

ĐẠI HỌC BÁCH KHOA HÀ NỘI
TRƯỜNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

BÁO CÁO BÀI TẬP LỚN

Học phần: IT4341 - Nhập môn trí tuệ nhân tạo

Đề tài: Nhận diện cảm xúc thông qua
tín hiệu điện não EEG

Học viên	Mã học viên
Vũ Đình Hiếu	20232300M
Đoàn Đức Cường	20232040M
Nguyễn Trần Hải Phong	20232019M

Giảng viên hướng dẫn TS. Trần Thế Hùng

HÀ NỘI, 06/2024

PHÂN CÔNG CÔNG VIỆC

Thành viên	Công việc
Vũ Đình Hiếu	Nghiên cứu bài toán, tổng hợp các phương pháp, lý thuyết, xây dựng phương pháp và mô hình nhận diện cảm xúc
Đoàn Đức Cường	Xây dựng mô đun tiền xử lý dữ liệu: trích chọn đặc trưng PSD, xây dựng ma trận 3 chiều đầu vào
Nguyễn Trần Hải Phong	Xây dựng mô đun tiền xử lý dữ liệu: tiền xử lý, trích chọn đặc trưng DE, xây dựng ma trận 2 chiều đầu vào

MỤC LỤC

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI.....	1
1.1 Giới thiệu bài toán	1
1.2 Mục tiêu và định hướng giải pháp	1
1.3 Bố cục báo cáo	2
CHƯƠNG 2. CƠ SỞ LÝ THUYẾT.....	3
2.1 Nhận diện cảm xúc (Emotion Recognition)	3
2.1.1 Các mô hình biểu diễn cảm xúc	3
2.1.2 Các phương thức phân tích cảm xúc.....	4
2.1.3 Tín hiệu điện não đồ (EEG)	4
2.1.4 Các phương pháp học máy và học sâu trong nhận diện cảm xúc	7
2.2 Nơ-ron nhân tạo.....	8
2.3 Mạng nơ-ron nhân tạo	9
2.4 Mạng nơ-ron tích chập	10
CHƯƠNG 3. PHƯƠNG PHÁP ĐỀ XUẤT.....	13
3.1 Tổng quan giải pháp.....	13
3.2 Tập dữ liệu	13
3.3 Xử lý dữ liệu	14
3.3.1 Trích chọn đặc trưng.....	14
3.3.2 Dựng ma trận đầu vào.....	15
3.4 Mô hình nhận diện cảm xúc.....	17
CHƯƠNG 4. ĐÁNH GIÁ THỰC NGHIỆM.....	18
4.1 Các tham số đánh giá	18
4.1.1 Binary Cross Entropy	18
4.1.2 Accuracy, Recall, Precision.....	18

4.2 Phương pháp thực nghiệm	19
4.3 Kết quả thực nghiệm	20
CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	21
5.1 Kết luận	21
5.2 Hướng phát triển.....	21
TÀI LIỆU THAM KHẢO.....	23

DANH MỤC HÌNH VẼ

Hình 2.1	Bánh xe cảm xúc của Plutchik và Ekman [1]	3
Hình 2.2	Mô hình theo chiều dữ liệu [1]	4
Hình 2.3	Một số phương thức phân tích cảm xúc [1]	5
Hình 2.4	Minh họa các thùy trên vỏ đại não [3]	5
Hình 2.5	Sơ đồ đặt các nút điện cực đo EEG [2]	6
Hình 2.6	Một số phương pháp học máy và học sâu nhận diện cảm xúc[1]	6
Hình 2.7	So sánh giữa một tế bào nơ-ron sinh học và một đơn vị nơ-ron nhân tạo [14]	8
Hình 2.8	Minh họa một mạng nơ-ron nhân tạo đơn giản với một lớp ẩn .	10
Hình 2.9	Minh họa một mạng nơ-ron tích chập	11
Hình 2.10	Minh họa thực hiện phép tích chập	12
Hình 3.1	Minh họa dữ liệu thuộc tính 'data' và 'label'	14
Hình 3.2	Tương quan trung bình giữa các chiều đo valence, arousal và liking (yêu thích) với phổ năng lượng của từng dải tần [17]	14
Hình 3.3	Minh họa ánh xạ ma trận 2 chiều [18]	16
Hình 3.4	Minh họa dựng ma trận 3 chiều [9]	16
Hình 3.5	Kiến trúc mạng tích chập	17

DANH MỤC BẢNG BIỂU

Bảng 2.1	Bảng tổng hợp một số bài nghiên cứu sử dụng mô hình học máy	7
Bảng 2.2	Bảng tổng hợp một số bài nghiên cứu sử dụng mô hình học sâu	7

DANH MỤC THUẬT NGỮ VÀ TỪ VIẾT TẮT

Thuật ngữ	Ý nghĩa
API	Giao diện lập trình ứng dụng (Application Programming Interface)
CVRP	Bài toán định tuyến phương tiện có giới hạn về tải (Capacitated Vehicle Routing Problem)
SBRP	Bài toán định tuyến xe buýt trường học (School Bus Routing Problem)
VRP	Bài toán định tuyến phương tiện giao thông (Vehicle Routing Problem)

CHƯƠNG 1. GIỚI THIỆU ĐỀ TÀI

1.1 Giới thiệu bài toán

Khi đời sống vật chất ngày càng được cải thiện, chúng ta bắt đầu đi sâu hơn về đời sống tinh thần của mỗi người, đặc biệt là về vấn đề cảm xúc. Cảm xúc không chỉ đơn thuần là trạng thái tâm lý của cảm giác, suy nghĩ hay hành vi mà còn là những phản ứng về mặt tâm lý, sinh lý đối với các kích thích. Cảm xúc đóng vai trò quan trọng trong đời sống của con người. Và việc nhận diện cảm xúc là bài toán cần thiết cho nhiều lĩnh vực đời sống. Rất nhiều lĩnh vực ứng dụng bài toán nhận diện cảm xúc như trong (i) tâm lý học, (ii) tương tác người máy, (iii) y tế,... Ví dụ cơ bản như trong y tế, các bác sĩ có thể thông qua việc nhận diện cảm xúc để đánh giá mức độ một số bệnh liên quan đến thần kinh ở bệnh nhân cũng như khả năng đáp ứng các phương pháp điều trị của họ.

Hiện nay, các mô hình nhận diện cảm xúc được phân thành 2 loại: (i) các phương pháp dựa trên các tín hiệu sinh lý, (ii) các phương pháp dựa trên các đặc điểm tín hiệu không sinh lý (non-physiological). Trong đó, phương pháp dựa trên các tín hiệu sinh lý cho thấy ưu điểm vượt trội và có thể xác định trạng thái cảm xúc đúng đắn hơn do ít bị ảnh hưởng bởi một số yếu tố chủ quan. Vậy nên, các phương pháp dựa trên loại tín hiệu này được sử dụng tương đối rộng, đáng tin.

Tín hiệu điện não đồ EEG là tín hiệu sinh lý được sử dụng phổ biến bởi tính tự phát, khách quan và khả năng phản ánh đúng về trạng thái cảm xúc con người. Các phương pháp nhận diện cảm xúc dựa trên tín hiệu điện não EEG chủ yếu được phát triển theo hai hướng: (i) sử dụng các mô hình học máy truyền thống, (ii) sử dụng các mô hình học sâu. Tuy nhiên, vì tín hiệu điện não EEG là dạng dữ liệu phi tuyến, nhiều chiều và phức tạp, nên hiện nay các phương pháp dựa trên học sâu được nghiên cứu tương đối sâu và cho kết quả rất tích cực.

Dựa trên những hiện trạng đó, đề tài của nhóm tập trung vào bài toán nhận diện cảm xúc dựa trên tín hiệu điện não EEG. Từ việc nghiên cứu báo cáo tổng hợp các phương pháp xử lý bài toán này của các tác giả tại [1], nhóm định hướng sử dụng kiến trúc mạng nơ-ron nhân tạo để xử lý bài toán với mục tiêu là nhận diện cảm xúc tích cực và tiêu cực theo độ đo Valence trong mô hình biểu diễn cảm xúc nhiều chiều.

1.2 Mục tiêu và định hướng giải pháp

Tại đề tài môn học, nhóm tiến hành trích chọn hai đặc trưng *Differential entropy* - *DE* và *Power spectral density* - *PSD* từ tập dữ liệu DEAP - tập dữ liệu tín hiệu

điện EEG. Sau khi dữ liệu được xử lý, nhóm sử dụng mô hình mạng nơ-ron tích chập để huấn luyện nhận diện cảm xúc tích cực và tiêu cực theo độ đo Valence của mô hình mô tả cảm xúc nhiều chiều. Các thuật ngữ và phương pháp xử lý sẽ được giới thiệu chi tiết lần lượt tại chương 2 và chương 3.

1.3 Bố cục báo cáo

Phần còn lại của báo cáo đề tài này được tổ chức như sau.

Chương 2 trình bày về cơ sở lý thuyết về các mô hình cảm xúc, tín hiệu điện não EEG; các phương pháp hiện tại được sử dụng cho bài toán nhận diện cảm xúc dựa trên tín hiệu điện não EEG; lý thuyết cơ bản về mạng nơ-ron nhân tạo và mạng nơ-ron tích chập.

Trong Chương 3, nhóm trình bày về phương pháp sử dụng để giải quyết bài toán

Tiếp theo, chương 4 trình bày về phương pháp thực nghiệm và kết quả.

Cuối cùng, chương 5 kết luận và định hướng phát triển về sau của đề tài.

CHƯƠNG 2. CƠ SỞ LÝ THUYẾT

2.1 Nhận diện cảm xúc (Emotion Recognition)

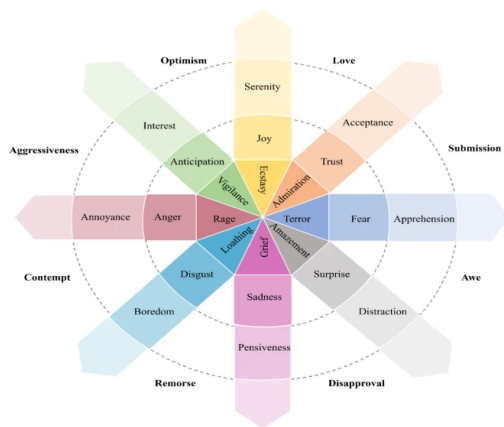
Cảm xúc là một trạng thái bao gồm cảm giác, suy nghĩ, hành động, là phản ứng về sinh lý của con người đối với các ảnh hưởng bên trong và bên ngoài. Những cảm xúc này có ảnh hưởng lớn đến việc giao tiếp, nhận thức và đưa ra quyết định. Nhận diện cảm xúc là bước vô cùng quan trọng trong lĩnh vực giao tiếp người máy.

2.1.1 Các mô hình biểu diễn cảm xúc

Các mô hình biểu diễn cảm xúc thường được chia thành 2 loại: Mô hình rời rạc (Discrete model) và Mô hình chiều dữ liệu (Dimensional model)

a, Mô hình rời rạc

Mô hình rời rạc hay còn gọi là mô hình cảm xúc theo phân loại (categorical emotion model), phân chia cảm xúc về các phân loại nhỏ hơn. Hai mô hình bánh xe cảm xúc tại hình 2.1a và hình 2.1b là hai mô hình được sử dụng phổ biến. Trong đó 6 cảm xúc cơ bản trong hình của thường được sử dụng, bao gồm: tức giận, ghê tởm, sợ hãi, hạnh phúc, buồn bã và ngạc nhiên [1].



(a) Mô hình của Plutchik



(b) Mô hình của Ekman

Hình 2.1: Bánh xe cảm xúc của Plutchik và Ekman [1]

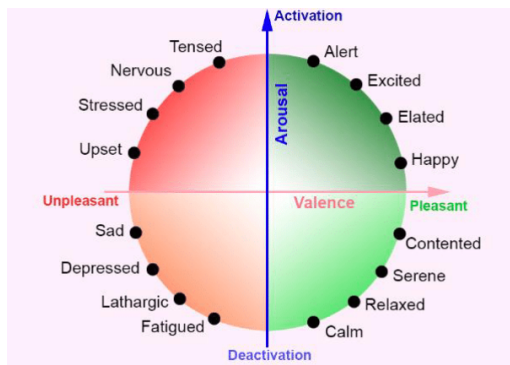
b, Mô hình chiều dữ liệu

Mô hình chiều dữ liệu sử dụng kết hợp các miền dữ liệu nhiều chiều liên tục để xác định các trạng thái cảm xúc khác nhau.

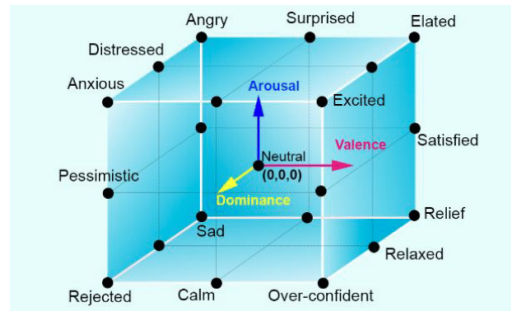
Mô hình 2 chiều như hình 2.2a minh họa thường sử dụng cặp đại lượng valence và arousal để biểu diễn cảm xúc (trong đó valence thể hiện trạng thái cảm xúc tích cực hay tiêu cực, còn arousal thể hiện mức độ tích cực hay tiêu cực như thế nào).

Mô hình 3 chiều bổ sung thêm dominance như mô hình 2.2b để đánh giá khả

năng kiểm soát cảm xúc của con người.



(a) Mô hình 2 chiều



(b) Mô hình 3 chiều

Hình 2.2: Mô hình theo chiều dữ liệu [1]

2.1.2 Các phương thức phân tích cảm xúc

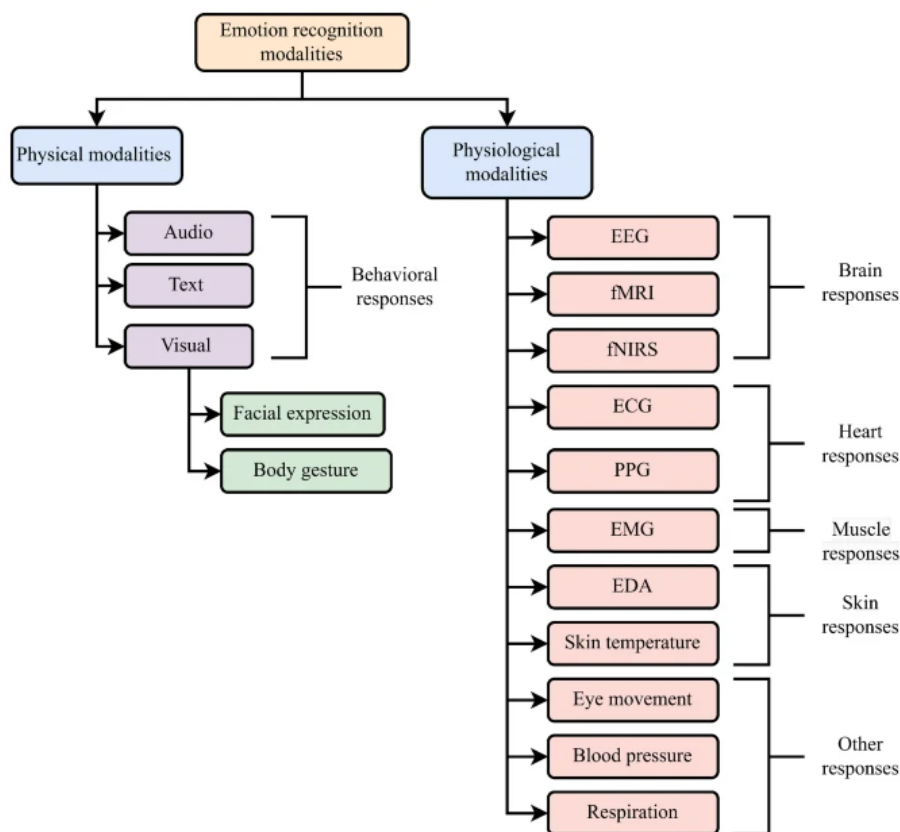
Để phân tích cảm xúc, ta có thể áp dụng nhiều phương thức dựa trên các đặc điểm vật lý cũng như sinh lý khác nhau. Như mô tả tại hình 2.3, các phương thức dựa trên các đặc điểm vật lý có thể bao gồm nhận diện qua (i) âm thanh, (ii) văn bản, (iii) các đặc điểm trực quan. Các đặc điểm trực quan có thể chia nhỏ hơn như gồm (i) biểu hiện qua nét mặt hay (ii) biểu hiện qua ngôn ngữ cơ thể. Tuy nhiên để đạt độ chính xác cao hơn, các nhà nghiên cứu sử dụng thêm các phương thức phân tích thông qua các đặc điểm sinh lý như (i) tín hiệu điện não đồ (EEG), (ii) cộng hưởng từ chức năng (fMRI), (iii) quang phổ cận hồng ngoại (fNIRS), (iv) tín hiệu điện tim (ECG), (v) cảm biến nhịp tim quang học (PPG),...

2.1.3 Tín hiệu điện não đồ (EEG)

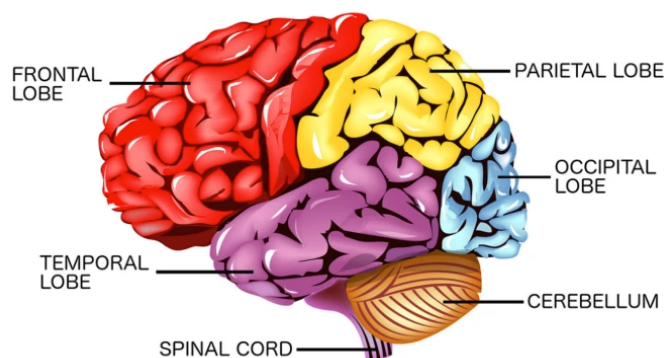
Bộ não là bộ phận phức tạp nhất trên cơ thể con người. Vỏ đại não là phần chiếm nhiều nhất, phụ trách các chức năng bậc cao bao gồm ngôn ngữ, ghi nhớ, học tập, cảm xúc, đưa ra quyết định, thông minh, ... Vỏ đại não phát ra tín hiệu điện sinh học gọi là tín hiệu điện não (brain electrical signal). Vỏ đại não chia thành 2 bán cầu và mỗi bán cầu chia thành 4 thùy gồm: (i) thùy trán (frontal), (ii) thùy đỉnh (parietal), (iii) thùy thái dương (temporal) và (iv) thùy chẩm (occipital) như mô tả tại hình 2.4.

Thông thường, bộ não người phát ra dòng điện rất nhỏ khoảng vài mi-li-vôn. Sự biến động điện thế này được gây ra bởi dòng ion di chuyển trong não, giữa các nơ-ron. Để thu thập các tín hiệu này, các nhà nghiên cứu sử dụng phương pháp điện não đồ (electroencephalographic - EEG). Đây là phương pháp không xâm lấn đo lường hoạt động điện của bộ não, sử dụng các điện cực được đặt tại một số điểm cố định trên đầu người đo[2]. Hình 2.5 minh họa cách đặt các điện cực trong EEG.

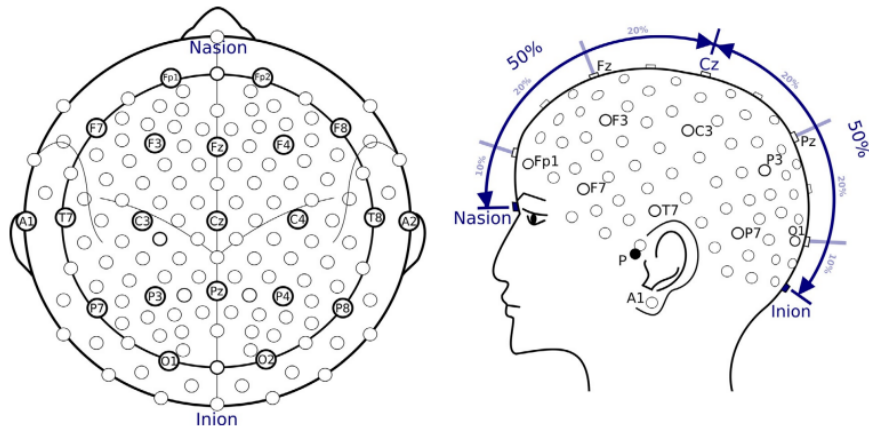
EEG là một trong những công cụ giúp quan sát hoạt động não hay sóng não hiệu quả nhất. Tín hiệu điện não đồ (EEG) đóng vai trò quan trọng trong việc tìm hiểu các quá trình sinh lý diễn ra trong bộ não người. Tín hiệu EEG có một số đặc điểm chủ yếu: (i) tín hiệu bị ảnh hưởng rất nhiều bởi nhiễu từ môi trường, cũng như từ các tín hiệu điện khác từ các mô, các bộ phận khác; cũng do bị ảnh hưởng bởi tín hiệu điện từ nhiều yếu tố bên ngoài nên tín hiệu EEG là (ii) phi tuyến và (iii) không ổn định; đặc biệt (iv) tín hiệu này có đặc trưng lớn ở miền tần số cụ thể, tập trung là ở miền tần số 0.5-30Hz.



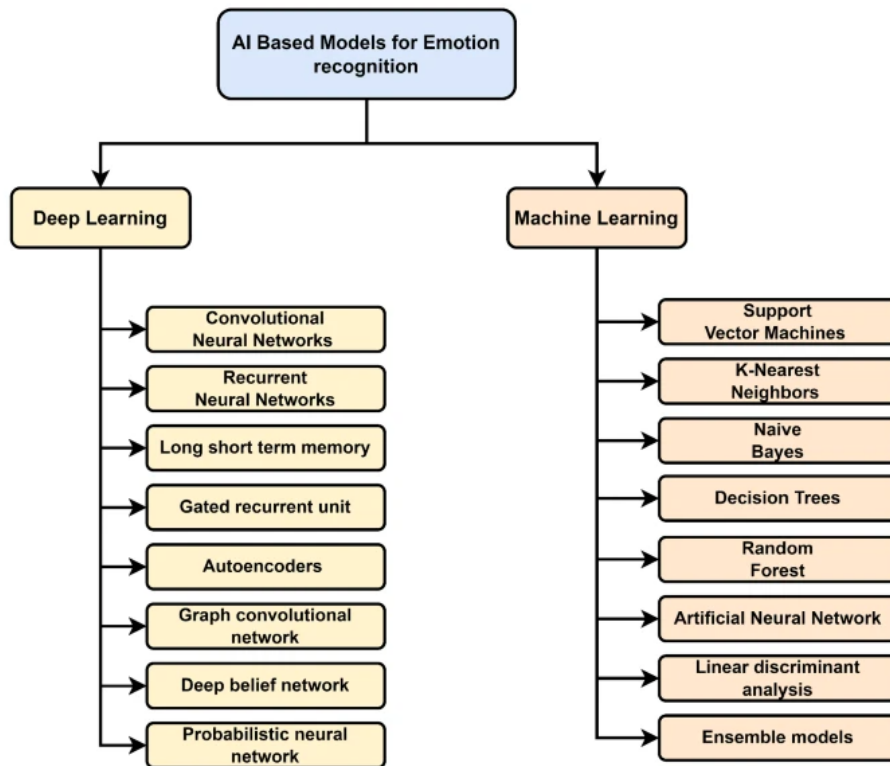
Hình 2.3: Một số phương thức phân tích cảm xúc [1]



Hình 2.4: Minh họa các thùy trên vỏ đại não [3]



Hình 2.5: Sơ đồ đặt các nút điện cực đo EEG [2]



Hình 2.6: Một số phương pháp học máy và học sâu nhận diện cảm xúc[1]

Tín hiệu EEG thường được chia thành hai nhóm là tự phát (spontaneous) hoặc được kích thích (evoked). Tín hiệu tự phát là tín hiệu tạo ra từ nội tại bản thân hệ thống nơ-ron thần kinh mà không phụ thuộc vào kích thích bên ngoài. Tín hiệu được kích thích được thu thập khi sử dụng các kích thích bên ngoài, tác động lên vỏ đại não. Ngoài ra, tín hiệu EEG còn được chia thành 5 dải tần nhỏ hơn dựa trên đặc trưng tần số bao gồm: (i) sóng Gamma (30-200Hz), (ii) sóng Beta (13-30Hz), (iii) sóng Alpha (8-13Hz), (iv) sóng Theta (4-8Hz), (v) sóng Delta (khoảng dưới 1-4Hz) [2].

2.1.4 Các phương pháp học máy và học sâu trong nhận diện cảm xúc

Học máy và học sâu là 2 phương pháp hiệu quả được sử dụng trong nhận diện cảm xúc dựa trên tín hiệu điện não đồ EEG. Trong [1], tác giả đã tổng hợp một số phương pháp bổ biến trong học máy và học sâu được sử dụng cho bài toán, được mô tả tại hình 2.6.

a, Các phương pháp học máy

Từ nghiên cứu của các tác giả tại [1], nhóm tổng hợp tại bảng 2.1 một số nghiên cứu sử dụng phương pháp học máy: (i) máy vector hỗ trợ (SVM), (ii) K - láng giềng gần nhất (KNN), (iii) rừng ngẫu nhiên (RD), (iv) cây quyết định (DT), (v) mạng nơ-ron nhân tạo (ANN).

Bảng 2.1: Bảng tổng hợp một số bài nghiên cứu sử dụng mô hình học máy

Tham chiếu	Tập dữ liệu	Bộ phân loại	Kết quả
Zhang cùng cộng sự [4]	DEAP	SVM	Hai lớp: 94,98% Nhiều lớp: 93,20%
Galvão cùng cộng sự [5]	DEAP	KNN, RF	Valence: 89,94% Arousal: 89,83%
Zhang cùng cộng sự [6]	DEAP	RF	Valence: 88,54% Arousal: 89,58%
Wagh và Vasanth [7]	SEED	DT, SVM, KNN	71,52%
Khubani và Kulkarni [8]	SEED, DEAP	ANN	94,23%

b, Các phương pháp học sâu

Từ nghiên cứu của các tác giả tại [1], nhóm tổng hợp tại bảng 2.2 một số nghiên cứu sử dụng phương pháp học sâu: (i) mạng nơ-ron tích chập (CNN), (ii) bộ nhớ dài ngắn hạn (LSTM), (iii) mạng nơ-ron hồi quy (RNN).

Bảng 2.2: Bảng tổng hợp một số bài nghiên cứu sử dụng mô hình học sâu

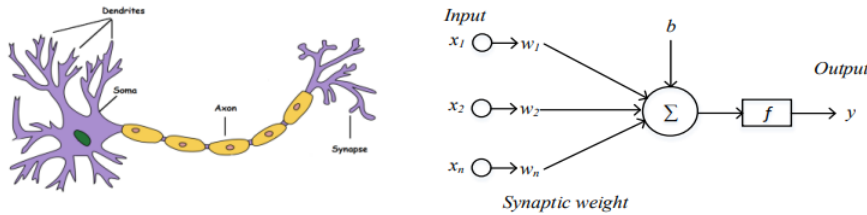
Tham chiếu	Dữ liệu	Bộ phân loại	Kết quả
Rudakov cùng cộng sự [9]	DEAP	MT-CNN	Valence: 96,28% Arousal: 96,62%
Iyer cùng cộng sự [10]	SEED, DEAP	CNN, LSTM	SEED: 97,16% DEAP: 65%
Chowdary cùng cộng sự [11]	EEG Brain wave	RNN, LSTM, GRU	97%

Continued on next page

Bảng 2.2: Bảng tổng hợp một số bài nghiên cứu sử dụng mô hình học sâu (Continued)

Yang cùng cộng sự [12]	DEAP	CNN	Valence: 90,01% Arousal: 90,65%
Pisipati và Nandy [13]	DREAMER	CNN	99,90%

2.2 Nơ-ron nhân tạo



Hình 2.7: So sánh giữa một tế bào nơ-ron sinh học và một đơn vị nơ-ron nhân tạo [14]

Nơ-ron là đơn vị cơ bản trong bộ não con người. Các nơ-ron kết nối với nhau tạo nên một mạng lưới các nơ-ron trong não giúp con người thực hiện các chức năng từ nhận thức, học, ghi nhớ đến hành động. Mạng nơ-ron nhân tạo (Artificial Neural Network - ANN) là một hệ thống tính toán được lấy ý tưởng từ mạng nơ-ron sinh học trong não bộ con người, tập hợp từ các đơn vị nơ-ron nhân tạo. Hình 2.7 minh họa so sánh giữa một tế bào nơ-ron sinh học và một đơn vị nơ-ron nhân tạo. Tùy thuộc vào cách các nơ-ron kết nối với nhau mà ta có các kiến trúc mạng khác nhau. Các kết nối trong mạng cho phép dữ liệu đầu vào được xử lý lần lượt và cho kết quả đầu ra tương ứng.

Mỗi nơ-ron nhân tạo là một mô hình toán của một nơ-ron sinh học cho phép tạo ra giá trị đầu ra dựa trên dữ liệu đầu vào. Mô hình toán của đơn vị này gồm 3 thành phần chính: (i) bộ nhân (multiplication), (ii) bộ cộng (summation) và (iii) hàm kích hoạt (activation). Các tín hiệu ngõ vào được nhân với trọng số tương ứng (weight), sau đó, các kết quả được tổng hợp, cộng lại và đưa qua hàm kích hoạt để thu đầu ra cuối cùng [14].

Kết quả đầu ra của nơ-ron nhân tạo được thực hiện bởi hàm kích hoạt. Mô hình toán học được mô tả bởi phương trình:

$$y = f \left(\sum_{i=1}^n w_i x_i + b \right) \quad (2.1)$$

Trong đó, y là đầu ra, f là hàm kích hoạt, x_i là dữ liệu đầu vào, w_i là các trọng số

tương ứng, b là giá trị độ lệch.

Hàm kích hoạt có thể là hàm ngưỡng (trong trường hợp đơn):

$$y = \begin{cases} 1, & \sum_{i=1}^n w_i x_i + b \geq threshold \\ 0, & \sum_{i=1}^n w_i x_i + b < threshold \end{cases} \quad (2.2)$$

hoặc có thể là các hàm phi tuyến như hàm Sigmoid, Tanh, Softmax hoặc hàm chỉnh lưu tuyến tính ReLU,...

Các nơ-ron ở lớp đầu ra trong bài toán phân lớp thường sử dụng hàm tạo ra phân bố xác suất như hàm Softmax [14].

$$Softmax(x) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (2.3)$$

Trước đây, các hàm sigmoid và tanh thường được dùng để huấn luyện mạng, tuy nhiên từ đầu những năm 2000, hàm ReLU thường được sử dụng thường xuyên hơn với các biến thể Leaky ReLU, PReLU và ELU [15].

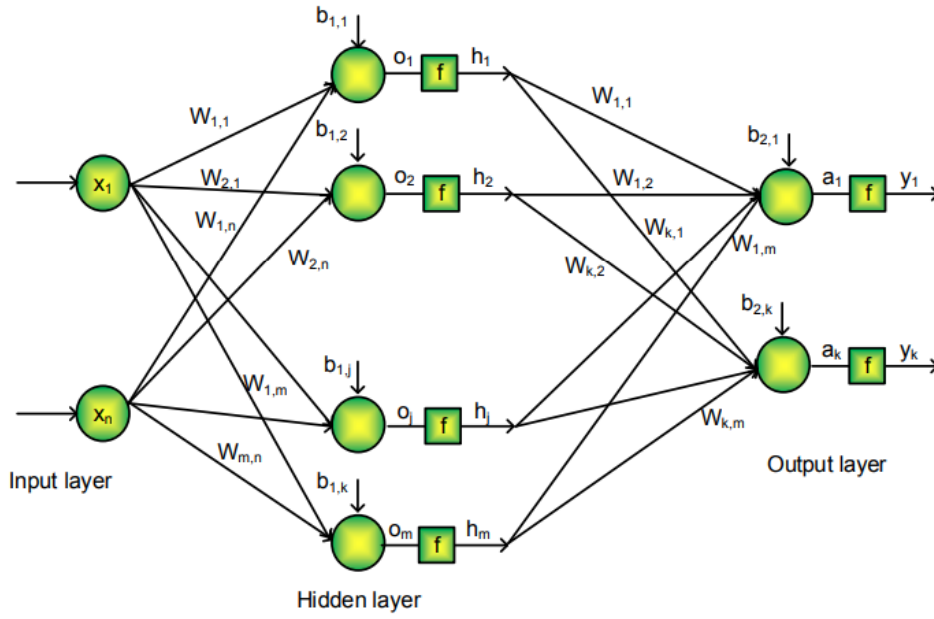
$$\sigma(s) = \frac{1}{1 + e^{-s}} \quad (2.4)$$

$$\tanh(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}} \quad (2.5)$$

$$ReLU(s) = \begin{cases} 0, & s < 0 \\ s, & s \geq 0 \end{cases} \quad (2.6)$$

2.3 Mạng nơ-ron nhân tạo

Mạng nơ-ron nhân tạo được hình thành từ việc kết nối các nơ-ron nhân tạo lại với nhau theo kiến trúc cụ thể nào đó. Mỗi kiến trúc sẽ có sự khác biệt về tính năng khác nhau. Kiểu kết nối đơn giản nhất với kết nối lan truyền thẳng được minh họa tại hình 2.8 với một lớp đầu vào, một lớp ẩn và một lớp đầu ra.



Hình 2.8: Minh họa một mạng nơ-ron nhân tạo đơn giản với một lớp ẩn

Giá trị đầu ra cuối cùng của mạng được xác định thông qua quá trình tính toán cho từng nút mạng từ đầu vào đến đầu ra. Minh họa toán học các bước tính toán theo hình 2.8:

$$o_i = \sum_{j=1}^n x_j W_{i,j} + b_{1,i} \quad (2.7)$$

$$h_i = f(o_i) \quad (2.8)$$

$$a_i = \sum_{j=1}^m h_j W_{i,j} + b_{2,i} \quad (2.9)$$

$$y_i = f(a_i) \quad (2.10)$$

Trong đó, o_i, a_i là giá trị đầu ra trước khi qua hàm kích hoạt lần lượt tại lớp ẩn và lớp đầu ra, f là hàm kích hoạt, h_i, y_i là giá trị đầu ra cuối cùng lần lượt tại lớp ẩn và lớp đầu ra, $W_{i,j}$ là các trọng số tương ứng, $b_{i,j}$ là các giá trị độ lệch tương ứng tại từng lớp.

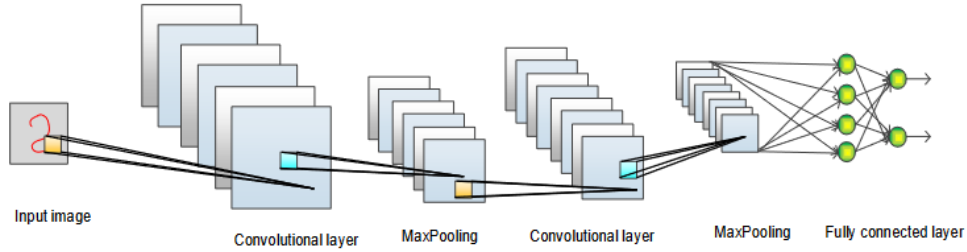
2.4 Mạng nơ-ron tích chập

Mạng nơ-ron tích chập (Convolutional Neural Network - CNN) dựa trên phép tích chập giữa ảnh với các bộ lọc đã được huấn luyện để tách các đặc trưng của

ảnh. Phương trình 2.11 mô tả phép tích chập giữa ảnh $I(i, j)$ và bộ lọc $K(u, v)$ kích thước $L \times L$

$$I'(i, j) = \sum_{u=-\frac{(L-1)}{2}}^{\frac{(L-1)}{2}} \sum_{v=-\frac{(L-1)}{2}}^{\frac{(L-1)}{2}} I(i-u, j-v)K(u, v) \quad (2.11)$$

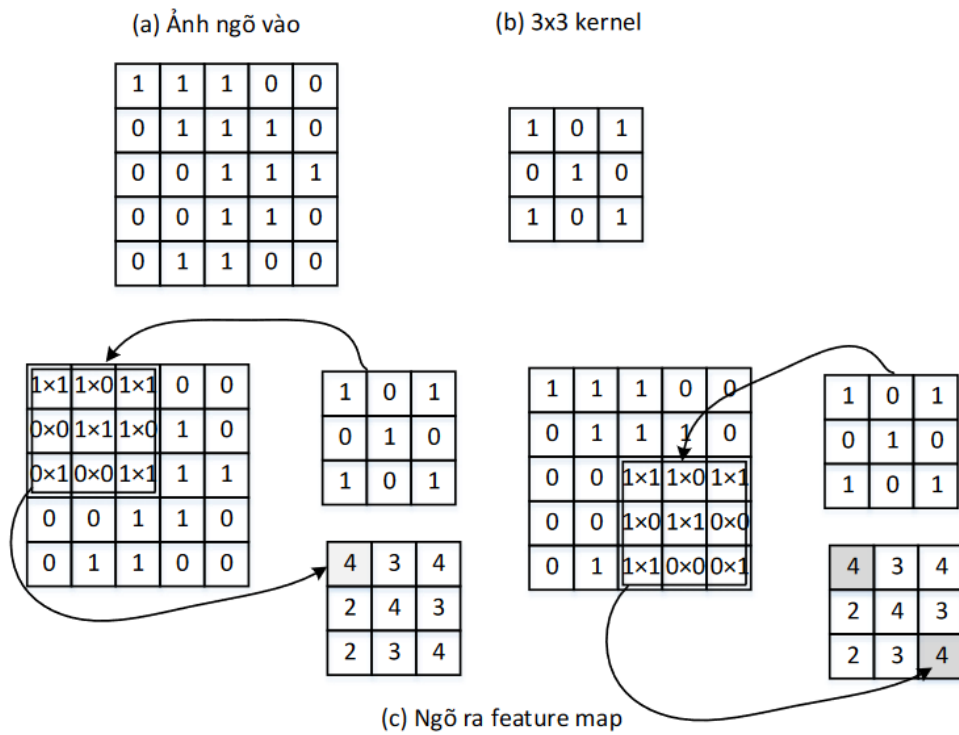
Mạng nơ-ron tích chập là sự kết hợp của mạng nơ-ron nhân tạo với các lớp kết nối đầy đủ và các lớp chập. Minh họa tại hình 2.9 là một mạng nơ-ron tích chập có kiến trúc bao gồm 2 lớp chập (convolutional layer), 2 lớp gộp (pooling layer) và lớp kết nối đầy đủ (fully-connected layer).



Hình 2.9: Minh họa một mạng nơ-ron tích chập

Dữ liệu đầu vào có thể như ví dụ tại hình 2.10 là ma trận nhị phân 2 chiều, bộ lọc (kernel) có kích thước 3×3 có các giá trị nhị phân. Hình minh họa các bước thực hiện phép tích chập trên ảnh và thu được đầu ra là ảnh đặc trưng (feature map) với kích thước nhỏ hơn. Ảnh đặc trưng có kích thước nhỏ hơn nhưng bao hàm được thông tin của ảnh gốc ban đầu, giúp việc tính toán hiệu quả, nên mạng CNN có thể xây dựng với nhiều lớp ẩn mà vẫn đáp ứng yêu cầu tính toán ở mức phù hợp.

Mạng nơ-ron tích chập điển hình có 3 lớp cơ bản: lớp tích chập, lớp gộp (pooling), lớp kết nối đầy đủ. Lớp tích chập sử dụng toán tử tích chập giữa một bộ lọc kernel với ma trận đầu vào. Lớp gộp có nhiệm vụ nhằm làm giảm kích thước nơ-ron của lớp trước, giảm độ phức tạp tính toán nhưng vẫn đảm bảo biểu diễn thông tin đặc trưng của đầu vào. Lớp kết nối đầy đủ thường được dùng làm lớp đầu ra với kiến trúc tương tự với mạng nơ-ron nhân tạo truyền thống cổ điển kết hợp với hàm kích hoạt, có thể thể là hàm phân bố xác suất Softmax [16].



Hình 2.10: Minh họa thực hiện phép tích chập

Ngoài ra, tùy theo nghiệp vụ, khi thiết kế mạng nơ-ron tích chập, một số lớp bổ sung khác được thêm vào nhằm tăng hiệu quả huấn luyện mô hình như: lớp Dropout, lớp chuẩn hóa. Lớp chuẩn hóa nhằm hiệu chỉnh giá trị kết quả từ lớp đầu vào trước khi sang lớp kế tiếp. Lớp chuẩn hóa giúp ổn định việc huấn luyện, cho phép nhiều mức học và tăng tốc huấn luyện mô hình [15]. Lớp Dropout được sử dụng để ngẫu nhiên bỏ đi một số nút riêng biệt theo xác suất là $1 - p$ để tránh vấn đề quá khớp dữ liệu.

CHƯƠNG 3. PHƯƠNG PHÁP ĐỀ XUẤT

3.1 Tổng quan giải pháp

Từ phần tổng hợp lý thuyết, các nghiên cứu trong chương 2 và mô hình đã nghiên cứu của các tác giả trong [9], nhóm lựa chọn phương án giải quyết bài toán gồm 2 phần tiền xử lý dữ liệu và mô hình mạng nơ-ron tích chập.

Nhóm sử dụng tập dữ liệu DEAP - đo lường tín hiệu điện não EEG theo các cảm xúc khác nhau. Sử dụng độ đo Valence theo mô hình cảm xúc 2 chiều để đánh giá cảm xúc là tích cực (1) hay tiêu cực (0). Các tín hiệu đo được sẽ được tách kênh lọc theo các dải tần, trích chọn đặc trưng và sử dụng để dựng ma trận 3 chiều đầu vào của mạng nơ-ron tích chập. Mạng tích chập nhóm sử dụng có kiến trúc đơn giản gồm nhiều lớp ẩn, nhận đầu vào là ma trận 3 chiều và đầu ra là xác suất cho thấy cảm xúc là tích cực hay tiêu cực.

3.2 Tập dữ liệu

Tập dữ liệu sử dụng là tập dữ liệu DEAP. Tập dữ liệu này sử dụng mô hình biểu diễn cảm xúc 2 chiều (valence-arousal). DEAP được thu thập từ việc đo điện não đồ của 32 tình nguyện viên với các kích thích cảm xúc khác nhau. Mỗi tình nguyện viên được xem 40 đoạn video ngắn, mỗi video dài 1 phút, được lựa chọn kỹ lưỡng để kích thích một cảm xúc nhất định. Trong quá trình xem, các tình nguyện viên sẽ đeo các điện cực đo EEG tại 32 điểm khác nhau trên đầu, nhằm ghi lại những diễn biến tín hiệu điện ứng với mỗi cảm xúc [17]. Dữ liệu thu được là giá trị tín hiệu điện trong khoảng thời gian dài 63 giây (3 giây trước khi xem video và 60 giây xem video) ứng với mỗi điểm đo ở mỗi tình nguyện viên với mỗi video. Vì để phù hợp với tính toán, DEAP được các tác giả lấy mẫu với tần số 128Hz (tức 8064 mẫu với đoạn ghi 63 giây), nên kích thước mỗi mẫu dữ liệu là $32 \times 40 \times 32 \times 8064$ tương ứng là (số tình nguyện viên \times số video \times số điểm đo \times số mẫu). Mỗi mẫu dữ liệu được gán nhãn giá trị theo thang điểm từ 1-9 theo 4 độ đo Valence, Arousal, Dominance và Liking theo mô hình biểu diễn cảm xúc nhiều chiều. Trong bài này, nhóm sử dụng độ đo Valence, với số điểm dưới 5 là cảm xúc tiêu cực và tích cực với số điểm ngược lại.

Dữ liệu từ tập dữ liệu DEAP được cho dưới dạng các file định dạng .mat với 32 file ứng với 32 tình nguyện viên được khảo sát. Mỗi tập dữ liệu có định dạng gồm 4 thuộc tính `__header__`, `__version__`, `__globals__`, `labels`, `data`, trong đó `data`, `label` lần lượt là dữ liệu tín hiệu EEG đo được và giá trị đo theo 4 độ đo Valence, Arousal, Dominance và Liking theo mô hình biểu diễn cảm xúc nhiều chiều. Hình 3.1 minh họa dữ liệu của hai thuộc tính này.

```
array([[ 9.48231681e-01,  1.65333533e+00,  3.01372577e+00, ...,
        -2.82648937e+00, -4.47722969e+00, -3.67692812e+00],
       [ 1.24706590e-01,  1.39008270e+00,  1.83509881e+00, ...,
        -2.98702069e+00, -6.28780884e+00, -4.47429041e+00],
       [-2.21651099e+00,  2.29201682e+00,  2.74636923e+00, ...,
        -2.63707760e+00, -7.40651010e+00, -6.75590441e+00],
       ...])
```

(a) Dữ liệu thuộc tính 'data'

```
array([[7.71, 7.6, 6.9, 7.83],
       [8.1, 7.31, 7.28, 8.47],
       [8.58, 7.54, 9., 7.08],
       [4.94, 6.01, 6.12, 8.06],
       [6.96, 3.92, 7.19, 6.05],
       [8.27, 3.92, 7., 8.03],
       ...])
```

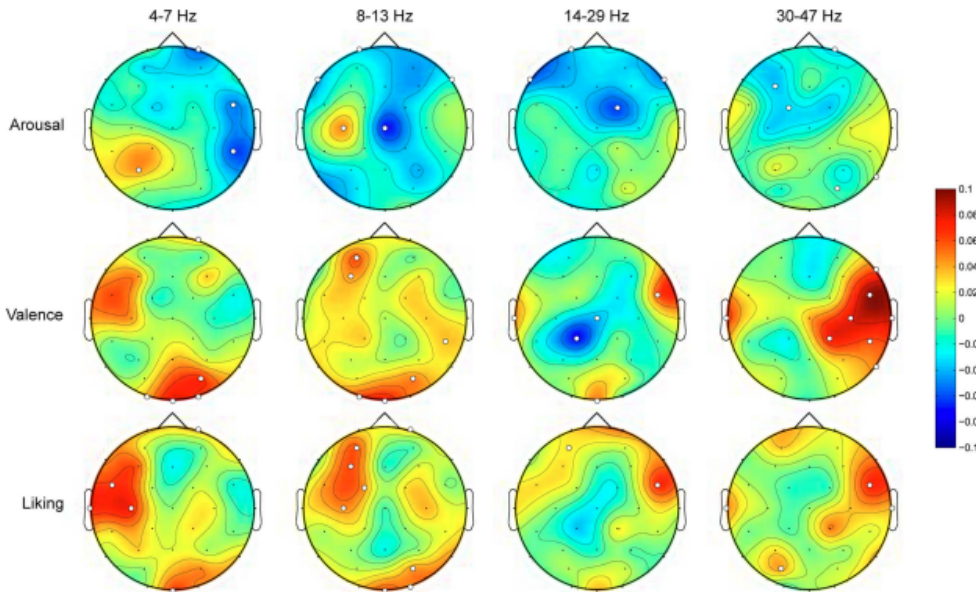
(b) Dữ liệu thuộc tính 'label'

Hình 3.1: Minh họa dữ liệu thuộc tính 'data' và 'label'

3.3 Xử lý dữ liệu

3.3.1 Trích chọn đặc trưng

Dựa trên kết quả từ bộ dữ liệu DEAP, ta có trong hình 3.2 là tương quan trung bình giữa các chiều đo valence, arousal và liking (yêu thích) với phổ năng lượng của từng dải tần ($\alpha, \beta, \gamma, \theta$). Từ đây, các tác giả trong có ý tưởng ánh xạ dữ liệu tại 32 đầu đo thành một ma trận 2 chiều và xử lý giống như một "bức ảnh" của các tín hiệu EEG. Thêm vào đó, vì các mẫu tuy rời rạc nhưng xuất phát từ việc đo liên tục trong thời gian khoảng 1 phút, các mẫu 2 chiều tiếp đến được ghép lại với nhau thành mẫu 3 chiều, và sử dụng đầu vào 3 chiều với mô hình mạng nơ-ron tích chập phía sau.



Hình 3.2: Tương quan trung bình giữa các chiều đo valence, arousal và liking (yêu thích) với phổ năng lượng của từng dải tần [17]

Tuy nhiên, dữ liệu từ DEAP cần phải tiền xử lý và trích chọn đặc trưng phù hợp trước khi đưa vào dựng "ảnh" 3 chiều. Theo các tác giả từ [17] và [18], vì phần lớn năng lượng tập trung tại 4 dải tần ($\theta(3 - 7Hz)$, $\alpha(8 - 13Hz)$, $\beta(14 - 29Hz)$, $\gamma(30 - 45/47Hz)$) nên dữ liệu tín hiệu EEG tại các điểm đo được trích xuất phần dải tần từ 3-45Hz dựa trên bộ lọc *ButterWorth*. ButterWorth là bộ lọc thông thấp, tuy nhiên

sử dụng kết hợp các bộ lọc thông thấp để tạo bộ lọc thông dải [19], tách dải tần 3-45Hz.

Dữ liệu sau khi đã tách các thành phần dải tần được các tác giả trong [9], đề xuất trích chọn hai đặc trưng là entropy vi sai (*Differential entropy* - DE) và phổ mật độ năng lượng (*Power spectral density* - PSD). Về đặc trưng DE, các tác giả tại [20], đã chứng minh rằng tín hiệu EEG tuân theo phân phối Gaussian tại mỗi bước 2Hz hay mỗi đoạn 0,5 giây, nên giá trị DE tại mỗi đoạn tín hiệu này, tại một dải tần cố định i được tính bởi công thức:

$$h_i(X) = \frac{1}{2} \log(2\pi e \sigma_i^2) \quad (3.1)$$

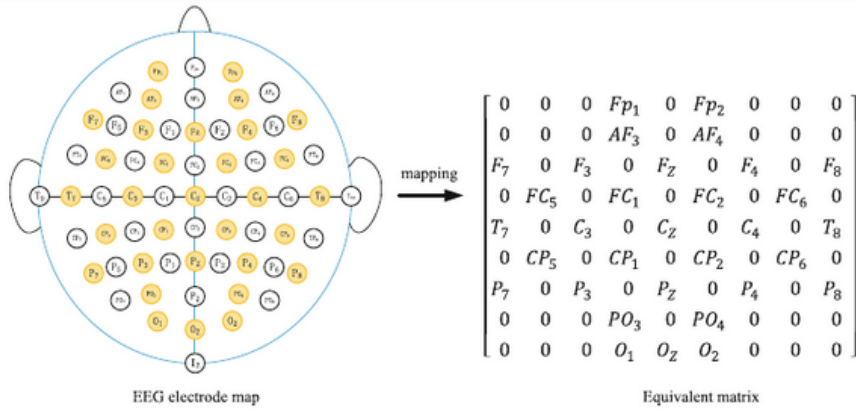
trong đó h_i là giá trị entropy vi sai và σ_i^2 là phương sai của tín hiệu. Vì vậy khi trích chọn đặc trưng, tất cả tín hiệu EEG sau khi tách dải sẽ được chia thành các đoạn 0.5 giây. Về đặc trưng PSD, theo quan hệ *Parseval*, ta có tổng năng lượng của tín hiệu bằng phổ mật độ năng lượng của tín hiệu đó [19]. Cộng thêm dữ liệu được chia thành các đoạn 0,5 giây khi tính DE nên ta có công thức tính PSD tại mỗi đoạn có N mẫu tín hiệu $x(n)$ là:

$$PSD = \sum_{n=0}^{N-1} x(n)^2 \quad (3.2)$$

Sau đó, các tín hiệu đo được tách phần năng lượng từ đường cơ sở (baseline) (là phần tín hiệu 3 giây đầu trước khi xem video của tình nguyện viên). Tiếp theo, toàn bộ tín hiệu điện đo khi xem video sẽ được trừ đi phần đường cơ sở để thu được phần biến động tương ứng khi tình nguyện viên tiếp nhận kích thích, đây gọi là *Baseline Correction*. Baseline Correction là cần thiết, vì khi đo tín hiệu dễ bị chông nhiễu từ nhiều nguồn bên ngoài cũng như từ bản thân người đo ví dụ từ tín hiệu điện của các mô da đầu, các mô xung quanh phần vỏ đại não,...[21].

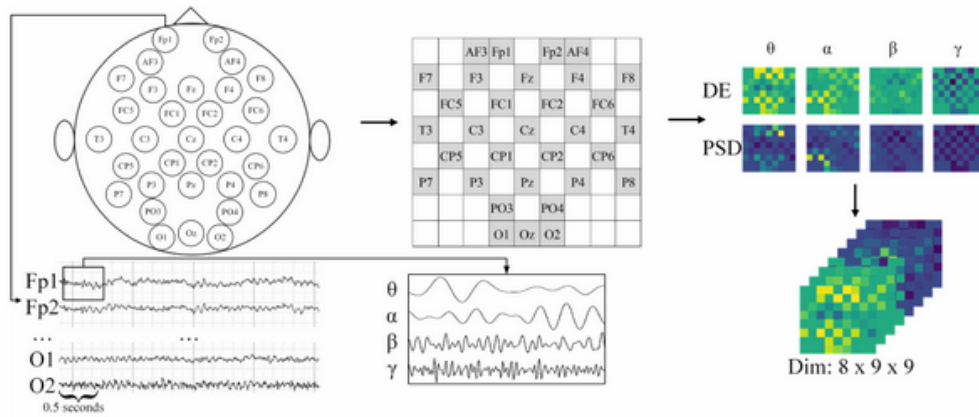
3.3.2 Dựng ma trận đầu vào

Hệ thống 10-20 quốc tế là hệ thống được công nhận về mô tả các vị trí đặt nút điện cực đo EEG. Sơ đồ bên trái của hình là mô tả hệ thống này, trong đó, các điểm tô vàng là điểm đặt đầu đo cho bộ dữ liệu DEAP. Vì để giữ được đặc trưng về mặt không gian, các tác giả trong [18], đề xuất biến đổi vector một chiều đầu vào $v \in \mathbb{R}^{32}$ thành ma trận 2 chiều $\mathbb{R}^{h \times w}$. Ví dụ trong hình 3.3, với $h = w = 9$, vector đặc trưng đầu vào được ánh xạ vào một ma trận hai chiều 9×9 .



Hình 3.3: Minh họa ánh xạ ma trận 2 chiều [18]

Như đã nêu trong 3.3.1, các tín hiệu đầu vào được tách thành 4 dải tần $\alpha, \beta, \gamma, \theta$. Điều này tương tự như tách kênh ảnh RGB, vậy nên các tác giả trong [18], có ý tưởng là ghép các ma trận 2 chiều vừa thu được ứng với từng dải tần thành ma trận 3 chiều. Ngoài ra, vì trong 3.3.1 ta sử dụng 2 đặc trưng DE và PSD nên trong [9] các tác giả ghép cả 2 đặc trưng kèm 4 dải tần thành ma trận 3 chiều $8 \times 9 \times 9$ như minh họa tại hình 3.4.

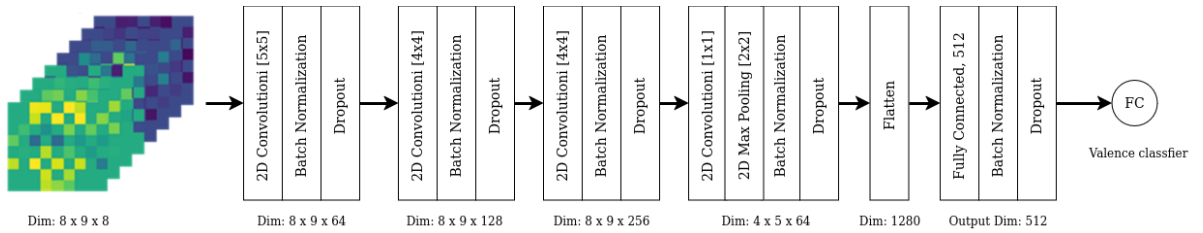


Hình 3.4: Minh họa dựng ma trận 3 chiều [9]

Như vậy dữ liệu đầu vào, xét đối với 1 tình nguyện viên có kích thước $40 - video \times 32 - channel \times 8064 - data$ sau khi tách 3 giây đầu còn $40 - video \times 32 - channel \times 7680 - data$ chia thành các đoạn 0,5 giây còn lại $40 - video \times 32 - channel \times 120 - data$. Sau khi tách thành 4 dải tần ta có kích thước $40 - video \times 32 - channel \times 120 - data \times 4 - band$ sau đó gộp thành $4800 - data \times 32 - channel \times 4 - band$. Sau đó dựng ma trận 3 chiều ta có kích thước dữ liệu đầu vào đối với 1 tình nguyện viên là $4800 - sample \times (4 \times 8 \times 9) - data$. (Kích thước ma trận là $4 \times 8 \times 9$ là theo tác giả [9] còn tác giả tại j[18] sử dụng kích thước $4 \times 9 \times 9$, nhóm sử dụng theo tác giả [9] nên lấy kích thước $4 \times 8 \times 9$).

3.4 Mô hình nhận diện cảm xúc

Mô hình mạng nơ-ron tích chập được dựa trên kiến trúc đã được xây dựng của các tác giả trong [9]. Chi tiết mô hình được nhóm sử dụng được cho tại hình 3.5. Trong đó các lớp tích chập được dùng để trích chọn đặc trưng, các lớp *Batch Normalization* (chuẩn hóa theo lô) được dùng để tăng tính ổn định của đặc trưng khi đi qua các lớp, giúp tổng thể mô hình thống nhất và nhanh hội tụ hơn. Các lớp *Dropout* được dùng để tránh quá khớp. Các tác giả trong [18] gọi đây là mô hình liên tục vì không có lớp *pooling* (để giảm chiều dữ liệu) tại một số lớp đầu tiên sau lớp tích chập. Điều này được các tác giả giải thích do dữ liệu đầu vào có kích thước nhỏ nên không cần thiết phải có thêm lớp *pooling*. Tuy nhiên để đảm bảo đặc trưng gần các cạnh, các lớp tích chập sử dụng *ze-ro padding* để không thay đổi kích thước hai chiều của dữ liệu (3 khối đầu tiên giữ kích thước 8×9 và tăng chiều sâu). Ngoài ra mô hình sử dụng hàm kích hoạt ReLU tại các khối ẩn và Sigmoid tại lớp đầu ra, tính xác suất cảm xúc của tín hiệu đầu vào là tích cực hay tiêu cực theo độ đo Valence.



Hình 3.5: Kiến trúc mạng tích chập

CHƯƠNG 4. ĐÁNH GIÁ THỰC NGHIỆM

4.1 Các tham số đánh giá

4.1.1 Binary Cross Entropy

Cross entropy loss hay còn gọi là *Log loss*, được tham số được dùng để đánh giá sự khác biệt giữa xác suất được dự đoán so với giá trị thực trong bài toán phân loại. Giá trị này càng thấp cho thấy mô hình phân loại càng tốt. Khi bài toán chỉ phân loại hai lớp, ta có trường hợp đặc biệt của cross-entropy loss là *Binary Cross Entropy* và công thức tính là:

$$BCE = -\frac{1}{N} \sum_{i=1}^N (y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))) \quad (4.1)$$

trong đó y_i là giá trị thực nhận hai giá trị 0,1, $p(y_i)$ là giá trị dự đoán.

4.1.2 Accuracy, Recall, Precision

a, Accuracy

Accuracy là tham số đánh giá độ chính xác của mô hình dựa trên tỉ lệ dự đoán đúng trên tổng số mẫu dự đoán. Công thức tính độ đo này là:

$$Accuracy = \frac{\sum \text{Dự đoán đúng}}{\sum \text{Tổng số mẫu}} \quad (4.2)$$

b, Recall và Precision

Độ đo *Recall* và *Precision* đều được dùng để đánh giá về độ chính xác của mô hình. Công thức của Recall và Precision lần lượt là:

$$Recall = \frac{TP}{TP + FN} \quad (4.3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4.4)$$

trong đó TP là *True Positive* là dữ liệu thuộc lớp 1 và dự đoán là 1 (1 coi là Positive và 0 là Negative), TN là *True Negative* là dữ liệu thuộc lớp 0 và dự đoán là 0, FP là *False Positive* là dữ liệu thuộc lớp 0 nhưng dự đoán là lớp 1, FN là *False Negative* là dữ liệu thuộc lớp 1 nhưng dự đoán là lớp 0. Như vậy, Recall đánh giá trong số những dữ liệu lớp 1, mô hình dự đoán đúng bao nhiêu và Precision đánh giá trong những dự đoán là 1 có bao nhiêu dự đoán đúng.

4.2 Phương pháp thực nghiệm

Nhóm thực hiện huấn luyện mô hình trên tập dữ liệu DEAP theo các phương pháp đã nêu tại chương 3. Trong đó, tập dữ liệu sau khi tiền xử lý gồm có 153.600 mẫu, mỗi mẫu có kích thước $8 \times 9 \times 8$ nhóm chia thành 2 phần để huấn luyện và kiểm thử theo tỉ lệ: 80% huấn luyện và 20% kiểm thử. Khi huấn luyện, nhóm sử dụng *K-fold cross-validation*, chia tập huấn luyện thành 5 phần, 4 phần dùng để huấn luyện và 1 phần còn lại để đánh giá mô hình, sau đó xáo trộn 5 phần để mỗi phần để được sử dụng để đánh giá 4 phần còn lại. Mỗi bước huấn luyện, mô hình sẽ xử lý theo 1 lô gồm 64 mẫu. Mô hình được huấn luyện với 50 *epoch*, mỗi lần chạy 1 epoch là 1 lần duyệt toàn bộ tập huấn luyện.

Nhóm sử dụng thư viện TensorFlow phiên bản 2.16.0 để xây dựng mô hình, cụ thể như sau:

```
seq.add(Conv2D(64, 5, activation="relu", padding="same",
              name="conv1", ))
if True:
    seq.add(BatchNormalization())
seq.add(Dropout(dropout_rate))
seq.add(Conv2D(128, 4, activation="relu", padding="same",
              name="conv2"))
if True:
    seq.add(BatchNormalization())
seq.add(Dropout(dropout_rate))
seq.add(Conv2D(256, 4, activation="relu", padding="same",
              name="conv3"))
if True:
    seq.add(BatchNormalization())
seq.add(Dropout(dropout_rate))
seq.add(Conv2D(64, 1, activation="relu", padding="same",
              name="conv4"))
seq.add(MaxPooling2D(2, 2, name="pool1"))
if True:
    seq.add(BatchNormalization())
seq.add(Dropout(dropout_rate))
seq.add(Flatten(name="fla1"))
seq.add(Dense(512, activation="relu", name="dense1"))
seq.add(Reshape((1, 512), name="reshape"))
if True:
```

```
seq.add(BatchNormalization())  
seq.add(Dropout(dropout_rate))
```

4.3 Kết quả thực nghiệm

Kết quả theo các tham số đánh giá sau khi huấn luyện 50 epoch trên tập dữ liệu kiểm thử của mô hình như sau:

- **Binary Cross Entropy:** 14,25%
- **Accuracy:** 94,69%
- **Recall:** 94,99%
- **Precision:** 95,44%

CHƯƠNG 5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1 Kết luận

Kết quả mô hình mạng nơ-ron tích chập đã nhận diện tốt cảm xúc tích cực và tiêu cực dựa trên tín hiệu điện não EEG. Kết quả nhận diện trên tập kiểm thử và trong quá trình huấn luyện cho độ chính xác đều trên 90%. Tuy nhiên do nhiều hạn chế, mô hình chỉ dừng lại ở nhận diện cảm xúc tích cực và tiêu cực theo độ đo Valence. Điều này là chưa tận dụng được tối đa các độ đo của bộ dữ liệu DEAP, cần phải cải tiến trong những phát triển về sau.

5.2 Hướng phát triển

Tuy rằng mô hình đã cho kết quả tốt với tín hiệu điện não EEG, nhưng mô hình hiện tại vẫn còn hạn chế là chỉ nhận diện cảm xúc tích cực và tiêu cực theo độ đo Valence. Vì vậy trong những hướng phát triển trong tương lai, mô hình cần cải tiến để mở rộng nhận diện theo các độ đo khác trong bộ dữ liệu DEAP như Arousal, Dominance, Liking. Bên cạnh đó, trong các mô hình biểu diễn cảm xúc nhiều chiều, để mô tả hay biểu diễn chính xác cảm xúc cần nhận diện cùng lúc nhiều độ đo, nên thay vì chỉ dự đoán một trong các độ đo đã nêu, mô hình cần mở rộng để nhận diện nhiều độ đo cùng một lúc. Ngoài ra, tín hiệu EEG tuy cho kết quả nhận diện cảm xúc rất tốt nhưng thực nghiệm để thu thập dữ liệu tương đối phức tạp, điều này đặt ra trong những hướng phát triển tương lai, cần nghiên cứu kết hợp các tín hiệu khác để đảm bảo nhận diện đúng nhưng có thể dễ dàng ứng dụng thực nghiệm.

TÀI LIỆU THAM KHẢO

- [1] P. Samal and M. F. Hashmi, “Role of machine learning and deep learning techniques in eeg-based bci emotion recognition system: A review,” *Artificial Intelligence Review*, vol. 57, no. 50, pp. 1–66, 2024.
- [2] M. J. Aminoff, F. Boller, and D. F. Swaab, *Handbook of Clinical Neurology*. Elsevier, 2020.
- [3] H. Yang, J. Han, and K. Min, “Emotions recognition using eeg signals: A survey,” *IEEE Transactions on Affective Computing*, vol. 10, no. 3, 2019.
- [4] Y. Zhang, X. Ji, and S. Zhang, “An approach to eeg-based emotion recognition using combined feature extraction method,” *Neuroscience Letters*, vol. 633, pp. 152–157, 2016.
- [5] F. Galvão, S. M. Alarcão, and M. J. Fonseca, “Predicting exact valence and arousal values from eeg,” *Sensors*, vol. 21, no. 10, 2021.
- [6] S. Zhang, B. Hu, J. Bian, M. Zhang, and X. Zheng, “A novel emotion recognition method incorporating mst-based brain network and fvmd-gampe,” in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Houston, TX, USA: IEEE, 2021, pp. 1153–1158.
- [7] K. P. Wagh and K. Vasanth, “Performance evaluation of multi-channel electroencephalogram signal (eeg) based time frequency analysis for human emotion recognition,” *Biomedical Signal Processing and Control*, vol. 78, 2022.
- [8] J. Khubani and S. Kulkarni, “Optimized artificial neural network for eeg-based emotion recognition,” in *2022 6th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, 2022, pp. 1624–1633.
- [9] E. Rudakov, L. Laurent, V. Cousin, *et al.*, “Multi-task cnn model for emotion recognition from eeg brain maps,” in *2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART)*, Paris / Créteil, France: IEEE, 2021, pp. 1–4.
- [10] A. Iyer, S. S. Das, R. Teotia, S. Maheshwari, and R. R. Sharma, “Cnn and lstm based ensemble learning for human emotion recognition using eeg recordings,” *Multimodal Interaction and IoT Applications*, vol. 82, pp. 4883–4896, 2023.
- [11] M. K. Chowdary, J. Anitha, and D. J. Hemanth, “Emotion recognition from eeg signals using recurrent neural networks,” *Electronics*, vol. 11, no. 15, 2022.

- [12] H. Yang, J. Han, and K. Min, “A multi-column cnn model for emotion recognition from eeg signals,” *Sensors*, vol. 19, no. 21, 2019.
- [13] M. Pisipati and A. Nandy, “Human emotion recognition using eeg signal in music listening,” in *2021 IEEE 18th India Council International Conference (INDICON)*, Guwahati, India, IEEE, 2021, pp. 1–6.
- [14] P. N. Sơn and L. Minh, *Huấn luyện mạng nơ-ron học sâu*. Nhà xuất bản Đại học Quốc Gia, 2022.
- [15] L. M. Hà, *Giáo trình mạng nơ-ron học sâu và ứng dụng*. Nhà xuất bản Đại học Quốc Gia, 2019.
- [16] H. V. Dũng, *Giáo trình nhận dạng và xử lý ảnh*. Nhà xuất bản Khoa học và Kỹ thuật, 2018.
- [17] S. Koelstra, C. Muhl, M. Soleymani, *et al.*, “Deap: A database for emotion analysis using physiological signals,” *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2011.
- [18] Y. Yang, Q. Wu, Y. Fu, and X. Chen, “Continuous convolutional neural network with 3d input for eeg-based emotion recognition,” in *Neural Information Processing. ICONIP 2018*, S. O. Long Cheng Andrew Chi Sing Leung, Ed., ser. Lecture Notes in Computer Science(), Siem Reap, Cambodia: Springer, Cham, 2018.
- [19] N. Q. Trung, *Xử lý tín hiệu và lọc số*. Nhà xuất bản Khoa học và Kỹ thuật, 2008.
- [20] L.-C. Shi, Y.-Y. Jiao, and B.-L. Lu, “Differential entropy feature for eeg-based vigilance estimation,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan: IEEE, 2013.
- [21] G. G. Berntson, J. T. Cacioppo, and L. G. Tassinary, *Handbook of Psychophysiology*. Cambridge University Press, 2016.