi chăn hoặc mền [53,54]; đồng thời việc sử dụng camera trong phòng ngủ làm phát sinh các vấn đề về quyền riêng tư. Trong khi đó, cảm biến gia tốc đeo trên cơ thể không phụ thuộc vào điều kiện chiếu sáng hoặc tầm nhìn trực tiếp, đồng thời có thể phản ánh sự thay đổi hướng và độ nghiêng cơ thể thông qua thành phần trọng trường. Vì vậy, luận án lựa chọn cảm biến gia tốc ba trục nhằm nghiên cứu bài toán phân loại tư thế ngủ chi tiết với dữ liệu đầu vào có độ phức tạp thấp, đồng thời phù hợp với định hướng phát triển hệ thống trên thiết bị đeo và thiết bị biên có tài nguyên hạn chế.

Bảng 1.1: Các công trình liên quan sử dụng gia tốc kế để phân loại tư thế ngủ

Phương pháp của	Năm	Vị trí đặt cảm biến	Số lượng cảm biến	Số lượng tư thế	Tần số lấy mẫu (Hz)	Số lượng người tham gia	Độ chính xác (%)	Mô hình
Jiang et.al. [56]	2016	ngực	2	4	100	20	–	Ngưỡng (Threshold)
Fallman et.al. [57]	2017	ngực, cổ chân	3	8	52	6	83.62	GML VQ
Poyuan et al. [58]	2017	ngực, cổ tay	2	4	–	–	–	Ngưỡng (Threshold)
Jeon et al. [59]	2017	cổ tay	2	3	100	–	88.02	CNN
Emer et al. [60]	2020	ngực, bụng	2	5	125	11	83.00	Hồi quy Logistic
Jeng et al. [61]	2021	cổ tay, ngực	2	4	60	–	85.00	SVM, RF
Rawan et al. [62]	2022	cổ	1	4	60	–	99.00	Extra-Tree
Rawan et al. [63]	2023	cổ	1	4	5,10,25, 50	18	98.00	LSTM

Các phương pháp dựa trên thiết bị đeo sử dụng các cảm biến gắn trên cơ thể để nắm bắt chuyển động và hướng nhằm nhận dạng tư thế ngủ. Trong số các hệ thống này, các thiết bị đeo dựa trên gia tốc kế được sử dụng rộng rãi nhất do kích thước nhỏ gọn, chi phí thấp và hiệu quả trong việc giám sát liên tục động lực học của cơ thể. Nhiều nghiên cứu đã nghiên cứu các biến thể trong vị trí đặt cảm biến, số lượng đơn vị cảm biến và các thuật toán được áp dụng để phân loại tư thế, như được tóm tắt trong Bảng 1.1.

Jiang và cộng sự [56] đã đặt các gia tốc kế trên ngực và sử dụng phương pháp dựa trên ngưỡng để phân loại bốn tư thế. Fallmann và cộng sự [57] đã mở rộng điều này bằng cách gắn các cảm biến vào ngực và cổ chân, nhận dạng mười tư thế với độ chính xác 83,62% bằng cách sử dụng Gaussian Mixture Learning Vector Quantization. Tương tự, Poyuan và cộng sự [58] đã áp dụng phương pháp dựa trên ngưỡng với các gia tốc kế trên ngực và cổ tay để phân loại bốn tư thế. Các mô hình tiên tiến hơn đã được giới thiệu trong các công trình sau này. Jeon và cộng sự [59] đã sử dụng các gia tốc kế đeo ở cổ tay với mạng nơ-ron tích chập (CNN), báo cáo độ chính xác 88,02% trên ba tư thế. Doheny và cộng sự [60] đã áp dụng hồi quy logistic với các cảm biến trên ngực và bụng, nhận dạng năm tư thế với độ chính xác 83%. Jeng và cộng sự [61] đã kết hợp các cảm biến cổ tay và ngực với các bộ phân loại SVM và Random Forest,

đạt được độ chính xác 85% cho bốn tư thế.

Các nghiên cứu gần đây đã nghiên cứu các thiết lập đơn cảm biến với các kỹ thuật học sâu. Abdulsadig và cộng sự [62] đã đạt được độ chính xác 99% bằng cách sử dụng một gia tốc kế trên cổ với bộ phân loại Extra-Tree. Tương tự, Abdulsadig và cộng sự [63] đã báo cáo độ chính xác 98% cho bốn tư thế bằng cách sử dụng LSTM với các gia tốc kế gắn trên cổ ở các tần số lấy mẫu khác nhau. các hệ thống đeo trên người dựa trên gia tốc kế đã chứng minh hiệu suất đáng tin cậy trong việc nhận dạng tư thế ngủ. Độ chính xác được báo cáo thay đổi tùy thuộc vào vị trí đặt cảm biến, số lượng cảm biến và sự lựa chọn các mô hình học máy hoặc học sâu. Tuy nhiên, mặc dù hiệu quả, các nghiên cứu hiện tại hầu như chỉ tập trung vào các 4 tư thế cơ bản như nằm ngửa, nằm sấp, nghiêng trái và nghiêng phải.

Như đã thảo luận trong Phần 1.1.2, bằng chứng lâm sàng đã chứng minh rõ ràng nhu cầu theo dõi tư thế chi tiết trong việc đánh giá chứng ngưng thở khi ngủ do tắc nghẽn phụ thuộc tư thế. Công trình trước đây của Tate và các đồng nghiệp [1] nhấn mạnh tầm quan trọng của việc phân loại hướng thân mình ở độ phân giải nhỏ tới $\pm 1.5^\circ$ so với các 4 tư thế truyền thống để nắm bắt các biến đổi có ý nghĩa lâm sàng trong các biến cố hô hấp. Tuy nhiên, chưa có nghiên cứu nào được công bố, đặc biệt là trong các hệ thống dựa trên thiết bị đeo, chứng minh được khả năng phát hiện tư thế ngủ với độ chính xác về góc như vậy. Sự thiếu hụt việc phát hiện tư thế bằng thiết bị đeo độ phân giải cao này nhấn mạnh một khoảng cách trống trong các tài liệu hiện có và thúc đẩy hơn nữa việc phát triển các hệ thống cảm biến có khả năng đáp ứng các nhu cầu về tư thế ngủ chi tiết đã được xác thực về mặt lâm sàng này.

1.3 Tổng quan về các phương pháp học máy cho phân loại tư thế ngủ

1.3.1 Các tiếp cận học máy truyền thống

Các phương pháp học máy từ lâu đã chứng minh khả năng phân loại mạnh mẽ thông qua khả năng học các mẫu trực tiếp từ dữ liệu. Trong bối cảnh phân loại tư thế ngủ, việc sử dụng các tín hiệu gia tốc làm đầu vào cho các mô hình học máy đã trở nên phổ biến trong nhiều nghiên cứu. Vì dữ liệu gia tốc vốn là các tín hiệu chuỗi thời gian, các bước tiền xử lý, đặc biệt là phân đoạn thành các cửa sổ trượt đóng vai trò quan trọng trong việc chuyển đổi các chuỗi dữ liệu cảm biến liên tục thành các mẫu dữ liệu có nhiều ý nghĩa. Sau khi phân đoạn, các đặc trưng liên quan được trích xuất từ

mỗi cửa sổ và sau đó được đưa vào các bộ phân loại học máy. Trong một nghiên cứu sơ bộ (CT4) được thực hiện vào năm 2023, nhóm nghiên cứu của chúng tôi đã đánh giá một số thuật toán học máy bằng cách sử dụng dữ liệu gia tốc kế được thu thập từ một cảm biến đặt trên bụng. Tập dữ liệu bao gồm bốn tư thế ngủ cơ bản cùng với một tư thế ngồi bổ sung. Trong số các mô hình được kiểm tra, Extra Trees và Hồi quy Logistic đã đạt được hiệu suất đặc biệt cao, với điểm chính xác lần lượt là 99,24% và 98,83%. Những phát hiện này xác nhận tính hiệu quả đáng kể của các phương pháp học máy truyền thống đối với việc phân loại tư thế ngủ cơ bản, và nhất quán với các kết quả được báo cáo trong các nghiên cứu trước đây sử dụng các phương thức cảm biến và kỹ thuật học máy dựa trên trích xuất đặc trưng tương tự [40,62,63].

Các phương pháp học máy truyền thống đã cho thấy hiệu suất đáng tin cậy trong việc phân loại tư thế và hoạt động ở mức độ 4 tư thế cơ bản, nơi các ranh giới lớp là rõ rệt và các đặc trưng thủ công có tính phân biệt có thể được thiết kế hiệu quả. Tuy nhiên, có hai hạn chế lớn của các mô hình học máy dựa trên đặc trưng trở nên rõ ràng khi chuyển sang phân loại tư thế ngủ với độ phân giải cao hơn [64].

Thứ nhất, các đặc trưng thủ công vốn dĩ hạn chế khả năng của các mô hình học máy trong việc nắm bắt các biến đổi tinh vi, phi tuyến tính và có cấu trúc theo thời gian trong các dữ liệu tín hiệu [65,66]. Các nghiên cứu về nhận dạng hành động đã kết luận rằng các đặc trưng được thiết kế thủ công gặp khó khăn với các phụ thuộc thời gian–không gian phức tạp và sự khác biệt chuyển động chi tiết [67,68], đặc biệt là khi các biến đổi tín hiệu nhỏ so với nhiễu hoặc các chuyển động do hô hấp gây ra [69,70]. Những ràng buộc này có hệ quả lớn đối với phân tích tư thế ngủ, nơi các thay đổi hướng có ý nghĩa lâm sàng chỉ $\pm 15^\circ$ tạo ra những thay đổi gia tốc kế nhỏ mà các đặc trưng thủ công không được thiết kế rõ ràng để mã hóa [71,72]. Một nghiên cứu gần đây của Fonseca và cộng sự [64] đã kiểm tra hiệu suất của các thuật toán học máy trong một nhiệm vụ phân loại tư thế trên giường với độ chi tiết tư thế cao bất thường (28 tư thế riêng biệt) và đánh giá tác động của độ phân giải đầu vào đối với độ chính xác. Sử dụng một tập dữ liệu lớn gồm 60 đối tượng và một tấm cảm biến bản đồ áp suất 27×64 , họ báo cáo rằng các bộ phân loại học máy phổ biến (k-NN, SVM, cây quyết định, rừng ngẫu nhiên, MLP) đạt được độ chính xác rất cao (98–99%) khi chỉ phân loại bốn tư thế kinh điển (nằm ngửa, nằm sấp, nghiêng trái, nghiêng phải). Tuy nhiên, độ chính xác đã giảm mạnh xuống còn 59–68% khi số lượng lớp mở rộng lên 28 biến thể chi tiết. Các tác giả còn quan sát thấy rằng việc giảm độ phân giải cảm

biến (tức là ít số lượng cảm biến hơn) chỉ ảnh hưởng nhỏ đến hiệu suất cho nhiệm vụ phân loại 4 lớp, nhưng làm giảm đáng kể độ chính xác trong kịch bản 28 lớp.

Nghiên cứu trên đã cung cấp một minh chứng thực nghiệm rõ ràng rằng, ngay cả khi sử dụng bản đồ áp suất có độ trung thực tương đối cao và các thuật toán ML đã được thiết lập tốt, việc phân loại các tư thế chi tiết vẫn đặt ra những thách thức nghiêm trọng cho các phương pháp học máy truyền thống. Nói cách khác, sự sụt giảm độ chính xác được minh họa trong nghiên cứu của Fonseca và cộng sự làm nổi bật cụ thể hạn chế của ML khi đối mặt với phân loại tư thế có độ chi tiết cao, từ đó củng cố lập luận rằng các phương pháp có khả năng xử lý các phân biệt tinh vi, độ phân giải cao chẳng hạn như học sâu là cần thiết cho việc phân loại tư thế ngũ chi tiết có liên quan về mặt lâm sàng.

Thứ hai, các mô hình học máy dựa trên việc trích xuất đặc trưng có khả năng tổng quát hóa hạn chế giữa các đối tượng, sự khác nhau về mặt hình thái cơ thể và vị trí đặt cảm biến. Bởi vì các đặc trưng thủ công phụ thuộc nhiều vào các giả định cụ thể của miền, hiệu suất của chúng bị suy giảm khi đối mặt với sự thay đổi giữa các đối tượng, điều kiện thí nghiệm hoặc những khác biệt nhỏ trong vị trí đặt cảm biến, là những yếu tố không thể tránh khỏi trong giám sát giấc ngủ thực tế [73,74].

Nghiên cứu của Bento và cộng sự [75] nghiên cứu khả năng tổng quát hóa của học máy dựa trên đặc trưng thủ công so với các biểu diễn mạng học sâu trong nhận dạng hoạt động người (HAR) bằng thiết bị đeo. Kết quả của họ cho thấy rằng mặc dù các đặc trưng tốt cho các nhiệm vụ trong miền (tức là khi các đối tượng huấn luyện và kiểm tra, vị trí thiết bị và định nghĩa hoạt động khớp nhau chặt chẽ), hiệu suất của chúng giảm đáng kể khi chuyển miền chẳng hạn như người dùng mới, vị trí đặt cảm biến chưa từng thấy hoặc các biến đổi tinh vi trong việc thực hiện hoạt động. Ngược lại, các mạng học sâu đã chứng minh tính mạnh mẽ vượt trội khi thay đổi miền, đạt được hiệu suất ổn định hơn trong các điều kiện không đồng nhất. Mặc dù nghiên cứu này không nhắm mục tiêu rõ ràng vào phân loại tư thế ngũ, nhưng hiện tượng này cơ bản là có liên quan trực tiếp: trong bối cảnh ước tính tư thế ngũ chi tiết, các thay đổi hướng cơ thể có thể là tối thiểu và vị trí đặt cảm biến hoặc giải phẫu người dùng có thể thay đổi. Khi đó, các quy trình trích xuất đặc trưng thủ công sẽ gặp khó khăn với việc tổng quát hóa miền. Do đó, những phát hiện của Bento và cộng sự củng cố quyết định theo đuổi học sâu như là cách tiếp cận mô hình hóa nền tảng trong luận án này, đặc biệt là với mục tiêu phát hiện các độ lệch góc tinh vi trong bối cảnh thiết bị đeo.

Tổng kết lại chúng ta thấy, các tác động lâm sàng như đã thảo luận trong Phần 1.1.2, các nghiên cứu như Tate và cộng sự [1] cho thấy hướng thân mình trong khoảng $\pm 15^\circ$ so với các tư thế kinh điển ảnh hưởng đáng kể đến chỉ số ngưng thở-giảm thở. Việc đáp ứng yêu cầu lâm sàng này đòi hỏi các mô hình có khả năng giải quyết các biến đổi tư thế tinh vi, chính xác mà các mô hình học máy kết hợp trích xuất đặc trưng thủ công xử lý kém. Học máy truyền thống tốt cho việc nhận dạng tư thế ngủ ở mức độ cơ bản, nhưng tính phù hợp của nó đối với việc ước tính góc chi tiết theo yêu cầu lâm sàng vẫn chưa được chứng minh. Ngược lại, học sâu cung cấp khả năng biểu diễn cần thiết để mô hình hóa các biến đổi thời gian-không gian tinh vi trong các tín hiệu gia tốc. Vì những lý do này, luận án này áp dụng học sâu làm khung phương pháp luận chính để phát triển các mô hình phân loại tư thế ngủ chi tiết dựa trên thiết bị đeo.

1.3.2 Học sâu cho dữ liệu cảm biến chuỗi thời gian

Học sâu cung cấp một cách tiếp cận để mô hình hóa dữ liệu cảm biến chuỗi thời gian bằng cách học nhiều cấp độ biểu diễn trực tiếp từ các đầu vào thô. Thay vì dựa vào các đặc trưng thủ công, các mạng học sâu kết hợp các phép biến đổi phi tuyến tính liên tiếp để trích xuất các đặc trưng ngày càng trừu tượng và có tính phân biệt từ dữ liệu [76]. Trong bối cảnh chuỗi thời gian, điều này có nghĩa là các hoạt động cấp thấp (ví dụ: các sai phân đơn giản hoặc các xu hướng cục bộ) có thể được kết hợp qua các lớp thành các mẫu cấp cao hơn mà không cần xác định rõ ràng các cấu trúc này trước đó. Các bài tổng quan về học sâu cho phân loại chuỗi thời gian nhấn mạnh việc học biểu diễn đầu-cuối (end-to-end) này là một sự chuyển hướng quan trọng so với các quy trình kỹ thuật đặc trưng truyền thống [77].

Một lớp kiến trúc quan trọng cho chuỗi thời gian là mạng nơ-ron tích chập một chiều (1D CNNs). Bằng cách trượt các bộ lọc có thể học được trên trục thời gian, CNN khai thác các tương quan thời gian cục bộ và tính dừng trong tín hiệu: mỗi bộ lọc học cách phát hiện một mẫu thời gian ngắn (chẳng hạn như một hình dạng dạng sóng cụ thể hoặc thay đổi cục bộ) độc lập với vị trí tuyệt đối của nó trong chuỗi [2]. Chia sẻ trọng số theo thời gian làm cho các bộ lọc này hiệu quả và thực thi một cách tự nhiên một dạng bất biến tịnh tiến, sao cho cùng một mẫu có thể được nhận dạng bất kể nó xảy ra ở đâu [78]. Các nghiên cứu thực nghiệm [79] về phân loại chuỗi thời gian liên tục báo cáo CNN là một trong những kiến trúc hiệu quả nhất, một phần vì chúng học

các bộ lọc bất biến theo thời gian trực tiếp từ dữ liệu chuỗi thời gian thô và tránh được nhu cầu thiết kế thủ công các đặc trưng miền thời gian hoặc miền tần số [77].

Mạng nơ-ron hồi quy (RNNs), và đặc biệt là các mô hình LSTM và GRU, giải quyết một khía cạnh trung tâm khác của dữ liệu chuỗi thời gian: sự phụ thuộc thời gian [80]. Bằng cách duy trì một trạng thái ẩn được cập nhật tại mỗi bước thời gian, các mô hình này có thể tích hợp thông tin qua các khoảng thời gian dài và nắm bắt các phụ thuộc dạng chuỗi vượt ra ngoài khả năng xử lý dữ liệu cục bộ của một phép tích chập [81]. Các cơ chế cổng trong LSTM và GRU làm giảm bớt các vấn đề biến mất gradient, cho phép mạng học cả các mối quan hệ thời gian ngắn hạn và dài hạn [82]. Trong lĩnh vực cảm biến đeo trên người, Ordóñez và Roggen [83] đã chỉ ra rằng một kiến trúc kết hợp CNN-LSTM có thể đồng thời học các đặc trưng cục bộ và mô hình hóa một cách rõ ràng động lực học thời gian của các kích hoạt đặc trưng, đồng thời vượt qua các mô hình sâu không hồi quy và loại bỏ nhu cầu về các đặc trưng do chuyên gia thiết kế. Hammerla và cộng sự [84] cũng báo cáo rằng các mô hình sâu hồi quy đạt được hiệu suất vượt trội so với các cách tiếp cận nông hơn trên nhiều tập dữ liệu chuyển động được thu thập từ các thiết bị đeo.

Ngoài các kiến trúc riêng lẻ, các đánh giá thực nghiệm quy mô lớn hỗ trợ hiệu quả của học sâu cho các nhiệm vụ chuỗi thời gian một cách rộng rãi hơn. Fawaz và cộng sự [77] đã thực hiện một thử nghiệm so sánh đa dạng các mạng học sâu cho phân loại chuỗi thời gian trên 97 tập dữ liệu và chỉ ra rằng các mô hình sâu thường khớp hoặc vượt quá hiệu suất của các phương pháp học máy tiên tiến nhất, trong khi hoạt động theo cách thức đầu-cuối mà không cần kỹ thuật đặc trưng thủ công. Các khảo sát gần đây hơn về học sâu cho dự báo và phân loại chuỗi thời gian cũng kết luận rằng các kiến trúc sâu đặc biệt phù hợp để nắm bắt động lực học phi tuyến tính, các mẫu thời gian đa quy mô và các phụ thuộc phức tạp trong dữ liệu chuỗi thời gian đa chiều.

1.3.3 Mạng nơ-ron tích chập cho tín hiệu cảm biến

Để khắc phục những hạn chế của các kỹ thuật đặc trưng truyền thống, Kiranyaz và cộng sự [85] đã giới thiệu các mạng 1D CNN nhỏ gọn và thích ứng đầu tiên vào năm 2015 để phân loại ECG đặc thù cho từng bệnh nhân. Kể từ đó, 1D CNN đã được áp dụng thành công trong nhiều lĩnh vực khác nhau, bao gồm phát hiện rối loạn nhịp tim ECG [85-87], phân tích rung động cho các hệ thống cơ khí [88], và chẩn đoán lỗi trong các máy điện.

Các nghiên cứu gần đây chỉ ra rằng các 1D CNN nhỏ gọn hoạt động đặc biệt tốt trong các kịch bản có dữ liệu được gán nhãn khan hiếm và độ biến thiên cao trong các nguồn tín hiệu, chẳng hạn như tín hiệu cảm biến đeo trên người, dữ liệu phát xạ âm thanh trong các cấu trúc dân dụng, các mẫu rung động động cơ hoặc đo xa từ thiết bị công nghiệp. Hình 1.2 [2] minh họa rằng một 1D CNN thường chứa hai loại lớp chính:

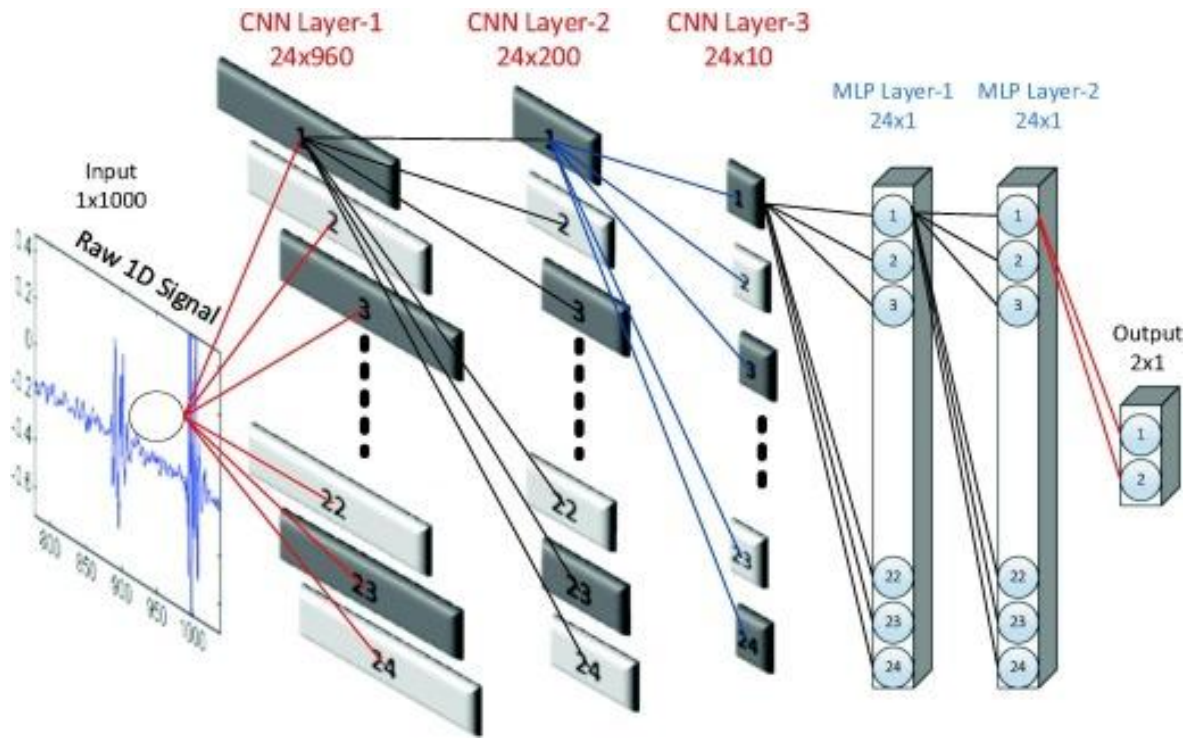
1. **Các lớp CNN**, kết hợp tích chập 1D, các hàm kích hoạt phi tuyến tính và các hoạt động gộp (pooling)/lấy mẫu phụ (sub-sampling).
2. **Các lớp kết nối đầy đủ (MLP)**, tương đương về mặt chức năng với các Perceptron đa lớp tiêu chuẩn, do đó được gọi là các lớp MLP.

Việc thiết kế một 1D CNN phụ thuộc vào một số siêu tham số [2]:

1. Số lượng các lớp CNN và MLP ẩn và các nơ-ron của chúng (ví dụ: trong Hình 1.2, có ba lớp CNN và hai lớp MLP).
2. Kích thước hạt nhân (kernel size) trong mỗi lớp CNN (ví dụ: 41 trong tất cả các lớp CNN trong Hình 1.2).
3. Hệ số lấy mẫu phụ trong các lớp CNN (ví dụ: hệ số 4 trong Hình 1.2).
4. Các lựa chọn hàm gộp và hàm kích hoạt.

Tương tự như 2D CNN, lớp đầu vào nhận tín hiệu 1D thô, trong khi lớp đầu ra là một MLP với các nơ-ron tương ứng với số lượng lớp mục tiêu. Hình 1.2 cho thấy ba lớp CNN liên tiếp, trong đó các bộ lọc có kích thước 3 và hệ số lấy mẫu phụ là 2. Mỗi nơ-ron trong một lớp ẩn l thực hiện các phép tích chập, chuyển kết quả qua một hàm kích hoạt, và sau đó trải qua quá trình lấy mẫu phụ.

Không giống như 2D CNN, 1D CNN hoạt động trên các mảng một chiều cho cả hạt nhân và bản đồ đặc trưng. Các lớp CNN tự động trích xuất các đặc trưng từ các tín hiệu thô, sau đó được phân loại bởi các lớp MLP. Sự tích hợp trích xuất đặc trưng và phân loại vào một khung có thể huấn luyện duy nhất cho phép tối ưu hóa đầu-cuối và độ chính xác cao. Hơn nữa, các phép tích chập 1D là các tổng trọng số tuyến tính của các mảng 1D, làm cho chúng hiệu quả về mặt tính toán và có khả năng song song hóa cao trong quá trình huấn luyện và suy luận.



Hình 1.2: Một cấu hình 1D CNN mẫu với 3 lớp CNN và 2 lớp MLP. [2]

1.3.4 Mạng nơ-ron hồi quy cho mô hình hóa chuỗi thời gian

Mạng bộ nhớ dài hạn (LSTM) là một sự lựa chọn mạnh mẽ cho chuỗi thời gian cảm biến, đặc biệt là các luồng gia tốc kế và IMU được sử dụng trong nhận dạng hoạt động người (HAR), bởi vì chúng nắm bắt được các phụ thuộc thời gian tầm xa và động lực học liên tục mà không yêu cầu các đặc trưng thủ công. Một nghiên cứu HAR có tính khởi đầu của Ordóñez và Roggen đã chỉ ra rằng việc thêm tính hồi quy LSTM vào các quy trình cảm biến đeo trên người giúp cải thiện đáng kể việc nhận dạng các hoạt động phức tạp từ dữ liệu gia tốc kế và con quay hồi chuyển, thiết lập LSTM như một mô hình tuần tự cốt lõi cho cảm biến chuyển động [83]. Công trình tiếp theo tập trung cụ thể vào dữ liệu gia tốc kế điện thoại thông minh đã đánh giá các biến thể LSTM cơ sở và báo cáo rằng các mô hình LSTM có tính cạnh tranh và thường vượt trội trong số các họ RNN cho HAR dựa trên cảm biến, đồng thời chi tiết hóa các lựa chọn thiết kế thực tế như phân cửa sổ và hợp nhất đa kênh [89]. Các khảo sát gần đây về học sâu cho HAR bằng thiết bị đeo đã xác nhận những phát hiện này ở quy mô lớn, làm nổi bật tính mạnh mẽ của LSTM đối với các tín hiệu chuyển động nhiễu, không dừng và khả năng tổng quát hóa của nó giữa các người dùng và vị trí thiết bị khi được hiệu chỉnh (regularized) đúng cách [90].

Ngoài các cảm biến chuyển động, LSTM cũng đã chứng minh hiệu quả trên các

chuỗi thời gian y sinh. Đối với ECG, Saadatnejad và cộng sự đã xây dựng một quy trình đa LSTM nhẹ (kết hợp với các đặc trưng wavelet) đạt được phân loại rối loạn nhịp tim theo thời gian thực chính xác trên phần cứng đeo trên người, bằng chứng cho thấy LSTM có thể đáp ứng các ràng buộc về độ trễ và tài nguyên trong cảm biến nhúng [91]. Đối với EEG, Ozal Yildirim và cộng sự đã chứng minh rằng việc phân loại giai đoạn giấc ngủ đơn kênh được hưởng lợi từ khả năng mô hình hóa cấu trúc thời gian tầm xa của LSTM, trong khi các nghiên cứu khác đã sử dụng các mô hình LSTM xếp chồng hoặc hai chiều để phát hiện động kinh, báo cáo độ chính xác mạnh mẽ ngay cả với các kênh hạn chế [92].

Nhìn chung, các tài liệu hỗ trợ LSTM như một cơ sở có nguyên tắc cho HAR gia tốc kế/IMU và là một lựa chọn thực tế cho phân tích ECG/EEG, với hiệu suất và khả năng triển khai được định hình bởi các yếu tố như độ dài cửa sổ, tần số lấy mẫu và các lựa chọn kiến trúc nhẹ (ví dụ: các chồng nông, các biến thể có cổng, hoặc nén sau huấn luyện) phù hợp với thiết bị mục tiêu.

Trong khi CNN xuất sắc trong việc nắm bắt các phụ thuộc thời gian cục bộ, mạng nơ-ron hồi quy (RNN) hiệu quả hơn trong việc mô hình hóa các mẫu tuần tự dài hạn trong dữ liệu chuỗi thời gian. Trong số đó, LSTM đã được áp dụng rộng rãi cho phân loại dựa trên cảm biến do khả năng khắc phục vấn đề triệt tiêu gradient. Các nghiên cứu đã chứng minh rằng các kiến trúc dựa trên LSTM có thể đạt được hiệu suất mạnh mẽ trong nhận dạng hoạt động, giám sát tín hiệu sinh lý và phát hiện tư thế ngủ, đặc biệt là khi dữ liệu có tính tuần tự và phụ thuộc vào ngữ cảnh.

Các đơn vị hồi quy có cổng (GRUs) đã được áp dụng như một giải pháp thay thế cho LSTM trong mô hình hóa dữ liệu cảm biến tuần tự do cơ chế cổng đơn giản hơn và giảm bớt tham số hóa. Điều này thường dẫn đến sự hội tụ nhanh hơn và tiêu thụ năng lượng thấp hơn, điều quan trọng đối với các hệ thống cảm biến di động và đeo trên người. Một số nghiên cứu đã khám phá việc sử dụng các kiến trúc dựa trên GRU như là mô hình tuần tự chính mà không dựa vào các thành phần CNN hoặc LSTM bổ sung.

Đối với dữ liệu gia tốc kế và IMU, Mohsen [93] đã nghiên cứu một khung chỉ sử dụng GRU cho nhận dạng hoạt động người (HAR) trên tập dữ liệu WISDM. Nghiên cứu đã chứng minh rằng các kiến trúc GRU thuần túy có thể đạt được độ chính xác phân loại mạnh mẽ trong khi vẫn duy trì hiệu quả tính toán, khiến chúng phù hợp cho HAR dựa trên điện thoại thông minh. Tương tự, Zhong và cộng sự [94] đã giới thiệu

một mạng GRU sâu để phát hiện căng thẳng tâm lý bằng tín hiệu ECG, nơi GRU đã mô hình hóa hiệu quả các đặc trưng biến thiên nhịp tim theo thời gian và cung cấp các kết quả phân loại căng thẳng mạnh mẽ. Trong một nghiên cứu ECG khác, Ramaraj và cộng sự [95] đã đề xuất một quy trình dựa trên GRU kết hợp với bộ phân loại học máy cực hạn (ELM). Ở đây, GRU đóng vai trò là bộ trích xuất đặc trưng thời gian chính, làm nổi bật đặc trưng của chúng trong các nhiệm vụ phát hiện rối loạn nhịp tim với độ phức tạp tính toán giảm so với LSTM.

Hơn nữa, Rivas và cộng sự [96] đã thực hiện một so sánh thực nghiệm trực tiếp giữa các mô hình LSTM và GRU để dự đoán trạng thái chú ý và thiền định bằng dữ liệu EEG thô từ các tai nghe giá rẻ. Nghiên cứu của họ tiết lộ rằng trong khi LSTM đạt được RMSE thấp hơn một chút trong dự đoán thiền định, GRU lại vượt trội trong dự đoán sự chú ý và cung cấp thời gian suy luận nhanh hơn (dưới 50 mili giây), khẳng định tính thực tế của GRU cho các ứng dụng giao diện não-máy tính (BCI) thời gian thực.

1.3.5 Kiến trúc Lai CNN-RNN cho dữ liệu cảm biến

Trong nhận dạng hoạt động người (HAR) bằng thiết bị đeo và giám sát tư thế ngủ, việc phân loại các tín hiệu quán tính đòi hỏi sự mô hình hóa kết hợp các mẫu tín hiệu cục bộ và các phụ thuộc thời gian dài hạn [67,90]. Các kiến trúc lai mạng nơ-ron tích chập (CNN) với mạng nơ-ron hồi quy (RNN) đã nổi lên như một giải pháp đáng tin cậy cho vấn đề này [97-99]. Phần này thảo luận về cơ sở khái niệm của các mô hình lai CNN-RNN trong bối cảnh phân loại tư thế dựa trên gia tốc kế.

Trích xuất đặc trưng qua 1D-CNN. Gia tốc kế tạo ra dữ liệu chuỗi thời gian đa trục, trong đó các mẫu phạm vi ngắn phản ánh hướng cơ thể tức thời hoặc các chuyển đổi [100]. Mạng nơ-ron tích chập một chiều (1D-CNN) thường được sử dụng làm giai đoạn xử lý đầu tiên để tự động trích xuất các biểu diễn cục bộ này mà không cần kỹ thuật đặc trưng thủ công [101]. Bằng cách tích chập trên chiều thời gian, 1D-CNN nắm bắt các biểu diễn thời gian-không gian như các gợi ý hướng trọng lực và các thay đổi đột ngột liên quan đến sự dịch chuyển tư thế. 1D CNN tiếp tục thúc đẩy tính bất biến tịnh tiến và giảm thiểu nhiễu cảm biến tần số cao, mang lại một biểu diễn tiềm ẩn nhỏ gọn và có khả năng chống nhiễu [102].

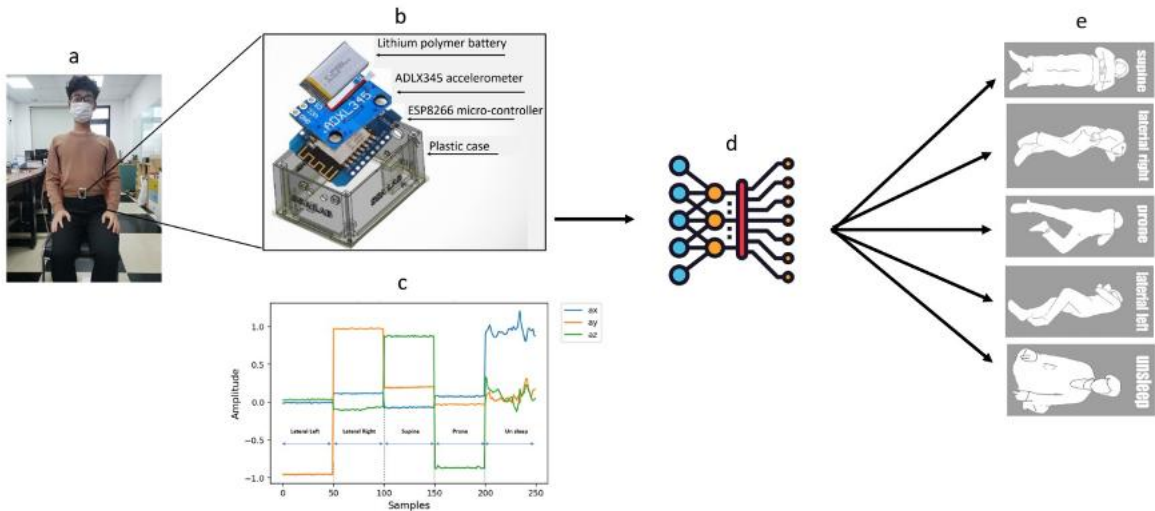
Mô hình hóa chuỗi thời gian qua RNN. Mặc dù mạng nơ-ron tích chập nắm bắt hiệu quả các đặc trưng cục bộ, tuy nhiên chúng không có khả năng ghi nhớ dài

hạn [103]. Để giải quyết vấn đề này, các đặc trưng do CNN tạo ra được cung cấp làm đầu vào cho mạng nơ-ron hồi quy, đặc biệt là các mô hình LSTM hoặc GRU [97, 104, 105]. Trong phân loại tư thế ngủ, tính chuỗi theo thời gian là quan trọng để phân biệt các dịch chuyển nhất thời với các khoảng thời gian tư thế ổn định [106]. Các lớp RNN tổng hợp thông tin thời gian để học các tư thế, trong khi các biến thể hai chiều (ví dụ: BiLSTM) có thể kết hợp cả ngữ cảnh quá khứ và tương lai [107, 108].

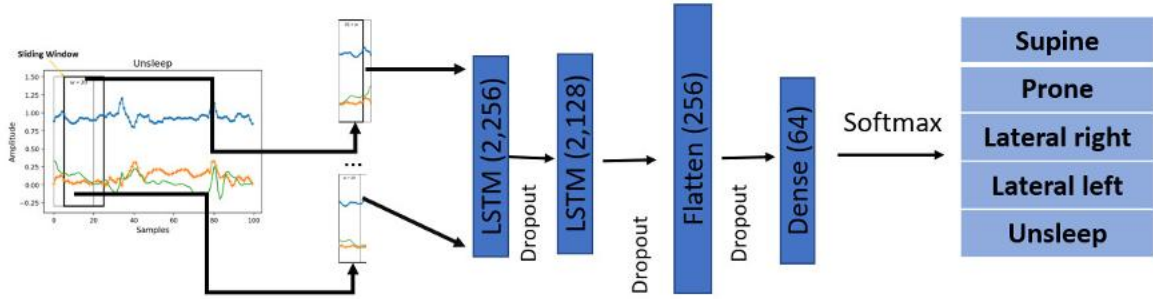
Sự liên quan đến phân loại tư thế chi tiết. Do khả năng mô hình hóa kết hợp các gợi ý hướng cục bộ và hành vi thời gian dài hạn, các kiến trúc lai CNN-RNN hình thành một nền tảng có nguyên tắc cho việc phân loại tư thế ngủ chi tiết [109]. Do đó, một số kiến trúc được phát triển trong các nghiên cứu gần đây—bao gồm cả những kiến trúc được nghiên cứu trong các chương sau của luận án này được xây dựng dựa trên nền tảng thiết kế lai này.

1.4 Thử nghiệm sơ bộ: xác thực tính khả thi

Trước khi đi vào các mô hình học sâu được đề xuất trong luận án, một số nghiên cứu sơ bộ đã được thực hiện nhằm kiểm chứng tính khả thi của hướng tiếp cận sử dụng cảm biến gia tốc đeo trên người cho bài toán phân loại tư thế ngủ. Các nghiên cứu này không nhằm mục tiêu giải quyết đầy đủ bài toán tổng quát hóa trong điều kiện triển khai thực tế, mà chủ yếu đóng vai trò kiểm chứng ban đầu về khả năng thu nhận tín hiệu, khả năng phân biệt tư thế từ dữ liệu gia tốc và tiềm năng triển khai mô hình trên các thiết bị có tài nguyên hạn chế.



Hình 1.3: Hệ thống phân loại tư thế ngủ



Hình 1.4: Kiến trúc của mô hình phân loại LSTM tuần tự

Bảng 1.2: So sánh với các mô hình cơ sở và học sâu cho nhiệm vụ phân loại 12 tư thế.

Mô hình	Độ chính xác (%)	AUC	Số tham số	Thời gian (s)
MLP	98.86	0.9993	9,093	
CNN	99.22	0.9995	25,669	144
LSTM_con	99.28	0.9996	471,941	78
LSTM_sq	99.60	0.9999	627,589	340
Abdulsadig [62]	99.24	0.9921	955,653	14.1
Doheny [60]	98.83	0.9924	15	21

Trong nghiên cứu sơ bộ đầu tiên, được báo cáo trong CT4, một nhóm nhỏ gồm bốn người tham gia đã được thu thập dữ liệu ở bốn tư thế ngủ cơ bản: nằm ngửa, nằm sấp, nghiêng trái và nghiêng phải. Mục tiêu của thử nghiệm này không phải là tối ưu hóa độ chính xác phân loại, mà là đánh giá xem cấu hình cảm biến gia tốc và các đặc điểm theo thời gian của tín hiệu ba trục có chứa đủ thông tin để phân biệt các tư thế ngủ cơ bản hay không.

Trong CT4, mô hình LSTM_sq với 2 tầng LSTM nối tiếp nhau được sử dụng để mô hình hóa đặc trưng chuỗi thời gian của tín hiệu gia tốc ba trục, như minh họa trong Hình 1.3 và Hình 1.4

Bảng 1.2 cho thấy các mô hình học sâu có xu hướng đạt hiệu quả cao hơn so với một số mô hình cơ sở trong bài toán phân loại 12 tư thế ngủ. Kết quả này củng cố tính khả thi của hướng tiếp cận sử dụng học sâu cho dữ liệu gia tốc chuỗi thời gian, đồng thời tạo cơ sở cho việc phát triển các kiến trúc chuyên biệt hơn ở các chương tiếp theo. Mặc dù tập dữ liệu còn nhỏ và số lượng tư thế còn hạn chế, kết quả thu được cho thấy tín hiệu gia tốc có khả năng phản ánh sự thay đổi về độ nghiêng và hướng cơ thể thông qua các thành phần trọng lực trong dữ liệu cảm biến. Điều này xác nhận rằng cảm biến gia tốc đeo trên người là một lựa chọn khả thi cho bài toán giám sát tư thế ngủ.

Tiếp theo, CT5 được thực hiện như một nghiên cứu phụ trợ nhằm kiểm chứng khả năng mở rộng hệ thống sang bài toán phân loại tư thế ngủ chi tiết hơn và đánh

giá tính khả thi của việc triển khai trên thiết bị biên. Khác với CT4, CT5 xem xét số lượng tư thế lớn hơn và sử dụng mô hình học máy truyền thống được tối ưu hóa, cụ thể là Random Forest kết hợp giải thuật di truyền. Mục tiêu chính của CT5 là trả lời câu hỏi liệu khi số lớp tư thế được mở rộng, hệ thống phần cứng và mô hình học máy có vẫn đủ khả năng xử lý bài toán theo thời gian thực trên các thiết bị có hiệu năng hạn chế hay không.

Kết quả từ CT5 cho thấy các mô hình học máy truyền thống, khi được tối ưu phù hợp, có thể xử lý hiệu quả bài toán phân loại tư thế ngủ mở rộng và có tiềm năng triển khai trong các hệ thống nhúng hoặc thiết bị biên. Điều này cung cấp bằng chứng quan trọng rằng bài toán phân loại tư thế ngủ chi tiết không chỉ khả thi về mặt thuật toán, mà còn khả thi về mặt triển khai hệ thống. Nói cách khác, CT5 cho thấy thiết bị và mô hình học máy có thể đáp ứng yêu cầu suy luận nhanh trong bối cảnh tài nguyên tính toán bị giới hạn.

Tuy nhiên, cần nhấn mạnh rằng CT4 và CT5 đều sử dụng các thiết lập đánh giá mang tính tiền khả thi, trong đó dữ liệu có thể được chia theo hướng trộn giữa các đối tượng trước khi tách thành tập huấn luyện và tập kiểm tra. Do đó, các kết quả này chưa nhằm trả lời câu hỏi về khả năng tổng quát hóa của mô hình đối với người dùng mới, tức là những người chưa từng xuất hiện trong quá trình huấn luyện. Đây là một yêu cầu quan trọng hơn trong bối cảnh triển khai thực tế, nơi hệ thống cần hoạt động ổn định trên các cá nhân có đặc điểm hình thái cơ thể, thói quen nằm ngủ và cách đeo cảm biến khác nhau.

Từ các kết quả sơ bộ trên, có thể rút ra hai nhận định quan trọng. Thứ nhất, tín hiệu gia tốc từ một cảm biến đeo trên người có đủ thông tin để phục vụ bài toán phân loại tư thế ngủ. Thứ hai, khi mở rộng số lượng tư thế, các mô hình học máy vẫn có khả năng xử lý bài toán với chi phí tính toán phù hợp cho thiết bị biên. Tuy nhiên, câu hỏi về khả năng khái quát hóa trong kịch bản độc lập theo đối tượng vẫn còn để mở.

Chính vì vậy, luận án tiếp tục chuyển trọng tâm sang bài toán khó hơn và gần với thực tế triển khai hơn: phân loại 12 tư thế ngủ trong thiết lập đánh giá độc lập theo đối tượng. Theo hướng này, các nghiên cứu tiếp theo, bao gồm CT3 và hai mô hình đề xuất chính trong luận án là AnpoNet và CGB-Net, tập trung vào việc đánh giá khả năng tổng quát hóa của các mô hình học sâu trên những người dùng chưa từng xuất hiện trong quá trình huấn luyện. Các chi tiết về bộ dữ liệu 12 tư thế, giao thức

đánh giá và các mô hình đề xuất được trình bày trong các chương tiếp theo.

1.5 Các chỉ số đánh giá

Để đánh giá hiệu suất của các mô hình phân loại tự thể ngữ, luận án sử dụng các độ đo phổ biến trong bài toán phân loại gồm: Accuracy, Precision, Recall, F1-score và Confusion Matrix. Các độ đo này được tính toán dựa trên kết quả dự đoán của mô hình so với nhãn thực tế của tập dữ liệu.

1.5.1 Confusion Matrix (ma trận nhầm lẫn)

Bên cạnh các độ đo tổng hợp như Accuracy và F1-score, **Confusion Matrix** (ma trận nhầm lẫn) được sử dụng để phân tích chi tiết hành vi của mô hình phân loại.

Đối với bài toán phân loại nhị phân, Confusion Matrix bao gồm bốn thành phần cơ bản:

- **TP (True Positive)**: số mẫu thuộc lớp dương tính và được dự đoán đúng.
- **FP (False Positive)**: số mẫu không thuộc lớp dương tính nhưng bị dự đoán sai là dương tính.
- **TN (True Negative)**: số mẫu không thuộc lớp dương tính và được dự đoán đúng.
- **FN (False Negative)**: số mẫu thuộc lớp dương tính nhưng bị dự đoán sai thành lớp khác.

Tổng số mẫu trong tập dữ liệu được tính như sau:

$$N = TP + TN + FP + FN \quad (1.1)$$

Trong bài toán phân loại đa lớp với K lớp, Confusion Matrix được biểu diễn dưới dạng ma trận kích thước $K \times K$:

$$C = \begin{bmatrix} c_{11} & c_{12} & \dots & c_{1K} \\ c_{21} & c_{22} & \dots & c_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ c_{K1} & c_{K2} & \dots & c_{KK} \end{bmatrix} \quad (1.2)$$

Trong đó c_{ij} biểu thị số lượng mẫu có nhãn thực là lớp i nhưng được mô hình dự đoán là lớp j .

Các phần tử trên đường chéo chính (c_{ii}) biểu thị các dự đoán đúng, trong khi các phần tử ngoài đường chéo biểu thị các trường hợp phân loại sai.

Confusion Matrix cung cấp cái nhìn chi tiết về cách mô hình phân loại từng lớp, cho phép xác định các lớp dễ bị nhầm lẫn với nhau. Điều này đặc biệt hữu ích trong bài toán phân loại tư thế ngủ, nơi một số tư thế có thể có đặc trưng tương đồng và dễ bị mô hình nhầm lẫn.

1.5.2 Accuracy

Accuracy (độ chính xác) đo lường tỷ lệ các mẫu được dự đoán đúng trên tổng số mẫu trong tập dữ liệu.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1.3)$$

Trong đó:

- TP (True Positive): số mẫu thuộc lớp dương tính được dự đoán đúng.
- TN (True Negative): số mẫu thuộc lớp âm tính được dự đoán đúng.
- FP (False Positive): số mẫu không thuộc lớp dương tính nhưng bị dự đoán sai là dương tính.
- FN (False Negative): số mẫu thuộc lớp dương tính nhưng bị dự đoán sai thành lớp khác.

Giá trị của Accuracy nằm trong khoảng:

$$0 \leq Accuracy \leq 1$$

hoặc có thể biểu diễn dưới dạng phần trăm từ 0% đến 100%.

Accuracy càng cao cho thấy mô hình phân loại càng chính xác trên toàn bộ tập dữ liệu. Chỉ số này đặc biệt phù hợp khi dữ liệu được phân bố tương đối cân bằng giữa các lớp.

Tuy nhiên, Accuracy có thể gây hiểu nhầm khi dữ liệu bị mất cân bằng giữa các lớp. Trong trường hợp đó, mô hình có thể đạt Accuracy cao chỉ bằng cách dự đoán phần lớn mẫu thuộc lớp chiếm ưu thế, trong khi vẫn dự đoán kém các lớp còn lại.

1.5.3 Precision

Precision đo lường mức độ chính xác của các dự đoán dương tính, tức là tỷ lệ các mẫu được dự đoán là dương tính và thực sự thuộc lớp đó.

$$Precision = \frac{TP}{TP + FP} \quad (1.4)$$

Giá trị của Precision nằm trong khoảng:

$$0 \leq Precision \leq 1$$

Precision cao cho thấy mô hình có ít dự đoán sai dương tính (false positive). Chỉ số này đặc biệt quan trọng trong các bài toán mà việc dự đoán sai dương tính có thể gây hậu quả nghiêm trọng.

Tuy nhiên, Precision cao không đảm bảo rằng mô hình phát hiện được tất cả các mẫu dương tính, vì mô hình có thể bỏ sót nhiều mẫu thật, dẫn đến Recall thấp.

1.5.4 Recall

Recall (còn gọi là Sensitivity hoặc True Positive Rate) đo lường khả năng của mô hình trong việc phát hiện đúng tất cả các mẫu thuộc lớp dương tính.

$$Recall = \frac{TP}{TP + FN} \quad (1.5)$$

Giá trị của Recall nằm trong khoảng:

$$0 \leq Recall \leq 1$$

Recall cao cho thấy mô hình có khả năng phát hiện hầu hết các mẫu thuộc lớp cần quan tâm. Chỉ số này đặc biệt quan trọng trong các bài toán mà việc bỏ sót mẫu dương tính là không mong muốn.

Tuy nhiên, Recall cao có thể đi kèm với Precision thấp nếu mô hình dự đoán quá nhiều mẫu là dương tính, dẫn đến tăng số lượng false positive.

1.5.5 F1-score

F1-score là trung bình điều hòa của Precision và Recall, được sử dụng để cân bằng giữa hai chỉ số này trong một thước đo duy nhất.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (1.6)$$

Giá trị của F1-score nằm trong khoảng:

$$0 \leq F1 \leq 1$$

Đối với bài toán phân loại đa lớp trong luận án, F1-score được tính theo trung bình macro (Macro F1-score). Trước hết, F1-score được xác định riêng cho từng lớp k dựa trên Precision và Recall của lớp đó:

$$F1_k = 2 \frac{Precision_k \times Recall_k}{Precision_k + Recall_k}. \quad (1.7)$$

Với K lớp tư thế, Macro F1-score được tính bằng trung bình số học của F1-score trên tất cả các lớp:

$$F1_{\text{macro}} = \frac{1}{K} \sum_{k=1}^K F1_k. \quad (1.8)$$

Trong luận án này, $K = 12$. Việc sử dụng Macro F1-score bảo đảm rằng mỗi lớp tư thế có trọng số như nhau trong chỉ số tổng hợp, không phụ thuộc vào số lượng mẫu của từng lớp. Do đó, chỉ số này bổ sung cho Accuracy trong việc đánh giá khả năng phân loại cân bằng giữa các tư thế, đặc biệt đối với các lớp có mức độ nhầm lẫn cao hơn. F1-score đạt giá trị cao khi cả Precision và Recall đều cao. Do đó, đây là chỉ số phù hợp để đánh giá hiệu quả của mô hình trong các bài toán có dữ liệu mất cân bằng.

1.6 Kết luận chương 1

Chương này đã xem xét bài toán phân loại tư thế ngủ chi tiết, ý nghĩa lâm sàng của nó và các thách thức kỹ thuật chính. Chúng tôi đã khảo sát cả các phương pháp tiếp xúc và không tiếp xúc, làm nổi bật sự đánh đổi vốn có giữa sự thoải mái, quyền riêng tư và tính mạnh mẽ. Trong số đó, các hệ thống đeo trên người dựa trên gia tốc

kể nổi lên như một hướng đi đầy hứa hẹn do tính di động và chi phí thấp, trong khi các phương pháp học sâu chứng minh khả năng mạnh mẽ trong việc trích xuất đặc trưng và phân loại tự động.

Những hạn chế này phát sinh chủ yếu từ cả thách thức về dữ liệu và mô hình hóa. Việc thu thập các tập dữ liệu quy mô lớn với các biến thể tư thế đa dạng và có liên quan về mặt lâm sàng vốn dĩ khó khăn do nhu cầu ghi âm kéo dài, gán nhãn chính xác và các điều kiện thực nghiệm được kiểm soát. Do đó, nhiều công trình giới hạn phạm vi của họ trong bốn tư thế kinh điển để đơn giản hóa việc thu thập dữ liệu và giảm độ phức tạp của việc chú thích. Hơn nữa, việc nhận dạng tư thế chi tiết làm tăng đáng kể sự tương đồng giữa các lớp trong tín hiệu cảm biến, điều này gây khó khăn cho các mô hình thông thường, đặc biệt là khi sử dụng các phương thức cảm biến hạn chế. Ngoài ra, các giao thức đánh giá chéo người dùng thường bị tránh vì sự khác biệt giữa các đối tượng về hình dáng cơ thể, hành vi ngủ và vị trí đặt cảm biến làm suy giảm đáng kể hiệu suất mô hình, dẫn đến việc các nhà nghiên cứu ưu tiên các thiết lập phụ thuộc vào đối tượng mang lại độ chính xác cao hơn nhưng ít có khả năng tổng quát hóa hơn.

Để lấp đầy những khoảng trống này, chúng tôi đề xuất một chuẩn so sánh mới với 12 nhân tư thế và thực thi các giao thức đánh giá độc lập với người dùng. Xây dựng trên tập dữ liệu này, chương tiếp theo (Chương 2) sẽ chi tiết hóa quá trình thu thập dữ liệu và giới thiệu một kiến trúc học sâu lai tận dụng các thể mạnh bổ trợ của 1D CNN cho việc trích xuất đặc trưng thời gian cục bộ và LSTM cho việc mô hình hóa các phụ thuộc tuần tự dài hạn. Sau đó, Chương 3 mở rộng hướng nghiên cứu này bằng cách khám phá các kiến trúc lai thay thế, tích hợp 1D CNN với các mô-đun GRU và BiLSTM để tăng cường hơn nữa việc học biểu diễn và hiệu suất phân loại.

CHƯƠNG 2. ĐỀ XUẤT MÔ HÌNH ANPONET VỚI SỐ LƯỢNG THAM SỐ RÀNG BUỘC CHO BÀI TOÁN PHÂN LOẠI TƯ THẾ NGỦ

Chương này có 2 mục tiêu chính. Mục tiêu thứ nhất là thiết lập nền tảng phương pháp luận để xây dựng một hệ thống nhận dạng tư thế ngủ chi tiết dựa vào cảm biến gia tốc gắn trên người. Dựa trên các kết quả đánh giá tính khả thi được trình bày trong Chương 1, chương này giới thiệu một bộ dữ liệu bao gồm 12 tư thế ngủ khác nhau, bao gồm cả các tư thế nằm nghiêng bán phần thể hiện góc lệch của cơ thể trong không gian. Chương này trình bày chi tiết về quá trình thu thập dữ liệu, quy trình gán nhãn và tiền xử lý, và các độ đo được sử dụng trong luận án.

Mục tiêu thứ hai là thiết kế mô hình học sâu AnpoNet cho bài toán phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc 3 trục, trong đó kiến trúc và bộ tham số được xây dựng có chủ đích nhằm kiểm soát độ phức tạp mô hình, hướng tới khả năng triển khai trên các thiết bị biên có tài nguyên hạn chế.

2.1 Phương pháp luận cho việc thiết kế bộ dữ liệu

2.1.1 Mục tiêu và phạm vi của bộ dữ liệu

Mục đích của việc xây dựng bộ dữ liệu này là cung cấp một nền tảng vững chắc về mặt kỹ thuật và mang ý nghĩa lâm sàng cho việc phân loại 12 tư thế ngủ, sử dụng cảm biến gia tốc gắn trên người. Khác với các nghiên cứu thông thường tập trung vào nhận dạng cơ bản (ví dụ: nằm ngửa, nằm sấp, nằm nghiêng trái và phải), luận án này nhắm đến phân loại 12 tư thế ngủ, đòi hỏi phải sử dụng mô hình học sâu kết hợp. Những phân biệt tư thế ngủ ở mức độ chi tiết như vậy có ý nghĩa lâm sàng đối với các rối loạn như hội chứng ngưng thở khi ngủ do tắc nghẽn, trào ngược dạ dày thực quản, loét do tì đè và phục hồi chức năng sau đột quỵ, nơi các sai lệch góc cụ thể có thể gây ra các rủi ro cho sức khỏe hoặc các liệu pháp điều trị.

Việc phát triển bộ dữ liệu này tuân theo một lộ trình nghiên cứu mang tính kế thừa và phát triển, dựa trên các công trình đã được công bố trước đó. Một nghiên cứu tiền đề được báo cáo trong công trình CT4 đã xác nhận tính khả thi của việc sử dụng gia tốc kế ba trục để phân biệt tư thế thông qua mô hình 2 lớp LSTM áp dụng cho phân loại bốn tư thế cơ bản. Mặc dù nghiên cứu trong CT4 đã chứng minh khả năng phân tách giữa các lớp tốt, nhưng sự đa dạng về tư thế và nhóm đối tượng là không

đủ cho các phân tích mang ý nghĩa lâm sàng. Tiếp đó, nghiên cứu trong CT3 đã giới thiệu một hệ thống cho thu thập dữ liệu và phân loại 12 tư thế phù hợp cho lâm sàng. Đóng góp về mặt phương pháp luận này đã xác nhận rằng tín hiệu gia tốc kế có thể mã hóa những sự thay đổi nhỏ trong góc tư thế, qua đó hỗ trợ việc mở rộng sang bài toán phân loại chi tiết.

Dựa trên những phát hiện đó, mục tiêu đặt ra khi thiết kế bộ dữ liệu trong luận án này bao gồm:

- Thiết lập một chiến lược về thu thập dữ liệu đảm bảo sự nhất quán về vị trí cảm biến, thực hiện tư thế, thời lượng ghi và hiệu chuẩn giữa các người tham gia.
- Thu thập các bộ dữ liệu với quy mô tăng dần số lượng đối tượng tham gia, cho phép cả đánh giá cơ sở và nghiên cứu khả năng mở rộng. Hai bộ dữ liệu dưới đây được thu thập một cách độc lập với nhau bao gồm:
 - *Dataset A*: 15 người tham gia được thu thập theo giao thức chuẩn hóa và được sử dụng để đánh giá mô hình học sâu cơ sở (AnpoNet).
 - *Dataset B*: 18 người tham gia với cùng giao thức trên được sử dụng trong Chương 3 để đánh giá các kiến trúc lai có khả năng biểu diễn tốt hơn.
- Hỗ trợ đánh giá chéo giữa các đối tượng, đảm bảo rằng hiệu suất phân loại phản ánh đúng sự biến thiên thực tế về đặc điểm nhân trắc học và hành vi ngủ.

Bộ dữ liệu thu được không chỉ đóng vai trò nền tảng để huấn luyện các kiến trúc học sâu mà còn đóng góp một phương pháp luận đáng tin cậy cho các nghiên cứu trong tương lai về lĩnh vực giám sát giấc ngủ dựa trên thiết bị đeo. Nhờ việc xây dựng dựa trên kết quả từ các nghiên cứu trong CT4 và CT3, luận án này thiết kế một bộ dữ liệu thống nhất và có thể mở rộng, cho phép tái lập các đánh giá trên mô hình cơ sở và các mô hình mở rộng, bắt đầu với khung thuật toán AnpoNet.

2.1.2 Cấu hình cảm biến và thiết lập phần cứng

Thiết bị đeo được đặt trên vùng bụng của tình nguyện viên (Hình 2.1) sẽ liên tục thu thập dữ liệu gia tốc ba trục ở tần số 50 Hz. Cảm biến ADXL345 được tích hợp vào vi điều khiển ESP8266 để thu thập dữ liệu, pin lithium-ion để cấp nguồn. Một vỏ nhựa bảo vệ đảm bảo độ bền. Thẻ nhớ microSD hỗ trợ lưu trữ dữ liệu tạm thời trong trường hợp kết nối bị gián đoạn. Bảng 2.1 cho thấy cấu hình phần cứng được

lựa chọn có kích thước nhỏ gọn, chi phí thấp và phù hợp với yêu cầu thu thập dữ liệu gia tốc trong bối cảnh thiết bị đeo. Việc sử dụng cảm biến ADXL345 và vi điều khiển ESP8266 cho phép hệ thống ghi nhận tín hiệu ba trục với tần số lấy mẫu phù hợp, đồng thời vẫn bảo đảm khả năng triển khai trong các kịch bản tài nguyên hạn chế.

ESP8266 là một vi điều khiển tích hợp Wi-Fi sử dụng bộ xử lý Tensilica Xtensa LX106. Thiết bị hỗ trợ chuẩn IEEE 802.11 b/g/n với tốc độ truyền dữ liệu lên tới 460,800 bps, cung cấp 4 MB bộ nhớ Flash và 80 KB SRAM. Với kích thước nhỏ gọn, mức tiêu thụ năng lượng thấp và khả năng hỗ trợ nhiều giao diện ngoại vi, ESP8266 phù hợp với vai trò thu nhận, lưu trữ tạm thời và truyền dữ liệu trong hệ thống nghiên cứu này.

Trong hệ thống hiện tại, ESP8266 được sử dụng chủ yếu cho nhiệm vụ thu thập, lưu trữ tạm thời và truyền dữ liệu gia tốc từ cảm biến ADXL345, chưa thực hiện trực tiếp quá trình suy luận của các mô hình học sâu đề xuất. Do đó, các thông số phần cứng của ESP8266 trong Bảng 2.1 được sử dụng để mô tả nền tảng thu nhận dữ liệu và làm cơ sở tham chiếu cho định hướng triển khai trên thiết bị biên, thay vì được xem là nền tảng đã được kiểm chứng cho quá trình suy luận mô hình.

ADXL345: ADXL345 là một gia tốc kế ba trục hiệu suất cao, vượt trội trong các ứng dụng chạy thời gian dài và ít tiêu thụ điện năng. Mức tiêu thụ điện năng cực thấp 23 μA ở chế độ đo và 0.1 μA ở chế độ chờ (điện hình ở 2.5 V) khiến cảm biến đo này trở nên lý tưởng cho các thiết bị chạy bằng pin. Ngoài ra, điện năng tiêu thụ có thể tự động thay đổi theo băng thông, giúp tối ưu hóa hơn nữa tuổi thọ pin. Hệ thống này sử dụng ADXL345 với gia tốc $\pm 2g$ ¹, cảm biến thể hiện nhiều lợi thế trong việc giám sát tư thế ngủ nhờ tập trung vào việc đo các chuyển động tinh tế đặc trưng của các tư thế ngủ phổ biến như nằm ngửa, nằm nghiêng và nằm sấp. Trong phạm vi hẹp này, ADXL345 cung cấp độ phân giải nâng cao, cho phép phân biệt rõ ràng giữa các tư thế ngủ.

Việc sử dụng ADXL345 ở dải hẹp $\pm 2g$ giúp tăng độ phân giải, điều này rất quan trọng để phân biệt các khác biệt nhỏ về độ nghiêng trong các tư thế ngủ chi tiết. Trong khi đó, ESP8266 cho phép truyền dữ liệu liên tục qua Wi-Fi đồng thời cung cấp bộ nhớ cục bộ qua microSD để ngăn mất dữ liệu khi kết nối mạng bị gián đoạn.

¹ $g=9.8 \text{ m/s}^2$

Bảng 2.1: Thông số kỹ thuật phần cứng của thiết bị đeo sử dụng trong nghiên cứu này.

Thành phần	Thông số kỹ thuật
Bộ vi điều khiển (MCU)	ESP8266 (32-bit Tensilica Xtensa LX106)
Tiêu chuẩn Wi-Fi	IEEE 802.11 b/g/n
Bộ nhớ	4 MB Flash, 80 KB SRAM
Tốc độ truyền dữ liệu	Lên tới 460,800 bps
Cảm biến	Gia tốc kế ba trục ADXL345 MEMS
Dải đo	$\pm 2g$
Độ phân giải	13-bit (3.9 mg/LSB)
Mức tiêu thụ điện năng	23 μA (chế độ đo), 0.1 μA (chế độ chờ)
Tần số lấy mẫu	50 Hz
Bộ nhớ cục bộ	Thẻ nhớ microSD
Vị trí đo	Vùng bụng
Nguồn điện	Pin sạc Lithium-ion
Vỏ bảo vệ	Vỏ nhựa cứng

Cơ sở lựa chọn và các yếu tố ảnh hưởng của vị trí đặt cảm biến Trong nghiên cứu này, cảm biến gia tốc ba trục được đặt tại vùng bụng nhằm thu nhận thông tin định hướng của phần thân trung tâm bằng một cấu hình chỉ sử dụng một cảm biến. Thành phần gia tốc trọng trường cho phép phản ánh sự thay đổi độ nghiêng của thân người giữa các tư thế nằm ngửa, nằm sấp và nằm nghiêng. Việc sử dụng một vị trí duy nhất cũng giúp giảm độ phức tạp phần cứng và mức độ bất tiện đối với người sử dụng.

Tuy nhiên, vị trí vùng bụng tồn tại một số nguồn biến thiên cần được xem xét. Thứ nhất, chuyển động của thành bụng do hô hấp có thể tạo ra thành phần dao động trong tín hiệu gia tốc ngay cả khi tư thế tổng thể không thay đổi. Các nghiên cứu trước đã cho thấy tín hiệu gia tốc đặt tại vùng bụng có thể được sử dụng để ước lượng nhịp thở, xác nhận rằng chuyển động hô hấp là một thành phần có thể đo được tại vị trí này. Vì vậy, trong bài toán phân loại tư thế, chuyển động hô hấp được xem là một nguồn biến thiên động chồng lên thành phần định hướng tĩnh của trọng trường.

Thứ hai, đặc điểm hình thể của từng người có thể ảnh hưởng đến tư thế tương đối giữa cảm biến và thân người. Sự khác biệt về vòng eo, độ cong của thành bụng và phân bố mô mềm có thể làm thay đổi góc tiếp xúc của thiết bị với bề mặt cơ thể. Ngoài ra, mức độ thắt đai quá chặt hoặc quá lỏng có thể làm thay đổi độ ổn định cơ học của cảm biến; nếu cảm biến bị xô dịch hoặc xoay tương đối so với thân người thì vectơ gia tốc đo được có thể thay đổi ngay cả khi tư thế cơ thể không đổi. Do đó, giao thức thu nhận yêu cầu cảm biến được gắn tại cùng một vùng giải phẫu, với hướng trục được xác định thống nhất và được cố định trong suốt phiên ghi nhằm giảm biến thiên

do cách đeo.

Thứ ba, tư thế nằm sắp tạo ra điều kiện cơ học khác biệt so với các tư thế còn lại do thiết bị có thể chịu áp lực trực tiếp giữa cơ thể và bề mặt giường. Áp lực này có khả năng làm thay đổi tư thế tương đối của cảm biến hoặc gây dịch chuyển cục bộ nếu thiết bị không được cố định chắc chắn. Vì vậy, đây là một nguồn sai số tiềm tàng của thiết kế hiện tại. Trong quá trình thu thập dữ liệu, cảm biến được cố định vào cơ thể và người tham gia được hướng dẫn duy trì tư thế ổn định nhằm hạn chế sự thay đổi vị trí cảm biến trong suốt thời gian ghi.

Mặc dù tồn tại các nguồn biến thiên trên, vùng bụng vẫn là một vị trí khả thi đối với giám sát tư thế ngủ. Doheny và cộng sự đã sử dụng gia tốc kế ba trục tại vùng ngực và bụng trên trong các phiên polysomnography và cho thấy tín hiệu gia tốc có thể đồng thời được sử dụng để ước lượng nhịp thở và phân loại tư thế ngủ. Điều này cho thấy chuyển động hô hấp không làm mất hoàn toàn thông tin định hướng của thân người; thay vào đó, tín hiệu đo tại vùng bụng chứa đồng thời thành phần định hướng tương đối chậm và thành phần dao động liên quan đến hô hấp.

Ngoài ra, các nghiên cứu về nhận dạng hoạt động sử dụng cảm biến đeo cho thấy các vị trí gần trung tâm cơ thể, như eo hoặc hông, thường phản ánh tốt chuyển động tổng thể do nằm gần tâm khối lượng cơ thể. Điều này cho thấy hông hoặc eo là những vị trí thay thế có cơ sở và có thể có lợi thế về độ ổn định cơ học trong một số ứng dụng. Tuy nhiên, nghiên cứu hiện tại chưa tiến hành đo đồng thời tại bụng, eo và hông trên cùng một nhóm đối tượng; do đó, chưa có đủ bằng chứng thực nghiệm để kết luận rằng một vị trí cụ thể là tối ưu cho bài toán phân loại 12 tư thế ngủ.

Vì vậy, lựa chọn vùng bụng trong luận án được xem là một lựa chọn thiết kế phù hợp với mục tiêu đánh giá khả năng phân loại tư thế ngủ chi tiết bằng một cảm biến gia tốc duy nhất, chứ không phải là khẳng định rằng đây là vị trí tối ưu tuyệt đối. Việc so sánh định lượng giữa các vị trí bụng, eo và hông, đồng thời đánh giá ảnh hưởng của đặc điểm hình thể, độ chặt của đai và áp lực tiếp xúc ở tư thế nằm sắp, là một hướng nghiên cứu tiếp theo cần được thực hiện.

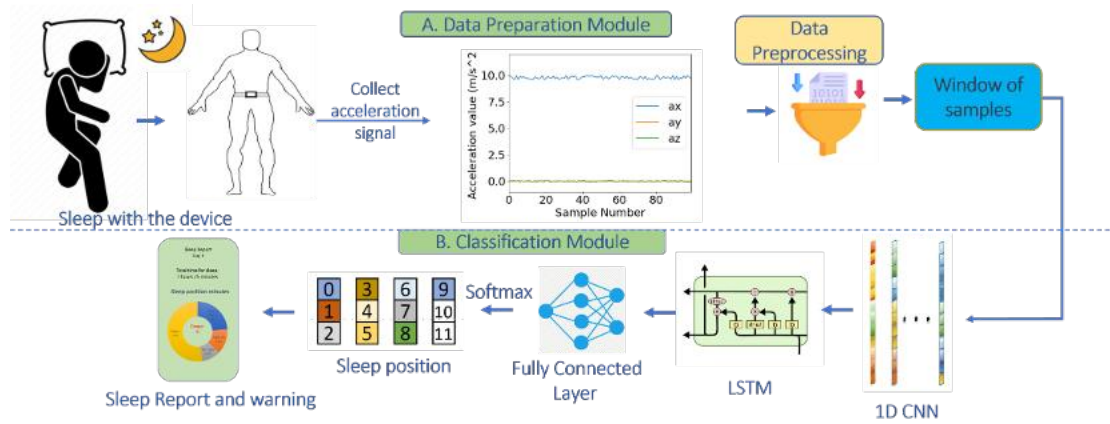
Cảm biến trong nghiên cứu này được đặt ở vùng bụng, dựa trên cơ sở sinh lý học và các bằng chứng từ những nghiên cứu gần đây về giám sát giấc ngủ bằng thiết bị đeo. Vùng bụng là một vị trí giải phẫu trung tâm phản ánh độ nghiêng tổng thể của cơ thể đồng thời nắm bắt được các chuyển động liên quan đến hô hấp. Các nghiên cứu về phát hiện ngưng thở khi ngủ cũng ủng hộ tính phù hợp trong cả hai khía cạnh này.

Ví dụ, Nghiên cứu của Hayano và cộng sự (2025) [110] đã chứng minh rằng một IMU tích hợp trong điện thoại thông minh đặt trên bụng trong quá trình đa ký giấc ngủ (polysomnography) có thể ghi lại liên tục các tín hiệu gia tốc và con quay hồi chuyển ba trục suốt cả đêm, qua đó xác nhận rằng chuyển động quán tính vùng bụng đủ ổn định cho việc giám sát giấc ngủ kéo dài. Tương tự, Dang và cộng sự (2025) [111] đã giới thiệu một thuật toán phát hiện OSA tự động sử dụng tín hiệu chuyển động vùng bụng được thu thập qua thiết bị không dây đeo trên bụng, qua đó củng cố tính phù hợp của vị trí này cho việc giám sát hô hấp và tư thế qua thiết bị đeo mà không gây ảnh hưởng đến nhịp sinh lý và trạng thái ngủ tự nhiên.

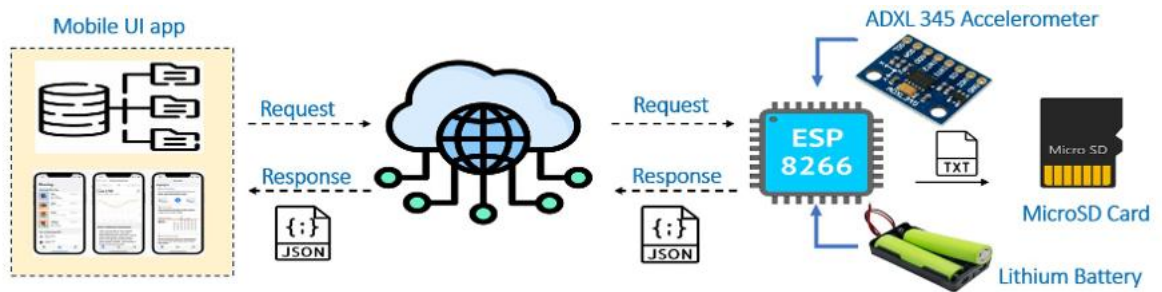
Về mặt sinh lý, vùng bụng trải qua sự dịch chuyển đồng bộ với sự giãn nở của lồng ngực và chuyển động của cơ hoành trong khi ngủ. Lin và cộng sự (2016) [112] đã chỉ ra rằng các tín hiệu chuyển động vùng bụng khi kết hợp với hoạt động lồng ngực sẽ có khả năng nắm bắt các kiểu hô hấp liên quan đến các sự kiện ngưng thở. Điều này ngụ ý rằng vùng bụng cung cấp không chỉ thông tin tư thế mà còn các dấu hiệu sinh lý liên quan đến nỗ lực hô hấp và định hướng thân mình, đều là những yếu tố có giá trị cho các ứng dụng đa phương thức trong tương lai.

Hơn nữa, vùng bụng ít bị ảnh hưởng bởi các nhiễu chuyển động hơn so với các vị trí khác như cổ tay hoặc mắt cá chân, vốn chủ yếu nắm bắt chuyển động cục bộ của tay hoặc chân thay vì tư thế toàn thân. So với ngực, vùng bụng tránh được việc tiếp xúc thường xuyên với giường và áp lực cơ bản từ khối lượng của chân trong khi vẫn xấp xỉ định hướng tổng thể của thân. Sự nhất quán của vị trí đặt ở bụng giữa các đối tượng cũng thúc đẩy sự đồng nhất trong các chỉ số cảm biến, một điều kiện thiết yếu cho đánh giá chéo giữa các đối tượng trong phân loại dựa trên học máy.

Dựa trên những cơ sở trên, việc gắn ở bụng mang lại sự cân bằng tối ưu giữa độ ổn định tín hiệu, sự phù hợp lâm sàng và sự thoải mái thực tế. Do đó, gia tốc kế sử dụng trong luận án này được đặt trên bụng để nắm bắt các dấu hiệu định hướng mạnh mẽ cho việc nhận dạng tư thế ở mức độ chi tiết.



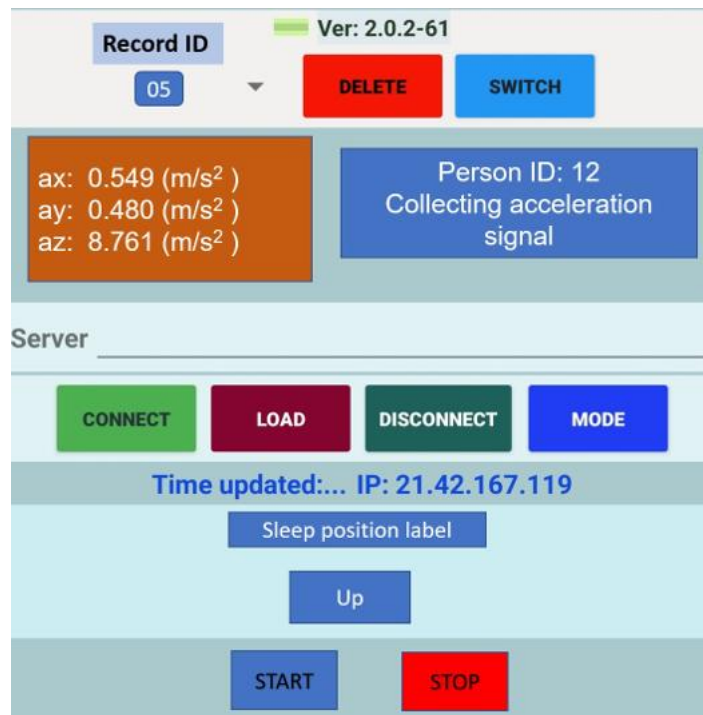
Hình 2.1: Hệ thống hai mô-đun cho phân loại 12 Tư thế Ngủ (12-SPSC)



Hình 2.2: Sơ đồ hệ thống thu thập dữ liệu gia tốc: ngoại tuyến và trực tuyến với ứng dụng di động.

Hình 2.1 cho thấy hệ thống được tổ chức thành hai mô-đun chính, bao gồm mô-đun thu thập dữ liệu từ cảm biến đeo và mô-đun xử lý/phân loại tư thế ngủ. Cách tổ chức này giúp tách biệt rõ quá trình ghi nhận tín hiệu gia tốc với quá trình xử lý dữ liệu và huấn luyện mô hình, qua đó thuận tiện cho việc triển khai, mở rộng và kiểm thử hệ thống.

Mô-đun thu thập dữ liệu giao tiếp không dây với một ứng dụng di động thân thiện với người dùng thông qua mạng cục bộ như trong Hình 2.2. Vì điều khiển hoạt động như một máy chủ, truyền dữ liệu định dạng JSON qua UDP khi có yêu cầu từ mô-đun ứng dụng ở phía người thu thập dữ liệu. Ứng dụng này được thiết kế để thu thập và xử lý dữ liệu trực quan như trong Hình 2.3.



Hình 2.3: Giao diện người dùng ứng dụng di động cho mô-đun thu thập dữ liệu.

2.1.3 Giao thức thu thập dữ liệu được chuẩn hóa

Việc lựa chọn 12 tư thế trong luận án tương ứng với việc lượng tử hóa một vòng quay của thân người thành các khoảng góc 30° . Cách lựa chọn này được xây dựng trên sự cân bằng giữa hai yêu cầu. Một mặt, các nghiên cứu lâm sàng cho thấy những thay đổi tư thế ở mức vài chục độ có thể dẫn đến khác biệt đáng kể về phân bố áp lực, mức độ ngưng thở khi ngủ hoặc hiệu quả của các chiến lược định vị cơ thể. Đặc biệt, góc nghiêng khoảng 30° thường được sử dụng như một mốc có ý nghĩa trong các nghiên cứu về giảm áp lực và định vị người bệnh.

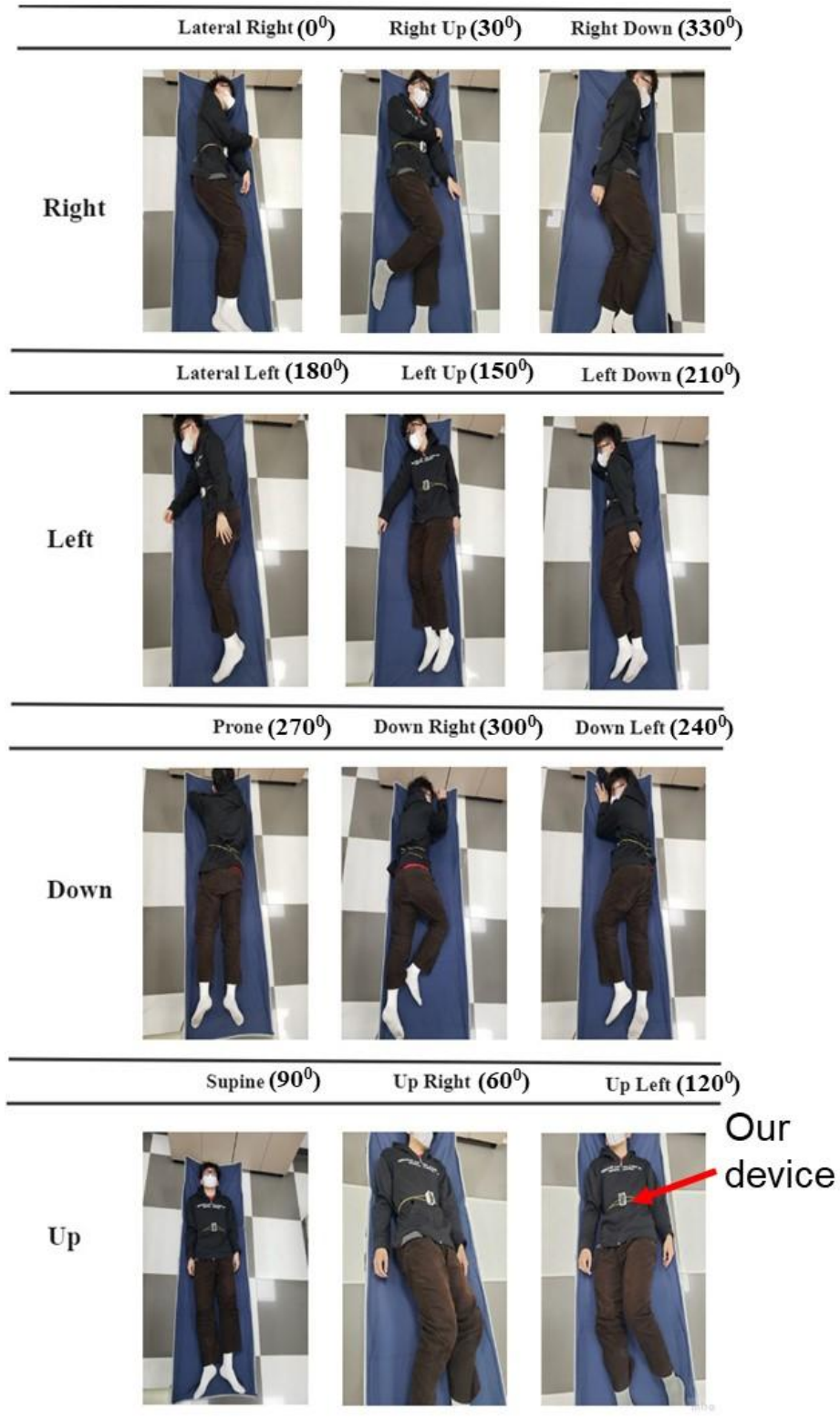
Mặt khác, việc sử dụng độ phân giải góc nhỏ hơn sẽ làm tăng nhanh số lượng lớp và mức độ chồng lấp giữa các phân bố tín hiệu gia tốc, trong khi độ phân giải quá thô sẽ làm mất đi các tư thế trung gian có ý nghĩa. Do đó, khoảng góc 30° được lựa chọn như một mức phân giải thực nghiệm hợp lý, tạo ra 12 lớp phủ toàn bộ miền định hướng 0° – 360° , vừa cho phép mô tả chi tiết hơn bốn tư thế cơ bản, vừa giữ số lượng lớp ở mức có thể thu thập, gán nhãn và đánh giá một cách ổn định trong phạm vi nghiên cứu.

Cần lưu ý rằng cách chia 30° không được xem là một ngưỡng sinh lý tuyệt đối cho mọi ứng dụng lâm sàng, mà là một lựa chọn thiết kế của luận án dựa trên các bằng chứng về ý nghĩa của sai lệch góc tư thế và yêu cầu cân bằng giữa độ phân giải tư thế

với độ phức tạp của bài toán phân loại.

Để đảm bảo việc thu thập dữ liệu nhất quán và có thể tái lập, một giao thức chuẩn hóa đã được thiết lập để đo tín hiệu gia tốc kể từ 12 tư thế ngủ được xác định trước. Tất cả người tham gia được hướng dẫn cách thực hiện từng tư thế một cách chính xác. Thiết bị cảm biến được gắn chắc chắn vào vùng bụng (Hình 2.4) để nắm bắt độ nghiêng của cơ thể, trong khi hệ tọa độ tham chiếu vẫn cố định trong suốt quá trình thu nhận. Trục x được xác định dọc theo mặt phẳng ngang, và góc tư thế được xác định bởi định hướng của một vectơ vuông góc với cơ thể, kéo dài từ lưng về phía bụng.

4 tư thế cơ bản bao gồm: nằm ngửa, nằm sấp, nằm nghiêng phải, và nằm nghiêng trái, tương ứng với các vị trí góc là 90° , 270° , 0° , và 180° . Ngoài các tư thế cơ bản này, tám biến thể ở các góc độ khác nhau cũng được thu thập để nắm bắt các sai lệch góc có liên quan về mặt lâm sàng. Các định hướng chi tiết này được xác định theo góc giữa trục Ox cố định và vectơ vuông góc của cơ thể, như được tóm tắt trong Bảng 2.2.



Hình 2.4: Các hình ảnh minh họa của 12 tư thế ngủ.

Bảng 2.2: Tóm tắt các tư thế ngủ dựa trên góc giữa hướng của bụng và trục x của hệ tọa độ gốc

Tên nhãn gốc	Góc quy ước (Độ)
Right (Lateral Right)	0
Right Up	30
Up Right	60
Up (Supine)	90
Up Left	120
Left Up	150
Left (Lateral Left)	180
Left Down	210
Down Left	240
Down (Prone)	270
Down Right	300
Right Down	330

Việc thu thập dữ liệu được thực hiện trong cùng một môi trường trong nhà có kiểm soát nhằm giảm ảnh hưởng của nhiều môi trường và sự khác biệt về điều kiện giường ngủ. Mỗi người tham gia được hướng dẫn duy trì từng tư thế trong thời gian 60 giây. Với tần số lấy mẫu 50 Hz, mỗi phiên tư thế tạo ra khoảng 3.000 mẫu gia tốc ba trục.

Thời lượng 60 giây được lựa chọn như một sự cân bằng giữa yêu cầu thu được đủ số mẫu đại diện cho trạng thái tư thế ổn định và yêu cầu hạn chế mệt mỏi, khó chịu hoặc dịch chuyển ngoài ý muốn của người tham gia khi phải duy trì một tư thế cố định. Với tần số lấy mẫu 50 Hz, một phút cung cấp 3.000 thời điểm đo cho mỗi tư thế và mỗi đối tượng, cho phép tiếp tục phân đoạn thành nhiều cửa sổ tín hiệu phục vụ quá trình huấn luyện và đánh giá mô hình. Do đó, việc kéo dài mỗi tư thế lên 3 hoặc 5 phút chủ yếu làm tăng số lượng mẫu liên tiếp của cùng một trạng thái tĩnh, nhưng không nhất thiết làm tăng tương ứng mức độ đa dạng của thông tin về tư thế nếu người tham gia vẫn được yêu cầu duy trì cùng một trạng thái cơ thể.

Việc sử dụng khoảng thời gian một phút cho mỗi tư thế cũng đã xuất hiện trong một số nghiên cứu thực nghiệm về nhận dạng tư thế ngủ. Chẳng hạn, một nghiên cứu sử dụng cảm biến đặt dưới giường trên 56 người tham gia yêu cầu mỗi đối tượng nằm

yên trong một phút ở từng tư thế nằm ngửa, nằm sấp, nghiêng trái và nghiêng phải.

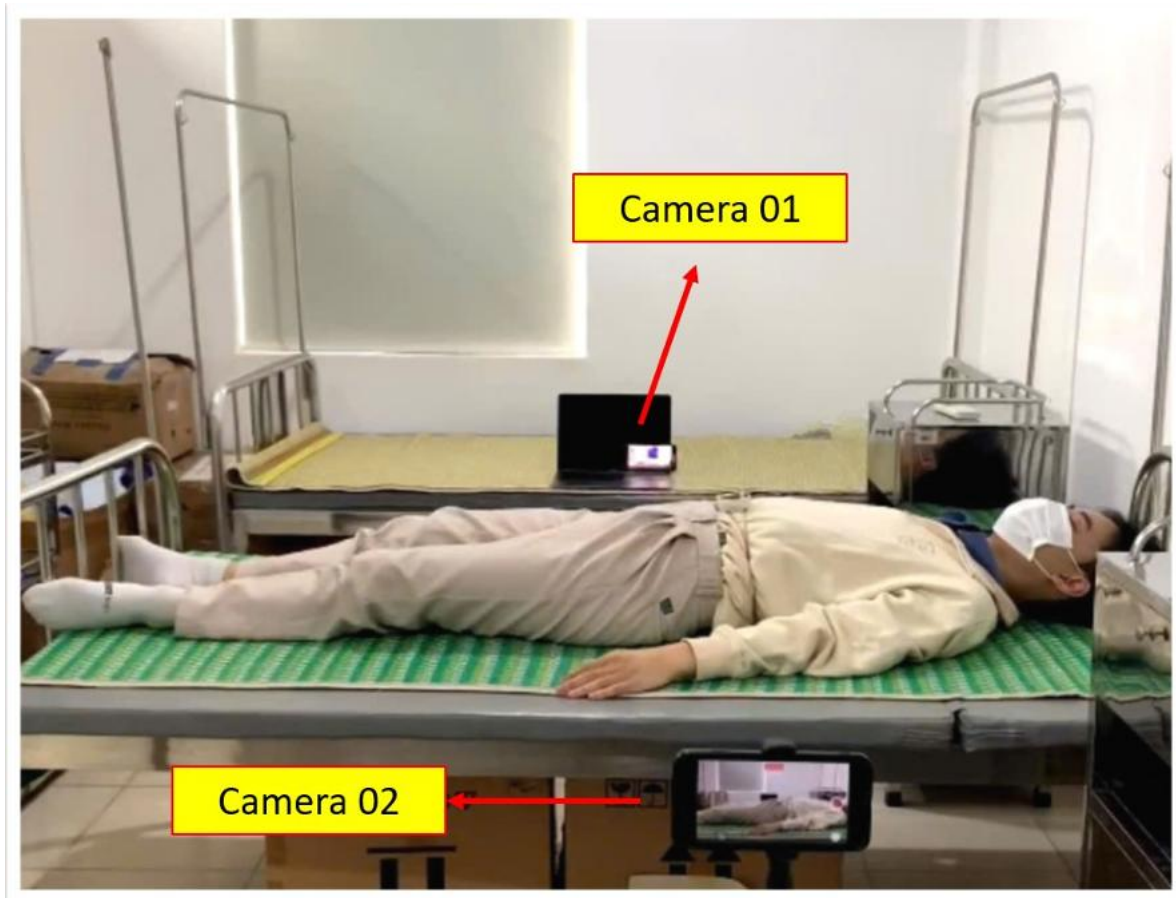
Tuy nhiên, thời lượng 60 giây trong nghiên cứu này không được xem là mô hình hóa thời gian duy trì tư thế trong giấc ngủ tự nhiên. Mục đích của giao thức là thu nhận một đoạn tín hiệu đủ dài và tương đối ổn định cho từng lớp tư thế đã xác định trước. Trong giấc ngủ qua đêm, một tư thế có thể được duy trì trong thời gian dài hơn nhiều, đồng thời xen kẽ các cử động ngắn, tư thế trung gian và quá trình chuyển tiếp giữa các tư thế. Do đó, việc tăng thời lượng của một tư thế tính từ 1 phút lên 3 hoặc 5 phút không giải quyết trực tiếp khoảng cách này, bởi dữ liệu thu thêm vẫn chủ yếu thuộc cùng một trạng thái ổn định.

Để cải thiện khả năng tổng quát hóa trong điều kiện thực tế, vấn đề quan trọng hơn không chỉ là kéo dài thời gian ghi của từng tư thế tĩnh mà là thu thập tín hiệu liên tục, trong đó người tham gia được phép chuyển đổi tự nhiên giữa các tư thế. Thiết kế như vậy sẽ bổ sung các cửa sổ chứa trạng thái chuyển tiếp, các tư thế trung gian và các chuyển động không ổn định, qua đó phản ánh sát hơn đặc điểm của tín hiệu trong giấc ngủ tự nhiên qua đêm.

Để bảo toàn chất lượng dữ liệu, các đối tượng có tình trạng cơ thể gây cản trở việc gắn cảm biến vùng bụng hoặc gây nhiễu tín hiệu chuyển động đã bị loại trừ. Bộ dữ liệu thu thập được phản ánh chuyển động bụng mà không có các nhiễu gây nhầm lẫn, cung cấp dữ liệu huấn luyện đáng tin cậy cho việc phân loại tư thế ngủ.

Giao thức được mô tả ở trên lần đầu tiên được giới thiệu như một phần của nền tảng phương pháp luận trong CT3 và được áp dụng thống nhất trên cả hai biến thể bộ dữ liệu (Dataset A và Dataset B), đảm bảo rằng sự khác biệt về hiệu suất sau đó phát sinh từ khả năng của mô hình thay vì sự không nhất quán trong thu nhận.

2.1.4 Quy trình gán nhãn và tiền xử lý



Hình 2.5: Quá trình gán nhãn tư thế

Để đảm bảo độ chính xác của việc gán nhãn tư thế, một quy trình gán nhãn với sự tham gia của ba người đã được áp dụng. Một người gán nhãn quan sát người tham gia từ xa nhằm giảm thiểu nguy cơ thiên kiến hành vi. Đồng thời, hai máy quay ghi lại video mỗi phiên thu thập từ các góc nhìn khác nhau như trong Hình 2.5.

Các tư thế được thiết lập theo các góc cách nhau 30° , tương ứng với 12 nhãn tư thế được xác định trước. Trong quá trình thu thập, người tham gia được hướng dẫn điều chỉnh thân người theo tư thế mục tiêu dưới sự quan sát của người thực hiện thí nghiệm. Hai nguồn video ở các góc nhìn khác nhau được sử dụng để hỗ trợ kiểm tra tư thế và gán nhãn sau quá trình thu thập.

Trong nghiên cứu này, góc tư thế được sử dụng như giá trị tham chiếu để xác định lớp tư thế, thay vì được đo liên tục bằng một thiết bị đo góc chuyên dụng trong từng thời điểm. Vì vậy, luận án không giả định rằng tất cả người tham gia có thể duy trì chính xác tuyệt đối góc danh định trong toàn bộ thời gian ghi. Các sai lệch nhỏ do

hình thể, cách nằm hoặc dịch chuyển tự nhiên của cơ thể được xem là một phần của biến thiên giữa các đối tượng.

Để hạn chế sai lệch trong quá trình gán nhãn, hai người đánh giá độc lập quan sát hai nguồn video riêng biệt, sau đó kết quả được đối chiếu cùng với người giám sát quá trình thu thập. Chỉ các đoạn dữ liệu đạt được sự đồng thuận về nhãn tư thế mới được sử dụng trong tập dữ liệu cuối cùng.

Dữ liệu gia tốc ba trục (các trục x , y , và z) được biểu diễn bởi các giá trị tương ứng ax , ay , và az . Các giá trị này có xu hướng biến thiên theo mỗi trục tùy thuộc vào tư thế ngủ cụ thể. Sau khi thu thập từ người tham gia, dữ liệu sẽ được tiền xử lý bằng cách phân đoạn thành các tập huấn luyện và kiểm tra. Tiếp theo, một kỹ thuật cửa sổ trượt với kích thước xác định sẽ được áp dụng cho dữ liệu thô để trích xuất các mẫu chồng lấp. Trước khi trích xuất đặc trưng, một bộ lọc thông thấp sẽ được áp dụng cho mỗi cửa sổ của các tín hiệu ax , ay , và az để làm suy giảm nhiễu tần số cao và các nhiễu chuyển động. Cách tiếp cận dựa trên cửa sổ kết hợp với lọc thông thấp này cho phép nhận biết các mối tương quan giữa các đoạn dữ liệu liên kế trên toàn bộ dữ liệu trong khi tập trung vào các thành phần tần số thấp hơn liên quan đến tư thế ngủ. Sau quá trình gán nhãn, dữ liệu gia tốc ba trục được xử lý theo một quy trình thống nhất trước khi đưa vào mô hình. Trước hết, các chuỗi tín hiệu được kiểm tra và tổ chức theo từng đối tượng và từng nhãn tư thế. Tín hiệu sau đó được xử lý nhằm giảm ảnh hưởng của nhiễu tần số cao trước khi được phân đoạn bằng cửa sổ trượt. Phép chuẩn hóa z-score được áp dụng theo công thức:

$$z = \frac{x - \mu}{\sigma}, \quad (2.1)$$

trong đó x là giá trị gia tốc ban đầu, μ và σ lần lượt là giá trị trung bình và độ lệch chuẩn được sử dụng cho phép chuẩn hóa. Việc chuẩn hóa đưa các thành phần gia tốc về thang giá trị tương đồng, qua đó giảm ảnh hưởng của sự khác biệt về biên độ giữa các trục và hỗ trợ quá trình tối ưu mô hình.

Sau bước chuẩn hóa, chuỗi tín hiệu được chia thành các cửa sổ có độ dài cố định với tỷ lệ chồng lấp được xác định trước. Các cửa sổ này tạo thành đơn vị đầu vào cho mô hình học sâu. Việc chia tập huấn luyện, xác thực và kiểm tra được thực hiện theo đối tượng nhằm bảo đảm rằng dữ liệu của đối tượng kiểm tra không xuất hiện trong quá trình huấn luyện.

2.1.5 Các bộ dữ liệu

Phần này mô tả đặc điểm của nhóm đối tượng tương ứng với Dataset A (15 người tham gia), được sử dụng trong Chương 2 để phát triển và đánh giá mô hình cơ sở phân loại tư thế ngủ. Nghiên cứu đã nhận được sự chấp thuận về đạo đức từ Hội đồng Đạo đức của Đại học Phenikaa (Mã số: 024.22/DHP-HDDD). Tình nguyện viên được tuyển dụng để tham gia thu thập dữ liệu, với độ tuổi từ 19-24 tuổi (trung bình: 20.56 ± 1.10 tuổi). Các đặc điểm thể chất như chiều cao (162 ± 4.14 cm), cân nặng (52.2 ± 6.34 kg), BMI (19.9 ± 2.62 kg/m²), và vòng eo (68.8 ± 9.53 cm; phạm vi: 58–85 cm) đã được ghi lại, như tóm tắt trong Bảng 2.3. Những số liệu này phản ánh các đặc điểm hình thể đa dạng, phù hợp cho việc huấn luyện một bộ phân loại cơ sở dưới các kích thước thân người khác nhau. Dataset A bao gồm 15 người tham gia có các thông tin nhân trắc học được ghi nhận bao gồm chiều cao, cân nặng, chỉ số BMI và vòng eo, như trình bày trong Bảng 2.3. Các thông số này cho thấy có sự biến thiên nhất định về đặc điểm hình thể giữa các đối tượng trong nhóm khảo sát.

Tuy nhiên, do số lượng người tham gia còn hạn chế và khoảng tuổi tương đối hẹp, Dataset A chưa đại diện cho sự đa dạng của quần thể người dùng nói chung, đặc biệt đối với các nhóm tuổi cao hơn, người có đặc điểm hình thể khác biệt hoặc các nhóm có bệnh lý liên quan đến giấc ngủ. Vì vậy, các kết quả thu được trên Dataset A chủ yếu được sử dụng để đánh giá khả năng của mô hình trong phạm vi nhóm đối tượng được khảo sát, thay vì khẳng định khả năng tổng quát hóa cho toàn bộ quần thể.

Cần lưu ý rằng Dataset A và Dataset B được xây dựng từ hai nhóm đối tượng độc lập và không trùng lặp. Mặc dù hai bộ dữ liệu được thu thập theo cùng một giao thức, sự khác biệt về thành phần đối tượng, đặc điểm nhân trắc học và phân bố tín hiệu có thể ảnh hưởng đến độ khó của bài toán và hiệu suất tuyệt đối của mô hình. Vì vậy, các giá trị độ chính xác và F1-Score được báo cáo trong Chương 2 và Chương 3 không nên được so sánh trực tiếp như một phép đối chứng tuyệt đối giữa AnpoNet và CGB-Net.

Bảng 2.3: Thống kê đặc điểm hình thể của các đối tượng tham gia trong Dataset A

Đặc điểm	Giá trị trung bình	Độ lệch chuẩn
Tuổi (năm)	20.56	1.10
Chiều cao (cm)	162.00	4.14
Cân nặng (kg)	52.30	6.34
Chỉ số BMI (kg/m^2)	19.90	2.62
Số đo vòng eo (cm)	68.80	9.53

Bên cạnh Dataset A, một nhóm người tham gia lớn hơn bao gồm 18 đối tượng không trùng lặp với Dataset A đã được thu thập để tạo nên bộ dữ liệu Dataset B, sẽ được sử dụng trong Chương 3. Việc đánh giá sự khác biệt giữa AnpoNet và CGB-Net chỉ có ý nghĩa khi hai mô hình được huấn luyện và đánh giá trên cùng một bộ dữ liệu, cùng cách chia đối tượng và cùng thiết lập tiền xử lý. Trong Chương 3, AnpoNet được sử dụng như một mô hình đối chứng và được đánh giá lại trên cùng Dataset B với CGB-Net. Do đó, các so sánh giữa hai kiến trúc trong Chương 3 phản ánh trực tiếp hơn ảnh hưởng của thiết kế mô hình, trong khi các kết quả riêng biệt giữa Chương 2 và Chương 3 chủ yếu được sử dụng để mô tả hiệu suất của từng mô hình trong phạm vi bộ dữ liệu tương ứng.

Giới hạn về quy mô mẫu và biến thiên hình thể của nhóm đối tượng

Một vấn đề cần được xem xét khi diễn giải kết quả của luận án là quy mô nhóm đối tượng còn tương đối hạn chế. Dataset A bao gồm 15 người tham gia, trong khi Dataset B được xây dựng từ 18 người tham gia khác và không trùng lặp với Dataset A. Do đó, các bộ dữ liệu hiện tại phù hợp hơn với vai trò của một tập dữ liệu thực nghiệm phục vụ đánh giá tính khả thi của phương pháp hơn là một mẫu đại diện cho quần thể người dùng rộng.

Quy mô này không phải là ngoại lệ trong các nghiên cứu về nhận dạng tư thế ngủ sử dụng cảm biến đeo hoặc cảm biến quán tính. Chẳng hạn, Yoon và cộng sự đánh giá phương pháp ước lượng tư thế ngủ sử dụng cảm biến gia tốc ba trục trên 13 đối tượng trong các phiên PSG ban đêm và báo cáo mức độ đồng thuận trung bình 99.16% với Cohen's kappa bằng 0.98 so với nhãn tham chiếu từ cảm biến tư thế và video [113]. Một nghiên cứu khác về hệ thống cảm biến đeo theo dõi tư thế và chất lượng giấc ngủ tuyển 16 người tham gia, trong đó 10 người thực hiện protocol tư thế, và đạt độ chính

xác 99.5% đối với bốn tư thế chính và 92.5% đối với tám tư thế [114]. Gần đây hơn, một phương pháp nhận dạng tư thế ngủ dựa trên đai cảm biến mềm được đánh giá trên 15 đối tượng và đạt độ chính xác 96.02% [115]. Các ví dụ này cho thấy việc sử dụng cohort có quy mô khoảng 10–20 người là tương đối phổ biến trong các nghiên cứu khả thi hoặc nghiên cứu xác thực ban đầu về cảm biến tư thế ngủ. Tuy nhiên, quy mô mẫu tương đương với các công trình trước không đồng nghĩa với việc bộ dữ liệu đã đủ đại diện về mặt dân số.

Đối với Dataset A, sự khác biệt hình thể giữa các đối tượng được phản ánh một phần thông qua các thông số nhân trắc học như chiều cao, cân nặng, BMI và vòng eo. Đặc biệt, vòng eo dao động từ 58 đến 85 cm. Sự khác biệt này có ý nghĩa trực tiếp đối với bài toán sử dụng cảm biến gia tốc gắn tại vùng bụng, bởi hình dạng cơ thể, kích thước vòng eo, độ cong bề mặt bụng và mức độ tiếp xúc giữa cảm biến với cơ thể có thể làm thay đổi góc đặt cảm biến cũng như phân bố tín hiệu gia tốc thu được. Vì vậy, biến thiên giữa các cá nhân cần được xem như một nguồn biến thiên của miền tín hiệu, chứ không chỉ là khác biệt về đặc điểm nhân trắc học.

Trong nghiên cứu này, ảnh hưởng của các đặc trưng riêng theo từng cá nhân được hạn chế thông qua thiết kế đánh giá độc lập theo đối tượng. Dữ liệu của các đối tượng dùng để kiểm thử không xuất hiện trong quá trình huấn luyện, qua đó giảm nguy cơ mô hình ghi nhớ các đặc trưng riêng của từng người. Cách đánh giá này nghiêm ngặt hơn so với việc chia ngẫu nhiên các cửa sổ tín hiệu từ cùng một đối tượng vào cả tập huấn luyện và tập kiểm thử. Các nghiên cứu trước cũng cho thấy sự khác biệt đáng kể giữa hai cách đánh giá này. Chẳng hạn, trong nghiên cứu phân loại tư thế ngủ trên 58 đối tượng sử dụng cảm biến đặt dưới giường, độ chính xác có thể đạt tới 100% trong một số cấu hình khi sử dụng k-fold cross-validation, trong khi hiệu suất giảm rõ rệt khi chuyển sang Leave-One-Subject-Out (LOSO), cho thấy đánh giá độc lập theo đối tượng phản ánh thách thức khái quát hóa sang người dùng mới một cách thực tế hơn [114].

Trong luận án, Dataset A cũng được đánh giá bằng kiểm định chéo ở cấp độ người tham gia và các thí nghiệm trên các đối tượng chưa xuất hiện trong huấn luyện. Việc lặp lại đánh giá trên các phép chia đối tượng khác nhau và qua nhiều lần khởi tạo ngẫu nhiên giúp kiểm tra mức độ ổn định của kết quả trong phạm vi cohort hiện tại. Tuy nhiên, các thủ tục này không làm tăng kích thước mẫu và cũng không loại bỏ hạn chế về tính đại diện của quần thể.

Do đó, các giá trị accuracy và F1-Score được báo cáo trong luận án nên được diễn giải là ước lượng hiệu năng của mô hình trong phạm vi nhóm đối tượng được khảo sát và dưới giao thức đánh giá độc lập theo đối tượng, thay vì bằng chứng thống kê cho khả năng tổng quát hóa trên toàn bộ quần thể. Việc xác nhận khả năng khái quát hóa ở cấp quần thể cần được thực hiện trên cohort lớn hơn, với độ tuổi rộng hơn và có phân tầng rõ theo giới tính, BMI, vòng eo và các đặc điểm hình thể có thể ảnh hưởng đến vị trí gắn cảm biến.

Trong các nghiên cứu tiếp theo, việc mở rộng quy mô mẫu cần đi kèm với báo cáo khoảng tin cậy của các chỉ số hiệu năng và phân tích theo các nhóm nhân trắc học. Khi số lượng đối tượng đủ lớn, có thể kiểm tra sự khác biệt về hiệu năng giữa các nhóm BMI, vòng eo hoặc độ tuổi khác nhau, từ đó đánh giá liệu mô hình có duy trì hiệu suất ổn định giữa các nhóm người dùng hay không.

Khoảng cách giữa dữ liệu tư thế tĩnh và dữ liệu ngủ tự nhiên

Bộ dữ liệu được sử dụng trong luận án được xây dựng nhằm tạo ra một thiết lập thực nghiệm có kiểm soát cho bài toán phân loại chi tiết 12 tư thế ngủ. Trong giao thức thu thập, người tham gia được hướng dẫn duy trì từng tư thế xác định trước trong thời gian một phút, với tần số lấy mẫu 50 Hz. Cách tổ chức này tạo ra các đoạn tín hiệu có nhân tương đối ổn định, giảm mức độ nhập nhằng giữa các lớp và cho phép đánh giá có kiểm soát khả năng phân biệt các tư thế có sai khác góc nhỏ. Tuy nhiên, dữ liệu thu được chủ yếu biểu diễn các trạng thái tư thế ổn định và tồn tại một khoảng cách đáng kể so với dữ liệu được ghi nhận trong giấc ngủ tự nhiên qua đêm.

Thứ nhất, sự khác biệt quan trọng nằm ở *tính liên tục của chuyển động*. Trong giao thức thực nghiệm, mỗi tư thế được duy trì tương đối cố định trong một khoảng thời gian xác định. Ngược lại, trong giấc ngủ tự nhiên, sự thay đổi tư thế thường diễn ra dưới dạng một quá trình liên tục, bao gồm giai đoạn bắt đầu chuyển động, xoay thân, điều chỉnh tay hoặc chân, dịch chuyển cảm biến và trạng thái ổn định sau chuyển động. Do đó, tín hiệu trong quá trình chuyển tiếp có thể không tương ứng rõ ràng với bất kỳ một trong 12 nhân tư thế rời rạc đã được định nghĩa.

Thứ hai, dữ liệu ngủ thực tế chứa các *trạng thái trung gian và nhập nhằng*. Hệ thống nhãn của luận án mô tả các tư thế tại những góc quy ước xác định, trong khi một người ngủ thực tế có thể duy trì cơ thể ở bất kỳ góc trung gian nào. Một cửa sổ tín hiệu vì vậy có thể nằm giữa hai lớp lân cận, chẳng hạn giữa tư thế nằm ngửa và

tư thế nghiêng 30 độ. Trong dữ liệu được kiểm soát, sự nhập nhằng này được hạn chế nhờ hướng dẫn người tham gia duy trì tư thế định trước; trong dữ liệu tự nhiên, đây lại là một đặc tính vốn có của bài toán.

Thứ ba, dữ liệu ngủ tự nhiên có mức *biến thiên động học* cao hơn. Bên cạnh chuyển đổi tư thế chính, tín hiệu gia tốc có thể chịu ảnh hưởng của các chuyển động cục bộ của tay, chân và thân mình, các cử động ngắn khi ngủ, thay đổi nhịp thở, rung động từ giường hoặc người ngủ cùng, cũng như sự dịch chuyển tương đối của cảm biến trên cơ thể. Các thành phần này có thể làm thay đổi tín hiệu tức thời ngay cả khi tư thế cơ thể tổng thể không thay đổi. Trong khi đó, giao thức thu thập có kiểm soát chủ động hạn chế các nguồn biến thiên này nhằm tạo dữ liệu có chất lượng và nhãn rõ ràng.

Thứ tư, tồn tại khác biệt về *phân bố thời gian và phân bố lớp*. Trong bộ dữ liệu thực nghiệm, thời lượng của từng tư thế được quy định trước, dẫn đến số lượng mẫu giữa các lớp tương đối cân bằng. Trong một giấc ngủ qua đêm, thời gian duy trì từng tư thế phụ thuộc mạnh vào từng cá nhân và có thể mất cân bằng đáng kể. Một số tư thế có thể được duy trì hàng chục phút, trong khi một số tư thế trung gian chỉ xuất hiện trong thời gian rất ngắn. Vì vậy, phân bố lớp trong dữ liệu thực tế có thể khác đáng kể so với phân bố của bộ dữ liệu thực nghiệm.

Thứ năm, giao thức hiện tại chưa mô hình hóa đầy đủ *sự phụ thuộc dài hạn theo thời gian*. Mỗi đoạn dữ liệu được thu thập trong một tư thế xác định và sau đó được phân đoạn thành các cửa sổ để huấn luyện mô hình. Điều này cho phép đánh giá khả năng nhận dạng một tư thế từ một đoạn tín hiệu cục bộ, nhưng chưa kiểm chứng khả năng của mô hình trong việc theo dõi một chuỗi tư thế liên tục kéo dài nhiều giờ, trong đó thời điểm trước và sau của một cửa sổ có thể chứa các tư thế khác nhau.

Những khác biệt trên có ý nghĩa trực tiếp đối với cách diễn giải kết quả thực nghiệm. Các giá trị accuracy và F1-Score đạt được trong luận án phản ánh khả năng phân biệt 12 tư thế ổn định trong điều kiện dữ liệu được kiểm soát và đánh giá độc lập theo đối tượng; chúng chưa thể được diễn giải trực tiếp như hiệu suất của một hệ thống giám sát tư thế trong giấc ngủ tự nhiên qua đêm. Đặc biệt, hiệu suất thực tế có thể suy giảm tại các thời điểm chuyển tiếp, với các tư thế trung gian, hoặc khi xuất hiện các chuyển động ngoài phạm vi dữ liệu huấn luyện.

Tuy nhiên, bộ dữ liệu tĩnh vẫn có vai trò như một bước nghiên cứu cơ sở. Việc kiểm soát tư thế, thời lượng ghi và điều kiện thu nhận cho phép tách biệt tương đối rõ

ảnh hưởng của kiến trúc mô hình khỏi các yếu tố gây nhiễu của môi trường, đồng thời cung cấp một thiết lập phù hợp để đánh giá khả năng phân biệt các tư thế có sai khác góc nhỏ. Vì vậy, kết quả của luận án nên được xem là bằng chứng về tính khả thi của phương pháp đối với bài toán nhận dạng tư thế ngủ chi tiết, thay vì bằng chứng hoàn chỉnh về hiệu năng trong điều kiện ngủ tự nhiên.

Để thu hẹp khoảng cách này, hướng nghiên cứu tiếp theo cần xây dựng dữ liệu được ghi liên tục trong các phiên ngủ kéo dài hoặc qua đêm, trong đó quá trình thay đổi tư thế diễn ra tự nhiên. Việc gán nhãn có thể cần kết hợp tín hiệu cảm biến với video hồng ngoại hoặc một phương thức tham chiếu độc lập để xác định chính xác thời điểm bắt đầu và kết thúc của các chuyển tiếp. Bộ dữ liệu mở rộng cũng cần bao gồm các tư thế trung gian, các cửa sổ chứa nhiều hơn một trạng thái tư thế và các chuyển động không liên quan trực tiếp đến thay đổi tư thế. Trên cơ sở đó, bài toán có thể được mở rộng từ phân loại các cửa sổ tư thế tĩnh sang nhận dạng liên tục, phát hiện chuyển tiếp và ước lượng tư thế trong điều kiện ngủ tự nhiên.

2.1.6 Thiết lập thực nghiệm

Tất cả các thí nghiệm trong luận án này được thực hiện trong hai môi trường tính toán riêng biệt: một môi trường được sử dụng riêng cho việc huấn luyện và tối ưu hóa mô hình, và môi trường còn lại dành riêng cho việc đánh giá hiệu năng suy luận của mô hình. Việc tách biệt các môi trường này đảm bảo tính minh bạch trong đo lường và tránh các sai lệch khi so sánh hiệu quả tính toán giữa các mô hình khác nhau.

Môi trường huấn luyện

Việc huấn luyện và tinh chỉnh siêu tham số cho các mô hình đề xuất được thực hiện trên một hệ thống tính toán hiệu năng cao (HPC) được trang bị một GPU NVIDIA RTX A5000 (24 GB VRAM). Tất cả các mô hình học sâu được triển khai bằng Python sử dụng khung học sâu TensorFlow (phiên bản 2.10). Kích thước mini-batch thống nhất là 512 được áp dụng cho tất cả các thí nghiệm. Việc sử dụng GPU đảm bảo tối ưu hóa ổn định, hội tụ nhanh hơn và điều kiện huấn luyện nhất quán giữa cả các kiến trúc đề xuất và các mô hình cơ sở.

Môi trường suy diễn

Để đánh giá và so sánh độ trễ suy luận giữa các mô hình trong cùng một điều kiện phần cứng, tất cả các thử nghiệm suy luận được thực hiện trên một hệ thống máy tính xách tay chỉ sử dụng CPU. Thiết bị được trang bị bộ xử lý Intel Core i5-8350U (tần số cơ bản 1.70 GHz, xung nhịp tối đa 1.90 GHz), RAM 32 GB và hệ điều hành Windows 64-bit. Trong quá trình suy luận không sử dụng GPU và không áp dụng các kỹ thuật tối ưu hóa chuyên biệt cho phần cứng, chẳng hạn TensorRT hoặc biên dịch đồ thị.

Thời gian suy luận của mỗi mô hình được đo trực tiếp trên nền tảng phần cứng nêu trên và được xác định bằng độ trễ trung bình của 1,000 lần truyền thẳng độc lập với kích thước batch bằng 1. Thiết lập batch size bằng 1 được lựa chọn nhằm mô phỏng cách xử lý tuần tự từng cửa sổ tín hiệu trong hệ thống giám sát tư thế ngủ, thay vì xử lý đồng thời một số lượng lớn mẫu như trong quá trình huấn luyện. Tất cả các mô hình được đánh giá với cùng cấu hình phần cứng, phần mềm và giao thức đo nhằm bảo đảm tính nhất quán và công bằng khi so sánh độ trễ suy luận.

Cần lưu ý rằng các giá trị thời gian suy luận được báo cáo trong luận án phản ánh hiệu năng của các mô hình trên nền tảng CPU Intel Core i5-8350U và không phải là kết quả đo trực tiếp trên ESP8266, vi điều khiển hoặc thiết bị đeo công suất thấp. Vì vậy, các giá trị này chủ yếu được sử dụng để so sánh tương đối chi phí thực thi giữa các mô hình trong cùng môi trường thử nghiệm. Khả năng đáp ứng thời gian thực trên thiết bị biên có tài nguyên hạn chế cần được đánh giá riêng thông qua triển khai và đo trực tiếp trên nền tảng phần cứng mục tiêu.

2.1.7 Xử lý dữ liệu đầu vào và tổ chức tập dữ liệu

Dữ liệu đầu vào của hệ thống là các chuỗi gia tốc ba trục được thu thập từ cảm biến gia tốc gắn trên cơ thể người tham gia thí nghiệm. Quá trình thu thập được thực hiện với tần số lấy mẫu cố định $f = 50\text{Hz}$. Mỗi đối tượng thực hiện tổng cộng 12 tư thế ngủ khác nhau; mỗi tư thế được duy trì liên tục trong 60 giây. Do đó, đối với mỗi đối tượng, dữ liệu thô được tổ chức thành 12 tệp riêng biệt (định dạng CSV), mỗi tệp tương ứng với một tư thế ngủ và chứa xấp xỉ

$$60 \times 50 = 3000$$

mẫu theo thời gian.

Mỗi tệp CSV lưu trữ chuỗi gia tốc ba trục dưới dạng bảng, trong đó mỗi hàng tương ứng với một mẫu gia tốc tại một thời điểm lấy mẫu cụ thể. Cấu trúc dữ liệu của mỗi hàng bao gồm: (i) dấu thời gian (timestamp), (ii) gia tốc theo trục x , (iii) gia tốc theo trục y , (iv) gia tốc theo trục z , và (v) nhãn tư thế ngủ tương ứng. Các tư thế ngủ được mã hóa bằng các nhãn số nguyên từ 0 đến 11 theo một ánh xạ được xác định trước.

Trong quá trình thu thập, toàn bộ các mẫu trong cùng một tệp CSV thuộc về cùng một tư thế ngủ. Do đó, nhãn tư thế được gán ở mức tệp và được kế thừa cho tất cả các mẫu trong tệp đó. Việc biểu diễn nhãn ở cấp độ từng mẫu thời gian trong bảng dữ liệu chỉ nhằm mục đích minh họa và thuận tiện cho quá trình xử lý dữ liệu, không hàm ý rằng nhãn được suy ra tại từng thời điểm riêng lẻ.

Bảng 2.4 minh họa một đoạn dữ liệu gia tốc ba trục tương ứng với một tư thế ngủ cụ thể (ví dụ nhãn 3). Cột *Timestamp* được tính dựa trên tần số lấy mẫu $f = 50\text{Hz}$, tương ứng với khoảng thời gian 0.02 giây giữa hai mẫu liên tiếp. Các cột a_x , a_y và a_z biểu diễn giá trị gia tốc đo được theo ba trục của cảm biến. Cột *Nhãn tư thế* nhận giá trị không đổi trong toàn bộ tệp, phản ánh tư thế ngủ đang được thực hiện trong khoảng thời gian thu thập đó.

Bảng 2.4: Ví dụ minh họa cấu trúc dữ liệu gia tốc ba trục cho một tư thế ngủ

Timestamp (s)	a_x (m/s ²)	a_y (m/s ²)	a_z (m/s ²)	Nhãn tư thế
0.00	0.12	-9.68	0.45	3
0.02	0.11	-9.70	0.47	3
0.04	0.13	-9.69	0.46	3
0.06	0.12	-9.71	0.44	3
0.08	0.10	-9.70	0.45	3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
59.96	0.14	-9.67	0.48	3
59.98	0.13	-9.69	0.46	3

Sau khi đọc các tệp CSV và gán nhãn tư thế tương ứng, dữ liệu được tổ chức theo đối tượng để phục vụ thiết lập phân loại độc lập theo đối tượng. Cụ thể, toàn bộ các tệp của một đối tượng chỉ thuộc về một tập duy nhất (huấn luyện hoặc kiểm tra). Việc phân đoạn tín hiệu thành các cửa sổ thời gian, chuẩn hóa dữ liệu và tạo các mẫu

huấn luyện/kiểm tra chỉ được thực hiện sau khi đã hoàn tất bước tách theo đối tượng, nhằm tránh rò rỉ thông tin giữa các tập dữ liệu.

2.2 Thiết kế mô hình học sâu hướng tới thực thi trên thiết bị biên

Trong những năm gần đây, các hệ thống theo dõi và phân tích giấc ngủ ngày càng được quan tâm nhờ vai trò quan trọng của giấc ngủ đối với sức khỏe con người. Trong đó, phân loại tư thế ngủ dựa trên dữ liệu cảm biến gia tốc đeo người là một hướng tiếp cận có nhiều tiềm năng nhờ tính không xâm lấn, chi phí thấp và khả năng triển khai liên tục trong môi trường thực tế. Tuy nhiên, bài toán phân loại tư thế ngủ đặt ra nhiều thách thức, đặc biệt khi số lượng tư thế cần phân biệt lớn và các tư thế có đặc điểm chuyển động gần nhau.

Các phương pháp học sâu gần đây đã cho thấy hiệu quả vượt trội trong xử lý dữ liệu chuỗi thời gian từ cảm biến. Tuy nhiên, phần lớn các mô hình đề xuất có kiến trúc phức tạp, số lượng tham số lớn và yêu cầu tài nguyên tính toán cao, điều này gây khó khăn cho việc triển khai trên các thiết bị biên như thiết bị đeo người hoặc hệ thống giám sát nhúng. Trong bối cảnh đó, việc thiết kế các mô hình học sâu gọn nhẹ, vừa đảm bảo hiệu quả phân loại vừa phù hợp với các ràng buộc về bộ nhớ và tính toán, vẫn là một bài toán mở và có ý nghĩa thực tiễn cao.

Đối với bài toán phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc 3 trục, yêu cầu về mô hình càng trở nên khắt khe hơn do số lượng lớp lớn và sự chồng lấn đặc trưng giữa các tư thế. Điều này đòi hỏi mô hình không chỉ có khả năng trích chọn đặc trưng cục bộ theo thời gian mà còn phải mô hình hóa được các phụ thuộc tuần tự của chuỗi, trong khi vẫn kiểm soát được độ phức tạp tổng thể của mô hình.

Xuất phát từ những thách thức trên, Chương này tập trung vào việc thiết kế mô hình học sâu AnpoNet theo định hướng hướng tới thiết bị biên, trong đó kiến trúc và các tham số của mô hình được xây dựng có chủ đích nhằm giảm số lượng tham số và chi phí tính toán. Mục tiêu đặt ra là đạt được sự cân bằng giữa khả năng biểu diễn cần thiết cho bài toán phân loại 12 tư thế ngủ và yêu cầu gọn nhẹ của mô hình, tạo tiền đề cho khả năng triển khai trên các nền tảng thiết bị biên có tài nguyên hạn chế.

Các kết quả được trình bày trong chương này đã được công bố trong một bài báo tạp chí SCIE, tương ứng với CT1 trong danh sách các đóng góp khoa học của tác giả.

2.2.1 Phương pháp tiếp cận của AnpoNet

Tư thế ngủ là một nguồn thông tin quan trọng trong các hệ thống giám sát sức khỏe thông minh, do mối liên hệ của chúng với nhiều tình trạng y tế khác nhau như trào ngược dạ dày thực quản (GERD), ngưng thở khi ngủ do tắc nghẽn (OSA), loét do tì đè và quá trình phục hồi sau đột quỵ. Tuy nhiên, xét dưới góc độ xử lý tín hiệu và học máy, phân loại tư thế ngủ không thể được xem như một trường hợp đặc biệt đơn giản của bài toán nhận dạng hoạt động dựa trên cảm biến quán tính.

Khác với các hoạt động sinh hoạt thường ngày, vốn đặc trưng bởi các chuyển động có động học mạnh và biến thiên biên độ lớn, tư thế ngủ có bản chất là các trạng thái gần như tĩnh, được duy trì trong khoảng thời gian dài, với sự khác biệt giữa các lớp chủ yếu thể hiện qua các sai lệch góc nhỏ giữa các trục gia tốc. Mặc dù các sai lệch này có thể mang ý nghĩa y sinh quan trọng, chúng lại chỉ tạo ra những thay đổi rất tinh tế trong tín hiệu đo, khiến việc phân biệt các biến thể tư thế chi tiết trở nên đặc biệt thách thức từ góc độ mô hình hóa.

Gia tốc kế cho phép thu nhận tín hiệu liên tục, không xâm lấn và phù hợp cho giám sát dài hạn, do đó thường được lựa chọn trong các hệ thống phân loại tư thế ngủ. Tuy nhiên, việc chuyển đổi dữ liệu gia tốc thô thành các dự đoán tư thế đáng tin cậy đặt ra một số thách thức cơ bản. Thứ nhất, thông tin về tư thế được mã hóa chủ yếu trong thành phần trọng lực tĩnh của tín hiệu, đòi hỏi các phương pháp trích xuất đặc trưng phải ưu tiên các mẫu không gian ổn định thay vì các đặc trưng chuyển động ngắn hạn. Thứ hai, các chuyển tiếp tư thế khi ngủ xảy ra không thường xuyên và ít lặp lại, làm giảm hiệu quả của các mô hình được thiết kế cho nhận dạng hoạt động, vốn thường dựa vào các biến đổi động rõ rệt. Thứ ba, sự khác biệt giữa các đối tượng, bao gồm hình thái cơ thể, thói quen ngủ và sai lệch trong vị trí gắn cảm biến, làm gia tăng yêu cầu đối với các biểu diễn học được có khả năng tổng quát hóa và ít phụ thuộc vào cá thể cụ thể.

Xuất phát từ các đặc điểm trên, việc thiết kế mô hình cho bài toán phân loại tư thế ngủ cần tuân thủ một số nguyên tắc cốt lõi sau:

- **Phân biệt định hướng trong không gian:** Mô hình cần nắm bắt chính xác các quan hệ hình học giữa các trục gia tốc, do sự khác biệt giữa các tư thế chủ yếu được thể hiện qua các sai lệch góc nhỏ và ổn định theo thời gian.
- **Mô hình hóa tính ổn định theo thời gian:** Do tư thế ngủ thường được duy

trì trong các khoảng thời gian dài, quá trình phân loại cần nhấn mạnh tính nhất quán dài hạn của tín hiệu, đồng thời triệt tiêu các dao động ngắn hạn và nhiễu cảm biến. Vì vậy, các cơ chế mô hình hóa thời gian dựa trên bộ nhớ, chẳng hạn như mạng nơ-ron hồi quy, là cần thiết để tổng hợp thông tin theo chuỗi.

- **Học biểu diễn độc lập đối tượng:** Trong bối cảnh hệ thống không yêu cầu hiệu chuẩn riêng cho từng người dùng, các đặc trưng học được cần có khả năng tổng quát hóa tốt trên nhiều đối tượng khác nhau, bất chấp sự khác biệt về hình thái cơ thể và vị trí gắn cảm biến.

Các nguyên tắc trên định hướng việc lựa chọn và xây dựng các mô hình học sâu lai, trong đó các mô-đun tích chập đảm nhiệm vai trò trích xuất đặc trưng không gian cục bộ, còn các mô-đun hồi quy chịu trách nhiệm mô hình hóa sự ổn định và quan hệ theo thời gian. Phần tiếp theo sẽ chính thức hóa bài toán phân loại tư thế ngủ trong bối cảnh độc lập đối tượng, đồng thời trình bày các cơ sở toán học được sử dụng xuyên suốt luận án.

2.2.2 Mô hình hóa bài toán

Bài toán phân loại tư thế ngủ có thể được xem như một bài toán phân loại chuỗi thời gian đa biến, trong đó mục tiêu là ánh xạ các đoạn tín hiệu cảm biến quán tính thu được từ cơ thể người sang một tập hữu hạn các nhãn tư thế rời rạc. Trong luận án này, bài toán được nghiên cứu trong bối cảnh *phân loại độc lập theo đối tượng*, tức là mô hình được huấn luyện trên dữ liệu thu thập từ một nhóm đối tượng và được đánh giá trên các đối tượng hoàn toàn mới, chưa từng xuất hiện trong quá trình huấn luyện. Thiết lập này phản ánh yêu cầu tổng quát hóa trong các hệ thống thực tế, nơi mô hình không được hiệu chỉnh hay huấn luyện lại riêng cho từng người dùng.

Về mặt hình thức, giả sử mỗi đối tượng $s \in \mathcal{S}$ được trang bị một gia tốc kế ba trục gắn tại vùng bụng trong suốt quá trình thu nhận dữ liệu. Chuỗi tín hiệu gia tốc liên tục theo thời gian của đối tượng s được biểu diễn như sau:

$$\mathbf{x}^{(s)}(t) \in \mathbb{R}^3, \quad t = 1, \dots, T_{\text{tot}}^{(s)}. \quad (2.2)$$

Chuỗi tín hiệu này sau đó được phân đoạn thành các cửa sổ có độ dài cố định T , có thể chồng lấp hoặc không chồng lấp, nhằm tạo ra các mẫu đầu vào cho mô hình

học máy. Mỗi mẫu được biểu diễn dưới dạng một ma trận:

$$\mathbf{x}_i \in \mathbb{R}^{T \times d}, \quad d = 3, \quad (2.3)$$

trong đó $\mathbf{x}_i$ tương ứng với cửa sổ thứ i của tín hiệu gia tốc kế ba trục.

Mỗi cửa sổ dữ liệu $\mathbf{x}_i$ được gán một nhãn tư thế ngủ:

$$y_i \in \mathcal{Y} = \{1, 2, \dots, 12\}, \quad (2.4)$$

tương ứng với một trong mười hai tư thế được định nghĩa trước trong tập dữ liệu, như được minh họa trong Hình 2.4. Do đó, mỗi mẫu huấn luyện bao gồm cặp dữ liệu–nhãn $(\mathbf{x}_i, y_i)$ cùng với thông tin đối tượng tương ứng.

Gọi tập huấn luyện là:

$$\mathcal{D}_{\text{train}} = \{(\mathbf{x}_i, y_i, s_i)\}_{i=1}^{\mathbf{N}_{\text{train}}}, \quad (2.5)$$

trong đó $s_i \in \mathcal{S}_{\text{train}}$ và $\mathcal{S}_{\text{train}} \subset \mathcal{S}$ là tập các đối tượng được sử dụng trong quá trình huấn luyện.

Tương tự, tập kiểm tra được ký hiệu là:

$$\mathcal{D}_{\text{test}} = \{(\mathbf{x}_j, y_j, s_j)\}_{j=1}^{\mathbf{N}_{\text{test}}}, \quad (2.6)$$

với $s_j \in \mathcal{S}_{\text{test}}$ và $\mathcal{S}_{\text{test}} \subset \mathcal{S}$.

Trong thiết lập đánh giá độc lập theo đối tượng, hai tập đối tượng dùng cho huấn luyện và kiểm tra phải thỏa mãn điều kiện:

$$\mathcal{S}_{\text{train}} \cap \mathcal{S}_{\text{test}} = \emptyset. \quad (2.7)$$

Điều kiện trong Công thức (2.7) đảm bảo rằng quá trình đánh giá phản ánh khả năng tổng quát hóa của mô hình đối với các đối tượng chưa từng xuất hiện trong quá trình huấn luyện.

Bài toán phân loại khi đó được phát biểu như việc học một hàm ánh xạ:

$$f_{\boldsymbol{\theta}} : \mathbb{R}^{T \times 3} \rightarrow \Delta^{11}, \quad (2.8)$$

trong đó f_{θ} là mô hình được tham số hóa bởi bộ tham số θ , và Δ^{12} là đơn hình xác suất tương ứng với 12 lớp tư thế ngủ.

Với một cửa sổ đầu vào $\mathbf{x}$, mô hình sinh ra một véc-tơ xác suất:

$$\hat{\mathbf{p}} = f_{\theta}(\mathbf{x}), \quad \hat{\mathbf{p}} \in \mathbb{R}^{12}, \quad \hat{p}_k \geq 0, \quad \sum_{k=1}^{12} \hat{p}_k = 1, \quad (2.9)$$

trong đó mỗi thành phần $\hat{p}_k$ biểu thị xác suất dự đoán của lớp tư thế thứ k .

Nhãn tư thế dự đoán cuối cùng được xác định bởi:

$$\hat{y} = \arg \max_{k \in \{1, \dots, 12\}} \hat{p}_k. \quad (2.10)$$

Các tham số của mô hình được học bằng cách tối thiểu hóa hàm mất mát phân loại thực nghiệm trên tập huấn luyện. Trong luận án này, hàm mất mát thường được lựa chọn là hàm cross-entropy:

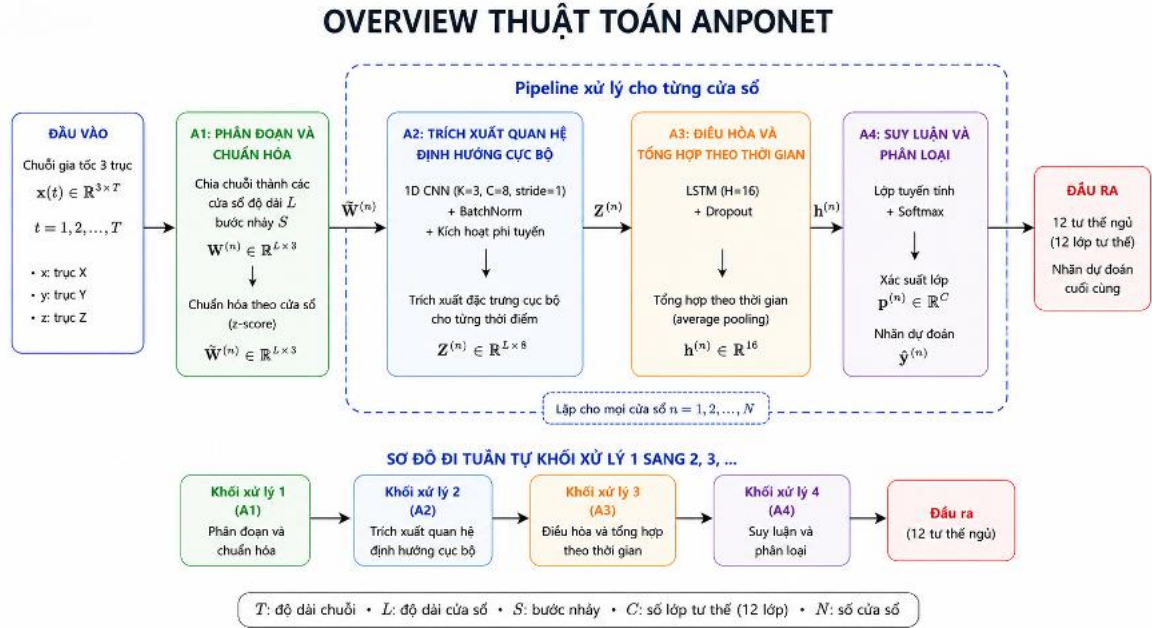
$$\theta^* = \arg \min_{\theta} \frac{1}{N_{\text{train}}} \sum_{i=1}^{N_{\text{train}}} \mathcal{L}(f_{\theta}(\mathbf{x}_i), y_i). \quad (2.11)$$

Sau khi huấn luyện, mô hình được đánh giá trên tập kiểm tra $\mathcal{D}_{\text{test}}$, chỉ bao gồm dữ liệu từ các đối tượng chưa từng xuất hiện trong quá trình học, như đã xác định trong Công thức (2.7). Do đó, việc đạt được hiệu suất phân loại cao trong thiết lập này đòi hỏi mô hình phải học được các biểu diễn đặc trưng có khả năng tổng quát hóa tốt, đồng thời đủ nhạy để phân biệt các khác biệt tinh vi giữa các tư thế ngủ trong điều kiện có sự biến thiên đáng kể giữa các đối tượng.

2.2.3 Thiết kế khung thuật toán AnpoNet

Tiểu mục này trình bày khung thuật toán AnpoNet, được đề xuất như một chiến lược mô hình hóa tổng hợp cho bài toán phân loại tư thế ngủ từ dữ liệu gia tốc ba trục trong bối cảnh tài nguyên tính toán hạn chế. Khác với các tiếp cận tập trung vào việc gia tăng độ sâu hoặc độ phức tạp của từng thành phần mạng, AnpoNet được tổ chức thành một chuỗi các khối xử lý có chức năng phân biệt rõ ràng. Mỗi khối đảm nhiệm một giai đoạn cụ thể của quá trình xử lý và được thiết kế dưới các ràng buộc về số lượng tham số cũng như chi phí tính toán.

Hình 2.6 minh họa tổng thể khung xử lý của AnpoNet, từ tín hiệu gia tốc ba



Hình 2.6: Tổng thể luồng xử lý của AnpoNet

trục đầu vào đến nhãn tư thế đầu ra. Khung xử lý gồm bốn khối chức năng chính, ký hiệu A1–A4, được tổ chức tuần tự. Trước hết, tín hiệu được phân đoạn thành các cửa sổ và chuẩn hóa nhằm hạn chế ảnh hưởng của sai lệch biên độ và sự khác biệt thống kê giữa các đoạn tín hiệu. Tiếp theo, khối trích xuất đặc trưng cục bộ khai thác các mẫu biến thiên ngắn hạn và mối quan hệ giữa các thành phần gia tốc trên ba trục. Chuỗi đặc trưng thu được sau đó được đưa vào khối mô hình hóa theo thời gian nhằm khai thác sự phụ thuộc tuần tự trong tín hiệu. Cuối cùng, biểu diễn đặc trưng ở mức cửa sổ được ánh xạ sang không gian 12 lớp tư thế thông qua khối phân loại đầu ra.

Việc phân tách chức năng giữa các khối nhằm hạn chế sự gia tăng độ phức tạp mô hình không cần thiết. Trong đó, thành phần tích chập một chiều tập trung vào khai thác các đặc trưng cục bộ đa trục, trong khi LSTM đảm nhiệm việc mô hình hóa sự phụ thuộc theo thời gian của chuỗi đặc trưng. Cách tổ chức này cho phép AnpoNet đồng thời khai thác thông tin cục bộ và đặc tính tuần tự của tín hiệu gia tốc, trong khi vẫn duy trì số lượng tham số và chi phí tính toán ở mức thấp.

Mỗi khối xử lý trong AnpoNet được thiết kế để giải quyết một khía cạnh cụ thể của bài toán. Khối xử lý A1 thực hiện phân đoạn và chuẩn hóa tín hiệu trong từng cửa sổ, tạo đầu vào có phân bố ổn định hơn cho các khối phía sau. Khối xử lý A2 trích xuất các đặc trưng cục bộ đa trục bằng tích chập một chiều với số kênh đầu ra nhỏ, qua đó tạo ra biểu diễn gọn nhẹ nhưng vẫn duy trì thông tin cần thiết để phân biệt các tư thế ngủ. Khối xử lý A3 mô hình hóa và tổng hợp thông tin theo thời gian bằng

LSTM, dropout và phép lấy trung bình theo thời gian, cho phép khai thác phụ thuộc tuần tự đồng thời hạn chế sự phụ thuộc quá mức vào các đáp ứng tức thời tại từng thời điểm. Khối xử lý A4 thực hiện ánh xạ biểu diễn thời gian sang không gian lớp và đưa ra dự đoán tư thế cuối cùng.

Điểm khác biệt cốt lõi của AnpoNet so với các mô hình học sâu có độ phức tạp cao nằm ở việc các ràng buộc tài nguyên được xem xét ngay từ giai đoạn thiết kế kiến trúc. Thay vì tối đa hóa năng lực biểu diễn của từng khối riêng lẻ, AnpoNet ưu tiên sự phù hợp giữa độ phức tạp mô hình và hàm lượng thông tin của tín hiệu đầu vào. Cách tiếp cận này hướng tới sự cân bằng giữa hiệu năng phân loại và hiệu quả sử dụng tài nguyên, phù hợp với các kịch bản xử lý trên thiết bị biên và góp phần hạn chế nguy cơ quá khớp trong thiết lập đánh giá độc lập theo đối tượng.

Các khối xử lý A1–A4 được trình bày chi tiết trong các phần tiếp theo, kèm theo quy trình tính toán và các biểu thức toán học cần thiết nhằm bảo đảm tính chặt chẽ và khả năng tái lập của phương pháp đề xuất.

Thuật toán 1 Khung xử lý AnpoNet cho phân loại tư thế ngũ hướng tới thiết bị biên

Input: Chuỗi gia tốc ba trục $\mathbf{x}(t) \in \mathbb{R}^3$

Output: Tập dự đoán tư thế theo cửa sổ

Phân đoạn và chuẩn hóa tín hiệu thành $\mathcal{W} = \{\tilde{\mathbf{W}}^{(n)}\}$ bằng Khối xử lý A1

foreach $\tilde{\mathbf{W}}^{(n)} \in \mathcal{W}$ **do**

$\mathbf{Z}^{(n)} \leftarrow$ Khối xử lý A2
 $\mathbf{h}^{(n)} \leftarrow$ Khối xử lý A3
 $(\mathbf{p}^{(n)}, \hat{y}^{(n)}) \leftarrow$ Khối xử lý A4

return $\{(\mathbf{p}^{(n)}, \hat{y}^{(n)})\}$

Khối xử lý A1: Phân đoạn và chuẩn hóa chuỗi gia tốc

Input: $\mathbf{x}(t) \in \mathbb{R}^3$, độ dài cửa sổ L , bước nhảy S . **Output:** $\mathcal{W} = \{\tilde{\mathbf{W}}^{(n)}\}_{n=1}^N$.

Số cửa sổ:

$$N = \left\lfloor \frac{T-L}{S} \right\rfloor + 1.$$

Với $n = 1, \dots, N$:

1. Xác định vị trí bắt đầu

$$s = 1 + (n-1)S.$$

2. Trích xuất cửa sổ

$$\mathbb{W}^{(n)} = [\mathbf{x}(s), \dots, \mathbf{x}(s + L - 1)].$$

3. Tính $\boldsymbol{\mu}^{(n)}$ và $\boldsymbol{\sigma}^{(n)}$ theo từng kênh, sau đó chuẩn hóa:

$$\tilde{\mathbb{W}}^{(n)} = \frac{\mathbb{W}^{(n)} - \boldsymbol{\mu}^{(n)}}{\boldsymbol{\sigma}^{(n)} + \epsilon}.$$

Return: $\mathcal{W}$.

Khối xử lý A2: Trích xuất đặc trưng cục bộ bằng 1D CNN và Batch-Norm

Input: $\tilde{\mathbb{W}} \in \mathbb{R}^{L \times 3}$. **Output:** $\mathbf{Z} \in \mathbb{R}^{L \times 8}$.

Chuyển dữ liệu sang dạng kênh-trước $\mathbf{U} \in \mathbb{R}^{3 \times L}$. Sử dụng $C = 8$ bộ lọc, kernel $K = 3$, stride = 1 và padding “same”.

Với $c = 1, \dots, 8$ và $t = 1, \dots, L$:

$$a_{c,t} = \sum_{i=1}^3 \sum_{k=-1}^1 W_{c,i,k} u_{i,t-k} + b_c.$$

Batch Normalization theo từng kênh:

$$\hat{a}_{c,t} = \frac{a_{c,t} - \mu_c}{\sqrt{\nu_c + \epsilon}}, \quad \tilde{a}_{c,t} = \gamma_c \hat{a}_{c,t} + \beta_c.$$

Áp dụng hàm kích hoạt:

$$z_{c,t} = \phi(\tilde{a}_{c,t}).$$

Đặt

$$\mathbf{Z}[t, c] = z_{c,t}, \quad \mathbf{Z} \in \mathbb{R}^{L \times 8}.$$

Return: $\mathbf{Z}$.

Khối xử lý A3: Mô hình hóa và tổng hợp đặc trưng theo thời gian bằng LSTM

Input: $\mathbf{Z} \in \mathbb{R}^{L \times 8}$, tỷ lệ dropout p . **Output:** $\mathbf{k} \in \mathbb{R}^{16}$.

Khởi tạo $\mathbf{k}_0 = \mathbf{0}$, $\mathbf{c}_0 = \mathbf{0}$. Với $t = 1, \dots, L$ và $\mathbf{z}_t = \mathbf{Z}[t]$:

$$\dot{\mathbf{i}}_t = \sigma \left(\mathbb{W}_i \mathbf{z}_t + \mathbb{U}_i \mathbf{k}_{t-1} + \mathbf{b}_i \right),$$

$$\mathbf{f}_t = \sigma \left(\mathbb{W}_f \mathbf{z}_t + \mathbb{U}_f \mathbf{k}_{t-1} + \mathbf{b}_f \right),$$

$$\mathbf{o}_t = \sigma \left(\mathbb{W}_o \mathbf{z}_t + \mathbb{U}_o \mathbf{k}_{t-1} + \mathbf{b}_o \right),$$

$$\tilde{\mathbf{c}}_t = \tanh \left(\mathbb{W}_c \mathbf{z}_t + \mathbb{U}_c \mathbf{k}_{t-1} + \mathbf{b}_c \right).$$

Cập nhật trạng thái:

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \dot{\mathbf{i}}_t \odot \tilde{\mathbf{c}}_t,$$

$$\mathbf{k}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t).$$

Trong huấn luyện, áp dụng dropout:

$$\bar{\mathbf{k}}_t = \frac{1}{1-p} \mathfrak{M}_t \odot \mathbf{k}_t, \quad \mathfrak{M}_t \sim \text{Bernoulli}(1-p).$$

Trong suy luận, $\bar{\mathbf{k}}_t = \mathbf{k}_t$.

Biểu diễn thời gian được tổng hợp bởi

$$\mathbf{k} = \frac{1}{L} \sum_{t=1}^L \bar{\mathbf{k}}_t.$$

Return: $\mathbf{k}$.

Khối xử lý A4: Suy luận và phân loại tư thế ngủ

Input: $\mathbf{k} \in \mathbb{R}^{16}$. **Output:** $\mathbf{p} \in \mathbb{R}^{12}$, $\hat{y}$.

Tính logits:

$$\mathbf{s} = \mathbb{W} \mathbf{k} + \mathbf{b}, \quad \mathbb{W} \in \mathbb{R}^{12 \times 16}.$$

Xác suất hậu nghiệm:

$$p_k = \frac{\exp(s_k)}{\sum_{j=1}^{12} \exp(s_j)}, \quad k = 1, \dots, 12.$$

Nhân dự đoán:

$$\hat{y} = \arg \max_{k \in \{1, \dots, 12\}} p_k.$$

Return: $(\mathbf{p}, \hat{y})$.

2.2.4 Bài toán huấn luyện và hàm mất mát

Cho tập huấn luyện

$$\mathcal{D}_{\text{train}} = \{(\mathbf{x}_i, y_i)\}_{i=1}^N,$$

trong đó $y_i \in \{1, \dots, 12\}$ là nhãn tư thế ngủ tương ứng với cửa sổ $\mathbf{x}_i$. Mô hình AnpoNet được ký hiệu bởi f_{θ} và ánh xạ mỗi cửa sổ đầu vào $\mathbf{x}_i$ sang véc-tơ xác suất hậu nghiệm

$$\hat{\mathbf{p}}_i = f_{\theta}(\mathbf{x}_i).$$

Quá trình huấn luyện được phát biểu dưới dạng bài toán tối ưu hóa thực nghiệm, trong đó bộ tham số θ được ước lượng bằng cách tối thiểu hóa hàm mất mát cross-entropy trung bình trên tập huấn luyện:

$$\mathcal{L}_{\text{CE}}(\theta) = -\frac{1}{N} \sum_{i=1}^N \log \hat{p}_{i, y_i}.$$

Để hạn chế hiện tượng quá khớp và kiểm soát độ lớn của các tham số mô hình, thành phần điều chuẩn ℓ_2 được bổ sung vào hàm mục tiêu:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{CE}}(\theta) + \lambda \|\theta\|_2^2,$$

trong đó $\lambda > 0$ là hệ số điều chuẩn.

Thuật toán 2 Tối ưu hóa tham số AnpoNet bằng Adam

Input: Gradient $\mathbf{g}_t = \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}_t)$, bước lặp t , tham số hiện tại $\boldsymbol{\theta}_t$

Output: Tham số cập nhật $\boldsymbol{\theta}_{t+1}$

Cập nhật moment bậc nhất:

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t$$

Cập nhật moment bậc hai:

$$\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^{\odot 2}$$

Hiệu chỉnh sai lệch:

$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{1 - \beta_1^t}, \quad \hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_2^t}$$

Cập nhật tham số:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t} + \epsilon}$$

return $\boldsymbol{\theta}_{t+1}$

2.2.5 Phân tích độ phức tạp và đặc tính tính toán của AnpoNet

Tiểu mục này phân tích AnpoNet từ góc độ độ phức tạp tính toán và số lượng tham số, qua đó làm rõ ảnh hưởng của các quyết định thiết kế đối với khả năng triển khai trong điều kiện tài nguyên hạn chế. Phân tích tập trung vào bốn khối xử lý A1–A4 và xác định thành phần chi phối chi phí tính toán của toàn bộ mô hình.

Phân tích độ phức tạp tính toán

Giả sử mỗi cửa sổ tín hiệu có độ dài L , số kênh đặc trưng đầu ra của khối tích chập là C , kích thước kernel là K , số chiều đầu vào ban đầu là $D = 3$, và số chiều trạng thái ẩn của LSTM là H . Trong cấu hình cụ thể của AnpoNet, các tham số được thiết lập là $C = 8$, $K = 3$ và $H = 16$, phản ánh chủ đích kiểm soát độ phức tạp ngay từ giai đoạn thiết kế.

Khối xử lý A1 – Phân đoạn và chuẩn hóa. Đối với mỗi cửa sổ, A1 thực hiện tính trung bình, độ lệch chuẩn và chuẩn hóa trên ba kênh gia tốc. Vì số kênh đầu vào $D = 3$ là cố định, số phép toán tỷ lệ tuyến tính với số mẫu trong cửa sổ. Do đó, độ phức tạp của A1 có thể được biểu diễn dưới dạng

$$\mathcal{O}(LD) = \mathcal{O}(L).$$

Chi phí của khối này nhỏ so với các thành phần học đặc trưng và mô hình hóa thời gian ở phía sau.

Khối xử lý A2 – Trích xuất đặc trưng cục bộ. A2 sử dụng một lớp tích chập một chiều với $D = 3$ kênh đầu vào, C kênh đầu ra và kernel có kích thước K . Với mỗi vị trí thời gian, mỗi bộ lọc thực hiện khoảng DK phép nhân-cộng. Do đó, độ phức tạp của phép tích chập trên toàn bộ cửa sổ là

$$\mathcal{O}(LDCK).$$

Vì $D = 3$ và $K = 3$ là các hằng số nhỏ, có thể rút gọn về

$$\mathcal{O}(LC).$$

Batch Normalization và hàm kích hoạt đều có độ phức tạp

$$\mathcal{O}(LC),$$

do đó không làm thay đổi bậc độ phức tạp tổng thể của A2.

Khối xử lý A3 – Mô hình hóa và tổng hợp theo thời gian. A3 sử dụng LSTM có đầu vào kích thước C và trạng thái ẩn kích thước H . Tại mỗi thời điểm, bốn phép biến đổi tương ứng với các cổng đầu vào, quên, đầu ra và trạng thái ứng viên yêu cầu các phép nhân ma trận với cả véc-tơ đầu vào và trạng thái ẩn trước đó. Vì vậy, chi phí tại mỗi bước thời gian có bậc

$$\mathcal{O}(H(C + H)).$$

Trên toàn bộ cửa sổ có độ dài L , độ phức tạp của LSTM là

$$\mathcal{O}(LH(C + H)).$$

Dropout và temporal average pooling chỉ yêu cầu các phép toán tuyến tính trên các trạng thái ẩn, với chi phí

$$\mathcal{O}(LH),$$

nên không làm thay đổi bậc độ phức tạp tổng thể. Do đó,

$$\mathcal{C}_{A3} = \mathcal{O}(LH(C + H)).$$

Trong cấu hình hiện tại, A3 là thành phần chi phối chi phí tính toán của AnpoNet. Tuy nhiên, việc sử dụng $H = 16$ giúp giữ chi phí tuyệt đối ở mức thấp.

Khối xử lý A4 – Suy luận và phân loại. A4 ánh xạ véc-tơ đặc trưng có kích thước H sang $M = 12$ lớp thông qua một lớp tuyến tính. Độ phức tạp của phép ánh xạ này là

$$\mathcal{O}(HM).$$

Phép Softmax bổ sung chi phí

$$\mathcal{O}(M).$$

Vì $M = 12$ là cố định, chi phí của A4 nhỏ và không phụ thuộc vào độ dài của số L .

Tổng hợp các thành phần, độ phức tạp tính toán của AnpoNet trên một cửa sổ có thể biểu diễn bởi

$$\mathcal{C}_{\text{AnpoNet}} = \mathcal{O}(LDCK) + \mathcal{O}(LH(C + H)) + \mathcal{O}(HM).$$

Với D, K, C, H và M được giữ cố định trong cấu hình triển khai, độ phức tạp theo độ dài của số được rút gọn thành

$$\mathcal{C}_{\text{AnpoNet}} = \mathcal{O}(L).$$

Như vậy, chi phí tính toán của AnpoNet tăng tuyến tính theo độ dài của số tín hiệu. Đặc tính này, kết hợp với số kênh tích chập và số chiều ẩn nhỏ, giúp duy trì chi phí suy luận có thể dự đoán được và phù hợp hơn với các kịch bản triển khai trên thiết

bị có tài nguyên hạn chế.

Phân tích số lượng tham số

Số lượng tham số là một trong những ràng buộc thiết kế trung tâm của AnpoNet. Khác với các mô hình học sâu có độ biểu diễn cao, trong đó số tham số thường tăng nhanh theo độ sâu và số chiều ẩn, AnpoNet được thiết kế để duy trì tổng số tham số ở mức tối thiểu cần thiết.

Cụ thể, khối trích xuất định hướng A2 với $C = 8$ kênh đầu ra và kernel $K = 3$ có tổng cộng $8 \times (3 \times 3) + 8 = 80$ tham số, bao gồm trọng số và bias. Khối mô hình hóa thời gian A3 với LSTM có $H = 16$ nút ẩn và đầu vào $C = 8$ có $4(HC + H^2 + H) = 1600$ tham số, chiếm tỷ trọng lớn nhất trong toàn bộ mô hình. Lớp phân loại A4 có $12 \times 16 + 12 = 204$ tham số. Tổng số tham số của AnpoNet là 1916, cho thấy mô hình đạt được năng lực phân loại đáng kể với một ngân sách tham số rất nhỏ so với các kiến trúc học sâu phổ biến.

Việc phân lớn tham số tập trung trong khối A3 phản ánh vai trò trung tâm của mô hình hóa thời gian trong bài toán phân loại tư thế ngủ, trong khi các khối còn lại được giữ ở mức gọn nhẹ nhằm hạn chế sự gia tăng không cần thiết của độ phức tạp.

Phân tích hành vi của các khối xử lý và tính phù hợp với thiết bị biên

Ngoài độ phức tạp tính toán và số lượng tham số, cách thức hoạt động của các khối xử lý trong AnpoNet cũng đóng vai trò quan trọng đối với hiệu quả biểu diễn và khả năng khái quát hóa của mô hình.

Khối xử lý A1 thực hiện phân đoạn và chuẩn hóa tín hiệu trong từng cửa sổ, qua đó hạn chế ảnh hưởng của sự khác biệt về biên độ và phân bố thống kê giữa các đoạn tín hiệu. Bước tiền xử lý này giúp tạo đầu vào ổn định hơn cho các khối học đặc trưng phía sau và giảm mức độ phụ thuộc của mô hình vào các sai lệch mang tính cục bộ của tín hiệu cảm biến.

Khối xử lý A2 khai thác các mẫu biến thiên cục bộ và mối quan hệ giữa các thành phần gia tốc trên ba trục bằng tích chập một chiều. Với kích thước kernel nhỏ và số kênh đầu ra được giới hạn, khối này đóng vai trò như một phép biến đổi đặc trưng gọn nhẹ, tập trung vào các cấu trúc cục bộ có giá trị phân biệt mà không làm gia tăng đáng kể độ phức tạp của mô hình.

Khối xử lý A3 đảm nhiệm việc mô hình hóa và tổng hợp thông tin theo thời

gian. Các trạng thái ẩn của LSTM cho phép tích hợp thông tin từ chuỗi đặc trưng theo thứ tự thời gian, trong khi dropout góp phần điều chuẩn mô hình trong quá trình huấn luyện. Sau đó, phép lấy trung bình theo thời gian tổng hợp chuỗi trạng thái ẩn thành một véc-tơ có số chiều cố định. Cách tổ chức này giúp duy trì kích thước biểu diễn đầu ra độc lập với độ dài chuỗi và tránh làm gia tăng số lượng tham số của khối phân loại phía sau.

Khối xử lý A4 sử dụng một phép ánh xạ tuyến tính kết hợp với Softmax để chuyển biểu diễn thời gian sang không gian 12 lớp tư thể. Việc sử dụng một bộ phân loại đầu ra có cấu trúc đơn giản dựa trên giả định rằng các khối A2 và A3 đã thực hiện phần lớn quá trình biến đổi và tổng hợp đặc trưng cần thiết. Nhờ đó, số lượng tham số tại tầng đầu ra được giữ ở mức thấp, đồng thời hạn chế sự gia tăng không cần thiết của độ phức tạp mô hình.

Xét trên toàn bộ chuỗi xử lý, AnpoNet không hướng tới việc tối đa hóa năng lực biểu diễn của từng khối riêng lẻ mà ưu tiên sự phối hợp giữa các khối có độ phức tạp thấp và chức năng bổ trợ cho nhau. Khối A1 ổn định dữ liệu đầu vào, khối A2 khai thác đặc trưng cục bộ, khối A3 mô hình hóa phụ thuộc theo thời gian và khối A4 thực hiện phân loại đầu ra. Sự phân tách chức năng này cho phép kiểm soát số lượng tham số và chi phí tính toán tại từng giai đoạn của mô hình.

Từ góc độ triển khai, việc sử dụng một lớp tích chập một chiều với số kênh nhỏ, LSTM có số chiều ẩn giới hạn và một lớp phân loại tuyến tính giúp duy trì yêu cầu bộ nhớ và số phép toán ở mức tương đối thấp. Kết hợp với đặc tính độ phức tạp tăng tuyến tính theo độ dài của số tín hiệu, cấu trúc này phù hợp với định hướng triển khai AnpoNet trên các thiết bị biên có tài nguyên tính toán và bộ nhớ hạn chế.

2.3 Thực nghiệm và đánh giá

2.3.1 Các mô hình so sánh

Trong luận án này, thuật ngữ *baseline* được sử dụng để chỉ các mô hình tham chiếu dùng cho mục đích so sánh hiệu năng với phương pháp đề xuất. Việc giữ nguyên thuật ngữ *baseline* thay vì dịch sang tiếng Việt nhằm đảm bảo tính ngắn gọn, nhất quán và phù hợp với các quy ước được sử dụng rộng rãi trong cộng đồng nghiên cứu học máy và học sâu. Các baseline được hiện thực lại dựa trên các công trình liên quan, có điều chỉnh cần thiết để phù hợp với bài toán phân loại 12 tư thế ngủ và tập dữ liệu

được sử dụng trong luận án.

Baseline ExtraTrees (Rawan et al. [62]): Baseline này được xây dựng dựa trên phương pháp ExtraTrees do Rawan et al. (2022) đề xuất cho bài toán phân loại tư thế ngủ từ dữ liệu gia tốc gắn ở cổ. Trong công trình gốc, mô hình được sử dụng để phân loại bốn tư thế ngủ. Trong luận án này, ExtraTrees được hiện thực lại theo cùng nguyên lý nhưng được mở rộng để phân loại 12 tư thế ngủ, nhằm đảm bảo sự công bằng trong so sánh với các mô hình đề xuất. Mô hình được triển khai bằng thư viện **Scikit-Learn**.

Baseline LSTM (Rawan et al. [63]): Rawan et al. (2023) đã đề xuất sử dụng mạng Long Short-Term Memory (LSTM) cho bài toán phân loại tư thế ngủ từ dữ liệu gia tốc. Kiến trúc gốc bao gồm một lớp LSTM với chiều ẩn là 25 nhằm mô hình hóa các phụ thuộc dài hạn trong chuỗi tín hiệu. Theo sau lớp LSTM là một lớp kết nối đầy đủ, trong đó số chiều đầu ra được điều chỉnh từ bốn (trong công trình gốc) lên mười hai để phù hợp với bài toán trong luận án này. Đầu ra của mô hình được xử lý bằng hàm Log-Softmax, và hàm mất mát negative log-likelihood được sử dụng để huấn luyện mô hình với bộ tối ưu hóa Adam.

Baseline LSTM (CT4): Trong nghiên cứu CT4 trước đó, bài toán phân loại tư thế ngủ được giải quyết bằng mô hình LSTM với năm nhân đầu ra. Trong luận án này, kiến trúc cơ bản của mô hình được giữ nguyên, bao gồm hai lớp LSTM và hai lớp kết nối đầy đủ. Một điều chỉnh được đưa vào là việc tích hợp các lớp Batch Normalization sau mỗi lớp LSTM và lớp kết nối đầy đủ nhằm cải thiện tính ổn định của quá trình huấn luyện và giảm hiện tượng quá khớp.

Các baseline học máy truyền thống: Để đánh giá toàn diện hiệu quả của phương pháp đề xuất, luận án sử dụng một số thuật toán học máy truyền thống làm baseline, bao gồm Decision Tree [116] và Extreme Gradient Boosting (XGBoost) [117]. Các mô hình này được triển khai bằng thư viện **Scikit-Learn** với các cấu hình phù hợp cho bài toán phân loại đa lớp.

Baseline MLP: Một mạng Multilayer Perceptron (MLP) được xây dựng với hai lớp ẩn và một lớp đầu ra. Số chiều của các lớp ẩn lần lượt là 64 và 32, trong khi lớp đầu ra có 12 chiều tương ứng với số lượng tư thế ngủ. Các lớp Batch Normalization được tích hợp trước mỗi lớp ẩn, và cơ chế Dropout với tỷ lệ 0.4 được sử dụng nhằm giảm thiểu hiện tượng quá khớp.

Baseline 1D CNN: Mô hình này bao gồm hai lớp tích chập một chiều với số

lượng bộ lọc lần lượt là 8 và 16, sử dụng kích thước kernel là 3. Các đặc trưng sau tích chập được làm phẳng và đưa qua một mạng MLP gồm một lớp ẩn 16 nơ-ron và một lớp phân loại đầu ra 12 chiều. Các lớp Batch Normalization được tích hợp sau mỗi lớp (trừ lớp Flatten và lớp phân loại), và Dropout với tỷ lệ 0.4 được sử dụng để điều chuẩn mô hình.

Baseline 1D CNN + GRU: Mô hình này là một kiến trúc lai giữa mạng tích chập một chiều và mạng hồi quy Gated Recurrent Unit (GRU) nhằm khai thác đồng thời các đặc trưng cục bộ theo thời gian và các phụ thuộc chuỗi trong dữ liệu gia tốc. Đầu vào có dạng $(seq_len, 3)$, được xử lý trước bởi một lớp tích chập một chiều với 8 bộ lọc, kích thước kernel là 5 và bước nhảy bằng 1, theo sau là một lớp chuẩn hóa. Chuỗi đặc trưng sau tích chập được đưa qua hai lớp GRU một chiều xếp chồng với số nút ẩn lần lượt là 16 và 32; mỗi lớp GRU được kết hợp với Batch Normalization và MaxPooling (kích thước 3, bước nhảy 1) để ổn định huấn luyện và giảm độ phân giải theo thời gian. Sau đó, mô hình sử dụng Global Average Pooling để tổng hợp thông tin theo toàn bộ chiều thời gian, áp dụng Dropout cho điều chuẩn, và kết thúc bằng một lớp kết nối đầy đủ 12 nút với hàm kích hoạt softmax cho bài toán phân loại 12 tư thế ngủ.

2.3.2 Thiết kế thực nghiệm

Trong các thí nghiệm, công trình này đánh giá hiệu suất của các mô hình phân loại thông qua hai chiến lược đánh giá độc lập theo đối tượng, nhằm phản ánh sát các kịch bản triển khai thực tế và giảm thiểu nguy cơ rò rỉ thông tin giữa các tập dữ liệu. Việc phân chia dữ liệu ở cấp độ đối tượng đặc biệt quan trọng trong nghiên cứu này do tín hiệu gia tốc thu được tại vùng bụng có thể chịu ảnh hưởng của đặc điểm hình thể và cách cảm biến tiếp xúc với cơ thể. Bằng cách bảo đảm dữ liệu của đối tượng kiểm thử không xuất hiện trong quá trình huấn luyện, giao thức này hạn chế nguy cơ mô hình khai thác các đặc trưng riêng của từng cá nhân và tạo ra một phép đánh giá nghiêm ngặt hơn về khả năng khái quát hóa sang người dùng mới.

Chiến lược thứ nhất là đánh giá theo một phép chia cố định dựa trên kiểm định chéo độc lập theo đối tượng. Cụ thể, tập dữ liệu gồm 15 tình nguyện viên được chia ở cấp độ đối tượng theo tỷ lệ 4:1, trong đó dữ liệu của 12 tình nguyện viên (tương đương 80%) được sử dụng cho quá trình huấn luyện và xác thực, còn dữ liệu của 3 tình nguyện viên (tương đương 20%) được giữ hoàn toàn cho tập kiểm tra. Trong tập

huấn luyện/xác thực, dữ liệu tiếp tục được chia theo tỷ lệ 8:2, tương ứng với việc sử dụng dữ liệu của khoảng 9–10 tình nguyện viên cho huấn luyện và phần còn lại cho xác thực. Cách chia này đảm bảo rằng dữ liệu của mỗi tình nguyện viên chỉ xuất hiện trong đúng một tập (huấn luyện, xác thực hoặc kiểm tra), qua đó duy trì thiết lập phân loại độc lập theo đối tượng.

Chiến lược thứ hai sử dụng phương pháp kiểm định chéo 5-fold ở cấp độ đối tượng (5-Fold Subject-Independent Cross-Validation). Trong thiết lập này, tập dữ liệu được chia thành năm fold không chồng lấp, mỗi fold bao gồm dữ liệu của ba tình nguyện viên. Tại mỗi vòng lặp, một fold (ba tình nguyện viên) được sử dụng làm tập kiểm tra, trong khi bốn fold còn lại (mười hai tình nguyện viên) được sử dụng cho huấn luyện và xác thực. Quá trình này được lặp lại cho đến khi mỗi fold đã từng đóng vai trò là tập kiểm tra đúng một lần. Hiệu suất cuối cùng của mô hình được tính bằng cách lấy trung bình kết quả trên năm vòng lặp.

Việc kết hợp hai chiến lược đánh giá này cho phép đánh giá mô hình vừa trong một kịch bản chia dữ liệu cố định, vừa trong một thiết lập kiểm định chéo toàn diện hơn. Cách tiếp cận này giúp giảm sự phụ thuộc của kết quả vào một cách chia dữ liệu cụ thể, đồng thời cung cấp cái nhìn khách quan hơn về khả năng tổng quát hóa của mô hình đối với các đối tượng chưa từng xuất hiện trong quá trình huấn luyện.

Bảng 2.5: Kích thước của các tập dữ liệu huấn luyện, xác thực và kiểm tra được xây dựng với độ dài của sổ cố định là 100 mẫu và tỷ lệ chồng lấp là 0.4.

Tư thế	Tập huấn luyện	Tập xác thực	Tập kiểm tra
Lateral Right (LR)	483	105	196
Right Up (RU)	480	120	200
Up Right (UR)	474	126	200
Up (Supine) (U)	473	127	200
Up Left (UL)	482	118	200
Left Up (LU)	476	124	200
Lateral Left (LL)	482	118	200
Left Down (LD)	479	121	250
Down Left (DL)	472	128	150
Down (Prone) (D)	478	122	200
Down Right (DR)	488	112	200
Right Down (RD)	481	119	200
Tổng	5748	1440	2396
Trung bình	479	120	199.6
Độ lệch chuẩn	4.47	6.24	20.44

Bảng 2.5 báo cáo số lượng mẫu của sổ thu được ở mỗi lớp thuộc 12 tư thế trong

tập huấn luyện, xác thực và kiểm tra, với độ dài cửa sổ là 100 mẫu và tỷ lệ chồng lấp là 0.4. Sau khi phân đoạn, các bộ dữ liệu chứa 5,748 cửa sổ huấn luyện, 1,440 cửa sổ xác thực và 2,396 cửa sổ kiểm tra, trong đó mỗi cửa sổ đại diện cho một đoạn dữ liệu gia tốc kế ba trục có độ dài cố định.

Tập huấn luyện cho thấy số lượng mẫu gần như bằng nhau cho tất cả các lớp tư thế, với trung bình khoảng 470 cửa sổ mỗi lớp và độ lệch chuẩn chỉ là 4.47. Điều này chỉ ra rằng dữ liệu huấn luyện được cân bằng tốt và không có lớp tư thế nào bị đại diện quá mức hoặc dưới mức. Một mô hình tương tự được quan sát trong tập xác thực, chứa khoảng 120 cửa sổ mỗi lớp (độ lệch chuẩn 6.4), đảm bảo rằng việc đánh giá không bị thiên lệch về bất kỳ tư thế cụ thể nào.

Để tối ưu hóa các tham số của mô hình, công trình này đã áp dụng lần lượt các cửa sổ trượt với độ dài từ 10 đến 140 mẫu. Phạm vi này được chọn để xác định kích thước cửa sổ tối ưu cho hiệu suất mô hình, vì chưa có nghiên cứu nào trước đây khám phá một cách có hệ thống các khoảng nhỏ hơn trong phạm vi này. Ngoài ra, công trình này đã điều tra tác động của tỷ lệ chồng lấp bằng cách thử nghiệm các giá trị từ 10% đến 70% độ dài cửa sổ. Kích thước cửa sổ nhỏ hơn không chỉ cải thiện khả năng phát hiện các chuyển động chi tiết của mô hình mà còn tăng cường tiềm năng cho các ứng dụng thời gian thực bằng cách giảm độ trễ.

Lựa chọn hạt giống khởi tạo ngẫu nhiên (Random Seed): Để đảm bảo độ tin cậy thống kê, mỗi thí nghiệm được thực hiện với năm hạt giống (seed) ngẫu nhiên khác nhau. Qua các lần chạy, nghiên cứu này thu được giá trị trung bình, độ lệch chuẩn của độ đo Accuracy và F1-Score.

Phương pháp tối ưu hóa. Trong nghiên cứu này, bộ tối ưu hóa Adam [118] được sử dụng cho tất cả các mô hình học sâu được giới thiệu. Cụ thể, tốc độ học ban đầu là 0.001, mỗi mô hình được huấn luyện trong 200 lượt với kích thước batch là 512.

2.4 Kết quả thực nghiệm và thảo luận

2.4.1 Phân tích hiệu suất

Bảng 2.6: So sánh hiệu suất của các mô hình trên tập dữ liệu kiểm tra

Mô hình	Accuracy (%)	F1-Score (%)
Baseline Decision Tree	81.82	78.67
Baseline XGBoost	85.34	82.94
Baseline MLP	81.25	72.96
Baseline 1D CNN	87.08	78.19
1D CNN + GRU	90.81	72.64
Baseline ExtraTrees [62]	86.50	84.68
Baseline LSTM (Rawan et al [63])	68.84	65.02
Baseline LSTM (CT4)	87.44	72.30
AnpoNet	94.67	92.94
AnpoNet w/o BN	71.11	54.88

Các kết quả trong Bảng 2.6 cho thấy AnpoNet đạt độ chính xác 94.67% và F1-Score 92.94%, cao nhất trong các mô hình được so sánh. Điểm đáng chú ý không nằm ở việc “tăng” một vài phần trăm độ chính xác, mà ở chỗ F1-Score của AnpoNet tăng đồng thời và tiệm cận độ chính xác, hàm ý rằng mô hình không chỉ dự đoán đúng nhiều mẫu hơn mà còn duy trì được sự cân bằng giữa độ chính xác và độ phủ trên nhiều lớp tư thế. Trong bối cảnh phân loại 12 tư thế với các biến thể góc gần nhau, cải thiện 3.86 điểm phần trăm về độ chính xác so với mô hình học sâu tốt thứ hai (1D CNN + GRU đạt 90.81%) thường tương ứng với việc giảm một lượng đáng kể số nhầm lẫn giữa các lớp “lân cận” về định hướng. Về mặt cơ chế, AnpoNet phù hợp với bản chất tín hiệu gia tốc khi phân loại tư thế ngủ: thành phần trọng lực mang thông tin định hướng, trong khi nhiễu ngắn hạn và dao động nhỏ cần được triệt tiêu theo thời gian. Việc kết hợp trích xuất đặc trưng cục bộ bằng tích chập 1D và mô hình hóa tính ổn định bằng LSTM tạo điều kiện để mô hình tích lũy bằng chứng theo chuỗi, thay vì ra quyết định dựa trên các biến thiên tức thời. Đây cũng là xu hướng đã được ghi nhận trong các nghiên cứu về nhận dạng tư thế ngủ bằng cảm biến đeo, nơi việc nhận dạng “fine-grained” thường đòi hỏi biểu diễn vừa nhạy với định hướng vừa ổn định theo thời gian.

Đối với nhóm các mô hình học máy, Decision Tree đạt 81.82% (F1-Score 78.67%) và XGBoost đạt 85.34% (F1-Score 82.94%). Các mô hình cây quyết định có thể mạnh

trong việc học các ranh giới phân chia phi tuyến từ đặc trưng đầu vào, và trong một số kịch bản nhận dạng tư thế thô (ví dụ 4 tư thế) chúng có thể đạt kết quả tốt. Tuy nhiên, trong bài toán 12 lớp với biến thể góc nhỏ, các mô hình này thiếu cơ chế mô hình hóa phụ thuộc theo thời gian; chúng có xu hướng “xem” mỗi cửa sổ như một điểm dữ liệu tĩnh, nên khó tích lũy bằng chứng khi tín hiệu thay đổi rất ít và bị nhiễu bởi dao động tư thế nhỏ. Thực tế, công trình của Abdulsadig (Rawan) [62] và cộng sự cho thấy các phương pháp học máy có thể đạt hiệu quả cao cho nhận dạng tư thế với cảm biến gia tốc gắn cổ trong thiết lập đơn giản hơn, nhưng bài toán fine-grained đòi hỏi mô hình phải khai thác tốt hơn cấu trúc chuỗi và ổn định theo thời gian.

Đối với MLP (81.25%, F1-Score 72.96%), khoảng cách giữa độ chính xác và F1-Score lớn hơn so với nhóm cây cho thấy mô hình có thể đang dự đoán tốt một số lớp nhưng kém đồng đều trên các lớp còn lại. Với dữ liệu cảm biến, MLP thường phụ thuộc mạnh vào cách biểu diễn đầu vào (flatten hoặc đặc trưng thủ công), nên dễ bỏ qua cấu trúc tuần tự nội tại. Khi số lớp tăng và sự khác biệt giữa các lớp nằm ở sai lệch góc nhỏ, việc thiếu cơ chế trích xuất mẫu cục bộ theo thời gian và thiếu bộ nhớ chuỗi khiến MLP khó duy trì hiệu suất ổn định trên toàn bộ hệ thống phân loại.

Nhóm mô hình học sâu tham chiếu phản ánh rõ vai trò của từng thành phần kiến trúc. 1D CNN đạt 87.08% (F1-Score 78.19%), cho thấy tích chập một chiều có khả năng học các mẫu cục bộ hữu ích từ chuỗi gia tốc và phần nào làm nổi bật thông tin định hướng. Tuy vậy, khoảng cách đáng kể giữa độ chính xác và F1-Score gợi ý rằng chỉ riêng tích chập chưa đủ để “làm trơn” quyết định theo thời gian đối với các tư thế duy trì lâu và khác nhau bởi biến thiên góc rất nhỏ. Khi bổ sung GRU (1D CNN + GRU đạt 90.81%), độ chính xác tăng lên, nghĩa là mô hình đã tận dụng được phụ thuộc thời gian tốt hơn. Tuy nhiên, F1-Score của mô hình này vẫn thấp (72.64%), cho thấy lợi ích tăng độ chính xác có thể đến từ việc dự đoán tốt hơn các lớp dễ/chiếm ưu thế, trong khi các lớp khó (các biến thể tư thế gần nhau) vẫn bị nhầm lẫn đáng kể. Đây là một hiện tượng thường gặp khi mô hình chuỗi có năng lực vừa phải: nó giúp tăng độ đúng tổng thể, nhưng nếu biểu diễn đặc trưng đầu vào chưa đủ “định hướng” hoặc cơ chế điều chuẩn chưa đủ mạnh, sai số vẫn tập trung ở các lớp khó và làm F1-Score giảm. Việc lựa chọn GRU cũng có cơ sở trong văn liệu: GRU thường cho hiệu năng tương đương LSTM trên nhiều bài toán chuỗi nhưng với cấu trúc công đơn giản hơn, do đó tiết kiệm tham số và chi phí tính toán.

Mô hình ExtraTrees (Rawan et al.) đạt 86.50% và F1-Score 84.68%, với F1-

Score tương đối sát độ chính xác. Điều này cho thấy mô hình cây ngẫu nhiên có thể phân loại khá đồng đều khi ranh giới giữa các lớp đủ rõ theo không gian đặc trưng. Tuy nhiên, so với AnpoNet, hạn chế cốt lõi vẫn là thiếu cơ chế biểu diễn chuỗi theo thời gian ở mức mô hình (thay vì trông chờ vào việc cửa sổ hóa “tự mang” thông tin). Trong khi đó, AnpoNet dùng LSTM để tích lũy bằng chứng ổn định theo thời gian, nên có lợi thế rõ rệt ở các tư thế có sai lệch góc nhỏ và dễ nhiễu.

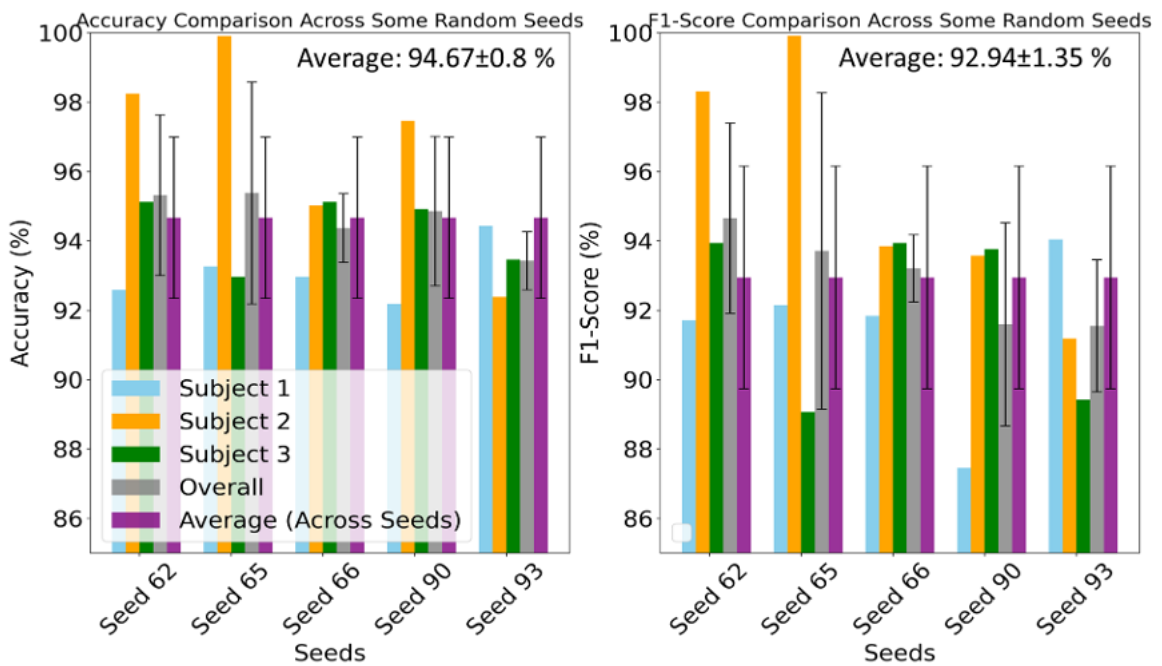
Hai baseline LSTM cho thấy một thông điệp quan trọng: hiệu năng không chỉ phụ thuộc vào việc “có RNN”, mà phụ thuộc vào tính phù hợp giữa kiến trúc, dữ liệu và cách điều chuẩn. LSTM (Rawan et al.) có độ chính xác 68.84% và F1-Score 65.02% thấp hơn đáng kể so với các mô hình khác trong bảng. Một lý do hợp lý là mô hình này được đề xuất trong bối cảnh cảm biến gắn cổ và nhiệm vụ ban đầu đơn giản hơn, nên khi chuyển sang dữ liệu/thiết lập khác và mở rộng số lớp, các giả định kiến trúc (ví dụ cách lấy đặc trưng, cách tổng hợp theo thời gian, cấu hình lớp cuối) có thể không còn phù hợp tối ưu. Ngược lại, LSTM (CT4) đạt 87.44% (F1-Score 72.30%), cho thấy việc dùng RNN sâu hơn có thể cải thiện độ chính xác tổng thể, nhưng vẫn chưa giải quyết tốt tính đồng đều theo lớp, thể hiện qua F1-Score còn thấp. Điều này củng cố lập luận rằng với bài toán tư thế ngủ fine-grained, cần một thiết kế chú trọng đồng thời: (i) trích xuất định hướng cục bộ ổn định, (ii) mô hình hóa sự ổn định theo thời gian, và (iii) điều chuẩn phân bố đặc trưng để tối ưu hóa ổn định.

Kết quả loại bỏ Batch Normalization cung cấp bằng chứng trực tiếp về vai trò của điều chuẩn trong AnpoNet. Khi bỏ Batch Normalization, độ chính xác giảm xuống 71.11% và F1-Score giảm mạnh xuống 54.88%. Mức suy giảm lớn như vậy cho thấy mô hình trở nên khó tối ưu và kém tổng quát hóa, thay vì chỉ “giảm nhẹ” do mất một kỹ thuật điều chuẩn phụ trợ. Về mặt lý thuyết và thực nghiệm, Batch Normalization được chứng minh giúp ổn định phân bố kích hoạt trong quá trình huấn luyện, cho phép tối ưu hóa với tốc độ học lớn hơn, giảm nhạy với khởi tạo và có thể đóng vai trò điều chuẩn. Trong bối cảnh dữ liệu cảm biến nhiều nhiễu và khác biệt giữa đối tượng, cơ chế này đặc biệt quan trọng vì nó giúp giảm dao động gradient và hạn chế sự lệch phân bố đặc trưng giữa các mini-batch, qua đó cải thiện hiệu năng khi đánh giá trên đối tượng chưa thấy.

Tổng hợp các luận giải trên, có thể rút ra rằng ưu thế của AnpoNet không đến từ việc tăng độ phức tạp kiến trúc, mà đến từ sự phù hợp giữa cấu trúc mô hình và đặc điểm thông tin của tín hiệu gia tốc khi phân loại tư thế ngủ fine-grained. AnpoNet

tận dụng tích chập một chiều để học đặc trưng định hướng gọn nhẹ từ chuỗi ba trục, dùng LSTM để tích lũy bằng chứng theo thời gian nhằm giảm nhiễu lẫn giữa các góc gần nhau, và dùng Batch Normalization để ổn định tối ưu hóa trong thiết lập độc lập theo đối tượng. Vì vậy, kết quả thực nghiệm trong Bảng 2.6 có thể được xem là hệ quả hợp lý của các quyết định thiết kế đã nêu trong Chương 2, đồng thời củng cố định hướng xây dựng mô hình gọn nhẹ nhưng có cơ chế hóa phù hợp để triển khai trong các kịch bản tài nguyên hạn chế.

2.4.2 Đánh giá dưới các hạt giống ngẫu nhiên khác nhau



Hình 2.7: Độ chính xác và F1-Score của ba đối tượng trong tập dữ liệu kiểm tra qua các hạt giống ngẫu nhiên khác nhau.

Hình 2.7 đánh giá độ nhạy của AnpoNet đối với các khởi tạo ngẫu nhiên tham số qua nhiều hạt giống. Kết quả chỉ ra rằng hiệu suất của mô hình duy trì độ ổn định cao trước các thay đổi ngẫu nhiên khi khởi tạo tham số ban đầu. Qua tất cả các hạt giống, độ chính xác trung bình là $94.67\% \pm 0.80\%$ và F1-Score trung bình là $92.94\% \pm 1.35\%$. Khi biểu diễn dưới dạng hệ số biến thiên, các giá trị này tương ứng với 0.85% (Accuracy) và 1.45% (F1), đều đủ thấp để cho thấy mức độ phân tán quanh giá trị trung bình là không đáng kể. Do đó, không có cấu hình hạt giống nào tạo ra hiệu suất chênh lệch đáng kể so với xu hướng chung.

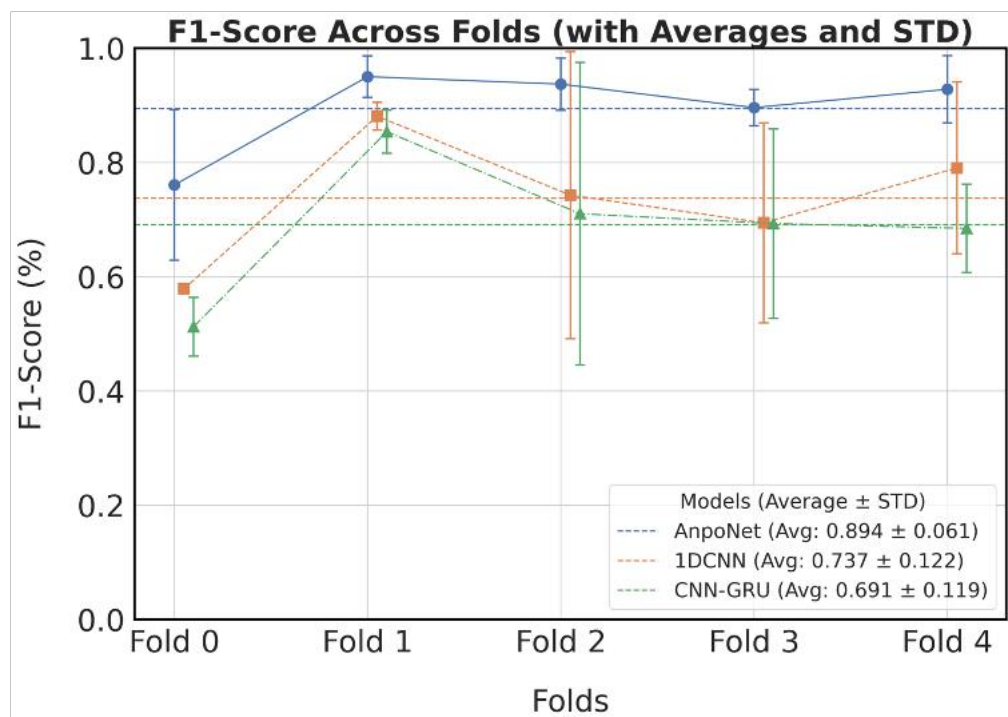
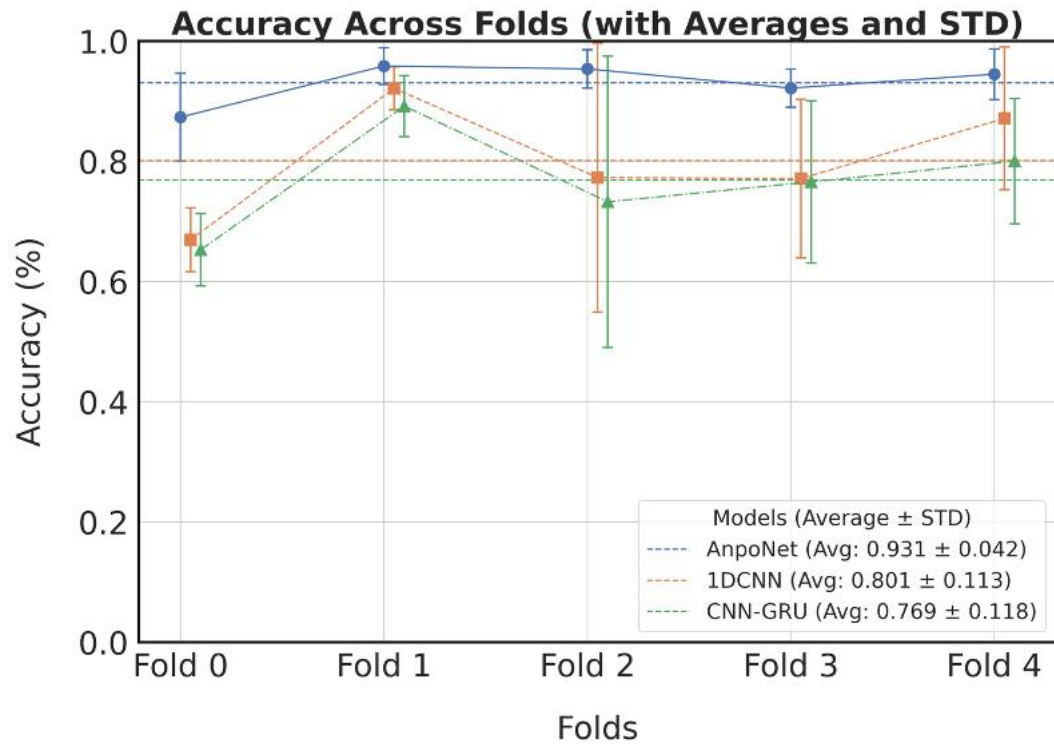
Mặc dù các hạt giống riêng lẻ tạo ra các giá trị cao hơn hoặc thấp hơn một

chút (ví dụ: Seed 65 luôn mang lại kết quả cao nhất, trong khi Seed 93 thấp hơn một chút), biên độ quan sát được vẫn hẹp, thường ở mức 2–3 phần trăm. Biên độ hạn chế này cho thấy động lực học của mô hình không phụ thuộc mạnh vào phân phối trọng số ban đầu. Các thanh tổng hợp trong Hình 2.7, mô tả mức trung bình qua tất cả các hạt giống cùng với độ lệch chuẩn, củng cố thêm rằng nhiều khởi tạo không làm nhiều loạn đáng kể các chỉ số hiệu suất toàn cục.

Các hành vi của đối tượng theo từng hạt giống thể hiện những xu hướng hợp lý thay vì các hiện tượng bất thường. Một đối tượng đo đạt độ chính xác và F1-Score cao nhất nhưng cho thấy mức độ dao động lớn hơn, có khả năng do các thay đổi góc rõ ràng hơn hoặc sự khác biệt trong cách thực hiện tư thế, vốn tương tác mạnh với các bộ lọc tích chập ban đầu. Một đối tượng đo khác hiển thị hiệu suất cao nhất quán với độ phân tán giảm, phản ánh các mẫu tín hiệu ổn định hơn. Đối tượng thứ ba duy trì điểm số vừa phải nhưng rất ổn định, chỉ ra rằng ngay cả những cá nhân ít đặc biệt hơn cũng không gây ra kết quả học tập biến động. Những quan sát này phù hợp với sự khác nhau liên quan đến thể trạng cơ thể giữa các đối tượng, các chuyển động nhỏ do hô hấp và các sai lệch nhỏ về vị trí cảm biến trên bụng.

Từ góc độ kiến trúc, sự bền vững này có thể được quy cho ba đặc tính cấu trúc của AnpoNet: (i) thành phần 1D CNN tạo ra các đặc trưng bất biến cục bộ giúp giảm thiểu tác động của các dao động tín hiệu phụ thuộc pha hoặc góc; (ii) LSTM biểu diễn hiệu quả các phụ thuộc thời gian, làm mịn các nhiễu do hạt giống gây ra trong các biểu diễn đặc trưng thấp hơn; và (iii) chuẩn hóa và dropout hoạt động như các cơ chế ổn định, ngăn chặn các khởi tạo bất lợi lan truyền qua mạng. Ngoài ra, tính chất gọn nhẹ của mô hình (1,916 tham số) hạn chế quá khớp, làm cho nó vốn ít nhạy cảm hơn với các biến thể ngẫu nhiên trong động lực huấn luyện ban đầu. Tóm lại, các yếu tố này chỉ ra rằng hiệu suất mạnh mẽ của AnpoNet không phải là kết quả của việc khởi tạo thuận lợi, mà là một đặc tính của kiến trúc mô hình và sự phù hợp của nó với dữ liệu tư thế dựa trên gia tốc kế.

2.4.3 Đánh giá dựa trên phương pháp K-Folds

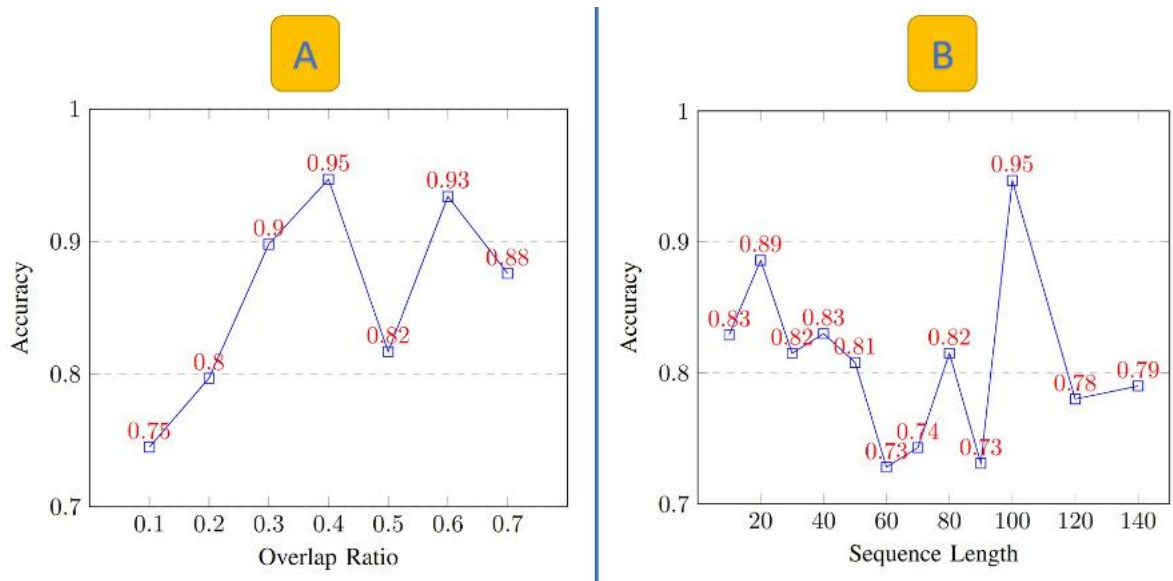


Hình 2.8: Độ chính xác và F1-Score của AnpoNet trong tập dữ liệu kiểm tra qua các fold khác nhau.

Hình 2.8 cung cấp một phân tích hiệu suất so sánh của các mô hình AnpoNet, 1D CNN và CNN-GRU sử dụng kiểm định chéo 5-fold trên thước đo độ chính xác, chứng minh sự vượt trội rõ ràng của kiến trúc AnpoNet được đề xuất. Nhìn chung, AnpoNet đạt được độ chính xác trung bình cao nhất là 0.931 ± 0.042 , vượt trội đáng kể so với cả 1D CNN (0.801 ± 0.113) và CNN-GRU (0.769 ± 0.118). Quan trọng hơn, phân tích không chỉ tập trung vào hiệu năng trung bình mà khi xét đến độ ổn định của mô hình, AnpoNet cho thấy sự bền vững đặc biệt, được chứng minh bởi độ lệch chuẩn thấp hơn đáng kể (0.042). Điều này chỉ ra hiệu suất rất nhất quán qua các phân vùng dữ liệu khác nhau, không giống như các mô hình 1D CNN và CNN-GRU thể hiện độ lệch chuẩn lớn hơn nhiều (lần lượt là 0.113 và 0.118) và các dao động hiệu suất đáng kể giữa các fold, cho thấy sự nhạy cảm của chúng đối với các tập huấn luyện và xác thực cụ thể. Do đó, biểu đồ này cung cấp bằng chứng thực nghiệm thuyết phục rằng AnpoNet không chỉ là mô hình chính xác nhất mà còn ổn định và đáng tin cậy nhất, sở hữu khả năng tổng quát hóa mạnh mẽ quan trọng cho các ứng dụng thực tế.

Hình 2.8 mở rộng phân tích so sánh bằng cách đánh giá các mô hình AnpoNet, 1D CNN và CNN-GRU dựa trên F1-Score qua kiểm định chéo 5-fold, củng cố thêm sự vượt trội của kiến trúc AnpoNet. Kết quả cho thấy AnpoNet đạt được F1-Score trung bình cao nhất là 0.894 ± 0.061 , chứng minh sự cân bằng vượt trội giữa độ chính xác (precision) và độ phủ so với 1D CNN (0.737 ± 0.122) và CNN-GRU (0.691 ± 0.119). Hơn nữa, AnpoNet thể hiện sự ổn định đáng chú ý với độ lệch chuẩn (0.061) chỉ bằng khoảng một nửa so với các mô hình khác, chỉ ra hiệu suất mạnh mẽ và nhất quán của nó bất kể phân vùng dữ liệu. Ngược lại hoàn toàn, phương sai cao và các sụt giảm hiệu suất đáng kể được quan sát thấy đối với 1D CNN và CNN-GRU qua các fold khác nhau nhấn mạnh sự không ổn định và độ nhạy của chúng đối với phân phối dữ liệu cơ bản. Do đó, phân tích F1-Score này cung cấp bằng chứng mạnh mẽ rằng AnpoNet không chỉ có hiệu suất cao nhất mà còn là mô hình ổn định và đáng tin cậy nhất, khiến nó trở thành lựa chọn hiệu quả nhất cho nhiệm vụ phân loại này.

2.4.4 Ảnh hưởng của độ dài chuỗi, tỷ lệ chồng lấp đến hiệu suất mô hình



Hình 2.9: A. Độ chính xác của AnpoNet với tỷ lệ chồng lấp thay đổi; B. Độ chính xác của AnpoNet với độ dài chuỗi thay đổi

Hình 2.9A minh họa kết quả của một thí nghiệm tinh chỉnh siêu tham số quan trọng, nhằm phân tích ảnh hưởng của tỷ lệ chồng lấp cửa sổ đến độ chính xác phân loại của mô hình AnpoNet khi áp dụng cho dữ liệu gia tốc kế ba trục được phân đoạn bằng cửa sổ trượt. Kết quả cho thấy mối quan hệ giữa tỷ lệ chồng lấp và hiệu suất mô hình là không đơn điệu, phản ánh sự đánh đổi phức tạp giữa lượng thông tin ngữ cảnh và mức độ dư thừa dữ liệu.

Cụ thể, trong giai đoạn đầu, độ chính xác thể hiện mối tương quan dương rõ rệt với tỷ lệ chồng lấp, tăng từ 0.75 tại mức chồng lấp 0.1 lên giá trị cực đại 0.95 tại mức 0.4. Quan sát này cho thấy rằng tỷ lệ chồng lấp 40% cung cấp đủ thông tin ngữ cảnh liên cửa sổ, cho phép mô hình học hiệu quả các đặc trưng ổn định theo thời gian—đặc biệt là các đặc trưng có thể trải dài qua ranh giới giữa các cửa sổ liền kề.

Tuy nhiên, khi tiếp tục tăng tỷ lệ chồng lấp vượt quá ngưỡng này, hiệu suất mô hình suy giảm đáng kể, thể hiện qua việc độ chính xác giảm xuống 0.82 tại mức chồng lấp 0.5. Mặc dù độ chính xác có sự phục hồi cục bộ lên 0.93 tại mức 0.6, xu hướng chung cho thấy việc chồng lấp quá mức làm gia tăng sự dư thừa dữ liệu, dẫn đến các mẫu huấn luyện có tính tương quan cao và làm suy giảm khả năng tổng quát hóa của mô hình. Do đó, tỷ lệ chồng lấp 0.4 được lựa chọn như một giá trị tối ưu, phản ánh sự cân bằng hợp lý giữa việc bảo toàn thông tin thời gian và kiểm soát độ dư thừa trong

dữ liệu đầu vào.

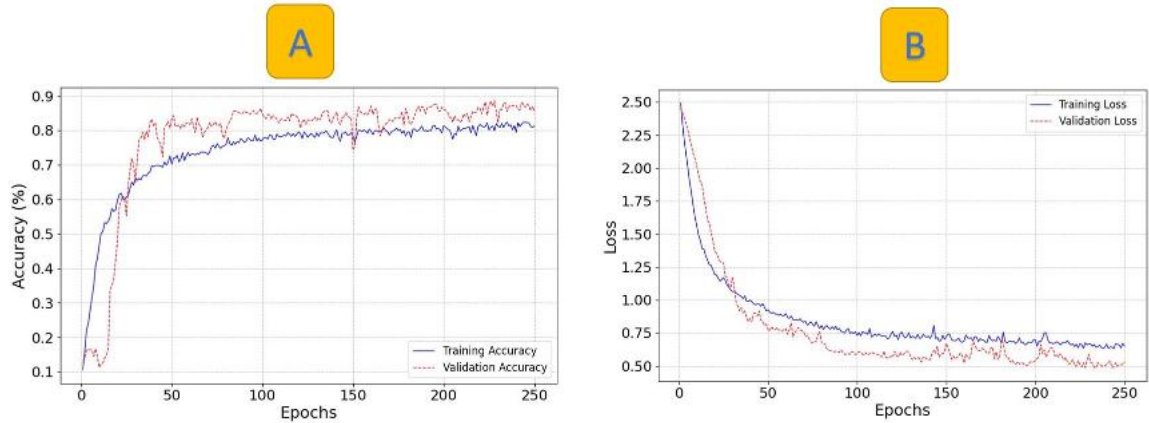
Hình 2.9.B trình bày kết quả tinh chỉnh độ dài chuỗi đầu vào, qua đó làm rõ ảnh hưởng phi tuyến và biến động mạnh của siêu tham số này đến hiệu suất phân loại của AnpoNet. Độ chính xác quan sát được dao động đáng kể trên phạm vi độ dài chuỗi được khảo sát, với sự xuất hiện của nhiều đỉnh và đáy cục bộ. Chẳng hạn, một giá trị cực đại cục bộ 0.89 đạt được tại độ dài chuỗi khoảng 20, sau đó hiệu suất suy giảm xuống 0.73 tại độ dài khoảng 60.

Kết quả quan trọng nhất là việc xác định độ dài chuỗi tối ưu tại $L = 100$, nơi mô hình đạt được độ chính xác cao nhất là 0.95. Vượt quá giá trị này, hiệu suất suy giảm rõ rệt, với độ chính xác giảm xuống 0.78 tại độ dài chuỗi 120. Hiện tượng này cho thấy rằng các chuỗi đầu vào quá dài có thể bao gồm các đoạn tín hiệu dư thừa hoặc nhiễu, làm loãng các đặc trưng phân biệt và gây khó khăn cho quá trình mô hình hóa phụ thuộc thời gian.

Những kết quả trên nhấn mạnh rằng độ dài chuỗi là một siêu tham số đặc biệt nhạy cảm trong bài toán phân loại tư thế ngủ từ dữ liệu gia tốc kế. Độ dài chuỗi $L = 100$ dường như phù hợp với thang thời gian đặc trưng của các mẫu ổn định trong tư thế ngủ, cho phép mô hình nắm bắt đầy đủ cấu trúc thời gian có ý nghĩa mà không đưa vào các thông tin không cần thiết. Ngược lại, các chuỗi ngắn hơn có nguy cơ cắt ngắn thông tin thời gian quan trọng, trong khi các chuỗi dài hơn làm gia tăng độ phức tạp và ảnh hưởng tiêu cực đến khả năng tổng quát hóa.

Tương tự, các kết quả tinh chỉnh này cho thấy rằng việc lựa chọn các siêu tham số trong giai đoạn tiền xử lý và phân đoạn tín hiệu đóng vai trò then chốt đối với hiệu suất cuối cùng của mô hình, đặc biệt trong các thiết lập hướng tới thiết bị biên, nơi yêu cầu cân bằng chặt chẽ giữa độ chính xác và hiệu quả tính toán.

2.4.5 Phân tích quá trình huấn luyện



Hình 2.10: Các đường cong huấn luyện: độ chính xác và giá trị hàm mất mát trên các tập dữ liệu huấn luyện (Training) và xác thực (Validation).

Hình 2.10 minh họa quá trình huấn luyện của mô hình. Cả hai đường cong đều thể hiện quỹ đạo đi xuống nhất quán, chỉ ra rằng mô hình đang tối ưu hóa hiệu quả các tham số của nó và dần dần học được các biểu diễn có ý nghĩa từ dữ liệu đầu vào. Giá trị mất mát huấn luyện hội tụ mượt mà, phản ánh quá trình lan truyền gradient và cập nhật tham số ổn định mà không xuất hiện dao động đáng kể, điều này là dấu hiệu của tốc độ học phù hợp và quy trình tối ưu hóa được điều kiện tốt.

Tuy nhiên, ta quan sát thấy một khoảng cách tương đối giữa mất mát huấn luyện và xác thực trong suốt quá trình huấn luyện. Sự khác biệt này có thể diễn tả một sự quá khớp, mô hình học thuộc các mẫu cụ thể trong tập huấn luyện mà không hoàn toàn tổng quát hóa được cho các mẫu xác thực chưa từng thấy. Hiện tượng này khá phổ biến trong các mô hình học sâu được huấn luyện trên dữ liệu cảm biến đặc thù miền, nơi các mẫu nhiễu nhỏ và sự dao động nội lớp có thể vô tình được học. Mặc dù khoảng cách không quá lớn, sự hiện diện của nó nhấn mạnh tầm quan trọng của việc sử dụng các chiến lược điều chuẩn—như dropout, batch normalization và tăng cường dữ liệu—để duy trì khả năng tổng quát hóa mạnh mẽ.

Hình 2.10.A mô tả các đường cong độ chính xác huấn luyện và xác thực tương ứng. Độ chính xác cải thiện mạnh mẽ trong giai đoạn huấn luyện ban đầu (trong vòng 50 epoch đầu tiên), chỉ ra rằng mô hình nhanh chóng học được các đặc trưng phân biệt cấp cao. Tiếp theo là giai đoạn cải thiện độ chính xác chậm hơn, với cả hai đường cong dần hội tụ về điểm tối ưu. Độ chính xác xác thực đạt đỉnh vào khoảng epoch 150, sau đó nó đi ngang và có dấu hiệu dao động nhỏ, có thể do các biến thể ngẫu nhiên

Bảng 2.7: Độ phức tạp mô hình và hiệu quả tính toán

Mô hình	#Số lượng tham số	#FLOPs (M)	Thời gian suy luận (s)	Bộ nhớ tiêu thụ (kB)
MLP	22,124	0.0433	0.304	116
1D CNN	25,436	0.1410	0.204	135
1D CNN + GRU	16,700	1.4633	1.481	134
LSTM (Rawan et al.)	3,212	0.2806	0.404	31
LSTM (CT4)	27,964	0.2548	0.897	154
AnpoNet (đề xuất)	1,916	0.1603	0.081	34

trong quá trình tối ưu hóa hoặc nhiều trong tập xác thực.

Từ góc độ tinh chỉnh siêu tham số, các kết quả này làm nổi bật sự cần thiết của việc lựa chọn epoch cẩn thận để ngăn chặn việc huấn luyện quá mức. Huấn luyện kéo dài vượt quá epoch tối ưu có thể dẫn đến mức độ cải thiện hiệu suất giảm dần hoặc thậm chí suy giảm độ chính xác, đặc biệt nếu quá khớp gia tăng. Việc xác định epoch 150 là điểm có độ chính xác xác thực cao nhất gợi ý rằng dừng sớm (early stopping) có thể là một cơ chế phù hợp để tối ưu hóa hiệu suất mô hình trong khi giảm thiểu chi phí tính toán của các epoch không cần thiết.

Nhìn chung, các đường cong huấn luyện trong hình 2.10 cung cấp bằng chứng mạnh mẽ về việc học hiệu quả, nhưng cũng nhấn mạnh sự cân bằng tinh tế giữa việc tối đa hóa độ chính xác và đảm bảo khả năng tổng quát hóa của mô hình. Các xu hướng quan sát được cho thấy sự cần thiết của việc áp dụng các tiêu chí dừng thích ứng và điều chuẩn một cách có mục tiêu vào quy trình huấn luyện của AnpoNet cho việc triển khai thực tế.

2.4.6 Phân tích độ phức tạp mô hình và hiệu quả tính toán

Bảng 2.7 trình bày phân tích so sánh về độ phức tạp kiến trúc và hiệu quả sử dụng tài nguyên của các mô hình được đánh giá. Các chỉ số như số lượng tham số, FLOPs, thời gian suy luận, dung lượng bộ nhớ có ý nghĩa đặc biệt quan trọng trong bối cảnh triển khai trên thiết bị biên và thiết bị đeo — nơi tài nguyên tính toán, bộ nhớ và pin đều bị giới hạn nghiêm ngặt.

Về kích thước mô hình, AnpoNet là kiến trúc gọn nhẹ nhất với chỉ 1,916 tham số, thấp hơn đáng kể so với các mô hình còn lại. Nhờ đó, bộ nhớ yêu cầu chỉ 34 kB — một mức rất phù hợp với các vi điều khiển hoặc nền tảng edge có dung lượng lưu trữ hạn chế. Đặc tính này giúp giảm áp lực lên bộ nhớ RAM và Flash, vốn là rào cản chính trong triển khai AI nhúng.

Xét về chi phí tính toán, AnpoNet duy trì mức FLOPs 0.1603M — thấp hơn đáng kể so với kiến trúc lai 1D CNN + GRU (1.4633M FLOPs) và ở mức tương đương các CNN nhẹ. Quan trọng hơn, mô hình đạt thời gian suy luận nhanh nhất (0.081 s), cho thấy hiệu quả thực thi cao trên phần cứng hạn chế. Đối với hệ thống edge thời gian thực, độ trễ thấp là yếu tố then chốt nhằm đảm bảo phản hồi tức thời và giảm nguy cơ nghẽn xử lý.

Đáng chú ý, mặc dù một số mô hình như MLP có FLOPs thấp hơn, chúng lại yêu cầu số lượng tham số và bộ nhớ lớn hơn, đồng thời không đạt được thời gian suy luận tối ưu. Điều này cho thấy việc tối ưu hóa cho edge không chỉ là giảm một chỉ số đơn lẻ mà là cân bằng đồng thời kích thước mô hình, chi phí tính toán.

Tổng thể, các kết quả trong Bảng 2.7 cho thấy AnpoNet được thiết kế có chủ đích hướng tới triển khai trên thiết bị biên: nhỏ gọn, tiêu thụ năng lượng thấp và độ trễ suy luận tối thiểu. Sự kết hợp này làm cho mô hình đặc biệt phù hợp với các ứng dụng AI nhúng và cảm biến đeo được trong thực tế, nơi yêu cầu vận hành ổn định, phản hồi nhanh là những tiêu chí cốt lõi.

2.4.7 So sánh hiệu suất với các mô hình tương đương về kiến trúc

Bảng 2.8 trình bày so sánh hiệu suất trực tiếp giữa mô hình AnpoNet đề xuất và một mô hình LSTM nhẹ (Small LSTM) được thiết kế với số lượng tham số tương đương để đảm bảo đánh giá công bằng về hiệu quả kiến trúc. Mặc dù kích thước mô hình tương tự, AnpoNet vượt trội đáng kể so với mô hình Small LSTM, đạt độ chính xác 94.67% và F1-Score 92.94%, so với 84.28% và 61.91%, tương ứng, cho Small LSTM.

Sự chênh lệch F1-Score đáng kể (+31.03% cải thiện tuyệt đối) chỉ ra rằng AnpoNet không chỉ đạt được khả năng phân loại đúng tổng thể cao hơn mà còn mang lại hiệu suất cân bằng hơn trên tất cả các lớp, giảm nguy cơ thiên lệch về phía các lớp chiếm đa số—một hạn chế thường thấy trong các kiến trúc hồi quy với dung lượng hạn chế. Sự cải tiến này có thể được giải thích bởi việc tối ưu hóa trích xuất đặc trưng tích chập và các cơ chế dựa trên chú ý của AnpoNet, giúp nắm bắt hiệu quả cả các phụ thuộc thời gian cục bộ và toàn cục trong tín hiệu gia tốc kế, mà không gặp phải các vấn đề triệt tiêu gradient và nút thắt tuần tự thường gặp phải bởi các mạng LSTM.

Quan trọng hơn, so sánh này tách biệt tác động của thiết kế kiến trúc khỏi tác động của dung lượng mô hình, chứng minh rằng những lợi ích hiệu suất của AnpoNet

Bảng 2.8: So sánh AnpoNet và LSTM với số lượng tham số tương tự (gọi tắt là Small LSTM)

Model	Accuracy (%)	F1-Score (%)
AnpoNet	94.67	92.94
Small LSTM	84.28	61.91

không chỉ đơn thuần là kết quả của việc tăng tham số mà bắt nguồn từ một cấu trúc liên kết mạng hiệu quả hơn và phù hợp với bài toán hơn. Những phát hiện này xác nhận thêm sự phù hợp của AnpoNet cho việc triển khai trong các kịch bản hạn chế về tính toán, nơi cả sự nhỏ gọn và độ tin cậy dự đoán đều là thiết yếu.

2.4.8 Tác động của nhiễu trắng đến độ bền vững của mô hình

Bảng 2.9 đánh giá độ bền vững của mô hình AnpoNet đề xuất trước sự tác động của nhiễu trắng được bổ sung vào tín hiệu gia tốc kế đầu vào. Trong điều kiện không có nhiễu, AnpoNet đạt độ chính xác 94.67% và F1-Score 92.94%. Khi tín hiệu đầu vào bị nhiễu trắng, độ chính xác của mô hình giảm nhẹ xuống 92.44%, tương ứng với mức suy giảm tương đối 2.35%, trong khi F1-Score giảm mạnh hơn xuống 80.20% (mức suy giảm tương đối 13.69%).

Sự suy giảm lớn hơn của F1-Score so với độ chính xác tổng thể cho thấy nhiễu trắng chủ yếu ảnh hưởng đến khả năng duy trì hiệu suất phân loại cân bằng giữa các lớp. Cụ thể, nhiễu có xu hướng làm gia tăng tỷ lệ phân loại sai đối với các lớp có số lượng mẫu hạn chế hoặc các lớp có đặc trưng động học tương đồng, trong khi các lớp chiếm ưu thế vẫn được nhận dạng tương đối ổn định. Hiện tượng này phù hợp với bản chất của dữ liệu gia tốc kế trong bài toán phân loại tư thế ngủ, nơi các khác biệt phân biệt thường thể hiện qua các đặc trưng thời gian tinh tế và các biến đổi biên độ nhỏ.

Mặc dù hiệu suất bị suy giảm, AnpoNet vẫn duy trì độ chính xác ở mức cao, phản ánh một mức độ bền vững nội tại trước nhiễu của kiến trúc đề xuất. Đặc tính này có thể được lý giải bởi vai trò của khối tích chập một chiều ở giai đoạn đầu, vốn hoạt động như một bộ lọc cục bộ giúp làm trơn tín hiệu và giảm ảnh hưởng của nhiễu cao tần. Bên cạnh đó, cơ chế tổng hợp theo thời gian ở các khối phía sau giúp làm nổi bật các đặc trưng ổn định và triệt tiêu các dao động ngẫu nhiên ngắn hạn do nhiễu gây ra.

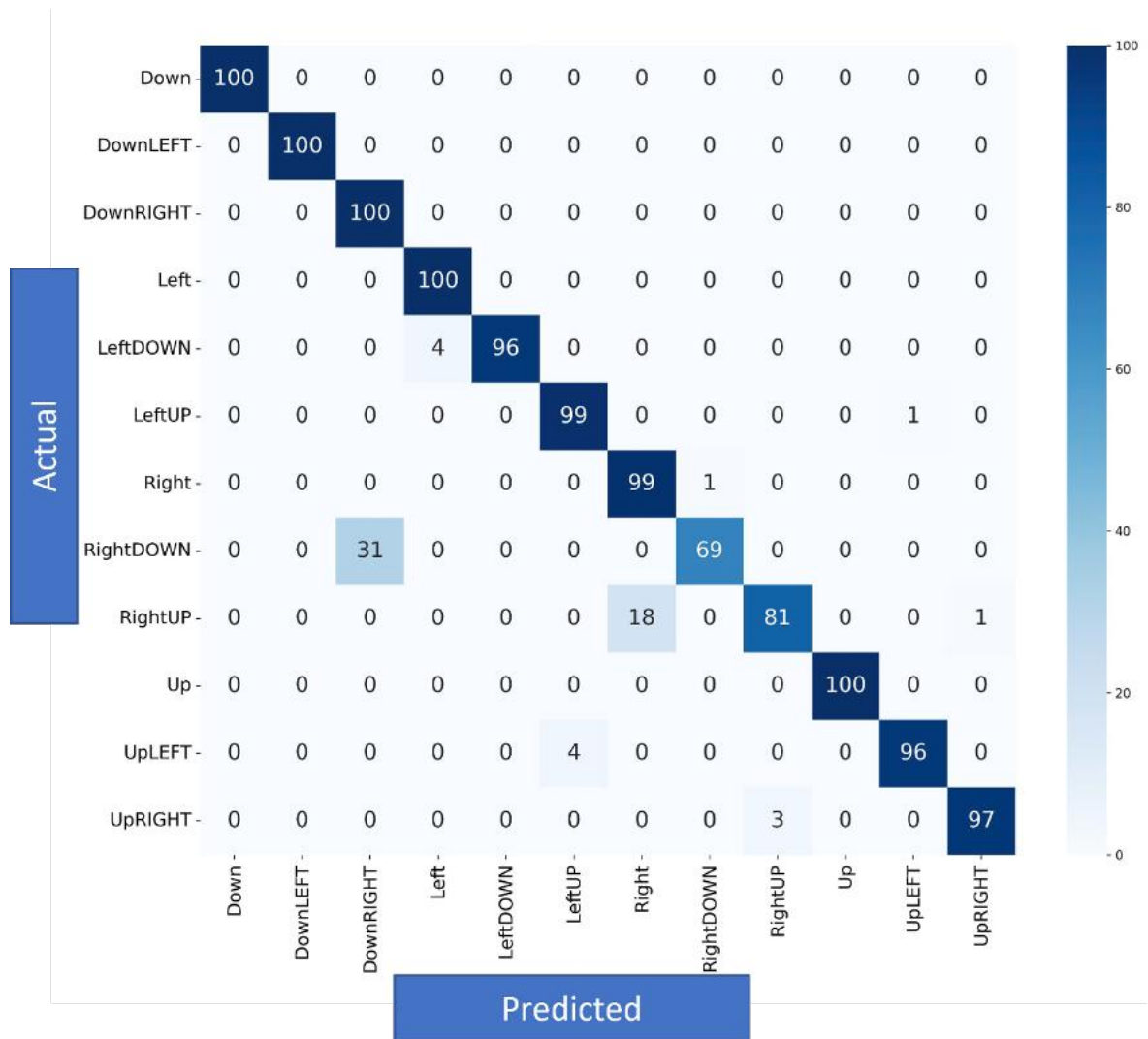
Tuy nhiên, sự suy giảm đáng kể về F1-Score cũng cho thấy rằng khả năng phân biệt các lớp khó vẫn còn nhạy cảm với nhiễu. Điều này gợi mở tiềm năng của việc tích

Bảng 2.9: So sánh AnpoNet và AnpoNet với Nhiễu Trắng

Model	Accuracy (%)	F1-Score (%)
AnpoNet	94.67	92.94
AnpoNet with Noise	92.44	80.20

hợp các chiến lược huấn luyện tăng cường độ bền vững, chẳng hạn như tăng cường dữ liệu với nhiễu tổng hợp hoặc các kỹ thuật điều chuẩn theo miền tín hiệu, nhằm cải thiện hơn nữa khả năng tổng quát hóa của mô hình trong các kịch bản triển khai thực tế, nơi nhiễu cảm biến là không thể tránh khỏi.

2.4.9 Hiệu suất theo từng lớp



Hình 2.11: Ma trận nhầm lẫn của mô hình AnpoNet đề xuất

Hình 2.11 trình bày hiệu suất cấp lớp thông qua ma trận nhầm lẫn và Bảng 2.10 minh họa các mẫu precision-recall tương ứng. Kết quả chứng minh rằng AnpoNet đạt được điểm số phân loại hoàn hảo (precision và recall là 1.00) cho phần lớn các lớp, bao gồm D, DL, DR, LL, U, và UR. Điều này làm nổi bật khả năng mạnh mẽ của mô hình trong việc nắm bắt các đặc điểm không gian-thời gian đặc biệt từ dữ liệu gia tốc kể. Tuy nhiên, ma trận nhầm lẫn cũng tiết lộ một điểm yếu cụ thể: mô hình thể hiện một mẫu phân loại sai giữa lớp RD và lớp D, dẫn đến giá trị recall giảm đáng kể xuống 0.63 cho lớp RD. Một mức độ nhầm lẫn nhỏ cũng được quan sát thấy đối với lớp RU. Những kết quả này cung cấp cái nhìn sâu sắc chi tiết về hành vi của mô hình, cho phép xác định các khu vực mục tiêu để cải thiện thêm.

Bảng 2.10: Precision và Recall cho mỗi Lớp

Lớp	Precision	Recall
D	1.00	1.00
DL	1.00	1.00
DR	0.86	1.00
LL (L)	0.91	1.00
LD	1.00	0.89
LU	0.98	1.00
LR (R)	0.91	1.00
RD	1.00	0.63
RU	1.00	0.82
U	1.00	1.00
UL	1.00	0.98
UR	1.00	1.00

2.5 Kết luận chương 2

Chương này đã trình bày phương pháp xây dựng và đánh giá mô hình AnpoNet cho bài toán phân loại tư thế ngủ từ dữ liệu gia tốc kể ba trục đeo trên cơ thể. Mô hình được xây dựng theo hướng kết hợp các lớp tích chập một chiều và mô-đun hồi quy nhằm khai thác đặc trưng cục bộ và đặc trưng theo thời gian của tín hiệu gia tốc. Bên cạnh yêu cầu về hiệu quả phân loại, thiết kế mô hình cũng xem xét đến số lượng tham số và chi phí tính toán, phù hợp với định hướng triển khai trên các thiết bị có tài nguyên hạn chế.

Về kiến trúc, AnpoNet sử dụng các lớp tích chập một chiều để trích xuất đặc trưng từ tín hiệu gia tốc trong từng cửa sổ dữ liệu, kết hợp với các lớp chuẩn hóa nhằm hỗ trợ quá trình huấn luyện. Mô-đun hồi quy được sử dụng để mô hình hóa sự phụ

thuộc theo thời gian của các đặc trưng đã trích xuất. Cách thiết kế này giúp tổ chức quá trình xử lý tín hiệu theo từng bước, từ dữ liệu đầu vào đến biểu diễn đặc trưng và nhận dự đoán. Mô hình có khoảng 2,000 tham số có thể huấn luyện, cho thấy mức độ gọn nhẹ so với một số kiến trúc học sâu thông dụng được sử dụng để so sánh trong chương này.

Hiệu quả của AnpoNet được đánh giá trong thiết lập độc lập theo đối tượng, bao gồm kiểm định chéo 5-fold ở cấp độ người tham gia và các thí nghiệm trên các đối tượng không xuất hiện trong quá trình huấn luyện. Trong thiết lập kiểm định chéo, AnpoNet đạt độ chính xác trung bình 94.67% và F1-Score 92.94%. Kết quả này cao hơn các mô hình baseline được xem xét trong cùng điều kiện thực nghiệm. Bên cạnh đó, kết quả trên các đối tượng kiểm tra chưa xuất hiện trong quá trình huấn luyện cho thấy mô hình có khả năng duy trì hiệu quả phân loại trong phạm vi dữ liệu được khảo sát.

Ngoài các chỉ số phân loại, chương này cũng xem xét số lượng tham số và chi phí tính toán của AnpoNet. Các kết quả thu được cho thấy mô hình đạt được sự cân bằng tương đối giữa hiệu quả phân loại và độ phức tạp mô hình trong thiết lập thực nghiệm của luận án. Đây là cơ sở để tiếp tục nghiên cứu các kiến trúc học sâu cho bài toán phân loại tư thế ngủ sử dụng cảm biến gia tốc, đặc biệt trong các bối cảnh có yêu cầu về mô hình gọn nhẹ và khả năng triển khai trên thiết bị biên.

Các kết quả trong chương này đã được công bố trong một bài báo tạp chí quốc tế được bình duyệt, tương ứng với công trình CT1 trong danh sách công bố khoa học của tác giả. Trên cơ sở các kết quả thu được với AnpoNet, chương tiếp theo tiếp tục xem xét bài toán phân loại tư thế ngủ trên tập dữ liệu mở rộng với 18 người tham gia, đồng thời trình bày mô hình CGB-Net nhằm khảo sát khả năng cải thiện hiệu quả phân loại trong bối cảnh dữ liệu mở rộng.

CHƯƠNG 3. ĐỀ XUẤT MÔ HÌNH CGB-NET NHẪM TĂNG CƯỜNG ĐỘ CHÍNH XÁC TRONG PHÂN LOẠI TƯ THỂ NGŨ

Nội dung trong Chương 3 này trình bày mô hình CGB-Net nhằm tăng cường độ chính xác trong bài toán phân loại 12 tư thể ngũ từ dữ liệu cảm biến gia tốc ba trục. Khác với Chương 2, nơi mô hình AnpoNet được thiết kế theo hướng gọn nhẹ và kiểm soát số lượng tham số, chương này tập trung vào việc nâng cao năng lực biểu diễn đặc trưng không gian-thời gian của tín hiệu nhằm cải thiện hiệu quả phân loại trong điều kiện dữ liệu mở rộng. Trước hết, chương phân tích bài toán và các vấn đề nghiên cứu còn tồn tại khi phân loại các tư thể ngũ chi tiết theo thiết lập độc lập theo đối tượng. Trên cơ sở đó, mô hình CGB-Net được đề xuất bằng cách kết hợp mạng nơ-ron tích chập một chiều, GRU và BiLSTM để khai thác đặc trưng cục bộ, quan hệ tuần tự và ngữ cảnh hai chiều của tín hiệu gia tốc. Chương cũng trình bày khung thuật toán, quy trình huấn luyện và các thiết lập thực nghiệm được sử dụng để đánh giá mô hình. Các kết quả thực nghiệm được phân tích thông qua độ chính xác, F1-score, ma trận nhầm lẫn, độ ổn định theo các lần khởi tạo ngẫu nhiên, quá trình huấn luyện, cũng như ảnh hưởng của độ dài chuỗi và tỷ lệ chồng lấp. Qua đó, chương làm rõ hiệu quả của CGB-Net trong việc cải thiện độ chính xác và khả năng tổng quát hóa cho bài toán phân loại tư thể ngũ chi tiết.

3.1 Bài toán và vấn đề nghiên cứu

Chương 2 đã trình bày mô hình AnpoNet cho bài toán phân loại tư thể ngũ từ dữ liệu gia tốc kế ba trục trong thiết lập độc lập theo đối tượng. Kết quả thực nghiệm cho thấy mô hình có thể đạt hiệu quả phân loại tương đối tốt với số lượng tham số nhỏ, phù hợp với định hướng xây dựng các mô hình gọn nhẹ cho thiết bị có tài nguyên hạn chế. Tuy nhiên, AnpoNet được thiết kế theo hướng đơn giản hóa kiến trúc, trong đó quá trình trích xuất đặc trưng cục bộ và mô hình hóa phụ thuộc thời gian được thực hiện bởi một chuỗi xử lý tương đối gọn.

Trong bài toán phân loại 12 tư thể ngũ, sự khác biệt giữa một số tư thể có thể không lớn, đặc biệt đối với các tư thể bán nghiêng hoặc các tư thể có góc xoay gần nhau. Khi đó, tín hiệu gia tốc giữa các lớp có thể có mức độ tương đồng cao, trong khi dữ liệu giữa các đối tượng lại có sự khác biệt do hình thái cơ thể, cách đeo cảm biến và

dao động nhỏ trong quá trình thực hiện tư thế. Vì vậy, ngoài yêu cầu về mô hình gọn nhẹ, bài toán cũng đặt ra nhu cầu khảo sát các kiến trúc có khả năng biểu diễn đặc trưng không gian–thời gian tốt hơn trong điều kiện đánh giá độc lập theo đối tượng.

Từ vấn đề trên, Chương 3 xem xét một hướng tiếp cận khác so với Chương 2. Thay vì chỉ tập trung vào việc giảm số lượng tham số, chương này nghiên cứu khả năng kết hợp nhiều thành phần học sâu khác nhau nhằm khai thác đặc trưng của tín hiệu gia tốc ở nhiều mức. Cụ thể, mô hình CGB-Net được xây dựng dựa trên sự kết hợp giữa khối tích chập một chiều, khối GRU và khối BiLSTM. Trong đó, khối tích chập một chiều được sử dụng để trích xuất đặc trưng cục bộ từ cửa sổ tín hiệu, GRU được sử dụng để mô hình hóa phụ thuộc thời gian theo chuỗi đặc trưng, còn BiLSTM được sử dụng để khai thác thêm thông tin ngữ cảnh theo hai chiều trong cửa sổ quan sát.

Vấn đề nghiên cứu chính của chương này là đánh giá xem việc bổ sung các thành phần mô hình hóa thời gian và ngữ cảnh hai chiều có thể cải thiện hiệu quả phân loại tư thế ngủ hay không, đặc biệt trong bối cảnh tập dữ liệu được mở rộng và mô hình được đánh giá trên các đối tượng không xuất hiện trong quá trình huấn luyện. Do đó, CGB-Net không được xem là sự thay thế trực tiếp cho AnpoNet theo hướng mô hình gọn nhẹ, mà là một kiến trúc bổ sung nhằm khảo sát sự đánh đổi giữa hiệu quả phân loại và độ phức tạp mô hình.

Để làm rõ kiến trúc đề xuất, chương này trình bày CGB-Net dưới dạng một hệ thống xử lý hoàn chỉnh, từ dữ liệu gia tốc đầu vào đến nhãn tư thế đầu ra. Các khối chức năng trong chương được tổ chức theo trình tự logic tương ứng với các khối chức năng của mô hình. Thuật toán 3 mô tả khung xử lý tổng thể của CGB-Net, trong khi các khối xử lý A5-A9 lần lượt thể hiện các bước xử lý thành phần, bao gồm phân đoạn dữ liệu, trích xuất đặc trưng bằng 1D-CNN, mô hình hóa phụ thuộc thời gian bằng GRU và học biểu diễn ngữ cảnh bằng BiLSTM. Trên cơ sở biểu diễn thu được, tầng phân loại cuối cùng được sử dụng để suy luận nhãn tư thế ngủ.

Thông qua các thí nghiệm trong Chương 3, luận án phân tích hiệu quả của CGB-Net trong cùng định hướng bài toán phân loại 12 tư thế ngủ từ cảm biến gia tốc. Các kết quả được sử dụng để xem xét mức độ cải thiện về hiệu quả phân loại, đồng thời đánh giá sự đánh đổi giữa độ chính xác, độ phức tạp mô hình và khả năng áp dụng trong các hệ thống giám sát tư thế ngủ sử dụng thiết bị đeo.

3.2 Phương pháp tiếp cận mô hình đề xuất

Quyết định tích hợp các lớp CNN, GRU và BiLSTM vào một kiến trúc duy nhất được củng cố thêm bởi một lượng lớn nghiên cứu về mô hình hóa chuỗi thời gian trên nhiều lĩnh vực khác nhau. Các mô hình học sâu kết hợp trích xuất đặc trưng tích chập với nhiều cơ chế hồi quy đã nhiều lần chứng minh hiệu suất vượt trội so với các kiến trúc chỉ dựa trên một mô hình thời gian duy nhất.

Ví dụ, một nghiên cứu gần đây [119] về dự đoán lưu lượng giao thông đã đề xuất một mô hình lai CNN-GRU-LSTM được thiết kế để cùng học các mối tương quan không gian giữa các đoạn đường và các phụ thuộc thời gian ở nhiều mức độ khác nhau trong môi trường giao thông biến đổi liên tục. Các tác giả báo cáo rằng CNN đã nắm bắt hiệu quả các mẫu không gian, trong khi các lớp GRU và LSTM mô hình hóa thông tin thời gian ngắn hạn và dài hạn tương ứng, dẫn đến giảm lỗi dự đoán lên đến 30-35% so với các mô hình cơ bản hiện đại. Điều này chứng tỏ rằng việc sử dụng GRU và LSTM cùng nhau có thể mang lại lợi thế khi dữ liệu thể hiện các phụ thuộc thời gian ở các quy mô khác nhau.

Một minh chứng khác đến từ lĩnh vực nhận dạng cảm xúc qua giọng nói (SER), nơi các tín hiệu thời gian không ổn định đóng vai trò then chốt. Phương pháp học kết hợp [120] tích hợp CNN, LSTM và GRU đã cải thiện đáng kể độ chính xác trên nhiều tập dữ liệu tiêu chuẩn. Trong đó, GRU mô hình hóa hiệu quả các biến thể ngữ cảnh ngắn, còn LSTM nắm bắt các phụ thuộc dài hạn. Các tác giả kết luận rằng việc tích hợp CNN với các kiến trúc mạng nơ-ron hồi quy khác nhau dẫn đến biểu diễn đặc trưng mạnh mẽ hơn, đặc biệt trong các trường hợp liên quan đến dữ liệu nhiều hoặc hạn chế — những điều kiện tương tự như môi trường cảm biến đeo.

Nghiên cứu của [121] trong dự báo thủy văn cũng sử dụng mô hình kết hợp CNN, LSTM và GRU để dự đoán dòng chảy. CNN trích xuất đặc trưng từ đầu vào thời gian dài, LSTM xử lý các đặc trưng dài hạn và GRU xử lý các biến động ngắn hạn. Một mô-đun có tham số được sử dụng riêng để hợp nhất những ưu điểm của cả hai nhánh hồi quy. Mô hình này vượt trội hơn các mạng học sâu truyền thống, khẳng định rằng mô hình hóa hồi quy đa nhánh có thể nắm bắt các cấu trúc thời gian phức tạp tốt hơn các mô hình đơn lẻ.

Tóm lại, các nghiên cứu này cho thấy bằng chứng nhất quán rằng việc kết hợp CNN với cả kiến trúc GRU và LSTM mang lại những lợi ích từ thế mạnh của mỗi

mô-đun đơn lẻ: CNN trích xuất các mẫu phân biệt cục bộ, GRU nắm bắt hiệu quả các phụ thuộc tầm ngắn đến tầm trung với ít tham số hơn, và LSTM (hoặc BiLSTM) nắm bắt ngữ cảnh thời gian hai chiều dài hạn hơn. Do đó, thiết kế kiến trúc của CGB-Net phù hợp với các nguyên tắc đã được kiểm chứng thực nghiệm trong mô hình hóa chuỗi thời gian và được củng cố bởi các tài liệu trước đây chứng minh hiệu quả của các khung kết hợp CNN–RNN trong nhiều ứng dụng thực tế khác nhau. Động lực thiết kế kiến trúc CGB-Net xuất phát từ ba mục tiêu phương pháp luận rõ ràng, nhằm tăng cường năng lực biểu diễn theo thời gian trong khi vẫn kiểm soát chặt chẽ chi phí tính toán.

Thứ nhất, mô-đun tích chập một chiều được sử dụng như một khối trích xuất đặc trưng ban đầu để làm nổi bật các mẫu cục bộ ổn định trong tín hiệu gia tốc kế ba trục. Vai trò của khối này không chỉ là giảm nhiễu và làm mịn tín hiệu, mà còn chuyển đổi dữ liệu thô sang một không gian đặc trưng phù hợp hơn cho các mô-đun mô hình hóa thời gian phía sau.

Thứ hai, lớp Gated Recurrent Unit được tích hợp như một tầng mô hình hóa thời gian trung gian, tập trung vào việc học các phụ thuộc ngắn và trung hạn có tính lặp lại trong chuỗi đặc trưng. Việc lựa chọn GRU thay vì LSTM ở giai đoạn này cho phép giảm số lượng tham số và chi phí huấn luyện, đồng thời tránh đặt toàn bộ gánh nặng mô hình hóa thời gian lên một khối hồi quy duy nhất.

Thứ ba, lớp Bidirectional LSTM được sử dụng ở giai đoạn sau cùng nhằm nắm bắt ngữ cảnh thời gian hai chiều trên toàn bộ chuỗi quan sát. Thiết kế này cho phép mô hình khai thác đồng thời thông tin từ các bước thời gian trước và sau, đặc biệt hữu ích trong các trường hợp ranh giới giữa các tư thế có tính mờ và các khác biệt chỉ thể hiện qua các biến thiên thời gian tinh tế.

Thông qua chiến lược phân rã mô hình hóa thời gian theo từng cấp, CGB-Net kết hợp khả năng trích xuất đặc trưng phân cấp của các lớp tích chập, hiệu quả tham số của GRU và khả năng học ngữ cảnh toàn cục của BiLSTM. Cách tiếp cận này không nhằm đơn thuần tăng độ sâu mạng, mà hướng tới việc phân vai chức năng rõ ràng cho từng mô-đun, qua đó tăng cường khả năng tổng quát hóa trong các thiết lập đánh giá độc lập theo đối tượng, đồng thời vẫn duy trì định hướng triển khai trên thiết bị biên.

Các kết quả và phân tích liên quan đến kiến trúc CGB-Net đã được công bố trong một bài báo khoa học quốc tế được bình duyệt (SCIE), tương ứng với Publication CT2 trong danh sách các đóng góp khoa học của tác giả. Ấn phẩm này xác nhận giá trị phương pháp luận của chiến lược thiết kế kiến trúc được đề xuất, cũng như tính phù

hợp của nó đối với các bài toán phân loại chuỗi thời gian trong bối cảnh tài nguyên hạn chế.

3.3 Khung thuật toán đề xuất CGB-Net

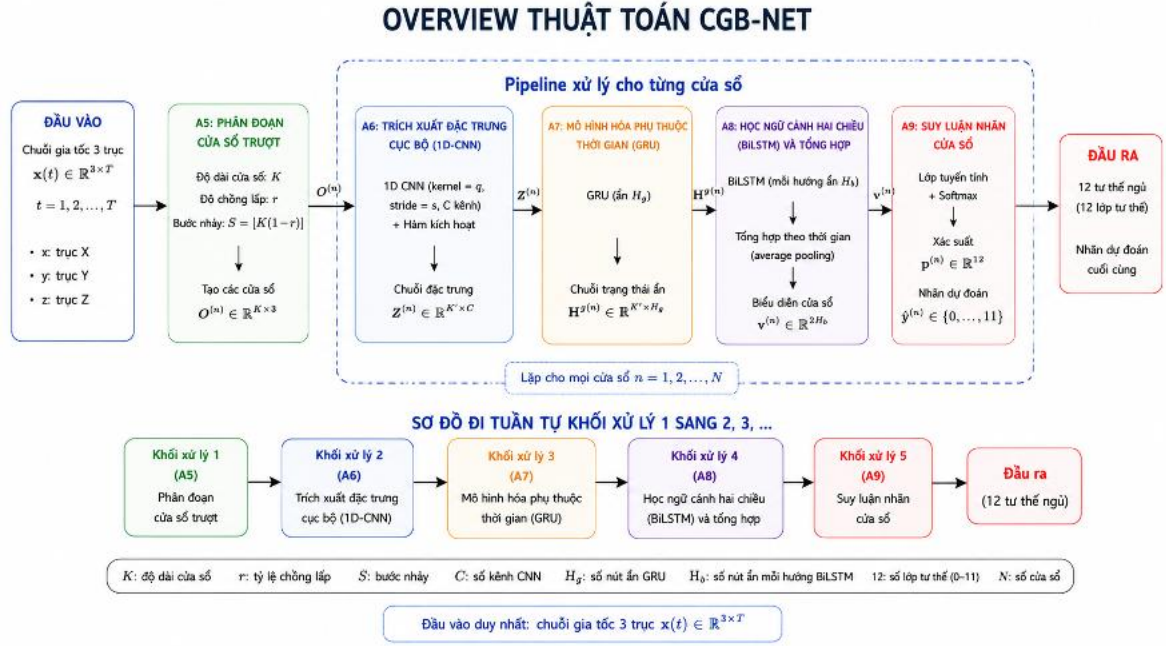
Hình 3.1 trình bày luồng xử lý tổng thể của CGB-Net, trong đó tín hiệu gia tốc ba trục lần lượt được xử lý bởi các khối 1D CNN, GRU và BiLSTM trước khi đưa đến lớp phân loại. Các thành phần này không chỉ được ghép nối nhằm tăng độ sâu của mạng mà được thiết kế với các vai trò bổ trợ khác nhau.

Khối 1D CNN thực hiện trích xuất các đặc trưng cục bộ từ tín hiệu gia tốc, đặc biệt là các biến thiên ngắn hạn và mối quan hệ giữa các thành phần tín hiệu trong từng cửa sổ quan sát. Biểu diễn này sau đó được đưa vào GRU để mô hình hóa các phụ thuộc tuần tự ngắn và trung hạn với chi phí tham số tương đối thấp. Trên cơ sở biểu diễn đã được xử lý bởi GRU, BiLSTM tiếp tục khai thác ngữ cảnh thời gian theo hai chiều nhằm tăng cường khả năng phân biệt các tư thế có đặc trưng gia tốc gần nhau.

Việc kết hợp đồng thời GRU và BiLSTM trong CGB-Net không nhằm đơn thuần tăng độ sâu của mạng mà được thiết kế theo hướng phân cấp chức năng. Sau khi các lớp 1D CNN trích xuất các đặc trưng cục bộ từ tín hiệu gia tốc, GRU được sử dụng như một tầng mô hình hóa tuần tự trung gian để tổng hợp các phụ thuộc thời gian ngắn và trung hạn. Với cấu trúc cổng tương đối đơn giản, GRU có khả năng tạo ra biểu diễn tuần tự gọn hơn trước khi chuyển sang tầng BiLSTM.

BiLSTM được đặt sau GRU nhằm tiếp tục khai thác ngữ cảnh thời gian theo cả hai chiều trên chuỗi đặc trưng đã được tinh lọc. Cơ chế hai chiều cho phép biểu diễn tại một thời điểm được xây dựng từ thông tin ở cả phía trước và phía sau trong cửa sổ quan sát, qua đó hỗ trợ phân biệt các tư thế có tín hiệu gia tốc tương đối gần nhau. Như vậy, GRU chủ yếu đảm nhiệm việc tổng hợp và nén thông tin tuần tự, trong khi BiLSTM tăng cường khai thác ngữ cảnh thời gian hai chiều.

Sự kết hợp này tạo thành một cơ chế mô hình hóa thời gian theo nhiều mức, trong đó 1D CNN khai thác đặc trưng cục bộ, GRU mô hình hóa phụ thuộc tuần tự và BiLSTM khai thác ngữ cảnh rộng hơn. Tuy nhiên, việc bổ sung đồng thời GRU và BiLSTM làm tăng số lượng tham số và chi phí tính toán. Vì vậy, hiệu quả của CGB-Net cần được đánh giá đồng thời trên hai khía cạnh: mức cải thiện về hiệu suất phân loại và phần chi phí tính toán tăng thêm so với các kiến trúc đơn giản hơn.



Hình 3.1: Tổng thể luồng xử lý của CGB-Net

Thuật toán 3 Khung thuật toán CGB-Net cho phân loại tư thế ngủ**Input:** Chuỗi gia tốc ba trục $\mathbf{x}(t) \in \mathbb{R}^3$ **Output:** Tập dự đoán tư thế theo cửa sổPhân đoạn tín hiệu thành các cửa sổ $\{\mathbf{O}^{(n)}\}$ bằng Khối xử lý A5**foreach** $\mathbf{O}^{(n)}$ **do** $\mathbf{Z}^{(n)} \leftarrow$ Khối xử lý A6 $\mathbf{H}^{g(n)} \leftarrow$ Khối xử lý A7 $\mathbf{v}^{(n)} \leftarrow$ Khối xử lý A8 $(\mathbf{p}^{(n)}, \hat{y}^{(n)}) \leftarrow$ Khối xử lý A9**return** $\{(\mathbf{p}^{(n)}, \hat{y}^{(n)})\}$ **Khối xử lý A5: Nạp dữ liệu và phân đoạn cửa sổ trượt theo đối tượng****Input:** root, S , K , r , $(\mathcal{S}_{train}, \mathcal{S}_{val}, \mathcal{S}_{test})$. **Output:** $\mathcal{D}_{train}$, $\mathcal{D}_{val}$, $\mathcal{D}_{test}$.

Bước nhảy:

$$\Delta = \max\{1, \lfloor K(1-r) \rfloor\}.$$

Với mỗi $s \in \mathcal{S}$, nạp

$$(\mathbf{x}^{(s)}, \mathbf{y}^{(s)}) = \text{load_subject}(\text{root}, s), \quad \mathbf{x}^{(s)} \in \mathbb{R}^{T_s \times 3}.$$

Số của số:

$$M_s = \left\lfloor \frac{T_s - K}{\Delta} \right\rfloor + 1.$$

Với $m = 1, \dots, M_s$, đặt

$$u_m = 1 + (m - 1)\Delta, \quad \mathbf{O}^{(s,m)} = \mathbf{K}^{(s)}[u_m : u_m + K - 1],$$

$$y^{(s,m)} = \mathbf{mode}(\mathbf{y}^{(s)}[u_m : u_m + K - 1]).$$

Đưa $(\mathbf{O}^{(s,m)}, y^{(s,m)})$ vào $\mathcal{D}_{train}$, $\mathcal{D}_{val}$ hoặc $\mathcal{D}_{test}$ theo tập đối tượng chứa s .

Return: $\mathcal{D}_{train}$, $\mathcal{D}_{val}$, $\mathcal{D}_{test}$.

Khối xử lý A6: Trích xuất đặc trưng cục bộ bằng 1D-CNN

Input: $\mathbf{O} \in \mathbb{R}^{K \times 3}$, C , q , s , $(\mathbf{W}, \mathbf{b})$, ϕ . **Output:** $\mathbf{Z} \in \mathbb{R}^{K' \times C}$.

Chuyển dữ liệu sang dạng kênh-trước $\mathbf{O} \in \mathbb{R}^{3 \times K}$ và tính

$$K' = \left\lfloor \frac{K - q}{s} \right\rfloor + 1.$$

Với $c = 1, \dots, C$ và $t = 1, \dots, K'$:

$$a_{c,t} = b_c + \sum_{i=1}^3 \sum_{k=0}^{q-1} W_{c,i,k} U_{i,(t-i)s+k+1},$$

$$z_{c,t} = \phi(a_{c,t}).$$

Đặt

$$\mathbf{Z} = [z_{c,t}]^\top \in \mathbb{R}^{K' \times C}.$$

Return: $\mathbf{Z}$.

Khối xử lý A7: Mô hình hóa phụ thuộc theo thời gian bằng GRU

Input: $\mathbf{Z} \in \mathbb{R}^{K' \times C}$, H_g , $(W_z, W_r, W_h, b_z, b_r, b_h)$. **Output:** $\mathbf{h}^g = \{\mathbf{h}_t^g\}_{t=1}^{K'}$.

Khởi tạo

$$\mathbf{h}_0^g = \mathbf{0}.$$

Với $t = 1, \dots, K'$ và $\mathbf{z}_t = \mathbf{Z}[t]$:

$$\mathbf{u}_t = \sigma(W_z[\mathbf{h}_{t-1}^g; \mathbf{z}_t] + b_z),$$

$$\mathbf{r}_t = \sigma \left(W_r [\mathbf{k}_{t-i}^g; \mathbf{z}_t] + b_r \right),$$

$$\tilde{\mathbf{k}}_t^g = \tanh \left(W_h [\mathbf{r}_t \odot \mathbf{k}_{t-i}^g; \mathbf{z}_t] + b_h \right),$$

$$\mathbf{k}_t^g = (\mathbf{1} - \mathbf{u}_t) \odot \mathbf{k}_{t-i}^g + \mathbf{u}_t \odot \tilde{\mathbf{k}}_t^g.$$

Return: $\mathbf{I}^g$.

Khối xử lý A8: Học ngữ cảnh hai chiều bằng BiLSTM và tổng hợp theo thời gian

Input: $\mathbf{I}^g = \{\mathbf{k}_t^g\}_{t=i}^{K'}, H_b$. **Output:** $\mathbf{v} \in \mathbb{R}^{2H_b}$.

Với mỗi thời điểm, LSTM chuẩn thực hiện

$$\mathbf{i}_t = \sigma(W_i \mathbf{x}_t + U_i \mathbf{k}_{t-i} + \mathbf{b}_i),$$

$$\mathbf{f}_t = \sigma(W_f \mathbf{x}_t + U_f \mathbf{k}_{t-i} + \mathbf{b}_f),$$

$$\mathbf{o}_t = \sigma(W_o \mathbf{x}_t + U_o \mathbf{k}_{t-i} + \mathbf{b}_o),$$

$$\tilde{\mathbf{c}}_t = \tanh(W_c \mathbf{x}_t + U_c \mathbf{k}_{t-i} + \mathbf{b}_c),$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-i} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t,$$

$$\mathbf{k}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t).$$

BiLSTM được tính theo hai hướng với $\mathbf{x}_t = \mathbf{k}_t^g$. Tại mỗi thời điểm:

$$\mathbf{k}_t^{bi} = [\overrightarrow{\mathbf{k}}_t || \overleftarrow{\mathbf{k}}_t].$$

Biểu diễn của số:

$$\mathbf{v} = \frac{1}{K'} \sum_{t=1}^{K'} \mathbf{k}_t^{bi}.$$

Return: $\mathbf{v}$.

Khối xử lý A9: Phân loại tư thế ngủ bằng lớp tuyến tính và Softmax

Input: $\mathbf{v} \in \mathbb{R}^{2H_b}$, (W_o, b_o) , $K_c = 12$. **Output:** $\mathbf{p} \in \mathbb{R}^{12}$, $\hat{y}$.

Tính logits:

$$\mathbf{s} = W_o \mathbf{v} + b_o.$$

Xác suất hậu nghiệm:

$$p_k = \frac{\exp(s_k)}{\sum_{j=1}^{12} \exp(s_j)}, \quad k = 1, \dots, 12.$$

Nhân dự đoán:

$$\hat{y} = \arg \max_{k \in \{1, \dots, 12\}} p_k.$$

Return: $(\mathbf{p}, \hat{y})$.

Phân tích tính khả thi của mô hình CGB-Net

Mục tiêu của CGB-Net không phải là “tăng tầng” để tăng độ chính xác dựa trên thực nghiệm, mà là tổ chức lại quá trình học biểu diễn chuỗi thời gian theo một chiến lược có thể lập luận được: tách phần học đặc trưng cục bộ (local representation) khỏi phần tích hợp phụ thuộc theo thời gian (temporal integration), sau đó bổ sung ngữ cảnh hai chiều khi bài toán có tính chất gần như phi tuyến. Cách tiếp cận này phù hợp với các nghiên cứu tổng quan trong phân loại chuỗi thời gian, nơi các kiến trúc dựa trên tích chập và các biến thể lai (CNN + RNN) thường mang lại biểu diễn ổn định hơn so với việc chỉ dùng một mô-đun tuần tự đơn lẻ trên tín hiệu thô [77].

Phân rã ánh xạ học (functional decomposition). Với một cửa sổ đầu vào $\mathbf{O} \in \mathbb{R}^{K \times 3}$, bài toán học có thể biểu diễn dưới dạng một ánh xạ

$$f : \mathbb{R}^{K \times 3} \rightarrow \Delta^{12},$$

trong đó Δ^{i2} là đơn hình xác suất 12 chiều. Thay vì học trực tiếp f bằng một mô hình tuần tự duy nhất, CGB-Net triển khai hợp thành các ánh xạ con:

$$f = f_{\text{cls}} \circ f_{\text{pool}} \circ f_{\text{bi}} \circ f_{\text{gru}} \circ f_{\text{cnn}}.$$

Cách phân rã này có hai lợi ích phương pháp luận: (i) giảm độ khó tối ưu ở từng tầng bằng cách thu hẹp không gian hàm cần học theo từng giai đoạn; (ii) làm rõ vai trò của từng khối theo đúng chức năng, từ đó tránh diễn giải mô hình như một phép ghép nối tùy ý.

1D-CNN như một toán tử trích xuất đặc trưng cục bộ cho chuỗi thời gian.

Trong phân loại chuỗi thời gian, tích chập một chiều thường được dùng như một toán tử học được để phát hiện mẫu cục bộ và tăng tính bất biến đối với nhiễu/dao động nhỏ theo trục thời gian [77]. Với kernel kích thước q , 1D-CNN thực hiện phép biến đổi:

$$\mathbf{z}_t = \phi \left(\sum_{k=0}^{q-1} W_k \mathbf{o}_{t+k} + \mathbf{b} \right),$$

tạo ra chuỗi đặc trưng $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_{K'}]$. Điểm quan trọng ở đây là CNN không “thay thế” mô hình thời gian mà đóng vai trò chuẩn hóa biểu diễn: thay vì đưa trực tiếp (ax, ay, az) vào RNN, CNN tạo ra một không gian đặc trưng mà ở đó các mẫu cục bộ liên quan (ví dụ dao động ngắn hạn do hô hấp hoặc nhiễu cảm biến) được ánh xạ thành các đặc trưng ổn định hơn trước khi tích hợp theo thời gian.

GRU: cơ chế ghi nhớ hiệu quả tham số cho phụ thuộc ngắn–trung hạn.

GRU được đề xuất như một biến thể cổng hóa giúp học phụ thuộc theo thời gian với cấu trúc đơn giản hơn LSTM. Về mặt cấu trúc, GRU chỉ sử dụng hai cổng (update và reset), trong khi LSTM có thêm trạng thái ô nhớ và nhiều cổng hơn. Do đó, với cùng kích thước ẩn, GRU thường có số tham số thấp hơn và ít phép toán hơn, điều này được phân tích rõ trong các so sánh hình thức giữa GRU và LSTM. Trong CGB-Net, GRU được đặt ngay sau CNN để tổng hợp phụ thuộc ngắn–trung hạn trên chuỗi đặc trưng (thay vì trên tín hiệu thô), giúp giảm nguy cơ quá khớp khi số mẫu huấn luyện theo đối tượng không lớn.

BiLSTM: khai thác ngữ cảnh hai chiều Mạng hồi quy hai chiều (Bidirectional RNN) được giới thiệu nhằm tận dụng thông tin từ cả quá khứ và tương lai bằng cách chạy RNN theo hai hướng thời gian. Với giả định suy luận theo cửa sổ (window-based) và xử lý ngoại tuyến/không ràng buộc nhân quả trong từng cửa sổ, việc bổ sung chiều nghịch cho phép mỗi thời điểm t được biểu diễn bởi

$$\mathbf{k}_t^{bi} = [\overrightarrow{\mathbf{k}}_t \parallel \overleftarrow{\mathbf{k}}_t],$$

từ đó tạo ra biểu diễn giàu ngữ cảnh hơn so với RNN một chiều. Mục đích của BiLSTM trong CGB-Net không phải là “tăng độ sâu”, mà là tăng năng lực biểu diễn theo thời gian bằng cách mở rộng miền ngữ cảnh có thể được mã hóa trong cùng một cửa sổ.

Tổng hợp theo thời gian và ổn định hóa biểu diễn đầu ra. Sau BiLSTM, CGB-Net dùng phép tổng hợp theo thời gian (mean pooling / global average pooling trên trục thời gian) để thu vector biểu diễn cố định:

$$\mathbf{v} = \frac{1}{K'} \sum_{t=1}^{K'} \mathbf{k}_t^{bi}.$$

Trong phân loại chuỗi thời gian, chiến lược tổng hợp kiểu này được sử dụng phổ biến nhằm giảm độ nhạy của mô hình đối với nhiễu cục bộ tại một vài thời điểm, đồng thời tạo biểu diễn gọn cho lớp phân loại; chẳng hạn InceptionTime sử dụng global average pooling trước lớp Softmax. Đối với bài toán tư thế ngủ, tổng hợp theo thời gian phù hợp với giả thiết rằng thông tin tư thế là trạng thái tương đối ổn định trong phần lớn cửa sổ, thay vì phụ thuộc vào một điểm thời gian đơn lẻ.

3.4 Thực nghiệm

3.4.1 Các mô hình so sánh

Để đánh giá toàn diện hiệu suất của CGB-Net được đề xuất, các so sánh đã được thực hiện với một số mô hình cơ sở, bao gồm các phương pháp tiên tiến gần đây và các kiến trúc mạng nơ-ron thường được sử dụng cho các nhiệm vụ phân loại chuỗi thời gian. Các mô hình cơ sở này thuộc hai loại: (1) các mô hình đã được thiết lập từ các sách chính thống và (2) các kiến trúc học sâu chuẩn được điều chỉnh cho nhiệm vụ phân loại tư thế ngủ 12 lớp.

Các mô hình từ nghiên cứu trước đó

- **Mô hình LSTM của Rawan và cộng sự [63]:** Rawan et al. (2023) đã sử dụng mạng LSTM để phân loại tư thế ngủ bằng cách sử dụng dữ liệu cảm biến chuỗi thời gian. Mô hình này đóng vai trò là một ví dụ tiêu biểu về kiến trúc mạng nơ-ron hồi quy thuần túy cho mô hình hóa thời gian trong nhận dạng tư thế.
- **AnpoNet:** AnpoNet tích hợp mạng CNN 1D với lớp LSTM, được bổ sung bởi mạng Bayes (BN) sau phép toán tích chập. Kiến trúc này được thiết kế để nắm bắt cả các mẫu thời gian cục bộ và các phụ thuộc tầm xa trong tín hiệu gia tốc kế.

Các kiến trúc nền tảng

Một số mô hình được đánh giá trong nghiên cứu này có thể được hiểu là các biến thể loại bỏ ngấm của CGB-Net. Các kiến trúc này giúp đánh giá đóng góp riêng lẻ của các thành phần chính như lớp CNN 1D, các lớp mạng nơ-ron hồi quy (GRU và BiLSTM). Tất cả các mô hình đều có cùng cấu trúc đầu vào-đầu ra và quy trình huấn luyện để đảm bảo tính nhất quán và so sánh công bằng.

1D CNN: Mô hình này tương ứng với phiên bản CGB-Net đã loại bỏ cả hai lớp *GRU* và *BiLSTM*. Nó bao gồm hai lớp tích chập 1D với 8 và 16 bộ lọc, mỗi lớp được theo sau bởi Batch Normalization và Dropout (tỷ lệ = 0.4). Một lớp Fully Connected với 16 đơn vị và đầu ra softmax 12 nhãn ở cuối mô hình. Mô hình cơ bản này chỉ nắm bắt các đặc trưng không gian, mà không mô hình hóa các phụ thuộc thời gian.

1D CNN + GRU: Biến thể này tái sử dụng mô hình hóa thời gian thông qua lớp GRU nhưng bỏ qua khối BiLSTM từ CGB-Net. Nó bắt đầu với một lớp tích chập 1D (8 bộ lọc), tiếp theo là Chuẩn hóa theo lô (Batch Normalization), Gộp theo giá trị tối đa (MaxPooling) và Gộp theo giá trị trung bình toàn cục (Global Average Pooling - GAP). Chuỗi thời gian được xử lý bằng GRU, và phân loại được thực hiện bởi một lớp Fully Connected với 12 đơn vị đầu ra. Thiết kế này giúp cô lập tác động của mô hình hóa thời gian một chiều dựa trên GRU.

1D CNN + BiLSTM: Mô hình này bao gồm cả lớp CNN 1D và lớp BiLSTM nhưng không bao gồm lớp GRU. Nó bắt đầu với một lớp tích chập 1D (8 bộ lọc), tiếp theo là hai lớp BiLSTM xếp chồng lên nhau để mô hình hóa các phụ thuộc thời gian hai chiều. Chuẩn hóa theo lô (Batch Normalization), gộp tối đa (MaxPooling) và GAP

được áp dụng trước bộ phân loại Fully Connected cuối cùng.

MLP: Mô hình MLP đóng vai trò là mô hình cơ sở tối giản, loại bỏ tất cả các module tích chập và mạng nơ-ron hồi quy. Nó bao gồm hai lớp ẩn (64 và 32 đơn vị), mỗi lớp đều được theo sau bởi Batch Normalization và Dropout (tỷ lệ = 0.4). Mô hình trực tiếp ánh xạ các cửa sổ đầu vào được làm phẳng đến xác suất của từng nhãn thông qua lớp đầu ra softmax.

3.4.2 Huấn luyện và tối ưu hóa mô hình

Quy trình huấn luyện và tối ưu hóa của CGB-Net được xây dựng trên cùng một khung phương pháp luận đã được thiết lập trong Chương 2 nhằm đảm bảo tính nhất quán và khả năng so sánh công bằng giữa các kiến trúc được nghiên cứu. Cụ thể, hàm mất mát cross-entropy đa lớp và bộ tối ưu Adam được sử dụng xuyên suốt cho cả AnpoNet và CGB-Net, với các siêu tham số huấn luyện được giữ cố định trong tất cả các thí nghiệm.

Việc sử dụng chung khung tối ưu hóa không chỉ giúp giảm sự phức tạp trình bày mà còn có ý nghĩa phương pháp luận quan trọng: nó cho phép cô lập tác động của thiết kế kiến trúc lên hiệu suất mô hình, loại bỏ các yếu tố gây nhiễu xuất phát từ chiến lược huấn luyện hoặc bộ tối ưu. Do đó, mọi khác biệt về độ chính xác, độ ổn định hoặc khả năng tổng quát hóa quan sát được giữa AnpoNet và CGB-Net có thể được quy cho năng lực biểu diễn nội tại của kiến trúc, thay vì các lựa chọn huấn luyện khác nhau.

Cụ thể, đối với mỗi cửa sổ đầu vào được phân đoạn theo Khối xử lý A5, mô hình CGB-Net thực hiện lan truyền tiến thông qua chuỗi các mô-đun được mô tả trong Khối xử lý A6 đến A9 để sinh ra phân bố xác suất trên mười hai lớp tư thế ngủ. Hàm mất mát cross-entropy được tính giữa phân bố dự đoán và nhãn thực, sau đó các gradient được lan truyền ngược qua toàn bộ mạng để cập nhật tham số bằng bộ tối ưu Adam. Các thuật toán chi tiết cho hàm mất mát và cập nhật tham số đã được trình bày trong Thuật toán 2 của Chương 2 và được áp dụng trực tiếp trong chương này mà không có thay đổi.

Cách tiếp cận này đảm bảo rằng CGB-Net được đánh giá trong cùng điều kiện huấn luyện với AnpoNet, từ đó cung cấp một cơ sở thực nghiệm đáng tin cậy để phân tích các đánh đổi giữa năng lực mô hình hóa thời gian, độ phức tạp kiến trúc và hiệu quả triển khai. Trong bối cảnh đó, Chương 3 không nhằm đề xuất một chiến lược huấn

luyện mới, mà tập trung vào việc mở rộng không gian thiết kế kiến trúc và đánh giá một cách có kiểm soát tác động của các lựa chọn kiến trúc nâng cao đối với bài toán phân loại tư thế ngủ độc lập theo đối tượng. Trong nghiên cứu này, hai chiến lược đánh giá đã được áp dụng để đánh giá hiệu suất của mô hình. Cách tiếp cận đầu tiên liên quan đến sơ đồ kiểm định chéo, trong đó tập dữ liệu thu thập từ 18 người tham gia được chia thành các tập con huấn luyện, xác thực và kiểm tra. Đối với mỗi lần chia, dữ liệu từ 15 người tham gia (khoảng 83%) được phân bổ cho huấn luyện và xác thực, trong khi 3 người tham gia còn lại (khoảng 17%) được dành cho kiểm tra. Trong tập huấn luyện, một phép chia nội bộ được thực hiện theo tỷ lệ 80:20 để tạo ra tập xác thực.

Chiến lược thứ hai bao gồm bốn thí nghiệm độc lập. Trong mỗi thí nghiệm, ba người tham gia được chọn ngẫu nhiên làm tập dữ liệu thử nghiệm, trong khi 15 người tham gia còn lại được sử dụng để huấn luyện và xác thực. Hiệu suất cuối cùng được báo cáo là kết quả trung bình trên ba đối tượng thử nghiệm trong mỗi thí nghiệm, cung cấp một đánh giá mạnh mẽ và mang tính đại diện về khả năng khái quát hóa của mô hình.

Để tối ưu hóa các tham số thời gian cho việc tạo cửa sổ, một phương pháp cửa sổ trượt dựa trên thời lượng thay vì số lượng mẫu thô đã được áp dụng. Với tần số lấy mẫu là 50 Hz, độ dài cửa sổ từ 0.5 đến 4.0 giây—tương ứng với 25, 50, 75, ..., lên đến 200 mẫu—đã được đánh giá. Phạm vi này được chọn để xác định kích thước cửa sổ tối ưu để thu thập các mẫu hoạt động, vì các nghiên cứu trước đây chưa khám phá một cách có hệ thống các khoảng thời gian ngắn hơn và trung gian.

Ngoài ra, ảnh hưởng của việc thay đổi tỷ lệ chồng lấp giữa các cửa sổ liên tiếp cũng được xem xét, với các giá trị từ 0.2 đến 0.8 theo bước tăng 0.2 (tức là chồng lấp 20%, 40%, 60% và 80%). Chiều dài cửa sổ nhỏ hơn kết hợp với độ chồng lấp cao hơn có thể tăng cường khả năng phản hồi và độ phân giải thời gian của mô hình, điều này đặc biệt có lợi cho các ứng dụng thời gian thực yêu cầu độ trễ thấp và phát hiện chi tiết.

Để giảm thiểu hiện tượng quá khớp, CGB-Net kết hợp một số chiến lược điều chỉnh, bao gồm các thành phần phạt L1 và L2, cũng như các lớp Dropout được phân bố khắp kiến trúc.

Thuật toán tối ưu: Tất cả các mô hình học sâu đều được huấn luyện bằng thuật toán tối ưu Adam [118], được khởi tạo với tốc độ học là 0.001. Quá trình huấn

luyện được thực hiện trong 250 epoch với kích thước batch là 512.

Lựa chọn Random Seed: Phần này tiến hành năm lần chạy độc lập sử dụng các Seed được chọn ngẫu nhiên để tính đến sự biến động về hiệu suất. Kết quả được báo cáo thể hiện giá trị trung bình và độ lệch chuẩn trên năm lần chạy này. Không có việc lựa chọn Seed nào dựa trên hiệu suất được áp dụng.

Chi tiết kiến trúc: Mô hình CGB-Net được đề xuất bao gồm một lớp CNN 1D với 8 bộ lọc và kích thước kernel là 3 (80 tham số), tiếp theo là Batch Normalization (32 tham số) để ổn định quá trình học. Một lớp GRU với 16 đơn vị (2,496 tham số) nắm bắt các phụ thuộc thời gian tiến về phía trước trong chuỗi. Tiếp theo là một lớp Bidirectional LSTM (BiLSTM) với 16 đơn vị mỗi hướng (6,272 tham số), cho phép mô hình kết hợp thông tin ngữ cảnh từ cả các bước thời gian trong quá khứ và tương lai. Các lớp Batch Normalization và MaxPooling bổ sung giúp tinh chỉnh thêm các đặc trưng. Mạng kết thúc với global average pooling, một lớp Dropout (tỷ lệ = 0.4) và một lớp đầu ra Dense để phân loại 12 lớp (396 tham số). Tổng số tham số có thể huấn luyện là 9,532.

3.5 Kết quả thực nghiệm và thảo luận

Phần này tóm tắt các kết quả chính thu được từ các thí nghiệm, tập trung vào độ chính xác của mô hình, F1-score và các chỉ số về hiệu suất của mô hình. Cần lưu ý rằng kết quả của AnpoNet trong Chương 2 và kết quả so sánh giữa AnpoNet với CGB-Net trong Chương 3 được thực hiện trên các thiết lập dữ liệu khác nhau. Do đó, việc so sánh trực tiếp hiệu suất tuyệt đối giữa hai chương là không phù hợp; thay vào đó, hiệu quả của CGB-Net cần được đánh giá trong cùng các thiết lập thí nghiệm của Chương 3.

3.5.1 Phân tích hiệu suất của mô hình

Để đánh giá hiệu suất của các mô hình học sâu khác nhau trong các thiết lập thử nghiệm khác nhau, nghiên cứu này đã tiến hành phân tích so sánh các kiến trúc, bao gồm MLP, 1D CNN, 1D CNN + BiLSTM, 1D CNN + GRU, phương pháp của Rawan et al. (LSTM), AnpoNet và mô hình đề xuất của chúng tôi, CGB-Net.

Bảng 3.1 và Bảng 3.2 trình bày kết quả đánh giá các mô hình theo bốn thiết lập thí nghiệm khác nhau, sử dụng hai chỉ số chính là độ chính xác (Accuracy) và F1-score. Trong khi độ chính xác phản ánh tỷ lệ dự đoán đúng tổng thể, F1-score cung cấp một

Bảng 3.1: Độ chính xác (%) của các mô hình theo bốn thiết lập thí nghiệm

Mô hình	Thiết lập 1	Thiết lập 2	Thiết lập 3	Thiết lập 4
MLP	82.02	81.25	88.07	78.18
1D CNN	86.38	77.84	89.52	83.20
1D CNN + BiLSTM	85.99	75.96	88.04	77.91
1D CNN + GRU	85.70	77.07	88.49	79.88
Rawan và cộng sự (LSTM)	77.45	78.97	90.63	88.60
AnpoNet	85.20	81.63	92.18	81.23
CGB-Net	89.34	84.39	92.70	83.95

Bảng 3.2: Chỉ số F1 (%) của các mô hình theo bốn thiết lập thí nghiệm

Mô hình	Thiết lập 1	Thiết lập 2	Thiết lập 3	Thiết lập 4
MLP	75.86	72.18	77.71	68.68
1D CNN	82.20	66.93	85.57	78.64
1D CNN + BiLSTM	80.11	68.48	83.20	69.20
1D CNN + GRU	79.98	69.07	85.85	69.90
Rawan và cộng sự (LSTM)	65.87	67.40	85.75	75.80
AnpoNet	76.10	70.54	88.21	72.01
CGB-Net	86.24	79.41	89.71	78.17

thước đo cân bằng hơn giữa precision và recall, đặc biệt quan trọng trong các bài toán phân loại đa lớp có thể tồn tại sự mất cân bằng giữa các lớp. Việc phân tích đồng thời hai chỉ số này giúp đánh giá toàn diện hơn hiệu suất của các mô hình.

Xét tổng thể, CGB-Net đạt hiệu suất cao nhất trong hầu hết các thiết lập đối với cả hai chỉ số. Trong Thiết lập 1, CGB-Net đạt độ chính xác 89.34% và F1-score 86.24%, cao hơn đáng kể so với các mô hình cơ sở. So với AnpoNet (85.20% accuracy và 76.10% F1-score), CGB-Net cải thiện khoảng 4.14 điểm phần trăm về độ chính xác và hơn 10 điểm phần trăm về F1-score. Sự chênh lệch lớn hơn ở F1-score cho thấy CGB-Net không chỉ tăng tỷ lệ dự đoán đúng tổng thể mà còn cải thiện đáng kể sự cân bằng giữa các lớp tư thế, giảm hiện tượng thiên lệch về phía các lớp dễ phân loại.

Xu hướng tương tự cũng được quan sát trong Thiết lập 2. Trong thiết lập này, các kiến trúc CNN-RNN đơn lẻ như CNN 1D + BiLSTM và CNN 1D + GRU đạt F1-score lần lượt là 68.48% và 69.07%, thấp hơn đáng kể so với CGB-Net (79.41%). Sự cải thiện gần 10 điểm phần trăm về F1-score cho thấy CGB-Net có khả năng phân biệt tốt hơn giữa các tư thế tương tự ngay cả khi dữ liệu kiểm thử có sự khác biệt lớn so với dữ liệu huấn luyện. Điều này phản ánh lợi thế của kiến trúc kết hợp nhiều cơ chế mô hình hóa thời gian trong việc học các biểu diễn ổn định hơn.

Trong Thiết lập 3, hầu hết các mô hình đều đạt hiệu suất cao hơn so với các

thiết lập còn lại, với nhiều kiến trúc đạt độ chính xác trên 88% và F1-score vượt 83%. Điều này cho thấy Thiết lập 3 có khả năng là kịch bản đánh giá ít thách thức hơn đối với bài toán phân loại. Một nguyên nhân hợp lý là sự tương đồng cao hơn giữa phân bố dữ liệu huấn luyện và kiểm thử, giúp các đặc trưng học được có thể được tái sử dụng hiệu quả hơn trong quá trình suy luận. Trong bối cảnh này, ngay cả các mô hình đơn giản hơn như 1D CNN cũng đạt độ chính xác 89.52% và F1-score 85.57%. Tuy nhiên, CGB-Net vẫn đạt kết quả cao nhất với 92.70% accuracy và 89.71% F1-score, cho thấy kiến trúc này vẫn duy trì lợi thế ngay cả trong các điều kiện đánh giá thuận lợi.

Một điểm đáng chú ý khác là hiệu suất tương đối cao của mô hình LSTM của Rawan và cộng sự trong Thiết lập 4, với độ chính xác 88.60% và F1-score 75.80%. Hiện tượng này có thể được lý giải từ bản chất của tín hiệu gia tốc trong bài toán tư thế ngủ. Khi cơ thể ở trạng thái gần tĩnh, tín hiệu gia tốc chủ yếu phản ánh sự thay đổi định hướng của véc-tơ trọng lực theo thời gian. Trong những trường hợp dữ liệu có cấu trúc thời gian ổn định và ít nhiễu chuyển động, các mô hình tuần tự thuần túy như LSTM có thể khai thác hiệu quả tính liên tục của chuỗi tín hiệu mà không cần đến các tầng tích chập để trích xuất đặc trưng cục bộ. Điều này giúp mô hình LSTM đạt hiệu suất tốt trong một số kịch bản cụ thể.

Một quan sát quan trọng khác là mức cải thiện của CGB-Net so với AnpoNet lớn hơn rõ rệt ở chỉ số F1-score so với độ chính xác. Chẳng hạn, trong Thiết lập 1, F1-score của CGB-Net cao hơn AnpoNet hơn 10 điểm phần trăm, trong khi độ chính xác chỉ tăng khoảng 4 điểm phần trăm. Xu hướng tương tự cũng xuất hiện trong Thiết lập 2. Điều này cho thấy lợi thế của CGB-Net không chỉ nằm ở việc tăng tỷ lệ dự đoán đúng tổng thể, mà quan trọng hơn là cải thiện khả năng phân biệt giữa các lớp tư thế.

Trong bài toán phân loại tư thế ngủ với nhiều lớp có sự khác biệt góc nhỏ, một số tư thế có thể dễ phân biệt hơn trong khi các tư thế khác dễ bị nhầm lẫn do đặc điểm tín hiệu tương tự. Trong những trường hợp như vậy, một mô hình có thể đạt độ chính xác tương đối cao bằng cách dự đoán tốt các lớp phổ biến hoặc dễ nhận dạng, nhưng vẫn có F1-score thấp nếu các lớp khó bị phân loại sai nhiều hơn. Việc F1-score của CGB-Net tăng đáng kể cho thấy kiến trúc này giúp giảm tỷ lệ nhầm lẫn giữa các tư thế có đặc trưng tương tự.

Sự cải thiện này có thể được giải thích bởi thiết kế kiến trúc phân cấp của CGB-Net. Trong khi AnpoNet sử dụng một lớp LSTM duy nhất để mô hình hóa phụ thuộc thời gian, CGB-Net kết hợp GRU và BiLSTM trong một cấu trúc hai tầng. GRU đóng

vai trò nắm bắt các phụ thuộc thời gian ngắn và trung hạn với chi phí tham số thấp, trong khi BiLSTM khai thác ngữ cảnh hai chiều của chuỗi tín hiệu. Nhờ đó, mô hình có thể học được các biểu diễn thời gian đa thang (multi-scale temporal representations), cho phép khai thác hiệu quả hơn các biến thiên nhỏ trong tín hiệu gia tốc, chẳng hạn như các khác biệt tinh tế trong định hướng trọng lực hoặc dao động do hô hấp.

Tổng thể, các kết quả trong Bảng 3.1 và Bảng 3.2 cho thấy CGB-Net đạt được sự cân bằng tốt giữa năng lực biểu diễn và khả năng tổng quát hóa. Việc tích hợp các mô-đun CNN, GRU và BiLSTM trong một kiến trúc phân cấp giúp mô hình khai thác hiệu quả cả các đặc trưng cục bộ và cấu trúc thời gian dài hạn của tín hiệu gia tốc kế, qua đó cải thiện cả độ chính xác tổng thể lẫn sự cân bằng giữa các lớp trong bài toán phân loại tư thế ngủ.

Bảng 3.3: So sánh hiệu suất trung bình và chi phí tính toán của các mô hình trên tập kiểm thử.

Mô hình	Độ chính xác (%)	F1-score (%)	Số tham số	FLOPS (M)	Bộ nhớ(kB)
MLP	82.38 ± 4.14	73.61 ± 4.01	22,124	0.0433	116
1D CNN	84.24 ± 4.98	78.34 ± 8.11	25,436	0.1410	135
1D CNN + BiLSTM	81.98 ± 5.93	75.25 ± 7.51	21,372	1.9530	152
1D CNN + GRU	82.79 ± 5.23	76.20 ± 8.12	16,700	1.4633	134
Rawan et al. (LSTM) [63]	83.91 ± 6.67	73.71 ± 9.14	3,212	0.2806	31
AnpoNet	85.06 ± 5.07	76.72 ± 8.02	1,916	0.1603	34
CGB-Net (Đề xuất)	87.60 ± 4.19	83.38 ± 5.51	9,532	1.72	88

Bảng 3.3 cho thấy CGB-Net đạt hiệu suất phân loại trung bình cao nhất trong số các mô hình được đánh giá, với độ chính xác $87,60\% \pm 4,19\%$ và Macro F1-score $83,38\% \pm 5,51\%$. So với AnpoNet, CGB-Net cải thiện độ chính xác trung bình khoảng 2,54 điểm phần trăm và Macro F1-score khoảng 6,66 điểm phần trăm. Mức cải thiện lớn hơn ở Macro F1-score cho thấy lợi ích của CGB-Net không chỉ nằm ở việc tăng tỷ lệ dự đoán đúng tổng thể mà còn ở khả năng duy trì hiệu quả phân loại cân bằng hơn giữa các lớp tư thế. Điều này đặc biệt có ý nghĩa đối với bài toán 12 lớp, trong đó một số tư thế trung gian có đặc trưng gia tốc gần nhau và khó phân biệt hơn so với các tư thế cơ bản.

Tuy nhiên, sự cải thiện về hiệu suất đi kèm với mức tăng đáng kể về độ phức tạp tính toán. CGB-Net sử dụng 9.532 tham số, cao hơn khoảng 4,98 lần so với 1.916 tham số của AnpoNet. Chi phí tính toán tăng từ 0,1603 M FLOPs của AnpoNet lên 1,72 M FLOPs của CGB-Net, tương đương khoảng 10,7 lần, trong khi mức sử dụng bộ nhớ tăng từ 34 kB lên 88 kB, tương đương khoảng 2,6 lần. Do đó, CGB-Net không

thể được xem là vượt trội tuyệt đối so với AnpoNet trên tất cả các tiêu chí. Thay vào đó, hai mô hình thể hiện hai điểm vận hành khác nhau trong sự đánh đổi giữa hiệu suất phân loại và yêu cầu tài nguyên.

Trong nhóm các mô hình có độ phức tạp thấp, AnpoNet thể hiện ưu thế rõ rệt về tính gọn nhẹ với chỉ 1.916 tham số, 0,1603 M FLOPs và mức sử dụng bộ nhớ 34 kB, trong khi vẫn đạt độ chính xác 85,06% và Macro F1-score 76,72%. Mô hình LSTM của Rawan et al. cũng có quy mô tương đối nhỏ với 3.212 tham số, 0,2806 M FLOPs và mức sử dụng bộ nhớ 31 kB, nhưng đạt độ chính xác 83,91% và Macro F1-score 73,71%. Các kết quả này cho thấy việc giảm kích thước mô hình và chi phí tính toán có thể đạt được với mức hiệu suất phân loại tương đối tốt, nhưng vẫn tồn tại khoảng cách nhất định so với CGB-Net, đặc biệt ở Macro F1-score.

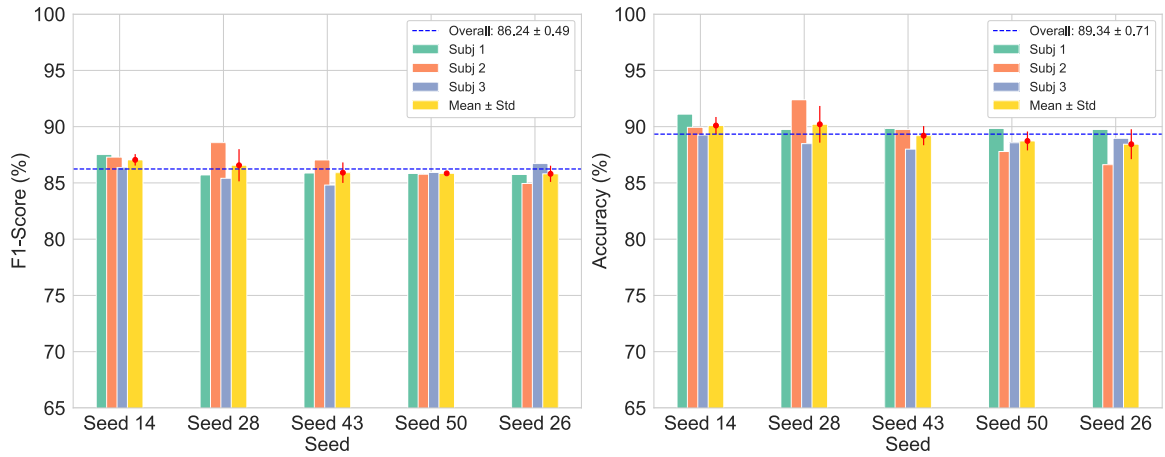
Đối với các kiến trúc cơ sở còn lại, MLP và các biến thể CNN-RNN đạt độ chính xác trung bình trong khoảng 81,98%–84,24%, trong khi Macro F1-score dao động từ 73,61% đến 78,34%. Đáng chú ý, việc bổ sung riêng một mô-đun hồi quy vào 1D CNN không mặc nhiên dẫn đến cải thiện hiệu suất. Chẳng hạn, 1D CNN đạt độ chính xác 84,24% và Macro F1-score 78,34%, trong khi 1D CNN+BiLSTM đạt lần lượt 81,98% và 75,25%, còn 1D CNN+GRU đạt 82,79% và 76,20%. Kết quả này cho thấy lợi ích của CGB-Net không đơn thuần bắt nguồn từ việc tăng số lượng các lớp hồi quy, mà liên quan đến cách tổ chức phân cấp và vai trò hỗ trợ giữa các thành phần 1D CNN, GRU và BiLSTM.

Trong CGB-Net, 1D CNN trước hết trích xuất các đặc trưng cục bộ từ tín hiệu gia tốc; GRU tiếp tục mô hình hóa và tổng hợp các phụ thuộc thời gian ngắn và trung hạn với cấu trúc cổng tương đối gọn; sau đó BiLSTM khai thác ngữ cảnh theo cả hai chiều trên chuỗi đặc trưng đã được xử lý. Việc sắp xếp này tạo ra một quá trình mô hình hóa thời gian theo nhiều mức thay vì sử dụng một thành phần hồi quy đơn lẻ. Kết quả thực nghiệm cho thấy cách tổ chức đó đi kèm với mức cải thiện rõ hơn về Macro F1-score, nhưng đồng thời làm tăng số lượng tham số, FLOPs và yêu cầu bộ nhớ.

Tổng hợp các kết quả trên, CGB-Net phù hợp với kịch bản ưu tiên hiệu suất phân loại và khả năng phân biệt cân bằng giữa các lớp, trong khi AnpoNet phù hợp hơn với các kịch bản có ràng buộc chặt về tài nguyên tính toán và bộ nhớ. Vì vậy, kết quả của Bảng 3.3 nên được diễn giải như một sự đánh đổi giữa năng lực biểu diễn và chi phí tính toán, thay vì xem một kiến trúc là tối ưu cho mọi điều kiện triển khai. Kết

quả này đồng thời làm rõ vai trò của thiết kế phân cấp trong CGB-Net: việc kết hợp GRU và BiLSTM mang lại lợi ích về hiệu suất, đặc biệt đối với Macro F1-score, nhưng lợi ích đó đạt được với chi phí tài nguyên cao hơn so với kiến trúc gọn nhẹ AnpoNet.

3.5.2 Đánh giá với Seed ngẫu nhiên



Hình 3.2: Độ chính xác và F1-Score của CGB-Net qua các lần khởi tạo seed ngẫu nhiên.

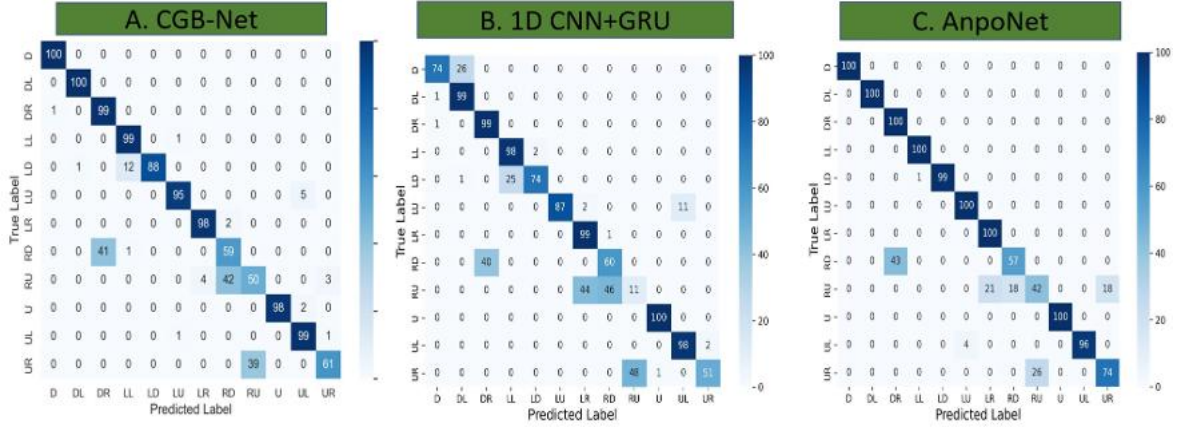
Để minh họa hành vi khái quát hóa của mô hình đề xuất, nghiên cứu này trình bày kết quả đại diện của CGB-Net trong một thiết lập đánh giá cụ thể (Thiết lập 1), trong đó ba đối tượng được chọn làm tập kiểm tra, trong khi các đối tượng còn lại được sử dụng cho huấn luyện và xác thực. Thiết lập này được lặp lại với năm seed ngẫu nhiên khác nhau (14, 28, 43, 50 và 26) nhằm quan sát mức độ nhạy của mô hình đối với sự thay đổi trong quá trình khởi tạo và phân chia dữ liệu.

Hiệu suất của CGB-Net trong thiết lập minh họa này được thể hiện trong Hình 3.2. Biểu đồ con bên trái trình bày F1-score, trong khi biểu đồ bên phải trình bày độ chính xác, tương ứng với từng seed và từng đối tượng trong tập kiểm tra. Các thanh màu vàng biểu thị giá trị trung bình $\pm$ độ lệch chuẩn trên ba đối tượng kiểm tra trong mỗi lần thử, trong khi đường chấm màu xanh lam thể hiện giá trị trung bình tổng thể.

Trong thiết lập minh họa này, CGB-Net đạt F1-score trung bình là $86.24\% \pm 0.49$ và độ chính xác trung bình là $89.34\% \pm 0.71$. Độ lệch chuẩn thấp trên các seed cho thấy mô hình thể hiện hành vi ổn định trước sự thay đổi của khởi tạo ngẫu nhiên và lựa chọn đối tượng kiểm tra. Mặc dù chỉ trình bày một thiết lập cụ thể, các kết quả này cung cấp bằng chứng định tính về khả năng học được các biểu diễn không-thời

gian ổn định của CGB-Net trong kịch bản đánh giá độc lập theo đối tượng.

3.5.3 Phân tích ma trận nhầm lẫn



Hình 3.3: Ma trận nhầm lẫn (Confusion Matrix) của CGB-Net, 1D CNN + GRU và AnpoNet

Ma trận nhầm lẫn của ba mô hình được đánh giá (CGB-Net, 1D CNN+GRU và AnpoNet) được minh họa trong Hình 3.3. CGB-Net đạt hiệu suất cao trên phần lớn các lớp tư thế, với tỷ lệ phân loại đúng 100% đối với D và DL; 99% đối với DR, LL và UL; 98% đối với U và LR; 95% đối với LU; và 88% đối với LD. Tuy nhiên, hiệu suất giảm đáng kể ở ba lớp RD, RU và UR, với tỷ lệ phân loại đúng lần lượt là 59%, 50% và 61%. Kết quả này cho thấy độ khó của bài toán không phân bố đồng đều giữa 12 lớp mà tập trung chủ yếu ở một số tư thế trung gian có đặc trưng định hướng tương đối gần nhau.

Đối với lớp RD, CGB-Net phân loại đúng 59% số mẫu, trong khi 41% bị nhầm sang DR. Xu hướng tương tự cũng xuất hiện ở các mô hình đối chứng, trong đó 1D CNN+GRU đạt 60% và AnpoNet đạt 57% trên lớp RD. Hai tư thế DR và RD tương ứng với các góc 300° và 330° , tức chỉ cách nhau 30° . Việc cả ba mô hình đều gặp khó khăn đối với cùng cặp lớp cho thấy đây không chỉ là hạn chế của riêng CGB-Net mà còn có thể phản ánh mức độ chồng lấp của dữ liệu khi hai tư thế có khoảng cách góc nhỏ.

Một cơ sở để lý giải hiện tượng này xuất phát từ bản chất của tín hiệu gia tốc trong các tư thế tương đối tĩnh. Trong điều kiện cơ thể ít chuyển động, tín hiệu đo bởi accelerometer chịu ảnh hưởng đáng kể bởi thành phần trọng lực và hướng tương đối giữa hệ trục cảm biến với cơ thể. Vì vậy, khi hai tư thế có góc nghiêng gần nhau, hình

chiều của vectơ trọng trường lên ba trục cảm biến cũng có thể gần nhau, làm giảm mức độ phân tách giữa các lớp. Các nghiên cứu về nhận dạng hoạt động sử dụng cảm biến gia tốc cũng chỉ ra rằng vị trí cảm biến và tập đặc trưng thu được tại từng vị trí có ảnh hưởng trực tiếp đến khả năng phân biệt các trạng thái hoạt động [122].

Trong thiết kế dữ liệu của luận án, các lớp liên kế được bố trí với khoảng cách 30° . Cụ thể, RU và UR tương ứng với 30° và 60° , U ở 90° ; trong khi DR và RD tương ứng với 300° và 330° , trước khi trở về tư thế R ở 0° . Do đó, RD, RU và UR đều nằm trong các vùng trung gian giữa những hướng tư thế chính. So với các tư thế cơ bản có sự khác biệt định hướng lớn hơn, các lớp trung gian này có khoảng cách hình học nhỏ hơn và có thể dễ bị ảnh hưởng hơn bởi các biến thiên nhỏ trong quá trình thu thập dữ liệu.

Đối với lớp RU, CGB-Net phân loại đúng 50% số mẫu, trong đó khoảng 42% bị nhầm sang RD. Mặc dù RU và RD cách nhau khoảng 60° , cả hai đều thuộc vùng định hướng nghiêng về phía bên phải. Điều này cho thấy khoảng cách góc danh định không phải là yếu tố duy nhất quyết định mức độ phân biệt của tín hiệu. Cách vectơ trọng trường được chiếu lên hệ trục cảm biến thực tế còn phụ thuộc vào hướng gắn cảm biến trên cơ thể. Một sai lệch nhỏ về vị trí hoặc góc xoay của thiết bị có thể làm thay đổi đồng thời các thành phần a_x , a_y và a_z , qua đó làm tăng mức độ chồng lấp giữa các lớp có định hướng gần nhau.

Nhận định này phù hợp với kết quả của Atallah và cộng sự, trong đó vị trí đặt accelerometer được xác định là một yếu tố quan trọng trong bài toán nhận dạng hoạt động. Nghiên cứu cho thấy vị trí cảm biến và các đặc trưng phù hợp cho từng vị trí có vai trò đáng kể đối với khả năng phân biệt các nhóm hoạt động [122]. Vì vậy, trong bài toán phân loại 12 tư thế với độ phân giải góc 30° , một sai lệch nhỏ về cách đặt hoặc hướng của cảm biến có thể gây ảnh hưởng rõ hơn so với bài toán chỉ phân biệt một số tư thế cơ bản.

Bên cạnh vị trí cảm biến, sự khác biệt giữa các đối tượng cũng có thể góp phần làm giảm hiệu suất ở các lớp trung gian. Cùng một tư thế danh định nhưng các đối tượng khác nhau có thể có hình thể, cách tiếp xúc với giường, góc nghiêng thực tế và vị trí tương đối của cảm biến khác nhau. Farrahi và cộng sự, trong một tổng quan hệ thống về các phương pháp học máy sử dụng accelerometer, cho thấy hiệu suất của mô hình có thể chịu ảnh hưởng bởi nhiều yếu tố như vị trí cảm biến, tần số lấy mẫu, kích thước cửa sổ, loại đặc trưng và điều kiện thu thập. Đáng chú ý, các mô hình được hiệu

chuẩn trong điều kiện phòng thí nghiệm không phải lúc nào cũng duy trì được hiệu suất tương đương khi chuyển sang điều kiện tự nhiên, một phần do xuất hiện các hoạt động chưa quan sát và sự khác biệt trong tín hiệu gia tốc [123].

Kết quả này đặc biệt liên quan đến thiết lập đánh giá độc lập theo đối tượng của luận án. Trong thiết lập này, dữ liệu của người được kiểm tra không xuất hiện trong quá trình huấn luyện. Do đó, mô hình không chỉ cần phân biệt 12 lớp tư thế mà còn phải duy trì khả năng phân loại khi đặc điểm tín hiệu của đối tượng mới khác với các đối tượng đã sử dụng để huấn luyện. Đối với các lớp có khoảng cách đặc trưng lớn, sự biến thiên liên đối tượng có thể chưa đủ để làm thay đổi quyết định phân loại; ngược lại, đối với các lớp trung gian có ranh giới gần nhau như RD, RU và UR, những biến thiên tương đối nhỏ cũng có thể làm mẫu dịch chuyển sang vùng đặc trưng của lớp lân cận.

Một giới hạn khác cần được xem xét là hệ thống chỉ sử dụng một cảm biến gia tốc ba trục tại một vị trí trên cơ thể. Cảm biến này cung cấp thông tin tốt về định hướng của vùng thân tại vị trí đặt cảm biến, nhưng không thể mô tả đầy đủ cấu hình không gian của toàn bộ cơ thể. Hai tư thế có thể khác nhau về vai, hông hoặc chi nhưng vẫn tạo ra hướng tương đối gần nhau tại vị trí cảm biến. Do đó, khi tăng độ phân giải từ các tư thế cơ bản lên 12 lớp, lượng thông tin thu được từ một vị trí cảm biến duy nhất có thể chưa đủ để tạo khoảng cách lớn giữa mọi cặp lớp.

Bằng chứng trực tiếp hơn đối với bài toán tư thế ngủ được trình bày trong nghiên cứu của Doheny và cộng sự [60]. Nghiên cứu sử dụng accelerometer ba trục đặt tại vùng ngực và vùng bụng, đồng thời sử dụng dữ liệu polysomnography làm tham chiếu. Bộ phân loại tư thế dựa trên các đặc trưng từ tín hiệu gia tốc và thông tin định hướng đạt ROC-AUC khoảng 0,87 khi phân biệt tư thế nằm ngửa với nhóm nằm nghiêng trái/phải và khoảng 0,94 khi phân biệt trái với phải. Kết quả này cho thấy ngay cả với bài toán chỉ gồm các nhóm tư thế tương đối thô, tín hiệu accelerometer vẫn chưa tạo ra sự phân tách hoàn toàn giữa các lớp.

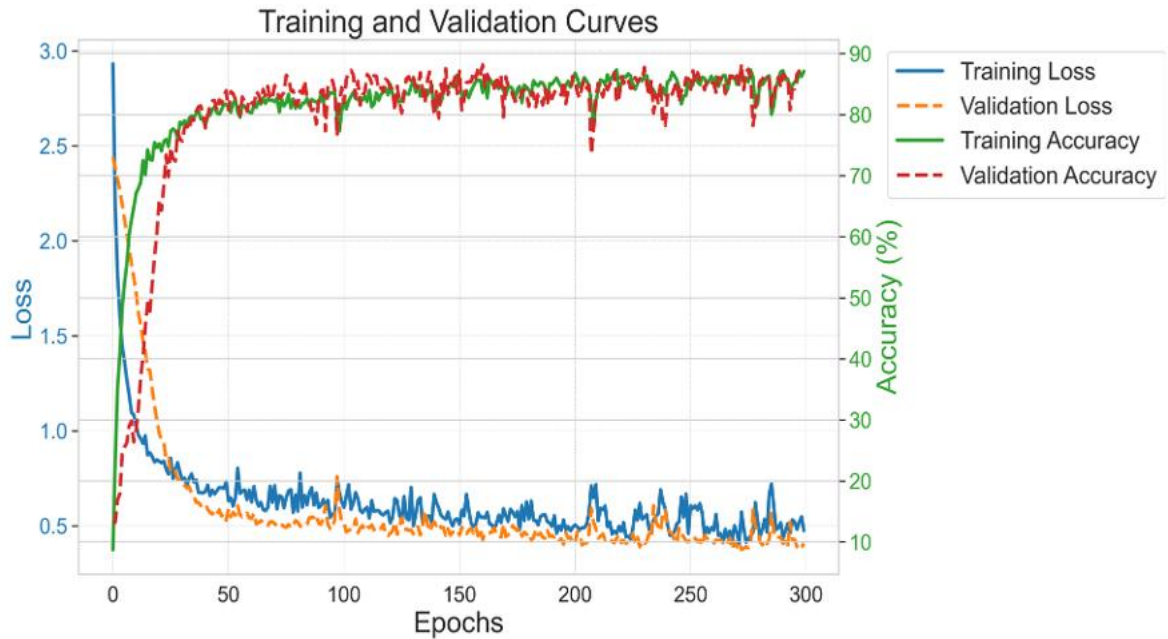
Do đó, khi bài toán được mở rộng thành 12 lớp với khoảng cách danh định chỉ 30° , việc các tư thế trung gian trở nên khó phân biệt hơn là một kết quả có cơ sở. Đối với lớp UR, tỷ lệ phân loại đúng 61% cũng có thể được giải thích theo cơ chế tương tự. UR tương ứng với góc 60° , nằm giữa RU ở 30° và U ở 90° . Vì vậy, đây là một lớp trung gian có thể đồng thời mang đặc trưng của hướng nghiêng phải và hướng nằm ngửa. Khi có biến thiên về góc thực tế của cơ thể hoặc hướng cảm biến, mẫu UR có

thể dịch chuyển về phía một trong các lớp lân cận, làm tăng khả năng phân loại sai.

Từ các kết quả trên, hiệu suất thấp ở RD, RU và UR nhiều khả năng không xuất phát từ một nguyên nhân đơn lẻ mà là kết quả tổng hợp của nhiều yếu tố, bao gồm: (i) khoảng cách góc nhỏ giữa các lớp làm giảm mức độ phân tách của tín hiệu; (ii) sự chồng lấp của các thành phần gia tốc giữa các tư thế có định hướng gần nhau; (iii) biến thiên giữa các đối tượng trong thiết lập đánh giá độc lập theo đối tượng; (iv) sai lệch về vị trí hoặc hướng đặt cảm biến; và (v) giới hạn thông tin khi chỉ sử dụng một accelerometer ba trục tại một vị trí trên cơ thể. Các cơ chế này phù hợp với các kết quả đã được ghi nhận trong văn liệu về nhận dạng hoạt động và tư thế bằng accelerometer [60,122,123].

Tuy nhiên, cần lưu ý rằng các nguyên nhân trên được sử dụng để lý giải kết quả dựa trên đặc điểm hình học của 12 tư thế, dạng nhằm lẫn quan sát được trong ma trận nhầm lẫn và các bằng chứng từ văn liệu; luận án chưa thực hiện các thí nghiệm riêng để định lượng mức đóng góp của từng nguyên nhân. Trong nghiên cứu tiếp theo, các giả thuyết này có thể được kiểm chứng bằng cách ghi nhận liên tục góc tư thế bằng một hệ tham chiếu độc lập, khảo sát có kiểm soát ảnh hưởng của sai lệch vị trí và hướng cảm biến, thực hiện tăng cường dữ liệu mô phỏng sự xoay của hệ trục cảm biến, hoặc bổ sung một vị trí cảm biến thứ hai để đánh giá trực tiếp mức cải thiện đối với các lớp RD, RU và UR.

3.5.4 Quá trình huấn luyện



Hình 3.4: Đường cong hàm mất mát (loss) và độ chính xác của CGB-Net

Hình 3.4 trình bày đồng thời sự biến thiên của hàm mất mát và độ chính xác trên tập huấn luyện và tập xác thực theo số epoch. Trong hình, hàm mất mát (*loss*) được biểu diễn theo trục tung bên trái, trong khi độ chính xác (*accuracy*) được biểu diễn theo trục tung bên phải. Các đường tương ứng với tập huấn luyện và tập xác thực được phân biệt thông qua kiểu đường thể hiện trong phần chú giải. Do *loss* và *accuracy* là hai đại lượng có bản chất và thang đo khác nhau, hai nhóm đường cong được phân tích riêng biệt.

Đối với hàm mất mát, Training Loss giảm nhanh từ khoảng 2,9 trong giai đoạn đầu xuống dưới 1,0 sau một số epoch đầu tiên, sau đó tiếp tục giảm chậm và dao động quanh vùng 0,5 ở giai đoạn cuối. Validation Loss cũng giảm từ khoảng 2,4 xuống vùng xấp xỉ 0,4–0,5 và duy trì tương đối ổn định trong phần lớn quá trình huấn luyện. Xu hướng giảm đồng thời của hai đường loss cho thấy quá trình tối ưu hóa dần tiến tới trạng thái hội tụ.

Đối với độ chính xác, Training Accuracy và Validation Accuracy đều tăng nhanh trong giai đoạn đầu của quá trình huấn luyện. Sau khoảng 40–50 epoch, tốc độ cải thiện giảm dần và cả hai đường chuyển sang vùng tương đối ổn định, chủ yếu dao động trong khoảng 80–87%. Validation Accuracy xuất hiện một số dao động cục bộ tại một số epoch nhưng sau đó nhanh chóng quay trở lại vùng ổn định.

Khoảng cách giữa kết quả trên tập huấn luyện và tập xác thực nhìn chung không có xu hướng tăng dần theo số epoch. Đồng thời, Validation Loss không tăng liên tục trong khi Training Loss tiếp tục giảm. Do đó, trong phạm vi quá trình huấn luyện được quan sát, chưa xuất hiện dấu hiệu quá khớp rõ rệt. Các dao động cục bộ trên tập xác thực phản ánh sự biến thiên trong quá trình tối ưu hóa nhưng không làm thay đổi xu hướng hội tụ chung của mô hình.

3.5.5 Ảnh hưởng của độ dài chuỗi và tỷ lệ chồng lấp

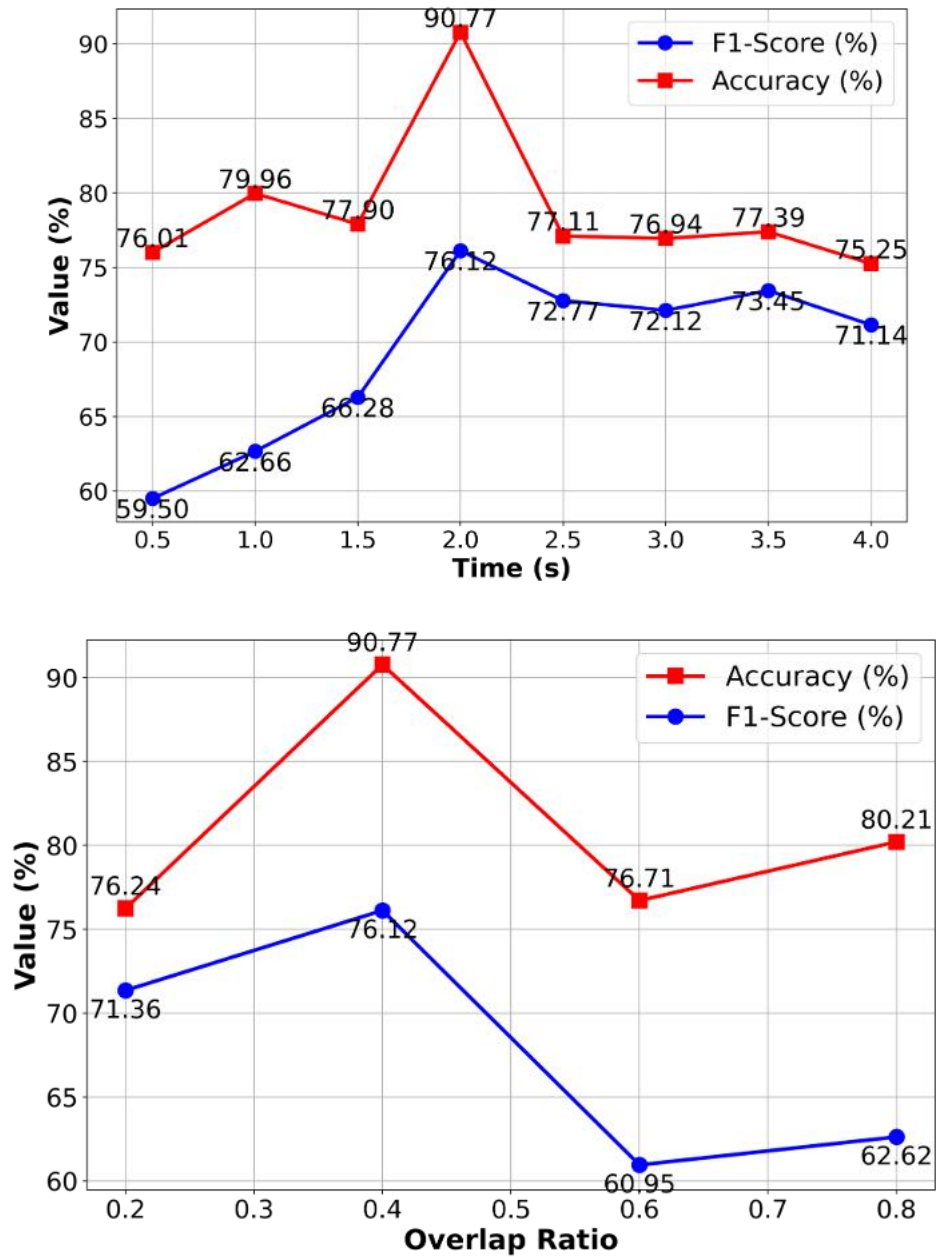
Hình minh họa [3.5](#) cho thấy tác động của việc thay đổi thời lượng cửa sổ từ 0.5 đến 4.0 giây lên hiệu suất của mô hình, được đánh giá bằng Độ chính xác (%) và F1-score (%). Kết quả trình bày ở đây tương ứng với một Seed duy nhất trong một thiết lập cụ thể và cung cấp cái nhìn chi tiết ở cấp độ trường hợp về hành vi của mô hình.

Khi thời lượng cửa sổ tăng lên, cả hai chỉ số đều cho thấy xu hướng tăng chung, đạt đỉnh điểm ở 2.0 giây, trong đó Độ chính xác đạt 90.77% và F1-score đạt đỉnh điểm ở mức 76.12%. Điều này cho thấy rằng các cửa sổ có độ dài vừa phải cho phép mô hình nắm bắt đủ ngữ cảnh thời gian, từ đó nâng cao khả năng phân loại của nó.

Ở các khoảng thời gian ngắn hơn—cụ thể là 0.5 và 1.0 giây—mô hình chỉ đạt được độ chính xác 76.00% và 79.96%, với F1-score tương ứng là 59.50% và 62.66%, cho thấy hiện tượng thiếu khớp do thông tin thời gian hạn chế. Có một sự cải thiện nhỏ ở 1.5 giây (Độ chính xác: 77.90%, F1: 66.28%), nhưng sự cải thiện hiệu suất đáng kể nhất xảy ra khi tăng cửa sổ lên 2.0 giây.

Vượt quá điểm tối ưu này, hiệu suất bắt đầu giảm sút. Ở 2.5 giây, độ chính xác giảm xuống còn 77.11%, trong khi F1-score giảm xuống còn 72.77%. Xu hướng này tiếp tục diễn ra với các cửa sổ dài hơn: Độ chính xác dao động quanh mức 76.94%, 77.39% và 75.25% lần lượt ở 3.0, 3.5 và 4.0 giây, trong khi F1-score là 72.12%, 73.45% và 71.14%. Những kết quả này cho thấy các cửa sổ quá dài có thể chứa thông tin dư thừa hoặc nhiễu, làm mờ các mẫu thời gian chứa thông tin của tư thế liên quan và dẫn đến hiệu suất giảm, có thể do hiện tượng quá khớp (overfitting). Kết quả từ trường hợp cụ thể này cho thấy 2.0 giây là thời lượng cửa sổ tối ưu—cân bằng giữa độ phân giải thời gian và độ phong phú ngữ cảnh—cho cấu hình cụ thể này. Mặc dù việc khái quát hóa rộng hơn sẽ cần tổng hợp trên nhiều seed và thiết lập khác nhau, trường hợp này nhấn mạnh tính nhạy cảm của mô hình đối với độ dài cửa sổ và tầm quan trọng

của việc lựa chọn độ phân giải thời gian phù hợp.



Hình 3.5: Hiệu suất CGB-Net theo kích thước cửa sổ và tỷ lệ chồng lấp.

Hình 3.5 cũng trình bày hiệu suất của mô hình dưới các tỷ lệ chồng lấp khác nhau là 0.2, 0.4, 0.6 và 0.8. Độ chính xác cao nhất đạt được ở tỷ lệ chồng lấp 0.4, cho thấy mức độ chồng lấp vừa phải mang lại sự cân bằng tốt nhất giữa tính liên tục theo thời gian và sự dư thừa dữ liệu.

F1-score cũng thay đổi theo tỷ lệ chồng lấp, cho thấy sự giảm mạnh ở tỷ lệ chồng lấp 0.6, điều này có thể là do sự dư thừa dữ liệu quá mức. Ở các tỷ lệ chồng lấp thấp hơn (ví dụ: 0.2), sự gián đoạn theo thời gian có thể cản trở khả năng nhận diện các chuyển đổi hoặc mẫu một cách hiệu quả của mô hình. Những kết quả này nhấn

mạnh tầm quan trọng của việc lựa chọn tỷ lệ chồng lấp phù hợp để đảm bảo hiệu suất mạnh mẽ mà không làm giảm hiệu quả dữ liệu hoặc tạo ra sự tương quan quá mức giữa các mẫu.

3.6 Kết luận chương 3

Chương này đã trình bày CGB-Net, một kiến trúc học sâu lai được đề xuất nhằm cải thiện khả năng phân loại chi tiết mười hai tư thế ngủ từ dữ liệu gia tốc kế ba trục gắn tại vùng bụng. Khác với Chương 2, trọng tâm của chương này không nằm ở tối ưu hóa tài nguyên hay thiết kế hướng thiết bị biên, mà tập trung vào việc nâng cao khả năng tổng quát hóa và độ ổn định của mô hình trong các kịch bản đánh giá độc lập theo đối tượng nghiêm ngặt hơn.

CGB-Net được xây dựng trên cơ sở tích hợp có chủ đích các mô-đun học sâu bổ sung, bao gồm 1D-CNN để trích xuất các đặc trưng cục bộ theo thời gian, GRU để mô hình hóa hiệu quả các phụ thuộc tuần tự với chi phí tham số vừa phải, và BiLSTM để khai thác ngữ cảnh thời gian hai chiều. Thiết kế phân cấp này cho phép mô hình học được các biểu diễn không-thời gian phong phú hơn, qua đó giảm độ nhạy đối với sự biến thiên giữa các đối tượng và các sai khác nhỏ trong cách thực hiện tư thế.

Các kết quả thực nghiệm, được đánh giá thông qua giao thức loại trừ đối tượng và kiểm định chéo, cho thấy CGB-Net đạt độ chính xác trung bình 87.60% và F1-score trung bình 83.38%, vượt trội so với các mô hình cơ sở được xem xét. Quan trọng hơn, mô hình thể hiện hành vi ổn định trên các seed ngẫu nhiên và các cách chia dữ liệu khác nhau, phản ánh khả năng học được các biểu diễn có tính khái quát cao thay vì phụ thuộc vào các đặc trưng đặc thù của từng nhóm đối tượng.

Những kết quả đạt được trong chương này bổ sung và mở rộng các phát hiện của Chương 2, đồng thời đặt nền tảng cho các nghiên cứu tiếp theo hướng tới các kịch bản triển khai thực tế và các hệ thống giám sát giấc ngủ có khả năng thích ứng cao.

KẾT LUẬN VÀ KIẾN NGHỊ

Luận án tập trung nghiên cứu bài toán phân loại chi tiết 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục đặt tại vùng bụng. Khác với nhiều nghiên cứu chỉ xem xét một số tư thế ngủ cơ bản, bài toán trong luận án bao gồm cả các tư thế cơ bản và các tư thế trung gian có sự khác biệt về góc nghiêng cơ thể.

Trong phạm vi thực nghiệm của luận án, các mô hình được xây dựng và đánh giá trên dữ liệu tư thế ngủ tĩnh được thu thập trong điều kiện có kiểm soát. Giao thức đánh giá độc lập theo đối tượng được sử dụng nhằm kiểm tra khả năng phân loại trên những người không xuất hiện trong quá trình huấn luyện. Vì vậy, các kết luận của luận án tập trung vào hiệu quả phân loại, khả năng tổng quát hóa giữa các đối tượng trong phạm vi hai bộ dữ liệu được xây dựng và các đặc trưng về độ phức tạp tính toán của các mô hình đề xuất.

Với mục tiêu nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục, luận án có hai đóng góp chính như sau:

- Thứ nhất, luận án đề xuất mô hình AnpoNet nhằm phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục với số lượng tham số được kiểm soát. Mô hình AnpoNet được xây dựng trên cơ sở kết hợp khả năng trích xuất đặc trưng cục bộ của mạng nơ-ron tích chập một chiều với khả năng mô hình hóa phụ thuộc thời gian của mạng hồi quy. Hướng tiếp cận này nhằm khai thác đặc trưng không gian–thời gian của tín hiệu gia tốc, đồng thời giảm độ phức tạp mô hình để phù hợp với các thiết bị cảm biến đeo hoặc thiết bị có tài nguyên hạn chế. Kết quả thực nghiệm cho thấy AnpoNet đạt độ chính xác 94,67% và F1-score 92,94% trên tập kiểm tra. Bên cạnh hiệu quả phân loại, mô hình chỉ sử dụng 1.916 tham số, yêu cầu bộ nhớ khoảng 34 kB, có chi phí tính toán 0,1603M FLOPs, thời gian suy luận 0,081 giây cho mỗi lần suy luận. Các kết quả này cho thấy AnpoNet đạt được sự cân bằng giữa độ chính xác phân loại và độ gọn nhẹ của mô hình, qua đó có giá trị đối với định hướng triển khai hệ thống giám sát tư thế ngủ trên thiết bị đeo.

- Thứ hai, luận án đề xuất mô hình CGB-Net nhằm nâng cao độ chính xác và

khả năng tổng quát hóa trong bài toán phân loại 12 tư thế ngủ. Mô hình CGB-Net được thiết kế bằng cách kết hợp ba thành phần gồm mạng nơ-ron tích chập một chiều, GRU và BiLSTM. Trong đó, mạng tích chập một chiều được sử dụng để trích xuất đặc trưng cục bộ của tín hiệu gia tốc, GRU hỗ trợ mô hình hóa quan hệ tuần tự với chi phí tham số hợp lý, còn BiLSTM khai thác ngữ cảnh thời gian hai chiều nhằm tăng cường khả năng phân biệt giữa các tư thế ngủ có đặc trưng gần nhau. Kết quả thực nghiệm cho thấy CGB-Net đạt độ chính xác trung bình 87,60% và F1-score trung bình 83,38% trong các thiết lập đánh giá độc lập theo đối tượng và kiểm định chéo. Ở thiết lập thí nghiệm thứ nhất, CGB-Net đạt độ chính xác 89,34% và F1-score 86,24%, cao hơn AnpoNet khoảng 4,14 điểm phần trăm về độ chính xác và hơn 10 điểm phần trăm về F1-score. Kết quả này cho thấy CGB-Net không chỉ cải thiện tỷ lệ dự đoán đúng tổng thể mà còn nâng cao khả năng phân loại cân bằng giữa các lớp tư thế ngủ, đặc biệt trong điều kiện dữ liệu có sự khác biệt giữa các cá nhân.

Nhìn chung, hai mô hình đề xuất trong luận án có vai trò bổ sung cho nhau. AnpoNet phù hợp với hướng thiết kế mô hình gọn nhẹ, có số lượng tham số và chi phí tính toán thấp, hướng tới triển khai trên thiết bị tài nguyên hạn chế. Trong khi đó, CGB-Net tập trung vào việc tăng cường năng lực biểu diễn đặc trưng không gian-thời gian nhằm cải thiện độ chính xác và khả năng tổng quát hóa. Các kết quả đạt được cho thấy việc sử dụng một cảm biến gia tốc ba trục duy nhất gắn tại vùng bụng là khả thi cho bài toán nhận dạng chi tiết 12 tư thế ngủ. Đây là hướng tiếp cận có ưu điểm về tính đơn giản, chi phí thấp, ít gây bất tiện cho người dùng và phù hợp với các hệ thống theo dõi giấc ngủ tại nhà. Mặc dù đạt được các kết quả khả quan, luận án vẫn còn một số hạn chế cần tiếp tục nghiên cứu. Dữ liệu trong luận án chủ yếu là các tư thế ngủ tĩnh được thu thập trong điều kiện có kiểm soát, trong đó mỗi tư thế được duy trì trong một khoảng thời gian xác định. Thiết lập này giúp bảo đảm tính ổn định của nhãn và sự nhất quán giữa các đối tượng, nhưng chưa phản ánh đầy đủ tính liên tục của hành vi ngủ trong điều kiện tự nhiên. Trong một giấc ngủ qua đêm, người dùng có thể thay đổi tư thế nhiều lần, xuất hiện các giai đoạn chuyển tiếp có thời lượng ngắn, các tư thế trung gian không hoàn toàn trùng với các nhãn rời rạc đã định nghĩa, cũng như các chuyển động của tay, chân, thân mình và sự thay đổi vị trí cảm biến. Vì vậy, hiệu suất đạt được trên dữ liệu tư thế tĩnh chưa thể được diễn giải trực tiếp như hiệu suất của hệ thống trong điều kiện ngủ tự nhiên qua đêm.

Ngoài ra, khả năng triển khai trên thiết bị biên hiện mới được đánh giá thông

qua các chỉ số gián tiếp như số lượng tham số, chi phí tính toán, yêu cầu bộ nhớ và thời gian suy luận. Mặc dù hệ thống thu thập dữ liệu của luận án sử dụng vi điều khiển ESP8266 kết hợp với cảm biến gia tốc ADXL345, ESP8266 hiện chỉ đảm nhiệm chức năng thu nhận, lưu trữ tạm thời và truyền dữ liệu, chưa thực hiện trực tiếp quá trình suy luận của các mô hình AnpoNet hoặc CGB-Net. Do đó, các kết quả hiện tại chưa đủ để khẳng định khả năng triển khai hoàn chỉnh của mô hình trên vi điều khiển hoặc thiết bị đeo công suất thấp trong môi trường thực tế. Trong thời gian tới, hướng nghiên cứu thứ nhất là xây dựng bộ dữ liệu gần với điều kiện sử dụng thực tế hơn. Thay vì yêu cầu người tham gia duy trì từng tư thế cố định trong các phiên ghi nhận, dữ liệu có thể được thu thập liên tục trong các phiên ngủ kéo dài hoặc qua đêm. Tín hiệu gia tốc cần được đồng bộ với một nguồn tham chiếu độc lập, chẳng hạn video hồng ngoại, camera chiều sâu hoặc hệ thống theo dõi giấc ngủ phù hợp, để xác định thời điểm bắt đầu, kết thúc và quá trình chuyển đổi giữa các tư thế. Trên cơ sở đó, dữ liệu có thể được gán nhãn không chỉ cho 12 tư thế ổn định mà còn cho các trạng thái chuyển tiếp giữa các tư thế, các đoạn chuyển động và các khoảng dữ liệu không xác định rõ tư thế.

Một vấn đề cần được nghiên cứu trong thiết lập này là cách định nghĩa và gán nhãn cho tư thế chuyển tiếp. Thay vì gán định mỗi cửa sổ tín hiệu chỉ thuộc duy nhất một lớp tư thế ổn định, có thể xây dựng thêm nhãn “transition” cho các cửa sổ chứa quá trình thay đổi tư thế, hoặc sử dụng chiến lược gán nhãn theo tỷ lệ thời gian của các tư thế xuất hiện trong cửa sổ. Việc thay đổi độ dài cửa sổ và tỷ lệ chồng lấp cũng cần được khảo sát lại, bởi các cửa sổ dài có thể bao gồm đồng thời hai hoặc nhiều trạng thái tư thế, trong khi cửa sổ quá ngắn có thể không cung cấp đủ ngữ cảnh để nhận dạng ổn định. Các mô hình khi đó cần được đánh giá không chỉ bằng độ chính xác và Macro F1-score trên các đoạn tư thế ổn định mà còn thông qua khả năng phát hiện đúng thời điểm chuyển đổi, tỷ lệ nhầm lẫn giữa các tư thế liên kề và độ ổn định của dự đoán theo thời gian.

Song song với việc mở rộng điều kiện thu thập, quy mô và mức độ đa dạng của nhóm tham gia cũng cần được tăng lên. Các nghiên cứu tiếp theo nên xem xét các nhóm tuổi, đặc điểm hình thể và thói quen ngủ khác nhau nhằm đánh giá ảnh hưởng của sự khác biệt giữa các cá nhân đến tín hiệu gia tốc. Đồng thời, cần thực hiện các giao thức đánh giá độc lập theo đối tượng trên quy mô lớn hơn và đánh giá chéo bộ dữ liệu (cross-dataset) khi có bộ dữ liệu độc lập phù hợp. Trong thiết lập này, mô hình có

thể được huấn luyện hoàn toàn trên bộ dữ liệu của nghiên cứu hiện tại và kiểm tra trực tiếp trên một bộ dữ liệu bên ngoài sau khi thống nhất hoặc ánh xạ các nhãn tư thế tương thích. Mức suy giảm Accuracy và Macro F1-score so với đánh giá trong cùng bộ dữ liệu, cùng với ma trận nhầm lẫn theo lớp, có thể được sử dụng để định lượng ảnh hưởng của sự khác biệt về miền dữ liệu. Nếu hiệu suất suy giảm đáng kể, các kỹ thuật như chuẩn hóa theo miền, tăng cường dữ liệu, fine-tuning hoặc domain adaptation có thể được khảo sát nhằm cải thiện khả năng tổng quát hóa.

Ngoài ra, trong các nghiên cứu tiếp theo, việc so sánh hiệu năng giữa các mô hình cần được bổ sung bằng các kiểm định thống kê suy luận phù hợp. Bên cạnh việc báo cáo giá trị trung bình và độ lệch chuẩn qua nhiều lần khởi tạo hoặc nhiều phân hoạch dữ liệu, các kiểm định thống kê có thể được sử dụng để đánh giá liệu sự khác biệt về Accuracy và Macro F1-score giữa các mô hình có ý nghĩa thống kê hay chỉ phản ánh dao động ngẫu nhiên giữa các lần thực nghiệm. Việc lựa chọn kiểm định cần phù hợp với thiết kế thí nghiệm và đặc điểm phân phối của các kết quả thu được.

Hướng nghiên cứu thứ hai là chuyển từ đánh giá độ phức tạp gián tiếp sang triển khai và kiểm chứng trực tiếp trên phần cứng nhúng. Trước hết, mô hình cần được chuyển đổi sang định dạng phù hợp với nền tảng suy luận mục tiêu và tối ưu bằng các kỹ thuật như lượng tử hóa trọng số và kích hoạt, cắt tỉa các tham số có đóng góp thấp hoặc nén mô hình. Đối với lượng tử hóa, có thể khảo sát các biểu diễn số nguyên có độ chính xác thấp hơn, chẳng hạn INT8, và đánh giá mức suy giảm hiệu suất phân loại so với mô hình sử dụng số dấu phẩy động. Mục tiêu của bước này là giảm kích thước mô hình, yêu cầu bộ nhớ và số phép toán trong khi vẫn duy trì hiệu suất phân loại ở mức chấp nhận được.

Sau quá trình tối ưu, mô hình cần được triển khai trực tiếp trên một nền tảng vi điều khiển hoặc thiết bị biên có tài nguyên hạn chế. Khi đó, việc đánh giá không chỉ dựa trên kích thước tham số hoặc FLOPs mà cần đo trực tiếp bộ nhớ Flash dành cho mô hình, SRAM cực đại trong quá trình suy luận, thời gian xử lý cho mỗi cửa sổ dữ liệu và mức tiêu thụ năng lượng thực tế. Đặc biệt, bộ nhớ dành cho các tensor kích hoạt trung gian, bộ đệm dữ liệu, ngăn xếp chương trình và các thành phần phục vụ truyền thông cũng cần được tính đến, bởi kích thước tham số của mô hình không phản ánh đầy đủ yêu cầu bộ nhớ khi suy luận trên phần cứng nhúng.

Cuối cùng, mô hình đã triển khai cần được tích hợp với chuỗi xử lý trực tuyến gồm thu nhận tín hiệu từ cảm biến, phân đoạn cửa sổ, tiền xử lý, suy luận và truyền

hoặc lưu trữ kết quả. Thử nghiệm liên tục trong thời gian dài sẽ cho phép đánh giá khả năng đáp ứng thời gian thực, tính ổn định của hệ thống, mức tiêu thụ pin và ảnh hưởng của truyền thông không dây đến thời gian hoạt động. Trên cơ sở các đánh giá này, có thể xác định rõ hơn sự đánh đổi giữa hiệu suất phân loại, độ phức tạp mô hình và mức tiêu thụ tài nguyên, từ đó lựa chọn kiến trúc phù hợp cho từng kịch bản triển khai thực tế.

DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ

- [CT1] Hoang-Dieu Vu, Duc-Nghia Tran, Huy-Hieu Pham, Dinh-Dat Pham, Khanh-Ly Can, To-Hieu Dao, and Duc-Tan Tran, “Human sleep position classification using a lightweight model and acceleration data,” *Sleep and Breathing*, vol. 29, Article 95, 2025. **(SCIE, Scopus Q1, IF: 2.4)**.
- [CT2] Hoang-Dieu Vu, Duc-Nghia Tran, Quang-Tu Pham, Ngoc-Linh Nguyen, and Duc-Tan Tran, “CGB-Net: A Novel Convolutional Gated Bidirectional Network for Enhanced Sleep Posture Classification,” *CMC: Computers, Materials & Continua*, vol. 85, no. 2, pp. 2819–2835, 2025. **(SCIE indexed, Scopus Q2, IF: 1.4)**.
- [CT3] Hoang-Dieu Vu, Duc-Nghia Tran, Pham Quang Huy, and Duc-Tan Tran, “Hy-PoNet: Fine-Grained Sleep Posture Recognition from a Single Abdominal Accelerometer,” *REV Journal on Electronics and Communications*, vol. 15, no. 3, p. 80, 2025.
- [CT4] Hoang-Dieu Vu, Duc-Nghia Tran, Khanh-Ly Can, To-Hieu Dao, Dinh-Dat Pham, and Duc-Tan Tran, “Enhancing sleep postures classification by incorporating acceleration sensor and LSTM model,” in *Proceedings of the 2023 IEEE Statistical Signal Processing Workshop (SSP)*, 2023, pp. 661–665. **(Scopus)**.
- [CT5] Hoang-Dieu Vu, Duc-Nghia Tran, Quang-Tu Pham, and Duc-Tan Tran, “GA-SPARF: Genetic algorithm-tuned random forest for sleep position monitoring using accelerometer,” *Sleep and Breathing*, vol. 30, p. 78, 2026, doi: 10.1007/s11325-026-03584-4. **(SCIE, Scopus Q1, IF: 2.4)**.

TÀI LIỆU THAM KHẢO

- [1] A. Tate, J. Walsh, V. Kurup, B. Shenoy, D. Mann, C. Freakley, P. Eastwood, and P. Terrill, “An emerging technology for the identification and characterization of postural-dependent obstructive sleep apnea,” *Journal of Clinical Sleep Medicine*, vol. 16, no. 2, pp. 309–318, Feb. 2020.
- [2] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, “1d convolutional neural networks and applications: A survey,” *Mechanical Systems and Signal Processing*, vol. 151, p. 107398, Apr. 2021.
- [3] S. M. Ali, S. P. Arjunan, J. Peter, L. Perju-Dumbrava, C. Ding, M. Eller, S. Raghav, P. Kempster, M. A. Motin, P. J. Radcliffe, and D. K. Kumar, “Wearable accelerometer and gyroscope sensors for estimating the severity of essential tremor,” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 194–203, 2024.
- [4] C. Mascolo, “Ai on wearable sensing data for human health and fitness,” *Computer*, vol. 58, no. 10, pp. 102–105, Oct. 2025.
- [5] P. Alinia, A. Samadani, M. Milosevic, H. Ghasemzadeh, and S. Parvaneh, “Pervasive lying posture tracking,” *Sensors*, vol. 20, no. 20, p. 5953, Oct. 2020.
- [6] C. Intongkum, Y. Sasiwat, K. Sengchuai, A. Booranawong, and P. Phukpattaranont, “Monitoring and classification of human sleep postures, seizures, and falls from bed using three-axis acceleration signals and machine learning,” *SN Computer Science*, vol. 5, no. 1, Dec. 2023.
- [7] C. He, Z. Fang, S. Liu, H. Wu, X. Li, Y. Wen, and J. Lin, “A smart detection method for sleep posture based on a flexible sleep monitoring belt and vital sign signals,” *Heliyon*, vol. 10, no. 11, p. e31839, Jun. 2024.

- [8] H. Pickenbrock, V. U. Ludwig, and A. Zapf, “Support pressure distribution for positioning in neutral versus conventional positioning in the prevention of decubitus ulcers: a pilot study in healthy participants,” *BMC Nursing*, vol. 16, no. 1, Oct. 2017.
- [9] A. Khan, C. Kim, J.-Y. Kim, A. Aziz, and Y. Nam, “Sleep posture classification using rgb and thermal cameras based on deep learning model,” *Computer Modeling in Engineering amp; Sciences*, vol. 140, no. 2, pp. 1729–1755, 2024.
- [10] S. C. Liew and T. Aung, “Sleep deprivation and its association with diseases- a review,” *Sleep Medicine*, vol. 77, pp. 192–204, Jan. 2021.
- [11] A. Menon and M. Kumar, “Influence of body position on severity of obstructive sleep apnea: A systematic review,” *ISRN Otolaryngology*, vol. 2013, pp. 1–7, Oct. 2013.
- [12] T. Saldiran, Kara, S. K. Yikilmaz, and M. Durgun, “Influence of body posture and apnea severity on the tone and elasticity of upper airway muscles in awake patients with obstructive sleep apnea: A cross-sectional study,” *Journal of Manipulative and Physiological Therapeutics*, vol. 45, no. 5, pp. 365–377, Jun. 2022.
- [13] T. Penzel, M. Möller, H. F. Becker, L. Knaack, and J.-H. Peter, “Effect of sleep position and sleep stage on the collapsibility of the upper airways in patients with sleep apnea,” *Sleep*, vol. 24, no. 1, pp. 90–95, Jan. 2001.
- [14] T. Leppänen, J. Töyräs, A. Muraja-Murro, S. Kupari, P. Tiihonen, E. Mervaala, and A. Kulkas, “Length of individual apnea events is increased by supine position and modulated by severity of obstructive sleep apnea,” *Sleep Disorders*, vol. 2016, pp. 1–13, 2016.
- [15] M. Strohm, A. Daboul, A. Obst, A. Weihs, C.-J. Busch, T. Bremert, J. Fanghänel, T. Ivanovska, I. Fietze, T. Penzel, R. Ewert, and M. Krüger, “Association between sleep position, obesity, and obstructive sleep apnea severity,” *Journal of Personalized Medicine*, vol. 14, no. 11, p. 1087, Nov. 2024.

- [16] K.-I. Lee and J. H. Choi, "Positional therapy for obstructive sleep apnea: Therapeutic modalities and clinical effects," *Sleep Medicine Research*, vol. 14, no. 3, pp. 129–134, Sep. 2023.
- [17] L. Li, M. Shi, D. M. Umbach, K. Bricker, and Z. Fan, "Sex- and age-differences in supine positional obstructive sleep apnea in children and adults," *Sleep and Breathing*, vol. 29, no. 1, Feb. 2025.
- [18] Y. Gao, S. Zhu, W. Li, and Y. Lai, "Comparative efficacy of sleep positional therapy, oral appliance therapy, and cpap in obstructive sleep apnea: a meta-analysis of mean changes in key outcomes," *Frontiers in Medicine*, vol. 12, Feb. 2025.
- [19] S. C. Ha, H. W. Hirai, and K. K. Tsoi, "Comparison of positional therapy versus continuous positive airway pressure in patients with positional obstructive sleep apnea: A meta-analysis of randomized trials," *Sleep Medicine Reviews*, vol. 18, no. 1, pp. 19–24, Feb. 2014.
- [20] H. Barnes, B. A. Edwards, S. A. Joosten, M. T. Naughton, G. S. Hamilton, and E. Dabscheck, "Positional modification techniques for supine obstructive sleep apnea: A systematic review and meta-analysis," *Sleep Medicine Reviews*, vol. 36, pp. 107–115, Dec. 2017.
- [21] S. A. Joosten, B. A. Edwards, A. Wellman, A. Turton, E. M. Skuza, P. J. Berger, and G. S. Hamilton, "The effect of body position on physiological factors that contribute to obstructive sleep apnea," *Sleep*, vol. 38, no. 9, pp. 1469–1478, Sep. 2015.
- [22] T. Ingman, "Cephalometric comparison of pharyngeal changes in subjects with upper airway resistance syndrome or obstructive sleep apnoea in upright and supine positions," *The European Journal of Orthodontics*, vol. 26, no. 3, pp. 321–326, Jun. 2004.
- [23] L. Messineo, S. Joosten, and E. Perger, "Mechanisms relating to sleeping position to the endotypes of sleep disordered breathing," *Current Opinion in Pulmonary Medicine*, vol. 29, no. 6, pp. 543–549, Aug. 2023.

- [24] D. M. Simadibrata, E. Lesmana, B. R. Amangku, M. P. Wardoyo, and M. Simadibrata, “Left lateral decubitus sleeping position is associated with improved gastroesophageal reflux disease symptoms: A systematic review and meta-analysis,” *World Journal of Clinical Cases*, vol. 11, no. 30, pp. 7329–7336, Oct. 2023.
- [25] J. M. Schuitenmaker, M. van Dijk, R. A. Oude Nijhuis, A. J. Smout, and A. J. Bredenoord, “Associations between sleep position and nocturnal gastroesophageal reflux: A study using concurrent monitoring of sleep position and esophageal ph and impedance,” *American Journal of Gastroenterology*, vol. 117, no. 2, pp. 346–351, Dec. 2021.
- [26] J. M. Schuitenmaker, T. Kuipers, R. A. Oude Nijhuis, M. P. Schijven, A. J. M. Smout, P. Fockens, and A. J. Bredenoord, “Sleep positional therapy for nocturnal gastroesophageal reflux: A double-blind, randomized, sham-controlled trial,” *Clinical Gastroenterology and Hepatology*, vol. 20, no. 12, pp. 2753–2762.e2, Dec. 2022.
- [27] R. M. Khoury, L. Camacho-Lobato, P. O. Katz, M. A. Mohiuddin, and D. O. Castell, “Influence of spontaneous sleep positions on nighttime recumbent reflux in patients with gastroesophageal reflux disease,” *American Journal of Gastroenterology*, vol. 94, no. 8, pp. 2069–2073, Aug. 1999.
- [28] E. Person, C. Rife, J. Freeman, A. Clark, and D. O. Castell, “A novel sleep positioning device reduces gastroesophageal reflux: A randomized controlled trial,” *Journal of Clinical Gastroenterology*, vol. 49, no. 8, pp. 655–659, Sep. 2015.
- [29] S. Allampati, R. Lopez, P. N. Thota, M. Ray, S. Birgisson, and S. L. Gabbard, “Use of a positional therapy device significantly improves nocturnal gastroesophageal reflux symptoms: Short-term improvement of nocturnal GERD symptoms,” *Diseases of the Esophagus*, Sep. 2016.
- [30] A. Oksenberg and D. Silverberg, “The effect of body posture on sleep-related breathing disorders: facts and therapeutic implications,” *Sleep Medicine Reviews*, vol. 2, no. 3, pp. 139–162, 8 1998.
- [31] D. Alfaiate, A. Pereira, P. Guerra, and N. Pinto, “Body posture as a factor determining sleep quality in patients using non-invasive ventilation,” *Sleep and Breathing*, vol. 27, no. 6, pp. 2437–2442, Mar. 2023.

- [32] K. Zhu, T. Bradley, M. Patel, and H. Alshaer, "Influence of head position on obstructive sleep apnea severity," *Sleep and Breathing*, vol. 21, no. 4, pp. 821–828, Jun. 2017.
- [33] Y. Saini, A. Rai, and S. Sen, "Relationship between sleep posture and low back pain: A systematic review," *Musculoskeletal Care*, vol. 23, no. 2, p. 23, May 2025.
- [34] C. He, S. Liu, Z. Fang, H. Wu, M. Liang, S. Deng, and J. Lin, "A novel detection method for heart rate variability and sleep posture based on a flexible sleep monitoring belt," *IEEE Sensors Journal*, vol. 25, no. 3, pp. 5178–5191, Feb. 2025.
- [35] D. Hu, W. Gao, K. Ang, M. Hu, G. Chuai, and R. Huang, "Smart sleep monitoring: Sparse sensor-based spatiotemporal cnn for sleep posture detection," *Sensors*, vol. 24, no. 15, p. 4833, Jul. 2024.
- [36] D. Chakupurakal, D. Liendo, D. Hinds, D. Andry, D. Santibanez, and D. Cherynshev, "1151 lateral head position improves supine predominant obstructive sleep apnea," *Sleep*, vol. 41, no. suppl₁, pp. A425 – –A425, Apr. 2018.
- [37] A. Tate, V. Kurup, B. Shenoy, C. Freakley, P. Eastwood, J. Walsh, and P. Terrill, "Influence of head flexion and rotation on obstructive sleep apnea severity during supine sleep," *Journal of Sleep Research*, vol. 30, no. 5, p. 30, Feb. 2021.
- [38] I. K. A. Priastana, L. K. C. Dewi, and D. P. H. Kusumaningtiyas, "Effectiveness of the 30-degree lateral sleeping position control pillow® in enhancing sleep quality among older adults with stroke: A quasi-experimental study," *Nursing and Midwifery Studies*, vol. 14, no. 3, pp. 192–200, 2025. [Online]. Available: https://nmsjournal.kaums.ac.ir/article_227367.html
- [39] N. H. Do, D. Y. Kim, J.-H. Kim, J. H. Choi, S. Y. Joo, N. K. Kang, and Y. S. Baek, "Effects of a continuous lateral turning device on pressure relief," *Journal of Physical Therapy Science*, vol. 28, no. 2, pp. 460–466, 2016.
- [40] Z. Hussain, Q. Z. Sheng, W. E. Zhang, J. Ortiz, and S. Pouriyeh, "Non-invasive techniques for monitoring different aspects of sleep: A comprehensive review," *ACM Transactions on Computing for Healthcare*, vol. 3, no. 2, pp. 1–26, Mar. 2022.

- [41] M. DE ZAMBOTTI, N. CELLINI, A. GOLDSTONE, I. M. COLRAIN, and F. C. BAKER, “Wearable sleep technology in clinical and research settings,” *Medicine and Science in Sports and Exercise*, vol. 51, no. 7, pp. 1538–1557, Feb. 2019.
- [42] Z. Li, Y. Zhou, and G. Zhou, “A dual fusion recognition model for sleep posture based on air mattress pressure detection,” *Scientific Reports*, vol. 14, no. 1, May 2024.
- [43] M. L. Hoang, G. Matrella, D. Giannetto, P. Craparo, and P. Ciampolini, “Benchmarking accelerometer and cnn-based vision systems for sleep posture classification in healthcare applications,” *Sensors*, vol. 25, no. 12, p. 3816, Jun. 2025.
- [44] S. Z. Gurbuz and M. G. Amin, “Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring,” *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 16–28, Jul. 2019.
- [45] H. Hong, L. Zhang, H. Zhao, H. Chu, C. Gu, M. Brown, X. Zhu, and C. Li, “Microwave sensing and sleep: Noncontact sleep-monitoring technology with microwave biomedical radar,” *IEEE Microwave Magazine*, vol. 20, no. 8, pp. 18–29, Aug. 2019.
- [46] M. Piriya-jitakonkij, P. Warin, P. Lakhan, P. Leelaarporn, N. Kumchaiseemak, S. Suwanajakorn, T. Pianpanit, N. Niparnan, S. C. Mukhopadhyay, and T. Wilaiprasitporn, “Sleepposenet: Multi-view learning for sleep postural transition recognition using uwb,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 4, pp. 1305–1314, Apr. 2021.
- [47] S. Liu and S. Ostadabbas, “Seeing under the cover: A physics guided learning approach for in-bed pose estimation,” pp. 236–245, Oct. 2019.
- [48] X. Liu, J. Cao, S. Tang, and J. Wen, “Wi-sleep: Contactless sleep monitoring via wifi signals,” pp. 346–355, Dec. 2014.
- [49] J. Liu, Y. Chen, Y. Wang, X. Chen, J. Cheng, and J. Yang, “Monitoring vital signs and postures during sleep using wifi signals,” *IEEE Internet of Things Journal*, vol. 5, no. 3, pp. 2071–2084, Jun. 2018.
- [50] S. Yue, Y. Yang, H. Wang, H. Rahul, and D. Katabi, “Bodycompass: Monitoring sleep posture with wireless signals,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 4, no. 2, pp. 1–25, Jun. 2020.

- [51] S. M. M. Islam and V. M. Lubecke, "Sleep posture recognition with a dual-frequency microwave doppler radar and machine learning classifiers," *IEEE Sensors Letters*, vol. 6, no. 3, pp. 1–4, Mar. 2022.
- [52] J. E. Kiriazi, S. M. M. Islam, O. Boric-Lubecke, and V. M. Lubecke, "Sleep posture recognition with a dual-frequency cardiopulmonary doppler radar," *IEEE Access*, vol. 9, pp. 36 181–36 194, 2021.
- [53] S. Liu and S. Ostadabbas, "A vision-based system for in-bed posture tracking," pp. 1373–1382, Oct. 2017.
- [54] S. Akbarian, G. Delfi, K. Zhu, A. Yadollahi, and B. Taati, "Automated non-contact detection of head and body positions during sleep," *IEEE Access*, vol. 7, pp. 72 826–72 834, 2019.
- [55] T. Grimm, M. Martinez, A. Benz, and R. Stiefelhagen, "Sleep position classification from a depth camera using bed aligned maps," pp. 319–324, Dec. 2016.
- [56] P. Jiang and R. Zhu, "Dual tri-axis accelerometers for monitoring physiological parameters of human body in sleep," *Proceedings of IEEE Sensor*, pp. 1–3, Oct. 2016.
- [57] S. Fallmann, R. van Veen, L. Chen, D. Walker, F. Chen, and C. Pan, "Wearable accelerometer based extended sleep position recognition," *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, Oct. 2017.
- [58] P. Jeng and L.-C. Wang, "An accurate, low-cost, easy-to-use sleep posture monitoring system," *2017 International Conference on Applied System Innovation (ICASI)*, pp. 903–905, May 2017.
- [59] S. Jeon, A. Paul, H. Lee, Y. Bun, and S. H. Son, "sleeps: Sleep position tracking system for screening sleep quality by wristbands," *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 3141–3146, Oct. 2017.
- [60] E. P. Doheny, M. M. Lowery, A. Russell, and S. Ryan, "Estimation of respiration rate and sleeping position using a wearable accelerometer," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC*, pp. 4668–4671, Jul. 2020.

- [61] P.-Y. Jeng, L.-C. Wang, C.-J. Hu, and D. Wu, “A wrist sensor sleep posture monitoring system: An automatic labeling approach,” *Sensors*, vol. 21, no. 1, p. 258, Jan. 2021.
- [62] R. S. Abdulsadig, S. Singh, Z. Patel, and E. Rodriguez-Villegas, “Sleep posture detection using an accelerometer placed on the neck,” *2022 44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2430–2433, Jul. 2022.
- [63] R. S. Abdulsadig and E. Rodriguez-Villegas, “Sleep posture monitoring using a single neck- situated accelerometer: A proof-of-concept,” *IEEE Access*, vol. 11, pp. 17 693–17 706, Jan. 2023.
- [64] L. Fonseca, F. Ribeiro, and J. Metrôlho, “Effects of the number of classes and pressure map resolution on fine-grained in-bed posture classification,” *Computation*, vol. 11, no. 12, p. 239, Dec. 2023.
- [65] Y. Wang and Y. Zhao, “General pre-trained inertial signal feature extraction based on temporal memory fusion,” *Information Fusion*, vol. 123, p. 103274, Nov. 2025.
- [66] F. Xu, X. Gao, and W. Wang, “A human activity recognition model based on deep neural network integrating attention mechanism,” *Scientific Reports*, vol. 15, no. 1, Jul. 2025.
- [67] W. Gomaa and M. A. Khamis, “A perspective on human activity recognition from inertial motion data,” *Neural Computing and Applications*, vol. 35, no. 28, pp. 20 463–20 568, Jul. 2023.
- [68] W. Sousa Lima, E. Souto, K. El-Khatib, R. Jalali, and J. Gama, “Human activity recognition using inertial sensors in a smartphone: An overview,” *Sensors*, vol. 19, no. 14, p. 3213, Jul. 2019.
- [69] M. Saif, T. H. T. M., K. B. Kishor, M. R. A.K.M., A. A. M., S. Mohammad, A. H. K. Muhammad, and A. A. Amin, “Human activity recognition from wearable sensor data using self-attention,” 2020.
- [70] J. Shin, N. Hassan, A. S. M. Miah, and S. Nishimura, “A comprehensive methodological survey of human activity recognition across diverse data modalities,” *Sensors*, vol. 25, no. 13, p. 4028, Jun. 2025.

- [71] H. Haresamudram, I. Essa, and T. Plötz, “Contrastive predictive coding for human activity recognition,” *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–26, Jun. 2021.
- [72] K. Sundararajan, S. Georgievska, B. H. W. te Lindert, P. R. Gehrman, J. Ramautar, D. R. Mazzotti, S. Sabia, M. N. Weedon, E. J. W. van Someren, L. Ridder, J. Wang, and V. T. van Hees, “Sleep classification from wrist-worn accelerometer data using random forests,” *Scientific Reports*, vol. 11, no. 1, Jan. 2021.
- [73] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, “Deep learning for sensor-based activity recognition: A survey,” *Pattern Recognition Letters*, vol. 119, pp. 3–11, Mar. 2019.
- [74] Y. Hao, R. Zheng, and B. Wang, “Invariant feature learning for sensor-based human activity recognition,” *IEEE Transactions on Mobile Computing*, pp. 1–1, 2021.
- [75] N. Bento, J. Rebelo, M. Barandas, A. V. Carreiro, A. Campagner, F. Cabitza, and H. Gamboa, “Comparing handcrafted features and deep neural representations for domain generalization in human activity recognition,” *Sensors*, vol. 22, no. 19, p. 7324, Sep. 2022.
- [76] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [77] H. Ismail Fawaz, G. Forestier, J. Weber, L. Idoumghar, and P.-A. Muller, “Deep learning for time series classification: a review,” *Data Mining and Knowledge Discovery*, vol. 33, no. 4, pp. 917–963, Mar. 2019.
- [78] N. Mohammadi Foumani, L. Miller, C. W. Tan, G. I. Webb, G. Forestier, and M. Salehi, “Deep learning for time series classification and extrinsic regression: A current survey,” *ACM Computing Surveys*, vol. 56, no. 9, pp. 1–45, Apr. 2024.
- [79] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, “Convolutional neural networks for time series classification,” *Journal of Systems Engineering and Electronics*, vol. 28, no. 1, pp. 162–169, Feb. 2017.
- [80] A. Yunita, M. I. Pratama, M. Z. Almuzakki, H. Ramadhan, E. A. P. Akhir, A. B. Firdausiah Mansur, and A. H. Basori, “Performance analysis of neural network architectures for time series forecasting: A comparative study of rnn, lstm, gru, and hybrid models,” *MethodsX*, vol. 15, p. 103462, Dec. 2025.

- [81] I. D. Mienye, T. G. Swart, and G. Obaido, “Recurrent neural networks: A comprehensive review of architectures, variants, and applications,” *Information*, vol. 15, no. 9, p. 517, Aug. 2024.
- [82] T. Can, K. Krishnamurthy, and D. J. Schwab, “Gating creates slow modes and controls phase-space complexity in grus and lstms,” in *Proceedings of Mathematical and Scientific Machine Learning, MSML 2020, 20-24 July 2020, Virtual Conference / Princeton, NJ, USA*, ser. Proceedings of Machine Learning Research, J. Lu and R. A. Ward, Eds., vol. 107. PMLR, 2020, pp. 476–511. [Online]. Available: <http://proceedings.mlr.press/v107/can20a.html>
- [83] F. Ordóñez and D. Roggen, “Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition,” *Sensors*, vol. 16, no. 1, p. 115, Jan. 2016.
- [84] N. Y. Hammerla, S. Halloran, and T. Plötz, “Deep, convolutional, and recurrent models for human activity recognition using wearables,” in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, ser. IJCAI’16. AAAI Press, 2016, p. 1533–1540.
- [85] S. Kiranyaz, T. Ince, R. Hamila, and M. Gabbouj, “Convolutional neural networks for patient-specific ecg classification,” pp. 2608–2611, Aug. 2015.
- [86] S. Kiranyaz, T. Ince, and M. Gabbouj, “Real-time patient-specific ecg classification by 1-d convolutional neural networks,” *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, Mar. 2016.
- [87] —, “Personalized monitoring and advance warning system for cardiac arrhythmias,” *Scientific Reports*, vol. 7, no. 1, Aug. 2017.
- [88] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, and D. J. Inman, “Wireless and real-time structural damage detection: A novel decentralized method for wireless sensor networks,” *Journal of Sound and Vibration*, vol. 424, pp. 158–172, Jun. 2018.
- [89] S. Mekruksavanich and A. Jitpattanakul, “Lstm networks using smartphone data for sensor-based human activity recognition in smart homes,” *Sensors*, vol. 21, no. 5, p. 1636, Feb. 2021.

- [90] Y. Zhang, A. Xiao, T. Zheng, H. Xiao, and R. Huang, "The relationship between sleeping position and sleep quality: A flexible sensor-based study," *Sensors*, vol. 22, no. 16, p. 6220, Aug. 2022.
- [91] S. Saadatnejad, M. Oveisi, and M. Hashemi, "Lstm-based ecg classification for continuous monitoring on personal wearable devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 515–523, Feb. 2020.
- [92] O. Yildirim, U. B. Baloglu, and U. R. Acharya, "A deep learning model for automated sleep stages classification using psg signals," *International Journal of Environmental Research and Public Health*, vol. 16, no. 4, p. 599, Feb. 2019.
- [93] S. Mohsen, "Recognition of human activity using gru deep learning algorithm," *Multimedia Tools and Applications*, vol. 82, no. 30, pp. 47 733–47 749, May 2023.
- [94] J. Zhong, Y. Liu, X. Cheng, L. Cai, W. Cui, and D. Hai, "Gated recurrent unit network for psychological stress classification using electrocardiograms from wearable devices," *Sensors*, vol. 22, no. 22, p. 8664, Nov. 2022.
- [95] C. V. S. and R. E., "A novel deep learning based gated recurrent unit with extreme learning machine for electrocardiogram (ecg) signal recognition," *Biomedical Signal Processing and Control*, vol. 68, p. 102779, Jul. 2021.
- [96] F. Rivas, J. E. Sierra-Garcia, and J. M. Camara, "Comparison of lstm- and gru-type rnn networks for attention and meditation prediction on raw eeg data from low-cost headsets," *Electronics*, vol. 14, no. 4, p. 707, Feb. 2025.
- [97] M. A. Khatun, M. A. Yousuf, S. Ahmed, M. Z. Uddin, S. A. Alyami, S. Al-Ashhab, H. F. Akhdar, A. Khan, A. Azad, and M. A. Moni, "Deep cnn-lstm with self-attention model for human activity recognition using wearable sensor," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 1–16, 2022.
- [98] W. C. Carlos, A. Copetti, L. Bertini, L. B. Moreira, and O. d. S. M. Gomes, "Human activity recognition: an approach 2d cnn-lstm to sequential image representation and processing of inertial sensor data," *AIMS Bioengineering*, vol. 11, no. 4, pp. 527–560, 2024.

- [99] T. Y. Wang, J. Cui, and Y. Fan, “A wearable-based sports health monitoring system using cnn and lstm with self-attentions,” *PLOS ONE*, vol. 18, no. 10, p. e0292012, Oct. 2023.
- [100] A. Jalal, M. A. K. Quaid, S. B. u. d. Tahir, and K. Kim, “A study of accelerometer and gyroscope measurements in physical life-log activities detection systems,” *Sensors*, vol. 20, no. 22, p. 6670, Nov. 2020.
- [101] J. Kim, H. Kim, H. Kim, D. Lee, and S. Yoon, “A comprehensive survey of deep learning for time series forecasting: architectural diversity and open challenges,” *Artificial Intelligence Review*, vol. 58, no. 7, Apr. 2025.
- [102] Y. J. Luwe, C. P. Lee, and K. M. Lim, “Wearable sensor-based human activity recognition with hybrid deep learning model,” *Informatics*, vol. 9, no. 3, p. 56, Jul. 2022.
- [103] M. Waqas and U. W. Humphries, “A critical review of rnn and lstm variants in hydrological time series predictions,” *MethodsX*, vol. 13, p. 102946, Dec. 2024.
- [104] W. Ahmad, B. M. Kazmi, and H. Ali, “Human activity recognition using multi-head cnn followed by lstm,” pp. 1–6, Dec. 2019.
- [105] P. Lalwani and R. Ganeshan, “A novel cnn-bilstm-gru hybrid deep learning model for human activity recognition,” *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, Nov. 2024.
- [106] D. Hu, W. Gao, K. K. Ang, M. Hu, G. Chuai, and R. Huang, “Smart sleep monitoring: Sparse sensor-based spatiotemporal cnn for sleep posture detection,” *Sensors*, vol. 24, no. 15, p. 4833, Jul. 2024.
- [107] R. L. Abduljabbar, H. Dia, and P.-W. Tsai, “Development and evaluation of bidirectional lstm freeway traffic forecasting models using simulation data,” *Scientific Reports*, vol. 11, no. 1, Dec. 2021.
- [108] S. Natha, F. Ahmed, M. Siraj, M. Lagari, M. Altamimi, and A. A. Chandio, “Deep bilstm attention model for spatial and temporal anomaly detection in video surveillance,” *Sensors*, vol. 25, no. 1, p. 251, Jan. 2025.
- [109] Z. Chen, M. Wu, W. Cui, C. Liu, and X. Li, “An attention based cnn-lstm approach for sleep-wake detection with heterogeneous sensors,” *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 9, pp. 3270–3277, Sep. 2021.

- [110] J. Hayano, M. Takeshima, A. Imanishi, M. Ogasawara, Y. Yamada, E. Yuda, and K. Mishima, "Detection of sleep apnea using smartphone-embedded inertial measurement unit," *Scientific Reports*, vol. 15, no. 1, Apr. 2025.
- [111] T. H. Dang, S.-m. Kim, M.-s. Choi, S.-n. Hwan, H.-k. Min, and F. Bien, "An automated algorithm for obstructive sleep apnea detection using a wireless abdomen-worn sensor," *Sensors*, vol. 25, no. 8, p. 2412, Apr. 2025.
- [112] Y.-Y. Lin, H.-T. Wu, C.-A. Hsu, P.-C. Huang, Y.-H. Huang, and Y.-L. Lo, "Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezoelectric bands," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1533–1545, Nov. 2017.
- [113] H. Yoon, S. Hwang, D. Jung, S. Choi, K. Joo, J. Choi, Y. Lee, D.-U. Jeong, and K. Park, "Estimation of sleep posture using a patch-type accelerometer based device," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, Aug. 2015, p. 4942–4945. [Online]. Available: <http://dx.doi.org/10.1109/EMBC.2015.7319500>
- [114] R. M. Kwasnicki, G. W. V. Cross, L. Geoghegan, Z. Zhang, P. Reilly, A. Darzi, G. Z. Yang, and R. Emery, "A lightweight sensing platform for monitoring sleep quality and posture: a simulated validation study," *European Journal of Medical Research*, vol. 23, no. 1, May 2018. [Online]. Available: <http://dx.doi.org/10.1186/s40001-018-0326-9>
- [115] C. He, Z. Fang, S. Liu, H. Wu, X. Li, Y. Wen, and J. Lin, "A smart detection method for sleep posture based on a flexible sleep monitoring belt and vital sign signals," *Heliyon*, vol. 10, no. 11, p. e31839, Jun. 2024.
- [116] J. R. Quinlan, "On explaining decision trees," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 20, no. 2, pp. 339–346, Aug. 2021.
- [117] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 785–794, Aug. 2016.
- [118] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations (ICLR)*, pp. 1–15, Dec. 2015.

- [119] V. Singh, S. K. Sahana, and V. Bhattacharjee, “A novel cnn-gru-lstm based deep learning model for accurate traffic prediction,” *Discover Computing*, vol. 28, no. 1, Apr. 2025.
- [120] M. Rayhan Ahmed, S. Islam, A. Muzahidul Islam, and S. Shatabda, “An ensemble 1d-cnn-lstm-gru model with data augmentation for speech emotion recognition,” *Expert Systems with Applications*, vol. 218, p. 119633, May 2023.
- [121] Z. Yao, Z. Wang, D. Wang, J. Wu, and L. Chen, “An ensemble cnn-lstm and gru adaptive weighting model based improved sparrow search algorithm for predicting runoff using historical meteorological and runoff data as input,” *Journal of Hydrology*, vol. 625, p. 129977, Oct. 2023.
- [122] L. Atallah, B. Lo, R. King, and G.-Z. Yang, “Sensor positioning for activity recognition using wearable accelerometers,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 4, pp. 320–329, 2011.
- [123] V. Farrahi, M. Niemelä, M. Kangas, R. Korpelainen, and T. Jämsä, “Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches,” *Gait & Posture*, vol. 68, pp. 285–299, 2019.

Số 576 - QĐ/HVKHCN

Hà Nội, ngày 25 tháng 6 năm 2026

QUYẾT ĐỊNH
về việc thành lập Hội đồng đánh giá luận án tiến sĩ cấp Học viện

- Căn cứ Quyết định số 175-QĐ/VHLKHCNVN ngày 28/04/2026 của Chủ tịch Viện Hàn lâm Khoa học và Công nghệ Việt Nam ban hành Quy chế Tổ chức và hoạt động của Học viện Khoa học và Công nghệ;

- Căn cứ Thông tư số 18/2021/TT-BGDĐT ngày 28/06/2021 của Bộ trưởng Bộ Giáo dục và Đào tạo về việc ban hành Quy chế tuyển sinh và đào tạo trình độ tiến sĩ;

- Căn cứ Quyết định số 1968/QĐ-HVKHCN ngày 28/12/2021 của Giám đốc Học viện Khoa học và Công nghệ về việc ban hành Quy chế tuyển sinh và đào tạo trình độ tiến sĩ;

- Căn cứ Quyết định số 1999/QĐ-HVKHCN ngày 30/11/2022 của Giám đốc Học viện Khoa học và Công nghệ về việc công nhận nghiên cứu sinh đợt 2 năm 2022;

- Căn cứ Quyết định số 575/QĐ-HVKHCN ngày 24/06/2026 của Giám đốc Học viện Khoa học và Công nghệ về việc gia hạn thời gian học tập lần 2 cho nghiên cứu sinh Vũ Hoàng Diệu;

- Xét đề nghị của Trưởng phòng Đào tạo.

GIÁM ĐỐC HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ
QUYẾT ĐỊNH

Điều 1. Thành lập Hội đồng đánh giá luận án tiến sĩ cấp Học viện cho nghiên cứu sinh Vũ Hoàng Diệu với đề tài:

Nâng cao hiệu quả phân loại tư thế ngũ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

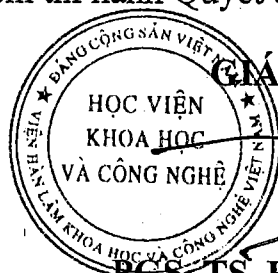
Danh sách thành viên Hội đồng đánh giá luận án kèm theo Quyết định này.

Điều 2. Hội đồng có trách nhiệm đánh giá luận án tiến sĩ theo đúng quy chế hiện hành của Bộ Giáo dục và Đào tạo, của Học viện Khoa học và Công nghệ. Quyết định có hiệu lực tối đa 90 ngày kể từ ngày ký. Hội đồng tự giải thể sau khi hoàn thành nhiệm vụ.

Điều 3. Trưởng phòng Tổ chức - Hành chính, Trưởng phòng Đào tạo, Trưởng phòng Kế toán, các thành viên có tên trong danh sách Hội đồng và nghiên cứu sinh có tên tại Điều 1 chịu trách nhiệm thi hành Quyết định này.

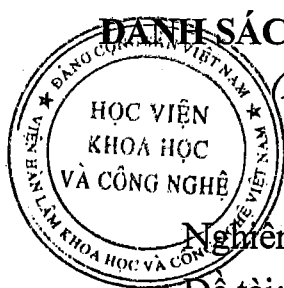
Nơi nhận:

- Như Điều 3;
- Lưu hồ sơ NCS;
- Lưu: VT, ĐT.PQ.15.



GIÁM ĐỐC

PGS. TS. Hoàng Anh Sơn



ĐANH SÁCH HỘI ĐỒNG ĐÁNH GIÁ LUẬN ÁN TIẾN SĨ CẤP HỌC VIỆN

(Kèm theo Quyết định số 576 - QĐ/HVKHCN ngày 25/06/2026
của Giám đốc Học viện Khoa học và Công nghệ)

Nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia
tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn chính: TS. Trần Đức Nghĩa, Viện Công nghệ thông tin,
Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Người hướng dẫn phụ: GS.TS. Trần Đức Tân, Trường Kỹ thuật Phenikaa,
Đại học Phenikaa

TT	Họ và tên, học hàm, học vị	Ngành	Cơ quan công tác	Trách nhiệm trong Hội đồng
1	PGS.TS. Nguyễn Việt Anh	Hệ thống thông tin	Viện Công nghệ thông tin, Viện Hàn lâm KHCN VN	Chủ tịch
2	PGS.TS. Trần Giang Sơn	Khoa học máy tính	Trường Đại học Khoa học và Công nghệ Hà Nội, Viện Hàn lâm KHCN VN	Phản biện 1
3	TS. Ngô Hải Anh	Hệ thống thông tin	Viện Công nghệ thông tin, Viện Hàn lâm KHCN VN	Phản biện 2
4	TS. Nguyễn Như Sơn	Hệ thống thông tin	Viện Công nghệ thông tin, Viện Hàn lâm KHCN VN	Ủy viên - Thư ký
5	PGS.TS. Nguyễn Hà Nam	Khoa học máy tính	Đại học Quốc gia Hà Nội	Ủy viên
6	PGS.TS. Nguyễn Long Giang	Hệ thống thông tin	Viện Công nghệ thông tin, Viện Hàn lâm KHCN VN	Ủy viên
7	PGS.TS. Phạm Văn Hải	Khoa học máy tính	Trường Công nghệ thông tin và Truyền thông, Đại học Bách khoa Hà Nội	Ủy viên

Hội đồng gồm 07 thành viên. *Pha*

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: PGS. TS. Nguyễn Việt Anh, Chủ tịch hội đồng

Cơ quan công tác của người nhận xét: Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Luận án tập trung vào bài toán nhận dạng tư thế ngủ bằng cảm biến gia tốc, trong đó nhấn mạnh yêu cầu phân biệt các tư thế ở mức chi tiết và khả năng làm việc với người dùng mới. Cách tiếp cận này phù hợp với xu hướng phát triển các hệ thống theo dõi sức khỏe dựa trên cảm biến đeo, đồng thời đặt ra yêu cầu rõ ràng về độ chính xác và độ gọn của mô hình. Kết quả nghiên cứu có thể tạo cơ sở cho các nghiên cứu tiếp theo về giám sát giấc ngủ tự động trong môi trường ngoài phòng thí nghiệm.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Luận án xây dựng quy trình nghiên cứu có tính liên kết từ xác định khoảng trống, thu thập và tổ chức dữ liệu, thiết kế mô hình đến kiểm chứng thực nghiệm. Việc đánh giá mô hình trên các đối tượng không xuất hiện trong tập huấn luyện là một

lựa chọn phù hợp, giúp kết quả gần với yêu cầu sử dụng thực tế hơn so với cách chia mẫu ngẫu nhiên thông thường. Ngoài kết quả tổng thể, luận án còn xem xét ảnh hưởng của cấu hình dữ liệu, độ ổn định qua nhiều lần huấn luyện và sai số theo từng lớp, qua đó tăng tính thuyết phục của phần thực nghiệm.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án đã xây dựng được hai hướng mô hình hóa có mục tiêu khác nhau nhưng bổ trợ cho nhau. AnpoNet chú trọng giảm số lượng tham số và chi phí tính toán, trong khi CGB-Net tập trung nâng cao khả năng khai thác thông tin theo thời gian và cải thiện hiệu quả phân loại. Việc đạt được kết quả tốt trong bài toán 12 lớp cho thấy các kiến trúc đề xuất có giá trị tham khảo đối với các bài toán nhận dạng tín hiệu cảm biến đa lớp có yêu cầu tổng quát hóa cao.

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Luận án có bố cục tương đối hợp lý, nội dung các chương được tổ chức theo trình tự từ tổng quan, xây dựng mô hình đến thực nghiệm và đánh giá kết quả. Hệ thống bảng, hình và các kết quả thực nghiệm khá phong phú, hỗ trợ tốt cho việc trình bày nội dung nghiên cứu. Tuy nhiên, tác giả nên tiếp tục rà soát và chỉnh sửa một số vấn đề về hình thức trình bày để luận án hoàn thiện hơn.

Một số ý kiến chỉnh sửa, bổ sung như sau:

- Rà soát lại toàn bộ lỗi chính tả, lỗi đánh máy, dấu câu và khoảng trắng; đặc biệt cần thống nhất cách viết các thuật ngữ tiếng Anh, chữ viết tắt và tên các mô hình trong toàn luận án.
- Cần thống nhất cách Việt hóa và sử dụng thuật ngữ chuyên môn, tránh trường hợp cùng một khái niệm nhưng được diễn đạt theo nhiều cách khác nhau ở các chương.
- Một số bảng có nhiều thông tin và số liệu, nên cân nhắc điều chỉnh cách căn chỉnh, số chữ số thập phân và cách nhấn mạnh kết quả tốt nhất để người đọc dễ theo dõi, so sánh.

- Cần kiểm tra lại tiêu đề, chú thích và cách đánh số của các bảng và hình vẽ, bảo đảm tất cả các bảng, hình đều được dẫn chiếu và giải thích đầy đủ trong phần nội dung.
- Một số hình biểu diễn kết quả thực nghiệm nên tăng kích thước chữ, nhãn trục và chú giải; đồng thời thống nhất cách trình bày giữa các hình có cùng loại kết quả.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét, nội dung và các kết quả chính của luận án không trùng lặp với các luận án, luận văn hay công trình khoa học đã công bố trong và ngoài nước. Các mô hình được đề xuất và cách tổ chức thực nghiệm thể hiện hướng nghiên cứu riêng của tác giả. Không nhận thấy dấu hiệu sao chép hoặc lặp lại một cách không phù hợp các kết quả nghiên cứu đã công bố.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Các kết quả của luận án đã được công bố trong 05 công trình khoa học, trong đó có các bài báo thuộc danh mục SCI/ISI, Scopus Q1, Q2 và hội nghị quốc tế. Các công bố có chất lượng tốt và phù hợp với hướng nghiên cứu của luận án. Nội dung của các bài báo phản ánh các kết quả chủ yếu liên quan đến mô hình AnpoNet, CGB-Net và bài toán phân loại tư thế ngủ, do đó luận án đáp ứng yêu cầu về công bố khoa học.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

Các kết quả nghiên cứu của luận án đã được trích dẫn trung thực trong các công trình công bố. Trích dẫn tài liệu tham khảo được thực hiện một cách rõ ràng, đầy đủ và cập nhật.

8. Kết luận

- Luận án đáp ứng được các yêu cầu về nội dung và hình thức đối với một luận án tiến sĩ, ngành Khoa học máy tính
- Đồng ý cho NCS được bảo vệ tại Hội đồng chấm luận án cấp Học viện.

Hà nội, ngày 26 tháng 07 năm 2026

Người nhận xét



PGS. TS. Nguyễn Việt Anh

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: PGS. TS. Trần Giang Sơn, Ủy viên hội đồng

Cơ quan công tác của người nhận xét: Trường Đại học Khoa học và Công nghệ Hà Nội, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Luận án đã chỉ ra được tầm quan trọng của việc theo dõi tư thế ngủ chi tiết đối với nhiều bệnh lý. Ví dụ, tư thế nằm ngửa làm tăng mức độ nghiêm trọng của hội chứng ngưng thở khi ngủ (OSA), trong khi tư thế nằm nghiêng trái (LLD) giúp giảm thiểu trào ngược dạ dày thực quản (GERD). Luận án cho rằng phân loại 4 tư thế cơ bản (ngửa, sấp, nghiêng trái, nghiêng phải) là không đủ, khi các sai lệch góc nhỏ (ví dụ 15-30 độ) có tác động đáng kể đến sinh lý, thần kinh và hiệu quả trị liệu (như giảm áp lực vùng xương cụt để chống loét).

Luận án cũng giải quyết bài toán chuỗi thời gian bằng các kiến trúc học sâu lai (hybrid deep learning) được tối ưu cho thiết bị biên sử dụng dữ liệu từ cảm biến gia tốc vùng bụng, thay vì sử dụng các mô hình học máy truyền thống phụ thuộc vào việc trích xuất đặc trưng thủ công.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Phương pháp nghiên cứu của luận án được xây dựng tương đối hợp lý: kết hợp giữa việc tự xây dựng bộ dữ liệu thực nghiệm và đề xuất các kiến trúc mạng học sâu. Phương pháp nghiên cứu gồm thiết kế thí nghiệm, thu thập dữ liệu, tiền xử lý, gán nhãn, đề xuất và huấn luyện mô hình học sâu, đánh giá thực nghiệm.

Nhìn chung, phương pháp nghiên cứu của luận án có độ tin cậy, được thiết kế chặt chẽ để giảm các sai số thống kê thường gặp trong xử lý tín hiệu chuỗi thời gian. Dữ liệu thực nghiệm được thiết kế để tránh tình trạng rò rỉ dữ liệu.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án có 3 đóng góp chính:

- Xây dựng dữ liệu
- Xây dựng mô hình AnpoNet tích hợp mạng 1D CNN để trích xuất đặc trưng và LSTM để xử lý bài toán chuỗi thời gian. Kiến trúc nhỏ gọn, đạt độ chính xác cao

Luận án đã đề xuất hai mô hình học sâu là AnpoNet và CGB-Net cho bài toán phân loại 12 tư thế ngủ từ tín hiệu gia tốc. Các kết quả thực nghiệm cho thấy AnpoNet đạt được sự cân bằng tốt giữa độ chính xác và độ phức tạp mô hình, trong khi CGB-Net cải thiện khả năng phân loại và tổng quát hóa. Các đóng góp này có giá trị cả về mặt phương pháp và định hướng triển khai trên các thiết bị có tài nguyên hạn chế.

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Ưu điểm: bố cục hợp lý, đề xuất phương pháp rõ ràng, đã xây dựng mô hình toán học, phân tích độ phức tạp của thuật toán, so sánh kết quả với các phương pháp tiên tiến cho từng mô hình đề xuất. Kết quả đạt được cũng tương đối cạnh tranh.

Điểm cần bổ sung: nên bổ sung so sánh với các mô hình sử dụng thị giác máy tính để nhận diện tư thế ngủ (không xâm lấn, không gây ảnh hưởng đến giấc ngủ...). Một số hình và bảng chưa rõ, khó diễn giải (ví dụ hình 2.6, hình 3.1, hình 3.3, bảng 3.3). Việc dùng nét đứt để thể hiện loss và accuracy trong các biểu đồ Training and Validation làm biểu đồ thiếu tính chính xác. Với định hướng triển khai trên thiết bị cảm biến đeo hoặc thiết bị edge có tài nguyên hạn chế, luận án nên bổ sung phân tích kết quả thực nghiệm trên các thiết bị phân cứng yếu như ESP.

Dữ liệu hiện tại chủ yếu là các tư thế ngủ tĩnh, được thu thập trong điều kiện có kiểm soát (60s), chưa phản ánh đầy đủ các tư thế chuyển tiếp hay các chuyển động phức tạp trong một giấc ngủ qua đêm tự nhiên.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét, luận án là công trình khoa học độc lập, có tính nguyên bản, không có dấu hiệu trùng lặp, sao chép hay đạo văn từ các đồ án, luận văn hay công trình khoa học khác trong và ngoài nước. Các nội dung kế thừa đều được trích dẫn chuẩn mực và các nội dung lỗi đều được bảo chứng bởi các bài báo quốc tế uy tín do chính tác giả công bố.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Kết quả chính của luận án được công bố trong 05 công trình khoa học, trong đó:

- 02 bài báo trên tạp chí Sleep and Breathing (SCIE/Scopus Q1),
- 01 bài trên tạp chí CMC (SCIE/SCOPUS Q2),
- 02 bài trên kỷ yếu hội nghị phản biện.

Các công trình có chất lượng, phù hợp với hướng nghiên cứu của luận án.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

Việc trích dẫn kết quả nghiên cứu và tài liệu tham khảo được thực hiện trung thực, rõ ràng, đầy đủ và cập nhật. Các công bố cũng như luận án đều trích dẫn đầy đủ, phù hợp.

8. Kết luận

- Luận án đáp ứng được các yêu cầu về nội dung và hình thức đối với một luận án tiến sĩ, ngành Khoa học máy tính
- Đồng ý cho NCS được bảo vệ tại Hội đồng chấm luận án cấp Học viện.

Hà Nội, ngày 25 tháng 7 năm 2026

Người nhận xét



PGS. TS. Trần Giang Sơn

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: TS. Ngô Hải Anh, Phản biện

Cơ quan công tác của người nhận xét: Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Luận án nghiên cứu bài toán phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục, tập trung vào phân loại chi tiết 12 tư thế bằng một cảm biến đặt tại vùng bụng. Đây là bài toán có ý nghĩa trong xử lý tín hiệu cảm biến, học máy và các hệ thống theo dõi sức khỏe dựa trên thiết bị đeo. Việc đồng thời xem xét khả năng tổng quát hóa cho người dùng mới và độ phức tạp của mô hình là hướng nghiên cứu có giá trị thực tiễn.

Phạm vi thực nghiệm của luận án chủ yếu là các tư thế tĩnh được thu thập trong điều kiện có kiểm soát; chưa có dữ liệu ngủ tự nhiên qua đêm và chưa xác nhận trên nhóm bệnh nhân. Vì vậy, các nhận định về ứng dụng lâm sàng hoặc triển khai thực tế nên được trình bày ở mức “có tiềm năng” hoặc “hướng tới ứng dụng”, phù hợp với phạm vi mà luận án đã thực hiện.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Luận án sử dụng kết hợp tổng quan tài liệu, xây dựng dữ liệu thực nghiệm, thiết kế mô hình học sâu và đánh giá định lượng. Cách tiếp cận này phù hợp với mục tiêu nghiên cứu. Dataset A gồm 15 người tham gia; Dataset B gồm 18 người và không trùng với Dataset A. Luận án chú ý đến đánh giá độc lập theo đối tượng, đồng thời khảo sát nhiều seed, kiểm

định chéo, ma trận nhầm lẫn, kích thước cửa sổ và tỷ lệ chồng lấp. Đây là những điểm tích cực.

Để tăng khả năng tái lập, NCS nên mô tả thật rõ trình tự chia dữ liệu, phân đoạn cửa sổ và chuẩn hóa; cách xác định các tham số chuẩn hóa; thông số bộ lọc; random seed và nguyên tắc lựa chọn siêu tham số. Riêng với tín hiệu gia tốc dùng để nhận biết tư thế, cần làm rõ việc chuẩn hóa có bảo toàn thông tin định hướng của thành phần trọng lực hay không. Nhìn chung, phương pháp nghiên cứu là hợp lý và các kết quả có cơ sở để sử dụng trong đánh giá các mô hình đề xuất.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án có hai kết quả chính. Thứ nhất, đề xuất mô hình AnpoNet kết hợp 1D CNN và LSTM cho phân loại 12 tư thế ngủ theo hướng gọn nhẹ. Trong thí nghiệm kiểm định chéo ở cấp đối tượng, luận án báo cáo độ chính xác trung bình 94,67% và F1-score 92,94%, với khoảng 2.000 tham số có thể huấn luyện. Kết quả cho thấy mô hình đạt được sự cân bằng tương đối tốt giữa hiệu quả phân loại và độ phức tạp.

Thứ hai, luận án đề xuất CGB-Net kết hợp 1D CNN, GRU và BiLSTM nhằm tăng khả năng biểu diễn đặc trưng và cải thiện hiệu quả phân loại. Trong các thí nghiệm của Chương 3, CGB-Net cho kết quả trung bình cao hơn các mô hình được so sánh trên cùng thiết lập, đổi lại số tham số và chi phí tính toán tăng so với AnpoNet. Hai mô hình vì vậy có thể xem là hai hướng bổ sung: AnpoNet ưu tiên tính gọn nhẹ, còn CGB-Net ưu tiên hiệu quả phân loại. Tính mới của luận án chủ yếu nằm ở thiết kế kiến trúc, tổ chức bài toán 12 tư thế và quy trình đánh giá độc lập theo đối tượng; phù hợp với một luận án theo hướng nghiên cứu ứng dụng.

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Về tổng thể, luận án có mạch nghiên cứu tương đối rõ: Chương 1 xác lập bài toán và cơ sở lựa chọn phương pháp; Chương 2 phát triển mô hình gọn nhẹ AnpoNet; Chương 3 phát triển CGB-Net theo hướng tăng năng lực biểu diễn. Tuy nhiên, do Chương 2 và Chương 3 sử dụng hai tập đối tượng độc lập, luận án cần nhấn mạnh hơn rằng các kết quả tuyệt đối giữa hai chương không nên so sánh trực tiếp; việc so sánh AnpoNet và CGB-Net chỉ có ý nghĩa khi hai mô hình được đánh giá trên cùng thiết lập dữ liệu.

Đối với Chương 1: Phần tổng quan khá đầy đủ, nhưng khoảng trống nghiên cứu nên được kết luận cô đọng hơn và liên hệ trực tiếp với hai đóng góp của luận án. Lý do lựa chọn đúng 12 tư thế, tương ứng các mức góc 30° , cần được giải thích chặt hơn. Phần chỉ số đánh giá cũng cần nêu rõ F1-score sử dụng trong bài toán đa lớp là macro, micro hay weighted để tránh cách hiểu khác nhau.

Đối với Chương 2: Đây là chương được tổ chức tương đối đầy đủ từ thiết bị, dữ liệu đến mô hình và thực nghiệm. Tuy nhiên, Dataset A chỉ gồm 15 người, độ tuổi 19-24, nên

các nhận xét về tính đa dạng của nhóm khảo sát cần thận trọng hơn. NCS nên mô tả rõ hơn sai số cho phép khi thiết lập/gán nhãn góc tư thế, quy trình tiền xử lý và chuẩn hóa. Các số liệu về bộ nhớ, thời gian suy luận và năng lượng cần nói rõ là đo trực tiếp hay ước lượng, trên nền tảng phần cứng nào. Vì chưa triển khai trực tiếp trên vi điều khiển hoặc thiết bị đeo công suất thấp, nên thống nhất dùng cách diễn đạt “hướng tới triển khai trên thiết bị biên”.

Đối với Chương 3: NCS cần giải thích rõ hơn vai trò của việc sử dụng đồng thời GRU và BiLSTM và lợi ích thu được so với mức tăng về độ phức tạp. Kết quả nên được thảo luận như một sự đánh đổi giữa độ chính xác/F1 và chi phí tính toán. Một số lớp tư thế vẫn có tỷ lệ nhầm lẫn cao hơn các lớp khác; đây là giới hạn nên được nêu thẳng và phân tích gắn với sự gần nhau về góc, khác biệt giữa đối tượng và sai lệch vị trí cảm biến.

Đối với phần Kết luận: Nên tách rõ những nội dung đã được chứng minh bằng thực nghiệm với những khả năng dự kiến trong tương lai. Các kết luận về ứng dụng lâm sàng, theo dõi ngủ tự nhiên và triển khai trên phần cứng nhúng cần giữ đúng phạm vi của nghiên cứu hiện tại.

Về hình thức trình bày: Bản thảo còn một số lỗi biên tập và thiếu nhất quán cần rà soát toàn văn. Ví dụ: mục lục ghi “LỜI CẢM ƠN iii” trong khi trang Lời cảm ơn là ii; “Confusion Matrix ((ma trận nhầm lẫn)” thừa dấu ngoặc; có lỗi dạng “§ố tham số”; cách viết Seed/seed, K-fold/K-Folds, F1-score/F1-Score, FLOPs/FLOPS chưa thống nhất. Cũng cần thống nhất cách dùng Accuracy và Precision trong tiếng Việt, cách ghi đơn vị, dấu %, ký hiệu góc và tên hình/bảng. Đây là các lỗi không ảnh hưởng đến kết quả nghiên cứu nhưng cần được sửa trong bản hoàn thiện.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét và trên cơ sở các tài liệu được cung cấp, chưa thấy các kết quả chính của luận án trùng lặp với các luận án, luận văn hoặc công trình khoa học đã công bố theo cách ảnh hưởng đến tính mới của luận án.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

NCS có 05 công trình được liệt kê liên quan đến hướng nghiên cứu của luận án. Trong đó, CT1 về AnpoNet được công bố trên Sleep and Breathing và CT2 về CGB-Net được công bố trên Computers, Materials & Continua; đây là hai công trình phản ánh trực tiếp hai đóng góp chính của luận án. Các công trình còn lại thể hiện quá trình phát triển hướng nghiên cứu về phân loại tư thế ngủ sử dụng cảm biến gia tốc và các phương pháp học máy/học sâu. Theo tôi, các công bố chứa đựng những nội dung chủ yếu của luận án và NCS đáp ứng yêu cầu về công bố khoa học.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

Qua bản luận án được cung cấp, các công trình của NCS và tài liệu tham khảo được trích dẫn tương đối đầy đủ. Tôi không thấy vấn đề đáng lưu ý về tính trung thực trong việc sử dụng và trích dẫn tài liệu.

8. Kết luận

Luận án đã thực hiện một khối lượng nghiên cứu tương đối đầy đủ về phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc. NCS đã xây dựng dữ liệu thực nghiệm, đề xuất hai mô hình AnpoNet và CGB-Net, tiến hành các nhóm thí nghiệm đánh giá và có các công bố khoa học liên quan trực tiếp đến nội dung luận án.

Luận án còn một số điểm cần chỉnh sửa về cách mô tả phương pháp, mối liên hệ giữa hai mô hình, phạm vi diễn giải của kết quả và hình thức trình bày. Tuy nhiên, các tồn tại này không làm thay đổi các kết quả và đóng góp chính của luận án.

Tôi đánh giá luận án đáp ứng các yêu cầu về nội dung và hình thức đối với một luận án tiến sĩ ngành Hệ thống thông tin. Tôi đồng ý thông qua luận án và đề nghị Hội đồng xem xét công nhận kết quả bảo vệ, trên cơ sở NCS tiếp thu các ý kiến của Hội đồng và hoàn thiện bản luận án trước khi nộp bản cuối.

Hà Nội, ngày 24 tháng 7 năm 2026

Người nhận xét



TS. Ngô Hải Anh

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: TS. Nguyễn Như Sơn, Thư ký hội đồng

Cơ quan công tác của người nhận xét: Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Theo dõi tư thế ngủ có ý nghĩa quan trọng trong hỗ trợ chẩn đoán các rối loạn giấc ngủ, đặc biệt là hội chứng ngưng thở khi ngủ và các bệnh lý hô hấp liên quan. Luận án lựa chọn một bài toán có tính ứng dụng thực tế cao trong y tế số, chăm sóc sức khỏe từ xa và các hệ thống theo dõi sức khỏe thông minh. Đề tài có ý nghĩa khoa học trong việc nghiên cứu bài toán phân loại chi tiết 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục bằng các mô hình học sâu. Nghiên cứu góp phần bổ sung hướng tiếp cận phân loại tư thế ngủ có độ phân giải cao, thay vì chỉ tập trung vào một số tư thế cơ bản. Về thực tiễn, kết quả có tiềm năng ứng dụng trong các hệ thống theo dõi giấc ngủ và chăm sóc sức khỏe sử dụng thiết bị đeo.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Phương pháp nghiên cứu của luận án được xây dựng tương đối hợp lý, kết hợp giữa tổng quan tài liệu, xây dựng dữ liệu thực nghiệm, mô hình hóa bằng học sâu và

đánh giá định lượng. Các chỉ số đánh giá như accuracy, precision, recall, F1-score và ma trận nhầm lẫn được sử dụng phù hợp với bài toán phân loại đa lớp. Đặc biệt, cách đánh giá độc lập theo đối tượng giúp nâng cao độ tin cậy và phản ánh tốt hơn khả năng tổng quát hóa của mô hình.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án đã đề xuất hai mô hình học sâu là AnpoNet và CGB-Net cho bài toán phân loại 12 tư thế ngủ từ tín hiệu gia tốc.

- Đề xuất mô hình AnpoNet nhằm phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục với số lượng tham số được kiểm soát. Mô hình AnpoNet được xây dựng trên cơ sở kết hợp khả năng trích xuất đặc trưng cục bộ của mạng nơ-ron tích chập một chiều với khả năng mô hình hóa phụ thuộc thời gian của mạng hồi quy.
- Đề xuất mô hình CGB-Net nhằm nâng cao độ chính xác và khả năng tổng quát hóa trong bài toán phân loại 12 tư thế ngủ. Mô hình CGB-Net được thiết kế bằng cách kết hợp ba thành phần gồm mạng nơ-ron tích chập một chiều, GRU và BiLSTM.

Các kết quả thực nghiệm cho thấy AnpoNet đạt được sự cân bằng tốt giữa độ chính xác và độ phức tạp mô hình, trong khi CGB-Net cải thiện khả năng phân loại và tổng quát hóa. Các đóng góp này có giá trị cả về mặt phương pháp và định hướng triển khai trên các thiết bị có tài nguyên hạn chế.

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Luận án có bố cục tương đối hợp lý, nội dung nghiên cứu được trình bày có hệ thống và các kết quả thực nghiệm được phân tích khá đầy đủ.

Một số nhận xét có thể bổ sung chỉnh sửa như sau:

- Dữ liệu hiện chủ yếu được thu thập trong điều kiện có kiểm soát, chưa phản ánh đầy đủ các tư thế chuyển tiếp và giấc ngủ tự nhiên qua đêm. Hầu hết các kết quả đánh giá được thực hiện trên bộ dữ liệu do nhóm nghiên cứu xây dựng. Nếu mở rộng ra các tập dữ liệu khác thì kết quả sẽ như thế nào?
- Làm rõ thêm về tính mới về mặt phương pháp khi kết hợp các thành phần trong các mô hình đề xuất?.
- Khả năng triển khai thực tế trên thiết bị biên, thiết bị đeo công suất thấp?.
- Nên bỏ Kiến nghị ở phần Kết luận cuối, hướng phát triển hiện khá ngắn gọn, có thể bổ sung đầy đủ hơn?.
- Các công trình công bố nên viết SCIE/Scopus nếu trong các danh mục này.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét, nội dung và các kết quả chính của luận án không trùng lặp với các luận án, luận văn hay công trình khoa học đã công bố trong và ngoài nước. Các mô hình được đề xuất và cách tổ chức thực nghiệm thể hiện hướng nghiên cứu riêng của tác giả.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Các kết quả của luận án đã được công bố trong 05 công trình khoa học, trong đó có 02 bài báo thuộc danh mục SCIE-Scopus Q1, 01 bài SCIE Q2, 01 tạp chí và 01 hội nghị quốc tế. Các công bố có chất lượng khá tốt và phù hợp với hướng nghiên cứu của luận án. Nội dung của các bài báo phản ánh các kết quả chủ yếu liên quan đến mô hình AnpoNet, CGB-Net và bài toán phân loại tư thế ngủ, NCS là tác giả đầu của tất cả 05 công bố, do đó luận án đáp ứng yêu cầu về công bố khoa học.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

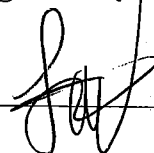
Các kết quả nghiên cứu của luận án đã được trích dẫn trung thực trong các công trình công bố. Trích dẫn tài liệu tham khảo được thực hiện một cách rõ ràng, đầy đủ và cập nhật.

8. Kết luận

Luận án của NCS Vũ Hoàng Diệu đáp ứng các yêu cầu đối với một luận án tiến sĩ chuyên ngành Khoa học Máy tính. Bản tóm tắt luận án phản ánh trung thành nội dung cơ bản của luận án. Luận án có thể đưa ra bảo vệ cấp Học viện để nhận học vị Tiến sĩ.

Hà nội, ngày 23 tháng 07 năm 2026

Người nhận xét



TS. Nguyễn Như Sơn

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: PGS. TS. Nguyễn Hà Nam, Ủy viên

Cơ quan công tác của người nhận xét: Ban Khoa học và Đổi mới sáng tạo, Đại học Quốc gia Hà Nội

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Phân loại tư thế ngủ là bài toán có ý nghĩa lâm sàng rõ ràng, liên quan trực tiếp đến hội chứng ngưng thở khi ngủ do tắc nghẽn (OSA), trào ngược dạ dày thực quản (GERD), phòng ngừa loét tì đè và phục hồi chức năng thần kinh. Phần lớn các hệ thống công bố trước đây chỉ dừng ở việc phân biệt 3-4 tư thế cơ bản (nằm ngửa, sấp, nghiêng trái, nghiêng phải), trong khi nhiều can thiệp y học đòi hỏi độ chi tiết cao hơn (ví dụ góc nghiêng bán phần 30°). Luận án lựa chọn hướng tiếp cận dùng một cảm biến gia tốc ba trục duy nhất, gắn tại vùng bụng, để phân loại 12 tư thế ngủ - vừa mở rộng độ chi tiết phân loại so với mặt bằng chung của lĩnh vực, vừa giữ được ưu điểm chi phí thấp, không xâm lấn và khả năng triển khai trên thiết bị đeo tài nguyên hạn chế so với các giải pháp dùng camera, cảm biến áp lực hoặc mảng đa cảm biến. Đây là khoảng trống nghiên cứu cụ thể, có cơ sở lý luận và thực tiễn xác đáng.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Về tính hiện đại, hai kiến trúc lai CNN-LSTM và CNN-GRU-BiLSTM là các lựa chọn kỹ thuật hợp lý, có cơ sở lý thuyết rõ ràng cho bài toán phân loại chuỗi thời gian đa lớp từ dữ liệu cảm biến. Thiết kế thực nghiệm của luận án khá kỹ lưỡng so với mặt bằng chung của lĩnh vực: đánh giá độc lập theo đối tượng (subject-independent), kiểm định chéo k-fold, chạy lặp lại với nhiều seed ngẫu nhiên và báo cáo mean $\pm$ độ lệch chuẩn, phân tích ma trận nhầm lẫn theo từng lớp tư thế, khảo sát ảnh hưởng của độ dài cửa sổ và tỷ lệ chồng lấp, đồng thời đo đạc các chỉ số triển khai thiết bị biên (số tham số, FLOPs, dung lượng bộ nhớ, thời gian suy luận, năng lượng tiêu thụ ước tính). Đây là mức độ chi tiết thực nghiệm hiếm gặp và đáng ghi nhận trong một luận án ứng dụng học sâu.

Đặc biệt, tại Chương 3 (trang 99), luận án chủ động cảnh báo rằng kết quả AnpoNet ở Chương 2 và kết quả so sánh AnpoNet - CGB-Net ở Chương 3 được thực hiện trên các thiết lập dữ liệu khác nhau nên không thể so sánh trực tiếp hiệu suất tuyệt đối giữa hai chương - đây là biểu hiện của sự cẩn trọng khoa học, không phải NCS nào cũng chủ động nêu ra.

Mặc dù vậy, người nhận xét có một số vấn đề cụ thể về độ tin cậy cần được giải đáp:

- (i) Thiếu kiểm định ý nghĩa thống kê. Bảng 3.1 cho thấy CGB-Net vượt AnpoNet với chênh lệch chỉ 0,52-4,14 điểm phần trăm ở 3/4 thiết lập thí nghiệm (riêng Thiết lập 3 chỉ chênh 0,52 điểm %: 92,70% so với 92,18%). Trong khi đó, chính kết quả kiểm định chéo 5-fold của AnpoNet ở Chương 2 (mục 2.4.3) cho thấy độ lệch chuẩn giữa các fold lên tới 0,042-0,118 (tức 4,2-11,8 điểm phần trăm) - lớn hơn nhiều so với mức chênh lệch CGB-Net/AnpoNet nêu trên. Luận án không cung cấp kiểm định thống kê (t-test, Wilcoxon signed-rank...) để khẳng định các cải thiện này nằm ngoài phạm vi biến động ngẫu nhiên. Đáng lưu ý, bài báo CT1 (Sleep and Breathing) có sử dụng phép kiểm Wilcoxon signed-rank, nhưng cho mục đích khác (kiểm tra tính nhất quán giữa các đối tượng thử nghiệm), không phải để so sánh mô hình đề xuất với các mô hình cơ sở, và kiểm định này cũng không được đưa vào luận án.

- (ii) Cỡ mẫu thực nghiệm nhỏ và đồng nhất về nhân khẩu học. Hai chương chính chỉ sử dụng 15 (Chương 2) và 18 (Chương 3) người tham gia, độ tuổi 19-24, thể trạng khỏe mạnh. Trong khi động cơ khoa học nêu ở Chương 1 gắn liền với các nhóm bệnh lý (OSA, GERD, người cao tuổi, bệnh nhân nằm liệt giường), thực nghiệm chưa có đối tượng thuộc các nhóm này. NCS đã tự nhận thức hạn chế này ở phần Kết luận, tuy vậy cần làm rõ thêm mức độ ảnh hưởng đến khả năng khái quát hóa cho ứng dụng lâm sàng thực tế.
- (iii) Đóng góp biên của thành phần hồi quy trong AnpoNet chưa được làm rõ triệt để. Tại Bảng 3.1, mô hình 1D CNN đơn thuần (không có thành phần hồi quy) đạt độ chính xác khá cạnh tranh và thậm chí vượt AnpoNet ở tình huống 1 (86,38% so với 85,20%) và tình huống 4 (83,20% so với 81,23%). Điều này đặt câu hỏi về hiệu quả biên thực sự của khối LSTM trong AnpoNet so với chi phí tham số bổ sung, cần một nghiên cứu loại bỏ thành phần (ablation study) tường minh hơn.
- (iv) Khả năng triển khai thiết bị biên mới dừng ở ước lượng lý thuyết (số tham số, FLOPs, bộ nhớ, năng lượng tính toán), chưa có thực nghiệm đo đạc trên phần cứng nhúng thực tế, trong khi định hướng "thiết bị biên" được nhấn mạnh xuyên suốt Chương 2.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án có hai đóng góp chính, số liệu đã được người nhận xét đối chiếu và xác nhận thống nhất giữa nội dung luận án và các bài báo công bố tương ứng:

- Đóng góp 1 (Chương 2): Mô hình AnpoNet (1D CNN + LSTM) cho bài toán phân loại 12 tư thế ngủ, đạt độ chính xác $94,67\% \pm 0,80\%$ và F1-score $92,94\% \pm 1,35\%$ trên tập kiểm tra, chỉ dùng 1.916 tham số (~34 kB bộ nhớ), 0,1603 triệu FLOPs, thời gian suy luận 0,081 giây và năng lượng ước tính 1,122 mJ/lần suy luận. Đây là kết quả có ý nghĩa, thể hiện sự cân bằng tốt giữa độ chính xác và độ gọn nhẹ mô hình, phù hợp định hướng thiết bị đeo tài nguyên

hạn chế. Kết quả được công bố tại [CT1], tạp chí Sleep and Breathing (SCI/ISI, Scopus Q1, IF 2,4).

- Đóng góp 2 (Chương 3): Mô hình CGB-Net (1D CNN + GRU + BiLSTM) nhằm tăng cường độ chính xác và khả năng tổng quát hóa trong điều kiện đánh giá độc lập theo đối tượng nghiêm ngặt hơn, đạt độ chính xác trung bình 87,60% và F1- score trung bình 83,38% qua bốn thiết lập thí nghiệm, vượt AnpoNet (được đánh giá lại trong cùng thiết lập) khoảng 4,14 điểm % về độ chính xác và hơn 10 điểm % về F1-score ở thiết lập thứ nhất. Kết quả được công bố tại [CT2], tạp chí CMC: Computers, Materials & Continua (SCI/ISI, Scopus Q2, IF 1,4).

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Ưu điểm:

- Tên đề tài nêu rõ cách tiếp cận "kết hợp các mô hình học sâu", phù hợp với nội dung cốt lõi của luận án: mô hình lai AnpoNet (1D CNN + LSTM, Chương 2) và mô hình lai CGB-Net (1D CNN + GRU + BiLSTM, Chương 3). Nội dung luận án - thiết kế kiến trúc mạng nơ-ron, xử lý tín hiệu chuỗi thời gian, đánh giá thực nghiệm định lượng - phù hợp với ngành Khoa học máy tính, mã số 9 48 01 01. Đây là điểm cộng rõ ràng so với một số luận án có tên đề tài dùng thuật ngữ không khớp nội dung kỹ thuật thực tế.
- Cấu trúc luận án mạch lạc, mạch đóng góp xuyên suốt từ tổng quan (Chương 1) đến hai mô hình đề xuất (Chương 2, 3) và kết luận. Quy trình xây dựng bộ dữ liệu, thu thập, gán nhãn và tiền xử lý được mô tả chi tiết, có khả năng tái lập.
- Thực nghiệm được thiết kế với mức độ chính xác cao hơn mặt bằng chung: đánh giá độc lập theo đối tượng, kiểm định chéo, phân tích độ ổn định theo seed, phân tích lỗi theo từng lớp và đo đặc chi phí triển khai.
- Luận án cũng thể hiện sự trung thực khoa học khi chủ động phân biệt kết quả từ các nghiên cứu tiền khả thi (CT4, CT5 - dữ liệu trộn đối tượng) với kết quả

chính thức (Chương 2, 3 - đánh giá độc lập theo đối tượng), thay vì gộp chung để tạo ấn tượng tốt hơn.

Thiếu sót và những điểm cần bổ sung, sửa chữa:

- Thứ nhất, luận án chưa sử dụng kiểm định thống kê để củng cố các so sánh hiệu năng giữa mô hình đề xuất và mô hình cơ sở, trong khi độ biến động giữa các lần chạy không nhỏ (mục 2.i).
- Thứ hai, cỡ mẫu thực nghiệm còn khiêm tốn và đồng nhất về nhân khẩu học so với phạm vi ứng dụng lâm sàng được đặt ra ở phần mở đầu.
- Thứ ba, mối quan hệ giữa 5 công trình công bố và 3 chương chính thức chưa được trình bày tường minh trong luận án (mục 3), gây khó khăn cho người đọc khi đối chiếu.
- Thứ tư, dữ liệu thực nghiệm là các tư thế tĩnh thu thập có kiểm soát, chưa phản ánh dữ liệu ngẫu nhiên qua đêm - hạn chế này đã được chính NCS nêu ở phần Kết luận, thể hiện sự tự nhận thức tốt, song cũng cho thấy khoảng cách còn khá xa giữa kết quả labo và mục tiêu ứng dụng lâm sàng ban đầu.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Hai đóng góp chính của luận án (AnpoNet, CGB-Net) đã được công bố trên hai tạp chí SCI/ISI độc lập; người nhận xét chưa phát hiện dấu hiệu trùng lặp nội dung với các công trình khác, tuy nhiên đề nghị cơ sở đào tạo đối chiếu bổ sung bằng phần mềm kiểm tra trùng lặp chuyên dụng theo quy định hiện hành.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Danh mục công trình công bố của NCS gồm 5 bài báo, trong đó chỉ 2 bài (CT1 tương ứng Chương 2, CT2 tương ứng Chương 3) cấu thành trực tiếp nội dung hai chương chính. Hai công trình CT3 (HypoNet) và CT5 (GA-SPARF, dùng Random Forest tối ưu bằng giải thuật di truyền - một phương pháp học máy truyền thống, không phải học sâu) không được trình bày như một chương hay mục riêng trong luận án, dù CT5 được nhắc đến ngắn gọn ở Chương 1 như một "nghiên cứu phụ trợ". Đề nghị NCS làm rõ vai trò chính thức của CT3 và CT5 trong cấu trúc tổng thể của luận

án, để tránh gây hiểu nhầm khi đối chiếu danh mục công bố với nội dung ba chương chính.

Luận án có 05 công trình liên quan, trong đó 02 công trình tương ứng trực tiếp với hai chương chính:

- [CT1] Hoang-Dieu Vu và cộng sự (2025), "Human sleep position classification using a lightweight model and acceleration data", Sleep and Breathing, vol. 29, Article 95 (SCI/ISI, Scopus Q1, IF 2,4) — tương ứng Chương 2.
- [CT2] Hoang-Dieu Vu và cộng sự (2025), "CGB-Net: A Novel Convolutional Gated Bidirectional Network for Enhanced Sleep Posture Classification", CMC: Computers, Materials & Continua, vol. 85, no. 2 (SCI/ISI, Scopus Q2, IF 1,4) - tương ứng Chương 3.
- [CT3] HyPoNet, REV Journal on Electronics and Communications (2025) - công trình hỗ trợ, chưa xác định là tạp chí thuộc danh mục SCI/ISI/Scopus.
- [CT4] Hội nghị IEEE Statistical Signal Processing Workshop (SSP) 2023 (Scopus) - nghiên cứu sơ bộ, được dùng làm cơ sở thúc đẩy Chương 2/3.
- [CT5] GA-SPARF, Sleep and Breathing (2026), vol. 30 (SCI/ISI, Scopus Q1) - nghiên cứu hỗ trợ dùng học máy truyền thống (Random Forest tối ưu bằng giải thuật di truyền), không thuộc nội dung ba chương chính.

Hai công trình chính (CT1, CT2) đều thuộc danh mục SCI/ISI (Scopus Q1 và Q2), vượt yêu cầu tối thiểu về công bố theo quy định hiện hành đối với NCS ngành Khoa học máy tính, và là các tạp chí có mức độ liên quan trực tiếp đến chủ đề luận án (Sleep and Breathing là tạp chí chuyên ngành y học giấc ngủ, CMC là tạp chí điện toán uy tín). Đây là điểm mạnh rõ rệt của luận án về mặt công bố khoa học.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

Các kết quả nghiên cứu của luận án đã được trích dẫn trung thực trong các công trình công bố.

Luận án có danh mục 110 tài liệu tham khảo (không tính các mục [CT]), trong đó khoảng 61/110 (~55%) được công bố trong giai đoạn 2021-2026, phần còn lại chủ yếu là các tài liệu lâm sàng nền tảng về sinh lý bệnh OSA/GERD/tư thế ngủ (một số từ 1998-2014) và các công trình phương pháp luận kinh điển của lĩnh vực học sâu (LeCun và cộng sự 2015; Kingma & Ba 2015 - thuật toán Adam; Chen & Guestrin 2016 - XGBoost). Việc trích dẫn các tài liệu nền tảng này là hợp lý về học thuật dù

không thuộc 5 năm gần nhất, vì chúng là nguồn gốc trực tiếp của các phương pháp được sử dụng. Nhìn chung tỷ lệ tài liệu cập nhật (>50% trong 5 năm) ở mức chấp nhận được đối với một luận án CNTT ứng dụng.

Qua rà soát đối chiếu ngẫu nhiên một số trích dẫn với nội dung sử dụng trong luận án, người nhận xét chưa phát hiện trường hợp trích dẫn sai lệch nội dung hoặc trích dẫn không phù hợp ngữ cảnh. Trích dẫn tài liệu tham khảo được thực hiện một cách rõ ràng, đầy đủ và cập nhật.

8. Kết luận

Xét tổng thể, luận án đáp ứng yêu cầu về nội dung khoa học, tính mới và số lượng/chất lượng công bố của một luận án tiến sĩ ngành Khoa học máy tính theo quy định của Bộ Giáo dục và Đào tạo và của Viện Hàn lâm Khoa học và Công nghệ Việt Nam. Phương pháp nghiên cứu được trình bày rõ ràng, thực nghiệm được thiết kế với mức độ nghiêm cần đáng ghi nhận so với mặt bằng chung của các luận án ứng dụng học sâu. Luận án đủ điều kiện đưa ra bảo vệ trước Hội đồng đánh giá luận án cấp Học viện. NCS cần giải đáp thỏa đáng các câu hỏi nêu dưới đây và chỉnh sửa luận án theo các ý kiến ở mục 6.

Hà nội, ngày 22 tháng 07 năm 2026

Người nhận xét



PGS. TS. Nguyễn Hà Nam

Câu hỏi

Câu hỏi 1: Trong Bảng 3.1/3.2, mô hình 1D CNN đơn thuần (không có thành phần hồi quy) đạt độ chính xác cạnh tranh hoặc vượt AnpoNet ở Thiết lập 1 và Thiết lập 4. NCS đánh giá thế nào về đóng góp thực sự của thành phần LSTM trong AnpoNet, xét trên góc độ đánh đổi giữa số lượng tham số bổ sung và mức cải thiện hiệu năng thu được?

Câu hỏi 2: Công trình [CT5] (GA-SPARF) của chính NCS báo cáo mô hình Random Forest tối ưu bằng giải thuật di truyền đạt 99,79% dưới kiểm định chéo 5-fold tiêu chuẩn, nhưng tự thừa nhận có "sự sụt giảm hiệu suất đáng kể" khi đánh giá theo Leave-One-Subject-Out. NCS có số liệu định lượng cụ thể của mô hình học máy truyền thống này khi được đánh giá dưới đúng thiết lập độc lập theo đối tượng như Chương 2/3 của luận án hay không, để làm rõ luận điểm về sự cần thiết của kiến trúc học sâu lai so với học máy truyền thống đã được tối ưu kỹ lưỡng?

Câu hỏi 3: Khả năng triển khai trên thiết bị biên của AnpoNet mới được đánh giá qua các chỉ số gián tiếp (số tham số, FLOPs, bộ nhớ, năng lượng ước tính). NCS có kế hoạch hoặc kết quả thực nghiệm nào trên phần cứng nhúng thực tế (ví dụ vi điều khiển dạng ESP32/ESP8266 như đã sử dụng trong CT5) để kiểm chứng các con số lý thuyết này hay chưa?

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: PGS.TS. Nguyễn Long Giang, Ủy viên hội đồng

Cơ quan công tác của người nhận xét: Viện Công nghệ thông tin, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Đề tài có ý nghĩa khoa học trong việc nghiên cứu bài toán phân loại chi tiết 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục bằng các mô hình học sâu. Nghiên cứu góp phần bổ sung hướng tiếp cận phân loại tư thế ngủ có độ phân giải cao, thay vì chỉ tập trung vào một số tư thế cơ bản. Về thực tiễn, kết quả có tiềm năng ứng dụng trong các hệ thống theo dõi giấc ngủ và chăm sóc sức khỏe sử dụng thiết bị đeo.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Phương pháp nghiên cứu của luận án được xây dựng tương đối hợp lý, kết hợp giữa tổng quan tài liệu, xây dựng dữ liệu thực nghiệm, mô hình hóa bằng học sâu và đánh giá định lượng. Các chỉ số đánh giá như accuracy, precision, recall, F1-score và ma trận nhầm lẫn được sử dụng phù hợp với bài toán phân loại đa lớp. Đặc biệt, cách đánh giá độc lập theo đối tượng giúp nâng cao độ tin cậy và phản ánh tốt hơn khả năng tổng quát hóa của mô hình.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án đã đề xuất hai mô hình học sâu là AnpoNet và CGB-Net cho bài toán phân loại 12 tư thế ngủ từ tín hiệu gia tốc. Các kết quả thực nghiệm cho thấy AnpoNet đạt được sự cân bằng tốt giữa độ chính xác và độ phức tạp mô hình, trong khi CGB-Net cải thiện khả năng phân loại và tổng quát hóa. Các đóng góp này có giá trị cả về mặt phương pháp và định hướng triển khai trên các thiết bị có tài nguyên hạn chế.

4. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

Luận án có bố cục tương đối hợp lý, nội dung nghiên cứu được trình bày có hệ thống và các kết quả thực nghiệm được phân tích khá đầy đủ. Tuy nhiên, dữ liệu hiện chủ yếu được thu thập trong điều kiện có kiểm soát, chưa phản ánh đầy đủ các tư thế chuyển tiếp và giấc ngủ tự nhiên qua đêm. Tác giả nên làm rõ hơn các hạn chế này, đồng thời bổ sung thảo luận về khả năng triển khai thực tế trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.

5. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét, nội dung và các kết quả chính của luận án không trùng lặp với các luận án, luận văn hay công trình khoa học đã công bố trong và ngoài nước. Các mô hình được đề xuất và cách tổ chức thực nghiệm thể hiện hướng nghiên cứu riêng của tác giả. Không nhận thấy dấu hiệu sao chép hoặc lặp lại một cách không phù hợp các kết quả nghiên cứu đã công bố.

6. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Các kết quả của luận án đã được công bố trong 05 công trình khoa học, trong đó có các bài báo thuộc danh mục SCI/ISI, Scopus Q1, Q2 và hội nghị quốc tế. Các công bố có chất lượng tốt và phù hợp với hướng nghiên cứu của luận án. Nội dung của các bài báo phản ánh các kết quả chủ yếu liên quan đến mô hình AnpoNet, CGB-Net và bài toán phân loại tư thế ngủ, do đó luận án đáp ứng yêu cầu về công bố khoa học.

7. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

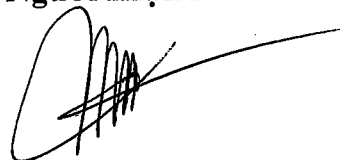
Các kết quả nghiên cứu của luận án đã được trích dẫn trung thực trong các công trình công bố. Trích dẫn tài liệu tham khảo được thực hiện một cách rõ ràng, đầy đủ và cập nhật.

8. Kết luận

- Luận án đáp ứng được các yêu cầu về nội dung và hình thức đối với một luận án tiến sĩ, ngành Khoa học máy tính
- Đồng ý cho NCS được bảo vệ tại Hội đồng chấm luận án cấp Học viện.

Hà nội, ngày 30 tháng 07 năm 2026

Người nhận xét



PGS. TS. Nguyễn Long Giang

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BẢN NHẬN XÉT LUẬN ÁN TIẾN SĨ
(Cấp Học viện)

Tên nghiên cứu sinh: Vũ Hoàng Diệu

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa, GS. TS. Trần Đức Tân

Người nhận xét: PGS.TS. Phạm Văn Hải, Ủy viên hội đồng

Cơ quan công tác của người nhận xét: Trường Công nghệ Thông tin và Truyền thông, Đại học Bách khoa Hà Nội.

NỘI DUNG NHẬN XÉT

1. Ý nghĩa khoa học và thực tiễn của đề tài luận án

Đề tài có ý nghĩa khoa học trong việc nghiên cứu bài toán phân loại chi tiết 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục bằng các mô hình học sâu. Nghiên cứu góp phần bổ sung hướng tiếp cận phân loại tư thế ngủ có độ phân giải cao, thay vì chỉ tập trung vào một số tư thế cơ bản. Về thực tiễn, kết quả có tiềm năng ứng dụng trong các hệ thống theo dõi giấc ngủ và chăm sóc sức khỏe sử dụng thiết bị đeo.

Luận án tập trung vào việc nâng cao hiệu quả nhận dạng các tư thế ngủ chi tiết, tăng khả năng tổng quát hóa trong điều kiện độc lập theo đối tượng, đồng thời kiểm soát độ phức tạp của mô hình theo định hướng triển khai trên các hệ thống cảm biến đeo có tài nguyên tính toán hạn chế.

2. Sự hợp lý và độ tin cậy của các phương pháp nghiên cứu

Phương pháp nghiên cứu của luận án được xây dựng tương đối hợp lý, kết hợp giữa tổng quan tài liệu, xây dựng dữ liệu thực nghiệm, mô hình hóa bằng học sâu và đánh giá định lượng.

Các chỉ số đánh giá như accuracy, precision, recall, F1-score và ma trận nhầm lẫn được sử dụng phù hợp với bài toán phân loại đa lớp. Đặc biệt, cách đánh giá độc lập theo đối tượng giúp nâng cao độ tin cậy và phản ánh tốt hơn khả năng tổng quát hóa của mô hình.

3. Đánh giá các kết quả đạt được, nêu những đóng góp mới và giá trị của các đóng góp đó

Luận án đã đề xuất hai mô hình học sâu là AnpoNet và CGB-Net cho bài toán phân loại 12 tư thế ngủ từ tín hiệu gia tốc. Các kết quả thực nghiệm cho thấy AnpoNet đạt được sự cân bằng tốt giữa độ chính xác và độ phức tạp mô hình, trong khi CGB-Net cải thiện khả năng phân loại và tổng quát hóa.

Đề xuất mô hình AnpoNet, một kiến trúc học sâu lai kết hợp 1D CNN và LSTM cho bài toán phân loại 12 tư thế ngủ. Mô hình được thiết kế nhằm khai thác các đặc trưng cục bộ và phụ thuộc thời gian của tín hiệu gia tốc, đồng thời kiểm soát số lượng tham số và chi phí tính toán theo định hướng triển khai trên thiết bị tài nguyên hạn chế.

Mô hình đề xuất được xây dựng nhằm giải quyết đồng thời hai yêu cầu: nâng cao hiệu quả nhận dạng các tư thế ngủ chi tiết và kiểm soát độ phức tạp mô hình theo định hướng triển khai trên thiết bị cảm biến đeo hoặc thiết bị có tài nguyên hạn chế. AnpoNet kết hợp khả năng trích xuất đặc trưng cục bộ của mạng nơ-ron tích chập một chiều với khả năng mô hình hóa phụ thuộc thời gian của mạng hồi quy, qua đó khai thác được đặc trưng không gian–thời gian của tín hiệu gia tốc.

Đề xuất mô hình CGB-Net, một kiến trúc học sâu lai kết hợp 1D CNN, GRU và BiLSTM nhằm tăng cường năng lực biểu diễn đặc trưng không gian–thời gian của

tín hiệu gia tốc. Mô hình này được xây dựng nhằm cải thiện hiệu quả phân loại trong điều kiện dữ liệu mở rộng và được so sánh với các mô hình cơ sở cũng như mô hình AnpoNet.

Mô hình CGB-Net được thiết kế trên cơ sở kết hợp ba thành phần gồm mạng nơ-ron tích chập một chiều, GRU và BiLSTM. Trong đó, mạng tích chập một chiều được sử dụng để trích xuất đặc trưng cục bộ từ tín hiệu gia tốc, GRU hỗ trợ mô hình hóa quan hệ tuần tự với chi phí tham số hợp lý, còn BiLSTM khai thác ngữ cảnh thời gian hai chiều nhằm tăng cường khả năng phân biệt giữa các tư thế có đặc trưng gần nhau.

Các đóng góp này có giá trị cả về mặt phương pháp và định hướng triển khai trên các thiết bị có tài nguyên hạn chế.

4. Đánh giá về sự trùng lặp của luận án so với các đồ án, luận văn hay công trình khoa học đã công bố trong và ngoài nước

Theo hiểu biết của người nhận xét, nội dung và các kết quả chính của luận án không trùng lặp với các luận án, luận văn hay công trình khoa học đã công bố trong và ngoài nước. Không nhận thấy dấu hiệu sao chép hoặc lặp lại một cách không phù hợp các kết quả nghiên cứu đã công bố.

5. Nhận xét về chất lượng các bài báo khoa học đã được công bố và khẳng định các bài báo đó chứa đựng nội dung chủ yếu của luận án

Các kết quả của luận án đã được công bố trong 05 công trình khoa học, trong đó có các bài báo thuộc danh mục SCI/ISI, Scopus Q1, Q2 và hội nghị quốc tế.

Các công bố có chất lượng tốt và phù hợp với hướng nghiên cứu của luận án. Nội dung của các bài báo phản ánh các nội dung trong luận án.

6. Tính trung thực trong việc trích dẫn các công trình đã được NCS công bố trong và ngoài nước, tài liệu tham khảo

Các kết quả nghiên cứu của luận án đã được trích dẫn trung thực trong các công trình công bố. Trích dẫn tài liệu tham khảo được thực hiện một cách rõ ràng, đầy đủ và cập nhật.

7. Những ưu điểm và thiếu sót, những điểm cần được bổ sung và sửa chữa

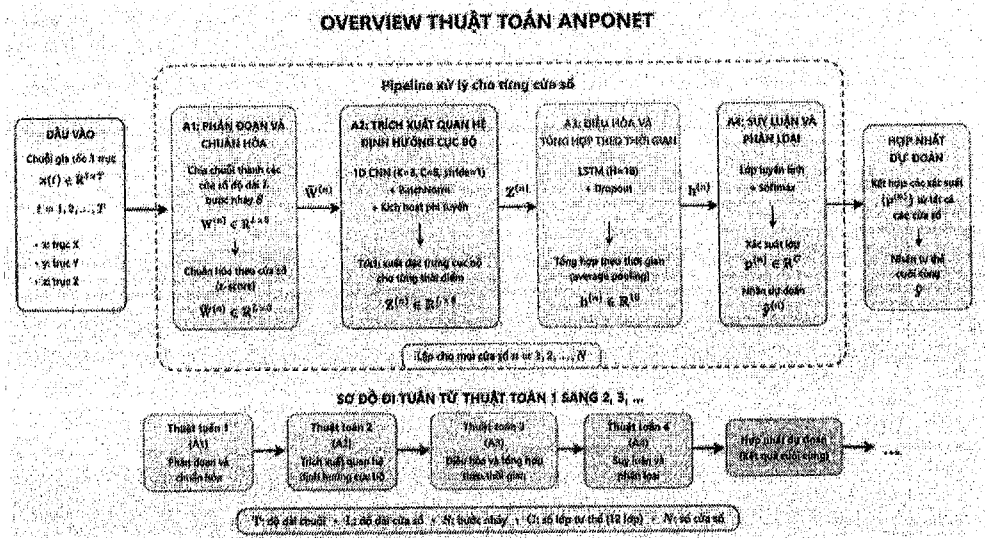
Luận án có bố cục tương đối hợp lý, nội dung nghiên cứu được trình bày có hệ thống và các kết quả thực nghiệm được phân tích khá đầy đủ. Một số cần giải thích:

Môi trường suy diễn

Để thu được các thông số đo độ trễ thực tế sát với môi trường triển khai, tất cả các thử nghiệm suy luận đều được thực hiện trên một hệ thống máy tính xách tay chỉ dùng CPU. Thiết bị được cài đặt Windows 10 (tần số cơ bản 1.70 GHz, xung nhịp tăng tốc 2.70 GHz), RAM 16GB, ổ cứng SSD 1TB, hệ điều hành Windows 64-bit. Không có GPU hỗ trợ (như chuyển đổi TensorBoard sang GPU). Môi trường suy diễn cần nêu khái niệm và đặt tên cho phù hợp cho mô hình suy luận quá trình suy luận.

Thời gian suy luận của mỗi mô hình là độ trễ trung bình của quá trình truyền tải qua 1,000 lần chạy độc lập, sử dụng kích thước batch là 1. Giao thức này phản ánh các điều kiện xử lý luồng thực tế trong các ứng dụng của thiết bị đeo, luồng dữ liệu từ thế đến tuần tự thay vì theo các batch lớn. Tất cả các mô hình được so sánh trong luận án này đều được đánh giá trong cùng một môi trường suy luận để đảm bảo tính công bằng khi báo cáo hiệu quả tính toán và chi phí thời gian chạy.

Cần trình bày rõ hơn sự kết nối các thuật toán, các bước trung gian, vẽ lại sơ đồ dưới đây, hoặc bổ sung thêm trong mô hình kiến trúc hệ thống của các thuật toán này:



Hình 2.6: Tổng quan sự kết hợp giữa các thuật toán con để tạo nên khung thuật toán AnpoNet

Trong hình 2.6 thuật toán 5 chưa thể hiện trong hình và cần mô tả rõ các đầu vào, đầu ra

Thuật toán 5 A4: Suy luận và phân loại tư thế ngã

Input: Biểu diễn thời gian $h \in \mathbb{R}^{16}$

Output: Xác suất hậu nghiệm $p \in \mathbb{R}^{12}$ và nhân dự đoán $\hat{y}$

Lớp ánh xạ tuyến tính: Khởi tạo trọng số $W \in \mathbb{R}^{12 \times 16}$ và bias $b \in \mathbb{R}^{12}$

$$s = Wh + b$$

Chuẩn hóa xác suất (Softmax): for $k \leftarrow 1$ to 12 do

$$p_k = \frac{\exp(s_k)}{\sum_{j=1}^{12} \exp(s_j)}$$

Dự đoán nhân:

$$\hat{y} = \arg \max_{k \in \{1, \dots, 12\}} p_k$$

return ($p, \hat{y}$)

Thuật toán 6 cần mô tả đầu vào, đầu ra và kết nối vào trong mô hình kiến trúc hệ thống

Thuật toán 6 Tối ưu hóa tham số AnpoNet bằng Adam

Input: Gradient $\mathbf{g}_t = \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}_t)$, bước t , tham số hiện tại $\boldsymbol{\theta}_t$

Output: Tham số cập nhật $\boldsymbol{\theta}_{t+1}$

Cập nhật moment bậc nhất:

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t$$

Cập nhật moment bậc hai:

$$\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^{\odot 2}$$

Hiệu chỉnh sai lệch:

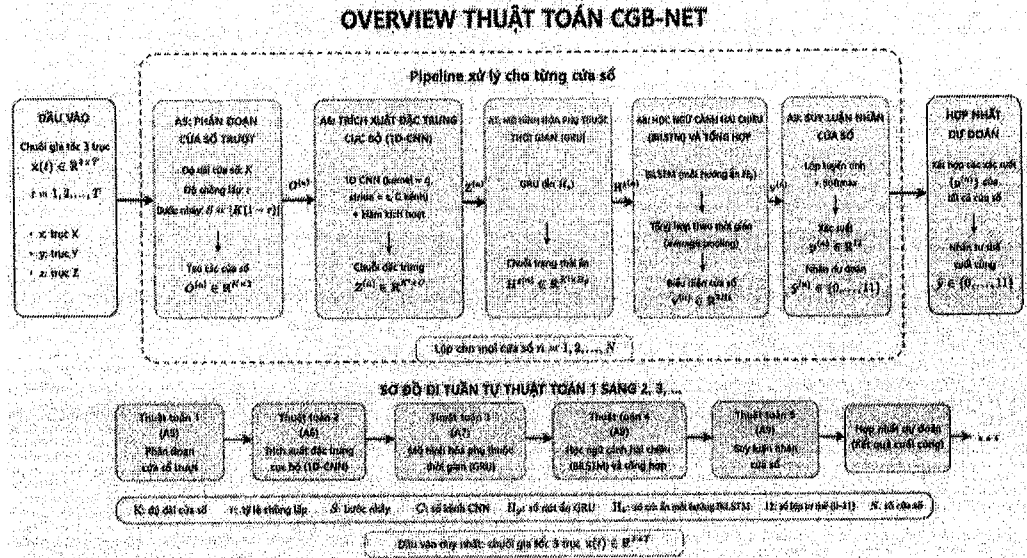
$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{1 - \beta_1^t}, \quad \hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_2^t}$$

Cập nhật tham số:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t} + \epsilon}$$

return $\boldsymbol{\theta}_{t+1}$

Chương 3 cũng cần mô hình kiến trúc hệ thống có thể thay thế hoặc bổ sung thêm so với hình dưới đây:



Hình 3.1: Tổng thể luồng khung thuật toán CGB-Net

Chưa rõ các kết nối trung gian giữa các thuật toán, cách lưu trữ các thuật toán và xử lý. Các thuật toán 7,8,9, 10, 11 chưa rõ sự kết nối và mô tả trong hình 3.1

Sau mỗi chương, nên có nhận xét, đánh giá các ưu điểm kết quả thực nghiệm và nêu rõ những nhược điểm, cách khắc phục

Tác giả nên làm rõ hơn các hạn chế này, đồng thời bổ sung thảo luận về khả năng triển khai thực tế trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.

Câu hỏi:

- 1- Thu thập dữ liệu ngủ tự nhiên qua đêm, xem xét các tư thế chuyển tiếp và các chuyển động không ổn định trong khi ngủ thì các mô hình sẽ triển khai như nào? Hãy giải thích?
- 2- So sánh các thiết bị đeo tay như đồng hồ thông minh, kết nối qua điện thoại có thể hơn nhằm hỗ trợ theo dõi sức khỏe và đánh giá hành vi ngủ của người dùng, các mô hình đề xuất có điểm gì khác biệt? Hãy giải thích
- 3- Các thuật toán trình bày trong các mô hình có thể hiện các kết quả đầu vào đầu ra thì kết quả trung gian có xử lý thêm bằng thao tác hay thực hiện tự động? Ưu và nhược của các thuật toán này là gì?

8. Kết luận

- Luận án đáp ứng được các yêu cầu về nội dung và hình thức đối với một luận án tiến sĩ, ngành Khoa học máy tính
 - Đồng ý cho NCS được bảo vệ tại Hội đồng chấm luận án cấp Học viện.
-

Hà nội, ngày 25 tháng 07 năm 2026

Người nhận xét



PGS. TS. Phạm Văn Hải

Hà Nội, ngày 26 tháng 08 năm 2026

**BIÊN BẢN CỦA
HỘI ĐỒNG ĐÁNH GIÁ LUẬN ÁN TIẾN SĨ CẤP HỌC VIỆN**

Căn cứ quyết định số 576/QĐ-HVKHCN ngày 25 tháng 06 năm 2026 của Giám đốc Học viện Khoa học và Công nghệ về việc thành lập Hội đồng đánh giá luận án tiến sĩ cấp Học viện, Hội đồng đã họp vào hồi 9 giờ ngày 26 tháng 08 năm 2026 tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, số 18 đường Hoàng Quốc Việt, Cầu Giấy, Hà Nội để đánh giá luận án tiến sĩ.

Họ và tên NCS: **Vũ Hoàng Diệu**

Tên đề tài luận án: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa và GS.TS. Trần Đức Tân

THAM DỰ BUỔI BẢO VỆ GỒM CÓ

- Đại diện cơ sở đào tạo:
 1. PGS.TS. Nguyễn Xuân Nhiệm – Phó Giám đốc Học viện KH&CN
- Đại diện Viện Công nghệ thông tin:
 1. PGS.TS. Nguyễn Trường Thắng - Viện trưởng Viện Công nghệ thông tin
- Đại diện Cơ quan chủ quản của NCS:
 1. GS.TS. Nguyễn Văn Hiếu, Phó giám đốc Đại học Phenikaa
- Thành viên Hội đồng có mặt: 7/7 thành viên
 1. PGS.TS. Nguyễn Việt Anh, Chủ tịch Hội đồng
 2. PGS.TS. Trần Giang Sơn, Phản biện 1



3. TS. Ngô Hải Anh, Phản biện 2
 4. TS. Nguyễn Như Sơn, Thư ký Hội đồng
 5. PGS.TS. Nguyễn Hà Nam, Ủy viên
 6. PGS.TS. Nguyễn Long Giang, Ủy viên
 7. PGS.TS. Phạm Văn Hải, Ủy viên
- Thành viên Hội đồng vắng mặt:
 - Đại diện tập thể cán bộ hướng dẫn: TS. Trần Đức Nghĩa, GS.TS. Trần Đức Tân
 - Cùng tham dự buổi bảo vệ còn có nhiều cán bộ nghiên cứu khoa học trong và ngoài Học viện.

TIẾN TRÌNH BUỔI BẢO VỆ

1. *Đại diện cơ sở đào tạo*, cô Phạm Thị Như Quỳnh, tuyên bố lý do, giới thiệu đại biểu và đọc quyết định số 576/QĐ-HVKHCN ngày 25 tháng 06 năm 2026 của Giám đốc HVKHCN về việc thành lập Hội đồng đánh giá luận án tiến sĩ cấp Học viện cho NCS Vũ Hoàng Diệu và đề nghị Chủ tịch Hội đồng điều khiển phiên họp.
2. *Chủ tịch Hội đồng*, PGS.TS. Nguyễn Việt Anh, công bố danh sách thành viên có mặt là 07, thông qua chương trình buổi bảo vệ, đề nghị Thư ký thông báo các điều kiện chuẩn bị cho buổi bảo vệ và đọc lý lịch khoa học của NCS.
3. *Thư ký Hội đồng*, TS. Nguyễn Như Sơn thông báo các điều kiện cho buổi bảo vệ
 - Đọc lý lịch khoa học của NCS Vũ Hoàng Diệu.
 - Đã nhận đủ 07 nhận xét của các phản biện và các thành viên HĐ.
 - Lịch bảo vệ của NCS đã được đăng trên Cổng thông tin điện tử Học viện Khoa học và Công nghệ ngày 31/07/2026.
 - Các giấy tờ cần thiết khác.

NCS Vũ Hoàng Diệu có đủ các điều kiện về thủ tục để bảo vệ luận án trước Hội đồng đánh giá luận án cấp Học viện.
4. Các thành viên hội đồng và những người tham dự thông qua về lý lịch khoa học và quá trình đào tạo của nghiên cứu sinh.
5. Nghiên cứu sinh Vũ Hoàng Diệu trình bày nội dung luận án trong 30 phút trước Hội đồng. Báo cáo của NCS bao gồm các nội dung chính như sau:

- Giới thiệu bài toán phân loại tư thế ngủ
 - Xây dựng bộ dữ liệu và đề xuất mạng học sâu kết hợp AnpoNet hướng tới ứng dụng trên thiết bị hiệu năng thấp
 - Đề xuất mạng học sâu kết hợp CGB-Net nâng cao hiệu quả phân loại tư thế ngủ
 - Kết quả và đóng góp
 - Danh mục các công bố.
6. Các phản biện đọc bản nhận xét và đặt câu hỏi đánh giá luận án của NCS Vũ Hoàng Diệu

1) **Phản biện 1, PGS.TS. Trần Giang Sơn**, đọc nhận xét đánh giá luận án và kết luận (có văn bản kèm theo).

Ưu điểm:

Bổ cục hợp lý, đề xuất phương pháp rõ ràng, đã xây dựng mô hình toán học, phân tích độ phức tạp của thuật toán, so sánh kết quả với các phương pháp tiên tiến cho từng mô hình đề xuất. Kết quả đạt được cũng tương đối cạnh tranh.

Nhược điểm:

- Nên bổ sung so sánh với các mô hình sử dụng thị giác máy tính để nhận diện tư thế ngủ (không xâm lấn, không gây ảnh hưởng đến giấc ngủ...).
- Một số hình và bảng chưa rõ, khó diễn giải (ví dụ hình 2.6, hình 3.1, hình 3.3, bảng 3.3). Việc dùng nét đứt để thể hiện loss và accuracy trong các biểu đồ Training and Validation làm biểu đồ thiếu tính chính xác.
- Với định hướng triển khai trên thiết bị cảm biến đeo hoặc thiết bị edge có tài nguyên hạn chế, luận án nên bổ sung phân tích kết quả thực nghiệm trên các thiết bị phần cứng yếu như ESP.
- Dữ liệu hiện tại chủ yếu là các tư thế ngủ tĩnh, được thu thập trong điều kiện có kiểm soát (60s), chưa phản ánh đầy đủ các tư thế chuyển tiếp hay các chuyển động phức tạp trong một giấc ngủ qua đêm tự nhiên.

2) **Phản biện 2, TS. Ngô Hải Anh**, đọc nhận xét đánh giá luận án và kết luận (có văn bản kèm theo).

Ưu điểm:

Về tổng thể, luận án có mạch nghiên cứu tương đối rõ: Chương 1 xác lập bài toán và cơ sở lựa chọn phương pháp; Chương 2 phát triển mô hình gọn nhẹ AnpoNet; Chương 3 phát triển CGB-Net theo hướng tăng năng lực biểu diễn.

Nhược điểm:

- Chương 2 và Chương 3 sử dụng hai tập đối tượng độc lập, luận án cần nhấn mạnh hơn rằng các kết quả tuyệt đối giữa hai chương không nên so sánh trực tiếp; việc so sánh AnpoNet và CGB-Net chỉ có ý nghĩa khi hai mô hình được đánh giá trên cùng thiết lập dữ liệu.
- Đối với Chương 1: Phần tổng quan khá đầy đủ, nhưng khoảng trống nghiên cứu nên được kết luận cô đọng hơn và liên hệ trực tiếp với hai đóng góp của luận án. Lý do lựa chọn đúng 12 tư thế, tương ứng các mức góc 30° , cần được giải thích chặt hơn. Phần chỉ số đánh giá cũng cần nêu rõ F1-score sử dụng trong bài toán đa lớp là macro, micro hay weighted để tránh cách hiểu khác nhau.
- Đối với Chương 2: Đây là chương được tổ chức tương đối đầy đủ từ thiết bị, dữ liệu đến mô hình và thực nghiệm. Tuy nhiên, Dataset A chỉ gồm 15 người, độ tuổi 19-24, nên các nhận xét về tính đa dạng của nhóm khảo sát cần thận trọng hơn. NCS nên mô tả rõ hơn sai số cho phép khi thiết lập/gán nhãn góc tư thế, quy trình tiền xử lý và chuẩn hóa. Các số liệu về bộ nhớ, thời gian suy luận và năng lượng cần nói rõ là đo trực tiếp hay ước lượng, trên nền tảng phần cứng nào. Vì chưa triển khai trực tiếp trên vi điều khiển hoặc thiết bị đeo công suất thấp, nên thống nhất dùng cách diễn đạt “hướng tới triển khai trên thiết bị biên”.
- Đối với Chương 3: NCS cần giải thích rõ hơn vai trò của việc sử dụng đồng thời GRU và BiLSTM và lợi ích thu được so với mức tăng về độ phức tạp. Kết quả nên được thảo luận như một sự đánh đổi giữa độ chính xác/F1 và chi phí tính toán. Một số lớp tư thế vẫn có tỷ lệ nhầm lẫn cao hơn các lớp khác; đây là giới hạn nên được nêu thẳng và phân tích gắn với sự gần nhau về góc, khác biệt giữa đối tượng và sai lệch vị trí cảm biến.

- Đối với phần Kết luận: Nên tách rõ những nội dung đã được chứng minh bằng thực nghiệm với những khả năng dự kiến trong tương lai. Các kết luận về ứng dụng lâm sàng, theo dõi ngu tự nhiên và triển khai trên phần cứng nhúng cần giữ đúng phạm vi của nghiên cứu hiện tại.
- Về hình thức trình bày: Bản thảo còn một số lỗi biên tập và thiếu nhất quán cần rà soát toàn văn. Ví dụ: mục lục ghi “LỜI CẢM ƠN iii” trong khi trang Lời cảm ơn là ii; “Confusion Matrix ((ma trận nhầm lẫn)” thừa dấu ngoặc; có lỗi dạng “§ố tham số”; cách viết Seed/seed, K-fold/K-Folds, F1-score/F1-Score, FLOPs/FLOPS chưa thống nhất. Cũng cần thống nhất cách dùng Accuracy và Precision trong tiếng Việt, cách ghi đơn vị, dấu %, ký hiệu góc và tên hình/bảng. Đây là các lỗi không ảnh hưởng đến kết quả nghiên cứu nhưng cần được sửa trong bản hoàn thiện.

NCS Vũ Hoàng Diệu tiếp thu ý kiến nhận xét của các phản biện và trả lời đầy đủ câu hỏi của các uỷ viên phản biện.

7. Các thành viên khác trong Hội đồng đưa ra ý kiến nhận xét; Hội đồng và những người tham dự đặt câu hỏi

- PGS.TS. Nguyễn Việt Anh

Luận án có bố cục tương đối hợp lý, nội dung các chương được tổ chức theo trình tự từ tổng quan, xây dựng mô hình đến thực nghiệm và đánh giá kết quả. Hệ thống bảng, hình và các kết quả thực nghiệm khá phong phú, hỗ trợ tốt cho việc trình bày nội dung nghiên cứu. Tuy nhiên, tác giả nên tiếp tục rà soát và chỉnh sửa một số vấn đề về hình thức trình bày để luận án hoàn thiện hơn.

Một số ý kiến chỉnh sửa, bổ sung như sau:

- Rà soát lại toàn bộ lỗi chính tả, lỗi đánh máy, dấu câu và khoảng trắng; đặc biệt cần thống nhất cách viết các thuật ngữ tiếng Anh, chữ viết tắt và tên các mô hình trong toàn luận án.
- Cần thống nhất cách Việt hóa và sử dụng thuật ngữ chuyên môn, tránh trường hợp cùng một khái niệm nhưng được diễn đạt theo nhiều cách khác nhau ở các chương.

- Một số bảng có nhiều thông tin và số liệu, nên cân nhắc điều chỉnh cách căn chỉnh, số chữ số thập phân và cách nhấn mạnh kết quả tốt nhất để người đọc dễ theo dõi, so sánh.
- Cần kiểm tra lại tiêu đề, chú thích và cách đánh số của các bảng và hình vẽ, bảo đảm tất cả các bảng, hình đều được dẫn chiếu và giải thích đầy đủ trong phần nội dung.
- Một số hình biểu diễn kết quả thực nghiệm nên tăng kích thước chữ, nhấn trục và chú giải; đồng thời thống nhất cách trình bày giữa các hình có cùng loại kết quả.

- TS. Nguyễn Như Sơn

Luận án có bố cục tương đối hợp lý, nội dung nghiên cứu được trình bày có hệ thống và các kết quả thực nghiệm được phân tích khá đầy đủ.

Một số nhận xét có thể bổ sung chỉnh sửa như sau:

- Dữ liệu hiện chủ yếu được thu thập trong điều kiện có kiểm soát, chưa phản ánh đầy đủ các tư thế chuyển tiếp và giấc ngủ tự nhiên qua đêm. Hầu hết các kết quả đánh giá được thực hiện trên bộ dữ liệu do nhóm nghiên cứu xây dựng. Nếu mở rộng ra các tập dữ liệu khác thì kết quả sẽ như thế nào?
- Làm rõ thêm về tính mới về mặt phương pháp khi kết hợp các thành phần trong các mô hình đề xuất.
- Khả năng triển khai thực tế trên thiết bị biên, thiết bị đeo công suất thấp.
- Nên bỏ Kiến nghị ở phần Kết luận cuối, hướng phát triển hiện khá ngắn gọn, có thể bổ sung đầy đủ hơn.
- Các công trình công bố nên viết SCIE/Scopus nếu trong các danh mục này.

- PGS.TS. Nguyễn Hà Nam

Ưu điểm:

- Cấu trúc luận án mạch lạc, mạch đóng góp xuyên suốt từ tổng quan (Chương 1) đến hai mô hình đề xuất (Chương 2, 3) và kết luận.
- Quy trình xây dựng bộ dữ liệu, thu thập, gán nhãn và tiền xử lý được mô tả chi tiết, có khả năng tái lập.
- Thực nghiệm được thiết kế với mức độ chính xác cao hơn mặt bằng chung: đánh giá độc lập theo đối tượng, kiểm định chéo, phân tích độ ổn định theo seed, phân tích lỗi theo từng lớp và đo đặc chỉ phí triển khai.

- Luận án cũng thể hiện sự trung thực khoa học khi chủ động phân biệt kết quả từ các nghiên cứu tiền khả thi (CT4, CT5 - dữ liệu trộn đối tượng) với kết quả chính thức (Chương 2, 3 - đánh giá độc lập theo đối tượng), thay vì gộp chung để tạo ấn tượng tốt hơn.

Nhược điểm:

- Luận án chưa sử dụng kiểm định thống kê để củng cố các so sánh hiệu năng giữa mô hình đề xuất và mô hình cơ sở, trong khi độ biến động giữa các lần chạy không nhỏ (mục 4.i).
- Cỡ mẫu thực nghiệm còn khiêm tốn và đồng nhất về nhân khẩu học so với phạm vi ứng dụng lâm sàng được đặt ra ở phần mở đầu.
- Mối quan hệ giữa 5 công trình công bố và 3 chương chính thức chưa được trình bày tường minh trong luận án (mục 3), gây khó khăn cho người đọc khi đối chiếu.
- Dữ liệu thực nghiệm là các tư thế tĩnh thu thập có kiểm soát, chưa phản ánh dữ liệu ngủ tự nhiên qua đêm - hạn chế này đã được chính NCS nêu ở phần Kết luận, thể hiện sự tự nhận thức tốt, song cũng cho thấy khoảng cách còn khá xa giữa kết quả labo và mục tiêu ứng dụng lâm sàng ban đầu.

- PGS.TS. Nguyễn Long Giang

Luận án có bố cục tương đối hợp lý, nội dung nghiên cứu được trình bày có hệ thống và các kết quả thực nghiệm được phân tích khá đầy đủ. Tuy nhiên, dữ liệu hiện chủ yếu được thu thập trong điều kiện có kiểm soát, chưa phản ánh đầy đủ các tư thế chuyển tiếp và giấc ngủ tự nhiên qua đêm. Tác giả nên làm rõ hơn các hạn chế này, đồng thời bổ sung thảo luận về khả năng triển khai thực tế trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.

- PGS.TS. Phạm Văn Hải

Cần trình bày rõ hơn sự kết nối các thuật toán, các bước trung gian, vẽ lại sơ đồ dưới đây, hoặc bổ sung thêm trong mô hình kiến trúc hệ thống của các thuật toán.

Sau mỗi chương, nên có nhận xét, đánh giá các ưu điểm kết quả thực nghiệm và nêu rõ những nhược điểm, cách khắc phục.

Tác giả nên làm rõ hơn các hạn chế này, đồng thời bổ sung thảo luận về khả năng triển khai thực tế trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.

Câu hỏi:

- 1- Thu thập dữ liệu ngẫu nhiên qua đêm, xem xét các tư thế chuyển tiếp và các chuyển động không ổn định trong khi ngủ thì các mô hình sẽ triển khai như nào? Hãy giải thích?
 - 2- So sánh các thiết bị đeo tay như đồng hồ thông minh, kết nối qua điện thoại có thể hơn nhằm hỗ trợ theo dõi sức khỏe và đánh giá hành vi ngủ của người dùng, các mô hình đề xuất có điểm gì khác biệt? Hãy giải thích
 - 3- Các thuật toán trình bày trong các mô hình có thể hiện các kết quả đầu vào đầu ra thì kết quả trung gian có xử lý thêm bằng thao tác hay thực hiện tự động? Ưu và nhược của các thuật toán này là gì?
8. Thư ký đọc nhận xét của 02 chuyên gia phản biện độc lập.
 9. NCS Vũ Hoàng Diệu tiếp thu ý kiến nhận xét của các thành viên của Hội đồng. NCS trả lời các câu hỏi của các thành viên Hội đồng và trình bày giải trình nhận xét của 02 chuyên gia phản biện độc lập.
 10. Đại diện tập thể hướng dẫn phát biểu ý kiến bằng văn bản.
 11. Hội đồng tiến hành họp riêng để bầu ban kiểm phiếu, bỏ phiếu kín và thảo luận thông qua quyết nghị của Hội đồng.
 - 1) Bầu ban kiểm phiếu gồm:
 - Trưởng ban: PGS.TS. Phạm Văn Hải
 - Ủy viên: PGS.TS. Nguyễn Hà Nam
 - Ủy viên: TS. Nguyễn Như Sơn
 - 2) Bỏ phiếu kín và thảo luận thông qua Quyết nghị của Hội đồng.
 - Trưởng ban kiểm phiếu, PGS.TS. Phạm Văn Hải công bố kết quả kiểm phiếu (có biên bản kiểm phiếu).
 - Chủ tịch Hội đồng, PGS.TS. Nguyễn Việt Anh, thông qua Quyết nghị (có văn bản kèm theo).
 - 3) Tóm tắt nghị quyết của Hội đồng
 - 3.1. *Tính phù hợp của tên đề tài và sự không trùng lặp về nội dung luận án*
 - Tên đề tài, nội dung và kết quả nghiên cứu của luận án phù hợp với Ngành đào tạo “Khoa học máy tính”, mã số “9 48 01 01”.
 - Nội dung của luận án không trùng lặp với các luận án đã bảo vệ và các kết quả nghiên cứu đã công bố trong và ngoài nước.

- Các tài liệu tham khảo của luận án có nội dung phù hợp và đã được trích dẫn trong luận án.

3.2. Các kết quả chính của luận án

Luận án nghiên cứu về các mô hình học sâu lai ghép ứng dụng cho phân loại các tư thế ngủ sử dụng cảm biến gia tốc. Kết quả nghiên cứu của Luận án có ý nghĩa khoa học và thực tiễn, cụ thể:

- Đề xuất mạng kết hợp các mô hình học sâu (AnpoNet) nhằm phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục, hướng tới thực thi trên các thiết bị có tài nguyên hạn chế.
- Đề xuất mạng kết hợp các mô hình học sâu (CGB-Net) nhằm cải thiện độ chính xác phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục.

3.3. Các điểm cần bổ sung chỉnh sửa trước khi nộp luận án cho Thư viện Quốc gia Việt Nam

NCS cần tiếp thu, rà soát, chỉnh sửa, bổ sung nội dung luận án theo ý kiến đóng góp trong bản nhận xét của các thành viên Hội đồng và Biên bản của Hội đồng đánh giá luận án tiến sĩ cấp Học viện trước khi nộp luận án cho Thư viện Quốc gia Việt Nam.

3.4. Mức độ đáp ứng yêu cầu của luận án tiến sĩ về cả nội dung và hình thức

- Luận án của NCS Vũ Hoàng Diệu đáp ứng yêu cầu của luận án tiến sĩ ngành “Khoa học máy tính”, mã số “9 48 01 01” về nội dung và hình thức theo các qui chế hiện hành về đào tạo tiến sĩ của Bộ Giáo dục và Đào tạo, của Học viện Khoa học và Công nghệ.
- Đề nghị Học viện Khoa học và Công nghệ công nhận kết quả bảo vệ và cấp bằng tiến sĩ cho NCS Vũ Hoàng Diệu sau khi chỉnh sửa, bổ sung luận án theo các góp ý của Hội đồng.

12. Tổng kết

- Trưởng ban kiểm phiếu, PGS.TS. Phạm Văn Hải, công bố kết quả bỏ phiếu đánh giá luận án.
- Chủ tịch Hội đồng, PGS.TS. Nguyễn Việt Anh, đọc Quyết nghị của Hội đồng.

- Chủ tịch Hội đồng tuyên bố Hội đồng đã hoàn thành nhiệm vụ và trao lại quyền điều khiển cho Cơ sở đào tạo.
- Các đại biểu và NCS phát biểu ý kiến.
- Đại diện cơ sở đào tạo tuyên bố kết thúc buổi bảo vệ luận án tiến sĩ.

Biên bản họp Hội đồng này được 7/7 ủy viên Hội đồng biểu quyết công khai thông qua.

Buổi họp Hội đồng đánh giá luận án tiến sĩ cấp Học viện kết thúc vào lúc 10 giờ 30 phút, ngày 26 tháng 08 năm 2026.

Thư ký Hội đồng



TS. Nguyễn Như Sơn

Chủ tịch Hội đồng



PGS.TS. Nguyễn Việt Anh

XÁC NHẬN CỦA CƠ SỞ ĐÀO TẠO

**KT. GIÁM ĐỐC
PHÓ GIÁM ĐỐC**



Nguyễn Thị Trung



Hà Nội, ngày 26 tháng 08 năm 2026

**QUYẾT NGHỊ CỦA
HỘI ĐỒNG ĐÁNH GIÁ LUẬN ÁN TIẾN SĨ CẤP HỌC VIỆN**

Căn cứ quyết định số 576/QĐ-HVKHCN ngày 25 tháng 06 năm 2026 của Giám đốc Học viện Khoa học và Công nghệ về việc thành lập Hội đồng đánh giá luận án tiến sĩ cấp Học viện, Hội đồng đã họp vào hồi 9 giờ 00 ngày 26 tháng 08 năm 2026 tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, số 18 đường Hoàng Quốc Việt, Cầu Giấy, Hà Nội để đánh giá luận án tiến sĩ.

Họ và tên NCS: **Vũ Hoàng Diệu**

Tên đề tài luận án: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn: TS. Trần Đức Nghĩa và GS.TS. Trần Đức Tân

**HỘI ĐỒNG ĐÁNH GIÁ LUẬN ÁN TIẾN SĨ CẤP HỌC VIỆN CỦA
NCS VŨ HOÀNG DIỆU KẾT LUẬN**

1. Tính phù hợp của tên đề tài và sự không trùng lặp về nội dung luận án

- Tên đề tài, nội dung và kết quả nghiên cứu của luận án phù hợp với ngành đào tạo Hệ thống thông tin, mã số 9 48 01 01.

- Nội dung của luận án không trùng lặp với các luận án đã bảo vệ và các kết quả nghiên cứu đã công bố trong và ngoài nước.

- Các tài liệu tham khảo của luận án có nội dung phù hợp và đã được trích dẫn trong luận án.

2. Kết quả, ý nghĩa khoa học, thực tiễn của đề tài

Luận án nghiên cứu về các mô hình học sâu lai ghép ứng dụng cho phân loại các tư thế ngủ sử dụng cảm biến gia tốc. Kết quả nghiên cứu của Luận án có ý nghĩa khoa học và thực tiễn, cụ thể:



- Đề xuất mạng kết hợp các mô hình học sâu (AnpoNet) nhằm phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục, hướng tới thực thi trên các thiết bị có tài nguyên hạn chế.
- Đề xuất mạng kết hợp các mô hình học sâu (CGB-Net) nhằm cải thiện độ chính xác phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục.

3. Những thiếu sót của luận án, vấn đề cần bổ sung, sửa chữa

NCS cần tiếp thu, rà soát, chỉnh sửa, bổ sung nội dung luận án theo ý kiến đóng góp trong bản nhận xét của các thành viên Hội đồng và Biên bản của Hội đồng đánh giá luận án tiến sĩ cấp Học viện trước khi nộp luận án cho Thư viện Quốc gia Việt Nam.

4. Mức độ đáp ứng yêu cầu của luận án tiến sĩ về cả nội dung và hình thức theo các quy chế hiện hành về đào tạo tiến sĩ của Bộ Giáo dục và Đào tạo

Luận án của NCS Vũ Hoàng Diệu đáp ứng yêu cầu của một luận án tiến sĩ ngành “Khoa học máy tính”, mã số 9 48 01 01 về nội dung và hình thức theo các quy chế hiện hành về đào tạo tiến sĩ của Bộ Giáo dục và Đào tạo và của Học viện Khoa học và Công nghệ.

Kết luận:

Kết quả bỏ phiếu đánh giá luận án của Hội đồng: 7/7 thành viên tán thành.

Hội đồng kết luận thông qua luận án, đề nghị Học viện Khoa học và Công nghệ công nhận kết quả bảo vệ và cấp bằng tiến sĩ cho NCS Vũ Hoàng Diệu.

Quyết nghị này được 7/7 thành viên Hội đồng biểu quyết công khai thông qua.

THƯ KÝ



TS. Nguyễn Như Sơn

CHỦ TỊCH



PGS.TS. Nguyễn Việt Anh

XÁC NHẬN CỦA CƠ SỞ ĐÀO TẠO

KT. GIÁM ĐỐC
PHÓ GIÁM ĐỐC



Nguyễn Thị Trung

VIỆN HÀN LÂM
KHOA HỌC VÀ CÔNG NGHỆ VN
HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

ĐẢNG CỘNG SẢN VIỆT NAM

**BẢN GIẢI TRÌNH CHỈNH SỬA, BỔ SUNG LUẬN ÁN TIẾN SĨ
CẤP HỌC VIỆN**

Ngày 26 tháng 08 năm 2026, Học viện Khoa học và Công nghệ đã tổ chức đánh giá luận án tiến sĩ cấp Học viện cho nghiên cứu sinh Vũ Hoàng Diệu theo Quyết định số 576/QĐ-HVKHCN ngày 25 tháng 06 năm 2026 của Giám đốc Học viện.

Đề tài: Nâng cao hiệu quả phân loại tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục theo cách tiếp cận kết hợp các mô hình học sâu

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Người hướng dẫn khoa học: TS. Trần Đức Nghĩa, GS.TS. Trần Đức Tân

Theo Biên bản của Hội đồng, NCS phải bổ sung và chỉnh sửa luận án các điểm sau đây:

STT	Nội dung đề nghị chỉnh sửa, bổ sung	Nội dung đã được chỉnh sửa, bổ sung (Ghi rõ số trang/chương/mục... đã được chỉnh sửa)
1	Nên bổ sung so sánh với các mô hình sử dụng thị giác máy tính để nhận diện tư thế ngủ (không xâm lấn, không gây ảnh hưởng đến giấc ngủ...).	NCS tiếp thu ý kiến của Hội đồng và đã bổ sung tại Mục 1.2 phần tổng quan và so sánh các hệ thống nhận dạng tư thế ngủ dựa trên thị giác sử dụng camera RGB, camera hồng ngoại và camera chiều sâu. Luận án đã làm rõ ưu điểm của phương pháp thị giác là không yêu cầu gắn cảm biến trực tiếp lên cơ thể, đồng thời phân tích các hạn chế liên quan đến điều kiện ánh sáng, che khuất bởi chăn/mền và quyền riêng tư. Luận án cũng bổ sung so sánh định lượng từ nghiên cứu [43], trong đó phương pháp thị giác VGG16 đạt độ chính xác 93,49%, trong khi phương pháp dựa trên cảm biến gia tốc đạt trên 99,8%; dữ liệu ảnh và thời gian huấn luyện cũng lớn hơn đáng kể. Các kết quả này được diễn giải trong phạm vi thiết lập thực nghiệm tương ứng, không xem là kết luận tuyệt đối giữa hai phương thức.
2	Một số hình và bảng chưa rõ, khó diễn giải (ví dụ hình 2.6, hình 3.1, hình 3.3, bảng 3.3). Việc dùng nét đứt để thể hiện loss và	NCS tiếp thu ý kiến của Hội đồng và đã bổ sung phần diễn giải chi tiết cho các hình và bảng được góp ý. Đối với Hình 2.6, luận án làm rõ luồng xử lý và vai trò của các nhóm



	accuracy trong các biểu đồ Training and Validation làm biểu đồ thiếu tính chính xác.	chức năng trong AnpoNet; đối với Hình 3.1, làm rõ vai trò hỗ trợ của 1D CNN, GRU và BiLSTM; đối với Bảng 3.3, bổ sung phân tích đồng thời hiệu suất phân loại và chi phí tính toán; đối với Hình 3.3, bổ sung phân tích các lớp có mức nhầm lẫn cao. Đối với các đường cong huấn luyện, luận án bổ sung giải thích các trục biểu diễn loss và accuracy, cách phân biệt các đường huấn luyện/xác thực và xu hướng hội tụ của mô hình.
3	Với định hướng triển khai trên thiết bị cảm biến đeo hoặc thiết bị edge có tài nguyên hạn chế, luận án nên bổ sung phân tích kết quả thực nghiệm trên các thiết bị phần cứng yếu như ESP.	NCS tiếp thu ý kiến của Hội đồng và đã làm rõ vai trò của ESP8266 trong hệ thống. ESP8266 hiện chỉ được sử dụng cho thu nhận, lưu trữ tạm thời và truyền dữ liệu, chưa thực hiện suy luận mô hình. Luận án đã bổ sung thông tin về môi trường đo thời gian suy luận trên CPU Intel Core i5-8350U và nêu rõ đây không phải kết quả đo trên ESP8266 hoặc vi điều khiển. Khả năng triển khai trực tiếp trên phần cứng nhúng, sau các bước lượng tử hóa, cắt tia và nén mô hình, được xác định là hướng nghiên cứu tiếp theo.
4	Dữ liệu hiện tại chủ yếu là các tư thế ngủ tĩnh, được thu thập trong điều kiện có kiểm soát (60s), chưa phản ánh đầy đủ các tư thế chuyển tiếp hay các chuyển động phức tạp trong một giấc ngủ qua đêm tự nhiên.	NCS tiếp thu ý kiến của Hội đồng và đã làm rõ phạm vi của dữ liệu thực nghiệm: 12 tư thế ngủ tĩnh được thu thập trong điều kiện có kiểm soát, chưa bao gồm đầy đủ các tư thế chuyển tiếp, chuyển động phức tạp và dữ liệu ngủ tự nhiên qua đêm. Hạn chế này được nêu tại phần Đối tượng và phạm vi nghiên cứu và được phân tích lại trong phần Kết luận. Đồng thời, luận án đã bổ sung hướng nghiên cứu tiếp theo về thu thập dữ liệu liên tục qua đêm, xây dựng nhãn cho các trạng thái chuyển tiếp và đánh giá tính ổn định của dự đoán theo thời gian.
5	Chương 2 và Chương 3 sử dụng hai tập đối tượng độc lập, luận án cần nhấn mạnh hơn rằng các kết quả tuyệt đối giữa hai chương không nên so sánh trực tiếp; việc so sánh AnpoNet và CGB-Net chỉ có ý nghĩa khi hai mô hình được đánh giá trên cùng thiết lập dữ liệu.	NCS tiếp thu ý kiến của Hội đồng và đã làm rõ Dataset A và Dataset B được xây dựng từ hai nhóm đối tượng độc lập, không trùng lặp. Vì vậy, các giá trị hiệu suất tuyệt đối giữa Chương 2 và Chương 3 không được sử dụng như một phép so sánh trực tiếp giữa AnpoNet và CGB-Net. Trong Chương 3, AnpoNet được đánh giá lại trên cùng Dataset B, cùng cách chia đối tượng và thiết lập tiền xử lý với

		CGB-Net; do đó, các so sánh giữa hai kiến trúc trong Chương 3 có ý nghĩa trực tiếp hơn.
6	Đối với Chương 1: Phần tổng quan khá đầy đủ, nhưng khoảng trống nghiên cứu nên được kết luận cô đọng hơn và liên hệ trực tiếp với hai đóng góp của luận án. Lý do lựa chọn đúng 12 tư thế, tương ứng các mức góc 30°, cần được giải thích chặt hơn. Phần chỉ số đánh giá cũng cần nêu rõ F1-score sử dụng trong bài toán đa lớp là macro, micro hay weighted để tránh cách hiểu khác nhau.	NCS tiếp thu ý kiến của Hội đồng và đã cô đọng khoảng trống nghiên cứu theo ba vấn đề chính: số lượng tư thế còn hạn chế trong nhiều nghiên cứu trước, thiếu đánh giá độc lập theo đối tượng và nhu cầu về mô hình có độ phức tạp được kiểm soát. Các khoảng trống này được liên hệ trực tiếp với hai nhóm đóng góp AnpoNet và CGB-Net. Luận án đồng thời bổ sung cơ sở lựa chọn 12 tư thế với độ phân giải góc 30° và làm rõ F1-score sử dụng trong bài toán đa lớp là Macro F1-score, được tính bằng trung bình F1-score của 12 lớp.
7	Đối với Chương 2: Đây là chương được tổ chức tương đối đầy đủ từ thiết bị, dữ liệu đến mô hình và thực nghiệm. Tuy nhiên, Dataset A chỉ gồm 15 người, độ tuổi 19-24, nên các nhận xét về tính đa dạng của nhóm khảo sát cần thận trọng hơn. NCS nên mô tả rõ hơn sai số cho phép khi thiết lập/gán nhãn góc tư thế, quy trình tiền xử lý và chuẩn hóa. Các số liệu về bộ nhớ, thời gian suy luận và năng lượng cần nói rõ là đo trực tiếp hay ước lượng, trên nền tảng phần cứng nào. Vì chưa triển khai trực tiếp trên vi điều khiển hoặc thiết bị đeo công suất thấp, nên thống nhất dùng cách diễn đạt “hướng tới triển khai trên thiết bị biên”.	NCS tiếp thu ý kiến của Hội đồng và đã điều chỉnh cách diễn giải về Dataset A theo hướng thận trọng, nêu rõ nhóm 15 người với khoảng tuổi tương đối hẹp chưa đại diện cho quần thể người dùng nói chung. Luận án đã mô tả rõ hơn quy trình gán nhãn có nhiều người đánh giá và nguồn video tham chiếu, quy trình tiền xử lý, chuẩn hóa và phân đoạn dữ liệu; đồng thời bổ sung thông tin chi tiết về môi trường huấn luyện và môi trường đo thời gian suy luận. Các kết quả về độ trễ được xác định là phép đo trực tiếp trên CPU thử nghiệm, không phải đo trên vi điều khiển hoặc thiết bị đeo công suất thấp; do đó luận án thống nhất sử dụng cách diễn đạt “hướng tới triển khai trên thiết bị biên”.
8	Đối với Chương 3: NCS cần giải thích rõ hơn vai trò của việc sử dụng đồng thời GRU và BiLSTM và lợi ích thu được so với mức tăng về độ phức tạp. Kết quả nên được thảo luận như một sự đánh đổi giữa độ chính xác/F1 và chi phí tính toán. Một số lớp tư thế vẫn có tỷ lệ nhầm lẫn cao hơn các lớp khác; đây là giới hạn nên được nêu thẳng và phân tích gắn với sự gần nhau về góc, khác biệt giữa đối tượng và sai lệch vị trí cảm biến.	NCS tiếp thu ý kiến của Hội đồng và đã bổ sung phân tích vai trò của các thành phần trong CGB-Net. 1D CNN được sử dụng để trích xuất đặc trưng cục bộ, GRU mô hình hóa các phụ thuộc tuần tự với cấu trúc tương đối gọn và BiLSTM khai thác ngữ cảnh hai chiều. Bảng 3.3 được thảo luận theo hướng đánh đổi giữa hiệu suất và chi phí: CGB-Net cải thiện khoảng 2,54 điểm phần trăm Accuracy và 6,66 điểm phần trăm Macro F1-score so với AnpoNet trên cùng thiết lập nhưng sử dụng nhiều tham số, FLOPs và bộ nhớ hơn. Luận

		án đồng thời phân tích các lớp RD, RU và UR có hiệu suất thấp hơn và liên hệ với khoảng cách góc nhỏ, sự khác biệt giữa đối tượng và sai lệch vị trí/hướng cảm biến.
9	Đối với phần Kết luận: Nên tách rõ những nội dung đã được chứng minh bằng thực nghiệm với những khả năng dự kiến trong tương lai. Các kết luận về ứng dụng lâm sàng, theo dõi ngủ tự nhiên và triển khai trên phần cứng nhưng cần giữ đúng phạm vi của nghiên cứu hiện tại.	NCS tiếp thu ý kiến của Hội đồng và đã chỉnh sửa phần Kết luận theo hướng phân biệt rõ kết quả đã được kiểm chứng với các khả năng ứng dụng trong tương lai. Các kết luận thực nghiệm được giới hạn ở phân loại 12 tư thế ngủ tĩnh, đánh giá độc lập theo đối tượng trong phạm vi hai bộ dữ liệu và các đặc trưng về độ phức tạp tính toán. Theo dõi ngủ tự nhiên qua đêm, xác nhận lâm sàng và triển khai hoàn chỉnh trên phần cứng nhúng được xác định là các hướng nghiên cứu tiếp theo, không phải kết quả đã được chứng minh trong luận án.
10	Về hình thức trình bày: Bản thảo còn một số lỗi biên tập và thiếu nhất quán cần rà soát toàn văn. Ví dụ: mục lục ghi “LỜI CẢM ƠN iii” trong khi trang Lời cảm ơn là ii; “Confusion Matrix ((ma trận nhầm lẫn)” thừa dấu ngoặc; có lỗi dạng “Số tham số”; cách viết Seed/seed, K-fold/K-Folds, F1-score/F1-Score, FLOPs/FLOPS chưa thống nhất. Cũng cần thống nhất cách dùng Accuracy và Precision trong tiếng Việt, cách ghi đơn vị, dấu %, ký hiệu góc và tên hình/bảng. Đây là các lỗi không ảnh hưởng đến kết quả nghiên cứu nhưng cần được sửa trong bản hoàn thiện.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát toàn văn về lỗi biên tập và tính nhất quán trình bày. Các lỗi về số trang Lời cảm ơn trong Mục lục, dấu ngoặc tại tên Mục 1.5.1, lỗi ký tự trong bảng, cách viết thuật ngữ, đơn vị, dấu phần trăm và ký hiệu góc đã được chỉnh sửa. Các thuật ngữ F1-score, FLOPs, seed và kiểm định chéo 5-fold được thống nhất trong toàn luận án; Accuracy, Precision và Recall được sử dụng nhất quán sau khi được định nghĩa.
11	Dữ liệu hiện chủ yếu được thu thập trong điều kiện có kiểm soát, chưa phản ánh đầy đủ các tư thế chuyển tiếp và giấc ngủ tự nhiên qua đêm. Hầu hết các kết quả đánh giá được thực hiện trên bộ dữ liệu do nhóm nghiên cứu xây dựng. Nếu mở rộng ra các tập dữ liệu khác thì kết quả sẽ như thế nào?	NCS tiếp thu ý kiến của Hội đồng. Luận án đã làm rõ rằng kết quả hiện tại được đánh giá trên hai bộ dữ liệu do nhóm nghiên cứu xây dựng và chưa chứng minh khả năng tổng quát hóa chéo bộ dữ liệu. Nội dung đánh giá cross-dataset đã được bổ sung vào hướng nghiên cứu tiếp theo: mô hình sẽ được huấn luyện trên dữ liệu hiện tại và kiểm tra trên bộ dữ liệu độc lập sau khi ánh xạ các nhãn tương thích; mức suy giảm

		Accuracy, Macro F1-score và ma trận nhầm lẫn sẽ được sử dụng để đánh giá ảnh hưởng của sự khác biệt miền dữ liệu.
12	Làm rõ thêm về tính mới về mặt phương pháp khi kết hợp các thành phần trong các mô hình đề xuất. Khả năng triển khai thực tế trên thiết bị biên, thiết bị đeo công suất thấp.	NCS tiếp thu ý kiến của Hội đồng và đã làm rõ tính mới về mặt thiết kế của hai mô hình. AnpoNet được xây dựng dưới ràng buộc tài nguyên ngay từ giai đoạn thiết kế, với các khối có chức năng rõ ràng nhằm cân bằng khả năng biểu diễn và độ gọn nhẹ. CGB-Net sử dụng cấu trúc phân cấp 1D CNN-GRU-BiLSTM, trong đó mỗi thành phần đảm nhiệm một vai trò bổ trợ thay vì chỉ tăng độ sâu mạng. Khả năng triển khai trên thiết bị biên hiện được đánh giá gián tiếp qua số tham số, FLOPs, bộ nhớ và độ trễ; triển khai trực tiếp trên thiết bị đeo công suất thấp được xác định là hướng nghiên cứu tiếp theo.
13	Nên bỏ Kiến nghị ở phần Kết luận cuối, hướng phát triển hiện khá ngắn gọn, có thể bổ sung đầy đủ hơn.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát, mở rộng đáng kể phần hạn chế và hướng nghiên cứu tiếp theo. Các hướng phát triển được cụ thể hóa gồm dữ liệu ngủ tự nhiên qua đêm, tư thế chuyển tiếp, mở rộng quy mô và mức độ đa dạng của đối tượng, đánh giá cross-dataset, kiểm định thống kê suy luận, lượng tử hóa, cắt tia, nén mô hình và triển khai trực tiếp trên phần cứng nhúng hoặc thiết bị đeo công suất thấp.
14	Các công trình công bố nên viết SCIE/Scopus nếu trong các danh mục này.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát, cập nhật tình trạng chỉ mục của các công trình trong Danh mục công trình của tác giả. Các công trình thuộc Science Citation Index Expanded được thống nhất ghi là SCIE, đồng thời thông tin Scopus và phân hạng tạp chí được ghi tương ứng đối với từng công trình.
15	Luận án chưa sử dụng kiểm định thống kê để củng cố các so sánh hiệu năng giữa mô hình đề xuất và mô hình cơ sở, trong khi độ biến động giữa các lần chạy không nhỏ (mục 4.i).	NCS tiếp thu ý kiến của Hội đồng. Trong phạm vi luận án hiện tại, độ ổn định của mô hình được đánh giá thông qua nhiều lần khởi tạo ngẫu nhiên, kiểm định chéo và báo cáo giá trị trung bình cùng độ lệch chuẩn; luận án chưa thực hiện kiểm định thống kê suy luận để xác định mức ý nghĩa của chênh lệch giữa các mô hình.

		Nội dung này đã được bổ sung vào hướng nghiên cứu tiếp theo, trong đó các kiểm định thống kê phù hợp sẽ được sử dụng để đánh giá liệu sự khác biệt về Accuracy và Macro F1-score có ý nghĩa thống kê hay chỉ phản ánh dao động ngẫu nhiên giữa các lần thực nghiệm.
16	Cỡ mẫu thực nghiệm còn khiêm tốn và đồng nhất về nhân khẩu học so với phạm vi ứng dụng lâm sàng được đặt ra ở phần mở đầu.	NCS tiếp thu ý kiến của Hội đồng và đã làm rõ giới hạn về quy mô và mức độ đa dạng của mẫu nghiên cứu. Dataset A gồm 15 người với khoảng tuổi tương đối hẹp và chưa đại diện cho quần thể nói chung; luận án vì vậy giới hạn các kết luận trong phạm vi nhóm đối tượng được khảo sát và nêu rõ chưa có xác nhận lâm sàng trên các nhóm bệnh nhân. Việc mở rộng số lượng người tham gia, nhóm tuổi, đặc điểm hình thể và thói quen ngủ được đưa vào hướng nghiên cứu tiếp theo.
17	Mối quan hệ giữa 5 công trình công bố và 3 chương chính thức chưa được trình bày tường minh trong luận án (mục 3), gây khó khăn cho người đọc khi đối chiếu.	NCS tiếp thu ý kiến của Hội đồng. Mối quan hệ giữa các công trình công bố và nội dung luận án đã được làm rõ tại các phần tương ứng: CT4 và CT5 được trình bày như các nghiên cứu tiền đề tại Chương 1; CT3 gắn với quá trình xây dựng hệ thống thu thập và bộ dữ liệu; CT1 gắn trực tiếp với AnpoNet và Chương 2; CT2 gắn trực tiếp với CGB-Net và Chương 3. Các chương cũng dẫn chiếu trực tiếp đến công trình tương ứng nhằm thể hiện tiến trình kế thừa và phát triển của nghiên cứu.
18	Dữ liệu thực nghiệm là các tư thế tĩnh thu thập có kiểm soát, chưa phản ánh dữ liệu ngủ tự nhiên qua đêm - hạn chế này đã được chính NCS nêu ở phần Kết luận, thể hiện sự tự nhận thức tốt, song cũng cho thấy khoảng cách còn khá xa giữa kết quả labo và mục tiêu ứng dụng lâm sàng ban đầu.	NCS tiếp thu ý kiến của Hội đồng và đã nêu rõ dữ liệu thực nghiệm hiện gồm các tư thế tĩnh được thu thập có kiểm soát, chưa phản ánh đầy đủ hành vi ngủ tự nhiên qua đêm. Luận án cũng giới hạn các bằng chứng lâm sàng ở vai trò động cơ khoa học, không xem kết quả phân loại hiện tại là bằng chứng về chẩn đoán hoặc điều trị. Khoảng cách giữa kết quả phòng thí nghiệm và ứng dụng thực tế được thảo luận trong phần hạn chế và được cụ thể hóa thành các hướng nghiên cứu tiếp theo.

19	Rà soát lại toàn bộ lỗi chính tả, lỗi đánh máy, dấu câu và khoảng trắng; đặc biệt cần thống nhất cách viết các thuật ngữ tiếng Anh, chữ viết tắt và tên các mô hình trong toàn luận án.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát toàn bộ luận án, chỉnh sửa các lỗi chính tả, lỗi đánh máy, dấu câu và khoảng trắng; đồng thời kiểm tra lại cách viết các thuật ngữ tiếng Anh, chữ viết tắt và tên các mô hình trong toàn văn.
20	Cần thống nhất cách Việt hóa và sử dụng thuật ngữ chuyên môn, tránh trường hợp cùng một khái niệm nhưng được diễn đạt theo nhiều cách khác nhau ở các chương.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát, thống nhất cách sử dụng và Việt hóa các thuật ngữ chuyên môn trong toàn luận án. Các thuật ngữ tiếng Anh được giữ lại đối với những khái niệm phổ biến trong học máy/học sâu và được giải thích ở lần xuất hiện đầu tiên, hạn chế việc sử dụng nhiều cách diễn đạt khác nhau cho cùng một khái niệm.
21	Một số bảng có nhiều thông tin và số liệu, nên cân nhắc điều chỉnh cách căn chỉnh, số chữ số thập phân và cách nhấn mạnh kết quả tốt nhất để người đọc dễ theo dõi, so sánh.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát cách trình bày các bảng số liệu, thống nhất cách ghi đơn vị, số chữ số thập phân và cách trình bày các chỉ số nhằm thuận tiện hơn cho việc đối chiếu giữa các mô hình. Đối với Bảng 3.3, luận án đồng thời bổ sung phân phân tích chi tiết sự đánh đổi giữa Accuracy/Macro F1-score và số lượng tham số, FLOPs, bộ nhớ của các kiến trúc.
22	Cần kiểm tra lại tiêu đề, chú thích và cách đánh số của các bảng và hình vẽ, bảo đảm tất cả các bảng, hình đều được dẫn chiếu và giải thích đầy đủ trong phần nội dung.	NCS tiếp thu ý kiến của Hội đồng và đã rà soát hệ thống tiêu đề, chú thích, đánh số và dẫn chiếu hình/bảng. Đồng thời, luận án bổ sung phần diễn giải chi tiết cho các hình/bảng có nội dung phức tạp, tiêu biểu là Hình 2.6, Hình 3.1, Bảng 3.3, Hình 3.3 và Hình 3.4, nhằm giúp người đọc hiểu rõ hơn vai trò của các thành phần, ý nghĩa các trục và xu hướng kết quả.
23	Một số hình biểu diễn kết quả thực nghiệm nên tăng kích thước chữ, nhãn trục và chú giải; đồng thời thống nhất cách trình bày giữa các hình có cùng loại kết quả.	NCS tiếp thu ý kiến của Hội đồng. Đối với các hình biểu diễn kết quả thực nghiệm, NCS đã rà soát cách trình bày và bổ sung phần diễn giải trực tiếp trong nội dung nhằm làm rõ ý nghĩa của các trục, đường biểu diễn, chú giải và xu hướng kết quả. Các hình có cùng loại kết quả được rà soát về cách dẫn chiếu và thuật ngữ để bảo đảm tính nhất quán trong toàn luận án.

NCS cần kiểm tra toàn bộ luận án để sửa các lỗi hình thức trình bày	NCS đã kiểm tra cẩn thận toàn bộ luận án và chỉnh sửa các lỗi liên quan đến hình thức/trình bày
---	---

Nghiên cứu sinh chân thành cảm ơn Quý thầy, cô trong Hội đồng đánh giá luận án tiến sĩ cấp Học viện đã góp ý và tạo cơ hội cho NCS hoàn thiện luận án của mình.

Xin trân trọng cảm ơn./.

Hà Nội, ngày 28 tháng 8 năm 2026

TẬP THỂ HƯỚNG DẪN

(Trường hợp có 02 người hướng dẫn xin chữ ký cả 02 người, ký và ghi rõ họ tên)

TS. Trần Đức Nghĩa GS. TS. Trần Đức Tân

CHỦ TỊCH HỘI ĐỒNG



PGS. TS. Nguyễn Việt Anh

NGHIÊN CỨU SINH



Vũ Hoàng Diệu

THƯ KÝ HỘI ĐỒNG



TS. Nguyễn Như Sơn

**XÁC NHẬN CỦA HỌC VIỆN
KHOA HỌC VÀ CÔNG NGHỆ**



Trần Thị Phương Anh

