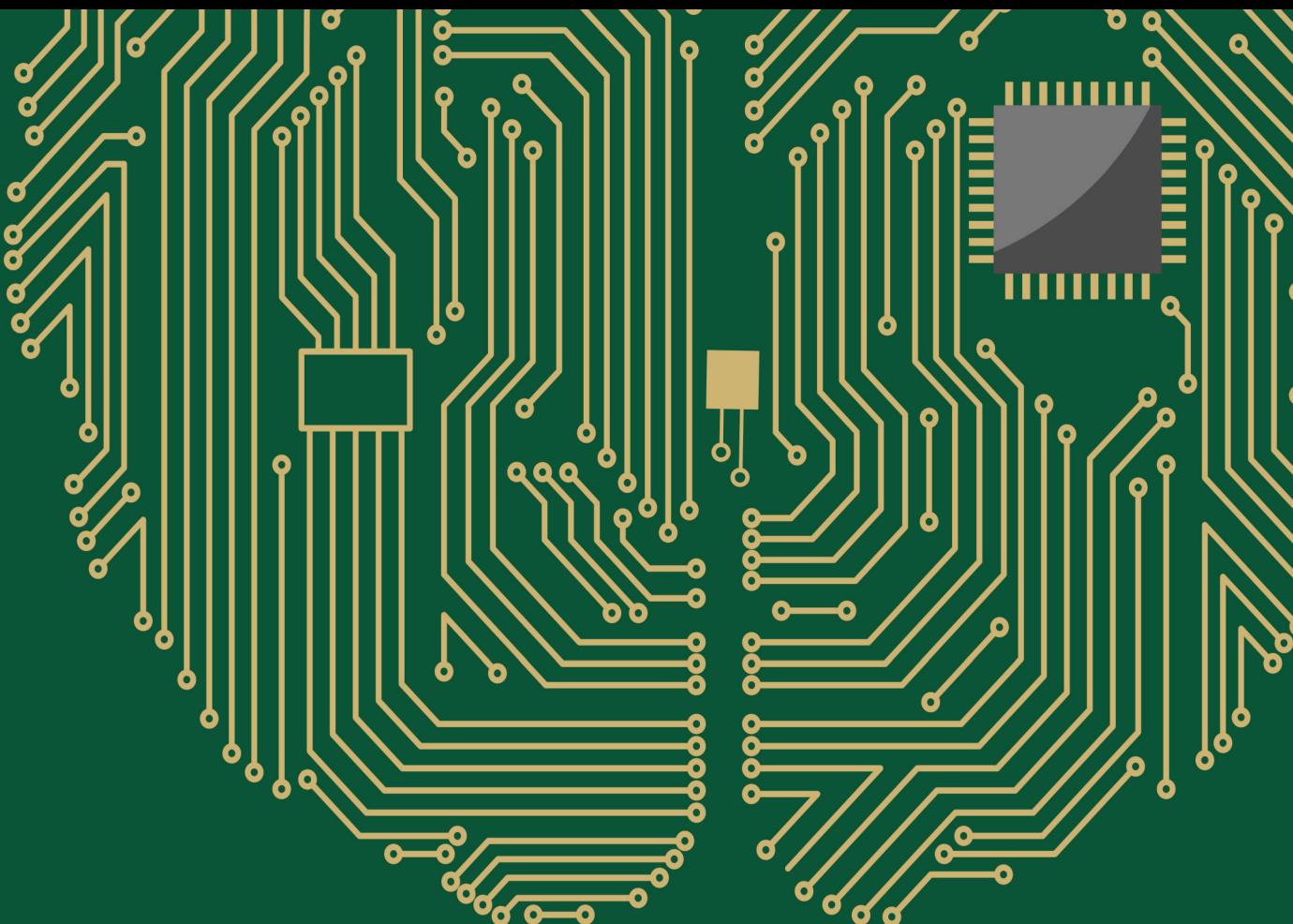


Ergonomic Issues in Brain-Computer Interface Technologies: Current Status, Challenges, and Future Direction

Lead Guest Editor: Hyun J. Baek

Guest Editors: Hyun S. Kim, Minkyu Ahn, Hohyun Cho, and Sangtae Ahn

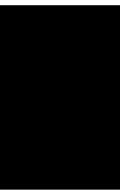


Ergonomic Issues in Brain-Computer Interface Technologies: Current Status, Challenges, and Future Direction

Ergonomic Issues in Brain-Computer Interface Technologies: Current Status, Challenges, and Future Direction

Lead Guest Editor: Hyun J. Baek

Guest Editors: Hyun S. Kim, Minkyu Ahn, Hohyun Cho, and Sangtae Ahn



Copyright © 2020 Hindawi Limited. All rights reserved.

This is a special issue published in "Computational Intelligence and Neuroscience." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chief Editor

Andrzej Cichocki, Japan

Editorial Board

- Ricardo Aler, Spain
Amparo Alonso-Betanzos, Spain
Pietro Aricò, Italy
Hasan Ayaz, USA
Sylvain Baillet, Canada
Theodore W. Berger, USA
Daniele Bibbo, Italy
Vince D. Calhoun, USA
Francesco Camasta, Italy
Michela Chiappalone, Italy
Jens Christian Claussen, United Kingdom
Silvia Conforto, Italy
Justin Dauwels, Singapore
Christian W. Dawson, United Kingdom
Carmen De Maio, Italy
Sergio Decherchi, Italy
Paolo Del Giudice, Italy
Maria Jose del Jesus, Spain
Arnaud Delorme, France
Thomas DeMarse, USA
Anastasios D. Doulamis, Greece
António Dourado, Portugal
Steven L. Fernandes, USA
Juan Carlos Fernández, Spain
Piotr Franaszczuk, USA
Leonardo Franco, Spain
Paolo Gastaldo, Italy
Samanwoy Ghosh-Dastidar, USA
Manuel Graña, Spain
Rodolfo E. Haber, Spain
Dominic Heger, Germany
Stephen Helms Tillery, USA
José Alfredo Hernández-Pérez, Mexico
Luis Javier Herrera, Spain
Etienne Hugues, USA
Ryotaro Kamimura, Japan
Pasi A. Karjalainen, Finland
Elpida Keravnou, Cyprus
Raşit Köker, Turkey
Dean J. Krusinski, USA
Fabio La Foresta, Italy
Antonino Laudani, Italy
Maciej Lawrynczuk, Poland
Mikhail A. Lebedev, USA
Cheng-Jian Lin, Taiwan
Giosuè Lo Bosco, Italy
Ezequiel López-Rubio, Spain
Bruce J. MacLennan, USA
Reinoud Maex, Belgium
Kezhi Mao, Singapore
Sergio Martinoia, Italy
Laura Marzetti, Italy
Elio Masciari, Italy
Paolo Massobrio, Italy
Gerard McKee, Nigeria
Michele Migliore, Italy
Paulo Moura Oliveira, Portugal
Debajyoti Mukhopadhyay, India
Klaus Obermayer, Germany
Karim G. Oweiss, USA
Massimo Panella, Italy
Fivos Panetsos, Spain
David M Powers, Australia
Sandhya Samarasinghe, New Zealand
Saeid Sanei, United Kingdom
Friedhelm Schwenker, Germany
Victor R. L. Shen, Taiwan
Fabio Solari, Italy
Jussi Tohka, Finland
Carlos M. Travieso-González, Spain
Lefteri Tsoukalas, USA
Pablo Varona, Spain
Roberto A. Vazquez, Mexico
Meel Velliste, USA
Mario Versaci, Italy
Francois B. Vialatte, France
Thomas Villmann, Germany
Ivan Volosyak, Germany
Cornelio Yáñez-Márquez, Mexico
Michał Zochowski, USA
Rodolfo Zunino, Italy

Contents

Ergonomic Issues in Brain-Computer Interface Technologies: Current Status, Challenges, and Future Direction

Hyun Jae Baek , Hyun Seok Kim , Minkyu Ahn, Hohyun Cho , and Sangtae Ahn 
Editorial (2 pages), Article ID 4876397, Volume 2020 (2020)

Evaluating a Semiautonomous Brain-Computer Interface Based on Conformal Geometric Algebra and Artificial Vision

Mauricio Adolfo Ramírez-Moreno and David Gutiérrez 
Research Article (19 pages), Article ID 9374802, Volume 2019 (2019)

Advances in Hybrid Brain-Computer Interfaces: Principles, Design, and Applications

Zina Li, Shuqing Zhang, and Jiahui Pan 
Review Article (9 pages), Article ID 3807670, Volume 2019 (2019)

Driving Fatigue Detection from EEG Using a Modified PCANet Method

Yuliang Ma , Bin Chen, Rihui Li, Chushan Wang, Jun Wang, Qingshan She , Zhizeng Luo , and Yingchun Zhang 
Research Article (9 pages), Article ID 4721863, Volume 2019 (2019)

Covert Intention to Answer “Yes” or “No” Can Be Decoded from Single-Trial Electroencephalograms (EEGs)

Jeong Woo Choi  and Kyung Hwan Kim 
Research Article (11 pages), Article ID 4259369, Volume 2019 (2019)

Comparison of Visual Stimuli for Steady-State Visual Evoked Potential-Based Brain-Computer Interfaces in Virtual Reality Environment in terms of Classification Accuracy and Visual Comfort

Kang-min Choi , Seonghun Park , and Chang-Hwan Im 
Research Article (7 pages), Article ID 9680697, Volume 2019 (2019)

Impact of Speller Size on a Visual P300 Brain-Computer Interface (BCI) System under Two Conditions of Constraint for Eye Movement

R. Ron-Angevin , L. Garcia, Á. Fernández-Rodríguez, J. Saracco , J. M. André, and V. Lespinet-Najib
Research Article (16 pages), Article ID 7876248, Volume 2019 (2019)

Enhancing the Usability of Brain-Computer Interface Systems

Hyun Jae Baek , Min Hye Chang , Jeong Heo , and Kwang Suk Park 
Review Article (12 pages), Article ID 5427154, Volume 2019 (2019)

Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning

Awwal Muhammad Dawud , Kamil Yurtkan , and Huseyin Oztoprak
Research Article (12 pages), Article ID 4629859, Volume 2019 (2019)

**An Optimized Channel Selection Method Based on Multifrequency CSP-Rank for Motor Imagery-Based
BCI System**

Jian Kui Feng, Jing Jin , Ian Daly, Jiale Zhou, Yugang Niu, Xingyu Wang, and Andrzej Cichocki
Research Article (10 pages), Article ID 8068357, Volume 2019 (2019)

Editorial

Ergonomic Issues in Brain-Computer Interface Technologies: Current Status, Challenges, and Future Direction

Hyun Jae Baek ,¹ Hyun Seok Kim ,² Minkyu Ahn,³ Hohyun Cho ,⁴ and Sangtae Ahn ,⁵

¹Department of Medical and Mechatronics Engineering, Soonchunhyang University, Asan, Republic of Korea

²Institute for Brain and Cognitive Engineering, Korea University, Seoul 02841, Republic of Korea

³School of Computer Science and Electrical Engineering, Handong Global University, Pohang, Republic of Korea

⁴National Center for Adaptive Neurotechnologies, Department of Neuroscience and Experimental Therapeutics, Albany Medical College, Albany, NY, USA

⁵Department of Psychiatry, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Correspondence should be addressed to Hyun Jae Baek; hjbaek@sch.ac.kr

Received 15 January 2020; Accepted 16 January 2020; Published 7 February 2020

Copyright © 2020 Hyun Jae Baek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The brain-computer interface (BCI) is defined as a computer technology that can interact with the nervous system by decoding information from neural activity. It is an increasingly becoming reliable piece of technology that involves the brain to sense, analyze, and translate signals into commands for the purpose of providing a human computer interaction that will change our life, particularly of patients who suffer from paralysis or similar neurological conditions. BCI enables the communication with computers and controls of external devices such as robotic agents.

Since BCI research emerged a few decades ago, it has mainly focused on signal processing methods to increase its accuracy for its applications such as virtual keyboards, wheelchair controls, or rehabilitation training. BCI is now maturing towards being more realistic and practically plausible; therefore, it is time to take ergonomics (or human factors) into account as part of the BCI design processes. In addition to improving BCI system performance such as accuracy, brain signal measurements should be more convenient and easier, and the BCI paradigm including audio or visual stimulation should not make the user tired.

The goal of this special issue is to share cutting-edge research and applications on user-friendly BCI solutions such as new concept paradigms, innovative sensors, devices, and signal-processing algorithms. To achieve this goal, the editorial team focused on the core technologies that could contribute to the implementation of possible future BCI

solutions and identified the nine representative manuscripts submitted to the special issue.

This special issue includes 2 review papers and 7 research papers on state-of-the-art BCI technologies that are being studied and developed for enhancing user acceptance. In the review article entitled “Enhancing the Usability of Brain-Computer Interface Systems,” H. J. Baek et al. investigated the EEG electrode technologies for wearable measurements. They also showed new paradigms for minimizing visual or auditory fatigue in BCI process. In the review article entitled “Advances in Hybrid Brain-Computer Interfaces: Principles, Design, and Applications,” Z. Li et al. discussed the research progress of hybrid BCI to improve BCI performance and achieve multifunctional control. In the article entitled “Evaluating a Semiautonomous Brain-Computer Interface Based on Conformal Geometric Algebra and Artificial Vision,” M. A. Ramírez-Moreno and D. Gutiérrez proposed a semiautonomous approach based on a conformal geometric algebra model that solves the inverse kinematics of the robot on the fly to control a robotic arm with less mental fatigue. In the article entitled “Driving Fatigue Detection from EEG Using a Modified PCANet Method,” Y. Ma et al. presented a novel feature extraction strategy developed by integrating the principle component analysis and a deep learning model to achieve high classification accuracy and efficiency in using EEG for driving fatigue detection. In the article entitled “Covert Intention to Answer “Yes” or “No” Can Be Decoded from Single-Trial Electroencephalograms (EEGs),” J. W.

Choi and K. H. Kim showed possibility to decode binary intentions from multichannel single-trial EEGs while covertly answering to self-referential questions with either “yes” or “no.” In the article entitled “Comparison of Visual Stimuli for Steady-State Visual Evoked Potential-Based Brain-Computer Interfaces in Virtual Reality Environment in terms of Classification Accuracy and Visual Comfort,” K. Choi et al. described the importance of an appropriate visual stimulus to enhance the overall performance of the steady-state visual evoked potential-based BCIs in virtual reality (VR) environments. In the article entitled “Impact of Speller Size on a Visual P300 Brain-Computer Interface (BCI) System under Two Conditions of Constraint for Eye Movement,” R. Ron-Angevin et al. investigated the effect of speller size on P300-based BCI usability, measured in terms of effectiveness, efficiency, and satisfaction under overt and covert attention conditions. In the article entitled “Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning,” A. M. Dawud et al. proposed deep learning approach to solve the problem of identifying brain haemorrhage from brain computer tomography images. In the article entitled “An Optimized Channel Selection Method Based on Multifrequency CSP-Rank for Motor Imagery-Based BCI System,” J. K. Feng et al. proposed a common spatial pattern rank channel selection method for multifrequency band EEG to prevent deteriorated BCI performance by redundant information.

Conflicts of Interest

The editors declare that there are no conflicts of interest regarding the publication of this issue.

Acknowledgments

The editors would like to thank all the authors and reviewers who made this special issue available. They hope that this collection of articles will be useful for the industrialization of BCI research and related technologies.

*Hyun Jae Baek
Hyun Seok Kim
Minkyu Ahn
Hohyun Cho
Sangtae Ahn*

Research Article

Evaluating a Semiautonomous Brain-Computer Interface Based on Conformal Geometric Algebra and Artificial Vision

Mauricio Adolfo Ramírez-Moreno and David Gutiérrez 

Centro de Investigación y de Estudios Avanzados (Cinvestav), Unidad Monterrey, Apodaca, Nuevo Leon 66600, Mexico

Correspondence should be addressed to David Gutiérrez; dgtz@ieee.org

Received 10 June 2019; Accepted 30 October 2019; Published 27 November 2019

Guest Editor: Hyun S. Kim

Copyright © 2019 Mauricio Adolfo Ramírez-Moreno and David Gutiérrez. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we evaluate a semiautonomous brain-computer interface (BCI) for manipulation tasks. In such a system, the user controls a robotic arm through motor imagery commands. In traditional process-control BCI systems, the user has to provide those commands continuously in order to manipulate the effector of the robot step-by-step, which results in a tiresome process for simple tasks such as pick and replace an item from a surface. Here, we take a semiautonomous approach based on a conformal geometric algebra model that solves the inverse kinematics of the robot on the fly, and then the user only has to decide on the start of the movement and the final position of the effector (goal-selection approach). Under these conditions, we implemented pick-and-place tasks with a disk as an item and two target areas placed on the table at arbitrary positions. An artificial vision (AV) algorithm was used to obtain the positions of the items expressed in the robot frame through images captured with a webcam. Then, the AV algorithm is integrated into the inverse kinematics model to perform the manipulation tasks. As proof-of-concept, different users were trained to control the pick-and-place tasks through the process-control and semiautonomous goal-selection approaches so that the performance of both schemes could be compared. Our results show the superiority in performance of the semiautonomous approach as well as evidence of less mental fatigue with it.

1. Introduction

A brain-computer interface (BCI) is a system that enables a real-time user-device communication pathway through brain activity. Through the years, research and development on BCI has mainly been oriented to the creation of rehabilitation systems as well as systems that help disabled patients regain to some extent their lost or diminished capabilities [1]. Some reported devices that have been successfully controlled using BCIs are spellers, electric wheelchairs, robotic arms, electric prostheses, and humanoid robots [2–5]. In BCI studies, the most common technique used to acquire brain noninvasively is electroencephalography (EEG).

In order to manipulate the device through brain activity, the design of the BCI must include the following stages: signal acquisition, filtering, feature extraction, classification, device modeling, and control [6]. During the filtering stage,

unwanted noise and artifacts are removed from the signals using temporal and spatial filters. Then, temporal or spatial features of interest are extracted from the signals to build feature vectors. These vectors are formed by characteristic components of the signals, which are then used in the classification stage to decipher user intention. Lastly, the device is manipulated based on the result of the classification algorithm. Depending on the device and the complexity of the system, a model of the system is needed to perform with precision the desired tasks. BCIs can be divided into two groups based on their control strategy: *process control* and *goal selection*. In the process-control strategy, users are continuously controlling each part of the process. This is done by performing low-level commands continuously through the BCI, with no additional assistance. On the other hand, in the goal-selection strategy, users are responsible for selecting their desired goal and the system provides assistance to successfully perform the tasks with minimum effort

[7]. In this case, the user performs high-level tasks by sending simple commands through the BCI.

Common paradigms used as control commands in BCI include steady-state visual evoked potentials (SSVEPs), P300 waveform, and motor imagery (MI). SSVEP is a resonance phenomenon occurring at occipital and parietal lobes as a result of oscillatory visual stimulus presented to a user at a constant frequency [8]. The P300 is an EEG signal component that appears 300 ms after an event of voluntary attention, and it is usually observed during visual or auditory stimulus presentation [9]. MI presents as an event-related desynchronization (ERD) found at the sensorimotor areas, which generates a contralateral power decrease in a frequency range from 8–13 Hz (also known as the μ band) [10]. Controlling a BCI with SSVEP and P300 requires less training in comparison to MI, as the first represents an involuntary response to a stimulus. However, its use in BCI is limited due to its requirement of a stimulus presentation device. The training process to control MI-based BCIs (MI-BCI) might involve stimulus presentation as well. However, it can be excluded for its final application on the BCI. Even though MI-BCIs require longer training periods, they are better suited for close-to-real-life environments and self-paced BCIs [11].

Several studies present successfully implemented ERD-based BCIs, most of them using a process-control strategy [12–14]. Some goal-selection BCIs have been reported as well [15, 16]. In [7], users were trained on process-control and goal-selection MI-BCIs to perform one-dimensional cursor movements on a screen. The results suggest that users performing on goal-selection strategy showed higher accuracy and faster learning in comparison to the process-control approach. However, the authors state that a direct comparison of goal selection and process control in a more complicated (real-world) scenario has not yet been presented. In the proposed study, three-dimensional object manipulation tasks through a robotic arm are implemented in a MI-BCI. The complexity of the three-dimensional movements on real objects is higher than the one-dimensional movements on virtual objects presented in [7]. In [17], a semiautonomous BCI is implemented to manipulate a robotic arm to perform tasks such as pouring a beverage inside a glass in a tray, through SSVEP. In future research, similar tasks as in [17] could be implemented in our BCI using MI instead, allowing a more natural execution of daily-life context tasks without the need of a stimulus presentation screen.

In a typical process-control MI-BCI, the user controls the direction of the final effector of a robotic arm through low-level commands, which means that the user has to maneuver the robot in a three-dimensional space to reach for a desired target. Clearly, the user remains in a high attention state during the maneuvers, as he/she is continuously aware of the final effector position during the whole task. This continuous awareness might lead to generation of mental fatigue or frustration, which is undesirable as it can directly affect user performance and learning [18]. The analysis of P300 features, such as amplitude and latency, has been shown to be useful in identifying the depth of cognitive

information processing [19]. The amplitude of P300 waveform tends to decrease when users encounter cognitive tasks with high difficulty [20]. On the other hand, P300 latency has shown to increase when the stimulus is cognitively difficult to process [21]. Another study has reported correlation between changes in the P300 component and BCI performance [22]. The evidence provided by these studies might suggest that the analysis of P300 could be implemented as a mental fatigue indicator during BCI training and control.

In order to diminish mental fatiguing in BCI systems, a semiautonomous BCI using a goal-selection strategy is here proposed. This system assists the user to perform a specific task by calculating all the variables needed to successfully execute it. Some studies have previously presented BCI designs focusing on this semiautonomous approach with successful results on performance, accuracy, and comfort for the user [17, 23, 24]. Therefore, this paper presents the implementation of a traditional low-level MI-BCI and a semiautonomous MI-BCI designed to perform object manipulation tasks with a robotic arm. In the process-control strategy MI-BCI, the user commands the final effector of the robot to move in a three-dimensional space to reach for a target placed on a table. In the semiautonomous MI-BCI, one small disk and two target areas are placed on a table. Here, the robot reaches for the disk and places it on a specific target, which is selected by the user. As proof-of-concept, two volunteers were trained on each BCI system, and their performance was evaluated and compared. A statistical P300 analysis was performed on all users in order to observe mental fatigue differences induced by the operation of low-level and semiautonomous BCIs.

In order to model the robot used in this experiment, a conformal geometric algebra (CGA) model was implemented in both the traditional and semiautonomous BCIs to solve the inverse kinematics of the robotic arm, i.e., obtaining the joint angles needed for a specific position of the final effector. Additionally, an artificial vision (AV) algorithm was integrated into the semiautonomous BCI in order to provide information about the positions of the items on the table referenced to robot frame. As the implementation of the semiautonomous BCI implies a higher computational load, the CGA model was chosen for the solution of the inverse kinematics. CGA has shown to represent an operation reduction and in some cases, a decrease in computational load when compared to traditional inverse kinematics solution [25].

This paper is organized as follows. The CGA model and AV algorithm are described in Section 2, and the design of both BCIs is explained in Section 3. Evaluations on both algorithms and performance results of users controlling both BCIs are presented in Section 4. Preliminary short reports of the system's implementation (but not its evaluation) have been presented in [26] and [27].

2. Robot Modeling and Artificial Vision

In this section, we describe each of the components required to compute the inverse kinematics of a robotic arm by using CGA. Furthermore, here we explain in detail the AV

algorithm used to obtain the positions of the objects to be manipulated by the robot.

2.1. Conformal Geometric Algebra. Traditional methods to solve the inverse kinematics of robots include several matrix operations as well as many trigonometric expressions. All this can result in a quite complex solution depending on the modeled robot [28]. In this study, a conformal geometric algebra (CGA) model is proposed instead, as it is considered to be computationally lighter, easier to implement, and highly intuitive. CGA has proved to be a powerful tool when solving the inverse kinematics of robotic arms [29, 30]. It also offers an operation reduction when compared to traditional methods and provides efficient runtime solutions. More information on computational efficiency characteristics can be found in [31].

With this model, the joint angles of the robot are obtained for a specific position of the final effector. In CGA, two new dimensions (e_0, e_∞) are defined, representing a point in the origin and a point in the infinity, respectively, in addition to the three-dimensional Euclidean space (e_1, e_2, e_3) [29]. In this space, geometric entities (points, lines, circles, planes, and spheres) and calculations involving them (distances and intersections) can be represented with simple algebraic equations.

Also, the geometric product between two vectors a and b is defined as a combination of the inner product and the outer product:

$$ab = a \cdot b + a \wedge b. \quad (1)$$

The inner product is used to calculate distances between elements, and the outer product generates a bivector, which is an element occupying the space spanned by both vectors. It is also used to find the intersection between two elements. The intersection M of two geometric objects A and B represented in CGA is given by $M^* = A^* \wedge B^*$ or $M^* = A^* \cdot B$. The element A^* is the duality of A and is expressed as

$$A^* = AI_c^{-1}, \quad (2)$$

where $I_c^{-1} = e_0e_3e_2e_1e_\infty$, which allows for a change in representation of the same element. Standard and dual representations of commonly used geometrical objects in CGA are shown in Table 1. There, x and n are points represented as a linear combination of the 3D base vectors:

$$x = x_1e_1 + x_2e_2 + x_3e_3. \quad (3)$$

There are two possible representations of the same element, as shown in Table 1. A circle can be represented as the space spanned by three points in space as well as the intersection of two spheres. Also, a line can be expressed as the intersection of two planes as well as the space spanned by two points expanded to the infinity.

Making use of the previous equations and relationships, a CGA model to solve the inverse kinematics of a manipulator robot was obtained following the proposed method in [32]. The modeled robot was the Dynamixel AX-18A Smart robotic arm, which is a five-degree-of-freedom (5-DOF)

TABLE 1: Representations of the conformal geometric entities.

Entity	Standard	Dual
Point	$P = x + (1/2)x^2e_\infty + e_0$	
Point pair	$Pp = s_1 \wedge s_2 \wedge s_3$	$Pp^* = x_1 \wedge x_2$
Line	$l = \pi_1 \wedge \pi_2$	$l^* = x_1 \wedge x_2 \wedge e_\infty$
Circle	$c = s_1 \wedge s_2$	$c^* = x_1 \wedge x_2 \wedge x_3$
Sphere	$s = P - (1/2)r^2e_\infty$	$s^* = x_1 \wedge x_2 \wedge x_3 \wedge x_4$

manipulator robot. Figure 1 shows the modeled robot as well as its joints and links. The DOF of this robot corresponds to its shoulder rotation, elbow flexion-extension, wrist flexion-extension, wrist rotation, and hand open-close function [33]. The inverse kinematics solution was obtained for joints J_0 , J_2 , and J_3 . For the particularities of the manipulation tasks, joints J_4 and J_5 were not considered for simplicity.

2.2. Our CGA Model. Next, we describe the required CGA model that we implemented specifically for our system.

2.2.1. Fixed Joints and Planes. The origin of the CGA model was located at joint J_0 , located at the center of the rotational base of the robot; therefore, $J_0 = e_0$. Joint J_1 is also a fixed joint with constant position, found directly above joint J_0 . The position for joint J_1 was defined as $x_1 = [0, 0, 0.036]$. Now, let us consider the desired final effector position as a point in space x_e . Then, a vertical plane π_e representing the direction of the final effector is described as

$$\pi_e = e_0 \wedge e_3 \wedge x_e \wedge e_\infty, \quad (4)$$

where e_0 represents the origin in robot frame and e_3 the Euclidian z axis. As the position of the final effector is used to define π_e , the direction of plane changes consistently with x_e . A plane π_b , representing the rotational base of the robot, is defined as

$$\pi_b = e_0 \wedge e_1 \wedge e_2 \wedge e_\infty, \quad (5)$$

where e_1 and e_2 represent the Euclidean x and y axes. Planes π_e and π_b are shown in Figure 2.

2.2.2. Calculation of Joint's Position. In a kinematic chain model of a robotic arm using CGA, the implemented method to find joint J_n is based on the intersection of two spheres centered at joints J_{n-1} and J_{n+1} with radii equal to the lengths of the links connecting J_{n-1} with J_n and J_n with J_{n+1} , respectively. The intersection of both spheres results in a circle, which is then intersected to the plane of the final effector to obtain a point pair representing two possible configurations for joint J_n . One point is then selected as J_n , depending on the desired configuration. The process requires the following:

- (i) Spheres centered at point P with radius r are given by

$$s = P - \frac{1}{2}r^2e_\infty. \quad (6)$$

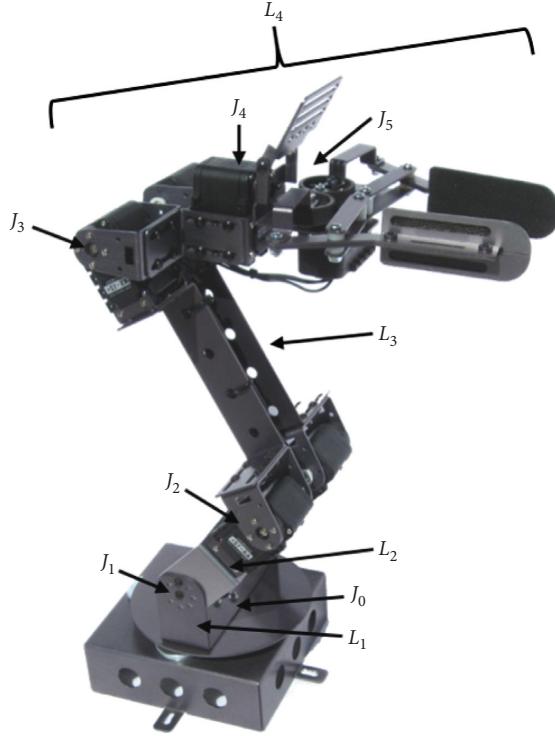


FIGURE 1: Joints and links of our 5-DOF Dynamixel AX-18A Smart robotic arm.

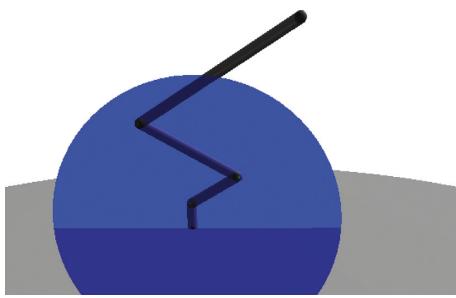


FIGURE 2: Planes π_e and π_b representing the orientation of the final effector and the robot base, respectively.

- (ii) There are two methods for creating a circle. We can either intersect two spheres s_j and s_k by

$$c = (s_j^* \wedge s_k^*)^*, \quad (7)$$

or we can intersect a plane π and a sphere s by

$$c = (\pi^* \wedge s^*)^*. \quad (8)$$

- (iii) The intersection of a circle c and a plane π to create a point pair Pp is given by

$$Pp = (c^* \wedge \pi^*)^*. \quad (9)$$

- (iv) Finally, to obtain a point P from Pp , we have

$$P = \frac{Pp \pm \sqrt{Pp^2}}{-e_\infty \cdot Pp}. \quad (10)$$

Based on the previous expressions, and in order to find the position of joint J_2 in our modeled robot, two spheres must be constructed, and they have to be centered at J_1 and J_3 . However, the position of joint J_3 is yet unknown in our model. A similar situation occurs if the desired position is instead joint J_3 . In this particular case, x_e is known but not J_2 . Given this situation, another approach was implemented in order to find joint J_2 .

2.2.3. Position of Joint J_2 . Using (6), sphere s_1 was centered at x_1 with radius equal to the length of link L_2 . Hence, in order to find joint J_2 , another sphere s_h must be intersected to s_1 . In order to construct s_h , its center must be defined. This is achieved by first creating an auxiliary sphere s_0 , centered at the origin with radius L_a equal to the horizontal component of the distance from J_0 to J_2 . This is valid as the distance from J_0 to J_2 is constant for any position of the final effector x_e .

Then, using (8), s_0 is intersected to plane π_e to obtain circle c_0 . Next, using (9), c_0 is intersected to plane π_b to produce point pair Pp_0 , from which one point is selected as x_h using (10). The procedure to find point x_h , which corresponds to the center of the desired sphere to be intersected with s_1 , is shown in Figure 3.

Using (6), sphere s_h is centered at x_h with radius L_b equal to the vertical component of the distance from J_0 to J_2 . Then, the intersection of spheres s_1 and s_h is given by (7), which results in circle c_2 . Using (9), the intersection of c_2 with plane π_e renders point pair Pp_2 . Finally, the position of J_2 is obtained from Pp_2 given by (10). The whole procedure previously detailed to obtain the position of joint J_2 is represented in Figure 4.

2.2.4. Position of Joint J_3 . The procedure to find the position of joint J_3 is straight forward once the position of joint J_2 is calculated. For that, two spheres s_2 and s_e are defined using (6), centered at x_2 and x_e , with radii equal to the length of links L_3 and L_4 , respectively. Both spheres are intersected to obtain circle c_3 using (7). With (9), c_3 is then intersected to plane π_e to obtain point pair Pp_3 . From Pp_3 , J_3 is easily obtained using (10). A representation of the procedure to find joint J_3 is shown in Figure 5.

2.2.5. Angle Calculation. In order to calculate the angles formed by two vectors α and β , their corresponding unit vectors are defined as $\hat{\alpha} = \alpha / \|\alpha\|$ and $\hat{\beta} = \beta / \|\beta\|$. The normalized bivector spanning the space formed by those vectors is expressed as

$$\hat{N} = \pm \frac{\hat{\alpha} \wedge \hat{\beta}}{\|\hat{\alpha} \wedge \hat{\beta}\|}. \quad (11)$$

As explained in [32], the angle θ between α and β is given by

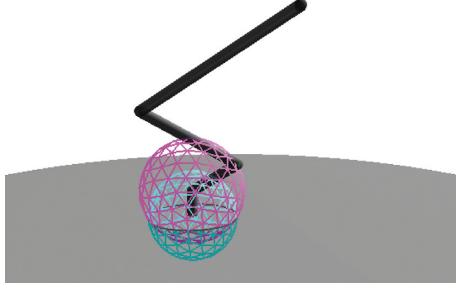


FIGURE 3: The intersection of spheres s_0 (bottom) and s_1 (top) define the circle c_0 , from which we find the position of point x_e .

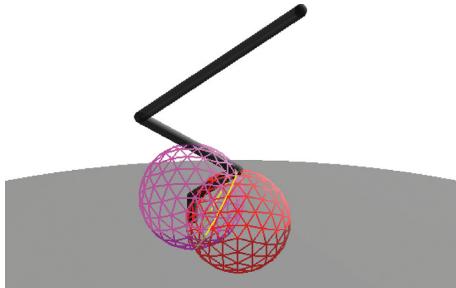


FIGURE 4: The intersection of spheres s_1 (left) and s_h (right) define the circle c_2 , from which we find the position of joint J_2 .

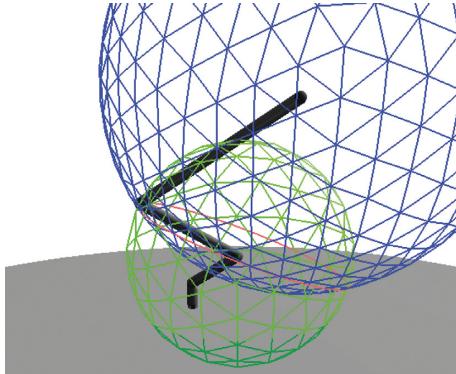


FIGURE 5: The intersection of spheres s_2 (bottom) and s_e (top) define the circle c_3 , from which it is possible to find the position of joint J_3 .

$$\theta = A \tan 2 \left[\frac{(\alpha \wedge \beta)}{\hat{N}}, \alpha \cdot \beta \right], \quad (12)$$

where $A \tan 2$ corresponds to the *four-quadrant* inverse tangent. This operator gathers information on the signs of its two arguments in order to return the appropriate quadrant of the computed angle [34]. Such result is not possible to be obtained from the conventional single-argument arctan function. Also, note that the plus sign in (11) applies if the rotation from α to β is counter-clockwise, while the minus sign applies in the opposite rotation.

In order to find the joint angles using (12), vectors formed by the links of the robot need to be calculated. First, lines representing each link are defined:

$$\begin{aligned} l_{01} &= e_0 \wedge J_1 \wedge e_\infty, \\ l_{12} &= J_1 \wedge J_2 \wedge e_\infty, \\ l_{23} &= J_2 \wedge J_3 \wedge e_\infty, \\ l_{3e} &= J_3 \wedge J_e \wedge e_\infty. \end{aligned} \quad (13)$$

The previous expressions define lines passing through links L_1 , L_2 , L_3 , and L_4 , respectively (see Figure 1). L_4 was considered a straight line from joint J_3 to the final effector x_e , i.e., we ignored wrist rotation and hand open-close joints.

In (12), the parameters α and β need to be *directional* vectors for the purpose of computing our joint angles. Therefore, the directional vectors of plane π_e , as well as lines l_{23} and l_{3e} , were calculated, which represent the base and links of the robot, respectively. From a given line l , its directional vector can be obtained as

$$(l \cdot e_0) \cdot e_\infty, \quad (14)$$

and the directional vector normal to a plane π is given by

$$(\pi^* \wedge e_\infty) \cdot e_0. \quad (15)$$

Based on all the previously defined elements, the vectors involved in the calculation of joint angles θ_k , for $k = 0, 2, 3$, are summarized in Table 2. Then, α and β in (11) and (12) are replaced by α_k and β_k , respectively, to calculate θ_k . Note that, as joint J_1 is fixed, θ_1 does not need to be calculated.

2.3. Artificial Vision Algorithm. An AV algorithm was implemented to calculate the positions of items on a table, so the robotic arm could perform the desired manipulation tasks. An ATW-1200 Acteck web camera was used to record images at 30 fps with a resolution of 640×480 pixels. The acquired images were processed and analyzed in real time using the OpenCV library (<https://www.opencv.org>) from Python.

The robotic arm was fixed on a white table, centered at one end of it. A plane was delimited on the table, defined as $400 \times 400 \text{ mm}^2$. Four $30 \times 30 \text{ mm}^2$ markers of different colors (cyan, orange, magenta, and yellow) were placed inside the delimited square, one at each corner. A blue disk with height of 6 mm and radius of 13 mm was used as the item to be picked, while two stickers with radius of 42 mm (green and red) were used to indicate target areas. The camera was fixed in a high angle so that all markers and items were inside its field of view. The setup of the robotic arm and items in the table are shown in Figure 6.

In order to perform object manipulation tasks, the *real-world* coordinates of the plane (in reference to robot frame) had to be obtained from the *image* coordinates obtained by the camera. To achieve this, a homography transformation was performed on the acquired images. In general, a two-dimensional point (u, v) in an image can be represented as a three-dimensional vector (x, y, z) by letting $u = x/z$ and $v = y/z$. This is called the homogeneous representation of a point, and it lies on the projective plane P^2 [35]. Homography is invertible mapping of points and lines on the projective plane P^2 , thus allowing to obtain the real-world

TABLE 2: Parameters for joint angles calculation.

k	α_k	β_k
0	e_2	$(\pi_e^* \wedge e_\infty) \cdot e_0$
2	$(L_{12} \cdot e_0) \cdot e_\infty$	$(L_{23} \cdot e_0) \cdot e_\infty$
3	$(L_{23} \cdot e_0) \cdot e_\infty$	$(L_{3e} \cdot e_0) \cdot e_\infty$

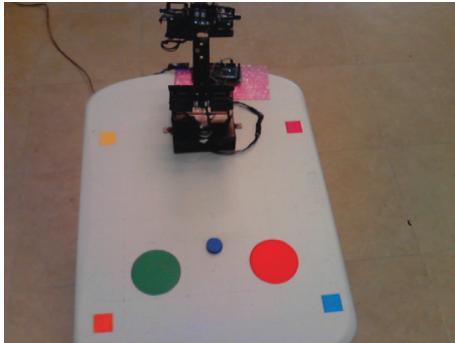


FIGURE 6: Robot and items on the table as seen by the camera during semiautonomous BCI trials.

coordinates of features in an image from its image coordinates.

In our case, the desired transformation is such that the image obtained from the camera is turned into a two-dimensional view of the same setup. In this transformation, the image shows a planar representation of the original view, as if the camera was placed directly above the delimited square. In order to obtain this representation, the following homography transformation was applied [35]:

$$\begin{bmatrix} u \\ v \end{bmatrix} = H \begin{bmatrix} x \\ y \end{bmatrix}, \quad (16)$$

where vectors $[uv]^T$ and $[xy]^T$ represent the positions of selected points in the image and their corresponding positions in real-world coordinates, respectively, $H = K[R|t]$ is the homography matrix that defines the desired perspective change to be performed on the image, and K is the calibration matrix which contains the intrinsic parameters of the camera, while R and t are, respectively, the rotation matrix and translation vector applied on the camera in order to perform this transformation view. In (16), z is ignored as all items are considered to be at $z = 0$.

In order to compute matrix H , both real-world and image coordinates of the centroids of the square markers were obtained. First, markers were detected through color segmentation and binarization, as shown in Figure 7(a). This process was performed separately on each marker, and their contours were detected. After that, the centroids of the markers in the image were calculated. The contours and centroids of each marker are shown in Figure 7(b).

Since the markers have known dimensions (30×30 mm), the positions of their centroids in real-world coordinates relative to the plane are known as well. These positions were defined as cyan at [15, 15] mm, orange at [385, 15] mm, magenta at [15, 385] mm, and yellow at [385, 385] mm, all inside the available 400×400 mm area of

the table. Then, both sets of coordinates are used to obtain H with OpenCV's command `findHomography`, and the resulting matrix is applied to transform the image, as shown in Figure 7(c).

Then, using the same procedure as with the markers, the centroids of the disk and targets in the new image were calculated. However, the reference frame from the image is different from the reference frame from the robot. Therefore, the first was transformed by applying the following rotation matrix:

$$R = \begin{bmatrix} \cos \pi & -\sin \pi \\ \sin \pi & \cos \pi \end{bmatrix}. \quad (17)$$

Furthermore, a translation vector $[-200 - 400]^T$ was applied as well as a sign switching of the x axis to obtain the desired positions. In robot frame, the x axis of the delimited square goes from -20 to 20 , while the y axis goes from 0 to 40 , and the robot is located at the origin. After applying all those transformations, the centroids of all items are finally expressed in robot frame, and they can be detected by the AV system together with the contours of all items. This is shown in Figure 8.

3. Implementation of BCI Systems

As proof-of-concept, four participants volunteered in this study (two females and two males, with average age of 22.25 years, SD = ± 0.95). The experimental protocol was divided into three stages for both the process-control and goal-selection BCIs: (i) training, (ii) cued manipulation, and (iii) uncued manipulation. Both BCIs were MI-based; therefore, users were trained to control the corresponding μ band desynchronization at will. In all trials, volunteers sat in front of a computer screen first showing a black screen (baseline) in which the user was meant to be in a resting state. Then, different types of stimulus were presented to the user, representing each a different command. The duration for the baseline (15 seconds) and stimulus presentation (4 seconds) was the same for all trials and stages. During stimulus presentation, users were expected to react accordingly, either by imagining the movement of either left or right hand, or by remaining in a resting state. In training trials, EEG signals were acquired and analyzed offline to build and evaluate the performance of classifiers, which were then used online during the manipulation trials. In cued manipulation trials, the user was expected to manipulate the device as indicated by the stimuli. On the other hand, the user was encouraged to manipulate the device at will during uncued manipulation trials.

3.1. Training Trials. The training protocol was identical for both the process-control and goal-selection BCIs. Three types of stimuli were presented to the user: *right hand imaginary movement* (RHIM), *left hand imaginary movement* (LHIM), and *rest*. A total of 30 stimuli (10 for each command) were randomly presented to the user. Stimuli were represented in the computer screen with a red arrow pointing to the right (for RHIM), pointing to the left (for

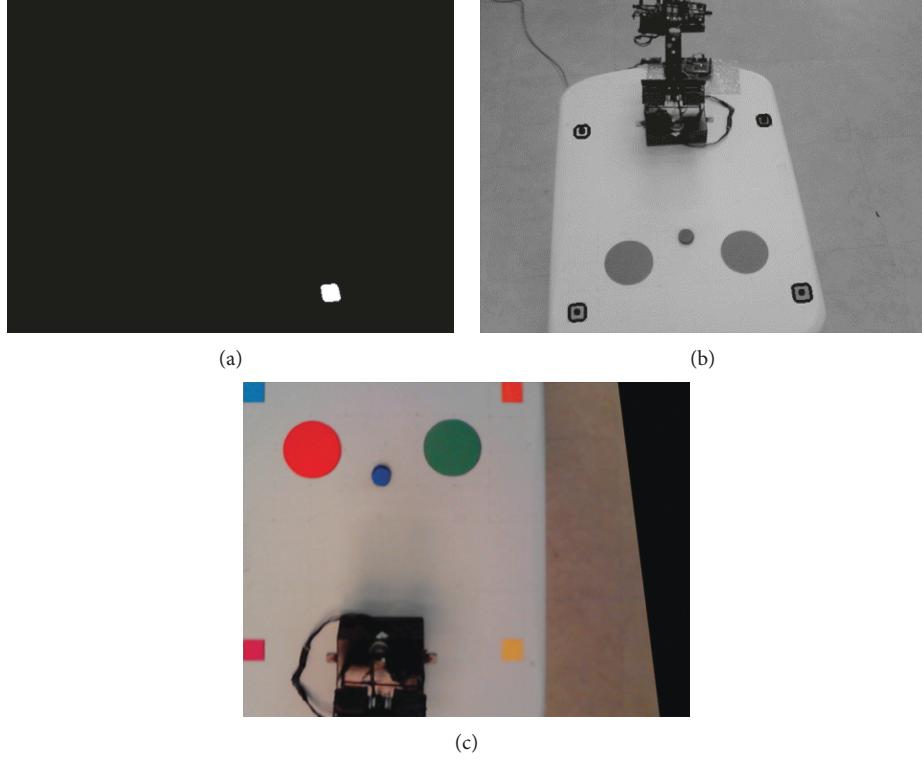


FIGURE 7: Required steps of the AV algorithm. (a) Segmentation and binarization. (b) Centroid calculation. (c) Homography transformation.

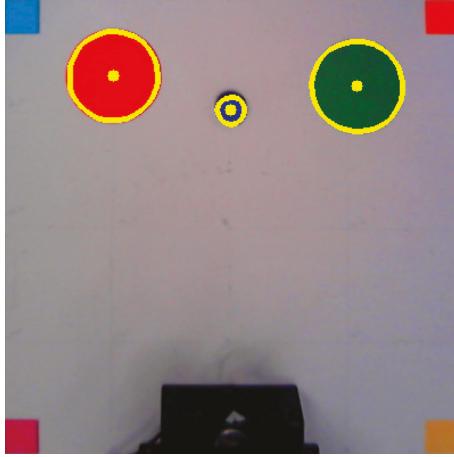


FIGURE 8: Visual representation of contours and centroids of the items in the table calculated by the AV algorithm in order to obtain their real-world coordinates.

LHIM), and a black screen for rest. A 2 s green cross appeared before all stimuli as a prestimulus, and there was a variable interstimulus resting period of 2–4 seconds between stimuli. Users underwent three training sessions on different days, each comprising five repetitions of the mentioned experimental protocol, while EEG recordings were obtained.

3.2. Signal Acquisition. EEG signals were recorded with the Mobita equipment from TMSi systems, using a measuring

cap of 19 channels: FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Cz, Fz, and Pz. Impedance of all electrodes was set below $5\text{ k}\Omega$ for all experiments. Signals were acquired with a sampling frequency of 1000 Hz. Recordings were band-pass filtered with a fourth-order 1–100 Hz Butterworth filter and a 60 Hz notch filter to eliminate power line interference. The OpenViBE software was used for the BCI design and implementation. More information about this software can be found in [36].

3.3. Classification Algorithm. Feature extraction was performed using the BCI2000 offline analysis tool (https://www.bci2000.org/mediawiki/index.php/User_Reference:BCI2000_Offline_Analysis), where the r^2 value was calculated. A higher r^2 value is related to a higher discrimination of a signal under two stimulus conditions. More details about the statistic meaning of r^2 can be found at <https://www.bci2000.org/mediawiki/index.php/Glossary>. After each training session, signals from the five training trials were used to calculate r^2 . Three r^2 maps (one per stimulus combination) were obtained per training session, showing the r^2 values in the 19 available channels and frequencies ranging from 1 to 70 Hz. Each map, as the one shown in Figure 9, represents the channels and frequencies which, for a specific combination of conditions, showed higher discrimination. Through this procedure, the selected channels and frequencies were used as features for the classification algorithm.

Signals were spatially filtered using a Laplacian filter on the selected channels as well as through a fourth-order

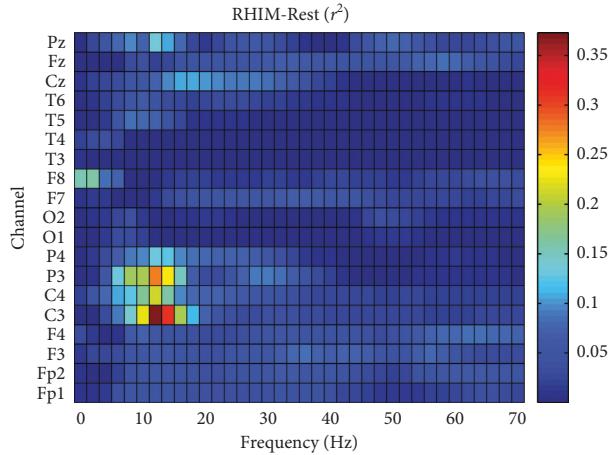


FIGURE 9: Representative r^2 map obtained during one training session. r^2 values shown here were measured under conditions RHIM-Rest for all channels and frequencies.

Butterworth band-pass filter tuned to the selected frequencies. Power values were then obtained from the filtered signals to build the feature vectors, which then became the input for a linear discriminant analysis (LDA) classifier, which separates data representing different classes by finding a hyperplane which maximizes the distances between the means of the classes, while minimizing the variance within classes [37].

In our case, three pairwise classifiers per training session were obtained using this procedure: LHIM versus RHIM, LHIM versus rest, and RHIM versus rest. The three classifiers were tested online on the recorded signals to evaluate their performance as a percentage of correctly classified stimuli. The classification was performed on each four-second stimulus, divided into overlapped subepochs using a window function. Each four-second epoch was formed by 64 subepochs of two seconds, separated by 0.0625 seconds. One pairwise classifier labeled each subblock as one of the two possible classes, and the four-second epoch was classified as the mode of the classification result for all its subblocks. Then, one general classifier was built, based on the results of the three pairwise classifiers. Here, the four-second epoch of each stimulus was labeled as class $I = 1, 2$, or 3 (LHIM, rest, or RHIM, respectively), if two out of the three pairwise classifiers labeled the same epoch identically. The mean performance of the general classifiers across trials is shown in Table 3 for all subjects and training sessions as well as their selected features. After training sessions, each user proceeded to perform the subsequent trials using the classifier with the highest performance obtained at the last training session.

3.4. Process-Control BCI. The process-control BCI was designed in such a manner that users were able to perform three-dimensional movements to complete reaching tasks. In this system, the position of the final effector as well as the desired axis in which the effector moves can be controlled through low-level MI-based commands. To

achieve this, the user has two choices: moving along a selected axis (y -axis at the initial step) or change between axes. In the design of this BCI, the classification of a LHIM results in a -10 mm displacement, while the classification of a RHIM results in a $+10$ mm displacement on the selected axis. The classification of a rest event holds the position of the final effector with no displacement. The consecutive classification of two rest events in a row allowed the user for a change of axis. This change of axis takes place in the following sequence: $y \rightarrow z$, $z \rightarrow x$, and $x \rightarrow y$.

3.4.1. Cued Manipulation. In these trials, users sat in front of a computer showing three windows on the screen. The first window was used for stimulus presentation, the second was used to display in which axis the movement of the robot took place, and the third was used to visualize the robot and its movements. The setup for these experiments is shown in Figure 10. After the baseline period, 15 random stimuli (5 for each type) were presented to the user. Prestimulus, stimulus, and interstimulus duration was the same as in training trials (see Section 3.1). After the stimulus was presented, the user was expected to emit the instructed command through the BCI. Then, the robot performed a specific movement based on the classification result. In these trials, performance was evaluated as the percentage of correctly classified stimuli. The intention of these trials was to get the users acquainted with the BCI, and they were performed immediately before the uncued manipulation trials. Users performed three sessions on different days, each formed by three repetitions of this protocol.

3.4.2. Uncued Manipulation. The same screen display was used as in cued trials, but here subjects were asked to complete reaching tasks on their own. At the start of each trial, the final effector was fixed at home position $[0, 155.5, 284.3]$ and a target was placed at $[0, 300, -49]$. At this initial step, the distance of the final effector to the target was 360 mm. Note that the target is placed at $z = -49$, as the robot base is 49 mm above the table. A baseline period was followed by the presentation of 20 stimuli showing the word “Imagine,” in which the user was expected to emit MI commands through the BCI. The duration of prestimulus, stimulus, and interstimulus periods was the same as in training trials (see Section 3.1). The user was instructed to move the final effector as close as possible to the target within the 20 stimuli, using the protocol described in Section 3.4. Performance was evaluated as the percentage of stimuli where the user moved the final effector closer to the target and changed successfully to the y -axis. Users performed three sessions on different days, each formed by five repetitions of the described protocol.

3.5. Goal-Selection BCI. The goal-selection BCI was designed to perform in a semiautonomous way pick-and-place tasks with the disk and two possible targets. Users were able to

TABLE 3: Features: EEG channels and frequency range (in Hz) and mean accuracy of LDA classifiers for all subjects and training sessions.

Subject	Session	Features	Accuracy (%)
S_1	1	C4 (11–15), P3 (11–15), C3 (7–11), Fp2 (13–17)	65
	2	C3 (7–13), P3 (9–13), P4 (9–13), P4 (13–17), Cz (9–13)	64
	3	C4 (9–15), C3 (13–17), P3 (23–27), P4 (25–29)	63
S_2	1	C4 (9–13), P4 (11–15), F4 (17–21), F4 (11–15), Cz (11–15)	65
	2	C4 (15–19), C3 (19–23), P4 (21–25), Cz (19–23), FP1 (25–29)	62
	3	C4 (19–23), C4 (17–21), Cz (19–23), P3 (21–25), F3 (19–23)	60
S_3	1	F4 (15–21), F3 (9–11), P4 (9–13), F4 (15–21)	56
	2	C3 (11–15), C4 (7–11), P3 (15–19), P4 (11–15)	61
	3	C3 (9–15), C4 (11–15), P3 (11–15)	78
S_4	1	C4 (9–13), C3 (9–13), P4 (17–21), F4 (19–23), F3 (19–23)	73
	2	F3 (11–15), C4 (7–13), P4 (17–21), P3 (17–21), F4 (19–23)	72
	3	C4 (7–11), C3 (13–17), F4 (17–21), C4 (11–15)	60

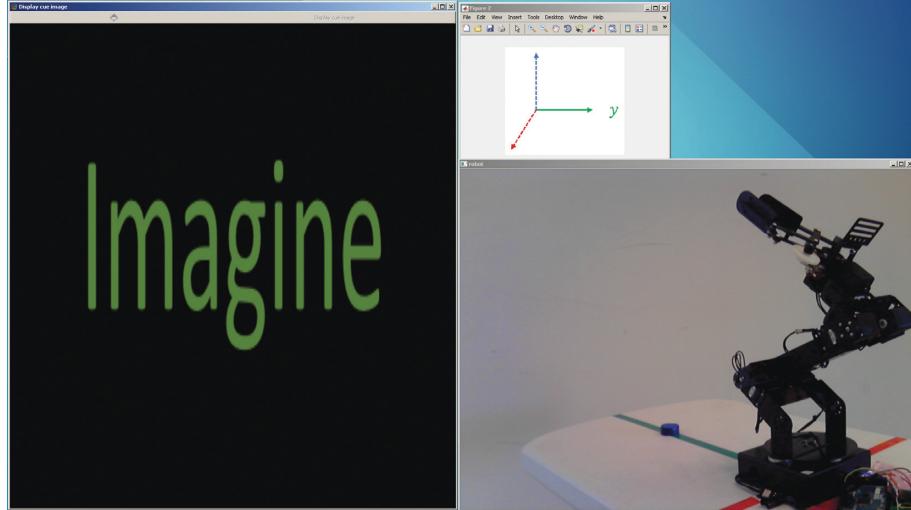


FIGURE 10: Setup of the process-control BCI. The windows shown in the screen are used for visualization of stimuli, indicating the current axis of the movement, and viewing of the robot performing the manipulation tasks.

perform these tasks for any position of the items (randomly chosen before a trial), inside of the robot workspace. The centroids $C = [C_x, C_y]$ of the two target stickers were calculated in these trials by the AV algorithm. In this case, the classification of three types of events resulted in different manipulation tasks:

- (i) If an event was classified as RHIM, the robot reached for the disk, placed it on the target located to the right (greater C_x component), and returned to home position
- (ii) If an event was classified as LHIM, the robot reached for the disk, placed it on the target located to the left (smaller C_x component), and returned to home position
- (iii) If an event was classified as rest, the robot remained at home position

After the robot performed a manipulation task, all the items in the table were manually changed to random positions, in preparation for the next trial.

3.5.1. Cued Manipulation Trials. In these trials, the subject sat in front of a computer screen which showed two screens. The first one was used for stimulus presentation, while the second was used to present the transformed image, as shown in Figure 11. After the baseline period, a stimulus (RHIM, LHIM, or rest) was randomly presented. A total of 15 stimuli (5 for each type) were presented in each trial. A one-second *beep* sound followed a two-second green cross as prestimulus, with a 27–29 seconds interstimulus period. Manipulation tasks were performed according to the result of the classification, and performance was evaluated as the percentage of correctly classified stimuli. The total duration of these trials was considerably longer than in the low-level BCI. This is mainly due to the longer interstimulus period, in which the manipulation tasks took place. Users underwent three sessions on different days, performing five trials in each session.

3.5.2. Uncued Manipulation Trials. For uncued manipulation trials, all stimuli were replaced with the word “Imagine,”

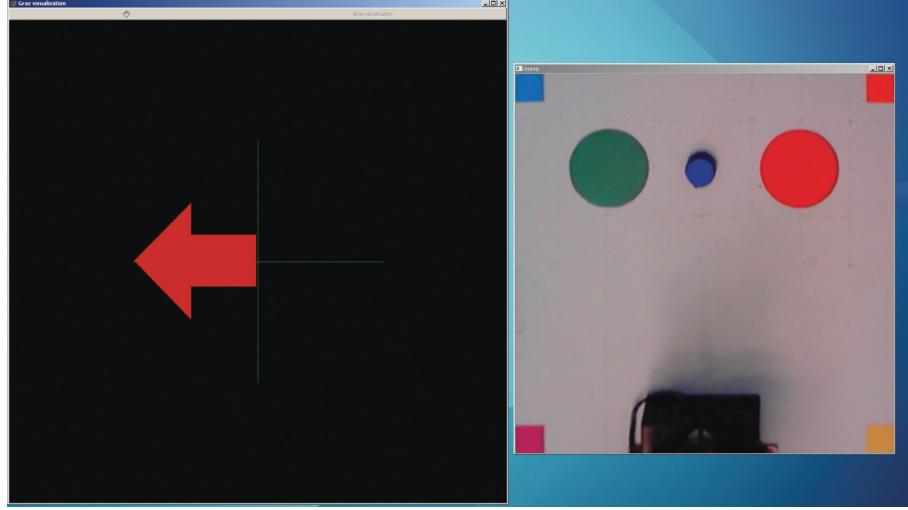


FIGURE 11: Setup of the semiautonomous goal-selection BCI, as seen by the user. The windows are used for stimulus presentation and visualization of the manipulation tasks.

and the user freely decided the task to perform, as explained in Section 3.5. A total of 15 stimuli were presented in each trial. The stimulus, prestimulus, and interstimulus duration was the same as in the goal-selection BCI cued manipulation trials (see Section 3.5.1). Immediately after the classification was performed, and before the robot executed the task, the user was asked the type of intended stimulus to emit. In these trials, performance was evaluated as the percentage of coincidences between the intended and the classified stimulus type.

3.6. Analysis of Data through P300 Estimation. Reported assessments of mental fatigue through P300 amplitude and latency can be found in [38] and [19]. In [19], mental fatigue was evaluated through EEG measurements. Participants' P300 were measured during a modified Erikson flanker task, replacing word stimuli with arrows, before and after performing mental arithmetic tasks. A decreased P300 amplitude and an increased latency were observed after performing arithmetic tasks, when users were mentally fatigued. Statistical analysis revealed the most significant changes in amplitude and latency at channels O1, O2 and Pz, probably as a reflection of visual processing during stimulus presentation of arrows. Similar to the protocol used in [19] to assess mental fatigue, signals were segmented into 1 s stimulus-locked EEG epochs from 200 ms before and 800 ms after stimulus presentation. These epochs were obtained for the presentation of the word "Imagine" during uncued manipulation trials for both the process-control and goal-selection BCIs. For each trial, a representative waveform was obtained by averaging the epochs from all stimuli. Then, the averaged waveforms were band-pass filtered at 1–10 Hz and were used to calculate P300 amplitude and latency. The amplitude was considered as the most positive peak within a 200–500 ms window immediately after stimulus presentation. Latency was obtained as the time this peak appeared. Amplitude and latency values were obtained through this procedure for all trials, sessions, and subjects, in channels O1, O2, and Pz. A representation of an obtained

P300 waveform is shown in Figure 12 for these three channels.

In order to examine the differences of mental fatigue within and between users in relationship with the use of our two different BCI schemes, two-way ANOVA tests were performed on all users: one for amplitude and one for latency. In these tests, influence of *trial repetition* (1–5), *channel location* (O1, O2, and Pz), and their interaction were analyzed on both P300 features. The number of replications was considered as three, representing the three uncued manipulation sessions performed by the users. To further analyze mental fatigue related to continuous BCI manipulation, one-way ANOVA ($p < 0.05$) tests were performed on each subject. Six one-way ANOVA tests were performed per subject: three channels (O1, O2, and Pz) \times two P300 features (amplitude and latency). These tests were performed in order to find which channel showed significant relationship to the trial repetition factor. Then, amplitude and latency values of all users were compared using the most significant channel from this analysis.

4. Results

A preliminary validation of our CGA model and AV algorithm can be found in [26] and [27], respectively; hence, we omit those details here. Therefore, this section shows the results of evaluating the whole system in the context of our BCI implementations for four subjects (two on each BCI type). Performance values were obtained for all subjects in training, cued, and uncued trials, according to the particularities of each experimental protocol. For training trials, performance values correspond to the classifier accuracies shown in Table 3. Performance for cued and uncued manipulation trials was obtained as explained in Sections 3.4 and 3.5. Performance values included in these results represent the average across trials for each session.

4.1. Performance of Process-Control BCI. Subject S_1 reached an accuracy level of 65% at its first training session, 64% at the

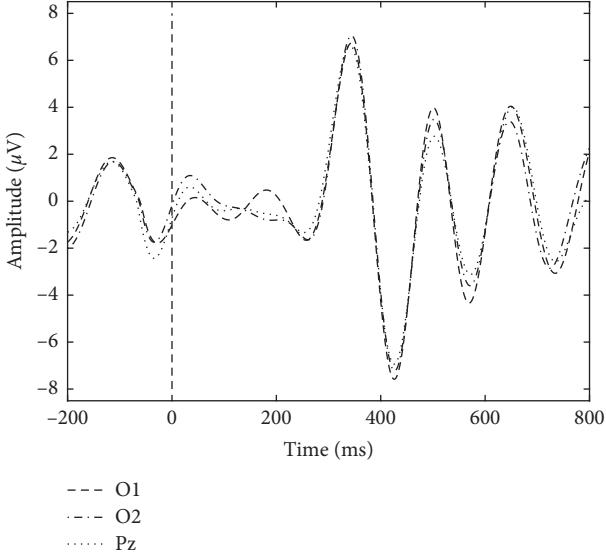


FIGURE 12: Representation of a P300 waveform calculated for channels O1, O2, and Pz.

second, and 63% by the third. During cued manipulation trials, performance started at 18% and then increased to 25% and 29% by the second and thirds sessions, respectively. For uncued manipulation trials, the user only moved far from the target at the first session (0%). For the second and third sessions, user S_1 obtained performances of 14% and 17%. Subject S_2 showed a similar behavior to S_1 during training trials, starting at 65% and decreasing to 62% and 60% by the second and third sessions. In cued manipulation trials, performance started at 33% and then increased to 37% by the second session and 45% by the third. For uncued manipulation trials, performance values started at 28% for the first session, then decreased to 17% at the second, and increased to 37% by the third. Results for process-control BCI performance are shown in Figure 13 for users S_1 and S_2 , respectively.

4.2. Performance of Goal-Selection BCI. Subject S_3 started the training sessions at 56% of accuracy, 58% at the second session, and reached 78% at the third session. Performance for the cued manipulation trials started at 40%, increasing to 56% at the second session and decreasing to 49% by the third. During uncued manipulation trials, performance started at 60% accuracy by the first session, 53% at the second, and 67% by the third. User S_4 obtained performance values of 73% at the first session, 72% at the second, and decreased to 60% at the third. During cued manipulation trials, the subject obtained performance values of 45% for the first session, 41% for the second, and 46% for the third. For uncued manipulation trials, user performance started at 30% and increased to 38% and 48% by the second and third sessions, respectively. Results for goal-selection BCI performance are shown in Figure 14 for users S_3 and S_4 , respectively.

4.3. P300 Analysis. The results for the two-way ANOVA tests are presented in Table 4. The results for the P300 latency

two-way ANOVA showed statistical significance for subjects S_1 ($p = 0.0147$) and S_2 ($p = 0.0001$) in the trial factor, but no significance was observed for channel and interaction factors. Users S_3 and S_4 showed no statistical significance for any of the analyzed factors. For the P300 amplitude two-way ANOVA, users showed smaller p values in trial when compared to channel and interaction. However, our tests did not show statistical significance for any factor or interaction.

The results for the one-way ANOVA tests are shown in Table 5. The results for P300 latency one-way ANOVA showed statistical significance for user S_1 at channel O1 ($p = 0.0476$) and for user S_4 at channel O2 ($p = 0.0242$). Regarding users S_2 and S_3 , p values were not significant at any channel. For the P300 amplitude one-way ANOVA, user S_4 showed statistical significance at channel Pz ($p = 0.0019$). The tests for S_1 , S_2 , and S_3 revealed no statistical significance at none of the three analyzed channels.

The results of the performed statistical tests allowed to observe differences between analyzing latency and amplitude. Among all tests, greater changes were found in latency rather than in amplitude. Based on these results, an evaluation and comparison on amplitude and latency values was performed. These values were considered as those corresponding to the channel with the lowest p value on the one-way latency ANOVA results. The selected channels were O1 for S_1 , Pz for S_2 and S_3 , and O2 for S_4 .

Amplitude values calculated for all uncued manipulation trials are shown in Figure 15 for each session and user. Users S_1 and S_4 showed a similar behavior: a decreasing P300 amplitude trend in all sessions. In this case, the amplitude observed at the first trial was higher than that of the last one. S_2 showed a decreasing trend as well for the first and second sessions, yet the opposite was observed during the third session. S_3 presented an increasing amplitude trend for all sessions. Here, the amplitude obtained at the last trial was higher than the one at the first trial.

Latency values can be observed in Figure 16 for all users and sessions. Subjects S_1 and S_3 showed an increasing P300 latency trend during the first and third sessions. A decreasing trend was observed during the second session for these users. User S_2 presented an increasing latency trend for all sessions. User S_4 showed an increase in latency during the first and second sessions and a decrease at the third.

5. Discussion

The implementation and integration of the CGA model and the AV algorithm allowed to successfully design a MI-based semiautonomous BCI for manipulation tasks. When compared against a low-level system, both BCIs were similar in terms of training protocol and control commands; however, the complexity of the executed tasks was different. The semiautonomous goal-selection BCI was superior in task complexity when compared to the process-control BCI, even though both systems used the same control commands as input. While the process-control BCI might be used to perform more general tasks, it demands a continuous awareness state from the user. Its output are discrete low-level commands which in the long run might lead the user to

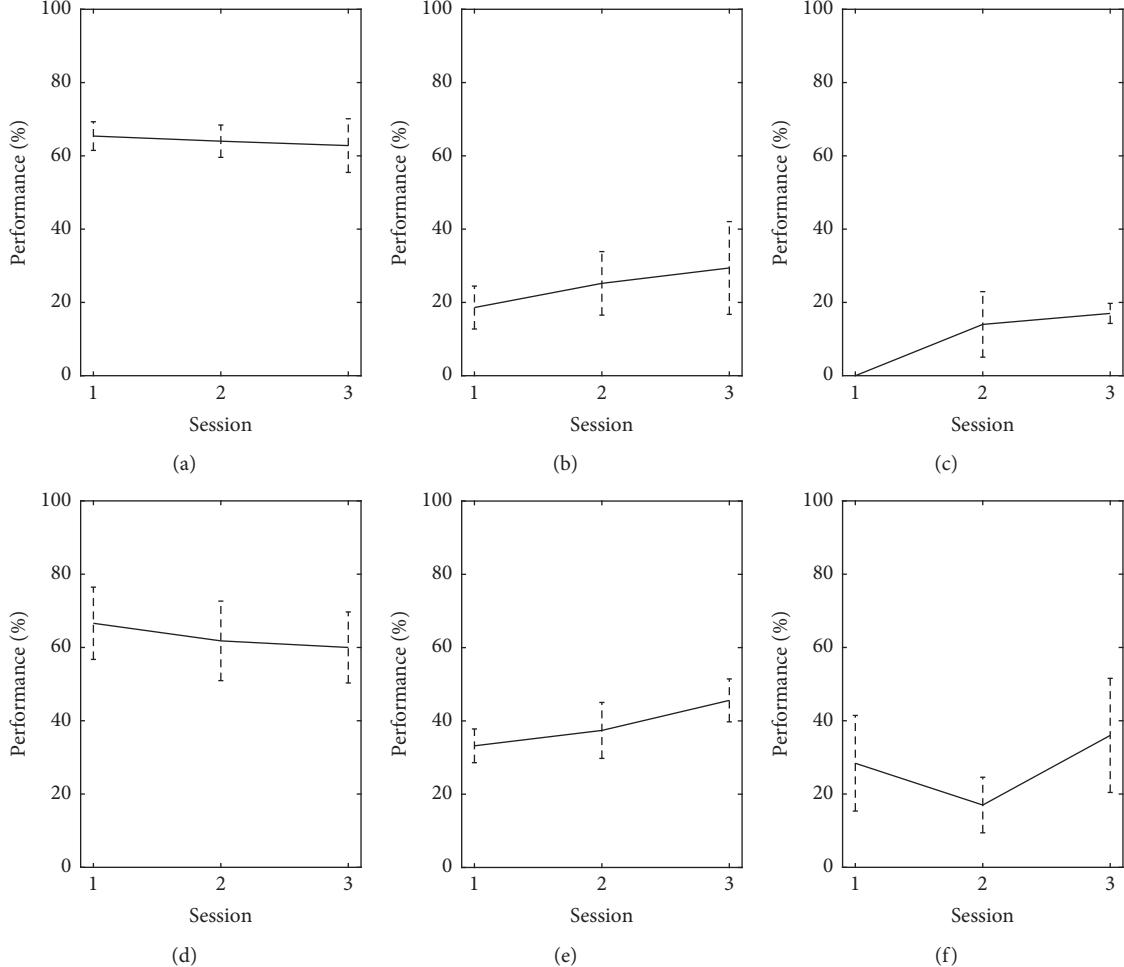


FIGURE 13: Performance for users S_1 (top) and S_2 (bottom) in process-control BCI during training, cued, and uncued manipulation trials (left, middle, and right columns, respectively). Bars indicate one standard deviation. (a) S_1 , training. (b) S_1 , cued manipulation. (c) S_1 , uncued manipulation. (d) S_2 , training. (e) S_2 , cued manipulation. (f) S_2 , uncued manipulation.

a mental fatigue condition. Although the semiautonomous BCI is goal specific, it requires user attention only during short time periods, making it theoretically less fatiguing. The semiautonomous goal-selection BCI works, in essence, in a more natural way to the user than the process-control BCI. This is because when performing reaching tasks, people think on the main goal and the cerebellum processes the necessary information to successfully achieve it, rather than executing several discrete low-level movements [39].

The selected features for the general classifiers of the users were mainly frontal, central, and parietal electrodes in the μ (8–13 Hz) and β (13–30 Hz) brain rhythms, which are known to be physiologically involved in the imaginary movement process. The selected channels for the classifiers are consistent with reports of central activity as a reflection of motor cortex contralateral desynchronization during imaginary movement [10] and frontoparietal activation related to control of spatial attention and motor planning during reaching tasks [40, 41].

Even though all users underwent the same training protocol, differences among them were observed. Across training sessions, S_1 and S_2 maintained a relatively constant performance, while S_3 showed a more notorious improvement.

S_4 showed a relatively high performance at the first and second session, but it decreased at the third. During cued manipulation trials, all users obtained low performance levels and none of them showed a significant improvement across sessions. S_1 obtained below chance level (33%) performance during all sessions. Performance of users S_2 , S_3 , and S_4 was in general above chance level, but always remained below 60%. During uncued manipulation trials, users S_1 and S_2 presented the lowest performance values, close to and below chance level. This indicates that these users were faced with difficulty while controlling the process-control BCI. Performance of S_3 and S_4 during uncued manipulation trials was higher (around 40–60%) when compared to S_1 and S_2 . Mean performance values across trials of users S_3 and S_4 failed to reach 70% considered as the theoretical threshold for practical MI-BCI use [42]. However, their performance was evidently higher than the one obtained by users performing on the process-control BCI. This might suggest that the designed semi-autonomous goal-selection BCI was easier to manipulate than the process-control BCI. Future research will address classification optimization to increase system accuracy and ease of use.

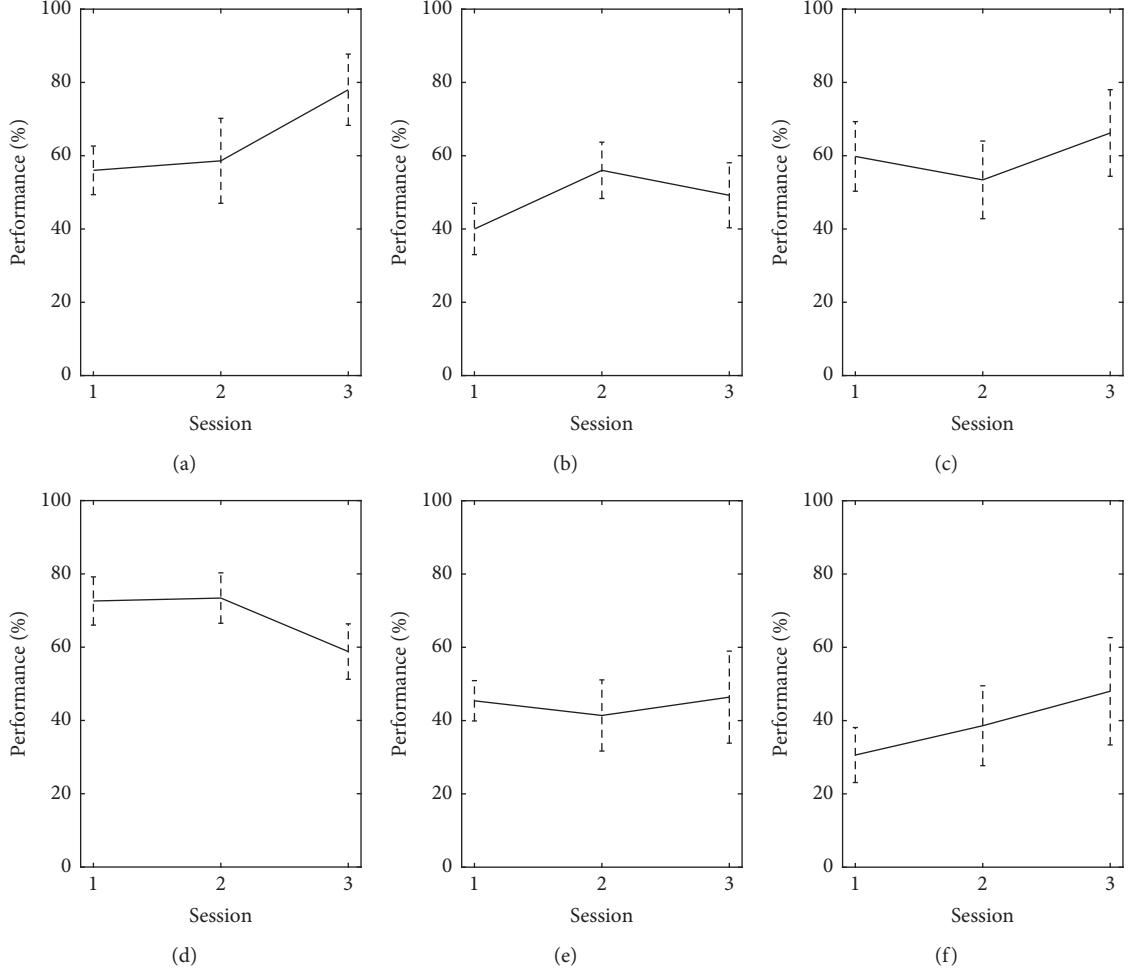


FIGURE 14: Performance for users S_3 (top) and S_4 (bottom) in semiautonomous goal-selection BCI during training, cued, and uncued manipulation trials (left, middle, and right columns, respectively). Bars indicate one standard deviation. (a) S_3 , training. (b) S_3 , cued manipulation. (c) S_3 , uncued manipulation. (d) S_4 , training. (e) S_4 , cued manipulation. (f) S_4 , uncued manipulation.

TABLE 4: Two-way ANOVA results for P300 latency and amplitude.

Subject	Trial	Latency			Amplitude		
		Channel	Interaction	Trial	Channel	Interaction	
S_1	$F = 3.69$	$F = 0.25$	$F = 0.69$	$F = 2.45$	$F = 3.08$	$F = 0.42$	
	$p = 0.0147$	$p = 0.782$	$p = 0.6994$	$p = 0.0676$	$p = 0.0609$	$p = 0.8969$	
S_2	$F = 9.33$	$F = 0.13$	$F = 0.1$	$F = 1.26$	$F = 0.48$	$F = 0.14$	
	$p = 0.0001$	$p = 0.8816$	$p = 0.999$	$p = 0.3074$	$p = 0.6217$	$p = 0.9970$	
S_3	$F = 0.03$	$F = 0.59$	$F = 0.44$	$F = 0.83$	$F = 0.01$	$F = 0$	
	$p = 0.9983$	$p = 0.5604$	$p = 0.8891$	$p = 0.5191$	$p = 0.9924$	$p = 1$	
S_4	$F = 1.05$	$F = 1.5$	$F = 0.3$	$F = 1.81$	$F = 0.17$	$F = 0.13$	
	$p = 0.4003$	$p = 0.24$	$p = 0.9589$	$p = 0.1534$	$p = 0.8405$	$p = 0.9972$	

Bold values highlight those for which $p < 0.05$.

As shown in Table 3, selected channels and frequencies for feature extraction showed changes across sessions for all users. This might suggest that the used channel/frequency selection method is sensitive to intra- and intersubject brain variability. After training trials, a classifier with fixed parameters was selected per subject and used in all BCI trials. Yet, constant adaptation of the classifier parameters is

required for optimal operation. Hence, an optimized feature selection algorithm should be implemented to address this issue and increase efficiency in our proposed semi-autonomous BCI. Such optimization was out of the scope of our work, but reports on how optimized correlation-based feature selection methods are used in MI-BCIs can be found in [43, 44].

TABLE 5: One-way ANOVA results for P300 latency and amplitude on channels O1, O2, and Pz.

Subject	Latency			Amplitude		
	O1	O2	Pz	O1	O2	Pz
S_1	$F = 3.54$ p = 0.0476	$F = 0.76$ $p = 0.5767$	$F = 0.55$ $p = 0.7055$	$F = 1.79$ $p = 0.2066$	$F = 1.61$ $p = 0.246$	$F = 0.82$ $p = 0.5421$
S_2	$F = 2.1$ $p = 0.1554$	$F = 2.13$ $p = 0.1518$	$F = 2.5$ $p = 0.1091$	$F = 0.3$ $p = 0.8705$	$F = 0.23$ $p = 0.9171$	$F = 0.1$ $p = 0.9806$
S_3	$F = 0.37$ $p = 0.8238$	$F = 0.12$ $p = 0.9731$	$F = 0.54$ $p = 0.7089$	$F = 1.31$ $p = 0.3322$	$F = 0.13$ $p = 0.9687$	$F = 0.41$ $p = 0.8043$
S_4	$F = 1.68$ $p = 0.2311$	$F = 4.52$ p = 0.0242	$F = 2.45$ $p = 0.1138$	$F = 2.99$ $p = 0.0728$	$F = 2.27$ $p = 0.1336$	$F = 9.5$ p = 0.0019

Bold values highlight those for which $p < 0.05$.

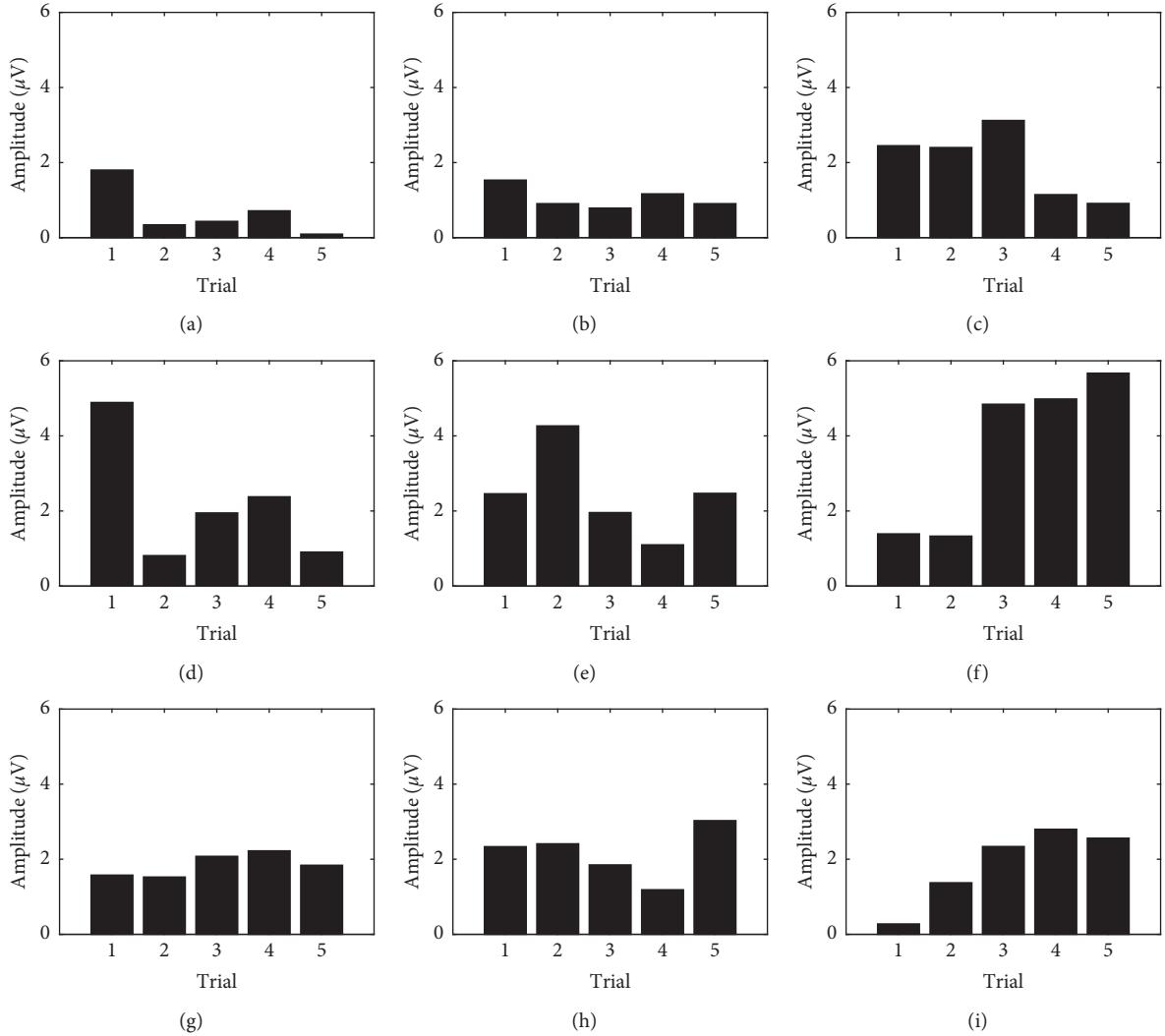


FIGURE 15: Continued.

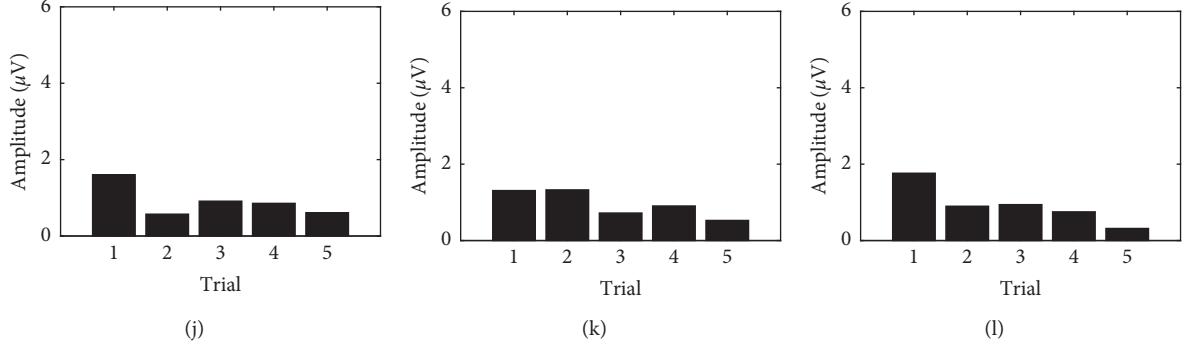


FIGURE 15: Amplitude of P300 waveform during all trials and experiments for all subjects during uncued manipulation trials. (a) Subject S_1 , session 1. (b) Subject S_1 , session 2. (c) Subject S_1 , session 3. (d) Subject S_2 , session 1. (e) Subject S_2 , session 2. (f) Subject S_2 , session 3. (g) Subject S_3 , session 1. (h) Subject S_3 , session 2. (i) Subject S_3 , session 3. (j) Subject S_4 , session 1. (k) Subject S_4 , session 2. (l) Subject S_4 , session 3.

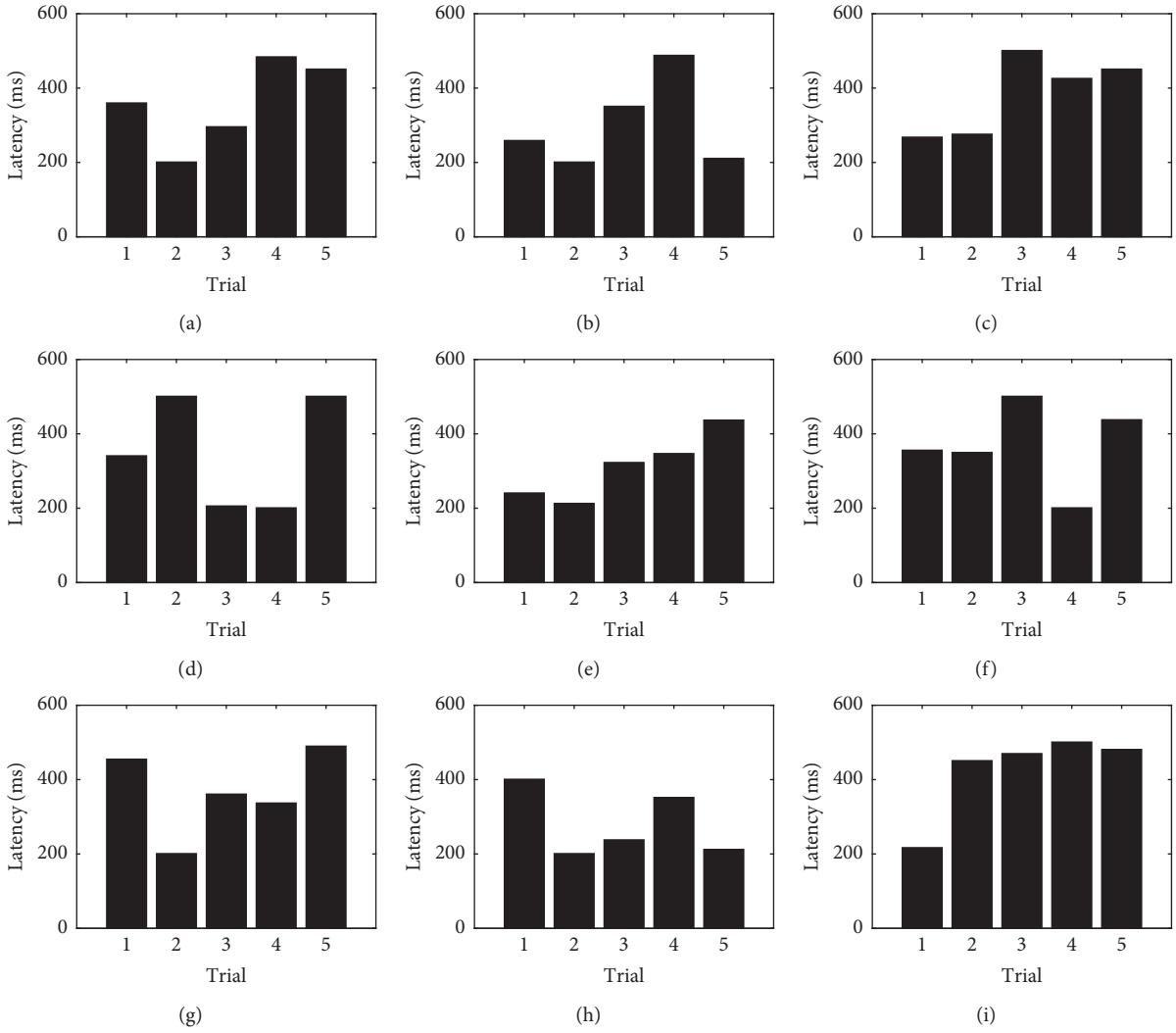


FIGURE 16: Continued.

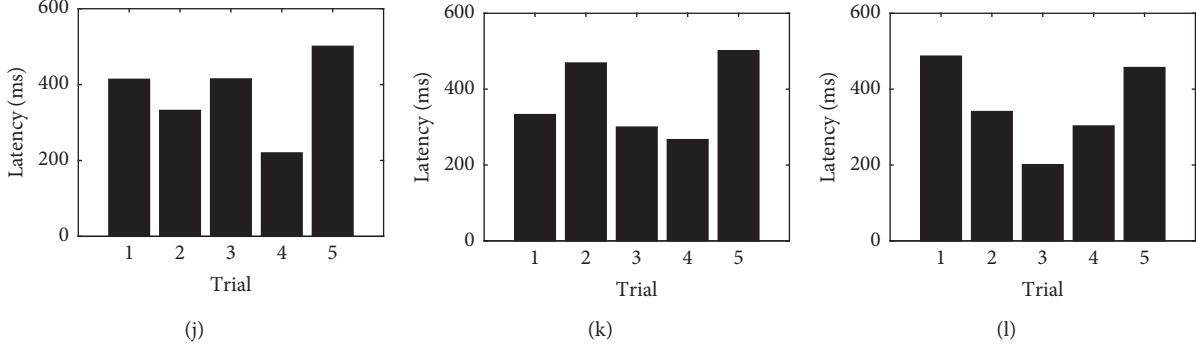


FIGURE 16: Latency of P300 waveform during all trials and experiments for all subjects during uncued manipulation trials. (a) Subject S_1 , session 1. (b) Subject S_1 , session 2. (c) Subject S_1 , session 3. (d) Subject S_2 , session 1. (e) Subject S_2 , session 2. (f) Subject S_2 , session 3. (g) Subject S_3 , session 1. (h) Subject S_3 , session 2. (i) Subject S_3 , session 3. (j) Subject S_4 , session 1. (k) Subject S_4 , session 2. (l) Subject S_4 , session 3.

Another efficient approach for feature selection is the partial directed coherence (PDC) analysis, which could help identify relevant channels and features. Recently, a PDC-based analysis was proposed in [45] to identify relevant features for MI tasks, and efficient classifiers were built based on this procedure. Even more recently, a review on EEG classification algorithms highlights Riemannian geometry-based classifiers as promising, as well as adaptive classification algorithms [46]. A simple implementation of an adaptive classifier for MI tasks was described in [47], which showed an encouraging increase on classification accuracy. More novel classifiers based on Riemannian geometry have shown good results in classifying brain-related MI tasks [48].

In regards to our selection of the P300 component to evaluate mental fatigue, such component is not exclusively presented during nonfrequent stimulus, rather its amplitude is enhanced, which makes it a suitable control command for BCIs. P300 amplitude is larger during nonfrequent stimuli, and it is typically used/analyzed based on this argument. However, it has been demonstrated that P300 responses can be observed to both frequent and nonfrequent stimuli [49, 50]. In fact, under a reaction-time regime, P300 is elicited on both predictable and unpredictable stimuli. Task demands increase in this scenario, as users must decide when to respond in a fast and correct manner. This leads to an enhancement of P300, independently of stimulus predictability [49]. In our study, users were instructed to perform MI commands after stimulus presentation of the word Imagine, and P300 components were analyzed immediately after stimulus onset. Although stimulus presentation during uncued manipulation trials could be considered as predictable, P300 analysis holds validity, as it was executed under a reaction-time regime.

Under those conditions, the results of the two-way ANOVA and one-way ANOVA tests showed statistically significant changes in P300 latency for users S_1 , S_2 , and S_4 . Except for S_4 , the tests revealed no statistical significance for P300 amplitude. When comparing the amplitude and latency values from Figures 15 and 16, a general trend was found among users: a decrease in amplitude and an increase in latency. These trends in P300 features were presented along trial repetition, that is, after continuous manipulation

of the BCI. These changes in amplitude and latency might be related to the generation of mental fatigue, as they are presented after a continuous execution of manipulation tasks through the BCI. It has been shown that a decrease in P300 amplitude and an increase in latency reflect decreased cognitive processing and lower attention levels [19]. Similar results have been found on a P300-BCI evaluated under different levels of mental workload and fatigue [51]. When comparing subjects performing on the same BCI type, the user with the lowest performance exhibited lower amplitude and higher latency values than the user with the highest performance (although it was more evident for amplitude values). This was observed when comparing both S_1 - S_2 and S_3 - S_4 . Subject S_3 showed an interesting behavior: an increasing amplitude trend as well as being the only subject which did not show statistical significance on any P300 test. At the same time, it was the subject with the highest performance values on uncued manipulation trials. A possible explanation to this particular case is that after performing manipulation trials on the BCI, mental fatigue affected differently user S_3 than the rest of the users. This difference in mental fatigue generation was reflected as nonsignificant changes in P300 parameters during the tests as well as higher performance values.

6. Conclusions

Two BCI systems, a process-control BCI and semiautonomous goal-selection BCI, were implemented and compared in terms of performance and mental fatigue. The process-control BCI allowed users to perform three-dimensional movements on a robotic arm to reach for a target. The semiautonomous BCI allowed the user to execute manipulation tasks, using the same robotic arm, which include reaching, picking, and placing movements successfully. The increase of task complexity represented by the semiautonomous BCI was achieved without compromising the simplicity of the control procedure, as both BCIs were controlled through MI commands. Users performing on semiautonomous BCI obtained higher performance values when compared to users performing on low-level BCI. The difference in task complexity also represented a difference in

the mental fatigue experienced by the users on different systems. A P300 amplitude decrease and a latency increase were found as users performed continuous BCI trials, which is consistent with reports of mental fatigue detection on EEG.

We also present strong evidence of the advantages of semiautonomous BCI in terms of performance and mental fatigue. It is also important to address the potential use of the P300 waveform as an indicator of mental fatigue during BCI testing, training, and evaluation. Techniques to further reduce mental fatigue while using BCI systems might provide an increase in BCI patient acceptance rate as well as a possible path to tackle BCI illiteracy. It is of great importance that the user finds the system as nonfatiguing and easy to use in order to provide a more comfortable and efficient assistance. This also facilitates the user in the process of learning how to control the BCI, which can be used together with different strategies to further personalize the system (see, e.g., a previous work by our group in how to select a feedback modality that better enhances the volunteer's capacity to operate a BCI system [52]).

The development of more advanced semiautonomous BCI systems which provide information about the environment during specific tasks will allow to further enhance performance and usability. Semiautonomous BCIs offer users the possibility to perform more complex tasks in a simple, less fatiguing way. In our system, the integration of the AV and CGA algorithms provided a real-time calculation of the robot's inverse model, offering flexibility to implement more complex object manipulation tasks in a dynamic environment. The use of a higher DOF robotic arm as well as the implementation of object recognition techniques might improve the complexity of the manipulation tasks to be performed while using the same MI commands to control the BCI, ensuring control simplicity to the users.

Data Availability

The electroencephalography datasets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Supplementary Materials

Video #1 (file: VID_20180629_133018082~2.mp4): this video shows the operation of the semiautonomous BCI system, specifically the part of the robotic arm performing the process of picking the disk and replacing it over one of the target areas. The pick-and-replace process is repeated once the position of all items is changed randomly. Incidentally, the robot places the disk in the green (left) target in both occasions, as that was the one selected by the user in the two trials shown here. Video #2 (file: VID_20180629_133317483~2.mp4): this video shows the user, in a contiguous room, interacting with the BCI system.

The brain activity of the user is measured with the EEG system, and a stimulus is provided through a red arrow in the screen which indicates the MI that the user has to perform. Also, the user receives video feedback of the robot movement. Video #3 (file: VID_20180629_133853176~2.mp4): this video shows the screen of the computer that processes all the data that are being acquired both by the EEG system and the AV algorithm. The screen also shows the CGA model that is adjusted on the fly based on the positions of the items detected by the AV algorithm and the selected target area where the disk will be placed. (*Supplementary Materials*)

References

- [1] L. Bi, X.-A. Fan, and Y. Liu, "EEG-based brain-controlled mobile robots: a survey," *IEEE Transactions on Human-Machine Systems*, vol. 43, no. 2, pp. 161–176, 2013.
- [2] H. Cecotti, "A self-paced and calibration-less SSVEP-based brain-computer interface speller," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 127–133, 2010.
- [3] C. Wang, B. Xia, J. Li et al., "Motor imagery BCI-based robot arm system," in *Proceedings of the 2011 Seventh International Conference on Natural Computation*, pp. 181–184, IEEE, Shanghai, China, July 2011.
- [4] D. P. Murphy, O. Bai, A. S. Gorgey et al., "Electroencephalogram-based brain-computer interface and lower-limb prosthesis control: a case study," *Frontiers in Neurology*, vol. 8, no. 696, 2017.
- [5] C. J. Bell, P. Shenoy, R. Chalodhorn, and R. P. N. Rao, "Control of a humanoid robot by a noninvasive brain-computer interface in humans," *Journal of Neural Engineering*, vol. 5, no. 2, pp. 214–220, 2008.
- [6] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [7] A. S. Royer, M. L. Rose, and B. He, "Goal selection versus process control while learning to use a brain-computer interface," *Journal of Neural Engineering*, vol. 8, no. 3, p. 036012, 2011.
- [8] Z. İşcan and V. V. Nikulin, "Steady state visual evoked potential (SSVEP) based brain-computer interface (BCI) performance under different perturbations," *PLOS ONE*, vol. 13, no. 1, Article ID e0191673, 2018.
- [9] R. Fazel-Rezai, B. Z. Allison, C. Guger, E. W. Sellers, S. C. Kleih, and A. Kübler, "P300 brain-computer interface: current challenges and emerging trends," *Frontiers in Neuroengineering*, vol. 5, no. 14, pp. 1–14, 2012.
- [10] D. J. McFarland, L. A. Miner, T. M. Vaughan, and J. R. Wolpaw, "Mu and beta rhythm topographies during motor imagery and actual movements," *Brain Topography*, vol. 12, no. 3, pp. 177–186, 2000.
- [11] R. Leeb, D. Friedman, G. R. Muller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 79642, 8 pages, 2007.
- [12] N. Birbaumer, N. Ghanayim, T. Hinterberger et al., "A spelling device for the paralysed," *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [13] J. R. Wolpaw and D. J. McFarland, "Control of a two-dimensional movement signal by a noninvasive brain-computer

- interface in humans," *Proceedings of the National Academy of Sciences*, vol. 101, no. 54, pp. 17849–17854, 2004.
- [14] G. Pfurtscheller, C. Neuper, G. R. Muller et al., "Graz-BCI: state of the art and clinical applications," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 1–4, 2003.
- [15] D. J. McFarland, D. J. Krusienski, W. A. Sarnacki, and J. R. Wolpaw, "Emulation of computer mouse control with a noninvasive brain-computer interface," *Journal of Neural Engineering*, vol. 5, no. 2, pp. 101–110, 2008.
- [16] E. V. C. Friedrich, D. J. McFarland, C. Neuper, T. M. Vaughan, P. Brunner, and J. R. Wolpaw, "A scanning protocol for a sensorimotor rhythm-based brain-computer interface," *Biological Psychology*, vol. 80, no. 2, pp. 169–175, 2009.
- [17] D. Valbuena, M. Cyriacks, O. Friman, I. Volosyak, and A. Gräser, "Brain-computer interface for high-level control of rehabilitation robotic systems," in *Proceedings of the 2007 IEEE 10th International Conference on Rehabilitation Robotics*, pp. 619–625, IEEE, Noordwijk, The Netherlands, June 2007.
- [18] B. Graimann, B. Allison, C. Mandel, T. Lüth, D. Valbuena, and A. Gräser, *Non-Invasive Brain-Computer Interfaces for Semi-Autonomous Assistive Devices*, Springer London, London, UK, 2008.
- [19] S.-Y. Cheng and H.-T. Hsueh, *Mental Fatigue Measurement Using EEG*, Intech Open, London, UK, 2011.
- [20] J. B. Isreal, G. L. Chesney, C. D. Wickens, and E. Donchin, "P300 and tracking difficulty: evidence for multiple resources in dual-task performance," *Psychophysiology*, vol. 17, no. 3, pp. 259–273, 1980.
- [21] A. Murata, A. Uetake, and Y. Takasawa, "Evaluation of mental fatigue using feature parameter extracted from event-related potential," *International Journal of Industrial Ergonomics*, vol. 35, no. 8, pp. 761–770, 2005.
- [22] J. N. Mak, D. J. McFarland, T. M. Vaughan et al., "EEG correlates of P300-based brain-computer interface (BCI) performance in people with amyotrophic lateral sclerosis," *Journal of Neural Engineering*, vol. 9, no. 2, Article ID 026014, 2012.
- [23] X. Perrin, R. Chavarriaga, F. Colas, R. Siegwart, and J. D. R. Millán, "Brain-coupled interaction for semi-autonomous navigation of an assistive robot," *Robotics and Autonomous Systems*, vol. 58, no. 12, pp. 1246–1255, 2010.
- [24] D. Göhring, D. Latotzky, M. Wang, and R. Rojas, "Semi-autonomous car control using brain-computer interfaces," in *Intelligent Autonomous Systems 12 of Advances in Intelligent Systems and Computing*, S. Lee, H. Cho, K.-J. Yoon, and J. Lee, Eds., vol. 2, pp. 393–408, Springer, Berlin, Germany, 2013.
- [25] D. Hildenbrand, D. Fontijne, Y. Wang, M. Alexa, and L. Dorst, *Competitive Runtime Performance for Inverse Kinematics Algorithms Using Conformal Geometric Algebra*, EUROGRAPHICS, Aire-la-Ville, Switzerland, 2006.
- [26] M. A. Ramírez-Moreno and D. Gutiérrez-Ruiz, "Modeling a robotic arm with conformal geometric algebra in a brain-computer interface," in *Proceedings of the 2018 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, pp. 11–17, IEEE, Cholula, Mexico, February 2018.
- [27] M. A. Ramírez-Moreno, S. M. Orozco-Soto, J. M. Ibarra-Zannatha, and D. Gutiérrez-Ruiz, "Artificial vision algorithm for object manipulation with a robotic arm in a semi-autonomous brain-computer interface," in *Wearable Robotics: Challenges and Trends of Biosystems & Biorobotics*, S. Maria, S. Carrozza, and J. L. Pons, Eds., vol. 22, pp. 187–191, Springer, Cham, Switzerland, 2019.
- [28] M. W. Spong, S. Hutchinson, and M. Vidyasagar, "Robot modeling and control," in *Chapter Forward and Inverse Kinematics*, pp. 85–98, John Wiley & Sons Inc, Hoboken, NJ, USA, 1st edition, 2005.
- [29] O. Carbajal-Espinosa, L. González-Jiménez, J. Oviedo-Barriaga, B. Castillo-Toledo, A. Loukianov, and E. Bayro-Corrochano, "Modeling and pose control of robotic manipulators and legs using conformal geometric algebra," *Computación Y Sistemas*, vol. 19, no. 3, pp. 475–486, 2015.
- [30] D. Hildenbrand, J. Zamora, and E. Bayro-Corrochano, "Inverse kinematics computation in computer graphics and robotics using conformal geometric algebra," *Advances in Applied Clifford Algebras*, vol. 18, no. 3-4, pp. 699–713, 2008.
- [31] C. Perwass, "Geometric algebra with applications in engineering, geometry and computing," in *Chapter Introduction*, pp. 1–23, Springer-Verlag, Berlin, Germany, 2009.
- [32] A. L. Kleppe and O. Egeland, "Inverse kinematics for industrial robots using conformal geometric algebra," *Modeling, Identification and Control: A Norwegian Research Bulletin*, vol. 37, no. 1, pp. 63–75, 2016.
- [33] L. Griggs and F. Fahimi, "Introduction and testing of an alternative control approach for a robotic prosthetic arm," *The Open Biomedical Engineering Journal*, vol. 8, no. 1, pp. 93–105, 2014.
- [34] E. I. Organick, *A FORTRAN IV Primer*, Addison-Wesley, Boston, MA, USA, 1st edition, 1966.
- [35] E. Dubrofsky, *Homography Estimation*, Master's thesis, The University of British Columbia, Vancouver, Canada, 2009.
- [36] Y. Renard, F. Lotte, G. Gibert et al., "OpenViBE: an open-source software platform to design, test, and use brain-computer interfaces in real and virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 19, no. 1, pp. 35–53, 2010.
- [37] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 2, pp. R1–R13, 2007.
- [38] A. Uetake and A. Murata, "Assessment of mental fatigue during VDT task using event-related potential (P300)," in *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, pp. 235–240, IEEE, Osaka, Japan, September 2000.
- [39] P. J. E. Attwell, S. F. Cooke, and C. H. Yeo, "Cerebellar function in consolidation of a motor memory," *Neuron*, vol. 34, no. 6, pp. 1011–1020, 2002.
- [40] P. Praamstra, L. Boutsen, and G. W. Humphreys, "Frontoparietal control of spatial attention and motor intention in human EEG," *Journal of Neurophysiology*, vol. 94, no. 1, pp. 764–774, 2005.
- [41] J. R. Naranjo, A. Brovelli, R. Longo, R. Budai, R. Kristeva, and P. P. Battaglini, "EEG dynamics of the frontoparietal network during reaching preparation in humans," *NeuroImage*, vol. 34, no. 4, pp. 1673–1682, 2007.
- [42] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 145–147, 2003.
- [43] J. Jin, Y. Miao, I. Daly, C. Zuo, D. Hu, and A. Cichocki, "Correlation-based channel selection and regularized feature optimization for MI-based BCI," *Neural Networks*, vol. 118, pp. 262–270, 2019.

- [44] J. K. Feng, J. Jin, I. Daly, J. Zhou, X. Wang, and A. Cichocki, "An optimized channel selection method based on multi-frequency CSP-rank for motor imagery-based BCI system," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 8068357, 10 pages, 2019.
- [45] J. A. Gaxiola-Tirado, R. Salazar-Varas, and D. Gutiérrez, "Using the partial directed coherence to assess functional connectivity in electroencephalography data for brain-computer interfaces," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 3, pp. 776–783, 2018.
- [46] F. Lotte, L. Bougrain, A. Cichocki et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, Article ID 031005, 2018.
- [47] P. Shenoy, M. Krauledat, B. Blankertz, R. P. Rao, and K.-R. Müller, "Towards adaptive classification for BCI," *Journal of Neural Engineering*, vol. 3, no. 1, pp. R13–R23, 2006.
- [48] S. Guan, K. Zhao, and S. Yang, "Motor imagery EEG classification based on decision tree framework and Riemannian geometry," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 5627156, 13 pages, 2019.
- [49] E. Donchin, M. Kubovy, M. Kutas, R. Johnson, and R. I. Tterning, "Graded changes in evoked response (P300) amplitude as a function of cognitive activity," *Perception & Psychophysics*, vol. 14, no. 2, pp. 319–324, 1973.
- [50] R. Verleger and K. Śmigajewicz, "Do rare stimuli evoke large P3S by being unexpected? A comparison of oddball effects between standard-oddball and prediction-oddball tasks," *Advances in Cognitive Psychology*, vol. 12, no. 2, pp. 88–104, 2016.
- [51] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface," *Biological Psychology*, vol. 102, pp. 118–129, 2014.
- [52] I. N. Angulo-Sherman and D. Gutiérrez, "A link between the increase in electroencephalographic coherence and performance improvement in operating a brain-computer interface," *Computational Intelligence and Neuroscience*, vol. 2015, Article ID 824175, 11 pages, 2015.

Review Article

Advances in Hybrid Brain-Computer Interfaces: Principles, Design, and Applications

Zina Li, Shuqing Zhang, and Jiahui Pan 

South China Normal University, Guangzhou 510631, China

Correspondence should be addressed to Jiahui Pan; panjh82@qq.com

Received 20 June 2019; Revised 9 September 2019; Accepted 17 September 2019; Published 8 October 2019

Guest Editor: Hyun J. Baek

Copyright © 2019 Zina Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Conventional brain-computer interface (BCI) systems have been facing two fundamental challenges: the lack of high detection performance and the control command problem. To this end, the researchers have proposed a hybrid brain-computer interface (hBCI) to address these challenges. This paper mainly discusses the research progress of hBCI and reviews three types of hBCI, namely, hBCI based on multiple brain models, multisensory hBCI, and hBCI based on multimodal signals. By analyzing the general principles, paradigm designs, experimental results, advantages, and applications of the latest hBCI system, we found that using hBCI technology can improve the detection performance of BCI and achieve multidegree/multifunctional control, which is significantly superior to single-mode BCIs.

1. Introduction

Brain-computer interface (BCI) is a technology that translates signals generated by brain activity into control signals without the involvement of peripheral nerves and muscles and uses these signals to control external devices [1]. In recent years, BCI has attracted increasing attention from academia and the public due to its potential clinical application. For example, BCI can provide augmented or repaired motor function, which can be of great help to patients with severe motor impairment. The most commonly used methods of extracting brain signals are nonimplanting, including functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), electroencephalography (EEG), and functional near-infrared spectroscopy (fNIRS) [2]. Although EEG has low signal-to-noise ratio and spatial resolution, it has been widely used in BCI because of its noninvasiveness, portability, low cost, good performance, real-time response, and technical requirements lower than other brain signals. This paper mainly describes the BCI based on EEG. Brain models used in EEG-based hybrid BCIs typically include the P300 visual-evoked potential proposed by Farwell and Donchin in 1988 [3], the steady-state-evoked

potential (such as the steady-state visual-evoked potential (SSVEP)) [4] and event-related desynchronization/synchronization (ERD/ERS) generated by motor imagination (MI) [5].

Conventional EEG-based BCI generally relies solely on a single-signal input (such as EEG, electromyography (EMG), and electro-oculogram (EOG)), single sensory stimulus (such as visual only, auditory only, and tactile only), or single brain pattern (such as the above P300 potential and SSVEP). The single-mode BCI system has achieved great progress in paradigm design, brain signal processing algorithms, and applications. However, these BCI systems have been facing multiple challenges, including low information transfer rates (ITRs), low man-machine adaptability, and high dynamics/nonstationarity of brain signals [6, 7]. Here, we mainly consider two fundamental challenges and introduce a hybrid BCI technique intended to address these challenges:

- (1) Multidegree/multifunction control: multidegree/multifunctional control is necessary for many devices, such as wheelchair, robots, or artificial limbs. For instance, the wheelchair control includes speed, direction, and start/stop functions. However, it is

- difficult for a conventional simple BCI to generate effective multiple control signals [8].
- (2) Improvement of detection performance: over the years, although many efforts have been made to improve the detection performance of BCI, the detection performance in terms of classification accuracy, information transfer rate (ITR), and false-positive rate (FPR) is still far from practice in many applications, especially for patients. Approximately 13% of healthy users suffer from BCI illiteracy and do not reach the criterion for controlling a BCI application [9]. Moreover, user acceptability and complexity of the BCI systems should be reported as important performance criteria.

To conquer the above two fundamental challenges, some researchers have proposed a hybrid BCI (hBCI). As described by Allison [8], an hBCI system consists of a BCI system and an add-on system, which can be a second BCI system, but designed to perform specific goals better than a conventional BCI. The main goal of hBCI is to overcome the existing limitations and disadvantages of the conventional BCI systems. In this paper, the recent progress in hBCIs was reviewed to illustrate how hBCI techniques could be implemented to address these challenges. The definition of hybrid BCIs was updated and extended, and three main types of hBCIs have been devised. For each type of hybrid BCIs, the principle was summarized and several representative hybrid BCI systems were highlighted by analyzing their paradigm designs, control methods, and experimental results. Finally, the future prospect and research direction of hBCI were discussed.

2. Hybrid BCI Overview

Although the concept of hBCI emerged before 2010, its development has become more and more rapid in recent years. Based on the search engine “Web of Science,” and title-abstract-keyword (“brain-computer interface” or “BCI”) and (“hybrid” or “multimodal”), the number of journal papers found before 2010 was only three. However, this number rose to 148 and 293 in the two periods of 2010–2014 and 2015–2019, respectively. It is evident that the number of publications on hBCI has grown rapidly in recent years. Note that those studies of single BCI combining only features and algorithms also can improve performance are excluded. In fact, “Hybrid BCI” and “multimodal BCI” are two highly related concepts. Li et al. [9] even considered that “hybrid BCI” and “multimodal BCI” to be interchangeable terms with the same BCI definition.

Pfurtscheller et al. [10] believed that in addition to the simple combination of different BCIs, the type of hBCI should meet the following four criteria: (1) the activity comes directly from the brain; (2) at least one brain signal acquisition method should be used to capture this activity, and the brain signal acquisition method can be in the form of electrical, magnetic, or hemodynamic changes; (3) the signal must be processed in real time/online to establish

communication between the brain and the computer to generate control commands; (4) feedback must be provided according to the results of brain activity for communication and control.

The signal flow of an hBCI system is as described in Figure 1, which includes two stages of brain signal processing. (1) In the signal acquisition, the signal input can be from multiple signals (e.g., EEG and NIRS) or multiple brain patterns (e.g., P300 and SSVEP), which are evoked by multisensory stimuli (e.g., audiovisual stimuli). (2) In the signal processing, an hBCI system can provide only a single-output/control signal or multiple-output/control signals. In the former case, when multiple brain patterns or multiple signals are involved, data fusion is generally required at the feature or decision level. In the latter case, multiple control signals may be separately manipulated by different brain patterns detected by the system, and the fusion of these brain patterns is generally not necessary. As shown in Figure 1, the hBCI can be divided into three main categories:

- (1) hBCI based on multiple brain patterns: it uses at least two brain modes (e.g., P300 and SSVEP or MI and P300). In this type of hBCI, multiple brain patterns are induced by a single sensory stimulus. Several studies have indicated that hybrid integration associated with multimodal stimuli has the potential to enhance brain patterns, which may be beneficial for BCI performance [11].
- (2) hBCI with multisensory stimuli: its brain pattern is simultaneously induced by multiple sensory stimuli, such as audiovisual stimuli. In this hBCI, one or more brain patterns are induced by multisensory stimuli. Some researchers believed multisensory BCIs may offer more versatile and user-friendly paradigms for control and feedback [12].
- (3) hBCI based on multiple signals: in this hBCI, two or more input signals are typically combined with a hybrid BCI system, such as EEG, MEG, fMRI, fNIRS, EOG, or EMG. Different brain signals have different signal characteristics and can be used for different functions.

The state-of-the-art of the above three types of hBCI is introduced in the following sections, including their general principles, stimuli paradigm, control methods, corresponding experimental results, and advantages.

3. hBCI Based on Multiple Brain Patterns

The first class of hBCIs combines multiple brain patterns, such as P300, SSVEP, and MI. It has been designed for a variety of applications, such as speller [13], idle state detection [14], orthotics [15], the wheelchair navigation, and control of computer components, which include two-dimensional (2D) cursor [16], mouse [17], or mail client [18]. Table 1 lists the representative hBCI applications of multiple brain patterns in recent years. In this section, we mainly describe hBCI based on P300 and SSVEP, hBCI based on MI and SSVEP, and hBCI based on MI and P300.

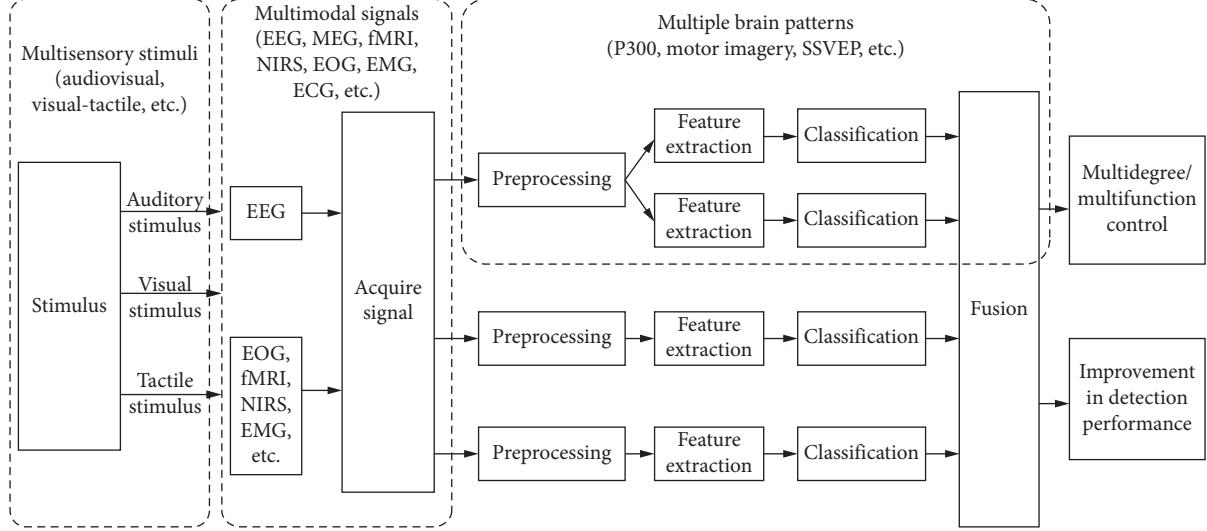


FIGURE 1: The signal flow of hybrid brain-computer interface discussed in this paper.

3.1. P300- and SSVEP-Based hBCIs. Both P300 potential and SSVEP can be elicited by visual stimuli, allowing subjects to evoke both brain patterns by performing a visual attention task without extra mental load. The P300 and SSVEP features are located in different domains (time domain versus frequency domain), and both brain patterns have significant independence. The improvement in performance may result from the utilization of both P300 and SSVEP features. The addition of the EEG feature may provide additional information that facilitates the classification of a target versus a nontarget.

Bi et al. [22] proposed a hybrid paradigm based on SSVEP and P300 for developing speed-direction-based cursor control. In this study, the stimulation of the P300 was distributed on the upper and lower sides of the screen, and the stimulus for detecting SSVEP (which can rotate the control device clockwise or counterclockwise) was displayed on the left and right sides of the screen. The results using the method based on the support vector machine classification showed that the accuracy of the hBCI was higher than 90%.

Pan et al. [29] detected consciousness in eight patients with disorders of consciousness (DOC) by using a hybrid paradigm of SSVEP and P300. Following the instructions, the left- and right-hand photos flickered on a black background with fixed frequencies of 6.0 and 7.5 Hz, respectively, to evoke the patient's SSVEP. Meanwhile, each of the two photo frames was randomly presented five times to evoke P300, with each appearance lasting 200 ms and the interval between two consecutive appearances being 800 ms. The BCI system used the characteristics of P300 and SSVEP to detect which photo the patient had noticed. Eight patients (four in the vegetative state (VS), three in the minimally conscious state (MCS), and one in the locked-in syndrome (LIS)) participated in the experiment. Using the SVM-based classifier, one VS patient, one MCS patient, and one LIS patient were able to select photos of themselves or others (classification accuracy, 66%–100%), which indicates that the patient command can be followed using an hybrid BCI and further proves that they have certain cognitive abilities and awareness.

3.2. MI- and SSVEP-Based hBCIs. There are four reasons to combine SSVEP and MI: (1) SSVEP- and MI-related brain patterns were produced simultaneously; (2) SSVEP is an evoked potential that can be stably detected in unfamiliar subjects with little training, but for most new users, it is difficult to adapt to the process of completing MI task; (3) SSVEP can detect by a single trial based on EEG data, and the detection does not require an averaging process; (4) nonvisual training will frustrate subjects, while SSVEP provides a possible solution to attract subjects to participate in MI task.

Based on the above principles, Yu et al. [26] combined SSVEP and MI to provide effective continuous feedback for MI training in 24 subjects. Initially, the classifier assigns a greater weight to the SSVEP in order to get the correct feedback at the beginning of the training. As the training goes on, participants reduced their visual attention to SSVEP stimuli but maintained sustained attention to MI mental tasks. When subjects adapt to rhythmic activities, the classifier shifts the weight to MI. The result showed that an hBCI can be used to improve MI training and produce distinguishable brain patterns after only five sessions (about 1.5 hours).

3.3. MI- and P300-Based hBCIs. An important aspect of the EEG-based BCI system is multidimensional control, which involves multiple independent control signals. These control signals can be obtained from multiple brain patterns, such as MI and P300. P300 represents the reliable type of brain pattern used to generate discrete control output commands, and MI is more effective against generating sequential control commands.

Li and colleagues [16] proposed hBCI combining MI brain patterns and P300 potentials for 2D cursor control and target selection. The GUI is shown in Figure 2, in which the circle and square represent the cursor and target, respectively, with the initial position of the cursor and the initial position and color (green or blue) of the target are

TABLE 1: Representative hBCI applications of multiple brain patterns.

Reference	Hybrid mode	Application	Classifiers	Commands	Accuracy (%)	Improvements
[19]	SSVEP, P300, MI	Humanoid machine navigation	CCA	6	P300: 84.6, SSVEP: 84.1	Better commands performance in navigation and exploration
[20]	SSVEP, P300	Wheelchair control with stop command	SVM	2	>80	Higher detection accuracy and low response time
[21]	SSVEP, P300	Target selection speller	SW-LDA	9	93.3	More effective in target discrimination
[22]	SSVEP, P300	Cursor control	SVM	9	>90	Higher accuracy and better commands performance
[11]	SSVEP, P300	Multiple option selection	CCA, LDA	4	P300: 99.9 SSVEP: 67.2	Better performance and user-friendly
[23]	P300, SSVEP	Speller	SW-LDA	36	93.85	Higher accuracy
[24]	MI, SSVEP	Play Tetris games in MI-SSVEP paradigm	LDA, CSP, CCA	4	MI: 87.01 SSVEP: 90.26	Higher accuracy
[25]	MI, SSVEP	Hybrid BCI system of MI and SSVEP	LDC	2	85.6 ± 7.7	Better classification performance
[9]	MI, SSVEP, visual, and auditory	Wheelchair control	SVM	6	—	Multidegree control commands
[26]	MI, SSVEP	Hybrid BCI system with feedback	LDA	2	≥83	Better MI training performance
[27]	SSVEP, MI	Control commands	CCA	5	MI: 93.3 SSVEP: 89	Better performance and easiness for users
[16]	MI, P300	2-D cursor control	SVM	2	>80	Multiple-degree control
[17]	P300, MI	BCI mouse-based web browser	SVM	3	93.21	Multidegree control with a feasible BCI mouse
[28]	P300, MI	BCI wheelchair with direction and speed control	LDA	4	83.10 ± 2.12	Direction and speed control

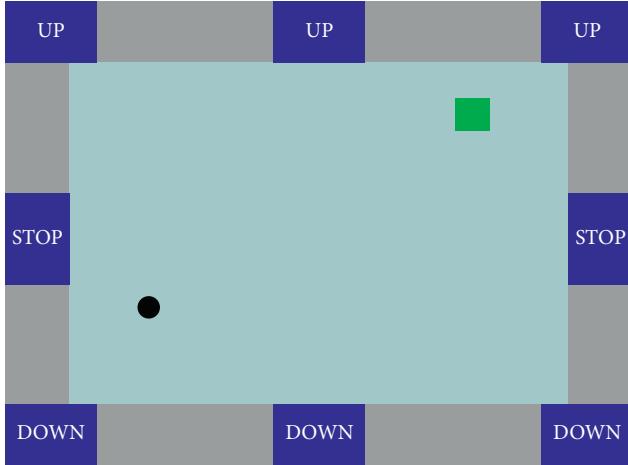


FIGURE 2: GUI of 2D cursor control and target selection of a hBCI system [16], which combines MI and P300 potential, including one cursor (black circle), one object (green square), and eight flashing buttons (three “UP,” three “DOWN,” and two “STOP” buttons).

randomly provided. The three “UP” buttons, three “DOWN” buttons, and two “STOP” buttons flash in a random order to evoke P300 potentials. The task of the user is to move the cursor to the target and then to select or reject the green/blue target. The control strategy of the user is described below. The user can move the cursor to the left or right by imagining his or her own left- or right-hand movement, respectively, and the user can move the cursor up or down by focusing on one of the three flashing “UP” or “DOWN” buttons to evoke

P300 potentials. If the user does not intend to move the cursor in the vertical direction, then the user can focus on one of the two “STOP” buttons.

To further implement a BCI mouse, target selection and rejection functions are required. Specifically, once the cursor hits the target of interest (green square), the user can select the target by focusing the attention on a flashing “STOP” button and simultaneously maintaining an idle state of motor imagery. If the target is not of interest (blue square), the user can reject it by continuing to imagine left- or right-hand movement without focusing on any flashing buttons.

The algorithm for the 2D cursor control includes two parts: P300 detection for vertical movement control and motor-imagery detection for horizontal movement control, with the details presented in [19]. The signal processing procedure for P300 detection consists of three stages: low-pass filtering, P300 feature extraction, and SVM classification. For motor-imagery detection, the signal processing stages include common average reference (CAR) spatial filtering, band-pass filtering of the specific mu rhythm band (8–13 Hz), feature extraction based on a CSP algorithm, and SVM classification. The algorithm for target selection or rejection was based on the hybrid features of P300 potentials and MI. After extracting the features of the P300 potentials and MI using the same algorithms described above, a hybrid feature vector for each trial is constructed by concatenating the feature vector of the MI with the feature vector of the P300 potentials, which is then fed into the SVM for classification.

Eleven healthy subjects attended the online experiment, which included one session of 80 trials for each subject. Each trial included two sequential tasks. During the first task, subjects were instructed to move the cursor to a target that was presented at a randomized position on the screen. After the cursor hit the target, the subject was instructed to perform the second task of selecting or rejecting the target according to the color of the target (green for selection and blue for rejection). The time interval for the second task was set to 2 s. Among all subjects, the average time for one trial was 18.96 s, the average accuracy for successful trials was 92.84%, and the average for target selection accuracy given that the cursor was successfully moved to the target was 93.99%. Additionally, several datasets were also collected for offline analysis to demonstrate the advantage of P300 potential and MI hybrid features for target selection/rejection compared with the use of P300 potential or MI features alone. The experimental results showed that the accuracy for use of the hybrid features was significantly higher than for use of only the MI or P300 potential features (hybrid features: $83.10 \pm 2.12\%$; MI features: $71.68 \pm 2.41\%$; P300 features: $80.44 \pm 1.82\%$). Based on the BCI cursor, Long et al. [28] proposed a hybrid BCI paradigm based on MI and P300 potential to operate actual wheelchairs by providing direction (left or right) and speed control (acceleration and deceleration) commands with 5 subjects.

All of these hybrid systems have three advantages. First, two independent control signals are generated based on MI and P300 potential. Second, the user can move the cursor from any position to a randomly located target. Third, the hybrid control strategy using MI and P300 potential provides better identification performance than the control strategy using MI-only or P300-only.

4. Multisensory hBCIs

Humans have multiple senses that provide pathways for processing information on the reality. The integration of multiple sensory stimuli enhances top-down attention, and these enhanced effects may be conducive to improve the performance of BCI systems. Taken into this consideration, hBCI based on audiovisual and visual-tactile was proposed, in which bimodal stimulation was used to improve system performance. Table 2 lists the representative applications of multisensory hBCIs in recent years.

4.1. Audiovisual hBCIs. Belitski et al. [30] proposed an offline audiovisual-based P300 speller and corresponding data analysis results. Their study of 7 healthy subjects showed that the intensity of P300 reaction was higher in audiovisual conditions than in visual or auditory conditions alone. Similarly, An et al. [32] explored parallel spellers for BCI unrelated to gaze for healthy subjects, where the auditory and visual domains are independent of each other. Their results showed that 15 users can spell online, with an average accuracy rate of 87.7%. These existing results suggest that audiovisual integration may

be a potential way to enhance brain patterns and further improve BCI performance. Wang et al. [33] proposed a novel audiovisual BCI system, which is based on time-synchronous visual and auditory stimuli. In the GUI of this audiovisual BCI, there are two number buttons (two numbers randomly drawn from 0 to 9) located on the left and right sides, and two speakers are placed laterally to the monitor. The two buttons flash in an alternative manner. When a number button is visually intensified, the corresponding spoken number is presented from the ipsilateral speaker. In this way, the user is presented with a temporally, spatially, and semantically congruent audiovisual stimulus that lasts for 300 ms, where the interstimulus interval is randomized from 700 to 1500 ms. Ten healthy subjects participated in the experiment. The experiment consisted of three sessions administered in a random order, corresponding to the visual-only, auditory-only, and audiovisual conditions. In each session, the subject first performed a training run of 10 trials and then a test run of 30 trials. The online average accuracy of audiovisual, visual-only, and auditory-only sessions for all healthy subjects was 95.67%, 86.33%, and 62.33%, respectively. The audiovisual BCI significantly outperformed the visual-only and auditory-only BCIs. This audiovisual hBCI system was then applied to the consciousness detection of 7 patients with DOC. The experimental results indicated that the audiovisual BCI can provide more sensitive results than the behavioral observation scale.

4.2. Audio-Tactile hBCIs. The above bimodal BCI requires visual interaction to focus on stimuli and feedback, which limits their applicability to users with good vision and complete gaze control. Since the user does not require visual interaction when operating auditory or tactile BCI, a bimodal auditory/tactile-based manner may allow visual scanning of unrelated BCI. Yin et al. [34] proposed a dual-mode P300 BCI with the same direction, which was presented simultaneously with auditory and tactile stimuli from the same spatial direction. Rutkowski and Mori [35] studied the tactile and auditory BCI of 11 users with vision and hearing impairment.

These existing results reveal the several advantages of BCI auditory-tactile. First, the auditory-tactile dual-mode BCI has better overall system performance than the auditory or tactile single-mode P300 BCI. Second, in visual computer applications, auditory-tactile hBCI offers an attractive possibility of target sensory fields that can induce potential without relying on visual stimuli, although the performance achieved by using this system is lower than that of BCI dependent on gaze transfer. Third, visual-tactile hBCI is an alternative for users with impaired vision.

5. hBCI Based on Multimodal Signals

hBCI systems can be constructed using multimodal signals, including EEG, MEG, fMRI, EOG, fNIRS, and EMG. Different brain signals have different signal characteristics and

TABLE 2: Representative applications of multisensory hBCIs.

Reference	Hybrid mode	Application	Classifiers	Commands	Accuracy (%)	Improvements
[30]	P300, visual, audio	P300 audiovisual speller	Regularized linear LR	—	>80	Improvement in performance
[31]	Visual, audio	Consciousness detection in patients with DOC	SVM	2	>64	Better performance and feasible to patients with DOC
[32]	Visual, audio	Visual-auditory speller	LDA	30	87.7 (chance level <3%)	Better BCI performance
[33]	Visual, audio	Awareness detection	SVM	2	95.67	Better performance over auditory-only and visual-only systems
[34]	Auditory, tactile, visual, P300	Visual saccade-independent BCI	BLDA	4	88.67	Better online performance
[35]	Auditory, tactile, P300	Tactile and bone-conduction BCI	SW-LDA	6	70	Higher classification accuracy
[36]	Audio, tactile	Robot gesture	FGMMs, SVM	10	92.75	Better performance over framework

can be used for different functions. Recently, several hybrid BCIs based on multiple signals have been reported in the following. Table 3 lists the representative hBCI applications based on multimodal signals in recent years.

5.1. EEG- and EMG-Based hBCIs. Leeb et al. [50] proposed an hBCI combining EEG and EMG. In each trial, 12 healthy subjects were instructed to repeat the exercise for five seconds with their left or right hand (holding the hand with the fist) based on visual cues (arrows to the left or right). The researchers processed and classified EEG and EMG signals separately and then fused them. Canonical variable analysis was used to select subject-specific features that maximized separability between different tasks, and stable features were determined by cross validation of a Gaussian classifier based on training data. The resulting features were given threshold, normalized, and classified based on maximum distance in a subject-specific manner. Finally, the Bayesian method was used to fuse the probabilities of two classifiers to generate a control signal. The accuracy of a single EEG activity was 73% and single EMG activity was 87%. However, the accuracy of the hBCI was improved to 91%. In addition, to simulate tired muscles, the amplitude of the EMG channel decreased during operation (from 10% to 100%), and EEG activity is increasingly important in fused data as EMG muscles become more tired. The results showed a significant advantage for EEG- and EMG-based BCI systems.

5.2. EEG- and EOG-Based hBCIs. Recently, some studies have combined EEG and EOG to construct an hBCI. Since many people with disabilities are able to control their eye movements, EOG signals are an appropriate choice for many users of the BCI system. Lee et al. [41] employed hBCIs based on EEG-EOG to a speller system with fast typing speed. The hBCI system comprised a conventional ERP-based speller, an EOG-based command detector, and a visual feedback module. The online ERP speller was used to compute the classification probabilities for all candidate

characters from EEG epoch. The character of highest probability was selected as visual feedback based on the probabilities sorting. The accuracy of the novel speller system was 97.6%, and its ITR is 39.6 ± 13.2 bits/min across 20 participants. The result showed that this EEG- and EOG-based speller has better performance than the conventional ERP-based speller.

5.3. Other hBCIs Based on Multimodal Signals. Other hybrid BCIs based on multiple signals have also been reported. A way to make full use of the spatial and temporal information of brain activity is to combine the fMRI with EEG in BCIs. A key advantage of EEG-fMRI combined BCI is that EEG can provide online slow cortical potential (SCP) feedback to subjects. It also reveals the basic psychophysiological mechanisms, such as the correlation between local BOLD-responses and the SCP changes, which helps to develop new training procedures and paradigms. Although fNIRS has poor spatial resolution compared to fMRI, it is portable and reflects the hemodynamic response of brain activity.

The authors in [45] have proved that the performance of an MI-based BCI was improved significantly by combining EEG and NIRS. It allows those who are unable to run EEG-based BCI alone to achieve meaningful classification rates. EEG is easily distorted by the inhomogeneities of the extracerebral tissues, while MEG is not affected as long as the electric inhomogeneities are concentric. Therefore, MEG signals are more local than the corresponding EEG signals and can provide more spatial information. In [47], the MEG and EEG signals generated in the sensorimotor cortex were used to index the finger movements for three tetraplegics.

6. Discussion and Conclusion

This paper focuses on several hBCI types and different stimulus designs and their performance analysis. To begin with, we summarized three classes of hBCIs: hBCIs based on multiple brain patterns, multisensory hBCIs, and hBCIs

TABLE 3: Representative applications of hBCI of multimodal signals.

Reference	Hybrid mode	Application	Classifiers	Commands	Accuracy (%)	Improvements
[37]	EMG, EEG	A motor imagery hybrid BCI speller	GMM	2	End-users: 91 Able-bodied users: 94	Better performance over command accuracy
[38]	EEG, EMG	Home environmental control system	CCA	4	96.3	Higher control accuracy, security, and interactivity
[39]	EEG, EOG	AIDS recovery	AR	4	62.28	Substantially better control over assistive devices
[40]	EEG, EOG	Mobile robot control	LDA	9	87.3	Reduce the best completion time
[41]	EEG, EOG	Hybrid speller system	LDA	1	97.6	Better performance and usability
[42]	fNIRS, EEG, eye movement	Control a quadcopter online	LDA	8	fNIRS: 75.6 EEG: 86	Higher accuracy on decoding
[43]	EEG, fNIRS	Hand movement and recognition	LDA	2	94.2	Reduce fNIRS delay time in detection
[44]	EEG, fNIRS	Left- and right-hand motion imagination	DL	2	—	Reduce response time
[45]	EEG, NIRS	Decoding of four movements	LDA	5	>80	Higher classification accuracy
[46]	EEG, NIRS	Mental state recognition	Meta	6	65.6	Better performance on mental states classification
[47]	EEG, MEG	Left- and right-hand motor imagery	CSP, LR	2	MEG: 70.6 EEG: 67.7	Better performance over good within-subject accuracy
[48]	EEG, NIRS	Classification of mental arithmetic, MI, and idle state	sLDA	3	82.2 ± 10.2	Higher classification accuracy
[49]	EEG, MEG	Intersubject decoding of left- vs. right-hand motor imagery	LR, L2, 1-norm regularization	4	MEG: 70 EEG: 67.7	Higher within-subject accuracy

based on multimodal signals. For each type of hBCIs, we reviewed several representative hybrid BCI systems, including their design principles, stimuli paradigms, control methods, experimental results, and corresponding advantages. In the following, we will elaborate concluding remarks regarding the benefits of hybrid BCI systems and future studies.

Following consideration of the three types of hybrid BCI and their respective applications, we can summarize the advantages of hybrid BCI in two aspects. First, the hBCI system can provide only a single control signal or output to improve the classification performance. The two main strategies for bringing about these improvements are as follows: (1) the combination of multiple brain patterns (such as MI, P300, and SSVEP) or the fusion of multiple signals (such as EEG, EMG, EOG, and NIRS) can be performed at the feature level; and (2) enhancing brain patterns by presenting multiple sensory stimuli, such as audiovisual stimuli. Second, when multiple control signals or outputs are available, hBCI systems attempt to implement multi-degree object control. In this paper, the multi-dimensional or functional control method based on hybrid BCIs and some application systems are presented. Two main methods can be adopted: (1) combining multiple brain patterns to obtain multiple independent control signals, such as 2D cursor control based on MI and P300 and orthopedic control based on MI and SSVEP; (2) using different signal characteristics to perform different functions, such as robot control based on EEG and EOG.

Here, we consider several challenging problems for further study.

6.1. Design and Implementation for hBCIs. From the user's point of view, the complexity of the hBCI system is usually higher than that of the conventional simple BCI. User acceptability is an important performance criterion that needs to be carefully considered in hBCI design and implementation. In the design of an hBCI based on brain patterns, one of the challenges is how to determine the best combination of brain patterns to achieve the desired goals, and the combination can vary from user to user. For example, it should be considered that long-term use of SSVEP and P300 will increase visual fatigue. While designing a couple sensory hBCI, the challenge is to ensure that the desired brain patterns are enhanced by multiple sensory stimuli. Previous studies [33] have found that combining audio stimuli with natural spoken words in a visual P300-based BCI can help reduce the burden of mental work. Therefore, we can consider more combinations of multiple sensory stimuli involving auditory and tactile patterns in future research. For the hBCI based on multiple signals, one challenge is how to make full use of the characteristics of different signals to achieve the greatest improvement in system performance. In addition, when designing the real-time hBCI based on EEG and fMRI, the high noise, slow response and high dimensionality of EEG data (generated by fMRI scanner), and the low temporal resolution of fMRI data are not negligible.

6.2. Brain Mechanisms for hBCIs. The hBCI system may involve multiple brain modes, multiple sensory modes, or multimode signal inputs. To ensure that these components are effectively coordinated in the hBCI system, it is necessary

to study the relevant brain mechanisms. For example, cross-modal integration/interaction in the brain can provide a brain mechanism for multisensory BCI. However, there have been few studies on the brain mechanism of hBCI so far.

6.3. Clinical Application. Until now, most hBCI systems (such as BCI browsers and BCI wheelchairs) were designed for healthy subjects. It needs to be extended to patients and extend their value to clinical applications. In recent years, more and more hBCIs have been used in clinical applications, such as in the rehabilitation and treatment of patients with hemiplegia [51, 52] and DOC [53]. When designing these hBCI systems for patients, the differences between them and healthy subjects need to be fully considered. In some cases, even a single patient design is necessary. The application of hBCI to patients with DOC has just begun, and hBCI-based communication and rehabilitation is an important topic for our future research. In addition, a variety of intelligent technologies, such as automatic navigation systems and intelligent robots, have been combined with BCI. This combination not only greatly reduces the user's workload but also makes the BCI system more reliable, flexible, and powerful by allowing the subject to focus on the final goal and to ignore the low-level details associated with the execution of the action. This is promising for patients with low recognition and control capabilities. Therefore, future research should focus on such systems developed for patients.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (Grant no. 61876067), the Pearl River S and T Nova Program of Guangzhou (201710010038), and Guangdong Natural Science Foundation (Grant 2014A030310244).

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] S. Fazli, J. Mehner, J. Steinbrink et al., "Enhanced performance by a hybrid NIRS-EEG brain computer interface," *NeuroImage*, vol. 59, no. 1, pp. 519–529, 2012.
- [3] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [4] G. R. Müllerputz, R. Scherer, C. Neuper, and G. Pfurtscheller, "Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces?," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 1, pp. 30–37, 2006.
- [5] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [6] K. S. Hong and M. J. Khan, "Hybrid brain–computer interface techniques for improved classification accuracy and increased number of commands: a review," *Frontiers in Neurorobotics*, vol. 11, p. 35, 2017.
- [7] A. Vučković and F. Sepulveda, "A two-stage four-class BCI based on imaginary movements of the left and the right wrist," *Medical Engineering & Physics*, vol. 34, no. 7, pp. 964–971, 2012.
- [8] B. Z. Allison, "Toward ubiquitous BCIs," in *Brain-Computer Interfaces: Revolutionizing Human-Computer Interaction*, B. Graimann, G. Pfurtscheller, and B. Allison, Eds., pp. 357–387, Springer, Berlin, Germany, 2010.
- [9] J. Li, H. Ji, L. Cao et al., "Evaluation and application of a hybrid brain computer interface for real wheelchair parallel control with multi-degree of freedom," *International Journal of Neural Systems*, vol. 24, no. 4, Article ID 1450014, 2014.
- [10] G. Pfurtscheller, B. Z. Allison, C. Brunner et al., "The hybrid BCI," *Frontiers in Neuroscience*, vol. 4, no. 3, 2010.
- [11] B. Z. Allison, J. Jin, Y. Zhang, and X. Wang, "A four-choice hybrid P300/SSVEP BCI for improved accuracy," *Brain-Computer Interfaces*, vol. 1, no. 1, pp. 17–26, 2014.
- [12] I. C. Wagner, I. Daly, and A. Välijamäe, "Non-visual and Multisensory BCI Systems: Present and Future," *Towards Practical Brain-Computer Interfaces*, Springer, Berlin, Germany, 2012.
- [13] R. C. Panicker, S. Puthusserypaday, and Y. Sun, "An asynchronous P300 BCI with SSVEP-based control state detection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1781–1788, 2011.
- [14] Y. Li, J. Pan, J. Long et al., "Multimodal BCIs: target detection, multidimensional control, and awareness evaluation in patients with disorder of consciousness," *Proceedings of the IEEE*, vol. 104, no. 2, pp. 332–352, 2016.
- [15] G. Pfurtscheller, T. Solis-Escalante, R. Ortner, P. Linortner, and G. R. Müller-Putz, "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based "brain switch": a feasibility study towards a hybrid BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 409–414, 2010.
- [16] Y. Li, J. Long, T. Yu et al., "An EEG-based BCI system for 2-D cursor control by combining mu/beta rhythm and P300 potential," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2495–2505, 2010.
- [17] T. Yu, Y. Li, J. Long, and Z. Gu, "Surfing the internet with a BCI mouse," *Journal of Neural Engineering*, vol. 9, no. 3, Article ID 036012, 2012.
- [18] T. Yu, Y. Li, J. Long, and F. Li, "A hybrid brain-computer interface-based mail client," *Computational and Mathematical Methods in Medicine*, vol. 2013, Article ID 750934, 9 pages, 2013.
- [19] B. Choi and S. Jo, "A low-cost EEG system-based hybrid brain-computer interface for humanoid robot navigation and recognition," *PLoS One*, vol. 8, no. 9, Article ID e74583, 2013.
- [20] Y. Li, J. Pan, F. Wang, and Z. Yu, "A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 11, pp. 3156–3166, 2013.
- [21] M. Xu, H. Qi, B. Wan, T. Yin, Z. Liu, and D. Ming, "A hybrid BCI speller paradigm combining P300 potential and the SSVEP blocking feature," *Journal of Neural Engineering*, vol. 10, no. 2, Article ID 026001, 2013.
- [22] L. Bi, J. Lian, K. Jie, R. Lai, and Y. Liu, "A speed and direction-based cursor control system with P300 and SSVEP," *Bio-medical Signal Processing and Control*, vol. 14, pp. 126–133, 2014.

- [23] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain-computer interface based on the fusion of P300 and SSVEP scores," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 693–701, 2015.
- [24] Z. Wang, Y. Yu, M. Xu, Y. Liu, E. Yin, and Z. Zhou, "Towards a hybrid BCI gaming paradigm based on motor imagery and SSVEP," *International Journal of Human-Computer Interaction*, vol. 35, no. 3, pp. 197–205, 2019.
- [25] L.-W. Ko, S. S. K. Ranga, O. Komarov, and C.-C. Chen, "Development of single-channel hybrid BCI system using motor imagery and SSVEP," *Journal of Healthcare Engineering*, vol. 2017, Article ID 3789386, 7 pages, 2017.
- [26] T. Yu, J. Xiao, F. Wang et al., "Enhanced motor imagery training using a hybrid BCI with feedback," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1706–1717, 2015.
- [27] F. Duan, D. Lin, W. Li, and Z. Zhang, "Design of a multimodal EEG-based hybrid BCI system with visual servo module," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 4, pp. 332–341, 2015.
- [28] J. Long, Y. Li, H. Wang, T. Yu, J. Pan, and F. Li, "A hybrid brain computer interface to control the direction and speed of a simulated or real wheelchair," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 5, pp. 720–729, 2012.
- [29] J. Pan, Q. Xie, Y. He et al., "Detecting awareness in patients with disorders of consciousness using a hybrid brain-computer interface," *Journal of Neural Engineering*, vol. 11, no. 5, Article ID 056007, 2014.
- [30] A. Belitski, J. Farquhar, and P. Desain, "P300 audio-visual speller," *Journal of Neural Engineering*, vol. 8, no. 2, Article ID 025022, 2011.
- [31] J. Pan, Q. Xie, H. Huang et al., "Emotion-related consciousness detection in patients with disorders of consciousness through an EEG-based BCI system," *Frontiers in Human Neuroscience*, vol. 12, p. 198, 2018.
- [32] X. An, J. Höhne, D. Ming, and B. Blankertz, "Exploring combinations of auditory and visual stimuli for gaze-independent brain-computer interfaces," *PLoS One*, vol. 9, no. 10, Article ID e111070, 2014.
- [33] F. Wang, Y. He, J. Pan et al., "Erratum: a novel audiovisual brain-computer interface and its application in awareness detection," *Scientific Reports*, vol. 5, no. 1, p. 9962, 2015.
- [34] E. Yin, T. Zeyl, R. Saab, D. Hu, Z. Zhou, and T. Chau, "An auditory-tactile visual saccade-independent P300 brain-computer interface," *International Journal of Neural Systems*, vol. 26, no. 1, Article ID 1650001, 2016.
- [35] T. M. Rutkowski and H. Mori, "Tactile and bone-conduction auditory brain computer interface for vision and hearing impaired users," *Journal of Neuroscience Methods*, vol. 244, pp. 45–51, 2015.
- [36] Z. Ju and H. Liu, "Human hand motion analysis with multisensory information," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 2, pp. 456–466, 2014.
- [37] S. Perdikis, R. Leeb, J. Williamson et al., "Clinical evaluation of BrainTree, a motor imagery hybrid BCI speller," *Journal of Neural Engineering*, vol. 11, no. 3, Article ID 036003, 2014.
- [38] X. Chai, Z. Zhang, Y. Lu, G. Liu, T. Zhang, and H. Niu, "A hybrid BCI-based environmental control system using SSVEP and EMG signals," in *World Congress on Medical Physics and Biomedical Engineering 2018*, vol. 2019, Springer, Singapore, 2019.
- [39] S. R. Soekadar, M. Witkowski, N. Vitiello, and N. J. B. E. Birbaumer, "An EEG/EOG-based hybrid brain-neural computer interaction (BNCI) system to control an exoskeleton for the paralyzed hand," *Biomedical Engineering*, vol. 60, no. 3, pp. 199–205, 2015.
- [40] J. Ma, Y. Zhang, A. Cichocki, and F. Matsuno, "A novel EOG/EEG hybrid human-machine interface adopting eye movements and ERPs: application to robot control," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 3, pp. 876–889, 2015.
- [41] M. H. Lee, J. Williamson, D. O. Won, S. Fazli, and S. W. Lee, "A high performance spelling system based on EEG-EOG signals with visual feedback," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, 2018.
- [42] M. J. Khan and K. S. Hong, "Hybrid EEG-fNIRS-based eight-command decoding for BCI: application to quadcopter control," *Frontiers in Neurorobotics*, vol. 11, p. 6, 2017.
- [43] A. P. Buccino, H. O. Keles, and A. Omurtag, "Hybrid EEG-fNIRS asynchronous brain-computer interface for multiple motor tasks," *PLoS One*, vol. 11, no. 1, Article ID e0146610, 2016.
- [44] A. M. Chiarelli, P. Croce, A. Merla, and F. Zappasodi, "Deep learning for hybrid EEG-fNIRS brain-computer interface: application to motor imagery classification," *Journal of Neural Engineering*, vol. 15, no. 3, Article ID 036028, 2018.
- [45] M. J. Khan, M. J. Hong, and K. S. J. FiH. N. Hong, "Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface," *Frontiers in Human Neuroscience*, vol. 8, no. 1, p. 244, 2014.
- [46] J. Shin, A. von Luhmann, B. Blankertz et al., "Open access dataset for EEG+NIRS single-trial classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1735–1745, 2017.
- [47] H. L. Halme and L. Parkkonen, "Across-subject offline decoding of motor imagery from MEG and EEG," *Scientific Reports*, vol. 8, no. 1, p. 10087, 2018.
- [48] J. Shin, J. Kwon, and C.-H. Im, "A ternary hybrid EEG-NIRS brain-computer interface for the classification of brain activation patterns during mental arithmetic, motor imagery, and idle state," *Frontiers in Neuroinformatics*, vol. 12, no. 5, 2018.
- [49] H.-L. Halme and L. Parkkonen, "Across-subject offline decoding of motor imagery from MEG and EEG," *Scientific Reports*, vol. 8, no. 1, p. 10087, 2018.
- [50] R. Leeb, H. Sagha, R. Chavarriaga, and J. del R Millán, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities," *Journal of Neural Engineering*, vol. 8, no. 2, Article ID 025011, 2011.
- [51] M. Hassan, H. Kadone, T. Ueno, Y. Hada, Y. Sankai, and K. Suzuki, "Feasibility of synergy-based exoskeleton robot control in hemiplegia," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 6, pp. 1233–1242, 2018.
- [52] K. Kato, K. Takahashi, N. Mizuguchi, and J. Ushiba, "Online detection of amplitude modulation of motor-related EEG desynchronization using a lock-in amplifier: comparison with a fast Fourier transform, a continuous wavelet transform, and an autoregressive algorithm," *Journal of Neuroscience Methods*, vol. 293, pp. 289–298, 2018.
- [53] F. Wang, Y. He, J. Qu et al., "Enhancing clinical communication assessments using an audiovisual BCI for patients with disorders of consciousness," *Journal of Neural Engineering*, vol. 14, no. 4, Article ID 046024, 2017.

Research Article

Driving Fatigue Detection from EEG Using a Modified PCANet Method

Yuliang Ma ,¹ **Bin Chen**,^{1,2} **Rihui Li**,² **Chushan Wang**,³ **Jun Wang**,³ **Qingshan She** ,¹
Zhizeng Luo ,¹ and **Yingchun Zhang** ,²

¹*Intelligent Control & Robotics Institute, College of Automation, Hangzhou Dianzi University, Hangzhou, China*

²*Department of Biomedical Engineering, University of Houston, Houston, Texas, USA*

³*Guangdong Provincial Work Injury Rehabilitation Hospital, Guangzhou, China*

Correspondence should be addressed to Yingchun Zhang; yzhang94@uh.edu

Received 11 February 2019; Revised 28 April 2019; Accepted 19 June 2019; Published 14 July 2019

Guest Editor: Sangtae Ahn

Copyright © 2019 Yuliang Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid development of the automotive industry has brought great convenience to our life, which also leads to a dramatic increase in the amount of traffic accidents. A large proportion of traffic accidents were caused by driving fatigue. EEG is considered as a direct, effective, and promising modality to detect driving fatigue. In this study, we presented a novel feature extraction strategy based on a deep learning model to achieve high classification accuracy and efficiency in using EEG for driving fatigue detection. EEG signals were recorded from six healthy volunteers in a simulated driving experiment. The feature extraction strategy was developed by integrating the principal component analysis (PCA) and a deep learning model called PCA network (PCANet). In particular, the principal component analysis (PCA) was used to preprocess EEG data to reduce its dimension in order to overcome the limitation of dimension explosion caused by PCANet, making this approach feasible for EEG-based driving fatigue detection. Results demonstrated high and robust performance of the proposed modified PCANet method with classification accuracy up to 95%, which outperformed the conventional feature extraction strategies in the field. We also identified that the parietal and occipital lobes of the brain were strongly associated with driving fatigue. This is the first study, to the best of our knowledge, to practically apply the modified PCANet technique for EEG-based driving fatigue detection.

1. Introduction

As a leading factor in traffic accidents, driving fatigue accounts for 14%–20% of motor vehicle accidents that cause serious injuries and fatalities to human life [1]. Consequently, growing attention has been paid to driving safety in recent years. Driving safety is thought to be affected by multiple factors, including monotonous environments, sleep deprivation, chronic sleepiness, drug and alcohol use [2, 3], wherein the most common factor is driving fatigue [4, 5]. In such situation, drivers are most likely to fall asleep and drive unconsciously, which is not only a serious threat to the driver's own life and safety, but also a serious threat to the involved pedestrians and other vehicles. Therefore, detection of driving fatigue is of great importance to reduce the frequency and severity of traffic accidents [6].

In recent years, a variety of methods have been proposed to detect driving fatigue. For instance, Hiesh and Tai developed an infrared light-based digital signal processing (DSP) embedded system to capture driver's face and detect driving fatigue by identifying the opening and closing of eyes [7]. In another study, a calculation method named as the improved percentage of eyelid closure over the pupil overtime (PERCLOS) was employed as a standard criteria to judge whether the driver was tired or not [8]. Although this method is convenient in fatigue detection, it is vulnerable to environmental illumination such as the brightness, resulting in poor detection performance [9]. Later on, a sensor-based device called the steering wheel monitoring system (SAM) was developed to detect driving fatigue by monitoring the abnormal movement of the vehicle's steering wheel [10]. Although detection of steering wheel features good

real-time performance and low cost, it also suffers poor anti-interference ability and low reliability [11].

Recently, fatigue detection based on physiological signals such as electroencephalogram (EEG), electrocardiogram (ECG), electromyogram (EMG), and electrocardiogram (ECG) signals has been increasingly investigated in this field [4, 12–19]. Among them, EEG has some major assets in detecting driving fatigue due to its high temporal resolution, high portability, and good sensitivity to fatigue. With this in mind, a variety of studies have attempted to perform EEG-based classification using different signal processing techniques to accurately detect the fatigue during driving. For instance, Yang et al. implemented the detection of driving fatigue using information fusion and dynamic Bayesian neural network [20]. In another study, Zhao et al. demonstrated that wavelet packet can be used to extract features from EEG signal and classify the driving condition by a support vector machine (SVM) [21]. However, the EEG-based fatigue detection during driving is still faced with challenges. For instance, EEG signals are usually collected with low signal to noise ratio (SNR), which requires large effort in preprocessing raw EEG data. Moreover, the widely employed EEG-based classification techniques depend heavily on handcrafted features, which is time-consuming and highly relies on skilled person in the domain before predictions are performed. Therefore, there is a clear need to develop a new strategy to improve the robustness and efficiency of EEG-based driving fatigue detection.

As a simplified deep learning model based on convolution neural network (CNN), principal component analysis network (PCANet) has been developed and widely used for feature extraction in two-dimensional image processing [22]. Referring to a previous study that performed EEG-based lie detection [23], PCANet was proved to be highly effective in classification problem as it automatically extracted features from multichannel EEG data based on the deep learning technique rather than extracting handcrafted features in conventional ways. However, PCANet may be subject to the phenomenon of dimension explosion when the dimensionality of input data is large, which dramatically increases the complexity and cost of computation, rendering it more difficult to be effectively employed in multichannel EEG signal processing.

To address this challenge, this study aimed to adapt the PCANet algorithm to enhance the efficiency of conventional EEG-based driving fatigue detection by incorporating the principal component analysis (PCA) with the PCANet technique. Specifically, PCA is used to preprocess the multichannel EEG signals and reduce the dimensionality of data prior to the PCANet processing. The performance of the proposed modified PCANet method in driving fatigue detection was evaluated by comparing to traditional PCANet and other conventional features extraction methods widely used in driving fatigue detection [4].

2. Methods and Material

2.1. General Structure and Purpose. The main structure of this study is demonstrated in Figure 1. EEG data was first

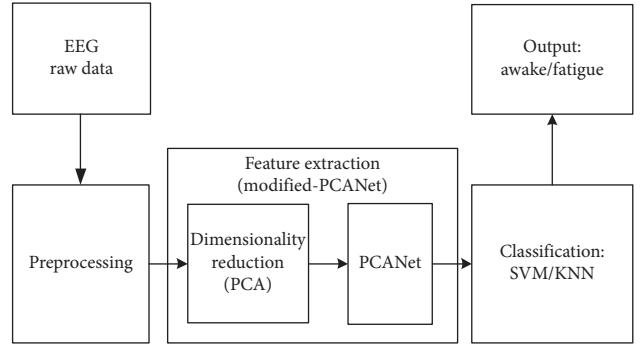


FIGURE 1: The overall schematic for the proposed EEG-based driving fatigue classification.

collected in a simulated driving environment, and followed by conventional signal preprocessing procedures. PCA was then used to reduce the dimensionality of the preprocessed signals, which were fed into the PCANet for feature extraction. Finally, the extracted features were classified using a support vector machine (SVM) and a K-Nearest Neighbors (KNN) classifier with respect to the driving condition.

2.2. Participants. Six male volunteers (right handed, aged 25.00 ± 2.00 years) with valid driver's licenses were recruited to participate in the experiment. All subjects were physically and psychologically healthy without any sleep disorders. The experiment was approved by the research ethics board of Guangdong Provincial Work Injury Rehabilitation Center and performed in accordance with the Declaration of Helsinki. Each subject was fully informed about the purpose of the research and provided written, informed consent prior to the start of the experiment.

2.3. Experiment Design and Data Acquisition. A 32-channel EEG acquisition system (Brain Products GmbH, Germany) was utilized to collect EEG signals with the sampling frequency set to 500 Hz. EEG electrodes were placed on the scalp according to the international 10–20 standard system. An advanced driving simulation system (Shanghai Infrared Automobile Simulator Driving Equipment Co., Ltd., China) was used in this study to simulate a real driving environment. Briefly, the driving simulation system could imitate the real driving scenarios with dynamically changing representations of the car and surrounding traffic. As shown in Figure 2, the driving simulation system consisted of a fixed car steering wheel, the brake and accelerator pedals, three large screens, a high-performance computer, driving simulation software, and a multifunctional data acquisition board. This system can be adapted to measure the EEG signal in different driving states in real time.

All participants were given sufficient time to practice and get familiar with the driving simulation system prior to the experiment. Datasets in two states were collected for each subject in this study, including the awake state and fatigue state. To collect the data in the awake state, all subjects were required to maintain a natural and adequate sleep for about 8 hours during the night before the experiment. EEG data



FIGURE 2: The setup of the experiment, including the driving platform and EEG recording system.

was collected at 9 a.m. on the next day for about 30–60 minutes while subjects were driving in the simulated environment. To collect the data for the fatigue state, all subjects were allowed to sleep for only 4 hours during the night before the experiment. The EEG data was then measured at 9 a.m. for 30 to 60 minutes while subjects were driving in the simulated environment. Specifically, the driving path was set to relatively long, straight with smooth curves and almost no pedestrian to increase the drowsiness of the subjects in fatigue group, while the path was set to relatively complicated to avoid the drowsiness of the subjects in the awake group. During the recording, an observer seated 2 meters beside the subjects monitored the subject's behavior without causing any disturbance to the subjects. The observer decided whether the subject was in a fatigue state or an awake state by observing the subject's drowsy signs (more than two seconds eye closure and head nodding, large deviation off the road) [4, 24]. Data recording was terminated 30 minutes after the subjects began to show fatigue signs. If the subject did not show any sign of drowsiness after 60 minutes of measurement, the experiment was terminated and data from the subject were excluded from further analysis. The experiment was conducted in a quiet, undisturbed laboratory with ambient temperatures around 22°C.

2.4. Data Preprocessing and Analysis. All the data analyses in this study were implemented using MATLAB (2014a, MathWorks, Natick, Massachusetts). 20-minute EEG signals in each state were selected for analysis. The raw multichannel EEG signals were first downsampled from 500 Hz to 200 Hz. A third-order bandpass filter (0.1–45 Hz) was then applied to remove artifacts such as slow drift, high-frequency noise, and powerline interference. The 20-minute preprocessed EEG data for each state were then segmented by a 10-second time window, resulting in 120 samples for each state and each subject. It is worthy of noting that in this study one sample was a two-dimension matrix (32 channels × 2000 data points). Overall, a total of 1440 samples were obtained from all subjects for classification (720 samples for awake and 720 samples for fatigue).

2.5. PCA Dimensionality Reduction and Extraction of Main Features. In this study, the proposed modified PCANet

consisted of two steps, PCA-based dimensionality reduction and PCANet-based feature extraction (Figure 1). To overcome the dimension explosion problem caused by conventional PCANet, PCA was first employed to reduce the dimensionality of each EEG sample. Briefly, for a given EEG sample data ($32\text{-channel} \times 2000\text{-point}$), the PCA transformed the data to 2000 linearly uncorrelated components known as principal components, ordered by the amount of variance of the original data that each component accounts for [25, 26]. By keeping the first r components with the largest variances and removing the remaining components, the size of original sample data could be reduced to 32 by r . In order to keep the original signal information as intact as possible, we kept the previous r components that accounted for at least 99% of the original signal as threshold, which was 20 for all samples in this study. Therefore, size of each sample was reduced to 32 by 20 after PCA optimization, and the optimized EEG data was analyzed using PCANet for feature extraction in the next step.

2.6. PCANet-Based Feature Extraction. As we introduced previously, PCANet is widely used in 2-D image processing, such as face recognition [27]. In this study, each optimized EEG sample was treated as a 2-D data matrix (32×20) and fed into PCANet for feature extraction.

The workflow of the PCANet network is shown in Figure 3, and details of the algorithm can be found in [22, 28]. In brief, the PCANet consists of two PCA-based filtering layers and an output layer that includes processing of binary hashing and blockwise histogram.

Assume we have N input samples after the EEG data preprocessing. Here, each EEG sample is treated as a two-dimensional signal of size $m \times n$ (channel number × sample number). Given an input EEG sample X_i , a sliding window of size $k_1 \times k_2$ is used to centralize the EEG sample by subtracting the mean value of each window's data from the corresponding window. Each centralized window is further vectorized into a single column, from which the input EEG sample is converted to a new 2-D matrix \bar{X}_i consisting of all centralized windows. The same processing is applied to all input EEG samples ($\{I_i\}_{i=1}^N$) to obtain the following data structure:

$$X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in R^{k_1 k_2 \times Nc}, \quad (1)$$

where $c = (m - k_1 + 1)(n - k_2 + 1)$ is the number of columns in \bar{X}_i .

The new vectorized matrix X is then used to perform the PCA filtration in the first layer. Specifically, the covariance matrix of X , denoted as $X_{\text{cov}} = XX^T/Nc$, is computed and applied to select the eigenvectors corresponding to L_1 principal eigenvalues as PCA filters W_l^1 . For the i th EEG sample, the output of the first PCA layer is then given by the convolution of the input EEG sample and the PCA filters:

$$I_i^l = \bar{I}_i * W_l^1, \quad i = 1, 2, \dots, N \text{ and } l = 1, 2, \dots, L_1, \quad (2)$$

where I_i^l denotes the l th output of the i th EEG sample and \bar{I}_i is the zero-padded form of I_i to ensure the same size of I_i^l and I_i .

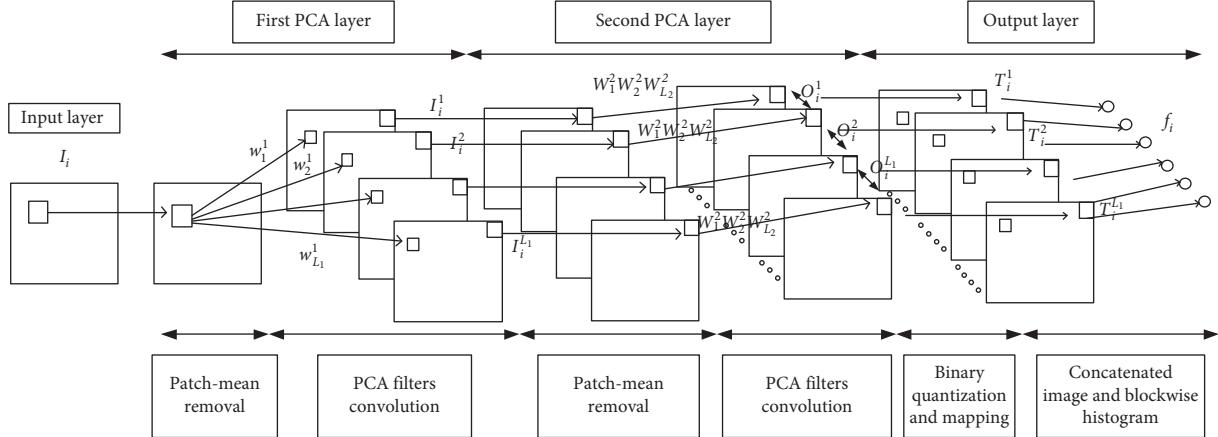


FIGURE 3: The PCANet network structure.

The second PCA layer is similar to the first layer. The output of the first PCA layer is centralized with the same sliding window and applied to select the eigenvectors corresponding to L_2 principal eigenvalues as PCA filters W_p^2 . The output of the second PCA layer is given as

$$O_i^l \doteq \vec{I}_i * W_p^2, \quad i = 1, 2, \dots, N \text{ and } p = 1, 2, \dots, L_2. \quad (3)$$

With the result obtained from the filtrations of two PCA layers, the output of the PCANet (T_i^l) is further processed by binary hashing as

$$T_i^l = \sum_{p=1}^{L_2} 2^{p-1} H(\vec{I}_i * W_p^2), \quad l = 1, 2, \dots, L_1, \quad (4)$$

where $H(\cdot)$ is a Heaviside step function that sets positive values as one and zero for others.

Finally, for the i th input EEG sample, each of the L_1 components in T_i^l is partitioned into B blocks. The histogram (with 2^{L_2} bins) of the decimal values in each block is computed and concatenated into one vector represented as Bhist (T_i^l). The PCANet-derived feature of the i th EEG sample is then denoted as

$$f_i = [\text{Bhist}(T_i^1), \dots, \text{Bhist}(T_i^{L_1})]^T \in R^{(2^{L_2})L_1B}. \quad (5)$$

The PCANet processing is applied to each EEG sample for feature extraction.

2.7. Classification. Support vector machine (SVM) and K-Nearest Neighbors (KNN) were employed as classifiers in the classification of awake and fatigue states for each subject. The performance of each classifier was evaluated using a 10-fold cross-validation strategy. At each iteration, 90% of the samples were randomly selected as the training set, and the accuracy, defined as the ratio between correct predictions and the total number of predictions, was computed on the remaining 10% of the samples, the testing set. Accuracies among the 10 steps of the cross validation were then averaged, yielding the mean accuracy for each subject. To evaluate the superiority of the proposed method, the obtained accuracy was compared to the performance obtained

from the traditional PCANet method and two commonly used feature extraction methods, i.e., the power spectral density (PSD) [29, 30] and wavelet packet decomposition (WPD) [31].

In this study, the PSD features of each EEG channel in a segmented EEG sample ($32\text{-channel} \times 2000\text{-point}$) were estimated through Short-Time Fourier transform (STFT) with a 128-point Hanning window and 50% overlap rate. For each EEG channel, the PSD feature of a specific frequency band was computed by averaging all squared magnitude values of STFT within the corresponding frequency range. In this study, five typical EEG bands were investigated, including delta (0.1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–20 Hz), and gamma (20–45 Hz). This resulted in 160 PSD features for each EEG sample (5-band \times 32-channel).

The WPD features of each EEG channel in a segmented EEG sample ($32\text{-channel} \times 2000\text{-point}$) were calculated based on discrete wavelet decomposition (DWT). In brief, the DWT decomposed the selected EEG signal into a number of layers by filtering the signal with quadrature mirror filters (a low-pass filter and a high-pass filter). The output of each layer was a series of detail coefficients (from the high-pass filter) and approximation coefficients (from the low-pass filter), which were extracted as features for classification [32]. In this study, we decomposed each EEG channel data (2000 points) with a 3-layer “Daubechies” wavelet, resulting in 8 groups of coefficients (256 points). Therefore, there were in total 65536 (8-groups \times 256-point \times 32-channel) features extracted for each EEG sample.

3. Results

Referring to previous studies [4, 24], it was found that the alteration in brain regions during awake and fatigue states were more prominent at the parietal lobe at alpha and beta frequency bands [33, 34]. Figure 4 shows the group-averaged PSD distribution of the relevant EEG signals in alpha (8–13 Hz) and beta (14–20 Hz) bands in awake and fatigue states. The PSD values were computed based on the average of all 10-second EEG samples in each state for each subject. In brief, in the fatigue state, the PSD in the parietal and

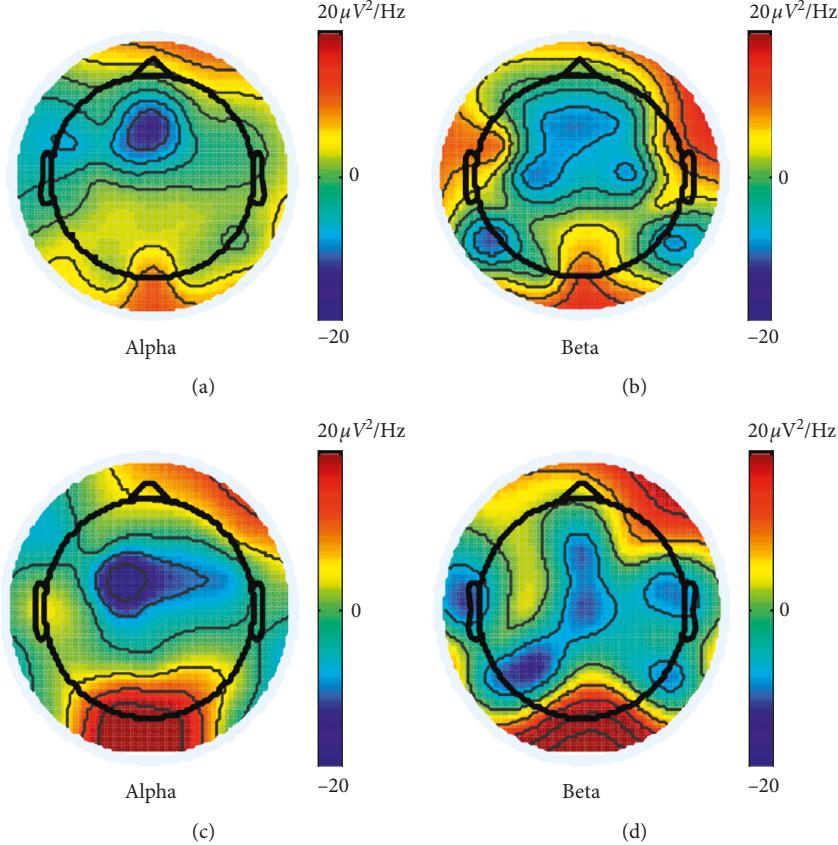


FIGURE 4: PSD distributions of alpha and beta bands for awake (a, b) and fatigue (c, d) states.

occipital lobes of the brain was more pronounced compared to the PSD distribution in the awake state.

In order to determine the optimal number of PCA filters (L_1 and L_2) when using the PCANet, the classification performance varied with the number of PCA filters was acquired for each subject. Figure 5 shows the variation of classification performance for each subject when using a SVM, with number of PCA filters increasing from 2 to 14 for both layers, respectively. Overall, for most subjects (except sub. 3 and sub. 4), the classification performance was enhanced as the filter number increased, and gradually decreased when filter number was over 10 or 12.

Additionally, the performance of classification using two classifiers and various feature extraction strategies, including the traditional PCANet, WPD, PSD, and the proposed modified PCANet method, is shown in Figure 6 and summarized in Tables 1 and 2. Overall, for both classifiers, when selecting the optimal PCA filter number for each individual subject, the traditional PCANet and the proposed modified PCANet method achieved better performance in the classification of awake and fatigue states. The result of the paired-*t* test between four feature extraction methods revealed that the PCANet-based methods significantly outperformed the other two methods ($p < 0.005$), as shown in Figure 6 and Table 2. Although no significant difference in classification accuracy was observed

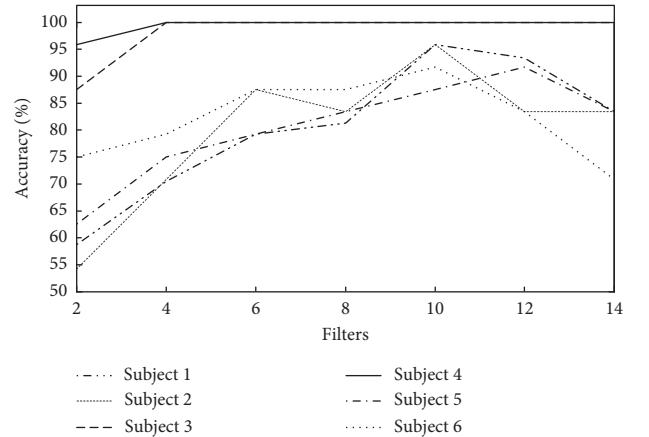


FIGURE 5: The classification performance of different PCA filter numbers when using SVM.

between the traditional PCANet and the modified PCANet, the time used in the feature extraction, model training, and testing was drastically reduced when using the modified PCANet method, indicating the high efficiency of this method (Table 3).

In addition, the area under the curve (AUC) of the receiver operating characteristic (ROC) curve [35], which evaluates how well a model separates the groups being classified, was employed to assess the performance of

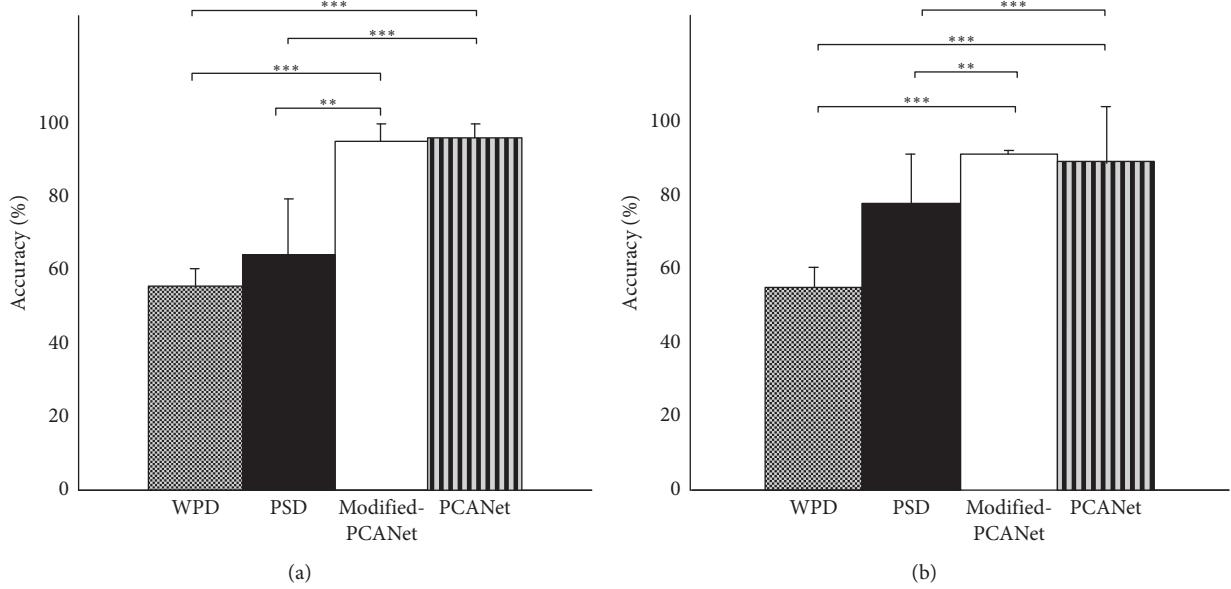


FIGURE 6: The average accuracies obtained by SVM (a) and KNN (b) when using different features extracted through WPD, PSD, modified PCANet, and PCANet. “**” denotes significantly different from controls ($p < 0.005$). “***” denotes significantly different from controls ($p < 0.0001$).

TABLE 1: The classification performance using different feature extraction approaches.

Methods	Classifiers	Classification performance	
		Accuracy (%)	AUC
WPD	SVM	55.42 ± 5.09	0.51 ± 0.11
	KNN	54.00 ± 5.00	0.46 ± 0.08
PSD	SVM	64.44 ± 15.06	0.55 ± 0.12
	KNN	76.00 ± 13.00	0.55 ± 0.08
Modified-PCANet	SVM	95.14 ± 4.87	0.97 ± 0.04
	KNN	89.00 ± 10.00	0.89 ± 0.12
PCANet	SVM	96.00 ± 4.00	0.98 ± 0.03
	KNN	87.00 ± 15.00	0.91 ± 0.10

TABLE 2: The summary of the statistical analysis (t -test) of the classification performance between all feature extraction methods.

Methods	Classifiers	p values	
		Accuracy (%)	AUC
Modified-PCANet-WPD	SVM	$1.36e - 07$	$1.02E - 04$
	KNN	$6.15e - 05$	$5.18E - 04$
Modified-PCANet-PSD	SVM	0.0023	$2.24E - 04$
	KNN	0.0018	$3.10E - 03$
Modified-PCANet-PCANet	SVM	0.3541	0.6612
	KNN	0.6823	0.2884

different feature extraction strategies. As summarized in Tables 1 and 2, for both classifiers, the AUC values obtained from the traditional PCANet and the proposed modified PCANet method are significantly higher compared to PSD and WPD features ($p < 0.005$). Similarly, no significant difference in AUC values was observed between the traditional PCANet and the modified PCANet.

4. Discussion

This study sought to validate the feasibility of using modified PCANet to enhance the performance of EEG-based driving fatigue detection. The neuronal electrical activity was recorded using EEG in a simulated driving environment with subjects experienced both awake and fatigue states. We employed PCA to alleviate the dimension explosion caused by PCANet before classification. The experimental results indicated a significantly enhanced performance in the fatigue detection compared to the traditional PCANet and other conventional approaches.

Alterations in low and high frequency bands were previously observed by EEG in the drowsy state [36]. In summary, investigations that included the transition from awake to sleepy states have demonstrated an increase in the alpha rhythm [24]. The alteration of the alpha band during drowsiness in both simulated and actual driving conditions was also reported in a previous study [37]. In this study, we compared the PSD between both states for alpha and beta frequency bands and found an increased PSD at occipital and parietal areas in both alpha and beta bands (Figure 4). This finding is in line with the results from these studies, demonstrating the possibility of using EEG as a portable and reliable approach to monitor and detect the driving fatigue.

In order to monitor the brain state during driving, it is of great importance to achieve high accuracy and reliability in detecting the driver’s fatigue state. With high classification accuracy between the awake and fatigue states achieved by the modified PCANet approach, our study proved the usefulness of EEG to study driving fatigue. In particular, the substantial increase in classification accuracy using the proposed method, compared to conventional feature extraction methods, offers a new perspective to deal with classification problem when using multichannel biosignals

TABLE 3: The average time (seconds) used in the feature extraction, model training, and testing between the traditional PCANet and the proposed modified PCANet method.

Steps	Methods	Number of PCA filters					
		2	4	6	8	10	12
Features extraction	Modified-PCANet	0.89	1.46	2.18	2.91	4.22	7.71
	PCANet	125.43	219.94	320.39	449.50	651.95	1202.40
Model training	Modified-PCANet	SVM	0.05	0.70	5.67	10.34	13.25
		KNN	0.02	0.30	2.20	4.10	5.12
	PCANet	SVM	7.02	126.66	302.52	736.79	1813.40
		KNN	2.70	49.01	118.60	285.04	697.46
Model testing	Modified-PCANet	SVM	0.25	0.26	0.26	0.28	0.34
		KNN	0.10	0.12	0.13	0.15	0.13
	PCANet	SVM	1.16	2.94	5.34	11.10	20.84
		KNN	0.45	1.14	2.10	4.22	8.02
							14.06

such as EEG and EMG signals. It is noteworthy that PCA was necessarily adopted before PCANet was employed to extract features in this study. As shown in Table 3, the proposed modified PCANet method remarkably reduced the time for the classification while maintaining a comparable performance relative to the traditional PCANet approach. This provides evidence that PCA is able to alleviate the curse of dimensionality induced by PCANet, reducing the computational cost when using the traditional PCANet. By taking this great advantage, the proposed method is considered a more effective strategy in a practical scenario such as monitoring driving fatigue in real time. In addition, the components compressed by PCA retains the main characteristics of original signals, which is the inherent benefit offered by PCA. The pre-refined signals can be further improved by PCANet to achieve the significantly enhanced classification accuracy. Particularly, the classification performance across all subjects not only exhibited high accuracy, but also yielded lower variance, demonstrating the good robustness of the proposed method.

Despite the improvement achieved in this study, there are still several limitations in this study. Firstly, only off-line analysis and small sample size were elected in this study. Real-time fatigue classification should be conducted on larger population base in the future to validate the potential of the proposed approach in actual driving environment. In addition, two-layer PCA structure was applied for feature extraction in this study, and 8–12 filters were considered optimal setting for achieving best performance. Apparently, the number of PCA filters within each layer affected the quality of feature extraction, which significantly affected the performance and efficiency of the fatigue classification. Further investigation on how to automatically select the best filter number for each subject is required. Finally, in the present study we solely focused on adapting a deep learning-based technique to the conventional EEG-based driving fatigue classification and provide a new perspective to deal with classification problem when using multichannel biosignals. Although compelling result was achieved in current study, it is expected that future work will evaluate and employ the state-of-the-art algorithms to enhance the performance of this application. Even though the mentioned limitations may prevent us drawing a solid conclusion, the

preliminary results demonstrate the capability of the proposed PCANet-based algorithm to monitor and detect the driving fatigue in advance so that it can prevent motor vehicle collision caused by driver drowsiness.

5. Conclusion

In this paper, a novel feature extraction strategy incorporating the PCA and PCANet techniques was proposed to enhance the classification performance in EEG-based driving fatigue detection. Significantly enhanced classification performance was achieved using the proposed modified PCANet method compared to the traditional PCANet algorithm and two conventional feature extraction strategies. Additionally, the power spectrum analysis of EEG signals indicated a higher power alteration at occipital and parietal areas in alpha and beta bands. The findings in this study not only demonstrated the effectiveness of using EEG to monitor driving fatigue but also provided a new perspective to adapt a novel machine learning algorithm to investigate the nature of philological signals.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China (no. 61871427, 61671197, and 61372023) and the University of Houston.

References

- [1] A. Fletcher, K. McCulloch, S. D. Baulk, and D. Dawson, “Countermeasures to driver fatigue: a review of public awareness campaigns and legal approaches,” *Australian & New Zealand Journal of Public Health*, vol. 29, no. 5, pp. 471–476, 2005.

- [2] P.-H. Ting, J.-R. Hwang, J.-L. Doong, and M.-C. Jeng, "Driver fatigue and highway driving: a simulator study," *Physiology & Behavior*, vol. 94, no. 3, pp. 448–453, 2008.
- [3] H. Summala, H. Hakkanen, T. Mikkola, and J. Sinkkonen, "Task effects on fatigue symptoms in overnight driving," *Ergonomics*, vol. 42, no. 6, pp. 798–806, 1999.
- [4] S. Ahn, T. Nguyen, H. Jang, J. G. Kim, and S. C. Jun, "Exploring neuro-physiological correlates of drivers' mental fatigue caused by sleep deprivation using simultaneous EEG, ECG, and fNIRS data," *Frontiers in Human Neuroscience*, vol. 10, no. 848, 2016.
- [5] G. Zhang, K. K. W. Yau, X. Zhang, and Y. Li, "Traffic accidents involving fatigue driving and their extent of casualties," *Accident Analysis & Prevention*, vol. 87, pp. 34–42, 2016.
- [6] H. Wang, C. Zhang, T. Shi, F. Wang, and S. Ma, "Real-time EEG-based detection of fatigue driving danger for accident prediction," *International Journal of Neural Systems*, vol. 25, no. 2, article 1550002, 2015.
- [7] C.-S. Hsieh and C.-C. Tai, "An improved and portable eye-blink duration detection system to warn of driver fatigue," *Instrumentation Science & Technology*, vol. 41, no. 5, pp. 429–444, 2013.
- [8] J. J. Yan, H.-H. Kuo, Y.-F. Lin, and T.-L. Liao, "Real-time driver drowsiness detection system based on PERCLOS and grayscale image processing," in *Proceedings of the International Symposium on Computer*, Seoul, South Korea, June 2016.
- [9] J. Jimenez-Pinto and M. Torres-Torriti, "Face salient points and eyes tracking for robust drowsiness detection," *Robotica*, vol. 30, no. 5, pp. 731–741, 2012.
- [10] T. Åkerstedt, B. Peters, A. Anund, and G. Kecklund, "Impaired alertness and performance driving home from the night shift: a driving simulator study," *Journal of Sleep Research*, vol. 14, no. 1, pp. 17–20, 2010.
- [11] A. Mittal, K. Kumar, S. Dhamija, and M. Kaur, "Head movement-based driver drowsiness detection: a review of state-of-art techniques," in *Proceedings of the IEEE International Conference on Engineering & Technology*, Coimbatore, India, March 2016.
- [12] R. Chai, G. R. Naik, T. N. Nguyen et al., "Driver fatigue classification with independent component by entropy rate bound minimization analysis in an EEG-based system," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 715–724, 2017.
- [13] J. Hu, "Comparison of different features and classifiers for driver fatigue detection based on a single EEG channel," *Computational and Mathematical Methods in Medicine*, vol. 2017, Article ID 5109530, 9 pages, 2017.
- [14] J. X. Ma, L. C. Shi, and B. L. Lu, "An EOG-based vigilance estimation method applied for driver fatigue detection," *Neuroscience & Biomedical Engineering*, vol. 2, no. 1, pp. 41–51, 2015.
- [15] W. L. Zheng and B. L. Lu, "A multimodal approach to estimating vigilance using EEG and forehead EOG," *Journal of Neural Engineering*, vol. 14, no. 2, article 026017, 2017.
- [16] L. Boon-Leng, L. Dae-Seok, and L. Boon-Giin, "Mobile-based wearable-type of driver fatigue detection by GSR and EMG," in *Proceedings of the Tencon IEEE Region 10 Conference*, Xi'an, China, January 2016.
- [17] A. Tjolleng, K. Jung, W. Hong et al., "Classification of a driver's cognitive workload levels using artificial neural network on ECG signals," *Applied Ergonomics*, vol. 59, pp. 326–332, 2017.
- [18] F. Wang, H. Wang, and R. Fu, "Real-time ECG-based detection of fatigue driving using sample entropy," *Entropy*, vol. 20, no. 3, p. 196, 2018.
- [19] J. F. O'Hanlon and G. R. Kelley, "Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation," in *Vigilance*, pp. 87–109, Springer, Boston, MA, USA, 1977.
- [20] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Information Sciences*, vol. 180, no. 10, pp. 1942–1954, 2010.
- [21] C. Zhao, C. Zheng, and M. Zhao, J. Liu, Y. Tu, "Automatic classification of driving mental fatigue with EEG by wavelet packet energy and KPCA-SVM," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 3, pp. 1157–1168, 2011.
- [22] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "PCANet: a simple deep learning baseline for image classification?," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5017–5032, 2015.
- [23] L.-Y. Gu, W. Z. Lv, Y. Yang, J. F. Gao, J. A. Guan, and D. Zhou, "Deception detection study based on PCANet and support vector machine," *Acta Electronica Sinica*, vol. 44, no. 8, pp. 1969–1973, 2016.
- [24] T. Nguyen, S. Ahn, H. Jang, S. C. Jun, and J. G. Kim, "Utilization of a combined EEG/NIRS system to predict driver drowsiness," *Scientific Reports*, vol. 7, no. 1, 2017.
- [25] S. Poorna, V. V. Arsha, P. T. A. Aparna, P. Gopal, and G. J. Nair, "Drowsiness detection for safe driving using PCA EEG signals," in *Progress in Computing, Analytics and Networking*, pp. 419–428, Springer, Singapore, 2018.
- [26] R. K Chaurasiya, N. D. Londhe, and S. Ghosh, "Statistical wavelet features, PCA, and SVM based approach for EEG signals classification," *International Journal of Computer and Information Engineering*, vol. 9, no. 2, pp. 182–186, 2015.
- [27] L. Tian, C. Fan, and Y. Ming, "Multiple scales combined principle component analysis deep learning network for face recognition," *Journal of Electronic Imaging*, vol. 25, no. 2, article 023025, 2016.
- [28] B. Li, Y. Dong, D. Zhao, Z. Wen, and L. Yang, "A PCANet based method for vehicle make recognition," in *Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, IEEE, Rio de Janeiro, Brazil, November 2016.
- [29] C. Kim, J. Sun, D. Liu, Q. Wang, and S. Paek, "An effective feature extraction method by power spectral density of EEG signal for 2-class motor imagery-based BCI," *Medical & Biological Engineering & Computing*, vol. 56, no. 9, pp. 1645–1658, 2018.
- [30] N. G. Ozmen, L. Gumusel, and Y. Yang, "A biologically inspired approach to frequency domain feature extraction for EEG classification," *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 9890132, 10 pages, 2018.
- [31] Y. Zhang, B. Liu, X. Ji, and D. Huang, "Classification of EEG signals based on autoregressive model and wavelet packet decomposition," *Neural Processing Letters*, vol. 45, no. 2, pp. 365–378, 2017.
- [32] R. Li, T. Potter, W. Huang, and Y. Zhang, "Enhancing performance of a hybrid EEG-fNIRS system using channel selection and early temporal features," *Frontiers in Human Neuroscience*, vol. 11, p. 462, 2017.
- [33] J. P. Liu, C. Zhang, and C. X. Zheng, "Estimation of the cortical functional connectivity by directed transfer function

- during mental fatigue," *Applied Ergonomics*, vol. 42, no. 1, pp. 114–121, 2011.
- [34] F. Gharagozlu, G. Nasl Saraji, A. Mazloumi et al., "Detecting driver mental fatigue based on EEG alpha power changes during simulated driving," *Iranian Journal of Public Health*, vol. 44, no. 12, pp. 1693–1700, 2015.
- [35] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.
- [36] I. Käthner, S. C. Wriessnegger, G. R. Müller-Putz, A. Kübler, and S. Halder, "Effects of mental workload and fatigue on the P300, alpha and theta band power during operation of an ERP (P300) brain-computer interface," *Biological Psychology*, vol. 102, no. 5, pp. 118–129, 2014.
- [37] M. A. Schier, "Changes in EEG alpha power during simulated driving: a demonstration," *International Journal of Psychophysiology*, vol. 37, no. 2, pp. 155–162, 2000.

Research Article

Covert Intention to Answer “Yes” or “No” Can Be Decoded from Single-Trial Electroencephalograms (EEGs)

Jeong Woo Choi^{1,2} and Kyung Hwan Kim¹

¹Department of Biomedical Engineering, Yonsei University, Wonju 26493, Republic of Korea

²Department of Neurosurgery, University of California, Los Angeles, CA 90095, USA

Correspondence should be addressed to Kyung Hwan Kim; khkim0604@yonsei.ac.kr

Received 19 March 2019; Revised 14 May 2019; Accepted 13 June 2019; Published 10 July 2019

Guest Editor: Hyun J. Baek

Copyright © 2019 Jeong Woo Choi and Kyung Hwan Kim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Interpersonal communication is based on questions and answers, and the most useful and simplest case is the binary “yes or no” question and answer. The purpose of this study is to show that it is possible to decode intentions on “yes” or “no” answers from multichannel single-trial electroencephalograms, which were recorded while covertly answering to self-referential questions with either “yes” or “no.” The intention decoding algorithm consists of a common spatial pattern and support vector machine, which are employed for the feature extraction and pattern classification, respectively, after dividing the overall time-frequency range into subwindows of $200\text{ ms} \times 2\text{ Hz}$. The decoding accuracy using the information within each subwindow was investigated to find useful temporal and spectral ranges and found to be the highest for 800–1200 ms in the alpha band or 200–400 ms in the theta band. When the features from multiple subwindows were utilized together, the accuracy was significantly increased up to ~86%. The most useful features for the “yes/no” discrimination was found to be focused in the right frontal region in the theta band and right centroparietal region in the alpha band, which may reflect the violation of autobiographic facts and higher cognitive load for “no” compared to “yes.” Our task requires the subjects to answer self-referential questions just as in interpersonal conversation without any self-regulation of the brain signals or high cognitive efforts, and the “yes” and “no” answers are decoded directly from the brain activities. This implies that the “mind reading” in a true sense is feasible. Beyond its contribution in fundamental understanding of the neural mechanism of human intention, the decoding of “yes” or “no” from brain activities may eventually lead to a natural brain-computer interface.

1. Introduction

The most fundamental linguistic communication consists of questions and answers, and the simplest one is the binary “yes or no” question and answer. This enables fundamental interpersonal communications (e.g., “Is your name John?” “Yes” or “Do you want to drink water?” “No”). So, by decoding the intentions to answer either “yes” or “no” from brain activities, a natural interpersonal communication tool, which does not require any operant training or heavy cognitive efforts, may be developed. As the first step toward this, here we tried to demonstrate that it is possible to decode the intentions to answer “yes” or “no” in response to self-referential questions from noninvasive electroencephalograms (EEGs) on single-trial basis. This was motivated by

our recent studies which showed that the intentions to answer “yes” or “no” to self-referential questions is represented significantly differently in event-related EEGs, particularly in alpha-band activities [1, 2].

Direct decoding of “yes” and “no” intentions may eventually lead to advancement of the brain-computer interface (BCI), which is a technological means to deliver user’s intention to the external world (device or other people) without behavioral outputs, by direct interpretation of brain activities. The most important target of the BCI is the patients with severe motor impairment, who are unable to communicate with others including those in the completely locked-in state (CLIS) due to amyotrophic lateral sclerosis, spinal cord injury, and brainstem stroke [3–6]. One of the most crucial technologies to enable the BCI is to read

or “decode” the users’ intention from their brain activities. Two major approaches have been pursued for the intention decoding. The first is based on voluntary self-regulation of specific brain signals such as slow cortical potential [7] and sensorimotor rhythms [8]. This requires extensive operant training using feedback and reward. Unfortunately, many people are unable to regulate the brain activities as required, which is known as “BCI illiteracy” [9, 10]. The other approach utilizes evoked brain activities such as P300 event-related potential (ERP) [11, 12] and steady-state evoked potential [13, 14]. The operant training is not required, but sustained attention is needed to induce discriminable brain response increases, resulting in significant cognitive workload.

Both approaches may not be so successful for the patients with CLIS [15]. It is speculated that the failure is due to the extinction of goal-directed cognition and thought in the CLIS patients [15]. An alternative approach for the mind reading is crucial, which does not require volitional and highly cognitive efforts. Birbaumer and colleagues suggested an approach based on classical conditioning [16–18]. They tried to associate language stimuli with unpleasant and painful sensory stimuli so that cortical responses to these nonlanguage stimuli are conditioned according to the language stimuli. This is remarkable considering that language is the most natural means of communication.

The specific aim of this study is to show that it is feasible to decode “yes” and “no” answers in mind from single-trial EEGs. We demonstrated that mind reading in a true sense, which is based on the prediction of the intentions to answer the questions from brain activities, is achievable. For the intention decoding, the discriminative characteristics of EEGs that we found in our previous study were utilized to find the time-frequency features for “yes/no” decoding. The decoding algorithm was developed based on the same data used in our previous study [2].

2. Materials and Methods

2.1. Subjects. 23 subjects with no record of neurological or psychiatric illness participated in the experiment (age: 23.13 ± 2.97 years, 12 males). All the subjects were undergraduate students of Yonsei University and right-handed native Korean speakers. Written informed consent was obtained from each subject before the experiment. The experimental procedure was approved by the Yonsei University Wonju Institutional Review Board (IRB). All experiments were performed in accordance with the guidelines and regulations of the IRB.

2.2. Experimental Task. Before the experiment, all subjects completed a written questionnaire on their autobiographical facts (e.g., job, name, age, and date of birth). We generated two opposite types of questions from a single autobiographical fact; one question should be answered “yes,” and the other (i.e., autobiographical fact violation (AFV)) should be answered “no.” These two questions were almost the same except one critical word (*italicized* word in the example below), which determined whether the question agreed with

the subject’s identity or not. For example, if the subject’s job was a student, the two questions were as follows:

Type (a), to be answered “yes”: Is your job a *student*?

Type (b), to be answered “no”: Is your job a *teacher*?

In total, 40 type (a) questions and 40 type (b) questions were generated based on the questionnaire for each subject. Each question was composed of 2 or 3 Korean words, and the average number of characters (Korean “Hangul”) in each critical word was 3.18 ± 1.02 . Each character had 3.3 cm width and 4.27 cm height.

All questions were presented visually through commercial software (PRESENTATION; Neurobehavioral systems, Berkeley, CA). After explaining the detailed procedure of the experimental task, we requested the subjects to watch each word presented on a 17 inch computer screen carefully so that they can make immediate response as soon as possible to the critical words. The distance between the subject’s eyes and the monitor was set to ~ 0.75 m. Each word in a question was presented sequentially one by one on the center of the monitor, as described below.

Figure 1(a) illustrates the experimental procedure. A cross mark (“+”) for the fixation appeared for 1000 ms and a black screen followed for 300 ms. And then, each word in a question was presented sequentially for 300 ms, with a black screen for 300 ms between the words. The last word in the question is referred to the critical word (CW), which was presented for 300 ms along with a question mark. Although this question mark may naturally induce decision of “yes” or “no” and thus evoke answer automatically, we instructed the subjects not to make any response neither covertly nor overtly but to retain the answer in mind during the 1 s blank period. This would enable us to explore the cortical activity during retaining the information on “yes” or “no” in working memory (WM). Finally, when “Please respond” (in Korean) was presented for 300 ms, the subjects were requested to respond covertly in mind with either “yes” or “no” without any behavioral responses.

Figure 1(b) illustrates expected temporal sequence of cognitive processing following the CW onset until the “Please respond” cue appeared, which was based on our previous studies on cortical information processing of intention [1, 2], which showed that the brain activities differed between “yes” and “no” answers at both early (0~600 ms) and late periods (600~1300 ms) relative to the CW onset. We found that the early period was associated with semantic processing and automatic decision to answer [1] (denoted by a red box in Figure 1(b)), while the late period was involved in the retention of the answer in memory (denoted by a blue box in Figure 1(b)) until the “Respond cue” appeared (denoted by a yellow box in Figure 1(b)) [2]. Thus, the temporal period of interest for decoding the intentions to answer “yes” or “no” was the late period, corresponding to retain the intention in mind (600~1300 ms).

Each subject performed two blocks of tasks. Each block included all questions generated based on the questionnaire (i.e., 40 type (a) and 40 type (b) questions), and 10 of 40 questions for each question type were randomly selected and

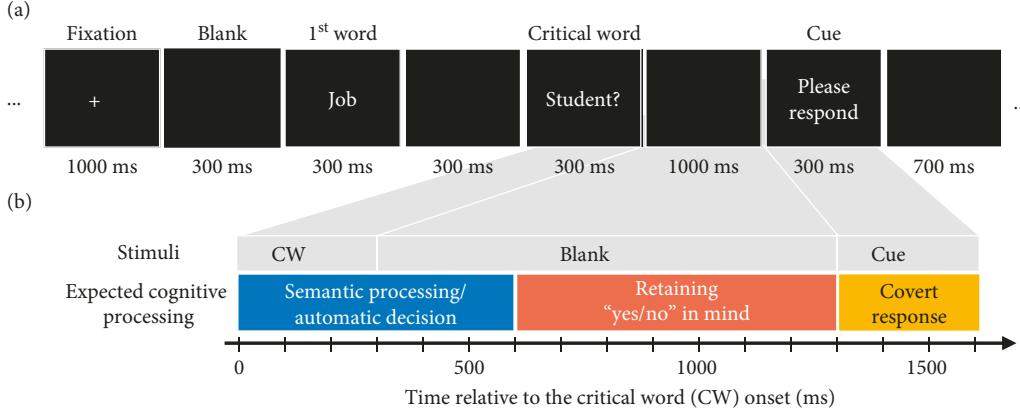


FIGURE 1: Experimental task: (a) presentation of stimuli; (b) expected temporal sequence of cognitive processing.

presented once again. Consequently, each block included 50 type (a) and 50 type (b) questions in total. The average duration of each single trial (i.e., one question and answer) was 4380 ± 274.95 ms. The total time for performing the tasks was approximately 20 minutes including at least 5 minutes of rest between blocks.

2.3. Electroencephalogram (EEG) Recording and Data Analysis. Sixty channel EEGs were recorded based on the 10–10 system using an EEG amplifier (Brain Products GmbH, Munich, Germany) with an Ag/AgCl electrode cap (EASYCAP, FMS, Munich, Germany). The ground and reference electrodes were at AFz and linked mastoids, respectively. The impedances of all electrodes were kept under $10\text{ k}\Omega$. The sampling rate was 500 samples/s. A bandpass filter (0.03–100 Hz) and a notch filter (60 Hz) were applied in order to reduce background noise and powerline interferences.

An open source toolbox EEGLAB was used for the whole procedure of preprocessing [19]. First, single-trial EEGs were segmented during the $-500\text{--}1300$ ms period relative to the critical word onset. By visual inspection, we removed the single-trial waveforms contaminated severely from non-stereotyped artifacts such as drifts and discontinuity. Then, an independent component analysis (ICA) was employed to the remaining single-trial EEGs in order to correct the ocular and muscular artifacts [20]. The group-averaged percentage of the number of epochs remaining per subject was $98.88 \pm 3.08\%$ and $97.96 \pm 5.86\%$ for “yes” and “no” questions, respectively.

2.4. Yes/No Decoding. Figure 2(a) illustrates the structure of “yes/no” intention decoding algorithm using local time-frequency information. First, we selected 29 channels out of 60 (i.e., Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, and O2), following the standard 10–20 system. This was based on a recent simulation study which showed that the decoding accuracy with the common spatial pattern (CSP) spatial filtering was optimized when ~ 30 channels were used and decreased for more channels [21]. The overall

time-frequency range ($0\text{--}1200$ ms, 4–50 Hz) was divided into subwindows of $200\text{ ms} \times 2$ Hz. The intention decoding within each of the local time-frequency subwindows was performed as follows. Single-trial EEGs were bandpass filtered in the frequency range of the subwindow using a linear-phase finite impulse response filter (the number of the filter order: 512, bandwidth: 2 Hz). The multichannel bandpass-filtered signals within the temporal period of the subwindow were subsequently projected to the lower dimension (four dimensions) by the CSP algorithm [22]. The four time series obtained from the CSP spatial filter were used to construct a four-dimensional feature vector, which was passed to a support vector machine (SVM) classifier. The final output of the classifier was either “yes” or “no,” a decision of answer for each single trial.

The performance of the trained classifier was validated by 10-fold cross-validation as follows: First, for each class, we randomly split all the trials into 10-folds with the same number of trials (i.e., ~ 10 trials per fold for each class). Then, we randomly selected one fold (k^{th} , where $k = 1, 2, \dots, 10$) as a test data (10%) and trained the classifier using the rest of data (i.e., 9 folds excepting the k^{th} fold, 90%). In order to keep a balance between the numbers of “yes” and “no” trials, the training/testing data were selected within each class (i.e., “yes” or “no”), as shown in Figure 2(b). The ground truth for each single trial was determined whether the question in the single trial was including AFV or not. The decoding accuracy for each subject was estimated by averaging the ratio of correct classification from 10 repetitions (i.e., $k = 1, 2, \dots, 10$) of this procedure.

Additionally, we also made effective use all the features obtained from multiple time-frequency subwindows, in order to investigate whether more accurate decoding is possible by combining useful features each of which was localized in the time-frequency domain. The time-frequency subwindows were selected if the decoding accuracies for a specific subwindow were higher than a predetermined threshold ($2 \times$ standard deviation above the mean among all time-frequency subwindows). And then, the classifier was trained and tested as described above, with input feature vectors obtained by combining all the selected subwindows.

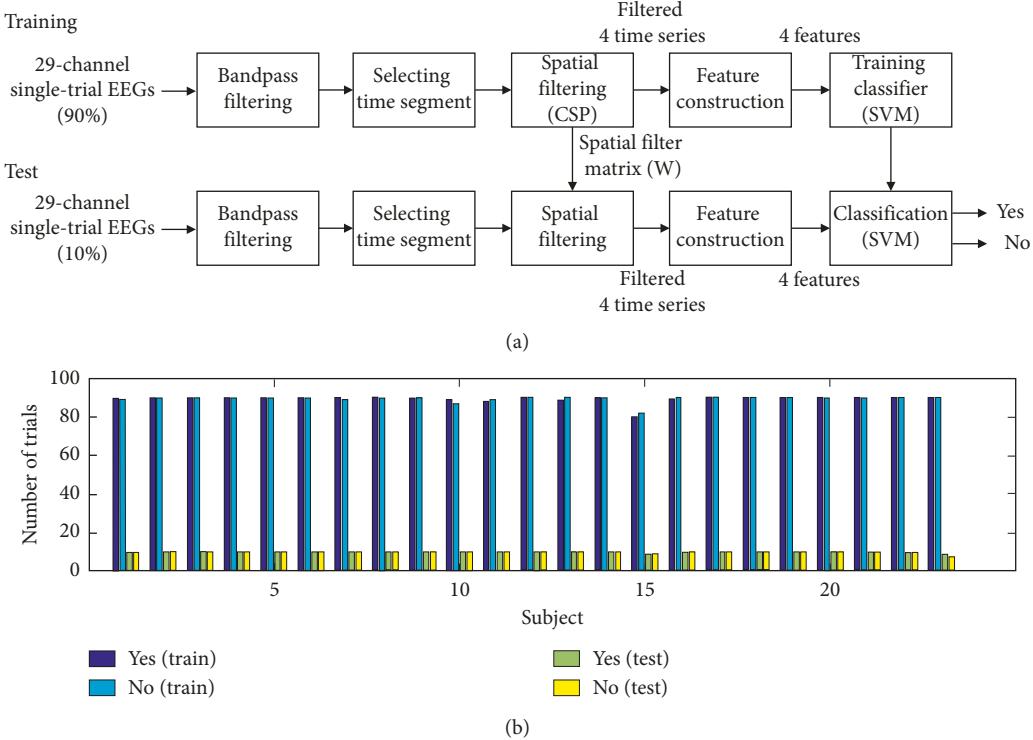


FIGURE 2: (a) Block diagram of the intention decoding algorithm; (b) the number of trials which was selected as training and testing data within each answer type for each subject.

2.5. Event-Related Spectral Perturbation (ERSP) Analysis. The time-frequency activation patterns, i.e., ERSPs, were investigated to reveal statistical differences between “yes” and “no” to find the time and frequency ranges of interests for effective classification. A continuous wavelet transform (CWT) based on a complex Morlet wavelet was used for the ERSP analysis [23]. The number of cycles for the CWT linearly increased according to the frequency from 4 to 13.5, at the lowest (1 Hz) and the highest frequencies (100 Hz), respectively [19]. This method provides better frequency resolution at high frequencies, and it is better matched to the linear scale that we adopted to visualize the time-frequency map [19]. The induced spectral power was calculated by averaging the ERSP patterns of each single trial [24]. The time-frequency distribution of ERSP patterns was represented as the ratio of the relative change to the power in a baseline interval from -300 to 0 ms prior to stimulus onset, to reduce intersubject variability and to normalize power changes across different frequency bands.

We employed the mass-univariate approach with the cluster-based permutation test for correcting multiple comparisons [25] in order to find the time, frequency, and electrode showing significant differences between “yes” and “no” without *a priori* knowledge. Detailed procedure is as follows:

- (1) A large number of paired-sample t -tests were applied to the data for all time-frequency-electrode bins within the range of $0\text{--}1200$ ms (time), $5\text{--}30$ Hz (frequency), and 29 electrodes. The number of bins was $181,714 = 241 \times 26 \times 29$ since there were 241 time samples, 26 frequency points, and 29 electrodes. The

electrodes showing high t values were selected, and the average power spectral power was calculated over the selected electrodes, as follows. First, from spatial distribution of the t values averaged within the frequency band of interest (e.g., theta band: 4–8 Hz; alpha band: 8–13 Hz) during the overall time period ($0\text{--}1200$ ms), the electrodes showing higher p values above a predetermined threshold (the upper 10% highest value) were selected. The average power spectral power was calculated over the selected electrodes for the next step.

- (2) After significant locations were found in step 1, time-frequency bins were screened to be significant among all 6,266 ($=241 \times 26$) bins if p values were below a predetermined threshold ($p < 0.05$). A cluster of time-frequency bins was formed if more than two successive bins were selected along either time or frequency axis. Sum of t values within the cluster, t_{mass} , was then calculated and compared with the null distribution of surrogate data to determine statistical significance of the cluster (above the highest 5% of the null distribution). The null distribution of t_{mass} was obtained from the largest values of t_{mass} for each of 5,000 surrogate data, which were derived by random permutation of “yes” and “no” answers.

2.6. Feature Extraction by Common Spatial Pattern (CSP) Filtering. CSP is a mathematical procedure to derive a spatial filter which separates a multichannel signal into

additive subcomponents so that the differences of variances are maximized between two classes. That is, the most discriminative features between two classes are obtained by maximizing the variance of the spatially filtered signal of one class while minimizing that of the other class [22]. The CSP algorithm is recognized to be effective for the discrimination of mental states from event-related EEG spectral powers [26]. The results of the CSP can be visualized as a topographic map on the scalp, which facilitates interpretation of functional neuroanatomical meanings [26].

The CSP spatial filter, \mathbf{W} , can be obtained by simultaneous diagonalization of two covariance matrices of classes 1 and 2 as follows:

$$\begin{aligned}\mathbf{W}^T \Sigma_1 \mathbf{W} &= \Lambda_1, \\ \mathbf{W}^T \Sigma_2 \mathbf{W} &= \Lambda_2,\end{aligned}\quad (1)$$

where $\Lambda_1 + \Lambda_2 = \mathbf{I}$. Σ_1 and Σ_2 represent the spatial covariance matrices averaged over all single-trial EEGs for each class, and Λ_1 and Λ_2 denote the diagonal matrices. The projection vector, w (column vectors of \mathbf{W}), can be obtained from a generalized eigenvalue decomposition as follows:

$$\Sigma_1 \mathbf{w}_k = \lambda_k \Sigma_2 \mathbf{w}_k, \quad (2)$$

where \mathbf{w}_k ($k = 1, \dots, C$, where C is the number of channels) is the generalized eigenvector, and $\lambda_{1,k} = \mathbf{w}_k^T \Sigma_1 \mathbf{w}_k$ and $\lambda_{2,k} = \mathbf{w}_k^T \Sigma_2 \mathbf{w}_k$ are defined as the k^{th} diagonal element of Λ_1 and Λ_2 , respectively, where $\lambda_k = \lambda_{1,k}/\lambda_{2,k}$. Importantly, $\lambda_{1,k}$ and $\lambda_{2,k}$ (ranges from 0 to 1) reflect the variance for each class and $\lambda_{1,k} + \lambda_{2,k} = 1$. Thus, if $\lambda_{1,k}$ is close to 1, $\lambda_{2,k}$ should be close to 0. This means that corresponding projection vector, \mathbf{w}_k , shows high variance in class 1 but low variance in class 2. The difference in variances between these two classes enables discriminating one class from another. The eigenvalues are sorted in the descending order during calculation, meaning that the first projection vector yields the highest variance for class 1 (but the lowest for class 2), whereas the last projection vector yields the highest variance for class 2 (but lowest for class 1). Thus, the first and last projection vectors are the most useful for the discrimination [22].

The spatial filter \mathbf{W} provides the decomposition of a single-trial multichannel EEG, \mathbf{E} , as $\mathbf{Z} = \mathbf{W}^T \mathbf{E}$, where \mathbf{E} is represented as a matrix with C (the number of channels) rows and T (the number of time samples) columns. The columns of \mathbf{W}^{-1} form the common spatial patterns and can be visualized as topographies on scalp. The variances of the spatially filtered time-series \mathbf{Z} are calculated as features for the classification as follows:

$$f_p = \log \left(\frac{\text{var}(\mathbf{Z}_p)}{\sum_{i=1}^{2m} \text{var}(\mathbf{Z}_i)} \right), \quad \text{where } p = 1, 2, \dots, 2m, \quad (3)$$

where p is the number of features. m was set to 2 which means that the first 2 and last 2 projection vectors were used as features, and thus, the number of features p was 4 for all classifications. The log transformation was adopted to approximate the normal distribution of the data.

2.7. Pattern Classification Using Support Vector Machine (SVM). SVM has been recognized to be a practical and robust method for the classification of human brain signals [27, 28]. The SVM is trained to determine an optimal hyperplane by which the distance to the support vectors (closest to the separating boundary) is maximized [29, 30]. In the case of the linear SVM classification, the hyperplane $\mathbf{a}^T \mathbf{x} + b$ satisfies

$$y_i (\mathbf{a}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \quad \text{for } i = 1, \dots, N, \quad (4)$$

where $\mathbf{x}_i = \{f_{p,i}\}$ denotes a feature vector (in which $p = 1, \dots, 4$) which can be obtained from the CSP algorithm and $y_i \in \{+1, -1\}$ denotes its correct class label. N and ξ_i denote the total number of training samples and the deviation from the optimal condition of linear separability, respectively. The pair of hyperplanes that provide the maximum separating margin can be found by minimizing the cost function $(1/2)\mathbf{a}^T \mathbf{a} + P \sum_{i=1}^N \xi_i$ subject to the constraints

$$\begin{aligned}y_i (\mathbf{a}^T \mathbf{x}_i + b) &\geq 1 - \xi_i, \\ \xi_i &\geq 0, \quad \text{for } i = 1, \dots, N,\end{aligned}\quad (5)$$

where $P > 0$ represents a regularization penalty parameter of the error term. By transforming this optimization problem into its dual problem, the solution may be determined as $\mathbf{a} = \sum_{i=1}^N \alpha_i y_i \mathbf{x}_i$ and achieves equality for nonzero values of α_i only. The corresponding data samples are referred to as support vectors, which are crucial to identify the decision boundary. Instead of the basic linear SVM, we used a radial basis function (RBF) kernel which nonlinearly projects the feature vectors onto a higher dimensional space and thus is better suited for nonlinear relationships between features and class labels [29]. The detailed parameters of the SVM including the RBF kernel parameter and regularization penalty were determined by trial-and-error.

3. Results

3.1. Yes/No Decoding. Figure 3 shows the time-frequency representation of the “yes/no” decoding accuracy averaged over all subjects for each time-frequency subwindow. The time-frequency map of decoding accuracy was generated by representing the decoding accuracies averaged over all subjects within each time-frequency subwindow, which enables estimation of the decoding accuracies over all time-frequency ranges. We used two criteria to define the most important time-frequency subwindows showing high decoding accuracies. The first was to use the threshold level of a decoding accuracy of 75%, determined by the theoretical 95% confidence limits of the chance level when 10 trials per class are used for testing [31]. Another criterion was that the decoding accuracy should be above the mean $+ 2 \times$ standard deviation value (79.34% here). The high decoding accuracies above these two threshold levels were obtained for three subwindows in the alpha and theta bands (as denoted by the three boxes in Figure 3) for both early and late periods. The highest and second highest decoding accuracies were found

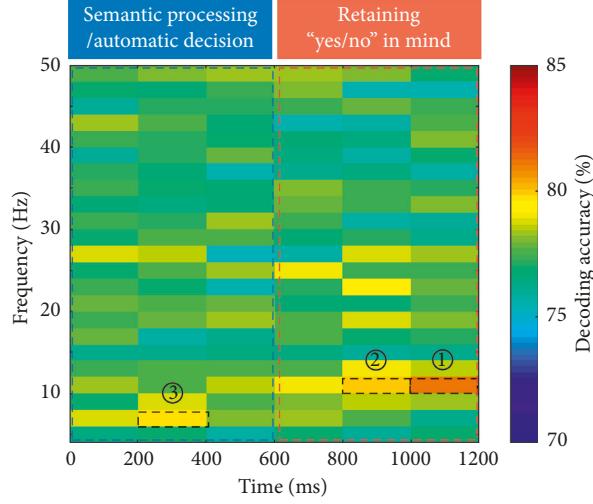


FIGURE 3: Decoding accuracies of time-frequency subwindows. Color code denotes the decoding accuracy averaged over 23 subjects within each time-frequency subwindow. The best three subwindows are denoted by dashed boxes (①: 10–12 Hz, 1000–1200 ms; ②: 10–12 Hz, 800–1000 ms; ③: 6–8 Hz, 200–400 ms).

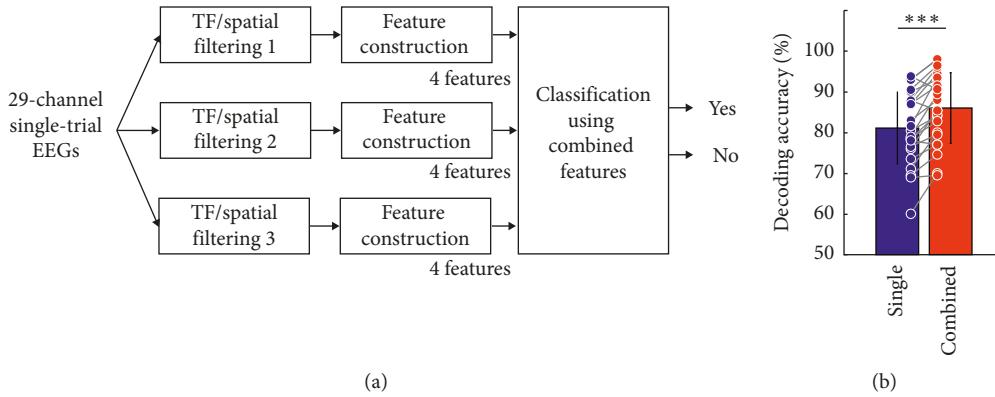


FIGURE 4: Intention decoding using the features from multiple time-frequency subwindows: (a) intention decoder using the features from multiple time-frequency subwindows; (b) statistical comparison of the decoding accuracies between the cases of using single and multiple time-frequency subwindows ($^{***} p < 0.001$, by the paired-sample t -test; error bar: standard deviation).

in the upper alpha band (10–12 Hz) at late epoch (box ①: $81.08 \pm 8.89\%$ at the 1000–1200 ms, box ②: $79.99 \pm 8.99\%$ at the 800–1000 ms). Also, the third highest decoding accuracy was found in the upper theta band (6–8 Hz) at the early period (box ③: $79.76 \pm 10.21\%$ at the 200–400 ms). When all 12 features within these three best time-frequency subwindows were used together, the decoding accuracy was drastically enhanced compared to the best subwindow (10–12 Hz, 1000–1200 ms), as shown in Figure 4(a) (single: $81.08 \pm 8.89\%$, combined: $86.03 \pm 8.69\%$, $t(22) = -5.95$, $p < 0.001$, by the paired-sample t -test). The individual decoding accuracies are presented in Table 1. The sensitivity and specificity values for each time-frequency subwindow are presented in Supplementary Figure 1.

3.2. Spatial Patterns. Figure 5 shows the difference between the most important common spatial patterns for “no” and “yes” answers within the three time-frequency subwindows

(averaged over all subjects). Each topography was obtained from the difference between the last (“no” answer) and first columns (“yes” answer) of the inverse matrix of the projection matrix, \mathbf{W} , for each subject (Supplementary Figure 2), which was calculated in each time-frequency subwindow and then averaged over all subjects. The difference between the most important common spatial patterns in the alpha band showed the strongest coefficient at the right centroparietal region at both 1000–1200 ms and 800–1000 ms periods (the leftmost and middle panels in Figure 5, respectively). The difference between the most important common spatial pattern in the theta band at the 200–400 ms period was focused in the right frontal regions (the rightmost panel in Figure 5).

3.3. Event-Related Spectral Perturbation (ERSP) Analysis. Figure 6(a) shows the topographical distributions of t values averaged within the theta band (4–8 Hz) in 0–1200 ms. The 3

TABLE 1: Decoding accuracies for individual subjects.

Subject	TF subwindow ①	TF subwindow ②	TF subwindow ③	Combined features
1	86.89 ± 6.52 (78.95, 100)	88.95 ± 4.35 (80.00, 95.00)	87.37 ± 9.47 (70.00, 100)	94.50 ± 3.50 (90.00, 100)
2	71.00 ± 6.63 (55.00, 80.00)	74.50 ± 10.11 (55.00, 90.00)	78.00 ± 9.80 (65.00, 95.00)	80.00 ± 10.49 (60.00, 95.00)
3	76.00 ± 10.20 (60.00, 90.00)	79.00 ± 6.63 (65.00, 90.00)	82.00 ± 7.14 (65.00, 90.00)	85.50 ± 7.89 (65.00, 95.00)
4	80.50 ± 8.79 (65.00, 95.00)	75.50 ± 7.23 (65.00, 90.00)	71.00 ± 8.00 (60.00, 85.00)	84.00 ± 3.74 (80.00, 90.00)
5	88.00 ± 3.32 (85.00, 95.00)	89.00 ± 10.44 (65.00, 100)	92.00 ± 7.48 (75.00, 100)	96.00 ± 4.90 (85.00, 100)
6	90.00 ± 7.07 (75.00, 100)	90.50 ± 4.15 (85.00, 95.00)	92.00 ± 5.57 (85.00, 100)	93.00 ± 4.00 (85.00, 100)
7	78.47 ± 13.21 (45.00, 94.74)	77.45 ± 10.00 (65.00, 100)	77.42 ± 9.75 (60.00, 90.00)	84.45 ± 8.17 (65.00, 95.00)
8	83.00 ± 5.57 (75.00, 95.00)	75.50 ± 7.57 (65.00, 85.00)	80.50 ± 7.57 (65.00, 90.00)	88.00 ± 8.72 (75.00, 100)
9	73.50 ± 9.76 (55.00, 85.00)	70.00 ± 8.37 (60.00, 85.00)	68.00 ± 6.40 (55.00, 80.00)	77.00 ± 8.12 (65.00, 90.00)
10	78.64 ± 10.95 (60.00, 94.74)	68.54 ± 8.60 (60.00, 89.47)	76.09 ± 12.77 (50.00, 94.74)	82.73 ± 6.25 (70.00, 90.00)
11	76.54 ± 7.58 (63.16, 85.00)	76.34 ± 9.54 (60.00, 89.47)	67.59 ± 6.00 (60.00, 75.00)	79.54 ± 6.25 (66.67, 90.00)
12	92.95 ± 3.37 (89.47, 100)	94.00 ± 7.35 (75.00, 100)	95.97 ± 3.75 (90.00, 100)	97.97 ± 2.48 (94.74, 100)
13	87.89 ± 6.54 (78.95, 100)	82.84 ± 9.76 (65.00, 95.00)	74.42 ± 12.27 (50.00, 90.00)	85.45 ± 8.17 (70.00, 100)
14	60.00 ± 7.75 (45.00, 75.00)	65.00 ± 10.25 (45.00, 75.00)	69.50 ± 14.04 (50.00, 90.00)	70.00 ± 10.49 (50.00, 85.00)
15	69.58 ± 14.98 (44.44, 88.89)	79.53 ± 10.70 (55.56, 94.44)	58.96 ± 13.75 (33.33, 77.78)	74.55 ± 10.76 (57.89, 94.44)
16	68.95 ± 8.49 (60.00, 90.00)	64.51 ± 6.80 (50.00, 75.00)	67.45 ± 8.21 (55.00, 80.00)	69.48 ± 7.00 (60.00, 85.00)
17	79.00 ± 9.43 (65.00, 90.00)	77.00 ± 7.14 (65.00, 85.00)	83.00 ± 9.00 (70.00, 95.00)	89.50 ± 7.89 (80.00, 100)
18	81.50 ± 8.67 (65.00, 100)	80.00 ± 9.22 (60.00, 90.00)	88.00 ± 7.14 (70.00, 95.00)	93.50 ± 5.50 (80.00, 100)
19	93.00 ± 5.10 (85.00, 100)	97.00 ± 4.00 (90.00, 100)	94.50 ± 2.69 (90.00, 100)	98.00 ± 2.45 (95.00, 100)
20	90.50 ± 4.72 (85.00, 100)	89.50 ± 6.10 (80.00, 100)	88.50 ± 6.73 (75.00, 100)	96.50 ± 4.50 (85.00, 100)
21	87.00 ± 7.48 (70.00, 95.00)	83.00 ± 9.27 (65.00, 100)	83.50 ± 5.50 (75.00, 95.00)	91.50 ± 5.94 (80.00, 100)
22	78.00 ± 8.72 (60.00, 90.00)	73.50 ± 8.08 (55.00, 85.00)	70.00 ± 8.66 (55.00, 85.00)	77.00 ± 4.00 (70.00, 85.00)
23	93.82 ± 5.51 (87.50, 100)	88.62 ± 8.34 (66.67, 100)	88.77 ± 6.86 (75.00, 100)	90.61 ± 9.01 (73.33, 100)
Ave	81.08 ± 8.89 (60.00, 93.82)	79.99 ± 8.99 (64.51, 97.00)	79.76 ± 10.21 (67.45, 95.97)	86.03 ± 8.69 (69.48, 98.00)

Value: mean ± standard deviation (SD) of decoding accuracy (%). The range of decoding accuracies was in parenthesis. Abbreviation: TF, time frequency; Ave, average values over all subjects.

electrodes (FC2, FC6, and C4) showing high t values above a predetermined threshold ($t > 1.62$, corresponding to the highest 10%) were selected over the right frontal region (denoted by black dots in the left panel in Figure 6(a)). Significant “yes/no” difference was found within a single time-frequency range around 200–800 ms in the upper theta and lower alpha bands (6–10 Hz), which was stronger for “no” compared to “yes” (denoted by a solid contour in the left panel in Figure 6(b)).

In the alpha band, 3 electrodes in right parietal area (CP2, Pz, and P4) with high t values were selected as described above ($t > 1.52$, the highest 10%) as denoted by black dots in the right panel in Figure 6(a). The “yes/no” difference in spectral power in this region was significant within a single time-frequency range (300–1200 ms, 9–12 Hz), where the alpha-band power was stronger for “no” compared to “yes” (denoted by a solid contour in the right panel in Figure 6(b)).

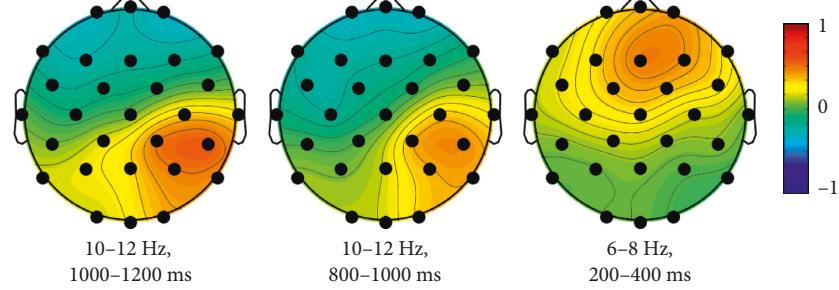


FIGURE 5: Difference between the most important common spatial patterns for “no” and “yes” answers averaged over all subjects within 3 time-frequency subwindows. The topography was obtained from the difference between the last (“no” answer) and first (“yes” answer) columns of the inverse of the matrix, \mathbf{W} , for each subject and then averaged over all subjects.

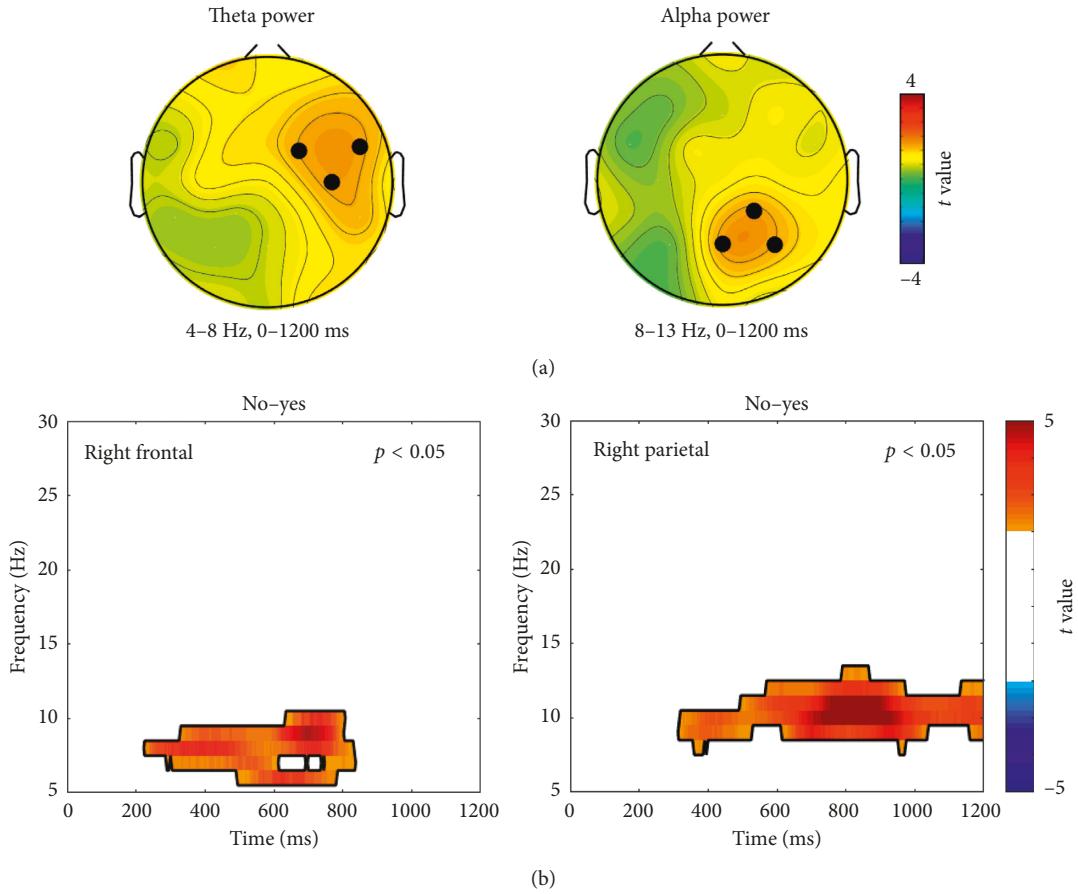


FIGURE 6: Statistical comparisons in spectral power between “yes” and “no.” (a) Topographical distributions of t values (left: theta band (4–8 Hz), 0–1200 ms; right: alpha band (8–13 Hz), 0–1200 ms). Black dots (FC2, FC6, and C4 in the left panel and CP2, Pz, and P4 in the right panel) denote high t values above a predetermined threshold, corresponding to the highest 10%. (b) The clusters in the time-frequency domain showing significant differences between “yes” and “no” in right frontal (left panel) and right parietal regions (right panel) (denoted by black contours).

4. Discussion

We showed that it is possible to decode the intentions to answer “yes” and “no” with high accuracy from single-trial EEGs. The best decoding accuracy averaged over 23 subjects was as high as 86.03% when useful features in multiple time-frequency subwindows were all combined. The decoding accuracy was above 70% for most of the subjects

(22 out of 23 subjects), which is considered as a reasonable accuracy for the binary classification [32]. We decoded the “yes” and “no” answers directly from the brain activities representing the two different answers, which implies that the “mind reading” in a true sense is feasible. The experimental paradigm of our study is based on a natural task which required the subjects to answer self-referential questions as in conversation with others, without any self-

regulation of the brain signals or high cognitive efforts. No unpleasant stimuli and volition or high cognitive efforts are required since our approach is based on a direct decoding of “yes” and “no” without any self-regulation of the brain signals. Birbaumer’s group has suggested a new alternative approach based on classical conditioning to solve the problem of conventional BCI in the CLIS patients [16–18]. For the training, two distinct unconditioned stimuli are presented to the subjects immediately after the simple “yes/no” questions (corresponding to the conditioned stimuli) so that the cortical responses can be conditioned differently for yes and no. The unconditioned stimuli include auditory pink noise and white noise [16, 18] and weak electrical stimulation to the thumb [17]. The main idea of this approach is to modulate the users’ brain activities indirectly through the unconditioned stimuli so that “yes” and “no” can be easily discriminated from neural signals responding to the sensory stimuli, rather than to read the users’ answers from neural signals. This approach may provide an alternative to the conventional BCI approaches in that volition, or high cognitive efforts are not required. However, it remains unclear how long the conditioned cortical response can be maintained considering the extinction effect of classical conditioning [33]. Moreover, unconditioned stimuli such as auditory noise or electrical stimulation can evoke significant displeasure.

Recently, a more natural approach for the “yes/no” decoding was demonstrated based on functional near-infrared spectroscopy (fNIRS) in the CLIS patients [34]. They achieved “yes/no” decoding accuracy over 70% based on fNIRS signals, which were recorded, while the patients answered “yes” or “no” to personal and open questions in minds repeatedly. Interestingly, for the same experimental protocol, they reported that EEG-based decoding yielded accuracy below the chance level. This study employed a natural question/answer task which does not require high cognitive efforts or volition, just as ours. But due to the slow nature of hemodynamics, the duration of each trial for the decoding was quite long (>10 sec). Here, we showed the possibility of “yes/no” decoding from considerably shorter signal recording, which is more beneficial for a practical BCI communication tool.

We took a systematic approach of finding features of brain activities reflecting “yes/no” answers in minds and then developing the decoding algorithm by utilizing these features. Further studies may be necessary to investigate whether the patients, who would potentially benefit from the BCI, can hold the intentions to answer in minds for a short time and to validate our method on the patients’ data.

In this study, the intentions regarding self-referential questions based on the autobiographic facts were investigated. It is important to further try decoding the intentions to answer various types of questions including desire, feeling, and preference. In addition, our questions were presented only in visual stimuli. Neurological patients may have an abnormal visual function such as disability to fix their gaze on specific visual stimuli [35]. Different sensory modality such as auditory stimuli has been tried for the BCI communication tools [34, 36]. It would be beneficial if our

approach can be validated with auditory stimuli such as voice, considering that a high decoding accuracy above 80% was obtained even when the brain activities during the period of retaining the decision in minds (10–12 Hz, 1000–1200 ms) used for the decoding. Thus, we expect that it is possible to decode the “yes” and “no” intentions in a similar way, even if other types of questions and/or the auditory stimuli are employed in the further studies. In addition, here, we did not try to optimize the detailed parameters of the SVM, including the RBF kernel parameter and regularization penalty. The use of the best parameters of the SVM, for example, by using the “grid-search” method [37], may be obviously helpful for better results.

We found two time-frequency regions containing useful information for the “yes/no” decoding, in early theta and late alpha bands. The useful features for the “yes/no” decoding in the alpha band were found to be concentrated in the parietal region at 800–1200 ms from the CSP algorithm. Recently, we showed that the alpha rhythms in the right parietal region are differentiated between the intentions to answers either “yes” or “no” in minds, presumably due to the difference in cognitive loads for the WM retention [2]. Several previous studies showed that the higher parietal alpha power reflects increased memory load [38, 39] or attentional demand [40, 41] during WM retention. The higher alpha power is attributed to active inhibitory control to block incoming stimuli during WM retention, for efficient cortical information processing [38, 39, 42, 43]. Our results showed higher parietal alpha power for “no” compared to “yes,” which may imply higher cognitive load during retaining “no” in minds compared to “yes” [2]. The greater increase in alpha-band activity for “no” may reflect the increased WM load during the intention retention. In Korean language, “yes,” is the one-character word, “네,” and “no” is three-character word, “아니오.” It is plausible that the higher WM load is required to represent intention to respond “no” than “yes” due to the length of the Korean words, resulting in the higher alpha rhythm. This assumption is supported by an ERP study which reported that greater alpha-band power was induced for retaining longer word [44].

It can also be interpreted that the significantly higher alpha-band activity in the centroparietal region for “no” is due to the higher attentional demand [40, 45], and this contributed to the high decoding accuracy. This is also in agreement with a recent study [46], which reported that a higher alpha rhythm was identified in the right parietal cortex for a higher internal attention condition during a divergent thinking task. Our result of greater alpha power for “no” than for “yes” may imply a stronger inhibition of the outer stimuli by the bottom-up attention network for “no,” induced by higher internal attentional demand. This is supported by psychophysical which showed that saying “no” requires more effortful reconsideration after comprehending a sentence and a longer response time for saying “no” than “yes” [47, 48].

The theta-band activity in the frontal region in 200–500 ms was another major feature for “yes/no” decoding. The theta ERS showed topography focused on midline frontal and lateral temporal regions. The difference

between “yes” and “no” was also most prominent in these regions. Hald et al. reported that temporal and frontal theta-band activity in 300–800 ms was significantly higher in semantically incongruent compared to congruent sentences [49]. This is commensurate with our result in that “no” stimuli are incongruent with autobiographic facts. The increase of theta-band activity for semantic incongruence was interpreted to reflect the general error detection mechanism, which is associated with error-related negativity (ERN) [50]. Interestingly, Luu and Tucker showed that frequency domain analysis of the ERN yields theta-band activity in the midfrontal region [50]. A related study reported higher theta oscillation for syntactic violation as well [51]. We observed that frontal theta power in 200–500 ms contributed to high decoding accuracy. Considering the location and frequency band, our result on the usefulness of frontal theta power in 200–500 ms can be interpreted as another evidence, suggesting that error-related frontal theta oscillation is a general phenomenon underlying processing of incoming stimuli containing violation with internal information.

Data Availability

The data used to support the findings of this study have not been made available because some participants of this study did not agree to distribute their physiological signals.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (no. 2017M3C7A1029485).

Supplementary Materials

The supplementary figures demonstrate the sensitivity and specificity of “yes/no” decoding for each time-frequency subwindow (Supplementary Figure 1) and the difference between the most important spatial patterns discriminating “yes” and “no” for the individual subjects at the three time-frequency subwindows (Supplementary Figure 2). (*Supplementary Materials*)

References

- [1] J. W. Choi, K. S. Cha, K. Y. Jung, and K. H. Kim, “Gamma-band neural synchrony due to autobiographical fact violation in a self-referential question,” *Brain Research*, vol. 1662, pp. 39–45, 2017.
- [2] J. W. Choi, K. S. Cha, and K. H. Kim, “Covert intention to answer to self-referential questions is represented in alpha-band local and interregional neural synchronies,” *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 7084186, 11 pages, 2019.
- [3] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, “Brain-computer interfaces for communication and rehabilitation,” *Nature Reviews Neurology*, vol. 12, no. 9, pp. 513–525, 2016.
- [4] G. Gallegos-Ayala, A. Furdea, K. Takano, C. A. Ruf, H. Flor, and N. Birbaumer, “Brain communication in a completely locked-in patient using bedside near-infrared spectroscopy,” *Neurology*, vol. 82, no. 21, pp. 1930–1932, 2014.
- [5] N. Birbaumer, N. Ghanayim, T. Hinterberger et al., “A spelling device for the paralysed,” *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [6] D. Beukelman, S. Fager, and A. Nordness, “Communication support for people with ALS,” *Neurology Research International*, vol. 2011, Article ID 714693, 6 pages, 2011.
- [7] N. Birbaumer, T. Elbert, A. G. Canavan, and B. Rockstroh, “Slow potentials of the cerebral cortex and behavior,” *Physiological Reviews*, vol. 70, no. 1, pp. 1–41, 1990.
- [8] G. Pfurtscheller and A. Aranibar, “Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement,” *Electroencephalography and Clinical Neurophysiology*, vol. 46, no. 2, pp. 138–146, 1979.
- [9] C. Guger, G. Edlinger, W. Harkam, I. G. Niedermayer, and G. Pfurtscheller, “How many people are able to operate an eeg-based brain-computer interface (bci)?,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 145–147, 2003.
- [10] C. Vidaurre and B. Blankertz, “Towards a cure for BCI illiteracy,” *Brain Topography*, vol. 23, no. 2, pp. 194–198, 2010.
- [11] C. Guger, S. Daban, E. Sellers et al., “How many people are able to control a P300-based brain-computer interface (BCI)?,” *Neuroscience Letters*, vol. 462, no. 1, pp. 94–98, 2009.
- [12] F. Nijboer, E. W. Sellers, J. Mellinger et al., “A P300-based brain-computer interface for people with amyotrophic lateral sclerosis,” *Clinical Neurophysiology*, vol. 119, no. 8, pp. 1909–1916, 2008.
- [13] C. Guger, B. Z. Allison, B. Großwindhager et al., “How many people could use an SSVEP BCI?,” *Frontiers in Neuroscience*, vol. 6, p. 169, 2012.
- [14] G. R. Müller-Putz, R. Scherer, C. Brauneis, and G. Pfurtscheller, “Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components,” *Journal of Neural Engineering*, vol. 2, no. 4, pp. 123–130, 2005.
- [15] A. Kübler and N. Birbaumer, “Brain-computer interfaces and communication in paralysis: extinction of goal directed thinking in completely paralysed patients?,” *Clinical Neurophysiology*, vol. 119, no. 11, pp. 2658–2666, 2008.
- [16] A. Furdea, C. A. Ruf, S. Halder et al., “A new (semantic) reflexive brain-computer interface: in search for a suitable classifier,” *Journal of Neuroscience Methods*, vol. 203, no. 1, pp. 233–240, 2012.
- [17] D. De Massari, T. Matuz, A. Furdea, C. A. Ruf, S. Halder, and N. Birbaumer, “Brain-computer interface and semantic classical conditioning of communication in paralysis,” *Biological Psychology*, vol. 92, no. 2, pp. 267–274, 2013.
- [18] C. A. Ruf, D. De Massari, A. Furdea et al., “Semantic classical conditioning and brain-computer interface control: encoding of affirmative and negative thinking,” *Frontiers in Neuroscience*, vol. 7, p. 23, 2013.
- [19] A. Delorme and S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.

- [20] T.-P. Jung, S. Makeig, C. Humphries et al., "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.
- [21] G. Huang, G. Liu, J. Meng, D. Zhang, and X. Zhu, "Model based generalization analysis of common spatial pattern in brain computer interfaces," *Cognitive Neurodynamics*, vol. 4, no. 3, pp. 217–223, 2010.
- [22] H. Ramoser, J. G. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 4, pp. 441–446, 2000.
- [23] C. Tallon-Baudry, O. Bertrand, C. J. Delpuech, and J. Pernier, "Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human," *Journal of Neuroscience*, vol. 16, no. 13, pp. 4240–4249, 1996.
- [24] C. S. Herrmann, M. H. J. Munk, and A. K. Engel, "Cognitive functions of gamma-band activity: memory match and utilization," *Trends in Cognitive Sciences*, vol. 8, no. 8, pp. 347–355, 2004.
- [25] E. Maris and R. Oostenveld, "Nonparametric statistical testing of EEG- and MEG-data," *Journal of Neuroscience Methods*, vol. 164, no. 1, pp. 177–190, 2007.
- [26] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 41–56, 2008.
- [27] J. R. Sato, A. Fujita, C. E. Thomaz et al., "Evaluating SVM and MLDA in the extraction of discriminant regions for mental state prediction," *Neuroimage*, vol. 46, no. 1, pp. 105–114, 2009.
- [28] F. Lotte, L. Bougrain, A. Cichocki et al., "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, article 031005, 2018.
- [29] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 1–43, 1998.
- [30] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [31] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random? A closer look on BCI results," *International Journal of Bioelectromagnetism*, vol. 10, pp. 52–5, 2008.
- [32] J. Perelmouter and N. Birbaumer, "A binary spelling interface with random errors," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 227–232, 2000.
- [33] S. J. Shettleworth, *Cognition, Evolution, and behavior*, Oxford University Press, Oxford, UK, 1998.
- [34] U. Chaudhary, B. Xia, S. Silvoni, L. G. Cohen, and N. Birbaumer, "Brain-computer interface-based communication in the completely locked-in state," *PLOS Biology*, vol. 15, no. 1, article e1002593, 2017.
- [35] R. Sharma, S. Hicks, C. M. Berna, C. Kennard, K. Talbot, and M. R. Turner, "Oculomotor dysfunction in amyotrophic lateral sclerosis," *Archives of Neurology*, vol. 68, no. 7, pp. 857–861, 2011.
- [36] N. J. Hill, E. Ricci, S. Haider et al., "A practical, intuitive brain-computer interface for communicating "yes" or "no" by listening," *Journal of Neural Engineering*, vol. 11, no. 3, article 035003, 2014.
- [37] C. W. Hsu, C. C. Chang , Lin et al., "A practical guide to support vector classification," Technical Report, pp. 1–12,
- [38] Department of Computer Science and Information Engineering, University of National Taiwan, Taipei, Taiwan, 2003.
- [39] O. Jensen, J. Gelfand, J. Kounios, and E. Lisman John, "Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task," *Cerebral Cortex*, vol. 12, no. 8, pp. 877–882, 2002.
- [40] W. Klimesch, P. Sauseng, and S. Hanslmayr, "EEG alpha oscillations: the inhibition-timing hypothesis," *Brain Research Reviews*, vol. 53, no. 1, pp. 63–88, 2007.
- [41] N. R. Cooper, R. J. Croft, S. J. J. Dominey, A. P. Burgess, and J. H. Gruzelier, "Paradox lost? exploring the role of alpha oscillations during externally vs. internally directed attention and the implications for idling and inhibition hypotheses," *International Journal of Psychophysiology*, vol. 47, no. 1, pp. 65–74, 2003.
- [42] M. Benedek, R. J. Schickel, E. Jauk, A. Fink, and A. C. Neubauer, "Alpha power increases in right parietal cortex reflects focused internal attention," *Neuropsychologia*, vol. 56, pp. 393–400, 2014.
- [43] C. S. Herrmann, D. Senkowski, and S. Röttger, "Phase-locking and amplitude modulations of EEG alpha: two measures reflect different cognitive processes in a working memory task," *Experimental Psychology*, vol. 51, no. 4, pp. 311–318, 2004.
- [44] P. Sauseng, W. Klimesch, M. Doppelmayr, T. Pecherstorfer, R. Freunberger, and S. Hanslmayr, "EEG alpha synchronization and functional coupling during top-down processing in a working memory task," *Human Brain Mapping*, vol. 26, no. 2, pp. 148–155, 2005.
- [45] L. Michels, M. Moazami-Goudarzi, D. Jeanmonod, and J. Sarnthein, "EEG alpha distinguishes between cuneal and precuneal activation in working memory," *Neuroimage*, vol. 40, no. 3, pp. 1296–1310, 2008.
- [46] W. Ray and H. Cole, "EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes," *Science*, vol. 228, no. 4700, pp. 750–752, 1985.
- [47] M. Benedek, E. Jauk, A. Fink et al., "To create or to recall? Neural mechanisms underlying the generation of creative new ideas," *NeuroImage*, vol. 88, pp. 125–133, 2014.
- [48] E. S. Knowles and C. A. Condon, "Why people say "yes": a dual-process theory of acquiescence," *Journal of Personality and Social Psychology*, vol. 77, no. 2, pp. 379–386, 1999.
- [49] M. R. Sheridan and K. A. Flowers, "Reaction times and deception - the lying constant," *International Journal of Psychological Studies*, vol. 2, no. 2, pp. 41–51, 2010.
- [50] L. A. Hald, M. C. M. Bastiaansen, and P. Hagoort, "EEG theta and gamma responses to semantic violations in online sentence processing," *Brain and Language*, vol. 96, no. 1, pp. 90–105, 2006.
- [51] P. Luu and D. M. Tucker, "Regulating action: alternating activation of midline frontal and motor cortical networks," *Clinical Neurophysiology*, vol. 112, no. 7, pp. 1295–1306, 2001.
- [52] M. C. M. Bastiaansen, J. J. A. van Berkum, and P. Hagoort, "Syntactic processing modulates the θ rhythm of the human EEG," *Neuroimage*, vol. 17, no. 3, pp. 1479–1492, 2002.

Research Article

Comparison of Visual Stimuli for Steady-State Visual Evoked Potential-Based Brain-Computer Interfaces in Virtual Reality Environment in terms of Classification Accuracy and Visual Comfort

Kang-min Choi , Seonghun Park , and Chang-Hwan Im 

Department of Biomedical Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

Correspondence should be addressed to Chang-Hwan Im; ich@hanyang.ac.kr

Received 26 March 2019; Accepted 3 June 2019; Published 1 July 2019

Guest Editor: Hyun J. Baek

Copyright © 2019 Kang-min Choi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Recent studies on brain-computer interfaces (BCIs) based on the steady-state visual evoked potential (SSVEP) have demonstrated their use to control objects or generate commands in virtual reality (VR) environments. However, most SSVEP-based BCI studies performed in VR environments have adopted visual stimuli that are typically used in conventional LCD environments without considering the differences in the rendering devices (head-mounted displays (HMDs) used in the VR environments). The proximity between the visual stimuli and the eyes in HMDs can readily cause eyestrain, degrading the overall performance of SSVEP-based BCIs. Therefore, in the present study, we have tested two different types of visual stimuli—pattern-reversal checkerboard stimulus (PRCS) and grow/shrink stimulus (GSS)—on young healthy participants wearing HMDs. Preliminary experiments were conducted to investigate the visual comfort of each participant during the presentation of the visual stimuli. In subsequent online avatar control experiments, we observed considerable differences in the classification accuracy of individual participants based on the type of visual stimuli used to elicit SSVEP. Interestingly, there was a close relationship between the subjective visual comfort score and the online performance of the SSVEP-based BCI: most participants showed better classification accuracy under visual stimulus they were more comfortable with. Our experimental results suggest the importance of an appropriate visual stimulus to enhance the overall performance of the SSVEP-based BCIs in VR environments. In addition, it is expected that the appropriate visual stimulus for a certain user might be readily selected by surveying the user's visual comfort for different visual stimuli, without the need for the actual BCI experiments.

1. Introduction

Electroencephalography (EEG) has been the most widely used neural signal for brain-computer interfaces (BCIs), whose main aim is to provide the paralyzed or disabled with new means of communication with the external environment [1]. Typical paradigms for EEG-based BCIs include motor imagery (MI), P300, and steady-state visual evoked potential (SSVEP) [2]. Among these, an SSVEP-based BCI paradigm has been widely employed because of its robustness to external noises and very little training requirement [3]. Owing to its advantages over the other paradigms and recent development of advanced analysis

methods [4, 5], the SSVEP-based BCIs have been implemented for a variety of applications including assistive and rehabilitation tools for the disabled [6] and practical applications for the healthy, such as car navigation [7] and entertainment [8]. Furthermore, with the rapid advancements in the virtual reality (VR) technology, the SSVEP-based BCIs have been successfully applied to VR applications with hand-free control of the VR objects or speechless communications [9–11].

Although most VR devices currently employ head-mounted displays (HMDs), no previous SSVEP-based BCI study has considered the environmental differences between the VR-HMDs and conventional LCD monitors. Since the

traditional SSVEP-based BCIs have used an LCD monitor as the rendering device to present visual stimuli for the past decades, a number of studies have already been conducted on the influence of the various parameters of this visual stimulus on the performance of the BCIs; these parameters include spatial frequency [12], temporal frequencies [13], colors [14], data recording channels [15], and time window sizes [16, 17]. On the contrary, the SSVEP-based BCIs implemented in VR environments have employed visual stimuli identical to those used in conventional LCD monitor environments, without any major modification. In other words, all SSVEP-based BCI studies performed in VR environments assumed that the presentation of visual stimuli on HMD is not significantly different from that on an LCD monitor. For example, MindBalance game [9], a 3D video game using SSVEP-based BCIs in VR environments, employed pattern-reversal checkerboard stimulus (PRCS) to elicit SSVEP response. A recently developed neuro-optical diagnostic tool using the VR headset [18] also employed the conventional PRCS. However, it is well known that an experiment in the VR environment is highly vulnerable to visual fatigue than that in the LCD environment; this is mainly due to the image distortion, or crosstalk, in the stereoscopic viewing [19] as well as the proximity between the source of illumination and the eyes [20].

In the present study, we have used two different types of visual stimuli—PRCS and grow/shrink stimulus (GSS)—both of which are known to effectively elicit SSVEP responses in the LCD monitor environment, on 14 participants wearing HMDs. The performance of the two representative visual stimuli was then investigated in terms of individual classification accuracy and subjective visual comfort scores. After the survey of the visual comfort of the participants in the preliminary offline experiments, the performance of SSVEP-based BCIs was investigated through online avatar control experiments in a VR environment.

2. Materials and Methods

2.1. Participants. Sixteen young, healthy people (10 males and 6 females, aged 20.5 ± 1.6 years) with normal or corrected-to-normal vision participated in our experiment. All participants were informed of the details of the experiments and had given their written consent. The data of two participants were excluded in further analyses: the first was excluded owing to the frequent blinking of the eyes during the presentation of the visual stimuli (eye blinks contaminated 14 out of the total 40 trials) and the second owing to the nonexistence of spectral peaks in the recorded EEG. The eye blinks were identified by visually inspecting vertical electrooculogram (EOG) recorded during the offline experiment. This so-called “BCI-illiteracy” is a well-known issue in EEG-based BCIs [21]. This experiment was approved by the institutional review board of Hanyang University, Republic of Korea (IRB HYI-14-167-11).

2.2. Visual Stimuli. Two different types of visual stimuli were employed to elicit SSVEP responses: a PRCS and a GSS. The

PRCS is a traditional visual stimulus, which is used most frequently to elicit SSVEP responses in LCD monitor environments; this stimulus alternately presents two checkerboard patterns with 180° phase difference [7] (Figure 1(a)). The GSS is a new visual stimulus that changes both luminance and size to elicit SSVEP responses. This stimulus was based on previous studies, which reported that motional changes can also elicit periodic VEP responses (often referred to as steady-state motion visual evoked potential or SSMVEP) [22, 23] (Figure 1(a)). These stimuli were presented in a VR environment using an HMD of the HTC VIVE™ VR system (HTC Co., Ltd., Xindian District, New Taipei City, Taiwan). Both visual stimuli were modulated to elicit SSVEP responses corresponding to four frequencies, namely, 6, 7.5, 9, and 10 Hz. These frequencies were determined by considering the refresh rate of the rendering device (90 Hz), which is an integer multiple of each of the four target frequencies. In the offline experiments, the visual angle of the PRCS was fixed at 14° , while that of the GSS was varied between 8° and 16° . In the online experiments, the visual angle of the PRCS was reduced to 6° and that of GSS was varied between 4° and 8° in order to validate the feasibility and usability of the visual stimuli in a realistic VR environment in which large-sized stimuli cannot be generally employed. Note that according to previous reports, visual stimuli with visual angles greater than 3.8° would produce similar levels of SSVEP responses [24].

2.3. Experimental Paradigm. In the preliminary offline experiments, each stimulus type was presented in a randomly shuffled order to each participant. In each trial, four visual stimuli with different frequencies were presented for 4 s, as shown in Figure 1(a). The interstimulus interval (ISI) was set to 2 s, during which one of the numbers presented on the screen was colored green and flickered at 1 Hz to indicate the stimulus that the participant should focus on during the next stimulus interval. Each visual stimulus in each stimulus type appeared 20 times (five times for each frequency), and thus, the total number of trials was 40. The EEG signals were recorded; however, no immediate feedback was delivered to the participants during the experiment. At the end of the preliminary offline experiment, the participants were asked to subjectively rate their visual comfort with the two stimulus types on a scale ranging from 0 (very uncomfortable) to 10 (very comfortable).

In the online experiments, the participants who also participated in the preliminary offline experiments were asked to control a human full body avatar standing on a virtual road in a VR environment. The avatar could move in four directions: top, bottom, left, and right. Four visual stimuli with the frequencies used in the offline experiment were presented at the top, bottom, left, and right of the avatar to indicate the possible movement directions of the avatar (Figure 1(b)). Each participant was asked to sequentially move the avatar in a correct direction following the given path. A total of three different paths, each consisting of 20 movement steps, were created. For all 60 movement steps, the numbers of each directional step were counterbalanced.

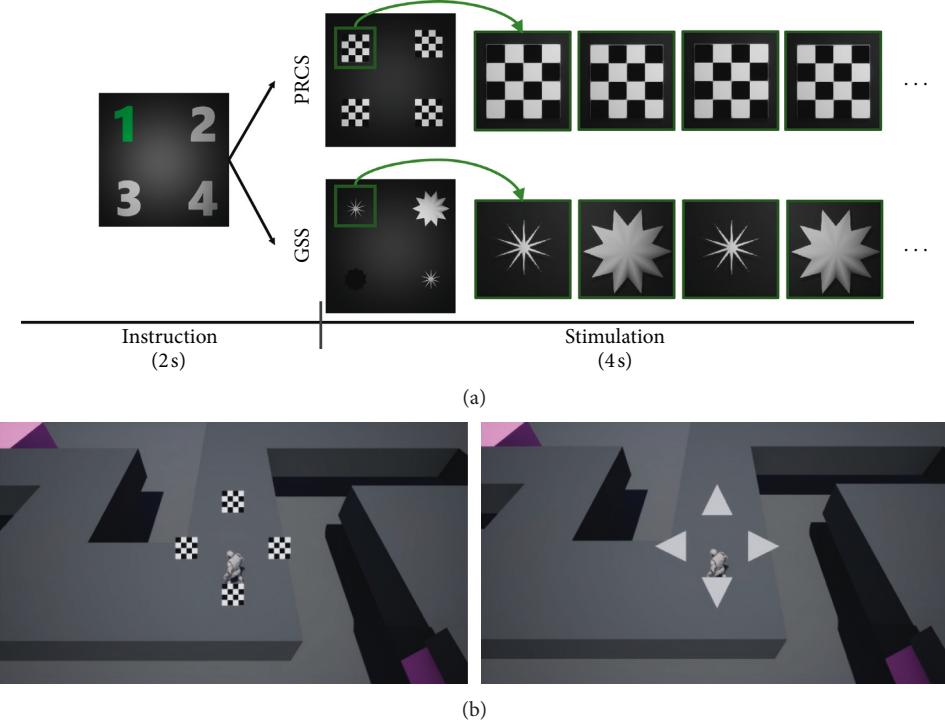


FIGURE 1: (a) Overall timeline for a single trial in offline experiments. In each trial, after a 2 s instruction period to inform participants of the location of a target visual stimulus, four visual stimuli, each of which was either PRCS or GSS, were presented for 4 s. (b) Screenshots of online experiments. The left picture was taken when the PRCS was employed, while the right picture was taken when the GSS was employed. The human avatar needed to move along a designated path (e.g., in the left direction in both figures).

For each path, the same paradigm was repeated twice with either PRCS or GSS, when the presentation order of the visual stimuli was randomly determined for each participant. The avatar could move a step forward only when the classification result (direction) coincided with the correct direction of the path. Consequently, the minimum number of trials required to complete each session was 20, when the classification accuracy was 100%. Each trial lasted for 5 s, including 2 s for the presentation of the visual stimuli, 1 s for avatar's movement, and 2 s for ISI to give participants the time to shift their gaze for the next movement. A video clip showing the online experiment of a participant is attached to this manuscript as a Supplementary Movie, and its high resolution version can be found at YouTube™ (<https://youtu.be/TC4QMPPhW6y8>).

2.4. Biosignal Acquisition and Preprocessing. The EEG data were recorded from seven electrodes (C_z , PO_3 , PO_z , PO_4 , O_1 , O_z , and O_2) using a commercial biosignal recording system (ActiveTwo, BioSemi, Amsterdam, and the Netherlands). In addition, a pair of electrodes was attached above and below the right eye to acquire the vertical EOG data. The sampling rate was set at 2,048 Hz. The recorded EEG data were re-referenced to C_z [4, 25] and then band-pass filtered at 6 and 50 Hz using a zero-phase Chebyshev type I infinite impulse response filter implemented in MATLAB (MathWorks, Inc., Natick, MA, USA). The program to analyze data in real time was developed using the FieldTrip toolbox [26].

2.5. Data Analysis and Statistical Analysis. For the classification of the SSVEP responses, we adopted a recently introduced algorithm called the extension of the multivariate synchronization index (EMSI) [5], which has exhibited outstanding performance compared to the conventional classification methods [27].

The Wilcoxon signed-rank test was employed for the statistical analysis because the classification accuracies with respect to the two visual stimulus types did not follow normal distribution as assessed by the Kolmogorov-Smirnov test.

3. Results

In the offline experiment, the GSS outperformed the PRCS in both classification accuracy and information transfer rate (ITR) for all window sizes (Figures 2 and 3); ITR was calculated as follows:

$$ITR = \frac{(\log_2 N + P \log_2 P + (1 - P) \log_2 (1 - P)) / (N - 1)}{C}, \quad (1)$$

where N denotes the number of stimuli, P denotes the classification accuracy ranging from 0 to 1, and C denotes the time needed to classify a single trial [28]. Statistical analysis using the Wilcoxon signed-rank test also showed statistically significant difference in the performance of the GSS and PRCS (Bonferroni-corrected $p < 0.005$ for both classification accuracy and ITR for all window sizes).

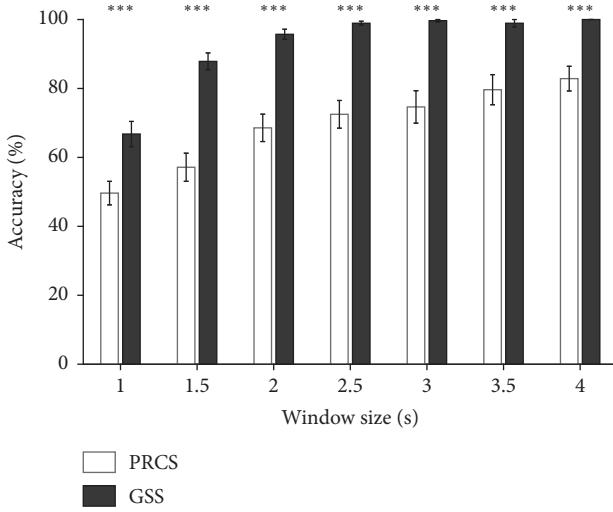


FIGURE 2: Comparison of offline experimental results between the PRCS and GSS in terms of the average classification accuracy across participants. The error bars indicate the standard errors. *** $p < 0.005$.

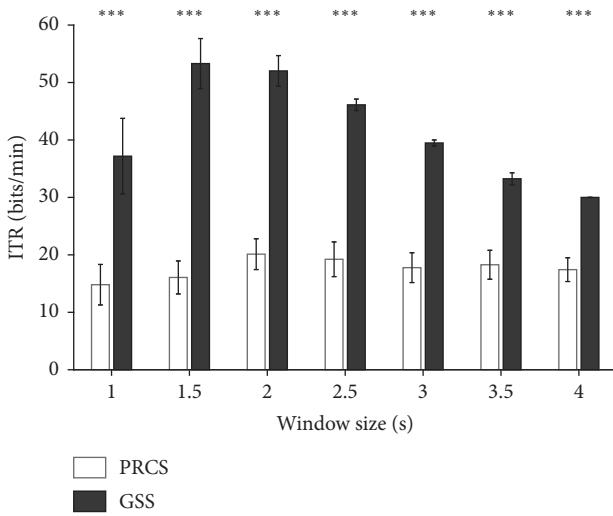


FIGURE 3: Comparison of offline experimental results between the PRCS and GSS in terms of the average ITR across participants. The error bars indicate the standard errors. *** $p < 0.005$.

Although a window size of 1.5 s showed the highest ITR (Figure 3), 2 s epochs were used for the classification in the online experiments. This was because the difference between the ITRs for the 1.5 s and 2 s epochs was not big, but the improvement in the classification accuracy was relatively distinct for the 2 s epoch compared with the 1.5 s epoch.

Table 1 shows the classification accuracy of each participant in the online experiment. Unlike the preliminary offline experiment, no statistical significance was observed between the classification accuracies for the PRCS and GSS ($p = 0.424$; Wilcoxon signed-rank test) in the online experiment although the average classification accuracy for the GSS was higher than that for the PRCS.

TABLE 1: Comparison of online classification accuracies between the PRCS and GSS.

Group	Participant	PRCS accuracy (%)	GSS accuracy (%)
Group 1	P6	84.5	58.8
	P8	92.3	85.7
	P10	98.4	92.3
Group 2	P2	100	100
	P3	66.7	98.4
	P5	96.8	76.9
	P7	100	98.4
	P11	89.6	98.4
	P12	80.0	80.0
Group 3	P13	84.5	95.2
	P1	70.6	89.6
	P4	84.5	90.9
	P9	75.9	92.3
	P14	83.3	96.8
	Average \pm std.	86.2 ± 10.7	89.6 ± 11.3

Group 1 includes participants who rated PRCS as more comfortable to their eyes than GSS. Group 2 includes participants who rated GSS as more comfortable than PRCS. The remaining participants who gave the same score to both stimuli are categorized as Group 3.

The possible reasons for the difference between the two cases, i.e., the offline and online experiments, will be discussed in Discussion.

For further analyses, all participants were divided into three groups based on the subjective visual comfort ratings for the two visual stimulus types that were obtained right after the preliminary offline experiment. The participants who were more comfortable with the PRCS were categorized as Group 1, and those who were more comfortable with the GSS were categorized as Group 2. The participants who rated both stimuli equally were categorized as Group 3 and excluded from further analyses. Interestingly, all three participants (i.e., P6, P8, and P10) in Group 1 exhibited higher classification accuracies for the PRCS than for the GSS, while most participants in Group 2, with the exception of only one participant (i.e., P5), exhibited higher or equivalent classification accuracies for the GSS than for the PRCS. These results suggest that the performance of the SSVEP-based BCIs in VR environments might be potentially improved by selecting the best stimulus type for each individual, which would be readily chosen by inspecting the individual's subjective visual comfort for different visual stimuli types.

4. Discussion

The performances of the reactive BCI systems are highly dependent on the types of stimuli used to elicit specific EEG responses. Although a series of studies has been performed to find an optimal visual stimulus for the conventional SSVEP-based BCIs in the LCD monitor environment, no study has yet been reported on the influence of visual stimuli on the performance of the SSVEP-based BCIs in VR-HMD environments. We hypothesized that the PRCS, which are widely used in the SSVEP-based BCIs, might not be the optimal visual stimulus in a VR-HMD environment

because the images displayed on the HMDs are closer to the eyes than those on the LCD monitors, and thus, the PRCS might be too intense for the eyes. Therefore, in this study, we tested another type of visual stimulus called the GSS that changes both size and luminance in VR environments and compared the BCI performances with the PRCS.

In the offline experimental results, the GSS outperformed the PRCS in terms of classification accuracy; however, the difference in the performance was considerably reduced in the online experiments. This phenomenon is thought to originate from several factors: first, the spatial frequency of the PRCS in the offline experiment was different from that in the online experiment. The spatial frequency changed from 0.25 cycle/deg in the offline experiment to 0.5 cycle/deg in the online experiment. According to a previous report [12], spatial frequency of PRCS has close relationship with the performance of SSVEP-based BCIs. The second reason might be the difference in the background; for instance, in the offline experiment, a monotonous dark grey background was used, while in the online experiment, a relatively complicated background with many distractors was employed (Figure 1(b)). This complicated background might have distracted the elicitation of the SSMVEP because the border of the GSS sometimes becomes obscure owing to the background images. On the contrary, the PRCS would be less affected by the background because this stimulus maintains its size during the presentation.

Our online experiments demonstrated that the SSVEP-based BCI with a visual stimulus that was more comfortable for the user generally outperformed that with the other stimulus in VR environment. This finding is not in line with previous reports showing that a visual stimulus evoking stronger SSVEP responses induced the severer visual fatigue [29–31] when an LCD monitor was used for presenting visual stimuli. However, there are also some evidences showing that the relationship between visual comfort and BCI performance is dependent upon the stimulation rendering device (e.g., light emitting diodes: LEDs) or stimulus types (e.g., SSMVEP) [32, 33]. Our results also suggest that a user's optimal visual stimulus in VR environments might be readily determined by rating the subjective visual comfort of the user even before the main BCI experiment. This strategy might considerably alleviate the necessity of a series of offline BCI experiments to determine an optimal visual stimulus for the user in the VR environment.

In the offline experiment, four participants rated the same visual comfort score for both PRCS and GSS. Interestingly, they commonly achieved better classification accuracies in the GSS than in the PRCS. Although the limited sample size makes it hard to generalize, selecting GSS rather than PRCS might yield better classification accuracies in cases when there is no difference in the subjective visual comfort ratings. However, further investigations are required to formulate a more generalized rule for selecting the optimal visual stimulus for the SSVEP-based BCIs in VR environments. In addition, in the present study, we tested only two types of visual stimuli; however, more types of

visual stimuli need to be developed and tested in VR environments in future studies.

5. Conclusions

To the best of our knowledge, this is the first study that has compared different types of visual stimuli for the SSVEP-based BCIs in VR environments. Our study demonstrated that selection of an optimal visual stimulus for an individual could improve the overall performance of the SSVEP-based BCIs and reduce visual fatigue in VR environment. A close association between the performance of the SSVEP-based BCIs and subjective visual comfort was observed, suggesting that the selection of an appropriate visual stimulus via a simple pre-experimental inspection of the individual's preference toward the visual stimuli might help to enhance the performance of the SSVEP-based BCIs in VR environments.

Data Availability

The data used to support the findings of this study can be made available form the corresponding author upon request.

Disclosure

Kang-min Choi and Seonghun Park are the co-first authors.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Kang-min Choi and Seonghun Park equally contributed to this study.

Acknowledgments

This work was supported by the Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (2017-0-00432, Development of non-invasive integrated BCI SW platform to control home appliances and external devices by user's thought via AR/VR interface).

Supplementary Materials

A movie clip showing the online experiment is attached to this manuscript. Two consecutive trials for each type of stimuli were recorded as a video clip to illustrate the brief paradigm of this experiment: one for the case of the classification result being correct, while the other for that being incorrect. As mentioned in Materials and Methods, the avatar moved only in cases when the classification result coincided with the direction of the path. (*Supplementary Materials*)

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] E. Yin, T. Zeyl, R. Saab, T. Chau, D. Hu, and Z. Zhou, "A hybrid brain-computer interface based on the fusion of P300 and SSVEP scores," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 4, pp. 693–701, 2015.
- [3] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method," *Journal of Neural Engineering*, vol. 6, no. 4, article 046002, 2009.
- [4] Y. Zhang, E. Yin, F. Li et al., "Two-stage frequency recognition method based on correlated component analysis for SSVEP-based BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, pp. 1314–1323, 2018.
- [5] Y. Zhang, D. Guo, D. Yao, and P. Xu, "The extension of multivariate synchronization index method for SSVEP-based BCI," *Neurocomputing*, vol. 269, pp. 226–231, 2017.
- [6] D. Lesenfants, D. Habbal, Z. Lugo et al., "An independent SSVEP-based brain-computer interface in locked-in syndrome," *Journal of Neural Engineering*, vol. 11, no. 3, article 035002, 2014.
- [7] P. Martinez, H. Bakardjian, and A. Cichocki, "Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 94561, 9 pages, 2007.
- [8] M. Van Vliet, A. Robben, N. Chumerin et al., "Designing a brain-computer interface controlled video-game using consumer grade EEG hardware," in *Proceedings of the 2012 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*, pp. 1–6, IEEE, Rio de Janeiro, Brazil, 2012.
- [9] E. C. Lalor, S. P. Kelly, C. Finucane et al., "Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, no. 19, pp. 3156–3164, 2005.
- [10] C. G. Coogan and B. He, "Brain-computer interface control in a virtual reality environment and applications for the internet of things," *IEEE Access*, vol. 6, pp. 10840–10849, 2018.
- [11] J. Faller, B. Z. Allison, C. Brunner et al., "A feasibility study on SSVEP-based interaction with motivating and immersive virtual and augmented reality," 2017, <https://arxiv.org/abs/1701.03981>.
- [12] N. R. Waytowich, Y. Yamani, and D. J. Krusienski, "Optimization of checkerboard spatial frequencies for steady-state visual evoked potential brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 6, pp. 557–565, 2017.
- [13] N. R. Waytowich and D. J. Krusienski, "Novel characterization of the steady-state visual evoked potential spectrum of EEG," in *Proceedings of ACM SIGKDD: Brain KDD Workshop*, New York City, NY, USA, 2014.
- [14] M. Gerloff and M. Schilling, "Subject response variability in terms of colour and frequency of capacitive SSVEP measurements," *Biomedical Engineering/Biomedizinische Technik*, vol. 57, pp. 95–98, 2012.
- [15] H.-J. Hwang, J.-H. Lim, Y.-J. Jung, H. Choi, S. W. Lee, and C.-H. Im, "Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard," *Journal of Neuroscience Methods*, vol. 208, no. 1, pp. 59–65, 2012.
- [16] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain-computer interface (BCI) speller," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1447–1456, 2015.
- [17] J. Jiang, E. Yin, C. Wang, M. Xu, and D. Ming, "Incorporation of dynamic stopping strategy into the high-speed SSVEP-based BCIs," *Journal of Neural Engineering*, vol. 15, no. 4, article 046025, 2018.
- [18] C. Versek, A. Rissmiller, A. Tran et al., "Portable system for neuro-optical diagnostics using virtual reality display," *Military Medicine*, vol. 184, pp. 584–592, 2019.
- [19] T. Bando, A. Iijima, and S. Yano, "Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: a review," *Displays*, vol. 33, no. 2, pp. 76–83, 2012.
- [20] J. Guo, D. Weng, H. B.-L. Duh, Y. Liu, and Y. Wang, "Effects of using HMDs on visual fatigue in virtual environments," in *Proceedings of the 2017 IEEE Virtual Reality (VR)*, pp. 249–250, Los Angeles, CA, USA, March 2017.
- [21] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, "BCI demographics: how many (and what kinds of) people can use an SSVEP BCI?," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 107–116, 2010.
- [22] J. Xie, G. Xu, J. Wang, F. Zhang, and Y. Zhang, "Steady-state motion visual evoked potentials produced by oscillating Newton's rings: implications for brain-computer interfaces," *PLoS One*, vol. 7, no. 6, Article ID e39707, 2012.
- [23] J. Xie, G. Xu, J. Wang et al., "Effects of mental load and fatigue on steady-state evoked potential based brain computer interface tasks: a comparison of periodic flickering and motion-reversal based visual attention," *PLoS One*, vol. 11, no. 9, Article ID e0163426, 2016.
- [24] K. B. Ng, A. P. Bradley, and R. Cunnington, "Stimulus specificity of a steady-state visual-evoked potential-based brain-computer interface," *Journal of Neural Engineering*, vol. 9, no. 3, article 036008, 2012.
- [25] S. M. Lai, Z. Zhang, Y. S. Hung, Z. Niu, and C. Chang, "A chromatic transient visual evoked potential based encoding/decoding approach for brain-computer interface," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 578–589, 2011.
- [26] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Computational Intelligence and Neuroscience*, vol. 2011, Article ID 156869, 9 pages, 2011.
- [27] Y. Zhang, P. Xu, K. Cheng, and D. Yao, "Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface," *Journal of Neuroscience Methods*, vol. 221, pp. 32–39, 2014.
- [28] J. R. Wolpaw, H. Ramoser, D. J. McFarland, and G. Pfurtscheller, "EEG-based communication: improved accuracy by response verification," *IEEE Transactions on Rehabilitation Engineering*, vol. 6, no. 3, pp. 326–333, 1998.
- [29] A. Duszyk, M. Bierzyńska, Z. Radzikowska et al., "Towards an optimization of stimulus parameters for brain-computer interfaces based on steady state visual evoked potentials," *PLoS One*, vol. 9, no. 11, Article ID e112099, 2014.
- [30] A. M. Dreyer, C. S. Herrmann, and J. W. Rieger, "Tradeoff between user experience and BCI classification accuracy with frequency modulated steady-state visual evoked potentials," *Frontiers in Human Neuroscience*, vol. 11, 2017.

- [31] T. Sakurada, T. Kawase, T. Komatsu, and K. Kansaku, "Use of high-frequency visual stimuli above the critical flicker frequency in a SSVEP-based BMI," *Clinical Neurophysiology*, vol. 126, no. 10, pp. 1972–1978, 2015.
- [32] S. Mouli and R. Palaniappan, "Toward a reliable PWM-based light-emitting diode visual stimulus for improved SSVEP response with minimal visual fatigue," *Journal of Engineering*, vol. 2017, no. 2, pp. 7–12, 2017.
- [33] W. Yan, G. Xu, M. Li et al., "Steady-state motion visual evoked potential (SSMVEP) based on equal luminance colored enhancement," *PLoS One*, vol. 12, Article ID e0169642, 2017.

Research Article

Impact of Speller Size on a Visual P300 Brain-Computer Interface (BCI) System under Two Conditions of Constraint for Eye Movement

R. Ron-Angevin¹, L. Garcia², Á. Fernández-Rodríguez¹, J. Saracco^{1,4}, J. M. André², and V. Lespinet-Najib²

¹Departamento de Tecnología Electrónica, ETSI Telecomunicación, Universidad de Málaga, Málaga, Spain

²IMS UMR 5218, CIH, ENSC, Bordeaux INP, Bordeaux, France

³Univ. Bordeaux, CNRS, Bordeaux INP, IMB, UMR 5251, F-33400 Talence, France

⁴INRIA, IMB, UMR 5251, F-33400 Talence, France

Correspondence should be addressed to R. Ron-Angevin; rra@dte.uma.es

Received 5 April 2019; Accepted 28 May 2019; Published 1 July 2019

Guest Editor: Minkyu Ahn

Copyright © 2019 R. Ron-Angevin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The vast majority of P300-based brain-computer interface (BCI) systems are based on the well-known P300 speller presented by Farwell and Donchin for communication purposes and an alternative to people with neuromuscular disabilities, such as impaired eye movement. The purpose of the present work is to study the effect of speller size on P300-based BCI usability, measured in terms of effectiveness, efficiency, and satisfaction under overt and covert attention conditions. To this end, twelve participants used three speller sizes under both attentional conditions to spell 12 symbols. The results indicated that the speller size had, in both attentional conditions, a significant influence on performance. In both conditions (covert and overt), the best performances were obtained with the small and medium speller sizes, both being the most effective. The speller size did not significantly affect workload on the three speller sizes. In contrast, covert attention condition produced very high workload due to the increased resources expended to complete the task. Regarding users' preferences, significant differences were obtained between speller sizes. The small speller size was considered as the most complex, the most stressful, the less comfortable, and the most tiring. The medium speller size was always considered in the medium rank, which is the speller size that was evaluated less frequently and, for each dimension, the worst one. In this sense, the medium and the large speller sizes were considered as the most satisfactory. Finally, the medium speller size was the one to which the three standard dimensions were collected: high effectiveness, high efficiency, and high satisfaction. This work demonstrates that the speller size is an important parameter to consider in improving the usability of P300 BCI for communication purposes. The obtained results showed that using the proposed medium speller size, performance and satisfaction could be improved.

1. Introduction

Evoked brain signals using different stimulus modalities can be employed to translate human intentions into external actions (motor outputs) through brain-computer interface (BCI) systems [1, 2]. These systems enable a nonmuscular channel of communication between the user and his environment, which could be especially useful for people diagnosed with severe motor disorders such as amyotrophic

lateral sclerosis (ALS). So these systems are the only option that some patients have to communicate and get autonomy.

BCI systems based on electroencephalographic (EEG) signal recording are the most widely studied. In spite of existing different EEG-based BCI systems, a visual P300-based BCI speller is the main interface used for communication and control purposes and represents a reliable real option to supply the needs of people with neurological dysfunctionality [3, 4]. The P300 signal is a type of event-

related potential (ERP) which is mostly recorded over the central and parietal regions. Specifically, it is a positive deflection of brain activity which occurs about 300 ms after an odd stimulus presentation.

There are several types of P300-based spellers, such as auditory, tactile, or visual. The main advantage of auditory and tactile modalities is that they do not depend on the ocular capacity (e.g., [5]). However, if the user retains some residual ocular mobility, the visual interfaces will show a better performance (see Rezeika et al. [6] for a review of P300-based spellers). The vast majority of P300-based BCI spellers studied and developed are based on the one proposed by Farwell and Donchin in 1988 [7]. Farwell and Donchin's speller consisted of a 6×6 matrix of characters. Its rows and columns are flashed (i.e., stimulated) pseudorandomly, one by one, while the user pays attention on his target element from the matrix. The stimulation of the user's target element represents the "rare event" of the oddball paradigm and elicits the P300. After a specified number of row and column stimulations, the computer identifies the matrix element to which the user is attending as the intersection of the row and column that elicits the largest P300, and this symbol is shown on the screen.

Given a number of studies which were carried out on both healthy subjects and patients affected by some motor disability, the effectiveness of the mentioned P300-based spellers is proved [4]. Overall, these studies agree on that the P300 speller gives an effective communication channel to those patients who have almost—or completely—lost the possibility to write or speak. However, as it is proposed in [8], the BCI spellers' usability is still needed to get better. The current definition of usability given by the International Organization for Standardization (ISO 9241-11) involves three measures: (i) effectiveness (i.e., accuracy and completeness of the system with which users achieve set goals), (ii) efficiency (i.e., resources expended to complete goals), and (iii) satisfaction (i.e., users' attitude to complete a given task) [9–11]. Froekjaer et al. came to the conclusion that these measures should be considered as independent usability aspects [12]. The efficiency and satisfaction could be basically measured through different subjective aspects: mental workload, fatigue, motivation, comfort, pleasure to use, and so on [13, 14].

The P300 signal amplitude and latency can be influenced by many factors, for instance, the mental fatigue after a long use [15, 16], the level of kept attention to a desired symbol [17], the user's motivation [15, 18], or the user's frustration (see [19] for a review). In this regard, the attention of researchers is increasingly focused on the effect on the user performance given several temporal and spatial aspects of the speller interfaces [20].

Although a vast majority of papers have focused on signal-processing algorithms in order to improve the performance of the P300 BCI system, there are several researches that have studied parameters which might have an influence on the user performance. Some of these parameters are the stimulus timing features [21–23], the effect of luminosity contrast [24], and the influence of interface colour contrast [25]. Regarding the effect of matrix

configuration, the research is limited. Some studies have demonstrated how the user performance is affected by the matrix size. Specifically, Allison and Pineda [26] made a study where three matrix sizes (4×4 , 8×8 , and 12×12) were compared. The results indicated that larger P300 amplitudes were evoked by larger matrices and the user performance or preference was not significantly affected by the matrix size. On the contrary, a study where two different matrix sizes were compared (3×3 and 6×6) showed that the 3×3 matrix achieved higher accuracy, whereas the P300 amplitude was higher for the 6×6 matrix condition [22]. In both studies, symbol size was the same in different matrices, and thus, the larger matrices were presented larger on the monitor by only increasing the distance between symbols. As a result of these studies, some P300 BCI spellers used a reduced matrix to increase the writing speed [27]. Salvaris and Sepulveda studied the effects on classification of changes in the dimensions of the symbols, the distance between the symbols, and the background colours [28]. In this study, only two different values of each parameter were compared: small symbol size versus large symbol size, small intersymbol distance versus large intersymbol distance, and black background versus white background. The worst performance was obtained with the small symbol size.

There have been no studies related to the effect of speller size, apart from the mentioned studies about matrix size and symbol size. In [29], three different screen sizes were tested: a computer monitor, a global positioning system (GPS) screen, and a mobile phone screen. Nevertheless, no symbol size information was provided. According to the information provided about the screen resolution and the distance from the participants to each screen, the visual field for the computer monitor was 6.4° , for the GPS screen 3.7° , and for the mobile phone screen 3.56° , where two smallest screens had almost the same visual fields. However, since this study did not provide information regarding the speller size, it is unclear how their results are related to speller size. Actually, this study's main purpose was to assess BCI performance when these three specific screens were used but not to study the effect of screen size. To evaluate the effect of speller size, different visual fields should be proposed in terms of both symbol size and symbol distance, as they are crucial to confirm the proposals of different speller sizes.

Most of the P300 visual spellers are used in the overt attention mode, that is, allowing the subjects to fixate the target with their eyes. However, several studies have also proposed P300 visual speller usage in the covert attention mode [30–33] as an alternative communication aid for completely locked people. In this mode, subjects have to fixate the centre of the screen while paying attention to the target using visual periphery. Effectively, unfortunately, some of the potential users of a BCI speller, that is, ALS patients, could have impaired visual function, not allowing to gaze different targets. In our study, the covert attention mode has been employed to replicate the lack of ocular mobility that is suffered for some patients with severe motor disorders. Some of these proposed studies [30, 33] have clearly demonstrated that the performance of the classical speller in the covert attention condition considerably

decreases compared to the overt attention condition. Degradation of spatial acuity in the peripheral vision is one of the effects that contribute to reduce this performance [31]. Human detail vision is limited to the fovea (centre), where visual acuity is 100%. As the distance from the fovea increases (eccentricity), the visual acuity drops rapidly to approximately 60% at 1° eccentricity, 50% at 2°, 30% at 7°, and 20% at 10° [34]. This degradation of visual acuity as a function of eccentricity should be taken into account in the design of BCI P300 speller in the covert attention condition. In a classical speller, one way to prevent the detrimental effects of declining visual acuity is to reduce the speller size. However, the symbol sizes would also reduce and be more difficult to distinguish. In this sense, it would also be interesting to study the effect of different speller sizes in the covert attention condition.

A published study by Brunner et al. [33] investigated the extent to which the performance of a classical P300 BCI speller depends on eye gaze. To this end, they evaluated the offline performance of 17 healthy subjects under overt and covert attention conditions. The obtained results showed a significant reduction in the classification accuracy in the covert attention condition compared to the overt attention condition. As it is mentioned in this paper, further studies are necessary to evaluate the effect of online feedback (online performance). On the contrary, it would be interesting to evaluate the workload required in both conditions.

The goal of this study was to explore the extent to which the performance of a classical P300 speller depends on speller size. The purpose was to better understand how easily a user can carry out the speller task comfortably and efficiently by analysing the same speller system type while using different BCI speller sizes. In this sense, we evaluated the usability of different speller sizes in terms of effectiveness, efficiency, and satisfaction [10, 35]. Effectively, the obtained performance is not a sufficient criterion to determine whether a user would want to use an interface. To this end, it is necessary to take into account these three parameters (effectiveness, efficiency, and satisfaction), making it possible not only to predict the user's intention [36] and the degree of acceptance of an interface [37, 38] but also to offer a better user experience [39].

2. Methods

2.1. Participants. Twelve French university students (seven males and five females; age range 19–25 years (20.6 ± 0.9 years)) participated in the present study (S1–S12), which consisted of six sessions, one for each speller size (i.e., small, medium, and large) and for each attentional condition (i.e., overt and covert). According to self-reports, all participants had no history of neurological or psychiatric illness and had normal or corrected to normal vision. Every participant gave informed consent through a protocol reviewed by the ENSC-IMS cognitive team. None of them had previous experience with BCI systems. The study was approved by the Ethics Committee of the University of Malaga and met the ethical standards of the Helsinki Declaration.

2.2. EEG Data and Processing. EEG recording and amplifying was through a 16-channel biosignal amplifier (g.BSamp, Guger Technologies) of gold electrodes. According to the 10/20 international system, the electrodes were placed at positions Fz, Cz, Pz, Oz, P3, P4, PO7, and PO8. The channels' reference was the right earlobe, and FPz was used as ground. Through the amplifier settings, the signal was bandpass filtered at 0.5 and 100 Hz, the notch filter (at 50 Hz) was on, and the sensitivity was $500 \mu\text{V}$. Next, the EEG data were digitized at a rate of 256 Hz by a 12-bit resolution NI-USB-6210 data acquisition card (National Instruments). Every aspect of EEG data recording and processing was controlled by the BCI2000 software [20].

2.3. BCI Speller. The BCI speller used was the classical Farwell and Donchin [7] speller, which consists of a 6×6 matrix of symbols (36 alphanumeric letters and numbers) arranged in rows and columns. The temporal parameter values for all the spellers were based on those used by Donchin et al. [40]. Specifically, each row and each column were intensified (i.e., flashed) pseudorandomly 10 times, and thus, each character was intensified 20 times. Both the stimulus presentation duration (i.e., the duration of each flash) and the interstimulus interval (ISI) pause between stimulus presentations were 125 ms. A pause of 6 sec was used following each sequence of flashes (i.e., pause between each character selection). This pause duration was selected to give the subject time to look for the new target character and gaze it. In the covert attention condition, subjects were also allowed to gaze the new target character during this pause. Considering these temporary parameter values, each symbol needed a time of 36 s to be selected (as it is 10 times the sum of the flash duration—125 ms—of six rows and six columns with an ISI of 125 ms, plus the 6 s after the sequences of flashes).

2.4. Speller Size. Three different speller sizes were proposed. The screen used to present the spellers was 17" TFT with a refresh rate of 60 Hz and a resolution of $1440 \times 900 \text{ px}^2$. Each speller consisted of a 6×6 matrix of 36 characters which was centred on the screen.

- (i) The speller size used in [30] was chosen as the largest size of the present study because it is frequently applied by other researchers (e.g., [33]). The matrix subtended $\pm 6.98^\circ$ of the visual field both horizontally and vertically. The size of each character was $1.12^\circ \text{ W} \times 1.12^\circ \text{ H}$, with the horizontal separation between columns being 1.46° W and the vertical separation between rows being 1.46° H .
- (ii) The smallest speller size was chosen according to what was reported by Salvaris and Sepulveda [28] as the smallest symbol size which could be used without loss of spelling performance. As the subjects were seated 1 m from the screen in their experiments and the smallest symbol size was $0.7 \text{ cm W} \times 0.8 \text{ cm H}$ in [28], in terms of visual field, the symbol size was equivalent to $0.4^\circ \text{ W} \times 0.45^\circ \text{ H}$. To

maintain the same characteristics regarding the size of different spellers, the symbol size used in our experiment should be the same in height and width. Finally, each character's size on the smallest speller was $0.4^\circ W \times 0.4^\circ H$. This small symbol size ($0.4^\circ W \times 0.4^\circ H$) represents a reduction of 35.89% of the largest symbol size chosen ($1.12^\circ W \times 1.12^\circ H$). In order to preserve the same proportions between speller sizes, the smallest speller subtended a visual field of $\pm 2.51^\circ$ both horizontally and vertically, and the horizontal and vertical separation between columns and rows, respectively, was 0.52° (i.e., 35.89% less compared to the largest size). With this size and according to the study in [33], in the cover attention condition, the visual acuity would drop to approximately 50% for targets located further away from the centre.

- (iii) The medium size was selected calculating the middle value between the large and small speller sizes. Then, the matrix subtended $\pm 4.75^\circ$ of the visual field both horizontally and vertically, the intermediate symbol size was $0.75^\circ W \times 0.75^\circ H$, and the horizontal and the vertical separation between columns and rows, respectively, was 1° .

Taking into account that subjects were situated at a distance of about 60 cm from the screen, the measures of each speller are presented in Figure 1 and Table 1.

2.5. Experimental Design. The three different speller sizes were tested by every participant following a within-subject design. Thus, the experiment consisted of three sessions, wherein each of them tested one speller. Sessions were carried out on different days, and the time interval between each session was between three and five days (both included). The order in which the spellers were assessed was counterbalanced over participants to control for the potential effects of experience. In order to measure eye gaze, an eye tracker (Tobii X1, Tobii Technology) was mounted under the screen. All sessions took place inside an isolated experimental room.

2.6. Task and Procedure. Prior to the experiment, instructions were given in written and verbal forms. All participants were seated at a distance of approximately 60 cm from the screen, which is the optimal operational range for the eye tracker (60 ± 10 cm). Each subject used the three speller sizes (small, medium, and large) under two attentional conditions (overt and covert attention). In the overt attention condition, the participant was asked to gaze at the target (i.e., the desired symbol that should be written). In the covert attention condition, the subject had to gaze only at a yellow dot located in the centre of the screen while counting the intensification of the desired character. Likewise, each session consisted of two tasks for each condition: a calibration task and an online task. Before the calibration task of the BCI speller, the eye tracker was calibrated.

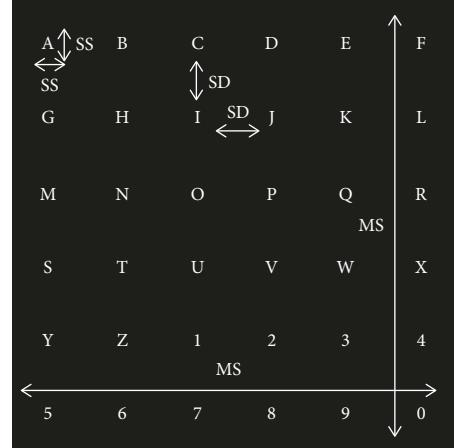


FIGURE 1: Size metrics determined for the visual protocol.

TABLE 1: Different parameters used in different speller sizes.

Parameters		Size		
		Small	Medium	Large
Matrix size (MS)	cm	5.27	9.98	14.69
	inch	2.26	3.92	5.78
Symbol size (SS)	cm	0.42	0.79	1.17
	inch	0.16	0.31	0.46
Symbol distance (SD)	cm	0.55	1.04	1.53
	inch	0.22	0.41	0.61

Before the calibration task started, the participants were informed that he/she would see 10 random short intensifications (i.e., flashes) of rows and columns. Each time all the rows and columns were flashed 10 times, a sequence was completed. As every row and column flashed 10 times (i.e., 10 sequences), each character was flashed 20 times. The mental task for participants to type a letter or a number was to mentally count every time that his/her desired symbol was flashed before the 6 sec pause. During the calibration task, the participants did not receive any feedback and were asked to focus consecutively on 16 characters to spell three French words and a number, all of them with four characters (four runs). The spelled words were “LUNE,” “FEUX,” and “KILO” and the number was “2015.” While they were doing the task, their EEG data were recorded in order to analyse them afterwards. The calibration task took approximately 10 minutes. At the end of the calibration task, a stepwise linear discriminant analysis of the last three runs was performed to get the weights of the P300 classifier.

Once the matrix of weights of the classifier was loaded to the system, the online task started. For the online task, the words and the number asked to spell were the following: “CHAT,” “PURE,” and “1935,” one after the other without spaces. They were instructed to continue without correcting the mistake in case a wrong letter was chosen by the classifier. Before each word’s (or number’s) set of flashes started, it was presented for 1 second on the screen (between seconds 2 and 3 of the 6 s pause). This time, the characters spelled appeared in a typing bar placed below the matrix of characters (Figure 2). The time required to choose a character



FIGURE 2: Subject during the experimental test. A Tobii eye tracker is mounted under the screen.

was 36 s. Thus, the time taken to write the four characters of each word (or number) was 2 min and 24 s. The experimental design timing is shown in Figure 3.

The choice of different characters during the online task was established so as to ensure that each target was located at different distances from the centre and at different directions. Thus, there are three layers in the 6×6 matrix from the edges to the middle (Figure 4). The three sequences of characters (two words and one number) to be spelt were selected so that each layer was covered by a similar percentage of characters (6/20 in layer 1, 4/12 in layer 2, and 2/4 in layer 3). Besides, in each layer, different characters were established so as to ensure that the user had to make, in the overt attention condition, gaze movements in different directions. The distribution of these characters is represented in bold in Figure 4.

After the online tasks, participants were asked to complete a visual analogue scale (VAS) of fatigue and the NASA-TLX test [41] and to answer a short questionnaire related to the speller tested in that session. This last questionnaire included three statements related to some features of speller size: (i) *statement 1*, difficulty perceiving different characters; (ii) *statement 2*, difficulty perceiving characters away from the centre; and (iii) *statement 3*, difficulty distinguishing different rows and columns. The participant expressed his/her level of agreement to each statement given a 10-point Likert scale (1 = very easy and 10 = very difficult).

The NASA-TLX test is a multidimensional rating questionnaire with six subscales (*mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration*) which are scored between 0 and 100 and where higher values are related to higher levels of workload. This test consisted of two phases: In the first phase, participants give a rating to assign a magnitude to each subscale. In the second phase, 15 pairs of subscales were obtained after combining the six subscales, so the subjects could compare each pair to indicate and identify the subscale which affected their workload more. The overall workload was computed given a weighting average technique which considers the particular contribution of every subscale to the *total workload*. The overall workload values indicated the speller size requiring the most mental workload, while the weighted subscale scores identified the workload factors that made the greater contributions to each speller size. The highest possible score for the overall workload is 100, while the highest score for the weighted subscales is 33.3. The endpoints for each subscale are “very low/very high” except for the performance subscale, which has “perfect/failure” endpoints.

At the end of the third session, every subject was asked to compare the three speller sizes regarding his/her preferences. A comparative questionnaire adapted from the System Usability Scale (SUS) [42] allowed us to evaluate six dimensions: *favourite*, *complex*, *comfortable*, *stressful*, *controllable*, and *tiring*. For each dimension, the three speller sizes were ranked between them. Three ranks were proposed for each dimension: *rank 1*, the least; *rank 2*, intermediate; and *rank 3*, the most.

2.7. Parameter and Statistical Analysis. As mentioned in Introduction, the main objective of this study was to evaluate the usability of different speller sizes under two different conditions: covert and overt attention. The employed usability approach includes three dimensions: effectiveness, efficiency, and satisfaction. Effectiveness is related to the accuracy with which a user can complete tasks. In order to study the effectiveness, different results are required: (i) classification *accuracy* and *number of flashes* required to select a symbol during the calibration task; (ii) *error performance* in the online task writing all the words; and (iii) *amplitude* of the P300 signals during the online task. Efficiency is related to the resources expended to complete a task, i.e., user’s effort and time required. In order to study the efficiency, the following metrics were provided: (i) the subjective workload assessed using NASA-TLX; (ii) the VAS of fatigue; and (iii) subjective questionnaires related to some features of speller size. Satisfaction is related to the users’ attitude, i.e., the perceived comfort and acceptability while using the system. Results related to preference and subjective feelings regarding different speller sizes were analysed through the comparative questionnaire adapted from the SUS. Table 2 summarises different objective and subjective metrics used to evaluate the three usability dimensions.

The analysis of variance (ANOVA) was used to analyse different evaluation metrics for effectiveness and efficiency dimensions. Additionally, multiple *R-squared* values were calculated for evaluating how well the model fits the data. Only results associated with a model with an *R-squared* value above 0.25 are reported here. A 5% threshold was considered significant for different Fisher’s tests. Regarding the satisfaction dimension, in order to class the user’s preference for each parameter (*favourite*, *complex*, *comfortable*, *stressful*, *controllable*, and *tiring*), Fisher’s exact test has been used for each condition. The main reason to select Fisher’s exact test instead of the χ^2 test is the low sample size.

2.8. Gaze Direction Control. In order to validate the experiments carried out under the covert attention condition, it was necessary to verify that the subjects fixated his/her gaze on a yellow dot placed at the centre of the screen. To this end, the gaze data were analysed offline during the evaluation phase (i.e., the online task). The instructions demanded that subjects had to fixate the centre of the screen during the flash period for selecting a letter (30 s) and could gaze the target to locate it during the 6 s pause between flash periods. Due to participants having to spell 12 characters, the trial time period for one experiment was 432 s. Of this period,

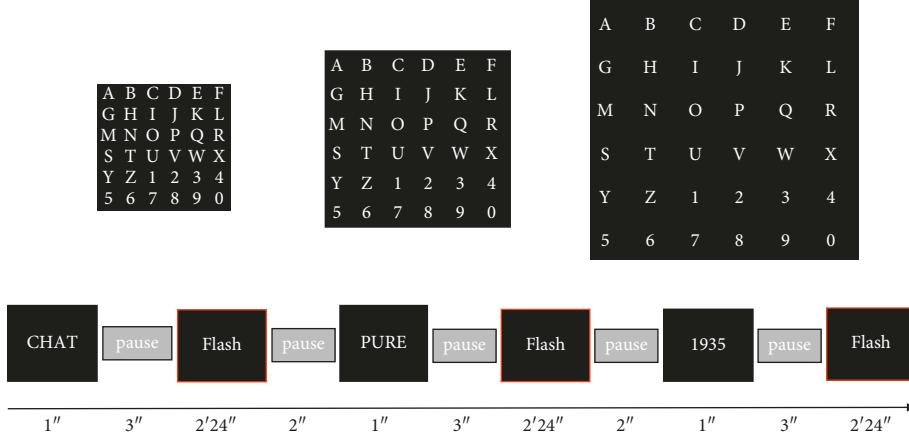


FIGURE 3: Temporal sequence employed on the online task for the three different speller sizes.

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z	1	2	3	4
5	6	7	8	9	0

FIGURE 4: Distribution of the three layers according to their distance from the centre.

TABLE 2: Different evaluation metrics used to analyse the three usability dimensions.

Effectiveness	(i) Classification accuracy and number of flashes required to select a symbol during calibration
	(ii) Error performance according to the online task
	(iii) P300 waveforms (amplitude and latency)
Efficiency	(i) NASA-TLX
	(ii) VAS fatigue
	(iii) Perception subjective questionnaires
Satisfaction	(i) A comparative questionnaire adapted from the SUS for six dimensions (favourite, complex, comfortable, stressful, controllable, and tiring)

subjects had to gaze the centre of the screen for 360 s (30 s × 12 characters), enabling them to use the 6 s pause between flash periods to shift attention and gaze at the new character, that is, during 72 s (6 s × 12 period flashes) of the total trial time. In order to check if subjects gazed the centre of the screen, we obtained, for each subject and each speller size (i.e., for each trial time), the total time that the gaze data, in samples of 0.5 s, was at a distance of $\pm 2^\circ$ from the centre. Taking into account that subjects had to gaze the centre of the screen for 360 s, the percentage of time above or below this period was calculated for each trial. In case of gazing the centre of the screen for all the trial time (i.e., for 432 s), the percentage should be 20%. In case of gazing the centre of the screen for less than 360 s, the percentage should be negative.

Finally, the percentage of time ranged from 8.16 to 17.78% (due to technical difficulties, gaze data were not available for two participants); that is, all the subjects gazed the centre of the screen for more than 360 s per trial. The average \pm standard deviation (SD) percentages of all subjects were 12.68 ± 3.32 , 13.19 ± 2.43 , and $13.41 \pm 2.53\%$ for small, medium, and large speller sizes, respectively. These results show that the subjects followed the instructions; that is, they looked at the yellow dot under covert attention condition.

3. Results

3.1. Effectiveness Dimension

3.1.1. Classification Accuracy during Calibration Task. Figure 5 shows the averages of the classification *accuracy* achieved by the participants, under both attentional conditions and the three tested speller sizes, which is related to the number of times that a row and a column in the matrix were intensified (i.e., the number of flashes, with one flash being the intensification of one row and one column). Regardless of the attentional condition or the speller size, the classification *accuracy* increases with the number of flashes.

In the overt attention condition, all subjects, except subject S4 who had 94% *accuracy* for small and medium speller sizes, obtained 100% of classification accuracy at the end of the ten flashes. In the covert attention condition, classification *accuracy* did not reach 100% for any subject (except for subjects S2, S3, and S9 using the medium speller size). In this condition, after 10 sequences (i.e., 20 flashes), the classification *accuracy* ranged from 63 to 94% for the small speller size, 63 to 100% for the intermediate speller size, and 50 to 94% for the large speller size, with averages (\pm standard error, SE) of 74.58 ± 13.09 , 80.27 ± 12.95 , and $70.33 \pm 12.87\%$ for small, medium, and large speller sizes, respectively.

For comparative results between speller sizes and attentional conditions, we calculated the average of the *highest classification accuracy* obtained during the calibration task and the *minimum number of flashes* required to obtain these accuracies. Results are depicted in Figures 6(a) and 6(b), respectively.

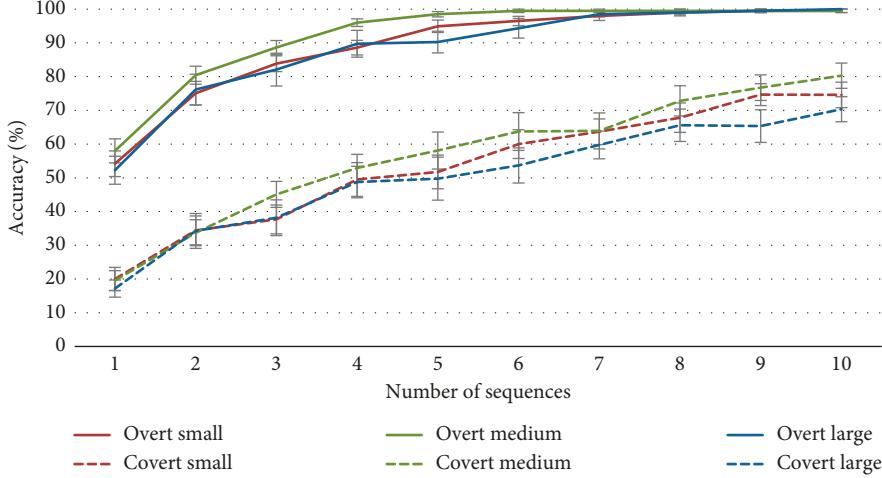


FIGURE 5: Average (\pm SE) of classification accuracy of the three speller sizes over the number of flash sequences for overt (solid line) and covert (dashed line) attention conditions.

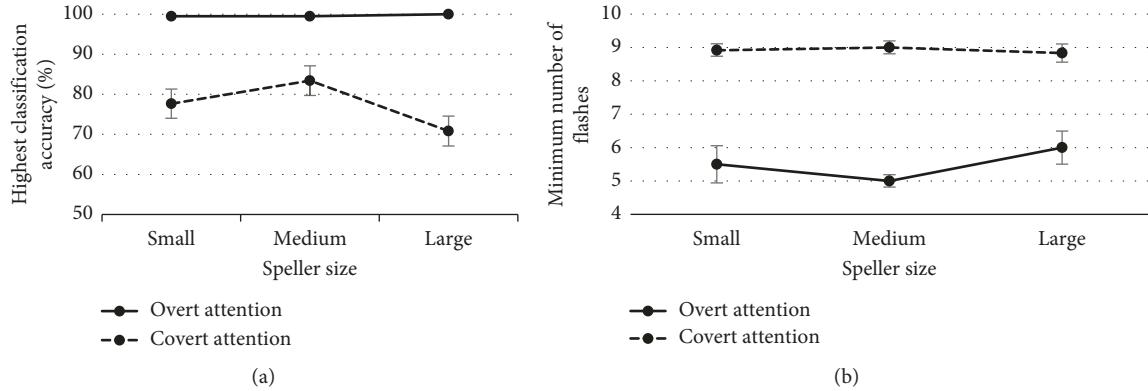


FIGURE 6: Mean (\pm SE) of the highest classification accuracy achieved by participant (a) and the minimum number of flashes required to reach that accuracy (b) during calibration task for each condition and each speller size.

According to the *highest classification accuracy*, a two-way ANOVA (3×2) with factors *Speller Size* and *Attention* showed a significant interaction between these factors [$F(2, 65) = 3.07; p = 0.05$]. In the overt attention condition, analysis showed no significant differences between speller sizes; however, in the covert attention condition, the mean *accuracy* for medium speller size ($83.4 \pm 3.7\%$) was the highest, offering significant differences versus the large speller size ($70.8 \pm 3.7\%$) ($p = 0.01$). The ANOVA showed a significant main effect of factor *Attention* [$F(1, 65) = 101; p < 0.001$]. Independent of the speller size, the accuracy was significantly lower for the covert attention than for the overt attention condition. These results showed that calibration did not represent a problem under the overt attention condition, independent of the speller size. However, in the covert attention condition, *accuracy* did not reach 100% for any subject, which will affect the online performance. In spite of that, it is important to mention that medium speller size may have an influence on *highest classification accuracy*. Thus, the best classification accuracies were offered by the medium speller size.

According to the *minimum number of flashes* required to obtain the *highest classification accuracy*, the two-way ANOVA with factors *Speller Size* and *Attention* did not show any significant interaction between these factors. Thus, the differences in the *minimum number of flashes* required for each condition of *Attention* (i.e., overt and covert) between the factor *Size* (i.e., small, medium, and large conditions) were similar. Only the *Attention* factor was significant [$F(1, 69) = 116; p < 0.001$]. Therefore, independent of the speller size, the *minimum number of flashes* required was higher for covert attention than for overt attention. In the overt attention condition, the averages obtained for *minimum number of flashes* were 5.8 ± 0.5 , 4.9 ± 0.2 , and 5.8 ± 0.6 for the small, medium, and large sizes, respectively. However, most of participants (S1–S3, S6, S8–S10, and S12) required a lower *minimum number of flashes* to obtain 100% of *accuracy* with the medium speller size. In the covert attention condition, the averages for *minimum number of flashes* were 8.9 ± 0.3 , 9.0 ± 0.3 , and 8.8 ± 0.3 for the small, medium, and large sizes, respectively. However, as mentioned before, the accuracies did not reach

100%. It is important to notice that a reduced *minimum number of flashes* required to select a symbol would lessen the time to select it, increasing the efficiency of the system.

3.1.2. Classification Accuracy during Online Task. Table 3 shows the overall performance results on the online task for each subject and each speller size. The “Mean” column represents the average \pm standard error in *error performance* between speller sizes. Additionally, Figure 7 has been presented to visually observe the differences.

The *error performance*, calculated as the inverse of *accuracy* (i.e. $1 - \text{accuracy}$) was analysed using a two-way ANOVA (3×2), with factors *Attention* and *Speller Size*. The obtained results did not show any significant interaction between these factors. However, a significant effect was obtained for *Attention* [$F(1, 67) = 83.5; p < 0.001$] and for *Speller Size* [$F(2, 67) = 3.56; p = 0.03$]. Regarding the factor *Attention*, analysis revealed higher *error performances* for the covert attention than for the overt attention, independent of the speller sizes. Regarding the factor *Speller Size*, for both conditions, the *error performances* were significantly higher with the large speller size than for both the medium ($p = 0.02$) and small speller sizes ($p = 0.02$). Additionally, *error performances* were not different between the small and medium speller sizes. It is important to notice that, under the overt attention, the total number of subjects with 0% error was 9, 8, and 3 for the small, medium, and large sizes, respectively. However, the total number of subjects with more than 10% error was only 1 for the small and medium sizes and 5 for the large size. Otherwise, under the covert attention, only 2 participants achieved a 0% error performance (below 10%) and both using the same size, the large one.

To study the effect of the *Speller Size* and *Attention* factors on different layers of a speller (Table 4), three two-way ANOVA (3×2) have been carried out, one for each layer. First, the *Speller Size* factor has shown significant differences in *layer 1* [$F(2, 67) = 4.4; p = 0.01$] and *layer 2* [$F(2, 67) = 3.4; p = 0.03$]. On the one hand, for *layer 1*, we observed that the large size has obtained a worse performance compared to the medium size ($p = 0.01$) and the small size ($p = 0.01$). On the other hand, for *layer 2*, the same significant differences were obtained, with the large speller offering the highest percentage of *error performance* compared to the medium ($p = 0.01$) and small ($p = 0.03$) sizes. Second, in reference to *Attention* factor, the condition of covert attention has shown a significantly lower performance than the overt attention condition in all layers (*layer 1* [$F(1, 67) = 81.7, p < 0.001$], *layer 2* [$F(1, 67) = 67.9, p < 0.001$], and *layer 3* [$F(1, 69) = 13.4, p < 0.001$]). Finally, no interaction effect was found between the *Attention* factor and *Speller Size* in any layer.

3.1.3. P300 Waveform Analysis. Figure 8 shows the overall grand average event-related potential (ERP) waveforms obtained as a response to target and nontarget stimuli for the eight electrodes and as a function of the speller size for overt and covert attention conditions. For the three speller sizes

and attentional conditions, different responses have positive peaks between 300 and 500 ms.

In order to investigate the effects of the overt versus covert attention and the speller size over the *latency* and *amplitude* of the P300, two three-way ANOVA ($3 \times 2 \times 8$), one for each dependent variable, were performed using the following factors: *Speller Size* (small, medium, and large), *Attention* (overt and covert), and *Channel* (Fz, Cz, Pz, Oz, P3, P4, PO7, and PO8). Regarding the *latency*, no main or interaction effects between factors were found. On the contrary, in reference to the *amplitude* of the P300, only an interaction effect of *Attention* \times *Channel* [$F(7, 77) = 3.456; p = 0.023$] was found. As such, paired *t*-test analysis for each channel comparing the *amplitude* between both overt and covert attentions (*Attention* factor) was carried out. These analyses have shown that significant differences have been restricted only to channels P4 [$t(11) = 2.444; p = 0.033$] and PO8 [$t(11) = 3.121; p = 0.01$], where the overt attention condition presented higher values (Figure 9).

3.2. Efficiency Dimension

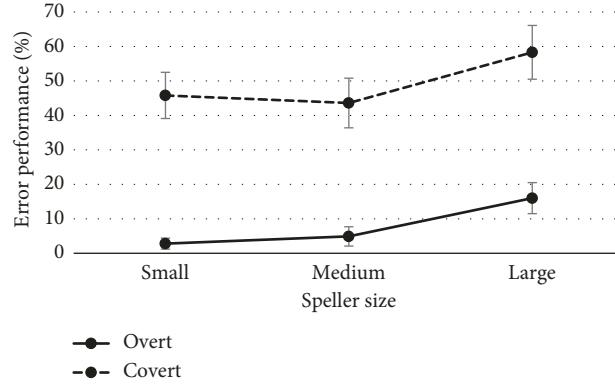
3.2.1. Workload and VAS Fatigue. Besides the NASA-TLX test, to evaluate the workload provoked by the use of the speller, overall fatigue was rated on a VAS ranging from 0 to 10 [43]. Table 5 shows the contributions of *VAS fatigue*, *total workload* (NASA-TLX global score ranged from 0 to 100), and dimensions to assess the subjective workload (*mental demand*, *physical demand*, *temporal demand*, *effort*, *performance*, and *frustration*, ranging from 0 to 33.3) for each speller size. The obtained values correspond to the average score among participants.

Another two-way ANOVA (3×2) was carried out in order to study the effect of *Speller Size* and *Attention* in the workload and fatigue. There was no main effect of the *Speller Size* or interaction effect between *Speller Size* and *Attention* in any variable. However, the workload and fatigue were generally higher for participants in the covert attention condition compared to the overt attention condition. Specifically, the covert attention condition offered significantly higher values for the following dimensions: *VAS fatigue* [$F(1, 69) = 6.08; p = 0.01$], *total workload* (NASA-TLX) [$F(1, 69) = 26.27; p < 0.001$], *mental demand* [$F(1, 69) = 6.98; p = 0.01$], *temporal demand* [$F(1, 69) = 6.29; p = 0.01$], *effort* [$F(1, 69) = 9.71; p = 0.002$], and *performance* [$F(1, 69) = 6.03; p = 0.01$]. Interestingly, *physical demand* and *frustration* dimensions did not seem to be influenced by the *Attention* factor.

3.2.2. Perception Subjective Questionnaires. The answers given by the participants at the end of each session in the usability questionnaire related to some features of the speller size are shown in Table 6. In this table, only the median of the answers in the sample of participants is given (ranging from 1 to 10: 1 = very easy and 10 = very difficult). Three two-way ANOVA were carried out to study the effects of *Speller Size* and *Attention* factors in different statements: *statement 1*, the difficulty perceiving different characters; *statement 2*, the

TABLE 3: Error performance (%), mean \pm SE) results of the online task for each participant.

Participant	Overt attention				Covert attention			
	Small	Medium	Large	Mean	Small	Medium	Large	Mean
S1	16.7	0	41.7	19.4 \pm 12.1	66.7	33.3	66.7	55.6 \pm 11.1
S2	8.3	0	0	2.8 \pm 2.7	83.3	41.7	50	58.3 \pm 12.7
S3	0	0	41.7	13.9 \pm 13.9	16.7	16.7	8.3	13.9 \pm 2.8
S4	0	8.3	8.3	5.6 \pm 2.8	16.7	66.7	41.7	41.7 \pm 14.4
S5	0	0	8.3	2.78 \pm 2.8	50	25	33.3	36.1 \pm 7.4
S6	0	33.3	0	11.1 \pm 11.1	41.7	75	100	72.2 \pm 16.9
S7	0	0	16.7	5.6 \pm 5.6	75	58.3	66.7	66.7 \pm 4.8
S8	0	0	0	0 \pm 0	25	25	33.3	27.8 \pm 2.8
S9	0	8.3	25	11.1 \pm 7.4	25	25	100	50 \pm 25
S10	8.3	0	33.3	13.9 \pm 10	33.3	75	58.3	55.5 \pm 12.1
S11	0	8.3	8.3	5.6 \pm 2.8	50	75	75	66.7 \pm 8.3
S12	0	0	8.3	2.8 \pm 2.8	66.7	8.3	66.7	47.2 \pm 19.5
Mean	2.8 \pm 1.6	4.9 \pm 2.8	16.0 \pm 4.5	7.87	45.83 \pm 6.7	43.75 \pm 7.2	58.33 \pm 7.8	49.3

FIGURE 7: Average (\pm SE) error performance (%) by speller size in overt and covert attention modes during the online task.TABLE 4: Error performance (%), mean \pm SE) results according to different layers of the matrix.

Layer	Overt attention				Covert attention			
	Small	Medium	Large	Mean	Small	Medium	Large	Mean
Layer 1 (red)	2.8 \pm 1.9	6.9 \pm 4.3	16.7 \pm 4.6	8.8 \pm 4.1	50 \pm 7.1	44.5 \pm 8.5	66.7 \pm 7.7	53.7 \pm 6.7
Layer 2 (green)	2.1 \pm 2.1	2.1 \pm 2.1	12.5 \pm 5.8	5.6 \pm 3.5	49.3 \pm 9.1	45.8 \pm 10.6	70.8 \pm 9.7	55.3 \pm 7.8
Layer 3 (yellow)	4.2 \pm 4.2	4.2 \pm 4.2	20.8 \pm 9.7	9.7 \pm 5.5	29.2 \pm 9.7	41.7 \pm 8.3	33.3 \pm 11.2	34.7 \pm 3.7

difficulty perceiving characters away from the centre; and *statement 3*, the difficulty distinguishing different rows and columns.

The *Speller Size* factor showed significant differences in *statement 1* [$F(2, 22) = 4.322$; $p = 0.026$] and *statement 3* [$F(2, 22) = 5.204$; $p = 0.014$]. Thus, it could be affirmed that the *Speller Size* affects the difficulty perceiving different characters and distinguishing different rows and columns. Specifically, in *statement 1*, the medium speller size offered better values (i.e., lower scores in difficulty) than small ($p = 0.025$) and large ($p = 0.019$) speller sizes. On the contrary, for *statement 3*, the small speller size offered worse significant results than medium ($p = 0.023$) and large sizes ($p = 0.048$). Regarding the *Attention* factor, we found that the covert attention condition showed significantly higher scores, i.e., worse rating, for all statements (*statement 1* [$F(1, 44) = 44.295$; $p < 0.001$], *statement 2* [$F(1, 11) = 88.393$; $p < 0.001$], and *statement 3* [$F(1, 11) = 44.044$; $p < 0.001$]). In

addition, an interaction effect was observed in *statement 2* between *Speller Size* and *Attention*. In order to study this interaction, two one-way ANOVA using the *Speller Size* factor were carried out for each attentional condition (i.e., overt and covert). No significant differences between speller sizes for the overt attention condition were found. However, there was significant differences between speller sizes under covert attention [$F(2, 22) = 4.32$; $p = 0.026$], offering the medium speller size a lower difficulty perceiving characters away from the centre than the large one ($p = 0.017$).

3.3. Satisfaction Dimension. Table 7 shows the usability questionnaire used for evaluating different dimensions concerning the participants' preference. Three ranks were proposed for each dimension: *rank 1*, the least; *rank 2*, intermediate; and *rank 3*, the most. The obtained results

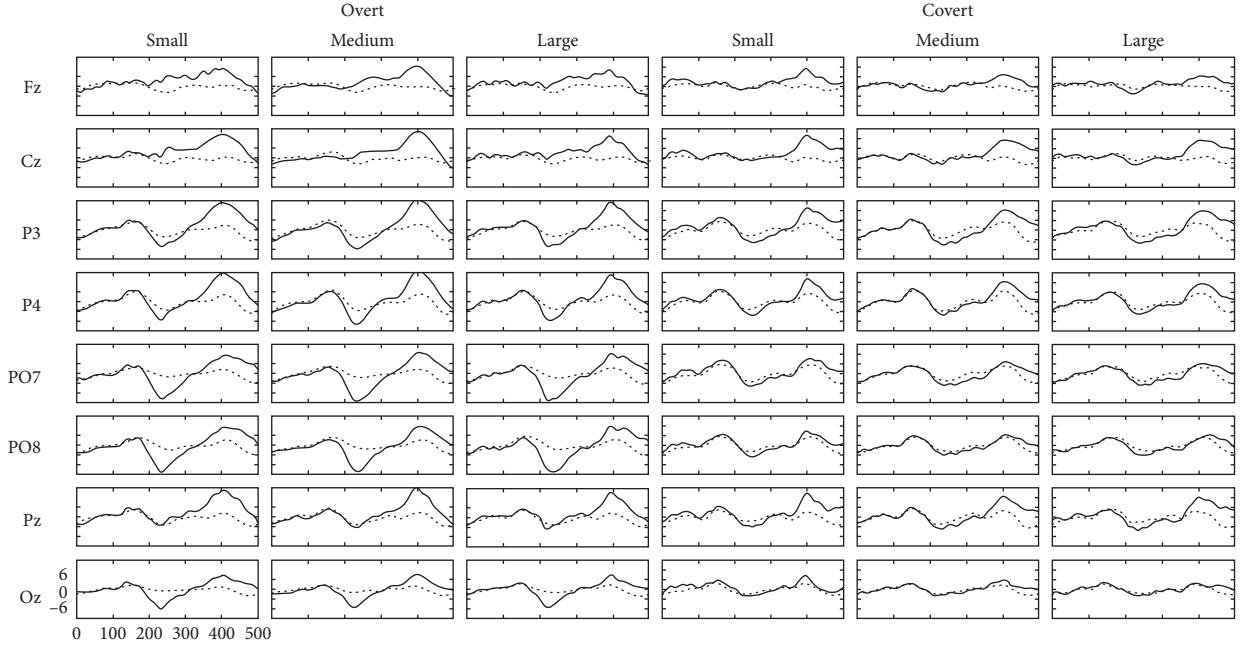


FIGURE 8: Grand average P300 waveforms (amplitude in μV for y axis and time in ms for x axis) for target stimuli (solid) and nontarget stimuli (dashed) for the eight electrodes used (Fz, Cz, P3, P4, PO7, PO8, Pz, and Oz) as a function of the attention condition (overt and covert) and speller size (small, medium, and large).

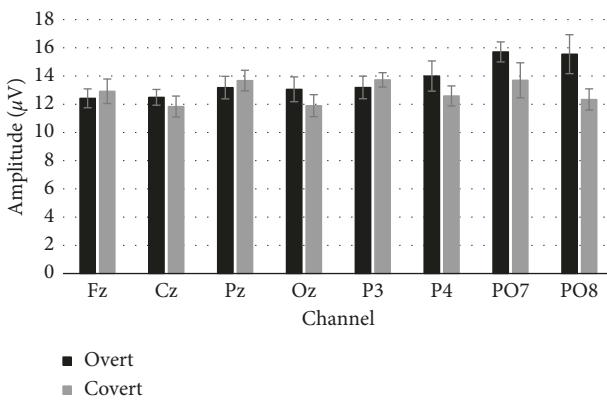


FIGURE 9: P300 peak amplitude for target and nontarget stimuli for each channel.

correspond to the subjects' distribution according to the rank in function of the speller size for each dimension. To determine whether the users' preference for speller size was significant, Fisher's exact test was conducted. The results relative to the overt attention condition were presented first, followed by the results relative to the covert attention condition. Finally, both attentional conditions were considered together to offer a general perspective.

3.3.1. Overt Attention Condition. Regarding the result of overt attention, according to the test, no significant differences between speller sizes mobile for the *favourite* and *controllable* dimensions were obtained. However, Fisher's exact test showed a significant relation between the speller

size and preference (order rank) for the following dimensions:

(i) Complex:

- 50% of subjects chose the large speller size as the least *complex* (*rank 1*)
- 75% of subjects chose the small speller size as the most *complex* (*rank 3*)
- 50% of subjects chose the medium speller size as the intermediate (*rank 2*)

(ii) Comfortable:

- 67% of subjects chose the small speller size as the least *comfortable* (*rank 1*)
- 42% of subjects chose the large speller size as the most *comfortable* (*rank 3*)
- 75% of subjects chose the medium speller size as the intermediate (*rank 2*)

(iii) Stressful:

- 50% of subjects chose the small speller size as the least *stressful* (*rank 1*)
- 50% of subjects chose also the small speller size as the most *stressful* (*rank 3*)
- 67% of subjects chose the medium speller size as the intermediate (*rank 2*)

(iv) Tiring:

- 58% of subjects chose the large speller size as the least *tiring* (*rank 1*)
- 75% of subjects chose the small speller size as the most *tiring* (*rank 3*)

TABLE 5: VAS fatigue and NASA-TLX scores (mean \pm SE), including six different dimensions such as mental demand, physical demand, temporal demand, effort, performance, and frustration.

Parameters	Overt attention				Covert attention			
	Small	Medium	Large	Mean	Small	Medium	Large	Mean
VAS fatigue	4 \pm 0.6	2.8 \pm 0.7	4.5 \pm 1	3.8 \pm 0.5	5.42 \pm 0.7	4.8 \pm 0.7	5.3 \pm 0.8	5.2 \pm 0.2
Total workload	40.4 \pm 7.2	38.22 \pm 4.8	41.2 \pm 6.4	39.9 \pm 0.9	65.1 \pm 5.4	60.4 \pm 6.4	66.9 \pm 4.9	64.1 \pm 1.9
Mental demand	9.9 \pm 2.8	11.7 \pm 2.9	12.5 \pm 2.7	11.4 \pm 0.8	16.7 \pm 2.6	16.8 \pm 2.7	18.0 \pm 3	17.2 \pm 0.4
Physical demand	3.8 \pm 2.1	4.2 \pm 1.7	6.3 \pm 2.3	4.8 \pm 0.8	6.9 \pm 2.7	3.31 \pm 1.1	6.39 \pm 1.9	5.5 \pm 1.1
Temporal demand	7.3 \pm 1.3	9.2 \pm 1.8	6.8 \pm 1.8	7.8 \pm 0.7	13.2 \pm 2.9	9.6 \pm 1.9	13.4 \pm 2.4	12.1 \pm 1.2
Effort	9.5 \pm 2.7	7.9 \pm 1.7	7.5 \pm 1.8	8.3 \pm 0.6	13.3 \pm 2.3	13.4 \pm 2.1	14.7 \pm 2.4	13.8 \pm 0.4
Performance	4.8 \pm 1.8	3.8 \pm 1.6	8.6 \pm 2.7	5.7 \pm 1.5	10.1 \pm 1.8	13.6 \pm 3.3	7.61 \pm 1.9	10.4 \pm 1.7
Frustration	4.8 \pm 2.3	1.2 \pm 0.7	3.5 \pm 1.2	3.2 \pm 1.1	4.95 \pm 2.5	3.3 \pm 1.1	6.78 \pm 2.0	5.01 \pm 1

TABLE 6: Average scores (\pm SE) of answers to the usability questionnaire.

Statements	Overt attention			Covert attention		
	Small	Medium	Large	Small	Medium	Large
Statement 1: difficulty perceiving different characters	2.4 \pm 0.8	1.3 \pm 0.3	1.8 \pm 0.5	5.75 \pm 0.7	4 \pm 0.4	5.75 \pm 0.5
Statement 2: difficulty perceiving characters away from the centre	2.2 \pm 0.7	1.2 \pm 0.4	1.8 \pm 0.5	6.5 \pm 0.6	5.9 \pm 0.5	7.8 \pm 0.4
Statement 3: difficulty distinguishing different rows and columns	2.7 \pm 0.8	1.5 \pm 0.4	1.6 \pm 0.5	6.8 \pm 0.7	5 \pm 0.7	5.8 \pm 0.8

TABLE 7: Subjects' distribution for each dimension.

Dimension	Overt attention			Statistics (Fisher's exact test)	Covert attention			Statistics (Fisher's exact test)
	Rank	Small	Medium		Small	Medium	Large	
Favourite	1	6	3	3	$F = 6.337; p = 0.182$	5 (42%)	3 (25%)	4 (33%)
	2	2	3	7		1 (8%)	3 (25%)	8 (67%)
	3	4	6	2		6 (50%)	6 (50%)	0 (0%)
Complex	1	1 (8%)	5 (42%)	6 (50%)	$F = 13.715; p = 0.006$	5	5	2
	2	2 (17%)	6 (50%)	4 (33%)		3	5	4
	3	9 (75%)	1 (8%)	2 (17%)		4	2	6
Comfortable	1	8 (67%)	0 (0%)	4 (33%)	$F = 19.358; p < 0.001$	7 (58%)	0 (0%)	5 (42%)
	2	0 (0%)	9 (75%)	3 (25%)		1 (8%)	8 (67%)	3 (25%)
	3	4 (33%)	3 (25%)	5 (42%)		4 (33%)	4 (33%)	4 (33%)
Stressful	1	6 (50%)	1 (8%)	5 (42%)	$F = 13.715; p = 0.006$	7	3	2
	2	0 (0%)	8 (67%)	4 (33%)		1	6	5
	3	6 (50%)	3 (25%)	3 (25%)		4	3	5
Controllable	1	6	2	4	$F = 6.24; p = 0.205$	4	4	4
	2	1	6	5		1	6	5
	3	5	4	3		7	2	3
Tiring	1	2 (17%)	3 (25%)	7 (58%)	$F = 21.288; p < 0.001$	4 (33%)	3 (25%)	5 (42%)
	2	1 (8%)	9 (75%)	2 (17%)		1 (8%)	8 (67%)	3 (25%)
	3	9 (75%)	0 (0%)	3 (25%)		7 (58%)	1 (8%)	4 (33%)

Percentages are indicated for those dimensions with significant differences. Ranks are ordered as follows: rank 1, the least; rank 2, intermediate; and rank 3, the most. Significant results have been denoted in bold.

75% of subjects chose the medium speller size as the intermediate (rank 2)

In the overt attention condition, the following results could be observed: the small size was considered as the most *complex* and *tiring*, and the less *comfortable*; the medium size was classified as moderate, without presenting any negative score (i.e., the most *complex*, *stressful* or *tiring*, but the least *comfortable*); and the large size obtained the best value for *complex*, *comfortable*, and *tiring*. Additionally, it should be remarked that the medium size was the condition in which the users showed a greater agreement scoring it.

3.3.2. Covert Attention Condition. In reference to the results relative to the covert attention condition, according to the test, no significant differences among speller sizes for the *complex*, *stressful*, and *controllable* dimensions were obtained. However, Fisher's exact test showed a significant relation between the speller sizes, under covert attention, and preference (order rank) for the following dimensions:

(i) Favourite:

42% of subjects chose the small speller size as the least *favourite* (rank 1)

50% of subjects chose the small and the medium speller sizes as the most *favourite* (*rank 3*)
 67% of subjects chose the large speller size as the intermediate (*rank 2*)

(ii) Comfortable:

58% of subjects chose the small speller size as the least *comfortable* (*rank 1*)
 33% of subjects chose the large or the small or the medium speller size as the most *comfortable* (*rank 3*)
 67% of subjects chose the medium speller size as the intermediate (*rank 2*)

(iii) Tiring:

42% of subjects chose the large speller size as the least *tiring* (*rank 1*)
 58% of subjects chose the small speller size as the most *tiring* (*rank 3*)
 67% of subjects chose the medium speller size as the intermediate (*rank 2*)

In general, under the covert attention condition, the following results have been obtained: the small size was considered as the most *favourite* (tied with the medium size) and the most *tiring*; the medium size was scored as the most *favourite* (tied with the small size); and the large size was considered as the least *tiring*.

3.3.3. Overt and Covert Attention Conditions. To offer a global perspective, each variable of the satisfaction construct was classified as negative (*complex*, *stressful*, and *tiring*) or positive (*controllable*, *comfortable*, and *favourite*). Thanks to this classification, it can be affirmed that the small speller size presented more negative than positive dimensions since it was considered as very *complex*, not very *comfortable* and *tiring* (Figure 10). The large speller size obtained as much positive as negative dimensions because it was considered moderately *comfortable* and not very *tiring*. By contrast, most dimensions of the medium speller size presented mainly positive values, and it was classified largely in *rank 2*. Thus, in general, the medium speller size was chosen as enough satisfactory by users.

4. Discussion

In this study, the impact of speller sizes has been evaluated under constraint or nonconstraint conditions (i.e., covert and overt attention conditions, respectively) on both objective and subjective parameters. It has been shown that constraint to eye movement represents an important effort that is correlated with lower performance and higher workload. The usability measures suggested that medium speller size is the most convenient to ensure comfort and control using a visual P300-based speller.

Several studies are focused on defining optimal parameters to improve speller design in order to be more useful. Effects of matrix size [22, 26], interstimulus interval [22], luminosity contrast [24], and interface colour contrast

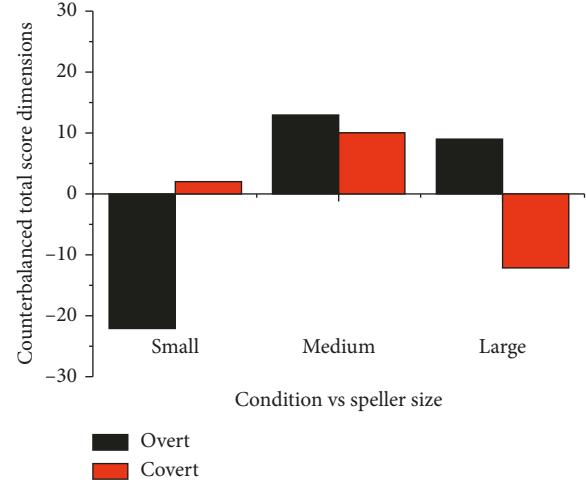


FIGURE 10: Representation of counterbalanced global dimensions to evaluate satisfaction of different interfaces. Rank was considered as a factor, which was multiplied by the subject's number that classified each dimension in ranks 1, 2, or 3. Each rank was translated in a positive or negative value (rank 1, ± 1 ; rank 2, ± 2 ; and rank 3, ± 3). Positive values were assigned to *favourite*, *comfortable*, and *controllable* dimensions. Negative values were attributed to *complex*, *stressful*, and *tiring* dimensions.

[25] on different factors, such as P300 event-related potential or subjective measures, have been considered. The main objective of this research was to design BCI systems that are not only accurate but also easily usable for healthy subjects and, especially, for patients.

The effect of matrix size has already been studied. However, speller size has not been considered when designing a P300-based speller. Different speller sizes have been proposed in the literature but without knowing the effects they could have on usability. The inclusion of usability analysis at the early stage could be beneficial for the progression of the ALS's research.

In this study, three usability factors were evaluated: effectiveness, efficiency, and satisfaction [12]. The main objective was to determine which speller size obtained the best degree of usability according to the three dimensions under each attentional condition, overt and covert attention. This study has demonstrated that speller size has an important influence on user performance and must be considered when a BCI system is being designed. The results presented above are discussed in the following paragraphs.

4.1. Effect of Speller Size and Attention on Effectiveness Dimension. The present study has replicated the results of previous works, in which it was shown that performance in classification was severely impaired due to the lack of ocular mobility (i.e., under covert attention condition) [30, 33]. However, the effect of Speller Size factor had not adequately been studied previously; therefore, we will focus our discussion on this factor.

On the one hand, for the calibration task in the overt attention condition, there were no significant differences in classification accuracy between sizes (Figure 6(a)). On the

other hand, for the covert attention, the medium speller size ($83.4 \pm 3.7\%$) seemed to show the best results, especially against the large size ($70.8 \pm 3.7\%$). These results could lead to the conclusion that the *Speller Size* factor is important to achieve a good *accuracy* in the case of users who do not have ocular mobility but less relevant when the user can control his/her eyes movements.

Regarding the *minimum number of flashes* required to achieve the maximum accuracy, it can be observed that more flashes were needed under covert attention (Figure 6(b)), presenting all sizes a similar *minimum number of flashes*. Under overt attention, all subjects obtained 100% *accuracy* before the 10 required flashes (one flash is the intensification of one row and one column), except for subject S4, who reached only 94% *accuracy* with the small and medium speller sizes. Contrastingly, under covert attention, the 100% *accuracy* was reached only by three participants (S2, S3, and S9) using the medium speller size.

In spite of having the possibility of reducing the time taken to select a letter during the online task (at least under overt attention), because the main objective of the study was to compare different speller sizes, we decided not to modify this parameter. For this reason, because the time required to select a symbol was always the same, the information transfer rate (ITR) to compare performances was not used. Nevertheless, it is important to notice that, under overt attention, a vast majority of subjects (8 subjects) was needed to calibrate the system with a reduced number of flashes when using the medium speller size. In this sense, the medium speller size under overt attention seems to require less number of flashes (4.9 ± 0.2 *minimum number of flashes*) and thus, less time to select a character. However, it was not the case for results obtained in constraint conditions where a higher number of flashes to obtain maximum classification performance were observed for all sizes (*minimum number of flashes*: small, 8.9 ± 0.3 ; medium, 9.0 ± 0.3 ; and large, 8.8 ± 0.3).

In reference to the performance in the online task, as expected, the *error performance* was affected by the attention condition (overt, 7.9%; covert, 49.3%). In addition, the *error performance* associated with the large speller size obtained the worst performance for both attentional conditions (Figure 7). In fact, the same trend was presented by overt and covert attention: the small (overt, $2.8 \pm 1.6\%$; covert, $45.8 \pm 6.7\%$) and medium (overt, $4.9 \pm 2.8\%$; covert, $43.7 \pm 7.2\%$) sizes have shown similar results, while the largest size offered the worst *error performance* regardless of *Attention* factor (overt, $16 \pm 4.5\%$; covert, $58.3 \pm 7.8\%$). In general, these results could lead us to think that there is a size from which the capacity to obtain an adequate *accuracy/error performance* decreases rapidly, i.e., the size between the medium and the large speller conditions.

In contrast with another study [29], in which participants had higher accuracy on a computer monitor than that on a mobile phone screen, the present work showed that the *error performance* was significantly higher for the large speller size compared to the small and medium speller sizes. In this sense, the worst performance is obtained when using the large speller size compared to the other two speller sizes.

Effectively, except for participants S2 and S6, the remaining users obtained worse or equal error percentages when using the large speller. It is important to mention that, in the other study [29], the only information provided about different parameters of the matrix size was the visual angle of the screen, which was 3.7° and 3.56° for global positioning system (GPS) and cell phone screen, respectively. No information was provided regarding speller size, symbol size, or distance between columns and rows. Probably, these parameter values were lower than those used in our small speller size, making it very difficult to identify different characters. Otherwise, in our experiment, the smallest speller size was chosen according to the smallest symbol size so that it could be used without loss of performance. This minimum size was reported in a different study [28]. Therein, symbol sizes would decrease the performance considerably, being probably the reason of the low performance obtained with the study on mobile phone and the GPS screens [29].

Sellers et al. [22] have reported that matrix size, i.e., the number of elements, has a significant effect on performance: a 3×3 matrix offered better results than a 6×6 matrix. In the present work, the results obtained suggest that speller size can also have a significant effect on user performance, and thus it is an important factor to consider it in the design. The obtained results show how *error performance* increases when using the large speller size, which has frequently been used by other researchers (e.g., [30, 33]). The best performances were obtained when using the small and medium speller sizes, achieving similar performance.

The analysis of layers in the online task, which is closely related to the size, has shown how the *Attention* factor presented an effect on all layers, not only on the outer ones. These results show the importance of ocular mobility, even when the stimuli are close to the point of view. On the contrary, the *Speller Size* factor has exclusively influenced *layer 1* and *layer 2*, the two most external layers. Specifically, in these layers (*layer 1* and *layer 2*), we observed that the large size showed a worse performance than the medium and small sizes. Thus, the worst combination is the large size in the external layers, that is, where the character is placed at the furthest position from the centre of the screen.

Ultimately, it seems clear that the large size is associated with a worse performance, especially under covert attention. Therefore, it should be recommended to avoid the large size in order to achieve a good level of *accuracy*, especially in the case of patients with impaired ocular mobility.

Finally, in reference to the analysis of the P300 signal, while matrix size affects P300 peak amplitude due to different target probabilities according to different matrices [22, 26], the obtained results showed no significant main effects on P300 response for speller size. Moreover, because the three speller sizes have the same matrix size (6×6), different P300 amplitudes for the target and nontarget conditions were similar for different speller sizes. Similar conclusions were obtained in another study [28]. However, the attention factor showed significant differences in amplitude levels of the target stimulus at channels P4 (overt, $14 \pm 1.07 \mu\text{V}$; covert, $12.59 \pm 0.71 \mu\text{V}$) and PO8 (overt,

$15.55 \pm 1.38 \mu\text{V}$; covert, $12.34 \pm 0.75 \mu\text{V}$). Thus, it has been corroborated that the performance differences previously presented between covert and overt attention conditions have a neural correlate [30].

4.2. Effect of Speller Size and Attention on Efficiency Dimension. In the present work, the fatigue and workload to efficiency assessments have been studied, including in the workload are the following dimensions: *total workload*, *mental demand*, *physical demand*, *temporal demand*, *effort*, *performance*, and *frustration*. Results showed that the effect of *Attention* factor has been important since it has shown significant differences in all measured dimensions, with the exception of *physical demand* and *frustration*. On the contrary, although the ANOVA did not show significant differences for the *Speller Size*, it can be observed in Table 5 that the medium speller has shown the lowest *total workload* score for both conditions, i.e., overt and covert attention conditions.

Regarding the questions related to perception (*statement 1*, *statement 2* and *statement 3*), it should be noted that the medium speller size has obtained the most appropriate scores for each statement in both conditions of the *Attention* factor (Table 6). In the same way, it can be observed how the scores under the covert attention condition have been higher for each of different speller sizes used. Therefore, it seems clear that the inability to move the eyes significantly affects the necessary cognitive resources used to control the interface. This disability should be considered in the case of several patients who cannot perform such action by using interfaces with an adequate size, since the medium size presented the best results despite the lack of significances, or new features that reduce the cognitive resources needed to use the speller (e.g., [44]).

In the previously mentioned work [29], subjects reported that smaller screens (GPS and especially mobile phone screens) were too difficult to read due to the difficulty to perceive the target symbol. However, the varying results between that work and our study could be explained by the differences in the employed sizes or other characteristics of the interface, such as the screen resolution.

In reference to the difficulty perceiving the presented stimuli (i.e., *statement 1*, *statement 2*, and *statement 3*), thanks to the related results with the *Attention* factor, it could be concluded that the inability to move the eyes is a key factor that provokes a significant impairment in the facility to perform the task. Regarding the speller size, the medium speller presented the lowest difficulty perceiving the characters (*statement 1*) in contrast to small and large spellers, while the small size speller showed the worst results in difficulty distinguishing different row and columns (*statement 3*) in contrast to medium and large spellers. Thus, the medium speller size was always related to the best, or lower, scores in difficulty since no other speller size obtained better significant results versus it. Additionally, the interaction effect between *Speller Size* and *Attention* in *statement 2* (the medium speller size was better than the large one under covert attention, but not under overt

attention) shows that the size is not important under the overt attention condition. This factor should instead be considered under covert attention.

4.3. Effect of Speller Size and Attention on Satisfaction Dimension. Satisfaction has been studied according to six dimensions: *favourite*, *complex*, *comfortable*, *stressful*, *controllable*, and *tiring*. For each dimension, the participants had to rank the three speller sizes. On the one hand, the most remarkable point regarding the overt attention condition was that, according to Table 7, the small speller size was the worst rated of all, while the medium and large speller sizes obtained the best scores. On the other hand, for the covert attention condition, the large speller was the worst valued in general, with the medium speller being the best valued. Therefore, three conclusions could be obtained: (i) the small speller was the worst rated in the overt attention condition, (ii) the large speller was the worst rated in the covert condition, and (iii) the medium speller has obtained the best rating in both conditions. Thus, as it was depicted in Figure 10, the medium speller could be denoted as the most satisfactory, regardless of the condition, since it has shown a tendency to offer the best results.

5. Conclusions

The present work has investigated the usability of three speller sizes under overt and covert attention handling a P300-based BCI speller. The obtained results showed that, in both attentional conditions (i.e. overt and covert attention), the speller size had significant effects, or trends that should not be ignored, on user usability considering the effectiveness, efficacy, and satisfaction. Regarding the effectiveness, the large speller size offered the worst results under overt and covert attention, while medium and small offered similar results, with a slight superiority of the medium size. In reference to efficacy, the large speller size offered a trend in which it gathered the worst values according to different NASA-TLX workload measures and fatigue. Finally, regarding the satisfaction dimension, the medium speller size was the best rated, while the larger speller size obtained the worst general results under covert attention because the gaze movements were restricted and the distance between symbols was larger. Additionally, the small speller size offered the worst results under overt attention, due perhaps to the denoted tiredness provoked by this size. Therefore, based on the trends offered by the medium speller size and the lack of worst obtained results when using it among all usability dimensions, this size may be the most recommended to employ.

In short, we have demonstrated that the speller size should be considered in the usability of a P300-based BCI speller, although it may also depend on whether or not the participant has the ability of gaze control. For future works, we recommend to continue the study relative to the speller size to be able to confirm the present results found. Some examples of possible studies include testing other sizes, such as those under our small speller size or the size in which the

performance in *accuracy/error performance* was significantly decreased between the medium and large speller sizes. In addition, in order to fulfill this aim, it could be convenient to increase the sample size and the number of letters written in the calibration and online tasks, as well as to assess this effect of the speller sizes in patients, instead of able-bodied participants, under the covert attention condition.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partially supported by the Spanish Ministry of Economy and Competitiveness through the project LICO (DPI2015-67064-R), the European Regional Development Fund (ERDF), and the University of Malaga. Moreover, the authors would like to thank all participants for their cooperation.

References

- [1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain-computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [2] N. Birbaumer, "Breaking the silence: brain-computer interfaces (BCI) for communication and motor control," *Psychophysiology*, vol. 43, no. 6, pp. 517–532, 2006.
- [3] E. W. Sellers and E. Donchin, "A P300-based brain-computer interface: initial tests by ALS patients," *Clinical Neurophysiology*, vol. 117, no. 3, pp. 538–548, 2006.
- [4] M. Marchetti and K. Priftis, "Effectiveness of the P3-speller in brain-computer interfaces for amyotrophic lateral sclerosis patients: a systematic review and meta-analysis," *Frontiers in Neuroengineering*, vol. 7, pp. 1–6, 2014.
- [5] E. Yin, T. Zeyl, R. Saab, D. Hu, Z. Zhou, and T. Chau, "An auditory-tactile visual saccade-independent P300 brain-computer interface," *International Journal of Neural Systems*, vol. 26, no. 1, article 1650001, 2016.
- [6] A. Rezeika, M. Benda, P. Stawicki, F. Gembler, A. Saboor, and I. Volosyak, "Brain-computer interface spellers: a review," *Brain Sciences*, vol. 8, no. 4, p. 57, 2018.
- [7] L. A. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988.
- [8] V. Guy, M.-H. Soriani, M. Bruno, T. Papadopoulos, C. Desnuelle, and M. Clerc, "Brain computer interface with the P300 speller: usability for disabled people with amyotrophic lateral sclerosis," *Annals of Physical and Rehabilitation Medicine*, vol. 61, no. 1, pp. 5–11, 2018.
- [9] ISO9241-8, *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)-Part 8: Requirements for Displayed Colours, Part 11 Guid Usability*, vol. 11, p. 22, 1997.
- [10] J. Nielsen, "What is usability?," in *User Experience Re-Mastered*, pp. 3–22, Elsevier, Burlington, NJ, USA, 1st edition, 2010.
- [11] J. Nielsen, "Heuristic evaluation," in *Usability Inspection Methods*, J. Nielsen and R. L. Mack, Eds., pp. 25–62, John Wiley & Sons, New York, NY, USA, 1994.
- [12] E. Frækjaer, M. Hertzum, and K. Hornbaek, "Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?," in *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, Hague, Netherlands, April 2000.
- [13] J. Jeng, "Usability assessment of academic digital libraries: effectiveness, efficiency, satisfaction, and learnability," *Libri*, vol. 55, no. 2-3, 2005.
- [14] M. Hertzum, "Images of usability," *International Journal of Human-Computer Interaction*, vol. 26, no. 6, pp. 567–600, 2010.
- [15] H. Kececi, Y. Degirmenci, and S. Atakay, "Habituation and dishabituation of P300," *Cognitive and Behavioral Neurology*, vol. 19, no. 3, pp. 130–134, 2006.
- [16] A. Murata and A. Uetake, "Evaluation of mental fatigue in human-computer interaction," in *Proceedings 10th IEEE International Workshop on Robot and Human Interactive Communication*, pp. 630–635, Bordeaux, France, 2001.
- [17] G. R. Mangun and L. A. Buck, "Sustained visual-spatial attention produces costs and benefits in response time and evoked neural activity," *Neuropsychologia*, vol. 36, no. 3, pp. 189–200, 1998.
- [18] S. C. Kleih, F. Nijboer, S. Halder, and A. Kübler, "Motivation modulates the P300 amplitude during brain-computer interface use," *Clinical Neurophysiology*, vol. 121, no. 7, pp. 1023–1031, 2010.
- [19] J. Polich and A. Kok, "Cognitive and biological determinants of P300: an integrative review," *Biological Psychology*, vol. 41, no. 2, pp. 103–146, 1995.
- [20] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.
- [21] J. Lu, W. Speier, X. Hu, and N. Pouratian, "The effects of stimulus timing features on P300 speller performance," *Clinical Neurophysiology*, vol. 124, no. 2, pp. 306–314, 2013.
- [22] E. W. Sellers, D. J. Krusinski, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "A P300 event-related potential brain-computer interface (BCI): the effects of matrix size and inter stimulus interval on performance," *Biological Psychology*, vol. 73, no. 3, pp. 242–252, 2006.
- [23] D. J. McFarland, W. A. Sarnacki, G. Townsend, T. Vaughan, and J. R. Wolpaw, "The P300-based brain-computer interface (BCI): effects of stimulus rate," *Clinical Neurophysiology*, vol. 122, no. 4, pp. 731–737, 2011.
- [24] Y. Li, S. Bahn, C. S. Nam, and J. Lee, "Effects of luminosity contrast and stimulus duration on user performance and preference in a P300-based brain-computer interface," *International Journal of Human-Computer Interaction*, vol. 30, no. 2, pp. 151–163, 2014.
- [25] C. S. Nam, Y. Li, and S. Johnson, "Evaluation of P300-based brain-computer interface in real-world contexts," *International Journal of Human-Computer Interaction*, vol. 26, no. 6, pp. 621–637, 2010.
- [26] B. Z. Allison and J. A. Pineda, "ERPs evoked by different matrix sizes: implications for a brain computer interface (BCI) system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 11, no. 2, pp. 110–113, 2003.

- [27] R. Ron-Angevin, S. Varona-Moya, and L. d. Silva-Sauer, “Initial test of a T9-like P300-based speller by an ALS patient,” *Journal of Neural Engineering*, vol. 12, no. 4, p. 046023, 2015.
- [28] M. Salvaris and F. Sepulveda, “Visual modifications on the P300 speller BCI paradigm,” *Journal of Neural Engineering*, vol. 6, no. 4, article 046011, 2009.
- [29] Y. Li, C. S. Nam, B. B. Shadden, and S. L. Johnson, “A p300-based brain-computer interface: effects of interface type and screen size,” *International Journal of Human-Computer Interaction*, vol. 27, no. 1, pp. 52–68, 2011.
- [30] M. S. Treder and B. Blankertz, “(C)overt attention and visual speller design in an ERP-based brain-computer interface,” *Behavioral and Brain Functions*, vol. 6, no. 1, p. 28, 2010.
- [31] F. Aloise, P. Aricò, F. Schettini et al., “A covert attention P300-based brain-computer interface: Geospell,” *Ergonomics*, vol. 55, no. 5, pp. 538–551, 2012.
- [32] M. S. Treder, N. M. Schmidt, and B. Blankertz, “Gaze-independent brain-computer interfaces based on covert attention and feature attention,” *Journal of Neural Engineering*, vol. 8, no. 6, article 066003, 2011.
- [33] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, “Does the ‘P300’ speller depend on eye gaze?,” *Journal of Neural Engineering*, vol. 7, no. 5, p. 056013, 2010.
- [34] G. Westheimer, “Visual acuity,” *Annual Review of Psychology*, vol. 16, no. 1, pp. 359–380, 1965.
- [35] ISO 13407, “ISO 13407: human centred design process for interactive systems,” *Human Factors*, 1999.
- [36] V. Venkatesh, M. G. Morris, and G. B. Davis, “User acceptance of information technology: toward a unified view,” *MIS Quarterly*, 2017.
- [37] E. Brangier and J. Barcenilla, “Qu'est-ce que l' utilisabilité,” in *Concevoir un Produit Facile à Utiliser*, Éditions d'Organisation, Paris, 2003.
- [38] A. Dillon and M. Morris, “Power, perception and performance: from usability engineering to Technology acceptance with the P3 model of user response,” in *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, Houston, TX, USA, 1999.
- [39] V. Roto, “User experience from product creation perspective,” in *Towards a UX Manifesto Workshop, Conjunction with HCI 2007*, Lancaster, UK, 2007.
- [40] E. Donchin, K. M. Spencer, and R. Wijesinghe, “The mental prosthesis: assessing the speed of a P300-based brain-computer interface,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 174–179, 2000.
- [41] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (task load index): results of empirical and theoretical research,” *Advances in Psychology*, vol. 52, pp. 139–183, 1988.
- [42] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” *International Journal of Human-Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.
- [43] E. Kim, J. Lovera, L. Schaben, J. Melara, D. Bourdette, and R. Whitham, “Novel method for measurement of fatigue in multiple sclerosis: real-time digital fatigue score,” *Journal of Rehabilitation Research and Development*, vol. 47, no. 5, p. 477, 2010.
- [44] J. Qu, F. Wang, Z. Xia et al., “A novel three-dimensional P300 speller based on stereo visual stimuli,” *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 4, pp. 392–399, 2018.

Review Article

Enhancing the Usability of Brain-Computer Interface Systems

Hyun Jae Baek ,¹ Min Hye Chang ,² Jeong Heo ,³ and Kwang Suk Park ,⁴

¹Department of Medical and Mechatronics Engineering, Soonchunhyang University, Asan, Republic of Korea

²Korea Electrotechnology Research Institute (KERI), Ansan, Republic of Korea

³Artificial Intelligence Laboratory, Software Center, LG Electronics, Seoul, Republic of Korea

⁴Department of Biomedical Engineering, College of Medicine, Seoul National University, Seoul, Republic of Korea

Correspondence should be addressed to Kwang Suk Park; kspark@bmsil.snu.ac.kr

Received 21 March 2019; Revised 2 May 2019; Accepted 14 May 2019; Published 16 June 2019

Academic Editor: Pietro Aricò

Copyright © 2019 Hyun Jae Baek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Brain-computer interfaces (BCIs) aim to enable people to interact with the external world through an alternative, nonmuscular communication channel that uses brain signal responses to complete specific cognitive tasks. BCIs have been growing rapidly during the past few years, with most of the BCI research focusing on system performance, such as improving accuracy or information transfer rate. Despite these advances, BCI research and development is still in its infancy and requires further consideration to significantly affect human experience in most real-world environments. This paper reviews the most recent studies and findings about ergonomic issues in BCIs. We review dry electrodes that can be used to detect brain signals with high enough quality to apply in BCIs and discuss their advantages, disadvantages, and performance. Also, an overview is provided of the wide range of recent efforts to create new interface designs that do not induce fatigue or discomfort during everyday, long-term use. The basic principles of each technique are described, along with examples of current applications in BCI research. Finally, we demonstrate a user-friendly interface paradigm that uses dry capacitive electrodes that do not require any preparation procedure for EEG signal acquisition. We explore the capacitively measured steady-state visual evoked potential (SSVEP) response to an amplitude-modulated visual stimulus and the auditory steady-state response (ASSR) to an auditory stimulus modulated by familiar natural sounds to verify their availability for BCI. We report the first results of an online demonstration that adopted this ergonomic approach to evaluating BCI applications. We expect BCI to become a routine clinical, assistive, and commercial tool through advanced EEG monitoring techniques and innovative interface designs.

1. Introduction

An episode of the TV series *Star Trek* first shown in 1966 describes a man, Captain Pike, who suffers from locked-in syndrome. He is cognitively intact, but his body is paralyzed, leaving him confined to a wheelchair controlled by his brain wave responses to flashing lights that indicate “yes” and “no.” Dr. Octopus, the villain in the movie *Spiderman 2*, which premiered in 2004, used brain-controlled equipment to operate four mechanical arms designed with tentacle-like flexibility and gripping capabilities. That machine was controlled by thought through an interface at the spinal cord level. Jake Sully, the protagonist of *Avatar*, which premiered in 2009, was in a wheelchair in his human body, but he could

walk, run, and jump in the form of his avatar, a 10-foot alien. All those characters are fictional representations of the ultimate goal of research into brain-computer interfaces (BCIs), sometimes called brain-machine interfaces. Human physical interactions, including communication, require motor control processes that use neuromuscular systems to activate and coordinate muscle movements. An individual’s intent triggers the activation of a specific brain area, sending signals through the peripheral nervous system to muscles that perform the movement necessary to complete the intended task. During the past several decades, much research has been done to bypass nonfunctional muscular control channels, attempting to translate a person’s intent by analyzing brain signals to empower patients with cognitive

or sensorimotor problems [1–4]. The primary goal of BCI technology is to provide communication capabilities that can improve the quality of life for people severely disabled by neuromuscular impairments including amyotrophic lateral sclerosis (ALS), brainstem stroke, cerebral palsy, or spinal cord injury. For example, Sellers and Donchin [5] and Nijboer et al. [6] have reported that ALS patients can communicate using a P300 speller. Pfurtscheller et al. trained a quadriplegic patient to control an electrically driven hand orthosis using EEG signals recorded through the sensorimotor cortex [7]. In a case study performed by Leeb et al., the spinal-cord-injured subject was able to generate bursts of beta oscillations in an EEG by imagining the movements of his paralyzed feet, and those oscillations were used for self-paced BCI control of a wheelchair in virtual reality [8]. Rapid progress toward those goals is being made by many BCI research groups, allowing BCI research to expand from communication to rehabilitation medicine through neurofeedback for stroke, autism, attention deficit hyperactivity disorder, and other disorders [9–12]. However, although significant advancement has been made in BCI research, much more progress must be made before BCI can have a significant effect in real-world environments [13–15]. In particular, previous research has generally lacked understanding of (or perhaps simply failed to pay attention to) ergonomics issues, such as aesthetic designs, user-friendly interaction methods, and usability. Progress in BCI research has mainly been made by enhancing BCI performance with respect to accuracy, information transfer rate (ITR), or the number of possible selections. Thus, most of the BCI literature focuses on advanced signal processing methods or new applications or task designs, and the applications controlled by state-of-the-art interfaces have largely been intended for research-oriented environments. Most current BCI techniques face two major challenges that prevent them from being useful in real-world settings.

(1) Advanced monitoring of brain activity: BCI research has used various neural signals that can be recorded noninvasively, such as electroencephalograms (EEGs), magnetoencephalograms, functional magnetic resonance imaging, and near-infrared spectroscopy. Of those, the EEG is the most commonly used method because it is noninvasive, offers high temporal resolution and portability, and has a reasonable cost. Conventional wet Ag/AgCl electrodes are used most frequently to measure EEG signals because their characteristics have been widely studied and discussed in detail [16–18]. The quality of EEG signals measured using those electrodes with skin preparation techniques and conductive gels is excellent. However, the requirement for wet electrodes greatly limits the applicability of BCIs in everyday use [19–22]. For practical use, BCIs should avoid typical EEG preparation procedures, including head measurement for accurate electrode placement and scalp preparation that requires an abrasive paste or gel to reduce skin-electrode impedance. Furthermore, electrodes for daily BCI use should not

make users feel uncomfortable or look unusual. It should be possible to take EEG signal measurements from simple caps that contain EEG electrodes in appropriate montages without the need to remove or specially treat scalp hair.

(2) Interface paradigm design: In a state-of-the-art BCI system, the control command, such as moving a cursor, is typically assigned to a specific mental state. The subject needs to perform a specific mental task to encode the desired control command through an attention shift or other voluntary regulation of the EEG. Currently, several types of EEG signals have been recognized: sensorimotor rhythm (also known as the μ/β rhythm), slow cortical potential, steady-state visual evoked potential (SSVEP), and event-related P300 potential. According to the literature survey performed by Hwang et al. [23], the most frequently used BCI paradigm is the motor imagery method. During actual or imagined movement, slow negative voltage shifts occur in EEGs recorded over the sensorimotor cortex, and the intention of a subject can be detected using that voltage shift. Recent motor imagery-based BCIs have used rhythmic EEG activity called event-related desynchronization/synchronization. During actual or imagined movement, event-related desynchronization occurs predominantly over the contralateral brain motor area, making it useful as a signal for a BCI system. Those methods do not require external stimuli to induce the desired EEG response. However, some drawbacks, including poor multidimensional control, high probability of error, and a need for long-term training, have led to a decrease in the use of motor imagery from 2007 to 2011. During that time, the proportion of research into the visual P300 and SSVEP paradigms has increased significantly. P300 and SSVEP require little training time and show a relatively high information transfer rate compared with other BCI paradigms. However, the requirement for a visual stimulus, such as flashing digits, letters, or other symbols that a user has to watch, has limited the flexibility, accessibility, and usability of those BCIs in real-world applications. Ideally, human-computer interactions should be free from sudden changes in luminance or high-contrast visual objects to minimize user visual fatigue and discomfort, especially for long-term use.

In 2012, Liao et al. proposed the concept of augmented BCI (ABCi) that would be appropriate for everyday use [24]. ABCi aims to expand the application of BCI technologies from their current laboratory or clinical settings to normal daily life by making them function while people move and interact with their environment. According to Liao et al.'s definition, ABCi includes nonintrusive and rapid-setup EEG solutions that require no or minimal training and thereby provide stability, robustness, comfort, and longevity for accurate long-term data collection. It also includes advanced

algorithmic approaches to analyzing and interpreting brain signals measured under noisy, real-world conditions. In this review, we emphasize new paradigm designs that fit the scope of ABCI and will not induce fatigue or discomfort during everyday, long-term use. First, we survey BCI articles that discuss ABCI research. Then, we demonstrate SSVEP-based and auditory steady-state response- (ASSR-) based BCIs that use recently developed polymer foam-based capacitively coupled EEG electrodes. Our demonstration study was approved by the Institutional Review Board of the Seoul National University College of Medicine, Seoul, Korea.

2. Advanced EEG Monitor

Technological advancements have greatly simplified the measurement and assessment of biopotential signals, particularly electrocardiograms. However, the sites for EEG electrodes are mostly covered with hair, and EEG signals are weaker than those used in other bio-potential measurement tools, which makes the use of dry electrodes in EEG difficult. Most dry EEG electrodes make signal measurements by penetrating the outermost layer of the skin, the stratum corneum, using microelectromechanical or carbon nanotube (CNT) techniques [25–28]. However, those types of dry electrodes are somewhat invasive, and electrodes that penetrate tissue always carry the risk of infection. In addition, those techniques do not allow EEG signals to be recorded through hair, and therefore, hair and scalp preparation is still required.

Another approach is to use electrode-finger-based sensors for EEG acquisition over hair (Figure 1(a)) [29, 36, 37]. This kind of electrode offers high geometric conformity between the electrode fingers and the irregular scalp surface, thus maintaining low electrode impedance. Additionally, the flexible substrate in which the spring probes are inserted permits the attachment of the sensor to the scalp without pain when force is applied. Similarly, a flexible, low-cost electrode about the size of a toothbrush made of polymer silver-coated bristles was suggested in 2011 (Figure 1(b)) and showed BCI feasibility using the motor imagery and oddball paradigms [30]. The main drawback of those electrodes is that they still require skin preparation to ensure contact between the finger-electrode and the scalp. Also, some participants reported prickling and other uncomfortable sensations. Figure 1(c) shows a reverse-curve-arch-shaped dry EEG electrode 3D-printed from sterling silver to increase the skin-electrode contact area over hair [31]. The curvature of the arches was designed to match the curvature of the scalp to maximize the contact area and disperse the pressure, thereby lessening the pain induced by conventional finger-type EEG electrodes.

Lee et al. proposed an electrode composed of a CNT/aPDMS (adhesive polydimethylsiloxane) nanocomposite material (Figure 1(d)) [32]. This electrode is elastic, highly conductive, self-adhesive, and able to make conformational contact with and attachment to a hairy scalp. Hundreds of conductive cylindrical pillars coated with a Parylene C insulation layer were fabricated on a conductive disk. The CNT/aPDMS layer was then attached to the disk to transmit

the EEG signal to the pillar and eliminate the air gap caused by hair, thereby maximizing the capacitance between the electrode and the scalp. The top of the disk was designed to be solderable, enabling the electrode to be connected to a variety of commercial EEG acquisition systems. Even though positive results have been published, these electrodes still have some drawbacks: they require multistep preparations and obtrusive wiring interfaces.

Epidermal electronics is an emerging class of integrated electronic systems that achieve thicknesses, effective elastic moduli, bending stiffness, and areal mass densities that match the skin [38, 39]. This technology has been proved feasible for many medical applications, such as monitoring vital signs. In particular, Norton et al. demonstrated an ultrathin, foldable neural electrode platform that could measure EEG signals from the surfaces of the outer ear (the auricle) and adjacent regions (the mastoid), as shown in Figure 1(e) [33]. The epidermal EEG electrode offers conformal contact and adequate adhesion based on van der Waals interactions alone in a manner that is mechanically unnoticeable to the user. It stayed well attached to skin with a complex surface topology (the auricle and mastoid) for more than 2 weeks, offering continuous monitoring without frequent removal or reapplication. That demonstration also showed BCI capabilities using SSVEP and P300. The electrode is soft, stretchable, and lightweight, so it can offer long-term, high-fidelity recording of EEG signals in daily life without user discomfort. However, the current sensing platform requires careful device mounting and complete dissolution of the polymer backing to allow successful EEG acquisition. Further progress in this area should consider people who do not want stick on temporary tattoos on their bodies.

A capacitive noncontact electrode that can measure EEG signals despite insulation by hair is another promising alternative to conventional wet electrodes for next-generation EEG measurement. An EEG could be measured through hair using displacement current with capacitive coupling. According to an electronic model of capacitive EEG measurement described in various studies, the capacitive measurement of EEG signals is characterized by very high electrode impedance created by the insulating effect of hair between the sensor and the scalp. Because the electrode impedance in a capacitive electrode is much higher than that in conventional wet electrodes, a high-input impedance amplifier is used as the active electrode in each sensor to convert the displacement current into voltage. Designing such an input impedance amplifier is a major challenge. Chi et al. developed a capacitive electrode that uses a custom integrated, high-impedance, low noise analog front-end [40]. The amplifier fully bootstraps both the internal and external parasitic impedances by including a low-leakage on-chip biasing network without external resistors that operates from hundreds of giga ohms to tera ohms. They also demonstrated an SSVEP-based phone dialing application that used the developed capacitive EEG electrode with two subjects [41]. The result was feasible, but the average ITR was lower than that with conventional wet electrodes. Around the same time, Baek et al. suggested a polymer foam-based

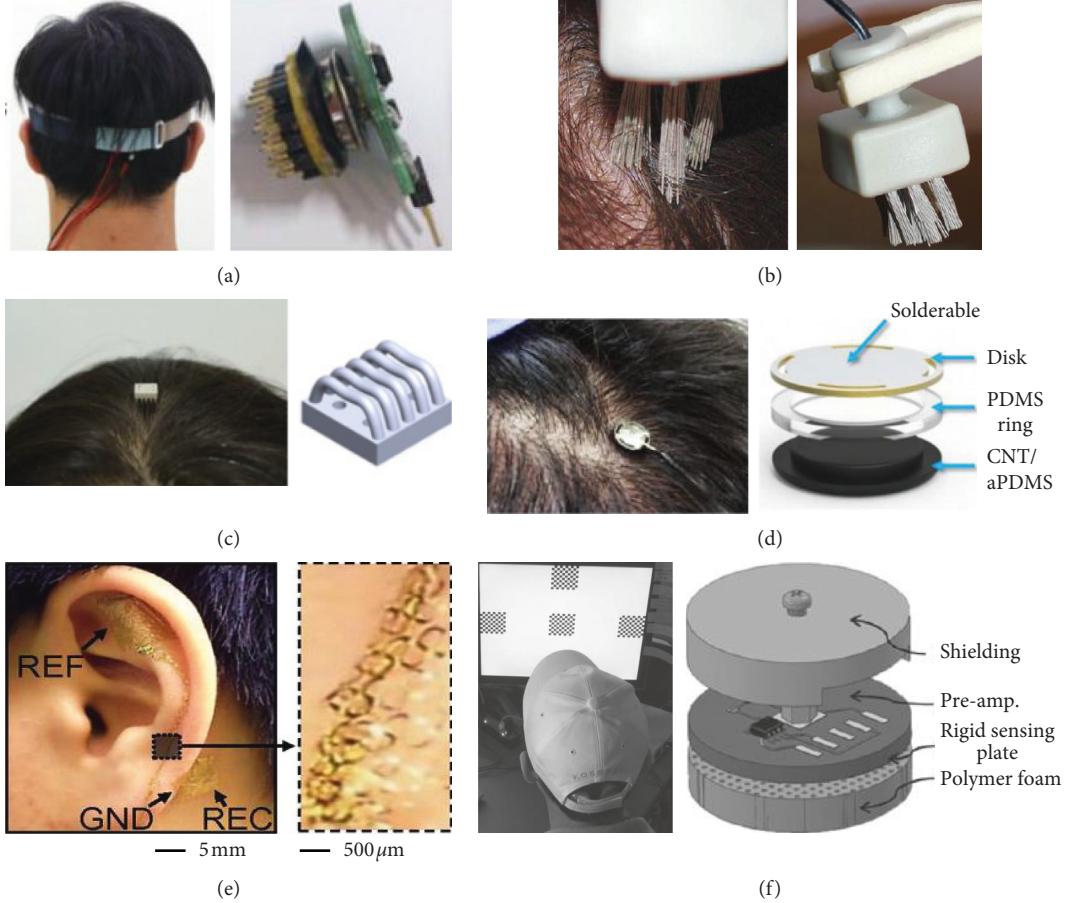


FIGURE 1: Several types of EEG electrodes: (a) active comb-shaped (electrode finger type) electrode [29], (b) bristle electrode [30], (c) reverse-curve-arch-shaped electrode [31], (d) carbon nanotube-based capacitive electrode [32], (e) epidermal electrode on the auricle [33], and (f) foam-surfaced capacitive electrode for use over hair [34, 35].

capacitive EEG electrode that combines an electrode face with polymer foam adaptive to head topography (Figure 1(f)) [34]. The rigid surface of previous conventional capacitive electrodes cannot adapt to head curvature and the hair-made irregular surface that produces hundreds of micrometer-wide air gaps between the scalp and the electrode face. The use of foam minimized the loss of electrode contact area and generated increased contact impedance. The soft foam used in Baek’s study enabled intimate electrode contact on the hairy scalp topography, thereby increasing the effective contact area. In addition, the foam-surfaced electrode maintained stable contact during motion, minimizing how much the electrode slid over the hair through its cushioning effect and textures. This electrode also showed BCI feasibility under SSVEP and ASSR [35]. The current capacitive electrode designs involve bulkier structures than the wet EEG electrodes widely used both clinically and academically. To translate laboratory nonclinical work into real-world clinical applications, studies should consider methodologies that maximize coupling capacitance while using small capacitive electrodes because the size of an EEG electrode is directly related to the spatial resolution of the EEG. Table 1 summarizes the features, strengths, and drawbacks of the dry EEG electrodes reviewed above.

New EEG electrodes will improve the state of the art and increase practicality, efficacy, and ease of use. Aesthetic perspectives also should be considered. For applications outside of hospital and laboratory environments, EEG measuring devices should not make users look strange.

3. Interface Paradigm Design

The SSVEP approach has been widely used in BCI systems because it is simple and precise about the stimulus frequency. SSVEP-based BCIs provide high ITRs with minimal user training and require fewer EEG channels than other techniques. However, they can be annoying or fatiguing for some users, which makes them impractical. Some efforts to alleviate visual fatigue have created higher-frequency SSVEP-based BCIs that use a stimulus frequency of more than 35 Hz to decrease the feeling of flickering. However, more people were unable to complete BCI tasks with high-frequency SSVEPs than with low-frequency SSVEPs [42–45]. In 2014, Chang et al. proposed amplitude-modulated (AM) visual stimuli to elicit integer and noninteger harmonic SSVEPs, including both low- and high-frequency bands [46]. As shown in Figure 2(c), the AM signal was presented as an amplitude variation in a carrier signal, as

TABLE 1: Summary of representative dry EEG electrodes.

Type	Fabrication	Flexibility	Ref.	BCI application	Strengths	Drawbacks
	Copper pin	Stiff	[29]	SSVEP	(i) Simple manufacturing procedure for mass production (ii) Small size ($d = 15$ mm) for good spatial resolution (i) Small size ($d = 15$ mm) for the good spatial resolution (ii) High level of geometric conformity between the sensor and the scalp surface	Subjects felt pain or discomfort from the pressure
Electrode finger	Spring loaded pin coated with gold	Soft	[36]	N/A		Equipment required for electrode fixation
	Flexible polymer pin	Soft	[37]	N/A	Subjects reported that they were more comfortable than the conventional EEG system	Slight erythema was found after 10–35 h, but it faded rapidly after the electrode was removed
Toothbrush	Silver-coated bristles	Soft	[30]	(i) Motor imagery (ii) P300 (iii) N100	Better comfort than wet or pin-based electrodes	(i) Requires multistep preparations (ii) Some subjects felt prickling sensation
Reverse-curve-arch-shaped	Sterling silver using 3D printer	Stiff	[31]	N/A	Maximized contact area and dispersed pressure	(i) Obtrusive wiring interfaces. (ii) Equipment required for electrode fixation
Epidermal (tattoo)	Microfabrication with polyimide	Soft	[33]	(i) SSVEP (ii) P300	(i) Soft, stretchable, and lightweight (ii) Conformal contact, adequate adhesion	(i) Cannot be applied to hairy scalp (ii) Sensor design problem for people who do not want stick-on, temporary tattoos
	CNT/aPDMS	Stiff	[32]	(i) SSVEP (ii) N100	(i) Electrode could be autonomously attached to the scalp without the need for additional equipment. (ii) Small size ($d = 6$ mm) for good spatial resolution	Require multistep preparations, obtrusive wiring interfaces
Capacitive	Custom-integrated AFE	Stiff	[40, 41]	SSVEP	No need for external G- or T-ohm for biasing network	(i) Weak for motion artifacts (ii) Poor coupling interface through dry hair
	Polymer foam surfaced	Soft	[34, 35]	(i) SSVEP (ii) ASSR	(i) EEG measurement through hair (ii) Comfortable for users (iii) Higher signal quality than with rigid capacitive electrodes	(i) Weak for motion artifacts (ii) Low spatial resolution due to relatively large size ($r = 36$ mm)

described in equation (1). Different combinations of carrier and modulation frequencies elicited different harmonic frequencies from the low- to high-frequency range, while the visual stimulus actually flickered at a high frequency. Their experiments demonstrated that AM SSVEP with an optimized combination of harmonic frequencies performed as well as a typical SSVEP. Subject evaluations indicated reduced eye fatigue and less flickering sensation. Similarly, in 2015, Dreyer and Herrmann showed frequency-modulated (FM) visual stimuli for SSVEP BCI [47]. The FM signal, simply expressed in equation (2), encodes stimulation in a carrier wave by varying the instantaneous frequency of the wave (Figure 2(d)). This contrasts with AM, which varies the amplitude of the carrier wave while the frequency remains constant. In their experiment, they used different FM stimulation combinations that all had their lower sidebands at 10 Hz, allowing them to use FM stimulation to evoke a

10 Hz SSVEP peak without the conscious perception of a 10 Hz flicker. FM-SSVEPs with different carrier and modulation frequencies can reliably be evoked with spectral peaks at the low FM sideband of 10 Hz. Subjective perceptibility ratings for flickering decreased as the FM carrier frequencies increased, while the peak amplitude and signal-to-noise ratio remained the same.

$$S_{AM} = -\frac{1}{2} [\cos(2\pi(f_c + f_m)t) + \cos(2\pi(f_c - f_m)t)], \quad (1)$$

$$S_{FM} = \sin(2\pi f_c t + M \sin(2\pi f_m t)). \quad (2)$$

Several recent studies have proposed a half-field stimulation pattern based on the brain mechanism of visual selective attention [48–51]. The user is expected to concentrate their eyes on a fixation point in the middle of two

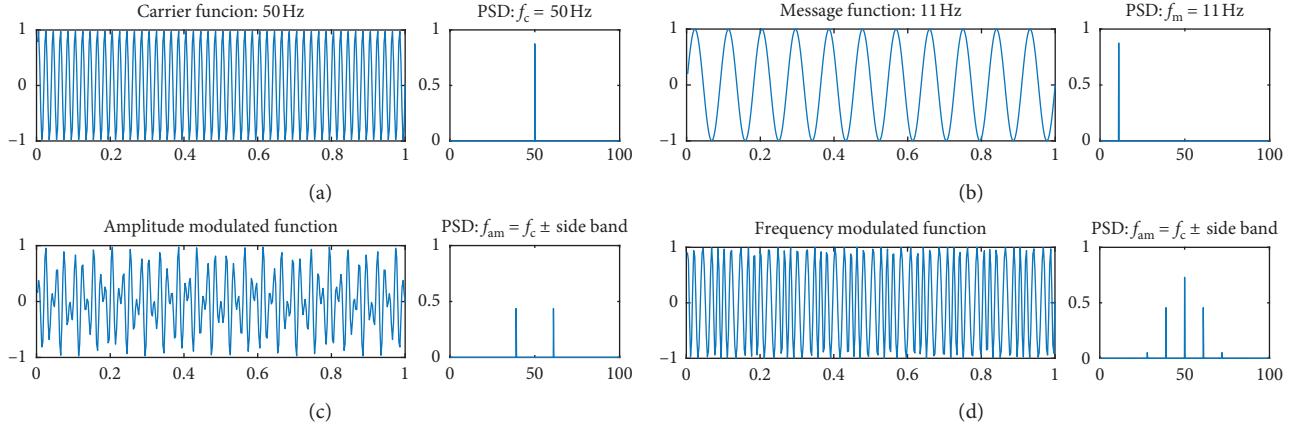


FIGURE 2: Example of time series simulation waves and their spectra: (a) a sinusoidal carrier wave at 50 Hz, (b) a message wave at 11 Hz, and (c) an amplitude-modulated signal, and (d) a frequency-modulated signal. Note that the spectra for the modulated signals show peaks at the sidebands at the carrier frequency \pm message frequency.

flickers modulated to specific frequencies. Considering the role of the optic chiasm, SSVEP was found to be strongly modulated by spatial selective attention. The two stimulus frequency components could be extracted from the contralateral occipital regions because SSVEP enlarges substantially in response to a flickering stimulus at an attended versus an unattended location. Yan et al. showed results from a multicommand, half-field SSVEP BCI. The visual display contained 9 visual targets with 18 flickers that were realized by combining 3 stimulation frequencies. Test results from 8 subjects showed an average classification accuracy of 75.8% [50]. Punsawad and Wongsawat also showed a half-field SSVEP BCI, but they used only one visual stimulus with two black boxes on both sides of the flicker to generate 4 commands by focusing to either side of the black box or on the flicker or by closing their eyes. In this study, the average classification accuracy was approximately 77% for 4 volunteers [51].

Many researchers have tried to use auditory signals instead of visual ones, especially for severely impaired users who have difficulty in controlling their voluntary extraocular movements or fixing their gaze on specific visual stimuli. As with the visual paradigm, auditory BCIs have also used unpleasant auditory stimuli that might be annoying or fatiguing to users. Research using spoken or sung syllables or polyphonic musical or even natural sounds has shown that such stimuli are perceived as more pleasant, and in some cases, they even lead to better classification performance. Lopez-Gordo et al. presented a novel fully auditory BCI based on a dichotic listening paradigm using human voices (two distinct streams of letters or sentences) as the stimulus [52]. The stimuli were read out simultaneously to the subjects for binary classification using selective attention. Prior to stimulation onset, an auditory question was read to the subject by an experimenter, followed by a beep sound indicating the beginning of the dichotic listening task. The subjects were asked to pay attention to the stimulus delivered to the left ear if the correct answer to the auditory question given before the beep sound was “yes/true” and to the

stimulus delivered to the right ear if the answer was “no/false.” The classification was established by recognition of the early component of human auditory evoked potentials, namely, the N1 and P2. Based on experimental results with 12 participants, they concluded that an auditory BCI evoked by natural speech showed promising results in terms of performance, usability, training, and cognitive effort. Höhne et al. explored spoken and sung syllables as the auditory stimuli [53]. Syllables that contained the vowels “i,” “æ,” or “o” were recorded by three different speakers and presented from the left ear only, right ear only, or both ears. This made a 3×3 matrix auditory paradigm, and a nine-class auditory BCI experiment was conducted with 9 healthy subjects, as described in Figure 3. Compared with a conventional, artificially generated monotone, the spoken or sung stimuli were expected to contain richer internal classification cues, including harmonics, pitch, and voice characteristics, as well as higher levels of variance and jitter in the auditory event-related potential (ERP) responses. The experimental results showed better classification performance when using the syllables than that when using the monotone and an increase in the subjective ergonomic ratings. Treder et al. used polyphonic music as the auditory stimulus in a multi-streamed oddball paradigm [54]. The subject was asked to shift their selective attention to one of three different musical instruments, bass, drums, or keyboard, in a musical audio clip. The attended instrument could be classified with an average accuracy of 91% among 11 participants. Heo et al. proposed a novel stimulation method to minimize auditory stress by replacing the monotone carrier with familiar music and natural sounds in an ASSR-based paradigm [55]. The sounds of a violin and a piano were used as music carriers, and a cicada singing and water streaming were used as natural sounds. The violin and water streaming sounds were amplitude-modulated with a 38 Hz message frequency, and the piano and cicada singing sounds were modulated with a 42 Hz message frequency. The experimental results with six healthy subjects demonstrated that a high ergonomic rating could be acquired while maintaining high average binary

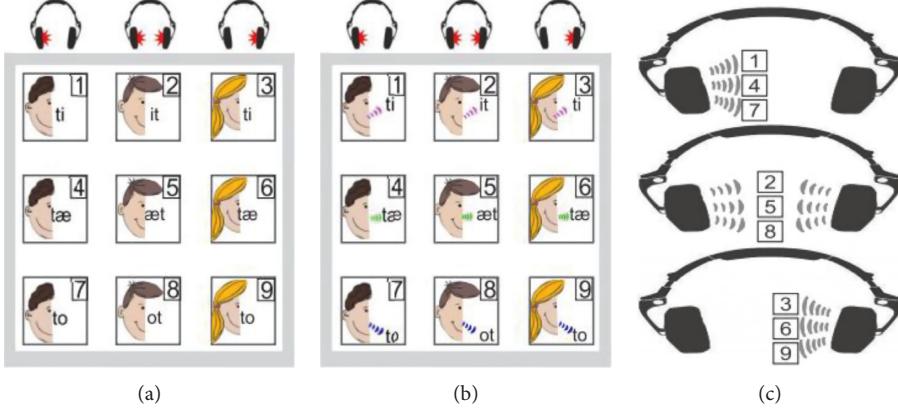


FIGURE 3: Graphical representation of the auditory stimuli set composed of spoken and sung syllables proposed by Höhne et al. [53].

classification accuracies, 74%, 89.67%, and 87.67%, for the monotone, music, and natural sound carrier waves, respectively. In conclusion, the use of pleasant sounds including a human voice or polyphonic musical or natural sounds instead of the conventional unpleasant monotone beep sound needs to be considered because the sounds can affect the level user interest and cognitive effort, even leading to improved classification performance.

4. Experiments and Results

In the current experiments, we used dry EEG electrodes [34] with BCIs using [46] SSVEP and [55] ASSR. Three healthy subjects (all males aged 26 to 30) who had no history of neurological disease and no neuropathological abnormalities agreed to carry out capacitive EEG measurements under new ABCI paradigms. The EEG data were recorded using foam-surfaced capacitively coupled electrodes at the O1 and O2 sites in a normal baseball cap for SSVEP detection and at the Oz, Cz, T7, and T8 sites for ASSR detection. All signals were recorded with a reference electrode at A2 and a grounding electrode at FpZ. The reference and grounding electrodes were not active capacitive electrodes but passive dry electrodes that did not require conduction gel or paste. Signals were measured through hair and transmitted through a hardware module composed of a high-pass filter (HPF), a low-pass filter (LPF), a 60 Hz notch filter, and an amplifier with a gain of 10,000. The HPF and LPF were used to reduce fluctuation and for antialiasing, respectively, and were designed as 4th-order Butterworth filters from 0.05 to 30 Hz. EEGs were digitized at a 512 Hz sampling rate using an analog-to-digital converter (NI-DAQ Pad 6015, National Instruments Co., TX, USA) and recorded using a Matlab data acquisition toolbox (Matlab2008b, Mathworks, Inc., Natick, MA, USA).

For the SSVEP-based ABCI application, four visual stimuli were positioned around an LCD monitor in two LED arrays (SMD 5050-3, Korea) with a diffusion film. Four targets (left, up, right, and down) flickered in an amplitude-modulated sinusoidal wave with different combinations of carrier and modulation frequency. The AM stimulus was digitally generated in eight bits at 1000 Hz using a

microcontroller unit (ATmega128, Atmel, USA) and then converted into an analog signal again to operate the LEDs using a digital-to-analog converter (LTC1657CN, Texas Instrument, USA). From equation (1), the spectrum of $S(t)$ has a peak frequency of $f_c + f_m$ and $f_c - f_m$. In this study, the f_{cs} were high frequencies (50 and 51 Hz) to reduce eye fatigue, and the f_{ms} were low frequencies near the α -band (9–12 Hz) to achieve a large SSVEP amplitude and allow high-frequency stimuli carrying low-frequency information to be generated. Figures 2(a)–2(c) provide examples of $c(t)$, $m(t)$, and $S(t)$ and their spectra. When f_c and f_m were 50 and 11 Hz, respectively, spectral peaks of $S(t)$ appeared at 39 Hz ($= (50 - 11)$ Hz) and 61 Hz ($= (50 + 11)$ Hz). All subjects performed an offline experiment first to determine the optimal EEG analysis time window size when measuring AM-SSVEP after exposure to only visual stimulation without an actual BCI application. Using programmed auditory instructions, the subjects were asked to focus on one of the four targets for 15 s. Each run contained 40 trials, and 2 runs with a 10 min break between them constituted the offline experiment. Each target was attended equally, 10 times per run. Time window sizes from 5 to 10 s with 1 s resolution were tested to investigate time-sensitive changes in the AM-SSVEP-based ABCI system performance. EEG frequency recognition under AM-SSVEP was performed using a canonical correlation analysis (CCA) to find the maximal correlation between the EEG electrode signal and signals from a matrix of templates corresponding to the AM-SSVEP stimulus frequencies, $f_c + f_m$ and $f_c - f_m$. Figure 4 shows the classification accuracies and ITR for each participant with respect to different time window sizes. Based on the trade-off relationship between the time window size/accuracy and ITR, we selected 7 or 8 s as the optimal analysis window size. Then, the online experiments were conducted using electrode-equipped caps. The task for the online experiment was a simple 2D maze in which a cursor (blue dot) could be moved in four possible directions toward a target position (red dot) (Figure 5(a)). The movement direction was determined by the SSVEP responses from focusing on one of four targets, which were oriented in the four cardinal directions. Figures 5(b)–5(d) illustrate the movement map that each participant actually performed using AM-SSVEP

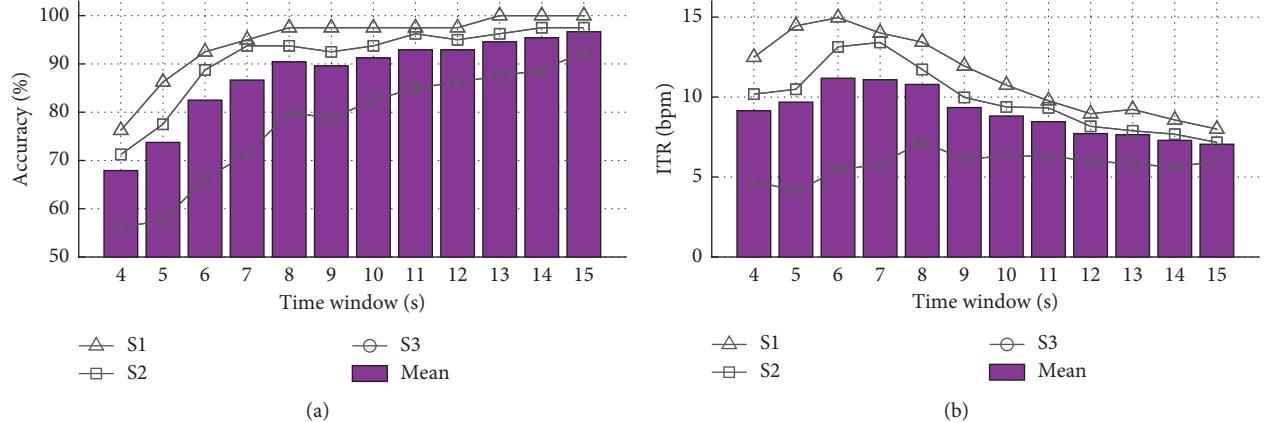


FIGURE 4: Offline AM-SSVEP analysis: (a) classification accuracy and (b) ITR for each participant with respect to different time window sizes. Bar graph indicates accuracy and ITR averaged over three subjects.

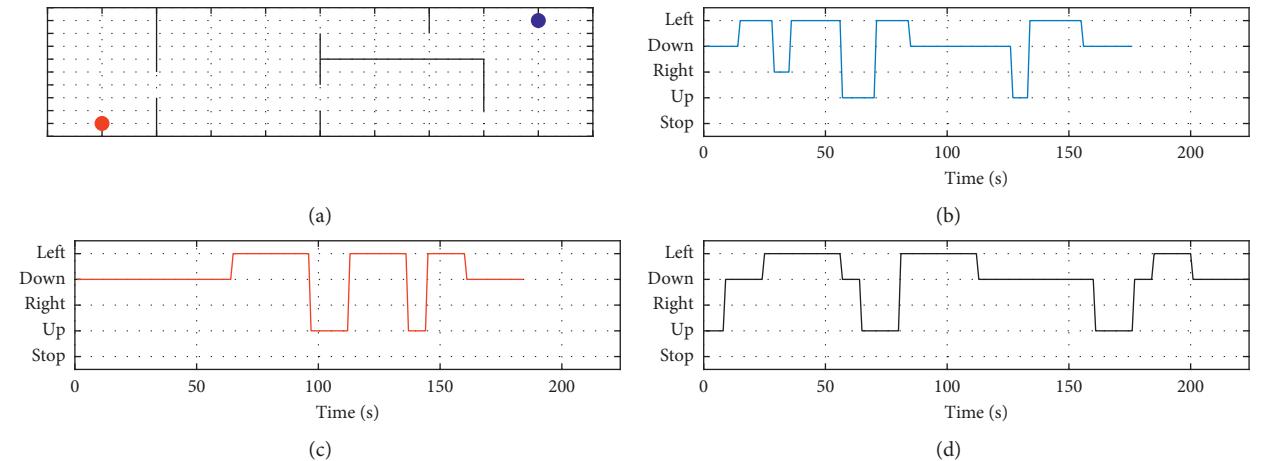


FIGURE 5: Graphical representation of online AM-SSVEP BCI task: (a) 2D maze in which a cursor (blue dot) can be moved along four possible directions to the target position (red dot); (b-d) movement map that each participant actually performed using the AM-SSVEP BCI.

BCI, and Table 2 summarizes the online experimental results for each subject. Efficiency (EFF) was defined as the minimum number of commands necessary to reach the target position divided by the number of commands issued during the run [56]. All subjects successfully completed the task, and average accuracy (ACC), EFF, and ITR were 86.07%, 79.46%, and 8.78 bits per minute, respectively. The AM stimuli were perceived as less annoying than conventional visual stimuli, but we found no other differences between the experimental conditions. The feasibility of the AM-SSVEP BCI using foam-surfaced capacitive EEG electrodes for successful BCI performance and low eye fatigue was confirmed using our current offline and online experiments. A demonstration video can be found at <https://youtu.be/YYbHM4HDTeg>.

For the ASSR-based ABCI, 37 and 42 Hz were selected as the message frequencies because the optimal modulation frequency for ASSR has been reported to be around 40 Hz. A water stream and insect sound were chosen as alternatives to the conventional pure tone burst to provide a natural and pleasant sensation. Also, subjects could easily distinguish the

different sound streams. The water stream, amplitude-modulated with a 37 Hz message frequency, was presented in the left sound field, and the insect sound, amplitude-modulated with a 42 Hz message frequency, was presented in the right sound field. Each of our three participants sat in a comfortable chair in front of a pair of commercial speakers (BR-2100S, Britz International, Paju, Korea) while wearing a cap containing the foam-surfaced capacitive electrodes. First, subjects performed an offline experiment. Following programmed visual and auditory instructions provided right before the onset of each stimulus, each participant was asked to concentrate on one of the stimuli (L or R) for 20 s. This process was repeated 50 times, and ten-fold cross validation was applied to compare our results with the performance reported in [35], which used the same EEG system but a pure-tone sinusoidal carrier sound for an ASSR BCI. We calculated the frequency spectra using a nonparametric periodogram method with a 1 s sliding time window and 50% overlap. The spectral density of each electrode over the stimulus message frequency ± 1 Hz range was extracted from the averaged frequency spectra and

TABLE 2: Results of the AM-SSVEP-based online BCI experiments (ACC: accuracy in %, ITR: information transfer rate in $\text{bit} \cdot \text{min}^{-1}$, LPM: letters per minute in letter min^{-1} , and EFF: efficiency in %).

Sub.	Time window	Output path (error underlined)	ACC (%)	EFF (%)	ITR (bits/min)
S1	7 s	DDDDDDDDLL <u>L</u> UULLLULLDDD	84	80	8.01
S2	8 s	DDLL <u>R</u> LLL <u>U</u> ULLLD <u>DDDD</u> DULLDDD	95.65	86.96	12
S3	8 s	<u>U</u> DD <u>LL</u> <u>LD</u> U <u>LL</u> LL <u>DDDD</u> D <u>U</u> UDLLDDD	78.57	71.43	6.34
Mean			86.07	79.46	8.78

fed into a linear discriminant analysis classifier as a feature vector. Classification accuracy and ITR with respect to the time window size are presented in Figure 6. The average value (bold line) showed a pattern similar to that in [35]: accuracy increased linearly with respect to window size. The online experiments were performed after allowing the subjects to have a brief rest. During the resting time, participants took off the electrode cap briefly, and then they put it back on right before starting the online experiment. An analysis window size of 14 s, derived from [35], was used in the online experiments. Ten trials of selective attention to either the left or right stimulus were performed during the online experiments, and the results are shown in Table 3: the number of correct decisions (NUM) 8/10, specificity (SPEC) 0.82, sensitivity (SENS) 0.79, and ITR 1.33 bits per minute. SPEC and SENS were calculated by assuming positive to be left (L) and negative to be right (R). These results are comparable to previously reported results from a study of the same EEG system and ASSR-based technique using conventional pure tone carrier sounds (NUM 7.2/10, SPEC 0.64, SENS 0.76, ITR 0.7 bits per minute). A video demonstration can be found at https://youtu.be/uPF_MjNEefA.

5. Discussion and Conclusion

This paper has focused on the challenges faced when moving from BCI systems designed for experimental use in laboratory settings to those intended for use in real-world environments. We have discussed the problems with EEG sensing technologies and new BCI paradigms and explored representative methods for handling laboratory, more realistic, or real-world settings.

We also demonstrated results from an ABCI system that uses foam-surfaced capacitive EEG electrodes with the AM-SSVEP and natural ASSR paradigms. In our AM-SSVEP experiments, we found that classification accuracy was increased with the analysis time window size. However, compared with Chang's result derived using conventional Ag/AgCl electrodes [46], an increased time window was required to sufficiently reject the extra noise seen with capacitive measurements. EEG signals longer than 4 s were suitable for reliable AM-SSVEP BCIs in Chang's study, whereas 8 s were required for the capacitive AM-SSVEP BCI in this study. In our experience, a time window of 8 s is the longest used to determine a single command by SSVEP response. That long time window is required to capacitively measure AM-SSVEP because of the low signal-to-noise ratio. A time window of 6 s was enough for a successful conventional SSVEP-based spelling task with capacitive EEG measurement performed by Baek et al. in 2013 [35]. For online

BCI, a direct comparison between the online BCI performance in the current study and Chang et al.'s result [46] cannot be made due to the duration of processing, the classified command, and our use of a different task with a different number of possible selections. Chang et al. stored EEG data in a 4 s data buffer every 0.5 s. The existing data were shifted, removing the initial 0.5 s of data to generate a new 4 s segment. Then, the AM-SSVEP was recognized within the 4 s EEG signal by using the CCA method every 0.5 s. If four consecutive temporal decisions were the same, the corresponding decision was selected as the final decision. In addition, a customized frequency was combined with the best performance of each subject to create a CCA reference signal for frequency recognition. In our experiment, decisions were made at every fixed time window without any data shifting or window sliding. Also, the frequency of the reference signal in the CCA was not optimized for each subject. We chose those parameters for convenience in the experimental setting. Nevertheless, all the participants succeeded in carrying out the cursor navigation task with relative ease. All of the subjects indicated that they preferred working with the AM-SSVEP approach despite the lower accuracy rates and reduced speed of operation because the flickering was less tiring and required less effort from the eyes. The ASSR-based BCI paradigm was used in 2011 by Kim et al. with conventional EEG electrodes and auditory stimuli generated using periodic amplitude-modulated and pure sinusoidal tones [57]. They implemented a pilot online ASSR-based BCI and tested it with one subject. Their experimental result showed a classification accuracy of 71.4%. Our previous study, performed in 2013 by Baek et al., used the same auditory stimulation but foam-surfaced capacitive electrodes instead of gel-based electrodes for the EEG sensing; we had an average accuracy of 72% for 5 subjects [35]. Compared with previous studies, we found a fair average classification accuracy of 80% for 3 participants under ABCI conditions in this study, and we acquired the EEG signals over hair using foam-surfaced capacitive electrodes and natural sound.

Although the number of subjects in the current demonstrations is too small to be representative of the general population, our results are in general agreement with published values for healthy adults, and we suggest that our demonstration offers sufficient power to assess the feasibility of ABCI applications. However, our results represent just a small sample of the broad future potential of ABCI technologies. As BCIs become more popular with different user groups, including healthy people, their increasing commercial possibilities will likely encourage new applied research efforts that will make BCIs even more practical. Consumer demand for reduced cost, increased performance,

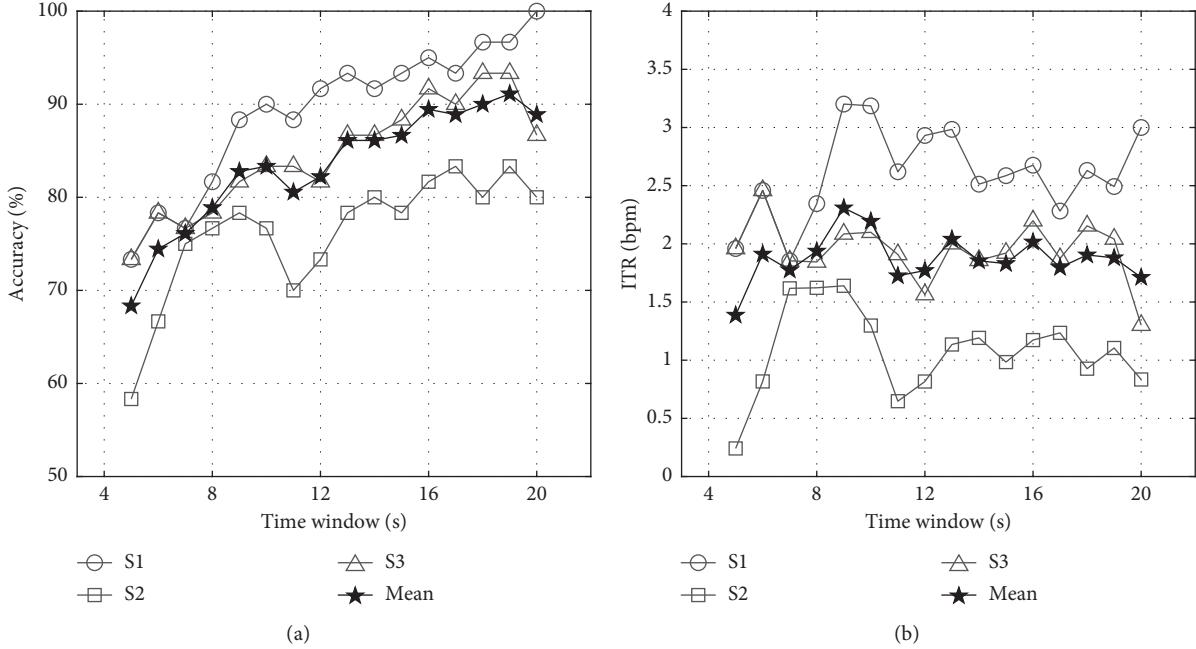


FIGURE 6: Offline, natural-sound ASSR analysis: (a) classification accuracy and (b) ITR for each participant with respect to different time window sizes. The thick line indicates accuracy and ITR averaged over three subjects.

TABLE 3: Results from natural-sound ASSR-based online BCI experiments (NUM: number of correct classifications per total number of trials, SEPC: specificity, SENS: sensitivity, and ITR: information transfer rate in $\text{bit} \cdot \text{min}^{-1}$). Inaccurate classifications are underlined.

Sub.	Time window	Task	Classification results	NUM (correct/total)	SPEC	SENS	ITR (bits/min)
S1	14 s	LLRLRLRLRR	LLRR <u>RLRLRR</u>	9/10	1.0	0.83	2.28
S2	14 s	LRRLLRLRLL	<u>RRLRL</u> LLRLR <u>L</u>	7/10	0.67	0.75	0.51
S3	14 s	RRRLLRLRLR	RRR <u>LLRLRLR</u> L	8/10	0.8	0.8	1.19
Mean				8/10	0.82	0.79	1.33

and greater flexibility and robustness could contribute substantially to making BCIs mainstream tools. The development of ABCIs requires clear validation of their real-life value in terms of efficacy, practicality, and impact on quality of life. Future BCI systems should (1) be comfortable, convenient, and offer aesthetically acceptable mountings, (2) be easy to set up, (3) function for many hours without maintenance, (4) perform well in all environments, (5) operate by telemetry instead of requiring wiring, and (6) interface easily with a wide range of applications. Before the results of ongoing and planned research efforts for ABCI become available, BCIs using various methods remain a fascinating research toy. If the intensive research into various aspects of ABCI continues to increase exponentially, as it has done recently, BCI systems could become routine clinical, assistive, and commercial tools in the not-too-distant future.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Soonchunhyang University Research Fund.

References

- [1] N. Birbaumer, N. Ghanayim, T. Hinterberger et al., “A spelling device for the paralysed,” *Nature*, vol. 398, no. 6725, pp. 297–298, 1999.
- [2] N. Birbaumer, A. Kubler, N. Ghanayim et al., “The thought translation device (TTD) for completely paralyzed patients,” *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 2, pp. 190–193, 2000.
- [3] L. R. Hochberg, M. D. Serruya, G. M. Friehs et al., “Neuronal ensemble control of prosthetic devices by a human with tetraplegia,” *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.
- [4] E. Buch, C. Weber, L. G. Cohen et al., “Think to move: a neuromagnetic brain-computer interface (BCI) system for chronic stroke,” *Stroke*, vol. 39, no. 3, pp. 910–917, 2008.
- [5] E. W. Sellers and E. Donchin, “A P300-based brain-computer interface: initial tests by ALS patients,” *Clinical Neurophysiology*, vol. 117, no. 3, pp. 538–548, 2006.
- [6] F. Nijboer, E. W. Sellers, J. Mellinger et al., “A P300-based brain-computer interface for people with amyotrophic lateral sclerosis,” *Clinical Neurophysiology*, vol. 119, no. 8, pp. 1909–1916, 2008.
- [7] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper, “Brain oscillations control hand orthosis in a tetraplegic,” *Neuroscience Letters*, vol. 292, no. 3, pp. 211–214, 2000.

- [8] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 79642, 8 pages, 2007.
- [9] A. K. Keng and G. Cuntai, "Brain-computer interface in stroke rehabilitation," *Journal of Computing Science and Engineering*, vol. 7, no. 2, pp. 139–146, 2013.
- [10] G. Cheron, M. Duvinage, C. De Saedeleer et al., "From spinal central pattern generators to cortical network: integrated BCI for walking rehabilitation," *Neural Plasticity*, vol. 2012, Article ID 375148, 13 pages, 2012.
- [11] J. Luauté, D. Morlet, and J. Mattout, "BCI in patients with disorders of consciousness: clinical perspectives," *Annals of Physical and Rehabilitation Medicine*, vol. 58, no. 1, pp. 29–34, 2015.
- [12] Y. Liu, M. Li, H. Zhang et al., "A tensor-based scheme for stroke patients' motor imagery EEG analysis in BCI-FES rehabilitation training," *Journal of Neuroscience Methods*, vol. 222, pp. 238–249, 2014.
- [13] P. Aricó, G. Borghini, G. D. Flumeri, N. Sciaraffa, and F. Babiloni, "Passive BCI beyond the lab: current trends and future directions," *Physiological Measurement*, vol. 39, no. 8, article 08TR02, 2018.
- [14] T. O. Zander and C. Kothe, "Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general," *Journal of Neural Engineering*, vol. 8, no. 2, article 025005, 2011.
- [15] P. Arico, G. Borghini, G. Di Flumeri, N. Sciaraffa, A. Colosimo, and F. Babiloni, "Passive BCI in operational environments: insights, recent advances, and future trends," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 7, pp. 1431–1436, 2017.
- [16] A. Searle and L. Kirkup, "A direct comparison of wet, dry and insulating bioelectric recording electrodes," *Physiological Measurement*, vol. 21, no. 2, pp. 271–283, 2000.
- [17] N. Meziane, J. G. Webster, M. Attari, and A. J. Nimunkar, "Dry electrodes for electrocardiography," *Physiological Measurement*, vol. 34, no. 9, pp. R47–R69, 2013.
- [18] J. D. Slater, G. P. Kalamangalam, and O. Hope, "Quality assessment of electroencephalography obtained from a "dry electrode" system," *Journal of Neuroscience Methods*, vol. 208, no. 2, pp. 134–137, 2012.
- [19] G. D. Flumeri, P. Aricó, G. Borghini, N. Sciaraffa, A. D. Florio, and F. Babiloni, "The dry revolution: evaluation of three different EEG dry electrode types in terms of signal spectral features, mental state classification and usability," *Sensors*, vol. 19, no. 6, p. 1365, 2019.
- [20] C. Guger, G. Krausz, B. Z. Allison, and G. Edlinger, "Comparison of dry and gel based electrodes for P300 brain-computer interfaces," *Frontiers in Neuroscience*, vol. 7, 2012.
- [21] J. Minguillon, M. A. Lopez-Gordo, and F. Pelayo, "Trends in EEG-BCI for daily-life: requirements for artifact removal," *Biomedical Signal Processing and Control*, vol. 31, pp. 407–418, 2017.
- [22] D. J. McFarland and T. M. Vaughan, "BCI in practice," in *Progress in Brain Research*, vol. 228, pp. 389–404, Elsevier, Amsterdam, Netherlands, 2016.
- [23] H.-J. Hwang, S. Kim, S. Choi, and C.-H. Im, "EEG-based brain-computer interfaces: a thorough literature survey," *International Journal of Human-Computer Interaction*, vol. 29, no. 12, pp. 814–826, 2013.
- [24] L.-D. Liao, C.-T. Lin, K. McDowell et al., "Biosensor technologies for augmented brain-computer interfaces in the next decades," *Proceedings of the IEEE*, vol. 100, pp. 1553–1566, 2012.
- [25] P. Griss, P. Enoksson, H. K. Tolvanen-Laakso, P. Merilainen, S. Ollmar, and G. Stemme, "Micromachined electrodes for biopotential measurements," *Journal of Microelectromechanical Systems*, vol. 10, no. 1, pp. 10–16, 2001.
- [26] J.-C. Chiou, L.-W. Ko, C.-T. Lin et al., "Using novel MEMS EEG sensors in detecting drowsiness application," in *Proceedings of the 2006 IEEE Biomedical Circuits and Systems Conference*, pp. 33–36, London, UK, November 2006.
- [27] G. Ruffini, S. Dunne, L. Fuentemilla et al., "First human trials of a dry electrophysiology sensor using a carbon nanotube array interface," *Sensors and Actuators A: Physical*, vol. 144, no. 2, pp. 275–279, 2008.
- [28] M. Arai, Y. Nushinaka, and N. Miki, "Electroencephalogram measurement using polymer-based dry microneedle electrode," *Japanese Journal of Applied Physics*, vol. 54, no. 6, article 06FP14, 2015.
- [29] Y.-J. Huang, C.-Y. Wu, A. M.-K. Wong, and B.-S. Lin, "Novel active comb-shaped dry electrode for EEG measurement in hairy site," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 1, pp. 256–263, 2015.
- [30] C. Grozea, C. D. Voinescu, and S. Fazli, "Bristle-sensors-low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications," *Journal of Neural Engineering*, vol. 8, no. 2, article 025008, 2011.
- [31] J. S. Lee, K. S. Park, J. H. Kim, and C. M. Han, "Reverse-curve-arch-shaped dry EEG electrode for increased skin-electrode contact area on hairy scalps," *Electronics Letters*, vol. 51, no. 21, pp. 1643–1645, 2015.
- [32] S. M. Lee, J. H. Kim, C. Park et al., "Self-adhesive and capacitive carbon nanotube-based electrode to record electroencephalograph signals from the hairy scalp," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 1, pp. 138–147, 2016.
- [33] J. J. S. Norton, D. S. Lee, J. W. Lee et al., "Soft, curved electrode systems capable of integration on the auricle as a persistent brain-computer interface," *Proceedings of the National Academy of Sciences*, vol. 112, no. 13, pp. 3920–3925, 2015.
- [34] H. J. Baek, H. J. Lee, Y. G. Lim, and K. S. Park, "Conductive polymer foam surface improves the performance of a capacitive EEG electrode," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 12, pp. 3422–3431, 2012.
- [35] H. J. Baek, H. S. Kim, J. Heo, Y. G. Lim, and K. S. Park, "Brain-computer interfaces using capacitive measurement of visual or auditory steady-state responses," *Journal of Neural Engineering*, vol. 10, no. 2, article 024001, 2013.
- [36] L.-D. Liao, S.-L. Wu, C.-H. Liou et al., "A novel 16-channel wireless system for electroencephalography measurements with dry spring-loaded sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 6, pp. 1545–1555, 2014.
- [37] Y.-H. Chen, M. de Beeck, L. Vanderheyden et al., "Soft, comfortable polymer dry electrodes for high quality ECG and EEG recording," *Sensors*, vol. 14, no. 12, pp. 23758–23780, 2014.
- [38] D.-H. Kim, N. Lu, R. Ma et al., "Epidermal electronics," *Science*, vol. 333, no. 6044, pp. 838–843, 2011.
- [39] W. H. Yeo, Y. S. Kim, J. Lee et al., "Multifunctional epidermal electronics printed directly onto the skin," *Advanced Materials*, vol. 25, no. 20, pp. 2773–2778, 2013.
- [40] Y. M. Chi, C. Maier, and G. Cauwenberghs, "Ultra-high input impedance, low noise integrated amplifier for noncontact

- biopotential sensing," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 526–535, 2011.
- [41] Y. M. Chi, Y.-T. Wang, Y. Wang, C. Maier, T.-P. Jung, and G. Cauwenberghs, "Dry and noncontact EEG sensors for mobile brain-computer interfaces," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 2, pp. 228–235, 2012.
- [42] G. G. Molina and V. Mihajlovic, "Spatial filters to detect steady-state visual evoked potentials elicited by high frequency stimulation: BCI application," *Biomedizinische Technik/Biomedical Engineering*, vol. 55, no. 3, pp. 173–182, 2010.
- [43] S. M. T. Muller, P. F. Diez, T. F. Bastos-Filho, M. Sarcinelli-Filho, V. Mut, and E. Laciar, "SSVEP-BCI implementation for 37–40 Hz frequency range," in *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6352–6355, Boston, MA, USA, August 2011.
- [44] I. Volosyak, D. Valbuena, T. Luth, T. Malechka, and A. Graser, "BCI demographics II: how many (and what kinds of) people can use a high-frequency SSVEP BCI?," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 3, pp. 232–239, 2011.
- [45] T. Sakurada, T. Kawase, T. Komatsu, and K. Kansaku, "Use of high-frequency visual stimuli above the critical flicker frequency in a SSVEP-based BMI," *Clinical Neurophysiology*, vol. 126, no. 10, pp. 1972–1978, 2015.
- [46] M. H. Chang, H. J. Baek, S. M. Lee, and K. S. Park, "An amplitude-modulated visual stimulation for reducing eye fatigue in SSVEP-based brain-computer interfaces," *Clinical Neurophysiology*, vol. 125, no. 7, pp. 1380–1391, 2014.
- [47] A. M. Dreyer and C. S. Herrmann, "Frequency-modulated steady-state visual evoked potentials: a new stimulation method for brain-computer interfaces," *Journal of Neuroscience Methods*, vol. 241, pp. 1–9, 2015.
- [48] S. T. Morgan, J. C. Hansen, and S. A. Hillyard, "Selective attention to stimulus location modulates the steady-state visual evoked potential," *Proceedings of the National Academy of Sciences*, vol. 93, no. 10, pp. 4770–4774, 1996.
- [49] A. Materka and M. Byczuk, "Alternate half-field stimulation technique for SSVEP-based brain-computer interfaces," *Electronics Letters*, vol. 42, no. 6, pp. 321–322, 2006.
- [50] Z. Yan, X. Gao, G. Bin, and B. Hone, "A half-field stimulation pattern for SSVEP-based brain-computer interface," in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6361–6464, Minneapolis, MN, USA, September 2009.
- [51] Y. Punyawad and Y. Wongsawat, "Multi-command SSVEP-based BCI system via single flickering frequency half-field stimulation pattern," in *Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1101–1104, Boston, MA, USA, August 2011.
- [52] M. A. Lopez-Gordo, E. Fernandez, S. Romero, F. Pelayo, and A. Prieto, "An auditory brain-computer interface evoked by natural speech," *Journal of Neural Engineering*, vol. 9, no. 3, article 036013, 2012.
- [53] J. Höhne, K. Krenzlin, S. Dähne, and M. Tangermann, "Natural stimuli improve auditory BCIs with respect to ergonomics and performance," *Journal of Neural Engineering*, vol. 9, no. 4, article 045003, 2012.
- [54] M. S. Treder, H. Purwins, D. Miklody, I. Sturm, and B. Blankertz, "Decoding auditory attention to instruments in polyphonic music using single-trial EEG classification," *Journal of Neural Engineering*, vol. 11, no. 2, article 026009, 2014.
- [55] J. Heo, H. J. Baek, S. Hong, M. H. Chang, J. S. Lee, and K. S. Park, "Music and natural sounds in an auditory steady-state response based brain-computer interface to increase user acceptance," *Computers in Biology and Medicine*, vol. 84, pp. 45–52, 2017.
- [56] B. Allison, T. Luth, D. Valbuena, A. Teymourian, I. Volosyak, and A. Graser, "BCI demographics: how many (and what kinds of) people can use an SSVEP BCI?," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 107–116, 2010.
- [57] D.-W. Kim, H.-J. Hwang, J.-H. Lim, Y.-H. Lee, K.-Y. Jung, and C.-H. Im, "Classification of selective attention to auditory stimuli: toward vision-free brain-computer interfacing," *Journal of Neuroscience Methods*, vol. 197, no. 1, pp. 180–185, 2011.

Research Article

Application of Deep Learning in Neuroradiology: Brain Haemorrhage Classification Using Transfer Learning

Awwal Muhammad Dawud , Kamil Yurtkan , and Huseyin Oztoprak

Department of Computer Engineering, Cyprus International University, Nicosia, Cyprus

Correspondence should be addressed to Awwal Muhammad Dawud; awwal125@gmail.com

Received 17 March 2019; Accepted 30 April 2019; Published 3 June 2019

Guest Editor: Hyun S. Kim

Copyright © 2019 Awwal Muhammad Dawud et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we address the problem of identifying brain haemorrhage which is considered as a tedious task for radiologists, especially in the early stages of the haemorrhage. The problem is solved using a deep learning approach where a convolutional neural network (CNN), the well-known AlexNet neural network, and also a modified novel version of AlexNet with support vector machine (AlexNet-SVM) classifier are trained to classify the brain computer tomography (CT) images into haemorrhage or nonhaemorrhage images. The aim of employing the deep learning model is to address the primary question in medical image analysis and classification: can a sufficient fine-tuning of a pretrained model (transfer learning) eliminate the need of building a CNN from scratch? Moreover, this study also aims to investigate the advantages of using SVM as a classifier instead of a three-layer neural network. We apply the same classification task to three deep networks; one is created from scratch, another is a pretrained model that was fine-tuned to the brain CT haemorrhage classification task, and our modified novel AlexNet model which uses the SVM classifier. The three networks were trained using the same number of brain CT images available. The experiments show that the transfer of knowledge from natural images to medical images classification is possible. In addition, our results proved that the proposed modified pretrained model “AlexNet-SVM” can outperform a convolutional neural network created from scratch and the original AlexNet in identifying the brain haemorrhage.

1. Introduction

Intracranial haemorrhage (ICH) reveals as a bleeding within the intracranial vault [1]. Weak blood vessels, hypertension, trauma, and drug abuse are generally what trigger such a medical condition. ICH is a neurologic emergency in which it can have several subtypes such as basal ganglia, caudate nucleus, or pons. The types of haemorrhage are generally dependent on the anatomic location of bleeding [2]. According to the American Heart Association and American Stroke Association, the early and timely diagnosis of ICH is significant as this condition can commonly deteriorate the affected patients within the first few hours after occurrence [3]. Noncontrast head computer tomography (CT) is the imaging modality used to detect haemorrhage due its wide availability and speed. This modality has shown a high sensitivity and specificity in detecting acute haemorrhage [2].

Recently, deep learning has risen rapidly and effectively. Deep learning-based networks have shown a great generalization capability when applied to solve challenging medical problems such as medical image classification [4, 5], medical image analysis [6], medical organs detection [7], and disease detection [8]. Convolutional neural networks were the most effective networks among deep networks, for they own the paradigms of more biologically inspired structures than other traditional networks [9].

Eventually, various convolutional neural networks were developed such as AlexNet [10], VGG-NET [11], and ResNet [12]; these deep networks are all extensively trained on a large database named ImageNet, Large-Scale Visual Recognition Challenge [13], and they were considered as the state of the art in image classification [11–13]. These networks are considered as machine learning methods that can learn features hierarchically

from lower level to higher level by building a deep architecture of the input data.

The rise in deep convolutional neural networks performance, due to their abstractions of different levels of features, motivated many researchers to transfer the knowledge acquired by these networks, when trained on millions of images into new tasks such as medical image classification, to benefit from their learned parameters, in particular, weights.

These convolutional neural networks models use fully connected layers, which represent a feedforward neural network trained using the conventional backpropagation algorithm. This means that these models may have the same drawbacks of the conventional simple neural network.

An effective neural network model is the one that performs well during both training and testing datasets; a good balance between variance error and bias error must be struck [14]. For simple models, a high bias and a low variance situation reveals when training these models; that is called underfitting. For more complex neural network models, the progress of training may let the model enter a region of low variance and bias; this can be considered as a good fit. However, as the training progresses further (more complex models), the model may go through a high variance and low bias, that is called overfitting. This is considered a major problem in training a complex neural network model.

There are many approaches for alleviating this problem [15]. These approaches include early stopping, weights penalization, weights pretraining, and dropout of hidden neurons. However, in our study, we ought to avoid these problems by replacing the SoftMax neural network with a multiclass SVM that acts as a classifier for both pretrained employed models. There have been many conducted studies [16–18] that attempt to find an alternative to SoftMax function for classification tasks. All these studies concluded that the support vector machine (SVM) might be the appropriate alternative as it may slightly boost the performance of neural network compared to the conventional SoftMax function.

Thus, in this paper, we aim to transfer the knowledge acquired by AlexNet into a new target task: classifying the CT brain haemorrhage into haemorrhage or non-haemorrhage images. Moreover, a CNN is created from scratch and a modified AlexNet combined with SVM are also employed to perform the same classification task. The goal of employing one CNN created from scratch and fine-tuning a pretrained model for the same classification task is to show that transfer learning-based network can perform better when data are not much. Also, it is aimed to show that sufficient fine-tuning of a pretrained model can eliminate the need for training a deep CNN from scratch which usually takes long time and requires large number of images to learn. Note that in this research, the CNN created from scratch is denoted as CNN, the pretrained model that uses original AlexNet architecture is denoted as AlexNet, and the modified model is denoted as AlexNet-SVM.

The paper is structured as follows: Section 1 is an introduction of the work. Section 3 is a brief explanation of the convolutional neural networks basics, while Section 4 explains the transfer learning concept including AlexNet.

Section 5.3 discusses the training of the two employed deep networks in which the data used for training are described. Section 6 discusses the networks performances and compares the results of both models. Finally, Section 8 is conclusion of the paper.

2. Related Work

Convolutional neural networks have been employed to overcome big medical challenges like image segmentation [19] and control for people with disabilities [20]. Hussain et al. [19] have developed a convolutional neural network designed for the segmentation of the most common brain tumor, i.e., glioma tumor. The authors proposed a system composed of two networks, stacked together to form a new ILinear nexus architecture. This new architecture was capable of achieving the best results among all the proposed and related architectures. Another study by Abiyev and Arslan [20] showed that convolutional neural networks can also be used as supporting elements for people with disabilities. The authors proposed a human-machine interface based on two convolutional neural networks designed for disabled people with spinal cord, to control mouse by eye movements. Their work was validated and tested by a handcrafted dataset, and results showed that the network's performance outscored many other related works.

Furthermore, deep learning techniques were employed by Helwan et al. [21] to classify brain computer tomography (CT) images into haemorrhage or healthy. The authors used autoencoders and deep convolutional neural networks to perform this task. As authors claimed, the employed models performed differently when trained and tested on 2527 images. It was found that the stacked autoencoder used in their paper consists of three hidden layers and outperformed other employed networks, where it achieved the highest classification rate and the lowest MSE. The authors concluded that the possible reason of this outperformance on the stacked autoencoder over convolutional neural network is due to the small number of data used for training, as a CNN needs large amount of training examples in order to converge.

In another study by Mahajan and Mahajan [22], brain haemorrhage was examined in more refined manner by feeding using the watershed algorithm along with artificial neural network (ANN) for CT identification of brain haemorrhage type. The authors of this work used features extraction before feeding images to the neural classifier, in which different features were extracted using grey-level co-occurrence matrix (GLCM). Features were then classified by a conventional backpropagation neural network used to identify the type of haemorrhage. They found that adequate image processing techniques such as noise removal and high segmentation methods are required for accurate identification of haemorrhage.

Furthermore, Gong et al. [23] focused on dividing brain CT images into regions, where each region could either be normal or haemorrhage. For images containing haemorrhage, the regions which did not include haemorrhage were treated as normal regions resulting in a highly imbalanced

dataset. The researcher had utilized an image segmentation scheme that used ellipse fitting, background removal, and wavelet decomposition technique. The weighted precision and recall value for this approach were approximately 83.6% and 88.5%, respectively.

3. Convolutional Neural Network

Convolutional neural network (CNN) is a well-employed network for several tasks in machine vision and medicine [24, 25]. Generally, the CNN relies on architectural features which include the receptive field, weight sharing, and pooling operation to take into account the 2D characteristic of structured data such as images [26]. The concept of weight sharing for convolution maps drastically reduces model parameters; this has the important implications that the model is less prone to overfitting as compared to fully connected models of comparable size. The pooling operation essentially reduces the spatial dimension of input maps and allows the CNN to learn some invariance to moderate distortions in the training; this feature enhances the generalization of the CNN at test time as the model is more tolerant to moderate distortion in the test data [27]. The typical CNN is shown in Figure 1. Essentially, convolution layers, pooling layers, and the fully connected layers are shown. For example, layer 1 employs n convolution filters of size $a \times a$ to generate a bank of n convolution maps (C_1) of size $i \times i$; this is followed by a pooling (subsampling) operation on the convolution maps with a window size of $b \times b$. Therefore, the pooling layer (S_1) composes n feature maps of size $j \times j$, where, $j = i/b$ [25]. The convolution layer performs feature extraction on the incoming inputs via a convolution filter of specified size. The pooling operation pools features across input maps using a window of specified size; common pooling operations used in applications are the average and max pooling [28]. In average pooling, the average value of the inputs captured by the pooling window is taken, while, in max pooling, the maximum value of the inputs captured by the pooling window is taken. For learning the classifier model, features are forward-propagated through the network to the fully connected layer with an output layer of units. Then, the backpropagation learning algorithm can be employed to update the model parameters via the gradient descent update rule [29].

4. Transfer Learning

In medical image analysis and processing, a most common issue is that the number of available data for research purposes is limited and small. Hence, training a fully deep network structure like CNN with small number of data may result in overfitting, which is usually the reason of low performance and generalization power [30]. Transfer learning is a solution to this problem where the learned parameters of effective and well-trained networks on a very large dataset are shared. The concept of transfer learning is the use of a pretrained model that is already trained on large datasets and transfers its pretrained learning parameters, in particular weights, to the targeted network model. To be able

to use the network for another problem, the last fully connected layers are then trained with initial random weights on the new dataset. Although the dataset is different than the one that the network was trained on, the low-level features are similar. Thus, the parameters' transfer of the pretrained model may provide the new target model with a powerful feature extraction capability and reduce its training computations and memory cost. Transfer learning has been used extensively in medical imaging, and it showed a great efficacy in terms of accuracy, training time, and error rates [10, 31, 32]. In this paper, we present a modified pretrained model, AlexNet, that has been employed for the classification of CT brain haemorrhage images into normal and abnormal classes.

4.1. AlexNet. AlexNet is the first convolutional neural network that achieved the highest classification accuracy at the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 [10]. This deep structure is comprised of eight main layers; the first five layers are mainly convolutions, while the last three are fully connected layers. Each convolutional layer is followed by an activation function layer, i.e., rectified linear units layer (ReLU), proposed to improve the performance of the network by making the training faster than equivalents of "tanh" activation functions [10]. After each convolution layer, a max pooling is used in AlexNet, in order to reduce the network size. Moreover, a dropout layer is added after the first two fully connected layer which helps to reduce the number of neurons and prevent overfitting [33]. Finally, a layer is added after the last layer to classify the input given data. Figure 1 shows the structure of the AlexNet.

5. Materials and Methods

This work addresses the problem of the classification of the CT brain images into normal or haemorrhage, which can be a hard task for some junior radiologists and doctors. The problem is addressed by the implementation of a deep learning network trained extensively to acquire the power of extracting low to high levels of features from normal brain CT images and others with haemorrhage medical conditions using its designed and trained filters. These features are then what distinguishes the class of the brain images, i.e., haemorrhage or not. Nonetheless, the transfer of knowledge from original to target task, which is here Haemorrhage identification, is also considered by transferring the knowledge of a pretrained model known as AlexNet, into a new classification task and testing it by the same number of images used for testing the CNN created from scratch. In this manner, we aim to address the central issue in medical image analysis and diagnosis: training deep CNN from scratch is not needed; instead, use a pretrained modified AlexNet by adding SVM classifier to transfer its knowledge to a new target task with sufficient fine-tuning. Our conducted experiment on the CT brain haemorrhage classification using a CNN created from scratch and the pretrained models will demonstrate the truth and accuracy behind this central issue.

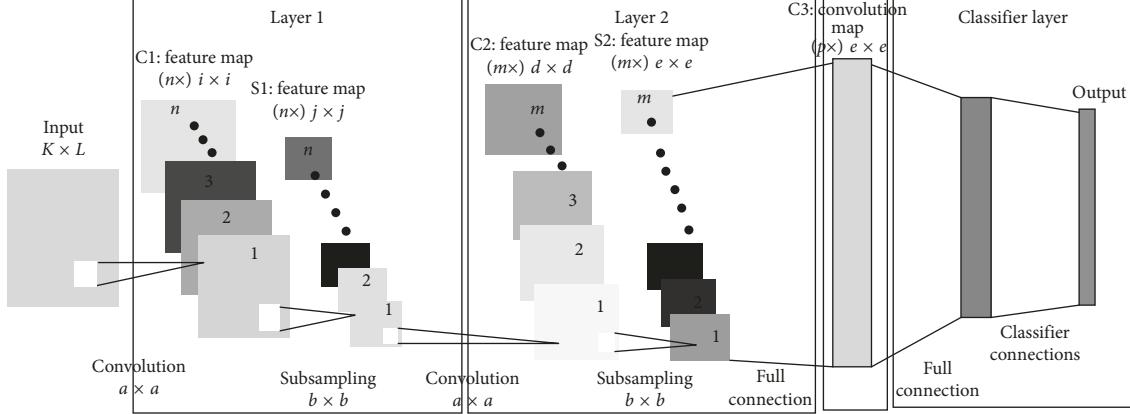


FIGURE 1: Convolutional neural network.

5.1. Data. The two employed models are trained and tested using normal and diseased brain computer tomography (CT) images collected from the Aminu Kano Teaching Hospital, Nigeria [34]. It is important to note that the abnormal images collected from this database are of different types of haemorrhage, but they were all labeled as haemorrhage, because this work aims to classify whether the CT slice contains haemorrhage or not; haemorrhage identification from set of images regardless of the haemorrhage pathology type it may have is feasible [35].

5.2. Data Augmentation. Deep networks are data-hungry systems [36], hence the more data you feed them, the more powerful and accurate they become. Therefore, in this work we decided to use data augmentation in order to multiply the number of images collected for the database, which can help in preventing the overfitting that may be encountered during training [37]. Thus, each image is first rotated left and right and then flipped 70, 160, and 270 degrees. Overall, a total number of 12635 normal and haemorrhage CT brain images are obtained. Note that 70% of the data are used for training the employed networks while 30% are used for testing, i.e., 8855 and 3790 images, respectively. Table 1 shows the learning scheme that is used in this work.

Figure 2 shows some normal and haemorrhage CT slices of the brain that are the used for training and testing the deep networks.

The images of this database are originally of size $1024 * 1024 * 1$ pixels; hence, they were first downsampled to $227 * 227 * 1$ pixels to fit the input layer of the pretrained model: AlexNet which does not accept other input data sizes. Note that we decided to use the same input images size for the CNN created from scratch, only for networks performance comparison purposes, although any size could be used. Moreover, the images of the database are of grayscale type, and since the AlexNet model requires 3-channels input data, images were all converted to RGB by concatenating their grayscale channel for three times to become $227 * 227 * 3$.

5.3. Training the Network Models. The two employed deep models are simulated using MATLAB environment. The

TABLE 1: Learning scheme of the networks.

	Total number of images
Train	8855
Test	3790
Total	12635

networks were trained on a Windows 64-bit desktop computer with an Intel Core i7 4770 central processing unit (CPU) and 16 GB random access memory. It is important to mention that there was no graphical processing unit (GPU) available in the used desktop.

The performance evaluation of the networks was carried out using a held-out test set 30% of the data. The calculation of the loss and accuracy was achieved as follows:

$$\text{Loss} = -\left(\frac{1}{n}\right) \sum_{i=1}^n \log P(C), \quad (1)$$

$$\text{Accuracy} = \frac{C}{N},$$

where $P(C)$ is the probability of the correctly classified images, n is the number of images, while N is the total number of images during the training and/or testing phases.

5.3.1. CNN Training. The model architecture and training settings for the CNN employed to perform the classification of brain haemorrhage are presented in this section. Extensive tests are performed to determine the best learning parameters that optimize the neural network. Note that out of the retrieved 12635 brain CT images, 8855 images are used for training and 3790 images are used for validating the trained network.

The CNN architecture employed for the classification of brain haemorrhage images is shown in Figure 3, where “Conv” denotes a convolution layer, “BN” denotes batch normalization, “FM” denotes feature maps, and “FC” denotes fully connected layer. In this paper, all convolution operations are performed using convolution filters of size 3×3 with zero padding; all pooling operations are

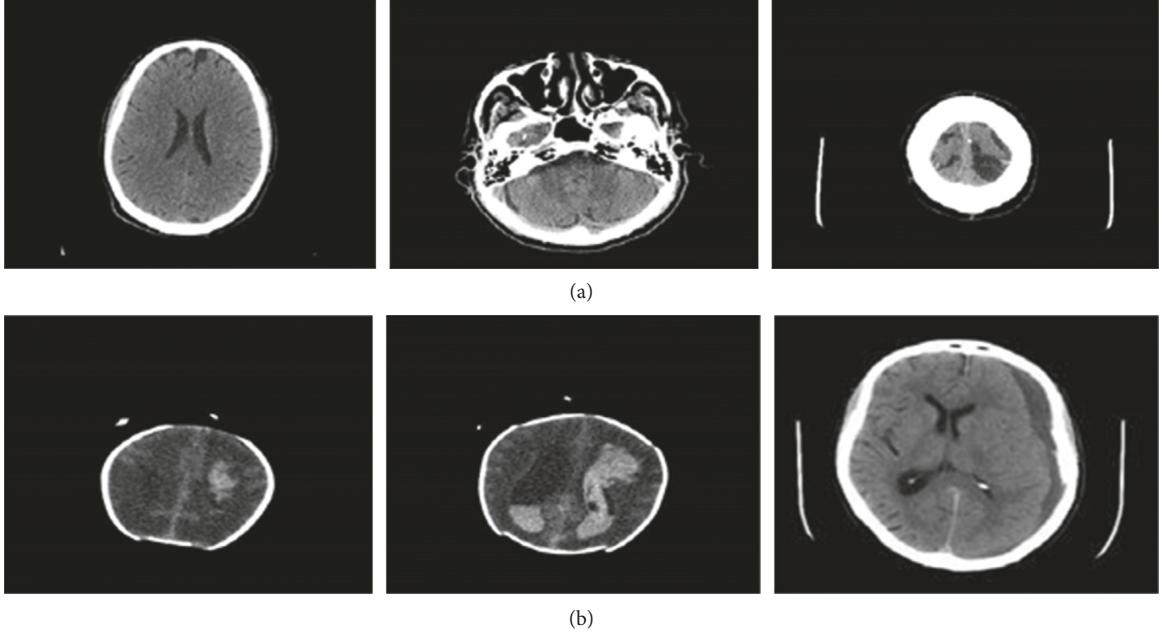


FIGURE 2: Sample of the databases training and validating images. (a) Haemorrhage images; (b) normal images.

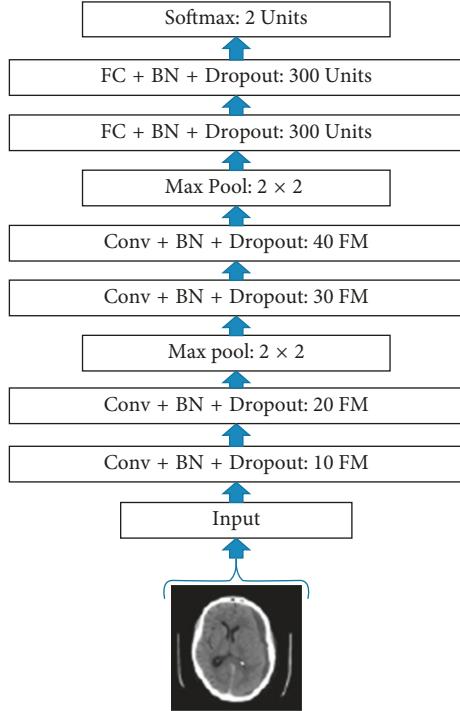


FIGURE 3: Proposed CNN architecture.

performed using max pooling windows of size 2×2 ; the input images to the model are of size 32×32 .

For designing the proposed architecture, we take into consideration the size of available (i.e., limited) training data for constructing a learning model that is considerably regularized. For example, we employ batch normalization and dropout training schemes which have been shown to improve model generalization [38–40]. For optimizing the

proposed model, we employ minibatch optimization via gradient descent; we use a batch size of 60. In addition, we use a learning rate of 0.001 and train the model for 100 epochs. The learning curve for the trained CNN is shown in Figure 4; a validation accuracy of 90.65% is achieved.

In addition, we observe a slight drop in validation performance when dropout and batch normalization are not employed for training the model; a validation accuracy of 87.33% is obtained. The overall proposed system for brain haemorrhage identification is tested using few CT brain haemorrhage images obtained from different sources available online. From the aforementioned database, we collect CT brain images of subjects with different haemorrhage conditions as test images. i.e., Figure 5. Experimental results show that the developed haemorrhage identification deep framework is capable of effectively classifying the haemorrhage within the test images with an accuracy of 87.13%.

We note that in contrast to other works that train and test the proposed approach on the same dataset, the proposed pipeline in this paper has been trained and validated on one dataset and achieved promising results when tested again on a completely different dataset. This shows the robustness of the deep CNN that is designed for such classification task.

5.3.2. AlexNet Training. AlexNet is the pretrained model selected to be used in this research because of its effective power in feature extraction. As can be seen in Figure 5, this deep convolutional neural network is comprised of 5 convolutional layers denoted as CONV1 to CONV5. These layers are followed by 3 fully connected layers denoted as FC1 to FC3, along with a Softmax activation function in the output layer (multinomial logistic regression).

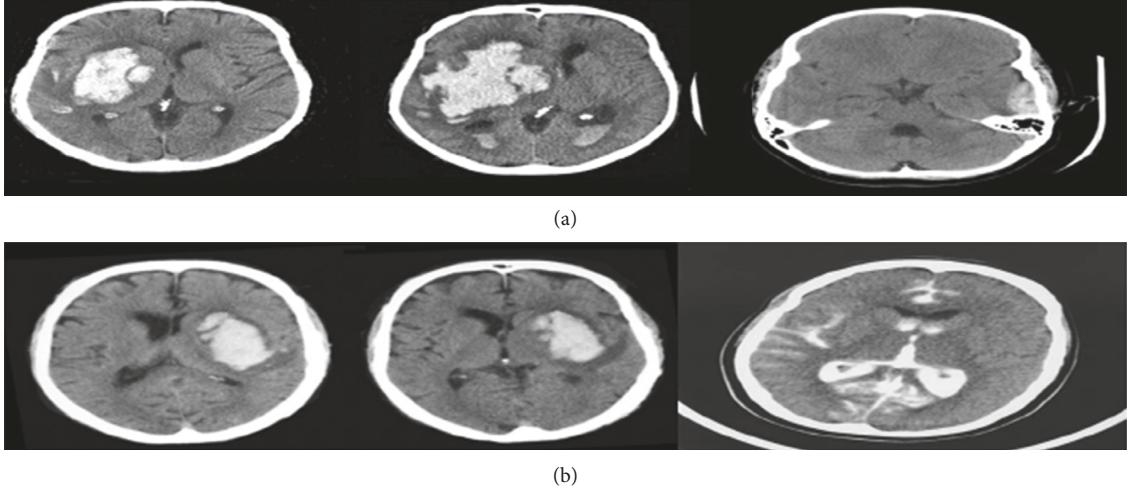


FIGURE 4: A sample of the brain images collected from the Internet to test the robustness of the system [41].

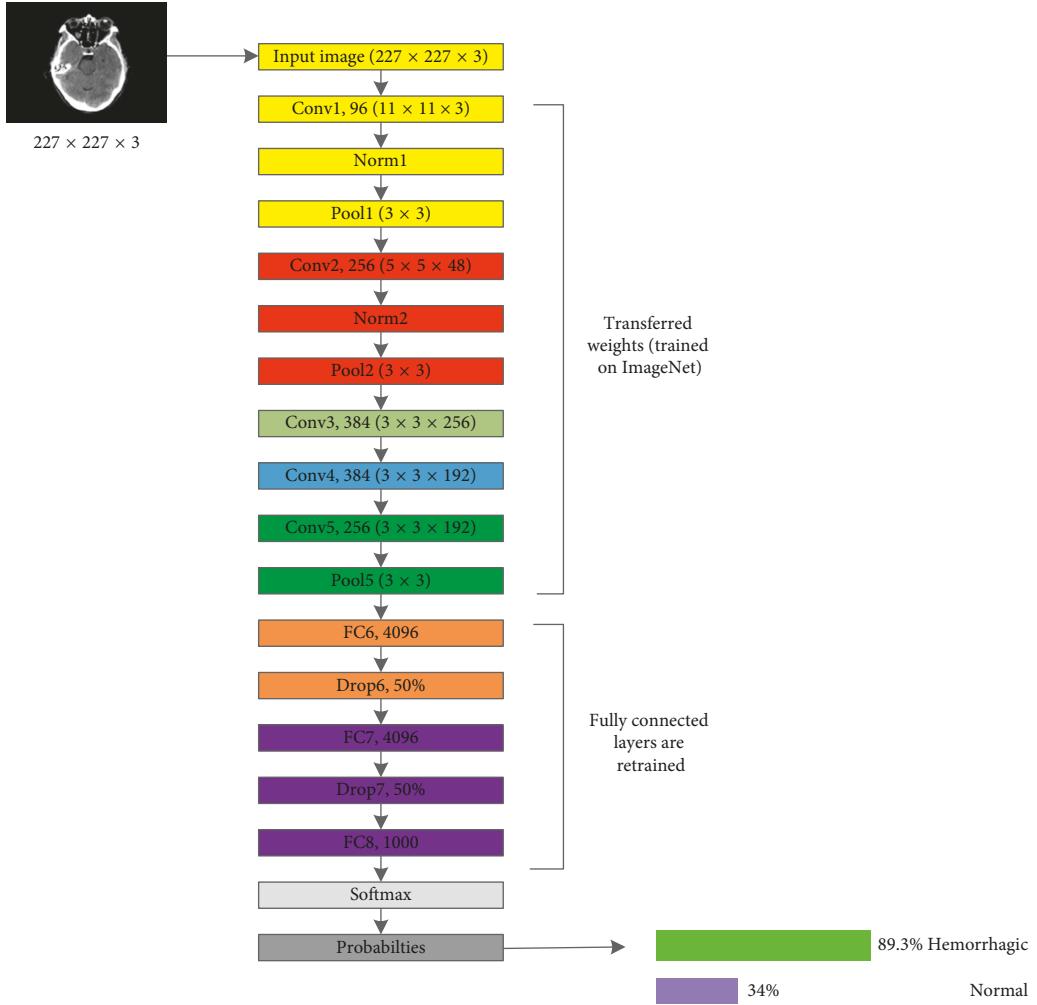


FIGURE 5: AlexNet proposed transfer learning network for the haemorrhage classification.

In this research, the publicly available weights of the network trained against the ILSVRC12 are used. As a pretrained model is employed (AlexNet), the final fully connected layer (FC8) was

disconnected in order to add a new layer having 2 output neurons corresponding to the two CT brain images' categories. Note that the weights of this layer are initialized at random.

Contrarily, the remaining five convolutional layers are kept in the network for sharing the learned parameters, in particular, weights. These weights are already trained on large datasets, ImageNet, to extract high-level features of the input data. Thus, when transferring the knowledge of AlexNet to haemorrhage classification task, these weights can act as a powerful extractor of different levels of abstractions from input data features.

The network is trained using minibatch of size 200 images of each iteration via stochastic gradient descent SGD [42]. Also, an initial learning rate is set to 0.01 to the fully connected layers (FC6, FC7, and FC8) and a reducing factor of 0.1 after 2000 iterations. Wherefore, this may fasten the learning of the network for the final fully connected layer (FC8). Table 2 shows the networks parameters during training and the result of the classification task. As seen, AlexNet has reached average training and testing accuracy of 94.12% and 92.13%, respectively.

An image from the test dataset is selected to evaluate the performance of the network in the classification pathway. Table 3 shows the mean square error (MSE) loss after each convolutional layer being trained.

5.3.3. Proposed AlexNet-SVM Training. Figure 6 shows the architecture of the modified version of AlexNet, in which an SVM classifier is used instead of a neural network. Similarly, this modified network, AlexNet-SVM, is also trained with the same conditions and same number of images except for the number of iterations which is here 140.

As seen in Figure 6 AlexNet-SVM's training parameters were similar to the parameters of AlexNet; however, it is noted that their performance was different. AlexNet-SVM was trained and it reached a lower MSE (0.054) compared to other networks. In addition, AlexNet-SVM achieved higher accuracies during training and testing with values of 96.34% and 93.48%, respectively.

6. Results and Discussion

Once trained, all network models are tested on 30% of the available data. Table 4 shows the performances of each model during testing. As can be seen, the CNN, AlexNet, and AlexNet-SVM achieved different accuracies of 90.65%, 92.13%, and 93.48%, respectively. AlexNet-SVM was capable of achieving more accurate generalizing power on unseen data. However, a larger number of epochs was required to achieve such accuracy, which is relatively higher than that needed for CNN and AlexNet to achieve their highest accuracy. It is also noted that AlexNet-SVM reached a lower mean square error (MSE) (0.054) than that reached by AlexNet (0.087) and CNN (0.092); however, this also required longer training time. The learning curves of the trained models are shown in Figures 7–9. The figures show the variations accuracy with respect to the increase of the number of epochs. Consequently, it is seen that all models are trained well, but the increase of depth of AlexNet and AlexNet-SVM makes it more difficult to train, i.e., it required longer time and more epochs to reach the minimum square

TABLE 2: Models learning parameters.

	CNN	AlexNet	AlexNet-SVM
Learning parameters	Values	Values	Values
Training ratio (%)	80	80	80
Initial learning rates	0.001	0.01	0.01
Number of epochs	100	200	140
Training accuracy (%)	92.89	94.12	96.34
Testing accuracy (%)	90.65	92.13	93.48
Achieved mean square error (MSE)	0.092	0.087	0.054

TABLE 3: Loss at each convolutional layer of CNN.

Layer	CONV1	CONV2	CONV3	CONV4	CONV5
Loss	0.186	0.341	0.412	0.46	0.51

error (MSE) and converge. Furthermore, it is important to mention that due to this difference in time and epoch number, the classifier of AlexNet-SVM resulted in a lower MSE and higher recognition rate than that scored by AlexNet and CNN. As a result, to understand the learning performance of networks, we have an insight into the different levels features learned by the employed models, by visualizing the learned kernels or features in the convolutional layers, shown in Figures 10 and 11.

Figures 10 and 11 show the learned features of CNN and AlexNet, respectively. From Figure 6, it can be seen that neurons in the first convolution layer are the mostly active neurons in capturing good features in the training data. However, from Figure 11, it is seen that the neurons of the last convolutional layer of AlexNet are the most active neurons in capturing descriptive and different levels features. In addition, compared to CNN, this layer has an improved activity as observed in the learned features. Lastly, it can be noted that the neurons of the first and last convolutional layers of both networks have learned different and interesting representation of the input images. Generally, networks that tend to learn more descriptive and different levels features tend to perform better at run time, as the good knowledge acquired in the unsupervised pretraining contributes to better fine-tuning and classification.

Table 5 shows a comparison of the developed networks with some previous works that were proposed to classify brain haemorrhage using deep learning. Note that we ought to compare our approach with the deep networks and pretrained model researches that provide explicitly achieved accuracies and number of data. Firstly, a general analysis of the table shows that the pretrained models (transfer learning-based networks) achieved higher accuracies when compared to those that were created from scratch. The proposed AlexNet_SVM employed in this research achieved more powerful generalization capabilities than other AlexNet that use neural network classifiers like the networks employed in this research and also in other researches [43]. Moreover, AlexNet-SVM outperformed the networks that were created from scratch such as convolutional neural networks and autoencoders [21]. Furthermore, it is seen that the employed pretrained model (AlexNet) achieved a higher recognition rate (92.13%) than

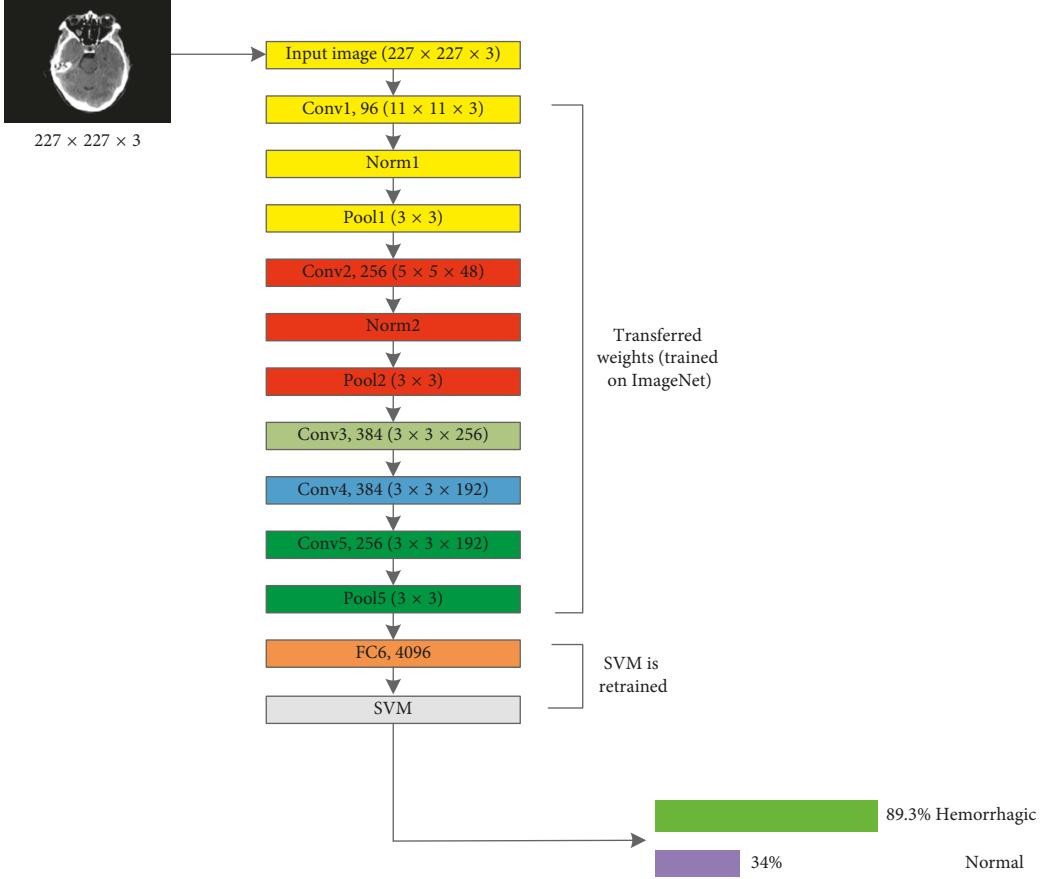


FIGURE 6: Modified AlexNet (AlexNet-SVM).

TABLE 4: Performance comparison of the employed networks.

	CNN	AlexNet	AlexNet-SVM
Testing images	3790	3790	3790
Number of correctly classified images	3436	3492	3543
Accuracy (%)	90.65	92.13	93.48

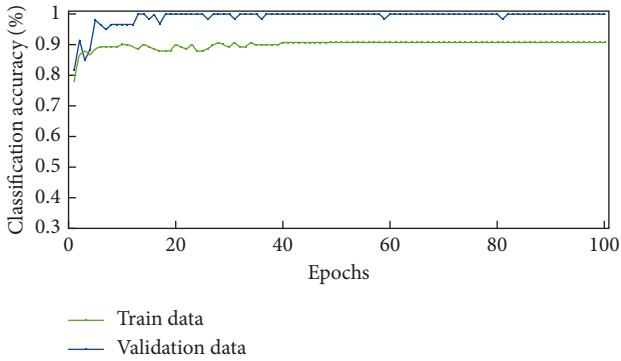


FIGURE 7: Learning curve for the trained CNN.

other earlier research works such as CNN created from scratch on less number of images [21]. Also, this model has outperformed other types of deep networks such as autoencoder (88.3%) and stacked autoencoder (90.9%) [21].

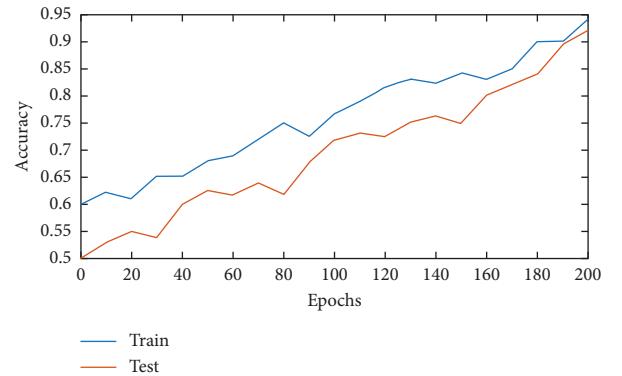


FIGURE 8: Learning curves of AlexNet.

This can probably be due to the deficiency of newly born networks in extracting the important features from input images which is a result of the small number of images used for training them in addition to their depth.

Overall, the application of pretrained models to solve haemorrhage classification challenge can end up with satisfying results since these deep structures have gained powerful feature extraction capabilities as they were trained using huge databases such as ImageNet [13]. The obtained results of applying the proposed AlexNet-SVM, AlexNet and CNN in this research show that applying deep CNNs to the

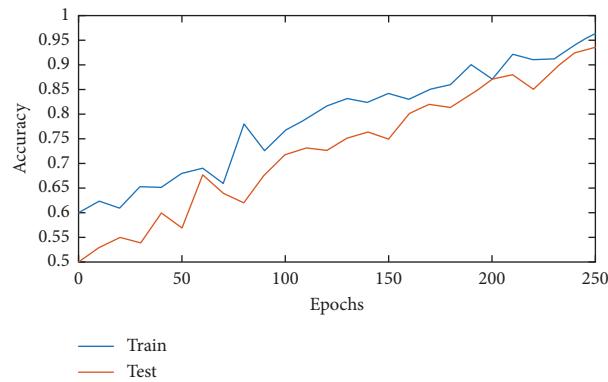


FIGURE 9: Learning curves of AlexNet-SVM.

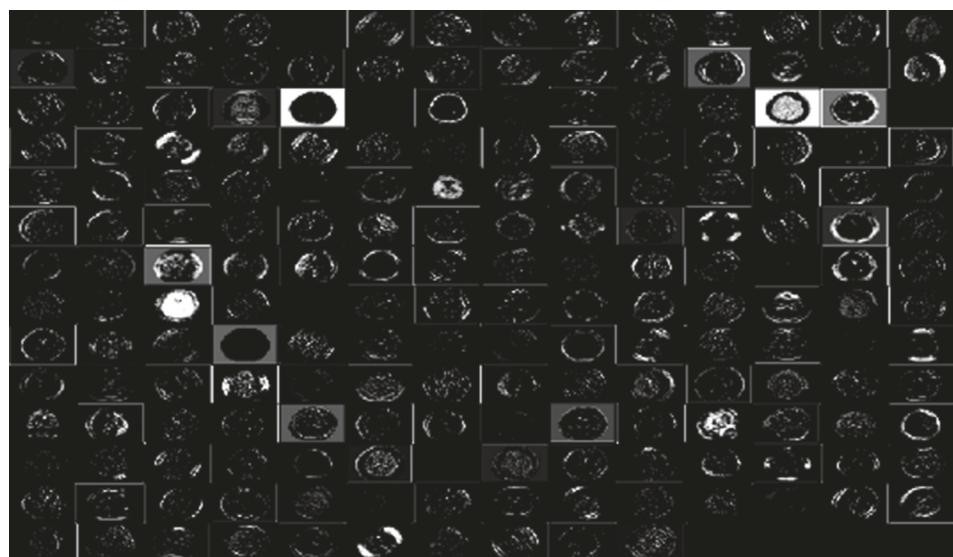


FIGURE 10: Learned kernels of CNN.

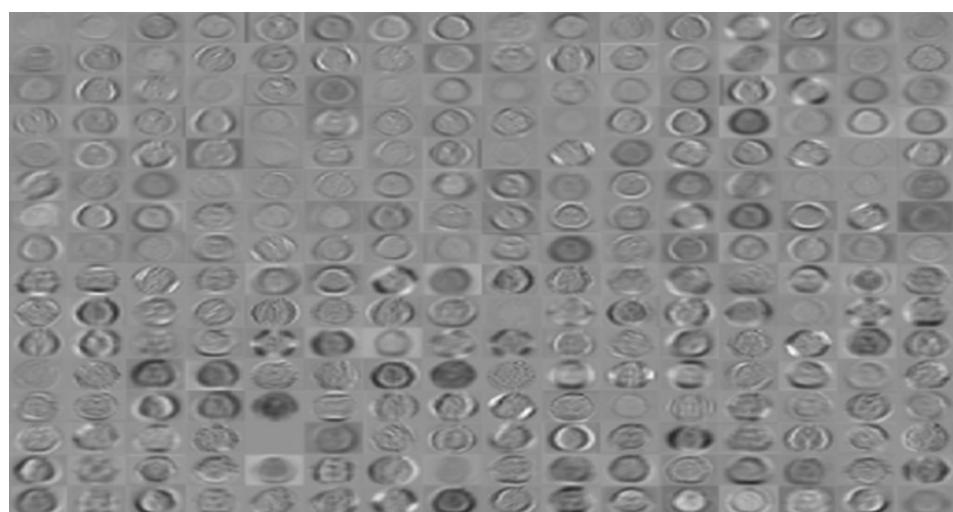


FIGURE 11: Learned kernels of AlexNet.

TABLE 5: Performance metrics of the networks.

Network model	CNN	AlexNet	AlexNet-SVM
Accuracy (%)	89	91	93
Sensitivity (%)	90	93	95
Specificity (%)	86	88	90
Misclassified (%)	11	9	7

TABLE 6: Results comparison with earlier works.

Network models	CNN	AlexNet	AlexNet-SVM	AlexNet [43]	CNN [19]	AE [19]	SAE [19]
Number of images	12635	12635	12635	11,088	2527	2527	2527
Accuracy (%)	90.65	92.13	93.48	92	89.6	88.3	90.9

problem of brain haemorrhage is promising, in a way that a haemorrhage can be identified by a deep neural network with low margins of error.

6.1. Performance Evaluation Metrics. These metrics are derived from classification of the tested sampling images, as shown in Table 6, being derived by a contingency table which is called confusion matrix [13]. Accuracy indicates the percentage of rightly classified image samples, without considering their class labels. For a binary classification that concludes on positive and negative classes, sensitivity is the percentage of correctly classified samples and specificity is the number of correctly negative samples classified:

$$\begin{aligned} \text{Accuracy} &= \frac{(TP + TN)}{TN + TP + FP + FN}, \\ \text{Sensitivity} &= \frac{TP}{TP + FN}, \\ \text{Specificity} &= \frac{TN}{TN + FP}. \end{aligned} \quad (2)$$

6.2. Models Comparison. In this section, the comparison of the conventional AlexNet and the proposed AlexNet-SVM is explained, in order to show the advantages of the fusion of AlexNet and SVM, in addition to the possible reasons of AlexNet-SVM outperformance. As seen in Table 5, the fusion of AlexNet and SVM resulted in a slight boost of accuracy by 0.934. This outperformance is mainly due to the use of a different optimization criterion that the SVM uses. This algorithm is used to minimize the prediction loss on the training set of the neural network. However, in practice, there are two challenges with this risk. First is the convexity; it is not convex which means that many local minimums may exist. Second problem is the smoothness; it is not smooth, which means it may not be practically minimized. In contrast, SVM aims to minimize the generalization error by using structural risk minimization principles for the testing set. As a result of a maximized margin, the generalization ability of SVM is greater than that of the other classifiers.

7. Limitations

The effectiveness of deep learning in medical applications is great and improving with time; however, it still encounters some drawbacks, in particular, the availability data. The variability of data (e.g., contrast, noise, and resolution) can be one of the main barriers of the adaptation of deep learning in medicine. These intelligent models can suffer from poor generalization if data contain some noise and when they are generated from different modalities. Moreover, deep learning models are data-driving systems; the more the data, the more efficient they become. The problem is very few data are not publicly available in the medical field due to privacy issues as in most cases, the data contain sensitive information. Thus, we and many other researchers prefer to use transfer learning based models which usually require less number of data to learn, as they are already trained using large amounts of data. Hence, the system is capable of learning different levels of features, which helps in adapting the new task accurately, even if the data are not large.

8. Conclusion

In this research, the detection of brain haemorrhage in CT images problem is solved using neural networks and the results sound robust and promising. One of the motivations behind this research is to address and attempt to overcome the difficulties that radiologists might encounter when diagnosing brain haemorrhage suspected images. Hence, we investigated the use of a potential deep convolutional neural network that can help the medical experts in making more accurate decisions. As a result, this may reduce the diagnosis error and boost the accuracy of haemorrhage identification made by medical experts. The paper proposes a pretrained modified network “AlexNet-SVM” for the same classification task. The three models including the proposed model were trained on a relatively small database in order to examine the network performance. It is obvious that the application of deep learning networks in medical image analysis encounters several challenges. The most common challenge is the lack of large training data sets which can be considered as an obstacle. The experiments conducted in this study demonstrated that the transfer of knowledge into medical images can be possible, even though the deep

networks are originally trained on natural images. The proposed model using the SVM classifier helps in improving the performance of AlexNet. Moreover, it was manifested that small number of data can be enough for fine-tuning a pretrained model, in contrast to a CNN created from scratch which needs a large number of data to be trained. Thus, the proposed model's performance is an indicator of how transfer learning-based networks can be considered in brain haemorrhage identification.

Data Availability

The brain haemorrhage data used to support the findings of this study may be released upon application to the Aminu Kano Teaching Hospital, Kano, Nigeria, at <http://akth.org.ng/index.php/contact>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] U. Balasooriya and M. S. Perera, "Intelligent brain haemorrhage diagnosis using artificial neural networks," in *Proceedings of the Business Engineering and Industrial Applications Colloquium (BEIAC)*, pp. 128–133, IEEE, Kuala Lumpur, Malaysia, September 2012.
- [2] R. Badenes and F. Bilotta, "Neurocritical care for intracranial haemorrhage: a systematic review of recent studies," *British Journal of Anaesthesia*, vol. 115, no. 2, pp. 68–74, 2015.
- [3] L. B. Morgenstern, J. C. Hemphill, C. Anderson et al., "Guidelines for the management of spontaneous intracerebral haemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association," *Stroke*, vol. 46, pp. 2032–2060, 2010.
- [4] R. H. Abiyev and M. K. S. Ma'aitah, "Deep convolutional neural networks for chest diseases detection," *Journal of Healthcare Engineering*, vol. 2018, Article ID 4168538, 11 pages, 2018.
- [5] A. Helwan and R. Abiyev, "Shape and texture features for the identification of breast cancer," in *Proceedings of the World Congress on Engineering and Computer Science*, vol. 2, pp. 19–21, San Francisco, USA, October 2016.
- [6] S. U. Akram, J. Kannala, L. Eklund, and J. Heikkilä, "Cell segmentation proposal network for microscopy image analysis," in *Proceedings of the International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, pp. 21–29, Springer International Publishing, Athens, Greece, October 2016.
- [7] A. Helwan and D. Uzun Ozsahin, "Sliding window based machine learning system for the left ventricle localization in MR cardiac images," *Applied Computational Intelligence and Soft Computing*, vol. 2017, Article ID 3048181, 9 pages, 2017.
- [8] O. K. Oyedotun, E. O. Olaniyi, A. Helwan, and A. Khashman, "Hybrid auto encoder network for iris nevus diagnosis considering potential malignancy," in *Proceedings of the 2015 International Conference on Advances in Biomedical Engineering (ICABME)*, pp. 274–277, Beirut, Lebanon, September 2015.
- [9] A. Mnih and G. E. Hinton, "Ascalable hierarchical distributed language model," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1081–1088, Vancouver, Canada, December 2009.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, Lake Tahoe, NV, USA, December 2012.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <http://arxiv.org/abs/1409.1556>.
- [12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Las Vegas, NV, USA, June 2016.
- [13] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large Scale visual recognition challenge," *International Journal of Computer Vision (IJCV)*, vol. 115, no. 3, pp. 211–252, 2015.
- [14] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer-Verlag, New York, USA, 2006.
- [15] O. K. Oyedotun, E. O. Olaniyi, and A. Khashman, "A simple and practical review of over-fitting in neural network learning," *International Journal of Applied Pattern Recognition*, vol. 4, no. 4, pp. 307–328, 2017.
- [16] Abien Fred Agarap, "A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data," 2017, <http://arxiv.org/abs/1709.03082>.
- [17] A. Alalshekmubarak and L. S. Smith, "A novel approach combining recurrent neural network and support vector machines for time series classification," in *Proceedings of the 2013 9th International Conference on Innovations in Information Technology (IIT)*, pp. 42–47, IEEE, Abu Dhabi, UAE, March 2013.
- [18] Y. Tang, "Deep learning using linear support vector machines," 2013, <http://arxiv.org/abs/1306.0239>.
- [19] S. Hussain, S. M. Anwar, and M. Majid, "Segmentation of glioma tumors in brain using deep convolutional neural network," *Neurocomputing*, vol. 282, pp. 248–261, 2018.
- [20] R. H. Abiyev and M. Arslan, "Head mouse control system for people with disabilities," *Expert Systems*, 2019.
- [21] A. Helwan, G. El-Fakhri, H. Sasani, and D. Uzun Ozsahin, "Deep networks in identifying CT brain hemorrhage," *Journal of Intelligent & Fuzzy Systems*, vol. 35, no. 2, pp. 2215–2228, 2018.
- [22] R. Mahajan and P. M. Mahajan, "Survey on diagnosis of brain haemorrhage by using artificial neural network," *International Journal of Scientific Research Engineering & Technology*, vol. 5, no. 6, pp. 378–381, 2016.
- [23] T. Gong, R. Liu, C. L. Tan et al., "Classification of CT brain images of head trauma," in *Proceedings of the IAPR International Workshop on Pattern Recognition in Bio-informatics*, pp. 401–408, Springer, Melbourne, Australia, October 2007.
- [24] O. K. Oyedotun and A. Khashman, "Deep learning in vision-based static hand gesture recognition," *Neural Computing and Applications*, vol. 28, no. 12, pp. 3941–3951, 2017.
- [25] H.-C. Shin, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
- [26] O. K. Oyedotun and K. Dimililer, "Pattern recognition: invariance learning in convolutional auto encoder network," *International Journal of Image, Graphics and Signal Processing*, vol. 8, no. 3, pp. 19–27, 2016.

- [27] M. D. Zeiler and R. Fergus, "Stochastic pooling for regularization of deep convolutional neural networks," 2013, <http://arxiv.org/abs/1301.3557>.
- [28] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" *Journal of Machine Learning Research*, vol. 11, pp. 625–660, 2010.
- [29] O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, and G. Penn, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'12)*, pp. 4277–4280, IEEE, Kyoto, Japan, March 2012.
- [30] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," in *Proceedings of the 32nd International Conference on Machine Learning*, pp. 97–105, Lille, France, July 2015.
- [31] P. M. Cheng and H. S. Malhi, "Transfer learning with convolutional neural networks for classification of abdominal ultrasound images," *Journal of digital imaging*, vol. 30, no. 2, pp. 234–243, 2017.
- [32] H. Lei, T. Han, F. Zhou et al., "A deeply supervised residual network for HEp-2 cell classification via cross-modal transfer learning," *Pattern Recognition*, vol. 79, pp. 290–302, 2018.
- [33] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1058, 2014.
- [34] Aminu Kano Teaching Hospital, Nigeria, <http://akth.org.ng>.
- [35] M. Al-Ayyoub, D. Alawad, K. Al-Darabsah, and I. Aljarrah, "Automatic detection and classification of brain hemorrhages," *WSEAS Transactions on Computers*, vol. 12, no. 10, pp. 395–405, 2013.
- [36] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [37] J. Weston, F. Ratle, H. Mobahi, and R. Collobert, "Deep learning via semi-supervised embedding," in *Neural Networks: Tricks of the Trade*, pp. 639–655, Springer, Berlin, Heidelberg, 2012.
- [38] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the International Conference on Machine Learning*, pp. 448–456, Lille, France, June 2015.
- [39] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15)*, pp. 5353–5360, Boston, MA, USA, June 2015.
- [40] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: a convolutional neural-network approach," *IEEE Transactions on Neural Networks*, vol. 8, no. 1, pp. 98–113, 1997.
- [41] F. Gaillard, "Intracranial hemorrhage, radiology reference article," 2018, <https://radiopaedia.org/articles/intracranial-haemorrhage>.
- [42] R. G. J. Wijnhoven and P. H. N. de With, "Fast training of object detection using stochastic gradient descent," in *Proceedings of the International Conference on Pattern Recognition (ICPR)*, pp. 424–427, Istanbul, Turkey, August 2010.
- [43] V. Desai, A. E. Flanders, and P. Lakhani, "Application of deep learning in neuroradiology: automated detection of basal ganglia haemorrhage using 2D-convolutional neural networks," 2017, <http://arxiv.org/abs/1710.03823>.

Research Article

An Optimized Channel Selection Method Based on Multifrequency CSP-Rank for Motor Imagery-Based BCI System

Jian Kui Feng,¹ Jing Jin¹, Ian Daly,² Jiale Zhou,¹ Yugang Niu,¹ Xingyu Wang,¹ and Andrzej Cichocki^{3,4,5}

¹Key Laboratory of Advanced Control and Optimization for Chemical Processes, Ministry of Education, East China University of Science and Technology, Shanghai, China

²Brain-Computer Interfacing and Neural Engineering Laboratory, School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK

³Skolkovo Institute of Science and Technology (SKOLTECH), 143026 Moscow, Russia

⁴Systems Research Institute PAS, Warsaw, Poland

⁵Nicolaus Copernicus University (UMK), Torun, Poland

Correspondence should be addressed to Jing Jin; jinjingat@gmail.com

Received 29 October 2018; Revised 7 March 2019; Accepted 18 April 2019; Published 13 May 2019

Guest Editor: Minkyu Ahn

Copyright © 2019 Jian Kui Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Due to the redundant information contained in multichannel electroencephalogram (EEG) signals, the classification accuracy of brain-computer interface (BCI) systems may deteriorate to a large extent. Channel selection methods can help to remove task-independent electroencephalogram (EEG) signals and hence improve the performance of BCI systems. However, in different frequency bands, brain areas associated with motor imagery are not exactly the same, which will result in the inability of traditional channel selection methods to extract effective EEG features. **New Method.** To address the above problem, this paper proposes a novel method based on common spatial pattern- (CSP-) rank channel selection for multifrequency band EEG (CSP-R-MF). It combines the multiband signal decomposition filtering and the CSP-rank channel selection methods to select significant channels, and then linear discriminant analysis (LDA) was used to calculate the classification accuracy. **Results.** The results showed that our proposed CSP-R-MF method could significantly improve the average classification accuracy compared with the CSP-rank channel selection method.

1. Introduction

Brain-computer interface (BCI) technology enables the human brain to communicate directly with the outside world through electroencephalogram (EEG) signals and has attracted considerable attention in recent years [1]. BCIs can translate brain signals into output commands that allow users to control external auxiliary devices (such as wheelchairs, robotic arms, etc.) [2, 3]. Since BCIs provide an alternative way for people to communicate without use of peripheral nerves and muscles, they show a great value of, for instance, helping patients with severe neuromuscular disorders such as spinal cord injury or amyotrophic lateral sclerosis to restore their communication pathways and control their environment [4, 5].

Compared with evoked potential-based BCIs, motor imagery- (MI-) based BCIs have the advantages of being independent of external stimuli and easier to operate. MI-based BCI has been shown to be suitable for mechanical control and exercise rehabilitation training [6]. However, the brain signals used to control MI movements suffer from a number of problems including, but not limited to, low spatial resolution and low signal-to-noise ratios and are susceptible to strong artifacts [7, 8]. Researchers have used many feature extraction methods to address these difficulties, such as wavelet transforms, wavelet packet transforms, autoregressive (AR) models, and common spatial patterns (CSP). CSP, in particular, has been widely used for feature extraction to improve the performance of MI-based BCIs [9]. The use of multichannel signals tends to achieve good

classification performance [10]. Nevertheless, multichannel signals typically carry a large amount of redundant information, which introduces additional noise sources and may decrease EEG-based motor imagery classification accuracy when compared to a small set of optimal EEG channels [11]. Channel selection can effectively exclude the redundant channels and select the optimal brain areas for MI-based BCIs. Therefore, channel selection is an important method of feature extraction for MI-based BCI [12].

In recent years, researchers have proposed many algorithms for channel selection, such as sequential floating forward selection (SFFS) [12], the mutual information-based channel selection method [13], support vector machine recursive feature elimination (SVM-rfe) [14], and CSP-rank [15, 16]. Among them, CSP-rank is one of the most frequently used channel selection methods [15, 16]. CSP-rank uses the projection matrix obtained by the CSP algorithm to sort and select the channels.

EEG is inherently noisy, and the signal-to-noise ratio is an important factor affecting the performance of BCIs [17, 18]. Many algorithms have been proposed to improve the signal-to-noise ratio of EEG. For example, the spatially sparsed common spatial pattern (SSCSP) method was proposed in [19], which has strong weights within the area of motor cortex and smooth weights elsewhere. The work in [20] proposed the spatially regularized common spatial pattern (SRCSP) method. Selim et al. [21] used root mean square (RMS) feature as inputs to an LDA classifier. Furthermore, Dai et al. [22] proposed the transfer kernel common spatial pattern (TKCSP) method to define the kernel of the domain-invariant by matching the division among source and target subjects. Park et al. [23] used noise-assisted multivariate extensions of empirical mode decomposition (NA-MEMD) to achieve a highly localized time-frequency representation. In [12], Qiu et al. proposed the improved sequential floating forward selection (SFFS) method, which combined the distribution of channels and an intelligent selection method (SFFS) to select channels for CSP in MI-based BCI. Finally, Feng et al. [8] designed a novel correlation-based time window selection (CTWS) algorithm for MI-based BCIs to address the time latency variation during an MI period between trials for each participant. These algorithms each worked in different ways to successfully improve the signal-to-noise ratio of EEG.

The cortical locations that are most heavily involved in motor control vary across EEG frequency bands [24]. Therefore, selecting channels in across different frequency bands could increase the discriminability of the extracted features further. However, traditional CSP-rank channel selection methods and the previous feature extraction algorithms did not consider the difference of channel configuration in different frequency bands [12, 15, 16] and could not suppress the interference from EEG features of different frequency bands, resulting in the degradation of their performance. In this paper, a new CSP-rank channel selection method is proposed, which considers the channel configuration in different frequency bands to increase the discriminability of the extracted features. The CSP-rank channel selection method was used to select the channels

under a certain frequency band, and then the features are extracted by CSP using the selected channels. The extracted features from all frequency bands were concatenated to form one feature vector, which was improved further by the least absolute shrinkage and selection operator (LASSO). Linear discriminant analysis (LDA) was used as the classifier to demonstrate the performance of the method in terms of its impact on classification accuracy.

The paper is structured as follows: Section 2 describes the applied datasets and the proposed method. Section 3 shows the evaluation results. Section 4 presents the discussion about our method. Finally, concluding remarks are given in Section 5.

2. Methods

2.1. Description of the Data. Dataset 1 (BCI Competition III datasets IVa): this dataset was recorded from 5 participants with 118 EEG channels. Visual cues were displayed for 3.5 s at a random interval uniformly drawn from the range 1.75 to 2.25 s. In the experiments, participants were instructed to perform three classes of MI movements: left hand, right hand, and right foot movements. Only data for the classes “right hand” and “right foot” were provided for the evaluation purpose. Each participant performed 140 trials for each class, respectively. A time window from 0 and 3.5 seconds, relative to cue presentation time, was used for feature extraction. The experiment process is illustrated in Figure 1(a). More details about the dataset can be found in the following website: http://www.bbci.de/competition/iii/desc_IVa.html.

Dataset 2 (BCI Competition IV dataset 1): the dataset was recorded from 7 participants with 59 EEG channels with a sampling rate of 100 Hz, including four healthy individuals and three artificially generated “participants.” For the purpose of the present study, only the calibration data (consisting of two runs totaling 200 trials) from the four healthy individuals were used. Each participant selected two classes of motor imagery from three available classes *left hand*, *right hand*, and *foot* motor imagery. Each trial started from a visual cue pointing left, right, or down. The cue was displayed for a period of 4 s, during which the participant was instructed to perform the cued motor imagery task. These periods were interleaved with 2 s of blank screen with a fixation cross shown in the center of the screen. The fixation cross was superimposed on the cues, i.e., it was shown for 6 s. Time window between 0.5 and 2.5 s was used for feature extraction. The experiment process is illustrated in Figure 1(b). More details can be found in the following website: http://www.bbci.de/competition/iv/desc_1.html.

2.2. CSP-Rank. CSP-rank was proposed based on the sorting of the CSP filter [13, 25]. CSP seeks a projection matrix to maximize the variance for one class and minimize the variance for another class at the same time in order to maximize the discriminability of the dataset across classes to aid classification. The CSP operation was achieved as follows:

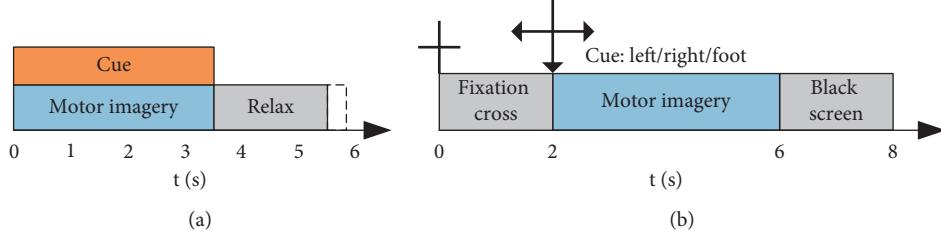


FIGURE 1: Illustration of the experimental protocol for a trial in dataset 1 (a) and dataset 2 (b). (b) is replaced from Feng et al. [8] (under the creative commons attribution license/public domain).

$$\arg \max \frac{w^T C_1 w}{w^T (C_1 + C_2) w}, \quad (1)$$

where w represented the projection vector and C_1 and C_2 represented the spatial covariance matrices of the two classes, respectively. We could regard it as the problem of finding generalized eigenvalues:

$$C_1 w = (C_1 + C_2) w D, \quad (2)$$

where D denotes the diagonal matrix containing the eigenvalues of C_1 . We select eigenvectors SF1 and SF2 corresponding to the largest and smallest eigenvalues, respectively, from w as the projection matrix. These filter coefficients were used to assign different weights to different electrodes based on their importance. If the coefficient of a particular electrode was large, then that means the electrode was more important.

The original EEG data need to be filtered before selecting the channels. To achieve this, a 5th order Butterworth bandpass filter from 8 Hz to 30 Hz was used to filter EEG data [12]. The CSP-rank method first found the two CSP filters SF1 and SF2, then sorted the absolute value of the filter coefficients in SF1 and SF2, respectively, and took the electrode with the next largest coefficient in turn from the two spatial filters. If an electrode was already taken, then the procedure simply moved on to the next coefficient in the same spatial filter until a new electrode was reached. The search process did not stop until a stopping criterion was fulfilled. The stopping criterion selected was that accuracy no longer increased with the number of selected channels.

2.3. Channel Selection Method Based on CSP-Rank for Multiple Frequency Bands. Figure 2 shows the structure diagram of the proposed method for optimizing EEG channel selection. Multiband signal decomposition filtering [25] was applied to the EEG signals recorded from all channels. Thereafter, the proposed CSP-rank method for reducing redundant channels was applied, and features were extracted with CSP from each frequency band. All features extracted from all frequency bands were concatenated to form one feature vector. The discriminant features were selected by LASSO from the feature vector. Finally, LDA was employed for the model training and accuracy calculation. The following subsections provide more details about the CSP-RMF approach.

2.3.1. Multiband Signal Decomposition Filter. A fixed time window was extracted from the trials of all the participants considered from the dataset. This time window began after the cue (beginning of motion imagination by the participants) and extended for 3.5 seconds for dataset 1 and 2 seconds for dataset 2 (the best window lengths for classification of dataset 2 were found to be among 1 s, 1.5 s, or 2 s [26]). Thereafter, seven frequency bands were considered covering the range 8–30 Hz. A fifth-order Butterworth filter was utilized to extract each band. The considered bandwidth was 4 Hz. Thus, the seven bands were defined as 8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz, 28–30 Hz, and 8–30 Hz.

For each subject, CSP-rank selection was used to remove several channels on each frequency band of the seven considered frequency bands. Thereafter, CSP was also applied on each frequency. Next, the CSP features were concatenated to form a high dimensional feature vector, which suffered from redundancies and irrelevant information. Such a feature vector might confuse the classifier. Therefore, an efficient feature selection algorithm was needed to select only the most relevant features.

2.3.2. Least Absolute Shrinkage and Selection Operator. LASSO is a filter-based feature selection method and does not depend on any classifier; instead, they selected features according to statistical criteria. This selection method was on average more time efficient and more resistant to overfitting compared to wrapper-based feature selection methods [27]. It has also been shown to be efficient in feature selection for MI-based BCIs [28]. The goal of the algorithm was to minimize the residual sum of squared errors with a bound on the sum of absolute values of linear regression coefficients that had to be less than a given constant. The features would automatically be discarded corresponding to coefficients that were exactly 0. LASSO minimized the function:

$$\left(\frac{1}{2N} \sum_{i=1}^N (y_i - \alpha_0 - x_i^T \alpha)^2 + \lambda \sum_{j=1}^n |\alpha_j| \right), \quad (3)$$

where N was the sample of samples, y_i was the response for sample i , x_i was the n -dimensional input vector for sample i , λ was a nonnegative regularization parameter, and α_0 and α were regression parameters (α_0 was a scalar; α was a n -dimensional vector). As λ increased, the number of nonzero components of α decreased. The values of

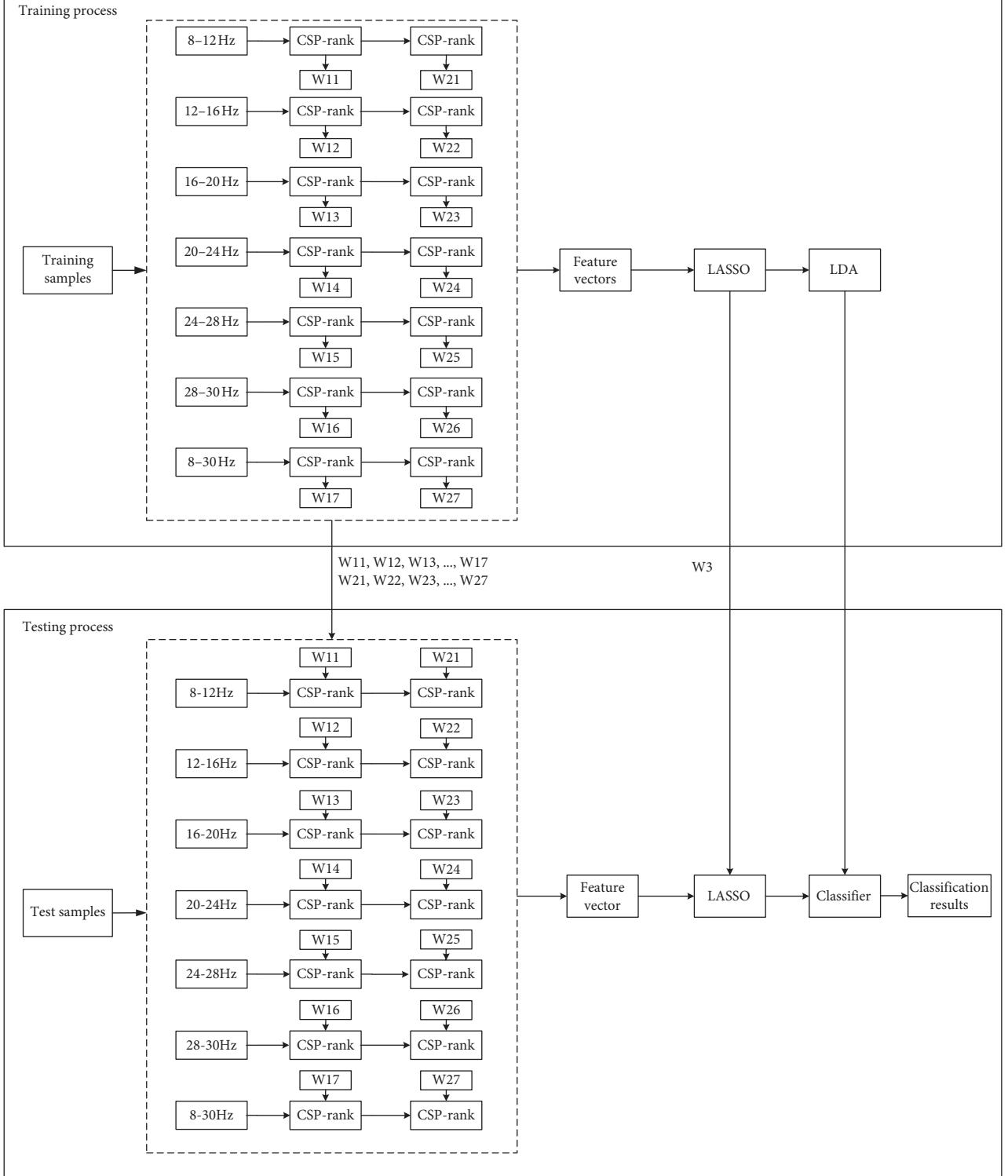


FIGURE 2: Structure diagram of the CSP-R-MF algorithm.

parameters mentioned were set as default values in MATLAB R2015b built-in LASSO functions.

2.4. Classification Scheme. We used n -fold cross validation for performance validation. All samples were divided into n blocks for each fold cross validation. Nine blocks were used

as training data, and the remaining one block was used as test data. Finally, the average classification accuracy of n -fold cross validation was selected as the estimation criteria. For dataset 1, each participant needs to perform 140 trials for each class; we selected 20 trials of 140 trials without repetition as the test data and the remaining trials as the training data for each cross validation. So, 7-fold cross validation was

finally selected as the performance evaluation. However, each subject needed to perform 100 trials for each class in dataset 2; we selected 10 trials of 100 trials without repetition as the test data and the remaining trials as the training data for each cross validation. So, 10-fold cross validation was finally selected as the performance evaluation method. LDA was selected as the classifier. It classified samples by maximizing the distance between classes and minimizing intraclass variance and is often used in research of motor imagery-based BCI system [5].

2.5. Filter Bank Common Spatial Patterns. Filter bank common spatial pattern (FBCSP) was proposed by Kai et al. [29] to perform autonomous selection of key temporal spatial discriminative EEG characteristics. FBCSP comprises four progressive stages of EEG measurements processing: frequency filtering (the 5-fold Butterworth was selected in this paper), spatial filtering (CSP was selected in this paper), feature selection (LASSO was selected in this paper), and classification (LDA was selected in this paper). In the first stage, multiband signal decomposition filter was applied to the EEG signals. In the second stage, CSP was used to extract features of the EEG signals from each frequency band. In addition, all features extracted from all frequency bands were concatenated to form one feature vector. In the third stage, the discriminative parts of the feature vector were automatically selected with LASSO. In the fourth stage, LDA was selected to classify the discriminative parts of CSP features.

3. Results

The overall accuracy behavior averaged from all the subjects in each dataset is shown in Figure 3. It may be observed that the CSP-R-MF achieves better performance in terms of classification accuracy. In the beginning, the classification accuracy increasingly corresponds to the increase in the number of selected channels. The classification accuracy would not increase further or would decrease a little, when more and more channels were selected. The peak point was 30 out of 118 channels for dataset 1, and the peak point was 24 out of 59 channels for dataset 2. So, we selected, respectively, 24 out of 59 channels for dataset 1 and 30 out of 118 channels for dataset 2 for later EEG analysis in each frequency band. Group level statistics are not reported as we do not have sufficient participants in this paper.

Figure 4 shows ROC curves for the two methods. The red line represents the accuracy achieved by the CSP-R-MF method proposed in this paper, and the blue line represents accuracy achieved by the CSP-rank method. Obviously, compared with the blue line, the red line generally tends to be in the upper left corner of the graph. This means that the model obtained by the method proposed in this paper has better performance than the model obtained by CSP-rank.

Figure 5 shows the classification accuracies achieved with each of the two methods. The horizontal axis indicates the subject, and the vertical axis indicates the classification

accuracy. The red bar represents CSP-R-MF results, and the blue bar represents CSP-rank results. Compared to CSP-rank, the classification accuracy of CSP-R-MF improved by 7.05% (75.43% VS 82.48%) for dataset 1 and 7% (70.75% VS 77.75%) for dataset 2. P value expresses the significance level. The smaller the P value, the higher the significance. In dataset 1, compared to CSP-rank, the classification accuracy of subjects aa, al, av, aw, and ay was, respectively, improved by 15.36%, 0.62%, 6.07%, 8.57%, and 4.64% with CSP-R-MF. In dataset 2, compared to CSP-rank, the classification accuracy of subjects S1, S2, S3, and S4 was, respectively, improved by 5%, 5.5%, 14.5%, and 3% with CSP-R-MF.

We also compared the performance of CSP-R-MF and FBCSP. The comparison results are shown in Table 1. Compared to FBCSP, the classification accuracy of CSP-R-MF improved by 2.91% for dataset 1 and 7% for dataset 2. Table 2 shows the correspondence between electrodes and numbers.

4. Discussion

It is well established that motor imagery produces an event-related de\synchronization (ERD\S) over the sensorimotor areas within the mu rhythm (8–13 Hz) and the beta band (13–30 Hz) [7]. Therefore, the bandpass filter used for measuring the ERD/S was between 8 to 30 Hz [8, 12]. This frequency band may be further divided into 7 subbands (4–8 Hz, 8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz, 28–32 Hz, 32–36 Hz, and 36–40 Hz) to study the effect of frequency band selection on motor imagery-based BCI control, as proposed in work on subband common spatial pattern (SBCSP) [30] and FBCSP. It was found that the performance of the MI-based BCI correlated with these frequency bands. However, the best frequency band was not exactly the same for all subjects. In this paper, we selected 7 bands (8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz, 28–30 Hz, and 8–30 Hz) to study movement-related brain activity during motor imagery in different frequency bands.

Feature extraction is one of the most important steps for the classification of motor imagery EEG [31]. In particular, CSP, a spatial feature extraction method, has become the most commonly used method in MI-based BCI systems [12, 32]. However, the performance of the CSP is susceptible to interference signals. The channel selection algorithm selects the channels associated with motor imagery and removes the channels that do not contribute significantly to classification of motor imagery activity. This improves the signal-to-noise ratio of the EEG signals. Brain areas associated with motor imagery are not exactly the same in different frequency bands [17]. Therefore, the channels selected will not be the same in different frequency bands.

In this study, the proposed CSP-R-MF algorithm considered variations of brain areas involved in motor imagery in different frequency bands and automatically selected channels for each frequency band via the CSP-rank method. Figure 6 shows the topographic maps of channels selected by CSP-R-MF in different frequency bands. The selected channels were marked with different colors according to the number of times each channel was selected.

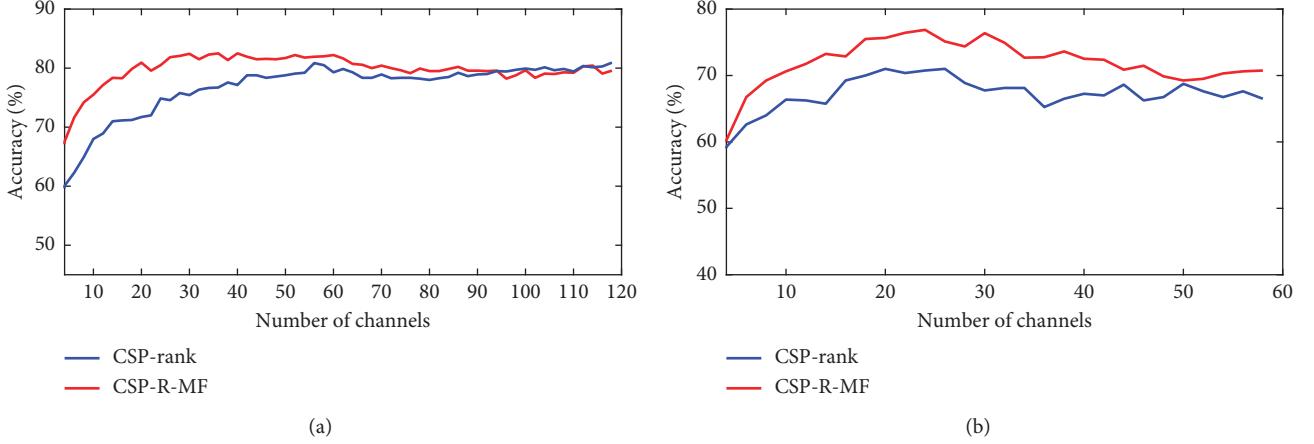


FIGURE 3: Overall accuracy curves showing the accuracy behavior with the varying numbers of channel in dataset 1 (a) and dataset 2 (b). The numbers of channels is the stopping criterion of the CSP-rank algorithm. The overall accuracy curves are the average over all subjects in each dataset. The red and blue lines denote the mean validation accuracy curves of two methods, respectively.

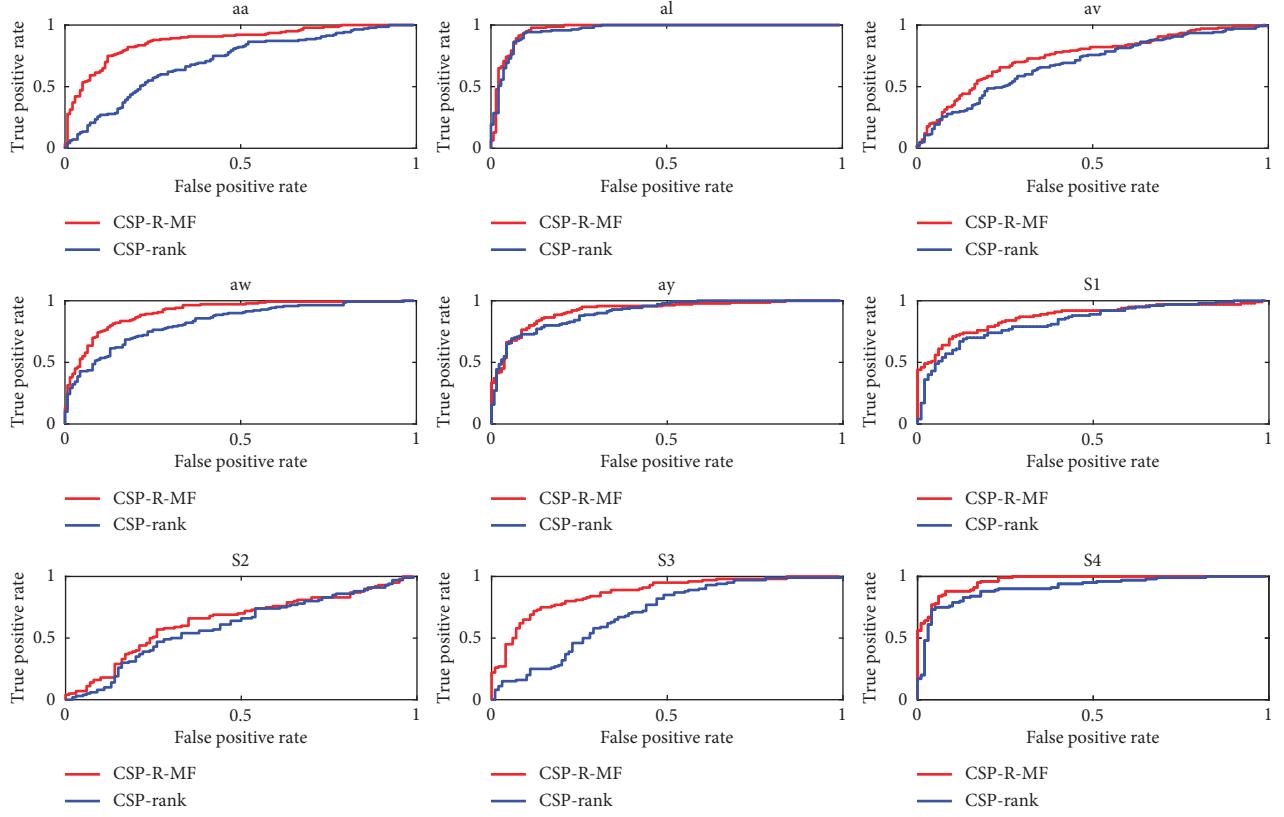


FIGURE 4: ROC curves of two methods for dataset 1 and dataset 2.

As shown in Figure 6, the distributions of selected channels were different under different frequency bands. However, they were basically distributed in the motor areas of the cerebral cortex. This reflects the frequently observed fact that motor imagery involves EEG activity in relevant areas of the motor cortex [31, 33]. Channels selected under a few frequency bands were not mainly distributed in the motion areas. For example, the selected channels in the 16–20 Hz and 28–30 Hz bands were not mainly distributed

in the motor areas for dataset 1. However, most of the selected channels under these two bands for dataset 2 were located in the motor area.

The ROC curve is a tool for evaluating the generalization performance of models. We compared the performance of two models based on CSP-rank and CSP-R-MF, respectively. Figures 3 and 5 showed that the performance of the model obtained by CSP-R-MF was better than the model obtained by CSP-rank. Figure 3 shows the accuracy behavior of CSP-

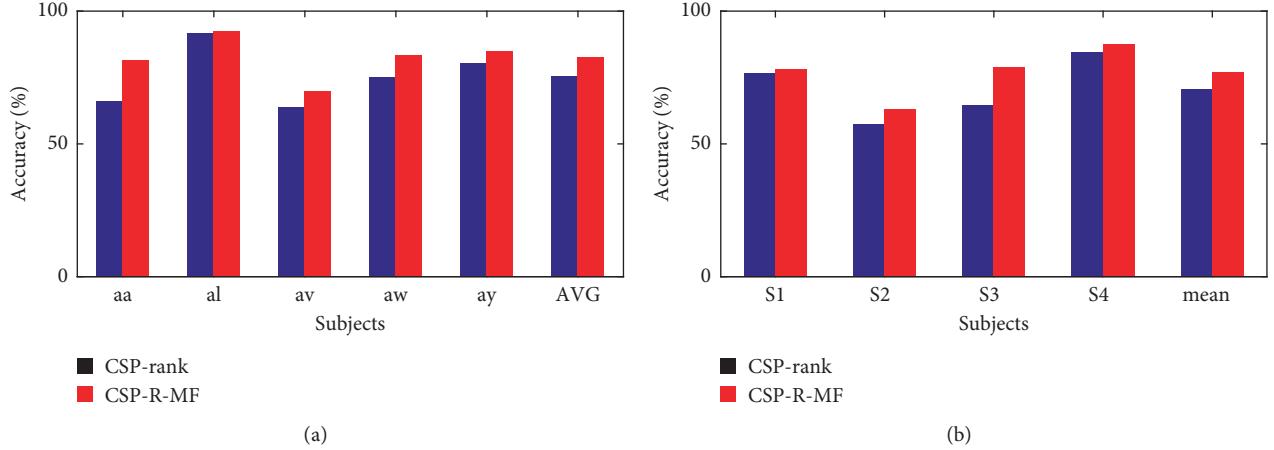


FIGURE 5: Performance comparison of the proposed algorithm CSP-R-MF with CSP-rank in dataset 1 (a) and dataset 2 (b).

TABLE 1: Classification accuracy of the proposed approach and FBCSP.

Subject	Dataset 1		Subject	Dataset 2	
	FBCSP	CSP-R-MF		FBCSP	CSP-R-MF
aa	79.64	81.43			
al	96.79	92.41	S1	75	81.5
av	51.79	70	S2	54	63
aw	90.71	83.57	S3	61.5	79
ay	78.93	85	S4	92.5	87.5
Average	79.57	82.48	Average	70.75	77.75

TABLE 2: Correspondence table between electrodes and numbers.

Mark : electrode									
1:PO3	2:PO1	3:POZ	4:PO2	5:PO4	6:OPO1	7:OPO2	8:O1	9:O2	10:l1
11:Ol1	12:OZ	13:Ol2	14:l2	15:PO7	16:PO8	17:AF7	18:Fp1	19:Fp2	20:AF8
21:AFP1	22:AFP2	23:AF3	24:AF4	25:FAF5	26:FAF1	27:FAF2	28:FAF6	29:F7	30:F5
31:F3	32:F1	33:Fz	34:F2	35:F4	36:F6	37:F8	38:FFC7	39:FFC5	40:FFC3
41:FFC1	42:FFC2	43:FFC4	44:FFC6	45:FFC8	46:FT9	47:FT7	48:FC5	49:FC3	50:FC1
51:FCz	52:FC2	53:FC4	54:FC6	55:FT8	56:FT10	57:CFC7	58:CFC5	59:CFC3	60:CFC1
61:CFC2	62:CFC4	63:CFC8	64:CFC8	65:T7	66:C5	67:C3	68:C1	69:Cz	70:C2
71:C4	72:C6	73:T8	74:CCP7	75:CCP5	76:CCP3	77:CCP1	78:CCP2	79:CCP4	80:CCP6
81:CCP8	82:TP9	83:TP7	84:CP5	85:CP3	86:CP1	87:CPz	88:CP2	89:CP4	90:CP6
91:TP8	92:TP10	93:PCP7	94:PCP5	95:PCP3	96:PCP1	97:PCP2	98:PCP4	99:PCP6	100:PCP8
101:P9	102:P7	103:P5	104:P3	105:P1	106:Pz	107:P2	108:P4	109:P6	110:P8
111:P10	112:PPO1	113:PPO5	114:PPO1	115:PPO2	116:PPO6	117:PPO8	118:FPz	—	—

Mark: electrode denotes electrode and its corresponding number. Mark: the number corresponding to the electrode.

R-MF and CSP-rank with the varying numbers of channel. The classification accuracy would increase corresponding to the increase in the number of selected channels in the beginning. The classification accuracy would decrease when more and more channels were selected. It proved the accuracy of the model has a globally optimal global region instead of falling into local optimum because of overfitting.

5. Conclusion

The brain areas associated with motor imagery are not exactly the same in different frequency bands [17]. Current state-of-the-art channel selection methods do not consider

this problem. In this study, we proposed a novel method based on common spatial pattern- (CSP-) rank channel selection method for multifrequency band (CSP-R-MF) selection. In our approach, a 5th order Butterworth filter was used to achieve multiband signal decomposition filtering. After that, CSP-rank and CSP were used to extract features of filtered EEG samples. The discriminant characteristics were selected with LASSO. Finally, classification algorithms were employed to calculate the classification accuracy. Experimental results showed that the CSP-R-MF algorithm could improve performance of MI-based BCIs compared to the CSP-rank algorithm. More specifically, the average classification accuracy improved 7.05% on dataset 1

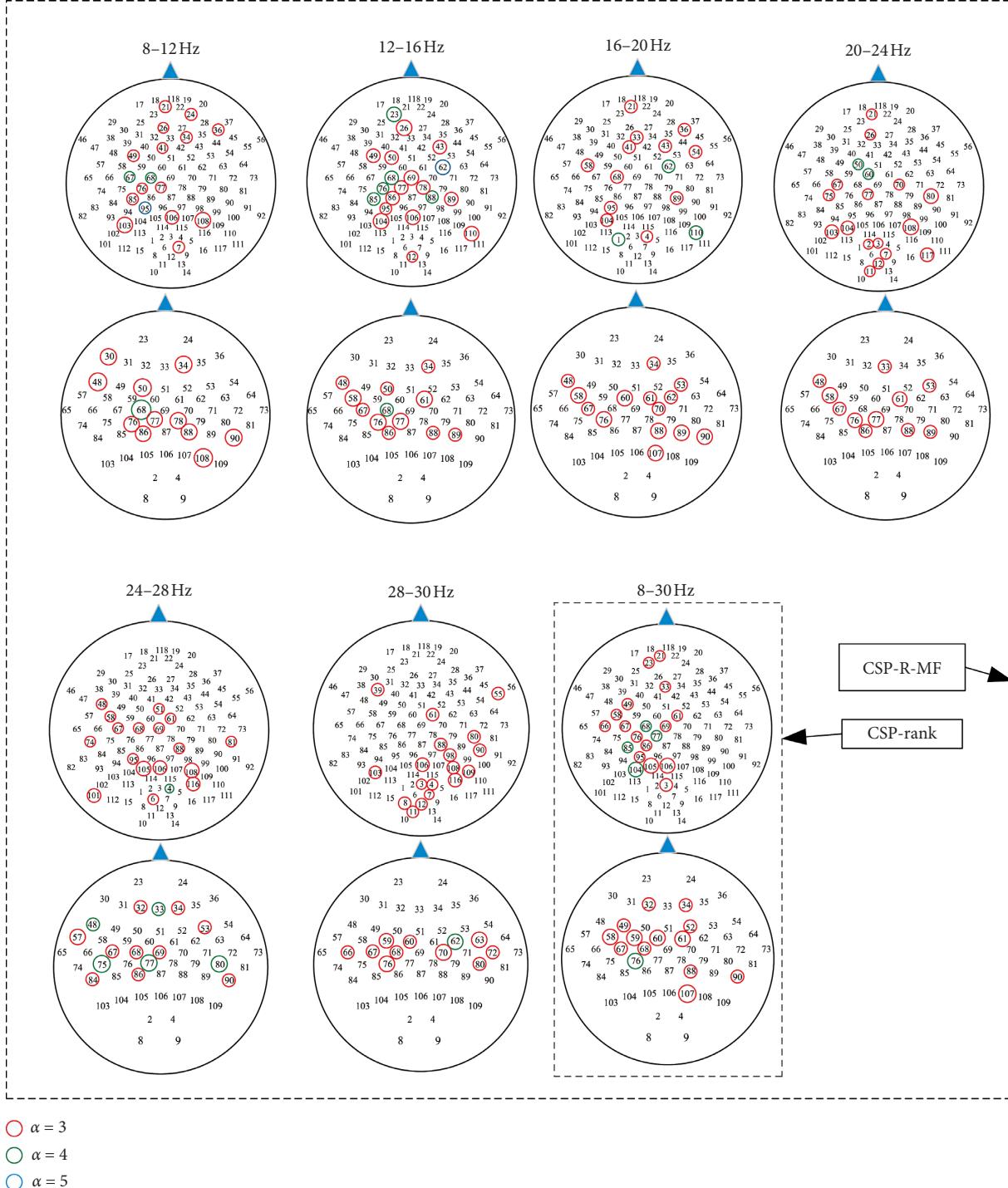


FIGURE 6: The topographic maps of the selected channels for two datasets under different frequency bands. Channels selected more than three times out of all subjects were marked. The channels selected by CSP-R-MF algorithm in 8–30 Hz frequency band were the same as those selected by CSP-rank algorithm. The electrode could be found in Table 2 with the number in topographic maps.

(BCI Competition III datasets IVa) and 7% on dataset 2 (BCI Competition IV dataset 1), when using the proposed method compared to using CSP-rank. In this paper, we set the same stopping criterion for all searching processes based on CSP-rank. However, the stopping criterion may be not exactly the same for different frequency bands and subjects. In future work, to further improve the performance of the CSP-R-MF

algorithm, we will consider setting different stopping criteria for different frequency bands and subjects.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Key Research and Development Program (2017YFB13003002). This work was also supported in part by the National Natural Science Foundation of China under grant nos. 61573142, 61773164, and 91420302 and the Programme of Introducing Talents of Discipline to Universities (the 111 Project) under grant no. B17017.

References

- [1] J. Jin, E. W. Sellers, S. Zhou, Y. Zhang, X. Wang, and A. Cichocki, "A P300 brain-computer interface based on a modification of the mismatch negativity paradigm," *International Journal of Neural Systems*, vol. 25, no. 3, article 1550011, 2015.
- [2] I. Daly, S. J. Nasuto, and K. Warwick, "Single tap identification for fast BCI control," *Cognitive Neurodynamics*, vol. 5, no. 1, pp. 21–30, 2011.
- [3] J. Jin, H. Zhang, D. Ian et al., "An improved P300 pattern in BCI to catch user's attention," *Cognitive Neurodynamics*, vol. 14, no. 3, article 036001, 2017.
- [4] X.-Y. Wang, J. Jin, Y. Zhang, and B. Wang, "Brain control: human-computer integration control based on brain-computer interface approach," *Acta Automatica Sinica*, vol. 39, no. 3, pp. 208–221, 2013.
- [5] J. Jin, B. Z. Allison, E. W. Sellers et al., "Adaptive P300 based control system," *Journal of Neural Engineering*, vol. 8, no. 3, article 036006, 2011.
- [6] T. Yu, Y. Li, J. Long, and Z. Gu, "Surfing the internet with a BCI mouse," *Journal of Neural Engineering*, vol. 9, no. 3, article 036012, 2012.
- [7] D. Thiyam, S. Cruces, J. Olias, and A. Cichocki, "Optimization of alpha-beta log-det divergences and their application in the spatial filtering of two class motor imagery movements," *Entropy*, vol. 19, no. 3, p. 89, 2017.
- [8] J. Feng, E. Yin, J. Jin et al., "Towards correlation-based time window selection method for motor imagery BCIs," *Neural Networks*, vol. 102, pp. 87–95, 2018.
- [9] L. F. Nicolas-Alonso, R. Corrales, J. Gomez-Pilar, D. Alvarez, and R. Hornero, "Adaptive semi-supervised classification to reduce intersession non-stationarity in multiclass motor imagery-based brain-computer interfaces," *Neurocomputing*, vol. 159, pp. 186–196, 2015.
- [10] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. Muller, "The Berlin brain-computer interface: accurate performance from first-session in BCI-naive subjects," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 10, pp. 2452–2462, 2008.
- [11] M. Arvaneh, C. Guan, K. A. Kai et al., "Optimizing the channel selection and classification accuracy in EEG-based BCI," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1865–1873, 2011.
- [12] Z. Qiu, J. Jin, H.-K. Lam, Y. Zhang, X. Wang, and A. Cichocki, "Improved SFFS method for channel selection in motor imagery based BCI," *Neurocomputing*, vol. 207, pp. 519–527, 2016.
- [13] L. Tian, D. Erdogmus, A. Adami et al., "Channel Selection and feature projection for cognitive load estimation using ambulatory EEG," *Computational Intelligence and Neuroscience*, vol. 2007, Article ID 74895, no. 7–8, 12 pages, 2007.
- [14] H. Shan, H. Xu, S. Zhu, and B. He, "A novel channel selection method for optimal classification in different motor imagery BCI paradigms," *BioMedical Engineering OnLine*, vol. 14, no. 1, pp. 1–18, 2015.
- [15] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selection in motor imagery based brain-computer interface," in *Proceedings of the IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 5392–5395, Shanghai, China, September 2005.
- [16] W. K. Tam, Z. Ke, and K. Y. Tong, "Performance of common spatial pattern under a smaller set of EEG electrodes in brain-computer interface on chronic stroke patients: a multi-session dataset study," in *Proceeding of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6344–6347, Boston, MA, USA, 2011.
- [17] D. Gutierrez, "Designing a spatial filter to improve SNR in two-class discrimination problems for BCI applications," in *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, pp. 372–377, Pacific Grove, CA, USA, October 2008.
- [18] J. Armstrong, *Random inter stimulus interval increases signal-to-noise ratio*, East Tennessee State University, Johnson City, TN, USA, 2012.
- [19] M. Arvaneh, C. Guan, K. A. Kai et al., "Spatially sparsified common spatial pattern to improve BCI performance," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2412–2415, Brighton, UK, May 2011.
- [20] F. Lotte and C. Cuntai Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 355–362, 2011.
- [21] S. Selim, M. Tantawi, H. Shedeed, and A. Badr, "Reducing execution time for real-time motor imagery based BCI systems," in *Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016*, Springer International Publishing, Cairo, Egypt, October 2016.
- [22] M. Dai, D. Zheng, S. Liu, and P. Zhang, "Transfer kernel common spatial patterns for motor imagery brain-computer interface classification," *Computational and Mathematical Methods in Medicine*, vol. 2018, Article ID 9871603, 9 pages, 2018.
- [23] C. Park, D. Looney, N. R. Rehman, A. Ahrabian, and D. P. Mandic, "Classification of motor imagery BCI using multivariate empirical mode decomposition," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 21, no. 1, pp. 10–22, 2013.
- [24] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
- [25] M. Arvaneh, C. Guan, K. K. Ang et al., *Optimizing EEG Channel Selection by Regularized Spatial Filtering and Multi Band Signal Decomposition*, Acta Press, Calgary, Canada, 2010.
- [26] I. Rejer, "Genetic algorithm with aggressive mutation for feature selection in BCI feature space," *Pattern Analysis and Applications*, vol. 18, no. 3, pp. 485–492, 2015.
- [27] K. Belwafi, O. Romain, S. Gannouni et al., "An embedded implementation based on adaptive filter bank for brain-computer interface systems," *Journal of Neuroscience Methods*, vol. 305, pp. 1–16, 2018.

- [28] B. A. S. Hasan, J. Q. Gan, and Q. Zhang, "Multi-objective evolutionary methods for channel selection in brain-computer interfaces: some preliminary experimental results," in *Proceedings of the IEEE Congress on Evolutionary Computation*, pp. 1–6, Barcelona, Spain, July 2010.
- [29] K. A. Kai, Y. C. Zheng, H. Zhang, and C. Guan, "Filter Bank Common spatial pattern (fbcsp) in brain-computer interface," in *Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, pp. 2390–2397, Wuhan, China, October 2008.
- [30] Q. Novi, C. Guan, T. H. Dat, and P. Xue, "Sub-band common spatial pattern (SBCSP) for brain-computer interface," in *Proceedings of the 2007 3rd International IEEE/EMBS Conference on Neural Engineering*, pp. 204–207, Kohala Coast, HI, USA, May 2007.
- [31] G. Pfurtscheller, C. Brunner, A. Schlogl, and F. H. Lopes da Silva, "Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks," *NeuroImage*, vol. 31, no. 1, pp. 153–159, 2006.
- [32] S. Selim, H. A. Tantawi, and A. Badr, "A CSP\AM-BA-SVM approach for motor imagery BCI system," *IEEE Access*, vol. 6, pp. 49192–49208, 2018.
- [33] G. Pfurtscheller, "EEG event-related desynchronization (ERD) and synchronization (ERS)," *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 1, p. 26, 1997.