Computer Methods and Programs in Biomedicine 255 (2024) 108332

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.sciencedirect.com/journal/computer-methods-and-programs-in-biomedicine

Design of a brain-machine interface for reducing false activations of a lower-limb exoskeleton based on error related potential

P. Soriano-Segura [a,b], M. Ortiz [a,b,*], E. Iáñez [a,b], J.M. Azorín [a,b,c]

[a] Brain-Machine Interface Systems Lab, Miguel Hernández University of Elche, Spain
[b] Engineering Research Institute of Elche - I3E, Miguel Hernández University of Elche, Spain
[c] Valencian Graduate School and Research Network of Artificial Intelligence–ValgrAI, Valencia, Spain

ARTICLE INFO

Keywords:
Brain-Machine Interface (BMI)
Error Related Potential (ErrP)
EEG signals
Exoskeleton
Neurorehabilitation

ABSTRACT

Background and objective: Brain-Machine Interfaces (BMIs) based on a motor imagination paradigm provide an intuitive approach for the exoskeleton control during gait. However, their clinical applicability remains difficulted by accuracy limitations and sensitivity to false activations. A proposed improvement involves integrating the BMI with methods based on detecting Error Related Potentials (ErrP) to self-tune erroneous commands and enhance not only the system accuracy, but also its usability. The aim of the current research is to characterize the ErrP at the beginning of the gait with a lower limb exoskeleton to reduce the false starts in the BMI system. Furthermore, this study is valuable for determining which type of feedback, Tactile, Visual, or Visuo-Tactile, achieves the best performance in evoking and detecting the ErrP.

Methods: The initial phase of the research concentrates on detecting ErrP at the beginning of gait to improve the efficiency of an asynchronous BMI based on motor imagery (BMI-MI) to control a lower limb exoskeleton. Initially, an experimental protocol is designed to evoke ErrP at the start of gait, employing three different stimuli: Tactile, Visual, and Visuo-Tactile. An iterative selection process is then utilized to characterize ErrP in both time and frequency domains and fine-tune various parameters, including electrode distribution, feature combinations, and classifiers. A generic classifier with 6 subjects is employed to configure an ensemble classification system for detecting ErrP and reducing the false starts.

Results: The ensembled system configured with the selected parameters yields an average correction of false starts of 72.60 % ± 10.23, highlighting its corrective efficacy. Tactile feedback emerges as the most effective stimulus, outperforming Visual and Visuo-Tactile feedback in both training types.

Conclusions: The results suggest promising prospects for reducing the false starts when integrating ErrP with BMI-MI, employing Tactile feedback. Consequently, the security of the system is enhanced. Subsequent, further research efforts will focus on detecting error potential during movement for gait stop, in order to limit undesired stops.

1. Introduction

In the recent years, the efforts in the field of neurorehabilitation have contributed to the raise of new technologies for the rehabilitation and assistance of patients with spinal cord injuries (SCI).

Brain-Machine Interfaces (BMI) are among the emerging technologies, as they are capable of decoding brain patterns from electroencephalographic (EEG) signals into commands to execute actions in other devices, such as robotic devices [1,2]. The combination of a BMI with an exoskeleton is a promising approach, not only because of the active implication of the subject in their rehabilitation process, but also its potential for neuroplasticity recovering in patients with motor impairments [3]. For these reasons, BMI based on Motor Imagery paradigm (BMI-MI) are preferred for this application, where the patient imagines how the movement feels and, the exoskeleton gives them the feedback of the real movement [4,5]. Furthermore, a BMI-MI functioning asynchronously allows patients to perform movements voluntarily whenever they decide, rather than being constrained by external factors, such as tasks and cues.

However, the effectiveness of BMI-MI for gait rehabilitation do not meet the desired standards, since the brain regions in charge of lower limbs control, such as the primary motor cortex and supplementary

* Corresponding author at: Brain-Machine Interface Systems Lab, Miguel Hernández University of Elche, Spain.
E-mail address: mortiz@umh.es (M. Ortiz).

https://doi.org/10.1016/j.cmpb.2024.108332
Received 11 April 2024; Received in revised form 8 July 2024; Accepted 17 July 2024
Available online 18 July 2024

motor area, are located deep within the interhemispheric fissure [6]. Consequently, the collected signals are often low quality, making it difficult to decode the commands [7,8]. Thus, the safety of the patient is compromised, affecting its practical use in the clinical setting.

Some studies have explored invasive methods to address this issue, which involve implanting electrodes directly into the brain to target these deep brain regions. These methods show a promising improve in signal quality and control accuracy [9]. Nevertheless, these types of invasive interfaces are often discarded due to the risks associated with surgical implantation and future damages, making them less viable for widespread clinical use.

However, the application of Error Related Potentials (ErrP) for self-tuning the wrong commands has recently gained significant interest as a potential solution to tackle this challenge and enhance the systems' accuracy with non-invasive methods [10–16]. Hence, an ErrP is a brain response elicited when the subject detects an error in the system performance.

For instance, in the context of using a BMI with a lower limb exoskeleton for gait rehabilitation, an ErrP can be elicited when the exoskeleton performs an undesired movement. A user wants to remain stationary, but the exoskeleton suddenly starts moving, or the user wants to keep walking, but the exoskeleton unexpectedly stops. In these cases, the user's brain detects the discrepancy between the intended action and the exoskeleton's response, triggering an ErrP.

Generally, this evoked potential manifests as two positive and two negative peaks appearing approximately within the first half second after the error occurrence, and it remains stable over time [11,14,17]. Nevertheless, the shape and timing of the ErrP may change depending on the stimulus used to elicit it.

In the literature, the most common approach for evoking the ErrP is visually [10,18,19], by means of monitoring tasks of cursors or robots, where the potential appears when the movement is not in the desired direction. Another method, presented by [20] is adding tactile feedback to the monitoring tasks, where vibration bracelets inform the participants about the direction before the cursor moves in the screen. In addition, some authors proposed combining both stimuli [21], tactile and visual, to elicit the ErrP. As a result of their research, all of them coincide in the fact that Tactile stimulus, individually or in combination, produces a ErrP potential with higher amplitude, but it also produces a slight delay in the presentation of the ErrP [21,22]. Additionally, studies [23,24] analyze the ErrP when is elicited with the proprioceptive feedback of an exoskeleton.

The aim of the current research is to enhance the efficiency of an asynchronous BMI-MI to control a lower limb exoskeleton during gait by means of the ErrP. In a first stage of this investigation, an experimental protocol is designed to evoke the ErrP at the beginning of the gait, where the subject is standing still and must activate the exoskeleton by imagining walking. Therefore, three different stimuli are studied to elicit this error potential: Tactile, Visual and Visuo-Tactile.

Furthermore, an iterative selection process is employed to characterize ErrP, in both time and frequency domain, and to fine-tune diverse parameters, including electrodes distribution, features combinations or classifiers. This process evaluates these different parameters iteratively, and their combinations, in order to identify the optimal set that reaches the highest scores. As a result of all this process, an ensembled classification system to detect the ErrP at the start of the gait is configured.

2. Materials and methods

In this section, the protocol designed to evoke ErrP with different stimuli and the materials employed to carry it out are described. Then, the iterative selection process implemented to select the best parameters of the ensembled system to discriminate the error from the correct commands is explained in detail.

2.1. Subjects

In the current experiments, 6 subjects have participated: 3 females and 3 males, with ages ranging from 20 to 32 years old (25.5 ± 3.64). All participants were able-bodied and did not take any type of medication. In addition, the experiments were explained to them, and they provided written informed consent in accordance with the Helsinki Declaration. The study was approved by the Responsible Research Office of Miguel Hernández University of Elche (Spain) (DIS.JAP.03.18).

2.2. Experimental protocol

This research is focused on detecting the Error Related Potential (ErrP) to self-tune the wrong activation commands in a BMI based on a motor imagery paradigm [7]. The protocol looks for generating the evoked potentials after an activation command is issued, but still has not been executed by the exoskeleton, assessing the EEG signals to identify the evoked potential as a correct activation by the BMI (NoErrP) or a wrong one (ErrP).

In the experiment, the participants must perform two different mental tasks while the exoskeleton is standing: keep an idle state (relax) or imagine how the first step of the gait feels in their own muscles and joints when the exoskeleton activates (kinesthetic motor imagery).

Throughout the experimental procedure, data is labelled during four different periods of time depending on the mental task (Relax or Imagination) the subject is performing and the period before and after feedback (stimulus appearance). When the stimulus is not provided, there are correct relax and erroneous imagination periods and when the feedback activates, erroneous relax and correct imagination periods. Feedback appearance is randomized in a time between 4 s and 8 s after the cue, remaining for 2 s (Fig. 1).

Therefore, the ErrP is evoked when the classifier detects an activation during the relax task, against subjects' will. In the literature, diverse authors have reported that an error ratio higher than a 30 % causes a degradation of the potential, because the subject gets used to the error situation and perceives it as normal [10]. For this reason, all trials are already set up from the beginning of the experiment with a 70 % of correct events and 30 % of error events and, consequently, the subject will never be in control of the system.

Furthermore, eliciting the ErrP through the proprioceptive feedback by the exoskeleton was not considered as possible feedback, as its control does not allow to reverse an activation command without completing a step and the objective was to correct false activations before they occur. Hence, the event must be evoked by an external stimulus (Tactile, Visual, Visuo-Tactile) that activates 2 s before the exoskeleton begins to walk, to inform the subject about the imminent movement. Consequently, a potential is elicited, which may be classified as either an ErrP or a NoErrP depending on the task is being performed. Concretely, Visual feedback delivers a blue light within the subject's field of view, while with Tactile feedback the subject experiments a vibration in both wrists. Meanwhile, VisuoTactile feedback emits both stimuli simultaneously. Obviously, the evoked potential caused by the stimulus will appear in both ErrP and NoErrP events. However, the analysis focuses not on the stimulus-evoked potential itself, but rather on identifying the variations between the responses to error or correct events.

A session consists of 12 trials with 10 interleaved mental tasks, each with 5 relax and 5 motor imagery instructions. Thus, the experiment is similar to the future real-time application where, after a Relax phase, a Motor Imagery task is always expected to activate the exoskeleton. In addition, performing the same mental task consecutively could cause mental fatigue in the subject. Among these 120 mental tasks, the correct periods randomly occur 70 % of the time and 30 % erroneously [10]. Consequently, a total of 18 periods for each error event and 42 periods for each correct event will be registered.

Notice that only the periods associated with an evoked potential are

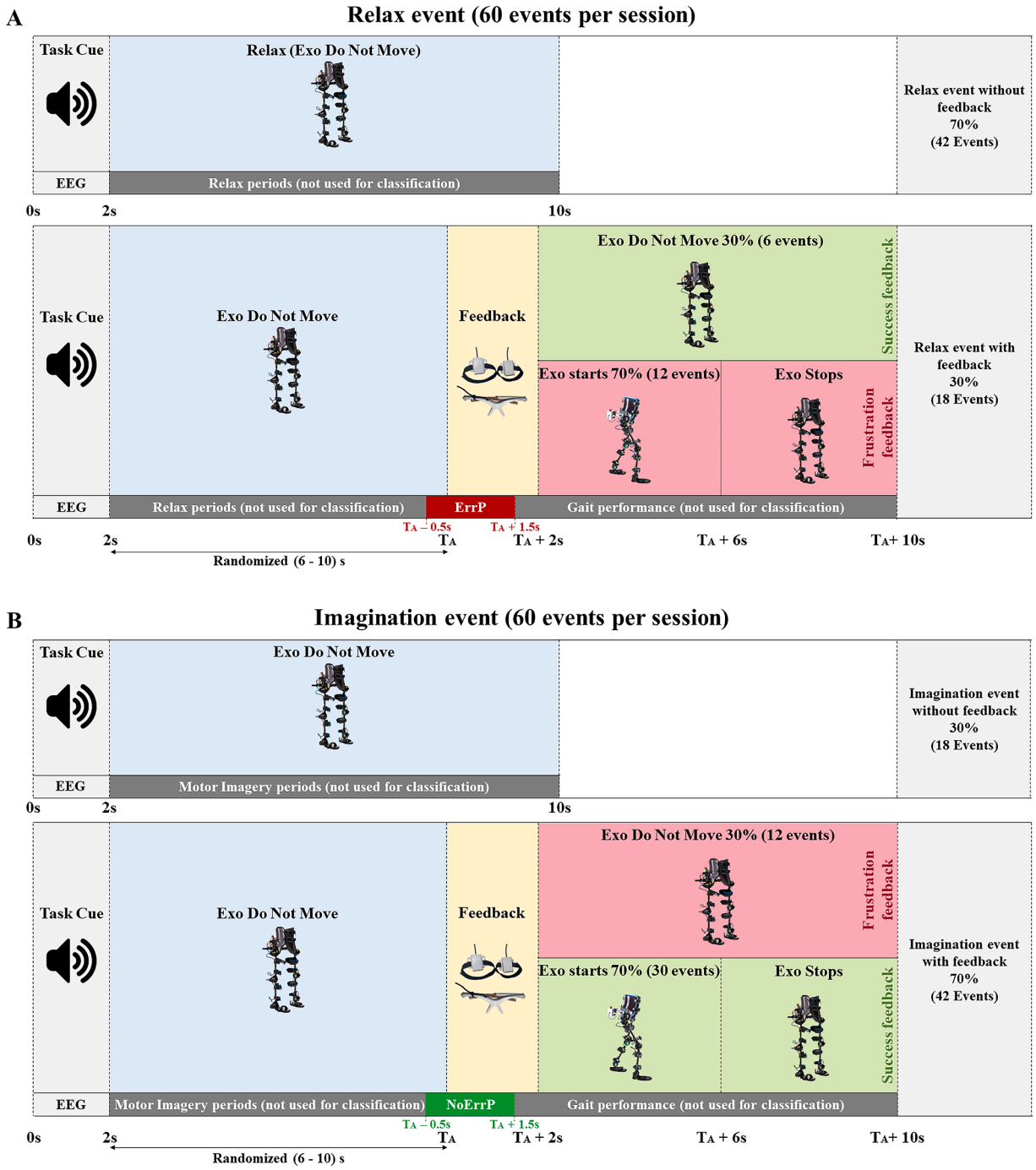

Fig. 1. Experimental protocol to elicit ErrP potential for activations. A: Relax event, B: Imagination event. Periods before feedback (blue) represent data matched to the mental task and are not used for classification. Thus, exoskeleton no-activation events are correct Relax and erroneous Motor Imagery, and they present a constant duration of 10 s. A 30 % of Relax events and a 70 % of Imagination events, the feedback activates (yellow) for 2 s to evoke the potentials, which can be an ErrP (dark red) in Relax event or a NoErrP (dark green) in an Imagination event. The EEG signals of these potentials are used for classification in a window ranged between 0.5 s before feedback turns on and 1.5 s after. In these cases, the mental task (blue) duration (T_A) is variable and randomized between 6 s and 10 s until the stimulus appearance. After feedback turns off, the exoskeleton starts walking (green) for a gait cycle (4 s) and then stops in a 70 % of the events when the feedback activates, while in the remaining 30 %, the exoskeleton does not move (light red). Thus, the randomized ratio of exoskeleton activation provides frustration/success feeling in a realistic way.

relevant for the study: erroneous relax for ErrP and correct imagination for NoErrP (red and green periods in EEG blocks of Fig. 1). The other periods are needed to enhance the realism of the experiment and prevent the subject from an expected activation of the exoskeleton. Actual exoskeleton activation after the feedback is also randomized at a 30/70 % ratio to provide a realistic feeling of success/frustration.

Each participant performed three sessions, one per type of feedback: Tactile, Visual and Visuo-Tactile, and the order of the sessions was randomized. Each session lasts approximately 45 min, with an additional 45 min for instrumentation. In order to prevent fatigue, each session took place on different days.

2.3. Equipment

In this research, the brain signals are recorded using the g.Nautilus EEG electrodes cap (g.tec medical engineering GmbH, Austria), operating at a sample rate of 500 Hz. The cap presents 28 non-invasive wet electrodes following the international system distribution of 10–10: AF3, F3, FZ, FC3, FC1, FCZ, C5, C3, C1, CZ, CP3, CP1, CPZ, P3, PZ, PO3, AF4, F4, FC2, FC4, C2, C4, C6, CP2, CP4, P4, POZ, PO4, where the reference is set on the right ear lobe and the ground on AFZ electrode. Furthermore, electrooculographic (EOG) signals are registered by means of 4 additional electrodes: VU (Vertical Up), VD (Vertical Down), HL (Horizontal Left) and HR (Horizontal Right).

The subjects' movements are assisted by the H3 exoskeleton (Technaid, Spain), which receives commands to start and stop walking through Bluetooth.

The feedback is elicited by a pair of glasses and two bracelets, one per wrist. Inside the glasses, manufactured by 3D printing, a blue LED is integrated within the users' field of view to produce visual feedback. Two vibration motors are embedded inside the bracelets to provide tactile feedback. The behavior of these gadgets is programmed on an Arduino Uno R3 (Arduino Org, Italy). Fig. 2 shows a subject wearing equipments prepared for starting an experiment.

2.4. Signal processing

The registered EEG signals go through different filters to remove noise and irrelevant information, emphasizing only those bands where the ErrP is clearer. First, a high pass filter of 1 Hz eliminates the noise of lower frequencies, which can affect negatively in the next filters. Then, Hinfinity (H_∞) filter removes artifacts caused by eye blinks and movements from signals [25]. Suddenly, the spatial filter Common Average Reference (CAR) is applied to mitigate common interferences and enhance the specificity of neural activity. Finally, a band pass filter between 1 Hz and 10 Hz is employed to focus the analysis on the specific frequency range for ErrP.

The evoked potential that appears under ErrP and NoErrP periods is analyzed through different features in a windows time from [−0.5 to 1.5] seconds of the feedback appearance, although each of them can use a different lapse within it.

Fig. 2. Illustration of a subject during an experimental session.

2.5. Characterization of ErrP

Multiple types of features have been considered to determine which are the more suitable to identify the ErrP potentials in both time and frequency domain regarding NoErrP evoked potentials. The feature definition approach is based on the characterization of ErrP events that has been defined in the literature [10,11,18,19,26]. This definition will be tested against the evoked potential that appears under correct activations, being the hypothesis that the evoked potential that appears under a correct activation (NoErrP) is different than the evoked potential that appears under a wrong activation (ErrP).

Additionally, time-frequency features, such as the Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT), were initially considered as features too. However, since the BMI system operates with a shift interval of 0.5 s, all steps (preprocessing, feature extraction, classification, etc.) must be completed well within this timeframe. During preliminary tests, obtaining these features alone exceeded the 0.5 s shift interval, which significantly increased the system's latency. For this reason, they were neglected for this study.

2.5.1. Time domain features

In time domain, multiple features describing the shape of ErrP are explained in detail within three different groups: ErrP peaks information at P300 and N500, the signal window shape and statistical values. a) Evoked Response Peaks Features

Since the ErrP is an evoked response which has been defined in the literature as a signal with increments and decrements at specific time points [10], for instance a positive peak around 300 ms or a negative peak at approximately 500 ms, a group of features that describe the potential shape and time appearance is extracted.

The theory reports that around 300 ms after the feedback, a positive peak is shown (P300) and then a deflection is produced around 500 ms (N500) [11], as it is illustrated in Fig. 3. The signals are filtered between 1 Hz and 10 Hz and P300 and N500 are looked in two windows after the stimuli: [200–500]ms and [300, 600]ms. Both peaks are characterized in value and time as *TimePeaksP300N500* and *AmplitudPeaksP300N500*.

On the one hand, *TimePeaksP300N500* presents three values: T_{P300}, T_{N500}, $T_{P300—N500}$, where T_{P300} (yellow) and T_{N500} (pink) are the time point of P300 and N500 peaks while $T_{P300—N500}$ (blue) is the time difference between both peaks. On the other hand, *AmplitudPeaksP300N500* is also expressed with three values: A_{P300}, A_{N500}, $A_{P300—N500}$, where A_{P300} (yellow) and A_{N500} (pink) are the amplitudes of the signal from 0μV to the positive or negative peak amplitude value. Then, $A_{P300—N500}$ (blue) is the total amplitude between both peaks. Furthermore, a feature called *PeaksP300N500* combines both peaks information with the six values described for each event (Table 1). b) Optimized Sampling Frequency

This feature uses the own shape of the transient as a feature. To limit the size of the feature vector, the signal is downsampled to the half (125 Hz). The feature is referred to as *Downsampling*. c) Statistical Features

Moreover, a variety of statistical features are extracted with the aim of characterizing the signal, such as the mean, median, standard deviation, complexity, mobility, kurtosis and energy. These formulas are applied to a 1 s segment of the signal starting from the instant when the stimulus turns on. Then, the resulting features vector is named as *StatsFeatures* and has a length of seven samples.

2.5.2. Frequency domain features

Exploring features in frequency domain, two different approaches have been employed: Welch and Fast Fourier Transform (FFT). On the one hand, Welch divides the signal into segments of 500 samples (twice the sampling rate) with overlap of 250 samples and estimates the smoothed Power Spectral Density (PSD), a feature referred to as *WelchPSD*. Furthermore, the feature *PeaksPSD* characterizes this PSD in diverse bands as delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta low (12–20 Hz), beta high (20–30 Hz) by calculating the power

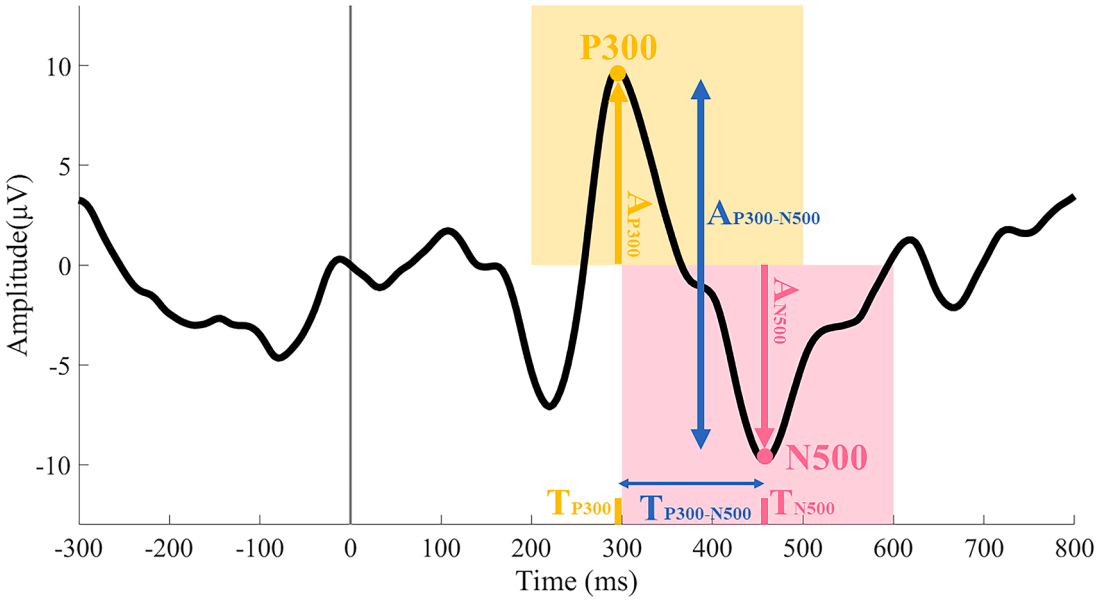

Fig. 3. Features extracted from the evoked potential in time (T) and amplitude (A). The yellow region indicates the signal segment where the maximum (P300) is extracted, and the pink region indicates where the minimum (N500) is extracted. Additionally, in yellow, features related to the P300 (TP300, AP300) are shown, in pink, features of the N500 (TN500, AN500), and in blue, the difference between the characteristics of each peak (TP300-N500, AP300-N500).

Table 1
Features considered for the classification of the feedback evoked potentials. They are based on time and frequency domain analysis of the EEG data in different periods of time matched to the feedback activation.

Feature abbreviation	Feature	Feature Domain	Band Pass Filter	Window size	Vector length (samples)
PPN	PeaksP300N500	Time	1 – 10 Hz	[0.2, 0.6] s	6
APPN	AmplitudePeaksP300N500	Time	1 – 10 Hz	[0.2, 0.6] s	3
TPPN	TimePeaksP300N500	Time	1 – 10 Hz	[0.2, 0.6] s	3
DS	Downsampling	Time	1 – 10 Hz	(0, 1.5] s	188
SF	StatsFeatures	Statistics	1 – 10 Hz	(0, 1.5] s	7
PS	PowerSpectrum	Frequency	1 – 10 Hz	(0, 1.5] s	187
WPSD	WelchPSD	Frequency	1 – 250 Hz	(0, 1.5] s	68
PPSD	PeaksPSD	Frequency	1 – 250 Hz	(0, 1.5] s	10

summatory within each band and identifying the maximum power reached in each respective range. On the other hand, the direct application of the FFT followed by the absolute square, referred to as *PowerSpectrum*, is computationally more efficient than Welch but also more sensible to noise.

2.6. Classifiers

In this research several classifiers are tested: Support Vector Machine (SVM), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Logistic Regression (LR), Stochastic Gradient Descent (SGD), Perceptron, Multi-Layer Perceptron (MLP), Gaussian Naïve Bayes (NB), Decision Tree (DT), Gradient Boosting (GB), and Random Forest

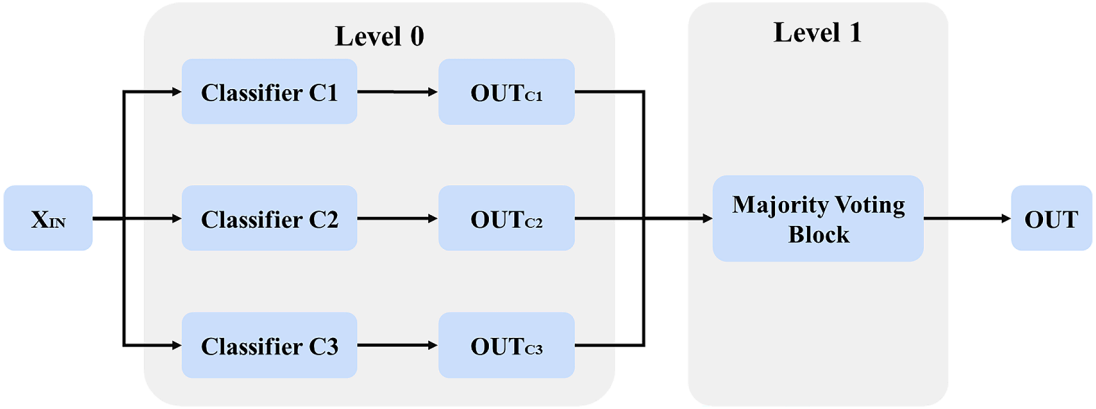

Fig. 4. Two level classification used in the research. X_{IN} is the features vector and is used as input for the Level. Level 0 is composed of 3 classifiers (C1, C2, C3) and each one produces a prediction output (OUT_{C1}, OUT_{C2}, OUT_{C3}). These Level 0 outputs are the inputs for the Level 1 majority Voting Block, whose output (OUT) is the definitive prediction of the system for a given features vector.

(RF). Classification is established at two levels using groups of three different classifiers and majority voting, see Fig. 4.

2.7. Metrics

The current binary classification system to detect ErrP discriminates between two classes. Consequently, the predictions are determined as:

- True Positives (TP): The actual class is ErrP, and it is correctly predicted as ErrP.
- False Positives (FP): The actual class is NoErrP, but the classifier wrongly detects it as ErrP.
- True Negatives (TN): There is no ErrP, and the classifier confirms the absence of error.
- False Negatives (FN): There is an ErrP, but the classifier fails to detect it.

As a result, the classical metrics commonly used in binary classification models are estimated from these four prediction possibilities.

Accuracy is a fundamental metric in any system evaluation, measuring the overall correct detections made by the model:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

Precision, or positive predictive value, indicates the percentage of correctly identified ErrP instances out of all ErrP predictions, whether right or wrong:

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall, or True Positive Ratio (TPR), estimates the proportion of correctly predicted ErrP samples out of the total actual ErrP samples:

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

F1-Score, or Effectively, is a combination of Precision and Recall in one single metric:

$$F1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

However, the aim of the ErrP detection is to enhance the precision of the BMI by correcting as many wrong commands as it is possible, avoiding canceling correct ones due to FP detections. For this reason, a goodness score is calculated as the weighted sum of individual means of the preceding metrics (μ_i), where each metric is multiplied by its corresponding weight (ω_i):

$$Goodness\ Score\ (GS)_i = \sum_{i=1}^{n} \omega_i * \mu_i - 0.1 * \sum_{i=1}^{n} \omega_i * \sigma_i \tag{5}$$

The weights (ω_i) are chosen to give more importance to those metrics that measure the model's ability to classify correctly (F1 Score), maximizing TP (Recall) and correcting false starts. For this reason, the F1-Score is assigned the highest weight of 0.5, Recall a 0.3, and both Accuracy and Precision a 0.1. Notice, that a penalty factor is subtracted as a 10 % of the sum of the weighted standard deviations of the metrics (σ_i).

2.8. Ensembled classification system design

In the literature, various authors have employed an ensembled system in EEG signals to enhance predictions of ErrP [27,28]. In this section, the complete process carried out to design the ensembled system to detect the ErrP is explained. Multiple features, their combinations, electrodes distributions and trios of classifiers have been evaluated to determine the optimal set for the system. For this reason, a process to optimize the selection of these parameters is executed.

This implementation is based on a heuristic brute-force approach to select different parameters of the system. Thus, an exhaustive exploration of the parameters is undertaken, yet it is constrained by specific decisions, significantly reducing options to limit the required computing time. The whole process is applied independently for each feedback.

In Fig. 5, a global overview of the selection process is shown. While the number of possible combinations among parameters is extensive, certain parameters have been pre-defined for the initial selection phase such as the initial trio of classifiers and their parameters: SVM, LDA and SGD.

Step 1. Parameter to test: Electrodes selection per feature.

First, the optimal electrodes distribution for each feature (E_{F_i}) is determined, maximizing the Goodness Score (5) for the initial trio of classifiers. This optimal combination of electrodes is associated with each individual feature for the rest of steps.

Step 2. Parameter to test: Individual features.

Once selected the best selection of electrodes per feature, they are tested. As the process is applied three times for each feedback type, the scores of each stimulus are normalized by their respective maximum score. The final score for each feature (F_i) is estimated by summing up the normalized scores across the three feedback types (6).

$$Total\ Score\ (TS)_i = \frac{\left(\sum_{j=1}^{3} \frac{F_{ij}}{\max(F_j)}\right)}{3}\ i \in \{1,\ 2, ..,\ 8\}\ features,\ j \in \{1,\ 2,\ 3\}\ feedbacks \tag{6}$$

Consequently, those features that have reached the highest scores with their respective optimal electrodes they are preselected as the best 5 features (F_S), ready to be combined with each other.

Step 3. Parameter to test: Features combination.

The next step involves the combination of the preselected features in groups up to 5 features. Then, the three combinations (FC_m) with the highest TS (6) are chosen for the next ensembled system steps.

Step 4. Parameter to test: Trio of classifiers.

First, a grid search fine-tunes the parameters of the classifiers according to the selected combination of features by testing all possible combinations of specific parameters' values to identify the set that achieves the best performance. Therefore, this process enhances the classifiers' performance and ensures equitable conditions for selection. This is crucial as the default classifier parameters may not be optimal for the current dataset. Given that approximately 10 classifiers are used, each with multiple hyperparameters, detailing every parameter combination would be impractical due to the sheer volume of possibilities. However, the grid search is conducted thoroughly to cover a comprehensive range of values for each classifier, ensuring optimal performance.

Finally, the classifiers (C) for the ensembled system are swept in combinations of three, based on their performance and adaptability to the optimized feature set and parameter configuration.

At the end of the iterative selection, the Ensembled System is configured with all the select parameters (FC, E_{FC},C) through the process and the feature combination with higher score after the classifiers selection.

2.8.1. Selection process of a parameter

Each of the steps of the previous section sweeps a group of variables fixing another group. The test of a certain variable requires a classification of the evoked potentials to obtain a GS value and compare it with the others under consideration. This section explains how the individual computation of a given combination of parameters is carried out in order to avoid an excessive number of iterations. A diagram with the selection process of a parameter is shown in Fig. 6.

First, an initial list of individual variables to evaluate a given parameter is generated. This requires keeping the variables of the rest of parameters fixed as the ensembled system steps indicate. For each of the variables under iteration the process is similar: features are extracted,

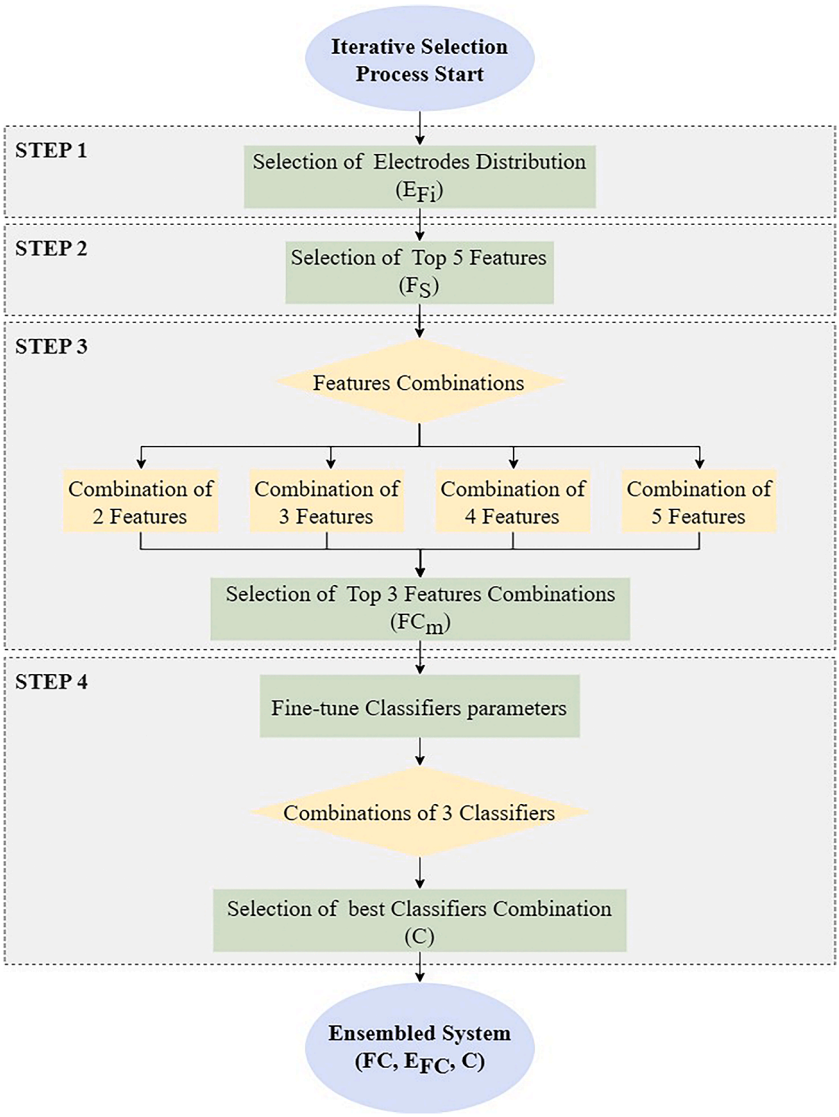

Fig. 5. General schema of the iterative selection process to select the parameters of the ensembled system. Process is carried out in several steps, choosing the optimal electrodes distribution (E_{F_i}), features (F_S), combination of features (FC), classifiers parameters and trio of classifiers (C).

classifiers trained, and scores assessed for the test vectors following a cross-validation approach.

This requires that once all parameters are set up, the dataset is divided into training and test samples. As the number of samples for model creation is limited, the training conducted is a generic, which means that each of the training models contains also the information of the rest of the subjects, balancing the samples per class for each subject, as the ErrP samples per subject (18 samples) is different to the NoErrP samples (42 samples). This approach means that when a specific subject is evaluated, their data is split and included in both the training and test datasets, and data from the other subjects is also incorporated into the training dataset. The main reason for applying a leave-one-out cross-validation is to ensure that all 18 ErrP test samples per subject are used. For instance, in the initial fold, the test dataset includes the first ErrP event and a randomly selected NoErrP event, whereas the training dataset comprises the remaining 17 ErrP samples and 17 randomly selected samples of the NoErrP class, along with balanced data (18 ErrP/18 NoErrP) from the other subjects. Then, the test dataset for the next fold includes the second ErrP event and another randomly selected NoErrP event, with the training set comprising the remaining 17 ErrP samples and another 17 randomly selected NoErrP samples, plus the balanced data from the other subjects. This process continues until each of the 18 ErrP samples are used once as part of the test set.

The balancing undersampling method employed consists of equalizing both classes by reducing the number of samples of NoErrP class in training and test groups. Previously, other techniques, such as weighting to the unbalanced classes, were tried but the results achieved less than a 50 % of the accuracy. Furthermore, the balanced data is also standard scaled to a range from 0 to 1 to enhance the classifiers' performance.

Once the data is already prepared, the models are trained and, by the end of the procedure, the four metrics Accuracy (1), Precision (2), Recall (3) and F1Score (4) are calculated. Furthermore, with the purpose of reaching a stable result and trying to include all the NoErrP data left out by the undersampling balance, the cross-validation process is repeated 20 times selecting 20 randomized vectors of the NoErrP class. Therefore, the mean and standard deviation of each of the metrics is calculated and weighted to estimate the Goodness Score (5).

After computing the GS (5), for each of the variables to evaluate, the one with the highest score is initially selected and systematically paired with the remaining non-selected variables. In the case that variables could be combined, i.e. electrodes, features, classifiers, the training process to estimate the GS of each variable combination begins again. This incremental loop continues as long as the GS reached by the combinations of variables remains higher than the best score from the

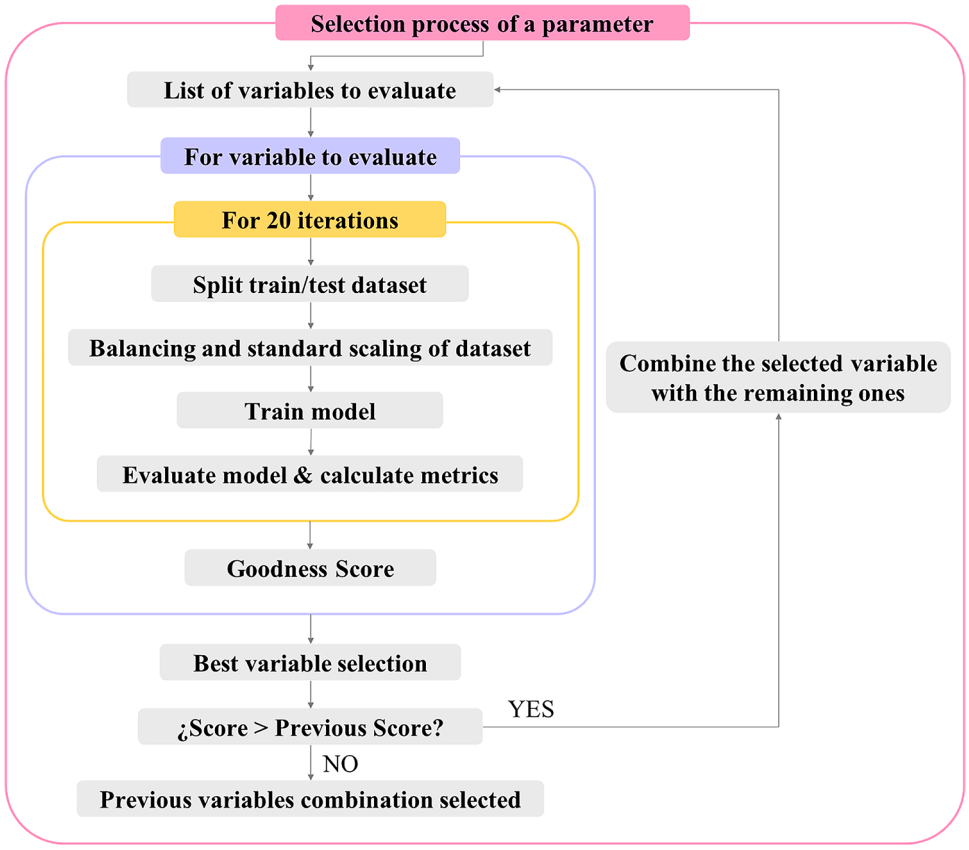

Fig. 6. General schema of the iterative selection process of a given variable.

previous loop iteration. If the new score is lower than the score preceding iteration, the combination loop breaks, and the previous combination is selected for the parameter, limiting the number of iterations.

3. Results

The current section provides the results of the different analysis carried out during the research. First, the ensembled classification results by step are presented. Then, an evaluation of the system for the

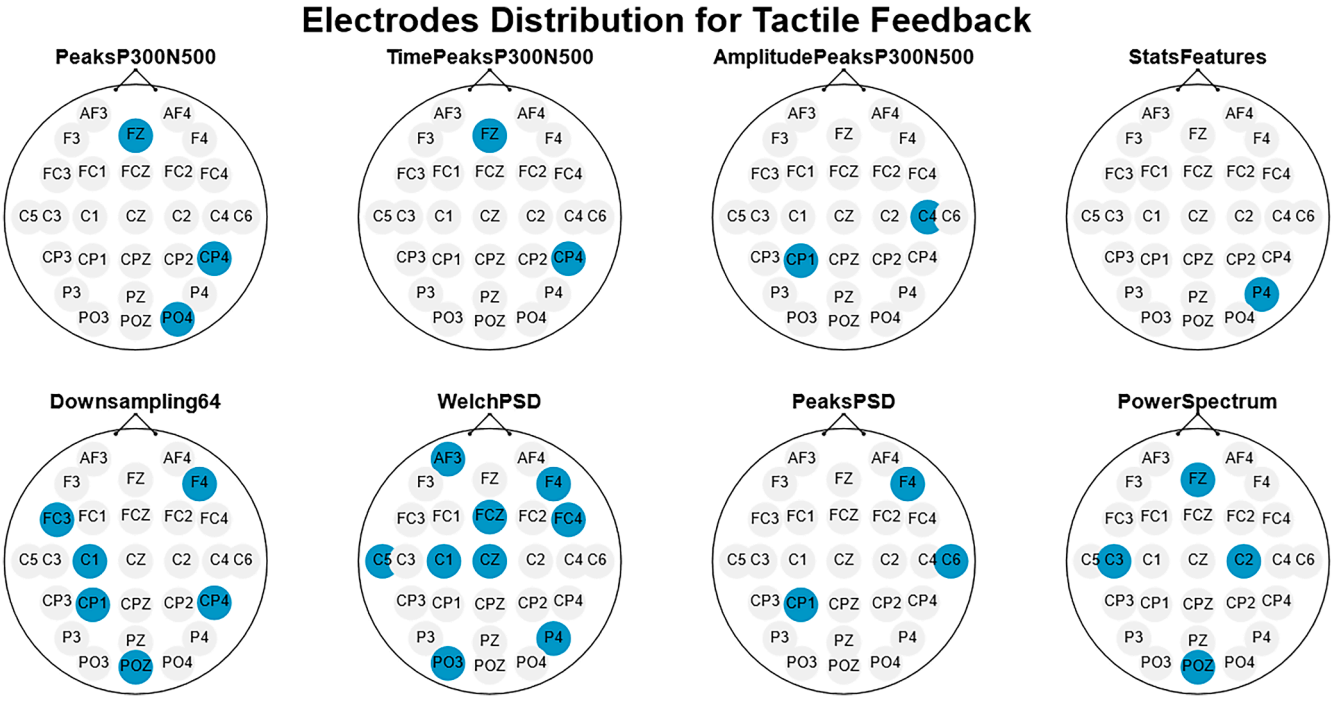

Fig. 7. Electrode distribution selection for features in Tactile feedback. The blue circles represent the electrodes with a higher GS (5).

parameters chosen by the ensembled system as a BMI for the correction of undesired activations of the exoskeleton is provided.

3.1. Iterative selection process

This section provides the results of the ensembled system during the different steps associated with the selection of electrodes, features, features combinations and classifiers.

3.1.1. Electrodes selection per feature

The first step is the selection of the electrodes more suitable for each feature. The process is carried out by the ensembled system in an iterative way, increasing the configuration of electrodes till the GS is not improved. This means that if, for instance, PowerSpectrum feature obtained a GS of 67 % with Fz+POz as best combination of two electrodes, all the combinations of Fz+POz+XX, with XX being the rest electrodes, are tested and a 3 electrodes combination is selected only if they improve the 67 % result and a four electrodes combination will then be tested.

Fig. 7 shows the results for Tactile feedback. Regarding time domain features, the electrodes more relevant are shown around Fz, CP4 and parieto-occipital electrodes. Shape features, such as Downsampling, also indicate CP4 and POz relevance, but without a predominance of right hemisphere. Statistical features only present one electrode on P4. In frequency domain, features related to Welch show a very different behavior with a spread of electrodes in the case of WelchPSD and a major presence of right hemisphere for PeaksPSD with a CP1 presence. In contrast, PowerSpectrum electrodes exhibit similarities to those selected in the peaks features of P300 and N500 along with C2 and C3. The number of electrodes selected for each feature is also very variable with features with just one or two electrodes to others with nine.

Notice that in this study the distribution is not indicating the electrodes which define ErrP, but those that discriminate with a higher score between ErrP and NoErrP classes.

3.1.2. Individual features

Once selected the combination of electrodes per feature, their GS are compared in order to select the more discriminant features. Fig. 8 shows the GS of each feature per feedback (5) and their normalized TS as defined in Eq. (6).

GS provides results around 60 %. Regarding stimuli, their performance depends on the feature under consideration. However, Visual achieves higher results for five of the eight features. This means that a combination of features could provide a better discrimination of the ErrP. The five features to test in the next step of the ensembled system are selected based on the TS.

3.1.3. Features combination

Once selected the five features with higher TS, they are combined with each other up to groups of five features. Results in Fig. 9 are improved just by a little for the combination of features in comparison to the individual characterization (left part of Fig. 8). However, there is a clear dominance for Tactile feedback in the results, which is not present in the individual features of Fig. 8. As expected by the similarity of results of the individual features of Section 3.1.2. Individual features, the highest TS are achieved by the combination of the majority of features.

3.1.4. Trio of classifiers

The last step in the ensembled system is choosing the trio of classifiers to use. Before this step, an optimization of hyperparameters per classifier was carried out by means of a grid search with all possible combinations of each classifier parameters, as before this step the initial trio (SVM, LDA and SGD) was using default parameters.

Given the extense number of hyperparameters involved, only the grid-search results for the classifiers SVM, LDA and LR using FC1 feature combination under Tactile feedback are specified. For SVM, the optimal hyperparameters are a C value of 10, a gamma of 0.001, and an RBF kernel. For LDA, the best configuration includes an auto shrinkage and an LSQR solver. For LR, the optimal parameters consist of an L1 penalty and a SAGA solver.

Fig. 10 illustrates the classifier analysis for each feedback type. The

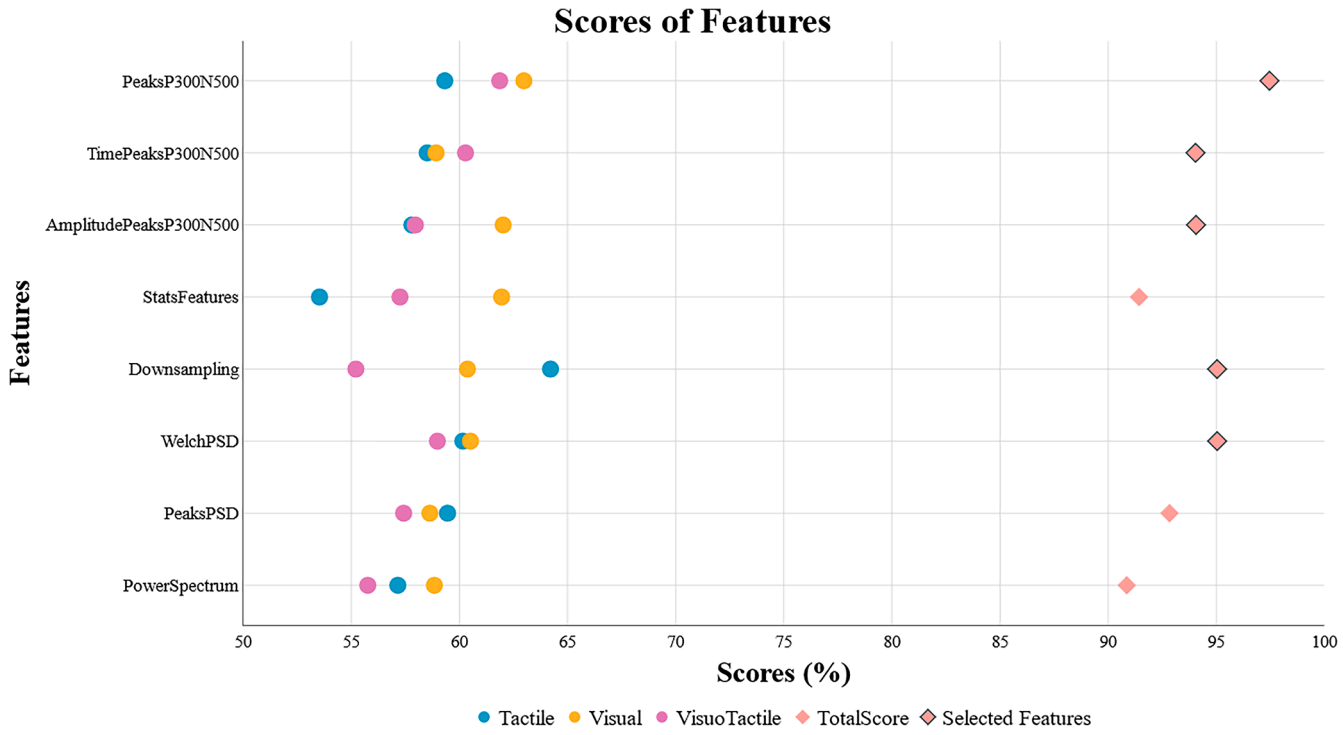

Fig. 8. Averaged results of the subjects for each feature and feedback. The GS (5) of each stimulus is indicated as circular markers in blue (Tactile), yellow (Visual) and pink (Visuo-Tactile). In addition, diamonds in salmon represent the TS of each feature (6). The five selected features appear as black bolded diamonds.

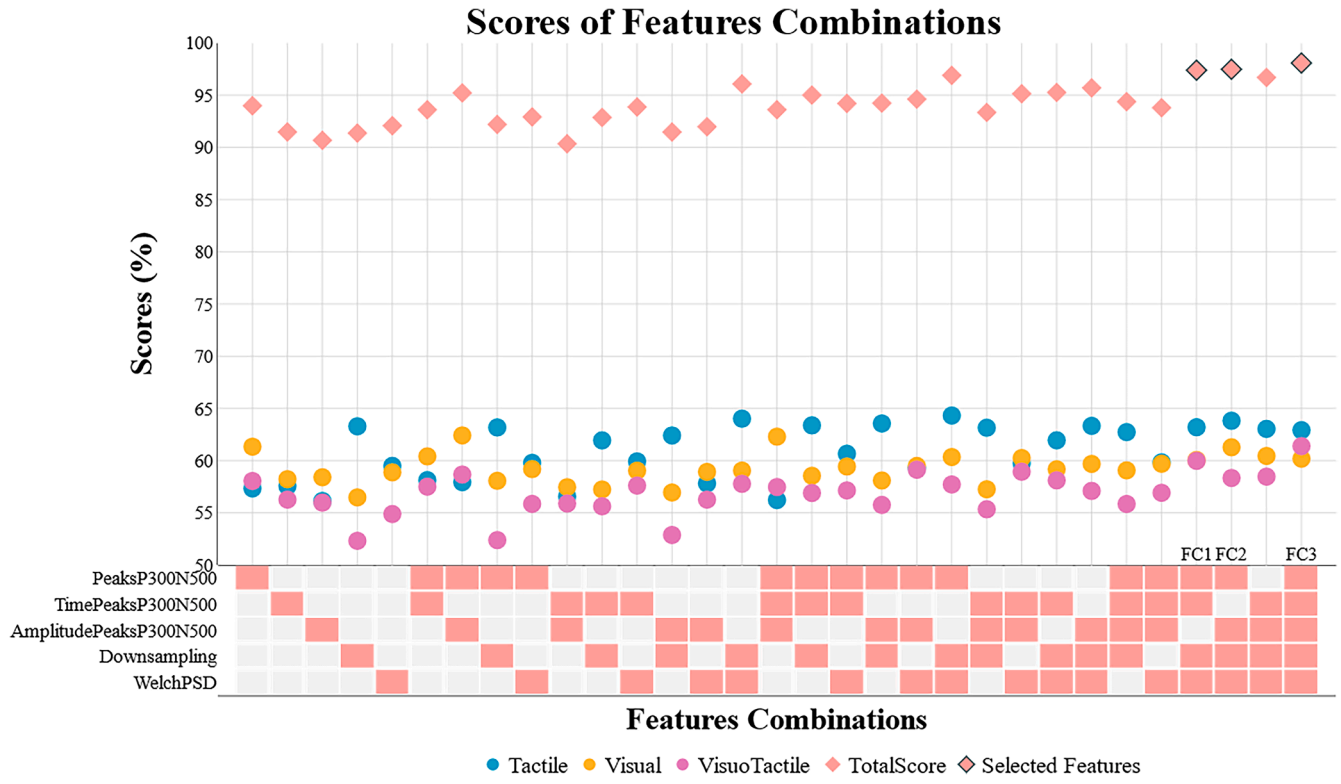

Fig. 9. Averaged scores of the six subjects for the five selected features and their combinations. The features combinations are marked by light red. The GS (5) of each stimulus is indicated as circular markers in blue (Tactile), yellow (Visual) and pink (Visuo-Tactile). In addition, diamonds in salmon represent the TS of each combination (6). The three selected combinations (FC1–3) are indicated by bolded black diamonds.

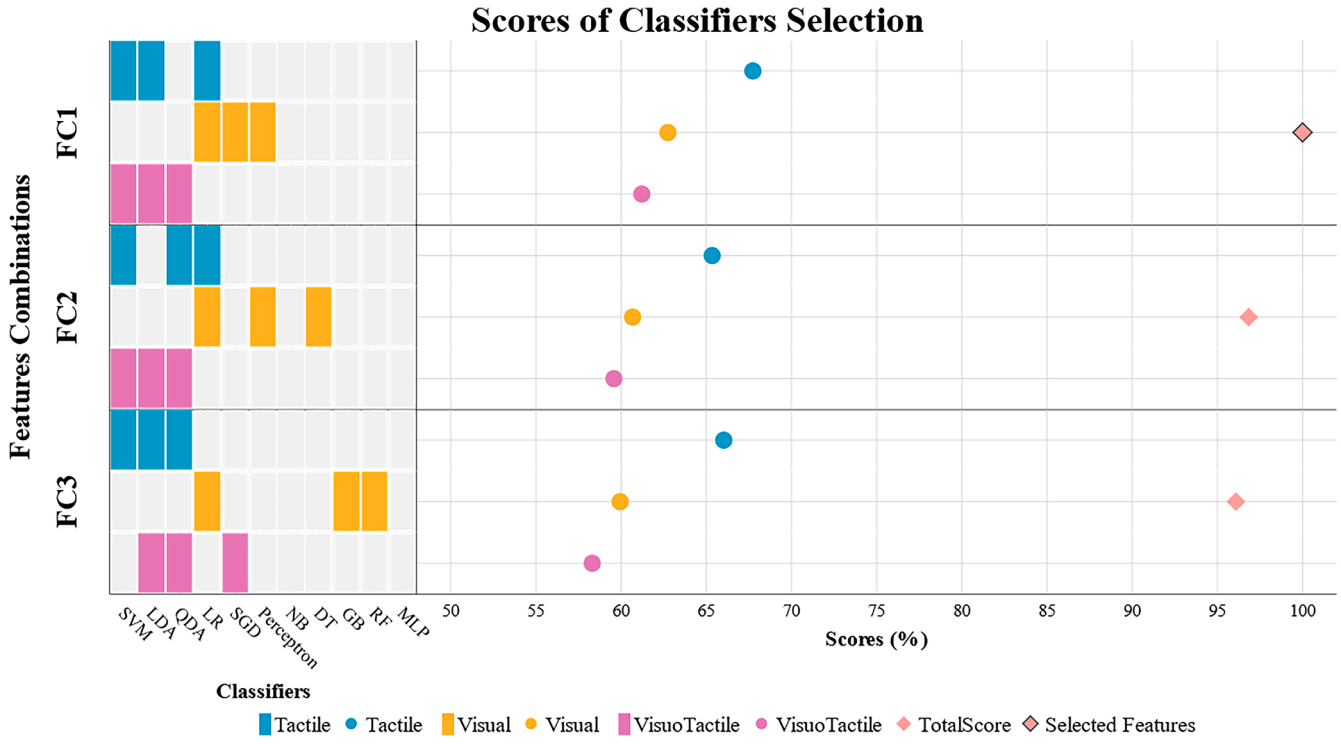

Fig. 10. Combination of classifiers with the highest subject averaged GS for each of the top three features combinations (FC). The selections and results are color coded for the three feedbacks: blue (Tactile), yellow (Visual) and pink (Visuo-Tactile). GS are represented in round spots and TS in diamond. The highest TS under bolded black diamond is achieved for the FC1 with Tactile feedback as highest GS.

choice of classifiers remains quite similar across feature combinations within the same feedback type. SVM, LDA, QDA, and LR consistently emerge as the most frequently selected classifiers. Analyzing scores the first combination (FC1), which contains both time and frequency domains features, such as PeaksP300N500, TimePeaksP300N500, Downsampling, and WelchPSD achieves the highest TS. Regarding feedback types, Tactile is the one with the highest GS with a 67.73 %.

3.2. Subject analysis

EEG analysis usually has a high subject dependency. Up to this point all the step decisions are based on the average scores of the six subjects. This section will present the individual subject results for the FC1 and the selected trio of classifiers per feedback.

As expected, subject dependency is high as Fig. 11 results show. There is also different subject behavior dependent on the feedback type. However, Tactile feedback is the only one which achieves GS over 60 % for all the subjects. GS index is a combination of accuracy, precision, recall and F1Score metrics and deviation. These indices can be seen in Table 2.

3.3. False starts correction

Although GS (5) and TS (6) have been used as a metrics for the ensembled system decision making, they do not provide a realistic idea of the behavior of the tool as a BMI for the correction of false activations of a lower-limb exoskeleton. Table 2 shows two new metrics for this purpose: True Positive Ratio (TPR) and False Positive Ratio (FPR). TPR represents in percentage the number of corrections to wrong activations of the exoskeleton and FPR the number of correct activations that are spoiled. This means that TPR should be as higher as possible and FPR as lower as it could. The results for the designed BMI show a TPR of 72.60 % and a FPR of 40.92 % in average.

4. Discussion

4.1. Iterative selection process

The different steps of the iterative parameter selection provide different results to discuss.

4.1.1. Brain locations

Electrode selection does not seem to indicate a certain brain zone configuration. As the selection is not trying to characterize the ErrP, but the electrodes that are best suited to differentiate between ErrP and NoErrP, electrode selection varies from different types of features. Nevertheless, there are some brain zones that appear in most of the features. Fz electrode due to its relationship with the ErrP characterization [10,29–31], parieto-occipital due to its relationship with visual stimuli processing [32] and right tempo-parietal brain zones due to the individual attention to new stimuli [33].

4.1.2. Features

The combinations with highest scores select a mix of time and frequency domain features like PeaksP300N500, TimePeaksP300N500, Downsampling, and WelchPSD. This is something that is reported by other investigations, such as [20], where it is remarked the importance of combining time and frequency domain features to improve ErrP detection. Nevertheless, as the discrimination between the two evoked potentials seem to be related to different brain processes, the use of different features provides better information to the classifiers.

4.1.3. Classifiers

The fine-tuning of the parameters along the iteration of the trio of classifiers allows a marginal increase in the scores as it is seen in the comparison of Fig. 9 and Fig. 10. From the inspection of the results, it is clear that Tactile and Visuo-Tactile have a similar selection of classifiers in comparison to visual, with SVM, LDA, QDA, and LR as the more repeated ones.

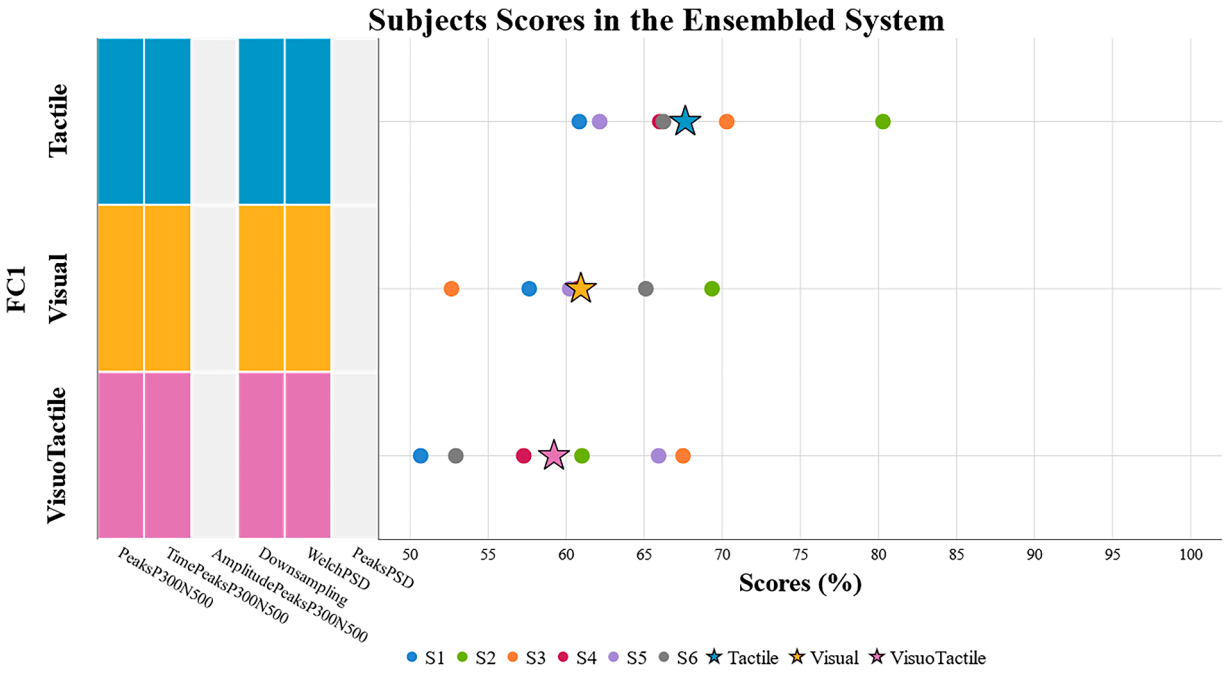

Fig. 11. Individual subjects' GS for FC1 and its best classifier combination per feedback. Individual subject scores are color coded in spots (S1: blue, S2: green, S3: orange, S4: magenta, S5: purple, S6: gray) while averaged feedback appears in blue stars for Tactile, yellow stars for Visual and pink stars for Visuo-Tactile feedback.

Table 2
Individual classifier statistic metrics obtained for Tactile feedback, using FC1 and SVM, LDA and LR classifier trio.

Subjects	Accuracy (%)	Precision (%)	Recall (%)	F1Score (%)	GS (%)	TPR (%)	FPR (%)
S1	64.18	66.78	60.08	61.95	60.84	60.00	31.70
S2	76.78	72.56	88.75	79.26	80.29	88.80	35.20
S3	66.18	64.21	78.10	69.70	70.28	78.10	45.80
S4	60.35	58.53	74.85	65.16	65.99	74.90	54.20
S5	59.28	58.41	67.95	62.22	62.14	67.90	49.40
S6	68.35	70.63	65.95	67.33	66.23	65.90	29.20
MEAN	65.85 ± 6.36	65.19 ± 5.96	72.61 ± 10.19	67.60 ± 6.44	67.63 ± 7.04	72.60 ± 10.23	40.92 ± 10.27

4.2. Feedback

Once the ensembled system has fixed the main parameters, Fig. 9 and Fig. 10 show Tactile feedback as the one with the highest scores, this is also referred by the individual results by subject shown in Fig. 11. Therefore, a statistical analysis is performed. Specifically, a Shapiro-Wilks normality test is conducted, and the results indicate that the p-values of all variables are greater than 0.05, following a normal distribution. For this reason, a multivariate analysis of variance (MANOVA) test is applied to further analyze the differences among feedback types in the obtained results and confirm the significance of the results. The MANOVA test, based on Pillai's Trace, show significant differences among the feedback types, as the p-values < 0.05 ($V = 1.160$, $F(12,22) = 2.531$, $p = 0.029$, $\eta^2 = 0.580$), confirming the statistical significance of the Tactile feedback's superior performance. These results are along to previous studies that show Tactile feedback as a better option to characterize ErrP [20,22].

Tactile feedback appears to be the best in terms of classification performance for several reasons. First, vibration stimulus is processed through the somatosensory system, which may provide a more direct and immediate response compared to Visual feedback. In the literature, studies have shown that tactile feedback can enhance the perception of movement and improve motor control accuracy [34,35]. Furthermore, tactile stimuli are less likely to cause cognitive overload, allowing more efficient processing of error [36,37]. Additionally, in our experiments the instrumentation of the feedback, as a less intrusive gadget, could affect in a lesser way to the conducted mental tasks.

4.3. Subject dependency

Subject dependency is widely reported in the literature [38]. This is something that is clearly seen in Table 2 and Fig. 11 results, which means that not all the subjects are adapted to the protocol or modulate their evoked potentials in the same way with TPRs from 59 % to 89 %. In addition, some authors have reported that subjects with low accuracy in one feedback, also present low accuracy in the rest of the feedbacks, as is the case of S1 [22].

Due to the limited number of ErrP samples the classifier model uses information of all the subjects, which has its pros and cons. A personalized model with data of just the subject under analysis would be desired, but it would require long training protocols which would make the BMI unusable. Nevertheless, as it would be seen in the next section, even low percentage over the 50 % would improve the BMI global behavior.

4.4. False starts correction

The aim of the research is to develop a new BMI that corrects false starts of a lower-limb exoskeleton working in real-time closed-loop control. Usually, accuracy percentage is reported as acceptable at a 70 % threshold [39]. However, it is rare to achieve those percentages in lower-limb real-time closed-loop BMIs based on MI [7].

For this reason, Table 3 shows a simulation of the expected outcome of the system in real conditions. As a MI classifier is not applied in this part of the research, but the final application will, an acceptable accuracy ratio for the MI classifier is considered to be 70 %. Therefore, this classifier would result in a 30 % rate of undesired starts that need correction, which is where the ErrP detection layer operates.

Table 3
Variations in the percentage of false and correct starts, as well as the global accuracy of a BMI based on MI combined with the ErrP detection achieved in Table 2 considering a BMI-MI original accuracy of 70 %.

Subjects	MI+ErrP false start (%)	MI+ErrP correct start (%)	MI+ErrP Global Accuracy (%)
S1	11.98	47.81	65.83
S2	3.36	45.36	72.00
S3	6.57	37.94	61.37
S4	7.53	32.06	54.53
S5	9.62	35.42	55.81
S6	10.22	49.56	69.35
MEAN	8.21±3.06	41.36±7.19	63.15±7.15

The new metrics of the simulated BMI-MI are calculated employing TPR and FPR metrics exhibited in Table 2. Thus, MI-ErrP false starts (FS) percentage is calculated as *(1–0.7)*(1-TPR)* and the percentage of correct starts (CS) is estimated as *0.7–0.7*FPR*. Then, the global accuracy of the BMI-MI combined with ErrP is *CS+0.3-FS*.

The combination with the ErrP, based on the exposed TPR and FPR, would correct the false starts from 30 % to the Table 3 new metrics. Therefore, this could make the BMI less sensitive, spoiling some of the correct starts from the previous 70 %. However, although only S2 would improve the global accuracy, the significant reduction of the false starts ratio enhances the security of the system, while a worsening of the correct start ratio would cause just longer periods of time to start, but without safety problems.

4.5. Strengths and limitations

This study presents several notable strengths that underscore its contribution to the field of BMIs to control lower limb exoskeletons. One of the primary strengths is the large dataset that includes three different types of feedback: Tactile, Visual, and Visuo-Tactile. This dataset is particularly valuable as it allows to carry out a comparative analysis of the different stimuli to evoke and detect the ErrP, providing also insights that guide the future research.

Additionally, this study is pioneering in its investigation of ErrP to begin the gait with exoskeleton, as there are no previous works in this specific area. Furthermore, the promising results highlight the potential of ErrP for increasing both, the efficiency and security of BMI-MI for gait control. Therefore, the self-correction of false activations is another significant strength, as it enhances the reliability and user-friendliness of the BMI system.

Despite these strengths, the study also shows diverse limitations that must be acknowledged. One of the most limiting points of the research is the low amount of useful data generated during each session, which poses challenges for classification models. Machine learning models have been used for the detection of ErrP, but there is a concern that these models used may be overfitted to the session dataset, which could affect its generalizability to other datasets or real-time applications. Moreover,

each feedback is tested in just one session with a pretty reduce number of test samples.

Although the number of subjects is considerable for a BMI study, as referenced in [40], a large number of samples would be desirable to confirm the indicated findings. High variability between subjects' data is a common issue in BMI research, and this variability can be mitigated by increasing the sample size. Additionally, with a larger dataset, the challenges associated with applying deep learning methods and the risk of the classifiers being overfitted can be addressed more effectively.

During the experiments, subjects experienced fatigue which may have influenced their performance and the reliability of the data collected. The monotony of the sessions can also cause subjects to lose focus towards the final trials, although it closely reproduces the real-time conditions, ensuring that the study's findings are relevant to practical applications.

Another significant limitation is the high false positive rate (FPR) of 40.92 %, which induces frustration in subjects due to the correction of correct starts resulting from erroneous ErrP detections. This issue can negatively impact the user experience and trust in the BMI system. However, the high FRP primarily leads to slightly longer start times, but it does not compromise the subject's safety. For instance, in subject S6, this would occur less than 3 times out of 10 trials, which is not a significant volume to cause considerable frustration.

Despite the high rate of FPR, the results are comparable to those reported in the literature. For instance, [20] employs Tactile feedback to evoke ErrP and achieves a TPR of approximately 73 % and a FPR of 12 %. Similarly, [41] trains the model with generic dataset, reaching around 74 % of TPR and 28 % of FPR. These results are similar to TPR (72.60 %) in Table 2 of the current study. Furthermore, [14] discusses the improvement of a BCI by detecting ErrP, achieving an error rate of approximately 7 %, which aligns closely with the results of the current research. In that study, the rejection rate is around 35 %, showing similarities with the results presented in Table 3.

To sum up, while the study demonstrates important advancements in BMI-MI technology and offers valuable data for further research, it also highlights several areas where improvements are needed to enhance the system's reliability and user experience.

4.6. Clinical applicability

The integration of ErrP detection to self-tune the exoskeleton's movements ensures that patients can rely on the technology to safely and accurately assist their mobility needs. The proposed system shows significant potential for clinical applicability, particularly in the rehabilitation of patients with SCI within hospital settings. By correcting the erroneous movements before they are executed in real-time, this technology can greatly enhance the effectiveness of gait rehabilitation programs. Additionally, the system could be useful for assisting patients in their daily routines, offering support in maintaining balance and correcting gait, which would contribute to giving SCI patients more independence and quality of life. Nevertheless, exoskeletons are still pretty expensive, making it difficult for individuals to afford and employ them at home.

4.7. Future work

Future research will focus on several key areas to further improve and validate the system. One of the primary goals is to explore the detection of ErrP during gait to reduce false stops, enhancing the fluidity and reliability of the BMI. Additionally, efforts will be made to increase TPR detections and minimize FPR to the greatest extent possible. Finally, the integration of both start and stop ErrP signals with the BMI-MI will be tested to evaluate the system's real-time performance comprehensively. These advancements will be tested with SCI patients to assess the practicality and effectiveness of the system in the clinical settings for their rehabilitation process.

5. Conclusion

The study has presented a BMI which discriminates false starts in a BMI based on MI for the control of a lower-limb exoskeleton. Even though the results percentages vary from subjects, the correction of false starts improve the security of the system at the cost of sensitivity to start.

The number of ErrP events that can be achieved by session and subject is limited. This is caused by the recommended 30 % error ratio at which subjects get used to errors and no longer perceive them as such [10]. As six subjects during three sessions participated in the study, the number or ErrP generated was of 108 per feedback type and 324 in total. This could be a low value, but there is not a similar study in the literature and the usual number of subjects for a lower-limb exoskeleton investigations usually present one subject and studies over ten are a rarity [40]. Thus, the provided database and the described protocol, which compares three different feedback types, should be very valuable for the study of ErrP in other investigations. Furthermore, after the feedback comparative and the selection of Tactile feedback for further research, increasing this feedback dataset is a priority to enhance the robustness of the system.

Future research would focus on the analysis of the ErrP during motion, in order to correct false stops during the gait, and the integration of the start and stop ErrP BMI with the BMI-MI to test its real-time performance.

Availability

Data related to correct and error evoked potentials by three different feedback types (Tactile, Visual and Visuo-Tactile) before a lower-limb exoskeleton activates to start the gait is available in the Zenodo repository (10.5281/zenodo.10828803).

CRediT authorship contribution statement

P. Soriano-Segura: Methodology, Investigation, Software, Visualization, Formal analysis, Writing – original draft. **M. Ortiz:** Conceptualization, Methodology, Data curation, Writing – review & editing, Supervision. **E. Iáñez:** Conceptualization, Resources, Supervision, Data curation. **J.M. Azorín:** Funding acquisition, Project administration, Supervision, Conceptualization.

Declaration of competing interest

The authors declare that they do not have competing financial interests or personal relationships that could represent a conflict of interest with the submitted work.

Acknowledgements

This publication is part of grant PID2021-124111OB-C31, funded by MICIU/AEI/10.13039/501100011033 and by ERDF, EU. This work is also supported by the Valencian Graduate School and Research Network of Artificial Intelligence (ValgrAI), Generalitat Valenciana and European Union.

References

[1] D.J. McFarland, J.R. Wolpaw, Brain-computer interfaces for communication and control, Commun. ACM 54 (2011) 60–66, https://doi.org/10.1145/1941487.1941506.
[2] M. Ortiz, K. Nathan, J.M. Azorín, J.L. Contreras-Vidal, Brain-machine interfaces for neurorobotics, in. Handbook of Neuroengineering, Springer Nature Singapore, Singapore, 2023, pp. 1817–1857, https://doi.org/10.1007/978-981-16-5540-1_52.
[3] K.K. Ang, C. Guan, Brain-computer interface in stroke rehabilitation, J. Comput. Sci. Eng. 7 (2013) 139–146, https://doi.org/10.5626/JCSE.2013.7.2.139.
[4] M. Ortiz, E. Ianez, J. Gaxiola, A. Kilicarslan, J.L. Contreras-Vidal, J.M. Azorin, Assessment of motor imagery in gamma band using a lower limb exoskeleton, in:

Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2019, pp. 2773–2778, https://doi.org/10.1109/SMC.2019.8914483.
[5] M. Rodríguez-Ugarte, E. Iáñez, M. Ortiz, J.M. Azorín, Improving real-time lower limb motor imagery detection using tDCS and an exoskeleton, Front. Neurosci. 12 (2018), https://doi.org/10.3389/fnins.2018.00757.
[6] M. Bakker, F.P. De Lange, R.C. Helmich, R. Scheeringa, B.R. Bloem, I. Toni, Cerebral correlates of motor imagery of normal and precision gait, Neuroimage 41 (2008) 998–1010, https://doi.org/10.1016/j.neuroimage.2008.03.020.
[7] L. Ferrero, V. Quiles, M. Ortiz, E. Iáñez, J.M. Azorín, A BMI based on motor imagery and attention for commanding a lower-limb robotic exoskeleton: a case study, Appl. Sci. 11 (2021) 4106, https://doi.org/10.3390/app11094106.
[8] L. Ferrero, E. Ianez, V. Quiles, J.M. Azorin, M. Ortiz, Adapting EEG based MI-BMI depending on alertness level for controlling a lower-limb exoskeleton, in: Proceedings of the IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), IEEE, 2022, pp. 399–403, https://doi.org/10.1109/MetroXRAINE54828.2022.9967639.
[9] H. Lorach, A. Galvez, V. Spagnolo, F. Martel, S. Karakas, N. Intering, M. Vat, O. Faivre, C. Harte, S. Komi, J. Ravier, T. Collin, L. Coquoz, I. Sakr, E. Baaklini, S. D. Hernandez-Charpak, G. Dumont, R. Buschman, N. Buse, T. Denison, I. van Nes, L. Asboth, A. Watrin, L. Struber, F. Sauter-Starace, L. Langar, V. Auboiroux, S. Carda, S. Chabardes, T. Aksenova, R. Demesmaeker, G. Charvet, J. Bloch, G. Courtine, Walking naturally after spinal cord injury using a brain–spine interface, Nature 618 (2023) 126–133, https://doi.org/10.1038/s41586-023-06094-5.
[10] R. Chavarriaga, J. del R Millan, Learning from EEG error-related potentials in noninvasive brain-computer interfaces, IEEE Trans. Neural Syst. Rehabil. Eng. 18 (2010) 381–388, https://doi.org/10.1109/TNSRE.2010.2053387.
[11] R. Chavarriaga, A. Sobolewski, J. del R Millán, Errare machinale est: the use of error-related potentials in brain-machine interfaces, Front. Neurosci. 8 (2014), https://doi.org/10.3389/fnins.2014.00208.
[12] M. Mousavi, L.R. Krol, V.R. de Sa, Hybrid brain-computer interface with motor imagery and error-related brain activity, J. Neural Eng. 17 (2020) 056041, https://doi.org/10.1088/1741-2552/abaa9d.
[13] A. Cruz, G. Pires, U.J. Nunes, Double ErrP detection for automatic error correction in an ERP-based BCI speller, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (2018) 26–36, https://doi.org/10.1109/TNSRE.2017.2755018.
[14] P.W. Ferrez, J. del R Millan, Simultaneous real-time detection of motor imagery and error-related potentials for improved BCI accuracy, in: Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course, 2008, pp. 197–202. https://www.researchgate.net/publication/41386722.
[15] V. Mondini, A.I. Sburlea, G.R. Müller-Putz, Towards unlocking motor control in spinal cord injured by applying an online EEG-based framework to decode motor intention, trajectory and error processing, Sci. Rep. 14 (2024) 4714, https://doi.org/10.1038/s41598-024-55413-x.
[16] A. Xavier Fidêncio, C. Klaes, I. Iossifidis, Error-related potentials in reinforcement learning-based brain-machine interfaces, Front. Hum. Neurosci. 16 (2022), https://doi.org/10.3389/fnhum.2022.806517.
[17] P.W. Ferrez, J. del R Millan, Error-related EEG potentials generated during simulated brain–computer interaction, IEEE Trans. Biomed. Eng. 55 (2008) 923–929, https://doi.org/10.1109/TBME.2007.908083.
[18] C. Lopes-Dias, A.I. Sburlea, G.R. Müller-Putz, Online asynchronous decoding of error-related potentials during the continuous control of a robot, Sci. Rep. 9 (2019) 17596, https://doi.org/10.1038/s41598-019-54109-x.
[19] J. Omedes, I. Iturrate, R. Chavarriaga, L. Montesano, Asynchronous decoding of error potentials during the monitoring of a reaching task, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, IEEE, 2015, pp. 3116–3121, https://doi.org/10.1109/SMC.2015.541.
[20] B. Ahkami, F. Ghassemi, Adding tactile feedback and changing ISI to improve BCI systems' robustness: an error-related potential study, Brain Topogr. 34 (2021) 467–477, https://doi.org/10.1007/s10548-021-00840-6.
[21] L. Schiatti, G. Barresi, J. Tessadori, L.C. King, L.S. Mattos, The effect of vibrotactile feedback on ErrP-based adaptive classification of motor imagery, in: Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2019, pp. 6750–6753, https://doi.org/10.1109/EMBC.2019.8857192.
[22] J. Tessadori, L. Schiatti, G. Barresi, L.S. Mattos, Does tactile feedback enhance single-trial detection of error-related eeg potentials?, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC) IEEE, 2017, pp. 1417–1422, https://doi.org/10.1109/SMC.2017.8122812.
[23] Y. Zhang, W. Chen, C.L. Lin, J. Chu, F. Meng, Research on command confirmation unit based on motor imagery EEG signal decoding feedback in brain-computer interface, in: Proceedings of the 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), IEEE, 2018, pp. 1923–1928, https://doi.org/10.1109/ICARCV.2018.8581088.
[24] P. Soriano-Segura, L. Ferrero, M. Ortiz, E. Iáñez, J.M. Azorín, Analysis of error potentials generated by a lower limb exoskeleton feedback in a BMI for gait control *, in: Proceedings of the 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2023, pp. 1–4, https://doi.org/10.1109/EMBC40787.2023.10340275.
[25] A. Kilicarslan, R.G. Grossman, J.L. Contreras-Vidal, A robust adaptive denoising framework for real-time artifact removal in scalp EEG measurements, J. Neural Eng. 13 (2016) 026013, https://doi.org/10.1088/1741-2560/13/2/026013.
[26] J. Omedes, I. Iturrate, L. Montesano, J. Minguez, Using frequency-domain features for the generalization of EEG error-related potentials among different tasks, in: Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 5263–5266, https://doi.org/10.1109/EMBC.2013.6610736.
[27] T. Tao, Y. Gao, Y. Jia, R. Chen, P. Li, G. Xu, a multi-channel ensemble method for error-related potential classification using 2D EEG images, Sensors 23 (2023) 2863, https://doi.org/10.3390/s23052863.
[28] S. Bhattacharyya, A. Konar, D.N. Tibarewala, M. Hayashibe, A generic transferable EEG decoder for online detection of error potential in target selection, Front. Neurosci. 11 (2017), https://doi.org/10.3389/fnins.2017.00226.
[29] I. Iturrate, L. Montesano, J. Minguez, Single trial recognition of error-related potentials during observation of robot operation, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 4181–4184, https://doi.org/10.1109/IEMBS.2010.5627380.
[30] E. Lopez-Larraz, I. Iturrate, L. Montesano, J. Minguez, Real-time recognition of feedback error-related potentials during a time-estimation task, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, IEEE, 2010, pp. 2670–2673, https://doi.org/10.1109/IEMBS.2010.5626623.
[31] C.S. Carter, T.S. Braver, D.M. Barch, M.M. Botvinick, D. Noll, J.D. Cohen, Anterior cingulate cortex, error detection, and the online monitoring of performance, Science 280 (1998) 747–749, https://doi.org/10.1126/science.280.5364.747.
[32] C. Babiloni, F. Vecchio, M. Miriello, G.L. Romani, P.M. Rossini, Visuo-spatial consciousness and parieto-occipital areas: a high-resolution EEG study, Cereb. Cortex 16 (2006) 37–46, https://doi.org/10.1093/cercor/bhi082.
[33] M. Corbetta, J.M. Kincade, J.M. Ollinger, M.P. McAvoy, G.L. Shulman, Voluntary orienting is dissociated from target detection in human posterior parietal cortex, Nat. Neurosci. 3 (2000) 292–297, https://doi.org/10.1038/73009.
[34] S. Park, J. Ha, L. Kim, Improving performance of motor imagery-based brain–computer interface in poorly performing subjects using a hybrid-imagery method utilizing combined motor and somatosensory activity, IEEE Trans. Neural Syst. Rehabil. Eng. 31 (2023) 1064–1074, https://doi.org/10.1109/TNSRE.2023.3237583.
[35] A.M. Savić, M. Novičić, O. Đorđević, L. Konstantinović, V. Miler-Jerković, Novel electrotactile brain-computer interface with somatosensory event-related potential based control, Front. Hum. Neurosci. 17 (2023), https://doi.org/10.3389/fnhum.2023.1096814.
[36] A. Kuc, I. Skorokhodov, A. Semirechenko, G. Khayrullina, V. Maksimenko, A. Varlamov, S. Gordleeva, A. Hramov, Oscillatory responses to tactile stimuli of different intensity, Sensors 23 (2023) 9286, https://doi.org/10.3390/s23229286.
[37] L. Yao, N. Mrachacz-Kersting, X. Sheng, X. Zhu, D. Farina, N. Jiang, A. multi-class BCI based on somatosensory imagery, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (2018) 1508–1515, https://doi.org/10.1109/TNSRE.2018.2848883.
[38] A. Melnik, P. Legkov, K. Izdebski, S.M. Kärcher, W.D. Hairston, D.P. Ferris, P. König, Systems, subjects, sessions: to what extent do these factors influence EEG data? Front. Hum. Neurosci. 11 (2017) https://doi.org/10.3389/fnhum.2017.00150.
[39] L. Yao, N. Jiang, N. Mrachacz-Kersting, X. Zhu, D. Farina, Y. Wang, Performance variation of a somatosensory BCI based on imagined sensation: a large population study, IEEE Trans. Neural Syst. Rehabil. Eng. 30 (2022) 2486–2493, https://doi.org/10.1109/TNSRE.2022.3198970.
[40] P. Wierzgała, D. Zapała, G.M. Wojcik, J. Masiak, Most popular signal processing methods in motor-imagery BCI: a review and meta-analysis, Front. Neuroinform. 12 (2018), https://doi.org/10.3389/fninf.2018.00078.
[41] C. Lopes-Dias, A.I. Sburlea, G.R. Muller-Putz, A generic error-related potential classifier offers a comparable performance to a personalized classifier, in: Proceedings of the 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), IEEE, 2020, pp. 2995–2998, https://doi.org/10.1109/EMBC44109.2020.9176640.