

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



DƯƠNG THỊ MAI THƯƠNG

**NGHIÊN CỨU MỘT SỐ PHƯƠNG PHÁP HỌC SÂU
TRONG BÀI TOÁN TƯƠNG TÁC NÃO MÁY DỰA TRÊN
TÍN HIỆU ĐIỆN NÃO**

TÓM TẮT LUẬN ÁN TIẾN SĨ KHOA HỌC MÁY TÍNH

Chuyên ngành: Khoa học máy tính
Mã số: 9.48.01.01

Thái Nguyên – Năm 2025

ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG



DƯƠNG THỊ MAI THƯƠNG

**NGHIÊN CỨU MỘT SỐ PHƯƠNG PHÁP HỌC SÂU
TRONG BÀI TOÁN TƯƠNG TÁC NÃO MÁY DỰA TRÊN
TÍN HIỆU ĐIỆN NÃO**

TÓM TẮT LUẬN ÁN TIẾN SĨ KHOA HỌC MÁY TÍNH

Chuyên ngành: Khoa học máy tính
Mã số: 9.48.01.01

NGƯỜI HƯỚNG DẪN KHOA HỌC

PGS.TS. Phùng Trung Nghĩa

Thái Nguyên – Năm 2025

MỞ ĐẦU

1. Tính cấp thiết của luận án

Ngày nay, với sự tiến bộ của công nghệ, rất nhiều hệ thống máy móc thông minh được ra đời với mục tiêu hỗ trợ con người nâng cao chất lượng cuộc sống. Điều này góp phần hình thành nên một lĩnh vực nghiên cứu quan trọng, đó là tương tác giữa người và máy (Human-Computer Interaction – HCI) [1], [2].

HCI là một lĩnh vực nghiên cứu liên ngành, đề cập đến việc trao đổi thông tin giữa con người và máy tính, bao gồm cả việc máy tính cung cấp thông tin cho con người thông qua các thiết bị đầu ra và việc con người nhập thông tin liên quan vào máy tính qua các thiết bị đầu vào [2]. Nhờ những đóng góp to lớn của trí tuệ nhân tạo (Artificial Intelligence - AI), rất nhiều nghiên cứu đã kết hợp được giọng nói, hình ảnh, văn bản và các loại thông tin khác để thúc đẩy sự nâng cấp của HCI, làm cho HCI được ứng dụng thành công trong rất nhiều lĩnh vực lĩnh vực như chăm sóc sức khỏe, giáo dục và công nghiệp...[1].

Trong những năm gần đây, với sự phát triển của nhiều thiết bị điện tử khác nhau, con người đã dành nhiều thời gian hơn cho việc sử dụng mạng xã hội, chơi game trực tuyến, mua sắm trực tuyến và sử dụng các sản phẩm điện tử khác. Tuy nhiên, hầu hết các hệ thống HCI vẫn chưa có khả năng xử lý và hiểu được dữ liệu về cảm xúc, và thiếu tính thông minh cảm xúc. Chúng không thể nhận biết được cảm xúc của con người và sử dụng dữ liệu cảm xúc đó để ra quyết định và hành động. Trong khi đó, việc giải quyết sự thiếu hụt mối quan hệ giữa con người và máy móc là rất quan trọng. Bất kỳ hệ thống HCI nào bỏ qua trạng thái cảm xúc của con người đều sẽ không thể phản ứng phù hợp với những cảm xúc đó. Để giải quyết vấn đề này, đã có rất nhiều nhà khoa học trên thế giới quan tâm đến việc giải quyết bài toán nhận dạng cảm xúc (Emotion Recognition - ER) khi xây dựng các hệ thống HCI.

Việc được trang bị thêm khả năng nhận dạng cảm xúc một cách chính xác có thể giúp các hệ thống HCI trở nên thông minh hơn, ra quyết định chính xác hơn, phục vụ tốt hơn các nhu cầu của con người [3], [4]

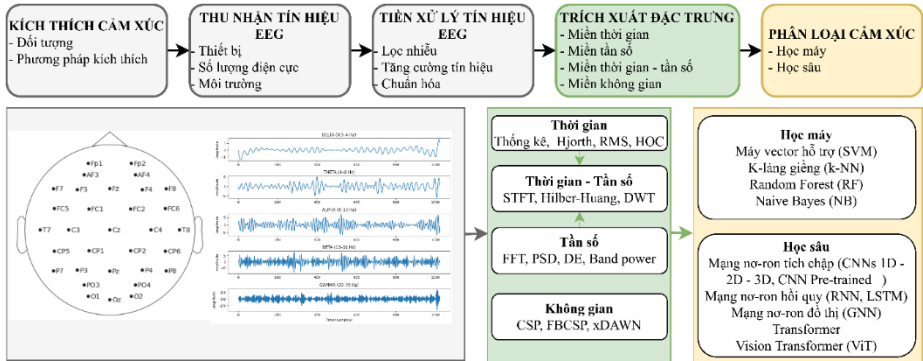
Cũng như con người, các hệ thống HCI có thể nhận dạng được cảm xúc thông qua các tín hiệu phi sinh lý, ví dụ như tiếng nói, nét mặt và tư thế cơ thể [5], [6]. Tuy nhiên, tính chính xác của những phương pháp này có thể bị ảnh hưởng bởi nhiều yếu tố, từ trạng thái tâm trí, giới tính, trình độ học vấn, độ tuổi và khả năng che giấu cảm xúc. Do các yếu tố này thường mang yếu tố chủ quan, việc đánh giá cảm xúc của người đó có thể trở nên khó khăn, thiếu chính xác.

Bên cạnh đó, một số nghiên cứu đã cho thấy các tín hiệu sinh lý và biểu hiện cảm xúc có liên hệ chặt chẽ với nhau. Do đó, các phương pháp sử dụng tín hiệu sinh lý như nhịp tim, trở kháng của da, hô hấp và điện não đồ (Electroencephalogram – EEG) đã được đề xuất để xác định các trạng thái cảm xúc [7]. Trong đó, việc nghiên cứu các cơ chế nhận thức cảm xúc và nhận biết trạng thái cảm xúc bằng cách sử dụng tín hiệu EEG đặc biệt quan trọng, nhất là với những người không thể biểu đạt cảm xúc thông qua lời nói tự nhiên, nét mặt hoặc tư thế cơ thể.

EEG có thể ghi lại các tín hiệu điện được tạo ra bởi các hoạt động bên trong của não bộ, đưa ra phép đo trực tiếp và không xâm lấn về chức năng não. Tín hiệu này phản ánh các quá trình nhận thức và tình cảm liên quan đến các trạng thái cảm xúc khác nhau. Điều này làm cho việc nhận dạng cảm xúc dựa trên tín hiệu điện não (EEG based Emotion Recognition – EEG_ER) trở thành một con đường đầy hứa hẹn để phân tích cảm xúc một cách chính xác hơn [8].

Hình 0.1 mô tả một hệ thống EEG_ER điển hình. Hệ thống này bao gồm các khâu kích thích cảm xúc, thu thập tín hiệu EEG, tiền xử lý tín hiệu EEG, trích xuất các đặc trưng EEG liên quan đến cảm xúc và huấn luyện mô hình nhận dạng cảm xúc [9]. Qua đó có thể thấy rằng, hiệu quả hoạt động của một mô hình EEG_ER phụ thuộc vào rất nhiều

yếu tố như thiết bị thu thập tín hiệu EEG, phương pháp trích xuất đặc trưng đặc trưng, kỹ thuật phân lớp cảm xúc.



Hình 0.1. Các bước thực hiện của bài toán nhận dạng cảm xúc

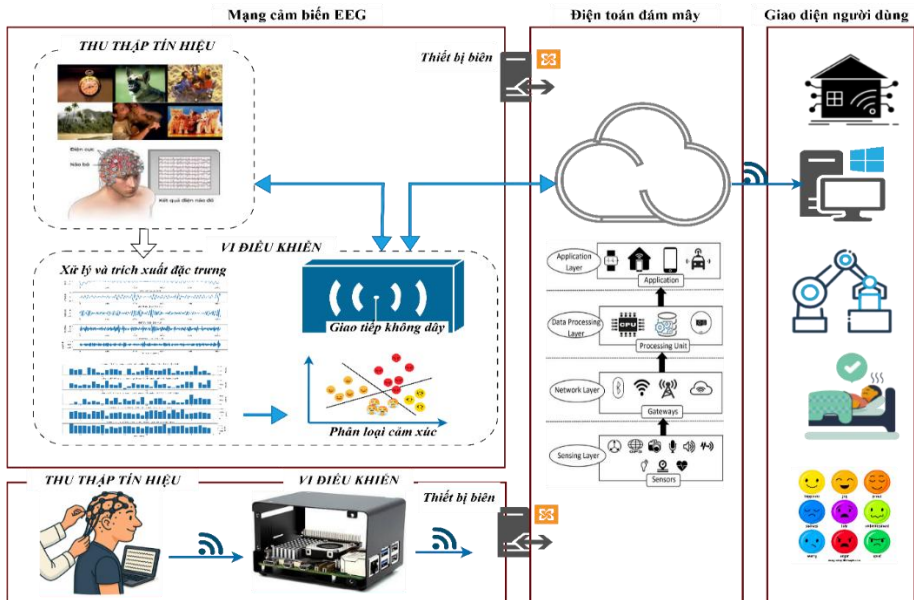
Về thiết bị thu thập tín hiệu EEG, trên thực tế có thể phân ra hai loại thiết bị. Thiết bị có số lượng điện cực EEG lớn (Biosemi ActiveTwo, ESI NeuroScan System) và thiết bị có số lượng điện cực EEG nhỏ (Emotiv EPOC) [9], [10]. Các thiết bị có số lượng điện cực EEG lớn thường, có độ chính xác cao, đắt tiền, sử dụng chủ yếu trong các phòng thí nghiệm và không thuận tiện cho việc di chuyển đi lại. Trái lại, các thiết bị có số lượng điện cực nhỏ có thể đeo được, di động, không dây và được thương mại nên chi phí phù hợp thì lại cho độ chính xác thấp hơn [10]. Do vậy, *khi triển khai ứng dụng thực tế, cần quan tâm xây dựng mô hình EEG_ER sao cho mô hình này vẫn hoạt động tốt với số lượng điện cực EEG được sử dụng là thấp.*

Hiệu quả hoạt động của hệ thống EEG_ER còn phụ thuộc rất nhiều vào các đặc trưng của tín hiệu EEG đầu vào. Từ dữ liệu EEG thô ban đầu, rất nhiều phương pháp trích, chọn đặc trưng được đề xuất bao gồm các đặc trưng về thời gian, đặc trưng về tần số hoặc là kết hợp cả đặc trưng về thời gian tần số [11]. Do vậy, *việc lựa chọn được đặc trưng phù hợp sẽ góp phần nâng cao hiệu suất nhận dạng của các hệ thống EEG_ER.*

Về kỹ thuật phân lớp cảm xúc, các phương pháp nhận dạng cảm xúc dựa trên EEG trước đây chủ yếu được phát triển dựa trên học máy truyền thống [12]. Gần đây, lấy cảm hứng từ khả năng học của các kỹ thuật học sâu, *những mô hình học sâu tiên tiến khác nhau đã được sử dụng ngày càng nhiều trong nhận dạng cảm xúc EEG* [13], [14], [15]. Với yêu cầu đòi hỏi ngày càng cao của cuộc sống, các ứng dụng EEG_ER chỉ thực sự phù hợp và có ý nghĩa thực tế khi được triển khai theo định hướng trí tuệ nhân tạo vạn vật (Artificial Intelligence of Things – AIoT) [16]. Hình 0.2 mô tả kiến trúc thực tế của một hệ thống EEG_ER theo định hướng AIoT. Trong đó, để triển khai bài toán EEG_ER trên thiết bị biên của mạng AIoT, các thiết bị thu thập tín hiệu EEG ngoài tính di động được và kết nối không dây thì thường có số lượng điện cực nhỏ nhằm đảm bảo tính cơ động, nhỏ gọn và tiết kiệm năng lượng. Các thuật toán xử lý tín hiệu, trích chọn đặc trưng và phân lớp được cài đặt trên một vi điều khiển nhỏ gọn để nhận dạng cảm xúc ngay tại thiết bị biên sau đó gửi kết quả lên đám mây rồi thực thi các tác vụ thông qua giao diện người dùng. Chính vì vậy, bộ phân lớp dựa trên các thuật toán học sâu cũng cần phải được điều chỉnh sao cho đạt được kiến trúc với số lượng tham số nhỏ trong khi vẫn phải duy trì tốt hiệu quả nhận dạng để có thể *triển khai được theo thời gian thực trên những vi điều khiển thông dụng, có kích thước nhỏ gọn, tài nguyên hạn chế* (như Raspberry PI, STM32Cube.AI...).

Từ các phân tích ở trên cho thấy vấn đề cấp thiết đặt ra là cần nghiên cứu để tìm kiếm một hệ thống EEG_ER hoạt động hiệu quả đối với các ứng dụng AIoT. Tức là cần xác định được kiến trúc mạng học sâu và loại đầu vào đặc trưng phù hợp. Điều này có nghĩa là bên cạnh việc lựa chọn đầu vào đặc trưng, kiến trúc học sâu được lựa chọn phải có kích thước nhỏ gọn, có thể hoạt động theo thời gian thực trên các vi điều khiển thông dụng. Đồng thời, kiến trúc này cũng phải cho hiệu quả nhận dạng cảm xúc cao trong khi các thiết bị thu thập tín hiệu EEG đầu

vào chỉ là các thiết bị có số lượng điện cực nhỏ, di động được, giao tiếp không dây và chi phí thấp. Đây cũng là mục đích chính của luận án.



Hình 0.2. Hệ thống EEG_ER theo tiếp cận AIoT

2. Mục tiêu của luận án

Để đạt được mục đích nêu trên, luận án cần phải thực hiện **bốn mục tiêu chính theo thứ tự lần lượt như sau:**

- 1) Cập nhật các kiến thức tổng quan về bài toán EEG_ER;
- 2) Lựa chọn kiến trúc học sâu phù hợp
- 3) Thiết kế mạng học sâu nhẹ
- 4) Tối ưu kiến trúc

3. Đối tượng, phạm vi và phương pháp nghiên cứu

3.1. Đối tượng nghiên cứu:

3.2. Phạm vi nghiên cứu:

3.3. Phương pháp nghiên cứu

4. Các đóng góp của luận án

- 1) Trên cơ sở phân tích các công trình công bố gần nhất cũng như tiến hành các mô phỏng đánh giá cần thiết, NCS đã chỉ ra được việc sử dụng kiến trúc mạng học sâu CNN 1D với đầu vào là các đặc trưng FFT của tín hiệu EEG là phù hợp nhất để triển khai bài toán EEG_ER. Công bố kết quả này trong các công trình CB [2], CB [3], CB [4].
- 2) Để xây dựng một hệ thống EEG_ER phù hợp với các ứng dụng AIoT, NCS đã đề xuất cải tiến hiệu quả hoạt động của mạng CNN 1D truyền thống bằng cách áp dụng hai kỹ thuật mới (Inception và Squeeze and Excitation). Các cải tiến này làm cho kiến trúc mạng học sâu thu được có kích thước đủ nhỏ (chỉ khoảng 40.14MB) trong khi vẫn duy trì hiệu suất nhận dạng đủ lớn mà không tốn nhiều tài nguyên. Kiến trúc này cho kết quả hoạt động tốt (trên hai lớp bài toán phân lớp nhị phân và đa lớp) ngay cả khi số lượng đầu vào chỉ EEG với 5 điện cực. Kết quả nghiên cứu này được công bố trong các công trình CB [5] và CB [6].
- 3) Nhằm nâng cao hơn nữa hiệu quả hoạt động của mạng học sâu ở trên đôi với các ứng dụng AIoT, NCS đã đề xuất áp dụng kỹ thuật học chuyên giao được gọi là “Tự chưng cất tri thức” (Self-Knowledge Distillation - SKD). Kỹ thuật này cho phép mô hình tự học từ các nhãn mềm mà nó sinh ra, giúp mô hình hội tụ nhanh hơn vì mô hình đã có nền tảng từ các dự đoán ban đầu, từ đó giảm thời gian suy luận, số lượng các phép toán dấu phẩy động (FLOPs) khi triển khai thực nghiệm. Thêm vào đó, trong kỹ thuật SKD, NCS cũng đưa ra giải pháp cải tiến bằng việc áp dụng kỹ thuật tăng cường dữ liệu, trọng số (Scaled weight) trong quá trình tính hàm Loss của SKD. Chứng minh mô hình EEG-SKDNet đạt hiệu suất vượt trội so với các phương pháp KD khác với số lượng tham số và FLOPs nhỏ, độ ổn định cao, thích hợp triển khai trên thiết bị AIoT. Kết quả nghiên cứu này sẽ được công bố trong CB [7].

5. Ý nghĩa khoa học và thực tiễn

5.1. Ý nghĩa khoa học

5.2. Ý nghĩa thực tiễn

6. Bố cục của luận án

Chương 1: Tổng quan về bài toán nhận dạng cảm xúc dựa trên tín hiệu điện não

Chương 2: Lựa chọn kiến trúc học sâu nền tảng và đặc trưng đầu vào cho bài toán EEG-ER

Chương 3: Thiết kế kiến trúc mạng học sâu nhẹ cho hệ thống EEG_ER

Chương 4: Tối ưu hóa mô hình EEG_ER sử dụng kỹ thuật “Tự chưng cất tri thức”

CHƯƠNG 1

TỔNG QUAN BÀI TOÁN NHẬN DẠNG CẢM XÚC DỰA TRÊN TÍN HIỆU ĐIỆN NÃO

Chương 1 được bố cục để thể hiện việc hoàn thành mục tiêu số 1 đã đề ra của luận án. Do vậy, phần đầu chương 1 sẽ trình bày các kiến thức tổng quan về bài toán nhận dạng cảm xúc. Trong đó, luận án nhấn mạnh tầm quan trọng của việc nhận dạng cảm xúc khi xây dựng các hệ thống HCI, đồng thời nêu rõ cơ sở của việc xác định các cảm xúc thực tế thông qua tổ hợp các thành phần trong mô hình cảm xúc. Tiếp theo, luận án tập trung vào phân tích thể mạnh của định hướng EEG_ER và mô tả chi tiết các khâu cơ bản trong một hệ thống EEG_ER, cập nhật các nghiên cứu mới nhất về các kỹ thuật quan trọng cần triển khai trong các khâu thu thập, trích chọn đặc trưng tín hiệu EEG và phân lớp cảm xúc. Phần cuối chương tổng hợp các công trình nghiên cứu nhằm xác định rõ các thách thức cần phải vượt qua khi xây dựng một hệ thống EEG_ER theo định hướng AIoT để từ đó làm nền tảng cho việc tìm kiếm một mô hình học sâu hiệu quả.

1.1 Tổng quan về bài toán nhận dạng cảm xúc

1.1.1 Cảm xúc là gì?

1.1.2 Tầm quan trọng của việc nhận dạng cảm xúc trong giao tiếp người - máy

1.1.3 Các mô hình cảm xúc

1.2 Nhận dạng cảm xúc dựa trên tín hiệu điện não đồ EEG

1.2.1 Giới thiệu về EEG

1.2.2 Hệ thống nhận dạng cảm xúc dựa trên tín hiệu điện não

BCI là một hệ thống cho phép giao tiếp trực tiếp giữa não bộ và máy tính, giúp người dùng có thể điều khiển thiết bị hoặc tương tác với môi trường mà không cần sử dụng hệ thần kinh ngoại vi hay cơ bắp. Ban đầu, các hệ thống BCI chủ yếu được thiết kế để hỗ trợ người bị khuyết tật vận động hoặc các bệnh lý thần kinh nghiêm trọng như chấn thương tủy sống, đa xơ cứng. Tuy nhiên, với sự phát triển mạnh mẽ của khoa học dữ liệu và trí tuệ nhân tạo, ứng dụng của BCI đã mở rộng sang nhiều lĩnh vực khác nhau, bao gồm chẩn đoán y tế, phục hồi chức năng, giám sát trạng thái tâm lý, nhà thông minh và điều khiển thiết bị điện tử. Thu nhận và xử lý số tín hiệu EEG.

1.2.3 Kích thích cảm xúc

1.2.4 Thu nhận và xử lý số tín hiệu EEG

1.3.2.1. Thiết bị thu thập

1.3.2.2. Thu nhận và xử lý tín hiệu EEG

1.2.5 Một số bộ dữ liệu cho bài toán EEG_ER

1.3 Trích xuất đặc trưng tín hiệu EEG

1.3.1 Trích xuất miền thời gian

1.3.2 Trích xuất miền tần số

1.3.3 Trích xuất miền thời gian-tần số

1.3.4 Trích xuất miền không gian

1.4 Phân lớp tín hiệu EEG trong bài toán nhận dạng cảm xúc

1.4.1 Các kỹ thuật học máy truyền thống

1.4.2 Các kỹ thuật học sâu

Không giống học máy truyền thống, các phương pháp học sâu tự học đặc trưng trực tiếp từ dữ liệu qua nhiều lớp mạng neuron, giảm thiểu sự cần thiết của trích chọn đặc trưng thủ công. Với khả năng học phi tuyến mạnh, học sâu đang trở thành xu hướng chính trong phân loại cảm xúc EEG [54], [55]. Dưới đây giới thiệu một số mô hình sâu đã được nghiên cứu và áp dụng với EEG:

1.5.2.1. Mạng nơ-ron tích chập (Convolutional Neural Network - CNN)

1.5.2.2. Mạng nơ-ron hồi quy

1.5.2.3. Cập nhật một số kết quả nghiên cứu về ứng dụng các kỹ thuật học sâu trong bài toán EEG_ER

1.5 Phát triển các hệ thống nhận dạng tín hiệu EEG theo định hướng AIoT

1.6 Kết luận chương 1

CHƯƠNG 2

LỰA CHỌN KIẾN TRÚC HỌC SÂU NỀN TẢNG VÀ ĐẶC TRƯNG ĐẦU VÀO CHO BÀI TOÁN EEG-ER

Biểu diễn và chuẩn hóa dữ liệu

2.1.1 Biểu diễn tín hiệu EEG dưới dạng ma trận

2.1.2 Biểu diễn tín hiệu EEG theo batch và cửa sổ thời gian

2.2 Lựa chọn các phương pháp trích chọn đặc trưng

2.2.1 Cơ sở lựa chọn các phương pháp trích chọn đặc trưng

Có nhiều phương pháp được sử dụng để trích xuất đặc trưng từ tín hiệu EEG trong miền tần số, bao gồm: Biến đổi Fourier nhanh (FFT), Mật độ phổ công suất (PSD), Biến đổi Wavelet (WT), Biến đổi Fourier ngắn hạn (STFT), Hilbert-Huang Transform (HHT), và Empirical Mode

Decomposition (EMD). Tuy nhiên, không phải phương pháp nào cũng phù hợp cho các ứng dụng yêu cầu hiệu suất tính toán cao và triển khai trên thiết bị di động. Do đó, việc lựa chọn phương pháp trích xuất đặc trưng trong nghiên cứu này tuân theo ba tiêu chí sau:

1. Các phương pháp được lựa chọn phải được sử dụng rộng rãi trong nghiên cứu liên quan đến nhận dạng cảm xúc từ EEG nhằm đảm bảo độ tin cậy.

2. Do mục tiêu nghiên cứu là triển khai mô hình trên thiết bị di động, các phương pháp trích xuất đặc trưng cần có độ phức tạp tính toán hợp lý để đảm bảo tốc độ xử lý gần thời gian thực.

3. Các phương pháp phải đại diện tốt cho phổ tần số của tín hiệu EEG, và có thể giữ lại thông tin thời gian quan trọng.

2.2.2 Trích xuất đặc trưng bằng biến đổi Fourier nhanh (FFT)

2.2.3 Trích xuất năng lượng tín hiệu EEG bằng Mật độ phổ công suất (PSD)

2.2.4 Trích xuất đặc trưng thời gian - tần số EEG bằng Biến đổi

2.2.5 Wavelet (WT)

2.3 Lựa chọn kiến trúc mô hình nền tảng

2.3.1 Cơ sở lựa chọn mô hình học sâu

2.3.2 Mạng nơ-ron tích chập (CNN - Convolutional Neural Networks)

2.3.3 Mạng nơ-ron học sâu BiLSTM

2.3.4 Mô hình kết hợp CNN và BiLSTM

2.4 Tiến hành thực nghiệm và đánh giá kết quả

2.4.1 Thiết lập thực nghiệm

2.4.1.1. Lựa chọn bộ dữ liệu chuẩn

Để đánh giá hiệu quả của các kiến trúc học sâu và phương pháp trích xuất đặc trưng trong bài toán nhận diện cảm xúc EEG, nghiên cứu sử dụng ba bộ dữ liệu phổ biến và có độ tin cậy cao gồm: DEAP,

AMIGOS, và DREAMER. Các bộ dữ liệu này đều được thu thập trong môi trường thực nghiệm có kiểm soát, với thông tin EEG được đồng bộ hóa cùng các phản hồi cảm xúc tự đánh giá của người tham gia (Self-Assessment).

2.4.1.2. *Cài đặt các thuật toán trích chọn đặc trưng*

2.4.1.3. *Thiết lập kiến trúc mạng*

2.4.2 *Thực nghiệm và đánh giá kết quả*

Để có thể đối sánh với các kết quả của các công bố trong [13], NCS sử dụng hai độ đo là độ chính xác trung bình kết quả nhận dạng của 4 nhãn (Acc) và hàm tổn thất trung bình (loss). Bên cạnh đó, để có thể đánh giá đồng thời 2 chỉ số acc và loss, với mục tiêu tối đa acc và tối thiểu loss, NCS sử dụng hệ số score, đây là chỉ số tổng hợp được tạo ra dựa trên độ chính xác và hàm tổn thất.

$$Score = Accuracy - \lambda \times Loss \quad (2.1)$$

Bảng 2.1. So sánh hiệu quả nhận dạng của mạng CNN đề xuất với các mạng CNN tiêu biểu

Mô hình	Độ chính xác trung bình (%)	Độ tổn thất trung bình (%)	Số lượng tham số
Alextnet	90.95	0.3211	22.178.186
Resnet	92.25	0.3088	4.323.394
GoogleNet	90.21	0.3689	11.647.042
VGG16	92.39	0.3484	32.219.466
MobileNet	22.34	0.2080	3.216.714
DenseNet	91.09	0.3463	3.727.562
Mạng CNN đề xuất	95.17	0.1886	16.740.778

NCS cũng thực hiện đánh giá hiệu quả hoạt động của mạng CNN đề xuất so với một số mạng CNN tiêu biểu được giới thiệu trong trong việc nhận dạng dữ liệu DEAP [76], [97]. Kết quả thể hiện trong Bảng 2.1 cho thấy mô hình mạng CNN được đề xuất cho kết quả nhận dạng

tốt nhất và có số lượng tham số ở mức trung bình. Điều này cho thấy hoàn toàn có thể triển khai kiến trúc mạng CNN này cho các ứng dụng nhận dạng cảm xúc trong thực tế với 32 kênh EEG tương tự như dữ liệu DEAP.

2.5 Kết luận chương 2

CHƯƠNG 3

THIẾT KẾ KIẾN TRÚC MẠNG HỌC SÂU NHẸ CHO HỆ THỐNG EEG_ER

3.1 Phân tích yêu cầu khi triển khai bài toán EEG_ER trên thiết bị có tài nguyên hạn chế

3.2 Các kỹ thuật tối ưu hóa mô hình học sâu

3.2.1 Kỹ thuật giảm *overfitting* và cải thiện tổng quát hóa

3.2.1.1. *Regularization (L1, L2)*

3.2.1.2. *Dropout*

3.2.1.3. *Batch Normalization (BN)*

3.2.1.4. *Early Stopping*

3.2.2 Kỹ thuật cải thiện khả năng học sâu và hội tụ nhanh

3.2.2.1. *Điều chỉnh tốc độ học (Learning Rate Scheduling)*

3.2.2.2. *Kỹ thuật tối ưu hóa dựa trên biến đổi động lực*

3.2.3 Nâng cao khả năng học đặc trưng trong EEG

3.2.3.1. *Tập trung vào các kênh EEG quan trọng (Channel Attention)*

Dữ liệu EEG có tính chất không đồng nhất giữa các kênh, nghĩa là không phải tất cả các kênh đều có mức độ quan trọng như nhau trong việc nhận diện cảm xúc. Một số kênh chứa thông tin quan trọng hơn về hoạt động thần kinh liên quan đến cảm xúc, trong khi các kênh khác có thể chứa nhiều hoặc thông tin không hữu ích. Tuy nhiên, các mô hình học sâu truyền thống thường xử lý tất cả các kênh EEG như nhau, dẫn đến việc không tận dụng tối đa các kênh quan trọng và có thể bị ảnh hưởng bởi nhiễu. Để giải quyết vấn đề này, [105] (SE) Block là một cơ

chế giúp tối ưu hóa việc học đặc trưng giữa các kênh EEG bằng cách tập trung vào các kênh quan trọng nhất.

3.2.3.2. *Học đa tầng (Multi-Scale Feature Learning)*

Học đa tầng (Multi-Scale Feature Learning) là một kỹ thuật quan trọng giúp mô hình có thể học các đặc trưng EEG ở nhiều cấp độ khác nhau. Tín hiệu EEG là một dạng dữ liệu có cấu trúc phức tạp, chứa nhiều thành phần tần số khác nhau, mỗi thành phần phản ánh một khía cạnh khác nhau của hoạt động thần kinh. Các dải tần như Alpha, Theta, Beta, Gamma có liên quan đến các trạng thái nhận thức và cảm xúc khác nhau, do đó việc trích xuất đặc trưng từ nhiều dải tần số này là một yếu tố quan trọng để cải thiện hiệu suất nhận diện cảm xúc. Tuy nhiên, các mô hình CNN truyền thống thường sử dụng một bộ lọc tích chập duy nhất với kích thước cố định để trích xuất đặc trưng, điều này có thể làm mất đi thông tin quan trọng từ các thang đo thời gian khác nhau.

3.2.3.3. *Kết hợp nhiều nguồn đặc trưng (Feature Fusion)*

3.3 Đề xuất mô hình học sâu có kích thước nhỏ cho bài toán EEG_ER

3.3.1 *Mô tả chung*

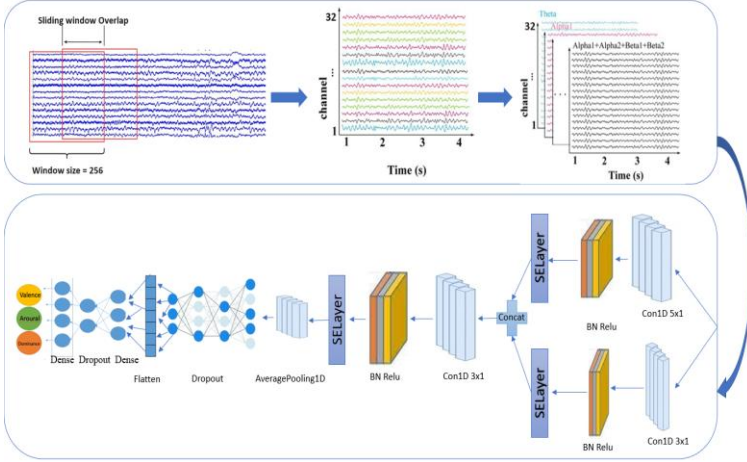
3.3.2 *Kiến trúc chi tiết*

Nhằm tối ưu hóa khả năng nhận diện cảm xúc từ tín hiệu EEG trong điều kiện số kênh thấp và tài nguyên tính toán hạn chế, luận án đề xuất kiến trúc chung cho EEG_SICNet và EEG_MCISNet. Cả hai mô hình đều được xây dựng dựa trên nền tảng CNN1D, kết hợp với Squeeze-and-Excitation (SE) Block và Inception Module nhằm tăng cường khả năng học đặc trưng EEG hiệu quả hơn.

Dựa trên những đặc điểm của bài toán, hai mô hình EEG_SICNet và EEG_MCISNet được thiết kế tương ứng với hai mục tiêu triển khai khác nhau.

EEG_MCISNet hướng đến bài toán phân loại nhị phân, giúp triển khai trên các hệ thống tài nguyên hạn chế như thiết bị di động hoặc

AIoT. Mô hình được tinh giản để tối ưu hóa tốc độ suy luận trong khi vẫn giữ được khả năng nhận diện chính xác.



EEG_SICNet được tối ưu hóa cho bài toán phân loại đa lớp, nhằm nhận diện nhiều trạng thái cảm xúc với độ chính xác cao nhất.

3.3.3 Đánh giá cấu trúc và hiệu suất thực nghiệm của EEG_MCSNet, EEG_SICNet

3.3.3.1. Đánh giá thực nghiệm EEG_SICNet

Bảng 3.1 cho thấy rằng mạng EEG_SICNet đạt hiệu suất cao nhất trên bộ dữ liệu DEAP với 32 kênh. Cụ thể, đối với nhãn Arousal, độ chính xác đạt $95.56 \pm 0.78\%$, F1 Score là $95.41 \pm 0.97\%$. Khi số lượng kênh giảm xuống còn 14, hiệu suất trên DEAP vẫn duy trì ở mức cao với Accuracy là $93.87 \pm 1.02\%$, cho thấy tính tổng quát hóa tốt của mô hình. Bảng 3.1 cung cấp một cái nhìn tổng quan về hiệu năng của mạng EEG_SICNet khi giảm số lượng kênh đầu vào xuống còn 5 kênh, nhằm kiểm chứng tính khả dụng của.

Bảng 3.2 cung cấp một cái nhìn tổng quan về hiệu năng của mạng EEG_SICNet khi giảm số lượng kênh đầu vào xuống còn 5 kênh, nhằm kiểm chứng tính khả dụng của mô hình trong bối cảnh sử dụng các thiết bị EEG di động hoặc giá rẻ với số lượng kênh hạn chế. Kết quả thu được

cho thấy rằng dù số lượng kênh giảm, các chỉ số Accuracy và F1-score vẫn duy trì ở mức chấp nhận được trên tất cả các bộ dữ liệu, cụ thể:

Bảng 3.1 Hiệu suất mạng EEG_SICNet trên các bộ dữ liệu EEG

Bộ dữ liệu	Nhãn	Accuracy (%)	F1 Score (%)	Recall (%)	Precision (%)
DEAP (32)	Arousal	95.56 ± 0.78	95.41 ± 0.97	95.44 ± 0.72	95.42 ± 0.86
	Valence	95.92 ± 0.81	95.81 ± 0.88	95.88 ± 0.62	95.92 ± 0.64
	Dominance	95.76 ± 0.69	95.80 ± 0.77	95.73 ± 0.67	95.74 ± 0.63
DEAP (14)	Arousal	93.87 ± 1.02	93.87 ± 1.11	93.80 ± 1.04	93.82 ± 1.07
	Valence	93.79 ± 0.85	93.91 ± 0.99	93.85 ± 0.87	93.99 ± 1.05
	Dominance	93.81 ± 0.80	93.88 ± 0.93	93.85 ± 0.96	93.82 ± 0.96
DREAMER (14)	Arousal	92.41 ± 0.44	92.43 ± 0.50	92.41 ± 0.51	92.41 ± 0.51
	Valence	92.56 ± 0.54	92.55 ± 0.53	92.56 ± 0.55	92.56 ± 0.54
	Dominance	92.28 ± 0.55	92.25 ± 0.56	92.22 ± 0.57	92.29 ± 0.59
AMIGOS (14)	Arousal	90.53 ± 0.66	90.54 ± 0.67	90.53 ± 0.65	90.54 ± 0.65
	Valence	90.74 ± 1.01	90.77 ± 0.74	90.75 ± 0.70	90.75 ± 0.70
	Dominance	90.50 ± 0.77	90.48 ± 0.70	90.47 ± 0.65	90.51 ± 0.65

3.3.3.2. *Đánh giá thực nghiệm EEG_MCIS*

Các số liệu thể hiện trên Bảng 3.7 cũng cho thấy mô hình mạng do NCS đề xuất có số lượng tham số đủ nhỏ, luôn hội tụ nhanh hơn (số Epoch luôn nhỏ hơn) với độ chính xác nhận dạng lớn hơn so với các mô hình còn lại.

Bảng 3.2 Kết quả hoạt động của một số mô hình học sâu trên bộ dữ liệu DEAP

No de	Mod el	No. Para ms (M)	Arousal			Valence			Dominance		
			Loss	Acc %	Epo ch	Loss	Acc %	Epo ch	Loss	Acc %	Epo ch
32	EEG_MCIS	12.701	0.0111	99.69	30	0.0116	99.62	31	0.0108	99.70	30
	M2	2.494	0.0221	99.10	100	0.0218	99.16	100	0.0216	99.25	100
	CNN	6.253	0.0243	99.17	32	0.0237	99.15	33	0.0254	99.23	35
14	EEG_MCIS	6.063	0.0191	99.55	19	0.0092	99.62	19	0.0112	99.67	20
	M2	1.154	0.0206	97.80	100	0.0206	99.22	100	0.0206	98.51	100
	CNN	4.493	0.0200	99.23	26	0.0195	99.26	26	0.0198	99.12	28
5	EEG_MCIS	2.967	0.0188	99.32	63	0.0137	98.51	43	0.0114	99.57	37
	M2	4.940	0.0521	97.24	100	0.0505	97.87	100	0.0511	97.87	100
	CNN	3.533	0.0674	97.12	66	0.0695	99.26	67	0.0617	96.97	70

3.4 Kết luận chương 3

CHƯƠNG 4

TỐI ƯU HÓA MÔ HÌNH EEG_ER SỬ DỤNG KỸ THUẬT “TỰ CHUNG CẤT TRI THỨC”

4.1 Tổng quan về kỹ thuật chung cất tri thức và kỹ thuật tự chung cất tri thức

4.1.1 Chung cất tri thức (Knowledge Distillation – KD)

4.1.2 Tự chung cất tri thức (Self-Knowledge Distillation - SKD)

4.2 Kỹ thuật tối ưu hóa mô hình EEG-ER sử dụng SKD và

4.3 Scaled Weights

4.3.1 Thiết lập bài toán

Hàm mất mát cross-entropy giữa nhãn thực y và phân phối xác suất đầu ra p được ký hiệu là: $CE(y, p)$

$$CE(y, p) = - \sum_{k=1}^K y_k \log p_k \quad (4.1)$$

Gọi độ đo mất mát Kullback-Leibler (KL) giữa hai phân phối xác suất p và q là:

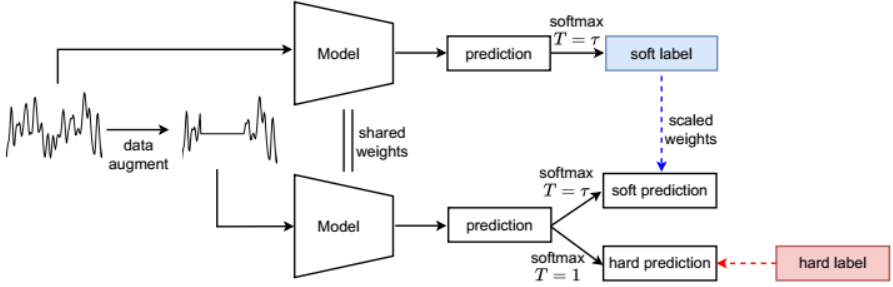
$$KL(p, q) = \sum_{k=1}^K p_k \log \frac{p_k}{q_k} \quad (4.2)$$

4.3.2 Đề xuất kiến trúc EEG-SKDnet

Tổng quan về phương pháp đề xuất của NCS được minh họa trong Hình 4.1. Trong đó: Đường gạch đỏ thể hiện hàm mất mát cross-entropy tiêu chuẩn, Đường gạch xanh thể hiện hàm mất mát knowledge distillation, T biểu thị giá trị của hệ số nhiệt độ (temperature value).

Cụ thể, với tín hiệu EEG đầu vào là x , NCS áp dụng các chiến lược tăng cường dữ liệu để tạo ra một tín hiệu tổng hợp mới ký hiệu là $\tilde{x}$. Gọi

mô hình là với θ là tập các trọng số có thể huấn luyện của mô hình. Các dự đoán của với hai đầu vào khác nhau là và tức là:



Hình 4.1 Tổng quan về phương pháp được đề xuất

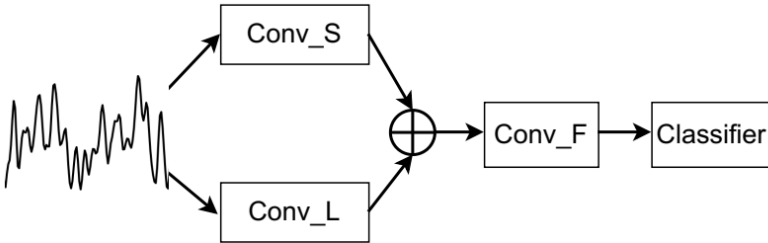
4.3.3 Trọng số có điều chỉnh (Scaled-Weights)

Trong nghiên cứu này, NCS áp dụng khung tăng cường dữ liệu được đề xuất trong [124] để sinh ra các mẫu EEG tổng hợp, từ đó nâng cao tính đa dạng và khả năng kháng nhiễu của tập dữ liệu huấn luyện:

$$x'_{ij} = \mathcal{T}(x_{ij}) \quad (4.3)$$

4.3.4 Chứng minh toán học về tác động của hệ số nhiệt và trọng số có điều chỉnh lên phân phối xác suất đầu ra trong EEG-SKNet

4.3.5 Thiết lập kiến trúc mạng



Hình 4.2 Mạng nhẹ được sử dụng trong EEG-SKNet

4.4 Thực nghiệm

4.4.1 Bộ dữ liệu

4.4.2 Chi tiết triển khai

4.4.3 So sánh và đánh giá với phương pháp cơ sở

Bảng 4.1 liệt kê hiệu quả hoạt động của các các kịch bản KD bao gồm phương pháp cơ sở, KD dựa trên đầu ra (Response-Based KD), KD dựa trên đặc trưng (Feature-Based KD), Progressive Self-Knowledge Distillation (PS-KD), Multi-Teacher Knowledge Distillation (MT-KD), và phương pháp được đề xuất với các cấu hình khác nhau. Tất cả kết quả được báo cáo trên tập DEAP(32), với các siêu tham số được giữ nguyên

Bảng 4.1 So sánh giữa các kịch bản KD khác nhau

Method	Acc	F1-score
Baseline	93.44 ± 0.78	93.43 ± 0.65
Response-based KD	93.15 ± 1.14	93.18 ± 1.17
Feature-based KD	90.87 ± 0.98	90.87 ± 0.98
PS-KD	94.02 ± 0.77	94.02 ± 0.79
MT-KD	93.99 ± 1.28	93.99 ± 1.30
EEG-SKDNet <i>w/o wscaled</i>	95.88 ± 0.53	95.87 ± 0.45
EEG-SKDNet	96.48 ± 0.35	96.48 ± 0.36

Bảng 4.2 So sánh độ chính xác top-1 giữa mô hình cơ sở và EEG-SKDNet trên các thành phần cảm xúc trong ba bộ dữ liệu: DEAP(32), DEAP(14), và DREAMER.

Component	Method	DEAP (32)	DEAP (14)	DREAMER
Valence	Baseline	95.92 ± 0.81	93.29 ± 0.85	91.16 ± 0.54
	EEG-SKDNet	96.67 ± 0.34	93.72 ± 0.27	92.41 ± 0.32

Arousal	Baseline	95.56 ± 0.78	93.27 ± 1.02	91.21 ± 0.44
	EEG-SKDNet	96.32 ± 0.36	94.01 ± 0.26	92.88 ± 0.30
Dominance	Baseline	95.76 ± 0.69	93.21 ± 0.82	91.28 ± 0.55
	EEG-SKDNet	96.45 ± 0.36	93.79 ± 0.28	92.39 ± 0.31
All	Baseline	95.75 ± 0.76	93.26 ± 0.89	91.22 ± 0.51
	EEG-SKDNet	96.48 ± 0.35	93.84 ± 0.27	92.56 ± 0.31

4.4.4 So sánh với các phương pháp tiên tiến (SOTA)

Bảng 4.3 So sánh hiệu suất giữa các phương pháp tiên tiến trên các tập dữ liệu phổ biến.

Method	Acc	Input shape	#classes	#params (M)	MFLOPs	Inference time(ms)
DEAP (32)						
MT-KD [125]	65.00	(128, 64)	2	—	—	—
SDC-GAT [126]	88.17	(32×6)	2	—	—	—
M1 [72]	99.22	(160, 1)	2	9.49	27.44	35.66
CNN & LSTM [127]	99.22	(256, 1)	2	—	—	475.14
TC_Net [128]	98.82	(48, 128, 32)	2	1.45	7.22	122.97
ResNet34 [118]	83.9	(32, 28)	2	7.29	17.55	75.04
MSDA-SFE [129]	70.10	(48, 128, 14)	2	17.02	242.02	3445.77

TDMNN [130]	99.00	(9,8,32)	2	15.25	—	—
DSE-Mixer [131]	95.19	(9,8,32)	2	0.32	90.00	—
STILN [132]	68.31	—	2	—	67.27	—
EEG-SKDNet (ours)	99.18	(160, 1)	2	2.63	9.33	22.12
Rhythm-specific CNN [133]	95.00	(1280, 32)	9	0.19	153.09	40.524
MTL_MSRN [134]	71.92	(32, 32)	9	1.00	350.00	—
EEG-SKDNet (ours)	96.48	(160, 1)	9	2.64	9.34	22.46
DREAMER						
Rhythm-specific CNN [133]	94.45	(1280, 14)	2	0.18	132.45	1280.3
TC_Net [128]	98.67	(48, 128, 14)	2	1.44	7.26	122.07
TDMNN [130]	95.00	(9,8,14)	2	15.25	31.92	—
EEG-SKDNet (ours)	98.76	(105, 1)	2	1.72	6.10	20.62

4.4.5 Nghiên cứu loại bỏ (Ablation Study)

Bảng 4.4 Độ chính xác Top-1 với các cấu hình khác nhau của hai siêu tham số chính α và τ trong mô hình EEG-SKDNet trên tập dữ liệu DEAP (32 kênh).

τ	α			
	$\alpha=0.2$	$\alpha=0.3$	$\alpha=0.5$	$\alpha=0.7$
3	90.06	92.12	94.72	95.71
5	90.15	92.34	94.05	96.02
7	91.32	92.50	94.62	96.17
10	91.45	92.68	95.00	96.48

4.5 Kết luận chương 4

KẾT LUẬN VÀ NHỮNG KIẾN NGHỊ NGHIÊN CỨU TIẾP THEO

1. Một số kết luận

Xây dựng kiến trúc mạng học sâu hiệu quả cho việc triển khai bài toán EEG_ER cho các ứng dụng HCI thực tế, triển khai tại các thiết bị biên của mạng AIoT là một xu hướng tất yếu. Tuy nhiên, khó khăn gặp phải là cần tìm kiếm một kiến trúc học sâu gọn nhẹ sao cho có thể cài đặt trên các thiết bị có khả năng lưu trữ và xử lý hạn chế trong khi vẫn phải duy trì được hiệu quả nhận dạng cao. Để giải quyết vấn đề này, trước tiên chúng ta cần phải xác định được kiến trúc học sâu nền tảng. Tiếp đó, cần áp dụng một số biện pháp kỹ thuật để cải tiến kiến trúc này sao cho đạt được mục tiêu về kích thước mô hình cũng như hiệu quả nhận dạng. So với mục tiêu đề ra, luận án đã thực hiện được các nội dung chủ yếu sau:

- Cập nhật các kiến thức tổng quan về bài toán EEG_ER, tập trung sâu vào nghiên cứu các kỹ thuật quan trọng cần triển khai trong các khâu thu thập, trích chọn đặc trưng tín hiệu EEG và phân lớp cảm xúc; Xác định rõ các thách thức cần phải vượt qua khi xây dựng một hệ thống EEG_ER theo định hướng AIoT. Một phần của nội dung này cũng đã được công bố trong CB [1].
- Nghiên cứu khả năng ứng dụng của các kỹ thuật học sâu khác nhau trong bài toán EEG_ER; Cập nhật các công trình nghiên

cứu gần nhất ứng dụng kỹ thuật học sâu để giải quyết bài toán EEG_ER; Tiến hành các thử nghiệm kiểm chứng để từ đó, xác định được *kiến trúc học sâu CNN 1D* cùng với *dạng đầu vào đặc trưng FFT* là phù hợp nhất đối với định hướng triển khai bài toán EEG_ER cho trên ứng dụng AIoT. Công bố các kết quả nghiên cứu trong các bài báo CB [2], CB [3], CB [4]

- Nghiên cứu *áp dụng hai kỹ thuật “Inception” và “Squeeze and Excitation”* để *cải tiến kiến trúc học sâu CNN 1D truyền thống* sao cho *phù hợp hơn với các ứng dụng AIoT*. Đề xuất hai mô hình học sâu có kích thước nhỏ nhưng hiệu quả nhận dạng cao. Chứng minh được hiệu quả hoạt động của hai mô hình này thông qua các công cụ mô phỏng. Công bố kết quả nghiên cứu trên các công trình CB [5] (mô hình EEG_MCSNet cho bài toán phân loại nhị phân) và CB [6] (Mô hình EEG_SICNet cho bài toán phân loại đa lớp).
- Dựa trên mô hình học sâu EEG_SICNet đã thu được, nghiên cứu *áp dụng kỹ thuật học chuyển giao SKD* nhằm *tăng hiệu suất nhận dạng, tăng tốc độ suy luận mà không cần tăng kích thước mô hình*. Trong đó, kỹ thuật Scaled weight cho student loss giúp mô hình tập trung hơn vào các tín hiệu quan trọng sau khi trích xuất FFT, giúp tăng hiệu suất và độ chính xác mà không cần tăng kích thước mô hình. Hiệu quả hoạt động của kiến trúc đạt được (mô hình EEG-SKDNet) đã được chứng minh thông qua các công cụ mô phỏng. Công bố kết quả nghiên cứu trên công trình CB [7].

2. Những kiến nghị nghiên cứu tiếp theo

Theo quan điểm của học viên, *luận án còn có một số hướng phát triển sau:*

- Nghiên cứu triển khai thử nghiệm cho một số bài toán hỗ trợ quyết định dựa trên thông tin từ tín hiệu EEG như điều khiển xe lăn điện, điều khiển nhà thông minh...

- Nghiên cứu áp dụng các giải pháp kỹ thuật mới nhằm nâng cao hơn nữa hiệu quả hoạt động của kiến trúc học sâu cho bài toán EEG_ER khi triển khai trên các ứng dụng AIoT thực tế.
- Đặc biệt đối với bài toán EEG_ER, nội dung của luận án mới chỉ tập trung vào việc tìm kiếm kiến trúc học sâu hiệu quả. Bên cạnh đó, dữ liệu cũng đóng vai trò vô cùng quan trọng. Việc thiếu hụt dữ liệu cũng như sự không đồng nhất và mất cân bằng dữ liệu là những vấn đề lớn cần phải giải quyết. Cần tìm kiếm các giải pháp mở rộng và làm giàu dữ liệu hiện có, sử dụng kỹ thuật bán giám sát để gán nhãn tự động, phát triển bộ dữ liệu kiểm tra chuẩn nhằm nâng cao độ chính xác và khả năng tổng quát hóa của các mô hình nhận dạng cảm xúc.

Do giới hạn về thời gian nghiên cứu và kiến thức của bản thân, luận án khó có thể tránh khỏi một số sai sót nhất định. NCS rất mong nhận được sự đóng góp ý kiến của các thầy cô, các nhà khoa học quan tâm để luận án được hoàn thiện hơn.

Một lần nữa NCS xin được cảm ơn Thầy giáo **PGS.TS. Phùng Trung Nghĩa** đã tận tình giúp đỡ, hướng dẫn trong thời gian thực hiện luận án, cảm ơn sự giúp đỡ của gia đình, bạn bè và các đồng nghiệp trong thời gian qua.

Thái Nguyên, ngày 20 tháng 5 năm 2025

Tác giả luận án

Dương Thị Mai Thương