

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

NGÔ BÁ VIỆT

KẾT HỢP TÍN HIỆU EEG, CAMERA VÀ VẬT MỐC
ĐỂ ĐỊNH VỊ, ĐIỀU KHIỂN XE LĂN ĐIỆN ĐẾN ĐÍCH
DỰA VÀO BẢN ĐỒ

LUẬN ÁN TIẾN SĨ
NGÀNH: KỸ THUẬT ĐIỆN TỬ

Tp. Hồ Chí Minh, tháng 1/2024

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

NGÔ BÁ VIỆT

KẾT HỢP TÍN HIỆU EEG, CAMERA VÀ VẬT MỐC
ĐỂ ĐỊNH VỊ, ĐIỀU KHIỂN XE LĂN ĐIỆN ĐẾN ĐÍCH
DỰA VÀO BẢN ĐỒ

NGÀNH: KỸ THUẬT ĐIỆN TỬ - 9520203

Hướng dẫn khoa học: PGS.TS. NGUYỄN THANH HẢI

Phản biện 1:

Phản biện 2:

Phản biện 3:

Tp. Hồ Chí Minh, tháng 1/2024

QUYẾT ĐỊNH GIAO ĐỀ TÀI

BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT
THÀNH PHỐ HỒ CHÍ MINH

Số: 2424/QĐ-ĐHSPKT

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

Tp. Hồ Chí Minh, ngày 03 tháng 10 năm 2016

QUYẾT ĐỊNH

V.v giao đề tài luận án và người hướng dẫn nghiên cứu sinh khóa 2016-2019

HIỆU TRƯỞNG TRƯỜNG ĐẠI HỌC SƯ PHẠM KỸ THUẬT TP. HỒ CHÍ MINH

Căn cứ Quyết định số 118/2000/QĐ-TTg ngày 10 tháng 10 năm 2000 của Thủ tướng Chính phủ về việc thay đổi tổ chức của Đại học Quốc gia TP. Hồ Chí Minh, tách Trường Đại học Sư phạm Kỹ thuật TP. Hồ Chí Minh trực thuộc Bộ Giáo dục và Đào tạo;

Căn cứ Quyết định số 70/2014/QĐ-TTg ngày 10/12/2014 của Thủ tướng Chính phủ về ban hành Điều lệ trường Đại học;

Căn cứ Thông tư số 10/2009/TT-BGDĐT ngày 07/5/2009 của Bộ Giáo dục và Đào tạo về việc Ban hành Quy chế đào tạo trình độ tiến sĩ;

Căn cứ Thông tư số 05/2012/TT-BGDĐT ngày 15/02/2012 của Bộ Giáo dục và Đào tạo về việc sửa đổi, bổ sung một số điều của Quy chế đào tạo trình độ tiến sĩ ban hành kèm theo Thông tư số 10/2009/TT-BGDĐT ngày 07/5/2009 của Bộ trưởng Bộ Giáo dục và Đào tạo;

Xét nhu cầu công tác và khả năng cán bộ;

Xét đề nghị của Trường phòng Đào tạo,

QUYẾT ĐỊNH:

Điều 1. Giao đề tài luận án tiến sĩ và người hướng dẫn cho:

Nghiên cứu sinh : Ngô Bá Việt

Ngành : Kỹ thuật điện tử

Khoá: 2016 – 2019

Tên luận án : Kết hợp tín hiệu EEG, Camera và vật mốc để định vị, điều khiển xe lăn điện đến đích dựa vào bản đồ

Người HD thứ nhất (HD chính) : TS. Nguyễn Thanh Hải

Người HD thứ hai : GS.TS Huỳnh Hữu Tuệ

Thời gian thực hiện : 03/10/2016 đến 03/10/2019

Điều 2. Giao cho Phòng Đào tạo quản lý, thực hiện theo đúng Quy chế đào tạo trình độ tiến sĩ của Bộ Giáo dục & Đào tạo đã ban hành.

Điều 3. Trưởng các đơn vị, phòng Đào tạo, các Khoa quản ngành tiến sĩ và các Ông (Bà) có tên tại Điều 1 chịu trách nhiệm thi hành quyết định này.

Quyết định có hiệu lực kể từ ngày ký. /

Nơi nhận :

- BGH (để biết);
- Như điều 2, 3;
- Lưu: VT, SDH (4b).



LÝ LỊCH CÁ NHÂN

I. THÔNG TIN CÁ NHÂN

Họ và tên: Ngô Bá Việt

Giới tính: Nam

Ngày, tháng, năm sinh: 18 -04-1987

Nơi sinh: Bình Định

Địa chỉ nhà: 04.2 Lô B chung cư Thủ Thiêm Star, Phường Bình Trưng Đông, Thành Phố Thủ Đức, Thành Phố Hồ Chí Minh.

Điện thoại: 0907689357

E-mail: vietnb@hcmute.edu.vn

Cơ quan - nơi làm việc: Trường Đại Học Sư Phạm Kỹ Thuật Tp. Hồ Chí Minh.

Địa chỉ cơ quan: 01 Võ Văn Ngân, Thành Phố Thủ Đức, Thành Phố Hồ Chí Minh. Điện thoại: (+84.8) 37225766; Website: www.hcmute.edu.vn

II. QUÁ TRÌNH ĐÀO TẠO

- Từ 2005 – 2010: Sinh viên ngành Kỹ thuật điện – điện tử, Trường Đại học Sư phạm Kỹ thuật Tp.HCM.

- Từ 2011 – 2013: Học viên cao học ngành Kỹ thuật điện tử, Trường Đại học Sư phạm Kỹ thuật Tp.HCM.

- Từ 2016 – nay: Nghiên cứu sinh ngành Kỹ thuật điện tử, Trường Đại học Sư phạm Kỹ thuật Tp.HCM.

III. QUÁ TRÌNH CÔNG TÁC

Thời gian	Nơi công tác	Công việc đảm nhiệm
03/2007- 10/2010	Công Ty GreyStone Data Systems Việt Nam	Kỹ sư lập trình
08/2011- 01/2019	Trường Cao Đẳng Kỹ Thuật Cao Thắng	Giảng viên
02/2019- nay	Trường Đại Học Sư Phạm Kỹ Thuật Thành Phố Hồ Chí Minh	Giảng viên

Tp. Hồ Chí Minh, ngày 18 tháng 1 năm 2024

NGÔ BÁ VIỆT

LỜI CAM ĐOAN

Tôi cam đoan đây là công trình nghiên cứu của tôi.

Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được công bố trong bất kỳ công trình nào khác

Tp. Hồ Chí Minh, ngày 18 tháng 1 năm 2024

NGÔ BÁ VIỆT

LỜI CẢM ƠN

Đầu tiên, tôi xin gửi lời cảm ơn chân thành và sâu sắc đến Thầy PGS. TS. Nguyễn Thanh Hải, người đã luôn nhiệt tình hướng dẫn, góp ý và định hướng giúp tôi đạt được những kết quả tốt nhất trong suốt quá trình thực hiện luận án.

Tôi xin gửi lời cảm ơn đến Ban Giám Hiệu Trường Đại Học SPKT Tp.HCM, Ban Chủ Nhiệm Khoa Điện – Điện tử, những Thầy/Cô và đồng nghiệp đã hỗ trợ tôi trong quá trình thực hiện luận án này.

Cuối cùng, tôi xin được gửi lời cảm ơn đến gia đình tôi, chỗ dựa tinh thần và là nguồn động viên vô cùng to lớn, giúp tôi có thể thực hiện tốt công việc học tập và nghiên cứu của mình.

Xin chân thành cảm ơn!

Tp. Hồ Chí Minh, ngày 18 tháng 1 năm 2024

NGÔ BÁ VIỆT

TÓM TẮT

Trong những năm gần đây, số lượng người gặp vấn đề về vận động đã tăng lên ở Việt Nam và trên thế giới. Đối với người già và người khuyết tật, các chức năng thể chất của họ bị giảm sút nghiêm trọng, và do đó xe lăn điện gần như là phương tiện duy nhất để duy trì tính cơ động. Luận án này đã đề xuất một số phương pháp cho việc điều khiển bán tự động cho xe lăn điện dựa vào tín hiệu điện não đồ (EEG), camera, các vật mốc và bản đồ. Cụ thể, để điều khiển bán tự động cho xe lăn đến đích, một quá trình gồm 3 giai đoạn được thực hiện, bao gồm: (1) người dùng chọn vị trí điểm đích trên giao diện máy tính bằng tín hiệu EEG; (2) xe lăn tự xác định vị trí bắt đầu trên bản đồ dựa vào các vật mốc trong môi trường; và (3) hệ thống điều khiển xe lăn điện tự động đến đích dựa vào thông tin điểm bắt đầu và đích đến.

Trong môi trường trong nhà, để bắt đầu cho một lộ trình di chuyển tự động, một điểm đích cần phải được lựa chọn. Khi người khuyết tật bị hạn chế về vận động chẳng hạn như không thể cử động tay hoặc đầu, các hoạt động nháy mắt là phù hợp để người dùng ra các lệnh lựa chọn đích đến trên một giao diện máy tính đã được thiết kế trước với các điểm đích. Từ đó, luận án đề xuất hai phương pháp phân loại các hoạt động nháy mắt từ tín hiệu EEG gồm phương pháp ngưỡng biên độ và mô hình học sâu CNN-1D. Ưu điểm của phương pháp phân loại các hoạt động mắt như nháy mắt trái, nháy mắt phải, dùng thuật toán ngưỡng biên độ là độ chính xác cao, trên 97% cho mỗi loại, và có thể xử lý trực tiếp trên tín hiệu mà không cần phải được huấn luyện trước. Với phương pháp phân loại các hoạt động mắt dùng mạng học sâu CNN-1D, tập dữ liệu huấn luyện cần được thu thập trước theo một quy trình được chuẩn hóa. Với đặc điểm của các hoạt động mắt, chỉ có tín hiệu từ 4 điện cực trong tổng số 14 điện cực thu được từ thiết bị Emotiv Epoc+ được trích xuất và ghép lại tạo thành 1 tín hiệu cho huấn luyện. Các tín hiệu này được làm trơn dùng bộ lọc Savitzky-Golay trước khi đưa vào huấn luyện. Kết quả phân loại cho 5 loại nháy mắt gồm nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt hai lần liên tiếp và không nháy mắt có độ chính xác trung bình trên 97%.

TÓM TẮT

Sau khi người dùng đã chọn được điểm đích, vị trí ban đầu của xe lăn trên bản đồ cần được xác định. Với xe lăn điện được trang bị camera, thông tin của các vật mốc tự nhiên bao gồm hình ảnh vật mốc và vị trí của nó trong môi trường, là rất cần thiết cho việc xác định vị trí của xe lăn. Để định vị trên bản đồ dựa vào vật mốc, xe lăn cần thu thập thông tin vật mốc và lưu vào cơ sở dữ liệu, sau đó trong quá trình di chuyển xe lăn sẽ xác định vị trí dựa vào các vật mốc này. Để thu thập thông tin vật mốc trong môi trường, luận án kiến nghị phương pháp mật độ điểm đặc trưng lớn nhất cho nhận biết vật mốc tự nhiên và phương pháp xác định vị trí của vật mốc dựa vào vị trí xe lăn và thông tin 3D từ camera. Cụ thể, hình ảnh môi trường thu thập được từ camera sẽ được trích xuất đặc trưng, và sau đó các thuật toán hình thái học được thực hiện để kết nối các điểm đặc trưng này lại tạo thành các đối tượng trong ảnh. Mật độ điểm đặc trưng cho từng đối tượng này sẽ được tính toán và đối tượng nào có mật độ điểm đặc trưng lớn nhất được chọn là vật mốc. Với vật mốc đã được lựa chọn, thông tin 3D của vật mốc và vị trí của xe lăn trong môi trường được dùng để xác định vị trí vật mốc. Bên cạnh đó, luận án cũng kiến nghị phương pháp xác định vị trí xe lăn trên bản đồ lưới 2D ảo từ thông tin vật mốc trong môi trường thực, nhằm phục vụ cho việc điều khiển xe lăn. Cụ thể, sau khi nhận dạng được các vật mốc có trong cơ sở dữ liệu đã thu thập trước đó, thông tin tọa độ vị trí của các vật mốc này trong không gian môi trường và trong không gian của camera sẽ được dùng để tính toán vị trí xe lăn trong môi trường thực và trên bản đồ 2D lưới ảo bằng các phương trình lượng giác.

Để giảm bớt việc tham gia điều khiển của người sử dụng và tăng sự an toàn, luận án đề xuất mô hình điều khiển thực - ảo cho xe lăn điện dựa vào bản đồ lưới 2D ảo, cho phép người khuyết tật đến bất kỳ điểm đến định sẵn nào trên bản đồ lưới này. Cụ thể, bản đồ lưới 2D ảo được xây dựng từ môi trường thực bằng cách chia thành các ô lưới có thể chứa các khoảng trống hoặc chướng ngại vật. Sau đó, bản đồ với các ô lưới được vẽ này sẽ được mô phỏng để tìm các đường đi tối ưu để đến các vị trí đích bằng mô hình Deep Q-Networks (DQNs) với hàm kích hoạt PreLU. Bên trong mô hình DQNs-PreLU là một mạng nơ-ron truyền thẳng sử dụng phương pháp lan

TÓM TẮT

truyền ngược để cập nhật các thông số mạng. Ngõ vào mô hình là bản đồ lưới và ngõ ra là các hành động của xe lăn tương ứng trên bản đồ bao gồm Lên, Xuống, Trái, Phải. Với các thí nghiệm đã được thực hiện, mô hình DQNs-PreLU cho thấy thời gian huấn luyện ngắn hơn rất nhiều so với các mô hình khác và các thông số mô hình có thể lưu trữ lại để dùng cho điều khiển xe lăn trong môi trường thực. Để có thể điều khiển xe lăn trong môi trường thực, một thuật toán mới được đề xuất để chuyển đổi các hành động của xe lăn từ ngõ ra của mô hình DQNs-PreLU khi mô phỏng với bản đồ lưới 2D ảo thành các lệnh điều khiển thực tế cho xe lăn. Kết quả thí nghiệm đã cho thấy mô hình điều khiển đề xuất có thể tự động điều khiển xe lăn đến đích mong muốn với tính ổn định và an toàn hơn so với khi người dùng tự điều khiển.

Từ những kết quả này, một mô hình xe lăn điện bán tự động cho người khuyết tật trong môi trường trong nhà có thể được ứng dụng trong thực tế. Hiệu quả của các phương pháp đề xuất đã được đánh giá thông qua các kết quả thí nghiệm. Các phương pháp và mô hình được kiến nghị cùng với các kết quả thu được trong luận án này đã được đăng trong các kỷ yếu hội thảo khoa học và tạp chí khoa học. Với mô hình xe lăn điện bán tự động được đề xuất cùng với các kết quả thử nghiệm, xe lăn điện với chi phí thấp có thể được sản xuất trong tương lai gần, hướng đến nhu cầu di chuyển thiết yếu và an toàn cho người khuyết tật.

ABSTRACT

In recent years, the number of people facing mobility issues has increased in Vietnam and worldwide. For the elderly and people with disabilities, their physical functions are severely compromised, and therefore, electric wheelchairs are nearly the sole means to maintain mobility. This thesis has proposed several methods for semi-automatic control of an electric wheelchair based on electroencephalogram (EEG) signals, cameras, landmarks, and maps. Specifically, to semi-automatically control the wheelchair to the destination, a three-stage process is implemented, including (1) users selecting the destination on a computer interface using EEG signals; (2) the wheelchair autonomously determining a starting position on the map based on landmarks in an indoor environment; and (3) the wheelchair system automatically controlling the electric wheelchair to reach the desired destination based on the starting position and destination information.

In an indoor environment, to initiate an automatic travel route, a destination needs to be selected. In the case of disabled people with limited mobility such as being unable to move an electric wheelchair by their hands or head, blinking operations are suitable for moving by selecting commands on a computer interface pre-designed with destinations. Therefore, the thesis proposes two methods of classifying blink activities from EEG signals, including the amplitude thresholding method and the CNN-1D deep learning model. In particular, the advantage of the first method with eye activities such as left-eye blink, and right-eye blink is that the amplitude threshold algorithm has a high accuracy, over 97% for each type, and can allow to directly process on the signals without pre-training required. With the second method using a CNN-1D deep learning network is that the training data set needs to be pre-collected according to a normalized procedure. With the characteristics of eye activities, only the EEG signals obtained from 4 electrodes of the Emotiv EPOC+ system with 14-electrodes are extracted and reassembled to produce one signal for training. Moreover, the EEG signals are smoothed using Savitzky-Golay filters

before training and this can produce the better accuracy. Finally, the classification results for 5 types of blinking activities, including left-eye blink, right-eye blink, two-eye blink, double two-eye blink, and neutral have an accuracy of over 97%.

After the user has selected a desired destination on the map, the wheelchair's initial position needs to be determined. With an electric wheelchair equipped with a camera, the information of natural landmarks, including the image of the landmark and its position in the environment, are essential for determining the position of the wheelchair. To determine the location based on these landmarks, the wheelchair needs to collect landmark information and then stores them in a database during the movement. Therefore, to collect information of landmarks in an indoor environment, the thesis proposes the methods of the maximum feature density for the identification of natural landmarks and determining the positions of landmarks based on the wheelchair's position and 3D information from the camera. In particular, environmental images collected from the camera will be extracted features, and then morphological algorithms are performed to connect these feature points for creating objects in the image. The feature density for each of these objects will be calculated and the object with the highest feature density is selected as the landmark. With the landmark selected, the 3D information of that landmark and the wheelchair's position in the environment are used for locating the landmark. In addition, the thesis proposes the method for locating the wheelchair on a virtual 2D grid map from landmark information in the real environment for controlling the mobile wheelchair. In particular, after identifying landmarks in the collected database, the coordinate of these landmarks in the environmental and camera spaces will be used to calculate the wheelchair position in the real environment and the virtual 2D grid map using the trigonometric equations.

To reduce the user's participation in control and increase safety, the thesis proposes a virtual-real control model for electric wheelchairs based on a virtual 2D grid map, allowing people with disabilities to reach any destination pre-designed on this grid map. In particular, the virtual 2D grid map is built from the real environment

by dividing it into grid cells that may contain free spaces or obstacles. This map with the plotted grid cells is then simulated for finding the optimal paths to reach the desired destination using a Deep Q-Networks (DQNs) model with the PreLU activation function. Therefore, the DQNs-PreLU model is designed with a feedforward neural network using backpropagation for updating the network parameters. The network algorithm is that the input is the grid map and the output is the wheelchair actions including Up, Down, Left, and Right. With the experiments performed, the DQNs-PreLU model shows that the training time is much shorter than other models and the model parameters can be stored for controlling the wheelchair in the real environment. In addition, to control the wheelchair in the real environment better, a new algorithm was proposed to convert the wheelchair actions from the output of the DQNs-PreLU model after simulated with the virtual 2D grid map into real control commands. Experimental results show that the proposed control model can allow to automatically control the wheelchair to reach the desired destination with more stability and safety compared to the wheelchair controlled by user.

From these results, the semi-automatic electric wheelchair model for the disabled people can be applied in practice. In addition, the effectiveness of the proposed methods has been evaluated through experimental results. In particular, these proposed methods with the results in this thesis have been published in the proceedings of scientific conferences and scientific journals. With the proposed semi-autonomous electric wheelchair model along with the test results, a low-cost electric wheelchair with the safe mobility can be produced to support the disabled people.

MỤC LỤC

QUYẾT ĐỊNH GIAO ĐỀ TÀI.....	iii
LÝ LỊCH CÁ NHÂN	iv
LỜI CAM ĐOAN	v
LỜI CẢM ƠN	vi
TÓM TẮT	vii
ABSTRACT	x
MỤC LỤC.....	xiii
DANH SÁCH CÁC TỪ VIẾT TẮT	xvii
DANH SÁCH CÁC HÌNH.....	xix
DANH SÁCH CÁC BẢNG.....	xxiv
CHƯƠNG 1.....	1
TỔNG QUAN	1
1.1. TÍNH CẤP THIẾT CỦA ĐỀ TÀI LUẬN ÁN	1
1.2. MỤC TIÊU NGHIÊN CỨU CỦA LUẬN ÁN	10
1.3. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU.....	10
1.4. CÁCH TIẾP CẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU.....	11
1.5. ĐÓNG GÓP VỀ KHOA HỌC CỦA LUẬN ÁN	12
1.6. CẤU TRÚC CỦA LUẬN ÁN	14
CHƯƠNG 2.....	16
CƠ SỞ LÝ THUYẾT	16
2.1. TỔNG QUAN VỀ TÍN HIỆU EEG.....	16
2.2. PHÂN LOẠI HOẠT ĐỘNG DỰA VÀO TÍN HIỆU EEG.....	17
2.2.1. Phát hiện hoạt động mắt dựa vào ngưỡng biên độ tín hiệu EEG	17
2.2.2. Phân loại tín hiệu EEG dùng mạng nơ-ron.....	18
2.2.3. Phân loại tín hiệu EEG dùng mạng nơ-ron tích chập	19
2.3. GIAO TIẾP GIỮA NÃO NGƯỜI VÀ MÁY TÍNH	20
2.4. MÔ HÌNH XE LĂN ĐIỆN CHO NGƯỜI KHUYẾT TẬT	21
2.4.1. Xe lăn điện thông minh	21
2.4.2. Xe lăn điện với bộ điều khiển robot	22
2.4.3. Xe lăn điện tích hợp với môi trường thông minh	23

MỤC LỤC

2.4.4.	Xe lăn điện với tính năng tránh chướng ngại vật	24
2.4.5.	Hệ thống chia sẻ điều khiển cho xe lăn điện	25
2.5.	PHƯƠNG PHÁP XÂY DỰNG BẢN ĐỒ LƯỚI 2D ĐỂ ĐIỀU HƯỚNG ROBOT TRONG NHÀ.....	27
2.6.	CÁC PHƯƠNG PHÁP ĐỊNH VỊ CHO ROBOT DI ĐỘNG	29
2.6.1.	Phương pháp đoán định vị trí	29
2.6.2.	Phương pháp định vị robot dùng vật mốc	30
2.6.3.	Phương pháp định vị cho robot dùng hệ thống WIFI.....	30
2.7.	CÁC PHƯƠNG PHÁP NHẬN DẠNG VẬT THỂ	31
2.7.1.	Phương pháp nhận dạng dựa trên diện mạo	31
2.7.2.	Phương pháp nhận dạng dựa trên các điểm đặc trưng.....	32
2.7.3.	Nhận dạng vật thể theo phương pháp máy học	32
2.8.	MÔ HÌNH HÓA VÀ ĐIỀU KHIỂN XE LĂN ĐIỆN	33
2.9.	PHƯƠNG PHÁP LẬP KẾ HOẠCH ĐƯỜNG ĐI CHO ROBOT DI ĐỘNG.....	35
2.9.1.	Phương pháp A*	35
2.9.2.	Phương pháp học tăng cường	36
CHƯƠNG 3.....	39	
PHÂN LOẠI TÍN HIỆU EEG CỦA HOẠT ĐỘNG MẮT CHO ỨNG DỤNG GIAO TIẾP GIỮA NÃO NGƯỜI VÀ MÁY TÍNH.....	39	
3.1.	TÍN HIỆU EEG CỦA HOẠT ĐỘNG NHÁY MẮT	39
3.2.	THU THẬP DỮ LIỆU	42
3.2.1.	Quy trình thu thập dữ liệu.....	42
3.2.2.	Tập dữ liệu gốc của tín hiệu EEG.....	44
3.3.	XỬ LÝ TÍN HIỆU	48
3.3.1.	Lọc nhiễu dùng bộ lọc Hamming	48
3.3.2.	Làm trơn tín hiệu dùng bộ lọc Savitzky – Golay	49
3.4.	PHÂN LOẠI TÍN HIỆU EEG CỦA HOẠT ĐỘNG MẮT	52
3.4.1.	Phân loại hoạt động nháy mắt theo phương pháp ngưỡng biên độ	52
3.4.2.	Phân loại tín hiệu hoạt động mắt dùng mô hình CNN-1D	56
CHƯƠNG 4.....	74	

NHẬN DẠNG VÀ XÁC ĐỊNH VỊ TRÍ VẬT MỐC TỰ NHIÊN TRONG MÔI TRƯỜNG TRONG NHÀ	74
4.1. VAI TRÒ CỦA VẬT MỐC ĐỐI VỚI QUÁ TRÌNH ĐỊNH VỊ VÀ ĐIỀU KHIỂN XE LĂN ĐIỆN	74
4.2. PHƯƠNG PHÁP MẬT ĐỘ ĐIỂM ĐẶC TRƯNG LỚN NHẤT CHO NHẬN DẠNG CÁC VẬT MỐC TRONG TỰ NHIÊN.....	76
4.2.1. Phát hiện các điểm đặc trưng.....	77
4.2.2. Kết nối các điểm đặc trưng trong các đối tượng	79
4.2.3. Nhận biết vật mốc tự nhiên.....	79
4.3. XÁC ĐỊNH VỊ TRÍ VẬT MỐC TRONG MÔI TRƯỜNG.....	80
4.3.1. Vị trí của xe lăn trong môi trường	80
4.3.2. Xác định vị trí vật mốc	81
4.4. KẾT QUẢ NHẬN BIẾT VÀ THU THẬP VẬT MỐC TỰ NHIÊN.....	83
4.5. KẾT QUẢ XÁC ĐỊNH VỊ TRÍ VẬT MỐC.....	93
CHƯƠNG 5.....	98
MÔ HÌNH ĐIỀU KHIỂN XE LĂN ĐIỆN KẾT HỢP TÍN HIỆU EEG VÀ CAMERA DỰA VÀO BẢN ĐỒ	98
5.1. HỆ THỐNG ĐIỀU KHIỂN THỰC - ẢO CHO XE LĂN ĐIỆN DỰA VÀO BẢN ĐỒ LƯỚI 2D ẢO	98
5.1.1. Cấu trúc hệ thống điều khiển thực -ảo cho xe lăn điện	98
5.1.2. Bản đồ lưới 2D ảo.....	100
5.1.3. Giao diện lựa chọn đích đến cho xe lăn.....	102
5.1.4. Mô hình DQNs lập kế hoạch đường đi tối ưu cho xe lăn.....	103
5.2. ĐỊNH VỊ XE LĂN ĐIỆN TRONG BẢN ĐỒ LƯỚI 2D ẢO DỰA VÀO VẬT MỐC TỰ NHIÊN.....	107
5.3. ĐIỀU HƯỚNG XE LĂN ĐIỆN TRONG MÔI TRƯỜNG THỰC.....	109
5.4. PHƯƠNG PHÁP TRÁNH VẬT CẢN DỰA VÀO THÔNG TIN 3D MÔI TRƯỜNG.....	112
5.5. KẾT QUẢ THÍ NGHIỆM ĐIỀU KHIỂN XE LĂN ĐIỆN	113
5.5.1. Mô phỏng huấn luyện tìm đường đi cho xe lăn dựa vào bản đồ lưới 2D ảo	113
5.5.2. Kết quả nhận dạng vật mốc	119

MỤC LỤC

5.5.3. Xác định vị trí xe lăn trên bản đồ lưới 2D ảo dựa vào vật mốc	122
5.5.4. Xe lăn di chuyển đến đích mong muốn dựa vào bản đồ lưới 2D ảo	125
CHƯƠNG 6.....	138
KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	138
6.1. KẾT LUẬN.....	138
6.2. HƯỚNG PHÁT TRIỂN.....	141
TÀI LIỆU THAM KHẢO	142
CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ	154

DANH SÁCH CÁC TỪ VIẾT TẮT

ACC	Accuracy	Độ chính xác
APF	Angle-based Potential Field	Vùng góc tiềm năng
ALS	Amyotrophic Lateral Sclerosis	Bệnh xơ cứng teo cơ bên
BCI	Brain – Computer Interface	Giao tiếp não người và máy tính
BRISK	Binary Robust Invariant Scalable Keypoints	Các điểm đặc trưng có thể mở rộng bất biến bền vững bằng nhị phân
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
DQNs	Deep Q-networks	Mạng Q sâu.
DWA	Dynamic Window Approach	Phương pháp cửa sổ động
D-DWA	Double Dynamic Window Approach	Phương pháp cửa sổ động kép
DSS	Drive-Safe System	Hệ thống lái xe an toàn
DRL	Deep Reinforcement Learning	Học tăng cường sâu
EEG	Electroencephalography	Điện não đồ
EMG	Electromyography	Điện cơ
ECG	Electrocardiogram	Điện tâm đồ
FAST	Features from Accelerated Segment Test	Các đặc trưng từ kiểm tra phân đoạn tăng tốc
GPS	Global Positioning System	Hệ thống định vị toàn cầu
HOG	Histogram of Oriented Gradients	Biểu đồ Gradient định hướng
HHT	Hilbert–Huang Transform	Biến đổi Hilbert–Huang
IOU	Intersection over Union	Vùng giao nhau tại những liên kết

DANH SÁCH CÁC TỪ VIẾT TẮT

IMU	Inertial Measurement Unit	Đơn vị đo lường quán tính
LIDAR	Light Detection and Ranging	Đo khoảng cách bằng ánh sáng
NN	Neural Network	Mạng nơ-ron
PCA	Principal Component Analysis	Phân tích thành phần chính
PRE	Precision	Độ chính xác
PWM	Pulse Width Modulation	Điều chế độ rộng xung
RGB	Red – Green - Blue	Đỏ - Lục - Lam
RGB-D	Red – Green - Blue - Depth	Đỏ - Lục – Lam - Độ sâu
RFID	Radio Frequency Identification	Nhận dạng tần số vô tuyến điện
RL	Reinforcement Learning	Học tăng cường
ROC	Receiver Operating Characteristic	Đặc tính hoạt động của bộ thu
SEN	Sensitivity	Độ nhạy
SLAM	Simultaneous Localization and Mapping	Bản đồ hóa và bản địa hóa đồng thời
SURF	Speeded Up Robust Features	Tăng tốc đặc tính bền vững
SIFT	Scale-Invariant Feature Transform	Biến đổi đặc trưng bất biến theo tỷ lệ
SSVEP	Steady-State Visual Evoked Potential	Tiềm năng kích thích trực quan ở trạng thái ổn định
VFH	Vector Field Histogram	Biểu đồ trường vectơ
WHO	World Health Organization	Tổ chức Y tế Thế giới
3D	Three - Dimensional	3 chiều
2D	Two - Dimensional	2 chiều
1D	One - Dimensional	1 chiều

DANH SÁCH CÁC HÌNH

Hình	Trang
Hình 1.1. Sơ đồ khối quá trình điều khiển xe lăn điện dựa vào tín hiệu EEG, camera, vật mốc và bản đồ	9
Hình 2.1. Tín hiệu EEG được phát hiện với các vùng chuyển động của mắt	18
Hình 2.2. Sơ đồ điều khiển kết hợp của BCI và xe lăn	22
Hình 2.3. Sơ đồ bộ điều khiển xe lăn di động kết hợp cánh tay robot... ..	23
Hình 2.4. Mô hình hệ thống xe lăn đa phương thức	24
Hình 2.5. Hệ thống xe lăn điện với chức năng định vị và tránh vật cản	24
Hình 2.6. Hệ thống chia sẻ điều khiển cho xe lăn điện	25
Hình 2.7. Mô hình chia sẻ điều khiển xe lăn điện sử dụng BCI và cảm biến	27
Hình 2.8. Biểu diễn chùm sonar trên bản đồ lưới 2D	28
Hình 2.9. Phương pháp xây dựng bản đồ lưới kết hợp ngữ nghĩa đối tượng	29
Hình 2.10. Định vị robot dùng các vật mốc trên trần nhà... ..	30
Hình 2.11. Sử dụng Wifi để định vị vị trí robot	31
Hình 2.12. Hệ quy chiếu toàn cục và hệ quy chiếu cục bộ của xe lăn hoặc robot ...	34
Hình 3.1. Một số tín hiệu cơ bản của chuyển động mắt	40
Hình 3.2. Tín hiệu EEG khi nháy mắt	41
Hình 3.3: Hai loại tín hiệu EEG của hoạt động nháy mắt	42
Hình 3.4. Thiết bị Emotiv Epoch+	43
Hình 3.5. Hướng dẫn đeo thiết bị và kiểm tra kết nối	43
Hình 3.6. Quy trình thực hiện thí nghiệm... ..	44
Hình 3.7. Tín hiệu gốc thu được từ 4 kênh của hoạt động nháy mắt trái	45
Hình 3.8. Tín hiệu gốc thu được từ 4 kênh của hoạt động nháy mắt phải	46
Hình 3.9. Tín hiệu gốc thu được từ 4 kênh của hoạt động nháy hai mắt	46
Hình 3.10. Tín hiệu gốc thu được từ 4 kênh của hoạt động nháy hai mắt hai lần liên tiếp	47

DANH SÁCH CÁC HÌNH

Hình 3.11. Tín hiệu gốc thu được từ 4 kênh của hoạt động không nháy mắt	47
Hình 3.12. Tín hiệu EEG ở kênh F7 của hoạt động nháy mắt trái trước và sau khi lọc dùng bộ lọc Hamming	49
Hình 3.13. Biểu diễn tín hiệu EEG tại kênh F7 trước và sau khi làm trơn bằng bộ lọc Savitzky-Golay	51
Hình 3.14. Tín hiệu EEG của hoạt động nháy hai mắt tại kênh F7 được lọc Hamming và làm trơn với bộ lọc Savitzky-Golay	51
Hình 3.15. Mô tả cách thức chia tín hiệu EEG thành các khung dữ liệu	52
Hình 3.16. Mô tả tín hiệu nháy mắt tự nguyện.....	53
Hình 3.17. Kết quả nhận dạng hoạt động mắt	55
Hình 3.18. Tín hiệu nháy mắt trái thu được từ 4 kênh AF3, F7, AF4, F8	57
Hình 3.19. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy mắt trái	58
Hình 3.20. Tín hiệu nháy mắt phải thu được từ 4 kênh AF3, F7, AF4, F8	58
Hình 3.21. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy mắt phải	59
Hình 3.22. Dạng tín hiệu nháy hai mắt được thu ở 4 kênh AF3, F7, AF4, F8.....	59
Hình 3.23. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy hai mắt.....	60
Hình 3.24. Dạng dữ liệu chớp hai mắt hai lần liên tiếp được thu ở 4 kênh AF3, F7, AF4, F8	60
Hình 3.25. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy hai mắt hai lần liên tiếp	61
Hình 3.26. Dạng dữ liệu không nháy mắt được thu ở 4 kênh AF3, F7, AF4, F8	62
Hình 3.27. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động không nháy mắt	62
Hình 3.28. Mô hình CNN-1D cho phân loại tín hiệu EEG của hoạt động mắt	63
Hình 3.29. Ma trận nhầm lẫn.....	65

DANH SÁCH CÁC HÌNH

Hình 3.30. Tổ chức dữ liệu để huấn luyện mô hình CNN-1D	66
Hình 3.31. Đồ thị biểu diễn hiệu suất huấn luyện mô hình CNN-1D phân loại hoạt động mắt	67
Hình 3.32. Kết quả phân loại các hoạt động mắt của mô hình CNN-1D trên tập kiểm tra	69
Hình 3.33. Kết quả phân loại các hoạt động mắt với hai loại nháy mắt trái và phải dùng tín hiệu không ghép kênh	71
Hình 4.1. Các bước xác định vật mốc tự nhiên	76
Hình 4.2. Mô hình xe lăn di động hai bánh vi sai và hai bánh tự do.....	80
Hình 4.3. Ước tính vị trí mốc trong không gian 2D	82
Hình 4.4. Xe lăn với camera RGB-D, Encoder và máy tính	83
Hình 4.5. Các đối tượng chứa các điểm đặc trưng	84
Hình 4.6. Biểu diễn sự giãn nở với các lần lặp khác nhau bằng cách sử dụng mặt nạ 3×3	85
Hình 4.7. Ảnh chứa đường bao đối tượng và khung đặc trưng	85
Hình 4.8. Quá trình nhận dạng vật mốc trong phòng thí nghiệm ở khoảng cách 2 m từ camera đến các đối tượng	86
Hình 4.9. Quá trình nhận dạng vật mốc trong phòng thí nghiệm ở khoảng cách 1 m từ camera đến các đối tượng	87
Hình 4.10. IOU cho các hộp giới hạn.....	87
Hình 4.11. Nhận biết mốc tự nhiên trong môi trường phòng thí nghiệm với các điều kiện ánh sáng khác nhau	89
Hình 4.12. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 0°	90
Hình 4.13. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 45°	90
Hình 4.14. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 30°	91
Hình 4.15. Các vật mốc tự nhiên được phát hiện từ các khu vực khác nhau	91

DANH SÁCH CÁC HÌNH

Hình 4.16. Biểu đồ sai số tương đối của phép đo khoảng cách tới các vật mốc được phát hiện	93
Hình 4.17. Biểu diễn quỹ đạo chuyển động của xe lăn đến các vị trí đặt trước.....	94
Hình 5.1. Hệ thống điều khiển thực - ảo cho xe lăn điện dựa trên bản đồ lưới 2D ảo	99
Hình 5.2. Quá trình tạo bản đồ lưới 2D ảo từ môi trường thực.....	101
Hình 5.3. Bản đồ lưới 2D ảo của môi trường thực.....	101
Hình 5.4. Quy trình thực hiện lựa chọn đích đến dùng BCI	102
Hình 5.5. Giao diện người dùng để chọn điểm đến mong muốn	103
Hình 5.6. Giao diện người dùng chọn điểm đến mong muốn “Phòng ngủ” bằng lệnh EEG	103
Hình 5.7. Cấu trúc mạng nơ-ron trong mô hình DQNs.....	106
Hình 5.8. Các mẫu vật mốc trong môi trường trong nhà	107
Hình 5.9. Hệ tọa độ của camera RGB-D	107
Hình 5.10. Vị trí của xe lăn với bốn hướng.....	109
Hình 5.11. Tọa độ của xe lăn, vật mốc và đích đến trong bản đồ lưới 2D ảo	110
Hình 5.12. Biểu diễn bộ chuyển đổi các lệnh điều khiển thực tế từ mô phỏng	110
Hình 5.13. Ảnh độ sâu được chụp từ camera RGB-D.....	112
Hình 5.14. Mô tả chuyển đổi bản đồ điểm 3D sang 2D với độ cao được định trước	113
Hình 5.15. Môi trường huấn luyện	114
Hình 5.16. So sánh tỷ lệ Win khi huấn luyện mô hình DQNs với 2 loại hàm kích hoạt trong trường hợp bản đồ lưới 8x11	115
Hình 5.17. So sánh tỷ lệ Win khi huấn luyện mô hình DQNs với hai loại hàm kích hoạt trong trường hợp bản đồ lưới 11x33	116
Hình 5.18. Các thành phần vectơ mô tả của vùng con trong ảnh.....	119
Hình 5.19. Bốn loại vật mốc khác nhau được nhận dạng dựa trên các vật mốc được lưu trữ trong cơ sở dữ liệu.....	120

DANH SÁCH CÁC HÌNH

Hình 5.20. Sai số tuyệt đối của phép đo X_a từ camera đến các vật mốc tại các vị trí khác nhau	122
Hình 5.21. Sai số tuyệt đối của khoảng cách Z_a từ camera đến các vật mốc tại các vị trí khác nhau	122
Hình 5.22. Môi trường thí nghiệm trong nhà	123
Hình 5.23. Bốn vị trí của xe lăn trên bản đồ lưới 2D với các vật mốc.....	124
Hình 5.24. Vị trí xe lăn trên bản đồ lưới 2D với 3 vật mốc	125
Hình 5.25. Hệ thống xe lăn với các camera RGB-D, Emotiv EPOC+, bộ điều khiển và máy tính	126
Hình 5.26. Môi trường thí nghiệm	127
Hình 5.27. Đường di chuyển của xe lăn trong môi trường thực tế	128
Hình 5.28. Quy trình điều hướng xe lăn điện ở chế độ bán tự động.....	129
Hình 5.29. Biểu diễn đường đi mô phỏng và đường đi thực của xe lăn bằng điều khiển bán tự động.....	130
Hình 5.30. So sánh chuyển động của xe lăn theo hai phương pháp điều khiển (điều khiển bán tự động và tự điều khiển).....	132
Hình 5.31. Môi trường trong thí nghiệm thứ hai.....	133
Hình 5.32. So sánh chuyển động của xe lăn trong hai phương pháp điều khiển (điều khiển bán tự động và tự điều khiển).....	134
Hình 5.33. Chuyển động của xe lăn khi có vật cản.....	135

DANH SÁCH CÁC BẢNG

Bảng	Trang
Bảng 2.1. Các dạng sóng EEG	17
Bảng 3.1. Các trường hợp phân loại hoạt động mắt	55
Bảng 3.2. Tên và nhãn tín hiệu của các hoạt động mắt.....	63
Bảng 3.3. Mô tả dữ liệu huấn luyện phân loại hoạt động mắt	66
Bảng 3.4. Hiệu suất mô hình CNN-1D khi huấn luyện phân loại các hoạt động mắt	68
Bảng 3.5. Xác thực chéo 5 lần đối với bộ phân loại hoạt động nháy mắt	70
Bảng 3.6. Các nghiên cứu phân loại hoạt động mắt	72
Bảng 3.7. Thời gian xử lý phân loại tín hiệu EEG của hoạt động mắt.....	73
Bảng 4.1. Kết quả trích đặc trưng dùng các phương pháp khác nhau	84
Bảng 4.2. IOU với các khoảng cách khác nhau	88
Bảng 4.3. IOU với điều kiện ánh sáng khác nhau	89
Bảng 4.4. Thời gian xử lý nhận dạng vật mốc của phương pháp đề xuất	92
Bảng 4.5. Đánh giá sai số vị trí của xe lăn	95
Bảng 4.6. Kết quả định vị vật mốc dựa vào vị trí xe lăn	95
Bảng 4.7. Các nghiên cứu nhận dạng, thu thập thông tin vị trí đối tượng tự động	96
Bảng 5.1. Thông số mô hình huấn luyện	114
Bảng 5.2. Hiệu suất của các mô hình DQNs được đề xuất	117
Bảng 5.3. So sánh hiệu suất của các mô hình	117
Bảng 5.4. Độ chính xác khi nhận dạng các vật mốc bằng phương pháp SURF ...	121
Bảng 5.5. Độ chính xác của việc định vị xe lăn	124
Bảng 5.6. Độ chính xác khi định vị xe lăn bằng 3 vật mốc	125
Bảng 5.7. Các giá trị và trạng thái điều khiển xe lăn điện	126
Bảng 5.8. Các lệnh điều khiển xe lăn được chuyển đổi từ các lệnh mô phỏng	131
Bảng 5.9. Sai số đường đi của xe lăn trong các môi trường thí nghiệm	135

CHƯƠNG 1

TỔNG QUAN

1.1. TÍNH CẤP THIẾT CỦA ĐỀ TÀI LUẬN ÁN

Trong xã hội hiện nay, người khuyết tật luôn gặp phải những khó khăn và thiệt thòi cả về thể chất và tinh thần hơn so với những người bình thường khác. Những khiếm khuyết trên cơ thể tạo ra những suy giảm đáng kể và ảnh hưởng lâu dài, trực tiếp đến khả năng tư duy, hoạt động, sinh hoạt hàng ngày của họ, gây nên những khó khăn nhất định trong cuộc sống. Thống kê của Tổ chức y tế thế giới (WHO) năm 2022 cho thấy 16% dân số thế giới tương đương 1,3 tỷ người bị khuyết tật và ngày càng tăng lên [1]. Theo Tổng cục Thống kê năm 2019, Việt Nam là quốc gia có số lượng người khuyết tật khá lớn so với tổng dân số trong khu vực châu Á-Thái Bình Dương, có khoảng 6,2 triệu người khuyết tật, chiếm 7,06% dân số từ 2 tuổi trở lên, trong đó có 58% là nữ, 28,3% là trẻ em, gần 29% là người khuyết tật nặng và đặc biệt nặng [2]. Ngoài ra, thống kê cũng cho thấy tỷ lệ người khuyết tật vận động là 29,41%, khuyết tật nghe nói 9,32%; khuyết tật nhìn 13,84%, khuyết tật thần kinh và tâm thần 16,83%, khuyết tật trí tuệ 6,52% và khuyết tật khác 24,08%. Người khuyết tật gặp rất nhiều khó khăn trong cuộc sống, đặc biệt đối với người khuyết tật về vận động. Khoảng 7% trong số những người khuyết tật về vận động trên thế giới cần dùng đến xe lăn [3]. Thị trường xe lăn điện đã đạt doanh thu 2,89 tỷ USD vào năm 2021 và dự kiến đạt 5,27 tỷ USD. Thị trường xe lăn điện dự kiến sẽ tăng trưởng gần 10,76% trong giai đoạn 2022-2027 [4].

Vận hành xe lăn truyền thống là một công việc đầy khó khăn đối với người khuyết tật. Do đó, việc sử dụng xe lăn điện để mang lại khả năng cơ động dễ dàng là cần thiết. Xe lăn điện được trang bị các thiết bị công nghệ cao có thể giúp cho người

khuyết tật di chuyển mà không cần sự trợ giúp. Đơn giản nhất là xe lăn điện được trang bị một cần điều khiển mà người sử dụng có thể điều khiển hướng di chuyển của xe lăn theo ý muốn của mình [5]. Trong những năm gần đây, nhiều phương pháp đã được giới thiệu để phát triển hệ thống xe lăn thông minh để phù hợp với người khuyết tật. Xu hướng phát triển có thể được phân loại thành ba lĩnh vực chính: 1) Cải tiến công nghệ hỗ trợ [6], 2) Cải tiến giao diện vật lý của người dùng [7], 3) Cải thiện điều khiển chia sẻ giữa người dùng và máy [8]. Một trong những vấn đề quan trọng của xe lăn thông minh là cung cấp khả năng di chuyển độc lập cho người tàn tật nặng, những người không thể điều khiển xe lăn bằng cần điều khiển tiêu chuẩn. Do đó, việc phục hồi khả năng vận động của họ có thể cải thiện đáng kể chất lượng cuộc sống.

Xe lăn thông minh được phát triển phụ thuộc nhiều vào tình trạng sức khỏe người dùng, tức là khả năng và khuyết tật, và không có giải pháp duy nhất phù hợp cho tất cả người dùng. Bệnh nhân bị suy yếu vận động thường thiếu kiểm soát cơ bắp và trong trường hợp xấu nhất họ không thể cử động cánh tay và chân. Để hỗ trợ khả năng di chuyển của những bệnh nhân này, tín hiệu hoặc hành động từ giọng nói, lưỡi, có thể tạo ra các lệnh điều khiển [9, 10]. Điều hướng bằng giọng nói đòi hỏi môi trường yên tĩnh và có thể không tốt để sử dụng trong môi trường ồn ào. Hơn nữa, giải pháp này đôi khi được coi là bất lịch sự khi nói to trong một khu vực yên tĩnh. Jin Sun Ju và các cộng sự đã sử dụng một camera nhận dạng các cử chỉ trên khuôn mặt của người sử dụng để điều khiển xe lăn đi thẳng, đi lùi, rẽ trái, rẽ phải [11]. Trong nghiên cứu của Y. Zhang và các cộng sự [12], một camera nhận dạng các cử chỉ của bàn tay để điều khiển hướng đi cho xe lăn điện. Sadi [13] cùng với Mhaske và Chhaya.G.Patil [14] dùng camera nhận dạng số ngón tay được giơ lên để thực hiện các lệnh điều khiển cho xe lăn. Một cách điều khiển xe lăn khác sử dụng cử chỉ của người sử dụng là nhận dạng sự di chuyển của mắt, với độ chính xác phân loại là 99,3% và có thể đưa ra dự đoán trong khoảng 1,57 ms [15]. Mặc dù các nghiên cứu dùng camera để phát hiện các cử chỉ của người sử dụng có thời gian xử lý nhanh và độ chính xác phân loại cao, nhưng chỉ phù hợp với một điều kiện ánh sáng chuẩn. Độ

chính xác sẽ giảm đáng kể khi bị quá sáng hoặc quá tối, hoặc vị trí của khuôn mặt, mắt, bàn tay bị lệch khỏi vùng ghi hình của camera.

Đối với việc sử dụng EEG trong những năm gần đây đã trở thành một chủ đề được quan tâm để điều khiển máy móc cho những đối tượng tàn tật nặng không thể dùng giọng nói hay vận động các chi hoặc đầu để ra lệnh điều khiển [16]. Để làm như vậy, các mẫu tín hiệu EEG cần được phân loại và nhóm thành các hành động dự định. Xe lăn điều khiển bằng EEG là thiết bị thích hợp cho những bệnh nhân bị liệt hoàn toàn với bộ não khỏe mạnh để điều hướng môi trường của họ [17]. Người sử dụng sẽ được gắn các điện cực lên đầu để thu thập các tín hiệu EEG và truyền về máy tính. Tín hiệu EEG thu được là dạng tín hiệu ngẫu nhiên khá phức tạp, do đó cần sử dụng các bộ lọc và các thuật toán để trích các đặc trưng của tín hiệu, cho biết tín hiệu nào liên quan đến hoạt động của cơ thể. Nhóm nghiên cứu của Nguyễn Thanh Hải [18, 19] thu thập tín hiệu EEG từ hoạt động của mắt (mở mắt, chớp mắt, liếc trái, liếc phải) và sử dụng bộ lọc Hamming để chia tín hiệu thành các dải tần khác nhau cho từng hoạt động. Sau đó, thông qua mạng nơ-ron để huấn luyện các tín hiệu thành bốn lệnh điều khiển cho xe lăn như đi tới, đi lùi, rẽ trái, rẽ phải với tốc độ di chuyển 5km/h trong môi trường trong nhà. Bên cạnh đó, các nghiên cứu về việc sử dụng tín hiệu EEG để điều khiển xe lăn cho người khuyết tật cũng được quan tâm phát triển với mục đích điều khiển thuận tiện cho người sử dụng với chi phí đầu tư thấp [20]. Tuy nhiên, khi sử dụng tín hiệu EEG cho điều khiển xe lăn, người dùng phải có sự kiểm soát cảm xúc và sự tập trung tốt để điều khiển hiệu quả. Đây là một gánh nặng cho người dùng mặc dù phương pháp này có thể là một lựa chọn tốt cho những người có cơ thể bị tê liệt hoàn toàn. Có lẽ giải pháp tốt nhất cho trường hợp này là dựa vào các phương pháp sử dụng nhiều tín hiệu từ người dùng và môi trường xung quanh để phân tích trước khi đưa ra lệnh mong muốn [21]. Sử dụng chiến lược này, sẽ áp đặt ít gánh nặng hơn cho người dùng so với trường hợp chỉ dựa vào một đầu vào.

Nhằm giảm bớt sự điều khiển của người dùng và đảm bảo an toàn, xe lăn được điều khiển tự động hoặc bán tự động là một giải pháp cần thiết. Chế độ tự động của xe lăn đồng nghĩa với việc xe lăn phải tự di chuyển đến điểm đích mà người sử dụng

mong muốn, đồng thời phải tránh được những chướng ngại vật trên đường đi. Trong những năm gần đây, đã có nhiều phương pháp được áp dụng để xe lăn tự hành như di chuyển theo đường đi đã được vạch sẵn [22], di chuyển theo tọa độ sử dụng định vị GPS và bộ lọc Kalman [23], di chuyển theo bản đồ [24]. Bên cạnh đó, xe lăn thông minh cũng đã được cải thiện về các bộ phận điều hướng, chẳng hạn như tự động tránh chướng ngại vật, phương thức giao tiếp giữa người dùng và hệ thống điều khiển và sự thích ứng với mức độ khuyết tật [25]. Trong thập kỷ qua, một số thuật toán điều hướng đã được nghiên cứu cho các xe lăn điện thông minh và hầu hết trong số các nghiên cứu này sử dụng các cảm biến khác nhau để phát hiện và tránh chướng ngại vật. Xe lăn được phát triển cho người cao tuổi và có thể nhận ra các chướng ngại vật khác nhau bằng cách sử dụng cảm biến siêu âm [26]. Từ đó, phương pháp kết hợp các cảm biến sẽ hỗ trợ người dùng điều hướng trong các không gian hẹp. Xe lăn điện được thiết kế có khả năng tránh chướng ngại vật và tự động đi đến các điểm đích được chọn bởi người dùng [27, 28]. Xe lăn này sử dụng một hệ thống điều khiển an toàn DSS đã được triển khai để hỗ trợ cho người dùng khiếm thị và cho phép xe lăn di chuyển theo một bức tường và vượt qua các cửa bên trong tòa nhà. Malek Njah [29] đã sử dụng đồng thời bộ điều khiển mờ để tránh chướng ngại vật và bộ lọc Kalman mở rộng để tổng hợp dữ liệu từ cảm biến siêu âm và encoder mang lại độ chính xác cao cho hệ thống định vị. Mặc dù, kỹ thuật tổng hợp dữ liệu từ nhiều loại cảm biến giúp cải thiện độ chính xác của dữ liệu và thông tin thu thập được, nhưng để triển khai và duy trì một hệ thống kết hợp nhiều cảm biến có thể tốn kém vì yêu cầu phần cứng và phần mềm phức tạp. Hơn nữa, việc xử lý và kết hợp dữ liệu từ nhiều nguồn có thể tạo ra độ trễ trong việc tính toán, điều này là quan trọng trong các ứng dụng yêu cầu độ phản ứng nhanh [30, 31].

Điều hướng dựa trên thị giác máy ngày càng được sự chú ý mạnh mẽ như là một giải pháp thay thế cho điều hướng dựa trên cảm biến. Các phương pháp này được phân loại thành các phương pháp dựa trên thị giác lập thể (stereovision) và thị giác một mắt (monocular vision). Các phương pháp sử dụng các kỹ thuật stereovision phân biệt chướng ngại vật bằng cách sử dụng thông tin độ sâu ba chiều [32]. Hạn chế đáng

Chương 1: Tổng quan

kẻ của các phương pháp này là thời gian tính toán nhiều và chi phí phần cứng cao. Ngược lại, các hệ thống điều hướng dựa trên monocular vision sử dụng các kỹ thuật xử lý ảnh để nhận biết các chướng ngại vật, liên quan đến tất cả các vật thể khác nhau trong môi trường [33]. Gần đây, LiDAR và camera ngày càng được sử dụng nhiều do khả năng cung cấp thông tin 3 chiều (3D) về môi trường so với các loại cảm biến khác chỉ cung cấp thông tin 2 chiều (2D) như siêu âm, laser 2D. LiDAR cung cấp thông tin 3D chính xác với khoảng cách xa so với camera do ít bị ảnh hưởng bởi điều kiện ánh sáng của môi trường xung quanh, trong khi đó camera có thể cung cấp nhiều thông tin về môi trường hơn như hình ảnh, màu sắc, và ngữ nghĩa [34]. Ngoài ra, LiDAR có giá thành cao hơn nhiều so với camera. Nhóm nghiên cứu của J. Hoey [35] và nhóm của Jordan S. Nguyen [36] đã sử dụng camera với các thuật toán áp dụng vào việc phát hiện vật cản, tìm khoảng trống để xe lăn điện vượt qua vật cản. Hệ thống này có thể giúp người sử dụng cảm thấy an toàn hơn khi di chuyển trong khu vực có nhiều chướng ngại vật.

Lĩnh vực vẽ bản đồ và định vị đã và đang được nghiên cứu rộng rãi cho xe lăn hay robot tự hành. Xe lăn hay robot tự hành phải được cung cấp chi tiết về bản đồ di chuyển để có thể được định vị trong không gian di chuyển. Hơn nữa, tọa độ hiện tại của xe lăn hay robot được dùng làm cơ sở thu thập thêm những thông tin mới trong quá trình di chuyển [37]. Các thuật toán vẽ bản đồ dần được phát triển như thuật toán bản đồ địa hóa và bản đồ hóa đồng thời (SLAM) được áp dụng để vẽ bản đồ 3D [38]. Để định vị xe lăn hay robot trong không gian di chuyển, có nhiều phương pháp được thực hiện. Đầu tiên là những phương pháp định vị 2D sử dụng những thông tin khoảng cách thu về từ sóng Wifi [39], cảm biến laser kết hợp với RFID [40], định vị bằng phương pháp tìm những đường thẳng tương đồng trong không gian 2D [41]. Hiện nay, các phương pháp xử lý ảnh dần được sử dụng trong định vị robot [42]. Phương pháp xử lý ảnh được sử dụng để nhận biết các vật mốc nhân tạo cố định được xây dựng sẵn trong không gian di chuyển [43], từ đó áp dụng vào việc xác định vị trí hiện tại của xe lăn trên một bản đồ đã được xây dựng trước trong quá trình di chuyển. Alcantarilla và các cộng sự đề xuất phương pháp mạnh mẽ và nhanh chóng định vị

cho xe lăn dựa vào thị giác máy tính, trong đó các đặc trưng của hình ảnh được trích xuất, kết hợp với các thành phần bản đồ để đưa ra vị trí hiện tại của robot tự hành [44]. Trong thực tế, bản đồ hóa cho robot di động trong môi trường là một thách thức lớn do dữ liệu thu được từ môi trường và thuật toán được áp dụng trên chúng [45, 46]. Với các robot di động dựa trên thông tin mang tính bước ngoặt, việc chọn các vật mốc như cửa ra vào, cầu thang, tường, trần nhà, sàn nhà và trích xuất các đặc trưng của chúng để nhận dạng đóng vai trò quan trọng [47, 48]. Do đó, để phát hiện các đối tượng dựa trên các đặc trưng, người ta có thể dựa trên màu sắc, kết cấu, độ sáng, kích thước đối tượng.

Từ các phân tích trên, sự kết hợp của một chiếc xe lăn điện, hệ thống máy tính, thiết bị thu thập tín hiệu điện não EEG, hệ thống camera 3D và bản đồ cần thiết được nghiên cứu để cho phép người sử dụng xe lăn cảm thấy thuận lợi hơn trong việc điều khiển và an toàn hơn khi di chuyển trong môi trường trong nhà. Người khuyết tật có thể tự điều khiển xe lăn thông qua tín hiệu EEG, tránh vật cản tự động trong những trường hợp khẩn cấp khi người sử dụng không thể phản ứng kịp thời thông qua hệ thống camera. Với hệ thống camera và bản đồ, xe lăn có thể nhận dạng vật mốc, tính toán khoảng cách và góc lệch giữa vật mốc và định vị trên bản đồ dựa vào vật mốc. Với vị trí được xác định trên bản đồ, xe lăn có thể đưa ra quyết định di chuyển phù hợp để đến đích mong muốn.

❖ CÁC KẾT QUẢ TRONG VÀ NGOÀI NƯỚC ĐÃ ĐƯỢC CÔNG BỐ

Để hỗ trợ những người gặp vấn đề nghiêm trọng về di chuyển giảm bớt hoặc loại bỏ việc điều khiển xe lăn điện, đã có nhiều công trình nghiên cứu trong và ngoài nước nghiên cứu về công nghệ xe lăn thông minh. Dưới đây trình bày tóm tắt một số kết quả nghiên cứu đã được thực hiện.

Trong đề tài nghiên cứu **Phát triển xe lăn điện thông minh dùng kỹ thuật điện não EEG và cảm biến camera cho người tàn tật nặng** của Tiến sĩ Nguyễn Thanh Hải năm 2013, một mô hình xe lăn bán tự động được thiết kế bao gồm sự kết hợp giữa người điều khiển bằng điện não EEG và chế độ điều khiển tự động tránh vật

cản [49]. Đề tài sử dụng tín hiệu EEG trong điều khiển các chuyển động đi tới - lui hoặc rẽ trái - phải của xe lăn dựa vào chớp mắt-mở mắt hay liếc mắt trái - phải. Trong quá trình di chuyển, những lệnh điều khiển của người sử dụng bằng điện não có thể không chắc chắn. Điều này có nghĩa là người sử dụng không được an toàn vì có thể va chạm vào vật cản trên đường đi. Do đó, một khoảng cách an toàn giữa camera và xe lăn được cài đặt trước và xe lăn luôn được camera dò tìm các vật cản ở phía trước để tránh va chạm nếu quá gần. Để cung cấp thông tin 3D cho phát hiện khoảng trống và những vật cản, hệ thống stereo camera “Bumblebee” được gắn trên một chiếc xe lăn điện. Cụ thể, một thuật toán tổng sai lệch tuyệt đối được sử dụng để tính toán sai lệch tối ưu giữa vị trí của một điểm ảnh trên ảnh trái và ảnh phải từ stereo camera. Dựa vào sự sai lệch này, bản đồ điểm 3D và bản đồ khoảng cách 2D được tạo ra cho mục đích tìm khoảng trống và tránh vật cản.

Luận án tiến sĩ của tác giả Lâm Quang Chuyên với đề tài **“Mạng neural trong hệ thống điều khiển xe lăn cho người tàn tật nặng sử dụng điện não (EEG) và camera”** thực hiện năm 2020 đã phân tích ba phương pháp tiền xử lý tín hiệu từ EEG, dùng biến đổi Fourier, phép biến đổi Wavelet, thuật toán biến đổi Hilbert Huang (HHT), để biến đổi thành 5 dạng sóng cơ bản Delta, Theta, Alpha, Beta, Gamma, sau đó sử dụng kỹ thuật gom cụm dữ liệu trước khi đưa vào mạng neuron để phân loại thành 5 tín hiệu mong muốn chuyển động [50]. Các mạng nơ-ron được thử nghiệm từ mạng đơn lớp đến mạng đa lớp cụ thể trong luận án này là 3 lớp. Hệ thống xử lý tín hiệu EEG bằng tiền xử lý HHT, dùng mạng nơ-ron cùng với camera được thử nghiệm trên mô hình thực tế điều khiển xe lăn đã cho những kết quả chính xác tốt nhất đến 92,4% đối với nhóm 20 người được chọn cho thực nghiệm.

Bên cạnh các nghiên cứu trong nước về xây dựng hệ thống xe lăn dùng tín hiệu EEG và cảm biến, các nghiên cứu trên thế giới cũng đã đề xuất nhiều mô hình kết hợp điều khiển hoặc chia sẻ quyền điều khiển xe lăn cho người khuyết tật. Trong nghiên cứu năm 2016, Ana Lopes đã đề xuất một mô hình chia sẻ điều khiển giữa hệ thống giao tiếp giữa não người và máy tính (BCI) P300 và thuật toán lập kế hoạch để điều khiển xe lăn điện trong môi trường thực tế trong nhà và theo thời gian thực [51].

Chương 1: Tổng quan

Công cụ lập kế hoạch đường dẫn toàn cục 3D dựa trên thuật toán A* cải tiến, sau đó nội suy đường dẫn được tạo bởi các điểm tham chiếu 3D (x_i, y_i, θ_i) . Công cụ lập kế hoạch đường dẫn toàn cục 3D chỉ hoạt động khi mục tiêu được cung cấp. Công cụ lập kế hoạch cục bộ sử dụng phương pháp cửa sổ động đôi (D-DWA), thuật toán này có được bằng cách áp dụng phương pháp cửa sổ động (DWA) tại hai lần điều khiển khác nhau. Bộ điều khiển chia sẻ cho phép người sử dụng chỉ cần ra một lệnh điều khiển như đi thẳng, rẽ trái, rẽ phải, quay lại mà không yêu cầu phản ứng liên tục hoặc theo thời gian. Hệ thống sẽ vẫn hoạt động mà không đặt người điều khiển vào tình huống nguy hiểm trong trường hợp người điều khiển không thể phản hồi vì bận thực hiện các nhiệm vụ khác. Trong nghiên cứu khác năm 2016 của Zhijun Li, một phương pháp điều khiển kết hợp giữa người và máy được đề xuất để điều khiển chuyển hướng của xe lăn, bao gồm chế độ điều khiển BCI và chế độ điều khiển tự động [52]. Trong chế độ điều khiển BCI, một giao diện não người và máy tính sử dụng phương pháp phát hiện sự tập trung ổn định vào hình ảnh trên giao diện máy tính (SSVEP) được trình bày. Trong chế độ điều khiển tự động, kỹ thuật trường tiềm năng dựa trên góc (APF) và kỹ thuật SLAM dựa trên thị giác máy tính được đề xuất để hướng dẫn xe lăn điều hướng giữa các chướng ngại vật.

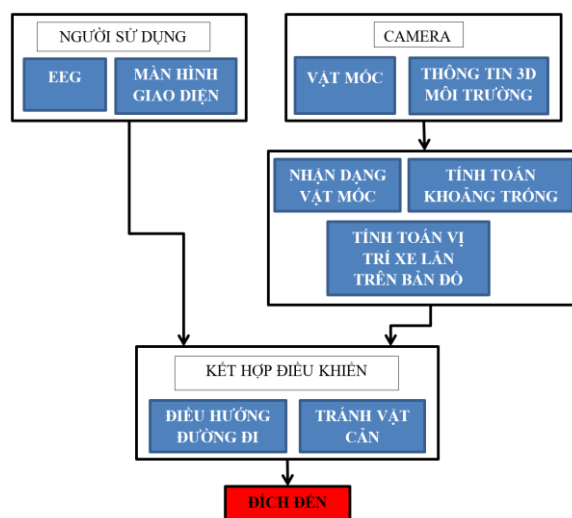
Jingsheng Tang đã đề xuất một cấu trúc di động cải tiến được trang bị cho xe lăn bao gồm cánh tay robot nhẹ, mô-đun nhận dạng mục tiêu và mô-đun điều khiển tự động trong nghiên cứu của mình năm 2018 [53]. Dựa trên thuật toán YOLO, trong thời gian thực, hệ thống này có thể nhận dạng và định vị các mục tiêu trong môi trường bằng camera Kinect và người dùng xác nhận một mục tiêu thông qua giao diện BCI – P300. Hệ thống dùng cảm biến LIDAR để tạo ra một bản đồ, dùng phương pháp DWA lập kế hoạch đường đi cục bộ cho xe lăn, và công cụ lập kế hoạch đường đi toàn cục dựa trên thuật toán A*. Trong quá trình di chuyển, mục tiêu cũng được theo dõi bằng cách sử dụng kỹ thuật theo dõi hình ảnh. Mô-đun nhận dạng mục tiêu sẽ nhận dạng và định vị mục tiêu trong môi trường trực tuyến, người dùng xác nhận một mục tiêu bằng cách chọn mục tiêu này trực tiếp thông qua hệ thống BCI và mô-

đun điều hướng tự động điều khiển xe lăn đến đích được chỉ định. Ngoài ra, xe lăn còn được trang bị một cánh tay robot nhẹ giúp người dùng lấy các đồ vật.

Với những ưu điểm của việc sử dụng tín hiệu EEG, camera, vật mốc và bản đồ cho xe lăn điện bán tự động, đề tài “**Kết hợp tín hiệu EEG, Stereo camera và vật mốc để định vị, điều khiển xe lăn điện đến đích dựa vào bản đồ**” được lựa chọn để nghiên cứu.

❖ CẤU TRÚC HỆ THỐNG XE LĂN ĐIỆN KẾT HỢP TÍN HIỆU EEG, CAMERA, VẬT MỐC VÀ BẢN ĐỒ

Sự kết hợp của một chiếc xe lăn điện, hệ thống máy tính, thiết bị thu thập tín hiệu điện não EEG, hệ thống camera 3D và bản đồ cho phép người sử dụng xe lăn cảm thấy thuận lợi hơn trong việc điều khiển và an toàn hơn khi di chuyển trong môi trường trong nhà. Người khuyết tật lựa chọn đích đến thông qua tín hiệu EEG và giao diện BCI. Với hệ thống camera và bản đồ, xe lăn có thể nhận dạng các vật mốc trong môi trường và xác định vị trí hiện tại trên bản đồ trong quá trình di chuyển. Vị trí hiện tại và đích đến của xe lăn sẽ được cung cấp cho hệ thống điều khiển tự động để lập lộ trình. Ngoài ra, trong quá trình di chuyển, xe lăn có thể tránh vật cản tự động thông qua hệ thống camera. Quá trình thu thập dữ liệu từ camera, tín hiệu EEG và bản đồ để điều khiển xe lăn điện được thực hiện như hình 1.1.



Hình 1.1. Sơ đồ khối quá trình điều khiển xe lăn điện dựa vào tín hiệu EEG, camera, vật mốc và bản đồ.

1.2. MỤC TIÊU NGHIÊN CỨU CỦA LUẬN ÁN

Mục tiêu của luận án là thiết kế và xây dựng hệ thống điều khiển bán tự động cho xe lăn điện trong môi trường trong nhà dựa vào sự kết hợp tín hiệu điện não EEG, hệ thống camera, vật mốc và bản đồ nhằm giảm mức độ điều khiển người khuyết tật và tăng độ an toàn. Để thực hiện được mục tiêu chính này, các mục tiêu cụ thể cần thực hiện là:

1. Thiết kế một giao diện giao tiếp giữa người và máy tính thông qua tín hiệu EEG từ hoạt động nháy mắt để chọn đích đến mong muốn. Trước tiên, các thành phần cơ bản của một tín hiệu EEG của hoạt động nháy mắt được phân tích chi tiết. Từ đó xây dựng một bộ lọc nhiễu và làm trơn tín hiệu EEG, cũng như trích xuất các thành phần đặc trưng của tín hiệu để có được một tín hiệu EEG của hoạt động nháy mắt tốt nhất. Cuối cùng, các bộ phân loại các hoạt động nháy mắt từ tín hiệu EEG với độ chính xác cao và thời gian xử lý nhanh, và sau đó gán các lệnh điều khiển giao tiếp với máy tính, cũng cần được nghiên cứu thiết kế.

2. Đề xuất phương pháp định vị xe lăn điện trên bản đồ dựa vào vị trí các vật mốc trong môi trường. Để thực hiện việc này, các vật mốc với thông tin vị trí của nó trong môi trường cần phải được thu thập. Do đó, phương pháp nhận biết các vật mốc trong môi trường tự nhiên và xác định vị trí của vật mốc cần được nghiên cứu. Hơn nữa, với cơ sở dữ liệu về vật mốc đã thu thập được, phương pháp định vị vị trí của xe lăn trên bản đồ với độ chính xác cao cũng cần được đề xuất nghiên cứu.

3. Xây dựng mô hình điều khiển bán tự động cho xe lăn đến đích mong muốn, trong đó giảm thiểu tối đa sự điều khiển của người sử dụng. Cụ thể, với vị trí của xe lăn đã được xác định dựa vào các vật mốc và vị trí điểm đích mà người dùng lựa chọn, một phương pháp điều khiển tự động với tính ổn định và độ an toàn cao được đề xuất nghiên cứu.

1.3. ĐỐI TƯỢNG VÀ PHẠM VI NGHIÊN CỨU

- **Đối tượng nghiên cứu**

- Nghiên cứu phương pháp phân loại hoạt động nháy mắt từ tín hiệu EEG.

- Nghiên cứu thuật toán nhận dạng vật mốc trong môi trường tự nhiên.
- Nghiên cứu thuật toán định vị dựa vào vật mốc trong môi trường tự nhiên.
- Nghiên cứu thuật toán điều khiển tự động cho xe lăn dựa vào sự kết hợp tín hiệu EEG, vật mốc và bản đồ.
 - **Phạm vi nghiên cứu**
- Luận án tập trung nghiên cứu hệ thống điều khiển xe lăn điện trong không gian trong nhà, với người sử dụng là những người khuyết tật bị hạn chế vận động tay, chân hoặc đầu nhưng mắt vẫn còn khỏe.

1.4. CÁCH TIẾP CẬN VÀ PHƯƠNG PHÁP NGHIÊN CỨU

Dựa trên đối tượng nghiên cứu, cũng như mục tiêu đã được đặt ra của luận án. Cách tiếp cận và phương pháp nghiên cứu trong luận án này được mô tả theo từng bước cụ thể như bên dưới:

- Phương pháp khảo sát lý thuyết: tìm hiểu, phân tích và tổng hợp các cơ sở lý thuyết liên quan đến luận án cũng như tổng hợp phương pháp và kết quả của các công trình nghiên cứu trước đó đã được công bố trong và ngoài nước về điều khiển xe lăn điện trong môi trường trong nhà. Từ những cơ sở lý thuyết được khảo sát và phân tích sẽ là tiền đề cho việc xây dựng mục tiêu nghiên cứu.

- Phương pháp tính toán, thiết kế và hiệu chỉnh: từ những lý thuyết liên quan được khảo sát, một hệ thống sẽ được tính toán và thiết kế để điều khiển xe lăn điện. Quá trình thiết kế sẽ được kiểm tra và hiệu chỉnh để đạt được kết quả tốt nhất.

- Phương pháp mô phỏng: từ hệ thống được thiết kế, sử dụng các chương trình mô phỏng có độ tin cậy cao để đánh giá hệ thống.

- Phương pháp phân tích và tổng hợp: tổng hợp kết quả đã đạt được trong luận án và so sánh với các kết quả đã được công bố trước đó để bình luận và phân tích ưu nhược điểm của phương pháp được đề xuất trong luận án.

- Phương pháp thực nghiệm: tiến hành các thí nghiệm trong môi trường thực để quan sát, thu thập thông tin, từ đó kiểm chứng và đánh giá tính thực tế của các phương pháp được đề xuất trong luận án.

1.5. ĐÓNG GÓP VỀ KHOA HỌC CỦA LUẬN ÁN

❖ Đóng góp mới về khoa học của luận án

Luận án tập trung vào nghiên cứu xây dựng mô hình hệ thống điều khiển bán tự động cho xe lăn điện trong môi trường trong nhà dựa vào sự kết hợp tín hiệu điện não EEG, hệ thống camera, vật mốc và bản đồ nhằm giảm mức độ điều khiển người khuyết tật và tăng độ an toàn. Vì vậy, những đóng góp mới về mặt khoa học của luận án gồm:

- Đề xuất phương pháp phân loại tín hiệu EEG của hoạt động nháy mắt dựa vào ngưỡng biên độ và phương pháp phân loại tín hiệu EEG của hoạt động nháy mắt dựa vào mạng học sâu CNN-1D. Cụ thể, tín hiệu EEG của hoạt động nháy mắt được thu thập và tiền xử lý sẽ được thu thập cho quá trình phân loại. Phương pháp phân loại tín hiệu EEG của hoạt động nháy mắt dựa vào ngưỡng biên độ cho phép xử lý trực tiếp tín hiệu EEG với thời gian ngắn và độ chính xác cao. Với mô hình mạng học sâu CNN-1D, dữ liệu cần được thu thập trước cho quá trình huấn luyện. Tuy nhiên, mạng CNN-1D cho phép phân loại được nhiều loại nháy mắt hơn, với độ chính xác cao hơn.

- Đề xuất phương pháp nhận dạng vật mốc tự nhiên và xác định vị trí vật mốc trong môi trường trong nhà. Trong quá trình di chuyển, với phương pháp này, các vật mốc không cần phải được học trước mà xe lăn sẽ tự nhận biết và lựa chọn vật mốc dựa vào mật độ đặc trưng của các đối tượng có trong ảnh môi trường. Từ đó, xe lăn sẽ tính toán vị trí của vật mốc, và sau đó thu thập vào cơ sở dữ liệu. Quá trình thu thập các vật mốc và thông tin vị trí của nó được thực hiện với thời gian ngắn với độ chính xác cao, làm cơ sở cho việc định vị xe lăn trên bản đồ.

- Với thông tin của vật mốc, luận án đề xuất phương pháp định vị xe lăn trên bản đồ lưới 2D ảo, giúp cho quá trình điều khiển xe lăn đến đích được chính xác và nhanh chóng hơn. Cụ thể, bằng việc xây dựng một bản đồ lưới 2D ảo từ môi trường thực với các ô trống và ô vật cản, xe lăn cần được xác định vị trí để hệ thống điều khiển đề xuất đường đi tối ưu đến đích. Vị trí này được tính toán từ vị trí của vật mốc trong không gian môi trường và trong không gian 3D của camera. Việc sử dụng một vật mốc để định vị xe lăn được kiến nghị vì cho độ chính xác cao hơn so với việc dùng 3 vật mốc.

- Đề xuất mô hình điều khiển thực - ảo cho xe lăn điện bán tự động. Trong đó, mô hình DQNs-PreLU được kiến nghị để huấn luyện tìm đường đi tối ưu cho xe lăn dựa vào bản đồ lưới 2D ảo. Mô hình DQNs-PreLU với các thông số được lựa chọn giúp giảm thời gian huấn luyện nhưng vẫn đảm bảo độ chính xác. Hơn nữa, thuật toán điều khiển xe lăn trong môi trường thực từ các đường đi mô phỏng trên bản đồ lưới cũng được đề xuất.

❖ Ý nghĩa thực tiễn của luận án

Với các phân tích về tính cấp thiết đã được trình bày ở phần trên, luận án có ý nghĩa thực tiễn trong việc xây dựng một mô hình xe lăn điện bán tự động cho những người khuyết tật bị hạn chế về vận động tay, chân hoặc đầu nhưng mắt vẫn còn khỏe. Cụ thể, mô hình xe lăn điện bán tự động với sự kết hợp của giao diện giao tiếp giữa người và máy tính thông qua tín hiệu EEG cho phép người dùng chọn đích đến mong muốn, và phương pháp điều khiển xe lăn tự động dựa vào bản đồ lưới 2D, camera và các vật mốc trong môi trường trong nhà sẽ có ý nghĩa quan trọng trong việc giải quyết các vấn đề sau: (1) Giảm nỗ lực điều khiển xe lăn của người khuyết tật; (2) Đảm bảo an toàn trong quá trình di chuyển; (3) Giảm giá thành sản phẩm với việc sử dụng phần cứng có chi phí thấp và phần mềm mã nguồn mở; (4) Hệ thống được thiết kế có tính di động cho người dùng. Ngoài ra, kết quả nghiên cứu còn được áp dụng để giảng dạy cho các sinh viên chuyên ngành Kỹ Thuật Y Sinh tại Bộ Môn Điện Tử Công Nghiệp – Y Sinh của Trường Đại Học Sư Phạm Kỹ Thuật Tp.HCM.

1.6. CẤU TRÚC CỦA LUẬN ÁN

Chương 1: Tổng quan, chương này giới thiệu tổng quan về đề tài, mục tiêu, đối tượng nghiên cứu, phương pháp nghiên cứu, các đóng góp mới về khoa học.

Chương 2: Cơ sở lý thuyết, chương này trình bày các vấn đề liên quan đến việc thiết kế một hệ thống xe lăn điện bán tự động cho người khuyết tật. Cụ thể, giao tiếp giữa não người và máy tính BCI thông qua việc sử dụng tín hiệu điện não EEG và các ứng dụng của nó trong thiết kế mô hình xe lăn điện cho người khuyết tật sẽ được trình bày. Bên cạnh đó, các phương pháp xây dựng bản đồ lưới 2D môi trường trong nhà và thuật toán học tăng cường cho điều khiển robot di động cũng sẽ được tóm lược.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính, chương này trình bày phương pháp phân loại các hoạt động nháy mắt dựa vào tín hiệu EEG. Quá trình thu thập dữ liệu EEG từ thiết bị Emotiv EPOC+ được trình bày bao gồm các phương pháp tiền xử lý tín hiệu, phân tích tín hiệu và gán nhãn dữ liệu. Ngoài ra, các phương pháp phân loại tín hiệu EEG của các hoạt động mắt cũng được đề xuất trong chương này bao gồm phương pháp phân loại theo ngưỡng biên độ và phương pháp phân loại dựa vào mô hình mạng học sâu CNN-1D. Các thí nghiệm sẽ được trình bày để đánh giá hiệu suất phân loại tín hiệu.

Chương 4: Nhận dạng và xác định vị trí vật mốc tự nhiên trong môi trường trong nhà, chương này trình bày phương pháp mật độ điểm đặc trưng lớn nhất để nhận biết các vật mốc tự nhiên trong môi trường trong nhà. Hơn nữa, phương pháp định vị vật mốc dựa vào vị trí xe lăn và thông tin 3D của vật mốc có được từ camera cũng được trình bày cụ thể. Hiệu suất nhận dạng vật mốc được đo lường bằng hệ số IOU, và độ chính xác của vị trí vật mốc được đánh giá thông qua tính toán sai số tuyệt đối và tương đối giữa vị trí thực và vị trí tính toán.

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ, chương này trình bày mô hình điều khiển bán tự động có sự kết hợp tín hiệu EEG, camera, vật mốc và bản đồ lưới 2D ảo để điều khiển xe lăn

đến đích mong muốn. Mô hình điều khiển thực - ảo bao gồm công cụ lập kế hoạch di chuyển (MP) và bộ chuyên đổi hành động của xe lăn (WAC) được trình bày. Hơn nữa, xe lăn cần xác định vị trí hiện tại của nó trong cả môi trường thực và ảo sử dụng thông tin của các vật mốc tự nhiên. Ngoài ra, mô hình Deep Q-Networks (DQNs) và bản đồ lưới 2D ảo để hoạch định đường đi tối ưu cho xe lăn được đề xuất. Sau đó, phương pháp điều khiển xe lăn điện trong môi trường thực cùng với các thí nghiệm mô phỏng và thí nghiệm trong môi trường thực cũng được trình bày.

Chương 6: Kết luận và hướng phát triển, chương này trình bày những kết quả đã đạt được của luận án và đề xuất phương hướng phát triển các nghiên cứu của luận án trong tương lai.

CHƯƠNG 2

CƠ SỞ LÝ THUYẾT

Chương này trình bày các vấn đề liên quan đến việc thiết kế một hệ thống xe lăn điện bán tự động cho người khuyết tật. Cụ thể, giao tiếp giữa não người và máy tính BCI thông qua việc sử dụng tín hiệu điện não EEG và các ứng dụng của nó trong thiết kế mô hình xe lăn điện cho người khuyết tật sẽ được trình bày. Đây là nền tảng cơ sở để xây dựng mô hình điều khiển xe lăn điện trong luận án này. Bên cạnh đó, các phương pháp xây dựng bản đồ lưới 2D môi trường trong nhà và các phương pháp lập kế hoạch đường đi cũng sẽ được tóm lược, hướng đến hình thành phương pháp điều khiển xe lăn điện trong môi trường trong nhà.

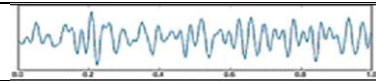
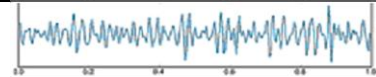
2.1. TỔNG QUAN VỀ TÍN HIỆU EEG

EEG là tín hiệu điện của hoạt động vỏ não được phát hiện bởi Hans Berger năm 1924. Tín hiệu này được đo bằng một dụng cụ đo dòng điện với điện cực được gắn trên đầu và ghi lại những dao động điện. Các hoạt động của não có mối liên quan đến các tín hiệu từ vỏ não [54]. Do đó, việc phân tích tín hiệu EEG có thể phát hiện những bất thường trong hoạt động của não [55]. Tín hiệu điện não có biên độ trong khoảng 5-200 μV . Tần số thay đổi trong khoảng 1-100 Hz. Vì biên độ nhỏ nên rất dễ bị nhiễu, chủ yếu là các thành phần tín hiệu điện cơ (EMG) và tín hiệu điện tim (ECG) [56]. Phương pháp đo điện não EEG có thể phát hiện ra những thay đổi trong thời gian vài mili giây, trong khi đó điện thế hoạt động truyền giữa các nơ-ron mất xấp xỉ từ 0,5 - 130 ms nên hoàn toàn có thể ghi nhận được. Ngoài ra, tín hiệu EEG cũng cho thấy sự khác nhau về cường độ và vị trí của hoạt động tại mỗi khu vực của não.

Các dạng sóng của tín hiệu EEG thường được phân loại theo tần số, biên độ, hình dạng cũng như vị trí của các điện cực trên da đầu. Sự phân loại của các dạng

sóng như Alpha, Beta, Theta, Delta và Gamma dựa trên tần số tín hiệu. Một số dạng sóng được nhận biết dựa trên hình dạng, phân bố đỉnh và tính chất đối xứng của chúng. Tần số EEG khác nhau tương ứng với các hành vi và trạng thái tinh thần khác nhau của não [54] như mô tả ở bảng 2.1.

Bảng 2.1. Các dạng sóng EEG.

Loại sóng	Tần số (Hz)	Đặc điểm hành vi	Dạng sóng
Delta	0.3 – 4	Ngủ sâu	
Theta	4 – 8	Thiền sâu	
Alpha	8 – 13	Nhắm mắt trong lúc tỉnh táo	
Beta	13 – 30	Mở mắt và suy nghĩ	
Gamma	Trên 30	Thống nhất ý thức	

2.2. PHÂN LOẠI HOẠT ĐỘNG DỰA VÀO TÍN HIỆU EEG

Phân loại tín hiệu EEG nói chung hay phân loại tín hiệu EEG về hoạt động mắt nói riêng là mấu chốt quan trọng trong việc hỗ trợ người không có khả năng tự chủ vận động. Hiện nay, có rất nhiều nghiên cứu phân loại các dạng tín hiệu EEG đã thành công về mặt nghiên cứu và ứng dụng. Điều này đã giúp người khuyết tật được chủ động hơn và cảm thấy hạnh phúc hơn trong cuộc sống của mình.

2.2.1. Phát hiện hoạt động mắt dựa vào ngưỡng biên độ tín hiệu EEG

Việc phát hiện các hoạt động của mắt như nháy mắt dựa vào ngưỡng biên độ tín hiệu EEG cho thấy sự hiệu quả với thời gian xử lý nhanh chóng và không yêu cầu đào tạo người dùng trước khi thực thi thuật toán [57, 58]. Các tín hiệu EEG được lọc và được làm mịn trước khi phát hiện nháy mắt. Thuật toán phát hiện nháy mắt tập trung vào việc tính toán đỉnh chính xác dựa vào đặc tính ngưỡng và biên độ đỉnh [57]. Thuật toán có thể hoạt động chính xác trên một hoặc hai kênh tín hiệu EEG. Các bộ dữ liệu EEG khác nhau đã được sử dụng để đánh giá độ chính xác của thuật toán cho

thấy rằng thuật toán có độ chính xác trung bình trên 96% cho phân loại nháy mắt trái và nháy mắt phải.

Bên cạnh phương pháp phát hiện đỉnh của tín hiệu EEG cho phân loại hoạt động mắt, nghiên cứu của nhóm Trung Nguyễn [58] đã sử dụng giá trị trung bình của tín hiệu EEG khi mở mắt để xác định ngưỡng cho việc phân loại mở mắt và nháy mắt. Công thức (2.1) mô tả cách xác định ngưỡng trung bình như sau:

$$THR = M - a \times SD \quad (2.1)$$

trong đó a là hệ số của độ lệch chuẩn SD , M là giá trị trung bình của tín hiệu mở mắt. Độ lệch chuẩn và giá trị trung bình của tín hiệu mở mắt được tính như sau:

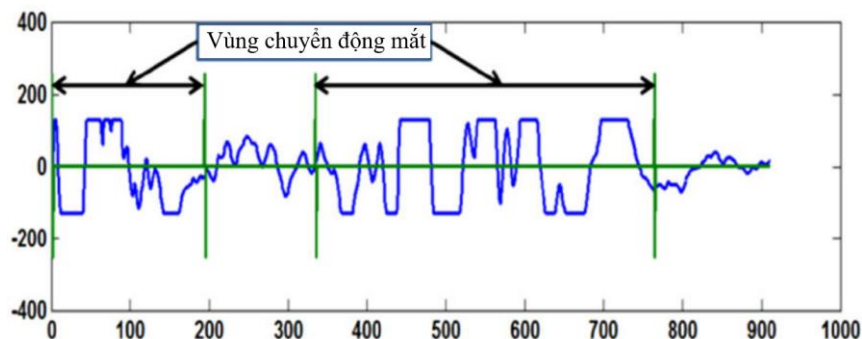
$$M = \frac{\sum_{n=1}^N y(n)}{N} \quad (2.2)$$

$$SD = \frac{\sqrt{\sum_{n=1}^N (y(n) - M)^2}}{N} \quad (2.3)$$

với $y(n)$ là tập hợp các tín hiệu EEG và N là số lượng mẫu.

2.2.2. Phân loại tín hiệu EEG dùng mạng nơ-ron

Tibdewal và cộng sự trình bày một mạng nơ-ron nhân tạo đơn giản và hiệu quả với tên gọi là thuật toán Biên độ thời gian [59]. Bằng cách áp dụng thuật toán Biên độ thời gian trên tín hiệu được xác định, chuyển động mắt ở các vùng khác nhau được tự động phát hiện và đánh dấu chính xác như mô tả ở hình 2.1. Thời gian thực hiện thuật toán Biên độ thời gian để tự động phát hiện chuyển động của mắt là rất ngắn, khoảng 4,3 ms và độ chính xác 97,5%.



Hình 2.1. Tín hiệu EEG được phát hiện với các vùng chuyển động của mắt.

Nghiên cứu của nhóm Nguyễn Thanh Hải [18] sử dụng đầu vào của mạng nơ-ron là các đặc trưng được trích xuất từ tín hiệu EEG bằng mô hình tự hồi quy cho phân loại hoạt động mắt như mở mắt, nháy mắt, liếc mắt với độ chính xác 94%. Nghiên cứu của tác giả Lâm Quang Chuyên năm 2020 sử dụng dữ liệu từ 10 kênh của thiết bị Emotiv EPOC [50]. Dữ liệu thô để huấn luyện mô hình mạng nơ-ron là 5 dạng sóng Delta, Theta, Alpha, Beta, Gamma có được khi các đối tượng thí nghiệm nhìn vào 5 loại hình ảnh khác nhau. Sau đó, tác giả sử dụng biến đổi FFT và kỹ thuật gom cụm để trích đặc trưng tín hiệu cho việc huấn luyện, với độ chính xác của mô hình phân loại là 92,68%.

Đặc điểm chung của các nghiên cứu phân loại hoạt động mắt dùng mạng nơ-ron là cần phải thực hiện trích xuất các đặc trưng của tín hiệu EEG. Quá trình này sẽ cần nhiều thời gian để phân tích và tạo ra cơ sở dữ liệu cho việc huấn luyện, đồng thời độ chính xác của mô hình mạng nơ-ron phụ thuộc nhiều vào độ chính xác của quá trình trích đặc trưng.

2.2.3. Phân loại tín hiệu EEG dùng mạng nơ-ron tích chập

Trong những năm gần đây, các nghiên cứu đã tập trung vào việc ứng dụng mô hình mạng nơ-ron tích chập CNN để phân loại các tín hiệu EEG. Trong nghiên cứu của Abdul Qayyum [60], tín hiệu EEG thu được từ phép biến đổi Wavelet rời rạc của tín hiệu gốc được sử dụng làm đầu vào cho mô hình CNN-1D để phân biệt các cấp độ khác nhau của trạng thái nghỉ ngơi và học tập. Z. Pei và các cộng sự đã sử dụng các đặc trưng của tín hiệu EEG từ một kênh kênh và nhiều kênh để phân loại nhiều nhiệm vụ công việc với độ chính xác là 85% [61]. Mô hình CNN-3D lặp lại sử dụng các đặc trưng từ phổ tần số, thời gian và không gian của tín hiệu EEG để phân loại mức độ khối lượng công việc trí óc cao và thấp cho 20 đối tượng với độ chính xác 88,9% được đề xuất bởi P. Zhang [62]. Tuy nhiên, hầu hết các nghiên cứu này sử dụng bước tiền xử lý để cải thiện độ chính xác phân loại hoặc biểu diễn tần số, thời gian của tín hiệu EEG dưới dạng đầu vào 2D hoặc 3D cho mô hình CNN, làm cho kiến trúc CNN tổng thể trở nên phức tạp và tăng thời gian tính toán của toàn hệ thống.

Mô hình phân loại của Rajendra Acharya và cộng sự năm 2017 sử dụng CNN-1D để phân biệt 3 loại tín hiệu EEG của người bình thường và người co giật [63]. Cấu trúc mô hình CNN-1D gồm 13 lớp, trong đó có 5 lớp tích chập, 5 lớp max-pooling và 3 lớp kết nối đầy đủ với 3 đầu ra tương ứng cho 3 tín hiệu cần nhận dạng. Các phân đoạn điện não đồ được sử dụng trong nghiên cứu này là những phân đoạn được thu thập bởi Andrzejak và cộng sự tại Đại học Bonn, Đức. Tập dữ liệu thu được từ 5 bệnh nhân, gồm ba tập dữ liệu là bình thường, trước đột quỵ, và động kinh, với 100 tín hiệu EEG trong mỗi tập dữ liệu. Thời gian thu tín hiệu trung bình là 23,6s. Độ chính xác của mô hình phân loại đạt 88,67%.

2.3. GIAO TIẾP GIỮA NÃO NGƯỜI VÀ MÁY TÍNH

Giao tiếp giữa não người và máy tính BCI là một trong những công nghệ hứa hẹn nhất và ngày càng phổ biến để hỗ trợ và cải thiện khả năng giao tiếp hoặc kiểm soát đối với chứng tê liệt vận động chẳng hạn như liệt nửa người hoặc liệt tứ chi do đột quỵ, chấn thương tủy sống, bại não và xơ cứng teo cơ một bên (ALS) [64]. BCI về cơ bản liên quan đến quá trình chuyển đổi hoạt động của bộ não con người thành hành động bên ngoài bằng cách gửi các lệnh thần kinh đến các thiết bị [65]. Đặc biệt, nghiên cứu gần đây về kiểm soát nhận thức và vận động nhằm cải thiện hệ thống BCI để tăng cường sức khỏe cho người cao tuổi đã được chỉ ra [64]. Theo nghiên cứu này, hệ thống BCI có thể hữu ích cho người cao tuổi trong việc rèn luyện khả năng vận động và nhận thức để ngăn ngừa tác động của quá trình lão hóa. Do đó, BCI có thể giúp họ điều khiển dễ dàng hơn các thiết bị gia dụng, trao đổi thông tin trong sinh hoạt hàng ngày.

Trong nghiên cứu khác về BCI [66], các tác giả trình bày các nguyên tắc vật lý của BCI và các phương pháp cơ bản mới để thu nhận và phân tích các tín hiệu EEG để điều khiển liên quan đến các hoạt động của não. Trong đó, hệ thống BCI được phân thành ba loại chính bao gồm chủ động, phản ứng và thụ động. Đối với BCI dạng chủ động, người dùng giao diện thần kinh điều khiển một thiết bị phức tạp như xe lăn thông qua một loạt các thành phần chức năng của hệ thống điều khiển và xem kết quả

của điều khiển này trên màn hình. BCI phản ứng kế thừa nhiều tính năng của BCI chủ động, với một sự thay đổi đáng kể để thực hiện một hệ thống điều khiển dựa trên việc phân loại các phản ứng lại của não đối với các kích thích như hình ảnh, âm thanh và xúc giác. BCI thụ động được thiết kế để theo dõi hoạt động hiện tại của não và do đó cung cấp thông tin quan trọng về trạng thái tinh thần của người vận hành, ý định của người dùng và diễn giải tình huống.

Xe lăn điều khiển bằng EEG là một ứng dụng điển hình của BCI, có thể giúp những người khuyết tật giao tiếp với môi trường bên ngoài. Trong nghiên cứu của H. Wang năm 2021 [67], xe lăn điều khiển bằng tín hiệu EEG đã được khai thác từ nhiều khía cạnh, bao gồm các loại tín hiệu EEG, các lệnh được sử dụng cho hệ thống điều khiển và phương pháp điều khiển. Hơn nữa, các tác giả đã tóm tắt sự phát triển gần đây của xe lăn điều khiển bằng tín hiệu EEG, thể hiện chủ yếu ở ba khía cạnh: từ điện cực ướt đến điện cực khô; từ chế độ đơn sang đa chế độ; từ điều khiển đồng bộ sang điều khiển bất đồng bộ. Nghiên cứu này chỉ ra rằng các chức năng mới đã được sử dụng trong xe lăn điều khiển bằng tín hiệu EEG để tăng tính ổn định và độ chính xác.

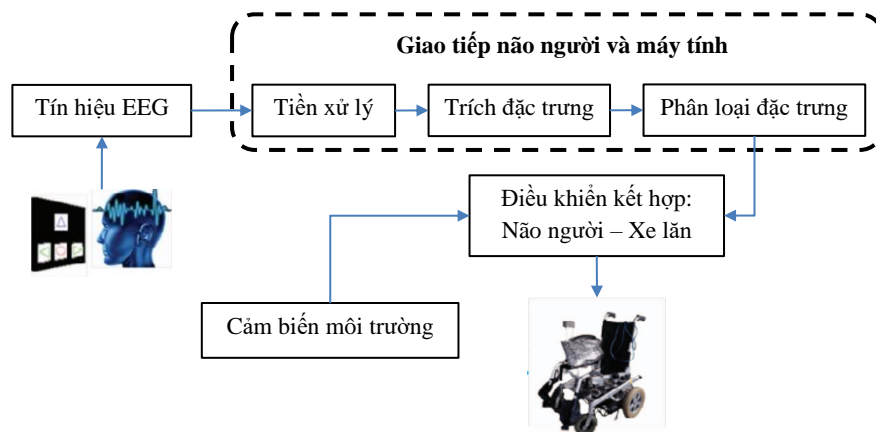
2.4. MÔ HÌNH XE LĂN ĐIỆN CHO NGƯỜI KHUYẾT TẬT

Xe lăn cho phép người sau đột quỵ hoặc người tàn tật nặng trở nên độc lập trong các hoạt động hàng ngày của họ, chẳng hạn như di chuyển. Hiện nay, nhu cầu sử dụng xe lăn cho bệnh nhân bại liệt và người già tăng nhanh do dân số già hóa. Đối với những người có vấn đề về thể chất, việc dựa vào xe lăn đẩy là không thoải mái. Do đó, xe lăn điện đã được phát minh cho người khuyết tật để tăng khả năng cơ động. Ngoài ra, một tỷ lệ lớn bệnh nhân bị chấn thương tủy sống và rối loạn thần kinh cơ chủ yếu dựa vào xe lăn điện để di chuyển. Do đó, xe lăn điện được coi là thiết bị hữu ích cho bệnh nhân bại liệt để họ có thể di chuyển được [67].

2.4.1. Xe lăn điện thông minh

Khả năng cơ động của xe lăn cơ học là một nhiệm vụ đầy thách thức đối với người già và người khuyết tật. Do đó, việc thiết kế một chiếc xe lăn thông minh mang

lại khả năng cơ động dễ dàng là cần thiết. Để giúp người khuyết tật thực hiện các hoạt động hàng ngày của họ, xe lăn thông minh được trang bị bộ điều khiển thông minh, cảm biến tiệm cận nhận dạng môi trường và động cơ đã được thiết kế như mô tả ở hình 2.2 [68]. Cộng đồng khoa học đã dành sự chú ý đáng kể cho việc phát minh ra xe lăn thông minh. Mục tiêu chính là giúp người dùng giảm hoặc loại bỏ sự tham gia lái xe lăn. Gần đây, một loạt các nghiên cứu về công nghệ xe lăn thông minh đã được tiến hành cho những người gặp vấn đề nghiêm trọng về di chuyển [15, 69].



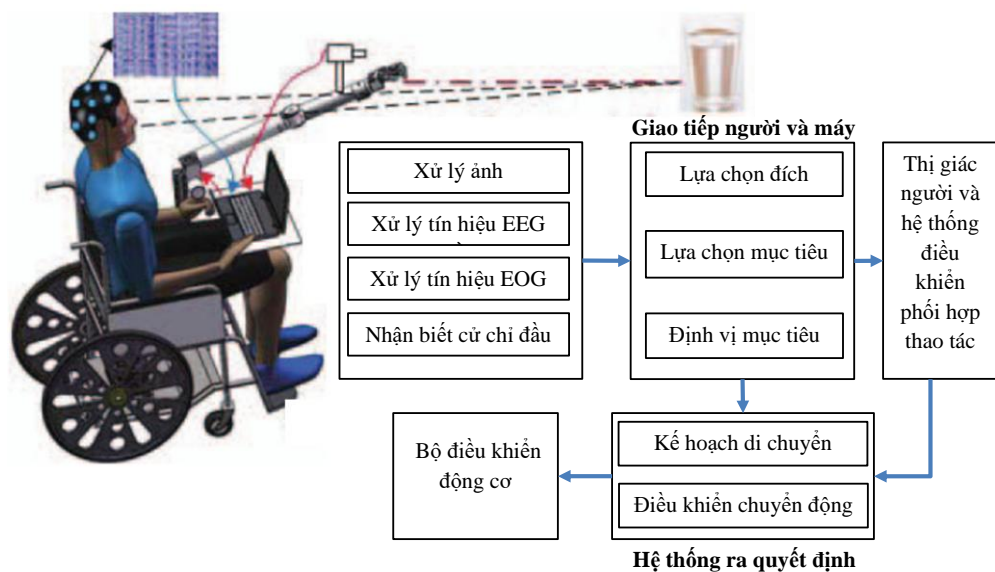
Hình 2.2. Sơ đồ điều khiển kết hợp của BCI và xe lăn.

Một số nghiên cứu tiên tiến trong lĩnh vực kỹ thuật y sinh để phát triển các thiết bị hỗ trợ cũng đã được tiến hành, bao gồm việc sử dụng tín hiệu sinh học, giọng nói và màn hình cảm ứng làm đầu vào thiết bị để cải thiện khả năng vận động của bệnh nhân bị liệt [70]. Xe lăn điều khiển bằng tín hiệu điện não là thiết bị thích hợp cho những bệnh nhân bị liệt hoàn toàn với bộ não khỏe mạnh để điều hướng xe lăn trong môi trường của họ [71]. Trong những năm gần đây, xe lăn thông minh đã được cải thiện về các tính năng với chi phí thấp và hiệu suất cao, chẳng hạn như tự động di chuyển, tránh chướng ngại vật, an toàn, linh hoạt, phương thức giao tiếp giữa người và máy, và thích ứng với mức độ khuyết tật [25].

2.4.2. Xe lăn điện với bộ điều khiển robot

Kết quả thống kê thu được từ các nghiên cứu về người khuyết tật cho thấy các cử động của các chi trên của con người như cử chỉ và nhật đồ vật chiếm ưu thế nhất

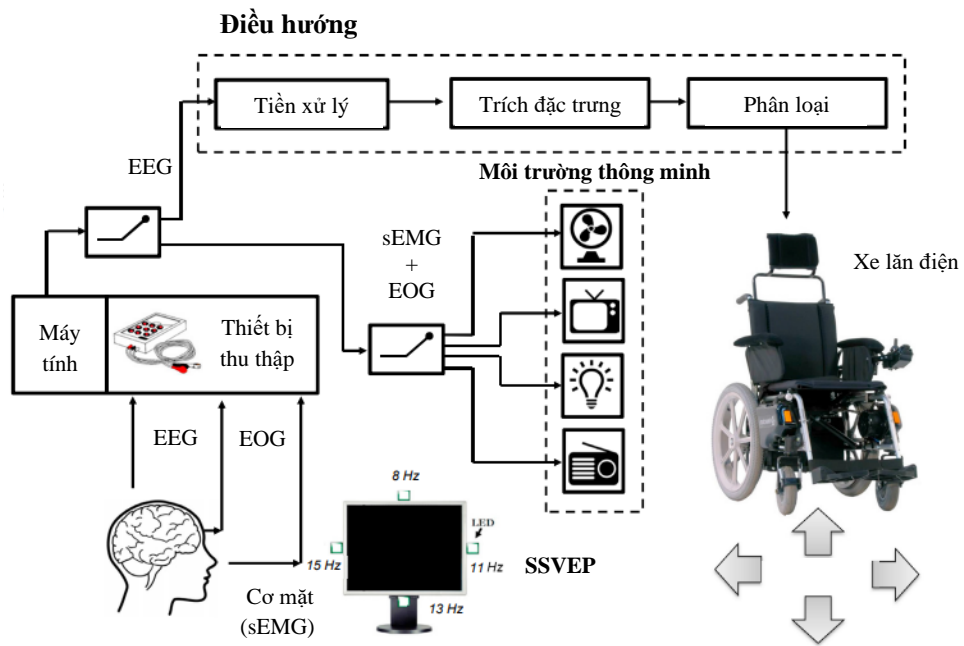
trong các hoạt động đời sống hàng ngày. Do đó, việc tích hợp một bộ điều khiển robot vào xe lăn điện có thể hỗ trợ người già và bệnh nhân bại liệt tương tác với môi trường của họ. Ngoài ra, một số loại robot điều khiển đã được phát triển, chẳng hạn như Victoria, Raptor, FRIEND, MANUS và Rancho Golden, để hỗ trợ di chuyển như minh họa ở hình 2.3 [72]. Kết quả thu được từ thực nghiệm chứng minh khả năng sử dụng và lựa chọn một số lệnh di chuyển để điều khiển xe lăn bằng tay máy robot. Đặc biệt, xe lăn tích hợp với bộ điều khiển robot có thể hỗ trợ bệnh nhân bị liệt do tổn thương tủy sống ở mức độ nhẹ và nặng [73].



Hình 2.3. Sơ đồ bộ điều khiển xe lăn di động kết hợp cánh tay robot.

2.4.3. Xe lăn điện tích hợp với môi trường thông minh

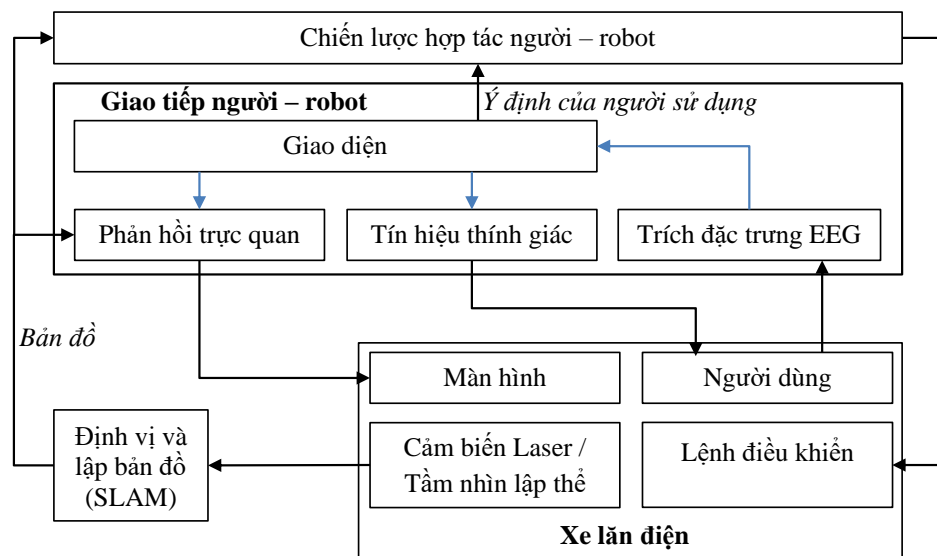
Gần đây, nhiều loại thiết bị trợ giúp đã ra đời nhằm cải thiện chất lượng cuộc sống của người khuyết tật, đặc biệt là khuyết tật vận động, giúp họ độc lập trong các hoạt động hàng ngày. Do đó, một sơ đồ tích hợp với các giao diện điều khiển đa phương thức như giọng nói, âm thanh và EEG, rất có giá trị để điều khiển một hoặc nhiều thiết bị hỗ trợ chẳng hạn như tivi, điều hòa không khí và xe lăn điện [74]. Mô hình xe lăn điện dựa trên các tín hiệu sinh học khác nhau được sử dụng để điều khiển một xe lăn tích hợp trong môi trường thông minh có thể cho phép người khuyết tật điều khiển các thiết bị gia dụng thiết yếu cũng như xe lăn của họ như hình 2.4 [75].



Hình 2.4. Mô hình hệ thống xe lăn đa phương thức.

2.4.4. Xe lăn điện với tính năng tránh chướng ngại vật

An toàn là yếu tố rất quan trọng trong các hệ thống điều hướng xe lăn thông minh. Các nghiên cứu về khả năng sử dụng xe lăn của người khuyết tật đã ước tính rằng khoảng hàng triệu cá nhân có thể sử dụng các hệ thống này với sự hiện diện của hệ thống tự định vị và mô-đun tránh chướng ngại vật như mô tả ở hình 2.5 [76].



Hình 2.5. Hệ thống xe lăn điện với chức năng định vị và tránh vật cản.

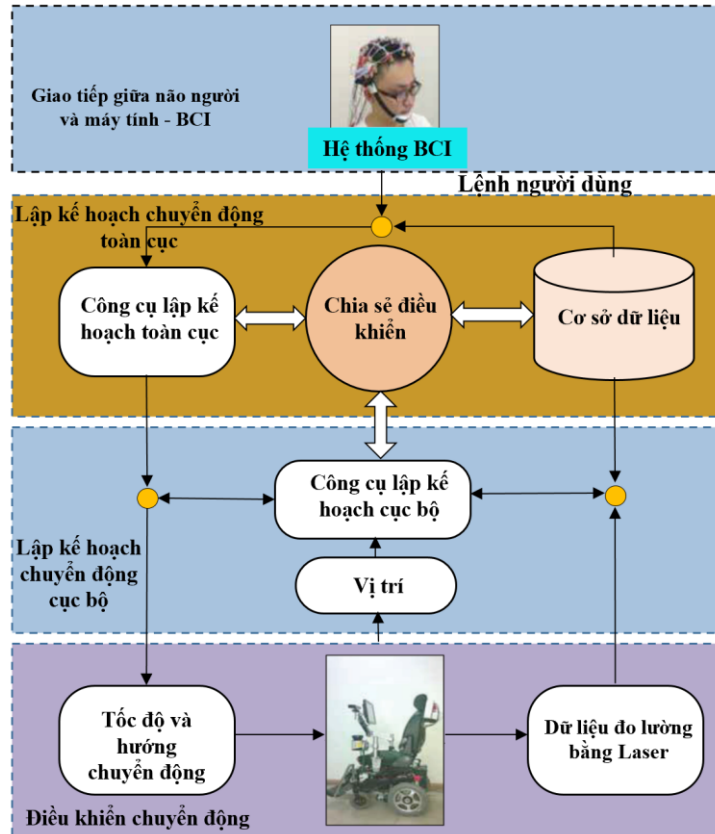
định trước, chẳng hạn như tránh chướng ngại vật hoặc đi theo bức tường. Nhiều xe lăn điện được kết hợp với các phương pháp như DWA [51, 53] hoặc biểu đồ trường vectơ (VFH) [82], được sử dụng khi người dùng phát hiện hoặc kích hoạt các tình huống đã xác định.

Cách tiếp cận thứ hai tương ứng với các hệ thống chia sẻ điều khiển dựa trên mục tiêu [83]. Các hệ thống này trước tiên ước tính một mục tiêu hoặc mục tiêu phụ tiềm năng từ tín hiệu đầu vào của người dùng. Sau đó, một số đường đi khả thi được tạo bởi các bộ quy hoạch toàn cục và cục bộ, và đường đi khả thi nhất được chọn dựa trên thông tin đầu vào của người dùng. Mặc dù hệ thống này có tính đến ý định của người dùng, nhưng xe lăn điện chỉ cần chọn đường dẫn đã chuẩn bị sẵn do người lập kế hoạch tính toán, do đó hạn chế đáng kể quyền điều khiển của người dùng.

Cách tiếp cận thứ ba tương ứng với các hệ thống chia sẻ điều khiển liên tục [84]. Với các loại hệ thống này, hiệu suất của người dùng được đánh giá trực tuyến thông qua các hàm chi phí được xác định trước và quyền điều khiển của người dùng sau đó được xác định bằng kết quả đánh giá trực tuyến. Tuy nhiên, điều khiển xe lăn điện là một quá trình phức tạp, lâu dài, trong đó các hoạt động khác nhau ngay cả trong cùng điều kiện. Tối ưu hóa thời gian thực giới hạn hiệu suất của người dùng và việc xây dựng các hàm chi phí luôn gặp khó khăn.

Phương pháp chia sẻ điều khiển hai lớp được đề xuất để điều khiển xe lăn thông minh [52, 77]. Lớp đầu tiên là lớp quyết định máy chịu trách nhiệm cho phép hoặc vô hiệu hóa lệnh BCI trong một số bối cảnh nhất định, lớp thứ hai được sử dụng để tạo ra lệnh chuyển động phù hợp với khả năng lái xe lăn, cũng như tình hình nhận thức về các hướng di chuyển tiềm năng trong một môi trường được biết trước. Phương pháp này cung cấp cách điều hướng hệ thống để xác nhận các lệnh người dùng. Bộ điều khiển dùng chung được đề xuất bao gồm hai chế độ điều khiển: điều khiển dựa trên bộ máy và tự trị điều khiển. Trong hình 2.7, kiến trúc của hệ thống chia sẻ điều khiển bao gồm bốn phần chính: 1) Giao diện người và máy, 2) Lập kế hoạch chuyển động toàn cục, 3) Lập kế hoạch chuyển động cục bộ và 4) Điều khiển chuyển động

[77]. Khối điều khiển chuyển động nhận các lệnh từ khối lập kế hoạch chuyển động cục bộ. Bên trong khối điều khiển chuyển động là hệ thống điều khiển vòng kín dùng để truyền động trực tiếp cho động cơ.

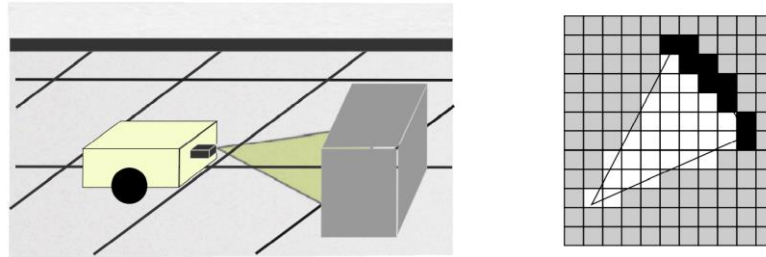


Hình 2.7. Mô hình chia sẻ điều khiển xe lăn điện sử dụng BCI và cảm biến.

2.5. PHƯƠNG PHÁP XÂY DỰNG BẢN ĐỒ LƯỚI 2D ĐỂ ĐIỀU HƯỚNG ROBOT TRONG NHÀ

Bản đồ lưới 2D là một trong những phương pháp biểu diễn môi trường quan trọng trong lĩnh vực robot di động. Bản đồ lưới phân chia không gian thành nhiều lưới với các thuộc tính như không xác định, khoảng trống và có vật cản. Gần đây, các phương pháp được sử dụng để tạo bản đồ lưới cho môi trường quy mô nhỏ 2 chiều đã hoàn thiện [85]. Dựa trên bản đồ này, robot có thể hoàn thành nhiệm vụ điều hướng từ điểm này sang điểm khác. So với các loại bản đồ khác được sử dụng trong SLAM, bản đồ lưới có những ưu điểm sau: (1) bản đồ lưới sử dụng phù hợp với môi trường xây dựng phi cấu trúc, (2) sơ đồ lưới được sử dụng đơn giản, (3) tốc độ lập bản đồ

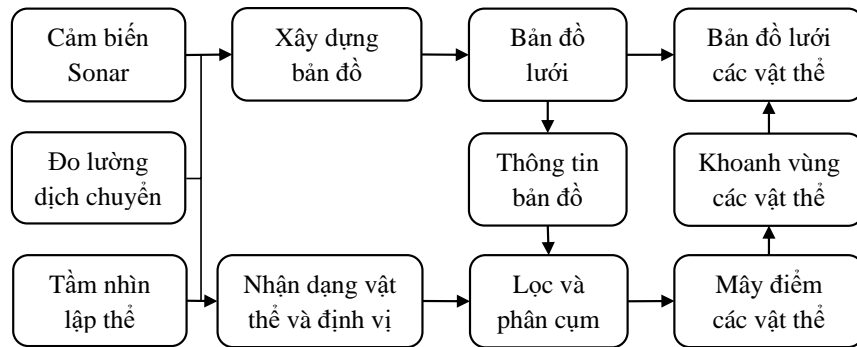
nhanh vì không có đặc điểm môi trường nào được trích xuất và (4) độ phân giải của bản đồ lưới có thể dễ dàng điều chỉnh bằng cách thay đổi kích thước của lưới. Mặt khác, một thiếu sót của bản đồ lưới là nếu chọn độ phân giải bản đồ cao hoặc môi trường lớn thì cần có không gian lưu trữ lớn và khả năng tính toán cao [86].



Hình 2.8. Biểu diễn chùm sonar trên bản đồ lưới 2D.

Hình 2.8 cho thấy cách biểu diễn một phần môi trường trong bản đồ lưới với dữ liệu thu được từ chùm sonar [87]. Ở đó, các ô màu đen biểu thị các đối tượng được sonar phát hiện, các ô màu trắng biểu thị các vùng trống hoặc các vùng không bị chiếm giữ bởi các đối tượng, và cuối cùng, các ô màu xám biểu thị các vùng chưa được ánh xạ. Mô hình không gian xác suất cho các ô lưới được trình bày dưới dạng bản đồ có thể được sử dụng trực tiếp cho các nhiệm vụ điều hướng, chẳng hạn như lập kế hoạch đường đi, tránh chướng ngại vật và ước tính vị trí. Sử dụng cảm biến laser để xây dựng bản đồ dạng lưới và camera để xác định các đối tượng, Meger và cộng sự đã xây dựng một bản đồ chứa cả sự sắp xếp không gian và ngữ nghĩa của đối tượng [88]. Pillai và Leonard đã kết hợp SLAM và nhận dạng đối tượng để tạo ra một bản đồ đám mây điểm với ngữ nghĩa đối tượng [89]. Ruiz-Sarmiento và cộng sự đã sử dụng cảm biến RGB-D để xây dựng bản đồ ngữ nghĩa chứa các mối quan hệ không gian và độ không đảm bảo dựa trên biểu tượng [90]. Pronobis và Jens đã hợp nhất nhiều thông tin cảm biến và sử dụng phương pháp khả năng xác suất để tạo bản đồ nhiều lớp [91]. Phần quan trọng của các phương pháp này là lấy thông tin ngữ nghĩa của các môi trường. Vì vậy, nhận dạng đối tượng, nhận dạng cảnh, cơ sở dữ liệu của mô hình hình học, tương tác giữa người và máy tính, điểm mốc nhân tạo được sử dụng để có được các khái niệm ngữ nghĩa. Trong những năm gần đây, với sự tiến bộ của học sâu, các phương pháp dựa trên mạng nơ-ron chiếm vị trí hàng đầu.

Nghiên cứu của Xianyu Qi mô tả một phương pháp lập bản đồ lưới ngữ nghĩa mới với các cảm biến sonar và stereo camera cho robot gia đình [92] như mô tả hình 2.9. Phương pháp trong nghiên cứu này trực tiếp bổ sung ngữ nghĩa trên bản đồ lưới và xây dựng không gian cấu trúc liên kết của các đối tượng, tận dụng các thuật toán điều hướng cổ điển để điều hướng robot mạnh mẽ và thân thiện với con người.



Hình 2.9. Phương pháp xây dựng bản đồ lưới kết hợp ngữ nghĩa đối tượng.

2.6. CÁC PHƯƠNG PHÁP ĐỊNH VỊ CHO ROBOT DI ĐỘNG

2.6.1. Phương pháp đoán định vị trí

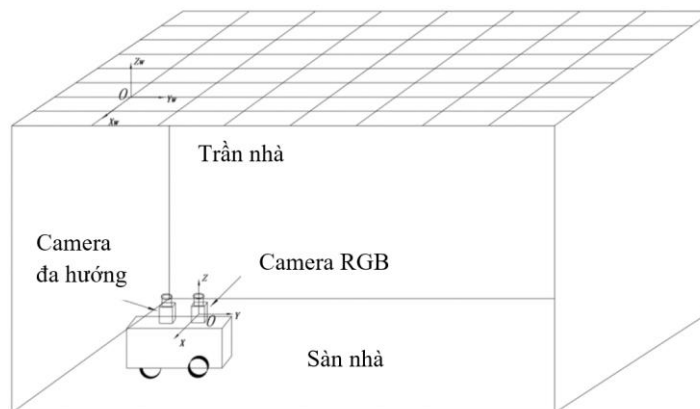
Đoán định vị trí (Dead-Reckoning) là phương pháp dẫn đường được sử dụng rộng rãi nhất [93, 94]. Phương pháp này có độ chính xác cao, mất ít thời gian xử lý và chi phí thấp. Nguyên tắc quan trọng nhất của phương pháp này là tích lũy thông tin về gia tốc chuyển động theo thời gian, từ đó dẫn tới sự tích lũy sai số. Sai số tích lũy dẫn đến sai số vị trí lớn, làm tăng tỉ lệ với khoảng cách chuyển động của robot. Tuy nhiên phương pháp này vẫn được coi là một phần quan trọng của hệ thống dẫn đường của robot. Các lệnh điều hướng sẽ được đơn giản hóa nếu độ chính xác của phương pháp này được cải thiện.

Máy đo tốc độ siêu âm được sử dụng để cải thiện hiệu suất của phương pháp Dead-Reckoning trong môi trường trong nhà [93]. Máy đo tốc độ siêu âm ngăn chặn sự phân kỳ vận tốc của robot di động. Để định vị robot di động trong nhà, máy đo tốc độ siêu âm đo chuyển động tương đối của không khí khi robot di chuyển trong không

khí tĩnh. Dữ liệu vận tốc từ máy đo và dữ liệu gia tốc và vận tốc góc từ cảm biến IMU được kết hợp thông qua bộ lọc Kalman.

2.6.2. Phương pháp định vị robot dùng vật mốc

Các thông tin từ vật mốc mà robot đã được học, thông qua webcam hay camera, bộ điều khiển sẽ sử dụng các công thức toán học được lập trình và tính ra vị trí của robot hoặc khoảng cách từ vật mốc đến robot. Để dự đoán vị trí của robot, một số hệ thống kết hợp các phép đo quãng đường đi được, so khớp các vật mốc với các quan sát về môi trường từ camera [95]. Nghiên cứu của Xuequn Zhang sử dụng các đặc điểm tự nhiên của trần nhà và môi trường xung quanh để ước tính vị trí robot [96]. Một camera RGB và camera đa hướng được gắn trên đỉnh của robot, như thể hiện trong hình 2.10. Trong môi trường trong nhà, trần nhà bao gồm các khối bàn cờ và vị trí robot được tính toán từ các đặc điểm tương đồng trong hai hình ảnh liên tiếp. Hệ thống định vị được đề xuất trong nghiên cứu này bao gồm bốn phần: (1) thiết lập thư viện các vật mốc; (2) nhận dạng các mốc tự nhiên; (3) đo lường trực quan dựa trên các đặc trưng từ trần nhà; và (4) định vị dựa trên vật mốc tự nhiên.

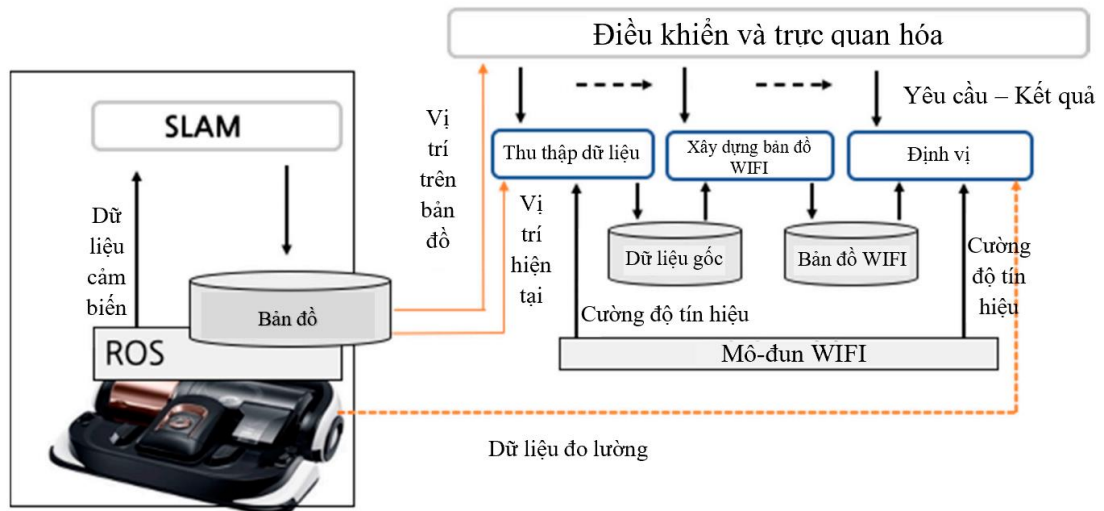


Hình 2.10. Định vị robot dùng các vật mốc trên trần nhà.

2.6.3. Phương pháp định vị cho robot dùng hệ thống WIFI

Định vị robot sử dụng Wifi được dùng nhiều trong các dự án về robot tự hành [39]. Hầu hết các bộ xử lý tích hợp trong robot đều có thể thu phát WIFI, do đó có thể xác định vị trí của robot dựa vào cường độ tín hiệu WIFI mà robot thu về tại các

vị trí khác nhau, từ đó suy ra vị trí của robot. Các robot sẽ định vị chính xác vị trí khi ở trong môi trường trong nhà, vì lúc này cường độ WIFI là mạnh nhất, và không bị ngăn cản bởi tường hay các vật cản. Hình 2.11 cho thấy cấu trúc tổng thể của một hệ thống định vị robot dùng WIFI. Thuật toán theo dõi kết hợp thuật toán Viterbi mở rộng và thông tin đo quãng đường được đề xuất để cải thiện độ chính xác của việc định vị robot.



Hình 2.11. Sử dụng Wifi để định vị vị trí robot.

2.7. CÁC PHƯƠNG PHÁP NHẬN DẠNG VẬT THỂ

2.7.1. Phương pháp nhận dạng dựa trên diện mạo

Phương pháp nhận dạng dựa trên diện mạo sử dụng tập ảnh về hình dạng vật mẫu để so sánh nên phương pháp này có thể nhận dạng vật thể đa dạng hơn [97]. Phương pháp này bao gồm hai giai đoạn thực hiện: 1) *Giai đoạn huấn luyện*: Trong giai đoạn này các tập ảnh mẫu của vật thể được thu thập. Tập ảnh vật thể được chụp dưới nhiều góc nhìn và cường độ, hướng ánh sáng khác nhau hoặc nhiều đối tượng cùng một lớp vật thể. Tập ảnh thường có sự tương quan cao nên để giảm số lượng ảnh trong tập ảnh, phương pháp phân tích thành phần chính (PCA) thường sử dụng [98]; 2) *Giai đoạn nhận dạng*: Các ảnh đầu vào được xử lý và sau đó đem so sánh với tập vật mẫu. Các ảnh này thường có kích thước bằng với ảnh tập mẫu.

2.7.2. Phương pháp nhận dạng dựa trên các điểm đặc trưng

Phương pháp này được xây dựng dựa trên việc so khớp giữa các điểm đặc trưng, là những điểm ảnh nổi bật có cường độ ánh sáng bị gián đoạn. Những điểm này là bất biến dù cho tỉ lệ kích thước hoặc nguồn ánh sáng thay đổi. Harris [99] và FAST [100] là các bộ phát hiện điểm đặc trưng, hay cụ thể là bộ dò góc với tốc độ xử lý nhanh. Thuật toán SIFT [101] bao gồm cả bộ phát hiện đặc trưng và bộ mô tả. Bộ phát hiện đặc trưng trong phương pháp SIFT tìm sự thay đổi về cường độ của các điểm ảnh sử dụng hàm Gaussian. Bộ mô tả của SIFT dựa trên biểu đồ định hướng độ dốc. Thuật toán SURF [102] sử dụng một xấp xỉ nhanh của SIFT. Thuật toán BRISK [103], cũng giống như SIFT và SURF, bao gồm bộ dò tìm điểm đặc trưng và bộ mô tả, với bộ dò tìm điểm đặc trưng là một bộ dò góc và bộ mô tả là một chuỗi nhị phân biểu thị các dấu hiệu của sự khác biệt giữa các cặp điểm ảnh nhất định xung quanh điểm quan tâm. Các bước cơ bản trong phương pháp nhận dạng dựa trên đặc điểm cục bộ bao gồm: Bộ dò tìm điểm đặc trưng, Bộ mô tả, Chỉ số hóa, So khớp, Kiểm tra.

2.7.3. Nhận dạng vật thể theo phương pháp máy học

Nhận dạng vật thể bằng các phương pháp máy học đang cho thấy tốc độ phát triển mang tính cách mạng nhanh chóng trong lĩnh vực thị giác máy tính. Phát hiện vật thể trong ảnh kết hợp phân loại đối tượng và bản địa hóa đối tượng là vấn đề thách thức nhất trong thị giác máy tính. Cụ thể, mục tiêu của kỹ thuật phát hiện này là xác định vị trí của các đối tượng trong một ảnh nhất định được gọi là bản địa hóa đối tượng và sau đó mỗi đối tượng được xếp thuộc loại nào, gọi là phân loại đối tượng. Mạng nơ-ron tích chập theo vùng R-CNN là một hướng tiếp cận tiên phong ứng dụng mô hình học sâu cho bài toán phát hiện vật thể. Mô hình R-CNN (Region-Based Convolutional Network) là sự kết hợp của các đề xuất vùng với mạng tích chập CNN [104]. R-CNN giúp định vị các đối tượng với một mạng học sâu và huấn luyện một mô hình chỉ với một lượng nhỏ dữ liệu có chú thích. Mô hình này đạt được độ chính xác cao khi phát hiện đối tượng bằng cách sử dụng mạng ConvNet sâu để phân loại

các đối tượng được đề xuất. R-CNN có khả năng mở rộng hàng nghìn lớp đối tượng mà không cần dùng các kỹ thuật gán đúng.

Mô hình Fast R-CNN sử dụng một thuật toán huấn luyện để phát hiện đối tượng. Thuật toán này chủ yếu khắc phục những nhược điểm của R-CNN, đồng thời cải thiện tốc độ và độ chính xác [105]. Ưu điểm của Fast R-CNN là chỉ số độ chính xác trung bình (mAP) cao hơn R-CNN. Hơn nữa, mô hình không cần bộ nhớ lưu trữ cho bộ nhớ đệm đặc trưng. Khác với Fast R-CNN và R-CNN, mô hình SSD (Single Shot Detector) sử dụng một phương pháp phát hiện các đối tượng trong hình ảnh với một mạng học sâu duy nhất [106]. SSD loại bỏ hoàn toàn các đề xuất đối tượng và các giai đoạn lấy mẫu lại, đồng thời đưa tất cả các tính toán vào trong một mạng duy nhất. SSD có ưu điểm là tốc độ xử lý nhanh, dễ huấn luyện và dễ tích hợp vào các hệ thống yêu cầu có bộ phát hiện đối tượng. Ngoài ra, SSD có độ chính xác cao tương đương với các phương pháp sử dụng các đề xuất đối tượng bổ sung. Một mô hình phổ biến để phát hiện đối tượng với tốc độ cực kỳ nhanh được nhiều nghiên cứu sử dụng là YOLO [107]. Mô hình YOLO xử lý ảnh trong thời gian thực với tốc độ 45 hình/giây, trong khi đó Fast YOLO với kiến trúc mạng nhỏ hơn, cho tốc độ xử lý 155 hình/giây mà vẫn đạt được gấp đôi mAP so với các mô hình khác.

2.8. MÔ HÌNH HÓA VÀ ĐIỀU KHIỂN XE LĂN ĐIỆN

2.8.1. Mô hình động lực học

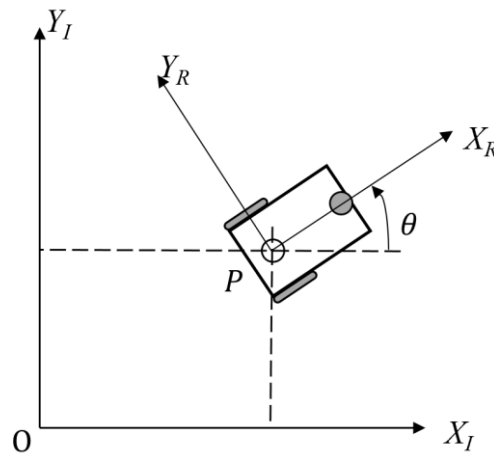
Mô hình động lực học của xe lăn hoặc robot di động là mô hình mô tả cách thức hoạt động của xe lăn hoặc robot di động. Xe lăn hoặc robot có thể di chuyển không bị ràng buộc đối với môi trường, nghĩa là không thể trực tiếp đo lường vị trí của chúng. Vị trí phải được xác định theo thời gian, dẫn đến sự thiếu chính xác trong ước tính vị trí và chuyển động. Mỗi bánh xe góp phần vào chuyển động của xe lăn hoặc robot, và áp đặt các ràng buộc đối với chuyển động. Những ràng buộc này phải được thể hiện thông qua khung tham chiếu (khung cơ khí của xe lăn hoặc robot).

Hệ quy chiếu của xe lăn hoặc robot là ba chiều bao gồm vị trí trên mặt phẳng và hướng như trong hình 2.12, trong đó các trục (X_R , Y_R) xác định hệ quy chiếu so

với điểm P trên khung robot và các trục (X_I, Y_I) xác định hệ quy chiếu toàn cục với gốc tọa độ O [108]. Nếu vị trí của P trong hệ quy chiếu toàn cục được xác định bởi tọa độ (x, y) và góc θ giữa X_I và X_R , vị trí và vận tốc của xe lăn hoặc robot có thể được xác định là một vectơ $[x, y, \theta]^T$ gồm ba thông số này. Phương trình động học của robot di động dẫn động vi sai được mô tả như sau:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 \\ \sin\theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix} \quad (2.6)$$

với ω là vận tốc góc và v là vận tốc chuyển động của xe lăn hoặc robot.



Hình 2.12. Hệ quy chiếu toàn cục và hệ quy chiếu cục bộ của xe lăn hoặc robot.

2.8.2. Điều khiển chuyển động của xe lăn điện

Điều khiển chuyển động của xe lăn điện là quá trình điều khiển vận tốc của các bánh xe để xe lăn di chuyển từ vị trí hiện tại đến vị trí mong muốn. Về cơ bản, có hai phương pháp điều khiển chuyển động như sau:

- Điều khiển vòng hở (điều khiển theo quỹ đạo): Vị trí xe lăn tính toán được không được đưa trở lại để điều khiển vận tốc hoặc vị trí. Vấn đề điều khiển là tính toán trước quỹ đạo để điều khiển xe lăn từ vị trí ban đầu đến vị trí cuối cùng. Phương pháp này đơn giản nhưng có nhiều nhược điểm. Đầu tiên, việc tính toán trước một quỹ đạo khả thi không dễ dàng nếu các hạn chế và ràng buộc của vận tốc và gia tốc của xe lăn được xem xét. Thứ hai,

xe lăn không đáp ứng hoặc bị thay đổi quỹ đạo nếu xảy ra những thay đổi từ môi trường. Cuối cùng, các quỹ đạo thường không trơn tru vì sự chuyển đổi từ một đoạn quỹ đạo này sang một đoạn quỹ đạo khác không liên tục.

- Điều khiển vòng kín (điều khiển phản hồi): Vị trí xe lăn sau khi tính toán được lấy làm phản hồi cho điều khiển vận tốc hoặc vị trí. Vì vậy, mục tiêu của bộ điều khiển chuyển động là ước tính vận tốc của xe lăn để giảm thiểu sai số vị trí.

2.9. PHƯƠNG PHÁP LẬP KẾ HOẠCH ĐƯỜNG ĐI CHO ROBOT DI ĐỘNG

Lập kế hoạch đường đi là nhiệm vụ cơ bản và quan trọng của robot di động. Mục tiêu chính của tìm đường là tìm một quỹ đạo an toàn cho robot di động để di chuyển từ điểm bắt đầu đến điểm kết thúc mà không va chạm với chướng ngại vật. Tùy thuộc vào môi trường, việc tìm đường có thể được phân loại thành: 1) Lập kế hoạch đường đi toàn cục; trong một môi trường tĩnh và có cấu trúc, nơi robot đã biết vị trí của các chướng ngại vật và mô hình của môi trường; 2) Quy hoạch đường đi cục bộ; trong môi trường động, nơi robot khám phá môi trường nhất định bằng cách thực hiện các hành động và sau đó sử dụng thông tin để lập kế hoạch đường đi [109].

2.9.1. Phương pháp A*

Phương pháp A* được sử dụng để lập kế hoạch đường đi nhằm tìm kiếm đường đi ngắn nhất và duy nhất, ước tính khoảng cách tối thiểu giữa hai vị trí để chọn đường đi tối ưu [110]. A* là giải thuật tìm kiếm trong đồ thị, tìm đường đi từ một vị trí hiện tại đến vị trí đích có sử dụng hàm ước lượng khoảng cách hay còn gọi là hàm Heuristic. Từ vị trí hiện tại, A* xây dựng tất cả các đường đi có thể đi và dùng hàm Heuristic để đánh giá đường đi tốt nhất. A* luôn tìm được đường đi ngắn nhất và lưu giữ một tập các đường đi thông qua dạng đồ thị. Thứ tự ưu tiên cho một đường đi được quyết định bởi hàm Heuristic được tính toán như sau:

$$f(n) = g(n) + h(n) \quad (2.7)$$

trong đó, $g(n)$ là chi phí của đường đi từ vị trí hiện tại tới vị trí n , $h(n)$ là hàm ước lượng chi phí từ vị trí hiện tại đến vị trí đích, thường là khoảng cách Euclidean.

Mặc dù thuật toán A^* có nhiều ưu điểm nhưng cũng có một số hạn chế. Đầu tiên là A^* có thể tốn kém về mặt tính toán trong một số tình huống nhất định, đặc biệt khi không gian tìm kiếm rộng và số lượng đường đi có thể lớn. Thuật toán có thể tiêu tốn đáng kể bộ nhớ và tài nguyên xử lý. Một hạn chế khác là A^* phụ thuộc rất nhiều vào chất lượng của hàm Heuristic. Nếu Heuristic được thiết kế kém hoặc không ước tính chính xác khoảng cách đến mục tiêu, thì hiệu suất và tính tối ưu của thuật toán có thể bị ảnh hưởng. Ngoài ra, A^* có thể gặp khó khăn với một số loại biểu đồ hoặc không gian tìm kiếm có cấu trúc bất thường hoặc không thể đoán trước.

2.9.2. Phương pháp học tăng cường

Học tăng cường (RL) là một trong ba kiểu học máy chính bên cạnh học giám sát và học không giám sát. Bản chất của RL là thử và sai, nghĩa là thử đi thử lại và rút ra kinh nghiệm sau mỗi lần thử như vậy. RL đã được áp dụng cho tự điều hướng của robot di động dựa trên thông tin hình ảnh đầu vào và đã đạt được thành công đáng kể [111]. Các tác giả của nghiên cứu này đã phân tích hành vi của robot trong các mê cung tĩnh với hình học phức tạp, vị trí bắt đầu và hướng ngẫu nhiên, trong khi vị trí mục tiêu có thể thay đổi. Kết quả cho thấy phương pháp RL này có thể cho phép robot điều hướng trong môi trường rộng lớn và trực quan, trong đó có vị trí xuất phát và đích thay đổi thường xuyên nhưng bố cục mê cung luôn cố định. Yuke Zhu và các cộng sự đã cố gắng tìm chuỗi hành động với khoảng cách tối thiểu để di chuyển robot từ vị trí hiện tại đến mục tiêu được chỉ định bởi hình ảnh RGB [112]. Điều này có nghĩa là họ phải thu thập một số lượng lớn các hình ảnh khác nhau để xử lý trước khi huấn luyện mô hình điều hướng.

Thuật toán học tăng cường sâu (DRL) đã được áp dụng để giải quyết nhiều thách thức liên quan đến tự động điều khiển [113]. Cụ thể, thuật toán Q-Learning kết hợp với mạng nơ-ron, trong đó DRL giải quyết các vấn đề với không gian quan sát nhiều chiều bằng cách sử dụng mạng nơ-ron để ước tính giá trị Q cho các hành động

tương ứng. Hầu hết, DRL được thực hiện trong môi trường ảo, vì quá trình đào tạo sử dụng phương pháp thử và sai có thể dẫn đến hư hỏng robot trong các nhiệm vụ.

❖ Thuật toán Q-Learning

RL bao gồm 7 khái niệm chính là Agent, Environment, State, Action, Reward, Episode, Policy. Agent sẽ tương tác với Environment bằng các Actions. Sau mỗi Action, Environment trả lại cho Agent một State và Reward tương ứng với State đó. Một loạt các tương tác giữa Agent và Environment từ thời điểm bắt đầu đến khi kết thúc được gọi là một Episode. Trong một Episode, Agent sẽ cố gắng chọn ra các Actions tối ưu để tối đa hóa Reward nhận được sau mỗi Episode. Cách mà Agent chọn những Actions đó gọi là Policy. Mục đích của RL là tìm ra Policy tốt nhất.

Để Agent biết phải chọn Action nào để đạt được Reward lớn nhất, một giá trị gọi là giá trị Q được tính bằng công thức:

$$Q(s, a) = r(s, a) + \gamma \max_a Q(s', a) \quad (2.8)$$

trong đó, $Q(s, a)$ là giá trị Q khi thực hiện Action a tại State s ; $r(s, a)$ là Reward nhận được; s' là State kế tiếp, γ là hệ số discount, đảm bảo càng xa đích thì giá trị Q càng nhỏ [114].

Công thức này cho thấy giá trị Q của Action a tại State s bằng Reward $r(s, a)$ cộng với giá trị Q lớn nhất của các State s' tiếp theo khi thực hiện các Action a . Như vậy có thể tạo ra một ma trận State-Action như một bảng tìm kiếm. Từ đó với mỗi State, Agent chỉ cần tìm Action nào có giá trị Q lớn nhất. Tuy nhiên, RL là một quá trình ngẫu nhiên nên giá trị Q ở thời điểm trước và sau khi thực hiện Action sẽ khác nhau. Khác biệt này gọi là khác biệt tạm thời TD [115], và được tính toán theo công thức:

$$TD(a, s) = R(s, a) + \gamma \max_{a'} Q(s', a') - Q_{t-1}(s, a) \quad (2.9)$$

Như vậy, ma trận $Q(s, a)$ cần phải cập nhật trọng số dựa trên TD :

$$Q_t(s, a) = Q_{t-1}(s, a) + \alpha TD(a, s) \quad (2.10)$$

trong đó, α là tốc độ học. Qua các lần Agent thực hiện Action, $Q(s, a)$ sẽ dần hội tụ.

❖ Thuật toán SARSA

Thuật toán SARSA là một biến thể nhỏ của thuật toán Q-Learning phổ biến [116]. Đối với một Agent trong bất kỳ thuật toán học tăng cường nào, chính sách của nó có thể có hai loại: (1) On Policy: Agent học hàm giá trị theo Action hiện tại xuất phát từ chính sách hiện đang được sử dụng, (2) Off Policy: Agent học hàm giá trị theo Action xuất phát từ chính sách khác.

Q-Learning là một kỹ thuật Off Policy và sử dụng cách tiếp cận tham lam để tìm giá trị Q . Ngược lại, SARSA là On Policy và sử dụng Action được thực hiện bởi chính sách hiện tại để tìm giá trị Q . Sự khác biệt này có thể nhìn thấy trong sự khác biệt của phương trình cập nhật cho Q :

$$Q_t(s_t, a_t) = Q_t(s_t, a_t) + \alpha(r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q_t(s_t, a_t)) \quad (2.11)$$

Phương trình cập nhật cho SARSA phụ thuộc vào State hiện tại, Action hiện tại, Reward nhận được, State tiếp theo và Action tiếp theo. Quan sát này dẫn đến việc đặt tên cho kỹ thuật học tập là SARSA, viết tắt của State – Action – Reward – State – Action, tượng trưng cho tập hợp (s, a, r, s', a') .

❖ Mô hình Deep Q-Learning

Mục đích của RL là chọn ra Action thích hợp cho một State nào đó; hay nói cách khác, State làm đầu vào và đầu ra là một Action. Như vậy có thể thay thế một bảng tìm kiếm $Q(s, a)$ bằng một mạng nơ-ron đơn giản [117]. Mạng nơ-ron sẽ được huấn luyện để học cách ước lượng giá trị Q cho các Action một cách chính xác. Do đó, hàm Loss sẽ là sai số giữa giá trị Q thực tế và dự đoán. Hàm Loss này được viết dưới dạng đầy đủ theo phương trình:

$$Loss = (r + \gamma \max_{a'} Q(s', a', \theta') - Q(s, a, \theta))^2 \quad (2.12)$$

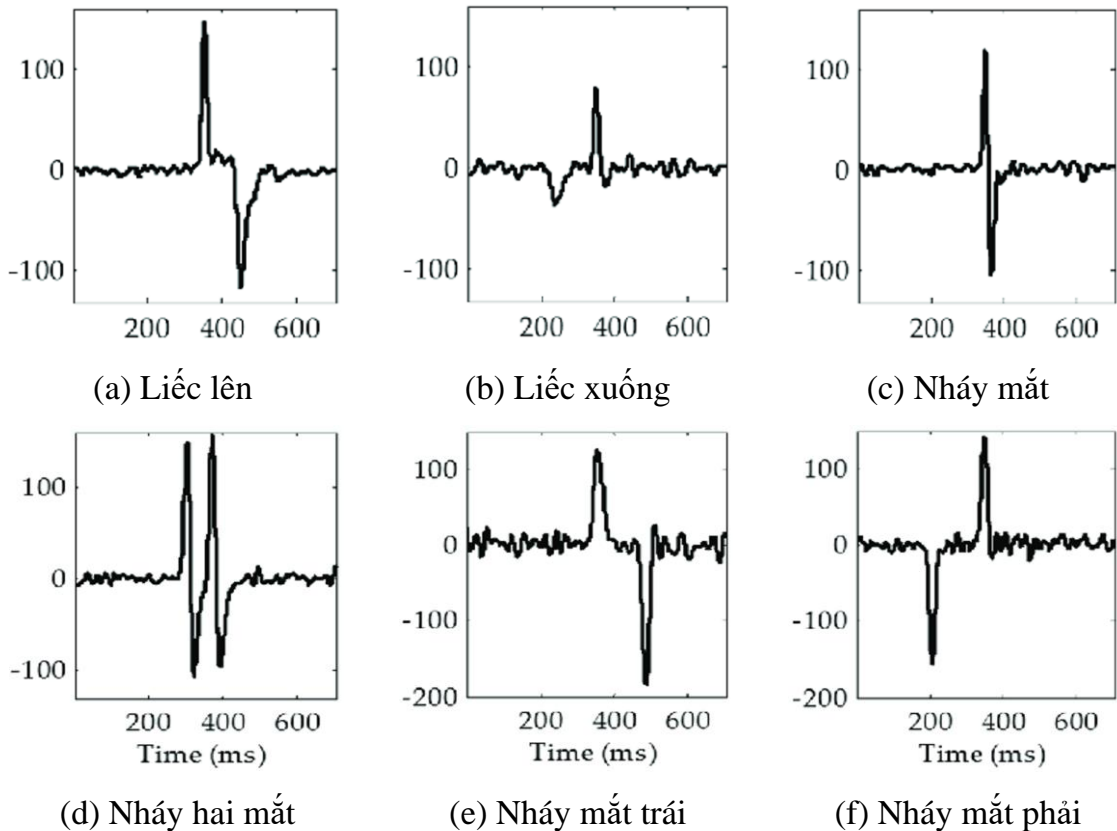
CHƯƠNG 3

PHÂN LOẠI TÍN HIỆU EEG CỦA HOẠT ĐỘNG MẮT CHO ỨNG DỤNG GIAO TIẾP GIỮA NÃO NGƯỜI VÀ MÁY TÍNH

Phương pháp phân loại các hoạt động mắt bao gồm nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt 2 lần liên tiếp và không nháy mắt dựa vào tín hiệu EEG sẽ được trình bày ở chương này. Quá trình thu thập dữ liệu EEG từ thiết bị Emotiv EPOC+ được trình bày bao gồm các phương pháp tiền xử lý tín hiệu, phân tích tín hiệu và gán nhãn dữ liệu. Ngoài ra, các phương pháp phân loại tín hiệu EEG của các hoạt động mắt cũng được đề xuất trong chương này bao gồm phương pháp phân loại theo ngưỡng biên độ và phương pháp phân loại dựa vào mô hình mạng học sâu CNN-1D. Các thí nghiệm cho các trường hợp khác nhau được thực hiện để đánh giá hiệu suất phân loại tín hiệu, từ đó làm cơ sở cho việc xây dựng các ứng dụng giao tiếp giữa não người và máy tính trong các chương tiếp theo.

3.1. TÍN HIỆU EEG CỦA HOẠT ĐỘNG NHÁY MẮT

Tín hiệu EEG của hoạt động nháy mắt thường được xem là nhiễu và được loại bỏ vì hầu hết các nghiên cứu điện não đồ chỉ tập trung vào hoạt động của não [118]. Các hoạt động nháy mắt ảnh hưởng đến dạng sóng Theta và Alpha của EEG, thường ở được thu thập tại các vị trí trước trán gây khó khăn trong vấn đề phân tích dữ liệu khi nghiên cứu các hoạt động của não bộ. Hình 3.1 thể hiện một số tín hiệu cơ bản của chuyển động mắt như liếc lên, liếc xuống, nháy mắt, nháy hai mắt, nháy mắt trái, và nháy mắt phải.



Hình 3.1. Một số tín hiệu cơ bản của chuyển động mắt.

Các tín hiệu nháy mắt được đặc trưng bởi các đỉnh với giá trị điện áp cao hơn tín hiệu EEG thông thường. Tín hiệu nháy mắt thường được phân loại bằng ngưỡng biên độ. Hơn nữa, biên độ của các tín hiệu nháy mắt phụ thuộc vào từng cá nhân cụ thể và có rất nhiều thay đổi khác giữa các đối tượng. Các tín hiệu chuyển động mắt thường có điện áp rất lớn, khoảng gấp 10 lần các tín hiệu EEG thông thường. Vì thế có thể dễ dàng quan sát các tín hiệu EEG của chuyển động mắt bằng mắt thường. Nháy mắt là một chức năng cơ bản của mắt, có cơ chế bán tự động, có thể kích thích có điều kiện hoặc không điều kiện. Với chức năng chính là bảo vệ mắt khỏi các tác nhân từ môi trường, gạt nước mắt loại bỏ các bụi bẩn trên mắt. Có 3 loại nháy mắt là phản xạ nháy mắt, nháy mắt tự phát và nháy mắt tự nguyện.

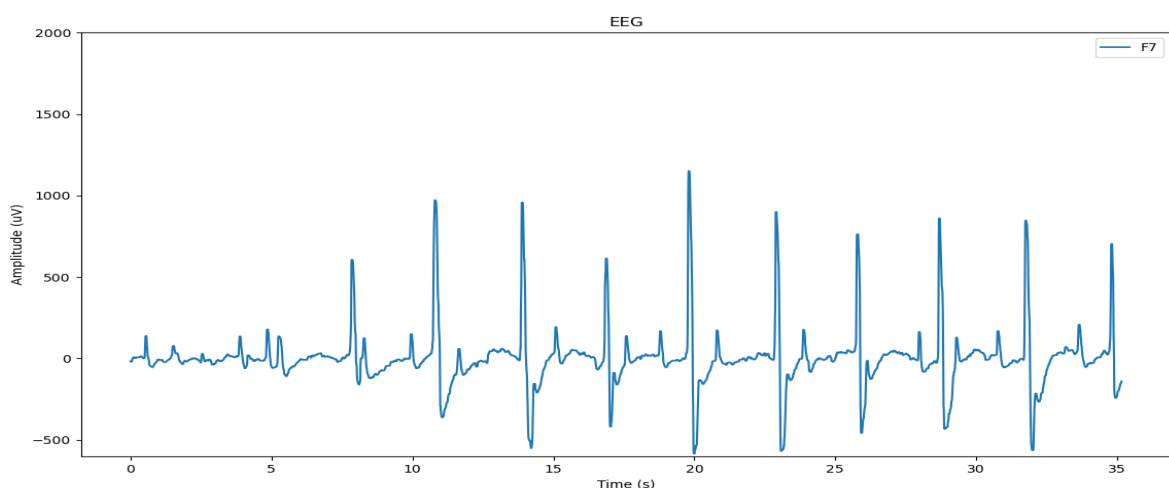
- *Phản xạ nháy mắt:* Phản xạ nháy mắt là phản xạ không điều kiện xảy ra để phản hồi lại với kích thích từ bên ngoài, chẳng hạn như có vật thể di chuyển nhanh qua mắt.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

- *Nháy mắt tự phát*: Nháy mắt tự phát xảy ra không cần kích thích từ bên ngoài hay từ bên trong mà được điều khiển bởi vùng vỏ não tiền vận động, xảy ra mà không cần ý thức giống như thở và co bóp dạ dày.
- *Nháy mắt tự nguyện*: Nháy mắt tự nguyện là hành vi nháy mắt do người thực hiện với mục đích theo ý muốn. Tín hiệu EEG của nháy mắt tự nguyện thường có cường độ lớn hơn so với phản xạ nháy mắt và nháy mắt tự phát.

Tín hiệu nháy mắt là hoạt động của cơ mắt, có độ lớn từ 50 – 3500 uV với tần số hoạt động từ 0,5 -100 Hz. Ngoài ra, độ lớn của tín hiệu nháy mắt thay đổi tuyến tính theo góc nhìn của mắt. Nháy mắt thường được đặc trưng bởi các đỉnh điện áp nên thường được xác định bằng cách đặt ngưỡng và phân loại cho tất cả các hoạt động vượt quá ngưỡng giá trị. Nháy mắt có thể được phân loại là nháy mắt ngắn nếu thời gian nháy mắt nhỏ hơn 200 ms hoặc nháy mắt dài nếu lớn hơn hoặc bằng 200 ms.

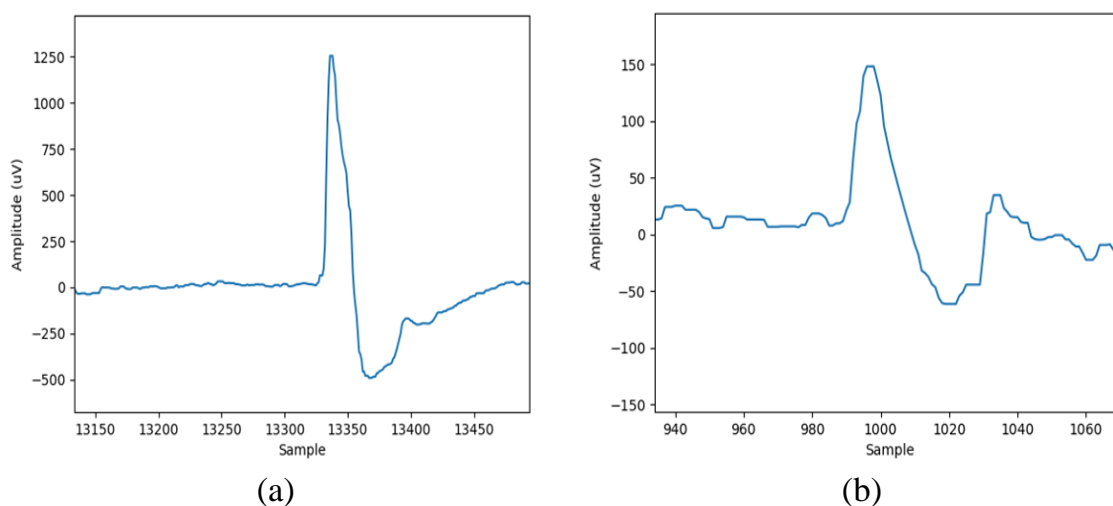
Nhiều trong tín hiệu EEG thường đặc trưng bởi biên độ cao và ảnh hưởng đến việc phân tích tín hiệu. Chẳng hạn như trong quá trình thu EEG, khi đối tượng mở mắt sẽ chứa nhiều tín hiệu nhiễu do nháy mắt. Tuy nhiên trong một số trường hợp, tín hiệu nhiễu nháy mắt có thể có ích. Có thể dùng nhiễu nháy mắt thu được bằng điện cực để xác định được hoạt động nháy mắt. Hình 3.2 mô tả dạng sóng tín hiệu EEG của hoạt động nháy mắt.



Hình 3.2. Tín hiệu EEG khi nháy mắt.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

Tín hiệu EEG khi nháy mắt tự nguyện như trong hình 3.3a thường được sử dụng cho các ứng dụng BCI vì có biên độ lớn hơn nhiều so với tín hiệu EEG khi nháy mắt tự phát như trong hình 3.3b. Nháy mắt tự nguyện có biên độ lớn nhất trong 3 loại nháy mắt với độ lớn trong khoảng -200 uV đến 1200 uV. Ngược lại, tín hiệu nháy mắt tự phát có biên độ nhỏ hơn nhiều so với nháy mắt tự nguyện với biên độ tín hiệu trong khoảng -50 uV đến 200 uV.



Hình 3.3. Hai loại tín hiệu EEG của hoạt động nháy mắt: (a) Tín hiệu nháy mắt tự nguyện, (b) Tín hiệu nháy mắt tự phát.

3.2. THU THẬP DỮ LIỆU

3.2.1. Quy trình thu thập dữ liệu

Thiết bị thu tín hiệu EEG được sử dụng trong nghiên cứu này là Emotiv EPOC+ có kích thước $9 \times 15 \times 15$ cm, khối lượng 170g, hỗ trợ kết nối không dây, như mô tả ở hình 3.4. Emotiv EPOC+ hỗ trợ đo 14 kênh bao gồm AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, với vị trí các điện cực như hình 3.4b. Dữ liệu được thu thập với tần số lấy mẫu 128 Hz. Tín hiệu thu được từ các điện cực sẽ được đưa về bộ chuyển đổi ADC 24-bit và sau đó được truyền tới máy tính.

Các điện cực nằm ở vị trí trước trán (AF3, F7, F8 và AF4) cho tín hiệu EEG có sự thay đổi cường độ rất lớn đối với các chuyển động mắt. Các điện cực AF3 và

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

F7 thể hiện cho tín hiệu của các chuyển động mắt trái, AF4 và F8 tương ứng với các chuyển động của mắt phải. Dữ liệu EEG dùng trong nghiên cứu này được thu từ các đối tượng có sức khỏe tốt, không mắc phải các tật mắt. Tín hiệu được thu lại ở trạng thái tốt nhất của cơ thể, các cá nhân thực hiện thí nghiệm đều ở trạng thái thư giãn trước khi thu. Không gian thu dữ liệu yên tĩnh, không bị gây nhiễu bởi các âm thanh môi trường. Độ tuổi tham gia thí nghiệm là các sinh viên ngành Kỹ Thuật Y Sinh Trường Đại Học Sư Phạm Kỹ Thuật Tp. HCM từ 18 đến 23 tuổi.



Hình 3.4. Thiết bị Emotiv Epoch+: (a) Mũ Emotiv Epoch+, (b) Vị trí các điện cực

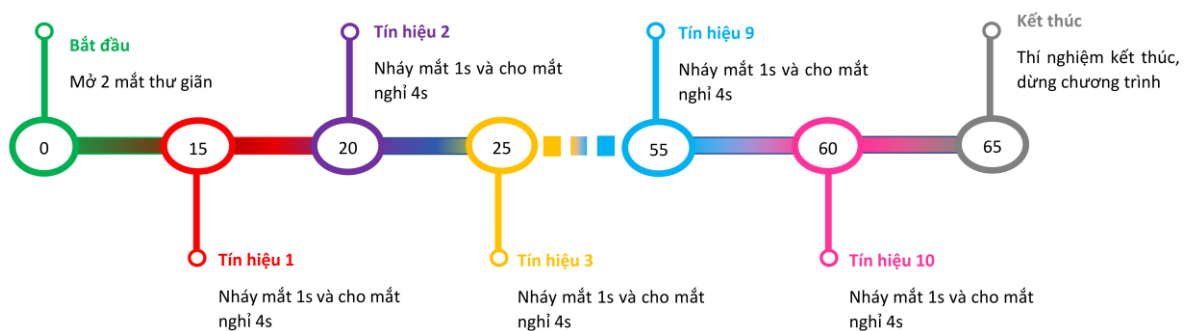
Người thực hiện thí nghiệm sẽ được đội mũ có gắn các điện cực theo đúng vị trí hướng dẫn của phần mềm Emotiv và sau đó tiến hành kết nối kiểm tra chất lượng kết nối tại các vị trí lắp điện cực như mô tả ở hình 3.5. Với độ tương thích khoảng 100%, việc thu dữ liệu theo giao thức sẽ được bắt đầu. Các tình nguyện viên tham gia thí nghiệm được phổ biến rõ yêu cầu và quy trình thí nghiệm. Đồng thời, các tình nguyện viên này đều đồng ý tiến hành thí nghiệm một cách tự nguyện. Các tình nguyện viên sẽ được xem trước các video về hoạt động ghi dữ liệu trước khi tiến hành thí nghiệm. Việc này giúp tín hiệu ghi được ở trạng thái tốt nhất, hạn chế các lỗi trong quy trình thí nghiệm làm ảnh hưởng đến chất lượng dữ liệu.



Hình 3.5. Hướng dẫn đeo thiết bị và kiểm tra kết nối: (a) Hướng dẫn đeo thiết bị Emotiv Epoch+; (b) Kiểm tra chất lượng kết nối

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

Quá trình thu dữ liệu được biểu diễn ở hình 3.6. Trong quá trình thu tín hiệu, người thực hiện thí nghiệm sẽ ngồi im và nhìn thẳng màn hình 15 giây, nhằm mục đích thư giãn mắt và tập trung cho quá trình thực hiện vận động. Thời gian nhìn vào màn hình máy tính được chọn là 15 giây vì nếu nhìn lâu hơn mắt dễ bị khô và mỏi, các tín hiệu thu được cũng chứa nhiều nhiễu hơn do các cử động nháy mắt theo phản xạ làm ảnh hưởng. Các tình nguyện viên sẽ thực hiện 4 hoạt động mắt, bao gồm nháy mắt trái, nháy mắt phải, nháy hai mắt một lần, nháy hai mắt hai lần. Mỗi cử động mắt có thời gian kéo dài là 5 giây và tổng thời gian của thí nghiệm là 65 giây. Cả 4 loại thí nghiệm để thu tín hiệu EEG hoạt động nháy mắt đều được thực hiện cùng một giao thức giống nhau, nhằm thống nhất độ dài của các tín hiệu, phục vụ cho việc xử lý và phân loại tín hiệu sau này.

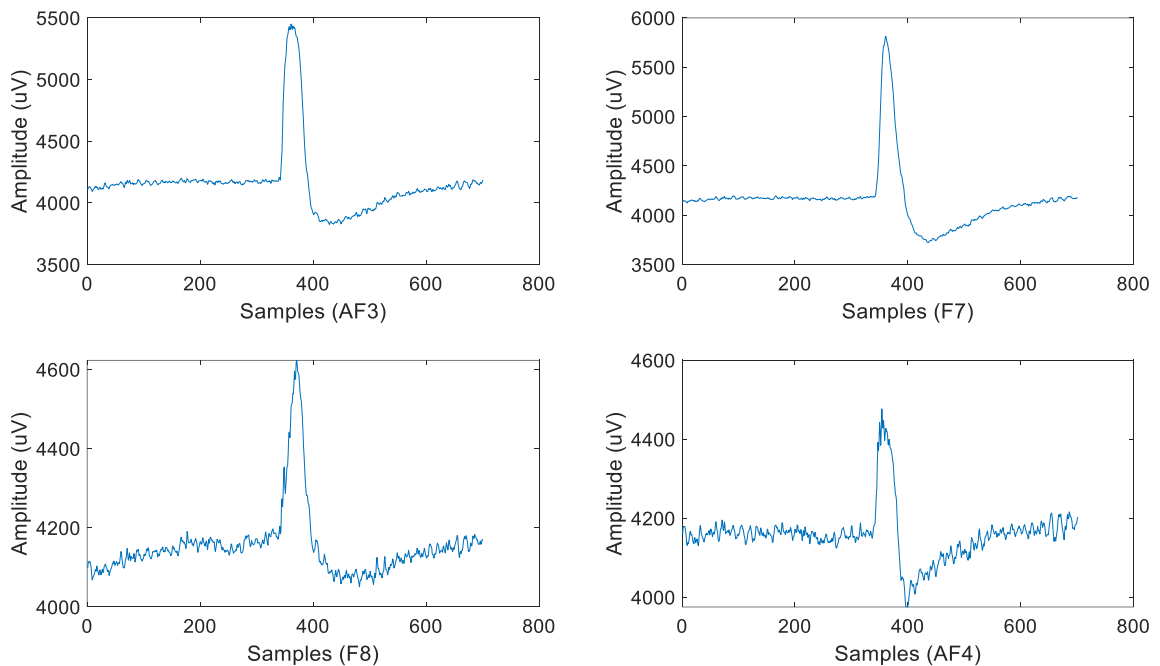


Hình 3.6. Quy trình thực hiện thí nghiệm.

3.2.2. Tập dữ liệu gốc của tín hiệu EEG

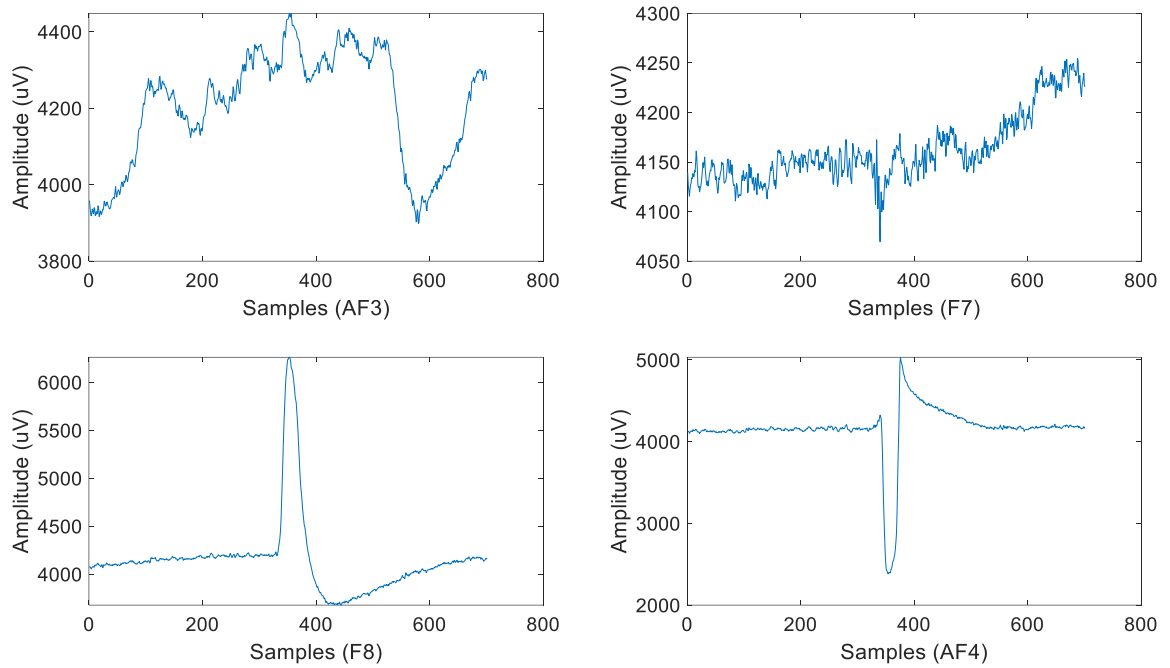
Các tín hiệu chuyển động mắt sẽ lấy từ 4 kênh AF3, F7, F8 và AF4. Trong đó, các tín hiệu nháy của mắt trái thể hiện rõ ở 2 kênh AF3 và F7. Và các tín hiệu nháy mắt phải thể hiện rõ ở F8 và AF4. Mỗi tín hiệu cử động nháy mắt riêng lẻ khoảng 5 giây, độ dài của một tín hiệu là 701 mẫu. Biên độ tín hiệu EEG từ khoảng 200 uV-12000 uV và biên độ này có thể thay đổi theo từng đối tượng, tùy thuộc thể trạng, sức khỏe, điện cực cảm biến và độ mạnh cử động nháy mắt. Trong khoảng 2000 mẫu đầu tiên, người tham gia thí nghiệm ở trạng thái thư giãn và nghỉ ngơi, nên phần tín hiệu sau đó sẽ được sử dụng cho quá trình huấn luyện và phân loại dữ liệu.

Hình 3.7 cho thấy tín hiệu EEG của hoạt động nháy mắt trái thu được từ hai điện cực AF3 và F7 có biên độ cao từ 5000 uV – 6000 uV. Trên hình 3.8 là dạng sóng tín hiệu EEG khi nháy mắt phải. Tín hiệu EEG của hoạt động nháy mắt phải thu được từ hai điện cực F8 và AF4 có biên độ cao từ 4500 uV – 6500 uV. Các tín hiệu gốc thu được vẫn còn chứa nhiều nhiễu, có thể ảnh hưởng đến hình dạng của tín hiệu nháy mắt. Hình 3.9 là dạng sóng tín hiệu EEG khi nháy hai mắt một lần. Tín hiệu EEG của hoạt động nháy hai mắt thu được từ bốn điện cực AF3, F7, F8 và AF4 trong thí nghiệm này có biên độ cao từ 4500 uV – 7000 uV. Dạng sóng của tín hiệu EEG khi nháy hai mắt hai lần liên tiếp được mô tả như hình 3.10. Tín hiệu này cũng có biên độ cao, từ 4400 uV – 5000 uV, và dạng sóng có dạng gai đôi, với điểm kết thúc của gai trong lần nháy mắt đầu là điểm bắt đầu của gai tín hiệu nháy mắt thứ hai. Dạng sóng EEG khi không nháy mắt được mô tả như hình 3.11. Dữ liệu thu được cho 5 tập dữ liệu này có số lượng khoảng 300 tín hiệu trên mỗi loại nháy mắt.

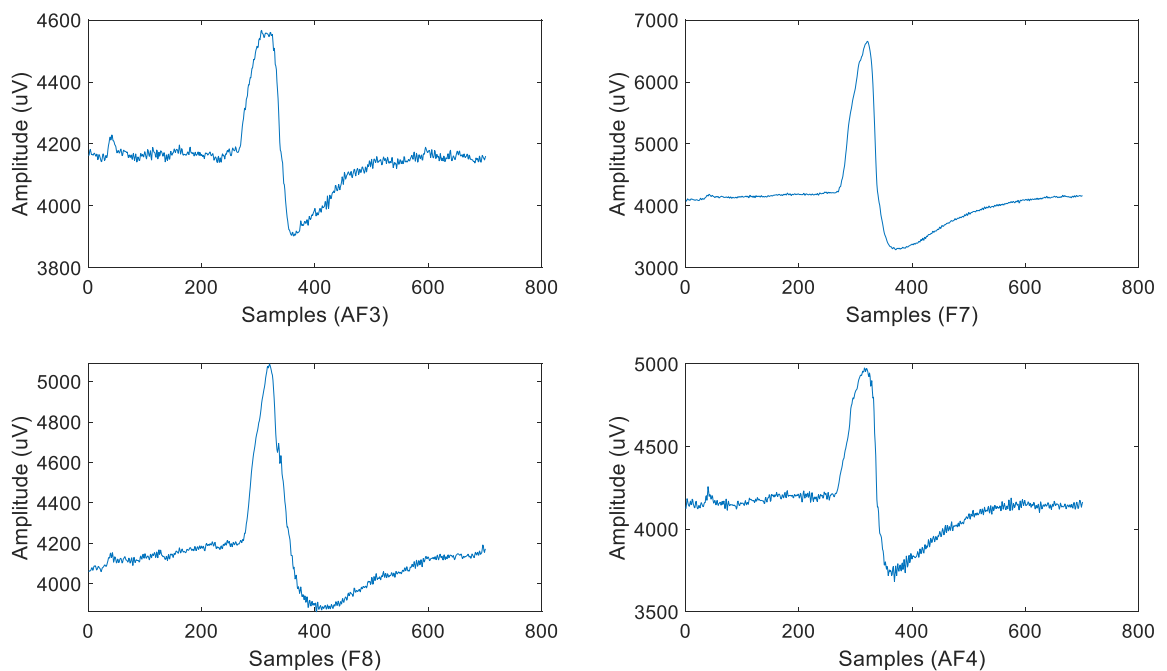


Hình 3.7. Tín hiệu gốc thu được từ 4 kênh AF3, F7, AF4, F8 của hoạt động nháy mắt trái.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

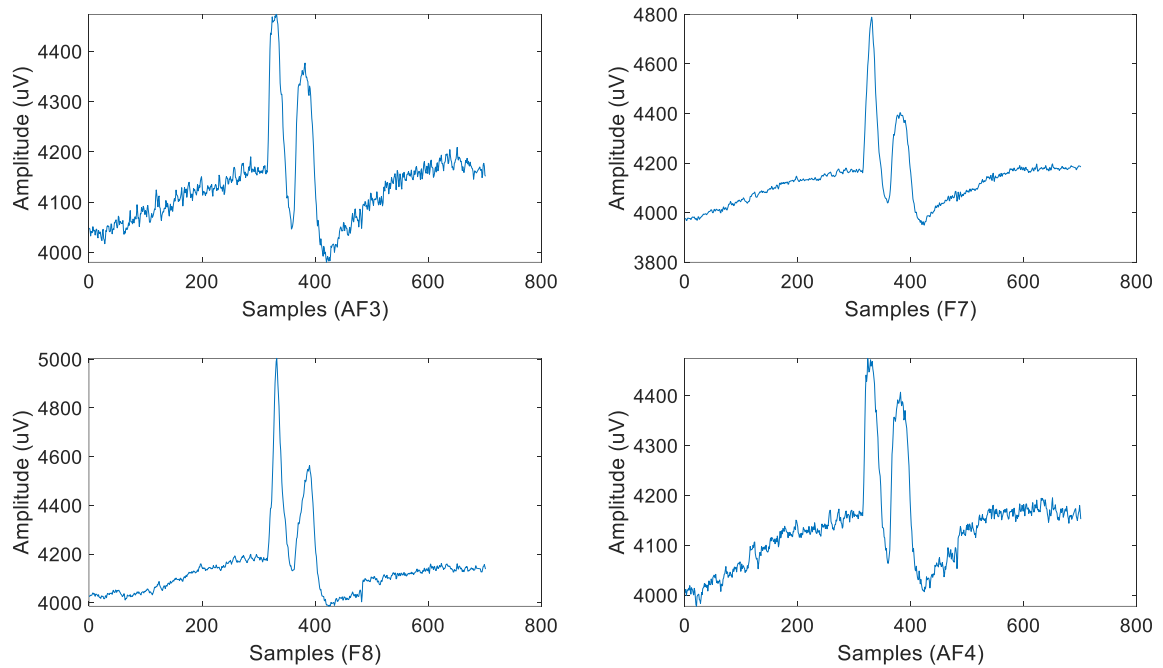


Hình 3.8. Tín hiệu gốc thu được từ 4 kênh AF3, F7, AF4, F8 của hoạt động nháy mắt phải.

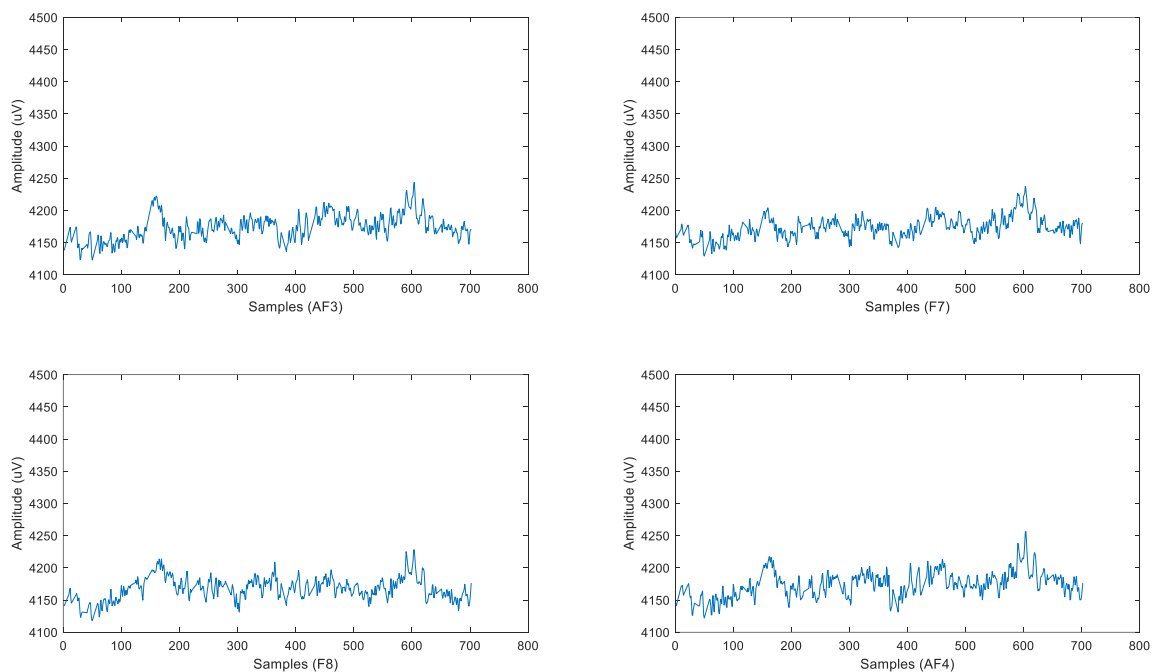


Hình 3.9. Tín hiệu gốc thu được từ 4 kênh AF3, F7, AF4, F8 của hoạt động nháy hai mắt.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.10. Tín hiệu gốc thu được từ 4 kênh AF3, F7, AF4, F8 của hoạt động nháy hai mắt hai lần liên tiếp.



Hình 3.11. Tín hiệu gốc thu được từ 4 kênh AF3, F7, AF4, F8 của hoạt động không nháy mắt.

3.3. XỬ LÝ TÍN HIỆU

3.3.1. Lọc nhiễu dùng bộ lọc Hamming

Các tín hiệu EEG thu được bằng cách đặt các điện cực trên da đầu. Nguyên nhân chính của hiện tượng trôi đường cơ sở và nhiễu trong tín hiệu EEG là do trong quá trình thu nhận tín hiệu EEG, mồ hôi và da ướt gây ra sự tiếp xúc kém của các điện cực và sự thay đổi trở kháng của điện cực. Ngoài ra, các yếu tố khác như hoạt động cơ bắp, môi trường thí nghiệm, dung dịch dẫn điện quá thừa hoặc quá ít, hoặc một số điện cực do sử dụng thời gian dài bị oxi hóa cũng gây ra nhiễu. Vì vậy, các tín hiệu EEG này cần phải được loại bỏ nhiễu.

Tín hiệu EEG của hoạt động nháy mắt thường có tần số thấp, nằm trong phạm vi từ 0.5 – 12 Hz [119, 120]. Tín hiệu EEG gốc sẽ được số hóa ở tần số lấy mẫu 128Hz. Trong nghiên cứu này, một bộ lọc thông dải Hamming được sử dụng với tần số cắt lần lượt là 1 Hz và 10 Hz. Trong tiền xử lý tín hiệu EEG, tín hiệu gốc $x[n]$ sẽ được đưa qua bộ lọc thông dải Hamming và có được tín hiệu đầu ra $g[n]$ như sau:

$$g[n] = h_H[n] * x[n] = \sum_{k=0}^{N-1} h_H[k]x(n-k) \quad (3.1)$$

với $h_H[n]$ là đáp ứng của bộ lọc ($n, k = 0, 1, 2, \dots, N-1$) và N là độ dài của tín hiệu ngõ vào $x[n]$. Đáp ứng xung thực tế của bộ lọc Hamming được tính toán như sau:

$$h_H[n] = \begin{cases} h[n]w[n] & 0 \leq n \leq N-1 \\ 0 & , \text{trường hợp khác} \end{cases} \quad (3.2)$$

trong đó $h[n]$ là đáp ứng xung lý tưởng và $w[n]$ của sổ Hamming [121] được mô tả như sau:

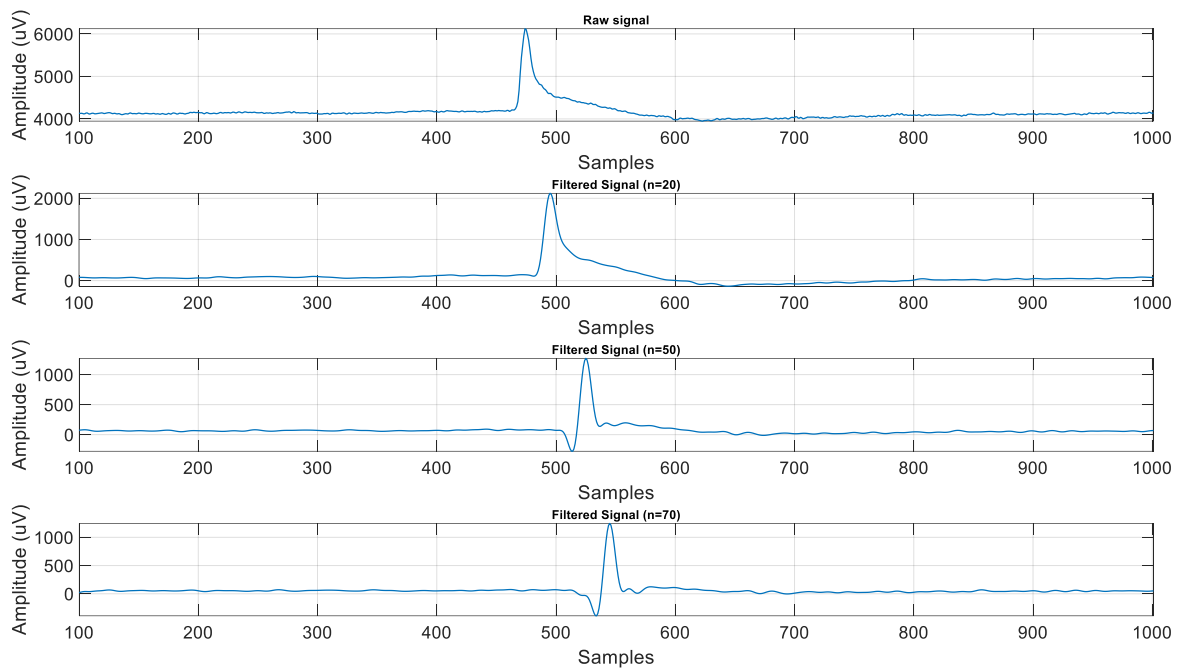
$$w[n] = 0,54 - 0,46 \cos\left(2\pi \frac{n}{N}\right), 0 \leq n \leq N \quad (3.3)$$

Để loại bỏ ảnh hưởng của độ lệch điện áp, tín hiệu sau khi lọc $y[n]$ được tính toán sử dụng công thức sau:

$$y[n] = g[n] - \frac{\sum_{n=0}^{N-1} g[n]}{N} \quad (3.4)$$

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

Hình 3.12 biểu diễn kết quả lọc tín hiệu EEG gốc dùng bộ lọc Hamming. Tín hiệu EEG gốc của hoạt động nháy mắt trái thu được ở kênh F7 có biên độ dao động xung quanh mức 4000 uV. Tín hiệu gốc sau khi lọc với bậc bộ lọc khác nhau đã loại bỏ độ lệch điện áp và cho các kết quả khác nhau. Cụ thể, bộ lọc Hamming với bậc 20 cho dạng tín hiệu sau lọc không thay đổi nhiều so với tín hiệu gốc. Trong khi đó, bộ lọc Hamming với bậc 50 và 70 cho dạng tín hiệu sau lọc giống với dạng tín hiệu nháy mắt trái. Vì vậy, để đảm bảo thu được tín hiệu nháy mắt, bộ lọc Hamming với bậc 50 sẽ được chọn. Tuy nhiên, tín hiệu EEG sau khi lọc sẽ tiếp tục được làm trơn sử dụng bộ lọc Savitzky-Golay nhằm làm có được tín hiệu tốt hơn, giúp tăng hiệu quả của việc phân loại tín hiệu.



Hình 3.12. Tín hiệu EEG ở kênh F7 của hoạt động nháy mắt trái trước và sau khi lọc dùng bộ lọc Hamming.

3.3.2. Làm trơn tín hiệu dùng bộ lọc Savitzky – Golay

Bộ lọc Savitzky–Golay [122] đã được sử dụng để làm cho tín hiệu EEG trở nên mượt mà hơn trong miền thời gian. Bộ lọc này sử dụng một cửa sổ trượt trên các

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

mẫu tín hiệu EEG ngõ vào y_i và kết hợp tuyến tính của một số điểm lân cận để thu được các giá trị điểm z_i ở đầu ra của bộ lọc bằng phương trình sau:

$$z_i = \sum_{n=-n_L}^{n_R} s_n y_{i+n} \quad (3.5)$$

trong đó n_L và n_R lần lượt là số điểm dữ liệu EEG của bên trái và bên phải và s_n là hệ số của bộ lọc.

Khi $n_L = n_R$, thì $s_n = 1/(n_L + n_R + 1)$. Trong bộ lọc Savitzky – Golay, với mỗi điểm y_i , một đa thức gồm $(n_L + n_R + 1)$ điểm (kích thước cửa sổ) trong cửa sổ di chuyển bằng phương pháp bình phương nhỏ nhất và z_i là giá trị của đa thức đó tại vị trí i . Do đó, đa thức bậc M được chuyển đổi từ $a_0 + a_1 i + \dots + a_M j^M$ thành y_{-n_L}, \dots, y_{n_R} . Cụ thể, ma trận thiết kế cho trường hợp này như sau:

$$A_{ij} = i^j \quad (3.6)$$

trong đó $i = -n_L, \dots, n_R$ và $j = 0, 1, 2, \dots, M$.

Để xác định các hệ số đa thức $\mathbf{a} = a_0, a_1, \dots, a_M$, có thể biểu diễn phương trình chuẩn hóa như sau:

$$(\mathbf{A}^T \mathbf{A}) \mathbf{a} = \mathbf{A}^T \mathbf{y} \quad (3.7)$$

với $\mathbf{y} = y_{-n_L}, \dots, y_{n_R}$ là các điểm dữ liệu ngõ vào.

Phương trình chuẩn hóa cho các điểm dữ liệu ngõ vào \mathbf{y} được viết như sau:

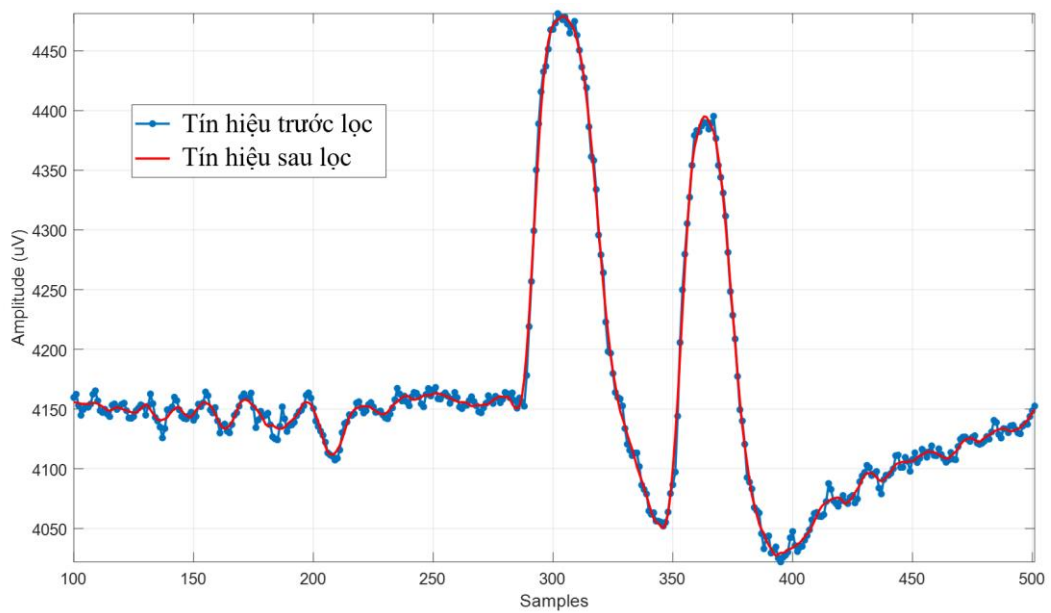
$$\{\mathbf{A}^T \mathbf{y}\}_j = \sum_{k=-n_L}^{n_R} A_{ij} y_k = \sum_{k=-n_L}^{n_R} k^j y_k \quad (3.8)$$

Vì s_n là thành phần a_0 khi \mathbf{y} được thay thế bằng vector đơn vị \mathbf{e}_n , nên s_n được tính theo phương trình sau:

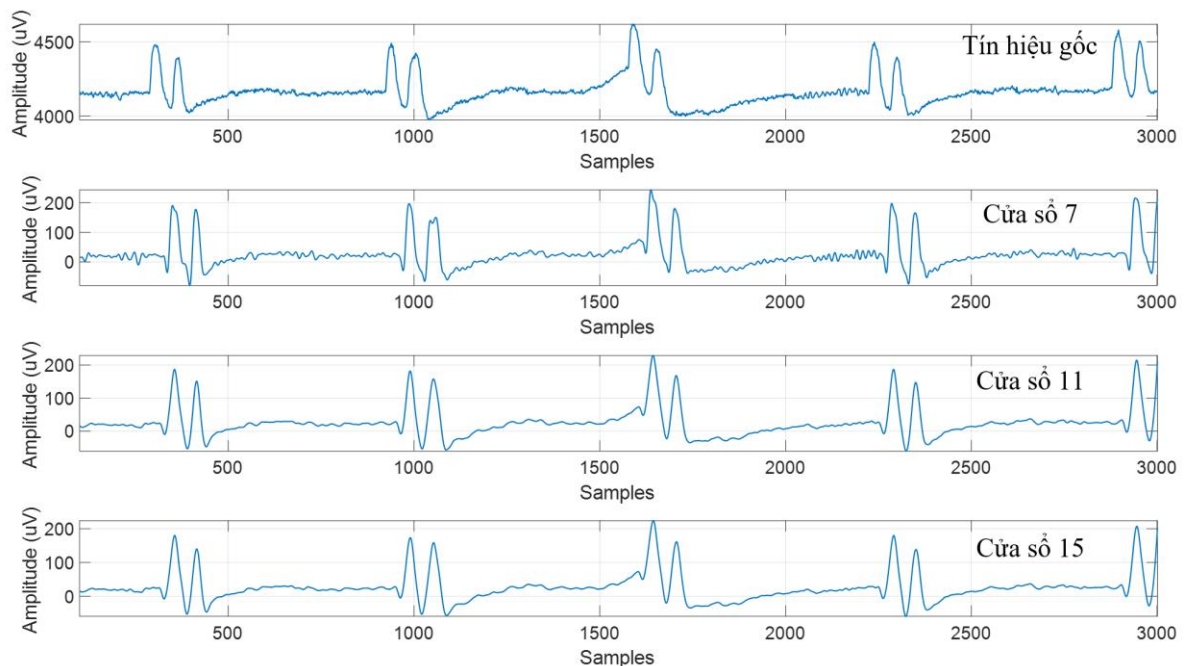
$$s_n = \{(\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{A}^T \mathbf{e}_n)\}_0 = \sum_{m=0}^M \{(\mathbf{A}^T \mathbf{A})^{-1}\}_{0m} n^m \quad (3.9)$$

Tín hiệu EEG trước và sau khi được lọc dùng bộ lọc Savitzky-Golay bậc $M = 2$ và kích thước cửa sổ là 11 ($n_L = 5, n_R = 5$) như hình 3.13. Từ hình 3.13 có thể thấy, dạng tín hiệu EEG sau khi lọc (đường màu đỏ) có dạng sóng trơn tru hơn so với tín hiệu EEG chưa lọc (đường màu xanh).

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.13. Biểu diễn tín hiệu EEG tại kênh F7 trước và sau khi làm trơn bằng bộ lọc Savitzky-Golay.



Hình 3.14. Tín hiệu EEG của hoạt động nháy hai mắt tại kênh F7 được lọc Hamming và làm trơn với bộ lọc Savitzky-Golay.

Hình 3.14 mô tả tín hiệu EEG của hoạt động nháy hai mắt đã được lọc thông dải dùng bộ lọc Hamming và làm trơn dùng bộ lọc Savitzky-Golay bậc $M = 2$, với

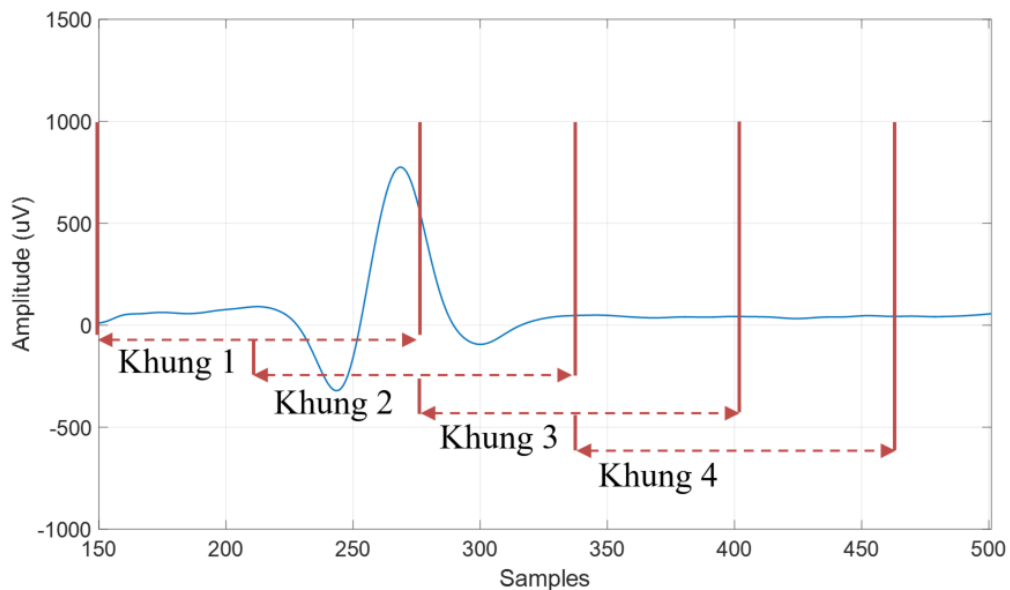
kích thước cửa sổ lần lượt là 7, 11 và 15. Các tín hiệu sau khi lọc được làm tròn một cách đáng kể. Kết quả thu được từ bộ lọc bậc $M = 2$ có kích thước cửa sổ lần lượt là 11 và 15 tốt hơn nhiều so với bộ lọc với kích thước cửa sổ là 7, nhưng tín hiệu sau khi lọc dùng 2 kích thước này không cho thấy sự khác biệt. Vì vậy, để đảm bảo thời gian xử lý và chất lượng của tín hiệu, cửa sổ có kích thước là 11 sẽ được chọn.

3.4. PHÂN LOẠI TÍN HIỆU EEG CỦA HOẠT ĐỘNG MẮT

3.4.1. Phân loại hoạt động nháy mắt theo phương pháp ngưỡng biên độ

❖ Phương pháp ngưỡng biên độ

Việc phát hiện các hoạt động của mắt như nháy mắt dựa vào ngưỡng biên độ tín hiệu EEG cho thấy sự hiệu quả với thời gian xử lý nhanh chóng và không yêu cầu đào tạo người dùng trước khi thực thi thuật toán. Các tín hiệu EEG được lọc và được làm mịn trước khi phát hiện nháy mắt. Tín hiệu nháy mắt sẽ được phân tách thành các khung, trong đó hình 3.15 mô tả cách tín hiệu được chia thành nhiều khung để làm dữ liệu đầu vào thuật toán nhận dạng. Mỗi khung tín hiệu có độ dài 1 giây tương ứng với 128 mẫu tín hiệu. Khung sau sẽ lặp lại khung trước đó 50% để giảm thiểu khả năng mất tín hiệu khi chia tín hiệu để xử lý.

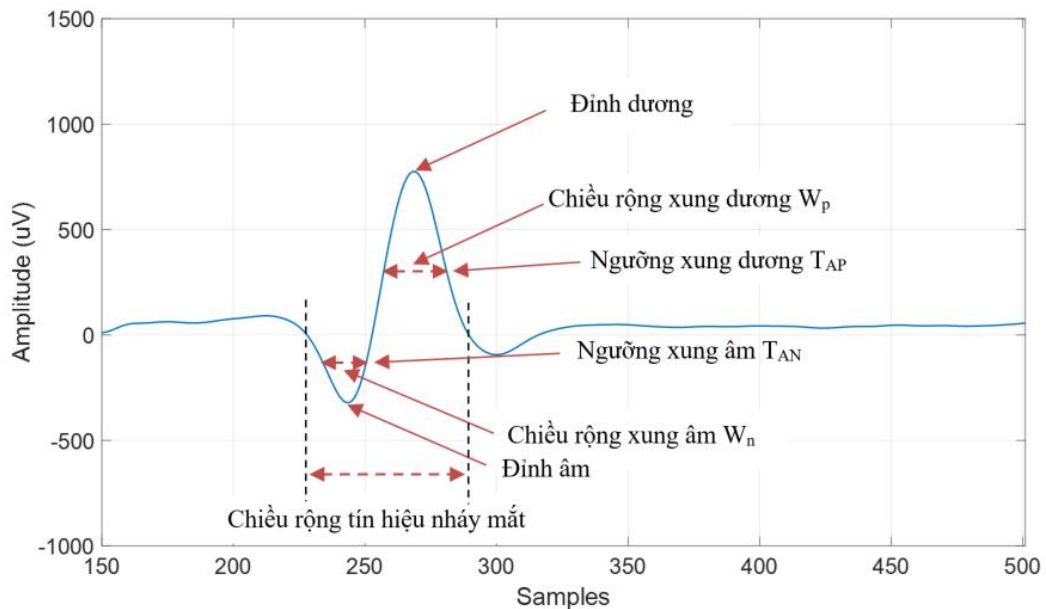


Hình 3.15. Mô tả cách thức chia tín hiệu EEG thành các khung dữ liệu.

Các tín hiệu nháy mắt tự nguyện bao gồm các đặc điểm như được mô tả trong hình 3.16, bao gồm đỉnh dương và đỉnh âm [123]. Chiều rộng của xung dương W_P được tính bằng khoảng cách giữa hai điểm với giá trị biên độ lớn hơn so với ngưỡng T_{AP} . Tương tự, cho việc tính toán chiều rộng xung âm W_N với ngưỡng T_{AN} . Chiều rộng của tín hiệu nháy mắt được tính bằng tổng chiều rộng của các xung âm và dương của tín hiệu. Nếu $Y[n]$ được gọi là tín hiệu EEG của một khung, các ngưỡng T_{AP} và T_{AN} được tính toán bằng các công thức sau:

$$T_{AP} = \frac{\max(Y[n]) + \min(Y[n])}{2}, \text{ với } Y[n] \geq 0 \quad (3.10)$$

$$T_{AN} = \frac{\max(Y[n]) + \min(Y[n])}{2}, \text{ với } Y[n] < 0 \quad (3.11)$$



Hình 3.16. Mô tả tín hiệu nháy mắt tự nguyện.

Thuật toán 3.1 mô tả cách phát hiện hoạt động nháy mắt từ tín hiệu EEG. Với $Y[n]$ là tín hiệu trong một khung, tín hiệu này sẽ được kiểm tra để xem có đủ các đỉnh dương và âm không. Nếu điều kiện này được đáp ứng, nó được gọi là tín hiệu nháy mắt. Để xác định đỉnh dương, chiều rộng của xung dương W_P của $Y[n]$ sẽ được so sánh với ngưỡng dương T_{WP} . Nếu chiều rộng này lớn hơn ngưỡng, xung được gọi là đỉnh dương. Tương tự cho việc xác định các đỉnh âm, trong đó W_N là chiều rộng của xung âm và T_{WN} là ngưỡng âm.

Thuật toán 3.1: Phát hiện hoạt động nháy mắt

```
1: Input:  
   -  $Y[n]$ : Tín hiệu EEG sau khi xử lý cho một khung  
   -  $T_{AP}$ : Ngưỡng biên độ cho xung dương  
   -  $T_{AN}$ : Ngưỡng biên độ cho xung âm  
   -  $T_{WP}$ : Ngưỡng độ rộng cho xung dương  
   -  $T_{WN}$ : Ngưỡng độ rộng cho xung âm  
2: for  $n = 1: N$  do  
3:   if  $Y[n] \geq T_{AP}$  then  
4:      $Timer\_count\_positive ++$   
5:   else if  $Y[n] \leq T_{AN}$  then  
6:      $Timer\_count\_negative ++$   
7:   end for  
8:      $W_P = Timer\_count\_positive$   
      $W_N = Timer\_count\_negative$   
9:   if  $W_P \geq T_{WP}$  then  
10:     $peak^+ = 1$   
11:   end if  
12:   if  $W_N \geq T_{WN}$  and  $peak^+ = 1$  then  
13:    Nháy mắt  
    Else  
    Không nháy mắt  
14:   end if  
15: Output: Hoạt động nháy mắt.
```

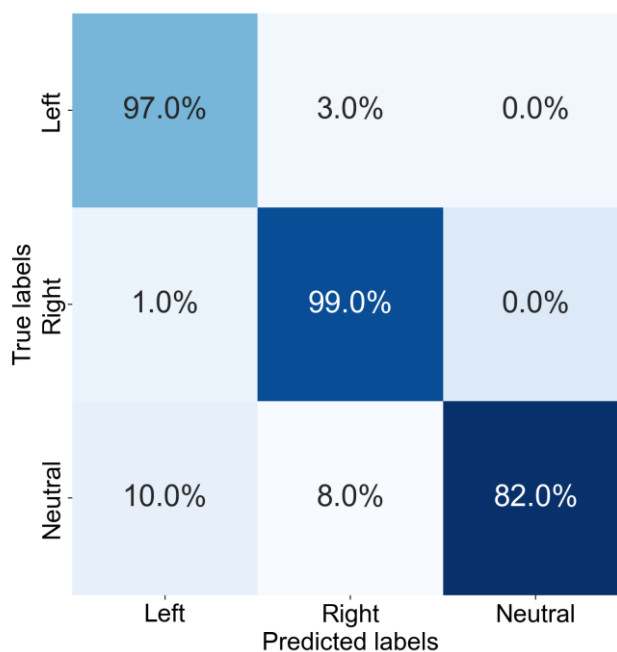
❖ **Kết quả phân loại hoạt động nháy mắt theo phương pháp ngưỡng biên độ**

Sau khi được tiền xử lý, tín hiệu từ hai điện cực F7 và F8 sẽ được sử dụng để xác định tín hiệu nháy mắt trái và phải tương ứng với mỗi kênh. Khi nháy mắt phải, tín hiệu từ kênh F8 nổi bật hơn kênh F7 và ngược lại. Kết hợp các kết quả nhận dạng nháy mắt trong mỗi kênh, có 3 trường hợp có thể xảy ra khi phân loại hoạt động mắt được liệt kê trong bảng 3.1.

Bảng 3.1. Các trường hợp phân loại hoạt động mắt.

TT	Nháy mắt ở kênh F7	Nháy mắt ở kênh F8	Loại hoạt động mắt
1	Có	Không	Nháy mắt trái
2	Không	Có	Nháy mắt phải
3	Không	Không	Không nháy mắt

Các thí nghiệm được thực hiện bởi các tình nguyện viên được sắp xếp ngồi trước màn hình máy tính. Trên màn hình, một giao diện người dùng đồ họa, trong đó có một vòng tròn có thể thay đổi màu sắc và các tình nguyện viên có thể kiểm tra bằng cách nháy mắt trái hoặc phải theo màu tương ứng. Quá trình thí nghiệm có thể kéo dài trong khoảng thời gian 95 giây và lặp lại 3 lần, với thời gian nghỉ giữa mỗi lần khoảng 3 đến 10 phút. Các thí nghiệm trong nghiên cứu này đã được thực hiện trên 10 đối tượng, với sức khỏe tốt, không có tổn thương chi, mắt khỏe mạnh và từ 20 đến 23 tuổi. Ngoài ra, những đối tượng này hiểu rõ về quá trình thí nghiệm và đồng ý tham dự các thí nghiệm này.



Hình 3.17. Kết quả nhận dạng hoạt động mắt.

Sau khi thực hiện các thí nghiệm trên 10 đối tượng, kết quả thống kê được chia thành 3 phân là nháy mắt trái, nháy mắt phải và không nháy mắt như trong hình 3.17. Kết quả cho thấy độ chính xác khi xác định nháy mắt trái là 97% và nháy mắt phải 99%. Tuy nhiên, trong trường hợp không nháy mắt (Neutral) có độ chính xác thấp chỉ với 82%. Nguyên nhân của sự nhận dạng sai là do trong quá trình chuyển động sự tiếp xúc giữa điện cực và da đầu bị giảm, gây ra nhiễu có biên độ lớn tương tự như tín hiệu nháy mắt. Kết quả phân loại nháy mắt này cao hơn kết quả phân loại trong nghiên cứu [57], với độ chính xác lần lượt là 95,1% đối với nháy mắt trái và 96,1% cho nháy mắt phải, tương ứng.

3.4.2. Phân loại tín hiệu hoạt động mắt dùng mô hình CNN-1D

Các phương pháp dựa trên máy học sử dụng các đặc trưng được trích xuất thủ công phụ thuộc vào các tiêu chí [18, 124], trong khi các phương pháp học sâu tốn nhiều thời gian để huấn luyện và yêu cầu nhiều lớp hơn để học các đặc trưng của các dạng sóng EEG phức tạp [61, 62]. Hơn nữa, các phương pháp dựa vào các đặc trưng có thể không khái quát hóa mối quan hệ giữa các đối tượng phân loại và cơ sở dữ liệu do khả năng thay đổi cấu trúc tín hiệu và đặc tính không ổn định của tín hiệu EEG. Vấn đề này có thể được giải quyết bằng cách sử dụng các phương pháp học sâu trong đó các đặc trưng được học tự động từ dữ liệu thô trong quá trình huấn luyện. Một trong những mô hình học sâu phổ biến nhất là CNN-1D, đã được áp dụng thành công cho các vấn đề phân loại khác nhau liên quan đến tín hiệu EEG, bao gồm phát hiện động kinh, nhận dạng cảm xúc và phân loại mức độ công việc trí óc [60, 63]. Bên cạnh đó, ưu điểm của mô hình CNN-1D là có thể hoạt động trên dữ liệu gốc và không tốn nhiều thời gian cho các bước tiền xử lý. Do đó, luận án này đề xuất sử dụng mô hình CNN-1D để huấn luyện và phân loại các hoạt động mắt dựa vào tín hiệu EEG.

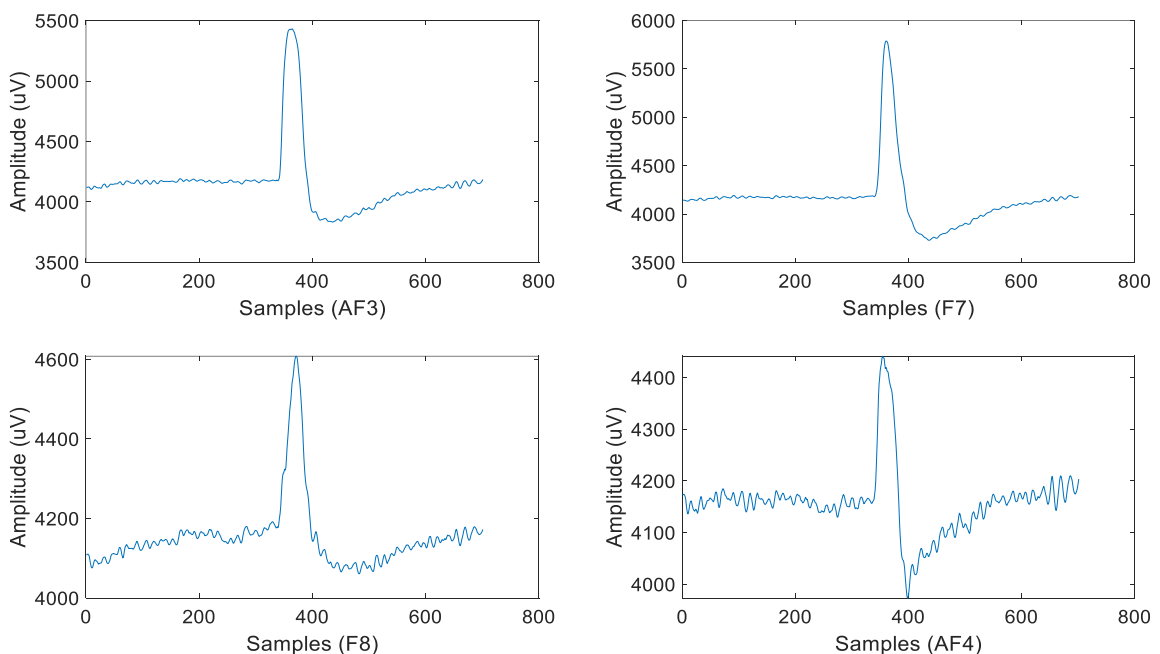
❖ Xây dựng cơ sở dữ liệu

Với dữ liệu thu được từ 2 kênh AF3 và AF4 hoặc F7 và F8, có thể phân loại được các dạng tín hiệu khác nhau. Tuy nhiên, việc sử dụng cùng lúc 4 điện cực, trong

đó sử dụng cùng lúc hai kênh thu từ mỗi một bên mắt, sẽ hạn chế được sai sót của dữ liệu do nhiễu hoặc do độ tiếp xúc điện cực không tốt. Độ dài của mỗi tín hiệu nháy mắt từ một kênh là 701 mẫu. Để tạo ra một tín hiệu lưu vào tập dữ liệu để huấn luyện phân loại hoạt động nháy mắt, đầu tiên các tín hiệu EEG từ 4 kênh AF3, F7, AF4 và F8 của các hoạt động nháy mắt, với mỗi tín hiệu cùng độ dài 701 mẫu, vị trí bắt đầu và kết thúc của các tín hiệu này là giống nhau sẽ được thu thập. Sau đó, các tín hiệu này sẽ được làm trơn bằng bộ lọc Savitzky – Golay và thực hiện ghép lại với nhau tạo thành một tín hiệu, gán nhãn và lưu vào cơ sở dữ liệu.

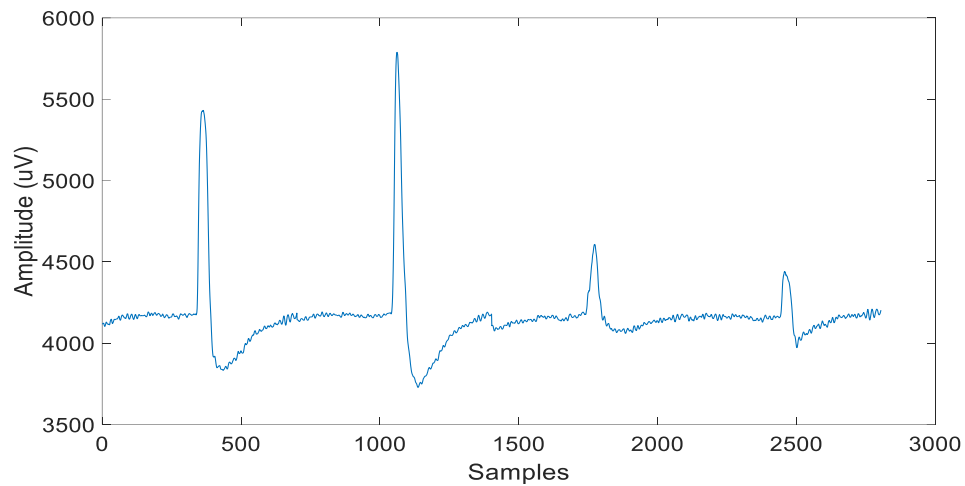
- **Dạng tín hiệu nháy mắt trái**

Một chuyển động nháy mắt trái sẽ cho dạng tín hiệu EEG từ 4 kênh, mỗi tín hiệu có độ dài 701 mẫu như hình 3.18. Hình 3.18 cho thấy sự chênh lệch giữa biên độ đỉnh của 2 tín hiệu từ kênh AF3, F7 (khoảng từ 5500uV đến 5800uV) so với biên độ của 2 tín hiệu AF4, F8 (khoảng từ 4400uV đến 4600uV) tại cùng một thời điểm nháy mắt trái. Hình 3.19 cho thấy tín hiệu được ghép từ 4 kênh AF3, F7, F8, và AF4, có độ dài 2804 mẫu.



Hình 3.18. Tín hiệu nháy mắt trái thu được từ 4 kênh AF3, F7, AF4, F8.

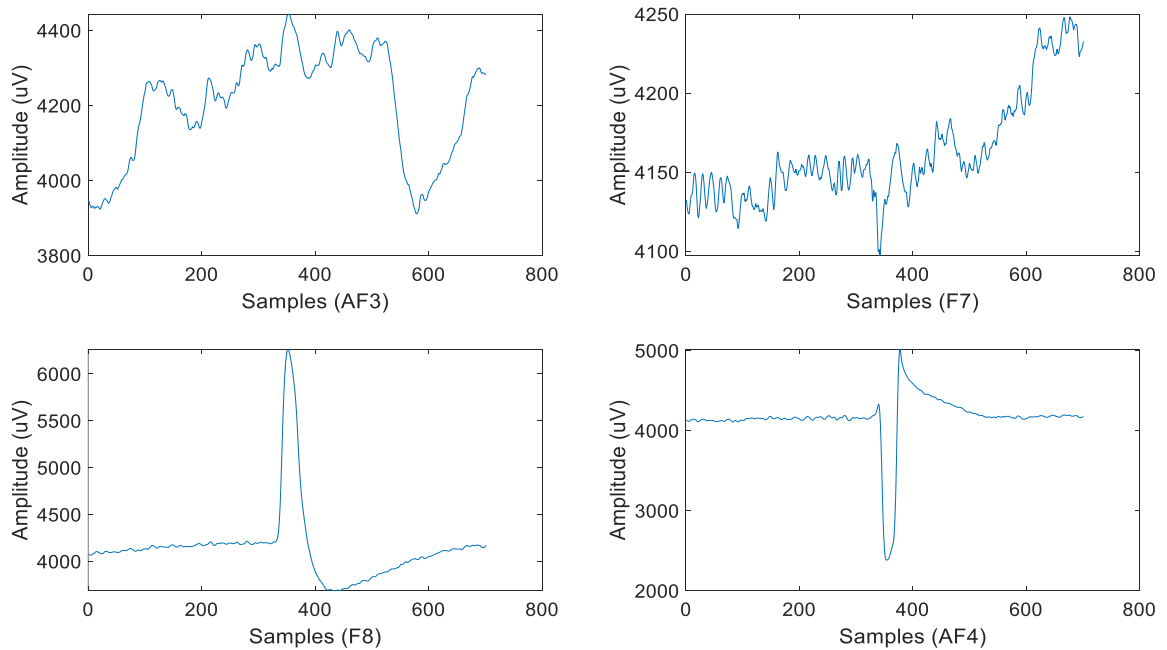
Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.19. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy mắt trái.

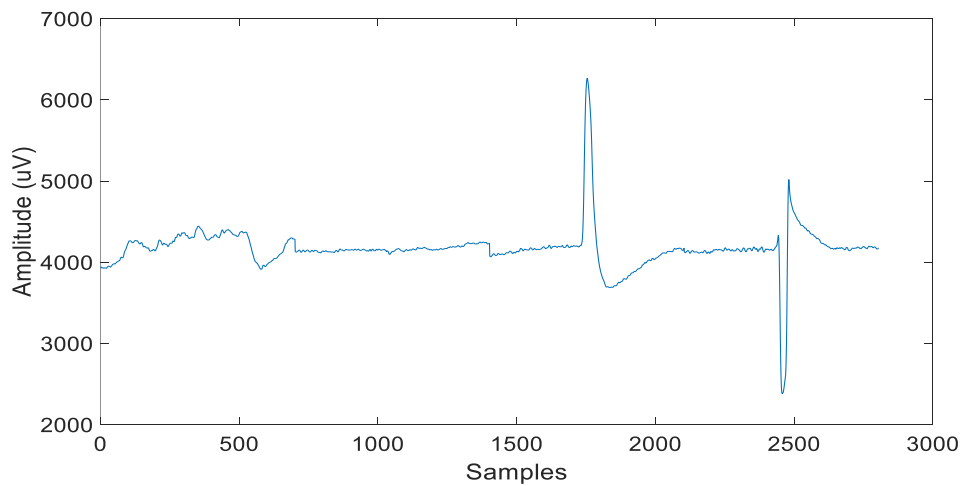
- **Dạng tín hiệu nháy mắt phải**

Một chuyển động nháy mắt phải sẽ cho dạng tín hiệu EEG từ 4 kênh, mỗi tín hiệu có độ dài 701 mẫu như hình 3.20. Kênh F8 thể hiện rõ dạng tín hiệu nháy mắt phải đặc trưng với các biên độ đỉnh sóng cao, khoảng 6300uV, so với 2 kênh AF3 và F7 được thấy ở cùng một thời điểm thực hiện hành động nháy mắt phải. Hình 3.21 cho thấy tín hiệu được ghép từ 4 kênh AF3, F7, F8, và AF4, có độ dài 2804 mẫu.



Hình 3.20. Tín hiệu nháy mắt phải thu được từ 4 kênh AF3, F7, AF4, F8.

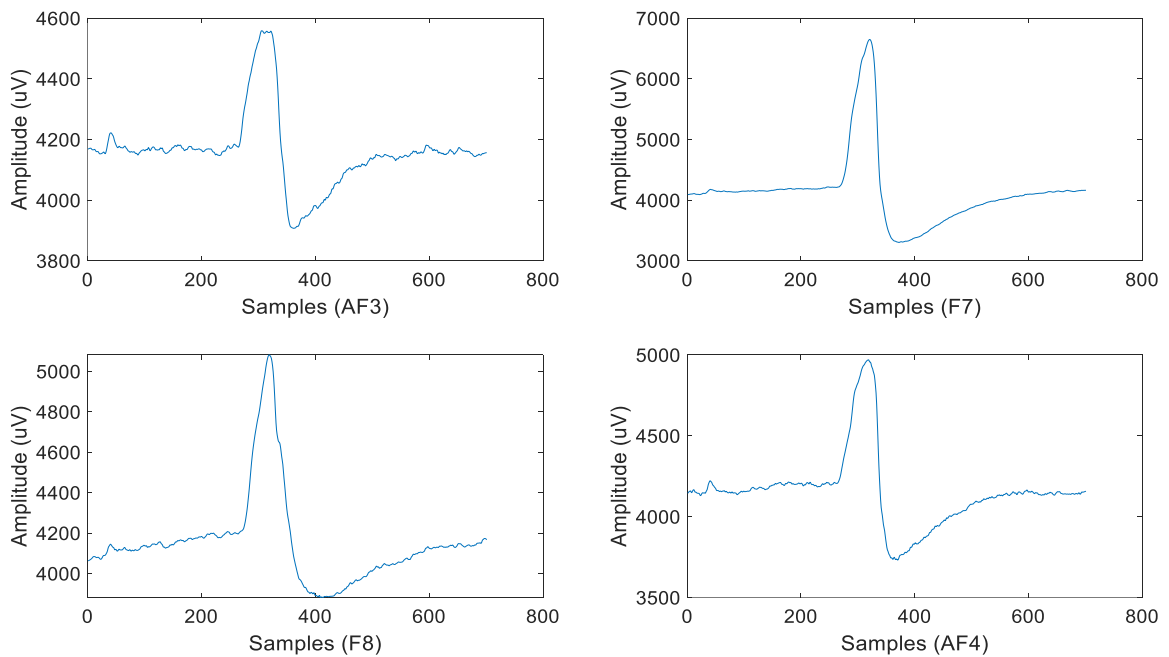
Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.21. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy mắt phải.

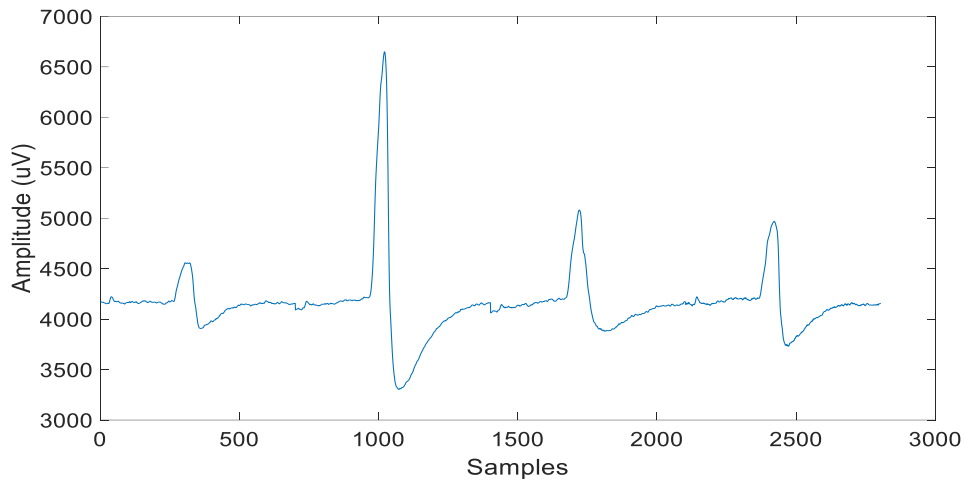
- **Dạng tín hiệu nháy hai mắt cùng lúc**

Dạng tín hiệu nháy hai mắt cùng lúc được trích xuất như hình 3.22. Có thể thấy biên độ tín hiệu EEG ở các kênh AF3, F7, AF4, F8 đều cao hơn so với lúc bình thường. Đây là dạng tín hiệu đặc trưng cho dạng tín hiệu nháy mắt với đỉnh dương cao và một đỉnh âm giảm nhẹ xuống. Hình 3.23 cho thấy tín hiệu được ghép từ 4 kênh AF3, F7, F8, và AF4, có độ dài 2804 mẫu với 4 đỉnh dương rất rõ ràng ở cả 4 kênh.



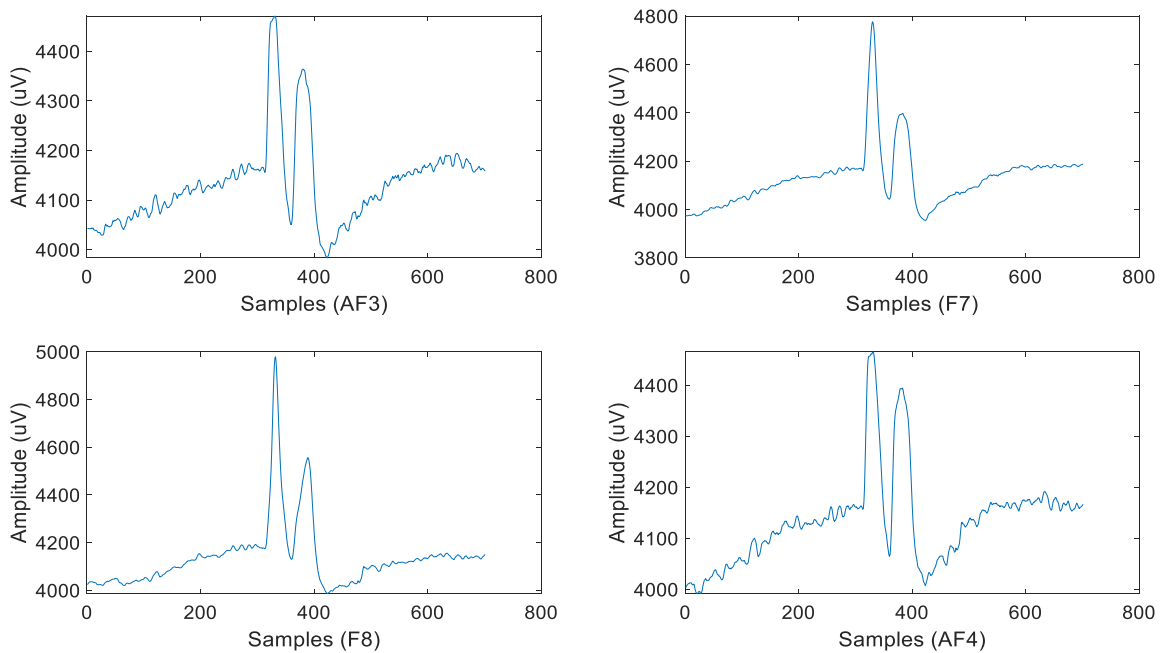
Hình 3.22. Dạng tín hiệu nháy hai mắt được thu ở 4 kênh AF3, F7, AF4, F8.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.23. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy hai mắt.

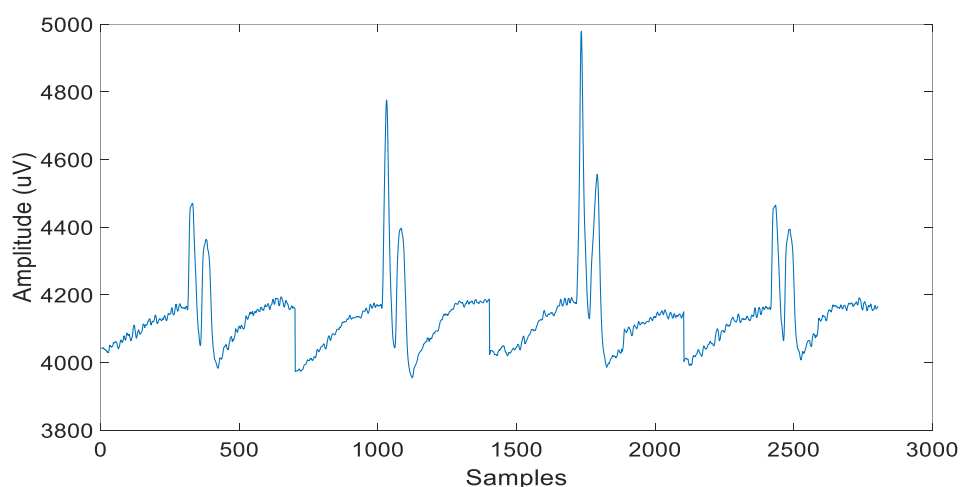
- **Dạng tín hiệu nháy hai mắt hai lần liên tiếp**



Hình 3.24. Dạng dữ liệu nháy hai mắt hai lần liên tiếp được thu ở 4 kênh AF3, F7, AF4, F8.

Tín hiệu nháy hai mắt hai lần liên tiếp thể hiện ở hình 3.24. Hình 3.24 cho thấy có hai gai sóng dương liên tục trong cùng một tín hiệu EEG. Một số nghiên cứu thực hiện hoạt động nháy một mắt hai lần cách xa nhau để thay thế cho hoạt động nháy

hai mắt hai lần liên tiếp [18]. Tuy nhiên, việc nháy mắt 2 lần cách xa nhau có thể gây nhầm lẫn với hoạt động nháy mắt 1 lần. Hơn nữa, với thói quen phản xạ mắt hoặc các yếu tố bên ngoài khác, cũng có thể gây nhầm lẫn trong việc phân loại hoạt động nháy mắt. Do đó, hoạt động nháy mắt hai lần liên tiếp được chọn để thu thập dữ liệu. Đối với hoạt động này, người thực hiện thí nghiệm sẽ phải nháy hai mắt liên tục hai lần trong 5 giây. Hình 3.25 cho thấy tín hiệu được ghép từ 4 kênh AF3, F7, F8, và AF4, có độ dài 2804 mẫu với 8 đỉnh dương rất rõ ràng ở cả 4 kênh.

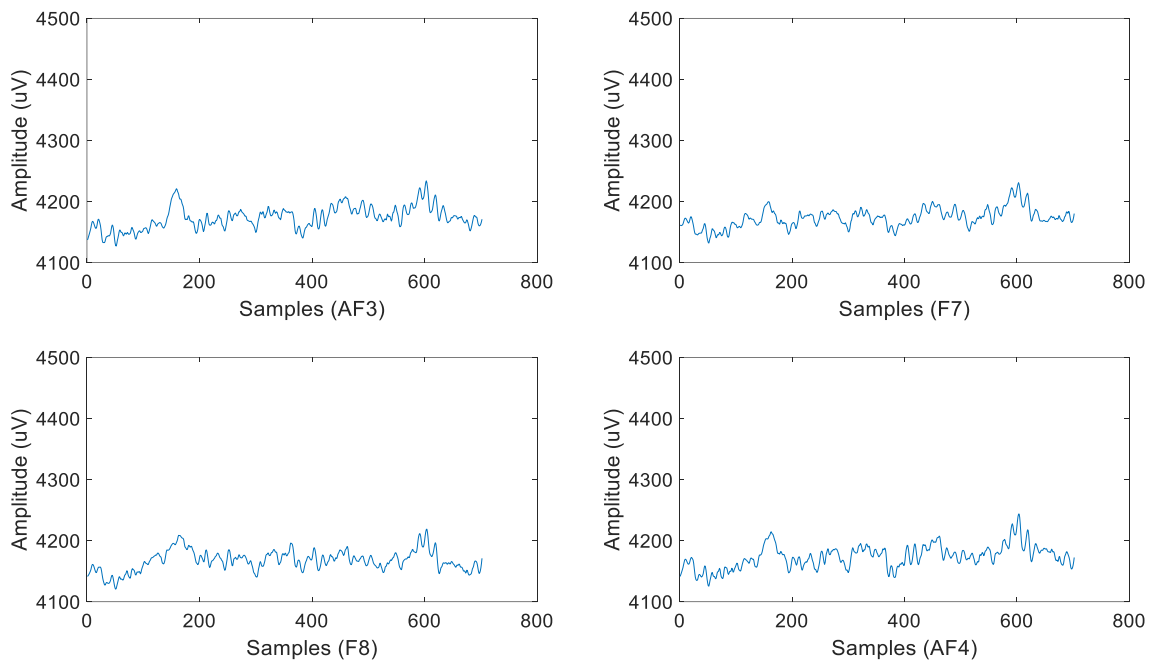


Hình 3.25. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động nháy hai mắt hai lần liên tiếp.

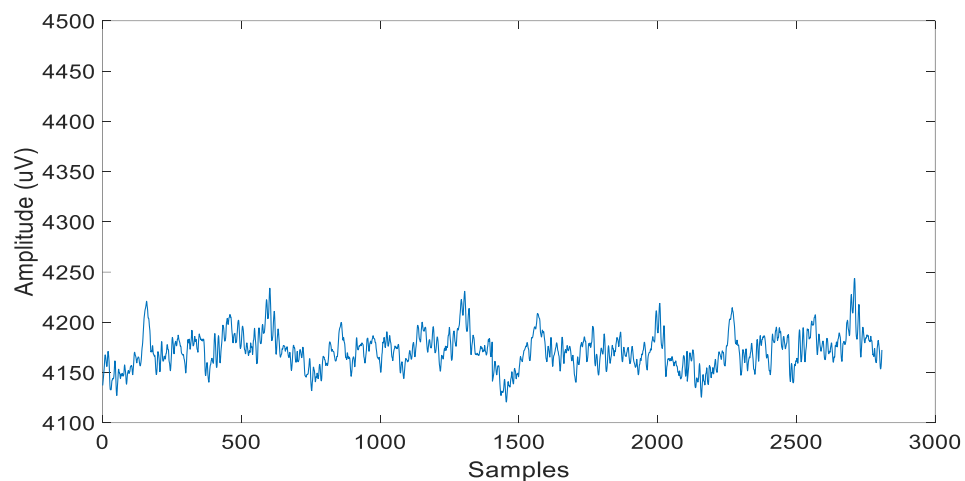
- **Dạng tín hiệu không nháy mắt**

Trong quá trình phân loại các hoạt động mắt, hoạt động không nháy mắt cũng cần được thu thập. Đối với hoạt động này, người thực hiện thí nghiệm sẽ thư giãn và mở hai mắt liên tục trong 15 giây. Sau khi thu thập được tín hiệu không nháy mắt, dữ liệu sẽ được chia ra thành các khung, mỗi khung tương ứng với một tín hiệu nháy mắt mới với độ dài theo thời gian là 5 giây. Hình 3.26 cho thấy tín hiệu EEG của hoạt động không nháy mắt thu được từ 4 kênh AF3, F7, F8, và AF4 sau khi được trích xuất và làm trơn. Hình 3.27 là tín hiệu được ghép từ 4 kênh AF3, F7, F8, và AF4, có độ dài 2804 mẫu.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.26. Dạng dữ liệu không nháy mắt được thu ở 4 kênh AF3, F7, AF4, F8.



Hình 3.27. Tín hiệu ghép từ 4 kênh độ dài 2804 mẫu của hoạt động không nháy mắt.

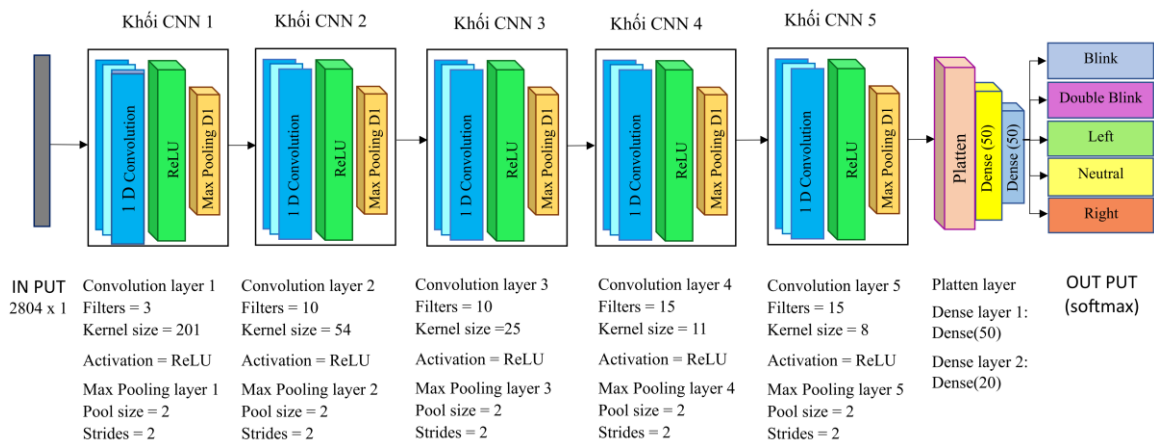
Sau quá trình chọn lọc các tín hiệu tốt và loại bỏ các dữ liệu thu lỗi, bộ dữ liệu gồm 1500 tín hiệu trên 5 tập dữ liệu hoạt động mắt sẽ được gán nhãn, phục vụ việc phân tích, huấn luyện trong mô hình phân loại tín hiệu. Bảng 3.2 mô tả cách thức gán nhãn cho các loại dữ liệu. Dữ liệu sau khi gán nhãn được lưu dưới dạng file “.csv”.

Bảng 3.2. Tên và nhãn tín hiệu của các hoạt động mắt.

Tín hiệu	Nhãn dữ liệu
Nháy mắt trái (Left)	L
Nháy mắt phải (Right)	R
Chớp 2 mắt 1 lần (Blink)	B
Chớp 2 mắt 2 lần liên tiếp (Double Blink)	DB
Không nháy mắt (Neutral)	N

❖ **Mô hình CNN-1D**

Luận án này đề xuất một mô hình mạng CNN-1D với cấu trúc như hình 3.28. Hình 3.28 cho thấy biểu diễn cấu trúc CNN-1D với độ dài tín hiệu đầu vào 2804 mẫu, bao gồm năm lớp tích chập, năm lớp max-pooling và ba lớp kết nối đầy đủ. Bước trượt được đặt ở 1 và 2 để tích chập các tín hiệu EEG và thực hiện max-pooling tương ứng. Các tham số và kích thước kernel trong hình 3.28 có được thông qua thử và sai. Các lớp max-pooling kích thước 2 được áp dụng cho toàn mô hình. Sau khi thực hiện qua 5 lớp tích chập và lớp max-pooling, các đặc trưng của tín hiệu được kết nối đầy đủ đến 50 nơ-ron ở Lớp 11 và Lớp 11 được kết nối hoàn toàn với 20 nơ-ron ở Lớp 12. Cuối cùng, Lớp 12 được kết nối với lớp cuối cùng gồm 5 nơ-ron đầu ra, đại diện cho các loại là nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt hai lần liên tiếp và không nháy mắt.



Hình 3.28. Mô hình CNN-1D cho phân loại tín hiệu EEG của hoạt động mắt.

Trong mô hình CNN-1D như mô tả ở hình 3.29, hàm kích hoạt được sử dụng tại các lớp tích chập là hàm ReLU để loại bỏ các giá trị đầu vào nhỏ hơn 0. Việc sử dụng hàm kích hoạt là để đảm bảo tính phi tuyến của mô hình sau khi trải qua nhiều phép toán tuyến tính ở lớp tích chập. Ưu điểm khi sử dụng hàm ReLU là tốc độ hội tụ và tính toán nhanh. Công thức của hàm kích hoạt ReLU như sau:

$$f(x) = \max(0, x) \quad (3.12)$$

Ngoài ra, trước khi các đặc trưng được trích xuất từ các lớp tích chập được đưa vào lớp kết nối đầy đủ (lớp 11), dữ liệu đầu vào đặc trưng sẽ được biến đổi thành một vec-tơ một chiều. Một hệ số dropout được thêm vào giúp mô hình mạng tạo ra sự khác biệt tốt bằng cách giảm quá khớp trong quá trình huấn luyện [42]. Trong mô hình đề xuất này, tỷ lệ dropout là 0,25. Cuối cùng, ở lớp ngõ ra sử dụng kích hoạt softmax để dự đoán tín hiệu EEG đầu vào thuộc về loại nào. Hàm kích hoạt softmax được định nghĩa:

$$p_i = \frac{e^{x_i}}{\sum_1^k e^{x_k}}, \text{ với } i = 1, \dots, k \quad (3.13)$$

trong đó x là đầu vào và các giá trị đầu ra của p_i nằm trong khoảng từ 0 đến 1 với tổng giá trị các ngõ ra bằng 1.

❖ Phương pháp đánh giá mô hình phân loại

Trong phạm vi và mục đích nghiên cứu của luận án này, ma trận nhầm lẫn được sử dụng để đánh giá độ chính xác của mô hình phân loại, và sau đó, đề xuất cấu hình huấn luyện tốt nhất. Ma trận nhầm lẫn là một bảng mô tả mối quan hệ giữa các lớp thực tế và các lớp dự đoán, trong đó có một biến quan sát và một biến dự đoán. Mỗi hàng trong ma trận nhầm lẫn đại diện cho một lớp quan sát, mỗi cột đại diện cho một lớp dự đoán và mỗi ô là số lượng mẫu giao nhau giữa hai lớp. Hình 3.29 biểu diễn một ma trận nhầm lẫn được sử dụng trong hệ thống phân loại hoạt động nháy mắt từ tín hiệu EEG.

		Lớp dự đoán	
		<i>Positive</i>	<i>Negative</i>
Lớp thực tế	<i>Positive</i>	True Positive (TP)	False Negative (FN)
	<i>Negative</i>	False Positive (FP)	True Negative (TN)

Hình 3.29. Ma trận nhầm lẫn.

Từ ma trận nhầm lẫn, rất nhiều chỉ số đánh giá được tính toán. Độ chính xác (ACC) của mô hình được tính như sau:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.14)$$

Tỷ lệ bộ phân loại dự đoán đúng kết quả với thực tế so với tổng số lượng tín hiệu dự đoán (PRE) được tính từ ma trận nhầm lẫn như sau:

$$PRE = \frac{TP}{TP+FP} \quad (3.15)$$

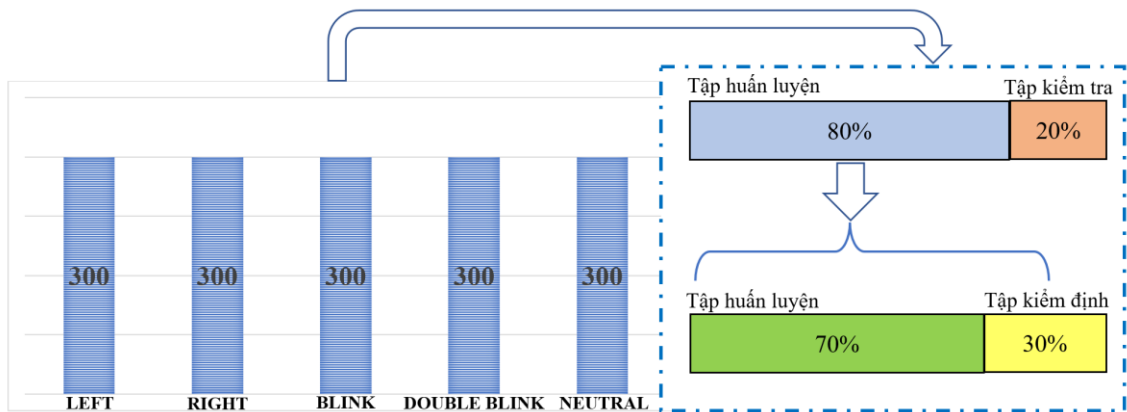
Tỷ lệ bộ phân loại dự đoán đúng một loại so với tổng số lượng tín hiệu thực tế của loại đó (SEN) được tính như sau:

$$SEN = \frac{TP}{TP+FN} \quad (3.16)$$

❖ Kết quả phân loại tín hiệu EEG dùng mô hình CNN-1D

Dữ liệu để huấn luyện mô hình mạng CNN-1D phân loại hoạt động mắt gồm 1500 tín hiệu, trong đó mỗi loại hoạt động mắt có 300 tín hiệu và được chia theo tỷ lệ như mô tả ở hình 3.30. Cụ thể là, dữ liệu sẽ được chia làm 2 tập huấn luyện và kiểm tra theo tỷ lệ 80–20. Trong đó, dữ liệu từ tập huấn luyện sẽ tiếp tục được chia thành 2 tập là huấn luyện và thẩm định theo tỷ lệ 70–30 và đưa vào huấn luyện mô hình. Quá trình huấn luyện sẽ thực hiện lần lượt các thí nghiệm cho trường hợp phân loại 3 loại hoạt động mắt, 4 loại hoạt động mắt và 5 loại hoạt động mắt. Bảng 3.3 mô tả chi tiết sự phân bố dữ liệu dùng cho huấn luyện và kiểm tra trong các trường hợp huấn luyện với các loại hoạt động mắt khác nhau.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.30. Tổ chức dữ liệu để huấn luyện mô hình CNN-1D.

Bảng 3.3. Mô tả dữ liệu huấn luyện phân loại hoạt động mắt.

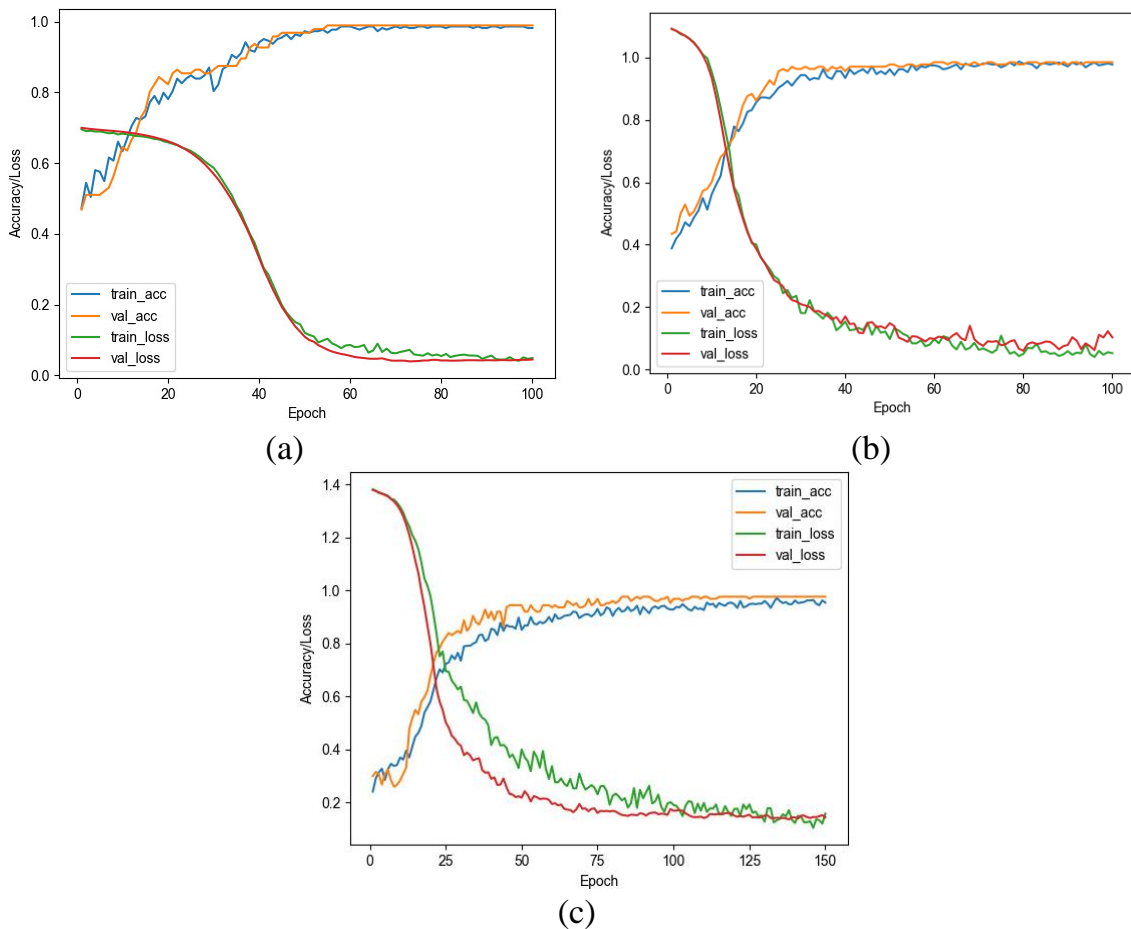
Phân loại hoạt động mắt	Số lượng dữ liệu huấn luyện					Số lượng dữ liệu kiểm tra				
	L	R	B	DB	N	L	R	B	DB	N
L-R-N	240	240			240	60	60			60
L-R-B-N	240	240	240		240	60	60	60		60
L-R-B-DB-N	240	240	240	240	240	60	60	60	60	60

Trong trường hợp huấn luyện phân loại 3 loại hoạt động mắt là nháy mắt trái, nháy mắt phải và không nháy mắt (L-R-N), dữ liệu huấn luyện bao gồm tập 720 tín hiệu, trong đó mỗi loại có 240 tín hiệu. Dữ liệu này được phân chia 70% dữ liệu cho việc huấn luyện và 30% cho việc thẩm định. Dữ liệu sẽ được gán nhãn trước khi đưa vào mô hình, với dữ liệu nháy mắt trái là L, dữ liệu nháy mắt phải là R và dữ liệu không nháy mắt là N. Dữ liệu dùng để kiểm tra mô hình gồm 180 tín hiệu cho 3 loại hoạt động mắt này.

Trong thí nghiệm huấn luyện phân loại 4 loại hoạt động mắt gồm nháy mắt trái, nháy mắt phải, nháy cùng lúc 2 mắt và không nháy mắt (L-R-B-N), số lượng tín hiệu dùng để huấn luyện cho mỗi loại là 240 tín hiệu, số lượng tín hiệu dùng để kiểm tra cho mỗi loại là 60 tín hiệu. Tương tự, với thí nghiệm với phân loại 5 loại hoạt động mắt gồm nháy mắt trái, nháy mắt phải, nháy cùng lúc 2 mắt, nháy cùng lúc 2 mắt 2 lần liên tiếp và không nháy mắt (L-R-B-DB-N), phân bố dữ liệu huấn luyện và

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

kiểm tra cho mỗi loại lần lượt là 240 – 240 – 240 – 240 – 240 tín hiệu và 60 – 60 – 60 – 60 – 60 tín hiệu.



Hình 3.31. Đồ thị biểu diễn hiệu suất huấn luyện mô hình CNN-1D phân loại hoạt động mắt: (a) Ba loại nháy mắt, (b) Bốn loại nháy mắt, (c) Năm loại nháy mắt

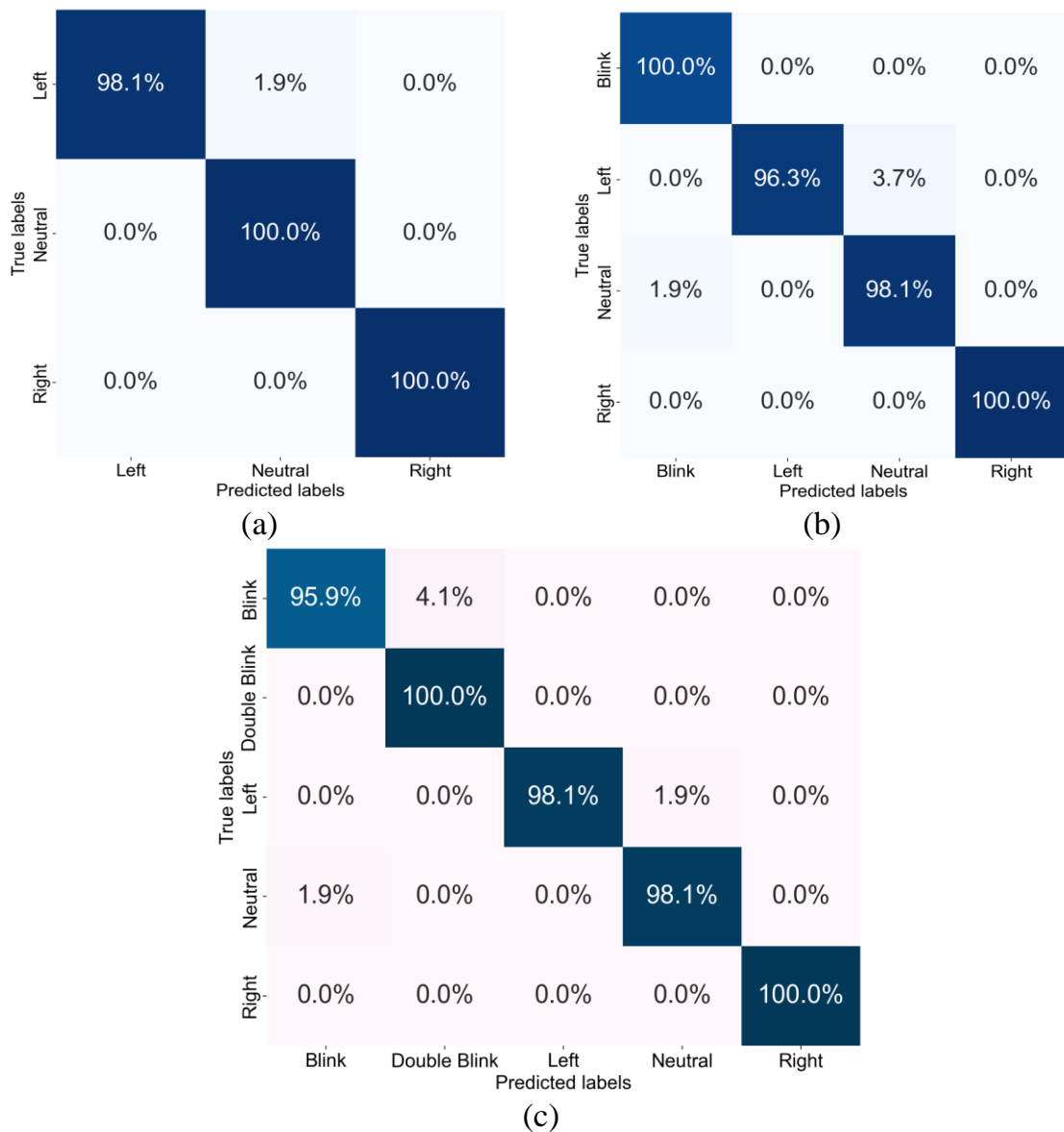
Hình 3.31 mô tả hiệu suất huấn luyện mô hình phân loại cho ba trường hợp nháy mắt trái, nháy mắt phải và không nháy mắt (hình 3.31a); nháy mắt trái, nháy mắt phải, nháy 2 mắt và không nháy mắt (hình 3.31b); nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt hai lần liên tiếp và không nháy mắt (hình 3.31c). Từ hình 3.31 có thể thấy, trong cả ba trường hợp, độ chính xác huấn luyện (đường màu xanh lam) và độ chính xác khi thẩm định (đường màu cam) là tương đồng nhau. Hơn nữa, sai số khi huấn luyện (đường màu xanh lá) và sai số khi thẩm định (đường màu đỏ) cũng ít có sự chênh lệch, cho thấy mô hình huấn luyện đạt chất lượng tốt, không có hiện tượng quá khớp khi huấn luyện các tập dữ liệu.

Bảng 3.4. Hiệu suất mô hình CNN-1D khi huấn luyện phân loại các hoạt động mắt.

Loại tín hiệu EEG	Chu kỳ huấn luyện [Epochs]	Tốc độ học	ACC (%)	PRE (%)	SEN (%)
L – R – N	150	0,0001	99,38	99,39	99,38
			98,15	98,25	98,15
			98,77	98,81	98,77
			98,77	98,81	98,77
			98,77	98,77	98,77
L – R – B – N	150	0,0001	97,16	97,15	97,17
			98,58	98,59	98,61
			98,10	98,11	98,15
			97,16	97,23	97,22
			97,63	97,62	97,64
L – R – B – DB – N	150	0,0001	96,60	96,61	96,63
			98,49	98,50	98,44
			97,74	97,72	97,74
			97,36	97,41	97,33
			97,36	97,50	97,33

Bảng 3.4 mô tả chi tiết hiệu suất của mô hình CNN-1D với các trường hợp khác nhau, mỗi trường hợp thực hiện huấn luyện 5 lần. Trong bảng 3.4, có thể thấy rằng độ chính xác (ACC) trung bình của 5 lần huấn luyện mô hình với tốc độ học 0,0001 thấp nhất là 97,51% đạt được khi huấn luyện phân loại năm loại nháy mắt và cao nhất là 98,77% đạt được khi huấn luyện phân loại hai loại nháy mắt trái và nháy mắt phải. Ngoài ra, các hệ số đánh giá PRE và SEN trung bình cho các trường hợp lần lượt là 98,81% - 98,77% cho ba loại nháy mắt, 97,74% - 97,76% cho bốn loại nháy mắt và 97,55% - 97,50% cho năm loại nháy mắt. Với các chỉ số đánh giá này, có thể thấy mô hình phân loại tín hiệu EEG của các hoạt động mắt cho độ chính xác trên 97% và như vậy trúc mô hình CNN-1D đề xuất với số lượng lớp tích chập và max-pooling là phù hợp cho việc huấn luyện phân loại các loại hoạt động mắt.

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính



Hình 3.32. Kết quả phân loại các hoạt động mắt của mô hình CNN-1D trên tập kiểm tra: (a) Ba loại nháy mắt, (b) Bốn loại nháy mắt, (c) Năm loại nháy mắt.

Kết quả phân loại sử dụng ma trận nhầm lẫn như hình 3.32 cho các trường hợp phân loại 3 loại nháy mắt (hình 3.32a), 4 loại nháy mắt (hình 3.32b) và 5 loại nháy mắt (hình 3.32c). Hình 3.32a cho thấy độ chính xác khi phân loại nháy mắt trái là 98,1%, nháy mắt phải là 100% và không nháy mắt là 100%. Hình 3.32b cho thấy độ chính xác khi phân loại nháy mắt trái là 96,3%, nháy mắt phải là 100%, nháy hai mắt là 100% và không nháy mắt là 98,1%. Với 5 loại nháy mắt, độ chính xác khi phân loại nháy hai mắt đạt 95,9%, nháy hai mắt liên tục hai lần đạt 100%, nháy mắt trái

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

đạt 98,1%, nháy mắt phải đạt 100% và không nháy mắt là 98,1%. Từ các thí nghiệm có thể thấy, việc phân loại nháy mắt phải có độ chính xác cao nhất, tiếp đến là nháy hai mắt liên tục hai lần, vì biên độ của các tín hiệu nháy mắt này thường cao hơn và rõ ràng hơn so với các dạng nháy mắt khác. Ngoài ra, độ chính xác khi phân loại nháy mắt trái đạt tỷ lệ cao, trên 96%.

Trong nghiên cứu này, khi lượng dữ liệu EEG bị hạn chế do khó khăn trong quá trình thu thập dữ liệu EEG. Do đó, trường hợp thu thập dữ liệu riêng từng kênh F7, AF3, F8, AF4 mà ko ghép tín hiệu từ 4 kênh này để tạo thành 1 tín hiệu huấn luyện cũng được thực hiện để so sánh [125]. Hiệu suất khi huấn luyện mô hình phân loại hai loại nháy mắt trái và nháy mắt phải được trình bày ở bảng 3.5. Từ bảng 3.5 có thể thấy, độ chính xác khi huấn luyện đạt tốt nhất với tỷ lệ học là 0,001 lần lượt cho ACC, SEN và PRE là 92,9%, 94,1% và 88,9%. Kết quả này là thấp hơn so với việc ghép 4 kênh tín hiệu F7, AF3, F8, AF4 để tạo thành 1 tín hiệu huấn luyện mô hình phân loại đã trình bày ở trên. Tuy nhiên, độ chính xác này cũng có thể chấp nhận được trong những trường hợp khi số lượng dữ liệu huấn luyện ít.

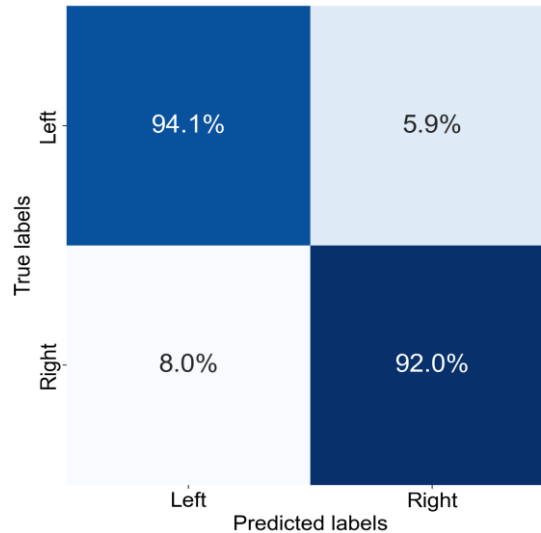
Bảng 3.5. Xác thực chéo 5 lần đối với bộ phân loại hoạt động nháy mắt.

	Lần thực hiện				
	1	2	3	4	5
ACC (%)	92,9	92,9	92,9	90,5	92,9
SEN (%)	94,1	88,9	94,4	89,5	94,1
PRE (%)	88,9	94,1	89,5	89,5	88,9

Hình 3.33 mô tả kết quả của hệ thống phân loại hoạt động nháy mắt trái và nháy mắt phải khi mô hình huấn luyện sử dụng tín hiệu riêng lẻ từ 4 kênh. Từ hình 3.33, mặc dù lượng dữ liệu huấn luyện không nhiều, kết quả phân loại của hệ thống phân loại nháy mắt đề xuất vẫn đạt hiệu suất cao. Trong thí nghiệm này, 42 tín hiệu được sử dụng để đánh giá mô hình phân loại bao gồm 17 tín hiệu nháy mắt trái và 25 tín hiệu nháy mắt phải. Từ hình 3.33, có thể thấy rằng 94,1% tín hiệu EEG của nháy mắt trái được phân loại chính xác là tín hiệu EEG của nháy mắt trái. Hơn nữa, một tỷ

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

lệ nhỏ 5,9% tín hiệu EEG của nháy mắt trái đã bị phân loại sai. Ngoài ra, tỷ lệ cao là 92% tín hiệu EEG của nháy mắt phải được phân loại chính xác, với 8% bị phân loại sai là nháy mắt trái. Độ chính xác trung bình của bộ phân loại đạt 92,9%.



Hình 3.33. Kết quả phân loại các hoạt động mắt với hai loại nháy mắt trái và phải dùng tín hiệu không ghép kênh.

Bảng 3.6 thống kê kết quả của các nghiên cứu phân loại hoạt động mắt dùng EEG trong những năm gần đây. Bằng cách sử dụng ngưỡng cực đại để phát hiện và loại bỏ các hoạt động nháy mắt trong tín hiệu EEG, độ chính xác của thuật toán cho công cụ 'BLINKER' của K. Kleifges và cộng sự [126] là 93,46%. K. Kleifges và cộng sự đã áp dụng bộ lọc thông dải với khoảng tần số 1 - 20 Hz trước khi phát hiện nháy mắt. Nghiên cứu này không sử dụng thuật toán tìm ngưỡng để phát hiện giá trị ngưỡng cực đại của tín hiệu. Do đó, giá trị cho ngưỡng cực đại này có thể chỉ dựa trên giá trị quan sát, trong khi giá trị ngưỡng trong phương pháp ngưỡng biên độ mà luận án đề xuất sẽ thay đổi tùy thuộc vào tín hiệu ngõ vào. M. Benda và I. Volosyak đã áp dụng bộ lọc thông dải trong khoảng từ 2 Hz đến 60 Hz cho tín hiệu EEG sau khi đã loại bỏ nhiễu nguồn điện 50 Hz [127]. Bộ lọc Savitzky–Golay có độ rộng cửa sổ 15 và bậc 9 cũng được áp dụng để làm trơn tín hiệu. Kết quả phát hiện nháy mắt dựa vào ngưỡng cực đại của các nghiên cứu này cho thấy độ chính xác là 89,69%. Trong nghiên cứu [18], mô hình mạng nơ-ron ANN dựa trên hệ số AR đã được áp dụng để nhận dạng

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

các hoạt động mắt. Đầu tiên, tín hiệu EEG được lọc bằng bộ lọc Hamming và sau đó các đặc trưng được trích xuất dựa vào hệ số AR và đưa vào mô hình ANN để huấn luyện. Độ chính xác trung bình khi sử dụng mô hình ANN để phân loại 4 hoạt động mắt như mở mắt, nháy hai mắt, liếc trái, liếc phải là 94%.

Bảng 3.6. Các nghiên cứu phân loại hoạt động mắt.

Nghiên cứu	Loại hoạt động mắt	Kỹ thuật phân loại	Độ chính xác
Dang-Khoa Tran [57]	Nháy mắt trái	Ngưỡng đỉnh	95,1%
	Nháy mắt phải		96,1%
	Nháy hai mắt		97,2%
Kleifges K [126]	Nháy mắt	Ngưỡng đỉnh	93,46%
M. Benda [127]	Nháy mắt	Phát hiện đỉnh alpha	89.69%
Thanh-Hai Nguyen [18]	Mở mắt	Mạng NN	90%
	Nháy hai mắt		97%
	Liếc mắt trái		92%
	Liếc mắt phải		95%
Phương pháp đề xuất 1	Nháy mắt trái	Ngưỡng biên độ	97%
	Nháy mắt phải		99%
	Không nháy mắt		82%
Phương pháp đề xuất 2	Nháy mắt trái	Mạng CNN-1D	98,1%
	Nháy mắt phải		100%
	Nháy hai mắt một lần		95,9%
	Nháy hai mắt hai lần		100%
	Không nháy mắt		98,1%

Bảng 3.7 mô tả thời gian phân loại tín hiệu EEG khi sử dụng 2 phương pháp phân loại hoạt động mắt khác nhau. Từ bảng 3.7 có thể thấy phương pháp ngưỡng biên độ có thời gian xử lý nhanh hơn khoảng 4 lần so với phương pháp phân loại dùng mô hình mạng CNN-1D, lần lượt là 1,21 giây và 5,11 giây tương ứng. Trong thực tế, quá trình thu thập dữ liệu chiếm nhiều thời gian nhất do phải xử lý một khung dữ liệu. Cụ thể là, với phương pháp ngưỡng biên độ, thời gian thu dữ liệu mất 1 giây cho một khung dữ liệu, trong khi đó mô hình CNN-1D phải cần 5 giây cho một khung. Ngoài ra, bảng 3.7 cho thấy thời gian tiền xử lý khi dùng phương pháp ngưỡng biên

Chương 3: Phân loại tín hiệu EEG của hoạt động mắt cho ứng dụng giao tiếp giữa não người và máy tính

độ là 200 ms, gấp 2 lần khi dùng mô hình CNN-1D. Nguyên nhân là do phương pháp ngưỡng biên độ dùng cả 2 bộ lọc Hamming và Savitky-Golay để tiền xử lý tín hiệu EEG, còn mô hình CNN-1D chỉ sử dụng một bộ lọc làm trơn tín hiệu Savitky-Golay. Thời gian phân loại hoạt động của mắt là 10 ms cho cả 2 phương pháp.

Bảng 3.7. Thời gian xử lý phân loại tín hiệu EEG của hoạt động mắt. – Đơn vị (ms)

Phương pháp	Thu thập dữ liệu	Tiền xử lý	Phân loại	Tổng thời gian
Phương pháp ngưỡng biên độ	1000	200	10	1210
CNN-1D	5000	100	10	5110

❖ Kết luận

Phương pháp phân loại hoạt động nháy mắt dựa trên tín hiệu EEG đã được đề xuất với hiệu suất cao bao gồm phương pháp phân loại theo ngưỡng biên độ tín hiệu [123] và phương pháp phân loại dùng mô hình mạng học sâu CNN-1D [125]. Tín hiệu EEG sau khi được thu nhận ở các kênh AF3, F7, AF4 và F8 sẽ được tiền xử lý dùng bộ lọc Hamming và bộ lọc Savitzky-Golay nhằm nâng cao hiệu suất phân loại. Với phương pháp phân loại hoạt động nháy mắt trái và nháy mắt phải dùng thuật toán ngưỡng biên độ, kết quả phân loại tốt với độ chính xác lần lượt là 97% và 99% cho mỗi loại và thời gian xử lý nhanh, khoảng 1,2 ms. Với mô hình CNN-1D gồm 13 lớp, tín hiệu thu được từ 4 kênh AF3, F7, AF4 và F8 được ghép lại thành một tín hiệu duy nhất giúp mô hình trích xuất được nhiều đặc trưng hơn từ tín hiệu EEG. Ngoài ra, mô hình CNN-1D có thể phân loại được nhiều hoạt động nháy mắt hơn so với phương pháp phân loại dựa vào ngưỡng biên độ. Mô hình phân loại 5 tín hiệu nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt hai lần liên tiếp và không nháy mắt có độ chính xác lần lượt là 98,1%, 100%, 95,9%, 100% và 98,1%. Mặc dù, hiệu suất phân loại hoạt động nháy mắt dùng mô hình CNN-1D là cao nhưng thời gian xử lý vẫn còn chậm. Với các kết quả đạt được, các phương pháp đề xuất có thể được sử dụng để phân loại các hoạt động nháy mắt, từ đó gán các lệnh điều khiển cho các ứng dụng BCI trên các nền tảng robot di động như xe lăn điện.

CHƯƠNG 4

NHẬN DẠNG VÀ XÁC ĐỊNH VỊ TRÍ VẬT MỐC TỰ NHIÊN TRONG MÔI TRƯỜNG TRONG NHÀ

Việc nhận dạng và xác định vị trí các vật mốc tự nhiên để bản địa hóa các nền tảng di động đã thu hút các nhà nghiên cứu trong nhiều năm gần đây. Trích xuất đặc trưng đóng một vai trò quan trọng trong việc nhận dạng các vật mốc tự nhiên này dựa trên mật độ đặc trưng. Chương này trình bày việc nhận dạng và xác định vị trí các vật mốc tự nhiên trong môi trường trong nhà với các đối tượng khác nhau bằng cách sử dụng hệ thống camera RGB-D được cài đặt trên xe lăn điện. Bộ dò tìm ORB được áp dụng để trích xuất đặc trưng của tất cả các đối tượng trong ảnh RGB, từ đó tạo ra nhiều khung đặc trưng nhị phân tương ứng với các đối tượng trong ảnh. Một trong các khung đặc trưng chứa mật độ đặc trưng lớn nhất được xác định là một vật mốc tự nhiên. Để đánh giá độ chính xác của nhận dạng vật mốc, hệ số IOU được sử dụng để tính toán tỷ lệ trùng khớp của giữa vật mốc nhận dạng được và vật mốc tham chiếu. Ngoài ra, vị trí của vật mốc cũng được xác định dựa vào thông tin 3D từ camera RGB-D và từ encoder được gắn ở các bánh xe lăn. Các thí nghiệm được trình bày ở cuối chương để kiểm chứng độ chính xác của phương pháp đề xuất.

4.1. VAI TRÒ CỦA VẬT MỐC ĐỐI VỚI QUÁ TRÌNH ĐỊNH VỊ VÀ ĐIỀU KHIỂN XE LĂN ĐIỆN

Trong quá trình điều khiển xe lăn điện tự động đi đến đích, việc định vị chính xác vị trí của xe lăn trong môi trường là rất quan trọng. Định vị có thể phân thành hai loại dựa theo môi trường hoạt động là định vị ngoài trời và định vị trong nhà. Việc định vị ngoài trời thường được thực hiện bằng cách sử dụng GPS. Tuy nhiên, việc định vị trong nhà sử dụng GPS sẽ cho thông tin vị trí thường không chính xác, cần

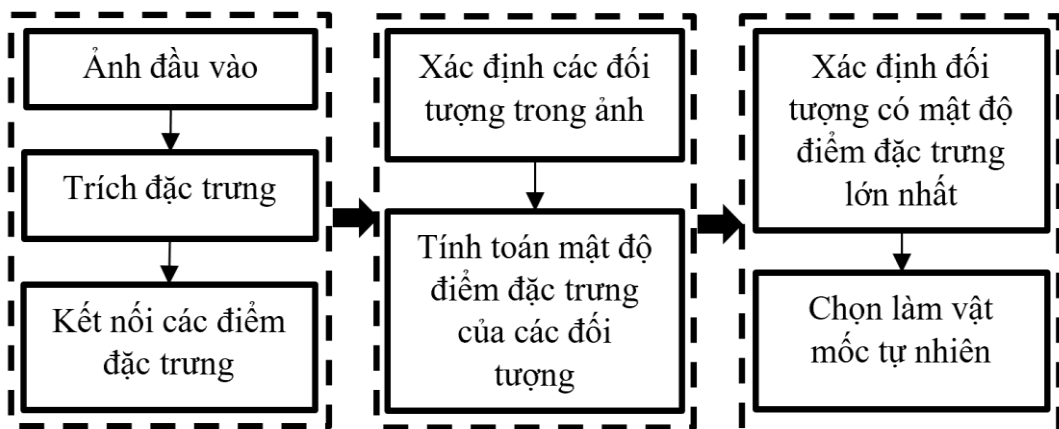
phải kết hợp bộ lọc Kalman để tăng độ chính xác [23]. Một phương pháp định vị khác là định vị bằng vô tuyến, sử dụng Bluetooth, sóng radio, NFC, RFID, hay Wifi [39, 40, 94]. Trong trường hợp này, robot tự hành được trang bị một thiết bị thu sóng tín hiệu từ các điểm phát sóng tín hiệu có vị trí biết trước được bố trí xung quanh môi trường làm việc của nó. Thiết bị thu sóng đặt trên robot tự hành tính toán vị trí của nó nhờ vào cường độ tín hiệu thu được từ ít nhất ba điểm phát sóng và vị trí của ba điểm phát sóng đó. Tuy nhiên, nhược điểm của các phương pháp này là giá thành cao, ít linh động vì cần trang bị các thiết bị thu phát sóng, và đa số các phương pháp này chỉ dùng để định vị hoặc định hướng, không thể kết hợp với việc tìm đường đi cho xe lăn điện hay robot. Phương pháp Odometry thường được sử dụng cho việc định vị xe lăn điện hoặc robot trong môi trường trong nhà vì dễ dàng thực hiện [5]. Phương pháp này xác định vị trí của xe lăn hoặc robot tự hành dựa vào thông tin của encoder gắn ở bánh xe. Các thông tin từ encoder được đưa vào phương trình động lực học của xe lăn hoặc robot tự hành để tính toán vị trí của nó. Tuy nhiên, phương pháp này cần phải thực hiện hiệu chỉnh các thông số trong mô hình động lực học để giảm sai số tích lũy trong quá trình tính toán vị trí.

Xe lăn điện hoặc robot được thiết kế để tự động di chuyển dựa trên thông tin môi trường được thu thập từ các cảm biến được nghiên cứu nhiều trong những năm gần đây. Với các nền tảng di động dựa trên thông tin mang tính bước ngoặt, việc chọn các vật mốc như cửa ra vào, cầu thang, tường, trần nhà, sàn nhà và trích xuất các đặc trưng của chúng để nhận dạng đóng vai trò quan trọng [47, 48]. Bằng cách sử dụng các vật mốc tự nhiên hoặc nhân tạo, xe lăn hoặc robot có thể ước tính vị trí chính xác hơn trong môi trường di chuyển nhờ vào sự kết hợp của tọa độ của từng vật mốc được phát hiện và các thông tin từ cảm biến. Để phát hiện các vật mốc dựa trên các đặc trưng, các nghiên cứu có thể dựa trên màu sắc, kết cấu, độ sáng, kích thước đối tượng. Trong thực tế, các vật mốc nhân tạo được tạo ra trong môi trường nơi robot di chuyển để giúp robot đi đến mục tiêu mong muốn [128, 129]. Trong các nghiên cứu khác, các nền tảng di động được thiết kế để di chuyển trong môi trường tự nhiên, trong đó các vật mốc là đèn trên trần nhà [130, 131]. Trong các môi trường không có cấu trúc động như khu công nghiệp, văn phòng, trường đại học, nơi có nhiều vật mốc tự nhiên,

các nền tảng di động được thiết kế để tự động định vị để di chuyển dựa trên dữ liệu 3D [132]. X. Chai và các cộng sự đã giới thiệu một cách tiếp cận địa phương hóa tự nhiên cho robot di động bằng cách xác định các đường thẳng và màu sắc đồng nhất trong môi trường tự nhiên [133]. Trong môi trường này với các vùng đặc trưng, phương pháp kết hợp các đường thẳng để tạo thành các đa giác lỗi đã được áp dụng để xác định các đối tượng là các vật mốc tự nhiên cho quỹ đạo của nó.

4.2. PHƯƠNG PHÁP MẬT ĐỘ ĐIỂM ĐẶC TRƯNG LỚN NHẤT CHO NHẬN DẠNG CÁC VẬT MỐC TRONG TỰ NHIÊN

Để nhận biết các vật mốc tự nhiên trong một ảnh, phương pháp mật độ điểm đặc trưng lớn nhất được áp dụng, có thể được chia thành ba giai đoạn sau: chuyển đổi ảnh RGB sang ảnh xám; phát hiện các đặc trưng; và tìm kiếm vùng có mật độ điểm đặc trưng lớn nhất như hình 4.1 [134]. Đầu tiên, các đặc trưng của các đối tượng được trích xuất bằng cách sử dụng phương pháp phát hiện các điểm góc FAST trong bộ tìm đặc trưng ORB [135]. Tiếp theo, các đặc trưng được giữ lại để tạo ra một đám mây các điểm đặc trưng chính cho phát hiện đối tượng. Do đó, kỹ thuật giãn nở được áp dụng cho đám mây điểm đặc trưng chính để tạo liên kết giữa các điểm đặc trưng chính của đối tượng và sau đó vẽ các hộp giới hạn. Cuối cùng, các hộp giới hạn chứa các điểm quan trọng nhất để nhận biết vật mốc được dán nhãn tự động.



Hình 4.1. Các bước xác định vật mốc tự nhiên.

4.2.1. Phát hiện các điểm đặc trưng

Phát hiện các điểm đặc trưng để nhận dạng vật mốc trong môi trường đóng một vai trò quan trọng. Để phát hiện điểm đặc trưng, bộ dò tìm ORB được áp dụng để tăng tốc độ trích xuất đặc trưng do quá trình nhận dạng và xác định vị trí vật mốc cần thực hiện trong thời gian thực [136]. Bộ dò tìm ORB dựa vào công cụ dò điểm góc FAST [137] để tìm các điểm đặc trưng với một số cải tiến để có được các điểm đặc trưng tốt nhất. Cụ thể, FAST không tạo ra thước đo độ góc và các điểm góc thường tập trung dọc theo các cạnh. Do đó, thước đo góc Harris [99] được sử dụng để sắp xếp các điểm đặc trưng có được sau khi áp dụng FAST.

Đối với số lượng điểm đặc trưng N được yêu cầu, trước tiên công cụ dò điểm góc FAST được sử dụng để lấy hơn N điểm bằng cách đặt ngưỡng đủ thấp. Cụ thể, một vòng tròn pixel sẽ được kiểm tra xung quanh mỗi ứng viên điểm góc. Ứng cử viên này được coi là một điểm góc nếu cường độ của các pixel liền kề xung quanh pixel trung tâm ở trên hoặc dưới ngưỡng cường độ của pixel trung tâm. Do đó, các pixel được trích xuất thường nằm trong các vùng đặc biệt và có độ tương phản cao của ảnh. Bộ phát hiện điểm góc FAST được mô tả theo các bước sau:

Bước 1: Trong tiêu chí kiểm tra, một vòng tròn m pixel xung quanh một ứng viên điểm góc p . Mỗi pixel trong vòng tròn được dán nhãn từ 1 đến m pixel theo chiều kim đồng hồ. Cường độ pixel trong vòng tròn này liên quan đến ứng viên p được ký hiệu là I_x , có thể có một trong ba trạng thái: sáng hơn, tối hơn hoặc tương tự. Phương pháp kiểm tra tốc độ cao được áp dụng để loại bỏ các điểm không phải là điểm góc, trong đó một số pixel được sử dụng để xác định cường độ của các pixel liền kề so với pixel trung tâm để phát hiện ứng viên điểm góc. Thuật toán 4.1 mô tả cách thức tìm kiếm điểm góc của bộ phát hiện điểm góc FAST.

Bước 2: Hàm F_{score} được tính toán để giảm bớt các điểm góc:

$$F_{score} = \max\left(\sum_{x \in S_{brighter}} |I_x - I_p|, \sum_{x \in S_{darker}} |I_p - I_x|\right) \quad (4.1)$$

trong đó I_x là cường độ pixel trong vòng tròn và I_p là cường độ của pixel.

Cuối cùng, khi đã có được các điểm đặc trưng, thước đo góc Harris được dùng để đo độ góc của các điểm đặc trưng. Độ góc Harris được tính bằng cách chọn một cửa sổ 7×7 xung quanh mỗi điểm đặc trưng và tính toán độ dốc trong vùng đó. Khi độ góc của các điểm đặc trưng được tính toán và sắp xếp theo thứ tự, N điểm đặc trưng tốt nhất được lựa chọn.

Thuật toán 4.1: Tìm kiếm điểm góc

- 1: **Input:**
 - I_p : Cường độ điểm p .
 - I_x : Cường độ điểm ảnh trên vòng tròn xung quanh điểm p .
 - t : Giá trị ngưỡng (20% giá trị cường độ điểm p).
 - m : Số điểm trên vòng tròn xung quanh điểm p .
 - 2: **for** $x = 1; x \leq m$ **do**
 - 3: $\Delta I_x = I_x - I_p$
 - 4: **if** $\Delta I_x \geq t$ **then**
 - 5: $x \in S_{brighter}$
 - 6: **else if** $\Delta I_x \leq -t$ **then**
 - 7: $x \in S_{darker}$
 - 8: **else if** $-t < \Delta I_x < t$ **then**
 - 9: $x \in S_{similar}$
 - 10: **end if**
 - 11: **end for**
 - 12: **if** (Ít nhất ba trong số bốn pixel ở phía dưới, ở hai bên và ở trên cùng của vòng tròn là thành viên của tập $S_{similar}$) **then**
 - 13: p không phải là điểm góc.
 - 14: **else if** (Ít nhất ba trong số bốn pixel ở phía dưới, ở hai bên và ở đầu vòng tròn là thành viên của tập $S_{brighter}$ hoặc S_{darker}) **then**
 - 15: Kiểm tra tất cả m pixels.
 - 16: **end if**
 - 17: **if** (m pixels trên vòng tròn thuộc tập $S_{brighter}$ or S_{darker}) **then**
 - 18: p được xem như là một điểm góc.
 - 19: **end if**
 - 20: **Output:** trạng thái của điểm p : điểm góc hoặc không phải điểm góc.
-

4.2.2. Kết nối các điểm đặc trưng trong các đối tượng

Sau quá trình trích xuất đặc trưng, việc lưu giữ các điểm đặc trưng được thực hiện và tất cả các điểm ảnh khác không phải là điểm đặc trưng sẽ bị xóa. Do đó, hình ảnh nhị phân \mathbf{A} chỉ chứa các điểm đặc trưng (tiền cảnh) hoặc điểm không phải là đặc trưng (nền). Các điểm đặc trưng của đối tượng trong ảnh được kết nối lại bằng cách làm giãn nở chúng ra. Cụ thể, phép giãn nở ảnh nhị phân \mathbf{A} với một phần tử cấu trúc \mathbf{K} được thực hiện. Hình ảnh sau khi giãn nở \mathbf{D} được tính như sau:

$$\mathbf{D} = \mathbf{A} \oplus \mathbf{K} = \left\{ \mathbf{z} \mid \left(\hat{\mathbf{K}} \right)_{\mathbf{z}} \cap \mathbf{A} \neq \Phi \right\} \quad (4.2)$$

trong đó $\left(\hat{\mathbf{K}} \right)_{\mathbf{z}}$ là phép chiếu của \mathbf{K} từ góc tọa độ và dịch chuyển theo \mathbf{Z} . Do đó, phép giãn của \mathbf{A} với \mathbf{K} là tập hợp của tất cả các phép chiếu \mathbf{Z} , sao cho $\left(\hat{\mathbf{K}} \right)_{\mathbf{z}}$ và \mathbf{A} trùng nhau tại ít nhất một phần tử [138].

Để kết nối đặc trưng của một đối tượng trong ảnh, sự giãn nở làm cho vật thể có màu trắng lớn hơn. Do đó, để thực hiện sự giãn nở, một phần tử cấu trúc có kích thước 3×3 được sử dụng như sau:

$$\mathbf{K} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.3)$$

4.2.3. Nhận biết vật mốc tự nhiên

Sau khi thực hiện kết nối các điểm đặc trưng của đối tượng trong ảnh, đối tượng được tách ra bằng kỹ thuật tìm đường bao của đối tượng. Đường bao là đường cong kết nối tất cả các điểm liên tục dọc theo đường biên dựa trên cùng một màu hoặc cường độ. Do đó, nếu tìm thấy một đường bao trong ảnh nhị phân, thì đường bao của một đối tượng trong ảnh sẽ được tìm thấy. Rõ ràng rằng đường bao là một công cụ hữu ích để phân tích hình dạng và phát hiện đối tượng.

Vật mốc tự nhiên được xác định là đối tượng chứa nhiều điểm đặc trưng nhất trong một ảnh. Do đó, các đối tượng được phát hiện với các đặc trưng này sẽ được

tính toán để chọn ra đối tượng có nhiều đặc trưng nhất. Cụ thể, hình ảnh giãn nở D được xử lý để chỉ chứa 2 giá trị 0 và 1, và tổng các điểm ảnh màu trắng theo r hàng và c cột được xác định như sau:

$$\chi = \sum_{x=0}^r \sum_{y=0}^c D(x, y) \quad (4.4)$$

Tương tự, tổng các điểm ảnh màu trắng trong các đối tượng của ảnh O_i sau khi vẽ các đường bao được tính như sau:

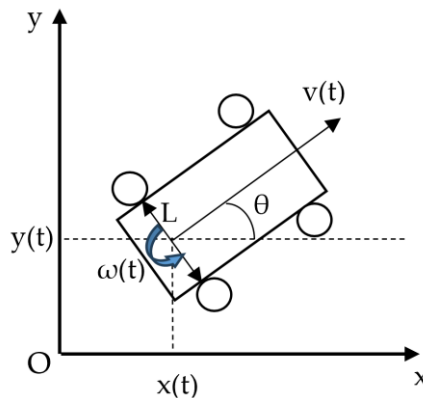
$$\chi_i = \sum_{x_i=0}^h \sum_{y_i=0}^w O_i(x_i, y_i) \quad (4.5)$$

Từ phương trình (4.4) và (4.5), hệ số mật độ δ_i của các điểm đặc trưng trên một đối tượng so với tổng số điểm đặc trưng của ảnh gốc được xác định theo phương trình (4.6). Sau đó, đối tượng nào có hệ số δ_i lớn nhất được chọn là vật mốc tự nhiên trong ảnh gốc.

$$\delta_i = \frac{\chi_i}{\chi} \quad (4.6)$$

4.3. XÁC ĐỊNH VỊ TRÍ VẬT MỐC TRONG MÔI TRƯỜNG

4.3.1. Vị trí của xe lăn trong môi trường



Hình 4.2. Mô hình xe lăn di động hai bánh vi sai và hai bánh tự do.

Xe lăn điện là một cấu trúc cơ học trong đó các thành phần được liên kết với nhau và di chuyển được bằng động cơ. Các bộ mã hóa (Encoder) được gắn vào trục động cơ để xác định vị trí bánh xe. Để điều khiển xe lăn, cần biểu diễn trạng thái của xe lăn bằng các thông số định lượng và từ đó có thể lập mô hình chuyển động của xe

lăn bằng phương trình động lực học. Mô hình xe lăn di động trong luận án này là loại có hai bánh xe sau truyền động vi sai và hai bánh xe trước tự do, trong đó vị trí của xe lăn được tính toán dựa vào chuyển động giữa bánh xe sau và mặt đất với điều kiện các bánh xe chỉ có chuyển động lăn mà không trượt, được mô tả trong hình 4.2.

Mô hình xe lăn có thể chuyển động và định hướng bằng hai bánh và đổi hướng dựa trên sự thay đổi vận tốc của bánh bên trái $v_l(t)$ và bánh bên phải $v_r(t)$. Với L là khoảng cách giữa hai bánh xe, θ là góc tạo bởi trục của khung xe và trục hoành. Vận tốc góc $\omega(t)$ và vận tốc di chuyển $v(t)$ của xe lăn được tính như sau:

$$\omega(t) = \frac{v_r(t) - v_l(t)}{L} \quad (4.7)$$

$$v(t) = \frac{v_r(t) + v_l(t)}{2} \quad (4.8)$$

Tất cả các biến trong phương trình động lực học của xe lăn được xét tại thời điểm t . Cụ thể, tọa độ $x(t)$, $y(t)$ và góc định hướng $\theta(t)$ tại thời điểm t liên quan đến vận tốc góc $\omega(t)$ và vận tốc $v(t)$ của xe lăn như trong công thức sau:

$$\begin{bmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{\theta}(t) \end{bmatrix} = \begin{bmatrix} \cos \theta(t) & 0 \\ \sin \theta(t) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v(t) \\ \omega(t) \end{bmatrix} \quad (4.9)$$

Tọa độ của xe lăn tại thời điểm $t = k + 1$ được mô tả như sau:

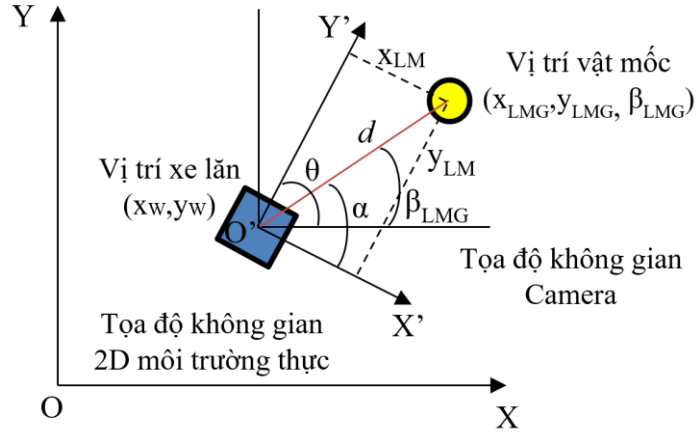
$$\begin{bmatrix} x(k+1) \\ y(k+1) \\ \theta(k+1) \end{bmatrix} = \begin{bmatrix} \cos \theta(k+1) & 0 \\ \sin \theta(k+1) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{d_r(k+1) + d_l(k+1)}{2} \\ \frac{d_r(k+1) - d_l(k+1)}{L} \end{bmatrix} + \begin{bmatrix} x(k) \\ y(k) \\ \theta(k) \end{bmatrix} \quad (4.10)$$

trong đó $d_r(k+1)$ và $d_l(k+1)$ lần lượt là khoảng cách của bánh xe bên phải và bên trái từ thời điểm thứ k đến $(k+1)$. Phương trình (4.10) cho thấy tọa độ của xe lăn tại thời điểm $(k+1)$ được xác định dựa trên quãng đường di chuyển của bánh trái và phải từ thời điểm k đến $(k+1)$ và tọa độ của xe lăn tại thời điểm k .

4.3.2. Xác định vị trí vật mốc

Trong quá trình di chuyển, vị trí của các vật mốc sẽ được phát hiện và gán nhãn bất cứ khi nào xe lăn đến gần các vật mốc. Sau khi phát hiện một vật mốc tự

nhiên, tâm của vật mốc sẽ được xác định bằng cách ánh xạ lên không gian 3D của camera RGB-D và xác định tọa độ của tâm vật mốc (x_{LM}, y_{LM}) , trong đó độ lệch x_{LM} của tâm vật mốc được tính từ tâm camera sẽ nhỏ hơn 0 nếu nó nằm bên phải camera và lớn hơn 0 nếu vật mốc nằm bên trái camera. Hình 4.3 mô tả vị trí của xe lăn và vật mốc trong không gian, trong đó tọa độ của xe lăn trong môi trường 2D là (x_W, y_W) và tọa độ của vật mốc trong không gian camera là (x_{LM}, y_{LM}) .



Hình 4.3. Ước tính vị trí mốc trong không gian 2D.

Từ hình 4.3, tọa độ vật mốc trong môi trường 2D được tính toán dựa vào tọa độ xe lăn như sau:

$$x_{LMG} = x_W + d \cdot \cos \beta_{LMG} \quad (4.11)$$

$$y_{LMG} = y_W + d \cdot \sin \beta_{LMG} \quad (4.12)$$

$$\beta_{LMG} = \alpha + \theta - 90 \quad (4.13)$$

trong đó θ là góc của xe lăn chuyển động sử dụng phương trình (4.10). Góc của vật mốc α được tính từ khoảng cách d từ vật mốc đến vị trí camera được mô tả như sau:

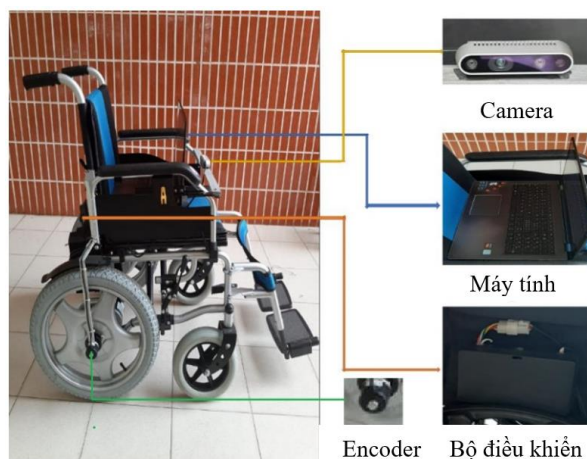
$$\alpha = \begin{cases} \arctan\left(\frac{y_{LM}}{x_{LM}}\right), & x_{LM} < 0 \\ 180 - \arctan\left(\frac{y_{LM}}{x_{LM}}\right), & x_{LM} \geq 0 \end{cases} \quad (4.14)$$

$$d = \sqrt{x_{LM}^2 + y_{LM}^2} \quad (4.15)$$

Sau khi xác định vị trí vật mốc, số thứ tự và tọa độ của vật mốc trong hình ảnh được đánh dấu cho hình ảnh vật mốc đó và được lưu lại vào cơ sở dữ liệu phục vụ cho việc định vị vị trí của xe lăn.

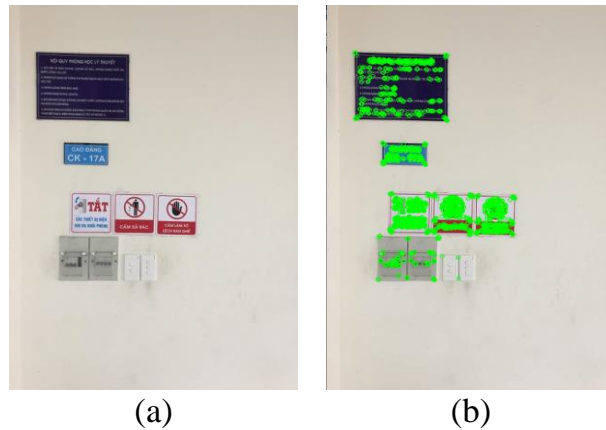
4.4. KẾT QUẢ NHẬN BIẾT VÀ THU THẬP VẬT MỐC TỰ NHIÊN

Để đánh giá phương pháp đề xuất, các thử nghiệm đã được thực hiện. Trong các thí nghiệm, xe lăn được trang bị các thiết bị như mô tả trong hình 4.4. Camera RGB-D được sử dụng trong thí nghiệm này là Intel RealSense D435, có kích thước là $90\text{ mm} \times 25\text{ mm} \times 25\text{ mm}$; khoảng cách độ sâu tối thiểu là 0,105 m và phạm vi tối đa 10 m, thay đổi tùy theo hiệu suất, ánh sáng, với tầm hoạt động tốt nhất là từ 0,3 m đến 3,0 m. Camera RGB có độ phân giải 1280x960 và tốc độ khung hình lên đến 90 khung hình/giây. Ảnh độ sâu thu được từ cảm biến IR, sau đó tính toán bản đồ độ sâu. Bánh xe có đường kính 0,47 m và khoảng cách giữa hai bánh là $L = 0,55\text{ m}$, 2 Encoder được gắn đồng trục với hai bánh xe, với thông số 2048 xung/vòng. Phòng học và hành lang được sử dụng làm thí nghiệm với các vật mốc tự nhiên. Thí nghiệm được thực hiện với các trường hợp khác nhau của các vật mốc, bao gồm ít vật mốc, nhiều vật mốc, góc di chuyển, khoảng cách đến vật mốc.



Hình 4.4. Xe lăn với camera RGB-D, Encoder và máy tính.

Một hình ảnh đầu vào với bức tường trắng và bốn đối tượng bao gồm các bảng thông tin và chỉ dẫn, các bảng công tắc điện như trong hình 4.5a. Trong hình 4.5b, các vòng tròn nhỏ màu xanh lục trên các đối tượng là các đặc trưng được trích xuất từ hình ảnh đầu vào bằng bộ dò tìm ORB. Từ hình 4.5b, có thể thấy rằng các điểm đặc trưng là những điểm có màu sắc khác biệt nhất hoặc nổi bật nhất so với những điểm còn lại. Các điểm đặc trưng đại diện cho các vật mốc và được sử dụng để nhận dạng các vật mốc tương ứng để định vị và lập kế hoạch di chuyển cho xe lăn điện.



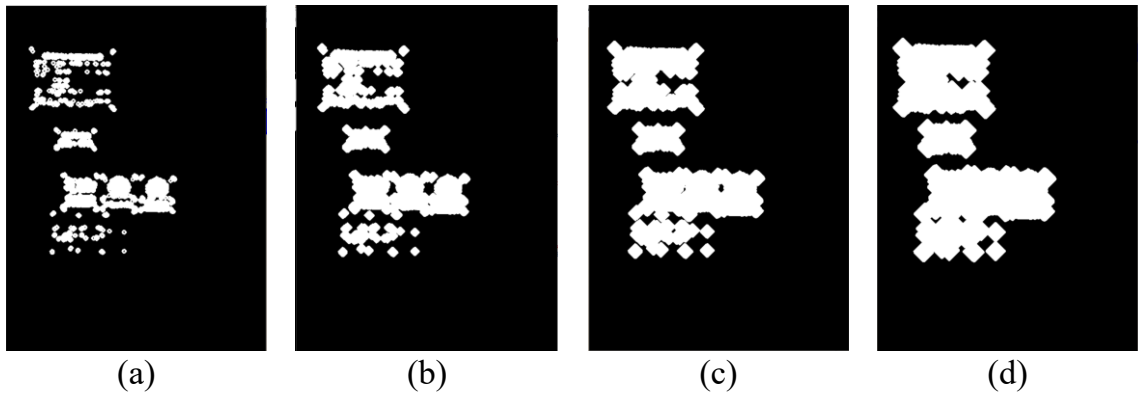
Hình 4.5. Các đối tượng chứa các điểm đặc trưng: (a) Ảnh RGB, (b) Ảnh chứa các điểm đặc trưng.

Đối với ảnh đầu vào của hệ thống nhận dạng đối tượng theo thời gian thực, các đặc trưng được trích xuất bằng các phương pháp khác nhau để so sánh và đánh giá. Bảng 4.1 cho thấy hiệu suất của các bộ dò tìm đặc trưng SIFT, SURF và ORB khi thực hiện trích xuất đặc trưng cho hình 4.5a. Các bộ dò tìm SIFT, SURF và ORB có thời gian khác nhau để xác định các điểm đặc trưng ban đầu, trong đó tốc độ tính toán cho một điểm đặc trưng bằng thuật toán SIFT và SURF gần như tương tự nhau và chậm hơn so với ORB.

Bảng 4.1. Kết quả trích đặc trưng dùng các phương pháp khác nhau.

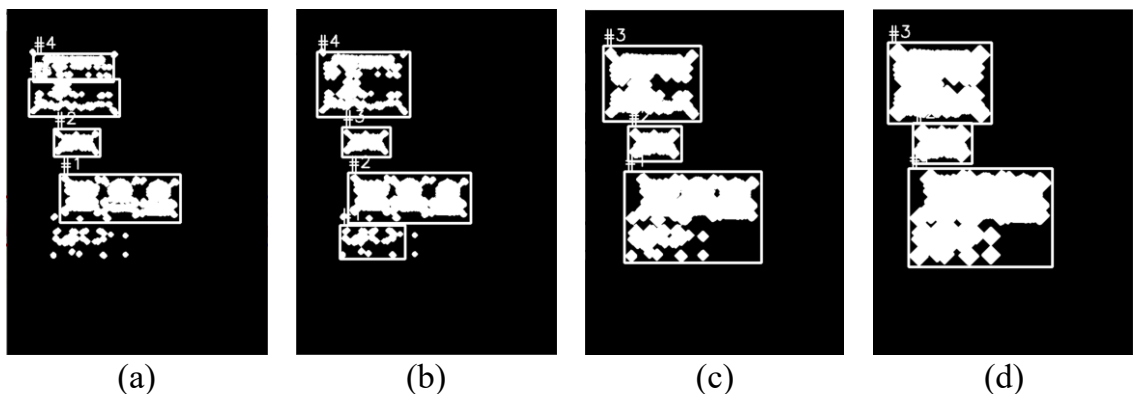
Bộ dò tìm đặc trưng	Thời gian xử lý cho một ảnh (ms)	Thời gian tính toán cho một điểm đặc trưng (ms)	Số lượng điểm đặc trưng trong một ảnh
SIFT [101]	31,08	0,07	426
SURF [102]	17,55	0,08	230
ORB [134]	3,74	0,002	1850

Hình 4.6 thể hiện sự kết nối của các điểm đặc trưng của các đối tượng trong hình 4.5, trong đó sự giãn nở của các điểm đặc trưng được thực hiện theo các lần lặp khác nhau. Với sự lặp lại khác nhau, các điểm đặc trưng của đối tượng sẽ bị giãn nở ra và cho thấy các vấn đề về như sự rời rạc hoặc chồng chéo lên nhau của các điểm đặc trưng. Cụ thể, với số lần lặp lại là 1 cho thấy sự rời rạc của các điểm đặc trưng của đối tượng như hình 4.6a, trong khi với lần lặp lại tăng lên sẽ cho thấy sự chồng lấp của các điểm đặc trưng, được thể hiện trong hình 4.6b, hình 4.6c và hình 4.6d.



Hình 4.6. Biểu diễn sự giãn nở với các lần lặp khác nhau bằng cách sử dụng mặt nạ 3×3 : (a) Một lần lặp, (b) Năm lần lặp, (c) Mười lần lặp, (d) Mười lăm lần lặp.

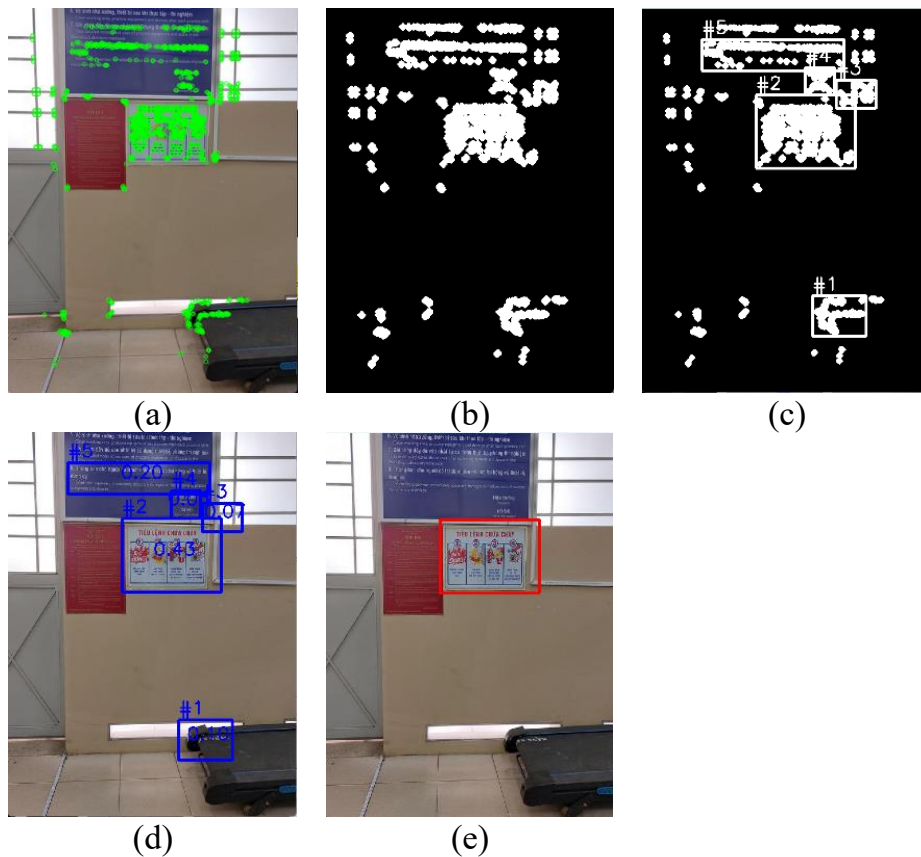
Từ kết quả của sự lặp lại các lần giãn nở các điểm đặc trưng, số lượng các đối tượng được tìm thấy dựa trên độ dài của các đường bao cũng thay đổi theo như mô tả ở hình 4.7. Cụ thể là, hình 4.7a và hình 4.7b cho thấy 4 hộp giới hạn được xác định và được gán các ký hiệu (từ # 1 đến # 4), trong khi đó chỉ có 3 hộp giới hạn được phát hiện cho một khung đặc trưng như trong hình 4.7c và hình 4.7d (từ # 1 đến # 3). Có thể thấy được khi số lần lặp là 1 thì các đối tượng được xác định sai và không đầy đủ. Cụ thể là thuật toán không thể xác định được các bảng công tắc điện, trong khi bảng thông báo màu xanh da trời đậm thì lại được nhận biết thành 2 đối tượng. Trong trường hợp số lần lặp lớn, cụ thể là 10 và 15, thì số đối tượng được phát hiện cũng sai, đó là các bảng công tắc điện và bảng chỉ dẫn bị gộp chung thành một đối tượng. Có thể thấy rằng số lần lặp là 5 phù hợp nhất cho việc xác định các đối tượng trong ảnh, làm cơ sở để chọn ra vật mốc tiêu biểu dựa mật độ điểm đặc trưng.



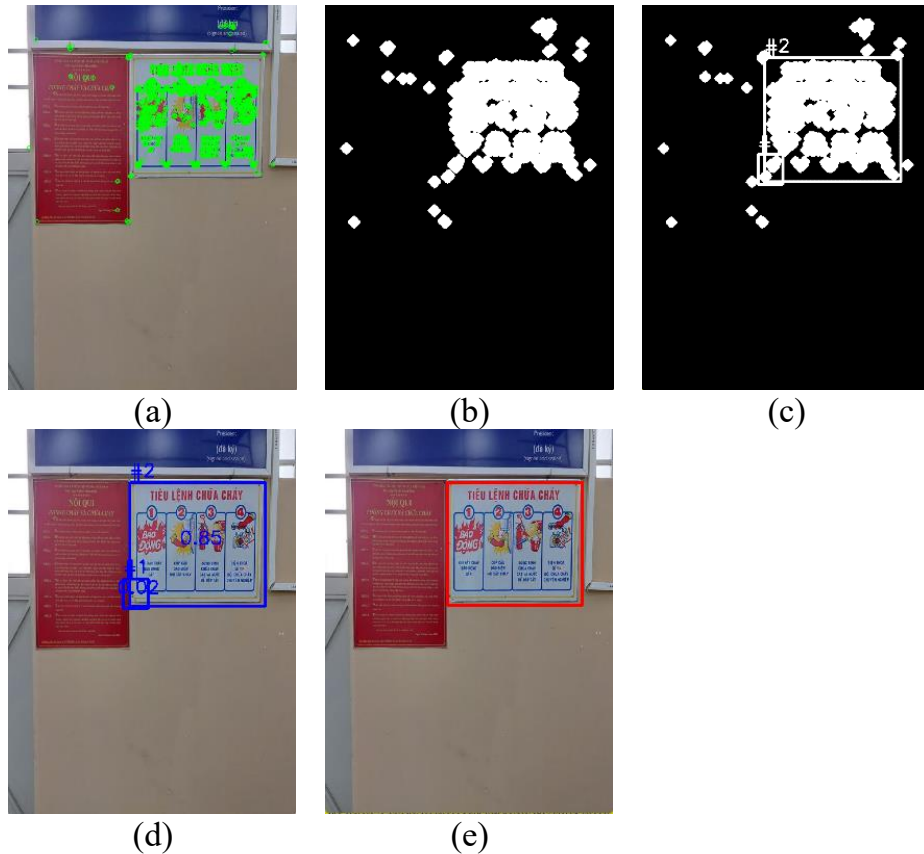
Hình 4.7. Ảnh chứa đường bao đối tượng và khung đặc trưng: (a) Một lần lặp, (b) Năm lần lặp, (c) Mười lần lặp, (d) Mười lăm lần lặp.

Chương 4: Nhận dạng và xác định vị trí vật mốc tự nhiên trong môi trường trong nhà

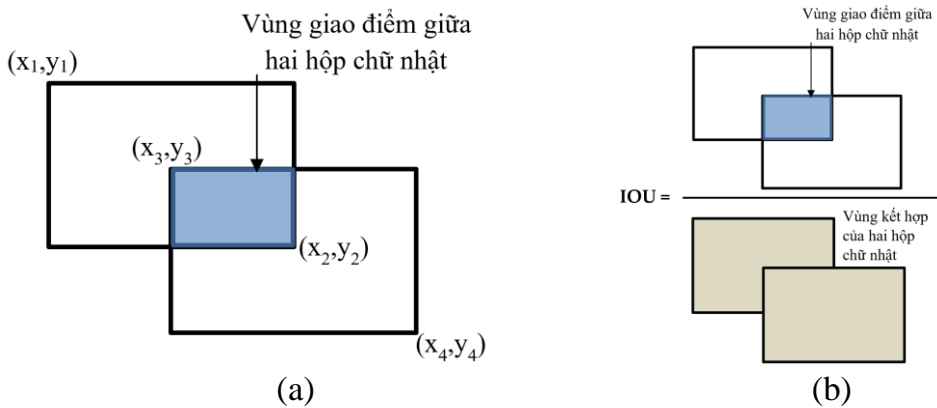
Trong quá trình di chuyển bằng xe lăn, hình ảnh với các vật thể tự nhiên được chụp ở các khoảng cách khác nhau từ hệ thống camera RGB-D. Trong hình 4.8, một hình ảnh môi trường với 5 đối tượng và bức tường được chụp ở khoảng cách 2 m từ camera đến bức tường, trong đó 5 đối tượng hiển thị mật độ đặc điểm khác nhau. Trong đó, đối tượng là bảng hướng dẫn màu trắng có mật độ đặc trưng lớn nhất là $\delta_2 = 0,43$ sử dụng công thức (4.10) nên hệ thống ghi nhận sẽ chọn bảng hướng dẫn này làm vật mốc. Hình 4.9 cho thấy hệ thống nhận dạng hình ảnh với 2 đối tượng ở khoảng cách 1m từ camera đến tường. Hình 4.9c cho thấy đối tượng thứ hai được chọn làm vật mốc tương ứng với mật độ $\delta_2 = 0,85$ lớn nhất so với đối tượng còn lại. Có thể thấy trong cả hai thí nghiệm, kết quả nhận dạng đều cho ra cùng một vật mốc giống nhau.



Hình 4.8. Quá trình nhận dạng vật mốc trong phòng thí nghiệm ở khoảng cách 2 m từ camera đến các đối tượng: (a) Ảnh có các đặc trưng; (b) Hình ảnh giãn nở; (c) Khung đặc trưng; (d) Hình ảnh các đối tượng; (e) Hình ảnh vật mốc.



Hình 4.9. Quá trình nhận dạng vật mốc trong phòng thí nghiệm ở khoảng cách 1 m từ camera đến các đối tượng: (a) Ảnh có các đặc trưng; (b) Hình ảnh giãn nở; (c) Khung đặc trưng; (d) Hình ảnh các đối tượng; (e) Hình ảnh vật mốc.



Hình 4.10. IOU cho các hộp giới hạn: (a) Giao điểm của hai hộp; (b) Công thức tính IOU.

Một thí nghiệm khác nhằm kiểm chứng độ chính xác của thuật toán phát hiện các vật mốc tự nhiên là đo tỉ lệ chồng lấp IOU [129] giữa 2 hộp (1 hộp chứa đối tượng thật và 1 hộp chứa đối tượng được chọn là vật mốc từ thuật toán đề xuất). Vùng chồng

lấp càng lớn thì IOU càng lớn, cho thấy khả năng chọn đối tượng làm vật mốc là đúng với thực tế và kết quả dự đoán tốt nếu IOU lớn hơn 0,5. Giả sử rằng hộp 1 được biểu diễn bởi $[x_1, y_1, x_2, y_2]$ và hộp 2 được biểu diễn bởi $[x_3, y_3, x_4, y_4]$ như hình 4.10a, IOU được tính như mô tả ở hình 4.10b.

Bảng 4.2 biểu diễn giá trị IOU khi thực hiện khi nhận biết các vật mốc ở các khoảng cách khác nhau. Từ bảng 4.2, có thể thấy giá trị IOU trung bình của 5 lần thí nghiệm với 5 khoảng cách khác nhau là lớn hơn 0,5, trong đó các thí nghiệm nhận dạng vật mốc ở khoảng cách 1 m và 2 m có giá trị IOU bằng nhau và lớn nhất với hơn 0,86. Cụ thể, trong 10 lần thực hiện thí nghiệm với khoảng cách 1 m và 2 m, giá trị IOU nằm trong khoảng từ 0,8 đến 0,92, trong khi đó thì giá trị IOU nằm trong khoảng từ 0,53 đến 0,72 với khoảng cách 3 m. Đối với khoảng cách 4 m và 5 m, giá trị IOU trung bình tương đối thấp với 0,56 và 0,52 tương ứng. Điều này cho thấy rằng, ở các khoảng cách khác nhau, thuật toán vẫn cho kết quả nhận dạng vật mốc tốt. Tuy nhiên, với các khoảng cách gần dưới 2 m thì kết quả nhận dạng có độ chính xác cao nhất.

Bảng 4.2. IOU với các khoảng cách khác nhau.

TT	1m	2m	3m	4m	5m
1	0,86	0,83	0,72	0,56	0,68
2	0,90	0,92	0,57	0,61	0,50
3	0,91	0,83	0,53	0,55	0,48
4	0,86	0,85	0,56	0,54	0,47
5	0,89	0,91	0,55	0,63	0,60
6	0,80	0,90	0,55	0,58	0,28
7	0,80	0,80	0,57	0,49	0,56
8	0,88	0,83	0,55	0,47	0,48
9	0,86	0,86	0,55	0,55	0,60
10	0,80	0,85	0,63	0,62	0,56
Trung bình	0,86	0,86	0,58	0,56	0,52

Bảng 4.3 mô tả thí nghiệm phát hiện các vật mốc với các điều kiện ánh sáng khác nhau và ở các khoảng cách khác nhau. Từ bảng này có thể thấy, ở khoảng cách

Chương 4: Nhận dạng và xác định vị trí vật mốc tự nhiên trong môi trường trong nhà

1 m và 2 m, giá trị IOU đều lớn hơn 0,8. Điều này có nghĩa là kết quả nhận dạng là tốt và các đặc trưng được trích xuất bằng bộ dò tìm ORB có tính bất biến tốt đối với những thay đổi về độ sáng. Ngoài ra, khi khoảng cách từ vật mốc đến camera lớn hơn 2 m, giá trị IOU nhỏ hơn 0,8 và giảm theo cường độ ánh sáng. Hình 4.11 mô tả hai vật mốc tự nhiên được phát hiện trong phòng thí nghiệm được mô tả trong bảng 4.3.

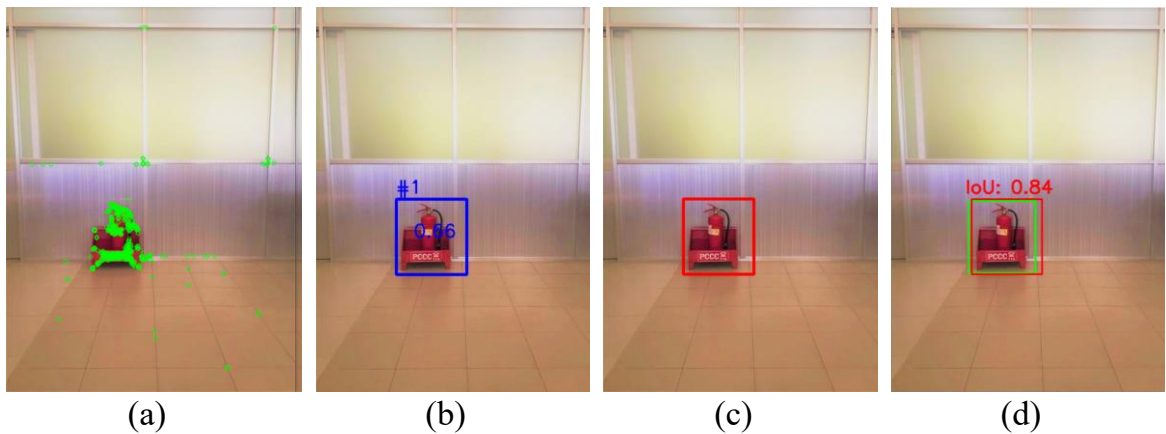


Hình 4.11. Nhận biết mốc tự nhiên trong môi trường phòng thí nghiệm với các điều kiện ánh sáng khác nhau: (a) Ảnh vật mốc 1; (b) Ảnh vật mốc 2.

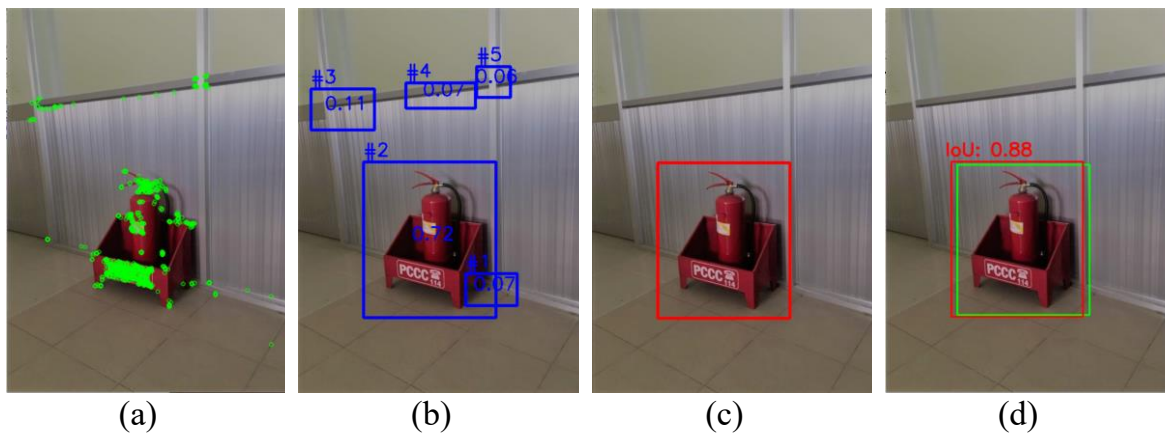
Bảng 4.3. IOU với điều kiện ánh sáng khác nhau.

Khoảng cách	Ánh sáng chuẩn		Thiếu sáng		Tối	
	Vật mốc 1	Vật mốc 2	Vật mốc 1	Vật mốc 2	Vật mốc 1	Vật mốc 2
1 m	0,84	0,86	0,82	0,85	0,84	0,86
2 m	0,84	0,82	0,90	0,80	0,90	0,82
3 m	0,77	0,59	0,70	0,50	0,51	0,45

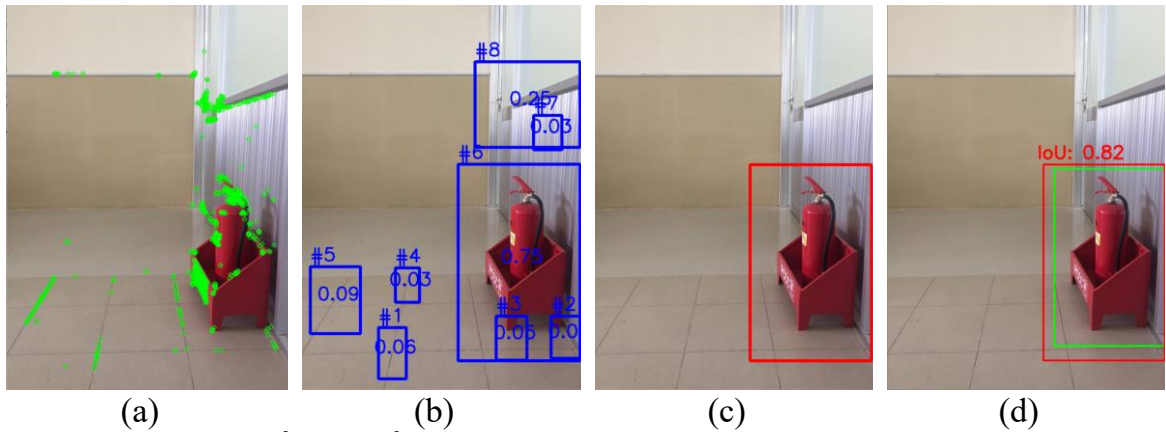
Các thí nghiệm với các góc khác nhau giữa vị trí camera và các vật mốc được thực hiện như hình 4.12, hình 4.13 và hình 4.14. Cụ thể, hình ảnh các đối tượng được chụp bằng hệ thống camera ở các góc lần lượt là 0° , 35° và 45° . Kết quả cho thấy với các góc chụp khác nhau nhưng thuật toán đề xuất đều nhận dạng cùng một đối tượng, và sẽ được chọn làm vật mốc. Ngoài ra, hình 4.12d, hình 4.13d và hình 4.14d còn chỉ ra rằng hộp giới hạn mà thuật toán dự đoán vật mốc (hình chữ nhật màu đỏ) chồng lấp lên gần hết hộp giới hạn của vật mốc thực tế (hình chữ nhật màu xanh lá). Giá trị IOU trong cả 3 trường hợp này đều lớn hơn 0,8, cho thấy độ chính xác là rất cao.



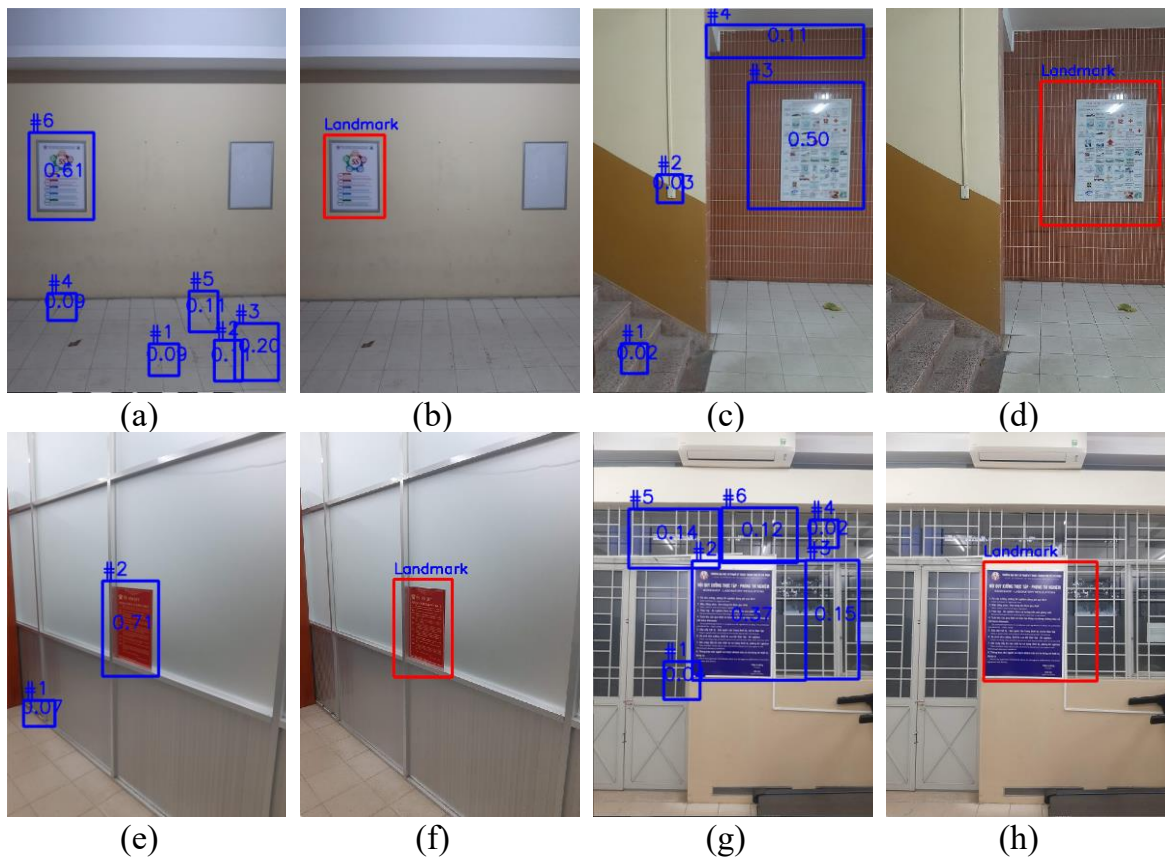
Hình 4.12. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 0° : (a) Ảnh có các điểm đặc trưng, (b) Ảnh các đối tượng, (c) Ảnh vật mốc nhận dạng được, (d) Hộp giới hạn với hệ số IOU



Hình 4.13. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 45° : (a) Ảnh có các điểm đặc trưng, (b) Ảnh các đối tượng, (c) Ảnh vật mốc nhận dạng được, (d) Hộp giới hạn với hệ số IOU



Hình 4.14. Nhận biết vật mốc tự nhiên trong môi trường phòng thí nghiệm theo góc 30° : (a) Ảnh có các điểm đặc trưng, (b) Ảnh các đối tượng, (c) Ảnh vật mốc nhận dạng được, (d) Hộp giới hạn với hệ số IOU



Hình 4.15. Các vật mốc tự nhiên được phát hiện từ các khu vực khác nhau: (a) Ảnh một bức tường trong hành lang; (b) Vật mốc được nhận dạng trên một bức tường trong hành lang; (c) Ảnh một bức tường khác trong hành lang; (d) Vật mốc được nhận dạng trên bức tường khác trong hành lang; (e) Ảnh bức tường phòng thí nghiệm; (f) Vật mốc được nhận dạng trên một bức tường của phòng thí nghiệm; (g) Ảnh bức tường khác của phòng thí nghiệm; (h) Vật mốc được nhận dạng trên một bức tường khác của phòng thí nghiệm.

Kết quả nhận dạng các vật mốc tự nhiên trong môi trường trong nhà được thể hiện trong hình 4.15. Cụ thể, các hình ảnh từ trái sang phải và các đối tượng được nhận dạng trong mỗi cặp ảnh này được viền bởi hình chữ nhật màu xanh lam và các đối tượng nổi bật nhất được chọn để là các vật mốc được đánh dấu bằng các đường viền màu đỏ. Hơn nữa, thời gian trung bình để xử lý mỗi bước được thể hiện trong bảng 4.4, trong đó mất nhiều thời gian nhất để tính toán xác định một vật mốc bằng cách tính mật độ đặc trưng của đối tượng. Với tổng thời gian nhận dạng một vật mốc trung bình là 38,08 ms, hệ thống cho thấy khả năng xử lý nhanh với thời gian thực.

Bảng 4.4. Thời gian xử lý nhận dạng vật mốc của phương pháp đề xuất.

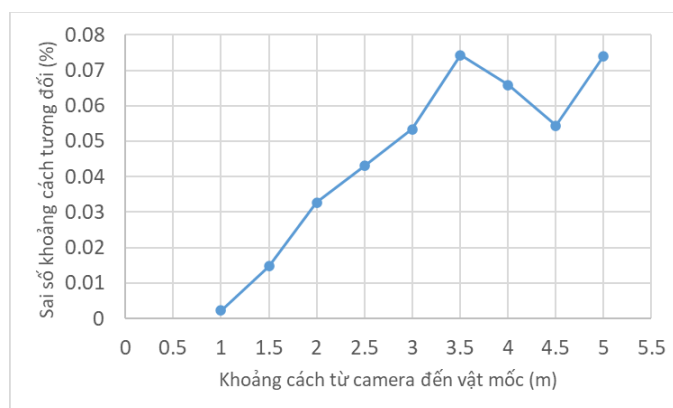
Bước thực hiện	Hình 4.15a	Hình 4.15c	Hình 4.15e	Hình 4.15g
Trích đặc trưng [ms]	2,90	3,77	2,89	4,77
Phát hiện đối tượng và nhận dạng vật mốc [ms]	37,10	37,93	36,12	41,16
Tổng [ms]	40,00	41,70	39,01	45,93

Trong môi trường trong nhà, các vật mốc nổi bật có thể được chọn bằng cách sử dụng phương pháp đề xuất và các vật mốc khác như cửa ra vào, cầu thang [132] và các khu vực có màu sắc đồng nhất đã được nhận dạng [133]. Do đó, xe lăn di động có thể dễ dàng nhận ra các loại vật mốc này trong môi trường tự nhiên phức tạp. Phương pháp mà luận án đề xuất và phương pháp trong các nghiên cứu [132, 133] đều nhằm mục đích thực hiện trực tuyến mà không cần các bước tiền huấn luyện khi phát hiện mốc. Cụ thể, nghiên cứu [132] đã đề xuất một phương pháp mạnh mẽ để phát hiện các vật mốc tự nhiên bằng cách sử dụng dữ liệu đám mây điểm 3D với độ chính xác trên 90% khi nhận dạng cửa ra vào và cầu thang. Ngoài ra, một kỹ thuật phân đoạn đối tượng dựa trên thị giác máy tính đã được đề xuất bởi nghiên cứu [133] để phát hiện các vật mốc tự nhiên cho robot di động trong môi trường trong nhà. Phương pháp này có thể phát hiện tường, cửa, sàn nhà và trần nhà bằng cách tìm các đường thẳng và tạo ra các đa giác lồi, với toàn bộ quá trình của phương pháp mất chưa đến 80 ms. Khác với các nghiên cứu này, phương pháp đề xuất của luận án chủ yếu tập trung vào việc phát hiện các đối tượng trong môi trường tự nhiên với nhiều vật mốc khác nhau. Hơn nữa, so với các phương pháp phát hiện đối tượng trong các

nghiên cứu [139, 140], phương pháp đề xuất của luận án dựa trên mật độ điểm đặc trưng của các vùng đối tượng đề xuất nên sẽ tìm ra ít đối tượng hơn trong ảnh. Nhờ đó, thời gian xử lý phát hiện đối tượng trong ảnh để lựa chọn và thu thập mốc sẽ nhanh hơn và có tiềm năng cho các ứng dụng thời gian thực.

4.5. KẾT QUẢ XÁC ĐỊNH VỊ TRÍ VẬT MỐC

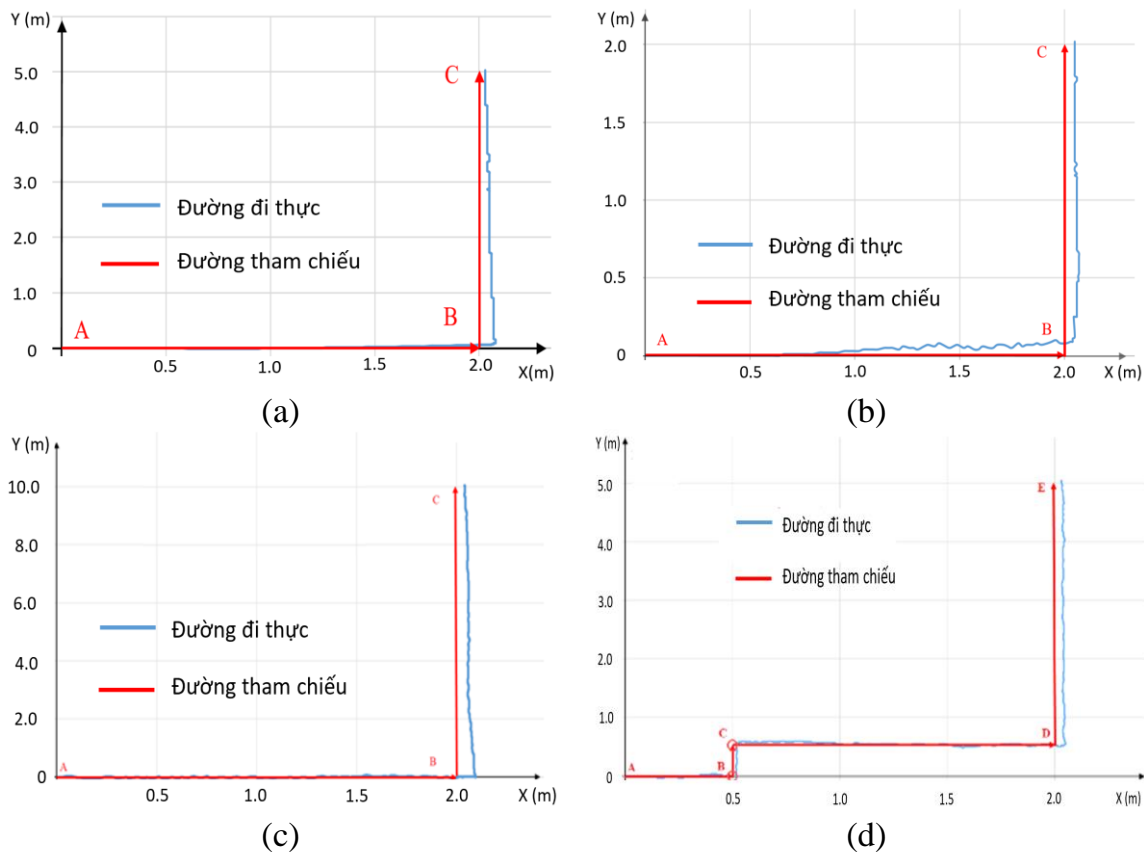
Phép đo khoảng cách từ camera đến các vật mốc bằng camera RealSense D435 được so sánh với vị trí thực của các vật mốc đó trong môi trường phòng thí nghiệm được thực hiện để đánh giá sai số phép đo. Hình 4.16 thể hiện sai số tương đối giữa khoảng cách đo trung bình và khoảng cách các vật mốc thực tế, trong đó mỗi vị trí được đo 100 lần. Rõ ràng là sai số khoảng cách tương đối có sự thay đổi lớn, từ 0,01% đến 0,07%. Khi thuật toán phát hiện đưa ra dự đoán vật mốc với giá trị IOU thấp, giá trị độ sâu có thể không còn chính xác. Điều này đúng với các vật thể nhỏ và ở xa, và được ghi nhận ở các đỉnh sai số tại khoảng cách 3,5 m và 5,0 m như trong hình 4.16. Hơn nữa, hình 4.16 còn cho thấy vùng hoạt động ổn định của camera là dưới 3,5 m với sai số phép đo là tuyến tính. Khi khoảng cách đo lớn hơn 3,5 m thì cảm biến hồng ngoại đo khoảng cách của camera không còn hoạt động ổn định nữa, dẫn đến giá trị các phép đo không còn chính xác. Do đó, thông qua thí nghiệm này chỉ ra rằng khoảng cách đo đến các vật mốc để có được độ chính xác cao nhất là dưới 3,5 m.



Hình 4.16. Biểu đồ sai số tương đối của phép đo khoảng cách tới các vật mốc.

Thí nghiệm tiếp theo là xác định sai số vị trí của xe lăn trong không gian 2D được kiểm tra bằng cách sử dụng dữ liệu từ encoder. Trong thí nghiệm này, xe lăn được di chuyển dọc theo một đường tham chiếu (đường màu đỏ) và sau đó dữ liệu vị

trí của xe lăn được ghi lại (đường màu xanh) như trong hình 4.17. Như hình 4.17a, xe lăn di chuyển đến vị trí $x = 2\text{ m}$, $y = 5\text{ m}$ trong không gian 2D theo một đường kẻ sẵn. Giả sử tọa độ ban đầu của xe lăn là $x = 0$, $y = 0$, như hình 4.17b, xe lăn đã dịch chuyển đến vị trí $x = 2\text{ m}$, $y = 2\text{ m}$. Trên hình 4.17c, xe lăn di chuyển đến vị trí $x = 2\text{ m}$, $y = 10\text{ m}$. Do đó, quỹ đạo của xe lăn chuyển động đến vị trí $x = 2\text{ m}$, $y = 5\text{ m}$ với các điểm vuông góc với nhau được mô tả trên hình 4.17d.



Hình 4.17. Biểu diễn quỹ đạo chuyển động của xe lăn đến các vị trí đặt trước:

(a) Thí nghiệm 1 (b) Thí nghiệm 2, (c) Thí nghiệm 3, (d) Thí nghiệm 4.

Sai số vị trí của xe lăn trong ba thí nghiệm như hình 4.17a, hình 4.17b và hình 4.17c được trình bày trong bảng 4.5. Cụ thể, vị trí xe lăn được tính toán có sai số theo trục X nhỏ hơn hoặc bằng 4,0 cm và nhỏ hơn hoặc bằng 2,0 cm theo trục Y. Sai số này của xe lăn là do trong quá trình di chuyển xe lăn bị đổi hướng gây ra sai số vị trí do quán tính của xe lăn. Hình 4.17d mô tả quỹ đạo đường đi thực tế (màu xanh) và đường tham chiếu (màu đỏ) của xe lăn. Khi di chuyển từ A đến B và từ B đến C, sai số vị trí trung bình của xe lăn nhỏ hơn 1,0 cm. Khi di chuyển từ C đến D, vị trí xe lăn

Chương 4: Nhận dạng và xác định vị trí vật mốc tự nhiên trong môi trường trong nhà

có sai lệch khoảng 4,0 cm theo trục X. Khi kết thúc tại điểm E, tọa độ của xe lăn được tính toán có sai số so với tọa độ thực là 1,5 cm theo trục X và 2,0 cm theo trục Y.

Bảng 4.5. Đánh giá sai số vị trí của xe lăn – Đơn vị: cm

TT	Vị trí thực tế	Vị trí tính toán	$ \Delta x $	$ \Delta y $
1	(200,0; 500,0)	(203,0; 502,0)	3,0	2,0
2	(200,0; 200,0)	(204,0; 201,0)	4,0	1,0
3	(200,0; 100,0)	(201,0; 99,0)	1,0	1,0

Bảng 4.6 mô tả kết quả xác định vị trí vật mốc với các thí nghiệm khác nhau. Cụ thể là, xe lăn di chuyển đến các vị trí khác nhau và thu thập các vật mốc. Vị trí vật mốc được tính toán dựa vào phương pháp đề xuất và so sánh với vị trí thực của nó trong môi trường. Từ bảng 4.6 có thể thấy với khoảng cách từ camera đến vật mốc dưới 200 cm, vị trí vật mốc có sai số nhỏ hơn 3,0 cm theo trục X và nhỏ hơn 2,0 cm theo trục Y. Trong trường hợp khoảng cách từ camera đến vật mốc lớn hơn 200 cm, có thể thấy sai số cũng tăng lên, cụ thể là lớn hơn 15,0 cm theo trục X và lớn hơn 7.0 cm theo trục Y. Điều này cho thấy nên thu thập và gán nhãn vị trí các vật mốc trong môi trường với khoảng cách nhỏ hơn 200 cm so với camera để có thể đạt được độ chính xác cao cho vị trí của vật mốc.

Bảng 4.6. Kết quả định vị vật mốc dựa vào vị trí xe lăn – Đơn vị: cm

Vị trí xe lăn (x_w, y_w, θ_w)	Khoảng cách đến vật mốc	Vị trí vật mốc thực tế	Vị trí vật mốc tính toán theo phương pháp đề xuất		
				$ \Delta x $	$ \Delta y $
(30,0; 30,0; 90)	96,8	(60,6; 121,1)	(63,1; 120,9)	2,5	0,2
(30,0; 30,0; 45)	85,9	(90,7; 90,7)	(92,2; 89,2)	1,5	1,5
(40,0; 40,0; 60)	112,7	(120,0; 120,0)	(119,5; 119,9)	0,5	0,1
(40,0; 40,0; 30)	111,0	(120,0; 120,0)	(128,6; 106,9)	8,6	13,1
(80,0; 40,0; 120)	89,1	(40,0; 120,0)	(38,7; 119,0)	1,3	1,0
(60,5; 60,5; 90)	94,4	(85,0; 151,4)	(82,4; 152,3)	2,6	0,9
(324,8; 116,6; 45)	207,0	(495,2; 124,2)	(511,1; 206,9)	15,9	7,3

Với kết quả từ các thí nghiệm đã cho thấy phương pháp nhận dạng vật mốc tự nhiên và xác định vị trí của vật mốc trong môi trường là rất tiềm năng cho việc định vị xe lăn hoặc robot di động trong môi trường trong nhà. Hơn nữa, phương pháp đề

Chương 4: Nhận dạng và xác định vị trí vật mốc tự nhiên trong môi trường trong nhà

xuất có thể ứng dụng vào việc xây dựng một công cụ gán nhãn vị trí tự động cho các vật mốc trong trong nhà, lĩnh vực mà các nghiên cứu khác cũng đang hướng tới. Bảng 4.7 liệt kê các nghiên cứu về nhận dạng các đối tượng trong ảnh và thu thập thông tin vị trí của các đối tượng một cách tự động. Cụ thể là, nghiên cứu [141] đề xuất một công cụ chú thích mới dựa trên AR tự động để phát hiện và gán nhãn các đối tượng nổi bật như bàn, ghế, trần nhà thấp cho các ứng dụng điều hướng trong nhà. Nghiên cứu này đã dùng YOLOv3 để phát hiện đối tượng, định vị theo kinh độ và vĩ độ, và gán nhãn đối tượng vào cơ sở dữ liệu. Tuy nhiên, việc dùng vị trí vật thể theo kinh độ và vĩ độ như trong [141] phù hợp với môi trường rộng lớn, còn với môi trường nhỏ cần độ chính xác cao hơn thì phương pháp định vị vị trí của các vật mốc trong không gian 2D của luận án sẽ phù hợp hơn.

Bảng 4.7. Các nghiên cứu nhận dạng, thu thập thông tin vị trí đối tượng tự động

Nghiên cứu	Đối tượng	Kỹ thuật nhận dạng đối tượng	Tập dữ liệu huấn luyện	Thời gian xử lý	Mô tả tập dữ liệu thu thập được
X. Chai [133]	Cửa, tường, trần và sàn	Phân đoạn đối tượng	Không sử dụng	75 ms	Vật mốc trong môi trường trong nhà
P. Du [141]	Bàn, ghế và trần nhà thấp	YOLOv3	Millar Library	-	Các đối tượng cùng với kinh độ và vĩ độ của chúng
Apud Baca [142]	Đồ chơi trẻ em	CNNs	MS COCO	40 s	Tư thế sáu bậc tự do (6-DoF) của đối tượng
O. Deane [143]	Dữ liệu theo dõi mắt di động	Mask R-CNN	MS COCO	1,5 s	Tọa độ ánh mắt
García-Aguilar [144]	Phương tiện giao thông	EfficientDet D4	COCO	-	Các đối tượng cùng với vị trí trong ảnh của đối tượng
Phuong pháp đề xuất	Vật mốc tự nhiên	Mật độ đặc trưng lớn nhất	Không sử dụng	41,66 ms	Vật mốc và vị trí trong môi trường trong nhà của vật mốc

Cùng với nghiên cứu [142-144], nghiên cứu của luận án này giúp quy trình thu thập và gán nhãn đối tượng được thực hiện tự động, giúp giảm rất nhiều thời gian

thực hiện so với quy trình gán nhãn thủ công. Cụ thể, nghiên cứu [142] trình bày một cách tiếp cận mới giúp tự động hóa quá trình thu thập và chú thích dữ liệu về tư thế sáu bậc tự do (6-DoF) của các đối tượng trong ảnh, và giúp giảm thiểu thời lượng thực hiện từ khoảng 4.5 giờ xuống còn 40 giây. Điều này là cần thiết cho nhiều ứng dụng như chế tạo robot, lái xe tự động và điều hướng dựa trên thị giác máy tính, trí tuệ nhân tạo và tự hành cho các hệ thống máy bay không người lái (UAS). Nghiên cứu [143] chỉ ra rằng các phương pháp trước đây dựa vào chú thích thủ công tốn thời gian của dữ liệu theo dõi bằng mắt và đề xuất một hệ thống sử dụng thuật toán phát hiện đối tượng dựa trên học sâu Mask R-CNN, kết hợp với tọa độ ánh nhìn theo từng khung hình được đo bằng thiết bị theo dõi ánh mắt, để phát hiện và chú thích mà không cần bất kỳ sự can thiệp thủ công nào. Ngoài ra, kỹ thuật phân đoạn đối tượng [133] đã được áp dụng để phát hiện mốc tự nhiên trên robot di động trong nhà và thời gian thực hiện phát hiện đối tượng là khoảng 75 ms, so với 41,66 ms khi sử dụng phương pháp đề xuất của luận án.

❖ Kết luận

Phương pháp phát hiện các vật mốc tự nhiên và xác định vị trí của các vật mốc trong môi trường trong nhà với độ chính xác cao và thời gian xử lý nhanh đã được đề xuất trong chương này [134, 137]. Cụ thể, các vật mốc tự nhiên được phát hiện dựa trên mật độ của các điểm đặc trưng lớn nhất của các đối tượng trong ảnh. Ngoài ra, thông tin vị trí của các vật mốc được xác định dựa vào thông tin 3D của camera RGB-D được lắp đặt trên xe lăn và từ vị trí của xe lăn. Trong các thí nghiệm, các vật mốc được phát hiện có độ chính xác cao thông qua việc tính toán các giá trị IOU. Kết quả nhận dạng cho thấy giá trị IOU lớn hơn 0,8 ở khoảng cách dưới 2 m từ camera đến vật mốc, đạt hiệu suất phát hiện tốt nhất. Hơn nữa, vị trí của các vật mốc được tính toán từ hệ thống này có sai số nhỏ so với vị trí thực tế và tốc độ xử lý của hệ thống phù hợp với quá trình xử lý thời gian thực trong chuyển động của xe lăn. Tuy nhiên, để hệ thống đạt được hiệu suất tốt nhất thì môi trường hoạt động phải đủ sáng. Các hình ảnh và vị trí của vật mốc trong môi trường tự nhiên sau khi được thu thập, sẽ được sử dụng để định vị xe lăn trên bản đồ, phục vụ cho việc điều hướng xe lăn điện.

CHƯƠNG 5

MÔ HÌNH ĐIỀU KHIỂN XE LĂN ĐIỆN KẾT HỢP TÍN HIỆU EEG VÀ CAMERA DỰA VÀO BẢN ĐỒ

Chương này trình bày mô hình điều khiển bán tự động có sự kết hợp tín hiệu EEG, camera, vật mốc và bản đồ lưới 2D ảo để điều khiển xe lăn đến đích mong muốn. Hệ thống điều khiển thực - ảo bao gồm công cụ lập kế hoạch di chuyển (MP) và bộ chuyển đổi hành động của xe lăn (WAC) được trình bày. Cụ thể, trong giai đoạn đầu tiên, bản đồ lưới 2D ảo được xây dựng dựa trên môi trường thực, bao gồm không gian trống, chướng ngại vật, vật mốc và những điểm đích. Hơn nữa, xe lăn cần xác định vị trí hiện tại của nó trong cả môi trường thực và ảo sử dụng thông tin của các vật mốc tự nhiên. Với vị trí bắt đầu và đích đến được xác định, MP sử dụng mô hình Deep Q-Networks (DQNs) và bản đồ lưới 2D ảo để hoạch định đường đi tối ưu cho xe lăn. Sau đó, WAC sẽ đưa ra các lệnh điều khiển thực tế cho xe lăn từ đường đi mà MP đề xuất để xe lăn có thể hoàn thành lịch trình của nó. Cuối cùng, các thí nghiệm trong mô phỏng và trong môi trường thực được thực hiện để đánh giá hiệu suất của mô hình đề xuất.

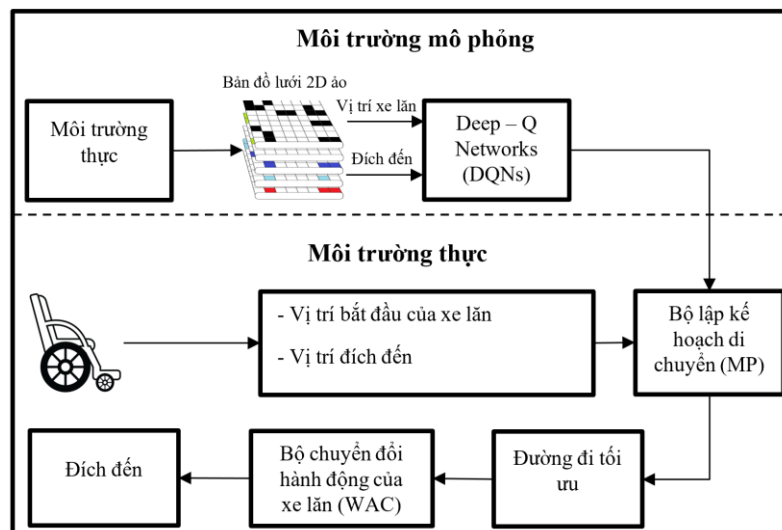
5.1. HỆ THỐNG ĐIỀU KHIỂN THỰC - ẢO CHO XE LĂN ĐIỆN DỰA VÀO BẢN ĐỒ LƯỚI 2D ẢO

5.1.1. Cấu trúc hệ thống điều khiển thực -ảo cho xe lăn điện

Trong hệ thống điều khiển tự động cho xe lăn điện đến đích mong muốn, việc tìm đường đi phù hợp có khả năng dẫn tới đích và là đường đi ngắn nhất đóng vai trò quan trọng. Đường đi này có thể được tạo ra bằng cách lập trình trước hoặc cũng có thể có được bằng cách cho xe lăn dò đường. Tuy nhiên, nhược điểm của các phương

pháp này là đường đi có thể không đúng hoặc không phải là đường đi ngắn nhất. Bên cạnh đó, bản đồ môi trường càng phức tạp thì độ khó trong việc tìm đường càng tăng lên. Ngoài các phương pháp trên, các thuật toán tự động có thể được áp dụng để tìm đường đi cho xe lăn. Ưu điểm của phương pháp này là có thể đưa ra một lộ trình di chuyển một cách tự động, và có thể tìm được đường đi ngắn nhất. Tốc độ của việc tìm đường sẽ phụ thuộc vào loại thuật toán được áp dụng. Tuy nhiên, để có thể áp dụng được thuật toán tìm đường tự động thì cần phải có thông tin về bản đồ.

Trong luận án này, một hệ thống điều khiển thực - ảo cho xe lăn điện được đề xuất để điều hướng xe lăn di chuyển đến đích mong muốn [145]. Cấu trúc hệ thống này bao gồm hai giai đoạn dành cho xe lăn điện trong môi trường trong nhà được mô tả trong hình 5.1. Ở giai đoạn đầu tiên, bản đồ lưới 2D ảo với các ô lưới được mô phỏng dựa trên một môi trường thực trong nhà sẽ cung cấp thông tin về tọa độ vị trí hiện tại và đích cần đến của xe lăn, làm đầu vào của mô hình DQNs. Sau khi được huấn luyện, mô hình DQNs sẽ có các tham số tối ưu để có thể ước tính giá trị Q của tất cả các hành động có thể xảy ra đối với các vị trí của xe lăn trên bản đồ lưới 2D ảo. Do đó, mô hình DQNs sẽ có 4 đầu ra tương ứng với 4 hành động (Lên, Xuống, Trái, Phải). Mỗi bản đồ lưới 2D ảo chỉ được xây dựng cho một đích đến trong một môi trường thực trong nhà, do đó, mỗi mô hình DQNs thu được cho một MP.



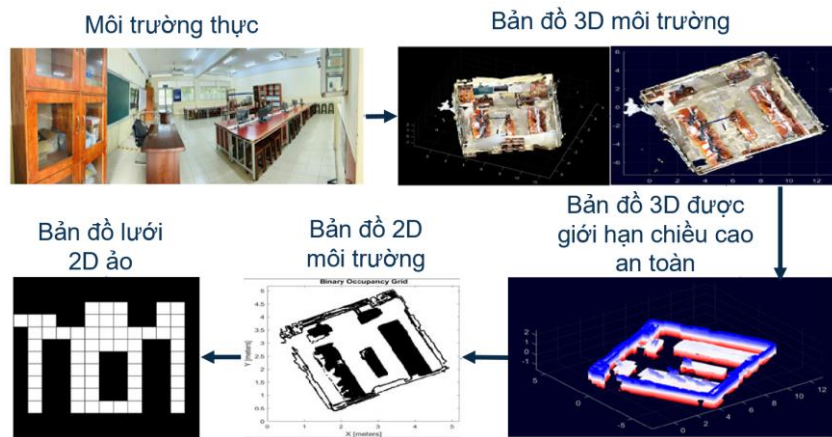
Hình 5.1. Hệ thống điều khiển thực - ảo cho xe lăn điện dựa trên bản đồ lưới 2D ảo.

Ở giai đoạn thứ hai, xe lăn sẽ được điều khiển để đi đến đích mong muốn trong một môi trường thực tế trong nhà. Tại thời điểm xuất phát, hệ thống camera sẽ giúp xe lăn nhận biết các vật cản tự nhiên, định vị vị trí của nó trong môi trường thực, và từ đó sẽ tính toán được vị trí của xe lăn trên bản đồ lưới 2D. Bên cạnh đó, người dùng sẽ lựa chọn đích đến của xe lăn bằng tín hiệu EEG trên giao diện máy tính. Khi nhận được vị trí ban đầu của xe lăn và điểm đích đến, mô hình DQNs sẽ ước tính giá trị Q của 4 đầu ra tương ứng với 4 hành động (Lên, Xuống, Trái, Phải). Do đó, hành động có giá trị Q cao nhất sẽ được chọn. Với hành động này, một vị trí mới trên bản đồ lưới sẽ được cập nhật và sau đó vị trí mới này sẽ là đầu vào cho mô hình DQNs và cũng sẽ chọn một hành động tương ứng. Quá trình này sẽ lặp lại và chỉ kết thúc khi vị trí là đích đến. Sau khi có được đường đi tối ưu để đến đích mong muốn, MP với một chuỗi hành động (Phải, Trái, Lên, Xuống) và WAC sẽ cho phép xe lăn di chuyển theo đường đi tối ưu này.

5.1.2. Bản đồ lưới 2D ảo

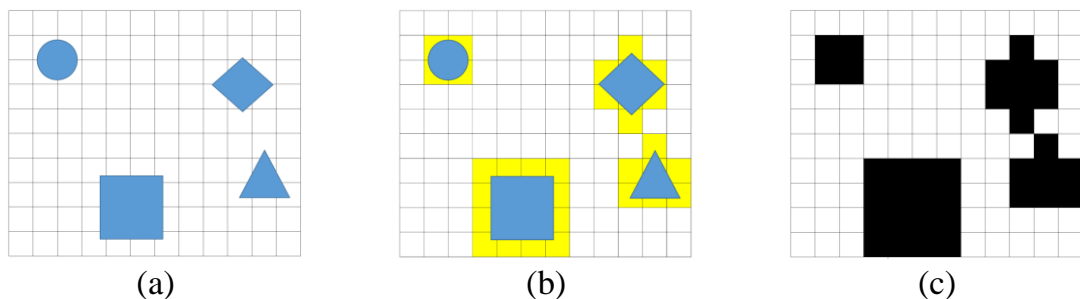
Trong quy hoạch đường đi tối ưu, bản đồ lưới 2D ảo đóng một vai trò rất quan trọng vì biểu diễn các đường đi tối ưu để điều hướng xe lăn điện đến các điểm đích. Cụ thể, một bản đồ lưới 2D ảo được mô phỏng dựa trên nhiều thông tin liên quan đến một môi trường thực. Điều đó có nghĩa là xe lăn trong quá trình di chuyển trong một môi trường thực tế có thể sử dụng các tham số và vị trí để điều hướng xe lăn. Hình 5.2 mô tả quá trình tạo ra bản đồ lưới 2D ảo từ môi trường thực. Để tạo ra một bản đồ lưới 2D ảo mà xe lăn có thể di chuyển an toàn, đầu tiên cần xây dựng một bản đồ 3D môi trường [30, 146]. Với bản đồ 3D này, các điều kiện giới hạn chiều cao an toàn cho xe lăn được áp dụng để loại bỏ các vùng có chiều cao lớn hơn chiều cao của xe lăn. Từ đó, một phép chiếu hình học được thực hiện để chiếu bản đồ 3D thành bản đồ 2D bao gồm các vùng trống (an toàn về chiều cao) và vùng vật cản [147]. Cuối cùng, bản đồ 2D này sẽ được chia thành nhiều ô, mỗi ô có thể là một khoảng trống hoặc một vùng chiếm chỗ (chướng ngại vật) để tạo ra bản đồ lưới 2D ảo. Do đó, giả

sử rằng một chiếc xe lăn có thể được điều khiển để đi qua các khu vực không gian trống này để đến được một điểm đích mong muốn.



Hình 5.2. Quá trình tạo bản đồ lưới 2D ảo từ môi trường thực.

Hình 5.3 mô tả chi tiết về một bản đồ lưới 2D ảo bao gồm $m \times n$ ô lưới trong môi trường trong nhà mà xe lăn có thể di chuyển qua để đến đích. Cụ thể, môi trường thực với các đối tượng (màu xanh) như hình 5.3a được đo đạc và chia thành các ô có kích thước bằng kích thước xe lăn. Tiếp theo, bản đồ với các ô được chia sẽ được lấp đầy các ô liên quan đến chướng ngại vật (màu vàng). Do đó, bản đồ trong hình 5.3b được xấp xỉ để tạo bản đồ lưới 2D ảo như được mô tả trong hình 5.3c. Các ô trong bản đồ lưới 2D ảo được gán là 1 để biểu thị không gian bị chiếm dụng (chướng ngại vật) và được gán là 0 cho không gian trống. Do đó, bản đồ lưới 2D ảo này được coi là bản đồ nhị phân với các ô đen trắng và tọa độ gốc của bản đồ ảo nằm ở góc trên cùng bên trái với vị trí đầu tiên (0,0). Rõ ràng bản đồ ảo này cho chúng ta biết thông tin tất cả các vị trí và được sử dụng để tìm đường đi tối ưu bằng mô hình DQNs.

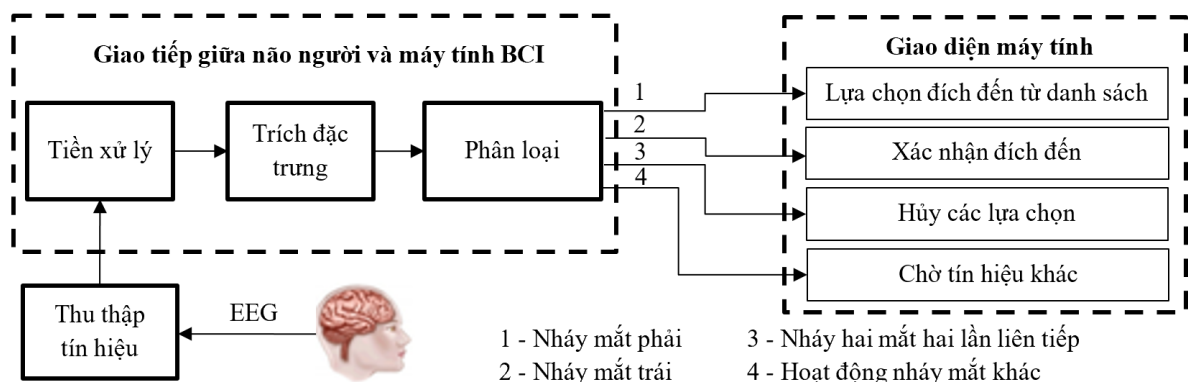


Hình 5.3. Bản đồ lưới 2D ảo của môi trường thực: (a) Bản đồ môi trường với các chướng ngại vật và vùng trống; (b) Các ô vật cản liên quan đến các chướng ngại vật thực sự; (c) Bản đồ lưới 2D ảo với các ô vật cản màu đen.

5.1.3. Giao diện lựa chọn đích đến cho xe lăn

Trong hệ thống xe lăn bán tự động, việc xây dựng và lựa chọn điểm đến trong bản đồ lưới cho người khuyết tật nặng là một nhiệm vụ hết sức quan trọng. Đối với những người khuyết tật nặng không thể sử dụng các điều khiển thông thường như bấm nút, điều khiển cần điều khiển hay chạm vào màn hình điều khiển thì tín hiệu điện não đồ EEG để điều khiển xe lăn bán tự động là một lựa chọn hữu ích. Việc sử dụng tín hiệu EEG để điều khiển trực tiếp xe lăn bán tự động có thể gây căng thẳng do phải tập trung trong thời gian dài, do đó người dùng có thể lựa chọn điểm đến mong muốn thông qua giao diện màn hình với các câu lệnh được thiết kế phù hợp với môi trường thực tế của mình [123]. Các lệnh trên màn hình giao diện được chỉ định dựa trên loại tín hiệu EEG từ hoạt động mắt của người dùng.

Hình 5.4 mô tả quá trình thu thập, xử lý và phân loại tín hiệu EEG để thực hiện các lệnh điều khiển liên quan đến giao diện người dùng. Tín hiệu EEG được thu thập từ thiết bị Emotiv EPOC+ tại các điện cực nằm ở vỏ não trước trán được coi là những tín hiệu đáng tin cậy nhất. Sau đó, tín hiệu EEG được chuyển đến khối tiền xử lý tín hiệu để lọc nhiễu trước khi gửi đến khối trích xuất đặc trưng. Tín hiệu EEG sau đó sẽ được phân loại để tạo ra các lệnh điều khiển [80, 123, 125]. Điều này có nghĩa là người dùng có thể sử dụng các lệnh điều khiển để chọn một trong các điểm đến trên bản đồ môi trường.



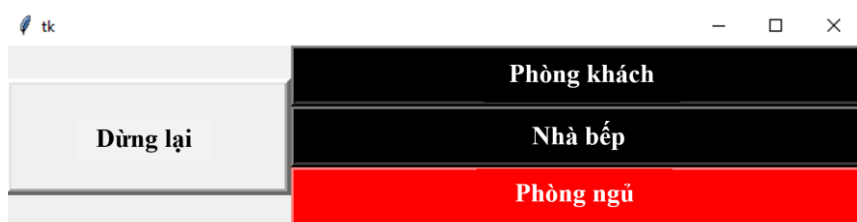
Hình 5.4. Quy trình thực hiện lựa chọn đích đến dùng BCI.

Giao diện người dùng luôn được thiết kế đơn giản và dễ dàng cho người khuyết tật, đặc biệt tất cả các lệnh chỉ có thể được thao tác bằng BCI như mô tả trong hình

5.5. Trên giao diện, người dùng sẽ thấy một danh sách dọc các biểu tượng tên đích đến. Tên trong danh sách này là các điểm đến được xác định trước như phòng khách, nhà bếp và phòng ngủ. Để điều khiển các lệnh lựa chọn đích đến, hoạt động nháy mắt của người dùng được thực hiện. Dựa trên kết quả phân loại tốt nhất của các hoạt động mắt được trình bày ở chương 3, ba loại nháy mắt là nháy mắt trái, nháy mắt phải và nháy hai mắt hai lần liên tiếp được lựa chọn. Trong đó, người dùng cần nháy mắt phải trong để có thể di chuyển con trỏ trên màn hình đến điểm cần đến rồi nháy mắt trái để xác nhận điểm đến mong muốn như hình 5.6. Nếu người dùng muốn hủy các lệnh đã chọn hoặc hủy đích đã chọn, người dùng cần thực hiện thao tác nháy hai mắt hai lần liên tiếp. Nếu chương trình của giao diện nhận được tín hiệu nháy mắt khác với ba loại nháy mắt trên, chương trình sẽ không thực hiện các lệnh điều khiển và chờ tín hiệu nháy mắt từ người dùng. Tất cả các thao tác được chọn để điều khiển giao diện người dùng với các đích đến được thiết kế đã được thử nghiệm trên nhiều người dùng và kết quả thực tế sử dụng cho độ chính xác cao.



Hình 5.5. Giao diện người dùng để chọn điểm đến mong muốn.



Hình 5.6. Giao diện người dùng chọn điểm đến mong muốn “Phòng ngủ” bằng lệnh EEG.

5.1.4. Mô hình DQNs lập kế hoạch đường đi tối ưu cho xe lăn

Mô hình DQNs được ứng dụng để tìm đường đi tối ưu dựa trên bản đồ lưới 2D ảo thông qua mô phỏng, trong đó mỗi đích đến sử dụng một bản đồ lưới ảo và

cũng tìm được nhiều đường đi tối ưu để một xe lăn có thể đến đích từ bất kỳ điểm xuất phát nào. Bản đồ lưới 2D ảo sẽ bao gồm các chướng ngại vật và không gian trống. Với các vị trí bắt đầu và đích đến, nhiệm vụ của xe lăn là đến được ô đích. Ngoài ra, xe lăn di chuyển trong môi trường dựa trên các hành động (Trái, Phải, Lên, Xuống). Sau mỗi hành động, xe lăn sẽ có vị trí mới $S_t = (x_t, y_t)$ trong môi trường, được gọi là vị trí xe lăn tại thời điểm t , với tọa độ lưới (x, y) , và điểm thưởng R tương ứng với vị trí đó. Ngoài ra, vị trí của xe lăn là giới hạn, $S_t \in S$, với kích thước S là $m \times n$ được xác định trước. Xe lăn thường được đặt ở giữa các ô lưới để có thể di chuyển theo cả bốn hướng.

Trong mô hình này, các vị trí trên bản đồ lưới 2D bao gồm ba loại là chướng ngại vật S_o , không gian trống S_f và đích đến S_g . Tại mỗi thời điểm t , xe lăn tại vị trí S_t và cần chọn một hành động từ một tập cố định các hành động. Do đó, quyết định chọn hành động nào cho việc di chuyển chỉ phụ thuộc vào vị trí hiện tại, không phụ thuộc vào lịch sử các hành động trước đó. Hơn nữa sau mỗi hành động, xe lăn sẽ thực hiện di chuyển từ vị trí hiện tại S_t sang vị trí mới S_{t+1} tại thời điểm $(t + 1)$, sau đó phần thưởng nhận được sau mỗi hành động là $R(s_t, a_t) \in [-1, 1]$ được tính toán theo quy tắc:

$$R(s_t, a_t) = \begin{cases} R_f & \text{với } a_t = s_t \rightarrow s_f \\ R_g & \text{với } a_t = s_t \rightarrow s_g \\ R_o & \text{với } a_t = s_t \rightarrow s_o \end{cases} \quad (5.1)$$

Mỗi lần di chuyển của xe lăn từ ô này sang ô liền kề trên bản đồ lưới sẽ mất R_f điểm để ngăn xe lăn đi lang thang, và có thể đến mục tiêu mong muốn bằng con đường ngắn nhất. Trong thuật toán này, phần thưởng tối đa là R_g điểm cho chuyển động của xe lăn đến đúng đích. Nếu xe lăn cố gắng đi vào một ô có chướng ngại vật, nó sẽ bị trừ R_o điểm. Đây là một hình phạt nghiêm trọng nên xe lăn sẽ học cách tránh và do đó việc nỗ lực di chuyển đến ô có chướng ngại vật là không thể thực hiện được. Quy tắc tương tự đối với nỗ lực di chuyển ra ngoài ranh giới bản đồ lưới, xe lăn sẽ bị trừ R_b điểm. Xe lăn cũng sẽ bị trừ R_p điểm cho bất kỳ chuyển động nào đến ô đã đi qua. Ngoài ra, để tránh các vòng lặp vô hạn trong quá trình huấn luyện sử dụng DQNs, tổng phần thưởng đạt được phải lớn hơn ngưỡng âm ($thr \times m \times n$) và xe lăn có thể di

chuyên bình thường. Ngược lại, xe lăn có thể bị lạc không thể đến đích và mắc nhiều lỗi nên việc huấn luyện cần được thực hiện lại cho đến khi đủ tổng điểm thưởng.

Trong mô hình DQNs này, mạng nơ-ron truyền thẳng (FWNN) với thuật toán huấn luyện lan truyền ngược được sử dụng. Cụ thể là, các vị trí của xe lăn trong môi trường là đầu vào của mạng và đầu ra sẽ là phần thưởng cho mỗi vectơ hành động. Mục tiêu của xe lăn là di chuyển theo bản đồ theo một chiến lược để nhận được phần thưởng tối đa từ môi trường. Do đó, chiến lược π cho vị trí S_t sẽ đưa ra một hành động sao cho tổng phần thưởng Q mà xe lăn nhận được lớn nhất và được tính theo phương trình sau:

$$\pi(s_t) = \arg \max_{i=0,1,\dots,n} Q(s_t, a_i) \quad (5.2)$$

$$Q(s_t, a_t) = R(s_t, a_t) + \gamma \cdot \max_{i=0,1,\dots,n} Q(s_{t+1}, a_i) \quad (5.3)$$

trong đó $Q(s_t, a_i)$ là giá trị điểm thưởng khi thực hiện các hành động a_i ($i = 0, 1, \dots, (n-1)$) tại vị trí S_t ; n là số lượng hành động; s_{t+1} là trạng thái tiếp theo; γ là hệ số chiết khấu đảm bảo rằng khi xe lăn ở xa ô đích thì giá trị Q càng nhỏ.

Để tính gần đúng $Q(s_t, a_t)$, FWNN có đầu vào là vị trí của xe lăn trên bản đồ lưới và đầu ra của nó là vectơ giá trị Q , trong đó mỗi giá trị Q tương ứng với hành động n . Ngoài ra, Q_i là giá trị xấp xỉ của $Q(s_t, a_{ti})$ cho mỗi hành động a_{ti} . Khi mạng nơ-ron được huấn luyện đầy đủ và chính xác, nó sẽ được sử dụng trong mô hình hoạch định đường đi tối ưu để lựa chọn chiến lược π như phương trình sau:

$$\pi(s_t) = a_j \quad (5.4)$$

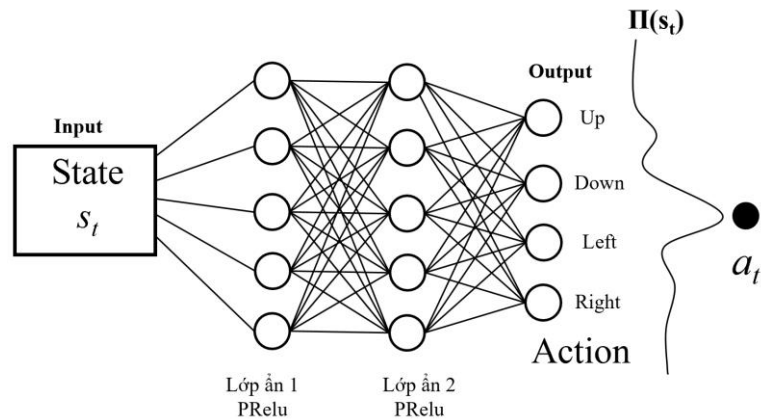
$$j = \arg \max_{i=0,1,\dots,n} Q_i \quad (5.5)$$

trong đó giá trị j được xác định dựa trên giá trị Q lớn nhất.

Mục đích của mạng nơ-ron là ước tính chính xác giá trị Q cho các hành động, vì vậy hàm mục tiêu được áp dụng ở đây là tính toán bình phương sai số giữa giá trị thực tế và giá trị dự đoán Q và được mô tả bằng phương trình sau:

$$Loss = \left(R(s_t, a_t) + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t) \right)^2 \quad (5.6)$$

Ngoài ra, mô hình FWNN có đầu vào là vị trí hiện tại và đầu ra là giá trị Q của các hành động. Tuy nhiên, nếu đầu vào của FWNN liên tục được đẩy vào từng vị trí của xe lăn trên bản đồ lưới thì rất dễ xảy ra hiện tượng quá khớp vì các vị trí thường giống nhau và tuyến tính. Để loại bỏ vấn đề quá khớp trong mô hình FWNN, kỹ thuật Experience Replay [117] được áp dụng. Cụ thể là, thay vì với mỗi vị trí, thông số mạng được cập nhật một lần, thì các vị trí sẽ được lưu vào bộ nhớ sau đó được lấy mẫu thành các lô nhỏ kết nối với đầu vào của FWNN để huấn luyện. Việc làm này có thể giúp đa dạng hóa đầu vào cho FWNN và cũng tránh được vấn đề quá khớp. Trong trường hợp này, mô hình huấn luyện sẽ quên các mẫu cũ không tốt cho quá trình huấn luyện và sau đó chúng sẽ bị xóa khỏi bộ nhớ.



Hình 5.7. Cấu trúc mạng nơ-ron trong mô hình DQNs.

Mô hình FWNN được sử dụng trong huấn luyện có hai lớp ẩn với số nút bằng số ô trong bản đồ lưới 2D ảo được xây dựng từ môi trường trong nhà như hình 5.7. Ngoài ra, kích thước của lớp đầu vào bằng kích thước lớp ẩn, trong khi lớp đầu ra có cùng số nút bằng với số lượng hành động (bốn hành động được sử dụng trong luận án này). Cuối cùng, mô hình FWNN sẽ chọn giá trị Q lớn nhất để thực hiện hành động tương ứng cho vị trí tiếp theo tiếp theo. Trong nghiên cứu này, hàm kích hoạt PreLU như mô tả ở phương trình (5.7), phương pháp tối ưu RMSProp và hàm mất mát của MSE được áp dụng.

$$f(y_i) = \max(0, y_i) + a_i \min(0, y_i) \quad (5.7)$$

Trong phương trình (5.7), y_i là đầu vào bất kỳ trên lớp thứ i và a_i là hệ số góc âm, là tham số có thể học được.

5.2. ĐỊNH VỊ XE LĂN ĐIỆN TRONG BẢN ĐỒ LƯỚI 2D ẢO DỰA VÀO VẬT MỐC TỰ NHIÊN

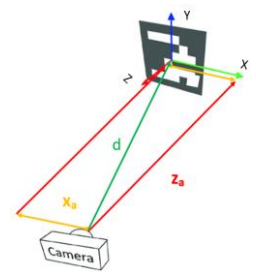
Hình ảnh với các vật mốc tự nhiên thường chứa nhiều điểm đặc trưng. Vì vậy, khi xe lăn được lắp đặt hệ thống camera để phát hiện các vật mốc trong quá trình di chuyển là cần thiết. Đặc biệt, chỉ những vật mốc tự nhiên được chọn mới có nhiều đặc điểm nổi bật giống với đặc điểm của các vật mốc trong cơ sở dữ liệu. Với các vật mốc được chọn như trong hình 5.8, thông tin vị trí của chúng cũng sẽ được lưu trữ trong cơ sở dữ liệu để phục vụ cho quá trình định vị xe lăn điện đang di chuyển. Trong quá trình xe lăn di chuyển trong môi trường thực với các vật mốc đã chọn trước đó, nếu camera nhận ra các vật mốc này, nó sẽ cung cấp thông tin về vị trí của mốc so với vị trí của camera như hình 5.9 [115]. Ngoài ra, khoảng cách từ camera đến vật mốc được phát hiện sẽ được tính toán trong điều kiện ánh sáng phù hợp. Từ hình 5.9, khoảng cách d từ tâm camera đến tâm mốc được tính như sau:

$$d = \sqrt{x_a^2 + z_a^2} \quad (5.8)$$

trong đó z_a là khoảng cách từ camera đến mặt phẳng chứa vật mốc, x_a là độ lệch của tâm mốc so với tâm camera, trong đó giá trị $x_a > 0$, nếu vật mốc ở bên phải camera và $x_a < 0$ nếu vật mốc nằm bên trái camera.



Hình 5.8. Các mẫu vật mốc trong môi trường trong nhà.



Hình 5.9. Hệ tọa độ của camera RGB-D

Hình 5.10 cho thấy hệ tọa độ OXY trong mặt phẳng 2D và hệ tọa độ camera O'X'Y'. Để biết vị trí của xe lăn trong mặt phẳng 2D, thông tin từ hệ tọa độ O'X'Y' của camera cần được chuyển đổi sang hệ tọa độ OXY. Hơn nữa, (X_M, Y_M) là tọa độ của vật mốc trong mặt phẳng 2D và được xác định dựa trên thông tin vật mốc. Ngoài ra, mỗi vật mốc trong môi trường đều mang thông tin về hướng của nó trên bản đồ lưới, cụ thể gồm có 4 hướng được đặt tên là (Up, Down, Left, Right). Trong nghiên cứu này, vị trí của xe lăn trong mặt phẳng 2D theo hệ tọa độ OXY được tính toán theo 4 trường hợp được mô tả trên hình 5.10a – 5.10d như sau:

- Hướng của vật mốc là “Up” trong bản đồ lưới 2D:

$$\begin{aligned} X_W &= X_M - x_a \\ Y_W &= Y_M - z_a \end{aligned} \quad (5.9)$$

- Hướng của vật mốc là “Down” trong bản đồ lưới 2D:

$$\begin{aligned} X_W &= X_M + x_a \\ Y_W &= Y_M + z_a \end{aligned} \quad (5.10)$$

- Hướng của vật mốc là “Right” trong bản đồ lưới 2D:

$$\begin{aligned} X_W &= X_M - z_a \\ Y_W &= Y_M + x_a \end{aligned} \quad (5.11)$$

- Hướng của vật mốc là “Left” trong bản đồ lưới 2D:

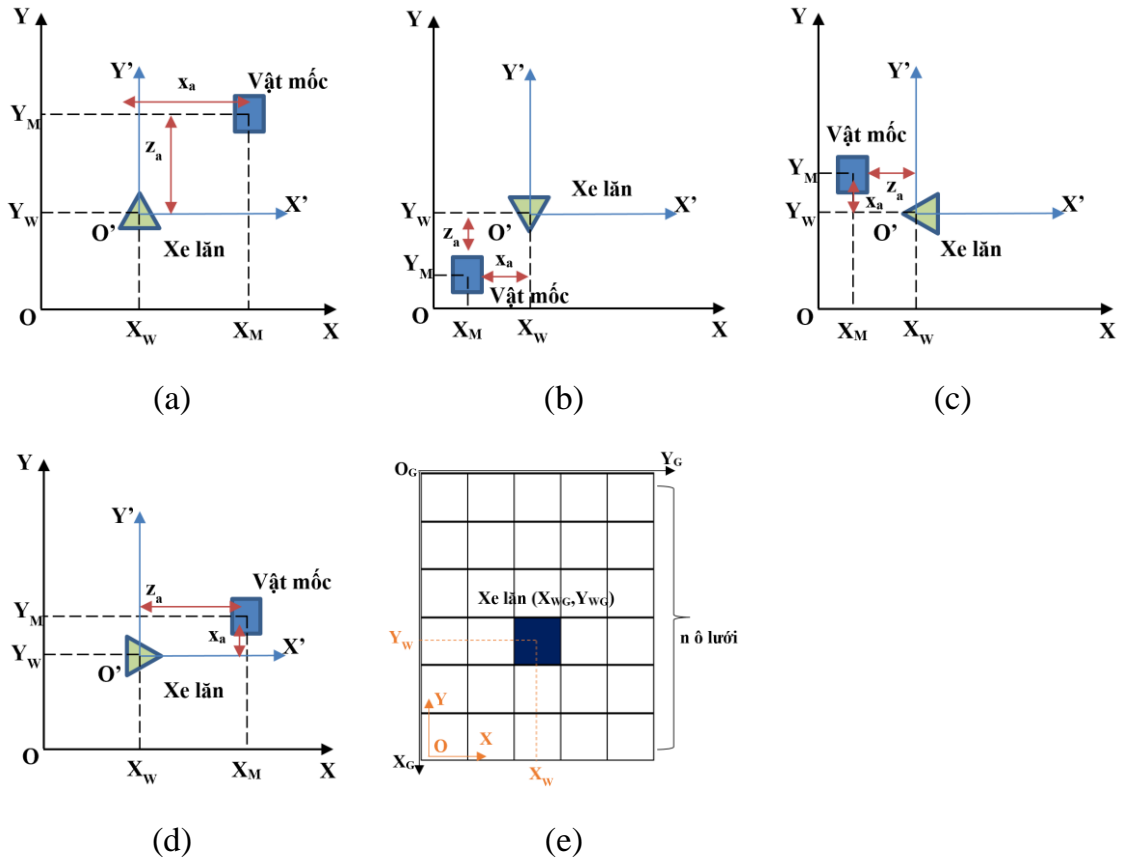
$$\begin{aligned} X_W &= X_M + z_a \\ Y_W &= Y_M - x_a \end{aligned} \quad (5.12)$$

với (X_W, Y_W) là tọa độ của xe lăn trong mặt phẳng OXY; (X_M, Y_M) là tọa độ của vật mốc trong mặt phẳng OXY; x_a là độ lệch từ tâm camera đến vật mốc theo phương ngang; z_a là khoảng cách từ tâm camera đến mặt phẳng chứa vật mốc theo phương dọc.

Sau khi có được vị trí của xe lăn trong mặt phẳng OXY, vị trí của xe lăn trong bản đồ lưới 2D ảo sẽ được xác định. Giả sử rằng bản đồ lưới 2D ảo như trong hình 5.8e có kích thước mỗi ô là $(a \times a)$ và gốc tọa độ $(0,0)$ nằm ở góc trên cùng bên trái, vị trí của xe lăn (X_G, Y_G) trong bản đồ lưới 2D ảo được tính toán bằng các công thức:

$$\begin{aligned} X_G &= n - \text{round}\left(\frac{Y_W}{a}\right) \\ Y_G &= \text{round}\left(\frac{X_W}{a}\right) - 1 \end{aligned} \quad (5.13)$$

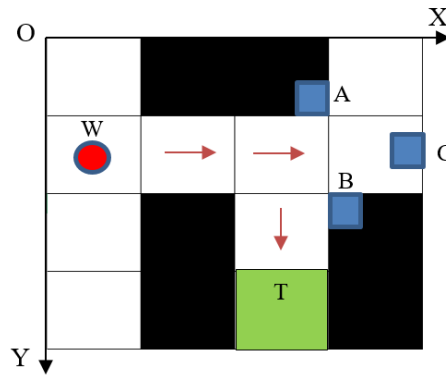
trong đó n là tổng số ô dọc theo trục tung của bản đồ lưới 2D.



Hình 5.10. Vị trí của xe lăn với bốn hướng: (a) Hướng “Up” của vật mốc, (b) Hướng “Down” của vật mốc, (c) Hướng “Left” của vật mốc, (d) Hướng “Right” của vật mốc, (e) Bản đồ lưới 2D trong mặt phẳng OXY.

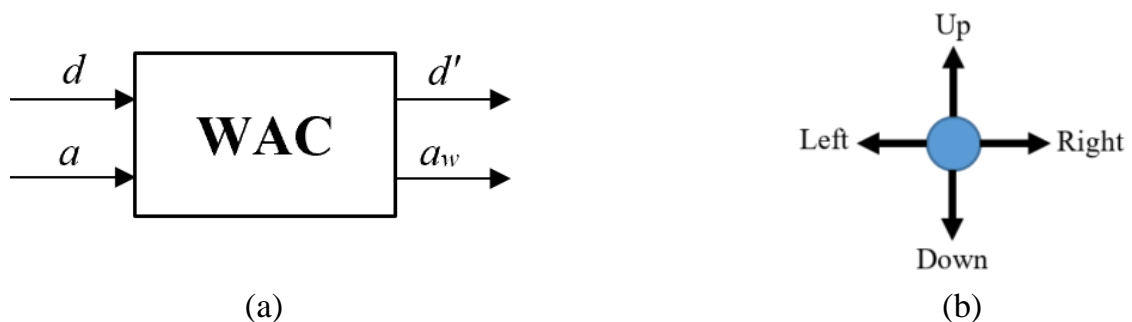
5.3. ĐIỀU HƯỚNG XE LĂN ĐIỆN TRONG MÔI TRƯỜNG THỰC

Ở mô hình này, xe lăn sẽ định vị trên bản đồ lưới 2D ảo thông qua các vật mốc bao gồm vị trí và hướng di chuyển của xe lăn. Việc cập nhật vị trí của xe lăn được thực hiện khi bắt đầu di chuyển lần đầu tiên. Trong phương pháp này, chỉ có vị trí xe lăn được đưa vào đầu vào của khối MP để xác định đường đi tối ưu và sau đó hiển thị các hành động cụ thể với vị trí đó của xe lăn.



Hình 5.11. Tọa độ của xe lăn, vật mốc và đích đến trong bản đồ lưới 2D ảo

Trong nghiên cứu này, khi xe lăn ở trong môi trường thực tế như mô tả trong hình 5.11, vị trí ban đầu của xe lăn được xác định dựa trên các vật mốc được chụp từ hệ thống camera được lắp đặt trên xe lăn. Với điểm bắt đầu $S_w(1,0) \in S_f$ và đích đến $T_i(3,2) \in T$ có được dựa trên bản đồ được xây dựng trước, trong đó S_f là tập hợp các ô trống và T là tập hợp của các đích đến đã biết. Do đó, MP đưa ra một đường đi tối ưu là một tập hợp các hành động bao gồm (Right, Right, Down, Down) như trong hình 5.11. Một thực tế là xe lăn không thể di chuyển dựa trên các hành động này do xe lăn điện không phải là một mô hình điều khiển đa hướng. Trong hình 5.12a, khối chuyển đổi WAC gồm hai đầu vào là hành động a được xác định từ đầu ra MP và hướng ban đầu d của xe lăn bao gồm bốn hướng (Up, Down, Left, Right) như được mô tả trong hình 5.12b. Như vậy, đầu ra của khối WAC là hành động a_w để phù hợp với hướng di chuyển của xe lăn trong môi trường thực tế.



Hình 5.12. Biểu diễn bộ chuyển đổi các lệnh điều khiển thực tế từ mô phỏng.
 (a) Bộ chuyển đổi WAC với đầu vào mô phỏng và đầu ra thực tế; (b) Biểu diễn bốn hướng điều khiển

Quá trình huấn luyện tìm đường đi tối ưu sẽ tạo ra một chuỗi hành động với các trạng thái khác nhau, trong đó các hành động này sẽ tạo ra đường đi tối ưu phụ thuộc vào vị trí ban đầu của xe lăn. Do đó, sau mỗi hành động a , hướng của xe lăn d sẽ thay đổi thành hướng mới d' . Đối với chuyển động của xe lăn, nghiên cứu đề xuất một thuật toán mới dựa trên WAC như được mô tả trong hình 5.12. Cụ thể, hành động của xe lăn a_w và hướng mới $d' = a$ trong quá trình di chuyển của nó trong môi trường thực cần được xác định và thuật toán này được thể hiện như sau:

$$\begin{cases} a_w = \begin{cases} Forward & \text{if } a = Up \\ Backward & \text{if } a = Down \\ Left - Forward & \text{if } a = Left \\ Right - Forward & \text{if } a = Right \end{cases} \\ d = Up \end{cases} \quad (5.14a)$$

$$\begin{cases} a_w = \begin{cases} Forward & \text{if } a = Down \\ Backward & \text{if } a = Up \\ Left - Forward & \text{if } a = Right \\ RightandForward & \text{if } a = Left \end{cases} \\ d = Down \end{cases} \quad (5.14b)$$

$$\begin{cases} a_w = \begin{cases} Forward & \text{if } a = Left \\ Backward & \text{if } a = Right \\ Left - Forward & \text{if } a = Down \\ Right - Forward & \text{if } a = Up \end{cases} \\ d = Left \end{cases} \quad (5.14c)$$

$$\begin{cases} a_w = \begin{cases} Forward & \text{if } a = Right \\ Backward & \text{if } a = Left \\ Left - Forward & \text{if } a = Up \\ Right - Forward & \text{if } a = Down \end{cases} \\ d = Right \end{cases} \quad (5.14d)$$

trong đó a và d là các tham số được xác định dựa trên hành động và hướng trong MP.

Trong phương trình (5.14a) – (5.14d), hành động của xe lăn a_w được định nghĩa:

$a_w = Forward$: Xe lăn sẽ đi thẳng

$a_w = Backward$: Xe lăn sẽ đi lùi

$a_w = Left-Forward$: Xe lăn sẽ xoay trái rồi đi thẳng

$a_w = Right-Forward$: Xe lăn sẽ xoay phải rồi đi thẳng

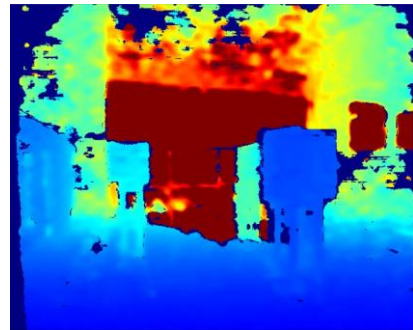
$a_w = Stop$: Xe lăn sẽ dừng lại

5.4. PHƯƠNG PHÁP TRÁNH VẬT CẢN DỰA VÀO THÔNG TIN 3D MÔI TRƯỜNG

Một bản đồ khoảng cách 2D cung cấp các giá trị khoảng cách của chướng ngại vật và khoảng trống là cần thiết để đưa ra các phương án điều khiển xe lăn cho những người khuyết tật [80]. Hình 5.13 mô tả ảnh độ sâu của môi trường. Trong hình 5.13b có thể thấy các vùng màu xanh dương là thể hiện cho các điểm ảnh ở gần camera và vùng màu đỏ thể hiện cho các điểm ảnh ở xa camera nhất. Từ thông tin độ sâu này, có thể tính toán các vùng vật cản và vùng trống, giúp xe lăn có thể di chuyển an toàn trong môi trường.



(a)



(b)

Hình 5.13. Ảnh độ sâu được chụp từ camera RGB-D: (a) Môi trường có các chướng ngại vật, (b) Ảnh độ sâu của môi trường.

Bản đồ khoảng cách 2D được chuyển đổi từ bản đồ điểm 3D được cung cấp bởi camera RGB-D sử dụng các phép chiếu hình học. Đầu tiên, bản đồ điểm 3D sẽ được loại bỏ các điểm có độ cao lớn hơn độ cao an toàn của xe lăn (chỉ xem xét các điểm ở phía trước xe lăn) và chia thành các cột theo trục Y . Tiếp đến các điểm có độ sâu nhỏ nhất theo trục Z trên cột thứ i (Z_{imin}) được lựa chọn cùng với giá trị độ rộng của điểm đó theo trục X . Nếu cột nào không có điểm ảnh trong vùng quan sát của camera thì vị trí cột đó xem như vị trí khoảng trống. Việc chuyển đổi bản đồ 2D, được tính toán như sau:

$$Z_{imin} = \min(Z_{ij}) \quad (j = \overline{0, n}) \quad (5.15)$$

trong đó các giá trị Z_{imin} được lựa chọn tương ứng với giá trị Y_{jmin} phụ thuộc vào chiều cao của xe lăn, hoặc chiều cao của người sử dụng xe lăn. Bản đồ khoảng cách 2D cuối cùng (X_i, Z_{imin}) sẽ phụ thuộc vào giá trị Y_{jmin} .

Để tìm ra điểm ảnh có độ sâu nhỏ nhất Z_{min} (gần camera nhất), tiến hành so sánh các Z_{imin} ở các cột theo phương trình sau:

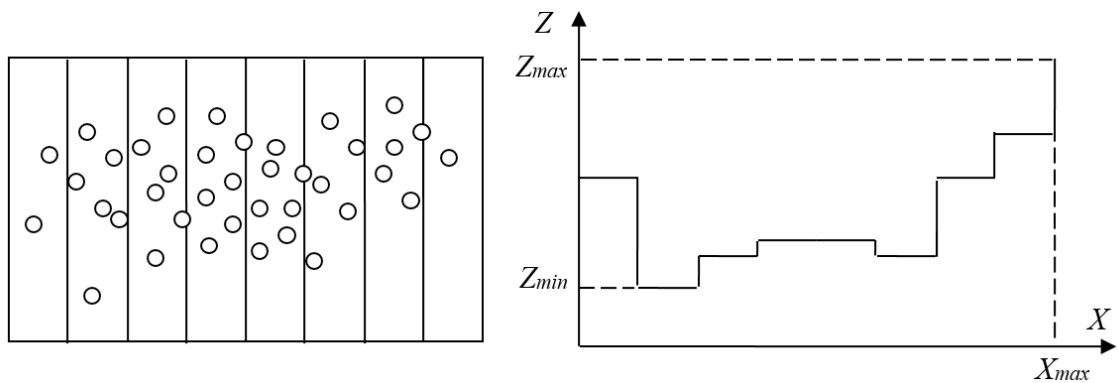
$$Z_{min} = \min(Z_{imin}) \quad (j = \overline{0, m}) \quad (5.16)$$

Độ rộng a_v của khoảng trống ($v = 1, 2, \dots$) trong bản đồ khoảng cách 2D (X_i, Z_{imin}) và được biểu diễn như sau:

$$a_v = |X_{k1} - X_{k2}| \quad (5.17)$$

với giá trị k_1 và k_2 là các phần tử đầu tiên và cuối cùng trên trục X của khoảng trống thứ v tại cột có độ sâu $Z \geq Z_{min}$.

Bản đồ 2D sau khi được xây dựng sẽ thể hiện các vùng vật cản và các khoảng trống phía trước camera như hình 5.14. Độ sâu Z_{min} của vùng vật cản sẽ được sử dụng để thiết lập khoảng cách an toàn cho xe lăn, đồng thời độ rộng các khoảng trống sẽ được dùng để kiểm tra với đường kính của xe lăn và quyết định lựa chọn phương án di chuyển qua các khoảng trống.



Hình 5.14. Mô tả chuyển đổi bản đồ điểm 3D sang 2D với độ cao được định trước.

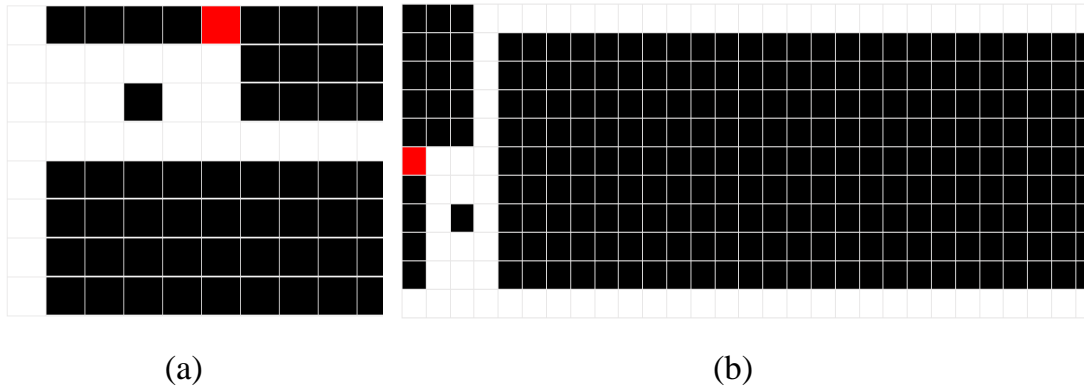
5.5. KẾT QUẢ THÍ NGHIỆM ĐIỀU KHIỂN XE LĂN ĐIỆN

5.5.1. Mô phỏng huấn luyện tìm đường đi cho xe lăn dựa vào bản đồ lưới 2D ảo

Hai bản đồ lưới 2D ảo mô tả môi trường trong nhà như trong hình 5.15 được xây dựng, trong đó các ô màu trắng là vùng trống, các ô màu đen là các chướng ngại vật và các ô màu đỏ là các điểm đích. Máy tính với hệ điều hành Window, vi xử lý

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

Intel(R) Core (TM) i5- 6300U, 2.4 GHz, RAM 16 GB được sử dụng trong quá trình huấn luyện và kiểm tra mô hình đề xuất. Trong mỗi lần huấn luyện, vị trí xuất phát được chọn ngẫu nhiên trên bản đồ và đảm bảo không trùng với ô chướng ngại vật. Bảng 5.1 mô tả các tham số mô hình được sử dụng khi huấn luyện cho các trường hợp như được mô tả trong hình 5.15.



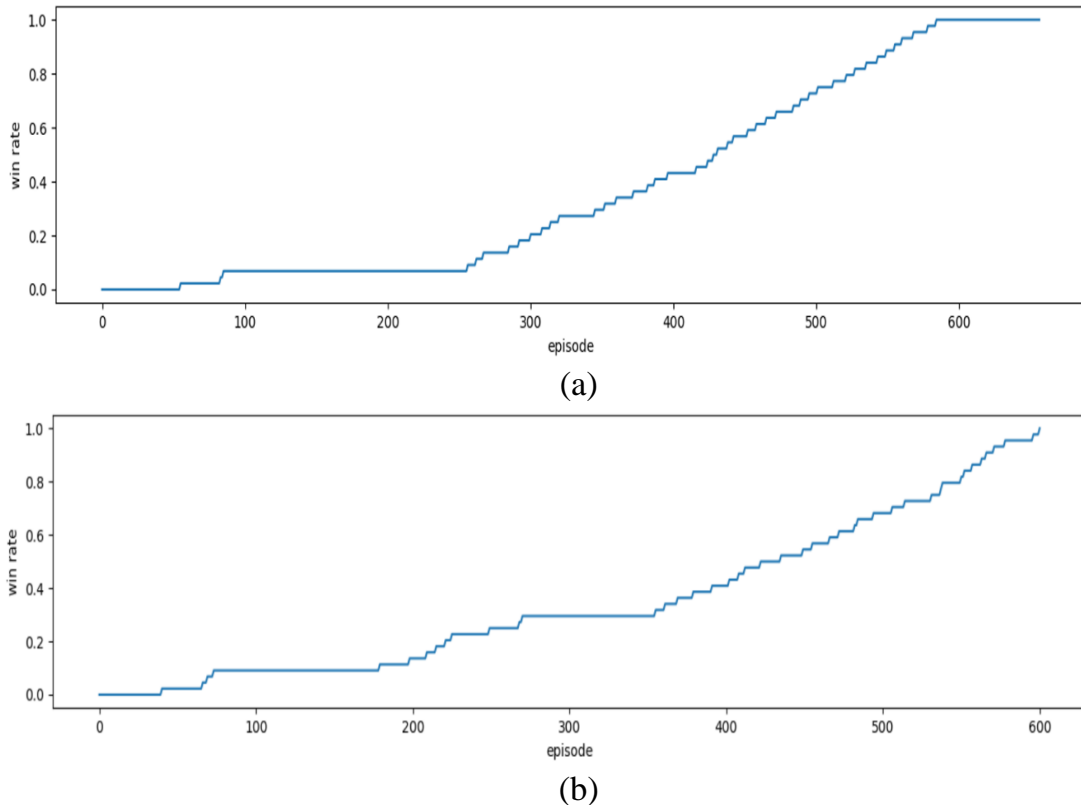
Hình 5.15. Môi trường huấn luyện. (a) Bản đồ lưới 8×11 ; (b) Bản đồ lưới 11×33

Bảng 5.1. Thông số mô hình huấn luyện.

Thông số	Giá trị
Tốc độ học	0,00001
Hệ số chiết khấu γ	0,8
Hệ số khám phá	0,1
Kích thước Mini-batch	32
Kích thước bộ nhớ phát lại	100
Điểm thưởng khi di chuyển ra ngoài bản đồ R_b	-0,8
Điểm thưởng khi di chuyển vào ô trống R_f	-0,4
Điểm thưởng khi di chuyển vào ô vật cản R_o	-0,75
Điểm thưởng khi di chuyển vào ô đích R_g	1

Để đánh giá hiệu quả của mô hình DQNs, các môi trường khác nhau được thực hiện. Kết quả sau khi huấn luyện của mô hình DQNs được thể hiện trong hình 5.16 và hình 5.17 cho từng môi trường, trong đó trục hoành là số lượng episode và trục tung là tỷ lệ Win. Tỷ lệ Win được tính dựa trên số lần hoàn thành di chuyển đến đích

thành công trên tổng số vị trí đã chọn để bắt đầu di chuyển trong một episode. Từ hình 5.16, có thể thấy tỷ lệ Win có thể tăng, giảm hoặc giữ nguyên sau mỗi episode. Ngoài ra, việc sử dụng mô hình DQNs với 2 loại hàm kích hoạt PreLU và ReLU cũng được thực hiện để so sánh hiệu suất giữa 2 mô hình này.

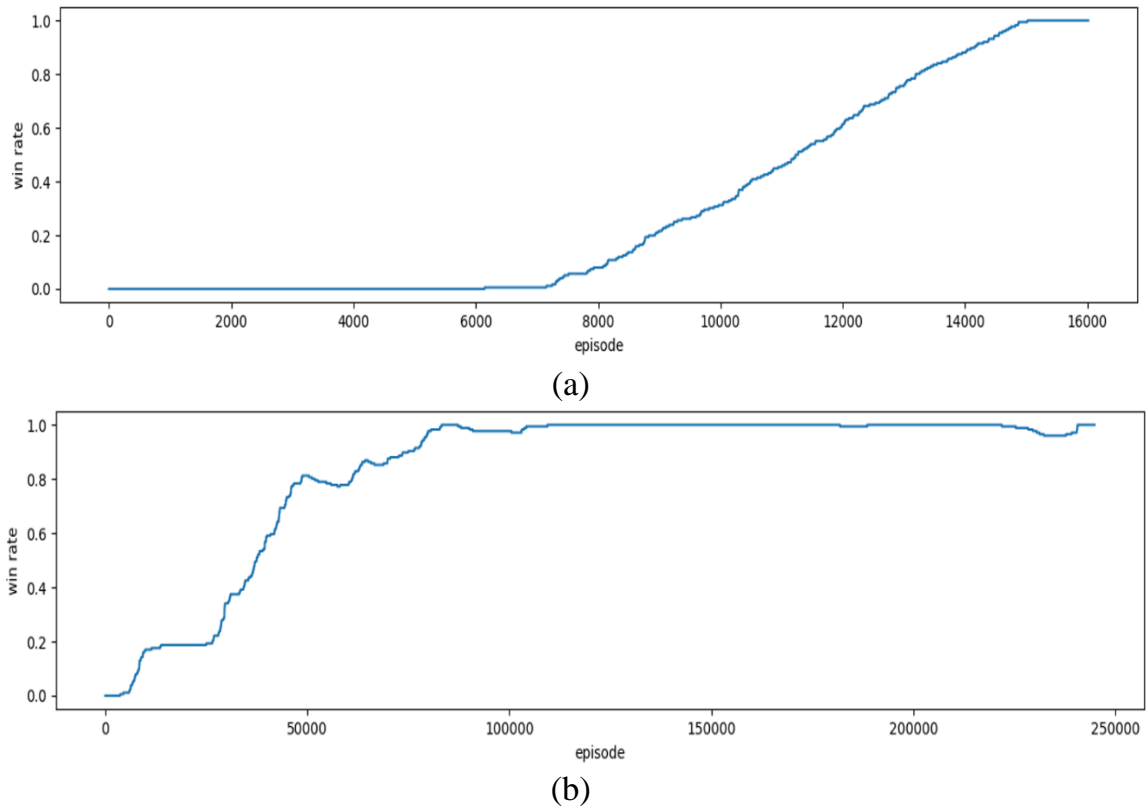


Hình 5.16. So sánh tỷ lệ Win khi huấn luyện mô hình DQNs với 2 loại hàm kích hoạt trong trường hợp bản đồ lưới 8×11 . (a) Mô hình DQNs với hàm kích hoạt PreLU; (b) Mô hình DQNs với hàm kích hoạt ReLU

Theo kết quả ở hình 5.16 với môi trường nhỏ, hai mô hình DQNs – PreLU và DQNs – ReLU có lộ trình tăng trưởng tỷ lệ Win giống nhau và cùng đạt ngưỡng tối đa là 1,0 sau khoảng 600 episode. Hình 5.17 cho thấy mức tăng trưởng của tỷ lệ Win của môi trường lớn hơn với 2 mô hình được chọn. Với kết quả của mô hình DQNs – PreLU trong hình 5.17a, khi số episode trên 7000, tỷ lệ Thắng bắt đầu tăng mạnh và sau đó đạt ngưỡng cực đại ở tập 15000. Do đó, tỷ lệ Win đạt bão hòa và điều này cho thấy mô hình đã đạt yêu cầu huấn luyện và kết thúc. Ngược lại, theo kết quả như hình 5.17b, khi sử dụng mô hình DQNs – ReLU, tỷ lệ Win bắt đầu tăng mạnh khi số

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

episode trên 25000 và đạt ngưỡng cực đại khi số episode là 240000. Sau đó, tỷ lệ Win đạt bão hòa và nghĩa là mô hình đã đạt yêu cầu huấn luyện và kết thúc. Như vậy có thể thấy trong môi trường rộng lớn, mô hình DQNs – PreLU nhanh chóng đạt điểm tối đa hơn so với mô hình DQNs – ReLU.



Hình 5.17. So sánh tỷ lệ Win khi huấn luyện mô hình DQNs với hai loại hàm kích hoạt trong trường hợp bản đồ lưới 11×33 . (a) Mô hình DQNs với hàm kích hoạt PreLU; (b) Mô hình DQNs với hàm kích hoạt ReLU

Ngoài ra, kết quả thu được có sự so sánh về thời gian huấn luyện và số episode của mô hình DQNs với 2 loại hàm kích hoạt như trong bảng 5.2. Cụ thể, trong môi trường nhỏ 8×11 , sự khác biệt về thời gian huấn luyện không quá lớn, lần lượt là 36.3 giây so với 42.3 giây cho hai hàm kích hoạt ReLU và PreLU. Số episode của 2 mô hình DQNs – ReLU và DQNs – PreLU dùng cho môi trường huấn luyện này không khác nhau nhiều, lần lượt là 601 và 607. Tuy nhiên, với môi trường lớn hơn 11×33 , sẽ có sự khác biệt lớn về thời gian huấn luyện và số episode giữa hai mô hình. Đặc biệt, thời gian đào huấn của mô hình DQNs – ReLU lớn hơn gần 4 lần so với mô hình

DQNs – PreLU. Ngoài ra, số episode trung bình trên mỗi lần huấn luyện khi sử dụng mô hình DQNs – ReLU nhiều hơn 15 lần so với mô hình DQNs – PreLU. Điều này có nghĩa là mô hình DQNs – PreLU cho hiệu năng tốt hơn so với mô hình DQNs – ReLU khi sử dụng cho môi trường lớn này.

Bảng 5.2. Hiệu suất của các mô hình DQNs được đề xuất.

Môi trường	Mô hình	Số lượng episode trung bình	Thời gian huấn luyện trung bình
Môi trường nhỏ (8×11)	DQNs – ReLU	601	36,3 giây
	DQNs – PreLU	657	42,3 giây
Môi trường lớn (11×33)	DQNs – ReLU	244879	6,05 giờ
	DQNs – PreLU	16015	35,24 phút

Bảng 5.3. So sánh hiệu suất của các mô hình.

Môi trường	Mô hình	Số lượng episode trung bình	Thời gian huấn luyện trung bình
Môi trường nhỏ (8×11)	Q-Learning truyền thống	60	198,4 giây
	SARSA	75	223,9 giây
Môi trường lớn (11×33)	Q-Learning truyền thống	235	1,45 giờ
	SARSA	275	57,23 phút

Bảng 5.3 mô tả số lượng episode và thời gian huấn luyện của 2 mô hình khác trong huấn luyện 2 môi trường (Nhỏ và Lớn). Mô hình Q-Learning truyền thống [114] học từ các hành động được thực hiện ngẫu nhiên, tức là không có trước một chiến lược cụ thể. Từ đó, quá trình học của Q-Learning là để tìm ra được một chiến lược có thể tối đa hoá được tổng phần thưởng. Giá trị Q đại diện cho chất lượng hay độ tốt của việc thực hiện một hành động để làm tăng phần thưởng trong dài hạn. Do đó, những giá trị Q này liên tục được thay đổi trong quá trình huấn luyện. State Action Reward State Action (SARSA) là một trong những thuật toán học tăng cường, là phiên bản sửa đổi của thuật toán Q-learning [116]. Điểm chính phân biệt thuật toán

SARSA với thuật toán Q-learning là thuật toán này không tối đa hóa phần thưởng cho giai đoạn tiếp theo của hành động được thực hiện và cập nhật giá trị Q của các vị trí tương ứng. Các mô hình với các thuật toán này được huấn luyện cho từng môi trường trong 10 lần để lấy thời gian huấn luyện trung bình và số episode trung bình. Trong quá trình huấn luyện, tùy theo môi trường lớn hay nhỏ mà thời gian huấn luyện và số lượng episode sẽ khác nhau.

Cụ thể, với môi trường nhỏ thì thời gian huấn luyện và số lượng episode ít hơn so với môi trường lớn như ở bảng 5.2 và bảng 5.3. Ngoài ra, ở bảng 5.3, các mô hình có số episode ít và nhiều thời gian vì Q-Learning truyền thống hoạt động dựa trên việc tìm kiếm phần thưởng tối đa cho mỗi hành động và số vị trí trên bản đồ lưới càng nhiều thì bảng Q càng lớn nên việc tính toán sẽ mất nhiều thời gian. Trong khi ở bảng 5.2, DQNs có nhiều episode nhưng lại tốn ít thời gian tính toán hơn do DQNs chọn một số quyết định ngẫu nhiên và mạo hiểm để nhanh chóng nhận được phần thưởng cao và sẽ chấp nhận mất một số lượng episode nhất định.

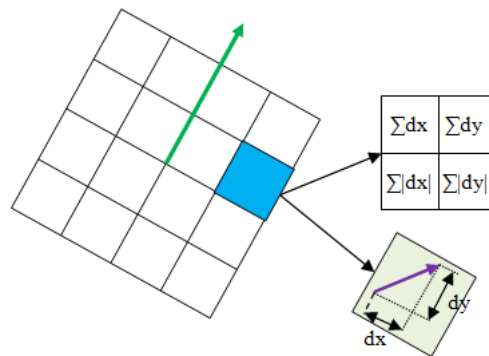
Với kết quả thống kê ở bảng 5.2 và bảng 5.3, mặc dù số episode trong quá trình huấn luyện lớn hơn nhiều so với mô hình dựa trên Q-Table ở bảng 5.3 nhưng mô hình DQNs – PreLU ở bảng 5.2 lại có thời gian huấn luyện ít hơn trong hai trường hợp huấn luyện cho cả môi trường nhỏ và lớn. Đặc biệt, đối với môi trường nhỏ, mô hình DQNs – PreLU có số episode gấp khoảng 10 lần so với mô hình Q-Learning truyền thống và SARSA nhưng thời gian đào tạo lại ít hơn gần 5 lần. Bên cạnh đó, với môi trường lớn, DQNs – PreLU có số lượng lớn khoảng 16015 episode, gấp gần 60 lần so với mô hình Q-Learning truyền thống, gần 70 lần so với mô hình SARSA. Tuy nhiên, thời gian huấn luyện giảm đáng kể khoảng 35,24 phút so với hai mô hình trong bảng 5.3 lần lượt là 1,45 giờ và 57,23 phút. Ngoài ra, với tính năng bổ sung sau khi học, mô hình DQNs có thể lưu bộ thông số đã được huấn luyện vào ổ cứng để có thể sử dụng sau này. Điều này cho thấy mô hình DQNs được huấn luyện riêng biệt cho bản đồ lưới mô phỏng môi trường thực cần được thực hiện trước và sau đó có thể ứng dụng các thông số đã được huấn luyện vào điều khiển xe lăn trong thực tế.

5.5.2. Kết quả nhận dạng vật mốc

Từ các vật mốc tự nhiên đã được nhận biết và thu thập vào cơ sở dữ liệu, phương pháp SURF được sử dụng để nhận dạng các vật mốc này trong môi trường, nhằm phục vụ cho quá trình định vị xe lăn điện. Cụ thể, phương pháp SURF bao gồm ba bước: phát hiện các điểm nổi bật, mô tả các điểm nổi bật và mô tả phù hợp để nhận dạng đối tượng. Hơn nữa, ma trận Hessian và ảnh tích phân được sử dụng để phát hiện các điểm đặc trưng vì thuật toán theo thời gian thấp và độ chính xác cao của nó. Giả sử rằng xem xét điểm $M(x, y)$ trong ảnh \mathbf{B} cần phát hiện các điểm đặc trưng, ma trận đạo hàm Hessian $\mathbf{H}(M, \sigma)$ tại điểm M có tỷ lệ σ được xác định như sau:

$$\mathbf{H}(M, \sigma) = \begin{bmatrix} L_{xx}(M, \sigma) & L_{xy}(M, \sigma) \\ L_{xy}(M, \sigma) & L_{yy}(M, \sigma) \end{bmatrix} \quad (5.18)$$

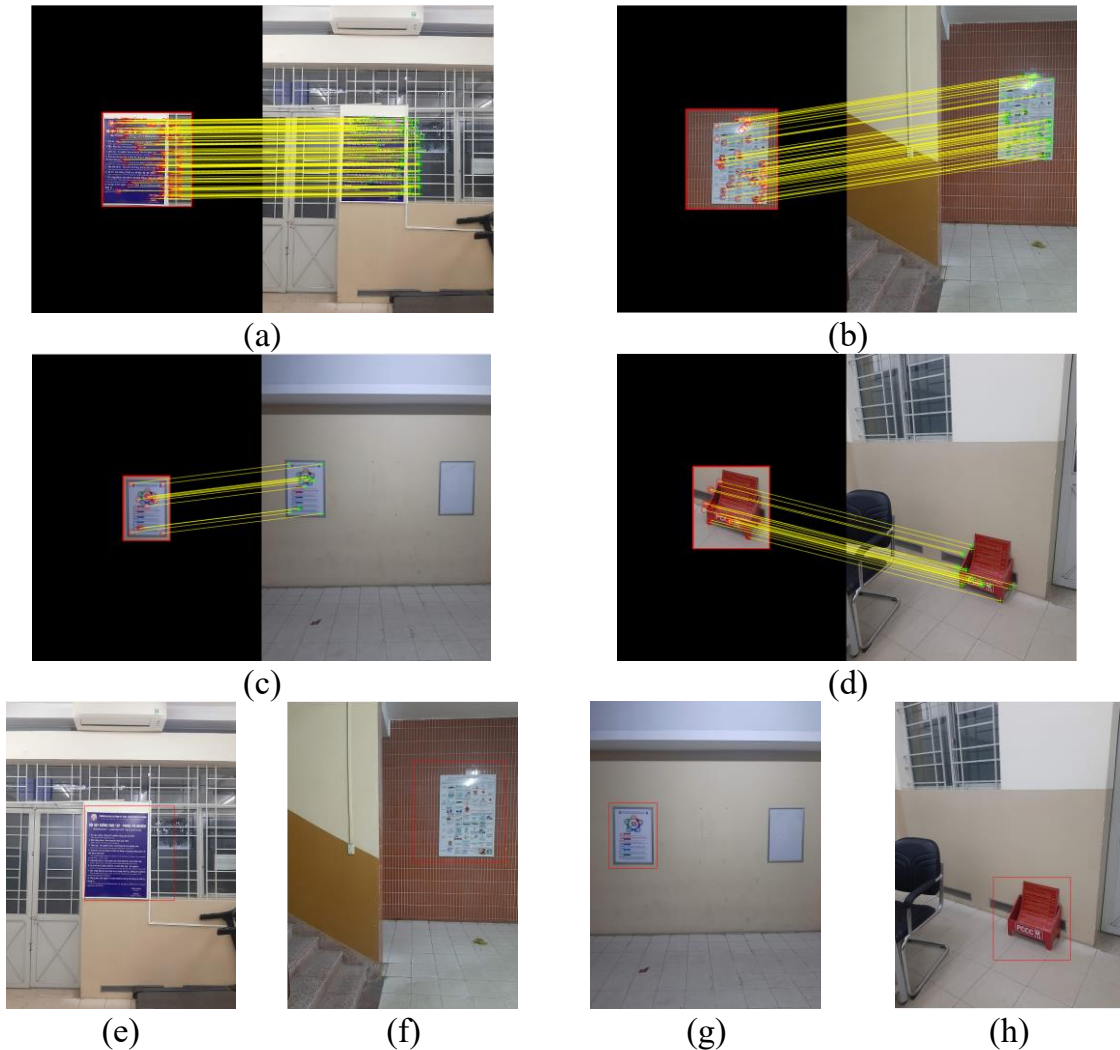
trong đó, $L_{xx}(M, \sigma)$, $L_{xy}(M, \sigma)$, $L_{yy}(M, \sigma)$ lần lượt là tích chập của đạo hàm bậc hai hàm Gausse $\frac{\partial^2}{\partial x^2} g(\sigma)$, $\frac{\partial^2}{\partial x \partial y} g(\sigma)$, $\frac{\partial^2}{\partial y^2} g(\sigma)$ với ảnh \mathbf{B} tại điểm M .



Hình 5.18. Các thành phần vector mô tả của vùng con trong ảnh.

Bộ mô tả đặc điểm thể hiện mối quan hệ về mặt không gian giữa điểm nổi bật và các điểm lân cận xung quanh nhằm xác định vị trí của các điểm nổi bật trong ảnh. Một vùng hình vuông xung quanh điểm nổi bật, có hướng trùng với hướng của vector đặc điểm đã được tìm ở trên sẽ được chia thành 4×4 vùng con. Trong mỗi vùng con sẽ tiếp tục áp dụng bộ lọc Haar Wavelet để tính toán vector mô tả. Vector mô tả của một vùng con $v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|)$ với dx là tổng đáp ứng theo trục hoành, dy tổng đáp ứng theo trục tung, $|dx|$ tổng giá trị tuyệt đối của đáp ứng theo

trục hoành và $|dy|$ tổng giá trị tuyệt đối của đáp ứng theo trục tung như hình 5.18. Trục tung có hướng trùng với hướng vectơ đặc điểm. Bước cuối cùng trong phương pháp nhận dạng SURF là so khớp các bộ mô tả giữa trong ảnh với các bộ mô tả trong tập dữ liệu để nhận dạng vật thể. Bước so khớp này dựa vào khoảng cách giữa các vectơ trong bộ mô tả, như khoảng cách Mahalanobis hoặc Oclit.



Hình 5.19. Bốn loại vật mốc khác nhau được nhận dạng dựa trên các vật mốc được lưu trữ trong cơ sở dữ liệu: (a)-(d) So khớp điểm đặc trưng của vật mốc 1, vật mốc 2, vật mốc 3 và vật mốc 4; (e)-(h) Vật mốc 1, Vật mốc 2, Vật mốc 3 và Vật mốc 4 được xác định.

Hình 5.19 thể hiện việc nhận dạng 4 vật mốc dùng phương pháp SURF, trong đó hình 5.19a – 5.19d biểu diễn sự so khớp các đặc trưng giữa vật mốc được lưu trữ trong cơ sở dữ liệu với các đặc trưng trong ảnh môi trường thực, hình 5.19e – 5.19h

là kết quả các vật mốc được tìm thấy trong ảnh thực tế (đường viền hình chữ nhật màu đỏ). Các đường màu vàng hiển thị kết nối giữa các điểm đặc trưng giống nhau khi được so khớp. Để đánh giá hiệu suất nhận dạng vật mốc, có thể dựa vào số lượng đặc trưng phù hợp và không phù hợp, độ chính xác và sai số được tính toán theo công thức sau:

$$R_t = \frac{TF}{SF} \times 100\% \quad (5.19)$$

$$R_f = \frac{SF-TF}{SF} \times 100\% \quad (5.20)$$

trong đó SF là số điểm đặc trưng của vật mốc đã xác định, TF là số điểm đặc trưng trùng khớp với số điểm đặc trưng của vật mốc đã xác định, R_t là tỷ lệ chính xác, R_f là sai số của quá trình nhận dạng vật mốc.

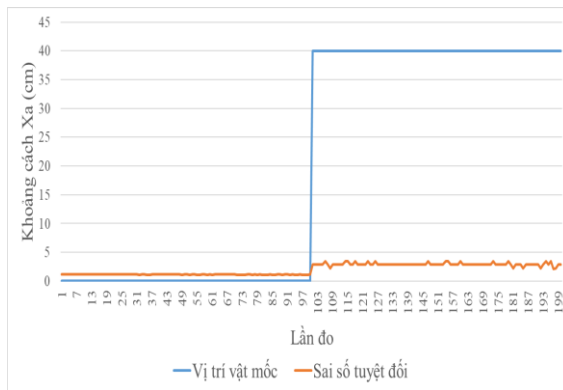
Bảng 5.4. Độ chính xác khi nhận dạng các vật mốc bằng phương pháp SURF.

Hình	Thời gian nhận dạng (ms)	SF	TF	R_t (%)	R_f (%)
Hình 5.19a	370,6	206	190	92,2	7,8
Hình 5.19b	320,4	64	58	90,6	9,4
Hình 5.19c	296,9	18	13	72,2	27,8
Hình 5.19d	228,0	30	26	86,7	13,3

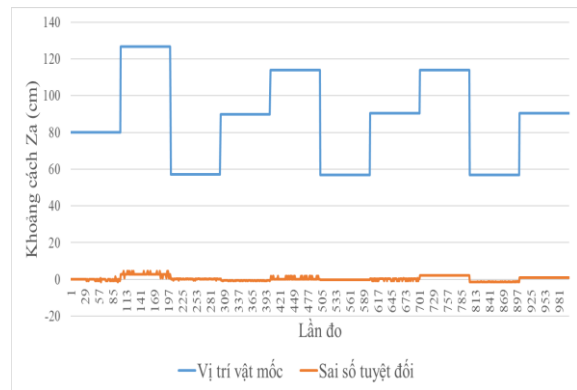
Bảng 5.4 cho thấy các vật mốc có nhiều đặc trưng (giá trị SF lớn) cho tỷ lệ chính xác trên 90%. Trong khi đó, các vật mốc ít đặc trưng (giá trị SF nhỏ) có sai số nhận dạng lớn ($R_f > 10\%$). Điều này cho thấy độ chính xác của nhận dạng vật mốc có mật độ đặc trưng lớn nhất tốt hơn so với nhận dạng vật mốc có ít điểm đặc trưng. Ngoài ra, bảng 5.4 cho thấy thời gian xử lý trung bình là 304 ms và điều này chứng tỏ xe lăn điện và hệ thống camera có thể phát hiện các vật mốc trong quá trình di chuyển với độ chính xác cao và thời gian phát hiện ngắn. Khi đã xác định được vật mốc, thông tin vị trí của vật mốc cũng được cung cấp từ cơ sở dữ liệu thu thập được, phục vụ cho quá trình định vị xe lăn.

5.5.3. Xác định vị trí xe lăn trên bản đồ lưới 2D ảo dựa vào vật mốc

Hình 5.20 mô tả độ chính xác của phép đo khoảng cách từ camera đến vật mốc. Trong thí nghiệm này, các vật mốc khác nhau được sử dụng để nhận dạng và đo khoảng cách tại các vị trí khác nhau nhằm chứng minh tính chính xác của phương pháp đề xuất. Đầu tiên, các vật mốc được đặt theo phương ngang ở cùng vị trí với camera, sau đó vị trí của các vật mốc được thay đổi theo phương dọc. Trong hình 5.20, đường màu xanh biểu thị vị trí của các vật mốc trong môi trường thực và đường màu đỏ biểu thị sai số tuyệt đối từ phép đo. Rõ ràng là phép đo khoảng cách theo phương ngang có sai số nhỏ trong khoảng $\pm 1,2$ cm. Tiếp theo, các vật mốc được đặt cách camera 40 cm theo phương ngang rồi thay đổi vị trí của nó theo phương dọc. Với thí nghiệm này, độ lệch của khoảng cách đo được tăng lên $\pm 2,8$ cm. Tuy nhiên, độ lệch này không quá lớn và có thể chấp nhận được.



Hình 5.20. Sai số tuyệt đối của phép đo khoảng cách X_a từ camera đến các vật mốc tại các vị trí khác nhau



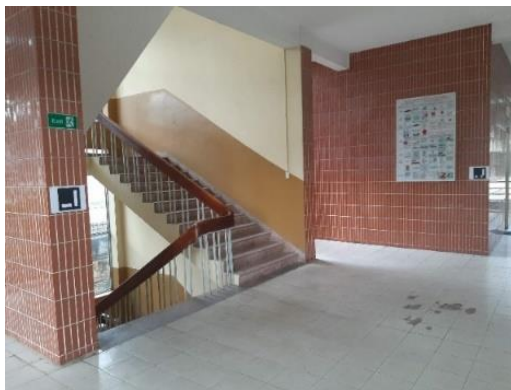
Hình 5.21. Sai số tuyệt đối của khoảng cách Z_a từ camera đến các vật mốc tại các vị trí khác nhau

Hình 5.21 cho thấy sai số tuyệt đối của phép đo khoảng cách từ camera đến vật mốc với các khoảng cách khác nhau theo phương dọc. Trong hình 5.21, đường màu xanh biểu thị vị trí của các mốc trong môi trường thực và đường màu đỏ biểu thị sai số tuyệt đối từ phép đo. Từ kết quả này, có thể thấy việc đo khoảng cách sẽ chính xác hơn khi vị trí giữa camera và mốc nhỏ hơn 1 m, nhưng vẫn cho sai số chấp nhận

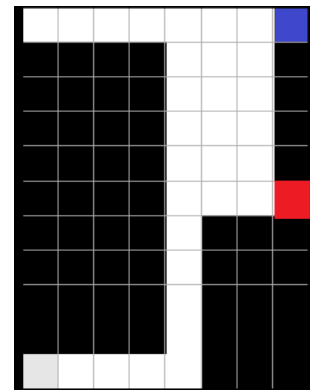
Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

được trong khoảng ± 2 cm với khoảng cách lớn hơn 1 m. Điều đó có nghĩa là xe lăn không thể hoạt động trong một khu vực quá rộng.

Một thử nghiệm khác đã được thực hiện tập trung vào việc định vị xe lăn trên bản đồ lưới 2D. Trong đó, thí nghiệm được thực hiện trong môi trường có diện tích $126,72 \text{ m}^2$, được chia thành bản đồ lưới 2D gồm (8×11) ô lưới, trong đó mỗi ô lưới có kích thước $1,2 \text{ m} \times 1,2 \text{ m}$ như trong hình 5.22. Điểm gốc $O(0,0)$ của bản đồ lưới có màu xanh ở góc trên cùng bên trái và đích là một ô màu đỏ có tọa độ $C(0,5)$. Hình 5.22b là bản đồ lưới 2D thu được từ môi trường thực tế như được mô tả trong hình 5.22a và các thử nghiệm đã được thực hiện cho các vị trí khác nhau của xe lăn để kiểm tra độ chính xác của vị trí trên bản đồ lưới 2D dựa trên các vật mốc.



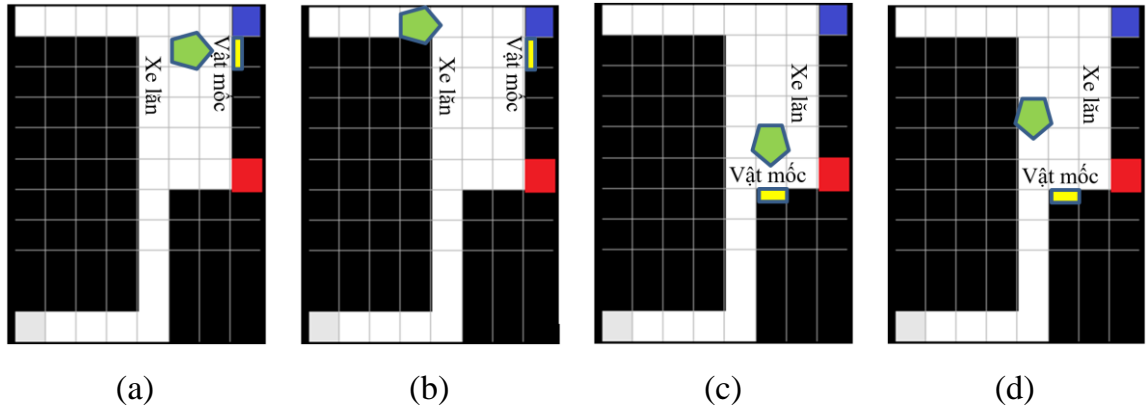
(a)



(b)

Hình 5.22. Môi trường thí nghiệm trong nhà: (a) Môi trường thực, (b) Bản đồ lưới.

Hình 5.23 biểu diễn các vị trí của xe lăn được gán biểu tượng màu xanh lá cây và các vật mốc (biểu tượng màu vàng) được áp dụng cho 4 thí nghiệm khác nhau, bao gồm vật mốc bên phải bản đồ lưới như hình 5.23a – 5.23b và vật mốc ở phía dưới của bản đồ lưới như hình 5.23c – 5.23d. Bảng 5.5 thể hiện kết quả định vị xe lăn trên bản đồ lưới 2D với độ chính xác của các thí nghiệm. Cụ thể, sau khi xác định được các vật mốc, xe lăn xác định được vị trí của nó trong mặt phẳng OXY là (X_w, Y_w) , và sau đó xác định được vị trí của xe lăn trên lưới bản đồ (X_{wg}, Y_{wg}) . Ngoài việc xác định và tính toán vị trí của xe lăn trên bản đồ lưới, hướng của xe lăn trên bản đồ lưới cũng được xác định dựa trên thông tin về hướng của các vật mốc trên bản đồ lưới. Kết quả từ bảng 5.5 cho thấy sai số tuyệt đối của xe lăn trong môi trường thực có thể chấp nhận được và xe lăn có thể xác định chính xác vị trí của mình trên bản đồ lưới.

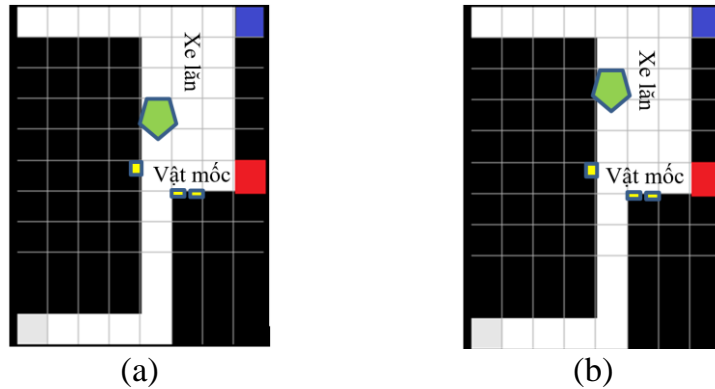


Hình 5.23. Bốn vị trí của xe lăn trên bản đồ lưới 2D với các vật mốc: (a) Vị trí thứ 1, (b) Vị trí thứ 2, (c) Vị trí thứ 3, d) Vị trí thứ 4

Bảng 5.5. Độ chính xác của việc định vị xe lăn.

TT	Vị trí thực (X_0, Y_0)	Vị trí tính toán (X_w, Y_w)	$ X_w - X_0 $ (cm)	$ Y_w - Y_0 $ (cm)	Vị trí trên bản đồ lưới (X_{WG}, Y_{WG})	Hướng trên bản đồ lưới
1	(180, 720)	(182, 718)	2	2	(2,1)	Up
2	(60, 480)	(58, 477)	2	3	(4,0)	Up
3	(600, 660)	(596, 665)	4	5	(2,4)	Right
4	(480, 540)	(474, 538)	6	2	(3,3)	Right

Hình 5.24 thể hiện các vị trí xe lăn và các vật mốc trong hai thí nghiệm khác nhau: xe lăn ở vị trí (3,3) trong sơ đồ lưới như trong hình 5.24a và vị trí (3,2) như trong hình 5.24b. Bảng 5.6 thể hiện kết quả định vị xe lăn trên bản đồ lưới 2D theo phương pháp định vị dựa vào 3 vật mốc [148]. Cụ thể, sau khi xác định được 3 vật mốc và đo khoảng cách từ xe lăn đến các vật mốc, vị trí xe lăn được tính toán trong mặt phẳng OXY là (X_w, Y_w), sau đó xác định được vị trí xe lăn trên bản đồ lưới (X_{WG}, Y_{WG}). Kết quả bảng 5.6 cho thấy sai số tuyệt đối của vị trí xe lăn trong môi trường thực tế khi sử dụng 3 vật mốc lớn hơn so với khi sử dụng 1 vật mốc. Hơn nữa, trong thí nghiệm này, nếu cả 3 vật mốc cùng nằm trên một mặt phẳng hoặc hệ thống không tìm được cả 3 vật mốc thì không xác định được vị trí của xe lăn. Vì vậy, đây là một hạn chế của việc sử dụng 3 vật mốc so với 1 mốc.



Hình 5.24. Vị trí xe lăn trên bản đồ lưới 2D với 3 vật mốc: (a) Xe lăn tại vị trí (3,3), (b) Xe lăn tại vị trí (3,2)

Bảng 5.6. Độ chính xác khi định vị xe lăn bằng 3 vật mốc.

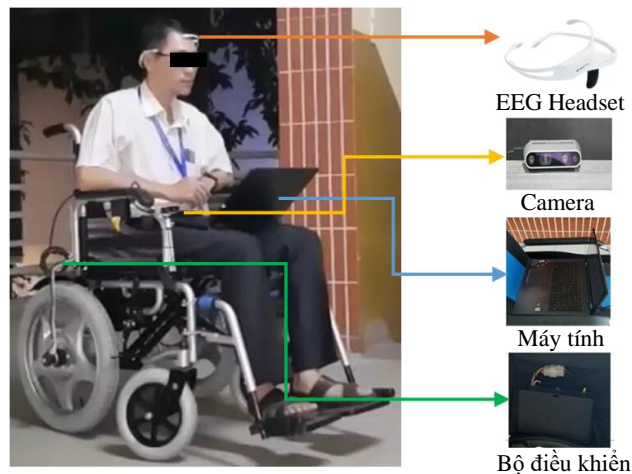
TT	Vị trí thực (X_0, Y_0)	Vị trí tính toán (X_w, Y_w)	$ X_w - X_0 $ (cm)	$ Y_w - Y_0 $ (cm)
1	(480, 540)	(489, 564)	8	14
2	(360, 540)	(348, 534)	12	6

5.5.4. Xe lăn di chuyển đến đích mong muốn dựa vào bản đồ lưới 2D ảo

Xe lăn điện được lắp đặt hệ thống camera RGB-D và các thiết bị khác như trong hình 5.25. Thí nghiệm được thực hiện trong môi trường có diện tích 126,72 m² được chia thành các ô lưới để tạo thành một bản đồ lưới có kích thước 8x11 ô lưới, trong đó mỗi ô lưới có kích thước 1,2 m×1,2 m như trong hình 5.26. Xe lăn được thiết lập để có thể di chuyển với tốc độ tối đa 3 km/h cho phù hợp với tốc độ xử lý của hệ thống. Thông tin về môi trường xung quanh thu được từ hệ thống camera sẽ được máy tính xử lý sau đó chuyển đến bộ điều khiển động cơ của xe lăn để điều khiển chuyển động. Xe lăn điện sử dụng hai động cơ một chiều kích từ độc lập với công suất 240W. Tốc độ động cơ được điều khiển bằng phương pháp điều chế rộng xung PWM, với tần số 1KHz. Khi độ rộng xung thay đổi thì điện áp cấp cho động cơ sẽ thay đổi, làm cho tốc độ động cơ thay đổi theo. Ngoài ra, hướng di chuyển của xe lăn được điều khiển bằng cách thay chiều quay và đổi tốc độ ở hai bánh xe gắn vào hai trục động cơ, như mô tả ở bảng 5.7.

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

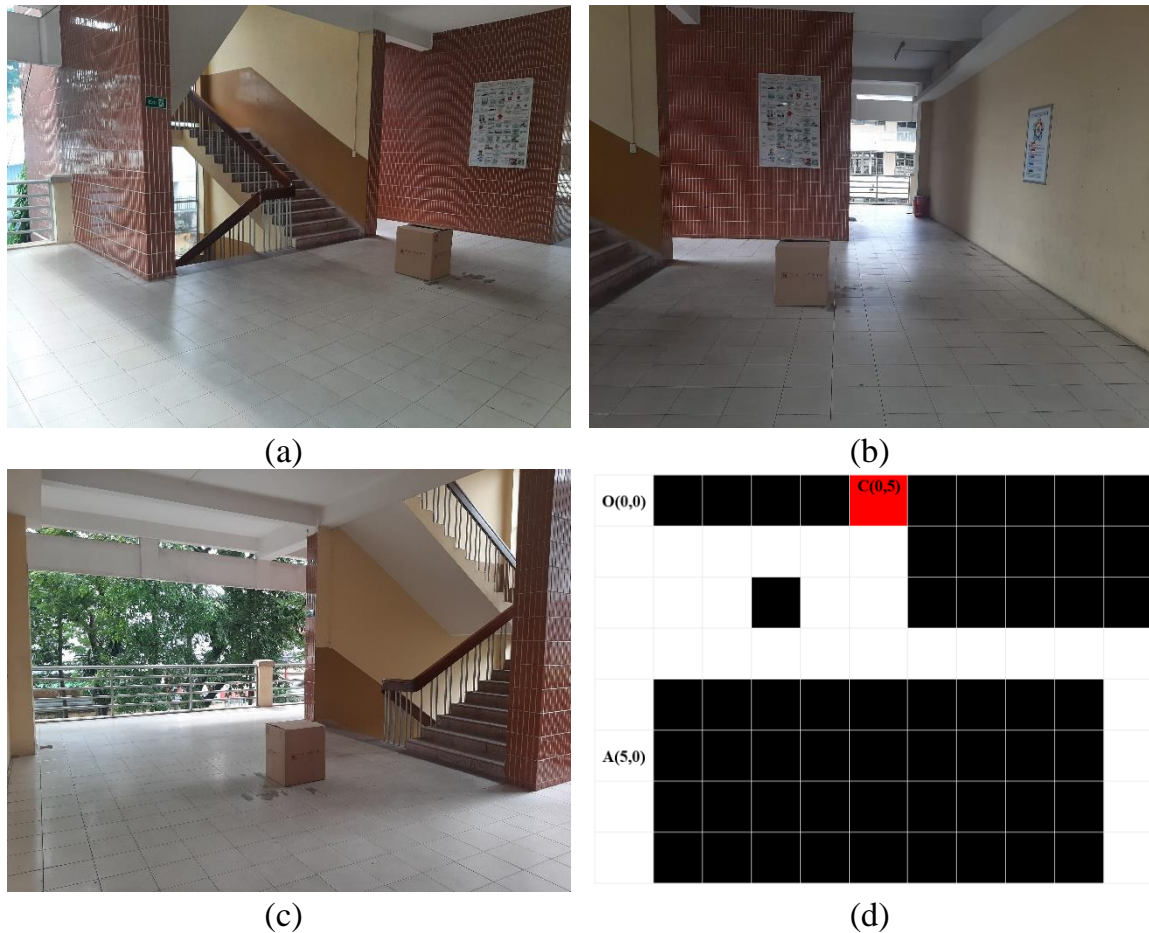
Ngoài ra, nghiên cứu này đã thực hiện 2 thí nghiệm, bao gồm người dùng tự điều khiển và người dùng điều khiển bán tự động. Nghĩa là ở mô hình người dùng tự điều khiển, người dùng có thể ra các lệnh điều khiển như tiến, lùi, rẽ phải, trái trong quá trình xe lăn di chuyển bằng BCI. Trong khi đó, với chế độ người dùng điều khiển bán tự động, người dùng có thể chọn một trong các đích đến bằng cách sử dụng tín hiệu EEG [36] và hệ thống điều khiển xe lăn sẽ tính toán và ra các lệnh để xe lăn có thể tự động tiếp cận điểm đích đã chọn.



Hình 5.25. Hệ thống xe lăn với các camera RGB-D, Emotiv Epoc+, bộ điều khiển và máy tính

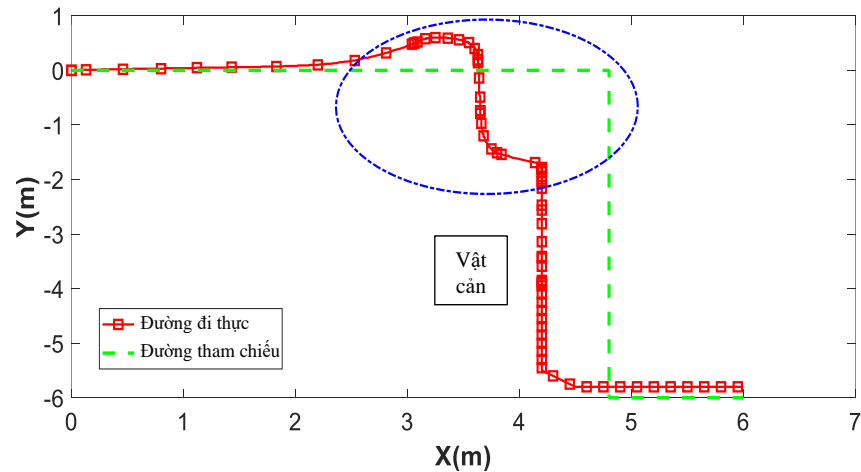
Bảng 5.7. Các giá trị và trạng thái điều khiển xe lăn điện

Hướng di chuyển	Chiều quay động cơ		Giá trị PWM (%)	
	Động cơ trái	Động cơ phải	Động cơ trái	Động cơ phải
Đi thẳng	Quay thuận	Quay thuận	12	12
Rẽ trái	Quay nghịch	Quay thuận	9	9
Rẽ phải	Quay thuận	Quay nghịch	9	9
Dừng	Dừng	Dừng	0	0
Đi lùi	Quay nghịch	Quay nghịch	12	12



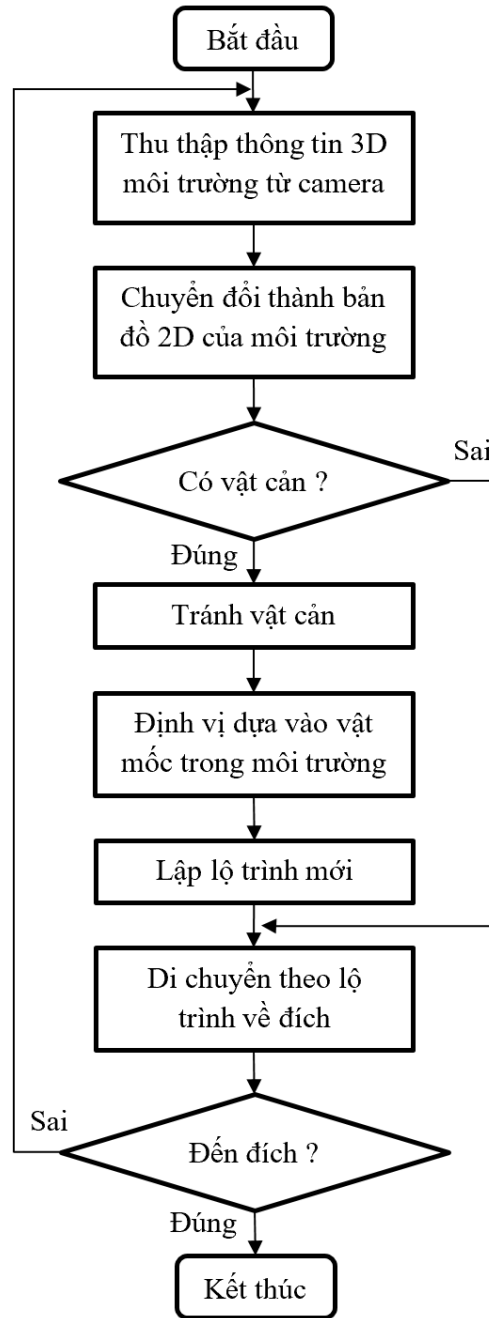
Hình 5.26. Môi trường thí nghiệm. (a)-(c) Môi trường thực; (d) Bản đồ lưới 2D

Hình 5.27 cho thấy đường màu xanh lục không liên tục là đường đi mong muốn trong môi trường thực tế mà xe lăn cần đi theo để đến đích, trong khi đường màu đỏ là đường mà xe lăn được điều khiển bởi chế độ tự điều khiển sử dụng tín hiệu EEG để đi thẳng, rẽ trái, phải khi đến đích [80]. Với thí nghiệm sử dụng chế độ tự điều khiển, xe lăn di chuyển theo đường màu đỏ rồi rẽ sang hướng không mong muốn (đường nằm trong elip màu xanh lam). Có thể thấy trong trường hợp này, xe lăn rất dễ xảy ra va chạm với chướng ngại vật. Ngoài ra, với chế độ tự điều khiển, xe lăn chuyển động không ổn định và không liên tục như hình 5.27. Cụ thể, xe lăn đi thẳng rồi dừng lại, rẽ phải rồi bị ngắt quãng liên tục trong suốt thời gian di chuyển. Rõ ràng là người dùng đã cố gắng hết sức để điều khiển xe lăn rẽ phải hoặc trái và đi thẳng.



Hình 5.27. Đường di chuyển của xe lăn trong môi trường thực tế.

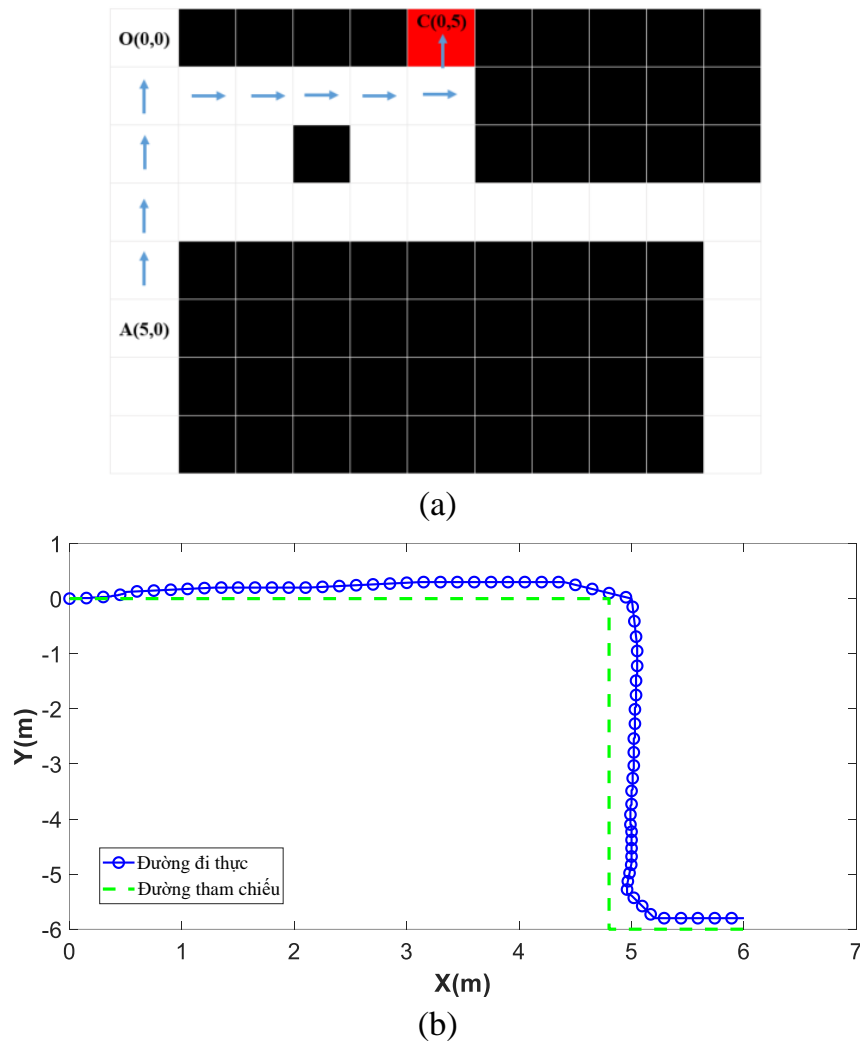
Để cải thiện khả năng điều khiển xe lăn, mô hình điều khiển đề xuất trong nghiên cứu này được áp dụng để điều khiển bán tự động cho xe lăn. Với chế độ này, người dùng chỉ cần chọn một đích đến bằng lệnh EEG, sau đó xe lăn sẽ tự động di chuyển đến đích mong muốn với độ ổn định cao và an toàn. Hình 5.28 mô tả qui trình điều khiển xe lăn điện ở chế độ bán tự động. Khi xe lăn di chuyển trong môi trường, thông tin 3D của môi trường được thu thập từ camera sẽ được chuyển đổi thành bản đồ 2D. Với bản đồ 2D này, hệ thống sẽ xác định không gian an toàn phía trước xe lăn. Với thời gian xử lý từ 3D sang 2D khoảng 220 ms và tốc độ di chuyển tối đa của xe lăn được thiết lập là 3 km/h, khoảng cách từ xe lăn đến chướng ngại vật lớn hơn 1,0 m được xem là an toàn. Nếu phía trước xe lăn là không gian an toàn, hệ thống sẽ di chuyển tự động theo đường đi tối ưu đã được bộ MP đề xuất. Trong tình huống xe lăn đi vào khu vực không an toàn, chế độ tránh vật cản được kích hoạt, trong đó xe lăn sẽ tính toán và di chuyển về phía khoảng trống lớn nhất ở phía trước dựa vào bản đồ 2D. Sau khi tránh vật cản, xe lăn sẽ thực hiện định vị vị trí hiện tại dựa vào các vật mốc xung quanh. Trên cơ sở đó, bộ MP sẽ đề xuất đường đi mới để đến đích mong muốn cho xe lăn.



Hình 5.28. Quy trình điều hướng xe lăn điện ở chế độ bán tự động

Cụ thể, với bản đồ môi trường ở hình 5.26a – 5.26c, các đường đi thực tế của xe lăn sau khi di chuyển đến đích được mô tả như hình 5.29b. Quá trình di chuyển được vẽ lại nhằm mục đích so sánh với các đường đi mô phỏng (mũi tên xanh) như hình 5.29a. Điểm xuất phát của xe lăn là ngẫu nhiên và xe lăn sẽ tự động xác định vị trí của nó trên bản đồ bằng cách nhận dạng các vật mốc trong môi trường. Cụ thể,

trong trường hợp này, xe lăn đã xác định vị trí của nó trên sơ đồ lưới tại tọa độ $A(5,0)$ và hướng của xe lăn d là Up.



Hình 5.29. Biểu diễn đường đi mô phỏng và đường đi thực của xe lăn bằng điều khiển bán tự động. (a) Đường dẫn mũi tên màu xanh được mô phỏng bằng DQNs; (b) Lộ trình di chuyển của xe lăn trong môi trường thực sử dụng DQNs

Với xe lăn bán tự động, người khuyết tật có thể điều khiển xe lăn bằng cách sử dụng tín hiệu EEG để lựa chọn một trong các lệnh trên màn hình giao diện được cài đặt một ký hiệu tương ứng với điểm đích $C(0,5)$. Với điểm bắt đầu $A(5,0)$ và điểm đích $C(0,5)$ được chọn, mô hình DQNs sẽ tạo ra một chuỗi các lệnh điều khiển hướng đi và sau đó các lệnh này được chuyển đổi thành các lệnh điều khiển trong môi trường thực cho xe lăn bằng cách sử dụng phương trình (5.14a) – (5.14d) như bảng 5.7.

Bảng 5.8. Các lệnh điều khiển xe lăn được chuyển đổi từ các lệnh mô phỏng.

Vị trí xe lăn	Hướng hiện tại	Hành động mô phỏng	Hướng mới	Hành động xe lăn
	D	A	d'	a_w
(5,0) → (4,0)	Up	Up	Up	Forward
(4,0) → (3,0)	Up	Up	Up	Forward
(3,0) → (2,0)	Up	Up	Up	Forward
(2,0) → (1,0)	Up	Up	Up	Forward
(1,0) → (1,1)	Up	Right	Right	Right-Forward
(1,1) → (1,2)	Right	Right	Right	Forward
(1,2) → (1,3)	Right	Right	Right	Forward
(1,3) → (1,4)	Right	Right	Right	Forward
(1,4) → (1,5)	Right	Right	Right	Forward
(1,5) → (0,5)	Right	Up	Up	Left-Forward

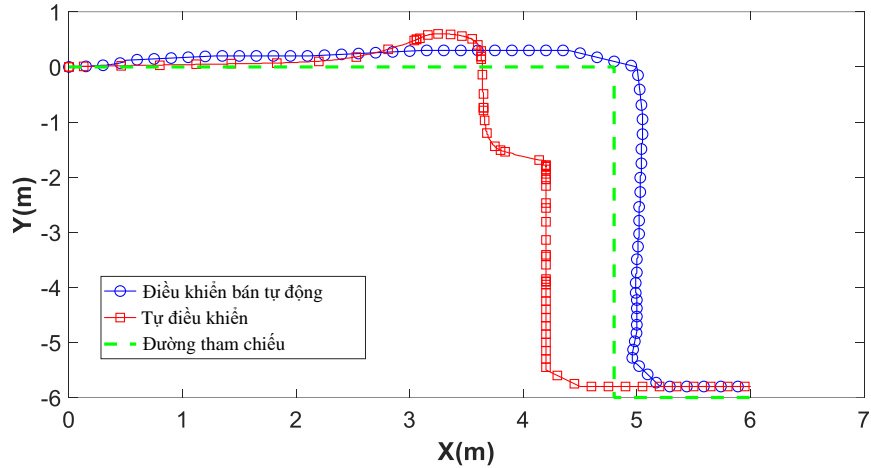
Từ bảng 5.8, có thể thấy hành động của xe lăn sẽ bao gồm 3 loại là Forward, Left và Right. Trong thí nghiệm này, khi xe lăn thực hiện hành động Forward, xe sẽ tiến lên phía trước một đoạn $\Delta S = 1,2$ m, tương ứng với việc di chuyển từ ô lưới này sang ô lưới kia trên bản đồ lưới 2D ảo. Dựa vào mô hình động lực học của xe lăn được trình bày trong chương 4, quãng đường đi được của xe lăn được tính như sau:

$$\Delta S = \frac{p_r + p_l}{2N_e} d_w \pi \quad (5.20)$$

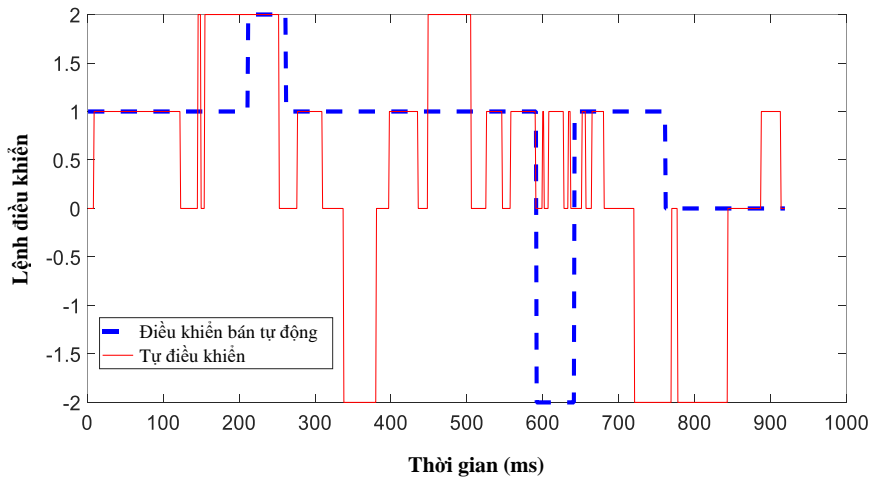
Trong đó, đường kính bánh xe $d_w = 0,457$ m; khoảng cách giữa hai bánh xe $L = 0,535$ m; hai encoder được gắn đồng trục với hai bánh xe có độ phân giải: $N_e = 2048$ xung/vòng; số xung đọc được từ bánh trái p_l ; số xung đọc được bánh phải p_r . Để điều khiển xe thực hiện đi thẳng, quay trái 90° hoặc quay phải 90° , góc quay của xe lăn $\Delta\theta$ được tính theo công thức (5.21). Ngoài ra, trong thí nghiệm này, đường đi thực tế của xe lăn với phương pháp DQNs đề xuất (đường màu xanh lam) được so sánh với đường dẫn tham chiếu (đường nét đứt màu xanh lá cây) như trong

hình 5.29b để đánh giá. Kết quả cho thấy xe lăn có thể di chuyển đến đích mong muốn với sai số trung bình 0,2 m ở trục X và 0,2 m ở trục Y.

$$\Delta\theta = \frac{p_r - p_l}{N_e L} d_w \pi \quad (5.21)$$



(a)



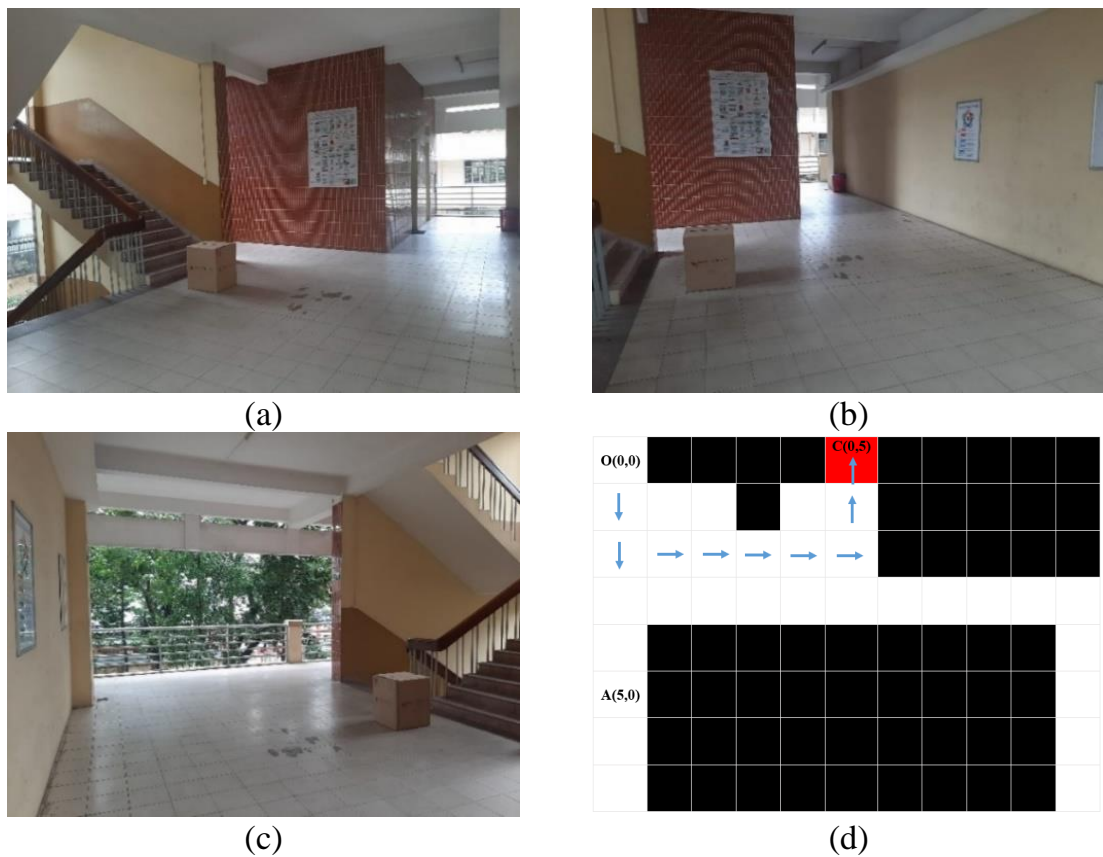
(b)

Hình 5.30. So sánh chuyển động của xe lăn theo hai phương pháp điều khiển (điều khiển bán tự động và tự điều khiển). (a) Đường đi thực tế của hai phương pháp điều khiển và đường tham chiếu; (b) Các lệnh điều khiển của hai phương pháp.

Mục đích của các thí nghiệm này là để so sánh kết quả của điều khiển bán tự động bằng mô hình DQNs với tự điều khiển của người dùng bằng tín hiệu EEG. Cụ thể, hình 5.30a thể hiện 2 đường chuyển động của xe lăn, trong đó đường màu xanh lam là của phương pháp bán tự động và đường màu đỏ là của phương pháp tự điều khiển. Từ hình 5.30a có thể thấy đường đi của xe lăn khi được điều khiển bằng

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

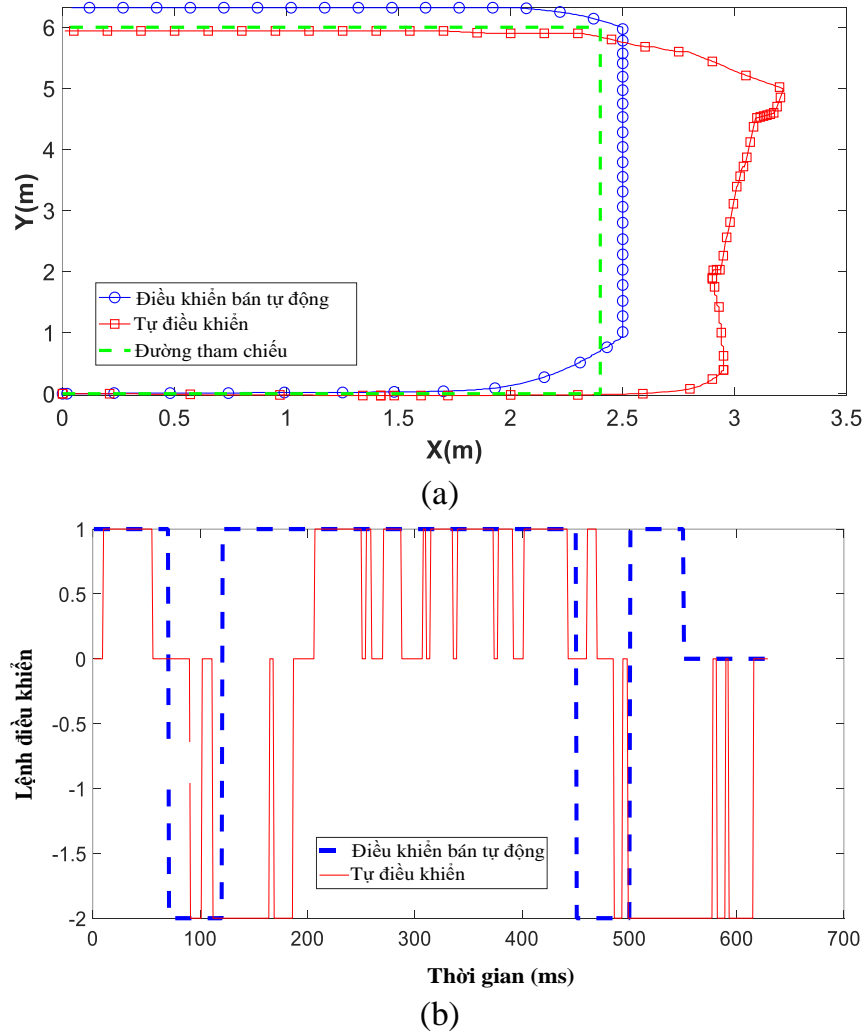
phương pháp bán tự động sẽ gần với đường tham chiếu hơn so với tự điều khiển. Ngoài ra, sự di chuyển của xe lăn sử dụng điều khiển bán tự động mượt mà và liên tục hơn so với tự điều khiển. Để làm rõ hai phương pháp điều khiển, các lệnh điều khiển xe lăn trong quá trình di chuyển để đến đích được ghi lại. Trên hình 5.30b, các lệnh điều khiển được thể hiện tại trục tung với các giá trị -2, 0, 1, 2 tương ứng với các lệnh rẽ trái, dừng, đi thẳng và rẽ phải. Do đó, có thể thấy xe lăn di chuyển với độ ổn định cao trong trường hợp điều khiển bán tự động so với khi người dùng tự điều khiển. Ngoài ra, kết quả cho thấy người sử dụng điều khiển bán tự động tốn ít thời gian hơn để về đích, với trung bình khoảng 80 giây so với khi người sử dụng tự điều khiển với thời gian trung bình khoảng 95 giây.



Hình 5.31. Môi trường trong thí nghiệm thứ hai: (a)-(c) Môi trường thực; (d) Đường dẫn mũi tên màu xanh lam được mô phỏng bằng DQNs.

Trong một thí nghiệm khác với môi trường thực như hình 5.31a – 5.31c. Hình 5.31d cho thấy các đường mô phỏng (mũi tên xanh) của xe lăn dựa trên bản đồ 2D ảo tương ứng với môi trường thực khi xe lăn di chuyển từ vị trí O(0,0) đến đích C(0,5).

Hình 5.32 mô tả sự khác nhau về đường đi và lệnh điều khiển của xe lăn điện với phương pháp điều khiển bán tự động và tự điều khiển để di chuyển trong môi trường thí nghiệm thứ 2 được mô tả ở hình 5.31.



Hình 5.32. So sánh chuyển động của xe lăn trong hai phương pháp điều khiển (điều khiển bán tự động và tự điều khiển): (a) Đường đi thực tế của hai phương pháp điều khiển và đường tham chiếu; (b) Các lệnh điều khiển của hai phương pháp

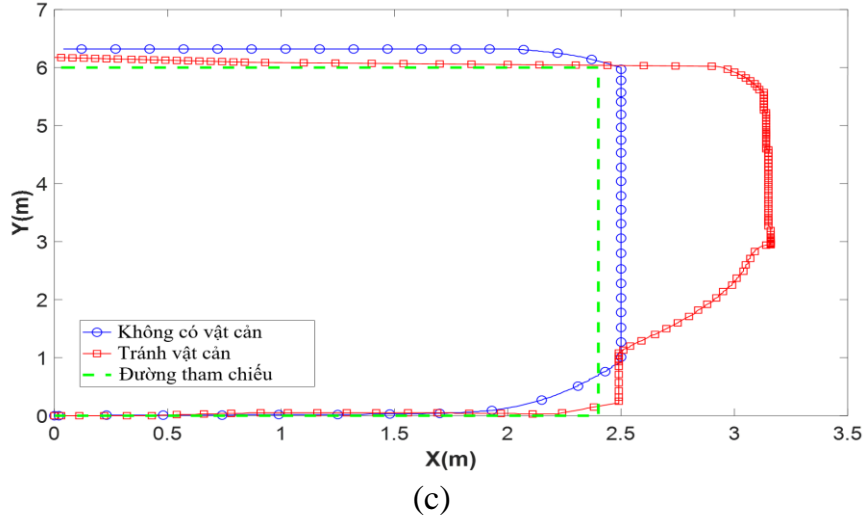
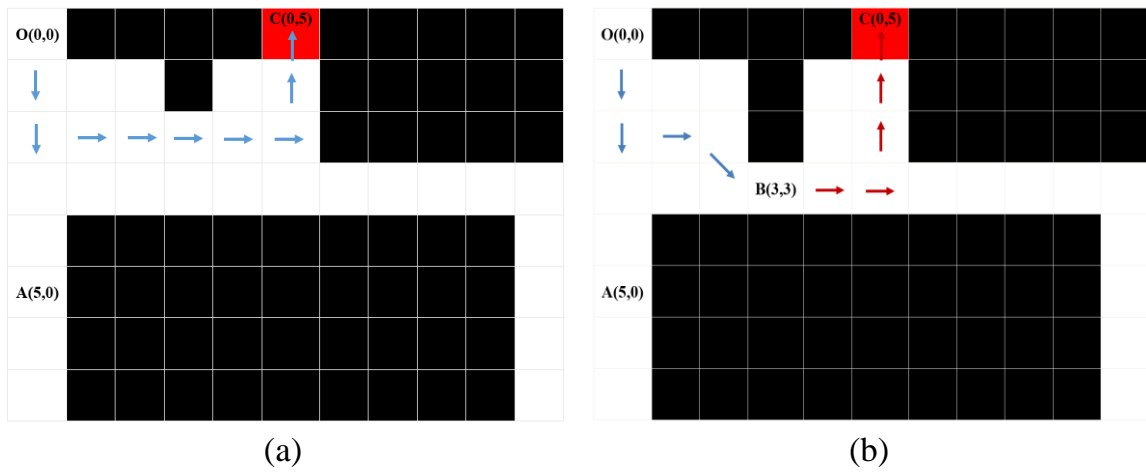
Từ hình 5.32a có thể thấy, đường đi của xe lăn được điều khiển bằng phương pháp điều khiển bán tự động ngắn hơn và mượt mà hơn so với tự điều khiển. Hơn nữa, như mô tả ở bảng 5.9, phương pháp điều khiển bán tự động có sai số trung bình là 0,1 m ở trục X và 0,3 m ở trục Y so với sai số 0,5 m ở trục X và 0,5 m ở trục Y của phương pháp tự điều khiển. Với hình 5.32b, có thể thấy xe lăn di chuyển với độ ổn

Chương 5: Mô hình điều khiển xe lăn điện kết hợp tín hiệu EEG và camera dựa vào bản đồ

định cao trong trường hợp điều khiển bán tự động so với chế độ người dùng tự điều khiển.

Bảng 5.9. Sai số đường đi của xe lăn trong các môi trường thí nghiệm.

Môi trường thí nghiệm	Tự điều khiển		Điều khiển bán tự động	
	Sai số trục X	Sai số trục Y	Sai số trục X	Sai số trục Y
	(m)	(m)	(m)	(m)
1	0,7	0,9	0,2	0,2
2	0,5	0,5	0,1	0,3



Hình 5.33. Chuyển động của xe lăn khi có vật cản: (a) Đường đi mô phỏng của xe lăn khi không có vật cản; (b) Đường đi mô phỏng của xe lăn khi có vật cản; (c) Biểu diễn đường đi thực của xe lăn trong trường hợp có và không có vật cản.

Hình 5.33 mô tả thí nghiệm có vật cản xuất hiện trên đường đi đã được lập kế hoạch trước. Trong thí nghiệm này, xe lăn ở vị trí $O(0,0)$ và chọn điểm đích là $C(0,5)$. Dựa vào vị trí bắt đầu và kết thúc trên bản đồ 2D lưới ảo, mô hình DQNs đề xuất đường đi mô phỏng như hình 5.33a (mũi tên màu xanh). Tuy nhiên, trong quá trình di chuyển, trên lộ trình này xuất hiện một vật cản như hình 5.33b. Do đó, hệ thống xe lăn điện tự động phát hiện vật cản và tìm khoảng trống, từ đó di chuyển về vị trí $B(3,3)$. Tại đây, hệ thống xe lăn điện sẽ tiến hành định vị lại vị trí dựa vào vật mốc trong môi trường, làm cơ sở để mô hình DQNs đề xuất đường đi mới để đến đích $C(0,5)$ như mong muốn. Hình 5.33b với các mũi tên màu đỏ mô tả cho đường đi mới được đề xuất cho lộ trình di chuyển từ $B(3,3)$ đến $C(0,5)$. Hình 5.33c mô tả đường đi thực tế được vẽ lại để kiểm chứng cho đường đi mô phỏng. Kết quả từ thí nghiệm này cho thấy, hệ thống xe lăn có khả năng tránh vật cản và tìm đường đi mới để đến đích mong muốn.

Mặc dù phương pháp điều khiển bán tự động cho xe lăn điện có nhiều ưu điểm nhưng vẫn còn một số vấn đề cần được cải tiến. Đầu tiên là độ sai lệch đường đi của xe lăn khi di chuyển ở chế độ bán tự động. Sai lệch này do nguyên nhân chủ quan như thiết kế cơ khí của xe lăn, phương pháp điều khiển động cơ, hoặc do nguyên nhân khách quan như môi trường di chuyển, vị trí ngồi của người sử dụng. Về phần cơ khí của xe lăn, nếu các bánh xe có đường kính không đồng đều sẽ dẫn đến quãng đường di chuyển của mỗi bánh sẽ khác nhau với cùng một tốc độ điều khiển động cơ, do đó hướng di chuyển của xe lăn sẽ sai lệch so với hướng đi ban đầu. Bên cạnh đó, vị trí người dùng cũng ảnh hưởng đến trọng tâm của xe và môi trường di chuyển không bằng phẳng cũng dẫn đến sự khác biệt về vận tốc của các bánh xe, làm sai lệch về đường đi. Để khắc phục vấn đề này, thiết kế cơ khí và bộ điều khiển của xe lăn cần được cải tiến để tăng độ chính xác.

❖ Kết luận

Mô hình điều khiển thực - ảo cho xe lăn điện bán tự động kết hợp với hệ thống camera RGB-D, giao tiếp BCI sử dụng tín hiệu EEG từ hoạt động nháy mắt với giao

diện người dùng đồ họa và các vật mốc tự nhiên đã được đề xuất để thực hiện lập kế hoạch đường đi tối ưu cho việc điều hướng xe lăn điện bán tự động [115, 145]. Cụ thể, bản đồ lưới 2D ảo được tạo ra từ môi trường thực được sử dụng để huấn luyện tạo đường đi ngắn nhất đến đích, trong đó mô hình DQNs giúp quá trình huấn luyện đạt hiệu quả cao. Sau khi mô hình được huấn luyện, người khuyết tật có thể chọn đích đến mong muốn trên một giao diện người dùng bằng cách sử dụng tín hiệu EEG từ hoạt động nháy mắt. Hơn nữa, xe lăn bán tự động tự định vị dựa trên các vật mốc tự nhiên trong quá trình di chuyển theo lộ trình tối ưu từ bộ lập kế hoạch MP trong môi trường thực tế trong nhà. Các thí nghiệm đã được thực hiện và chỉ ra rằng hệ thống vật mốc được sử dụng để định vị xe lăn trong môi trường trong nhà là hiệu quả với độ chính xác cao và tốc độ xử lý nhanh, đáp ứng với thời gian thực. Từ những thí nghiệm này, hệ thống camera RGB-D là một lựa chọn tốt trong nghiên cứu này vì chi phí thấp, phổ biến và dễ phát triển với độ phân giải cao để cung cấp thông tin 3D của môi trường trong nhà.

Với phương pháp đề xuất đường đi tối ưu dựa trên DQNs, xe lăn bán tự động có thể vận hành tốt để đến được điểm đích mong muốn với sai số nhỏ so với quỹ đạo mô phỏng, cũng như quỹ đạo của người dùng tự điều khiển bằng EEG. Rõ ràng là, với đường đi tối ưu được đề xuất và phương pháp điều khiển bán tự động, chuyển động của xe lăn ổn định hơn, an toàn hơn và mất ít thời gian hơn để di chuyển. Tuy nhiên, hệ thống xe lăn bán tự động vẫn còn một số hạn chế về môi trường hoạt động, quỹ đạo di chuyển vẫn còn sai số. Hệ thống điều khiển thực - ảo cho xe lăn điện cần tiếp tục được phát triển để áp dụng cho các môi trường phức tạp hơn với nhiều chướng ngại vật trong tương lai.

CHƯƠNG 6

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

6.1. KẾT LUẬN

Xe lăn điện là phương tiện cần thiết cho những người tàn tật nặng như bị mất chi, bị chấn thương tủy sống hoặc rối loạn thần kinh cơ, bởi nó cho phép họ có thể di chuyển độc lập trong các hoạt động hàng ngày. Luận án này đã kiến nghị một số phương pháp cho việc điều khiển bán tự động cho xe lăn điện trong môi trường trong nhà dựa vào tín hiệu điện não đồ EEG, camera, các vật mốc và bản đồ. Trong mô hình điều khiển xe lăn điện mà luận án kiến nghị gồm có các thành phần quan trọng là mô hình điều khiển thực - ảo sử dụng mô hình DQNs và bản đồ lưới 2D ảo để tìm đường đi tối ưu cho xe lăn, phương pháp nhận dạng vật mốc tự nhiên và xác định vị trí của các vật mốc trong môi trường. Từ đó, những thành phần này làm cơ sở để định vị xe lăn điện trên bản đồ lưới 2D, và phương pháp phân loại hoạt động mắt dựa vào tín hiệu EEG cho các ứng dụng giao tiếp giữa não người và máy tính nhằm xác định nhu cầu di chuyển của người dùng.

Với đặc điểm của các đối tượng tàn tật nặng với những phần cơ thể khác nhưng mắt vẫn còn hoạt động tốt, tín hiệu EEG từ các hoạt động mắt là phù hợp cho việc giao tiếp giữa người dùng và hệ thống điều khiển. Từ đó, luận án **đề xuất hai phương pháp phân loại các hoạt động mắt từ tín hiệu điện não đồ EEG gồm phương pháp ngưỡng biên độ và mô hình mạng học sâu CNN-1D**. Ưu điểm của phương pháp phân loại các hoạt động mắt như nháy mắt trái, nháy mắt phải, dùng thuật toán ngưỡng biên độ là độ chính xác cao với tỷ lệ lượt là 97% và 99% cho mỗi loại. Hơn nữa, phương pháp này có thể xử lý trực tiếp trên tín hiệu thu được từ các điện cực F7 và F8 mà không cần phải được huấn luyện trước. Trong khi với phương pháp phân

loại các hoạt động mắt dùng mạng học sâu CNN-1D, tập dữ liệu huấn luyện cần được thu thập trước theo một quy trình được chuẩn hóa. Dữ liệu được thu thập từ thiết bị Emotiv EPOC+ cho 14 kênh tín hiệu EEG, trong đó 4 điện cực F7, AF3, F8, AF4 được sử dụng chính cho thu tín hiệu. Cụ thể, với đặc điểm của các hoạt động mắt, chỉ có tín hiệu từ 4 điện cực F7, AF3, F8, AF4 được trích xuất và ghép lại tạo thành 1 tín hiệu cho huấn luyện. Các tín hiệu này được làm trơn dùng bộ lọc Savitzky-Golay trước khi đưa vào huấn luyện để tăng độ chính xác. Kết quả phân loại cho 5 loại nháy mắt gồm nháy mắt trái, nháy mắt phải, nháy hai mắt, nháy hai mắt hai lần liên tiếp và không nháy mắt lần lượt là 98,1%, 100%, 95,9%, 100% và 98,1%. So với phương pháp ngưỡng biên độ thì việc dùng mạng học sâu CNN-1D có lợi thế ở việc có thể mở rộng số lượng loại hoạt động nháy mắt tùy thuộc vào số lượng lệnh điều khiển của người dùng trong các ứng dụng giao tiếp giữa não người và máy tính hoặc hệ thống điều khiển.

Luận án còn **đề xuất phương pháp mật độ điểm đặc trưng lớn nhất cho việc nhận dạng các vật mốc tự nhiên trong môi trường và phương pháp xác định vị trí của vật mốc dựa vào vị trí xe lăn và thông tin 3D từ camera**. Cụ thể, hình ảnh môi trường thu thập được từ camera sẽ được trích xuất đặc trưng dùng bộ dò tìm ORB. Với các điểm đặc trưng này, các thuật toán hình thái học như giãn nở được thực hiện để kết nối các điểm đặc trưng này lại với nhau, giúp tạo thành các đối tượng trong ảnh. Sau đó, thuật toán mật độ điểm đặc trưng lớn nhất được áp dụng để tính toán mật độ điểm đặc trưng cho từng đối tượng đã được xác định trước đó và lựa chọn đối tượng nào có mật độ điểm đặc trưng lớn nhất làm vật mốc. Độ chính xác của việc nhận dạng các vật mốc tự nhiên được đánh giá bằng hệ số chồng lấp IOU giữa vật mốc thực và vật mốc được lựa chọn từ phương pháp mật độ điểm đặc trưng lớn nhất. Kết quả cho thấy khoảng cách dưới 2 m từ camera đến vật mốc, đạt độ chính xác cao nhất với giá trị IOU lớn hơn 0,8. Hơn nữa, thời gian xử lý nhận dạng vật mốc trung bình là 38,08 ms, rất phù hợp với các tác vụ thời gian thực. Với vật mốc đã nhận dạng được, thông tin 3D của vật mốc và vị trí của xe lăn trong môi trường được dùng để xác định vị trí vật mốc. Cụ thể, vị trí của xe lăn được xác định trong quá trình di

chuyển dựa vào phương trình động lực học của xe lăn và encoder được gắn ở hai bánh xe, và thông tin 3D của vật mốc được xác định dựa vào camera RGB-D. Kết quả cho thấy vị trí vật mốc với khoảng cách từ camera đến vật mốc dưới 200 cm có sai số nhỏ hơn 3,0 cm theo phương ngang và nhỏ hơn 2,0 cm theo phương dọc.

Cho việc điều khiển xe lăn điện trong môi trường trong nhà, **luận án đã kiến nghị mô hình điều khiển thực - ảo cho xe lăn điện, trong đó áp dụng mô hình DQNs-PreLU với cấu trúc và tham số mô hình đã được lựa chọn tối ưu nhất cho việc tìm đường đi cho xe lăn dựa vào bản đồ lưới 2D ảo. Bên cạnh đó, phương pháp định vị xe lăn trên bản đồ lưới 2D dựa vào vật mốc và phương pháp điều khiển xe lăn điện trong môi trường thực dựa vào lộ trình di chuyển từ mô hình DQNs-PreLU đã được đề xuất.** Cụ thể, với bản đồ lưới 2D ảo có được từ môi trường thực bao gồm các ô trống, ô vật cản và ô đích, mô hình DQNs với hàm kích hoạt PreLU được lựa chọn để huấn luyện để đề xuất đường đi tối ưu cho xe lăn đến đích mong muốn từ bất kỳ vị trí bắt đầu nào của xe lăn trên bản đồ lưới 2D. Với mô hình DQNs – PreLU, thời gian huấn luyện cho bản đồ lưới (8×11) ít hơn gần 5 lần so với mô hình Q-Learning và SARSA, và ít hơn gần 12 lần so với mô hình DQNs – ReLU, gần 3 lần so với mô hình Q-Learning, gần 2 lần so với mô hình SARSA khi huấn luyện cho bản đồ lưới (11×33). Ngoài thời gian huấn luyện ít hơn so với các mô hình khác, mô hình DQNs – PreLU còn có thể lưu lại bộ thông số sau khi huấn luyện và dùng cho việc tìm đường đi cho xe lăn trong môi trường thực. Ngoài ra, để mô hình DQNs có thể đề xuất đường đi cho xe lăn thì cần phải cung cấp thông tin vị trí hiện tại của xe lăn và vị trí điểm đích mong muốn. Với kết quả đã đạt được từ việc phân loại các hoạt động nháy mắt từ tín hiệu EEG, một giao diện người dùng đã được thiết kế cho việc lựa chọn các điểm đích đã được định trước trên bản đồ lưới 2D ảo. Bằng việc nháy mắt trái, nháy mắt phải và nháy hai mắt, người dùng có thể lựa chọn đích đến mong muốn trên giao diện. Cùng với đó, vị trí xe lăn trên bản đồ lưới cũng được xác định thông qua phương pháp định vị xe lăn dựa vào vật mốc để làm đầu vào cho mô hình DQNs tìm đường đi đến đích mong muốn. Kết quả thí nghiệm cho thấy sai số vị trí của xe lăn so với vị trí thực trong không gian OXY là nhỏ, từ 2 – 6 cm theo

trục X và 2 – 5 cm theo trục Y, và vị trí của xe lăn trên bản đồ lưới 2D là hoàn toàn trùng khớp với vị trí thực. Hơn nữa, các thí nghiệm về sự di chuyển của xe lăn theo phương pháp điều khiển đề xuất cho thấy xe lăn có thể di chuyển tự động đến đích mong muốn với quỹ đạo di chuyển liên tục, không bị gián đoạn như khi để người dùng tự điều khiển.

Với những mô hình và phương pháp đề xuất cùng kết quả đạt được trong luận án này, hệ thống điều khiển xe lăn điện bán tự động cho người khuyết tật trong môi trường trong nhà đã được hoàn thành. Ngoài những kết quả đạt được, mô hình vẫn còn một số hạn chế như môi trường hoạt động chưa phức tạp, tốc độ xử lý tín hiệu EEG vẫn còn chậm, quỹ đạo di chuyển vẫn có sai lệch do thiết kế cơ khí của mô hình xe lăn. Mặc dù vậy, mô hình xe lăn điện mà luận án này kiến nghị là rất thực tiễn, hữu ích và có ý nghĩa cho việc hỗ trợ di chuyển đối với những người bị hạn chế về vận động tay, chân hoặc đầu nhưng mắt vẫn còn khỏe. Với các thiết bị được sử dụng trong nghiên cứu này có chi phí phù hợp, giá thành của mô hình xe lăn điện được đề xuất thấp hơn so với các mô hình xe lăn thông minh khác, và do đó người dùng sẽ dễ dàng tiếp cận hơn. Điều quan trọng nhất là người dùng sẽ chủ động và tự tin hơn trong cuộc sống của mình, cảm thấy thoải mái và an toàn với các công nghệ mà họ được hỗ trợ.

6.2. HƯỚNG PHÁT TRIỂN

Ngày nay, các hệ thống nhúng dựa trên nền tảng máy tính nhỏ gọn với tốc độ xử lý cao đã có sự phát triển rất nhanh chóng. Do vậy, những giải thuật và phương pháp được đề xuất trong luận án có thể tiếp tục nghiên cứu và tích hợp vào một thiết bị chuyên dụng có kích thước nhỏ gọn, giúp cho việc kết nối và tương thích với các loại xe lăn điện hiện có trên thị trường được nhanh chóng hơn, từ đó giúp giảm giá thành sản phẩm. Hơn nữa, với sự phát triển của hệ thống IoT và hệ thống y tế điện tử, hệ thống quản lý, điều phối và cảnh báo cần được nghiên cứu và áp dụng trong các bệnh viện, viện dưỡng lão để giám sát và đảm bảo sự an toàn cho các bệnh nhân sử dụng xe lăn điện.

TÀI LIỆU THAM KHẢO

- [1] "Global report on health equity for persons with disabilities," *World Health Organization*, pp. 1-312, 2022.
- [2] "Điều tra Quốc gia về Người Khuyết tật tại Việt Nam," *Tổng cục Thống kê*, 2019.
- [3] H. A. Lamti, P. Gorce, M. M. Ben Khelifa, and A. M. Alimi, "When mental fatigue maybe characterized by Event Related Potential (P300) during virtual wheelchair navigation," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 19, no. 16, pp. 1749-1759, 2016.
- [4] "Electric Wheelchair Market - Growth, Trends, COVID-19 Impact, and Forecast (2022 - 2027)," *Mordor Intelligence*, pp. 1-90, 2022.
- [5] J. H. Choi, Y. Chung, and S. Oh, "Motion control of joystick interfaced electric wheelchair for improvement of safety and riding comfort," *Mechatronics*, vol. 59, pp. 104-114, 2019.
- [6] E. Y. Kim, "Wheelchair Navigation System for Disabled and Elderly People," *Sensors*, vol. 16, no. 11, pp. 1-24, 2016.
- [7] X. Xu, Y. Zhang, Y. Luo, and D. Chen, "Robust Bio-Signal Based Control of an Intelligent Wheelchair," *Robotics*, vol. 2, no. 4, pp. 187-197, 2013.
- [8] L. Xi and M. Shino, "Shared Control of an Electric Wheelchair Considering Physical Functions and Driving Motivation," *International Journal of Environmental Research and Public Health*, vol. 17, no. 15, 2020.
- [9] M. M. Abdulghani, K. M. Al-Aubidy, M. M. Ali, and Q. J. Hamarsheh, "Wheelchair Neuro Fuzzy Control and Tracking System Based on Voice Recognition," *Sensors*, vol. 20, no. 10, 2020.
- [10] J. Kim *et al.*, "The tongue enables computer and wheelchair control for people with spinal cord injury," *Science translational medicine*, vol. 5, no. 213, pp. 1-22, 2013.
- [11] J. S. Ju, Y. H. Shin, and E. Y. Kim, "Intelligent Wheelchair (IW) Interface using face and mouth recognition," in *IUI09: 14th International Conference on Intelligent User Interfaces*, Sanibel Island, Florida, USA, pp. 307-314, 2009.
- [12] Y. Zhang, J. Zhang, and Y. Luo, "A novel intelligent wheelchair control system based on hand gesture recognition," in *The 2011 IEEE/ICME International Conference on Complex Medical Engineering*, pp. 334-339, 2011.
- [13] M. S. Sadi, M. Alotaibi, M. R. Islam, M. S. Islam, T. Alhmiedat, and Z. Bassfar, "Finger-Gesture Controlled Wheelchair with Enabling IoT," *Sensors*, vol. 22, no. 22, pp. 1-16, 2022.
- [14] G. Patil, K. S. Gharge, V. Modhave, and Y. Angal, "Design of Wheelchair using Finger Operation with Image Processing Algorithms," *International*

- Journal of Research in Engineering and Technology*, vol. 03, no. 02, pp. 232-237, 2014.
- [15] M. Dahmani *et al.*, "An Intelligent and Low-Cost Eye-Tracking System for Motorized Wheelchair Control," *Sensors*, vol. 20, no. 14, pp. 1-27, 2020.
- [16] R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: control signals review," *Neurocomputing*, vol. 223, pp. 26-44, 2017.
- [17] W. Alvarado, A. Roman-Gonzalez, and B. Meneses, "Implementation of a Brain-machine interface for controlling a wheelchair," in *CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, Pucon, Chile, pp. 1-6, 2017.
- [18] H. T. Nguyen, N. Trung, V. Toi, and V. S. Tran, "An autoregressive neural network for recognition of eye commands in an EEG-controlled wheelchair," in *2013 International Conference on Advanced Technologies for Communications (ATC 2013)*, pp. 333-338, 2013.
- [19] H. T. Tran, H. T. Nguyen, H. V. Phan, V. Van Toi, T. V. Ngo, and C. Bui-Thu, "An EEG-Controlled Wheelchair Using Eye Movements," in *5th International Conference on Biomedical Engineering in Vietnam*, Cham, pp. 470-473, 2015.
- [20] I. A. Mirza *et al.*, "Mind-controlled wheelchair using an EEG headset and arduino microcontroller," in *2015 International Conference on Technologies for Sustainable Development (ICTSD)*, pp. 1-5, 2015.
- [21] A. L. Barriuso, J. Pérez-Marcos, D. M. Jiménez-Bravo, G. Villarrubia González, and J. F. De Paz, "Agent-Based Intelligent Interface for Wheelchair Movement Control," *Sensors*, vol. 18, no. 5, pp. 1-31, 2018.
- [22] Q. Zeng, C. L. Teo, B. Rebsamen, and E. Burdet, "A Collaborative Wheelchair System," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 16, no. 2, pp. 161-170, 2008.
- [23] M. Zacharie, "GPS and Discrete Kalman Filter for Indoor Robot Navigation," *International Journal of Electrical and Computer Engineering*, vol. 60, pp. 1635 - 1641, 2011.
- [24] A. Murarka, J. Modayil, and B. Kuipers, "Building Local Safety Maps for a Wheelchair Robot using Vision and Lasers," in *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, pp. 25-25, 2006.
- [25] F. B. Taher, N. B. Amor, and M. Jallouli, "EEG control of an electric wheelchair for disabled persons," in *2013 International Conference on Individual and Collective Behaviors in Robotics (ICBR)*, pp. 27-32, 2013.
- [26] J. Pu, Y. Jiang, X. Xie, X. Chen, M. Liu, and S. Xu, "Low cost sensor network for obstacle avoidance in share-controlled smart wheelchairs under daily scenarios," *Microelectronics Reliability*, vol. 83, pp. 180-186, 2018.
- [27] V. Sharma, E. F. Simpson Rc Fau - LoPresti, M. LoPresti Ef Fau - Schmeler, and M. Schmeler, "Clinical evaluation of semiautonomous smart wheelchair architecture (Drive-Safe System) with visually impaired individuals," *The*

- Journal of Rehabilitation Research and Development*, vol. 49, no. 1, pp. 35-50, 2012.
- [28] E. F. Lopresti, R. C. Sharma V Fau - Simpson, L. C. Simpson Rc Fau - Mostowy, and L. C. Mostowy, "Performance testing of collision-avoidance system for power wheelchairs," *Journal of rehabilitation research and development*, vol. 48, no. 5, pp. 529-544, 2011.
- [29] M. Njah and M. Jallouli, "Fuzzy-EKF Controller for Intelligent Wheelchair Navigation," *J. Intell. Syst.*, vol. 25, no. 2, pp. 107-121, 2016.
- [30] T.-H. Nguyen, B.-V. Ngo, and T.-H. Quach, "Optimization of Transformation Matrix for 3D Cloud Mapping Using Sensor Fusion," *American Journal of Signal Processing*, vol. 8, no. 1, pp. 9-19, 2018.
- [31] L. Zhang, X. Wu, R. Gao, L. Pan, and Q. Zhang, "A multi-sensor fusion positioning approach for indoor mobile robot using factor graph," *Measurement*, vol. 216, pp. 1-12, 2023.
- [32] J. S. Nguyen, T. N. Nguyen, Y. Tran, S. W. Su, A. Craig, and H. T. Nguyen, "Real-time performance of a hands-free semi-autonomous wheelchair system using a combination of stereoscopic and spherical vision," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3069-3072, 2012.
- [33] V. B. Bastos, A. F. P. Tavares, C. H. C. Quiroz, P. R. M. Silva, M. V. P. Lima, and P. R. G. Kurka, "Monocular visual odometry for robotic wheelchair in a virtual environment," in *2018 Annual IEEE International Systems Conference (SysCon)*, pp. 1-8, 2018.
- [34] H. Liu, C. Wu, and H. Wang, "Real time object detection using LiDAR and camera fusion for autonomous driving," *Scientific Reports*, vol. 13, no. 1, pp. 1-12, 2023.
- [35] J. Hoey, D. Gunn, A. Mihailidis, and P. Elinas, "Obstacle Avoidance Wheelchair System," in *International Conference on Robotics and Automation*, Orlando, Florida, USA, pp. 1-4, 2006.
- [36] J. S. Nguyen, T. H. Nguyen, and H. T. Nguyen, "Semi-autonomous wheelchair system using stereoscopic cameras," in *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5068-5071, 2009.
- [37] J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha, "Visual simultaneous localization and mapping: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 55-81, 2015.
- [38] J. L. Cras and J. Paxman, "A modular hybrid SLAM for the 3D mapping of large scale environments," in *2012 12th International Conference on Control Automation Robotics & Vision (ICARCV)*, pp. 1036-1041, 2012.
- [39] G. Lee, B.-C. Moon, S. Lee, and D. Han, "Fusion of the SLAM with Wi-Fi-Based Positioning Methods for Mobile Robot-Based Learning Data Collection, Localization, and Tracking in Indoor Spaces," *Sensors*, vol. 20, no. 18, pp. 1-20, 2020.

- [40] Y. Zhou and W. Liu, "Preliminary Research on Indoor Mobile Robot Localization using Laser-activated RFID," in *2007 IEEE International Conference on RFID*, pp. 78-85, 2007.
- [41] R. C. Luo, L. Jian-Xian, and C. T. Chen, "Indoor localization using line based map for autonomous mobile robot," in *2008 IEEE Workshop on Advanced robotics and Its Social Impacts*, pp. 1-6, 2008.
- [42] R. Lin, Z. Wang, R. Sun, and L. Sun, "Image features-based mobile robot localization," in *2012 IEEE International Conference on Information and Automation*, pp. 304-310, 2012.
- [43] S. B. Han, J. H. Kim, and H. Myung, "Landmark-Based Particle Localization Algorithm for Mobile Robots With a Fish-Eye Vision System," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 6, pp. 1745-1756, 2013.
- [44] P. F. Alcantarilla, S. M. Oh, G. L. Mariottini, L. M. Bergasa, and F. Dellaert, "Learning visibility of landmarks for vision-based localization," in *2010 IEEE International Conference on Robotics and Automation*, pp. 4881-4888, 2010.
- [45] A. Basu, S. K. Ghosh, and S. Sarkar, "Autonomous navigation and 2d mapping using SONAR," in *2016 5th International Conference on Wireless Networks and Embedded Systems (WECON)*, pp. 1-5, 2016.
- [46] X. Liu, B. Guo, and C. Meng, "A method of simultaneous location and mapping based on RGB-D cameras," in *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 1-5, 2016.
- [47] X. Zhong, Y. Zhou, and H. Liu, "Design and recognition of artificial landmarks for reliable indoor self-localization of mobile robots," *International Journal of Advanced Robotic Systems*, vol. 14, no. 1, pp. 1-13, 2017.
- [48] T. Yu and Y. Shen, "Asymptotic Performance Analysis for Landmark Learning in Indoor Localization," *IEEE Communications Letters*, vol. 22, no. 4, pp. 740-743, 2018.
- [49] N. T. Hai, "Development of a smart wheelchair using EEG technology and stereoscopic camera for disabled people," *International University*, vol. C2013-28-06, 2013.
- [50] L. Q. Chuyên, "Mạng neural trong hệ thống điều khiển xe lăn cho người tàn tật nặng sử dụng điện não (EEG) và camera," *Luận Án Tiến Sĩ - Trường Đại Học Giao Thông Vận Tải Tp.HCM*, 2020.
- [51] A. Lopes, J. Rodrigues, J. Perdigo, G. Pires, and U. Nunes, "A New Hybrid Motion Planner: Applied in a Brain-Actuated Robotic Wheelchair," *IEEE Robotics & Automation Magazine*, vol. 23, no. 4, pp. 82-93, 2016.
- [52] Z. Li, S. Zhao, J. Duan, C. Y. Su, C. Yang, and X. Zhao, "Human Cooperative Wheelchair With Brain-Machine Interaction Based on Shared Control Strategy," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 185-195, 2017.
- [53] J. Tang, Y. Liu, D. Hu, and Z. Zhou, "Towards BCI-actuated smart wheelchair system," *BioMedical Engineering OnLine*, vol. 17, no. 1, pp. 1-22, 2018.

- [54] H. N. Oon, A. Saidatul, and Z. Ibrahim, "Analysis on Non-Linear Features of Electroencephalogram (EEG) Signal for Neuromarketing Application," in *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, pp. 1-8, 2018.
- [55] D.-K. Tran *et al.*, "Evaluation Between Visual And Auditory EEG Responses On Upper Limb Movement in The Temporal and Spectral Domains," in *2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*, pp. 817-824, 2022.
- [56] A. J. N. John G. Webster, "Medical Instrumentation: Application and Design, 5th Edition," *Hoboken, NJ : John Wiley & Sons*, pp. 126-188, 2020.
- [57] D.-K. Tran, T.-H. Nguyen, and T.-N. Nguyen, "Detection of EEG-Based Eye-Blinks Using A Thresholding Algorithm," *European Journal of Engineering and Technology Research*, vol. 6, no. 4, pp. 6-12, 2021.
- [58] T. Nguyen, T. H. Nguyen, K. Q. D. Truong, and T. Van Vo, "A Mean Threshold Algorithm for Human Eye Blinking Detection Using EEG," in *4th International Conference on Biomedical Engineering in Vietnam*, Berlin, Heidelberg, pp. 275-279, 2013.
- [59] M. N. Tibdewal, R. R. Fate, M. Mahadevappa, A. K. Ray, and M. Malokar, "Classification of artifactual EEG signal and detection of multiple eye movement artifact zones using novel Time-amplitude algorithm," *Signal, Image and Video Processing*, vol. 11, no. 2, pp. 333-340, 2017.
- [60] A. Qayyum, M. K. A. A. Khan, M. Mazher, and M. Suresh, "Classification of EEG Learning and Resting States using 1D-Convolutional Neural Network for Cognitive Load Assesment," in *2018 IEEE Student Conference on Research and Development (SCORED)*, pp. 1-5, 2018.
- [61] Z. Pei, H. Wang, A. Bezerianos, and J. Li, "EEG-Based Multiclass Workload Identification Using Feature Fusion and Selection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-8, 2021.
- [62] P. Zhang, X. Wang, W. Zhang, and J. Chen, "Learning Spatial–Spectral–Temporal EEG Features With Recurrent 3D Convolutional Neural Networks for Cross-Task Mental Workload Assessment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 1, pp. 31-42, 2019.
- [63] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals," *Computers in Biology and Medicine*, vol. 100, pp. 270-278, 2018.
- [64] A. N. Belkacem, N. Jamil, J. A. Palmer, S. Ouhbi, and C. Chen, "Brain Computer Interfaces for Improving the Quality of Life of Older Adults and Elderly Patients," *Frontiers in Neuroscience*, vol. 14, pp. 1-11, 2020.
- [65] C. Chao *et al.*, "Quadcopter Robot Control Based on Hybrid Brain–Computer Interface System," *Sensors and Materials*, vol. 32, no. 3, pp. 991-1004, 2020.

- [66] A. E. Hramov, V. A. Maksimenko, and A. N. Pisarchik, "Physical principles of brain–computer interfaces and their applications for rehabilitation, robotics and control of human brain states," *Physics Reports*, vol. 918, pp. 1-133, 2021.
- [67] H. Wang *et al.*, "Brain-Controlled Wheelchair Review: From Wet Electrode to Dry Electrode, From Single Modal to Hybrid Modal, From Synchronous to Asynchronous," *IEEE Access*, vol. 9, pp. 55920-55938, 2021.
- [68] A. Widyotriatmo, Suprijanto, and S. Andronicus, "A collaborative control of brain computer interface and robotic wheelchair," in *2015 10th Asian Control Conference (ASCC)*, pp. 1-6, 2015.
- [69] F. B. Taher, N. B. Amor, and M. Jallouli, "A multimodal wheelchair control system based on EEG signals and Eye tracking fusion," in *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, pp. 1-8, 2015.
- [70] Y. Xie and X. Li, "A brain controlled wheelchair based on common spatial pattern," in *2015 International Symposium on Bioelectronics and Bioinformatics (ISBB)*, pp. 19-22, 2015.
- [71] P. Ivan Halim, S. Jennifer, B. Widodo, and G. Alexander Agung Santoso, "Fast brain control systems for electric wheelchair using support vector machine," in *First International Workshop on Pattern Recognition*, Tokyo, Japan, vol. 10011, pp. 1-6, 2016.
- [72] C. Naijian, H. Xiangdong, W. Yantao, C. Xinglai, and C. Hui, "Coordination control strategy between human vision and wheelchair manipulator based on BCI," in *2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1872-1875, 2016.
- [73] F. Achic, J. Montero, C. Penalozza, and F. Cuellar, "Hybrid BCI system to operate an electric wheelchair and a robotic arm for navigation and manipulation tasks," in *2016 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)*, pp. 249-254, 2016.
- [74] U. Qidwai, E. M. Hassan, R. M. A. Halabi, and M. Shakir, "Device interface for people with mobility impairment," in *2013 7th IEEE GCC Conference and Exhibition (GCC)*, pp. 177-181, 2013.
- [75] R. J. M. G. Tello, A. L. C. Bissoli, F. Ferrara, S. Müller, A. Ferreira, and T. F. Bastos-Filho, "Development of a Human Machine Interface for Control of Robotic Wheelchair and Smart Environment," *IFAC-PapersOnLine*, vol. 48, no. 19, pp. 136-141, 2015.
- [76] W. Jia, D. Huang, O. Bai, H. Pu, X. Luo, and X. Chen, "Reliable planning and execution of a human-robot cooperative system based on noninvasive brain-computer interface with uncertainty," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3798-3805, 2012.
- [77] J. Duan, Z. Li, C. Yang, and P. Xu, "Shared control of a brain-actuated intelligent wheelchair," in *Proceeding of the 11th World Congress on Intelligent Control and Automation*, pp. 341-346, 2014.

- [78] A. H. Aziz F, Mokhtar N, Mubin M, "HMM based automated wheelchair navigation using EOG traces in EEG," *Journal of Neural Engineering*, vol. 11, no. 5, pp. 1-16, 2014.
- [79] T. Carlson, R. Leeb, R. Chavarriaga, and J. d. R. Millán, "The birth of the brain-controlled wheelchair," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5444-5445, 2012.
- [80] B.-V. Ngo, T.-H. Nguyen, V.-T. Ngo, D.-K. Tran, and T.-D. Nguyen, "Wheelchair Navigation System using EEG Signal and 2D Map for Disabled and Elderly People," in *2020 5th International Conference on Green Technology and Sustainable Development (GTSD)*, pp. 219-223, 2020.
- [81] A. Poncela, C. Urdiales, E. J. Perez, and F. Sandoval, "A New Efficiency-Weighted Strategy for Continuous Human/Robot Cooperation in Navigation," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 3, pp. 486-500, 2009.
- [82] R. C. Simpson, S. P. Levine, D. A. Bell, L. A. Jaros, Y. Koren, and J. Borenstein, "NavChair: An assistive wheelchair navigation system with automatic adaptation," in *Assistive Technology and Artificial Intelligence: Applications in Robotics, User Interfaces and Natural Language Processing*, V. O. Mittal, H. A. Yanco, J. Aronis, and R. Simpson, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 235-255, 1998.
- [83] E. Demeester, A. Hüntemann, D. Vanhooydonck, G. Vanacker, H. Van Brussel, and M. Nuttin, "User-adapted plan recognition and user-adapted shared control: A Bayesian approach to semi-autonomous wheelchair driving," *Autonomous Robots*, vol. 24, no. 2, pp. 193-211, 2008.
- [84] C. Urdiales *et al.*, "A new multi-criteria optimization strategy for shared control in wheelchair assisted navigation," *Autonomous Robots*, vol. 30, no. 2, pp. 179-197, 2011.
- [85] C. Cadena *et al.*, "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309-1332, 2016.
- [86] Z. Jiang, J. Zhu, Y. Li, J. Wang, Z. Li, and H. Lu, "Simultaneous Merging Multiple Grid Maps Using the Robust Motion Averaging," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 3, pp. 655-668, 2019.
- [87] A. M. Santana, K. R. T. Aires, R. M. S. Veras, and A. A. D. Medeiros, "An Approach for 2D Visual Occupancy Grid Map Using Monocular Vision," *Electronic Notes in Theoretical Computer Science*, vol. 281, pp. 175-191, 2011.
- [88] D. Meger *et al.*, "Curious George: An attentive semantic robot," *Robotics and Autonomous Systems*, vol. 56, no. 6, pp. 503-511, 2008.
- [89] S. Pillai and J. J. Leonard, "Monocular SLAM Supported Object Recognition," in *Robotics: Science and Systems*, Rome, Italy, pp. 1-9, 2015.

- [90] J.-R. Ruiz-Sarmiento, C. Galindo, and J. Gonzalez-Jimenez, "Building Multiversal Semantic Maps for Mobile Robot Operation," *Knowledge-Based Systems*, vol. 119, pp. 257-272, 2017.
- [91] A. Pronobis and P. Jensfelt, "Large-scale semantic mapping and reasoning with heterogeneous modalities," in *2012 IEEE International Conference on Robotics and Automation*, pp. 3515-3522, 2012.
- [92] X. Qi *et al.*, "Building semantic grid maps for domestic robot navigation," *International Journal of Advanced Robotic Systems*, vol. 17, no. 1, pp. 1-12, 2020.
- [93] W. Seo and K.-R. Baek, "Indoor Dead Reckoning Localization Using Ultrasonic Anemometer with IMU," *Journal of Sensors*, vol. 2017, pp. 1-12, 2017.
- [94] H. Rashid and A. K. Turuk, "Dead reckoning localisation technique for mobile wireless sensor networks," *IET Wireless Sensor Systems*, vol. 5, no. 2, pp. 87-96, 2015.
- [95] G. N. Desouza and A. C. Kak, "Vision for mobile robot navigation: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 237-267, 2002.
- [96] X. Zhang, S. Zhu, Z. Wang, and Y. Li, "Hybrid visual natural landmark-based localization for indoor mobile robots," *International Journal of Advanced Robotic Systems*, vol. 15, no. 6, pp. 1-18, 2018.
- [97] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, pp. 511-518, 2001.
- [98] C. Wang, P. Wu, L. Yan, Z. Ye, H. Chen, and H. Ling, "Image classification based on principal component analysis optimized generative adversarial networks," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 9687-9701, 2021.
- [99] L. Gueguen and M. Pesaresi, "Multi scale Harris corner detector based on Differential Morphological Decomposition," *Pattern Recognition Letters*, vol. 32, no. 14, pp. 1714-1719, 2011.
- [100] J. Xu, H. w. Chang, S. Yang, and M. Wang, "Fast feature-based video stabilization without accumulative global motion estimation," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 3, pp. 993-999, 2012.
- [101] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.
- [102] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006*, Berlin, Heidelberg, pp. 404-417, 2006.
- [103] Y. Liu, H. Zhang, H. Guo, and N. N. Xiong, "A FAST-BRISK Feature Detector with Depth Information," *Sensors*, vol. 18, no. 11, pp. 1-19, 2018.

- [104] Y. Qian, H. Zheng, D. He, Z. Zhang, and Z. Zhang, "R-CNN Object Detection Inference With Deep Learning Accelerator," in *2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, pp. 297-302, 2018.
- [105] R. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1440-1448, 2015.
- [106] K. Desarda, S. Gulve, S. Hase, and R. Raut, "Single Shot Detector Based Object Detection For Visually Impaired People," in *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, pp. 1378-1385, 2022.
- [107] C. Liu, Y. Tao, J. Liang, K. Li, and Y. Chen, "Object Detection Based on YOLO Network," in *2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 799-803, 2018.
- [108] E. H. C. Harik and A. Korsaeht, "The Heading Weight Function: A Novel LiDAR-Based Local Planner for Nonholonomic Mobile Robots," *Sensors*, vol. 19, no. 16, pp. 1-17, 2019.
- [109] R. D. Phalgun Chintala, and Thomas Hanne, "Robotic Path Planning by Q Learning and a Performance Comparison with Classical Path Finding Algorithms," *International Journal of Mechanical Engineering and Robotics Research*, vol. 11, no. 6, pp. 373-378, 2022.
- [110] D. Foad, A. Ghifari, M. B. Kusuma, N. Hanafiah, and E. Gunawan, "A Systematic Literature Review of A* Pathfinding," *Procedia Computer Science*, vol. 179, pp. 507-514, 2021.
- [111] A. I. Panov, K. S. Yakovlev, and R. Suvorov, "Grid Path Planning with Deep Reinforcement Learning: Preliminary Results," *Procedia Computer Science*, vol. 123, pp. 347-353, 2018.
- [112] Y. Zhu *et al.*, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3357-3364, 2017.
- [113] P.-G. Ó *et al.*, "Deep Reinforcement Learning based control algorithms: Training and validation using the ROS Framework in CARLA Simulator for Self-Driving applications," in *2021 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1268-1273, 2021.
- [114] S. Wen, X. Chen, C. Ma, H. K. Lam, and S. Hua, "The Q-learning obstacle avoidance algorithm based on EKF-SLAM for NAO autonomous walking under unknown environments," *Robotics and Autonomous Systems*, vol. 72, pp. 29-36, 2015.
- [115] B.-V. Ngo and T.-H. Nguyen, "Positioning an electric wheelchair in 2D grid map based on natural landmarks for navigation using Q-Learning," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, no. 1, pp. 115-125, 2023.

- [116] H.-Y. Ryu, J.-S. Kwon, J.-H. Lim, A. H. Kim, S.-J. Baek, and J.-W. Kim, "Development of an Autonomous Driving Smart Wheelchair for the Physically Weak," *Applied Sciences*, vol. 12, no. 1, pp. 1-18, 2022.
- [117] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529-533, 2015.
- [118] A. K. Maddirala and K. C. Veluvolu, "Eye-blink artifact removal from single channel EEG with k-means and SSA," *Scientific Reports*, vol. 11, no. 1, p. 11043, 2021.
- [119] A. K. Maddirala and K. C. Veluvolu, "Eye-blink artifact removal from single channel EEG with k-means and SSA," *Scientific Reports*, vol. 11, no. 1, pp. 1-14, 2021.
- [120] A. A.-O. X. Maddirala and K. A.-O. Veluvolu, "SSA with CWT and k-Means for Eye-Blink Artifact Removal from Single-Channel EEG Signals," *Sensors*, vol. 22, no. 3, pp. 1-17, 2022.
- [121] K. F. L. Gonçalves, L. Kupssinskü, and M. Segalotto, "The effects of applying filters on EEG signals for classifying developers," *Journal of Applied Research and Technology*, vol. 19, no. 6, pp. 584-602, 2021.
- [122] H. Nguyen, C. Ngo, K. Truong Quang Dang, and V. T. Vo, "Temporal hemodynamic classification of two hands tapping using functional near— infrared spectroscopy," *Frontiers in Human Neuroscience*, vol. 7, pp. 1-12, 2013.
- [123] B.-V. Ngo, T.-H. Nguyen, D.-K. Tran, and D.-D. Vo, "Control of a Smart Electric Wheelchair Based on EEG Signal and Graphical User Interface for Disabled People," in *2021 International Conference on System Science and Engineering (ICSSE)*, pp. 257-262, 2021.
- [124] M. Rashida and M. A. Habib, "Quantitative EEG features and machine learning classifiers for eye-blink artifact detection: A comparative study," *Neuroscience Informatics*, vol. 3, no. 1, pp. 1-20, 2023.
- [125] B.-V. Ngo, T.-H. Nguyen, and T.-N. Nguyen, "EEG Signal-Based Eye Blink Classifier Using Convolutional Neural Network For BCI Systems," in *2021 15th International Conference on Advanced Computing and Applications (ACOMP)*, pp. 176-180, 2021.
- [126] B.-S. N. Kleifges K, Kerick SE and Robbins KA, "BLINKER: Automated Extraction of Ocular Indices from EEG Enabling Large-Scale Analysis," *Frontiers in Neuroscience*, vol. 11, no. 12, pp. 1-17, 2017.
- [127] M. Benda and I. Volosyak, "Peak Detection with Online Electroencephalography (EEG) Artifact Removal for Brain-Computer Interface (BCI) Purposes," *Brain Sci*, vol. 9, no. 12, pp. 1-21, 2019.
- [128] Y. Xu *et al.*, "Design and recognition of monocular visual artificial landmark based on arc angle information coding," in *2018 33rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 722-727, 2018.

- [129] D. Kartashov, A. Huletski, and K. Krinkin, "Fast artificial landmark detection for indoor mobile robots," in *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 209-214, 2015.
- [130] H. Wang and T. Ishimatsu, "Vision-Based Navigation for an Electric Wheelchair Using Ceiling Light Landmark," *Journal of Intelligent and Robotic Systems*, vol. 41, no. 4, pp. 283-314, 2005.
- [131] J. Vidal and C. Lin, "Simple and robust localization system using ceiling landmarks and infrared light," in *2016 12th IEEE International Conference on Control and Automation (ICCA)*, pp. 583-587, 2016.
- [132] L. A. V. Souto, A. Castro, L. M. G. Gonçalves, and T. P. Nascimento, "Stairs and Doors Recognition as Natural Landmarks Based on Clouds of 3D Edge-Points from RGB-D Sensors for Mobile Robot Localization," *Sensors*, vol. 17, no. 8, pp. 1-16, 2017.
- [133] X. Chai, F. Wen, and K. Yuan, "Fast vision-based object segmentation for natural landmark detection on Indoor Mobile Robot," in *2011 IEEE International Conference on Mechatronics and Automation*, pp. 2232-2237, 2011.
- [134] B.-V. Ngo, T.-H. Nguyen, and C.-C. Vu, "Automatic Labeling of Natural Landmarks for Wheelchair Motion Planning," *Electronics*, vol. 12, no. 14, pp. 1-20, 2023.
- [135] Y. Xie, Q. Wang, Y. Chang, and X. Zhang, "Fast Target Recognition Based on Improved ORB Feature," *Applied Sciences*, vol. 12, no. 2, pp. 1-14, 2022.
- [136] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *2011 International Conference on Computer Vision*, pp. 2564-2571, 2011.
- [137] B.-V. Ngo and T.-H. Nguyen, "Dense Feature-based Landmark Identification for Mobile Platform Localization," *International Journal of Computer Science and Network Security*, vol. 18, no. 12, pp. 186-200, 2018.
- [138] E. W. Richard, P.A.W, "Digital Image Processing," *4th ed.; Pearson India: Noida, India*, 2019.
- [139] I. Endres and D. Hoiem, "Category-Independent Object Proposals with Diverse Ranking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 2, pp. 222-234, 2014.
- [140] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, "Selective Search for Object Recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154-171, 2013.
- [141] P. Du and N. Bulusu, "An automated AR-based annotation tool for indoor navigation for visually impaired people," in *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*, Virtual Event, USA, pp. 1-4, 2021.
- [142] J. G. Apud Baca *et al.*, "Automated Data Annotation for 6-DoF AI-Based Navigation Algorithm Development," *Journal of Imaging*, vol. 7, no. 11, pp. 236-253, 2021.

- [143] O. Deane, E. Toth, and S.-H. Yeo, "Deep-SAGA: a deep-learning-based system for automatic gaze annotation from eye-tracking data," *Behavior Research Methods*, vol. 55, pp. 1372–1391, 2022.
- [144] I. García-Aguilar, J. García-González, R. M. Luque-Baena, and E. López-Rubio, "Automated labeling of training data for improved object detection in traffic videos by fine-tuned deep convolutional neural networks," *Pattern Recognition Letters*, vol. 167, pp. 45-52, 2023.
- [145] B.-V. Ngo and T.-H. Nguyen, "A Semi-Automatic Wheelchair with Navigation Based on Virtual-Real 2D Grid Maps and EEG Signals," *Applied Sciences*, vol. 12, pp. 1-20, 2022.
- [146] B.-V. Ngo, T.-H. Nguyen, and D.-D. Vo, "An MICP Algorithm for 3D Map Reconstruction Based on 3D Cloud Information of Landmarks," *European Journal of Engineering and Technology Research*, vol. 6, no. 3, pp. 130-138, 2021.
- [147] P. Cosimo, R. Vito, M. Nicola, S. Maria di, and N. Massimiliano, "A robust method for 2D occupancy map building for indoor robot navigation," in *Multimodal Sensing and Artificial Intelligence: Technologies and Applications II*, pp. 1-14, 2021.
- [148] N. T. H. Nguyen Tan Nhu, "Landmark-Based Robot Localization Using a Stereo Camera System," *American Journal of Signal Processing*, vol. 5, no. 2, pp. 40-50, 2015.

CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

Danh mục các kết quả nghiên cứu trong tạp chí và hội thảo đã được công bố trong quá trình thực hiện luận án:

Tạp chí quốc tế trong hệ thống ISI/Scopus

- [1] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Chi-Cuong Vu, " Automatic Labeling of Natural Landmarks for Wheelchair Motion Planning ", *Electronics*, vol. 12, pp. 1-20, 2023. (**SCIE-Q2-IF 2.9**)
- [2] **Ba-Viet Ngo**, Thanh-Hai Nguyen, "Positioning an electric wheelchair in 2D grid map based on natural landmarks for navigation using Q-Learning", *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 31, pp. 115-125, 2023. (**Scopus-Q3**)
- [3] **Ba-Viet Ngo**, Thanh-Hai Nguyen, "A Semi-Automatic Wheelchair with Navigation Based on Virtual-Real 2D Grid Maps and EEG Signals", *Applied Sciences*, vol.12, pp. 1-20, 2022. (**SCIE-Q2-IF 2.7**)
- [4] **Ba-Viet Ngo**, Thanh-Hai Nguyen, "Dense Feature-based Landmark Identification for Mobile Platform Localization", *International Journal of Computer Science and Network Security*, vol. 18, pp. 186-200, 2018 (**ESCI**)

Hội nghị quốc tế

- [5] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Thanh-Nghia Nguyen, "EEG Signal-Based Eye Blink Classifier Using Convolutional Neural Network For BCI Systems", in *Proceedings of 15th International Conference on Advanced Computing and Applications (ACOMP)*, pp. 176-180, 2022.
- [6] Dang-Khoa Tran; Thanh-Hai Nguyen; **Ba-Viet Ngo**; Minh-Ly Lai; Thi-My-Kieu Ong; Thi-Minh Vu, Van-Minh Nguyen, "Evaluation Between Visual And Auditory EEG Responses On Upper Limb Movement in The Temporal and Spectral Domains," in *Proceedings of 6th International Conference on Green Technology and Sustainable Development (GTSD)*, pp. 817-824, 2022.
- [7] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Dang-Khoa Tran, Duc-Dung Vo, "Control of a Smart Electric Wheelchair Based on EEG Signal and Graphical User Interface for Disabled People", in *Proceedings of International Conference on System Science and Engineering (ICSSE)*, pp. 257-262, 2021.
- [8] Dang-Khoa Tran, Thanh-Hai Nguyen, **Ba-Viet Ngo**, "Amplitude Thresholding of EEG Signals For Eye Blink and Saccade Detection", in *Proceedings of International Conference on System Science and Engineering (ICSSE)*, pp. 268-273, 2021.
- [9] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Van-Thuyen Ngo, Dang-Khoa Tran, Truong-Duy Nguyen, " Wheelchair Navigation System using EEG Signal and 2D Map for Disabled and Elderly People ", in *Proceedings of 5th International Conference on Green Technology and Sustainable Development (GTSD)*, pp. 219-223, 2020.
- [10] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Van-Thuyen Ngo, "Hands-free control of an electric wheelchair using face behaviors ", in *Proceedings of International Conference on System Science and Engineering (ICSSE)*, pp. 29-33, 2017.

Tạp chí quốc tế khác

- [11] **Ba-Viet Ngo**, Thanh-Hai Nguyen, Duc-Dung Vo, "An MICP Algorithm for 3D Map Reconstruction Based on 3D Cloud Information of Landmarks", *European Journal of Engineering and Technology Research*, vol. 6, pp. 130-138, 2021.
- [12] Thanh-Hai Nguyen, **Ba-Viet Ngo**, Thanh-Hai Quach, " Optimization of Transformation Matrix for 3D Cloud Mapping Using Sensor Fusion ", *American Journal of Signal Processing*, vol. 8, pp. 9-19, 2018.