Endless Runner Game in Virtual Reality Controlled by a Self-paced Brain-Computer Interface Based on EEG and Motor Imagery

Endless Runner Game in Virtual Reality Controlled by a Self-Paced Brain-Computer Interface Based on EEG and Motor Imagery

Pasquale Arpaia[1,2], Antonio Esposito[1], Enza Galasso[1], Fortuna Galdieri[1], Angela Natalizio[3], Marco Parvis[3], Michael Sommeling[1], and Mattia Volpe[1]

[1] Department of Electrical Engineering and Information Technology (DIETI), University of Naples Federico II, Naples, Italy

[2] Interdepartmental Center for Research on Management and Innovation in Healthcare (CIRMIS), University of Naples Federico II, Naples, Italy

[3] Department of Electronics and Telecommunications (DET), Polytechnic University of Turin, Turin, Italy

Abstract. A prototype system for an endless runner game controlled by a self-paced brain-computer interface (BCI) based on electroencephalography (EEG) and motor imagery (MI) is presented. In self-paced BCI systems, brain activity can be distinguished between control and non-control states, allowing the user to continuously engage with the application. The continuous nature of the system enhances the user experience and broadens the experimental setting to more real-world applications. Additionally, metrics for assessing the player's performance during the endless runner game are introduced, including the number of collected coins, distance from the coins, and the efficiency of the avatar's path. The system was evaluated on six subjects with varying levels of experience with MI-BCI. The results demonstrated that the proposed system can be used feasibly with as few as three calibration runs and a highly wearable low-density (8-channel) EEG cap. Furthermore, participants familiar with MI were observed to have better calibration sessions, and subsequently exhibited greater control of the endless runner game.

Keywords: Brain-Computer Interface · BCI · EEG · Motor Imagery · Neurofeedback · Endless Runner · Serious Game · Extended Reality

1 Introduction

Motor imagery (MI) involves the mental simulation of a movement without its physical execution [1]. It activates neural systems and specific brain regions similarly to actual movement. Brain-computer interfaces (BCIs) based on MI exploit this phenomenon, enabling users to control external devices through imagined movements [2]. These systems acquire brain signals, process them with machine learning algorithms, and convert them into control signals. The brain signals are often acquired using non-invasive electroencephalography (EEG), where the recording devices measure the brain electrical activity through electrodes placed on the scalp. EEG-based BCIs are widely used in research due to their non-invasiveness and high temporal resolution [3].

Currently, most MI-BCI studies use synchronous operating modes, where the execution of MI tasks is externally paced by a computer system [1]. These "cue-paced" BCIs limit the autonomy of the users, as they are instructed when to start performing the MI and what type of movement to imagine. Furthermore, the EEG signal has to be analyzed in predefined time windows. An advancement in MI-BCI technology is the development of self-paced systems, which can be controlled by users without any external cues [4]. In this case, the EEG data is continuously analyzed. This autonomy enhances the naturalness and usability of BCIs, making them more suitable for real-world applications. A key element in effective self-paced BCIs is the incorporation of a neurofeedback, which provides users with real-time sensory feedback based on their brain activity [5]. Neurofeedback is crucial for improving user engagement, learning efficiency, and overall BCI performance. MI-BCIs have been used as a promising technology in multiple fields of application, e.g. as assistive tools for disabled people [6], for rehabilitation [7], wheelchair control [8], cursor control [9], and spelling systems [10], as well as for not medical purposes including virtual reality, gaming, robotic arm control, and navigation in 2D and 3D environments [11]. In this framework, virtual reality (VR) can create engaging environments that provide real-time sensory feedback in response to users' MI tasks [12]. VR technologies have been used extensively for entertainment and gaming, and they also hold significant potential for applications in education, medicine, and industry. Integrating VR with self-paced MI-BCIs represents a promising avenue for developing interactive and engaging experiences. In this context, users control the game not through conventional input devices like a mouse, keyboard, or joystick, but by modulating their mental rhythms. This modulation is detected by the BCI and translated into control commands that result in actions in VR or movements of their avatar [4].

Self-paced MI-BCI systems have been previously proposed. However, hybrid systems are often used to enhance system performance and increase the number of available commands. These systems typically combine multiple biosignals, such as near-infrared spectroscopy (NIRS) and EEG [13], or use multiple BCI paradigms in conjunction with MI. Examples include systems that integrate motor imagery with P300 potentials [14] or steady-state visual evoked potential [15]. Voluntary artifacts, such as eye blinking, are also used to this end [16,17]. Another limitation of current systems is the number of electrodes required to acquire EEG signals, which typically ranges from 10 to 30 electrodes [17,18]. This increases the system setup time and limits its usability. Finally, previous studies on self-paced MI-BCI systems often disregarded detailed data about online usage, thus limiting a thorough performance assessment [16], [19].

In this paper, a two-level endless runner VR game of increasing difficulty is proposed. This is controlled via a self-paced MI-BCI. To ensure the exploitability of the data acquired during the game, all in-game events (e.g., appearance time and position of targets, player position, targets collected/missed) were synchronized with the EEG and stored for a performance assessment regarding online usage. Additionally, the proposed system is highly wearable since an EEG device with only eight electrodes was employed. This reduced setup complexity and user discomfort. Finally, control commands were managed exclusively through motor imagery and relaxation states of the user, thus enhancing the naturalness and ease of interaction.

The paper is organized as follows: Sec. 2 details the architecture of the BCI system and the developed virtual environments, Sec. 3 presents the results of the study and provides a discussion of the findings, and conclusions are outlined in Sec. 4.

2 Materials and Methods

2.1 System Overview

The proposed system architecture is shown in Fig. 1. The EEG signal from the partecipant is acquired with a wearable cap. These raw signals are then continuously sent to a computer via Bluetooth. On the same computer, Matlab scripts are used for processing the data offline and analogous Simulink models are used online. Finally, Unity applications are employed either during the calibration phase and the gaming phase. The developed virtual environments were shown to the participant through the PC screen. In order to use the processed data as an input signals, a UDP connection is exploited to connect Matlab/Simulink and Unity. At first, a communication string is sent from Simulink to Unity for establishing a connection. After this, Simulink seamlessly listens for Unity data consisting of timestamps and game-related metrics. During the gaming phase, Simulink also sends to Unity control signals resulting from the EEG online processing.

The timestamps from Unity are used for marking key events such as the beginning, the end, and in-game events, while the metrics are used for assessing participants' performances during the gaming phase. These data from Unity are properly matched with the EEG data received in Simulink. Whenever the Unity app is closed, the data of the entire process is stored. Communication latencies between the different system components were estimated to be below 100 ms.

Fig. 1: Architecture of the self-paced brain-computer interface integrated with a virtual reality game.

2.2 EEG Device

The Unicorn Hybrid Black cap (g.tec medical engineering GmbH, Austria) was used to acquire the EEG signals. This involves eight hybrid EEG electrodes, namely the signal can be acquired with or without the use of conductive gels. In this study, wet electrodes were used. The electrodes were placed according to the international 10-20 EEG system at the Fz, FC3, FC4, C5, Cz, C6, CP3, and CP4 standard positions as highlighted in Fig. 2. This configuration is proposed to cover the sensorimotor area, mostly involved in MI phenomena [20]. The ground and reference were placed on the participant's right and left mastoid, respectively, by using disposable surface electrode pads. Data were collected at a sampling rate of 250 Sa/s and with a resolution of 24 bits.

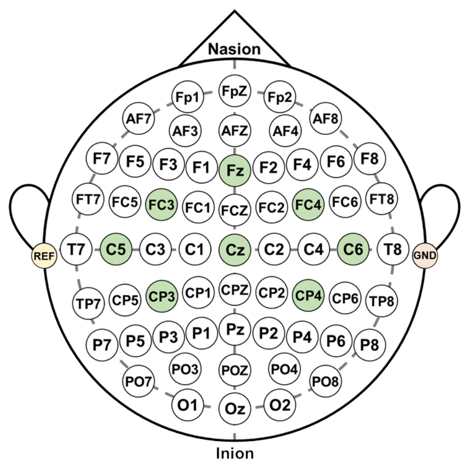

Fig. 2: Electrodes configuration for the EEG acquisition device.

2.3 Calibration and Gaming Applications

Two different applications were developed and used during the experiments. Firstly, a calibration application was used for gathering cue-paced data, which are needed for training the EEG online processing algorithm. Next, a gaming application was used for exploiting the processed EEG to control an avatar. In addition, the gaming application allowed to assess the user's performance during a self-paced experiment. Both applications were created via Unity Development Platform[4].

The calibration application was used to display cues and instructions on the screen during the EEG acquisition. After opening the application, a connection

[4] www.unity.com

message notified the user on the status of the connection with the Simulink model. The model was adopted for acquiring the EEG, labeling it, and saving the data in accordance with the cued instructions. In order to acquire data during various MI tasks, the calibration phase was divided into runs consisting of several trials. The sequence of scenes related to a single trial within the calibration is represented in Fig. 3. At the beginning of each trial, a fixation cross was displayed on the screen for 2.00 s. Such a cross was used to indicate an idling state during which the user should focus the gaze. This was followed by a visual cue, displayed for 1.25 s and accompanied by the "GET READY" indication. The cue consisted of an arrow pointing to the left or to the right direction, indicating to imaging left or right upper limb movements, respectively. Subsequently, the MI task had to commence and last for 3.00 s. This was pointed out by the on-screen "GO!" in substitution of the previous "GET READY". Following the mental task, the user could relax for a variable time window of 2.00 s to 3.00 s. At the end of a complete run, an acoustic cue informed the participant about its conclusion.

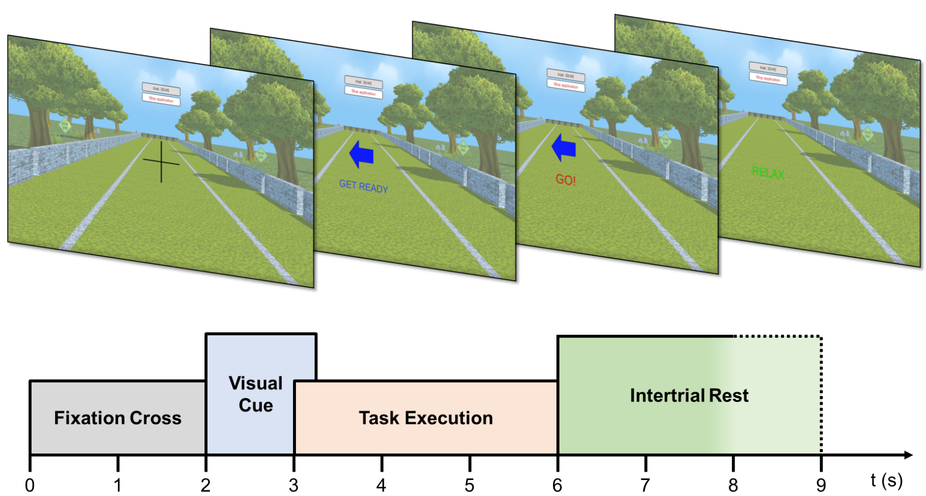

Fig. 3: Timeline and scenes associated with the application used during the calibration phase.

On the other side, the gaming application consisted of a virtual environment in which an avatar resembling a ghost could be moved horizontally in a road confined by walls and divided into three lanes. The gaming interface is shown in Fig. 4. In the top right corner, a red text next to a numerical value constantly monitored the user's score throughout the gaming session. On the other hand, in the top left corner a similar indication tracked the playing time. The user's objective was to achieve the highest score possible, in a limited period of time, by collecting coins spread among the lanes. The score went up by 10 every time the avatar picked up a coin. Differently from the majority of MI-based BCI applications, the avatar could be moved with continuity in a self-paced manner.

Fig. 4: Virtual interface with the brain-controlled avatar exploited during the gaming phase.

The proposed game has a multi-level structure. In the first level, coins are about 10 s apart, so that the user has sufficient time to move the avatar between two successive coins. Moreover, this first level has a fixed duration of 100 s. The second level presents similar characteristics but, in order to increase the difficulty, coins were 5 s apart and the road sliding speed was changed dynamically in proportion to the in-game score. The aim of being adaptive is to make the experience challenging but not frustrating, so as to maintain high engagement. In this case, due to the higher effort required to the user, the level could last up to 80 s.

2.4 Experimental Protocol

The experimental protocol consisted of two different phases that took place on the same day: a calibration phase and a gaming one. Before the actual experiments took place, participants were verbally informed about the procedures and the purpose of the test. Subsequently, the EEG cap was mounted on their head and they did not remove the cap until the end of the whole experiment.

The calibration phase was required for training the system on the user's EEG patterns. Each session consisted of three runs, each comprising 45 trials and lasting approximately 6 min. Between two consecutive runs, an additional minute was required to properly save the acquired data and let the user additional relax time. The whole calibration session lasted about 30 min. During each run, participants were asked to imagine movements or remain in an idle state depending on the visual cue appearing on the screen. The user could imagine the preferred left and right upper limb movements (e.g: grasping, squeezing, lifting weights, or moving the whole arm) with the constraint to be consistent over both phases. Each trial was carried out as presented in Sec. 2.3 and each kind of cue was presented 15 times per run in random order. This was designed to prevent the user from getting used to a fixed presentation order of the visual cues and avoid predicting the next one. A total of 135 trials were thus acquired in the calibration phase of a single participant.

The gaming phase, instead, involved the use of the virtual reality application with the avatar. The users controlled the avatar by modulating their MI-related brain activity thus implying the movement of the ghost accordingly. Fig. 5 shows an example of an user playing with the avatar in virtual reality, as well as a representation of the different scenes of the application.

Fig. 5: Gaming scenes and crisp example: (a) user playing with the avatar, (b)-(c)-(d)-(e) timeline of the scenes associated with the gaming application. Remarkably, the subfigure (a) involves the scene also shown in subfigure (d).

2.5 Signal Processing

Offline Processing The data collected during the three runs of the calibration phase were processed offline with two main purposes related to the subsequent phase: (i) finding the optimal time window associated with the best classification accuracy, and (ii) finding the optimal features. In accordance with previous studies [21], the EEG signals collected in a single trial were analyzed with a sliding window of 2.00 s shifting of 0.25 s. In order to increase the statistical reliability of the results, a k-fold cross-validation [22] was exploited to calculate the classification accuracy window-by-window. This consisted of randomly splitting the data into k equally sized folds (or partitions) and then applying the classification algorithm k times, each time holding out one of the subsets from the training and using it for validation. In the current study a k-value of 5 was used.

The pipeline used to process the data is based on the Filter Bank Common Spatial Pattern (FBCSP) [23] algorithm for feature extraction, and on a Minimal-Redundancy-Maximal-Relevance (MRMR) for features selection [24]. The acquired EEG signals were filtered channel-by-channel in four non-overlapped standard bands ($8-12$ Hz, $12-18$ Hz, $18-28$ Hz, and $28-40$ Hz corresponding to μ, low β, high β, and low γ bands, respectively) [18] by using band-pass type II Chebyshev filter modules with a band-stop attenuation of 20 dB. The Common Spatial Patterns (CSP) method was used to create spatial filters that increase the separability between two classes by maximizing the variance of band-pass filtered EEG signals from one class, while minimizing their variance from the other classes [25]. The pre-processed EEG signals were projected into a new vector space defined by the filters using the linear transformation matrix given by CSP. The number of selected CSP filter pairs for each frequency band was set to three and then the time-varying log-variances of each CSP-filtered epoch were calculated. Next, a combination of Linear Discriminant Analysis (LDA) [26] was employed to classify the extracted and selected features. Three binary classifiers were used, namely: left vs rest (L vs X), right vs rest (R vs X), left vs right (L vs R). The global outcome of the classification relied on the outputs of the three binary classifiers, but this was not used in the offline analysis.

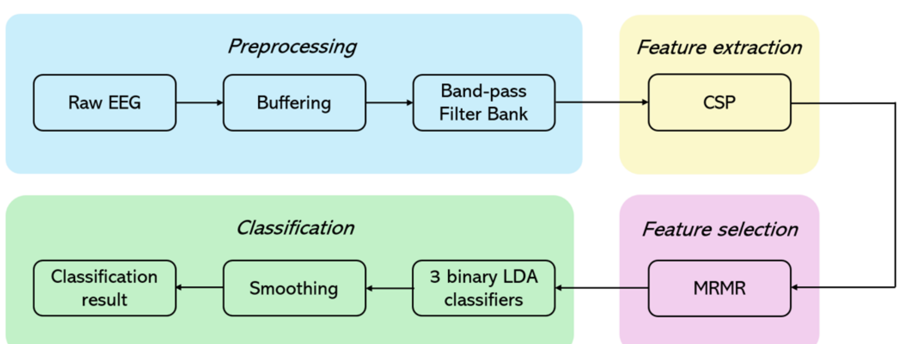

Fig. 6: Online version of the proposed FBCSP classification algorithm; CSP: Common Spatial Pattern; MRMR: Minimal-Redundancy-Maximal-Relevance; LDA: Linear Discriminant Analysis.

Online Processing The module for EEG data acquisition and online processing was developed in Simulink[5] and it consisted of an online version of the already described offline processing. For each channel, a 2.00 s epoch of EEG data was buffered and processed, while the time window sliding was 0.25 s. Previous investigations [27] demonstrated that the optimal classification results are obtained when the epoch size is about 2 seconds. This implies that the feedback from the avatar's movement is only provided after such an interval has elapsed and it has been analyzed. To account for this latency, the coins were placed at a minimum distance of 8 s from each other, so as to provide sufficient time for the system to properly process the EEG signal. The online processing algorithm is depicted in Fig. 6. In the figure, the same steps of the offline pipeline appear, while the raw data were acquired during the game. As an additional step for the online processing, the outputs of the LDA classifiers were sent to a moving average filter to smooth out short-term fluctuations. This served to increase the robustness of classifiers outputs. Finally, while in the offline analysis the classifiers outputs could be considered separately, the online usage required a final decision about the global outcome in order to send a command to the avatar. The adopted rule was:

- if the outputs of two classifiers were 'right', the output command was 'go right';
- if the outputs of two classifiers were 'left', the output command was 'go left';
- in any other case, the predicted class was 'rest'.

Ultimately, the avatar moved by following the prediction results.

2.6 Evaluation Metrics

During each level of the game, EEG data were complemented by recording the position of the avatar, the position of the coins on screen, and labels about collected or missed coins. Storing this information was useful for ex-post analysis of participants' performance, even by a synthetic metric. The metric proposed for performance evaluation was named *CoinError* and it was calculated starting from the distances between the avatar and each coins when the avatar was closest. If the avatar took the coin, the inherent distance was forced to zero. In the remaining cases, the Euclidean distance was taken into account. The synthetic metric was obtained by averaging and then normalizing such distance, so that the metric values ranged between 0 and 1, with lower values indicating better performance. A standard deviation could be associated to the mean too. The mathematical expression of the synthetic metric (mean value) is

$$CoinError = \frac{1}{maxError}\frac{1}{N}\sum_{n=1}^{N}\sqrt{(x_{coin,n} - x_{avt,n})^2 + (y_{coin,n} - y_{avt,n})^2} \quad (1)$$

[5] Simulink for Matlab (The MathWorks, Inc.) 2024.

where $maxError$ is the maximum possible mean distance between ideal path and avatar path, N is the number of coins, and the sum considers the point-by-point distance between the closest position of the avatar with respect to the n-th coin and the avatar position.

2.7 Questionnaires

The results gathered through the above described Simulink-Unity connection are objective data useful for assessing the system performance and identifying critical issues. On the other side, the subjective user experience was taken into account with respect to the proposed system. For this reason, two questionnaires were administered to each participant immediately after the completion of the experimental protocol: the SUS (System Usability Scale) and the NASA-TLX (NASA-TaskLoadIndex) questionnaires.

The SUS questionnaire is currently one of the most widely used psychometric tools to assess users' perceptions regarding systems usability. It consists of 10 sections with 5-point scales: "1" corresponds to "strongly disagree" and 5 to "completely agree." Each section questioned the subject regarding various aspects of the proposed system: its complexity and laboriousness, the possible need for expert support during its use, inconsistency among the functions presented, level of confidence perceived and predisposition to use the system frequently.

The NASA-TLX is also employed as a tool to test the subjective workload perceived by subjects participating in an activity. In this paper, the questionnaire is presented without a weighting scheme for greater simplicity and intuitiveness. The NASA-TLX questionnaire has six sections: mental demand, physical demand, temporal demand, performance, effort and frustration. Subjectively, each subject can indicate, on a 20-point scale, their estimate for the corresponding section.

3 Results and Discussion

3.1 Participants

For a preliminary evaluation of the performance of the proposed self-paced MI-BCI, six healthy subjects (S01 to S06) participated in the experiment. The participants, all males, were aged between 25 and 32. No subject had motor problems, brain injuries, or other medical conditions of interest for the experiments. The experiments were carried out at the Augmented Reality for Health Monitoring Laboratory (ARHeMLab), University of Naples Federico II, Italy. Two participants (S03, S04) had no prior experience with BCIs, three participants (S02, S05, S06) had used a MI-BCI at least once before, and one participant (S01) had experience with various BCI paradigms.

3.2 Calibration Phase

The acquisition of EEG data during the calibration phase was used to identify the optimal time window and number of features for training the online processing

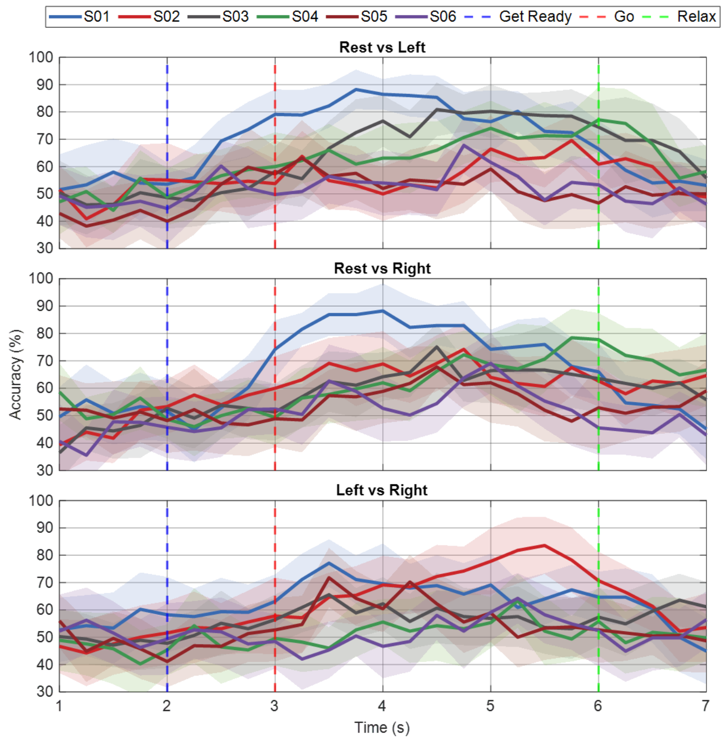

Fig. 7: Time trend of the cross-validation accuracies for each subject. Each bounded line indicates the mean and standard deviation of the cross-validation accuracies calculated with a sliding time window.

algorithm. Due to inter-subject variability in brain signals, acquiring data from scratch was necessary for each subject.

Fig. 7 shows the cross-validation accuracy for each subject across the three pairs of tasks. In accordance with Sec. 2.5, a 2.00 s-wide window shifted by 0.25 s was used to obtain the time trend of the classification accuracy. Bounded lines representing the mean cross-validation accuracy and standard deviation across trials are shown. Noteworthy time instants are also marked in the figures, namely the cue appearance (blue), the starting of the task (red), and the ending of the task with starting of the relax (green).

Peak accuracy was mainly achieved within the window related to the task. Per each subject, the 2.00 s-wide window associated with maximum accuracy was selected to train the algorithm. Overall, the highest classification accuracies were obtained with up to 25 features, after which the accuracies remained roughly constant for all three classifiers.

Tab. 1 summarizes the classification accuracies for the three binary classifications for each subject. The reported results refer to the classification accuracies obtained in cross-validation with the best time window and the best number of features for each classifier and subject. They are reported in terms of mean accuracies and associated standard deviations.

Table 1: Mean and standard deviation of the cross-validation accuracies calculated for each subject during the calibration phase.

Mean Accuracy ± Standard Deviation			
Subject	***Rest vs Left***	***Rest vs Right***	***Left vs Right***
S01	88 % ± 7 %	88 % ± 10 %	77 % ± 8 %
S02	69 % ± 9 %	74 % ± 9 %	83 % ± 10 %
S03	80 % ± 9 %	75 % ± 8 %	65 % ± 10 %
S04	77 % ± 11 %	78 % ± 10 %	63 % ± 9 %
S05	63 % ± 13 %	67 % ± 7 %	71 % ± 10 %
S06	67 % ± 10 %	69 % ± 11 %	64 % ± 9 %

As shown in both Tab. 1 and Fig. 7, subjects S01 and S02 achieved high accuracies across all three couple of tasks. Subjects S03 and S04 achieved good accuracies in rest vs. left MI and rest vs. right MI, but not in left MI vs. right MI. Subject S05 achieved good accuracy in left MI vs. right MI, but not in the other two cases. Subject S06 did not achieve good accuracy in any task. Therefore, it was expected that subjects S01 and S02 would exhibit good control of the player during the subsequent evaluation session, while the other subjects could control it worse.

3.3 Gaming Phase

Tab. 2 presents the performance of each subject evaluated on the gaming phase. The table reports the number of collected coins out of the total number of available one, as well as the *CoinError* metric introduced in Sec. 2.6. Next to subjects' names, "L1" and "L2" indicate the level (first or second) for which results are reported. At this preliminary stage, only subjects with good control at the first level were allowed to proceed to the second level. Hence, S01 was the only subject to advance to the second level of the game. Indeed, subject S01 collected the highest number of coins during the first level (66 %). This outcome was traced back to the fact that S01 achieved a relatively high calibration accuracy greater for all the three binary classifiers. Next, during the second level, S01 continued to exhibit good control, collecting 6 out of 16 coins. The number of collected coins is actually complemented by the *CoinError* metric to give a better indication of how close the avatar was to collecting coins during the game. Such a metric indicates that the actual performance of the self-paced BCI is better than what the mere success rate suggests. Notably, S02, S04, and S05 collected the same number of coins, but S02 and S04 were closer to collecting more of them. Moreover, even S01 correctly controlled his MI by getting closer to all the coins, despite the fact that there were actually six coins in both levels.

Table 2: Performance of each subject during the gaming phase.

Run	Collected Coins	Coin Error
S01_L1	6/9 (66 %)	0.06 ± 0.10
S01_L2	6/16 (37 %)	0.23 ± 0.23
S02_L1	3/9 (33 %)	0.31 ± 0.34
S03_L1	2/9 (22 %)	0.41 ± 0.30
S04_L1	3/9 (33 %)	0.31 ± 0.24
S05_L1	3/9 (33 %)	0.62 ± 0.50
S06_L1	1/9 (11 %)	0.62 ± 0.40

As a further insight, Fig. 8 provides a reconstruction of the online events with subject S01. The avatar's trajectory is shown in blue, the ideal trajectory in dashed red, missed coins in black, and collected coins in yellow. From the graph, it is evident that the player often approached the coins but did not get close enough to collect them. Additionally, there was evidence of trajectory correction during both levels, indicating excellent control.

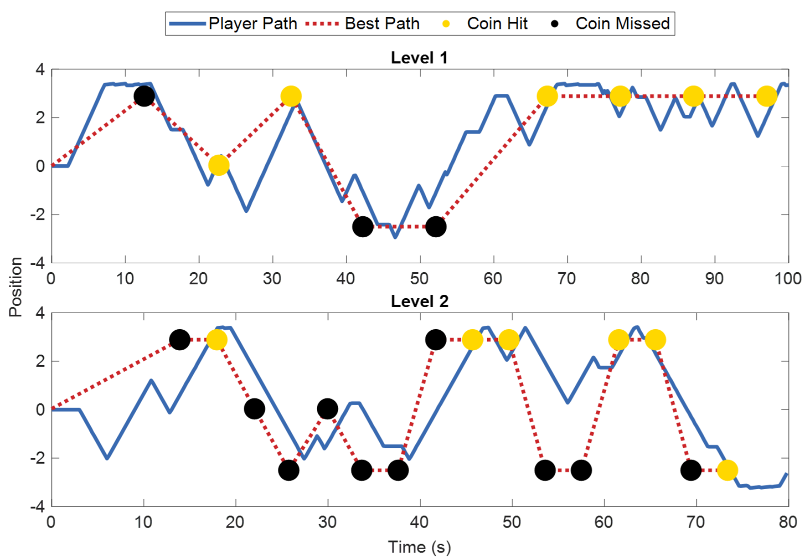

Fig. 8: Example of a reconstructed trajectory for subject S01.

Overall, the subjects' performance during the gaming phase appears correlated with the calibration accuracy as expected. Specifically, higher algorithm accuracy in distinguishing all three class pairs led to improved recognition of the subjects' mental states and consequently better player control.

3.4 User-perceived system usability

Tab. 3 and Tab. 4 report the scores given by participants when they were administered the System Usability Scale (SUS) and NASA-TLX questionnaires. Regarding the SUS, the lowest average score concerns the inconsistency of features presented by the system. Other low scores concern the complexity of the system and the need for expert user support. Although no score, reached average values equal to 4 - 5, participants showed a positive attitude toward frequent use of the system and by a potentially wider audience.

On the other hand, the NASA-TLX scores show that physical workload, temporal demand, and sense of frustration were not significant for the subjects. Conversely, the required mental workload was definitely perceived as high. Instead, the perceived effort and performance were rated as medium ones according to the scale presented in the questionnaire.

Table 3: Scores given by each subject to all sections of the SUS questionnaire.

Sections	S01	S02	S03	S04	S05	S06	Mean ± std
Repeated usability	4	3	3	4	4	5	4 ± 1
Unnecessary complexity	2	1	2	2	2	3	2 ± 1
Ease of use	3	5	4	3	3	2	3 ± 1
Need for support	2	2	1	2	2	4	2 ± 1
Features integration	3	3	4	4	4	3	3 ± 1
Features inconsistency	1	1	1	2	1	2	1 ± 1
Ease for wide audience	4	5	4	4	3	2	4 ± 1
System laborousiness	1	3	2	3	3	4	3 ± 1
Level of confidence	4	2	3	3	1	3	3 ± 1
System intuitiveness	2	1	1	4	4	2	2 ± 1

3.5 Discussion

This study proposed an endless runner game in VR controlled by a self-paced MI-BCI. The game consists of two levels of increasing difficulty, although only one user was able to play the second level. The system appeared user-friendly due to the wearability of the adopted EEG device and due to an intuitive gamified

Table 4: Scores given by each subject to all sections of the NASA-TLX questionnaire.

Sections	S01	S02	S03	S04	S05	S06	Mean ± std
Mental demand	16	13	6	17	16	18	14 ± 4
Physical demand	3	8	2	12	8	4	6 ± 4
Temporal demand	7	12	4	14	4	3	7 ± 5
Performance	5	8	14	10	18	11	11 ± 4
Effort	15	4	4	17	18	12	12 ± 6
Frustration	4	13	1	13	12	8	8 ± 5

virtual interface. Moreover, the naturalness in using the interface was leveraged by solely relying on the users' motor imagery and state of relaxation, while avoiding the usage of voluntary artifacts recognition or other BCI paradigms.

The electrodes type played a crucial role. Indeed, using dry electrodes (no gels) should be preferred for user comfort, but preliminary experiments confirmed that achieving a proper signal-to-noise ratio is challenging and prevents proper functionality of the interface. This implies that a substantial investigation of artifact removal in this context would be needed [28]. Meanwhile, electrodes with gel ensured a better and consistent signal quality over time and allowed to still minimize the setup burden thanks to a limited number of electrodes.

The proposed system was designed to allow a comprehensive reconstruction of the in-game events. This approach allows for the eventual re-use of data and addresses a common issue in self-paced BCIs, namely the inability to re-use unlabeled data acquired during the online phase [29].

The results from both the calibration and gaming phases confirm that an effective usage of MI-BCIs requires a properly calibrated system. Moreover, users with distinguishable brain patterns are required, especially for discriminating between left MI and right MI. Indeed, experienced users like S01 can achieve meaningful control of the avatar because of the capability of voluntarily modulate brain activity. As a rule of thumb, successful gaming sessions and meaningful control of the avatar were achieved when calibration accuracies overcame 70 % for each classifier. Consequently, user training could enhance the overall performance [30]. In these regards, transfer learning techniques will be also investigated in the future to reduce (or ideally zero) the calibration phase.

The proposed $CoinError$ metric ultimately allowed a deeper understanding of user control over the avatar despite the collected coins. This can be used as supplementary information to coin collection accuracy to summarize online performance. Moreover, inherent results suggest possible improvements, such as adjusting the application to collect the coin not only when the avatar hits the target but also when it is close enough to that. This would also help more users reaching the second level of the gaming without being frustrated. Finally, more extended reality games could be developed to enhance the immersiveness of the gameplay experience.

Finally, opinions given by users about their subjective experience showed that the system was generally positively perceived. Although future improvements should enhance the intuitiveness of the system, this already appeared as user-friendly and consistent in its features. Despite this, prolonged use of the system resulted in high mental fatigue in all subjects.

4 Conclusion

Self-paced paradigms represents a significant advance in the field of active brain-computer interfaces based on motor imagery. In these paradigms, the user has the autonomy to decide when and what to imagine while neural data are continuously analyzed. Such features enhance the naturalness and usability of the system, making them quite suitable for real-world and extended reality applications.

The accessibility of the proposed system leverages on the wearability and low number of electrodes with which brain signals are acquired, as well as the development of virtual applications that can engage the user in the use of these systems. Notably, this study proposed a prototypical virtual reality game controlled by imagery and relaxation states, in which electroencephalographic data are acquired via a wearable cap involving only eight wet electrodes. The gaming application consists of an endless running avatar with the objective of collecting scattered coins along the road. Moreover, the game was structured in levels of increasing difficulty.

Tests carried out with six subjects yielded encouraging outcomes. As expected, the performance of a subject during the online gaming phase was correlated with the quality of the data acquired during the calibration phase and the inherent calibration accuracy. In particular, higher classification accuracies in distinguishing the three class pairs (involving rest, left motor imagery, right motor imagery) led to better avatar control. Achieved results and limitations appeared reasonable since self-paced control is usually suitable for users capable of modulating brain rhythms.

It is important to note that the work proposes a preliminary study, whereby only six young, male subjects were recruited. It has been hypothesised that they are supposed to have less hair and thus favour electrode adhesion to the scalp, as they are more likely to want to play with an app gaming. This allowed the functionality of the system to be tested, but future work will focus on enlarging the number of subjects for experimentation, to give greater significance to the results achieved.

The introduction of the metric “CoinError” also allowed deeper insights on the user’s performance during the gaming phase. This metric appeared as more closely related to the ability to modulate the motor imagination than the mere number of collected coins and it pointed out possible improvements for the next game application versions. Moreover, although the proposed metrics are closely related to the gaming application, the outcomes are promising in terms of reusing the data collected during the online phase. This notably paves the way to labeling online data and re-used them for further analysis and/or system re-calibration.

To enhance the robustness of the system, future works will focus on achieving greater control over the avatar by improving the online classification. This will be mainly investigated by focusing on transfer learning approaches, with the aim to ease the reuse of previous data (from previous sessions and/or other subjects). Indeed, avoiding to calibrate from scratch at each system usage would also reduce the mental workload. A further improvement could be the utilization of dry electrodes, but degradation of signal quality must be prevented to avoid affecting classification performance.

Acknowledgements

This work was financially supported by Italian Ministry for Universities and Research (MUR) through the project "AGE-IT - Ageing Well in an ageing society: A novel public-private alliance to generate socioeconomic, biomedical and technological solutions for an inclusive Italian ageing society (Spoke: Ageing and clinical practice)", PNRR PE8, CUP E63C22002050006.

Sommeling gratefully acknowledges financial support for this project by the Fulbright U.S. Student Program, sponsored by the U.S. Department of State, the U.S.-Italy Fulbright Commission, and the Fondazione Con il Sud. The paper contents are solely the responsibility of the authors and do not necessarily represent the official views of the Fulbright Program, the Government of the United States, the U.S.-Italy Fulbright Commission, or the Fondazione Con il Sud.

References

1. A. Singh, A. A. Hussain, S. Lal, and H. W. Guesgen, "A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface," *Sensors*, vol. 21, no. 6, p. 2173, 2021.
2. P. Arpaia, A. Esposito, A. Natalizio, and M. Parvis, "How to successfully classify eeg in motor imagery bci: a metrological analysis of the state of the art," *Journal of Neural Engineering*, vol. 19, no. 3, p. 031002, 2022.
3. R. A. Ramadan and A. V. Vasilakos, "Brain computer interface: control signals review," *Neurocomputing*, vol. 223, pp. 26–44, 2017.
4. B. Alchalabi, J. Faubert, and D. R. Labbe, "A multi-modal modified feedback self-paced bci to control the gait of an avatar," *Journal of Neural Engineering*, vol. 18, no. 5, p. 056005, 2021.
5. C. Jeunet, B. Glize, A. McGonigal, J.-M. Batail, and J.-A. Micoulaud-Franchi, "Using eeg-based brain computer interface and neurofeedback targeting sensorimotor rhythms to improve motor skills: Theoretical background, applications and prospects," *Neurophysiologie Clinique*, vol. 49, no. 2, pp. 125–136, 2019.
6. 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.
7. K. K. Ang and C. Guan, "Eeg-based strategies to detect motor imagery for control and rehabilitation," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 4, pp. 392–401, 2016.

8. J. Li, J. Liang, Q. Zhao, J. Li, K. Hong, and L. Zhang, "Design of assistive wheelchair system directly steered by human thoughts," *International journal of neural systems*, vol. 23, no. 03, p. 1350013, 2013.
9. D. D. Chakladar and S. Chakraborty, "Multi-target way of cursor movement in brain computer interface using unsupervised learning," *Biologically Inspired Cognitive Architectures*, vol. 25, pp. 88–100, 2018.
10. L. Cao, B. Xia, O. Maysam, J. Li, H. Xie, and N. Birbaumer, "A synchronous motor imagery based neural physiological paradigm for brain computer interface speller," *Frontiers in human neuroscience*, vol. 11, p. 274, 2017.
11. R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain–computer communication: motivation, aim, and impact of exploring a virtual apartment," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 4, pp. 473–482, 2007.
12. G. Riva, B. K. Wiederhold, and F. Mantovani, "Neuroscience of virtual reality: from virtual exposure to embodied medicine," *Cyberpsychology, behavior, and social networking*, vol. 22, no. 1, pp. 82–96, 2019.
13. B. Koo, H.-G. Lee, Y. Nam, H. Kang, C. S. Koh, H.-C. Shin, and S. Choi, "A hybrid nirs-eeg system for self-paced brain computer interface with online motor imagery," *Journal of neuroscience methods*, vol. 244, pp. 26–32, 2015.
14. Y. Yu, Z. Zhou, Y. Liu, J. Jiang, E. Yin, N. Zhang, Z. Wang, Y. Liu, X. Wu, and D. Hu, "Self-paced operation of a wheelchair based on a hybrid brain-computer interface combining motor imagery and p300 potential," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 12, pp. 2516–2526, 2017.
15. E. Dong, H. Zhang, L. Zhu, S. Du, and J. Tong, "A multi-modal brain–computer interface based on threshold discrimination and its application in wheelchair control," *Cognitive Neurodynamics*, vol. 16, no. 5, pp. 1123–1133, 2022.
16. L. Yang and M. M. Van Hulle, "Real-time navigation in google street view® using a motor imagery-based bci," *Sensors*, vol. 23, no. 3, p. 1704, 2023.
17. J. Choi, K. T. Kim, J. H. Jeong, L. Kim, S. J. Lee, and H. Kim, "Developing a motor imagery-based real-time asynchronous hybrid bci controller for a lower-limb exoskeleton," *Sensors*, vol. 20, no. 24, p. 7309, 2020.
18. A. Korik, K. McCreadie, N. McShane, N. Du Bois, M. Khodadadzadeh, J. Stow, J. McElligott, Á. Carroll, and D. Coyle, "Competing at the cybathlon championship for people with disabilities: long-term motor imagery brain–computer interface training of a cybathlete who has tetraplegia," *Journal of NeuroEngineering and Rehabilitation*, vol. 19, no. 1, p. 95, 2022.
19. Y. Yu, Z. Zhou, E. Yin, J. Jiang, J. Tang, Y. Liu, and D. Hu, "Toward brain-actuated car applications: Self-paced control with a motor imagery-based brain-computer interface," *Computers in biology and medicine*, vol. 77, pp. 148–155, 2016.
20. G. Pfurtscheller and F. L. Da Silva, "Event-related eeg/meg synchronization and desynchronization: basic principles," *Clinical neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999.
21. P. Arpaia, D. Coyle, F. Donnarumma, A. Esposito, A. Natalizio, and M. Parvis, "Visual and haptic feedback in detecting motor imagery within a wearable brain–computer interface," *Measurement*, vol. 206, p. 112304, 2023.
22. A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *International journal of remote sensing*, vol. 39, no. 9, pp. 2784–2817, 2018.

23. K. K. Ang, Z. Y. Chin, C. Wang, C. Guan, and H. Zhang, "Filter bank common spatial pattern algorithm on BCI competition IV datasets 2a and 2b," *Frontiers in neuroscience*, vol. 6, p. 39, 2012.
24. J. Pohjalainen, O. Räsänen, and S. Kadioglu, "Feature selection methods and their combinations in high-dimensional classification of speaker likability, intelligibility and personality traits," *Computer Speech & Language*, vol. 29, no. 1, pp. 145–171, 2015.
25. F. Jamaloo and M. Mikaeili, "Discriminative common spatial pattern sub-bands weighting based on distinction sensitive learning vector quantization method in motor imagery based brain-computer interface," *Journal of medical signals and sensors*, vol. 5, no. 3, p. 156, 2015.
26. S. N. Abdulkader, A. Atia, and M.-S. M. Mostafa, "Brain computer interfacing: Applications and challenges," *Egyptian Informatics Journal*, vol. 16, no. 2, pp. 213–230, 2015.
27. P. Arpaia, A. Esposito, N. Moccaldi, A. Natalizio, and M. Parvis, "Online processing for motor imagery-based brain-computer interfaces relying on eeg," in *2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2023, pp. 01–06.
28. G.-L. Li, J.-T. Wu, Y.-H. Xia, Q.-G. He, and H.-G. Jin, "Review of semi-dry electrodes for eeg recording," *Journal of Neural Engineering*, vol. 17, no. 5, p. 051004, 2020.
29. J. D. Cunha, S. Perdikis, S. Halder, and R. Scherer, "Post-adaptation effects in a motor imagery brain-computer interface online coadaptive paradigm," *IEEE Access*, vol. 9, pp. 41 688–41 703, 2021.
30. V. Kaiser, G. Bauernfeind, A. Kreilinger, T. Kaufmann, A. Kübler, C. Neuper, and G. R. Müller-Putz, "Cortical effects of user training in a motor imagery based brain–computer interface measured by fnirs and eeg," *Neuroimage*, vol. 85, pp. 432–444, 2014.