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 THẾ NGỦ TỪ DỮ LIỆU CẢM BIẾN GIA TỐC BA TRỤC THEO CÁCH TIẾP CẬN KẾT HỢP CÁC MÔ HÌNH HỌC SÂU

TÓM TẮT LUẬN ÁN TIẾN SĨ MÁY TÍNH

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

Hà Nội - Năm 2026

Công trình được hoàn thành tại: Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Người hướng dẫn khoa học:
TS. Trần Đức Nghĩa, Viện Hàn lâm Khoa học và Công nghệ Việt Nam
GS. TS. Trần Đức Tân, Đại học Phenikaa

Phản biện 1: .....

Phản biện 2: .....

Phản biện 3: .....

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi ……. giờ ……, ngày …….. tháng …….. năm ……..

Có thể tìm hiểu luận án tại:
1. Thư viện Học viện Khoa học và Công nghệ
2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tính cấp thiết của luận án

Sự phát triển của cảm biến đeo, cảm biến quán tính và các hệ thống chăm sóc sức khỏe thông minh đã tạo điều kiện cho việc theo dõi giấc ngủ liên tục trong môi trường tại nhà. Trong các cảm biến thường dùng, cảm biến gia tốc ba trục có ưu điểm kích thước nhỏ, tiêu thụ năng lượng thấp, chi phí hợp lý và có khả năng ghi nhận hướng cơ thể mà không gây nhiều bất tiện cho người sử dụng. Vì vậy, tín hiệu gia tốc là nguồn dữ liệu phù hợp cho bài toán nhận dạng tư thế ngủ.

Tư thế ngủ có liên quan tới nhiều vấn đề sức khỏe như ngưng thở khi ngủ do tắc nghẽn, trào ngược dạ dày thực quản, đau cơ xương khớp và loét do tỳ đè. Tuy nhiên, nhiều nghiên cứu trước đây chủ yếu chỉ phân loại ba hoặc bốn tư thế cơ bản, gồm nằm ngửa, nằm sấp, nghiêng trái và nghiêng phải. Cách mô tả thô này chưa đủ để phản ánh các sai lệch góc nhỏ nhưng có thể mang ý nghĩa lâm sàng. Do đó, luận án tập trung vào bài toán phân loại 12 tư thế ngủ tĩnh từ một cảm biến gia tốc ba trục gắn tại vùng bụng.

Một yêu cầu quan trọng khác là đánh giá mô hình theo kịch bản độc lập theo đối tượng. Trong thực tế, hệ thống phải hoạt động với người dùng mới, không xuất hiện trong tập huấn luyện. Vì vậy, nếu cửa sổ tín hiệu của cùng một người vừa nằm trong tập huấn luyện vừa nằm trong tập kiểm tra, kết quả có thể quá lạc quan. Luận án sử dụng cách chia dữ liệu theo cấp độ người tham gia nhằm phản ánh đúng hơn điều kiện triển khai.

2. Mục tiêu nghiên cứu

Mục tiêu tổng quát của luận án là phát triển và đánh giá các mô hình học sâu cho phân loại 12 tư thế ngủ từ dữ liệu cảm biến gia tốc ba trục đeo tại vùng bụng. Các mục tiêu cụ thể gồm: xây dựng bài toán phân loại chuỗi thời gian đa lớp; đề xuất mô hình AnpoNet kết hợp 1D CNN và LSTM với số lượng tham số được kiểm soát; đề xuất mô hình

CGB-Net kết hợp 1D CNN, GRU và BiLSTM nhằm tăng cường năng lực biểu diễn; xây dựng quy trình thực nghiệm độc lập theo đối tượng; phân tích ảnh hưởng của cấu hình cửa sổ, tỷ lệ chồng lấp, độ ổn định và lỗi theo từng lớp tư thế.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu là bài toán phân loại tư thế ngủ của con người từ dữ liệu gia tốc ba trục. Trọng tâm của luận án là nhận dạng 12 tư thế ngủ tĩnh bằng một cảm biến gắn tại vùng bụng, trong điều kiện thu thập dữ liệu có kiểm soát. Các cảm biến khác như con quay hồi chuyển, cảm biến áp lực, cảm biến hô hấp hoặc các hệ đa cảm biến không thuộc phạm vi chính của luận án.

Phạm vi nghiên cứu được giới hạn ở phân tích ngoại tuyến trên dữ liệu đã thu thập và phân đoạn. Luận án chưa đánh giá dữ liệu ngủ tự nhiên qua đêm, tư thế chuyển tiếp, chuyển động không ổn định trong khi ngủ, cũng như chưa xác nhận lâm sàng trên nhóm bệnh nhân. Khả năng triển khai trên thiết bị biên được xem xét thông qua các chỉ số gián tiếp như số lượng tham số, FLOPs, bộ nhớ sử dụng, thời gian suy luận và năng lượng ước tính.

4. Nội dung nghiên cứu

Luận án gồm các nội dung chính: khảo sát các nghiên cứu về theo dõi tư thế ngủ dựa trên cảm biến; xây dựng bộ dữ liệu tư thế ngủ từ cảm biến gia tốc ba trục gắn ở bụng; đề xuất AnpoNet hướng tới mô hình gọn nhẹ; đề xuất CGB-Net nhằm nâng cao độ chính xác; đánh giá các mô hình trong thiết lập độc lập theo đối tượng; tổng hợp ưu điểm, hạn chế và hướng phát triển tiếp theo.

5. Phương pháp nghiên cứu

Luận án kết hợp tổng quan tài liệu, thiết kế dữ liệu thực nghiệm, mô hình hóa bằng học sâu và đánh giá định lượng. Dữ liệu sau khi thu thập được chuẩn hóa, phân đoạn thành các cửa sổ tín hiệu và tổ chức theo cấp độ đối tượng. Các mô hình được huấn luyện trực tiếp trên chuỗi gia tốc, thay vì chỉ dựa vào đặc trưng thủ công. Hiệu năng được đánh

giá bằng accuracy, precision, recall, F1-score, ma trận nhầm lẫn, kiểm tra nhiều seed ngẫu nhiên và kiểm định chéo theo đối tượng.

6. Đóng góp chính

Luận án có hai đóng góp chính: (i) đề xuất AnpoNet cho phân loại 12 tư thế ngủ từ dữ liệu gia tốc ba trục, hướng tới thiết bị tài nguyên hạn chế; (ii) đề xuất CGB-Net nhằm cải thiện độ chính xác và khả năng phân loại cân bằng giữa các lớp. Ngoài ra, luận án đóng góp quy trình đánh giá độc lập theo đối tượng cho bài toán nhận dạng tư thế ngủ chi tiết bằng một cảm biến đeo.

7. Cấu trúc của luận án

Ngoài phần Mở đầu, Kết luận và Kiến nghị, luận án gồm ba chương. Chương 1 trình bày tổng quan về bài toán phân loại tư thế ngủ, các hệ thống dựa trên cảm biến và các phương pháp học máy, học sâu cho dữ liệu chuỗi thời gian. Chương 2 mô tả bộ dữ liệu, quy trình tiền xử lý và đề xuất mô hình AnpoNet với số lượng tham số được kiểm soát. Chương 3 đề xuất CGB-Net nhằm tăng cường khả năng biểu diễn đặc trưng không gian-thời gian và cải thiện hiệu năng phân loại.

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

1.1. Bài toán phân loại tư thế ngủ

Tư thế ngủ là một biến hành vi đơn giản nhưng có khả năng phản ánh nhiều hiện tượng sinh lý trong khi ngủ. Với bệnh nhân ngưng thở khi ngủ do tắc nghẽn, tư thế nằm ngửa thường làm tăng nguy cơ xẹp đường hô hấp trên, làm tăng chỉ số ngưng thở-giảm thở và gây giảm oxy máu nghiêm trọng hơn. Trong bệnh trào ngược dạ dày thực quản, tư thế nằm nghiêng trái có thể làm giảm thời gian tiếp xúc axit thực quản vào ban đêm. Với người nằm lâu hoặc sau phẫu thuật, góc nghiêng thân có thể ảnh hưởng trực tiếp đến phân bố áp lực và nguy cơ loét do tỳ đè.

Vì vậy, theo dõi tư thế ngủ không chỉ nhằm mô tả người dùng nằm ở đâu, mà còn có thể hỗ trợ phân tầng nguy cơ, đánh giá tuân thủ liệu pháp tư thế và cung cấp dữ liệu dài hạn tại nhà. Tuy nhiên, các tư thế trung gian là thách thức lớn vì khác biệt tín hiệu giữa chúng nhỏ hơn nhiều so với khác biệt giữa các tư thế cơ bản.

Trong điều kiện tĩnh, cảm biến gia tốc đo thành phần trọng lực chiếu lên ba trục cảm biến. Hướng của véc-tơ trọng lực liên quan trực tiếp đến hướng cơ thể. Dù vậy, tín hiệu thực tế chịu ảnh hưởng của vị trí gắn cảm biến, độ chặt của dây đeo, chuyển động hô hấp và khác biệt hình thái giữa các cá nhân. Do đó, một mô hình phân loại tốt phải vừa học được đặc trưng định hướng vừa có khả năng tổng quát hóa sang người dùng mới.

1.2. Tư thế độ phân giải cao và tư thế trung gian

Luận án nhấn mạnh rằng tư thế ngủ không chỉ gồm bốn nhóm cơ bản. Nhiều nghiên cứu lâm sàng cho thấy các góc nghiêng như 15 độ, 30 độ hoặc 45 độ có thể tạo ra sự khác biệt về hô hấp, áp lực và cảm giác thoải mái. Điều này đặt ra yêu cầu phân loại tư thế ở độ phân giải cao hơn.

Trong bối cảnh luận án, 12 tư thế được xây dựng bằng cách kết hợp các hướng cơ bản và các biến thể trung gian. Cách tiếp cận này giữ được tính rời rạc để thuận lợi cho bài toán học máy, đồng thời tăng độ chi tiết so với các bài toán ba hoặc bốn lớp. Đây là tiền đề để phát triển các hệ thống theo dõi tư thế ngủ có giá trị thực tiễn hơn.

1.3. Tổng quan các hệ thống phân loại tư thế ngủ

Các hệ thống phân loại tư thế ngủ có thể được chia thành nhiều nhóm theo phương thức cảm biến. Hệ thống dùng camera cung cấp thông tin không gian phong phú, nhưng gặp vấn đề về riêng tư, ánh sáng và che khuất bởi chăn gối. Hệ thống dùng cảm biến áp lực trên đệm ít gây khó chịu nhưng phụ thuộc vào môi trường giường ngủ và khó sử dụng trong các bối cảnh khác nhau. Hệ thống radar hoặc sóng vô tuyến có ưu điểm không tiếp xúc, song thường phức tạp hơn về phần cứng và xử lý tín hiệu.

Thiết bị đeo là hướng tiếp cận phù hợp cho theo dõi dài hạn vì có thể hoạt động ở nhiều môi trường. Trong các vị trí gắn cảm biến, vùng bụng là lựa chọn hợp lý vì gần trung tâm thân người, phản ánh hướng toàn thân và vẫn đảm bảo tính thoải mái khi ngủ. Luận án sử dụng một cảm biến gia tốc ba trục tại vùng bụng nhằm giữ hệ thống đơn giản, chi phí thấp và thuận lợi cho ứng dụng tại nhà.

Từ tổng quan, luận án nhận diện ba khoảng trống: số lớp tư thế còn hạn chế, nhiều nghiên cứu chưa đánh giá độc lập theo đối tượng, và độ phức tạp mô hình chưa được xem xét đầy đủ trong định hướng triển khai trên thiết bị đeo.

1.4. Các phương pháp học máy cho phân loại tư thế ngủ

Các phương pháp học máy truyền thống thường bắt đầu bằng việc trích xuất đặc trưng thủ công từ từng cửa sổ tín hiệu, như trung bình, phương sai, năng lượng, biên độ, góc nghiêng hoặc hệ số tương quan giữa các trục. Sau đó, các đặc trưng này được đưa vào cây quyết định, SVM, Random Forest, XGBoost, MLP hoặc các bộ phân loại tương tự. Cách tiếp cận này đơn giản và dễ giải thích, nhưng có thể chưa đủ

mạnh khi cần phân biệt 12 lớp có ranh giới gần nhau.

Học sâu có khả năng học đặc trưng trực tiếp từ chuỗi thời gian. Mạng 1D CNN phù hợp để trích xuất các mẫu cục bộ và quan hệ giữa các trục gia tốc. LSTM và GRU thích hợp cho mô hình hóa phụ thuộc thời gian, giúp nhận dạng sự ổn định của tư thế trong một cửa sổ. BiLSTM mở rộng LSTM bằng cách khai thác ngữ cảnh theo hai chiều trong chuỗi, có thể hữu ích khi toàn bộ cửa sổ tín hiệu đã sẵn có trong phân tích ngoại tuyến.

Kiến trúc lai CNN-RNN kết hợp ưu điểm của hai nhóm mô hình. CNN trích xuất đặc trưng cục bộ từ tín hiệu gia tốc, còn RNN tổng hợp thông tin theo thời gian. Đây là nền tảng cho hai mô hình đề xuất trong luận án: AnpoNet ở Chương 2 và CGB-Net ở Chương 3.

1.5. Thử nghiệm sơ bộ và các chỉ số đánh giá

Thử nghiệm sơ bộ trong luận án xác nhận rằng phân loại 12 tư thế bằng một cảm biến gia tốc là khả thi, nhưng không đơn giản. Các tư thế gần nhau về góc nghiêng dễ gây nhầm lẫn; vì vậy cần mô hình có khả năng khai thác cả đặc trưng định hướng và sự ổn định thời gian.

Các chỉ số đánh giá gồm accuracy, precision, recall, F1-score và ma trận nhầm lẫn. Accuracy phản ánh tỷ lệ dự đoán đúng tổng thể. Precision cho biết mức độ tin cậy khi mô hình dự đoán một lớp. Recall cho biết khả năng nhận ra đầy đủ các mẫu của một lớp. F1-score là trung bình điều hòa giữa precision và recall, phù hợp khi cần đánh giá cân bằng giữa các lớp. Ma trận nhầm lẫn giúp chỉ ra những cặp tư thế dễ nhầm, đặc biệt là các tư thế trung gian.

1.6. Tóm tắt Chương 1

Chương 1 đã trình bày cơ sở khoa học và kỹ thuật của luận án. Về mặt ứng dụng, tư thế ngủ có ý nghĩa đối với theo dõi sức khỏe và một số vấn đề lâm sàng. Về mặt kỹ thuật, cảm biến gia tốc ba trục là lựa chọn phù hợp cho hệ thống đeo gọn nhẹ. Về mặt mô hình, học sâu, đặc biệt là các kiến trúc lai CNN-RNN, có tiềm năng học đặc trưng trực tiếp từ tín hiệu chuỗi thời gian.

Tuy nhiên, tổng quan cũng cho thấy ba thách thức chính: cần phân loại tư thế chi tiết hơn bốn lớp cơ bản; cần đánh giá mô hình theo thiết lập độc lập theo đối tượng; và cần kiểm soát độ phức tạp để hướng tới thiết bị tài nguyên hạn chế. Các vấn đề này được giải quyết lần lượt trong Chương 2 và Chương 3.

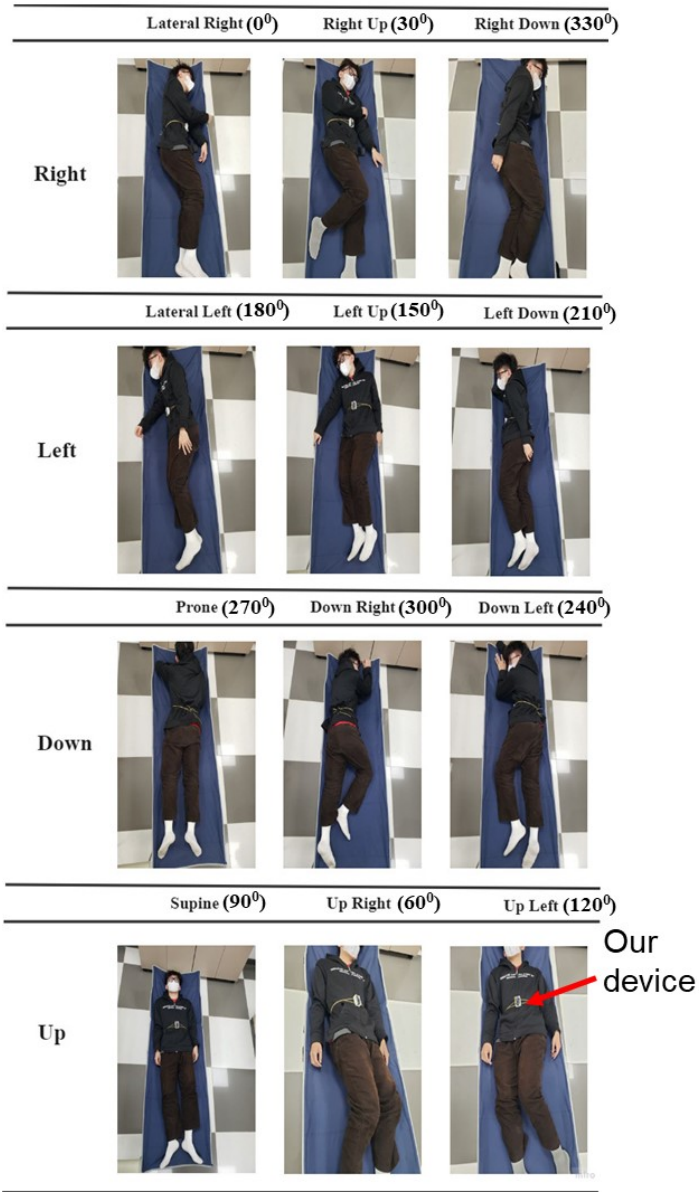
CHƯƠNG 2. ĐỀ XUẤT MÔ HÌNH ANPONET VỚI SỐ LƯỢNG THAM SỐ RÀNG BUỘC CHO BÀI TOÁN PHÂN LOẠI TƯ THẾ NGỦ.

2.1. Phương pháp luận thiết kế bộ dữ liệu

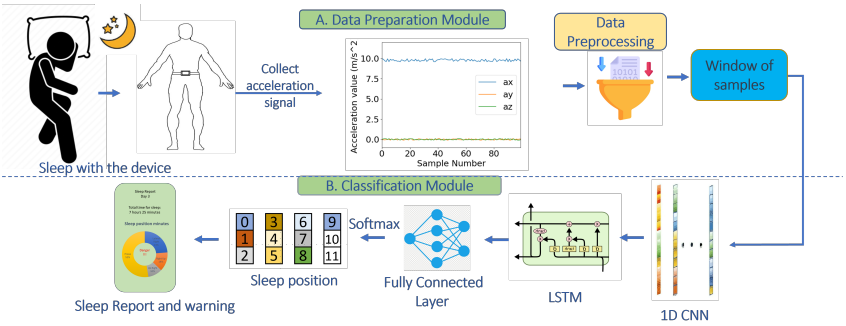
Bộ dữ liệu trong Chương 2 được xây dựng nhằm phục vụ bài toán phân loại 12 tư thế ngủ tĩnh từ một cảm biến gia tốc ba trục gắn tại vùng bụng. Mục tiêu là tạo ra dữ liệu có nhãn đáng tin cậy, có giao thức thu thập thống nhất và có thể chia theo đối tượng để đánh giá khả năng tổng quát hóa.

Mỗi người tham gia thực hiện 12 tư thế ngủ theo một quy trình chuẩn. Mỗi tư thế được duy trì trong khoảng 60 giây, tần số lấy mẫu là 50 Hz, nên mỗi bản ghi tư thế có khoảng 3.000 mẫu theo thời gian. Dữ liệu thô gồm thời gian, gia tốc theo trục x, y, z và nhãn tư thế. Các nhãn được gán ở mức bản ghi vì toàn bộ mẫu trong một bản ghi tương ứng với một tư thế được hướng dẫn.

Vị trí gắn cảm biến được chọn tại vùng bụng. Đây là vị trí có ý nghĩa vì phản ánh hướng thân người và có thể đeo trong khi ngủ. Việc chỉ sử dụng một cảm biến cũng làm giảm chi phí, giảm độ phức tạp phần cứng và thuận lợi cho triển khai thực tế.



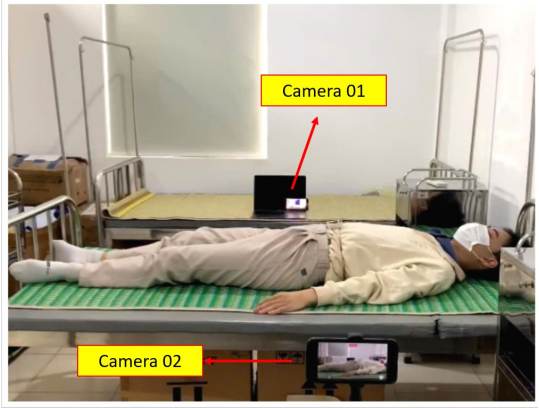
Hình 1: Các hình ảnh minh họa của 12 tư thế ngủ.



Hình 2: Hệ thống hai mô-đun cho phân loại 12 Tư thế Ngủ (12-SPSC)

2.1.1. Giao thức thu thập, tiền xử lý và tổ chức dữ liệu

Quy trình thu thập dữ liệu kiểm soát thứ tự tư thế, vị trí cảm biến, thời lượng ghi và tần số lấy mẫu. Sau khi thu thập, dữ liệu được chuẩn hóa và phân đoạn thành các cửa sổ trượt. Cấu hình điển hình sử dụng cửa sổ 100 mẫu và tỷ lệ chồng lấp 0,4. Với tần số 50 Hz, cửa sổ 100 mẫu tương ứng 2 giây.



Hình 3: Quá trình gán nhãn tư thế

Điểm quan trọng là dữ liệu được tổ chức theo đối tượng trước khi chia tập. Nếu chia ngẫu nhiên theo cửa sổ sau khi phân đoạn, cửa sổ của cùng một người có thể xuất hiện trong cả tập huấn luyện và tập

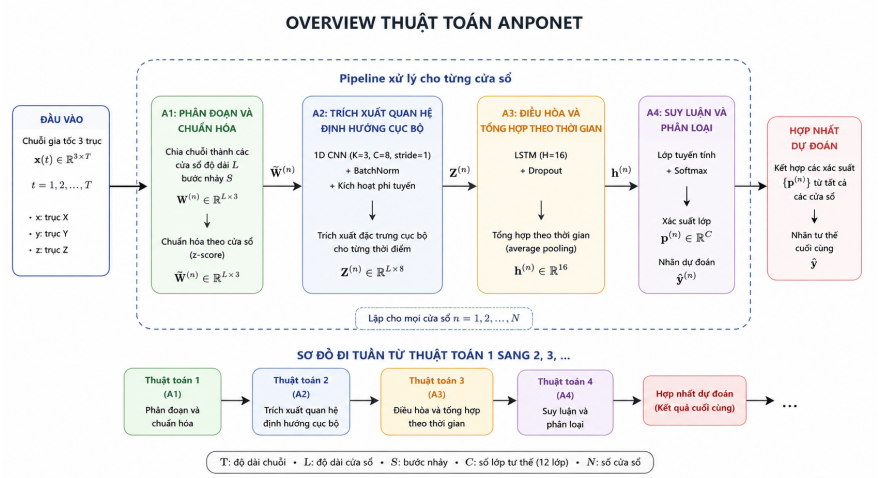
kiểm tra, gây rò rỉ thông tin. Luận án tránh vấn đề này bằng cách đảm bảo dữ liệu của người kiểm tra không xuất hiện trong huấn luyện.

Bộ dữ liệu A được dùng để phát triển và đánh giá AnpoNet. Các đặc điểm hình thể của người tham gia được ghi nhận nhằm mô tả sự khác biệt về chiều cao, cân nặng và vòng bụng, là các yếu tố có thể ảnh hưởng tới vị trí cảm biến và tín hiệu.

2.2. Thiết kế mô hình AnpoNet hướng tới thiết bị biên

Phân loại tư thế ngủ khác với nhận dạng hoạt động mạnh của con người. Thông tin quan trọng chủ yếu là hướng tĩnh của cơ thể, được phản ánh qua véc-tơ trọng lực trong ba trục gia tốc. Vì vậy, một mô hình quá lớn có thể không cần thiết và khó triển khai trên thiết bị đeo. AnpoNet được thiết kế theo hướng gọn nhẹ nhưng vẫn khai thác được đặc trưng cục bộ và phụ thuộc thời gian.

AnpoNet gồm các bước chính: phân đoạn và chuẩn hóa cửa sổ tín hiệu; trích xuất quan hệ định hướng cục bộ bằng 1D CNN; điều hòa và tổng hợp theo thời gian bằng LSTM; suy luận và phân loại qua lớp tuyến tính và softmax. Batch Normalization được sử dụng để ổn định phân bố kích hoạt và hỗ trợ quá trình huấn luyện.



Hình 4: Tổng quan thuật toán AnpoNet với các mô-đun xử lý theo từng cửa sổ tín hiệu.

Hình minh họa cho thấy AnpoNet không phải là một mạng sâu lớn, mà là một chuỗi mô-đun được lựa chọn phù hợp với đặc tính của tín hiệu gia tốc. CNN học các mẫu cục bộ giữa ba trục, LSTM kiểm tra tính ổn định theo thời gian và lớp softmax đưa ra xác suất cho 12 lớp tư thế.

2.2.1. Mô hình hóa bài toán và hàm mất mát

Một cửa sổ gia tốc được ký hiệu là $X \in \mathbb{R}^{T\times 3}$, trong đó T là số mẫu thời gian và ba cột tương ứng với ba trục gia tốc. Mô hình học ánh xạ $f(X;\theta)$ từ cửa sổ đầu vào đến một trong $C = 12$ nhãn tư thế. Lớp đầu ra sử dụng softmax:

$$\hat{y}_c = \frac{\exp(z_c)}{\sum_{j=1}^{C} \exp(z_j)}, \quad c = 1, \ldots, C.$$

Nhãn dự đoán là lớp có xác suất lớn nhất. Mục tiêu huấn luyện là tìm tham số θ để mô hình dự đoán đúng nhãn trên tập huấn luyện và vẫn tổng quát tốt cho đối tượng chưa thấy.

Hàm mất mát sử dụng categorical cross-entropy:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{12} y_{i,c}\log(\hat{y}_{i,c}).$$

Bộ tối ưu Adam được dùng với tốc độ học ban đầu 0,001, kích thước mini-batch 512 và tối đa 200 epoch. Nhiều seed ngẫu nhiên được sử dụng để kiểm tra độ ổn định của mô hình.

2.2.2. Phân tích độ phức tạp

AnpoNet được đánh giá không chỉ bằng độ chính xác mà còn bằng số tham số, FLOPs, bộ nhớ, thời gian suy luận và năng lượng ước tính. Các chỉ số này phù hợp với mục tiêu hướng tới thiết bị biên, nơi tài nguyên tính toán và năng lượng bị giới hạn. Một mô hình có độ chính xác cao nhưng quá chậm hoặc quá lớn có thể không phù hợp cho thiết bị đeo.

2.3. Thực nghiệm và các mô hình so sánh

AnpoNet được so sánh với nhiều mô hình cơ sở. Nhóm học máy truyền thống gồm Decision Tree, XGBoost, MLP và ExtraTrees. Nhóm học sâu gồm 1D CNN, 1D CNN + GRU, các biến thể LSTM và phiên bản AnpoNet không có Batch Normalization. Việc so sánh này giúp xác định liệu cải thiện đến từ kiến trúc đề xuất hay chỉ từ số lượng tham số.

Các thí nghiệm sử dụng cách chia dữ liệu độc lập theo đối tượng. Luận án khảo sát ảnh hưởng của độ dài chuỗi từ 10 đến 140 mẫu và tỷ lệ chồng lấp từ 10% đến 70%. Accuracy và F1-score là hai chỉ số chính; precision, recall và ma trận nhầm lẫn được dùng cho phân tích sâu hơn.

2.4. Kết quả thực nghiệm và thảo luận

Kết quả chính của Chương 2 cho thấy AnpoNet đạt hiệu năng cao nhất trên tập kiểm tra, với độ chính xác 94,67% và F1-score 92,94%. Mô hình 1D CNN + GRU đạt 90,81% accuracy nhưng F1-score chỉ 72,64%, cho thấy accuracy cao chưa chắc phản ánh khả năng phân loại cân bằng giữa các lớp.

Bảng 2.2. So sánh hiệu suất của các mô hình trên tập kiểm tra

Mô hình	Accuracy (%)	F1-score (%)
Decision Tree	81,82	78,67
XGBoost	85,34	82,94
MLP	81,25	72,96
1D CNN	87,08	78,19
1D CNN + GRU	90,81	72,64
ExtraTrees baseline	86,50	84,68
LSTM baseline	68,84	65,02
LSTM (CT4)	87,44	72,30
AnpoNet	**94,67**	**92,94**
AnpoNet không BN	71,11	54,88

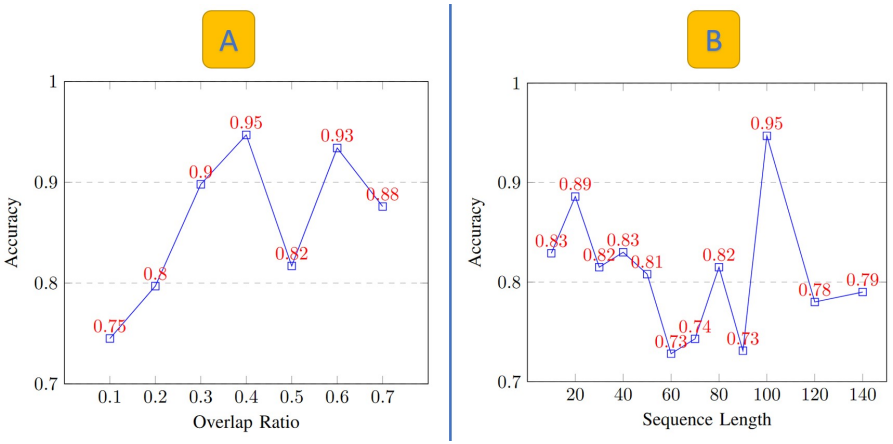
Kết quả loại bỏ Batch Normalization làm giảm mạnh hiệu năng, từ 94,67% xuống 71,11% accuracy. Điều này cho thấy Batch Normalization đóng vai trò quan trọng trong ổn định huấn luyện và tăng khả năng tổng quát hóa.

2.4.1. Độ ổn định theo seed và kiểm định chéo

Thử nghiệm với nhiều seed ngẫu nhiên cho thấy AnpoNet duy trì accuracy và F1-score cao, chứng tỏ kết quả không phải là một lần chạy may mắn. Kiểm định chéo theo đối tượng tiếp tục xác nhận mô hình học được các đặc trưng có thể chuyển giao giữa các người tham gia.

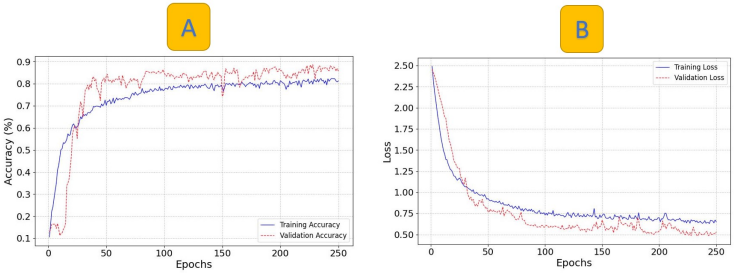
2.4.2. Ảnh hưởng của độ dài cửa sổ và tỷ lệ chồng lấp

Độ dài cửa sổ ảnh hưởng đến cân bằng giữa độ chính xác và độ trễ. Cửa sổ quá ngắn có thể không đủ thông tin thời gian; cửa sổ quá dài làm tăng độ trễ và có thể chứa thông tin dư thừa. Luận án cho thấy cửa sổ 100 mẫu, tương ứng 2 giây, là cấu hình hợp lý. Tỷ lệ chồng lấp 0,4 giúp tăng tính liên tục của dữ liệu mà không tạo dư thừa quá lớn.



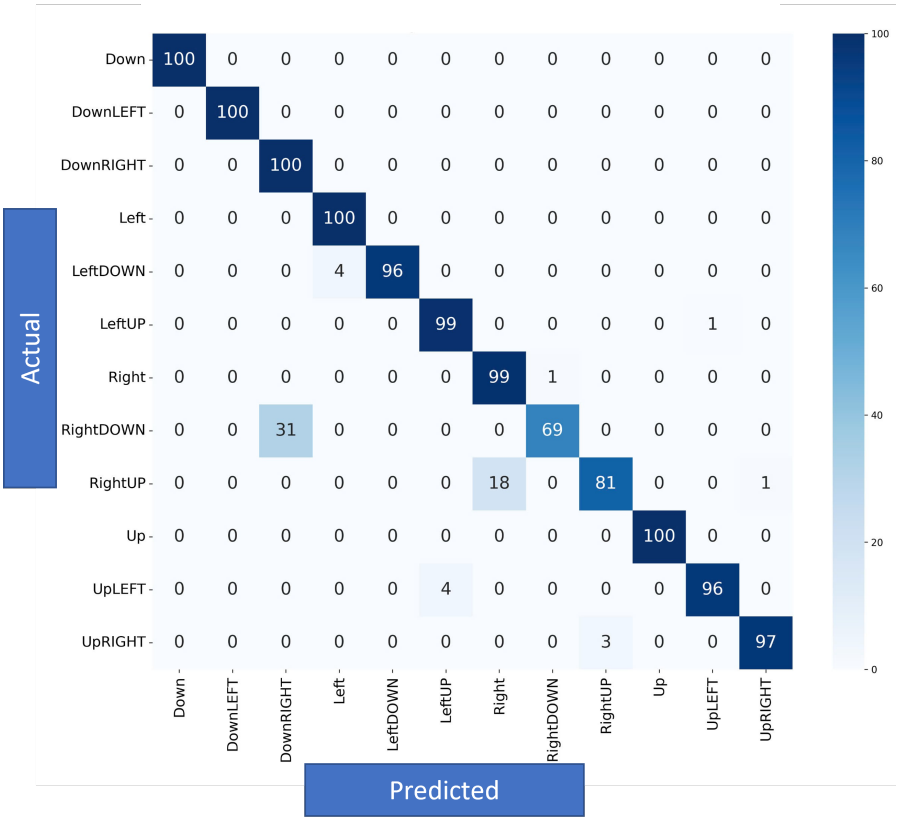
Hình 5: A. Độ chính xác của AnpoNet với tỷ lệ chồng lấp thay đổi; B. Độ chính xác của AnpoNet với độ dài chuỗi thay đổi

2.4.3. Phân tích quá trình huấn luyện



Hình 6: Đường cong huấn luyện minh họa sự thay đổi của accuracy và loss trong quá trình học.

Các đường cong huấn luyện cho thấy mô hình học hiệu quả: accuracy tăng trong các epoch đầu, loss giảm dần và tập xác thực đạt hiệu năng tốt trước khi huấn luyện quá lâu. Điều này gợi ý rằng early stopping có thể hữu ích để tránh quá khớp, đặc biệt khi kích thước dữ liệu còn hạn chế.



Hình 7: Ma trận nhầm lẫn của mô hình AnpoNet đề xuất

2.4.4. Độ phức tạp mô hình và hiệu quả tính toán

AnpoNet là mô hình nhỏ nhất trong nhóm so sánh, chỉ có 1.916 tham số và yêu cầu khoảng 34 kB bộ nhớ. Thời gian suy luận trên môi trường CPU-only là 0,081 giây, nhanh hơn các mô hình so sánh. Đây là kết quả quan trọng vì mục tiêu của Chương 2 không chỉ là đạt độ chính xác cao, mà còn hướng tới khả năng triển khai trên thiết bị có tài nguyên hạn chế.

Bảng 2.3. Độ phức tạp mô hình và hiệu quả tính toán

Mô hình	Tham số	FLOPs (M)	Thời gian (s)	Bộ nhớ (kB)	Năng lượng (mJ)
MLP	22.124	0,0433	0,304	116	0,304
1D CNN	25.436	0,1410	0,204	135	0,989
1D CNN + GRU	16.700	1,4633	1,481	134	10,257
LSTM baseline	3.212	0,2806	0,404	31	1,968
LSTM (CT4)	27.964	0,2548	0,897	154	1,787
AnpoNet	**1.916**	**0,1603**	**0,081**	**34**	**1,122**

Khi so sánh với một LSTM nhỏ có số tham số tương đương, AnpoNet vẫn vượt trội về accuracy và F1-score. Điều này chứng tỏ ưu thế của mô hình không chỉ đến từ kích thước, mà từ cách kết hợp CNN, LSTM và chuẩn hóa phù hợp với đặc tính tín hiệu gia tốc.

2.4.5. Độ bền vững và hiệu suất theo từng lớp

Để đánh giá độ bền vững, luận án thêm nhiễu trắng vào tín hiệu. AnpoNet khi có nhiễu đạt 92,44% accuracy và 80,20% F1-score, so với 94,67% và 92,94% trong điều kiện không nhiễu. Accuracy giảm vừa phải, cho thấy mô hình vẫn nhận dạng được nhiều tư thế khi tín hiệu bị nhiễu.

Phân tích theo từng lớp cho thấy AnpoNet đạt precision và recall rất cao ở nhiều tư thế. Một số lỗi tập trung ở các tư thế trung gian hoặc tư thế có góc gần nhau, phản ánh bản chất khó của bài toán phân loại chi tiết.

Bảng 2.4. Precision và recall theo từng lớp tư thế

Lớp	Precision	Recall	Lớp	Precision	Recall
D	1,00	1,00	LD	1,00	0,89
DL	1,00	1,00	LU	0,98	1,00
DR	0,86	1,00	LR (R)	0,91	1,00
LL (L)	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 2 chứng minh rằng một kiến trúc lai gọn nhẹ có thể phân loại 12 tư thế ngủ với hiệu năng cao trong thiết lập độc lập theo đối tượng. AnpoNet đạt sự cân bằng giữa độ chính xác, F1-score, số tham

số thấp, thời gian suy luận ngắn và bộ nhớ nhỏ. Mô hình phù hợp với định hướng giám sát tư thế ngủ bằng thiết bị đeo tài nguyên hạn chế.

CHƯƠNG 3. ĐỀ XUẤT MÔ HÌNH CGB-NET NHẰM TĂNG CƯỜNG ĐỘ CHÍNH XÁC TRONG PHÂN LOẠI TƯ THẾ NGỦ

3.1. Bài toán và vấn đề nghiên cứu

Chương 3 tiếp tục xử lý bài toán phân loại 12 tư thế ngủ nhưng tập trung vào việc nâng cao độ chính xác và khả năng tổng quát hóa khi dữ liệu được mở rộng. Dù AnpoNet gọn nhẹ và hiệu quả, một số cặp tư thế vẫn khó phân biệt do chênh lệch góc nhỏ và tín hiệu gia tốc gần nhau. Khi số đối tượng tăng, khác biệt hình thái cơ thể, cách nằm và vị trí cảm biến làm bài toán khó hơn.

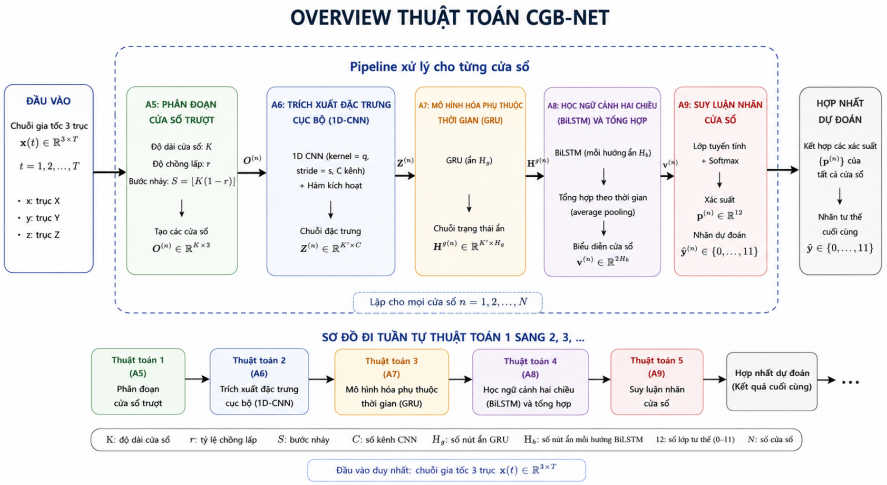
Vấn đề nghiên cứu là thiết kế một kiến trúc có năng lực biểu diễn mạnh hơn nhưng vẫn giữ độ phức tạp ở mức hợp lý. Mô hình cần học đặc trưng cục bộ, quan hệ tuần tự và ngữ cảnh hai chiều trong từng cửa sổ tín hiệu. Do đó, Chương 3 đề xuất CGB-Net, kết hợp 1D CNN, GRU và BiLSTM.

3.2. Động cơ của mô hình đề xuất

1D CNN có khả năng trích xuất đặc trưng cục bộ và quan hệ giữa các trục gia tốc. GRU mô hình hóa phụ thuộc theo thời gian với chi phí tham số thấp hơn LSTM trong nhiều trường hợp. BiLSTM khai thác thông tin theo cả chiều tiến và chiều lùi trong cửa sổ, giúp tăng khả năng phân biệt các tư thế có đặc trưng gần nhau. CGB-Net không nhằm là mô hình nhỏ nhất, mà tìm kiếm cân bằng giữa độ chính xác, F1-score và độ phức tạp.

3.3. Khung thuật toán CGB-Net

CGB-Net nhận cửa sổ gia tốc ba trục làm đầu vào. Tầng đầu sử dụng tích chập một chiều để học đặc trưng không gian-thời gian cục bộ. Đầu ra của CNN được đưa vào GRU nhằm tổng hợp quan hệ tuần tự. Sau đó, BiLSTM khai thác ngữ cảnh hai chiều để làm giàu biểu diễn



Hình 8: Tổng thể luồng khung thuật toán CGB-Net

trước khi lớp phân loại softmax dự đoán nhãn tư thế.

3.4. Thực nghiệm

Các mô hình so sánh trong Chương 3 gồm MLP, 1D CNN, 1D CNN + BiLSTM, 1D CNN + GRU, mô hình LSTM của Rawan và cộng sự, AnpoNet và CGB-Net. Nhóm mô hình này giúp đánh giá giá trị của từng thành phần trong kiến trúc. So sánh với 1D CNN cho thấy vai trò của ngữ cảnh thời gian; so sánh với CNN-GRU và CNN-BiLSTM cho thấy vai trò của việc kết hợp hai cơ chế hồi quy; so sánh với AnpoNet cho thấy lợi ích của tăng năng lực biểu diễn trong dữ liệu mở rộng.

Quy trình huấn luyện vẫn tuân theo nguyên tắc độc lập theo đối tượng. Accuracy và F1-score là chỉ số chính; ma trận nhầm lẫn, đường cong huấn luyện, đánh giá nhiều seed, độ dài cửa sổ và tỷ lệ chồng lấp được phân tích để hiểu sâu hơn hành vi của mô hình.

3.5. Kết quả thực nghiệm và thảo luận

CGB-Net đạt hiệu năng tốt nhất trong hầu hết các thiết lập. Bốn thiết lập thí nghiệm cho thấy CGB-Net đạt accuracy lần lượt 89,34%, 84,39%, 92,70% và 83,95%; F1-score lần lượt 86,24%, 79,41%, 89,71%

và 78,17%. Mức tăng F1-score so với AnpoNet đặc biệt có ý nghĩa vì nó cho thấy mô hình phân loại cân bằng hơn giữa các lớp khó.

Bảng 3.1. Accuracy (%) của các mô hình theo bốn thiết lập thí nghiệm

Mô hình	TH1	TH2	TH3	TH4
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 et al. 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. F1-score (%) của các mô hình theo bốn thiết lập thí nghiệm

Mô hình	TH1	TH2	TH3	TH4
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 et al. 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**

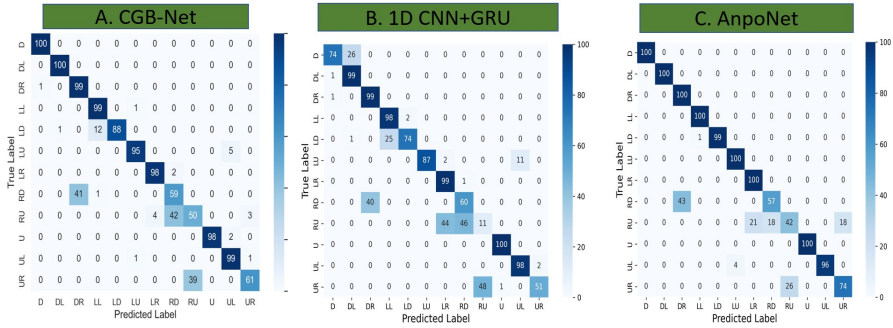
3.5.1. Hiệu năng trung bình và chi phí tính toán

CGB-Net đạt accuracy trung bình 87,60% và F1-score trung bình 83,38%. Mặc dù có độ phức tạp cao hơn AnpoNet, mô hình vẫn chỉ có 9.532 tham số và bộ nhớ khoảng 88 kB, nhỏ hơn nhiều so với nhiều kiến trúc học sâu lớn. Điều này cho thấy việc tăng năng lực biểu diễn được thực hiện ở mức vừa phải và có kiểm soát.

Bảng 3.3. So sánh hiệu suất trung bình và chi phí tính toán

Mô hình	Accuracy TB (%)	F1 TB (%)	Tham số	Bộ nhớ (kB)
MLP	82,38	73,61	22.124	116
1D CNN	84,24	78,34	25.436	135
1D CNN + GRU	82,79	76,20	16.700	134
AnpoNet	85,06	76,72	1.916	34
CGB-Net	**87,60**	**83,38**	**9.532**	**88**

3.5.2. Đánh giá với seed ngẫu nhiên và ma trận nhầm lẫn



Hình 9: Ma trận nhầm lẫn (Confusion Matrix) của CGB-Net, 1D CNN + GRU và AnpoNet

Thử nghiệm nhiều seed ngẫu nhiên cho thấy CGB-Net duy trì hiệu năng ổn định. Điều này quan trọng vì dữ liệu theo đối tượng có thể gây biến động lớn giữa các lần huấn luyện. Sự ổn định cho thấy kiến trúc đề xuất không chỉ tốt trong một cấu hình ngẫu nhiên cụ thể.

Ma trận nhầm lẫn được dùng để so sánh CGB-Net, 1D CNN + GRU và AnpoNet. Kết quả cho thấy CGB-Net giảm một số nhầm lẫn giữa các tư thế gần nhau. Điều này phù hợp với giả thuyết rằng kết hợp CNN, GRU và BiLSTM giúp mô hình tận dụng tốt hơn ngữ cảnh không gian-thời gian.

3.5.3. Quá trình huấn luyện

Các đường cong loss và accuracy của CGB-Net cho thấy mô hình học ổn định, hội tụ sau một số epoch và có khả năng tránh dao động lớn. Việc kết hợp các tầng hồi quy giúp mô hình học tốt hơn các phụ thuộc trong cửa sổ nhưng vẫn cần kiểm soát quá khớp bằng lựa chọn số epoch và tham số huấn luyện phù hợp.

3.5.4. Ảnh hưởng của độ dài chuỗi và tỷ lệ chồng lấp

Tương tự Chương 2, độ dài cửa sổ ảnh hưởng rõ tới hiệu năng của CGB-Net. Trong trường hợp phân tích, cửa sổ 2 giây cho kết quả tốt nhất với accuracy 90,77% và F1-score 76,12%. Các cửa sổ ngắn hơn

như 0,5 giây hoặc 1 giây không cung cấp đủ ngữ cảnh thời gian, trong khi cửa sổ quá dài có thể tăng độ trễ và thông tin dư thừa.

Tỷ lệ chồng lấp cũng ảnh hưởng đến sự liên tục giữa các cửa sổ. Tỷ lệ 0,4 cho kết quả tốt trong phân tích, cân bằng giữa tăng số mẫu huấn luyện, giữ tính liên tục và tránh dư thừa quá mức. Khi triển khai thực tế, lựa chọn cửa sổ và chồng lấp cần cân nhắc giữa độ chính xác, độ trễ, chi phí tính toán và mức tiêu thụ năng lượng.

3.6. Kết luận Chương 3

Chương 3 đã đề xuất CGB-Net cho bài toán phân loại 12 tư thế ngủ. Mô hình kết hợp 1D CNN, GRU và BiLSTM để học đặc trưng không gian-thời gian từ dữ liệu gia tốc đeo tại vùng bụng. Kết quả thực nghiệm cho thấy CGB-Net cải thiện accuracy và đặc biệt là F1-score trong nhiều thiết lập đánh giá độc lập theo đối tượng.

So với AnpoNet, CGB-Net sử dụng nhiều tham số hơn nhưng vẫn ở mức vừa phải. Vì vậy, hai mô hình có vai trò bổ sung: AnpoNet phù hợp khi ưu tiên độ gọn nhẹ, còn CGB-Net phù hợp khi ưu tiên độ chính xác và sự cân bằng giữa các lớp. Kết quả của Chương 3 củng cố tính khả thi của việc phân loại tư thế ngủ chi tiết bằng một cảm biến gia tốc ba trục duy nhất.

KẾT LUẬN VÀ KIẾN NGHỊ

Luận án đã tập trung nghiên cứu bài toán phân loại tư thế ngủ chi tiết từ dữ liệu cảm biến gia tốc ba trục đeo tại vùng bụng. Khác với nhiều nghiên cứu trước đây chỉ phân loại một số tư thế ngủ cơ bản, luận án hướng tới bài toán phân loại 12 tư thế ngủ, bao gồm các tư thế cơ bản và các tư thế trung gian có sự khác biệt về góc nghiêng cơ thể. Bài toán này có ý nghĩa trong các hệ thống theo dõi giấc ngủ sử dụng thiết bị đeo, đặc biệt trong các ứng dụng cần nhận dạng tư thế ngủ ở mức độ chi tiết hơn nhằm hỗ trợ theo dõi sức khỏe và đánh giá hành vi ngủ của người dùng. Các thí nghiệm trong luận án được thiết kế theo hướng đánh giá độc lập theo đối tượng, trong đó dữ liệu của đối tượng kiểm tra không xuất hiện trong quá trình huấn luyện, nhằm phản ánh sát hơn điều kiện triển khai thực tế.

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 và mức tiêu thụ năng lượng khoảng 1,122 mJ 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, chưa phản ánh đầy đủ các chuyển động tự nhiên, tư thế chuyển tiếp và biến thiên phức tạp trong giấc ngủ qua đêm. Ngoài ra, khả năng triển khai trên thiết bị biên mới được đánh giá thông qua các chỉ số gián tiếp như số lượng tham số, chi phí tính toán, bộ nhớ và thời gian suy luận, chưa được kiểm chứng hoàn chỉnh trên phần cứng nhúng hoặc thiết bị đeo công suất thấp trong môi trường thực tế.

Trong thời gian tới, các hướng nghiên cứu có thể tiếp tục phát triển bao gồm mở rộng bộ dữ liệu với số lượng đối tượng đa dạng hơn, 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. Bên cạnh đó, các mô hình đề xuất có thể được tối ưu thêm bằng các kỹ thuật như lượng tử hóa, cắt tỉa tham số hoặc nén mô hình, sau đó triển khai và kiểm chứng trực tiếp trên thiết bị đeo công suất thấp nhằm đánh giá đầy đủ hơn khả năng ứng dụng trong thực tế.

DANH MỤC CÁC BÀI BÁO ĐÃ XUẤT BẢN
LIÊN QUAN ĐẾN LUẬN ÁN

[CT1] Hoang-Dieu Vu, Duc-Nghia Tran, Huy-Hieu Pham, Dinh-Dat Pham, Khanh Ly Can, To-Hieu Dao, Duc-Tan Tran. Human sleep position classification using a lightweight model and acceleration data. Sleep and Breathing, Volume 29, 2025, Article 95. (SCI/ISI indexed, Scopus Q1).

[CT2] Hoang-Dieu Vu, Duc-Nghia Tran, Quang-Tu Pham, Ngoc-Linh Nguyen, 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 Q2).

[CT3] Hoang-Dieu Vu, Duc-Nghia Tran, Pham Quang Huy, Duc-Tan Tran. HyPoNet: Fine-Grained Sleep Posture Recognition from a Single Abdominal Accelerometer. REV Journal on Electronics and Communications, vol. 15, no. 3, 2025

[CT4] Hoang-Dieu Vu, Duc-Nghia Tran, Khanh-Ly Can, To-Hieu Dao, Dinh-Dat Pham, Duc-Tan Tran, "Enhancing sleep postures classification by incorporating acceleration sensor and LSTM model," in 2023 IEEE Statistical Signal Processing Workshop (SSP), 2023, pp. 661–665, doi: 10.1109/SSP53291.2023.10208083. (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. (SCI/ISI indexed, Scopus Q1).