Introducing a high-accuracy brain-computer interface (BCI) for intelligent wheelchairs

Nick Amvazas
Computer Science and Engineering Department
University of Ioannina
Ioannina, Greece
nikosamvazas@gmail.com

Spyridon Moschopoulos
Information Technologies Institute
Centre for Research and Technology Hellas
Ioannina, Greece
spirosmos@iti.gr

Kyriakos Koritsoglou
Computer Science and Engineering Department
University of Ioannina
Ioannina, Greece
kkoritsoglou@uoi.gr

Giorgios Tatsis
Information Technologies Institute
Centre for Research and Technology Hellas
Ioannina, Greece
gtatsis@iti.gr

Ioannis Fudos
Computer Science and Engineering Department
University of Ioannina
Ioannina, Greece
fudos@uoi.gr

Dimitrios Tzovaras
Information Technologies Institute
Centre for Research and Technology Hellas
Ioannina, Greece
Dimitrios.Tzovaras@iti.gr

***Abstract*—We report on the development of a cutting-edge brain-computer interface (BCI) network that leverages the sensor output of a low cost electroencephalogram (EEG) headband to detect specific eye and head movements, enabling intelligent wheelchair navigation. Using a hybrid CNN-LSTM architecture, our method achieves high accuracy classification of these movements while maintaining low inference time on the complete EEG signal. To validate our approach, we conducted a comprehensive experimental evaluation.**

***Index Terms*—brain computer interface, deep learning, wheelchair navigation, EEG**

I. Introduction

Modern neuroscience, combined with the power of EEG signals and machine learning, has become an exciting avenue for investigating human movement. This study not only incorporates EEG signals but also integrates gyroscope data from a cost-effective EEG headband, employing advanced machine learning techniques. The overarching objective is to accurately identify and categorize eye and head movements with the ambition of enabling navigation within an intelligent wheelchair system. Through the utilization of a hybrid CNN-LSTM architecture, our methodology not only achieves high accuracy in classifying these eye and head movements but also exhibits low inference times, making real time application feasible.

ASONAM '23, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
http://dx.doi.org/10.1145/3625007.3627480

II. Related Work

Numerous studies have explored the use of affordable EEG headbands in developing BCI systems. Devices like the Emotiv EPOC, paired with machine learning techniques, have been employed for BCI-driven wheelchair navigation [1]. Similarly, the MindWave Mobile was leveraged by [2] to design a wheelchair navigation simulator. Despite its affordability and user-friendly design, the Muse 2 headband is not commonly employed for real time applications in intelligent wheelchairs [3]. Techniques using multilayer perceptrons(MLP), and Long Short Term Memory (LSTM) networks for classification, achieve high accuracy in areas like attention and emotional state [4], [5]. There have been suggestions for using brainwave-generated blinks to enable the control of household appliances [6]. In [7], [8] BCI-integrated games based on eye blinks and movement show significant promises. This research introduces a hybrid network that interprets signals of eye/head movement from Muse's sensors, with the objective of assessing its potential in a BCI system for navigation in an intelligent wheelchair. The BCI will be integrated with [9] and [10] to form a multi-system for an intelligent wheelchair, further enhancing its capabilities.

III. Detecting Eye and Head Movements

The workflow of the system that performs eye and head movement detection (Figure 1) involves data acquisition, subsequent pre-processing and the development of fundamental neural network architectures for signal classification. The system undergoes an evaluation process that encompasses accuracy and performance assessment.

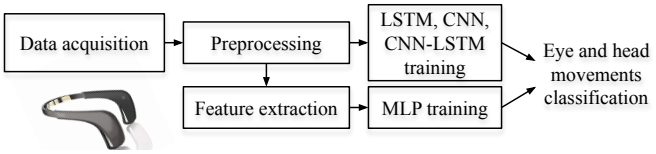

Fig. 1. Workflow diagram for system that performs eye and head movement detection.

A. Muse 2 Headband

Muse 2[1] is a low cost EEG headband that can be connected to a computer or a smartphone via Bluetooth. It is equipped with four EEG sensors placed at frontal and temporal positions (TP9, AF7, AF8, TP10) based on the international 10-20 system. These sensors effectively capture the electrical activity of the brain at a sampling rate of 256Hz. In addition, Muse 2 includes a photoplethysmogram (PPG) sensor with three channels, each sampled at 64Hz. Also, it incorporates an accelerometer and a three-axis gyroscope, each with three channels and a sampling rate of 52Hz. Muse 2 offers the advantage of not requiring conductive gel application, ensuring ease of use and comfortable electrode-to-skin contact. The MuseLSL library [11] was employed for establishing a connection with the Muse device and for acquiring real time sensor data.

B. Datasets

A total of 7 people consisting of 2 females and 5 males between the ages of 27 to 37 volunteered to take part in the data acquisition process. During the procedure, the Muse 2 headband was utilized to record the four EEG channels as well as the X and Z channels of the gyroscope. These particular channels were deemed essential to accurately identify the movements of interest. Participants were instructed to wear the headband comfortably, ensuring proper contact between the sensors and the scalp. Each participant underwent two sessions for each movement type, resulting in a total of six actions per session. Visual and sound feedback was utilized to prompt participants to perform the specified actions. Actions were initiated every 9 seconds to ensure consistent timing throughout the experiment and included 'double blink' (DB), 'look right' (LR), 'look left' (LF), 'look up' (LU), 'look down' (LD), 'rotate head right' (RR), and 'rotate head left' (RL). During the DB action, participants were instructed to perform two fast blinks with both eyes. For LR, LF, LU and LD, participants were required to keep their heads steady and move their eyes as far as possible in the specified direction, before resetting their gaze to the center. In the case of RR and RL, participants were instructed to rotate their heads more than 45 degrees to the right or left, respectively, and then reset them to the center position. They were provided with clear instructions to remain calm, still, and focused between the requested actions and avoid any unnecessary eye or head movements unless explicitly instructed to do so.

A method was developed to select and label the six classes mentioned above by visualizing the recorded signals in timelines and picking relevant segments. Each sample was defined with a three-second window, which served as the upper limit for performing an action. Additionally, two more categories have been added to the final dataset. The first category ('Inactivity') comprises signals recorded during periods between the execution of actions, when participants were calm and focused. The second category ('Unknown') consists of noise signals or signals with irrelevant actions that were not part of the prescribed set of movements. These additions enhance the dataset, providing a comprehensive representation of various signal types. After labeling and extracting all signals corresponding to each category a dataset of 1098 samples was generated (Figure 2). The training set consisted of 70 samples from each class, except the 'Unknown' class, which comprises 140 samples to enhance diversity and improve the generalization on unseen data. Additionally, due to the abundant moments of inactivity during data acquisition, the test set for the 'Inactivity' class was significantly larger than the rest of the classes.

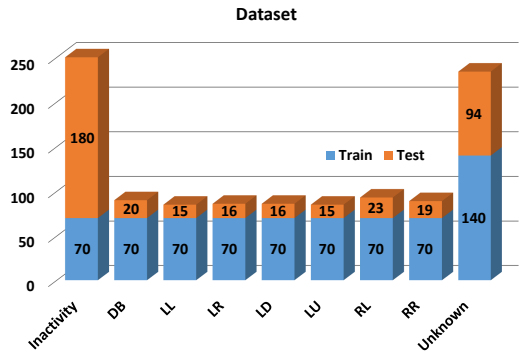

Fig. 2. Train and Test set

In the preprocessing stage, a low-pass filter was applied to the EEG signals to eliminate noise while preserving the useful frequency bands (Delta, Theta, Alpha, Beta). The same low-pass filter is applied to the gyroscope signals despite the absence of high-frequency noise (Figure 3). The filter implementation was based on the Finite Impulse Response (FIR) design technique, employing a cutoff frequency of 30Hz to permit frequencies below this limit to pass through with minimal changes. In addition, to address the DC offset error, a filter was applied to remove the average amplitude of the signal and center it to zero. Finally, the EEG signals were subsampled from 256Hz to 64Hz while the gyroscope signals interpolated

[1]Muse™, InteraXon Inc, Ontario, Canada

from a frequency 52Hz to a frequency of 64Hz. This results in achieving consistency across all channels and also a reduction of computational demands for subsequent stages.

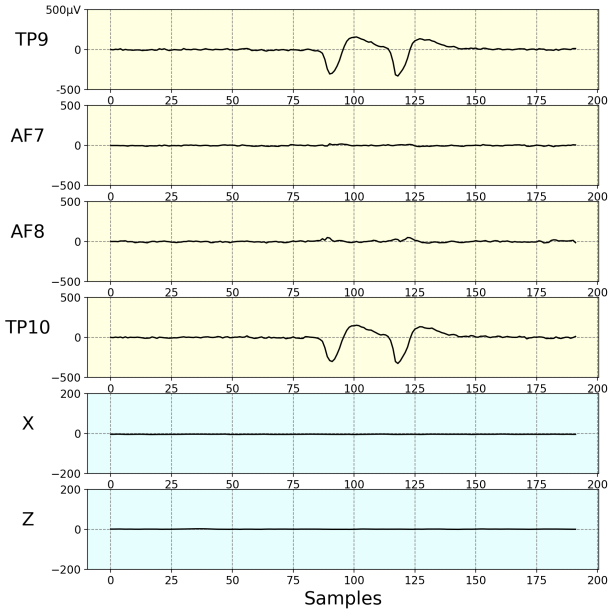

Fig. 3. A preprocessed DB action (3 seconds, 192 values x 6 channels).

One of the classifiers that were used incorporates time-domain features extracted from the processed signals of all six channels. These features encompass the minimum and maximum values of each channel, along with the skewness and kurtosis, capturing vital distribution characteristics [4]. Additionally, the energy and peak-to-peak range for each channel is calculated [5]. Finally, the Hjorth parameters (activity, mobility, complexity) [12] are extracted, contributing diverse information essential for the classification analysis. From the feature extraction process, a vector with 54 elements (6 channels x 9 features) is initialized for each sample.

C. Network Architectures

Four different neural network models were utilized for classification, with each model customized using various parameters like activation functions, dropout layers, and filter sizes to optimize its performance. The set of models used comprises a Multi-Layer Perceptron (MLP), a 1D Convolutional Neural Network (1D_CNN), a Long Short-Term Memory (LSTM) network, and a hybrid 1D CNN and LSTM (1D_CNN-LSTM) (Table I). All models accept as input preprocessed signals with dimensions (192, 6), except for the MLP, which uses extracted features (54) derived from the preprocessed signals.

- **MLP**: Operates on extracted features obtained from preprocessed signals. The architecture consists of multiple Dense layers with ReLU activation, promoting feature learning.
- **1D_CNN**: Captures spatial dependencies in preprocessed signals using Conv1D layers with ReLU activation. These layers are adept at identifying local patterns and relevant features.

TABLE I
NETWORK ARCHITECTURES

	MLP	**LSTM**
Input Shape	(54,)	(192,6)
Hidden Layers	Dense (256, ReLU) Dropout (0.2) Dense (128, ReLU) Dropout (0.2) Dense (64, ReLU) Dropout (0.2)	Lstm(128, tanh) Lstm(64, tanh) Dropout (0.2) Dense (32, ReLU) Dropout (0.2)
Output	Dense (9, SoftMax)	Dense (9, SoftMax)
	1D_CNN	**1D_CNN-LSTM**
Input Shape	(192,6)	(192,6)
Hidden Layers	Conv1D (128, ReLU) MaxPooling (2) Dropout (0.2) Conv1D (64, ReLU) MaxPooling (4) Dropout (0.2) Flatten() Dense (64, ReLU) Dropout (0.2)	Conv1D (128, ReLU) MaxPooling (2) Conv1D (64, ReLU) MaxPooling (4) Dropout (0.2) Lstm (32, tanh) Lstm (16, tanh) Dropout (0.2)
Output	Dense (9, SoftMax)	Dense (9, SoftMax)

- **LSTM**: Utilizes LSTM units to capture sequential patterns and dependencies within the input data. The LSTM units retain information over time, making them well-suited for time-series data like sequential signals.
- **1D_CNN-LSTM**: Harnesses the advantages of both 1D CNN and LSTM architectures. The 1D CNN block specializes in detecting spatial patterns, while the subsequent LSTM layers handle sequential dependencies.

D. Ablation Study

The training data was split into training and validation sets with a 20% validation split. The four networks were trained using the Adam optimizer. To determine the initial learning rate and epoch range for training, we experimented with various values and observed the convergence behavior of the training curves. After the experiments, the learning rate was set to 0.0001, and the number of epochs was set to 200. For every different combination of learning rate and epoch range, accuracies were lower. Also, due to the small size of the training set, a batch size of 16 was selected. Finally, the categorical cross-entropy loss function emerged as the most appropriate choice due to its effectiveness in optimizing classification tasks.

The training curves of all four networks (Figure 4) exhibit training accuracies above 85% for the validation set, and the losses converge to low values without displaying any signs of over-fitting or under-fitting. Also, the training process indicates that the 1D_CNN and 1D_CNN-LSTM networks demonstrate faster learning and achieve higher accuracy in classifying categories compared to the other two networks.

IV. Experimental Evaluation

A. Accuracy

In this study, we evaluated the performance of four neural network models on a test set containing 398 samples. The

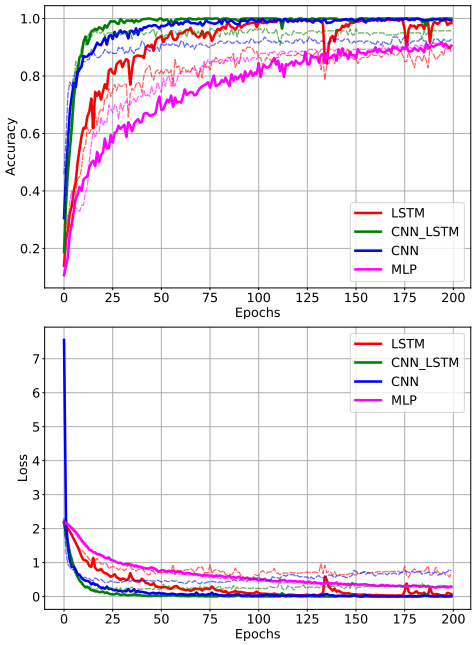

Fig. 4. Accuracy and Loss of training procedure for the four networks (fat lines: training set, faint lines: validation set).

evaluation was based on two key metrics (Table II). F1-score of each class and the total accuracy for the test set with classification probability above 80%.

TABLE II
F1 SCORE AND ACCURACY OF EACH NETWORK

	MLP	LSTM	1D_CNN	1D_CNN-LSTM
	F1 score			
Inactivity	0.98	0.96	0.98	0.99
DB	0.88	0.86	0.83	0.92
LL	0.81	0.88	1.00	1.00
LR	0.84	0.81	1.00	1.00
LD	0.78	0.72	0.97	1.00
LU	0.69	0.59	0.76	0.88
RL	0.91	0.98	1.00	1.00
RR	0.93	0.88	1.00	1.00
Unknown	0.90	0.82	0.88	0.94
	Accuracy			
Total test set	0.859	0.877	0.935	0.967

The 1D_CNN-LSTM network demonstrated the highest performance among all models, achieving an accuracy of 0.967. Furthermore, its F1 scores ranged from 0.88 to 1.00 for each class, indicating consistent and accurate classification of actions. The combination of both CNN and LSTM networks demonstrates superior performance in this dataset by effectively capturing both spatial and temporal dependencies in the EEG and gyroscope signals. The other networks also performed reasonably well, with accuracy values varying between 0.859 and 0.935. As a result, the 1D_CNN-LSTM architecture is recommended for future applications that require accurate and efficient classification of these signals.

B. Performance

In order to evaluate the real-time performance of the four neural networks, we quantized their model weights from 32-bit precision floats to 8-bit precision integers. We conducted tests on a Raspberry Pi 4, a single-board computer that can be easily integrated into a wheelchair due to its minimal power supply and space requirements (3-6 Watt, 56.5 x 85.6 x 11 mm). The Raspberry Pi 4 also supports the MuseLSL library.

The purpose of this evaluation is to assess the suitability of these models for navigation purposes in an intelligent wheelchair through a BCI, where fast execution time and high classifier accuracy are crucial. Quantization had no adverse effects on the accuracy of the models, as they maintained their original performance levels on the test set. Moreover, the mean inference times on the Raspberry Pi 4 were highly efficient (Table III). Also, the signal filtering preprocessing step took approximately 3.81 ms, while feature extraction (only for MLP) required around 9.64 ms.

TABLE III
SIGNAL PROCESSING TIME IN RASPBERRY PI 4.

	MLP	LSTM	1D_CNN	1D_CNN-LSTM
Signal filtering	3.81 ms			
Feature extraction	9.64 ms	-	-	-
Inference time	0.079 ms	47 ms	0.91ms	6.9 ms
Total signal processing time	13.53 ms	50.81 ms	4.72 ms	10.71 ms

The 1D_CNN-LSTM network emerged as the top performer, showcasing faster signal processing time compared to LSTM and MLP models. Although 1D_CNN was faster, its lower accuracy and considering that 10.71 ms per prediction is still tolerable in real-time applications, 1D_CNN-LSTM remains the most suitable choice for a BCI. It excels in accuracy, efficient inference, and satisfactory preprocessing performance, making it ideal for deployment in intelligent wheelchair scenarios.

V. DISCUSSION

The study has identified specific challenges that require further investigation. The constrained dataset size, especially in terms of testing samples, raises concerns about the models' ability to generalize effectively. In scenarios involving individuals outside the training set, potentially a retraining must be employed. To address these limitations in the future a more comprehensive and diverse dataset will be used, allowing for broader model evaluation and improved generalizability. The hybrid CNN-LSTM architecture shows potential for offline performance, but it is important to experimentally evaluate and fine-tune its online behavior to ensure high accuracy in real-time applications.

VI. CONCLUSIONS

Our study presents a cost-effective approach to detecting eye and head movements using an EEG headband. By utilizing machine learning techniques, we were able to achieve accurate results in real-time performance. This method allows for the

navigation of an intelligent wheelchair solely through the use of eye and head movements.

Our experiments have involved networks that analyze the entire EEG signal, rather than a specific set of signal features, in contrast to previous approaches.

Additional experiments are required to confirm the effectiveness of our approach towards navigating intelligent wheelchairs online.

ACKNOWLEDGMENT

This research has been co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call Research—Create—Innovate (project code: T2EDK-02438).

The publication of the article in OA mode was financially supported by HEAL-Link.

REFERENCES

[1] W. Zgallai, J. T. Brown, A. Ibrahim, F. Mahmood, K. Mohammad, M. Khalfan, M. Mohammed, M. Salem, and N. Hamood, "Deep learning ai application to an eeg driven bci smart wheelchair," in *2019 Advances in Science and Engineering Technology International Conferences (ASET)*, 2019, pp. 1–5.

[2] O. R. Pinheiro, L. R. G. Alves, M. F. M. Romero, and J. R. de Souza, "Wheelchair simulator game for training people with severe disabilities," in *2016 1st International Conference on Technology and Innovation in Sports, Health and Wellbeing (TISHW)*, 2016, pp. 1–8.

[3] N. Jamil, A. N. Belkacem, S. Ouhbi, and A. Lakas, "Noninvasive electroencephalography equipment for assistive, adaptive, and rehabilitative brain–computer interfaces: A systematic literature review," *Sensors*, vol. 21, no. 14, 2021.

[4] J. Bird, D. Faria, L. Manso, A. Ekart, and C. Buckingham, "A deep evolutionary approach to bioinspired classifier optimisation for brain-machine interaction," *Complexity*, vol. 2019, 03 2019.

[5] A. Arsalan and M. Majid, "A study on multi-class anxiety detection using wearable eeg headband," *Journal of Ambient Intelligence and Humanized Computing*, vol. 13, 04 2021.

[6] S. Poveda Zavala, J. L. León Bayas, A. Ulloa, J. Sulca, J. L. Murillo López, and S. G. Yoo, "Brain computer interface application for people with movement disabilities," in *Human Centered Computing*, Y. Tang, Q. Zu, and J. G. Rodríguez García, Eds. Springer International Publishing, 2019, pp. 35–47.

[7] G. Prapas, K. Glavas, A. T. Tzallas, K. D. Tzimourta, N. Giannakeas, and M. G. Tsipouras, "Motor imagery approach for bci game development," in *2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, 2022, pp. 1–5.

[8] G. Prapas, K. Glavas, K. D. Tzimourta, A. T. Tzallas, and M. G. Tsipouras, "Mind the move: Developing a brain-computer interface game with left-right motor imagery," *Information*, vol. 14, no. 7, 2023.

[9] G. Tatsis, K. Koritsoglou, I. Fudos, E. Karvounis, K. Votis, and D. Tzovaras, "Design and implementation of obstacle detection system for powered wheelchairs," in *2022 7th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, 2022, pp. 1–4.

[10] S. Moschopoulos, I. Fudos, K. Koritsoglou, G. Tatsis, and D. Tzovaras, "A Real-time Voice Interface for Intelligent Wheelchairs," in *International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET)*. The Eurographics Association, 2023.

[11] A. Barachant, D. Morrison, H. Banville, J. Kowaleski, U. Shaked, S. Chevallier, and J. J. T. Tresols, "muse-lsl," May 2019. [Online]. Available: https://doi.org/10.5281/zenodo.3228861

[12] B. Hjorth, "Eeg analysis based on time domain properties," *Electroencephalography and Clinical Neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.